最优估计理论及其应用

——建模、滤波、信息融合估计

邓自立　著

Optimal Estimation Theory with Applications

——Modeling, Filtering, and Information Fusion Estimation

DENG ZILI

哈尔滨工业大学出版社
Harbin Institute of Technology Press
哈尔滨·**Harbin**

内 容 简 介

本书用邓自立教授独创的现代时间序列分析方法提出了关于系统状态或信号的最优估计和最优融合估计的新理论、新方法和新算法，并给出在目标跟踪系统中的仿真应用。

全书共分七章，包括时间序列 ARMA 模型和状态空间模型，最小二乘法参数估计，ARMA 时间序列预报，经典 Kalman 滤波理论及多传感器最优信息融合 Kalman 滤波理论，基于现代时间序列分析方法的最优滤波理论及最优信息融合滤波理论。内容新颖，理论严谨，并含有大量仿真例子。

本书可作为高等学校控制理论与控制工程、信号处理、检测与估计等专业的研究生及本科高年级学生教材，也可供在信号处理、控制、通信、航天、制导、雷达跟踪、石油地震勘探、故障诊断、卫星测控、GPS 定位、多传感器信息融合、机器人、经济、生物医学等领域工作的科技人员参考。

图书在版编目(CIP)数据

最优估计理论及其应用：建模、滤波、信息融合估计/邓自立著.—哈尔滨：哈尔滨工业大学出版社，2005.6

ISBN 7-5603-2152-6

Ⅰ.最…　Ⅱ.邓…　Ⅲ.最佳化-估计-理论　Ⅳ.0211.67

中国版本图书馆 CIP 数据核字(2005)第 041648 号

出版发行　哈尔滨工业大学出版社
社　　址　哈尔滨市南岗区复华四道街 10 号　邮编 150006
网　　址　http://hitpress.hit.edu.cn
传　　真　0451-86414049
印　　刷　肇东粮食印刷厂
开　　本　787×1092　1/16　印张 31.25　字数 708 千字
版　　次　2005 年 6 月第 1 版　2005 年 6 月第 1 次印刷
书　　号　ISBN 7-5603-2152-6/TP·215
印　　数　1～1 000
定　　价　45.00 元

前　　言

最优估计问题有三类。第一类是建模或系统辨识中的模型参数估计问题。第二类是时间序列、信号或状态预报、滤波和平滑问题，简称最优滤波问题。第三类是多传感器信息融合状态或信号估计问题，简称融合估计问题。

传统的最优估计理论主要包括最小二乘法和经典 Kalman 滤波理论。本书以作者独创的现代时间序列分析方法作为方法论，系统地提出了最优估计新理论、新方法和新算法及其在跟踪系统中的应用，构成了现代最优估计理论，可应用于信号处理、控制、通信、航天、制导、雷达跟踪、石油地震勘探、故障诊断、卫星测控、GPS 定位、多传感器信息融合、机器人、经济、生物医学等领域。

全书分为七章。第一章介绍 ARMA 模型和状态空间模型。第二章介绍最小二乘法及由作者提出的多种两段最小二乘法算法。第三章介绍经典 Kalman 滤波理论及其新进展，特别介绍了由作者提出的统一和通用的白噪声估计理论和时域 Wiener 方法。第四章介绍经典的 ARMA 时间序列预报方法，特别介绍了由作者提出的带观测噪声的 ARMA 时间序列预报方法及非平稳时间序列预报方法。第五章介绍由作者独创的最优滤波新的方法论——现代时间序列分析方法。第六章介绍由作者提出的在线性最小方差意义下的三种最优信息融合加权准则及基于经典 Kalman 滤波的多传感最优信息融合滤波理论。第七章介绍由作者提出的基于现代时间序列分析方法的多传感器最优信息融合滤波理论。书中给出了大量仿真例子和算例，特别以目标跟踪系统为应用背景，给出了许多信息融合跟踪滤波器的仿真例子，说明了所提出的结果的有效性和可应用性。

本书以现代时间序列分析方法、白噪声估计理论和所提出的多传感器最优信息融合滤波新理论为特色。

继 Kalman 滤波方法和 Wiener 滤波方法之后，现代时间序列分析方法是解决最优滤波问题新的方法论。作者的专著《最优滤波理论及其应用——现代时间序列分析方法》（哈尔滨：哈尔滨工业大学出版社，2000）、《卡尔曼滤波与维纳滤波——现代时间序列分析方法》（哈尔滨：哈尔滨工业大学出版社，2001）和《自校正滤波理论及其应用——现代时间序列分析方法》（哈尔滨：哈尔滨工业大学出版社，2003），以及本书构成了现代时间序列分析方法的完整的理论体系。它以 ARMA 新息模型和白噪声估计理论作为基本工具解决最优估计问题。该方法是作者于 1989 年在专著《现代时间序列分析及其应用——建模、滤波、去卷、预报和控制》（北京：知识出版社，1989）中提出的，已故中科院院士张钟俊教授曾给予高度评价（张钟俊．一门新兴边缘学科——现代时间序列分析．信息与控制，1988，17（4）：62～63）。同经典 Kalman 滤波方法和现代 Wiener 滤波方法相比，经典 Kalman 滤波方法是以 Riccati 方程作为基本工具解决最优估计问题，而现代 Wiener 滤波方法（即多项式方法）是以谱分解和 Diophantine 方程作为基本工具解决最优估计问题。对同一类滤波问

题，用三种方法论所得滤波器的公式在形式上可能是完全不同的，但它们是等价的，即它们在数值上是相同的。书中用大量数值仿真例子验证了这种等价性。作为新的方法论，用现代时间序列分析方法可以解决用 Kalman 滤波方法和 Wiener 滤波方法没有解决或不容易解决的许多最优和自校正估计或融合估计问题，显示了其强大的生命力。

多传感器信息融合是 20 世纪 70 年代后产生的一门新兴边缘学科。它随着电子技术和计算机应用技术的发展，特别是随着高科技武器（例如精确制导、远程打击及导弹拦截武器）的出现，应运而生，目前已发展成十分活跃的热门领域，广泛应用于指挥、控制、通信和情报系统。本书主要研究最优信息融合状态或信号估计问题。

本书系由作者主持的“多传感器信息融合最优和自校正滤波新理论和新方法”科研项目（国家自然科学基金资助项目。项目批准号：60374026；2004 年 1 月起至 2006 年 12 月止）的研究成果。信息融合的核心问题是最优融合规则的确定。本书提出了按矩阵加权、按对角阵加权和按标量加权三种分布式最优融合规则。解决最优融合估计问题的难点和关键技术在于如何计算各局部估计误差方差和互协方差。这些信息被用于计算最优加权。本书攻克了这一难题，提出了两种新的协方差信息融合滤波理论。

作者感谢国家自然科学基金委的资助。

作者深深感激已故中国科学院院士张钟俊教授生前对作者的鼓励和帮助。

还要感谢由作者指导的历届 50 余名研究生们，其中包括高媛、王欣、李云、毛琳、杜洪越、石莹、孟华等，他们对本书提出的新理论和新方法做了大量的仿真研究工作。

由于水平所限，书中缺点和疏漏之处在所难免，望读者批评指正。

著　者

2005 年元旦于哈尔滨

目　　录

绪　　论

最优估计理论要解决的问题有三类.第一类最优估计问题是模型参数估计问题.建立数学模型是对时间序列、信号或系统状态进行估计的基础.模型参数估计的最基本的方法是最小二乘法(Least Squares Method).由于它的原理直观,算法简单,收敛性能好,且不要求先验的统计知识,因而广泛被应用.最小二乘法是在1795年由大数学家高斯(C.F. Gauss)研究天体运动轨道问题时提出的.它的基本原理是实际观测值与模型计算值的误差的平方和最小原理,由此得名"最小二乘"法.最小二乘法原理的启发性例子如下.

【例0.1】　动态快速椭圆检测.[20]

在图像处理、机器人等领域,需要对运动图像中的椭圆曲线进行快速检测,这个问题类似于当年高斯提出用最小二乘法确定天体运动轨道.几何上这个问题归结为确定其标准型的五个参数,即椭圆中心位置(x_c, y_c),长短轴a,b和旋转角θ,见图0.1所示.

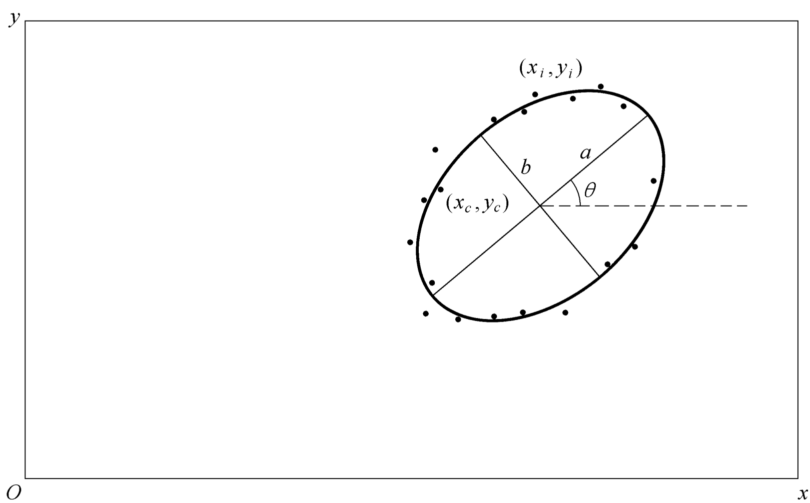


图0.1　椭圆曲线和检测点(x_i, y_i)

椭圆和其他二次曲线方程的一般形式为

$$x^2 + gxy + cy^2 + dx + ey + f = 0 \tag{0.1}$$

为了确定椭圆方程的未知参数(g, c, d, e, f),如果能精确地检测到椭圆曲线上五个点的坐标,则将它们代入上式即可得五个方程,解线性方程组即可得椭圆参数,然而通常检测椭圆上的点的坐标是带有测量误差的,通常是微小的随机误差,用上述方法只能粗略地得到椭圆参数,为此人们希望利用椭圆曲线上更多的点的坐标的检测得到较精确的椭圆参数估计.设已知椭圆上N个点的坐标的检测值(x_i, y_i)(含有检测误差),$i = 1, 2, \cdots, N$,将每组检测值(x_i, y_i)代入上式,则有方程误差ε_i,即

$$x_i^2 + gx_iy_i + cy_i^2 + dx_i + ey_i + f = \varepsilon_i, \quad i = 1, \cdots, N \tag{0.2}$$

方程误差 ε_i 是由于对椭圆上点的坐标的检测误差引起的.通常 N 远大于 5.最小二乘法原理就是用极小化方程误差的平方和来确定未知椭圆模型参数（g, c, d, e, f），即它们极小化性能指标

$$J = \sum_{i=1}^{N} \varepsilon_i^2 = \sum_{i=1}^{N} (x_i^2 + g x_i y_i + c y_i^2 + d x_i + e y_i + f)^2 \tag{0.3}$$

由极值原则，置

$$\frac{\partial J}{\partial g} = 0, \quad \frac{\partial J}{\partial c} = 0, \quad \frac{\partial J}{\partial d} = 0, \quad \frac{\partial J}{\partial e} = 0, \quad \frac{\partial J}{\partial f} = 0 \tag{0.4}$$

可得关于（g, c, d, e, f）的线性方程组，从而可解出（g, c, d, e, f），进而由有关公式可立刻求出标准型椭圆参数（a, b, x_c, y_c, θ）.

因为采用了极小化方程误差平方和的原理，因此得名“最小二乘”法.

【例 0.2】 对一个未知长度为 θ 的物体进行 N 次测量，设每次测量物体长度为 l_i，$i = 1, \cdots, N$，我们来求真实物体长度 θ 的估值.设每次测量误差为 ε_i，则有关系

$$l_i = \theta + \varepsilon_i, \quad i = 1, 2, \cdots, N \tag{0.5}$$

最小二乘法是选择 θ 的估值极小化测量误差平方和，即

$$\min J = \sum_{i=1}^{N} \varepsilon_i^2 = \sum_{i=1}^{N} (l_i - \theta)^2 \tag{0.6}$$

置 J 关于 θ 的偏导数为零，即

$$\frac{\partial J}{\partial \theta} = -2 \sum_{i=1}^{N} (l_i - \theta) = 0 \tag{0.7}$$

则有 θ 的最小二乘法估值为

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^{N} l_i \tag{0.8}$$

这是 N 次测量结果的算术平均值，与常识是一致的.

第二类最优估计问题是时间序列、信号或状态的最优估计问题.时间序列分析（Time Series Analysis）是概率统计学科中的一个重要分支，广泛应用于气象、水文、金融、经济、信号处理、通信和控制领域.时间序列分析的经典著作是 G.E.P.Box 和 G.M.Jenkins 的书《Time Series Analysis, Forecasting and Control》.[1]经典时间序列分析的主要内容是对时间序列的建模及基于时间序列模型对时间序列进行预报和控制。依离散时间顺序排列的观测序列 z_t: $z_1, z_2, \cdots, z_n, \cdots$，叫时间序列.例如，某地降雨量时间序列，我国国民经济年增长率时间序列，按天记股票价格时间序列，按秒采样导弹位置时间序列等.这些时间序列的取值均带有随机性，因而叫统计时间序列.

由时间序列目前和过去的观测历史 $z_t, z_{t-1}, z_{t-2}, \cdots$ 预报估计它的将来值 z_{t+k}（$k > 0$）叫预报，预报值记为 $\hat{z}_{t+k|t}$.例如气象预报（包括气温、降雨、降雪、沙尘暴等预报），水文预报（包括水位、洪峰、河流流量预报等），经济预报（包括商品销量、产量、经济指标、股市行情预报等），过程控制、目标跟踪、制导中的预报（包括温度、压力、体积、流量、产量、位置、速度等的预报）.在控制领域有一个新分支叫预测控制，就是以预报作为基础的控制理论.

最优预报 $\hat{z}_{t+k|t}$ 是指稳态线性最小方差预报，即理论上假设初始观测时刻 $t_o = -\infty$，即已知无限过去观测历史，且预报估值（也叫预报器）$\hat{z}_{t+k|t}$ 是由已知观测值 $z_t, z_{t-1}, \cdots$ 的线性组合构成，即

$$\hat{z}_{t+k|t} = \sum_{j=0}^{\infty} \alpha_j z_{t-j} \tag{0.9}$$

应选择系数 α_j 极小化预报均方误差

$$\min J = \mathrm{E}[(z_{t+k} - \hat{z}_{t+k|t})^2] \tag{0.10}$$

这个预报器 $\hat{z}_{t+k|t}$ 由 $z_t, z_{t-1}, \cdots$ 的无穷级数产生，从应用角度是不可取的，因为它要求存贮全部无限个历史数据. 为了克服这个缺点，人们给出了等价的递推预报器.

重要的最优预报方法有 Box - Jenkins[1] 的递推预报方法和 Åström[2] 的预报方法，其中 Box - Jenkins 递推预报器应用最广泛，但在理论研究中 Åström 预报器应用较多.

除了时间序列最优预报外还有信号和状态估计，也称最优滤波.

由被噪声污染的观测信号中，过滤噪声，求未知真实信号或状态最优估值叫滤波. "滤波"这一术语最初来自无线电领域.

1941 年，在第二次世界大战期间，以研究火炮打飞机控制系统为应用背景，控制论创始人 Wiener[3] 提出了信号的 Wiener 滤波理论. 经典 Wiener 滤波方法是一种频域方法，其局限性是限于处理平稳时间序列的滤波、预报问题. 缺点是不能处理多变量、时变、非平稳时间序列，且算法是非递推的，要求存贮全部历史数据，不便于工程应用. 但自 1979 年以来流行的现代 Wiener 滤波方法——多项式方法[16] 可处理多维非平稳时间序列滤波问题.

【例 0.3】 Wiener 滤波问题.

典型的 Wiener 信号滤波问题如图 0.2 所示. 其中未知真实信号 $s(t)$ 被观测噪声 $v(t)$ 污染，因而已知观测信号 $z(t)$，即

$$z(t) = s(t) + v(t) \tag{0.11}$$

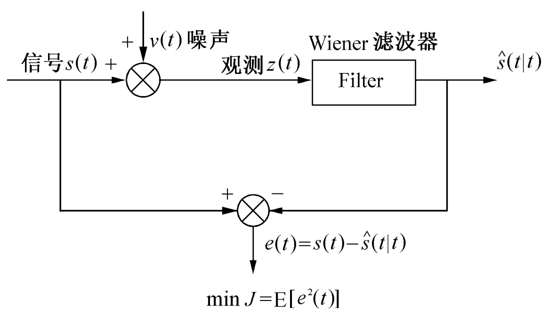


图 0.2　信号 Wiener 滤波问题

问题是如何由观测信号 $z(t)$ 中，过滤噪声 $v(t)$，在线性最小均方误差准则下，设计 Wiener 滤波器 $\hat{s}(t|t)$，它是 $z(t), z(t-1), \cdots$ 的线性函数，且极小化均方误差 $J = \mathrm{E}[e^2(t)]$，其中 E 为均值号，$e(t) = s(t) - \hat{s}(t|t)$ 为滤波误差.

图 0.3 为二维信号 $s_1(t)$ 和 $s_2(t)$ Wiener 滤波，其中图 (a) 中的 $s_i(t)$ 为原始信号，图 (b) 中 $v_i(t)$ 为观测噪声，图 (c) 中 $y_i(t)$ 为观测信号，完全淹没了真实信号，图 (d) 为信号 $s_i(t)$ 的 Wiener 滤波器 $\hat{s}_i(t|t)$，它有效地过滤了观测噪声，还信号 $s_i(t)$ 的本来面目.

1960 年美国数学家和控制论学者 Kalman[4] 针对 Wiener 滤波理论的上述缺点和局限性，以及由于电子技术和计算机应用技术发展的需要，提出了 Kalman 滤波理论（状态估计理论）. Kalman 滤波方法是一种时域方法，它基于状态空间模型和射影理论解决状态估计问题. Kalman 滤波算法是递推算法，便于在计算机上实现，且可处理多变量、时变、非平稳

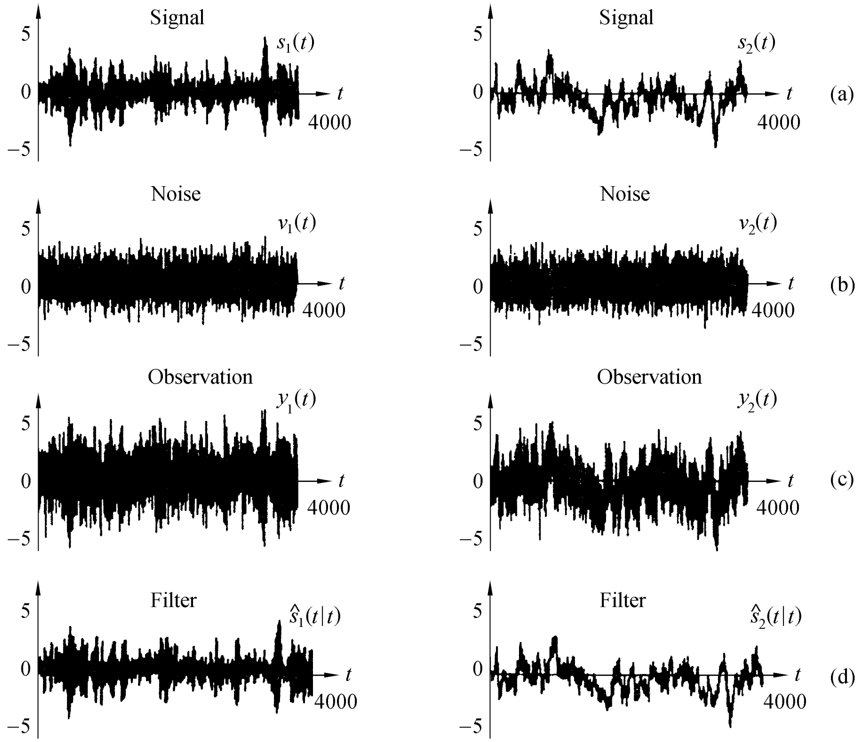


图 0.3 信号 Wiener 滤波问题仿真例子

时间序列滤波问题，克服了 Wiener 滤波理论的局限性. Kalman 滤波被广泛应用于各种领域，例如惯性导航、制导、GPS 定位、目标跟踪、通信、信号处理、控制等. 在 Kalman 滤波理论中，系统状态可视具体问题来规定和定义，特别信号也可视为状态或状态的分量，因而 Kalman 滤波也可解决信号滤波问题. 阿波罗登月计划和 C－5A 飞机导航系统的设计是 Kalman 早期应用中最成功的实例.

在 20 世纪 60 年代初由于电子计算机运算速度和存贮量的限制，要求能实时、快速实现滤波算法，要求存贮量小、计算量小的滤波算法. 满足这些要求的算法就是递推滤波算法. 以例 0.2 动态测量长度为 θ 的物体为例，记基于 N 个测量值对 θ 的估值为

$$\hat{\theta}(N) = \frac{1}{N}\sum_{i=1}^{N} l_i \tag{0.12}$$

当测量次数 N 不断增加，即进行动态测量时，则基于 $(N+1)$ 个测量值对 θ 的估值为

$$\hat{\theta}(N+1) = \frac{1}{N+1}\sum_{i=1}^{N+1} l_i \tag{0.13}$$

这种计算是非递推的，即彼此独立地计算估值 $\hat{\theta}(N)$ 和 $\hat{\theta}(N+1)$. 当 N 很大时，计算量增加，而且计算 $\hat{\theta}(N)$ 与计算 $\hat{\theta}(N+1)$ 有重复的加法运算. 为了减小计算负担，是否能在 $\hat{\theta}(N)$ 基础上来计算 $\hat{\theta}(N+1)$？这就是递推算法的思想. 事实上，

$$\hat{\theta}(N+1) = \frac{N}{N+1}\cdot\frac{1}{N}\left(\sum_{i=1}^{N} l_i + l_{N+1}\right) = \left(1-\frac{1}{N+1}\right)\hat{\theta}(N) + \frac{1}{N+1}l_{N+1} \tag{0.14}$$

即有递推公式

$$\hat{\theta}(N+1) = \hat{\theta}(N) + \frac{1}{N+1}[l_{N+1} - \hat{\theta}(N)] \tag{0.15}$$

因而在估值 $\hat{\theta}(N)$ 的基础上，只需计算上式第二项就立刻得到估值 $\hat{\theta}(N+1)$，避免了非递推算法（0.13）的重复加法运算，大大减小了计算量和存贮量．对于非递推算法，计算机需存贮 $(N+1)$ 个测量数据 l_i，而对递推算法（0.15），每次测量仅需存贮两个数据 $\hat{\theta}(N)$ 和 l_{N+1} 就可实现估值 $\hat{\theta}(N+1)$ 的计算．在（0.15）中第二项为校正量，它是根据误差

$$\varepsilon_{N+1} = l_{N+1} - \hat{\theta}(N) \tag{0.16}$$

的大小来进行校正估值 $\hat{\theta}(N)$ 的．因为估值 $\hat{\theta}(N)$ 已包含了前 N 次测量的信息，而 l_{N+1} 是第 $(N+1)$ 次测量值，估误差 $\varepsilon_{N+1} = l_{N+1} - \hat{\theta}(N)$ 包含了从第 $(N+1)$ 次测量中去掉了前 N 次测量的信息剩下的新的信息，故称为“新息”（Innovation）．于是我们最终得到新息校正形式的递推估值公式

$$\hat{\theta}(N+1) = \hat{\theta}(N) + K(N+1)\varepsilon_{N+1} \tag{0.17}$$

其中 $K(N+1) = 1/(N+1)$ 叫做新息校正系数或波滤增益．（0.17）是递推 Kalman 滤波算法的基本思想．

Kalman 滤波方法的基本特征和关键技术之一是状态空间模型．Kalman 滤波基于状态空间模型设计 Kalman 滤波器．而 Wiener 滤波理论采用的是传递函数模型．状态空间模型可用机理或物理、运动定律导出，也可用系统辨识方法得到．

Kalman 滤波问题可用如下启发性的例子来说明．

【例 0.4】 动态测量系统 Kalman 滤波问题，继例 0.2．

考虑例 0.2，对未知长度为 θ 的物体进行动态测量，即测量次数 t 是变化的，$t=1,2,\cdots,N,N+1,\cdots$．因长度 θ 为未知常数，故有 θ 的动态方程为

$$\theta(t+1) = \theta(t) \tag{0.18}$$

而第 t 次对 θ 的测量值 $l(t)$ 含有随机误差 $\varepsilon(t)$，故有对 θ 的观测方程为

$$l(t) = \theta(t) + \varepsilon(t) \tag{0.19}$$

我们可将未知长度 θ 定义为系统的状态，则（0.18）称为状态方程，它描写 θ 随 t 变化的规律，（0.18）说明长度 θ 不随 t 而变化，即 θ 为常数．而（0.19）则是对状态 θ 的观测方程，观测误差 $\varepsilon(t)$ 通常为零均值、方差为 σ_{ε}^2 的正态白噪声．（0.18）和（0.19）构成最简单的状态空间模型．Kalman 滤波问题：基于 t 次观测 $(l(1),\cdots,l(t))$ 求 θ 的线性最小方差估值 $\hat{\theta}(t)$．它与 θ 的最小二乘估值（0.8）不同的是：最小二乘估值（0.8）不要求已知观测误差 $\varepsilon(t)$ 的统计（均值和方差）．而实现 Kalman 滤波则要求已知这些统计知识．

【例 0.5】 雷达跟踪系统 Kalman 滤波问题．

考虑雷达跟踪系统，由运动定律有关系

$$s(t+1) = s(t) + \dot{s}(t)T + \frac{T^2}{2}w(t) \tag{0.20}$$

$$\dot{s}(t+1) = \dot{s}(t) + Tw(t) \tag{0.21}$$

$$y(t) = s(t) + v(t) \tag{0.22}$$

其中 T 为采样周期，$s(t)$，$\dot{s}(t)$ 和 $w(t)$ 各为在时刻 tT 运动目标（例如导弹、飞机、坦克、船舰、汽车等）的位置、速度和加速度．$y(t)$ 是对位置 $s(t)$ 的观测信号，$v(t)$ 为观测噪声．假设 $w(t)$ 和 $v(t)$ 都是白噪声（即不相关的随机序列）．问题是基于到时刻 t 为止的观测 $(y(t),y(t-1),\cdots,y(1))$ 求运动目标位置 $s(t)$ 和速度 $\dot{s}(t)$ 的最优估值 $\hat{s}(t\mid t)$ 和 $\hat{\dot{s}}(t\mid t)$．

定义系统的状态变量 $\boldsymbol{x}(t)$ 为

$$\boldsymbol{x}(t)=\begin{bmatrix} s(t) \\ \dot{s}(t) \end{bmatrix}=\begin{bmatrix} \text{位置} \\ \text{速度} \end{bmatrix} \tag{0.23}$$

则等价地有状态方程和观测方程

$$\begin{bmatrix} s(t+1) \\ \dot{s}(t+1) \end{bmatrix}=\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}\begin{bmatrix} s(t) \\ \dot{s}(t) \end{bmatrix}+\begin{bmatrix} \dfrac{T^2}{2} \\ T \end{bmatrix}w(t) \tag{0.24}$$

$$y(t)=[1\quad 0]\begin{bmatrix} s(t) \\ \dot{s}(t) \end{bmatrix}+v(t) \tag{0.25}$$

即我们有状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}\boldsymbol{x}(t)+\boldsymbol{\Gamma}w(t)\quad\text{(状态方程)} \tag{0.26}$$

$$y(t)=\boldsymbol{H}\boldsymbol{x}(t)+v(t)\quad\text{(观测方程)} \tag{0.27}$$

$$\boldsymbol{\Phi}=\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix} \dfrac{T^2}{2} \\ T \end{bmatrix},\quad \boldsymbol{H}=[1\quad 0] \tag{0.28}$$

上述状态方程和观测方程统称为状态空间模型，Kalman 滤波问题就是基于到时刻 t 观测 $(y(t),y(t-1),\cdots,y(1))$ 求状态 $\boldsymbol{x}(j)$ 的最优(线性最小方差)估值 $\hat{\boldsymbol{x}}(j|t)$，对 $j=t$，$j<t$ 或 $j>t$，分别称 $\hat{\boldsymbol{x}}(j|t)$ 为 Kalman 滤波器、平滑器或预报器.

【例 0.6】 宇宙飞船制导问题.[5]

下图表示宇宙飞船发射的雷达制导系统，目的是把载人航天器送至位于指定点 P 的特定轨道上去.首先把来自宇宙飞船的遥测数据及雷达跟踪数据(例如距离、距离变化率、俯仰角和方位角等测量数据)进行滤波处理，以估计出宇宙飞船的状态，然后把这些估计结果提供给具有一定算法的控制器，经过计算产生制导命令.最后按制导命令去遥控宇宙飞船.这个例子说明对系统进行状态估计的重要性.因为来自地面上或海面上对宇宙飞船的遥测数据是含有噪声的，必须进行滤波处理.

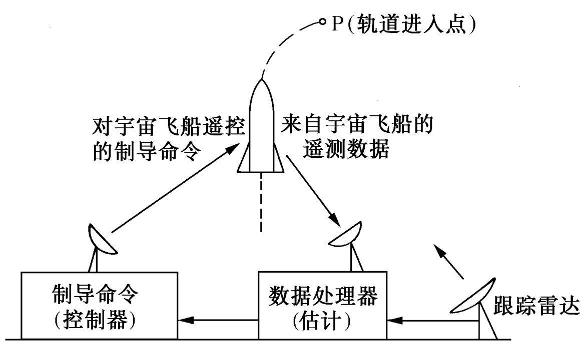


图 0.4　雷达制导方块图

第三类最优估计问题是最优信息融合估计问题.随着电子技术和计算机应用技术的发展和现代电子与信息战争及国防军事上的需要，为了提高对运动目标(导弹、飞机、卫星、坦克、车辆、船舰等)的跟踪精度或对动态系统状态的估计精度，大量涌现具有不同应用背景的多传感器系统.对目标跟踪而言，有各种类型测量运动目标位置、速度或加速度的传感器.特别是 20 世纪 70 年代后，由于高技术武器的出现，尤其是由于精确制导武器、

远程打击和导弹拦截武器的出现,使得依靠单传感器提供的信息很难满足目标跟踪或状态估计精度的要求,因此必须对每个传感器提供的信息按某种最优融合准则进行最优融合,才能提高对目标跟踪或状态估计的精度.早在20世纪70年代初美国海军就发现,对多个独立的声纳信号进行融合处理后,能更准确地探测出敌方潜艇的位置.这一发现对现代电子和信息战争产生了重大影响,早在1988年美国国防部就把信息融合技术列为90年代重点研究开发的二十项关键技术之一,且列为最优先发展的A类.近年来每年用于信息融合技术的研究费用达上亿美元.信息融合技术在海湾战争、科索沃战争及伊拉克战争中发挥了重要作用.在上述应用背景下,一门新兴学科——多传感器信息融合(Multi-sensor Information Fusion)应运而生,30多年来已发展成为倍受人们关注的热门领域,且在军事领域现代C^3I(指挥、控制、通信和情报)系统中得到广泛应用.

在例0.6中就遇到对宇宙飞船状态的融合估计问题,因为在地面上或海面上同时有多个遥测器(传感器)对飞船的状态进行观测.

【例0.7】 多传感器分布式信息融合Kalman滤波器原理.

对于状态估计而言,每个传感器可用一个观测方程代表,问题是对一个状态方程在有多个观测方程情形下,如何利用基于每个观测方程得到的局部Kalman滤波器进行加权融合得到在某种性能指标下的最优融合Kalman滤波器?它的精度应当比每个局部Kalman滤波器的精度高.多传感器分布式信息融合Kalman滤波器原理如图0.5所示.

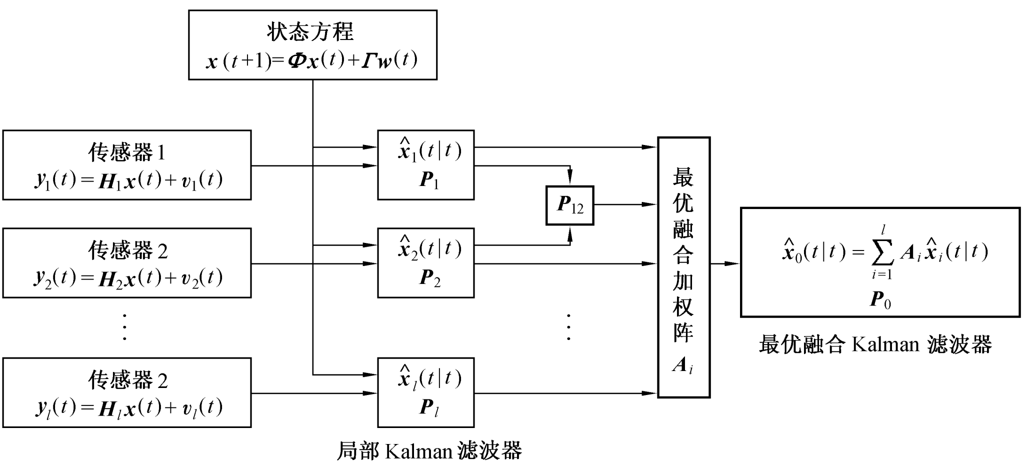


图0.5 多传感器分布式信息融合Kalman滤波器原理

图中$\hat{x}_i(t|t)$为对状态$x(t)$的局部Kalman滤波器,P_i为相应的局部滤波误差方差阵,P_{ij}为第i传感器滤波误差与第j传感器滤波误差的互协方差阵.这些方差阵P_i和互协方差阵P_{ij}信息被用于计算最优融合加权阵A_i。

对于第一类最优估计问题,本书重点介绍递推最小二乘法参数估计方法,特别介绍由作者提出的单变量和多变量ARMA模型的两段最小二乘法。此外,还介绍MA模型参数估计的Gevers-Wouters算法,并首次从理论上证明了该算法的一致性和指数收敛性.Gevers-Wouters算法是建立ARMA新息模型的重要工具,在现代时间序列分析方法[8,13~15]中起重要作用.

对于第二类最优估计问题,本书除了介绍如经典Kalman滤波方法和理论外,重点介

绍由作者提出的现代时间序列分析方法[8,13~15]，它为解决最优滤波问题提供了新的方法论，不同于基于 Riccati 方程的经典 Kalman 滤波方法[4,9~11,18].现代时间序列分析方法以 ARMA 新息模型作为基本工具解决状态或信号最优滤波问题.

对于第三类最优估计问题，本书分别介绍了首次由作者提出的基于经典 Kalman 滤波的多传感器协方差信息融合滤波理论和基于现代时间序列分析方法的多传感器协方差信息融合滤波理论.它们利用局部估计误差协方差信息来计算最优融合估计的加权阵，因此称为协方差信息融合滤波理论.书中以目标跟踪系统为应用背景，给出了大量的多传感器最优信息融合跟踪滤波器的仿真应用例子。

参 考 文 献

1 Box G E P，Jenkins G M. Time Series Analysis，Forecasting，and Control. San Francisco: Holden - Day，1970

2 Åström K J. 随机控制理论导论.潘裕焕译.北京：科学出版社，1983

3 Wiener N. Extrapolation，Interpolation，and Smoothing of Stationary Time Series with Engineering Application，New York: John Wiley & Sons，1949

4 Kalman R E. A New Approach to Linear Filtering and Prediction Problems. Trans. ASME，J. Basic Eng.，1961，83D:95 ~ 101

5 Meditch J S. 随机最优线性估计与控制.赵希人译.哈尔滨：黑龙江人民出版社，1981

6 何友，王国宾，陆大绘等.多传感器信息融合及应用.北京：电子工业出版社，2000

7 Chen C T. Linear System Theory and Design. New York: Holt，Rinehart and Winston，1984

8 邓自立.卡尔曼滤波与维纳滤波——现代时间序列分析方法.哈尔滨：哈尔滨工业大学出版社，2001

9 Lewis F L. Optimal Estimation. New York: John Wiley & Sonc，1986

10 Kamen E W，Su J K. Introduction to Optimal Estimation. Springer - Verlag London Limited，1999

11 Kailath T，Sayed A H，Hassibi B. Linear Estimation. New Jersey: Prentice - Hall，Inc.，2000

12 Mendel J M. Lessons in Estimation Theory for Signal Processing，Communications，and Control. New Jersey: Prentice - Hall，1995

13 邓自立.最优滤波理论及其应用——现代时间序列分析方法.哈尔滨：哈尔滨工业大学出版社，2000

14 邓自立.自校正滤波理论及其应用——现代时间序列分析方法.哈尔滨：哈尔滨工业大学出版社，2003

15 邓自立，郭一新.现代时间序列分析及其应用——建模、滤波、去卷、预报和控制.北京：知识出版社，1989

16 付梦印，邓志江，张继伟.Kalman 滤波理论及其在导航系统中的应用.北京：科学出版社，2003

17 Ahlen A，Sternad M. Wiener Filter Design Using Polynomial Equations. IEEE Trans. Signal Processing，1991，39(11):2387 ~ 2399

18 Abderson B D O，Moore J B. Optinal Filtering. New Jersey: Prentice - Hall，Inc.，1979

19 Söderström T. Discrete - Time Stochastic Systems，Estimation and Control. New York: Prentice Hall，1994

20 王忠立，高文.基于最小二乘预测的动态快速椭圆检测.信息与控制，2003，32(7):729 ~ 733

第一章　ARMA 模型和状态空间模型

1.1　引　　言

时间序列最优预报的基本数学模型是自回归滑动平均(Autoregressive Moving Average)模型,简称 ARMA 模型.ARMA 模型不仅是时间序列预报的基本工具和方法论,而且也是信号与状态最优估计的基本模型.在信号 Wiener 滤波问题中,被估信号和观测噪声,以及观测信号均为随机信号[1],通常用 ARMA 模型来描写.ARMA 模型可以描写一大类平稳和非平稳信号[2].

时间序列 z_t 的 ARMA 模型具有形式:

$$z_t - \varphi_1 z_{t-1} - \cdots - \varphi_p z_{t-p} = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \tag{1.1.1}$$

其中 a_t 为零均值、方差为 σ_a^2 的不相关随机序列,称为白噪声,即

$$\mathrm{E} a_t = 0, \quad \mathrm{E}[a_t a_s] = \sigma_a^2 \delta_{ts} \tag{1.1.2}$$

其中 E 为均值号,$\delta_{tt}=1$,$\delta_{tj}=0\,(t\neq j)$,而 φ_i 和 θ_i 为模型参数,p,q 为模型的阶次.

由(1.1.1)看到 z_t 与过去的历史 $z_{t-1},\cdots,z_{t-p}$ 有关,这就是“自回归”一词的直观含义.而(1.1.1)右端的随机项恰好为白噪声项 $a_t, a_{t-1},\cdots,a_{t-q}$ 的加权(线性组合),这就是“滑动平均”或“滑动和”一词的直观含义.用 ARMA 模型描写一个时间序列,关键在于确定模型参数 φ_i,θ_i 和阶次 p,q,以及白噪声方差 σ_a^2.本书第二章将介绍 ARMA 模型参数 φ_i,θ_i 和 σ_a^2 估计的递推最小二乘法及其改进算法.

状态空间模型是 Kalman 滤波或状态估计的基本数学模型,类似于绪论中例 0.5 雷达跟踪系统的状态空间模型.一般的离散时间随机系统的状态空间模型具有如下形式:

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)\boldsymbol{u}(t) + \boldsymbol{\Gamma}(t)\boldsymbol{w}(t) \tag{1.1.3}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}(t)\boldsymbol{x}(t) + \boldsymbol{v}(t) \tag{1.1.4}$$

其中系统状态 $\boldsymbol{x}(t)\in R^n$,t 为离散时间,$\boldsymbol{y}(t)\in R^m$ 为观测信号,$\boldsymbol{v}(t)\in R^m$ 为观测噪声,$\boldsymbol{w}(t)\in R^r$ 为输入白噪声,$\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 是零均值、方差阵各为 $\boldsymbol{Q}(t)$ 和 $\boldsymbol{R}(t)$ 的独立的或相关的白噪声,$\boldsymbol{u}(t)\in R^s$ 是已知的控制输入.$\boldsymbol{\Phi}(t)$,$\boldsymbol{B}(t)$,$\boldsymbol{\Gamma}(t)$,$\boldsymbol{H}(t)$,$\boldsymbol{Q}(t)$,$\boldsymbol{R}(t)$ 各为 $n\times n$,$n\times s$,$n\times r$ 和 $m\times n$,$r\times r$ 和 $m\times m$ 时变矩阵.因而称(1.1.3)和(1.1.4)为时变系统,$\Phi(t)$ 被称为状态转移阵,$H(t)$ 被称为观测阵.若 $\boldsymbol{\Phi}$,$\boldsymbol{B}$,$\boldsymbol{\Gamma}$,$\boldsymbol{H}$,$\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 均为常阵,则称(1.1.3)和(1.1.4)为时不变系统或定常系统.Kalman 滤波问题就是基于被观测噪声 $\boldsymbol{v}(t)$ 污染的观测信号 $\boldsymbol{y}(t)$,$\boldsymbol{y}(t-1)$,$\cdots$,和控制 $\boldsymbol{u}(t-1)$,$\boldsymbol{u}(t-2)$,$\cdots$,求系统在时刻 j 处的状态 $\boldsymbol{x}(j)$ 的最优估计 $\hat{\boldsymbol{x}}(j|t)$.

线性系统理论[7]证明了 ARMA 模型与状态空间模型是可以相互转化的.这种相互转化是解决具体最优估计问题的有力杠杆.可以将 ARMA 信号最优滤波问题转化为状态估

计问题,也可将状态估计问题转化为构造 ARMA 新息模型[1].

本章介绍 ARMA 模型和状态空间模型,以及它们的基本性质及相互转化.

1.2 随机过程

本书所研究的时间序列是统计时间序列,即随机过程.随时间顺序演化的随机现象叫随机过程,若时间是离散的,则也称为随机序列.

1.2.1 随机过程概念

【例 1.2.1】 电网电压随机过程.

我们考察在相同一段时间 T 内电网电压 $X(t)$ 的波动,虽然标准电压是 220 V,但实际电压却随机地在 220 V 上下波动,这是因为各用电单位、各家庭、各用电的人的用电时间、用电量都带有随机性,造成电网电压的随机波动,每做一次观察可得一条普通的定义在 T 上的电压波动曲线 $x(t)$,叫样本函数或样本曲线,也叫实现.电压波动随机过程 $X(t)$ 可看成是它的所有可能的定义在时间区间 T 上的实现族 $\{x(t)\}$ 或样本函数族.样本族可能有无穷个样本曲线.从另一观点来看,对于任意固定时刻 $t_1\in T$,电压值 $X(t_1)$ 是随机变量.因为在这一时刻没到来之前,人们不可能预知在 t_1 时刻的电压的准确值.因此当我们在时间区间 T 上做一次观察时,由于每一随机变量 $X(t)$,$t\in T$,均有一个具体的观测值,这些观测值构成了在时间 T 上的一条样本曲线.

【定义 1.2.1】 随机过程 $X(t)$,$t\in T$,是定义在时间集合 T 上的随机变量族,记为 $\{X(t),t\in T\}$,等价地,随机过程 $X(t)$ 是定义在时间集合 T 上的所有样本函数族 $\{x(t),t\in T\}$.

时间 T 可以是区间,也可以是离散时间集合,可以是有限集合,也可以是无限集合.例如常用的时间集合有

$$T=[a,b],\quad T=(-\infty,+\infty),$$
$$T=\{1,2,\cdots,n\},\quad T=[0,1,2,\cdots],\quad T=\{\cdots,-2,-1,0,1,2,\cdots\}$$

当 T 为离散时间集时,也称随机过程 $X(t)$ 为随机序列.

从随机过程观点出发,时间序列可定义如下.

【定义 1.2.2】 依时间次序排列的随机变量序列 Z_t,$t\in T$,T 为离散时间集合,称为时间序列.如果用

$$z_1,z_2,\cdots,z_N \tag{1.2.1}$$

分别表示相应随机变量 $Z_1,Z_2,\cdots,Z_N$ 的观测值,则称其为时间序列 Z_t 的一个容量为 N 的样本,也称长度为 N 的样本.

在实际应用问题中人们能得到的只是时间序列的有限的观测样本,甚至在某些应用问题中只能得到时间序列的一个长度为 N 的有限的样本.例如,自有记录以来某地年降雨量、某地年最高气温等时间序列,因为时间是不可倒退的,人们只能得到一个有限长度的观测样本.时间序列分析的方法论就是根据它的一个有限长样本建立它的统计模型,然后利用模型实现预报、滤波或控制的目的.

1.2.2 随机过程的数学期望(均值)函数、方差函数和相关函数

随机过程的数学期望是随机变量数学期望(均值)的推广,它由随机过程在每时刻的均值构成的定义在时间区间 T 上的均值曲线来定义,它从总体上刻画随机过程取值的平均.

【定义 1.2.3】 已给随机过程 $X(t), t\in T$,它的数学期望(均值)定义为在 T 上的一个确定的时间函数,记为 $\mathrm{E}X(t)$或 $m(t)$,

$$m(t)=\mathrm{E}[X(t)], \quad t\in T \tag{1.2.2}$$

其中 E 为数学期望符号.

【定义 1.2.4】 已给随机过程 $X(t), t\in T$,它的方差函数定义为

$$\sigma^2(t)=D[X(t)]=\mathrm{E}[(X(t)-m(t))^2], \quad t\in T \tag{1.2.3}$$

它是在 T 上的确定性的时间函数,其中 D 为方差符号,$\sigma(t)$叫标准方差函数.

方差函数刻划了随机过程 $X(t)$偏离其均值 $m(t)$的误差的平方的平均状况.

1.2.3 随机过程的相关函数

数学期望和方差描写了随机过程在任意时刻 t 的集中和离散程度,而随机过程的相关函数则反映随机过程在任意两个不同时刻相应随机变量之间的联系.相关性概念表征了随机过程在两时刻之间的关联程度,进而说明随机过程波动的快慢.

【定义 1.2.5】 已给随机过程 $X(t), t\in T$,它的相关函数 $R(t_1,t_2), t_1\in T, t_2\in T$,定义为

$$R(t_1,t_2)=\mathrm{E}[(X(t_1)-m(t_1))(X(t_2)-m(t_2))] \tag{1.2.4}$$

1.2.4 平稳随机过程

有一类随机过程(例如电网电压、海浪波动幅度等)的样本函数具有如下特征,每个样本函数都在某一固定值附近波动,且波动的平均偏差和快慢不变,即它的统计特性不随时间推移而改变.这类随机过程叫平稳随机过程.

【定义 1.2.6】 已给随机过程 $X(t), t\in T$,若

$$E[X(t)]=\text{常数}=m \tag{1.2.5}$$

且相关函数 $R(t_1,t_2)$仅是 $\tau=t_2-t_1$ 的函数,记为

$$R(t_1,t_2)=R(t_2-t_1)=R(\tau) \tag{1.2.6}$$

则称 $X(t)$为广义平稳随机过程,简称平稳随机过程.

【推论 1.2.1】 平稳随机过程有常数方差,即

$$\sigma^2(t)=\text{常数}=\sigma^2 \tag{1.2.7}$$

证明 对任意 $t\in T$ 有 $\sigma^2(t)=\mathrm{E}[(X(t)-m)^2]=R(t,t)=R(0)=\sigma^2$ □

不满足(1.2.5)和(1.2.6)的随机过程叫非平稳随机过程.

在现实中会遇到大量平稳和非平稳过程.常常一个随机过程在开始阶段和结尾阶段呈现非平稳性,而在中间阶段具有平稳性,例如地震波.导弹飞行轨迹随目标而定,不是平稳过程.股市行情曲线总体上呈非平稳性(带有一定趋势性),机械振动开始和结束阶段是非平稳的,中间阶段是平稳的.人们利用平稳性和非平稳性的转化可以进行故障诊断.

1.2.5 平稳随机过程的遍历性

对于随机过程的数学期望和相关函数的近似计算,涉及到大量样本的统计平均,能否寻求更简单的计算方法?对于一个平稳随机过程,能否由它的一个样本函数取时间平均来计算它的均值和相关函数?这要看这个样本函数是否具有代表性?是否能充分体现过程的统计特性.若平稳随机过程的每个样本函数都经历了它的各种可能状态,能充分代表过程的统计特性,则称它为各态历经平稳随机过程.

用小写 $x(t)$ 表示平稳随机过程 $X(t)$ 的一个样本函数.

【定义 1.2.7】 已给平稳随机过程 $X(t)$,$t\in[0,\infty]$,$x(t)$ 是它的任一个样本函数,若

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)\mathrm{d}t=\mathrm{E}X(t)=m \tag{1.2.8}$$

以概率 1 成立,则称 $X(t)$ 的均值具有遍历性或各态历经性.若以概率 1 成立

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)x(t+\tau)\mathrm{d}t=\mathrm{E}[X(t)X(t+\tau)] \tag{1.2.9}$$

则称 $X(t)$ 的相关函数具有遍历性.若 $X(t)$ 的均值和相关函数均具有遍历性,则称 $X(t)$ 是各态历经平稳过程.

【定义 1.2.8】 已给平稳随机序列 $X(t)$,$t\in\{1,2,\cdots\}$,若 $x(t)$ 是它的一个样本序列,若

$$\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}x(i)=\mathrm{E}X(t)=m \tag{1.2.10}$$

$$\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}x(i)x(i+\tau)=\mathrm{E}[X(t)X(t+\tau)] \tag{1.2.11}$$

以概率 1 成立,则称它具有遍历性.

在较弱的条件下可证明平稳过程具有遍历性,大量的现实平稳过程均可近似看成各态历经过程.例如某地年降雨量、心电图、脑电图、机械振动、海浪、电网电压、量测误差等.

1.2.6 平稳随机序列

今后我们主要研究离散时间随机过程,即随机序列,记为 z_t,$t\in T=\{\cdots,-1,0,1,2,\cdots\}$,时间集 T 遍历了从 $-\infty$ 到 $+\infty$ 的所有整数.应注意,在不引起混淆情形下,今后将不再用大字母 Z_t 表示随机序列 $\{Z_t,t\in T\}$,用小写字母 z_t 表示它的实现(样本)$\{z_t,t\in T\}$,而统一用小写字母 z_t 表示随机序列及其实现。这里 $\{z_t,t\in T\}$ 既可表示随机序列,也可表示它的实现.$(z_1,z_2,\cdots,z_N)$ 既可看成随机向量,也可看成 $\{z_t,t\in T\}$ 的容量为 N 的样本.平稳随机序列的相关函数记为

$$\gamma_k=\mathrm{E}[(z_t-m)(z_{t+k}-m)],\quad m=\mathrm{E}z_t \tag{1.2.12}$$

不失一般性,我们可假设平稳随机序列 z_t 的均值为零,即

$$m=\mathrm{E}[z_t]=0 \tag{1.2.13}$$

标准相关函数定义为

$$\rho_k = \frac{\gamma_k}{\gamma_0} \tag{1.2.14}$$

其中 $\gamma_0 = \mathrm{E}[z_t^2]$为 z_t 的方差，即 $\gamma_0 = \sigma^2$，相关函数 γ_k 具有如下性质：

⑴ $\gamma_0 \geqslant 0$ (1.2.15)

⑵ $\gamma_k = \gamma_{-k}$（对称性） (1.2.16)

⑶ $|\gamma_k| \leqslant \gamma_0$ (1.2.17)

⑷对任意自然数 n 和不全为零的实数 $l_1, l_2, \cdots, l_n$，有

$$\sum_{i,j=1}^{n} l_i l_j \gamma_{i-j} \geqslant 0 \tag{1.2.18}$$

即如下矩阵 Γ_n 为非负定阵

$$\Gamma_n = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \vdots & & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0 \end{bmatrix} \geqslant 0 \tag{1.2.19}$$

证明

$$\gamma_0 = \mathrm{E}[z_t^2] \geqslant 0 \tag{1.2.20}$$

$$\gamma_k = \mathrm{E}[z_t z_{t+k}] = \mathrm{E}[z_{t+k} z_t] = \gamma_{-k} \tag{1.2.21}$$

$$|\gamma_k| = |\mathrm{E}[z_t z_{t+k}]| \leqslant \sqrt{\mathrm{E}[z_t^2] \cdot \mathrm{E}[z_{t+k}^2]} = \gamma_0 \tag{1.2.22}$$

令 $L_t = l_1 z_t + \cdots + l_n z_{t-n+1}$，则有

$$\mathrm{E}[L_t^2] = \sum_{i,j=1}^{n} l_i l_j \gamma_{|i-j|} \geqslant 0 \tag{1.2.23}$$

易知标准相关函数 ρ_k 有性质：

⑴ $\rho_0 = 1$ (1.2.24)

⑵ $|\rho_k| \leqslant 1, k = 0, 1, \cdots$ (1.2.25)

⑶ $\rho_k = \rho_{-k}$ (1.2.26)

⑷ $$\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & & \rho_{n-2} \\ \vdots & \vdots & & \vdots \\ \rho_{n-1} & \rho_{n-2} & \cdots & 1 \end{bmatrix} \geqslant 0 \tag{1.2.27}$$

【定义 1.2.9】 对于平稳随机序列 z_t 长度为 N 的一个样本 $(z_1, z_2, \cdots, z_N)$，定义：

采样均值 $$\bar{z} = \frac{1}{N} \sum_{t=1}^{N} z_t \tag{1.2.28}$$

采样方差 $$\hat{\sigma} = \frac{1}{N} \sum_{t=1}^{N} (z_t - \bar{z})^2 \tag{1.2.29}$$

采样相关函数 $$\hat{\gamma}_k = \frac{1}{N} \sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z}), \quad k = 0, 1, 2, \cdots \tag{1.2.30}$$

采样标准相关函数 $$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}, \quad k = 0, 1, 2, \cdots \tag{1.2.31}$$

假如 z_t 具有遍历性，则由随机过程理论当 $N \to \infty$ 以概率 1 成立

$$\bar{z} \to m, \quad \hat{\sigma}^2 \to \sigma^2, \quad \hat{\gamma}_k \to \gamma_k, \quad \hat{\rho}_k \to \rho_k \tag{1.2.32}$$

【定义 1.2.10】 设 $a_t, t \in T=\{\cdots,-1,0,1,\cdots\}$，是带零均值、方差为 σ^2 的不相关随机序列，即

$$\mathrm{E}a_t=0, \quad \mathrm{E}[a_t a_s]=\sigma^2\delta_{ts} \tag{1.2.33}$$

其中 $\delta_{tt}=1, \delta_{ts}=0\,(t\neq s)$，则称 a_t 是白噪声.

显然，白噪声相关函数为

$$\gamma_k=\mathrm{E}[z_t z_{t+k}]=\begin{cases}\sigma^2 & k=0\\ 0 & k\neq 0\end{cases} \tag{1.2.34}$$

白噪声的标准相关函数为

$$\rho_k=\begin{cases}1 & k=0\\ 0 & k\neq 0\end{cases} \tag{1.2.35}$$

因此白噪声是最简单的平稳随机序列.

【定义 1.2.11】 如果一个白噪声序列 a_t 同时又是正态随机序列，即在每时刻 t，a_t 均服从正态分布，则称其为正态白噪声.

白噪声在理论和应用上具有重要意义. 在理论上，用白噪声的线性运算可构造一般的平稳时间序列模型；在应用上，许多系统的观测噪声是白噪声，例如仪器或仪表的量测噪声. 图 1.2.1 分别给出了正态 $N(0,1)$ 白噪声和正态 $N(0,0.64)$ 白噪声的一个样本曲线.

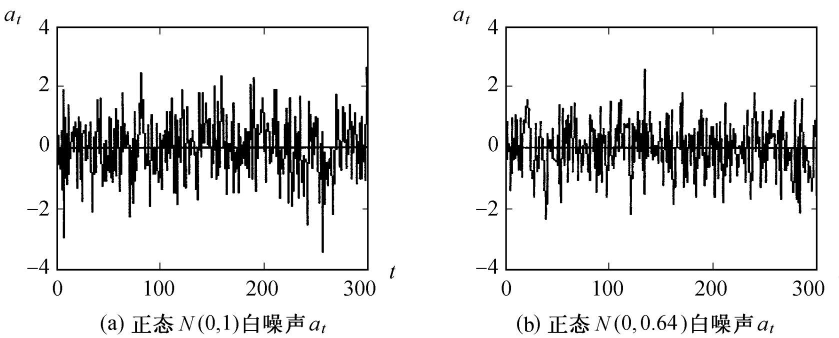


图 1.2.1　正态白噪声样本

1.3　自回归滑动平均（ARMA）模型

对时间序列估计（预报和滤波）的基本方法就是建模. 时间序列的预报器和滤波器的设计是基于它的模型. 本节介绍应用最广泛的时间序列模型——自回归滑动平均（ARMA）模型. 它可由白噪声的线性运算生成.

1.3.1　自回归滑动平均模型

【例 1.3.1】 股市行情建模和预报问题.

每天股票价格的波动带有很大随机性，通常它也呈现出一定趋势性，但这种趋势也是随机的和变化莫测的. 股票价格波动的幅度常常也是很大的. 因而股票价格时间序列本质上是一个非平稳时间序列. 用建模方法进行预报成为一个有较大难度的问题. 关键在于建立何种模型. 有一种经验性预报方法是“对明天股票价格最好的预报是今天股票的价格”，

这等价于用如下随机游动模型描写股票价格时间序列 z_t，即

$$z_t = z_{t-1} + a_t \tag{1.3.1}$$

其中 a_t 是白噪声.这是说，明天股票价格 z_{t+1}等于在今天股票价格 z_t 基础上加上随机波动 a_{t+1}.因为白噪声是不可预测的，因此由上面的随机游动模型引出明天股票价格最优预报值 $\hat{z}_{t+1|t}$为

$$\hat{z}_{t+1|t} = z_t \tag{1.3.2}$$

即“对明天股票价格最优的预报就是今天股票的价格”.因为模型(1.3.1)描写了时间序列现在的值与它自身过去的值之间的关系，故称(1.3.1)为简单的自回归模型.因为现在的值 z_t 仅与前一时刻值z_{t-1}有关，故称其为一阶自回归模型。“自回归”顾名思义就是回顾时间序列自身的历史，考虑它现在和过去的联系.

Box 和 Jenkins[2]考察了美国 IBM 普通股 1961 年 5 月 17 日至 1962 年 11 月 2 日的收盘价格 z_t 有 369 个数据，它呈现带随机趋势性的非平稳性，为了消除趋势性，化为平稳序列，引入一阶差分序列为

$$y_t = z_t - z_{t-1} \tag{1.3.3}$$

计算 y_t 的采样相关函数$\hat{\rho}_k$，发现 $\hat{\rho}_0 = 1$，$\hat{\rho}_1 \neq 0$，但对 $k > 1$，$\hat{\rho}_k$ 的绝对值很小，因而可近似认为 $\hat{\rho}_k = 0\,(k > 1)$.因此可认为 y_t 是一个一步相关序列，即可用模型

$$y_t = a_t + \theta a_{t-1} \tag{1.3.4}$$

描写 y_t，这里 a_t 是零均值、方差为 σ^2 的白噪声，θ 为模型参数.容易验证模型(1.3.4)的相关函数 ρ_k 在 $k=1$ 处截尾，因而它是一步相关序列.我们称模型(1.3.4)为一阶滑动平均模型，它由白噪声 a_t 和 a_{t-1}的线性组合生成，叫“滑动和”或“滑动平均”或“加权和”.

将(1.3.3)代入(1.3.4)可得改进的股票价格时间序列 z_t 的模型

$$z_t - z_{t-1} = a_t + \theta a_{t-1} \tag{1.3.5}$$

称(1.3.5)为一阶自回归滑动平均模型.它的左边含有自回归项，它的右边含有滑动平均项.它改进了随机游动模型(1.3.1).

在许多时间序列预报问题中，例如气象和水文预报，我们要预报某地明年年降雨量.周知，明年年降雨量通常不仅与今年降雨量有关，而且与去年和前年降雨量均有关，这引出如下 n 阶自回归模型.

【定义 1.3.1】 若时间序列 z_t 有模型

$$z_t = \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \cdots + \varphi_p z_{t-p} + a_t \tag{1.3.6}$$

其中模型线差 a_t 是零均值、方差为 σ^2 的白噪声，$\varphi_1, \cdots, \varphi_p$ 为模型参数，则称式(1.3.6)为 p 阶自回归模型，记为 AR(p)，p 为它的阶次.

将非平稳序列用差分方法化为平稳序列后，这个平稳序列的相关函数有时是 q 步相关的，即它的相关函数 ρ_k 具有性质$\rho_k \neq 0$，$k = 0, 1, \cdots, q$；$\rho_k = 0$，$k > q$，即它的相关函数在 $k = q$ 处截尾，这种平稳序列可用如下滑动平均模型来描写.

【定义 1.3.2】 若时间序列 z_t 有模型

$$z_t = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \tag{1.3.7}$$

其中 a_t 是零均值、方差为 σ^2 的白噪声，$\theta_1, \cdots, \theta_q$ 是模型参数，则称式(1.3.7)为 q 阶滑动平均模型，记为 MA(q)，q 为它的阶次.

容易验证 z_t 的相关函数 ρ_k 在 $k=q$ 处截尾.

推广模型(1.3.5)有一般的自回归滑动平均(Autoregressive Moving Average)模型的定义.

【定义 1.3.3】 若时间序列 z_t 有模型

$$z_t - \varphi_1 z_{t-1} - \cdots - \varphi_p z_{t-p} = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \tag{1.3.8}$$

其中 a_t 是零均值、方差为 σ^2 的白噪声,则称(1.3.8)为自回归滑动平均模型,简记 ARMA(p,q)模型,p,q 为它的阶次.

引入单位滞后算子 q^{-1},$q^{-1}x_t = x_{t-1}$,$q^{-i}x_t = x_{t-i}$,定义自回归和滑动平均算子多项式分别为

$$\begin{aligned}\varphi(q^{-1}) &= 1 - \varphi_1 q^{-1} - \cdots - \varphi_p q^{-p},\\ \theta(q^{-1}) &= 1 - \theta_1 q^{-1} - \cdots - \theta_q q^{-q}\end{aligned} \tag{1.3.9}$$

则 ARMA(p,q)模型(1.3.8)可简写为

$$\varphi(q^{-1})z_t = \theta(q^{-1})a_t \tag{1.3.10}$$

AR(p)模型(1.3.6)可写为

$$\varphi(q^{-1})z_t = a_t \tag{1.3.11}$$

MA(q)模型(1.3.7)可写为

$$z_t = \theta(q^{-1})a_t \tag{1.3.12}$$

1.3.2 ARMA 模型的平稳性

【例 1.3.2】 AR(1)模型的平稳性.

考虑 AR(1)模型

$$z_t = \varphi z_{t-1} + a_t \tag{1.3.13}$$

其中 a_t 是零均值、方差为 σ_a^2 的白噪声.

在什么条件下,AR(1)模型(1.3.13)描写一个平稳随机序列?(1.3.13)可写为

$$(1-\varphi q^{-1})z_t = a_t \tag{1.3.14}$$

$$z_t = \frac{1}{1-\varphi q^{-1}}a_t \tag{1.3.15}$$

应用几何级数求和公式有无穷级数展式

$$z_t = \sum_{j=0}^{\infty}\varphi^j q^{-j}a_t = \sum_{j=0}^{\infty}\varphi^j a_{t-j} \tag{1.3.16}$$

这相当于把 z_t 表为无穷阶次滑动平均模型 MA(∞).

要使(1.3.16)代表一个平稳序列,它必须在均方收敛意义下收敛,这要求如下数值级数收敛,即

$$\sum_{j=0}^{\infty}\varphi^{2j} < \infty \tag{1.3.17}$$

这等价于要求

$$|\varphi| < 1 \tag{1.3.18}$$

注意本例自回归多项式,$\varphi(q^{-1}) = 1 - \varphi q^{-1}$,这等价于要求 $\varphi(x)$的零点位于单位圆外.

当$|\varphi|<1$ 时,我们有级数(1.3.16)均方收敛.因为由均方收敛的 Cauchy 准则,当 $m\to\infty$,$n\to\infty$ 时有

$$\mathrm{E}\Big(\sum_{j=m}^{n}\varphi^{j}a_{t-j}\Big)^{2} = \sigma_a^2\sum_{j=m}^{n}\varphi^{2j}\to 0 \tag{1.3.19}$$

其中 E 为数学期望符号.于是我们有

$$\mathrm{E}z_t = \mathrm{E}\sum_{j=0}^{\infty}\varphi^{j}a_{t-j} = \sum_{j=0}^{\infty}\varphi^{j}\mathrm{E}a_{t-j} = 0,$$

$$\gamma_{\mathrm{k}} = \mathrm{E}\left[z_t z_{t+k}\right] = \mathrm{E}\Big(\sum_{j=0}^{\infty}\varphi^{j}a_{t-j}\Big)\Big(\sum_{i=0}^{\infty}\varphi^{i}a_{t+k-i}\Big) =$$

$$\sigma_a^2\sum_{j=0}^{\infty}\varphi^{j}\varphi^{j+k} = \sigma_a^2\varphi^{k}\sum_{j=0}^{\infty}\varphi^{2j} = \frac{\sigma_a^2\varphi^{k}}{1-\varphi^{2}} \tag{1.3.20}$$

这证明了在 $|\varphi|<1$ 条件下,z_t 是一个带零均值、相关函数为 γ_k 的平稳序列.取 $k=0$ 可得 z_t 的方差为

$$\sigma_z^2 = \gamma_0 = \frac{\sigma_a^2}{1-\varphi^2} \tag{1.3.21}$$

从而有 z_t 的标准相关函数为

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \varphi^k \tag{1.3.22}$$

可证明:对一般的 ARMA 模型(1.3.10),它是平稳的(即它描写一个平稳随机序列),充分条件是多项式 $\varphi(x)$ 的零点位于单位圆外,称这样的 $\varphi(q^{-1})$ 是稳定的多项式.

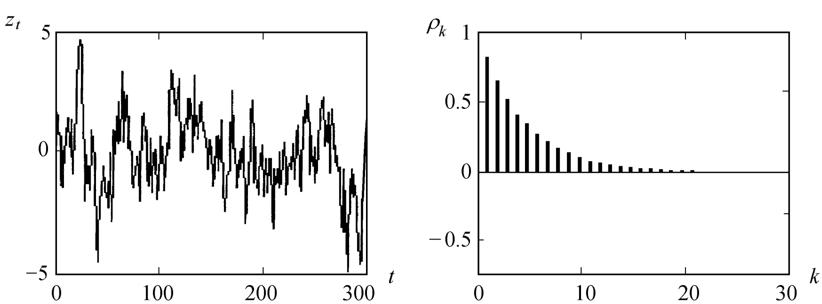
【例 1.3.3】 考虑 AR(1)模型

$$z_t = 0.8z_{t-1} + a_t \tag{1.3.23}$$

其中 a_t 是零均值、方差为 $\sigma_a^2=1$ 的正态白噪声.由式(1.3.22)有 z_t 的标准相关函数为

$$\rho_k = (0.8)^k \tag{1.3.24}$$

z_t 的一个样本函数和 z_t 的相关函数图形如图 1.3.1 所示.



(a) 样本函数 z_t　　(b) 标准相关函数 ρ_k

图 1.3.1　$z_t = 0.8z_{t-1} + a_t$ 的样本函数和标准相关函数

【例 1.3.4】 考虑 AR(1)模型

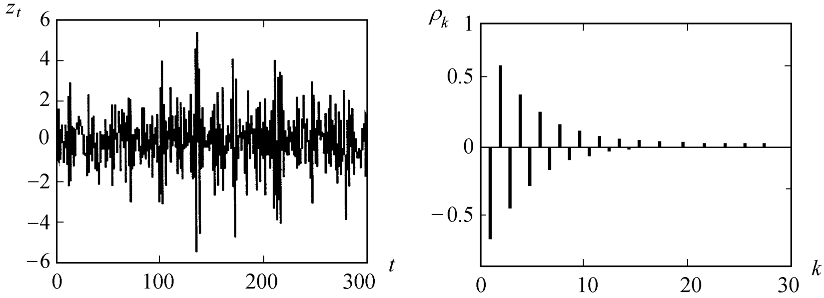
$$z_t = -0.8z_{t-1} + a_t \tag{1.3.25}$$

其中 a_t 是零均值、方差为 $\sigma_a^2=1$ 的正态白噪声.由式(1.3.22)有 z_t 的相关函数为

$$\rho_k = (-0.8)^k \tag{1.3.26}$$

z_t 的一个样本函数和相关函数图形如图 1.3.2 所示.对比图 1.3.1 和图 1.3.2 可看到,由于两个自回归系数 φ 的符号相反,导致相应的标准相关函数有完全不同的特征.对例

1.3.3，ρ_k 呈指数衰减趋于零，对例1.3.4，ρ_k 呈正、负交错指数衰减趋于零. 因前者 $\rho_k>0$，故均为正相关，这引出它的样本函数波动较慢，而后者 ρ_k 的符号正负交错，这引出它的样本函数波动很快，呈来回正负交错激烈振荡形式.



(a) 样本函数 z_t　　(b) 标准相关函数 ρ_k

图 1.3.2　$z_t=-0.8z_{t-1}+a_t$ 的样本函数和标准相关函数

1.3.3　ARMA 模型的可逆性

【例 1.3.5】　考虑 MA(1)模型

$$z_t=a_t-\theta a_{t-1} \tag{1.3.27}$$

其中 a_t 是零均值、方差为 σ_a^2 的白噪声. 容易验证 z_t 是一个平稳时间序列. 事实上

$$\mathrm{E}z_t=\mathrm{E}a_t-\theta\mathrm{E}a_{t-1}=0 \tag{1.3.28}$$

$$\gamma_k=\mathrm{E}[z_tz_{t+k}]=\mathrm{E}[(a_t-\theta a_{t-1})(a_{t+k}-\theta a_{t+k-1})]=\begin{cases}(1+\theta^2)\sigma_a^2 & k=0\\ -\theta\sigma_a^2 & k=1\\ 0 & k>1\end{cases} \tag{1.3.29}$$

这引出它的标准相关函数为

$$\rho_k=\begin{cases}1 & k=0\\ -\dfrac{\theta}{1+\theta^2} & k=1\\ 0 & k>1\end{cases} \tag{1.3.30}$$

可逆性是指：在什么条件下，白噪声 a_t 可通过 z_t 来表示？由式(1.3.27)有

$$z_t=(1-\theta q^{-1})a_t \tag{1.3.31}$$

形式上 a_t 可通过 z_t 表示为

$$a_t=\frac{1}{1-\theta q^{-1}}z_t \tag{1.3.32}$$

应用几何级数公式，有展式

$$a_t=\sum_{j=0}^{\infty}\theta^jq^{-j}z_t=\sum_{j=0}^{\infty}\theta^jz_{t-j} \tag{1.3.33}$$

这相当于将 z_t 表为无穷阶自回归模型 AR(∞).

由(1.3.27)引出 z_t 的方差 σ_z^2 是有界的，事实上

$$\sigma_z^2=\mathrm{E}[z_t^2]=\mathrm{E}[(a_t-\theta a_{t-1})^2]=(1+\theta^2)\sigma_a^2 \tag{1.3.34}$$

由均方收敛的 Cauchy 准则，级数(1.3.33)均方收敛的充要条件为当 $m\to\infty$，$n\to\infty$ 时，

$$\mathrm{E}\Big(\sum_{j=m}^{n}\theta^{j}z_{t-j}\Big)^{2}\rightarrow 0 \tag{1.3.35}$$

事实上

$$\mathrm{E}\Big(\sum_{j=m}^{n}\theta^{j}z_{t-j}\Big)^{2}=\mathrm{E}\Big[\Big(\sum_{j=m}^{n}\theta^{j}z_{t-j}\Big)\Big(\sum_{s=m}^{n}\theta^{s}z_{t-s}\Big)\Big]=\sum_{j,s=m}^{n}\theta^{j}\theta^{s}\gamma_{j-s} \tag{1.3.36}$$

因由(1.3.29)知 γ_k 是有界的,即

$$|\gamma_k|\leqslant r,\quad k=0,\pm 1,\pm 2,\cdots \tag{1.3.37}$$

因此由(1.3.36)引出(1.3.35)成立的充要条件是

$$|\theta|<1 \tag{1.3.38}$$

此时当 m,n 充分大后,(1.3.36)右边的项的值可任意小,即(1.3.35)成立.

条件$|\theta|<1$等价于滑动平均多项式 $\theta(q^{-1})=1-\theta q^{-1}$的零点(即 $\theta(x)$的零点)位于单位圆外.

对于一般的 ARMA 模型(1.3.10),可证明它可逆的充分条件为 $\theta(x)$的零点位于单位圆外,即 $\theta(q^{-1})$是一个稳定的多项式.

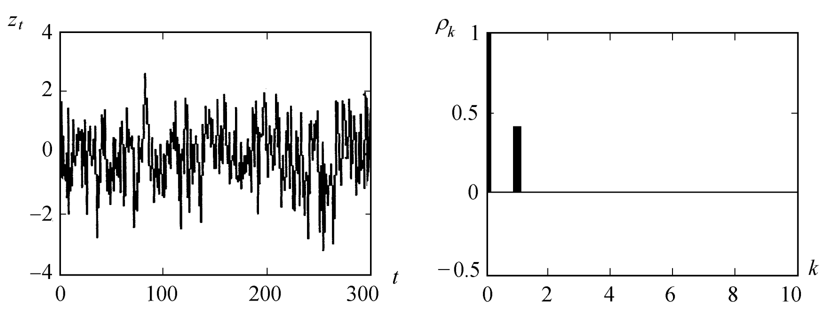
【例 1.3.6】 考虑 MA(1)模型

$$z_t=a_t+0.5a_{t-1} \tag{1.3.39}$$

其中 a_t 是零均值、方差为 $\sigma_a^2=1$ 的正态白噪声.显然模型(1.3.39)是可逆的.由(1.3.30)有样本函数 z_t 的标准相关函数为

$$\rho_k=\begin{cases}1 & k=0\\ 0.4 & k=1\\ 0 & k>1\end{cases} \tag{1.3.40}$$

z_t 的一个样本函数和标准相关函数 ρ_k 图形如图 1.3.3 所示.



(a) 样本函数 z_t (b) 标准相关函数 ρ_k

图 1.3.3 $z_t=a_t+0.5z_{t-1}$的样本函数和标准相关函数

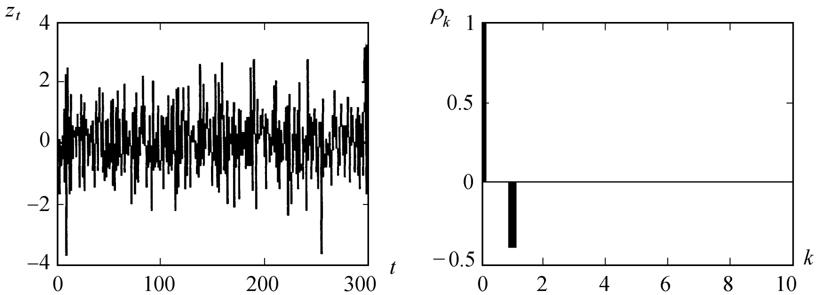
【例 1.3.7】 考虑 MA(1)模型

$$z_t=a_t-0.5a_{t-1} \tag{1.3.41}$$

其中 a_t 是零均值、方差为 $\sigma^2=1$ 的正态白噪声.显然模型(1.3.41)是可逆的.由(1.3.30)有 z_t 的标准相关函数

$$\rho_k=\begin{cases}1 & k=0\\ -0.4 & k=1\\ 0 & k>1\end{cases} \tag{1.3.42}$$

z_t 的一个样本函数和标准相关函数 ρ_k 图形如图 1.3.4 所示.



(a) 样本函数 z_t　　(b) 标准相关函数 ρ_k

图 1.3.4　$z_t = a_t - 0.5z_{t-1}$的样本函数和标准相关函数

比较图 1.3.3 和图 1.3.4 可看到,由于例 1.3.6 中 $\rho_1 = 0.4 > 0$,故 z_t 在相邻时刻取值正相关,这引出 z_t 取值的正负波动较缓慢,而由于例 1.3.7 中 $\rho_1 = -0.4 < 0$,故 z_t 在相邻时刻取值负相关,这引出 z_t 取值的正负波动很快.

1.3.4　几个平稳和非平稳 ARMA 模型的样本曲线

用 MATLAB 语言和标准正态分布 $N(0,1)$ 随机数在计算机上可用递推计算生成 ARMA模型的样本函数.

【例 1.3.8】　非平稳 AR(1)模型——随机游动模型

$$(1 - q^{-1}) z_t = a_t \tag{1.3.43}$$

其中 a_t 是零均值、方差为 $\sigma_a^2 = 0.1$ 的正态白噪声.因 $\varphi(q^{-1}) = 1 - q^{-1}$有零点 $q^{-1} = 1$,故 z_t 是非平稳的.它的一个样本函数如图 1.3.5 所示.

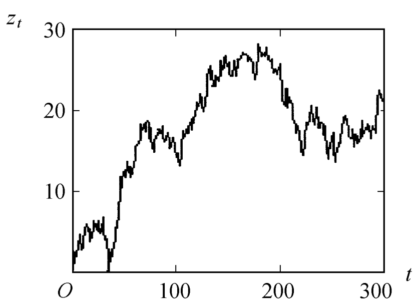


图 1.3.5　随机游动 $z_t = z_{t-1} + a_t$ 的样本函数

【例 1.3.9】　平稳可逆的 ARMA(1,1)模型

$$(1 - 0.9q^{-1}) z_t = (1 - 0.5q^{-1}) a_t \tag{1.3.44}$$

其中 a_t 是零均值、方差为 $\sigma_a^2 = 0.36$ 的正态白噪声.因 $\varphi = 0.9$,$\theta = 0.5$,它们的绝对值小于 1,故 z_t 是平稳、可逆的.它的一个样本函数如图 1.3.6 所示.

【例 1.3.10】　平稳 AR(2)模型

$$(1 - 0.9q^{-1})(1 - 0.5q^{-1}) z_t = a_t \tag{1.3.45}$$

其中 a_t 是零均值、方差为 $\sigma_a^2 = 1$ 的正态白噪声.它的一个样本函数如图 1.3.7 所示.

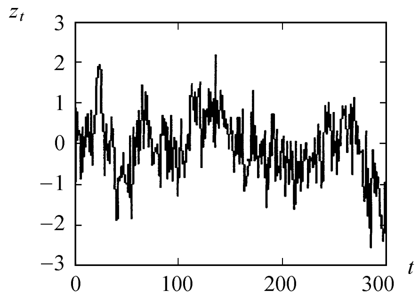


图 1.3.6　$(1-0.9q^{-1})z_t=(1-0.5q^{-1})a_t$ 的一个样本函数

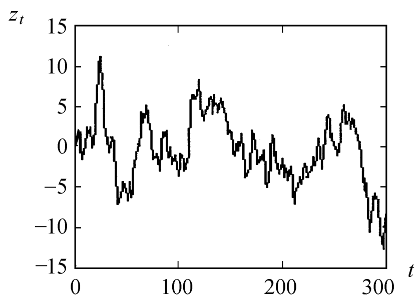


图 1.3.7　平稳模型 $(1-0.9q^{-1})(1-0.5q^{-1})z_t=a_t$ 的一个样本函数

【例 1.3.11】　平稳、可逆的 ARMA(2.1)模型

$$(1-0.8q^{-1})(1-0.5q^{-1})z_t=(1-0.3q^{-1})a_t \tag{1.3.46}$$

其中 a_t 是零均值、方差为 $\sigma_a^2=0.64$ 的正态白噪声. 它的一个样本函数如图 1.3.8 所示.

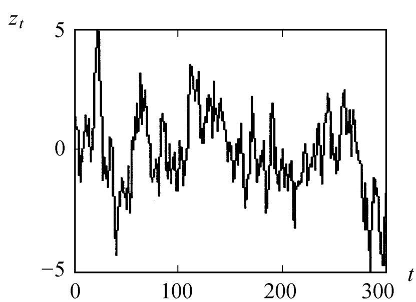


图 1.3.8　平稳可逆模型 $(1-0.8q^{-1})(1-0.5q^{-1})z_t=(1-0.3q^{-1})a_t$ 的一个样本函数

【例 1.3.12】　非平稳 ARMA(2,1)模型

$$(1-0.8q^{-1})(1-q^{-1})z_t=(1-0.3q^{-1})a_t \tag{1.3.47}$$

其中 a_t 是零均值、方差为 $\sigma_a^2=0.64$ 的正态白噪声,因 $\varphi(q^{-1})=(1-0.8q^{-1})(1-q^{-1})$有零点 $q^{-1}=1$,故式(1.3.47)是非平稳的. 它的一个样本函数如图 1.3.9 所示.

1.3.5　应用实例

【例 1.3.13】　对某化学反应过程每两小时做一次观测,依次得到溶液浓度 z_t 的 197 个数据[3],如图 1.3.10 所示. 显然 z_t 是非平稳序列,因为它有曲线趋势. 为了消除曲线趋势,化 z_t 为平稳序列,引入它的差分序列 y_t 为

$$y_t=z_t-z_{t-1} \tag{1.3.48}$$

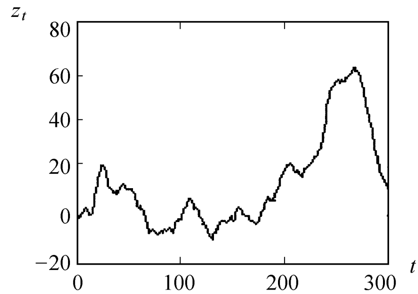


图 1.3.9　非平稳模型$(1-0.8q^{-1})(1-q^{-1})z_t=(1-0.3q^{-1})a_t$的一个样本函数

它的图形如图 1.3.11 所示，可看到 y_t 具有平稳序列的特征.

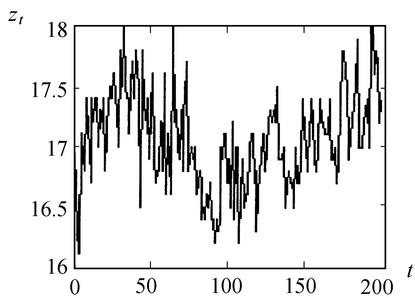


图 1.3.10　某化学反应过程中浓度 z_t 的 197 个数据

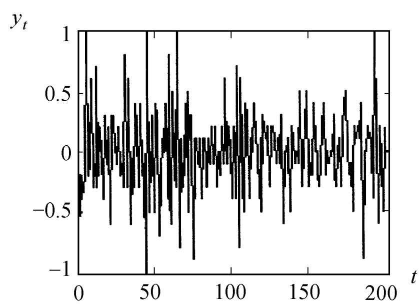


图 1.3.11　某化学反应过程中浓度差分序列 $y_t=z_t-z_{t-1}$

计算差分序列 y_t 的采样相关函数和标准相关函数

$$\hat{\gamma} = \frac{1}{197}\sum_{t=2}^{197-k}(y_t-\bar{y})(y+k-\bar{y}), \quad \hat{\rho}_k = \hat{\gamma}_k/\hat{\gamma}_0,$$

$$\bar{y} = \sum_{t=2}^{197} y_t/197 = 0.002 \tag{1.3.49}$$

计算结果为[3]

$$\begin{aligned}
&\hat{\rho}_0=1, \quad \hat{\rho}_1=-0.412\,9,\\
&\hat{\rho}_2=0.020\,1, \quad \hat{\rho}_3=-0.068\,0, \quad \hat{\rho}_4=-0.008\,7,\\
&\hat{\rho}_5=-0.076\,6, \quad \hat{\rho}_6=-0.008\,3, \quad \hat{\rho}_7=0.135\,0,\\
&\qquad\vdots
\end{aligned} \tag{1.3.50}$$

显然 $\hat{\rho}_1$ 是非零的，而对 $k>1$，$\hat{\rho}_k$ 的绝对值很小，故可近似认为 $\hat{\rho}_k=0\ (k>1)$. 于是可将 y_t 看成是一步相关序列，即它的相关函数在 $k=1$ 处截尾，换言之，y_t 的标准相关函数可看成是

$$\rho_0=1,\quad \rho_1\neq 0,\quad \rho_k=0\quad (k>1) \tag{1.3.51}$$

显然它可用 MA(1)模型来描写，即

$$y_t=a_t-\theta a_{t-1} \tag{1.3.52}$$

其中 a_t 是零均值、方差为 σ_a^2 的白噪声，θ 是模型参数. σ_a^2 和 θ 是未知的.

我们可用如下方法求方差 σ_a^2 和 θ. 由(1.3.30)MA(1)模型，当 $k=1$ 时有

$$\rho_1=-\frac{\theta}{1+\theta^2} \tag{1.3.53}$$

由(1.3.50)，将 $\rho_1=-0.412\,7$ 代入(1.3.53)，有关 θ 的一元二次方程为

$$\theta^2-2.423\,067\,6\theta+1=0 \tag{1.3.54}$$

解之可得使 MA(1)模型(1.3.52)可逆的 θ 的估值为

$$\hat{\theta}=0.527\,564\,4 \tag{1.3.55}$$

又由 y_t 数据可算出在 $k=1$ 时的相关函数值 γ_1 的估值

$$\hat{\gamma}_1=-0.526 \tag{1.3.56}$$

由(1.3.29)有关系

$$\hat{\gamma}_1=-\hat{\theta}\hat{\sigma}_a^2 \tag{1.3.57}$$

这引出 σ_a^2 的估值为

$$\hat{\sigma}_a^2=-\hat{\gamma}_1/\hat{\theta}=0.106\,527\,2 \tag{1.3.58}$$

因而最终得到原非平稳序列 z_t 的 ARMA 模型为

$$z_t-z_{t-1}=a_t-0.527\,564\,4a_{t-1} \tag{1.3.59}$$

其中 a_t 是零均值、方差为 $\sigma_a^2=0.106\,529\,2$ 的白噪声.

本例的建模过程完全类似于在例 1.3.1 中的对 IBM 股票收盘价格时间序列的建模.

1.4 ARMA 过程的展式

用 ARMA 模型描写的随机序列称为 ARMA 过程. ARMA 过程本质上可看成是以白噪声作为输入的一个动态系统的输出，输入与输出递推关系用 ARMA 模型表示. 本节研究 ARMA 过程的非递推无穷级数表达式问题，可更进一步揭示 ARMA 过程的实质：平稳的 ARMA 过程可用白噪声的线性运算来生成，这叫平稳性，即平稳 ARMA 过程可展为白噪声的无穷级数. 反之，一个平稳可逆的 ARMA 模型的输入白噪声总可用它的输出的线性运算生成，这叫可逆性. 这种关系级数展式在理论和应用研究中具有重要意义.

对于 ARMA 过程从一种形态的表达式到另一种形态的表达式的转化，不是数学游戏，而是解决问题强有力的杠杆. 我们将揭示 ARMA 过程、AR 过程和 MA 过程三者之间的转化关系. 将引入无穷阶 AR 模型 AR(∞)和无穷阶 MA 模型 MA(∞)概念，并指出平稳可逆的 ARMA 模型等价于 AR(∞)模型，也等价于 MA(∞)模型. 我们指出，平稳的 ARMA 模型总可用高阶 AR 模型逼近到任意精度. 这就是为什么工程上乐于采用 AR 模型近似描写

平稳时间序列.我们也指出,平稳、可逆的 ARMA 模型总可用高阶 MA 模型逼近到任意精度.

线性运算也叫线性“滤波器”(Filter).平稳可逆 ARMA 过程与白噪声关系如图 1.4.1 所示,其中由白噪声生成 ARMA 过程的滤波器叫“成形滤波器”,由 ARMA 过程生成白噪声的滤波器叫“白化滤波器”.图中 ψ_j 和 β_j 为滤波器系数.我们将给出它们的计算公式.

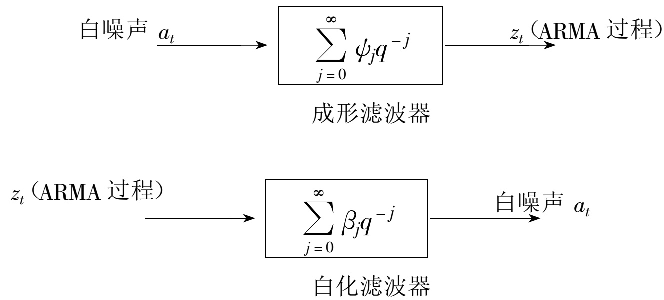


图 1.4.1　ARMA 过程与白噪声关系

我们将证明平稳可逆 ARMA 模型、AR 模型和 MA 模型的关系如下:

ARMA 模型 = MA(∞)模型

ARMA 模型 ≈ 高阶 MA(q)

ARMA 模型 = AR(∞)模型

ARMA 模型 ≈ 高阶 AR(p)

AR 模型 = MA(∞)模型

MA 模型 = AR(∞)模型

AR 模型 ≈ 高阶 MA(q)模型

MA 模型 ≈ 高阶 AR(p)模型

上述精确表达式在理论分析时有重要作用,近似表达式在工程应用中和建模中有重要意义.上述关系体现了有限与无限的转化、近似的与精确的转化.

1.4.1　平稳 ARMA 过程的 MA(∞)展式

【例 1.4.1】　继例 1.3.2　考虑平稳 AR(1)过程

$$(1-\varphi q^{-1})z_t = a_t,\quad |\varphi|<1 \tag{1.4.1}$$

其中 a_t 是零均值、方差为 σ_a^2 的白噪声,现在我们用不同于例 1.3.2 的方法求 z_t 的展式.

设 z_t 的展式为

$$z_t = \frac{1}{1-\varphi q^{-1}}a_t = \sum_{j=0}^{\infty}\psi_j q^{-j}a_t = \sum_{j=0}^{\infty}\psi_j a_{t-j} \tag{1.4.2}$$

其中 ψ_j 为待定系数.由(1.4.2)有关系

$$\frac{1}{1-\varphi q^{-1}} = \sum_{j=0}^{\infty}\psi_i q^{-j} \tag{1.4.3}$$

$$1 = (1-\varphi q^{-1})\sum_{j=0}^{\infty}\psi_i q^{-j} \tag{1.4.4}$$

比较(1.4.4)两边 q^{-j}系数有关系

$$\psi_0 = 1, \quad \psi_j - \varphi\psi_{j-1} = 0, \quad j = 1, 2, \cdots \tag{1.4.5}$$

这引出

$$\psi_1 = \varphi, \quad \psi_2 = \varphi^2, \cdots, \quad \psi_j = \varphi^j \tag{1.4.6}$$

故有展式

$$z_t = \sum_{j=0}^{\infty} \varphi^j a_{t-j} \tag{1.4.7}$$

它相同于例 1.3.2 用几何级数求和方法所得结果，这种方法叫“待定系数法”.

考虑一般平稳 ARMA (p, q)过程

$$\varphi(q^{-1}) z_t = \theta(q^{-1}) a_t \tag{1.4.8}$$

其中 a_t 为零均值、方差为 σ_a^2 的白噪声，即

$$\mathrm{E}a_t = 0, \quad \mathrm{E}[a_t a_s] = \sigma_a^2 \delta_{ts} \tag{1.4.9}$$

其中 E 为均值号，$\delta_{tt} = 1$，$\delta_{ts} = (t \neq s)$，q^{-1}为单位滞后算子，

$$\begin{aligned} \varphi(q^{-1}) &= 1 - \varphi_1 q^{-1} - \cdots - \varphi_p q^{-p} \\ \theta(q^{-1}) &= 1 - \theta_1 q^{-1} - \cdots - \theta_q q^{-q} \end{aligned} \tag{1.4.10}$$

设过程是平稳的，即多项式 $\varphi(x)$的零点位于单位圆外，考虑展式

$$z_t = \frac{\theta(q^{-1})}{\varphi(q^{-1})} a_t = \sum_{j=0}^{\infty} \psi_j q^{-j} a_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} \tag{1.4.11}$$

它是一个无穷阶滑动平均过程，记为 MA(∞).

如何求待定系统 ψ_j？受例 1.4.1 的启发，由(1.4.11)有关系

$$\frac{\theta(q^{-1})}{\varphi(q^{-1})} = \sum_{j=0}^{\infty} \psi_j q^{-j} \tag{1.4.12}$$

$$\theta(q^{-1}) = \varphi(q^{-1}) \sum_{j=0}^{\infty} \psi_j q^{-j} \tag{1.4.13}$$

即

$$1 - \theta_1 q^{-1} - \cdots - \theta_q q^{-q} = (1 - \varphi_1 q^{-1} - \cdots - \varphi_p q^{-p})(\psi_0 + \psi_1 q^{-1} + \cdots + \psi_j q^{-j} + \cdots) \tag{1.4.14}$$

比较上式两边 q^{-j}系数引出递推关系

$$\begin{aligned} &\psi_0 = 1 \\ &\psi_1 - \varphi_1 = -\theta_1 \\ &\vdots \\ &\psi_j - \varphi_1 \psi_{j-1} - \cdots - \varphi_p \psi_{j-p} = -\theta_j \end{aligned} \tag{1.4.15}$$

这引出如下定理.

【定理 1.4.1】 平稳 ARMA 过程(1.4.8)有无穷阶 MA(∞)模型展式

$$z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} \tag{1.4.16}$$

其中系数 ψ_j 有递推计算公式

$$\psi_j = \varphi_1 \psi_{j-1} + \cdots + \varphi_p \psi_{j-p} - \theta_j \tag{1.4.17}$$

其中规定 $\psi_0 = 1$，$\psi_j = 0(j < 0)$，$\theta_j = 0(j > q)$. 特别，当 $j > q$，有齐次差分方程

$$\psi_j = \varphi_1\psi_{j-1} + \cdots + \varphi_p\psi_{j-p}, \quad j > q \tag{1.4.18}$$

即

$$\varphi(q^{-1})\psi_j = 0, \quad j > q \tag{1.4.19}$$

【推论 1.4.1】 平稳 AR(p)过程 $\varphi(q^{-1})z_t = a_t$ 有 MA(∞)展式(1.4.16)，其中系数 ψ_j 可递推计算为

$$\psi_j = \varphi_1\psi_{j-1} + \cdots + \varphi_p\psi_{j-p}, \quad j > 0 \tag{1.4.20}$$

其中规定 $\psi_0 = 1$，$\psi_j = 0 (j < 0)$.

注意由平稳性($\varphi(x)$的零点位于单位圆外)假设引出如下级数收敛

$$\sum_{j=0}^{\infty}\psi_j^2 < \infty \tag{1.4.21}$$

从而展式(1.4.16)均方收敛，由(1.4.21)引出当 $j \to \infty$ 时

$$\psi_j \to 0 \tag{1.4.22}$$

因而当取 n_0 充分大时有近似 MA(n_0)展式为

$$z_t \approx \sum_{j=0}^{n_0}\psi_j a_{t-j} \tag{1.4.23}$$

这引出如下重要定理.

【定理 1.4.2】 平稳 ARMA 过程(1.4.8)可用高阶 MA(n_0)过程(1.4.23)逼近到任意精度，即当 n_0 充分大，平稳 ARMA 过程(1.4.8)可用高阶滑动平均过程 MA(n_0)(1.4.23)近似表示.

【例 1.4.2】 考虑平稳 AR(1)模型

$$(1 - 0.1q^{-1})z_t = a_t \tag{1.4.24}$$

它有 MA(∞)展式

$$z_t = \frac{1}{1 - 0.1q^{-1}}a_t = \sum_{j=0}^{\infty}(0.1)^j a_{t-j} \tag{1.4.25}$$

于是有 z_t 的近似 MA(1)模型为

$$z_t = (1 + 0.1q^{-1})a_t \tag{1.4.26}$$

注意，定理 1.4.1 的重要意义在于：任何平稳 ARMA 过程 z_t 可用白噪声的线性运算(无穷级数)来生成. 定理 1.4.2 有重要的工程应用意义. 在应用中人们常常用更简单的近似模型来描写一个时间序列.

1.4.2 平稳、可逆的 ARMA 过程的 AR(∞)展式

由式(1.4.8)有 a_t 可表为

$$a_t = \frac{\varphi(q^{-1})}{\theta(q^{-1})}z_t = \left(1 - \sum_{j=1}^{\infty}\pi_j q^{-j}\right)z_t \tag{1.4.27}$$

其中 π_j 为待定系数. 由式(1.4.27)有关系

$$\frac{\varphi(q^{-1})}{\theta(q^{-1})} = -\sum_{j=0}^{\infty}\pi_j q^{-j}, \quad \pi_0 = -1 \tag{1.4.28}$$

用待定系数法，类似于定理 1.4.1 的推导有如下定理.

【定理 1.4.3】 平稳、可逆的 ARMA 过程(1.4.8)有无穷阶自回归模型 AR(∞)展式

$$z_t = \sum_{j=1}^{\infty} \pi_j z_{t-j} + a_t \tag{1.4.29}$$

其中系数 π_j 可递推计算为

$$\pi_j = \theta_1 \pi_{j-1} + \cdots + \theta_q \pi_{j-q} + \varphi_j, \quad j > 0 \tag{1.4.30}$$

其中规定 $\pi_0 = -1$，$\pi_j = 0\ (j < 0)$，$\varphi_j = 0\ (j > p)$.

【推论 1.4.2】 MA(q)过程 $z_t = \theta(q^{-1}) a_t$ 有 AR(∞)模型展式(1.4.29)，其中系数 π_j 有递推计算公式

$$\theta(q^{-1}) \pi_j = 0, \quad j > 0 \tag{1.4.31}$$

其中规定 $\pi_0 = -1$，$\pi_j = 0\ (j < 0)$.

注意由均方收敛的展式(1.4.29)引出收敛级数

$$\sum_{j=1}^{\infty} \pi_j^2 < \infty \tag{1.4.32}$$

这引出

$$\pi_j \to 0 \qquad (j \to \infty) \tag{1.4.33}$$

因而当取 n_0 充分大有近似展式

$$z_t \approx \sum_{j=1}^{n_0} \pi_j z_{t-j} + a_t \tag{1.4.34}$$

这引出如下定理.

【定理 1.4.4】 平稳可逆 ARMA 过程(1.4.8)可用高阶 AR(n_0)过程逼近到任意精度，即当 n_0 充分大，ARMA 过程(1.4.8)可近似用高阶 AR(n_0)过程(1.4.34)表示.

这一定理有重要应用意义，在许多应用中常常用高阶 AR 模型代替一个 ARMA 模型，使处理问题简化.

【例 1.4.3】 考虑 MA(1)过程

$$z_t = (1 - 0.5q^{-1}) a_t \tag{1.4.35}$$

则有 AR(∞)展式

$$a_t = \frac{1}{1 - 0.5q^{-1}} z_t = \sum_{k=0}^{\infty} (0.5)^j z_{t-j} \tag{1.4.36}$$

取 $n_0 = 3$，则 z_t 有近似 AR(3)表达式

$$z_t = -0.5z_{t-1} - 0.25z_{t-2} - 0.125z_{t-3} + a_t \tag{1.4.37}$$

【例 1.4.4】 考虑平稳可逆的 ARMA(1,1)过程

$$(1 - \varphi q^{-1}) z_t = (1 - \theta q^{-1}) a_t, \quad |\varphi| < 1, |\theta| < 1 \tag{1.4.38}$$

由定理 1.4.3，它有 AR(∞)展式(1.4.29)，其中系数 π_j 由(1.4.30)递推计算为

$$\pi_0 = -1, \quad \pi_1 = \varphi - \theta, \quad \pi_2 = \theta(\varphi - \theta), \cdots, \pi_j = \theta^{j-1}(\varphi - \theta), \quad j \geqslant 1 \tag{1.4.39}$$

于是它有 AR(∞)展式

$$z_t = \sum_{j=1}^{\infty} \theta^{j-1} (\varphi - \theta) z_{t-j} + a_t \tag{1.4.40}$$

1.5 ARMA 过程的相关函数

研究不同类型平稳时间序列的相关函数特性和计算相关函数，具有重要的理论和应用意义. 在 ARMA 模型建模中，常常根据时间序列的采样相关函数特点来决定模型类型和阶次. 例如若发现某时间序列的采样相关函数具有截尾特性，则可判定该时间序列可用 MA 模型描写，其截尾长度 q 就是 MA 模型的阶次. 例如根据 AR 模型或 MA 模型的相关函数，人们可估计 AR 或 MA 模型参数. 本节介绍平稳可逆 ARMA 过程的相关函数非递推和递推形式的两种算法[2].

1.5.1 利用 ARMA 过程的 MA(∞)展式计算相关函数

考虑平稳、可逆的 ARMA(p, q)过程(1.4.8) ~ (1.4.10). 由定理 1.4.1，它有 MA(∞)均方收敛展式

$$z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}, \quad \psi_0 = 1 \tag{1.5.1}$$

其中系数 ψ_j 可递推计算为

$$\psi_j = \varphi_1 \psi_{j-1} + \cdots + \varphi_p \psi_{j-p} - \theta_j \tag{1.5.2}$$

其中规定 $\psi_0 = 1$，$\psi_j = 0\,(j < 0)$，$\theta_j = 0\,(j > q)$. 由均方收敛级数性质和(1.4.9)，这引出 z_t 的均值和相关函数分别为

$$\mathrm{E} z_t = \mathrm{E} \sum_{j=0}^{\infty} \psi_j a_{t-j} = \sum_{j=0}^{\infty} \psi_j \mathrm{E} a_{t-j} = 0 \tag{1.5.3}$$

$$\gamma_k = \mathrm{E}[z_t z_{t+k}] = \mathrm{E}\left[\left(\sum_{j=0}^{\infty} \psi_j a_{t-j}\right)\left(\sum_{i=0}^{\infty} \psi_i a_{t+k-i}\right)\right] = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \tag{1.5.4}$$

由 Cauchy 不等式

$$|\gamma_k| \leqslant \sigma_a^2 \sum_{j=0}^{\infty} |\psi_j| |\psi_{j+k}| \leqslant \left(\sum_{j=0}^{\infty} \psi_j^2\right)^{\frac{1}{2}} \left(\sum_{j=0}^{\infty} \psi_{j+k}^2\right)^{\frac{1}{2}} \tag{1.5.5}$$

由(1.5.1)均方收敛，引出

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty \tag{1.5.6}$$

从而由收敛级数性质有

$$\sum_{j=0}^{\infty} \psi_{j+k}^2 \to 0 \quad (k \to \infty) \tag{1.5.7}$$

故有

$$\gamma_k \to 0 \qquad (k \to \infty) \tag{1.5.8}$$

当 $k = 0$ 时，由(1.5.4)引出 z_t 的方差 $\sigma_z^2 = \gamma_0$ 为

$$\sigma_z^2 = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2 \tag{1.5.9}$$

综上所述有如下定理.

【定理 1.5.1】 平稳、可逆 ARMA 过程(1.4.8) ~ (1.4.10)有零均值、相关函数 γ_k 和

方差 σ_z^2 为

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \tag{1.5.10}$$

$$\sigma_z^2 = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2 \tag{1.5.11}$$

且相关函数有性质

$$\gamma_k \to 0 \qquad (k \to \infty) \tag{1.5.12}$$

它的标准相关函数 ρ_k 为

$$\rho_k = \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \Big/ \sum_{j=0}^{\infty} \psi_j^2 \tag{1.5.13}$$

且也有

$$\rho_k \to 0 \quad (k \to \infty) \tag{1.5.14}$$

注意，在应用中性质(1.5.12)常用来判断时间序列的平稳性．若某时间序列的采样标准相关函数 $\hat{\rho}_k$ 不趋于零，则可判定该时间序列是非平稳的．通常若 $\hat{\rho}_k$ 不很快趋于零，则可认为该时间序列是非平稳的．

【例 1.5.1】 考虑例 1.4.1 的 AR(1)过程

$$(1 - \varphi q^{-1}) z_t = a_t, \quad |\varphi| < 1 \tag{1.5.15}$$

由(1.4.7)，它有展式

$$z_t = \sum_{j=0}^{\infty} \varphi^j a_{t-j} \tag{1.5.16}$$

由(1.5.10)和(1.5.11)，它有相关函数 γ_k 和方差 $\sigma_z^2 = \gamma_0$ 为

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{\infty} \varphi^j \varphi^{j+k} = \sigma_a^2 \varphi^k \sum_{j=0}^{\infty} \varphi^{2j} = \frac{\sigma_a^2 \varphi^k}{1 - \varphi^2} \tag{1.5.17}$$

$$\sigma_z^2 = \gamma_0 = \frac{\sigma_a^2}{1 - \varphi^2} \tag{1.5.18}$$

它的标准相关函数为

$$\rho_k = \varphi^k \tag{1.5.19}$$

【例 1.5.2】 考虑 ARMA(1.1)过程

$$(1 - \varphi q^{-1}) z_t = (1 - \theta q^{-1}) a_t, \quad |\varphi| < 1, |\theta| < 1 \tag{1.5.20}$$

由(1.5.2)有

$$\psi_0 = 1, \quad \psi_1 = \varphi - \theta, \quad \psi_j = \varphi^{j-1}(\varphi - \theta) \tag{1.5.21}$$

于是由(1.5.1)有展式

$$z_t = a_t + \sum_{j=1}^{\infty} (\varphi - \theta) \varphi^{j-1} a_{t-j} \tag{1.5.22}$$

由(1.5.10)和(1.5.11)有相关函数和方差

$$\gamma_k = \sigma_a^2 \psi_k + \sum_{j=1}^{\infty} (\varphi - \theta) \varphi^{j-1} (\varphi - \theta) \varphi^{j+k-1} = \sigma_a^2 \left[\psi_k + (\varphi - \theta)^2 \varphi^k \frac{1}{1 - \varphi^2} \right] \tag{1.5.23}$$

$$\sigma_z^2 = \gamma_0 = \sigma_a^2 \left[1 + \frac{(\varphi - \theta)^2}{1 - \varphi^2} \right] \tag{1.5.24}$$

1.5.2 MA(q)过程相关函数

对于 MA(q)过程，展式(1.5.1)成为

$$z_t = \sum_{j=0}^{q} \psi_j a_{t-j}, \quad \psi_0 = 1 \tag{1.5.25}$$

这相当于在无穷展式(1.5.1)中 $\psi_j = 0 (j > q)$. 于是由定理 1.5.1 可得如下定理.

【定理 1.5.2】 对于 MA(q)过程 z_t,

$$z_t = a_t + \psi_1 a_{t-1} + \cdots + \psi_q a_{t-q} \tag{1.5.26}$$

有在 $k = q$ 处截尾的相关函数

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{q-k} \psi_j \psi_{j+k}, \quad k = 0, 1, \cdots, q \tag{1.5.27}$$

$$\gamma_k = 0, \quad k > q \tag{1.5.28}$$

方差 σ_z^2 为

$$\sigma_z^2 = \sigma_a^2 \sum_{j=0}^{q} \psi_j^2 \tag{1.5.29}$$

它的标准相关函数为

$$\rho_k = \sum_{j=0}^{q-k} \psi_j \psi_{j+k} \Big/ \sum_{j=0}^{q} \psi_j^2, \quad k = 0, 1, \cdots, q \tag{1.5.30}$$

$$\rho_k = 0, \quad k > q \tag{1.5.31}$$

对于 MA(q)过程(1.5.26)，它的相关函数在 $k = q$ 处具有截尾性质：$\gamma_k = 0 (k > q)$. 当已知 γ_k，利用(1.5.27)的$(q+1)$个方程，可解出$(q+1)$个未知参数 $\psi_j (j = 1, \cdots, k)$和 σ_a^2. 但(1.5.27)是非线性方程组，它的解是不惟一的. 但使 MA 多项式 $\psi(q^{-1}) = 1 + \psi_1 q^{-1} + \cdots + \psi_q q^{-q}$稳定(即 $\psi(x)$的零点全位于单位圆外)的(1.5.27)的解是惟一的[2]，即使MA(q)过程(1.5.26)可逆的参数 ψ_j 和σ_a^2 是惟一的. 一般，已知 γ_k，关于 ψ_j 和σ_a^2 的非线性方程组(1.5.27)没有解析解. 但可以证明，对 $q = 1$ 和 $q = 2$，非线性方程组(1.5.27)有解析解. 下面以 $q = 1$ 为例，我们来说明，如何由 γ_0 和 γ_1 求 MA(1)模型参数 ψ_1 和白噪声方差 σ_a^2.

【例 1.5.3】 考虑可逆的 MA(1)过程

$$z_t = a_t + \psi a_{t-1}, \quad |\psi| < 1 \tag{1.5.32}$$

已知它的相关函数值 γ_0, γ_1，试求它的模型参数 ψ 和白噪声 a_t 的方差σ_a^2.

由(1.5.27)有关于 ψ 和σ_a^2 的非线性方程组

$$\gamma_0 = \sigma_a^2 (1 + \psi^2) \tag{1.5.33}$$

$$\gamma_1 = \psi \sigma_a^2 \tag{1.5.34}$$

将(1.5.34)被(1.5.33)除，引出关系

$$\rho_1 = \frac{\psi}{1 + \psi^2} \tag{1.5.35}$$

其中 $\rho_1 = \gamma_1 / \gamma_0$. 这引出关于未知参数 ψ 的一元二次方程

$$\psi^2 - \frac{1}{\rho_1} \psi + 1 = 0 \tag{1.5.36}$$

它有判别式

$$\Delta = \frac{1}{\rho_1^2} - 4 = \frac{1 - 4\rho_1^2}{\rho_1^2} \tag{1.5.37}$$

下证 $\Delta \geqslant 0$，因而(1.5.36)有相异实根. 事实上，考察由(1.5.35)给出的函数 $\rho_1 = \rho_1(\psi)$，注意，置

$$\frac{\mathrm{d}\rho_1}{\mathrm{d}\psi} = \frac{(1+\psi^2) - 2\psi^2}{(1+\psi^2)^2} = \frac{1-\psi^2}{(1+\psi^2)^2} = 0 \tag{1.5.38}$$

有稳定点 $\psi = \pm 1$. 容易验证在 $\psi = 1$ 处有 $\mathrm{d}^2\rho_1/\mathrm{d}\psi^2 < 0$，故 $\psi = 1$ 为 $\rho_1(\psi)$ 的极大值点，同理 $\psi = -1$ 为 $\rho_1(\psi)$ 的极小值点. 注意当 $|\psi| < 1$ 时有 $\mathrm{d}\rho/\mathrm{d}\psi > 0$，因而 $\rho_1(\psi)$ 是单调增加函数. 因当 $\psi = \pm 1$ 时，由(1.5.35)有 $\rho_1 = \pm\frac{1}{2}$，故当 $|\psi| < 1$ 时有 $|\rho_1| < \frac{1}{2}$，因而 $1 - 4\rho_1^2 > 1 - 4\left(\frac{1}{2}\right)^2 = 0$，这引出 $\Delta > 0$. 方程(1.5.36)有相异实根，由根与系数关系，这两个相异实根之积等于常数项 1，因而方程(1.5.36)必有一个的绝对值小于 1，另一根的绝对值大于 1，取绝对值小于 1 的根为所求. 它保证 MA(1)过程(1.5.32)是可逆的，进而由式(1.5.34)可求得 σ_a^2 为

$$\sigma_a^2 = \gamma_1/\psi \tag{1.5.39}$$

【例 1.5.4】 考虑带观测噪声的平稳 AR(1)过程

$$(1 + \varphi q^{-1}) s_t = w_t, \quad |\varphi| < 1 \tag{1.5.40}$$

$$z_t = s_t + v_t \tag{1.5.41}$$

其中 w_t 和 v_t 是零均值、方差各为 σ_w^2 和 σ_v^2 的相互独立白噪声，s_t 是原始信号，y_t 是对 s_t 的观测信号，v_t 为观测噪声. 试求关于观测信号 z_t 的 ARMA 模型.

将(1.5.40)代入(1.5.41)引出关系

$$(1 + \varphi q^{-1}) z_t = w_t + (1 + \varphi q^{-1}) v_t \tag{1.5.42}$$

记(1.5.42)右边的随机过程为

$$m_t = w_t + (1 + \varphi q^{-1}) v_t \tag{1.5.43}$$

且记它的相关函数为 $\gamma_i = \mathrm{E}[m_t m_{t-i}]$，则有

$$\gamma_0 = \sigma_w^2 + (1 + \varphi^2)\sigma_v^2 \tag{1.5.44}$$

$$\gamma_1 = \varphi\,\sigma_v^2 \tag{1.5.45}$$

$$\gamma_i = 0 \quad (i > 1) \tag{1.5.46}$$

因此 γ_k 在 $k = 1$ 处截尾. 根据 MA 过程的特性，可将 m_t 用一个等价的可逆的 MA(1)过程表示为

$$(1 + dq^{-1}) \varepsilon_t = w_t + (1 + \varphi q^{-1}) v_t \tag{1.5.47}$$

其中 ε_t 是零均值、方差为 σ_ε^2 的白噪声，且我们要求多项式 $(1 + dq^{-1})$ 是稳定的，这等价于要求 $|d| < 1$. 在这里，我们定义两个随机过程等价是指它们具有相同的相关函数. 问题转化为如何求参数 d 和 σ_ε^2？

计算(1.5.47)两边随机过程的相关函数，有关系

$$(1 + d^2)\sigma_\varepsilon^2 = \gamma_0 \tag{1.5.48}$$

$$d\sigma_\varepsilon^2 = \gamma_1 \tag{1.5.49}$$

其中 γ_0, γ_1 由式(1.5.44)和(1.5.45)定义，消去 σ_ε^2 有一元二次方程

$$d^2 - (\gamma_0/\gamma_1)d + 1 = 0 \tag{1.5.50}$$

将(1.5.44)和(1.5.45)代入上式有一元二次方程

$$\varphi d^2 - [(1+\varphi^2) + \sigma_w^2/\sigma_v^2]d + \varphi = 0 \tag{1.5.51}$$

它有判别式 Δ 为

$$\begin{aligned}\Delta = [(1+\varphi^2)+\sigma_w^2/\sigma_v^2]^2 - 4\varphi^2 = [(1+\varphi^2)+\sigma_w^2/\sigma_v^2+2\varphi][(1+\varphi^2)+\sigma_w^2/\sigma_v^2-2\varphi] = \\ [(1+\varphi)^2+\sigma_w^2/\sigma_v^2][(1-\varphi^2)^2+(\sigma_w^2/\sigma_v^2)] > 0\end{aligned} \tag{1.5.52}$$

因而(1.5.51)有相异实根 d_1 和 d_2. 由根与系数关系有 $d_1 d_2 = \varphi/\varphi = 1$，因而必有一实根绝对值大于 1，另一实根绝对值小于 1，取 $|d|<1$ 的根为所求. 再将 d 代入(1.5.49)可得

$$\sigma_\varepsilon^2 = \gamma_1/d \tag{1.5.53}$$

于是最终由(1.5.42)和(1.5.41)可得 z_t 的平稳、可逆 ARMA 模型为

$$(1+\varphi q^{-1})z_t = (1+dq^{-1})\varepsilon_t \tag{1.5.54}$$

【例 1.5.5】 考虑 MA(2)过程

$$z_t = (1-0.6q^{-1})(1-0.5q^{-1})a_t \tag{1.5.55}$$

或

$$z_t = a_t - 1.1a_{t-1} + 0.3a_{t-2} \tag{1.5.56}$$

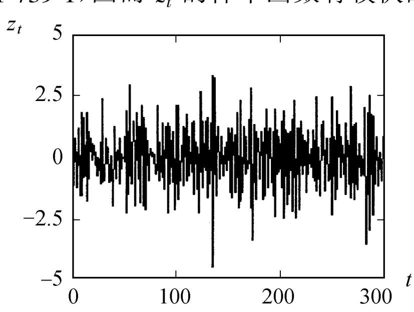
其中 a_t 是零均值、方差为 $\sigma_a^2 = 1$ 的正态白噪声. 容易求得它的相关函数 γ_k 为

$$\gamma_k = \begin{cases} 2.3 & k=0 \\ -1.43 & k=1 \\ 0.3 & k=2 \\ 0 & k>2 \end{cases} \tag{1.5.57}$$

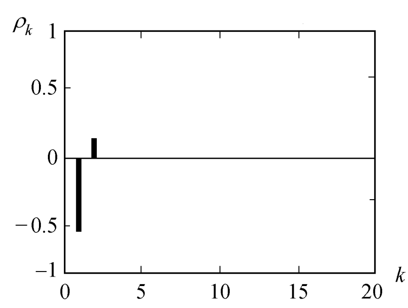
因而标准相关函数 $\rho_k = \gamma_k/\gamma_0$ 为

$$\rho_k = \begin{cases} 1 & k=0 \\ -0.621\,739\,1 & k=1 \\ 0.130\,431\,17 & k=2 \\ 0 & k>2 \end{cases} \tag{1.5.58}$$

z_t 及其标准相关函数 ρ_k 图形如图 1.5.1 所示，可看到 ρ_k 在 $k=2$ 处截尾. 由于 $\rho_1 = -0.621\,739\,1$，因而 z_t 的样本函数有较快的波动.



(a) MA(2) 过程 z_t 的一个样本曲线



(b) MA(2) 过程 z_t 的相关函数

图 1.5.1　MA(2)过程 z_t 的样本函数和标准相关函数

1.5.3 用差分方程递推计算平稳 AR(p)过程相关函数

考虑 AR(p)过程

$$\varphi(q^{-1})z_t = a_t \tag{1.5.59}$$

其中 $\varphi(q^{-1}) = 1 - \varphi_1 q^{-1} - \cdots - \varphi_p q^{-p}$，$a_t$ 是零均值、方差为 σ_a^2 的白噪声，它可写为

$$z_t = \varphi_1 z_{t-1} + \cdots + \varphi_p z_{t-p} + a_t \tag{1.5.60}$$

上式两边乘 z_{t-k}($k > 0$)取数学期望，有差分方程

$$\gamma_k = \varphi_1\gamma_{k-1} + \varphi_2\gamma_{k-2} + \cdots + \varphi_p\gamma_{k-p}, \quad k > 0 \tag{1.5.61}$$

其中用到事实 $\mathrm{E}[z_{t-k}a_t] = 0$($k > 0$). 这一事实由展式(1.5.1)和 a_t 为白噪声得证. 差分方程(1.5.61)可简写为

$$\varphi(q^{-1})\gamma_k = 0, \quad k > 0 \tag{1.5.62}$$

用 z_t 乘(1.5.60)后取数学期望，并注意 $\gamma_k = \gamma_{-k}$，$\mathrm{E}[z_t a_t] = \sigma_a^2$(这一事实由展式(1.5.1)得到)，则有 z_t 的方差 σ_z^2 为

$$\sigma_z^2 = \gamma_0 = \varphi_1\gamma_1 + \varphi_2\gamma_2 + \cdots + \varphi_p\gamma_p + \sigma_a^2 \tag{1.5.63}$$

由(1.5.61)有标准相关函数 $\rho_k = \gamma_k/\gamma_0$ 所满足的差分方程

$$\rho_k = \varphi_1\rho_{k-1} + \varphi_2\rho_{k-2} + \cdots + \varphi_p\rho_{k-p}, \quad k > 0 \tag{1.5.64}$$

或

$$\varphi(q^{-1})\rho_k = 0, \quad k > 0 \tag{1.5.65}$$

由(1.5.63)引出关系

$$1 = \varphi_1\rho_1 + \varphi_2\rho_2 + \cdots + \varphi_p\rho_p + \sigma_a^2/\sigma_z^2 \tag{1.5.66}$$

这引出 z_t 的方差的另一表达式为

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \varphi_1\rho_1 - \cdots - \varphi_p\rho_p} \tag{1.5.67}$$

在(1.5.64)中取 $k = 1, \cdots, p$，注意 $\rho_k = \rho_{-k}$，$\rho_0 = 1$，可得到著名的 Yule – Walker 方程

$$\begin{aligned}
\rho_1 &= \varphi_1\rho_0 + \varphi_2\rho_1 + \cdots + \varphi_p\rho_{p-1} \\
\rho_2 &= \varphi_1\rho_1 + \varphi_2\rho_0 + \cdots + \varphi_p\rho_{p-2} \\
&\vdots \\
\rho_p &= \varphi_1\rho_{p-1} + \varphi_2\rho_{p-2} + \cdots + \varphi_p\rho_0
\end{aligned} \tag{1.5.68}$$

引入记号

$$\boldsymbol{\rho} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}, \quad \boldsymbol{\varphi} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix}, \quad \boldsymbol{P}_\rho = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{bmatrix} \tag{1.5.69}$$

则 Yule – Walker 方程可简写为

$$\boldsymbol{P}_\rho\boldsymbol{\varphi} = \boldsymbol{\rho} \tag{1.5.70}$$

因而 AR 模型参数向量 $\boldsymbol{\varphi}$ 可由标准相关函数计算为

$$\boldsymbol{\varphi} = \boldsymbol{P}_\rho^{-1}\boldsymbol{\rho} \tag{1.5.71}$$

上述结果可概括为如下定理.

【定理 1.5.3】 平稳 AR(p)过程的相关函数满足差分方程(1.5.61)或(1.5.62),它的方差由(15.63)计算.它的模型参数可由 Yule – Walker 方程式(1.5.70)计算,计算公式为(1.5.71).差分方程(1.5.64)的初值($\rho_{p-1}\rho_{p-2}\cdots\rho_1$)可在(1.5.64)中置 $k=1,2,\cdots,p-1$ 后解线性方程组得到.

【例 1.5.6】 考虑例 1.5.1 的平稳 AR(1)过程

$$(1-\varphi q^{-1})z_t = a_t, \quad |\varphi|<1 \tag{1.5.72}$$

由定理 1.5.3,它的相关函数 ρ_k 满足差分方程

$$\rho_k = \varphi\rho_{k-1}, \quad \rho_0 = 1 \tag{1.5.73}$$

这引出差分方程的解为

$$\rho_k = \varphi^k \tag{1.5.74}$$

由(1.5.67)有 z_t 的方差为

$$\sigma_z^2 = \frac{\sigma_a^2}{1-\varphi\rho_1} = \frac{\sigma_a^2}{1-\varphi^2} \tag{1.5.75}$$

上述结果相同于例 1.5.1 用 MA(∞)展式所得结果.

【例 1.5.7】 考虑平稳 AR(2)过程

$$(1-\varphi_1 q^{-1}-\varphi_2 q^{-2})z_t = a_t \tag{1.5.76}$$

它的相关函数满足差分方程

$$\rho_k = \varphi_1\rho_{k-1}+\varphi_2\rho_{k-2}, \quad k>0 \tag{1.5.77}$$

置 $k=1$ 有线性方程

$$\rho_1 = \varphi_1\rho_0+\varphi_2\rho_1, \quad \rho_0 = 1 \tag{1.5.78}$$

可解出

$$\rho_1 = \frac{\varphi_1}{1-\varphi_2} \tag{1.5.79}$$

于是由(1.5.77)带初值 $\rho_1=\varphi_1/(1-\varphi_2)$ 和 $\rho_0=1$ 可递推计算 ρ_k.

【例 1.5.8】 考虑平稳 AR(2)过程 z_t,

$$(1-1.2q^{-1}+0.3q^{-2})z_t = a_t \tag{1.5.80}$$

其中 a_t 是零均值、方差为 $\sigma_a^2=1$ 的正态白噪声. $\varphi(q^{-1})$ 的零点为 $q^{-1}=1/0.4$,$q^{-1}=1/0.8$位于单位圆外.故 z_t 平稳.由(1.5.77),它有标准相关函数

$$\rho_k = 1.2\rho_{k-1}-0.3\rho_{k-2} \tag{1.5.81}$$

由(1.5.79)有

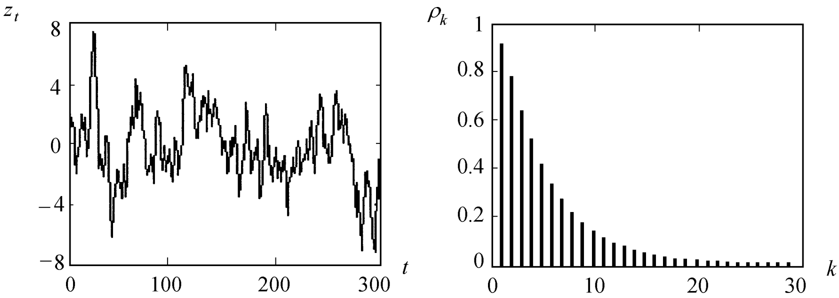
$$\rho_1 = \frac{1.2}{1+0.32} = 0.909 \tag{1.5.82}$$

注意 $\rho_0=1$,于是可用 ρ_0,ρ_1 作为初值由(1.5.81)递推计算 ρ_k. z_t 的一个样本曲线及它的标准相关函数ρ_k 图形如图 1.5.2 所示,可看到 ρ_k 呈指数衰减趋于零,因而具有拖尾性质.

【例 1.5.9】 考虑平稳 AR(2)过程 z_t,

$$(1-0.75q^{-1}+0.5q^{-2})z_t = a_t \tag{1.5.83}$$

其中 a_t 是零均值、方差为 $\sigma_a^2=1$ 的正态白噪声.容易验证 $\varphi(x)=0$ 有在单位圆外的复



(a) AR(2)过程 z_t 的一个样本曲线　　(b) AR(2)过程 z_t 的相关函数 ρ_k

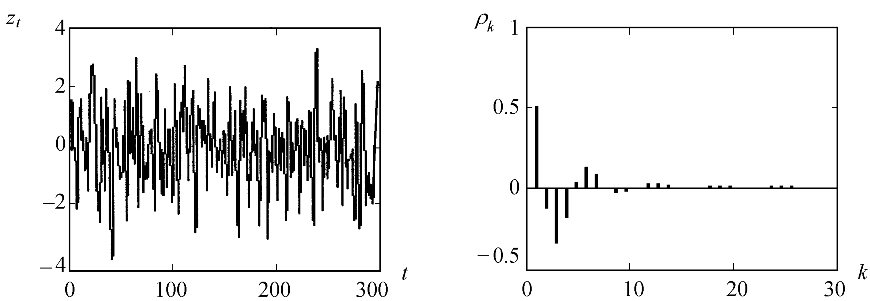
图 1.5.2　平稳 AR(2)过程 z_t 的样本函数和标准相关函数

根.类似于例 1.5.7 可求得它的标准相关函数 ρ_k 为

$$\rho_k = 0.75\rho_{k-1} - 0.5\rho_{k-2} \tag{1.5.84}$$

$$\rho_0 = 1,\quad \rho_1 = 0.5 \tag{1.5.85}$$

z_t 的一个样本曲线和 ρ_k 曲线如图 1.5.3 所示,可看到 ρ_k 呈指数正弦衰减振荡趋于零,因而也具有拖尾性质.



(a) AR(2)过程 z_t 的一个样本曲线　　(b) AR(2)过程 z_t 的相关函数 ρ_k

图 1.5.3　平稳 AR(2)过程 z_t 的样本函数和标准相关函数

1.5.4　ARMA 过程相关函数的递推计算

考虑平稳、可逆的 ARMA(p,q)(1.4.8)~(1.4.10),即

$$z_t = \varphi_1 z_{t-1} + \cdots + \varphi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \tag{1.5.86}$$

定义相关函数

$$\gamma_k = \mathrm{E}[z_{t-k} z_t],\quad \gamma_{za}(i) = \mathrm{E}[z(t-i)a_t] \tag{1.5.87}$$

用 z_{t-k} 左乘(1.5.86)后取数学期望有递推关系

$$\gamma_k = \varphi_1\gamma_{k-1} + \cdots + \varphi_p\gamma_{k-p} + \gamma_{za}(k) - \theta_1\gamma_{za}(k-1) - \cdots - \theta_q\gamma_{za}(k-q) \tag{1.5.88}$$

应用展式(1.5.1)有

$$\gamma_{za}(i) = \begin{cases} 0 & i > 0 \\ \psi_{-i}\sigma_a^2 & i \leqslant 0 \end{cases} \tag{1.5.89}$$

由上两式引出如下定理.

【定理 1.5.4】　平稳可逆 ARMA 过程(1.4.8)~(1.4.10)的相关函数 γ_k 可递推计算为

$$\gamma_k = \varphi_1\gamma_{k-1} + \cdots + \varphi_p\gamma_{k-p} + (\theta_k\psi_0 + \theta_{k+1}\psi_1 + \cdots + \theta_q\psi_{q-k})\sigma_a^2,$$
$$k = 0,1,\cdots,q \tag{1.5.90}$$

其中规定 $\theta_0 = -1, \psi_0 = 1$.

$$\gamma_k = \varphi_1\gamma_{k-1} + \cdots + \varphi_p\gamma_{k-p}, \quad k > q \tag{1.5.91}$$

在上两式中取 $k = 0,1,\cdots,p$,利用关系 $\gamma_k = \gamma_{-k}$,解线性方程组可解出 $\gamma_0,\gamma_1,\cdots,\gamma_p$,将它们作为(1.5.90)和(1.5.91)的初值,便可递推计算 γ_k.

当 $k=0$ 时有 z_t 的方差 σ_z^2 为

$$\sigma_z^2 = \gamma_0 = \varphi_1\gamma_1 + \cdots + \varphi_p\gamma_p + \sigma_a^2(1 - \theta_1\psi - \cdots - \theta_q\psi_q) \tag{1.5.92}$$

【例 1.5.10】 考虑 ARMA(1,1)过程 z_t

$$z_t - \varphi_1 z_{t-1} = a_t - \theta_1 a_{t-1}, \quad |\varphi_1| < 1, |\theta_1| < 1 \tag{1.5.93}$$

由定理 1.5.4 有

$$\gamma_0 = \varphi_1\gamma_1 - (-1 + \theta_1\psi_1)\sigma_a^2 \tag{1.5.94}$$

$$\gamma_1 = \varphi_1\gamma_0 - \theta_1\sigma_a^2 \tag{1.5.95}$$

$$\gamma_k = \varphi_1\gamma_{k-1}, \quad k > 1 \tag{1.5.96}$$

其中由(1.5.2)有 $\psi_1 = \gamma_1 - \theta_1$.解方程组(1.5.94)和(1.5.95)可求得 γ_0 和 γ_1,然后以 γ_1 作为初值由(1.5.96)可递推计算 γ_k.

1.6 状态空间模型

现代控制理论的创始人之一 R.E.Kalman[4]的重大贡献之一是提出了描述动态系统的状态空间方法.其核心思想有三点:(1)引入状态变量概念;(2)建立描述状态变化的模型——状态方程;(3)给出了对状态进行观测的观测方程.状态变量是对动态系统特征的概括,通常用 $n \times 1$ 列向量表示,n 叫状态的维数.在时刻 t 状态 $\boldsymbol{x}(t)$ 取值于 n 维欧氏空间 R^n 中的点,即 $\boldsymbol{x}(t) \in R^n$,称 R^n 为状态空间.为了直观理解状态空间方法,考虑如下启发性例子.

【例 1.6.1】 考虑雷达跟踪系统[1]

$$s(t+1) = s(t) + \dot{s}(t)T + [u + w(t)]T^2/2 \tag{1.6.1}$$

$$\dot{s}(t+1) = \dot{s}(t) + [u + w(t)]T \tag{1.6.2}$$

$$y(t) = s(t) + v(t) \tag{1.6.3}$$

其中 T 为采样周期,$s(t)$,$\dot{s}(t)$ 和 $[u + w(t)]$ 各为在时刻 tT 运动目标的位置、速度和加速度.其中 $w(t)$ 为随机加速度,它是零均值、方差为 σ_w^2 的白噪声,而 u 是常数机动加速度.$v(t)$ 为零均值、方差为 σ_v^2 的独立于 $w(t)$ 的观测白噪声,$y(t)$ 为对位置的观测信号.若我们感兴趣的问题是对运动目标的位置和速度的估计问题,则可引入状态变量为

$$\boldsymbol{x}(t) = [s(t), \dot{s}(t)]^{\mathrm{T}} \tag{1.6.4}$$

其中上角 T 为转置号.于是由(1.6.1)~(1.6.3)有状态方程和观测方程分别为

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi x}(t) + \boldsymbol{B}u + \boldsymbol{\Gamma}w(t) \tag{1.6.5}$$

$$y(t) = \boldsymbol{Hx}(t) + v(t) \tag{1.6.6}$$

其中 $\boldsymbol{\Phi}$ 叫状态转移阵,$\boldsymbol{H}$ 叫观测阵,且显然有

$$\boldsymbol{\Phi}=\begin{bmatrix}1 & T\\ 0 & 1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}T^2/2\\ T\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}T^2/2\\ T\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0] \tag{1.6.7}$$

因而雷达跟踪问题化为基于观测$(y(t),y(t-1),\cdots)$求状态$\boldsymbol{x}(t)$的最优估值器$\hat{\boldsymbol{x}}(t|t)$.

【例 1.6.2】 未知常量的测量估计问题.对某一未知常量c(c可代表长度、距离、半径、体积等)进行测量,设在离散时刻t对c的测量值为$y(t)$,其中含有观测噪声$v(t)$.取状态为$x(t)=c$,则有状态方程和观测方程分别为

$$x(t+1)=x(t) \tag{1.6.8}$$

$$y(t)=x(t)+v(t) \tag{1.6.9}$$

一般时不变(定常)线性离散随机系统的状态空间模型为

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{Bu}(t)+\boldsymbol{\Gamma w}(t) \tag{1.6.10}$$

$$\boldsymbol{y}(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t) \tag{1.6.11}$$

其中t为离散时间,状态$\boldsymbol{x}(t)\in R^n$,观测$\boldsymbol{y}(t)\in R^m$,控制$\boldsymbol{u}(t)\in R^p$,输入噪声$\boldsymbol{w}(t)\in R^r$和观测噪声$\boldsymbol{v}(t)\in R^m$是带零均值的相关白噪声

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix}[\boldsymbol{w}^{\mathrm{T}}(j),\boldsymbol{v}^{\mathrm{T}}(j)]\right\}=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}\delta_{tj} \tag{1.6.12}$$

其中$\boldsymbol{Q}_w$和$\boldsymbol{Q}_v$各为$\boldsymbol{w}(t)$和$\boldsymbol{v}(t)$的方差阵,$\boldsymbol{S}$为它们的相关阵,$\delta_{tt}=1$,$\delta_{tj}=0\ (t\neq j)$,E为数学期望号,称$\boldsymbol{\Phi}$为状态转移阵,$\boldsymbol{H}$为观测阵.$\boldsymbol{\Phi}$,$\boldsymbol{B}$,$\boldsymbol{\Gamma}$,$\boldsymbol{H}$分别为$n\times n$,$n\times p$,$n\times r$,$m\times n$阵.特别地若$\boldsymbol{u}(t)=\boldsymbol{0}$,则状态空间模型有形式

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma}w(t) \tag{1.6.13}$$

$$\boldsymbol{y}(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t) \tag{1.6.14}$$

1.6.1 状态变量的非递推表达式

对状态空间模型(1.6.13)和(1.6.14)进行迭代,可得$\boldsymbol{x}(t)$是$\boldsymbol{x}(0)$,$\boldsymbol{w}(0)$,$\cdots$,$\boldsymbol{w}(t-1)$的线性组合,即

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}^t\boldsymbol{x}(0)+\sum_{i=1}^{t}\boldsymbol{\Phi}^{t-i}\boldsymbol{\Gamma w}(i-1) \tag{1.6.15}$$

当$t>j$时,也类似可得到非递推表达式

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}^{t-j}\boldsymbol{x}(j)+\sum_{i=j+1}^{t}\boldsymbol{\Phi}^{t-i}\boldsymbol{\Gamma w}(i-1) \tag{1.6.16}$$

上述$\boldsymbol{x}(t)$的表达式与初始状态$\boldsymbol{x}(0)$或$\boldsymbol{x}(j)$有关.为了避免出现初始状态,设$t_0=-\infty$,考虑稳态情形.设$\boldsymbol{\Phi}$为稳定矩阵,由(1.6.13)有传递函数阵模型

$$\boldsymbol{x}(t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma w}(t-1) \tag{1.6.17}$$

利用展式

$$(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}=\sum_{j=0}^{\infty}\boldsymbol{\Phi}^j q^{-j} \tag{1.6.18}$$

则$\boldsymbol{x}(t)$可表为均方收敛的级数

$$\boldsymbol{x}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Phi}^j\boldsymbol{\Gamma w}(t-1-j) \tag{1.6.19}$$

1.6.2 可观性与可估计性

在什么条件下状态 $\boldsymbol{x}(t)$是可估计的？由状态空间模型(1.6.13)和(1.6.14)迭代有关系

$$\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{\beta-1} \end{bmatrix} \boldsymbol{x}(t) = \begin{bmatrix} \boldsymbol{y}(t) - \boldsymbol{v}(t) \\ \boldsymbol{y}(t+1) - \boldsymbol{H\Gamma w}(t) - \boldsymbol{v}(t+1) \\ \vdots \\ \boldsymbol{y}(t+\beta-1) - \sum_{i=0}^{\beta-2} \boldsymbol{H\Phi}^{\beta-2-i}\boldsymbol{\Gamma w}(t+i) - \\ \boldsymbol{v}(t+\beta-1) \end{bmatrix} \tag{1.6.20}$$

为了使状态 $\boldsymbol{x}(t)$能表为 $\boldsymbol{y}(t+i)$，$\boldsymbol{w}(t+i)$和 $\boldsymbol{v}(t+i)$的线性组合，要求 $\beta m \times n$ 阵

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{\beta-1} \end{bmatrix} \tag{1.6.21}$$

为列满秩，即 rank $\boldsymbol{\Omega} = n$，此时伪逆 $\boldsymbol{\Omega}^{+}$存在为 $\boldsymbol{\Omega}^{+} = (\boldsymbol{\Omega}^{\mathrm{T}}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathrm{T}}$. 将伪逆 $\boldsymbol{\Omega}^{+}$分块表示为

$$\boldsymbol{\Omega}^{+} = [\boldsymbol{\Omega}_0, \boldsymbol{\Omega}_1, \cdots, \boldsymbol{\Omega}_{\beta-1}] \tag{1.6.22}$$

其中 $\boldsymbol{\Omega_i}$ 为 $n \times m$ 阵，则 $\boldsymbol{x}(t)$有非递推表达式

$$\boldsymbol{x}(t) = \sum_{i=0}^{\beta-1} \boldsymbol{\Omega}_i \left[\boldsymbol{y}(t+i) - \sum_{j=0}^{i-1} \boldsymbol{H\Phi}^{i-1-j}\boldsymbol{\Gamma w}(t+j) - \boldsymbol{v}(t+i)\right] \tag{1.6.23}$$

其中规定 $\boldsymbol{\Phi}^i = \boldsymbol{0}(i<0)$，且 $j \geqslant 0$，即规定当 $i=0$，上式中括号内的第二项为零.

注意，由线性系统理论[5]$\boldsymbol{\Omega}$ 是系统的可观阵，rank $\boldsymbol{\Omega} = n$ 恰好是系统为完全可观的条件. 这里 β 是使 rank $\boldsymbol{\Omega} = n$ 的最小自然数，称为可观性指数. 因此，在系统完全可观条件下，表达式(1.6.23)成立. 因而，可将状态 $\boldsymbol{x}(t)$的估计问题转化为观测预报器和白噪声估值器的计算. 因此，系统完全可观条件也是系统状态可估计的充分条件. [1,11]

1.6.3 时变线性离散随机系统的非递推状态表达式

考虑时变线性随机系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}(t+1)\boldsymbol{x}(t) + \boldsymbol{\Gamma}(t)\boldsymbol{w}(t) \tag{1.6.24}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}(t)\boldsymbol{x}(t) + \boldsymbol{v}(t) \tag{1.6.25}$$

其中 $\boldsymbol{\Phi}(t)$，$\boldsymbol{\Gamma}(t)$和 $\boldsymbol{H}(t)$是时变的. 当 $t>j$ 时，由上两式迭代有非递推状态表达式

$$\boldsymbol{x}(t) = \boldsymbol{\Phi}(t,j)\boldsymbol{x}(j) + \sum_{i=j+1}^{t} \boldsymbol{\Phi}(t,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1) \tag{1.6.26}$$

其中定义 $\boldsymbol{\Phi}(t,t) = \boldsymbol{I}_n$

$$\boldsymbol{\Phi}(t,i) = \boldsymbol{\Phi}(t)\boldsymbol{\Phi}(t-1)\cdots\boldsymbol{\Phi}(i+1), \quad t>i \tag{1.6.27}$$

若将(1.6.24)改为如下形式

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}(t)\boldsymbol{x}(t) + \boldsymbol{\Gamma w}(t) \tag{1.6.28}$$

则 $\boldsymbol{\Phi}(t,i)$的定义应修改为

$$\boldsymbol{\Phi}(t,i)=\boldsymbol{\Phi}(t-1)\boldsymbol{\Phi}(t-2)\cdots\boldsymbol{\Phi}(i),$$
$$\boldsymbol{\Phi}(t,t)=\boldsymbol{I}_n \tag{1.6.29}$$

1.6.4 并联状态空间模型

考虑信号滤波模型

$$\boldsymbol{y}(t)=\boldsymbol{s}(t)+\boldsymbol{\eta}(t)+\boldsymbol{v}(t) \tag{1.6.30}$$

其中 $\boldsymbol{y}(t)$为观测信号，$\boldsymbol{s}(t)$为待估信号，$\boldsymbol{\eta}(t)$为有色观测噪声，$\boldsymbol{v}(t)$为白色观测噪声. 设 $\boldsymbol{s}(t)$与 $\boldsymbol{\eta}(t)$的状态空间模型分别为

$$\boldsymbol{a}(t+1)=\boldsymbol{\Phi a}(t)+\boldsymbol{\Gamma e}(t) \tag{1.6.31}$$
$$\boldsymbol{s}(t)=\boldsymbol{H}_1\boldsymbol{a}(t) \tag{1.6.32}$$
$$\boldsymbol{\beta}(t+1)=\boldsymbol{A\beta}(t)+\boldsymbol{C\xi}(t) \tag{1.6.33}$$
$$\boldsymbol{\eta}(t)=\boldsymbol{H}_2\boldsymbol{\beta}(t) \tag{1.6.34}$$

其中 $\boldsymbol{e}(t)$和 $\boldsymbol{\xi}(t)$为白噪声. 为了求 $\boldsymbol{s}(t)$的最优滤波器，引入增广状态

$$\boldsymbol{x}(t)=[\boldsymbol{a}^{\mathrm{T}}(t),\boldsymbol{\beta}^{\mathrm{T}}(t)]^{\mathrm{T}} \tag{1.6.35}$$

则有增广状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{Fx}(t)+\boldsymbol{Gw}(t) \tag{1.6.36}$$
$$y(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t) \tag{1.6.37}$$

其中定义

$$\boldsymbol{F}=\begin{bmatrix}\boldsymbol{\Phi} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}\end{bmatrix},\quad \boldsymbol{G}=\begin{bmatrix}\boldsymbol{\Gamma} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{C}\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{H}_1\quad \boldsymbol{H}_2],\quad \boldsymbol{w}(t)=\begin{bmatrix}\boldsymbol{e}(t)\\ \boldsymbol{\xi}(t)\end{bmatrix} \tag{1.6.38}$$

并称其为并联增广状态空间模型，其中观测信号含有两模型的输出相加的信息.

【例 1.6.3】 考虑雷达跟踪系统

$$s(t+1)=s(t)+\dot{s}(t)+e(t)T^2/2 \tag{1.6.39}$$
$$\dot{s}(t+1)=\dot{s}(t)+e(t)T \tag{1.6.40}$$
$$y(t)=s(t)+\eta(t)+v(t) \tag{1.6.41}$$
$$\eta(t+1)=a\eta(t)+\xi(t) \tag{1.6.42}$$

其中 T 为采样周期，$s(t)$，$\dot{s}$ 和 $e(t)$各为在时刻 tT 运动目标的位置、速度和加速度，$y(t)$为观测，$\eta(t)$为有色观测噪声，服从 AR(1)模型，$v(t)$为白色观测噪声. 问题是由被噪声污染的对位置的观测信号 $y(t)$求位置 $s(t)$的最优滤波器. 为此应建立增广状态空间模型. 由(1.6.39)和(1.6.40)有 $s(t)$的状态空间模型

$$\boldsymbol{a}(t+1)=\boldsymbol{\Phi a}(t)+\boldsymbol{\Gamma}e(t) \tag{1.6.43}$$
$$s(t)=\boldsymbol{H}_1\boldsymbol{a}(t) \tag{1.6.44}$$

其中定义

$$\boldsymbol{\Phi}=\begin{bmatrix}1 & T\\ 0 & 1\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}T^2/2\\ T\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0] \tag{1.6.45}$$

而有色观测噪声 $\eta(t)$有状态空间模型

$$\beta(t+1)=a\beta(t)+\xi(t) \tag{1.6.46}$$
$$\eta(t)=\beta(t) \tag{1.6.47}$$

引入增广状态 $\boldsymbol{x}(t)=[\boldsymbol{a}^{\mathrm{T}}(t),\beta(t)]^{\mathrm{T}}$ 和增广噪声 $\boldsymbol{w}(t)=[e(t),\xi(t)]^{\mathrm{T}}$ 则有并联增广状态空间模型

$$\boldsymbol{x}(t+1)=\begin{bmatrix}1 & T & 0\\ 0 & 1 & 0\\ 0 & 0 & a\end{bmatrix}x(t)+\begin{bmatrix}T^2/2 & 0\\ T & 0\\ 0 & 1\end{bmatrix}\boldsymbol{w}(t) \tag{1.6.48}$$

$$y(t)=[1\quad 0\quad 1]\boldsymbol{x}(t)+v(t) \tag{1.6.49}$$

1.6.5 串联状态空间模型

考虑信号 $\boldsymbol{s}(t)$的反卷积模型

$$\boldsymbol{a}(t+1)=\boldsymbol{\Phi a}(t)+\boldsymbol{\Gamma s}(t) \tag{1.6.50}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}_1\boldsymbol{a}(t)+\boldsymbol{v}(t) \tag{1.6.51}$$

$$\boldsymbol{\beta}(t+1)=\boldsymbol{A\beta}(t)+\boldsymbol{C\xi}(t) \tag{1.6.52}$$

$$\boldsymbol{s}(t)=\boldsymbol{H}_2\boldsymbol{\beta}(t) \tag{1.6.53}$$

其中(1.6.52)和(1.6.53)为信号状态空间模型,(1.6.50)和(1.6.51)为对信号 $\boldsymbol{s}(t)$的观测模型,而 $\boldsymbol{s}(t)$恰好为该观测模型的输入.因而由观测 $\boldsymbol{y}(t)$求 $\boldsymbol{s}(t)$的最优滤波问题叫反卷积.为此,应引入增广状态空间模型.引入增广状态 $\boldsymbol{x}(t)=[\boldsymbol{a}^{\mathrm{T}}(t),\boldsymbol{\beta}^{\mathrm{T}}(t)]^{\mathrm{T}}$,则有增广状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{Fx}(t)+\boldsymbol{G\xi}(t) \tag{1.6.54}$$

$$\boldsymbol{y}(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t) \tag{1.6.55}$$

其中定义

$$\boldsymbol{F}=\begin{bmatrix}\boldsymbol{\Phi} & \boldsymbol{\Gamma H}_2\\ \boldsymbol{0} & \boldsymbol{A}\end{bmatrix},\quad \boldsymbol{G}=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{C}\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{H}_1\quad \boldsymbol{0}], \tag{1.6.56}$$

并称其为串联增广状态空间模型,其中一个模型的输出恰好为另一模型的输入.

【例 1.6.4】 考虑雷达跟踪系统

$$\boldsymbol{a}(t+1)=\boldsymbol{\Phi a}(t)+\boldsymbol{\Gamma}e(t) \tag{1.6.57}$$

$$y(t)=\boldsymbol{Hx}(t)+v(t) \tag{1.6.58}$$

$$e(t+1)=ae(t)+\xi(t) \tag{1.6.59}$$

其中 T 为采样周期,$a(t)=[s(t),\dot{s}(t)]^{\mathrm{T}}$,$s(t)$,$\dot{s}(t)$和 $e(t)$各为在时刻 tT 运动目标的位置、速度和加速度,且

$$\boldsymbol{\Phi}=\begin{bmatrix}1 & T\\ 0 & 1\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}T^2/2\\ T\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0] \tag{1.6.60}$$

加速度 $e(t)$服从 AR(1)模型(1.6.59),$y(t)$为对位置 $s(t)$的观测信号,$\xi(t)$和 $v(t)$为白噪声.问题是由 $y(t)$求加速度 $e(t)$的最优滤波器.为此,应建立串联状态空间模型.引入增广状态 $\boldsymbol{x}(t)=[\boldsymbol{a}^{\mathrm{T}}(t),e(t)]^{\mathrm{T}}$,容易验证串联增广状态空间模型为

$$\boldsymbol{x}(t+1)=\begin{bmatrix}1 & T & T^2/2\\ 0 & 1 & T\\ 0 & 0 & a\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}0\\ 0\\ 1\end{bmatrix}\xi(t) \tag{1.6.61}$$

$$y(t)=[1\quad 0\quad 0]\boldsymbol{x}(t)+v(t) \tag{1.6.62}$$

1.6.6 解耦状态空间模型

若状态空间模型具有形式

$$\begin{bmatrix} \boldsymbol{x}_1(t+1) \\ \boldsymbol{x}_2(t+1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Phi}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1(t) \\ \boldsymbol{x}_2(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Gamma}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Gamma}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_1(t) \\ \boldsymbol{w}_2(t) \end{bmatrix} \tag{1.6.63}$$

$$\begin{bmatrix} \boldsymbol{y}_1(t) \\ \boldsymbol{y}_2(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1(t) \\ \boldsymbol{x}_2(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{v}_1(t) \\ \boldsymbol{v}_2(t) \end{bmatrix} \tag{1.6.64}$$

其中 $\boldsymbol{w}_1(t)$和 $\boldsymbol{w}_2(t)$不相关，噪声 $\boldsymbol{v}_1(t)$和 $\boldsymbol{v}_2(t)$不相关，则它可分解为两个独立的状态空间模型

$$\boldsymbol{x}_1(t+1) = \boldsymbol{\Phi}_1\boldsymbol{x}_1(t) + \boldsymbol{\Gamma}_1\boldsymbol{w}_1(t) \tag{1.6.65}$$

$$\boldsymbol{y}_1(t) = \boldsymbol{H}_1\boldsymbol{x}_1(t) + \boldsymbol{v}_1(t) \tag{1.6.66}$$

$$\boldsymbol{x}_2(t+1) = \boldsymbol{\Phi}_2\boldsymbol{x}_2(t) + \boldsymbol{\Gamma}_2\boldsymbol{w}_2(t) \tag{1.6.67}$$

$$\boldsymbol{y}_2(t) = \boldsymbol{H}_2\boldsymbol{x}_2(t) + \boldsymbol{v}_2(t) \tag{1.6.68}$$

称其为解耦状态空间模型.它具有降低状态估值器维数，可减小计算负担的优点.

1.6.7 状态空间模型转化为 ARMA 新息模型

解决状态估计问题的一种新途径是借助于等价于观测信息(观测模型)的 ARMA 新息模型.它是由作者[1,6~8]提出的现代时间序列分析方法的基本工具.

为简单计，考虑单输入单输出定常系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}\boldsymbol{x}(t) + \boldsymbol{\Gamma}w(t) \tag{1.6.69}$$

$$y(t) = \boldsymbol{H}\boldsymbol{x}(t) + v(t) \tag{1.6.70}$$

其中状态 $\boldsymbol{x}(t)\in R^n$，输出观测 $y(t)\in R^1$ 为标量，输入 $w(t)\in R^1$ 为标量，$\boldsymbol{\Phi}$，$\boldsymbol{\Gamma}$，$\boldsymbol{H}$ 各为 $n\times n$，$n\times 1$ 和 $1\times n$ 常阵，$w(t)$和 $v(t)$是零均值、方差各为 σ_w^2 和 σ_v^2 的独立白噪声.假设系统是完全可观、完全可控的[5].由(1.6.69)，$\boldsymbol{x}(t)$可表为

$$\boldsymbol{x}(t) = (\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}w(t-1) \tag{1.6.71}$$

其中 q^{-1}是单位滞后算子.将上式代入(1.6.70)有

$$y(t) = \boldsymbol{H}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}w(t-1) + v(t) \tag{1.6.72}$$

将矩阵求逆公式

$$(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1} = \mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})/\varphi(q^{-1}) \tag{1.6.73}$$

$$\varphi(q^{-1}) = \det(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi}) \tag{1.6.74}$$

代入(1.6.72)

$$y(t) = \frac{\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})\boldsymbol{\Gamma}}{\varphi(q^{-1})}w(t-1) + v(t) \tag{1.6.75}$$

由系统完全可观、完全可控的假设引出上式分母与分子无公因式相消.用 $\varphi(q^{-1})$乘上式两边引出

$$\varphi(q^{-1})y(t) = \boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})\boldsymbol{\Gamma}w(t-1) + \varphi(q^{-1})v(t) \tag{1.6.76}$$

上式右边的两个滑动平均(MA)过程可用一个等价的稳定的 MA 过程表示

$$\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})\boldsymbol{\Gamma}w(t-1) + \varphi(q^{-1})v(t) = D(q^{-1})\varepsilon(t) \tag{1.6.77}$$

其中 $\varepsilon(t)$是零均值、方差为 σ_ε^2 的白噪声，$D(q^{-1})$是首系数为 1 关于 q^{-1}的稳定多项式. 于是可得观测过程 $y(t)$的 ARMA 模型

$$\varphi(q^{-1})y(t)=D(q^{-1})\varepsilon(t) \tag{1.6.78}$$

由第三章射影理论可证明白噪声 $\varepsilon(t)$是 $y(t)$的新息过程，因而称(1.6.78)为 ARMA 新息模型. 其中 $D(q^{-1})$和 σ_ε^2 可用第二章给出的 Gevers – Wouters 算法求得.

【例 1.6.5】 考虑单输入单输出随机系统，它有状态空间模型

$$\boldsymbol{x}(t+1)=\begin{bmatrix}1 & 0\\ 2 & 0\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\ 1\end{bmatrix}w(t) \tag{1.6.79}$$

$$y(t)=[1 \quad 1]\boldsymbol{x}(t)+v(t) \tag{1.6.80}$$

其中 $w(t)$和 $v(t)$是零均值、方差各为 $\sigma_w^2=0.1$ 和 $\sigma_v^2=0.1$ 的独立白噪声. 容易验证该系统是完全可观、完全可控的，即

$$\operatorname{rank}\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{H\Phi}\end{bmatrix}=\operatorname{rank}\begin{bmatrix}1 & 1\\ 3 & 0\end{bmatrix}=2 \tag{1.6.81}$$

$$\operatorname{rank}[\boldsymbol{\Gamma} \quad \boldsymbol{\Phi\Gamma}]=\operatorname{rank}\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}=2 \tag{1.6.82}$$

应用矩阵求逆公式有

$$y(t)=[1 \quad 1]\begin{bmatrix}1-q^{-1} & 0\\ -2q^{-1} & 1\end{bmatrix}^{-1}\begin{bmatrix}1\\ 1\end{bmatrix}w(t-1)+v(t) \tag{1.6.83}$$

$$(1-q^{-1})y(t)=(2+q^{-1})w(t-1)+(1-q^{-1})v(t) \tag{1.6.84}$$

将上式右边两个 MA 过程之和同一个等价的稳定的 MA 过程表示

$$(1+dq^{-1})\varepsilon(t)=(2+q^{-1})w(t-1)+(1-q^{-1})v(t) \tag{1.6.85}$$

其中 $\varepsilon(t)$是零均值、方差为 σ_ε^2 的白噪声. 计算上式两边 MA 过程的相关函数有

$$(1+d^2)\sigma_\varepsilon^2=5\sigma_w^2+2\sigma_v^2 \tag{1.6.86}$$

$$d\sigma_\varepsilon^2=2\sigma_w^2-\sigma_v^2 \tag{1.6.87}$$

将 $\sigma_w^2=0.1$ 和 $\sigma_v^2=0.1$ 代入上两式并消去 σ_ε^2 可得一元二次方程

$$d^2-7d+1=0 \tag{1.6.88}$$

可解出使$(1+dq^{-1})$稳定的 d 为

$$d=\frac{7-\sqrt{49-4}}{2}=0.145\,898 \tag{1.6.89}$$

进而由(1.6.87)可得

$$\sigma_\varepsilon^2=0.685\,410\,3 \tag{1.6.90}$$

从而由(1.6.84)和(1.6.85)最后可得 ARMA 新息模型

$$(1-q^{-1})y(t)=(1+0.145\,898q^{-1})\varepsilon(t) \tag{1.6.91}$$

在 $D(q^{-1})=1+d_1q^{-1}+\cdots+d_{n_d}q^{-n_d}$的阶次 $n_d>2$ 时，没有系数 d_i 和σ_ε^2 的解析解，需用第二章的 Gevers – Wouters 迭代算法[1,6,9,10]求 $D(q^{-1})$和 σ_ε^2.

1.6.8 由 ARMA 模型转化为状态空间模型

为了把信号滤波问题转化为状态估计问题，需要将信号的 ARMA 模型化为状态空间

模型.

【例 1.6.6】 继例 0.4 的雷达跟踪系统.

引入单位滞后算子 q^{-1},则运动目标在时刻 tT 的位置 $s(t)$、速度 $\dot{s}(t)$ 和加速度 $w(t)$ 有模型

$$(1-q^{-1})s(t)=\dot{s}(t-1)T+w(t-1)T^2/2 \tag{1.6.92}$$

$$(1-q^{-1})\dot{s}(t)=w(t-1)T \tag{1.6.93}$$

$$y(t)=s(t)+v(t) \tag{1.6.94}$$

由上两式消去 $\dot{s}(t)$ 可得关于位置 $s(t)$ 的 Wiener 滤波问题:

$$(1-q^{-1})^2 s(t)=\frac{T^2}{2}w(t-1)+\frac{T^2}{2}w(t-2) \tag{1.6.95}$$

$$y(t)=s(t)+v(t) \tag{1.6.96}$$

若我们仅对目标位置 $s(t)$ 感兴趣,而不感兴趣它的速度,则可构造不同于(0.22)~(0.24)的状态空间模型.容易验证引入状态变量 $\boldsymbol{x}(t)=[x_1(t),x_2(t)]^{\mathrm{T}}$,其中

$$x_1(t)=s(t) \tag{1.6.97}$$

$$x_2(t)=-x_1(t-1)+w(t-1)T^2/2 \tag{1.6.98}$$

可将 ARMA 模型(1.6.95)化为等价的状态空间模型

$$\boldsymbol{x}(t+1)=\begin{bmatrix}2 & 1\\ -1 & 0\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}T^2/2\\ T^2/2\end{bmatrix}w(t) \tag{1.6.99}$$

$$y(t)=[1\quad 0]\boldsymbol{x}(t)+v(t) \tag{1.6.100}$$

即由(1.6.99)和(1.6.100)可导出(1.6.95)和(1.6.96).在这个状态空间模型中,第一个分量 $x_1(t)=s(t)$ 有明显的物理意义,但其第二个分量 $x_2(t)$ 的定义(1.6.98)却没有物理意义.然而对估计位置 $s(t)$ 而言,我们对 $x_2(t)$ 有无物理意义不感兴趣.故亦可采用状态空间模型(1.6.99)和(1.6.100)将位置 $s(t)$ 的 Wiener 滤波问题(1.6.95)和(1.6.96)转化为 Kalman 滤波问题.其中被估信号 $s(t)$ 作为状态 $\boldsymbol{x}(t)$ 的第一个分量.此例表明对同一个 ARMA 模型,它的等价状态空间模型不是惟一的.所谓等价,是指它们具有相同的输出 $y(t)$.

在一般情形下我们有如下等价表示定理.

【定理 1.6.1】 向量 ARMA 模型

$$\boldsymbol{s}(t)+\boldsymbol{A}_1\boldsymbol{s}(t-1)+\cdots+\boldsymbol{A}_n\boldsymbol{s}(t-n)=\boldsymbol{C}_0\boldsymbol{e}(t)+\boldsymbol{C}_1\boldsymbol{e}(t-1)+\cdots+\boldsymbol{C}_n\boldsymbol{e}(t-n) \tag{1.6.101}$$

其中 $\boldsymbol{s}(t)\in R^m$, $\boldsymbol{e}(t)\in R^m$, $\boldsymbol{A}_i$ 和 $\boldsymbol{C}_i$ 为 $m\times m$ 阵,它等价于如下块伴随形状态空间模型

$$\begin{bmatrix}\boldsymbol{x}_1(t+1)\\ \boldsymbol{x}_2(t+1)\\ \vdots\\ \boldsymbol{x}_n(t+1)\end{bmatrix}=\begin{bmatrix}-\boldsymbol{A}_1 & & & \\ -\boldsymbol{A}_2 & \boldsymbol{I}_{(n-1)m} & & \\ \vdots & & & \\ -\boldsymbol{A}_n & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_1(t)\\ \boldsymbol{x}_2(t)\\ \vdots\\ \boldsymbol{x}_n(t)\end{bmatrix}+\begin{bmatrix}\boldsymbol{C}_1-\boldsymbol{A}_1\boldsymbol{C}_0\\ \boldsymbol{C}_2-\boldsymbol{A}_2\boldsymbol{C}_0\\ \vdots\\ \boldsymbol{C}_n-\boldsymbol{A}_n\boldsymbol{C}_0\end{bmatrix}\boldsymbol{e}(t) \tag{1.6.102}$$

$$s(t)=[\boldsymbol{I}_m \quad \boldsymbol{0} \quad \cdots \quad \boldsymbol{0}]\begin{bmatrix}\boldsymbol{x}_1(t)\\ \boldsymbol{x}_2(t)\\ \vdots \\ \boldsymbol{x}_n(t)\end{bmatrix}+\boldsymbol{C}_0\boldsymbol{e}(t) \tag{1.6.103}$$

证明 由(1.6.103)有 $\boldsymbol{s}(t)=\boldsymbol{x}_1(t)+\boldsymbol{C}_0\boldsymbol{e}(t)$，利用(1.6.102)和(1.6.103)逐项代入法可得(1.6.101). 详细推导从略.

我们以下面简单例子来说明定理 1.6.1 的证明思路，用逐次迭代法.

【例 1.6.7】 纯量 AR(2)模型

$$(1+a_1q^{-1}+a_2q^{-2})y(t)=e(t) \tag{1.6.104}$$

可化为状态空间模型

$$\begin{bmatrix}x_1(t+1)\\ x_2(t+1)\end{bmatrix}=\begin{bmatrix}-a_1 & 1\\ -a_2 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}+\begin{bmatrix}-a_1\\ -a_2\end{bmatrix}e(t) \tag{1.6.105}$$

$$y(t)=[1 \quad 0]\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}+e(t) \tag{1.6.106}$$

证明 由上两式逐次迭代有

$$\begin{aligned}y(t)=x_1(t)+e(t)&=-a_1x_1(t-1)+x_2(t-1)-a_1e(t-1)+e(t)=\\ &-a_1x_1(t-1)-a_2x_1(t-2)-a_2e(t-2)-a_1e(t-1)+e(t)=\\ &-a_1[y(t-1)-e(t-1)]-a_2[y(t-2)-e(t-2)]-a_2e(t-2)-\\ &a_1e(t-1)+e(t)=\\ &-a_1y(t-1)-a_2y(t-2)+e(t)\end{aligned} \tag{1.6.107}$$

它相同于(1.6.104). 证毕. □

【定义 1.6.1】 我们称 $nm\times nm$ 矩阵 $\boldsymbol{A}$ 和 $m\times nm$ 矩阵 $\boldsymbol{H}$，

$$\boldsymbol{A}=\begin{bmatrix}-\boldsymbol{A}_1 & & & \\ \vdots & & \boldsymbol{I}_{(n-1)m} & \\ -\boldsymbol{A}_n & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix}, \quad \boldsymbol{H}=[\boldsymbol{I}_m \quad \boldsymbol{0} \quad \cdots \quad \boldsymbol{0}] \tag{1.6.108}$$

为块伴随形$(\boldsymbol{A},\boldsymbol{H})$，其中 $\boldsymbol{A}_i$ 为 $m\times m$ 矩阵，$\boldsymbol{0}$ 为 $m\times m$ 零阵，$\boldsymbol{I}_r$ 为 $r\times r$ 单位阵. 块伴随形$(\boldsymbol{A},\boldsymbol{H})$在最优估计理论中具有重要作用，它的如下重要性质是今后经常用到的.

【定理 1.6.2】 (块伴随形矩阵的重要性质)若矩阵$(\boldsymbol{A},\boldsymbol{H})$具有块伴随形(1.6.108)，若定义多项式矩阵

$$\boldsymbol{A}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{A}_1q^{-1}+\cdots+\boldsymbol{A}_nq^{-n} \tag{1.6.109}$$

则有关系

(i)
$$\det(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})=\mathrm{et}\boldsymbol{A}(q^{-1}) \tag{1.6.110}$$

(ii)
$$\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})=\mathrm{adj}\boldsymbol{A}(q^{-1})[\boldsymbol{I}_m,\boldsymbol{I}_mq^{-1},\cdots,\boldsymbol{I}_mq^{-(n-1)}] \tag{1.6.111}$$

证明 先证(i). 利用分块矩阵求逆公式[10]

$$\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{0} & \boldsymbol{A}_{22}\end{bmatrix}^{-1}=\begin{bmatrix}\boldsymbol{A}_{11}^{-1}, & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\\ \boldsymbol{0} & \boldsymbol{A}_{22}^{-1}\end{bmatrix} \tag{1.6.112}$$

首先我们用数学归纳法证明如下求逆公式：定义 $nm\times nm$ 矩阵 $\boldsymbol{J}_n$ 为

$$
\boldsymbol{J}_n=\left[\begin{array}{c|ccccc}
\boldsymbol{I}_m & -q^{-1}\boldsymbol{I}_m & \boldsymbol{0} & \cdots & \cdots & \boldsymbol{0} \\ \hline
\boldsymbol{0} & \boldsymbol{I}_m & -q^{-1}\boldsymbol{I}_m & & \boldsymbol{0} & \\
\vdots & & \ddots & \ddots & & \\
\vdots & \boldsymbol{0} & & \ddots & & -q^{-1}\boldsymbol{I}_m \\
\boldsymbol{0} & & & & & \boldsymbol{I}_m
\end{array}\right] \tag{1.6.113}
$$

则有它的逆矩阵为

$$
\boldsymbol{J}_n^{-1}=\left[\begin{array}{c|cccc}
\boldsymbol{I}_m & q^{-1}\boldsymbol{I}_m, & \cdots & \cdots & q^{-(n-1)}\boldsymbol{I}_m \\ \hline
\boldsymbol{0} & \boldsymbol{I}_m & q^{-1}\boldsymbol{I}_m & \cdots & q^{-(n-2)}\boldsymbol{I}_m \\
\vdots & & \ddots & \ddots & \vdots \\
\vdots & \boldsymbol{0} & & \ddots & q^{-1}\boldsymbol{I}_m \\
\boldsymbol{0} & & & & \boldsymbol{I}_m
\end{array}\right] \tag{1.6.114}
$$

显然，我们有关系

$$
\boldsymbol{J}_n=\left[\begin{array}{c|cccc}
\boldsymbol{I}_m & -q^{-1}\boldsymbol{I}_m & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \hline
\boldsymbol{0} & & & & \\
\vdots & & \boldsymbol{J}_{n-1} & & \\
\boldsymbol{0} & & & &
\end{array}\right] \tag{1.6.115}
$$

置 $n=2$，我们应用公式(1.6.112)有

$$
\boldsymbol{J}_2^{-1}=\left[\begin{array}{c|c}
\boldsymbol{I}_m & -q^{-1}\boldsymbol{I}_m \\ \hline
\boldsymbol{0} & \boldsymbol{I}_m
\end{array}\right]^{-1}=\left[\begin{array}{c|c}
\boldsymbol{I}_m & q^{-1}\boldsymbol{I}_m \\ \hline
\boldsymbol{0} & \boldsymbol{I}_m
\end{array}\right] \tag{1.6.116}
$$

即(1.6.113)成立.置 $n=3$，应用公式(1.6.112)有

$$
\boldsymbol{J}_3^{-1}=\left[\begin{array}{c|cc}
\boldsymbol{I}_m & -q^{-1}\boldsymbol{I}_m & \boldsymbol{0} \\ \hline
\boldsymbol{0} & \boldsymbol{J}_2 & \\
\boldsymbol{0} & &
\end{array}\right]^{-1}=\left[\begin{array}{c|cc}
\boldsymbol{I}_m & q^{-1}\boldsymbol{I}_m & q^{-2}\boldsymbol{I}_m \\ \hline
\boldsymbol{0} & \boldsymbol{I}_m & q^{-1}\boldsymbol{I}_m \\
\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_m
\end{array}\right] \tag{1.6.117}
$$

即(1.6.113)成立.

假设(1.6.113)对自然数 n 成立，则它对 $(n+1)$ 也成立.事实上应用(1.6.112)和(1.6.114)有

$$
\boldsymbol{J}_{n+1}^{-1}=\left[\begin{array}{c|cccc}
\boldsymbol{I}_m & -q^{-1}\boldsymbol{I}_m & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \hline
\boldsymbol{0} & & & & \\
\vdots & & \boldsymbol{J}_n & & \\
\boldsymbol{0} & & & &
\end{array}\right]=\left[\begin{array}{c|c}
\boldsymbol{I}_m & [q^{-1}\boldsymbol{I}_m\boldsymbol{0}\cdots\boldsymbol{0}]\boldsymbol{J}_n^{-1} \\ \hline
\boldsymbol{0} & \\
\vdots & \boldsymbol{J}_n^{-1} \\
\boldsymbol{0} &
\end{array}\right]=
$$

$$
\left[\begin{array}{c|cccc}
\boldsymbol{I}_m & q^{-1}\boldsymbol{I}_m & \cdots & \cdots & q^{-n}\boldsymbol{I}_m \\ \hline
\boldsymbol{0} & \boldsymbol{I}_m & q^{-1}\boldsymbol{I}_m & \cdots & q^{-(n-1)}\boldsymbol{I}_m \\
\vdots & & \ddots & \ddots & \vdots \\
\vdots & \boldsymbol{0} & & \ddots & q^{-1}\boldsymbol{I}_m \\
\boldsymbol{0} & & & & \boldsymbol{I}_m
\end{array}\right] \tag{1.6.118}
$$

由数学归纳法，(1.6.114)对任意自然数 n 成立(规定 $\boldsymbol{J}_1=\boldsymbol{I}_m$，$\boldsymbol{J}_1^{-1}=\boldsymbol{I}_m$).

注意 $\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A}$ 可分块为

$$\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A}=\left[\begin{array}{c|cccc}\boldsymbol{I}_m+\boldsymbol{A}_1q^{-1} & -q^{-1}\boldsymbol{I}_m & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \hline \boldsymbol{A}_2q^{-1} & & & & \\ \vdots & & \boldsymbol{J}_{n-1} & & \\ \boldsymbol{A}_nq^{-1} & & & & \end{array}\right]\triangleq\left[\begin{array}{c|c}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \hline \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{array}\right] \tag{1.6.119}$$

其中 $\boldsymbol{J}_n$ 由(1.6.113)定义.应用分块矩阵行列式公式[10]

$$\det(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})=\det\boldsymbol{A}_{22}\det(\boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}) \tag{1.6.120}$$

可得

$$\det(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})=\det\left(\boldsymbol{I}_m+\boldsymbol{A}_1q^{-1}-[-q^{-1}\boldsymbol{I}_m\boldsymbol{0}\cdots\boldsymbol{0}]\boldsymbol{J}_{n-1}^{-1}\begin{bmatrix}\boldsymbol{A}_2q^{-1}\\ \vdots\\ \boldsymbol{A}_nq^{-1}\end{bmatrix}\right)=$$

$$\det\left(\boldsymbol{I}_m+\boldsymbol{A}_1q^{-1}+[q^{-1}\boldsymbol{I}_m,\cdots,q^{-(n-1)}\boldsymbol{I}_m]\begin{bmatrix}\boldsymbol{A}_2q^{-1}\\ \vdots\\ \boldsymbol{A}_nq^{-1}\end{bmatrix}\right)=$$

$$\det(\boldsymbol{I}_m+\boldsymbol{A}_1q^{-1}+\boldsymbol{A}_2q^{-2}+\cdots+\boldsymbol{A}_nq^{-n})=\det\boldsymbol{A}(q^{-1}) \tag{1.6.121}$$

这证明了(i).下证(ii).

容易验证关系

$$[\boldsymbol{I}_m,\boldsymbol{I}_mq^{-1},\cdots,\boldsymbol{I}_mq^{-(n-1)}](\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})=[\boldsymbol{A}(q^{-1}),\boldsymbol{0}\cdots\boldsymbol{0}]=\boldsymbol{A}(q^{-1})\boldsymbol{H} \tag{1.6.122}$$

其中 $\boldsymbol{A},\boldsymbol{H},\boldsymbol{A}(q^{-1})$ 由(1.6.108)和(1.6.109)定义.于是用 $\mathrm{adj}\boldsymbol{A}(q^{-1})$ 左乘上式有

$$\mathrm{adj}\boldsymbol{A}(q^{-1})[\boldsymbol{I}_m,\boldsymbol{I}_mq^{-1},\cdots,\boldsymbol{I}_mq^{-(n-1)}](\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})=$$

$$\mathrm{adj}\boldsymbol{A}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{H}=\det\boldsymbol{A}(q^{-1})\boldsymbol{I}_m\boldsymbol{H} \tag{1.6.123}$$

又用 $(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})^{-1}$ 右乘上式有

$$\mathrm{adj}\boldsymbol{A}(q^{-1})[\boldsymbol{I}_m,\boldsymbol{I}_mq^{-1},\cdots,\boldsymbol{I}_mq^{-(n-1)}]=\det\boldsymbol{A}(q^{-1})\boldsymbol{H}(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})^{-1}=$$

$$\det\boldsymbol{A}(q^{-1})\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})/\det(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A}) \tag{1.6.124}$$

但由(i)有 $\det\boldsymbol{A}(q^{-1})=\det(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A})$,于是有

$$\mathrm{adj}\boldsymbol{A}(q^{-1})[\boldsymbol{I}_m,\boldsymbol{I}_mq^{-1},\cdots,\boldsymbol{I}_mq^{-(n-1)}]=\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_{nm}-q^{-1}\boldsymbol{A}) \tag{1.6.125}$$

这证明了(ii).证毕. □

【推论 1.6.1】 对在 $m=1$ 情形的伴随形

$$\boldsymbol{A}=\begin{bmatrix}-a_1 & & & \\ \vdots & \boldsymbol{I}_{n-1} & & \\ -a_n & 0 & \cdots & 0\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0\quad \cdots\quad 0] \tag{1.6.126}$$

定义标量多项式

$$A(q^{-1})=1+a_1q^{-1}+\cdots+a_nq^{-n} \tag{1.6.127}$$

则有

(i) $$\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})=A(q^{-1}) \tag{1.6.128}$$

(ⅱ) $$\boldsymbol{H}\mathrm{adj}(I_n - q^{-1}\boldsymbol{A}) = [1, q^{-1}, \cdots, q^{-(n-1)}] \tag{1.6.129}$$

证明 注意 $A^{-1}(q^{-1}) = \mathrm{adj}A(q^{-1})/\det A(q^{-1}) = 1/A(q^{-1})$，故有 $\mathrm{adj}A(q^{-1}) = 1$，$\det A(q^{-1}) = A(q^{-1})$，于是由定理 1.6.23 引出(1.6.128)和(1.6.129). 证毕. □

参 考 文 献

1 邓自立.卡尔曼滤波与维纳滤波——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2001

2 Box G E P，Jenkins G M. Time Series Analysis，Forecasting and Control. San Francisco：Holden－Day,1970

3 何书元.应用时间序列分析.北京:北京大学出版社,2003

4 Kalman R E. A New Approach to Linear Filtering and Prediction Theory. Trans. ASME. Journal of Basic Eng.，1960,82D:35～46

5 尤昌德.线性系统理论基础.北京:电子工业出版社,1985

6 邓自立.最优滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2000

7 邓自立.自校正滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2003

8 Gevers M，Wouters W R E. An. Innovation Approach to the Discrete－Time Realization Problem. Quarterly Journal of Automatic Control，1978,19(2):90～109

9 邓自立.估计 MA 参数的多维强 Gevers－Wouters 算法及其在构造 ARMA 新息模型中的应用.控制理论与应用,2001,18(5):737～740

10 须田信吴等.自动控制中的矩阵理论.北京:科学出版社,1979

11 Lewis F L. Optimal Estimation. New York：John Wiley & Sonc，1986

第二章　最小二乘法参数估计

最小二乘法是系统辨识中参数估计的最基本方法.它是由数学家高斯于1795年在研究星体运动轨道时提出的.由于最小二乘法要求系统的先验统计知识少,算法简单,计算量小,收敛性能好,因而被广泛应用于系统辨识、自适应信号处理、自适应滤波、自校正控制等领域.本章目的是在递推最小二乘(Recursive Least Squares)法基础上,提出标量和向量ARMA模型参数估计的多种改进的快速最小二乘法算法,可实时应用.所谓快速算法是指:同极大似然法等离线算法相比,计算量小,计算速度快,且收敛速度快,可实时(在线)应用.

所提出的许多新算法是基于ARMA模型可用高阶AR模型逼近的原理.大量仿真例子说明了新算法的有效性.

最小二乘法的基本原理是选择模型参数极小化模型误差(残差)平方和.并因此而得名"最小二乘"法.所谓模型误差是指由模型计算的值与观测值之差,它体现了模型的精度.因误差有正负号问题,为数学处理方便,故以误差平方大小衡量精度.为了在总体上选择最优模型参数,故以极小化模型误差平方和为性能指标来选择模型参数.这就是最小二乘法.

2.1　递推最小二乘(RLS)法

2.1.1　非递推最小二乘法

为了说明最小二乘法参数估计原理,首先考虑一个简单的回归分析问题.

【例2.1.1】　雷达跟踪测速问题.[34]

考虑雷达跟踪直线水平匀速飞行目标(飞机),要求估计飞行目标的速度 v.设目标初始位为坐标原点,每分钟观测目标位置一次,如图2.1.1所示,共计观测5次,位置观测值如表2.1.1所示。

表2.1.1　雷达跟踪目标位置测量

时间 t/min	1	2	3	4	5
位置观测值 y/km	9.6	20.3	30.4	39.5	50.2

求飞行目标速度 v 的最优估值.

由于雷达对目标位置的观测值带有随机误差(观测噪声),因而观测方程为

$$y(t) = vt + e(t) \tag{2.1.1}$$

其中 v 是待估的飞行目标速度,$e(t)$为观测误差,vt 为在时刻 t 位置真实值,$y(t)$为对飞

行目标位置的观测值. 按最小二乘法原理，速度 v 的最优估值应使观测误差 $e(t)$ 的平方和最小，即 v 应极小化性能指标

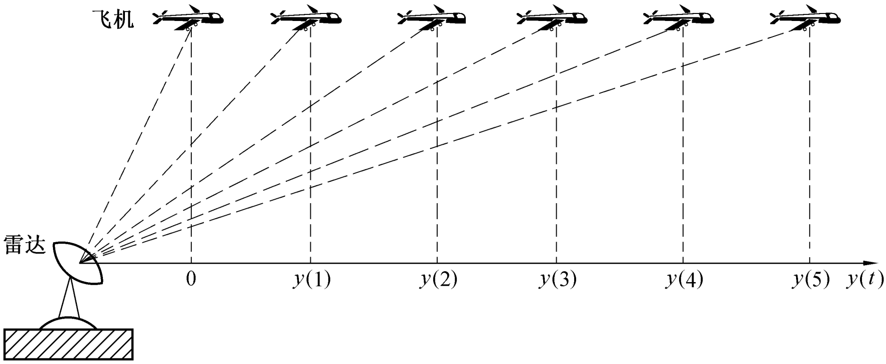


图 2.1.1　雷达跟踪测速问题

$$J = \sum_{t=1}^{5} e^2(t) = \sum_{t=1}^{5} [y(t) - vt]^2 \tag{2.1.2}$$

置$\partial J/\partial v = 0$ 有

$$-2\sum_{t=1}^{5} [y(t) - vt]t = 0 \tag{2.1.3}$$

即

$$\sum_{t=1}^{5} y(t)t = v\sum_{t=1}^{5} t^2 \tag{2.1.4}$$

这引出 v 的最小二乘估值为

$$\hat{v} = \sum_{t=1}^{5} y(t)t \bigg/ \sum_{t=1}^{5} t^2 \tag{2.1.5}$$

将观测值 $y(t)$代入上式可求得最小二乘估值 $\hat{v}$ 为

$$\hat{v} = 10.007\ 272\ \mathrm{km/min} \tag{2.1.6}$$

下面，利用上述原理来考虑 AR(n)模型的非递推最小二乘法参数估计问题. 考虑 AR(n)模型

$$y(t) = a_1 y(t-1) + \cdots + a_n y(t-n) + \varepsilon(t) \tag{2.1.7}$$

其中 $\varepsilon(t)$是零均值、方差为 $\sigma_\varepsilon^2 > 0$ 的白噪声，$y(t)$为观测，已知阶次 n，但参数 a_i 和σ_ε^2 是未知的. 问题是基于到时刻 t 的已知观测数据$(y(t), y(t-1), \cdots, y(0), \cdots, y(1-n))$，求 a_i 和σ_ε^2 的估值$\hat{a}_i(t)$和 $\hat{\sigma}_\varepsilon^2(t)$.

(2.1.7)可写为向量形式

$$y(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} + \varepsilon(t) \tag{2.1.8}$$

其中 T 为转置号，且定义向量

$$\boldsymbol{\theta} = [a_1, a_2, \cdots, a_n]^{\mathrm{T}}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t) = [y(t-1), \cdots, y(t-n)] \tag{2.1.9}$$

模型残差为

$$\varepsilon(t) = y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} \tag{2.1.10}$$

最小二乘法原理是寻求未知参数向量 $\boldsymbol{\theta}$ 的估值 $\hat{\boldsymbol{\theta}}(t)$，使其极小化模型残差平方和

$$J = \sum_{i=1}^{t} \varepsilon^2(i) = \sum_{i=1}^{t} [y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(i)\boldsymbol{\theta}]^2 \tag{2.1.11}$$

除了 $y(0),\cdots,y(1-n)$之外，由观测 $y(1),\cdots,y(t)$有 t 个等式

$$y(i) = \boldsymbol{\varphi}^{\mathrm{T}}(i)\boldsymbol{\theta} + \varepsilon(i), \quad i = 1,\cdots,t \tag{2.1.12}$$

可将它们合成写成矩阵向量形式

$$\boldsymbol{Y}(t) = \boldsymbol{\Phi}(t)\boldsymbol{\theta} + \boldsymbol{\Sigma}(t) \tag{2.1.13}$$

其中定义

$$\boldsymbol{\Phi}(t) = \begin{bmatrix} \boldsymbol{\varphi}^{\mathrm{T}}(1) \\ \vdots \\ \boldsymbol{\varphi}^{\mathrm{T}}(t) \end{bmatrix}, \quad \boldsymbol{Y}(t) = \begin{bmatrix} y(1) \\ \vdots \\ y(t) \end{bmatrix}, \quad \boldsymbol{\Sigma}(t) = \begin{bmatrix} \varepsilon(1) \\ \vdots \\ \varepsilon(t) \end{bmatrix} \tag{2.1.14}$$

于是性能指标 J 可写成

$$J = [\boldsymbol{Y}(t) - \boldsymbol{\Phi}(t)\boldsymbol{\theta}]^{\mathrm{T}}[\boldsymbol{Y}(t) - \boldsymbol{\Phi}(t)\boldsymbol{\theta}] \tag{2.1.15}$$

令 $\partial J/\partial\boldsymbol{\theta} = \boldsymbol{0}$，应用矩阵微分公式有

$$-2\boldsymbol{\Phi}^{\mathrm{T}}(t)[\boldsymbol{Y}(t) - \boldsymbol{\Phi}(t)\boldsymbol{\theta}] = \boldsymbol{0} \tag{2.1.16}$$

即 $\boldsymbol{\theta}$ 满足的正规方程

$$\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}(t)\boldsymbol{\theta} = \boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{Y}(t) \tag{2.1.17}$$

于是得到基于到时刻 t 的观测 $\boldsymbol{\theta}$ 的非递推最小二乘(LS)估值为

$$\hat{\boldsymbol{\theta}}(t) = [\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}(t)]^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{Y}(t) \tag{2.1.18}$$

【定理 2.1.1】 AR(n)模型(2.1.7)的非递推最小二乘法参数估值为

$$\hat{\boldsymbol{\theta}}(t) = [\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}(t)]^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{Y}(t) \tag{2.1.19}$$

其中假设矩阵 $\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}(t)$是非异的.

【注 2.1.1】 矩阵微分公式

标量对向量求导数定义.

设 $J = J(\boldsymbol{\theta})$，$J$ 为标量，$\boldsymbol{\theta}$ 为列向量，$\boldsymbol{\theta} = (\theta_1,\theta_2,\cdots,\theta_n)^{\mathrm{T}}$，定义 J 关于 $\boldsymbol{\theta}$ 的导数 $\partial J/\partial\boldsymbol{\theta}$ 为 J 的梯度，即

$$\frac{\partial J}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \dfrac{\partial J}{\partial \theta_1} \\ \vdots \\ \dfrac{\partial J}{\partial \theta_n} \end{bmatrix} \tag{2.1.20}$$

它是一个列向量.

向量对向量求导数定义.

设 $\boldsymbol{g}(\boldsymbol{\theta})$为 $m\times 1$ 维列向量，且 $\boldsymbol{\theta}$ 为 $n\times 1$ 维列向量，即

$$\boldsymbol{g}(\boldsymbol{\theta}) = \begin{bmatrix} g_1(\boldsymbol{\theta}) \\ \vdots \\ g_m(\boldsymbol{\theta}) \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \tag{2.1.21}$$

定义 $\boldsymbol{g}(\boldsymbol{\theta})$关于 $\boldsymbol{\theta}$ 的导数 $\mathrm{d}\boldsymbol{g}(\boldsymbol{\theta})/\mathrm{d}\boldsymbol{\theta}$ 为 $m\times n$ 矩阵，它的第(i,j)元素 $\partial g_i(\boldsymbol{\theta})/\partial\theta_j$，即

$$\frac{\mathrm{d}\boldsymbol{g}(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}} = \left(\frac{\partial g_i(\boldsymbol{\theta})}{\partial \theta_j}\right)_{m\times n} \tag{2.1.22}$$

显然有

$$\frac{\mathrm{d}\boldsymbol{A\theta}}{\mathrm{d}\boldsymbol{\theta}}=\boldsymbol{A},\quad \frac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}\boldsymbol{\theta}}=\mathbf{0} \tag{2.1.23}$$

其中 $\boldsymbol{A}$ 为常阵，$\boldsymbol{c}$ 为列向量.

复合函数求得公式

$$J=\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{\theta})\boldsymbol{g}(\boldsymbol{\theta}) \tag{2.1.24}$$

$$\frac{\partial J}{\partial\boldsymbol{\theta}}=2\left[\frac{\mathrm{d}\boldsymbol{g}(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}}\right]^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{\theta}) \tag{2.1.25}$$

2.1.2 递推最小二乘(RLS)法

非递推最小二乘估值 $\hat{\boldsymbol{\theta}}(t)$(2.1.18)的缺点是要求在每时刻计算 $n\times n$ 矩阵 $\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}(t)$的逆矩阵，引起较大计算负担. 从计算机实时应用观点，要求寻求估值 $\hat{\boldsymbol{\theta}}(t)$的递推算法和避免求逆矩阵. 本节应用矩阵求逆引理导出递推最小二乘法算法.

定义 $n\times n$ 矩阵 $\boldsymbol{P}(t)$为

$$\boldsymbol{P}(t)=[\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}(t)]^{-1} \tag{2.1.26}$$

则(2.1.18)成为

$$\hat{\boldsymbol{\theta}}(t)=\boldsymbol{P}(t)\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{Y}(t) \tag{2.1.27}$$

注意

$$\hat{\boldsymbol{\theta}}(t+1)=\boldsymbol{P}(t+1)\boldsymbol{\Phi}^{\mathrm{T}}(t+1)\boldsymbol{Y}(t+1) \tag{2.1.28}$$

$$\boldsymbol{\Phi}(t+1)=\begin{bmatrix}\boldsymbol{\Phi}(t)\\ \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\end{bmatrix},\quad \boldsymbol{Y}(t+1)=\begin{bmatrix}\boldsymbol{Y}(t)\\ y(t+1)\end{bmatrix} \tag{2.1.29}$$

于是有

$$\boldsymbol{P}(t+1)=\left[[\boldsymbol{\Phi}^{\mathrm{T}}(t),\boldsymbol{\varphi}(t+1)]\begin{bmatrix}\boldsymbol{\Phi}(t)\\ \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\end{bmatrix}\right]^{-1}=$$

$$[\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}(t)+\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)]^{-1} \tag{2.1.30}$$

这引出

$$\boldsymbol{P}(t+1)=[\boldsymbol{P}^{-1}(t)+\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)]^{-1} \tag{2.1.31}$$

为了由(2.1.31)导出 $\boldsymbol{P}(t+1)$的递推公式，应用如下矩阵求逆引理.

矩阵求逆引理 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 分别为 $n\times n$, $n\times m$, $n\times m$ 矩阵，且设 $\boldsymbol{A}^{-1}$及$(\boldsymbol{I}_m+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}$存在，则有

$$(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}^{\mathrm{T}})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I}_m+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1} \tag{2.1.32}$$

其中 $\boldsymbol{I}_m$ 为 $m\times m$ 单位阵.

证明 令 $\boldsymbol{D}=\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}^{\mathrm{T}}$，则有

$$\boldsymbol{I}_n=\boldsymbol{D}^{-1}\boldsymbol{D}=\boldsymbol{D}^{-1}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}^{\mathrm{T}})=\boldsymbol{D}^{-1}\boldsymbol{A}+\boldsymbol{D}^{-1}\boldsymbol{B}\boldsymbol{C}^{\mathrm{T}} \tag{2.1.33}$$

$$\boldsymbol{D}^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{D}^{-1}\boldsymbol{B}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1} \tag{2.1.34}$$

由(2.1.34)看到，为了求 $\boldsymbol{D}^{-1}$，要求计算 $\boldsymbol{D}^{-1}\boldsymbol{B}$. 为此，用 $\boldsymbol{B}$ 右乘(2.1.34)有

$$\boldsymbol{D}^{-1}\boldsymbol{B}=\boldsymbol{A}^{-1}\boldsymbol{B}-\boldsymbol{D}^{-1}\boldsymbol{B}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B} \tag{2.1.35}$$

$$\boldsymbol{D}^{-1}\boldsymbol{B}(\boldsymbol{I}_m+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B})=\boldsymbol{A}^{-1}\boldsymbol{B} \tag{2.1.36}$$

$$\boldsymbol{D}^{-1}\boldsymbol{B}=\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I}_m+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1} \tag{2.1.37}$$

将(2.1.37)代入(2.1.34)有

$$\boldsymbol{D}^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I}_m+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1} \tag{2.1.38}$$

这证明(2.1.32). □

应用矩阵求逆公式(2.1.32)，置 $\boldsymbol{A}=\boldsymbol{P}^{-1}(t)$，$\boldsymbol{B}=\boldsymbol{\varphi}(t+1)$，$\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$，则(2.1.31)化为递推形式，

$$\boldsymbol{P}(t+1)=\boldsymbol{P}(t)-\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))^{-1}\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t) \tag{2.1.39}$$

以下推导 $\hat{\boldsymbol{\theta}}(t+1)$ 的递推公式.由(2.1.28)和(2.1.29)有

$$\begin{aligned}
\hat{\boldsymbol{\theta}}(t+1)={}&\boldsymbol{P}(t+1)[\boldsymbol{\Phi}^{\mathrm{T}}(t),\boldsymbol{\varphi}(t+1)]\begin{bmatrix}\boldsymbol{Y}(t)\\ y(t+1)\end{bmatrix}=\\
&\boldsymbol{P}(t+1)[\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{Y}(t)+\boldsymbol{\varphi}(t+1)y(t+1)]=\\
&[\boldsymbol{P}(t)-\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))^{-1}\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)]\times\\
&[\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{Y}(t)+\boldsymbol{\varphi}(t+1)y(t+1)]=\hat{\boldsymbol{\theta}}(t)-\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))^{-1}\\
&\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}(t)+\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)y(t+1)-\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))^{-1}\\
&\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)y(t+1)=\hat{\boldsymbol{\theta}}(t)-\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\\
&(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))^{-1}\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}(t)+\\
&\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))^{-1}(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))y(t+1)-\\
&\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))^{-1}\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)y(t+1)=\\
&\hat{\boldsymbol{\theta}}(t)+\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))^{-1}[y(t+1)-\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}(t)]
\end{aligned} \tag{2.1.40}$$

上述推导可概括为如下定理.

【定理 2.1.2】 AR(n)模型(2.1.7)的递推最小二乘法(RLS)参数估值为

$$\hat{\boldsymbol{\theta}}(t+1)=\hat{\boldsymbol{\theta}}(t)+\frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}[y(t+1)-\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}(t)] \tag{2.1.41}$$

$$\boldsymbol{P}(t+1)=\boldsymbol{P}(t)-\frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{2.1.42}$$

$$\hat{\boldsymbol{\theta}}(0)=\boldsymbol{\theta}_0,\quad \boldsymbol{P}(0)=\boldsymbol{P}_0,\quad y(i)=0\quad(i\leqslant 0) \tag{2.1.43}$$

由(2.1.10)定义白噪声 $\varepsilon(t)$ 的估值 $\hat{\varepsilon}(t)$ 为

$$\hat{\varepsilon}(t)=y(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1) \tag{2.1.44}$$

$\varepsilon(t)$ 的方差 σ_ε^2 的采样方差估值 $\hat{\sigma}_\varepsilon^2(t)$ 定义为

$$\hat{\sigma}_\varepsilon^2(t)=\frac{1}{t}\sum_{i=1}^{t}\hat{\varepsilon}^2(i) \tag{2.1.45}$$

【定理 2.1.3】 AR(n)模型(2.1.7)的输入白噪声 $\varepsilon(t)$ 的方差 σ_ε^2 的递推估值器 $\hat{\sigma}_\varepsilon^2(t)$ 为

$$\hat{\sigma}_\varepsilon^2(t)=\hat{\sigma}_\varepsilon^2(t-1)+\frac{1}{t}[\hat{\varepsilon}^2(t)-\hat{\sigma}_\varepsilon^2(t-1)] \tag{2.1.46}$$

带初值 $\hat{\sigma}_\varepsilon^2(1)=\hat{\varepsilon}^2(1)$.

证明 由非递推公式(2.1.45)有

$$\hat{\sigma}_{\varepsilon}^{2}(t) = \frac{1}{t}\sum_{i=1}^{t}\hat{\varepsilon}^{2}(i) = \frac{1}{t}\Big[\sum_{i=1}^{t-1}\hat{\varepsilon}^{2}(i) + \hat{\varepsilon}^{2}(t)\Big] = \frac{t-1}{t}\cdot\frac{1}{t-1}\sum_{i=1}^{t-1}\hat{\varepsilon}^{2}(i) + \frac{1}{t}\hat{\varepsilon}^{2}(t) =$$

$$\Big(1-\frac{1}{t}\Big)\hat{\sigma}_{\varepsilon}^{2}(t-1) + \frac{1}{t}\hat{\varepsilon}^{2}(t) = \hat{\sigma}_{\varepsilon}^{2}(t-1) + \frac{1}{t}\big[\hat{\varepsilon}^{2}(t) - \hat{\sigma}_{\varepsilon}^{2}(t-1)\big] \tag{2.1.47}$$

即(2.1.46)成立. □

【注 2.1.2】 注意 $1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)$是标量,因而 RLS 避免了矩阵求逆运算.在无观测数据 $y(0), y(-1), \cdots, y(1-n)$时,可规定 $y(i)=0\,(i\leqslant 0)$,因而 $\boldsymbol{\varphi}(t), t=1,2,\cdots$是可计算的.

【注 2.1.3】 初值 $\hat{\boldsymbol{\theta}}(0)$和 $\boldsymbol{P}(0)$有两种选取方法.

一种是利用少量观测数据$(y(1),\cdots,y(t_0))$,利用非递推算法(2.1.26)和(2.1.27)预先求得

$$\boldsymbol{P}(t_0) = [\boldsymbol{\Phi}^{\mathrm{T}}(t_0)\boldsymbol{\Phi}(t_0)]^{-1} \tag{2.1.48}$$

$$\hat{\boldsymbol{\theta}}(t_0) = \boldsymbol{P}(t_0)\boldsymbol{\Phi}^{\mathrm{T}}(t_0)\boldsymbol{Y}(t_0) \tag{2.1.49}$$

然后重置 $\boldsymbol{P}(0)=\boldsymbol{P}(t_0)$, $\hat{\boldsymbol{\theta}}(0)=\hat{\boldsymbol{\theta}}(t_0)$用 RLS 算法递推计算.为了保证 $\boldsymbol{P}(t_0)$存在,要求 $t_0>n$.

另一种近似选取方法是取

$$\boldsymbol{P}(0) = a\boldsymbol{I}_n,\ a \text{ 为很大实数,例如取 } a=10^5,$$

$$\hat{\boldsymbol{\theta}}(0) = \boldsymbol{0} \tag{2.1.50}$$

这可证明如下:定义

$$\boldsymbol{\Phi}(0) = \begin{bmatrix} \boldsymbol{\varphi}(0) \\ \vdots \\ \boldsymbol{\varphi}^{\mathrm{T}}(-t_0) \end{bmatrix},\quad \boldsymbol{Y}(0) = \begin{bmatrix} y(0) \\ \vdots \\ y(-t_0) \end{bmatrix},\quad t_0>n \tag{2.1.51}$$

则由(2.1.26)和(2.1.27)有初始 LS 估值

$$\hat{\boldsymbol{\theta}}(0) = \boldsymbol{P}(0)\boldsymbol{\Phi}^{\mathrm{T}}(0)\boldsymbol{Y}(0),\quad \boldsymbol{P}(0) = [\boldsymbol{\Phi}^{\mathrm{T}}(0)\boldsymbol{\Phi}(0)]^{-1} \tag{2.1.52}$$

因而基于合成观测 $\boldsymbol{Y}(0)$和 $\boldsymbol{Y}(t)$有 LS 估值

$$\underline{\hat{\boldsymbol{\theta}}}(t) = \underline{\boldsymbol{P}}(t)[\boldsymbol{\Phi}^{\mathrm{T}}(t), \boldsymbol{\Phi}^{\mathrm{T}}(0)]\begin{bmatrix} \boldsymbol{Y}(t) \\ \boldsymbol{Y}(0) \end{bmatrix} = \underline{\boldsymbol{P}}(t)[\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{Y}(t) + \boldsymbol{P}^{-1}(0)\hat{\boldsymbol{\theta}}(0)],$$

$$\underline{\boldsymbol{P}}(t) = \left[[\boldsymbol{\Phi}^{\mathrm{T}}(t), \boldsymbol{\Phi}^{\mathrm{T}}(0)]\begin{bmatrix} \boldsymbol{\Phi}(t) \\ \boldsymbol{\Phi}(0) \end{bmatrix}\right]^{-1} = [\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}^{\mathrm{T}}(t) + \boldsymbol{P}^{-1}(0)]^{-1} \tag{2.1.53}$$

当取 $\boldsymbol{P}(0)=a\boldsymbol{I}$ 时, $a>0$ 很大,则 $\boldsymbol{P}^{-1}(0)$很小,且取 $\hat{\boldsymbol{\theta}}(0)=\boldsymbol{0}$,则由(2.1.52)近似有 $\underline{\hat{\boldsymbol{\theta}}}(t)\approx\hat{\boldsymbol{\theta}}(t)$, $\underline{\boldsymbol{P}}(t)=\boldsymbol{P}(t)$,其中 $\hat{\boldsymbol{\theta}}(t)$和 $\boldsymbol{P}(t)$由(2.1.26)和(2.1.27)定义.这种初值给定方法避免了 $n\times n$ 矩阵$\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}(t)$的求逆运算,因而适于微机应用.

【注 2.1.4】 由(2.1.8)在时刻 t 处$\varepsilon(i)$的估值也可定义为

$$\hat{\varepsilon}(i) = y(i) - \boldsymbol{\varphi}^{\mathrm{T}}(i)\hat{\boldsymbol{\theta}}(t),\quad i=1,\cdots,t \tag{2.1.54}$$

且有在时刻 t 处$\varepsilon(t)$的方差 σ_{ε}^{2} 的采样估值为

$$\hat{\sigma}_{\varepsilon}^{2}(t) = \frac{1}{t}\sum_{i=1}^{t}\hat{\varepsilon}^{2}(t) \tag{2.1.55}$$

为便于实时应用,当 $t>t_f$ 时, t_f 为截断长度,可用如下 σ_{ε}^{2} 的采样估值

$$\hat{\sigma}_{\varepsilon}^{2}(t) = \frac{1}{t_f}\sum_{i=t-t_f+1}^{t}\hat{\varepsilon}^{2}(i) \tag{2.1.56}$$

此时只需用(2.1.53)计算 $\hat{\varepsilon}(i)$，$i = t, t-1, \cdots, t-t_f+1$ 即可，而不必计算全部的 $\hat{\varepsilon}(i)$，$i = t, t-1, \cdots, 1$.

若在采样估值(2.1.55)中 $\hat{\varepsilon}(i)$ 不用(2.1.44)而用下式计算.

$$\hat{\varepsilon}(i) = y(i) - \boldsymbol{\varphi}^{\mathrm{T}}(i)\hat{\boldsymbol{\theta}}(i), \quad i = 1, 2, \cdots, t \tag{2.1.57}$$

则也有递推公式[11]

$$\hat{\sigma}_{\varepsilon}^{2}(t) = \hat{\sigma}_{\varepsilon}^{2}(t-1) + \frac{1}{t}\left[\hat{\varepsilon}^{2}(t) - \hat{\sigma}_{\varepsilon}^{2}(t-1)\right] \tag{2.1.58}$$

一般来说估值 $\hat{\varepsilon}(t)$ 的公式(2.1.57)要比(2.1.44)精度高些.

【注 2.1.5】 RLS 算法计算程序框图如图 2.1.2 所示。

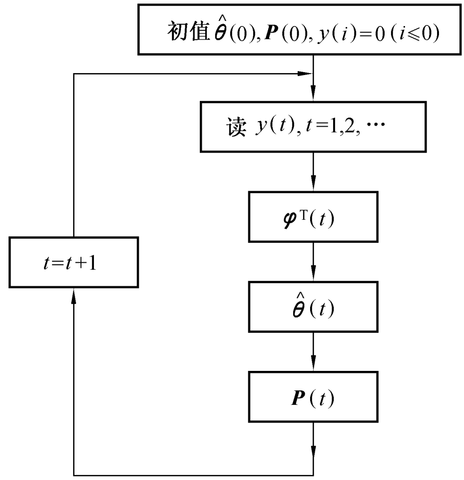


图 2.1.2　RLS 算法程序框图

2.1.3　指数加权最小二乘法——时变参数估计

当 AR(n) 模型参数随时间变化时，随观测数据增加用普通 RLS 算法参数估计误差较大. 这是因为新数据被旧数据淹没. 为了体现系统的时变性，应强调新近数据的作用，而渐渐遗忘陈旧数据的作用. 为此在性能指标(2.1.11)中对被求和的每项乘一个指数加权系数 λ^{t-i}，$0<\lambda\leqslant 1$，λ 叫遗忘因子，有改进的性能指标

$$J(\theta) = \sum_{i=1}^{t}\lambda^{t-i}\left[y(i) - \boldsymbol{\varphi}^{\mathrm{T}}(i)\boldsymbol{\theta}\right]^2 \tag{2.1.59}$$

当 λ 小于 1 时，体现了在残差平方和中强调新近数据的作用. 当 $\lambda = 1$ 时化为普通最小二乘法. 类似于定理 2.1.2 的推导有如下定理.

【定理 2.1.4】 带时变参数的 AR(n) 模型(2.1.7)的指数加权 RLS 参数估值公式为

$$\begin{aligned}
\hat{\boldsymbol{\theta}}(t+1) &= \hat{\boldsymbol{\theta}}(t) + \frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\left[\boldsymbol{y}(t+1) - \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}(t)\right]}{\lambda + \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}, \\
\boldsymbol{P}(t+1) &= \frac{1}{\lambda}\left\{\boldsymbol{P}(t) - \frac{\left[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\right]\left[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\right]^{\mathrm{T}}}{\lambda + \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}\right\}
\end{aligned} \tag{2.1.60}$$

带 $0<\lambda\leqslant 1$，$\hat{\boldsymbol{\theta}}(0) = 0$，$\boldsymbol{P}(0) = a\boldsymbol{I}_n$，$a>0$.

注意，对慢时变参数 θ，应取较大的 λ，例如 $\lambda = 0.99$，对快时变参数 θ，应取较小的 λ.

2.1.4　RLS 参数估计的收敛性

在什么条件下 RLS 参数估值收敛于相应的真实值？这是参数估计的一致性问题. 下面我们将证明：带白噪声 $\varepsilon(t)$ 的平稳 AR(n) 模型参数的非递推最小二乘法估值(2.1.18)的概率 1 和均方收敛意义下收敛于真实参数.

将(2.1.13)代入(2.1.18)有

$$\hat{\boldsymbol{\theta}}(t)=[\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{\Phi}(t)]^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(t)[\boldsymbol{\Phi}(t)\boldsymbol{\theta}+\boldsymbol{\Sigma}(t)] \tag{2.1.61}$$

置 $t=N$ 有

$$\hat{\boldsymbol{\theta}}(N)=\boldsymbol{\theta}+[\boldsymbol{\Phi}^{\mathrm{T}}(N)\boldsymbol{\Phi}(N)]^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(N)\boldsymbol{\Sigma}(N) \tag{2.1.62}$$

应用定义(2.1.14)有

$$\hat{\boldsymbol{\theta}}(N)-\boldsymbol{\theta}=\left[\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\right]^{-1}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\varepsilon(t) \tag{2.1.63}$$

将(2.1.63)变形为采样平均形式

$$\hat{\boldsymbol{\theta}}(N)-\boldsymbol{\theta}=\left[\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\right]^{-1}\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\varepsilon(t) \tag{2.1.64}$$

由 AR(n)模型平稳性假设，则 $y(t)$可表为在均方和以概率 1 收敛意义下的级数

$$y(t)=\sum_{j=0}^{\infty}\psi_j\varepsilon(t-j) \tag{2.1.65}$$

其中 ψ_j 可由(1.4.18)递推计算为

$$\psi_j=a_1\psi_{j-1}+\cdots+a_n\psi_{j-n},\quad j>0 \tag{2.1.66}$$

其中规定 $\psi_0=1$，$\psi_j=0\,(j<0)$.

由(2.1.65)引出

$$\mathrm{E}[y(t-i)\varepsilon(t)]=\sum_{j=0}^{\infty}\psi_j\mathrm{E}[\varepsilon(t-i-j)\varepsilon(t)]=0,\quad i>0 \tag{2.1.67}$$

其中 E 为均值号. 由平稳随机过程的遍历性，当 $N\to\infty$ 时有

$$\frac{1}{N}\sum_{t=1}^{N}y(t-i)\varepsilon(t)\longrightarrow\mathrm{E}[y(t-i)\varepsilon(t)]=0,\quad i>0 \tag{2.1.68}$$

$$\frac{1}{N}\sum_{t=1}^{N}y(t-i)y(t-j)\longrightarrow\mathrm{E}[y(t-i)y(t-j)]=R_y(j-i) \tag{2.1.69}$$

其中定义相关函数 $R_y(j-i)=\mathrm{E}[y(t-i)y(t-j)]$.

这引出当 $t\to\infty$ 时有

$$\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\varepsilon(t)=\frac{1}{N}\sum_{t=1}^{N}\begin{bmatrix}y(t-1)\\ \vdots\\ y(t-n)\end{bmatrix}\varepsilon(t)\to\begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix}=\mathbf{0} \tag{2.1.70}$$

$$\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\longrightarrow\mathrm{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)]=\boldsymbol{R} \tag{2.1.71}$$

且 $\boldsymbol{R}$ 为正定阵. 事实上由协方差阵定义，$R_y(j-i)$是 $\boldsymbol{R}$ 的第 j 行第 i 列元素，对任意实数 c_i，$i=1,\cdots,n$，注意

$$\mathrm{E}\left[\sum_{i=1}^{n}c_iy(t-i)\right]^2=\sum_{i,j=1}^{n}c_ic_jR_y(j-i)\geqslant 0 \tag{2.1.72}$$

带 $R_y(i)=R_y(-i)$，表明 $\boldsymbol{R}$ 是非负定的. 为了证明 $\boldsymbol{R}$ 的正定性，只要证明

$$\mathrm{E}\left[\sum_{i=1}^{n}c_iy(t-i)\right]^2=0 \tag{2.1.73}$$

的充要条件是 $c_1=c_2=\cdots=c_n=0$. 充分性是显然的. 现证必要性，假设(2.1.73)成立，应用(2.1.65)有关系

$$\mathrm{E}\left[\sum_{i=1}^{n}c_iy(t-i)\right]^2=\mathrm{E}[c_1\varepsilon(t-1)+l(\varepsilon(t-2),\cdots,\varepsilon(t-n),\cdots)]^2\geqslant c_1^2\sigma_\varepsilon^2\geqslant 0 \tag{2.1.74}$$

其中 $l(\varepsilon(t-2),\cdots,\varepsilon(t-n),\cdots)$ 为 $\varepsilon(t-2),\cdots,\varepsilon(t-n),\cdots$ 的线性组合项. 由(2.1.73)和(2.1.74)引出 $c_1=0$. 重复上述步骤可推出 $c_2=\cdots=c_n=0$. 这证明了必要性,因而 $\boldsymbol{R}$ 正定,故当 N 充分大在(2.1.64)中的逆矩阵存在. 由(2.1.64)和(2.1.71)及(2.1.70)引出 $(\hat{\boldsymbol{\theta}}(N)-\boldsymbol{\theta})\to\boldsymbol{0}(N\to\infty)$. 即在以概率 1 和均方意义下有 $\hat{\boldsymbol{\theta}}(N)\to\boldsymbol{\theta},N\to\infty$.

【定理 2.1.5】 对于带白噪声 $\varepsilon(t)$ 的平稳 AR(n)模型(2.1.7),带初值(2.1.49)的递推最小二乘法参数估值(2.1.41)和(2.1.42)在以概率 1 和均方意义下收敛于真实参数.

证明 在初值(2.1.49)下,RLS 参数估值(2.1.41)在数值上恒同于相应的非递推 LS 参数估值(2.1.19). 我们在上面证明了非递推 LS 参数估值(2.1.19)在以概率 1 和均方意义下收敛于真实值,故 RLS 参数估值也具有这种性质. □

【注 2.1.6】 RLS 算法和带遗忘因子的 RLS 算法也可推广到估计如下 CAR 模型的常的或时变参数:

$$y(t)=a_1y(t-1)+\cdots+a_ny(t-n)+b_0u(t)+\cdots+b_n(t-n)+\varepsilon(t) \tag{2.1.75}$$

其中 $u(t)$ 为已知的控制输入. 在这种情形下上式有 LS 结构

$$y(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}+\varepsilon(t) \tag{2.1.76}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t)=(y(t-1),\cdots,y(t-n),u(t),\cdots,u(t-n)),$$
$$\boldsymbol{\theta}^{\mathrm{T}}=(a_1,\cdots,a_n,b_0,\cdots,b_n) \tag{2.1.77}$$

RLS 算法(2.1.41)和(2.1.42)或(2.1.60)形式不变.

【例 2.1.2】 考虑 AR(2)模型

$$(1-1.3q^{-1}+0.4q^{-2})y(t)=\varepsilon(t) \tag{2.1.78}$$

其中 $\varepsilon(t)$ 是均值、方差为 $\sigma_\varepsilon^2=1.21$ 的高斯白噪声. 试由输出 $y(t)$ 用 RLS 算法估计 AR 参数 $a_1=-1.3,a_2=0.4$.

RLS 参数估计收敛曲线如图 2.13 和图 2.1.4 所示,其中直线代表真实值 a_1 或 a_2,曲线代表估值 $\hat{a}_1(t)$ 或 $\hat{a}_2(t)$. 可看到 RLS 算法有较好的收敛性能.

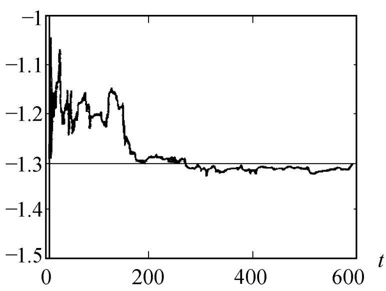


图 2.1.3 a_1 的 RLS 参数估计 $\hat{a}_1(t)$ 的收敛性

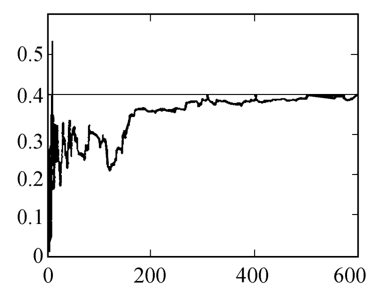


图 2.1.4 a_2 的 RLS 参数估计 $\hat{a}_2(t)$ 的收敛性

【例 2.1.3】 考虑带时变参数的 CAR 模型

$$y(t)=a(t)y(t-1)+b(t)u(t-1)+\varepsilon(t) \tag{2.1.79}$$

其中 $\varepsilon(t)$ 是零均值、方差为 $\sigma_\varepsilon^2=0.81$ 的高斯白噪声,控制 $u(t)$ 为在[-1,1]上服从均匀分布的随机变量的采样值,时变参数 $a(t)$ 和 $b(t)$ 按如下方式发生阶跃变化:

$$a(t)=\begin{cases}-0.9 & 1\leqslant t\leqslant 1\,000\\ 0.9 & 1\,000<t\leqslant 2\,000\end{cases}\qquad b(t)=\begin{cases}3 & 1\leqslant t\leqslant 1\,000\\ 1 & 1\,000<t\leqslant 2\,000\end{cases} \tag{2.1.80}$$

试由输出 $y(t)$ 用带遗忘因子的 RLS 算法估计时变参数 $a(t)$ 和 $b(t)$.

取遗忘因子 $\lambda = 0.99$，用 RLS 算法(2.1.60)仿真结果如图 2.1.5 所示. 可看到经过一段较短的过渡过程，RLS 参数估值 $\hat{a}(t)$和 $\hat{b}(t)$能较快地跟踪 $a(t)$和 $b(t)$的变化.

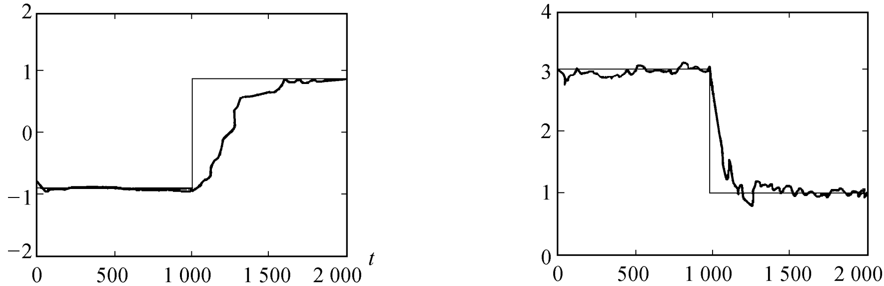


图 2.1.5　带遗忘因子的 RLS 算法参数估值对阶跃变化参数的跟踪性

【例 2.1.4】　考虑带时变参数的 CAR 模型

$$y(t) = a(t)y(t-1) + b(t)u(t-1) + \varepsilon(t) \tag{2.1.81}$$

其中 $\varepsilon(t)$是零均值、方差为 $\sigma_\varepsilon^2 = 1$ 的高斯白噪声，$u(t)$为如下方波

$$u(t) = \begin{cases} 1 & 0 \leqslant t < 100 \\ -1 & 100 \leqslant t < 200 \\ 1 & 200 \leqslant t < 300 \\ -1 & 300 \leqslant t < 400 \end{cases} \tag{2.1.82}$$

时变参数 $a(t)$和 $b(t)$按如下规律变化

$$\begin{aligned} a(t) &= 0.5 - 0.15\sin(2\pi t/2\,000), \\ b(t) &= 0.81 - 0.2\cos(2\pi t/2\,000) \end{aligned} \tag{2.1.83}$$

试用带遗忘因子 RLS 算法估计 $a(t)$和 $b(t)$.

取遗忘因子 $\lambda = 0.99$，应用(2.1.60)的 RLS 参数估计如图 2.1.6 所示，其中光滑曲线为 $a(t)$或 $b(t)$，带高频扰动的曲线为 $\hat{a}(t)$或 $\hat{b}(t)$. 可看到 RLS 估值能跟踪其实值的变化.

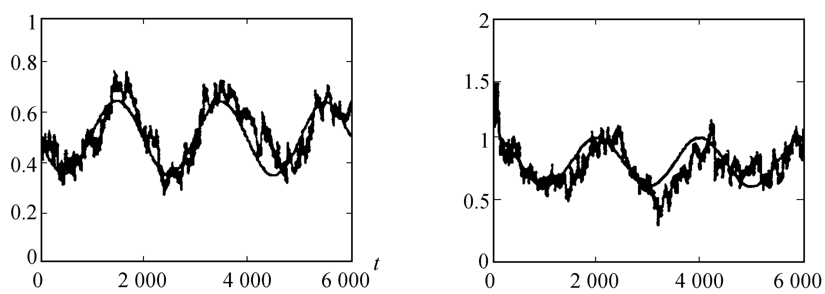


图 2.1.6　带遗忘因子的 RLS 算法参数估值对缓慢变化参数的跟踪性

2.2　递推增广最小二乘(RELS)法

考虑平稳可逆的 ARMA 模型

$$A(q^{-1})y(t) = D(q^{-1})\varepsilon(t) \tag{2.2.1}$$

其中 $\varepsilon(t)$是零均值、方差为 σ_ε^2 的白噪声，

$$A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a},$$

$$D(q^{-1}) = 1 + d_1 q^{-1} + \cdots + a_{n_d} q^{-n_d} \tag{2.2.2}$$

设阶次 n_a, n_d 已知,但参数 a_i, d_i 及 σ_ε^2 未知.为了应用 RLS 算法估计这些参数,将 $\varepsilon(t-i)$ 用其估值 $\hat{\varepsilon}(t-i)$ 近似代替,由(2.2.1)有 LS 结构

$$y(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} + \varepsilon(t) \tag{2.2.3}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t) = [-y(t-1), \cdots, y(t-n_a), \hat{\varepsilon}(t-1), \cdots, \hat{\varepsilon}(t-n_d)] \tag{2.2.4}$$

$$\boldsymbol{\theta} = [a_1, \cdots, a_{n_a}, d_1, \cdots, d_{n_d}]^{\mathrm{T}} \tag{2.2.5}$$

则由 RLS 算法引出递推增广最小二乘(Recursive Extended Least Squares)算法,简记 RELS 算法,即

$$\hat{\boldsymbol{\theta}}(t+1) = \hat{\boldsymbol{\theta}}(t) + \frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)[y(t+1) - \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}(t)]}{1 + \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{2.2.6}$$

$$\boldsymbol{P}(t+1) = \boldsymbol{P}(t) - \frac{[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)][\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)]^{\mathrm{T}}}{1 + \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{2.2.7}$$

其中 $\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$ 中的 $\hat{\varepsilon}(j), j = t, t-1, \cdots, t-1-n_d$,由下式计算

$$\hat{\varepsilon}(j) = y(j) - \boldsymbol{\varphi}^{\mathrm{T}}(j)\hat{\boldsymbol{\theta}}(j-1), \quad j = t, \cdots, t-n_d+1 \tag{2.2.8}$$

且带初值 $\hat{\boldsymbol{\theta}}(0) = 0, \boldsymbol{P}(0) = a\boldsymbol{I}$,$a$ 为很大正数,且规定

$$\hat{\varepsilon}(i) = 0 \ (i \leqslant 0), \quad y(i) = 0 \quad (i \leqslant 0) \tag{2.2.9}$$

可证明[10]:在 $D(q^{-1})$ 满足正实性条件下,RELS 算法参数估计是一致的,即 $\hat{\boldsymbol{\theta}}(t) \to \boldsymbol{\theta}, t \to \infty$.

在时刻 t 处 σ_ε^2 的估值 $\hat{\sigma}_\varepsilon^2(t)$ 为

$$\hat{\sigma}_\varepsilon^2(t) = \frac{1}{t}\sum_{i=1}^{t}\hat{\varepsilon}^2(j) \tag{2.2.10}$$

由(2.2.10)可引出递推公式

$$\hat{\sigma}_\varepsilon^2(t) = \hat{\sigma}_\varepsilon^2(t-1) + \frac{1}{t}[\hat{\varepsilon}^2(t) - \hat{\sigma}_\varepsilon^2(t-1)] \tag{2.2.11}$$

【注 2.2.1】 在上述 RELS 算法中,$\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$ 中的估值 $\hat{\varepsilon}(t), \hat{\varepsilon}(t-1), \cdots, \hat{\varepsilon}(t+1-n_d)$ 是用(2.2.8)计算的,有些文献[4]用(2.1.57)计算它们.

【注 2.2.2】 循环-递推增广最小二乘法.当数据长度 N 较短时,若基于 N 个观测数据所得 RELS 参数估值常常不能令人满意的,特别 MA 参数估值收敛较慢.为了改进 RELS 参数估值精度,可循环应用已知的 N 个数据.取第一个循环最后时刻估值 $\hat{\boldsymbol{\theta}}(N)$ 及 $\boldsymbol{P}(N)$ 作为第二个循环的初值 $\hat{\boldsymbol{\theta}}(0)$ 和 $\boldsymbol{P}(0)$,即取 $\hat{\boldsymbol{\theta}}(0) = \hat{\boldsymbol{\theta}}(N), \boldsymbol{P}(0) = \boldsymbol{P}(N)$.用这种循环-RELS 算法可充分利用已知信息改进估值精度.

【注 2.2.3】 RELS 算法的缺点是 ARMA 模型的 MA 参数估值收敛较慢.这是因为白噪声估值 $\hat{\varepsilon}(t)$ 与参数估值 $\hat{\theta}(t)$ 通过(2.2.6)和(2.2.7)与(2.2.8)有相互耦合作用.因此改进 RELS 算法的关键在于改进白噪声估值 $\hat{\varepsilon}(t)$ 的精度.

【注 2.2.4】 RELS 算法(2.2.6)~(2.2.8)也适用于 CARMA 模型

$$A(q^{-1})y(t) = B(q^{-1})u(t) + D(q^{-1})\varepsilon(t) \tag{2.2.12}$$

其中 $B(q^{-1}) = b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}$,$u(t)$ 是已知控制输入.此时在 LS 结构(2.2.3)中应

定义

$$\boldsymbol{\varphi}^{\mathrm{T}}(t) = [-y(t-1),\cdots,-y(t-n_a),u(t-1),\cdots,u(t-n_b),\hat{\varepsilon}(t-1),\cdots,\hat{\varepsilon}(t-n_d)] \tag{2.2.13}$$

$$\boldsymbol{\theta} = [a_1,\cdots a_{n_a},b_1,\cdots,b_{n_b},d_1,\cdots,d_{n_d}]^{\mathrm{T}} \tag{2.2.14}$$

【例 2.2.1】 考虑 ARMA(2.2)模型

$$(1+1.3q^{-1}+0.4q^{-2})y(t) = (1-1.1q^{-1}+0.3q^{-2})\varepsilon(t) \tag{2.2.15}$$

其中 $\varepsilon(t)$是零均值、方差为 $\sigma_\varepsilon^2=0.81$ 的高斯(正态)白噪声,试利用输出 $y(t)$求模型参数和噪声方差的 RELS 估值.注意输出 $y(t)$的采样值可由(2.2.15)通过高斯白噪声 $\varepsilon(t)$的采样值(可用计算机由正态 $N(0,1)$随机数生成)递推计算.

RELS 参数估计的收敛性如图 2.2.1 所示,其中直线代表参数的真实值,曲线代表参数估值.可看到 RELS 算法有良好的收敛性能.在 $t=600$ 步时,相应的参数估值为

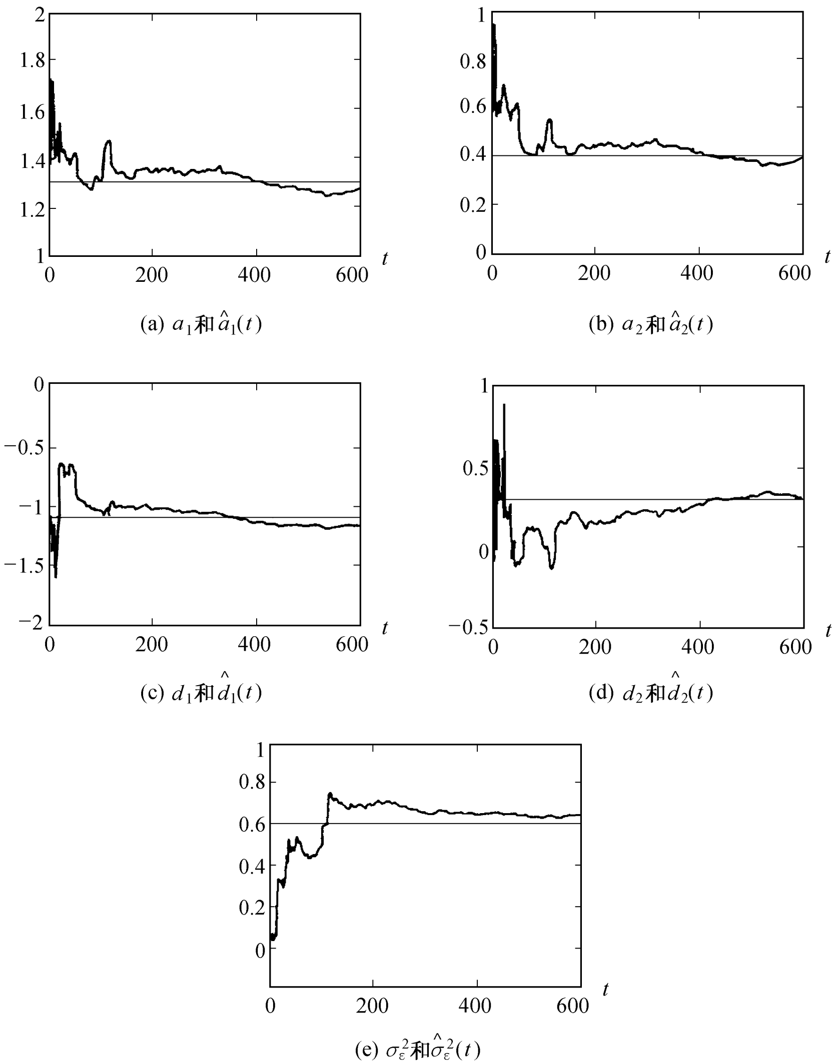


图 2.2.1　RELS 参数估计的收敛性

$$\hat{a}_1 = 1.278\,717, \quad \hat{a}_2 = 0.385\,523, \quad \hat{d}_1 = -1.158\,287, \quad \hat{d}_2 = 0.317\,733, \quad \hat{\sigma}_\varepsilon^2 = 0.853\,890, \tag{2.2.16}$$

2.3 ARMA 模型参数估计的两段 RLS－RELS 算法——改进的 RELS 算法

ARMA 模型广泛出现在系统辨识、信号处理、状态估计、时间序列分析等领域.特别在自校正滤波理论的应用中，ARMA 模型参数估计问题是一项关键技术.用常规 RELS 算法虽然简单，但其中采用了与 ARMA 参数估值耦合的白噪声估值，因而影响了算法的精度和收敛程度.通常 MA 参数收敛较慢，这是由于白噪声估值精度低引起的.因此本节作者[12]提出对 ARMA 模型用 RLS 算法拟合高阶 AR 模型单独得到白噪声估值，并将它们用于常规 RELS 算法中，得到一种改进的 RELS 算法，它由两段 RLS 算法组成，可实时实现，克服了常规 RELS 算法的上述缺点，便于工程应用.虽然作者在文献[1]中也提出了改进的递推增广最小二乘法，但它的算法是离线的，不能在线应用.

2.3.1 用 RLS 算法拟合高阶 AR 模型产生白噪声估值

考虑平稳可逆的 ARMA 模型(2.1.1)，假设、问题及记号同节 2.2.由平稳、可逆性假设有 $A(q^{-1})$和 $D(q^{-1})$是稳定的，则 ARMA 模型(2.1.1)等价于无穷阶 AR 模型

$$\varepsilon(t) = \frac{A(q^{-1})}{D(q^{-1})} y(t) = \sum_{j=0}^{\infty} \beta_j y(t-j) \tag{2.3.1}$$

其中 $\beta_0 = 1, \beta_j \to 0, j \to \infty$，$\beta_j$ 可用递推公式(1.1.45)计算.取 n_0 充分大，例如 $n_0 = 10 \sim 20$，则有近似的高阶 AR(n_0)模型

$$\sum_{j=0}^{n_0} \beta_j y(t-j) = \varepsilon(t) \tag{2.3.2}$$

它有 LS 结构

$$y(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\beta} + \varepsilon(t) \tag{2.3.3}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t) = [-y(t-1), \cdots, -y(t-n_0)] \tag{2.3.4}$$

$$\boldsymbol{\beta} = [\beta_1, \beta_2, \cdots, \beta_{n_0}]^{\mathrm{T}} \tag{2.3.5}$$

于是由(2.1.41)和(2.1.42)可得 $\boldsymbol{\beta}$ 的 RLS 估值 $\hat{\beta}(t+1)$为

$$\hat{\boldsymbol{\beta}}(t+1) = \hat{\boldsymbol{\beta}}(t) + r(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)[y(t+1) - \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\beta}}(t)] \tag{2.3.6}$$

$$\boldsymbol{P}(t+1) = \boldsymbol{P}(t) - r(t+1)[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)][\boldsymbol{P}(t+1)\boldsymbol{\varphi}(t)]^{\mathrm{T}} \tag{2.3.7}$$

$$r(t+1) = 1/[1 + \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t+1)\boldsymbol{\varphi}(t+1)] \tag{2.3.8}$$

带初值 $\hat{\boldsymbol{\beta}}(0) = 0, \boldsymbol{P}(0) = a\boldsymbol{I}, a = 10^5$，且规定 $y(i) = 0\ (i \leqslant 0)$.于是由(2.3.3)可得在时刻 $t+1$白噪声 $\varepsilon(j)$的平滑估值

$$\hat{\varepsilon}(j) = y(j) - \boldsymbol{\varphi}^{\mathrm{T}}(j)\hat{\boldsymbol{\beta}}(t+1), \quad j = t, t-1, \cdots, t+1-n_d \tag{2.3.9}$$

且方差 σ_ε^2 的估值$\hat{\sigma}_\varepsilon^2(t)$可用(2.1.46)计算.注意，基于到时刻 $t+1$ 的观测信息估计在过去时刻 $j\ (j \leqslant t)\ \varepsilon(j)$的估值 $\hat{\varepsilon}(j)$，称其为平滑估值.注意，为简单计，在(2.1.46)中，$\hat{\varepsilon}(t)$可

计算为

$$\hat{\varepsilon}(t)=y(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\beta}}(t),\quad t=1,2,\cdots \tag{2.3.10}$$

2.3.2 改进的递推增广最小二乘(RELS)算法

在 ARMA 模型的 LS 结构(2.2.3)~(2.2.5)中,若其中 $\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$中的白噪声平滑估值 $\hat{\varepsilon}(j)$, $j=t,t-1,\cdots,t-n_d+1$,用(2.3.9)计算,则由 RELS 算法(2.2.6)和(2.2.7)有改进的 RELS 算法

$$\hat{\boldsymbol{\theta}}(t+1)=\hat{\boldsymbol{\theta}}(t)+r(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)[y(t+1)-\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}(t)] \tag{2.3.11}$$

$$\boldsymbol{P}(t+1)=\boldsymbol{P}(t)-r(t+1)[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)][\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)]^{\mathrm{T}} \tag{2.3.12}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t+1)=[y(t),\cdots,y(t+1-n_d),\hat{\varepsilon}(t),\cdots,\hat{\varepsilon}(t+1-n_d)] \tag{2.3.13}$$

$$\hat{\varepsilon}(j)=y(j)-\boldsymbol{\varphi}^{\mathrm{T}}(j)\hat{\boldsymbol{\beta}}(t+1),\quad j=t,t-1,\cdots,t+1-n_d \tag{2.3.14}$$

其中 $\hat{\boldsymbol{\beta}}(t+1)$由(2.3.6)~(2.3.8)计算,且初值为 $\hat{\boldsymbol{\theta}}(0)=0$, $\boldsymbol{P}(0)=a\boldsymbol{I}$, $a=10^5$,且规定 $y(i)=0(i\leqslant 0)$, $\hat{\varepsilon}(i)=0(i\leqslant 0)$.

因此改进的 RELS 算法由两段 RLS 算法(2.3.6)~(2.3.8)和(2.3.11)~(2.3.14)组成.在每时刻 $t+1$,先实现(2.3.6)~(2.3.8),再实现(2.3.11)~(2.3.14),且后段不影响前段结果.

【注 2.3.1】 普通 RELS 算法与这里提出的改进的 RELS 算法的根本区别在于:在普通 RELS 算法中 $\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$中的白噪声估值 $\hat{\varepsilon}(j)$是用(2.2.8)计算的,它与参数估值 $\hat{\boldsymbol{\theta}}(j-1)$有相互耦合作用,即估值 $\hat{\boldsymbol{\theta}}(j-1)$通过(2.2.8)影响值 $\hat{\varepsilon}(j)$,又通过 RELS 算法(2.2.6)和(2.2.7)影响 $\hat{\boldsymbol{\theta}}(t+1)$.但在改进的 RELS 算法中 $\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$中的白噪声估值 $\hat{\varepsilon}(j)$是用(2.3.9)计算的,它与估值 $\hat{\boldsymbol{\theta}}(t)$无关,且被单独用 RELS 算法计算.只要所拟合的 AR(n_0)的阶次 n_0 充分大,则用(2.3.9)就可给出 $\varepsilon(j)$的高精度平滑估值 $\hat{\varepsilon}(j)$,因而用改进的 RELS 算法可加快 MA 参数收敛速度和提高 ARMA 模型参数估计精度.

【注 2.3.2】 上述改进的 RELS 算法是由两个简单的 RLS 参数估值器组成,因而可在线(实时)实现.

【注 2.3.3】 已知容量为 N 的一批观测数据 $y(i)$, $i=1,\cdots,N$,离线改进的 RELS 算法为在改进的 RELS 算法(2.3.11)~(2.3.14)中,$\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$中的 $\varepsilon(j)$的估值用下式计算[1]

$$\hat{\varepsilon}(j)-y(j)-\boldsymbol{\varphi}^{\mathrm{T}}(j)\hat{\boldsymbol{\beta}}(N),\quad j=N,\cdots,1 \tag{2.3.15}$$

其中 $\hat{\boldsymbol{\beta}}(N)$是用高阶 AR$(n_0)$模型(2.3.3)拟合 $y(t)$,在最终时刻 N 处的参数向量$\boldsymbol{\beta}$ 的 RLS 估值.

【例 2.3.1】 考虑 ARMA(2,1)模型

$$(1-1.4q^{-1}+0.45q^{-2})y(t)=(1-0.3q^{-1})\varepsilon(t) \tag{2.3.16}$$

其中 $\varepsilon(t)$是零均值、方差为 $\sigma_{\varepsilon}^2=1$ 的高斯白噪声,试由输出 $y(t)$用两段 RLS-RELS(也称改进的 RELS)算法估计模型(2.3.16)的参数.

模型参数的两段 RLS-RELS 估值如图 2.3.1 所示,其中直线代表参数真实值,曲线代表参数估值.在 $t=600$ 处的参数估值为

$$\hat{a}_1=-1.395\,686,\quad \hat{a}_2=0.448\,265,\quad \hat{d}_1=-0.365\,585,\quad \hat{\varepsilon}_{\varepsilon}^2=0.895\,489 \tag{2.3.17}$$

它们近似于相应的真实值 $a_1=-1.4a$，$a_2=0.45$，$d_1=-0.3$，$\sigma_\varepsilon^2=1$.

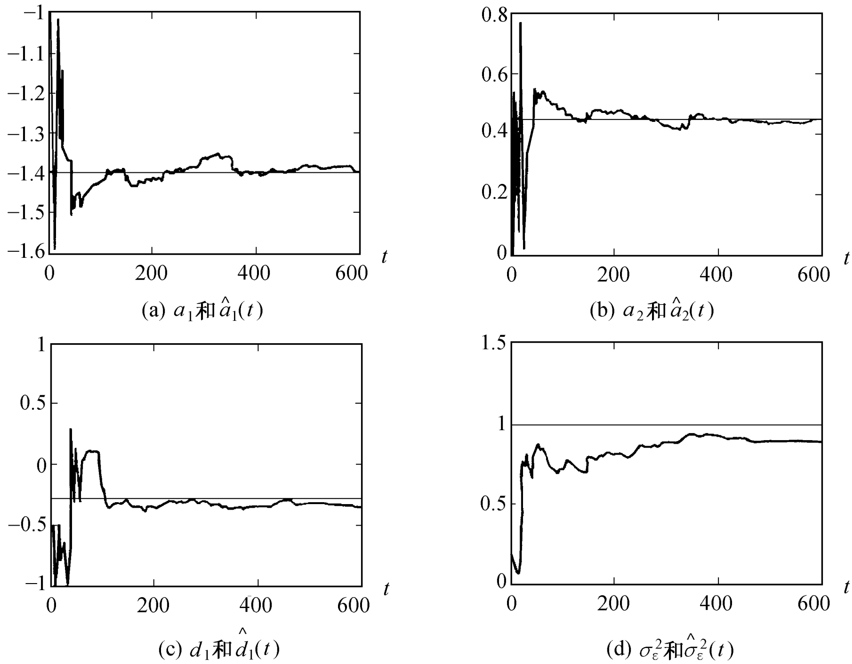


图 2.3.1　二段 RLS－RELS 参数估计的收敛性

2.3.3　CARMA 模型的两段 RLS——RELS 参数估计算法

上述 ARMA 模型的两段参数估计算法可推广到 CARMA 模型情形.考虑 CARMA 模型

$$A(q^{-1})y(t)=B(q^{-1})u(t)+D(q^{-1})\varepsilon(t) \tag{2.3.18}$$

其中 $y(t)$为观测，$u(t)$为已知控制输入，$A(q^{-1})$，$B(q^{-1})$，$D(q^{-1})$为 q^{-1}的多项式，系数 $a_0=1$，$b_0=0$，$d_0=1$，其他系数 a_i，b_i，d_i 和白噪声 $\varepsilon(t)$的方差 σ_ε^2 是未知的. 设 $A(q^{-1})$和 $D(q^{-1})$是稳定的，则由定理 1.4.4 CARMA 模型(2.3.18)可用如下高阶 CAR 模型近似代替.

$$\sum_{j=0}^{n_0}\alpha_j y(t-j)-\sum_{j=0}^{n_0}\beta_j u(t-j)=\varepsilon(t) \tag{2.3.19}$$

带 $\alpha_0=1$，$\beta_0=0$. 它有 LS 结构

$$y(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\rho}+\varepsilon(t) \tag{2.3.20}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t)=[-y(t-1),\cdots,-y(t-n_0),u(t-1),\cdots,u(t-n_0)] \tag{2.3.21}$$

$$\boldsymbol{\rho}=[\alpha_1,\cdots,\alpha_{n_0},\beta_1,\cdots,\beta_{n_0}]^{\mathrm{T}} \tag{2.3.22}$$

于是用 RLS 算法(2.1.41)和(2.1.42)可得 $\boldsymbol{\rho}$ 的 LS 估值 $\hat{\boldsymbol{\rho}}(t+1)$，从而可得 $\varepsilon(j)$的平滑估值 $\hat{\varepsilon}(j)$为

$$\hat{\varepsilon}(j)=y(j)-\boldsymbol{\varphi}^{\mathrm{T}}(j)\hat{\boldsymbol{\rho}}(t+1),\quad j=t,t-1,\cdots,t+1-n_d \tag{2.3.23}$$

且可用(2.1.58)得 σ_ε^2 的估值$\hat{\sigma}_\varepsilon^2(t)$.

另一方面(2.3.18)有 LS 结构

$$y(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}+\varepsilon(t) \tag{2.3.24}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t)=[-y(t-1),\cdots,-y(t-n_a),u(t-1),\cdots,u(t-n_b),\hat{\varepsilon}(t-1),\cdots,\hat{\varepsilon}(t-n_d)] \tag{2.3.25}$$

$$\boldsymbol{\theta}=[a_1,\cdots,a_{n_a},b_1,\cdots,b_{n_b},d_1,\cdots,d_{n_d}]^{\mathrm{T}} \tag{2.3.26}$$

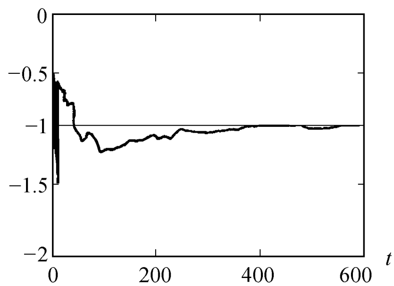
将由(2.3.23)计算的平滑估值 $\hat{\varepsilon}(t),\hat{\varepsilon}(t-1),\cdots,\hat{\varepsilon}(t+1-n_d)$代入 $\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$中，用 RELS 算法(2.2.6)和(2.2.7)得到改进的 RELS 参数估值 $\hat{\boldsymbol{\theta}}(t+1)$.

【例 2.3.2】 考虑 CARMA 模型

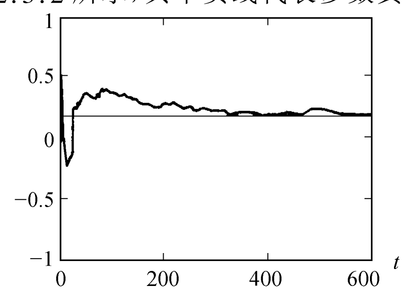
$$(1-0.9q^{-1}+0.2q^{-2})y(t)=0.5u(t-2)+(1-0.4q^{-1})\varepsilon(t) \tag{2.3.27}$$

其中 $\varepsilon(t)$是零均值、方差为 $\sigma_\varepsilon^2=1$ 的高斯(正态)白噪声，已知控制 $u(t)$是零均值、方差为 $\sigma_u^2=1$ 的高斯白噪声的采样值，试用两段 RLS－RELS 算法由输入 $u(t)$和输出 $y(t)$估计模型参数和噪声 $\varepsilon(t)$的方差.

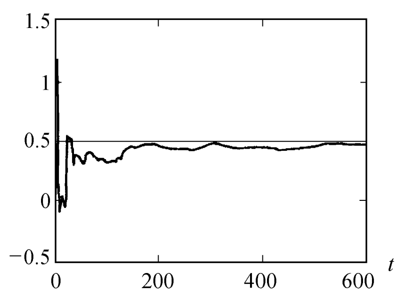
二段 RLS－RELS 参数估值的收敛性如图 2.3.2 所示，其中实线代表参数真实值，曲



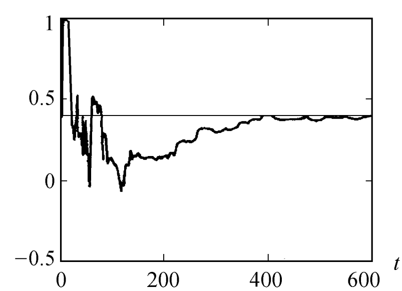
(a) a_1和$\hat{a}_1(t)$



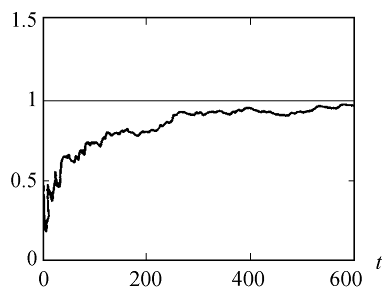
(b) a_2和$\hat{a}_2(t)$



(c) b_2和$\hat{b}_2(t)$



(d) d_1和$\hat{d}_1(t)$



(e) σ_ε^2和$\hat{\sigma}_\varepsilon^2(t)$

图 2.3.2　二段 RLS－RELS 参数估计的收敛性

线代表参数估值.在 $t=600$ 处的参数估值为

$$\begin{aligned}&\hat{a}_1=-0.917\ 169,\quad \hat{a}_2=0.210\ 361,\quad \hat{b}_2=0.465\ 976,\\&\hat{d}_1=0.377\ 60,\quad \hat{\sigma}_\varepsilon^2=0.966\ 575\end{aligned}\tag{2.3.28}$$

它们近似于相应的真实值

$$a_1=-0.9,\quad a_2=0.2,\quad b_2=0.5,\quad d_1=0.4,\quad \sigma_\varepsilon^2=1\tag{2.3.29}$$

2.4 ARMA 模型参数估计的两段 RLS – LS 算法

Graup[5,15]提出的 ARMA 模型参数估计方法是:用最小二乘法对 ARMA 模型拟合高阶 AR 模型,然后基于所拟合的 AR 模型参数,用解一个相容代数方程组得到 ARMA 模型参数估值.该方法的缺点是只能得到 ARMA 模型参数的粗糙估计.因为实际上所得到的线性代数方程组是不相容的,即超定矛盾方程组.因而文献[16]提出用最小二乘法(LS)求解超定不相容代数方程组来改进参数估计精度.这引出两段 RLS – LS 算法.当 ARMA 模型的 MA 参数个数不超过 3 时,这种两段 RLS – LS 算法还可实时实现.此时只需在线计算 2 阶或 3 阶矩阵的逆矩阵.

2.4.1 用 RLS 算法拟合高阶 AR 模型

考虑平稳、可逆的 ARMA 模型(2.2.1),其中假设和记号同节 2.2,则它等价于无穷阶 AR(∞)模型

$$\varepsilon(t)=\frac{A(q^{-1})}{D(q^{-1})}y(t)=\sum_{j=0}^{\infty}\beta_j q^{-j}y(t)\tag{2.4.1}$$

其中 $\beta_0=1$,$\beta_j\to0\ (j\to\infty)$.取 n_0 充分大,例如取 $n_0=10\sim20$,则有近似的高阶 AR(n_0)模型

$$\beta(q^{-1})y(t)=\varepsilon(t)\tag{2.4.2}$$

$$\beta(q^{-1})=1+\beta_1q^{-1}+\cdots+\beta_iq^{-n_0}\tag{2.4.3}$$

比较(2.4.1)和(2.4.2)有近似关系

$$D(q^{-1})\beta(q^{-1})=A(q^{-1})\tag{2.4.4}$$

由(2.4.2)有 LS 结构

$$y(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\beta}+\varepsilon(t)\tag{2.4.5}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t)=[-y(t-1),\cdots,-y(t-n_0)]\tag{2.4.6}$$

$$\boldsymbol{\beta}=[\beta_1,\beta_2,\cdots,\beta_{n_0}]^{\mathrm{T}}\tag{2.4.7}$$

于是用(2.3.6)~(2.3.8)可得在时刻 t 处 $\boldsymbol{\beta}$ 的 RLS 估值 $\hat{\boldsymbol{\beta}}(t)$,且可由(2.1.58)得方差 σ_ε^2 的估值 $\hat{\sigma}_\varepsilon^2(t)$.

2.4.2 用求不相容超定线性方程组的最小二乘解估计 ARMA 模型参数

比较(2.4.4)两边 q^{-i} 的系数有关系

$$\boldsymbol{a}=\underline{\boldsymbol{d}}+\boldsymbol{D}\,\underline{\boldsymbol{b}}\tag{2.4.8}$$

$$\boldsymbol{B d}=\boldsymbol{b} \tag{2.4.9}$$

其中定义 $d_i=0(i>n_d)$，$\beta_i=0(i<0$ 或 $i>n_0)$，且定义

$$\boldsymbol{a}=[a_1,a_2,\cdots,a_{n_a}]^{\mathrm{T}}$$

$$\underline{\boldsymbol{d}}=[d_1,d_2,\cdots,d_{n_a}]^{\mathrm{T}}$$

$$\underline{\boldsymbol{b}}=[\beta_1,\beta_2,\cdots,\beta_{n_a}]^{\mathrm{T}}$$

$$\boldsymbol{d}=[d_1,d_2,\cdots,d_{n_d}]^{\mathrm{T}}$$

$$\boldsymbol{b}=[-\beta_{n_{a+1}},\cdots\beta_{n_0},0\cdots0]^{\mathrm{T}}$$

$$\boldsymbol{D}=\begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ d_1 & 1 & & & \vdots \\ \vdots & \vdots & & & 0 \\ d_{n_a-1} & \cdots & \cdots & d_1 & 1 \end{bmatrix}$$

$$\boldsymbol{B}=\begin{bmatrix} \beta_{n_a} & \cdots & \cdots & \beta_{n_a-n_d+1} \\ \beta_{n_a+1} & \cdots & \cdots & \beta_{n_a-n_d+2} \\ \vdots & & & \vdots \\ \beta_{n_0} & \cdots & \cdots & \beta_{n_0-n_d+1} \\ \mathbf{0} & & & \vdots \\ 0 & \cdots & 0 & \beta_{n_0} \end{bmatrix} \tag{2.4.10}$$

注意(2.4.8)和(2.4.9)对 $n_a=n_d$，$n_a>n_d$ 和 $n_a<n_d$，三种情形均成立. 将 RLS 估值 $\hat{\boldsymbol{\beta}}(t)$ 代入(2.4.9)可得在时刻 t 的估值$\hat{\boldsymbol{B}}$，$\hat{\boldsymbol{d}}$ 和 $\hat{\boldsymbol{b}}$ 有关系

$$\hat{\boldsymbol{B}}\hat{\boldsymbol{d}}=\hat{\boldsymbol{b}} \tag{2.4.11}$$

其中 $\hat{\boldsymbol{B}}$ 为$(n_0+n_d-n_a)\times n_d$ 非方矩阵. 由于关系式(2.4.4)是近似成立的，因而等式(2.4.9)也近似成立，从而代入估值 $\hat{\boldsymbol{\beta}}(t)$后使(2.4.11)近似成立，即(2.4.11)是不相容线性方程组(也称矛盾方程组). 因 $n_0+d_d-n_a>n_d$，故它是超定的. 矛盾方程组(2.4.11)的最小二乘解 $\hat{d}$ 应极小化方程误差$(\hat{\boldsymbol{B}}\hat{\boldsymbol{d}}-\hat{\boldsymbol{b}})$的分量平方和，即 $\hat{d}$ 应极小化最小二乘性能指标

$$\min_{\hat{d}} J=(\hat{\boldsymbol{B}}\hat{\boldsymbol{d}}-\hat{\boldsymbol{b}})^{\mathrm{T}}(\hat{\boldsymbol{B}}\hat{\boldsymbol{d}}-\hat{\boldsymbol{b}}) \tag{2.4.12}$$

置$\partial J/\partial\hat{\boldsymbol{d}}=\mathbf{0}$，应用求导公式(2.1.25)有

$$\frac{\partial J}{\partial\hat{\boldsymbol{d}}}=2\hat{\boldsymbol{B}}^{\mathrm{T}}(\hat{\boldsymbol{B}}\hat{\boldsymbol{d}}-\hat{\boldsymbol{b}})=\mathbf{0} \tag{2.4.13}$$

这引出矛盾方程(2.4.11)的最小二乘解为

$$\hat{\boldsymbol{d}}=(\hat{\boldsymbol{B}}^{\mathrm{T}}\hat{\boldsymbol{B}})^{-1}\hat{\boldsymbol{B}}^{\mathrm{T}}\hat{\boldsymbol{b}} \tag{2.4.14}$$

由估值 $\hat{\boldsymbol{d}}$ 可得估值$\hat{\underline{\boldsymbol{d}}}$ 和$\hat{\boldsymbol{D}}$，由估值 $\hat{\boldsymbol{\beta}}(t)$可得估值 $\hat{\underline{\boldsymbol{b}}}$，进而由(2.4.8)可得在时刻 t 处 $\boldsymbol{a}$ 的估值$\hat{\boldsymbol{a}}$ 为

$$\hat{\boldsymbol{a}}=\hat{\underline{\boldsymbol{d}}}+\hat{\boldsymbol{D}}\hat{\underline{\boldsymbol{b}}} \tag{2.4.15}$$

注意$(\hat{\boldsymbol{B}}^{\mathrm{T}}\hat{\boldsymbol{B}})$是 $n_d\times n_d$ 方阵，通常 $n_d=1\sim3$，因而可在线计算逆矩阵$(\hat{\boldsymbol{B}}^{\mathrm{T}}\hat{\boldsymbol{B}})^{-1}$，从而

可在线计算 $\hat{\boldsymbol{d}}$ 和 $\hat{\boldsymbol{a}}$.

应指出 Graup[5]仅用线性方程组(2.4.11)的前 n_d 个方程求估值 $\hat{d}$，舍去了其余方程，损失了信息，因而只能得到 MA 参数 d 的粗糙估值，且其结果仅能处理 $n_a \geqslant n_d$ 情形.

特别，当 $A(q^{-1})=1$，即 $n_a=0$ 时，对纯量 MA 模型

$$y(t)=D(q^{-1})\varepsilon(t) \tag{2.4.16}$$

在(2.4.9)和(2.4.10)中有

$$\boldsymbol{B}=\begin{bmatrix} 1 & 0 & \cdots & 0 \\ \beta_1 & 1 & & 0 \\ \vdots & \vdots & & \vdots \\ \beta_{n_d-1} & \cdots & \beta_1 & 1 \\ \vdots & & & \vdots \\ \beta_{n_0} & \cdots & \cdots & \beta_{n_0-n_d+1} \\ 0 & \beta_{n_0} & \cdots & \beta_{n_0-n_d+2} \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & \beta_{n_0} \end{bmatrix}, \quad \boldsymbol{b}=\begin{bmatrix} -\beta_1 \\ -\beta_2 \\ \vdots \\ \vdots \\ -\beta_{n_0} \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix} \tag{2.4.17}$$

仍可用(2.4.14)得 MA 参数估值 $\hat{\boldsymbol{d}}$.[21]

特别对 MA(1)模型 $y(t)=(1+d_1q^{-1})\varepsilon(t)$有

$$\boldsymbol{B}=\begin{bmatrix} 1 \\ \beta_1 \\ \vdots \\ \beta_{n_0-1} \\ \beta_{n_0} \end{bmatrix}, \quad \boldsymbol{b}=\begin{bmatrix} -\beta_1 \\ -\beta_2 \\ \vdots \\ -\beta_{n_0} \\ 0 \end{bmatrix} \tag{2.4.18}$$

对 MA(2)模型 $y(t)=(1+d_1q^{-1}+d_2q^{-2})\varepsilon(t)$有

$$\boldsymbol{B}=\begin{bmatrix} 1 & 0 \\ \beta_1 & 1 \\ \beta_2 & \beta_1 \\ \vdots & \\ \beta_{n_0} & \beta_{n_0-1} \\ 0 & \beta_{n_0} \end{bmatrix}, \quad \boldsymbol{b}=\begin{bmatrix} -\beta_1 \\ -\beta_2 \\ \vdots \\ -\beta_{n_0} \\ 0 \\ 0 \end{bmatrix} \tag{2.4.19}$$

应用(2.4.14)和(2.4.18)对 MA(1)模型有简单的二段 RLS－LS 参数估计公式

$$\hat{d}_1(t)=-\sum_{i=0}^{n_0}\hat{\beta}_i(t)\hat{\beta}_{i+1}(t)\Big/\sum_{i=0}^{n_0}\hat{\beta}_i^2(t) \tag{2.4.20}$$

其中规定 $\hat{\beta}_0(t)=1$，$\hat{\beta}_i(t)=0\,(i>n_0)$.

【例 2.4.1】 考虑 ARMA(1,2)模型

$$(1-0.9q^{-1})y(t)=(1+0.2q^{-1}-0.15q^{-2})\varepsilon(t) \tag{2.4.21}$$

其中 $\varepsilon(t)$是零均值、方差为 $\sigma_\varepsilon^2=0.81$ 的高斯白噪声，试由输出 $y(t)$用两段 RLS－LS 算法求模型参数估值和噪声方差估值.

取 $n_0=10$，模型参数和噪声方差的两段 RLS－LS 估值的收敛性如图 2.4.1 所示，其中直线代表真实值，曲线代表估值.在 $t=720$ 处估值为

$$\hat{a}_1=-0.899\,095,\quad \hat{d}_1=0.208\,382,\quad \hat{d}_2=-0.146\,941,\quad \hat{\sigma}_\varepsilon^2=0.813\,721 \tag{2.4.22}$$

它们逼近于相应的真实值

$$a_1=-0.9,\quad d_1=0.2,\quad d_2=-0.15,\quad \sigma_\varepsilon^2=0.81 \tag{2.4.23}$$

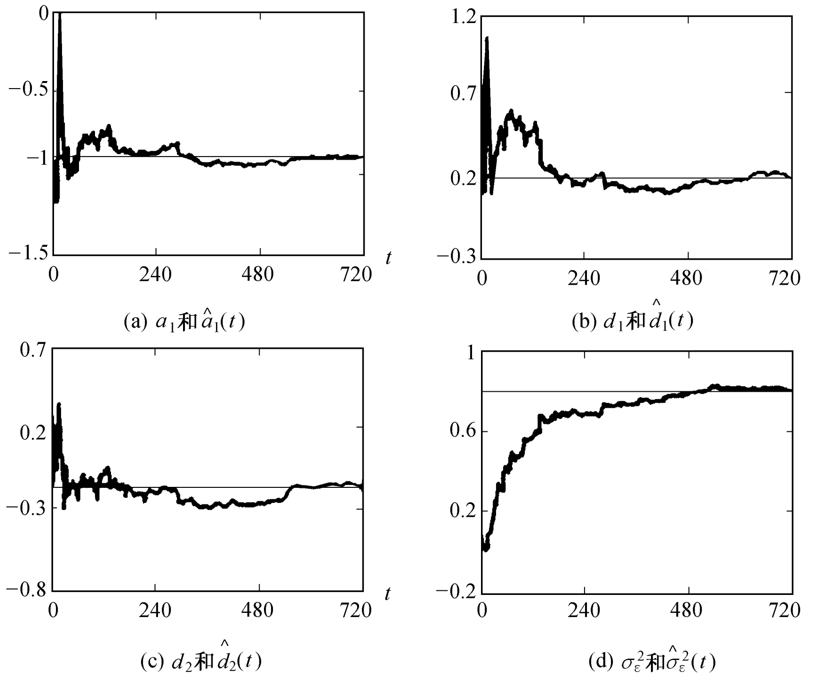


图 2.4.1　二段 RLS－LS 参数估计的收敛性

【例 2.4.2】　考虑 MA(3)模型

$$y(t)=(1-1.2q^{-1}+0.47q^{-2}-0.06q^{-3})\varepsilon(t) \tag{2.4.24}$$

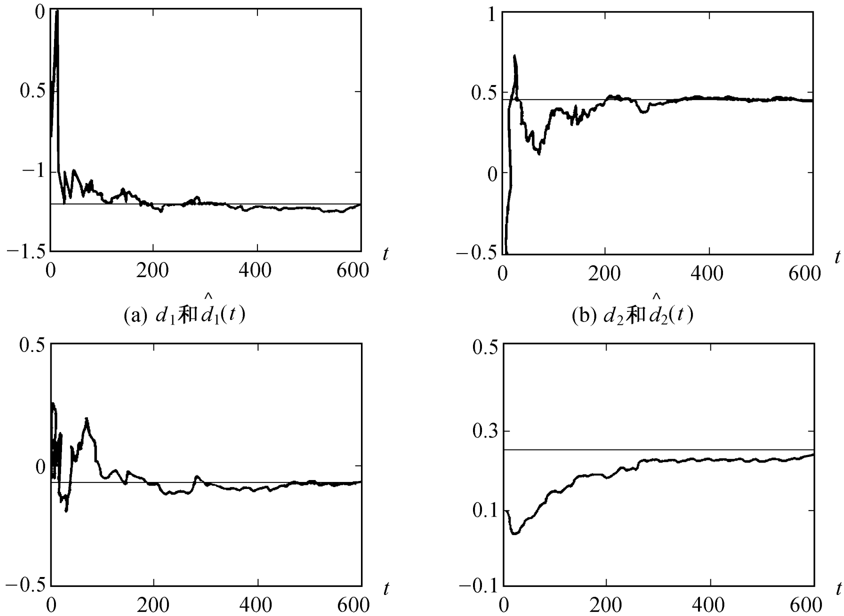
其中 $\varepsilon(t)$是零均值、方差为 $\sigma_\varepsilon^2=0.25$ 的高斯白噪声，试由输出 $y(t)$用两段 RLS－LS 算法求模型参数和方差的估值.

取 $n_0=10$，两段 RLS－LS 参数和方差估计收敛性如图 2.4.2 所示，其中直线代表真实值，曲线代表估值.在 $t=600$ 处相应的估值为

$$\hat{d}_1=-1.222\,427,\quad \hat{d}_2=0.464\,366,\quad \hat{d}_3=-0.062\,136,\quad \hat{\sigma}_\varepsilon^2=0.233\,041 \tag{2.4.25}$$

它们逼近于相应的真实值

$$d_1=-1.2,\quad d_2=0.47,\quad d_3=-0.06,\quad \sigma_\varepsilon^2=0.25 \tag{2.4.26}$$



(a) d_1和$\hat{d}_1(t)$　(b) d_2和$\hat{d}_2(t)$

(c) d_3和$\hat{d}_3(t)$　(d) σ_ε^2和$\hat{\sigma}_\varepsilon^2(t)$

图 2.4.2　二段 RLS – LS 参数估计的收敛性

2.5　CARMA 模型的三段 RLS – LS – LS 参数估计算法

考虑(2.3.18)式所示的 CARMA 模型

$$A(q^{-1})y(t) = B(q^{-1})u(t) + D(q^{-1})\varepsilon(t) \tag{2.5.1}$$

假设 $A(q^{-1})$和 $D(q^{-1})$是稳定的,但 $A(q^{-1})$,$B(q^{-1})$和 $D(q^{-1})$的参数 a_i,b_i,d_i 和白噪声 $\varepsilon(t)$的方差 σ_ε^2 是未知的.这里已知阶次 n_a,n_b,n_a,且 $a_0 = 1$,$d_0 = 1$,$b_0 = 0$.本节介绍用三段最小二乘法估计 CARMA 模型参数.第一段用 RLS 算法估计等价的高阶 CAR 模型参数;第二段用求解不相容线性方程组的 LS 算法估计 $A(q^{-1})$和 $B(q^{-1})$的参数;第三段用 LS 算法估计 $C(q^{-1})$的参数.

2.5.1　第一段——用 RLS 算法拟合高阶 CAR 模型

定义 MA 过程 $v(t) = D(q^{-1})\varepsilon(t)$,由 $D(q^{-1})$的稳定性有

$$\varepsilon(t) = D^{-1}(q^{-1})v(t) = \sum_{j=0}^{\infty} c_j q^{-j} v(t) \tag{2.5.2}$$

其中系数 c_j 由(1.4.30)递推计算,且 $c_0 = 1$,$c_j \to 0\,(j \to \infty)$.取 n_c 充分大,于是 $v(t)$可用高阶 AR(n_c)近似代替:

$$C(q^{-1})v(t) = \varepsilon(t), \quad C(q^{-1}) = 1 + c_1 q^{-1} + \cdots + c_n q^{-n_c} \tag{2.5.3}$$

由 $v(t)$的定义可将(2.5.1)化为

$$A(q^{-1})y(t) = B(q^{-1})u(t) + v(t) \tag{2.5.4}$$

由(2.5.3)和(2.5.4)CARMA 模型(2.5.1)可用如下高阶 CAR 模型近似代替

$$\alpha(q^{-1})y(t)=\beta(q^{-1})u(t)+\varepsilon(t) \tag{2.5.5}$$

$$\alpha(q^{-1})=A(q^{-1})C(q^{-1}),\quad \beta(q^{-1})=B(q^{-1})C(q^{-1}) \tag{2.5.6}$$

其中 $\alpha(q^{-1})$和 $\beta(q^{-1})$的系数分别为 α_i 和β_i，且阶次分别为 $n_\alpha=n_a+n_c$，$n_\beta=n_b+n_c$，且 $\alpha_0=1$，$\beta_0=0$.

对 CAR 模型(2.5.5)应用 RLS 算法可得在时刻 t 的参数估值$\hat{\alpha}_i(t)$，$\hat{\beta}_i(t)$及 $\hat{\sigma}_\varepsilon^2(t)$.

2.5.2 第二段——用 LS 算法估计 $A(q^{-1})$和 $B(q^{-1})$

由(2.5.6)消去 $C(q^{-1})$有关系

$$A(q^{-1})\beta(q^{-1})=B(q^{-1})\alpha(q^{-1}) \tag{2.5.7}$$

记未知参数向量 $\boldsymbol{\theta}$ 为

$$\boldsymbol{\theta}=[a_1,\cdots,a_{n_a},b_1,\cdots,b_{n_b}]^{\mathrm{T}} \tag{2.5.8}$$

比较(2.5.7)两边 q^{-i}的系数引出不相容线性方程组

$$\left[\begin{array}{ccc:ccc} 0 & & \mathbf{0} & 1 & & \mathbf{0} \\ -\beta_1 & \ddots & & \alpha_1 & \ddots & \\ -\beta_2 & \ddots & 0 & \alpha_2 & \ddots & 1 \\ \vdots & \ddots & -\beta_1 & \vdots & \ddots & \alpha_1 \\ -\beta_{n_\beta} & \ddots & -\beta_2 & \alpha_{n_\alpha} & \ddots & \alpha_2 \\ 0 & \ddots & \vdots & 0 & \ddots & \vdots \\ \mathbf{0} & & -\beta_{n_\beta} & \mathbf{0} & & \alpha_{n_\alpha} \end{array}\right]\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_a} \\ b_1 \\ b_2 \\ \vdots \\ b_{n_b} \end{bmatrix}=\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_{n_\beta} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.5.9}$$

简记(2.5.9)为

$$\boldsymbol{G\theta}=\boldsymbol{\rho} \tag{2.5.10}$$

则有最小二乘(LS)解为

$$\boldsymbol{\theta}=(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G})^{-1}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{\rho} \tag{2.5.11}$$

将估值 $\hat{\alpha}_i(t)$，$\hat{\beta}_i(t)$代入(2.5.11)可得在时刻 t 的估值

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{G}}^{\mathrm{T}}(t)\hat{\boldsymbol{G}}(t)]^{-1}\hat{\boldsymbol{G}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t) \tag{2.5.12}$$

2.5.3 第三段——用 LS 算法估计 $D(q^{-1})$

由(2.5.2)和(2.5.3)有关系

$$C(q^{-1})=D^{-1}(q^{-1}) \tag{2.5.13}$$

将其代入(2.5.6)有关系

$$\alpha(q^{-1})D(q^{-1})=A(q^{-1}),\quad \beta(q^{-1})D(q^{-1})=B(q^{-1}) \tag{2.5.14}$$

比较上式每个等式两边 q^{-1}系数列出不相容线性方程组

$$\boldsymbol{HD}=\boldsymbol{\delta} \tag{2.5.15}$$

其中定义

$$\boldsymbol{d}=[d_1,d_2,\cdots,d_{n_d}]^{\mathrm{T}},$$

$$\boldsymbol{\delta}=[a_1-\alpha_2,\cdots a_{n_a}-\alpha_{n_a},-\alpha_{n_a+1},\cdots,-\alpha_{n_a},0,\cdots,0,b_1-\beta_1,\cdots,$$

$$b_{n_b}-\beta_{n_b},-\beta_{n_b+1},\cdots,-\beta_{n_\beta},0,\cdots,0]^{\mathrm{T}},$$

$$\boldsymbol{H}=\begin{bmatrix} 1 & & \mathbf{0} \\ a_1 & \ddots & 0 \\ \vdots & \ddots & 1 \\ a_{n_a} & \ddots & a_1 \\ 0 & \ddots & \vdots \\ 0 & & a_{n_a} \\ 0 & \cdots & 0 \\ \beta_1 & & 0 \\ \vdots & \ddots & 0 \\ \beta_{n_\beta} & \ddots & \beta_1 \\ 0 & \ddots & \vdots \\ \mathbf{0} & & \beta_{n_\beta} \end{bmatrix} \tag{2.5.16}$$

于是 $\boldsymbol{d}$ 的最小二乘解为

$$\boldsymbol{d}=(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{\delta} \tag{2.5.17}$$

将估值 $\hat{a}_i(t)$和 $\hat{\beta}_i(t)$代入上式有估值

$$\hat{\boldsymbol{d}}(t)=[\hat{\boldsymbol{H}}^{\mathrm{T}}(t)\hat{\boldsymbol{H}}(t)]^{-1}\hat{\boldsymbol{H}}^{\mathrm{T}}(t)\hat{\boldsymbol{\delta}}(t) \tag{2.5.18}$$

【例 2.5.1】 考虑 CARMA 模型

$$\begin{aligned}(1-0.9q^{-1})y(t)=&(2q^{-1}+0.3q^{-2}-0.15q^{-3})u(t)+\\&(1-0.4q^{-1}-0.32q^{-2})\varepsilon(t)\end{aligned} \tag{2.5.19}$$

其中 $\varepsilon(t)$是零均值、方差为 $\sigma_\varepsilon^2=0.16$ 的高斯白噪声，$u(t)$为已知的在$[-1,1]$上服从均匀分布随机变量的采样值. 试用三段 RLS－LS－LS 算法估计该模型参数.

取 $n_0=20$，三段 RLS－LS－LS 参数估计收敛性如图 2.5.1 所示，其中直线代表真实值，曲线代表估值. 在 $t=1\ 500$ 处的参数估值为

$$\begin{aligned}&\hat{a}_1=-0.902\ 131,\quad \hat{b}_1=1.995\ 814,\quad \hat{b}_2=0.292\ 148,\\&\hat{b}_3=-0.153\ 12,\quad \hat{d}_1=-0.367\ 08,\quad \hat{d}_2=-0.317\ 095\end{aligned} \tag{2.5.20}$$

它们近似等于相应的真实值

$$a_1=-0.9,\quad b_1=2,\quad b_2=0.3,\quad b_3=-0.15,\quad d_1=-0.4,\quad d_2=-0.32 \tag{2.5.21}$$

2.6 向量 CAR 模型的多重 RLS 参数估计算法

考虑向量 CAR 模型

$$\boldsymbol{y}(t)=\sum_{j=1}^{n_a}\boldsymbol{A}_j\boldsymbol{y}(t-j)+\sum_{j=1}^{n_b}\boldsymbol{B}_j\boldsymbol{u}(t-j)+\boldsymbol{\varepsilon}(t) \tag{2.6.1}$$

其中 $\boldsymbol{y}(t)=[y_1(t),\cdots,y_m(t)]^{\mathrm{T}}$ 是 $m\times1$ 输出(观测)，$\boldsymbol{u}(t)=[u_1(t),\cdots,u_p(t)]^{\mathrm{T}}$ 是 $p\times1$ 控制输入，$\boldsymbol{\varepsilon}(t)=[\varepsilon_1(t),\cdots,\varepsilon_m(t)]^{\mathrm{T}}$ 是 $m\times1$ 带零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，T 为转置号. 问题是求未知参数阵 $\boldsymbol{A}_j$，$\boldsymbol{B}_j$ 和 $\boldsymbol{Q}_\varepsilon$ 的 RLS 估值. 记 $\boldsymbol{A}_j=(a_{is}^j)$，$\boldsymbol{B}_j=(b_{is}^j)$，其中 a_{is}^j和

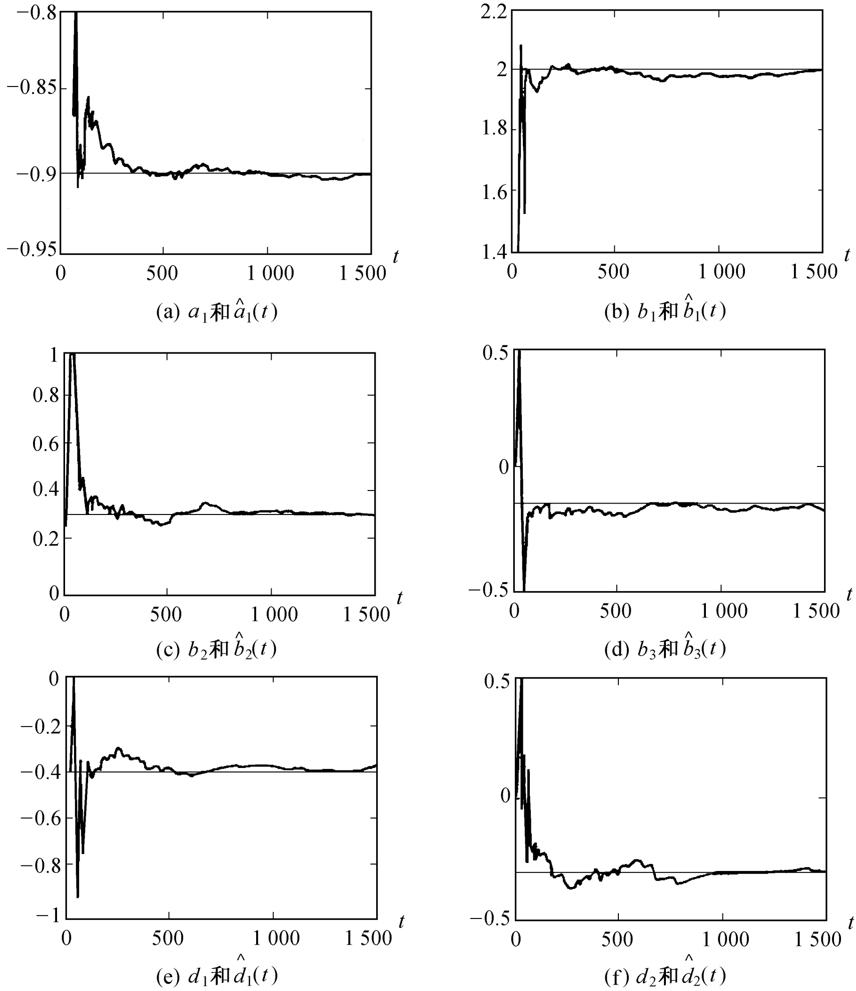


图 2.5.1　三段 RLS – LS – LS 参数估计的收敛性

b_{is}^i各为它们的第 i 行第 s 列元素.注意 $\boldsymbol{A}_j$ 和 $\boldsymbol{B}_j$ 分别为 $m\times m$ 和 $m\times p$ 矩阵.将(2.6.1)改写成分量形式有

$$y_i(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}_i+\varepsilon_i(t),\quad i=1,\cdots,m \tag{2.6.2}$$

其中定义

$$\begin{aligned}\boldsymbol{\theta}_i&=[a_{i1}^1,\cdots,a_{im}^1,\cdots,a_{i1}^{n_a},\cdots,a_{im}^{n_a};b_{i1}^1,\cdots,b_{ip}^1;\cdots;b_{i1}^{n_a},\cdots,b_{ip}^{n_b}]^{\mathrm{T}},\\ \boldsymbol{\varphi}^{\mathrm{T}}(t)&=[y^{\mathrm{T}}(t-1),\cdots,y^{\mathrm{T}}(t-n_a),u^{\mathrm{T}}(t-1),\cdots,u^{\mathrm{T}}(t-n_b)]\end{aligned} \tag{2.6.3}$$

对每个分量 $y_i(t)$,用 RLS 算法(2.1.41)和(2.1.42)可得 $\boldsymbol{\theta}_i$ 的估值

$$\begin{aligned}\hat{\boldsymbol{\theta}}_i(t+1)&=\hat{\boldsymbol{\theta}}_i(t)+\frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)[y_i(t+1)-\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}_i(t)]}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)},\\ \boldsymbol{P}(t+1)&=\boldsymbol{P}(t)-\frac{[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)][\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)]^{\mathrm{T}}}{[1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)]},\\ \hat{\boldsymbol{\theta}}_i(0)&=0,\quad \boldsymbol{P}(0)=a\boldsymbol{I},\quad a>0,i=1,2,\cdots,m\end{aligned} \tag{2.6.4}$$

注意(2.6.4)为 m 个 RLS 估值器,称为多重最小二乘法.

由(2.6.2)有 $\varepsilon_i(t)$ 和 $\boldsymbol{\varepsilon}(t)$ 的估值为

$$\hat{\varepsilon}_i(t)=y_i(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_i(t),\quad i=1,\cdots,m,$$

$$\hat{\boldsymbol{\varepsilon}}(t)=[\hat{\varepsilon}_1(t),\cdots,\hat{\varepsilon}_m(t)]^{\mathrm{T}} \tag{2.6.5}$$

$\boldsymbol{Q}_\varepsilon$ 的采样估值定义为

$$\hat{\boldsymbol{Q}}_\varepsilon(t)=\frac{1}{t}\sum_{j=1}^{t}\hat{\boldsymbol{\varepsilon}}(j)\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}(j) \tag{2.6.6}$$

它引出递推公式[11]

$$\hat{\boldsymbol{Q}}_\varepsilon(t)=\hat{\boldsymbol{Q}}_\varepsilon(t-1)+\frac{1}{t}[\hat{\boldsymbol{\varepsilon}}(t)\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}(t)-\hat{\boldsymbol{Q}}_\varepsilon(t-1)] \tag{2.6.7}$$

由单变量 RLS 算法参数估计的一致性引出上述向量 CAR 模型参数估计的多重 RLS 算法是一致的,即 RLS 估值收敛于相应真实值.

【例 2.6.1】 考虑二维 AR(1)模型

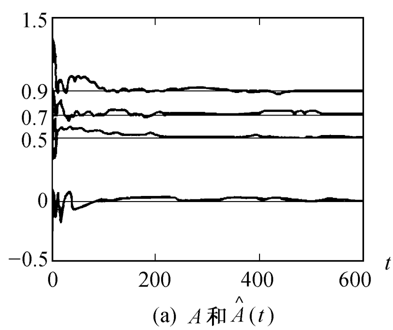
$$\boldsymbol{y}(t)=\boldsymbol{A}\boldsymbol{y}(t-1)+\boldsymbol{\varepsilon}(t) \tag{2.6.8}$$

其中 $\boldsymbol{y}(t)=[y_1(t),y_2(t)]^{\mathrm{T}}$,$\boldsymbol{\varepsilon}(t)=[\varepsilon_1(t),\varepsilon_2(t)]^{\mathrm{T}}$,$\boldsymbol{Q}_\varepsilon=\mathrm{diag}[\sigma_{\varepsilon1}^2,\sigma_{\varepsilon2}^2]$,

$$\boldsymbol{A}=\begin{bmatrix}0.9 & 0\\ 0.7 & 0.5\end{bmatrix},\quad \boldsymbol{Q}_\varepsilon=\begin{bmatrix}1 & 0\\ 0 & 0.81\end{bmatrix} \tag{2.6.9}$$

试用多重 RLS 算法估计 $\boldsymbol{A}$ 和 $\boldsymbol{Q}_\varepsilon$.

$\boldsymbol{A}$ 和 $\boldsymbol{Q}_\varepsilon$ 中的未知参数的多重 RLS 估值 $\hat{\boldsymbol{A}}(t)$ 和 $\hat{\boldsymbol{Q}}_\varepsilon$ 如图 2.6.1 所示,其中直线表示未知参数的真实值,曲线表示参数估值. 在 $t=600$ 处各参数的多重 RLS 估值为



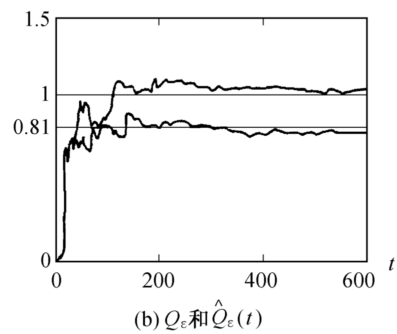


图 2.6.1 多段 RLS 参数估计的收敛性

$$\hat{a}_{11}=0.884\,707,\quad \hat{a}_{12}=0.004\,275,\quad \hat{a}_{21}=0.683\,781,$$

$$\hat{a}_{22}==0.518\,404,\quad \hat{\sigma}_{\varepsilon1}^2=1.041\,634,\quad \hat{\sigma}_{\varepsilon2}^2=0.780\,145 \tag{2.6.10}$$

它们与相应的真实值

$$a_{11}=0.9,\ a_{12}=0,\ a_{21}=0.7,\ a_{22}=0.5,\ \hat{\sigma}_{\varepsilon1}^2=1,\ \hat{\sigma}_{\varepsilon2}^2=0.8 \tag{2.6.11}$$

很相近.

2.7 向量 CAR 模型的多维 RLS 参数估计算法

在节 2.6 将向量 CAR 模型参数估计问题转化为 m 个标量 CAR 模型参数估计问题,引出了多重 RLS 参数估计算法. 本节用矩阵微分运算给出等价的多维 RLS 算法. 并证明

了它等价于多重 RLS 算法. 考虑向量 CAR 模型(2.6.1),引入未知参数阵 $\boldsymbol{\Theta}$ 和列向量 $\boldsymbol{\varphi}(t)$ 为

$$\boldsymbol{\Theta} = [\boldsymbol{A}_1, \cdots \boldsymbol{A}_{n_a}, \boldsymbol{B}_1, \cdots, \boldsymbol{B}_{n_b}],$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t) = [\boldsymbol{y}^{\mathrm{T}}(t-1), \cdots, \boldsymbol{y}^{\mathrm{T}}(t-n_a), \boldsymbol{u}^{\mathrm{T}}(t-1), \cdots, \boldsymbol{u}^{\mathrm{T}}(t-n_b)] \tag{2.7.1}$$

则向量 CAR 模型(2.6.1)可写为

$$\boldsymbol{y}(t) = \boldsymbol{\Theta\varphi}(t) + \boldsymbol{\varepsilon}(t) \tag{2.7.2}$$

基于观测 $(\boldsymbol{y}(j), \boldsymbol{u}(j)), j = 1, \cdots, t$, $\boldsymbol{\Theta}$ 的最小二乘估值 $\hat{\boldsymbol{\Theta}}(t)$ 极小化性能指标

$$\begin{aligned} J = & \sum_{j=1}^{t} [\boldsymbol{y}(j) - \boldsymbol{\Theta\varphi}(j)]^{\mathrm{T}} [\boldsymbol{y}(j) - \boldsymbol{\Theta\varphi}(j)] = \\ & \sum_{j=1}^{t} \mathrm{trace}[\boldsymbol{y}(j) - \boldsymbol{\Theta\varphi}(j)][\boldsymbol{y}(j) - \boldsymbol{\Theta\varphi}(j)]^{\mathrm{T}} = \\ & \sum_{j=1}^{t} \mathrm{trace}[\boldsymbol{\Theta\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\boldsymbol{\Theta}^{\mathrm{T}} + \boldsymbol{y}(j)\boldsymbol{y}^{\mathrm{T}}(j) - \boldsymbol{\Theta\varphi}(j)\boldsymbol{y}^{\mathrm{T}}(j) - \boldsymbol{y}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\boldsymbol{\Theta}^{\mathrm{T}}] \end{aligned} \tag{2.7.3}$$

应用矩阵迹的微分运算公式

$$\frac{\partial \mathrm{tr}(\boldsymbol{XAX}^{\mathrm{T}})}{\partial \boldsymbol{X}} = 2\boldsymbol{XA} \quad (\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}) \tag{2.7.4}$$

$$\frac{\partial \mathrm{tr}(\boldsymbol{BX}^{\mathrm{T}})}{\partial \boldsymbol{X}} = \boldsymbol{B} \tag{2.7.5}$$

$$\frac{\partial \mathrm{tr}(\boldsymbol{XB})}{\partial \boldsymbol{X}} = \boldsymbol{B}^{\mathrm{T}} \tag{2.7.6}$$

$$\frac{\partial c}{\partial X} = \boldsymbol{0} \tag{2.7.7}$$

其中 $\boldsymbol{X}, \boldsymbol{A}, \boldsymbol{B}$ 为矩阵,c 为常数.

置 $\partial J / \partial \boldsymbol{\Theta} = \boldsymbol{0}$

$$2\sum_{j=1}^{t} [\boldsymbol{\Theta\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j) - \boldsymbol{y}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)] = \boldsymbol{0} \tag{2.7.8}$$

这引出最小二乘估值

$$\hat{\boldsymbol{\Theta}}(t) = \left[\sum_{j=1}^{t} \boldsymbol{y}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\right]\left[\sum_{j=1}^{t} \boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\right]^{-1} \tag{2.7.9}$$

定义

$$\boldsymbol{P}(t) = \left[\sum_{j=1}^{t} \boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\right]^{-1} \tag{2.7.10}$$

则有

$$\hat{\boldsymbol{\Theta}}(t) = \left[\sum_{j=1}^{t} \boldsymbol{y}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\right]\boldsymbol{P}(t) \tag{2.7.11}$$

$$\boldsymbol{P}(t+1) = \left[\sum_{j=1}^{t} \boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j) + \boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\right]^{-1} \tag{2.7.12}$$

应用矩阵求逆公式(2.1.32)可导出递推公式

$$\boldsymbol{P}(t+1) = \boldsymbol{P}(t) - \boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)(1 + \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))^{-1}\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t) \tag{2.7.13}$$

或由 $\boldsymbol{P}(t)=\boldsymbol{P}^{\mathrm{T}}(t)$ 有

$$\boldsymbol{P}(t+1)=\boldsymbol{P}(t)-\frac{[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)][\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)]^{\mathrm{T}}}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{2.7.14}$$

另一方面由(2.7.11)有

$$\hat{\boldsymbol{\Theta}}(t+1)=\Big[\sum_{j=1}^{t}\boldsymbol{y}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)+\boldsymbol{y}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\Big]\boldsymbol{P}(t+1) \tag{2.7.15}$$

将(2.7.14)代入(2.7.15)有

$$\begin{aligned}
\hat{\boldsymbol{\Theta}}(t+1)=&\Big[\sum_{j=1}^{t}\boldsymbol{y}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)+\boldsymbol{y}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\Big]\times\\
&\Big[\boldsymbol{P}(t)-\frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}\Big]=\\
&\hat{\boldsymbol{\Theta}}(t)+\boldsymbol{y}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)-\frac{\sum_{j=1}^{t}\boldsymbol{y}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}-\\
&\frac{\boldsymbol{y}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}=\\
&\hat{\boldsymbol{\Theta}}(t)-\frac{\hat{\boldsymbol{\Theta}}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}+\\
&\frac{(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))\boldsymbol{y}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}-\\
&\frac{\boldsymbol{y}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}=\\
&\hat{\boldsymbol{\Theta}}(t)+\frac{[\boldsymbol{y}(t+1)-\hat{\boldsymbol{\Theta}}(t)\boldsymbol{\varphi}(t+1)]\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}
\end{aligned} \tag{2.7.16}$$

其中应用了定义:对矩阵 $\boldsymbol{A}=(a_{ij})$,$k\boldsymbol{A}=(ka_{ij})$,$k$ 为常数,且有

$$\begin{aligned}
&\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{y}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)=\\
&\boldsymbol{y}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)
\end{aligned} \tag{2.7.17}$$

因为 $\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)$ 是标量.

上述推导可概括为如下定理.

【定理 2.7.1】 向量 CAR 模型(2.6.1)参数的多维递推最小二乘估值为

$$\hat{\boldsymbol{\Theta}}(t+1)=\hat{\boldsymbol{\Theta}}(t)+\frac{[\boldsymbol{y}(t+1)-\hat{\boldsymbol{\Theta}}(t)\boldsymbol{\varphi}(t+1)]\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{2.7.18}$$

$$\boldsymbol{P}(t+1)=\boldsymbol{P}(t)-\frac{[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)][\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)]^{\mathrm{T}}}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{2.7.19}$$

带初值 $\hat{\boldsymbol{\Theta}}(0)=\mathbf{0}$,$\boldsymbol{P}(0)=a\boldsymbol{I}$,$a>0$ 为很大实数,规定 $\boldsymbol{y}(i)=\mathbf{0}(i\leqslant 0)$,$\boldsymbol{u}(i)=\mathbf{0}(i\leqslant 0)$.

下面我们证明一个关系

$$\frac{\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}=\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t+1) \tag{2.7.20}$$

事实上,应用(2.7.19)有

$$\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t+1)=\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\Big[\boldsymbol{P}(t)-\frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}\Big]=$$

$$\frac{\boldsymbol{\varphi}^{\mathrm{T}}(t+1)[(1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1))\boldsymbol{P}(t)-\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)]}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}=$$

$$\frac{\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)]}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}-$$

$$\frac{\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}=$$

$$\frac{\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{2.7.21}$$

这引出(2.7.20)成立.其中用到了事实

$$\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{P}(t)=$$
$$\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t) \tag{2.7.22}$$

这是因为 $\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)$ 是标量.因此定理 2.7.1 可简化为如下定理.

【定理 2.7.2】 CAR 模型(2.6.1)参数的多维递推最小二乘估值为

$$\hat{\boldsymbol{\Theta}}(t+1)=\hat{\boldsymbol{\Theta}}(t)+[\boldsymbol{y}(t+1)-\hat{\boldsymbol{\Theta}}(t)\boldsymbol{\varphi}(t+1)]\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t+1) \tag{2.7.23}$$

$$\boldsymbol{P}(t+1)=\boldsymbol{P}(t)-\frac{[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)][\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)]^{\mathrm{T}}}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{2.7.24}$$

带初值 $\hat{\boldsymbol{\Theta}}(0)=\boldsymbol{0}$,$\boldsymbol{P}(0)=a\boldsymbol{I}$,$a>0$ 为很大实数,规定 $\boldsymbol{y}(i)=\boldsymbol{0}(i\leqslant 0)$,$\boldsymbol{u}(i)=\boldsymbol{0}(i\leqslant 0)$.

对于时变参数 $\boldsymbol{\Theta}$,用指数加权选择 $\hat{\boldsymbol{\Theta}}(t)$ 极小化

$$J=\sum_{j=1}^{t}\lambda^{t-j}[\boldsymbol{y}(j)-\hat{\boldsymbol{\Theta}}\boldsymbol{\varphi}(j)]^{\mathrm{T}}[\boldsymbol{y}(j)-\hat{\boldsymbol{\Theta}}\boldsymbol{\varphi}(j)] \tag{2.7.25}$$

带遗忘因子 λ,$0<\lambda\leqslant 1$,平行于上述推导有如下定理.

【定理 2.7.3】 带时变参数 CAR 模型(2.6.1)的带遗忘因子 λ 的多维 RLS 参数估计算法为

$$\hat{\boldsymbol{\Theta}}(t+1)=\hat{\boldsymbol{\Theta}}(t)+[\boldsymbol{y}(t+1)-\hat{\boldsymbol{\Theta}}(t)\boldsymbol{\varphi}(t+1)]\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t+1)$$

$$\boldsymbol{P}(t+1)=\frac{1}{\lambda}\left\{\boldsymbol{P}(t)-\frac{[\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)][\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)]^{\mathrm{T}}}{\lambda+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}\right\} \tag{2.7.26}$$

上述多维 RLS 算法(2.7.18)~(2.7.26)的优点是不要求矩阵求逆运算,因而可实时应用.

【定理 2.7.4】 (多维 RLS 算法与多重 RLS 算法的等价性)多维 RLS 算法(2.7.23)和(2.7.24)等价于多重 RLS 算法(2.6.4).

证明 由 $\boldsymbol{\theta}_i$ 的定义(2.6.3)和 $\boldsymbol{\Theta}$ 的定义(2.7.1)有关系

$$\boldsymbol{\Theta}=\begin{bmatrix}\boldsymbol{\theta}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\theta}_m^{\mathrm{T}}\end{bmatrix},\quad \hat{\boldsymbol{\Theta}}(t)=\begin{bmatrix}\hat{\boldsymbol{\theta}}_1^{\mathrm{T}}(t)\\ \vdots\\ \hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\end{bmatrix} \tag{2.7.27}$$

应用(2.7.20)将(2.7.27)和(2.7.20)代入(2.7.23)中并对(2.7.23)两边转置,则可导出(2.6.4).证毕. □

定义 $\boldsymbol{\varepsilon}(t)$ 的估值为

$$\hat{\boldsymbol{\varepsilon}}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{\Theta}}(t)\boldsymbol{\varphi}(t) \tag{2.7.28}$$

它等价于定义(2.6.5),于是有 $\boldsymbol{Q}_{\varepsilon}$ 的估值(2.6.7).

【注 2.7.1】 Ljung[19]证明了多维 RLS 参数估计算法(2.7.11)和(2.7.12)是一致的，即在较弱条件下，估值 $\hat{\boldsymbol{\Theta}}(t)$以概率 1 收敛于真实值 $\boldsymbol{\Theta}$，$\hat{\boldsymbol{\Theta}}(t)\to\boldsymbol{\Theta}$，$t\to\infty$.

【例 2.7.1】 考虑二维 CAR 模型

$$\boldsymbol{y}(t)=\boldsymbol{A}\boldsymbol{y}(t-1)+\boldsymbol{B}\boldsymbol{u}(t-1)+\boldsymbol{\varepsilon}(t) \tag{2.7.29}$$

其中 $\boldsymbol{\varepsilon}(t)$是零均值、方差阵为 $\boldsymbol{Q}_{\varepsilon}$ 的高斯白噪声，已知控制输入 $\boldsymbol{u}(t)=[u_1(t),u_2(t)]^{\mathrm{T}}$为

$$u_1(t)=\sin(2\pi t/100),\quad u_2(t)=\cos(2\pi t/100) \tag{2.7.30}$$

且

$$\boldsymbol{A}=\begin{bmatrix}-0.8 & 0\\ -0.3 & 0.3\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}-0.3 & 0\\ 0.3 & 0.5\end{bmatrix},\quad \boldsymbol{Q}_{\varepsilon}=\begin{bmatrix}1 & 0\\ 0 & 0.81\end{bmatrix} \tag{2.7.31}$$

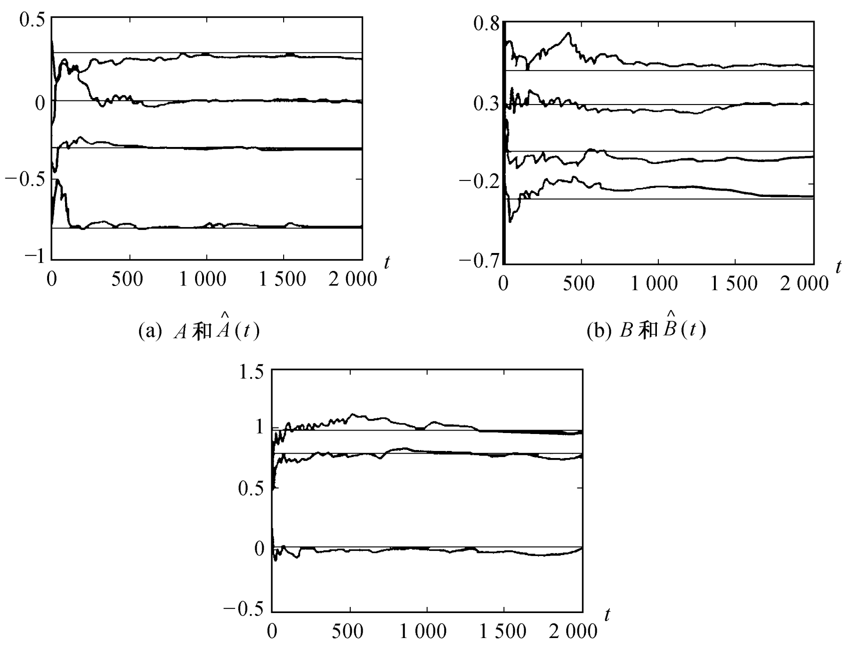
假设模型参数阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 及方差 $\boldsymbol{Q}_{\varepsilon}$ 是未知的. 试用多维 RLS 算法估计它们.

多维 RLS 参数估值 $\hat{\boldsymbol{A}}(t)$，$\hat{\boldsymbol{B}}(t)$和 $\hat{\boldsymbol{Q}}_{\varepsilon}(t)$的收敛性如图 2.7.1 所示，其中直线代表未知参数真实值，曲线表示估值. 在 $t=2\,000$ 处估值为

$$\hat{\boldsymbol{A}}=\begin{bmatrix}-0.793\,007 & -0.012\,611\\ -0.304\,050 & 0.284\,420\end{bmatrix},\quad \hat{\boldsymbol{B}}=\begin{bmatrix}-0.268\,339 & -0.028\,736\\ 0.305\,592 & 0.530\,194\end{bmatrix},$$

$$\hat{\boldsymbol{Q}}_{\varepsilon}=\begin{bmatrix}0.996\,070 & -0.006\,539\\ -0.006\,539 & 0.808\,750\end{bmatrix} \tag{2.7.32}$$

它们近似等于真实值(2.7.31).



(a) A 和 $\hat{A}(t)$

(b) B 和 $\hat{B}(t)$

(c) Q_{ε}和 $\hat{Q}_{\varepsilon}(t)$

图 2.7.1 多维 RLS 参数估计的收敛性

2.8 向量 CARMA 模型的多重和多维 RELS 参数估计算法

考虑向量 CARMA 模型

$$\boldsymbol{y}(t)=\boldsymbol{A}_1\boldsymbol{y}(t-1)+\cdots+\boldsymbol{A}_{n_a}\boldsymbol{y}(t-n_a)+\boldsymbol{B}_1\boldsymbol{u}(t-1)+\cdots+\boldsymbol{B}_{n_b}\boldsymbol{u}(t-n_b)+\boldsymbol{\varepsilon}(t)+\boldsymbol{D}_1\boldsymbol{\varepsilon}(t-1)+\cdots+\boldsymbol{D}_{n_d}\boldsymbol{\varepsilon}(t-n_d) \tag{2.8.1}$$

其中输出 $\boldsymbol{y}(t)\in R^m$，控制 $\boldsymbol{u}(t)\in R^p$，$\boldsymbol{\varepsilon}(t)\in R^m$ 是带零均值和未知方差阵为 $\boldsymbol{Q}_{\boldsymbol{\varepsilon}}$ 的白噪声。设 $y_i(t)$，$u_i(t)$，$\varepsilon_i(t)$分别为 $\boldsymbol{y}(t)$，$\boldsymbol{u}(t)$，$\boldsymbol{\varepsilon}(t)$的第 i 个分量，未知参数阵 $\boldsymbol{A}_i$，$\boldsymbol{B}_i$ 和 $\boldsymbol{D}_i$ 的第 r 行、第 s 列元素分别为 a_{rs}^i，b_{rs}^i，d_{rs}^i.

2.8.1 多重 RELS 参数估计算法

向量 CRARMA 模型(2.8.1)按分量可化为 m 个标量 CARMA 模型.它们有 LS 结构

$$y_i(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}_i+\boldsymbol{\varepsilon}_i(t),\quad i=1,\cdots,m \tag{2.8.2}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t)=[\boldsymbol{y}^{\mathrm{T}}(t-1),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots\boldsymbol{y}^{\mathrm{T}}(t-n_a),\boldsymbol{u}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n_b),\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}(t-1),\cdots,\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}(t-n_d)] \tag{2.8.3}$$

$$\boldsymbol{\theta}_i=[a_{11}^1,\cdots,a_{1m}^1,\cdots,a_{11}^{n_a},\cdots,a_{1m}^{n_a};b_{11}^1,\cdots,b_{1p}^1,\cdots,b_{11}^{n_b},\cdots,b_{1p}^{n_b};d_{11}^1,\cdots,d_{1m}^1,\cdots,d_{11}^{n_d},\cdots,d_{1m}^{n_d}]^{\mathrm{T}},\quad i=1,\cdots,m \tag{2.8.4}$$

于是用标量 RELS 算法(2.2.6)和(2.2.7)可得 RELS 估值 $\hat{\boldsymbol{\theta}}(t+1)$，其中在 $\hat{\boldsymbol{\varphi}}(t+1)$中 $\hat{\boldsymbol{\varepsilon}}(t),\cdots,\hat{\boldsymbol{\varepsilon}}(t+1-n_d)$用下式计算

$$\hat{\boldsymbol{\varepsilon}}(j)=[\hat{\varepsilon}_1(j),\cdots,\hat{\varepsilon}_m(j)]^{\mathrm{T}},\quad j=t,t-1,\cdots,t+1-n_d,$$

$$\hat{\varepsilon}(j)=y_i(j)-\boldsymbol{\varphi}^{\mathrm{T}}(j)\hat{\boldsymbol{\theta}}_i(j-1),\quad j=t,t-1,\cdots,t+1-n_d,i=1,\cdots,m \tag{2.8.5}$$

而 $\boldsymbol{Q}_\varepsilon$ 估值$\hat{\boldsymbol{Q}}_\varepsilon(t)$由(2.6.7)计算.

2.8.2 多维 RELS 参数估计算法

置未知参数阵为

$$\boldsymbol{\Theta}=[\boldsymbol{A}_i,\cdots,\boldsymbol{A}_{n_a},\boldsymbol{B}_i,\cdots,\boldsymbol{B}_{n_a},\boldsymbol{D}_i,\cdots,\boldsymbol{D}_{n_d}] \tag{2.8.6}$$

则(2.8.1)有 LS 结构

$$\boldsymbol{y}(t)=\boldsymbol{\Theta}\boldsymbol{\varphi}(t)+\boldsymbol{\varepsilon}(t) \tag{2.8.7}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t)=[\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{y}^{\mathrm{T}}(t-n_a),\boldsymbol{u}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n_b),\cdots,\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}(t-1),\cdots,\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}(t-n_d)] \tag{2.8.8}$$

则由(2.7.23)和(2.7.24)引出多维 RELS 估值 $\hat{\boldsymbol{\Theta}}(t)$，其中在 $\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$中的 $\hat{\boldsymbol{\varepsilon}}(t),\cdots,\hat{\boldsymbol{\varepsilon}}(t+1-n_d)$用下式计算

$$\hat{\boldsymbol{\varepsilon}}(j)=\boldsymbol{y}(j)-\hat{\boldsymbol{\Theta}}(j-1)\boldsymbol{\varphi}(j),\quad j=t,t-1,\cdots,t+1-n_d \tag{2.8.9}$$

估值 $\hat{\boldsymbol{Q}}_\varepsilon(t)$用(2.6.7)计算.

【例 2.8.1】 考虑二维 ARMA 模型

$$\boldsymbol{y}(t)=\boldsymbol{A}\boldsymbol{y}(\mathrm{t}-1)+\boldsymbol{\varepsilon}(t)+\boldsymbol{D}\boldsymbol{\varepsilon}(t-1) \tag{2.8.10}$$

其中 $\boldsymbol{\varepsilon}(t)$是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的高斯白噪声，$\boldsymbol{y}(t)=[y_1(t),y_2(t)]^{\mathrm{T}}$，$\boldsymbol{\varepsilon}(t)=[\varepsilon_1(t),$

$\varepsilon_2(t)]^{\mathrm{T}}$

$$\boldsymbol{A}=\begin{bmatrix}-0.8 & 0\\ -0.3 & 0.3\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}-0.3 & 0\\ 0.3 & 0.5\end{bmatrix},\quad \boldsymbol{Q}_{\varepsilon}=\begin{bmatrix}1 & 0\\ 0 & 1.62\end{bmatrix} \tag{2.8.11}$$

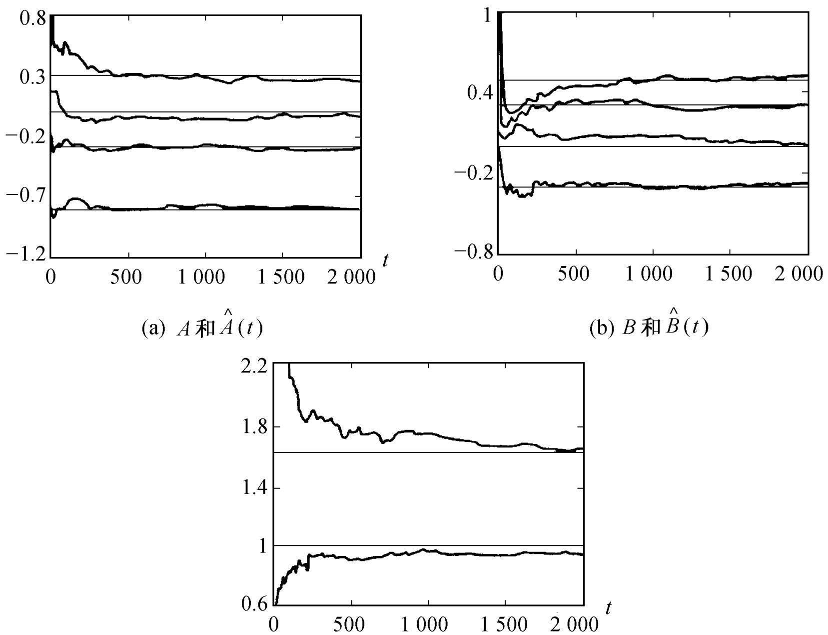
试用多重 RELS 算法求估值 $\hat{\boldsymbol{A}}(t),\hat{\boldsymbol{D}}(t),\hat{\boldsymbol{Q}}_{\varepsilon}(t)$.

多重 RELS 参数估值 $\hat{\boldsymbol{A}}(t)$，$\hat{\boldsymbol{D}}(t)$和 $\hat{\boldsymbol{Q}}_{\varepsilon}(t)$的收敛性如图 2.8.1 所示，其中直线代表未知参数真实值，曲线表示估值. 在 $t=2\ 000$ 处的估值为

$$\hat{\boldsymbol{A}}=\begin{bmatrix}-0.795\ 992 & -0.021\ 637\\ -0.309\ 511 & 0.269\ 013\end{bmatrix},\quad \hat{\boldsymbol{D}}=\begin{bmatrix}-0.282\ 508 & 0.028\ 847\\ 0.309\ 264 & 0.524\ 325\end{bmatrix},$$

$$\hat{\boldsymbol{Q}}_{\varepsilon}=\begin{bmatrix}0.952\ 198 & 0\\ 0 & 1.652\ 647\end{bmatrix} \tag{2.8.12}$$

它们逼近相应的真实值(2.8.11).



(a) A 和 $\hat{A}(t)$　　(b) B 和 $\hat{B}(t)$

(c) Q_{ε} 和 $\hat{Q}_{\varepsilon}(t)$

图 2.8.1　多重 RLS 参数估计的收敛性

2.9　向量 CARMA 模型的两段 RLS – RELS 参数估计算法

在节 2.8.2 给出的向量 CARMA 模型的多维 RELS 算法的缺点是由(2.8.8)计算 $\boldsymbol{\varphi}^{\mathrm{T}}(t)$时由(2.8.9)给出估值 $\hat{\boldsymbol{\varepsilon}}(j)$与参数估值 $\hat{\boldsymbol{\Theta}}(j-1)$是互耦的. 这将影响 $\hat{\boldsymbol{\varepsilon}}(j)$估值精度和值 $\hat{\boldsymbol{\Theta}}(t)$的收敛速度，特别将影响估值 $\hat{\boldsymbol{D}}(t)$收敛速度. 为此本节提出用一个单独的 RLS 估值器估计 $\boldsymbol{\varepsilon}(j)$，并利用所得估值 $\hat{\boldsymbol{\varepsilon}}(j)$到 RELS 算法中，从而引出两段 RLS – RELS 算法，可明显改进收敛速度和参数估计精度.

2.9.1 第一段——用 RLS 算法拟合高阶 CAR 模型产生白噪声估值

考虑向量 CARMA 模型(2.8.1),其中假设多项式矩阵 $\boldsymbol{A}(q^{-1})=\boldsymbol{I}_m-\boldsymbol{A}_1q^{-1}-\cdots-\boldsymbol{A}_{n_a}q^{-n_a}$和$\boldsymbol{D}(q^{-1})=\boldsymbol{I}_m-\boldsymbol{D}_1q^{-1}-\cdots-\boldsymbol{D}_{n_d}q^{-n_d}$是稳定的.于是由定理 1.1.10 向量 CARMA 模型可用高阶向量 CAR 模型近似代替为

$$\boldsymbol{y}(t)=\sum_{j=1}^{n_0}\boldsymbol{\Phi}_j\boldsymbol{y}(t-j)+\sum_{j=1}^{n_0}\boldsymbol{\Psi}_j\boldsymbol{u}(t-j)+\boldsymbol{\varepsilon}(t) \tag{2.9.1}$$

用多维 RLS 算法(2.7.23)和(2.7.24)可得估值 $\hat{\boldsymbol{\Phi}}_j(t+1)$,$\hat{\boldsymbol{\Psi}}_j(t+1)$,从而可得白噪声平滑估值

$$\hat{\boldsymbol{\varepsilon}}(i)=\boldsymbol{y}(i)-\sum_{j=1}^{n_0}\hat{\boldsymbol{\Phi}}(t+1)\boldsymbol{y}(i-j)+\sum_{j=1}^{n_0}\hat{\boldsymbol{\Psi}}(t+1)\boldsymbol{u}(i-j) \tag{2.9.2}$$

其中 $i=t,t-1,\cdots,t+1-n_d$.估值 $\hat{\boldsymbol{Q}}_\varepsilon(t+1)$可用(2.6.7)计算.

2.9.2 第二段——改进的多维 RELS 算法

向量 CARMA 模型(2.8.1)有 LS 结构(2.8.7),将用(2.9.2)计算的白噪声 $\varepsilon(t)$的平滑值 $\hat{\boldsymbol{\varepsilon}}(i)$,$i=,t,t-1,\cdots,t+1-n_d$,代入 $\boldsymbol{\varphi}^{\mathrm{T}}(t+1)$中,则由(2.7.23)和(2.7.24)得改进的多维 RELS 参数估值 $\hat{\boldsymbol{\Theta}}(t+1)$.

【例 2.9.1】 考虑二维 ARMA 模型

$$\boldsymbol{y}(t)=\boldsymbol{A}\boldsymbol{y}(t-1)+\boldsymbol{\varepsilon}(t)+\boldsymbol{D}_1\boldsymbol{\varepsilon}(t-1) \tag{2.9.3}$$

其中 $\boldsymbol{\varepsilon}(t)$是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的高斯白噪声,且

$$\boldsymbol{A}=\begin{bmatrix}-0.8 & 0\\ -0.3 & 0.2\end{bmatrix},\quad \boldsymbol{D}=\begin{bmatrix}-0.3 & 0\\ 0.2 & 0.6\end{bmatrix},\quad \boldsymbol{Q}_{\boldsymbol{\varepsilon}}=\begin{bmatrix}1 & 0\\ 0 & 0.81\end{bmatrix} \tag{2.9.4}$$

试用多维两段 RLS-RELS 算法求估值 $\hat{\boldsymbol{A}}(t)$,$\hat{\boldsymbol{D}}(t)$和 $\hat{\boldsymbol{Q}}_\varepsilon(t)$.

取 $n_0=10$,即对 $\boldsymbol{y}(t)$拟合 AR(10),所得 $\boldsymbol{Q}_\varepsilon$ 的估值$\hat{\boldsymbol{A}}(t)$如图 2.9.1(c)所示,多维两段 RLS-RELS 估值 $\hat{\boldsymbol{A}}(t)$,$\hat{\boldsymbol{D}}(t)$和 $\hat{\boldsymbol{Q}}_\varepsilon(t)$如图 2.9.1(a),(b),(d)所示.图中直线代表未知参数真实值,曲线代表参数估值。$t=2\,000$ 处的估值为

$$\hat{\boldsymbol{A}}=\begin{bmatrix}-0.799\,998 & -0.022\,234\\ -0.306\,413 & 0.180\,111\end{bmatrix},\quad \hat{\boldsymbol{D}}=\begin{bmatrix}-0.275\,445 & -0.026\,699\\ 0.210\,255 & 0.614\,487\end{bmatrix},$$

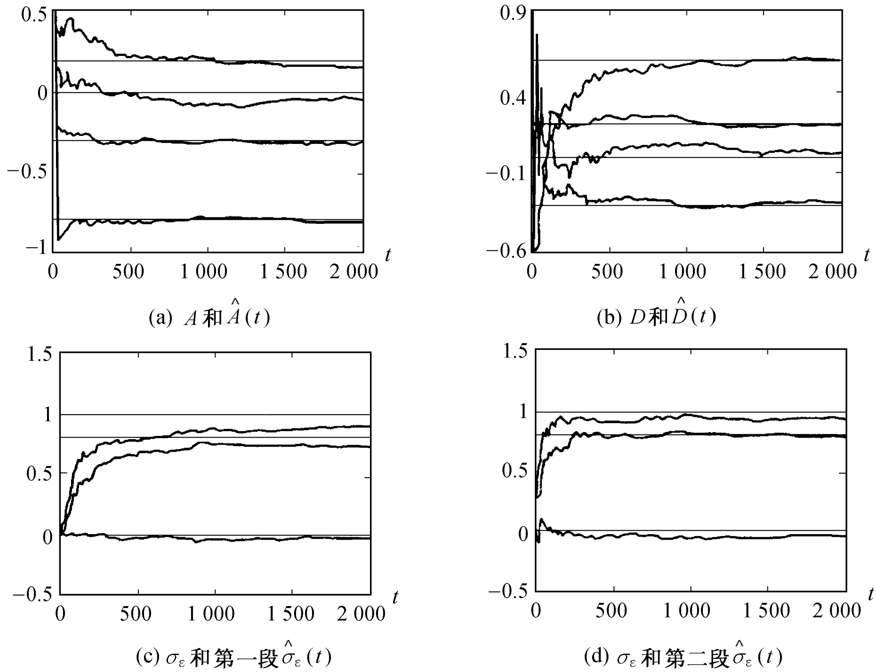
$$\hat{\boldsymbol{Q}}_\varepsilon=\begin{bmatrix}0.908\,458 & -0.018\,387\\ -0.018\,387 & 0.0.761\,772\end{bmatrix},\quad \hat{\boldsymbol{Q}}_\varepsilon=\begin{bmatrix}0.957\,332 & -0.024\,462\\ -0.024\,462 & 0.809\,274\end{bmatrix} \tag{2.9.5}$$

其中前一个 $\hat{\boldsymbol{Q}}_\varepsilon$ 为第一段 RLS 估值;后一个 $\hat{\boldsymbol{Q}}_\varepsilon$ 为第二段 RELS 估值.它们与相应的真实值(2.9.4)十分相近.

2.10 向量 ARMA 模型的两段 RLS-LS 参数估计算法

本节将把节 2.4 标量 ARMA 模型的两段 RLS-LS 参数估计算法推广到向量 ARMA 模型情形.考虑向量 ARMA 模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{2.10.1}$$



(a) A和$\hat{A}(t)$　(b) D和$\hat{D}(t)$

(c) σ_ε和第一段$\hat{\sigma}_\varepsilon(t)$　(d) σ_ε和第二段$\hat{\sigma}_\varepsilon(t)$

图 2.9.1　多维两段 RLS－RELS 算法的收敛性

其中 $\boldsymbol{\varepsilon}(t)\in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，观测 $\boldsymbol{y}(t)\in R^m$，且 $\boldsymbol{A}(q^{-1})$和 $\boldsymbol{D}(q^{-1})$是稳定的多项式矩阵，

$$\boldsymbol{A}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{A}_1q^{-1}+\cdots\boldsymbol{A}_{n_a}q^{-n_a},$$

$$\boldsymbol{D}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{D}_1q^{-1}+\cdots\boldsymbol{D}_{n_d}q^{-n_d}\tag{2.10.2}$$

$\boldsymbol{A}_i$ 和 $\boldsymbol{D}_i$ 为 $m\times m$ 参数阵. 问题是由观测 $(\boldsymbol{y}(t),\cdots,\boldsymbol{y}(1))$求未知参数阵 $\boldsymbol{A}_i,\boldsymbol{D}_i,\boldsymbol{Q}_\varepsilon$ 的估值.

2.10.1　用多维 RLS 算法拟合高阶向量 AR 模型

无穷阶向量 AR(∞)模型

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Pi}_jq^{-j}\boldsymbol{y}(t)\tag{2.10.3}$$

其中 $\boldsymbol{\Pi}_0=\boldsymbol{I}_m,\boldsymbol{\Pi}_j\to 0\,(j\to\infty)$. 取 n_0 充分大，则有近似的高阶 AR(n_0)模型

$$\boldsymbol{\Pi}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{\varepsilon}(t)\tag{2.10.4}$$

$$\boldsymbol{\Pi}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{\Pi}_1q^{-1}+\cdots+\boldsymbol{\Pi}_{n_0}q^{-n_0}\tag{2.10.5}$$

比较(2.10.3)与(2.10.4)有近似关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\Pi}(q^{-1})=\boldsymbol{A}(q^{-1})\tag{2.10.6}$$

注意(2.10.4)有 LS 结构

$$\boldsymbol{y}(t)=\boldsymbol{\Theta}\boldsymbol{\varphi}(t)+\boldsymbol{\varepsilon}(t)\tag{2.10.7}$$

$$\boldsymbol{\Theta}=[\boldsymbol{\Pi}_1,\boldsymbol{\Pi}_2,\cdots,\boldsymbol{\Pi}_{n_0}]\tag{2.10.8}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(\mathrm{t})=[-\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots-\boldsymbol{y}^{\mathrm{T}}(t-n_0)]\tag{2.10.9}$$

其中 T 为转置号，于是多维 RLS 算法(2.7.23)和(2.7.24)可得估值 $\hat{\boldsymbol{\Theta}}(t)$，即估值 $\hat{\boldsymbol{\Pi}}_i(t)$，$i=1,\cdots,n_0$，以及用(2.7.28)和(2.6.7)得估值 $\hat{\boldsymbol{Q}}_\varepsilon(t)$.

2.10.2 用求矩阵线性方程组最小二乘解估计向量 ARMA 模型参数

比较(2.10.6)两边 q^{-i}的系数阵有关系

$$\boldsymbol{\Theta}_A=\boldsymbol{\Gamma}+\boldsymbol{GP} \tag{2.10.10}$$

$$\boldsymbol{M\Theta}_D=\boldsymbol{\Delta} \tag{2.10.11}$$

其中定义 $\boldsymbol{D}_i=\boldsymbol{0}(i>n_d)$，$\boldsymbol{\Pi}_i=\boldsymbol{0}(i>n_0)$，$\boldsymbol{\Pi}_i=\boldsymbol{0}(i<n_d)$，

$$\boldsymbol{\Theta}_A=\begin{bmatrix}\boldsymbol{A}_1\\ \boldsymbol{A}_2\\ \vdots\\ \boldsymbol{A}_{n_a}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\boldsymbol{D}_1\\ \boldsymbol{D}_2\\ \vdots\\ \boldsymbol{D}_{n_a}\end{bmatrix},\quad \boldsymbol{G}=\begin{bmatrix}\boldsymbol{I}_m & \boldsymbol{0} & \cdots & \boldsymbol{0}\\ D_1 & \boldsymbol{I}_m & & \vdots\\ \vdots & & \ddots & \boldsymbol{0}\\ \boldsymbol{D}_{n_a-1} & \cdots & \boldsymbol{D}_1 & \boldsymbol{I}_m\end{bmatrix},\quad \boldsymbol{P}=\begin{bmatrix}\boldsymbol{\Pi}_1\\ \boldsymbol{\Pi}_2\\ \vdots\\ \boldsymbol{\Pi}_{n_a}\end{bmatrix},$$

$$\boldsymbol{\Theta}_D=\begin{bmatrix}\boldsymbol{D}_1^{\mathrm{T}}\\ \boldsymbol{D}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{D}_{n_d}^{\mathrm{T}}\end{bmatrix},\quad \boldsymbol{M}=\begin{bmatrix}\boldsymbol{\Pi}_{n_a}^{\mathrm{T}} & \cdots & \boldsymbol{\Pi}_{n_a-n_d+1}^{\mathrm{T}}\\ \boldsymbol{\Pi}_{n_a+1}^{\mathrm{T}} & \cdots & \boldsymbol{\Pi}_{n_a-n_d+2}^{\mathrm{T}}\\ \vdots & & \vdots\\ \boldsymbol{\Pi}_{n_d+n_0-1}^{\mathrm{T}} & \cdots & \boldsymbol{\Pi}_{n_0}^{\mathrm{T}}\end{bmatrix},\quad \boldsymbol{\Delta}=\begin{bmatrix}-\boldsymbol{\Pi}_{n_a+1}^{\mathrm{T}}\\ -\boldsymbol{\Pi}_{n_a+2}^{\mathrm{T}}\\ \vdots\\ -\boldsymbol{\Pi}_{n_d+n_0}^{\mathrm{T}}\end{bmatrix} \tag{2.10.12}$$

不相容矩阵线性方程组(2.10.11)的最小二乘解为

$$\boldsymbol{\Theta}_D=(\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M})^{-1}\boldsymbol{M}^{\mathrm{T}}\boldsymbol{\Delta} \tag{2.10.13}$$

将 RLS 估值 $\hat{\boldsymbol{\Pi}}_i(t)$代入 $\boldsymbol{M}$ 和 $\boldsymbol{\Delta}$ 可得估值$\hat{\boldsymbol{M}}(t)$和 $\hat{\Delta}(t)$，进而由(2.10.13)可得估值

$$\hat{\boldsymbol{\Theta}}_D(t)=[\hat{\boldsymbol{M}}^{\mathrm{T}}(t)\boldsymbol{M}(t)]^{-1}\hat{\boldsymbol{M}}^{\mathrm{T}}(t)\hat{\boldsymbol{\Delta}}(t) \tag{2.10.14}$$

即可得估值 $\hat{\boldsymbol{D}}_i(t)$. 再将估值 $\hat{\boldsymbol{D}}_i(t)$和 $\hat{\boldsymbol{\Pi}}_i(t)$代入(2.10.10)可得估值

$$\hat{\boldsymbol{\Theta}}_A(t)=\hat{\boldsymbol{\Gamma}}(t)+\hat{\boldsymbol{G}}(t)\hat{\boldsymbol{P}}(t) \tag{2.10.15}$$

即可得估值 $\hat{\boldsymbol{A}}_i(t)$.

特别，当 $\boldsymbol{A}(q^{-1})=\boldsymbol{I}_m$，即 $n_a=0$，对向量 MA 模型

$$\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{2.10.16}$$

仍可用(2.10.11)和(2.10.12)求 MA 参数估值 $\hat{\boldsymbol{\Theta}}_D(t)$，但应在(2.10.12)$\boldsymbol{M}$ 和 $\boldsymbol{\Delta}$ 的表达式中置 $n_a=0$，$\boldsymbol{\Pi}_i=0(\boldsymbol{i}<0)$和 $\boldsymbol{\Pi}_i=0(i>n_0)$.

【例 2.10.1】 考虑二级 ARMA 模型

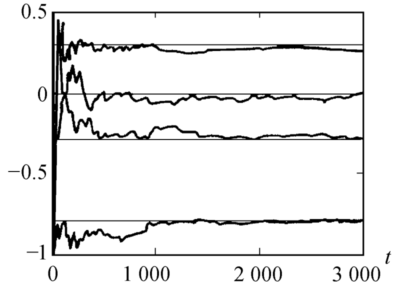
$$\boldsymbol{y}(t)=\boldsymbol{Ay}(t-1)+\boldsymbol{\varepsilon}(t)+\boldsymbol{D\varepsilon}(t-1) \tag{2.10.17}$$

其中 $\boldsymbol{\varepsilon}(t)$是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的高斯白噪声，且

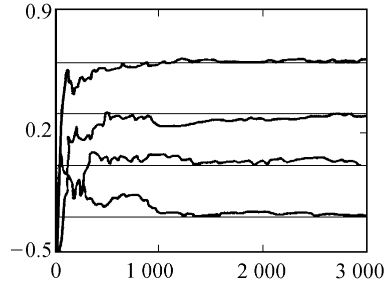
$$\boldsymbol{A}=\begin{bmatrix}-0.8 & 0\\ -0.3 & 0.3\end{bmatrix},\quad \boldsymbol{D}=\begin{bmatrix}-0.3 & 0\\ 0.3 & 0.6\end{bmatrix},\quad \boldsymbol{Q}_{\boldsymbol{\varepsilon}}=\begin{bmatrix}0.16 & 0\\ 0 & 0.09\end{bmatrix} \tag{2.10.18}$$

试用多维二段 RLS－L 算法估计 $\boldsymbol{A}$，$\boldsymbol{D}$ 和 $\boldsymbol{Q}_{\boldsymbol{\varepsilon}}$.

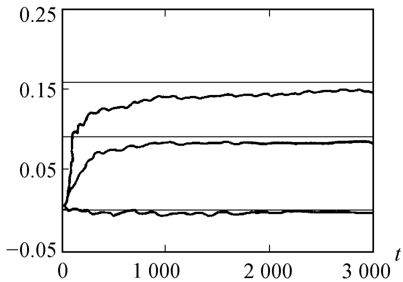
在第一段中取 $n_0=10$，即对 $\boldsymbol{y}(t)$拟合 AR(10). 二段 RLS－LS 估值 $\hat{\boldsymbol{A}}(t)$，$\hat{\boldsymbol{D}}(t)$和 $\hat{\boldsymbol{Q}}_\varepsilon(t)$的收敛性如图 2.10.1 所示，其中直线代表未知参数真实值，曲线表示估值. 在 $t=3\,000$处的估值为



(a) A和$\hat{A}(t)$



(b) D和$\hat{D}(t)$



(c) Q_ε和$\hat{Q}_\varepsilon(t)$

图 2.10.1　多维两段 RLS – LS 参数估计的收敛性

$$\hat{\boldsymbol{A}} = \begin{bmatrix} -0.801\,081 & 0.006\,086 \\ -0.285\,707 & 0.267\,245 \end{bmatrix},\quad \hat{\boldsymbol{D}} = \begin{bmatrix} -0.284\,104 & -0.006\,430 \\ 0.289\,012 & 0.624\,330 \end{bmatrix},$$

$$\hat{\boldsymbol{Q}}_\varepsilon = \begin{bmatrix} 0.150\,705 & -0.001\,815 \\ -0.001\,815 & 0.086\,393 \end{bmatrix} \tag{2.10.19}$$

它们近似等于相应的真实值(2.10.18).

【例 2.10.2】　考虑二维 MA 模型

$$\boldsymbol{y}(t) = (\boldsymbol{I}_2 + \boldsymbol{D}_1 q^{-1} + \boldsymbol{D}_2 q^{-2})\boldsymbol{\varepsilon}(t) \tag{2.10.20}$$

其中 $\boldsymbol{\varepsilon}(t)$是零均值、方差为 $\boldsymbol{Q}_\varepsilon$ 的高斯白噪声,且

$$\boldsymbol{D}_1 = \begin{bmatrix} -0.7 & 0.2 \\ 0.4 & -1.1 \end{bmatrix},\quad \boldsymbol{D}_2 = \begin{bmatrix} 0.14 & -0.1 \\ -0.24 & 0.3 \end{bmatrix},\quad Q_\varepsilon = \begin{bmatrix} 0.36 & 0 \\ 0 & 0.48 \end{bmatrix} \tag{2.10.21}$$

试用多维二段 RLS – LS 算法估计 $\boldsymbol{D}_1$,$\boldsymbol{D}_2$ 和 $\boldsymbol{Q}_\varepsilon$.

取 $n_0 = 10$,两段 RLS – LS 估值 $\hat{\boldsymbol{D}}_1(t)$,$\hat{\boldsymbol{D}}_2(t)$的收敛性如图 2.10.2 所示,直线代表未知参数真实值,曲线表示估值曲线.

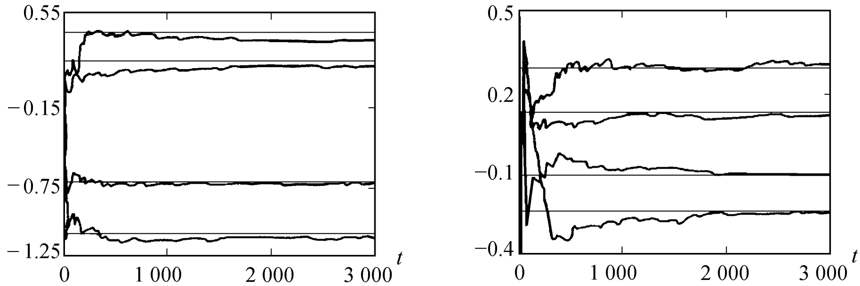
对本例用节 2.8 的多变量 RELS 算法估值收敛性如图 2.10.3 所示.可看到对估计 MA 模型参数而言,二段 RLS – LS 算法的精度高于 RELS 算法的精度.为了便于比较,在 $t =$ 3 000处两种算法估值分别为

$$\hat{\boldsymbol{D}}_1 = \begin{bmatrix} -0.714\,877 & 0.168\,900 \\ 0.360\,563 & -1.124\,661 \end{bmatrix},\quad \hat{\boldsymbol{D}}_2 = \begin{bmatrix} 0.131\,279 & -0.098\,294 \\ -0.235\,128 & 0.313\,231 \end{bmatrix},$$

$$\hat{\boldsymbol{Q}}_\varepsilon = \begin{bmatrix} 0.359\,733 & 0.007\,832 \\ 0.007\,832 & 0.474\,579 \end{bmatrix} \tag{2.10.22}$$

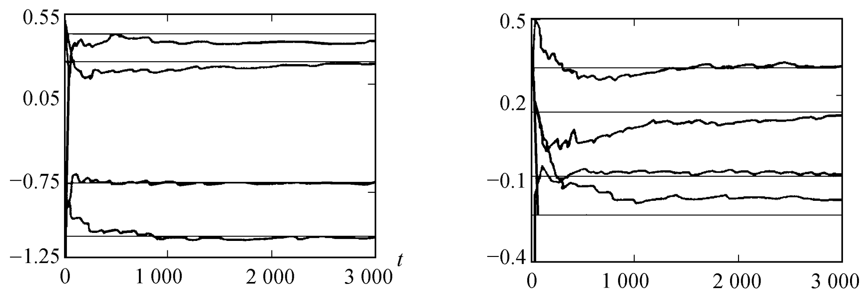
$$\hat{\boldsymbol{D}}_1=\begin{bmatrix}-0.712\,408 & 0.169\,195\\ 0.326\,461 & -1.114\,963\end{bmatrix},\quad \hat{\boldsymbol{D}}_2=\begin{bmatrix}0.107\,421 & -0.085\,795\\ -0.183\,257 & 0.305\,961\end{bmatrix},$$

$$\hat{\boldsymbol{Q}}_\varepsilon=\begin{bmatrix}0.375\,513 & 0.004\,350\\ 0.004\,350 & 0.486\,715\end{bmatrix} \tag{2.10.23}$$



(a) D_1和$\hat{D}_1(t)$ (b) D_2和$\hat{D}_2(t)$

图 2.10.2 多维两段 RLS－LS 算法 MA 参数估计的收敛性



(a) D_1和$\hat{D}_1(t)$ (b) D_2和$\hat{D}_2(t)$

图 2.10.3 多维 RELS 算法 MA 参数估计的收敛性

2.11 偏差补偿递推最小二乘(BCRLS)法

带观测噪声的 AR 模型参数估计问题出现在许多应用领域,例如语音增强[25~27].在噪声环境下说话,例如长途电话语音增强问题是过滤噪音,求真实语音信号的最优滤波器.为此需要由被噪声污染的观测信号估计原始语言信号(例如用 AR 模型描写)的模型参数.观测噪声(环境噪声或背景噪声)模型参数可以认为是已知的.因为在说话间断期间人们可检测(观测)到噪声信号真实值,从而可对噪声信号建模,得到噪声模型参数.

周知[5],当 AR 模型的输入噪声是白噪声时,用普通递推最小二乘法可得到 AR 参数的一致估计,但输入噪声为有色噪声时,用普通递推最小二乘法将引出有偏的 AR 参数估计。为此,文献[28,29]提出了带白色观测噪声的 AR 模型参数估计的偏差补偿最小二乘(Bias Compensated Least－Squares)法.作者在文献[7,30]提出了带有色观测噪声的 AR 模型参数估计的偏差补偿最小二乘法(BCLS),并严格证明了它的收敛性,推广了文献[28,29]的结果.本节介绍上述结果,并给出了仿真例子.

2.11.1 带白色观测噪声的 AR 模型的 BCRLS 参数估计

考虑带白色观测噪声的平稳 AR 模型

$$y(t) = a_1 y(t-1) + \cdots + a_n y(t-n) + \varepsilon(t) \tag{2.11.1}$$

$$z(t) = y(t) + v(t) \tag{2.11.2}$$

其中 $y(t)$为真实信号，$z(t)$为观测信号，$v(t)$为零均值、方差为 σ_v^2 的白色观测噪声，且独立于带零均值、方差为 σ_ε^2 的白噪声 $\varepsilon(t)$. 已知 σ_v^2，但 a_i 和 σ_ε^2 未知，问题是基于观测 $(z(1),\cdots z(t))$求 AR 参数估值 $\hat{a}_i(t)$和 $\hat{\sigma}_\varepsilon^2(t)$. 将(2.11.2)代入(2.11.1)有 LS 结构

$$z(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} + m(t) \tag{2.11.3}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t) = [z(t-1),\cdots,z(t-n)] \tag{2.11.4}$$

$$\boldsymbol{\theta} = [a_1, a_2,\cdots a_n]^{\mathrm{T}} \tag{2.11.5}$$

$$m(t) = \varepsilon(t) + v(t) - a_1 v(t-1) - \cdots - a_n v(t-n) \tag{2.11.6}$$

其中 $m(t)$是有色噪声. 于是用普通 RLS 可得有偏的 $\boldsymbol{\theta}$ 的估值$\hat{\boldsymbol{\theta}}_b(t)$为

$$\hat{\boldsymbol{\theta}}_b(t) = \hat{\boldsymbol{\theta}}_b(t-1) + \frac{\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)[z(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_b(t-1)]}{1 + \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)} \tag{2.11.7}$$

$$\boldsymbol{P}(t) = \boldsymbol{P}(t-1) - \frac{[\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)][\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)]^{\mathrm{T}}}{1 + \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)} \tag{2.11.8}$$

$$\hat{\boldsymbol{\theta}}_b(t) = \boldsymbol{0},\quad \boldsymbol{P}(0) = a\boldsymbol{I},\quad a > 0 \tag{2.11.9}$$

现在我们分析 RLS 估值 $\hat{\boldsymbol{\theta}}_b(t)$的渐近偏差. 应用(2.1.17)相应于(2.11.7)和(2.11.8)的非递推 LS 估值的正规方程为

$$\left[\sum_{j=1}^{t}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\right]\hat{\boldsymbol{\theta}}_b(t) = \sum_{j=1}^{t}\boldsymbol{\varphi}(j)z(j) \tag{2.11.10}$$

将(2.11.3)代入(2.11.10)有

$$\sum_{j=1}^{t}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)[\hat{\boldsymbol{\theta}}_b(t) - \boldsymbol{\theta}] = \sum_{j=1}^{t}\boldsymbol{\varphi}(j)m(j) \tag{2.11.11}$$

为了应用平稳随机序列的遍历性，上式两边除以 t 有

$$\frac{1}{t}\sum_{j=1}^{t}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)[\hat{\boldsymbol{\theta}}_b(t) - \boldsymbol{\theta}] = \frac{1}{t}\sum_{j=1}^{t}\boldsymbol{\varphi}(j)m(j) \tag{2.11.12}$$

由对 $y(t)$的平稳性假设，当 $t\to\infty$时以概率 1 有

$$\frac{1}{t}\sum_{j=1}^{t}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j) \to \boldsymbol{M} = \mathrm{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(\mathrm{t})] \tag{2.11.13}$$

注意，由(2.11.2)，(2.11.4)和(2.11.6)有

$$\frac{1}{t}\sum_{j=1}^{t}\boldsymbol{\varphi}(j)m(j) = \frac{1}{t}\sum_{j=1}^{t}\begin{bmatrix} y(j-1) + v(j-1) \\ \vdots \\ y(j-n) + v(j-n)\end{bmatrix}\times$$
$$[\varepsilon(j) + v(j) - a_1 v(j-1) - \cdots - a_n v(j-n)] \tag{2.11.14}$$

由白噪声 $\varepsilon(t)$与 $v(t)$独立及 $y(t)$的遍历性，并注意 $y(t)$可展为

$$y(t) = \sum_{j=0}^{\infty}\psi_j\varepsilon(t-j) \tag{2.11.15}$$

其中系数 ψ_j 由(1.4.20)递推计算为

$$\psi_j = a_1\psi_{j-1} + \cdots + a_n\psi_{j-n}, \quad j > 0 \tag{2.11.16}$$

其中规定 $\psi_0 = 1$，$\psi_j = 0\ (j > 0)$，当 $t \to \infty$ 时我们有关系[8]

$$\frac{1}{t}\sum_{j=1}^{t} y(j-i)\varepsilon(j) \to \mathrm{E}[y(t-i)\varepsilon(t)] = 0 \quad (i = 1,\cdots,n),$$

$$\frac{1}{t}\sum_{j=1}^{t} v(j-i)\varepsilon(j) \to \mathrm{E}[v(t-i)\varepsilon(t)] = 0 \quad (i = 1,\cdots,n),$$

$$\frac{1}{t}\sum_{j=1}^{t} v(j-i)v(j-i) \to \mathrm{E}[v^2(j-i)] = \sigma_v^2 \quad (i = 1,\cdots,n),$$

$$\frac{1}{t}\sum_{j=1}^{t} v(j-i)v(j-k) \to \mathrm{E}[v(t-i)v(t-k)] = 0 \quad (i \neq k) \tag{2.11.17}$$

$$\frac{1}{t}\sum_{j=1}^{t} y(j-i)v(j-k) \to \mathrm{E}[y(t-i)v(t-k)] = 0 \quad (\forall\, i,k)$$

于是由(2.11.14)和(2.11.17)引出，当 $t \to \infty$ 有

$$\frac{1}{t}\sum_{j=1}^{t} \boldsymbol{\varphi}(j)m(j) \to -\sigma_v^2\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = -\sigma_v^2\boldsymbol{\theta} \tag{2.11.18}$$

因而由(2.11.12)和(2.11.13)引出渐近偏差为

$$\lim_{t\to\infty}\hat{\boldsymbol{\theta}}_b(t) - \boldsymbol{\theta} = -\sigma_v^2\boldsymbol{M}^{-1}\boldsymbol{\theta} \tag{2.11.19}$$

由 $\boldsymbol{P}(t)$ 的定义(2.1.26)，当 $t \to \infty$ 有

$$t\boldsymbol{P}(t) = \left[\frac{1}{t}\sum_{j=1}^{t}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\right]^{-1} \to \boldsymbol{M}^{-1} \tag{2.11.20}$$

由(2.11.19)有

$$\boldsymbol{\theta} = \lim_{t\to\infty}\hat{\boldsymbol{\theta}}_b(t) + \sigma_v^2\boldsymbol{M}^{-1}\boldsymbol{\theta} \tag{2.11.21}$$

在上式中当 t 充分大，可近似用 $t\boldsymbol{P}(t)$ 代替 $\boldsymbol{M}^{-1}$，上式右边的 $\boldsymbol{\theta}$ 可用估值 $\hat{\boldsymbol{\theta}}(t-1)$ 近似代替，于是可得偏差补充递推最小二乘估值(BCRLS)为

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}_b(t) + \sigma_v^2 t\boldsymbol{P}(t)\hat{\boldsymbol{\theta}}(t-1) \tag{2.11.22}$$

【定理 2.11.1】 (BCRLS 算法)带白色观测噪声 AR(n)模型(2.11.1)和(2.11.2)，偏差补偿递推最小二乘参数估值 $\hat{\boldsymbol{\theta}}(t)$ 为

$$\hat{\boldsymbol{\theta}}_b(t) = \hat{\boldsymbol{\theta}}_b(t-1) + \frac{\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)[z(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_b(t-1)]}{1 + \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)} \tag{2.11.23}$$

$$\boldsymbol{P}(t) = \boldsymbol{P}(t-1) - \frac{[\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)][\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)]^{\mathrm{T}}}{1 + \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)} \tag{2.11.24}$$

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}_b(t) + \sigma_v^2 t\boldsymbol{P}(t)\hat{\boldsymbol{\theta}}(t-1) \tag{2.11.25}$$

带初值 $\hat{\boldsymbol{\theta}}_b(0) = \boldsymbol{0}$，$\boldsymbol{P}(0) = a\boldsymbol{I}$，$a > 0$，$\hat{\boldsymbol{\theta}}(0) = \boldsymbol{0}$.

可证明：[31,32]当 $t \to \infty$ 时，以概率 1 有 $\hat{\boldsymbol{\theta}}(t) \to \boldsymbol{\theta}$.

这里提出未知 σ_ε^2 的一种估计值算法. 由(2.11.3)

$$m(t)=z(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} \tag{2.11.26}$$

则有 $m(t)$为平稳随机过程,且有关系

$$m(t)=\varepsilon(t)+v(t)-a_1v(t-1)-\cdots-a_nv(t-n) \tag{2.11.27}$$

计算上式两边随机过程的方差,并注意 $\varepsilon(t)$与 $v(t)$独立,则有

$$\sigma_\varepsilon^2=\sigma_m^2-\sigma_v(1+a_1^2+\cdots a_n^2) \tag{2.11.28}$$

其中定义 $m(t)$的方差 σ_m^2 的采样估值为

$$\hat{\sigma}_m^2(t)=\frac{1}{t}\sum_{j=1}^{t}\hat{m}^2(j),\quad \hat{m}(j)=z(j)-\boldsymbol{\varphi}^{\mathrm{T}}(j)\hat{\boldsymbol{\theta}}(j) \tag{2.11.29}$$

则有递推公式

$$\hat{\sigma}_m^2(t)=\hat{\sigma}_m^2(t-1)+\frac{\hat{m}_m^2(t)-\hat{\sigma}_m^2(t-1)}{t} \tag{2.11.30}$$

将 $\hat{\sigma}_m^2(t)$和估值 $\hat{a}_i^2(t)$代入(2.11.28)有在时刻 t 的估值

$$\hat{\sigma}_\varepsilon^2(t)=\hat{\sigma}_m^2(t)-\sigma_v^2[1+\hat{a}_1^2(t)+\cdots+\hat{a}_n^2(t)] \tag{2.11.31}$$

2.11.2 带有色观测噪声的 AR 模型的 BCRLS 参数估计

考虑带零均值的有色观测噪声 $v(t)$的 AR 模型(2.11.1)和(2.11.2),其中 $v(t)$是平稳随机过程,它的相关函数 $R_v(i)=\mathrm{E}[v(t)v(t-i)]$是已知的,且可表为均方收敛的级数

$$v(t)=\sum_{j=0}^{\infty}\beta_jw(t-j),\quad \beta_0=1,\quad \sum_{j=0}^{\infty}\beta^2<\infty \tag{2.11.32}$$

其中 $w(t)$是零均值、方差为 σ_w^2 的独立于$\varepsilon(t)$的白噪声.条件(2.11.32)是很宽的,因为由定理 1.4.1 用 ARMA 模型描写的平稳随机过程 $v(t)$均可展为(2.11.32).这里不需要$v(t)$的具体表达式(2.11.32),只需已知其相关函数 $R_v(i)$.显然 $v(t)$与 $\varepsilon(t)$和 $y(t)$是不相关的.问题是由观测$(z(1),\cdots,z(t))$求 AR 参数估值 $\hat{a}_i(t)$及 $\hat{\sigma}_\varepsilon^2(t)$.

由(2.11.3)有带偏差的 $\boldsymbol{\theta}$ 的 RLS 估值 $\hat{\boldsymbol{\theta}}_b(t)$为(2.11.7)和(2.11.8).

下面来求 RLS 估值 $\hat{\boldsymbol{\theta}}_b(t)$的渐近偏差.由(2.1.17)递推 LS 估值 $\hat{\boldsymbol{\theta}}_b(t)$的正规方程为

$$\sum_{j=1}^{t}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\hat{\boldsymbol{\theta}}_b(t)=\sum_{j=1}^{t}\boldsymbol{\varphi}(j)z(j) \tag{2.11.33}$$

将(2.11.3)代入(2.11.33)后两边除以 t 有

$$\frac{1}{t}\sum_{j=1}^{t}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)(\hat{\boldsymbol{\theta}}_b(t)-\boldsymbol{\theta})=\frac{1}{t}\sum_{j=1}^{t}\boldsymbol{\varphi}(j)m(j) \tag{2.11.34}$$

将(2.11.6)代入上式,并注意在 $\boldsymbol{\varphi}(j)$中的 $z(j-i)=y(j-i)+v(j-i)$,则有(2.11.14)式成立.应用平稳随机序列的遍历性及(2.11.14)和(2.11.32),当 $t\to\infty$ 时有[8]

$$\frac{1}{t}\sum_{j=1}^{t}y(j-i)\varepsilon(j)\to\mathrm{E}[y(t-i)\varepsilon(t)]=0\qquad(i=1,\cdots,n),$$

$$\frac{1}{t}\sum_{j=1}^{t}v(j-i)\varepsilon(j)\to\mathrm{E}[v(t-i)\varepsilon(t)]=0\qquad(i=1,\cdots,n),$$

$$\frac{1}{t}\sum_{j=1}^{t}v(j-i)v(j-k)\to\mathrm{E}[v(t-i)v(t-k)]=R_v(k-i)=R_v(i-k)$$

$$\frac{1}{t}\sum_{j=1}^{t}y(j-i)v(j-k)\to\mathrm{E}[y(t-i)v(t-k)]=0\quad(\forall\ i,k) \tag{2.11.35}$$

应用(2.11.35)和(2.11.14)，当 $t\to\infty$ 有

$$\frac{1}{t}\sum_{j=1}^{t}\boldsymbol{\varphi}(j)m(j)\to\begin{bmatrix}R_v(1)\\ \vdots\\ R_v(n)\end{bmatrix}-\begin{bmatrix}R_v(0), & \cdots, & R_v(n-1)\\ \vdots & & \\ R_v(n-1), & \cdots, & R_v(0)\end{bmatrix}\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}\tag{2.11.36}$$

又由(2.11.2)有 $z(t)$ 也为平稳随机序列，故由遍历性有

$$\frac{1}{t}\sum_{j=1}^{t}\boldsymbol{\varphi}(j)\boldsymbol{\varphi}^{\mathrm{T}}(j)\to\mathrm{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)]=\boldsymbol{M}\tag{2.11.37}$$

当 $t\to\infty$ 时，由上两式和(2.11.34)引出极限关系

$$\boldsymbol{M}(\lim_{t\to\infty}\hat{\boldsymbol{\theta}}_b(t)-\boldsymbol{\theta})=\boldsymbol{r}-\boldsymbol{R}\boldsymbol{\theta}\tag{2.11.38}$$

其中定义

$$\boldsymbol{r}=\begin{bmatrix}R_v(1)\\ \vdots\\ R_v(n)\end{bmatrix},\quad \boldsymbol{R}=\begin{bmatrix}R_v(0), & \cdots, & R_v(n-1)\\ \vdots & & \vdots\\ R_v(n-1), & \cdots, & R_v(0)\end{bmatrix},\quad \boldsymbol{\theta}=\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}\tag{2.11.39}$$

因而有渐近偏差

$$\lim_{t\to\infty}\hat{\boldsymbol{\theta}}_b(t)-\boldsymbol{\theta}=\boldsymbol{M}^{-1}[\boldsymbol{r}-\boldsymbol{R}\boldsymbol{\theta}]\tag{2.11.40}$$

这引出关系

$$\boldsymbol{\theta}=\lim_{t\to\infty}\hat{\boldsymbol{\theta}}_b(t)-\boldsymbol{M}^{-1}\boldsymbol{r}+\boldsymbol{M}^{-1}\boldsymbol{R}\boldsymbol{\theta}\tag{2.11.41}$$

应用(2.11.20)，当 t 充分大用 $t\boldsymbol{P}(t)$ 近似代替 $\boldsymbol{M}^{-1}$，用 $\hat{\boldsymbol{\theta}}(t-1)$ 近似代替(2.11.41)右边的 $\boldsymbol{\theta}$，可得偏差补偿递推最小二乘估值为

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}_b(t)-t\boldsymbol{P}(t)\boldsymbol{r}+t\boldsymbol{P}(t)\boldsymbol{R}\hat{\boldsymbol{\theta}}(t-1)\tag{2.11.42}$$

这引出以下定理.

【定理 2.11.2】 (推广的 BCRLS 算法)带平稳有色观测噪声(2.11.32)的 AR(n)模型(2.11.1)和(2.11.2)，偏差补偿递推最小二乘(BCRLS)参数估值 $\hat{\boldsymbol{\theta}}(t)$ 为

$$\hat{\boldsymbol{\theta}}_b(t)=\hat{\boldsymbol{\theta}}_b(t-1)+\frac{\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)[z(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_b(t-1)]}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)}\tag{2.11.43}$$

$$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\frac{[\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)][\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)]^{\mathrm{T}}}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\boldsymbol{\varphi}(t)}\tag{2.11.44}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}_b(t)-t\boldsymbol{P}(t)\boldsymbol{r}+t\boldsymbol{P}(t)\boldsymbol{R}\hat{\boldsymbol{\theta}}(t-1)\tag{2.11.45}$$

带初值 $\hat{\boldsymbol{\theta}}_b(0)=\mathbf{0}$，$\boldsymbol{P}(0)=a\boldsymbol{I}$，$a>0$，$\hat{\boldsymbol{\theta}}(0)=\mathbf{0}$，$a=10^5$.

可证明:[31,32]当 $t\to\infty$ 时，以概率 1 有 $\hat{\boldsymbol{\theta}}(t)\to\boldsymbol{\theta}$.

【推论 2.11.1】 在白色观测噪声下，定理 2.11.1 成立.

证明 在白色观测噪声 $v(t)$ 下，有 $\boldsymbol{r}=0$，$\boldsymbol{R}=\sigma_v^2\boldsymbol{I}$，于是由(2.11.42)引出(2.11.22). 证毕. □

下面给出 σ_ε^2 的估值公式 $\sigma_\varepsilon^2(t)$. 当 $v(t)$ 为有色噪声时，计算(2.11.27)两边随机过程的方差引出

$$\sigma_\varepsilon^2=\sigma_m^2-\sum_{i,j=0}^{n}a_ia_jR_v(j-i)\tag{2.11.46}$$

将估值 $\hat{\sigma}_m^2(t)$，$\hat{a}_i(t)$代入上式可得估值 $\hat{\sigma}_\varepsilon^2(t)$.

【例 2.11.1】 考虑带有色观测噪声的 AR 模型

$$(1-1.3q^{-1}+0.4q^{-2})y(t)=\varepsilon(t) \tag{2.11.47}$$

$$z(t)=y(t)+v(t) \tag{2.11.48}$$

$$(1-0.3q^{-1})v(t)=\xi(t) \tag{2.11.49}$$

其中 $\varepsilon(t)$和 $\xi(t)$是零均值、方差各为 $\sigma_\varepsilon^2=1$，$\sigma_\xi^2=0.02$ 的独立的高斯白噪声，$a_1=1.3$，$a_2=0.4$ 和 $\sigma_\varepsilon^2=1$ 未知，但 $\sigma_\xi^2=0.02$，$p_1=-0.3$ 已知，试求 BCRLS 参数估计.

由(2.11.49)可求得有色观测噪声 $v(t)$的相关函数 $R_v(i)=\mathrm{E}[v(t)v(t-i)]$满足差分方程

$$R_v(i)=0.3R_v(i-1) \tag{2.11.50}$$

这引出

$$R_v(i)=(0.3)^iR_v(0) \tag{2.11.51}$$

而由(2.11.49)可知

$$R_v(0)=\sigma_v^2=\frac{\sigma_\xi^2}{1-(0.3)^2}=0.274\,725\,2 \tag{2.11.52}$$

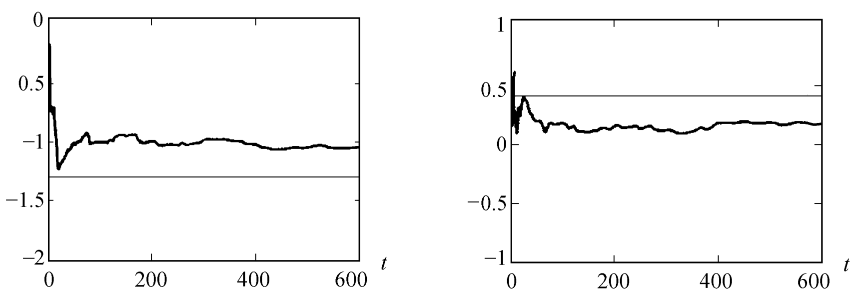
这是因为由(2.11.49) $v(t)$有展式

$$v(t)=\sum_{j=0}^{\infty}(0.3)^j\xi(t-j) \tag{2.11.53}$$

故有

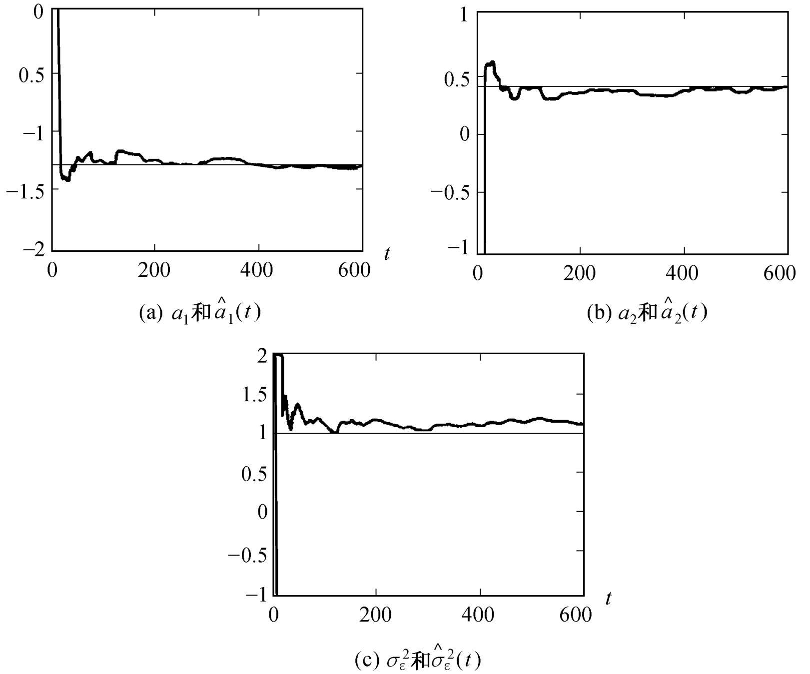
$$R_v(0)=\mathrm{E}[v^2(t)]=\sum_{j=0}^{\infty}(0.3)^{2j}\sigma_\xi^2=\frac{\sigma_\xi^2}{1-(0.3)^2} \tag{2.11.54}$$

由定理 2.11.2 可求得 BCRLS 参数估计，仿真结果如图 2.11.1 和图 2.11.2 所示，其中直线代表未知参数真实值，曲线代表估值.图 2.11.1 为应用普通 RLS 算法(2.11.43)和(2.11.44)的有偏参数估计 $\hat{a}_{b1}(t)$和 $\hat{a}_{b2}(t)$.图 2.11.2 为 BCRLS 参数估计 $\hat{a}_i(t)$及方差估计值 $\hat{\sigma}_\varepsilon^2(t)$.可见后者消除了偏差.



(a) a_1和$\hat{a}_{b1}(t)$　　(b) a_2和$\hat{a}_{b2}(t)$

图 2.11.1　普通 RLS 有偏参数估计



(a) a_1和$\hat{a}_1(t)$　　(b) a_2和$\hat{a}_2(t)$

(c) σ_ε^2和$\hat{\sigma}_\varepsilon^2(t)$

图 2.11.2　偏差补偿递推最小二乘法(BCRLS)参数估计的收敛性

2.12　带有色观测噪声的 AR 模型参数估计的 RELS 算法

考虑带 ARMA 有色观测噪声的 AR 模型参数估计问题:

$$A(q^{-1})y(t)=\varepsilon(t) \tag{2.12.1}$$

$$z(t)=y(t)+v(t) \tag{2.12.2}$$

$$P(q^{-1})v(t)=R(q^{-1})\xi(t) \tag{2.12.3}$$

其中 $z(t)$为观测,$y(t)$为输出,$v(t)$为 ARMA 有色观测噪声,$\varepsilon(t)$和 $\xi(t)$是带零均值、方差各为 σ_ε^2 和σ_ξ^2 的独立白噪声,$A(q^{-1})$,$P(q^{-1})$和 $R(\varepsilon^{-1})$是稳定的,且有形式

$$X(q^{-1})=x_0+x_1q^{-1}+\cdots+x_{n_x}q^{-n_x} \tag{2.12.4}$$

带 $a_0=1$,$p_0=1$,$r_0=1$. 已知系数 p_i,r_i 和σ_ξ^2. 问题是由观测($z(1)$,…,$z(t)$)求估值 $\hat{a}_i(t)$和 $\sigma_\varepsilon^2(t)$.

将(2.12.1)和(2.12.3)代入(2.12.2)可得 ARMA 新息模型

$$A(q^{-1})m(t)=D(q^{-1}e(t) \tag{2.12.5}$$

$$m(t)=P(q^{-1})z(t) \tag{2.12.6}$$

$$D(q^{-1})e(t)=P(q^{-1})\varepsilon(t)+A(q^{-1})R(q^{-1})\xi(t) \tag{2.12.7}$$

应用节 2.2 的 RELS 算法辨识 ARMA 模型(2.12.5)可得估值 $\hat{a}_i(t)$,$\hat{d}_i(t)$及 $\hat{\sigma}_e^2(t)$. 现在来求估值 $\hat{\sigma}_\varepsilon^2(t)$. 记 $\Psi(q^{-1})=A(q^{-1})R(q^{-1})$,则 $n_\psi=n_a+n_r$,且它的系数为 $\psi_0=1$,且

$$\psi_i=\sum_{i=0}^{j}a_ir_{j-i} \tag{2.12.8}$$

计算(2.12.7)两边 MA 过程的方差有

$$\sigma_\varepsilon^2 = \frac{\left[\sigma_e^2 \sum_{j=0}^{n_d} d_j^2 - \sigma_\xi^2 \sum_{j=0}^{n_\psi} \psi_j^2\right]}{\sum_{i=0}^{n_p} p_i^2} \tag{2.12.9}$$

将估值 $\hat{a}_i(t)$，$\hat{d}_i(t)$和 $\hat{\sigma}_e^2(t)$代入(2.12.8)和(2.12.9)可得估值 $\hat{\sigma}_\varepsilon^2(t)$.

【例 2.12.1】 考虑带有色观测噪声的 AR 模型

$$(1+1.2q^{-1}+0.35q^{-1})y(t)=\varepsilon(t) \tag{2.12.10}$$

$$z(t)=y(t)+v(t) \tag{2.12.11}$$

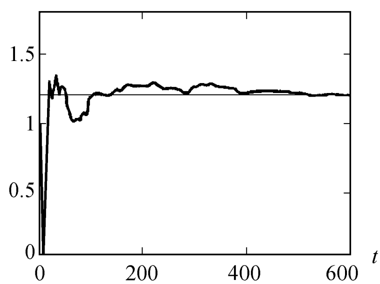
$$(1-0.2q^{-1})v(t)=(1-0.1q^{-1})\xi(t) \tag{2.12.12}$$

其中 $\varepsilon(t)$和 $\xi(t)$是零均值、方差各为 $\sigma_\varepsilon^2=1$，$\sigma_\xi^2=0.1$ 的独立的高斯白噪声，$a_1=1.2$，$a_2=0.35$，$\sigma_\varepsilon^2=1$ 未知，但 $p_1=-0.2$，$q_1=-0.1$，$\sigma_\xi^2=0.1$ 已知. 试求 AR 参数 a_1，a_2 和 σ_ε^2 的 RELS 估计.

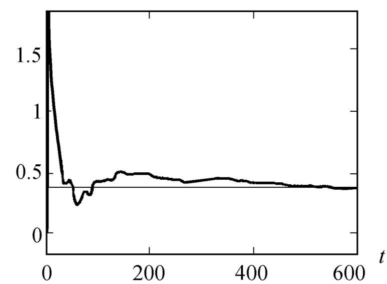
用本节方法的 RELS 参数估计收敛性如图 2.12.1 所示，其中直线表示参数真实值，曲线表示估值，在 $t=600$ 处参数估值为

$$\hat{a}_1=1.205\,767,\quad \hat{a}_2=0.372\,446,\quad \hat{\sigma}_\varepsilon^2=1.041\,246 \tag{2.12.13}$$

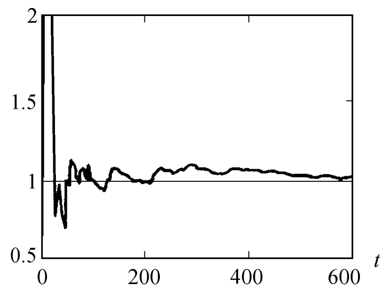
它们近似等于相应的真实值.



(a) a_1和$\hat{a}_1(t)$



(b) a_2和$\hat{a}_2(t)$



(c) σ_ε^2和$\hat{\sigma}_\varepsilon^2(t)$

图 2.12.1　AR 参数的 RELS 估计的收敛性

2.13　求 MA 模型参数的 Gevers – Wouters 算法

考虑可逆的 MA(n_d)过程

$$r(t)=\varepsilon(t)+d_1\varepsilon(t-1)+\cdots+d_{n_d}\varepsilon(t-n_d) \tag{2.13.1}$$

其中 $\varepsilon(t)$是零均值、方差为 σ_ε^2 的白噪声：

$$\mathrm{E}[\varepsilon(t)]=0,\quad \mathrm{E}[\varepsilon(t)\varepsilon(j)]=\sigma_\varepsilon^2\delta_{tj} \tag{2.13.2}$$

设初始时刻 $t_0=-\infty$，则 $r(t)$是一个平稳随机序列. 已知 $r(t)$的阶次数 n_d 和相关函数

$$R_r(k)=\mathrm{E}[r(t)r(t-k)],\quad k=0,1,\cdots,n_d \tag{2.13.3}$$

$$R_r(k)=0,\quad k>n_d \tag{2.13.4}$$

问题是由 $R_r(k)$求 MA 参数 $d_1,\cdots,d_{n_d}$和σ_ε^2.

启发性分析如下：

关键思想是：稳态关系可通过非稳态关系取极限来实现. 从而实现了稳态与非稳态的转化，平稳和非平稳的转化. 为此，先从分析稳态关系入手.

由(2.13.1)和(2.13.2)有

$$\mathrm{E}[r(t)\varepsilon(t-k)]=d_k\sigma_\varepsilon^2 \tag{2.13.5}$$

这引出

$$d_k=\frac{\mathrm{E}[r(t)\varepsilon(t-k)]}{\sigma_\varepsilon^2} \tag{2.13.6}$$

注意

$$\mathrm{E}[r(t)\varepsilon(t)]=\mathrm{E}[\varepsilon(t)\varepsilon(t)]=\sigma_\varepsilon^2 \tag{2.13.7}$$

引入记号

$$R_{r\varepsilon}(t,t-k)=\mathrm{E}[r(t)\varepsilon(t-k)] \tag{2.13.8}$$

$$R_\varepsilon(t,t)=\mathrm{E}[\varepsilon(t)\varepsilon(t)] \tag{2.13.9}$$

则有稳态关系

$$d_k=R_{r\varepsilon}(t,t-k)/\sigma_\varepsilon^2,\quad k=1,\cdots,n_d \tag{2.13.10}$$

$$\sigma_\varepsilon^2=R_{r\varepsilon}(t,t)=R_\varepsilon(t,t) \tag{2.13.11}$$

注意 $t_0=-\infty$，因此 $R_{r\varepsilon}(t,t-k)$，$R_{r\varepsilon}(t,t)$与时间 t 无关系.

由(2.13.1)有

$$r(t-k)=\varepsilon(t-k)+d_1\varepsilon(t-k-1)+\cdots+d_{n_d}\varepsilon(t-k-n_d) \tag{2.13.12}$$

$$\varepsilon(t-k)=r(t-k)-d_1\varepsilon(t-k-1)-\cdots-d_{n_d}\varepsilon(t-k-n_d) \tag{2.13.13}$$

这引出

$$\mathrm{E}[r(t-k)\varepsilon(t-k-i)]=d_i\mathrm{E}[\varepsilon(t-k-i)\varepsilon(t-k-i)] \tag{2.13.14}$$

$$d_i=\frac{R_{r\varepsilon}(t-k,t-k-i)}{R_\varepsilon(t-k-i,t-k-i)},\quad i=1,\cdots,n_d \tag{2.13.15}$$

由(2.13.13)和(2.13.15) $R_{r\varepsilon}(t,t-k)$有递推公式

$$R_{r\varepsilon}(t,t-k)=\mathrm{E}[r(t)\varepsilon(t-k)]=$$

$$\mathrm{E}\left\{r(t)\left[r(t-k)-\sum_{i=1}^{n_d}\frac{R_{r\varepsilon}(t-k,t-k-i)}{R_\varepsilon(t-k-i,t-k-i)}\varepsilon(t-k-i)\right]\right\}=$$

$$R_r(k)-\sum_{i=1}^{n_d-k}R_{r\varepsilon}(t,t-k-i)R_{r\varepsilon}(t-k,t-k-i)/R_\varepsilon(t-k-i,t-k-i)=$$

$$R_r(k)-\sum_{s=k+1}^{n_d} R_{r\varepsilon}(t,t-s)R_{r\varepsilon}(t-k,t-s)/R_{\varepsilon}(t-s,t-s) \tag{2.13.16}$$

其中用到事实：当 $j>n_d$，$R_{r\varepsilon}(t,t-j)=0$.

在稳态（$t_0\to-\infty$）上述推导的 d_k 和 σ_ε^2 与时间 t 无关. 这个极限过程等价于置 $t_0=0$，令 $t\to+\infty$.

设初始时刻 $t_0=0$，并设初值

$$\varepsilon(i)=0 \quad (i<0) \tag{2.13.17}$$

则由(2.13.1)有 $r(0)=\varepsilon(0)$，于是

$$\mathrm{E}[r(0)\varepsilon(0)]=\mathrm{E}[\varepsilon(0)\varepsilon(0)]=\mathrm{E}[r(0)r(0)] \tag{2.13.18}$$

因而有初值

$$R_{r\varepsilon}(0,0)=R_{r\varepsilon}(0,0)=R_r(0) \tag{2.13.19}$$

由(2.13.17)引出规定：

$$R_{r\varepsilon}(t,t-s)=0 \quad (s>t) \tag{2.13.20}$$

$$R_{\varepsilon}^{-1}(t-s,t-s)=0 \quad (s>t) \tag{2.13.21}$$

则在上述初值下由(2.13.16)递推计算的 $R_{r\varepsilon}(t,t-k)$ 是时变的，与时间 t 有关. 为了消除初值的影响，当 $t\to+\infty$，我们由(2.13.10)和(2.13.11)有

$$d_k=\lim_{t\to+\infty}\frac{R_{r\varepsilon}(t,t-k)}{R_{r\varepsilon}(t,t)}, \quad k=1,\cdots,n_d \tag{2.13.22}$$

$$\sigma_\varepsilon^2=\lim_{t\to+\infty}R_{r\varepsilon}(t,t) \tag{2.13.23}$$

综上所述有如下定理.

【定理 2.13.1】 （Gevers－Wouters[32]算法）考虑可逆的纯量 MA(n_d)过程(2.13.1)，已知它的相关函数 $R_r(k)$，$k=0,1,\cdots,n_d$，则可用如下 Gevers－Wouters 迭代算法求 MA 参数 d_i 和 σ_ε^2：

$$\sigma_\varepsilon^2=\lim_{t\to+\infty}R_{r\varepsilon}(t,t) \tag{2.13.24}$$

$$d_i=\lim_{t\to+\infty}\frac{R_{r\varepsilon}(t,t-i)}{R_{r\varepsilon}(t,t)}, \quad i=1,\cdots,n_d \tag{2.13.25}$$

$$R_{r\varepsilon}(t,t-i) = R_r(i)-\sum_{s=i+1}^{n_d}R_{r\varepsilon}(t,t-s)R_{r\varepsilon}(t-i,t-s)/R_{r\varepsilon}(t-s,t-s) \tag{2.13.26}$$

其中 $t=0,1,2,\cdots,i=t,t-1,\cdots,0$，且规定

$$\begin{gathered} R_{r\varepsilon}(0,0)=R_r(0), \\ R_{r\varepsilon}(t,t-s)=0 \quad (t<s) \\ R_{r\varepsilon}^{-1}(t-s,t-s),=0 \quad (t<s) \end{gathered} \tag{2.13.27}$$

注意，规定(2.13.27)可保证 $R_{r\varepsilon}(t,t-i)=0(i>n_d)$.

可以证明[36]：在较弱的条件下，上述极限关系成立，且可保证 MA(n_d)过程是可逆的. 应用上述 Gevers－Woutes 算法，取迭代次数 t 充分大，则有在时刻 t 估值

$$\hat{\sigma}_\varepsilon^2(t)=R_{r\varepsilon}(t,t) \tag{2.13.28}$$

$$\hat{d}_i(t) = R_{r\varepsilon}(t, t-i)/R_{r\varepsilon}(t, t), \quad i = 1, \cdots, n_d \tag{2.13.29}$$

通常迭代次数 $t = 50 \sim 100$ 便可达到参数估计误差不超 10^{-3}的满意精度.

上述单变量 MA 参数估计的 Gevers – Wouters 算法可平行推广到多变量情形.[32]

考虑可逆的多变量 MA 过程 $\boldsymbol{r}(t)$,

$$\boldsymbol{r}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{2.13.30}$$

其中 $\boldsymbol{r}(t) \in R^m$, $\boldsymbol{\varepsilon}(t) \in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声, q^{-1}为单位滞后算子,且多项式矩阵 $\boldsymbol{D}(q^{-1}) = \boldsymbol{I}_m + \boldsymbol{D}_1 q^{-1} + \cdots + \boldsymbol{D}_{n_d} q^{-n_d}$是稳定的(即 $\det\boldsymbol{D}(x)$的零点全位于单位圆外). 记它的相关函数为 $\boldsymbol{R}_r(i) = \mathrm{E}[\boldsymbol{r}(t)\boldsymbol{r}^{\mathrm{T}}(t-i)]$, $i = 0, 1, \cdots, n_d$; $\boldsymbol{R}_r(i) = 0 \ (i > n_d)$. 已知它的相关函数 $\boldsymbol{R}_r(i)$,但 $\boldsymbol{D}(q^{-1})$和 $\boldsymbol{Q}_\varepsilon$ 是未知的,下面的多变量 Gevers – Wouters 算法[31,32]解决由它的相关函数 $\boldsymbol{R}_r(i)$如何得到参数阵 $\boldsymbol{D}_i$ 和方差阵$\boldsymbol{Q}_\varepsilon$ 的一致估计.

【定理 2.13.2】 (多维 Gevers – Wouters 算法[32])若可逆的 MA(2.13.30)的谱阵

$$\boldsymbol{S}_r(q) = \sum_{j=-n_d}^{n_d} \boldsymbol{R}_r(j) q^{-j} \tag{2.13.31}$$

或等价地

$$\boldsymbol{S}_r(q) = \boldsymbol{D}(q^{-1})\boldsymbol{Q}_\varepsilon \boldsymbol{D}^{\mathrm{T}}(q) \tag{2.13.32}$$

在单位圆上处处非异,即 $\boldsymbol{S}_r(e^{iw})$, $-\pi \leqslant \omega \leqslant \pi$,为非异阵,则用如下多维 Gevers – Wouters 算法可得到 $\boldsymbol{D}_i$ 和 $\boldsymbol{Q}_\varepsilon$ 的一致估计

$$\boldsymbol{Q}_\varepsilon = \lim_{t\to\infty} \boldsymbol{R}_{r\varepsilon}(t, t) \tag{2.13.33}$$

$$\boldsymbol{D}_i = \lim_{t\to\infty} \boldsymbol{R}_{r\varepsilon}(t, t-i)\boldsymbol{R}_{r\varepsilon}^{-1}(t, t), \quad i = 1, \cdots, n_d \tag{2.13.34}$$

$$\boldsymbol{R}_{r\varepsilon}(t, t-i) = \boldsymbol{R}_r(i) - \sum_{s=i+1}^{n_d} \boldsymbol{R}_{r\varepsilon}(t, t-s)\boldsymbol{R}_{r\varepsilon}^{-1}(t-s, t-s)\boldsymbol{R}_{r\varepsilon}^{\mathrm{T}}(t-i, t-s) \tag{2.13.35}$$

其中 $t = 0, 1, 2, \cdots$, $i = t, t-1, \cdots 0$,且规定

$$\boldsymbol{R}_{r\varepsilon}(0, 0) = \boldsymbol{R}_r(0) \tag{2.13.36}$$

$$\boldsymbol{R}_{r\varepsilon}(t, t-j) = \boldsymbol{0} \quad (t < j) \tag{2.13.37}$$

$$\boldsymbol{R}_{r\varepsilon}{}^{-1}(t-s, t-s) = \boldsymbol{0} \quad (t < s) \tag{2.13.38}$$

可保证[31]这样得到的 MA 过程(2.13.30)具有已知的相关函数 $\boldsymbol{R}_r(i)$,且 $\boldsymbol{D}(q^{-1})$是稳定的, $r(t)$是可逆的.

详细推导和证明见文献[31,32],从略.

用 MATLAB 语言编制的 Gevers – Wouters 算法的通用程序如下.

为了编制 G – W 算法的 MATLAB 程序,首先来分析一下如何计算 $\boldsymbol{R}_{re}(t, t-i)$? 当 t 和 i变化时按什么次序计算 $\boldsymbol{R}_{re}(t, t-i)$? 以下设 $n_d \geqslant 1$,由(2.13.35)及规定 $\boldsymbol{R}_{re}(t, t-j) = \boldsymbol{0} \ (t < j)$, $\boldsymbol{R}_{re}^{-1}(t-s, t-s) = \boldsymbol{0} \ (t < s)$,容易得到

$$\boldsymbol{R}_{re}(0, 0) = \boldsymbol{R}_r(0), \quad \boldsymbol{R}_{re}(1, 0) = \boldsymbol{R}_r(1),$$

$$\boldsymbol{R}_{re}(1, 1) = \boldsymbol{R}_r(0) - \boldsymbol{R}_{re}(1, 0)\boldsymbol{R}_{re}^{-1}(0, 0)\boldsymbol{R}_{re}^{\mathrm{T}}(1, 0) \tag{2.13.39}$$

因此计算 $\boldsymbol{R}_{re}(t, t-i)$的次序为

$$\boldsymbol{R}_{re}(0, 0)$$

$$
\begin{aligned}
&\boldsymbol{R}_{re}(1,0),\boldsymbol{R}_{re}(1,1)\\
&\boldsymbol{R}_{re}(2,0),\boldsymbol{R}_{re}(2,1),\boldsymbol{R}_{re}(2,2)\\
&\qquad\vdots\\
&\boldsymbol{R}_{re}(t,0)\boldsymbol{R}_{re}(t,1),\cdots,\boldsymbol{R}_{re}(t,t)
\end{aligned}\tag{2.13.40}
$$

按上述思路可得到用 MATLAB 语言编制 G－W 算法的程序清单：以下程序为计算 Rre（t，t－1），其中下标中的字母 e 代表字母 ε.

```
    Rre(1:m,1:m) = R(:,1:m)
for t = 1:100
        Rre(t*m+1:t*m+m,1:m) = R(:,t*m+1:t*m+m);
  for i = t-1:-1:0    sum = zeros(m,m);
      for j = i+1:min(nn,t)
        sum = (sum + Rre(t*m+1:t*m+m,(t-j)*m+1:(t-j)*m+m)...
            * inv(Rre((t-j)*m+1:(t-j)*m+m,(t-j)*m+1:(t-j)*m+m))...
            * Rre((t-i)*m+1:(t-i)*m+m,(t-j)*m+1:(t-j)*m+m)';end
      Rre(t*m+1:t*m+m,(t-i)*m+1:(t-i)*m+m) = R(:,i*m+1:i*m+m) - sum;
  end
end
```

以下程序为计算 $t=100$ 步处 Q_ε 及 D 估值

```
        Qe = Rre(100*m+1:100*m+m,100*m+1:100*m+m);
        for i  = 0:nn
        D(1:m,i*m+1:i*m+m) = Rre(100*m+1:100*m+m,(100-i)*m+1:
        (100-i)*m+m)*inv(Qe);
        end
```

注意在上述程序中将阶次 n_d 记为 nn.

【例 2.13.1】 考虑 MA（3）模型

$$
\begin{aligned}
r(t)&=(1+0.5q^{-1})^3\varepsilon(t)=(1+1.5q^{-1}+0.75q^{-2}+0.125q^{-3})\varepsilon(t)=\\
&\varepsilon(t)+1.5\varepsilon(t-1)+0.75\varepsilon(t-2)+0.125\varepsilon(t-3)
\end{aligned}\tag{2.13.41}
$$

其中取 $\sigma_\varepsilon^2=1$，容易算出它的相关函数为

$$
\begin{aligned}
&R_r(0)=3.828\,125,\quad R_r(1)=2.718\,750\\
&R_r(2)=0.937\,500,\quad R_r(3)=0.125\,000
\end{aligned}\tag{2.13.42}
$$

MA 参数为

$$
d_1=1.5,\quad d_2=0.75,\quad d_3=0.125,\quad \sigma_\varepsilon^2=1 \tag{2.13.43}
$$

应用 Gevers－Wouters 算法（2.13.24）～（2.13.29）仿真结果如图 2.13.1 所示，其中迭代次数 $t=100$ 次. 可看到在 $t=20$ 以后估值 $\hat{d}_i(t)$ 和 $\hat{\sigma}_\varepsilon^2(t)$ 已几乎与相应的真实值重合. 在 $t=100$ 时，MA 参数估值为

$$
\hat{d}_1=1.500\,000,\quad \hat{d}_2=0.750\,000,\quad \hat{d}_3=0.125\,000,\quad \hat{\sigma}_\varepsilon^2=1.000\,000 \tag{2.13.44}
$$

【例 2.13.2】 考虑二维可逆的 MA（2）过程

$$
\boldsymbol{r}(t)=(\boldsymbol{I}_2+\boldsymbol{D}_1q^{-1}+\boldsymbol{D}_2q^{-2})\boldsymbol{\varepsilon}(t) \tag{2.13.45}
$$

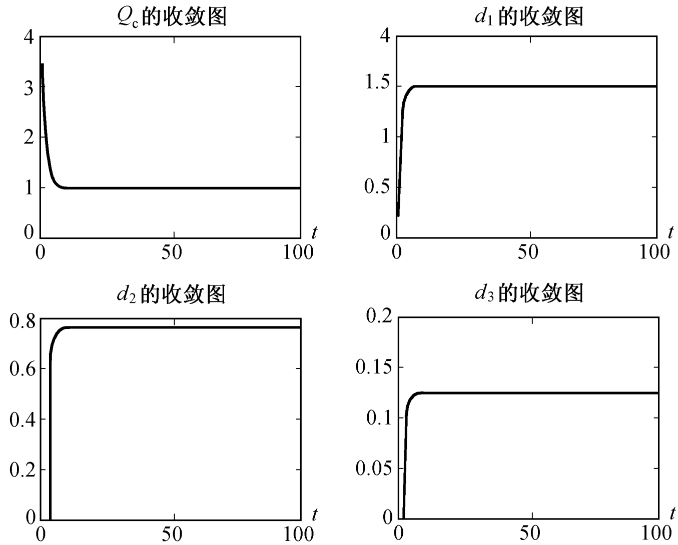


图 2.13.1 用 Gevers－Wouters 算法 MA 参数估计的收敛性

其中 $\boldsymbol{\varepsilon}(t)$是零均值、方差为 $\boldsymbol{Q}_{\varepsilon}$ 的白噪声，且

$$\boldsymbol{D}_1=\begin{bmatrix}0.2 & 0.9\\ 0.6 & -0.2\end{bmatrix},\quad \boldsymbol{D}_2=\begin{bmatrix}-0.35 & -0.45\\ 0.42 & -0.09\end{bmatrix},\quad Q_{\varepsilon}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \tag{2.13.46}$$

由 $R_r(i)$利用多维 Gevers－Wouters 算法(2.13.37)～(2.13.39)可得在 $t=100$ 处的 MA 参数估值为

$$\hat{\boldsymbol{Q}}_{\varepsilon}=\begin{bmatrix}1.000\,000 & -0.000\,000\\ -0.000\,000 & 1.000\,000\end{bmatrix},\quad \hat{\boldsymbol{D}}_1=\begin{bmatrix}0.200\,000 & 0.900\,00\\ 0.600\,000 & -0.200\,000\end{bmatrix},$$

$$\hat{\boldsymbol{D}}_2=\begin{bmatrix}-0.350\,000 & -0.450\,000\\ 0.420\,000 & -0.090\,000\end{bmatrix} \tag{2.13.47}$$

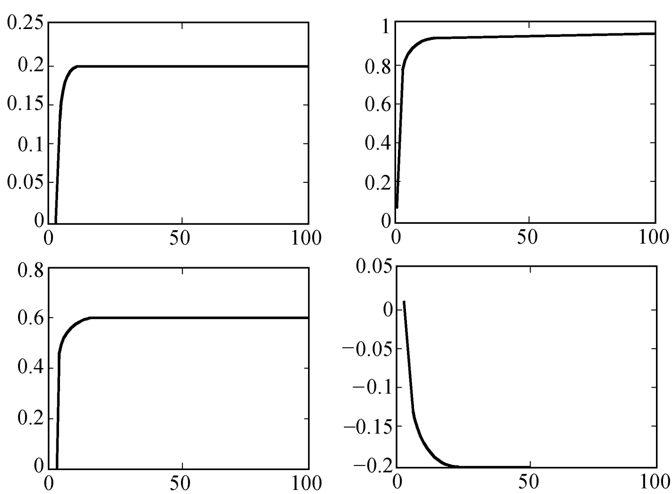


图 2.13.2 用 G－W 算法 MA 参数估值 $\hat{\boldsymbol{D}}_1$ 的收敛性

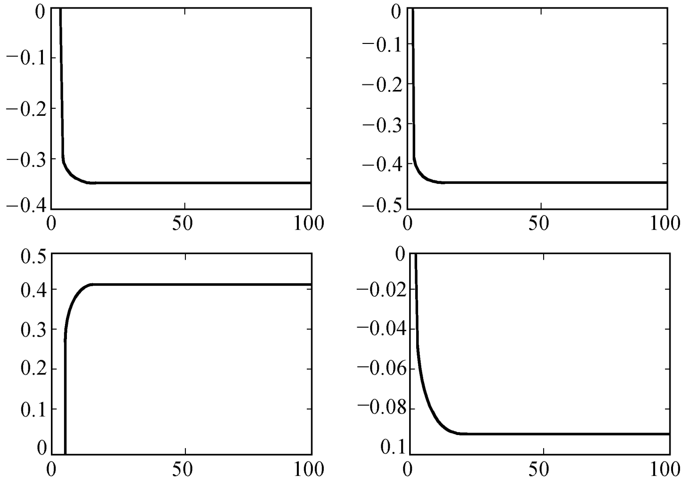


图 2.13.3　用 G－W 算法 MA 参数估值 $\hat{\boldsymbol{D}}_2$ 的收敛性

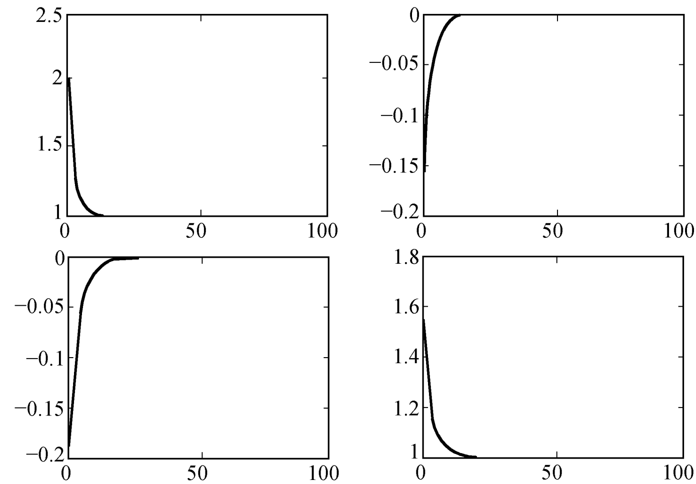


图 2.13.4　用 G－W 算法 MA 参数估值 $\hat{\boldsymbol{Q}}_\varepsilon$ 的收敛性

参　考　文　献

1　邓自立，郭一新.现代时间序列分析及其应用——建模、滤波、去卷、预报和控制.北京：知识出版社，1989

2　Sage A P，Melsa J L.估计理论及其在通讯与控制中的应用.北京：科学出版社，1983

3　徐宁寿.系统辨识技术及其应用.北京：机械工业出版社，1986

4　方崇智，萧德云.过程辨识.北京：科学出版社，1988

5　Graup D. Time Series Analysis，Identification and Adaptive filtering. Malabar，Florida：Robert E. Krieger Pulishing Company，Inc.，1984

6　Ljung L. 系统辨识——使用者的理论.袁震东，阮荣耀，陈树中译.上海：华东师范大学出版社，1990

7　邓自立，郭一新.动态系统分析及其应用——建模、滤波、预波、控制的新方法和程序库.沈阳：辽宁科学技术出版社，1985

8　Söderström T. Ergodicity Results for Sample Covariances. Problems of Control and Information Theory，1975，4：131～138

9 邓自立.卡尔曼滤波与维纳滤波——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2001
10 Ljung L. On Positive Real Transfer Function and Convergence of Some Recurisive Schemes. IEEE Trans. Automatic Control, 1977, 22:539~551
11 Kamen E W, Su J K. Introduction to Optimal Estimation. Spriger-Verlag London Berlin Heidelberg, 1999
12 邓自立,杜洪越,马建为.改进的递推增广最小二乘参数估计方法.科学技术与工程,2002,2(5):1~3
13 杜金观,项静怡,戴俭华.时间序列分析——建模与预报.合肥:安徽教育出版社,1991
14 Gregory C. Elements of Multivariate Time Series Analysis. Springer-Verlag New York Berlin Heidelberg, 1993
15 Graup D, Krause D J, Moore J B. Identification of ARMA Parameters of Time Sevies. IEEE Trans. Automatic Control, 1975,20:140~107
16 邓自立,马建为,杜洪越.ARMA模型参数估计的两段最小二乘法.科学技术与工程,2002,2(4)
17 杨虎引,马正午,孙宇.电子计算机应用.北京:冶金工业出版社,1979
18 Hetthessy J, and Keviczhy L. Minimum Variance Control of Multivariable Linear Systems. Problem of Control and Information Theory, 1977, 6:229~242
19 Ljung L. Consistency of the Least-Squares Identification Method. IEEE Trans. Automatic Control, 1976, 21(5):777~781
20 邓自立,杜洪越.带有色观测噪声ARMA模型参数估计的三段算法.科学技术与工程,2003,3(1):1~3
21 邓自立,马建为.MA模型参数估计的两段最小二乘法及其在自校正跟踪滤波器中的应用.科学技术与工程,2003,3(1):3~5
22 杜洪越,邓自立.多变量滑动平均模型参数估计两段最小二乘法.黑龙江大学自然科学学报,2003,20(2):55~58
23 邓自立,郭一新.单变量ARMAX模型结构辨识.控制理论与应用,1985,2(3):114~117
24 邓自立,郭一新.多变量CARMA模型结构辨结.自动化学报,1986,12(1):18~24
25 Gibson J D, Boneung K and Steven D G. Filtering of Colored Noise for Speech Enhancement and Coding. IEEE Trans, Signal Processing, 1991,39(8):1732~1741
26 Moir T J, Campbell D R, Dabis H S. A Polynomial Approach to Optimal and Adaptive Filtering with Application to Speech Enhacement. IEEE Trans. Signal Processing, 1991,39(5):577~598
27 Dabis H S, and Moir T J. A Unified Approach to Optimal Estimation Using Diophantine. Int. J. Control, 1993, 57(3):577~598
28 Sakai H, Arase M. Recursive Parameter Estimation of Autoregressive Process Disturbed by White Noise. Int. J. Control, 1979,30(6):949~966
29 Sagara S, and Wada K. On-Line Modified Least-Squares Estimation of Linear Dynamic Systems. Int. J. Control, 1977,25(3):329~343
30 邓自立.自回归模型补偿偏差最小二乘辨识.黑龙江大学自然科学学报,1981.2:10~20
31 邓自立.估计MA参数的多维强Gevers-Wouters算法及其在构造ARMA新息模型中的应用.控制理论与应用,2001,18(5):737~740
32 Gevers M, Wouters W R E. An Innovations Approach to Discrete-Time Stochastic Realiztion Problem. Ouartely Journal on Automatic Control, 1978, 19(2):90~110
33 邓自立.自校正滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2003
34 罗抟翼,程桂芳.随机信号处理与控制基础.北京:化学工业出版社,2002
35 Lewis F L. Optimal Estimation. New York: John Wiley & Sons, 1986
36 Deng Zi-Li. Convergence Analysis of Gevers-Wouters Algorithm for Scalar Speceral Factorization. Science Technology and Engineering, 2005,5(1):7~13

第三章　Kalman 滤波

3.1　引　　论

滤波问题是如何从被噪声污染的观测信号中过滤噪声，尽可能消除或减小噪声影响，求未知真实信号或系统状态的最优估计.在某些应用问题中甚至真实信号被噪声所淹没，滤波的目的就是过滤噪声，还真实信号本来面目.这类问题广泛出现在信号处理、通信、目标跟踪和控制等领域[6].通常噪声和真实信号或状态均为随机过程，因而滤波问题本质上是统计估计问题.常用的最优估计准则是线性最小方差估计，即要求信号或状态的最优估值应与相应的真实值的误差的方差最小.这种滤波也叫最优滤波.

处理滤波问题的简单的硬件方法是在某些系统中可近似认为信号和噪声均为由不同频率的周期函数的叠加合成的确定性信号，因而可用电阻、电容等器件组成低通滤波器等，使有用低频信号无衰减地通过，而高频噪声受到抑制.这类用硬件实现的滤波器叫模拟滤波器.它也可用集成电路芯片或用计算机通过算法实现，称为数字滤波器.

对于随机信号的统计滤波问题，用上述处理确定性信号滤波技术难以得到满意结果.因此在第二次世界大战期间，针对空防战斗的需要，Wiener 用频域方法提出了 Wiener 滤波器.它的缺点和局限性是适于处理一维平稳随机信号滤波问题，要求解维纳 – 霍普方程，计算量和存储量大，限制了其工程应用.

由于信号和噪声往往是多维非平稳随机过程，Wiener 滤波问题不能解决这类随机过程的滤波问题.因此 1960 年 Kalman[1]用时域上的状态空间方法提出了 Kalman 滤波理论，提出了便于计算机上递推实现的 Kalman 滤波算法，计算量和存储量小，克服了 Wiener 滤波理论的缺点，解决了多维非平稳随机信号的滤波问题.

Kalman 滤波在工程实践中获得了广泛的应用，例如应用于制导、全球定位系统、目标跟踪、石油地震勘探、通信和信号处理、信息融合等.

在状态空间方法中，引入了状态变量和状态空间概念.状态是比信号更广泛、更灵活的概念，它非常适合处理多变量系统，也非常适合处理信号估计问题，信号可视为状态或状态的分量.系统状态变量是能体现系统特征、特点和状况的变量.例如在目标跟踪问题中，最简单情形是可把运动目标的位置视为状态，稍复杂些也可将位置和速度两者视为状态，一般也可将位置、速度和加速度三者视为状态变量.状态变量的维数由具体问题和具体要求而定.一个 n 维状态变量的取值属于 n 维欧氏空间 R^n，即 n 维状态变量的取值是 R^n 中的“点”，称状态变量取值的欧氏空间 R^n 为状态空间.状态空间方法的关键技术包括状态空间模型和基于射影理论的状态估计方法.状态方程是描写状态变化规律的模型，它描写了相邻时刻的状态转移变化规律.观测方程描写对状态进行观测的信息，通常含有观测噪声，且通常只能对部分状态变量进行观测. Kalman 滤波问题就是由观测方程所得到的

观测信息求系统状态的最优估计.

从抽象的 Hilbert 空间角度,状态变量和观测信号均可看成是抽象的由随机变量的线性运算生成的 Hilbert 空间中的元素或“点”.因而 Kalman 滤波问题在几何上化为状态变量在由观测信号生成的子 Hilbert 空间上的射影.[22]

以上阐述的就是 Kalman 滤波方法的关键思想.为了说明 Kalman 滤波方法,以下给出三个启发性的应用例子.

【例 3.1.1】 潜艇声纳目标跟踪预报问题.[3,23]

如图 3.1.1 所示,潜艇在水下用声纳系统不断测量海面上沿 s 轴方向直线匀速运动的敌舰的位置 s,设测量位置的采样周期为 T,而敌舰运动的速度 $\dot{s}$(常数)是未知的.设 $s(t)$,$\dot{s}(t)$和 $y(t)$各为在时刻 tT 敌舰的位置、速度和位置的测量值.声纳系统可由在时刻 tT 处矢径 r 和方位角 θ 的观测值 $r_m(t)$和 $\theta_m(t)$算出位置 $s(t)$的测量值 $y(t)=r_m(t)\cos\theta_m(t)$.因此真实位置 $s(t)$与它的测量值 $y(t)$有关系

$$y(t)=s(t)+v(t) \tag{3.1.1}$$

其中 $v(t)$是在时刻 tT 的测量误差,也叫观测噪声,通常可认为它是零均值、方差为 σ_v^2 的正态白噪声.为了有效地发射鱼雷命中敌舰,问题是如何基于到时刻 t 为止对敌舰位置的观测数据($y(t),y(t-1),\cdots,y(1)$)求在未来时刻$(t+k)T$ 敌舰的真实位置 $s(t+k)$的最优预报 $\hat{s}(t+k\mid t)$.

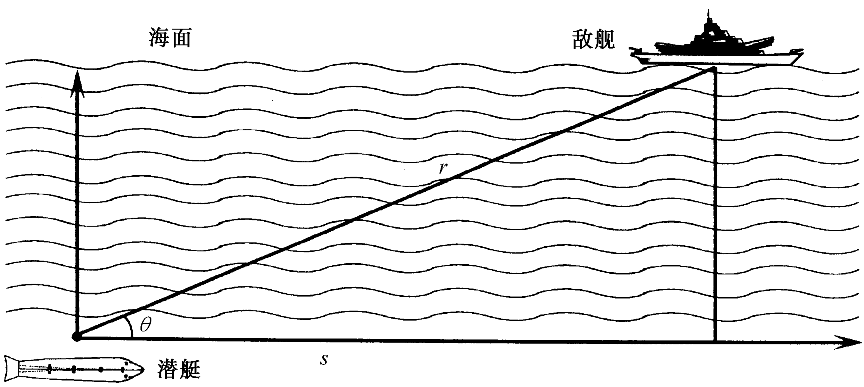


图 3.1.1　潜艇声纳目标跟踪预报问题

按均速运动的敌舰有运动方程

$$\begin{aligned}&s(t+1)=s(t)+T\dot{s}(t),\\&\dot{s}(t+1)=\dot{s}(t)\quad(\text{速度 }\dot{s}(t)\text{为常数})\end{aligned} \tag{3.1.2}$$

引入敌舰的状态变量为

$$\boldsymbol{x}(t)=\begin{bmatrix}s(t)\\\dot{s}(t)\end{bmatrix} \tag{3.1.3}$$

则有状态空间模型

$$\begin{aligned}&\begin{bmatrix}s(t+1)\\\dot{s}(t+1)\end{bmatrix}=\begin{bmatrix}1&T\\0&1\end{bmatrix}\begin{bmatrix}s(t)\\\dot{s}(t)\end{bmatrix},\\&y(t)=[1\quad 0]\begin{bmatrix}s(t)\\\dot{s}(t)\end{bmatrix}+v(t)\end{aligned} \tag{3.1.4}$$

即有状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}\boldsymbol{x}(t),$$
$$y(t)=\boldsymbol{H}\boldsymbol{x}(t)+v(t),$$
$$\boldsymbol{\Phi}=\begin{bmatrix}1 & T\\0 & 1\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0] \tag{3.1.5}$$

于是问题转化为对系统(3.1.5)求 $\boldsymbol{x}(t+k)$最优 Kalman 预报器 $\hat{\boldsymbol{x}}(t+k\mid t)(k>0)$,它的第一个分量就是 $\hat{s}(t+k\mid t)$.

现在讨论更实际的情形. 由于海面上海风和海浪的影响,虽然理论上敌舰按均速运动,但实际上敌舰运动的实际速度模型应由 $\dot{s}(t+1)=\dot{s}(t)$修改为

$$\dot{s}(t+1)=\dot{s}(t)+w(t) \tag{3.1.6}$$

其中附加噪声 $w(t)$体现了海风和海浪对速度的随机影响,通常假设 $w(t)$是白噪声,这种模型叫随机游运模型,它体现了在相邻采样间隔时间实际速度的随机变化. 通常 $w(t)$的方差 σ_w^2 很小,因而在相邻时间速度 $\dot{s}(t)$只有微小变化. 在这种情况下,我们有敌舰运动的状态空间模型

$$\begin{bmatrix}s(t+1)\\\dot{s}(t+1)\end{bmatrix}=\begin{bmatrix}1 & T\\0 & 1\end{bmatrix}\begin{bmatrix}s(t)\\\dot{s}(t)\end{bmatrix}+\begin{bmatrix}0\\1\end{bmatrix}w(t),$$
$$y(t)=[1\quad 0]\begin{bmatrix}s(t)\\\dot{s}(t)\end{bmatrix}+v(t) \tag{3.1.7}$$

即有状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}\boldsymbol{x}(t)+\boldsymbol{\Gamma}w(t),$$
$$y(t)=\boldsymbol{H}x(t)+v(t),$$
$$\boldsymbol{x}(t)=\begin{bmatrix}s(t)\\\dot{s}(t)\end{bmatrix},\quad \boldsymbol{\Phi}=\begin{bmatrix}1 & T\\0 & 1\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}0\\1\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0] \tag{3.1.8}$$

于是问题转化为对系统(3.1.8)求状态 $\boldsymbol{x}(t+k)$的最优 Kalman 预报估值 $\hat{\boldsymbol{x}}(t+k\mid t)$ $(k>0)$问题.

【例 3.1.2】 地对空战斗的雷达跟踪问题.[16]

考虑在对空战斗中,一架敌机正在径直向我方雷达俯冲过来,见图 3.1.2 所示. 此时,为了命中敌机,要求立即根据雷达发出的脉冲电磁波的反射时间的测量结果,对敌机的距离 $r(t)$和速度 $\dot{r}(t)$做出尽可能准确的估算,即要求得出它们的最优估计.

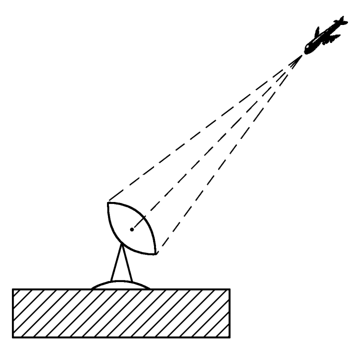


图 3.1.2 地对空战斗中的 Kalman 滤波问题

文献[16]用牛顿定律导出了敌机状态所满足的随机微分方程. 这里从运动定律出发导出不同于文献[16]的离散状态空间模型.

设采样周期为 T,用 $r(t)$,$\dot{r}(t)$和 $w(t)$分别表示在时刻 tT 敌机与雷达的径向距离、速度和加速度,则由运动定律有关系

$$r(t+1)=r(t)+T\dot{r}(t)+\frac{T^2}{2}w(t),$$

$$\dot{r}(t+1)=\dot{r}(t)+Tw(t),$$
$$y(t)=r(t)+v(t) \tag{3.1.9}$$
其中 $y(t)$为径向距离 $r(t)$的观测信号，$v(t)$为观测噪声(即观测误差)，通常视为白噪声.由于高空中风力变化的随机干扰，加上驾驶员为逃避我方火力，故意采取随机操纵作法，可将加速度 $w(t)$视为白噪声.

因为我们感兴趣飞机的距离 $r(t)$和速度 $\dot{r}(t)$，故本例可引入该系统的状态变量$x(t)$为

$$\boldsymbol{x}(t)=\begin{bmatrix} r(t) \\ \dot{r}(t) \end{bmatrix} \tag{3.1.10}$$

并且由(3.1.7)~(3.1.9)有状态方程和观测方程

$$\begin{bmatrix} r(t+1) \\ \dot{r}(t+1) \end{bmatrix}=\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}\begin{bmatrix} r(t) \\ \dot{r}(t) \end{bmatrix}+\begin{bmatrix} \dfrac{T^2}{2} \\ T \end{bmatrix}w(t) \tag{3.1.11}$$

$$y(t)=[1 \quad 0]\begin{bmatrix} r(t) \\ \dot{r}(t) \end{bmatrix}+v(t) \tag{3.1.12}$$

它可简写为状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma} w(t) \tag{3.1.13}$$
$$y(t)=\boldsymbol{Hx}(t)+v(t) \tag{3.1.14}$$

其中定义

$$\boldsymbol{\Phi}=\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma}=\begin{bmatrix} \dfrac{T^2}{2} \\ T \end{bmatrix}, \quad \boldsymbol{H}=[1 \quad 0] \tag{3.1.15}$$

因而 Kalman 滤波问题成为由状态空间模型(3.1.13)和(3.1.14)，基于到时刻 t 的观测$(y(t),y(t-1),\cdots)$，求状态 $\boldsymbol{x}(t)$的最优(线性最小方差)估计 $\hat{\boldsymbol{x}}(t|t)$.

【例 3.1.3】 GPS 车辆导航定位系统.[16]

全球定位系统(GPS—Global Positioning System)现已广泛应用于军事和国民经济各领域.GPS 车辆导航定位系统如图 3.1.3 所示.其原理是在地面上装有接收机的运动车辆(载体)可接收到在轨的 24 颗导航卫星播发的信号，通过采集安装于车辆上的接收机测得的数据，可能算出作为接收机载体的车辆当前的位置和速度.但美国出于自身利益的考虑，自 1991 年 7 月起对在轨的 24 颗导航卫星播发的信号加入高频振荡随机噪声干扰信号，致使所有派生的卫星信号均产生高频抖动，致使 GPS 用户二维定位精度只有 ±100m 左右.因此，提高 GPS 定位精度是倍受人们关注的重要问题.Kalman 滤波方法是解决此问题的有力工具.下面说明该问题如何化为 Kalman 滤波问题.

文献[16]的 GPS 定位模型没有考虑驾驶员的机动操作.车辆运动是一个控制系统，现在给出考虑驾驶员机动操作的 GPS 定位模型，它更符合于情况，并包括文献[16]的 GPS 定位模型作为特殊情形.

通常可近似认为在城市道路上行驶的车辆在二维平面内运动.设采样周期为 T，在图 3.1.4 中，车辆的运动可分解为东向运动和北向运动.车辆的运动状态变量可取为：

$x_e(t)$——在时刻 tT 东向位置；

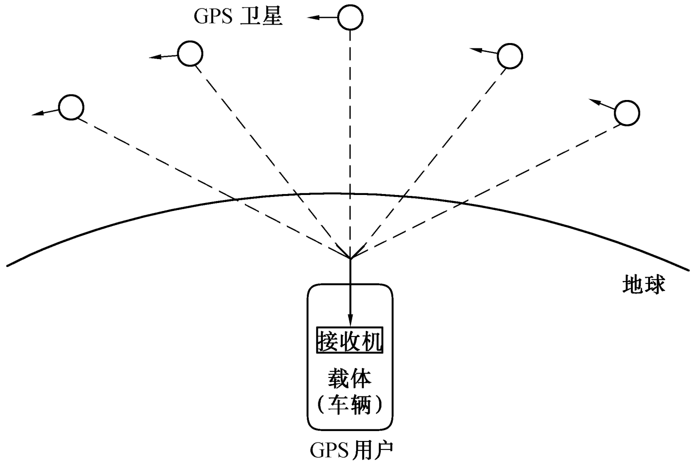


图 3.1.3　GPS 车辆定位导航系统

$\dot{x}_e(t)$——在时刻 tT 东向速度；

$\ddot{x}_e(t)$——在时刻 tT 东向加速度；

$x_n(t)$——在时刻 tT 北向位置；

$\dot{x}_n(t)$——在时刻 tT 北向速度；

$\ddot{x}_n(t)$—在时刻 tT 北向加速度。　　(3.1.16)

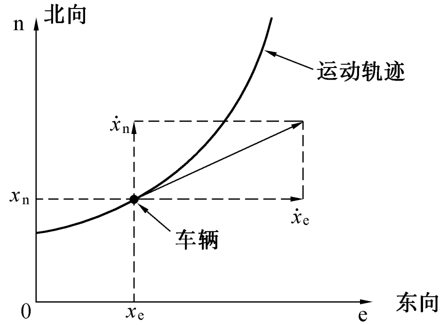


图 3.1.4　接收机载体(车辆)运动的分解

因而车辆运动的位置、速度和加速度完全被其东向和北向运动的位置、速度和加速度决定. 例如在时刻 tT 处，车辆位置坐标为 $(x_e(t), x_n(t))$，车辆距坐标原点距离为

$$\rho(t) = \sqrt{x_e^2(t) + x_n^2(t)} \tag{3.1.17}$$

在时刻 tT 处车辆的速度为

$$v(t) = \sqrt{\dot{x}_e^2(t) + \dot{x}_n^2(t)} \tag{3.1.18}$$

由均加速运动定律在东向和北向车辆的运动方程各为

$$\begin{aligned} x_e(t+1) &= x_e(t) + T\dot{x}_e(t) + \frac{T^2}{2}[\ddot{x}_e(t) + u_e] \\ \dot{x}_e(t+1) &= \dot{x}_e(t) + T[\ddot{x}_e(t) + u_e] \\ \ddot{x}_e(t+1) &= \ddot{x}_e(t) + w_e(t) \end{aligned} \tag{3.1.19}$$

$$\begin{aligned} y_{xe}(t) &= x_e(t) + v_{xe}(t), \\ y_{ve}(t) &= \dot{x}_e(t) + v_{ve}(t) \end{aligned} \tag{3.1.20}$$

和

$$\begin{aligned} x_n(t+1) &= x_n(t) + T\dot{x}_n(t) + \frac{T^2}{2}[\ddot{x}_n(t) + u_n] \\ \dot{x}_n(t+1) &= \dot{x}_n(t) + T[\ddot{x}_n(t) + u_n] \\ \ddot{x}_n(t+1) &= \ddot{x}_n(t) + w_n(t) \end{aligned} \tag{3.1.21}$$

$$y_{xn}(t) = x_n(t) + v_{xn}(t),$$
$$y_{vn}(t) = \dot{x}_n(t) + v_{vn}(t) \tag{3.1.22}$$

其中 u_e 和 u_n 各为东向和北向由驾驶员决定的常的机动加速度.(3.1.19)和(3.1.21)意味着假设东向和北向随机加速度服从随机游动模型,其中 $w_e(t)$ 和 $w_n(t)$ 为相互独立的白噪声.(3.1.20)和(3.1.22)为东向和北向位置和速度的观测方程,其中假设位置观测噪声 $v_{xe}(t)$ 和 $v_{xn}(t)$ 为相互独立的白噪声,且假设速度观测噪声 $v_{ve}(t)$ 和 $v_{vn}(t)$ 为相互独立的白噪声.因而有东向和北向的状态空间模型分别为

$$\begin{bmatrix} x_e(t+1) \\ \dot{x}_e(t+1) \\ \ddot{x}_n(t+1) \end{bmatrix} = \begin{bmatrix} 1 & T & 0.5T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_e(t) \\ \dot{x}_e(t) \\ \ddot{x}_e(t) \end{bmatrix} + \begin{bmatrix} 0.5T^2 \\ T \\ 0 \end{bmatrix} u_e + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} w_e(t) \tag{3.1.23}$$

$$\begin{bmatrix} y_{xe}(t) \\ y_{ve}(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_e(t) \\ \dot{x}_e(t) \\ \ddot{x}_e(t) \end{bmatrix} + \begin{bmatrix} v_{xe}(t) \\ v_{ve}(t) \end{bmatrix} \tag{3.1.24}$$

$$\begin{bmatrix} x_n(t+1) \\ \dot{x}_n(t+1) \\ \ddot{x}_n(t+1) \end{bmatrix} = \begin{bmatrix} 1 & T & 0.5T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_n(t) \\ \dot{x}_n(t) \\ \ddot{x}_n(t) \end{bmatrix} + \begin{bmatrix} 0.5T^2 \\ T \\ 0 \end{bmatrix} u_n + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} w_n(t) \tag{3.1.25}$$

$$\begin{bmatrix} y_{xn}(t) \\ y_{vn}(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_n(t) \\ \dot{x}_n(t) \\ \ddot{x}_n(t) \end{bmatrix} + \begin{bmatrix} v_{xn}(t) \\ v_{vn}(t) \end{bmatrix} \tag{3.1.26}$$

引入记号

$$\boldsymbol{X}_e(t) = \begin{bmatrix} x_e(t) \\ \dot{x}_e(t) \\ \ddot{x}_e(t) \end{bmatrix}, \quad \boldsymbol{Y}_e(t) = \begin{bmatrix} y_{xe}(t) \\ y_{ve}(t) \end{bmatrix}, \quad \boldsymbol{V}_e(t) = \begin{bmatrix} v_{xe}(t) \\ v_{ve}(t) \end{bmatrix} \tag{3.1.27}$$

$$\boldsymbol{X}_n(t) = \begin{bmatrix} x_n(t) \\ \dot{x}_n(t) \\ \ddot{x}_n(t) \end{bmatrix}, \quad \boldsymbol{Y}_n(t) = \begin{bmatrix} y_{xn}(t) \\ y_{vn}(t) \end{bmatrix}, \quad \boldsymbol{V}_n(t) = \begin{bmatrix} v_{xn}(t) \\ v_{vn}(t) \end{bmatrix} \tag{3.1.28}$$

和定义矩阵

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T & 0.5T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 0.5T^2 \\ T \\ 0 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{3.1.29}$$

则东向和北向的状态空间模型各为

$$\boldsymbol{X}_e(t+1) = \boldsymbol{\Phi}\boldsymbol{X}_e(t) + \boldsymbol{B}u_e + \boldsymbol{\Gamma}w_e(t) \tag{3.1.30}$$

$$\boldsymbol{Y}_e(t) = \boldsymbol{H}\boldsymbol{X}_e(t) + \boldsymbol{V}_e(t) \tag{3.1.31}$$

和

$$\boldsymbol{X}_n(t+1) = \boldsymbol{\Phi}\boldsymbol{X}_n(t) + \boldsymbol{B}u_n + \boldsymbol{\Gamma}w_n(t) \tag{3.1.32}$$

$$\boldsymbol{Y}_n(t) = \boldsymbol{H}\boldsymbol{X}_n(t) + \boldsymbol{V}_n(t) \tag{3.1.33}$$

它们具有相同的形式.从物理意义上看,东向和北向运动的位置、速度和加速度之间没有

必然联系，因而是解耦的. 故可对东向和北向运动单独分别处理. 车辆 GPS 导航定位 Kalman 滤波问题是：基于东向和北向观测数据 $\boldsymbol{Y}_e(t)$，$\boldsymbol{Y}_e(t-1)$，…和 $\boldsymbol{Y}_n(t)$，$\boldsymbol{Y}_n(t-1)$，…，分别求东向状态 $\boldsymbol{X}_e(t)$和北向状态 $\boldsymbol{X}_n(t)$的最优估值 $\hat{\boldsymbol{X}}_e(t\mid t)$和 $\hat{\boldsymbol{X}}_n(t\mid t)$. 注意在观测噪声 $\boldsymbol{V}_e(t)$和 $\boldsymbol{V}_n(t)$中考虑了 GPS 信号中含有的随机干扰噪声. 上述状态空间模型包括文献[16]的 GPS 定位模型作为特殊情形，应注意，随机加速度 $\ddot{x}_e(t)$和 $\ddot{x}_n(t)$由风的变化引起，而机动加速度 u_e 和 u_n 由驾驶员决定. 在文献[16]中取 $u_e=0$，$u_n=0$.

3.2　射影理论

Kalman 滤波器是线性最小方差估值器，也叫最优滤波器，在几何上 Kalman 滤波估值可看出是状态变量在由观测生成的线性空间上的射影. 因此射影理论是 Kalman 滤波推导的基本工具. 本节介绍射影和新息概念及其性质，以及射影公式.

3.2.1　线性最小方差估计和射影

【定义 3.2.1】　由 $m\times1$ 维随机变量 $\boldsymbol{y}\in R^m$ 的线性函数估计 $n\times1$ 维随机变量 $\boldsymbol{x}\in R^n$，记估值为

$$\hat{\boldsymbol{x}}=\boldsymbol{b}+\boldsymbol{A}\boldsymbol{y},\quad \boldsymbol{b}\in R^n,\quad \boldsymbol{A}\in R^{n\times m}\tag{3.2.1}$$

若估值 $\hat{\boldsymbol{x}}$ 极小化性能指标为 J，

$$J=\mathrm{E}[(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\hat{\boldsymbol{x}})]\tag{3.2.2}$$

则称 $\hat{\boldsymbol{x}}$ 为随机变量 $\boldsymbol{x}$ 的线性最小方差估计，其中 E 为均值号，T 为转置号.

如何求待定 $n\times1$ 向量 $\boldsymbol{b}$ 和 $n\times m$ 矩阵 $\boldsymbol{A}$？

分析　将(3.2.1)代入(3.2.2)有

$$J=\mathrm{E}[(\boldsymbol{x}-\boldsymbol{b}-\boldsymbol{A}\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{b}-\boldsymbol{A}\boldsymbol{y})]\tag{3.2.3}$$

应选择 $\boldsymbol{b}$ 和 $\boldsymbol{A}$ 极小化 J. 置 $\partial J/\partial\boldsymbol{b}=\boldsymbol{0}$ 有

$$\frac{\partial J}{\partial\boldsymbol{b}}=-2\mathrm{E}(\boldsymbol{x}-\boldsymbol{b}-\boldsymbol{A}\boldsymbol{y})=\boldsymbol{0}\tag{3.2.4}$$

这引出

$$\boldsymbol{b}=\mathrm{E}\boldsymbol{x}-\boldsymbol{A}\,\mathrm{E}\boldsymbol{y}\tag{3.2.5}$$

将(3.2.5)引入(3.2.3)并定义方差阵和协方差阵符号

$$\boldsymbol{P}_{xx}=\mathrm{E}[(\boldsymbol{x}-\mathrm{E}\boldsymbol{x})(\boldsymbol{x}-\mathrm{E}\boldsymbol{x})^{\mathrm{T}}]\tag{3.2.6}$$

$$\boldsymbol{P}_{xy}=\mathrm{E}[(\boldsymbol{x}-\mathrm{E}\boldsymbol{x})(\boldsymbol{y}-\mathrm{E}\boldsymbol{y})^{\mathrm{T}}]\tag{3.2.7}$$

易知有关系

$$\boldsymbol{P}_{xy}=\boldsymbol{P}_{yx}{}^{\mathrm{T}}\tag{3.2.8}$$

$$\begin{aligned}J=&\mathrm{E}[((\boldsymbol{x}-\mathrm{E}\boldsymbol{x})-\boldsymbol{A}(\boldsymbol{y}-\mathrm{E}\boldsymbol{y}))^{\mathrm{T}}((\boldsymbol{x}-\mathrm{E}\boldsymbol{x})-\boldsymbol{A}(\boldsymbol{y}-\mathrm{E}\boldsymbol{y}))]=\\&\mathrm{tr}\mathrm{E}[((\boldsymbol{x}-\mathrm{E}\boldsymbol{x})-\boldsymbol{A}(\boldsymbol{y}-\mathrm{E}\boldsymbol{y}))((\boldsymbol{x}-\mathrm{E}\boldsymbol{x})-\boldsymbol{A}(\boldsymbol{y}-\mathrm{E}\boldsymbol{y}))^{\mathrm{T}}]=\\&\mathrm{tr}[\boldsymbol{P}_{xx}-\boldsymbol{A}\boldsymbol{P}_{yx}-\boldsymbol{P}_{xy}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{A}\boldsymbol{P}_{yy}\boldsymbol{A}^{\mathrm{T}}]=\\&\mathrm{tr}\boldsymbol{P}_{xx}+\mathrm{tr}\boldsymbol{A}\boldsymbol{P}_{yx}-\mathrm{tr}\boldsymbol{P}_{xy}\boldsymbol{A}^{\mathrm{T}}+\mathrm{tr}\boldsymbol{A}\boldsymbol{P}_{yy}\boldsymbol{A}^{\mathrm{T}}\end{aligned}\tag{3.2.9}$$

置 $\partial J/\partial\boldsymbol{A}=\boldsymbol{0}$，应用矩阵迹求导公式有

$$\frac{\partial J}{\partial \boldsymbol{A}} = -\boldsymbol{P}_{yx}{}^{\mathrm{T}} - \boldsymbol{P}_{xy} + 2\boldsymbol{A}\boldsymbol{P}_{yy} = \boldsymbol{0} \tag{3.2.10}$$

利用关式(3.2.8)引出

$$\boldsymbol{A} = \boldsymbol{P}_{xy}\boldsymbol{P}_{yy}{}^{-1} \tag{3.2.11}$$

上述结果可概括为如下定理.

【定理 3.2.1】 由随机变量 $\boldsymbol{y} \in R^m$ 对随机变量 $\boldsymbol{x} \in R^n$ 的线性最小方差估值公式为

$$\hat{\boldsymbol{x}} = \mathrm{E}\boldsymbol{x} + \boldsymbol{P}_{xy}\boldsymbol{P}_{yy}{}^{-1}(\boldsymbol{y} - \mathrm{E}\boldsymbol{y}) \tag{3.2.12}$$

其中假设 $\mathrm{E}\boldsymbol{x}, \mathrm{E}\boldsymbol{y}, \boldsymbol{P}_{xy}, \boldsymbol{P}_{yy}$ 均存在.

【推论 3.2.1】 (无偏性)$\mathrm{E}\hat{\boldsymbol{x}} = \mathrm{E}\boldsymbol{x}$.

证明 $$\mathrm{E}\hat{\boldsymbol{x}} = \mathrm{E}[\mathrm{E}\boldsymbol{x} + \boldsymbol{P}_{xy}\boldsymbol{P}_{yy}{}^{-1}(\boldsymbol{y} - \mathrm{E}\boldsymbol{y})] = \mathrm{E}\boldsymbol{x} + \boldsymbol{P}_{xy}\boldsymbol{P}_{yy}{}^{-1}(\mathrm{E}\boldsymbol{y} - \mathrm{E}\boldsymbol{y}) = \mathrm{E}\boldsymbol{x} \tag{3.2.13}$$

【推论 3.2.2】 (正交性)$$\mathrm{E}[(\boldsymbol{x} - \hat{\boldsymbol{x}})\boldsymbol{y}^{\mathrm{T}}] = \boldsymbol{0} \tag{3.2.14}$$

证明 $$\begin{aligned}&\mathrm{E}[(\boldsymbol{x} - \mathrm{E}\boldsymbol{x} - \boldsymbol{P}_{xy}\boldsymbol{P}_{yy}{}^{-1}(\boldsymbol{y} - \mathrm{E}\boldsymbol{y}))\boldsymbol{y}^{\mathrm{T}}] = \\ &\mathrm{E}[(\boldsymbol{x} - \mathrm{E}\boldsymbol{x} - \boldsymbol{P}_{xy}\boldsymbol{P}_{yy}{}^{-1}(\boldsymbol{y} - \mathrm{E}\boldsymbol{y}))(\boldsymbol{y} - \mathrm{E}\boldsymbol{y})^{\mathrm{T}}] = \\ &\boldsymbol{P}_{xy} - \boldsymbol{P}_{xy}\boldsymbol{P}_{yy}{}^{-1}\boldsymbol{P}_{yy} = \boldsymbol{P}_{xy} - \boldsymbol{P}_{xy} = \boldsymbol{0}\end{aligned} \tag{3.2.15}$$

【推论 3.2.3】 $\tilde{\boldsymbol{x}} = \boldsymbol{x} - \hat{\boldsymbol{x}}$ 与 $\boldsymbol{y}$ 不相关.

证明 $\mathrm{E}[(\boldsymbol{x} - \hat{\boldsymbol{x}})(\boldsymbol{y} - \mathrm{E}\boldsymbol{y})^{\mathrm{T}}] = \mathrm{E}[(\boldsymbol{x} - \hat{\boldsymbol{x}})\boldsymbol{y}^{\mathrm{T}}] - \mathrm{E}[(\boldsymbol{x} - \hat{\boldsymbol{x}})(\mathrm{E}\boldsymbol{y})^{\mathrm{T}}] = \boldsymbol{0}$. □

【定义 3.2.2】 我们称 $\boldsymbol{x} - \hat{\boldsymbol{x}}$ 与 $\boldsymbol{y}$ 不相关为 $\boldsymbol{x} - \hat{\boldsymbol{x}}$ 与 $\boldsymbol{y}$ 正交(垂直),记为$(\boldsymbol{x} - \hat{\boldsymbol{x}}) \perp \boldsymbol{y}$,并称 $\hat{\boldsymbol{x}}$ 为 $\boldsymbol{x}$ 在 $\boldsymbol{y}$ 上的射影,记为 $\hat{\boldsymbol{x}} = \mathrm{proj}(\boldsymbol{x} | \boldsymbol{y})$.

线性最小方差估值 $\hat{\boldsymbol{x}}$ 的几何意义如图 3.2.1 所示.

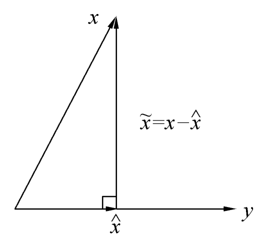


图 3.2.1 线性最小方差估计(射影)几何意义

【定义 3.2.3】 由随机变量 $\boldsymbol{y} \in R^m$ 张成的线性流形(线性空间)定义为如下形式随机变量 $\boldsymbol{z} \in R^n$ 的集合 $L(\boldsymbol{y})$,

$$L(\boldsymbol{y}) = \{\boldsymbol{z} | \boldsymbol{z} = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}, \forall \boldsymbol{b} \in R^n, \forall \boldsymbol{A} \in R^{n \times m}\} \tag{3.2.16}$$

【推论 3.2.4】 $(\boldsymbol{x} - \hat{\boldsymbol{x}}) \perp \boldsymbol{z}, \forall \boldsymbol{z} \in L(\boldsymbol{y})$,记为 $(\boldsymbol{x} - \hat{\boldsymbol{x}}) \perp L(\boldsymbol{y})$.

证明 $$\begin{aligned}&\mathrm{E}[(\boldsymbol{x} - \hat{\boldsymbol{x}})\boldsymbol{z}^{\mathrm{T}}] = \mathrm{E}[(\boldsymbol{x} - \hat{\boldsymbol{x}})(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b})^{\mathrm{T}}] = \\ &\mathrm{E}[(\boldsymbol{x} - \hat{\boldsymbol{x}})\boldsymbol{y}^{\mathrm{T}}]\boldsymbol{A}^{\mathrm{T}} + \mathrm{E}[(\boldsymbol{x} - \hat{\boldsymbol{x}})]\boldsymbol{b}^{\mathrm{T}} = \boldsymbol{0}\end{aligned} \tag{3.2.17}$$

其中应用了推论 3.2.1 和推论 3.2.2. □

设随机变量 $\boldsymbol{x} \in R^n$,随机变量 $\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k) \in R^m$,引入合成随机变量 $\boldsymbol{w}$ 为

$$\boldsymbol{w} = (\boldsymbol{y}^{\mathrm{T}}(1), \cdots, \boldsymbol{y}^{\mathrm{T}}(k))^{\mathrm{T}} \in R^{km} \tag{3.2.18}$$

【定义 3.2.4】 由 $\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k) \in R^m$ 张成的线性流形 $L((\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k))$ 定义为

$$L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k)) \triangleq L(\boldsymbol{w}) \tag{3.2.19}$$

$$L(\boldsymbol{w}) = \{\boldsymbol{y} | \boldsymbol{y} = A\boldsymbol{w} + \boldsymbol{b}, \forall \boldsymbol{b} \in R^n, \forall \boldsymbol{A} \in R^{n \times km}\} \tag{3.2.20}$$

引入分块矩阵

$$\boldsymbol{A} = [\boldsymbol{A}_1, \cdots, \boldsymbol{A}_k], \quad \boldsymbol{A}_i \in R^{n \times m} \tag{3.2.21}$$

则有

$$L(\boldsymbol{w}) = \{\boldsymbol{y} | \boldsymbol{y} = \sum_{i=1}^{k} \boldsymbol{A}_i \boldsymbol{y}(i) + \boldsymbol{b}, \forall \boldsymbol{A}_i \in R^{n \times m}, \forall \boldsymbol{b} \in R^n\} = L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k)) \tag{3.2.22}$$

【定义 3.2.5】 基于随机变量 $\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k)\in R^m$ 对随机变量 $\boldsymbol{x}\in R^n$ 的线性最小方差估计 $\hat{\boldsymbol{x}}$ 定义为

$$\hat{\boldsymbol{x}}=\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{w})\triangleq\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k)) \tag{3.2.23}$$

也称 $\hat{\boldsymbol{x}}$ 为 $\boldsymbol{x}$ 在线性流形 $L(\boldsymbol{w})$ 或 $L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k))$ 上的射影.

【推论 3.2.5】 设 $\boldsymbol{x}\in R^n$ 为零均值随机变量，$\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k)\in R^m$ 为零均值，互不相关（正交）的随机变量，则有

$$\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k))=\sum_{i=1}^{k}\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{y}(i)) \tag{3.2.24}$$

即 $\boldsymbol{x}$ 在由 $\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k)$ 张成的线性流形 $L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k))$ 上的射影等于它在由每一个 $\boldsymbol{y}(i)$ 张成的线性流形上的射影之和，即 $\boldsymbol{x}$ 在全空间上的射影等于它在相互正交的子空间上的射影之和. 这大大简化了射影的计算. 因此，若 $\boldsymbol{y}(1),\cdots\boldsymbol{y}(k)$ 是相关的（非正交的），问题是如何使它们正交化. 这将导致新息概念.

证明 记合成向量 $\boldsymbol{w}=(\boldsymbol{y}^{\mathrm{T}}(1),\cdots,\boldsymbol{y}^{\mathrm{T}}(k))^{\mathrm{T}}\in R^{km}$，注意 $\mathrm{E}\boldsymbol{x}=\boldsymbol{0}$，$\mathrm{E}\boldsymbol{w}=\boldsymbol{0}$，则应用射影公式(3.2.12)有

$$\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k))=\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{w})=\boldsymbol{P}_{xw}\boldsymbol{P}_{ww}^{-1}\boldsymbol{w}=$$

$$\mathrm{E}[\boldsymbol{x}(\boldsymbol{y}^{\mathrm{T}}(1),\cdots,\boldsymbol{y}^{\mathrm{T}}(k))]\begin{bmatrix}\boldsymbol{P}_{y(1)y(1)}^{-1} & & \boldsymbol{0}\\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{P}_{y(k)y(k)}^{-1}\end{bmatrix}\begin{bmatrix}\boldsymbol{y}(1)\\ \vdots\\ \boldsymbol{y}(k)\end{bmatrix}=$$

$$\sum_{i=1}^{k}\boldsymbol{P}_{xy(i)}\boldsymbol{P}_{y(i)y(i)}^{-1}\boldsymbol{y}(i)=\sum_{i=1}^{k}\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{y}(i)) \tag{3.2.25}$$

【推论 3.2.6】 设随机变量 $\boldsymbol{x}\in R^p$，$\boldsymbol{z}\in R^q$，随机变量 $\boldsymbol{Ax}+\boldsymbol{Bz}\in R^n$，$\boldsymbol{A}\in R^{n\times p}$，$\boldsymbol{B}\in R^{n\times q}$，随机变量 $\boldsymbol{y}\in R^m$，则有

$$\mathrm{proj}(\boldsymbol{Ax}+\boldsymbol{Bz}\mid\boldsymbol{y})=\boldsymbol{A}\,\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{y})+\boldsymbol{B}\,\mathrm{proj}(\boldsymbol{z}\mid\boldsymbol{y}) \tag{3.2.26}$$

证明 应用射影公式(3.2.12)得证，从略. □

【推论 3.2.7】 设随机变量 $\boldsymbol{x}\in R^n$，随机变量 $\boldsymbol{y}\in R^m$，记 x 的分量形式为

$$\boldsymbol{x}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} \tag{3.2.27}$$

则有关系

$$\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{y})=\begin{bmatrix}\mathrm{proj}(x_1\mid\boldsymbol{y})\\ \vdots\\ \mathrm{proj}(x_n\mid\boldsymbol{y})\end{bmatrix} \tag{3.2.28}$$

证明

$$\mathrm{proj}(\boldsymbol{x}\mid\boldsymbol{y})=\mathrm{E}\boldsymbol{x}+\boldsymbol{P}_{xy}\boldsymbol{P}_{yy}^{-1}(\boldsymbol{y}-\mathrm{E}\boldsymbol{y})=\begin{bmatrix}\mathrm{E}x_1\\ \vdots\\ \mathrm{E}x_1\end{bmatrix}+\mathrm{E}\begin{bmatrix}x_1-\mathrm{E}x_1\\ \vdots\\ x_n-\mathrm{E}x_n\end{bmatrix}(\boldsymbol{y}-\mathrm{E}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{P}_{yy}^{-1}(\boldsymbol{y}-\mathrm{E}\boldsymbol{y})=$$

$$\begin{bmatrix} \mathrm{E}x_1 \\ \vdots \\ \mathrm{E}x_1 \end{bmatrix} + \begin{bmatrix} \boldsymbol{P}_{x_1 y} \\ \vdots \\ \boldsymbol{P}_{x_n y} \end{bmatrix} \boldsymbol{P}_{yy}{}^{-1}(\boldsymbol{y} - \mathrm{E}\boldsymbol{y}) = \begin{bmatrix} \mathrm{E}x_1 + \boldsymbol{P}_{x_1 y}\boldsymbol{P}_{yy}{}^{-1}(\boldsymbol{y} - \mathrm{E}\boldsymbol{y}) \\ \vdots \\ \mathrm{E}x_n + \boldsymbol{P}_{x_n y}\boldsymbol{P}_{yy}{}^{-1}(\boldsymbol{y} - \mathrm{E}\boldsymbol{y}) \end{bmatrix} = \begin{bmatrix} \mathrm{proj}(x_1 | \boldsymbol{y}) \\ \vdots \\ \mathrm{proj}(x_n | \boldsymbol{y}) \end{bmatrix} \tag{3.2.29}$$

这个性质说明随机变量 $\boldsymbol{x}$ 在线性流形 $L(\boldsymbol{y})$上的射影的每个分量必为 $\boldsymbol{x}$ 的相应的分量在线性流形 $L(\boldsymbol{y})$上的射影. 换言之，随机变量 $\boldsymbol{x}$ 的线性最小方差估计的每个分量等于 $\boldsymbol{x}$ 的相应的分量的线性最小方差估计，即全局最优估计等价于局部最优估计.

3.2.2 新息序列

【定义 3.2.6】 设 $\boldsymbol{y}(1), \boldsymbol{y}(2), \cdots, \boldsymbol{y}(k), \cdots \in R^m$ 是存在二阶矩的随机序列，它的新息序列(新息过程)定义为

$$\boldsymbol{\varepsilon}(k) = \boldsymbol{y}(k) - \mathrm{proj}(\boldsymbol{y}(k) | \boldsymbol{y}(1), \cdots, \boldsymbol{y}(k-1)), \quad k = 1, 2, \cdots \tag{3.2.30}$$

并定义 $\boldsymbol{y}(k)$的一步最优预报估值为

$$\hat{\boldsymbol{y}}(k | k-1) = \mathrm{proj}(\boldsymbol{y}(k) | \boldsymbol{y}(1), \cdots, \boldsymbol{y}(k-1)) \tag{3.2.31}$$

因而新息列可定义为

$$\boldsymbol{\varepsilon}(k) = \boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k | k-1), \quad k = 1, 2, \cdots \tag{3.2.32}$$

其中规定 $\hat{\boldsymbol{y}}(1|0) = \mathrm{E}\boldsymbol{y}(1)$，这保证 $\mathrm{E}\boldsymbol{\varepsilon}(1) = \boldsymbol{0}$.

新息的几何意义如图 3.2.2 所示，可看到 $\boldsymbol{\varepsilon}(k) \perp L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k-1))$.

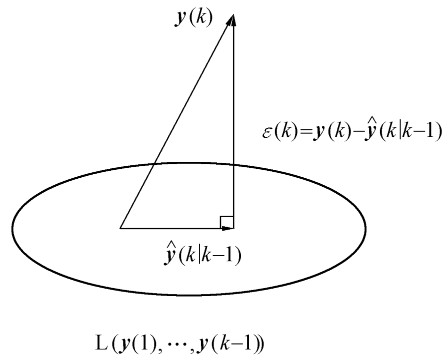


图 3.2.2 新息 $\varepsilon(k)$的几何意义

【定理 3.2.2】 新息序列 $\boldsymbol{\varepsilon}(k)$是零均值白噪声.

证明 由推论 3.2.1 射影是无偏差的，故有

$$\begin{aligned} \mathrm{E}\boldsymbol{\varepsilon}(k) = \mathrm{E}\boldsymbol{y}(k) - \mathrm{E}\hat{\boldsymbol{y}}(k | k-1) = \\ \mathrm{E}\boldsymbol{y}(k) - \mathrm{E}\boldsymbol{y}(k) = \boldsymbol{0}, \quad k \geqslant 1 \end{aligned} \tag{3.2.33}$$

设 $i \neq j$，不妨设 $i > j$，因 $\boldsymbol{\varepsilon}(i) \perp L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(i-1))$，且有 $L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(j)) \subset L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(i-1))$，故有 $\boldsymbol{\varepsilon}(i) \perp L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}((j))$. 而 $\boldsymbol{\varepsilon}(j) = \boldsymbol{y}(j) - \hat{\boldsymbol{y}}(j | j-1) \in L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(j))$，因而 $\boldsymbol{\varepsilon}(i) \perp \boldsymbol{\varepsilon}(j)$，即 $\mathrm{E}[\boldsymbol{\varepsilon}(i)\boldsymbol{\varepsilon}^{\mathrm{T}}(j)] = \boldsymbol{0}$，故 $\boldsymbol{\varepsilon}(i)$是白噪声. 证毕. □

定理 3.2.2 表明新息序列是正交(不相交)序列，通过引入新息序列实现了非正交(相关)随机序列的正交化. 由推论 3.2.5，在由新息序列张成的线性流形上的射影将大大简化射影的计算. 但问题是是否 $L(\boldsymbol{\varepsilon}(1), \cdots, \boldsymbol{\varepsilon}(k)) = L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k))$？即新息序列 $\boldsymbol{\varepsilon}(k)$与原序列 $\boldsymbol{y}(k)$是否含有相同的统计信息？下面定理回答了这个问题.

【定理 3.2.3】 新息序列 $\boldsymbol{\varepsilon}(k)$与原序列 $\boldsymbol{y}(k)$含有相同的统计信息，即 $(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k))$与 $(\boldsymbol{\varepsilon}(1), \cdots, \boldsymbol{\varepsilon}(k))$张成相同的线性流形，即

$$L(\boldsymbol{\varepsilon}(1), \cdots, \boldsymbol{\varepsilon}(k)) = L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k)), \quad k = 1, 2, \cdots \tag{3.2.34}$$

证明 由射影公式(3.2.12)和定义(3.2.31)，每个 $\boldsymbol{\varepsilon}(k)$是 $\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k)$的线性组合，这引出 $\boldsymbol{\varepsilon}(k) \in L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k))$，从而有 $L(\boldsymbol{\varepsilon}(1), \cdots, \boldsymbol{\varepsilon}(k)) \subset L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(k))$.

下证 $\boldsymbol{y}(k) \in L(\boldsymbol{\varepsilon}(1), \cdots, \boldsymbol{\varepsilon}(k))$.

事实上,我们由(3.2.31)用归纳法有

$$\begin{aligned}
&\boldsymbol{y}(1)=\boldsymbol{\varepsilon}(1)+\mathrm{E}\boldsymbol{y}(1)\in L(\boldsymbol{\varepsilon}(1)),\\
&\boldsymbol{y}(2)=\boldsymbol{\varepsilon}(2)+\operatorname{proj}(\boldsymbol{y}(2)\mid\boldsymbol{y}(1))\in L(\boldsymbol{\varepsilon}(1),\boldsymbol{\varepsilon}(2)),\\
&\quad\vdots\\
&\boldsymbol{y}(k)=\boldsymbol{\varepsilon}(k)+\operatorname{proj}(\boldsymbol{y}(k)\mid\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k-1))\in L(\boldsymbol{\varepsilon}(1),\boldsymbol{\varepsilon}(2),\cdots,\boldsymbol{\varepsilon}(k))
\end{aligned}\tag{3.2.35}$$

这引出 $L(\boldsymbol{y}(1),\cdots\boldsymbol{y}(k))\subset L(\boldsymbol{\varepsilon}(1),\cdots,\boldsymbol{\varepsilon}(k))$. 故有(3.2.34)成立. □

【推论 3.2.7】 设随机变量 $\boldsymbol{x}\in R^n$,则有

$$\operatorname{proj}(\boldsymbol{x}\mid\boldsymbol{y}(1),\cdots\boldsymbol{y}(k))=\operatorname{proj}(\boldsymbol{x}\mid\boldsymbol{\varepsilon}(1),\cdots,\boldsymbol{\varepsilon}(k))\tag{3.2.36}$$

由于新息序列的正交性,这一推论将大大简化射影的计算.

【定理 3.2.4】 (递推射影公式)设随机变量 $\boldsymbol{x}\in R^n$,随机序列 $\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k),\cdots\in R^m$,且它们存在二阶矩,则有递推射影公式

$$\operatorname{proj}(\boldsymbol{x}\mid\boldsymbol{y}(1),\cdots\boldsymbol{y}(k))=\operatorname{proj}(\boldsymbol{x}\mid\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k-1))+\mathrm{E}[\boldsymbol{x}\boldsymbol{\varepsilon}^{\mathrm{T}}(k)][\mathrm{E}(\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^{\mathrm{T}}(k)]^{-1}\boldsymbol{\varepsilon}(k)\tag{3.2.37}$$

证明 引入合成向量

$$\boldsymbol{\varepsilon}=\begin{bmatrix}\boldsymbol{\varepsilon}(1)\\ \vdots\\ \boldsymbol{\varepsilon}(k)\end{bmatrix}\tag{3.2.38}$$

应用(3.2.36)和射影公式,并注意 $\mathrm{E}\boldsymbol{\varepsilon}(i)=\mathbf{0}$,有

$$\begin{aligned}
&\operatorname{proj}(\boldsymbol{x}\mid\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k))=\operatorname{proj}(\boldsymbol{x}\mid\boldsymbol{\varepsilon}(1),\cdots,\boldsymbol{\varepsilon}(k))=\\
&\operatorname{proj}(\boldsymbol{x}\mid\boldsymbol{\varepsilon})=\mathrm{E}\boldsymbol{x}+\boldsymbol{P}_{x\varepsilon}\boldsymbol{P}_{\varepsilon\varepsilon}^{-1}\boldsymbol{\varepsilon}=\mathrm{E}\boldsymbol{x}+\mathrm{E}[(\boldsymbol{x}-\mathrm{E}\boldsymbol{x})(\boldsymbol{\varepsilon}^{\mathrm{T}}(1),\cdots,\boldsymbol{\varepsilon}^{\mathrm{T}}(k))]\times\\
&\begin{bmatrix}\mathrm{E}[\boldsymbol{\varepsilon}(1)\boldsymbol{\varepsilon}^{\mathrm{T}}(1)] & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathrm{E}[\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^{\mathrm{T}}(k)]\end{bmatrix}\begin{bmatrix}\boldsymbol{\varepsilon}(1)\\ \vdots\\ \boldsymbol{\varepsilon}(k)\end{bmatrix}=\\
&\mathrm{E}\boldsymbol{x}+\sum_{i=1}^{k}\mathrm{E}[\boldsymbol{x}\boldsymbol{\varepsilon}^{\mathrm{T}}(i)][\mathrm{E}[\boldsymbol{\varepsilon}(i)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]^{-1}\boldsymbol{\varepsilon}(i)=\\
&\mathrm{E}\boldsymbol{x}+\sum_{i=1}^{k-1}\mathrm{E}[\boldsymbol{x}\boldsymbol{\varepsilon}^{\mathrm{T}}(i)][\mathrm{E}[\boldsymbol{\varepsilon}(i)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]^{-1}\boldsymbol{\varepsilon}(i)+\mathrm{E}[\boldsymbol{x}\boldsymbol{\varepsilon}^{\mathrm{T}}(k)][\mathrm{E}(\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^{\mathrm{T}}(k)]^{-1}\boldsymbol{\varepsilon}(k)=\\
&\operatorname{proj}(\boldsymbol{x}\mid\boldsymbol{\varepsilon}(1),\cdots,\boldsymbol{\varepsilon}(k-1))+\mathrm{E}[\boldsymbol{x}\boldsymbol{\varepsilon}^{\mathrm{T}}(k)][\mathrm{E}[\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^{\mathrm{T}}(k)]^{-1}\boldsymbol{\varepsilon}(k)=\\
&\operatorname{proj}(\boldsymbol{x}\mid\boldsymbol{y}(1),\cdots,\boldsymbol{y}(k-1))+\mathrm{E}[\boldsymbol{x}\boldsymbol{\varepsilon}^{\mathrm{T}}(k)][\mathrm{E}[\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^{\mathrm{T}}(k)]^{-1}\boldsymbol{\varepsilon}(k)
\end{aligned}\tag{3.2.39}$$

递推射影公式(3.2.37)是推寻 Kalman 滤波器的递推算法的出发点.

3.3 Kalman 滤波器和预报器

考虑用如下状态空间模型描写的动态系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}\boldsymbol{x}(t)+\boldsymbol{\Gamma}\boldsymbol{w}(t)\tag{3.3.1}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}\boldsymbol{x}(t)+\boldsymbol{v}(t)\tag{3.3.2}$$

其中 t 为离散时间,系统在时刻 t 的状态为 $\boldsymbol{x}(t)\in R^n$,$\boldsymbol{y}(t)\in R^m$ 为对状态的观测信号,

$\boldsymbol{w}(t)\in R^r$ 为输入白噪声，$\boldsymbol{v}(t)\in R^m$ 为观测噪声．称（3.3.1）为状态方程，称（3.3.2）为观测方程．$\boldsymbol{\Phi}$，$\boldsymbol{\Gamma}$，$\boldsymbol{H}$ 分别为已知的 $n\times n$，$n\times r$ 和 $m\times n$ 矩阵．称 $\boldsymbol{\Phi}$ 为状态转移阵，$\boldsymbol{H}$ 为观测阵．

【假设 1】 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 是零均值、方差阵各为 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 的不相关白噪声

$$\mathrm{E}\boldsymbol{w}(t)=\mathbf{0},\quad \mathrm{E}\boldsymbol{v}(t)=\mathbf{0} \tag{3.3.3}$$

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(j)]=\boldsymbol{Q}\delta_{tj},\quad \mathrm{E}[\boldsymbol{v}(t)\boldsymbol{v}^{\mathrm{T}}(j)]=\boldsymbol{R}\delta_{tj} \tag{3.3.4}$$

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(j)]=\mathbf{0},\quad \forall\, t,j \tag{3.3.5}$$

其中 E 为均值号，T 为转置号，$\delta_{tt}=1$，$\delta_{tj}=0\,(t\neq j)$．

【假设 2】 $\boldsymbol{x}(0)$ 不相关于 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$，

$$\mathrm{E}\boldsymbol{x}(0)=\boldsymbol{\mu}_0,\ \mathrm{E}[(\boldsymbol{x}(0)-\boldsymbol{\mu}_0)(\boldsymbol{x}(0)-\boldsymbol{\mu}_0)^{\mathrm{T}}]=\mathbf{P}_0 \tag{3.3.6}$$

Kalman 滤波问题是：基于观测（$\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t)$），求状态 $\boldsymbol{x}(j)$ 的线性最小方差估值 $\hat{\boldsymbol{x}}(j\mid t)$，它极小化性能指标为

$$J=\mathrm{E}[(\boldsymbol{x}(j)-\hat{\boldsymbol{x}}(j\mid t))^{\mathrm{T}}(\boldsymbol{x}(j)-\hat{\boldsymbol{x}}(j\mid t))] \tag{3.3.7}$$

对于 $j=t$，$j>t$，$j<t$，各称 $\hat{\boldsymbol{x}}(j\mid t)$ 为 Kalman 滤波器、预报器和平滑器．

注意，在解决导弹拦截，卫星回收等问题时，涉及到导弹或卫星轨道预测，即涉及到 Kalman 预报问题．在解决卫星入轨初速度估计或卫星轨道重构问题时，涉及到 Kalman 平滑器．本节应用射影理论主要推导 Kalman 滤波器，并顺便给出 Kalman 预报器．

在性能指标（3.3.7）下，问题归结为求射影

$$\hat{\boldsymbol{x}}(j\mid t)=\mathrm{proj}(\boldsymbol{x}(j)\mid \boldsymbol{y}(1),\cdots,\boldsymbol{y}(t)) \tag{3.3.8}$$

由递推射影公式（3.2.37）有递推关系

$$\hat{\boldsymbol{x}}(t+1\mid t+1)=\hat{\boldsymbol{x}}(t+1\mid t)+\boldsymbol{K}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.3.9}$$

$$\boldsymbol{K}(t+1)=\mathrm{E}[\boldsymbol{x}(t+1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)][\mathrm{E}(\boldsymbol{\varepsilon}(t+1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)]^{-1} \tag{3.3.10}$$

称 $\boldsymbol{K}(t+1)$ 为 Kalman 滤波器增益．对（3.3.1）两边项取射影有

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t\mid t)+\boldsymbol{\Gamma}\mathrm{proj}(\boldsymbol{w}(t)\mid \boldsymbol{y}(1),\cdots,\boldsymbol{y}(t)) \tag{3.3.11}$$

由（3.3.1）迭代有

$$\boldsymbol{x}(t)\in L(\boldsymbol{w}(t-1),\cdots,\boldsymbol{w}(0),\boldsymbol{x}(0)) \tag{3.3.12}$$

且应用（3.3.2）有

$$\boldsymbol{y}(t)\in L(\boldsymbol{v}(t),\boldsymbol{w}(t-1),\cdots,\boldsymbol{w}(0),\boldsymbol{x}(0)) \tag{3.3.13}$$

这引出

$$L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t))\subset L(\boldsymbol{v}(t),\cdots,\boldsymbol{v}(1),\boldsymbol{w}(t-1),\cdots,\boldsymbol{w}(0),\boldsymbol{x}(0)) \tag{3.3.14}$$

由此式、假设 1 和假设 2 有

$$\boldsymbol{w}(t)\perp L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t)) \tag{3.3.15}$$

应用射影公式（3.2.12）及 $\mathrm{E}\boldsymbol{w}(t)=\mathbf{0}$ 可得

$$\mathrm{proj}(\boldsymbol{w}(t)\mid \boldsymbol{y}(1),\cdots,\boldsymbol{y}(t))=\mathbf{0} \tag{3.3.16}$$

于是（3.3.11）成为

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t\mid t) \tag{3.3.17}$$

对（3.3.2）取射影有

$$\hat{\boldsymbol{y}}(t+1\mid t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t+1\mid t)+\mathrm{proj}(\boldsymbol{v}(t+1)\mid \boldsymbol{y}(1),\cdots,\boldsymbol{y}(t)) \tag{3.3.18}$$

由假设 1、2 和（3.3.14）有

$$\boldsymbol{v}(t+1) \perp L(\boldsymbol{y}(1), \cdots, \boldsymbol{y}(t)) \tag{3.3.19}$$

故有

$$\operatorname{proj}(\boldsymbol{v}(t+1) \mid \boldsymbol{y}(1), \cdots, \boldsymbol{y}(t)) = \boldsymbol{0} \tag{3.3.20}$$

于是有

$$\hat{\boldsymbol{y}}(t+1 \mid t) = \boldsymbol{H}\hat{\boldsymbol{x}}(t+1 \mid t) \tag{3.3.21}$$

这引出新息表达式

$$\boldsymbol{\varepsilon}(t+1) = \boldsymbol{y}(t+1) - \hat{\boldsymbol{y}}(t+1 \mid t) = \boldsymbol{y}(t+1) - \boldsymbol{H}\hat{\boldsymbol{x}}(t+1 \mid t) \tag{3.3.22}$$

记滤波和预报估值误差及方差阵为

$$\tilde{\boldsymbol{x}}(t \mid t) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t \mid t) \tag{3.3.23}$$

$$\tilde{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{x}(t+1) - \hat{\boldsymbol{x}}(t+1 \mid t) \tag{3.3.24}$$

$$\boldsymbol{P}(t \mid t) = \mathrm{E}[\tilde{\boldsymbol{x}}(t \mid t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t \mid t)] \tag{3.3.25}$$

$$\boldsymbol{P}(t+1 \mid t) = \mathrm{E}[\tilde{\boldsymbol{x}}(t+1 \mid t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+1 \mid t)] \tag{3.3.26}$$

则由(3.3.2)和(3.3.22)有新息表达式

$$\boldsymbol{\varepsilon}(t+1) = \boldsymbol{H}\tilde{\boldsymbol{x}}(t+1 \mid t) + \boldsymbol{v}(t+1) \tag{3.3.27}$$

且由(3.3.1)和(3.3.17)有

$$\tilde{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{\Phi}\tilde{\boldsymbol{x}}(t \mid t) + \boldsymbol{\Gamma}\boldsymbol{w}(t) \tag{3.3.28}$$

由(3.3.9)有

$$\tilde{\boldsymbol{x}}(t+1 \mid t+1) = \tilde{\boldsymbol{x}}(t+1 \mid t) - \boldsymbol{K}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.3.29}$$

将(3.3.27)代入(3.3.29)引出

$$\tilde{\boldsymbol{x}}(t+1 \mid t+1) = [\boldsymbol{I}_n - K(t+1)\boldsymbol{H}]\tilde{\boldsymbol{x}}(t+1 \mid t) - \boldsymbol{K}(t+1)\boldsymbol{v}(t+1) \tag{3.3.30}$$

其中 $\boldsymbol{I}_n$ 为 $n \times n$ 单位阵.因为

$$\tilde{\boldsymbol{x}}(t \mid t) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t \mid t) \in \mathrm{L}(\boldsymbol{v}(t), \cdots, \boldsymbol{v}(1), \boldsymbol{w}(t-1), \cdots, \boldsymbol{w}(0), \boldsymbol{x}(0)) \tag{3.3.31}$$

故有

$$\boldsymbol{w}(t) \perp \tilde{\boldsymbol{x}}(t \mid t) \tag{3.3.32}$$

这引出

$$\mathrm{E}[\boldsymbol{w}(t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t \mid t)] = \boldsymbol{0} \tag{3.3.33}$$

于是由(3.3.28)得到

$$\boldsymbol{P}(t+1 \mid t) = \boldsymbol{\Phi}\boldsymbol{P}(t \mid t)\boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.3.34}$$

因为

$$\tilde{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{x}(t+1) - \hat{\boldsymbol{x}}(t+1 \mid t) \in L(\boldsymbol{v}(t), \cdots, \boldsymbol{v}(1), \boldsymbol{w}(t), \cdots, \boldsymbol{w}(0), \boldsymbol{x}(0)) \tag{3.3.35}$$

故有

$$\boldsymbol{v}(t+1) \perp \tilde{\boldsymbol{x}}(t+1 \mid t) \tag{3.3.36}$$

这引出

$$\mathrm{E}[\boldsymbol{v}(t+1)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+1 \mid t)] = \boldsymbol{0} \tag{3.3.37}$$

于是由(3.3.27)得到新息方差阵为

$$\mathrm{E}[\boldsymbol{\varepsilon}(t+1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)] = \boldsymbol{H}\boldsymbol{P}(t+1 \mid t)\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R} \tag{3.3.38}$$

且由(3.3.30)可得

$$P(t+1|t+1)=[I_n-K(t+1)H]P(t+1|t)[I_n-K(t+1)H]^{\mathrm{T}}+K(t+1)RK^{\mathrm{T}}(t+1) \tag{3.3.39}$$

下面求 Kalman 滤波器增益 $K(t+1)$. 为此，我们来求 $\mathrm{E}[x(t+1)\varepsilon^{\mathrm{T}}(t+1)]$. 注意，应用(3.3.27)有

$$\mathrm{E}[x(t+1)\varepsilon^{\mathrm{T}}(t+1)]=\mathrm{E}[(\hat{x}(t+1|t)+\tilde{x}(t+1|t))(H\tilde{x}(t+1|t)+v(t+1))^{\mathrm{T}}] \tag{3.3.40}$$

因为由射影正交性有

$$\hat{x}(t+1|t)\perp\tilde{x}(t+1|t) \tag{3.3.41}$$

且注意 $v(t+1)\perp\hat{x}(t+1|t)$，$v(t+1)\perp\tilde{x}(t+1|t)$，于是由(3.3.40)引出

$$\mathrm{E}[x(t+1)\varepsilon^{\mathrm{T}}(t+1)=P(t+1|t)H^{\mathrm{T}} \tag{3.3.42}$$

将上式和(3.3.38)代入(3.3.10)有增益

$$K(t+1)=P(t+1|t)H^{\mathrm{T}}[HP(t+1|t)H^{\mathrm{T}}+R]^{-1} \tag{3.3.43}$$

现在用 $K(t+1)$的表达式(3.3.43)简化 $P(t+1|t+1)$的公式(3.3.39). 暂时略去(3.3.39)右端的时标，将(3.3.43)代入(3.3.39)有

$$\begin{aligned}P(t+1|t+1)=&[I_n-KH]P-PH^{\mathrm{T}}K^{\mathrm{T}}+KHPH^{\mathrm{T}}K^{\mathrm{T}}+KRK^{\mathrm{T}}=\\&[I_n-KH]P-PH^{\mathrm{T}}K^{\mathrm{T}}+K(HPH^{\mathrm{T}}+R)K^{\mathrm{T}}=\\&[I_n-KH]P-PH^{\mathrm{T}}K^{\mathrm{T}}+PH^{\mathrm{T}}K^{\mathrm{T}}=\\&[I_n-KH]P\end{aligned}$$

即

$$P(t+1|t+1)=[I_n-K(t+1)H]P(t+1|t) \tag{3.3.44}$$

上述推导结果可概括为如下定理.

【定理 3.3.1】 (Kalman 滤波器)系统(3.3.1)和(3.3.2)在假设 1～2 下，递推 Kalman 滤波器为

$$\hat{x}(t+1|t+1)=\hat{x}(t+1|t)+K(t+1)\varepsilon(t+1) \tag{3.3.45}$$

$$\hat{x}(t+1|t)=\Phi\hat{x}(t|t) \tag{3.3.46}$$

$$\varepsilon(t+1)=y(t+1)-H\hat{x}(t+1|t) \tag{3.3.47}$$

$$K(t+1)=P(t+1|t)H^{\mathrm{T}}[HP(t+1|t)H^{\mathrm{T}}+R]^{-1} \tag{3.3.48}$$

$$P(t+1|t)=\Phi P(t|t)\Phi^{\mathrm{T}}+\Gamma Q\Gamma^{\mathrm{T}} \tag{3.3.49}$$

$$P(t+1|t+1)=[I_n-K(t+1)H]P(t+1|t) \tag{3.3.50}$$

$$\hat{x}(0|0)=\mu_0,\quad P(0|0)=P_0 \tag{3.3.51}$$

【推论 3.3.1】 在定理 3.3.1 条件下，Kalman 滤波器有另一种递推形式

$$\hat{x}(t+1|t+1)=[I_n-K(t+1)H]\Phi\hat{x}(t|t)+K(t+1)y(t+1) \tag{3.3.52}$$

证明 将(3.3.46)和(3.3.47)代入(3.3.45)得(3.3.52). □

【注 3.3.1】 初值取(3.3.51)是为了保证估值的无偏性. 事实上，将(3.3.46)代入(3.3.45)且与(3.3.1)相减有误差方程

$$\tilde{x}(t+1|t+1)=\Phi\tilde{x}(t|t)+\Gamma w(t)-K(t+1)\varepsilon(t+1) \tag{3.3.53}$$

这引出差分方程

$$\mathrm{E}\tilde{x}(t+1|t+1)=\Phi\mathrm{E}\tilde{x}(t|t) \tag{3.3.54}$$

取初值(3.3.51)有 $\mathrm{E}\tilde{\boldsymbol{x}}(0|0)=\boldsymbol{0}$,这引出 $\mathrm{E}\tilde{\boldsymbol{x}}(t|t)=\boldsymbol{0}(t>0)$,即 $\mathrm{E}\boldsymbol{x}(t)=\mathrm{E}\hat{\boldsymbol{x}}(t|t)$. 这表明估值是无偏的.

【注 3.3.2】 Kalman 滤波器计算程序框图如图 3.3.1 所示.

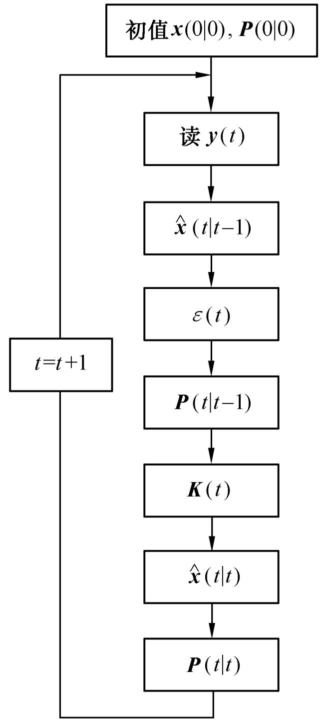


图 3.3.1 Kalman 滤波递推算法框图

【定理 3.3.2】 (Kalman 预报器系统(3.3.1)和(3.3.2)在假设 1~2 下,递推 Kalman 预报器为

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p(t)\boldsymbol{\varepsilon}(t) \tag{3.3.55}$$

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)-\boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1) \tag{3.3.56}$$

$$\boldsymbol{K}_p(t)=\boldsymbol{\Phi}\boldsymbol{K}(t) \tag{3.3.57}$$

或

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{\Phi}-\boldsymbol{K}_p(t)\boldsymbol{H})\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.3.58}$$

其中 $\boldsymbol{K}_p(t)$ 叫 Kalman 预报器增益. 预报误差方差阵满足 Riccati 方程

$$\boldsymbol{P}(t+1|t)=\boldsymbol{\Phi}[\boldsymbol{P}(t|t-1)-\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R})^{-1}\boldsymbol{H}\boldsymbol{P}(t|t-1)]\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.3.59}$$

带初值 $\boldsymbol{P}(1|0)=\boldsymbol{\Phi}\boldsymbol{P}(0|0)\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}$.

证明 将(3.3.45)代入(3.3.46)有

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}[\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}(t)\boldsymbol{\varepsilon}(t)] \tag{3.3.60}$$

这引出(3.3.55)~(3.3.57). 将(3.3.56)代入(3.3.55)得(3.3.58). 将(3.3.48)和(3.3.50)代入(3.3.49)得(3.3.59). □

【定理 3.3.3】 (超前 k 步 Kalman 预报器)系统(3.3.1)和(3.3.2)在假设 1~2 下,超前 $k>1$ 步 Kalman 预报器 $\hat{\boldsymbol{x}}(t+k|t)$ 为

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{\Phi}^{k-1}\hat{\boldsymbol{x}}(t+1|t),\quad k>1 \tag{3.3.61}$$

相应的 k 步预报误差 $\tilde{\boldsymbol{x}}(t+k|t)=\boldsymbol{x}(t+k)-\tilde{\boldsymbol{x}}(t+k|t)$ 方差阵 $\boldsymbol{P}(t+k|t)=\mathrm{E}[\tilde{\boldsymbol{x}}(t+k|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+k|t)]$ 为

$$\boldsymbol{P}(t+k|t)=\boldsymbol{\Phi}^{k-1}\boldsymbol{P}(t+1|t)\boldsymbol{\Phi}^{(k-1)\mathrm{T}}+\sum_{j=2}^{k}\boldsymbol{\Phi}^{k-j}\boldsymbol{\Gamma}Q_w\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Phi}^{(k-j)\mathrm{T}} \tag{3.3.62}$$

其中 $\boldsymbol{P}(t+1|t)$ 由(3.3.59)计算.

证明 由(3.3.1)和射影性质有

$$\hat{\boldsymbol{x}}(t+2|t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{\Gamma}\mathrm{proj}(\boldsymbol{w}(t+1)|\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t)) \tag{3.3.63}$$

由(3.3.14)有

$$\boldsymbol{w}(t+1)\perp L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t)) \tag{3.3.64}$$

这引出 $\mathrm{proj}(\boldsymbol{w}(t+1)|\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t))=\boldsymbol{0}$,因而有

$$\hat{\boldsymbol{x}}(t+2|t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t+1|t) \tag{3.3.65}$$

同理有

$$\hat{\boldsymbol{x}}(t+3|t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t+2|t)=\boldsymbol{\Phi}^2\hat{\boldsymbol{x}}(t+1|t) \tag{3.3.66}$$

用归纳法得(3.3.61).

由(3.3.1)迭代$(k-1)$次有非递推关系

$$\boldsymbol{x}(t+k)=\boldsymbol{\Phi}^{k-1}\boldsymbol{x}(t+1)+\sum_{j=2}^{k}\boldsymbol{\Phi}^{k-j}\boldsymbol{\Gamma}\boldsymbol{w}(t+j-1) \tag{3.3.67}$$

上式与(3.3.61)相减引出

$$\tilde{\boldsymbol{x}}(t+k\mid t)=\boldsymbol{\Phi}^{k-1}\tilde{\boldsymbol{x}}(t+1\mid t)+\sum_{j=2}^{k}\boldsymbol{\Phi}^{k-j}\boldsymbol{\Gamma}\boldsymbol{w}(t+j-1) \tag{3.3.68}$$

因为 $\boldsymbol{w}(t+j-1)\perp\tilde{\boldsymbol{x}}(t+1|t)\ (j\geqslant 2)$,故由(3.3.68)得(3.3.62).证毕. □

定理 3.3.1 可推出到时变系统,即矩阵 $\boldsymbol{\Phi},\boldsymbol{H},\boldsymbol{\Gamma},\boldsymbol{Q},\boldsymbol{R}$ 均为时变的情形,且可带控制项.系统(3.3.1)和(3.3.2)叫定常(非时变)系统.

考虑时变系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}(t)x(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{\Gamma}(t)w(t) \tag{3.3.69}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}(t)\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{3.3.70}$$

【假设 1】 $\mathrm{E}\boldsymbol{w}(t)=\boldsymbol{0},\mathrm{E}\boldsymbol{v}(t)=\boldsymbol{0},\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(j)]=\boldsymbol{0}$,,$\forall\ t,j$,且

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(j)]=\boldsymbol{Q}(t)\delta_{tj},\quad \mathrm{E}[\boldsymbol{v}(t)\boldsymbol{v}^{\mathrm{T}}(j)]=\boldsymbol{R}(t)\delta_{tj} \tag{3.3.71}$$

【假设 2】 $\boldsymbol{x}(0)$不相关于 $\boldsymbol{w}(t)$和 $\boldsymbol{v}(t)$,

$$\mathrm{E}\boldsymbol{x}(0)=\boldsymbol{\mu}_0,\quad \mathrm{E}[(\boldsymbol{x}(0)-\boldsymbol{\mu}_0)(\boldsymbol{x}(0)-\boldsymbol{\mu}_0)^{\mathrm{T}}]=\boldsymbol{P}_0 \tag{3.3.72}$$

【假设 3】 $\boldsymbol{u}(t)\in R^p$ 是已知的控制输入.$\boldsymbol{\Phi}(t),\boldsymbol{\Gamma}(t),\boldsymbol{H}(t),\boldsymbol{B}(t),\boldsymbol{Q}(t)$和 $\boldsymbol{R}(t)$是已知时变矩阵.

【定理 3.3.4】 在上述假设 1 ~ 3 下,时变系统(3.3.69)和(3.3.70)有 Kalman 滤波器

$$\hat{\boldsymbol{x}}(t+1|t)=\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.3.73}$$

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}(t)\hat{\boldsymbol{x}}(t|t)+\boldsymbol{B}(t)\boldsymbol{u}(t) \tag{3.3.74}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H}(t+1)\hat{\boldsymbol{x}}(t+1|t) \tag{3.3.75}$$

$$\boldsymbol{K}(t+1)=\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)[\boldsymbol{H}(t+1)\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)+\boldsymbol{R}(t+1)]^{-1} \tag{3.3.76}$$

$$\boldsymbol{P}(t+1|t)=\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t)\boldsymbol{\Phi}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.3.77}$$

$$\boldsymbol{P}(t+1|t+1)=[\boldsymbol{I}_n-\boldsymbol{K}(t+1)\boldsymbol{H}(t+1)]\boldsymbol{P}(t+1|t) \tag{3.3.78}$$

$$\hat{\boldsymbol{x}}(0|0)=\boldsymbol{\mu}_0,\quad \boldsymbol{P}(0|0)=\boldsymbol{P}_0 \tag{3.3.79}$$

证明 完全平行于定理 3.3.1,从略. □

【定理 3.3.5】 (Kalman 滤波与递推最小二乘法关系)AR(n)模型未知参数的 Kalman 滤波估值等价于递推最小二乘法估值.

证明 考虑 AR(n)模型

$$y(t)=a_1y(t-1)+\cdots+a_ny(t-n)+v(t) \tag{3.3.80}$$

其中 $v(t)$是零均值、方差为 σ_v^2 的白噪声,参数 $a_1,\cdots,a_n$ 是未知的.引入未知参数向量

$$\boldsymbol{\theta}=[a_1,\cdots,a_n]^{\mathrm{T}} \tag{3.3.81}$$

并引入向量

$$\boldsymbol{\varphi}^{\mathrm{T}}(t)=[y(t-1),\cdots,y(t-n)] \tag{3.3.82}$$

则(3.3.80)有状态空间模型

$$\boldsymbol{\theta}(t+1)=\boldsymbol{\theta}(t) \tag{3.3.83}$$

$$y(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}+v(t) \tag{3.3.84}$$

应用定理 3.3.4 有 $\boldsymbol{\theta}$ 的 Kalman 滤波估值器

$$\hat{\boldsymbol{\theta}}(t+1|t+1)=\hat{\boldsymbol{\theta}}(t+1|t)+\boldsymbol{K}(t+1)[y(t+1)-\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}(t+1|t)] \tag{3.3.85}$$

$$\boldsymbol{K}(t+1)=\boldsymbol{P}(t+1|t)\boldsymbol{\varphi}(t+1)[\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t+1|t)\boldsymbol{\varphi}(t+1)+\sigma_v^2]^{-1} \tag{3.3.86}$$

$$\hat{\boldsymbol{\theta}}(t+1|t)=\hat{\boldsymbol{\theta}}(t|t) \tag{3.3.87}$$

$$\boldsymbol{P}(t+1|t)=\boldsymbol{P}(t|t),$$

$$\boldsymbol{P}(t+1|t+1)=[\boldsymbol{I}_n-\boldsymbol{K}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)]\boldsymbol{P}(t+1|t) \tag{3.3.88}$$

引入记号

$$\boldsymbol{P}(t)=\boldsymbol{P}(t|t)/\sigma_v^2,\quad \hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t|t) \tag{3.3.89}$$

则有 RLS 算法

$$\hat{\boldsymbol{\theta}}(t+1)=\hat{\boldsymbol{\theta}}(t)+\boldsymbol{K}(t+1)[y(t+1)-\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\hat{\boldsymbol{\theta}}(t)] \tag{3.3.90}$$

$$\boldsymbol{K}(t+1)=\frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{3.3.91}$$

$$\boldsymbol{P}(t+1)=[\boldsymbol{I}_n-\boldsymbol{K}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)]\boldsymbol{P}(t) \tag{3.3.92}$$

或等价地有

$$\boldsymbol{P}(t+1)=\boldsymbol{P}(t)-\frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t+1)\boldsymbol{P}(t)\boldsymbol{\varphi}(t+1)} \tag{3.3.93}$$

这表明由 Kalman 滤波算法引出了 RLS 算法. 此外, 由(3.3.89)看到, 在 RLS 算法中 $\boldsymbol{P}(t)$ 的意义是它刻划了估值误差方差阵. 注意, 实现 Kalman 滤波算法要求已知噪声方差 σ_v^2, 而应用 RLS 算法不要求已知 σ_v^2. □

【例 3.3.1】 考虑随机系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}0.1 & 0.1 & 0\\ 2 & 0.2 & 0\\ 3 & 1 & 0.8\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\2\\3\end{bmatrix}w(t) \tag{3.3.94}$$

$$y(t)=[1\quad 1\quad 1]\boldsymbol{x}(t)+v(t) \tag{3.3.95}$$

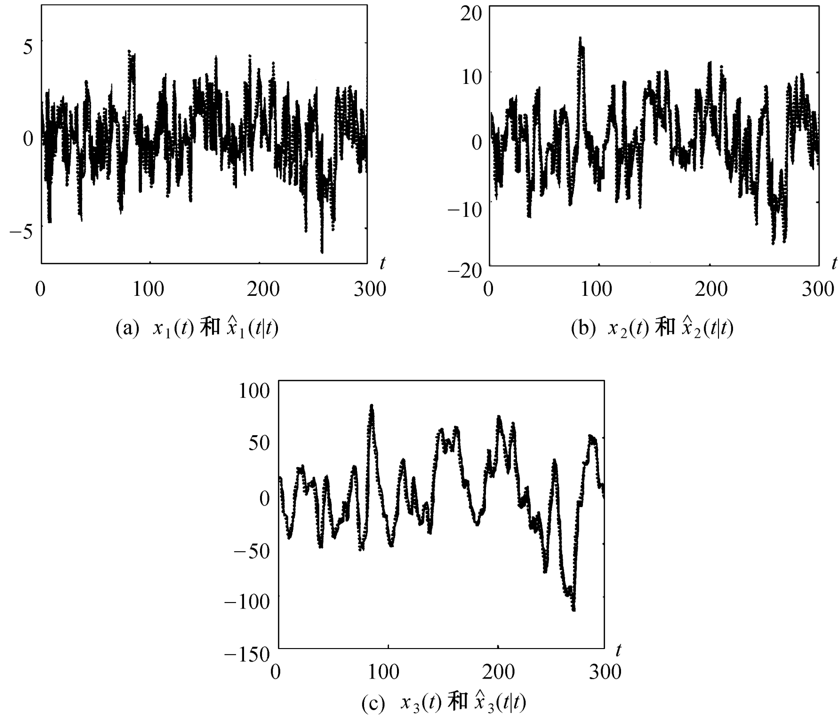
其中状态 $\boldsymbol{x}(t)=[x_1(t),x_2(t),x_3(t)]^{\mathrm{T}}$, $w(t)$和 $v(t)$是零均值、方差各为 $Q=0.81$ 和 $R=1$的独立高斯白噪声. 应用定理 3.3.1 最优 Kalman 滤波器 $\hat{\boldsymbol{x}}(t|t)$的仿真结果如图 3.3.2 所示, 其中实线表示真实状态 $x_i(t)$, 虚线表示 Kalman 滤波估值 $\hat{x}_i(t|t)$.

【例 3.3.2】 考虑系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}0.6 & -0.4\\ 0.2 & 1.1\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\2\end{bmatrix}w(t) \tag{3.3.96}$$

$$y(t)=[1\quad 1]\boldsymbol{x}(t)+v(t) \tag{3.3.97}$$

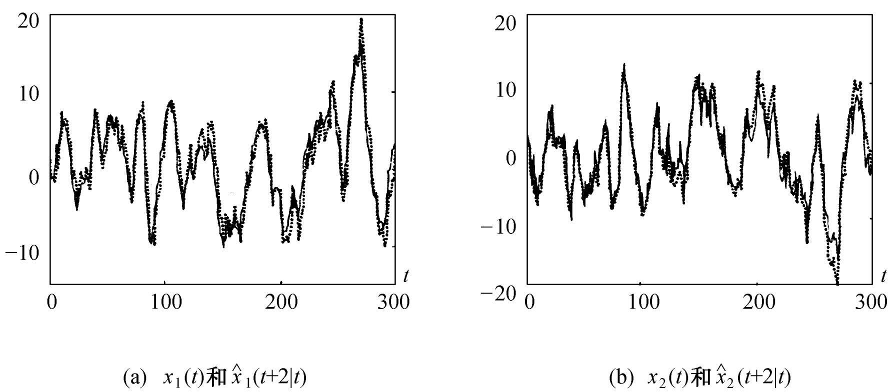
其中 $\boldsymbol{x}(t)=[x_1(t),x_2(t)]^{\mathrm{T}}$, $w(t)$和 $v(t)$是零均值、方差各为 $Q=0.64$ 和 $R=1$ 的独立白噪声. 应用定理 3.3.3 可求最优 Kalman 滤波器 $\hat{\boldsymbol{x}}(t+2|t)=[\hat{x}_1(t+2|t),\hat{x}_2(t+2|t)]^{\mathrm{T}}$. 仿真结果如图 3.3.3 所示, 其中实线代表真实值, 虚线代表预报估值.



(a) $x_1(t)$ 和 $\hat{x}_1(t|t)$　　(b) $x_2(t)$ 和 $\hat{x}_2(t|t)$

(c) $x_3(t)$ 和 $\hat{x}_3(t|t)$

图 3.3.2　状态 $\hat{\boldsymbol{x}}_i(t)$ 和最优 Kalman 滤波器 $\hat{\boldsymbol{x}}_i(t|t)$



(a) $x_1(t)$和$\hat{x}_1(t+2|t)$　　(b) $x_2(t)$和$\hat{x}_2(t+2|t)$

图 3.3.3　状态 $\boldsymbol{x}_i(t)$ 和最优 Kalman 预报器 $\hat{\boldsymbol{x}}_i(t+2|t)$

3.4　Kalman 平滑器

对于系统(3.3.1)和(3.3.2)在假设 1～2 下，寻求 Kalman 平滑估值 $\hat{\boldsymbol{x}}(j|t)$ $(j>t)$ 问题在许多应用问题中会遇到. 例如发射导弹的初速度估计问题、化学反应过程溶液初始浓度估计问题、发射人造地球卫星入轨初速度估计问题等，可归结为求平滑估值 $\hat{\boldsymbol{x}}(t_0|t_0+N)$，其中初始时刻 t_0 固定，但 N 是变化的，$N=1,2,\cdots$，这类平滑器叫固定点 Kalman 平滑器. 它是基于时刻 t_0+N 为止的观测信息求在时刻 t_0 处的状态 $\boldsymbol{x}(t_0)$ 的最优估值. t_0 叫固

定点.在某些应用问题中要求求平滑估值 $\hat{\boldsymbol{x}}(t|t+N)$，其中 t 是变化的，但 N 是固定的正整数.这类平滑器叫固定滞后 Kalman 平滑器，N 叫固定滞后.对于带有观测滞后 N 的系统，固定滞后平滑器 $\hat{\boldsymbol{x}}(t|t+N)$ 可改善 Kalman 滤波器 $\hat{\boldsymbol{x}}(t|t)$ 的精度.发射人造地球卫星后，需根据卫星环绕地球的轨道观测数据，重新更精确估算卫星的轨道，以便研究与预定轨道的偏差.这类问题叫轨道重构或叫固定区间平滑问题，即求平滑估值 $\hat{\boldsymbol{x}}(t|N)$，其中 N 是固定的，叫固定区间长度，$t=0,1,\cdots,N$ 是变化的，并称 $\hat{\boldsymbol{x}}(t|N)$ 为固定区间 Kalman 平滑器.下面统一用记号 $\hat{\boldsymbol{x}}(t|t+N)$，$(N>0)$ 表示固定滞后 Kalman 平滑器（N 固定，t 变化）或固定点 Kalman 平滑器（t 固定，N 变化），而用 $\hat{\boldsymbol{x}}(t|N)$ 表示固定区间 Kalman 平滑器（N 固定，t 变化，$t\leqslant N$）.

本节沿用节 3.3 的记号和公式.

【定理 3.4.1】 系统(3.3.1)和(3.3.2)在假设 1 ~ 2 下，最优递推固定点 Kalman 平滑器为

$$\hat{\boldsymbol{x}}(t|t+N)=\hat{\boldsymbol{x}}(t|t+N-1)+\boldsymbol{K}(t|t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.4.1}$$

带初值 $\hat{\boldsymbol{x}}(t|t)$，$N=1,2,\cdots$，且平滑增益 $\boldsymbol{K}(t|t+N)$ 为

$$\boldsymbol{K}(t|t+N)=\boldsymbol{P}(t|t-1)\left\{\prod_{i=0}^{N-1}[\boldsymbol{I}_n-\boldsymbol{K}(t+i)\boldsymbol{H}]^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}}\right\}\times \boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{P}(t+N|t+N-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}]^{-1} \tag{3.4.2}$$

其中 $\boldsymbol{\varepsilon}(t+N)$ 由 Kalman 滤波器计算，$\boldsymbol{K}(t+i)$ 为 Kalman 滤波器增益，$\boldsymbol{P}(t|t-1)$ 为 Kalman 预报误差方差阵，$\hat{\boldsymbol{x}}(t|t)$ 为 Kalman 滤波估值.且规定

$$\boldsymbol{K}(t|t)=\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}]=\boldsymbol{K}(t) \tag{3.4.3}$$

估值误差 $\tilde{\boldsymbol{x}}(t|t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t+N)$ 的方差阵 $\boldsymbol{P}(t|t+N)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t+N)]$ 可递推计算为

$$\boldsymbol{P}(t|t+N)=\boldsymbol{P}(t|t+N-1)-\boldsymbol{K}(t|t+N)[\boldsymbol{H}\boldsymbol{P}(t+N|t+N-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}]\boldsymbol{K}^{\mathrm{T}}(t|t+N) \tag{3.4.4}$$

带初值 $\boldsymbol{P}(t|t)$. $\boldsymbol{P}(t|t)$ 为 Kalman 滤波误差方差阵.

证明 应用递推射影公式(3.2.37)得到(3.4.1)，其中平滑增益为

$$\boldsymbol{K}(t|t+N)=\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)][\mathrm{E}(\boldsymbol{\varepsilon}(t+N)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N))]^{-1} \tag{3.4.5}$$

由(3.3.27)有

$$\boldsymbol{\varepsilon}(t+N)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t+N|t+N-1)+\boldsymbol{v}(t+N) \tag{3.4.6}$$

而由(3.3.38)有

$$\mathrm{E}[\boldsymbol{\varepsilon}(t+N)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]=\boldsymbol{H}\boldsymbol{P}(t+N|t+N-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R} \tag{3.4.7}$$

因此求平滑增益 $\boldsymbol{K}(t|t+N)$ 归结为求 $\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]$.

将(3.3.27)代入(3.3.55)后，并与(3.3.1)相减引出误差系统

$$\tilde{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{\Gamma}\boldsymbol{w}(t)-\boldsymbol{K}_p(t)\boldsymbol{v}(t) \tag{3.4.8}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}-\boldsymbol{K}_p(t)\boldsymbol{H}=\boldsymbol{\Phi}-\boldsymbol{\Phi}\boldsymbol{K}(t)\boldsymbol{H}=\boldsymbol{\Phi}[\boldsymbol{I}_n-\boldsymbol{K}(t)\boldsymbol{H}] \tag{3.4.9}$$

其中应用了关系(3.3.57).由(3.4.3)迭代 N 次有关系

$$\tilde{\boldsymbol{x}}(t+N|t+N-1)=\boldsymbol{\Psi}(t+N,t)\tilde{\boldsymbol{x}}(t|t-1)+\sum_{i=t+1}^{t+N}\boldsymbol{\Psi}(t+N,i)\times[\boldsymbol{\Gamma}\boldsymbol{w}(i-1)-\boldsymbol{K}_p(i-1)\boldsymbol{v}(i-1)] \tag{3.4.10}$$

其中定义 $\boldsymbol{\Psi}(t+N,t+N)=\boldsymbol{I}_n$，

$$\boldsymbol{\Psi}(t+N,i)=\boldsymbol{\Psi}_p(t+N-1)\cdots\boldsymbol{\Psi}_p(i),\quad i<t+N \tag{3.4.11}$$

应用(3.4.4)有

$$\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]=\mathrm{E}[\boldsymbol{x}(t)(\boldsymbol{H}\tilde{\boldsymbol{x}}(t+N|t+N-1)+\boldsymbol{v}(t+N))^{\mathrm{T}}] \tag{3.4.12}$$

注意关系

$$\boldsymbol{x}(t)=\hat{\boldsymbol{x}}(t|t-1)+\tilde{\boldsymbol{x}}(t|t-1) \tag{3.4.13}$$

$$\hat{\boldsymbol{x}}(t|t-1)\perp\tilde{\boldsymbol{x}}(t|t-1) \tag{3.4.14}$$

$$\tilde{\boldsymbol{x}}(t|t-1)\perp\boldsymbol{w}(j),\quad \tilde{\boldsymbol{x}}(t|t-1)\perp\boldsymbol{v}(j),\quad j\geqslant t \tag{3.4.15}$$

将(3.4.8)和(3.4.12)代入(3.4.11)可得

$$\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]=\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}^{\mathrm{T}}(t+N,t)\boldsymbol{H}^{\mathrm{T}} \tag{3.4.16}$$

应用(3.4.11)引出

$$\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]=\boldsymbol{P}(t|t-1)\left\{\prod_{i=0}^{N-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\right\}\boldsymbol{H}^{\mathrm{T}} \tag{3.4.17}$$

将(3.4.17)和(3.4.5)代入(3.4.3)得(3.4.2).

由(3.4.1)引出关系

$$\tilde{\boldsymbol{x}}(t|t+N)=\tilde{\boldsymbol{x}}(t|t+N-1)-\boldsymbol{K}(t|t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.4.18}$$

$$\tilde{\boldsymbol{x}}(t|t+N-1)=\tilde{\boldsymbol{x}}(t|t+N)+\boldsymbol{K}(t|t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.4.19}$$

由射影的正交性有

$$\tilde{\boldsymbol{x}}(t|t+N)\perp L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t+N)) \tag{3.4.20}$$

而 $L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t+N))=L(\boldsymbol{\varepsilon}(1),\cdots,\boldsymbol{\varepsilon}(t+N))$,故有

$$\tilde{\boldsymbol{x}}(t|t+N)\perp\boldsymbol{\varepsilon}(t+N) \tag{3.4.21}$$

因而由(3.4.19)得

$$\boldsymbol{P}(t|t+N-1)=\boldsymbol{P}(t|t+N)+\boldsymbol{K}(t|t+N)\mathrm{E}[\boldsymbol{\varepsilon}(t+N)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]\boldsymbol{K}^{\mathrm{T}}(t|t+N) \tag{3.4.22}$$

将(3.4.6)代入(3.4.22)引出(3.4.3).证毕. □

【定理 3.4.2】 系统(3.3.1)和(3.3.2)在假设 1~2 下,有最优固定滞后 Kalman 平滑器

$$\hat{\boldsymbol{x}}(t|t+N)=\hat{\boldsymbol{x}}(t|t-1)+\sum_{i=0}^{N}\boldsymbol{K}(t|t+i)\boldsymbol{\varepsilon}(t+i),\quad N>0 \tag{3.4.23}$$

其中定义 $\boldsymbol{K}(t|t)=\boldsymbol{K}(t)$,$\boldsymbol{K}(t)$为 Kalman 滤波器增益,而 $\boldsymbol{K}(t|t+i)$由(3.4.2)计算.平滑误差方差阵 $\boldsymbol{P}(t|t+N)$为

$$\boldsymbol{P}(t|t+N)=\boldsymbol{P}(t|t-1)-\sum_{i=0}^{N}\boldsymbol{K}(t|t+i)[\boldsymbol{H}\boldsymbol{P}(t+i|t+i-1)\boldsymbol{H}^{\mathrm{T}}\times +\boldsymbol{R}]\boldsymbol{K}^{\mathrm{T}}(t|t+i) \tag{3.4.24}$$

证明 (3.4.23)是(3.4.1)的非递推形式.类似于(3.4.4)的证明可证(3.4.24). □

【定理 3.4.3】 系统(3.3.1)和(3.3.2)在假设 1~2 下,有固定区间 Kalmwn 平滑器为

$$\hat{\boldsymbol{x}}(t|N)=\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{P}(t|t-1)\boldsymbol{r}(t|N) \tag{3.4.25}$$

其中 $t=1,2,\cdots,N$,且有反向递推公式

$$\boldsymbol{r}(t|N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\boldsymbol{r}(t+1|N)+\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}]^{-1}\boldsymbol{\varepsilon}(t) \tag{3.4.26}$$

其中 $t=N,N-1,\cdots,1$,且规定 $\boldsymbol{r}(N+1|N)=\boldsymbol{0}$.平滑误差 $\tilde{\boldsymbol{x}}(t|N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|N)$的方差阵 $\boldsymbol{P}(t|N)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|N)]$可递推计算为

$$\boldsymbol{P}(t|N)=\boldsymbol{P}(t|t-1)-\boldsymbol{P}(t|t-1)\boldsymbol{S}(t|N)\boldsymbol{P}(t|t-1) \tag{3.4.27}$$

带 $t=N,N-1,\cdots,1$，且有反向递推公式

$$\boldsymbol{S}(t|N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\boldsymbol{S}(t+1|N)\boldsymbol{\Psi}_p(t)+\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{HP}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}]^{-1}\boldsymbol{H} \tag{3.4.28}$$

带 $t=N,N-1,\cdots,1$，且规定 $\boldsymbol{S}(N+1|N)=\boldsymbol{0}$.

证明 注意 $\hat{\boldsymbol{x}}(t|N)=\hat{\boldsymbol{x}}(t|t+N-t)$，应用(3.4.23)有

$$\hat{\boldsymbol{x}}(t\mid N)=\hat{\boldsymbol{x}}(t\mid t-1)+\sum_{i=0}^{N-t}\boldsymbol{K}(t\mid t+i)\boldsymbol{\varepsilon}(t+i) \tag{3.4.29}$$

简记新息 $\boldsymbol{\varepsilon}(t)$ 的方差阵(3.3.38)为 $\boldsymbol{Q}_{\varepsilon}(t)$，即

$$\boldsymbol{Q}_{\varepsilon}(t)=\boldsymbol{HP}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R} \tag{3.4.30}$$

将(3.4.2)和(3.4.3)代入(3.4.29)有

$$\hat{\boldsymbol{x}}(t\mid N)=\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{P}(t\mid t-1)\Big\{\sum_{i=1}^{N-t}\Big[\prod_{j=0}^{i-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+j)\Big]\times \\ \boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+i)\boldsymbol{\varepsilon}(t+i)+\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t)\boldsymbol{\varepsilon}(t)\Big\} \tag{3.4.31}$$

定义 $\boldsymbol{r}(t|N)$ 为

$$\boldsymbol{r}(t\mid N)=\sum_{i=1}^{N-t}\Big[\prod_{j=0}^{i-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+j)\Big]\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+i)\boldsymbol{\varepsilon}(t+i)+\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t)\boldsymbol{\varepsilon}(t) \tag{3.4.32}$$

则(3.4.31)成为

$$\hat{\boldsymbol{x}}(t|N)=\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{P}(t|t-1)\boldsymbol{r}(t|N) \tag{3.4.33}$$

它即(3.4.25). 下证递推公式(3.4.26). 事实上，在(3.4.32)中置 $t=t+1$ 有

$$\boldsymbol{r}(t+1\mid N)=\sum_{i=1}^{N-t-1}\Big[\prod_{j=0}^{i-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1+j)\Big]\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1+i)\boldsymbol{\varepsilon}(t+1+i)+ \\ \boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.4.34}$$

而(3.4.32)可写为

$$\begin{aligned}\boldsymbol{r}(t\mid N)&=\sum_{i=2}^{N-t}\Big[\prod_{j=0}^{i-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+j)\Big]\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+i)\boldsymbol{\varepsilon}(t+i)+ \\ &\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\boldsymbol{\varepsilon}(t+1)+\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t)\boldsymbol{\varepsilon}(t)= \\ &\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\Big\{\sum_{i=2}^{N-t}\Big[\prod_{j=1}^{i-1}\boldsymbol{\Psi}_p(t+j)\Big]\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+i)\boldsymbol{\varepsilon}(t+i)+\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\boldsymbol{\varepsilon}(t+1)\Big\}+ \\ &\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t)\boldsymbol{\varepsilon}(t)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\sum_{i=1}^{N-t-1}\Big[\prod_{j=0}^{i-1}\boldsymbol{\Psi}_p(t+1+j)\Big]\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1+i)\boldsymbol{\varepsilon}(t+1+i)+ \\ &\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\boldsymbol{\varepsilon}(t+1)+\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t)\boldsymbol{\varepsilon}(t)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\boldsymbol{r}(t+1|N)+\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t)\boldsymbol{\varepsilon}(t)\end{aligned} \tag{3.4.35}$$

即(3.4.26)成立. 由(3.4.33)引出

$$\tilde{\boldsymbol{x}}(t|t-1)=\tilde{\boldsymbol{x}}(t|N)+\boldsymbol{P}(t|t-1)\boldsymbol{r}(t|N) \tag{3.4.36}$$

因为由(3.4.32)有

$$\boldsymbol{r}(t|N)\in L(\boldsymbol{\varepsilon}(t),\boldsymbol{\varepsilon}(t+1),\cdots,\boldsymbol{\varepsilon}(N)) \tag{3.4.37}$$

而由射影正交性有

$$\tilde{\boldsymbol{x}}(t|N)\perp L(\boldsymbol{\varepsilon}(1),\boldsymbol{\varepsilon}(2),\cdots,\boldsymbol{\varepsilon}(N)) \tag{3.4.38}$$

因此有

$$\boldsymbol{r}(t|N) \perp \tilde{\boldsymbol{x}}(t|N) \tag{3.4.39}$$

于是由(3.4.36)得

$$\boldsymbol{P}(t|t-1) = \boldsymbol{P}(t|N) + \boldsymbol{P}(t|t-1)\boldsymbol{S}(t|N)\boldsymbol{P}(t|t-1) \tag{3.4.40}$$

其中定义

$$\boldsymbol{S}(t|N) = \mathrm{E}[\boldsymbol{r}(t|N)\boldsymbol{r}^{\mathrm{T}}(t|N)] \tag{3.4.41}$$

(3.4.40)引出(3.4.27)成立.应用(3.4.26)有递推关系

$$\boldsymbol{S}(t|N) = \boldsymbol{\Psi}_p^{\mathrm{T}}(t)\boldsymbol{S}(t+1|N)\boldsymbol{\Psi}_p^{\mathrm{T}}(t) + \boldsymbol{H}[\boldsymbol{HP}(t|t-1)\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}]^{-1}\boldsymbol{H}^{\mathrm{T}} \tag{3.4.42}$$

其中应用了由(3.4.34)引出的事实

$$\boldsymbol{\varepsilon}(t) \perp \boldsymbol{r}(t+1|N) \tag{3.4.43}$$

在(3.4.25)中取 $t = N+1$ 引出应规定

$$\boldsymbol{r}(N+1|N) = \boldsymbol{0} \tag{3.4.44}$$

在(3.4.41)中置 $t = N+1$,引出应规定

$$\boldsymbol{S}(N+1|N) = \boldsymbol{0} \tag{3.4.45}$$

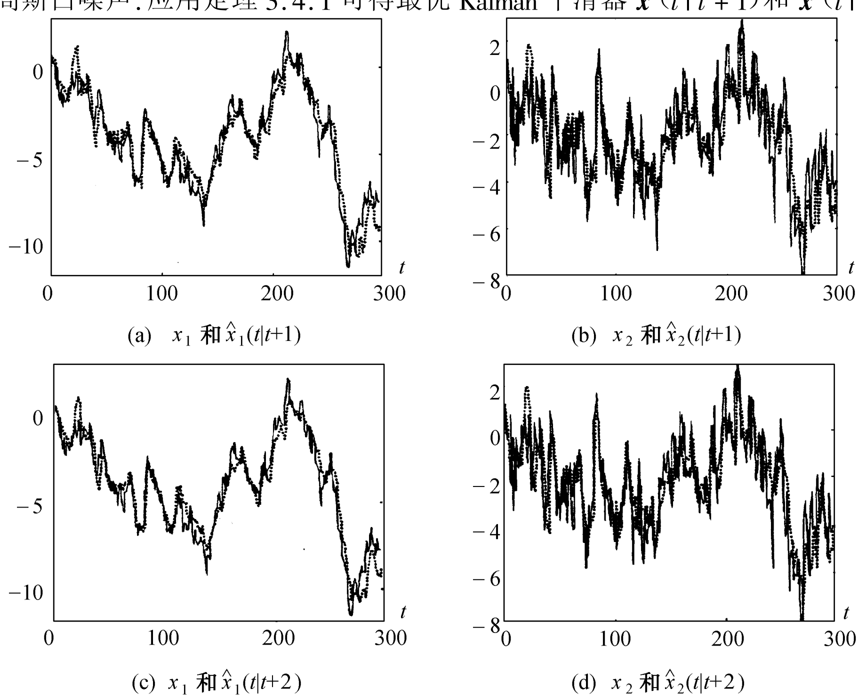
(3.4.44)也可由在(3.4.25)中置 $t = N$ 导出,此时应有 $\boldsymbol{r}(N|N) = \boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(N)\boldsymbol{\varepsilon}(N)$,这相当于在(3.4.26)中置 $\boldsymbol{r}(N+1|N) = \boldsymbol{0}$. □

【例 3.4.1】 考虑随机系统

$$\boldsymbol{x}(t+1) = \begin{bmatrix} 1 & 0 \\ 0.2 & 0.6 \end{bmatrix}\boldsymbol{x}(t) + \begin{bmatrix} 1 \\ 2 \end{bmatrix} w(t) \tag{3.4.46}$$

$$y(t) = [1 \quad 1]x(t) + v(t) \tag{3.4.47}$$

其中 $\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}$,$w(t)$ 和 $v(t)$ 是零均值、方差各为 $Q = 0.4$ 和 $R = 10$ 的独立的高斯白噪声.应用定理 3.4.1 可得最优 Kalman 平滑器 $\hat{\boldsymbol{x}}(t|t+1)$ 和 $\hat{\boldsymbol{x}}(t|t+2)$.仿



(a) x_1 和 $\hat{x}_1(t|t+1)$　(b) x_2 和 $\hat{x}_2(t|t+1)$

(c) x_1 和 $\hat{x}_1(t|t+2)$　(d) x_2 和 $\hat{x}_2(t|t+2)$

图 3.4.1 状态 $\hat{\boldsymbol{x}}_i(t)$ 和最优 Kalman 平滑器 $\hat{\boldsymbol{x}}_i(t|t+1)$ 与 $\hat{\boldsymbol{x}}_i(t|t+2)$

真结果如图 3.4.1 所示，其中实线代表真实值 $x_i(t)$，虚线代表其他的最优 Kalman 平滑器 $\hat{x}_i(t|t+1)$ 和 $\hat{x}_i(t|t+2)$，$i=1,2$.

3.5 白噪声估值器及其在信号处理中的应用

白噪声估计理论的一个重要应用背景是石油地震勘探信号处理问题.这一问题曾被美国学者 Mendel[10,11]深入研究.石油地震勘探原理是利用埋在地表下的炸药爆炸后产生的地震波在油层的反射系数序列提供的信息来判断是否有油田及油田几何形状大小.因为反射系数序列可用 Bernoulli - Gaussian 白噪声来表示，因而白噪声估计问题成为地震勘探的关键问题.Mendel[10,11]用 Kalman 滤波方法解决这个问题，提出了系统的输入白噪声估值器，也叫白噪声反卷积(Deconvolution)滤波器.但 Mendel 的白噪声估计理论的局限性是没有解决系统的观测白噪声估计问题，因而不能构成统一的白噪声估计理论，且不能用于解决状态估计问题.为了克服这一局限性，新近作者在文献[8,12,13]中系统地提出了基于 Kalman 滤波方法的统一的和通用的白噪声估计理论，其中解决了观测白噪声估计问题，并将其应用于解决状态和信号估计问题[8].

【例 3.5.1】 在石油地震勘探中的输入白噪声估计问题.

石油地震勘探原理如图 3.5.1，图 3.5.2 及图 3.5.3 所示.炸药埋在地表下爆炸后产生的反射地震波信号 $s(t)$ 由地面上的传感器接收，接收信号为 $y(t)$，其中被观测噪声 $v(t)$ 污染，假设 $v(t)$ 是零均值、方差为 σ_v^2 的白噪声，可用如下卷积模型描写地震勘探系统，[10,11]即

$$s(t) = \sum_{j=1}^{n} h_j w(t-j) \tag{3.5.1}$$

$$y(t) = s(t) + v(t) \tag{3.5.2}$$

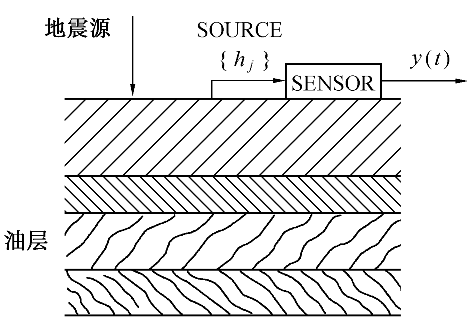


图 3.5.1 石油地震勘探原理

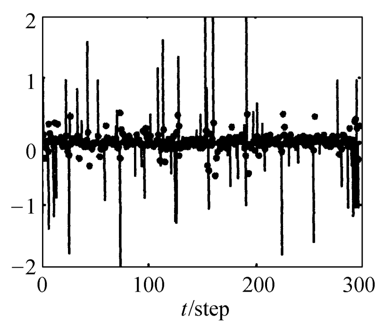


图 3.5.2 反射系数序列——Bernoulli - Gaussian 白噪声 $w(t)$

其中序列 h_j 是与传感器有关的脉冲响应序列，而 $w(t)$ 是油层的“反射系数序列”.通常 $w(t)$ 可用如下 Bernoulli - Gaussian 白噪声来描写

$$w(t) = b(t)g(t) \tag{3.5.3}$$

其中 $b(t)$ 为取值 1 和 0 的 Bernoulli 白噪声，取值概率为

$$P(b(t)=1)=\lambda, \quad P(b(t)=0)=1-\lambda, \quad 0<\lambda<1 \tag{3.5.4}$$

而 $g(t)$ 是零均值、方差为 σ_g^2 的独立于 $b(t)$ 的 Gaussian 白噪声. $w(t)$ 的图形如图 3.5.2 所示,其中实线端点纵坐标为 $w(t)$. 可看到 $w(t)$ 只在个别处不为零,且非零 $w(t)$ 的幅度和非零值出现的时刻 t 都是随机的. 由 (3.5.1) 知 $w(t)$ 取非零值的概率为 λ. 显然 $w(t)$ 有零均值,注意 $b(t)$ 的方差为 λ,故 $w(t)$ 的方差为 $\sigma_w^2 = \lambda\sigma_g^2$.

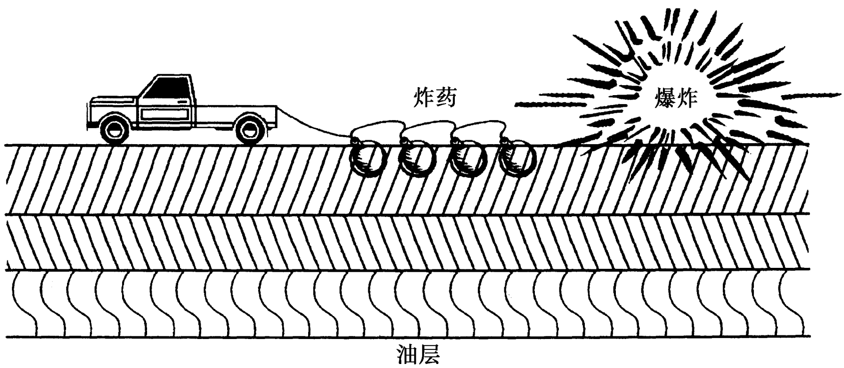


图 3.5.3　石油地震勘探作业

反射系数序列 $w(t)$ 提供了判断是否有油层及油层几何形状的重要信息. 因此石油地震勘探的目的是估计反射系数序列 $w(t)$,特别感兴趣估计 $w(t)$ 的非零值发生的时刻 t 和幅值. 在图 3.5.2 中用实圆点纵坐标表示 $w(t)$ 的估值. 这归结为由观测信号 $y(t)$ 估计系统 (3.5.1) 和 (3.5.2) 的输入白噪声的问题,即要求解决输入估计问题. 因 (3.5.1) 为卷积模型,故估计输入 $w(t)$ 也叫"反卷积"(Deconvolution).

注意系统 (3.5.1) 和 (3.5.2) 有状态空间模型

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi x}(t) + \boldsymbol{\Gamma} w(t) \tag{3.5.5}$$

$$y(t) = \boldsymbol{Hx}(t) + v(t) \tag{3.5.6}$$

其中定义 $\boldsymbol{x}(t) = [w(t-1), \cdots, w(t-n)]^{\mathrm{T}}$,T 为转置号,且 $s(t) = \boldsymbol{Hx}(t)$,

$$\boldsymbol{\Phi} = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots & \\ 0 & \cdots & \cdots & 1 & 0 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{H} = [h_1, \cdots h_n] \tag{3.5.7}$$

因此问题也可转化为估计状态空间模型 (3.5.5) 和 (3.5.6) 的输入白噪声 $w(t)$.

推广反卷积模型 (3.5.1) 和 (3.5.2),可假设如下一般的反卷积模型[19]

$$s(t) + a_1 s(t-1) + \cdots + a_n s(t-n) = c_1 w(t-1) + \cdots + c_n w(t-n) \tag{3.5.8}$$

$$y(t) = s(t) + v(t) \tag{3.5.9}$$

由定理 1.6.1 它也可写成状态空间模型 (3.5.5) 和 (3.5.6) 形式. 其中 $s(t) = \boldsymbol{Hx}(t)$,

$$\boldsymbol{\Phi} = \begin{bmatrix} -a_1 & & & \\ -a_2 & & \boldsymbol{I}_{n-1} & \\ \vdots & & & \\ -a_n & 0 & \cdots & 0 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, \quad \boldsymbol{H} = [1, 0, \cdots, 0] \tag{3.5.10}$$

可用图 3.5.4 表示石油地震勘探白噪声反射系数序列最优估计问题.

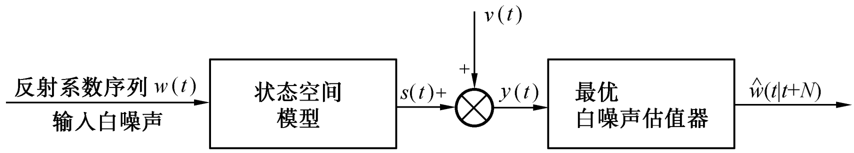


图 3.5.4 石油地震勘探白噪声反射系数序列最优估计问题

【例 3.5.2】 带白色观测噪声的 ARMA 信号最优滤波器和平滑器设计问题.

考虑典型的带白色观测噪声的多通道 ARMA 信号 Wiener 滤波问题:

$$\boldsymbol{s}(t)+\boldsymbol{A}_1\boldsymbol{s}(t-1)+\cdots+\boldsymbol{A}_n\boldsymbol{s}(t-n)=\boldsymbol{C}_1\boldsymbol{w}(t-1)+\cdots+\boldsymbol{C}_n\boldsymbol{w}(t-n) \tag{3.5.11}$$

$$\boldsymbol{y}(t)=\boldsymbol{s}(t)+\boldsymbol{v}(t) \tag{3.5.12}$$

其中 $\boldsymbol{s}(t)\in R^m$ 为待估信号,$\boldsymbol{y}(t)\in R^m$ 为对 $s(t)$的观测信号,$\boldsymbol{v}(t)\in R^m$ 为零均值、方差阵为 $\boldsymbol{R}$ 的白色观测噪声,$\boldsymbol{w}(t)\in R^r$ 是零值、方差阵为 $\boldsymbol{Q}$ 的独立于 $\boldsymbol{v}(t)$的输入白噪声,$\boldsymbol{A}_i\in R^{m\times m}$,$\boldsymbol{C}_i\in R^{m\times r}$是模型参数阵.问题是基于观测$(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots,\boldsymbol{y}(1))$求信号 $\boldsymbol{s}(t)$的最优滤波器和平滑器 $\hat{\boldsymbol{s}}(t|t+N)(N\geqslant 0)$.

对(3.5.12)两边的项取射影运算有

$$\boldsymbol{y}(\boldsymbol{t})=\hat{\boldsymbol{s}}(t|t+N)+\hat{\boldsymbol{v}}(t|t+N) \tag{3.5.13}$$

其中 $\hat{\boldsymbol{v}}(t|t+N)$为最优观测白噪声估值器,即

$$\hat{\boldsymbol{v}}(t|t+N)=\mathrm{proj}(\boldsymbol{v}(t)|\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t+N)) \tag{3.5.14}$$

因而有 $\boldsymbol{s}(t)$的最优滤波器和平滑器为

$$\hat{\boldsymbol{s}}(t|t+N)=\boldsymbol{y}(t)-\hat{\boldsymbol{v}}(t|t+N) \tag{3.5.15}$$

于是问题归结为求观测白噪声估值器 $\hat{\boldsymbol{v}}(t|t+N)$.

由定理 1.6.1,系统(3.5.11)和(3.5.12)可表为等价的状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}\boldsymbol{x}(t)+\boldsymbol{\Gamma}w(t) \tag{3.5.16}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{3.5.17}$$

其中定义

$$\boldsymbol{\Phi}=\begin{bmatrix}-\boldsymbol{A}_1 & & & \\ -\boldsymbol{A}_2 & & \boldsymbol{I}_{(n-1)m} & \\ \vdots & & & \\ -\boldsymbol{A}_n & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\boldsymbol{C}_1\\ \boldsymbol{C}_2\\ \vdots\\ \boldsymbol{C}_n\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{I}_m,\boldsymbol{0},\cdots,\boldsymbol{0}] \tag{3.5.18}$$

因此问题进一步转化为对用状态空间模型描写的随机系统,求观测白噪声滤波器和平滑器 $\hat{\boldsymbol{v}}(t|t+N)(N\geqslant 0)$的问题.

上述例 3.5.1 和例 3.5.2 提出的问题可统一归结为对随机系统(3.1.1)和(3.1.2)在假设 1~2 下,求输入白噪声估值器 $\hat{\boldsymbol{w}}(t|t+N)$和观测白噪声估值器 $\hat{\boldsymbol{v}}(t|t+N)$,$N\geqslant 0$.

【定理 3.5.1】 系统(3.3.1)和(3.3.2)在假设 1~2 下,有最优输入白噪声估值器

$$\hat{\boldsymbol{w}}(t|t)=\boldsymbol{0} \tag{3.5.19}$$

$$\hat{\boldsymbol{w}}(t|t+N)=\boldsymbol{0}\quad (N<0) \tag{3.5.20}$$

$$\hat{\boldsymbol{w}}(t|t+1)=\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.5.21}$$

其中记新息 $\boldsymbol{\varepsilon}(t)$的方差阵 $\boldsymbol{Q}_{\varepsilon}(t)$为

$$\boldsymbol{Q}_{\varepsilon}(t)=\boldsymbol{HP}(t\mid t-1)\boldsymbol{H}^{\mathrm{T}}++\boldsymbol{R} \tag{3.5.22}$$

且有估值误差方差阵 $\boldsymbol{P}_w(t\mid t+N)=\mathrm{E}[(\boldsymbol{w}(t)-\hat{\boldsymbol{w}}(t\mid t+N))(\boldsymbol{w}(t)-\hat{\boldsymbol{w}}(t\mid t+N))^{\mathrm{T}}]$分别为

$$\boldsymbol{P}_w(t\mid t+N)=\boldsymbol{Q}\quad(N\leqslant 0) \tag{3.5.23}$$

$$\boldsymbol{P}_w(t\mid t+1)=\boldsymbol{Q}-\boldsymbol{Q\Gamma}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\boldsymbol{H\Gamma Q} \tag{3.5.24}$$

其中 $\boldsymbol{P}(t+1)$，$\boldsymbol{Q}_{\varepsilon}(t+1)$，$\boldsymbol{\varepsilon}(t+1)$由定理 3.3.1 的 Kalman 滤波器计算.

证明 注意

$$L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t))\subset L(\boldsymbol{v}(t),\cdots,\boldsymbol{v}(1),\boldsymbol{w}(t-1),\cdots,\boldsymbol{w}(t-1),\cdots,\boldsymbol{w}(0),\boldsymbol{x}(0)) \tag{3.5.25}$$

因而由假设 1～2 有

$$\boldsymbol{w}(t)\perp L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t)) \tag{3.5.26}$$

故有

$$\hat{\boldsymbol{w}}(t\mid t)=\mathrm{proj}(\boldsymbol{w}(t)\mid \boldsymbol{y}(1),\cdots \boldsymbol{y}(t))=0 \tag{3.5.27}$$

注意 $\boldsymbol{w}(t)\perp L(\boldsymbol{y}(1),\cdots,\boldsymbol{y}(t+N))$，$N<0$，故有

$$\hat{\boldsymbol{w}}(t\mid t+N)=\mathrm{proj}(\boldsymbol{w}(t)\mid \boldsymbol{y}(1),\cdots \boldsymbol{y}(t+N))=0\quad(N<0) \tag{3.5.28}$$

这说明输入白噪声滤波器和预报器为零. 由递推射影公式有

$$\hat{\boldsymbol{w}}(t\mid t+1)=\hat{\boldsymbol{w}}(t\mid t)+\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)]\boldsymbol{Q}_{\boldsymbol{\varepsilon}}^{-1}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.5.29}$$

注意到

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t+1\mid t)+\boldsymbol{v}(t+1) \tag{3.5.30}$$

进而由(3.4.8)有

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{H}[\boldsymbol{\Psi}_p(t)\tilde{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{\Gamma w}(t)-\boldsymbol{K}_p(t)\boldsymbol{v}(t)]+\boldsymbol{v}(t+1) \tag{3.5.31}$$

因 $\boldsymbol{w}(t)\perp\tilde{\boldsymbol{x}}(t\mid t-1)$，应用假设 1～2 引出

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)]=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{H} \tag{3.5.32}$$

将它代入(3.5.29)，见应用(3.5.19)得(3.5.21). 将(3.5.21)代入 $\boldsymbol{P}_w(t\mid t+N)$定义中，利用(3.5.32)得(3.5.24). 证毕. □

【定理 3.5.2】 系统(3.3.1)和(3.3.2)在假设 1～2 下，有最优输入白噪声递推平滑器

$$\hat{\boldsymbol{w}}(t\mid t+N)=\hat{\boldsymbol{w}}(t\mid t+N-1)+\boldsymbol{M}_w(t\mid t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.5.33}$$

带初值 $\hat{\boldsymbol{w}}(t\mid t)=\boldsymbol{0}$，$N=1,2,\cdots$，且估值器增益为

$$\boldsymbol{M}_w(t\mid t+1)=\boldsymbol{Q\Gamma}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\boldsymbol{\varepsilon}}^{-1}(\boldsymbol{t}+1) \tag{3.5.34}$$

$$\boldsymbol{M}_w(t\mid t+N)=\boldsymbol{Q\Gamma}^{\mathrm{T}}\left\{\prod_{i=1}^{N-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\right\}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\boldsymbol{\varepsilon}}^{-1}(\boldsymbol{t}+\boldsymbol{N}) \tag{3.5.35}$$

且估值误差方差阵 $\boldsymbol{Q}_w(t\mid t+N)$可递推计算为

$$\boldsymbol{P}_w(t\mid t+N)=\boldsymbol{P}_w(t\mid t+N-1)-\boldsymbol{M}_w(t\mid t+N)\boldsymbol{Q}_{\varepsilon}(t+N)\boldsymbol{M}_w^{\mathrm{T}}(t\mid t+N) \tag{3.5.36}$$

带初值 $\boldsymbol{P}_w(t\mid t)=\boldsymbol{Q}$. $\boldsymbol{\Psi}_p(t)$由(3.4.9)定义.

证明 由递推射影公式有(3.5.33)，其中

$$\boldsymbol{M}_w(t\mid t+N)=\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]\boldsymbol{Q}_{\varepsilon}^{-1}(t+N) \tag{3.5.37}$$

应用(3.4.10)和(3.4.11)，且应用关系(3.4.6)有

$$\boldsymbol{\varepsilon}(t+N)=\boldsymbol{H}\{\boldsymbol{\Psi}(t+N,t)\tilde{\boldsymbol{x}}(t\mid t-1)+\sum_{i=t+1}^{t+N}\boldsymbol{\Psi}(t+N,i)\times [\boldsymbol{\Gamma}\boldsymbol{w}(i-1)-\boldsymbol{K}_p(i-1)\boldsymbol{v}(i-1)]\}+\boldsymbol{v}(t+N) \tag{3.5.38}$$

由假设 1～2 和上式有

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]=\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Psi}^{\mathrm{T}}(t+N,t+1)\boldsymbol{H}^{\mathrm{T}}=\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\left\{\prod_{\mathrm{i}=1}^{\mathrm{N}-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\right\}\boldsymbol{H}^{\mathrm{T}} \tag{3.5.39}$$

将(3.5.39)代入(3.5.37)得(3.5.35). 由(3.5.33)引出关系

$$\tilde{\boldsymbol{w}}(t\mid t+N)=\tilde{\boldsymbol{w}}(t\mid t+N-1)-\boldsymbol{M}_w(t\mid t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.5.40}$$

其中 $\tilde{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{w}(t)-\hat{\boldsymbol{w}}(t\mid t+N)$, $\tilde{\boldsymbol{w}}(t\mid t+N-1)=\boldsymbol{w}(t)-\hat{\boldsymbol{w}}(t\mid t+N-1)$. 改写上式为

$$\tilde{\boldsymbol{w}}(t\mid t+N-1)=\tilde{\boldsymbol{w}}(t\mid t+N)+\boldsymbol{M}_w(t\mid t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.5.41}$$

由射影正交性有 $\tilde{\boldsymbol{w}}(t\mid t+N)\perp\boldsymbol{\varepsilon}(t+N)$，故有

$$\boldsymbol{P}_w(t\mid t+N-1)=\boldsymbol{P}_w(t\mid t+N)+\boldsymbol{M}_w(t\mid t+N)\boldsymbol{Q}_\varepsilon(t+N)\boldsymbol{M}_w^{\mathrm{T}}(t\mid t+N) \tag{3.5.42}$$

这引出(3.5.36). 证毕. □

【推论 3.5.1】 在定理 3.5.2 条件下，有最优非递推输入白噪声平滑器

$$\hat{\boldsymbol{w}}(t\mid t+N)=\sum_{i=1}^{N}\boldsymbol{M}_w(t\mid t+i)\boldsymbol{\varepsilon}(t+i),\quad N>0 \tag{3.5.43}$$

且有误差方差阵

$$\boldsymbol{P}_w(t\mid t+N)=\boldsymbol{Q}-\sum_{i=1}^{N}\boldsymbol{M}_w(t\mid t+i)\boldsymbol{Q}_\varepsilon(t+i)\boldsymbol{M}_w^{\mathrm{T}}(t\mid t+i) \tag{3.5.44}$$

【定理 3.5.3】 系统(3.3.1)和(3.3.2)在假设 1～2 下，最优观测白噪声估值器 $\hat{\boldsymbol{v}}(t\mid t+N)$为

$$\hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{0}\quad(N<0) \tag{3.5.45}$$

$$\hat{\boldsymbol{v}}(t\mid t+N)=\hat{\boldsymbol{v}}(t\mid t+N-1)+\boldsymbol{M}_v(t\mid t+N)\boldsymbol{\varepsilon}(t+N),\quad N=1,2,\cdots \tag{3.5.46}$$

初值为

$$\hat{\boldsymbol{v}}(t\mid t)=\boldsymbol{R}\boldsymbol{Q}_\varepsilon^{-1}(t)\boldsymbol{\varepsilon}(t) \tag{3.5.47}$$

平滑增益 $\boldsymbol{M}_v(t\mid t+N)$为

$$\boldsymbol{M}_v(t\mid t+1)=-\boldsymbol{R}\boldsymbol{K}^{\mathrm{T}}(t)\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.5.48}$$

$$\boldsymbol{M}_v(t\mid t+N)=-\boldsymbol{R}\boldsymbol{K}^{\mathrm{T}}(t)\boldsymbol{\Phi}^{\mathrm{T}}\left\{\prod_{\mathrm{i}=1}^{\mathrm{N}-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\right\}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}(t+N) \tag{3.5.49}$$

且有误差方差阵

$$\boldsymbol{P}_v(t\mid t+N)=\boldsymbol{R}\quad(N<0) \tag{3.5.50}$$

$$\boldsymbol{P}_v(t\mid t+N)=\boldsymbol{P}_v(t\mid t+N-1)-\boldsymbol{M}_v(t\mid t+N)\boldsymbol{Q}_\varepsilon(t+N)\boldsymbol{M}_v^{\mathrm{T}}(t\mid t+N),\quad N\geqslant 0 \tag{3.5.51}$$

其中定义 $\boldsymbol{M}_v(t\mid t)$为

$$\boldsymbol{M}_v(t\mid t)=\boldsymbol{R}\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.5.52}$$

证明 类似于定理 3.5.2，留给读者练习，从略. □

【推论 3.5.2】 在定理 3.5.1 条件下，有非递推最优观测白噪声平滑器

$$\hat{\boldsymbol{v}}(t \mid t+N)=\sum_{i=0}^{N}\boldsymbol{M}_v(t \mid t+i)\boldsymbol{\varepsilon}(t+i) \tag{3.5.53}$$

误差方差阵为

$$\boldsymbol{P}_v(t \mid t+N)=\boldsymbol{R}-\sum_{i=0}^{N}\boldsymbol{M}_v(t \mid t+i)\boldsymbol{Q}_\varepsilon(t+i)\boldsymbol{M}_v^{\mathrm{T}}(t+i) \tag{3.5.54}$$

【定理 3.5.4】 系统(3.3.1)和(3.3.2)在假设 1～2 下，有固定区间白噪声估值器 $\hat{\boldsymbol{w}}(t|N)$和 $\hat{\boldsymbol{v}}(t|N)$分别为

$$\hat{\boldsymbol{w}}(t|N)=\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{r}(t+1|N) \tag{3.5.55}$$

$$\hat{\boldsymbol{v}}(t|N)=\hat{\boldsymbol{v}}(t|t)-\boldsymbol{R}\boldsymbol{K}^{\mathrm{T}}(t)\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{r}(t+1|N) \tag{3.5.56}$$

其中 $t=N,N-1,\cdots,1$，且有反向递推公式

$$\boldsymbol{r}(t+1|N)=\boldsymbol{\Psi}_p^{T}(t+1)\boldsymbol{r}(t+2|N)+\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.5.57}$$

带初值 $\boldsymbol{r}(N+1|N)=\mathbf{0}$，相应的估值误差方差阵各为

$$\boldsymbol{P}_w(t|N)=\boldsymbol{Q}-\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{S}(t+1|N)\boldsymbol{\Gamma}\boldsymbol{Q} \tag{3.5.58}$$

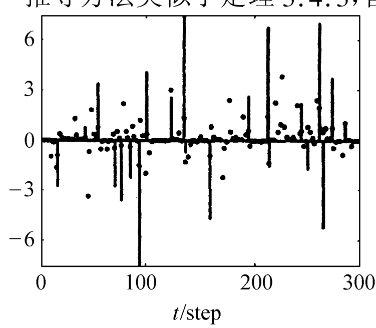
$$\boldsymbol{P}_v(t|N)=\boldsymbol{P}_v(t|t)-\boldsymbol{R}\boldsymbol{K}^{\mathrm{T}}(t)\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{S}(t+1|N)\boldsymbol{\Phi}\boldsymbol{K}(t)\boldsymbol{R} \tag{3.5.59}$$

其中 $\boldsymbol{S}(t+1|N)$可递推计算为

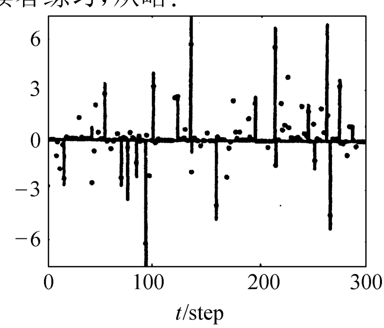
$$\boldsymbol{S}(t+1|N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)\boldsymbol{S}(t+2|N)\boldsymbol{\Psi}_p(t+1)+\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_\varepsilon(t+1)\boldsymbol{H} \tag{3.5.60}$$

带初值 $\boldsymbol{S}(N+1|N)=\mathbf{0}$，$t=N,N-1,\cdots,1$.

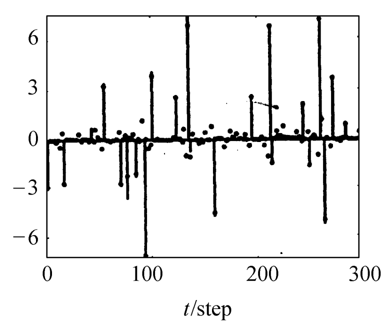
证明 推导方法类似于定理 3.4.3，留给读者练习，从略. □



(a) Bernoulli－Caussian **输入白噪声** $w(t)$ **和** Wiener **平滑器** $\hat{w}(t|t+1)$



(b) Bernoulli－Caussian **输入白噪声** $w(t)$ **和** Wiener **平滑器** $\hat{w}(t|t+2)$



(c) Bernoulli－Caussian **输入白噪声** $w(t)$ **和** Wiener **平滑器** $\hat{w}(t|t+3)$

图 3.5.5 Bernoulli－Gaussian 白噪声 $w(t)$平滑器 $\hat{w}(t|t+N)$，$N=1,2,3$

【例 3.5.3】 考虑随机系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}1 & 0\\ 0.3 & -0.5\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}-1\\ 2\end{bmatrix}w(t) \tag{3.5.61}$$

$$y(t)=[1\quad 1]\boldsymbol{x}(t)+v(t) \tag{3.5.62}$$

其中 $w(t)=b(t)g(t)$ 是 Bemoulli – Gaussian 白噪声. $b(t)$ 是取值 1 或 0 的 Bernoulli 白噪声取值概率为 $P(b(t)=1)=\lambda$, $P(b(t)=0)=1-\lambda$, $g(t)$ 是零、方差为 σ_g^2 的独立于 $b(t)$ 的高斯白噪声. 应用定理 3.5.1 和定理 3.5.2 可得 Bernoulli – Gaussian 白噪声反卷积平滑器 $\hat{w}(t|t+N)$, $N=1,2,3$. 仿真中取 $\lambda=0.3$, $\sigma_v^2=0.1$, $\sigma_g^2=7$, 仿真结果如图 3.5.5 所示, 其中实线端点纵标代表真实值, 圆点纵标代表平滑估值. 可看到 $\hat{w}(t|t+3)$ 精度较高.

3.6 稳态 Kalman 滤波

由节 3.3 我们看到, Kalman 滤波器是一种时变递推滤波器. 这里时变是指 Kalman 滤波器增益阵 $\boldsymbol{K}(t)$ 是时变的. 实现最优 Kalman 滤波器要求在每时刻计算增益 $\boldsymbol{K}(t)$ 阵, 带来较大计算负担. 因为计算 $\boldsymbol{K}(t)$ 要求计算 $m\times m$ 维逆矩阵 $[\boldsymbol{HP}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}]^{-1}$. 当观测 $\boldsymbol{y}(t)$ 的维数 m 较大时, 这将要求高维矩阵的逆矩阵, 需要较大计算量. 从工程应用观点, 这不便于实时应用. 简化 Kalman 滤波器增益计算的途径是考察 Kalman 滤波器增益阵是否趋于常阵? 即当 $t\to\infty$ 时增益 $\boldsymbol{K}(t)$ 是否有极限? 若 $\boldsymbol{K}(t)\to\boldsymbol{K}$, $t\to\infty$, $\boldsymbol{K}$ 为一个常阵, 则当 t 充分大, 则有 $\boldsymbol{K}(t)\approx\boldsymbol{K}$, 因而在工程应用中可用常阵 $\boldsymbol{K}$ 近似代替最优时变增益 $\boldsymbol{K}(t)$, 就可得到稳态 Kalman 滤波器, 它是带常增益阵 $\boldsymbol{K}$ 的 Kalman 滤波器, 是一种次优 Kalman 滤波器, 它不需要在每时刻计算增益阵, 因而大大减小在线计算负担, 便于工程应用. 当 t 增大时, 它是渐近最优的, 因为 $\boldsymbol{K}(t)$ 与 $\boldsymbol{K}$ 的误差可任意小.

由推论 3.3.1 最优时变 Kalman 滤波器为

$$\hat{\boldsymbol{x}}(t+1|t+1)=\boldsymbol{\Psi}_f(t)\hat{\boldsymbol{x}}(t|t)+\boldsymbol{K}(t+1)\boldsymbol{y}(t+1) \tag{3.6.1}$$

$$\boldsymbol{\Psi}_f(t)=[\boldsymbol{I}_n-\boldsymbol{K}(t+1)\boldsymbol{H}]\boldsymbol{\Phi} \tag{3.6.2}$$

它是时变滤波器, 因为增益 $\boldsymbol{K}(t)$ 和转移阵 $\boldsymbol{\Psi}_f(t)$ 均为时变矩阵. 假如增益 $\boldsymbol{K}(t)$ 存在极限

$$\lim_{t\to\infty}\boldsymbol{K}(t)=\boldsymbol{K} \tag{3.6.3}$$

则稳态 Kalman 滤波器为

$$\hat{\boldsymbol{x}}(t+1|t+1)=\boldsymbol{\Psi}_f\hat{\boldsymbol{x}}(t|t)+\boldsymbol{K}\boldsymbol{y}(t+1) \tag{3.6.4}$$

$$\boldsymbol{\Psi}_f=[\boldsymbol{I}_n-\boldsymbol{KH}]\boldsymbol{\Phi} \tag{3.6.5}$$

它是非时变滤波器, 因为增益 $\boldsymbol{K}$ 和转移阵 $\boldsymbol{\Psi}_f$ 是非时变的.

这里存在两个理论问题: 一个是什么条件下, 滤波增益 $\boldsymbol{K}(t)$ 存在极限? 这是稳态 Kalman 滤波器存在的问题. 注意稳态 Kalman 滤波器 (3.6.4) 是递推滤波器, 它要求设置初值 $\hat{\boldsymbol{x}}(0|0)$. 因此另一个问题是是否稳态 Kalman 滤波器的计算渐近地与初值 $\hat{\boldsymbol{x}}(0|0)$ 的选取无关? 这是滤波器的渐近稳定性问题.

为了解决这两个问题我们看下面的启发性例子.

【例 3.6.1】 一个启发性例子. 考虑一维系统

$$x(t+1)=x(t)+w(t) \tag{3.6.6}$$

$$y(t) = x(t) + v(t) \tag{3.6.7}$$

其中 $w(t)$和 $v(t)$是零均值、方差各为 $\sigma_w^2 = 1$ 和 $\sigma_v^2 = 1$ 的独立白噪声. 初值 $\mathrm{E}x(0) = 0$, $\mathrm{E}[x^2(0)] = 1$.

由定理 3.3.1 和推论 3.3.1 有 Kalman 滤波器

$$\hat{x}(t|t) = (1 - K(t))\hat{x}(t-1|t-1) + K(t)y(t) \tag{3.6.8}$$

$$\hat{x}(0|0) = 0, \quad P(0|0) = 1 \tag{3.6.9}$$

$$P(t|t-1) = P(t-1|t-1) + 1 \tag{3.6.10}$$

$$K(t) = \frac{P(t|t-1)}{1 + P(t|t-1)} \tag{3.6.11}$$

$$P(t|t) = [1 - K(t)]P(t|t-1) \tag{3.6.12}$$

这引出 Riccati 方程

$$P(t|t-1) = \frac{1 - 2P(t-1|t-2)}{1 + P(t-1|t-2)} \tag{3.6.13}$$

由(3.6.11)~(3.6.13), $K(t)$存在极限问题归结为 $P(t|t-1)$是否存在极限? 而且若 $P(t|t-1)$存在极限,则 $P(t|t)$也存在极限. 现考察 $P(t|t-1)$的变化趋势,见表 3.6.1.

表 3.6.1

t	$P(t\|t-1)$	$K(t)$	$P(t\|t)$
1	2	0.666	0.668
2	1.666	0.625	0.625
3	1.625	0.619	0.619
4	1.619	0.618	0.618
5	1.618	0.618	0.618

直观上由表可看到 $P(t|t-1)$, $K(t)$, $P(t|t)$三者均为单调下降,且均有一定变化趋势,且收敛很快. 可见预报误差方差 $P(t|t-1)$是非负递降的,因而存在极限

$$\lim_{t\to\infty} P(t|t-1) = \Sigma \tag{3.6.14}$$

在(3.6.13)两边取极限有稳态 Riccati 方程

$$\Sigma = \frac{1 - 2\Sigma}{1 + \Sigma} \tag{3.6.15}$$

它可化为一元二次方程

$$\Sigma^2 - \Sigma - 1 = 0 \tag{3.6.16}$$

它有非负解

$$\Sigma = (1 + \sqrt{5})/2 = 1.618\,033\,9 \tag{3.6.17}$$

这就是说

$$\lim_{t\to\infty} P(t|t-1) = 1.618\,033\,9 \tag{3.6.18}$$

因而由(3.6.11)Kalman 滤波器增益 $K(t)$有极限

$$\lim_{t\to\infty} K(t) = K \tag{3.6.19}$$

$$K = \frac{\Sigma}{1+\Sigma} = \frac{1+\sqrt{5}}{3+\sqrt{5}} = 0.618\,033\,9 \tag{3.6.20}$$

且由(3.6.12)滤波误差方差 $P(t\mid t)$ 也有极限

$$\lim_{t\to\infty} P(t\mid t) = P \tag{3.6.21}$$

$$P = (1-K)\Sigma = 0.618\,034 \tag{3.6.22}$$

我们称 K 为稳态 Kalman 滤波器增益，在时变 Kalman 滤波器(3.6.8)中用 K 近似代替$K(t)$ 有稳态 Kalman 滤波器

$$\hat{x}(t\mid t) = (1-K)\hat{x}(t-1\mid t-1) + Ky(t) \tag{3.6.23}$$

对于从有限初始时刻 $t_0 = 0$ 开始的最优 Kalman 滤波器(3.6.8)而言，稳态 Kalman 滤波器(3.6.23)是次优的不是最优的. 因为其增益 K 不是最优增益 $K(t)$. 滤波器的转移阵 $\Psi_f = 1-K$ 也不是最优转移阵 $\Psi(t) = 1-K(t)$. 但由于当 $t\to\infty$ 时，$K(t)\to K$，因而也有 $\Psi(t)\to\Psi$，从而当 $t\to\infty$ 时(3.6.23)是渐近最优的.

从工程应用观点，采用稳态 Kalman 滤波器的优点是避免了在线计算增益 $K(t)$，而稳态增益 K 可一次性离线计算，因而简化了 Kalman 滤波器的计算，减小了在线计算负担，便于实时应用.

下面考察稳态 Kalman 滤波器(3.6.23)的稳定性问题，即稳态 Kalman 滤波器(3.6.23)的计算与初值 $\hat{x}(0\mid 0)$ 的选取有何关系？初值 $\hat{x}(0\mid 0)$ 是否可任意选取？是否初值的影响可忽略不计？这个问题很重要，因为在实际应用中，滤波初值通常是未知的. 为此，任取(3.6.23)的两个初值 $\hat{x}^{(1)}(0\mid 0)$ 和 $\hat{x}^{(2)}(0\mid 0)$，记相应的稳态 Kalman 滤波器为 $\hat{x}^{(1)}(t\mid t)$ 和 $\hat{x}^{(2)}(t\mid t)$，则有

$$\hat{x}^{(1)}(t\mid t) = (1-K)\hat{x}^{(1)}(t-1\mid t-1) + Ky(t) \tag{3.6.24}$$

$$\hat{x}^{(2)}(t\mid t) = (1-K)\hat{x}^{(2)}(t-1\mid t-1) + Ky(t) \tag{3.6.25}$$

记两个稳态 Kalman 滤波估值误差为 $\delta(t)$，即

$$\delta(t) = \hat{x}^{(1)}(t\mid t) - \hat{x}^{(2)}(t\mid t) \tag{3.6.26}$$

由(3.6.24)减(3.6.25)引出差分方程

$$\delta(t) = (1-K)\delta(t-1) \tag{3.6.27}$$

由(3.6.27)迭代可得通解

$$\delta(t) = (1-K)^t\delta(0) \tag{3.6.28}$$

其中 $\delta(0) = \hat{x}^{(1)}(0\mid 0) - \hat{x}^{(2)}(0\mid 0)$. 因为 $0<1-K<1$，故

$$\lim_{t\to\infty}\delta(t) = 0 \tag{3.6.29}$$

这表明：无论稳态 Kalman 滤波器(3.6.23)的初值 $\hat{x}(t\mid t)$ 如何选取，当 $t\to\infty$ 时，它对稳态 Kalman 滤波器 $\hat{x}(t\mid t)$ 的影响将消失，即 $\hat{x}(t\mid t)$ 渐近地与初值选取无关，故称稳态 Kalman 滤波器(3.6.23)是渐近稳定的.

在一般情形下，稳态 Kalman 滤波器(3.6.4)是渐近稳定的充要条件为：滤波器的转移阵 $\boldsymbol{\Psi}_f$ 是一个稳定矩阵，即 $\boldsymbol{\Psi}_f$ 的所有特征值位于单位圆内. 事实上，任取(3.6.4)的两个初值 $\hat{\boldsymbol{x}}^{(1)}(0\mid 0)$ 和 $\hat{\boldsymbol{x}}^{(2)}(0\mid 0)$，记相应稳定 Kalman 滤波器 $\hat{\boldsymbol{x}}^{(1)}(t\mid t)$ 和 $\hat{\boldsymbol{x}}^{(2)}(t\mid t)$ 的值之差为 $\boldsymbol{\delta}(t)$，$\boldsymbol{\delta}(t) = \hat{\boldsymbol{x}}^{(1)}(t\mid t) - \hat{\boldsymbol{x}}^{(2)}(t\mid t)$，则类似于(3.6.24) ~ (3.6.27)的推导，可得向量差分方程

$$\boldsymbol{\delta}(t)=\boldsymbol{\Psi}_f\boldsymbol{\delta}(t-1) \tag{3.6.30}$$

它有通解

$$\boldsymbol{\delta}(t)=\boldsymbol{\Psi}_f^t\boldsymbol{\delta}(0) \tag{3.6.31}$$

由此式看到 $\boldsymbol{\delta}(t)\to\boldsymbol{0}$ 的充要条件是 $\boldsymbol{\Psi}_f^t\to\boldsymbol{0}(t\to\infty)$，而这个条件等价于 $\boldsymbol{\Psi}_f$ 是一个稳定矩阵.

经典 Kalman 滤波理论已解决了上述两个问题，给出了如下充分条件：若系统(3.3.1)和(3.3.2)是完全可观、完全可控的或 $\boldsymbol{\Phi}$ 为稳定矩阵，则稳态 Kalman 滤波器存在，且是渐近稳定的.

系统(3.3.1)和(3.3.2)是完全可观的等价于条件

$$\operatorname{rank}\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{H\Phi}\\ \vdots\\ \boldsymbol{H\Phi}^{n-1}\end{bmatrix}=n \tag{3.6.32}$$

系统(3.3.1)和(3.3.2)是完全可控的等价于条件

$$\operatorname{rank}[\boldsymbol{\Gamma},\boldsymbol{\Phi\Gamma},\cdots,\boldsymbol{\Phi}^{n-1}\boldsymbol{\Gamma}]=n \tag{3.6.33}$$

在定理 3.3.1 定理 3.3.2 和推论 3.3.1 中，令 $t\to\infty$，我们有如下稳态 Kalman 滤波定理.

【定理 3.6.1】 定常系统(3.3.1)和(3.3.2)在假设 1 ~ 2 下，若系统是完全可观、完全可控的，或 $\boldsymbol{\Phi}$ 为稳定矩阵，则对任意非负定矩阵 $\boldsymbol{P}(1|0)\geqslant0$，矩阵 Riccati 方程(3.3.59)

$$\boldsymbol{P}(t+1|t)=\boldsymbol{\Phi}[\boldsymbol{P}(t|t-1)-\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{HP}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R})^{-1}\boldsymbol{HP}(t|t-1)]\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{\Gamma Q\Gamma}^{\mathrm{T}} \tag{3.6.34}$$

总存在极限

$$\lim_{t\to\infty}\boldsymbol{P}(t|t-1)=\boldsymbol{\Sigma} \tag{3.6.35}$$

其中极限 $\boldsymbol{\Sigma}$ 是如下稳态 Riccati 方程的惟一非负定解：

$$\boldsymbol{\Sigma}=\boldsymbol{\Phi}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}}+\boldsymbol{R})^{-1}\boldsymbol{H\Sigma}]\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{\Gamma Q\Gamma}^{\mathrm{T}} \tag{3.6.36}$$

其中极限 $\boldsymbol{\Sigma}$ 与初值 $\boldsymbol{P}(1|0)$ 选取无关，且也存在极限

$$\lim_{t\to\infty}\boldsymbol{K}(t)=\boldsymbol{K},\quad \lim_{t\to\infty}\boldsymbol{P}(t|t)=\boldsymbol{P} \tag{3.6.37}$$

它们满足关系

$$\boldsymbol{\Sigma}=\boldsymbol{\Phi P\Phi}^{\mathrm{T}}+\boldsymbol{\Gamma Q\Gamma}^{\mathrm{T}} \tag{3.6.38}$$

$$\boldsymbol{K}=\boldsymbol{\Sigma H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\boldsymbol{R}]^{-1} \tag{3.6.39}$$

$$\boldsymbol{P}=[\boldsymbol{I}_n-\boldsymbol{KH}]\boldsymbol{\Sigma} \tag{3.6.40}$$

稳态 Kalman 滤波器为

$$\hat{\boldsymbol{x}}(t|t)=\boldsymbol{\Psi}_f\hat{\boldsymbol{x}}(t-1|t-1)+\boldsymbol{Ky}(t) \tag{3.6.41}$$

$$\boldsymbol{\Psi}_f=[\boldsymbol{I}_n-\boldsymbol{KH}]\boldsymbol{\Phi} \tag{3.6.42}$$

稳态 Kalman 预报器为

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\boldsymbol{\varepsilon}(t) \tag{3.6.43}$$

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)-\boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1) \tag{3.6.44}$$

$$\boldsymbol{K}_p=\boldsymbol{\Phi K} \tag{3.6.45}$$

其中 $\boldsymbol{K}_p$ 叫稳态 Kalman 预报器增益.稳态 Kalman 预报器也可表为

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\boldsymbol{y}(t) \tag{3.6.46}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H}=\boldsymbol{\Phi}[\boldsymbol{I}_n-\boldsymbol{K}\boldsymbol{H}] \tag{3.6.47}$$

且转移阵 $\boldsymbol{\Psi}_f$ 和 $\boldsymbol{\Psi}_p$ 均为稳定矩阵,它们有相同的特征值.因而稳态 Kalman 滤波器和预报器是渐近稳定的.

证明 这个定理的严格理论证明超出本书范围,从略,见参考文献[5]. □

【注 3.6.1】 由(3.6.34)给出的迭代法给出了稳态 Riccati 方程(3.6.36)的迭代解.当 t 充分大,有 $\boldsymbol{P}(t|t+1)\approx\boldsymbol{\Sigma}$.

【注 3.6.2】 稳态 Kalman 滤波是对定常(非时变)系统而言的.

【注 3.6.3】 上述稳定 Kalman 滤波器和预报器在初始观测时刻 $t_0=1$ 情形下,令 $t\to+\infty$ 产生的.相对而言,它们也可用固定时刻 t 组合 $t_0\to-\infty$ 的方式产生.此时我们也有

$$\lim_{t\to-\infty}\boldsymbol{P}(t+1|t)=\boldsymbol{\Sigma} \tag{3.6.48}$$

而 $\boldsymbol{\Sigma}$ 满足(3.6.36).此时稳态 Kalman 滤波器和预报器也是当 $t_0\to-\infty$ 时的最优 Kalman 滤波器和预报器.

类似地由节 3.4 和节 3.5 有关定理可得相应的稳态 Kalman 平滑器和稳态白噪声估值器.

【定理 3.6.2】 完全可观、完全可控或 $\boldsymbol{\Phi}$ 为稳定的定常系统(3.3.1)和(3.3.2)在假设 1~2 下,稳态固定滞后 Kalman 平滑器为

$$\hat{\boldsymbol{x}}(t\mid t+N)=\hat{\boldsymbol{x}}(t\mid t-1)+\sum_{i=0}^{N}\boldsymbol{K}_i\boldsymbol{\varepsilon}(t+i) \tag{3.6.49}$$

且稳态平滑增益 $\boldsymbol{K}_i$ 为

$$\boldsymbol{K}_i=\boldsymbol{\Sigma}(\boldsymbol{\Psi}_p^{\mathrm{T}})^i\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}=\boldsymbol{\Sigma}[(\boldsymbol{I}_n-\boldsymbol{K}\boldsymbol{H})^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}}]^i\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1} \tag{3.6.50}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}[\boldsymbol{I}_n-\boldsymbol{K}\boldsymbol{H}] \tag{3.6.51}$$

$$\boldsymbol{Q}_\varepsilon=\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R} \tag{3.6.52}$$

$$\boldsymbol{K}_0=\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}=\boldsymbol{K} \tag{3.6.53}$$

其中 $\hat{\boldsymbol{x}}(t|t-1)$,$\boldsymbol{\Sigma}$,$\boldsymbol{K}$,$\boldsymbol{\Psi}_p$,$\boldsymbol{Q}_\varepsilon$ 均由稳态 Kalman 预报器计算.稳态平滑误差方差阵为

$$\boldsymbol{P}_N=\boldsymbol{\Sigma}-\sum_{i=0}^{N}\boldsymbol{K}_i\boldsymbol{Q}_\varepsilon\boldsymbol{K}_i^{\mathrm{T}} \tag{3.6.54}$$

【定理 3.6.3】 完全可观、完全可控或 $\boldsymbol{\Phi}$ 为稳定的定常系统(3.3.1)和(3.3.2)在假设 1~2 下,有超前 N 步稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+N|t)=\boldsymbol{\Phi}^{N-1}\hat{\boldsymbol{x}}(t+1|t),\quad N>1 \tag{3.6.55}$$

且预报误差方差阵为

$$\boldsymbol{P}_N=\boldsymbol{\Phi}^{N-1}\boldsymbol{\Sigma}\boldsymbol{\Phi}^{(N-1)\mathrm{T}}+\sum_{j=2}^{N}\boldsymbol{\Phi}^{N-j}\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Phi}^{(N-j)\mathrm{T}} \tag{3.6.56}$$

其中 $\hat{\boldsymbol{x}}(t+1|t)$,$\boldsymbol{\Sigma}$ 由稳态 Kalman 预报器计算.

【定理 3.6.4】 完全可观、完全可控或 $\boldsymbol{\Phi}$ 是稳定的系统(3.3.1)和(3.3.2)在假设 1~2 下,稳态白噪声滤波器和平滑器为

$$\hat{\boldsymbol{w}}(t\mid t+N)=\sum_{i=1}^{N}\boldsymbol{M}_w(i)\boldsymbol{\varepsilon}(t+i),\quad N>0 \tag{3.6.57}$$

$$\hat{\boldsymbol{v}}(t \mid t+N)=\sum_{i=0}^{N} \boldsymbol{M}_v(i)\boldsymbol{\varepsilon}(t+i), \quad N \geqslant 0 \tag{3.6.58}$$

其中增益 $\boldsymbol{M}_w(t)$和 $\boldsymbol{M}_v(t)$分别为

$$\boldsymbol{M}_w(1)=\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1} \tag{3.6.59}$$

$$\boldsymbol{M}_w(i)=\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}(\boldsymbol{\Psi}_p^{\mathrm{T}})^{(i-1)}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1} \tag{3.6.60}$$

$$\boldsymbol{M}_v(0)=\boldsymbol{R}\boldsymbol{Q}_{\varepsilon}^{-1} \tag{3.6.61}$$

$$\boldsymbol{M}_v(1)=-\boldsymbol{R}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1} \tag{3.6.62}$$

$$\boldsymbol{M}_v(i)=-\boldsymbol{R}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\Psi}_p^{\mathrm{T}})^{(i-1)}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1} \tag{3.6.63}$$

且相应的稳态估值误差方差阵为

$$\boldsymbol{P}_w(N)=\boldsymbol{Q}-\sum_{i=1}^{N}\boldsymbol{M}_w(i)\boldsymbol{Q}_{\varepsilon}\boldsymbol{M}_w^{\mathrm{T}}(i) \tag{3.6.64}$$

$$\boldsymbol{P}_v(N)=\boldsymbol{R}-\sum_{i=0}^{N}\boldsymbol{M}_v(i)\boldsymbol{Q}_{\varepsilon}\boldsymbol{M}_v^{\mathrm{T}}(i) \tag{3.6.65}$$

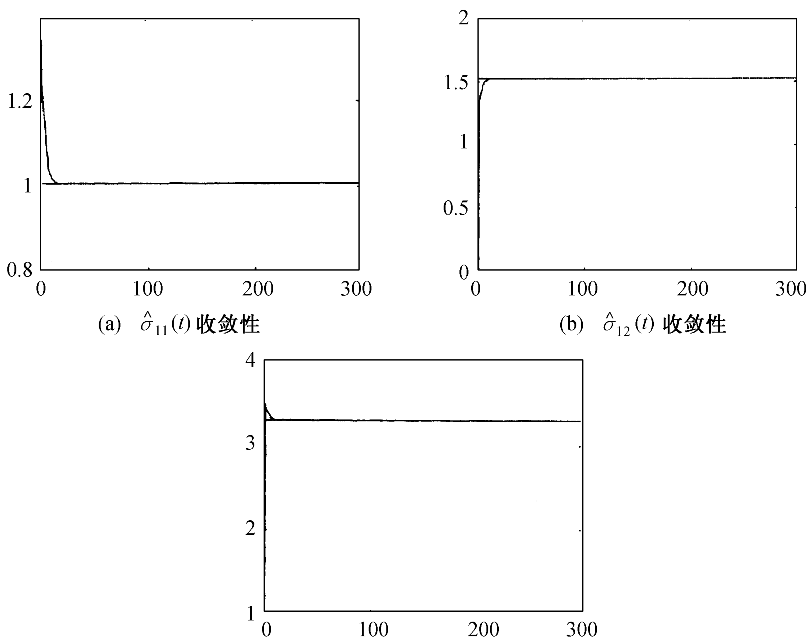
在上述公式中，$\boldsymbol{\varepsilon}(t)$，$\boldsymbol{K}$，$\boldsymbol{\Psi}_p$，$\boldsymbol{Q}_{\varepsilon}$ 均由稳态 Kalman 滤波器和预报器计算.

【例 3.6.2】 考虑完全可观、完全可控系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix} 0.9 & 0 \\ -0.5 & 0.2 \end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix} 1 \\ 2 \end{bmatrix}w(t) \tag{3.6.66}$$

$$y(t)=[1 \quad 1]\boldsymbol{x}(t)+v(t) \tag{3.6.67}$$

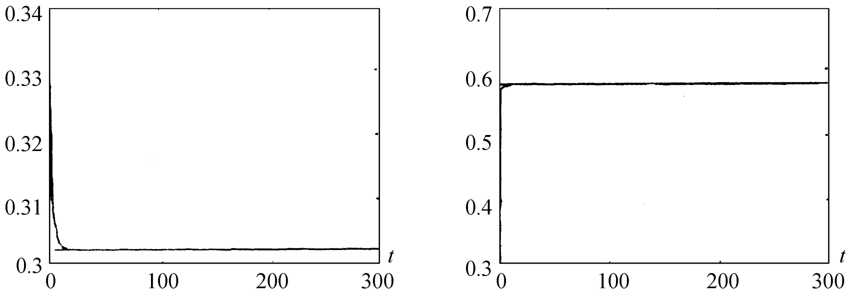
其中 $\boldsymbol{x}(t)=[(x_1(t), x_2(t)]^{\mathrm{T}}$，$w(t)$和 $v(t)$是零均值、方差各为 $Q=0.81$ 和 $R=1$ 的高斯白噪声.应用定理 3.6.1 可求得稳态 Kalman 滤波器 $\hat{x}_i(t \mid t)$.仿真结果如图 3.6.1 至图 3.6.3所示.图3.6.1表示用迭代算法(3.6.34)的收敛性，其中置



(a) $\hat{\sigma}_{11}(t)$ 收敛性 (b) $\hat{\sigma}_{12}(t)$ 收敛性 (c) $\hat{\sigma}_{22}(t)$ 收敛性

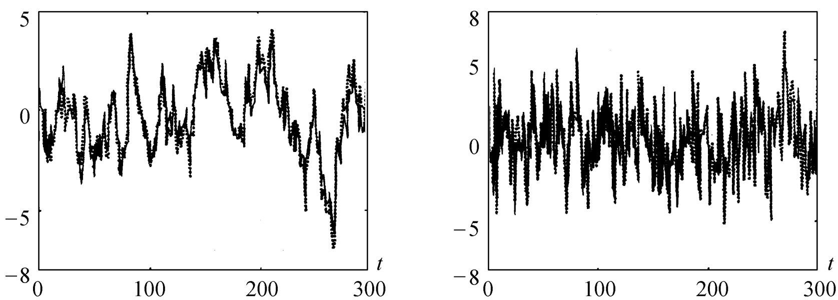
图 3.6.1 稳态 Riccati 方程迭代解的收敛性

$$\boldsymbol{P}(t \mid t-t)=\begin{bmatrix}\hat{\sigma}_{11}(t), & \hat{\sigma}_{12}(t)\\ \hat{\sigma}_{12}(t), & \hat{\sigma}_{22}(t)\end{bmatrix},\quad \boldsymbol{\Sigma}=\begin{bmatrix}\sigma_{11}, & \sigma_{12}\\ \sigma_{12}, & \sigma_{22}\end{bmatrix} \tag{3.6.68}$$



(a) 滤波增益 $k_1(t)$ 收敛性　　(b) 滤波增益 $k_2(t)$ 收敛性

图 3.6.2　最优时变 Kalman 滤波器增 $\boldsymbol{K}(t)=[k_1(t),k_2(t)]^{\mathrm{T}}$ 的收敛性



(a) $x_1(t)$ 和 $\hat{x}_1(t|t)$　　(b) $x_2(t)$ 和 $\hat{x}_2(t|t)$

图 3.6.3　状态 $x_i(t)$ 和稳态 Kalman 滤波器 $\hat{x}_j(t \mid t)$

在 $t=300$ 处可得稳态 Riccati 方程(3.6.36)的近似解 $\boldsymbol{\Sigma}$ 为

$$\boldsymbol{\Sigma}=\begin{bmatrix}1.004\,633\,013\,300\,49 & 1.523\,048\,938\,740\,52\\ 1.523\,049\,938\,740\,52 & 3.308\,286\,551\,886\,38\end{bmatrix} \tag{3.6.69}$$

应用(3.6.39)可求得稳态 Kalman 滤波器增益 K 为

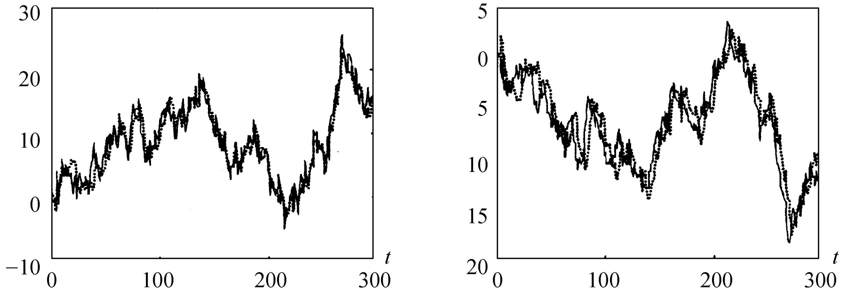
$$\boldsymbol{K}=\begin{bmatrix}k_1\\ k_2\end{bmatrix}=\begin{bmatrix}0.302\,389\,840\,597\,61\\ 0.577\,978\,874\,163\,58\end{bmatrix} \tag{3.6.70}$$

【例 3.6.3】　考虑随机系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}0.3 & -0.9\\ 0 & 1\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\ 1\end{bmatrix}w(t) \tag{3.6.71}$$

$$y(t)=[1\quad 0]\,\boldsymbol{x}(t)+v(t) \tag{3.6.72}$$

其中 $\boldsymbol{x}(t)=[(x_1(t),x_2(t)]^{\mathrm{T}}$，$w(t)$ 和 $v(t)$ 是零均值、方差各为 $Q=0.1$ 和 $R=6$ 的独立高斯白噪声．应用定理 3.6.3 可求得稳态 Kalman 预报器 $\hat{\boldsymbol{x}}(t+2 \mid t)$．仿真结果如图 3.6.4 所示，其中实线代表真实值，虚线代表 Kalman 预报估值．



(a) $x_1(t)$ 和 $\hat{x}_1(t+2|t)$　　(b) $x_2(t)$ 和 $\hat{x}_2(t+2|t)$

图 3.6.4　状态 $x_i(t)$和稳态 Kalman 预报器 $\hat{x}_i(t+2|t)$

【例 3.6.4】　考虑随机系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}1.3 & 1\\ -0.4 & 0\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\ -0.5\end{bmatrix}w(t) \tag{3.6.73}$$

$$y(t)=\begin{bmatrix}1 & 0\end{bmatrix}\boldsymbol{x}(t)+v(t) \tag{3.6.74}$$

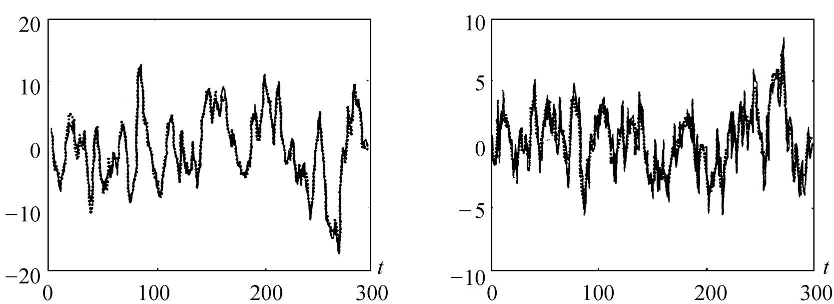
其中 $\boldsymbol{x}(t)=[(x_1(t),x_2(t)]^{\mathrm{T}}$，$w(t)$和 $v(t)$是零均值、方差各为 $Q=1$ 和 $R=1$ 的独立高斯白噪声. 应用定理 3.6.2 可求得稳态 Kalman 滤波器 $\hat{\boldsymbol{x}}(t|t+2)$. 在 $t=300$ 处可求得稳态 Riccati 方程解为

$$\boldsymbol{\Sigma}=\begin{bmatrix}3.694\,288\,238\,652\,3 & 0.544\,385\,825\,951\,77\\ 0.544\,385\,825\,951\,77 & 1.125\,916\,030\,607\,62\end{bmatrix} \tag{3.6.75}$$

应用(3.6.39)可求得稳态 Kalman 滤波器增益 K 为

$$\boldsymbol{K}=\begin{bmatrix}0.786\,975\,191\,297\,64\\ 0.115\,967\,686\,433\,65\end{bmatrix} \tag{3.6.76}$$

仿真结果如图 3.6.5 所示，其中实线代表真实值，虚线代表平滑估值.



(a) $x_1(t)$ 和 $\hat{x}_1(t|t+2)$　　(b) $x_2(t)$ 和 $\hat{x}_2(t|t+2)$

图 3.6.5　状态 $x_i(t)$和稳态 Kalman 平滑器 $\hat{x}_i(t|t+2)$

【例 3.6.5】　Bernoulli – Gaussian 稳态白噪声平滑器，考虑随机系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}0.1 & 1\\ 0.8 & 0\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\ 3\end{bmatrix}w(t) \tag{3.6.77}$$

$$y(t)=\begin{bmatrix}1 & 0\end{bmatrix}\boldsymbol{x}(t)+v(t) \tag{3.6.78}$$

其中 $w(t)=b(t)g(t)$为 Bernoulli – Gaussian 白噪声，$b(t)$是取值 1 或 0，Bernoulli 白噪声，

取值概率为 $P(b(t)=1)=\lambda$，$P(b(t)=1)=1-\lambda$，$g(t)$是零均值、方差为 σ_g^2 的独立于 $b(t)$的 Gaussian 白噪声.应用定理 3.6.4 可求稳态白噪声平滑 $\hat{w}(t|t+1)$和 $\hat{w}(t|t+2)$.在 $t=300$ 处可求得稳态 Riccati 方程解为

$$\boldsymbol{\Sigma}=\begin{bmatrix}2.706\,688\,621\,883\,8 & 0.903\,457\,521\,910\,50\\0.903\,457\,521\,910\,50 & 2.706\,376\,441\,910\,26\end{bmatrix} \tag{3.6.79}$$

进而可求得稳态 Kalman 滤波器增益为

$$\boldsymbol{K}=\begin{bmatrix}0.996\,319\,048\,478\,76\\0.332\,558\,333\,964\,95\end{bmatrix} \tag{3.6.80}$$

仿真结果如图 3.6.6 所示，其中实线端点纵标代表真实值 $w(t)$，圆点纵标代表平滑估值 $\hat{w}(t|t+1)$和 $\hat{w}(t|t+2)$.可看到 $\hat{w}(t|t+2)$的精度比 $\hat{w}(t|t+1)$高.

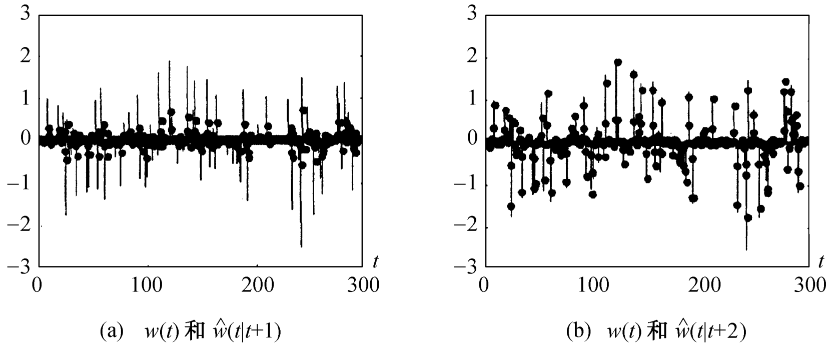


图 3.6.6 Bernoulli – Gaussian 白噪声 $w(t)$和稳态平滑器 $\hat{w}(t|t+1)$及 $\hat{w}(t|t+2)$

3.7 带相关噪声的时变系统最优 Kalman 滤波和最优白噪声估值器

本节将推广节 3.3 和节 3.4 的结果到一般带相关噪声的时变系统情形.给出了两套 Kalman 滤波、预报和平滑算法，并证明了它们的等价性.

3.7.1 最优 Kalman 滤波器和预报器

考虑线性离散时变随机控制系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t) \tag{3.7.1}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}(t)\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{3.7.2}$$

其中状态 $\boldsymbol{x}(t)\in R^n$，观测 $\boldsymbol{y}(t)\in R^m$，控制 $\boldsymbol{u}(t)\in R^p$，$\boldsymbol{\Phi}(t)$，$\boldsymbol{B}(t)$，$\boldsymbol{\Gamma}(t)$和 $\boldsymbol{H}(t)$是已知的时变的适当维数矩阵.

【假设 1】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}(t)\in R^m$ 是带零均值的相关白噪声：

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\\boldsymbol{v}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{w}^{\mathrm{T}}(j) & \boldsymbol{v}^{\mathrm{T}}(j)\end{bmatrix}\right\}=\begin{bmatrix}\boldsymbol{Q}_w(t) & \boldsymbol{S}(t)\\\boldsymbol{S}^{\mathrm{T}}(t) & \boldsymbol{Q}_v(t)\end{bmatrix}\delta_{tj} \tag{3.7.3}$$

其中 E 为均值号，T 为转置号，$\delta_{tt}=1$，$\delta_{tj}=0\ (t\neq j)$.

【假设 2】 控制 $\boldsymbol{u}(t)$是已知的时间序列，或是$(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$的线性函数(反馈

控制).

【假设3】 $\boldsymbol{x}(0)$不相关于$\boldsymbol{w}(t)$和$\boldsymbol{v}(t)$,且$\mathrm{E}\boldsymbol{x}(0)=\boldsymbol{\mu}$,$\mathrm{cov}\boldsymbol{x}(0)=\boldsymbol{P}_0$,其中cov为协方差号.

问题是基于观测$(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots\boldsymbol{y}(1),\boldsymbol{y}(0))$和$\boldsymbol{u}(t-1),\boldsymbol{u}(t-2),\cdots,\boldsymbol{u}(0)$求状态$\boldsymbol{x}(t)$的Kalman滤波器和Kalman预报器.

定义基于观测$(\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(0))$对状态$\boldsymbol{x}(t)$的线性最小方差估值器$\hat{\boldsymbol{x}}(t|t-1)$为Kalman预报器.由递推射影公式有

$$\hat{\boldsymbol{x}}(t+1|t)=\hat{\boldsymbol{x}}(t+1|t-1)+\mathrm{E}[\boldsymbol{x}(t+1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]\mathrm{E}[\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]^{-1}\boldsymbol{\varepsilon}(t) \tag{3.7.4}$$

其中$\boldsymbol{\varepsilon}(t)$是新息过程,对(3.7.2)两边取射影运算有

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{y}}(t|t-1)=\boldsymbol{y}(t)-\boldsymbol{H}(t)\hat{\boldsymbol{x}}(t|t-1) \tag{3.7.5}$$

记预报误差$\tilde{\boldsymbol{x}}(t|t-1)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t-1)$,由(3.7.2)有

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t) \tag{3.7.6}$$

由(3.7.1)和(3.7.6),并注意$\boldsymbol{x}(t)=\tilde{\boldsymbol{x}}(t|t-1)+\hat{\boldsymbol{x}}(t|t-1)$,有

$$\begin{aligned}\mathrm{E}[\boldsymbol{x}(t+1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]=\mathrm{E}[(\boldsymbol{\Phi}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t))(\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t))^{\mathrm{T}}]=\\ \boldsymbol{\Phi}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\end{aligned} \tag{3.7.7}$$

其中定义预报误差方差阵$\boldsymbol{P}(t|t-1)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1)]$,并应用了假设1和假设2,以及$\hat{\boldsymbol{x}}(t|t-1)$不相关于(正交)$\tilde{\boldsymbol{x}}(t|t-1)$的事实.由(3.7.6)有新息方差阵为

$$\boldsymbol{Q}_{\varepsilon}(t)=\mathrm{E}[\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]=\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{Q}_v(t) \tag{3.7.8}$$

对(3.7.1)取射影运算有

$$\hat{\boldsymbol{x}}(t+1|t-1)=\boldsymbol{\Phi}(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}(t)\boldsymbol{u}(t) \tag{3.7.9}$$

将(3.7.7)~(3.7.9)代入(3.7.4)有递推Kalman预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{K}_p(t)\boldsymbol{\varepsilon}(t) \tag{3.7.10}$$

$$\boldsymbol{K}_p(t)=[\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)]\boldsymbol{Q}_{\boldsymbol{\varepsilon}}^{-1}(\boldsymbol{t}) \tag{3.7.11}$$

由(3.7.1),(3.7.6),(3.7.10)引出预报误差递推方程

$$\tilde{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{K}_p(t)\boldsymbol{v}(t) \tag{3.7.12}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t) \tag{3.7.13}$$

这引出递推预报误差方差阵方程

$$\begin{aligned}\boldsymbol{P}(t+1|t)=[\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t)]\boldsymbol{P}(t|t-1)[\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t)]^{\mathrm{T}}+\\ [\boldsymbol{\Gamma}(t),-\boldsymbol{K}_p(t)]\begin{bmatrix}\boldsymbol{Q}_w(t) & \boldsymbol{S}(t)\\ \boldsymbol{S}^{\mathrm{T}}(t) & \boldsymbol{Q}_v(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Gamma}^{\mathrm{T}}(t)\\ -\boldsymbol{K}_p^{\mathrm{T}}(t)\end{bmatrix}\end{aligned} \tag{3.7.14}$$

或

$$\begin{aligned}\boldsymbol{P}(t+1|t)=[\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t)]\boldsymbol{P}(t|t-1)[\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t)]^{\mathrm{T}}+\\ \boldsymbol{\Gamma}(t)\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{K}_p(t)\boldsymbol{S}^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\\ \boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{K}_p^{\mathrm{T}}(t)+\boldsymbol{K}_p(t)\boldsymbol{Q}_v(t)\boldsymbol{K}_p^{\mathrm{T}}(t)\end{aligned} \tag{3.7.15}$$

将(3.7.1)减(3.7.10)引出

$$\tilde{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{K}_p(t)\boldsymbol{\varepsilon}(t) \tag{3.7.16}$$

因$\boldsymbol{\varepsilon}(t)\perp\tilde{\boldsymbol{x}}(t+1|t)$,将上式右边第三项移到左边,这引出

$$P(t+1|t)=\boldsymbol{\Phi}(t)P(t|t-1)\boldsymbol{\Phi}^{\mathrm{T}}(t)-K_p(t)Q_\varepsilon(t)K_p^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)Q_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.7.17}$$

将(3.7.11)代入(3.7.17)后，上述结果可概括为如下定理.

【定理 3.7.1】 系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下有最优递推 Kalman 预报器为

$$\hat{x}(t+1|t)=\boldsymbol{\Phi}(t)\hat{x}(t|t-1)+B(t)u(t)+K_p(t)\boldsymbol{\varepsilon}(t) \tag{3.7.18}$$

$$\boldsymbol{\varepsilon}(t)=y(t)-H(t)\hat{x}(t|t-1) \tag{3.7.19}$$

或

$$\hat{x}(t+1|t)=\boldsymbol{\Psi}_p(t)\hat{x}(t|t-1)+B(t)u(t)+K_p(t)y(t) \tag{3.7.20}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-K_p(t)H(t) \tag{3.7.21}$$

带初值 $\hat{x}(0|-1)=\boldsymbol{\mu}$. Kalman 预报器增益 $K_p(t)$ 为

$$K_p(t)=[\boldsymbol{\Phi}(t)P(t|t-1)H^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)S(t)]Q_\varepsilon^{-1}(t) \tag{3.7.22}$$

$$Q_\varepsilon(t)=H(t)P(t|t-1)H^{\mathrm{T}}(t)+Q_v(t) \tag{3.7.23}$$

预报误差方差阵 $P(t|t-1)$满足 Riccati 方程

$$\begin{aligned}P(t+1|t)=&\boldsymbol{\Phi}(t)P(t|t-1)\boldsymbol{\Phi}^{\mathrm{T}}(t)-[\boldsymbol{\Phi}(t)P(t|t-1)H^{\mathrm{T}}(t)+\\&\boldsymbol{\Gamma}(t)S(t)][H(t)P(t|t-1)H^{\mathrm{T}}(t)+Q_v(t)]^{-1}[\boldsymbol{\Phi}(t)P(t|t-1)\times\\&H^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)S(t)]^{\mathrm{T}}+\boldsymbol{\Gamma}(t)Q_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)\end{aligned} \tag{3.7.24}$$

带初值 $P(0|-1)=P_0$.

注意，取初值 $\hat{x}(0|-1)=\boldsymbol{\mu}$，$P(0|-1)=P_0$ 是为了保证估值器的无偏性：$\mathrm{E}\hat{x}(t|t-1)=\mathrm{E}x(t)$.

下面推导最优递推 Kalman 滤波器.

由递推射影公式有

$$\hat{x}(t+1|t+1)=\hat{x}(t+1|t)+\mathrm{E}[x(t+1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)]Q_\varepsilon^{-1}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.25}$$

应用 $x(t+1)=\hat{x}(t+1|t)+\tilde{x}(t+1|t)$及(3.7.6)引出

$$\mathrm{E}[x(t+1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)]=P(t+1|t)H^{\mathrm{T}}(t+1) \tag{3.7.26}$$

于是有

$$\hat{x}(t+1|t+1)=\hat{x}(t+1|t)+K_f(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.27}$$

其中滤波增益 $K_f(t+1)$为

$$K_f(t+1)=P(t+1|t)H^{\mathrm{T}}(t+1)Q_\varepsilon^{-1}(t+1) \tag{3.7.28}$$

由(3.7.7)有滤波误差 $\tilde{x}(t+1|t+1)=x(t+1)-\hat{x}(t+1|t+1)$与预报误差 $\tilde{x}(t+1|t)$则有关系

$$\tilde{x}(t+1|t+1)=K_f(t+1)\boldsymbol{\varepsilon}(t+1)+\tilde{x}(t+1|t) \tag{3.7.29}$$

利用 $\tilde{x}(t+1|t+1)$与 $\boldsymbol{\varepsilon}(t+1)$有正交性引出

$$\begin{aligned}P(t+1|t+1)=&P(t+1|t)-P(t+1|t)H^{\mathrm{T}}(t+1)\times\\&[H(t+1)P(t+1|t)H^{\mathrm{T}}(t+1)+Q_v(t+1)]^{-1}H(t+1)P(t+1|t)\end{aligned} \tag{3.7.30}$$

或

$$\begin{aligned}P(t+1|t+1)=&P(t+1|t)-P(t+1|t)H^{\mathrm{T}}(t+1)\times\\&Q_\varepsilon^{-1}(t+1)H(t+1)P(t+1|t)\end{aligned} \tag{3.7.31}$$

或

$$\boldsymbol{P}(t+1|t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\boldsymbol{P}(t+1|t) \tag{3.7.32}$$

【定理 3.7.2】 系统(3.7.1)和(3.7.2)在假设 1~3 下,有最优递推 Kalman 滤波器

$$\hat{\boldsymbol{x}}(t+1|t+1)=\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}_f(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.33}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H}(t+1)\hat{\boldsymbol{x}}(t+1|t) \tag{3.7.34}$$

$$\boldsymbol{K}_f(t+1)=\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)[\boldsymbol{H}(t+1)\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)+\boldsymbol{Q}_v(t+1)]^{-1} \tag{3.7.35}$$

$$\boldsymbol{P}(t+1|t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\boldsymbol{P}(t+1|t) \tag{3.7.36}$$

其中 $\hat{\boldsymbol{x}}(t+1|t)$,$\boldsymbol{P}(t+1|t)$由定理 3.7.1 计算.

下面我们给出另一种形式的等价的 Kalman 滤波器和预报器算法.推导原理是将带相关噪声系统转化为带不相关噪声系统.

在(3.7.1)右边形式地加上一项等于零的项:

$$\begin{aligned}\boldsymbol{x}(t+1)=&\boldsymbol{\Phi}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)+\\&\boldsymbol{J}(t)[\boldsymbol{y}(t)-\boldsymbol{H}(t)\boldsymbol{x}(t)-\boldsymbol{v}(t)]\end{aligned} \tag{3.7.37}$$

其中 $\boldsymbol{J}(t)$是待定矩阵,记

$$\begin{aligned}\bar{\boldsymbol{\Phi}}(t)&=\boldsymbol{\Phi}(t)-\boldsymbol{J}(t)\boldsymbol{H}(t),\\ \bar{\boldsymbol{w}}(t)&=\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{J}(t)\boldsymbol{v}(t)\end{aligned} \tag{3.7.38}$$

则状态方程(3.7.37)化为

$$\boldsymbol{x}(t+1)=\bar{\boldsymbol{\Phi}}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{J}(t)\boldsymbol{y}(t)+\bar{\boldsymbol{w}}(t) \tag{3.7.39}$$

观测方程仍为(3.7.2),注意 $\bar{\boldsymbol{w}}(t)$仍为白噪声,且有

$$\mathrm{E}\bar{\boldsymbol{w}}(t)=\boldsymbol{0},\quad \mathrm{E}[\bar{\boldsymbol{w}}(t)\boldsymbol{v}^{\mathrm{T}}(j)]=[\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)-\boldsymbol{J}(t)\boldsymbol{Q}_v(t)]\delta_{tj} \tag{3.7.40}$$

由上式看到,取 $\boldsymbol{J}(t)$为

$$\boldsymbol{J}(t)=\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t) \tag{3.7.41}$$

则 $\bar{\boldsymbol{w}}(t)$与 $\boldsymbol{v}(t)$不相关,即

$$\mathrm{E}[\bar{\boldsymbol{w}}(t)\boldsymbol{v}^{\mathrm{T}}(j)]=\boldsymbol{0},\quad \forall\, t,j \tag{3.7.42}$$

易知白噪声 $\bar{\boldsymbol{w}}(t)$的协方差阵为

$$\mathrm{E}[\bar{\boldsymbol{w}}(t)\bar{\boldsymbol{w}}^{\mathrm{T}}(j)]=\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t)-\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t)\delta_{tj} \tag{3.7.43}$$

对(3.7.39)取射影运算引出滤波与预报关系

$$\hat{\boldsymbol{x}}(t+1|t)=\bar{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{x}}(t|t)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{J}(t)\boldsymbol{y}(t) \tag{3.7.44}$$

由(3.7.39)减(3.7.44)引出预报误差与滤波误差的关系

$$\tilde{\boldsymbol{x}}(t+1|t)=\bar{\boldsymbol{\Phi}}(t)\tilde{\boldsymbol{x}}(t|t)+\bar{\boldsymbol{w}}(t) \tag{3.7.45}$$

于是有关系

$$\begin{aligned}\boldsymbol{P}(t+1|t)=&\bar{\boldsymbol{\Phi}}(t)\boldsymbol{P}(t|t)\bar{\boldsymbol{\Phi}}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t)-\\&\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t)\end{aligned} \tag{3.7.46}$$

由(3.7.38),(3.7.27),(3.7.44),(3.7.45)和(3.7.6)引出滤波误差递推方程

$$\begin{aligned}\tilde{\boldsymbol{x}}(t+1|t+1)=&\bar{\boldsymbol{\Phi}}(t)\tilde{\boldsymbol{x}}(t|t)+\bar{\boldsymbol{w}}(t)-\boldsymbol{K}_f(t+1)\times\\&\{\boldsymbol{H}(t+1)[\bar{\boldsymbol{\Phi}}(t)\tilde{\boldsymbol{x}}(t|t)+\bar{\boldsymbol{w}}(t)]+\boldsymbol{v}(t+1)\}\end{aligned} \tag{3.7.47}$$

即

$$\tilde{\boldsymbol{x}}(t+1|t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\bar{\boldsymbol{\Phi}}(t)\tilde{\boldsymbol{x}}(t|t)+$$
$$[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\bar{\boldsymbol{w}}(t)-\boldsymbol{K}_f(t+1)\boldsymbol{v}(t+1) \tag{3.7.48}$$

注意(3.7.27)和(3.7.6)有

$$\hat{\boldsymbol{x}}(t|t)=\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_f(t)[\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t)] \tag{3.7.49}$$

这引出

$$\tilde{\boldsymbol{x}}(t|t)=\tilde{\boldsymbol{x}}(t|t-1)-\boldsymbol{K}_f(t)[\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t)]=$$
$$[\boldsymbol{I}_n-\boldsymbol{K}_f(t)\boldsymbol{H}(t)]\tilde{\boldsymbol{x}}(t|t-1)-\boldsymbol{K}_f(t)\boldsymbol{v}(t) \tag{3.7.50}$$

将它代入(3.7.45)引出预报误差的另一种不同于(3.7.12)的递推方程

$$\tilde{\boldsymbol{x}}(t+1|t)=\bar{\boldsymbol{\Phi}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t)\boldsymbol{H}(t)]\tilde{\boldsymbol{x}}(t|t-1)-$$
$$\bar{\boldsymbol{\Phi}}(t)\boldsymbol{K}_f(t)\boldsymbol{v}(t)+\bar{\boldsymbol{w}}(t) \tag{3.7.51}$$

注意(3.7.39)和(3.7.40),(3.7.43),在(3.7.48)中 $\tilde{\boldsymbol{x}}(t|t)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t)$ 与 $\bar{\boldsymbol{w}}(t)$ 是不相关的,显然 $\boldsymbol{v}(t+1)$ 不相关于 $\bar{\boldsymbol{w}}(t)$ 和 $\tilde{\boldsymbol{x}}(t|t)$,因此可得递推滤波误差方差阵方程 $\boldsymbol{P}(t+1|t+1)=\mathrm{E}[\tilde{\boldsymbol{x}}(t+1|t+1)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+1|t+1)]$ 为

$$\boldsymbol{P}(t+1|t+1)=\boldsymbol{\Psi}_f(t+1)\boldsymbol{P}(t|t)\boldsymbol{\Psi}_f^{\mathrm{T}}(t+1)+[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\times$$
$$\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t)-\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]^{\mathrm{T}}+$$
$$\boldsymbol{K}_f(t+1)\boldsymbol{Q}_v(t+1)\boldsymbol{K}_f^{\mathrm{T}}(t+1) \tag{3.7.52}$$

其中定义

$$\boldsymbol{\Psi}_f(t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\bar{\boldsymbol{\Phi}}(t) \tag{3.7.53}$$

同理由(3.7.51)可得不同于(3.7.15)的另一种形式的预报误差方差阵递推方程

$$\boldsymbol{P}(t+1|t)=\boldsymbol{\Psi}_p(t)\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\bar{\boldsymbol{K}}_p(t)\boldsymbol{Q}_v(t)\bar{\boldsymbol{K}}_p^{\mathrm{T}}(t)+$$
$$\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t)-\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.7.54}$$

其中定义

$$\boldsymbol{\Psi}_p(t)=\bar{\boldsymbol{\Phi}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t)\boldsymbol{H}(t)]=\bar{\boldsymbol{\Phi}}(t)-\bar{\boldsymbol{K}}_p(t)\boldsymbol{H}(t) \tag{3.7.55}$$

$$\bar{\boldsymbol{K}}_p(t)=\bar{\boldsymbol{\Phi}}(t)\boldsymbol{K}_f(t) \tag{3.7.56}$$

由(3.7.27)和(3.7.44)还有另一种形式 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\bar{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{J}(t)\boldsymbol{y}(t)+\bar{\boldsymbol{K}}_p(t)\boldsymbol{\varepsilon}(t) \tag{3.7.57}$$

利用(3.7.19)有相同于(3.7.20)的表达式

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.7.58}$$

其中定义预报增益为

$$\boldsymbol{K}_p(t)=\bar{\boldsymbol{K}}_p(t)+\boldsymbol{J}(t) \tag{3.7.59}$$

比较(3.7.20)和(3.7.21)与(3.7.56)和(3.7.57)有关系

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t)=\bar{\boldsymbol{\Phi}}(t)-\bar{\boldsymbol{K}}_p(t)\boldsymbol{H}(t)=$$
$$\bar{\boldsymbol{\Phi}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t)\boldsymbol{H}(t)] \tag{3.7.60}$$

其中 $\boldsymbol{K}_p(t)$ 用(3.7.22)定义,且它等值于(3.7.59)定义的 $\boldsymbol{K}_p(t)$.

由(3.7.19),(3.7.20),(3.7.27)和(3.7.44)有递推 Kalman 滤波器

$$\hat{\boldsymbol{x}}(t+1|t+1)=\boldsymbol{\Psi}_f(t+1)\hat{\boldsymbol{x}}(t|t)+\boldsymbol{C}_1(t)\boldsymbol{u}(t)+$$
$$\boldsymbol{C}_2(t)\boldsymbol{y}(t)+\boldsymbol{K}_f(t+1)\boldsymbol{y}(t+1) \tag{3.7.61}$$

其中用(3.7.53)定义 $\boldsymbol{\Psi}_f(t)$，且定义

$$\boldsymbol{C}_1(t) = [\boldsymbol{I}_n - \boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\boldsymbol{B}(t),$$
$$\boldsymbol{C}_2(t) = [\boldsymbol{I}_n - \boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\boldsymbol{J}(t) \tag{3.7.62}$$

【定理 3.7.3】 系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下有最优递推 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t) = \boldsymbol{\Psi}_p(t)\hat{\boldsymbol{x}}(t|t-1) + \boldsymbol{B}(t)\boldsymbol{u}(t) + \boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.7.63}$$

$$\boldsymbol{\Psi}_p(t) = \bar{\boldsymbol{\Phi}}(t) - \bar{\boldsymbol{K}}_p(t)\boldsymbol{H}(t), \quad \bar{\boldsymbol{K}}_p(t) = \bar{\boldsymbol{\Phi}}(t)\boldsymbol{K}_f(t) \tag{3.7.64}$$

$$\boldsymbol{K}_p(t) = \bar{\boldsymbol{K}}_p(t) + \boldsymbol{J}(t) \tag{3.7.65}$$

且预报误差方差阵有递推方程

$$\boldsymbol{P}(t+1|t) = \boldsymbol{\Psi}_p(t)\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t) + \bar{\boldsymbol{K}}_p(t)\boldsymbol{Q}_v(t)\bar{\boldsymbol{K}}_p^{\mathrm{T}}(t) +$$
$$\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t) - \boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.7.66}$$

带初值 $\hat{\boldsymbol{x}}(0|-1) = \boldsymbol{\mu}_0, \boldsymbol{P}(0|-1) = \boldsymbol{P}_0$.

证明 见(3.7.51) ~ (3.7.60). □

【定理 3.7.4】 系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下有最优递推 Kalman 滤波器

$$\hat{\boldsymbol{x}}(t+1|t+1) = \hat{\boldsymbol{x}}(t+1|t) + \boldsymbol{K}_f(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.67}$$

$$\boldsymbol{\varepsilon}(t+1) = \boldsymbol{y}(t+1) - \boldsymbol{H}(t+1)\hat{\boldsymbol{x}}(t+1|t) \tag{3.7.68}$$

$$\boldsymbol{K}_f(t+1) = \boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)[\boldsymbol{H}(t+1)\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1) + \boldsymbol{Q}_v(t+1)]^{-1} \tag{3.7.69}$$

$$\boldsymbol{P}(t+1|t+1) = \boldsymbol{P}(t+1|t) - \boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)[\boldsymbol{H}(t+1)\boldsymbol{P}(t+1|t)\times$$
$$\boldsymbol{H}^{\mathrm{T}}(t+1) + \boldsymbol{Q}_v(t+1)]^{-1}\boldsymbol{H}(t+1)\boldsymbol{P}(t+1|t) \tag{3.7.70}$$

$$\hat{\boldsymbol{x}}(t+1|t) = \bar{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{x}}(t|t) + \boldsymbol{B}(t)\boldsymbol{u}(t) + \boldsymbol{J}(t)\boldsymbol{y}(t) \tag{3.7.71}$$

$$\boldsymbol{P}(t+1|t) = \bar{\boldsymbol{\Phi}}(t)\boldsymbol{P}(t|t)\bar{\boldsymbol{\Phi}}^{\mathrm{T}}(t) + \boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t) - \boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.7.72}$$

带初值 $\hat{\boldsymbol{x}}(0|-1) = \boldsymbol{\mu}, \boldsymbol{P}(0|-1) = \boldsymbol{P}_0$.

合并(3.7.70)和(3.7.72)有 $\boldsymbol{P}(t+1|t)$满足如下 Riccati 方程

$$\boldsymbol{P}(t+1|t) = \bar{\boldsymbol{\Phi}}(t)[\boldsymbol{P}(t|t-1) - \boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)(\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t) +$$
$$\boldsymbol{Q}_v(t))^{-1}\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)]\bar{\boldsymbol{\Phi}}^{\mathrm{T}}(t) +$$
$$\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t) - \boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.7.73}$$

它等价于 Riccati 方程(3.7.24).

证明 见(3.7.8)，(3.7.27)，(3.7.28)，(3.7.30)，(3.7.44)，(3.7.46). □

【定理 3.7.5】 系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下有最优递推 Kalman 滤波器

$$\hat{\boldsymbol{x}}(t+1|t+1) = \boldsymbol{\Psi}_f(t+1)\hat{\boldsymbol{x}}(t|t) + \boldsymbol{C}_1(t)\boldsymbol{u}(t) +$$
$$\boldsymbol{C}_2(t)\boldsymbol{y}(t) + \boldsymbol{K}_f(t+1)\boldsymbol{y}(t+1) \tag{3.7.74}$$

其中 $\boldsymbol{C}_1(t)$和 $\boldsymbol{C}_2(t)$由(3.7.62)定义，$\boldsymbol{\Psi}_f(t+1)$由(3.7.53)定义，且有滤波误差方差递推方程

$$\boldsymbol{P}(t+1|t+1) = \boldsymbol{\Psi}_f(t+1)\boldsymbol{P}(t|t)\boldsymbol{\Psi}_f^{\mathrm{T}}(t+1) + [\boldsymbol{I}_n - \boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\times$$
$$\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t) - \boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t)[\boldsymbol{I}_n - \boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]^{\mathrm{T}} +$$
$$\boldsymbol{K}_f(t+1)\boldsymbol{Q}_v(t+1)\boldsymbol{K}_f^{\mathrm{T}}(t+1) \tag{3.7.75}$$

其中 $\boldsymbol{\Psi}_f(t+1)$ 由(3.7.53)定义，初值 $\hat{\boldsymbol{x}}(0|0)$ 和 $\boldsymbol{P}(0|0)$ 由(3.7.67)和(3.7.70)计算.

证明 见(3.7.52)和(3.7.61). □

【定理 3.7.6】 下述两组表达式是等价(等值)的：

$$\boldsymbol{K}_p(t)=[\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)]\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.7.76}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t) \tag{3.7.77}$$

与

$$\boldsymbol{K}_p(t)=\bar{\boldsymbol{K}}_p(t)+\boldsymbol{J}(t),\bar{\boldsymbol{K}}_p(t)=\bar{\boldsymbol{\Phi}}(t)\boldsymbol{K}_f(t) \tag{3.7.78}$$

$$\boldsymbol{\Psi}_p(t)=\bar{\boldsymbol{\Phi}}(t)-\bar{\boldsymbol{K}}_p(t)\boldsymbol{H}(t) \tag{3.7.79}$$

其中 $\boldsymbol{K}_f(t)$ 由(3.7.69)定义，$\bar{\boldsymbol{\Phi}}(t)$，$\boldsymbol{J}(t)$ 由(3.7.38)和(3.7.41)定义.

证明 先证(3.7.76)与(3.7.78)相等. 由(3.7.76)和(3.7.28)有

$$\boldsymbol{K}_p(t)=\boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.7.80}$$

另一方面，由(3.7.78)，(3.7.38)和(3.7.41)有

$$\begin{aligned}\boldsymbol{K}_p(t)=\bar{\boldsymbol{\Phi}}(t)\boldsymbol{K}_f(t)+\boldsymbol{J}(t)=[\boldsymbol{\Phi}(t)-\boldsymbol{J}(t)\boldsymbol{H}(t)]\boldsymbol{K}_f(t)+\boldsymbol{J}(t)=\\ \boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)+\boldsymbol{J}(t)[\boldsymbol{I}_n-\boldsymbol{H}(t)\boldsymbol{K}_f(t)]=\\ \boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)[\boldsymbol{I}_n-\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{Q}_\varepsilon^{-1}(t)]=\\ \boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)[\boldsymbol{I}_m-(\boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t)+\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{Q}_\varepsilon^{-1}(t))+\\ \boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t)]=\boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t)=\\ \boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_\varepsilon^{-1}(t)\end{aligned} \tag{3.7.81}$$

其中用到了由(3.7.23)引出的事实

$$\begin{aligned}\boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t)+\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{Q}_\varepsilon^{-1}(t)=\\ [\boldsymbol{Q}_v(t)+\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)]\boldsymbol{Q}_\varepsilon^{-1}(t)=\boldsymbol{I}_m\end{aligned} \tag{3.7.82}$$

这证明了(3.7.76)等值于(3.7.78).

下证(3.7.77)等值于(3.7.79).

将(3.7.76)代入(3.7.77)有

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)+\boldsymbol{H}(t)\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_\varepsilon^{-1}(t)\boldsymbol{H}(t) \tag{3.7.83}$$

另一方面将(3.7.78)代入(3.7.79)，并应用定义(3.7.38)和(3.7.41)有

$$\begin{aligned}\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{J}(t)\boldsymbol{H}(t)-(\boldsymbol{\Phi}(t)-\boldsymbol{J}(t)\boldsymbol{H}(t))\boldsymbol{K}_f(t)\boldsymbol{H}(t)=\\ \boldsymbol{\Phi}(t)-\boldsymbol{J}(t)\boldsymbol{H}(t)-\boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)\boldsymbol{H}(t)+\boldsymbol{J}(t)\boldsymbol{H}(t)\boldsymbol{K}_f(t)\boldsymbol{H}(t)=\\ \boldsymbol{\Phi}(t)-\boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)\boldsymbol{H}(t)-\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\times\\ [\boldsymbol{I}_m-\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{Q}_\varepsilon^{-1}(t)]\boldsymbol{H}(t)=\\ \boldsymbol{\Phi}(t)-\boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)\boldsymbol{H}(t)-\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)[\boldsymbol{I}_m-(\boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t)+\\ \boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{Q}_\varepsilon^{-1}(t))+\boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t)]\boldsymbol{H}(t)=\\ \boldsymbol{\Phi}(t)-\boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)\boldsymbol{H}(t)-\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t)\boldsymbol{H}(t)=\\ \boldsymbol{\Phi}(t)-\boldsymbol{\Phi}(t)\boldsymbol{K}_f(t)\boldsymbol{H}(t)-\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_\varepsilon^{-1}(t)\boldsymbol{H}(t)\end{aligned} \tag{3.7.84}$$

比较(3.7.83)与(3.7.84)引出(3.7.77)等值于(3.7.79). 证毕. □

【推论 3.7.1】 系统(3.7.1)和(3.7.2)在假设 1~3 下，若噪声是不相关的，即 $\boldsymbol{S}(t)=\boldsymbol{0}$，则有最优递推 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.7.85}$$

$$\boldsymbol{K}_p(t)=\boldsymbol{\Phi}(t)\boldsymbol{K}_f(t) \tag{3.7.86}$$

$$\boldsymbol{K}_f(t)=\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)[\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{Q}_v(t)]^{-1} \tag{3.7.87}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t)=\boldsymbol{\Phi}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t)\boldsymbol{H}(t)] \tag{3.7.88}$$

$$\boldsymbol{P}(t+1|t)=\boldsymbol{\Psi}_p(t)\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\boldsymbol{K}_p(t)\boldsymbol{Q}_v(t)\boldsymbol{K}_p^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.7.89}$$

带初值 $\hat{\boldsymbol{x}}(0|-1)=\boldsymbol{\mu}$，$\boldsymbol{P}(0|-1)=\boldsymbol{P}_0$. 注意，将(3.7.86)和(3.7.87)代入(3.7.89)后的方程叫 Riccati 方程.

【推论 3.7.2】 系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下，若噪声是不相关的，即 $\boldsymbol{S}(t)=\boldsymbol{0}$，则有最优递推 Kalman 滤波器

$$\begin{aligned}\hat{\boldsymbol{x}}(t+1|t+1)=&\boldsymbol{\Psi}_f(t+1)\hat{\boldsymbol{x}}(t|t)+[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\boldsymbol{B}(t)\boldsymbol{u}(t)+\\&\boldsymbol{K}_f(t+1)\boldsymbol{y}(t+1)\end{aligned} \tag{3.7.90}$$

其中 $\boldsymbol{K}_f(t)$ 由(3.7.87)定义，$\boldsymbol{P}(t+1|t)$ 由(3.7.72)定义，且

$$\boldsymbol{\Psi}_f(t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\boldsymbol{\Phi}(t) \tag{3.7.91}$$

$$\begin{aligned}\boldsymbol{P}(t+1|t+1)=&\boldsymbol{\Psi}_f(t+1)\boldsymbol{P}(t|t)\boldsymbol{\Psi}_f^{\mathrm{T}}(t+1)+[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\times\\&\boldsymbol{\Gamma}(t)\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]^{\mathrm{T}}+\boldsymbol{K}_f(t+1)\boldsymbol{Q}_v(t+1)\boldsymbol{K}_f^{\mathrm{T}}(t+1)\end{aligned} \tag{3.7.92}$$

其中 $\hat{\boldsymbol{x}}(0|0)$ 和 $\boldsymbol{P}(0|0)$ 的计算同定理 3.7.4.

【推论 3.7.3】 系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下，若噪声是不相关的，即 $\boldsymbol{S}(t)=\boldsymbol{0}$，则有最优递推 Kalman 滤波器

$$\hat{\boldsymbol{x}}(t+1|t+1)=\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}_f(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.93}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H}(t+1)\hat{\boldsymbol{x}}(t+1|t) \tag{3.7.94}$$

$$\boldsymbol{K}_f(t)=\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)[\boldsymbol{H}(t+1)\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)+\boldsymbol{Q}_v(t+1)]^{-1}$$

$$\begin{aligned}\boldsymbol{P}(t+1|t+1)=&\boldsymbol{P}(t+1|t)-\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)[\boldsymbol{H}(t+1)\boldsymbol{P}(t+1|t)\times\\&\boldsymbol{H}^{\mathrm{T}}(t+1)+\boldsymbol{Q}_v(t+1)]^{-1}\boldsymbol{H}(t+1)]\boldsymbol{P}(t+1|t)\end{aligned} \tag{3.7.95}$$

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}(t)\hat{\boldsymbol{x}}(t|t)+\boldsymbol{B}(t)\boldsymbol{u}(t) \tag{3.7.96}$$

$$\boldsymbol{P}(t+1|t)=\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t)\boldsymbol{\Phi}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.7.97}$$

带初值 $\hat{\boldsymbol{x}}(0|-1)=\boldsymbol{\mu}$，$\boldsymbol{P}(0|-1)=\boldsymbol{P}_0$. $\boldsymbol{P}(t+1|t)$ 满足如下 Riccati 方程

$$\begin{aligned}\boldsymbol{P}(t+1|t)=&\boldsymbol{\Phi}(t)[\boldsymbol{P}(t+1|t)-\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)(\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\\&\boldsymbol{Q}_v(t))^{-1}\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)]\boldsymbol{\Phi}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)\end{aligned} \tag{3.7.98}$$

它可将(3.7.95)代入(3.7.97)得到，也可由(3.7.24)置 $\boldsymbol{S}(t)=\boldsymbol{0}$ 得到.

因为系统(3.7.39)与(3.7.2)为带不相关噪声系统，利用定理 3.3.4 的标准 Kalman 滤波器立刻得如下定理.

【定理 3.7.7】 带相关噪声的时变系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下，有最优 Kalman 滤波器

$$\hat{\boldsymbol{x}}(t+1|t+1)=\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}_f(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.99}$$

$$\hat{\boldsymbol{x}}(t+1|t)=\bar{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{x}}(t|t)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{J}(t)\boldsymbol{y}(t) \tag{3.7.100}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H}(t+1)\hat{\boldsymbol{x}}(t+1|t) \tag{3.7.101}$$

$$\boldsymbol{K}_f(t+1)=\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_{\varepsilon}^{-1}(t+1) \tag{3.7.102}$$

$$\boldsymbol{Q}_{\varepsilon}(t+1)=\boldsymbol{H}(t+1)\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}(t+1)+\boldsymbol{Q}_v(t+1) \tag{3.7.103}$$

$$\boldsymbol{P}(t+1|t)=\overline{\boldsymbol{\Phi}}(t)\boldsymbol{P}(t|t)\overline{\boldsymbol{\Phi}}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t)-\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.7.104}$$

$$\boldsymbol{P}(t+1|t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}(t+1)]\boldsymbol{P}(t+1|t) \tag{3.7.105}$$

带初值 $\hat{\boldsymbol{x}}(0|0)=\boldsymbol{\mu}$，$\boldsymbol{P}(0|0)=\boldsymbol{P}_0$. 且 $\overline{\boldsymbol{\Phi}}(t)$和 $\boldsymbol{J}(t)$由(3.7.38)和(3.7.41)定义. 此外，将(3.7.105)代入(3.7.104)可得 $\boldsymbol{P}(t+1|t)$满足递推 Riccati 方程

$$\boldsymbol{P}(t+1|t)=\overline{\boldsymbol{\Phi}}(t)[\boldsymbol{P}(t|t-1)-\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)[\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{Q}_v(t)]^{-1}\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)]\overline{\boldsymbol{\Phi}}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_w(t)-\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t)\boldsymbol{S}^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{3.7.106}$$

【推论 3.7.4】 对于定常系统(3.7.1)和(3.7.2)，即 $\boldsymbol{\Phi}(t)=\boldsymbol{\Phi}$，$\boldsymbol{B}(t)=\boldsymbol{B}$，$\boldsymbol{\Gamma}(t)=\boldsymbol{\Gamma}$，$\boldsymbol{H}(t)=\boldsymbol{H}$，$\boldsymbol{Q}_w(t)=\boldsymbol{Q}_w$，$\boldsymbol{Q}_v(t)=\boldsymbol{Q}_v$ 和 $\boldsymbol{S}(t)=\boldsymbol{S}$ 均为常阵，在假设 1 ~ 3 下，有最优 Kalman 滤波器

$$\hat{\boldsymbol{x}}(t+1|t+1)=\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.107}$$

$$\hat{\boldsymbol{x}}(t+1|t)=\overline{\boldsymbol{\Phi}}\hat{\boldsymbol{x}}(t|t)+\boldsymbol{B}\boldsymbol{u}(t)++\boldsymbol{J}\boldsymbol{y}(t) \tag{3.7.108}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H}\hat{\boldsymbol{x}}(t+1|t) \tag{3.7.109}$$

$$\boldsymbol{K}(t+1)=\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1) \tag{3.7.110}$$

$$\boldsymbol{Q}_{\varepsilon}(t+1)=\boldsymbol{H}\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v \tag{3.7.111}$$

$$\boldsymbol{P}(t+1|t)=\overline{\boldsymbol{\Phi}}\boldsymbol{P}(t|t)\overline{\boldsymbol{\Phi}}^{\mathrm{T}}+\boldsymbol{\Gamma}[\boldsymbol{Q}_w-\boldsymbol{S}\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}}]\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.7.112}$$

$$\boldsymbol{P}(t+1|t+1)=[\boldsymbol{I}_n-\boldsymbol{K}(t+1)\boldsymbol{H}]\boldsymbol{P}(t+1|t) \tag{3.7.113}$$

$$\overline{\boldsymbol{\Phi}}=\boldsymbol{\Phi}-\boldsymbol{J}\boldsymbol{H},\quad \boldsymbol{J}=\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_v^{-1} \tag{3.7.114}$$

且 $\boldsymbol{P}(t+1|t)$满足 Riccati 方程

$$\boldsymbol{P}(t+1|t)=\overline{\boldsymbol{\Phi}}[\boldsymbol{P}(t|t-1)-\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v)^{-1}\boldsymbol{H}\boldsymbol{P}(t|t-1)]\overline{\boldsymbol{\Phi}}^{\mathrm{T}}+\boldsymbol{\Gamma}(\boldsymbol{Q}_w-\boldsymbol{S}\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.7.115}$$

带初值 $\hat{\boldsymbol{x}}(0|0)=\boldsymbol{\mu}$，$\boldsymbol{P}(0|0)=\boldsymbol{P}_0$.

证明 见定理 3.7.4. □

【推论 3.7.5】 在推论 3.7.4 条件下有最优 Kalman 预报器.

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.7.116}$$

$$\boldsymbol{\Psi}_p(t)=\overline{\boldsymbol{\Phi}}-\overline{\boldsymbol{K}}_p(t)\boldsymbol{H} \tag{3.7.117}$$

$$\overline{\boldsymbol{K}}_p(t)=\overline{\boldsymbol{\Phi}}\boldsymbol{K}_f(t) \tag{3.7.118}$$

$$\boldsymbol{K}_p(t)=\overline{\boldsymbol{K}}_p(t)+\boldsymbol{J} \tag{3.7.119}$$

$$\boldsymbol{K}_f(t)=\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v]^{-1} \tag{3.7.120}$$

$\boldsymbol{P}(t|t-1)$满足 Riccati 方程(3.7.115)或

$$\boldsymbol{P}(t+1|t)=\boldsymbol{\Psi}_p(t)\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\overline{\boldsymbol{K}}_p(t)\boldsymbol{Q}_v\overline{\boldsymbol{K}}_p^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(\boldsymbol{Q}_w-\boldsymbol{S}\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.7.121}$$

证明 见定理 3.7.3. □

【推论 3.7.6】 在推论 3.7.4 下有最优 Kalman 预报器为

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.7.122}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}-\boldsymbol{K}_p(t)\boldsymbol{H} \tag{3.7.123}$$

$$\boldsymbol{K}_p(t)=[\boldsymbol{\Phi}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}\boldsymbol{S}]\boldsymbol{Q}_{\varepsilon}^{-1}(t) \tag{3.7.124}$$

$$\boldsymbol{Q}_{\varepsilon}(t)=\boldsymbol{H}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v \tag{3.7.125}$$

而 $\boldsymbol{P}(t+1|t)$满足 Riccati 方程

$$\begin{aligned}\boldsymbol{P}(t+1|t)=&\boldsymbol{\Phi}\boldsymbol{P}(t|t-1)\boldsymbol{\Phi}^{\mathrm{T}}-[\boldsymbol{\Phi}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S}][\boldsymbol{H}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v]^{-1}\times\\&[\boldsymbol{\Phi}\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S}]^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}\end{aligned} \tag{3.7.126}$$

带初值 $\hat{\boldsymbol{x}}(0|-1)=\boldsymbol{\mu}_0$, $\boldsymbol{P}(0|-1)=\boldsymbol{P}_0$. 而相应 Kalman 滤波器为

$$\hat{\boldsymbol{x}}(t+1|t+1)=\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}_f(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.127}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H}\hat{\boldsymbol{x}}(t+1|t) \tag{3.7.128}$$

$$\boldsymbol{K}_f(t+1)=\boldsymbol{P}(t+1|t)\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}(t+1) \tag{3.7.129}$$

$$\boldsymbol{P}(t+1|t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_f(t+1)\boldsymbol{H}]\boldsymbol{P}(t+1|t) \tag{3.7.130}$$

证明 见定理 3.7.1 和定理 3.7.2. □

3.7.2 超前 N 步最优 Kalman 预报器

考虑 $\boldsymbol{u}(t)=\boldsymbol{0}$ 的系统(3.7.1)和(3.7.2),即

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}(t)\boldsymbol{x}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t) \tag{3.7.131}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}(t)\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{3.7.132}$$

在假设 1 和假设 3 下,有一步最优 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.7.133}$$

其中 $\boldsymbol{\Psi}_p(t)$, $\boldsymbol{K}_p(t)$, $\boldsymbol{P}(t+1|t)$由推论 3.7.5 或推论 3.7.6 计算. 由式(3.7.131)和射影性质,注意 $\boldsymbol{w}(j)\perp L(\boldsymbol{y}(t),\boldsymbol{y}(t-1)\cdots,\boldsymbol{y}(0))$,可得多步 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+2|t)=\boldsymbol{\Phi}(t+1)\hat{\boldsymbol{x}}(t+1|t) \tag{3.7.134}$$

$$\hat{\boldsymbol{x}}(t+3|t)=\boldsymbol{\Phi}(t+2)\hat{\boldsymbol{x}}(t+2|t)=\boldsymbol{\Phi}(t+2)\boldsymbol{\Phi}(t+1)\hat{\boldsymbol{x}}(t+1|t) \tag{3.7.135}$$

用归纳法可得

$$\hat{\boldsymbol{x}}(t+N|t)=\boldsymbol{\Phi}(t+N,t+1)\hat{\boldsymbol{x}}(t+1|t),\quad N>1 \tag{3.7.136}$$

其中定义

$$\begin{gathered}\boldsymbol{\Phi}(t+N,t)=\boldsymbol{\Phi}(t+N-1)\boldsymbol{\Phi}(t+N-2)\cdots\boldsymbol{\Phi}(t),\\\boldsymbol{\Phi}(t+N,t+N)=\boldsymbol{I}_n\end{gathered} \tag{3.7.137}$$

另一方面由(3.7.131)迭代有

$$\boldsymbol{x}(t+N)=\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{x}(t+1)+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}(t+N,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1) \tag{3.7.138}$$

上式减(3.7.136)引出预报误差 $\tilde{\boldsymbol{x}}(t+N|t)=\boldsymbol{x}(t+N)-\hat{\boldsymbol{x}}(t+N|t)$为

$$\tilde{\boldsymbol{x}}(t+N|t)=\boldsymbol{\Phi}(t+N,t+1)\tilde{\boldsymbol{x}}(t+1|t)+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}(t+N,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1) \tag{3.7.139}$$

因 $\boldsymbol{w}(j)\perp\tilde{\boldsymbol{x}}(t+1|t)(j\geqslant t+1)$，于是有预报误差方差阵 $\boldsymbol{P}(t+N|t)=\mathrm{E}[\tilde{\boldsymbol{x}}(t+N|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+N|t)]$为

$$\boldsymbol{P}(t+N|t)=\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{P}(t+1|t)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+1)+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}(t+N,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{Q}_w(i-1)\boldsymbol{\Gamma}^{\mathrm{T}}(i-1)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,i)\tag{3.7.140}$$

最优预报器 $\hat{\boldsymbol{x}}(t|t+N)(N<-1)$可表为类似形式.由(3.7.131)迭代可得关系

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}(t,t+N+1)\boldsymbol{x}(t+N+1)+\sum_{i=t+N+2}^{t}\boldsymbol{\Phi}(t,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1)\tag{3.7.141}$$

类似可得预报器

$$\hat{\boldsymbol{x}}(t|t+N)=\boldsymbol{\Phi}(t,t+N+1)\hat{\boldsymbol{x}}(t+N+1|t+N),\quad N<-1\tag{3.7.142}$$

上两式相减引出预报误差 $\tilde{\boldsymbol{x}}(t|t+N)=\boldsymbol{x}(t)-\tilde{\boldsymbol{x}}(t|t+N)$方差阵 $\boldsymbol{P}(t|t+N)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t+N)]$为

$$\boldsymbol{P}(t|t+N)=\boldsymbol{\Phi}(t,t+N+1)\boldsymbol{P}(t+N+1|t+N)\boldsymbol{\Phi}^{\mathrm{T}}(t,t+N+1)+\sum_{i=t+N+2}^{t}\boldsymbol{\Phi}(t,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{Q}_w(i-1)\boldsymbol{\Gamma}^{\mathrm{T}}(i-1)\boldsymbol{\Phi}^{\mathrm{T}}(t,i)\tag{3.7.143}$$

【定理 3.7.8】 带相关噪声时变系统(3.7.131)和(3.7.132)在节 3.7 假设 1 和假设 3 下有超前 N 步 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+N|t)=\boldsymbol{\Phi}(t+N,t+1)\hat{\boldsymbol{x}}(t+1|t),\quad N>1\tag{3.7.144}$$

或等价地

$$\hat{\boldsymbol{x}}(t|t+N)=\boldsymbol{\Phi}(t,t+N+1)\hat{\boldsymbol{x}}(t+N+1|t+N),\quad N<-1\tag{3.7.145}$$

预报误差方差阵 $\boldsymbol{P}(t+N|t)(N>1)$或 $\boldsymbol{P}(t|t+N)(N<-1)$由(3.7.140)或(3.7.143)给出.

【推论 3.7.7】 对于定常系统(3.7.131)和(3.7.132)，即 $\boldsymbol{\Phi}(t)=\boldsymbol{\Phi},\boldsymbol{H}(t)=\boldsymbol{H},\boldsymbol{\Gamma}(t)=\boldsymbol{\Gamma},\boldsymbol{Q}_w(t)=\boldsymbol{Q}_w,\boldsymbol{Q}_v(t)=\boldsymbol{Q}_v$，定理 3.3.3 成立.

3.7.3 最优 Kalman 平滑器

类似于定理 3.4.1 的推导，可得如下最优 Kalman 平滑器.

【定理 3.7.9】 带相关噪声时变系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下，有最优 Kalman 滤波器和平滑器

$$\hat{\boldsymbol{x}}(t|t+N)=\hat{\boldsymbol{x}}(t|t+N-1)+\boldsymbol{K}(t|t+N)\boldsymbol{\varepsilon}(t+N),\quad N\geqslant 0\tag{3.7.146}$$

带初值 $\hat{\boldsymbol{x}}(t|t-1)$，$N=0,1,2,\cdots$，见规定 $\boldsymbol{K}(t|t)=\boldsymbol{K}_f(t)$，而 $\hat{\boldsymbol{x}}(t|t-1)$，$\boldsymbol{K}_f(t)$，$\boldsymbol{P}(t|t-1)$，$\boldsymbol{\varepsilon}(t+N)$由定理 3.7.3 和定理 3.7.4 计算.平滑增益为

$$\boldsymbol{K}(t|t+N)=\boldsymbol{P}(t|t-1)\left\{\prod_{i=0}^{N-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\right\}\boldsymbol{H}^{\mathrm{T}}(t+N)\boldsymbol{Q}_\varepsilon^{-1}(t+N),\quad N\geqslant 1\tag{3.7.147}$$

$$\boldsymbol{\Psi}_p(t)=\bar{\boldsymbol{\Phi}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t)\boldsymbol{H}(t)]\tag{3.7.148}$$

$$\boldsymbol{Q}_\varepsilon(t+N)=\boldsymbol{H}(t+N)\boldsymbol{P}(t+N|t+N-1)\boldsymbol{H}^{\mathrm{T}}(t+N)+\boldsymbol{Q}_v(t+N)\tag{3.7.149}$$

平滑误差 $\tilde{\boldsymbol{x}}(t|t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t+N)$ 方差阵 $\boldsymbol{P}(t|t+N)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t+N)]$ 为

$$\boldsymbol{P}(t|t+N)=\boldsymbol{P}(t|t+N-1)-\boldsymbol{K}(t|t+N)\boldsymbol{Q}_{\varepsilon}(t+N)\boldsymbol{K}^{\mathrm{T}}(t|t+N) \tag{3.7.150}$$

证明 由递推射影公式有(3.7.146),其中

$$\boldsymbol{K}(t|t+N)=\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]\boldsymbol{Q}_{\varepsilon}^{-1}(t+N) \tag{3.7.151}$$

由(3.7.6)有

$$\boldsymbol{\varepsilon}(t+N)=\boldsymbol{H}(t+N)\tilde{\boldsymbol{x}}(t+N|t+N-1)+\boldsymbol{v}(t+N) \tag{3.7.152}$$

由(3.7.51)有一步预报误差方程

$$\tilde{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\tilde{\boldsymbol{x}}(t|t-1)-\bar{\boldsymbol{\Phi}}(t)\boldsymbol{K}_f(t)\boldsymbol{v}(t)+\bar{\boldsymbol{w}}(t) \tag{3.7.153}$$

由此式迭代 N 次有关系

$$\tilde{\boldsymbol{x}}(t+N|t+N-1)=\boldsymbol{\Psi}(t+N,t)\tilde{\boldsymbol{x}}(t|t-1)+\sum_{i=t+1}^{t+N}\boldsymbol{\Psi}(t+N,i)\times$$
$$[\bar{\boldsymbol{w}}(i-1)-\bar{\boldsymbol{\Phi}}(i-1)\boldsymbol{K}_f(i-1)\boldsymbol{v}(i-1)] \tag{3.7.154}$$

其中定义

$$\boldsymbol{\Psi}(t+N,t)=\boldsymbol{\Psi}_p(t+N-1)\cdots\boldsymbol{\Psi}_p(t),\quad N>0$$
$$\boldsymbol{\Psi}(t+N,t+N)=\boldsymbol{I}_n \tag{3.7.155}$$

由于 $\boldsymbol{x}(t)=\tilde{\boldsymbol{x}}(t|t-1)+\hat{\boldsymbol{x}}(t|t-1)$,而 $\tilde{\boldsymbol{x}}(t|t-1)\perp\hat{\boldsymbol{x}}(t|t-1)$,又 $\tilde{\boldsymbol{x}}(t|t-1)\perp\bar{\boldsymbol{w}}(j)$ $(j\geqslant t)$,$\tilde{\boldsymbol{x}}(t|t-1)\perp v(j)$ $(j\geqslant t)$,将(3.7.154)代入(3.7.152),再将(3.7.152)代入(3.7.151)引出

$$\boldsymbol{K}(t|t+N)=\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}^{\mathrm{T}}(t+N,t)\boldsymbol{H}^{\mathrm{T}}(t+N)\boldsymbol{Q}_{\varepsilon}^{-1}(t+N) \tag{3.7.156}$$

由此立刻得(3.7.147).由(3.7.146)引出误差方程

$$\tilde{\boldsymbol{x}}(t|t+N)=\tilde{\boldsymbol{x}}(t|t+N-1)+\boldsymbol{K}(t|t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.7.157}$$

由于 $\tilde{\boldsymbol{x}}(t|t+N)\perp\boldsymbol{\varepsilon}(t+N)$,将上式右边第二项移到左边后可得(3.7.150).证毕. □

【推论 3.7.8】 在定理 3.7.9 下,非递推 Kalman 平滑器为

$$\hat{\boldsymbol{x}}(t|t+N)=\hat{\boldsymbol{x}}(t|t-1)+\sum_{i=0}^{N}\boldsymbol{K}(t|t+i)\boldsymbol{\varepsilon}(t+i) \tag{3.7.158}$$

且有平滑误差方差阵

$$\boldsymbol{P}(t|t+N)=\boldsymbol{P}(t|t-1)-\sum_{i=0}^{N}\boldsymbol{K}(t|t+i)\boldsymbol{Q}_{\varepsilon}(t+i)\boldsymbol{K}^{\mathrm{T}}(t|t+i) \tag{3.7.159}$$

【定理 3.7.10】 带相关噪声时变系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下,有最优固定区间 Kalman 平滑器

$$\hat{\boldsymbol{x}}(t|N)=\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{P}(t|t-1)\boldsymbol{r}(t|N) \tag{3.7.160}$$

其中 $t=1,2,\cdots,N$,且有反向递推公式

$$\boldsymbol{r}(t|N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\boldsymbol{r}(t+1|N)+\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{Q}_{\varepsilon}^{-1}(t)\boldsymbol{\varepsilon}(t) \tag{3.7.161}$$

带 $t=N,\cdots,1$,且规定 $\boldsymbol{r}(N+1|N)=\boldsymbol{0}$.估值误差 $\tilde{\boldsymbol{x}}(t|N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|N)$ 方差阵 $\boldsymbol{P}(t|N)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|N)]$ 为

$$\boldsymbol{P}(t|N)=\boldsymbol{P}(t|t-1)-\boldsymbol{P}(t|t-1)\boldsymbol{S}(t|N)\boldsymbol{P}(t|t-1) \tag{3.7.162}$$

带 $t=N,\cdots,1$，且有反向递推公式

$$\boldsymbol{S}(t|N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\boldsymbol{S}(t+1|N)\boldsymbol{\Psi}_p(t)+\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{Q}_\varepsilon^{-1}(t)\boldsymbol{H}(t) \tag{3.7.163}$$

带 $t=N,\cdots,1$，且规定 $\boldsymbol{S}(N+1|N)=\boldsymbol{0}$. 定理中 $\hat{\boldsymbol{x}}(t|t-1)$，$\boldsymbol{P}(t|t-1)$，$\boldsymbol{\varepsilon}(t)$，$\boldsymbol{\Psi}_p(t)$ 由定理 3.7.3 和定理 3.7.4 计算.

证明　完全类似于定理 3.4.3 的推导，从略. □

【定理 3.7.11】　对带相关噪声时变系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下，上述定理 3.7.9 和定理 3.7.10 仍成立，但其中 $\boldsymbol{\Psi}_p(t)$ 不用(3.7.148)计算，而用(3.7.13)计算，即

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t) \tag{3.7.164}$$

其中 $\boldsymbol{K}_p(t)$ 由(3.7.21)计算为

$$\boldsymbol{K}_p(t)=[\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)]\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.7.165}$$

而 $\boldsymbol{P}(t|t-1)$ 由(3.7.24)计算. 就是说，用定理 3.7.1 计算 $\hat{\boldsymbol{x}}(t|t-1)$，$\boldsymbol{P}(t|t-1)$，$\boldsymbol{K}_p(t)$，$\boldsymbol{\Psi}_p(t)$，$\boldsymbol{\varepsilon}(t)$. 由定理 3.7.6 知上述两种算法是等价(等值)的.

证明　应用一步预报误差(3.7.12)

$$\tilde{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{K}_p(t)\boldsymbol{v}(t) \tag{3.7.166}$$

其中 $\boldsymbol{\Psi}_p(t)$ 和 $\boldsymbol{K}_p(t)$ 由(3.7.164)和(3.7.165)计算，应用定理 3.7.1，平行于定理 3.7.9 和定理 3.7.10 的推导得定理 3.7.11. □

3.7.4　最优白噪声估值器

下面将定理 3.5.2 和定理 3.5.3 给出的白噪声估值器推广到带相关噪声的时变系统.[13]

【定理 3.7.12】[13]　带相关噪声的时变系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下，有统一的最优白噪声估值器 $\hat{\boldsymbol{w}}(t|t+N)$ 和 $\hat{\boldsymbol{v}}(t|t+N)$，统一记为 $\hat{\boldsymbol{\theta}}(t|t+N)$，$\boldsymbol{\theta}=\boldsymbol{w},\boldsymbol{v}$，

$$\hat{\boldsymbol{\theta}}(t|t+N)=\boldsymbol{0}\quad(N<0) \tag{3.7.167}$$

$$\hat{\boldsymbol{\theta}}(t|t+N)=\hat{\boldsymbol{\theta}}(t|t+N-1)+\boldsymbol{M}_\theta(t|t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.7.168}$$

其中 $\boldsymbol{\theta}=\boldsymbol{w},\boldsymbol{v}$，$N=0.1,2,\cdots$，且定义

$$\boldsymbol{M}_w(t|t)=\boldsymbol{S}(t)\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.7.169}$$

$$\boldsymbol{M}_v(t|t)=\boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.7.170}$$

$$\boldsymbol{M}_\theta(t|t+1)=\boldsymbol{D}_\theta(t)\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_\varepsilon^{-1}(t+1) \tag{3.7.171}$$

$$\boldsymbol{D}_w(t)=-\boldsymbol{S}(t)\boldsymbol{K}_f^{\mathrm{T}}(t)\overline{\boldsymbol{\Phi}}^{\mathrm{T}}(t)+\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{S}(t)\boldsymbol{J}^{\mathrm{T}}(t) \tag{3.7.172}$$

$$\boldsymbol{D}_v(t)=-\boldsymbol{Q}_v(t)\boldsymbol{K}_f^{\mathrm{T}}(t)\overline{\boldsymbol{\Phi}}^{\mathrm{T}}(t) \tag{3.7.173}$$

$$\boldsymbol{M}_\theta(t|t+N)=\boldsymbol{D}_\theta(t)\left\{\prod_{i=1}^{N-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\right\}\boldsymbol{H}^{\mathrm{T}}(t+N)\boldsymbol{Q}_\varepsilon^{-1}(t+N),\quad N>1 \tag{3.7.174}$$

$$\boldsymbol{\Psi}_p(t)=\overline{\boldsymbol{\Phi}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t)\boldsymbol{H}(t)] \tag{3.7.175}$$

其中 $\boldsymbol{Q}_\varepsilon(t+N)$，$\boldsymbol{K}_f(t)$，$\boldsymbol{\varepsilon}(t)$，$\boldsymbol{\Psi}_p(t)$ 由定理 3.7.7 或定理 3.7.3 和定理 3.7.4 计算，且定义

$$\overline{\boldsymbol{\Phi}}(t)=\boldsymbol{\Phi}(t)-\boldsymbol{J}(t)\boldsymbol{H}(t),\quad \boldsymbol{J}(t)=\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)\boldsymbol{Q}_v^{-1}(t) \tag{3.7.176}$$

估值误差方差阵 $\boldsymbol{P}_\theta(t|t+N)$ 为

$$\boldsymbol{P}_w(t|t+N)=\boldsymbol{Q}_w(t),\quad N<0 \tag{3.7.177}$$

$$\boldsymbol{P}_v(t|t+N)=\boldsymbol{Q}_v(t),\quad N<0 \tag{3.7.178}$$

$$\boldsymbol{P}_\theta(t|t+N)=\boldsymbol{P}_\theta(t|t+N-1)-\boldsymbol{M}_\theta(t|t+N)\boldsymbol{Q}_\varepsilon(t+N)\boldsymbol{M}_\theta^{\mathrm{T}}(t|t+N),\quad N\geqslant 0 \tag{3.7.179}$$

证明 置 $\boldsymbol{\theta}=\boldsymbol{v}$，因 $\boldsymbol{v}(t)\perp L(\boldsymbol{y}(t+N),\cdots,\boldsymbol{y}(1))$ $(N<0)$，由射影性质有

$$\hat{\boldsymbol{v}}(t|t+N)=\boldsymbol{0}\quad(N<0) \tag{3.7.180}$$

从而有 $\boldsymbol{P}_v(t|t+N)=\boldsymbol{Q}_v$ $(N<0)$. 由递推射影公式有

$$\hat{\boldsymbol{v}}(t|t)=\hat{\boldsymbol{v}}(t|t-1)+\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]\boldsymbol{Q}_\varepsilon^{-1}(t)\boldsymbol{\varepsilon}(t) \tag{3.7.181}$$

因 $\boldsymbol{v}(t)\perp\tilde{\boldsymbol{x}}(t|t-1)$，且注意(3.7.6)

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t) \tag{3.7.182}$$

可得

$$\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]=\boldsymbol{Q}_v(t) \tag{3.7.183}$$

这引出定理对 $\boldsymbol{\theta}=\boldsymbol{v}$ 和 $N=0$ 成立. 对 $N=1$，由射影公式有

$$\hat{\boldsymbol{v}}(t|t+1)=\hat{\boldsymbol{v}}(t|t)+\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)]\boldsymbol{Q}_\varepsilon^{-1}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.184}$$

注意

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{H}(t+1)\tilde{\boldsymbol{x}}(t+1|t)+\boldsymbol{v}(t+1) \tag{3.7.185}$$

而由(3.7.51)和(3.7.175)有

$$\tilde{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\tilde{\boldsymbol{x}}(t|t-1)-\bar{\boldsymbol{\Phi}}(t)\boldsymbol{K}_f(t)\boldsymbol{v}(t)+\bar{\boldsymbol{w}}(t) \tag{3.7.186}$$

因于 $\boldsymbol{v}(t)$ 与 $\bar{\boldsymbol{w}}(t)$ 是独立的，可得

$$\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)]=-\boldsymbol{Q}_v(t)\boldsymbol{K}_f^{\mathrm{T}}(t)\bar{\boldsymbol{\Phi}}^{\mathrm{T}}(t) \tag{3.7.187}$$

这引出(3.7.171)对于 $\boldsymbol{\theta}=\boldsymbol{v}$ 和 $N=1$ 成立. 由递推射影公式对一般情形有(3.7.168)对 $\boldsymbol{\theta}=\boldsymbol{v}$成立，且有

$$\boldsymbol{M}_v(t|t+N)=\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]\boldsymbol{Q}_\varepsilon^{-1}(t+N) \tag{3.7.188}$$

应用关系

$$\boldsymbol{\varepsilon}(t+N)=\boldsymbol{H}(t+N)\tilde{\boldsymbol{x}}(t+N|t+N-1)+\boldsymbol{v}(t+N) \tag{3.7.189}$$

及 $\tilde{\boldsymbol{x}}(t+N|t+N-1)$的表达式(3.7.154)，可得

$$\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]=-\boldsymbol{Q}_v(t)\boldsymbol{K}_f^{\mathrm{T}}(t)\bar{\boldsymbol{\Phi}}^{\mathrm{T}}(t)\boldsymbol{\Psi}^{\mathrm{T}}(t+N,t+1)\boldsymbol{H}^{\mathrm{T}}(t+N) \tag{3.7.190}$$

由定义(3.7.155)和上式得(3.7.174). 同理对 $\boldsymbol{\theta}=\boldsymbol{w}$，应用(3.7.186)和(3.7.154)，并注意定义(3.7.38)，$\bar{\boldsymbol{w}}(t)=\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{J}(t)\boldsymbol{v}(t)$，可证明定理对 $\boldsymbol{\theta}=\boldsymbol{w}$ 也成立. □

【推论 3.7.9】 在定理 3.7.11 条件下，有非递推白噪声估值器

$$\hat{\boldsymbol{w}}(t|t)=\boldsymbol{S}(t)\boldsymbol{Q}_\varepsilon^{-1}(t)\boldsymbol{\varepsilon}(t) \tag{3.7.191}$$

$$\hat{\boldsymbol{v}}(t|t)=\boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t)\boldsymbol{\varepsilon}(t) \tag{3.7.192}$$

$$\hat{\boldsymbol{\theta}}(t|t+N)=\sum_{i=0}^{N}\boldsymbol{M}_\theta(t|t+i)\boldsymbol{\varepsilon}(t+i),\quad N>0 \tag{3.7.193}$$

且有估值误差方差阵

$$\boldsymbol{P}_\theta(t\mid t+N)=\boldsymbol{Q}_\theta-\sum_{i=0}^{N}\boldsymbol{M}_\theta(t\mid t+i)\boldsymbol{Q}_\varepsilon(t+i)\boldsymbol{M}_\theta^{\mathrm{T}}(t\mid t+i) \tag{3.7.194}$$

下面给出不同于定理 3.7.11 的另一种白噪声估值器算法，两者不同之处在于在定理 3.7.12 中是基于定理 3.7.3 或定理 3.7.7 计算白噪声估值器. 而下面的定理是基于定理 3.7.1 来计算白噪声估值器，因而得出不同算法，但可证明它们在数值上是等价（等值）的.

【定理 3.7.13】 带相关噪声的时变系统(3.7.1)和(3.7.2)在假设 1 ~ 3 下，有统一的最优白噪声估值器

$$\hat{\boldsymbol{\theta}}(t\mid t+N)=\mathbf{0}\quad(N<0) \tag{3.7.195}$$

$$\hat{\boldsymbol{\theta}}(t\mid t+N)=\hat{\boldsymbol{\theta}}(t\mid t+N-1)+\boldsymbol{M}_\theta(t\mid t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.7.196}$$

其中 $\boldsymbol{\theta}=\boldsymbol{w},\boldsymbol{v}$，$N=1,2,\cdots$，且定义

$$\boldsymbol{M}_w(t\mid t)=\boldsymbol{S}(t)\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.7.197}$$

$$\boldsymbol{M}_v(t\mid t)=\boldsymbol{Q}_v(t)\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.7.198}$$

$$\boldsymbol{M}_\theta(t\mid t+1)=\boldsymbol{D}_\theta(t)\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_\varepsilon^{-1}(t+1) \tag{3.7.199}$$

$$\boldsymbol{D}_w(t)=\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{S}(t)\boldsymbol{K}_p^{\mathrm{T}}(t) \tag{3.7.200}$$

$$\boldsymbol{D}_v(t)=\boldsymbol{S}^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{Q}_v(t)\boldsymbol{K}_p^{\mathrm{T}}(t) \tag{3.7.201}$$

$$\boldsymbol{M}_\theta(t\mid t+N)=\boldsymbol{D}_\theta(t)\left\{\prod_{i=1}^{N-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\right\}\boldsymbol{H}^{\mathrm{T}}(t+N)\boldsymbol{Q}_\varepsilon^{-1}(t+N),\quad N>1 \tag{3.7.202}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t) \tag{3.7.203}$$

其中 $\boldsymbol{K}_p(t)$，$\boldsymbol{\varepsilon}(t)$由定理 3.7.1 计算.

证明 由定理 3.7.1，利用关系(3.7.166)及由它迭代引出的关系

$$\tilde{\boldsymbol{x}}(t+N\mid t+N-1)=\boldsymbol{\Psi}(t+N,t)\tilde{\boldsymbol{x}}(t\mid t-1)+\sum_{i=t+1}^{t+N}\boldsymbol{\Psi}(t+N,i)\times \\ [\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1)-\boldsymbol{K}_p(i-1)\boldsymbol{v}(i-1)] \tag{3.7.204}$$

得证，详细推导从略. □

【定理 3.7.14】 由定理 3.7.12 和定理 3.7.13 是等价的白噪声估计算法.

证明 只需证明由定理 3.7.12 计算的 $\boldsymbol{M}_\theta(t\mid t+N)$等于由定理 3.7.13 计算的 $\boldsymbol{M}_\theta(t\mid t+N)$. 由定理 3.7.6 我们已证明由(3.7.175)和(3.7.203)给出的 $\boldsymbol{\Psi}_p(t)$是等值的. 因此剩下的问题只需证明由定理 3.7.12 和定理 3.7.13 计算的 $\boldsymbol{D}_\theta(t)$（$\boldsymbol{\theta}=w,v$）是相等的. 由定理 3.7.12 的(3.7.172)和 $\boldsymbol{J}(t)$的定义(3.7.41)有

$$\begin{aligned}\boldsymbol{D}_w(t)=&-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{H}\boldsymbol{P}(\boldsymbol{\Phi}^{\mathrm{T}}-\boldsymbol{H}^{\mathrm{T}}\boldsymbol{J}^{\mathrm{T}})+\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{J}^{\mathrm{T}}=\\ &-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{H}\boldsymbol{P}\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}+\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{J}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{J}^{\mathrm{T}}=\\ &-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{H}\boldsymbol{P}\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}+\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}(\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v)\boldsymbol{J}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{Q}_v\boldsymbol{J}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{J}^{\mathrm{T}}=\\ &-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{H}\boldsymbol{P}\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}+\boldsymbol{S}\boldsymbol{J}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{Q}_v\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{J}^{\mathrm{T}}=\\ &-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{H}\boldsymbol{P}\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\end{aligned} \tag{3.7.205}$$

其中为推导简单计，省略了上式右端的时标 t.

另一方面，由定理 3.7.13 的(3.7.200)和(3.7.21)有

$$\boldsymbol{D}_w(t)=\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}(\boldsymbol{HP\Phi}^{\mathrm{T}}+\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}})=\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{HP\Phi}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.7.206}$$

上两式证明了两种算法给出相同的 $\boldsymbol{D}_w(t)$.

最后由定理 3.7.12 的(3.7.173)计算的 $\boldsymbol{D}_v(t)$为

$$\begin{aligned}\boldsymbol{D}_v(t)=&-\boldsymbol{D}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{HP}(\boldsymbol{\Phi}^{\mathrm{T}}-\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}})=\\&-\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{HP\Phi}^{\mathrm{T}}+\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{HPH}^{\mathrm{T}}\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}=\\&-\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{HP\Phi}^{\mathrm{T}}+\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}(\boldsymbol{HPH}^{\mathrm{T}}+\boldsymbol{Q}_v)\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}})-\\&\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{Q}_v\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}=\\&-\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{HP\Phi}^{\mathrm{T}}+\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\end{aligned} \tag{3.7.207}$$

其中应用了事实 $\boldsymbol{Q}_\varepsilon=\boldsymbol{HPH}^{\mathrm{T}}+\boldsymbol{Q}_v$，$\boldsymbol{J}=\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_v^{-1}$ 和 $\bar{\boldsymbol{\Phi}}=\boldsymbol{\Phi}-\boldsymbol{HJ}$. 为了简单计，上式右端的时标 t 被省略. 另一方面，由定理 3.7.13 的(3.7.201)和 $\boldsymbol{K}_p(t)$的表达式(3.7.21)有

$$\boldsymbol{D}_v(t)=\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}(\boldsymbol{HP\Phi}^{\mathrm{T}}+\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}})=-\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{HP\Phi}^{\mathrm{T}}+\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.7.208}$$

因此(3.7.207)和(3.7.208)是相同的. 证毕. □

【定理 3.7.15】[13] 带相关噪声系统(3.7.1)和(3.7.2)在假设 1～3 下，有统一的固定区间最优白噪声估值器 $\hat{\boldsymbol{w}}(t|N)$和 $\hat{\boldsymbol{v}}(t|N)$，统一记为 $\hat{\boldsymbol{\theta}}(t|N)$ $(\boldsymbol{\theta}=\boldsymbol{w},\boldsymbol{v})$为

$$\hat{\boldsymbol{\theta}}(t|N)=\hat{\boldsymbol{\theta}}(t|t)+\boldsymbol{D}_\theta(t)\boldsymbol{r}(t+1|N) \tag{3.7.209}$$

其中 $\boldsymbol{\theta}=w,v,t=N-1,\cdots,1$，且有反向递推式

$$\boldsymbol{r}(t+1|N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)\boldsymbol{r}(t+2|N)+\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_\varepsilon^{-1}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.7.210}$$

带初值 $\boldsymbol{r}(N+1|N)=\boldsymbol{0}$，且 $\boldsymbol{\Psi}_p^{\mathrm{T}}(t)$为

$$\boldsymbol{\Psi}_p^{\mathrm{T}}(t)=\bar{\boldsymbol{\Phi}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t)\boldsymbol{H}(t)] \tag{3.7.211}$$

其中 $\boldsymbol{K}_f(t)$，$\boldsymbol{Q}_\varepsilon(t)$，$\boldsymbol{\varepsilon}(t)$由定理 3.7.7 或定理 3.7.3 和定理 3.7.4 计算. 平滑误差方差阵为

$$\boldsymbol{P}_\theta(t|N)=\boldsymbol{P}_\theta(t|t)-\boldsymbol{D}_\theta(t)\boldsymbol{S}(t+1|N)\boldsymbol{D}_\theta^{\mathrm{T}}(t) \tag{3.7.212}$$

其中 $t=N-1,\cdots,1$，且有反向递推式

$$\boldsymbol{S}(t+1|N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)\boldsymbol{S}(t+2|N)\boldsymbol{\Psi}_p(t+1)+\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_\varepsilon^{-1}(t+1)\boldsymbol{H}(t+1) \tag{3.7.213}$$

带初值 $\boldsymbol{S}(N+1|N)=\boldsymbol{0}$. 初值 $\hat{\boldsymbol{\theta}}(t|t)$由(3.7.191)和(3.7.192)计算.

另一种等价的算法是

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t) \tag{3.7.214}$$

$$\boldsymbol{K}_p(t)=[\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)]\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.7.215}$$

其中 $\boldsymbol{P}(t|t-1)$，$\boldsymbol{\varepsilon}(t)$由定理 3.7.1 计算.

证明 应用定理 3.7.12 或推论 3.7.9 有

$$\hat{\boldsymbol{\theta}}(t|N)=\hat{\boldsymbol{\theta}}(t|t)+\sum_{i=1}^{N-t}\boldsymbol{M}_\theta^{\mathrm{T}}(t|t+i)\boldsymbol{\varepsilon}(t+i) \tag{3.7.216}$$

应用(3.7.174)$\boldsymbol{M}_\theta^{\mathrm{T}}(t|t+i)$的表达式可将上式表为(3.7.209)，其中定义

$$\boldsymbol{r}(t+1|N)=\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\boldsymbol{\varepsilon}(t+1)+$$
$$\sum_{i=2}^{N-t}\left\{\prod_{j=1}^{i-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+j)\right\}\boldsymbol{H}^{\mathrm{T}}(t+i)\boldsymbol{Q}_{\varepsilon}^{-1}(t+i)\boldsymbol{\varepsilon}(t+i)\tag{3.7.217}$$

注意

$$\sum_{i=2}^{N-t}\left\{\prod_{j=1}^{i-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+j)\right\}\boldsymbol{H}^{\mathrm{T}}(t+i)\boldsymbol{Q}_{\varepsilon}^{-1}(t+i)\boldsymbol{\varepsilon}(t+i)=$$
$$\boldsymbol{\Psi}_p(t+1)\boldsymbol{H}^{\mathrm{T}}(t+2)\boldsymbol{Q}_{\varepsilon}^{-1}(t+2)\boldsymbol{\varepsilon}(t+2)+$$
$$\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)\sum_{i=3}^{N-t}\left\{\prod_{j=2}^{i-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+j)\right\}\boldsymbol{H}^{\mathrm{T}}(t+i)\boldsymbol{Q}_{\varepsilon}^{-1}(t+i)\boldsymbol{\varepsilon}(t+i)=$$

$$\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)\left\{\boldsymbol{H}^{\mathrm{T}}(t+2)\boldsymbol{Q}_{\varepsilon}^{-1}(t+2)\boldsymbol{\varepsilon}(t+2)+\sum_{i=2}^{N-t-1}\left[\prod_{j=1}^{i-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1+j)\right]\times\right.$$

$$\left.\boldsymbol{H}^{\mathrm{T}}(t+1+i)\boldsymbol{Q}_{\varepsilon}^{-1}(t+1+i)\boldsymbol{\varepsilon}(t+1+i)\right\}=\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)\boldsymbol{r}(t+2|N)\tag{3.7.218}$$

这证明(3.7.210)成立.在(3.7.209)中置 $t=N$ 引出 $\boldsymbol{r}(N+1|N)=\mathbf{0}$.由(3.7.209)有误差关系

$$\tilde{\boldsymbol{\theta}}(t|t)=\tilde{\boldsymbol{\theta}}(t|N)+\boldsymbol{D}_{\theta}(t)\boldsymbol{r}(t+1|N)\tag{3.7.219}$$

其中 $\tilde{\boldsymbol{\theta}}(t|j)=\boldsymbol{\theta}(t)-\tilde{\boldsymbol{\theta}}(t|j)$.因 $\tilde{\boldsymbol{\theta}}(t|N)\perp L(\boldsymbol{y}(N),\cdots,\boldsymbol{y}(0))$,故有 $\tilde{\boldsymbol{\theta}}(t|N)\perp\boldsymbol{r}(t+1|N)$.这引出(3.7.212),其中定义

$$\boldsymbol{S}(t+1|N)=\mathrm{E}[\boldsymbol{r}(t+1|N)\boldsymbol{r}^{\mathrm{T}}(t+1|N)]\tag{3.7.220}$$

将(3.7.210)代入上式,注意 $\boldsymbol{\varepsilon}(t+1)\perp\boldsymbol{r}(t+2|N)$,可得(3.7.213),由上式和 $\boldsymbol{r}(N+1|N)=\mathbf{0}$引出 $\boldsymbol{S}(N+1|N)=\mathbf{0}$.证毕. □

3.8 带相关噪声定常系统稳态 Kalman 滤波和稳态白噪声估值器

考虑带相关噪声的定常(时不变)系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{Bu}(t)+\boldsymbol{\Gamma w}(t)\tag{3.8.1}$$
$$\boldsymbol{y}(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t)\tag{3.8.2}$$

其中状态 $\boldsymbol{x}(t)\in R^n$,观测 $\boldsymbol{y}(t)\in R^m$,控制 $\boldsymbol{u}(t)\in R^p$,$\boldsymbol{\Phi},\boldsymbol{B},\boldsymbol{\Gamma},\boldsymbol{H}$ 是已知适当维数矩阵.

【假设 1】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}(t)\in R^m$ 是带零均值的相关白噪声:

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\\boldsymbol{v}(t)\end{bmatrix}[\boldsymbol{w}^{\mathrm{T}}(j)\ \boldsymbol{v}^{\mathrm{T}}(j)]\right\}=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\\boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}\delta_{tj}\tag{3.8.3}$$

【假设 2】 控制 $\boldsymbol{u}(t)$是已知的时间序列,或是$(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$的线性函数(反馈控制).

【假设 3】[2] $(\bar{\boldsymbol{\Phi}},\boldsymbol{H})$为完全可观对,$(\bar{\boldsymbol{\Phi}},\boldsymbol{\Gamma}\bar{\boldsymbol{Q}}_w)$为完全能稳对,其中 $\bar{\boldsymbol{\Phi}}=\boldsymbol{\Phi}-\boldsymbol{JH}$,$\boldsymbol{J}=\boldsymbol{\Gamma S Q}_v^{-1}$,$\boldsymbol{Q}_w-\boldsymbol{SQ}_v^{-1}\boldsymbol{S}^{\mathrm{T}}=\bar{\boldsymbol{Q}}_w\bar{\boldsymbol{Q}}_w^{\mathrm{T}}$,或假设 $\bar{\boldsymbol{\Phi}}$ 是稳定的.

相应于定理 3.7.1 有如下稳态 Kalman 预报器.

【定理 3.8.1】 带相关噪声定常系统(3.8.1)和(3.8.2)在假设 1 ~ 3 下,有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}\boldsymbol{u}(t)+\boldsymbol{K}_p\boldsymbol{\varepsilon}(t) \tag{3.8.4}$$

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)-\boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1) \tag{3.8.5}$$

或在 Wiener 滤波器形式下的 Kalman 预报器为

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}\boldsymbol{u}(t)+\boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.8.6}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H} \tag{3.8.7}$$

稳态 Kalman 预报器增益 $\boldsymbol{K}_p$ 为

$$\boldsymbol{K}_p=(\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S})(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v)^{-1} \tag{3.8.8}$$

其中稳态 Kalman 预报器误差方差阵 $\boldsymbol{\Sigma}$ 是如下稳态 Riccati 方程的解,

$$\boldsymbol{\Sigma}=\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{\Phi}^{\mathrm{T}}-(\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S})(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v)^{-1}(\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S})+\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.8.9}$$

且稳态 Kalman 预报器(3.8.6)是渐近稳定的,即 $\boldsymbol{\Psi}_p$ 是一个稳定矩阵($\boldsymbol{\Psi}_p$ 的所有特征值位于单位圆内).

【定理 3.8.2】 在定理 3.8.1 条件下,稳态 Kalman 滤波器为

$$\hat{\boldsymbol{x}}(t+1|t+1)=\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}_f\boldsymbol{\varepsilon}(t+1) \tag{3.8.10}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H}\hat{\boldsymbol{x}}(t+1|t) \tag{3.8.11}$$

$$\boldsymbol{K}_f=\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v]^{-1} \tag{3.8.12}$$

稳态滤波误差方差阵 $\boldsymbol{P}$ 为

$$\boldsymbol{P}=[\boldsymbol{I}_n-\boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Sigma} \tag{3.8.13}$$

其中 $\hat{\boldsymbol{x}}(t+1|t)$和 $\boldsymbol{\Sigma}$ 由定理 3.8.1 计算.

【注 3.8.1】 Riccati 方程(3.8.9)可由非稳态 Riccati 方程(3.7.24)带任意初值 $\boldsymbol{P}(1|0)=a\boldsymbol{I}_n$, $a>0$,用迭代法求解.由关系

$$\lim_{t\to\infty}\boldsymbol{P}(t+1|t)=\boldsymbol{\Sigma} \tag{3.8.14}$$

当取 t 充分大,我们有近似解 $\boldsymbol{\Sigma}\approx\boldsymbol{P}(t+1|t)$.通常迭代次数 $t=50$ 左右即可达到满意的精度.

相应于定理 3.7.3、定理 3.7.4 和定理 3.7.7 有如下稳态 Kalman 预报器和滤波器.

【定理 3.8.3】 带相关噪声定常系统(3.8.1)和(3.8.2)在假设 1 ~ 3 下,有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{B}\boldsymbol{u}(t)+\boldsymbol{K}_p\boldsymbol{y}(t) \tag{3.8.15}$$

$$\boldsymbol{\Psi}_p=\bar{\boldsymbol{\Phi}}-\bar{\boldsymbol{K}}_p\boldsymbol{H},\quad \bar{\boldsymbol{K}}_p=\bar{\boldsymbol{\Phi}}\boldsymbol{K}_f \tag{3.8.16}$$

$$\boldsymbol{K}_p=\bar{\boldsymbol{K}}_p+\boldsymbol{J} \tag{3.8.17}$$

$$\bar{\boldsymbol{\Phi}}=\boldsymbol{\Phi}-\boldsymbol{J}\boldsymbol{H},\quad \boldsymbol{J}=\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_v^{-1} \tag{3.8.18}$$

$$\boldsymbol{K}_f=\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v]^{-1} \tag{3.8.19}$$

稳态预报误差方差阵 $\boldsymbol{\Sigma}$ 满足稳 Riccati 方程

$$\boldsymbol{\Sigma}=\bar{\boldsymbol{\Phi}}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}]\bar{\boldsymbol{\Phi}}+\boldsymbol{\Gamma}(\boldsymbol{Q}_w-\boldsymbol{S}\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.8.20}$$

它可由(3.7.73)用迭代法求解.稳态 Kalman 滤波器为

$$\hat{\boldsymbol{x}}(t+1|t+1)=\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}_f\boldsymbol{\varepsilon}(t+1) \tag{3.8.21}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H}\hat{\boldsymbol{x}}(t+1|t) \tag{3.8.22}$$

或由(3.7.53)和(3.7.61)有

$$\hat{\boldsymbol{x}}(t+1|t+1)=\boldsymbol{\Psi}_f\hat{\boldsymbol{x}}(t|t)+\boldsymbol{C}_1\boldsymbol{u}(t)+\boldsymbol{C}_2\boldsymbol{y}(t)+\boldsymbol{K}_f\boldsymbol{y}(t+1) \tag{3.8.23}$$

$$\boldsymbol{C}_1=[\boldsymbol{I}_n-\boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{B},\quad \boldsymbol{C}_2=[\boldsymbol{I}_n-\boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{J},\quad \boldsymbol{\Psi}_f=[\boldsymbol{I}_n-\boldsymbol{K}_f\boldsymbol{H}]\overline{\boldsymbol{\Phi}} \tag{3.8.24}$$

稳态滤波误差方差阵 $\boldsymbol{P}$ 为

$$\boldsymbol{P}=[\boldsymbol{I}_n-\boldsymbol{K}\boldsymbol{H}]\boldsymbol{\Sigma} \tag{3.8.25}$$

且(3.8.15)是渐近稳定的,即 $\boldsymbol{\Psi}_p$ 是一个稳定矩阵.[2]

相应于定理3.7.8有如下定理.

【定理3.8.4】 对于带相关噪声定常系统(3.8.1)和(3.8.2),设 $\boldsymbol{u}(t)=\boldsymbol{0}$,在假设1和假设3下,有一步稳态Kalman预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\boldsymbol{y}(t) \tag{3.8.26}$$

其中 $\boldsymbol{\Sigma}$, $\boldsymbol{\Psi}_p$ 和 $\boldsymbol{K}_p$ 由定理3.8.1或定理3.8.3计算.超前 N 步稳态Kalman预报器为

$$\hat{\boldsymbol{x}}(t+N|t)=\boldsymbol{\Phi}^{N-1}\hat{\boldsymbol{x}}(t+1|t) \tag{3.8.27}$$

稳态预报误差方差阵 $\boldsymbol{P}_N=\mathrm{E}[\tilde{\boldsymbol{x}}(t+N|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+N|t)]$, $\tilde{\boldsymbol{x}}(t+N|t)=\boldsymbol{x}(t+N)-\hat{\boldsymbol{x}}(t+N|t)$,为

$$\boldsymbol{P}_N=\boldsymbol{\Phi}^{N-1}\boldsymbol{\Sigma}(\boldsymbol{\Phi}^{N-1})^{\mathrm{T}}+\sum_{j=2}^{N}\boldsymbol{\Phi}^{N-j}\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^{N-j})^{\mathrm{T}},\quad N\geqslant 2 \tag{3.8.28}$$

证明 由(3.7.140)有

$$\boldsymbol{P}_N=\boldsymbol{\Phi}^{N-1}\boldsymbol{\Sigma}(\boldsymbol{\Phi}^{N-1})^{\mathrm{T}}+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}^{t+N-i}\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^{t+N-i})^{\mathrm{T}} \tag{3.8.29}$$

置变换 $j=i-t$,则由(3.8.29)立刻得(3.8.28). □

相应于定理3.7.9有如下稳态Kalman平滑器.

【定理3.8.5】 带相关噪声定常系统(3.8.1)和(3.8.2)在假设1~3下,有稳态Kalman滤波器和平滑器($N\geqslant 0$)为

$$\hat{\boldsymbol{x}}(t|t+N)=\hat{\boldsymbol{x}}(t|t+N-1)+\boldsymbol{K}_N\boldsymbol{\varepsilon}(t+N),\quad N\geqslant 0 \tag{3.8.30}$$

带初值 $\hat{\boldsymbol{x}}(t|t-1)$, $N=0,1,2,\cdots$,且规定 $\boldsymbol{K}_0=\boldsymbol{K}_f$,而 $\hat{\boldsymbol{x}}(t|t-1)$, $\boldsymbol{K}_f$, $\boldsymbol{\Psi}_p$, $\boldsymbol{\Sigma}$, $\boldsymbol{\varepsilon}(t+N)$ 由定理3.8.1或定理3.8.3计算.稳态平滑增益阵 $\boldsymbol{K}_N$ 为

$$\boldsymbol{K}_N=\boldsymbol{\Sigma}(\boldsymbol{\Psi}_p^{\mathrm{T}})^N\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1} \tag{3.8.31}$$

其中稳态新息方差 $\boldsymbol{Q}_\varepsilon$ 为

$$\boldsymbol{Q}_\varepsilon=\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v \tag{3.8.32}$$

其中由定理3.8.1

$$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H},\quad \boldsymbol{K}_p=(\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S})\boldsymbol{Q}_\varepsilon^{-1} \tag{3.8.33}$$

或由定理3.8.3

$$\boldsymbol{\Psi}_p=\overline{\boldsymbol{\Phi}}(\boldsymbol{I}_n-\boldsymbol{K}_f\boldsymbol{H}),\quad \boldsymbol{K}_f=\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1} \tag{3.8.34}$$

在(3.8.33)中 $\boldsymbol{\Sigma}$ 用Riccati方程(3.8.9)求解.在(3.8.34)中 $\boldsymbol{\Sigma}$ 由Riccati方程(3.8.20)求解.稳态平滑误差方差阵 $\boldsymbol{P}_N$ 为

$$\boldsymbol{P}_N=\boldsymbol{P}_{N-1}-\boldsymbol{K}_N\boldsymbol{Q}_\varepsilon\boldsymbol{K}_N^{\mathrm{T}} \tag{3.8.35}$$

非递推稳态Kalman平滑器为

$$\hat{\boldsymbol{x}}(t \mid t+N)=\hat{\boldsymbol{x}}(t \mid t-1)+\sum_{i=0}^{N} \boldsymbol{K}_i \boldsymbol{\varepsilon}(t+i), \quad N \geqslant 0 \tag{3.8.36}$$

具有稳态平滑误差方差阵

$$\boldsymbol{P}_N=\boldsymbol{\Sigma}-\sum_{i=0}^{N} \boldsymbol{K}_i \boldsymbol{Q}_{\varepsilon} \boldsymbol{K}_i^{\mathrm{T}} \tag{3.8.37}$$

相应于定理 3.7.12 应用定理 3.7.3 和定理 3.7.4 有稳态白噪声估值器如下.

【定理 3.8.6】 带相关噪声的定常系统(3.8.1)和(3.8.2)在假设 1 ~ 3 下,有统一的稳态白噪声估值器 $\hat{\boldsymbol{w}}(t \mid t+N)$和 $\hat{\boldsymbol{v}}(t \mid t+N)$,统一记 $\hat{\boldsymbol{\theta}}(t \mid t+N)$ $(\boldsymbol{\theta}=\boldsymbol{w}, \boldsymbol{v})$为

$$\hat{\boldsymbol{\theta}}(t \mid t+N)=\mathbf{0} \quad (N<0) \tag{3.8.38}$$

$$\hat{\boldsymbol{\theta}}(t \mid t+N)=\hat{\boldsymbol{\theta}}(t \mid t+N-1)+\boldsymbol{M}_\theta(N) \boldsymbol{\varepsilon}(t+N) \tag{3.8.39}$$

$$\hat{\boldsymbol{\theta}}(t \mid t+N)=\sum_{i=0}^{N} \boldsymbol{M}_\theta(i) \boldsymbol{\varepsilon}(t+i) \tag{3.8.40}$$

其中 $\boldsymbol{\theta}=\boldsymbol{w}, \boldsymbol{v}, N=0,1,\cdots$,且定义

$$\boldsymbol{M}_w(0)=\boldsymbol{S} \boldsymbol{Q}_{\varepsilon}^{-1}, \quad \boldsymbol{M}_v(0)=\boldsymbol{Q}_v \boldsymbol{Q}_{\varepsilon}^{-1} \tag{3.8.41}$$

$$\boldsymbol{M}_\theta(1)=\boldsymbol{D}_\theta \boldsymbol{H}^{\mathrm{T}} \boldsymbol{Q}_{\varepsilon}^{-1} \tag{3.8.42}$$

$$\boldsymbol{D}_w=-\boldsymbol{S} \boldsymbol{K}_f^{\mathrm{T}} \overline{\boldsymbol{\Phi}}^{\mathrm{T}}+\boldsymbol{Q}_w \boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{S} \boldsymbol{J}^{\mathrm{T}} \tag{3.8.43}$$

$$\boldsymbol{D}_v=-\boldsymbol{Q}_v \boldsymbol{K}_f^{\mathrm{T}} \overline{\boldsymbol{\Phi}}^{\mathrm{T}} \tag{3.8.44}$$

$$\boldsymbol{M}_\theta(N)=\boldsymbol{D}_\theta\left(\boldsymbol{\Psi}_p^{\mathrm{T}}\right)^{N-1} \boldsymbol{H}^{\mathrm{T}} \boldsymbol{Q}_{\varepsilon}^{-1}, \quad N>1 \tag{3.8.45}$$

$$\boldsymbol{\Psi}_p=\overline{\boldsymbol{\Phi}}\left[\boldsymbol{I}_n-\boldsymbol{K}_f \boldsymbol{H}\right] \tag{3.8.46}$$

$$\overline{\boldsymbol{\Phi}}=\boldsymbol{\Phi}-\boldsymbol{J} \boldsymbol{H}, \quad \boldsymbol{J}=\boldsymbol{\Gamma} \boldsymbol{S} \boldsymbol{Q}_v^{-1} \tag{3.8.47}$$

$$\boldsymbol{K}_f=\boldsymbol{\Sigma} \boldsymbol{H}^{\mathrm{T}} \boldsymbol{Q}_{\varepsilon}^{-1}, \quad \boldsymbol{Q}_{\varepsilon}=\boldsymbol{H} \boldsymbol{\Sigma} \boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v \tag{3.8.48}$$

其中 $\boldsymbol{\Sigma}$ 满足 Riccati 方程

$$\boldsymbol{\Sigma}=\overline{\boldsymbol{\Phi}}\left[\boldsymbol{\Sigma}-\boldsymbol{\Sigma} \boldsymbol{H}\left(\boldsymbol{H} \boldsymbol{\Sigma} \boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v\right)^{-1} \boldsymbol{H} \boldsymbol{\Sigma}\right] \overline{\boldsymbol{\Phi}}^{\mathrm{T}}+\boldsymbol{\Gamma}\left(\boldsymbol{Q}_w-\boldsymbol{S} \boldsymbol{Q}_{\varepsilon}^{-1} \boldsymbol{S}^{\mathrm{T}}\right) \boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.8.49}$$

它可由(3.7.73)用迭代法求解.而 $\boldsymbol{\varepsilon}(t)$由如下稳态 Kalman 预报器计算.

$$\hat{\boldsymbol{x}}(t+1 \mid t)=\boldsymbol{\Psi}_p \hat{\boldsymbol{x}}(t \mid t-1)+\boldsymbol{B} \boldsymbol{u}(t)+\boldsymbol{K}_p \boldsymbol{y}(t) \tag{3.8.50}$$

$$\boldsymbol{K}_p=\overline{\boldsymbol{\Phi}} \boldsymbol{K}_f+\boldsymbol{J} \tag{3.8.51}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H} \hat{\boldsymbol{x}}(t+1 \mid t) \tag{3.8.52}$$

相应的稳态估值器误差方差阵为

$$\boldsymbol{P}_w(N)=\boldsymbol{Q}_w, \quad \boldsymbol{P}_v(N)=\boldsymbol{Q}_v, \quad N<0 \tag{3.8.53}$$

$$\boldsymbol{P}_\theta(N)=\boldsymbol{P}_\theta(N-1)-\boldsymbol{M}_\theta(N) \boldsymbol{Q}_{\varepsilon} \boldsymbol{M}_\theta(N) \tag{3.8.54}$$

或

$$\boldsymbol{P}_\theta(N)=\boldsymbol{Q}_\theta-\sum_{i=0}^{N} \boldsymbol{M}_\theta(i) \boldsymbol{Q}_{\varepsilon} \boldsymbol{M}_\theta^{\mathrm{T}}(i), \quad N \geqslant 0 \tag{3.8.55}$$

相应于定理 3.7.13,应用定理 3.7.1 有如下另一种形式稳态白噪声估值器.

【定理 3.8.7】 带相关噪声系统(3.8.1)和(3.8.2)在假设 1 ~ 3 下,有统一的稳态白噪声估值器

$$\hat{\boldsymbol{\theta}}(t \mid t+N)=\mathbf{0} \quad (N<0) \tag{3.8.56}$$

$$\hat{\boldsymbol{\theta}}(t \mid t+N)=\hat{\boldsymbol{\theta}}(t \mid t+N-1)+\boldsymbol{M}_\theta(N) \boldsymbol{\varepsilon}(t+N) \tag{3.8.57}$$

或

$$\hat{\boldsymbol{\theta}}(t \mid t+N)=\sum_{i=0}^{N} \boldsymbol{M}_{\theta}(i) \boldsymbol{\varepsilon}(t+i) \tag{3.8.58}$$

其中 $\boldsymbol{\theta}=w, v, N=1,2,\cdots$，且定义

$$\boldsymbol{M}_{w}(0)=\boldsymbol{S} \boldsymbol{Q}_{\varepsilon}^{-1}, \quad \boldsymbol{M}_{v}(0)=\boldsymbol{Q}_{v} \boldsymbol{Q}_{\varepsilon}^{-1} \tag{3.8.59}$$

$$\boldsymbol{M}_{\theta}(N)=\boldsymbol{D}_{\theta}\left(\boldsymbol{\Psi}_{p}^{\mathrm{T}}\right)^{N-1} \boldsymbol{H}^{\mathrm{T}} \boldsymbol{Q}_{\varepsilon}^{-1}, \quad N \geqslant 1 \tag{3.8.60}$$

$$\boldsymbol{D}_{w}=\boldsymbol{Q}_{w} \boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{S} \boldsymbol{K}_{p}^{\mathrm{T}} \tag{3.8.61}$$

$$\boldsymbol{D}_{v}=\boldsymbol{S}^{\mathrm{T}} \boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{Q}_{v} \boldsymbol{K}_{p}^{\mathrm{T}} \tag{3.8.62}$$

$$\boldsymbol{\Psi}_{p}=\boldsymbol{\Phi}-\boldsymbol{K}_{p} \boldsymbol{H} \tag{3.8.63}$$

$$\boldsymbol{K}_{p}=\left[\boldsymbol{\Phi} \boldsymbol{\Sigma} \boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma} \boldsymbol{S}\right] \boldsymbol{Q}_{\varepsilon}^{-1}, \quad \boldsymbol{Q}_{\varepsilon}=\boldsymbol{H} \boldsymbol{\Sigma} \boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_{v} \tag{3.8.64}$$

其中 $\boldsymbol{\Sigma}$ 满足 Riccati 方程

$$\boldsymbol{\Sigma}=\boldsymbol{\Phi} \boldsymbol{\Sigma} \boldsymbol{\Phi}^{\mathrm{T}}-\left[\boldsymbol{\Phi} \boldsymbol{\Sigma} \boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma} \boldsymbol{S}\right]\left[\boldsymbol{H} \boldsymbol{\Sigma} \boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_{v}\right]^{-1}\left[\boldsymbol{\Phi} \boldsymbol{\Sigma} \boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma} \boldsymbol{S}\right]^{\mathrm{T}}+\boldsymbol{\Gamma} \boldsymbol{Q}_{w} \boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.8.65}$$

它可由(3.7.24)用迭代法求解：$\boldsymbol{\Sigma}=\lim\limits_{t \rightarrow \infty} \boldsymbol{P}(t+1 \mid t)$，因而当 t 充分大有 $\boldsymbol{\Sigma} \approx \boldsymbol{P}(t+1 \mid t)$. 新息 $\boldsymbol{\varepsilon}(t)$可由如下稳态 Kalman 预报器计算.

$$\hat{\boldsymbol{x}}(t+1 \mid t)=\boldsymbol{\Psi}_{p} \hat{\boldsymbol{x}}(t \mid t-1)+\boldsymbol{B} \boldsymbol{u}(t)+\boldsymbol{K}_{p} \boldsymbol{y}(t) \tag{3.8.66}$$

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)-\boldsymbol{H} \hat{\boldsymbol{x}}(t \mid t-1) \tag{3.8.67}$$

带任意初值 $\hat{\boldsymbol{x}}(1 \mid 0)$. 稳态白噪声估值器误差方差阵由(3.8.53)~(3.8.55)计算.

3.9 基于 Kalman 滤波的时域 Wiener 滤波器设计方法

所谓 Wiener 滤波器是指滤波器可表为以观测信号作为输入的传递函数形式. 经典 Wiener[12]滤波方法是频域方法. 本节介绍由作者[38]提出的基于经典 Kalman 滤波器的时域 Wiener 滤波器设计方法. 通过建立观测信号的 ARMA 新息模型，以 ARMA 新息模型作为桥梁，由稳态 Kalman 估值器可化为等价的 Wiener 估值器. 这种 Wiener 滤波器的优点是它避免了计算最优初值，它是渐近稳定的，且可化为分量的解耦 Wiener 滤波器. 这种 Wiener 滤波器是稳态滤波器，可统一处理滤波预报和平滑问题，是针对线性离散定常(时不变)系统设计的.

3.9.1 ARMA 新息模型

为简单计，本节考虑带相关噪声的定常系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi} \boldsymbol{x}(t)+\boldsymbol{\Gamma} \boldsymbol{w}(t) \tag{3.9.1}$$

$$\boldsymbol{y}(t)=\boldsymbol{H} \boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{3.9.2}$$

它不带控制项 $\boldsymbol{u}(t)$，即 $\boldsymbol{u}(t)=\mathbf{0}$. 假设该系统满足节 3.8 的假设 1 和假设 3，于是存在稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1 \mid t)=\boldsymbol{\Phi} \hat{\boldsymbol{x}}(t \mid t-1)+\boldsymbol{K}_{p} \boldsymbol{\varepsilon}(t) \tag{3.9.3}$$

$$\boldsymbol{y}(t)=\boldsymbol{H} \hat{\boldsymbol{x}}(t \mid t-1)+\boldsymbol{\varepsilon}(t) \tag{3.9.4}$$

其中稳态 Kalman 预报器增益 $\boldsymbol{K}_{p}$，由节 3.8 通过解稳态 Riccati 方程可用两种算法求 $\boldsymbol{K}_{p}$. 上两式用状态空间模型描写了新息 $\boldsymbol{\varepsilon}(t)$与观测 $y(t)$的关系，叫状态空间新息模型. 将(3.9.

4)代入(3.9.3)可得在 Wiener 滤波器形式下的稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\boldsymbol{y}(t) \tag{3.9.5}$$

或写在传递函数形式下有

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{y}(t) \tag{3.9.6}$$

其中 q^{-1}为单位滞后算子.因上式具有以观测 $\boldsymbol{y}(t)$作为输入、传递函数为$(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}$的表达式,因此是 Wiener 滤波器.易知上式中 $\boldsymbol{\Psi}_p$ 为

$$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H} \tag{3.9.7}$$

状态空间新息模型(3.9.3)和(3.9.4)可化为如下 ARMA 新息模型.

【定理 3.9.1】 带相关噪声的定常系统(3.9.1)和(3.9.2)在节 3.8 的假设 1 和假设 3 下,有 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\psi(q^{-1})\boldsymbol{\varepsilon}(t) \tag{3.9.8}$$

其中分别定义多项式 $\psi(q^{-1})$和多项式矩阵 $\boldsymbol{A}(q^{-1})$

$$\psi(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p) \tag{3.9.9}$$

$$\boldsymbol{A}(q^{-1})=\psi(q^{-1})\boldsymbol{I}_m-\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-\boldsymbol{\Psi}_p)\boldsymbol{K}_pq^{-1} \tag{3.9.10}$$

且 $\psi(q^{-1})$是稳定的多项式(即 $\psi(x)$的所有零点位于单位圆外).

证明 将(3.9.6)代入(3.9.4)有

$$\boldsymbol{y}(t)=\boldsymbol{H}(I_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{y}(t-1)+\boldsymbol{\varepsilon}(t) \tag{3.9.11}$$

将求逆公式$(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}=\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)/\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)$代入上式后,两边乘以$\det(I_n-q^{-1}\boldsymbol{\Psi}_p)$,整理后得(3.9.8)~(3.9.10).因为在稳态 Kalman 预报器(3.9.5)中 $\boldsymbol{\Psi}_p$是稳定矩阵[2],即 $\boldsymbol{\Psi}_p$ 的特征值位于单位圆内,故有 $\psi(q^{-1})$为稳定的多项式. □

【注 3.9.1】 注意 $\psi(q^{-1})$是标量多项式,记 $\psi(q^{-1})=1+\psi_1q^{-1}+\cdots+\psi_{n_\psi}q^{-n_\psi}$,而 $\varepsilon(t)$是 $m\times1$ 维向量,在式(3.9.8)中应定义 $\psi(q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{\varepsilon}(t)+\psi_1\boldsymbol{\varepsilon}(t-1)+\cdots+\psi_{n_\psi}\boldsymbol{\varepsilon}(t-n_\psi)$.即 $\psi(q^{-1}\boldsymbol{\varepsilon}(t)=\psi(q^{-1})\boldsymbol{I}_m\boldsymbol{\varepsilon}(t)$.

3.9.2 白噪声 Wiener 滤波器

由定理 3.8.6 或定理 3.8.7 的(3.8.40)有稳态白噪声估值器

$$\hat{\boldsymbol{\theta}}(t\mid t+N)=\sum_{i=0}^{N}\boldsymbol{M}_\theta(i)\boldsymbol{\varepsilon}(t+i),\quad N\geqslant0,\boldsymbol{\theta}=w,v \tag{3.9.12}$$

【定理 3.9.2】 带相关噪声定常系统(3.9.1)和(3.9.2)在节 3.8 的假设 1 和假设 3 下,有白噪声新息滤波器

$$\hat{\boldsymbol{\theta}}(t|t+N)=\boldsymbol{L}_N^\theta(q^{-1})\boldsymbol{\varepsilon}(t+N),\quad \boldsymbol{\theta}=w,v \tag{3.9.13}$$

它的输入为新息 $\boldsymbol{\varepsilon}(t+N)$,传递函数为

$$\boldsymbol{L}_N^\theta(q^{-1})=\sum_{i=0}^{N}\boldsymbol{M}_\theta(i)q^{i-N} \tag{3.9.14}$$

它是多项式矩阵.因为 $\hat{\boldsymbol{\theta}}(t|t+N)=\boldsymbol{0}(N<0)$,故可定义

$$\boldsymbol{L}_N^\theta(q^{-1})=\boldsymbol{0}\quad(N<0) \tag{3.9.15}$$

而 $\boldsymbol{M}_\theta(i)$由定理 3.8.6 或定理 3.8.7 计算.

稳态估值误差 $\tilde{\boldsymbol{\theta}}(t|t+N)=\boldsymbol{\theta}(t)-\hat{\boldsymbol{\theta}}(t|t+N)$的方差阵 $\boldsymbol{P}_\theta(N)$为(3.8.53),

$$\boldsymbol{P}_{\theta}(N) = \boldsymbol{Q}_{\theta} - \sum_{i=0}^{N} \boldsymbol{M}_{\theta}(i)\boldsymbol{Q}_{\varepsilon}\boldsymbol{M}_{\theta}^{\mathrm{T}}(i) \tag{3.9.16}$$

证明 由(3.9.12)有

$$\hat{\boldsymbol{\theta}}(t \mid t+N) = \sum_{i=0}^{N} \boldsymbol{M}_{\theta}(i) q^{i-N} \boldsymbol{\varepsilon}(t+N) \tag{3.9.17}$$

由定义(3.9.14)和上式得(3.9.13). □

【定理 3.9.3】 在定理 3.9.2 条件下,有白噪声 Wiener 滤波器

$$\hat{\boldsymbol{\theta}}(t \mid t+N) = \frac{\boldsymbol{L}_N^{\theta}(q^{-1})\boldsymbol{A}(q^{-1})}{\psi(q^{-1})}\boldsymbol{y}(t+N), \quad \boldsymbol{\theta} = w, v \tag{3.9.18}$$

或

$$\psi(q^{-1})\hat{\boldsymbol{\theta}}(t \mid t+N) = \boldsymbol{K}_N^{\theta}(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.19}$$

$$\boldsymbol{K}_N^{\theta}(q^{-1}) = \boldsymbol{L}_N^{\theta}(q^{-1})\boldsymbol{A}(q^{-1}) \tag{3.9.20}$$

且(3.9.18)或(3.9.19)是渐近稳定的,因而可任意取 $\hat{\boldsymbol{\theta}}(t \mid t+N)$ 的初值.

证明 由(3.9.8)有

$$\boldsymbol{\varepsilon}(t+N) = \frac{\boldsymbol{A}(q^{-1})}{\psi(q^{-1})}\boldsymbol{y}(t+N) \tag{3.9.21}$$

将其代入(3.9.13)得(3.9.18)和(3.9.19). 由 $\psi(q^{-1})$ 是稳定的多项式引出(3.9.18)或(3.9.19)的渐近稳定性. □

3.9.3 统一的 Wiener 状态滤波器

在定理 3.8.3~定理 3.8.5 基础上,有如下定理.

【定理 3.9.4】 带相关噪声的定常系统(3.9.1)和(3.9.2)在节 3.8 的假设 1 和假设 3 下,有统一的 Wiener 状态滤波器 $\hat{\boldsymbol{x}}(t \mid t+N)$ $(N=0, N>0$ 或 $N<0)$ 为

$$\boldsymbol{\Psi}_N(q^{-1})\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.22}$$

其中对 $N=0$ 定义

$$\boldsymbol{\Psi}_0(q^{-1}) = \boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_f,$$
$$\boldsymbol{K}_0(q^{-1}) = \boldsymbol{K}_f + [\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{J}q^{-1} \tag{3.9.23}$$

对 $N>0$ 定义

$$\boldsymbol{\Psi}_N(q^{-1}) = \psi(q^{-1}) \tag{3.9.24}$$

$$\boldsymbol{K}_N(q^{-1}) = \mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p q^{-1-N} + \boldsymbol{K}_N^{s}(q^{-1})\boldsymbol{A}(q^{-1}) \tag{3.9.25}$$

$$\boldsymbol{K}_N^{s}(q^{-1}) = \sum_{i=0}^{N} \boldsymbol{K}_i q^{i-N} \tag{3.9.26}$$

对 $N=-1$ 定义

$$\boldsymbol{\Psi}_{-1}(q^{-1}) = \boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p \tag{3.9.27}$$

$$\boldsymbol{K}_{-1}(q^{-1}) = \boldsymbol{K}_p \tag{3.9.28}$$

对 $N<-1$ 定义

$$\boldsymbol{\Psi}_N(q^{-1}) = \psi(q^{-1}) \tag{3.9.29}$$

$$\boldsymbol{K}_N(q^{-1}) = \boldsymbol{\Phi}^{-N-1}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p \tag{3.9.30}$$

其中 $\boldsymbol{\Psi}_p, \boldsymbol{K}_p$ 和 $\boldsymbol{\Psi}_f, \boldsymbol{K}_f$ 分别为稳态 Kalman 预报器和滤波器的转移阵和增益阵,$\boldsymbol{K}_i$ 为稳态

Kalman 平滑增益，详见定理 3.8.3 ~ 3.8.5. $\psi(q^{-1})$和 $\boldsymbol{A}(q^{-1})$的 ARMA 新息模型定义.

对 $N=-1$，一步稳态 Kalman 预报器误差方差阵为 $\boldsymbol{\Sigma}$. 对 $N=0$，稳态 Kalman 滤波器误差方差阵为 $\boldsymbol{P}$，且 $\boldsymbol{P}=[\boldsymbol{I}_n-\boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Sigma}$，它们已由定理 3.8.3 ~ 3.8.5 给出.

对 $N>0$，稳态平滑误差方差阵为

$$\boldsymbol{P}_N = \boldsymbol{\Sigma} - \sum_{i=0}^{N}\boldsymbol{K}_i\boldsymbol{Q}_\varepsilon\boldsymbol{K}_i^{\mathrm{T}} \tag{3.9.31}$$

对 $N<-1$，稳态 Kalman 预报误差方差阵为

$$\boldsymbol{P}_N = \boldsymbol{\Phi}^{-N-1}\boldsymbol{\Sigma}(\boldsymbol{\Phi}^{-N-1})^{\mathrm{T}} + \sum_{j=2}^{-N}\boldsymbol{\Phi}^{-N-j}\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^{-N-j})^{\mathrm{T}} \tag{3.9.32}$$

或等价地有

$$\boldsymbol{P}_N = \boldsymbol{\Phi}^{-N-1}\boldsymbol{\Sigma}(\boldsymbol{\Phi}^{-N-1})^{\mathrm{T}} + \sum_{i=0}^{-N-2}\boldsymbol{\Phi}^{i}\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Phi}^{i\mathrm{T}} \tag{3.9.33}$$

证明 由(3.8.15)和(3.8.23)对 $N=0$，$N=-1$ 定理显然成立. 对 $N>0$，(3.8.36)可写为

$$\hat{\boldsymbol{x}}(t|t+N) = \hat{\boldsymbol{x}}(t|t-1) + \boldsymbol{K}_N^s(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{3.9.34}$$

其中 $\boldsymbol{K}_N^s(q^{-1})$由(3.9.26)定义. 将(3.9.6)和(3.9.21)代入(3.9.34)得(3.9.24)和 3.9.25). 对 $N<-1$，由(3.9.1)和射影性质有

$$\begin{aligned}
&\hat{\boldsymbol{x}}(t|t-2) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t-1|t-2),\\
&\hat{\boldsymbol{x}}(t|t-3) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t-1|t-3) = \boldsymbol{\Phi}^2\hat{\boldsymbol{x}}(t-2|t-3),\\
&\qquad\vdots\\
&\hat{\boldsymbol{x}}(t|t+N) = \boldsymbol{\Phi}^{-N-1}\hat{\boldsymbol{x}}(t+N+1|t+N)
\end{aligned} \tag{3.9.35}$$

应用(3.9.6)有

$$\hat{\boldsymbol{x}}(t+N+1|t+N) = (\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{y}(t+N) \tag{3.9.36}$$

将上式代入(3.9.35)得(3.9.29)和(3.9.30). 而(3.9.31)和(3.9.32)见(3.8.35)和(3.8.28). □

3.9.4 状态分量解耦 Wiener 滤波器

注意，上述 Wiener 状态滤波器的一个重要应用是用于设计状态分量的解耦 Wiener 滤波器. 在许多应用问题中，人们只是感兴趣状态的一部分分量或某个分量的估计问题，特别对用增广状态方法引起的高维系统状态估计问题，就出现这种问题. 为了减小计算负担，人们只需计算感兴趣的状态分量 Wiener 滤波器即可，而不必计算其他状态分量的滤波器. 文献[18]用 z 变换和求解线性方程组解决了状态分量解耦 Wiener 滤波器设计问题，没有解决解耦 Wiener 预报器和平滑器设计问题. 文献[7]用现代时间序列分析方法统一解决了状态分量解耦 Wiener 预报器和平滑器设计问题. 下面提出了统一的基于经典 Kalman 方法的状态分量解耦 Wiener 滤波器，可统一处理滤波、预报和平滑问题.

【定理 3.9.5】 带相关噪声定常系统(3.9.1)和(3.9.2)在节 3.8 的假设 1 和假设 3 下，记状态 $\boldsymbol{x}(t)$的分量形式为 $\boldsymbol{x}(t)=[x_1(t),x_2(t),\cdots,x_n(t)]^{\mathrm{T}}$，则有状态分量 $x_i(t)$的解耦 Wiener 滤波器 $\hat{x}_i(t|t+N)$ $(N=0, N<0$ 或 $N>0)$为

$$\psi(q^{-1})\hat{x}_i(t|t+N) = \boldsymbol{K}_{Ni}(q^{-1})\boldsymbol{y}(t+N),\quad i=1,\cdots,n \tag{3.9.37}$$

其中对 $N>1$ 和 $N<-1$，$\boldsymbol{K}_N(q^{-1})$ 由定理 3.9.3 的(3.9.25)和(3.9.30)定义，且定义

$$\boldsymbol{K}_N(q^{-1})=\begin{bmatrix}\boldsymbol{K}_{N1}(q^{-1})\\ \vdots\\ \boldsymbol{K}_{Nn}(q^{-1})\end{bmatrix} \tag{3.9.38}$$

其中 $1\times m$ 行向量 $\boldsymbol{K}_{Ni}(q^{-1})$ 为 $1\times m$ 阵 $\boldsymbol{K}_N(q^{-1})$ 的第 i 行向量. 也可记为

$$\boldsymbol{K}_{Ni}(q^{-1})=\boldsymbol{e}_i\boldsymbol{K}_N(q^{-1}) \tag{3.9.39}$$

$$\boldsymbol{e}_i=[0\cdots 0\ 1\ 0\cdots 0] \tag{3.9.40}$$

其中定义 $1\times m$ 行向量 $\boldsymbol{e}_i$ 为第 i 个元素为 1，其余元素为零的向量.

对于 $N=0$，分量解耦 Wiener 状态滤波器为

$$\psi(q^{-1})\hat{x}_i(t\mid t)=\boldsymbol{K}_{0i}(q^{-1})\boldsymbol{y}(t),\quad i=1,\cdots,n \tag{3.9.41}$$

$$\boldsymbol{K}_{0i}(q^{-1})=\boldsymbol{e}_i\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_f)\boldsymbol{K}_0(q^{-1}) \tag{3.9.42}$$

其中 $\boldsymbol{K}_0(q^{-1})$ 由(3.9.23)定义.

对 $N=-1$，分量解耦 Wiener 状态预报器为

$$\boldsymbol{\psi}(q^{-1})\hat{x}_i(t\mid t-1)=\boldsymbol{K}_{-1i}(q^{-1})\boldsymbol{y}(t-1),\quad i=1,\cdots,n \tag{3.9.43}$$

$$\boldsymbol{K}_{-1i}(q^{-1})=\boldsymbol{e}_i\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p \tag{3.9.44}$$

证明 将定理(3.9.3)中对 $N>0$ 和 $N<-1$ 的(3.9.24)和(3.9.25)及(3.9.29)和(3.9.30)写成分量形式即得(3.9.37)～(3.9.42). 对 $N=0$，由(3.9.22)和(3.9.23)有稳态 Wiener 滤波器

$$\hat{\boldsymbol{x}}(t\mid t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_f)^{-1}\boldsymbol{K}_0(q^{-1})\boldsymbol{y}(t) \tag{3.9.45}$$

由(3.8.16)和(3.8.24)有

$$\boldsymbol{\Psi}_p=\overline{\boldsymbol{\Phi}}[\boldsymbol{I}_n-\boldsymbol{K}_f\boldsymbol{H}],\quad \boldsymbol{\Psi}_f=[\boldsymbol{I}_n-\boldsymbol{K}_f\boldsymbol{H}]\overline{\boldsymbol{\Phi}} \tag{3.9.46}$$

因而有

$$\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_f)=\psi(q^{-1}) \tag{3.9.47}$$

将 $(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_f)^{-1}=\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_f)/\psi(q^{-1})$ 代入(3.9.45)引出(3.9.41)和(3.9.42).

对 $N=-1$ 有 Wiener 预报器

$$\hat{\boldsymbol{x}}(t\mid t-1)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{y}(t-1) \tag{3.9.48}$$

类似地，这引出(3.9.43)和(3.9.44). □

【例 3.9.1】 考虑雷达跟踪系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}1 & T\\ 0 & \rho\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\ 0\end{bmatrix}w(t) \tag{3.9.49}$$

$$y(t)=[1\quad 0]\boldsymbol{x}(t)+v(t) \tag{3.9.50}$$

其中 $\boldsymbol{x}(t)=[s(t),\dot{s}(t)]^{\mathrm{T}}$，$\rho$ 为模型参数，$s(t)$ 和 $\dot{s}(t)$ 各为运动目标在时刻 tT 的位置和速度，T 为采样周期，$w(t)$ 和 $v(t)$ 是零均值、方差各为 σ_v^2 的独立白噪声. 容易导出在 Wiener 滤波器形式下的 Kalman 稳态滤波器为

$$\hat{\boldsymbol{x}}(t\mid t)=\boldsymbol{\Psi}_f\hat{\boldsymbol{x}}(t-1\mid t-1)+\boldsymbol{K}_f y(t) \tag{3.9.51}$$

$$\boldsymbol{K}_f=\begin{bmatrix}\alpha\\ \beta/T\end{bmatrix},\quad \boldsymbol{\Psi}_f=\begin{bmatrix}1-\alpha & (1-\alpha)T\\ -\beta/T & \rho-\beta\end{bmatrix} \tag{3.9.52}$$

并称其为 $\alpha-\beta$ 滤波器. 由定理 3.9.5 可得解耦 Wiener 滤波器为

$$[1-(1-\alpha+\rho-\beta)q^{-1}+\rho(1-\alpha)q^{-2}]\hat{s}(t|t)=[\alpha+(\beta-\alpha\rho)q^{-1}]y(t) \tag{3.9.53}$$

$$[1-(1-\alpha+\rho-\beta)q^{-1}+\rho(1-\alpha)q^{-2}]\hat{\dot{s}}(t|t)=[\beta(1-q^{-1})/T]y(t) \tag{3.9.54}$$

3.9.5 观测 Wiener 预报器

【定理 3.9.6】 带相关噪声定常系统(3.9.1)和(3.9.2)在节 3.8 的假设 1 和假设 3 下,有观测 $\boldsymbol{y}(t+N)$的 Wiener 预报器为

$$\psi(q^{-1})\hat{\boldsymbol{y}}(t+N|t)=\boldsymbol{J}_N(q^{-1})\boldsymbol{y}(t) \tag{3.9.55}$$

其中定义多项式矩阵 $\boldsymbol{J}_N(q^{-1})$为

$$\boldsymbol{J}_N(q^{-1})=\boldsymbol{H\Phi}^{N-1}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p,\quad N>0,$$

$$\boldsymbol{J}_N(q^{-1})=\psi(q^{-1})\boldsymbol{I}_m q^N,\quad N\leqslant 0 \tag{3.9.56}$$

稳态预报误差 $\tilde{\boldsymbol{y}}(t+N|t)=\boldsymbol{y}(t+N)-\hat{\boldsymbol{y}}(t+N|t)$的方差阵 $\boldsymbol{P}_{yN}=\mathrm{E}[\tilde{\boldsymbol{y}}(t+N|t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+N|t)]$为

$$\boldsymbol{P}_{yN}=\boldsymbol{HP}_N\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v,\quad N>1 \tag{3.9.57}$$

其中 $\boldsymbol{P}_N$为状态预报 $\hat{\boldsymbol{x}}(t+N|t)$的稳态方差阵,由(3.8.28)定义为

$$\boldsymbol{P}_N=\boldsymbol{\Phi}^{N-1}\boldsymbol{\Sigma}(\boldsymbol{\Phi}^{N-1})^{\mathrm{T}}+\sum_{j=2}^{N}\boldsymbol{\Phi}^{N-j}\boldsymbol{\Gamma Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Phi}^{(N-j)\mathrm{T}},\quad N>1 \tag{3.9.58}$$

对 $N=1$ 有

$$\boldsymbol{P}_{y1}=\boldsymbol{H\Sigma H}^{\mathrm{T}}+\boldsymbol{Q}_v \tag{3.9.59}$$

对 $N\leqslant 0$ 有

$$\boldsymbol{P}_{yN}=\boldsymbol{0},\quad N\leqslant 0 \tag{3.9.60}$$

证明 由(3.9.2)和射影性质,注意 $\boldsymbol{v}(t+N)\perp(\boldsymbol{y}(t),\cdots,\boldsymbol{y}(1))$,则有

$$\hat{\boldsymbol{y}}(t+N|t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t+N|t),\quad N>0 \tag{3.9.61}$$

而注意

$$\hat{\boldsymbol{x}}(t+N|t)=\boldsymbol{\Phi}^{N-1}\hat{\boldsymbol{x}}(t+1|t),$$

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{y}(t) \tag{3.9.62}$$

由此可得(3.9.55)和(3.9.56)的第一式. 对 $N\leqslant 0$,我们有 $\hat{\boldsymbol{y}}(t+N|t)=\boldsymbol{y}(t+N)$,这引出(3.9.55)和(3.9.56)的第二式. 对 $N>1$,由(3.9.2)和(3.9.61)相减引出

$$\tilde{\boldsymbol{y}}(t+N|t)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t+N|t)+\boldsymbol{v}(t+N) \tag{3.9.63}$$

易知 $\boldsymbol{v}(t+N)\perp\tilde{\boldsymbol{x}}(t+N|t)$,故有(3.9.59)成立. 对 $N\leqslant 0$ 有 $\hat{\boldsymbol{y}}(t+N|t)=\boldsymbol{y}(t+N)$,因而 $\tilde{\boldsymbol{y}}(t+N|t)=\boldsymbol{0}$,这引出(3.9.60)成立. □

3.9.6 多通道 ARMA 信号 Wiener 滤波器

利用基于 Kalman 滤波的白噪声 Wiener 滤波器和观测 Wiener 预报器,可以设计多通道 ARMA 信号 Wiener 滤波器[7,8]. 它可统一处理滤波、平滑和预报问题.

考虑多通道 ARMA 信号 Wiener 滤波问题

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{3.9.64}$$

$$\boldsymbol{y}(t)=\boldsymbol{s}(t)+\boldsymbol{v}(t) \tag{3.9.65}$$

其中 $\boldsymbol{s}(t)\in R^m$ 为待估的多通道 ARMA 信号，$\boldsymbol{y}(t)\in R^m$ 为对 $\boldsymbol{s}(t)$ 的观测信号，$\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}(t)\in R^m$ 为零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$ 的独立白噪声，q^{-1} 为单位滞后算子，

$$\boldsymbol{A}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{A}_1q^{-1}+\cdots+\boldsymbol{A}_{n_a}q^{-n_a},$$

$$\boldsymbol{C}(q^{-1})=\boldsymbol{C}_1q^{-1}+\cdots+\boldsymbol{C}_{n_c}q^{-n_c} \tag{3.9.66}$$

其中 $m\times m$ 阵 $\boldsymbol{A}_i$ 和 $m\times r$ 阵 $\boldsymbol{C}_i$ 为系数阵，设阶次 $n_a\geqslant n_c$，且多项式矩阵 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{C}(q^{-1})$ 左素．问题是基于观测 $(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots)$，设计信号 $\boldsymbol{s}(t)$ 的 Wiener 滤波器 $\hat{\boldsymbol{s}}(t\mid t+N)$．

由定理 1.6.1，式(3.9.64)和(3.9.65)有等价的状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t),$$

$$\boldsymbol{y}(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t) \tag{3.9.67}$$

其中定义 $\boldsymbol{C}_i=\boldsymbol{0}(i>n_c)$，且

$$\boldsymbol{\Phi}=\begin{bmatrix} -\boldsymbol{A}_1 & & & \\ \vdots & \boldsymbol{I}_{m(n_a-1)} & & \\ -\boldsymbol{A}_{n_a} & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\boldsymbol{C}_1\\ \vdots \\ \boldsymbol{C}_{n_a}\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{I}_m\ \boldsymbol{0}\ \cdots\ \boldsymbol{0}] \tag{3.9.68}$$

由 $(\boldsymbol{A}(q^{-1}),\boldsymbol{C}(q^{-1}))$ 左素的假设引出系统(3.9.66)和(3.9.67)是完全可观、完全可控的，因而存在稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{K}_p\boldsymbol{y}(t) \tag{3.9.69}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{\varepsilon}(t) \tag{3.9.70}$$

它引出 ARMA 新息模型

$$\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t)=\psi(q^{-1})\boldsymbol{\varepsilon}(t) \tag{3.9.71}$$

$$\psi(q^{-1})=\det(I_n-q^{-1}\boldsymbol{\Psi}_p) \tag{3.9.72}$$

$$\boldsymbol{\Lambda}(q^{-1})=\psi(q^{-1})\boldsymbol{I}_m-\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_pq^{-1} \tag{3.9.73}$$

其中 $\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H}$，$\boldsymbol{K}_p=\boldsymbol{\Phi K}_f$，$\boldsymbol{K}_f=\boldsymbol{\Sigma H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\boldsymbol{Q}_v]^{-1}$，且 $\boldsymbol{\Sigma}$ 满足如下稳态 Riccati 方程

$$\boldsymbol{\Sigma}=\boldsymbol{\Phi}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}}+\boldsymbol{Q}_v)^{-1}\boldsymbol{H\Sigma}]\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{\Gamma Q}_w\boldsymbol{\Gamma}^{\mathrm{T}} \tag{3.9.74}$$

它可用迭代法求解．由 $\boldsymbol{\Psi}_p$ 是稳定矩阵[2]引出 $\psi(q^{-1})$ 是稳定多项式．应用定理 3.9.3 可得白噪声 Wiener 滤波器

$$\psi(q^{-1})\hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{L}_N^v(q^{-1})\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.75}$$

由定理 3.9.6 可得 Wiener 观测预报器

$$\psi(q^{-1})\hat{\boldsymbol{y}}(t\mid t+N)=\boldsymbol{J}_{-N}(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.76}$$

由(3.9.63)和射影性质有

$$\hat{\boldsymbol{s}}(t\mid t+N)=\hat{\boldsymbol{y}}(t\mid t+N)-\hat{\boldsymbol{v}}(t\mid t+N) \tag{3.9.77}$$

将(3.9.75)和(3.9.76)代入(3.9.77)可得如下定理．

【定理 3.9.7】 多通道 ARMA 信号(3.9.62)和(3.9.63)有渐近稳定的 Wiener 滤波器

$$\psi(q^{-1})\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.78}$$

其中定义多项式矩阵 $\boldsymbol{K}_N(q^{-1})$ 为

$$\boldsymbol{K}_N(q^{-1})=\boldsymbol{J}_{-N}(q^{-1})-\boldsymbol{L}_N^v(q^{-1})\boldsymbol{\Lambda}(q^{-1}) \tag{3.9.79}$$

稳态滤波误差 $\tilde{\boldsymbol{s}}(t|t+N)=\boldsymbol{s}(t)-\hat{\boldsymbol{s}}(t|t+N)$ 方差阵 $\boldsymbol{P}_{sN}=\mathrm{E}[\tilde{\boldsymbol{s}}(t|t+N)\tilde{\boldsymbol{s}}^{\mathrm{T}}(t|t+N)]$ 为

$$\boldsymbol{P}_{sN}=\boldsymbol{H}\boldsymbol{P}_N\boldsymbol{H}^{\mathrm{T}} \tag{3.9.80}$$

其中状态估值误差 $\tilde{\boldsymbol{x}}(t|t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t+N)$ 的误差方差阵 $\boldsymbol{P}_N=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t+N)]$ 由定理 3.8.4 和定理 3.8.5 给出为

$$\boldsymbol{P}_N=\boldsymbol{\Sigma}-\sum_{i=0}^{N}\boldsymbol{K}_i\boldsymbol{Q}_\varepsilon\boldsymbol{K}_i^{\mathrm{T}},\quad N\geqslant 0 \tag{3.9.81}$$

$$\boldsymbol{P}_N=\boldsymbol{\Phi}^{-N-1}\boldsymbol{\Sigma}(\boldsymbol{\Phi}^{-N-1})^{\mathrm{T}}+\sum_{j=2}^{-N}\boldsymbol{\Phi}^{-N-j}\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^{-N-j})^{\mathrm{T}},\quad N<-1 \tag{3.9.82}$$

$$\boldsymbol{P}_{-1}=\boldsymbol{\Sigma} \tag{3.9.83}$$

其中 $\boldsymbol{Q}_\varepsilon, \boldsymbol{K}_f$ 由定理 3.8.5 计算.

证明 将(3.9.75)和(3.9.76)代入(3.9.77)得(3.9.78)和(3.9.79). 由 $\psi((q^{-1})$ 是稳定的多项式引出(3.9.78)是渐近稳定的. 比较(3.9.63)和(3.9.67)有

$$\boldsymbol{s}(t)=\boldsymbol{H}\boldsymbol{x}(t) \tag{3.9.84}$$

在上式两边取射影运算有

$$\hat{\boldsymbol{s}}(t|t+N)=\boldsymbol{H}\hat{\boldsymbol{x}}(t|t+N) \tag{3.9.85}$$

上两式相减引出

$$\tilde{\boldsymbol{s}}(t|t+N)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t|t+N) \tag{3.9.86}$$

这引出(3.9.80)成立. 于是应用定理 3.8.4 和定理 3.8.5 得(3.9.81)~(3.9.83). □

【定理 3.9.8】 多通道 ARMA 信号(3.9.64)和(3.9.65)有渐近稳定的 Wiener 滤波器

$$\psi(q^{-1})\hat{\boldsymbol{s}}(t|t+N)=\boldsymbol{K}_N^s(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.87}$$

$$\boldsymbol{K}_N^s(q^{-1})=\boldsymbol{H}\boldsymbol{K}_N(q^{-1}),\quad \boldsymbol{H}=[\boldsymbol{I}_m\ \boldsymbol{0}\ \cdots\ \boldsymbol{0}] \tag{3.9.88}$$

其中 $\boldsymbol{K}_N(q^{-1})$ 由定理 3.9.4 计算,且估值误差方差阵 $\boldsymbol{P}_{sN}$ 由(3.9.80)~(3.9.83)计算.

证明 由定理 3.9.4 有系统(3.9.67)的 Wiener 状态滤波器

$$\psi(q^{-1})\hat{\boldsymbol{x}}(t|t+N)=\boldsymbol{K}_N(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.89}$$

注意 $\boldsymbol{s}(t)=\boldsymbol{H}\boldsymbol{x}(t)$,因而有

$$\hat{\boldsymbol{s}}(t|t+N)=\boldsymbol{H}\hat{\boldsymbol{x}}(t|t+N) \tag{3.9.90}$$

将(3.9.89)代入(3.9.90)得(3.9.87)和(3.9.88). 由关系 $\tilde{\boldsymbol{s}}(t|t+N)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t|t+N)$ 引出(3.9.80)~(3.9.83). 证毕. □

【例 3.9.2】 考虑雷达跟踪系统

$$(\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1})\boldsymbol{s}(t)=\boldsymbol{C}_1q^{-1}w(t) \tag{3.9.91}$$

$$\boldsymbol{y}(t)=\boldsymbol{s}(t)+\boldsymbol{v}(t) \tag{3.9.92}$$

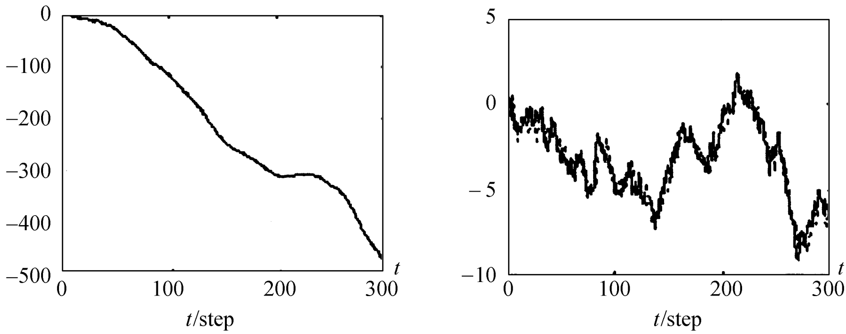
$$\boldsymbol{A}_1=-\begin{bmatrix}1 & T\\ 0 & 1\end{bmatrix},\quad \boldsymbol{C}_1=\begin{bmatrix}0.5T^2\\ T\end{bmatrix} \tag{3.9.93}$$

其中 $\boldsymbol{s}(t)=[s_1(t), s_2(t)]^{\mathrm{T}}$, $s_1(t)$, $s_2(t)$ 和 $w(t)$ 各为在时刻 tT 运动目标的位置、速度和加速度且 $\boldsymbol{y}(t)\in R^2$ 为对 $\boldsymbol{s}(t)$ 的观测信号, $\boldsymbol{v}(t)\in R^2$ 为观测噪声. 设 $w(t)$ 和 $v(t)$ 是零均值、方差阵各为 σ_w^2 和 $\boldsymbol{Q}_v$ 的独立白噪声. 问题是求 Wiener 信号平滑器 $\hat{\boldsymbol{s}}(t|t+1)$. 在仿真中取

$$T=0.5,\quad \sigma_w^2=1,$$

$$Q_v = \begin{bmatrix} 4 & 0 \\ 0 & 6.25 \end{bmatrix} \tag{3.9.94}$$

应用定理 3.9.7 仿真结果如图 3.9.1 所示，其中(a)为位置 $s_1(t)$和 Wiener 平滑器 $\hat{s}_1(t|t+1)$，(b)为速度 $s_2(t)$和 Wiener 平滑器 $\hat{s}_2(t|t+1)$，图中实线代表真实值，虚线代表平滑估值.



(a) 位置 $s_1(t)$和 Wiener 平滑器 $\hat{s}_1(t|t+1)$的比较图

(b) 位置 $s_2(t)$和 Wiener 平滑器 $\hat{s}_2(t|t+1)$的比较图

图 3.9.1　位置 $s_1(t)$和速度 $s_2(t)$及它们的 Wiener 跟踪平滑器

3.9.7　随机控制系统的状态分量解耦 Wiener 滤波器

在许多应用问题中，人们有时只对状态的某个分量或某几个分量的估计感兴趣.例如在以位置、速度和加速度为状态变量的雷达跟踪系统中，对目标位置跟踪而言，人们只对目标的位置的估计感兴趣.又例如在解决带观测噪声的 ARMA 信号 Wiener 滤波问题时，用状态空间方法构造相应的状态空间模型，待估 ARMA 信号只是状态的一个分量.因此用 Kalman 滤波器计算整个状态的估值会浪费了许多不必要的计算量，不便于实时应用，而只需计算相应于待估 ARMA 信号的那个状态分量的估值即可.这里将定理 3.9.5 的结果推广到随机控制系统.

考虑随机控制系统(3.8.1)和(3.8.2)在假设 1 ~ 3 下，有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t) = \boldsymbol{\Psi}_p \hat{\boldsymbol{x}}(t|t-1) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{K}_p \boldsymbol{y}(t) \tag{3.9.95}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1) + \boldsymbol{\varepsilon}(t) \tag{3.9.96}$$

【定理 3.9.9】　带相关噪声的定常随机控制系统(3.8.1)和(3.8.2)在假设 1 ~ 3 下，有受控的自回归滑动平均(CARMA)新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{B}(q^{-1})\boldsymbol{u}(t) + \psi(q^{-1})\boldsymbol{\varepsilon}(t) \tag{3.9.97}$$

其中定义

$$\psi(q^{-1}) = \det(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p) \tag{3.9.98}$$

$$\boldsymbol{A}(q^{-1}) = \psi(q^{-1})\boldsymbol{I}_m - \boldsymbol{H}\operatorname{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p q^{-1} \tag{3.9.99}$$

$$\boldsymbol{B}(q^{-1}) = \boldsymbol{H}\operatorname{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{B}q^{-1} \tag{3.9.100}$$

证明　由(3.9.95)有

$$\hat{\boldsymbol{x}}(t|t-1) = (\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)^{-1}[\boldsymbol{B}q^{-1}\boldsymbol{u}(t) + \boldsymbol{K}_p q^{-1}\boldsymbol{y}(t)] \tag{3.9.101}$$

将它代入(3.9.96)可得(3.9.97) ~ (3.9.100).　□

【定理 3.9.10】 带相关噪声的定常随机控制系统(3.8.1)和(3.8.2)在假设 1~3 下，记状态 $\boldsymbol{x}(t)$ 的分量形式为 $\boldsymbol{x}(t)=[x_1(t),x_2(t),\cdots,x_n(t)]^{\mathrm{T}}$，则有状态分量 $x_i(t)$ 的解耦 Wiener 滤波器和平滑器($N\geqslant 0$)为

$$\psi(q^{-1})\hat{x}_i(t\mid t+N)=\boldsymbol{K}_{Ni}^{u}(q^{-1})\boldsymbol{u}(t+N)+\boldsymbol{K}_{Ni}^{y}(q^{-1})\boldsymbol{y}(t+N),\quad i=1,\cdots,n \tag{3.9.102}$$

其中定义记号 $\boldsymbol{e}_i=[0\cdots0\ 1\ 0\cdots0]$，它是 $1\times n$ 行向量，它的第 i 个元素为 1，其余元素为零，且定义

$$\boldsymbol{K}_{Ni}^{u}(q^{-1})=\boldsymbol{e}_i\boldsymbol{K}_{N}^{u}(q^{-1}),\quad \boldsymbol{K}_{Ni}^{y}(q^{-1})=\boldsymbol{e}_i\boldsymbol{K}_{N}^{y}(q^{-1}) \tag{3.9.103}$$

$$\boldsymbol{K}_{N}^{u}(q^{-1})=\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{B}q^{-1-N}-\boldsymbol{B}\boldsymbol{K}_{N}^{s}(q^{-1})\boldsymbol{B}(q^{-1}) \tag{3.9.104}$$

$$\boldsymbol{K}_{N}^{y}(q^{-1})=\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p q^{-1-N}+\boldsymbol{K}_{N}^{s}(q^{-1})\boldsymbol{A}(q^{-1}) \tag{3.9.105}$$

$$\boldsymbol{K}_{N}^{s}(q^{-1})=\sum_{i=0}^{N}\boldsymbol{K}_i q^{i-N} \tag{3.9.106}$$

稳态分量估值误差 $\tilde{x}_i(t\mid t+N)=x_i(t)-\hat{x}_i(t\mid t+N)$ 的方差 $P_{Ni}=\mathrm{E}[\tilde{x}_i^2(t\mid t+N)]$($N\geqslant 0$)为

$$P_{Ni}=\boldsymbol{e}_i\boldsymbol{P}_N\boldsymbol{e}_i^{\mathrm{T}},\quad i=1,\cdots,n,N\geqslant 0 \tag{3.9.107}$$

即 P_{Ni} 为 $\boldsymbol{P}_N$ 的第(i,i)对角元素，$\boldsymbol{P}_N$ 为状态 $\boldsymbol{x}(t)$ 的估值误差 $\tilde{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t\mid t+N)$ 的误差方差阵，$\boldsymbol{P}_N$，$\boldsymbol{K}_i$，$\boldsymbol{\Psi}_p$ 均由节 3.8 有关公式计算.

对 $N=-1$ 有解耦 Wiener 状态分量预报器(3.9.102)，其中定义

$$\boldsymbol{K}_{-1i}^{u}(q^{-1})=\boldsymbol{e}_i\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{B} \tag{3.9.108}$$

$$\boldsymbol{K}_{-1i}^{y}(q^{-1})=\boldsymbol{e}_i\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p \tag{3.9.109}$$

稳态预报误差 $\tilde{x}_i(t\mid t-1)=x_i(t)-\hat{x}_i(t\mid t-1)$ 的方差 $P_{-1i}=\mathrm{E}[\tilde{x}_i^2(t\mid t-1)]$ 为

$$P_{-1i}=\boldsymbol{e}_i\boldsymbol{\Sigma}\boldsymbol{e}_i^{\mathrm{T}} \tag{3.9.110}$$

其中 $\boldsymbol{\Sigma}$ 为稳态 Kalman 预报误差 $\tilde{\boldsymbol{x}}(t\mid t-1)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t\mid t-1)$ 的方差阵，它由节 3.8 的稳态 Riccati 方程决定.这表明 P_{-1i} 为 $\boldsymbol{\Sigma}$ 的第(i,i)对角元素.

证明 对 $N\geqslant 0$ 由节 3.8 有

$$\hat{\boldsymbol{x}}(t\mid t+N)=\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{K}_{N}^{s}(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{3.9.111}$$

其中 $\boldsymbol{K}_{N}^{s}(q^{-1})$ 由(3.9.106)定义，由(3.9.97)有

$$\boldsymbol{\varepsilon}(t+N)=\psi^{-1}(q^{-1})[\boldsymbol{A}(q^{-1})\boldsymbol{y}(t+N)-\boldsymbol{B}(q^{-1})\boldsymbol{u}(t+N)] \tag{3.9.112}$$

将(3.9.101)和上式代入(3.9.111)可得 Wiener 滤波器和平滑器

$$\psi(q^{-1})\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{K}_{N}^{u}(q^{-1})\boldsymbol{u}(t+N)+\boldsymbol{K}_{N}^{y}(q^{-1})y(t+N),\quad N\geqslant 0 \tag{3.9.113}$$

将它按分量表示，可得(3.9.102)~(3.9.106)，对 $N=-1$，由(3.9.101)立即引出(3.9.108)和(3.9.109)，且有

$$\psi(q^{-1})\hat{x}_i(t\mid t-1)=\boldsymbol{K}_{-1i}^{u}(q^{-1})\boldsymbol{u}(t-1)+\boldsymbol{K}_{-1i}^{y}(q^{-1})\boldsymbol{y}(t-1),\quad i=1,\cdots,n \tag{3.9.114}$$

【例 3.9.3】 考虑带相关噪声的随机控制系统

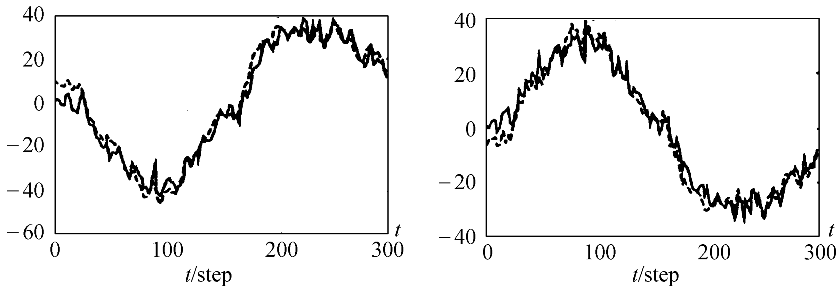
$$\boldsymbol{x}(t+1)=\begin{bmatrix}0.6 & -0.5\\ 0 & 0.9\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\3\end{bmatrix}u(t)+\begin{bmatrix}-1\\2\end{bmatrix}w(t) \tag{3.9.115}$$

$$y(t)=[1\quad 1]\boldsymbol{x}(t)+v(t) \tag{3.9.116}$$

$$v(t) = w(t) + \xi(t) \tag{3.9.117}$$

$$u(t) = \sin(2\pi t/300) \tag{3.9.118}$$

其中状态 $\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}$，$w(t)$和 $\xi(t)$是零均值、方差各为 $\sigma_{2_w} = 1$ 和 $\sigma_\xi^2 = 1$ 的独立高斯白噪声.显然有 $w(t)$和 $v(t)$是相关的白噪声，相关系数为 $\mathrm{E}[w(t)v(t)] = S = \sigma_w^2 = 1$.取 $N = 2$，应用定理 3.9.10 可求得分量解耦 Wiener 平滑器 $\hat{x}_i(t|t+2)$，$i = 1,2$.仿真结果如图 3.9.2 所示，其中虚线代表真实值，实线代表平滑估值.



(a) 分量 $x_1(t)$和解耦 Wiener 平滑器 $\hat{x}_1(t|t+2)$

(b) 分量 $x_2(t)$和解耦 Wiener 平滑器 $\hat{x}_2(t|t+2)$

图 3.9.2　解耦 Wiener 平滑器 $\hat{x}(t|t+2)$

3.9.8　带有色观测噪声随机控制系统状态分量解耦 Wiener 滤波器

考虑带有色观测噪声的定常随机控制系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi x}(t) + \boldsymbol{Bu}(t) + \boldsymbol{\Gamma w}(t) \tag{3.9.119}$$

$$\boldsymbol{z}(t) = \boldsymbol{H}_0\boldsymbol{x}(t) + \boldsymbol{\eta}(t) \tag{3.9.120}$$

$$\boldsymbol{\eta}(t+1) = \boldsymbol{A\eta}(t) + \boldsymbol{\xi}(t) \tag{3.9.121}$$

其中状态 $\boldsymbol{x}(t) \in R^n$，$\boldsymbol{x}(t) = [x_1(t), \cdots, x_n(t)]^{\mathrm{T}}$，观测 $\boldsymbol{z}(t) \in R^m$，控制 $\boldsymbol{u}(t) \in R^p$，$\boldsymbol{\eta}(t)$为有色观测噪声，服从模型(3.9.121)，$\boldsymbol{\xi}(t) \in R^m$，$\boldsymbol{w}(t) \in R^r$ 是零均值、方差阵各为 $\boldsymbol{Q}_\xi$ 和 $\boldsymbol{Q}_w$ 的相互独立白噪声，$\boldsymbol{\Phi}, \boldsymbol{B}, \boldsymbol{\Gamma}, \boldsymbol{H}_0, \boldsymbol{A}$ 是适当维数常阵.且设 $\boldsymbol{u}(t)$是有界的.问题是基于观测$(\boldsymbol{z}(t+N), \boldsymbol{z}(t+N-1), \cdots)$求状态分量 $x_i(t)$的解耦 Wiener 滤波器 $\hat{x}_i(t|t+N)$，$i = 1, \cdots, n$.

有色观测噪声 $\boldsymbol{\eta}(t)$可表为

$$(\boldsymbol{I}_m - q^{-1}\boldsymbol{A})\boldsymbol{\eta}(t+1) = \boldsymbol{\xi}(t) \tag{3.9.122}$$

在(3.9.120)中置 $t = t + 1$ 后用$(\boldsymbol{I}_m - q^{-1}\boldsymbol{A})$左乘，再利用(3.9.19)可得新的观测方程

$$\boldsymbol{y}(t) = \boldsymbol{H}x(t) + \boldsymbol{v}(t) \tag{3.9.123}$$

其中定义新的观测过程是 $\boldsymbol{y}(t)$，观测阵 $\boldsymbol{H}$ 和观测噪声 $\boldsymbol{v}(t)$为

$$\boldsymbol{y}(t) = (\boldsymbol{I}_m - q^{-1}\boldsymbol{A})\boldsymbol{z}(t+1) - \boldsymbol{H}_0\boldsymbol{Bu}(t) \tag{3.9.124}$$

$$\boldsymbol{H} = \boldsymbol{H}_0\boldsymbol{\Phi} - \boldsymbol{AH}_0 \tag{3.9.125}$$

$$\boldsymbol{v}(t) = \boldsymbol{H}_0\boldsymbol{\Gamma w}(t) + \boldsymbol{\xi}(t) \tag{3.9.126}$$

对于等价系统(3.9.119)和(3.9.123)，显然 $\boldsymbol{w}(t)$和 $\boldsymbol{v}(t)$是带零均值白噪声，且它们是相关的，即它们有相关阵 $\boldsymbol{S}$ 为

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(t)] = \boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{H}_0^{\mathrm{T}} \tag{3.9.127}$$

且 $\boldsymbol{v}(t)$ 有方差阵

$$\boldsymbol{Q}_v = \mathrm{E}[\boldsymbol{v}(t)\boldsymbol{v}^{\mathrm{T}}(t)] = \boldsymbol{H}_0\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{H}_0^{\mathrm{T}} + \boldsymbol{Q}_\xi \tag{3.9.128}$$

带相关噪声的定常系统(3.9.119)和(3.9.123)显然具有(3.8.1)和(3.8.2)形式.假设它满足节 3.8 的假设 1 ~ 3,则应用定理 3.9.10 有分量解耦 Wiener 滤波器为

$$\psi(q^{-1})\hat{x}_i^{\gamma}(t \mid t+N) = \boldsymbol{K}_{Ni}^{u}(q^{-1})\boldsymbol{u}(t+N) + \boldsymbol{K}_{Ni}^{\gamma}(q^{-1})\boldsymbol{y}(t+N),$$

$$i = 1,\cdots,n, N = -1 \text{ 或 } N \geqslant 0 \tag{3.9.129}$$

其中 $\hat{x}_i^{\gamma}(t \mid t+N)$ 表示基于观测 $(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots)$ 对 $x_i(t)$ 的估值.

用 $\hat{x}_i(t \mid t+N)$ 表示基于原始观测 $(\boldsymbol{z}(t+N),\boldsymbol{z}(t+N-1),\cdots)$ 对 $x_i(t)$ 的估值,由定义(3.9.124)有相等的线性流形 $L(\boldsymbol{z}(t+N),\boldsymbol{z}(t+N-1),\cdots) = L(\boldsymbol{y}(t+N-1),\boldsymbol{y}(t+N-2),\cdots)$,因而有

$$\hat{x}_i(t \mid t+N) = \hat{x}_i^{\gamma}(t \mid t+N-1) \tag{3.9.130}$$

于是有如下定理.

【定理 3.9.11】 带有色观测噪声定常系统(3.9.119)和(3.9.120),假设它的等价系统(3.9.119)和(3.9.123)满足节 3.8 的假设 1 ~ 3,则有状态分量 $x_i(t)$ 的解耦 Wiener 滤波器为

$$\psi(q^{-1})\hat{x}_i(t \mid t+N) = \boldsymbol{D}_{Ni}^{u}(q^{-1})\boldsymbol{u}(t+N) + \boldsymbol{D}_{Ni}^{\gamma}(q^{-1})\boldsymbol{z}(t+N), \quad i = 1,\cdots,n \tag{3.9.131}$$

其中定义多项式矩阵

$$\boldsymbol{D}_{Ni}^{u}(q^{-1}) = [\boldsymbol{K}_{(N-1)i}^{u}(q^{-1}) - \boldsymbol{K}_{(N-1)i}^{\gamma}(q^{-1})\boldsymbol{H}_0\boldsymbol{B}]q^{-1} \tag{3.9.132}$$

$$\boldsymbol{D}_{Ni}^{\gamma}(q^{-1}) = \boldsymbol{K}_{(N-1)i}^{\gamma}(q^{-1})(\boldsymbol{I}_m - q^{-1}\boldsymbol{A}) \tag{3.9.133}$$

其中 $\boldsymbol{K}_{Ni}^{u}(q^{-1})$,$\boldsymbol{K}_{Ni}^{\gamma}(q^{-1})$,$\psi(q^{-1})$ 应用定理 3.9.10 求得,滤波误差方差 P_{Ni} 由(3.9.107)给出.

证明 由(3.9.129)置 $N = N-1$ 有

$$\psi(q^{-1})\hat{x}_i^{\gamma}(t \mid t+N-1) = \boldsymbol{K}_{(N-1)i}^{u}(q^{-1})\boldsymbol{u}(t+N-1) + \boldsymbol{K}_{(N-1)i}^{\gamma}(q^{-1})\boldsymbol{y}(t+N-1) \tag{3.9.134}$$

由(3.9.124)有

$$\boldsymbol{y}(t+N-1) = (\boldsymbol{I}_m - q^{-1}\boldsymbol{A})\boldsymbol{z}(t+N) - \boldsymbol{H}_0\boldsymbol{B}\boldsymbol{u}(t+N-1) \tag{3.9.135}$$

将(3.9.135)代入(3.9.134)得(3.9.131) ~ (3.9.133).证毕. □

【例 3.9.4】 考虑带有色观测噪声的雷达跟踪控制系统

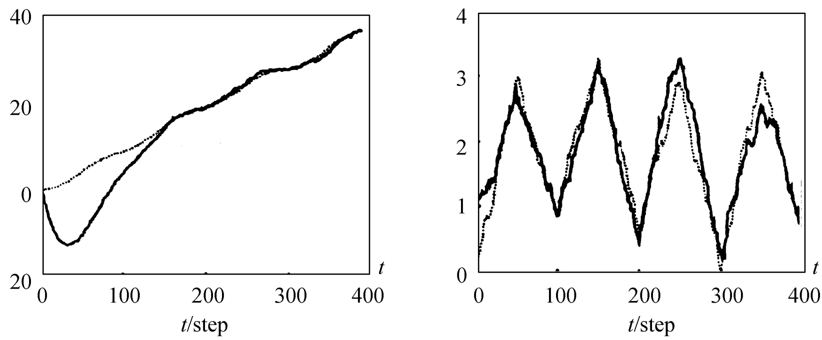
$$\boldsymbol{x}(t+1) = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}\boldsymbol{x}(t) + \begin{bmatrix} 0.5T^2 \\ T \end{bmatrix}u(t) + \begin{bmatrix} 0.5T^2 \\ T \end{bmatrix}w(t) \tag{3.9.136}$$

$$y(t) = [1 \quad 0]\boldsymbol{x}(t) + \eta(t) \tag{3.9.137}$$

$$\eta(t+1) = 0.8\eta(t) + \xi(t) \tag{3.9.138}$$

其中 T 为采样周期,$\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}$,$x_1(t)$,$x_2(t)$,$u(t)$ 和 $w(t)$ 分别为在时刻 tT 运动目标的位置、速度、机动加速度和随机加速度,且已知 $u(t)$ 是幅度为 ±1,宽度为 50 的方波,即 $u(t) = 1, 0 \leqslant t \leqslant 50$,$u(t) = -1, 50 \leqslant t \leqslant 100$,余类推.取 $T = 0.05$,$w(t)$ 和 $\xi(t)$ 是零均值、方差各为 $\sigma_{2_w} = 1$ 和 $\sigma_\xi^2 = 0.36$ 的独立高斯白噪声.取 $N = 2$ 应用定理 3.9.11 可求得分量解耦 Wiener 平滑器 $\hat{x}_i(t \mid t+2)$,$i = 1,2$.仿真结果如图 3.9.3 所示,其中虚线代

表真实值,实线代表估值.仿真中取初值 $\hat{x}_1(0|2)=1$,$\hat{x}_2(1|3)=1$.可看到在开始阶段平滑估值对真实值的跟踪有一个过程,过渡过程的长短不仅与初值有关,而且还与 $\boldsymbol{\Psi}_p$ 的特征值的位置有关.若 $\boldsymbol{\Psi}_p$ 的特征值均位于零点附近,则过渡过程很短,若 $\boldsymbol{\Psi}$ 有接近于单位圆周的特征值,则有较长的过渡过程.对本例 $\boldsymbol{\Psi}_p$ 有特征值 $\lambda_{1,2}=0.925\,8\pm0.019\,06i$ 接近单位圆周,故有较长过渡过程.



(a) 位置 $x_1(t)$ 和解耦 Wiener 平滑器 $\hat{x}_1(t|t+2)$

(b) 速度 $x_2(t)$ 和解耦 Wiener 平滑器 $\hat{x}_2(t|t+2)$

图 3.9.3 解耦 Wiener 平滑器 $\hat{x}(t|t+2)$

3.9.9 带有色观测噪声的多通道 ARMA 信号 Wiener 滤波器

考虑带有色观测噪声和白噪声的多通道 ARMA 信号

$$\boldsymbol{y}(t)=\boldsymbol{s}(t)+\boldsymbol{\eta}(t)+\boldsymbol{\xi}(t) \tag{3.9.139}$$

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{3.9.140}$$

$$\boldsymbol{P}(q^{-1})\boldsymbol{\eta}(t)=\boldsymbol{R}(q^{-1})\boldsymbol{n}(t) \tag{3.9.141}$$

其中 $\boldsymbol{y}(t)\in R^m$ 为观测信号,待估信号 $\boldsymbol{s}(t)\in R^m$ 服从向量 ARMA 模型(3.9.140),$\boldsymbol{\eta}(t)\in R^m$ 为有色观测噪声,服从 ARMA 模型(3.9.141),$\boldsymbol{\xi}(t)\in R^m$ 为白色观测噪声.$\boldsymbol{A}(q^{-1}),\cdots,\boldsymbol{R}(q^{-1})$ 为形如 $\boldsymbol{X}(q^{-1})=\boldsymbol{X}_0+\boldsymbol{X}_1q^{-1}+\cdots+\boldsymbol{X}_{n_x}q^{-n_x}$ 的多项式矩阵,n_x 为阶次,$\boldsymbol{X}_i$ 为系数阵.假设 $n_a\geqslant n_c$,$n_p\geqslant n_r$,$\boldsymbol{A}_0=\boldsymbol{I}_m$,$\boldsymbol{P}_0=\boldsymbol{I}_m$,$\boldsymbol{C}_0=\boldsymbol{I}_m$ 或 $\boldsymbol{C}_0=0$,$\boldsymbol{R}_0=\boldsymbol{I}_m$ 或 $\boldsymbol{R}_0=\boldsymbol{0}$.假设 $\boldsymbol{\xi}(t)$,$\boldsymbol{e}(t)\in R^r$ 和 $\boldsymbol{n}(t)\in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_\xi$,$\boldsymbol{Q}_e$ 和 $\boldsymbol{Q}_n$ 的相互独立白噪声.问题是基于观测 $(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots)$ 求信号 $\boldsymbol{s}(t)$ 的 Wiener 滤波器 $\hat{\boldsymbol{s}}(t|t+N)$,$N=0$,$N>0$ 或 $N<0$.这一 Wiener 滤波问题可用图 3.9.4 表示.

以下我们用状态空间方法基于 Kalman 滤波方法解决上述问题.

应用定理 1.6.1ARMA 模型(3.9.140)和(3.9.141)各有块伴随形状态空间模型

$$\boldsymbol{\alpha}(t+1)=\bar{\boldsymbol{A}}\boldsymbol{\alpha}(t)+\bar{\boldsymbol{C}}\boldsymbol{e}(t) \tag{3.9.142}$$

$$\boldsymbol{s}(t+1)=\bar{\boldsymbol{H}}_1\boldsymbol{\alpha}(t)+\boldsymbol{C}_0\boldsymbol{e}(t) \tag{3.9.143}$$

和

$$\boldsymbol{\beta}(t+1)=\bar{\boldsymbol{P}}\boldsymbol{\beta}(t)+\bar{\boldsymbol{R}}\boldsymbol{n}(t) \tag{3.9.144}$$

$$\boldsymbol{\eta}(t)=\bar{\boldsymbol{H}}_2\boldsymbol{\beta}(t)+\boldsymbol{R}_0\boldsymbol{n}(t) \tag{3.9.145}$$

其中规定 $\boldsymbol{C}_i=\boldsymbol{0}(i>n_c)$,$\boldsymbol{R}_i=\boldsymbol{0}(i>n_r)$,

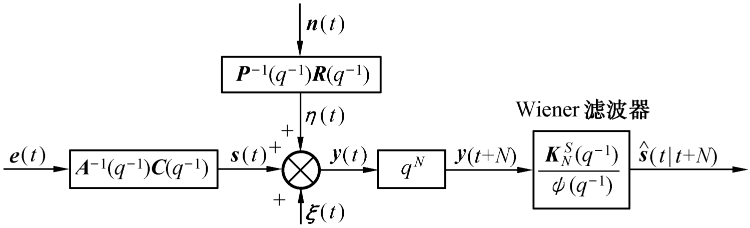


图 3.9.4　一般多通道 ARMA 信号 Wiener 滤波问题

$$\bar{\boldsymbol{A}}=\begin{bmatrix}-\boldsymbol{A}_1 & & \\ \vdots & \boldsymbol{I}_{m(n_a-1)} & \\ -\boldsymbol{A}_{n_a} & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix},\quad \bar{\boldsymbol{C}}=\begin{bmatrix}\boldsymbol{C}_1-\boldsymbol{A}_1\boldsymbol{C}_0\\ \vdots \\ \boldsymbol{C}_{n_a}-\boldsymbol{A}_{n_a}\boldsymbol{C}_0\end{bmatrix},\quad \bar{\boldsymbol{H}}_1=[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}] \tag{3.9.146}$$

$$\bar{\boldsymbol{P}}=\begin{bmatrix}-\boldsymbol{P}_1 & & \\ \vdots & \boldsymbol{I}_{m(n_p-1)} & \\ -\boldsymbol{P}_{n_p} & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix},\quad \bar{\boldsymbol{R}}=\begin{bmatrix}\boldsymbol{R}_1-\boldsymbol{P}_1\boldsymbol{R}_0\\ \vdots \\ \boldsymbol{R}_{n_p}-\boldsymbol{P}_{n_p}\boldsymbol{R}_0\end{bmatrix},\quad \bar{\boldsymbol{H}}_2=[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}] \tag{3.9.147}$$

引入增广状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t) \tag{3.9.148}$$

$$\boldsymbol{y}(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t) \tag{3.9.149}$$

其中定义增广状态 $\boldsymbol{x}(t)$，增广噪声 $\boldsymbol{w}(t)$和观测噪声 $\boldsymbol{v}(t)$为

$$\boldsymbol{x}(t)=\begin{bmatrix}\boldsymbol{\alpha}(t)\\ \boldsymbol{\beta}(t)\end{bmatrix},\quad \boldsymbol{w}(t)=\begin{bmatrix}\boldsymbol{e}(t)\\ \boldsymbol{n}(t)\end{bmatrix},\quad \boldsymbol{v}(t)=\boldsymbol{C}_0\boldsymbol{e}(t)+\boldsymbol{R}_0\boldsymbol{n}(t)+\boldsymbol{\xi}(t) \tag{3.9.150}$$

且定义

$$\boldsymbol{\Phi}=\begin{bmatrix}\bar{\boldsymbol{A}} & \boldsymbol{0}\\ \boldsymbol{0} & \bar{\boldsymbol{P}}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\bar{\boldsymbol{C}} & \boldsymbol{0}\\ \boldsymbol{0} & \bar{\boldsymbol{R}}\end{bmatrix},\quad \boldsymbol{H}=[\bar{\boldsymbol{H}}_1,\bar{\boldsymbol{H}}_2] \tag{3.9.151}$$

易知增广系统(3.9.148)和(3.9.149)是带相关噪声系统，易知 $\boldsymbol{w}(t)$和 $\boldsymbol{v}(t)$是带零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$，相关阵为 $\boldsymbol{S}$ 的白噪声，且有

$$\boldsymbol{Q}_w=\begin{bmatrix}\boldsymbol{Q}_e & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{Q}_n\end{bmatrix},\quad \boldsymbol{Q}_v=\boldsymbol{C}_0\boldsymbol{Q}_e\boldsymbol{C}_0^{\mathrm{T}}+\boldsymbol{R}_0\boldsymbol{Q}_n\boldsymbol{R}_0^{\mathrm{T}}+\boldsymbol{Q}_\xi,$$

$$\boldsymbol{S}=\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(t)]=\begin{bmatrix}\boldsymbol{Q}_e\boldsymbol{C}_0^{\mathrm{T}}\\ \boldsymbol{Q}_n\boldsymbol{R}_0^{\mathrm{T}}\end{bmatrix} \tag{3.9.152}$$

假设增广系统是完全可观、完全能稳[2]的或假设 $\boldsymbol{\Phi}$ 是稳定矩阵，则它存在稳态 Kalman 预报器，进而由本节前述定理可得增广系统的 Wiener 状态滤波器

$$\psi(q^{-1})\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.153}$$

且可求得白噪声 Wiener 滤波器

$$\psi(q^{-1})\hat{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{K}_N^w(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.154}$$

由(3.9.143)和射影性质引出

$$\hat{\boldsymbol{s}}(t\mid t+N)=\bar{\boldsymbol{H}}_1\hat{\boldsymbol{\alpha}}(t\mid t+N)+\boldsymbol{C}_0\hat{\boldsymbol{e}}(t\mid t+N) \tag{3.9.155}$$

注意定义(3.9.150)，它可写为

$$\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{H}_s\hat{\boldsymbol{x}}(t\mid t+N)+\boldsymbol{C}_0\boldsymbol{H}_e\hat{\boldsymbol{w}}(t\mid t+N) \tag{3.9.156}$$

其中定义 $m\times(n_a+n_p)$阵 $\boldsymbol{H}_s$ 和 $r\times(r+m)$阵 $\boldsymbol{H}_e$ 分别当

$$\boldsymbol{H}_s=[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}],\quad \boldsymbol{H}_e=[\boldsymbol{I}_r\quad \boldsymbol{0}] \tag{3.9.157}$$

【定理 3.9.12】 多通道 ARMA 信号 Wiener 滤波问题(3.9.139)~(3.9.141)有 Wiener 滤波器

$$\psi(q^{-1})\hat{\boldsymbol{s}}(t|t+N)=\boldsymbol{K}_N^s(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.158}$$

$$\boldsymbol{K}_N^s(q^{-1})=\boldsymbol{H}_s\boldsymbol{K}_N(q^{-1})+\boldsymbol{C}_0\boldsymbol{H}_e\boldsymbol{K}_N^w(q^{-1}) \tag{3.9.159}$$

记对增广系统状态 $\boldsymbol{x}(t)$的估值 $\hat{\boldsymbol{x}}(t|t+N)$和白噪声 $\boldsymbol{w}(t)$的估值 $\hat{\boldsymbol{w}}(t|t+N)$的误差方差阵各为 $\boldsymbol{P}_N$ 和 $\boldsymbol{P}_N^w$,则当 $N\geqslant 0$ 时,对信号 $\boldsymbol{s}(t)$的估值 $\hat{\boldsymbol{s}}(t|t+N)$的误差 $\tilde{\boldsymbol{s}}(t|t+N)$的方差阵 $\boldsymbol{P}_N^s=\mathrm{E}[\tilde{\boldsymbol{s}}(t|t+N)\tilde{\boldsymbol{s}}^{\mathrm{T}}(t|t+N)]$为

$$\boldsymbol{P}_N^s=\boldsymbol{H}_s\boldsymbol{P}_N\boldsymbol{H}_s^{\mathrm{T}}+\boldsymbol{C}_0\boldsymbol{H}_e\boldsymbol{P}_N^w\boldsymbol{H}_e^{\mathrm{T}}\boldsymbol{C}_0^{\mathrm{T}}+\boldsymbol{H}_s\boldsymbol{R}_{\tilde{x}\tilde{w}}\boldsymbol{H}_e^{\mathrm{T}}\boldsymbol{C}_0^{\mathrm{T}}+\boldsymbol{C}_0\boldsymbol{H}_e\boldsymbol{R}_{\tilde{x}\tilde{w}}^{\mathrm{T}}\boldsymbol{H}_s^{\mathrm{T}},\quad N\geqslant 0 \tag{3.9.160}$$

其中 $\boldsymbol{R}_{\tilde{x}\tilde{w}}=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t+N)\tilde{\boldsymbol{w}}^{\mathrm{T}}(t|t+N)]$为

$$\boldsymbol{R}_{\tilde{x}\tilde{w}}=\sum_{i=0}^{N}\sum_{j=0}^{N}[\boldsymbol{I}_n\delta_{0i},\ -\boldsymbol{K}_i]\begin{bmatrix}\boldsymbol{0} & \boldsymbol{\Sigma}(\boldsymbol{\Psi}_p^{\mathrm{T}})^j\boldsymbol{H}^{\mathrm{T}}\\ \boldsymbol{H}\boldsymbol{\Psi}_p^{i-1}\boldsymbol{D}_w^{\mathrm{T}} & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_n\delta_{0j}\\ -\boldsymbol{M}_w^{\mathrm{T}}(j)\end{bmatrix} \tag{3.9.161}$$

当 $N<0$ 时,有

$$\boldsymbol{P}_N^s=\boldsymbol{H}_s\boldsymbol{P}_N\boldsymbol{H}_s^{\mathrm{T}}+\boldsymbol{C}_0\boldsymbol{H}e\boldsymbol{Q}_w\boldsymbol{H}_e^{\mathrm{T}}\boldsymbol{C}_0^{\mathrm{T}} \tag{3.9.162}$$

其中 $\boldsymbol{\Psi}_p,\boldsymbol{K}_i,\boldsymbol{M}_w(j)$分别是为增广系统稳态 Kalman 预报器的转移阵,稳态 Kalman 平滑器的平滑增益阵及白噪声 $w(t)$的稳态平滑估值增益阵,$\boldsymbol{\Sigma}$ 和 $\boldsymbol{K}_p$ 为 Kalman 一步预报稳态误差主差阵和增益阵,$\boldsymbol{Q}_\varepsilon$ 为稳态新息方差阵,$\delta_{tj}=0(t\neq j)$,$\delta_{tt}=1$,且

$$\boldsymbol{D}_w=\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{S}\boldsymbol{K}_p^{\mathrm{T}} \tag{3.9.163}$$

证明 将(3.9.153)和(3.9.154)代入(3.9.156)得(3.9.158)和(3.9.159).由(3.9.143)有

$$\boldsymbol{s}(t)=\boldsymbol{H}_s\boldsymbol{x}(t)+\boldsymbol{C}_0\boldsymbol{H}_e\boldsymbol{w}(t) \tag{3.9.164}$$

其中 $\boldsymbol{H}_s,\boldsymbol{H}_e$ 由(3.9.157)定义,又对上式取射影运算有

$$\hat{\boldsymbol{s}}(t|t+N)=\boldsymbol{H}_s\hat{\boldsymbol{x}}(t|t+N)+\boldsymbol{C}_0\boldsymbol{H}_e\hat{\boldsymbol{w}}(t|t+N) \tag{3.9.165}$$

上两式相减有

$$\tilde{\boldsymbol{s}}(t|t+N)=\boldsymbol{H}_s\tilde{\boldsymbol{x}}(t|t+N)+\boldsymbol{C}_0\boldsymbol{H}_e\tilde{\boldsymbol{w}}(t|t+N) \tag{3.9.166}$$

当 $N\geqslant 0$ 时我们有

$$\boldsymbol{P}_N^s=\boldsymbol{H}_s\boldsymbol{P}_N\boldsymbol{H}_s^{\mathrm{T}}+\boldsymbol{C}_0\boldsymbol{H}_e\boldsymbol{P}_N^w\boldsymbol{H}_e^{\mathrm{T}}\boldsymbol{C}_0^{\mathrm{T}}+\boldsymbol{H}_s\boldsymbol{R}_{\tilde{x}\tilde{w}}\boldsymbol{H}_e^{\mathrm{T}}\boldsymbol{C}_0^{\mathrm{T}}+\boldsymbol{C}_0\boldsymbol{H}_e\boldsymbol{R}_{\tilde{x}\tilde{w}}^{\mathrm{T}}\boldsymbol{H}_s^{\mathrm{T}} \tag{3.9.167}$$

$$\boldsymbol{R}_{\tilde{x}\tilde{w}}=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t+N)\tilde{\boldsymbol{w}}^{\mathrm{T}}(t|t+N)] \tag{3.9.168}$$

注意关系

$$\tilde{\boldsymbol{x}}(t\mid t+N)=\tilde{\boldsymbol{x}}(t\mid t-1)-\sum_{i=0}^{N}\boldsymbol{K}_i\boldsymbol{\varepsilon}(t+i) \tag{3.9.169}$$

$$\hat{\boldsymbol{w}}(t\mid t+N)=\sum_{j=0}^{N}\boldsymbol{M}_w(j)\boldsymbol{\varepsilon}(t+j),\quad \tilde{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{w}(t)-\hat{\boldsymbol{w}}(t\mid t+N) \tag{3.9.170}$$

将上两式代入(3.9.168)中有

$$\boldsymbol{R}_{\tilde{x}\tilde{w}}=\mathrm{E}\left[\left(\sum_{i=0}^{N}[\boldsymbol{I}_n\delta_{0i},\ -\boldsymbol{K}_i]\begin{bmatrix}\tilde{\boldsymbol{x}}(t\mid t-1)\\ \boldsymbol{\varepsilon}(t+i)\end{bmatrix}\right)\left(\sum_{j=0}^{N}[\boldsymbol{I}_n\delta_{0j},\ -\boldsymbol{M}_w(j)]\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{\varepsilon}(t+j)\end{bmatrix}\right)^{\mathrm{T}}\right] \tag{3.9.171}$$

容易导出关系式

$$\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\boldsymbol{w}^{\mathrm{T}}(t)]=\mathbf{0} \tag{3.9.172}$$

$$\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+j)]=\boldsymbol{\Sigma}(\boldsymbol{\Psi}_p^{\mathrm{T}})^j\boldsymbol{H}^{\mathrm{T}} \tag{3.9.173}$$

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+i)]=\boldsymbol{D}_w(\boldsymbol{\Psi}_p^{\mathrm{T}})^{i-1}\boldsymbol{H}^{\mathrm{T}},\quad i>0,$$

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]=\boldsymbol{S} \tag{3.9.174}$$

将(3.9.172)~(3.9.174)代入(3.9.171)得(3.9.161).

当 $N<0$ 时，$\hat{\boldsymbol{w}}(t|t+N)=\mathbf{0}$，因而由 $\tilde{\boldsymbol{w}}(t|t+N)=\boldsymbol{w}(t)$，故(3.9.166)成为

$$\tilde{\boldsymbol{s}}(t|t+N)=\boldsymbol{H}_s\tilde{\boldsymbol{x}}(t|t+N)+\boldsymbol{C}_0\boldsymbol{H}_e\boldsymbol{w}(t) \tag{3.9.175}$$

注意，$\boldsymbol{w}(t)\perp\tilde{\boldsymbol{x}}(t|t+N)$，故有(3.9.162)成立. □

注意，当 $\boldsymbol{C}_0=\mathbf{0}$ 时，估值误差方差阵的计算大大简化，此时有

$$\boldsymbol{P}_N^s=\boldsymbol{H}_s\boldsymbol{P}_N\boldsymbol{H}_s^{\mathrm{T}}\quad(\boldsymbol{C}_0=\mathbf{0}) \tag{3.9.176}$$

3.9.10 带有色观测噪声系统 Wiener 状态滤波器

用增广状态方法设计带有色观测噪声系统 Kalman 滤波器，要求计算增广状态 Kalman 估值器，通常增广状态维数较大，因而对计算原始状态 Kalman 估值器而言，这增加了计算负担，计算了我们不感兴趣的状态部分分量的 Kalman 估值器，然而这对计算原始状态 Kalman 估值器而言，这又是必须的. 因为 Kalman 估值器对状态分量或部分分量而言是非解耦的. 这里我们用设计 Wiener 状态滤波器方法克服了上述缺点. 将直接给出原始系统状态的 Wiener 估值器，避免了计算其余增广状态分量的估值器.

考虑带有色观测噪声的定常随机系统

$$\boldsymbol{x}_0(t+1)=\boldsymbol{\Phi}_0\boldsymbol{x}_0(t)+\boldsymbol{\Gamma}_0\boldsymbol{w}_0(t) \tag{3.9.177}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}_0\boldsymbol{x}_0(t)+\boldsymbol{\eta}(t)+\boldsymbol{\xi}(t) \tag{3.9.178}$$

$$\boldsymbol{P}(q^{-1})\boldsymbol{\eta}(t+1)=\boldsymbol{R}(q^{-1})\boldsymbol{n}(t) \tag{3.9.179}$$

其中状态 $\boldsymbol{x}_0(t)\in R^n$，观测 $\boldsymbol{y}(t)\in R^m$，$\boldsymbol{\eta}(t)\in R^m$ 是 ARMA 有色观测噪声，$\boldsymbol{w}_0(t)\in R^r$，$\boldsymbol{\xi}(t)\in R^m$ 和 $\boldsymbol{n}(t)\in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_{w0}$，$\boldsymbol{Q}_\xi$ 和 $\boldsymbol{Q}_n$ 的相互独立白噪声，$\boldsymbol{\xi}(t)$是白色观测噪声，$\boldsymbol{H}_0$，$\boldsymbol{\Phi}_0$，$\boldsymbol{\Gamma}_0$ 为适当维常阵，且

$$\boldsymbol{P}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{P}_1q^{-1}+\cdots+\boldsymbol{P}_{n_p}q^{-n_p} \tag{3.9.180}$$

$$\boldsymbol{R}(q^{-1})=\boldsymbol{R}_0+\boldsymbol{R}_1q^{-1}+\cdots+\boldsymbol{R}_{n_p}q^{-n_p} \tag{3.9.181}$$

其中 $\boldsymbol{P}_i$，$\boldsymbol{R}_i$ 为系数阵，$\boldsymbol{R}_0=\boldsymbol{I}_m$ 或 $\boldsymbol{R}_0=\mathbf{0}$. 且设 $n_p\geqslant n_r$. 问题是基于观测($\boldsymbol{y}(t+N)$，$\boldsymbol{y}(t+N-1)$，…)求状态 $\boldsymbol{x}_0(t)$的 Wiener 滤波器 $\hat{\boldsymbol{x}}_0(t|t+N)$，$N=0$，$N>0$ 或 $N<0$.

(3.9.179)有状态空间模型

$$\boldsymbol{\beta}(t+1)=\bar{\boldsymbol{P}}\boldsymbol{\beta}(t)+\bar{\boldsymbol{R}}\boldsymbol{n}(t) \tag{3.9.182}$$

$$\boldsymbol{\eta}(t)=\boldsymbol{H}_1\boldsymbol{\beta}(t)+\boldsymbol{R}_0\boldsymbol{n}(t) \tag{3.9.183}$$

其中 $\boldsymbol{R}_i=\mathbf{0}(i>n_r)$，且

$$\bar{\boldsymbol{P}}=\begin{bmatrix}-\boldsymbol{P}_1 & & \\ \vdots & \boldsymbol{I}_{(n_p-1)m} & \\ -\boldsymbol{P}_{n_p} & \mathbf{0} & \cdots\ \mathbf{0}\end{bmatrix},\quad \bar{\boldsymbol{R}}=\begin{bmatrix}\boldsymbol{R}_1-\boldsymbol{P}_1\boldsymbol{R}_0\\ \vdots\\ \boldsymbol{R}_{n_p}-\boldsymbol{P}_{n_p}\boldsymbol{R}_0\end{bmatrix},\quad \bar{\boldsymbol{H}}_1=[\boldsymbol{I}_m\ \mathbf{0}\cdots\mathbf{0}] \tag{3.9.184}$$

引入增广状态 $\boldsymbol{x}(t)$，增广噪声 $\boldsymbol{w}(t)$和合成观测噪声 $\boldsymbol{v}(t)$为

$$\boldsymbol{x}(t)=\begin{bmatrix}\boldsymbol{x}_0(t)\\ \boldsymbol{\beta}(t)\end{bmatrix},\quad \boldsymbol{w}(t)=\begin{bmatrix}\boldsymbol{w}_0(t)\\ \boldsymbol{n}(t)\end{bmatrix},\quad \boldsymbol{v}(t)=\boldsymbol{R}_0\boldsymbol{n}(t)+\boldsymbol{\xi}(t) \tag{3.9.185}$$

则有增广状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t) \tag{3.9.186}$$

$$\boldsymbol{y}(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t) \tag{3.9.187}$$

其中定义

$$\boldsymbol{\Phi}=\begin{bmatrix}\boldsymbol{\Phi}_0 & \boldsymbol{0}\\ \boldsymbol{0} & \bar{P}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\boldsymbol{\Gamma}_0 & \boldsymbol{0}\\ \boldsymbol{0} & \bar{R}\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{H}_0,\boldsymbol{H}_2] \tag{3.9.188}$$

增广系统是带相关白噪声系统，且有 $\boldsymbol{w}(t)$和 $\boldsymbol{v}(t)$的方差阵 $\boldsymbol{Q}_w$，$\boldsymbol{Q}_v$ 及相关阵 $\boldsymbol{S}$ 为

$$\boldsymbol{Q}_w=\begin{bmatrix}\boldsymbol{Q}_{w0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{Q}_n\end{bmatrix},\quad \boldsymbol{Q}_v=\boldsymbol{R}_0\boldsymbol{Q}_n\boldsymbol{R}_0^{\mathrm{T}}+\boldsymbol{Q}_\xi,\quad \boldsymbol{S}=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{Q}_n\boldsymbol{R}_0^{\mathrm{T}}\end{bmatrix} \tag{3.9.189}$$

对增广系统假设它是完全可观、完全能稳[2]的，则可用有关定理得到它的 Wiener 状态滤波器

$$\psi(q^{-1})\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.190}$$

注意定义(3.9.185)我们有

$$\boldsymbol{x}_0(t)=\boldsymbol{H}_e\boldsymbol{x}(t),\quad \boldsymbol{H}_e=[\boldsymbol{I}_n\quad \boldsymbol{0}] \tag{3.9.191}$$

记 $\boldsymbol{P}_N$ 为 $\boldsymbol{x}(t)$的估值误差 $\tilde{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t\mid t+N)$的方差阵，则有如下定理成立.

【定理 3.9.13】 带有色观测噪声系统(3.9.177)~(3.9.179)有 Wiener 状态滤波器

$$\psi(q^{-1})\hat{\boldsymbol{x}}_0(t\mid t+N)=\boldsymbol{K}_N^0(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.192}$$

其中 $N=0$，$N>0$ 或 $N<0$，且

$$\boldsymbol{K}_N^0(q^{-1})=\boldsymbol{H}_e\boldsymbol{K}_N(q^{-1}) \tag{3.9.193}$$

且滤波误差 $\tilde{\boldsymbol{x}}_0(t\mid t+N)=\boldsymbol{x}_0(t)-\hat{\boldsymbol{x}}_0(t\mid t+N)$方差阵 $\boldsymbol{P}_N^0$为

$$\boldsymbol{P}_N^0=\boldsymbol{H}_e\boldsymbol{P}_N\boldsymbol{H}_e^{\mathrm{T}} \tag{3.9.194}$$

证明 由(3.9.191)有

$$\hat{\boldsymbol{x}}_0(t\mid t+N)=\boldsymbol{H}_e\hat{\boldsymbol{x}}(t\mid t+N) \tag{3.9.195}$$

$$\tilde{\boldsymbol{x}}_0(t\mid t+N)=\boldsymbol{H}_e\tilde{\boldsymbol{x}}(t\mid t+N) \tag{3.9.196}$$

由(3.9.190)有

$$\hat{\boldsymbol{x}}(t\mid t+N)=\frac{\boldsymbol{K}_N(q^{-1})}{\psi(q^{-1})}\boldsymbol{y}(t+N) \tag{3.9.197}$$

将上式代入(3.9.195)得(3.9.192)和(3.9.193). 由(3.9.196)得(3.9.194).

注意，(3.9.194)表明 $\boldsymbol{P}_N^0$是$\boldsymbol{P}_N$的 n 阶主对角线子阵. $\boldsymbol{K}_N^0(q^{-1})$是 $\boldsymbol{K}_N(q^{-1})$的首 n 行子阵. 由 $\boldsymbol{\Psi}_p$ 的稳定性引出 $\psi(q^{-1})=\det(I_n-q^{-1}\boldsymbol{\Psi}_p)$是稳定多项式，因而(3.9.192)是渐近稳定的. 证毕. □

3.9.11 多通道 Wiener 反卷积滤波器

多通道反卷积问题是估计一个动态随机系统的输入信号[20]，经常出现在通信、信号

处理和控制领域.文献[20]用频域多项式方法将问题归结为求解两个耦合的 Diophantine 方程.文献[7][8]用现代时间序列分析方法基于 ARMA 新息模型解决问题.这里基于 Kalman 滤波方法,基于 Riccati 方程设计多通道 Wiener 反卷积滤波器,且给出了滤波器误差方差阵的计算公式.

考虑多通道反卷积问题:

$$\boldsymbol{y}(t)=\boldsymbol{\Pi}^{-1}(q^{-1})\boldsymbol{\Psi}(q^{-1})\boldsymbol{s}(t)+\boldsymbol{L}^{-1}(q^{-1})\boldsymbol{M}(q^{-1})\boldsymbol{n}(t)+\boldsymbol{\xi}(t) \tag{3.9.198}$$

$$\boldsymbol{s}(t)=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{3.9.199}$$

其中 $\boldsymbol{y}(t)\in R^m$ 为观测,$\boldsymbol{s}(t)\in R^n$ 为输入信号,$\boldsymbol{\Pi}(q^{-1})$,$\boldsymbol{\Psi}(q^{-1})$,…,$\boldsymbol{C}(q^{-1})$为单位滞后算子 q^{-1}的多项式矩阵,形如 $\boldsymbol{X}(q^{-1})=\boldsymbol{X}_0+\boldsymbol{X}_1q^{-1}+\cdots+\boldsymbol{X}_{n_x}q^{-n_x}$,$\boldsymbol{\Pi}_0=\boldsymbol{I}_m$,$\boldsymbol{L}_0=\boldsymbol{I}_m$,$\boldsymbol{A}_0=\boldsymbol{I}_n$,$n_a\geqslant n_c$,$n_\pi\geqslant n_\psi$,$n_l\geqslant n_m$.假设 $\boldsymbol{e}(t)\in R^p$,$\boldsymbol{n}(t)\in R^m$ 是零均值的相关白噪声:

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{e}(t)\\ \boldsymbol{n}(t)\end{bmatrix}[\boldsymbol{e}^{\mathrm{T}}(j),\boldsymbol{n}^{\mathrm{T}}(j)]\right\}=\begin{bmatrix}\boldsymbol{Q}_e & \boldsymbol{S}_0\\ \boldsymbol{S}_0^{\mathrm{T}} & \boldsymbol{Q}_n\end{bmatrix}\delta_{tj} \tag{3.9.200}$$

且它们独立于带零均值、方差阵为 $\boldsymbol{Q}_\xi$ 的白噪声$\boldsymbol{\xi}(t)$.

引入左素分解

$$\boldsymbol{P}^{-1}(q^{-1})[\boldsymbol{B}(q^{-1}),\boldsymbol{R}(q^{-1})]=[\boldsymbol{\Pi}^{-1}(q^{-1})\boldsymbol{\Psi}(q^{-1}),\boldsymbol{L}^{-1}(q^{-1})\boldsymbol{M}(q^{-1})] \tag{3.9.201}$$

带 $\boldsymbol{P}_0=\boldsymbol{I}_m$,则系统(3.9.198)和(3.9.199)化为等价系统

$$\boldsymbol{P}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{B}(q^{-1})\boldsymbol{s}(t)+\boldsymbol{R}(q^{-1})\boldsymbol{n}(t)+\boldsymbol{P}(q^{-1})\boldsymbol{\xi}(t) \tag{3.9.202}$$

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{3.9.203}$$

问题是基于观测($\boldsymbol{y}(t+N)$,$\boldsymbol{y}(t+N-1)$,…)求输入 $\boldsymbol{s}(t)$的 Wiener 反卷积滤波器 $\hat{\boldsymbol{s}}(t|t+N)$,$N=0$,$N>0$ 或 $N<0$.这类反卷积问题如图 3.9.5 所示.

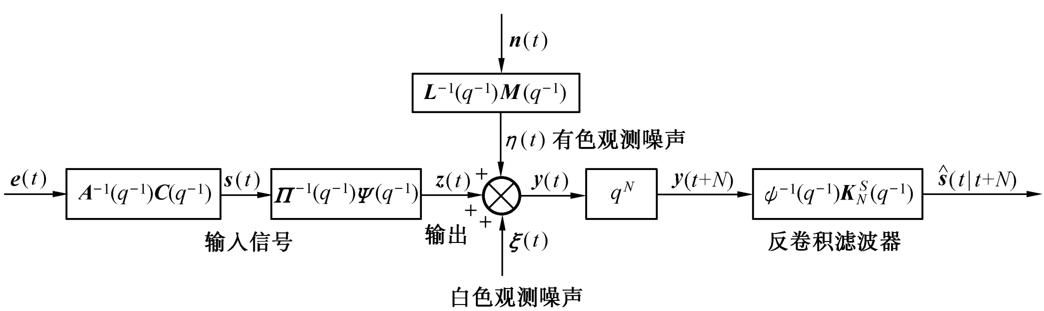


图 3.9.5　Wiener 反卷积滤波问题

【例 3.9.5】　考虑雷达跟踪系统

$$x(t+1)=x(t)+s(t)T+e(t)T^2/2 \tag{3.9.204}$$

$$s(t+1)=s(t)+e(t)T \tag{3.9.205}$$

$$y(t)=x(t)+\xi(t) \tag{3.9.206}$$

其中 T 为采样周期,$x(t)$,$s(t)$和 $e(t)$为在时刻 tT 运动目标的位置、速度和加速度,$y(t)$为对位置 $x(t)$的观测信号,$\xi(t)$为白色观测噪声.设 $e(t)$和 $\xi(t)$是零均值、方差各为 σ_e^2 和 σ_ξ^2 的独立白噪声.问题是求速度 $s(t)$的 Wiener 跟踪滤波器 $\hat{s}(t|t+N)$.

容易验证该问题等价于如下单通道 Wiener 反卷积滤波问题,即

$$y(t)=\frac{Tq^{-1}}{1-q^{-1}}s(t)+\frac{T^2q^{-1}}{2(1-q^{-1})}n(t)+\xi(t) \tag{3.9.207}$$

$$(1-q^{-1})s(t)=Tq^{-1}e(t) \tag{3.9.208}$$

$$n(t)=e(t) \tag{3.9.209}$$

因为 $n(t)=e(t)$，故有 $n(t)$与 $e(t)$是相关白噪声.

$$\mathrm{E}\left\{\begin{bmatrix} e(t) \\ n(t) \end{bmatrix}[e(j),n(j)]\right\}=\begin{bmatrix} \sigma_e^2 & \sigma_e^2 \\ \sigma_e^2 & \sigma_e^2 \end{bmatrix}\delta_{tj} \tag{3.9.210}$$

即相关系数 s_0 为

$$s_0=\mathrm{E}[e(t)n(t)]=\sigma_e^2 \tag{3.9.211}$$

由本例看到，在某些应用问题中假设相关噪声(3.9.200)是有意义的.

下面解决一般多通道 Wiener 反卷积问题.(3.9.203)有状态空间模型

$$\boldsymbol{\alpha}(t+1)=\bar{\boldsymbol{A}}\boldsymbol{\alpha}(t)+\bar{\boldsymbol{C}}\boldsymbol{e}(t) \tag{3.9.212}$$

$$\boldsymbol{s}(t)=\bar{\boldsymbol{H}}_1\boldsymbol{\alpha}(t)+\boldsymbol{C}_0\boldsymbol{e}(t) \tag{3.9.213}$$

(3.9.202)有状态空间模型

$$\boldsymbol{\beta}(t+1)=\bar{\boldsymbol{P}}\boldsymbol{\beta}(t)+\bar{\boldsymbol{B}}\boldsymbol{s}(t)+\bar{\boldsymbol{R}}\boldsymbol{n}(t) \tag{3.9.214}$$

$$\boldsymbol{y}(t)=\bar{\boldsymbol{H}}_2\boldsymbol{\beta}(t)+\boldsymbol{B}_0\boldsymbol{s}(t)+\boldsymbol{R}_0\boldsymbol{n}(t)+\boldsymbol{\xi}(t) \tag{3.9.215}$$

其中规定 $\boldsymbol{B}_i=\boldsymbol{0}(i>n_b)$，$\boldsymbol{R}_i=\boldsymbol{0}(i>n_r)$，

$$\bar{\boldsymbol{A}}=\begin{bmatrix} -\boldsymbol{A}_1 & & \\ \vdots & \boldsymbol{I}_{(n_a-1)m} & \\ -\boldsymbol{A}_{n_a} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix},\quad \bar{\boldsymbol{C}}=\begin{bmatrix} \boldsymbol{C}_1-\boldsymbol{A}_1\boldsymbol{C}_0 \\ \vdots \\ \boldsymbol{C}_{n_a}-\boldsymbol{A}_{n_a}\boldsymbol{C}_0 \end{bmatrix},\quad \bar{\boldsymbol{H}}_1=[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}] \tag{3.9.216}$$

$$\bar{\boldsymbol{P}}=\begin{bmatrix} -\boldsymbol{P}_1 & & \\ \vdots & \boldsymbol{I}_{(n_p-1)m} & \\ -\boldsymbol{P}_{n_p} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix},\quad \bar{\boldsymbol{R}}=\begin{bmatrix} \boldsymbol{R}_1-\boldsymbol{P}_1\boldsymbol{R}_0 \\ \vdots \\ \boldsymbol{R}_{n_p}-\boldsymbol{P}_{n_p}\boldsymbol{R}_0 \end{bmatrix},\quad \bar{\boldsymbol{H}}_2=[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}] \tag{3.9.217}$$

合并上述两个状态空间模型有增广系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}\boldsymbol{x}(t)+\boldsymbol{\Gamma}\boldsymbol{w}(t) \tag{3.9.218}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{3.9.219}$$

其中定义

$$\boldsymbol{x}(t)=\begin{bmatrix} \boldsymbol{\alpha}(t) \\ \boldsymbol{\beta}(t) \end{bmatrix},\quad \boldsymbol{w}(t)=\begin{bmatrix} \boldsymbol{e}(t) \\ \boldsymbol{n}(t) \end{bmatrix},\quad \boldsymbol{v}(t)=\boldsymbol{B}_0\boldsymbol{C}_0\boldsymbol{e}(t)+\boldsymbol{R}_0\boldsymbol{n}(t)+\boldsymbol{\xi}(t) \tag{3.9.220}$$

$$\boldsymbol{\Phi}=\begin{bmatrix} \bar{\boldsymbol{A}} & \boldsymbol{0} \\ \bar{\boldsymbol{B}}\bar{\boldsymbol{H}}_1 & \bar{\boldsymbol{P}} \end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix} \bar{\boldsymbol{C}} & \boldsymbol{0} \\ \bar{\boldsymbol{B}}\boldsymbol{C}_0 & \bar{\boldsymbol{R}} \end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{B}_0\bar{\boldsymbol{H}}_1,\bar{\boldsymbol{H}}_2] \tag{3.9.221}$$

显然，$\boldsymbol{w}(t)$和 $\boldsymbol{v}(t)$是带零均值的相关白噪声，且有方差阵 $\boldsymbol{Q}_w$，$\boldsymbol{Q}_v$ 和相关阵 $\boldsymbol{S}_0$ 分别为

$$\boldsymbol{Q}_w=\begin{bmatrix} \boldsymbol{Q}_e & \boldsymbol{S}_0 \\ \boldsymbol{S}_0^{\mathrm{T}} & \boldsymbol{Q}_n \end{bmatrix},\quad \boldsymbol{Q}_v=\boldsymbol{B}_0\boldsymbol{C}_0\boldsymbol{Q}_e\boldsymbol{C}_0^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}+\boldsymbol{R}_0\boldsymbol{Q}_n\boldsymbol{R}_0^{\mathrm{T}}+\boldsymbol{Q}_\xi+\boldsymbol{B}_0\boldsymbol{C}_0\boldsymbol{S}_0\boldsymbol{R}_0^{\mathrm{T}}+\boldsymbol{R}_0\boldsymbol{S}_0^{\mathrm{T}}\boldsymbol{C}_0^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}},$$

$$\boldsymbol{S}=\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(t)]=\begin{bmatrix} \boldsymbol{Q}_e\boldsymbol{C}_0^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}+\boldsymbol{S}_0\boldsymbol{R}_0^{\mathrm{T}} \\ \boldsymbol{S}_0\boldsymbol{C}_0^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}}+\boldsymbol{Q}_n\boldsymbol{R}_0^{\mathrm{T}} \end{bmatrix} \tag{3.9.222}$$

假设增广系统是完全可观、完全能稳的，则存在稳态 Kalman 预报器，在此基础上基于关于

稳态 Kalman 预报器误差方差阵 $\boldsymbol{\Sigma}$ 的 Riccati 方程可求得 Wiener 状态滤波器

$$\psi(q^{-1})\hat{\boldsymbol{x}}(t|t+N)=\boldsymbol{K}_N(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.223}$$

且可求得相应的稳态估值误差方差阵 $\boldsymbol{P}_N$，还可求得白噪声 $\boldsymbol{w}(t)$ 的 Wiener 滤波器

$$\psi(q^{-1})\hat{\boldsymbol{w}}(t|t+N)=\boldsymbol{K}_N^w(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.224}$$

及其他的稳态估值误差方差阵 $\boldsymbol{P}_N^w$.

【定理 3.9.14】 多通道反卷积系统(3.9.198)和(3.9.199)有 Wiener 反卷积滤波器

$$\psi(q^{-1})\hat{\boldsymbol{s}}(t|t+N)=\boldsymbol{K}_N^s(q^{-1})\boldsymbol{y}(t+N) \tag{3.9.225}$$

$$\boldsymbol{K}_N^s(q^{-1})=\boldsymbol{H}_s\boldsymbol{K}_N(q^{-1})+\boldsymbol{C}_0\boldsymbol{H}_e\boldsymbol{K}_N^w(q^{-1}) \tag{3.9.226}$$

$$\boldsymbol{H}_s=[\boldsymbol{I}_m\boldsymbol{0}\cdots\boldsymbol{0}],\quad \boldsymbol{H}_e=[\boldsymbol{I}_r\quad \boldsymbol{0}] \tag{3.9.227}$$

且有稳态估值误差方差阵 $\boldsymbol{P}_N^s$为(3.9.160)~(3.9.162).

证明 类似于定理 3.9.12，从略. □

【注 3.9.1】 在单通道情形($m=1$)Wiener 去卷问题为

$$y(t)=\frac{\Psi(q^{-1})}{\Pi(q^{-1})}s(t)+\frac{M(q^{-1})}{L(q^{-1})}n(t)+\xi(t) \tag{3.9.228}$$

$$A(q^{-1})s(t)=C(q^{-1})e(t) \tag{3.9.229}$$

引入多项式互质分解

$$\frac{1}{P(q^{-1})}[B(q^{-1}),R(q^{-1})]=\left[\frac{\Psi(q^{-1})}{\Pi(q^{-1})},\frac{M(q^{-1})}{L(q^{-1})}\right] \tag{3.9.230}$$

则 Wiener 去卷问题化为

$$P(q^{-1})y(t)=B(q^{-1})s(t)+R(q^{-1})n(t)+P(q^{-1})\xi(t) \tag{3.9.231}$$

$$A(q^{-1})s(t)=C(q^{-1})e(t) \tag{3.9.232}$$

它等价于如下 Wiener 去卷问题

$$P(q^{-1})z(t)=B(q^{-1})s(t)+R(q^{-1})n(t) \tag{3.9.233}$$

$$y(t)=z(t)+\xi(t) \tag{3.9.234}$$

$$A(q^{-1})s(t)=C(q^{-1})e(t) \tag{3.9.235}$$

这里 $z(t)$表示通道输出信号，而 $y(t)$表示观测信号，$\xi(t)$为白色观测噪声. 而 $B(q^{-1})/P(q^{-1})$为通道传递函数，$\eta(t)=[R(q^{-1})/P(q^{-1})]n(t)$为有色观测噪声. 这可由(3.9.233)的如下等价表示看出：

$$z(t)=\frac{B(q^{-1})}{P(q^{-1})}s(t)+\frac{P(q^{-1})}{P(q^{-1})}n(t) \tag{3.9.236}$$

特别当 $P(q^{-1})=B(q^{-1})$时，Wiener 反卷积问题化为 Wiener 滤波问题：

$$y(t)=s(t)+\frac{R(q^{-1})}{P(q^{-1})}n(t)+\xi(t) \tag{3.9.237}$$

$$A(q^{-1})s(t)=C(q^{-1})e(t) \tag{3.9.238}$$

或等价地，它等价于 Wiener 滤波问题

$$y(t)=s(t)+\eta(t)+\xi(t) \tag{3.9.239}$$

$$A(q^{-1})s(t)=C(q^{-1})e(t) \tag{3.9.240}$$

$$P(q^{-1})\eta(t)=R(q^{-1})n(t) \tag{3.9.241}$$

这里 $\eta(t)$代表有色观测噪声.

3.9.12 带有色观测噪声的单通道白噪声 Wiener 反卷积滤波器

在石油地震勘探中[10,11]最简单的情形归结为估计带白色观测噪声的单通道动态随机系统的输入白噪声 $\mu(t)$，即白噪声反卷积模型为

$$A(q^{-1})s(t)=C(q^{-1})\mu(t) \tag{3.9.242}$$

$$y(t)=s(t)+\xi(t) \tag{3.9.243}$$

$$A(q^{-1})=1+a_1q^{-1}+\cdots+a_{n_a}q^{-n_a},\quad C(q^{-1})=c_1q^{-1}+\cdots+c_{n_c}q^{-n_c} \tag{3.9.244}$$

假设系统还带有色观测噪声 $\eta(t)$，

$$P(q^{-1})\eta(t)=R(q^{-1})n(t) \tag{3.9.245}$$

$$P(q^{-1})=1+p_1q^{-1}+\cdots+p_{n_p}q^{-n_p},\quad R(q^{-1})=r_0+r_1q^{-1}+\cdots+r_{n_r}q^{-n_r} \tag{3.9.246}$$

因而观测方程成为

$$y(t)=s(t)+\eta(t)+\xi(t) \tag{3.9.247}$$

这里假设 $\mu(t)$，$\xi(t)$，$\eta(t)$是零均值、方差各为 σ_μ^2，σ_ξ^2 和 σ_n^2 的独立白噪声. 于是带白色和有色观测噪声的反卷积模型成为(2.9.247)，(2.9.242)和(2.9.245)，即

$$y(t)=\frac{C(q^{-1})}{A(q^{-1})}\mu(t)+\frac{R(q^{-1})}{P(q^{-1})}n(t)+\xi(t) \tag{3.9.248}$$

问题是基于观测$(y(t+N),y(t+N-1),\cdots)$求输入白噪声 $\mu(t)$的 Wiener 滤波器 $\hat{\mu}(t|t+N)$，$N=0$，$N>0$ 或 $N<0$. 这类白噪声反卷积问题如图 3.9.6 所示.

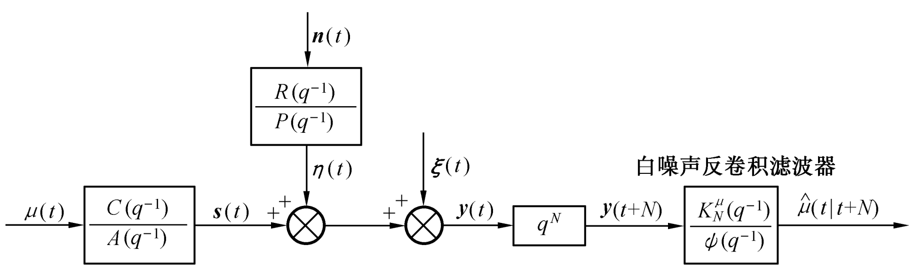


图 3.9.6 白噪声 $\mu(t)$的 Wiener 反卷积问题

(3.9.242)有状态空间模型

$$\boldsymbol{\alpha}(t+1)=\bar{\boldsymbol{A}}\boldsymbol{\alpha}(t)+\bar{\boldsymbol{C}}\boldsymbol{\mu}(t) \tag{3.9.249}$$

$$\boldsymbol{s}(t)=\boldsymbol{H}_1\boldsymbol{\alpha}(t) \tag{3.9.250}$$

其中规定 $c_i=0(i>n_c)$，且

$$\bar{\boldsymbol{A}}=\begin{bmatrix}-a_1 & & \\ \vdots & \boldsymbol{I}_{n_a-1} & \\ -a_{n_a} & 0 & \cdots & 0\end{bmatrix},\quad \bar{\boldsymbol{C}}=\begin{bmatrix}C_1\\ \vdots\\ C_{n_a}\end{bmatrix},\quad \boldsymbol{H}_1=[1\quad 0\cdots0] \tag{3.9.251}$$

(3.9.245)有状态空间模型

$$\boldsymbol{\beta}(t+1)=\bar{\boldsymbol{P}}\boldsymbol{\beta}(t)+\bar{\boldsymbol{R}}n(t) \tag{3.9.252}$$

$$\eta(t)=\boldsymbol{H}_2\boldsymbol{\beta}(t)+r_0n(t) \tag{3.9.253}$$

其中规定 $r_i=0(i>n_r)$，且

$$\bar{\boldsymbol{P}}=\begin{bmatrix} -p_1 & & \\ \vdots & \boldsymbol{I}_{n_p-1} & \\ -p_{n_p} & 0 & \cdots & 0\end{bmatrix},\quad \bar{\boldsymbol{R}}=\begin{bmatrix} r_1-p_1r_0 \\ \vdots \\ r_{n_p}-p_{n_p}r_0\end{bmatrix},\quad \boldsymbol{H}_2=[1\quad 0\cdots 0] \tag{3.9.254}$$

引入增广状态 $\boldsymbol{x}(t)$，增广输入噪声 $\boldsymbol{w}(t)$和合成观测噪声 $\boldsymbol{v}(t)$为

$$\boldsymbol{x}(t)=\begin{bmatrix}\alpha(t)\\ \beta(t)\end{bmatrix},\quad \boldsymbol{w}(t)=\begin{bmatrix}\mu(t)\\ n(t)\end{bmatrix},\quad v(t)=r_0n(t)+\xi(t) \tag{3.9.255}$$

则有增广系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t) \tag{3.9.256}$$

$$y(t)=\boldsymbol{Hx}(t)+v(t) \tag{3.9.257}$$

其中

$$\boldsymbol{\Phi}=\begin{bmatrix}\bar{A} & \boldsymbol{0}\\ \boldsymbol{0} & \bar{\boldsymbol{P}}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\bar{\boldsymbol{C}} & \boldsymbol{0}\\ \boldsymbol{0} & \bar{\boldsymbol{R}}\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{H}_1,\boldsymbol{H}_2] \tag{3.9.258}$$

注意增广系统中 $\boldsymbol{w}(t)$和 $v(t)$是带零均值、方差阵各为 $\boldsymbol{Q}_w$ 和σ_v^2，相关阵为 $\boldsymbol{S}$ 的相关白噪声，且有

$$\boldsymbol{Q}_w=\begin{bmatrix}\sigma_\mu^2 & 0\\ 0 & \sigma_n^2\end{bmatrix},\quad \sigma_v^2=r_0^2\sigma_n^2+\sigma_\xi^2,\quad \boldsymbol{S}=\mathrm{E}[\boldsymbol{w}(t)v(t)]=\begin{bmatrix}0\\ r_0\sigma_n^2\end{bmatrix} \tag{3.9.259}$$

假设增广系统完全可观、完全能稳[2]，则它存在稳态 Kalman 预报器.故基于 Riccati 方程可求得白噪声 Wiener 滤波器 $\hat{w}(t\mid t+N)$为

$$\psi(q^{-1})\hat{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{K}_N^w(q^{-1})y(t+N) \tag{3.9.260}$$

且可求得相应的估值误差方差阵 $\boldsymbol{P}_N^w=\mathrm{E}[(\boldsymbol{w}(t)-\hat{\boldsymbol{w}}(t\mid t+N))(\boldsymbol{w}(t)-\hat{\boldsymbol{w}}(t\mid t+N))^{\mathrm{T}}]$.

【定理 3.9.15】 带有色和白色观测噪声的单通道白噪声反卷积系统(3.9.248)有白噪声 Wiener 反卷积滤波器为

$$\psi(q^{-1})\hat{\mu}(t\mid t+N)=K_N^\mu(q^{-1})y(t+N) \tag{3.9.261}$$

$$K_N^\mu(q^{-1})=[1\quad 0]\boldsymbol{K}_N^w(q^{-1}) \tag{3.9.262}$$

且估值误差 $\tilde{\mu}(t\mid t+N)=\mu(t)-\hat{\mu}(t\mid t+N)$的方差 $P_N^\mu=\mathrm{E}[\tilde{\mu}^2(t\mid t+N)]$为

$$P_N^\mu=[1\quad 0]\boldsymbol{P}_N^w[1\quad 0]^{\mathrm{T}} \tag{3.9.263}$$

即 P_N^μ为 $\boldsymbol{P}_N^w$的第(1,1)元素.

证明 注意 $\mu(t)=[1\quad 0]\boldsymbol{w}(t)$，这引出(3.9.261)~(3.9.263). □

3.10 统一的和通用的 Kalman 滤波理论和白噪声估计理论

经典 Kalman 滤波理论[2~4]限于处理状态噪声和观测噪声均为白噪声，且它们在相同时刻是相关的系统，这种系统不能满足理论和应用的要求.虽然用增广状态方法可处理带有色噪声系统，但由于状态维数增加却大大增加了计算负担.为了避免增广状态，文献[21]直接用射影理论给出了一种推广的 Kalman 滤波器，它可处理系统噪声在相邻时刻是相关的，且系统噪声和观测噪声在相同时刻和相邻时刻是相关的随机系统.本节对这类系统用射影理论不仅给出了不同于文献[21]的 Kalman 滤波器和预报器新算法，而且还提出

了最优 Kalman 平滑器，且进一步提出了相应的稳态 Kalman 滤波器、预报器和平滑器，构成了统一的和通用的 Kalman 滤波理论[14]．在此基础上，文献[15]提出了统一和通用的白噪声估计理论，推广了文献[8，12，13]的结果．

3.10.1　统一和通用的 Kalman 滤波器和预报器

考虑一般的线性离散时变随机系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}(t)\boldsymbol{x}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t) \tag{3.10.1}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}(t)\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{3.10.2}$$

其中状态 $\boldsymbol{x}(t)\in R^n$，观测 $\boldsymbol{y}(t)\in R^m$，$\boldsymbol{\Phi}(t)$，$\boldsymbol{\Gamma}(t)$和 $\boldsymbol{H}(t)$是已知时变矩阵，系统噪声 $\boldsymbol{w}(t)\in R^r$ 是带零均值的有色噪声，它在相邻时刻是相关的，观测噪声 $\boldsymbol{v}(t)\in R^m$ 是带零均值白噪声，它在相同时刻和相邻时刻与 $w(t)$是相关的，即

$$\begin{gathered}
\mathrm{E}\boldsymbol{w}(t)=\boldsymbol{0},\quad \mathrm{E}\boldsymbol{v}(t)=\boldsymbol{0}\\
\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(t)]=\boldsymbol{Q}_w(t),\quad \mathrm{E}[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(t+1)]=\boldsymbol{Q}_{w1}(t),\\
\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(t+j)]=\boldsymbol{0}\quad (j\geqslant 2),\\
\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{v}^{\mathrm{T}}(j)]=\boldsymbol{Q}_v(t)\delta_{tj},\quad \delta_{tt}=1,\quad \delta_{tj}=0\,(t\neq j),\\
\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(t)]=\boldsymbol{S}(t),\quad \mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(t+1)]=\boldsymbol{S}_1(t),\\
\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(j)]=\boldsymbol{0},\quad \mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(t-j)]=\boldsymbol{0}\,(j\geqslant 1)
\end{gathered} \tag{3.10.3}$$

其中 E 为均值号，T 为转置号．初始观测时刻 $t_0=0$，且

$$\mathrm{E}\boldsymbol{x}(0)=\boldsymbol{\mu},\quad \mathrm{cov}\boldsymbol{x}(0)=\boldsymbol{P}_0 \tag{3.10.4}$$

其中 cov 为协方差号，且设 $\boldsymbol{x}(0)$不相关于 $\boldsymbol{w}(t)$和 $\boldsymbol{v}(t)$．

最优 Kalman 滤波问题是：基于观测$(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots \boldsymbol{y}(1))$求状态 $\boldsymbol{x}(t)$的最优(线性最小方差)估值器 $\hat{\boldsymbol{x}}(t\mid t+N)$．对 $N=0$，$N>0$ 或 $N<0$，分别称其为最优 Kalman 滤波器、平滑器或预报器．

【例 3.10.1】　一个纯量相邻时刻相关噪声的例子．

设 $g(t)$和 $\xi(t)$是零均值、方差各为 σ_g^2 和σ_ξ^2 的相互独立白噪声，构造

$$w(t)=g(t)+\beta g(t-1) \tag{3.10.5}$$

$$v(t)=\alpha g(t-1)+\xi(t) \tag{3.10.6}$$

其中 α,β 为常数，则有

$$\begin{gathered}
\mathrm{E}w(t)=0,\quad \sigma_w^2=\mathrm{E}[w^2(t)]=(1+\beta^2)\sigma_g^2,\\
\sigma_{w1}^2=\mathrm{E}[w(t)w(t+1)]=\beta\sigma_g^2,\\
\mathrm{E}[w(t)w(t+j)]=0,\quad j\geqslant 2,\\
\mathrm{E}v(t)=0,\quad \sigma_v^2=\mathrm{E}[v^2(t)]=\alpha^2\sigma_g^2+\sigma_\xi^2,\\
S=\mathrm{E}[w(t)v(t)]=\alpha\beta\sigma_g^2,\\
S_1=\mathrm{E}[w(t)v(t+1)]=\alpha\sigma_g^2,\\
\mathrm{E}[w(t)v(t+j)]=0\quad (j\geqslant 2),\\
\mathrm{E}[w(t)v(t-j)]=0\quad (j\geqslant 1)
\end{gathered} \tag{3.10.7}$$

注意在假设(3.10.3)中，$\boldsymbol{w}(t)$在相邻时刻的相关性是双向的，因为

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(t-j)]=\{\mathrm{E}[\boldsymbol{w}(t-j)\boldsymbol{w}^{\mathrm{T}}(t)]\}^{\mathrm{T}},\quad j>0 \tag{3.10.8}$$

而 $\boldsymbol{w}(t)$ 与 $\boldsymbol{v}(t)$ 在相邻时刻的相关性是单向的.

应用射影性质和(3.10.3)有如下重要关系

$$L(\boldsymbol{y}(t),\cdots,\boldsymbol{y}(1))=L(\boldsymbol{\varepsilon}(t),\cdots,\boldsymbol{\varepsilon}(1)),$$
$$\boldsymbol{\varepsilon}(t)\in L(\boldsymbol{y}(t),\cdots,\boldsymbol{y}(1)),$$
$$\boldsymbol{\varepsilon}(t)\perp L(\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(1)),$$
$$\boldsymbol{v}(t)\perp L(\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(1)),$$
$$\boldsymbol{w}(t-1)\perp L(\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(1)),$$
$$\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{y}}(t|t-1)=\boldsymbol{y}(t)-\boldsymbol{H}(t)\hat{\boldsymbol{x}}(t|t-1) \tag{3.10.9}$$

其中 $\boldsymbol{\varepsilon}(t)$ 是 $\boldsymbol{y}(t)$ 的新息过程.应用(3.10.1)引出

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t) \tag{3.10.10}$$

其中 $\tilde{\boldsymbol{x}}(t|t-1)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t-1)$.为了计算 $\boldsymbol{\varepsilon}(t)$ 的方差阵,注意由(3.10.1)有

$$\begin{aligned}&\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\boldsymbol{v}^{\mathrm{T}}(t)]=\mathrm{E}[(\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t-1))\boldsymbol{v}^{\mathrm{T}}(t)]=\\&\mathrm{E}\{[\boldsymbol{\Phi}(t-1)\boldsymbol{x}(t-1)+\boldsymbol{\Gamma}(t-1)\boldsymbol{w}(t-1)-\hat{\boldsymbol{x}}(t|t-1)]\boldsymbol{v}^{\mathrm{T}}(t)\}=\\&\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)\end{aligned} \tag{3.10.11}$$

其中应用了事实 $\boldsymbol{v}(t)\perp\boldsymbol{x}(t-1)$. $\boldsymbol{v}(t)\perp\hat{\boldsymbol{x}}(t|t-1)$.这是因为 $\boldsymbol{x}(t-1)\in L(\boldsymbol{w}(t-2),\cdots$ $\boldsymbol{w}(0),\boldsymbol{x}(0))$,$\hat{\boldsymbol{x}}(t|t-1)\in L(\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(1))$.于是有新息方差 $\boldsymbol{Q}_\varepsilon(t)$ 为

$$\begin{aligned}&\boldsymbol{Q}_\varepsilon(t)=\mathrm{E}[\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]=\mathrm{E}[(\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t))(\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t))^{\mathrm{T}}]=\\&\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{Q}_v(t)+\boldsymbol{H}(t)\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)+\boldsymbol{S}_1^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{H}^{\mathrm{T}}(t)\end{aligned} \tag{3.10.12}$$

其中记 $\boldsymbol{P}(t|t-1)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1)]$.由射影正交性有

$$\mathrm{E}[\hat{\boldsymbol{x}}(t|t-1)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1)]=\boldsymbol{0} \tag{3.10.13}$$

而注意应用(3.10.11)和(3.10.13)有

$$\begin{aligned}&\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]=\mathrm{E}[(\hat{\boldsymbol{x}}(t|t-1)+\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1))(\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t))^{\mathrm{T}}]=\\&\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)\end{aligned} \tag{3.10.14}$$

由射影性质有 Kalman 滤波器

$$\begin{aligned}&\hat{\boldsymbol{x}}(t|t)=\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}(t)\boldsymbol{\varepsilon}(t),\\&\boldsymbol{K}(t)=\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]\boldsymbol{Q}_\varepsilon^{-1}(t)=[\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)]\boldsymbol{Q}_\varepsilon^{-1}(t)\end{aligned} \tag{3.10.15}$$

在(3.10.1)中置 $t+1=t$ 后两边取射影运算有

$$\hat{\boldsymbol{x}}(t|t-1)=\boldsymbol{\Phi}(t-1)\hat{\boldsymbol{x}}(t-1|t-1)+\boldsymbol{\Gamma}(t-1)\hat{\boldsymbol{w}}(t-1|t-1) \tag{3.10.16}$$

由递推射影公式有白噪声滤波器

$$\hat{\boldsymbol{w}}(t-1|t-1)=\hat{\boldsymbol{w}}(t-1|t-2)+\mathrm{E}[\boldsymbol{w}(t-1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1)]\boldsymbol{Q}_\varepsilon^{-1}(t-1)\boldsymbol{\varepsilon}(t-1) \tag{3.10.17}$$

由(3.10.3)$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(t-j)]=0\ (j\geqslant 1)$引出 $\boldsymbol{w}(t-1)\perp L(\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(1))$.这引出

$$\hat{\boldsymbol{w}}(t-1|t-2)=\boldsymbol{0} \tag{3.10.18}$$

注意　$\boldsymbol{w}(t-1)\perp\boldsymbol{x}(t-2)$,故有

$$\begin{aligned}&\mathrm{E}[\boldsymbol{w}(t-1)\boldsymbol{x}^{\mathrm{T}}(t-1)]=\mathrm{E}[\boldsymbol{w}(t-1)(\boldsymbol{\Phi}(t-2)\boldsymbol{x}(t-2)+\\&\boldsymbol{\Gamma}(t-2)\boldsymbol{w}(t-2))^{\mathrm{T}}]=\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-2)\boldsymbol{\Gamma}^{\mathrm{T}}(t-2)\end{aligned} \tag{3.10.19}$$

且注意 $\boldsymbol{w}(t-1)\perp\hat{\boldsymbol{x}}(t-1|t-2)$，则有

$$\mathrm{E}[\boldsymbol{w}(t-1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1)]=\mathrm{E}[\boldsymbol{w}(t-1)(\boldsymbol{H}(t-1)\boldsymbol{x}(t-1)+\boldsymbol{v}(t-1)-\boldsymbol{H}(t-1)\hat{\boldsymbol{x}}(t-1|t-2))^{\mathrm{T}}]=\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-2)\boldsymbol{\Gamma}^{\mathrm{T}}(t-2)\boldsymbol{H}^{\mathrm{T}}(t-1)+\boldsymbol{S}(t-1) \tag{3.10.20}$$

这引出

$$\hat{\boldsymbol{w}}(t-1|t-1)=[\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-2)\boldsymbol{\Gamma}^{\mathrm{T}}(t-2)\boldsymbol{H}^{\mathrm{T}}(t-1)+\boldsymbol{S}(t-1)]\boldsymbol{Q}_{\varepsilon}^{-1}(t-1)\boldsymbol{\varepsilon}(t-1) \tag{3.10.21}$$

故由(3.10.16)有

$$\hat{\boldsymbol{x}}(t|t-1)=\boldsymbol{\Phi}(t-1)\hat{\boldsymbol{x}}(t-1|t-1)+\boldsymbol{K}_1(t-1)\boldsymbol{\varepsilon}(t-1) \tag{3.10.22}$$

$$\boldsymbol{K}_1(t-1)=\boldsymbol{\Gamma}(t-1)[\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-2)\boldsymbol{\Gamma}^{\mathrm{T}}(t-2)\boldsymbol{H}^{\mathrm{T}}(t-1)+\boldsymbol{S}(t-1)]\boldsymbol{Q}_{\varepsilon}^{-1}(t-1) \tag{3.10.23}$$

由(3.10.1)减(3.10.22)引出

$$\tilde{\boldsymbol{x}}(t|t-1)=\boldsymbol{\Phi}(t-1)\tilde{\boldsymbol{x}}(t-1|t-1)+\boldsymbol{\Gamma}(t-1)\boldsymbol{w}(t-1)-\boldsymbol{K}_2(t-1)\boldsymbol{\varepsilon}(t-1) \tag{3.10.24}$$

这引出预报误差方差阵 $\boldsymbol{P}(t|t-1)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1)]$为

$$\boldsymbol{P}(t|t-1)=[\boldsymbol{\Phi}(t-1),\boldsymbol{\Gamma}(t-1),-\boldsymbol{K}_1(t-1)]\times \mathrm{E}\left\{\begin{bmatrix}\tilde{\boldsymbol{x}}(t-1|t-1)\\ \boldsymbol{w}(t-1)\\ \boldsymbol{\varepsilon}(t-1)\end{bmatrix}[\tilde{\boldsymbol{x}}^{\mathrm{T}}(t-1|t-1),\boldsymbol{w}^{\mathrm{T}}(t-1),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1)]\right\}\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}(t-1)\\ \boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\\ -\boldsymbol{K}_1^{\mathrm{T}}(t-1)\end{bmatrix} \tag{3.10.25}$$

由(3.10.19)和(3.10.20)，并注意 $\hat{\boldsymbol{x}}(t-1|t-2)\perp\boldsymbol{w}(t-1)$，则有

$$\begin{aligned}&\mathrm{E}[\tilde{\boldsymbol{x}}(t-1|t-1)\boldsymbol{w}^{\mathrm{T}}(t-1)]=\\ &\mathrm{E}[(\boldsymbol{x}(t-1)-\hat{\boldsymbol{x}}(\boldsymbol{t}-1|\boldsymbol{t}-2)-\boldsymbol{K}(t-1)\boldsymbol{\varepsilon}(t-1))\boldsymbol{w}^{\mathrm{T}}(t-1)]=\\ &\mathrm{E}[\boldsymbol{x}(t-1)\boldsymbol{w}^{\mathrm{T}}(t-1)]-\boldsymbol{K}(t-1)\mathrm{E}[\boldsymbol{\varepsilon}(t-1)\boldsymbol{w}^{\mathrm{T}}(t-1)]=\\ &\boldsymbol{\Gamma}(t-2)\boldsymbol{Q}_{w1}(t-2)-\boldsymbol{K}(t-1)[\boldsymbol{H}(t-1)\boldsymbol{\Gamma}(t-2)\boldsymbol{Q}_{w1}(t-2)+\boldsymbol{S}^{\mathrm{T}}(t-1)]=\\ &[\boldsymbol{I}_n-\boldsymbol{K}(t-1)\boldsymbol{H}(t-1)]\boldsymbol{\Gamma}(t-2)\boldsymbol{Q}_{w1}(t-2)-\boldsymbol{K}(t-1)\boldsymbol{S}^{\mathrm{T}}(t-1)\end{aligned} \tag{3.10.26}$$

因 $\hat{\boldsymbol{x}}(t-1|t-2)\in L(\boldsymbol{y}(t-2),\cdots\boldsymbol{y}(1))=L(\boldsymbol{\varepsilon}(t-2),\cdots,\boldsymbol{\varepsilon}(1))$，故 $\hat{\boldsymbol{x}}(t-1|t-2)\perp\boldsymbol{\varepsilon}(t-1)$. 于是应用(3.10.15)有

$$\begin{aligned}&\mathrm{E}[\tilde{\boldsymbol{x}}(t-1|t-1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1)]=\\ &\mathrm{E}[(\boldsymbol{x}(t-1)-\hat{\boldsymbol{x}}(t-1|t-2)-\boldsymbol{K}(t-1)\boldsymbol{\varepsilon}(t-1))\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1)]=\\ &\mathrm{E}[\boldsymbol{x}(t-1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1)]-\boldsymbol{K}(t-1)\boldsymbol{Q}_{\varepsilon}(t-1)=\boldsymbol{0}\end{aligned} \tag{3.10.27}$$

将(3.10.20)，(3.10.26)，(3.10.27)代入(3.10.25)可得

$$\boldsymbol{P}(t|t-1)=[\boldsymbol{\Phi}(t-1),\boldsymbol{\Gamma}(t-1),-\boldsymbol{K}_1(t-1)]\times \begin{bmatrix}\boldsymbol{P}(t-1|t-1) & \boldsymbol{\Delta}_1(t-1) & \boldsymbol{0}\\ \boldsymbol{\Delta}_1^{\mathrm{T}}(t-1) & \boldsymbol{Q}_w(t-1) & \boldsymbol{\Delta}_2(t-1)\\ \boldsymbol{0} & \boldsymbol{\Delta}_2^{\mathrm{T}}(t-1) & \boldsymbol{Q}_{\varepsilon}(t-1)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}(t-1)\\ \boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\\ -\boldsymbol{K}_1^{\mathrm{T}}(t-1)\end{bmatrix} \tag{3.10.28}$$

其中由(3.10.26)和(3.10.20)定义

$$\boldsymbol{\Delta}_1(t-1)=[\boldsymbol{I}_n-\boldsymbol{K}(t-1)\boldsymbol{H}(t-1)]\boldsymbol{\Gamma}(t-2)\boldsymbol{Q}_{w1}(t-2)-\boldsymbol{K}(t-1)\boldsymbol{S}^{\mathrm{T}}(t-1) \tag{3.10.29}$$

$$\boldsymbol{\Delta}_2(t-1)=\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-2)\boldsymbol{\Gamma}^{\mathrm{T}}(t-2)\boldsymbol{H}^{\mathrm{T}}(t-1)+\boldsymbol{S}(t-1) \tag{3.10.30}$$

下面来求 $\boldsymbol{P}(t|t)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t)]$, $\tilde{\boldsymbol{x}}(t|t)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t)$. 应用(3.10.15)有

$$\begin{aligned}\boldsymbol{P}(t|t)=&\mathrm{E}[(\tilde{\boldsymbol{x}}(t|t-1)-\boldsymbol{K}(t)\boldsymbol{\varepsilon}(t))(\tilde{\boldsymbol{x}}(t|t-1)-\boldsymbol{K}(t)\boldsymbol{\varepsilon}(t))^{\mathrm{T}}]=\\&\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1)]+\boldsymbol{K}(t)\boldsymbol{Q}_{\varepsilon}(t)\boldsymbol{K}^{\mathrm{T}}(t)-\\&\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]\boldsymbol{K}^{\mathrm{T}}(t)-\boldsymbol{K}(t)\mathrm{E}[\boldsymbol{\varepsilon}(t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1)]\end{aligned} \tag{3.10.31}$$

因 $\boldsymbol{\varepsilon}(t)\perp\hat{\boldsymbol{x}}(t|t-1)\in L(\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(1))=L(\boldsymbol{\varepsilon}(t-1),\cdots,\boldsymbol{\varepsilon}(1))$,于是应用(3.10.14)有

$$\begin{aligned}\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]=&\mathrm{E}[(\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t-1))\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]=\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]=\\&\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)\end{aligned} \tag{3.10.32}$$

由上两式引出

$$\begin{aligned}\boldsymbol{P}(t|t)=&\boldsymbol{P}(t|t-1)+\boldsymbol{K}(t)\boldsymbol{Q}_{\varepsilon}(t)\boldsymbol{K}^{\mathrm{T}}(t)-[\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)]\times\\&\boldsymbol{K}^{\mathrm{T}}(t)-\boldsymbol{K}(t)[\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)]^{\mathrm{T}}\end{aligned} \tag{3.10.33}$$

上述结果可概括为如下定理.

【定理 3.10.1】 带相邻时刻相关噪声的系统(3.10.1)和(3.10.2),在假设(3.10.3)下,有最优 Kalman 滤波器 $\hat{\boldsymbol{x}}(t|t)$和预报器 $\hat{\boldsymbol{x}}(t|t-1)$为

$$\hat{\boldsymbol{x}}(t|t)=\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}(t)\boldsymbol{\varepsilon}(t) \tag{3.10.34}$$

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)-\boldsymbol{H}(t)\hat{\boldsymbol{x}}(t|t-1) \tag{3.10.35}$$

$$\boldsymbol{K}(t)=[\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)]\boldsymbol{Q}_{\varepsilon}^{-1}(t) \tag{3.10.36}$$

$$\begin{aligned}\boldsymbol{Q}_{\varepsilon}(t)=&\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{Q}_v(t)+\\&\boldsymbol{H}(t)\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)+\boldsymbol{S}_1^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{H}^{\mathrm{T}}(t)\end{aligned} \tag{3.10.37}$$

$$\hat{\boldsymbol{x}}(t|t-1)=\boldsymbol{\Phi}(t-1)\hat{\boldsymbol{x}}(t-1|t-1)+\boldsymbol{K}_1(t-1)\boldsymbol{\varepsilon}(t-1) \tag{3.10.38}$$

$$\boldsymbol{K}_1(t-1)=\boldsymbol{\Gamma}(t-1)[\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-2)\boldsymbol{\Gamma}^{\mathrm{T}}(t-2)\boldsymbol{H}^{\mathrm{T}}(t-1)+\boldsymbol{S}(t-1)]\boldsymbol{Q}_{\varepsilon}^{-1}(t-1) \tag{3.10.39}$$

$$\begin{aligned}\boldsymbol{P}(t|t-1)=&[\boldsymbol{\Phi}(t-1),\boldsymbol{\Gamma}(t-1),-\boldsymbol{K}_1(t-1)]\times\\&\begin{bmatrix}\boldsymbol{P}(t-1|t-1) & \boldsymbol{\Delta}_1(t-1) & \boldsymbol{0}\\ \boldsymbol{\Delta}_1^{\mathrm{T}}(t-1) & \boldsymbol{Q}_w(t-1) & \boldsymbol{\Delta}_2(t-1)\\ \boldsymbol{0} & \boldsymbol{\Delta}_2^{\mathrm{T}}(t-1) & \boldsymbol{Q}_{\varepsilon}(t-1)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}(t-1)\\ \boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\\ -\boldsymbol{K}_1^{\mathrm{T}}(t-1)\end{bmatrix}\end{aligned} \tag{3.10.40}$$

$$\boldsymbol{\Delta}_1(t-1)=[\boldsymbol{I}_n-\boldsymbol{K}(t-1)\boldsymbol{H}(t-1)]\boldsymbol{\Gamma}(t-2)\boldsymbol{Q}_w(t-2)-\boldsymbol{K}(t-1)\boldsymbol{S}^{\mathrm{T}}(t-1) \tag{3.10.41}$$

$$\boldsymbol{\Delta}_2(t-1)=\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-2)\boldsymbol{\Gamma}^{\mathrm{T}}(t-2)\boldsymbol{H}^{\mathrm{T}}(t-1)+\boldsymbol{S}(t-1) \tag{3.10.42}$$

$$\begin{aligned}\boldsymbol{P}(t|t)=&\boldsymbol{P}(t|t-1)+\boldsymbol{K}(t)\boldsymbol{Q}_{\varepsilon}(t)\boldsymbol{K}^{\mathrm{T}}(t)-\\&[\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)]\boldsymbol{K}^{\mathrm{T}}(t)-\\&\boldsymbol{K}(t)[\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)]^{\mathrm{T}}\end{aligned} \tag{3.10.43}$$

$$\hat{\boldsymbol{x}}(1|0)=\boldsymbol{\mu}_0,\quad \boldsymbol{P}(1|0)=\boldsymbol{P}_0 \tag{3.10.44}$$

下面提出一种封闭形式 Kalman 预报器算法.

应用(3.10.38)和(3.10.34)有

$$\begin{aligned}\hat{\boldsymbol{x}}(t+1|t)&=\boldsymbol{\Phi}(t)\hat{\boldsymbol{x}}(t|t)+\boldsymbol{K}_1(t)\boldsymbol{\varepsilon}(t)=\\ &\boldsymbol{\Phi}(t)(\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}(t)\boldsymbol{\varepsilon}(t))+\boldsymbol{K}_1(t)\boldsymbol{\varepsilon}(t)=\\ &\boldsymbol{\Phi}(t)\hat{\boldsymbol{x}}(t|t-1)+[\boldsymbol{\Phi}(t)\boldsymbol{K}(t)+\boldsymbol{K}_1(t)][\boldsymbol{y}(t)-\boldsymbol{H}(t)\hat{\boldsymbol{x}}(t|t-1)]=\\ &[\boldsymbol{\Phi}(t)-(\boldsymbol{\Phi}(t)\boldsymbol{K}(t)+\boldsymbol{K}_1(t))\boldsymbol{H}(t)]\hat{\boldsymbol{x}}(t|t-1)+[\boldsymbol{\Phi}(t)\boldsymbol{K}(t)+\boldsymbol{K}_1(t)]\boldsymbol{y}(t)\end{aligned}\tag{3.10.45}$$

定义

$$\boldsymbol{K}_p(t)=\boldsymbol{\Phi}(t)\boldsymbol{K}(t)+\boldsymbol{K}_1(t)\tag{3.10.46}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t)\tag{3.10.47}$$

则有封闭形式 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p(t)\boldsymbol{y}(t)\tag{3.10.48}$$

另一方面由(3.10.45)Kalman 预报器还有新息形式

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p(t)\boldsymbol{\varepsilon}(t)\tag{3.10.49}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{\varepsilon}(t)\tag{3.10.50}$$

由(3.10.10)有

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t)\tag{3.10.51}$$

应用(3.10.1)、(310.49)和(3.10.51)有

$$\begin{aligned}\tilde{\boldsymbol{x}}(t+1|t)&=\boldsymbol{\Phi}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{K}_p(t)\boldsymbol{\varepsilon}(t)=\\ &\boldsymbol{\Phi}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{K}_p(t)(\boldsymbol{H}(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{v}(t)))=\\ &(\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t))\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{K}_p(t)\boldsymbol{v}(t)\end{aligned}\tag{3.10.52}$$

即

$$\tilde{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{K}_p(t)\boldsymbol{v}(t)\tag{3.10.53}$$

于是有 $\boldsymbol{P}(t+1|t)$为

$$\begin{aligned}\boldsymbol{P}(t+1|t)&=[\boldsymbol{\Psi}_p(t),\boldsymbol{\Gamma}(t),-\boldsymbol{K}_p(t)]\times\\ &\mathrm{E}\left\{\begin{bmatrix}\tilde{\boldsymbol{x}}(t|t-1)\\ \boldsymbol{w}(t-1)\\ \boldsymbol{v}(t)\end{bmatrix}[\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1),\boldsymbol{w}^{\mathrm{T}}(t),\boldsymbol{v}^{\mathrm{T}}(t)]\right\}\begin{bmatrix}\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\\ \boldsymbol{\Gamma}^{\mathrm{T}}(t)\\ -\boldsymbol{K}_p^{\mathrm{T}}(t)\end{bmatrix}\end{aligned}\tag{3.10.54}$$

应用(3.10.3)和(3.10.53)有

$$\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\boldsymbol{w}^{\mathrm{T}}(t)]=\boldsymbol{\Gamma}(t-1)\boldsymbol{Q}_{w1}(t-1)\tag{3.10.55}$$

$$\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\boldsymbol{v}^{\mathrm{T}}(t)]=\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)\tag{3.10.56}$$

将上两式代入(3.10.54)得

$$\begin{aligned}\boldsymbol{P}(t+1|t)&=[\boldsymbol{\Psi}_p(t),\boldsymbol{\Gamma}(t),-\boldsymbol{K}_p(t)]\times\\ &\begin{bmatrix}\boldsymbol{P}(t|t-1) & \boldsymbol{\Gamma}(t-1)\boldsymbol{Q}_{w1}(t-1) & \boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)\\ \boldsymbol{Q}_{\mathrm{w1}}^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1) & \boldsymbol{Q}_w(t) & \boldsymbol{S}(t)\\ \boldsymbol{S}_1^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1) & \boldsymbol{S}^{\mathrm{T}}(t) & \boldsymbol{Q}_v(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Psi}_p^{\mathrm{T}}(t)\\ \boldsymbol{\Gamma}^{\mathrm{T}}(t)\\ -\boldsymbol{K}_p^{\mathrm{T}}(t)\end{bmatrix}\end{aligned}\tag{3.10.57}$$

【注 3.10.1】 因 $\boldsymbol{\Psi}_p(t)$和 $\boldsymbol{K}_p(t)$中均含有 $\boldsymbol{P}(t|t-1)$,将它们的表达式代入上式便

得到关于 $\boldsymbol{P}(t|t-1)$的 Riccati 方程.

上述结果可概括为如下定理.

【定理 3.10.2】 带相邻时刻相关噪声系统(3.10.1)和(3.10.2)在假设(3.10.3)下，有最优 Kalman 预报器 $\hat{\boldsymbol{x}}(t+1|t)$为

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.10.58}$$

或

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p(t)\boldsymbol{\varepsilon}(t) \tag{3.10.59}$$

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)-\boldsymbol{H}(t)\hat{\boldsymbol{x}}(t|t-1) \tag{3.10.60}$$

其中 $\boldsymbol{K}_p(t)$是预报增益,

$$\boldsymbol{K}_p(t)=\boldsymbol{\Phi}(t)\boldsymbol{K}(t)+\boldsymbol{K}_1(t) \tag{3.10.61}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_p(t)\boldsymbol{H}(t) \tag{3.10.62}$$

$$\boldsymbol{K}(t)=[\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)]\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.10.63}$$

$$\begin{aligned}\boldsymbol{Q}_\varepsilon(t)=&\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{Q}_v(t)+\\&\boldsymbol{H}(t)\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)+\boldsymbol{S}_1^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{H}^{\mathrm{T}}(t)\end{aligned} \tag{3.10.64}$$

$$\boldsymbol{K}_1(t)=\boldsymbol{\Gamma}(t)[\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{S}(t)]\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.10.65}$$

且 $\boldsymbol{P}(t|t-1)$由(3.10.57)递推计算,带初值 $\hat{\boldsymbol{x}}(1|0)=\boldsymbol{\mu}$,$\boldsymbol{P}(1|0)=\boldsymbol{P}_0$. 而 $\boldsymbol{P}(t|t)$由(3.10.43)计算.

【推论 3.10.1】 带相关噪声系统(3.10.1)和(3.10.2)在假设(3.10.3)下,若 $\boldsymbol{S}_1(t)=\boldsymbol{0}$,$\boldsymbol{Q}_{w1}(t)=\boldsymbol{0}$,则有常规的 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p(t)\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p(t)\boldsymbol{y}(t) \tag{3.10.66}$$

$$\boldsymbol{K}_p(t)=[\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)]\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.10.67}$$

$$\boldsymbol{Q}_\varepsilon(t)=\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{Q}_v(t) \tag{3.10.68}$$

$\boldsymbol{P}(t+1|t)$满足方程

$$\boldsymbol{P}(t+1|t)=\boldsymbol{\Psi}_p(t)\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+[\boldsymbol{\Gamma}(t),-\boldsymbol{K}_p(t)]\begin{bmatrix}\boldsymbol{Q}_w(t) & \boldsymbol{S}(t)\\ \boldsymbol{S}^{\mathrm{T}}(t) & \boldsymbol{Q}_v(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Gamma}^{\mathrm{T}}(t)\\ -\boldsymbol{K}_p^{\mathrm{T}}(t)\end{bmatrix} \tag{3.10.69}$$

它等价于 Riccati 方程

$$\begin{aligned}\boldsymbol{P}(t+1|t)=&\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t-1)\boldsymbol{\Phi}^{\mathrm{T}}(t)-[\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)]\times\\&[\boldsymbol{H}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{Q}_v(t)]^{-1}[\boldsymbol{\Phi}(t)\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{S}(t)]^{\mathrm{T}}+\\&\boldsymbol{\Gamma}(t)\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)\end{aligned} \tag{3.10.70}$$

带初值 $\hat{\boldsymbol{x}}(1|0)=\boldsymbol{\mu}$. $\boldsymbol{P}(1|0)=\boldsymbol{P}_0$.

【注 3.10.2】 在定理 3.10.1 和定理 3.10.2 中置 $\boldsymbol{S}(t)=\boldsymbol{0}$,$\boldsymbol{Q}_{w1}(t)=\boldsymbol{0}$,可得带滞后相关噪声($S_1(t)\neq 0$)系统 Kalman 滤波器和预报器[3].

下面推导最优 Kalman 平滑器.

由递推射影公式有

$$\hat{\boldsymbol{x}}(t|t+N)=\hat{\boldsymbol{x}}(t|t+N-1)+\boldsymbol{K}(t|t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.10.71}$$

$$\boldsymbol{K}(t|t+N)=\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N)]\boldsymbol{Q}_\varepsilon^{-1}(t+N) \tag{3.10.72}$$

对 $N=0$，我们有

$$\boldsymbol{K}(t|t)=\boldsymbol{K}(t)=[\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)]\boldsymbol{Q}_{\varepsilon}^{-1}(t) \tag{3.10.73}$$

其中 $\boldsymbol{K}(t)$由(3.10.36)定义.

对 $N=1$，注意

$$\boldsymbol{K}(t|t+1)=\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+1)]\boldsymbol{Q}_{\varepsilon}^{-1}(t+1) \tag{3.10.74}$$

且应用(3.10.51)和(3.10.53)有

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{H}(t+1)[\boldsymbol{\Psi}_p(t)\tilde{\boldsymbol{x}}(t|t-1)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{K}_p(t)\boldsymbol{v}(t)]+\boldsymbol{v}(t+1) \tag{3.10.75}$$

将上式代入(3.10.74)，注意

$$\mathrm{E}[\boldsymbol{x}(t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1)]=\boldsymbol{P}(t|t-1) \tag{3.10.76}$$

$$\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{w}^{\mathrm{T}}(t)]=\boldsymbol{\Gamma}(t-1)\boldsymbol{Q}_{w1}(t-1) \tag{3.10.77}$$

$$\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{v}^{\mathrm{T}}(t)]=\mathrm{E}[(\boldsymbol{\Phi}(t-1)\boldsymbol{x}(t-1)+\boldsymbol{\Gamma}(t-1)\boldsymbol{w}(t-1))\boldsymbol{v}^{\mathrm{T}}(t)]=\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1) \tag{3.10.78}$$

则有

$$\begin{aligned}\boldsymbol{K}(t|t+1)=[&\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{Q}_{w1}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\\&\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)\boldsymbol{K}_p^{\mathrm{T}}(t)]\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\end{aligned} \tag{3.10.79}$$

对 $N>1$ 情形，由(3.10.53)迭代 N 次有

$$\begin{aligned}\tilde{\boldsymbol{x}}(t+N\mid t+N-1)=\boldsymbol{\Psi}(t+N,t)\tilde{\boldsymbol{x}}(t\mid t-1)+\sum_{i=t+1}^{t+N}\boldsymbol{\Psi}(t+N,i)\times\\ [\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1)-\boldsymbol{K}_p(i-1)\boldsymbol{v}(i-1)]\end{aligned} \tag{3.10.80}$$

其中定义 $\boldsymbol{\Psi}(t+N,t+N)=\boldsymbol{I}_n$，且

$$\boldsymbol{\Psi}(t+N,i)=\boldsymbol{\Psi}_p(t+N-1),\cdots\boldsymbol{\Psi}_p(i),\quad i<t+N \tag{3.10.81}$$

由(3.10.51)有

$$\boldsymbol{\varepsilon}(t+N)=\boldsymbol{H}(t+N)\tilde{\boldsymbol{x}}(t+N|t+N-1)+\boldsymbol{v}(t+N) \tag{3.10.82}$$

将(3.10.80)代入(3.10.82)后，再代入(3.10.72)可得

$$\begin{aligned}\boldsymbol{K}(t|t+N)=\{&\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}^{\mathrm{T}}(t+N,t)+[\boldsymbol{\Gamma}(t-1)\boldsymbol{Q}_{w1}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\\&\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)\boldsymbol{K}_p^{\mathrm{T}}(t)]\boldsymbol{\Psi}^{\mathrm{T}}(t+N,t+1)\}\boldsymbol{H}^{\mathrm{T}}(t+N)\boldsymbol{Q}_{\varepsilon}^{-1}(t+N)\end{aligned} \tag{3.10.83}$$

【定理 3.10.3】 带相邻时刻相关噪声系统(3.10.1)和(3.10.2)在假设(3.10.3)下，有最优 Kalman 平滑器

$$\hat{\boldsymbol{x}}(t|t+N)=\hat{\boldsymbol{x}}(t|t+N-1)+\boldsymbol{K}(t|t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.10.84}$$

其中 $N=0,1,2,\cdots$，初值 $\hat{\boldsymbol{x}}(t|t-1)$由定理 3.10.2 给出，平滑增益 $\boldsymbol{K}(t|t+N)$为

$$\boldsymbol{K}(t\mid t+N)=\boldsymbol{D}_x(t)\left\{\prod_{i=1}^{N-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\right\}\boldsymbol{H}^{\mathrm{T}}(t+N)\boldsymbol{Q}_{\varepsilon}^{-1}(t+N),\quad N\geqslant 2 \tag{3.10.85}$$

$$\boldsymbol{D}_x(t)=\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t-1)\boldsymbol{Q}_{w1}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{\Gamma}(t-1)\boldsymbol{S}_1(t-1)\boldsymbol{K}_p^{\mathrm{T}}(t) \tag{3.10.86}$$

对 $N=0$ 和 $N=1$，$\boldsymbol{K}(t|t)$分别由(3.10.73)和(3.10.79)给出.

估值误差 $\tilde{\boldsymbol{x}}(t|t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t+N)$方差阵 $\boldsymbol{P}(t|t+N)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|$

$t+N)$]为

$$\boldsymbol{P}(t\mid t+N)=\boldsymbol{P}(t\mid t+N-1)-\boldsymbol{K}(t\mid t+N)\boldsymbol{Q}_{\varepsilon}(t+N)\boldsymbol{K}^{\mathrm{T}}(t\mid t+N) \tag{3.10.87}$$

带初值 $\boldsymbol{P}(t\mid t-1)$.

证明 由(3.10.81)和(3.10.83)得(3.10.85)和(3.10.86).由(3.10.71)有

$$\tilde{\boldsymbol{x}}(t\mid t+N)=\tilde{\boldsymbol{x}}(t\mid t+N-1)-\boldsymbol{K}(t\mid t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.10.88}$$

由 $\tilde{\boldsymbol{x}}(t\mid t+N)\perp\boldsymbol{\varepsilon}(t+N)$,将上式右边第二项移到左边后,直接可得(3.10.87). □

【推论 3.10.2】 在定理 3.10.3 条件下,有非递推最优平滑器

$$\hat{\boldsymbol{x}}(t\mid t+N)=\hat{\boldsymbol{x}}(t\mid t-1)+\sum_{i=0}^{N}\boldsymbol{K}(t\mid t+i)\boldsymbol{\varepsilon}(t+i),\quad N\geqslant 0 \tag{3.10.89}$$

且有估值误差方差阵

$$\boldsymbol{P}(t\mid t+N)=\boldsymbol{P}(t\mid t-1)-\sum_{i=0}^{N}\boldsymbol{K}(t\mid t+i)\boldsymbol{Q}_{\varepsilon}(t+i)\boldsymbol{K}^{\mathrm{T}}(t\mid t+i) \tag{3.10.90}$$

【定理 3.10.4】 带相邻时刻相关噪声系统(3.10.1)和(3.10.2)在假设(3.10.3)下,有最优固定区间 Kalman 平滑器($N>0$)

$$\hat{\boldsymbol{x}}(t\mid N)=\hat{\boldsymbol{x}}(t\mid t)+\boldsymbol{D}_x(t)\boldsymbol{r}(t+1\mid N) \tag{3.10.91}$$

$$\boldsymbol{r}(t+1\mid N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)\boldsymbol{r}(t+2\mid N)+\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{3.10.92}$$

带初值 $\boldsymbol{r}(N+1\mid N)=\mathbf{0}$, $t=N-1,\cdots,1$.平滑误差 $\tilde{\boldsymbol{x}}(t\mid N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t\mid N)$方差阵 $\boldsymbol{P}(t\mid N)$为

$$\boldsymbol{P}(t\mid N)=\boldsymbol{P}(t\mid t)-\boldsymbol{D}_x(t)\boldsymbol{S}(t+1\mid N)\boldsymbol{D}_x^{\mathrm{T}}(t) \tag{3.10.93}$$

$$\boldsymbol{S}(t+1\mid N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)\boldsymbol{S}(t+2\mid N)\boldsymbol{\Psi}_p(t+1)+\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_{\varepsilon}^{-1}(t+1)\boldsymbol{H}(t+1) \tag{3.10.94}$$

带初值 $\boldsymbol{S}(N+1\mid N)=\mathbf{0}$, $t=N-1,\cdots,1$.

证明 类似于定理 3.4.3 的证明,从略. □

【注 3.10.3】 (3.10.90)稍不同于定理 3.4.3,其中 $\hat{\boldsymbol{x}}(t\mid N)=\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{P}(t\mid t-1)\boldsymbol{r}(t\mid N)$.这是因为在(3.10.88)$\boldsymbol{K}(t\mid t)$中没有因式 $\boldsymbol{D}_x(t)$.为了使第二项能提出公因式 $\boldsymbol{D}_x(t)$,应将(3.10.88)改写为

$$\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{x}(t\mid t)+\sum_{i=1}^{N}\boldsymbol{K}(t\mid t+i)\boldsymbol{\varepsilon}(t+i) \tag{3.10.95}$$

这导至(3.10.90).

下面求 N 步 Kalman 预报器.

由(3.10.1)和射影性质有

$$\hat{\boldsymbol{x}}(t+2\mid t)=\boldsymbol{\Phi}(t+1)\hat{\boldsymbol{x}}(t+1\mid t)+\boldsymbol{\Gamma}(t+1)\hat{\boldsymbol{w}}(t+1\mid t) \tag{3.10.96}$$

因 $\boldsymbol{w}(t+1)\perp L(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(1))$,故有 $\hat{\boldsymbol{w}}(t+1\mid t)=\mathbf{0}$,于是有

$$\hat{\boldsymbol{x}}(t+2\mid t)=\boldsymbol{\Phi}(t+1)\hat{\boldsymbol{x}}(t+1\mid t) \tag{3.10.97}$$

用归纳法可求得

$$\hat{\boldsymbol{x}}(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\hat{\boldsymbol{x}}(t+1\mid t),\quad N\geqslant 2 \tag{3.10.98}$$

其中定义

$$\boldsymbol{\Phi}(t+N,t+1)=\boldsymbol{\Phi}(t+N-1)\cdots\boldsymbol{\Phi}(t+1),\quad \boldsymbol{\Phi}(t+N,t+N)=\boldsymbol{I}_m \tag{3.10.99}$$

另一方面(3.10.1)迭代有

$$\boldsymbol{x}(t+N)=\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{x}(t+1)+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}(t+N,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1) \tag{3.10.100}$$

这引出预报误差 $\tilde{\boldsymbol{x}}(t+N|t)=\boldsymbol{x}(t+N)-\hat{\boldsymbol{x}}(t+N|t)$为

$$\tilde{\boldsymbol{x}}(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\tilde{\boldsymbol{x}}(t+1\mid t)+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}(t+N,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1) \tag{3.10.101}$$

应用(3.10.53)且应用假设 $\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(t-j)]=\boldsymbol{0}(j\geqslant1)$可得

$$\mathrm{E}[\tilde{\boldsymbol{x}}(t+1|t)\boldsymbol{w}^{\mathrm{T}}(t+1)]=\boldsymbol{\Gamma}(t)\boldsymbol{Q}_{w1}(t) \tag{3.10.102}$$

于是由上两式可得预报误差方差阵 $\boldsymbol{P}(t+N|t)=\mathrm{E}[\tilde{\boldsymbol{x}}(t+N|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+N|t)]$为

$$\begin{aligned}\boldsymbol{P}(t+N|t)=&\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{P}(t+1|t)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+1)+\\&\sum_{i=t+1}^{t+N}\boldsymbol{\Phi}(t+N,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{Q}_w(i-1)\boldsymbol{\Gamma}^{\mathrm{T}}(i-1)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,i)+\\&\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{\Gamma}(t)\boldsymbol{Q}_{w1}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t+1)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+2)+\\&\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+2)\boldsymbol{\Gamma}(t+1)\boldsymbol{Q}_{w1}^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+1)\end{aligned} \tag{3.10.103}$$

上述推导可概括为如下定理.

【定理 3.10.5】 系统(3.10.1)和(3.10.2)在假设(3.10.3)下,有超前 N 步 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+N|t)=\boldsymbol{\Phi}(t+N,t+1)\hat{\boldsymbol{x}}(t+1|t),\quad N\geqslant2 \tag{3.10.104}$$

且预报误差方差阵 $\boldsymbol{P}(t+N|t)$由(3.10.103)给出.$\hat{\boldsymbol{x}}(t+1|t)$由定理 3.10.2 给出.

应用定理 3.10.2 的 Kalman 预报器和射影理论,可导出如下统一和通用的白噪声估计定理,证明从略.

【定理 3.10.6】 系统式(3.10.1)~(3.10.3)有统一的最优递推拟白噪声 $\boldsymbol{w}(t)$和白噪声 $\boldsymbol{v}(t)$的平滑器 $\hat{\boldsymbol{\theta}}(t|t+N)$,$\boldsymbol{\theta}=\boldsymbol{w},\boldsymbol{v}$,

$$\hat{\boldsymbol{\theta}}(t|t+N)=\boldsymbol{0},\quad N<0 \tag{3.10.105}$$

$$\hat{\boldsymbol{\theta}}(t|t+N)=\hat{\boldsymbol{\theta}}(t|t+N-1)+\boldsymbol{M}_\theta(t|t+N)\boldsymbol{\varepsilon}(t+N) \tag{3.10.106}$$

$$\boldsymbol{M}_\theta(t\mid t+N)=\boldsymbol{D}_\theta(t+1)\left\{\prod_{i=2}^{N-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\right\}\boldsymbol{H}^{\mathrm{T}}(t+N)\boldsymbol{Q}_\varepsilon^{-1}(t+N),\quad N\geqslant2 \tag{3.10.107}$$

其中 $N=0,1,\cdots$且规定当 $N=2$ 时,

$$\prod_{i=2}^{2-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)=\boldsymbol{I}_n \tag{3.10.108}$$

且有

$$\begin{aligned}\boldsymbol{D}_w(t+1)=&[\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\\&\boldsymbol{S}(t)\boldsymbol{K}_p^{\mathrm{T}}(t)]\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)+\boldsymbol{Q}_{w1}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t+1)-\boldsymbol{S}_1(t)\boldsymbol{K}_p^{\mathrm{T}}(t+1)\end{aligned} \tag{3.10.109}$$

$$\boldsymbol{D}_v(t+1)=[\boldsymbol{S}_1^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\boldsymbol{S}^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{Q}_v(t)\boldsymbol{K}_p^{\mathrm{T}}(t)]\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1) \tag{3.10.110}$$

且有

$$\boldsymbol{M}_w(t|t+1)=\{[\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{S}(t)\boldsymbol{K}_p^{\mathrm{T}}(t)]\times$$
$$\boldsymbol{H}^{\mathrm{T}}(t+1)+\boldsymbol{S}_1(t)\}\boldsymbol{Q}_\varepsilon^{-1}(t+1) \tag{3.10.111}$$

$$\boldsymbol{M}_v(t|t+1)=[\boldsymbol{S}_1^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+$$
$$\boldsymbol{S}^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{Q}_v(t)\boldsymbol{K}_p^{\mathrm{T}}(t)]\boldsymbol{H}^{\mathrm{T}}(t+1)\boldsymbol{Q}_\varepsilon^{-1}(t+1) \tag{3.10.112}$$

$$\boldsymbol{M}_w(t|t)=[\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{S}(t)]\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.10.113}$$

$$\boldsymbol{M}_v(t|t)=[\boldsymbol{S}_1^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{Q}_v(t)]\boldsymbol{Q}_\varepsilon^{-1}(t) \tag{3.10.114}$$

估值误差 $\tilde{\boldsymbol{\theta}}(t|t+N)=\boldsymbol{\theta}(t)-\hat{\boldsymbol{\theta}}(t|t+N)$的方差阵 $\boldsymbol{P}_\theta(t|t+N)=\mathrm{E}[\tilde{\boldsymbol{\theta}}(t|t+N)\tilde{\boldsymbol{\theta}}^{\mathrm{T}}(t|t+N)]$为

$$\boldsymbol{P}_\theta(t|t+N)=\boldsymbol{P}_\theta(t|t+N-1)-\boldsymbol{M}_\theta(t|t+N)\boldsymbol{Q}_\varepsilon(t+N)\boldsymbol{M}_\theta^{\mathrm{T}}(t|t+N) \tag{3.10.115}$$

带初值 $\boldsymbol{P}_\theta(t|t+N)=\boldsymbol{Q}_\theta(t)$, $N<0$, $\theta=w,v$.

【定理 3.10.7】 系统式(3.10.1)~(3.10.3)有统一的最优递推固定区间拟白噪声 $\boldsymbol{w}(t)$和白噪声 $\boldsymbol{v}(t)$平滑器 $\hat{\boldsymbol{\theta}}(t|N)$为

$$\hat{\boldsymbol{\theta}}(t|N)=\hat{\boldsymbol{\theta}}(t|t+1)+\boldsymbol{D}_\theta(t+1)\boldsymbol{r}(t+2|N),\quad \boldsymbol{\theta}=\boldsymbol{w},\boldsymbol{v} \tag{3.10.116}$$

$$\boldsymbol{r}(t+2|N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t+2)\boldsymbol{r}(t+3|N)+\boldsymbol{H}^{\mathrm{T}}(t+2)\boldsymbol{Q}_\varepsilon^{-1}(t+2)\boldsymbol{\varepsilon}(t+2) \tag{3.10.117}$$

带初值 $\boldsymbol{r}(N+1|N)=\boldsymbol{0}$,且定义

$$\boldsymbol{D}_w(t+1)=[\boldsymbol{Q}_{w1}^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\boldsymbol{Q}_w(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-$$
$$\boldsymbol{S}(t)\boldsymbol{K}_p^{\mathrm{T}}(t)]\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1)+\boldsymbol{Q}_{w1}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t+1)-\boldsymbol{S}_1(t)\boldsymbol{K}_p^{\mathrm{T}}(t+1) \tag{3.10.118}$$

$$\boldsymbol{D}_v(t+1)=[\boldsymbol{S}_1^{\mathrm{T}}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1)\boldsymbol{\Psi}_p^{\mathrm{T}}(t)+\boldsymbol{S}^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{Q}_v(t)\boldsymbol{K}_p^{\mathrm{T}}(t)]\boldsymbol{\Psi}_p^{\mathrm{T}}(t+1) \tag{3.10.119}$$

且有估值误差 $\tilde{\boldsymbol{\theta}}(t|N)=\boldsymbol{\theta}(t)-\hat{\boldsymbol{\theta}}(t|N)$的方差阵为

$$\boldsymbol{P}_\theta(t|N)=\boldsymbol{P}_\theta(t|t-1)-\boldsymbol{D}_\theta(t+1)\boldsymbol{S}(t+2|N)\boldsymbol{D}_\theta^{\mathrm{T}}(t+1) \tag{3.10.120}$$

$$\boldsymbol{S}(t+2|N)=\boldsymbol{\Psi}_p^{\mathrm{T}}(t+2)\boldsymbol{S}(t+3|N)\boldsymbol{\Psi}_p(t+2)+\boldsymbol{H}^{\mathrm{T}}(t+2)\boldsymbol{Q}_\varepsilon^{-1}(t+2)\boldsymbol{H}(t+2) \tag{3.10.121}$$

其中 $\boldsymbol{S}(t+2|N)=\mathrm{E}[\boldsymbol{r}(t+2|N)\boldsymbol{r}^{\mathrm{T}}(t+2|N)]$,初值为 $\boldsymbol{S}(N+1|N)=\boldsymbol{0}$.

注意,在上述两定理中置 $\boldsymbol{Q}_{w1}(t)=\boldsymbol{0}$, $\boldsymbol{S}_1(t)=\boldsymbol{0}$ 可引出文献[8]给出的带相关噪声系统白噪声估值器作为特例.

【例 3.10.2】 仿真例子.

考虑随机系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}1 & -0.1\\ 0 & 0.8\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\ 0.5\end{bmatrix}w(t) \tag{3.10.122}$$

$$y(t)=[1\quad 1]\boldsymbol{x}(t)+v(t) \tag{3.10.123}$$

$$w(t)=g(t)+\alpha g(t-1) \tag{3.10.124}$$

$$v(t)=g(t-1) \tag{3.10.125}$$

$$g(t)=b(t)\xi(t) \tag{3.10.126}$$

假设 $b(t)$ 是取值 1 或 0 的 Bernoulli 白噪声，取值概率为 $P(b(t)=1)=\lambda$，$P(b(t)=0)=1-\lambda$，$\xi(t)$ 是零均值、方差为 $\sigma_\xi^2=1.5$ 的独立于 $b(t)$ 的 Gaussian 白噪声. 求拟白噪声 $w(t)$ 和 Bernoulli－Gaussian 白噪声 $v(t)$ 的最优平滑器 $\hat{w}(t|t+3)$ 和 $\hat{v}(t|t+3)$. 应用定理 3.10.6 仿真结果如图 3.10.1 和图 3.10.2 所示，其中实线端点纵标为真实值 $w(t)$ 或 $v(t)$，实圆点纵标为平滑估值 $\hat{w}(t|t+3)$ 或 $\hat{v}(t|t+3)$，可看到它们有较高的精度.

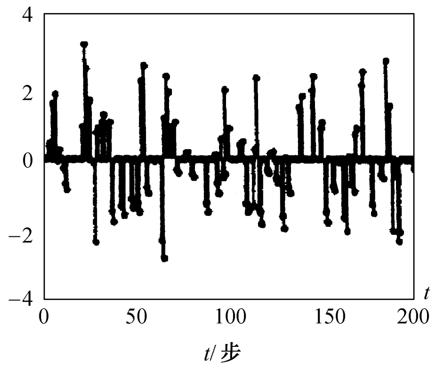


图 3.10.1　拟白噪声 $w(t)$ 和最优平滑器 $\hat{w}(t|t+3)$

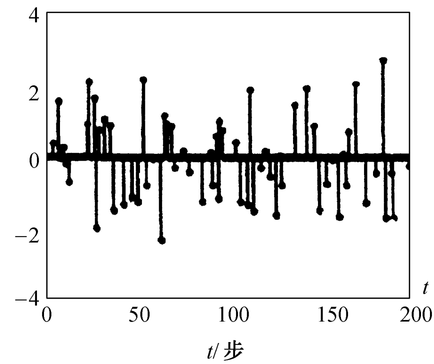


图 3.10.2　Bernoulli－Gaussian 观测白噪声 $v(t)$ 和最优平滑器 $\hat{v}(t|t+3)$

参　考　文　献

1　Kalman R E. A New Approach to Linear Filtering and Predictiong Problems. Trans. ASME，J. Basic Eng.，1960，82D:34～45

2　Anderson B D O，Moore J B. Optimal Filtering. New Jersey：Prentie－Hall，1979

3　Lewis F L Optimal Estimation. New York:John Wiley & Sons，1986

4　Kailath T，Sayed A，Hassibi B. Linear Estimation. New Jersey：Prentice－Hall，2000

5　Kamen E W，Su J K. Introduction to Optimal Estimation. Springer－Verlag London Limited，1999

6　Mendel J M. Lessons in Estimation Theory for Signal Processing，Communications，and Control. New Jersey：Prentice－Hall，1995

7　邓自立.最优滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2000

8　邓自立.卡尔曼滤波与维纳滤波——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2001

9　邓自立.自校正滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2003

10　Mendel J M. Optimal Seismic Deconvolution：An Estimation－Based Approach. New York：Academic Press，1983

11　Mendel J M. White Noise Estimators for Seismic Data Processing in Oil Exploration. IEEE Trans. Automatic Control，1977，AC－25(5)：694～706

12　邓自立,许燕.基于 Kalman 滤波的白噪声估计理论.自动化学报.2003,29(1):23～31

13　邓自立.时变系统的统一和通用最优白噪声估值器.控制理论与应用.2003,20(11):143～146

14　邓自立,孙书利,石莹.统一的和通用的 Kalman 滤波理论.科学技术与工程,2003,3(5):400～404

15　邓自立,孙书利,石莹.基于 Kalman 滤波的白噪声估计理论的推广.科学技术与工程,2003,3(6):521～524

16　徐守寿.随机信号估计与系统控制.北京:北京工业大学出版社,2001

17 Söderström T. Discrete – Time Stochastic System, Estimation and control. New York: Prentice – Hall, 1994
18 Chen C K, Chen C. Kalman Filtering with Real – Time Applications. Springer – Verlag, New York, 1987, 124 ~ 136
19 邓自立.解耦 Winer 状态估值器.中国学术期刊文摘(科技快报),2000,6(8):981 ~ 982
20 Grimble M J. Multichannel Optimal Linear Deconvolution Filters and Strip Thickness Estimation From Gauge Measurements. Trans, of the ASME, Journal of Dynamic Systems, Measurement, and Control, 1995, 117: 165 ~ 173
21 陈建国,师自强,王培德.Kalman 滤波理论的推广.控制理论与应用,1990,7(3):108 ~ 112
22 邓自立,郭一新.现代时间序列分析及其应用——建模、滤波、去卷、预报和控制.北京:知识出版社,1989
23 罗抟翼,程桂芳.随机信号处理与控制基础.北京:化学工业出版社,2002
24 邓自立,孙书利.基于 Kalman 滤波的带相关噪声系统统一的 Wiener 状态估值器.控制理论与应用,2003,20(4):573 ~ 576
25 邓自立,孙书利.极点配置固定滞后稳态 Kalman 平滑器.控制理论与应用,2003,20(5):802 ~ 804
26 Deng Zili, Sun Shuli. Wiener State Estimators Based on Kalman Filtering. Acta Automatica Sinica, 2004,30(1):126 ~ 130
27 Deng Zili, Xu Yan, Shi Ying, Sun Shuli. Unified Fast Suboptimal Fixed – Interval White Noise Wiener Smoothing Alqorithm. Acta Autovntica Sinica, 2004,30(2):224 ~ 228
28 Sun Shuli, Deng Zili. Pole – Assignment Fixed – Interval Kalman Smoother and Wiener Smoother. Acta Automatica Sinica, 2004,30(2):239 ~ 243
29 邓自立,石莹,孙书利,许燕.快速次优固定区间 Wiener 平滑算法.控制理论与应用,2004,21(2):275 ~ 278
30 邓自立.时域 Wiener 状态滤波新方法.控制理论与应用,2004,21(3):367 ~ 372
31 邓自立,许燕.基于 Kalman 滤波的通用的和统一的白噪声估计方法.控制理论与应用,2004,21(4):501 ~ 506
32 邓自立,高媛,王好谦.基于 Kalman 滤波的统一的 Wiener 状态滤波器.控制理论与应用,2004,21(6):1003 ~ 1006

第四章 ARMA 时间序列预报

已知时间序列的现在和过去的观测历史，预测（估计）其在将来时刻的取值，叫时间序列预报. 人们对这类预报问题有广泛的理论和应用兴趣. 例如气象预报、水文预报、商品销量预报、股市行情预报、目标跟踪预报、经济增长率预报等. 因此，ARMA 时间序列的预报是经典时间序列分析的基本课题之一.[1~6]

经典时间序列分析预报[2]的局限性是没有考虑带观测噪声的 ARMA 时间序列预报问题. 例如在对目标跟踪的雷达跟踪系统中，如何根据对目标位置 $s(t)$ 的观测信号 $y(t)$ 来预测运动目标将来的运动轨迹问题就属带观测噪声的 ARMA 信号预报问题. 事实上，设 T 为采样周期，$s(t)$，$\dot{s}(t)$ 和 $w(t)$ 各为在时刻 tT 运动目标的位置、速度和加速度，则有运动方程

$$s(t+1)=s(t)+\dot{s}(t)T+\frac{T^2}{2}w(t) \tag{4.0.1}$$

$$\dot{s}(t+1)=\dot{s}(t)+Tw(t) \tag{4.0.2}$$

$$y(t)=s(t)+v(t) \tag{4.0.3}$$

其中 $v(t)$ 是对位置的观测噪声，假设 $w(t)$ 和 $v(t)$ 是零均值、方差各为 σ_w^2 和 σ_v^2 的独立白噪声，问题是基于雷达对位置的观测信号 $(y(t),y(t-1)\cdots)$ 求运动目标真实位置 $s(t+N)$ 的最优预报器 $\hat{s}(t+N|t)$，$N=1,2,3\cdots$. 这类问题与导弹拦截或打击空中或海上运动目标（飞机、军舰等）有关. 运动模型 (4.0.1) ~ (4.0.3) 可用算子形式改写为

$$(1-q^{-1})s(t)=\dot{s}(t-1)T+\frac{T^2}{2}w(t-1) \tag{4.0.4}$$

$$(1-q^{-1})\dot{s}(t)=Tw(t-1) \tag{4.0.5}$$

$$y(t)=s(t)+v(t) \tag{4.0.6}$$

其中 q^{-1} 为单位滞后算子，$q^{-1}s(t)=s(t-1)$. 将 (4.0.5) 代入 (4.0.4) 消去 $\dot{s}(t-1)$ 可得运动目标位置 $s(t)$ 所满足的 ARMA 模型

$$(1-q^{-1})^2s(t)=\frac{T^2}{2}q^{-1}(1+q^{-1})w(t) \tag{4.0.7}$$

即问题转化为对带白色观测噪声 $v(t)$ 的 ARMA 信号 $s(t)$ 的预报问题：

$$(1-2q^{-1}+q^{-2})s(t)=\frac{T^2}{2}w(t)+\frac{T^2}{2}w(t-2) \tag{4.0.8}$$

$$y(t)=s(t)+v(t) \tag{4.0.9}$$

一般的带观测噪声的 ARMA 信号 $s(t)$ 的最优预报问题为

$$A(q^{-1})s(t)=C(q^{-1})w(t) \tag{4.0.10}$$

$$y(t)=s(t)+\eta(t)+\xi(t) \tag{4.0.11}$$

$$P(q^{-1})\eta(t)=R(q^{-1})n(t) \tag{4.0.12}$$

其中待估 ARMA 信号 $s(t)$ 带有白色观测噪声 $\xi(t)$ 和 ARMA 有色观测噪声 $\eta(t)$. 问题是

基于观测 $(y(t),y(t-1),\cdots)$ 求 $s(t+N)$ 的 N 步最优预报器 $\hat{s}(t+N|t)$. 利用 ARMA 模型与状态空间模型的转化,这个问题可用 Kalman 滤波方法解决.

本章除了介绍经典 ARMA 时间序列预报方法外,将提出解决这类预报问题的两种方法.此外本章也将介绍指数平滑预报方法,它的优点是不需要建立时间序列的数学模型,因而具有较大适用性.它的基本原理是由目前和过去观测数据 $(y(t),y(t-1),\cdots)$ 的加权和来求一步预报器 $\hat{y}(t+1|t)$,即

$$\hat{y}(t+1 \mid t)=\sum_{i=0}^{\infty} c_i y(t-i) \tag{4.0.13}$$

将用渐消记忆指数加权法求加权系数 c_i.

从历史上看,线性最小方差预报理论几乎同时被 Wiener[1] 和 Kolmosgorov[2] 提出. Wiener 预报理论的应用背景是在第二次世界大战期间 Wiener 研究火炮控制系统,要求对飞机的轨道预测.该项目完成于 1940 年,但由于军事上保密原因直到 1949 年才发表.[1] 但 Wiener 和 Kolmogorov[1,2] 的预报方法是频域方法.本章只介绍时域预报方法. 20 世纪 70 年代初,Box 和 Jenkins[3] 提出了著名的 Box - Jenkins 递推预报方法,Åström[8] 基于 Diophantine 方程提出了 Åström 预报方法,并将其应用于过程控制.还有一种时间序列的重要预报方法是 Kalman 预报方法.

经典时间序列预报方法[2,6] 是对单变量时间序列而言的,本章将介绍单变量和多变量 ARMA 时间序列预报,并给出在目标跟踪系统中的应用.特别,提出带观测噪声的 ARMA 时间序列的两种预报方法——ARMA 新息模型预报方法和 Kalman 预报方法.

本章作者将平稳时间序列预报理论和方法基于 Kalman 滤波理论和射影理论,推广到非平稳时间序列情形,提出了非平稳时间序列预报理论和方法.

4.1 Hilbert 空间中的射影运算

为了理论上的严谨,我们将用 Hilbert 空间中的射影理论解决最优预报问题.在考虑时间序列预报问题时,为了理论研究方便,假设已知到目前时刻 t 为止的全部过去观测历史,即初始观测时刻为 $t_0=-\infty$,这样我们已知在时刻 t 和时刻 t 以前的无穷个观测数据,基于这种信息来预报时间序列在将来时刻的值.这里我们考虑线性最优预报问题,即用这无限个观测数据的线性运算预报时间序列将来的值.从统计观点,时间序列是随机过程,也称随机序列,因而时间序列预报问题归结为构造基于到时刻 t 为止的无限个观测随机变量的线性函数作为时间序列将来值(随机变量)的最优预报器.这相当于在一个由无限个随机变量的线性函数构成的线性流形(线性空间或 Hilbert 空间)中求一个随机变量的最优估计.这等价于在无穷维 Hilbert 空间中的射影问题.它是第三章节 3.2 中基于有限个随机变量张成的线性流形中的射影理论的自然推广.

记带零均值且有二阶矩的随机变量全体所构成的集合为 H,对任意两个随机变量 $x,y\in H$,定义它们的内积为

$$(x,y)=\mathrm{E}(xy) \tag{4.1.1}$$

其中 E 为均值号,且定义随机变量 $x\in H$ 的范数 $\|x\|$ 为

$$\|x\|=(x,x)^{\frac{1}{2}}=(\mathrm{E}x^2)^{\frac{1}{2}} \tag{4.1.2}$$

在这种范数下，随机变量序列的收敛性是指均方收敛. 称这种规定范数的空间为 Hilbert 空间. 下面介绍 Hilbert 空间射影理论的概要[11,12]. 证明从略.

考虑 Hilbert 空间 H 中的时间序列 $\{y(t), y(t-1), \cdots\}$，设初始观测时刻为 $t_0=-\infty$，则这个时间序列含有无限个随机变量 $y(t), y(t-1), \cdots$. 记 $L=L(y(t), y(t-1), \cdots)$ 为由它们张成的子 Hlibert 空间，即 L 中的元素由 $y(t), y(t-1), \cdots$ 的线性组合及其均方极限组成. 设随机变量 $x\in H$，基于 $(y(t), y(t-1), \cdots)$ 对 x 的线性最小方差估计 $\hat{x}$ 是子空间 L 中的元素，它极小化均方误差

$$J=\min_{z\in L}\|x-z\| = \|x-\hat{x}\| \tag{4.1.3}$$

并称 $\hat{x}$ 为 x 在子空间 L 上的射影，记为

$$\hat{x}=\mathrm{proj}(x\mid y(t), y(t-1), \cdots)=\mathrm{proj}(x\mid L) \tag{4.1.4}$$

【定理 4.1.1】 （正交射影定理）$\hat{x}$ 为 x 在 L 上的射影的充要条件为

$$\mathrm{E}[(x-\hat{x})z]=0, \quad \forall z\in L \tag{4.1.5}$$

即 $(x-\hat{x})$ 与 z 的内积为零，$(x-\hat{x}, z)=0$，就是说 $(x-\hat{x})$ 不相关（正交）于 z，记为

$$(x-\hat{x})\perp z, \quad z\in L \tag{4.1.6}$$

记 $y(t)$ 在子 Hilbert 空间 $L(y(t-1), y(t-2), \cdots)$ 上的射影为 $\hat{y}(t\mid t-1)$，即

$$\hat{y}(t\mid t-1)=\mathrm{proj}(y(t)\mid y(t-1), y(t-2), \cdots)=\mathrm{proj}(y(t)\mid L) \tag{4.1.7}$$

由(4.1.3)称它是基于 $(y(t-1), y(t-2), \cdots)$ 对 $y(t)$ 的线性最小方差一步预报估值. 因为 L 是由 $y(t-1), y(t-2), \cdots$ 的有限线性组合及均方极限生成，且通常 $y(t), y(t-1), \cdots$ 为相关随机序列，求射影是困难和复杂的. 故人们希望引出等价于 $(y(t), y(t-1), \cdots)$ 的正交（不相关）序列. 称为新息序列. 它大大简化了对射影计算.

【定义 4.1.1】 随机序列 $y(t)$ 的新息序列定义为

$$\varepsilon(t)=y(t)-\hat{y}(t\mid t-1) \tag{4.1.8}$$

【定理 4.1.2】 新息序列 $\varepsilon(t), \varepsilon(t-1), \cdots$ 是带零均值的白噪声：

$$\mathrm{E}[\varepsilon(t)]=0, \quad \forall t, \quad \mathrm{E}[\varepsilon(t)\varepsilon(j)]=\sigma_{\varepsilon}^2(t)\delta_{tj} \tag{4.1.9}$$

其中 $\sigma_{\varepsilon}^2(t)$ 是 $\varepsilon(t)$ 的方差. 这表明新息序列 $\varepsilon(t), \varepsilon(t-1), \cdots$ 是 $L(y(t), y(t-1), \cdots)$ 中的正交序列，即 $\varepsilon(t)\perp\varepsilon(j)\ (t\neq j)$.

【定理 4.1.3】 （射影公式）随机变量 $x\in H$ 在子 Hilbert 空间 $L(\varepsilon(t), \varepsilon(t-1), \cdots)$ 上的射影为

$$\hat{x} = \sum_{i=0}^{\infty}\frac{\mathrm{E}[x\varepsilon(t-i)]}{\mathrm{E}[\varepsilon^2(t-i)]}\varepsilon(t-i) \tag{4.1.10}$$

【定理 4.1.4】 （递推射影公式）随机变量 $x\in H$ 在子 Hilbert 空间 $L(\varepsilon(t), \varepsilon(t-1), \cdots)$ 上的射影可递推计算为

$$\begin{aligned}\mathrm{proj}(x\mid\varepsilon(t), \varepsilon(t-1), \cdots)=\mathrm{proj}(x\mid\varepsilon(t-1), \varepsilon(t-2), \cdots)+\\ \mathrm{E}[x\varepsilon(t)][\mathrm{E}(\varepsilon^2(t))]^{-1}\varepsilon(t)\end{aligned} \tag{4.1.11}$$

【定义 4.1.2】 定义 $m\times 1$ 维带零均值和分量有二阶矩的时间序列 $(\boldsymbol{y}(t), \boldsymbol{y}(t-1), \cdots)$ 的新息序列 $\boldsymbol{\varepsilon}(t)$ 为

$$\begin{gathered}\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{y}}(t\mid t-1),\\ \hat{\boldsymbol{y}}(t\mid t-1)=[\hat{\boldsymbol{y}}_1(t\mid t-1), \cdots, \hat{\boldsymbol{y}}_m(t\mid t-1)]\end{gathered} \tag{4.1.12}$$

其中记 $\boldsymbol{y}(t)=[y_1(t), \cdots, y_m(t)]^{\mathrm{T}}$，T 为转置号，定义 $L(\boldsymbol{y}^{\mathrm{T}}(t-1), \boldsymbol{y}^{\mathrm{T}}(t-2), \cdots)$ 为由一

维随机序列 $\boldsymbol{y}^{\mathrm{T}}(t-1),\boldsymbol{y}^{\mathrm{T}}(t-2),\cdots$张成的 Hilbert 空间，且定义 $\hat{y}_i(t|t-1)$为 $y_i(t)$在一维子 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t-1),\boldsymbol{y}^{\mathrm{T}}(t-2),\cdots)$上的射影，

$$\hat{y}_i(t|t-1)=\mathrm{proj}(y_i(t)|\boldsymbol{y}^{\mathrm{T}}(t-1),\boldsymbol{y}^{\mathrm{T}}(t-2),\cdots) \tag{4.1.13}$$

【定理 4.1.5】 $m\times1$ 维新息序列$(\boldsymbol{\varepsilon}(t),\boldsymbol{\varepsilon}(t-1),\cdots)$是带零均值白噪声：

$$\mathrm{E}[\boldsymbol{\varepsilon}(t)]=\boldsymbol{0},\quad \forall t,\quad \mathrm{E}[\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(j)]=\boldsymbol{0},\quad \forall t\neq j \tag{4.1.14}$$

【定义 4.1.3】 设 $\boldsymbol{x}=[x_1,\cdots,x_n]^{\mathrm{T}}$ 为带零均值和有二阶矩的 $n\times1$ 维随机变量，定义它在 $m\times1$ 维随机序列$(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$张成的一维子 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)$上的射影为

$$\hat{\boldsymbol{x}}=[\hat{x}_1,\cdots,\hat{x}_n]^{\mathrm{T}} \tag{4.1.15}$$

其中定义 $\hat{x}_i(i=1,\cdots,n)$为分量 x_i 在一维子 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)$上的射影

$$\hat{x}_i=\mathrm{proj}(x_i|\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots) \tag{4.1.16}$$

也称 $\hat{\boldsymbol{x}}$ 为基于$(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$对 $\boldsymbol{x}$ 的线性最小方差估值，即

$$\hat{\boldsymbol{x}}=\mathrm{proj}(\boldsymbol{x}|\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots) \tag{4.1.17}$$

【定理 4.1.6】 （多维正交射影定理）$\hat{\boldsymbol{x}}$ 为线性最小方差估值

$$\min_{\hat{x}}\boldsymbol{J}=E[(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\hat{\boldsymbol{x}})] \tag{4.1.18}$$

的充要条件为

$$\boldsymbol{E}[(\boldsymbol{x}-\hat{\boldsymbol{x}})\boldsymbol{y}^{\mathrm{T}}(t-j)]=\boldsymbol{0},\quad j=0,1,2,\cdots \tag{4.1.19}$$

【定理 4.1.7】 （多维射影公式）$n\times1$ 维随机变量 $\boldsymbol{x}$ 在子 Hilbert 空间 $L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots)$上的射影为

$$\hat{\boldsymbol{x}}=\sum_{i=0}^{\infty}\mathrm{E}[\boldsymbol{x}\boldsymbol{\varepsilon}^{\mathrm{T}}(t-i)]E[\boldsymbol{\varepsilon}(t-i)\boldsymbol{\varepsilon}^{\mathrm{T}}(t-i))]^{-1}\boldsymbol{\varepsilon}(t-i) \tag{4.1.20}$$

【定理 4.1.8】 （多维递推射影公式）

$$\begin{aligned}\mathrm{proj}(\boldsymbol{x}|\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(\mathrm{t}-1),\cdots)&=\mathrm{proj}(\boldsymbol{x}|\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-2),\cdots)+\\&\mathrm{E}[\boldsymbol{x}\boldsymbol{\varepsilon}^{\mathrm{T}}(t)][\mathrm{E}(\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t))]^{-1}\boldsymbol{\varepsilon}(t)\end{aligned} \tag{4.1.21}$$

【定理 4.1.9】 设 $\boldsymbol{x}$ 为 $n\times1$ 随机变量，$\boldsymbol{z}$ 为 $p\times1$ 随机变量，$\boldsymbol{A},\boldsymbol{B}$ 为适当维矩阵，则

$$\begin{aligned}\mathrm{proj}(\boldsymbol{Ax}+\boldsymbol{Bz}|\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)&=\boldsymbol{A}\,\mathrm{proj}(\boldsymbol{x}|\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)+\\&\boldsymbol{B}\,\mathrm{proj}(\boldsymbol{z}|\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)\end{aligned} \tag{4.1.22}$$

【定义 4.1.4】 若由$(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)$张成的 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)$等于由它的新息序列$(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots)$张成的 Hilbert 空间 $L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots)$，即

$$L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)=L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots) \tag{4.1.23}$$

则称新息序列是完全的.

【定理 4.1.10】 若$(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)$的新息序列$(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots)$是完全的，则有

$$\mathrm{proj}(\boldsymbol{x}|\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)=\mathrm{proj}(\boldsymbol{x}|\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots) \tag{4.1.24}$$

【定理 4.1.11】 若随机序列 $y(t)$的新息序列 $\varepsilon(t)$是完全的，则有递推射影公式

$$\mathrm{proj}(\boldsymbol{x}|\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)=\mathrm{proj}(\boldsymbol{x}|\boldsymbol{y}^{\mathrm{T}}(t-1),\boldsymbol{y}^{\mathrm{T}}(t-2),\cdots)+$$

$$\mathrm{E}[\boldsymbol{x}\boldsymbol{\varepsilon}^{\mathrm{T}}(t)][\mathrm{E}(\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t))]^{-1}\boldsymbol{\varepsilon}(t) \tag{4.1.25}$$

注意，因为($\boldsymbol{\varepsilon}(t)$，$\boldsymbol{\varepsilon}(t-1)\cdots$)是正交(不相关)序列，这给计算射影带来很大方便，大大简化了射影计算，而且具有重要理论意义. 递推射影公式(4.1.25)是递推 Kalman 滤波器推导的出发点，它将问题归结为新息的计算.

4.2 单变量 ARMA 过程的 Wiener – Kolmogorov 预报器

这里用时域方法导出 Wiener – Kolmogorov 预报器[1,2].

考虑平稳、可逆的单变量 ARMA 过程 $y(t)$

$$A(q^{-1})y(t) = C(q^{-1})\varepsilon(t) \tag{4.2.1}$$

其中 $\varepsilon(t)$是零均值、方差为 σ_ε^2的白噪声，且

$$A(q^{-1}) = 1 + a_1q^{-1} + \cdots + a_{n_a}q^{-n_a},$$

$$C(q^{-1}) = 1 + c_1q^{-1} + \cdots + c_{n_c}q^{-n_c} \tag{4.2.2}$$

其中 q^{-1}为单位滞后算子. 假设 $A(q^{-1})$和 $C(q^{-1})$是稳定的，因而(4.2.1)是平稳、可逆的，故 $y(t)$是平稳时间序列. 设 $t_0 = -\infty$，即已知无限观测历史($y(t)$，$y(t-1)$，…). 问题是求 $y(t+k)$($k>0$)的线性最小方差预报器 $\hat{y}(t+k|t)$，它是($y(t)$，$y(t-1)$，…)的线性函数，且极小化性能指标

$$J = \mathrm{E}[(y(t+k) - \hat{y}(t+k|t))^2] \tag{4.2.3}$$

由平稳性有

$$y(t) = \frac{C(q^{-1})}{A(q^{-1})}\varepsilon(t) = \sum_{j=0}^{\infty}\alpha_j\varepsilon(t-j) = \alpha(q^{-1})\varepsilon(t) \tag{4.2.4}$$

其中定义级数

$$\alpha(q^{-1}) = \frac{C(q^{-1})}{A(q^{-1})} = \sum_{j=0}^{\infty}\alpha_jq^{-j} \tag{4.2.5}$$

其中系数 α_j 可递推计算为

$$\alpha_j = -a_1\alpha_{j-1} - \cdots - a_{n_a}\alpha_{j-n_a} + c_j \tag{4.2.6}$$

其中规定 $\alpha_0 = 1$，$\alpha_j = 0\,(j<0)$，$c_j = 0\,(j>n_c)$. 由可逆性有

$$\varepsilon(t) = \frac{A(q^{-1})}{C(q^{-1})}y(t) = \sum_{j=0}^{\infty}\beta_jy(t-j) = \beta(q^{-1})y(t) \tag{4.2.7}$$

其中定义级数

$$\beta(q^{-1}) = \frac{A(q^{-1})}{C(q^{-1})} = \sum_{j=0}^{\infty}\beta_j q^{-j} \tag{4.2.8}$$

其中系数 β_j 可递推计算为

$$\beta_j = -c_1\beta_{j-1} - \cdots - c_{n_c}\beta_{j-n_c} + a_j \tag{4.2.9}$$

其中规定 $\beta_0 = 1$，$\beta_j = 0\,(j<0)$，$a_j = 0\,(j>n_a)$.

由(4.2.4)和(4.2.7)引出由($y(t)$，$y(t-1)$，…)生成的 Hilber 空间 $L(y(t)$，$y(t-1)$，…)等于由($\varepsilon(t)$，$\varepsilon(t-1)$，…)生成的 Hilbert 空间，即

$$L(y(t), y(t-1), \cdots) = L(\varepsilon(t), \varepsilon(t-1), \cdots) \tag{4.2.10}$$

注意极小化(4.2.3)的最优预报器(线性最小方差预报器)$\hat{y}(t+k|t)$ $(k>0)$是$y(t+k)$在$L(y(t),y(t-1),\cdots)$上的射影,即

$$\hat{y}(t+k|t)=\mathrm{proj}(y(t+k)|y(t),y(t-1),\cdots)=\mathrm{proj}(y(t+k)|\varepsilon(t),\varepsilon(t-1),\cdots) \tag{4.2.11}$$

由(4.2.4)有

$$y(t+k)=\sum_{j=0}^{\infty}\alpha_j\varepsilon(t+k-j)=$$
$$\varepsilon(t+k)+\alpha_1\varepsilon(t+k-1)+\cdots+\alpha_{k-1}\varepsilon(t+1)+\alpha_k\varepsilon(t)+\alpha_{k+1}\varepsilon(t-1)+\cdots \tag{4.2.12}$$

利用$\varepsilon(t)$的正交(不相关)性有

$$\varepsilon(t+i)\perp L(\varepsilon(t),\varepsilon(t-1),\cdots),\quad i>0 \tag{4.2.13}$$

于是

$$\mathrm{proj}(\varepsilon(t+i)|\varepsilon(t),\varepsilon(t-1),\cdots)=0,\quad i>0 \tag{4.2.14}$$

取(4.2.12)两边各项在Hilbert空间$L(y(t),y(t-1),\cdots)=L(\varepsilon(t),\varepsilon(t-1)\cdots)$上射影引出最优预报器

$$\hat{y}(t+k|t)=\alpha_k\varepsilon(t)+\alpha_{k+1}\varepsilon(t-1)+\cdots=\sum_{j=k}^{\infty}\alpha_j\varepsilon(t+k-j) \tag{4.2.15}$$

最小均方预报误差为

$$\mathrm{E}[(y(t+k)-\hat{y}(t+k|t))^2]=\mathrm{E}\Big[\Big(\sum_{j=0}^{k-1}\alpha_j\varepsilon(t+k-j)\Big)^2\Big]=\sigma_\varepsilon^2\sum_{j=0}^{k-1}\alpha_j^2 \tag{4.2.16}$$

注意用q^k乘(4.2.5)两边有

$$q^k\alpha(q^{-1})=q^k+\alpha_1q^{k-1}+\cdots+\alpha_{k-1}q+\alpha_k+\alpha_{k+1}q^{-1}+\cdots \tag{4.2.17}$$

去掉上式q的正幂项,定义

$$[q^k\alpha(q^{-1})]_-=\alpha_k+\alpha_{k+1}q^{-1}+\alpha_{k+2}q^{-2}+\cdots \tag{4.2.18}$$

则最优预报器$\hat{y}(t+k|t)$可表为

$$\hat{y}(t+k|t)=[q^k\alpha(q^{-1})]_-\varepsilon(t) \tag{4.2.19}$$

应用(4.2.7)有

$$\hat{y}(t+k|t)=[q^k\alpha(q^{-1})]_-\beta(q^{-1})y(t) \tag{4.2.20}$$

由(4.2.5)和(4.2.8)有关系

$$\alpha(q^{-1})\beta(q^{-1})=1 \tag{4.2.21}$$

将(4.2.21)代入(4.2.20)有如下定理.

【定理4.2.1】 (Wiener-Kolmogorov预报公式)平稳可逆的ARMA过程(4.2.1)有超前k步$(k>0)$最优预报器

$$\hat{y}(t+k|t)=\frac{[q^k\alpha(q^{-1})]_-}{\alpha(q^{-1})}y(t) \tag{4.2.22}$$

且最小均方预报误差为(4.2.16).

【例4.2.1】 考虑平稳AR(1)过程$y(t)$,

$$(1-aq^{-1})y(t)=\varepsilon(t) \tag{4.2.23}$$

其中$|a|<1$，$\varepsilon(t)$是零均值、方差为σ_ε^2的白噪声. 求$y(t)$的超前k步预报器$\hat{y}(t+k|t)$.

应用几何级数求和公式有

$$\alpha(q^{-1}) = \frac{1}{1-aq^{-1}} = \sum_{j=0}^{\infty} a^j q^{-j} \tag{4.2.24}$$

$$[q^k \alpha(q^{-1})]_- = a^k + a^{k+1}q^{-1} + \cdots = a^k(1 + aq^{-1} + a^2q^{-2} + \cdots) = a^k \frac{1}{1-aq^{-1}} \tag{4.2.25}$$

于是由(4.2.22)有

$$\hat{y}(t+k|t) = a^k y(t) \tag{4.2.26}$$

且由(4.2.16)有预报误差方差为

$$\mathrm{E}[(y(t+k) - \hat{y}(t+k|t))^2] = \sigma_\varepsilon^2 \sum_{j=0}^{k-1} a^{2j} = \sigma_\varepsilon^2 \frac{1-a^{2k}}{1-a^2} \tag{4.2.27}$$

【例4.2.2】 考虑可逆的MA(1)过程$y(t)$，

$$y(t) = (1 + cq^{-1})\varepsilon(t) \tag{4.2.28}$$

其中$\varepsilon(t)$是零均值、方差为σ_ε^2的白噪声，$|c|<1$. 求最优k步预报器$\hat{y}(t+k|t)$.

注意

$$\alpha(q^{-1}) = 1 + cq^{-1} \tag{4.2.29}$$

故有

$$[q\alpha(q^{-1})]_- = [q + c]_- = c \tag{4.2.30}$$

$$[q^k\alpha(q^{-1})]_- = [q^k + cq^{k-1}]_- = 0, \quad k>1 \tag{4.2.31}$$

于是由(4.2.22)有

$$\hat{y}(t+1|t) = \frac{c}{1+cq^{-1}} y(t) \tag{4.2.32}$$

$$\hat{y}(t+k|t) = 0, \quad k>1 \tag{4.2.33}$$

且由(4.2.16)有预报误差方差

$$\mathrm{E}[(y(t+k) - \hat{y}(t+k|t))^2] = \begin{cases} \sigma_\varepsilon^2 & k=1 \\ \sigma_\varepsilon^2(1+c^2) & k>1 \end{cases} \tag{4.2.34}$$

【例4.2.3】 考虑平稳可逆的ARMA(1,1)过程$y(t)$，

$$(1 - aq^{-1})y(t) = (1 + cq^{-1})\varepsilon(t) \tag{4.2.35}$$

其中$|a|<1$，$|c|<1$，$\varepsilon(t)$是零均值、方差为σ_ε^2的白噪声. 求最优k步预报器$\hat{y}(t+k|t)$.

注意

$$\alpha(q^{-1}) = \frac{1+cq^{-1}}{1-aq^{-1}} \tag{4.2.36}$$

于是有

$$\left[\frac{q^k(1+cq^{-1})}{1-aq^{-1}}\right]_- = \left[q^k \sum_{j=0}^{\infty} a^j q^{-j}\right]_- + \left[q^k\left(cq^{-1}\sum_{j=0}^{\infty} a^j q^{-j}\right)\right]_- =$$

$$(a^k + a^{k+1}q^{-1} + \cdots) + c(a^{k-1} + a^k q^{-1} + \cdots) = (a^k + ca^{k-1})(1 + aq^{-1} + a^2q^{-2} + \cdots) =$$

$$\frac{a^k + ca^{k-1}}{1-aq^{-1}} \tag{4.2.37}$$

故由(4.2.22)和(4.2.36)有

$$\hat{y}(t+k|t)=\frac{a^k+ca^{k-1}}{1-aq^{-1}}\frac{1-aq^{-1}}{1+cq^{-1}}y(t)=\frac{a^{k-1}(a+c)}{1+cq^{-1}}y(t) \tag{4.2.38}$$

由(4.2.36) $\alpha(q^{-1})$ 有展式

$$\alpha(q^{-1})=\frac{1}{1-aq^{-1}}+\frac{cq^{-1}}{1-aq^{-1}}=\sum_{j=0}^{\infty}\alpha_j q^{-j}=$$

$$\sum_{j=0}^{\infty}a^j q^{-j}+cq^{-1}\sum_{j=0}^{\infty}a^j q^{-j} \tag{4.2.39}$$

用比较系数法有

$$\alpha_0=1,\quad \alpha_k=a^{k-1}(a+c),\quad k\geqslant 1 \tag{4.2.40}$$

于是由(4.2.16)有预报误差方差

$$\mathrm{E}[(y(t+1)-\hat{y}(t+1|t))^2]=\sigma_\varepsilon^2$$

$$\mathrm{E}[(y(t+k)-\hat{y}(t+k\mid t))^2]=\sigma_\varepsilon^2\Big[1+\sum_{j=1}^{k-1}a^{2(j-1)}(a+c)^2\Big]=$$

$$\sigma_\varepsilon^2\Big[1+\sum_{i=0}^{k-2}a^{2i}(a+c)^2\Big]=\sigma_\varepsilon^2\Big[1+(a+c)^2\frac{1-a^{2(k-1)}}{1-a^2}\Big] \tag{4.2.41}$$

4.3 单变量 Box – Jenkins 递推预报器

【定理 4.3.1】 对于平稳、可逆的 ARMA 过程(4.2.1)(即假设 $A(q^{-1})$ 和 $C(q^{-1})$ 是稳定的多项式,换言之,多项式 $A(x)$ 和 $C(x)$ 的零点位于单位圆外),白噪声 $\varepsilon(t)$ 是 $y(t)$ 的新息过程,即

$$\varepsilon(t)=y(t)-\hat{y}(t|t-1) \tag{4.3.1}$$

其中 $\hat{y}(t|t-1)$ 是基于观测 $(y(t-1),y(t-2),\cdots)$ 对 $y(t)$ 的一步最优(线性最小方差)预报估值. 因而(4.2.1)是 $y(t)$ 的 ARMA 新息模型.

证明 由平稳性有均方收敛的展式

$$y(t)=\frac{C(q^{-1})}{A(q^{-1})}\varepsilon(t)=\sum_{j=0}^{\infty}\alpha_j\varepsilon(t-j) \tag{4.3.2}$$

其中系数 α_j 由(4.2.6)计算. 由可逆性有展式

$$\varepsilon(t)=\frac{A(q^{-1})}{C(q^{-1})}y(t)=\sum_{j=0}^{\infty}\beta_j y(t-j) \tag{4.3.3}$$

其中系数 β_j 由(4.2.9)计算. 上两式表明

$$y(t)\in L(\varepsilon(t),\varepsilon(t-1),\cdots),$$
$$\varepsilon(t)\in L(y(t),y(t-1),\cdots) \tag{4.3.4}$$

因而 $(y(t),y(t-1),\cdots)$ 与 $(\varepsilon(t),\varepsilon(t-1),\cdots)$ 张成相同的 Hilbert 空间,即

$$L(y(t),y(t-1),\cdots)=L(\varepsilon(t),\varepsilon(t-1),\cdots) \tag{4.3.5}$$

取(4.2.1)两边各项到 Hilbert 空间 $L(y(t-1),y(t-2),\cdots)$ 上的射影,并注意(4.3.5),则有

$$\hat{y}(t|t-1) = -a_1 y(t-1) - \cdots - a_{n_a} y(t-n_a) + c_1 \varepsilon(t-1) + \cdots + c_{n_c} \varepsilon(t-n_c) \tag{4.3.6}$$

由(4.2.1)减(4.3.6)得(4.3.1).其中用到了事实

$$\mathrm{proj}(\varepsilon(t)|y(t-1),y(t-2),\cdots) = \mathrm{proj}(\varepsilon(t)|\varepsilon(t-1),\varepsilon(t-2),\cdots) = 0 \tag{4.3.7}$$

因为 $\varepsilon(t) \perp \varepsilon(t-i), i>0$,从而有 $\varepsilon(t) \perp L(\varepsilon(t-1),\varepsilon(t-2),\cdots)$.证毕. □

现在推导 Box - Jenkins 递推预报器.在(4.2.1)中置 $t = t+k$ 有

$$y(t+k) = -\sum_{i=1}^{n_a} a_i y(t+k-i) + \sum_{i=0}^{n_c} c_i \varepsilon(t+k-i), \quad c_0 = 1 \tag{4.3.8}$$

取上式两边各项在 Hilbert 空间 $L(y(t),y(t-1),\cdots)$上的射影,应用射影性质和(4.3.5)有

$$\begin{aligned}
&\mathrm{proj}(y(t+k-i)|y(t),y(t-1),\cdots) = y(t+k-i) && (t+k-i \leqslant t), \\
&\mathrm{proj}(\varepsilon(t+k-i)|y(t),y(t-1),\cdots) = \varepsilon(t+k-i) && (t+k-i \leqslant t), \\
&\mathrm{proj}(\varepsilon(t+k-i)|y(t),y(t-1),\cdots) = 0 && (t+k-i > t)
\end{aligned} \tag{4.3.9}$$

进而可得如下 Box - Jenkins 递推预报器.

【定理 4.3.2】 平稳、可逆的 ARMA 过程(4.2.1)有 Box - Jenkins 递推预报器

$$\hat{y}(t+k|t) = -\sum_{i=1}^{n_a} a_i \hat{y}(t+k-i|t) + \sum_{i=k}^{n_c} c_i \varepsilon(t+k-i), \quad k \leqslant n_c \tag{4.3.10}$$

$$\hat{y}(t+k|t) = -\sum_{i=1}^{n_a} a_i \hat{y}(t+k-i|t), \quad k > n_c \tag{4.3.11}$$

其中规定 $\hat{y}(t+k-i|t) = y(t+k-i)\ (t+k-i \leqslant t)$.

上述 Box - Jenkins 递推预报器可统一写为

$$A(\tilde{q}^{-1})\hat{y}(t+k|t) = C(q^{-1})\varepsilon(t+k) \tag{4.3.12}$$

其中规定 $\varepsilon(t+j) = 0\ (j>0)$,且定义 $A(\tilde{q}^{-1})$是只对 $\hat{y}(t+k|t)$的时标$(t+k)$运算的 $A(q^{-1})$,即

$$A(\tilde{q}^{-1})\hat{y}(t+k|t) = \hat{y}(t+k|t) + a_1\hat{y}(t+k-1|t) + \cdots + a_{n_a}\hat{y}(t+k-n_a|t) \tag{4.3.13}$$

【例 4.3.1】 考虑平稳 AR(1)过程 $y(t)$,

$$(1-aq^{-1})y(t) = \varepsilon(t) \tag{4.3.14}$$

其中$|a|<1$,$\varepsilon(t)$是零均值、方差为 σ_ε^2的白噪声.求最优 Box - Jenkins 预报器 $\hat{y}(t+k|t)$.

应用定理 4.3.2 有

$$\hat{y}(t+1|t) = ay(t) \tag{4.3.15}$$

$$\hat{y}(t+k|t) = a\hat{y}(t+k-1|t) \quad (k>1) \tag{4.3.16}$$

这引出

$$\hat{y}(t+k|t) = a^k y(t) \tag{4.3.17}$$

它相同于例 4.2.1 的结果.

【例 4.3.2】 考虑可逆的 MA(1)过程 $y(t)$,

$$y(t) = (1+cq^{-1})\varepsilon(t) \tag{4.3.18}$$

其中$|c|<1$,$\varepsilon(t)$是零均值、方差为 σ_ε^2的白噪声.求 Box - Jenkins 预报器 $\hat{y}(t+k|t)$.

应用定理 4.3.2 有 Box – Jenkins 预报器

$$\hat{y}(t+1|t) = c\,\varepsilon(t) \tag{4.3.19}$$

$$\hat{y}(t+k|t) = 0, \quad k > 1 \tag{4.3.20}$$

注意由(4.3.18)有

$$\varepsilon(t) = \frac{1}{1 + cq^{-1}} y(t) \tag{4.3.21}$$

将它代入(4.3.19)得

$$\hat{y}(t+1|t) = \frac{c}{1 + cq^{-1}} y(t) \tag{4.3.22}$$

它相同于例 4.2.2 的结果.

【例 4.3.3】 考虑平稳、可逆的 ARMA(1,1)过程 $y(t)$,

$$(1 - aq^{-1}) y(t) = (1 + cq^{-1}) \varepsilon(t) \tag{4.3.23}$$

其中$|a| < 1$,$|c| < 1$,$\varepsilon(t)$是零均值、方差为 σ_ε^2的白噪声. 求 Box – Jenkins 预报器 $\hat{y}(t+k|t)$.

应用定理 4.3.2 有 Box – Jenkins 递推预报器

$$\hat{y}(t+1|t) = ay(t) + c\,\varepsilon(t) \tag{4.3.24}$$

$$\hat{y}(t+k|t) = a\hat{y}(t+k-1|t), \quad k > 1 \tag{4.3.25}$$

这引出它的非递推形式为

$$\hat{y}(t+k|t) = a^{k-1}[ay(t) + c\,\varepsilon(t)], \quad k \geqslant 1 \tag{4.3.26}$$

由(4.3.23)有

$$\varepsilon(t) = \frac{1 - aq^{-1}}{1 + cq^{-1}} y(t) \tag{4.3.27}$$

将它代入(4.3.26)有

$$\hat{y}(t+k|t) = \frac{a^{k-1}(a+c)}{1 + cq^{-1}} y(t) \tag{4.3.28}$$

它相同于例 4.2.3 的结果.

4.4 单变量 Åström 预报方法

【例 4.4.1】 一个启发性例子.

考虑平稳、可逆的 ARMA(1,1)过程 $y(t)$,

$$(1 - aq^{-1}) y(t) = (1 + cq^{-1}) \varepsilon(t) \tag{4.4.1}$$

其中$|a| < 1$,$|c| < 1$,$\varepsilon(t)$是零均值、方差为 σ_ε^2的白噪声. 最优预报器 $\hat{y}(t+1|t)$和 $\hat{y}(t+2|t)$.

(4.4.1)可写成传递函数形式

$$y(t+1) = \frac{1 + cq^{-1}}{1 - aq^{-1}} \varepsilon(t+1) \tag{4.4.2}$$

为了预报 $y(t+1)$,用综合除法将 $y(t+1)$分解为两部分:一部分是已知的,一部分是未知的.

$$\frac{1+cq^{-1}}{1-aq^{-1}}=\frac{1-aq^{-1}+aq^{-1}+cq^{-1}}{1-aq^{-1}} \tag{4.4.3}$$

即有分解式

$$\frac{1+cq^{-1}}{1-aq^{-1}}=1+\frac{(c+a)q^{-1}}{1-aq^{-1}} \tag{4.4.4}$$

将其代入(4.4.2)可将 $y(t+1)$ 分解为

$$y(t+1)=\varepsilon(t+1)+\frac{c+a}{1-aq^{-1}}\varepsilon(t) \tag{4.4.5}$$

而 $\varepsilon(t)$ 是已知的,可由 $(y(t), y(t-1), \cdots)$ 计算为

$$\varepsilon(t)=\frac{1-aq^{-1}}{1+cq^{-1}}y(t) \tag{4.4.6}$$

将它代入(4.4.5)有分解

$$y(t+1)=\varepsilon(t+1)+\frac{c+a}{1+cq^{-1}}y(t) \tag{4.4.7}$$

取上式两边各项在 Hilbert 空间 $L(y(t), y(t-1), \cdots)=L(\varepsilon(t), \varepsilon(t-1), \cdots)$ 上的射影引出最优预报器

$$\hat{y}(t+1|t)=\frac{c+a}{1+cq^{-1}}y(t) \tag{4.4.8}$$

预报误差方差为

$$\mathrm{E}[(y(t+1)-\hat{y}(t+1|t))^2]=\mathrm{E}[\varepsilon^2(t+1)]=\sigma_\varepsilon^2 \tag{4.4.9}$$

类似地用综合除法进一步由(4.4.4)有

$$\frac{1+cq^{-1}}{1-aq^{-1}}=1+(c+a)q^{-1}\frac{1-aq^{-1}+aq^{-1}}{1-aq^{-1}}=1+(c+a)q^{-1}\left(1+\frac{aq^{-1}}{1-aq^{-1}}\right) \tag{4.4.10}$$

因而有分解式

$$\frac{1+cq^{-1}}{1-aq^{-1}}=1+(c+a)q^{-1}+\frac{a(c+a)q^{-2}}{1-aq^{-1}} \tag{4.4.11}$$

于是由(4.4.2) $y(t+2)$ 可分解为

$$y(t+2)=\varepsilon(t+2)+(c+a)\varepsilon(t+1)+\frac{a(c+a)}{1-aq^{-1}}\varepsilon(t) \tag{4.4.12}$$

应用(4.4.6)有分解式

$$y(t+2)=\varepsilon(t+2)+(c+a)\varepsilon(t+1)+\frac{a(c+a)}{1+cq^{-1}}y(t) \tag{4.4.13}$$

取上式两边各项在 Hilbert 空间 $L(y(t), y(t-1), \cdots)$ 上的射影引出二步最优预报器

$$\hat{y}(t+2|t)=\frac{a(c+a)}{1+cq^{-1}}y(t) \tag{4.4.14}$$

且有预报误差方差

$$\mathrm{E}[(y(t+2)-\hat{y}(t+2|t))^2]=\sigma_\varepsilon^2[1+(c+a)^2] \tag{4.4.15}$$

在一般情形下有如下 Åström 预报器.

【定理 4.4.1】 (Åström 预报器)平稳、可逆的 ARMA 过程(4.2.1)超前 k 步 Åström 最优预报器

$$\hat{y}(t+k\mid t)=\frac{G_k(q^{-1})}{C(q^{-1})}y(t) \tag{4.4.16}$$

它有递推形式

$$C(q^{-1})\hat{y}(t+k\mid t)=G_k(q^{-1})y(t) \tag{4.4.17}$$

其中 $G_k(q^{-1})$连同 $F_k(q^{-1})$由如下 Diophantine 方程决定

$$C(q^{-1})=A(q^{-1})F_k(q^{-1})+q^{-k}G_k(q^{-1}) \tag{4.4.18}$$

其中 $F_k(q^{-1})$为$(k-1)$阶多项式,$G_k(q^{-1})$的阶次为 $n_g=\max(n_a-1,n_c-k)$,

$$F_k(q^{-1})=f_{k0}+f_{k1}q^{-1}+\cdots+f_{k,k-1}q^{-(k-1)},$$

$$G_k(q^{-1})=g_{k0}+g_{k1}q^{-1}+\cdots+g_{k,n_g}q^{-n_g} \tag{4.4.19}$$

预报误差为

$$\tilde{y}(t+k\mid t)=y(t+k)-\hat{y}(t+k\mid t)=F_k(q^{-1})\varepsilon(t+k) \tag{4.4.20}$$

且有误报误差方差为

$$\mathrm{E}[(\tilde{y}(t+k\mid t))^2]=\sigma_\varepsilon^2\sum_{i=0}^{k-1}f_{ki}{}^2 \tag{4.4.21}$$

证明　由(4.2.1)有

$$y(t+k)=\frac{C(q^{-1})}{A(q^{-1})}\varepsilon(t+k) \tag{4.4.22}$$

用综合除法可将 $C(q^{-1})/A(q^{-1})$分解为

$$\frac{C(q^{-1})}{A(q^{-1})}=F_k(q^{-1})+q^{-k}\frac{G_k(q^{-1})}{A(q^{-1})} \tag{4.4.23}$$

用 $A(q^{-1})$乘上式两边引出 Diophantine 方程

$$C(q^{-1})=A(q^{-1})F_k(q^{-1})+q^{-k}G_k(q^{-1}) \tag{4.4.24}$$

注意 $F_k(q^{-1})$是$(k-1)$阶多项式和 $G_k(q^{-1})=q^k[C(q^{-1})-A(q^{-1})F_k(q^{-1})]$,则有 $n_g=\max(n_c-k,n_a+(k-1)-k)=\max(n_c-k,n_a-1)$.系数 f_{ki}和 g_{ki}可由 Diophantine 方程(4.4.24)用比较系数法求得.于是 $y(t+k)$可分解为

$$y(t+k)=\left[F_k(q^{-1})+q^{-k}\frac{G_k(q^{-1})}{A(q^{-1})}\right]\varepsilon(t+k)=F_k(q^{-1})\varepsilon(t+k)+\frac{G_k(q^{-1})}{A(q^{-1})}\varepsilon(t) \tag{4.4.25}$$

由(4.2.1)和可逆性有

$$\varepsilon(t)=\frac{A(q^{-1})}{C(q^{-1})}y(t) \tag{4.4.26}$$

由上两式引出分解式

$$y(t+k)=F_k(q^{-1})\varepsilon(t+k)+\frac{G_k(q^{-1})}{C(q^{-1})}y(t) \tag{4.4.27}$$

因 $F_k(q^{-1})\varepsilon(t+k)=f_{k0}\varepsilon(t+k)+f_{k1}\varepsilon(t+k-1)+\cdots+f_{k,k-1}\varepsilon(t+1)$是未知的,取(4.4.27)两边各项在 Hilbert 空间 $L(y(t),y(t-1),\cdots)$上的射影,有 Åström 最优预报器

$$\hat{y}(t+k\mid t)=\frac{G_k(q^{-1})}{C(q^{-1})}y(t) \tag{4.4.28}$$

其中用到了事实 $L(y(t),y(t-1),\cdots)=L(\varepsilon(t),\varepsilon(t-1),\cdots)$及事实

$$\frac{G_k(q^{-1})}{C(q^{-1})}y(t)\in L(y(t),y(t-1),\cdots) \tag{4.4.29}$$

因而它是已知的.故有

$$\operatorname{proj}\left(\frac{G_k(q^{-1})}{C(q^{-1})}y(t)\mid y(t),y(t-1),\cdots\right)=\frac{G_k(q^{-1})}{C(q^{-1})}y(t) \tag{4.4.30}$$

由(4.4.27)减(4.4.28)引出预报误差

$$\tilde{y}(t+k\mid t)=F_k(q^{-1})\varepsilon(t+k) \tag{4.4.31}$$

这引出预报误差方差

$$\mathrm{E}[(\tilde{y}(t+k\mid t))^2]=\mathrm{E}[(f_{k0}\varepsilon(t+k)+\cdots+f_{k,k-1}\varepsilon(t+1))^2]=\sigma_\varepsilon^2\sum_{i=0}^{k-1}f_{ki}^2 \tag{4.4.32}$$

证毕. □

在具体例子中,在不引起混淆情形下,简记 f_{ki} 为 f_i,g_{ki} 为 g_i.

【例 4.4.2】 对例 4.4.1 求 Åström 预报器 $\hat{y}(t+k\mid t)$,$k\geqslant 1$.

因 $n_a=1$,$n_c=1$,则有 $n_g=\max(0,1-k)=0$,故 $G(q^{-1})=g_0$. Diophantine 方程(4.4.24)成为

$$1+cq^{-1}=(1-aq^{-1})(f_0+f_1q^{-1}+\cdots+f_{k-1}q^{-(k-1)})+q^{-k}g_0 \tag{4.4.33}$$

用比较系数法,比较上式两边 q^{-i}的系数有关系

$$f_0=1,\quad c=f_1-af_0,\quad 0=f_i-af_{i-1},\quad i>1;\quad 0=g_0-af_{k-1} \tag{4.4.34}$$

这引出

$$f_0=1,\quad f_i=a^{i-1}(c+a),\quad i=1,\cdots,k-1;\quad g_0=a^{k-1}(c+a) \tag{4.4.35}$$

于是由(4.4.16)有 Åström 最优预报器

$$\hat{y}(t+k\mid t)=\frac{a^{k-1}(c+a)}{1+cq^{-1}}y(t) \tag{4.4.36}$$

它相同于例 4.3.3 的结果.预报误差方差为

$$\mathrm{E}[(y(t+k)-\hat{y}(t+k\mid t))^2]=\sigma_\varepsilon^2\left[1+\sum_{i=1}^{k-1}a^{2(i-1)}(c+a)^2\right]=$$

$$\sigma_\varepsilon^2\left[1+(c+a)^2\frac{1-a^{2(k-1)}}{1-a^2}\right] \tag{4.4.37}$$

它相同于例 4.2.3 的结果(4.2.41).

【例 4.4.3】 考虑平稳、可逆的 ARMA(1,2)过程,

$$(1+aq^{-1})y(t)=(1+c_1q^{-1}+c_2q^{-2})\varepsilon(t) \tag{4.4.38}$$

其中 $e(t)$是零均值、方差为 σ_ε^2的白噪声.下面用三种方法求它的最优预报器 $\hat{y}(t+1\mid t)$.

先用 Åström 方法.注意 $n_a=1$,$n_c=2$,$k=1$,故有 $n_g=\max(0,2-1)=1$,置 $F_k(q^{-1})=f_0$,$G_k(q^{-1})=g_0+g_1q^{-1}$,有 Diophantine 方程

$$1+c_1q^{-1}c_2q^{-2}=(1+aq^{-1})f_0+q^{-1}(g_0+g_1q^{-1}) \tag{4.4.39}$$

用比较系数法有

$$f_0=1,\quad g_0=c_1-a,\quad g_1=c_2 \tag{4.4.40}$$

从而有 Åström 预报器

$$\hat{y}(t+k|t)=\frac{c_1-a_1+c_2q^{-1}}{1+c_1q^{-1}+c_2q^{-2}}y(t) \tag{4.4.41}$$

其次用 Box – Jenkins 方法.由(4.4.38)有 Box – Jenkins 递推预报器

$$\hat{y}(t+1|t)=-ay(t)+c_1\varepsilon(t)+c_2\varepsilon(t-1) \tag{4.4.42}$$

由(4.4.38)有

$$\varepsilon(t)=\frac{1+aq^{-1}}{1+c_1q^{-1}+c_2q^{-2}}y(t) \tag{4.4.43}$$

将它代入(4.4.42)整理后可得(4.4.41).

最后用 Wiener – Kolmogorov 预报器也可得到相同结果,详细从略.

【例 4.4.4】 考虑平稳、可逆的 ARMA 过程 $y(t)$,

$$A(q^{-1})y(t)=C(q^{-1})\varepsilon(t) \tag{4.4.44}$$

其中 $\varepsilon(t)$是零均值、方差为 σ_ε^2的白噪声,$A(q^{-1})$和 $C(q^{-1})$由(4.2.2)定义.考虑由 $y(t)$生成的信号 $z(t)$,

$$z(t)=\frac{P(q^{-1})}{Q(q^{-1})}y(t) \tag{4.4.45}$$

其中 $P(q^{-1})$和 $Q(q^{-1})$为首系数为 1 的 q^{-1}的稳定的多项式.问题是基于观测($y(t)$,$y(t-1)$,…)求 $z(t+k)$的最优预报器 $\hat{z}(t+k|t)$.

将(4.4.45)代入(4.4.44)有 $z(t)$服从的平稳、可逆的 ARMA 模型

$$A(q^{-1})Q(q^{-1})z(t)=P(q^{-1})C(q^{-1})\varepsilon(t) \tag{4.4.46}$$

应用定理 4.4.1 有 Åström 预报器

$$\hat{z}(t+k|t)=\frac{G_k(q^{-1})}{P(q^{-1})C(q^{-1})}z(t) \tag{4.4.47}$$

其中 $G_k(k)$和 $F_k(q^{-1})$由如下 Diophantine 方程决定

$$\frac{P(q^{-1})C(q^{-1})}{A(q^{-1})Q(q^{-1})}=F_k(q^{-1})+q^{-k}\frac{G_k(q^{-1})}{A(q^{-1})Q(q^{-1})} \tag{4.4.48}$$

其中 $F_k(q^{-1})$的阶次为$(k-1)$,$n_g=(n_a+n_q-1,n_p+n_c-k)$.将(4.4.45)代入(4.4.47)有最优预报器

$$\hat{z}(t+k|t)=\frac{G_k(q^{-1})}{C(q^{-1})Q(q^{-1})}y(t) \tag{4.4.49}$$

预报误差方差为

$$\mathrm{E}[(z(t+k)-\hat{z}(t+k|t))^2]=\sigma_\varepsilon^2\sum_{i=0}^{k-1}f_{ki}{}^2 \tag{4.4.50}$$

其中 $F_k(q^{-1})=f_{k0}+f_{k1}q^{-1}+\cdots+f_{k,k-1}q^{-(k-1)}$.

【注 4.4.1】 注意节 4.2 ~ 节 4.4 介绍的三种最优预报方法是在假设 ARMA 过程是平稳可逆条件下用射影理论推导的.可以证明[17]对非平稳 ARMA 过程所得结果也成立.

现在我们给出不同于文献[17]的非平稳 ARMA 过程的 Wiener 预报器的推导和证明.

4.5 非平稳 ARMA 过程的 Wiener 预报器

周知,与经典 Wiener 滤波方法相比,Kalman 滤波方法的优点之一是可处理非平稳随

机过程.本节我们将用 Kalman 滤波方法对非平稳 ARMA 过程导出 Wiener 预报器,它具有 Åström 预报器的形式,并且给出计算预报器的算子多项式 $G_k(q^{-1})$ 的新算法.

考虑非平稳 ARMA 过程 $y(t)$

$$A(q^{-1})y(t)=C(q^{-1})e(t) \tag{4.5.1}$$

其中 $e(t)$ 是零均值、方差为 σ_e^2 的白噪声,$A(q^{-1})$ 是不稳定的多项式,$C(q^{-1})$ 是稳定的多项式,且

$$A(q^{-1})=1+a_1q^{-1}+\cdots+a_{n_a}q^{-n_a}$$

$$C(q^{-1})=1+c_1q^{-1}+\cdots+c_{n_c}q^{-n_c}$$

设 $(A(q^{-1}),C(q^{-1}))$ 互质,且初始观测时刻 $t_0=0$,问题是当 $t\to\infty$ 时求稳态 Wiener 预报器 $\hat{y}(t+k|t)$.等价地,当 $t_0\to-\infty$ 时,求稳态 Wiener 预报器 $\hat{y}(t+k|t)$,它具有以 $y(t)$ 作为输入的传递函数表达式.

由定理 1.6.1,(4.5.1)有状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{\Gamma}e(t) \tag{4.5.2}$$

$$y(t)=\boldsymbol{H}\boldsymbol{x}(t)+e(t) \tag{4.5.3}$$

其中 $(\boldsymbol{A},\boldsymbol{H})$ 为伴随形,$c_i=0\,(i>n_c)$,$a_i=0\,(i>n_a)$,且

$$\boldsymbol{A}=\begin{bmatrix}-a_1 & & & \\ \vdots & & \boldsymbol{I}_{n-1} & \\ -a_n & 0 & \cdots & 0\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}c_1-a_1\\ \vdots\\ c_n-a_n\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0\cdots 0] \tag{4.5.4}$$

其中 $n=\max(n_a,n_c)$.注意系统(4.5.2)和(4.5.3)是带相关的观测噪声 $e(t)$ 和输入噪声 $e(t)$ 的系统,它可化为等价的带不相关噪声系统[14],其状态转移阵 $\bar{\boldsymbol{\Phi}}=\boldsymbol{A}-\boldsymbol{J}\boldsymbol{H}$,$\boldsymbol{J}=\boldsymbol{\Gamma}$,即

$$\bar{\boldsymbol{\Phi}}=\begin{bmatrix}-c_1 & & & \\ \vdots & & \boldsymbol{I}_{n-1} & \\ -c_n & 0 & \cdots & 0\end{bmatrix} \tag{4.5.5}$$

注意 $\bar{\boldsymbol{\Phi}}$ 是稳定矩阵,这是因为 $C(q^{-1})$ 被假定是稳定的多项式且有关系 $\det(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{\Phi}})=C(q^{-1})$.因 $(\bar{\boldsymbol{\Phi}},\boldsymbol{H})$ 为完全可观对,这是因为 $(\bar{\boldsymbol{\Phi}},\boldsymbol{H})$ 为伴随形,因此系统(4.5.2)和(4.5.3)存在稳态 Kalman 预报器[15,16]

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{A}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\varepsilon(t) \tag{4.5.6}$$

$$y(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1)+\varepsilon(t) \tag{4.5.7}$$

其中白噪声 $\varepsilon(t)$ 是 $y(t)$ 的新息过程.我们在下面将证明 $\boldsymbol{K}_p=\boldsymbol{\Gamma}$.

现在我们揭示当 t_0 为有限时,由 ARMA 模型(4.5.1)取初值 $(y(t_0),\cdots,y(t_0+n-1),e(t_0),\cdots,e(t_0+n-1))$ 递推计算 $e(t)$,

$$e(t)=A(q^{-1})y(t)-c_1e(t-1)-\cdots-c_{n_c}e(t-n_c),\quad t=t_0+n,t_0+n+1,\cdots \tag{4.5.8}$$

其中 $n=\max(n_a,n_c)$,与由等价的状态空间模型(4.5.2)和(4.5.3)取初值 $x(t_0)$ 计算的 $e(t)$,

$$\boldsymbol{x}(t+1)=\boldsymbol{A}x(t)+\boldsymbol{\Gamma}e(t) \tag{4.5.9}$$

$$e(t)=y(t)-Hx(t),\quad t=t_0+1,t_0+2,\cdots \tag{4.5.10}$$

有何关系？以及与由状态空间新息模型(4.5.6)和(4.5.7)取初值 $\hat{\boldsymbol{x}}(t_0|t_0-1)$ 计算的新息 $\varepsilon(t)$

$$\hat{\boldsymbol{x}}(\boldsymbol{t}+1|\boldsymbol{t})=\boldsymbol{A}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\varepsilon(t),\quad \boldsymbol{K}_p=\boldsymbol{\Gamma} \tag{4.5.11}$$

$$\varepsilon(t)=y(t)-\boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1),\quad t=t_0+1,\quad t_0+2,\cdots \tag{4.5.12}$$

有何关系，事实上，由(4.5.2)和(4.5.3)迭代有关系

$$\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{HA}\\ \vdots\\ \boldsymbol{HA}^{n-1}\end{bmatrix}\boldsymbol{x}(t_0)=\begin{bmatrix}y(t_0)-e(t_0)\\ y(t_0+1)-\boldsymbol{H\Gamma}e(t_0)-e(t_0+1)\\ \vdots\\ y(t_0+n-1)-\sum_{i=0}^{n-2}\boldsymbol{HA}^{n-2-i}\boldsymbol{\Gamma}e(t_0+i)-e(t_0+n-1)\end{bmatrix} \tag{4.5.13}$$

因$(\boldsymbol{A},\boldsymbol{H})$为伴随形，则$(\boldsymbol{A},\boldsymbol{H})$为完全可观对，从而可解出 $\boldsymbol{x}(t_0)$与(4.5.1)初值关系

$$\boldsymbol{x}(t_0)=\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{HA}\\ \vdots\\ \boldsymbol{HA}^{n-1}\end{bmatrix}^{-1}\begin{bmatrix}y(t_0)-e(t_0)\\ y(t_0+1)-\boldsymbol{H\Gamma}e(t_0)-e(t_0+1)\\ \vdots\\ y(t_0+n-1)-\sum_{i=0}^{n-2}\boldsymbol{HA}^{n-2-i}\boldsymbol{\Gamma}e(t_0+i)-e(t_0+n-1)\end{bmatrix} \tag{4.5.14}$$

由此我们得出结论：当状态空间模型(4.5.2)和(4.5.3)的初值 $\boldsymbol{x}(t_0)$由 ARMA 模型(4.5.1)的初值$(y(t_0),\cdots,y(t_0+n-1),e(t_0),\cdots,e(t_0+n-1))$通过公式(4.5.14)计算时，则两种模型产生相同的 $e(t)$.

类似地，当取状态空间新息模型(4.5.6)和(4.5.7)的初值 $\hat{\boldsymbol{x}}(t_0|t_0-1)$为由(4.5.14)计算的 $\boldsymbol{x}(t_0)$时，即取

$$\hat{\boldsymbol{x}}(t_0|t_0-1)=\boldsymbol{x}(t_0) \tag{4.5.15}$$

其中 $\boldsymbol{x}(t_0)$由(4.5.14)计算，则由(4.5.11)和(4.5.12)(即(4.5.6)和(4.5.7))计算的非稳态新息 $\varepsilon(t)$，恒同于由前两种模型计算的 $e(t)$，即

$$\varepsilon(t)=e(t),\quad t=t_0,t_0+1,\cdots \tag{4.5.16}$$

因此我们可从 Kalman 滤波观点对非平稳 ARMA 过程(4.5.1)中的白噪声 $e(t)$意义给出如下解释：在(4.5.1)中取初值$(y(t_0),\cdots,y(t_0+n-1),e(t_0),\cdots,e(t_0+n-1))$通过(4.5.8)计算的 $e(t)$是 $y(t)$的非稳态新息过程(4.5.16). 非稳态新息 $e(t)$具有时变方差 $Q_e(t)=\mathrm{E}[e^2(t)]$等于由状态空间新息模型(4.5.11)和(4.5.12)计算的非稳态新息 $\varepsilon(t)$的时变方差 $Q_\varepsilon(t)=\mathrm{E}[\varepsilon^2(t)]$. 将(4.5.3)代入(4.5.12)引出

$$\varepsilon(t)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t|t-1)+e(t) \tag{4.5.17}$$

记 $\boldsymbol{P}(t|t-1)=\mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1)]$，$\tilde{\boldsymbol{x}}(t|t-1)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t|t-1)$，由(4.5.2)有 $e(t)\perp\boldsymbol{x}(t)$，$e(t)\perp\tilde{\boldsymbol{x}}(t|t-1)$，则有关系

$$Q_e(t)=Q_\varepsilon(t)=\boldsymbol{HP}(t|t-1)\boldsymbol{H}^{\mathrm{T}}+\sigma_e^2 \tag{4.5.18}$$

因$(\bar{\boldsymbol{\Phi}},\boldsymbol{H})$为伴随形，则$(\bar{\boldsymbol{\Phi}},\boldsymbol{H})$为完全可观对，因而存在稳态 Kalman 预报器(4.5.6)和(4.

5.7)，且由(3.8.20)稳态 Riccati 方程有

$$\boldsymbol{\Sigma}=\overline{\boldsymbol{\Phi}}\left[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\sigma_e^2)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}\right]\overline{\boldsymbol{\Phi}}^{\mathrm{T}}+\boldsymbol{\Gamma}(\sigma_e^2-S\sigma_e^{-2}S)\boldsymbol{\Gamma}^{\mathrm{T}} \tag{4.5.19}$$

因在(4.5.2)和(4.5.3)中输入噪声 $e(t)$和观测噪声 $e(t)$是相关的，即

$$S=\mathrm{E}\left[e(t)e(t)\right]=\sigma_e^2 \tag{4.5.20}$$

这引出(4.5.19)右边的第二项为零，从而稳态 Riccati 方程化为

$$\boldsymbol{\Sigma}=\overline{\boldsymbol{\Phi}}\left[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\sigma_e^2)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}\right]\overline{\boldsymbol{\Phi}} \tag{4.5.21}$$

它有非负定稳态解

$$\boldsymbol{\Sigma}=\mathbf{0} \tag{4.5.22}$$

于是由(3.8.8)有稳态 Kalman 预报器增益 $\boldsymbol{K}_p$ 为

$$\boldsymbol{K}_p=(\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{\Gamma}S)(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}+\sigma_e^2)^{-1}=\boldsymbol{\Gamma} \tag{4.5.23}$$

其中应用了 $\boldsymbol{\Sigma}=\mathbf{0}$. 由(4.5.2)和(4.5.3)有

$$\boldsymbol{x}(t+1)=(\boldsymbol{A}-\boldsymbol{\Gamma}\boldsymbol{H})\boldsymbol{x}(t)+\boldsymbol{\Gamma}y(t) \tag{4.5.24}$$

由(4.5.6)和(4.5.7)有

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{A}-\boldsymbol{K}_p\boldsymbol{H})\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_py(t) \tag{4.5.25}$$

注意 $\boldsymbol{K}_p=\boldsymbol{\Gamma}$，上两式相减有预报误差

$$\tilde{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\tilde{\boldsymbol{x}}(t|t-1) \tag{4.5.26}$$

其中由(4.5.23)$\boldsymbol{K}_p=\boldsymbol{\Gamma}$ 引出

$$\boldsymbol{\Psi}_p=\boldsymbol{A}-\boldsymbol{K}_p\boldsymbol{H}=\boldsymbol{A}-\boldsymbol{\Gamma}\boldsymbol{H} \tag{4.5.27}$$

于是由(4.5.26)有

$$\boldsymbol{P}(t+1|t)=\boldsymbol{\Psi}_p\boldsymbol{P}(t|t-1)\boldsymbol{\Psi}_p^{\mathrm{T}} \tag{4.5.28}$$

注意

$$\boldsymbol{\Psi}_p=\boldsymbol{A}-\boldsymbol{\Gamma}\boldsymbol{H}=\begin{bmatrix}-c_1 & & & \\ \vdots & & \boldsymbol{I}_{n-1} & \\ -c_n & 0 & \cdots & 0\end{bmatrix} \tag{4.5.29}$$

则有 $\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)=C(q^{-1})$，因 $C(q^{-1})$是稳定的多项式，则 $\boldsymbol{\Psi}_p$ 是稳定矩阵. 由(4.5.26)迭代$(t-t_0)$次有

$$\boldsymbol{P}(t|t-1)=\boldsymbol{\Psi}_p^{t-t_0}\boldsymbol{P}(t_0|t_0-1)\boldsymbol{\Psi}_p^{(t-t_0)\mathrm{T}} \tag{4.5.30}$$

由 $\boldsymbol{\Psi}_p$ 的稳定性引出

$$\lim_{t_0\to-\infty}\boldsymbol{P}(t|t-1)=\mathbf{0} \tag{4.5.31}$$

于是由(4.5.18)有

$$\sigma_\varepsilon^2\triangleq\lim_{t_0\to-\infty}Q_\varepsilon(t)=\lim_{t_0\to-\infty}Q_e(t)=\sigma_e^2 \tag{4.5.32}$$

其中 σ_ε^2 和 σ_e^2 为稳态新息方差.

当 t_0 为有限时刻时，由递推射影公式有最优 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+k|t)=\hat{\boldsymbol{x}}(t+k|t-1)+\mathrm{E}\left[\boldsymbol{x}(t+k)\varepsilon(t)\right]Q_\varepsilon^{-1}(t)\varepsilon(t),\quad k>0 \tag{4.5.33}$$

由(4.5.2)迭代有

$$\boldsymbol{x}(t+k)=\boldsymbol{A}^k\boldsymbol{x}(t)+\sum_{i=t+1}^{t+k}\boldsymbol{A}^{t+k-i}\boldsymbol{\Gamma}e(i-1) \tag{4.5.34}$$

注意关系(4.5.17),因为 $e(t)$ 是白噪声,$\boldsymbol{x}(t)=\tilde{\boldsymbol{x}}(t|t-1)+\hat{x}(t|t-1)$,这引出非稳态关系

$$\mathrm{E}[\boldsymbol{x}(t+k)\varepsilon(t)]=\boldsymbol{A}^{k-1}\boldsymbol{\Gamma}\sigma_e^2+\boldsymbol{A}^k\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}} \tag{4.5.35}$$

令 $t_0\to-\infty$,注意稳态新息 $\varepsilon(t)=e(t)$,并注意(4.5.31)和(4.5.32),因而当 $t_0\to-\infty$ 时,引出稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{A}\hat{\boldsymbol{x}}(t+k-1|t-1)+\boldsymbol{A}^{k-1}\boldsymbol{\Gamma}e(t),$$
$$\hat{y}(t+k|t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t+k|t) \tag{4.5.36}$$

其中第二式由(4.5.3)和射影性质得到.因为($\boldsymbol{A}$,$\boldsymbol{H}$)为伴随形,应用定理 1.6.1 有等价的稳态 ARMA 预报器

$$A(q^{-1})\hat{y}(t+k|t)=\bar{G}_k(q^{-1})e(t+1) \tag{4.5.37}$$

即有稳态 k 步预报器

$$A(q^{-1})\hat{y}(t+k|t)=G_k(q^{-1})e(t) \tag{4.5.38}$$

其中定义

$$\boldsymbol{A}^{k-1}\boldsymbol{\Gamma}=\begin{bmatrix}g_{k1}\\ \vdots\\ g_{kn}\end{bmatrix},$$

$$\bar{G}_k(q^{-1})=g_{k1}q^{-1}+\cdots+g_{kn}q^{-n},$$
$$G_k(q^{-1})=q\bar{G}_k(q^{-1})=g_{k1}+g_{k2}q^{-1}+\cdots+g_{kn}q^{-(n-1)} \tag{4.5.39}$$

注意(4.5.36)也可直接由关系

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{A}^{k-1}\hat{\boldsymbol{x}}(t+1|t),\quad k>0 \tag{4.5.40}$$

及 Kalman 预报器(4.5.6)导出为

$$\hat{\boldsymbol{x}}(t+k+1|t+1)=\boldsymbol{A}^{k-1}\hat{\boldsymbol{x}}(t+2|t+1)=\boldsymbol{A}^{k-1}[\boldsymbol{A}\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}_p\varepsilon(t+1)]=$$
$$\boldsymbol{A}\hat{\boldsymbol{x}}(t+k|t)+\boldsymbol{A}^{k-1}\boldsymbol{K}_p\varepsilon(t+1) \tag{4.5.41}$$

因 $C(q^{-1})$ 是稳定的,由(4.5.1)有

$$e(t)=\frac{1}{C(q^{-1})}(A(q^{-1})y(t)) \tag{4.5.42}$$

将它代入(4.5.37)后,两边消去 $A(q^{-1})$ 有稳态 k 步预报器

$$\hat{y}(t+k|t)=\frac{G_k(q^{-1})}{C(q^{-1})}y(t) \tag{4.5.43}$$

或表为 ARMA 递推形式

$$C(q^{-1})\hat{y}(t+k|t)=G_k(q^{-1})y(t) \tag{4.5.44}$$

上两式具有 Åström 预报器形式,但它适用于非平稳 ARMA 过程预报.

上述结果可概括为如下定理.

【定理 4.5.1】 对非平稳 ARMA 过程(4.5.1),设 $A(q^{-1})$ 与 $C(q^{-1})$ 互质,$A(q^{-1})$ 是不稳定的多项式,$C(q^{-1})$ 是稳定的多项式,$e(t)$ 是零均值、方差为 σ_ε^2 的白噪声,则有 k 步 Wiener 预报器

$$\hat{y}(t+k|t)=\frac{G_k(q^{-1})}{C(q^{-1})}y(t) \tag{4.5.45}$$

其中 $G_k(q^{-1})=g_{k1}+g_{k2}q^{-1}+\cdots+g_{kn}q^{-(n-1)}$，且它的系数可用如下公式计算

$$\begin{bmatrix} g_{k1} \\ \vdots \\ g_{kn} \end{bmatrix} = \boldsymbol{A}^{k-1}\begin{bmatrix} c_1-a_1 \\ \vdots \\ c_n-a_n \end{bmatrix} \tag{4.5.46}$$

置

$$\boldsymbol{G}_k=[g_{k1},\cdots,g_{kn}]^{\mathrm{T}} \tag{4.5.47}$$

则有递推公式

$$\begin{gathered} \boldsymbol{G}_{i+1}=\boldsymbol{A}\boldsymbol{G}_i, \quad i=1,\cdots,k-1 \\ \boldsymbol{G}_1=[c_1-a_1,\cdots,c_n-a_n]^{\mathrm{T}} \end{gathered} \tag{4.5.48}$$

预报误差 $\tilde{y}(t+k|t)=y(t+k)-\hat{y}(t+k|t)$方差为

$$\mathrm{E}[\tilde{\mathrm{y}}(t+k|t)]=\boldsymbol{H}\boldsymbol{P}_k\boldsymbol{H}^{\mathrm{T}}+\sigma_e^2 \tag{4.5.49}$$

其中稳态 Kalman 预报误差 $\tilde{\boldsymbol{x}}(t+k|t)=\boldsymbol{x}(t+k)-\hat{\boldsymbol{x}}(t+k|t)$方差阵 $\boldsymbol{P}_k=\mathrm{E}[\tilde{\boldsymbol{x}}(t+k|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+k|t)]$为

$$\boldsymbol{P}_k = \boldsymbol{A}^{k-1}\boldsymbol{\Sigma}(\boldsymbol{A}^{k-1})^{\mathrm{T}}+\sigma_e^2\sum_{j=0}^{k-2}\boldsymbol{A}^j\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{A}^{j\mathrm{T}} \tag{4.5.50}$$

证明 只需证(4.5.49).这只是要注意关系 $\tilde{y}(t+k|t)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t+k|t)+e(t+k)$即可.而(4.5.50)由第三章给出. □

【注 4.5.1】 这里给出多项式 $G_k(q^{-1})$的新算法，它不同于 Åström 预报器中用解 Diophantine方程方法求 $G_k(q^{-1})$，这里避免了求解 Diophantine 方程.

下面我们给出定理 4.5.1 的第二种推导方法.

在定理 4.5.1 条件下，存在稳态 Kalman 预报器.

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})^{-1}\boldsymbol{K}_pe(t) \tag{4.5.51}$$

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{A}\hat{\boldsymbol{x}}(t+k-1|t) \tag{4.5.52}$$

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{A}^{k-1}\hat{\boldsymbol{x}}(t+1|t) \tag{4.5.53}$$

$$\hat{y}(t+k|t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{H}\boldsymbol{A}^{k-1}\hat{\boldsymbol{x}}(t+1|t) \tag{4.5.54}$$

因而稳态 k 步预报器有传递函数表示

$$\hat{y}(t+k|t)=\boldsymbol{H}\boldsymbol{A}^{k-1}(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})^{-1}\boldsymbol{K}_pe(t) \tag{4.5.55}$$

应用矩阵求逆公式

$$\begin{gathered} (\boldsymbol{I}_n-q^{-1}\boldsymbol{A})^{-1}=\boldsymbol{F}(q^{-1})/A(q^{-1}), \\ A(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{A}), \\ \boldsymbol{F}(q^{-1})=\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})=\boldsymbol{I}_n+\boldsymbol{F}_1q^{-1}+\cdots+\boldsymbol{F}_{n-1}q^{-(n-1)} \end{gathered} \tag{4.5.56}$$

且有关系

$$A(q^{-1})\boldsymbol{I}_n=\boldsymbol{F}(q^{-1})(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})=(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})\boldsymbol{F}(q^{-1}) \tag{4.5.57}$$

这引出重要关系

$$\boldsymbol{A}\boldsymbol{F}_i=\boldsymbol{F}_i\boldsymbol{A} \tag{4.5.58}$$

$$\boldsymbol{A}^{k-1}\boldsymbol{F}(q^{-1})=\boldsymbol{F}(q^{-1})\boldsymbol{A}^{k-1} \tag{4.5.59}$$

故(4.5.55)成为

$$A(q^{-1})\hat{y}(t+k|t)=\boldsymbol{H}\boldsymbol{F}(q^{-1})\boldsymbol{A}^{k-1}\boldsymbol{K}_pe(t) \tag{4.5.60}$$

因为 $(\boldsymbol{A},\boldsymbol{H})$ 为伴随形(4.5.4),由推论 1.6.1 有

$$\boldsymbol{H}\boldsymbol{F}(q^{-1})=\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})=[1,q^{-1},\cdots,q^{-(n-1)}] \tag{4.5.61}$$

定义

$$\boldsymbol{A}^{k-1}\boldsymbol{K}_p=\begin{bmatrix} g_{k1}\\ \vdots \\ g_{kn}\end{bmatrix},$$

$$G_k(q^{-1})=g_{k1}+g_{k2}q^{-1}+\cdots+g_{kn}q^{-(n-1)} \tag{4.5.62}$$

则(4.5.60)化为稳态 k 步 ARMA 预报器

$$A(q^{-1})\hat{y}(t+k\mid t)=G_k(q^{-1})e(t) \tag{4.5.63}$$

将(4.5.42)代入上式引出

$$\hat{y}(t+k\mid t)=\frac{G_k(q^{-1})}{C(q^{-1})}y(t) \tag{4.5.64}$$

它相同于定理 4.5.1 的结果.

下面给出非平稳 ARMA 过程稳态 k 步 Åström 预报器的另一种形式.

【定理 4.5.2】 在定理 4.5.1 条件下,有非平稳 ARMA 过程的稳态 k 步预报器

$$\hat{y}(t+k\mid t)=\frac{G_k(q^{-1})}{C(q^{-1})}y(t) \tag{4.5.65}$$

其中多项式 $G_k(q^{-1})=g_{k1}+g_{k2}q^{-1}+\cdots+g_{kn}q^{-(n-1)}$ 定义为

$$G_k(q^{-1})=\boldsymbol{H}\boldsymbol{A}^{k-1}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p \tag{4.5.66}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{A}-\boldsymbol{K}_p\boldsymbol{H} \tag{4.5.67}$$

$$\boldsymbol{K}_p=[c_1-a_1,\cdots,c_n-a_n]^{\mathrm{T}} \tag{4.5.68}$$

其中 $\boldsymbol{A},\boldsymbol{H}$ 由(4.5.4)定义.预报误差方差为(4.5.49)和(4.5.50).

证明 由(4.5.6)和(4.5.7)有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{K}_py(t) \tag{4.5.69}$$

由(4.5.69)有

$$\hat{x}(t+1\mid t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_py(t) \tag{4.5.70}$$

将上式代入(4.5.36)和(4.5.40)有

$$\hat{y}(t+k\mid t)=\boldsymbol{H}\boldsymbol{A}^{k-1}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_py(t) \tag{4.5.71}$$

应用(4.5.4)和(4.5.15)有

$$\boldsymbol{\Psi}_p=\boldsymbol{A}-\boldsymbol{K}_p\boldsymbol{H}=\begin{bmatrix} -c_1 & & & \\ \vdots & & \boldsymbol{I}_{n-1} & \\ -c_n & 0 & \cdots & 0\end{bmatrix} \tag{4.5.72}$$

这引出[17]

$$\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi})=C(q^{-1}) \tag{4.5.73}$$

将 $(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}=\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)/C(q^{-1})$ 代入(4.5.71)引出(4.5.65)和(4.5.66).证毕. □

【定理 4.5.3】 非平稳 ARMA 过程(4.5.1)在定理 4.5.1 条件下,当 $t_0\to-\infty$ 或 $t\to+\infty$ 有稳定关系

$$\mathrm{E}[y(t+k)\varepsilon(t)] = \boldsymbol{H}\boldsymbol{A}^{k-1}\boldsymbol{\Gamma}\sigma_e^2 \tag{4.5.74}$$

证明 对有限的 t_0 和 t，由(4.5.33)有

$$y(t+k) = \boldsymbol{H}\boldsymbol{A}^k x(t) + \sum_{i=t+1}^{t+k}\boldsymbol{H}\boldsymbol{A}^{t+k-i}\boldsymbol{\Gamma}e(i-1) \tag{4.5.75}$$

利用关系

$$\varepsilon(t) = \boldsymbol{H}\tilde{\boldsymbol{x}}(t|t-1) + e(t) \tag{4.5.76}$$

其中 $e(t)$是零均值、方差为 σ_e^2 的白噪声，利用关系

$$\boldsymbol{x}(t) = \tilde{\boldsymbol{x}}(t|t-1) + \hat{\boldsymbol{x}}(t|t-1) \tag{4.5.77}$$

我们可得到

$$\mathrm{E}[y(t+k)\varepsilon(t)] = \boldsymbol{H}\boldsymbol{A}^k\boldsymbol{P}(t|t-1)\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{H}\boldsymbol{A}^{k-1}\boldsymbol{\Gamma}\sigma_e^2 \tag{4.5.78}$$

因为由(4.5.29)有 $\boldsymbol{P}(t|t-1)\to\boldsymbol{0}$，$t_0\to-\infty$或 $t\to+\infty$，故而引出

$$\mathrm{E}[y(t+k)\varepsilon(t)] = \boldsymbol{H}\boldsymbol{A}^{k-1}\boldsymbol{\Gamma}\sigma_e^2 \tag{4.5.79}$$

证毕. □

【例 4.5.1】 考虑非平稳 ARMA(1,1)过程 $y(t)$

$$(1-aq^{-1})y(t) = (1+cq^{-1})e(t),\quad |a|\geqslant 1, |c|<1 \tag{4.5.80}$$

其中 $e(t)$是零均值、方差为 σ_e^2的白噪声.求稳态 k 步预报器$\hat{y}(t+k|t)$.

(4.5.80)有状态空间模型

$$x(t+1) = ax(t) + (a+c)e(t) \tag{4.5.81}$$

$$y(t) = x(t) + e(t) \tag{4.5.82}$$

这引出它的稳态 Kalman 预报器增益为

$$K_p = c + a \tag{4.5.83}$$

应用(4.5.47)和(4.5.48)有

$$G_1 = c + a,\quad G_k = a^{k-1}(c+a) \tag{4.5.84}$$

从而

$$G_k(q^{-1}) = a^{k-1}(c+a) \tag{4.5.85}$$

于是应用(4.5.44)k 步稳态预报器为

$$\hat{y}(t+k|t) = \frac{a^{k-1}(c+a)}{1+cq^{-1}} \tag{4.5.86}$$

它与用平稳 ARMA 过程的 Åströmk 步预报器公式得到相同的结果.

4.6 带白色观测噪声的 ARMA 过程的 Wiener 预报器

考虑带白色观测噪声 $v(t)$的 ARMA 过程 $y(t)$，

$$A(q^{-1})y(t) = C(q^{-1})e(t) \tag{4.6.1}$$

$$z(t) = y(t) + v(t) \tag{4.6.2}$$

$$A(q^{-1}) = 1 + a_1q^{-1} + \cdots + a_{n_a}q^{-n_a},$$

$$C(q^{-1}) = c_1q^{-1} + \cdots + c_{n_c}q^{-n_c} \tag{4.6.3}$$

其中 q^{-1}为单位滞后算子，$A(q^{-1})$与 $C(q^{-1})$互质，且 $A(q^{-1})$为稳定的多项式，$v(t)$和

$e(t)$是零均值、方差为 σ_v^2 和 σ_e^2 的独立的白噪声. 设 $t_0 \to -\infty$，问题是基于观测 $(z(t), z(t-1), \cdots)$ 求超前 k 步稳态预报器 $\hat{y}(t+k|t)$，$k>0$.

本节将给出解决此问题两种预报方法，一种是基于 ARMA 新息模型，这种方法叫现代时间序列分析方法[14,17]. 另一种是基于稳态 Kalman 预报器.

4.6.1 基于 ARMA 新息模型的预报方法

将(4.6.1)代入(4.6.2)有关系

$$A(q^{-1})z(t) = C(q^{-1})e(t) + A(q^{-1})v(t) \tag{4.6.4}$$

上式右边两个 MA 过程可用一个等价的 MA 过程表示为

$$D(q^{-1})\varepsilon(t) = C(q^{-1})e(t) + A(q^{-1})v(t) \tag{4.6.5}$$

其中 $\varepsilon(t)$是零均值、方差为 σ_ε^2的白噪声，且由$(A(q^{-1}), C(q^{-1}))$互质引出 $D(q^{-1}) = 1 + d_1 q^{-1} + \cdots + d_{n_d} q^{-n_d}$，是稳定的. $D(q^{-1})$和 σ_ε^2可用定理 2.13.1 的 Gevers – Wouters 算法求得. 于是由(4.6.4)和(4.6.5)引出 $z(t)$的 ARMA 模型

$$A(q^{-1})z(t) = D(q^{-1})\varepsilon(t) \tag{4.6.6}$$

因 $A(q^{-1})$和 $D(q^{-1})$是稳定的，则它是平稳、可逆的. 因而 $\varepsilon(t)$是 $z(t)$的新息过程，故称(4.6.6)为 $z(t)$的 ARMA 新息模型. 注意由(4.6.5)及 $D(q^{-1})$的稳定性有展式

$$\varepsilon(t) = \frac{C(q^{-1})}{D(q^{-1})}e(t) + \frac{A(q^{-1})}{D(q^{-1})}v(t) = \sum_{j=0}^{\infty}\alpha_j e(t-j) + \sum_{j=0}^{\infty}\beta_j v(t-j) \tag{4.6.7}$$

其中 $\alpha_0 = 1, \beta_0 = 1, \alpha_j$ 和β_j 可用公式(1.4.17)和(1.4.30)递推计算. 由(4.6.2)和射影性质有

$$\hat{y}(t+k|t) = \hat{z}(t+k|t) - \hat{v}(t+k|t) \tag{4.6.8}$$

注意由白噪声 $e(t)$和 $v(t)$的独立性和(4.6.7)引出

$$\hat{v}(t+k|t) = \mathrm{proj}(v(t+k)|z(t), z(t-1), \cdots) = \mathrm{proj}(v(t+k)|\varepsilon(t), \varepsilon(t-1), \cdots) = 0 \tag{4.6.9}$$

其中用到了由(4.6.7)引出的事实 $v(t+k) \perp L(\varepsilon(t), \varepsilon(t-1), \cdots)(k>0)$. 于是有

$$\hat{y}(t+k|t) = \hat{z}(t+k|t) \tag{4.6.10}$$

基于 $z(t)$的平稳、可逆的 ARMA 模型，用 Box – Jenkins 递推预报方法或 Åström 预报方法可求得 $\hat{z}(t+k|t)$，从而可求得 $\hat{y}(t+k|t) = \hat{z}(t+k|t)$.

【定理 4.6.1】 带白色观测噪声的平稳 ARMA 过程(4.6.1)和(4.6.2)，有稳态 k 步预报器 $\hat{y}(t+k|t)$为

$$\hat{y}(t+k|t) = \hat{z}(t+k|t) \tag{4.6.11}$$

其中 $\hat{z}(t+k|t)$由 ARMA 新息模型(4.6.6)，

$$A(q^{-1})z(t) = D(q^{-1})\varepsilon(t) \tag{4.6.12}$$

可用 Åström 预报方法求得为

$$\hat{z}(t+k|t) = \frac{G_k(q^{-1})}{D(q^{-1})}z(t) \tag{4.6.13}$$

其中 $G_k(q^{-1})$和 $F_k(q^{-1})$由如下 Diophantine 方程决定

$$D(q^{-1}) = A(q^{-1})F_k(q^{-1}) + q^{-k}G_k(q^{-1}) \tag{4.6.14}$$

带阶次 $n_f = (k-1)$，$n_g = \max(n_a - 1, n_d - k)$. 从而有

$$\hat{y}(t+k|t)=\frac{G_k(q^{-1})}{D(q^{-1})}z(t) \tag{4.6.15}$$

或

$$D(q^{-1})\hat{y}(t+k|t)=G_k(q^{-1})z(t) \tag{4.6.16}$$

预报误差 $\tilde{y}(t+k|t)=y(t+k)-\hat{y}(t+k|t)$为

$$\tilde{y}(t+k|t)=F_k(q^{-1})\varepsilon(t+k)-v(t+k) \tag{4.6.17}$$

预报误差方差为

$$\mathrm{E}[\tilde{y}(t+k|t)]=\sigma_\varepsilon^2\sum_{i=0}^{k-1}f_{ki}{}^2-\sigma_v^2 \tag{4.6.18}$$

证明 由(4.6.10)和定理4.4.1得(4.6.11)~(4.6.16).由(4.6.2)和(4.6.11)引出

$$\tilde{y}(t+k|t)=\tilde{z}(t+k|t)-v(t+k) \tag{4.6.19}$$

其中 $\tilde{z}(t+k|t)=z(t+k)-\hat{z}(t+k|t)$, $\tilde{z}(t+k|t)=F_k(q^{-1})\varepsilon(t+k)$.应用(4.6.7)注意 $\mathrm{E}[\varepsilon(t+k)v(t+k)]=\sigma_v^2$,则引出(4.6.18).证毕. □

4.6.2 基于 Kalman 预报器的预报方法

由定理1.6.1,系统(4.6.1)~(4.6.3)有等价的状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{C}e(t) \tag{4.6.20}$$

$$z(t)=\boldsymbol{H}\boldsymbol{x}(t)+v(t) \tag{4.6.21}$$

其中 $n=\max(n_a,n_c)$, $a_i=0(i>n_a)$, $c_i=0(i>n_c)$且

$$\boldsymbol{A}=\begin{bmatrix}-a_1 & & & \\ \vdots & & \boldsymbol{I}_{n-1} & \\ -a_n & 0 & \cdots & 0\end{bmatrix},\quad \boldsymbol{C}=\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0\cdots 0] \tag{4.6.22}$$

由$(A(q^{-1}),C(q^{-1}))$互质引出系统(4.6.20)和(4.6.21)是完全可观、完全可控的,因而存在稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_pz(t) \tag{4.6.23}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{A}-\boldsymbol{K}_p\boldsymbol{H} \tag{4.6.24}$$

其中 $\boldsymbol{\Psi}_p$是一个稳定矩阵[15,16]且增益 $\boldsymbol{K}_p$ 为

$$\boldsymbol{K}_p=\boldsymbol{A\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2)^{-1} \tag{4.6.25}$$

其中 $\boldsymbol{\Sigma}$ 是如下稳态 Riccati 方程的解

$$\boldsymbol{\Sigma}=\boldsymbol{A}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2)^{-1}\boldsymbol{H\Sigma}]\boldsymbol{A}^{\mathrm{T}}+\sigma_e^2\boldsymbol{CC}^{\mathrm{T}} \tag{4.6.26}$$

它可用迭代法求解.[14]

置 $t=t+k$,取(4.6.12)两边各项在 Hilbert 空间 $L(z(t),z(t-1),\cdots)$上的射影有

$$\hat{z}(t+k|t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t+k|t) \tag{4.6.27}$$

类似由(4.6.2)有

$$\hat{y}(t+k|t)=\hat{z}(t+k|t) \tag{4.6.28}$$

从而有

$$\hat{y}(t+k|t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t+k|t) \tag{4.6.29}$$

注意系统(4.6.20)和(4.6.21)有稳态 k 步 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{A}^{k-1}\hat{\boldsymbol{x}}(t+1|t) \tag{4.6.30}$$

由(4.6.23)有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p z(t) \tag{4.6.31}$$

于是我们有

$$\hat{y}(t+k|t)=\boldsymbol{HA}^{k-1}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p z(t) \tag{4.6.32}$$

记

$$\psi(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p) \tag{4.6.33}$$

$$G_k(q^{-1})=\boldsymbol{HA}^{k-1}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p=g_{k1}+g_{k2}q^{-1}+\cdots+g_{kn}q^{-(n-1)} \tag{4.6.34}$$

将上两式代入(4.6.32)可得如下定理.

【定理 4.6.2】 带白色观测噪声的平稳或非平稳 ARMA 过程(4.6.1)~(4.6.3),其中假设 $A(q^{-1})$是稳定的或不稳定的多项式,$A(q^{-1})$和 $C(q^{-1})$互质,$e(t)$和 $v(t)$是零均值、方差各为 σ_e^2和σ_v^2的独立的白噪声,有渐近稳定的超前 k 步稳态 Wiener 预报器

$$\psi(q^{-1})\hat{y}(t+k|t)=G_k(q^{-1})z(t) \tag{4.6.35}$$

或

$$\hat{y}(t+k|t)=\frac{G_k(q^{-1})}{\psi(q^{-1})}z(t) \tag{4.6.36}$$

其中 $\psi(q^{-1})$和 $G_k(q^{-1})$由(4.6.33)和(4.6.34)定义. $\boldsymbol{\Psi}_p$ 和 $\boldsymbol{K}_p$ 及 $\boldsymbol{\Sigma}$ 由(4.6.24)~(4.6.26)计算.预报误差 $\tilde{y}(t+k|t)=y(t+k)-\hat{y}(t+k|t)$方差为

$$\mathrm{E}[\tilde{y}^2(t+k|t)]=\boldsymbol{HP}_k\boldsymbol{H}^{\mathrm{T}},\quad \boldsymbol{H}=[1,0\cdots0] \tag{4.6.37}$$

其中稳态状态预报误差 $\tilde{\boldsymbol{x}}(t+k|t)=\boldsymbol{x}(t+k)-\hat{\boldsymbol{x}}(t+k|t)$方差阵 $\boldsymbol{P}_k=\mathrm{E}[\tilde{\boldsymbol{x}}(t+k|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+k|t)]$为

$$\boldsymbol{P}_k=\boldsymbol{A}^{k-1}\boldsymbol{\Sigma}(\boldsymbol{A}^{k-1})^{\mathrm{T}}+\sigma_e^2\sum_{j=0}^{k-2}\boldsymbol{A}^j\boldsymbol{CC}^{\mathrm{T}}\boldsymbol{A}^{j\mathrm{T}} \tag{4.6.38}$$

证明 将$(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}=\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)/\psi(q^{-1})$代入(4.6.32)引出(4.6.35)或(4.6.36).因 $\boldsymbol{\Psi}_p$ 是一个稳定矩阵,[15,16]故 $\psi(q^{-1})$是一个稳定多项式,因而(4.6.35)是渐近稳定的.注意由(4.6.1)和(4.6.2)有

$$y(t)=\boldsymbol{Hx}(t) \tag{4.6.39}$$

由(4.6.29)和上式引出(4.6.37).由第三章有(4.6.38).证毕. □

由定理 4.6.1 和定理 4.6.2 给出的两种预报器 $\hat{y}(t+k|t)$是否等价?即它们是否是恒同的?下述引理和定理回答了这个问题.

【引理 4.6.1】 在定理 4.6.2 条件下,用 Kalman 滤波方法可得 $z(t)$ARMA 新息模型为

$$A(q^{-1})z(t)=\psi(q^{-1})\varepsilon(t) \tag{4.6.40}$$

$$\psi(q^{-1})=1+\psi_1q^{-1}+\cdots+\psi_nq^{-n} \tag{4.6.41}$$

$$\psi_i=k_{pi}+a_i,\quad i=1,\cdots,n \tag{4.6.42}$$

$$\boldsymbol{K}_p=\begin{bmatrix}k_{p1}\\ \vdots\\ k_{pn}\end{bmatrix} \tag{4.6.43}$$

其中稳态新息 $\varepsilon(t)$ 等价于由稳态 Kalman 预报器决定的稳态新息,

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{A}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\varepsilon(t) \tag{4.6.44}$$

$$\varepsilon(t)=z(t)-\boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1) \tag{4.6.45}$$

带任意初值 $\hat{\boldsymbol{x}}(1|0)$. $\boldsymbol{K}_p$ 由(4.6.25)和(4.6.26)给出.

证明 由上两式有

$$z(t)=\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})^{-1}\boldsymbol{K}_p\varepsilon(t-1)+\varepsilon(t) \tag{4.6.46}$$

注意(4.6.22)有

$$A(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{A}) \tag{4.6.47}$$

上两式引出

$$A(q^{-1})z(t)=\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})\boldsymbol{K}_p\varepsilon(t-1)+A(q^{-1})\varepsilon(t) \tag{4.6.48}$$

因$(\boldsymbol{A},\boldsymbol{H})$为伴随形(4.6.22),故有[17]

$$\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})=[1,q^{-1},\cdots,q^{-(n-1)}] \tag{4.6.49}$$

将此式和(4.6.43)代入(4.6.48)引出 ARMA 新息模型

$$A(q^{-1})z(t)=\psi(q^{-1})\varepsilon(t) \tag{4.6.50}$$

其中 $\psi(q^{-1})$, ψ_i 由(4.6.41)和(4.6.42)定义. 下面证明 $\psi(q^{-1})$是稳定的. 事实上

$$\psi(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p) \tag{4.6.51}$$

这是因为由(4.6.22)、(4.6.42)和(4.6.43)有

$$\boldsymbol{\Psi}_p=\boldsymbol{A}-\boldsymbol{K}_p\boldsymbol{H}=\begin{bmatrix}-(a_1+k_{p1}) & & & \\ \vdots & \boldsymbol{I}_{n-1} & & \\ -(a_n+k_{pn}) & 0 & \cdots & \cdots\end{bmatrix}=\begin{bmatrix}-\psi_1 & & & \\ \vdots & \boldsymbol{I}_{n-1} & & \\ -\psi_n & 0 & \cdots & \cdots\end{bmatrix} \tag{4.6.52}$$

这引出(4.6.51)成立. 因为 $\boldsymbol{\Psi}_p$ 为稳定矩阵,故 $\psi(q^{-1})$是稳定多项式,即 $\psi(x)$的零点全位于单位圆外. 证毕. □

比较(4.6.6)和(4.6.50)引出

$$\psi(q^{-1})=D(q^{-1}) \tag{4.6.53}$$

于是有如下定理.

【定理 4.6.3】 由定理 4.6.1 和定理 4.6.2 给出的稳态 k 步预报器(4.6.15)与(4.6.36)是等价的.

证明 由(4.6.53)得证. 对平稳可逆 ARMA 过程(4.6.1)~(4.6.3)情形,令 $t_0=-\infty$,则(4.6.15)和(4.6.36)均为稳态最优预报器. 由 Hilbert 空间射影惟一性引出这两个预报器在数值上是恒同的. 因而由(4.6.53)引出由(4.6.14)解 Diophantine 得到的 $G_k(q^{-1})$与由(4.6.34)Kalman 预报器得到的 $G_k(q^{-1})$是恒同的. 因而这两个 k 步预报器是等价的. 证毕. □

4.7 带有色观测噪声的 ARMA 过程的稳态最优预报器

考虑带白色和有色观测噪声的 ARMA 过程 $y(t)$,

$$z(t)=y(t)+\eta(t)+\xi(t) \tag{4.7.1}$$

$$A(q^{-1})y(t)=C(q^{-1})e(t) \tag{4.7.2}$$

$$P(q^{-1})\eta(t)=R(q^{-1})n(t) \tag{4.7.3}$$

其中 $y(t)$为被估 ARMA 过程，$z(t)$是对 $y(t)$的观测，有色观测噪声 $\eta(t)$服从 ARMA 模型(4.7.3)，$\xi(t)$是白色观测噪声.设$(A(q^{-1})P(q^{-1}),P(q^{-1})C(q^{-1}),A(q^{-1})R(q^{-1}))$互质，且 $\xi(t)$，$e(t)$和 $n(t)$是零均值、方差各为 σ_ε^2，σ_e^2和σ_n^2的相互独立白噪声，q^{-1}是单位滞后算子，$A(q^{-1})$，…，$R(q^{-1})$有形式 $X(q^{-1})=x_0+x_1q^{-1}+\cdots+x_{n_x}q^{-n_x}$，且 $a_0=1$，$p_0=1$，$c_0=0$，$r_0=0$.假设 $A(q^{-1})$，$P(q^{-1})$均为稳定的多项式，即 $y(t)$和 $\eta(t)$为平稳时间序列，且设初始观测时刻 $t_0=-\infty$.问题是基于观测$(z(t),z(t-1),\cdots)$，求 $y(t)$的稳态最优预报器 $\hat{y}(t+k|t)$，$k>0$.

本节推广节 4.6 的结果，将用两种方法解决这个问题.

4.7.1 基于 ARMA 新息模型的稳态最优预报器

将(4.7.2)和(4.7.3)代入(4.7.1)有

$$z(t)=\frac{C(q^{-1})}{A(q^{-1})}e(t)+\frac{R(q^{-1})}{P(q^{-1})}n(t)+\xi(t) \tag{4.7.4}$$

上式两边同乘 $A(q^{-1})P(q^{-1})$有

$$A(q^{-1})P(q^{-1})z(t)=P(q^{-1})C(q^{-1})e(t)+A(q^{-1})R(q^{-1})n(t)+A(q^{-1})P(q^{-1})\xi(t) \tag{4.7.5}$$

由互质性假设上式右边三个 MA 过程可用一个等价的稳定的 MA 过程表示为

$$D(q^{-1})\varepsilon(t)=\Phi(q^{-1})e(t)+\Psi(q^{-1})n(t)+\Lambda(q^{-1})\xi(t) \tag{4.7.6}$$

$$\Phi(q^{-1})=P(q^{-1})C(q^{-1}),\quad \Psi(q^{-1})=A(q^{-1})R(q^{-1}),\quad \Lambda(q^{-1})=A(q^{-1})P(q^{-1}) \tag{4.7.7}$$

其中稳定的多项式 $D(q^{-1})=1+d_1q^{-1}+\cdots+d_{n_d}q^{-n_d}$，$d_0=1$，$\varepsilon(t)$是零均值、方差为 σ_ε^2的白噪声，$D(q^{-1})$和 σ_ε^2可用定理 2.13.1 给出的 Gevers－Wouters 算法求得.比较(4.7.5)和(4.7.6)有 ARMA 模型

$$\Lambda(q^{-1})z(t)=D(q^{-1})\varepsilon(t) \tag{4.7.8}$$

因为 $A(q^{-1})$和 $P(q^{-1})$是稳定的，故 $\Lambda(q^{-1})$也是稳定的，因而(4.7.8)是平稳可逆的，且 $\varepsilon(t)$是 $z(t)$的新息过程，故称(4.7.8)为 $z(t)$的 ARMA 新息模型.

由射影理论有

$$\hat{y}(t+k|t)=\mathrm{proj}(y(t+k)|z(t),z(t-1),\cdots)=\mathrm{proj}(y(t+k)|\varepsilon(t),\varepsilon(t-1),\cdots) \tag{4.7.9}$$

由 $D(q^{-1})$的稳定性和(4.7.6)$\varepsilon(t)$可展为均方收敛的级数

$$\varepsilon(t)=\sum_{j=0}^{\infty}\alpha_je(t-j)+\sum_{j=0}^{\infty}\beta_jn(t-j)+\sum_{j=0}^{\infty}\gamma_j\xi(t-j) \tag{4.7.10}$$

其中系数 α_j，β_j 和 γ_j 可由(1.4.17)递推计算为

$$\begin{aligned}\alpha_j&=-d_1\alpha_{j-1}-\cdots-d_{n_d}\alpha_{j-n_d}+\varphi_j,\\ \beta_j&=-d_1\beta_{j-1}-\cdots-d_{n_d}\beta_{j-n_d}+\psi_j,\\ \gamma_j&=-d_1\gamma_{j-1}-\cdots-d_{n_d}\gamma_{j-n_d}+\lambda_j\end{aligned} \tag{4.7.11}$$

其中规定 $\alpha_0=0$，$\beta_0=0$，$\gamma_0=0$，$\alpha_j=0\,(j<0)$，$\beta_j=0\,(j<0)$，$\gamma_j=0\,(j<0)$，$\varphi_j=0\,(j>n_\varphi)$，$\psi_j=0\,(j>n_\psi)$，$\lambda_j=0\,(j>n_\lambda)$. 由(4.7.2)和(4.7.3)，$y(t)$和 $\eta(t)$有展式

$$y(t)=\frac{C(q^{-1})}{A(q^{-1})}e(t)=\sum_{j=0}^{\infty}\omega_j e(t-j) \tag{4.7.12}$$

$$\eta(t)=\frac{R(q^{-1})}{P(q^{-1})}n(t)=\sum_{j=0}^{\infty}\delta_j n(t-j) \tag{4.7.13}$$

其中系数 ω_j 和 δ_j 可递推计算为

$$\omega_j=-a_1\omega_{j-1}-\cdots-a_{n_a}\omega_{j-n_a}+c_j \tag{4.7.14}$$

$$\delta_j=-p_1\delta_{j-1}-\cdots-p_{n_p}\delta_{j-n_p}+r_j \tag{4.7.15}$$

其中 $\omega_0=0$，$\omega_j=0\,(j<0)$，$c_j=0\,(j>n_c)$；$\delta_0=0$，$\delta_j=0\,(j<0)$，$r_j=0\,(j>n_r)$.

由新息序列 $\varepsilon(t)$的正交性和射影性质有

$$\begin{aligned}\hat{y}(t+k\mid t)=&\sum_{s=0}^{\infty}\mathrm{proj}(y(t+k)\mid\varepsilon(t-s))=\\&\sum_{s=0}^{\infty}\mathrm{proj}\Big(\sum_{j=0}^{\infty}\omega_j e(t+k-j)\mid\varepsilon(t-s)\Big)=\\&\sum_{s=0}^{\infty}\sum_{j=0}^{\infty}\omega_j\mathrm{proj}(e(t+k-j)\mid\varepsilon(t-s))=\\&\sum_{s=0}^{\infty}\sum_{j=0}^{\infty}\omega_j\mathrm{E}[e(t+k-j)\mid\varepsilon(t-s)]\frac{1}{\sigma_\varepsilon^2}\varepsilon(t-s)\end{aligned} \tag{4.7.16}$$

应用(4.7.10)和 $e(t)$，$n(t)$和 $\xi(t)$的相互独立性有

$$\mathrm{E}[e(t+k-j)\varepsilon(t-s)]=\mathrm{E}[e(t+k-j)\sum_{i=0}^{\infty}\alpha_i e(t-s-i)]=\sigma_e^2\alpha_{j-k-s} \tag{4.7.17}$$

将上式代入(4.7.16)引出稳态 k 步最优预报器

$$\hat{y}(t+k\mid t)=\frac{\sigma_e^2}{\sigma_\varepsilon^2}\sum_{s=0}^{\infty}\sum_{j=0}^{\infty}\omega_j\alpha_{j-k-s}\varepsilon(t-s) \tag{4.7.18}$$

因 $\alpha_j=0\,(j<0)$，故有

$$\hat{y}(t+k\mid t)=\frac{\sigma_e^2}{\sigma_\varepsilon^2}\sum_{s=0}^{\infty}\sum_{j=k+s}^{\infty}\omega_j\alpha_{j-k-s}\varepsilon(t-s) \tag{4.7.19}$$

【定理 4.7.1】 带白色和有色观测噪声的 ARMA 过程(4.7.1)～(4.7.3)，若多项式 $A(q^{-1})$和 $P(q^{-1})$是稳定的，则有稳态最优 k 步预报器(4.7.19)，其中新息 $\varepsilon(t-s)$由 ARMA新息模型(4.7.8)计算为

$$\varepsilon(t)=\frac{A(q^{-1})}{D(q^{-1})}z(t)=\sum_{j=0}^{\infty}\gamma_j z(t-j) \tag{4.7.20}$$

其中 γ_j 由(4.7.11)计算. 而误差 $\tilde{y}(t+k\mid t)=y(t+k)-\hat{y}(t+k\mid t)$的方差 $P_k^y=\mathrm{E}[\tilde{y}^2(t+k\mid t)]$为

$$P_k^y=\sigma_e^2\sum_{j=0}^{\infty}\omega_j^2-\sigma_\varepsilon^2\sum_{s=0}^{\infty}(\varphi_s^{(k)})^2 \tag{4.7.21}$$

其中定义

$$\varphi_s^{(k)} = \frac{\sigma_e^2}{\sigma_\varepsilon^2}\sum_{j=0}^{\infty}\omega_j\alpha_{j-k-s} \tag{4.7.22}$$

证明　只需证(4.7.21). 由(4.7.18)$\hat{y}(t+k|t)$可表为

$$\hat{y}(t+k\mid t) = \sum_{s=0}^{\infty}\varphi_s^{(k)}\varepsilon(t-s) \tag{4.7.23}$$

其中 $\varphi_s^{(k)}$ 由(4.7.22)定义. 应用(4.7.12)和上两式,并应用表达式(4.7.10)及白噪声 $e(t)$,$n(t)$,$\xi(t)$的相互独立性有

$$\begin{aligned}
P_k^y &= \mathrm{E}[\tilde{y}^2(t+k\mid t)] = \mathrm{E}\Big[\Big(\sum_{j=0}^{\infty}\omega_j e(t+k-j) - \sum_{s=0}^{\infty}\varphi_s^{(k)}\varepsilon(t-s)\Big)^2\Big] = \\
&\sigma_e^2\sum_{j=0}^{\infty}\omega_j^2 + \sigma_\varepsilon^2\sum_{s=0}^{\infty}(\varphi_s^{(k)})^2 - 2\mathrm{E}\Big[\Big(\sum_{j=0}^{\infty}\omega_j e(t+k-j)\Big)\Big(\sum_{s=0}^{\infty}\varphi_s^{(k)}\varepsilon(t-s)\Big)\Big] = \\
&\sigma_e^2\sum_{j=0}^{\infty}\omega_j^2 + \sigma_\varepsilon^2\sum_{s=0}^{\infty}(\varphi_s^{(k)})^2 - 2\sigma_e^2\sum_{s=0}^{\infty}\sum_{j=0}^{\infty}\varphi_s^{(k)}\omega_j\alpha_{j-k-s} = \\
&\sigma_e^2\sum_{j=0}^{\infty}\omega_j^2 + \sigma_\varepsilon^2\sum_{s=0}^{\infty}(\varphi_s^{(k)})^2 - 2\sigma_\varepsilon^2\sum_{s=0}^{\infty}\varphi_s^{(k)}\frac{\sigma_e^2}{\sigma_\varepsilon^2}\sum_{j=0}^{\infty}\omega_j\alpha_{j-k-s} = \\
&\sigma_e^2\sum_{j=0}^{\infty}\omega_j^2 - \sigma_\varepsilon^2\sum_{s=0}^{\infty}(\varphi_s^{(k)})^2
\end{aligned} \tag{4.7.24}$$

即(4.7.21)成立. □

【注 4.7.1】　在应用中可用有限项和近似代替无穷级数(4.7.19),(4.7.21)和(4.7.22),即取自然 N 充分大有次优预报器

$$\hat{y}(t+k\mid t) = \frac{\sigma_e^2}{\sigma_\varepsilon^2}\sum_{s=0}^{N}\sum_{j=k+s}^{N}\omega_j\alpha_{j-k-s}\varepsilon(t-s),$$

$$P_k^y = \sigma_e^2\sum_{j=0}^{N}\omega_j^2 - \sigma_\varepsilon^2\sum_{s=0}^{N}(\varphi_s^{(k)})^2,$$

$$\varphi_s^{(k)} = \frac{\sigma_e^2}{\sigma_\varepsilon^2}\sum_{j=0}^{N}\omega_j\alpha_{j-k-s} \tag{4.7.25}$$

4.7.2　基于 Kalman 预报器的稳态 k 步预报器

以下推导的结果适用于非平稳 ARMA 过程 $y(t)$和非平稳有色观测噪声 $\eta(t)$,即 $A(q^{-1})$和 $P(q^{-1})$可以是不稳定的多项式. 考虑带白色和有色观测噪声的 ARMA 过程(4.7.1)~(4.7.3),(4.7.2)有状态空间模型

$$\begin{aligned}\boldsymbol{\alpha}(t+1) &= \bar{\boldsymbol{A}}\boldsymbol{\alpha}(t) + \bar{\boldsymbol{C}}e(t),\\ y(t) &= \bar{\boldsymbol{H}}_1\boldsymbol{\alpha}(t)\end{aligned} \tag{4.7.26}$$

其中 $n_1 = \max(n_a, n_c)$ $a_i = 0(i > n_a)$,$c_i = 0(i > n_c)$,且

$$\bar{\boldsymbol{A}} = \begin{bmatrix} -a_1 & & \\ \vdots & \boldsymbol{I}_{n_1-1} & \\ -a_{n_1} & 0 \quad \cdots & 0\end{bmatrix},\quad \bar{\boldsymbol{C}} = \begin{bmatrix} c_1 \\ \vdots \\ c_{n_1}\end{bmatrix},\quad \bar{\boldsymbol{H}}_1 = [1 \quad 0\cdots 0] \tag{4.7.27}$$

(4.7.3)有状态空间模型

$$\boldsymbol{\beta}(t+1) = \bar{\boldsymbol{P}}\boldsymbol{\beta}(t) + \bar{\boldsymbol{R}}n(t),$$

$$\eta(t)=\bar{\boldsymbol{H}}_2\boldsymbol{\beta}(t) \tag{4.7.28}$$

其中规定 $n_2=\max(n_p,n_r)$，$p_i=0(i>n_p)$，$r_i=0(i>n_r)$，且

$$\bar{\boldsymbol{P}}=\begin{bmatrix} -p_1 & & & \\ \vdots & \boldsymbol{I}_{n_2-1} & & \\ -p_{n_2} & 0 & \cdots & 0\end{bmatrix},\quad \bar{\boldsymbol{R}}=\begin{bmatrix} r_1\\ \vdots \\ r_{n_2}\end{bmatrix},\quad \bar{\boldsymbol{H}}_2=[1\quad 0\cdots 0] \tag{4.7.29}$$

于是(4.7.1)~(4.7.3)有增广状态空间模型

$$\begin{aligned}\boldsymbol{x}(t+1)&=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t),\\ z(t)&=\boldsymbol{Hx}(t)+\xi(t)\end{aligned} \tag{4.7.30}$$

其中定义 $n=n_1+n_2$，

$$\boldsymbol{x}(t)=\begin{bmatrix}\boldsymbol{\alpha}(t)\\ \boldsymbol{\beta}(t)\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\bar{\boldsymbol{C}} & \boldsymbol{0}\\ \boldsymbol{0} & \bar{\boldsymbol{R}}\end{bmatrix},\quad \boldsymbol{w}(t)=\begin{bmatrix}e(t)\\ n(t)\end{bmatrix},$$

$$\boldsymbol{\Phi}=\begin{bmatrix}\bar{\boldsymbol{A}} & \boldsymbol{0}\\ \boldsymbol{0} & \bar{\boldsymbol{P}}\end{bmatrix},\quad \boldsymbol{H}=[\bar{\boldsymbol{H}}_1,\bar{\boldsymbol{H}}_2] \tag{4.7.31}$$

由 $A(q^{-1})P(q^{-1})$，$P(q^{-1})C(q^{-1})$，$A(q^{-1})R(q^{-1})$互质的假设引出系统(4.7.30)和(4.7.31)是完全可观、完全可控的[14,12]，因而存在稳态 Kalman 预报器[15]

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p z(t) \tag{4.7.32}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H} \tag{4.7.33}$$

$$\boldsymbol{K}_p=\boldsymbol{\Phi\Sigma H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_\xi^2]^{-1} \tag{4.7.34}$$

其中稳态预报误差方差阵 $\boldsymbol{\Sigma}$ 是如下稳态 Riccati 方程的解：

$$\boldsymbol{\Sigma}=\boldsymbol{\Phi}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_\xi^2)^{-1}\boldsymbol{H\Sigma}]\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{\Gamma Q}_w\boldsymbol{\Gamma}^{\mathrm{T}} \tag{4.7.35}$$

$$\boldsymbol{Q}_w=\begin{bmatrix}\sigma_e^2 & 0\\ 0 & \sigma_n^2\end{bmatrix} \tag{4.7.36}$$

Riccati 方程(4.7.35)可用迭代法求解[14]. 可证明[15,16] $\boldsymbol{\Psi}_p$ 是一个稳定矩阵.

由(4.7.26)和(4.7.31)有

$$y(t)=\boldsymbol{H}_0x(t),\quad \boldsymbol{H}_0=[1\quad 0\cdots 0] \tag{4.7.37}$$

因而有

$$\hat{y}(t+k|t)=\boldsymbol{H}_0\hat{\boldsymbol{x}}(t+k|t),\quad k>0 \tag{4.7.38}$$

而由(4.7.30)有

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{\Phi}^{k-1}\hat{\boldsymbol{x}}(t+k|t),\quad k>0 \tag{4.7.39}$$

于是有稳态 k 步预报器

$$\hat{y}(t+k|t)=\boldsymbol{H}_0\boldsymbol{\Phi}^{k-1}\hat{\boldsymbol{x}}(t+1|t) \tag{4.7.40}$$

将(4.7.32)代入上式有

$$\hat{y}(t+k|t)=\boldsymbol{H}_0\boldsymbol{\Phi}^{k-1}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p z(t) \tag{4.7.41}$$

于是有稳态 k 步 Wiener 预报器

$$\hat{y}(t+k|t)=\frac{G_k(q^{-1})}{\psi(q^{-1})}z(t) \tag{4.7.42}$$

$$\psi(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p) \tag{4.7.43}$$

$$G_k(q^{-1}) = \boldsymbol{H}_0\boldsymbol{\Phi}^{k-1}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p = g_{k1} + g_{k2}q^{-1} + \cdots + g_{kn}q^{-(n-1)} \tag{4.7.44}$$

因 $\boldsymbol{\Psi}_p$ 是稳定矩阵，故 $\psi(q^{-1})$ 是稳定多项式，故(4.7.42)是渐近稳定的.

预报误差 $\tilde{y}(t+k|t) = y(t+k) - \hat{y}(t+k|t)$ 方差为

$$\mathrm{E}[\tilde{y}^2(t+k|t)] = \boldsymbol{H}_0\boldsymbol{P}_k\boldsymbol{H}_0^{\mathrm{T}} \tag{4.7.45}$$

其中状态稳态预报误差 $\tilde{\boldsymbol{x}}(t+k|t) = \boldsymbol{x}(t+k) - \hat{\boldsymbol{x}}(t+k|t)$ 方差阵 $\boldsymbol{P}_k = \mathrm{E}[\tilde{\boldsymbol{x}}(t+k|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+k|t)]$ 由第三章有

$$\boldsymbol{P}_k = \boldsymbol{\Phi}^{k-1}\boldsymbol{\Sigma}(\boldsymbol{\Phi}^{k-1})^{\mathrm{T}} + \sum_{j=0}^{k-2}\boldsymbol{\Phi}^j\boldsymbol{\Gamma}\boldsymbol{Q}_w\Gamma^{\mathrm{T}}\boldsymbol{\Phi}^{j\mathrm{T}}, \quad k > 1 \tag{4.7.46}$$

带 $\boldsymbol{P}_1 = \boldsymbol{\Sigma}$.

【定理 4.7.2】 带平稳或非平稳有色观测噪声的平稳或非平稳 ARMA 过程(4.7.1)~(4.7.3)有渐近稳定的稳态 k 步 Wiener 预报器 $\hat{y}(t+k|t)$ 为(4.7.42)，相应的预防误差方差为(4.7.45)和(4.7.46).

下面提出不同于定理 4.7.2 的另一种基于 Kalman 滤波的推导方法.

由递推射影公式有 k 步稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+k+1|t+1) = \hat{\boldsymbol{x}}(t+k+1|t) + \boldsymbol{K}_k\varepsilon(t+1), \quad k>0 \tag{4.7.47}$$

$$\boldsymbol{K}_k = \mathrm{E}[\boldsymbol{x}(t+k+1)\varepsilon^{\mathrm{T}}(t+1)]Q_\varepsilon^{-1} \tag{4.7.48}$$

$$Q_\varepsilon = \boldsymbol{H\Sigma H}^{\mathrm{T}} + \sigma_\xi^2 \tag{4.7.49}$$

其中 $\boldsymbol{\Sigma}$ 由(4.7.35)计算，$\varepsilon(t+1)$ 是新息，Q_ε 是稳定新息方差. 由(4.7.30)和射影性质有

$$\hat{\boldsymbol{x}}(t+k+1|t) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t+k|t), \quad k>0 \tag{4.7.50}$$

其中用到事实 $\boldsymbol{w}(t+k) \perp L(z(t), z(t-1), \cdots)$. 于是有递推 k 步 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+k+1|t+1) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t+k|t) + \boldsymbol{K}_k\varepsilon(t+1) \tag{4.7.51}$$

由定义(4.7.31)，且定义

$$\boldsymbol{K}_k = \begin{bmatrix} \boldsymbol{K}_{k\alpha} \\ \boldsymbol{K}_{k\beta} \end{bmatrix} \tag{4.7.52}$$

其中 $\boldsymbol{K}_{k\alpha}$ 是 $n_1 \times 1$ 子块，则有子预报器

$$\hat{\boldsymbol{\alpha}}(t+k+1|t+1) = \bar{\boldsymbol{A}}\hat{\boldsymbol{\alpha}}(t+k|t) + \boldsymbol{K}_{k\alpha}\varepsilon(t+1) \tag{4.7.53}$$

在(4.7.26)中置 $t = t+k$ 后取射影运算有

$$\hat{y}(t+k|t) = \bar{\boldsymbol{H}}_1\hat{\boldsymbol{\alpha}}(t+k|t) \tag{4.7.54}$$

应用定理(1.6.1)，因为$(\bar{\boldsymbol{A}}, \bar{\boldsymbol{H}}_1)$为块伴随形，则状态空间表示(4.7.53)和(4.7.54)等价于 ARMA 递推预报器

$$A(q^{-1})\hat{y}(t+k|t) = \bar{K}_{k\alpha}(q^{-1})\varepsilon(t+1) \tag{4.7.55}$$

$$\bar{K}_{k\alpha}(q^{-1}) = K_{k\alpha 1}q^{-1} + \cdots + K_{k\alpha n_1}q^{-n_1} \tag{4.7.56}$$

$$\boldsymbol{K}_{k\alpha} = \begin{bmatrix} K_{k\alpha 1} \\ \vdots \\ \boldsymbol{K}_{k\alpha n_1} \end{bmatrix} \tag{4.7.57}$$

其中标量 $K_{k\alpha i}$ 为 $\boldsymbol{K}_{k\alpha}$ 的分量. (4.7.55)可化为

$$A(q^{-1})\hat{y}(t+k|t) = K_{k\alpha}(q^{-1})\varepsilon(t) \tag{4.7.58}$$

$$K_{k\alpha}(q^{-1})=K_{k\alpha 1}+K_{k\alpha 2}q^{-1}+\cdots+K_{k\alpha n_1}q^{-(n_1-1)} \tag{4.7.59}$$

为了实现(4.7.58)的计算，要求计算预报增益阵 K_k，从而可得到 $K_{k\alpha}(q^{-1})$. 注意(4.7.48)，问题归结为 $\mathrm{E}[\boldsymbol{x}(t+k+1)\varepsilon(t+1)]$. 由(4.7.30)迭代有关系

$$\boldsymbol{x}(t+k+1)=\boldsymbol{\Phi}^k\boldsymbol{x}(t+1)+\sum_{i=t+2}^{t+k+1}\boldsymbol{\Phi}^{t+k+1-i}\boldsymbol{\Gamma}\boldsymbol{w}(i-1) \tag{4.7.60}$$

又由第三章有

$$\varepsilon(t+1)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t+1|t)+\xi(t+1) \tag{4.7.61}$$

其中 $\tilde{\boldsymbol{x}}(t+1|t)=\boldsymbol{x}(t+1)-\hat{\boldsymbol{x}}(t+1|t)$. 由(4.7.61)引出(4.7.49)成立. 将 $\boldsymbol{x}(t+1)=\hat{\boldsymbol{x}}(t+1|t)+\tilde{\boldsymbol{x}}(t+1|t)$代入(4.7.60)，并注意 $\hat{\boldsymbol{x}}(t+1|t)\perp\tilde{\boldsymbol{x}}(t+1|t)$，$\boldsymbol{w}(t+1)\perp\tilde{\boldsymbol{x}}(t+1|t)$，$\xi(t+1)\perp\boldsymbol{w}(t+i)\ (i>0)$，$\xi(t+1)\perp\boldsymbol{x}(t+1)$，应用上两式可得

$$\mathrm{E}[\boldsymbol{x}(t+k+1)\varepsilon^{\mathrm{T}}(t+1)]=\boldsymbol{\Phi}^k\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} \tag{4.7.62}$$

从而由(4.7.48)有预报增益阵

$$\boldsymbol{K}_k=\boldsymbol{\Phi}^k\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\sigma_\xi^2]^{-1} \tag{4.7.63}$$

增益 K_k 的另一种推导方法如下：

注意关系(4.7.39)

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{\Phi}^{k-1}\hat{\boldsymbol{x}}(t+1|t) \tag{4.7.64}$$

置 $t=t+1$ 有

$$\hat{\boldsymbol{x}}(t+k+1|t+1)=\boldsymbol{\Phi}^{k-1}\hat{\boldsymbol{x}}(t+2|t+1) \tag{4.7.65}$$

而在新息滤波器形式下的稳态 Kalman 预报器为

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\varepsilon(t) \tag{4.7.66}$$

其中一步 Kalman 预报器增益 K_p 由(4.7.34)计算. 将上式代入(4.7.65)引出

$$\begin{aligned}\hat{\boldsymbol{x}}(t+k+1|t+1)&=\boldsymbol{\Phi}^{k-1}[\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{K}_p\varepsilon(t+1)]=\\ &\boldsymbol{\Phi}^k\hat{\boldsymbol{x}}(t+1|t)+\boldsymbol{\Phi}^{k-1}\boldsymbol{K}_p\varepsilon(t+1)\end{aligned} \tag{4.7.67}$$

而注意(4.7.64)上式化为

$$\hat{\boldsymbol{x}}(t+k+1|t+1)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t+k|t)+\boldsymbol{K}_k\varepsilon(t+1) \tag{4.7.68}$$

$$\boldsymbol{K}_k=\boldsymbol{\Phi}^{k-1}\boldsymbol{K}_p \tag{4.7.69}$$

将(4.7.34)代入(4.7.69)引出

$$\boldsymbol{K}_k=\boldsymbol{\Phi}^k\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\sigma_\xi^2]^{-1} \tag{4.7.70}$$

它相同于(4.7.63).

注意递推预报器(4.7.58)是新息滤波器，它的缺点是当 $A(q^{-1})$不稳定(即 $y(t)$为非平稳 ARMA 信号)时，它是非渐近稳定的. 因而不能用于递推计算. 现在我们要将其化为渐近稳定的 Wiener 预报器. 这要求建立 $z(t)$的 ARMA 新息模型. 注意

$$z(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1)+\varepsilon(t) \tag{4.7.71}$$

将(4.7.66)代入上式有

$$z(t)=\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{K}_p\varepsilon(t-1)+\varepsilon(t) \tag{4.7.72}$$

由定义(4.7.31)有

$$\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi}=\begin{bmatrix}\boldsymbol{I}_{n_1}-q^{-1}\bar{\boldsymbol{A}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}_{n_2}-q^{-1}\bar{\boldsymbol{P}}\end{bmatrix} \tag{4.7.73}$$

而由(4.7.27)和(4.7.29)，$\bar{\boldsymbol{A}}$ 和 $\bar{\boldsymbol{P}}$ 具有伴随形，故有

$$\det(\boldsymbol{I_n}-\boldsymbol{q}^{-1}\boldsymbol{\Phi})=\det(\boldsymbol{I_n}-\boldsymbol{q}^{-1}\bar{\boldsymbol{A}})\det(\boldsymbol{I}_{\boldsymbol{n}_2}-\boldsymbol{q}^{-1}\bar{\boldsymbol{P}})=A(q^{-1})P(q^{-1}) \tag{4.7.74}$$

将$(\boldsymbol{I_n}-\boldsymbol{q}^{-1}\boldsymbol{\Phi})^{-1}=\mathrm{adj}(\boldsymbol{I_n}-\boldsymbol{q}^{-1}\boldsymbol{\Phi})/\det(\boldsymbol{I_n}-\boldsymbol{q}^{-1}\boldsymbol{\Phi})$代入(4.7.72)可得 ARMA 新息模型

$$A(q^{-1})P(q^{-1})z(t)=D(q^{-1})\varepsilon(t) \tag{4.7.75}$$

其中定义多项式 $D(q^{-1})$为

$$D(q^{-1})=\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})\boldsymbol{K}_pq^{-1}+A(q^{-1})P(q^{-1}) \tag{4.7.76}$$

对比(4.7.75)与(4.7.8)可证明[17]：$D(q^{-1})$是稳定的(即 $D(x)$的零点全位于单位圆外).由(4.7.75)有

$$\varepsilon(t)=\frac{A(q^{-1})P(q^{-1})}{D(q^{-1})}z(t) \tag{4.7.77}$$

将它代入(4.7.58)可得如下定理.

【定理 4.7.3】 带平稳或非平稳有色观测噪声的平稳或非平稳 ARMA 过程(4.7.1)~(4.7.3)有渐近稳定的 k 步 Wiener 预报器 $\hat{y}(t+k|t)$为

$$D(q^{-1})\hat{y}(t+k|t)=G_k(q^{-1})z(t) \tag{4.7.78}$$

其中定义多项式

$$G_k(q^{-1})=K_{k\alpha}(q^{-1})P(q^{-1}) \tag{4.7.79}$$

$$D(q^{-1})=\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})\boldsymbol{K}_pq^{-1}+A(q^{-1})P(q^{-1}) \tag{4.7.80}$$

且有预报误差方差为(4.7.45)和(4.7.46).

证明 将(4.7.77)代入(4.7.58)后两边消去 $A(q^{-1})$得(4.7.78)~(4.7.80).由 $D(q^{-1})$的稳定性引出(4.7.78)是渐近稳定的.证毕. □

下面我们将证明由定理 4.7.2 给出的 Wiener 预报器(4.7.42)是等价于由定理 4.7.3 给出的 Wiener 预报器(4.7.78).

【定理 4.7.4】 Wiener 预报器(4.7.42)与 Wiener 预报器(4.7.78)是等价的，即在数值上它们是相同的，但计算 $G_k(q^{-1})$和 $\psi(q^{-1})$或 $D(q^{-1})$的算法不同.

证明 只需证明 $\psi(q^{-1})=D(q^{-1})$即可.由(4.7.32)有

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_pz(t) \tag{4.7.81}$$

将它代入(4.7.71)有

$$z(t)=\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_pz(t-1)+\varepsilon(t) \tag{4.7.82}$$

将$(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}=\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)/\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)$代入上式得 ARMA 新息模型

$$\Lambda(q^{-1})z(t)=\psi(q^{-1})\varepsilon(t) \tag{4.7.83}$$

$$\psi(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p) \tag{4.7.84}$$

$$\Lambda(q^{-1})=\psi(q^{-1})-\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_pq^{-1} \tag{4.7.85}$$

比较(4.7.75)有

$$D(q^{-1})=\psi(q^{-1}) \tag{4.7.86}$$

$$A(q^{-1})P(q^{-1})=\Lambda(q^{-1}) \tag{4.7.87}$$

证毕. □

4.8 多变量 Box – Jenkins 递推预报器

考虑多变量平稳、可逆的 ARMA 过程

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{4.8.1}$$

其中 $\boldsymbol{y}(t)\in R^m$，白噪声 $\boldsymbol{\varepsilon}(t)\in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声：

$$\mathrm{E}[\boldsymbol{\varepsilon}(t)]=\boldsymbol{0},\quad \mathrm{E}[\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(j)]=\boldsymbol{Q}_\varepsilon\delta_{tj} \tag{4.8.2}$$

$$\boldsymbol{A}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{A}_1\boldsymbol{q}^{-1}+\cdots+\boldsymbol{A}_{n_a}\boldsymbol{q}^{-n_a},$$

$$\boldsymbol{C}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{C}_1q^{-1}+\cdots+\boldsymbol{C}_{n_c}q^{-n_c} \tag{4.8.3}$$

其中初始观测时刻 $t_0\to-\infty$，系数阵 $\boldsymbol{A}_i,\boldsymbol{C}_i$ 均为 $m\times m$ 阵. 假设多项式矩阵 $\boldsymbol{A}(q^{-1})$和 $\boldsymbol{C}(q^{-1})$是稳定的，即多项式 $\det\boldsymbol{A}(x)$和 $\det\boldsymbol{D}(x)$的零点全位于单位圆外. 于是有展式

$$\boldsymbol{y}(t)=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{\varepsilon}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Psi}_jq^{-j}\boldsymbol{\varepsilon}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Psi}_j\boldsymbol{\varepsilon}(t-j) \tag{4.8.4}$$

且有关系

$$\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})=\sum_{j=0}^{\infty}\boldsymbol{\Psi}_jq^{-j} \tag{4.8.5}$$

或

$$\boldsymbol{C}(q^{-1})=\boldsymbol{A}(q^{-1})\sum_{j=0}^{\infty}\boldsymbol{\Psi}_jq^{-j} \tag{4.8.6}$$

用比较上式两边 q^{-j}的系数阵方法引出 $\boldsymbol{\Psi}_j$ 可递推计算为

$$\boldsymbol{\Psi}_j=-\boldsymbol{A}_1\boldsymbol{\Psi}_{j-1}-\cdots-\boldsymbol{A}_{n_a}\boldsymbol{\Psi}_{j-n_a}+\boldsymbol{C}_j \tag{4.8.7}$$

其中规定 $\boldsymbol{\Psi}_j=\boldsymbol{0}(j<0)$，$\boldsymbol{C}_j=\boldsymbol{0}(j>n_c)$. 类似地有展式

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{C}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Pi}_jq^j\boldsymbol{y}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Pi}_j\boldsymbol{y}(t-j) \tag{4.8.8}$$

其中用比较系数法 $m\times m$ 系数阵$\boldsymbol{\Pi}_j$ 可递推计算为

$$\boldsymbol{\Pi}_j=-\boldsymbol{C}_1\boldsymbol{\Pi}_{j-1}-\cdots-\boldsymbol{C}_{n_c}\boldsymbol{\Pi}_{j-n_c}+\boldsymbol{A}_j \tag{4.8.9}$$

其中规定 $\boldsymbol{\Pi}_j=\boldsymbol{0}(j<0)$，$\boldsymbol{A}_j=\boldsymbol{0}(j>n_a)$.

由(4.8.4)和(4.8.8)引出由($\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots$)张成的 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)$相同于由($\boldsymbol{\varepsilon}(t),\boldsymbol{\varepsilon}(t-1),\cdots$)张成的 Hilbert 空间 $L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots)$，即

$$L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)=L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots) \tag{4.8.10}$$

由节 4.1 的射影理论有基于观测($\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots$)的稳态最优预报器

$$\begin{aligned}\hat{\boldsymbol{x}}(t+k\mid t)&=\mathrm{proj}(\boldsymbol{y}(t+k)\mid\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots)=\\&\mathrm{proj}(\boldsymbol{y}(t+k)\mid\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)\end{aligned} \tag{4.8.11}$$

它是线性最小方差预报器，极小化

$$J=E[(\boldsymbol{y}(t+k)-\hat{\boldsymbol{y}}(t+k\mid t))^{\mathrm{T}}(\boldsymbol{y}(t+k)-\hat{\boldsymbol{y}}(t+k\mid t)] \tag{4.8.12}$$

因 $\boldsymbol{\varepsilon}(t)\perp L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-2),\cdots)$取(4.8.1)两边各项到 $L(\boldsymbol{y}^{\mathrm{T}}(t-1),\boldsymbol{y}^{\mathrm{T}}(t-2),\cdots)$上射影引出

$$\hat{\boldsymbol{y}}(t|t-1)=-\boldsymbol{A}_1\boldsymbol{y}(t-1)-\cdots-\boldsymbol{A}_{n_a}\boldsymbol{y}(t-n_a)+\boldsymbol{C}_1\boldsymbol{\varepsilon}(t-1)+\cdots+\boldsymbol{C}_{n_c}\boldsymbol{\varepsilon}(t-n_c) \tag{4.8.13}$$

(4.8.1)减(4.8.13)引出

$$\boldsymbol{y}(t)-\hat{\boldsymbol{y}}(t|t-1)=\boldsymbol{\varepsilon}(t) \tag{4.8.14}$$

即 $\boldsymbol{\varepsilon}(t)$是 $\boldsymbol{y}(t)$的新息过程.

节 4.3 的单变量 Box – Jenkins 预报器容易平行推广到多变量情形.

【定理 4.8.1】 (多变量 Box – Jenkins 递推预报器)多变量平稳、可逆的 ARMA 过程(4.8.1) ~ (4.8.3)有 Box – Jenkins 递推预报器

$$\hat{\boldsymbol{y}}(t+k|t)=-\sum_{i=1}^{n_a}\boldsymbol{A}_i\hat{\boldsymbol{y}}(t+k-i|t)+\sum_{i=k}^{n_c}\boldsymbol{C}_i\boldsymbol{\varepsilon}(t+k-i),\quad k\leqslant n_c \tag{4.8.15}$$

$$\hat{\boldsymbol{y}}(t+k|t)=-\sum_{i=1}^{n_a}\boldsymbol{A}_i\hat{\boldsymbol{y}}(t+k-i|t),\quad k>n_c \tag{4.8.16}$$

其中规定 $\hat{\boldsymbol{y}}(t+k-i|t)=\boldsymbol{y}(t+k-i)\ (t+k-i\leqslant t)$,且 $\boldsymbol{\varepsilon}(t+k-i)$由(4.8.8)计算.在应用中 $\boldsymbol{\varepsilon}(t+k-i)$也可由(4.8.1)任取初值后递推计算为

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{y}(t)+\boldsymbol{A}_1\boldsymbol{y}(t-1)+\cdots+\boldsymbol{A}_{n_a}\boldsymbol{y}(t-n_a)-\boldsymbol{C}_1\boldsymbol{\varepsilon}(t-1)-\cdots-\boldsymbol{C}_{n_c}\boldsymbol{\varepsilon}(t-n_c),\quad t=1,2,\cdots \tag{4.8.17}$$

带初值$(\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(1-n_c),y(0),\cdots,y(1-n_a))$.

【注 4.8.1】 上述定理可推广到对非平稳 ARMA 过程也成立.

4.9 多变量 ARMA 过程的 Åström 预报器

本节用多项式矩阵伪交换原理将单变量 Åström 预报器推广到多变量情形.

考虑多变量 ARMA 过程(4.8.1) ~ (4.8.3),假设$(\boldsymbol{A}(q^{-1}),\boldsymbol{C}(q^{-1}))$左素[18],初始观测时刻 $t_0=-\infty$,且 $\boldsymbol{A}(q^{-1})$和 $\boldsymbol{C}(q^{-1})$是稳定的,问题基于观测$(y(t),y(t-1),\cdots)$求线性最小方差最优预报器 $\hat{\boldsymbol{y}}(t+k|t)$,$k>0$.

(4.8.1)可写为

$$\boldsymbol{y}(t+k)=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{\varepsilon}(t+k) \tag{4.9.1}$$

引入伪交换(右素分解)

$$\boldsymbol{C}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})=\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{C}}^{-1}(q^{-1}) \tag{4.9.2}$$

带 $\det\boldsymbol{C}(q^{-1})=\det\widetilde{\boldsymbol{C}}(q^{-1})$,且 $\widetilde{\boldsymbol{A}}_0=\boldsymbol{I}_m$,$\widetilde{\boldsymbol{C}}_0=\boldsymbol{I}_m$,$n_{\tilde a}=n_a$,$n_{\tilde c}=n_c$,

$$\widetilde{\boldsymbol{A}}(q^{-1})=\widetilde{\boldsymbol{A}}_0+\widetilde{\boldsymbol{A}}_1q^{-1}+\cdots+\widetilde{\boldsymbol{A}}_{n_{\tilde a}}q^{-n_{\tilde a}},$$

$$\widetilde{\boldsymbol{C}}(q^{-1})=\widetilde{\boldsymbol{C}}_0+\widetilde{\boldsymbol{C}}_1q^{-1}+\cdots+\widetilde{\boldsymbol{C}}_{n_{\tilde c}}q^{-n_{\tilde c}} \tag{4.9.3}$$

则(4.9.1)成为

$$\hat{\boldsymbol{y}}(t+k)=\widetilde{\boldsymbol{C}}(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\boldsymbol{\varepsilon}(t+k) \tag{4.9.4}$$

为了将上式右边分解为已知部分和未知部分,引入 Diophantine 方程

$$\widetilde{\boldsymbol{C}}(q^{-1})=\boldsymbol{F}_k(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})+q^{-k}\boldsymbol{G}_k(q^{-1}) \tag{4.9.5}$$

$$\boldsymbol{F}_k(q^{-1})=\boldsymbol{F}_{k0}+\boldsymbol{F}_{k1}q^{-1}+\cdots+\boldsymbol{F}_{k,k-1}q^{-(k-1)},\quad n_f=k-1 \tag{4.9.6}$$

$$G_k(q^{-1}) = G_{k0} + G_{k1}q^{-1} + \cdots + G_{kn_g}q^{-n_g}, \quad n_g = \max(n_a - 1, n_c - k) \tag{4.9.7}$$

其中 $F_k(q^{-1})$ 和 $G_k(q^{-1})$ 可用比较系统法求得. 将 (4.9.5) 代入 (4.9.4) 有

$$\boldsymbol{y}(t+k) = \boldsymbol{F}_k(q^{-1})\boldsymbol{\varepsilon}(t+k) + \boldsymbol{G}_k(q^{-1})\tilde{A}^{-1}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{4.9.8}$$

由 (4.8.1) 和 (4.9.2) 有

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{C}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \tilde{\boldsymbol{A}}(q^{-1})\tilde{\boldsymbol{C}}^{-1}(q^{-1})\boldsymbol{y}(t) \tag{4.9.9}$$

将它代入 (4.9.8) 引出

$$\boldsymbol{y}(t+k) = \boldsymbol{F}_k(q^{-1})\boldsymbol{\varepsilon}(t+k) + \boldsymbol{G}_k(q^{-1})\tilde{\boldsymbol{C}}^{-1}(q^{-1})\boldsymbol{y}(t) \tag{4.9.10}$$

注意上式右边第一项是 $\boldsymbol{\varepsilon}(t+1),\cdots,\boldsymbol{\varepsilon}(t+k)$ 的线性组合，是未知的，而第二项是已知的. 取上式两边各项到 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)$ 上的射影，注意 (4.8.2) 和 (4.8.10) 有

$$\hat{\boldsymbol{y}}(t+k|t) = \mathrm{proj}(\boldsymbol{y}(t+k)|\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots) \tag{4.9.11}$$

$$\boldsymbol{\varepsilon}(t+i) \perp L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots), \quad i>0 \tag{4.9.12}$$

$$L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots) = L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots) \tag{4.9.13}$$

因而稳态最优 k 步预报器为

$$\boldsymbol{y}(t+k|t) = \boldsymbol{G}_k(q^{-1})\tilde{\boldsymbol{C}}^{-1}(q^{-1})\boldsymbol{y}(t) \tag{4.9.14}$$

预报误差 $\tilde{\boldsymbol{y}}(t+k|t) = \boldsymbol{y}(t+k) - \hat{\boldsymbol{y}}(t+k|t)$ 为

$$\tilde{\boldsymbol{y}}(t+k|t) = \boldsymbol{F}_k(q^{-1})\boldsymbol{\varepsilon}(t+k) \tag{4.9.15}$$

预报误差方差阵为

$$\boldsymbol{P}_k = \mathrm{E}[\tilde{\boldsymbol{y}}(t+k|t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+k|t)] = \sum_{i=0}^{k-1}\boldsymbol{F}_{ki}Q_{\varepsilon}\boldsymbol{F}_{ki}^{\mathrm{T}} \tag{4.9.16}$$

上述结果可概括为如下定理.

【定理 4.9.1】 (多变量 Åström 预报器) 平稳、可逆的 $m\times 1$ 维 ARMA 过程 (4.8.1) ~ (4.8.3) 有渐近稳定的 Åström 稳态最优预报器

$$\hat{\boldsymbol{y}}(t+k|t) = \boldsymbol{G}_k(q^{-1})\tilde{\boldsymbol{C}}^{-1}(q^{-1})\boldsymbol{y}(t), \quad k>0 \tag{4.9.17}$$

带伪交换 (4.9.2)，且 $\boldsymbol{G}_k(q^{-1})$ 和 $\boldsymbol{F}_k(q^{-1})$ 由 Diophantine 方程 (4.9.5) 决定，预报误差为

$$\tilde{\boldsymbol{y}}(t+k|t) = \boldsymbol{F}_k(q^{-1})\boldsymbol{\varepsilon}(t+k) \tag{4.9.18}$$

预报误差方差阵为

$$\boldsymbol{P}_k = \sum_{i=0}^{k-1}\boldsymbol{F}_{ki}\boldsymbol{Q}_{\varepsilon}\boldsymbol{F}_{ki}^{\mathrm{T}} \tag{4.9.19}$$

【注 4.9.1】 因 $\det\tilde{\boldsymbol{C}}(q^{-1}) = \det\boldsymbol{C}(q^{-1})$，故由 $\boldsymbol{C}(q^{-1})$ 是稳定的引出 $\tilde{\boldsymbol{C}}(q^{-1})$ 是稳定的，从而 (4.9.17) 是渐近稳定的.

【注 4.9.2】 (4.9.17) 有递推形式

$$\det\tilde{\boldsymbol{C}}(q^{-1})\hat{\boldsymbol{y}}(t+k|t) = \boldsymbol{G}_k(q^{-1})\mathrm{adj}\tilde{\boldsymbol{C}}(q^{-1})\boldsymbol{y}(t) \tag{4.9.20}$$

【例 4.9.1】 考虑平稳、可逆的二维 ARMA(1,1) 过程 $\boldsymbol{y}(t)$，

$$(\boldsymbol{I}_2 + \boldsymbol{A}_1q^{-1})\boldsymbol{y}(t) = (\boldsymbol{I}_2 + \boldsymbol{C}_1q^{-1})\boldsymbol{\varepsilon}(t) \tag{4.9.21}$$

其中 $\boldsymbol{\varepsilon}(t)$ 是零均值、方差阵为 $\boldsymbol{Q}_{\varepsilon}$ 的白噪声，且

$$\boldsymbol{A}_1 = \begin{bmatrix} -0.85 & 0 \\ 0.85 & -0.79 \end{bmatrix}, \quad \boldsymbol{C}_1 = \begin{bmatrix} -0.422 & -0.452 \\ 0.143 & -0.236 \end{bmatrix},$$

$$\boldsymbol{Q}_\varepsilon=\begin{bmatrix}2.849 & -0.863\\ -0.863 & 2.799\end{bmatrix}\tag{4.9.22}$$

问题是求 Åström 预报器 $\hat{\boldsymbol{y}}(t+1\mid t)$.

注意 $\boldsymbol{A}(q^{-1})=\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1}$, $\boldsymbol{C}(q^{-1})=\boldsymbol{I}_2+\boldsymbol{C}_1q^{-1}$，置

$$\widetilde{\boldsymbol{A}}(q^{-1})=\boldsymbol{I}_1+\widetilde{\boldsymbol{A}}_1q^{-1},\quad \widetilde{\boldsymbol{C}}(q^{-1})=\boldsymbol{I}_2+\widetilde{\boldsymbol{C}}_1q^{-1}\tag{4.9.23}$$

由(4.9.2)有

$$(\boldsymbol{I}_2+\boldsymbol{C}_1q^{-1})(\boldsymbol{I}_2+\widetilde{\boldsymbol{A}}_1q^{-1})=(\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1})(\boldsymbol{I}_2+\widetilde{\boldsymbol{C}}_1q^{-1})\tag{4.9.24}$$

比较上式两边 q^{-i}系数阵引出关系

$$\boldsymbol{C}_1+\widetilde{\boldsymbol{A}}_1=\boldsymbol{A}_1+\widetilde{\boldsymbol{C}}_1\tag{4.9.25}$$

$$\boldsymbol{C}_1\widetilde{\boldsymbol{A}}_1=\boldsymbol{A}_1\widetilde{\boldsymbol{C}}_1\tag{4.9.26}$$

它等价于矩阵方程组

$$\begin{bmatrix}\boldsymbol{I}_2 & -\boldsymbol{I}_2\\ \boldsymbol{C}_1 & -\boldsymbol{A}_1\end{bmatrix}\begin{bmatrix}\widetilde{\boldsymbol{A}}_1\\ \widetilde{\boldsymbol{C}}_1\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_2-\boldsymbol{C}_1\\ \boldsymbol{0}\end{bmatrix}\tag{4.9.27}$$

$$\begin{bmatrix}\widetilde{\boldsymbol{A}}_1\\ \widetilde{\boldsymbol{C}}_1\end{bmatrix}=\begin{bmatrix}\boldsymbol{I}_2 & -\boldsymbol{I}_2\\ \boldsymbol{C}_1 & -\boldsymbol{A}_1\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{A}_1-\boldsymbol{C}_1\\ \boldsymbol{0}\end{bmatrix}\tag{4.9.28}$$

经计算有

$$\widetilde{\boldsymbol{A}}_1=\begin{bmatrix}-2.616 & 1.927\\ -1.672 & 0.976\end{bmatrix},\quad \widetilde{\boldsymbol{C}}_1=\begin{bmatrix}-2.188 & 1.475\\ -2.377 & 1.529\end{bmatrix}\tag{4.9.29}$$

解 Diophantine 方程(4.9.5)

$$(\boldsymbol{I}_2+\widetilde{\boldsymbol{C}}_1q^{-1})=\boldsymbol{F}_{10}(\boldsymbol{I}_2+\widetilde{\boldsymbol{A}}_1q^{-1})+q^{-1}\boldsymbol{G}_{10},\tag{4.9.30}$$

这引出

$$\boldsymbol{F}_{10}=0,\quad \boldsymbol{G}(q^{-1})=\boldsymbol{G}_{10}=\widetilde{\boldsymbol{C}}_1-\widetilde{\boldsymbol{A}}_1=\begin{bmatrix}0.4279 & -0.4517\\ -0.7066 & 0.5538\end{bmatrix}\tag{4.9.31}$$

故有 Åström 一步最优预报器

$$\hat{\boldsymbol{y}}(t+1\mid t)=\boldsymbol{G}_{10}(\boldsymbol{I}_2+\widetilde{\boldsymbol{C}}_1q^{-1})^{-1}\boldsymbol{y}(t)\tag{4.9.32}$$

它可递推计算为

$$\det(\boldsymbol{I}_2+\widetilde{\boldsymbol{C}}_1q^{-1})\hat{\boldsymbol{y}}(t+k\mid t)=\boldsymbol{G}_{10}\mathrm{adj}(\boldsymbol{I}_2+\widetilde{\boldsymbol{C}}_1q^{-1})\boldsymbol{y}(t)\tag{4.9.33}$$

4.10 多变量 Koivo 预报器

由 Åström 预报器(4.9.17)不能直接引出递推形式，它的递推形式(4.9.20)的缺点是预报器的阶次增加，并且要求逆矩阵 $\widetilde{\boldsymbol{C}}^{-1}(q^{-1})$. 本节给出可直接表为 ARMA 递推形式的 Koivo 预报器.

【定理 4.10.1】 (多变量 Koivo 预报器)[17]平稳、可逆的 $m\times1$ 维 ARMA 过程(4.8.1) ~ (4.8.3)有渐近稳定的 Koivo 稳态最优预报器

$$\hat{\boldsymbol{y}}(t+k\mid t)=\widetilde{\boldsymbol{C}}^{-1}(q^{-1})\widetilde{\boldsymbol{G}}_k(q^{-1})\boldsymbol{y}(t)\tag{4.10.1}$$

或表为 ARMA 递推形式

$$\widetilde{\boldsymbol{C}}(q^{-1})\hat{\boldsymbol{y}}(t+k\mid t)=\widetilde{\boldsymbol{G}}_k(q^{-1})\boldsymbol{y}(t)\tag{4.10.2}$$

其中 $\boldsymbol{F}_k(q^{-1})$ 和 $\boldsymbol{G}_k(q^{-1})$ 由如下 Diophantine 方程决定

$$\boldsymbol{C}(q^{-1})=\boldsymbol{A}(q^{-1})\boldsymbol{F}_k(q^{-1})+q^{-k}\boldsymbol{G}_k(q^{-1}) \tag{4.10.3}$$

其中 $\boldsymbol{F}_k(q^{-1})$ 和 $\boldsymbol{G}_k(q^{-1})$ 具有形式(4.9.6)和(4.9.7).引入伪交换

$$\boldsymbol{G}_k(q^{-1})\boldsymbol{F}_k^{-1}(q^{-1})=\widetilde{\boldsymbol{F}}_k^{-1}(q^{-1})\widetilde{\boldsymbol{G}}_k(q^{-1}) \tag{4.10.4}$$

带 $\det\boldsymbol{F}_k(q^{-1})=\det\widetilde{\boldsymbol{F}}_k(q^{-1}),\widetilde{\boldsymbol{F}}_{k0}=\boldsymbol{I}_m$.定义

$$\widetilde{\boldsymbol{C}}(q^{-1})=\widetilde{\boldsymbol{F}}_k(q^{-1})\boldsymbol{A}(q^{-1})+q^{-k}\widetilde{\boldsymbol{G}}_k(q^{-1}) \tag{4.10.5}$$

则有

$$\det\widetilde{\boldsymbol{C}}(q^{-1})=\det\boldsymbol{C}(q^{-1}) \tag{4.10.6}$$

$$\widetilde{\boldsymbol{C}}(q^{-1})\boldsymbol{F}_k(q^{-1})=\widetilde{\boldsymbol{F}}_k(q^{-1})\boldsymbol{C}(q^{-1}) \tag{4.10.7}$$

预报误差 $\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{y}(t+k)-\hat{\boldsymbol{y}}(t+k|t)$ 为

$$\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{F}_k(q^{-1})\boldsymbol{\varepsilon}(t+k) \tag{4.10.8}$$

预报误差方差阵为 $\boldsymbol{P}_k=\mathrm{E}[\tilde{\boldsymbol{y}}(t+k|t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+k|t)]$ 为

$$\boldsymbol{P}_k=\sum_{i=0}^{k-1}\boldsymbol{F}_{ki}\boldsymbol{Q}_\varepsilon\boldsymbol{F}_{ki}^{\mathrm{T}} \tag{4.10.9}$$

证明 由(4.10.3)和(4.10.5)有

$$\widetilde{\boldsymbol{F}}_k(q^{-1})\boldsymbol{C}(q^{-1})=\widetilde{\boldsymbol{F}}_k(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{F}_k(q^{-1})+q^{-k}\widetilde{\boldsymbol{F}}_k(q^{-1})\boldsymbol{G}_k(q^{-1}) \tag{4.10.10}$$

$$\widetilde{\boldsymbol{C}}(q^{-1})\boldsymbol{F}_k(q^{-1})=\widetilde{\boldsymbol{F}}_k(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{F}_k(q^{-1})+q^{-k}\widetilde{\boldsymbol{G}}_k(q^{-1})\boldsymbol{F}_k(q^{-1}) \tag{4.10.11}$$

应用(4.10.4)引出(4.10.7)成立.由(4.10.7)有

$$\det\widetilde{\boldsymbol{C}}(q^{-1})\det\boldsymbol{F}_k(q^{-1})=\det\widetilde{\boldsymbol{F}}_k(q^{-1})\det\boldsymbol{C}(q^{-1}) \tag{4.10.12}$$

由 $\det\boldsymbol{F}_k(q^{-1})=\det\widetilde{\boldsymbol{F}}_k(q^{-1})$ 引出(4.10.6)成立.

由(4.8.1)有

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t+k)=\boldsymbol{C}(q^{-1})\boldsymbol{\varepsilon}(t+k) \tag{4.10.13}$$

用 $\widetilde{\boldsymbol{F}}_k(q^{-1})$ 左乘(4.10.13)并应用(4.10.5)有

$$[\widetilde{\boldsymbol{C}}(q^{-1})-q^{-k}\widetilde{\boldsymbol{G}}_k(q^{-1})]\boldsymbol{y}(t+k)=\widetilde{\boldsymbol{F}}_k(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{\varepsilon}(t+k) \tag{4.10.14}$$

利用(4.10.7)有

$$\widetilde{\boldsymbol{C}}(q^{-1})\boldsymbol{y}(t+k)-\widetilde{\boldsymbol{G}}_k(q^{-1})\boldsymbol{y}(t)=\widetilde{\boldsymbol{C}}(q^{-1})\boldsymbol{F}_k(q^{-1})\boldsymbol{\varepsilon}(t+k) \tag{4.10.15}$$

这引出

$$\boldsymbol{y}(t+k)=\widetilde{\boldsymbol{C}}^{-1}(q^{-1})\widetilde{\boldsymbol{G}}_k(q^{-1})\boldsymbol{y}(t)+\boldsymbol{F}_k(q^{-1})\boldsymbol{\varepsilon}(t+k) \tag{4.10.16}$$

取上式两边各项到 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)$ 上的射影引出 Koivo 稳态最优预报器

$$\hat{\boldsymbol{y}}(t+k|t)=\widetilde{\boldsymbol{C}}^{-1}(q^{-1})\widetilde{\boldsymbol{G}}_k(q^{-1})\boldsymbol{y}(t) \tag{4.10.17}$$

或表为 ARMA 递推形式(4.10.2),且预报误差为(4.10.8),因而预报误差方差阵 P_k 为(4.10.9).由(4.10.6)知 $\widetilde{\boldsymbol{C}}(q^{-1})$ 是稳定的,故(4.10.1)是渐近稳定的.证毕. □

4.11 多变量非平稳 ARMA 过程的 Wiener 预报器

考虑多变量非平稳 ARMA 过程 $\boldsymbol{y}(t)$,

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{4.11.1}$$

其中 $\boldsymbol{y}(t) \in R^m$，其中 $\boldsymbol{e}(t) \in R^m$，$\boldsymbol{e}(t) \in R^m$，是零均值、方差阵为 $\boldsymbol{Q}_e$ 的白噪声，

$$\boldsymbol{A}(q^{-1}) = \boldsymbol{I}_m + \boldsymbol{A}_1 \boldsymbol{q}^{-1} + \cdots + \boldsymbol{A}_{n_a} \boldsymbol{q}^{-n_a},$$

$$\boldsymbol{C}(q^{-1}) = \boldsymbol{I}_m + \boldsymbol{C}_1 q^{-1} + \cdots + \boldsymbol{C}_{n_c} q^{-n_c} \tag{4.11.2}$$

其中 $\boldsymbol{A}_i$，$\boldsymbol{C}_i$ 为 $m \times m$ 系数阵，$\boldsymbol{A}(q^{-1})$ 是不稳定的(即 $\det\boldsymbol{A}(x)$ 的零点不全位于单位圆外)，而 $\boldsymbol{C}(q^{-1})$ 是稳定的. 问题是基于观测 $(\boldsymbol{y}(t), \boldsymbol{y}(t-1), \cdots)$ 求 Wiener 预报器 $\hat{\boldsymbol{y}}(t+k \mid t)$.

问题的难点在于 $\boldsymbol{A}(q^{-1})$ 是不稳定的，这引起当 $t \to +\infty$ 或 $t_0 \to -\infty$，$\boldsymbol{y}(t)$ 的方差阵是无界的，因而不能利用无穷维 Hilbert 空间射影理论解决问题. 这相当于处理不稳定系统的预报问题. 用稳态 Kalman 滤波理论可解决上述问题. 因为 Kalman 滤波理论可处理不稳定系统(状态转移阵是不稳定矩阵)，可处理非平稳随机信号，可处理完全可观的随机系统的状态估计问题.

非平稳 ARMA 过程(4,11,1)有状态空间模型

$$\boldsymbol{x}(t+1) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{\Gamma}\boldsymbol{e}(t) \tag{4.11.3}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{e}(t) \tag{4.11.4}$$

其中 $n = \max(n_a, n_c)$，$\boldsymbol{A}_i = 0\,(i > n_a)$，$\boldsymbol{C}_i = 0\,(i > n_c)$，且

$$\boldsymbol{A} = \begin{bmatrix} -\boldsymbol{A}_1 & & & \\ \vdots & & \boldsymbol{I}_{m(n-1)} & \\ -\boldsymbol{A}_n & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{C}_1 - \boldsymbol{A}_1 \\ \vdots \\ \boldsymbol{C}_n - \boldsymbol{A}_n \end{bmatrix}, \quad \boldsymbol{H} = [\boldsymbol{I}_m \ \boldsymbol{0} \cdots \boldsymbol{0}] \tag{4.11.5}$$

注意系统(4.11.3)和(4.11.4)是不稳定系统，因为 $\boldsymbol{A}$ 是不稳定矩阵. 这是由于

$$\det(\boldsymbol{I}_m - q^{-1}\boldsymbol{A}) = \det\boldsymbol{A}(q^{-1}) \tag{4.11.6}$$

而 $\boldsymbol{A}(q^{-1})$ 是不稳定的. 还应注意该系统是带相关噪声系统，观测噪声 $\boldsymbol{e}(t)$ 与输入白噪声 $\boldsymbol{e}(t)$ 相关，即

$$\boldsymbol{S} = \mathrm{E}[\boldsymbol{e}(t)\boldsymbol{e}^{\mathrm{T}}(t)] = \boldsymbol{Q}_e \tag{4.11.7}$$

该系统可化为带不相关噪声系统[14]，它的状态转移阵为 $\bar{\boldsymbol{\Phi}} = \boldsymbol{A} - \boldsymbol{J}\boldsymbol{H}$，$\boldsymbol{J} = \boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_e^{-1} = \boldsymbol{\Gamma}$，即

$$\bar{\boldsymbol{\Phi}} = \begin{bmatrix} -\boldsymbol{C}_1 & & & \\ \vdots & & \boldsymbol{I}_{m(n-1)} & \\ -\boldsymbol{C}_n & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix} \tag{4.11.8}$$

因 $(\bar{\boldsymbol{\Phi}}, \boldsymbol{H})$ 为块伴随形[14]，故系统(4.11.3)和(4.11.4)存在稳态 Kalman 预报器[15]

$$\hat{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{A}\hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{K}_p \boldsymbol{\varepsilon}(t) \tag{4.11.9}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{\varepsilon}(t) \tag{4.11.10}$$

其中白噪声 $\boldsymbol{\varepsilon}(t) \in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的 $\boldsymbol{y}(t)$ 的稳态新息过程. 因由(4.11.5)$(\boldsymbol{A}, \boldsymbol{H})$为块伴随形，故由定理 1.6.1 上两式等价于 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{4.11.11}$$

$$\boldsymbol{D}(q^{-1}) = \boldsymbol{I}_m + \boldsymbol{D}_1 q^{-1} + \cdots + \boldsymbol{D}_{n_d} q^{-n_d}, \quad n_d = n \tag{4.11.12}$$

$$\boldsymbol{K}_p = \begin{bmatrix} \boldsymbol{K}_{p1} \\ \vdots \\ \boldsymbol{K}_{pn} \end{bmatrix} \tag{4.11.13}$$

$$\boldsymbol{K}_{pi}=\boldsymbol{D}_i-\boldsymbol{A}_i,\quad i=1,\cdots,n \tag{4.11.14}$$

或

$$\boldsymbol{D}_i=\boldsymbol{K}_{pi}-\boldsymbol{A}_i,\quad i=1,\cdots,n \tag{4.11.15}$$

且矩阵

$$\boldsymbol{\Psi}_p=\boldsymbol{A}-\boldsymbol{K}_p\boldsymbol{H}=\begin{bmatrix} -\boldsymbol{D}_1 & & & \\ \vdots & \boldsymbol{I}_{m(n-1)} & & \\ -\boldsymbol{D}_n & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix} \tag{4.11.16}$$

是稳定矩阵[15,16].注意[17]

$$\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)=\det\boldsymbol{D}(q^{-1}) \tag{4.11.17}$$

因而 $\boldsymbol{D}(q^{-1})$是稳定的多项式矩阵.比较(4.11.1)与(4.11.11)有关系

$$\boldsymbol{C}(q^{-1})\boldsymbol{e}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{4.11.18}$$

因为两个 MA 过程 $\boldsymbol{C}(q^{-1})\boldsymbol{e}(t)$与 $\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)$有相同的相关函数,$\boldsymbol{C}(q^{-1})$和 $\boldsymbol{D}(q^{-1})$均为稳定的,且 $\boldsymbol{D}_0=\boldsymbol{I}_m$,$\boldsymbol{C}_0=\boldsymbol{I}_m$,因而在谱等价(有相同的二阶矩)意义下有[3]

$$\boldsymbol{C}(q^{-1})=\boldsymbol{D}(q^{-1}),\quad \boldsymbol{e}(t)=\boldsymbol{\varepsilon}(t),\quad \boldsymbol{Q}_e=\boldsymbol{Q}_\varepsilon \tag{4.11.19}$$

注意,其中当 $n_a>n_c$ 时,$n=n_a$,$\boldsymbol{D}_i=0\ (i>n_c)$.于是由(4.11.13)有

$$\boldsymbol{K}_p=\begin{bmatrix}\boldsymbol{C}_1-\boldsymbol{A}_1\\ \vdots\\ \boldsymbol{C}_n-\boldsymbol{A}_n\end{bmatrix} \tag{4.11.20}$$

注意(4.11.19)还可用另一种方法证明如下:

比较(4.11.1)与(4.11.11),令 $\boldsymbol{r}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)$,则 $\boldsymbol{r}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t)$且 $\boldsymbol{r}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)$,因而 $\boldsymbol{C}(q^{-1})\boldsymbol{e}(t)$与 $\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)$具有相同的相关函数,它们都等于 MA 过程 $\boldsymbol{r}(t)$的相关函数.另一方面,稳态一步 Kalman 预报误差方差阵 $\boldsymbol{\Sigma}$ 满足 Riccati 方程[14]

$$\boldsymbol{\Sigma}=\overline{\boldsymbol{\Phi}}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_e)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}]\overline{\boldsymbol{\Phi}}+\boldsymbol{\Gamma}(\boldsymbol{Q}_e-\boldsymbol{S}\boldsymbol{Q}_e^{-1}\boldsymbol{S}^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}} \tag{4.11.21}$$

因 $\boldsymbol{Q}_e-\boldsymbol{S}\boldsymbol{Q}_e^{-1}\boldsymbol{S}^{\mathrm{T}}=\boldsymbol{Q}_e-\boldsymbol{Q}_e\boldsymbol{Q}_e^{-1}\boldsymbol{Q}_e=\boldsymbol{0}$,则稳态 Riccati 方程(4.11.21)化为

$$\boldsymbol{\Sigma}=\overline{\boldsymbol{\Phi}}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_e)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}]\overline{\boldsymbol{\Phi}}^{\mathrm{T}} \tag{4.11.22}$$

它有非负定解

$$\boldsymbol{\Sigma}=\boldsymbol{0} \tag{4.11.23}$$

而稳态新息 $\boldsymbol{\varepsilon}(t)$的方差阵 $\boldsymbol{Q}_\varepsilon$ 为[14]

$$\boldsymbol{Q}_\varepsilon=\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_e \tag{4.11.24}$$

将(4.11.23)代入上式引出

$$\boldsymbol{Q}_\varepsilon=\boldsymbol{Q}_e \tag{4.11.25}$$

又 $\mathrm{E}[\boldsymbol{\varepsilon}(t)]=\mathrm{E}[\boldsymbol{e}(t)]=\boldsymbol{0}$,$\boldsymbol{\varepsilon}(t)$和 $\boldsymbol{e}(t)$均为 $m\times 1$ 维白噪声,因此在谱等价(即有相同的二阶矩)意义下有

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{e}(t) \tag{4.11.26}$$

再由 MA 过程 $\boldsymbol{C}(q^{-1})\boldsymbol{\varepsilon}(t)$与 $\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)$均为可逆的,即 $\boldsymbol{C}(q^{-1})$与 $\boldsymbol{D}(q^{-1})$均为稳定的,且 $\boldsymbol{C}_0=\boldsymbol{D}_0=\boldsymbol{I}_m$,故有[3]

$$\boldsymbol{D}(q^{-1})=\boldsymbol{C}(q^{-1}) \tag{4.11.27}$$

即首系数阵为 $\boldsymbol{I}_m$ 的具有已知相关函数的可逆的 MA 过程的 MA 多项式矩阵是唯一的[3].

注意稳态 Kalman 预报器(4.11.9)和(4.11.10)可写为 Wiener 预报器形式

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\boldsymbol{y}(t) \tag{4.11.28}$$

或表为传递函数形式

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{y}(t) \tag{4.11.29}$$

注意由(4.11.3)和(4.11.4)及射影性质引出稳态预报器

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{A}^{k-1}\hat{\boldsymbol{x}}(t+1|t),\quad k>0 \tag{4.11.30}$$

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t+k|t) \tag{4.11.31}$$

于是有

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{H}\boldsymbol{A}^{k-1}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{y}(t) \tag{4.11.32}$$

【定理 4.11.1】 多变量非平稳 ARMA 过程(4.11.1),其中假设 $\boldsymbol{A}(q^{-1})$ 是不稳定的多项式矩阵,而 $\boldsymbol{C}(q^{-1})$ 是稳定的多项式矩阵,有渐近稳定的 Wiener 预报器

$$\det\boldsymbol{C}(q^{-1})\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{G}_k(q^{-1})\boldsymbol{y}(t) \tag{4.11.33}$$

其中定义多项式矩阵 $\boldsymbol{G}_k(q^{-1})$ 为

$$\boldsymbol{G}_k(q^{-1})=\boldsymbol{H}\boldsymbol{A}^{k-1}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p=\boldsymbol{G}_{k0}+\boldsymbol{G}_{k1}q^{-1}+\cdots+\boldsymbol{G}_{k,n-1}q^{-(n-1)} \tag{4.11.34}$$

其中 $\boldsymbol{H},\boldsymbol{A},\boldsymbol{K}_p$ 分别由(4.11.5),(4.11.20)计算,而由(4.11.16) $\boldsymbol{\Psi}_p$ 为

$$\boldsymbol{\Psi}_p=\begin{bmatrix}-\boldsymbol{C}_1 & & & \\ \vdots & \boldsymbol{I}_{m(n-1)} & & \\ -\boldsymbol{C}_n & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix} \tag{4.11.35}$$

其中规定 $\boldsymbol{C}_i=\boldsymbol{0}\ (i>n_a)$. 预报误差 $\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{y}(t+k)-\hat{\boldsymbol{y}}(t+k|t)$ 方差阵 $P_k^{\gamma}=\mathrm{E}[\tilde{\boldsymbol{y}}(t+k|t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+k|t)$ 为

$$\boldsymbol{P}_k^{\gamma}=\boldsymbol{H}\boldsymbol{P}_k\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_e \tag{4.11.36}$$

其中稳态 k 步 Kalman 预报误差方差阵为 $\boldsymbol{P}_k=\mathrm{E}[\tilde{\boldsymbol{x}}(t+k|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+k|t)]$, $\tilde{\boldsymbol{x}}(t+k|t)=\boldsymbol{x}(t+k)-\hat{\boldsymbol{x}}(t+k|t)$,

$$\boldsymbol{P}_k=\boldsymbol{A}^{k-1}\boldsymbol{\Sigma}(\boldsymbol{A}^{k-1})^{\mathrm{T}}+\sum_{j=0}^{k-2}\boldsymbol{A}^j\boldsymbol{\Gamma}\boldsymbol{Q}_e\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{A}^{j\mathrm{T}},\quad k>1 \tag{4.11.37}$$

证明 将 $(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}=\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)/\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)$ 代入(4.11.32),并应用(4.11.16),(4.11.17),(4.11.27)得(4.11.33)~(4.11.35). 由(4.11.4)和(4.11.31)有

$$\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t+k|t)+\boldsymbol{e}(t+k) \tag{4.11.38}$$

易知 $\boldsymbol{e}(t+k)\perp\tilde{\boldsymbol{x}}(t+k|t)$,于是有(4.11.36)成立. 由第三章有关结果有(4.11.37). 由 $\boldsymbol{C}(q^{-1})$ 是稳定的有 $\det\boldsymbol{C}(q^{-1})$ 是稳定的多项式,故(4.11.33)是渐近稳定的. 证毕. □

【注 4.11.1】 Wiener 预报器(4.11.33)与多变量 Åström 预报器(4.9.20)具有相同的形式. 这只要注意在(4.9.20)中 $\det\tilde{\boldsymbol{C}}(q^{-1})=\det\boldsymbol{C}(q^{-1})$.

4.12 带白色观测噪声的多变量 ARMA 过程的 Wiener 预报器

本节把节 4.6 的带白色观测噪声的单变量 ARMA 过程 Wiener 预报器推广到多变量

情形.文献[32]曾用频域法解决这个问题,这里介绍由作者提出的两种时域方法.[14,17]

考虑带白色观测噪声的多变量 ARMA 过程 $y(t)$

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{4.12.1}$$

$$\boldsymbol{z}(t)=\boldsymbol{y}(t)+\boldsymbol{v}(t) \tag{4.12.2}$$

其中 $\boldsymbol{y}(t)\in R^m$ 为待估信号,$\boldsymbol{z}(t)\in R^m$ 为对 $y(t)$ 的观测信号,$\boldsymbol{v}(t)\in R^m$ 为白色观测噪声,$\boldsymbol{e}(t)\in R^r$ 和 $\boldsymbol{v}(t)$ 是带零均值、方差各为 $\boldsymbol{Q}_e$ 和 $\boldsymbol{Q}_v$ 的独立白噪声,q^{-1}为单位滞后算子,

$$\boldsymbol{A}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{A}_1\boldsymbol{q}^{-1}+\cdots+\boldsymbol{A}_{n_a}\boldsymbol{q}^{-n_a},$$

$$\boldsymbol{C}(q^{-1})=\boldsymbol{C}_1q^{-1}+\cdots+\boldsymbol{C}_{n_c}q^{-n_c} \tag{4.12.3}$$

其中 $\boldsymbol{A}_i$ 为 $m\times m$ 系数阵,$\boldsymbol{C}_i$ 为 $m\times r$ 系数阵.假设 $\boldsymbol{A}(q^{-1})$是稳定的,且($\boldsymbol{A}(q^{-1})$,$\boldsymbol{C}(q^{-1})$)左素,初始观测时刻 $t_0=-\infty$,问题是基于观测($\boldsymbol{z}(t)$,$\boldsymbol{z}(t-1)$,…)求最优 Wiener 预报器 $\hat{\boldsymbol{y}}(t+k|t)$.

4.12.1 基于多变量 ARMA 新息模型的预报方法

将(4.12.1)代入(4.12.2)可得 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{z}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{4.12.4}$$

其中 $\boldsymbol{D}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{D}_1q^{-1}+\cdots+\boldsymbol{D}_{n_d}q^{-n_d}$是稳定的,$n_d=\max(n_a,n_c-1)$,新息 $\boldsymbol{\varepsilon}(t)\in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声,且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t)+\boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{4.12.5}$$

$\boldsymbol{D}(q^{-1})$和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers - Wouters 算法求得.由 $\boldsymbol{A}(q^{-1})$和 $\boldsymbol{D}(q^{-1})$都是稳定的,故(4.12.4)是平稳、可逆的 ARMA 过程,于是有相同的 Hilbert 空间

$$L(\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)=L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots) \tag{4.12.6}$$

且有

$$\begin{aligned}\hat{\boldsymbol{y}}(t+k|t)=\mathrm{proj}(\boldsymbol{y}(t+k)|\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots)=\\ \mathrm{proj}(\boldsymbol{y}(t+k)|\boldsymbol{\varepsilon}^{\mathrm{T}}(t),\boldsymbol{\varepsilon}^{\mathrm{T}}(t-1),\cdots)\end{aligned} \tag{4.12.7}$$

由(4.12.2)引出

$$\hat{\boldsymbol{y}}(t+k|t)=\hat{\boldsymbol{z}}(t+k|t) \tag{4.12.8}$$

由定理(4.9.1),基于(4.12.4)有多变量 Åström 预报器为

$$\hat{\boldsymbol{z}}(t+k|t)=\boldsymbol{G}_k(q^{-1})\widetilde{\boldsymbol{D}}^{-1}q^{-1}\boldsymbol{z}(t),\quad k>0 \tag{4.12.9}$$

$$\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})=\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{4.12.10}$$

$$\widetilde{\boldsymbol{D}}(q^{-1})=\boldsymbol{F}_k(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})+q^{-k}\boldsymbol{G}_k(q^{-1}) \tag{4.12.11}$$

$$\det\boldsymbol{D}(q^{-1})=\det\widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{4.12.12}$$

于是有如下定理.

【定理 4.12.1】 带白色观测噪声的平稳的多变量平稳 ARMA 过程有渐近稳定的 Wiener 预报器

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{G}_k(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{z}(t) \tag{4.12.13}$$

预报误差 $\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{y}(t+k)-\hat{\boldsymbol{y}}(t+k|t)$为

$$\tilde{\boldsymbol{y}}(t+k|t)=\tilde{\boldsymbol{z}}(t+k|t)-\boldsymbol{v}(t+k) \tag{4.12.14}$$

$$\tilde{z}(t+k\mid t)=z(t+k)-\hat{z}(t+k\mid t)=$$
$$F_k(q^{-1})\boldsymbol{\varepsilon}(t+k)=\sum_{i=0}^{k-1}F_{ki}\boldsymbol{\varepsilon}(t+k-i) \tag{4.12.15}$$

其中 $\boldsymbol{\varepsilon}(t+k)$ 由 ARMA 新息模型(4.12.4)决定

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{z}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Pi}_j q^{-j}\boldsymbol{z}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Pi}_j\boldsymbol{z}(t-j) \tag{4.12.16}$$

其中系数阵 $\boldsymbol{\Pi}_j$ 可递推计算为

$$\boldsymbol{\Pi}_j=-\boldsymbol{D}_1\boldsymbol{\Pi}_{j-1}-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{\Pi}_{j-n_d}+\boldsymbol{A}_j \tag{4.12.17}$$

其中规定 $\boldsymbol{\Pi}_j=\boldsymbol{0}\,(j<0)$，$\boldsymbol{A}_j=\boldsymbol{0}\,(j>n_a)$. 预报误差方差阵 $\boldsymbol{P}_k^{y}=\mathrm{E}[\tilde{\boldsymbol{y}}(t+k\mid t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+k\mid t)]$ 为

$$\boldsymbol{P}_k^{y}=\sum_{i=0}^{k-1}\boldsymbol{F}_{ki}\boldsymbol{Q}_{\varepsilon}\boldsymbol{F}_{ki}^{\mathrm{T}}-\boldsymbol{Q}_v \tag{4.12.18}$$

证明 由(4.12.8)和定理(4.9.1)引出(4.12.13)和(4.12.15). 由(4.12.2)和(4.12.8)引出(4.12.14)和(4.12.18). 这里用到了事实

$$\mathrm{E}[\tilde{\boldsymbol{z}}(t+k\mid t)\boldsymbol{v}^{\mathrm{T}}(t+k)]=\boldsymbol{Q}_v,\quad k>0 \tag{4.12.19}$$

因为 $\tilde{\boldsymbol{z}}(t+k\mid t)=\boldsymbol{z}(t+k)-\hat{\boldsymbol{z}}(t+k\mid t)=\boldsymbol{y}(t+k)+\boldsymbol{v}(t+k)-\hat{\boldsymbol{z}}(t+k\mid t)$，而 $\boldsymbol{v}(t+k)\perp\boldsymbol{y}(t+k)$，$\boldsymbol{v}(t+k)\perp\hat{\boldsymbol{z}}(t+k\mid t)$. 证毕. □

【注 4.12.1】 由(4.12.13)有 ARMA 递推 Wiener 预报器

$$\det\tilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{y}}(t+k\mid t)=\boldsymbol{G}_k(q^{-1})\mathrm{adj}\tilde{\boldsymbol{D}}(q^{-1})\boldsymbol{z}(t) \tag{4.12.20}$$

因 $\det\boldsymbol{D}(q^{-1})$ 为稳定多项式，故(4.12.20)或(4.12.9)是渐近稳定的.

4.12.2 基于 Kalman 预报器的预报方法

考虑多变量 ARMA 过程(4.12.1)~(4.12.3)，这里可以假设过程是非平稳的，即假设 $\boldsymbol{A}(q^{-1})$ 是不稳定的. 由定理 1.6.1 该系统有状态空间模型

$$\boldsymbol{x}(t=1)=\bar{\boldsymbol{A}}\boldsymbol{x}(t)+\bar{\boldsymbol{C}}\boldsymbol{e}(t) \tag{4.12.21}$$
$$\boldsymbol{z}(t)=\boldsymbol{H}\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{4.12.22}$$

其中 $n=\max(n_a,n_c)$，$\boldsymbol{A}_i=\boldsymbol{0}\,(i>n_a)$，$\boldsymbol{C}_i=\boldsymbol{0}\,(i>n_c)$，且

$$\bar{\boldsymbol{A}}=\begin{bmatrix}-\boldsymbol{A}_1 & & \\ \vdots & \boldsymbol{I}_{m(n-1)} & \\ -\boldsymbol{A}_n & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix},\quad \bar{\boldsymbol{C}}=\begin{bmatrix}\boldsymbol{C}_1\\ \vdots\\ \boldsymbol{C}_n\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}] \tag{4.12.23}$$

由$(\boldsymbol{A}(q^{-1}),\boldsymbol{C}(q^{-1}))$左素的假设引出系统(4.12.21)和(4.12.22)是完全可观、完全可控的.[19]这是因为 $\boldsymbol{y}(t)=\boldsymbol{H}\boldsymbol{x}(t)$ 有传递函数表示

$$\boldsymbol{y}(t)=\boldsymbol{H}[\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})^{-1}\bar{\boldsymbol{C}}q^{-1}\boldsymbol{e}(t)=\frac{[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}]\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})\bar{\boldsymbol{C}}q^{-1}}{\det(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})}\boldsymbol{e}(t)$$
$$\frac{\mathrm{adj}\boldsymbol{A}(q^{-1})[\boldsymbol{I}_m,\boldsymbol{I}_m q^{-1},\cdots,\boldsymbol{I}_m q^{-(n-1)}]\bar{\boldsymbol{C}}q^{-1}}{\det\boldsymbol{A}(q^{-1})}=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{4.12.24}$$

其中注意$(\bar{\boldsymbol{A}},\boldsymbol{H})$为块伴随形，因而由定理 1.6.2 有等式

$$\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})=\mathrm{adj}\boldsymbol{A}(q^{-1})[\boldsymbol{I}_m,\boldsymbol{I}_m q^{-1},\cdots,\boldsymbol{I}_m q^{-(n-1)}],$$
$$\det(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})=\det\boldsymbol{A}(q^{-1}) \tag{4.12.25}$$

于是存在稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\boldsymbol{z}(t) \tag{4.12.26}$$

$$\boldsymbol{\Psi}_p=\bar{\boldsymbol{A}}-\boldsymbol{K}_p\boldsymbol{H} \tag{4.12.27}$$

$$\boldsymbol{K}_p=\bar{\boldsymbol{A}}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v]^{-1} \tag{4.12.28}$$

其中稳态预报误差方差阵 $\boldsymbol{\Sigma}$ 是如下稳态 Riccati 方程的惟一正定解:

$$\boldsymbol{\Sigma}=\bar{\boldsymbol{A}}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}]\bar{\boldsymbol{A}}^{\mathrm{T}}+\bar{\boldsymbol{C}}\boldsymbol{Q}_e\bar{\boldsymbol{C}}^{\mathrm{T}} \tag{4.12.29}$$

它可用迭代法求解.此外可证明[16] $\boldsymbol{\Psi}_p$ 是一个稳定矩阵.

比较(4.12.2)和(4.12.22)有关系

$$\boldsymbol{y}(t)=\boldsymbol{H}\boldsymbol{x}(t) \tag{4.12.30}$$

这引出

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t+k|t) \tag{4.12.31}$$

且有预报误差系

$$\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{H}\tilde{\boldsymbol{x}}(t+k|t) \tag{4.12.32}$$

其中 $\tilde{\boldsymbol{x}}(t+k|t)=\boldsymbol{x}(t+k)-\tilde{\boldsymbol{x}}(t+k|t)$. 又由(4.12.21)和射影性质有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+k|t)=\bar{\boldsymbol{A}}^{k-1}\hat{\boldsymbol{x}}(t+1|t) \tag{4.12.33}$$

由第三章有稳态误差方差阵 $\boldsymbol{P}_k=\mathrm{E}[\tilde{x}(t+k|t)\tilde{x}^{\mathrm{T}}(t+k|t)]$为

$$\boldsymbol{P}_k=\bar{\boldsymbol{A}}^{k-1}\boldsymbol{\Sigma}(\bar{\boldsymbol{A}}^{k-1})^{\mathrm{T}}+\sum_{j=0}^{k-2}\bar{\boldsymbol{A}}^{j}\bar{\boldsymbol{C}}\boldsymbol{Q}_e\bar{\boldsymbol{C}}^{\mathrm{T}}\bar{\boldsymbol{A}}^{j\mathrm{T}},\quad k>1 \tag{4.12.34}$$

且 $\boldsymbol{P}_1=\boldsymbol{\Sigma}$. 由(4.12.26)有 Wiener 预报器

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{z}(t) \tag{4.12.35}$$

将它代入(4.12.31)和(4.12.33)有

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{H}\bar{\boldsymbol{A}}^{k-1}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{z}(t) \tag{4.12.36}$$

定义

$$\psi(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p) \tag{4.12.37}$$

$$\boldsymbol{G}_k(q^{-1})=\boldsymbol{H}\bar{\boldsymbol{A}}^{k-1}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p \tag{4.12.38}$$

则有 Wiener 预报器

$$\psi(q^{-1})\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{G}_k(q^{-1})\boldsymbol{z}(t) \tag{4.12.39}$$

且有预报误差方差阵 $P_k^y=\mathrm{E}[\tilde{y}(t+k|t)\tilde{y}^{\mathrm{T}}(t+k|t)]$为

$$\boldsymbol{P}_k^y=\boldsymbol{H}\boldsymbol{P}_k\boldsymbol{H}^{\mathrm{T}} \tag{4.12.40}$$

其中 $\boldsymbol{P}_k$ 由(4.12.34)计算. 由 $\boldsymbol{\Psi}_p$ 为稳定矩阵引出 $\psi(q^{-1})$是一个稳定多项式,因而(4.12.39)是渐近稳定的.

上述结果可概括为如下定理.

【定理 4.12.2】 带白色观测噪声的平稳或非平稳 ARMA 过程(4.12.1)~(4.12.3)有渐近稳定的 Wiener 预报器(4.12.39),预报误差方差阵为(4.12.40)和(4.12.34).

下面我们证明:Wiener 预报器(4.12.39)与 Wiener 预报器(4.12.20)是等价的,即用两种方法引出相同的结果.

注意状态空间新息模型为

$$\hat{\boldsymbol{x}}(t+1|t)=\bar{\boldsymbol{A}}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\boldsymbol{\varepsilon}(t) \tag{4.12.41}$$

$$\boldsymbol{z}(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{\varepsilon}(t) \tag{4.12.42}$$

这引出

$$\boldsymbol{z}(t)=\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})^{-1}\boldsymbol{K}_p\boldsymbol{\varepsilon}(t)+\boldsymbol{\varepsilon}(t) \tag{4.12.43}$$

注意 $(\bar{\boldsymbol{A}},\boldsymbol{H})$ 为块伴随形，于是有

$$\det(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})=\det\boldsymbol{A}(q^{-1}) \tag{4.12.44}$$

$$\boldsymbol{H}\operatorname{adj}(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})=\operatorname{adj}\boldsymbol{A}(q^{-1})[\boldsymbol{I}_m,\boldsymbol{I}_m q^{-1},\cdots,\boldsymbol{I}_m q^{-(n-1)}] \tag{4.12.45}$$

引入分块表示

$$\boldsymbol{K}_p=\begin{bmatrix}\boldsymbol{K}_{p1}\\ \vdots\\ \boldsymbol{K}_{pn}\end{bmatrix} \tag{4.12.46}$$

其中 $\boldsymbol{K}_{pi}$ 为 $m\times m$ 阵. 将 $(\boldsymbol{I}_n-q^{-1}\boldsymbol{A})^{-1}=\operatorname{adj}(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})/\det(\boldsymbol{I}_n-q^{-1}\bar{\boldsymbol{A}})$ 代入 (4.12.43) 并利用上面三个式子引出 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{z}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{4.12.47}$$

其中 $\boldsymbol{D}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{D}_1q^{-1}+\cdots+\boldsymbol{D}_nq^{-n}$，且有关系

$$\boldsymbol{D}_i=\boldsymbol{K}_{pi}+\boldsymbol{A}_i,\quad i=1,\cdots,n \tag{4.12.48}$$

比较用两种不同方法引出的 ARMA 新息模型 (4.12.4) 与 (4.12.47)，在 (4.12.4) 中，$\boldsymbol{D}(q^{-1})$ 的阶次为 $n_d=\max(n_a,n_c-1)$，而在 (4.12.47) 中 $\boldsymbol{D}(q^{-1})$ 的阶次为 $n=\max(n_a,n_c)$，可知后者在 $n_c-1>n_a$ 情况下，将有 $\boldsymbol{D}_n=\boldsymbol{0}$. 注意，应用 (4.12.48) 有

$$\boldsymbol{\Psi}_p=\bar{\boldsymbol{A}}-\boldsymbol{K}_p\boldsymbol{H}=\begin{bmatrix}-(\boldsymbol{A}_1+\boldsymbol{K}_{p1}) & & \\ \vdots & \boldsymbol{I}_{m(n-1)} & \\ -(\boldsymbol{A}_n+\boldsymbol{K}_{pn}) & \boldsymbol{0} \quad \cdots & \boldsymbol{0}\end{bmatrix}=\begin{bmatrix}-\boldsymbol{D}_1 & & \\ \vdots & \boldsymbol{I}_{m(n-1)} & \\ -\boldsymbol{D}_n & \boldsymbol{0}\quad \cdots & \boldsymbol{0}\end{bmatrix} \tag{4.12.49}$$

这引出

$$\psi(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)=\det\boldsymbol{D}(q^{-1}) \tag{4.12.50}$$

于是 Wiener 预报器 (4.12.39) 成为

$$\det\boldsymbol{D}(q^{-1})\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{G}_k(q^{-1})\boldsymbol{z}(t) \tag{4.12.51}$$

它与 (4.12.20) 具有相同形式. 因为 $\det\tilde{\boldsymbol{D}}(q^{-1})=\det\boldsymbol{D}(q^{-1})$.

这引出如下定理.

【定理 4.12.3】 用两种不同方法给出的 Wiener 预报器 (4.12.20) 与 (4.12.39) 是等价的.

【例 4.12.1】 考虑两通道系统

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{4.12.52}$$

$$\boldsymbol{z}(t)=\boldsymbol{y}(t)+\boldsymbol{v}(t) \tag{4.12.53}$$

其中待估信号 $\boldsymbol{y}(t)=[y_1(t),y_2(t)]^{\mathrm{T}}$，观测信号 $\boldsymbol{z}(t)=[z_1(t),z_2(t)]^{\mathrm{T}}$，观测白噪声 $\boldsymbol{v}(t)=[v_1(t),v_2(t)]^{\mathrm{T}}$，$\boldsymbol{e}(t)=[e_1(t),e_2(t)]^{\mathrm{T}}$，和 $\boldsymbol{v}(t)$ 是零均值、方差阵各为 $\boldsymbol{Q}_\varepsilon$ 和 $\boldsymbol{Q}_v$ 的独立白噪声，且

$$\boldsymbol{Q}_e=0.0025\boldsymbol{I}_2,\quad \boldsymbol{Q}_v=\boldsymbol{I}_2 \tag{4.12.54}$$

$\boldsymbol{A}(q^{-1}) = \boldsymbol{I}_2 + \boldsymbol{A}_1 q^{-1}$，$\boldsymbol{C}(q^{-1}) = \boldsymbol{C}_1 q^{-1}$，且

$$\boldsymbol{A}_1 = \begin{bmatrix} 0.9 & 0.4 \\ 0.4 & -0.9 \end{bmatrix}, \quad \boldsymbol{C}_1 = \begin{bmatrix} 0.2 & 0.9 \\ 0.9 & 0.9 \end{bmatrix} \tag{4.12.55}$$

现在用节 4.12.1 的方法求 Wiener 预报器 $\hat{\boldsymbol{y}}(t+2|t) = [\hat{y}_1(t+2|t), \hat{y}_2(t+2|t)]^{\mathrm{T}}$. 可求得 ARMA 新息模型为

$$(\boldsymbol{I}_2 + \boldsymbol{A}_1 q^{-1})\boldsymbol{z}(t) = (\boldsymbol{I}_2 + \boldsymbol{D}_1 q^{-1})\boldsymbol{\varepsilon}(t) \tag{4.12.56}$$

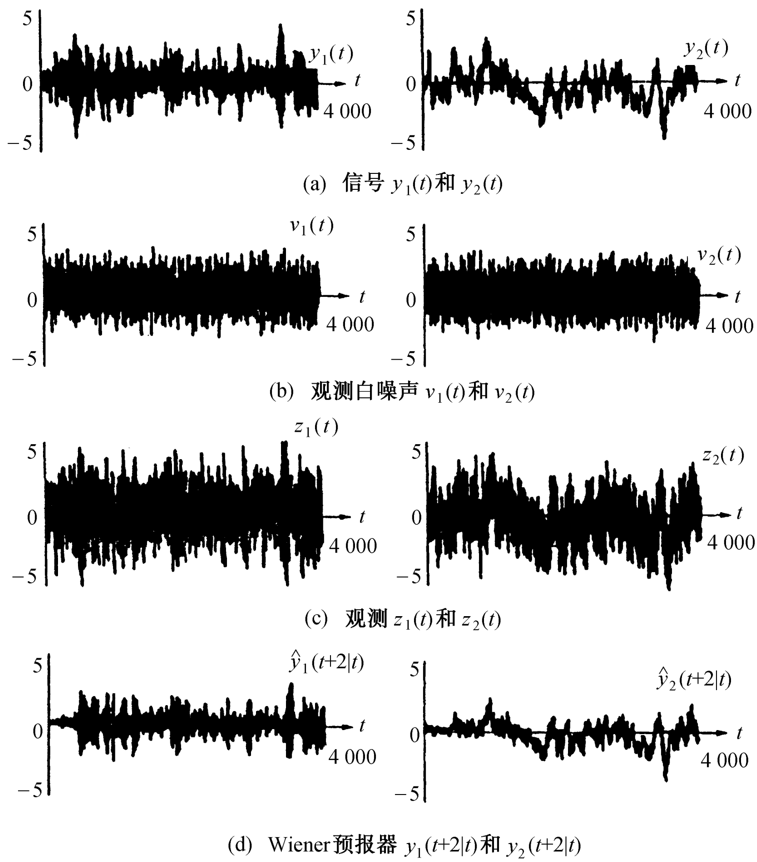
其中新息 $\boldsymbol{\varepsilon}(t) \in R^2$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，用 Gevers – Wouters 算法可求得

$$\boldsymbol{D}_1 = \begin{bmatrix} 0.764\,723 & 0.332\,451 \\ 0.352\,366 & -0.776\,113 \end{bmatrix}, \quad \boldsymbol{Q}_\varepsilon = \begin{bmatrix} 1.770\,021 & 0.015\,816 \\ 0.015\,816 & 1.166\,805 \end{bmatrix} \tag{4.12.57}$$

应用公式(4.12.20)可得 Wiener 预报器

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{y}}(t+2|t) = \boldsymbol{G}_2(q^{-1})\,\mathrm{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{z}(t) \tag{4.12.58}$$

其中 $\boldsymbol{D}(q^{-1}) = \boldsymbol{I}_2 + \boldsymbol{D}_1 q^{-1}$，$\widetilde{\boldsymbol{D}}(q^{-1})$ 和 $\boldsymbol{G}_2(q^{-1})$ 由(4.12.10)和(4.12.11)计算. 仿真结果如图 4.12.1 所示. 由图(a)和图(c)看到，真实信号 $y_1(t)$ 和 $y_2(t)$ 已分别被带白噪声的观测信号 $z_1(t)$ 和 $z_2(t)$ 所淹没. 由图(a)和图(d)看到，Wiener 预报器 $\hat{\boldsymbol{y}}_1(t+2|t)$ 和 $\hat{\boldsymbol{y}}_2(t+2|t)$ 已基本上还原始真实信号 $y_1(t)$ 和 $y_2(t)$ 的本来面目，已有效地减小了噪声 $v_1(t)$ 和 $v_2(t)$ 的影响.



(a) 信号 $y_1(t)$ 和 $y_2(t)$

(b) 观测白噪声 $v_1(t)$ 和 $v_2(t)$

(c) 观测 $z_1(t)$ 和 $z_2(t)$

(d) Wiener 预报器 $y_1(t+2|t)$ 和 $y_2(t+2|t)$

t/步

图 4.12.1　信号 $\boldsymbol{y}(t)$，观测 $\boldsymbol{z}(t)$，观测噪声 $\boldsymbol{v}(t)$ 和 Wiener 预报器 $\hat{\boldsymbol{y}}(t+2|t)$

4.13 带有色观测噪声的多变量 ARMA 过程的 Wiener 预报器

考虑带白色和有色观测噪声的多变量 ARMA 过程

$$\boldsymbol{z}(t)=\boldsymbol{y}(t)+\boldsymbol{\eta}(t)+\boldsymbol{\xi}(t) \tag{4.13.1}$$

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{4.13.2}$$

$$\boldsymbol{P}(q^{-1})\boldsymbol{\eta}(t)=\boldsymbol{R}(q^{-1})\boldsymbol{n}(t) \tag{4.13.3}$$

其中 $\boldsymbol{y}(t)\in R^m$ 为待估信号，$\boldsymbol{z}(t)\in R^m$ 为观测信号，有色观测噪声 $\boldsymbol{\eta}(t)\in R^m$ 服从 ARMA 模型(4.13.3)，$\boldsymbol{\xi}(t)\in R^m$ 为白色观测噪声. 设 $\boldsymbol{\xi}(t)$，$\boldsymbol{e}(t)\in R^r$ 和 $\boldsymbol{n}(t)\in R^s$ 是零均值、方差阵各为 $\boldsymbol{Q}_\xi$，$\boldsymbol{Q}_e$ 和 $\boldsymbol{Q}_n$ 的相互独立白噪声，q^{-1}为单位滞后算子，$\boldsymbol{A}(q^{-1})$，…，$\boldsymbol{R}(q^{-1})$是形如 $\boldsymbol{X}(q^{-1})=\boldsymbol{X}_0+\boldsymbol{X}_1q^{-1}+\cdots+\boldsymbol{X}_{n_x}q^{-n_x}$的多项式矩阵. 设 $\boldsymbol{A}_0=\boldsymbol{I}_m$，$\boldsymbol{P}_0=\boldsymbol{I}_m$，$\boldsymbol{C}_0=\boldsymbol{0}$ 或 $\boldsymbol{C}_0=\boldsymbol{I}_m$，$\boldsymbol{R}_0=0$ 或$_0=\boldsymbol{I}_m$. 问题是基于观测 $\boldsymbol{z}(t)$，$\boldsymbol{z}(t-1)$，…，求 $\boldsymbol{y}(t+k)$的 Wiener 预报器 $\hat{\boldsymbol{y}}(t+k\mid t)$，$k>0$.

本节用 Kalman 滤波方法解决上述问题，适用于非平稳有色观测噪声和非平稳 ARMA 信号，即 $\boldsymbol{A}(q^{-1})$和 $\boldsymbol{P}(q^{-1})$可以是不稳定的多项式矩阵.

应用定理 1.6.1，(4.13.2)和(4.13.3)有状态空间模型

$$\boldsymbol{\alpha}(t+1)=\bar{\boldsymbol{A}}\boldsymbol{\alpha}(t)+\bar{\boldsymbol{C}}\boldsymbol{e}(t) \tag{4.13.4}$$

$$\boldsymbol{y}(t)=\bar{\boldsymbol{H}}_1\boldsymbol{\alpha}(t)+\boldsymbol{C}_0\boldsymbol{e}(t) \tag{4.13.5}$$

$$\boldsymbol{\beta}(t+1)=\bar{\boldsymbol{P}}\boldsymbol{\beta}(t)+\bar{\boldsymbol{R}}\boldsymbol{n}(t) \tag{4.13.6}$$

$$\boldsymbol{\eta}(t)=\bar{\boldsymbol{H}}_2\boldsymbol{\beta}(t)+\boldsymbol{R}_0\boldsymbol{n}(t) \tag{4.13.7}$$

其中规定 $\boldsymbol{C}_i=\boldsymbol{0}(i>n_c)$，$\boldsymbol{R}_i=\boldsymbol{0}(i>n_r)$，$\boldsymbol{P}_i=\boldsymbol{0}(i>n_p)$，$\boldsymbol{A}_i=\boldsymbol{0}(i>n_a)$，$n_1=\max(n_a,n_c)$，$n_2=\max(n_p,n_r)$，

$$\bar{\boldsymbol{A}}=\begin{bmatrix}-\boldsymbol{A}_1 & & & \\ \vdots & \boldsymbol{I}_{(n_1-1)m} & & \\ -\boldsymbol{A}_{n_1} & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix},\quad \bar{\boldsymbol{C}}=\begin{bmatrix}\boldsymbol{C}_1-\boldsymbol{A}_1\boldsymbol{C}_0\\ \vdots\\ \boldsymbol{C}_{n_1}-\boldsymbol{A}_{n_1}\boldsymbol{C}_0\end{bmatrix},\quad \bar{\boldsymbol{H}}_1=[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}] \tag{4.13.8}$$

$$\bar{\boldsymbol{P}}=\begin{bmatrix}-\boldsymbol{P}_1 & & & \\ \vdots & \boldsymbol{I}_{(n_2-1)m} & & \\ -\boldsymbol{P}_{n_2} & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix},\quad \bar{\boldsymbol{R}}=\begin{bmatrix}\boldsymbol{R}_1-\boldsymbol{P}_1\boldsymbol{R}_0\\ \vdots\\ \boldsymbol{R}_{n_2}-\boldsymbol{P}_{n_2}\boldsymbol{R}_0\end{bmatrix},\quad \bar{\boldsymbol{H}}_2=[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}] \tag{4.13.9}$$

合并(4.13.4)~(4.13.7)有增广系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}\boldsymbol{x}(t)+\boldsymbol{\Gamma}\boldsymbol{w}(t) \tag{4.13.10}$$

$$\boldsymbol{z}(t)=\boldsymbol{H}\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{4.13.11}$$

其中定义

$$\boldsymbol{x}(t)=\begin{bmatrix}\boldsymbol{\alpha}(t)\\ \boldsymbol{\beta}(t)\end{bmatrix},\quad \boldsymbol{w}(t)=\begin{bmatrix}\boldsymbol{e}(t)\\ \boldsymbol{n}(t)\end{bmatrix},\quad \boldsymbol{v}(t)=\boldsymbol{C}_0\boldsymbol{e}(t)+\boldsymbol{R}_0\boldsymbol{n}(t)+\boldsymbol{\xi}(t) \tag{4.13.12}$$

$$\boldsymbol{\Phi}=\begin{bmatrix}\bar{\boldsymbol{A}} & \boldsymbol{0}\\ \boldsymbol{0} & \bar{\boldsymbol{P}}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\bar{\boldsymbol{C}} & \boldsymbol{0}\\ \boldsymbol{0} & \bar{\boldsymbol{R}}\end{bmatrix},\quad \boldsymbol{H}=[\bar{\boldsymbol{H}}_1,\bar{\boldsymbol{H}}_2] \tag{4.13.13}$$

显然增广系统是带相关噪声系统，白噪声 $\boldsymbol{w}(t)$和 $\boldsymbol{v}(t)$的方差阵 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$ 及相关阵 $\boldsymbol{S}$ 各为

$$\boldsymbol{Q}_w=\begin{bmatrix}\boldsymbol{Q}_e & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{Q}_n\end{bmatrix},\quad \boldsymbol{Q}_v=\boldsymbol{C}_0\boldsymbol{Q}_e\boldsymbol{C}_0^{\mathrm{T}}+\boldsymbol{R}_0\boldsymbol{Q}_n\boldsymbol{R}_0^{\mathrm{T}}+\boldsymbol{Q}_\xi,$$

$$\boldsymbol{S}=\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(t)]=\begin{bmatrix}\boldsymbol{Q}_e\boldsymbol{C}_0^{\mathrm{T}}\\ \boldsymbol{Q}_n\boldsymbol{R}_0^{\mathrm{T}}\end{bmatrix} \tag{4.13.14}$$

它可化为转移阵为 $\bar{\boldsymbol{\Phi}}$ 的带不相关噪声系统[14]，且

$$\bar{\boldsymbol{\Phi}}=\boldsymbol{\Phi}-\boldsymbol{J}\boldsymbol{H},\quad \boldsymbol{J}=\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_v^{-1} \tag{4.13.15}$$

假设 $(\bar{\boldsymbol{\Phi}},\boldsymbol{H})$ 为完全可观对，且设增广系统是完全能稳的，则增广系统存在稳态 Kalman 预报器[16]

$$\hat{\boldsymbol{x}}(t+1|t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{K}_p\boldsymbol{z}(t) \tag{4.13.16}$$

或在 Wiener 预报器形式下有

$$\hat{\boldsymbol{x}}(t+1|t)=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{z}(t) \tag{4.13.17}$$

其中 $n=n_1+n_2$ 为增广状态维数.可证明[15,16] $\boldsymbol{\Psi}_p$ 是一个稳定矩阵，且有

$$\boldsymbol{\Psi}_p=\bar{\boldsymbol{\Phi}}-\bar{\boldsymbol{K}}_p\boldsymbol{H} \tag{4.13.18}$$

$$\boldsymbol{K}_p=\bar{\boldsymbol{K}}_p+\boldsymbol{J} \tag{4.13.19}$$

$$\bar{\boldsymbol{K}}_p=\bar{\boldsymbol{\Phi}}\boldsymbol{K} \tag{4.13.20}$$

$$\boldsymbol{K}=\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v]^{-1} \tag{4.13.21}$$

其中稳态 Kalman 预报误差方差阵 $\boldsymbol{\Sigma}$ 满足稳态 Riccati 方程

$$\boldsymbol{\Sigma}=\bar{\boldsymbol{\Phi}}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}]\bar{\boldsymbol{\Phi}}^{\mathrm{T}}+\boldsymbol{\Gamma}(\boldsymbol{Q}_w-\boldsymbol{S}\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}} \tag{4.13.22}$$

它可用迭代法求解.

由(4.13.10)和射影性质有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+k|t)=\boldsymbol{\Phi}^{k-1}\hat{\boldsymbol{x}}(t+1|t) \tag{4.13.23}$$

且有预报误差 $\tilde{\boldsymbol{x}}(t+k|t)=\boldsymbol{x}(t+k)-\hat{\boldsymbol{x}}(t+k|t)$.方差阵 $\boldsymbol{P}_k=\mathrm{E}[\tilde{\boldsymbol{x}}(t+k|t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+k|t)]$为

$$\boldsymbol{P}_k=\boldsymbol{\Phi}^{k-1}\boldsymbol{\Sigma}(\boldsymbol{\Phi}^{k-1})^{\mathrm{T}}+\sum_{j=0}^{k-2}\boldsymbol{\Phi}^j\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Phi}^{j\mathrm{T}},\quad k>1 \tag{4.13.24}$$

由(4.13.5)和定义(4.13.12)有

$$\boldsymbol{y}(t)=\boldsymbol{H}_0\boldsymbol{x}(t)+\boldsymbol{C}_0\boldsymbol{e}(t),\quad \boldsymbol{H}_0=[\boldsymbol{I}_m\ \boldsymbol{0}\cdots\boldsymbol{0}] \tag{4.13.25}$$

这引出稳态 k 步预报器

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{H}_0\hat{\boldsymbol{x}}(t+k|t) \tag{4.13.26}$$

其中应用了事实 $\boldsymbol{e}(t+k)\perp L(\boldsymbol{z}^{\mathrm{T}}(t),\boldsymbol{z}^{\mathrm{T}}(t-1),\cdots)$, $k>0$.将(4.13.17)和(4.13.23)代入上式有

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{H}_0\boldsymbol{\Phi}^{k-1}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p\boldsymbol{z}(t) \tag{4.13.27}$$

这引出 Wiener 预报器

$$\psi(q^{-1})\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{G}_k(q^{-1})\boldsymbol{z}(t) \tag{4.13.28}$$

$$\psi(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p) \tag{4.13.29}$$

$$\boldsymbol{G}_k(q^{-1})=\boldsymbol{H}_0\boldsymbol{\Phi}^{k-1}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p \tag{4.13.30}$$

因 $\boldsymbol{\Psi}_p$ 为稳定矩阵，$\psi(q^{-1})$是一个稳定多项式，因而(4.13.28)是渐近稳定的.

由(4.13.25)和(4.13.26)引出预报误差 $\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{y}(t+k)-\tilde{\boldsymbol{y}}(t+k|t)$有关系

$$\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{H}_0\tilde{\boldsymbol{x}}(t+k|t)+\boldsymbol{C}_0\boldsymbol{e}(t+k) \tag{4.13.31}$$

易知 $\boldsymbol{e}(t+k)\perp\tilde{\boldsymbol{x}}(t+k|t)$，$k>0$，这引出预报误差方差阵 $\boldsymbol{P}_k^y=\mathrm{E}[\tilde{\boldsymbol{y}}(t+k|t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+k|t)]$为

$$\boldsymbol{P}_k^y=\boldsymbol{H}_0\boldsymbol{P}_k\boldsymbol{H}_0^{\mathrm{T}}+\boldsymbol{C}_0\boldsymbol{Q}_e\boldsymbol{C}_0^{\mathrm{T}} \tag{4.13.32}$$

上述结果可概括为如下定理.

【定理 4.13.1】 带平稳或非平稳有色观测噪声的平稳或非平稳多变量 ARMA 过程有渐近稳定的 k 步 Wiener 预报器(4.13.28)，且有预报误差方差阵(4.13.32).

【定理 4.13.2】 在定理 4.13.1 条件下，记 $y(t)$的分量表示为

$$\boldsymbol{y}(t)=\begin{bmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{bmatrix} \tag{4.13.33}$$

定义 $m\times m$ 多项式矩阵 $\boldsymbol{G}_k(q^{-1})$的分块表示为

$$\boldsymbol{G}_k(q^{-1})=\begin{bmatrix} \boldsymbol{G}_{k1}(q^{-1}) \\ \vdots \\ \boldsymbol{G}_{km}(q^{-1}) \end{bmatrix} \tag{4.13.34}$$

其中 $\boldsymbol{G}_{ki}(q^{-1})$为 $1\times m$ 向量，则有渐近稳定的解耦 Wiener 预报器

$$\psi(q^{-1})\hat{y}_i(t+k|t)=\boldsymbol{G}_{ki}(q^{-1})\boldsymbol{z}(t),\quad i=1,\cdots,m \tag{4.13.35}$$

【注 4.13.1】 该定理在应用中可减小计算负担，便于实时应用，因为在某些应用问题中我们只感兴趣 $\boldsymbol{y}(t)$的某个分量或某几个分量的预报，而不必计算其他分量的预报.

4.14 指数平滑预报器

周知，Box – Jenkins 预报器，Åström 预报器和 Kalman 预报器均要求已知时间序列的模型.若我们事先不知道时间序列的模型，如何利用它的目前和以往的观测值预报下一时刻(超前一步)时间序列的值？指数平滑预报方法的基本思想是利用对时间序列目前和以往的观测值的指数加权和来预报它的将来的值，故称为“指数平滑”预报方法.

4.14.1 指数平滑预报器的推导[10,11]

为理论推导方便，设初始观测时刻 $t_0\to-\infty$，即已知到时刻 t 为止的无限个观测数据$(y(t),y(t-1),\cdots)$，问题是基于它们的加权和求一步预报器 $\hat{y}(t+1|t)$，即 $\hat{y}(t+1|t)$可表为

$$\hat{y}(t+1\mid t)=\sum_{j=0}^{\infty}c_j y(t-j) \tag{4.14.1}$$

其中 c_j 为加权系数序列.根据渐消记忆方法，应加大新近观测数据的权系数，减小陈旧数据的权系数，以体现过程的时变性.例如对于短期预报商品销售而言，几年前的销量数据对下一个月销售预报提供的信息甚微，而近几个月的销售数据对下一个月销售的预报关系很大.因此应采用指数递降加权系数来满足上述要求，即

$$c_j = c_0\beta^j,\quad 0<\beta<1,\quad c_0>0,\quad \sum_{j=0}^{\infty}c_j = 1 \tag{4.14.2}$$

由几何级数求和公式和(4.14.2)有

$$c_0\sum_{j=0}^{\infty}\beta^j = \frac{c_0}{1-\beta} = 1,\quad c_0 = 1-\beta \tag{4.14.3}$$

因此指数平滑加权系数 c_j 为

$$c_j = (1-\beta)\beta^j,\quad j=0,1,2,\cdots \tag{4.14.4}$$

注意预报器(4.14.1)是目前和过去观测数据的加权平均,因而它具有"平滑"作用,故叫指数平滑预报器. β 叫遗忘因子, β 越小,则越强调新的数据作用,例如取 $\beta=0.1$,则 c_j 成为 0.9,0.09,0.009,…. 在极端情形下,取 $\beta=0$,则以往数据对下一步预报没有任何影响,此时

$$\hat{y}(t+1|t) = y(t) \tag{4.14.5}$$

这种预报器的一个典型例子是股票价格的预报. 预报器(4.14.5)是说,对明天股票价格的最好预报估值就是今天股票的价格.

上述推导可概括为如下定理.

【定理 4.14.1】 已知无限观测 $(y(t),y(t-1),\cdots)$ 时, $y(t+1)$ 的指数平滑预报器 $\hat{y}(t+1|t)$ 为

$$\hat{y}(t+1|t) = \sum_{j=0}^{\infty}(1-\beta)\beta^j y(t-j) \tag{4.14.6}$$

其中 $0<\beta<1$.

理论上,预报器(4.14.6)是无限个观测数据的加权和,然而在应用中仅有有限个观测数据是可利用的. 而且非递推预报器(4.14.6)在应用上不是方便的,它要求存贮全部过去的观测数据. 因此应推导等价的递推预报公式.

【定理 4.14.2】 递推指数平滑预报公式为

$$\hat{y}(t+1|t) = (1-\beta)y(t) + \beta\hat{y}(t|t-1) \tag{4.14.7}$$

或

$$\hat{y}(t+1|t) = \hat{y}(t|t-1) + (1-\beta)[y(t)-\hat{y}(t|t-1)] \tag{4.14.8}$$

带任意的初值 $\hat{y}(1|0)$,且 $\hat{y}(t+1|t)$ 关于初值 $\hat{y}(1|0)$ 是渐近稳定的,即 $\hat{y}(t+1|t)$ 的值渐近地与初值 $\hat{y}(1|0)$ 选取无关.

证明 由(4.14.6)有

$$\hat{y}(t+1|t) = (1-\beta)y(t) + \sum_{j=1}^{\infty}(1-\beta)\beta^j y(t-j) =$$

$$(1-\beta)y(t) + \beta\sum_{j=1}^{\infty}(1-\beta)\beta^{j-1}y(t-j) \tag{4.14.9}$$

令 $i=j-1$ 并应用(4.14.6)得(4.14.7). 任取(4.14.7)两个初值 $\hat{y}^{(i)}(1|0)$,设相应的指数平滑预报器为 $\hat{y}^{(i)}(t+1|t)$,令 $\delta(t) = \hat{y}^{(1)}(t+1|t) - \hat{y}^{(2)}(t+1|t)$,由(4.14.7)则有

$$(1-\beta q^{-1})\delta(t) = 0 \tag{4.14.10}$$

由(4.14.10)引出

$$\delta(t) = \beta^t\delta(0) \tag{4.14.11}$$

其中 $\delta(0)=\hat{y}^{(1)}(1|0)-\hat{y}^{(2)}(1|0)$. 因为 $0<\beta<1$，故有

$$\delta(t)\to 0,\quad t\to\infty \tag{4.14.12}$$

4.14.2 指数平滑预报器与 ARMA 过程预报器的关系

【定理 4.14.3】 递推指数平滑预报器(4.14.7)等价于如下非平稳 ARMA(1,1)过程

$$(1-q^{-1})y(t)=(1-\beta q^{-1})\varepsilon(t) \tag{4.14.13}$$

的一步最优 Box – Jenkins 预报器

$$\hat{y}(t+1|t)=y(t)-\beta\varepsilon(t) \tag{4.14.14}$$

$$\varepsilon(t)=y(t)-\hat{y}(t|t-1) \tag{4.14.15}$$

证明 由(4.14.7)有

$$\hat{y}(t+1|t)=y(t)-\beta[y(t)-\hat{y}(t|t-1)] \tag{4.14.16}$$

它引出(4.14.14)和(4.14.15). 证毕. □

定理【4.14.4】 递推指数平滑预报器(4.14.7)等价于非平稳 ARMA 过程(4.14.13) Åström 一步最优预报器

$$\hat{\boldsymbol{y}}(t+1|t)=\frac{1-\beta}{1-\beta q^{-1}}y(t) \tag{4.14.17}$$

证明 由(4.14.13)和(4.14.14)有 Åström 预报器

$$\hat{y}(t+1|t)=y(t)-\frac{\beta(1-q^{-1})}{1-\beta q^{-1}}y(t)=\frac{1-\beta}{1-\beta q^{-1}}y(t) \tag{4.14.18}$$

即(4.14.17)成立. 由(4.14.17)有

$$(1-\beta q^{-1})\hat{y}(t+1|t)=(1-\beta)y(t) \tag{4.14.19}$$

它等价于(4.14.7). 证毕. □

4.14.3 指数平滑预报器与 Kalman 预报器的关系

考虑随机系统

$$x(t+1)=x(t)+w(t) \tag{4.14.20}$$

$$y(t)=x(t)+v(t) \tag{4.14.21}$$

其中 $w(t)$和 $v(t)$是零均值、方差各为 σ_w^2和σ_v^2的独立白噪声，$x(t)$为状态，$y(t)$为观测.

【定理 4.14.5】 指数平滑预报器(4.14.8)等价于系统(4.14.20)和(4.14.21)的稳态 Kalman 预报器 $\hat{x}(t+1|t)$，即

$$\hat{y}(t+1|t)=\hat{x}(t+1|t) \tag{4.14.22}$$

$$\hat{x}(t+1|t)=\hat{x}(t|t-1)+K_p\varepsilon(t) \tag{4.14.23}$$

$$K_p=1-\beta \tag{4.14.24}$$

$$\varepsilon(t)=y(t)-\hat{y}(t|t-1)=y(t)-\hat{x}(t|t-1) \tag{4.14.25}$$

而遗忘因子 $0<\beta<1$ 是如下二次方程

$$\beta^2-\delta\beta+1=0 \tag{4.14.26}$$

的根，其中 $\delta=(\sigma_w^2+2\sigma_v^2)/\sigma_v^2=2+(\sigma_w^2/\sigma_v^2)$.

证明 将(4.14.21)代入(4.14.20)有

$$(1-q^{-1})y(t)=w(t-1)+(1-q^{-1})v(t) \tag{4.14.27}$$

引入等价的滑动平均过程 $(1-\beta q^{-1})\varepsilon(t)$,

$$(1-\beta q^{-1})\varepsilon(t)=w(t-1)+(1-q^{-1})v(t) \tag{4.14.28}$$

其中 $\varepsilon(t)$ 是零均值、方差为 σ_ε^2 的白噪声.计算上式两边随机过程的相关函数引出

$$(1+\beta^2)\sigma_\varepsilon^2=\sigma_w^2+2\sigma_v^2 \tag{4.14.29}$$

$$\beta\sigma_\varepsilon^2=\sigma_v^2 \tag{4.14.30}$$

上两式相除引出(4.14.26).容易证明(4.14.26)有相异实根,且有一根大于1.又由(4.14.26)知两根之积等于1,故另一根为 $0<\beta<1$.由(4.14.27)和(4.14.28)有 ARMA 新息模型

$$(1-q^{-1})y(t)=(1-\beta q^{-1})\varepsilon(t) \tag{4.14.31}$$

其中 $\varepsilon(t)=y(t)-\hat{y}(t|t-1)$.

因系统(4.14.20)和(4.14.21)是完全可观、完全可控的,故存在稳态 Kalman 预报器(4.14.23)和(4.14.25).将(4.14.23)代入(4.14.25)有

$$y(t)=\frac{K_p}{1-q^{-1}}\varepsilon(t-1)+\varepsilon(t) \tag{4.14.32}$$

这引出 ARMA 新息模型

$$(1-q^{-1})y(t)=K_p\varepsilon(t-1)+(1-q^{-1})\varepsilon(t) \tag{4.14.33}$$

比较(4.14.31)和(4.14.33)得

$$K_p-1=-\beta \quad 或 \quad K_p-1-\beta \tag{4.14.34}$$

由(4.14.21)和射影性质有

$$\hat{y}(t+1|t)=\hat{x}(t+1|t) \tag{4.14.35}$$

将它代入(4.14.23)得(4.14.8).证毕. □

4.14.4 指数平滑递推预报器与信号检测滤波器关系

在炼油、化工、冶金等许多生产过程控制系统中,由于各种随机因素的影响使工艺参数的检测信号被噪声污染.工艺参数 $x(t)$ 本身通常是低频信号,而对 $x(t)$ 的检测信号 $y(t)=x(t)+v(t)$ 中含有高频干扰信号 $v(t)$.为了更好地控制生产过程,需要由信号 $y(t)$ 得到真实信号 $x(t)$ 的最优估值 $\hat{x}(t|t)$.工艺参数 $x(t)$ 通常是慢时变的,因而近似用随机游动模型(4.14.20)来描写其变化,而(4.14.21)是观测方程,其中假设 $w(t)$ 和 $v(t)$ 是零均值、方差各为 σ_w^2 和 σ_v^2 的独立白噪声.

基于检测(观测)$(y(t),y(t-1),\cdots)$ 信号 $x(t)$ 的稳态 Kalman 滤波器为

$$\hat{x}(t|t)=\hat{x}(t|t-1)+K_f\varepsilon(t) \tag{4.14.36}$$

其中 K_f 为稳态滤波增益,易知有

$$K_f=K_p=1-\beta \tag{4.14.37}$$

其中 $\varepsilon(t)$,K_p 和 β 由定理 4.14.5 给出.注意由(4.14.20)有

$$\hat{x}(t|t-1)=\hat{x}(t-1|t-1) \tag{4.14.38}$$

于是有信号 $x(t)$ 的稳态 Kalman 滤波器

$$\hat{x}(t|t)=\hat{x}(t-1|t-1)+(1-\beta)[y(t)-\hat{x}(t-1|t-1)] \tag{4.14.39}$$

它具有指数平滑预报器(4.14.8)的形式

$$\hat{x}(t|t)=\beta\hat{x}(t-1|t-1)+(1-\beta)y(t) \tag{4.14.40}$$

带任意初值 $\hat{x}(0|0)$. 它是前一时刻信号估值与目前时刻信号观测的加权，且它具有一队滤波器形式

$$\hat{x}(t|t)=\frac{1-\beta}{1-\beta q^{-1}}y(t),\quad 0<\beta<1 \tag{4.14.41}$$

【定理 4.14.6】 服从随机游动模型(4.14.20)和(4.14.21)信号 $x(t)$，由被高频噪声 $v(t)$污染的信号 $y(t)$所得到的最优信号检测数字滤波器 $\hat{x}(t|t)$具有指数平滑预报器形式(4.14.40)或(4.14.41). 它关于初值 $\hat{x}(0|0)$是渐近稳定的.

4.14.5 仿真例子

设信号 $x(t)$和检测信号 $y(t)$服从随机游动模型(4.14.20)和(4.14.21)，其中 $w(t)$和 $v(t)$是零均值、方差各为 $\sigma_w^2=0.09$ 和 $\sigma_v^2=4$ 的独立高斯白噪声. 它的 ARMA 新息模型为

$$(1-q^{-1})y(t)=(1-\beta q^{-1})\varepsilon(t) \tag{4.14.42}$$

可求得真实值 $\beta=0.860\ 82$，于是它的 Box－Jenkins 一步预报器

$$\hat{y}(t+1|t)=y(t)-\beta\varepsilon(t) \tag{4.14.43}$$

$$\varepsilon(t)=y(t)-\hat{y}(t|t-1) \tag{4.14.44}$$

即最优指数平滑预报器，且有 $\hat{x}(t+1|t)=\hat{y}(t+1|t)$. 也可用(4.14.40)形式的指数平滑预报器

$$\hat{x}(t+1|t)=\beta\hat{x}(t|t-1)-(1-\beta)y(t) \tag{4.14.45}$$

仿真结果如图 4.14.1 和图 4.14.2 所示. 图 4.14.1 为信号，它是真实信号 $x(t)$与观测噪声 $v(t)$的叠加. 图 4.14.2 为真实信号 $x(t)$与指数平滑预报器 $\hat{x}(t+1|t)$，它们几乎重合，可看到指数平滑预报器 $\hat{x}(t+1|t)$能还真实信号 $x(t)$的本来面目，有效地过滤了噪声 $v(t)$.

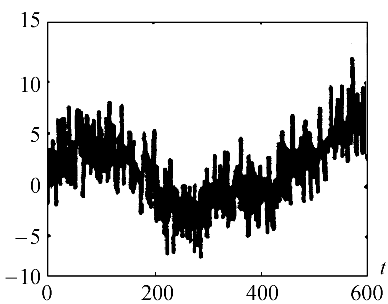


图 4.14.1 检测信号 $y(t)$

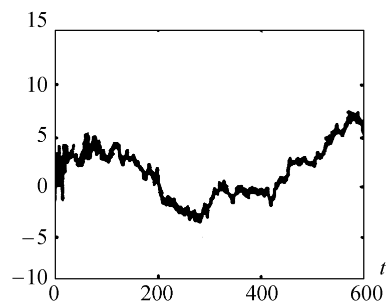


图 4.14.2 信号 $x(t)$和指数平滑预报器 $x(t+1|t)$

4.15 非平稳 ARMA 过程的 Box－Jenkins 递推预报器和 Åström 预报器

在节 4.5 我们用 Kalman 滤波方法由稳态 Kalman 预报器导出了 Åström 预报器，但 $G_k(q^{-1})$的计算公式有所不同. 本节我们用有限与无限的转化思想，通过对初始观测时刻

t_0 为有限的非平稳 ARMA 过程的线性最小方差预报器当 $t_0 \to -\infty$ 或 $t \to +\infty$ 取极限的方法得到稳态最优 Box – Jenkins 预报器和 Åström 预报器.实现了有限与无限的转化和稳态与非稳态的转化.但避免了 Kalman 滤波方法.

4.15.1 非平稳 ARMA 过程的 Box – Jenkins 预报器

【例 4.15.1】 一个启发性例子.

考虑非平稳可逆的 ARMA(1,1)过程 $y(t)$,

$$(1-q^{-1})y(t)=(1-cq^{-1})e(t),\quad |c|<1 \tag{4.15.1}$$

其中 $A(q^{-1})=1-q^{-1}$是不稳定的,它有在单位圆上的零点 $q^{-1}=1$,且设 $e(t)$是零均值、方差为 σ_e^2的白噪声.问题是当初始观测时刻 $t_0 \to -\infty$ 时,求它的稳态最优预报器 $\hat{y}(t+k|t)$.

先暂时设 t_0 是有限的,且设

$$\mathrm{E}[y(t_0)]=0,\quad \mathrm{E}[y^2(t_0)]<\infty \tag{4.15.2}$$

其中 E 为数学期望,则由(4.15.1)可递推引出

$$\mathrm{E}[y(t)]=0,\quad t>t_0 \tag{4.15.3}$$

但 $y(t)$的方差却是无界的.事实上,

$$\begin{aligned}\mathrm{E}[y^2(t)]&=\mathrm{E}[y(t-1)+e(t)-ce(t-1)][y(t-1)+e(t)-ce(t-1)]=\\ &\mathrm{E}[y^2(t-1)]+(1+c^2)\sigma_e^2-2\sigma_e^2=\mathrm{E}[y^2(t-1)]+(1-c)^2\sigma_e^2\end{aligned} \tag{4.15.4}$$

其中利用了事实

$$\begin{aligned}&\mathrm{E}[y(t-1)e(t-1)]=\mathrm{E}[(y(t-2)+e(t-1)-ce(t-2))e(t-1)]=\sigma_e^2,\\ &\mathrm{E}[y(t-1)e(t)]=\mathrm{E}[(y(t-2)+e(t-1)-ce(t-2))e(t)]=0\end{aligned} \tag{4.15.5}$$

于是我们有

$$\mathrm{E}[y^2(t)]=\mathrm{E}[y^2(t_0)]+(t-t_0)(1-c)^2\sigma_e^2 \tag{4.15.6}$$

因此当 $t_0 \to -\infty$ 时有

$$\mathrm{E}[y^2(t)] \to +\infty \tag{4.15.7}$$

即 $y(t)$有无穷大方差.因此我们不能直接应用节 4.1 无穷级 Hilbert 空间射影理论来解决当 $t_0=-\infty$ 时 $y(t+k)$的稳态最优预报问题.因为 Hilbert 空间理论要求随机变量 $y(t)$的方差是有限的.但我们可用有限维 Hilbert 空间取极限的方法得到稳态最优预报器.

在(4.15.1)中任取常数初值$(y(t_0-1),e(t_0-1))$,则当 $t \geqslant t_0$ 时,$y(t)$与 $e(t)$可相互线性表示,因而由$(y(t),\cdots,y(t_0))$张成的线性流形(有限维 Hilbert 空间)$L(y(t),\cdots,y(t_0))$相同于由$(e(t),\cdots,e(t_0))$张成的线性流形(有限维 Hilbert 空间)$L(y(t),\cdots,y(t_0))$相同于由$(e(t),\cdots,e(t_0))$张成的线性流形 $L(e(t),\cdots,e(t_0))$,即

$$L(y(t),\cdots,y(t_0))=L(e(t),\cdots,e(t_0)) \tag{4.15.8}$$

在上述线性流形中 $y(t),\cdots,y(t_0)$的方差均为有限的.

在(4.15.1)中取两端各项到线性流形 $L(y(t-1),\cdots,y(t_0))$上的射影可得基于有限个观测$(y(t-1),\cdots,y(t_0))$的最优一步预报器

$$\hat{y}(t|t-1)=y(t-1)-ce(t-1) \tag{4.15.9}$$

由(4.15.1)减上式引出

$$e(t)=y(t)-\hat{y}(t|t-1) \tag{4.15.10}$$

因而 $e(t)$是 $y(t)$的新息过程,但因 $e(t)$与 $y(t)$的计算均与初值$(y(t_0-1),e(t_0-1))$有关,故 $e(t)$是 $y(t)$的非稳态新息过程.显然,由(4.15.9)和(4.15.10)引出非稳态新息过程 $e(t)$也满足(4.15.1).

类似地由(4.15.1)可得非稳态 Box - Jenkins 递推预报器

$$\hat{y}(t+k|t)=\hat{y}(t+k-1|t),\quad k>1 \tag{4.15.11}$$

在(4.15.1)中置新的观测过程

$$z(t)=(1-q^{-1})y(t)=y(t)-y(t-1) \tag{4.15.12}$$

则非平稳过程 $y(t)$(4.15.1)转化为一个平稳 MA(1)过程 $z(t)$,

$$z(t)=(1+cq^{-1})e(t),\quad |c|<1 \tag{4.15.13}$$

因 $z(t)$是平稳 MA(1)过程,故 $z(t)$有界的方差

$$\mathrm{E}[z^2(t)]=(1+c^2)\sigma_e^2,\quad \forall\, t \tag{4.15.14}$$

且 $e(t)$是 $z(t)$的稳态新息过程,即当 $t_0\to-\infty$有均方收敛的级数

$$e(t)=\frac{1}{1-cq^{-1}}z(t)=\sum_{j=0}^{\infty}c^jq^{-j}z(t)=$$

$$\sum_{j=0}^{\infty}c^jz(t-j)=\sum_{j=0}^{\infty}c^j[y(t-j)-y(t-j-1)] \tag{4.15.15}$$

容易证明当 $t_0=-\infty$时,由(4.15.15)定义的 $e(t)$也是 $y(t)$的稳态新息过程.事实上,记由(4.15.15)计算的 $e(t)$为

$$e(t)\triangleq e_\infty(t) \tag{4.15.16}$$

则 $e_\infty(t)$满足(4.15.1),即

$$(1-q^{-1})y(t)=(1-cq^{-1})e_\infty(t) \tag{4.15.17}$$

另一方面带初值$(y(t_0-1),e(t_0-1))$由(4.15.1)计算的非稳态新息 $e(t)$也满足(4.15.1),即

$$(1-q^{-1})y(t)=(1-cq^{-1})e(t) \tag{4.15.18}$$

上两式相减,令 $\delta(t)=e(t)-e_\infty(t)$,则有

$$(1-cq^{-1})\delta(t)=0 \tag{4.15.19}$$

这引出

$$\delta(t)=c^{t-t_0}\delta(t_0) \tag{4.15.20}$$

因$|c|<1$,这引出

$$\delta(t)=e(t)-e_\infty(t)\to 0,\quad t_0\to-\infty \tag{4.15.21}$$

即 $e_\infty(t)$是 $y(t)$的稳态新息过程,

$$\lim_{t_0\to-\infty}e(t)=e_\infty(t) \tag{4.15.22}$$

当 $t_0\to-\infty$时在(4.15.9)和(4.15.11)两边取极限引出稳态最优 Box - Jenkins 递推预报器为

$$\begin{gathered}\hat{y}(t|t-1)=y(t-1)-ce_\infty(t-1),\\ \hat{y}(t+k|t)=\hat{y}(t+k-1|t),\quad k>1\end{gathered} \tag{4.15.23}$$

其中 $e_\infty(t)$由(4.15.15)定义,即

$$e_{\infty}(t) = \sum_{j=0}^{\infty} c^j z(t-j) = \sum_{j=0}^{\infty} c^j [y(t-j) - y(t-1-j)] \tag{4.15.24}$$

若用(4.15.15)和(4.15.16)定义 $e(t)$，则非平稳 ARMA 过程(4.15.1)的稳态 Box – Jenkins 递推预报器为

$$\hat{y}(t|t-1) = y(t-1) - ce(t-1) \tag{4.15.25}$$

$$\hat{y}(t+k|t) = \hat{y}(t+k-1|t), \quad k > 1 \tag{4.15.26}$$

它与平稳 ARMA 过程的 Box – Jenkins 递推预报器具有相同形式.

利用上述原理和方法，读者可容易证明如下一般非平稳 ARMA 过程 Box – Jenkins 递推预报器定理，它与平稳 ARMA 过程的 Box – Jenkins 递推预报器具有相同形式.

【定理 4.15.1】 (非平稳 ARMA 过程的 Box – Jenkins 的递推预报器)考虑非平稳、可逆的 ARMA 过程 $y(t)$

$$A(q^{-1})y(t) = C(q^{-1})e(t) \tag{4.15.27}$$

其中 $e(t)$是零均值、方差为 σ_e^2的白噪声，$A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}$是不稳定的，而 $C(q^{-1}) = 1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c}$是稳定的，设初始观测时刻 $t_0 = -\infty$，则有 Box – Jenkins 稳态最优递推预报器

$$\hat{y}(t+1 \mid t) = -\sum_{i=1}^{n_a} a_i y(t+1-i) + \sum_{i=1}^{n_c} c_i e(t+1-i),$$

$$\hat{y}(t+k \mid t) = -\sum_{i=1}^{n_a} a_i \hat{y}(t+k-i \mid t) + \sum_{i=k}^{n_c} c_i e(t+k-i), \quad 1 \leqslant k \leqslant n_c$$

$$\hat{y}(t+k \mid t) = -\sum_{i=1}^{n_a} a_i \hat{y}(t+k-i \mid t), \quad i > n_c \tag{4.15.28}$$

其中规定 $\hat{y}(t+k-i|t) = y(t+i-i)\ (t+k-i \leqslant t)$，且稳态新息为

$$e(t) = y(t) - \hat{y}(t|t-1) \tag{4.15.29}$$

$$e(t) = \frac{1}{C(q^{-1})} z_t = \sum_{j=0}^{\infty} \pi_j q^{-j} z_t = \sum_{j=0}^{\infty} \pi_j z_{t-j} \tag{4.15.30}$$

$$z(t) = A(q^{-1})y(t) \tag{4.15.31}$$

其中系数 π_j 可递推计算为

$$\pi_j = -c_1 \pi_{j-1} - \cdots - c_{n_c} \pi_{j-n_c}, \quad \pi_0 = 1 \tag{4.15.32}$$

而 $e(t)$可由(4.15.27)任取初值$(e(0), \cdots, e(1-n_c), y(0), \cdots, y(1-n_c))$递推近似计算为

$$e(t) = A(q^{-1})y(t) - c_1 e(t-1) - \cdots - c_{n_c} e(t-n_c), \quad t = 1, 2, \cdots \tag{4.15.33}$$

它是非稳态新息.上述结果与平稳 ARMA 过程的 Box – Jenkins 递推预报器具有相同形式.

注意，定理 4.15.1 也可平行推广到多变量非平稳 ARMA 过程的情形，可得到与多变量平稳 ARMA 过程的 Box – Jenkins 递推预报器具有相同的形式的稳态最优预报器.

【定理 4.15.2】 (多变量非平稳 ARMA 过程的 Box – Jenkins 递推预报器)考虑多变量非平稳 ARMA 过程

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{4.15.34}$$

其中 $\boldsymbol{y}(t) \in R^m$，$\boldsymbol{e}(t) \in R^m$ 是零均值、方差为 $\boldsymbol{Q}_e$ 的白噪声，$\boldsymbol{A}(q^{-1})$是不稳定的，$\boldsymbol{C}(q^{-1})$是稳定的，且 $\boldsymbol{A}(q^{-1}) = \boldsymbol{I}_m + \boldsymbol{A}_1 q^{-1} + \cdots + \boldsymbol{A}_{n_a} q^{-n_a}$，$\boldsymbol{C}(q^{-1}) = \boldsymbol{I}_m + \boldsymbol{C}_1 q^{-1} + \cdots + \boldsymbol{C}_{n_c} q^{-n_c}$，

系数阵 $\boldsymbol{A}_i, \boldsymbol{C}_i$ 为 $m\times m$ 阵，则当 $t_0=-\infty$ 时，有 $y(t+k)$ 的递推 Box－Jenkins 递推预报器

$$\hat{\boldsymbol{y}}(t+k\mid t)=-\sum_{i=1}^{n_a}\boldsymbol{A}_i\hat{\boldsymbol{y}}(t+k-i\mid t)+\sum_{i=k}^{n_c}\boldsymbol{C}_i\hat{\boldsymbol{e}}(t+k-i\mid t) \tag{4.15.35}$$

其中规定

$$\hat{\boldsymbol{y}}(t+k-i|t)=\boldsymbol{y}(t+k-i)(t+k-i\leqslant t),$$

$$\hat{\boldsymbol{e}}(t+k-i|t)=\begin{cases}\boldsymbol{e}(t+k-i) & t+k-i\leqslant t\\ 0 & t+k-i>t\end{cases} \tag{4.15.36}$$

稳态新息 $\boldsymbol{e}(t)$ 可由 $(\boldsymbol{y}(t), \boldsymbol{y}(t-1), \cdots)$ 计算为

$$\boldsymbol{e}(t)=\boldsymbol{C}^{-1}(q^{-1})\boldsymbol{z}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Pi}_j\boldsymbol{z}(t-j),$$

$$\boldsymbol{z}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) \tag{4.15.37}$$

系数阵 $\boldsymbol{\Pi}_j$ 可递推计算为

$$\boldsymbol{\Pi}_j=-\boldsymbol{C}_1\boldsymbol{\Pi}_{j-1}-\cdots-\boldsymbol{C}_{n_c}\boldsymbol{\Pi}_{j-n_c},\quad \boldsymbol{\Pi}_0=\boldsymbol{I}_m, \boldsymbol{\Pi}_j=\boldsymbol{0}(j<0) \tag{4.15.38}$$

在应用中，$\boldsymbol{e}(t)$ 可近似由(4.15.34)递推计算为

$$\boldsymbol{e}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)-\boldsymbol{C}_1\boldsymbol{e}(t-1)-\cdots-\boldsymbol{C}_{n_c}\boldsymbol{e}(t-n_c) \tag{4.15.39}$$

带初值 $(\boldsymbol{y}(0), \cdots, \boldsymbol{y}(1-n_a), \boldsymbol{e}(0), \cdots, \boldsymbol{e}(1-n_c)), t=1,2,\cdots$，它是非稳态新息.

证明　留给读者，从略. □

4.15.2　非平稳 ARMA 过程的 Åström 预报器

我们将证明非平稳 ARMA 过程的稳态 Åström 预报器具有与平稳 ARMA 过程的 Åström 预报器相同的形式.但这里不采用 Kalman 预报方法，而采用从有限维 Hilbert 空间射影出发，当 $t_0\to-\infty$ 时，取极限的方法.

考虑非平稳 ARMA 过程(4.15.27)，置 $t=t+k$ 有

$$y(t+k)=\frac{C(q^{-1})}{A(q^{-1})}e(t+k) \tag{4.15.40}$$

其中 $e(t)$ 是零均值、方差为 σ_e^2 的白噪声. $A(q^{-1})$ 是不稳定的.由多项式综合除法有

$$\frac{C(q^{-1})}{A(q^{-1})}=F(q^{-1})+q^{-k}\frac{G(q^{-1})}{A(q^{-1})} \tag{4.15.41}$$

其中 $F(q^{-1})=f_1+f_1q^{-1}+\cdots+f_{k-1}q^{-(k-1)}$, $G(q^{-1})=g_0+g_1q^{-1}+\cdots+g_{n_g}q^{-n_g}$, $n_g=\max(n_a-1, n_c-k)$，上式等价于 Diophantine 方程

$$C(q^{-1})=A(q^{-1})F(q^{-1})+q^{-k}G(q^{-1}) \tag{4.15.42}$$

暂时设初始观测时刻 t_0 是有限的，对(4.15.27)任取常数初值

$$(y(t_0-1), \cdots, y(t_0-n_a), e(t_0-1), \cdots, e(t_0-n_c)) \tag{4.15.43}$$

则由 $(y(t), y(t-1), \cdots, y(t_0))$ 张成的线性流形 $L(y(t), \cdots, y(t_0))$ 重合于由 $(e(t), \cdots, e(t_0))$ 张成的线性流形 $L(e(t), e(t-1), \cdots, e(t_0))$，即

$$L(y(t), y(t-1), \cdots, y(t_0))=L(e(t), e(t-1), \cdots, e(t_0)) \tag{4.15.44}$$

将(4.15.41)代入(4.15.40)有

$$y(t+k)=F(q^{-1})e(t+k)+\frac{G(q^{-1})}{A(q^{-1})}e(t) \tag{4.15.45}$$

定义一个 ARMA 过程 $\eta(t)$为

$$\eta(t)=\frac{G(q^{-1})}{A(q^{-1})}e(t) \tag{4.15.46}$$

任取常数实值$(\eta(t_0-1),\cdots,\eta(t_0-n_a))$，易知 $\eta(t)\in L(e(t),e(t-1),\cdots,e(t_0))$，取(4.15.41)两边的项在线性流形 $L(e(t),e(t-1),\cdots,e(t))$上射影有非稳态最优预报器 $\hat{y}(t+k|t)$为

$$\hat{y}(t+k|t)=\frac{G(q^{-1})}{A(q^{-1})}e(t) \tag{4.15.47}$$

这是因为它与初值有关.它不仅与初值$(e(t_0-1),\cdots,e(t_0-n_c))$有关，而且还与初值$(\hat{y}(t_0+k-1|t_0-1),\cdots,\hat{y}(t_0-n_c+k|t_0-n_c))$有关.对求稳态预报器而言，初值选取何值是非本质的，当 $t_0\to-\infty$时，任何初值的影响将消失.

现在在(4.15.47)中令 $t_0\to-\infty$，由 $C(q^{-1})$的稳定性(即 ARMA 模型(4.15.27)的可逆性)有稳态 $e(t)$为

$$e_\infty(t)=\frac{1}{C(q^{-1})}z(t)=\sum_{j=0}^{\infty}\pi_j q^{-j}z(t)=\sum_{j=0}^{\infty}\pi_j z(t-j) \tag{4.15.48}$$

其中

$$z(t)=A(q^{-1})y(t) \tag{4.15.49}$$

且系数 π_j 用(4.15.32)计算，则有稳态最优 Åström 预报器为

$$\hat{y}(t+k|t)=\frac{G(q^{-1})}{A(q^{-1})}e_\infty(t) \tag{4.15.50}$$

将(4.15.48)和(4.15.49)代入(4.15.50)消去 $A(q^{-1})$引出稳态最优 Åström 预报器为

$$\hat{y}(t+k|t)=\frac{G(q^{-1})}{C(q^{-1})}y(t) \tag{4.15.51}$$

且由(4.15.47)和(4.15.45)有稳态预报误差为

$$\tilde{y}(t+k|t)=y(t+k)-\hat{y}(t+k|t)=F(q^{-1})e_\infty(t+k) \tag{4.15.52}$$

若记用(4.15.48)定义稳态新息 $e_\infty(t)$为

$$e(t)=\frac{1}{C(q^{-1})}z(t)=\sum_{j=0}^{\infty}\pi_j z(t-j) \tag{4.15.53}$$

则稳态最优预报误差为

$$\tilde{y}(t+k|t)=F(q^{-1})e(t+k) \tag{4.15.54}$$

其中 $e(t+k)$(4.15.53)计算.于是稳态最优预报误差方差为

$$\mathrm{E}[\tilde{y}^2(t+k|t)]=\sigma_e^2\sum_{i=0}^{k-1}f_i^2,\quad f_0=1 \tag{4.15.55}$$

【定理 4.15.3】 (非平稳 ARMA 过程 Åström 预报器)非平稳 ARMA 过程(4.15.27)与平稳 ARMA 过程的 Åström 预报器具有相同的形式，当 $t_0=-\infty$时有

$$\hat{y}(t+k|t)=\frac{G(q^{-1})}{C(q^{-1})}y(t) \tag{4.15.56}$$

其中 $G(q^{-1})$和 $F(q^{-1})$由 Diophantine 方程(4.15.42)决定，且预报误差为

$$\tilde{y}(t+k|t)=F(q^{-1})e(t+k) \tag{4.15.57}$$

其中 $e(t)$由(4.15.53)计算，预报误差方差为(4.15.55).

证明 只需证明 $e_\infty(t)$ 是稳态新息. 因为其他结论上面已推导. 由(4.15.45)和(4.15.47)有非稳态最优预报器预报误差为 $\tilde{y}(t+k|t)=F(q^{-1})e(t+k)$，取 $k=1$，注意 $f_0=1$，因而有

$$\tilde{y}(t+1|t)=y(t+1)-\hat{y}(t+1|t)=e(t+1) \tag{4.15.58}$$

即 $e(t)$ 是非稳态新息. 下证

$$\lim_{t_0\to-\infty} e(t)=e_\infty(t) \tag{4.15.59}$$

因 $e(t)$ 和 $y(t)$ 均为由(4.15.27)取初值后递推计算的，即

$$y(t)=-a_1y(t-1)\cdots-a_{n_a}y(t-n_a)+C(q^{-1})e(t),$$

$$e(t)=A(q^{-1})y(t)-c_1e(t-1)\cdots-c_{n_c}e(t-n_c) \tag{4.15.60}$$

带初值 $(y(t_0-1),\cdots,y(t_0-n_a),e(t_0-1),\cdots,e(t_0-n_c))$，因而有 $e(t)$ 满足(4.15.27)即

$$A(q^{-1})y(t)=C(q^{-1})e(t) \tag{4.15.61}$$

另一方面由(4.15.48)和(4.15.49)有 $e_\infty(t)$ 满足(4.15.27)即

$$C(q^{-1})e_\infty(t)=z(t)=A(q^{-1})y(t) \tag{4.15.62}$$

或

$$A(q^{-1})y(t)=C(q^{-1})e_\infty(t) \tag{4.15.63}$$

引入误差 $\delta(t)=e(t)-e_\infty(t)$，由 4.15.61 减(4.15.63)引出

$$C(q^{-1})\delta(t)=0 \tag{4.15.64}$$

因 $C(q^{-1})$ 为稳定的多项式，则由差分方程理论有

$$\delta(t)\to 0 \quad (t_0\to-\infty) \tag{4.15.65}$$

即我们有

$$\lim_{t_0\to\infty} e(t)=e_\infty(t) \tag{4.15.66}$$

这表明 $e_\infty(t)$ 是非稳态新息 $e(t)$ 的稳态值，即稳态新息. 证毕. □

定理 4.15.3 可平行推广到多变量情形.

【定理 4.15.4】 (多变量非平稳 ARMA 过程的 Åström 预报器)考虑多变量非平稳可逆的 ARMA 过程(4.15.34)，其中 $\boldsymbol{A}(q^{-1})$ 是不稳定的，$\boldsymbol{C}(q^{-1})$ 是稳定的，则有当 $t_0=-\infty$ 时的 Åström 预报器

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{G}(q^{-1})\tilde{\boldsymbol{C}}^{-1}(q^{-1})\boldsymbol{y}(t) \tag{4.15.67}$$

$$\boldsymbol{C}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})=\tilde{\boldsymbol{A}}(q^{-1})\tilde{\boldsymbol{C}}^{-1}(q^{-1}) \tag{4.15.68}$$

$$\det\boldsymbol{C}(q^{-1})=\det\tilde{\boldsymbol{C}}(q^{-1}) \tag{4.15.69}$$

$$\tilde{\boldsymbol{C}}(q^{-1})=\boldsymbol{F}(q^{-1})\tilde{\boldsymbol{A}}(q^{-1})+q^{-k}\boldsymbol{G}(q^{-1}) \tag{4.15.70}$$

其中 $\boldsymbol{F}(q^{-1})=\boldsymbol{F}_0+\boldsymbol{F}_1q^{-1}+\cdots+\boldsymbol{F}_{k-1}q^{-(k-1)}$，$\boldsymbol{F}_0=\boldsymbol{I}_m$，$\boldsymbol{G}(q^{-1})=\boldsymbol{G}_0+\boldsymbol{G}_1q^{-1}+\cdots+\boldsymbol{G}_{n_g}q^{-n_g}$，$n_g=\max(n_a-1,n_c-k)$，

预报误差为

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{y}(t+k)-\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{F}(q^{-1})\boldsymbol{e}(t+k) \tag{4.15.71}$$

预报误差方差为

$$\mathrm{E}[\tilde{\boldsymbol{y}}(t+k\mid t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+k\mid t)]=\sum_{i=0}^{k-1}\boldsymbol{F}_i\boldsymbol{Q}_e\boldsymbol{F}_{i-1} \tag{4.15.72}$$

$e(t)$由(4.15.37)和(4.15.38)计算.特别由(4.15.37)给出的 $\boldsymbol{e}(t)$(即由(4.15.79)给出的 $\boldsymbol{e}_\infty(t)$)是稳态新息,即

$$\boldsymbol{y}(t+1)-\hat{\boldsymbol{y}}(t+1|t)=\boldsymbol{e}(t+1) \tag{4.15.73}$$

且有稳态新息方差阵为 $\boldsymbol{Q}_e$.

证明 暂时设 t_0 为有限的,对(4.15.34)取初值($\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(t_0-n_a),\boldsymbol{e}(t_0-1),\cdots,\boldsymbol{e}(t_0-n_c)$),则 $\boldsymbol{y}(t)$和 $\boldsymbol{e}(t)$可分别递推计算为

$$\begin{aligned}&\boldsymbol{y}(t)=-\boldsymbol{A}_1y(t-1)\cdots-\boldsymbol{A}_{n_a}\boldsymbol{y}(t-n_a)+\boldsymbol{C}(q^{-1})\boldsymbol{e}(t),\\&\boldsymbol{e}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)-\boldsymbol{C}_1e(t-1)\cdots-\boldsymbol{C}_{n_c}\boldsymbol{e}(t-n_c),\\&\qquad\qquad t=t_0,t_0+1\end{aligned} \tag{4.15.74}$$

且由它们生成相同的线性流形

$$L(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(t_0))=L(\boldsymbol{e}(t),\boldsymbol{e}(t-1),\cdots,\boldsymbol{e}(t_0)) \tag{4.15.75}$$

应用(4.15.68)和(4.15.70) $\boldsymbol{y}(t+k)$可分解为

$$\begin{aligned}\boldsymbol{y}(t+k)&=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{e}(t+k)=\widetilde{\boldsymbol{C}}(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\boldsymbol{e}(t+k)=\\&\boldsymbol{F}(q^{-1})\boldsymbol{e}(t+k)+\boldsymbol{G}(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\boldsymbol{e}(t)\end{aligned} \tag{4.15.76}$$

取上式两边各项在线性流形 $L(\boldsymbol{e}(t),\boldsymbol{e}(t-1),\cdots,\boldsymbol{e}(t_0))$上的射影,由 $\boldsymbol{e}(t)$为白噪声可得非稳态最优预报器

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{G}(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\boldsymbol{e}(t) \tag{4.15.77}$$

且非稳态预报误差为

$$\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{F}(q^{-1})\boldsymbol{e}(t+k) \tag{4.15.78}$$

这是因为上两式均与初值有关.因而是非稳态的.

现在令 $t_0\to-\infty$,由(4.15.34)和 $\boldsymbol{C}(q^{-1})$的稳定性引出白噪声 $\boldsymbol{e}(t)$的稳态值为

$$\boldsymbol{e}_\infty(t)=\boldsymbol{C}^{-1}(q^{-1})\boldsymbol{z}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Pi}_jq^{-j}\boldsymbol{z}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Pi}_j\boldsymbol{z}(t-j) \tag{4.15.79}$$

$$\boldsymbol{z}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) \tag{4.15.80}$$

事实上,令 $\boldsymbol{\delta}(t)=\boldsymbol{e}(t)-\boldsymbol{e}_\infty(t)$,因为 $\boldsymbol{e}(t)$满足(4.15.74)或(4.15.34),即

$$\boldsymbol{z}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{4.15.81}$$

又由(4.15.79)有

$$\boldsymbol{z}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}_\infty(t) \tag{4.15.82}$$

上两式相减引出

$$\boldsymbol{C}(q^{-1})\boldsymbol{\delta}(t)=\boldsymbol{0} \tag{4.15.83}$$

因 $\boldsymbol{C}(q^{-1})$的稳定性和差分方程理论引出

$$\boldsymbol{\delta}(t)\to\boldsymbol{0}\quad(t_0\to-\infty) \tag{4.15.84}$$

即我们有

$$\lim_{t_0\to\infty}\boldsymbol{e}(t)=\boldsymbol{e}_\infty(t) \tag{4.15.85}$$

于是在(4.15.77)和(4.15.78)中令 $t_0\to-\infty$有稳态最优预报器

$$\hat{\boldsymbol{y}}(t+k|t)=\boldsymbol{G}(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\boldsymbol{e}_\infty(t) \tag{4.15.86}$$

及稳态预报误差

$$\tilde{\boldsymbol{y}}(t+k|t)=\boldsymbol{F}(q^{-1})\boldsymbol{e}_{\infty}(t+k) \tag{4.15.87}$$

应用(4.15.70)、(4.15.79)和(4.15.80)，(4.15.86)可化为 Åström 预报器

$$\begin{aligned}\hat{\boldsymbol{y}}(t+k|t)&=\boldsymbol{G}(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\boldsymbol{C}^{-1}(q^{-1})\boldsymbol{z}(t)=\\ &\boldsymbol{G}(q^{-1})\widetilde{\boldsymbol{C}}(q^{-1})\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{G}(q^{-1})\widetilde{\boldsymbol{C}}^{-1}(q^{-1})\boldsymbol{y}(t)\end{aligned} \tag{4.15.88}$$

重新记 $\boldsymbol{e}_{\infty}(t)$为 $\boldsymbol{e}(t)$，其中 $\boldsymbol{e}(t)$由(4.15.37)~(4.15.38)计算，则由(4.15.86)有(4.15.71)和(4.15.72). 在(4.15.72)中置 $k=1$ 引出(4.15.73)，即 $\boldsymbol{e}(t)=\boldsymbol{e}_{\infty}(t)$是稳态新息过程. 下面证明

$$\boldsymbol{Q}_{e\infty}\triangleq \mathrm{E}[\boldsymbol{e}_{\infty}(t)\boldsymbol{e}_{\infty}^{\mathrm{T}}(t)]=\boldsymbol{Q}_e \tag{4.15.89}$$

事实上，由(4.15.79)和(4.15.80)引出稳态新息 $e_{\infty}(t)$满足

$$\boldsymbol{z}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}_{\infty}(t) \tag{4.15.90}$$

另一方面由(4.15.34)有

$$\boldsymbol{z}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{e}(t) \tag{4.15.91}$$

因 $\boldsymbol{C}(q^{-1})$是稳定的，且首系数阵 $\boldsymbol{C}_0=\boldsymbol{I}_m$. 因上述两个 MA 过程 $z(t)$有相同的相关函数，由定理 2.13.2 给出的 Gevers - Wouters 算法引出

$$\boldsymbol{Q}_{e\infty}=\mathrm{E}[\boldsymbol{e}_{\infty}(t)\boldsymbol{e}_{\infty}^{\mathrm{T}}(t)]=\boldsymbol{Q}_e \tag{4.15.92}$$

即稳态新息方差 $\boldsymbol{Q}_{e\infty}$为 $\boldsymbol{Q}_e$. 证毕. □

【定理 4.15.5】 多变量非平稳 ARMA 过程(4.15.34)在定理 4.15.2 条件下，当 $t_0=-\infty$时有关系

$$\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}^{\mathrm{T}}(j)]=\boldsymbol{F}_{t-j}\boldsymbol{Q}_e,\quad \forall\, t,j \tag{4.15.93}$$

其中 $m\times m$ 系数阵 $\boldsymbol{F}_i$ 由下式递推计算

$$\boldsymbol{F}_i=-\boldsymbol{A}_1\boldsymbol{F}_{i-1}-\cdots-\boldsymbol{A}_{n_a}\boldsymbol{F}_{i-n_a}+\boldsymbol{C}_i \tag{4.15.94}$$

其中规定 $\boldsymbol{F}_i=\boldsymbol{0}(i<0)$，$\boldsymbol{C}_i=\boldsymbol{0}(i>n_c)$. 稳定新息 $e(t)$由(4.15.79)置 $e(t)=e_{\infty}(t)$定义.

证明 暂时假设初始观测时刻 t_0 是有限的，取常的初值($\boldsymbol{y}(t_0-1),\cdots,\boldsymbol{y}(t_0-n_a)$，$\boldsymbol{e}(t_0-1),\cdots,\boldsymbol{e}(t_0-n_c)$)由(4.15.74)可生成相等的有限维线性流形

$$L(\boldsymbol{y}(t),\cdots,\boldsymbol{y}(t_0))=L(\boldsymbol{e}(t),\cdots,\boldsymbol{e}(t_0)) \tag{4.15.95}$$

因为 $\boldsymbol{y}(t)$是非平稳时间序列，取自然数 p 充分大，引入 Diophantine 方程

$$\boldsymbol{C}(q^{-1})=\boldsymbol{A}(q^{-1})\boldsymbol{F}(q^{-1})+q^{-p}\boldsymbol{G}(q^{-1}) \tag{4.15.96}$$

$$\boldsymbol{F}(q^{-1})=\boldsymbol{F}_0+\boldsymbol{F}_1q^{-1}+\cdots+\boldsymbol{F}_{p-1}q^{-(p-1)},$$

$$\boldsymbol{G}(q^{-1})=\boldsymbol{G}_0+\boldsymbol{G}_1q^{-1}+\cdots+\boldsymbol{G}_{n_g}q^{-n_g},$$

$$n_g=\max(n_a-1,n_c-p) \tag{4.15.97}$$

由(4.15.34)和(4.15.96) $\boldsymbol{y}(t)$有分解式

$$\boldsymbol{y}(t)=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{e}(t)=\boldsymbol{F}(q^{-1})\boldsymbol{e}(t)+q^{-p}\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{G}(q^{-1})\boldsymbol{e}(t) \tag{4.15.98}$$

在有限维线性流形(4.15.95)中因所有元素均有二阶矩，且因 $\boldsymbol{e}(t)$是零均值、方差阵为 $\boldsymbol{Q}_e$ 的白噪声，注意 $\boldsymbol{y}(t)\in L(\boldsymbol{e}(t),\cdots,\boldsymbol{e}(t_0))$，$\boldsymbol{e}(j)\perp L(\boldsymbol{e}(t),\cdots,\boldsymbol{e}(t_0))$ $(j>t)$，于是有数学期望

$$\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}^{\mathrm{T}}(j)]=\boldsymbol{0},\quad t<j \tag{4.15.99}$$

当 $j \leqslant t$ 时,取 p 充分大使 $t-p<j$,于是有

$$\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}^{\mathrm{T}}(j)]=\mathrm{E}[(\boldsymbol{F}(q^{-1})\boldsymbol{e}(t)+q^{-p}\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{G}(q^{-1})\boldsymbol{e}(t))\boldsymbol{e}^{\mathrm{T}}(j)]=$$

$$\mathrm{E}\Big[\sum_{i=0}^{p-1}\boldsymbol{F}_i\boldsymbol{e}(t-i)\boldsymbol{e}^{\mathrm{T}}(j)\Big]=\boldsymbol{F}_{t-j}\boldsymbol{Q}_e(j) \tag{4.15.100}$$

其中定义非稳态新息 $\boldsymbol{e}(t)$ 的方差阵为

$$\boldsymbol{Q}_e(t)=\mathrm{E}[\boldsymbol{e}(t)\boldsymbol{e}^{\mathrm{T}}(t)] \tag{4.15.101}$$

这是因为由(4.15.74)计算的非稳态新息 $\boldsymbol{e}(t)$ 与初值($\boldsymbol{y}(t_0-1),\cdots,\boldsymbol{y}(t_0-n_a),\boldsymbol{e}(t_0-1),\cdots,\boldsymbol{e}(t_0-n_c)$)有关,因而它的方差阵是时变的.当 $t_0\to-\infty$,因为

$$\lim_{t_0\to-\infty}\boldsymbol{e}(t)=\boldsymbol{e}_\infty(t) \tag{4.15.102}$$

其中 $\boldsymbol{e}_\infty(t)$ 是稳态新息,由(4.15.79)定义,从而由(4.15.101)和(4.15.92)有

$$\lim_{t_0\to-\infty}\boldsymbol{Q}_e(t)=\mathrm{E}[\boldsymbol{e}_\infty(t)\boldsymbol{e}_\infty^{\mathrm{T}}(t)]=\boldsymbol{Q}_e \tag{4.15.103}$$

在(4.15.100)中令 $t_0\to-\infty$ 并应用(4.15.103)有极限关系

$$\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}_\infty^{\mathrm{T}}(j)]\triangleq\lim_{t_0\to-\infty}\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}^{\mathrm{T}}(j)]=\lim_{t_0\to-\infty}\boldsymbol{F}_{t-j}\boldsymbol{Q}_e(j)=\boldsymbol{F}_{t-j}\boldsymbol{Q}_e \tag{4.15.104}$$

若将 $\boldsymbol{e}_\infty(t)$ 记为 $\boldsymbol{e}(t)$,$\boldsymbol{e}_\infty(t)\triangleq\boldsymbol{e}(t)$,但 $\boldsymbol{e}(t)$ 由(4.5.79)定义,则(4.15.104)即(4.15.93).因 $\boldsymbol{F}_j=\boldsymbol{0}\,(j<0)$,故(4.15.99)也是(4.15.93)的特例.证毕. □

【注 4.15.1】 定理 4.15.5 的(4.15.93)$\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}^{\mathrm{T}}(j)]=\boldsymbol{F}_{t-j}\boldsymbol{Q}_e$ 是一种从有限到无限的极限关系.当 $t_0\to-\infty$,因 $y(t)$ 有无穷大方差,故在通常意义下数学期望 $\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}^{\mathrm{T}}(j)]$ 是不存在的.只能在(4.15.104)的极限意义下来定义 $\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}^{\mathrm{T}}(j)]$.在 Kalman 滤波框架下的定理 4.5.3 也是在极限意义下定义 $\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}^{\mathrm{T}}(j)]$.

【注 4.15.2】 由 Diophantine 方程(4.15.96)决定的 $\boldsymbol{F}(q^{-1})$,随不同的 p,$\boldsymbol{F}(q^{-1})$ 的阶次($p-1$)是不同的,但 $\boldsymbol{F}(q^{-1})$ 的系数阵并不随 p 变化而变化,记取自然数 p 时由 Diophantine 方程(4.15.96)决定的 $\boldsymbol{F}(q^{-1})$ 和 $\boldsymbol{G}(q^{-1})$ 为 $\boldsymbol{F}_p(q^{-1})$ 和 $\boldsymbol{G}_p(q^{-1})$,即

$$\boldsymbol{C}(q^{-1})=\boldsymbol{A}(q^{-1})\boldsymbol{F}_p(q^{-1})+q^{-p}\boldsymbol{G}_p(q^{-1}) \tag{4.15.105}$$

$$\boldsymbol{F}_p(q^{-1})=\boldsymbol{F}_0+\boldsymbol{F}_1q^{-1}+\cdots+\boldsymbol{F}_{p-1}q^{-(p-1)}, \tag{4.15.106}$$

记取自然数 $p+1$ 由(4.15.96)决定的 $\boldsymbol{F}(q^{-1})$ 和 $\boldsymbol{G}(q^{-1})$ 为 $\boldsymbol{F}_{p+1}(q^{-1})$ 和 $\boldsymbol{G}_{p+1}(q^{-1})$,即

$$\boldsymbol{C}(q^{-1})=\boldsymbol{A}(q^{-1})\boldsymbol{F}_{p+1}(q^{-1})+q^{-p+1}\boldsymbol{G}_{p+1}(q^{-1}) \tag{4.15.107}$$

可以容易证明

$$\boldsymbol{F}_{p+1}(q^{-1})=\boldsymbol{F}_p(q^{-1})+\boldsymbol{F}_pq^{-p} \tag{4.15.108}$$

【注 4.15.3】 对多变量平稳 ARMA 过程(4.15.37)而言,定理 4.15.5 自然成立.因为 $A(q^{-1})$ 是稳定的,当 $t_0=-\infty$ 时有均方收敛的无穷级数展式

$$\boldsymbol{y}(t)=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{e}(t)=\sum_{j=0}^{\infty}\boldsymbol{F}_jq^{-j}\boldsymbol{e}(t)=\sum_{j=0}^{\infty}\boldsymbol{F}_j\boldsymbol{e}(t-j) \tag{4.15.109}$$

由等式

$$\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})=\sum_{j=0}^{\infty}\boldsymbol{F}_jq^{-j}\ \text{或}\ \boldsymbol{C}(q^{-1})=\boldsymbol{A}(q^{-1})\sum_{j=0}^{\infty}\boldsymbol{F}_jq^{-j} \tag{4.15.110}$$

用比较两边 q^{-j} 的系数阵可得系数阵 $\boldsymbol{F}_j$ 的递推公式为

$$\boldsymbol{F}_j=-\boldsymbol{A}_1\boldsymbol{F}_{j-1}-\cdots-\boldsymbol{A}_{n_a}\boldsymbol{F}_{j-n_a}+\boldsymbol{C}_j \tag{4.15.111}$$

其中规定 $\boldsymbol{F}_j=\boldsymbol{0}(j<0)$，$\boldsymbol{C}_j=\boldsymbol{0}(j>n_a)$. 于是有 $\boldsymbol{e}(t)$是白噪声有

$$\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{e}^{\mathrm{T}}(j)]=\mathrm{E}\Big[\sum_{i=0}^{\infty}\boldsymbol{F}_i\boldsymbol{e}(t-i)\boldsymbol{e}^{\mathrm{T}}(j)\Big]=\boldsymbol{F}_{t-j}\boldsymbol{Q}_e \tag{4.15.112}$$

即定理 4.15.5 成立.

在关键技术上,处理平稳和非平稳 ARMA 过程的区别是:对平稳情形我们可将 $\boldsymbol{y}(t)$展为 $\boldsymbol{e}(t)$的均方收敛的无穷级数,而对非平稳情形,由于 $\boldsymbol{A}(q^{-1})$不稳定,因而 $\boldsymbol{y}(t)$不能展成关于 $\boldsymbol{e}(t)$的均方收敛级数,只能通过 Diophantine 方程得到有限形式的展开. 前者等价于无穷级数展式

$$\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})=\sum_{j=0}^{\infty}\boldsymbol{F}_jq^{-j} \tag{4.15.113}$$

后者等价于有限形式展式

$$\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})=\sum_{j=0}^{p-1}\boldsymbol{F}_jq^{-j}+q^{-p}\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{G}(q^{-1}) \tag{4.15.114}$$

上式两边左乘 $A(q^{-1})$即 Diophantine 方程

$$\boldsymbol{C}(q^{-1})=\boldsymbol{A}(q^{-1})\boldsymbol{F}(q^{-1})+q^{-p}\boldsymbol{G}(q^{-1}) \tag{4.15.115}$$

这等价于综合除法.

【注 4.15.4】 本节的技术关键是从非稳态新息 $e(t)$到稳态新息 $e_\infty(t)$的转化. 这等价于在状态空间模型下由稳态 Kalman 预报器取初值后决定的非稳态新息到稳态新息的转化.

【例 4.15.2】 带位置和速度观测的非平稳雷达跟踪预报系统.

考虑带位置和速度观测的二维雷达跟踪系统

$$(\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1})\boldsymbol{x}(t)=\boldsymbol{C}_1q^{-1}w(t) \tag{4.15.116}$$

$$\boldsymbol{y}(t)=\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{4.15.117}$$

其中 T 为采样周期,$\boldsymbol{x}(t)=[x_1(t),x_2(t)]^{\mathrm{T}}$, $x_1(t)$, $x_2(t)$和 $w(t)$是运动目标(飞机、导弹、车辆等)在时刻 tT 的位置、速度和加速度,$\boldsymbol{y}(t)\in R^2$ 为对 $\boldsymbol{x}(t)$的观测信号,$\boldsymbol{v}(t)\in R^2$ 是观测噪声,设 $w(t)$和 $\boldsymbol{v}(t)$是零均值、方差阵各为 σ_w^2和 Q_v 的相互独立正态白噪声,且

$$\boldsymbol{A}_1=-\begin{bmatrix}1 & T\\ 0 & 1\end{bmatrix},\quad \boldsymbol{C}_1=\begin{bmatrix}0.5T^2\\ T\end{bmatrix},\quad \sigma_w^2=0.025,\quad \boldsymbol{Q}_v=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \tag{4.15.118}$$

因 $\boldsymbol{A}(q^{-1})=\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1}$是不稳定的多项式矩阵,即 $\det A(q^{-1})$有在单位圆上的零点 $q^{-1}=1$,故 $\boldsymbol{x}(t)$是非平稳 ARMA 过程. 问题是基于观测$(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$,求信号 $\boldsymbol{x}(t)$的稳态最优 Box – Jenkins 预报器 $\hat{\boldsymbol{x}}(t+2|t)$.

将(4.15.116)代入(4.15.117)引出关系

$$(\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1})\boldsymbol{y}(t)=\boldsymbol{C}_1q^{-1}w(t)+(\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1})\boldsymbol{v}(t) \tag{4.15.119}$$

将上式右边两个 MA 过程用一个等价的 MA 过程表示引出 ARMA 新息模型

$$(\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1})\boldsymbol{y}(t)=(\boldsymbol{I}_2+\boldsymbol{D}_1q^{-1})\boldsymbol{\varepsilon}(t) \tag{4.15.120}$$

其中新息 $\boldsymbol{\varepsilon}(t)\in R^2$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声,且有关系

$$(\boldsymbol{I}_2+\boldsymbol{D}_1q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{C}_1q^{-1}w(t)+(\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1})\boldsymbol{v}(t) \tag{4.15.121}$$

利用 Gevers – Wouters 算法可求得

$$\boldsymbol{D}_1=\begin{bmatrix}-0.720\,597 & -0.430\,682\\ 0.066\,495 & -0.967\,95\end{bmatrix},\quad \boldsymbol{Q}_\varepsilon=\begin{bmatrix}1.127\,430\,2 & 0.141\,021\\ 0.141\,021 & 1.027\,110\end{bmatrix} \tag{4.15.122}$$

注意由(4.15.117)有关系

$$\hat{\boldsymbol{x}}(t+2|t)=\hat{\boldsymbol{y}}(t+2|t) \tag{4.15.123}$$

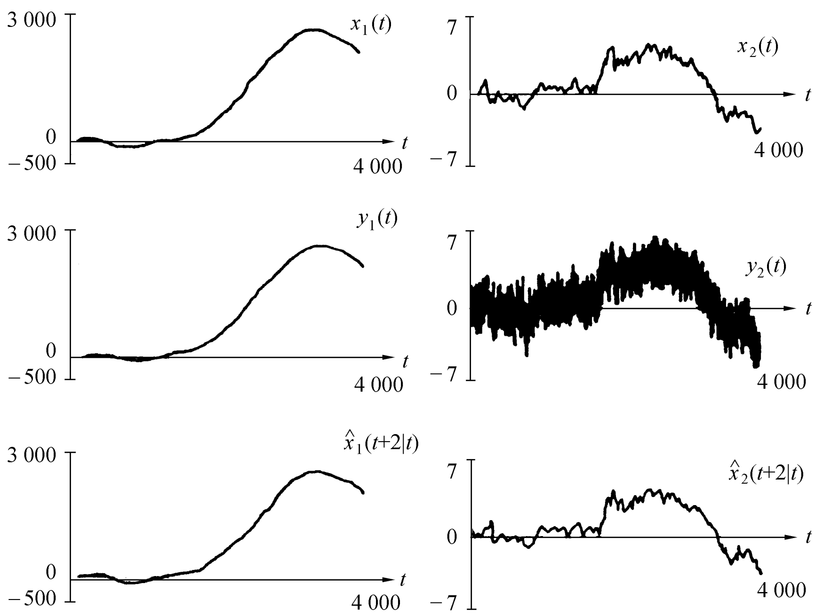
因此归结为对非平稳 ARMA 新息模型(4.15.120)求 Box - Jenkins 最优递推预报器

$$\hat{\boldsymbol{x}}(t+1|t)=\hat{\boldsymbol{y}}(t+1|t)=-\boldsymbol{A}_1\boldsymbol{y}(t)+\boldsymbol{D}_1\boldsymbol{\varepsilon}(t),$$

$$\hat{\boldsymbol{x}}(t+2|t)=\hat{\boldsymbol{y}}(t+2|t)=-\boldsymbol{A}_1\hat{\boldsymbol{y}}(t+1|t) \tag{4.15.124}$$

$$\boldsymbol{\varepsilon}(t)=(\boldsymbol{I}_2+\boldsymbol{A}_1q^{-1})\boldsymbol{y}(t)-\boldsymbol{D}_1\boldsymbol{\varepsilon}(t-1),\quad \boldsymbol{\varepsilon}(0)=\boldsymbol{0},\quad y(0)=\boldsymbol{0},\quad t=1,2,\cdots \tag{4.15.125}$$

仿真结果如图 4.15.1 所示,可看到 $\hat{x}_i(t+2|t)$很好地跟踪 $x_i(t)$的变化,特别 $\hat{x}_2(t+2|t)$明显地过滤了观测噪声 $v_2(t)$,从观测 $y_2(t)$中还 $x_2(t)$的本来面目.



(a) **位置** $x_1(t)$、**观测** $y_1(t)$**和** Box–Jenkins **预报器** $\hat{x}_1(t+2|t)$

(b) **速度** $x_2(t)$、**观测** $y_2(t)$**和** Box–Jenkins **预报器** $\hat{x}_2(t+2|t)$

图 4.15.1　非平稳过程雷达跟踪预报器

4.16　MA 参数估计的 Gevers - Wouters 算法收敛性分析

滑动平均(MA)模型广泛应用于时间序列分析、最优滤波、信号处理、通讯和系统辨识等领域. MA 参数估计问题等价于一个谱分解问题,是一个困难的非线性估计问题. 大量仿真结果表明,MA 参数估计的 Gevers - Wouters 迭代算法具有快的收敛性. 一般取迭代次数为 50 ~ 100 左右,MA 参数估计误差可小于 10^{-3}. 但是到目前为止,Gevers - Wouters 迭代算法的收敛性和一致性问题没有解决,即 MA 参数估值是否收敛到真实值? 收敛速度是

什么？收敛速度由什么决定？本节应用 Kalman 滤波方法证明了 MA 参数估计的 Gevers－Wouters 迭代算法的指数收敛性[24]，即 MA 参数估计误差以指数律衰减至零，且衰减速度由 MA 多项式的零点位置决定.

4.16.1 问题阐述

考虑可逆的 MA 模型

$$r(t)=e(t)+d_1e(t-1)+\cdots+d_ne(t-n) \tag{4.16.1}$$

其中 t 为离散时间，$r(t)$是输出(观测)，d_i 为模型参数。

【假设 1】 $e(t)$是零均值、方差 σ_e^2 的白噪声，即

$$\mathrm{E}e(t)=0,\quad \mathrm{E}[e(t)e(j)]=\sigma_e^2\delta_{tj} \tag{4.16.2}$$

其中 E 为均值号，$\delta_{tt}=1$，$\delta_{tj}=0\,(t\neq j)$.

式(4.16.1)可写为

$$r(t)=D(q^{-1})e(t) \tag{4.16.3}$$

其中 q^{-1}为单位滞后算子，且 MA 多项式 $D(q^{-1})$为

$$D(q^{-1})=1+d_1q^{-1}+\cdots+d_nq^{-n} \tag{4.16.4}$$

【假设 2】 $D(q^{-1})$是稳定的多项式，即 $D(x)$的所有零点位于单位圆外，换言之，MA 模型是可逆的.

【假设 3】 已知它的相关函数

$$R_r(i)=\mathrm{E}[r(t)r(t-i)],\quad i=0,1,\cdots,n$$
$$R_r(i)=0,\quad i>n \tag{4.16.5}$$

但参数 d_i 和σ_e^2 是未知的.

【假设 4】 初始时刻 $t_0=0$，且假设

$$e(i)=0,\quad i<0 \tag{4.16.6}$$

定义 $r(t)$的谱为

$$S(q)=\sum_{j=-n}^{n}R_r(j)q^{-j} \tag{4.16.7}$$

其中 $R_r(j)=R_r(-j)$. 容易验证

$$S(q)=D(q^{-1})\sigma_e^2D(q) \tag{4.16.8}$$

定义 $r(t)$与 $e(t-i)$的协方差函数为

$$R_{re}(t,t-i)=\mathrm{E}[r(t)e(t-i)],\quad i=0,1,\cdots,n$$
$$R_{re}(t,t-i)=0,\quad i>n \tag{4.16.9}$$

由假设 4 引出

$$R_{re}(0,0)=R_r(0) \tag{4.16.10}$$

问题是由已知的 $R_r(i)$求 MA 参数 d_i 和σ_e^2 的估值。这等价于在(4.16.8)中已知谱 $S(q)$求 $D(q^{-1})$和 σ_e^2. 这个问题也叫谱分解问题.

令 $t_0=-\infty$，在稳态情形，由(4.16.1)和(4.16.2)有 $R_{re}(t,t-i)$与 t 无关，且有

$$d_i=R_{re}(t,t-i)/R_{re}(t,t),\quad i=1,\cdots,n$$
$$\sigma_e^2=R_{re}(t,t) \tag{4.16.11}$$

取 $t_0=0$,在初始条件(4.16.6)和(4.16.10)下,$R_{re}(t,t-i)$和 $R_{re}(t,t)$与 t 有关,由节 2.13 可得到 MA 参数估值的 Gevers – Wouters 算法为

$$d_i=\lim_{t\to\infty}\frac{R_{re}(t,t-i)}{R_{re}(t,t)},\quad i=1,\cdots,n$$

$$\sigma_e^2=\lim_{t\to\infty}R_{re}(t,t) \tag{4.16.12}$$

其中 $R_{re}(t,t-i)$可用迭代法计算为

$$R_{re}(t,t-i)=R_r(i)-\sum_{s=i+1}^{n}\frac{R_{re}(t,t-s)R_{re}(t-i,t-s)}{R_{re}(t-s,t-s)} \tag{4.16.13}$$

其中定义

$$R_{re}(t,t-s)=0\quad(t<s)$$

$$R_{re}^{-1}(t-s,t-s)=0\quad(t<s) \tag{4.16.14}$$

其中 $t=0,1,2,\cdots,i=0,1,\cdots,n$.定义 d_i 和 σ_e^2 的估值 $\hat{d}_i(t)$和 $\hat{\sigma}_e^2(t)$及估值误差 $\tilde{d}_i(t)$和 $\tilde{\sigma}_e^2$ 分别为

$$\hat{d}_i(t)=R_{re}(t,t-i)/R_{re}(t,t),\quad \hat{\sigma}_e^2=R_{re}(t,t),$$

$$\tilde{d}_i(t)=d_i-\hat{d}_i(t),\quad \tilde{\sigma}_e^2(t)=\sigma_i^2-\hat{\sigma}_e^2(t) \tag{4.16.15}$$

其中 $R_{re}(t,t-i)(i=0,1,\cdots,n)$由(4.16.13)和(4.16.14)计算。于是问题归结为证明:当 $t\to\infty$ 时

$$\lim\tilde{d}_i(t)=0,\quad \lim\tilde{\sigma}_e^2(t)=0 \tag{4.16.16}$$

4.16.2 MA 模型与状态空间模型的等价性

(4.16.1)有状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{d}e(t) \tag{4.16.17}$$

$$r(t)=\boldsymbol{Hx}(t)+e(t) \tag{4.16.18}$$

其中

$$\boldsymbol{\Phi}=\begin{bmatrix}0 & & & \\ \vdots & \boldsymbol{I}_{n-1} & & \\ 0 & 0 & \cdots & 0\end{bmatrix},\quad \boldsymbol{d}=\begin{bmatrix}d_1\\ \vdots\\ d_n\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0\quad \cdots\quad 0] \tag{4.16.19}$$

将(4.16.18)代入(4.16.17)有状态空间模型

$$\boldsymbol{x}(t+1)=\bar{\boldsymbol{\Phi}}\boldsymbol{x}(t)+\boldsymbol{d}r(t) \tag{4.16.20}$$

$$r(t)=\boldsymbol{Hx}(t)+e(t) \tag{4.16.21}$$

$$\bar{\boldsymbol{\Phi}}=\boldsymbol{\Phi}-\boldsymbol{dH}=\begin{bmatrix}-d_1 & & & \\ \vdots & \boldsymbol{I}_{n-1} & & \\ -d_n & 0 & \cdots & 0\end{bmatrix} \tag{4.16.22}$$

注意$(\bar{\boldsymbol{\Phi}},\boldsymbol{H})$是伴随形,因而$(\bar{\boldsymbol{\Phi}},\boldsymbol{H})$为完全可观对.于是当 $t_0=-\infty$ 时存在稳态 Kalman 预报器[15]

$$\hat{\boldsymbol{x}}(t+1|t)=\bar{\boldsymbol{\Phi}}\hat{\boldsymbol{x}}(t|t-1)+\boldsymbol{d}r(t)+\bar{\boldsymbol{K}}_p\varepsilon(t) \tag{4.16.23}$$

$$r(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t|t-1)+\varepsilon(t) \tag{4.16.24}$$

$$\bar{\boldsymbol{K}}_p=\bar{\boldsymbol{\Phi}}\boldsymbol{\Sigma} \tag{4.16.25}$$

其中 $\varepsilon(t)$是 $r(t)$的稳态新息过程,它是带零均值、方差为 σ_ε^2 的白噪声,而 $\boldsymbol{\Sigma}$ 是稳态预报

误差方差阵，满足 Riccati 方程

$$\boldsymbol{\Sigma}=\overline{\boldsymbol{\Phi}}\left[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\sigma_e^2)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}\right]\overline{\boldsymbol{\Phi}}^{\mathrm{T}} \tag{4.16.26}$$

其中 T 为转置号，它有正半定解

$$\boldsymbol{\Sigma}=\mathbf{0} \tag{4.16.27}$$

注意 $(\overline{\boldsymbol{\Phi}},\boldsymbol{H})$ 为完全可观对保证了 Riccati 方程存在稳态解，[15]从而存在稳态 Kalman 预报器，于是由(4.16.25)和(4.16.27)引出

$$\overline{\boldsymbol{K}}_p=\mathbf{0} \tag{4.16.28}$$

将(4.16.24)代入(4.16.23)，且应用(4.16.22)和(4.16.27)，可得稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{d}\varepsilon(t) \tag{4.16.29}$$

$$r(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t\mid t-1)+\varepsilon(t) \tag{4.16.30}$$

它等价于 MA 模型

$$r(t)=\varepsilon(t)+d_1\varepsilon(t-1)+\cdots+d_n\varepsilon(t-n) \tag{4.16.31}$$

它有与(4.16.1)相同的形式，于是当 $t_0=-\infty$ 时有

$$\varepsilon(t)=e(t),\quad \forall t \tag{4.16.32}$$

这是因为由可逆性引出 $\varepsilon(t)$ 和 $e(t)$ 可表为关于 $r(t)$ 的相同的无穷级数展式. 于是有 $\sigma_\varepsilon^2=\sigma_e^2$.

当 $t_0=0$ 时，问题是由(4.16.29)和(4.16.30)取初值 $\hat{x}(0\mid -1)$ 得到的非稳态新息 $\varepsilon(t)$ 是否相同于由(4.16.1)取初值 $(e(0),\cdots,e(n-1))$ 得到的 $e(t)$？

【定理 4.16.1】 当 $t_0=0$ 时，由(4.16.1)取初值 $(e(0),\cdots,e(n-1))$ 计算的非稳态新息过程 $e(t)$，即

$$e(t)=r(t)-d_1e(t-1)-\cdots-d_ne(t-n),\quad t=n,n+1,\cdots \tag{4.16.33}$$

数值上相同于由(4.16.29)和(4.16.30)按如下初值 $\hat{x}(0\mid -1)$ 计算的非稳态新息过程 $\varepsilon(t)$，即

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{d}\varepsilon(t) \tag{4.16.34}$$

$$r(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t\mid t-1)+\varepsilon(t) \tag{4.16.35}$$

带初值

$$\hat{\boldsymbol{x}}(0\mid -1)=\begin{bmatrix} r(0)-e(0)\\ r(1)-\boldsymbol{H}\boldsymbol{d}e(0)-e(1)\\ \vdots\\ r(n-1)-\sum\limits_{i=0}^{n-2}\boldsymbol{H}\boldsymbol{\Phi}^{n-2-i}\boldsymbol{d}e(i)-e(n-1)\end{bmatrix} \tag{4.16.36}$$

即在上述初始条件下，有相同的非稳态新息过程

$$e(t)=\varepsilon(t),\quad t-0,1,\cdots,n-1,n,n+1,\cdots \tag{4.16.37}$$

证明 重复应用(4.16.29)和(4.16.30)引出关系

$$\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{H}\boldsymbol{\Phi}\\ \vdots\\ \boldsymbol{H}\boldsymbol{\Phi}^{n-1}\end{bmatrix}\hat{\boldsymbol{x}}(0\mid -1)=\begin{bmatrix} r(0)-\varepsilon(0)\\ r(1)-\boldsymbol{H}\boldsymbol{d}\varepsilon(0)-\varepsilon(1)\\ \vdots\\ r(n-1)-\sum\limits_{i=0}^{n-2}\boldsymbol{H}\boldsymbol{\Phi}^{n-2-i}\boldsymbol{d}\varepsilon(i)-\varepsilon(n-1)\end{bmatrix} \tag{4.16.38}$$

注意由(4.16.19)有

$$\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{n-1} \end{bmatrix} = \boldsymbol{I}_n \tag{4.16.39}$$

其中 $\boldsymbol{I}_n$ 为 $n \times n$ 单位阵. 于是有

$$\hat{\boldsymbol{x}}(0|-1) = \begin{bmatrix} r(0) - \varepsilon(0) \\ r(1) - \boldsymbol{Hd}\varepsilon(0) - \varepsilon(1) \\ \vdots \\ r(n-1) - \sum_{i=0}^{n-2} \boldsymbol{H\Phi}^{n-2-i}\boldsymbol{d}\varepsilon(i) - \varepsilon(n-1) \end{bmatrix} \tag{4.16.40}$$

注意(4.16.29)和(4.16.30)取初值(4.16.40)计算 $\varepsilon(t)$等价于(4.16.31)取初值($\varepsilon(0)$,…,$\varepsilon(n-1)$)计算 $\varepsilon(t)$. 还应注意(4.16.31)与(4.16.1)有相同的形式。因此我们有(4.16.37)成立,假如在(4.16.40)中取

$$\varepsilon(i) = e(i), \quad i = 0,1,\cdots,n-1 \tag{4.16.41}$$

将上式代入(4.16.40)引出(4.16.36). 证毕. □

4.16.3 用 Kalman 滤波方法的 Gevers – Wouters 算法的收敛性分析

当初始时刻 $t_0 = 0$ 时,在初值(4.16.6)下由(4.16.33)可递推得到($e(0)$,…,$e(n-1)$). 应用定理 4.16.1 和(4.16.37)有非稳态协方差函数

$$R_{re}(t+k,t) = \mathrm{E}[r(t+k)e(t)] = \mathrm{E}[r(t+k)\varepsilon(t)], \quad k \geqslant 0 \tag{4.16.42}$$

重复应用(4.16.17)引出

$$\boldsymbol{x}(t+k) = \boldsymbol{\Phi}^k\boldsymbol{x}(t) + \sum_{i=t+1}^{t+k} \boldsymbol{\Phi}^{t+k-i}\boldsymbol{d}e(i-1) \tag{4.16.43}$$

由(4.16.18)和上式有

$$r(t+k) = \boldsymbol{H\Phi}^k\boldsymbol{x}(t) + \sum_{i=t+1}^{t+k} \boldsymbol{H\Phi}^{t+k-i}\boldsymbol{d}e(i-1) + e(t+k) \tag{4.16.44}$$

将(4.16.18)代入(4.16.35)引出

$$\varepsilon(t) = \boldsymbol{H}\tilde{\boldsymbol{x}}(t|t-1) + e(t) \tag{4.16.45}$$

注意,在(4.16.44)和(4.16.45)中,像在(4.16.17)和(4.16.18)中一样,$e(t)$是带零均值、方差为 σ_e^2 白噪声,且 $\tilde{\boldsymbol{x}}(t|t-1) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t|t-1)$或

$$\boldsymbol{x}(t) = \tilde{\boldsymbol{x}}(t|t-1) + \hat{\boldsymbol{x}}(t|t-1) \tag{4.16.46}$$

其中 $\hat{\boldsymbol{x}}(t|t-1)$是带初值(4.16.36)的非稳态 Kalman 预报器(4.16.34)和(4.16.35). 将(4.16.46)代入(4.16.44),且将(4.16.44)和(4.16.45)代入(4.16.42),注意 $\tilde{\boldsymbol{x}}(t|t-1)$不相关于 $\hat{\boldsymbol{x}}(t|t-1)$和 $e(t+i)$ $(i \geqslant 0)$,则有

$$R_{re}(t+k,t) = \boldsymbol{H\Phi}^k\boldsymbol{\Sigma}(t|t-1)\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{H\Phi}^{k-1}\boldsymbol{d}\sigma_e^2, \quad k \geqslant 0 \tag{4.16.47}$$

其中定义预报误差方差阵为 $\boldsymbol{\Sigma}(t|t-1) = \mathrm{E}[\tilde{\boldsymbol{x}}(t|t-1)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t|t-1)]$.

【定理 4.16.2】 MA 模型(4.16.1)在假设 1 ~ 4 和初值($e(0)$,…,$e(n-1)$)下,非稳态协方差函数 $R_{re}(t+k,t)$为

$$R_{re}(t+k,t) = \boldsymbol{H\Phi}^k\boldsymbol{\Sigma}(t|t-1)\boldsymbol{H}^{\mathrm{T}} + d_k\sigma_e^2, \quad k \geqslant 0, d_0 = 1 \tag{4.16.48}$$

证明 应用(4.16.19)引出

$$H\Phi^{k-1} = [0\cdots0\ 1\ 0\cdots0] \tag{4.16.49}$$

它的第 k 列元素为 1. 应用(4.16.19)和(4.16.49),则(4.16.47)化为(4.16.48). 证毕. □

【定理 4.16.3】 在假设 1 ~ 4 下,$\boldsymbol{\Sigma}(t|t-1)$满足特殊的 Lyapunov 方程

$$\boldsymbol{\Sigma}(t|t-1) = \bar{\boldsymbol{\Phi}}\boldsymbol{\Sigma}(t|t-1)\bar{\boldsymbol{\Phi}}^{\mathrm{T}} \tag{4.16.50}$$

其中 $\bar{\boldsymbol{\Phi}}$ 由(4.16.22)定义,且它是一个稳定矩阵,而 $\bar{\boldsymbol{\Phi}}$ 的所有特征值 $\lambda_1,\cdots,\lambda_n$ 位于单位圆内。$\boldsymbol{\Sigma}(t|t-1)$指数地衰减到零,即

$$\|\boldsymbol{\Sigma}(t|t-1)\|_s \leqslant c\rho_0^t \tag{4.16.51}$$

其中 $\|\cdot\|_s$ 表示矩阵的谱范数,常数 $c>0$,$\rho_0=\rho^2$,而 ρ 是 $\bar{\boldsymbol{\Phi}}$ 的谱半径,即

$$0<\rho=\max(|\lambda_1|,\cdots,|\lambda_n|)<1 \tag{4.16.52}$$

(4.16.51)可用符号表示为

$$\|\boldsymbol{\Sigma}(t|t-1)\| = \mathrm{O}(\rho_0^t) \tag{4.16.53}$$

证明 将(4.16.35)代入(4.16.34)引出非稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1|t) = \bar{\boldsymbol{\Phi}}\hat{\boldsymbol{x}}(t|t-1) + \boldsymbol{d}r(t) \tag{4.16.54}$$

由(4.16.20)减(4.16.54)有

$$\tilde{\boldsymbol{x}}(t+1|t) = \bar{\boldsymbol{\Phi}}\tilde{\boldsymbol{x}}(t|t-1) \tag{4.16.55}$$

它引出(4.16.50). 由(4.16.22)我们得到

$$\det(\boldsymbol{I}_n - q^{-1}\bar{\boldsymbol{\Phi}}) = D(q^{-1}) \tag{4.16.56}$$

其中 $D(q^{-1})$由(4.16.4)定义. 因假设 $D(q^{-1})$是稳定的,故 $\bar{\boldsymbol{\Phi}}$ 是一个稳定矩阵,即它的特征值 λ_i 满足关系

$$0<|\lambda_i|<1,\quad i=1,\cdots,n \tag{4.16.57}$$

将(4.16.50)迭代 t 次有

$$\boldsymbol{\Sigma}(t|t-1) = \bar{\boldsymbol{\Phi}}^t\boldsymbol{\Sigma}(0|-1)\bar{\boldsymbol{\Phi}}^{Tt} \tag{4.16.58}$$

为证明简单起见,假设 $\bar{\boldsymbol{\Phi}}$ 有 n 个线性无关的特征向量,于是存在非异阵 $\boldsymbol{Q}$ 使

$$\boldsymbol{Q}^{-1}\bar{\boldsymbol{\Phi}}\boldsymbol{Q} = \mathrm{diag}(\lambda_1,\cdots,\lambda_n) \tag{4.16.59}$$

这引出

$$\bar{\boldsymbol{\Phi}}^t = \boldsymbol{Q}\,\mathrm{diag}(\lambda_1^t,\cdots,\lambda_n^t)\boldsymbol{Q}^{-1} \tag{4.16.60}$$

注意在矩阵谱范数意义下有

$$\|\mathrm{diag}(\lambda_1^{\mathrm{t}},\cdots,\lambda_{\mathrm{n}}^{\mathrm{t}})\|_s = \rho^t \tag{4.16.61}$$

其中 ρ 是 $\bar{\boldsymbol{\Phi}}$ 的谱半径。于是有

$$\|\bar{\boldsymbol{\Phi}}^t\|_s \leqslant \|\boldsymbol{Q}\|_s\|\mathrm{diag}(\lambda_1^{\mathrm{t}},\cdots,\lambda_{\mathrm{n}}^{\mathrm{t}})\|_s\|\boldsymbol{Q}^{-1}\|_s \tag{4.16.62}$$

它引出

$$\|\bar{\boldsymbol{\Phi}}^t\|_s \leqslant c_1\rho^t \tag{4.16.63}$$

带 $c_1=\|\boldsymbol{Q}\|_s\|\boldsymbol{Q}^{-1}\|_s$. 类似有

$$\|\bar{\Phi}^{Tt}\|_s \leqslant c_2\rho^t \tag{4.16.64}$$

带 $c_2=\|\boldsymbol{Q}^T\|_s\|\boldsymbol{Q}^{-T}\|_s$,其中记 $\boldsymbol{Q}^{-T}=(\boldsymbol{Q}^{-1})^{\mathrm{T}}$. 于是由(4.16.58)和上两式有

$$\|\boldsymbol{\Sigma}(t|t-1)\|_s \leqslant \|\bar{\boldsymbol{\Phi}}^t\|_s\|\boldsymbol{\Sigma}(0|-1)\|_s\|\bar{\boldsymbol{\Phi}}^{Tt}\|_s \leqslant c\rho_0^t \tag{4.16.65}$$

带 $c=c_1c_2\|\boldsymbol{\Sigma}(0|-1)\|_s$,$\rho_0=\rho^2$. 证毕. □

【定理 4.16.4】 MA 模型(4.16.1)在假设 1 ~ 4 下，Gevers - Wouters 算法(4.16.12) ~ (4.16.14)是一致的，且有指数收敛性，即 MA 参数估计误差以指数律衰减至零，即

$$\tilde{d}_i(t) = O(\rho_0^t), \quad \tilde{\sigma}_e^2(t) = O(\rho_0^t), \quad 0 < \rho_0 < 1 \tag{4.16.66}$$

证明 由(4.16.15)和(4.16.48)有

$$\hat{d}_i(t) = \frac{R_{re}(t, t-i)}{R_{re}(t,t)} = \frac{\boldsymbol{H\Phi}^i\boldsymbol{\Sigma}(t-i \mid t-i-1)\boldsymbol{H}^{\mathrm{T}} + d_i\sigma_e^2}{\boldsymbol{H\Phi}^i\boldsymbol{\Sigma}(t \mid t-1)\boldsymbol{H}^{\mathrm{T}} + \sigma_e^2} \tag{4.16.67}$$

$$\hat{\sigma}_e^2(t) = \boldsymbol{H\Sigma}(t \mid t-1)\boldsymbol{H}^{\mathrm{T}} + \sigma_e^2 \tag{4.16.68}$$

由(4.16.51)引出当 $t\to\infty$ 时有

$$\boldsymbol{\Sigma}(t \mid t-1) \to \boldsymbol{0} \tag{4.16.69}$$

于是由(4.16.67) ~ (4.16.69)引出当 $t\to\infty$ 时有

$$\begin{gathered}\hat{d}_i(t) \to d_i \quad (i = 1, \cdots, n) \\ \hat{\sigma}_e^2 \to \sigma_e^2\end{gathered} \tag{4.16.70}$$

这证明了算法的一致性.定义

$$\delta^{(i)}(t) = \boldsymbol{H\Phi}^i\boldsymbol{\Sigma}(t-i \mid t-i-1)\boldsymbol{H}^{\mathrm{T}}/\sigma_e^2, \quad i = 0, 1, \cdots, n \tag{4.16.71}$$

记 $\boldsymbol{\Sigma}(t \mid t-1)$ 为 $\boldsymbol{\Sigma}(t \mid t-1) = (\sigma_{ij}(t))$，其中 $\sigma_{ij}(t)$ 为其第 i 行第 j 列元素，将(4.16.19)和(4.16.49)代入(4.16.71)有

$$\begin{gathered}\delta^{(i)}(t) = \sigma_{i+1,1}(t-i)/\sigma_e^2, \quad i = 0, \cdots, n-1 \\ \delta^{(n)}(t) = 0\end{gathered} \tag{4.16.72}$$

定义矩阵范数 $\| \boldsymbol{\Sigma}(t \mid t-1) \|$ 为

$$\| \boldsymbol{\Sigma}(t \mid t-1) \| = \sum_{i=1}^{n}\sum_{j=1}^{n} | \sigma_{ij}(t) | \tag{4.16.73}$$

应用矩阵范数的等价性，存在 $\alpha > 0, \beta > 0$ 使

$$\alpha \| \boldsymbol{\Sigma}(t \mid t-1) \|_s \leqslant \| \boldsymbol{\Sigma}(t \mid t-1) \| \leqslant \beta \| \boldsymbol{\Sigma}(t \mid t-1) \|_s \tag{4.16.74}$$

应用(4.16.51)和上两式引出

$$| \sigma_{ij}(t) | \leqslant \| \boldsymbol{\Sigma}(t \mid t-1) \| \leqslant c_3\rho_0^t, \quad i, j = 1, \cdots, n \tag{4.16.75}$$

带 $c_3 = \beta c$.这意味着

$$\sigma_{ij}(t) = O(\rho_0^t), \quad i = 0, 1, \cdots, n \tag{4.16.76}$$

由(4.16.72)和(4.16.75)有

$$| \delta^{(i)}(t) | \leqslant c_{4i}\rho_0^t, \quad i = 0, 1, \cdots, n \tag{4.16.77}$$

带 $c_{4i} = c_3/\sigma_e^2\rho_0^i$.这引出

$$\delta^{(i)}(t) = O(\rho_0^t), \quad i = 0, 1, \cdots, n \tag{4.16.78}$$

由(4.16.67)，(4.16.68)和(4.16.71)有

$$\begin{gathered}\hat{d}_i(t) = \frac{\delta^{(i)}(t) + d_i}{\delta^{(0)}(t) + 1}, \\ \hat{\sigma}_e^2 = \sigma_e^2(1 + \delta^{(0)}(t))\end{gathered} \tag{4.16.79}$$

注意 $\sigma_{11}(t) \geqslant 0$，且由(4.16.72)有 $\delta^{(0)}(t) = \sigma_{11}(t)/\sigma_e^2 \geqslant 0$，于是有

$$| \tilde{d}_i(t) | = | d_i - \hat{d}_i(t) | = \frac{| d_i\delta^{(0)}(t) - \delta^{(i)}(t) |}{1 + \delta^{(0)}(t)} \leqslant | d_i | | \delta^{(0)}(t) | + | \delta^{(i)}(t) |$$

$$\tilde{\sigma}_e^2(t)=\sigma_e^2-\hat{\sigma}_e^2=-\sigma_e^2\delta^{(0)}(t) \tag{4.16.80}$$

由(4.16.77)和(4.16.78)引出

$$\tilde{d}_i(t)=O(\rho_0^t),\quad \tilde{\sigma}_e^2(t)=O(\rho_0^t) \tag{4.16.81}$$

证毕。 □

【注 4.16.1】 这里基于 Kalman 预报器的预报误差方差阵的指数收敛性(4.16.65)证明了 Gevers – Wouters 算法的指数收敛性，证明的核心思想是将 MA 模型表为等价的状态空间模型，进而将 MA 过程在 $t_0=0$ 和初值(4.16.6)下的非稳态新息 $e(t)$ 用在状态空间模型下的等价的非稳态新息 $\varepsilon(t)$ 代替，导出了时变协方差 $R_{re}(t+k,t)$ 的表达式(4.16.47)，为进一步分析收敛性提供了关键技术.

【注 4.16.2】 由(4.16.81)看到 Gevers – Wouters 算法的收敛速度由 $\rho_0=\rho^2$ 大小决定，而 ρ 是 $\overline{\boldsymbol{\Phi}}$ 的谱半径。而由(4.16.22)引出 $\overline{\boldsymbol{\Phi}}$ 的特征值即 $D(x)=0$ 的根的倒数. 因此假如 $D(x)$ 是稳定的，即 $D(x)=0$ 的根位于单位圆外，且它的根远离单位圆周，则有 $\overline{\boldsymbol{\Phi}}$ 的特征值位于单位圆内，且接近于原点附近，此时 $\overline{\boldsymbol{\Phi}}$ 的谱半径 ρ 将明显小于 1，因而可实现估值误差快速衰减至零. 这表明 Gevers – Wouters 算法收敛速度由 $D(q^{-1})$ 的零点位置决定.

【注 4.16.3】 本节结果可平行推广到多变量 MA 模型情形，有关结果将另文报道.

【注 4.16.4】 在定理 4.16.3 的证明中假设 $\overline{\boldsymbol{\Phi}}$ 有 n 个线性无关的特征向量. 在一般情况下可避免这个假设，用将 $\overline{\boldsymbol{\Phi}}$ 化为约当形的方法亦可证明定理 4.16.3 成立. 具体证明方法可参考文献[25].

参 考 文 献

1 Wiener N. THe Extrapolation, Interpolation and Smoothing of Stationary Time Series. The M.I.T. Press, Cambridge, Mass, 1949

2 Колмогоров А Н. Интерполирование и Экстраполирование Стапионарннх Случайных Последовательностей. Изв. АН СССР, сер. Матем., 1941

3 Box G E P, Jenkins G M. Time Series Analysis, Forecasting and Control. San Francisco: Holden – Day, 1970

4 Hamilton J D. Time Series Analysis. New Jersey: Princeton University Press, 1994

5 Reinsel G C. Elements of Multivariate Time Series Analysis. Springer – Verlag, New York, 1993

6 Bowerman B L, O Connell, T R. Forecasting and Time Series: An Applied Approach, Third Edition, Brooks/Cole, 1993

7 Favier G, Dubois D. A Review of k – Step – Ahead Predictors. Automatica, 1990, 26(1):75 ~ 84

8 Åström K J. 随机控制理论导论.北京:科学出版社,1983

9 Koivo H N. A Multivariable Self – Tuning Controller. Automatica, 1980, 16:351 ~ 366

10 邓自立.自校正滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2003

11 邓自立,郭一新.现代时间序列分析及其应用——建模、滤波、去卷、预报和控制.北京:知识出版社 1989

12 Jazwinski A H. Stochastic Processes and Filtering Theory. New York: Academic Press, 1970

13 Gevers M, Wouters W R E. An Innovations Approach to Discrete – Time Stochastic Realization Problem. Ouartely Journal on Automatic Control, 1978, 19(2):90 ~ 110

14 邓自立.卡尔曼滤波与维纳滤波——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2001
15 Lews F L. Optimal Estimation. New York: John Wiley & Sons, 1986
16 Anderson B D O, Moore J B. Optimal Filtering. New Jersey: Prentice Hall, 1979
17 邓自立.最优滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2000
18 韩京清,何关钰,许可康.线性系统理论代数基础.沈阳:辽宁科学技术出版社,1985
19 须田信英等.自动控制中的矩阵理论.北京:科学出版社,1979
20 Kamen E W, Su J K. Introduction to Optimal Estimation. Springer - Verlag London Berlin Heidelberg, 1999
21 邓自立.估计 MA 参数的多维强 Gevers - Wouters 算法及其在构造 ARMA 模型中的应用.控制理论与应用,2001,18(5):737 ~ 740
22 韩京清,许可康.线性控制系统理论——构造性方法.北京:科学出版社,2001
23 程云鹏.矩阵论(第二版).西安:西北工业大学出版社,2001
24 邓自立.Convergence Analysis of Gevers - Wouters Algorithm for Scalar Spectral Factorization. 科学技术与工程,2005,5(1):7 ~ 13
25 邓自立.Lyapunov 方程迭代解的指数收敛性.科学技术与工程,2005,5(9):542 ~ 546

第五章　现代时间序列分析方法及其应用

现代时间序列分析方法是由邓自立在专著《现代时间序列分析及其应用——建模、滤波、去卷、预报和控制》(北京:知识出版社,1989)中提出的.它是最优滤波的一种新的方法论.该书将现代控制理论与经典时间序列分析相结合,开拓了一门新兴边缘学科.已故中科院院士张钟俊先生曾为该书撰写序言和发表书评(张钟俊:一门新兴的边缘学科——现代时间序列分析、信息与控制,1988,17(4):62~63),给予高度评价.该书出版15年来,作者在该领域进行一系列开拓性和创造性工作,在国内外发表学术论文300余篇,其中所提出的白噪声估计理论[3]发表于自动控制理论国际权威刊物《Automatica》上.提出了关于状态和信号最优滤波的一系列新理论、新方法和新结果,显示了现代时间序列分析方法的强大的生命力.

现代时间序列分析方法的基本工具是自回归滑动平均(ARMA)新息模型[174],它的理论基础是白噪声估计理论[3].

“新息”(Innovation)概念在经典Kalman滤波方法中被用于推导Kalman滤波器,起到了重要作用.但经典Kalman滤波理论仅导出了刻划观测过程与它的新息过程之间关系的状态空间新息模型[197],而没有发现揭示观测过程与它的新息过程之间关系的时间序列ARMA新息模型.

在信号最优估计问题中,最早Hagander和Wittenmarle[160]用构造观测信号的ARMA新息模型成功地解决了带白色观测噪声的单通道ARMA信号的最优滤波器和平滑器设计问题.其中构造ARMA新息模型问题等价于一个谱分解问题.进一步,作者[161]将文献[160]的结果推广到多变量情形,问题归结为构造多变量ARMA新息模型.该文只采用辨识方法得到ARMA新息模型的参数估计,而没有解决ARMA新息模型参数的理论计算问题.直到文献[170]提出用Gevers-Wouters算法估计滑动平均(MA)参数,从而完全解决了ARMA新息模型的构造(理论计算)问题.Gevers-Wouters算法的优点是可保证MA多项式矩阵是稳定的[174],且具有快速指数收敛性.[211]这是构造ARMA新息模型的最简单的有效算法.构造ARMA新息模型本质上是一个谱分解问题[3],因此除了Gevers-Wouters算法外,其他有效算法[175]也可用于构造ARMA新息模型.本书第四章用Kalman滤波方法首次证明了Gevers-Wouters算法的指数收敛速度,为Gevers-Wouters算法的快速收敛性提供了理论根据.应指出,估计MA参数等价于解非线性方程组,其解是不惟一的.用牛顿-拉斐森迭代法[176]是不可取的,因为它不能保证所求得的MA参数使相应的MA多项式是稳定的.应注意,ARMA新息模型也可由状态空间新息模型通过变换被得到,因而也可用经典Kalman滤波方法,基于稳态Riccati方程的迭代解得到ARMA新息模型[171,177].以ARMA新息模型作为桥梁,由稳态Kalman估值器可引出等价的Wiener状态估值器,而从

构成时域 Wiener 滤波新方法[126,156].

以石油地震勘探为应用背景，Mendel[13~17]最早用经典 Kalman 滤波方法解决动态系统输入白噪声信号最优(线性最小方差)估计问题.著者[4]最早用 ARMA 新息模型解决稳态最优输入白噪声估计问题，其中提出了用 Gevers – Wouters 迭代算法构造 ARMA 新息模型.但其局限性是仅给出了输入白噪声估值器，没有给出观测白噪声估值器.在文献[2,3]中提出了基于 ARMA 新息模型的统一的和通用的白噪声估计理论.

基于 ARMA 新息模型解决 ARMA 信号最优反卷积滤波器设计问题见文献[19].

基于 ARMA 新息模型解决状态估计问题的早期工作见文献[159,170,2]，其中 ARMA 新息模型的构造具有特殊性和局限性，即利用矩阵求逆或 Fadeeva 公式导出的 ARMA 新息模型的自回归(AR)部分是标量多面式，适用于单输出单输入系统.更一般的 ARMA 新息模型导出方法应用了多项式矩阵左素分解，适用于多输入多输出系统[1,171,172].基于 ARMA 新息模型求稳态 Kalman 滤波器或预报器增益的早期工作为 Tajima[165]的稳态 Kalman 滤波器增益算法，以及文献[1,2,168,170,171,172]中给出的一些不同的新算法.但这些算法严格地说仅适用于单输入单输出系统.本章将给出对多输入多输出稳定或不稳定系统的稳态 Kalman 估值器(滤波、预报、平滑)增益的通用和统一的公式.

对于一个线性离散随机系统，文献[3]揭示了状态、观测、白噪声三者之间的关系，提出了基于白噪声估值器和观测预报器设计状态估值器的新的状态估计理论框架，在此框架下，文献[1,171,172]系统地提出了基于 ARMA 新息模型的新的状态估计理论和方法[18,30].

同经典 Kalman 滤波方法相比[173]，Kalman 滤波方法的基本工具是 Riccati 方程.对于稳态最优 Kalman 滤波，现代时间序列分析方法用构造 ARMA 新息模型取代了解稳态 Riccati 方程.在某些情形下构造 ARMA 新息模型同解 Riccati 方程相比可明显减小计算负担.在稳态最优信号估计问题中，用现代时间序列分析方法常常可更简单有效地解决问题，避免了用引入带高维状态变量的状态空间模型带来的计算上的复杂性和增加的计算负担.但经典 Kalman 滤波方法可处理时变系统或定常系统的非稳态最优滤波问题，而现代时间序列分析方法仅适用于定常系统(时不变系统)的稳态最优滤波问题.但它的优点是基于 ARMA 新息模型的递推参数估计可处理含有未知模型参数和噪声统计系统的自校正滤波问题[172].而经典 Kalman 滤波方法要求模型参数和噪声统计是完全精确已知的.

同现代 Wiener 滤波方法相比，现代 Wiener 滤波方法是一种频域方法，也称多项式方法[78~80,86~88]，它的基本工具是 Diphantine 方程.它将稳态最优滤波问题归结为解若干个互耦的 Diphantine 方程.求解过程较为复杂.对解决状态估计问题文献中报道甚少[178]，该方法的适用性有一定局限性.

专著《最优滤波理论及其应用》[1]，《卡尔曼滤波与维纳滤波》[171]，《自校正滤波理论及其应用》[172]及本书构成现代时间序列分析方法完整的理论体系，构成了独具特色的新的最优滤波理论.现代时间序列分析方法成为继 Wiener 滤波方法和 Kalman 滤波方法之后的最优滤波的一种新的方法论.

上述三种最优滤波方法是相互渗透的.白噪声估计理论是现代时间序列分析方法的理论基础，ARMA 新息模型是其基本工具.将它们渗透到经典 Kalman 滤波理论中，文献[115,119,154]进一步提出了基于经典 Kalman 滤波器的统一和通用的白噪声估计理论，发

展了 Mendel[13~17]的输入白噪声估值器，在此基础上，利用基于 Riccati 方程迭代解构造的 ARMA 新息模型，可得到一系列新理论、新方法和新结果，发展了经典 Kalman 滤波理论[126,156,131,132,134].

现代时间序列分析方法作为最优估计的一种新的方法论，产生的背景如下：

现代科学发展的重要特点是各学科、各领域、各分支之间的相互渗透，相互交叉，新兴边缘学科像雨后春笋似地相继出现.

著者将时间序列分析与现代控制理论相结合、相交叉，以时间序列 ARMA 模型与状态空间模型的相互转化作为基本出发点，提出以 ARMA 新息模型作为基本工具，以白噪声估计理论作为理论基础的新的时域方法论，开拓了传统时间序列分析与现代控制理论之间的新的边缘领域——现代时间序列分析.它从根本上更新了传统时间序列分析的研究对象、内容和方法论.传统时间序列分析只研究了没有观测噪声的 ARMA 时间序列预报问题，而现代时间序列分析主要研究带观测噪声的时间序列的最优滤波、平滑和预报问题.这些研究内容与现代控制理论中的状态估计理论（Kalman 滤波理论）有交叉，相重合，但采用了以 ARMA 新息模型作为基本工具的全新的时域方法——现代时间序列分析方法.

因为所提出的新的方法论是以 ARMA 新息模型作为基本工具，而 ARMA 模型是时间序列分析的基本模型，因此新方法论具有浓厚的时间序列分析色彩，故将新方法论取名为现代时间序列分析方法，以区别于状态空间方法（Kalman 滤波方法）及经典时间序列分析中的谱分析方法.

本章介绍现代时间序列分析方法的基本原理、理论基础及关键技术，并给出在跟踪系统中的应用.

在文献[161]和[162]中用现代时间序列分析方法得到的结果，其后被 Moir 和 Grimble[164]和 Moir[163]用频域 Wiener 滤波方法得到.文献[179~182]用现代时间序列分析方法提出了最优滤波与反卷积问题的 Diophantine 方程解，不同于用多项式方法所得到的解，避免了求解两个耦合的 Diophantine 方法，仅要求解单个 Diophantine 方程.而且用现代时间序列分析方法还发现了许多用其他方法不容易得到的新结果，例如文献[19]提出的多通道最优反卷积时域方法，文献[18]提出的广义 Wiener 状态估值器，文献[30]提出的广义 Kalman 估值器.其中都应用了 ARMA 新息模型和白噪声估值器.

现代时间序列分析方法的特色是 ARMA 新息模型和白噪声估值器.ARMA 新息模型含有最优估计所需的全部统计信息，它揭示了观测、新息、输入和观测白噪声之间的定量关系.这就是以 ARMA 新息模型作为最优滤波基本工具的理由.利用白噪声估值器允许我们将复杂的最优滤波问题归结为简单的白噪声估计问题.

本章介绍现代时间序列分析方法的基本原理，基本结果及关键技术，并给出在跟踪系统中的应用，并介绍现代时间序列分析方法的某些最新进展.例如将给出适用于不稳定系统的稳态 Kalman 估值器（滤波器、预报器、平滑器）增益公式的严格的证明和推导.将给出在典型的跟踪系统中所遇到的利用多项式矩阵左素分解方法构造正确的 ARMA 新息模型，并指出若不进行左素分解，即不消去两多项式矩阵极大左因式，则将得到错误的 ARMA 新息模型，进而将引出错误滤波结果.还将给出最优滤波、预报、平滑估计误差协方差阵的计算公式，可用于设计信息融合滤波器.[197~205]

5.1 统一的稳态最优白噪声估值器

本节介绍用现代时间序列分析方法对稳定和不稳定系统的输入白噪声和观测白噪声提出了统一的稳态最优的白噪声估值器[3,1,171].

考虑线性离散时间随机系统

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{B}(q^{-1})\boldsymbol{w}(t) + \boldsymbol{C}(q^{-1})\boldsymbol{v}(t) \tag{5.1.1}$$

其中 t 为离散时间,输出(观测)$\boldsymbol{y}(t) \in R^m$,输入白噪声 $\boldsymbol{w}(t) \in R^r$,观测白噪声 $\boldsymbol{v}(t) \in R^m$,$q^{-1}$ 为单位滞后算子,$q^{-1}\boldsymbol{y}(t) = \boldsymbol{y}(t-1)$,$\boldsymbol{A}(q^{-1})$,$\boldsymbol{B}(q^{-1})$,$\boldsymbol{C}(q^{-1})$ 为 q^{-1} 的多项式矩阵,形如

$$\boldsymbol{X}(q^{-1}) = \boldsymbol{X}_0 + \boldsymbol{X}_1 q^{-1} + \cdots + \boldsymbol{X}_{n_x} q^{-n_x} \tag{5.1.2}$$

其中 $\boldsymbol{X}_i$ 为系数阵,n_x 为阶次,规定 $\boldsymbol{X}_i = \boldsymbol{0}(i > n_x)$. 设 $\boldsymbol{A}_0 = I_m$,$\boldsymbol{C}_0 = \boldsymbol{I}_m$,$\boldsymbol{I}_m$ 为 $m \times m$ 单位阵.

【假设 1】 系统是稳定的,即 $\boldsymbol{A}(q^{-1})$ 是稳定的多项式矩阵(即 $\det\boldsymbol{A}(x)$ 的零点全位于单位圆外),或系统是不稳定的.

【假设 2】 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 是带零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$、相关阵为 $\boldsymbol{S}$ 的白噪声:

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix}[\boldsymbol{w}^{\mathrm{T}}(j), \boldsymbol{v}^{\mathrm{T}}(j)]\right\} = \begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}\delta_{tj}, \quad \begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix} > 0 \tag{5.1.3}$$

其中 E 为均值号,T 为转置号,$\delta_{tt} = 1$,$\delta_{tj} = 0(t \neq j)$.

【假设 3】 $(\boldsymbol{A}(q^{-1}), \boldsymbol{B}(q^{-1}), \boldsymbol{C}(q^{-1}))$ 左素,且 $(\boldsymbol{B}(q^{-1}), \boldsymbol{C}(q^{-1}))$ 左素或 $(\boldsymbol{B}(q^{-1}), \boldsymbol{C}(q^{-1}))$ 无行列式的零点在单位圆上的左因式.

【假设 4】 初始观测时刻 $t_0 = -\infty$,即已知无限观测历史.

稳态最优(线性最小方差)白噪声估计问题是:基于无限观测历史 $(\boldsymbol{y}(t+N), \boldsymbol{y}(t+N-1), \cdots)$ 分别求白噪声 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 的稳态最优(线性最小方差)估值器 $\hat{\boldsymbol{w}}(t \mid t+N)$ 和 $\hat{\boldsymbol{v}}(t \mid t+N)$,统一记为 $\hat{\boldsymbol{\theta}}(t \mid t+N)$,$\boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v}$,它们极小化性能指标

$$J = \mathrm{E}[(\boldsymbol{\theta}(t) - \hat{\boldsymbol{\theta}}(t \mid t+N))^{\mathrm{T}}(\boldsymbol{\theta}(t) - \hat{\boldsymbol{\theta}}(t \mid t+N))] \tag{5.1.4}$$

对 $N = 0$,$N > 0$ 或 $N < 0$,各称 $\hat{\boldsymbol{\theta}}(t \mid t+N)$,$\boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v}$ 为白噪声滤波器、平滑器或预报器.

5.1.1 ARMA 新息模型或 MA 新息模型

注意(5.1.1)右边为两个 MA 过程之和,记

$$\boldsymbol{r}(t) = \boldsymbol{B}(q^{-1})\boldsymbol{w}(t) + \boldsymbol{C}(q^{-1})\boldsymbol{v}(t) \tag{5.1.5}$$

记它的相关函数为

$$\boldsymbol{R}_r(i) = \mathrm{E}[\boldsymbol{r}(t)\boldsymbol{r}^{\mathrm{T}}(t-i)] \tag{5.1.6}$$

为了计算相关函数 $\boldsymbol{R}_r(t)$,可将(5.1.5)写成一个合成的 MA 过程

$$\boldsymbol{r}(t) = [\boldsymbol{B}(q^{-1}), \boldsymbol{C}(q^{-1})]\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix} \tag{5.1.7}$$

它可简写为

$$\boldsymbol{r}(t)=\boldsymbol{M}(q^{-1})\boldsymbol{\eta}(t) \tag{5.1.8}$$

$$\boldsymbol{M}(q^{-1})=[\boldsymbol{B}(q^{-1}),\boldsymbol{C}(q^{-1})],\quad \boldsymbol{\eta}(t)=\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix} \tag{5.1.9}$$

$$\boldsymbol{M}(q^{-1})=\boldsymbol{M}_0+\boldsymbol{M}_1q^{-1}+\cdots+\boldsymbol{M}_{n_m}q^{-n_m} \tag{5.1.10}$$

$$n_m=\max(n_b,n_c),\quad \boldsymbol{M}_i=[\boldsymbol{B}_i,\boldsymbol{C}_i] \tag{5.1.11}$$

$$\boldsymbol{B}_i=\boldsymbol{0}\quad(i>n_b),\quad \boldsymbol{C}_i=\boldsymbol{0}\quad(i>n_c) \tag{5.1.12}$$

由(5.1.3)和(5.1.8)可得 $\boldsymbol{R}_r(i)$ 为

$$\boldsymbol{R}_r(i)=\sum_{j=i}^{n_m}\boldsymbol{M}_j\boldsymbol{Q}_\eta\boldsymbol{M}_{j-i}^{\mathrm{T}},\quad i=0,1,\cdots,n_m \tag{5.1.13}$$

其中 $\boldsymbol{\eta}(t)$ 的方差阵 $\boldsymbol{Q}_\eta$ 为

$$\boldsymbol{Q}_\eta=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix} \tag{5.1.14}$$

将(5.1.11)和(5.1.14)代入(5.1.13)引出

$$\boldsymbol{R}_r(i)=\sum_{j=i}^{n_m}\boldsymbol{B}_j\boldsymbol{Q}_w\boldsymbol{B}_{j-i}^{\mathrm{T}}+\sum_{j=i}^{n_m}\boldsymbol{C}_j\boldsymbol{Q}_v\boldsymbol{C}_{j-i}^{\mathrm{T}}+\sum_{j=i}^{n_m}\boldsymbol{B}_j\boldsymbol{S}\boldsymbol{C}_{j-i}^{\mathrm{T}}+\sum_{j=i}^{n_m}\boldsymbol{C}_j\boldsymbol{S}^{\mathrm{T}}\boldsymbol{B}_{j-i}^{\mathrm{T}} \tag{5.1.15}$$

由假设 2 ~ 4，(5.1.5)右边的两个 MA 过程之和可用一个等价的稳定的 MA 过程 $\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)$ 表示为

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{B}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{C}(q^{-1})\boldsymbol{v}(t) \tag{5.1.16}$$

其中 $\boldsymbol{\varepsilon}(t)\in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，$\boldsymbol{D}_0=\boldsymbol{I}_m$，且它的阶次 n_d 为

$$n_d=\max(i\mid R_r(i)\neq 0)$$

即相关函数在 $i=n_d$ 处截尾. 由定义(5.1.5)和关系(5.1.16)引出

$$\boldsymbol{r}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.1.17}$$

问题归结为已知 MA 过程 $\boldsymbol{r}(t)$ 的相关函数 $\boldsymbol{R}_r(i)$(它由(5.1.15)计算)，如何求稳定的 $\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$. 这等价于一个如下的谱分解问题：

定义 $\boldsymbol{r}(t)$ 的谱阵为

$$\boldsymbol{S}_r(q)=\sum_{j=-n_d}^{n_d}\boldsymbol{R}_r(j)q^{-j} \tag{5.1.18}$$

容易验证 $\boldsymbol{r}(t)$ 的谱阵为

$$\boldsymbol{S}_r(q)=\boldsymbol{D}(q^{-1})\boldsymbol{Q}_\varepsilon\boldsymbol{D}^{\mathrm{T}}(q) \tag{5.1.19}$$

当 $\boldsymbol{R}_r(r)$ 已知时，由定义(5.1.18)谱阵 $\boldsymbol{S}_r(q)$ 也是已知的. 因此问题归结为已知谱阵 $\boldsymbol{S}_r(q)$ 求满足(5.1.19)的稳定的多项式矩阵 $\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$. 这叫谱分解问题[1,3,175].

谱阵 $\boldsymbol{S}_r(q)$ 可由(5.1.8)和(5.1.14)给出为

$$\boldsymbol{S}_r(q)=\boldsymbol{M}(q^{-1})\boldsymbol{Q}_\eta\boldsymbol{M}^{\mathrm{T}}(q) \tag{5.1.20}$$

应用定义(5.1.9)和(5.1.10)引出

$$\boldsymbol{S}_r(q)=\boldsymbol{B}(q^{-1})\boldsymbol{Q}_w\boldsymbol{B}(q)+\boldsymbol{C}(q^{-1})\boldsymbol{Q}_v\boldsymbol{C}^{\mathrm{T}}(q)+\boldsymbol{B}(q^{-1})\boldsymbol{S}\boldsymbol{C}^{\mathrm{T}}(q)+\boldsymbol{C}(q^{-1})\boldsymbol{S}^{\mathrm{T}}\boldsymbol{B}(q) \tag{5.1.21}$$

Gevers - Wouters[169] 算法是一种快速谱分解算法，在假设 1 ~ 4 下[1]，由 $\boldsymbol{R}_r(i)$ 用

Gevers－Wouters 算法(2.13.33)～(2.13.38)可求得使 $\boldsymbol{D}(q^{-1})$ 稳定的系数阵 $\boldsymbol{D}_i$ 和方差阵 $\boldsymbol{Q}_\varepsilon$. 于是由(5.1.1)和(5.1.15)引出观测过程 $y(t)$ 的 ARMA 模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.1.22}$$

可证明白噪声 $\boldsymbol{\varepsilon}(t)$ 是 $y(t)$ 的新息过程[1],故称(5.1.22)为 ARMA 新息模型. 由(5.1.22)和(5.1.16)看到,ARMA 新息模型揭示了观测 $\boldsymbol{y}(t)$,新息 $\boldsymbol{\varepsilon}(t)$ 及输入白噪声 $\boldsymbol{w}(t)$ 和观测白噪声 $\boldsymbol{v}(t)$ 四者之间的定量关系. 它是寻求白噪声估值器的基本工具.

由假设 1,若系统是稳定的,则(5.1.22)是平稳、可逆的 ARMA 模型,因而 $\boldsymbol{y}(t)$ 是平稳时间序列,且 $\boldsymbol{\varepsilon}(t)$ 恰好为 $\boldsymbol{y}(t)$ 新息过程,它可由 $(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$ 计算为

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Psi}_j\boldsymbol{y}(t-j) \tag{5.1.23}$$

其中由(4.8.9)系数阵 $\boldsymbol{\Psi}_j$ 可递推计算为

$$\boldsymbol{\Psi}_j=-\boldsymbol{D}_1\boldsymbol{\Psi}_{j-1}-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{\Psi}_{j-n_d}+\boldsymbol{A}_j \tag{5.1.24}$$

其中规定 $\boldsymbol{\Psi}_j=\boldsymbol{0}(j<0)$,$\boldsymbol{A}_j=\boldsymbol{0}(j>n_a)$.

在应用中,新息 $\boldsymbol{\varepsilon}(t)$ 也可由(5.1.22)任取初值 $(\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(n_d-1))$ 后简单地递推计算为

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)-\boldsymbol{D}_1\boldsymbol{\varepsilon}(t-1)-\cdots-\boldsymbol{D}_{n_d}(t-n_d) \tag{5.1.25}$$

其中 $t=n_d,n_d+1,\cdots$. 由 $\boldsymbol{D}(q^{-1})$ 的稳定性引出当 t 增大时新息初值 $(\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(n_d-1))$ 的影响将渐渐消失,即用(5.1.25)计算 $\varepsilon(t)$ 是渐近稳定的.

若系统(5.1.1)是不稳定的,即 $\boldsymbol{A}(q^{-1})$ 是不稳定的多项式矩阵,此时 $\boldsymbol{y}(t)$ 是非平稳时间序列. 当 $t_0=-\infty$ 时,对任意有限时刻 t,$\boldsymbol{y}(t)$ 有无穷大方差[1]. 为了解决此矛盾,引入新的观测过程 $\boldsymbol{z}(t)$ 为

$$\boldsymbol{z}(t)=\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) \tag{5.1.26}$$

则(5.1.1)化为

$$\boldsymbol{z}(t)=\boldsymbol{B}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{C}(q^{-1})\boldsymbol{v}(t) \tag{5.1.27}$$

由于 $z(t)$ 是两个 MA 过程之和,故 $z(t)$ 是平稳时间序列,且由(5.1.22)引出 $\boldsymbol{z}(t)$ 的 MA 新息模型

$$\boldsymbol{z}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.1.28}$$

此时 $\boldsymbol{\varepsilon}(t)$ 也是 $\boldsymbol{z}(t)$ 的新息过程,且 $\boldsymbol{\varepsilon}(t)$ 可由 $(\boldsymbol{z}(t),\boldsymbol{z}(t-1),\cdots)$ 计算为

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{z}(t)=\sum_{j=0}^{\infty}\boldsymbol{\Psi}_j\boldsymbol{z}(t-j) \tag{5.1.29}$$

其中系数阵 $\boldsymbol{\Psi}_j$ 可递推计算为

$$\boldsymbol{\Psi}_j=-\boldsymbol{D}\boldsymbol{\Psi}_{j-1}-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{\Psi}_{j-n_d} \tag{5.1.30}$$

其中规定 $\boldsymbol{\Psi}_0=\boldsymbol{I}_m$,$\boldsymbol{\Psi}=\boldsymbol{0}(j<0)$.

在应用中新息 $\boldsymbol{\varepsilon}(t)$ 可由(5.1.28)取初值 $(\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(n_d-1))$ 后递推计算为

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{z}(t)-\boldsymbol{D}_1\boldsymbol{\varepsilon}(t-1)-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{\varepsilon}(t-n_d),\quad t=n_d,n_d+1,\cdots \tag{5.1.31}$$

容易证明[1,3]:当 t 增大时 $\boldsymbol{\varepsilon}(t)$ 是渐近稳定的.

5.1.2 统一的稳态最优白噪声估值器

由(5.1.27)看到,对不稳定系统 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 仅与 $\boldsymbol{z}(t)$ 有关,因此基于 $(\boldsymbol{y}(t+N)$,

$\boldsymbol{y}(t+N-1),\cdots)$ 求最优白噪声估值器 $\hat{\boldsymbol{\theta}}(t\mid t+N)$，$\theta=w,v$，等价于基于 $(\boldsymbol{z}(t+N),\boldsymbol{z}(t+N-1),\cdots)$ 的最优估值器 $\hat{\boldsymbol{\theta}}(t\mid t+N)$. 注意对稳定系统(5.1.1)，子 Hilbert 空间有关系

$$L(\boldsymbol{y}^{\mathrm{T}}(t+N),\boldsymbol{y}^{\mathrm{T}}(t+N-1),\cdots)=L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N),\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N-1),\cdots) \tag{5.1.32}$$

而对不稳定系统，子 Hilbert 空间有关系

$$L(\boldsymbol{z}^{\mathrm{T}}(t+N),\boldsymbol{z}^{\mathrm{T}}(t+N-1),\cdots)=L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N),\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N-1),\cdots) \tag{5.1.33}$$

因此问题归结为求 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 在子 Hilbert 空间 $L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N),\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N-1),\cdots)$ 上的射影. 由射影公式(4.1.20)有

$$\hat{\boldsymbol{\theta}}(t\mid t+N)=\sum_{j=0}^{\infty}\mathrm{E}[\boldsymbol{\theta}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N-j)]\boldsymbol{Q}_{\varepsilon}^{-1}\boldsymbol{\varepsilon}(t+N-j),\quad \boldsymbol{\theta}=\boldsymbol{w},\boldsymbol{v} \tag{5.1.34}$$

由(5.1.16)和 $\boldsymbol{D}(q^{-1})$ 的稳定性有均方收敛的展式[171]

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{v}(t)=\sum_{j=0}^{\infty}\boldsymbol{F}_j\boldsymbol{w}(t-j)+\sum_{j=0}^{\infty}\boldsymbol{G}_j\boldsymbol{v}(t-j) \tag{5.1.35}$$

其中用像节 4.8 的比较系数阵法 $\boldsymbol{F}_j$ 和 $\boldsymbol{G}_j$ 可递推计算为

$$\begin{aligned}\boldsymbol{F}_j&=-\boldsymbol{D}_1\boldsymbol{F}_{j-1}-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{F}_{j-n_d}+\boldsymbol{B}_j,\\ \boldsymbol{G}_j&=-\boldsymbol{D}_1\boldsymbol{G}_{j-1}-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{G}_{j-n_d}+\boldsymbol{C}_j\end{aligned} \tag{5.1.36}$$

其中规定 $\boldsymbol{F}_j=\boldsymbol{0}(j<0)$，$\boldsymbol{G}_j=\boldsymbol{0}(j<0)$，$\boldsymbol{B}_j=\boldsymbol{0}(j>n_b)$，$\boldsymbol{C}_j=\boldsymbol{0}(j>n_c)$.

【定理 5.1.1】 系统(5.1.1)在假设 1 ~ 4 下，有关系

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]=\boldsymbol{\Lambda}_{i-t}^{w},\quad \boldsymbol{\Lambda}_i^{w}=\boldsymbol{Q}_w\boldsymbol{F}_i^{\mathrm{T}}+\boldsymbol{S}\boldsymbol{G}_i^{\mathrm{T}} \tag{5.1.37}$$

$$\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]=\boldsymbol{\Lambda}_{i-t}^{v},\quad \boldsymbol{\Lambda}_i^{v}=\boldsymbol{Q}_v\boldsymbol{G}_i^{\mathrm{T}}+\boldsymbol{S}^{\mathrm{T}}\boldsymbol{F}_i^{\mathrm{T}} \tag{5.1.38}$$

证明 应用(5.1.35)和假设 2 有

$$\begin{aligned}\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]&=\sum_{j=0}^{\infty}\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(i-j)]\boldsymbol{F}_j^{\mathrm{T}}=\sum_{j=0}^{\infty}\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{v}^{\mathrm{T}}(i-j)]\boldsymbol{G}_j^{\mathrm{T}}=\\ &\boldsymbol{Q}_w\boldsymbol{F}_{i-t}^{\mathrm{T}}+\boldsymbol{S}\boldsymbol{G}_{i-t}^{\mathrm{T}}=\boldsymbol{\Lambda}_{i-t}^{w}\end{aligned} \tag{5.1.39}$$

类似可证(5.1.38). □

【定理 5.1.2】 系统(5.1.1)在假设 1 ~ 4 下，有统一的稳态最优白噪声估值器

$$\hat{\boldsymbol{w}}(t\mid t+N)=\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{w}\boldsymbol{Q}_{\varepsilon}^{-1}\boldsymbol{\varepsilon}(t+i),\quad N\geqslant 0 \tag{5.1.40}$$

$$\hat{\boldsymbol{v}}(t\mid t+N)=\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{v}\boldsymbol{Q}_{\varepsilon}^{-1}\boldsymbol{\varepsilon}(t+i),\quad N\geqslant 0 \tag{5.1.41}$$

$$\hat{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{0},\quad N<0 \tag{5.1.42}$$

$$\hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{0},\quad N<0 \tag{5.1.43}$$

且稳态估值误差 $\tilde{\boldsymbol{\theta}}(t\mid t+N)=\boldsymbol{\theta}(t)-\hat{\boldsymbol{\theta}}(t\mid t+N)$ 的方差阵 $\boldsymbol{P}_{\theta}(t\mid t+N)=\mathrm{E}[\tilde{\boldsymbol{\theta}}(t\mid t+N)\tilde{\boldsymbol{\theta}}^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}_w(N)=\boldsymbol{Q}_w-\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{w}\boldsymbol{Q}_{\varepsilon}^{-1}\boldsymbol{\Lambda}_i^{w\mathrm{T}},\quad N\geqslant 0 \tag{5.1.44}$$

$$\boldsymbol{P}_v(N)=\boldsymbol{Q}_v-\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\Lambda}_i^{v\mathrm{T}},\quad N\geqslant 0 \tag{5.1.45}$$

$$\boldsymbol{P}_w(N)=\boldsymbol{Q}_w,\quad N<0 \tag{5.1.46}$$

$$\boldsymbol{P}_v(N)=\boldsymbol{Q}_v,\quad N<0 \tag{5.1.47}$$

证明　将(5.1.37)和(5.1.38)代入(5.1.34)引出

$$\hat{\boldsymbol{\theta}}(t\mid t+N)=\sum_{j=0}^{\infty}\boldsymbol{\Lambda}_{N-j}^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+N-j),\quad \boldsymbol{\theta}=\boldsymbol{w},\boldsymbol{v} \tag{5.1.48}$$

由 $\boldsymbol{F}_i=\boldsymbol{0}(i<0)$ 和 $\boldsymbol{G}_i=\boldsymbol{0}(i<0)$ 引出 $\boldsymbol{\Lambda}_i^{\theta}=\boldsymbol{0}(i<0)$. 于是(5.1.48)成为

$$\hat{\boldsymbol{\theta}}(t\mid t+N)=\sum_{j=0}^{N}\boldsymbol{\Lambda}_{N-j}^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+N-j)=\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i),\quad N\geqslant 0 \tag{5.1.49}$$

这证明了(5.1.40)和(5.1.41). 当 $N<0$ 时，由 $\boldsymbol{\Lambda}_{N-j}^{\theta}=\boldsymbol{0}$ 和(5.1.48)引出 $\hat{\boldsymbol{\theta}}(t\mid t+N)=\boldsymbol{0}$. 这证明了(5.1.42)和(5.1.43). 注意

$$\begin{aligned}\boldsymbol{P}_\theta(N)=&\mathrm{E}[\boldsymbol{\theta}(t)\boldsymbol{\theta}^{\mathrm{T}}(t)]-\mathrm{E}[\hat{\boldsymbol{\theta}}(t\mid t+N)\boldsymbol{\theta}^{\mathrm{T}}(t)]-\mathrm{E}[\boldsymbol{\theta}(t)\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t\mid t+N)]+\\&[\hat{\boldsymbol{\theta}}(t\mid t+N)\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t\mid t+N)],\quad N\geqslant 0\end{aligned} \tag{5.1.50}$$

将(5.1.49)代入(5.1.50)并利用(5.1.37)和(5.1.38)有

$$\begin{aligned}\boldsymbol{P}_\theta(N)=&\boldsymbol{Q}_\theta-\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\Lambda}_i^{\theta T}-\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\Lambda}_i^{\theta T}+\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\Lambda}_i^{\theta T}=\\&\boldsymbol{Q}_\theta-\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\Lambda}_,^{\theta T}\quad \boldsymbol{\theta}=\boldsymbol{w},\boldsymbol{v}\end{aligned} \tag{5.1.51}$$

这证明了(5.1.44)和(5.1.45)，将(5.1.42)和(5.1.43)代入 $\boldsymbol{P}_\theta(N)$ 中立刻得(5.1.46)和(5.1.47). □

【注 5.1.1】　(5.1.44)和(5.1.45)也可更简单地证明如下：

应用(5.1.49)引出

$$\tilde{\boldsymbol{\theta}}(t\mid t+N)=\boldsymbol{\theta}(t)-\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i) \tag{5.1.52}$$

由射影性质有 $\tilde{\boldsymbol{\theta}}(t\mid t+N)$ 正交于线性流形 $L(\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N),\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N-1),\cdots)$，注意(5.1.52)可写成

$$\boldsymbol{\theta}(t)=\tilde{\boldsymbol{\theta}}(t\mid t+N)-\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i) \tag{5.1.53}$$

由此立刻引出

$$\boldsymbol{Q}_\theta=\boldsymbol{P}_\theta(N)+\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\Lambda}_i^{\theta T},\quad N\geqslant 0 \tag{5.1.54}$$

这引出(5.1.44)和(5.1.45).

【定理 5.1.3】　系统(5.1.1)在假设 1 ~ 4 下，由(5.1.9)定义的增广白噪声 $\boldsymbol{\eta}(t)$ 的稳态最优估值器为

$$\hat{\boldsymbol{\eta}}(t\mid t+N)=\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\eta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i),\quad N\geqslant 0 \tag{5.1.55}$$

$$\hat{\boldsymbol{\eta}}(t\mid t+N)=\boldsymbol{0},\quad N<0 \tag{5.1.56}$$

其中定义

$$\boldsymbol{\Lambda}_i^{\eta} = \boldsymbol{Q}_{\eta}\boldsymbol{L}_i^{\mathrm{T}} \tag{5.1.57}$$

而 $\boldsymbol{Q}_{\eta}$ 由(5.1.14)定义,系数阵 $\boldsymbol{L}_i$ 可递推计算为

$$\boldsymbol{L}_j = -\boldsymbol{D}_1\boldsymbol{L}_{j-1} - \cdots - \boldsymbol{D}_{n_d}\boldsymbol{L}_{j-m_d} + \boldsymbol{M}_j \tag{5.1.58}$$

而 $\boldsymbol{M}_i$ 由(5.1.11)定义.稳态估值误差方差阵 $\boldsymbol{P}_{\eta}(N) = \mathrm{E}[(\boldsymbol{\eta}(t) - \hat{\boldsymbol{\eta}}(t \mid t+N))(\boldsymbol{\eta}(t) - \hat{\boldsymbol{\eta}}(t \mid t+N)^{\mathrm{T}}]$ 为

$$\boldsymbol{P}_{\eta}(N) = \boldsymbol{Q}_{\eta} - \sum_{i=0}^{N}\boldsymbol{\Lambda}_i^{\eta}\boldsymbol{Q}_{\varepsilon}^{-1}\boldsymbol{\Lambda}_i^{\eta T}, \quad N \geqslant 0 \tag{5.1.59}$$

$$\boldsymbol{P}_{\eta}(N) = \boldsymbol{Q}_{\eta}, \quad N < 0 \tag{5.1.60}$$

证明　由(5.1.8)、(5.1.16)和(5.1.17)有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \boldsymbol{M}(q^{-1})\boldsymbol{\eta}(t) \tag{5.1.61}$$

于是 $\boldsymbol{\varepsilon}(t)$ 可展可

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{D}^{-1}(q^{-1})\boldsymbol{M}(q^{-1})\boldsymbol{\eta}(t) = \sum_{j=0}^{\infty}\boldsymbol{L}_j\boldsymbol{\eta}(t-j) \tag{5.1.62}$$

其中 $\boldsymbol{L}_j$ 由(5.1.58)给出.应用射影公式(4.1.20)有

$$\hat{\boldsymbol{\eta}}(t \mid t+N) = \sum_{j=0}^{\infty}\mathrm{E}[\boldsymbol{\eta}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+N-j)]\boldsymbol{Q}_{\varepsilon}^{-1}\boldsymbol{\varepsilon}(t+N-j) \tag{5.1.63}$$

$$\hat{\boldsymbol{\eta}}(t \mid t+N) = \sum_{i=0}^{N}\boldsymbol{Q}_{\eta}\boldsymbol{L}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}\boldsymbol{\varepsilon}(t+i) \tag{5.1.64}$$

由定义(5.1.57)和(5.1.63)得(5.1.55).类似于注 5.1.1 的推导可得(5.1.59)和(5.1.60). □

【定理 5.1.4】　定理 5.1.3 与定理 5.1.2 是等价的,即由定理 5.1.3 的结果可推出定理 5.1.2 的结果,反之亦然.

证明　现在我们证明由定理5.1.3可引出定理5.1.2的结果.注意 $\boldsymbol{\eta}(t)$ 的定义(5.1.9)和射影性质有

$$\hat{\boldsymbol{\eta}}(t \mid t+N) = \begin{bmatrix} \hat{\boldsymbol{w}}(t \mid t+N) \\ \hat{\boldsymbol{v}}(t \mid t+N) \end{bmatrix} \tag{5.1.65}$$

由(5.1.58)和 $\boldsymbol{M}_j$ 的定义(5.1.11)引出

$$\boldsymbol{L}_j = [\boldsymbol{F}_j, \boldsymbol{G}_j] \tag{5.1.66}$$

从而(5.1.36)成立.由(5.1.14),(5.1.55)和(5.1.57)有

$$\hat{\boldsymbol{\eta}}(t \mid t+N) = \sum_{i=0}^{N}\begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v \end{bmatrix}\begin{bmatrix} \boldsymbol{F}_j^{\mathrm{T}} \\ \boldsymbol{G}_j^{\mathrm{T}} \end{bmatrix}\boldsymbol{Q}_{\varepsilon}^{-1}\boldsymbol{\varepsilon}(t+i), \quad N \geqslant 0 \tag{5.1.67}$$

应用定义(5.1.37)和(5.1.38)及关系(5.1.65)立刻得(5.1.40)和(5.1.41).由(5.1.56)和(5.1.65)立刻引出(5.1.42)和(5.1.43).

定义互协方差阵

$$\boldsymbol{P}_{wv}(N) = \mathrm{E}[(\boldsymbol{w}(t) - \hat{\boldsymbol{w}}(t \mid t+N))(\boldsymbol{v}(t) - \hat{\boldsymbol{v}}(t \mid t+N)^{\mathrm{T}}] \tag{5.1.68}$$

由(5.1.57)和(5.1.59)有分块表达式

$$\boldsymbol{P}_{\eta}(N) = \begin{bmatrix} \boldsymbol{P}_w(N) & \boldsymbol{P}_{wv}(N) \\ \boldsymbol{P}_{wv}^{\mathrm{T}}(N) & \boldsymbol{P}_v(N) \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v \end{bmatrix} -$$

$$\sum_{i=0}^{N}\begin{bmatrix} Q_w & S \\ S^{\mathrm{T}} & Q_v \end{bmatrix}\begin{bmatrix} F_j^{\mathrm{T}} \\ G_j^{\mathrm{T}} \end{bmatrix} Q_{\varepsilon}^{-1}[F_i \quad G_i]\begin{bmatrix} Q_w & S \\ S^{\mathrm{T}} & Q_v \end{bmatrix} \tag{5.1.69}$$

应用定义(5.1.37)和(5.1.38),这引出(5.1.44)和(5.1.45).应用(5.1.60)方刻得(5.1.46)和(5.1.47).反之,由定理5.1.2亦可推出定理5.1.3. □

【推论5.1.1】 若 $B_0 = 0, C_0 = I_m$,则有稳态最优白噪声滤波器

$$\hat{w}(t \mid t) = SQ_{\varepsilon}^{-1}\varepsilon(t) \tag{5.1.70}$$

$$\hat{v}(t \mid t) = Q_v Q_{\varepsilon}^{-1}\varepsilon(t) \tag{5.1.71}$$

证明 由 $B_0 = 0$ 引出 $F_0 = 0$,注意 $G_0 = I_m$,取 $N = 0$ 由定理5.1.1立刻得证. □

【推论5.1.2】 若 $w(t)$ 与 $v(t)$ 不相关,即 $S = 0$,则有稳态最优白噪声滤波器和平滑器

$$\hat{w}(t \mid t+N) = \sum_{i=0}^{N} Q_w F_i^{\mathrm{T}} Q_{\varepsilon}^{-1}\varepsilon(t+i), \quad N \geqslant 0 \tag{5.1.72}$$

$$\hat{v}(t \mid t+N) = \sum_{i=0}^{N} Q_v G_i^{\mathrm{T}} Q_{\varepsilon}^{-1}\varepsilon(t+i), \quad N \geqslant 0 \tag{5.1.73}$$

相应的稳态误差方差阵为

$$P_w(N) = Q_w - \sum_{i=0}^{N} Q_w F_i^{\mathrm{T}} Q_{\varepsilon}^{-1} F_i Q_w, \quad N \geqslant 0 \tag{5.1.74}$$

$$P_v(N) = Q_v - \sum_{i=0}^{N} Q_v G_i^{\mathrm{T}} Q_{\varepsilon}^{-1} G_i Q_v, \quad N \geqslant 0 \tag{5.1.75}$$

【推论5.1.3】 系统(5.1.1)在假设1 ~ 4下有稳态最优递推白噪声估值器为($N > 0$)

$$\hat{w}(t \mid t+N) = \hat{w}(t \mid t+N-1) + \Lambda_N^w Q_{\varepsilon}^{-1}\varepsilon(t+N) \tag{5.1.76}$$

$$\hat{v}(t \mid t+N) = \hat{v}(t \mid t+N-1) + \Lambda_N^v Q_{\varepsilon}^{-1}\varepsilon(t+N) \tag{5.1.77}$$

其中初值分别为 $\hat{w}(t \mid t)$ 和 $\hat{v}(t \mid t)$,$N = 1,2,\cdots$,且有递推误差方差阵

$$P_w(N) = P_w(N-1) - \Lambda_N^w Q_{\varepsilon}^{-1}\Lambda_N^{w\mathrm{T}} \tag{5.1.78}$$

$$P_v(N) = P_v(N-1) - \Lambda_N^v Q_{\varepsilon}^{-1}\Lambda_N^{v\mathrm{T}} \tag{5.1.79}$$

分别带初值 $P_w(0)$ 和 $P_v(0)$,且

$$P_w(0) = Q_w - \Lambda_0^w Q_{\varepsilon}^{-1}\Lambda_0^{w\mathrm{T}} \tag{5.1.80}$$

$$P_v(0) = Q_v - \Lambda_0^v Q_{\varepsilon}^{-1}\Lambda_0^{v\mathrm{T}} \tag{5.1.81}$$

其中 $G_0 = I_m, F_0 = B_0$.特别当 $S = 0$ 时有递推白噪声估值器($N > 0$)为

$$\hat{w}(t \mid t+N) = \hat{w}(t \mid t+N-1) + Q^w F_i^{\mathrm{T}} Q_{\varepsilon}^{-1}\varepsilon(t+N) \tag{5.1.82}$$

$$\hat{v}(t \mid t+N) = \hat{v}(t \mid t+N-1) + Q_v G_i^{\mathrm{T}} Q_{\varepsilon}^{-1}\varepsilon(t+N) \tag{5.1.83}$$

相应的估值误差方差阵各为

$$P_w(N) = P_w(N-1) - Q_w F_N^{\mathrm{T}} Q_{\varepsilon}^{-1} F_N Q_w \tag{5.1.84}$$

$$P_v(N) = P_v(N-1) - Q_v G_N^{\mathrm{T}} Q_{\varepsilon}^{-1} G_N Q_v \tag{5.1.85}$$

分别带初值 $P_w(0)$ 和 $P_v(0)$,且

$$P_w(0) = Q_w - Q^w F_0^{\mathrm{T}} Q_{\varepsilon}^{-1} F_0 Q_w \tag{5.1.86}$$

$$P_v(0) = Q_v - Q_v Q_{\varepsilon}^{-1} Q_v \tag{5.1.87}$$

【推论 5.1.4】 对单输入单输出系统(5.1.1)在假设 1 ~ 4 下,分别记 $\boldsymbol{F}_i$, $\boldsymbol{G}_i$, $\boldsymbol{Q}_w$, $\boldsymbol{Q}_v$, $\boldsymbol{S}$, $\boldsymbol{Q}_\varepsilon$ 为 f_i, g_i, σ_w^2, σ_v^2, s, σ_ε^2,则稳态最优白噪声估值器为($N \geqslant 0$)

$$\hat{w}(t \mid t+N) = \frac{\sigma_w^2}{\sigma_\varepsilon^2}\sum_{i=0}^{N} f_i\varepsilon(t+i) + \frac{s}{\sigma_\varepsilon^2}\sum_{i=0}^{N} g_i\varepsilon(t+i) \tag{5.1.88}$$

$$\hat{v}(t \mid t+N) = \frac{\sigma_v^2}{\sigma_\varepsilon^2}\sum_{i=0}^{N} g_i\varepsilon(t+i) + \frac{s}{\sigma_\varepsilon^2}\sum_{i=0}^{N} f_i\varepsilon(t+i) \tag{5.1.89}$$

特别若 $f_0 = 0(b_0 = 0)$, $c_0 = 1$,则有

$$\hat{w}(t \mid t) = \frac{s}{\sigma_\varepsilon^2}\varepsilon(t), \quad \hat{v}(t \mid t) = \frac{\sigma_v^2}{\sigma_\varepsilon^2}e(t) \tag{5.1.90}$$

且误差方差分别为

$$P_w(N) = \sigma_w^2 - \frac{1}{\sigma_\varepsilon^2}\sum_{i=0}^{N}(\sigma_w^2 f_i + s g_i)^2 \tag{5.1.91}$$

$$P_v(N) = \sigma_v^2 - \frac{1}{\sigma_\varepsilon^2}\sum_{i=0}^{N}(\sigma_v^2 g_i + s f_i)^2 \tag{5.1.92}$$

$$P_w(0) = \sigma_w^2 - \frac{s^2}{\sigma_\varepsilon^2}, \quad P_v(0) = \sigma_v^2 - \left(\frac{\sigma_v^2}{\sigma_\varepsilon^2}\right) \tag{5.1.93}$$

特别若 $w(t)$ 与 $v(t)$ 不相关,即 $s = 0$,则有统一的稳态最优白噪声估值器($N \geqslant 0$)为

$$\hat{w}(t \mid t+N) = \frac{\sigma_w^2}{\sigma_\varepsilon^2}\sum_{i=0}^{N} f_i\varepsilon(t+i) \tag{5.1.94}$$

$$\hat{v}(t \mid t+N) = \frac{\sigma_v^2}{\sigma_\varepsilon^2}\sum_{i=0}^{N} g_i\varepsilon(t+i) \tag{5.1.95}$$

$$\hat{w}(t \mid t) = 0, \quad \hat{v}(t \mid t) = \left(\frac{\sigma_v^2}{\sigma_\varepsilon^2}\right)\varepsilon(t) \tag{5.1.96}$$

相应的估值误差方差为

$$P_w(N) = \sigma_w^2\left[1 - \frac{\sigma_w^2}{\sigma_\varepsilon^2}\sum_{i=0}^{N} f_i^2\right] \tag{5.1.97}$$

$$P_v(N) = \sigma_v^2\left[1 - \frac{\sigma_v^2}{\sigma_\varepsilon^2}\sum_{i=0}^{N} g_i^2\right] \tag{5.1.98}$$

$$P_w(0) = \sigma_w^2, \quad P_v(0) = \sigma_v^2 - \left(\frac{\sigma_v^2}{\sigma_\varepsilon^2}\right) \tag{5.1.99}$$

且有平滑估值误差方差最大下界为

$$P_w(\infty) = \sigma_w^2\left[1 - \frac{\sigma_w^2}{\sigma_\varepsilon^2}\sum_{i=0}^{\infty} f_i^2\right] \tag{5.1.100}$$

$$P_v(\infty) = \sigma_v^2\left[1 - \frac{\sigma_v^2}{\sigma_\varepsilon^2}\sum_{i=0}^{\infty} g_i^2\right] \tag{5.1.101}$$

且有关系

$$P_w(\infty) < P_w(N), \quad P_v(\infty) < P_v(N) \tag{5.1.102}$$

这表明当 N 增加时,平滑估值器的精度不能无限制地改善.

【注 5.1.2】 考虑状态空间模型

$$x(t+1)=\Phi x(t)+\Gamma w(t) \tag{5.1.103}$$

$$y(t)=Hx(t)+v(t) \tag{5.1.104}$$

其中状态 $\boldsymbol{x}(t)\in R^n$，观测 $\boldsymbol{y}(t)\in R^m$，$\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}(t)\in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$，相关阵为 $\boldsymbol{S}$ 的白噪声. 问题是基于观测 $(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots)$ 求稳态最优白噪声估值器 $\hat{\boldsymbol{w}}(t\mid t+N)$ 和 $\hat{\boldsymbol{v}}(t\mid t+N)$.

该状态空间模型可写为

$$y(t)=H(I_n-q^{-1}\Phi)^{-1}\Gamma q^{-1}w(t)+v(t) \tag{5.1.105}$$

其中 q^{-1} 为单位滞后算子. 引入左素分解

$$H(I_n-q^{-1}\Phi)^{-1}\Gamma q^{-1}=A^{-1}(q^{-1})B(q^{-1}) \tag{5.1.106}$$

其中 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{B}(q^{-1})$ 为 q^{-1} 的多项式矩阵，$\boldsymbol{A}_0=\boldsymbol{I}_m,\boldsymbol{B}_0=\boldsymbol{0}$. 将上式代入(5.1.105)有动态系统(5.1.1)形式的动态系统

$$A(q^{-1})y(t)=B(q^{-1})w(t)+A(q^{-1})v(t) \tag{5.1.107}$$

于是进一步可用本节方法求白噪声估值器.

【注 5.1.3】 考虑带控项 $\boldsymbol{u}(t)\in R^p$ 的线性随机控制系统

$$x(t+1)=\Phi x(t)+Bu(t)+\Gamma w(t) \tag{5.1.108}$$

$$y(t)=Hx(t)+v(t) \tag{5.1.109}$$

其中 $\boldsymbol{\Phi},\boldsymbol{B},\boldsymbol{\Gamma},\boldsymbol{H}$ 为常阵，其他记号意义及假设同注 5.1.2. 问题是求白噪声估值器 $\hat{\boldsymbol{w}}(t\mid t+N)$ 和 $\hat{\boldsymbol{v}}(t\mid t+N)$. 该状态空间模型可写为

$$\begin{aligned} y(t)&=H(I_n-q^{-1}\Phi)^{-1}[Bq^{-1}u(t)+\Gamma q^{-1}w(t)]+v(t)= \\ &[H(I_n-q^{-1}\Phi)^{-1}Bq^{-1},H(I_n-q^{-1}\Phi)^{-1}\Gamma q^{-1}]\begin{bmatrix}u(t)\\ w(t)\end{bmatrix}+v(t)\end{aligned} \tag{5.1.110}$$

引入左素分解

$$[H(I_n-q^{-1}\Phi)^{-1}Bq^{-1},H(I_n-q^{-1}\Phi)^{-1}\Gamma q^{-1}]=A^{-1}(q^{-1})[R(q^{-1}),P(q^{-1})] \tag{5.1.111}$$

其中 $\boldsymbol{R}_0=0,\boldsymbol{P}_0=0$. 将它代入(5.1.110)引出动态系统

$$A(q^{-1})y(t)=R(q^{-1})u(t)+P(q^{-1})w(t)+A(q^{-1})v(t) \tag{5.1.112}$$

引入新的观测过程 $\boldsymbol{z}(t)$ 为

$$z(t)=A(q^{-1})y(t)-R(q^{-1})u(t) \tag{5.1.113}$$

形如(5.1.1)的动态系统

$$z(t)=P(q^{-1})w(t)+A(q^{-1})v(t) \tag{5.1.114}$$

则有 MA 新息模型

$$z(t)=D(q^{-1})\varepsilon(t) \tag{5.1.115}$$

其中新息 $\boldsymbol{\varepsilon}(t)$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，且有关系

$$D(q^{-1})\varepsilon(t)=P(q^{-1})w(t)+A(q^{-1})v(t) \tag{5.1.116}$$

由此出发用本节方法可求白噪声估值器 $\hat{\boldsymbol{w}}(t\mid t+N)$ 和 $\hat{\boldsymbol{v}}(t\mid t+N)$.

【注 5.1.4】 考虑如下动态系统的白噪声估计问题：

$$y(t)=\frac{B(q^{-1})}{A(q^{-1})}w(t)+\frac{Q(q^{-1})}{P(q^{-1})}v(t)+\xi(t) \tag{5.1.117}$$

其中 $y(t)$ 是标量观测(输出),$w(t)$,$v(t)$ 和 $\xi(t)$ 是标量白噪声,$A(q^{-1}),\cdots,Q(q^{-1})$ 是标量多项式,且首系数 $a_0=1,p_0=1$.问题是由观测$(y(t+N),y(t+N+1),\cdots)$求白噪声估值器 $\hat{w}(t\mid t+N)$,$\hat{v}(t\mid t+N)$ 和 $\hat{\xi}(t\mid t+N)$.

将(5.1.117)两边乘以 $A(q^{-1})P(q^{-1})$ 有动态系统

$$A(q^{-1})P(q^{-1})y(t)=P(q^{-1})B(q^{-1})w(t)+A(q^{-1})Q(q^{-1})v(t)+A(q^{-1})P(q^{-1})\xi(t) \tag{5.1.118}$$

假设多项式$(P(q^{-1})B(q^{-1}),A(q^{-1})Q(q^{-1}),A(q^{-1})P(q^{-1}))$互质,则上式右边三个 MA 过程之和可用一个等价的可逆的 MA 过程表示[1]

$$D(q^{-1})\varepsilon(t)=P(q^{-1})B(q^{-1})w(t)+A(q^{-1})Q(q^{-1})v(t)+A(q^{-1})P(q^{-1})\xi(t) \tag{5.1.119}$$

其中 $D(q^{-1})$ 是稳定的多项式,首系数 $d_0=1$,$\varepsilon(t)$ 是零均值、方差为 σ_ε^2 的白噪声. $D(q^{-1})$ 和 σ_ε^2 可用 Gevers – Wouters 算法求得.于是可得 ARMA 新息模型

$$A(q^{-1})P(q^{-1})y(t)=D(q^{-1})\varepsilon(t) \tag{5.1.120}$$

从(5.1.119)出发用类似于本节的方法可求白噪声估值器 $\hat{\theta}(t\mid t+N)$,$\theta=w,v,\xi$.详细推导留给读者.

【定理 5.1.5】 考虑动态系统

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{B}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{C}(q^{-1})\boldsymbol{v}(t) \tag{5.1.121}$$

其中观测 $\boldsymbol{y}(t)\in R^m$,输入白噪声 $\boldsymbol{w}(t)\in R^r$ 和观测白噪声 $\boldsymbol{v}(t)\in R^m$ 有非零均值

$$\mathrm{E}\boldsymbol{w}(t)=\boldsymbol{q}_w,\quad \mathrm{E}\boldsymbol{v}(t)=\boldsymbol{q}_v \tag{5.1.122}$$

且有协方差阵

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)-\boldsymbol{q}_w\\ \boldsymbol{v}(t)-\boldsymbol{q}_v\end{bmatrix}[\boldsymbol{w}^{\mathrm{T}}(j)-\boldsymbol{q}_w^{\mathrm{T}},\boldsymbol{v}^{\mathrm{T}}(j)-\boldsymbol{q}_v^{\mathrm{T}}]\right\}=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix} \tag{5.1.123}$$

q^{-1} 为单位滞后算子,$\boldsymbol{A}(q^{-1}),\boldsymbol{B}(q^{-1}),\boldsymbol{C}(q^{-1})$ 有形式 $\boldsymbol{X}(q^{-1})=\boldsymbol{X}_0+\boldsymbol{X}_1q^{-1}+\cdots+\boldsymbol{X}_{n_x}q^{-n_x}$,$\boldsymbol{A}_0=\boldsymbol{I}_m$,$\boldsymbol{C}_0=\boldsymbol{I}_m$.设$(\boldsymbol{A}(q^{-1}),\boldsymbol{B}(q^{-1}),\boldsymbol{C}(q^{-1}))$左素,且

$$\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}>0 \tag{5.1.124}$$

则用 Gevers – Wouters 算法有 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{\rho}+\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.1.125}$$

其中 $\boldsymbol{D}(q^{-1})$ 是稳定的,$\boldsymbol{D}_0=\boldsymbol{I}_m$,$\boldsymbol{\varepsilon}(t)\in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声,且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{B}(q^{-1})\boldsymbol{w}_c(t)+\boldsymbol{C}(q^{-1})\boldsymbol{v}_c(t) \tag{5.1.126}$$

$$\boldsymbol{w}_c(t)=\boldsymbol{w}(t)-\boldsymbol{q}_w,\quad \boldsymbol{v}_c(t)=\boldsymbol{v}(t)-\boldsymbol{q}_v,$$

$$\boldsymbol{\rho}=\boldsymbol{B}(1)\boldsymbol{q}_w+\boldsymbol{C}(1)\boldsymbol{q}_0 \tag{5.1.127}$$

稳态最优白噪声估值器$(N\geqslant 0)$为

$$\hat{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{q}_w+\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^w\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i) \tag{5.1.128}$$

$$\hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{q}_v+\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i) \tag{5.1.129}$$

其中 $\boldsymbol{\Lambda}_i^w,\boldsymbol{\Lambda}_i^v,\boldsymbol{F}_i,\boldsymbol{G}_i$,用公式(5.1.36),(5.1.37)和(5.1.38)计算.新息可由(5.1.125)置

$z(t)=A(q^{-1})y(t)$ 计算为

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{z}(t)-\boldsymbol{D}^{-1}(1)\boldsymbol{\rho}=\sum_{j=0}^{\infty}\boldsymbol{\Psi}_j\boldsymbol{z}(t-j)-\boldsymbol{D}^{-1}(1)\boldsymbol{\rho} \tag{5.1.130}$$

其中 $\boldsymbol{\Psi}_j$ 由(5.1.30)递推计算.在应用中 $\boldsymbol{\varepsilon}(t)$ 由(5.1.125)取初值($\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(n_d-1)$)后递推计算为

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{z}(t)-\boldsymbol{\rho}-\boldsymbol{D}_1\boldsymbol{\varepsilon}(t-1)-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{\varepsilon}(t-n_d),\quad t=n_d,n_d+1,\cdots \tag{5.1.131}$$

相应的估值误差方差阵各为

$$\boldsymbol{P}_w(N)=\boldsymbol{Q}_w-\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^w\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\Lambda}_i^{w\mathrm{T}},\quad N\geqslant 0 \tag{5.1.132}$$

$$\boldsymbol{P}_v(N)=\boldsymbol{Q}_v-\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\Lambda}_i^{v\mathrm{T}},\quad N\geqslant 0 \tag{5.1.133}$$

当 $N<0$ 时有

$$\hat{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{q}_w,\quad \hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{q}_v \tag{5.1.134}$$

$$\boldsymbol{P}_w(N)=\boldsymbol{Q}_w,\quad \boldsymbol{P}_v(N)=\boldsymbol{Q}_v,\quad N<0 \tag{5.1.135}$$

证明　类似于定理 5.1.2 的证明,从略. □

5.2　白噪声新息滤波器与 Wiener 滤波器

【定义 5.2.1】　一种估值器称为新息滤波器,假如它可表为以新息过程作为输入的传递函数形式.一种估值器称为 Wiener 滤波器,假如它可表为以观测信号作为输入的传递函数形式.

【定义 5.2.2】　称 Wiener 滤波器是渐近稳定的,假如它的值渐近地与滤波器初值选取无关,或它的值渐近地不依赖于滤波器初值的选取.

【定理 5.2.1】　系统(5.1.1)在假设 1 ~ 4 下有新息滤波器

$$\hat{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{L}_N^w(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.2.1}$$

$$\hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{L}_N^v(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.2.2}$$

其中新息过程 $\boldsymbol{\varepsilon}(t)$ 由(5.1.22)定义,且由(5.1.40)或(5.1.41)计算.定义传递函数阵 $\boldsymbol{L}_N^w(q^{-1})$ 和 $\boldsymbol{L}_N^v(q^{-1})$ 为

$$\boldsymbol{L}_N^w(q^{-1})=\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^w\boldsymbol{Q}_\varepsilon^{-1}q^{i-N},\quad N\geqslant 0 \tag{5.2.3}$$

$$\boldsymbol{L}_N^v(q^{-1})=\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^v\boldsymbol{Q}_\varepsilon^{-1}q^{i-N},\quad N\geqslant 0 \tag{5.2.4}$$

$$\boldsymbol{L}_N^w(q^{-1})=\boldsymbol{0}\quad(N<0) \tag{5.2.5}$$

$$\boldsymbol{L}_N^v(q^{-1})=\boldsymbol{0}\quad(N<0) \tag{5.2.6}$$

证明　引入单位滞后算子 q^{-1},$q^{-1}\boldsymbol{\varepsilon}(t)=\boldsymbol{\varepsilon}(t-1)$,并引入单位前进算子 q,$q\boldsymbol{\varepsilon}(t)=\boldsymbol{\varepsilon}(t+1)$.由(5.1.40)和(5.1.41)有

$$\hat{\boldsymbol{\theta}}(t\mid t+N)=\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^\theta\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i)=\sum_{i=0}^{N}\boldsymbol{\Lambda}_i^\theta\boldsymbol{Q}_\varepsilon^{-1}q^i\boldsymbol{\varepsilon}(t)=$$

$$\sum_{i=0}^{N} \boldsymbol{\Lambda}_i^{\theta} \boldsymbol{Q}_{\varepsilon}^{-1} q^{i-N} \boldsymbol{\varepsilon}(t+N) = \boldsymbol{L}_N^{\theta}(q^{-1}) \boldsymbol{\varepsilon}(t+N) \tag{5.2.7}$$

其中 $\boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v}$. 由(5.1.42)和(5.1.43)引出(5.2.5)和(5.2.6). □

【定理 5.2.2】 系统(5.1.1)在假设 1 ~ 4 下,有渐近稳定的最优白噪声 Wiener 滤波器($N \geqslant 0$)为

$$\det \boldsymbol{D}(q^{-1}) \hat{\boldsymbol{w}}(t \mid t+N) = \boldsymbol{L}_N^{w}(q^{-1}) \operatorname{adj} \boldsymbol{D}(q^{-1}) \boldsymbol{A}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.2.8}$$

$$\det \boldsymbol{D}(q^{-1}) \hat{\boldsymbol{v}}(t \mid t+N) = \boldsymbol{L}_N^{v}(q^{-1}) \operatorname{adj} \boldsymbol{D}(q^{-1}) \boldsymbol{A}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.2.9}$$

其中“det”表示矩阵的行列式,“adj”表示伴随阵.或等价地有

$$\hat{\boldsymbol{w}}(t \mid t+N) = \boldsymbol{L}_N^{w}(q^{-1}) \boldsymbol{D}^{-1}(q^{-1}) \boldsymbol{A}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.2.10}$$

$$\hat{\boldsymbol{v}}(t \mid t+N) = \boldsymbol{L}_N^{v}(q^{-1}) \boldsymbol{D}^{-1}(q^{-1}) \boldsymbol{A}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.2.11}$$

证明 由(5.1.22)有

$$\boldsymbol{\varepsilon}(t+N) = \boldsymbol{D}^{-1}(q^{-1}) \boldsymbol{A}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.2.12}$$

将它代入(5.2.1)和(5.2.2)得(5.2.10)和(5.2.11). 再将 $\boldsymbol{D}^{-1}(q^{-1}) = \operatorname{adj}\boldsymbol{D}(q^{-1}) / \det \boldsymbol{D}(q^{-1})$ 代入(5.2.10)和(5.2.11)得(5.2.8)和(5.2.9).由 $\boldsymbol{D}(q^{-1})$ 的稳定性引出 $\det \boldsymbol{D}(q^{-1})$ 是一个稳定的多项式(即 $\det \boldsymbol{D}(x)$ 的零点全位于单位圆外),这引出(5.2.8)和(5.2.9)是渐近稳定的. □

【推论 5.2.1】 对单输入单输出系统($m = 1, r = 1$)(5.1.1)在假设 1 ~ 4 下,有渐近稳定的白噪声 Wiener 滤波器($N \geqslant 0$)为

$$D(q^{-1}) \hat{w}(t \mid t+N) = L_N^{w}(q^{-1}) A(q^{-1}) y(t+N) \tag{5.2.13}$$

$$D(q^{-1}) \hat{v}(t \mid t+N) = L_N^{v}(q^{-1}) A(q^{-1}) y(t+N) \tag{5.2.14}$$

证明 由(5.2.10)和(5.2.11)两边乘 $D(q^{-1})$ 得证. □

【定理 5.2.3】 系统(5.1.1)在假设 1 ~ 4 下有渐近稳定的白噪声 Wiener 滤波器($N \geqslant 0$)为

$$\hat{\boldsymbol{w}}(t \mid t+N) = \boldsymbol{L}_N^{w}(q^{-1}) \widetilde{\boldsymbol{A}}(q^{-1}) \widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.2.15}$$

$$\hat{\boldsymbol{v}}(t \mid t+N) = \boldsymbol{L}_N^{v}(q^{-1}) \widetilde{\boldsymbol{A}}(q^{-1}) \widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.2.16}$$

其中 $\widetilde{\boldsymbol{A}}(q^{-1})$ 和 $\widetilde{\boldsymbol{D}}^{-1}(q^{-1})$ 由如下伪交换(右素分解)决定

$$\boldsymbol{D}^{-1}(q^{-1}) \boldsymbol{A}(q^{-1}) = \widetilde{\boldsymbol{A}}(q^{-1}) \widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{5.2.17}$$

带 $\det \widetilde{\boldsymbol{D}}(q^{-1}) = \det \boldsymbol{D}(q^{-1}), \widetilde{\boldsymbol{D}}_0 = I_m, \widetilde{\boldsymbol{A}}_0 = \boldsymbol{I}_m$.

证明 将(5.2.17)代入(5.2.10)和(5.2.11)得(5.2.15)和(5.2.16).由 $\boldsymbol{D}(q^{-1})$ 是稳定的,引出 $\widetilde{\boldsymbol{D}}(q^{-1})$ 也是稳定的,从而(5.2.15)和(5.2.16)是渐近稳定的. □

【推论 5.2.2】 系统(5.1.1)在假设 1 ~ 4 下,有渐近稳定的白噪声 Wiener 滤波器($N \geqslant 0$)为

$$\det \widetilde{\boldsymbol{D}}(q^{-1}) \hat{\boldsymbol{w}}(t \mid t+N) = \boldsymbol{L}_N^{w}(q^{-1}) \widetilde{\boldsymbol{A}}(q^{-1}) \operatorname{adj} \widetilde{\boldsymbol{D}}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.2.18}$$

$$\det \widetilde{\boldsymbol{D}}(q^{-1}) \hat{\boldsymbol{v}}(t \mid t+N) = \boldsymbol{L}_N^{v}(q^{-1}) \widetilde{\boldsymbol{A}}(q^{-1}) \operatorname{adj} \widetilde{\boldsymbol{D}}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.2.19}$$

【例 5.2.1】 考虑动态系统

$$y(t) = \frac{q^{-1}(1 + bq^{-1})}{(1 + a_1 q^{-1})(1 + a_2 q^{-1})} w(t) + v(t) \tag{5.2.20}$$

$$(1 + pq^{-1}) v(t) = \xi(t) \tag{5.2.21}$$

其中 $y(t)$ 为标量观测信号，输入白噪声 $w(t)$ 是 Bernoulli - Gaussian 白噪声

$$w(t) = b(t)g(t) \tag{5.2.22}$$

其中 $b(t)$ 是取值 1 或 0 的 Bernoulli 白噪声，取值概率为 $P(b(t)=1)=\lambda$，$P(b(t)=0)=\lambda$，$g(t)$ 是零均值、方差为 σ_g^2 的独立于 $b(t)$ 的 Gaussian 白噪声，有色观测噪声 $v(t)$ 服从 AR(1) 模型(5.2.21)，其中 $\xi(t)$ 是零均值、方差为 σ_ξ^2 独立于 $w(t)$ 的 Gaussian 白噪声. 容易知道 $w(t)$ 是零均值、方差为 $\sigma_w^2=\lambda\sigma_g^2$ 的取非零值稀疏的白噪声. 取 $N=3$，问题是求 $w(t)$ 的 Wiener 反卷积平滑器 $\hat{w}(t \mid t+3)$.

仿真中取 $\lambda=0.3$，$\sigma_g^2=1$，$a_1=0.5$，$a_2=0.1$，$b=0.2$，$p=0.1$，$\sigma_\xi^2=0.02$ 和 $\sigma_\xi^2=0.1$. 将(5.2.21)代入(5.2.20)通分后可化为(5.1.1)形式的动态系统. 应用推论(5.2.1)可得白噪声 Wiener 反卷积平滑器 $\hat{w}(t \mid t+3)$. 对 $\sigma_\xi^2=0.02$(情形1)和 $\sigma_\xi^2=0.1$(情形2)，仿真结果分别如图 5.2.1 和图 5.2.2 所示，其中实线纵坐标代表真实值 $w(t)$，圆点纵标代表 Wiener 平滑估值 $\hat{w}(t \mid t+3)$. 可看出取情形 1 比情形 2 精度高. 从理论上对情形 1 可求得方差 $P_w(3)=0.027\,384$，对情形 2 可求得 $P_w(3)=0.082\,101$.

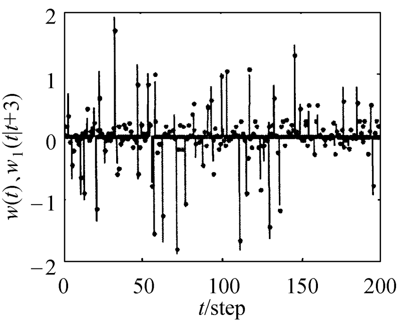


图 5.2.1　Bernoulli - Gaussian 白噪声 $w(t)$ 反卷积和平滑器 $\hat{w}(t \mid t+3)$ ($\sigma_\xi^2=0.02$ 情形)

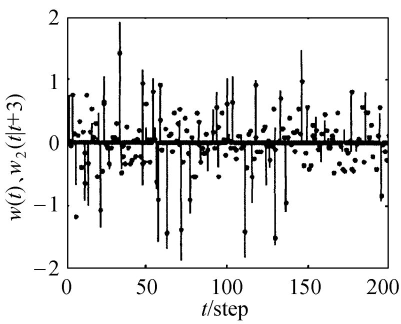


图 5.2.2　Bernoulli - Gaussian 白噪声 $w(t)$ 反卷积和平滑器 $\hat{w}(t \mid t+3)$ ($\sigma_\xi^2=0.1$ 情形)

【注 5.2.1】　Bernoulli - Gaussian 白噪声 $w(t)$ 估计问题在石油地震勘探中有重要的应用[13~17]. 在石油地震勘探信号处理中，埋于地表下的炸药爆炸后产生的地震波经油层反射后引出的反射系数系列可用 Bernoulli - Gaussian 白噪声来描写，它含有是否有油田和油田几何形状的重要信息. 因此估计 Bernoulli - Gaussian 白噪声反射系数序列具有重要的应用价值. Mendel[13~17] 用 Kalman 滤波方法解决白噪声估计问题，而这里用现代时间序列分析方法给出白噪声估值器. 两种方法本质不同是前者用 Riccati 方程而后者用 ARMA 新息模型来解决问题. 对处理单通道白噪声反卷积问题，用 Gevers - Wouters 算法可快速构造 ARMA 新息模型，其中避免了矩阵求逆和矩阵运算，仅需标量的四则运算，因此可减小计算负担.

5.3　多通道 ARMA 信号 Wiener 滤波器

多通道 ARMA 信号在白色观测噪声下的最优滤波、预报和平滑问题广泛应用于信号处理、通信、目标跟踪等领域[164,171,172]. 第四章我们介绍了单通道和多通道 ARMA 信号的

预报器，本节利用白噪声估值器将进一步给出滤波器和平滑器，并且可在统一框架下处理滤波、平滑和预报问题，并给出了相应的估计误差方差阵.虽然节 3.9.7 利用 Kalman 滤波方法基于 Riccati 方程和白噪声估值器给出了多通道 ARMA 信号 Wiener 滤波器，但采用等价于 ARMA 模型的状态空间模型增加了计算负担.

考虑多通道 ARMA 信号 Wiener 滤波问题：

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.3.1}$$

$$\boldsymbol{y}(t) = \boldsymbol{s}(t) + \boldsymbol{v}(t) \tag{5.3.2}$$

其中观测 $\boldsymbol{y}(t) \in R^m$，待估信号 $\boldsymbol{s}(t) \in R^m$，$\boldsymbol{v}(t) \in R^m$ 是观测噪声，$\boldsymbol{w}(t) \in R^r$ 是输入噪声，且 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 是零均值、方差各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$ 相互独立白噪声：

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix}[\boldsymbol{w}^{\mathrm{T}}(j), \boldsymbol{v}^{\mathrm{T}}(j)]\right\} = \begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{Q}_v\end{bmatrix}\delta_{tj} \tag{5.3.3}$$

其中 E 为均值号，T 为转置号，$\delta_{tt} = 1$，$\delta_{tj} = 0\ (t \neq j)$.假设 $(\boldsymbol{A}(q^{-1}), \boldsymbol{C}(q^{-1}))$ 左素，q^{-1} 是单位滞后算子，多项式矩阵 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{C}(q^{-1})$ 有形式 $\boldsymbol{X}(q^{-1}) = \boldsymbol{X}_0 + \boldsymbol{X}_1 q^{-1} + \cdots + \boldsymbol{X}_{n_x} q^{-n_x}$，且设 $n_a \geqslant n_c$，$\boldsymbol{A}_0 = \boldsymbol{I}_m \boldsymbol{C}_0 = \boldsymbol{0}$.设初始观测时刻 $t_0 = -\infty$，问题是基于观测 $(\boldsymbol{y}(t+N), \boldsymbol{y}(t+N-1), \cdots)$ 求信号 $\boldsymbol{s}(t)$ 的最优 Wiener 滤波器 $\hat{\boldsymbol{s}}(t \mid t+N)$.对 $N = 0$，$N > 0$ 或 $N < 0$，各称其为 Wiener 滤波器、平滑器或预报器.

将式(5.3.1)代入式(5.3.2)有

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.3.4}$$

由 $(\boldsymbol{A}(q^{-1}), \boldsymbol{C}(q^{-1}))$ 左素，存在一个可逆的 MA 过程 $\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)$ 使

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.3.5}$$

其中 $\boldsymbol{D}_0 = \boldsymbol{I}_m$，$\boldsymbol{D}(q^{-1})$ 是稳定的，新息 $\boldsymbol{\varepsilon}(t) \in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声. $\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers - Wouters 算法求得.于是由上两式引出 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.3.6}$$

【引理5.3.1】 对平稳或非平稳 ARMA 信号 $\boldsymbol{s}(t)$（即 $\boldsymbol{A}(q^{-1})$ 是稳定的或非稳定的多项式矩阵），当 $t_0 = -\infty$ 时，稳态最优滤波器 $\hat{\boldsymbol{y}}(t \mid t+N)$，$\hat{\boldsymbol{s}}(t \mid t+N)$ 和 $\hat{\boldsymbol{v}}(t \mid t+N)$（$N = 0$，$N > 0$ 或 $N < 0$）有关系

$$\hat{\boldsymbol{s}}(t \mid t+N) = \hat{\boldsymbol{y}}(t \mid t+N) - \hat{\boldsymbol{v}}(t \mid t+N) \tag{5.3.7}$$

证明 当 $\boldsymbol{A}(q^{-1})$ 稳定时，$\boldsymbol{s}(t)$，$\boldsymbol{y}(t)$ 和 $\boldsymbol{v}(t)$ 均为平稳时间序列，应用射影理论，取(5.3.2)两边各项在无穷维 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t+N), \boldsymbol{y}^{\mathrm{T}}(t+N-1), \cdots)$ 上的射影立刻得(5.3.7).当 $\boldsymbol{A}(q^{-1})$ 不稳定时，$\boldsymbol{s}(t)$ 和 $\boldsymbol{y}(t)$ 是非平稳时间序列，当 $t_0 = -\infty$ 时，$\boldsymbol{s}(t)$ 和 $\boldsymbol{y}(t)$ 的分量有无穷大方差[1,173]，因此不能直接应用射影理论.此时由 $(\boldsymbol{y}^{\mathrm{T}}(t+N), \boldsymbol{y}^{\mathrm{T}}(t+N-1), \cdots)$ 的线性运算不能构成无穷维 Hilbere 空间，这是因为 $\boldsymbol{y}(t)$ 的分量方差是无界的.而无穷维 Hilbere 空间是由具有二阶矩的随机变量生成的闭线性流形构成.在这种情形下，可用 Kalman 滤波理论证明关系(5.3.7).事实上(5.3.1)和(5.3.2)有状态空间模型

$$\boldsymbol{x}(t+1) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{C}\boldsymbol{w}(t) \tag{5.3.8}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}(t) \tag{5.3.9}$$

$$\boldsymbol{s}(t) = \boldsymbol{H}\boldsymbol{x}(t) \tag{5.3.10}$$

其中定义 $\boldsymbol{C}_i = \boldsymbol{0}(i > n_c)$，且

$$\boldsymbol{A} = \begin{bmatrix} -\boldsymbol{A}_1 & & \\ \vdots & \boldsymbol{I}_{(n_a-1)m} & \\ -\boldsymbol{A}_{n_a} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix},\quad \boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_1 \\ \vdots \\ \boldsymbol{C}_{n_a} \end{bmatrix},\quad \boldsymbol{H} = [\boldsymbol{I}_m \quad \boldsymbol{0}\cdots\boldsymbol{0}] \tag{5.3.11}$$

由 $(\boldsymbol{A}(q^{-1}),\boldsymbol{C}(q^{-1}))$ 左素的假设引出系统(5.3.8) ~ (5.3.10)是完全可观、完全可控的，故存在稳态 Kalman 估值器[171] $\hat{\boldsymbol{x}}(t \mid t+N)$，且有关系

$$\hat{\boldsymbol{y}}(t \mid t+N) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t+N) + \hat{\boldsymbol{v}}(t \mid t+N) \tag{5.3.12}$$

$$\hat{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t+N) \tag{5.3.13}$$

上述稳定关系可由取初始观测时刻 t_0 为有限时刻，基于有限个观测 $(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots,\boldsymbol{y}(t_0))$ 可生成有限维 Hilbert 空间 $L(\boldsymbol{y}(t+N),\cdots,\boldsymbol{y}(t_0))$，取(5.3.9)和(5.3.10)各项到 Hilbert 空间 $L(y(t+N),\cdots,y(t_0))$ 上射影引出非稳态关系(5.3.12)和(5.3.13)．再令 $t_0 \to -\infty$ 便得到稳态关系(5.3.12)和(5.3.13)．注意当取初值 $\boldsymbol{x}(t_0)$ 具有有限方差阵时，由(5.3.8)和(5.3.9)可递推得到具有有限方差阵的观测 $(\boldsymbol{y}(t+N),\cdots,\boldsymbol{y}(t_0))$．□

由引理 5.3.1 看到，为了求信号最优 Wiener 滤波器 $\hat{\boldsymbol{s}}(t \mid t+N)$ 要求计算观测预报器 $\hat{\boldsymbol{y}}(t \mid t+N)(N<0)$．当 $\boldsymbol{A}(q^{-1})$ 稳定时，平稳 ARMA 过程(5.3.6)的 Åström 预报器已在节 4.9 中用无穷维 Hilbert 空间射影理论被导出．现在证明 $\boldsymbol{A}(q^{-1})$ 不稳定时 Åström 预报器仍然成立．

【引理 5.3.2】 (多变量非平稳 ARMA 过程的 Åström 预报器)　考虑多变量非平稳 ARMA 过程(5.3.6)，其中观测 $\boldsymbol{y}(t) \in R^m$，$\boldsymbol{\varepsilon}(t) \in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声．假设 $\boldsymbol{A}(q^{-1})$ 是不稳定的，$\boldsymbol{D}(q^{-1})$ 是稳定的，则有基于无穷个观测 $(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots,)$ 的超前 $N>0$ 步稳态最优(线性最小方差)Wiener 预报器为

$$\hat{\boldsymbol{y}}(t+N \mid t) = \boldsymbol{J}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t) \tag{5.3.14}$$

带伪交换

$$\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1}) = \widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{5.3.15}$$

其中 $\widetilde{\boldsymbol{A}}_0 = \boldsymbol{I}_m, \widetilde{\boldsymbol{D}}_0 = \boldsymbol{I}_m$，且 $n_{\tilde{a}} = n_a, n_{\tilde{d}} = n_d$，

$$\det\boldsymbol{D}(q^{-1}) = \det\widetilde{\boldsymbol{D}}(q^{-1}) \tag{5.3.16}$$

而 $\boldsymbol{J}_N(q^{-1})$ 和 $\boldsymbol{E}_N(q^{-1})$ 由如下 Diophantine 方程决定：

$$\widetilde{\boldsymbol{D}}(q^{-1}) = \boldsymbol{E}_N(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) + q^{-N}\boldsymbol{J}_N(q^{-1}) \tag{5.3.17}$$

其中 $\boldsymbol{E}_0 = \boldsymbol{I}_m$

$$\begin{aligned} \boldsymbol{E}_N(q^{-1}) &= \boldsymbol{E}_0^{(N)} + \boldsymbol{E}_1^{(N)}q^{-1} + \cdots + \boldsymbol{E}_{N-1}^{(N)}q^{-(N-1)}, \\ \boldsymbol{J}_N(q^{-1}) &= \boldsymbol{J}_0^{(N)} + \boldsymbol{J}_1^{(N)}q^{-1} + \cdots + \boldsymbol{J}_{n_j}^{(N)}q^{-n_j} \end{aligned} \tag{5.3.18}$$

其中 $n_j = \max(n_a - 1, n_d - N)$．预报误差为

$$\tilde{\boldsymbol{y}}(t+N \mid t) = \boldsymbol{y}(t+N) - \hat{\boldsymbol{y}}(t+N \mid t) = \boldsymbol{E}_N(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.3.19}$$

预报误差方差阵 $\boldsymbol{P}_y(N) = \mathrm{E}[\tilde{\boldsymbol{y}}(t+N \mid t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+N \mid t)]$ 为

$$\boldsymbol{P}_y(N) = \sum_{i=0}^{N-1}\boldsymbol{E}_i^{(N)}\boldsymbol{Q}_\varepsilon\boldsymbol{E}_i^{(N)\mathrm{T}},\quad \boldsymbol{E}_0^{(N)} = \boldsymbol{I}_m \tag{5.3.20}$$

证明　由(5.3.6)，(5.3.15)和(5.3.17)引出

$$\boldsymbol{y}(t+N) = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t+N) = \widetilde{\boldsymbol{D}}(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\boldsymbol{\varepsilon}(t+N) =$$

$$
\begin{aligned}
&\boldsymbol{E}(q^{-1})\boldsymbol{\varepsilon}(t+N)+\boldsymbol{J}_N(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\boldsymbol{\varepsilon}(t)=\\
&\boldsymbol{E}(q^{-1})\boldsymbol{\varepsilon}(t+N)+\boldsymbol{J}_N(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\\
&\boldsymbol{E}(q^{-1})\boldsymbol{\varepsilon}(t+N)+\boldsymbol{J}_N(q^{-1})\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t)=\\
&\boldsymbol{E}(q^{-1})\boldsymbol{\varepsilon}(t+N)+\boldsymbol{J}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t)
\end{aligned}\tag{5.3.21}
$$

上式右边第一项是 $\boldsymbol{\varepsilon}(t+N),\cdots,\boldsymbol{\varepsilon}(t+1)$ 的线性组合，而第二项是渐近稳定的 Wiener 滤波器，可用已知观测 $(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$ 计算. 注意 $\boldsymbol{\varepsilon}(t+1),\cdots,\boldsymbol{\varepsilon}(t+N)$ 是白噪声在将来时刻的值，是不可进一步由 $(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$ 来估计的. 因为由(5.3.6)引出 $(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$ 只与 $(\boldsymbol{\varepsilon}(t+1),\cdots,\boldsymbol{\varepsilon}(t-1),\cdots)$ 有关. 这引出最优 Åström 预报器(5.3.14)和预报误差(5.3.19)及其方差阵(5.3.20). □

【注 5.3.1】 $\boldsymbol{A}(q^{-1})$ 不稳定情形下的 Åström 预报器(5.3.14)相同于 $\boldsymbol{A}(q^{-1})$ 稳定时由节 4.9 导出的 Åström 预报器. 但这里的证明没有用到射影理论.

【注 5.3.2】 当 $\boldsymbol{A}(q^{-1})$ 是不稳定时，$\boldsymbol{y}(t)$ 是非平稳随机序列. 但(5.3.14)给出了渐近稳定的 Wiener 预报器，即由(5.3.16)和 $\boldsymbol{D}(q^{-1})$ 的稳定性引出 $\det\widetilde{\boldsymbol{D}}(q^{-1})$ 是稳定的多项式，因而非平稳 ARMA 过程 $\boldsymbol{y}(t)$ 的 Åström 预报器可递推计算为

$$
\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{y}}(t+N\mid t)=\boldsymbol{J}_N(q^{-1})\operatorname{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t)\tag{5.3.22}
$$

由 $\det\boldsymbol{D}(q^{-1})$ 的稳定性引出(5.3.22)的计算渐近地与预报器初值选取无关. 这与不稳定系统的稳态 Kalman 一步预报器可表为渐近稳定的 Wiener 预报器形式类似. 还应注意，Åström 预报器(5.3.22)等价于在节 4.11 中用稳态 Kalman 滤波方法导出的 Wiener 预报器(4.11.33). 这是因为关系(5.3.16)成立.

【推论 5.3.1】 在引理 5.3.2 条件下，对任意整数 N，有 Åström 预报器

$$
\hat{\boldsymbol{y}}(t+N\mid t)=\boldsymbol{J}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t)\tag{5.3.23}
$$

其中定义 $\boldsymbol{J}_N(q^{-1})$ 为：由 $N>0$ 时，由(5.3.17)决定 $\boldsymbol{J}_N(q^{-1})$；若 $N\leqslant 0$，定义 $\boldsymbol{J}_N(q^{-1})$ 为

$$
\boldsymbol{J}_N(q^{-1})=\widetilde{\boldsymbol{D}}(q^{-1})q^N,\quad N\leqslant 0\tag{5.3.24}
$$

证明 当 $N<0$ 时有

$$
\hat{\boldsymbol{y}}(t+N\mid t)=\boldsymbol{y}(t+N)\tag{5.3.25}
$$

这引出(5.3.24). □

【推论 5.3.2】 对平稳或非平稳 ARMA 过程(5.3.6)有统一的 Åström 预报器

$$
\hat{\boldsymbol{y}}(t\mid t+N)=\boldsymbol{J}_{-N}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N)\tag{5.3.26}
$$

其中 $\boldsymbol{J}_{-N}(q^{-1})$ 定义为

$$
\begin{aligned}
&\widetilde{\boldsymbol{D}}(q^{-1})=\boldsymbol{E}_{-N}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})+q^N\boldsymbol{J}_{-N}(q^{-1}),\quad N<0\\
&\boldsymbol{J}_{-N}(q^{-1})=\widetilde{\boldsymbol{D}}(q^{-1})q^{-N},\quad N\geqslant 0
\end{aligned}\tag{5.3.27}
$$

带 $\boldsymbol{J}_{-N}(q^{-1})$ 的阶次 $n_j=\max(n_a-1,n_d+N)$，$n_a=n_{\tilde{a}}$，$n_d=n_{\tilde{d}}$，有

$$
\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})=\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\tag{5.3.28}
$$

且有预报误差为

$$
\begin{aligned}
&\tilde{\boldsymbol{y}}(t\mid t+N)=\boldsymbol{y}(t)-\hat{\boldsymbol{y}}(t\mid t+N)=\boldsymbol{E}_{-N}(q^{-1})\boldsymbol{\varepsilon}(t),\quad N<0,\\
&\tilde{\boldsymbol{y}}(t\mid t+N)=\boldsymbol{0},\quad N\geqslant 0
\end{aligned}\tag{5.3.29}
$$

相应的预报误差方差阵为

$$\boldsymbol{P}_y(N) = \sum_{i=0}^{-N-1} \boldsymbol{E}_i^{(-N)} \boldsymbol{Q}_\varepsilon \boldsymbol{E}_i^{(-N)\mathrm{T}}, \quad N < 0$$

$$\boldsymbol{P}_y(N) = \boldsymbol{0}, \quad N \geqslant 0 \tag{5.3.30}$$

【引理 5.3.3】 （多变量平稳或非平稳 ARMA 过程的 Koivo 预报器）对于平稳或非平稳多变量 ARMA 过程(5.3.6)，其中 $\boldsymbol{A}(q^{-1})$ 是稳定的或不稳定的多项式矩阵，$\boldsymbol{D}(q^{-1})$ 是稳定的，$\boldsymbol{y}(t) \in R^m$，$\boldsymbol{\varepsilon}(t) \in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，则有渐近稳定的 Koivo 稳态最优预报器

$$\hat{\boldsymbol{y}}(t+N \mid t) = \widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \widetilde{\boldsymbol{G}}_N(q^{-1}) \boldsymbol{y}(t), \quad N > 0 \tag{5.3.31}$$

或表为 ARMA 递推形式

$$\widetilde{\boldsymbol{D}}(q^{-1}) \hat{\boldsymbol{y}}(t+N \mid t) = \widetilde{\boldsymbol{G}}_N(q^{-1}) \boldsymbol{y}(t), \quad N > 0 \tag{5.3.32}$$

或

$$\det \boldsymbol{D}(q^{-1}) \hat{\boldsymbol{y}}(t+N \mid t) = \operatorname{adj} \widetilde{\boldsymbol{D}}(q^{-1}) \widetilde{\boldsymbol{G}}_N(q^{-1}) \boldsymbol{y}(t), \quad N > 0 \tag{5.3.33}$$

其中 $\boldsymbol{E}_N(q^{-1})$ 和 $\boldsymbol{G}_N(q^{-1})$ 由如下 Diophautime 方程决定：

$$\boldsymbol{D}(q^{-1}) = \boldsymbol{A}(q^{-1}) \boldsymbol{E}_N(q^{-1}) + q^{-N} \boldsymbol{G}_N(q^{-1}) \tag{5.3.34}$$

$$\boldsymbol{E}_N(q^{-1}) = \boldsymbol{E}_0^{(N)} + \boldsymbol{E}_1^{(N)} q^{-1} + \cdots + \boldsymbol{E}_{N-1}^{(N)} q^{-(N-1)},$$

$$\boldsymbol{G}_N(q^{-1}) = \boldsymbol{G}_0^{(N)} + \boldsymbol{G}_1^{(N)} q^{-1} + \cdots + \boldsymbol{G}_{n_g}^{(N)} q^{-n_g} \tag{5.3.35}$$

其中 $n_g = \max(n_a - 1, n_d - N)$. 引入伪交换

$$\boldsymbol{G}_N(q^{-1}) \boldsymbol{E}_N^{-1}(q^{-1}) = \widetilde{\boldsymbol{E}}_N^{-1}(q^{-1}) \widetilde{\boldsymbol{G}}_N(q^{-1}) \tag{5.3.36}$$

带 $\det \boldsymbol{E}_N(q^{-1}) = \det \widetilde{\boldsymbol{E}}_N(q^{-1})$，$\widetilde{\boldsymbol{E}}_0^{(N)} = \boldsymbol{I}_m$. 定义

$$\widetilde{\boldsymbol{D}}(q^{-1}) = \widetilde{\boldsymbol{E}}_N(q^{-1}) \boldsymbol{A}(q^{-1}) + q^{-N} \widetilde{\boldsymbol{G}}_N(q^{-1}) \tag{5.3.37}$$

则有

$$\det \widetilde{\boldsymbol{D}}(q^{-1}) = \det \boldsymbol{D}(q^{-1}),$$

$$\widetilde{\boldsymbol{D}}(q^{-1}) \boldsymbol{E}_N(q^{-1}) = \widetilde{\boldsymbol{E}}_N(q^{-1}) \boldsymbol{D}(q^{-1}) \tag{5.3.38}$$

预报误差 $\tilde{\boldsymbol{y}}(t+N \mid t) = \boldsymbol{y}(t+N) - \hat{\boldsymbol{y}}(t+N \mid t)$ 为

$$\tilde{\boldsymbol{y}}(t+N \mid t) = \widetilde{\boldsymbol{E}}_N(q^{-1}) \boldsymbol{\varepsilon}(t+N), \quad N > 0 \tag{5.3.39}$$

预报误差方差阵 $\boldsymbol{P}_y(N) = \mathrm{E}[\tilde{\boldsymbol{y}}(t+N \mid t) \tilde{\boldsymbol{y}}^{\mathrm{T}}(t+N \mid t)]$ 为

$$\boldsymbol{P}_y(N) = \sum_{i=0}^{N-1} \boldsymbol{E}_i^{(N)} \boldsymbol{Q}_\varepsilon \boldsymbol{E}_i^{(N)\mathrm{T}}, \quad \boldsymbol{E}_0^{(N)} = \boldsymbol{I}_m, \quad N > 0 \tag{5.3.40}$$

且当 $N \leqslant 0$ 时，有

$$\hat{\boldsymbol{y}}(t+N \mid t) = \boldsymbol{y}(t+N), \quad N \leqslant 0 \tag{5.3.41}$$

相应的预报误差方差阵为

$$\boldsymbol{P}_y(N) = \boldsymbol{0}, \quad N \leqslant 0 \tag{5.3.42}$$

证明　类似于定理 4.10.1 的推导和引理 5.3.2 的证明，得证. 详细推导从略. □

【引理 5.3.4】 带白色观测噪声的多通道 ARMA 信号(5.3.1)～(5.3.2)有观测白噪声 Wiener 滤波器

$$\hat{\boldsymbol{v}}(t+N \mid t) = \boldsymbol{L}_N^v(q^{-1}) \widetilde{\boldsymbol{A}}(q^{-1}) \widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.3.43}$$

或

$$\det \widetilde{\boldsymbol{D}}(q^{-1}) \hat{\boldsymbol{v}}(t+N \mid t) = \boldsymbol{L}_N^v(q^{-1}) \widetilde{\boldsymbol{A}}(q^{-1}) \operatorname{adj} \widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.3.44}$$

其中 $\widetilde{\boldsymbol{A}}(q^{-1})$ 和 $\widetilde{\boldsymbol{D}}(q^{-1})$ 由伪交换 (5.3.15) 定义

$$\boldsymbol{L}_N^v(q^{-1}) = \sum_{i=0}^{N} \boldsymbol{Q}_v \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{Q}_\varepsilon^{-1} q^{i-N}, \quad N \geqslant 0,$$

$$\boldsymbol{L}_N^v(q^{-1}) = \boldsymbol{0}, \quad N < 0 \tag{5.3.45}$$

其中系数阵 $\boldsymbol{G}_i$ 可递推计算为

$$\boldsymbol{G}_i = -\boldsymbol{D}_1 \boldsymbol{G}_{i-1} - \cdots - \boldsymbol{D}_{n_d} \boldsymbol{G}_{i-n_d} + \boldsymbol{A}_i \tag{5.3.46}$$

其中规定 $\boldsymbol{G}_i = \boldsymbol{0}(i < 0), \boldsymbol{A}_i = \boldsymbol{0}(i > n_a)$. 估值误差方差阵 $\boldsymbol{P}_v(N) = [\mathrm{E}(\boldsymbol{v}(t) - \hat{\boldsymbol{v}}(t \mid t + N))(\boldsymbol{v}(t) - \hat{\boldsymbol{v}}(t \mid t + N))^{\mathrm{T}}]$ 为

$$\boldsymbol{P}_v(N) = \boldsymbol{Q}_v - \sum_{i=0}^{N} \boldsymbol{Q}_v \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{Q}_\varepsilon^{-1} \boldsymbol{G}_i \boldsymbol{Q}_v, \quad N \geqslant 0,$$

$$\boldsymbol{P}_v(N) = \boldsymbol{Q}_v, \quad N < 0 \tag{5.3.47}$$

【引理 5.3.5】 带白色观测噪声的多通道 ARMA 信号 (5.3.1) ~ (5.3.3) 有观测白噪声 Wiener 滤波器

$$\det \boldsymbol{D}(q^{-1}) \hat{\boldsymbol{v}}(t \mid t + N) = \boldsymbol{L}_N^v(q^{-1}) \operatorname{adj} \boldsymbol{D}(q^{-1}) \boldsymbol{A}(q^{-1}) \boldsymbol{y}(t + N) \tag{5.3.48}$$

其中 $\boldsymbol{A}(q^{-1}), \boldsymbol{D}(q^{-1})$ 由 (5.3.6) 定义.

引理 5.3.4 和引理 5.3.5 的证明见节 5.2.

【定理 5.3.1】 带白色观测噪声的多通道 ARMA 信号 (5.3.1) ~ (5.3.3) 有渐近稳定的 Wiener 滤波器

$$\hat{\boldsymbol{s}}(t \mid t + N) = \boldsymbol{K}_N(q^{-1}) \widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \boldsymbol{y}(t + N) \tag{5.3.49}$$

其中多项式矩阵 $\boldsymbol{K}_N(q^{-1})$ 定义为

$$\boldsymbol{K}_N(q^{-1}) = \boldsymbol{J}_{-N}(q^{-1}) - \boldsymbol{L}_N^v(q^{-1}) \widetilde{\boldsymbol{A}}(q^{-1}) \tag{5.3.50}$$

(5.3.49) 有 ARMA 递推形式

$$\det \widetilde{\boldsymbol{D}}(q^{-1}) \hat{\boldsymbol{s}}(t \mid t + N) = \boldsymbol{K}_N(q^{-1}) \operatorname{adj} \widetilde{\boldsymbol{D}}(q^{-1}) \boldsymbol{y}(t + N) \tag{5.3.51}$$

其中 $\boldsymbol{A}(q^{-1}), \widetilde{\boldsymbol{A}}(q^{-1}), \boldsymbol{D}(q^{-1})$ 和 $\widetilde{\boldsymbol{D}}(q^{-1})$ 由 (5.3.6) 和 (5.3.15) 定义，且有滤波误差方差阵 $\boldsymbol{P}(N) = \mathrm{E}[(\boldsymbol{s}(t) - \hat{\boldsymbol{s}}(t \mid t + N))[(\boldsymbol{s}(t) - \hat{\boldsymbol{s}}(t \mid t + N))^{\mathrm{T}}]$ 为

$$\boldsymbol{P}(N) = \boldsymbol{Q}_v - \sum_{i=0}^{N} \boldsymbol{Q}_v \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{Q}_\varepsilon^{-1} \boldsymbol{G}_i \boldsymbol{Q}_v, \quad N \geqslant 0 \tag{5.3.52}$$

$$\boldsymbol{P}(N) = \sum_{i=0}^{-N-1} \boldsymbol{E}_i^{(-N)} \boldsymbol{Q}_\varepsilon \boldsymbol{E}_i^{(-N)\mathrm{T}} - \boldsymbol{Q}_v, \quad N < 0 \tag{5.3.53}$$

其中 $\boldsymbol{E}_i^{(-N)}$ 由 (5.3.17) 计算，$\boldsymbol{G}_i$ 由 (5.3.46) 计算.

证明　由 (5.3.2) 有

$$\hat{\boldsymbol{s}}(t \mid t + N) = \hat{\boldsymbol{y}}(t \mid t + N) - \hat{\boldsymbol{v}}(t \mid t + N) \tag{5.3.54}$$

将 (5.3.26) 和 (5.3.43) 代入上式得 (5.3.49) 和 (5.3.50)，进而可得 (5.3.51). 由 (5.3.7) 当 $N \geqslant 0$ 时有

$$\hat{\boldsymbol{s}}(t \mid t + N) = \boldsymbol{y}(t) - \hat{\boldsymbol{v}}(t \mid t + N), \quad N \geqslant 0 \tag{5.3.55}$$

应用 (5.3.55) 和 (5.3.2) 有估值误差

$$\tilde{\boldsymbol{s}}(t \mid t + N) = \boldsymbol{s}(t) - \hat{\boldsymbol{s}}(t \mid t + N) = -\boldsymbol{v}(t) - \hat{\boldsymbol{v}}(t \mid t + N) \tag{5.3.56}$$

于是应用 (5.3.47) 有

$$\boldsymbol{P}(N) = \boldsymbol{P}_v(N) = \boldsymbol{Q}_v - \sum_{i=0}^{N} \boldsymbol{Q}_v \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{Q}_\varepsilon^{-1} \boldsymbol{G}_i \boldsymbol{Q}_v, \quad N \geqslant 0 \tag{5.3.57}$$

当 $N < 0$ 时由引理 5.3.4 有 $\hat{\boldsymbol{v}}(t \mid t+N) = \boldsymbol{0}, N < 0$，于是由 (5.3.7) 引出

$$\hat{\boldsymbol{s}}(t \mid t+N) = \hat{\boldsymbol{y}}(t \mid t+N), \quad N < 0 \tag{5.3.58}$$

由此式和 (5.3.2) 有 $\tilde{\boldsymbol{s}}(t \mid t+N) = \tilde{\boldsymbol{y}}(t \mid t+N) - v(t)$，故应用 (5.3.29) 有

$$\boldsymbol{P}(N) = \boldsymbol{P}_y(N) = \sum_{i=0}^{-N-1} \boldsymbol{E}_i^{(-N)} \boldsymbol{Q}_\varepsilon \boldsymbol{E}_i^{(-N)\mathrm{T}} - \boldsymbol{Q}_v, \quad N < 0 \tag{5.3.59}$$

其中当 $N < 0$ 时应用了事实 $\mathrm{E}[\tilde{\boldsymbol{y}}(t \mid t+N)\boldsymbol{v}^{\mathrm{T}}(t)] = \mathrm{E}[\boldsymbol{y}(t)\boldsymbol{v}^{T}(t)] = \boldsymbol{Q}_v$. 证毕. □

【定理 5.3.2】 带白色观测噪声的多通道 ARMA 信号 (5.3.1) ~ (5.3.3) 有渐近稳定的 Wiener 滤波器

$$\det \boldsymbol{D}(q^{-1})\hat{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{M}_N(q^{-1})\boldsymbol{y}(t+N), \quad N \geqslant 0 \tag{5.3.60}$$

其中定义

$$\boldsymbol{M}_N(q^{-1}) = \det \boldsymbol{D}(q^{-1})\boldsymbol{I}_m q^{-N} - \boldsymbol{L}_N^v(q^{-1}) \operatorname{adj} \boldsymbol{D}(q^{-1})\boldsymbol{A}(q^{-1}) \tag{5.3.61}$$

且有渐近稳定的 Wiener 滤波器

$$\det \boldsymbol{D}(q^{-1})\hat{\boldsymbol{s}}(t \mid t+N) = \operatorname{adj} \widetilde{\boldsymbol{D}}(q^{-1})\widetilde{\boldsymbol{G}}_{-N}(q^{-1})\boldsymbol{y}(t+N), \quad N < 0 \tag{5.3.62}$$

其中 $\widetilde{\boldsymbol{D}}(q^{-1}), \widetilde{\boldsymbol{G}}_{-N}(q^{-1})$ 由 (5.3.34)，(5.3.36)，(5.3.37) 决定. 且有估值误差方差阵 $\boldsymbol{P}(N)$ 为

$$\boldsymbol{P}(N) = \boldsymbol{Q}_v - \sum_{i=0}^{N} \boldsymbol{Q}_v \boldsymbol{G}_i^{\mathrm{T}} \boldsymbol{Q}_\varepsilon^{-1} \boldsymbol{G}_i \boldsymbol{Q}_v, \quad N \geqslant 0 \tag{5.3.63}$$

$$\boldsymbol{P}(N) = \sum_{i=0}^{-N-1} \boldsymbol{E}_i^{(-N)} \boldsymbol{Q}_\varepsilon \boldsymbol{E}_i^{(-N)\mathrm{T}} - \boldsymbol{Q}_v, \quad N < 0 \tag{5.3.64}$$

其中 $\boldsymbol{E}_i^{(-N)}$ 由 (5.3.34) 计算，$\boldsymbol{G}_i$ 由 (5.3.46) 计算.

证明 当 $N \geqslant 0$ 时，由 (5.3.7) 和 (5.3.48) 有

$$\hat{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{y}(t) - \hat{\boldsymbol{v}}(t \mid t+N) = \boldsymbol{y}(t) - \frac{\boldsymbol{L}_N^v(q^{-1}) \operatorname{adj} \boldsymbol{D}(q^{-1})\boldsymbol{A}(q^{-1})}{\det \boldsymbol{D}(q^{-1})}\boldsymbol{y}(t+N) =$$

$$\frac{\det \boldsymbol{D}(q^{-1})\boldsymbol{I}_m q^{-N} - \boldsymbol{L}_N^v(q^{-1}) \operatorname{adj} \boldsymbol{D}(q^{-1})\boldsymbol{A}(q^{-1})}{\det \boldsymbol{D}(q^{-1})}\boldsymbol{y}(t+N) \tag{5.3.65}$$

这引出 (5.3.60) 和 (5.3.61).

当 $N < 0$ 时，由 (5.3.7) 和 (5.3.32) 有

$$\hat{\boldsymbol{s}}(t \mid t+N) = \hat{\boldsymbol{y}}(t \mid t+N) = \frac{1}{\det \boldsymbol{D}(q^{-1})} \operatorname{adj} \widetilde{\boldsymbol{D}}^{-1}(q^{-1})\widetilde{\boldsymbol{G}}_{-N}(q^{-1})\boldsymbol{y}(t+N) \tag{5.3.66}$$

其中应用了 (5.3.38). 这引出 (5.3.62). 由 (5.3.47) 和 (5.3.38) 引出 (5.3.63) 和 (5.3.64). □

【注 5.3.3】 定理 5.3.1 和定理 5.3.2 的区别为后者不用伪交换 (5.3.28)，但却用伪交换 (5.3.36). 前者应用 (5.3.43) 形式的白噪声估值器 $\hat{\boldsymbol{v}}(t \mid t+N)$，后者应用 (5.3.48) 形式的白噪声估值器. 前者用 Åström 预报器 $\hat{\boldsymbol{y}}(t \mid t+N)(N < 0)$，后者用 Koivo 预报器.

现在进一步考虑带白色观测噪声和常的观测偏差的多通道 ARMA 信号 Wiener 滤波问题：

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.3.67}$$

$$\boldsymbol{y}(t) = \boldsymbol{s}(t) + \boldsymbol{b} + \boldsymbol{v}(t) \tag{5.3.68}$$

除了(5.3.2)用(5.3.68)代替，其中 $\boldsymbol{b}$ 为常的观测偏差向量，$\boldsymbol{b} \in R^m$，其他假设同(5.3.1)和(5.3.3).

引入新的观测过程

$$\boldsymbol{z}(t) = \boldsymbol{y}(t) - \boldsymbol{b} \tag{5.3.69}$$

则有等价的化为(5.3.2)形式的观测方程

$$\boldsymbol{z}(t) = \boldsymbol{s}(t) + \boldsymbol{v}(t) \tag{5.3.70}$$

系统(5.3.67)和(5.3.70)同系统(5.3.1)和(5.3.2)本质上是相同的.因此利用定理5.3.1和定理5.3.2的结果有如下定理.

【定理5.3.3】 多通道(5.3.67)和(5.3.68)在变换 $\boldsymbol{z}(t) = \boldsymbol{y}(t) - \boldsymbol{b}$ 下有ARMA新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{z}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.3.71}$$

渐近稳定的Wiener滤波器为

$$\hat{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{z}(t+N)$$

或

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\operatorname{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{z}(t+N) \tag{5.3.72}$$

估计误差方差阵 $\boldsymbol{P}(N)$ 由(5.3.52)和(5.3.53)给出.所有符号含义同定理5.3.1.

【定理5.3.4】 多通道系统(5.3.67)和(5.3.68)在变换 $\boldsymbol{z}(t) = \boldsymbol{y}(t) - \boldsymbol{b}$ 下，有ARMA新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{z}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.3.73}$$

渐近稳定的Wiener滤波器为

$$\begin{aligned}\det\boldsymbol{D}(q^{-1})\hat{\boldsymbol{s}}(t \mid t+N) &= \boldsymbol{M}_N(q^{-1})\boldsymbol{z}(t+N), \quad N \geqslant 0,\\ \det\boldsymbol{D}(q^{-1})\hat{\boldsymbol{s}}(t \mid t+N) &= \operatorname{adj}\widetilde{\boldsymbol{D}}(q^{-1})\widetilde{\boldsymbol{G}}_{-N}(q^{-1})\boldsymbol{z}(t+N), \quad N < 0\end{aligned} \tag{5.3.74}$$

且误差方差阵为(5.3.63)和(5.3.64).所有符号含义同定理5.3.2.

在标量情形(单通道)，分别记 $\boldsymbol{E}_i^{(N)}, \boldsymbol{G}_i, \boldsymbol{Q}_w, \boldsymbol{Q}_v, \boldsymbol{Q}_\varepsilon$ 为 $e_i^{(N)}, g_i, \sigma_w^2, \sigma_v^2, \sigma_\varepsilon^2$，由定理5.3.1和定理5.3.2有如下推论.

【推论5.3.1】 对单通道($m=1$)ARMA信号(5.3.1)~(5.3.3)，有渐近稳定的Wiener滤波器

$$D(q^{-1})\hat{s}(t \mid t+N) = K_N(q^{-1})y(t+N) \tag{5.3.75}$$

其中定义多项式

$$K_N(q^{-1}) = J_{-N}(q^{-1}) - L_N^v(q^{-1})A(q^{-1}) \tag{5.3.76}$$

其中 $A(q^{-1}), D(q^{-1})$ 由(5.3.6)决定，$L_N^v(q^{-1})$ 定义为

$$\begin{aligned}L_N^v(q^{-1}) &= \sum_{i=0}^{N}\frac{\sigma_v^2}{\sigma_\varepsilon^2}g_i q^{i-N}, \quad N \geqslant 0\\ L_N^v(\varepsilon^{-1}) &= 0, \quad N < 0\end{aligned} \tag{5.3.77}$$

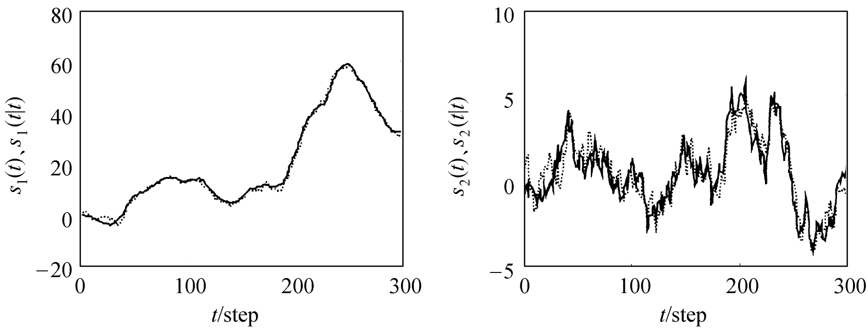
$J_{-N}(q^{-1})$ 由如下Diophantine方程决定：

$$D(q^{-1}) = E_{-N}(q^{-1})A(q^{-1}) + q^N J_{-N}(q^{-1}), \quad N < 0,$$
$$J_{-N}(q^{-1}) = D(q^{-1})q^{-N}, \quad N \geqslant 0 \tag{5.3.78}$$
其中 $E_{-N}(q^{-1})$ 和 $J_{-N}(q^{-1})$ 有形式(5.3.18),分别带系数 $e_i^{(N)}$ 和 $j_i^{(N)}$. 估值误差方差为
$$P(N) = \sigma_v^2\left[1 - \frac{\sigma_v^2}{\sigma_\varepsilon^2}\sum_{i=0}^{N} g_i^2\right], \quad N \geqslant 0 \tag{5.3.79}$$
$$P(N) = \sigma_\varepsilon^2 \sum_{i=0}^{-N-1}(e_i^{(-N)})^2, \quad N < 0 \tag{5.3.80}$$
证明 注意对单通道情形有 $\widetilde{A}(q^{-1}) = A(q^{-1})$, $\widetilde{D}(q^{-1}) = D(q^{-1})$,由定理 5.3.1 立刻得证. □

【推论 5.3.2】 对单通道情形由定理 5.3.2 可引出推论 5.3.1 的结果.

【推论 5.3.3】 对带白色观测噪声和常数观测偏差的单通道 ARMA 信号(5.3.67)和(5.3.68),引入新的观测 $z(t) = y(t) - b$ 后有 ARMA 新息模型(5.3.71),且有渐近稳定的 Wiener 滤波器为
$$D(q^{-1})\hat{s}(t \mid t + N) = K_N(q^{-1})z(t + N) \tag{5.3.81}$$
其中 $K_N(q^{-1})$ 由推论 5.3.1 计算,且有估值误差方差(5.3.79)和(5.3.80).

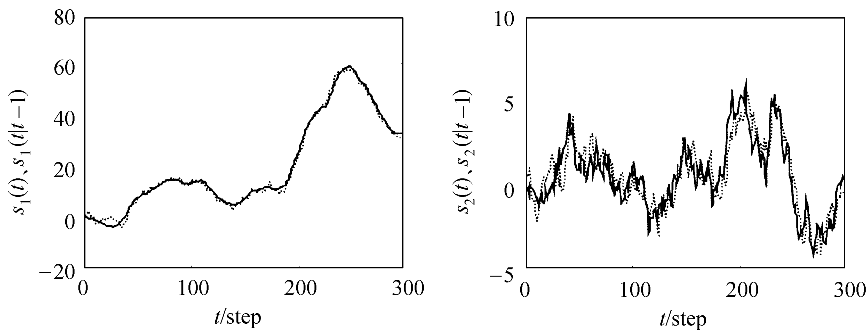
【例 5.3.1】 考虑二维跟踪系统
$$(\boldsymbol{I}_2 + \boldsymbol{A}_1 q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}_1 q^{-1}\boldsymbol{w}(t) \tag{5.3.82}$$
$$\boldsymbol{y}(t) = \boldsymbol{s}(t) + \boldsymbol{v}(t) + \boldsymbol{b} \tag{5.3.83}$$
$$\boldsymbol{A}_1 = -\begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{C}_1 = \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix} \tag{5.3.84}$$
其中 T_0 为采样周期, $\boldsymbol{b}$ 为观测系统偏差信号, $\boldsymbol{s}(t) = [s_1(t), s_2(t)]^{\mathrm{T}}$, $s_1(t)$, $s_2(t)$ 和 $w(t)$ 各为运动目标在时刻 tT_0 处的位置、速度和加速度, $\boldsymbol{y}(t)$ 和 $\boldsymbol{v}(t)$ 各为对 $\boldsymbol{s}(t)$ 的观测信号和观测噪声. 假设 $w(t)$ 和 $\boldsymbol{v}(t)$ 是零均值、方差阵各为 Q_w 和 $\boldsymbol{Q}_v$ 的相互独立的高斯白噪声,且
$$Q_w = \sigma_w^2, \quad \boldsymbol{Q}_v = \mathrm{diag}(\sigma_{v1}^2, \sigma_{v2}^2) \tag{5.3.85}$$
取 $N = 0$ 和 $N = -1$,问题是求最优 Wiener 滤波器 $\hat{\boldsymbol{s}}(t \mid t)$ 和预报器 $\hat{\boldsymbol{s}}(t \mid t - 1)$. 在仿真中取 $T_0 = 0.3$, $\sigma_w^2 = 6$,且 $\boldsymbol{Q}_v = \mathrm{diag}(5.6, 5.04)$, $\boldsymbol{b} = [0.3, 0]^{\mathrm{T}}$,可得 ARMA 新息模型
$$(\boldsymbol{I}_2 + \boldsymbol{A}_1 q^{-1})\boldsymbol{z}(t) = (\boldsymbol{I}_2 + \boldsymbol{D}_1 q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.3.86}$$
其中 $\boldsymbol{z}(t) = \boldsymbol{y}(t) - \boldsymbol{b}$,且
$$\boldsymbol{D}_1 = \begin{bmatrix} -0.7927 & -0.090997 \\ 0.096574 & -0.7932 \end{bmatrix}, \quad \boldsymbol{Q}_\varepsilon = \begin{bmatrix} 6.9671 & 0.84825 \\ 0.84825 & 6.4573 \end{bmatrix} \tag{5.3.87}$$
且应用定理 5.3.3 求得 $\hat{\boldsymbol{s}}(t \mid t)$ 和 $\hat{\boldsymbol{s}}(t \mid t - 1)$,且可求得相应的估计误差方差阵的迹为
$$\mathrm{tr}\boldsymbol{P}(0) = 2.0679, \quad \mathrm{tr}\boldsymbol{P}(-1) = 2.7843 \tag{5.3.88}$$
可见 $\mathrm{tr}\boldsymbol{P}(0) < \mathrm{tr}\boldsymbol{P}(-1)$. 这从理论上证明了 $\hat{\boldsymbol{s}}(t \mid t)$ 的精度比 $\hat{\boldsymbol{s}}(t \mid t - 1)$ 的精度高. 仿真结果如图 5.3.1 和图 5.3.2 所示,其中实线代表信号真实值,虚线代表估值 $\hat{\boldsymbol{s}}(t \mid t)$ 或 $\hat{\boldsymbol{s}}(t \mid t - 1)$. 我们可直观看到 $\hat{\boldsymbol{s}}(t \mid t)$ 的精度比 $\hat{\boldsymbol{s}}(t \mid t - 1)$ 高.



(a) 位置 $s_1(t)$ 和 Wiener 跟踪预报器 $\hat{s}_1(t|t)$

(b) 速度 $s_2(t)$ 和 Wiener 跟踪预报器 $\hat{s}_2(t|t)$

图 5.3.1　信号 $\boldsymbol{s}(t)$ 和 Wiener 跟踪滤波器 $\hat{\boldsymbol{s}}(t \mid t)$



(a) 位置 $s_1(t)$ 和 Wiener 跟踪预报器 $\hat{s}_1(t|t-1)$

(b) 速度 $s_2(t)$ 和 Wiener 跟踪预报器 $\hat{s}_2(t|t-1)$

图 5.3.2　信号 $\boldsymbol{s}(t)$ 和 Wiener 跟踪预报波器 $\hat{\boldsymbol{s}}(t \mid t-1)$

5.4　带 MA 有色观测噪声的多通道 ARMA 信号 Wiener 滤波器

本节推广节 5.3 的结果到带有色观测噪声情形. 对于带 MA 有色观测噪声情形，可直接利用白噪声估值器解决 Wiener 滤波问题. 考虑带 MA 有色观测噪声和带白噪声的多通道 ARMA 信号

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.4.1}$$

$$\boldsymbol{y}(t) = \boldsymbol{s}(t) + \boldsymbol{\eta}(t) + \boldsymbol{v}(t) \tag{5.4.2}$$

$$\boldsymbol{\eta}(t) = \boldsymbol{P}(q^{-1})\boldsymbol{\xi}(t) \tag{5.4.3}$$

其中观测 $\boldsymbol{y}(t) \in R^m$，待估信号 $\boldsymbol{s}(t) \in R^m$，$\boldsymbol{v}(t) \in R^m$ 是白色观测噪声，$\boldsymbol{\eta}(t) \in R^m$ 是有色观测噪声，服从 MA 模型 (5.4.3)，$\boldsymbol{w}(t) \in R^r$，$\boldsymbol{v}(t) \in R^m$ 和 $\boldsymbol{\xi}(t) \in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$，$\boldsymbol{Q}_v$，$\boldsymbol{Q}_\xi$ 的相互独立白噪声. $\boldsymbol{A}(q^{-1})$，$\boldsymbol{C}(q^{-1})$ 和 $\boldsymbol{P}(q^{-1})$ 为 q^{-1} 的多项式矩阵，形如 $\boldsymbol{X}(q^{-1}) = \boldsymbol{X}_0 + \boldsymbol{X}_1 q^{-1} + \cdots + \boldsymbol{X}_{n_x} q^{-n_x}$，且 $\boldsymbol{A}_0 = \boldsymbol{I}_m$，$\boldsymbol{C}_0 = \boldsymbol{0}$，$\boldsymbol{P}_0 = \boldsymbol{I}_m$，假设 $(\boldsymbol{A}(q^{-1}), \boldsymbol{C}(q^{-1}))$ 左素，初始观测时刻 $t_0 = -\infty$. 问题是基于观测 $(\boldsymbol{y}(t+N), \boldsymbol{y}(t+N-1), \cdots)$ 求信号 $\boldsymbol{s}(t)$ 的 Wiener 滤波器 $\hat{\boldsymbol{s}}(t \mid t+N)$.

将(5.4.1)和(5.4.3)代入(5.4.2)可得

$$\boldsymbol{y}(t)=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{P}(q^{-1})\boldsymbol{\xi}(t)+\boldsymbol{v}(t) \tag{5.4.4}$$

上式两边左乘 $\boldsymbol{A}(q^{-1})$ 引出 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.4.5}$$

其中 $\boldsymbol{D}(q^{-1})$ 是稳定的,$\boldsymbol{D}_0=\boldsymbol{I}_m$,新息 $\boldsymbol{\varepsilon}(t)\in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声,且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{A}(q^{-1})\boldsymbol{P}(q^{-1})\boldsymbol{\xi}(t)+\boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.4.6}$$

其中 $\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers - Wouters 算法求得.

【引理 5.4.1】 白噪声 Wiener 滤波器为

$$\hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{L}_N^v(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.4.7}$$

$$\hat{\boldsymbol{\xi}}(t\mid t+N)=\boldsymbol{L}_N^\xi(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.4.8}$$

带伪交换

$$\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})=\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{5.4.9}$$

带 $\widetilde{\boldsymbol{A}}_0=\boldsymbol{I}_m,\widetilde{\boldsymbol{D}}_0=\boldsymbol{I}_m,\det\widetilde{\boldsymbol{D}}(q^{-1})=\det\boldsymbol{D}(q^{-1})$,且

$$\begin{aligned}&\boldsymbol{L}_N^v(q^{-1})=\sum_{i=0}^{N}\boldsymbol{Q}_v\boldsymbol{G}_i^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}q^{i-N},\quad N\geqslant 0\\&\boldsymbol{L}_N^v(q^{-1})=\boldsymbol{0},\quad N<0\end{aligned} \tag{5.4.10}$$

$$\begin{aligned}&\boldsymbol{L}_N^\xi(q^{-1})=\sum_{i=0}^{N}\boldsymbol{Q}_\xi\boldsymbol{\Pi}_i^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}q^{i-N},\quad N\geqslant 0\\&\boldsymbol{L}_N^\xi(q^{-1})=\boldsymbol{0},\quad N<0\end{aligned} \tag{5.4.11}$$

其中系数阵 $\boldsymbol{G}_i$ 和 $\boldsymbol{\Pi}_i$ 可递推计算为

$$\boldsymbol{G}_i=-\boldsymbol{D}_1\boldsymbol{G}_{i-1}-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{G}_{i-n_d}+\boldsymbol{A}_i \tag{5.4.12}$$

$$\boldsymbol{\Pi}_i=-\boldsymbol{D}_1\boldsymbol{\Pi}_{i-1}-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{\Pi}_{i-n_d}+\bar{\boldsymbol{A}}_i \tag{5.4.13}$$

其中定义 $\bar{\boldsymbol{A}}(q^{-1})=\boldsymbol{A}(q^{-1})\boldsymbol{P}(q^{-1})$,且规定:$\boldsymbol{G}_i=\boldsymbol{0}(i<0)$,$\boldsymbol{A}_i=\boldsymbol{0}(i>n_a)$;$\boldsymbol{\Pi}_i=\boldsymbol{0}(i<0)$,$\bar{\boldsymbol{A}}_i=\boldsymbol{0}(i>n_a)$.白噪声新息滤波器为

$$\hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{L}_N^v(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.4.14}$$

$$\hat{\boldsymbol{\xi}}(t\mid t+N)=\boldsymbol{L}_N^\xi(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.4.15}$$

【引理 5.4.2】 下述公式成立

$$\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]=\boldsymbol{Q}_v\boldsymbol{G}_{i-t}^{\mathrm{T}} \tag{5.4.16}$$

$$\mathrm{E}[\boldsymbol{\xi}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]=\boldsymbol{Q}_\xi\boldsymbol{\Pi}_{i-t}^{\mathrm{T}} \tag{5.4.17}$$

【定理 5.4.1】 带 MA 有色观测噪声的多通道 ARMA 信号(5.4.1)~(5.4.3)有渐近稳定的 Wiener 滤波器

$$\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.4.18}$$

其中定义多项式矩阵 $\boldsymbol{K}_N(q^{-1})$ 为

$$\boldsymbol{K}_N(q^{-1})=\boldsymbol{J}_{-N}(q^{-1})-\sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{L}_{N+i}^\xi(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})-\boldsymbol{L}_N^v(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) \tag{5.4.19}$$

或表为 ARMA 递推形式

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\operatorname{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t+N) \tag{5.4.20}$$

其中 $\boldsymbol{J}_{-N}(q^{-1})$ 由推论 5.3.2 计算，而 $\boldsymbol{L}_{N+i}^{\xi}(q^{-1})$，$\boldsymbol{L}_N^{v}(q^{-1})$，$\widetilde{\boldsymbol{A}}(q^{-1})$，$\widetilde{\boldsymbol{D}}(q^{-1})$ 由引理 5.4.1 计算. 估值误差 $\tilde{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{s}(t)-\hat{\boldsymbol{s}}(t\mid t+N)$，当 $N\geqslant 0$ 时有表达式

$$\tilde{\boldsymbol{s}}(t\mid t+N)=-\sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{\xi}(t-i)+\sum_{i=0}^{n_p}\boldsymbol{P}_i\hat{\boldsymbol{\xi}}(t-i\mid t+N)-\boldsymbol{v}(t)+\hat{\boldsymbol{v}}(t\mid t+N),\quad N\geqslant 0 \tag{5.4.21}$$

它可表为

$$\tilde{\boldsymbol{s}}(t\mid t+N)=\sum_{j=0}^{N+n_p}[\boldsymbol{\alpha}_j\quad\boldsymbol{\beta}_j\quad\boldsymbol{\gamma}_j]\begin{bmatrix}\boldsymbol{\xi}(t-j)\\ \boldsymbol{v}(t-j)\\ \boldsymbol{\varepsilon}(t-n_p+j)\end{bmatrix},\quad N\geqslant 0 \tag{5.4.22}$$

其中应用引理 5.4.1 合并了 $\hat{\boldsymbol{\xi}}(t-i\mid t+N)$，$\hat{\boldsymbol{v}}(t\mid t+N)$ 中 $\boldsymbol{\varepsilon}(t-n_p+j)$ 的同类项可得系数阵 $\boldsymbol{\gamma}_j$，且定义 $\boldsymbol{\alpha}_j=-\boldsymbol{P}_j$，$\boldsymbol{P}_j=\boldsymbol{0}(j>n_p)$，$\boldsymbol{\beta}_0=\boldsymbol{I}_m$，$\boldsymbol{\beta}_j=\boldsymbol{0}(j>0)$. 当 $N<0$ 时有表达式

$$\tilde{\boldsymbol{s}}(t\mid t+N)=\tilde{\boldsymbol{y}}(t\mid t+N)-\sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{\xi}(t-i)+\sum_{i=0}^{n_p}\boldsymbol{P}_i\hat{\boldsymbol{\xi}}(t-i\mid t+N)-\boldsymbol{v}(t),\quad N<0 \tag{5.4.23}$$

应用(5.3.28)和(5.4.15)合并 $\boldsymbol{\varepsilon}(t-i)$ 的同类项引出表达式

$$\tilde{\boldsymbol{s}}(t\mid t+N)=-\sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{\xi}(t-i)-\boldsymbol{v}(t)+\sum_{i=0}^{n_h}\boldsymbol{h}_i\boldsymbol{\varepsilon}(t-i),\quad N<0 \tag{5.4.24}$$

其中 $n_h=\max(-N-1,n_p)$. 它可表为

$$\tilde{\boldsymbol{s}}(t\mid t+N)=\sum_{j=0}^{n_h}[\boldsymbol{\alpha}_j\quad\boldsymbol{\beta}_j\quad\boldsymbol{h}_j]\begin{bmatrix}\boldsymbol{\xi}(t-j)\\ \boldsymbol{v}(t-j)\\ \boldsymbol{\varepsilon}(t-j)\end{bmatrix},\quad N<0 \tag{5.4.25}$$

其中定义 $\boldsymbol{\alpha}_j=-\boldsymbol{P}_j$，$\boldsymbol{P}_j=\boldsymbol{0}(j>n_p)$；$\boldsymbol{\beta}_0=\boldsymbol{I}_m$，$\boldsymbol{\beta}_j=\boldsymbol{0}(j>0)$；$\boldsymbol{h}_j$ 由合并 $\boldsymbol{\varepsilon}(t-j)$ 的同类项得到. 估值误差方差阵 $\boldsymbol{P}(N)=\mathrm{E}[\tilde{\boldsymbol{s}}(t\mid t+N)\tilde{\boldsymbol{s}}^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}(N)=\sum_{i=0}^{N+n_p}\sum_{j=0}^{N+n_p}[\boldsymbol{\alpha}_i\quad\boldsymbol{\beta}_i\quad\boldsymbol{\gamma}_i]\boldsymbol{Q}\begin{bmatrix}\boldsymbol{\alpha}_j^{\mathrm{T}}\\ \boldsymbol{\beta}_j^{\mathrm{T}}\\ \boldsymbol{\gamma}_j^{\mathrm{T}}\end{bmatrix},\quad N\geqslant 0 \tag{5.4.26}$$

$$\boldsymbol{Q}=\begin{bmatrix}\boldsymbol{Q}_\xi\delta_{ij} & \boldsymbol{0} & \boldsymbol{Q}_\xi\boldsymbol{\Pi}_{i+j-n_p}^{\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{Q}_v\delta_{ij} & \boldsymbol{Q}_v\boldsymbol{G}_{i+j-n_p}^{\mathrm{T}}\\ \boldsymbol{\Pi}_{i+j-n_p}\boldsymbol{Q}_\xi & \boldsymbol{G}_{i+j-n_p}\boldsymbol{Q}_v & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix} \tag{5.4.27}$$

$$\boldsymbol{P}(N)=\sum_{i=0}^{n_h}\sum_{j=0}^{n_h}[\boldsymbol{\alpha}_i\quad\boldsymbol{\beta}_i\quad\boldsymbol{h}_i]\boldsymbol{Q}\begin{bmatrix}\boldsymbol{\alpha}_j^{\mathrm{T}}\\ \boldsymbol{\beta}_j^{\mathrm{T}}\\ \boldsymbol{\gamma}_j^{\mathrm{T}}\end{bmatrix},\quad N<0 \tag{5.4.28}$$

$$\boldsymbol{Q}=\begin{bmatrix}\boldsymbol{Q}_\xi\delta_{ij} & \boldsymbol{0} & \boldsymbol{Q}_\xi\boldsymbol{\Pi}_{i-j}^{\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{Q}_v\delta_{ij} & \boldsymbol{Q}_v\boldsymbol{G}_{i-j}^{\mathrm{T}}\\ \boldsymbol{\Pi}_{j-i}\boldsymbol{Q}_\xi & \boldsymbol{G}_{j-i}\boldsymbol{Q}_v & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix} \tag{5.4.29}$$

证明　由(5.4.2)有关系

$$\boldsymbol{s}(t)=\boldsymbol{y}(t)-\boldsymbol{\eta}(t)-\boldsymbol{v}(t) \tag{5.4.30}$$

$$\hat{\boldsymbol{s}}(t\mid t+N)=\hat{\boldsymbol{y}}(t\mid t+N)-\hat{\boldsymbol{\eta}}(t\mid t+N)-\hat{\boldsymbol{v}}(t\mid t+N) \tag{5.4.31}$$

由(5.4.3)有

$$\boldsymbol{\eta}(t)=\sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{\xi}(t-i) \tag{5.4.32}$$

应用引理5.4.1有

$$\hat{\boldsymbol{\eta}}(t\mid t+N)=\sum_{i=0}^{n_p}\boldsymbol{P}_i\hat{\boldsymbol{\xi}}(t-i\mid t+N)=\sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{L}_{N+i}^{\xi}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.4.33}$$

将(5.4.33),(5.3.24)和(5.4.7)代入(5.4.31)得(5.4.18)和(5.4.19),进而有(5.4.20). 将(5.4.30)与(5.4.31)相减引出估值误差关系

$$\tilde{\boldsymbol{s}}(t\mid t+N)=\tilde{\boldsymbol{y}}(t\mid t+N)-\tilde{\boldsymbol{\eta}}(t\mid t+N)-\tilde{\boldsymbol{v}}(t\mid t+N) \tag{5.4.34}$$

即

$$\tilde{\boldsymbol{s}}(t\mid t+N)=\tilde{\boldsymbol{y}}(t\mid t+N)-\sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{\xi}(t-i)+\sum_{i=0}^{n_p}\boldsymbol{P}_i\hat{\boldsymbol{\xi}}(t-i\mid t+N)-\boldsymbol{v}(t)+\hat{\boldsymbol{v}}(t\mid t+N) \tag{5.4.35}$$

而由(5.4.15)有

$$\hat{\boldsymbol{\xi}}(t-i\mid t+N)=\boldsymbol{L}_{N+i}^{\xi}(q^{-1})\boldsymbol{\varepsilon}(t+N)=\sum_{j=0}^{N+i}\boldsymbol{Q}_{\xi}\boldsymbol{\Pi}_j^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon}^{-1}q^{j-(N+i)}\boldsymbol{\varepsilon}(t+N) \tag{5.4.36}$$

当 $N\geqslant 0$ 时,$\tilde{\boldsymbol{y}}(t\mid t+N)=\boldsymbol{y}(t)-\hat{\boldsymbol{y}}(t\mid t+N)=\boldsymbol{y}(t)-\boldsymbol{y}(t)=\boldsymbol{0}$,合并 $\boldsymbol{\varepsilon}(t-n_p+i)$ 的同类项得(5.4.21)和(5.4.22). 当 $N<0$ 时,应用(5.3.29)有 $\tilde{\boldsymbol{y}}(t\mid t+N)$ 为

$$\tilde{\boldsymbol{y}}(t\mid t+N)=\boldsymbol{E}_{-N}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.4.37}$$

它是 $\boldsymbol{\varepsilon}(t),\boldsymbol{\varepsilon}(t-1),\cdots\boldsymbol{\varepsilon}(t+N-1)$ 的线性组合,当 $n_p\geqslant -N$,(5.4.35)的右边第三项是 $\boldsymbol{\varepsilon}(t-n_p),\cdots,\boldsymbol{\varepsilon}(t+N)$ 的线性组合,于是得(5.4.24)和(5.4.25). 进而应用引理5.4.2可得(5.4.26)~(5.4.29). 证毕. □

【推论5.4.1】　带MA有色观测噪声的单通道($m=1$)ARMA信号(5.4.1)~(5.4.3)有渐近稳定的Wiener滤波器

$$D(q^{-1})\hat{s}(t\mid t+N)=K_N(q^{-1})y(t+N) \tag{5.4.38}$$

其中定义多面式 $K_N(q^{-1})$ 为

$$K_N(q^{-1})=J_{-N}(q^{-1})-\sum_{i=0}^{n_p}P_iL_{N+i}^{\xi}(q^{-1})A(q^{-1})-L_N^{v}(q^{-1})A(q^{-1}) \tag{5.4.39}$$

且估值误差方差由(5.4.26)~(5.4.29)计算.

【推论5.4.2】　对带MA有色观测噪声的多通道ARMA信号

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.4.40}$$

$$\boldsymbol{y}(t)=\boldsymbol{s}(t)+\boldsymbol{\eta}(t) \tag{5.4.41}$$

$$\boldsymbol{\eta}(t)=\boldsymbol{P}(q^{-1})\boldsymbol{\xi}(t) \tag{5.4.42}$$

其中有关假设同对(5.4.1)～(5.4.3)的假设,有渐近稳定的 Wiener 滤波器

$$\hat{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.4.43}$$

或

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\mathrm{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t+N) \tag{5.4.44}$$

其中定义

$$\boldsymbol{K}_N(q^{-1}) = \boldsymbol{J}_{-N}(q^{-1}) - \sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{L}_{N+i}^{\xi}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) \tag{5.4.45}$$

且有估值误差为

$$\tilde{\boldsymbol{s}}(t \mid t+N) = -\sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{\xi}(t-i) + \sum_{i=0}^{n_p}\boldsymbol{P}_i\hat{\boldsymbol{\xi}}(t-i \mid t+N), \quad N \geqslant 0 \tag{5.4.46}$$

$$\tilde{\boldsymbol{s}}(t \mid t+N) = \tilde{\boldsymbol{y}}(t \mid t+N) - \sum_{i=0}^{n_p}\boldsymbol{P}_i\boldsymbol{\xi}(t-i) + \sum_{i=0}^{n_p}\boldsymbol{P}_i\hat{\boldsymbol{\xi}}(t-i \mid t+N), \quad N < 0 \tag{5.4.47}$$

合并同类项有表达式

$$\tilde{\boldsymbol{s}}(t \mid t+N) = \sum_{j=0}^{N+n_p}[\boldsymbol{\alpha}_j \quad \boldsymbol{\beta}_j]\begin{bmatrix}\boldsymbol{\xi}(t-j)\\ \boldsymbol{\varepsilon}(t-n_p+j)\end{bmatrix}, \quad N \geqslant 0 \tag{5.4.48}$$

其中规定:$\boldsymbol{\alpha}_j = -\boldsymbol{P}_j$,$\boldsymbol{\alpha}_j = \boldsymbol{0}(j > n_p)$,$\boldsymbol{\beta}_j$ 由合并 $\boldsymbol{\varepsilon}(t-n_p+j)$ 的同类项系数阵得到. 同理有

$$\tilde{\boldsymbol{s}}(t \mid t+N) = \sum_{j=0}^{n_h}[\boldsymbol{\alpha}_j \quad \boldsymbol{h}_j]\begin{bmatrix}\boldsymbol{\xi}(t-j)\\ \boldsymbol{\varepsilon}(t-j)\end{bmatrix}, \quad N < 0 \tag{5.4.49}$$

其中定义 $\boldsymbol{\alpha}_j = -\boldsymbol{P}_j$,$\boldsymbol{\alpha}_j = \boldsymbol{0}(j > n_p)$,$\boldsymbol{h}_j$ 由合并 $\boldsymbol{\varepsilon}(t-j)$ 系数阵得到. 估值误差方差阵为

$$\boldsymbol{P}(N) = \sum_{i=0}^{N+n_p}\sum_{j=0}^{N+n_p}[\boldsymbol{\alpha}_i \quad \boldsymbol{\beta}_i]\begin{bmatrix}\boldsymbol{Q}_\xi\delta_{ij} & \boldsymbol{Q}_\xi\boldsymbol{\Pi}_{i+j-n_p}^{\mathrm{T}}\\ \boldsymbol{\Pi}_{i+j-n_p}\boldsymbol{Q}_\xi & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\alpha}_j^{\mathrm{T}}\\ \boldsymbol{\beta}_j^{\mathrm{T}}\end{bmatrix}, \quad N \geqslant 0 \tag{5.4.50}$$

$$\boldsymbol{P}(N) = \sum_{i=0}^{n_h}\sum_{j=0}^{n_h}[\boldsymbol{\alpha}_i \quad \boldsymbol{h}_i]\begin{bmatrix}\boldsymbol{Q}_\xi\delta_{ij} & \boldsymbol{Q}_\xi\boldsymbol{\Pi}_{i-j}^{\mathrm{T}}\\ \boldsymbol{\Pi}_{j-i}\boldsymbol{Q}_\xi & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\alpha}_j^{\mathrm{T}}\\ \boldsymbol{h}_j^{\mathrm{T}}\end{bmatrix}, \quad N < 0 \tag{5.4.51}$$

【推论 5.4.3】 对带 MA 有色观测噪声的单通道($m=1$)ARMA 信号(5.4.40)～(5.4.42),分别记 $\boldsymbol{P}_i$,$\boldsymbol{Q}_\xi$,$\boldsymbol{Q}_\varepsilon$,$\boldsymbol{\Pi}_i$ 为 p_i,σ_ξ^2,σ_ε^2,π_i,则有渐近稳定的 Wiener 滤波器为

$$D(q^{-1})\hat{s}(t \mid t+N) = K_N(q^{-1})y(t+N) \tag{5.4.52}$$

其中定义多项式

$$K_N(q^{-1}) = J_{-N}(q^{-1}) - \sum_{i=0}^{n_p}p_iL_{N+i}^{\xi}(q^{-1})A(q^{-1}) \tag{5.4.53}$$

且有估值误差为

$$\tilde{s}(t \mid t+N) = \sum_{j=0}^{n_p}p_i\xi(t-j) + \sum_{j=0}^{N+n_p}\beta_i\varepsilon(t-n_p+j), \quad N \geqslant 0 \tag{5.4.54}$$

$$\tilde{s}(t \mid t+N) = \sum_{j=0}^{n_p}p_i\xi(t-j) + \sum_{j=0}^{n_h}h_j\varepsilon(t-j), \quad N < 0 \tag{5.4.55}$$

带 $n_h = \max(-N-1, n_p)$,其中 β_j,h_j 由合并同类项系数得到. 相应的估值误差方差为

$$P(N)=\sum_{j=0}^{n_p}p_j^2\sigma_\xi^2+\sum_{j=0}^{N+n_p}\beta_j^2\sigma_\varepsilon^2-2\sum_{j=0}^{n_p}\sum_{i=0}^{N+n_p}p_j\beta_i\sigma_\xi^2\pi_{i+j-n_p},\quad N\geqslant 0 \tag{5.4.56}$$

$$P(N)=\sum_{j=0}^{n_p}p_j^2\sigma_\xi^2+\sum_{j=0}^{n_h}h_j^2\sigma_\varepsilon^2-2\sum_{j=0}^{n_p}\sum_{i=0}^{n_h}p_jh_i\sigma_\xi^2\pi_{j-i},\quad N<0 \tag{5.4.57}$$

【例 5.4.1】 考虑系统

$$(1+aq^{-1})s(t)=w(t-1) \tag{5.4.58}$$

$$y(t)=s(t)+b+v(t) \tag{5.4.59}$$

$$v(t)=(1+pq^{-1})\xi(t) \tag{5.4.60}$$

其中 $s(t)$ 为待估信号，b 为观测系统偏差，$y(t)$ 为对 $s(t)$ 的观测信号，$v(t)$ 为 MA 有色观测噪声，$\xi(t)$ 和 $w(t)$ 是零均值、方差各为 σ_ξ^2 和 σ_w^2 的相互独立高斯白噪声，问题是求 Wiener 平滑器 $\hat{s}(t\mid t+2)$ 和预报器 $\hat{s}(t\mid t+1)$.

引入新的观测信号 $z(t)=y(t)-b$，则有观测方程 $z(t)=s(t)+v(t)$，于是系统(5.4.58)～(5.4.60)可化为系统(5.4.40)～(5.4.42)的形式. 仿真中取 $a=-0.95$，$b=0.2$，$\sigma_w^2=0.1$，$\sigma_\xi^2=1$，$p=0.98$. 可得 ARMA 新息模型

$$(1-0.95q^{-1})z(t)=(1+0.005\,582\,4q^{-1}-0.715\,16q^{-2})\varepsilon(t) \tag{5.4.61}$$

其中新息方差 $\sigma_\varepsilon^2=1.301\,8$. 应用推论 5.4.3 可求得 $\hat{s}(t\mid t+2)$ 和 $\hat{s}(t\mid t+1)$，且可求得估值误差方差各为

$$P(2)=0.360\,34,\quad P(-1)=1.809\,9 \tag{5.4.62}$$

于是 $P(2)<P(-1)$. 这说明 $\hat{s}(t\mid t+2)$ 的精度比 $\hat{s}(t\mid t-1)$ 高. 仿真结果如图 5.4.1 和图 5.4.2 所示，其中实线表示真实值 $s(t)$，虚线表示估值 $\hat{s}(t\mid t+2)$ 或 $\hat{s}(t\mid t-1)$. 比较这两个图形也可看到 $\hat{s}(t\mid t+2)$ 的精度高于 $\hat{s}(t\mid t-1)$ 的精度.

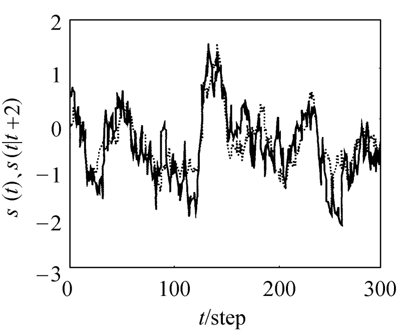


图 5.4.1 信号 $s(t)$ 和 Wiener 平滑器 $\hat{s}(t\mid t+2)$

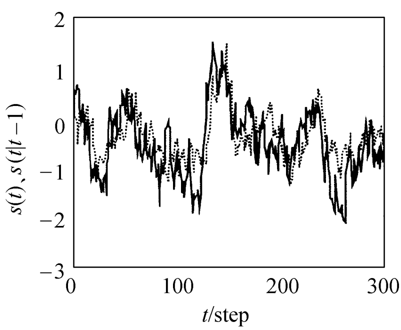


图 5.4.2 信号 $s(t)$ 和 Wiener 预报器 $\hat{s}(t\mid t-1)$

5.5 多通道 ARMA 信号 Wiener 反卷积滤波器

估计一个随机系统的输入信号叫反卷积(Deconvolution)或输入估计，广泛应用于信号处理、通信与控制领域，文献[13,16,19,25,29,31,71,79,87,88]已用 Kalman 滤波方法[13]，或多项式方法[79,87,88]，或现代时间序列分析方法[19]提出了不同的反卷积滤波器. 本节介绍文献[1]用现代时间序列分析方法，基于 ARMA 新息模型和白噪声估值器提出

的 Wiener 反卷积滤波器，并提出了估值误差方差阵的计算公式.

考虑多通道 ARMA 信号 Wiener 反卷积问题：

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.5.1}$$

$$\boldsymbol{P}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{B}(q^{-1})\boldsymbol{s}(t)+\boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \tag{5.5.2}$$

其中 $\boldsymbol{s}(t)\in R^m$ 为待估的ARMA信号，它作为输入被动态系统(5.5.2)观测，$\boldsymbol{y}(t)\in R^m$ 是观测信号，$\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}(t)\in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$ 的相互独立白噪声. q^{-1} 为单位滞后算子，多项式矩阵 $\boldsymbol{A}(q^{-1}),\cdots,\boldsymbol{R}(q^{-1})$ 有形式 $\boldsymbol{X}(q^{-1})=\boldsymbol{X}_0+\boldsymbol{X}_1q^{-1}+\cdots+\boldsymbol{X}_{n_x}q^{-n_x}$，且 $\boldsymbol{A}_0=\boldsymbol{I}_m,\boldsymbol{P}_0=\boldsymbol{I}_m$. 设初始观测时刻 $t_0=-\infty$，问题是基于观测 $(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots)$ 求 $\boldsymbol{s}(t)$ 的 Wiener 反卷积滤波器 $\hat{\boldsymbol{s}}(t\mid t+N)$. 对 $N=0,N>0$ 或 $N<0$，各称其为反卷积滤波器、平滑器或预报器.

将(5.5.1)代入(5.5.2)引出

$$\boldsymbol{y}(t)=\boldsymbol{P}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{P}^{-1}(q^{-1})\boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \tag{5.5.3}$$

引入左素分解

$$[\boldsymbol{P}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1}),\boldsymbol{P}^{-1}(q^{-1})\boldsymbol{R}(q^{-1})]=\boldsymbol{\Lambda}^{-1}(q^{-1})[\bar{\boldsymbol{B}}(q^{-1}),\bar{\boldsymbol{R}}(q^{-1})] \tag{5.5.4}$$

$$\boldsymbol{y}(t)=[\boldsymbol{P}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1}),\boldsymbol{P}^{-1}(q^{-1})\boldsymbol{R}(q^{-1})]\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix}=$$

$$\boldsymbol{\Lambda}^{-1}(q^{-1})[\bar{\boldsymbol{B}}(q^{-1}),\bar{\boldsymbol{R}}(q^{-1})]\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix} \tag{5.5.5}$$

这引出

$$\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t)=\bar{\boldsymbol{B}}(q^{-1})\boldsymbol{w}(t)+\bar{\boldsymbol{R}}(q^{-1})\boldsymbol{v}(t) \tag{5.5.6}$$

假设[1]$(\bar{\boldsymbol{B}}(q^{-1}),\bar{\boldsymbol{R}}(q^{-1}))$ 左素或无行列式的零点在单位圆上的左因式，则有 ARMA 新息模型

$$\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.5.7}$$

其中 $\boldsymbol{D}(q^{-1})$ 是稳定的，$\boldsymbol{D}_0=\boldsymbol{I}_m$，新息 $\boldsymbol{\varepsilon}(t)\in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\bar{\boldsymbol{B}}(q^{-1})\boldsymbol{w}(t)+\bar{\boldsymbol{R}}(q^{-1})\boldsymbol{v}(t) \tag{5.5.8}$$

【引理 5.5.1】 白噪声 Wiener 滤波器为

$$\hat{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{L}_N^w(q^{-1})\tilde{\boldsymbol{\Lambda}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.5.9}$$

$$\hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{L}_N^v(q^{-1})\tilde{\boldsymbol{\Lambda}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.5.10}$$

或

$$\hat{\boldsymbol{w}}(t\mid t+N)=\boldsymbol{L}_N^w(\varepsilon^{-1})\boldsymbol{\varepsilon}(t+N),\quad \hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{L}_N^v(\varepsilon^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.5.11}$$

带伪交换

$$\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{\Lambda}(q^{-1})=\tilde{\boldsymbol{\Lambda}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{5.5.12}$$

$\tilde{\boldsymbol{D}}_0=\boldsymbol{I}_m,\tilde{\boldsymbol{\Lambda}}_0=\boldsymbol{I}_m,\det\boldsymbol{D}(q^{-1})=\det\tilde{\boldsymbol{D}}(q^{-1}),n_d=n_{\tilde{d}},n_\lambda=n_{\tilde{\lambda}}$.

$$\boldsymbol{L}_N^w(q^{-1})=\sum_{i=0}^{N}\boldsymbol{Q}_w\boldsymbol{F}_i^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}q^{i-N},\quad N\geqslant 0$$

$$\boldsymbol{L}_N^v(q^{-1})=\sum_{i=0}^{N}\boldsymbol{Q}_v\boldsymbol{G}_i^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}q^{i-N},\quad N\geqslant 0$$

$$L_N^w(q^{-1}) = \mathbf{0}, \quad N < 0; \quad L_N^v(q^{-1}) = \mathbf{0}, \quad N < 0 \tag{5.5.13}$$

$\boldsymbol{F}_i$ 和 $\boldsymbol{G}_i$ 可递推计算为

$$\begin{aligned} \boldsymbol{F}_i &= -\boldsymbol{D}_1\boldsymbol{F}_{i-1} - \cdots - \boldsymbol{D}_{n_d}\boldsymbol{F}_{i-n_d} + \bar{\boldsymbol{B}}_i, \\ \boldsymbol{G}_i &= -\boldsymbol{D}_1\boldsymbol{G}_{i-1} - \cdots - \boldsymbol{D}_{n_d}\boldsymbol{G}_{i-n_d} + \bar{\boldsymbol{R}}_i \end{aligned} \tag{5.5.14}$$

其中规定 $\boldsymbol{F}_i = \mathbf{0}(i < 0)$，$\boldsymbol{G}_i = \mathbf{0}(i < 0)$，$\bar{\boldsymbol{B}}_i = \mathbf{0}(i < n_{\bar{b}})$，$\bar{\boldsymbol{R}}_i = \mathbf{0}(i < n_{\bar{r}})$.

【引理 5.5.2】 Åström 预报器为

$$\hat{\boldsymbol{y}}(t \mid t+N) = \boldsymbol{J}_{-N}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.5.15}$$

其中 $\boldsymbol{J}_{-N}(q^{-1})$ 定义为

$$\begin{aligned} \widetilde{\boldsymbol{D}}(q^{-1}) &= \boldsymbol{E}_{-N}(q^{-1})\widetilde{\boldsymbol{\Lambda}}(q^{-1}) + q^N\boldsymbol{J}_{-N}(q^{-1}), \quad N < 0 \\ \boldsymbol{J}_{-N}(q^{-1}) &= \widetilde{\boldsymbol{D}}(q^{-1})q^{-N}, \quad N \geqslant 0 \end{aligned} \tag{5.5.16}$$

预报误差 $\tilde{\boldsymbol{y}}(t \mid t+N)$ 计算为

$$\tilde{\boldsymbol{y}}(t \mid t+N) = \boldsymbol{y}(t) - \boldsymbol{y}(t \mid t+N) = \boldsymbol{E}_{-N}(q^{-1})\varepsilon(t) \tag{5.5.17}$$

为了求 $\hat{\boldsymbol{s}}(t \mid t+N)$，我们希望将问题归结为求白噪声估值器和预测预报器. 由 (5.5.1) 和 (5.5.2) 有关系

$$\boldsymbol{\Omega}\begin{bmatrix} \boldsymbol{s}(t) \\ \boldsymbol{s}(t-1) \\ \vdots \\ \vdots \\ \vdots \\ \boldsymbol{s}(t-n_a-n_b+1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \\ \vdots \\ \boldsymbol{C}(q^{-1})\boldsymbol{w}(t-n_b+1) \\ \boldsymbol{P}(q^{-1})\boldsymbol{y}(t) - \boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \\ \vdots \\ \boldsymbol{P}(q^{-1})\boldsymbol{y}(t-n_a+1) - \boldsymbol{R}(q^{-1})\boldsymbol{v}(t-n_a+1) \end{bmatrix} \tag{5.5.18}$$

其中 $\boldsymbol{\Omega}$ 是如下 $(n_a+n_b)m \times (n_a+n_b)m$ 矩阵：

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots\boldsymbol{A}_{n_a-1} & \boldsymbol{A}_{n_a} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_m & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots\boldsymbol{A}_{n_a-1} & \boldsymbol{A}_{n_a} & & \cdots \\ & \vdots & & & & & & \\ \mathbf{0} & \cdots\mathbf{0} & \boldsymbol{I}_m & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & & \boldsymbol{A}_{n_a} \\ \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots\boldsymbol{B}_{n_b-1} & \boldsymbol{B}_{n_b} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots\boldsymbol{B}_{n_b-1} & \boldsymbol{B}_{n_b} & & \cdots \\ & \vdots & & & & & & \\ \mathbf{0} & \cdots\mathbf{0} & \boldsymbol{B}_0 & \boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & & \boldsymbol{B}_{n_b} \end{bmatrix} \tag{5.5.19}$$

这是结式矩阵或 Sylvester 矩阵. 为了保证 $\boldsymbol{\Omega}$ 非异，在单通道情形 $(m = n = 1)$，由 $(A(q^{-1}), B(q^{-1}))$ 互质可引出 $\boldsymbol{\Omega}$ 是非异方阵. 在多通道情形，为了保证 $\boldsymbol{\Omega}$ 非异，应假设 $(\boldsymbol{A}(q^{-1}), \boldsymbol{B}(q^{-1}))$ 左素. 于是由 (5.5.18) 和 (5.5.19) 有如下引理.

【引理 5.5.3】 $s(t)$ 有非递推表达式

$$\boldsymbol{s}(t) = \sum_{i=0}^{n_b-1}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{C}(q^{-1})\boldsymbol{w}(t-i) + \sum_{i=0}^{n_a-1}\boldsymbol{\Omega}_i^{(2)}[\boldsymbol{P}(q^{-1})\boldsymbol{y}(t-i) - \boldsymbol{R}(q^{-1})\boldsymbol{v}(t-i)] \tag{5.5.20}$$

其中 $\boldsymbol{\Omega}_i^{(1)},\boldsymbol{\Omega}_i^{(2)}$ 为 $m\times m$ 阵，它们由如下分块表示定义：

$$[\boldsymbol{\Omega}_0^{(1)}\cdots\boldsymbol{\Omega}_{n_b-1}^{(1)}\ \boldsymbol{\Omega}_0^{(2)}\cdots\boldsymbol{\Omega}_{n_a-1}^{(1)}]=[\boldsymbol{I}_n\quad \boldsymbol{0}\cdots\boldsymbol{0}]\boldsymbol{\Omega}^{-1}\tag{5.5.21}$$

且有非递推最优反卷积滤波器

$$\begin{aligned}\hat{\boldsymbol{s}}(t\mid t+N)=&\sum_{i=0}^{n_b-1}\boldsymbol{\Omega}_i^{(1)}\sum_{j=0}^{n_c}\boldsymbol{C}_j\hat{\boldsymbol{w}}(t-i-j\mid t+N)+\sum_{i=0}^{n_a-1}\boldsymbol{\Omega}_i^{(2)}\times\\&[\sum_{j=0}^{n_p}\boldsymbol{P}_j\hat{\boldsymbol{y}}(t-i-j\mid t+N)-\sum_{j=0}^{n_r}\boldsymbol{R}_j\hat{\boldsymbol{v}}(t-i-j\mid t+N)]\end{aligned}\tag{5.5.22}$$

证明　由 $\boldsymbol{\Omega}$ 为非异的假设，对(5.5.18)取矩阵逆运算有

$$\begin{bmatrix}\boldsymbol{s}(t)\\\boldsymbol{s}(t-1)\\\vdots\\\boldsymbol{s}(t-n_a-n_b+1)\end{bmatrix}=\boldsymbol{\Omega}^{-1}\begin{bmatrix}\boldsymbol{C}(q^{-1})\boldsymbol{w}(t)\\\vdots\\\vdots\\\boldsymbol{P}(q^{-1})\boldsymbol{y}(t-n_a+1)-\boldsymbol{R}(q^{-1})\boldsymbol{v}(t-n_a+1)\end{bmatrix}\tag{5.5.23}$$

由定义(5.5.21)和(5.5.23)引出(5.5.20). 当 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{P}(q^{-1})$ 稳定时，对(5.5.20)取射影运算立刻得(5.5.22). 当 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{P}(q^{-1})$ 至少有一个不稳定时，对 $\boldsymbol{y}(t)$ 不能直接取射影运算，可类似于引理 5.3.2 求 $\boldsymbol{P}(q^{-1})\boldsymbol{y}(t-i)$ 的预报器，从而得到(5.5.22). □

【引理 5.5.4】　下述公式成立

$$\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]=\boldsymbol{Q}_v\boldsymbol{G}_{i-t}^{\mathrm{T}}\tag{5.5.24}$$

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]=\boldsymbol{Q}_w\boldsymbol{F}_{i-t}^{\mathrm{T}}\tag{5.5.25}$$

【定理 5.5.1】　多通道 ARMA 信号(5.5.1)和(5.5.2)有渐近稳定的 Wiener 反卷积滤波器

$$\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N)\tag{5.5.26}$$

或等价地表为 ARMA 递推形式

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\mathrm{adj}\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N)\tag{5.5.27}$$

其中多项式矩阵 $\boldsymbol{K}_N(q^{-1})$ 定义为

$$\begin{aligned}\boldsymbol{K}_N(q^{-1})=&\sum_{i=0}^{n_b-1}\boldsymbol{\Omega}_i^{(1)}\sum_{j=0}^{n_c}\boldsymbol{C}_j\boldsymbol{L}_{N+i+j}^{w}(q^{-1})\widetilde{\boldsymbol{\Lambda}}(q^{-1})+\\&\sum_{i=0}^{n_a-1}\boldsymbol{\Omega}_i^{(2)}[\sum_{j=0}^{n_p}\boldsymbol{P}_j\boldsymbol{J}_{-i-j-N}(q^{-1})-\sum_{j=0}^{n_r}\boldsymbol{R}_j\boldsymbol{L}_{N+i+j}^{v}(q^{-1})\widetilde{\boldsymbol{\Lambda}}(q^{-1})]\end{aligned}\tag{5.5.28}$$

证明　将(5.5.9)，(5.5.10)和(5.5.15)代入(5.5.22)引出(5.5.26)～(5.5.28). 由 $\det\boldsymbol{D}(q^{-1})=\det\widetilde{\boldsymbol{D}}(q^{-1})$ 和 $\boldsymbol{D}(q^{-1})$ 的稳定性引出 $\widetilde{\boldsymbol{D}}(q^{-1})$ 是稳定的. 因而(5.5.26)或(5.5.27)是渐近稳定的. □

【定理 5.5.2】　当 $N\geqslant 0$ 时估值误差 $\tilde{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{s}(t)-\hat{\boldsymbol{s}}(t\mid t+N)$ 为

$$\begin{aligned}\tilde{\boldsymbol{s}}(t\mid t+N)=&\sum_{i=0}^{n_b-1}\boldsymbol{\Omega}_i^{(1)}\sum_{j=0}^{n_c}\boldsymbol{C}_j[\boldsymbol{w}(t-i-j)-\hat{\boldsymbol{w}}(t-i-j\mid t+N)]-\\&\sum_{i=0}^{n_a-1}\boldsymbol{\Omega}_i^{(2)}\sum_{j=0}^{n_r}R_j[\boldsymbol{v}(t-i-j)-\hat{\boldsymbol{v}}(t-i-j\mid t+N)],\quad N\geqslant 0\end{aligned}\tag{5.5.29}$$

它可等价地表为

$$\tilde{\boldsymbol{s}}(t \mid t+N) = \sum_{j=0}^{n_1} \boldsymbol{\Omega}_j^{(3)} \boldsymbol{w}(t-j) + \sum_{j=0}^{n_2} \boldsymbol{\Omega}_j^{(4)} \boldsymbol{v}(t-j) + \sum_{j=0}^{n_0+N} \boldsymbol{\Omega}^{(5)} \boldsymbol{\varepsilon}(t-n_0+j), \quad N \geqslant 0 \tag{5.5.30}$$

其中 $\boldsymbol{\Omega}_j^{(i)}, i = 3,4,5$,由合并(5.5.29)同类项得到,且定义

$$n_1 = n_b + n_c - 1, \quad n_2 = n_a + n_r - 1, \quad n_0 = \max(n_1, n_2) \tag{5.5.31}$$

也可将(5.5.30)统一地表为

$$\tilde{\boldsymbol{s}}(t \mid t+N) = \sum_{j=0}^{n_0+N} [\boldsymbol{\Omega}_j^{(3)} \boldsymbol{w}(t-j) + \boldsymbol{\Omega}_j^{(4)} \boldsymbol{v}(t-j) + \boldsymbol{\Omega}_j^{(5)} \boldsymbol{\varepsilon}(t-n_0+j)], \quad N \geqslant 0 \tag{5.5.32}$$

其中规定 $\boldsymbol{\Omega}_j^{(3)} = \mathbf{0}(j > n_1), \boldsymbol{\Omega}_j^{(4)} = \mathbf{0}(j > n_2)$.

当 $N < 0$ 时估值误差 $\tilde{\boldsymbol{s}}(t \mid t+N)$ 有表达式

$$\tilde{\boldsymbol{s}}(t \mid t+N) = \sum_{j=0}^{n_1} \boldsymbol{\Omega}_j^{(3)} \boldsymbol{w}(t-j) + \sum_{j=0}^{n_2} \boldsymbol{\Omega}_j^{(4)} \boldsymbol{v}(t-j) + \sum_{j=0}^{n_3} \boldsymbol{\Omega}_j^{(6)} \boldsymbol{\varepsilon}(t-j), \quad N < 0, \quad n_3 = \max(-N-1, n_0) \tag{5.5.33}$$

或表为统一形式

$$\tilde{\boldsymbol{s}}(t \mid t+N) = \sum_{j=0}^{n_3} [\boldsymbol{\Omega}_j^{(3)} \boldsymbol{w}(t-j) + \boldsymbol{\Omega}_j^{(4)} \boldsymbol{v}(t-j) + \boldsymbol{\Omega}_j^{(6)} \boldsymbol{\varepsilon}(t-j)], \quad N < 0 \tag{5.5.34}$$

其中规定 $\boldsymbol{\Omega}_j^{(3)} = 0(j > n_1), \boldsymbol{\Omega}_j^{(4)} = 0(j > n_2)$.

证明 当 $N \geqslant 0$ 时,注意 $\hat{\boldsymbol{y}}(t-i-j \mid t+N) = \boldsymbol{y}(t-i-j), i,j = 0,1,\cdots$,于是由(5.5.20)和(5.5.22)引出(5.5.29).应用新息滤波器(5.5.11),合并(5.5.29)同类项系数阵容易得到(5.5.30),进而得(5.5.32).

当 $N < 0$ 时,由(5.5.20)和(5.5.22)有项

$$\sum_{i=0}^{n_a-1} \boldsymbol{\Omega}_i^{(2)} \sum_{j=0}^{n_p} \boldsymbol{P}_j \tilde{\boldsymbol{y}}(t-i-j \mid t+N) \tag{5.5.35}$$

是 $\boldsymbol{\varepsilon}(t+N+1), \cdots, \boldsymbol{\varepsilon}(t)$ 的线性组合.而项

$$-\sum_{j=0}^{n_1} \boldsymbol{\Omega}_i^{(3)} \hat{\boldsymbol{w}}(t-j \mid t+N) - \sum_{j=0}^{n_2} \boldsymbol{\Omega}_i^{(4)} \hat{\boldsymbol{v}}(t-j \mid t+N) \tag{5.5.36}$$

当 $n_0 \geqslant -N$ 时才不为零,且是 $\boldsymbol{\varepsilon}(t-n_0), \cdots, \boldsymbol{\varepsilon}(t+N)$ 的线性组合.于是由(5.5.20)和(5.5.22)及上两式得(5.5.34).证毕. □

【定理 5.5.3】 估值误差方差阵 $\boldsymbol{P}(N) = \mathrm{E}[\tilde{\boldsymbol{s}}(t \mid t+N)\tilde{\boldsymbol{s}}^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}(N) = \sum_{i=0}^{n_0+N} \sum_{j=0}^{n_0+N} [\boldsymbol{\Omega}_i^{(3)}, \boldsymbol{\Omega}_i^{(4)}, \boldsymbol{\Omega}_i^{(5)}] \begin{bmatrix} \boldsymbol{Q}_w \delta_{ij} & \mathbf{0} & \boldsymbol{Q}_w \boldsymbol{F}_{i+j-n_0}^{\mathrm{T}} \\ \mathbf{0} & \boldsymbol{Q}_v \delta_{ij} & \boldsymbol{Q}_v \boldsymbol{G}_{i+j-n_0}^{\mathrm{T}} \\ \boldsymbol{F}_{i+j-n_0} \boldsymbol{Q}_w & \boldsymbol{G}_{i+j-n_0} \boldsymbol{Q}_v & \boldsymbol{Q}_\varepsilon \delta_{ij} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega}_j^{(3)\mathrm{T}} \\ \boldsymbol{\Omega}_j^{(4)\mathrm{T}} \\ \boldsymbol{\Omega}_j^{(5)\mathrm{T}} \end{bmatrix}, \quad N \geqslant 0 \tag{5.5.37}$$

$$\boldsymbol{P}(N)=\sum_{i=0}^{n_3}\sum_{j=0}^{n_3}[\boldsymbol{\Omega}_i^{(3)},\boldsymbol{\Omega}_i^{(4)},\boldsymbol{\Omega}_i^{(6)}]\begin{bmatrix}\boldsymbol{Q}_w\delta_{ij} & \boldsymbol{0} & \boldsymbol{Q}_w\boldsymbol{F}_{i-j}^{\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{Q}_v\delta_{ij} & \boldsymbol{Q}_v\boldsymbol{G}_{i-j}^{\mathrm{T}}\\ \boldsymbol{F}_{j-i}\boldsymbol{Q}_w & \boldsymbol{G}_{j-i}\boldsymbol{Q}_v & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}_j^{(3)}\\ \boldsymbol{\Omega}_j^{(4)}\\ \boldsymbol{\Omega}_j^{(5)}\end{bmatrix},\quad N<0 \tag{5.5.38}$$

证明 当 $N\geqslant 0$ 时(5.5.32)可写为

$$\tilde{\boldsymbol{s}}(t\mid t+N)=\sum_{j=0}^{n_0+N}[\boldsymbol{\Omega}_j^{(3)},\boldsymbol{\Omega}_j^{(4)},\boldsymbol{\Omega}_j^{(5)}]\begin{bmatrix}\boldsymbol{w}(t-j)\\ \boldsymbol{v}(t-j)\\ \boldsymbol{\varepsilon}(t-n_0+j)\end{bmatrix},\quad N\geqslant 0 \tag{5.5.39}$$

而当 $N<0$ 时(5.5.35)可写为

$$\tilde{\boldsymbol{s}}(t\mid t+N)=\sum_{j=0}^{n_3}[\boldsymbol{\Omega}_j^{(3)},\boldsymbol{\Omega}_j^{(4)},\boldsymbol{\Omega}_j^{(6)}]\begin{bmatrix}\boldsymbol{w}(t-j)\\ \boldsymbol{v}(t-j)\\ \boldsymbol{\varepsilon}(t-j)\end{bmatrix},\quad N<0 \tag{5.5.40}$$

利用上两式和(5.5.24)和(5.5.25)得(5.5.37)和(5.5.38). □

考虑多通道 ARMA 信号 Wiener 滤波问题:

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.5.41}$$

$$\boldsymbol{y}(t)=\boldsymbol{s}(t)+\boldsymbol{\eta}(t) \tag{5.5.42}$$

$$\boldsymbol{P}(q^{-1})\boldsymbol{\eta}(t)=\boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \tag{5.5.43}$$

其中(5.5.41)同(5.5.1),$\boldsymbol{y}(t)\in R^m$ 为观测信号,$\boldsymbol{\eta}(t)$ 为服从 ARMA 模型(5.5.43)的有色观测噪声. $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 的假设同系统(5.5.1)和(5.5.2). 问题是基于理论($\boldsymbol{y}(t+N)$, $\boldsymbol{y}(t+N-1),\cdots$)求信号 $\boldsymbol{s}(t)$ 的 Wiener 滤波器 $\hat{\boldsymbol{s}}(t\mid t+N)$.

问题可化为特殊的反卷积问题. 事实上由(5.5.42) ~ (5.5.43)可得

$$\boldsymbol{P}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{P}(q^{-1})\boldsymbol{s}(t)+\boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \tag{5.5.44}$$

它是(5.5.2)取 $\boldsymbol{B}(q^{-1})=\boldsymbol{P}(q^{-1})$ 的特殊情形. 于是有如下定理.

【定理 5.5.4】 带 ARMA 有色观测噪声的多通道 ARMA 信号(5.5.41) ~ (5.5.43)有渐近稳定的 Wiener 滤波器(5.5.26) ~ (5.5.28),其中推导中应置 $\boldsymbol{B}(q^{-1})=\boldsymbol{P}(q^{-1})$. 而估值误差方差阵由(5.5.37)和(5.5.38)计算.

【注 5.5.1】 ARMA 新息模型(5.5.7)也可用引入如下伪交换得到:

$$\boldsymbol{B}(q^{-1})\boldsymbol{A}^{-1}(q^{-1})=\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) \tag{5.5.45}$$

带 $\widetilde{\boldsymbol{A}}_0=\boldsymbol{I}_m$, $\det\widetilde{\boldsymbol{A}}(q^{-1})=\det\widetilde{\boldsymbol{A}}(q^{-1})$. 设 $(\widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{P}(q^{-1}),\widetilde{\boldsymbol{B}}(q^{-1})\boldsymbol{C}(q^{-1}),\widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{R}(q^{-1}))$ 左素,且 $(\widetilde{\boldsymbol{B}}(q^{-1})\boldsymbol{C}(q^{-1}),\widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{R}(q^{-1}))$ 无行列式零点在单位圆上的左因式. 将(5.5.1)代入(5.5.2)有

$$\boldsymbol{P}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{B}(q^{-1})\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \tag{5.5.46}$$

将(5.5.45)代入上式有

$$\widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{P}(q^{-1})\boldsymbol{y}(t)=\widetilde{\boldsymbol{B}}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{w}(t)+\widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \tag{5.5.47}$$

这引出 ARMA 新息模型

$$\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.5.48}$$

$$\boldsymbol{\Lambda}(q^{-1})=\widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{P}(q^{-1}) \tag{5.5.49}$$

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\overline{\boldsymbol{B}}(q^{-1})\boldsymbol{w}(t)+\overline{\boldsymbol{R}}(q^{-1})\boldsymbol{v}(t) \tag{5.5.50}$$

$$\bar{\boldsymbol{B}}(q^{-1}) = \tilde{\boldsymbol{B}}(q^{-1})\boldsymbol{C}(q^{-1}), \quad \bar{\boldsymbol{R}}(q^{-1}) = \tilde{\boldsymbol{A}}(q^{-1})\boldsymbol{R}(q^{-1}) \tag{5.5.51}$$

【注 5.5.2】 对于 Wiener 滤波问题(5.5.41) ~ (5.5.43),引入伪交换

$$\boldsymbol{P}(q^{-1})\boldsymbol{A}^{-1}(q^{-1}) = \tilde{\boldsymbol{A}}^{-1}(q^{-1})\tilde{\boldsymbol{P}}(q^{-1}) \tag{5.5.52}$$

带 $\tilde{\boldsymbol{P}}_0 = \boldsymbol{I}_m$, $\det\boldsymbol{A}(q^{-1}) = \det\tilde{\boldsymbol{A}}(q^{-1})$,也可导出 ARMA 新息模型.事实上,将(5.5.41)和(5.5.43)代入(5.5.42)有

$$\boldsymbol{P}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{P}(q^{-1})\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{w}(t) + \boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \tag{5.5.53}$$

将(5.5.52)代入(5.5.53)有 ARMA 新息模型

$$\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.5.54}$$

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \bar{\boldsymbol{B}}(q^{-1})\boldsymbol{w}(t) + \bar{\boldsymbol{R}}(q^{-1})\boldsymbol{v}(t) \tag{5.5.55}$$

$$\boldsymbol{\Lambda}(q^{-1}) = \tilde{\boldsymbol{A}}(q^{-1})\boldsymbol{P}(q^{-1}) \tag{5.5.56}$$

$$\bar{\boldsymbol{B}}(q^{-1}) = \tilde{\boldsymbol{P}}(q^{-1})\boldsymbol{C}(q^{-1}), \quad \bar{\boldsymbol{R}}(q^{-1}) = \tilde{\boldsymbol{A}}(q^{-1})\boldsymbol{R}(q^{-1}) \tag{5.5.57}$$

其中假设$(\boldsymbol{\Lambda}(q^{-1}), \bar{\boldsymbol{B}}(q^{-1}), \bar{\boldsymbol{R}}(q^{-1}))$左素,且$(\bar{\boldsymbol{B}}(q^{-1}), \bar{\boldsymbol{R}}(q^{-1}))$无行列式的零点位于单位圆上的左因式.

【注 5.5.3】 用本节方法可进一步解决如下一般反卷积问题[1,88]:

$$\boldsymbol{s}(t) = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.5.58}$$

$$\boldsymbol{y}(t) = \boldsymbol{\Phi}^{-1}(q^{-1})\boldsymbol{\Psi}(q^{-1})\boldsymbol{s}(t) + \boldsymbol{L}^{-1}(q^{-1})\boldsymbol{M}(q^{-1})\boldsymbol{v}(t) + \boldsymbol{\xi}(t) \tag{5.5.59}$$

其中 $\boldsymbol{y}(t) \in R^m$, $\boldsymbol{s}(t) \in R^n$, $\boldsymbol{\Phi}(q^{-1}), \cdots, \boldsymbol{M}(q^{-1})$ 为 q^{-1} 的多项式矩阵,且设首系阵 $\boldsymbol{\Phi}_0 = \boldsymbol{I}_m$, $\boldsymbol{L}_0 = \boldsymbol{I}_m$, $\boldsymbol{A}_0 = \boldsymbol{I}_n$.假设 $\boldsymbol{w}(t) \in R^r$, $\boldsymbol{v}(t) \in R^m$, $\boldsymbol{\xi}(t) \in R^m$ 是带零均值白噪声,$\boldsymbol{\xi}(t)$ 独立于 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$,但 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 是相关白噪声:

$$\mathrm{E}\left\{\begin{bmatrix} \boldsymbol{w}(t) \\ \boldsymbol{v}(t) \end{bmatrix}\begin{bmatrix} \boldsymbol{w}^{\mathrm{T}}(j) & \boldsymbol{v}^{\mathrm{T}}(j) \end{bmatrix}\right\} = \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v \end{bmatrix}\delta_{tj}, \quad \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v \end{bmatrix} > 0 \tag{5.5.60}$$

引入左素分解

$$\boldsymbol{P}^{-1}(q^{-1})[\boldsymbol{B}(q^{-1}), \boldsymbol{R}(q^{-1})] = [\boldsymbol{\Phi}^{-1}(q^{-1})\boldsymbol{\Psi}(q^{-1}), \boldsymbol{L}^{-1}(q^{-1})\boldsymbol{M}(q^{-1})] \tag{5.5.61}$$

则化为等价的反卷积问题:

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t)$$

$$\boldsymbol{P}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{B}(q^{-1})\boldsymbol{s}(t) + \boldsymbol{R}(q^{-1})\boldsymbol{v}(t) + \boldsymbol{P}(q^{-1})\boldsymbol{\xi}(t) \tag{5.5.62}$$

可进一步用本节方法解决反卷积问题.

可用如下雷达跟踪系统[1]来说明反卷积问题(5.5.58) ~ (5.5.60)的应用背景:

$$x(t+1) = x(t) + s(t)T_0 + w(t)T_0^2/2 \tag{5.5.63}$$

$$s(t+1) = s(t) + w(t)T_0 \tag{5.5.64}$$

$$y(t) = x(t) + \xi(t) \tag{5.5.65}$$

其中 T_0 为采样周期,$x(t)$, $s(t)$, $w(t)$ 各为在时刻 tT_0 运动目标的位置、速度和加速度,设 $w(t)$ 和 $\xi(t)$ 是零均值、方差各为 σ_w^2 和σ_ξ^2 的独立白噪声.$y(t)$ 是对位置的观测信号,$\xi(t)$ 是观测噪声.问题是基于位置观测$(y(t+N), y(t+N-1), \cdots)$求速度 $s(t)$ 的 Wiener 滤波器 $\hat{s}(t \mid t+N)$.该问题等价于如下单通道反卷积问题:

$$y(t) = \frac{T_0 q^{-1}}{1-q^{-1}}s(t) + \frac{T_0^2 q^{-1}}{2(1-q^{-1})}v(t) + \xi(t) \tag{5.5.66}$$

$$(1-q^{-1})s(t) = T_0 w(t-1) \tag{5.5.67}$$

$$v(t)=w(t) \tag{5.5.68}$$

显然 $v(t)$ 与 $w(t)$ 是相关白噪声,即

$$\mathrm{E}[w(t)v(t)]=\sigma_w^2=s \tag{5.5.69}$$

即(5.5.60)成立.

对于单通道系统(5.5.1)和(5.5.2),设 $(A(q^{-1})P(q^{-1}),B(q^{-1})C(q^{-1}),A(q^{-1})R(q^{-1}))$ 互质,且 $(B(q^{-1})C(q^{-1}),A(q^{-1})R(q^{-1}))$ 无零点在单位圆上的左因式,由注 5.5.1,可得 ARMA 新息模型

$$\Lambda(q^{-1})y(t)=D(q^{-1})\varepsilon(t) \tag{5.5.70}$$

$$\begin{aligned}&\Lambda(q^{-1})=A(q^{-1})P(q^{-1}),\quad \bar{B}(q^{-1})=B(q^{-1})C(q^{-1}),\\&\bar{R}(q^{-1})=A(q^{-1})R(q^{-1})\end{aligned} \tag{5.5.71}$$

【定理 5.5.5】 对于单通道反卷积问题(5.5.1)和(5.5.2)有渐近稳定的 Wiener 反卷积滤波器

$$D(q^{-1})\hat{s}(t\mid t+N)=K_N(q^{-1})y(t+N) \tag{5.5.72}$$

其中多项式 $K_N(q^{-1})$ 由(5.5.28)定义,但应置 $\widetilde{\boldsymbol{A}}(q^{-1})=\Lambda(q^{-1})=A(q^{-1})P(q^{-1})$. 估计误差方差由定理 5.5.3 计算.

证明 注意对单通道有 $\widetilde{\boldsymbol{A}}(q^{-1})=A(q^{-1})$, $\widetilde{\boldsymbol{D}}(q^{-1})=D(q^{-1})$,由定理 5.5.1 得证.

【注 5.5.4】 定理 5.5.5 的方法和结果可推广到处理如下单通道反卷积问题:

$$y(t)=\frac{\Psi(q^{-1})}{\Phi(q^{-1})}s(t)+\frac{M(q^{-1})}{L(q^{-1})}v(t)+\xi(t) \tag{5.5.73}$$

$$A(q^{-1})s(t)=C(q^{-1})w(t) \tag{5.5.74}$$

$$\mathrm{E}\left\{\begin{bmatrix}w(t)\\v(t)\end{bmatrix}\begin{bmatrix}w(j) & v(j)\end{bmatrix}\right\}=\begin{bmatrix}\sigma_w^2 & s\\ s & \sigma_v^2\end{bmatrix}\delta_{tj} \tag{5.5.75}$$

(5.5.66)~(5.5.69)就是这类单通道系统. 引入多项式互质分解

$$\left[\frac{\Psi(q^{-1})}{\Phi(q^{-1})},\frac{M(q^{-1})}{L(q^{-1})}\right]=\frac{1}{P(q^{-1})}[B(q^{-1}),R(q^{-1})] \tag{5.5.76}$$

则原系统化为等价的反卷积系统

$$P(q^{-1})y(t)=B(q^{-1})s(t)+R(q^{-1})v(t)+P(q^{-1})\xi(t) \tag{5.5.77}$$

$$A(q^{-1})s(t)=C(q^{-1})w(t) \tag{5.5.78}$$

【例 5.5.1】 考虑跟踪系统

$$s(t+1)=s(t)+T_0\dot{s}(t) \tag{5.5.79}$$

$$\dot{s}(t)=a\dot{s}(t-1)+w(t) \tag{5.5.80}$$

$$y(t)=s(t)+\eta(t) \tag{5.5.81}$$

$$(1+pq^{-1})\eta(t)=(1+rq^{-1})v(t) \tag{5.5.82}$$

其中 T_0 为采样周期,$s(t)$ 和 $\dot{s}(t)$ 各为在采样时刻 tT_0 处运动目标的位置和速度,$y(t)$ 是对位置 $s(t)$ 的观测信号,$\eta(t)$ 是 ARMA 有色观测噪声,$w(t)$ 和 $v(t)$ 是零均值、方差各为 σ_w^2 和 σ_v^2 的相互独立高斯白噪声. 取 $N=1$ 和 $N-1$,问题是求位置信号 $s(t)$ 的 Wiener 跟踪滤波器 $\hat{s}(t\mid t)$ 和预报器 $\hat{s}(t\mid t-1)$.

将(5.5.80)代入(5.5.79)消去 $\dot{s}(t)$ 可得 $s(t)$ 的 ARMA 模型

$$(1 - aq^{-1})(1 - q^{-1})s(t) = T_0 w(t-1) \tag{5.5.83}$$

系统(5.5.83),(5.5.81)和(5.5.82)具有(5.5.41)~(5.5.43)的形式.仿真中取 $T_0 = 0.1$,$a = 0.98$,$p = 0.7$,$r = 0.8$,$\sigma_w^2 = 0.1$,$\sigma_v^2 = 0.6$,可得 ARMA 新息模型为

$$\begin{aligned}&(1 - 0.98q^{-1})(1 - q^{-1})(1 + 0.7q^{-1})y(t) = \\ &(1 - 0.921\,72q^{-1} - 0.621\,78q^{-2} + 0.604\,44q^{-3})\varepsilon(t)\end{aligned} \tag{5.5.84}$$

其中新息方差 $\sigma_\varepsilon^2 = 0.778\,24$.于是由定理5.5.3和定理5.5.4可求得 Wiener 跟踪滤波器 $\hat{s}(t \mid t)$ 和预报器 $\hat{s}(t \mid t-1)$,且可求得估值误差方差各为

$$P(0) = 0.152\,746, \quad P(-1) = 0.198\,12 \tag{5.5.85}$$

可见 $\hat{s}(t \mid t)$ 的精度高于 $\hat{s}(t \mid t-1)$ 的精度.仿真结果如图5.5.1和图5.5.2所示,其中实线代表真实值 $s(t)$,虚线代表估值 $\hat{s}(t \mid t)$ 或 $\hat{s}(t \mid t-1)$.

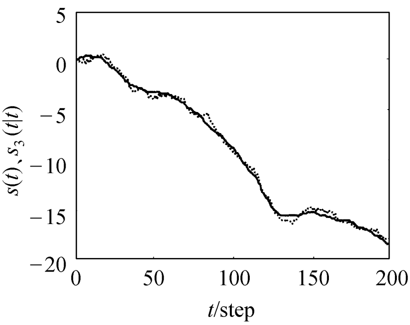


图5.5.1　信号 $s(t)$ 和 Wiener 跟踪滤波器 $\hat{s}(t \mid t)$

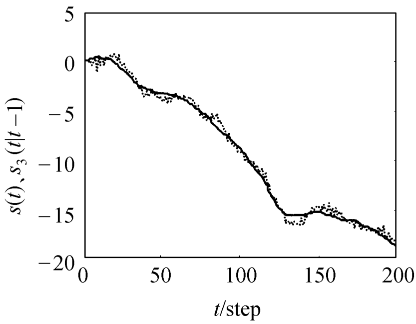


图5.5.2　信号 $s(t)$ 和 Wiener 跟踪预报器 $\hat{s}(t \mid t-1)$

【注5.5.1】　例5.5.1也可用状态空间方法基于经典稳态Kalman滤波处理.(5.5.83)有状态空间模型

$$\boldsymbol{\alpha}(t+1) = \begin{bmatrix} 1+a & 1 \\ -a & 0 \end{bmatrix}\boldsymbol{\alpha}(t) + \begin{bmatrix} T_0 \\ 0 \end{bmatrix} w(t) \tag{5.5.86}$$

$$s(t) = [1 \quad 0]\boldsymbol{\alpha}(t) \tag{5.5.87}$$

(5.5.82)有状态空间模型

$$\beta(t+1) = -p\beta(t) + (r-p)v(t) \tag{5.5.88}$$

$$\eta(t) = \beta(t) + v(t) \tag{5.5.89}$$

于是有增广系统

$$\boldsymbol{x}(t+1) = \begin{bmatrix} 1+a & 1 & 0 \\ -a & 0 & 0 \\ 0 & 0 & -p \end{bmatrix}\boldsymbol{x}(t) + \begin{bmatrix} T_0 & 0 \\ 0 & 0 \\ 0 & r-p \end{bmatrix}\begin{bmatrix} w(t) \\ v(t) \end{bmatrix} \tag{5.5.90}$$

$$y(t) = [1 \quad 0 \quad 1]\boldsymbol{x}(t) + v(t) \tag{5.5.91}$$

其中定义 $\boldsymbol{x}(t) = [\boldsymbol{\alpha}^{\mathrm{T}}(t), \beta^{\mathrm{T}}(t)]^{\mathrm{T}}$,且有关系

$$s(t) = [1 \quad 0 \quad 0]\boldsymbol{x}(t) \tag{5.5.92}$$

$$\hat{s}(t \mid t+N) = [1 \quad 0 \quad 0]\hat{\boldsymbol{x}}(t \mid t+N), \quad N = 0, -1 \tag{5.5.93}$$

记估值 $\hat{s}(t \mid t+N)$ 和 $\hat{\boldsymbol{x}}(t \mid t+N)$ 的误差方差阵各为 $P_s(N)$ 和 $\boldsymbol{P}_x(N)$,则有关系

$$P_s(N) = [1 \quad 0 \quad 0]\boldsymbol{P}_x(N)[1 \quad 0 \quad 0]^{\mathrm{T}} \tag{5.5.94}$$

用稳态 Kalman 滤波方法可求得

$$P_s(N)=0.152\,746,\quad P_s(-1)=0.198\,12 \tag{5.5.95}$$

它相同于应用定理 5.5.3 的结果(5.5.85).

5.6 统一的 Wiener 状态滤波器

经典稳态 Kalman 滤波要求解 Riccati 方程,且稳态 Kalman 滤波、平滑和预报尚无统一算法.本节介绍用现代时间序列分析方法[1],基于 ARMA 新息模型和白噪声估值器及观测预报器提出的统一的 Wiener 状态估值器[1,171,172],可统一处理状态滤波、平滑和预报问题.避免了 Riccati 方程.方法原理是利用系统可观性,可将系统状态表为白噪声和观测的线性组合,因而可将状态估计问题转化为白噪声估计和观测预报问题.有关结果可见文献[18,20,27,34,36,37,58,61,171,172].特别本节还进一步提出了状态估计误差方差阵的计算公式,可用于处理信息融合状态估计问题.[197~205]

考虑线性离散定常随机系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t) \tag{5.6.1}$$

$$\boldsymbol{y}(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t) \tag{5.6.2}$$

其中状态 $\boldsymbol{x}(t)\in R^n$,观测 $\boldsymbol{y}(t)\in R^m$,$\boldsymbol{\Phi},\boldsymbol{\Gamma},\boldsymbol{H}$ 为适当维数常阵.

【假设 1】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}(t)\in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$、相关阵为 $\boldsymbol{S}$ 的白噪声:

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix}[\boldsymbol{w}^{\mathrm{T}}(j),\boldsymbol{v}^{\mathrm{T}}(j)]\right\}=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}\delta_{tj},\quad \begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}>0 \tag{5.6.3}$$

【假设 2】 系统是完全可观的,即

$$\mathrm{rank}\boldsymbol{\Omega}=n,\quad \boldsymbol{\Omega}=[\boldsymbol{H}^{\mathrm{T}},(\boldsymbol{H\Phi})^{\mathrm{T}},\cdots,(\boldsymbol{H\Phi}^{\beta-1})^{\mathrm{T}}]^{\mathrm{T}} \tag{5.6.4}$$

其中 β 为可观性指数.还设系统是完全能稳的[173].

【假设 3】 初始观测时刻 $t_0=-\infty$.

Wiener 状态估计问题是基于观测 $(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots)$ 求状态 $\boldsymbol{x}(t)$ 的线性最小方差(最优)Wiener 状态估值器 $\hat{\boldsymbol{x}}(t\mid t+N)$,它具有以 $\boldsymbol{y}(t+N)$ 作为输入的传递函数形式.对 $N=0,N>0$ 或 $N<0$,各称其为 Wiener 状态滤波器、平滑器或预报器.

由(5.6.1)和(5.6.2)有

$$\boldsymbol{y}(t)=\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}\boldsymbol{w}(t)+\boldsymbol{v}(t) \tag{5.6.5}$$

其中 q^{-1} 为单位滞后算子.引入左素分解

$$\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1}) \tag{5.6.6}$$

带 $\boldsymbol{A}_0=\boldsymbol{I}_m,\boldsymbol{B}_0=\boldsymbol{0}$.将上式代入(5.6.4)引出

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{B}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.6.7}$$

由$(\boldsymbol{A}(q^{-1}),\boldsymbol{B}(q^{-1}))$左素和假设 1 引出[1]存在 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.6.8}$$

其中新息 $\boldsymbol{\varepsilon}(t)\in R^m$ 的零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声,$\boldsymbol{D}(q^{-1})$ 是稳定的,$\boldsymbol{D}_0=\boldsymbol{I}_m$,且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \boldsymbol{B}(q^{-1})\boldsymbol{w}(t) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.6.9}$$

$\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers－Wouters 算法求得.

从(5.6.7)～(5.6.9)出发可由节5.1.1和节5.1.2的公式求白噪声估值器 $\hat{\boldsymbol{w}}(t \mid t+N)$ 和 $\hat{\boldsymbol{v}}(t \mid t+N)$，且由(5.6.8)可求观测预报器 $\hat{\boldsymbol{y}}(t \mid t+N)(N \leqslant 0)$(5.3.23).

【引理5.6.1】 系统(5.6.1)和(5.6.2)在假设1～3下，状态 $\boldsymbol{x}(t)$ 有非递推表达式

$$\boldsymbol{x}(t) = \sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i\left[\boldsymbol{y}(t+i) - \sum_{j=0}^{i-1}\boldsymbol{H\Phi}^{i-1-j}\boldsymbol{\Gamma w}(t+j) - \boldsymbol{v}(t+i)\right] \tag{5.6.10}$$

其中规定 $\boldsymbol{\Phi}^i = \boldsymbol{0}(i<0)$ 且 $j \geqslant 0$，$n \times m$ 阵 $\boldsymbol{\Omega}_i$ 定义为

$$\boldsymbol{\Omega}^+ = (\boldsymbol{\Omega}^{\mathrm{T}}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathrm{T}} = [\boldsymbol{\Omega}_0, \boldsymbol{\Omega}_1, \cdots, \boldsymbol{\Omega}_{\beta-1}] \tag{5.6.11}$$

其中 $\boldsymbol{\Omega}$ 由(5.6.4)定义. 非递推最优状态估值器为

$$\begin{aligned}\hat{\boldsymbol{x}}(t \mid t+N) = &\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i\Big[\hat{\boldsymbol{y}}(t+i \mid t+N) - \sum_{j=0}^{i-1}\boldsymbol{H\Phi}^{i-1-j}\boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t+j \mid t+N) - \\ &\hat{\boldsymbol{v}}(t+i \mid t+N)\Big]\end{aligned} \tag{5.6.12}$$

证明 重复应用(5.6.1)和(5.6.2)引出关系

$$\begin{bmatrix}\boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{\beta-1}\end{bmatrix}\boldsymbol{x}(t) = \begin{bmatrix}\boldsymbol{y}(t) - \boldsymbol{v}(t) \\ \boldsymbol{y}(t+1) - \boldsymbol{H\Gamma w}(t) - \boldsymbol{v}(t+1) \\ \vdots \\ \boldsymbol{y}(t+\beta-1) - \sum_{i=0}^{\beta-2}\boldsymbol{H\Phi}^{\beta-2-i}\boldsymbol{\Gamma w}(t+i) - \boldsymbol{v}(t+\beta-1)\end{bmatrix} \tag{5.6.13}$$

应用假设2及分块表示(5.6.11)引出(5.6.10)，其中规定当 $i=0$ 时，(5.6.10)右边中括号内第二项为零，即规定 $\boldsymbol{\Phi}^i = \boldsymbol{0}(i<0)$ 且 $j \geqslant 0$.

对于稳定系统($\boldsymbol{\Phi}$ 稳定)，取(5.6.10)两边各项到无穷维 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t+N), \boldsymbol{y}^{\mathrm{T}}(t+N-1), \cdots)$ 上的射影得(5.6.12). 对不稳定系统($\boldsymbol{\Phi}$ 为不稳定矩阵)，先取 t_0 为有限初始时刻，取(5.6.10)两边各项到有限维 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t+N), \boldsymbol{y}^{\mathrm{T}}(t+N-1), \cdots, \boldsymbol{y}^{\mathrm{T}}(t_0))$ 上的射影可得非稳态关系(5.6.12)，然后令 $t_0 \to -\infty$ 引出稳态估值关系(5.6.12)，其中应用稳态 Kalman 滤波理论[173]引出(5.6.12)中各项稳态估值均存在. □

【引理5.6.2】 系统(5.6.1)和(5.6.2)在假设1～3下，有白噪声新息滤波器和 Wiener 滤波器各为

$$\hat{\boldsymbol{\theta}}(t \mid t+N) = \boldsymbol{L}_N^{\theta}(q^{-1})\boldsymbol{\varepsilon}(t+N), \quad \boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v} \tag{5.6.14}$$

$$\hat{\boldsymbol{\theta}}(t \mid t+N) = \boldsymbol{L}_N^{\theta}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N), \quad \boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v} \tag{5.6.15}$$

有观测预报器

$$\hat{\boldsymbol{y}}(t \mid t+N) = \boldsymbol{J}_{-N}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.6.16}$$

预报误差为

$$\tilde{\boldsymbol{y}}(t \mid t+N) = \boldsymbol{E}_{-N}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.6.17}$$

其中 $\boldsymbol{L}_N^{\theta}(q^{-1})$，$\widetilde{\boldsymbol{A}}(q^{-1})$，$\widetilde{\boldsymbol{D}}(q^{-1})$，$\boldsymbol{J}_{-N}(q^{-1})$，$\boldsymbol{E}_{-N}(q^{-1})$ 用节5.1～节5.3有关公式计算.

【引理5.6.3】 下述公式成立

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)] = \boldsymbol{\Lambda}_{i-t}^{w}, \quad \boldsymbol{\Lambda}_i^{w} = \boldsymbol{Q}_w\boldsymbol{F}_i^{\mathrm{T}} + \boldsymbol{S}\boldsymbol{G}_i^{\mathrm{T}}$$

$$\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)]=\boldsymbol{\Lambda}_{i-t}^{v},\quad \boldsymbol{\Lambda}_{i}^{v}=\boldsymbol{Q}_{v}\boldsymbol{G}_{i}^{\mathrm{T}}+\boldsymbol{S}^{\mathrm{T}}\boldsymbol{F}_{i}^{\mathrm{T}} \tag{5.6.18}$$

其中 $\boldsymbol{F}_i$ 和 $\boldsymbol{G}_i$ 由(5.1.36)置 $\boldsymbol{C}_j=\boldsymbol{A}_j$ 计算.

【定理5.6.1】 系统(5.6.1)和(5.6.2)在假设1 ~ 3下,有统一的渐近稳定的Wiener状态估值器

$$\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.6.19}$$

其中定义多项式矩阵 $K_N(q^{-1})$ 为

$$\boldsymbol{K}_N(q^{-1})=\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i\Big[\boldsymbol{J}_{i-N}(q^{-1})-\sum_{j=0}^{i-1}\boldsymbol{H}\boldsymbol{\Phi}^{i-1-j}\boldsymbol{\Gamma}\boldsymbol{L}_{N-j}^{w}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})-\boldsymbol{L}_{N-i}^{v}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})\Big] \tag{5.6.20}$$

其中规定 $\boldsymbol{\Phi}^i=\boldsymbol{0}(i<0)$ 且 $j\geqslant 0$. 它有ARMA递推形式

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\mathrm{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t+N) \tag{5.6.21}$$

证明 将(5.6.14) ~ (5.6.16)代入(5.6.12)得(5.6.19)和(5.6.20). 将 $\widetilde{\boldsymbol{D}}^{-1}(q^{-1})=\mathrm{adj}\widetilde{\boldsymbol{D}}(q^{-1})/\det\widetilde{\boldsymbol{D}}(q^{-1})$ 代入(5.6.19)得(5.6.21). 由 $\det\widetilde{\boldsymbol{D}}(q^{-1})=\det\boldsymbol{D}(q^{-1})$,由 $\boldsymbol{D}(q^{-1})$ 的稳定性引出 $\widetilde{\boldsymbol{D}}(q^{-1})$ 是稳定的,因而(5.6.19)或(5.6.21)是渐近稳定的. 证毕. □

在某些应用中人们仅对状态的某个或某些分量的估值器感兴趣,为了减小计算负担,只计算感兴趣状态分量估值器即可. 这引出如下解耦Wiener状态估值器. 引入状态分量表示

$$\boldsymbol{x}(t)=[x_1(t),\cdots,x_n(t)]^{\mathrm{T}} \tag{5.6.22}$$

并将 $n\times m$ 多项式矩阵 $K_N(q^{-1})$ 分块表示为

$$\boldsymbol{K}_N(q^{-1})=\begin{bmatrix}\boldsymbol{K}_{N1}(q^{-1})\\ \vdots\\ \boldsymbol{K}_{Nm}(q^{-1})\end{bmatrix} \tag{5.6.23}$$

其中 $\boldsymbol{K}_{Ni}(q^{-1})$ 为 $1\times m$ 多项式矩阵,由定理5.6.1立刻得如下分量解耦Wiener状态估值器.

【定理5.6.2】 系统(5.6.1)和(5.6.2)在假设1 ~ 3下,有渐近稳定的分量解耦Wiener状态估值器

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{x}_i(t\mid t+N)=\boldsymbol{K}_{Ni}(q^{-1})\mathrm{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t+N),\quad i=1,\cdots,n \tag{5.6.24}$$

其中 $\hat{x}_i(t\mid t+N)$ 是分量 $x_i(t)$ 的Wiener状态估值器.

【注5.6.1】 所谓分量解耦状态估值器是指计算分量估值器 $\hat{x}_i(t\mid t+N)$ 时,不需要计算 $\hat{x}_j(t-k\mid t+N-k)(j\neq i)$, $k=1,3,\cdots$.

下面来求估值误差表达式及估值误差方差阵.

由(5.6.10)减(5.6.12),用合并同类项 $\boldsymbol{w}(t+j)$ 和 $\hat{\boldsymbol{w}}(t+j\mid t+N)$ 的系数阵方法,可得估值误差 $\tilde{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t\mid t+N)$ 的表达式

$$\tilde{\boldsymbol{x}}(t\mid t+N)=\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i\tilde{\boldsymbol{y}}(t+i\mid t+N)+\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{w}\boldsymbol{w}(t+i)-\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i\boldsymbol{v}(t+i)-\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{w}\hat{\boldsymbol{w}}(t+i\mid t+N)+\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i\hat{\boldsymbol{v}}(t+i\mid t+N) \tag{5.6.25}$$

其中 $\boldsymbol{\Omega}_i$ 由(5.6.11)定义，$\boldsymbol{\Omega}_i^w$ 由合并(5.6.10)和(5.6.12)中的同类项 $\boldsymbol{w}(t+i)$ 的系数阵得到. 因为由引理 5.6.2 $\tilde{\boldsymbol{y}}(t+i \mid t+N)$，$\hat{\boldsymbol{w}}(t+i \mid t+N)$ 和 $\hat{\boldsymbol{v}}(t+i \mid t+N)$ 均可用新息 $\boldsymbol{\varepsilon}(t+i)$ 的线性组合表示，可得下述定理给出的表达式.

【定理 5.6.3】 系统(5.6.1)和(5.6.2)在假设 1 ~ 3 下，估值误差 $\tilde{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t \mid t+N)$ 有如下表达式：

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{\beta-2} \boldsymbol{\Omega}_i^w \boldsymbol{w}(t+i) - \sum_{i=0}^{\beta-1} \boldsymbol{\Omega}_i \boldsymbol{v}(t+i) + \sum_{i=0}^{n_0} \boldsymbol{\Omega}_i^c \boldsymbol{\varepsilon}(t+i), \quad N \geqslant 0 \tag{5.6.26}$$

其中定义 $n_0 = \max(\beta-1, N)$，$\boldsymbol{\Omega}_i^c$ 由合并同类项系数阵得到.

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{\beta-2} \boldsymbol{\Omega}_i^w \boldsymbol{w}(t+i) - \sum_{i=0}^{\beta-1} \boldsymbol{\Omega}_i \boldsymbol{v}(t+i) + \sum_{i=0}^{n_1} \boldsymbol{\Omega}_i^y \boldsymbol{\varepsilon}(t+N+1+i), \quad N < 0 \tag{5.6.27}$$

其中定义 $n_1 = \beta - 2 - N$，$\boldsymbol{\Omega}_i^y$ 由合并同类项系数阵得到.

(5.6.26)和(5.6.27)可表为统一形式

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_0} [\boldsymbol{\Omega}_i^w \boldsymbol{w}(t+i) - \boldsymbol{\Omega}_i \boldsymbol{v}(t+i) + \boldsymbol{\Omega}_i^c \boldsymbol{\varepsilon}(t+i)], \quad N \geqslant 0 \tag{5.6.28}$$

其中规定：$\boldsymbol{\Omega}_i^w = \boldsymbol{0}(i > \beta-2)$，$\boldsymbol{\Omega}_i = \boldsymbol{0}(i > \beta-1)$.

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_1} [\boldsymbol{\Omega}_i^w \boldsymbol{w}(t+i) - \boldsymbol{\Omega}_i \boldsymbol{v}(t+i) + \boldsymbol{\Omega}_i^y \boldsymbol{\varepsilon}(t+N+1+i)], \quad N < 0 \tag{5.6.29}$$

其中规定：$\boldsymbol{\Omega}_i^w = \boldsymbol{0}(i > \beta-2)$，$\boldsymbol{\Omega}_i = \boldsymbol{0}(i > \beta-1)$.

证明 当 $N \geqslant 0$ 时，(5.6.25)右边的第四项和第五项，应用引理 5.6.2，可表为 $\boldsymbol{\varepsilon}(t), \cdots, \boldsymbol{\varepsilon}(t+N)$ 的线性组合. 另一方面(5.6.25)右边第一项，当 $\beta - 1 > N$ 时由引理 5.6.2，可表为 $\boldsymbol{\varepsilon}(t+N+1), \cdots, \boldsymbol{\varepsilon}(t+\beta-1)$ 的线性组合. 合并这两种情形引出(5.6.26). 当 $N < 0$ 时，应用引理 5.6.2，(5.6.25)右边的第四、第五项为零，而右边第一项为 $\boldsymbol{\varepsilon}(t+N+1), \cdots, \boldsymbol{\varepsilon}(t+\beta-1)$ 的线性组合，这引出(5.6.27). 证毕. □

【定理 5.6.4】 系统(5.6.1)和(5.6.2)在假设 1 ~ 3 下，Wiener 状态估值器(5.6.19)的估值误差方差阵 $\boldsymbol{P}(N) = \mathrm{E}[\tilde{\boldsymbol{x}}(t \mid t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}(N) = \sum_{i=0}^{n_0} \sum_{j=0}^{n_0} [\boldsymbol{\Omega}_i^w, -\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_i^c] \begin{bmatrix} \boldsymbol{Q}_w \delta_{ij} & \boldsymbol{S}\delta_{ij} & \boldsymbol{\Lambda}_{j-i}^w \\ \boldsymbol{S}^{\mathrm{T}}\delta_{ij} & \boldsymbol{Q}_v \delta_{ij} & \boldsymbol{\Lambda}_{j-i}^v \\ \boldsymbol{\Lambda}_{i-j}^{w\mathrm{T}} & \boldsymbol{\Lambda}_{i-j}^{v\mathrm{T}} & \boldsymbol{Q}_\varepsilon \delta_{ij} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega}_j^{w\mathrm{T}} \\ -\boldsymbol{\Omega}_j^{\mathrm{T}} \\ \boldsymbol{\Omega}_j^{c\mathrm{T}} \end{bmatrix}, \quad N \geqslant 0 \tag{5.6.30}$$

$$\boldsymbol{P}(N) = \sum_{i=0}^{n_1} \sum_{j=0}^{n_1} [\boldsymbol{\Omega}_i^w, -\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_i^y] \begin{bmatrix} \boldsymbol{Q}_w \delta_{ij} & \boldsymbol{S}\delta_{ij} & \boldsymbol{\Lambda}_{N+1+j-i}^w \\ \boldsymbol{S}^{\mathrm{T}}\delta_{ij} & \boldsymbol{Q}_v \delta_{ij} & \boldsymbol{\Lambda}_{N+1+j-i}^v \\ \boldsymbol{\Lambda}_{N+1+i-j}^{w\mathrm{T}} & \boldsymbol{\Lambda}_{N+1+i-j}^{v\mathrm{T}} & \boldsymbol{Q}_\varepsilon \delta_{ij} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega}_j^{w\mathrm{T}} \\ -\boldsymbol{\Omega}_j^{\mathrm{T}} \\ \boldsymbol{\Omega}_j^{y\mathrm{T}} \end{bmatrix}, \quad N < 0 \tag{5.6.31}$$

其中 $n_0 = \max(\beta - 1, N)$，$n_1 = \beta - 2 - N$.

证明　由(5.6.28)和(5.6.29)有简化表达式

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_0} [\boldsymbol{\Omega}_i^w, -\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_i^c] \begin{bmatrix} \boldsymbol{w}(t+i) \\ \boldsymbol{v}(t+i) \\ \boldsymbol{\varepsilon}(t+i) \end{bmatrix}, \quad N \geqslant 0 \tag{5.5.32}$$

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_1} [\boldsymbol{\Omega}_i^w, -\boldsymbol{\Omega}_i, \boldsymbol{\Omega}_i^\gamma] \begin{bmatrix} \boldsymbol{w}(t+i) \\ \boldsymbol{v}(t+i) \\ \boldsymbol{\varepsilon}(t+N+1+i) \end{bmatrix}, \quad N < 0 \tag{5.5.33}$$

由此利用(5.6.3)和(5.6.18)引出(5.6.30)和(5.6.31). □

【注 5.6.1】　在(5.6.10)中合并有关 $\boldsymbol{w}(t+j)$ 的同类项可得到一般公式如下：记有关 $\boldsymbol{w}(t+j)$ 的项为

$$\boldsymbol{\omega}(t) = \sum_{i=1}^{\beta-1} \sum_{j=0}^{i-1} \boldsymbol{\Omega}_i \boldsymbol{H} \boldsymbol{\Phi}^{i-1-j} \boldsymbol{\Gamma} \boldsymbol{w}(t+j) \tag{5.6.34}$$

合并同类项后可表为

$$\boldsymbol{\omega}(t) = \sum_{\rho=0}^{\beta-2} \boldsymbol{K}_\rho \boldsymbol{w}(t+\rho) \tag{5.6.35}$$

其中系数阵 $\boldsymbol{K}_\rho$ 为

$$\boldsymbol{K}_\rho = \sum_{i=\rho+1}^{\beta-1} \boldsymbol{\Omega}_i \boldsymbol{H} \boldsymbol{\Phi}^{i-1-\rho} \boldsymbol{\Gamma}, \quad \rho = 0, \cdots, \beta - 2 \tag{5.6.36}$$

【例 5.6.1】　考虑带有色观测噪声的跟踪系统

$$\boldsymbol{x}(t) = \boldsymbol{\Phi x}(t) + \boldsymbol{\Gamma} w(t) \tag{5.6.37}$$

$$z(t) = \boldsymbol{H}_0 \boldsymbol{x}(t) + \eta(t) \tag{5.6.38}$$

$$\eta(t+1) = p\eta(t) + \xi(t) \tag{5.6.39}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix}, \quad \boldsymbol{H}_0 = [1 \quad 0] \tag{5.6.40}$$

其中 T_0 为采样周期，$\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}$ 是状态，$x_1(t)$，$x_2(t)$ 和 $w(t)$ 各为在采样时刻 tT_0 处运动目标的位置、速度和加速度，$z(t)$ 为对位置的观测信号，$\eta(t)$ 为有色观测噪声，$w(t)$ 和 $\xi(t)$ 是零均值、方差各为 σ_w^2 和 σ_ξ^2 的相互独立高斯白噪声. 取 $N = 0, -1, 1$，问题是求 Wiener 状态估值器 $\hat{\boldsymbol{x}}(t \mid t+N)$.

仿真中取 $T_0 = 0.2$，$p = 0.4$，$\sigma_w^2 = 1$，$\sigma_\xi^2 = 6$. 首先将上述带有色观测噪声系统化为带白色观测噪声系统. 为此，引入新的观测

$$y(t) = z(t+1) - pz(t) \tag{5.6.41}$$

可得带白色观测噪声的观测方程

$$y(t) = \boldsymbol{Hx}(t) + v(t) \tag{5.6.42}$$

其中定义

$$\boldsymbol{H} = \boldsymbol{H}_0 \boldsymbol{\Phi} - p\boldsymbol{H}_0$$

$$v(t) = \boldsymbol{H}_0 \boldsymbol{\Gamma} w(t) + \xi(t) \tag{5.6.43}$$

显然 $v(t)$ 是带零均值、方差为 σ_v^2 的白噪声

$$\sigma_v^2 = \boldsymbol{H}_0 \boldsymbol{\Gamma} \boldsymbol{\Gamma}^{\mathrm{T}} \boldsymbol{H}_0^{\mathrm{T}} \sigma_w^2 + \sigma_\xi^2 \tag{5.6.44}$$

且 $w(t)$ 与 $v(t)$ 的相关系数为

$$S = \mathrm{E}[w(t)v(t)] = \sigma_w^2 \boldsymbol{\Gamma}^{\mathrm{T}} \boldsymbol{H}_0^{\mathrm{T}} \tag{5.6.45}$$

对系统(5.6.37)和(5.6.42),可得 ARMA 新息模型

$$(1 - q^{-1})^2 y(t) = (1 + d_1 q^{-1} + d_2 q^{-2}) \varepsilon(t) \tag{5.6.46}$$

其中 $d_1 = -1.859\,9, d_2 = -0.869\,06, \sigma_\varepsilon^2 = 6.903\,8$.

应用定理 5.6.2 和定理 5.6.4 可得 Wiener 状态估值器 $\hat{\boldsymbol{x}}(t \mid t+N)(N=0,-1,1)$ 及相应的估值误差方差阵 $\boldsymbol{P}(N)$.

$$\boldsymbol{P}(0) = \begin{bmatrix} 1.731\,578\,318\,014\,57 & 0.592\,990\,881\,189\,91 \\ 0.592\,990\,881\,189\,90 & 0.487\,330\,235\,384\,23 \end{bmatrix} \tag{5.6.47}$$

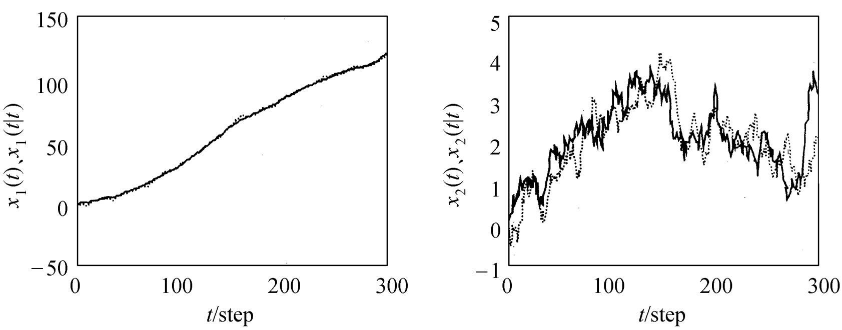
$$\boldsymbol{P}(-1) = \begin{bmatrix} 1.988\,501\,441\,519\,96 & 0.693\,594\,405\,103\,48 \\ 0.693\,594\,405\,103\,48 & 0.526\,723\,610\,669\,95 \end{bmatrix} \tag{5.6.48}$$

$$\boldsymbol{P}(1) = \begin{bmatrix} 1.512\,970\,821\,146\,33 & 0.501\,942\,645\,454\,88 \\ 0.501\,942\,645\,454\,87 & 0.449\,409\,389\,350\,93 \end{bmatrix} \tag{5.6.49}$$

这引出

$$\mathrm{tr}\boldsymbol{P}(0) = 2.218\,9, \quad \mathrm{tr}\boldsymbol{P}(-1) = 2.515\,2, \quad \mathrm{tr}\boldsymbol{P}(1) = 1.962\,3 \tag{5.6.50}$$

即平滑估值精度高于滤波估值精度,而滤波估值精度高于预报估值精度.仿真结果如图 5.6.1 至图 5.6.3 所示,其中实线代表真实值 $\boldsymbol{x}(t)$,虚线代表估值 $\hat{\boldsymbol{x}}(t \mid t+N)$.



(a) 位置 $x_1(t)$ 和 Wiener 跟踪预报器 $\hat{x}_1(t|t)$

(b) 速度 $x_2(t)$ 和 Wiener 跟踪预报器 $\hat{x}_2(t|t)$

图 5.6.1 状态 $\boldsymbol{x}(t)$ 和 Wiener 跟踪滤波器 $\hat{\boldsymbol{x}}(t \mid t)$

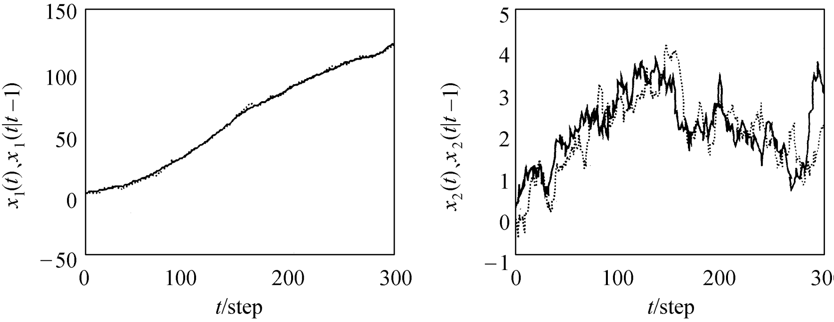
【注 5.6.1】 对系统(5.6.37)和(5.6.42)应用经典 Kalman 滤波方法可求得

$$\boldsymbol{P}(0) = \begin{bmatrix} 1.731\,578\,318\,016\,26 & 0.592\,990\,881\,189\,93 \\ 0.592\,990\,881\,189\,93 & 0.487\,330\,235\,388\,18 \end{bmatrix} \tag{5.6.51}$$

$$\boldsymbol{P}(-1) = \begin{bmatrix} 1.988\,501\,441\,522\,14 & 0.693\,594\,405\,103\,78 \\ 0.693\,594\,405\,103\,78 & 0.526\,723\,610\,674\,05 \end{bmatrix} \tag{5.6.52}$$

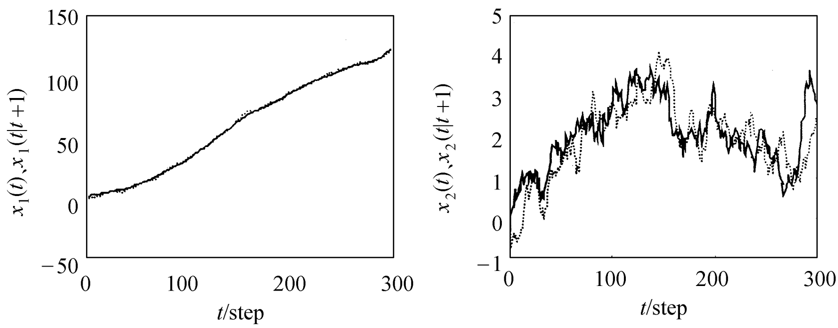
$$\boldsymbol{P}(1) = \begin{bmatrix} 1.512\,970\,821\,147\,69 & 0.501\,942\,645\,454\,60 \\ 0.501\,942\,645\,454\,60 & 0.449\,409\,389\,354\,71 \end{bmatrix} \tag{5.6.53}$$

可看到它们分别相同于(5.6.47)~(5.6.49),其中矩阵的每个元素的值仅末尾几位稍有差别.这说明用两种方法求估值误差方差阵是功能等价的.



(a) 位置 $x_1(t)$ 和 Wiener 跟踪预报器 $\hat{x}_1(t|t-1)$　　(b) 速度 $x_2(t)$ 和 Wiener 跟踪预报器 $\hat{x}_2(t|t-1)$

图 5.6.2　状态 $\boldsymbol{x}(t)$ 和 Wiener 跟踪滤波器 $\hat{\boldsymbol{x}}(t \mid t-1)$



(a) 位置 $x_1(t)$ 和 Wiener 跟踪平滑器 $\hat{x}_1(t|t+1)$　　(b) 速度 $x_2(t)$ 和 Wiener 跟踪平滑器 $\hat{x}_2(t|t+1)$

图 5.6.3　状态 $\boldsymbol{x}(t)$ 和 Wiener 跟踪平滑器 $\hat{\boldsymbol{x}}(t \mid t+1)$

5.7　带白色和有色观测噪声的多通道反卷积滤波器

文献[86 ~ 88]用多项式方法分别提出了不同的多通道 Wiener 反卷积滤波器.它们均需求解矩阵 Diophantine 方程.文献[19,29,1,171]用现代时间序列分析方法提出了新的多通道 Wiener 反卷积滤波器,只需构造 ARMA 新息模型,就可简单地设计多通道 Wiener 反卷积滤波器.节 5.5 已提出一种基于 ARMA 新息模型和白噪声估值器及观测预报器的 Wiener 反卷积滤波器设计方法.本节介绍基于 ARMA 新息模型和状态空间模型设计反卷积滤波器的方法[171].关键思想是将反卷积问题转化为状态估计问题.从而可利用节 5.6 的结果解决问题.

考虑一般的多通道反卷积系统

$$\boldsymbol{y}(t) = \boldsymbol{\Pi}^{-1}(q^{-1})\boldsymbol{\Psi}(q^{-1})\boldsymbol{s}(t) + \boldsymbol{L}^{-1}(q^{-1})\boldsymbol{M}(q^{-1})\boldsymbol{v}(t) + \boldsymbol{\xi}(t) \tag{5.7.1}$$

$$\boldsymbol{s}(t) = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.7.2}$$

其中 $\boldsymbol{y}(t) \in R^m$ 为观测,$\boldsymbol{s}(t) \in R^n$ 为输入信号,$\boldsymbol{\Pi}(q^{-1}), \cdots, \boldsymbol{C}(q^{-1})$ 为单位滞后算子 q^{-1} 的多项式矩阵,形如 $\boldsymbol{X}(q^{-1}) = \boldsymbol{X}_0 + \boldsymbol{X}_1 q^{-1} + \cdots + \boldsymbol{X}_{n_x} q^{-n_x}$,$\boldsymbol{\Pi}_0 = \boldsymbol{I}_m$,$\boldsymbol{L}_0 = \boldsymbol{I}_m$,$\boldsymbol{A}_0 = \boldsymbol{I}_n$,$n_a \geqslant n_c$,$n_\pi \geqslant n_\phi$,$n_l \geqslant n_m$.

【假设 1】 $\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}(t) \in R^m$ 是带零均值的相关白噪声

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix}[\boldsymbol{w}^{\mathrm{T}}(j), \boldsymbol{v}^{\mathrm{T}}(j)]\right\} = \begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}\delta_{tj}, \quad \begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix} > 0 \tag{5.7.3}$$

其中E为均值号，$\delta_{tt} = 1$，$\delta_{tj} = 0(t \neq j)$，T为转置号，且它们独立于带零均值、方差阵为 $\boldsymbol{Q}_\xi$ 的白噪声 $\xi(t)$.

引入左素分解

$$\boldsymbol{P}^{-1}(q^{-1})[\boldsymbol{B}(q^{-1}), \boldsymbol{R}(q^{-1})] = [\boldsymbol{\Pi}^{-1}(q^{-1})\boldsymbol{\Psi}(q^{-1}), \boldsymbol{L}^{-1}(q^{-1})\boldsymbol{M}(q^{-1})] \tag{5.7.4}$$

带 $\boldsymbol{P}_0 = \boldsymbol{I}_m$，则(5.7.1)和(5.7.2)化为等价系统

$$\boldsymbol{P}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{B}(q^{-1})\boldsymbol{s}(t) + \boldsymbol{R}(q^{-1})\boldsymbol{v}(t) + \boldsymbol{P}(q^{-1})\boldsymbol{\xi}(t) \tag{5.7.5}$$

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.7.6}$$

引入左素分解

$$\boldsymbol{B}(q^{-1})\boldsymbol{A}^{-1}(q^{-1}) = \widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) \quad (\widetilde{\boldsymbol{A}}_0 = \boldsymbol{I}_m) \tag{5.7.7}$$

【假设 2】 设 $(\widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{P}(q^{-1}), \widetilde{\boldsymbol{B}}(q^{-1})\boldsymbol{C}(q^{-1}), \widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{R}(q^{-1}))$ 左素，初始观测时刻 $t_0 = -\infty$.

Wiener 反卷积问题是：基于观测 $(\boldsymbol{y}(t+N), \boldsymbol{y}(t+N-1), \cdots)$ 求输入信号 $\boldsymbol{s}(t)$ 的 Wiener 反卷积滤波器 $\hat{\boldsymbol{s}}(t \mid t+N)$，$N = 0$，$N > 0$ 或 $N < 0$.

将(5.7.6)代入(5.7.5)并利用(5.7.7)，由假设1和假设2可得ARMA新息模型

$$\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.7.8}$$

$$\boldsymbol{\Lambda}(q^{-1}) = \widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{P}(q^{-1}) \tag{5.7.9}$$

其中 $\boldsymbol{D}(q^{-1})$ 是稳定的，$\boldsymbol{D}_0 = \boldsymbol{I}_m$，新息 $\boldsymbol{\varepsilon}(t) \in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \widetilde{\boldsymbol{B}}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{w}(t) + \widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{R}(q^{-1})\boldsymbol{v}(t) + \widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{P}(q^{-1})\boldsymbol{\xi}(t) \tag{5.7.10}$$

$\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers - Wouters[1] 算法求得.

【引理 5.7.1】 系统(5.7.1)和(5.7.2)在假设 1 ~ 2 下有白噪声滤波器

$$\hat{\boldsymbol{\theta}}(t \mid t+N) = \boldsymbol{L}_N^\theta(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.7.11}$$

$$\hat{\boldsymbol{\theta}}(t \mid t+N) = \boldsymbol{L}_N^\theta(q^{-1})\widetilde{\boldsymbol{\Lambda}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N), \quad \boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v}, \boldsymbol{\xi} \tag{5.7.12}$$

其中定义

$$\begin{aligned}
&\boldsymbol{L}_N^\theta(q^{-1}) = \sum_{i=0}^{N}\boldsymbol{\Lambda}_i^\theta\boldsymbol{Q}_\varepsilon^{-1}q^{i-N}, \quad N \geqslant 0, \theta = w, v, \xi,\\
&\boldsymbol{L}_N^\theta(q^{-1}) = \boldsymbol{0}, \quad N < 0,\\
&\boldsymbol{\Lambda}_i^w = \boldsymbol{Q}_w\boldsymbol{F}_i^{v\mathrm{T}} + \boldsymbol{S}\boldsymbol{F}_i^{v\mathrm{T}},\\
&\boldsymbol{\Lambda}_i^v = \boldsymbol{Q}_v\boldsymbol{F}_i^{v\mathrm{T}} + \boldsymbol{S}^{\mathrm{T}}\boldsymbol{F}_i^{w\mathrm{T}},\\
&\boldsymbol{\Lambda}_i^\xi = \boldsymbol{Q}_\xi\boldsymbol{F}_i^{\xi\mathrm{T}}
\end{aligned} \tag{5.7.13}$$

其中定义 $(\boldsymbol{F}_i^{\theta\mathrm{T}}) = (\boldsymbol{F}_i^\theta)^{\mathrm{T}}$，T为转置号，$\boldsymbol{F}_i^\theta$ 可递推计算为

$$\boldsymbol{F}_i^\theta = -\boldsymbol{D}_1\boldsymbol{F}_{i-1}^\theta - \cdots - \boldsymbol{D}_{n_d}\boldsymbol{F}_{i-n_d}^\theta + \boldsymbol{\Psi}_i^\theta, \quad \theta = w, v, \xi \tag{5.7.14}$$

其中规定 $\boldsymbol{F}_i^\theta = 0(i < 0)$，$\boldsymbol{\Psi}_i^\theta = 0(i > n_{\psi\theta})$，其中定义 $n_{\psi\theta}$ 为 $\boldsymbol{\Psi}^\theta(q^{-1})$ 的阶次，且 $\boldsymbol{\Psi}^\theta(q^{-1})$ 定义为

$$\boldsymbol{\Psi}^w(q^{-1}) = \widetilde{\boldsymbol{B}}(q^{-1})\boldsymbol{C}(q^{-1}),\boldsymbol{\Psi}^v(q^{-1}) = \widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{R}(q^{-1}),$$
$$\boldsymbol{\Psi}^\xi(q^{-1}) = \widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{P}(q^{-1}) \tag{5.7.15}$$

由如下伪交换定义 $\widetilde{\boldsymbol{\Lambda}}(q^{-1})$ 和 $\widetilde{\boldsymbol{D}}(q^{-1})$ 为

$$\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{\Lambda}(q^{-1}) = \widetilde{\boldsymbol{\Lambda}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{5.7.16}$$

带 $\widetilde{\boldsymbol{D}}_0 = \boldsymbol{I}_m, \det\widetilde{\boldsymbol{D}}(q^{-1}) = \det\boldsymbol{D}(q^{-1})$.

【引理 5.7.2】 系统(5.7.1)和(5.7.2)在假设 1 ~ 2 下有 Åström 预报器

$$\hat{\boldsymbol{y}}(t \mid t+N) = \boldsymbol{J}_{-N}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.7.17}$$

其中 $\boldsymbol{J}_{-N}(q^{-1})$ 定义为

$$\widetilde{\boldsymbol{D}}(q^{-1}) = \boldsymbol{E}_{-N}(q^{-1})\widetilde{\boldsymbol{\Lambda}}(q^{-1}) + q^N\boldsymbol{J}_{-N}(q^{-1}), \quad N < 0$$
$$\boldsymbol{J}_{-N}(q^{-1}) = \widetilde{\boldsymbol{D}}(q^{-1})q^{-N}, \quad N \geqslant 0 \tag{5.7.18}$$

证明 详见引理 5.3.2 和推论 5.3.2. □

【引理 5.7.3】 系统(5.7.1)和(5.7.2)在假设 1 ~ 2 下,有关系

$$\mathrm{E}[\boldsymbol{\theta}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)] = \boldsymbol{\Lambda}_{i-t}^{\theta}, \quad \boldsymbol{\theta} = \boldsymbol{w},\boldsymbol{v},\boldsymbol{\xi} \tag{5.7.19}$$

其中 $\boldsymbol{\Lambda}_i^\theta$ 由(5.7.13)给出.

下面将等价反卷积系统(5.7.5)和(5.7.6)用状态空间模型表示,从而可把反卷积问题归结为状态估计问题,进而可应用节 5.6 的 Wiener 状态估值器解决问题.

由定理 1.6.1(5.7.6)有状态空间模型

$$\boldsymbol{\alpha}(t+1) = \bar{\boldsymbol{A}}\boldsymbol{\alpha}(t) + \bar{\boldsymbol{C}}\boldsymbol{w}(t) \tag{5.7.20}$$
$$\boldsymbol{s}(t) = \bar{\boldsymbol{H}}_1\boldsymbol{\alpha}(t) + \boldsymbol{C}_0\boldsymbol{w}(t) \tag{5.7.21}$$

其中定义 $\boldsymbol{C}_i = \boldsymbol{0}(i > n_c)$,

$$\bar{\boldsymbol{A}} = \begin{bmatrix} -\boldsymbol{A}_1 & & & \\ \vdots & & \boldsymbol{I}_{(n_a-1)n} & \\ -\boldsymbol{A}_{n_a} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix}, \quad \bar{\boldsymbol{C}} = \begin{bmatrix} \boldsymbol{C}_1 - \boldsymbol{A}_1\boldsymbol{C}_0 \\ \vdots \\ \boldsymbol{C}_{n_a} - \boldsymbol{A}_{n_a}\boldsymbol{C}_0 \end{bmatrix}, \quad \bar{\boldsymbol{H}}_1 = [\boldsymbol{I}_n \quad 0 \cdots 0] \tag{5.7.22}$$

(5.7.5)有状态空间模型

$$\boldsymbol{\beta}(t+1) = \bar{\boldsymbol{P}}\boldsymbol{\beta}(t) + \bar{\boldsymbol{B}}\boldsymbol{s}(t) + \bar{\boldsymbol{R}}\boldsymbol{v}(t) \tag{5.7.23}$$
$$\boldsymbol{y}(t) = \bar{\boldsymbol{H}}_2\boldsymbol{\beta}(t) + \boldsymbol{B}_0\boldsymbol{s}(t) + \boldsymbol{R}_0\boldsymbol{v}(t) + \boldsymbol{\xi}(t) \tag{5.7.24}$$

其中定义 $\boldsymbol{B}_i = \boldsymbol{0}(i > n_b), \boldsymbol{R}_i = \boldsymbol{0}(i > n_r)$,且

$$\bar{\boldsymbol{P}} = \begin{bmatrix} -\boldsymbol{P}_1 & & & \\ \vdots & & \boldsymbol{I}_{(n_p-1)m} & \\ -\boldsymbol{P}_{n_p} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix}, \quad \bar{\boldsymbol{B}} = \begin{bmatrix} \boldsymbol{B}_1 - \boldsymbol{P}_1\boldsymbol{B}_0 \\ \vdots \\ \boldsymbol{B}_{n_p} - \boldsymbol{P}_{n_p}\boldsymbol{B}_0 \end{bmatrix},$$

$$\bar{\boldsymbol{R}} = \begin{bmatrix} \boldsymbol{R}_1 - \boldsymbol{P}_1\boldsymbol{R}_0 \\ \vdots \\ \boldsymbol{R}_{n_p} - \boldsymbol{P}_{n_p}\boldsymbol{R}_0 \end{bmatrix}, \quad \bar{\boldsymbol{H}}_2 = [\boldsymbol{I}_m \quad \boldsymbol{0} \cdots \boldsymbol{0}] \tag{5.7.25}$$

于是有增广状态空间模型

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}\boldsymbol{x}(t) + \boldsymbol{\Gamma}\boldsymbol{e}(t) \tag{5.7.26}$$
$$\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{\delta}(t) \tag{5.7.27}$$

其中定义

$$\boldsymbol{x}(t)=\begin{bmatrix}\boldsymbol{\alpha}(t)\\ \boldsymbol{\beta}(t)\end{bmatrix},\quad \boldsymbol{e}(t)=\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix}$$

$$\boldsymbol{\delta}(t)=\boldsymbol{B}_0\boldsymbol{C}_0\boldsymbol{w}(t)+\boldsymbol{R}_0\boldsymbol{v}(t)+\boldsymbol{\xi}(t)$$

$$\boldsymbol{\Phi}=\begin{bmatrix}\bar{\boldsymbol{A}} & \boldsymbol{0}\\ \bar{\boldsymbol{B}}\bar{\boldsymbol{H}}_1 & \bar{\boldsymbol{P}}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\bar{\boldsymbol{C}} & \boldsymbol{0}\\ \bar{\boldsymbol{B}}\boldsymbol{C}_0 & \bar{\boldsymbol{R}}\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{B}_0\bar{\boldsymbol{H}}_1,\bar{\boldsymbol{H}}_2] \tag{5.7.28}$$

可证明[31]增广系统(5.7.26)和(5.7.27)是完全可观、完全可控的，设可观性指数为β.因而存在稳态估值器$\hat{\boldsymbol{x}}(t\mid t+N)$.

由(5.7.28)中白噪声$\boldsymbol{e}(t)$和$\boldsymbol{\delta}(t)$的定义有

$$\hat{\boldsymbol{e}}(t\mid t+N)=\boldsymbol{L}_N^e(q^{-1})\tilde{\boldsymbol{\Lambda}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.7.29}$$

$$\boldsymbol{L}_N^e(q^{-1})=\begin{bmatrix}\boldsymbol{L}_N^w(q^{-1})\\ \boldsymbol{L}_N^v(q^{-1})\end{bmatrix} \tag{5.7.30}$$

$$\begin{aligned}\hat{\boldsymbol{\delta}}(t\mid t+N)&=\boldsymbol{B}_0\boldsymbol{C}_0\hat{\boldsymbol{w}}(t\mid t+N)+\boldsymbol{R}_0\hat{\boldsymbol{v}}(t\mid t+N)+\hat{\boldsymbol{\xi}}(t\mid t+N)=\\ &\boldsymbol{L}_N^\delta(q^{-1})\tilde{\boldsymbol{\Lambda}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N)\end{aligned} \tag{5.7.31}$$

$$\boldsymbol{L}_N^\delta(q^{-1})=\boldsymbol{B}_0\boldsymbol{C}_0\boldsymbol{L}_N^w(q^{-1})+\boldsymbol{R}_0\boldsymbol{L}_N^v(q^{-1})+\boldsymbol{L}_N^\xi(q^{-1}) \tag{5.7.32}$$

对增广系统应用定理5.6.1有Wiener状态估值器

$$\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.7.33}$$

其中多项式矩阵$\boldsymbol{K}_N(q^{-1})$定义为

$$\boldsymbol{K}_N(q^{-1})=\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i\Big[\boldsymbol{J}_{i-N}(q^{-1})-\sum_{j=0}^{i-1}\boldsymbol{H}\boldsymbol{\Phi}^{i-1-j}\boldsymbol{\Gamma}\boldsymbol{L}_{N-j}^e(q^{-1})\tilde{\boldsymbol{\Lambda}}(q^{-1})-\boldsymbol{L}_{N-i}^\delta(q^{-1})\tilde{\boldsymbol{\Lambda}}(q^{-1})\Big] \tag{5.7.34}$$

其中规定$\boldsymbol{\Phi}^i=\boldsymbol{0}(i<0)$且$j\geqslant 0$，$\beta$为该系统可观性指数，

$$\boldsymbol{\Omega}=[\boldsymbol{H}^{\mathrm{T}},(\boldsymbol{H}\boldsymbol{\Phi})^{\mathrm{T}},\cdots,(\boldsymbol{H}\boldsymbol{\Phi}^{\beta-1})^{\mathrm{T}}]^{\mathrm{T}} \tag{5.7.35}$$

且伪逆$\boldsymbol{\Omega}^+=(\boldsymbol{\Omega}^{\mathrm{T}}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathrm{T}}$分块表示为

$$\boldsymbol{\Omega}^+=[\boldsymbol{\Omega}_0,\boldsymbol{\Omega}_1,\cdots,\boldsymbol{\Omega}_{\beta-1}] \tag{5.7.36}$$

由(5.7.21)和射影性质引出

$$\hat{\boldsymbol{s}}(t\mid t+N)=[\boldsymbol{I}_n\quad \boldsymbol{0}]\hat{\boldsymbol{x}}(t\mid t+N)+\boldsymbol{C}_0\hat{\boldsymbol{w}}(t\mid t+N) \tag{5.7.37}$$

其中零为适当维数零阵.

【定理5.7.1】 多通道反卷积系统(5.7.1)和(5.7.2)或(5.7.5)和(5.7.6)在假设1～2下有渐近稳定的Wiener反卷积滤波器

$$\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{K}_N^s(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.7.38}$$

或表为ARMA递推形式

$$\det\tilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{K}_N^s(q^{-1})\mathrm{adj}\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.7.39}$$

其中多项式矩阵$\boldsymbol{K}_N^s(q^{-1})$定义为

$$\boldsymbol{K}_N^s(q^{-1})=[\boldsymbol{I}_n\quad \boldsymbol{0}]\boldsymbol{K}_N(q^{-1})+\boldsymbol{C}_0\boldsymbol{L}_N^w(q^{-1})\tilde{\boldsymbol{\Lambda}}(q^{-1}) \tag{5.7.40}$$

其中$\boldsymbol{K}_N(q^{-1})$由(5.7.34)定义.

证明 将(5.7.11)和(5.7.33)代入(5.7.37)有

$$\hat{\boldsymbol{s}}(t \mid t+N) = [\boldsymbol{I}_n \quad \boldsymbol{0}]\boldsymbol{K}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) + \boldsymbol{C}_0\boldsymbol{L}_N^w(q^{-1})\widetilde{\boldsymbol{\Lambda}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.7.41}$$

这引出(5.7.38)～(5.7.40).由 $\det\widetilde{\boldsymbol{D}}(q^{-1})$ 是稳定的多项式有(5.7.38)或(5.7.39)是渐近稳定的. □

估值误差 $\tilde{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{s}(t) - \hat{\boldsymbol{s}}(t \mid t+N)$ 有表达式

$$\tilde{\boldsymbol{s}}(t \mid t+N) = [\boldsymbol{I}_n \quad \boldsymbol{0}]\tilde{\boldsymbol{x}}(t \mid t+N) + \boldsymbol{C}_0\widetilde{\boldsymbol{w}}(t \mid t+N) \tag{5.7.42}$$

其中定义

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t \mid t+N),$$
$$\widetilde{\boldsymbol{w}}(t \mid t+N) = \boldsymbol{w}(t) - \hat{\boldsymbol{w}}(t \mid t+N) \tag{5.7.43}$$

对增广系统(5.7.26)～(5.7.28),由(5.6.28)和(5.6.29)有一般表达式

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_0}[\boldsymbol{\Omega}_i^e\boldsymbol{e}(t+i) - \boldsymbol{\Omega}_i\boldsymbol{\delta}(t+i) + \boldsymbol{\Omega}_i^c\boldsymbol{\varepsilon}(t+i)], \quad N \geqslant 0 \tag{5.7.44}$$

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_1}[\boldsymbol{\Omega}_i^e\boldsymbol{e}(t+i) - \boldsymbol{\Omega}_i\boldsymbol{\delta}(t+i) + \boldsymbol{\Omega}_i^y\boldsymbol{\varepsilon}(t+N+1+i)], N < 0 \tag{5.7.45}$$

其中 $n_0 = \max(\beta-1, N)$, $n_1 = \beta - 2 - N$.

由(5.7.11)有

$$\hat{\boldsymbol{w}}(t \mid t+N) = \sum_{i=0}^{N}\boldsymbol{\Lambda}_i^w\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i), \quad N \geqslant 0$$
$$\hat{\boldsymbol{w}}(t \mid t+N) = \boldsymbol{0}, \quad N < 0 \tag{5.7.46}$$

将(5.7.44)～(5.7.46)代入(5.7.42)并合并同类项可得一般表达式

$$\tilde{\boldsymbol{s}}(t \mid t+N) = \sum_{i=0}^{n_0}[\boldsymbol{\Omega}_i^{(1)}\boldsymbol{e}(t+i) - \boldsymbol{\Omega}_i^{(2)}\boldsymbol{\delta}(t+i) + \boldsymbol{\Omega}_i^{(3)}\boldsymbol{\varepsilon}(t+i)], \quad N \geqslant 0 \tag{5.7.47}$$

$$\tilde{\boldsymbol{s}}(t \mid t+N) = \sum_{i=0}^{n_1}[\boldsymbol{\Omega}_i^{(1)}\boldsymbol{e}(t+i) - \boldsymbol{\Omega}_i^{(2)}\boldsymbol{\delta}(t+i) + \boldsymbol{\Omega}_i^{(4)}\boldsymbol{\varepsilon}(t+N+1+i)], \quad N < 0 \tag{5.7.48}$$

注意(5.7.28)中 $\boldsymbol{e}(t)$ 和 $\boldsymbol{\delta}(t)$ 的定义,由(5.7.3)引出 $\boldsymbol{e}(t)$ 和 $\boldsymbol{\delta}(t)$ 是相关的白噪声,它们的方差阵 $\boldsymbol{Q}_e$, $\boldsymbol{Q}_\delta$ 和相关阵 $\boldsymbol{S}_0$ 分别为

$$\boldsymbol{Q}_e = \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v \end{bmatrix},$$
$$\boldsymbol{Q}_\delta = \boldsymbol{B}_0\boldsymbol{C}_0\boldsymbol{Q}_w\boldsymbol{C}_0^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}} + \boldsymbol{R}_0\boldsymbol{Q}_v\boldsymbol{R}_0^{\mathrm{T}} + \boldsymbol{Q}_\xi,$$
$$\boldsymbol{S}_0 = \mathrm{E}[\boldsymbol{e}(t)\boldsymbol{\delta}^{\mathrm{T}}(t)] = \begin{bmatrix} \boldsymbol{Q}_w\boldsymbol{C}_0^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}} + \boldsymbol{S}\boldsymbol{R}_0^{\mathrm{T}} \\ \boldsymbol{S}^{\mathrm{T}}\boldsymbol{C}_0^{\mathrm{T}}\boldsymbol{B}_0^{\mathrm{T}} + \boldsymbol{Q}_v\boldsymbol{R}_0^{\mathrm{T}} \end{bmatrix} \tag{5.7.49}$$

由(5.7.19)和(5.7.28)有

$$\mathrm{E}[\boldsymbol{e}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)] = \boldsymbol{\Lambda}_{i-t}^e, \quad \mathrm{E}[\boldsymbol{\delta}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(i)] = \boldsymbol{\Lambda}_{i-t}^\delta \tag{5.7.50}$$

$$\boldsymbol{\Lambda}_{i-t}^{e}=\begin{bmatrix}\boldsymbol{\Lambda}_{i-t}^{w}\\ \boldsymbol{\Lambda}_{i-t}^{v}\end{bmatrix},$$

$$\boldsymbol{\Lambda}_{i-t}^{\delta}=\boldsymbol{B}_0\boldsymbol{C}_0\boldsymbol{\Lambda}_{i-t}^{w}+\boldsymbol{R}_0\boldsymbol{\Lambda}_{i-t}^{v}+\boldsymbol{\Lambda}_{i-t}^{\xi}\tag{5.7.51}$$

【定理5.7.2】 在定理5.7.1条件下，当 $\boldsymbol{C}_0\neq\boldsymbol{0}$ 时，滤波误差方差阵 $\boldsymbol{P}_s(N)=\mathrm{E}[\tilde{\boldsymbol{s}}(t\mid t+N)\tilde{\boldsymbol{s}}^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}_s(N)=\sum_{i=0}^{n_0}\sum_{j=0}^{n_0}[\boldsymbol{\Omega}_i^{(1)},\boldsymbol{\Omega}_i^{(2)},\boldsymbol{\Omega}_i^{(3)}]\begin{bmatrix}\boldsymbol{Q}_e\delta_{ij} & \boldsymbol{S}_0\delta_{ij} & \boldsymbol{\Lambda}_{j-i}^{e}\\ \boldsymbol{S}_0^{\mathrm{T}}\delta_{ij} & \boldsymbol{Q}_\delta\delta_{ij} & \boldsymbol{\Lambda}_{j-i}^{\delta}\\ \boldsymbol{\Lambda}_{i-j}^{e\mathrm{T}} & \boldsymbol{\Lambda}_{i-j}^{\delta\mathrm{T}} & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}_j^{(1)\mathrm{T}}\\ \boldsymbol{\Omega}_j^{(2)\mathrm{T}}\\ \boldsymbol{\Omega}_j^{(3)\mathrm{T}}\end{bmatrix},\quad N\geqslant 0\tag{5.7.52}$$

$$\boldsymbol{P}_s(N)=\sum_{i=0}^{n_1}\sum_{j=0}^{n_1}[\boldsymbol{\Omega}_i^{(1)},\boldsymbol{\Omega}_i^{(2)},\boldsymbol{\Omega}_i^{(4)}]\begin{bmatrix}\boldsymbol{Q}_e\delta_{ij} & \boldsymbol{S}_0\delta_{ij} & \boldsymbol{\Lambda}_{N+1+j-i}^{e}\\ \boldsymbol{S}_0\delta_{ij} & \boldsymbol{Q}_\delta\delta_{ij} & \boldsymbol{\Lambda}_{N+1+j-i}^{\delta}\\ \boldsymbol{\Lambda}_{N+1+i-j}^{e\mathrm{T}} & \boldsymbol{\Lambda}_{N+1+i-j}^{\delta\mathrm{T}} & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}_j^{(1)\mathrm{T}}\\ \boldsymbol{\Omega}_j^{(2)\mathrm{T}}\\ \boldsymbol{\Omega}_j^{(3)\mathrm{T}}\end{bmatrix},\quad N<0\tag{5.7.53}$$

证明 由(5.7.47)～(5.7.51)类似于定理5.6.4的证明引出(5.7.52)和(5.7.53). □

【定理5.7.3】 对系统(5.7.1)和(5.7.2)或(5.7.5)和(5.7.6)，当 $\boldsymbol{C}_0=0$，Wiener反卷积滤波器(5.7.38)或(5.7.39)的滤波误差方差阵 $\boldsymbol{P}_s(N)$ 为

$$\boldsymbol{P}_s(N)=[\boldsymbol{I}_n\quad\boldsymbol{0}]\boldsymbol{P}(N)[\boldsymbol{I}_n\quad\boldsymbol{0}]^{\mathrm{T}}\tag{5.7.54}$$

其中 $\boldsymbol{P}(N)$ 为状态估计误差方差阵，由定理5.6.4计算，即由(5.6.30)和(5.6.31)计算.

证明 当 $\boldsymbol{C}_0=\boldsymbol{0}$ 时，(5.7.42)成为

$$\tilde{\boldsymbol{s}}(t\mid t+N)=[\boldsymbol{I}_n\quad\boldsymbol{0}]\tilde{\boldsymbol{x}}(t\mid t+N)\tag{5.7.55}$$

这引出(5.7.54). □

【推论5.7.1】 单通道反卷积系统(5.7.1)和(5.7.2)或(5.7.5)和(5.7.6)有渐近稳定的Wiener反卷积滤波器为

$$D(q^{-1})\hat{s}(t\mid t+N)=K_N^s(q^{-1})y(t+N)\tag{5.7.56}$$

其中多项式 $K_N^s(q^{-1})$ 定义为

$$K_N^s(q^{-1})=[1\quad 0\cdots 0]\boldsymbol{K}_N(q^{-1})+c_0L_N^w(q^{-1})\Lambda(q^{-1})\tag{5.7.57}$$

$$\boldsymbol{K}_N(q^{-1})=\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i\left[J_{i-N}(q^{-1})-\sum_{j=0}^{i-1}\boldsymbol{H}\boldsymbol{\Phi}^{i-1-j}\boldsymbol{T}\boldsymbol{L}_{N-j}^{e}(q^{-1})\Lambda(q^{-1})-L_{N-i}^{\delta}(q^{-1})\Lambda(q^{-1})\right]\tag{5.7.58}$$

且 $\Lambda(q^{-1})=A(q^{-1})P(q^{-1})$，ARMA新息模型为

$$\Lambda(q^{-1})y(t)=D(q^{-1})\varepsilon(t)\tag{5.7.59}$$

且有关系

$$D(q^{-1})\varepsilon(t)=B(q^{-1})C(q^{-1})w(t)+A(q^{-1})R(q^{-1})v(t)+A(q^{-1})P(q^{-1})\xi(t)\tag{5.7.60}$$

且估值误差方差 $P_s(N)=\mathrm{E}[\tilde{s}^2(t\mid t+N)]$ 由定理5.7.2和定理5.7.3计算.

证明 对单通道系统($n=1,m=1,r=1$)，有 $\widetilde{\boldsymbol{A}}(q^{-1})=A(q^{-1})$，$\widetilde{\boldsymbol{B}}(q^{-1})=$

$B(q^{-1}),\widetilde{\boldsymbol{D}}(q^{-1}) = D(q^{-1}),\widetilde{\boldsymbol{\Lambda}}(q^{-1}) = \Lambda(q^{-1})$，由定理 5.7.1 ~ 定理 5.7.3 得证. □

考虑如下多通道 Wiener 滤波问题：

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.7.61}$$

$$\boldsymbol{y}(t) = \boldsymbol{s}(t) + \boldsymbol{\eta}(t) + \boldsymbol{\xi}(t) \tag{5.7.62}$$

$$\boldsymbol{P}(q^{-1})\boldsymbol{\eta}(t) = \boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \tag{5.7.63}$$

其中 $\boldsymbol{s}(t) \in R^m$ 为待估信号，$\boldsymbol{y}(t) \in R^m$ 是观测信号，$\boldsymbol{\eta}(t) \in R^m$ 为服从 ARMA 模型(5.7.63)的有色观测噪声，且设 $\boldsymbol{\xi}(t) \in R^m$，$\boldsymbol{v}(t) \in R^m$ 和 $\boldsymbol{w}(t) \in R^r$ 是零均值、方差为 $\boldsymbol{Q}_\xi$，$\boldsymbol{Q}_v$ 和 $\boldsymbol{Q}_w$ 的相互独立白噪声：

$$\mathrm{E}\left\{\begin{bmatrix} \boldsymbol{w}(t) \\ \boldsymbol{v}(t) \\ \boldsymbol{\xi}(t) \end{bmatrix}\begin{bmatrix} \boldsymbol{w}^{\mathrm{T}}(j) & \boldsymbol{v}^{\mathrm{T}}(j) & \boldsymbol{\xi}^{\mathrm{T}}(j) \end{bmatrix}\right\} = \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_v & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Q}_\xi \end{bmatrix}\delta_{ij} \tag{5.7.64}$$

问题是当 $t_0 = -\infty$ 时基于观测 $(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots)$，求信号 $\boldsymbol{s}(t)$ 的渐近稳定的最优 Wiener 滤波器 $\hat{\boldsymbol{s}}(t \mid t+N)$.

Wiener 滤波问题(5.7.61) ~ (5.7.63) 可看成是特殊的反卷积问题. 事实上将(5.7.63)代入(5.7.62)有如下等价的特殊反卷积系统(5.7.5)和(5.7.6)

$$\boldsymbol{P}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{P}(q^{-1})\boldsymbol{s}(t) + \boldsymbol{R}(q^{-1})\boldsymbol{v}(t) + \boldsymbol{P}(q^{-1})\boldsymbol{\xi}(t) \tag{5.7.65}$$

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.7.66}$$

因此在定理 5.7.1 ~ 定理 5.7.3 中置 $\boldsymbol{B}(q^{-1}) = \boldsymbol{P}(q^{-1})$ 就可得到相应的渐近稳定的最优 Wiener 滤波器 $\hat{\boldsymbol{s}}(t \mid t+N)$. 由 $\boldsymbol{B}_0 = \boldsymbol{P}_0 = \boldsymbol{I}_m$，$\boldsymbol{B}_i = \boldsymbol{P}_i$，由(5.7.25)有 $\bar{\boldsymbol{B}} = \boldsymbol{0}$，故状态空间模型(5.7.23)和(5.7.24)成立

$$\boldsymbol{\beta}(t+1) = \bar{\boldsymbol{P}}\boldsymbol{\beta}(t) + \bar{\boldsymbol{R}}\boldsymbol{v}(t) \tag{5.7.67}$$

$$\boldsymbol{y}(t) = \bar{\boldsymbol{H}}_2\boldsymbol{\beta}(t) + \boldsymbol{R}_0\boldsymbol{v}(t) + \boldsymbol{s}(t) + \boldsymbol{\xi}(t) \tag{5.7.68}$$

且在增广状态空间模型(5.7.26) ~ (5.7.28)中有

$$\boldsymbol{\Phi} = \begin{bmatrix} \bar{\boldsymbol{A}} & \boldsymbol{0} \\ \boldsymbol{0} & \bar{\boldsymbol{P}} \end{bmatrix},\quad \boldsymbol{\Gamma} = \begin{bmatrix} \bar{\boldsymbol{C}} & \boldsymbol{0} \\ \boldsymbol{0} & \bar{\boldsymbol{R}} \end{bmatrix},\quad \boldsymbol{H} = [\bar{\boldsymbol{H}}_1,\bar{\boldsymbol{H}}_2],$$

$$\boldsymbol{\delta}(t) = \boldsymbol{C}_0\boldsymbol{w}(t) + \boldsymbol{R}_0\boldsymbol{v}(t) + \boldsymbol{\xi}(t) \tag{5.7.69}$$

伪交换(5.7.7)成为

$$\boldsymbol{P}(q^{-1})\boldsymbol{A}^{-1}(q^{-1}) = \widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{P}}(q^{-1}) \tag{5.7.70}$$

且由(5.7.57)在(5.7.13)中有

$$\boldsymbol{\Lambda}_i^w = \boldsymbol{Q}_w\boldsymbol{F}_i^{w\mathrm{T}},\quad \boldsymbol{\Lambda}_i^v = \boldsymbol{Q}_v\boldsymbol{F}_i^{v\mathrm{T}} \tag{5.7.71}$$

且在(5.7.15)中有

$$\boldsymbol{\Psi}^w(q^{-1}) = \widetilde{\boldsymbol{P}}(q^{-1})\boldsymbol{C}(q^{-1}) \tag{5.7.72}$$

上述推导可概括为如下定理.

【定理 5.7.5】 多通道 Wiener 滤波问题(5.7.61) ~ (5.7.64)有渐近稳定的 Wiener 滤波器

$$\hat{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{K}_N^s(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.7.73}$$

或表为 ARMA 递推形式

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{K}_N^s(q^{-1})\mathrm{adj}\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.7.74}$$

其中多项式矩阵 $\boldsymbol{K}_N^s(q^{-1})$ 定义为

$$\boldsymbol{K}_N^s(q^{-1}) = [\boldsymbol{I}_m \quad \boldsymbol{0}]\boldsymbol{K}_N(q^{-1}) + \boldsymbol{C}_0\boldsymbol{L}_N^w(q^{-1})\widetilde{\boldsymbol{\Lambda}}(q^{-1}) \tag{5.7.75}$$

而 $\boldsymbol{K}_N(q^{-1})$ 由(5.7.34)计算，其中 β 为增广系统(5.7.26)和(5.7.27)的可观性指数，$\boldsymbol{\Phi}$，$\boldsymbol{\Gamma}$，$\boldsymbol{H}$，$\boldsymbol{\delta}(t)$(5.7.69)定义，且应用(5.7.71)和(5.7.72)计算 $\boldsymbol{L}_{N-j}^w(q^{-1})$ 和 $\boldsymbol{L}_{N-i}^v(q^{-1})$. 滤波误差方差阵 $\boldsymbol{P}_s(N)$ 由定理5.7.2和定理5.7.3计算，但在有关公式(5.7.47)～(5.7.51)中应置 $\boldsymbol{B}_0 = \boldsymbol{I}_m$，$\boldsymbol{S} = \boldsymbol{0}$.

【推论 5.7.2】 单通道系统(5.7.61)～(5.7.64)有渐近稳定的 Wiener 滤波器

$$D(q^{-1})\hat{s}(t \mid t+N) = K_N^s(q^{-1})y(t+N) \tag{5.7.76}$$

它的 ARMA 新息模型为

$$\Lambda(q^{-1})y(t) = D(q^{-1})\varepsilon(t), \quad \Lambda(q^{-1}) = A(q^{-1})P(q^{-1}) \tag{5.7.77}$$

且有关系

$$D(q^{-1})\varepsilon(t) = P(q^{-1})C(q^{-1})w(t) + A(q^{-1})R(q^{-1})v(t) + A(q^{-1})P(q^{-1})\xi(t) \tag{5.7.78}$$

而多项式 $K_N^s(q^{-1})$ 为

$$K_N^s(q^{-1}) = [1 \quad 0 \cdots 0]\boldsymbol{K}_N(q^{-1}) + c_0L_N^w(q^{-1})\Lambda(q^{-1}) \tag{5.7.79}$$

其中多项式矩阵 $\boldsymbol{K}_N(q^{-1})$ 用(5.7.34)计算，但应置 $\widetilde{\boldsymbol{\Lambda}}(q^{-1}) = \Lambda(q^{-1})$. 滤波误差方差 $P_s(N)$ 由定理 5.7.5 计算.

【推论 5.7.3】 对于特殊的多通道反卷积问题

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{5.7.80}$$

$$\boldsymbol{P}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{B}(q^{-1})\boldsymbol{s}(t) + \boldsymbol{R}(q^{-1})\boldsymbol{v}(t) \tag{5.7.81}$$

其中观测 $\boldsymbol{y}(t) \in R^m$，输入信号 $\boldsymbol{s}(t) \in R^n$，$\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}(t) \in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$ 的独立白噪声，$\boldsymbol{A}(q^{-1}),\cdots,\boldsymbol{R}(q^{-1})$ 为形如 $\boldsymbol{X}(q^{-1}) = \boldsymbol{X}_0 + \boldsymbol{X}_1q^{-1} + \cdots + \boldsymbol{X}_{n_x}q^{-n_x}$ 的多项式矩阵，$\boldsymbol{A}_0 = \boldsymbol{I}_n$，$\boldsymbol{P}_0 = \boldsymbol{I}_m$，$n_a \geqslant n_c$，$n_p \geqslant n_b$，$n_p \geqslant n_r$. 设由伪交换(5.7.7)决定的 $\widetilde{\boldsymbol{A}}(q^{-1})$，$\widetilde{\boldsymbol{B}}(q^{-1})$ 使 $(\widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{P}(q^{-1}),\widetilde{\boldsymbol{B}}(q^{-1})\boldsymbol{C}(q^{-1}),\widetilde{\boldsymbol{A}}(q^{-1})\boldsymbol{R}(q^{-1}))$ 左素，则有渐近稳定的Wiener反卷积滤波器(5.7.38)～(5.7.40)，其中在有关公式计算中应置 $\boldsymbol{\xi}(t) = \boldsymbol{0}$，$\boldsymbol{S} = \boldsymbol{0}$，$\boldsymbol{L}_w^\xi(q^{-1}) = \boldsymbol{0}$，$\boldsymbol{Q}_\xi = \boldsymbol{0}$，且滤波误差方差阵由定理 5.7.2 和定理 5.7.3 计算.

【推论 5.7.4】 对单通道反卷积问题(5.7.80)和(5.7.81)有渐近稳定的Wiener反卷积滤波器(5.7.56)～(5.7.60)且估计误差方差 $P_s(N)$ 由定理 5.7.2 和定理 5.7.3 计算，但在有关公式中应置 $\boldsymbol{\xi}(t) = \boldsymbol{0}$，$\boldsymbol{Q}_\xi = \boldsymbol{0}$，$\boldsymbol{L}_N^\xi(q^{-1}) = \boldsymbol{0}$，$\boldsymbol{S} = \boldsymbol{0}$.

5.8 广义系统 Wiener 状态估值器

广义系统大量出现在电网络、机器人、经济系统、航空航天、航海等领域，因而近年来尤为人们所关注. 用经典 Kalman 滤波方法[98,99]解决广义系统状态估计问题报道很少，且所得结果有较大局限性，不能统一处理滤波、预报和平滑问题. 为此，作者用现代时间序列分析方法[1]已提出了广义系统稳态 Kalman 估值器的多种新方法和新结果[18,22,26,27,30,44,48,59,69]. 显示了现代时间序列分析方法作为最优滤波的一种新的方法论的优越性. 它能解决用经典 Kalman 滤波方法不能解决的难题. 另一方面用现代 Wiener 滤

波方法——多项式方法解决 Wiener 状态滤波问题报道也甚少，文献[103]用多项式方法仅给出了常规系统的 Wiener 状态滤波器设计方法，它要求解 Diophantine 方程，且不能处理状态平滑和预报问题.用多项式方法解决广义系统状态估计尚未见报道.本节介绍文献[18]提出的广义系统 Wiener 状态估值器.方法原理是：基于非递推稳态最优状态估值器的递推变形引出 Wiener 状态估值器，其中应用了白噪声估值器和观测预报器.考虑到信息融合理论的需要，将给出状态估计误差方差阵的计算公式.

考虑广义离散定常线性随机系统

$$\boldsymbol{M}\boldsymbol{x}(t+1)=\boldsymbol{\Phi}\boldsymbol{x}(t)+\boldsymbol{\Gamma}\boldsymbol{w}(t) \tag{5.8.1}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{5.8.2}$$

其中 t 为离散时间，$\boldsymbol{x}(t)\in R^n$ 状态，$\boldsymbol{y}(t)\in R^m$ 为观测，$\boldsymbol{M},\boldsymbol{\Phi},\boldsymbol{\Gamma},\boldsymbol{H}$ 为已知适当维数的常阵.

【假设 1】 $\boldsymbol{M}$ 为奇异方阵（$\det \boldsymbol{M}=0$）.

【假设 2】 系统是正则的，即 $\det(\boldsymbol{M}-q^{-1}\boldsymbol{\Phi})\not\equiv 0$.

【假设 3】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}(t)\in R^m$ 是零均值相关白噪声.

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix}\left[\boldsymbol{w}^{\mathrm{T}}(j),\boldsymbol{v}^{\mathrm{T}}(j)\right]\right\}=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}\delta_{tj},\quad \begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}>0 \tag{5.8.3}$$

其中 E 为均值号，T 为转置号，$\delta_{tt}=1,\delta_{tj}=0(t\neq j)$.

【假设 4】 系统是完全可观的，即对任意复数 z 有

$$\operatorname{rank}\begin{bmatrix}z\boldsymbol{M}-\boldsymbol{\Phi}\\ \boldsymbol{H}\end{bmatrix}=n,\quad \operatorname{rank}\begin{bmatrix}\boldsymbol{M}\\ \boldsymbol{H}\end{bmatrix}=n \tag{5.8.4}$$

【假设 5】 初始观测时刻 $t_0=-\infty$.

问题是基于观测$(\boldsymbol{y}(t+N),\boldsymbol{y}(t+N-1),\cdots)$求状态 $\boldsymbol{x}(t)$ 的稳态最优 Wiener 状态估值器 $\hat{\boldsymbol{x}}(t\mid t+N)$.对 $N=0,N>0$ 或 $N<0$，各称其为滤波器、平滑器或预报器.

假设 4 等价于如下可观阵 $\boldsymbol{\Omega}$ 为列满秩[99]：

$$\boldsymbol{\Omega}=\begin{bmatrix}-\boldsymbol{\Phi} & \boldsymbol{M} & \cdots & \boldsymbol{0} & \boldsymbol{0}\\ \vdots & \vdots & & \vdots & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & & -\boldsymbol{\Phi} & \boldsymbol{M}\\ \boldsymbol{H} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0}\\ \vdots & \vdots & & \vdots & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{H}\end{bmatrix}\begin{array}{l}\left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\}(\beta-1)n\ \text{rows}\\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\}\beta m\ \text{rows}\end{array} \tag{5.8.5}$$

其中 $\boldsymbol{\Omega}$ 为$[(\beta-1)n+\beta m]\times\beta n$ 矩阵，β 是使 $\boldsymbol{\Omega}$ 为列满秩的最小自然数，叫可观测性指数.

【注 5.8.1】 若 $\boldsymbol{M}=\boldsymbol{I}_n$，则(5.8.1)和(5.8.2)化为常规系统.若在(5.8.4)中取 $\boldsymbol{M}=\boldsymbol{I}_n$，则可得常规系统的可观性 PBH 判据

$$\operatorname{rank}\begin{bmatrix}z\boldsymbol{I}_n-\boldsymbol{\Phi}\\ \boldsymbol{H}\end{bmatrix}=n \tag{5.8.6}$$

【注 5.8.2】 可观性指数的引入具有减小计算负担的作用.

【引理 5.8.1】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 2 下存在左素分解

$$\boldsymbol{H}(\boldsymbol{M}-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})q^{\tau} \tag{5.8.7}$$

其中 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{B}(q^{-1})$ 是单位滞后算子 q^{-1} 的多项式矩阵，形如 $\boldsymbol{X}(q^{-1}) = \boldsymbol{X}_0 + \boldsymbol{X}_1 q^{-1} + \cdots + \boldsymbol{X}_{n_x} q^{-n_x}$，$q$ 为单位前进算子，$q\boldsymbol{w}(t) = \boldsymbol{w}(t+1)$，$\boldsymbol{A}_0 = \boldsymbol{I}_m$，$\boldsymbol{B}_0 \neq 0$，$\tau$ 为整数，$\tau = 0$，$\tau > 0$ 或 $\tau < 0$.

证明　由假设 2 引出存在非异矩阵 $\boldsymbol{R}$ 和 $\boldsymbol{P}$ 将原系统化为如下典范型

$$\bar{\boldsymbol{M}}\bar{\boldsymbol{x}}(t+1) = \bar{\boldsymbol{\Phi}}\bar{\boldsymbol{x}}(t) + \bar{\boldsymbol{\Gamma}}\boldsymbol{w}(t) \tag{5.8.8}$$

$$\boldsymbol{y}(t) = \bar{\boldsymbol{H}}\bar{\boldsymbol{x}}(t) + \boldsymbol{v}(t) \tag{5.8.9}$$

其中

$$\bar{\boldsymbol{x}}(t) = \boldsymbol{P}^{-1}\boldsymbol{x}(t),\quad \bar{\boldsymbol{H}} = \boldsymbol{HP},\quad \bar{\boldsymbol{\Gamma}} = \boldsymbol{R\Gamma},$$

$$\bar{\boldsymbol{M}} = \boldsymbol{RMP} = \begin{bmatrix} \boldsymbol{I}_{n_1} & \boldsymbol{0} \\ \boldsymbol{0} & \bar{\boldsymbol{N}} \end{bmatrix},\quad \bar{\boldsymbol{\Phi}} = \boldsymbol{R\Phi P} = \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{n_2} \end{bmatrix} \tag{5.8.10}$$

其中 $n = n_1 + n_2$，$\bar{\boldsymbol{N}}$ 是 $n_2 \times n_2$ 幂零阵，即存在自然数 k 使 $\bar{\boldsymbol{N}}^k = \boldsymbol{0}$，这引出 $\bar{\boldsymbol{N}}$ 的所有特征值为零[101]，即

$$\det(q^{-1}\boldsymbol{I}_{n_2} - \bar{\boldsymbol{N}}) = q^{-n_2} \tag{5.8.11}$$

这引出

$$\det(\bar{\boldsymbol{N}} - q^{-1}\boldsymbol{I}_{n_2}) = (-1)^{n_2} q^{-n_2} \tag{5.8.12}$$

容易验证

$$\boldsymbol{H}(\boldsymbol{M} - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1} = \bar{\boldsymbol{H}}(\bar{\boldsymbol{M}} - q^{-1}\bar{\boldsymbol{\Phi}})^{-1}\bar{\boldsymbol{\Gamma}}q^{-1} \tag{5.8.13}$$

注意

$$(\bar{\boldsymbol{M}} - q^{-1}\bar{\boldsymbol{\Phi}})^{-1} = \begin{bmatrix} (\boldsymbol{I}_{n_1} - q^{-1}\boldsymbol{\Phi}_1)^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & (\bar{\boldsymbol{N}} - q^{-1}\boldsymbol{I}_{n_2})^{-1} \end{bmatrix} \tag{5.8.14}$$

应用矩阵求逆公式有

$$(\bar{\boldsymbol{M}} - q^{-1}\bar{\boldsymbol{\Phi}})^{-1} = \frac{\bar{\boldsymbol{F}}_0 + \bar{\boldsymbol{F}}_1 q^{-1} + \cdots + \bar{\boldsymbol{F}}_{n_4} q^{-n_4}}{1 + \varphi_1 q^{-1} + \cdots + \varphi_{n_3} q^{-n_3}} q^{n_2} \tag{5.8.15}$$

其中 $\varphi_{n_3} \neq 0$，$n_3 \leqslant n_1$，且 $n_4 \leqslant n-1$. 于是有

$$\boldsymbol{H}(\boldsymbol{M} - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1} = \frac{\boldsymbol{F}_{i_0} + \boldsymbol{F}_{i_0+1} q^{-1} + \cdots + \boldsymbol{F}_{n_4} q^{-(n_4 - i_0)}}{1 + \varphi_1 q^{-1} + \cdots + \varphi_{n_3} q^{-n_3}} q^{\tau} \tag{5.8.16}$$

其中 $\boldsymbol{F}_i = \bar{\boldsymbol{H}}\bar{\boldsymbol{F}}_i\bar{\boldsymbol{\Gamma}}$，$i = 0, 1, \cdots, n_4$，$i_0 = \min(i \mid \boldsymbol{F}_i \neq \boldsymbol{0})$，且

$$\tau = n_2 - i_0 - 1 \tag{5.8.17}$$

(5.8.16) 可写成

$$\boldsymbol{H}(\boldsymbol{M} - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1} = [(1 + \varphi_1 q^{-1} + \cdots + \varphi_{n_3} q^{-n_3})\boldsymbol{I}_m]^{-1} \times \\ [\boldsymbol{F}_{i_0} + \boldsymbol{F}_{i_1} q^{-1} + \cdots + \boldsymbol{F}_{n_4} q^{-(n_4 - i_0)}] q^{\tau} \tag{5.8.18}$$

为了实现左素分解，可求得最大左因式为 $\boldsymbol{C}(q^{-1})$，即

$$(1 + \varphi_1 q^{-1} + \cdots + \varphi_{n_3} q^{-n_3})\boldsymbol{I}_m = \boldsymbol{C}(q^{-1})\bar{\boldsymbol{A}}(q^{-1}),$$

$$\boldsymbol{F}_{i_0} + \boldsymbol{F}_{i_1} q^{-1} + \cdots + \boldsymbol{F}_{n_4} q^{-(n_4 - i_0)} = \boldsymbol{C}(q^{-1})\bar{\boldsymbol{B}}(q^{-1}) \tag{5.8.19}$$

消去 $\boldsymbol{C}(q^{-1})$ 后有左素分解

$$\boldsymbol{H}(\boldsymbol{M}-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}=\bar{\boldsymbol{A}}^{-1}(q^{-1})\bar{\boldsymbol{B}}(q^{-1})q^{\tau} \tag{5.8.20}$$

由(5.8.19)有

$$\boldsymbol{C}_0\bar{\boldsymbol{A}}_0=\boldsymbol{I}_m,\quad \boldsymbol{C}_0\bar{\boldsymbol{B}}_0=\boldsymbol{F}_{i_0} \tag{5.8.21}$$

注意 $\det\boldsymbol{C}_0\det\bar{\boldsymbol{A}}_0=\det\boldsymbol{I}_m=1$，且 $\boldsymbol{F}_{i_0}\neq\boldsymbol{0}$ 于是 $\bar{\boldsymbol{A}}_0$ 和 $\boldsymbol{C}_0$ 是非异的，且 $\bar{\boldsymbol{B}}_0\neq\boldsymbol{0}$，置

$$\boldsymbol{A}(q^{-1})=\bar{\boldsymbol{A}}_0^{-1}\bar{\boldsymbol{A}}(q^{-1}),\quad \boldsymbol{B}(q^{-1})=\bar{\boldsymbol{A}}_0^{-1}\bar{\boldsymbol{B}}(q^{-1}) \tag{5.8.22}$$

则有左素分解

$$\boldsymbol{H}(\boldsymbol{M}-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})q^{\tau} \tag{5.8.23}$$

带 $\boldsymbol{A}_0=\boldsymbol{I}_m$，$\boldsymbol{B}_0=\bar{\boldsymbol{A}}_0^{-1}\bar{\boldsymbol{B}}_0=\boldsymbol{F}_{i_0}\neq\boldsymbol{0}$. 证毕. □

由(5.8.1)和(5.8.2)有

$$\boldsymbol{y}(t)=\boldsymbol{H}(\boldsymbol{M}-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}\boldsymbol{w}(t)+\boldsymbol{v}(t) \tag{5.8.24}$$

将(5.8.7)代入(5.8.24)有

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{B}(q^{-1})q^{\tau}\boldsymbol{w}(t)+\boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.8.25}$$

这引出 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.8.26}$$

其中 $\boldsymbol{D}(q^{-1})$ 是稳定的，$\boldsymbol{D}_0=\boldsymbol{I}_m$，新息 $\boldsymbol{\varepsilon}(t)$ 是零均值、方差阵为 $\boldsymbol{Q}_{\varepsilon}$ 的白噪声，且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{B}(q^{-1})q^{\tau}\boldsymbol{w}(t)+\boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.8.27}$$

其中 $\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_{\varepsilon}$ 可用 Gevers – Wouters 算法求得.

【引理 5.8.2】 广义系统(5.8.1)，(5.8.2)在假设 1 ~ 5 下，对任意整数 t,j 有关系

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(j)]=\boldsymbol{\Lambda}_{j-t}^{w} \tag{5.8.28}$$

$$\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(j)]=\boldsymbol{\Lambda}_{j-t}^{v} \tag{5.8.29}$$

$$\boldsymbol{\Lambda}_i^{w}=\boldsymbol{Q}_w\boldsymbol{F}_{i+(\tau\vee 0)}^{w\mathrm{T}}+\boldsymbol{S}\boldsymbol{F}_{i+(\tau\vee 0)}^{v\mathrm{T}} \tag{5.8.30}$$

$$\boldsymbol{\Lambda}_i^{v}=\boldsymbol{Q}_v\boldsymbol{F}_{i+(\tau\vee 0)}^{v\mathrm{T}}+\boldsymbol{S}^{\mathrm{T}}\boldsymbol{F}_{i+(\tau\vee 0)}^{w\mathrm{T}} \tag{5.8.31}$$

其中定义 $(\tau\vee 0)=\max(\tau,0)$，$\boldsymbol{F}_i^{\theta\mathrm{T}}=(\boldsymbol{F}_i^{\theta})^{\mathrm{T}}$，$\theta=w,v$，$\boldsymbol{F}_i^{\theta}$ 可递推计算为

$$\boldsymbol{F}_i^{\theta}=-\boldsymbol{D}_1\boldsymbol{F}_{i-1}^{\theta}-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{F}_{i-n_d}^{\theta}+\boldsymbol{X}_i^{\theta},\quad \theta=w,v \tag{5.8.32}$$

其中规定 $\boldsymbol{F}_i^{\theta}=\boldsymbol{0}\,(i<0)$，且定义多项式矩阵

$$\boldsymbol{X}^{w}(q^{-1})=\boldsymbol{B}(q^{-1})q^{(\tau\wedge 0)},\quad \boldsymbol{X}^{v}(q^{-1})=\boldsymbol{A}(q^{-1})q^{(-\tau\wedge 0)} \tag{5.8.33}$$

其中定义 $(\tau\wedge 0)=\min(\tau,0)$，且定义

$$\boldsymbol{X}^{\theta}(q^{-1})=\boldsymbol{X}_0^{\theta}+\boldsymbol{X}_1^{\theta}q^{-1}+\cdots+\boldsymbol{X}_{n_{x\theta}}^{\theta}q^{-n_{x\theta}},\quad \theta=w,v \tag{5.8.34}$$

并规定 $\boldsymbol{X}_i^{\theta}=\boldsymbol{0}\,(i>n_{x\theta})$.

证明 由(5.8.27)和定义(5.8.33)，类似于(5.1.35)，$\boldsymbol{\varepsilon}(t)$ 有展式

$$\begin{aligned}\boldsymbol{\varepsilon}(t)&=\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})q^{\tau}\boldsymbol{w}(t)+\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{v}(t)=\\&\sum_{k=0}^{\infty}\boldsymbol{F}_k^{w}\boldsymbol{w}(t+t_{\tau}-k)+\sum_{k=0}^{\infty}\boldsymbol{F}_k^{v}\boldsymbol{v}(t+t_{\tau}-k)\end{aligned} \tag{5.8.35}$$

其中 $\boldsymbol{F}_k^{w}$ 和 $\boldsymbol{F}_k^{v}$ 由(5.8.32)计算，且定义

$$t_{\tau}=\max(\tau,0)=(\tau\vee 0) \tag{5.8.36}$$

由定义(5.8.33)，展式(5.8.35)等价于当 $\tau>0$ 时，按下式求 $\boldsymbol{\varepsilon}(t)$ 展式：

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})\boldsymbol{w}(t+\tau)+\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})q^{-\tau}\boldsymbol{v}(t+\tau) \tag{5.8.37}$$

且当 $\tau \leqslant 0$ 时按下式求 $\boldsymbol{\varepsilon}(t)$ 展式:

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{D}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})q^{\tau}\boldsymbol{w}(t) + \boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.8.38}$$

应用(5.8.3)和(5.8.35)可得

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(j)] = \boldsymbol{Q}_w\boldsymbol{F}_{j-t+t_\tau}^{w\mathrm{T}} + \boldsymbol{S}\boldsymbol{F}_{j-t+t_\tau}^{v\mathrm{T}} \tag{5.8.39}$$

$$\mathrm{E}[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(j)] = \boldsymbol{Q}_v\boldsymbol{F}_{j-t+t_\tau}^{v\mathrm{T}} + \boldsymbol{S}^{\mathrm{T}}\boldsymbol{F}_{j-t+t_\tau}^{w\mathrm{T}} \tag{5.8.40}$$

由定义(5.8.36)和上两式得(5.8.28) ~ (5.8.31). 证毕. □

【引理 5.8.3】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有稳态最优白噪声估值器

$$\hat{\boldsymbol{w}}(t \mid t+N) = \sum_{i=-t_\tau}^{N}\boldsymbol{\Lambda}_i^w\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i), \quad N \geqslant -t_\tau \tag{5.8.41}$$

$$\hat{\boldsymbol{v}}(t \mid t+N) = \sum_{i=-t_\tau}^{N}\boldsymbol{\Lambda}_i^v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i), \quad N \geqslant -t_\tau \tag{5.8.42}$$

$$\hat{\boldsymbol{w}}(t \mid t+N) = \boldsymbol{0}, \quad N < -t_\tau \tag{5.8.43}$$

$$\hat{\boldsymbol{v}}(t \mid t+N) = \boldsymbol{0}, \quad N < -t_\tau \tag{5.8.44}$$

它们具有统一的新息滤波器形式

$$\hat{\boldsymbol{\theta}}(t \mid t+N) = \boldsymbol{L}_N^{\theta}(q^{-1})\boldsymbol{\varepsilon}(t+N), \quad \boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v} \tag{5.8.45}$$

其中定义多项式矩阵

$$\boldsymbol{L}_N^{\theta}(q^{-1}) = \sum_{i=-t_\tau}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}q^{i-N}, \quad N \geqslant -t_\tau \tag{5.8.46}$$

$$\boldsymbol{L}_N^{\theta}(q^{-1}) = \boldsymbol{0}, \quad N < -t_\tau, \theta = w, v \tag{5.8.47}$$

证明 应用射影公式(4.1.20)和引理 5.8.2,并注意 $\boldsymbol{F}_i^{\theta} = \boldsymbol{0}(i < 0)$,则有 $\boldsymbol{\Lambda}_i^{\theta} = \boldsymbol{0}(i < -t_\tau)$,于是当 $N \geqslant -t_\tau$ 有

$$\hat{\boldsymbol{\theta}}(t \mid t+N) = \sum_{i=-\infty}^{N}\mathrm{E}[\boldsymbol{\theta}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t+i)]\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i) = \sum_{i=-t_\tau}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t+i) \tag{5.8.48}$$

其中 $\boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v}$,且当 $N < -t_\tau$ 有

$$\hat{\boldsymbol{\theta}}(t \mid t+N) = \boldsymbol{0}, \quad \boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v} \tag{5.8.49}$$

注意 $q^i\boldsymbol{\varepsilon}(t) = \boldsymbol{\varepsilon}(t+i)$,故(5.8.45) ~ (5.8.47)成立. 证毕. □

【引理 5.8.4】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有白噪声 Wiener 滤波器

$$\hat{\boldsymbol{w}}(t \mid t+N) = \boldsymbol{L}_N^{w}(q^{-1})\tilde{\boldsymbol{A}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N),$$

$$\hat{\boldsymbol{v}}(t \mid t+N) = \boldsymbol{L}_N^{v}(q^{-1})\tilde{\boldsymbol{A}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.8.50}$$

带伪交换

$$\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1}) = \tilde{\boldsymbol{A}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{5.8.51}$$

其中 $\tilde{\boldsymbol{D}}_0 = \boldsymbol{I}_m$, $\det\boldsymbol{D}(q^{-1}) = \det\tilde{\boldsymbol{D}}(q^{-1})$.

【引理 5.8.5】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下白噪声估值误差 $\tilde{\boldsymbol{\theta}}(t \mid t+N) = \boldsymbol{\theta}(t) - \hat{\boldsymbol{\theta}}(t \mid t+N)(\boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v})$ 的方差阵 $\boldsymbol{P}_\theta(N) = \mathrm{E}[\tilde{\boldsymbol{\theta}}(t \mid t+N)\tilde{\boldsymbol{\theta}}^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}_\theta(N) = \boldsymbol{Q}_\theta - \sum_{i=-t_\tau}^{N}\boldsymbol{\Lambda}_i^{\theta}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\Lambda}_i^{\theta\mathrm{T}}, \quad N \geqslant -t_\tau$$

$$\boldsymbol{P}_\theta(N) = \boldsymbol{Q}_\theta, \quad N < -t_\tau \tag{5.8.52}$$

【引理 5.8.6】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有观测预报器

$$\hat{\boldsymbol{y}}(t+N \mid t) = \boldsymbol{J}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t) \tag{5.8.53}$$

其中 $\widetilde{\boldsymbol{A}}(q^{-1}),\widetilde{\boldsymbol{D}}(q^{-1})$ 由(5.8.41)定义,且 $\boldsymbol{J}_N(q^{-1})$ 由下式决定:

$$\begin{aligned} &\widetilde{\boldsymbol{D}}(q^{-1}) = \boldsymbol{E}_N(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) + q^{-N}\boldsymbol{J}_N(q^{-1}), \quad N > 0 \\ &\boldsymbol{J}_N(q^{-1}) = \widetilde{\boldsymbol{D}}(q^{-1})q^N, \quad N \leqslant 0 \end{aligned} \tag{5.8.54}$$

其中 $\boldsymbol{E}_N(q^{-1})$ 和 $\boldsymbol{J}_N(q^{-1})$ 的阶次见引理 5.3.2,预报误差为

$$\tilde{\boldsymbol{y}}(t+N \mid t) = \boldsymbol{y}(t+N) - \hat{\boldsymbol{y}}(t+N \mid t) = \boldsymbol{E}_N(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.8.55}$$

预报误差方差阵 $\boldsymbol{P}_y(N) = \mathrm{E}[\tilde{\boldsymbol{y}}(t+N \mid t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+N \mid t)]$ 为

$$\begin{aligned} &\boldsymbol{P}_y(N) = \sum_{i=0}^{N-1}\boldsymbol{E}_i^{(N)}\boldsymbol{Q}_\varepsilon\boldsymbol{E}_i^{(N)\mathrm{T}}, \quad \boldsymbol{E}_0^{(N)} = \boldsymbol{I}_m, \quad N > 0 \\ &\boldsymbol{P}_y(N) = \boldsymbol{0}, \quad N \leqslant 0 \end{aligned} \tag{5.8.56}$$

【定理 5.8.1】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有非递推状态表达式

$$\boldsymbol{x}(t) = \sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\boldsymbol{w}(t+i) + \sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i^{(2)}[\boldsymbol{y}(t+i) - \boldsymbol{v}(t+i)] \tag{5.8.57}$$

且有非递推稳态最优状态估值器

$$\begin{aligned} \hat{\boldsymbol{x}}(t \mid t+N) = &\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t+i \mid t+N) + \sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i^{(2)}[\hat{\boldsymbol{y}}(t+i \mid t+N) - \\ &\hat{\boldsymbol{v}}(t+i \mid t+N)] \end{aligned} \tag{5.8.58}$$

其中可观阵 $\boldsymbol{\Omega}$ 由(5.8.5)定义,由假设 4 存在伪逆[1]

$$\boldsymbol{\Omega}^+ = (\boldsymbol{\Omega}^{\mathrm{T}}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathrm{T}} \tag{5.8.59}$$

$\boldsymbol{\Omega}^+$ 的首 n 行分块表示为

$$[\boldsymbol{I}_n \quad \boldsymbol{0}]\boldsymbol{\Omega}^+ = [\boldsymbol{\Omega}_0^{(1)}\cdots\boldsymbol{\Omega}_{\beta-2}^{(1)}, \boldsymbol{\Omega}_0^{(2)}\cdots\boldsymbol{\Omega}_{\beta-1}^{(2)}] \tag{5.8.60}$$

其中 $\boldsymbol{\Omega}_i^{(1)}$ 为 $n \times n$ 阵,$\boldsymbol{\Omega}_i^{(2)}$ 为 $n \times m$ 阵,$\boldsymbol{0}$ 表示适当维数零阵.

证明 由(5.8.1)和(5.8.2)有关系

$$\boldsymbol{\Omega}\begin{bmatrix} \boldsymbol{x}(t) \\ \boldsymbol{x}(t+1) \\ \vdots \\ \vdots \\ \vdots \\ \boldsymbol{x}(t+\beta-1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Gamma}\boldsymbol{w}(t) \\ \vdots \\ \boldsymbol{\Gamma}\boldsymbol{w}(t+\beta-2) \\ \boldsymbol{y}(t) - \boldsymbol{v}(t) \\ \vdots \\ \boldsymbol{y}(t+\beta-1) - \boldsymbol{v}(t+\beta-1) \end{bmatrix} \tag{5.8.61}$$

上式两边对 $\boldsymbol{\Omega}$ 取伪逆运算引出

$$\boldsymbol{x}(t) = [\boldsymbol{I}_n \quad \boldsymbol{0}]\boldsymbol{\Omega}^+\begin{bmatrix} \boldsymbol{\Gamma}\boldsymbol{w}(t) \\ \vdots \\ \boldsymbol{\Gamma}\boldsymbol{w}(t+\beta-2) \\ \boldsymbol{y}(t) - \boldsymbol{v}(t) \\ \vdots \\ \boldsymbol{y}(t+\beta-1) - \boldsymbol{v}(t+\beta-1) \end{bmatrix} \tag{5.8.62}$$

应用分块表示(5.8.60)得(5.8.57).若 $\boldsymbol{A}(q^{-1})$ 是稳定的,即 $\boldsymbol{y}(t)$ 为平稳时间序列,则取(5.8.57)两边各项在无穷维 Hilbert 空间 $L(\boldsymbol{y}^{\mathrm{T}}(t+N),\boldsymbol{y}^{\mathrm{T}}(t+N-1),\cdots)$ 上的射影得(5.8.58).若 $\boldsymbol{A}(q^{-1})$ 是不稳定的,即 $\boldsymbol{y}(t)$ 为非平稳时间序列,此时应用引理5.3.2给出的非平稳时间序列 Åström 预报器,仍统一定义最优状态估值器为(5.8.58). □

下面求状态估值误差 $\tilde{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t\mid t+N)$ 的方差阵 $\boldsymbol{P}(N)=\mathrm{E}[\tilde{\boldsymbol{x}}(t\mid t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t\mid t+N)]$.

【定理5.8.2】 在定理5.8.1条件下,有状态估值误差表达式

$$\begin{aligned}\tilde{\boldsymbol{x}}(t\mid t+N)=&\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\boldsymbol{w}(t+i)-\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i^{(2)}\boldsymbol{v}(t+i)+\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i^{(2)}\tilde{\boldsymbol{y}}(t+i\mid t+N)-\\&\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t+i\mid t+N)+\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i^{(2)}\hat{\boldsymbol{v}}(t+i\mid t+N)\end{aligned}\tag{5.8.63}$$

进一步用合并同类项方法有表达式

$$\tilde{\boldsymbol{x}}(t\mid t+N)=\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\boldsymbol{w}(t+i)-\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i^{(2)}\boldsymbol{v}(t+i)+\sum_{i=0}^{t_\tau+\beta-1}\boldsymbol{\Omega}_i^{c}\boldsymbol{\varepsilon}(t-t_\tau+i),$$
$$t_\tau=\max(\tau,0),\quad N<0\tag{5.8.64}$$

$$\tilde{\boldsymbol{x}}(t\mid t+N)=\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\boldsymbol{w}(t+i)-\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i^{(2)}\boldsymbol{v}(t+i)+\sum_{i=0}^{n_0+t_\tau}\boldsymbol{\Omega}_i^{d}\boldsymbol{\varepsilon}(t-t_\tau+i),$$
$$n_0=\max(\beta-1,N),\quad N\geqslant 0\tag{5.8.65}$$

其中定义 $\boldsymbol{\Omega}_i^c,\boldsymbol{\Omega}_i^d$ 由合并同类项得到,合并(5.8.64)和(5.8.65)有

$$\tilde{\boldsymbol{x}}(t\mid t+N)=\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\boldsymbol{w}(t+i)-\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i^{(2)}\boldsymbol{v}(t+i)+\sum_{i=0}^{n_0+t_\tau}\boldsymbol{\Omega}_i^{e}\boldsymbol{\varepsilon}(t-t_\tau+i),$$
$$n_0=\max(\beta-1,N)\tag{5.8.66}$$

证明 当 $N<0$ 时,由(5.8.43)有 $\hat{\boldsymbol{w}}(t+i\mid t+N)=\boldsymbol{0}(i>N+t_\tau)$.因此当 $i\leqslant N+t_\tau(N-i\geqslant -t_\tau)$ 时,由(5.8.41)有 $\hat{\boldsymbol{w}}(t+i\mid t+N)=\hat{\boldsymbol{w}}(t+i\mid t+i+N-i)$ 是 $(\boldsymbol{\varepsilon}(t+i-t_\tau),\cdots,\boldsymbol{\varepsilon}(t+N))$ 的线性组合.故当 $N<0$ 时(5.8.63)右边第四项

$$\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t+i\mid t+N)=\sum_{i=0}^{p}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t+i\mid t+N),$$
$$p=\min(\beta-2,N+t_\tau),N<0\tag{5.8.67}$$

是 $(\boldsymbol{\varepsilon}(t-t_\tau),\cdots,\boldsymbol{\varepsilon}(t+N))$ 的线性组合.同理(5.8.63)右边第五项也是 $(\boldsymbol{\varepsilon}(t-t_\tau),\cdots,\boldsymbol{\varepsilon}(t+N))$ 的线性组合.

另一方面,当 $N<0$ 时,由(5.8.55)引出(5.8.63)右边第三项为 $(\boldsymbol{\varepsilon}(t+N+1),\cdots,\boldsymbol{\varepsilon}(t+\beta-1))$ 的线性组合.合并上述情形引出(5.8.64).

当 $N\geqslant 0$ 时,由(5.8.43)有 $\hat{\boldsymbol{w}}(t+i\mid t+N)=\boldsymbol{0}(i>N+t_\tau)$,故(5.8.63)右边第四项为

$$\sum_{i=0}^{\beta-2}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t+i\mid t+N)=\sum_{i=0}^{r}\boldsymbol{\Omega}_i^{(1)}\boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t+i\mid t+N),$$
$$r=\min(\beta-2,N+t_\tau),\quad N\geqslant 0\tag{5.8.68}$$

此时 $i\leqslant N+t_\tau$ 即 $N-i\geqslant -t_\tau$,由(5.8.41)引出上式是 $(\boldsymbol{\varepsilon}(t-t_\tau),\cdots,\boldsymbol{\varepsilon}(t+N))$ 的线性组合.

另一方面，当 $N \geqslant 0$ 时由(5.8.55)引出(5.8.63)右边第三项，当 $\beta - 1 \leqslant N$ 时为零，当 $\beta - 1 > N$ 时为 $(\boldsymbol{\varepsilon}(t+N+1), \cdots, \boldsymbol{\varepsilon}(t+\beta-1))$ 的线性组合. 合并上述情形引出(5.8.65). 因 $N < 0$ 时由(5.8.66)有 $n_0 = \beta - 1$，故(5.8.66)成立. 证毕. □

【推论 5.8.1】 广义系统状态估值误差(5.8.66)可写成统一形式

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_1} [\boldsymbol{\Omega}_i^w \boldsymbol{w}(t+i) + \boldsymbol{\Omega}_i^v \boldsymbol{v}(t+i) + \boldsymbol{\Omega}_i^\varepsilon \boldsymbol{\varepsilon}(t - t_\tau + i)],$$

$$n_1 = n_0 + t_\tau, n_0 = \max(\beta - 1, N), t_\tau = \max(\tau, 0) \tag{5.8.69}$$

其中置 $\boldsymbol{\Omega}_i^w = \boldsymbol{\Omega}_i^{(1)} \boldsymbol{\Gamma}, \boldsymbol{\Omega}_i^w = \mathbf{0}(i > \beta - 2); \boldsymbol{\Omega}_i^v = -\boldsymbol{\Omega}_i^{(2)}, \boldsymbol{\Omega}_i^v = \mathbf{0}(i > \beta - 1); \boldsymbol{\Omega}_i^\varepsilon = \boldsymbol{\Omega}_i^e$.

【定理 5.8.3】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下，稳态最优状态估值器(5.8.58)的误差方差阵 $\boldsymbol{P}(N)$ 为

$$\boldsymbol{P}(N) = \sum_{i=0}^{n_1} \sum_{j=0}^{n_1} [\boldsymbol{\Omega}_i^w, \boldsymbol{\Omega}_i^v, \boldsymbol{\Omega}_i^\varepsilon] \begin{bmatrix} \boldsymbol{Q}_w \delta_{ij} & \boldsymbol{S}\delta_{ij} & \boldsymbol{\Lambda}_{j-i-t_\tau}^w \\ \boldsymbol{S}^{\mathrm{T}} \delta_{ij} & \boldsymbol{Q}_v \delta_{ij} & \boldsymbol{\Lambda}_{j-i-t_\tau}^v \\ \boldsymbol{\Lambda}_{i-j-t_\tau}^{w\mathrm{T}} & \boldsymbol{\Lambda}_{i-j-t_\tau}^{v\mathrm{T}} & \boldsymbol{Q}_\varepsilon \delta_{ij} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega}_j^{w\mathrm{T}} \\ \boldsymbol{\Omega}_j^{v\mathrm{T}} \\ \boldsymbol{\Omega}_j^{\varepsilon\mathrm{T}} \end{bmatrix} \tag{5.8.70}$$

证明 由(5.8.69)有

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_1} [\boldsymbol{\Omega}_i^w, \boldsymbol{\Omega}_i^v, \boldsymbol{\Omega}_i^\varepsilon] \begin{bmatrix} \boldsymbol{w}(t+i) \\ \boldsymbol{v}(t+i) \\ \boldsymbol{\varepsilon}(t - t_\tau + i) \end{bmatrix} \tag{5.8.71}$$

应用(5.8.3)、(5.8.28)和(5.8.29)得(5.8.70). □

【定理 5.8.4】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下，有渐近稳定的 Wiener 状态估值器

$$\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1}) \tilde{\boldsymbol{D}}^{-1}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.8.72}$$

其中多项式矩阵 $\boldsymbol{K}_N(q^{-1})$ 定义为

$$\boldsymbol{K}_N(q^{-1}) = \sum_{i=0}^{\beta-2} \boldsymbol{\Omega}_i^{(1)} \boldsymbol{\Gamma} \boldsymbol{L}_{N-i}^w (q^{-1}) \tilde{\boldsymbol{A}}(q^{-1}) + \sum_{i=0}^{\beta-1} \boldsymbol{\Omega}_i^{(2)} [\boldsymbol{J}_{i-N}(q^{-1}) - \boldsymbol{L}_{N-i}^v (q^{-1}) \tilde{\boldsymbol{A}}(q^{-1})] \tag{5.8.73}$$

$\hat{\boldsymbol{x}}(t \mid t+N)$ 也可表为 ARMA 递推形式

$$\det \tilde{\boldsymbol{D}}(q^{-1}) \hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1}) \operatorname{adj} \tilde{\boldsymbol{D}}(q^{-1}) \boldsymbol{y}(t+N) \tag{5.8.74}$$

证明 将(5.8.50)和(5.8.53)代入(5.8.58)中引出(5.8.72) ~ (5.8.74). 因 $\boldsymbol{D}(q^{-1})$ 是稳定的，由关系式 $\det \tilde{\boldsymbol{D}}(q^{-1}) = \det \boldsymbol{D}(q^{-1})$ 引出 $\det \tilde{\boldsymbol{D}}(q^{-1})$ 是稳定的多项式，故(5.8.72)和(5.8.74)是渐近稳定的. □

【推论 5.8.2】 单输出广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有渐近稳定的 Wiener 状态估值器

$$D(q^{-1}) \hat{x}(t \mid t+N) = K_N(q^{-1}) y(t+N) \tag{5.8.75}$$

其中多项式 $K_N(q^{-1})$ 用(5.8.73)计算，但应置 $\tilde{\boldsymbol{A}}(q^{-1}) = A(q^{-1})$.

注意，完全可观性假设 4 等价于[99]如下 $[(\beta-1)n + \beta m] \times \beta n$ 矩阵 $\boldsymbol{\Theta}$ 为列满秩矩阵：

$$\boldsymbol{\Theta}=\begin{bmatrix}\boldsymbol{M} & -\boldsymbol{\Phi} & \cdots & \boldsymbol{0} & \boldsymbol{0}\\ \vdots & \vdots & & \vdots & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{M} & -\boldsymbol{\Phi}\\ \boldsymbol{H} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0}\\ \vdots & \vdots & & \vdots & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{H}\end{bmatrix}\tag{5.8.76}$$

其中 $\boldsymbol{\Theta}$ 为 $[(\beta-1)n+\beta m]\times\beta n$ 矩阵，β 是使 $\boldsymbol{\Theta}$ 为列满秩的最小自然数，叫可观测性指数.

类似于定理 5.8.1 有如下定理.

【定理 5.8.5】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有非递推状态表达式

$$\boldsymbol{x}(t)=\sum_{i=1}^{\beta-1}\boldsymbol{\Theta}_i^{(1)}\boldsymbol{\Gamma w}(t-i)+\sum_{i=0}^{\beta-1}\boldsymbol{\Theta}_i^{(2)}[\boldsymbol{y}(t-i)-\boldsymbol{v}(t-i)]\tag{5.8.77}$$

且有非递推稳态最优状态估值器

$$\begin{aligned}\hat{\boldsymbol{x}}(t\mid t+N)=&\sum_{i=1}^{\beta-1}\boldsymbol{\Theta}_i^{(1)}\boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t-i\mid t+N)+\sum_{i=0}^{\beta-1}\boldsymbol{\Theta}_i^{(2)}[\hat{\boldsymbol{y}}(t-i\mid t+N)-\\&\hat{\boldsymbol{v}}(t-i\mid t+N)]\end{aligned}\tag{5.8.78}$$

其中可观阵 $\boldsymbol{\Theta}$ 由(5.8.76)定义，且由假设 4 存在伪逆

$$\boldsymbol{\Theta}^{+}=(\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Theta})^{-1}\boldsymbol{\Theta}^{\mathrm{T}}\tag{5.8.79}$$

且将 $\boldsymbol{\Theta}^{+}$ 的首 n 行分块表示为

$$[\boldsymbol{I}_n\quad\boldsymbol{0}]\boldsymbol{\Theta}^{+}=[\boldsymbol{\Theta}_1^{(1)}\cdots\boldsymbol{\Theta}_{\beta-1}^{(1)},\boldsymbol{\Theta}_0^{(2)}\cdots\boldsymbol{\Theta}_{\beta-1}^{(2)}]\tag{5.8.80}$$

其中 $\boldsymbol{\Theta}_i^{(1)}$ 为 $n\times n$ 阵，$\boldsymbol{\Theta}_i^{(2)}$ 为 $n\times m$ 阵.

证明 由(5.8.1)和(5.8.2)有关系

$$\boldsymbol{\Theta}\begin{bmatrix}\boldsymbol{x}(t)\\ \boldsymbol{x}(t-1)\\ \vdots\\ \vdots\\ \vdots\\ \boldsymbol{x}(t-\beta+1)\end{bmatrix}=\begin{bmatrix}\boldsymbol{\Gamma w}(t-1)\\ \vdots\\ \boldsymbol{\Gamma w}(t-\beta+1)\\ \boldsymbol{y}(t)-\boldsymbol{v}(t)\\ \vdots\\ \boldsymbol{y}(t-\beta+1)-\boldsymbol{v}(t-\beta+1)\end{bmatrix}\tag{5.8.81}$$

其中 $\boldsymbol{\Theta}$ 由(5.8.76)定义，因 $\boldsymbol{\Theta}$ 为列满秩，对上式实行 $\boldsymbol{\Theta}$ 的伪逆运算得(5.8.77). 若 $\boldsymbol{y}(t)$ 是平稳时间序列，则对上式实行射影运算得(5.8.78). 若 $\boldsymbol{y}(t)$ 是非平稳时间序列，由非平稳时间序列的 Åström 预报器，仍定义 $\hat{\boldsymbol{x}}(t\mid t+N)$ 为(5.8.78). 证毕. □

【定理 5.8.6】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有渐近稳定的 Wiener 状态估值器

$$\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N)\tag{5.8.82}$$

其中定义多项式矩阵 $\boldsymbol{K}_N(q^{-1})$ 为

$$\boldsymbol{K}_N(q^{-1})=\sum_{i=1}^{\beta-1}\boldsymbol{\Theta}_i^{(1)}\boldsymbol{\Gamma L}_{N-i}^{w}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})+\sum_{i=0}^{\beta-1}\boldsymbol{\Theta}_i^{(2)}[\boldsymbol{J}_{-i-N}(q^{-1})-\boldsymbol{L}_{N+i}^{v}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})]\tag{5.8.83}$$

$\hat{\boldsymbol{x}}(t\mid t+N)$ 也可表为 ARMA 递推形式

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\operatorname{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t+N)\tag{5.8.84}$$

证明　类似于定理 5.8.4 的证明，从略. □

【定理 5.8.7】　广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有估值误差 $\tilde{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t \mid t+N)$ 的表达式

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=1}^{\beta-1} \boldsymbol{\Theta}_i^{(1)} \boldsymbol{\Gamma w}(t-i) - \sum_{i=0}^{\beta-1} \boldsymbol{\Theta}_i^{(2)} \boldsymbol{v}(t-i) + \sum_{i=0}^{\beta-1+t_\tau+N} \boldsymbol{\Theta}_i^{(3)} \boldsymbol{\varepsilon}(t-\beta+1-t_\tau+i), \quad N \geqslant 0 \tag{5.8.85}$$

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=1}^{\beta-1} \boldsymbol{\Theta}_i^{(1)} \boldsymbol{\Gamma w}(t-i) - \sum_{i=0}^{\beta-1} \boldsymbol{\Theta}_i^{(2)} \boldsymbol{v}(t-i) + \sum_{i=0}^{n_0} \boldsymbol{\Theta}_i^{(4)} \boldsymbol{\varepsilon}(t-i),$$
$$N < 0, \quad n_0 = \max(-N-1, \beta-1+t_r), t_r = \max(\tau, 0) \tag{5.8.86}$$

其中定义 $\boldsymbol{\Theta}_i^{(3)}$ 和 $\boldsymbol{\Theta}_i^{(4)}$ 由合并同类项得到.

证明　当 $N \geqslant 0$ 时有 $\hat{\boldsymbol{y}}(t-i \mid t+N) = \boldsymbol{y}(t-i)(i \geqslant 0)$，于是有

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=1}^{\beta-1} \boldsymbol{\Theta}_i^{(1)} \boldsymbol{\Gamma w}(t-i) - \sum_{i=0}^{\beta-1} \boldsymbol{\Theta}_i^{(2)} \boldsymbol{v}(t-i) - \sum_{i=1}^{\beta-1} \boldsymbol{\Theta}_i^{(1)} \boldsymbol{\Gamma} \hat{\boldsymbol{w}}(t-i \mid t+N) + \sum_{i=0}^{\beta-1} \boldsymbol{\Theta}_i^{(2)} \hat{\boldsymbol{v}}(t-i \mid t+N) \tag{5.8.87}$$

由(5.8.41)和(5.8.42)上式右边第三和第四项是 $\boldsymbol{\varepsilon}(t-\beta+1-t_\tau), \cdots, \boldsymbol{\varepsilon}(t+N)$ 的线性组合，合并同类项后得(5.8.85). 当 $N < 0$ 时有

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=1}^{\beta-1} \boldsymbol{\Theta}_i^{(1)} \boldsymbol{\Gamma w}(t-i) - \sum_{i=0}^{\beta-1} \boldsymbol{\Theta}_i^{(2)} \boldsymbol{v}(t-i) - \sum_{i=1}^{\beta-1} \boldsymbol{\Theta}_i^{(1)} \boldsymbol{\Gamma} \hat{\boldsymbol{w}}(t-i \mid t+N) + \sum_{i=0}^{\beta-1} \boldsymbol{\Theta}_i^{(2)} \tilde{\boldsymbol{y}}(t-i \mid t+N) + \sum_{i=0}^{\beta-1} \boldsymbol{\Theta}_i^{(2)} \hat{\boldsymbol{v}}(t-i \mid t+N) \tag{5.8.88}$$

由(5.8.55)上式右边第四项为 $\boldsymbol{\varepsilon}(t+N-1), \cdots, \boldsymbol{\varepsilon}(t)$ 的线性组合. 另一方面注意白噪声估值器

$$\hat{\boldsymbol{\theta}}(t-i \mid t+N) = \hat{\boldsymbol{\theta}}(t-i \mid t-i+N+i), \quad \boldsymbol{\theta} = \boldsymbol{w}, \boldsymbol{v} \tag{5.8.89}$$

由(5.8.41)和(5.8.42)，取 $i = \beta-1$，对 $\hat{\boldsymbol{w}}(t-\beta+i \mid t+N)$ 和 $\hat{\boldsymbol{v}}(t-\beta+i \mid t+N)$ 而言，当 $N+\beta-1 \geqslant -t_\tau$ 时，即 $\beta-1+t_\tau > -N-1$ 时，它们是 $\boldsymbol{\varepsilon}(t-\beta+1-t_\tau), \cdots, \boldsymbol{\varepsilon}(t+N)$ 的线性组合. 于是引出(5.8.86). 证毕. □

【推论 5.8.3】　广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有估值误差表达式

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_1} [\boldsymbol{\Theta}_i^w \boldsymbol{w}(t-i) + \boldsymbol{\Theta}_i^v \boldsymbol{v}(t-i) + \boldsymbol{\Theta}_i^{(5)} \boldsymbol{\varepsilon}(t-\beta_\tau+i)],$$
$$N \geqslant 0, n_1 = \beta-1+t_\tau+N, \beta_\tau = \beta-1+t_\tau \tag{5.8.90}$$

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_0} [\boldsymbol{\Theta}_i^w \boldsymbol{w}(t-i) + \boldsymbol{\Theta}_i^v \boldsymbol{v}(t-i) + \boldsymbol{\Theta}_i^{(6)} \boldsymbol{\varepsilon}(t-i)],$$
$$N < 0, n_0 = \max(-N-1, \beta-1+t_\tau), t_\tau = \max(\tau, 0) \tag{5.8.91}$$

其中置 $\boldsymbol{\Theta}_i^w = \boldsymbol{\Theta}_i^{(1)}(i = 1, \cdots, \beta-1)$；$\boldsymbol{\Theta}_0^w = \boldsymbol{0}$；$\boldsymbol{\Theta}_i^w = \boldsymbol{0}(i > \beta-1)$；$\boldsymbol{\Theta}_i^v = \boldsymbol{\Theta}_i^{(2)}(i = 0, \cdots, \beta-1)$；$\boldsymbol{\Theta}_i^v = \boldsymbol{0}(i > \beta-1)$；$\boldsymbol{\Theta}_i^{(5)} = \boldsymbol{\Theta}_i^{(3)}$；$\boldsymbol{\Theta}_i^{(6)} = \boldsymbol{\Theta}_i^{(4)}$.

证明　由(5.8.85)和(5.8.86)得证. □

【定理 5.8.8】　广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下，稳态最优状态估值器

(5.8.78) 的误差方差阵 $\boldsymbol{P}(N)=\mathrm{E}[\tilde{\boldsymbol{x}}(t\mid t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}(N)=\sum_{i=0}^{n_1}\sum_{j=0}^{n_1}[\boldsymbol{\Theta}_i^w,\boldsymbol{\Theta}_i^v,\boldsymbol{\Theta}_i^{(5)}]\begin{bmatrix}\boldsymbol{Q}_w\delta_{ij} & \boldsymbol{S}\delta_{ij} & \boldsymbol{\Lambda}_{i+j-\beta_\tau}^w \\ \boldsymbol{S}^{\mathrm{T}}\delta_{ij} & \boldsymbol{Q}_v\delta_{ij} & \boldsymbol{\Lambda}_{i+j-\beta_\tau}^v \\ \boldsymbol{\Lambda}_{i+j-\beta_\tau}^{w\mathrm{T}} & \boldsymbol{\Lambda}_{i+j-\beta_\tau}^{v\mathrm{T}} & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Theta}_j^{w\mathrm{T}} \\ \boldsymbol{\Theta}_j^{v\mathrm{T}} \\ \boldsymbol{\Theta}_j^{\varepsilon\mathrm{T}}\end{bmatrix},\quad N\geqslant 0 \tag{5.8.92}$$

$$\boldsymbol{P}(N)=\sum_{i=0}^{n_0}\sum_{j=0}^{n_0}[\boldsymbol{\Theta}_i^w,\boldsymbol{\Theta}_i^v,\boldsymbol{\Theta}_i^{(6)}]\begin{bmatrix}\boldsymbol{Q}_w\delta_{ij} & \boldsymbol{S}\delta_{ij} & \boldsymbol{\Lambda}_{i-j}^w \\ \boldsymbol{S}^{\mathrm{T}}\delta_{ij} & \boldsymbol{Q}_v\delta_{ij} & \boldsymbol{\Lambda}_{i-j}^v \\ \boldsymbol{\Lambda}_{j-i}^{w\mathrm{T}} & \boldsymbol{\Lambda}_{j-i}^{v\mathrm{T}} & \boldsymbol{Q}_\varepsilon\delta_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Theta}_j^{w\mathrm{T}} \\ \boldsymbol{\Theta}_j^{v\mathrm{T}} \\ \boldsymbol{\Theta}_j^{(6)\mathrm{T}}\end{bmatrix},\quad N<0 \tag{5.8.93}$$

证明 应用 (5.8.28) 和 (5.8.29)，由推论 5.8.3 类似定理 5.8.3 的证明得证. □

【引理 5.8.6】[187] 广义系统 (5.8.1) 和 (5.8.2) 在假设 1 ~ 5 下，且设在 (5.8.10) 中 $n_2>0$，则存在 $n\times m$ 矩阵 $\boldsymbol{K}$ 使

$$\det[z\boldsymbol{M}-(\boldsymbol{\Phi}+\boldsymbol{KH})]=\gamma\neq 0\quad(\text{常数}) \tag{5.8.94}$$

【推论 5.8.4】 广义系统 (5.8.1) 和 (5.8.2) 在假设 1 ~ 5 下，存在 $n\times m$ 矩阵 $\boldsymbol{K}$ 使

$$\det[\boldsymbol{M}-q^{-1}(\boldsymbol{\Phi}+\boldsymbol{KH})]=\gamma q^{-n},\quad \gamma\neq 0\quad(\text{常数}) \tag{5.8.95}$$

证明 由 (5.8.94)，注意关系

$$\det[q\boldsymbol{M}-(\boldsymbol{\Phi}+\boldsymbol{KH})]=\det[q(\boldsymbol{M}-q^{-1}(\boldsymbol{\Phi}+\boldsymbol{KH}))]=$$
$$q^n\det[\boldsymbol{M}-q^{-1}(\boldsymbol{\Phi}+\boldsymbol{KH})]=\gamma\neq 0 \tag{5.8.96}$$

这引出 (5.8.95). 证毕. □

【定理 5.8.9】 广义系统 (5.8.1) 和 (5.8.2) 在假设 1 ~ 5 下有状态表达式

$$\boldsymbol{x}(t)=\boldsymbol{R}(q^{-1})\boldsymbol{w}(t+n-1)-\boldsymbol{P}(q^{-1})\boldsymbol{y}(t+n-1)+\boldsymbol{P}(q^{-1})\boldsymbol{v}(t+n-1) \tag{5.8.97}$$

其中定义

$$\boldsymbol{R}(q^{-1})=\frac{\operatorname{adj}[\boldsymbol{M}-q^{-1}(\boldsymbol{\Phi}+\boldsymbol{KH})]\boldsymbol{\Gamma}}{\gamma} \tag{5.8.98}$$

$$\boldsymbol{P}(q^{-1})=\frac{\operatorname{adj}[\boldsymbol{M}-q^{-1}(\boldsymbol{\Phi}+\boldsymbol{KH})]\boldsymbol{K}}{\gamma} \tag{5.8.99}$$

其中 γ 和 $\boldsymbol{K}$ 由引理 5.8.6 决定，adj 为伴随阵符号.

证明 用 $\boldsymbol{K}$ 左乘 (5.8.2) 并与 (5.8.1) 相加有

$$\boldsymbol{Mx}(t+1)=(\boldsymbol{\Phi}+\boldsymbol{KH})\boldsymbol{x}(t)+\boldsymbol{\Gamma w}(t)-\boldsymbol{Ky}(t)+\boldsymbol{Kv}(t) \tag{5.8.100}$$

这引出

$$\boldsymbol{x}(t)=[\boldsymbol{M}-q^{-1}(\boldsymbol{\Phi}+\boldsymbol{KH})]^{-1}[\boldsymbol{\Gamma w}(t-1)-\boldsymbol{Ky}(t-1)+\boldsymbol{Kv}(t-1)] \tag{5.8.101}$$

利用 (5.8.95) 立刻得 (5.8.97) ~ (5.8.99). 证毕. □

【定理 5.8.10】 广义系统 (5.8.1) 和 (5.8.2) 在假设 1 ~ 5 下有非递推稳态最优状态估值器

$$\hat{\boldsymbol{x}}(t\mid t+N)=\sum_{i=0}^{n-1}\boldsymbol{R}_i\hat{\boldsymbol{w}}(t+n-1-i\mid t+N)-\sum_{i=0}^{n-1}\boldsymbol{P}_i\hat{\boldsymbol{y}}(t+n-1-i\mid t+N)+$$

$$\sum_{i=0}^{n-1} \boldsymbol{P}_i \hat{\boldsymbol{v}}(t+n-1-i \mid t+N) \tag{5.8.102}$$

证明　将(5.8.97)右边展开有

$$\boldsymbol{x}(t) = \sum_{i=0}^{n-1} \boldsymbol{R}_i \boldsymbol{w}(t+n-1-i) - \sum_{i=0}^{n-1} \boldsymbol{P}_i \boldsymbol{y}(t+n-1-i) + \sum_{i=0}^{n-1} \boldsymbol{P}_i \boldsymbol{v}(t+n-1-i) \tag{5.8.103}$$

若 $\boldsymbol{y}(t)$ 为平稳时间序列，则对上式两边取射影运算立刻得(5.8.102)．若 $\boldsymbol{y}(t)$ 是非平稳时间序列，应用引理 5.3.2 给出的非平稳时间序列的 Åström 预报器仍有(5.8.102)成立．□

【定理 5.8.11】　广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有状态估值误差 $\tilde{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t \mid t+N)$ 的表达式

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \sum_{i=0}^{n_1} [\boldsymbol{\Omega}_i^w \boldsymbol{w}(t+i) + \boldsymbol{\Omega}_i^v \boldsymbol{v}(t+i) + \boldsymbol{\Omega}_i^\varepsilon \boldsymbol{\varepsilon}(t-t_\tau+i) \tag{5.8.104}$$

其中定义 $\boldsymbol{\Omega}_i^w = \boldsymbol{R}_{n-1-i}, i = 0,\cdots,n-1, \boldsymbol{\Omega}_i^w = 0(i > n-1)$；$\boldsymbol{\Omega}_i^v = \boldsymbol{P}_{n-1-i}, i = 0,\cdots,n-1, \boldsymbol{\Omega}_i^v = \boldsymbol{0}(i > n-1)$；$\boldsymbol{\Omega}_i^\varepsilon$ 由合并同类项得到，且

$$n_1 = n_0 + t_\tau, n_0 = \max(n-1, N), \quad t_\tau = \max(\tau, 0) \tag{5.8.105}$$

证明　由定理 5.8.2 和推论 5.8.1 得证．□

【定理 5.8.12】　广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有稳态估值误差方差阵 $\boldsymbol{P}(N) = \mathrm{E}[\tilde{\boldsymbol{x}}(t \mid t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t \mid t+N)]$ 为(5.8.70)．

证明　应用(5.8.104)和定理 5.8.3 得证．□

【定理 5.8.13】　广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有渐近稳定的 Wiener 状态估值器

$$\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.8.106}$$

或表为 ARMA 递推形式

$$\det\tilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\mathrm{adj}\tilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t+N) \tag{5.8.107}$$

其中定义多项式矩阵 $\boldsymbol{K}_N(q^{-1})$ 为

$$\begin{aligned}\boldsymbol{K}_N(q^{-1}) = &\sum_{i=0}^{n-1} \boldsymbol{R}_i \boldsymbol{L}_{N-n+1+i}^w(q^{-1})\tilde{\boldsymbol{A}}(q^{-1}) - \sum_{i=0}^{n-1} \boldsymbol{P}_i \boldsymbol{J}_{n-1-i-N}(q^{-1}) + \\ &\sum_{i=0}^{n-1} \boldsymbol{P}_i \boldsymbol{L}_{N-n+1+i}^v(q^{-1})\tilde{\boldsymbol{A}}(q^{-1})\end{aligned} \tag{5.8.108}$$

其中 $N = 0, N > 0$ 或 $N < 0$．

证明　将(5.8.50)和(5.8.53)代入(5.8.102)得(5.8.106) ~ (5.8.108)．由 $\tilde{\boldsymbol{D}}(q^{-1})$ 的稳定性引出(5.8.106)或(5.8.107)是渐近稳定的．证毕．□

【推论 5.8.5】　单输出广义系统(5.8.1)和(5.8.2)在假设 1 ~ 4 下，有渐近稳定的 Wiener 状态估值器

$$D(q^{-1})\hat{x}(t \mid t+N) = K_N(q^{-1})y(t+N) \tag{5.8.109}$$

其中 $K_N(q^{-1})$ 为在(5.8.108)中置 $\tilde{\boldsymbol{A}}(q^{-1}) = A(q^{-1})$．

证明　对单输出系统有 $m = 1$，故有 $\tilde{\boldsymbol{A}}(q^{-1}) = A(q^{-1}), \tilde{\boldsymbol{D}}(q^{-1}) = D(q^{-1})$，由定理 5.8.13 立刻得证．□

【注 5.8.3】　广义系统(5.8.1)和(5.8.2)的 ARMA 新息模型可基于左素分解(5.8.7)

得到，也可用另一种方法得到，即基于表达式(5.8.97)构造 ARMA 新息模型. 事实上将(5.8.97)代入(5.8.2)有

$$\begin{aligned}\boldsymbol{y}(t) = &\boldsymbol{HR}(q^{-1})\boldsymbol{w}(t+n-1) - \boldsymbol{HP}(q^{-1})\boldsymbol{y}(t+n-1) + \\ &\boldsymbol{HP}(q^{-1})\boldsymbol{v}(t+n-1) + \boldsymbol{v}(t)\end{aligned} \tag{5.8.110}$$

即

$$\begin{aligned}\boldsymbol{y}(t) + \boldsymbol{HR}(q^{-1})\boldsymbol{y}(t+n-1) = &\boldsymbol{HP}(q^{-1})\boldsymbol{w}(t+n-1) + \\ &\boldsymbol{HP}(q^{-1})\boldsymbol{v}(t+n-1) + \boldsymbol{v}(t)\end{aligned} \tag{5.8.111}$$

注意 $\boldsymbol{P}(q^{-1})$ 和 $\boldsymbol{R}(q^{-1})$ 的首系数阵 $\boldsymbol{P}_0$ 和 $\boldsymbol{R}_0$ 可能为零阵. 将(5.8.111)左边的多项式矩阵的首系数阵(可证明它是非异方阵)化为单位阵后可得到同样的 ARMA 新息模型(5.8.26)和(5.8.27).

【注 5.8.4】 在引理 5.8.6 中矩阵 $\boldsymbol{K}$ 和常数 γ 是非惟一的，可由等式(5.8.95)用待定系数法求得. 详见仿真例子.

【例 5.8.1】 考虑带相关噪声的广义随机系统

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\boldsymbol{x}(t+1) = \begin{bmatrix} 0.9 & 0 & 0 \\ -0.5 & 0.5 & 0 \\ -0.25 & 0 & 0.25 \end{bmatrix}\boldsymbol{x}(t) + \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} w(t) \tag{5.8.112}$$

$$y(t) = [0 \quad 1 \quad 0]\boldsymbol{x}(t) + v(t) \tag{5.8.113}$$

$$v(t) = 0.5w(t) + \xi(t) \tag{5.8.114}$$

其中 $w(t)$ 和 $\xi(t)$ 是零均值相互独立的高斯白噪声，方差各为 $\sigma_w^2 = 1$ 和 $\sigma_v^2 = 1$. 容易应用(5.8.7)求得 ARMA 新息模型为

$$(1 - 0.9q^{-1})y(t) = (1 - 0.540\,5q^{-1})\varepsilon(t) \tag{5.8.115}$$

其中新息 $\varepsilon(t)$ 是零均值、方差为 $\sigma_\varepsilon^2 = 7.447\,1$ 的白噪声，且有关系

$$(1 - 0.540\,5q^{-1})\varepsilon(t) = (2 - 0.8q^{-1})w(t) + (1 - 0.9q^{-1})v(t) \tag{5.8.116}$$

记 $\boldsymbol{K} = [k_1, k_2, k_3]^{\mathrm{T}}$，容易求得

$$\det[\boldsymbol{M} - q^{-1}(\boldsymbol{\Phi} + \boldsymbol{KH})] = k_3q^{-1} + (0.125 + 0.25k_2 - 0.9k_3)q^{-2} - 0.125k_1q^{-3} \tag{5.8.117}$$

为了得到满足(5.8.95)的 $\boldsymbol{K}$，在上式中应取

$$k_3 = 0, \quad k_2 = -0.5 \tag{5.8.118}$$

而 k_1 可任意取，例如可取 $k_1 = 1$，则有 $\gamma = -0.125$，于是有

$$\boldsymbol{K} = [1 \quad -0.5 \quad 0]^{\mathrm{T}}, \quad \gamma = -0.125 \tag{5.8.119}$$

容易算得

$$\boldsymbol{R}(q^{-1}) = \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix} q^{-1} + \begin{bmatrix} -2 \\ 0.8 \\ -2 \end{bmatrix} q^{-2} \tag{5.8.120}$$

$$\boldsymbol{P}(q^{-1}) = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} q^{-1} + \begin{bmatrix} -1 \\ -0.1 \\ -1 \end{bmatrix} q^{-2} \tag{5.8.121}$$

注意 ARMA 新息模型(5.8.115)也可用注 5.8.3 的方法求得. 事实上，将(5.8.120)和

(5.8.121) 及 $\boldsymbol{H}=[0\quad 1\quad 0]$ 代入(5.8.111)有

$$-y(t+1)+0.9y(t)=(-2+0.8q^{-1})w(t+1)-v(t+1)+0.9v(t) \tag{5.8.122}$$

上式两边乘 -1 化为首系数为 1 的自回归多项式有

$$(1-0.9q^{-1})y(t)=(2-0.8q^{-1})w(t)+(1-0.9q^{-1})v(t) \tag{5.8.123}$$

它相同于(5.8.116),且由上式引出 ARMA 新息模型(5.8.115).

应用推论 5.8.5,取 $N=0$,可得广义 Wiener 状态滤波器

$$(1-0.540\,5q^{-1})\hat{x}(t\mid t)=K_0(q^{-1})y(t) \tag{5.8.124}$$

$$\boldsymbol{K}_0(q^{-1})=\begin{bmatrix}0.026\,5\\0.697\,9\\0.026\,5\end{bmatrix}+\begin{bmatrix}0.335\,7\\-0.268\,6\\0.335\,7\end{bmatrix}q^{-1} \tag{5.8.125}$$

仿真结果如图 5.8.1 至图 5.8.3 所示,其中实线为 $x_i(t)$,$i=1,2,3$,且 $\boldsymbol{x}(t)=[x_1(t),x_2(t),x_3(t)]^{\mathrm{T}}$,虚线为估值 $\hat{x}_i(t\mid t)$,$i=1,2,3$,且 $\hat{\boldsymbol{x}}(t\mid t)=[\hat{x}_1(t\mid t),\hat{x}_2(t\mid t),\hat{x}_3(t\mid t)]^{\mathrm{T}}$.

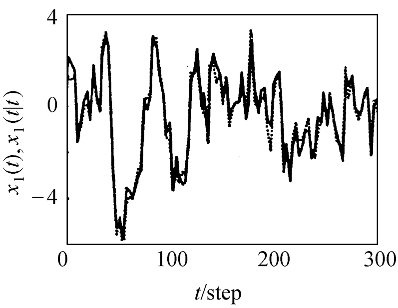


图 5.8.1　状态 $x_1(t)$ 和广义 Wiener 状态滤波器 $\hat{x}_1(t\mid t)$

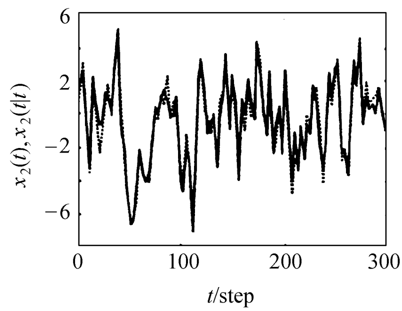


图 5.8.1　状态 $x_2(t)$ 和广义 Wiener 状态滤波器 $\hat{x}_2(t\mid t)$

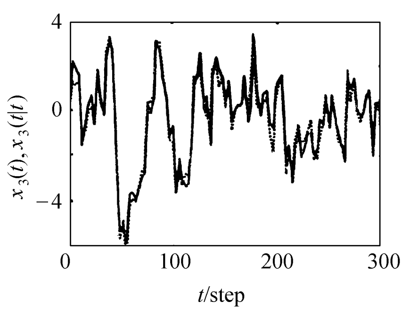


图 5.8.3　状态 $x_3(t)$ 和广义 Wiener 状态滤波器 $\hat{x}_3(t\mid t)$

【注 5.8.5】　考虑广义系统(5.8.112),带有观测方程

$$\boldsymbol{y}(t)=\begin{bmatrix}0&1&0\\1&0&0\end{bmatrix}\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{5.8.126}$$

其中 $\boldsymbol{v}(t)$ 是零均值、方差阵为 $\boldsymbol{Q}_v$ 的独立于 $w(t)$ 的白噪声,观测 $\boldsymbol{y}(t)$ 和观测噪声 $\boldsymbol{v}(t)$ 均为二维随机过程.为了求推论 5.8.4 中的 3×2 矩阵 $\boldsymbol{K}$,可取 K 有如下形式

$$\boldsymbol{K}=\begin{bmatrix}k_1&0\\k_2&0\\k_3&0\end{bmatrix} \tag{5.8.127}$$

其中 k_1,k_2,k_3 为待定系数,于是有

$$
\boldsymbol{KH} = \begin{bmatrix} k_1 & 0 \\ k_2 & 0 \\ k_3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & k_1 & 0 \\ 0 & k_2 & 0 \\ 0 & k_3 & 0 \end{bmatrix} \tag{5.8.128}
$$

由例 5.8.1 可求得

$$
k_1 = 1, \quad k_2 = -0.5, \quad k_3 = 0 \tag{5.8.129}
$$

于是可得满足(5.8.95)的 $\boldsymbol{K}$ 和 γ 为

$$
\boldsymbol{K} = \begin{bmatrix} 1 & 0 \\ -0.5 & 0 \\ 0 & 0 \end{bmatrix}, \quad \gamma = -0.125 \tag{5.8.130}
$$

5.9 广义系统降阶 Wiener 状态估值器

考虑广义系统(5.8.1)和(5.8.2),在节 5.8 的假设 1 ~ 5 下,我们考虑广义系统的降阶状态估计问题,即在一定变换下将广义系统分解为两个降阶的子系统,将广义系统状态估计问题转化为子系统的状态估计问题,从而可减小计算负担,且为广义系统状态估计提供了新的途径和方法.

由假设 1 ~ 2,存在非异矩阵 $\boldsymbol{P}$,$\boldsymbol{Q}$ 使

$$
\boldsymbol{PMQ} = \begin{bmatrix} \boldsymbol{I}_{n_1} & \boldsymbol{0} \\ \boldsymbol{0} & \bar{\boldsymbol{N}} \end{bmatrix}, \quad \boldsymbol{P\Phi Q} = \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{n_2} \end{bmatrix}, \quad \boldsymbol{P\Gamma} = \begin{bmatrix} \boldsymbol{\Gamma}_1 \\ \boldsymbol{\Gamma}_2 \end{bmatrix},
$$

$$
\boldsymbol{HQ} = [\boldsymbol{H}_1, \boldsymbol{H}_2] \tag{5.9.1}
$$

其中 $n_1 + n_2 = n$,$\bar{\boldsymbol{N}}$ 为幂零阵. 令

$$
\boldsymbol{x}(t) = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{x}_1(t) \\ \boldsymbol{x}_2(t) \end{bmatrix} \tag{5.9.2}
$$

则原广义系统(5.8.1)和(5.8.2)化为典范形

$$
\begin{bmatrix} \boldsymbol{I}_{n_1} & \boldsymbol{0} \\ \boldsymbol{0} & \bar{\boldsymbol{N}} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1(t+1) \\ \boldsymbol{x}_2(t+1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{n_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1(t) \\ \boldsymbol{x}_2(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Gamma}_1 \\ \boldsymbol{\Gamma}_2 \end{bmatrix} \boldsymbol{w}(t) \tag{5.9.3}
$$

其中 $\boldsymbol{x}_1(t) \in R^{n_1}$ 和 $\boldsymbol{x}_2(t) \in R^{n_2}$ 为子系统的状态. 由关系(5.9.2)求广义系统状态估值器 $\hat{\boldsymbol{x}}(t \mid t+N)$ 转化为求子系统状态估值器 $\hat{\boldsymbol{x}}_1(t \mid t+N)$ 和 $\hat{\boldsymbol{x}}_2(t \mid t+N)$,且有关系

$$
\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{Q}\begin{bmatrix} \hat{\boldsymbol{x}}_1(t \mid t+N) \\ \hat{\boldsymbol{x}}_2(t \mid t+N) \end{bmatrix} \tag{5.9.4}
$$

(5.9.3)可分解为如下两个子系统

$$
\boldsymbol{x}_1(t+1) = \boldsymbol{\Phi}_1 \boldsymbol{x}_1(t) + \boldsymbol{\Gamma}_1 \boldsymbol{w}(t) \tag{5.9.5}
$$

$$
\boldsymbol{N}\boldsymbol{x}_2(t+1) = \boldsymbol{x}_2(t) + \boldsymbol{\Gamma}_2 \boldsymbol{w}(t) \tag{5.9.6}
$$

$$
\boldsymbol{y}(t) = \boldsymbol{H}_1 \boldsymbol{x}_1(t) + \boldsymbol{H}_2 \boldsymbol{x}_2(t) + \boldsymbol{v}(t) \tag{5.9.7}
$$

因 $\bar{\boldsymbol{N}}$ 为幂零阵,则有它的全部特征值为零,即

$$
\det(q^{-1}\boldsymbol{I}_{n_2} - \bar{\boldsymbol{N}}) = q^{-n_2} \tag{5.9.8}
$$

由(5.9.6)有

$$\boldsymbol{x}_2(t) = -(q^{-1}\boldsymbol{I}_{n_2} - \bar{\boldsymbol{N}})^{-1}\boldsymbol{\Gamma}_2\boldsymbol{w}(t-1) = -\frac{\operatorname{adj}(q^{-1}\boldsymbol{I}_{n_2} - \bar{\boldsymbol{N}})\boldsymbol{\Gamma}_2}{\det(q^{-1}\boldsymbol{I}_{n_2} - \bar{\boldsymbol{N}})}\boldsymbol{w}(t-1) =$$

$$\bar{\boldsymbol{R}}(q^{-1})\boldsymbol{w}(t+n_2-1) \tag{5.9.9}$$

$$\bar{\boldsymbol{R}}(q^{-1}) = -\operatorname{adj}(q^{-1}\boldsymbol{I}_{n_2} - \bar{\boldsymbol{N}})\boldsymbol{\Gamma}_2 \tag{5.9.10}$$

它的阶次 $n_{\bar{r}} = n_2 - 1$. 将(5.9.9)代入(5.9.7)引出子系统(5.9.5)的观测方程

$$\boldsymbol{y}(t) = \boldsymbol{H}_1\boldsymbol{x}_1(t) + \boldsymbol{\eta}(t) \tag{5.9.11}$$

其中 MA 有色观测噪声 $\boldsymbol{\eta}(t)$ 为

$$\boldsymbol{\eta}(t) = \boldsymbol{R}(q^{-1})\boldsymbol{w}(t+n_2-1) + \boldsymbol{v}(t) \tag{5.9.12}$$

$$\boldsymbol{R}(q^{-1}) = \boldsymbol{H}_2\bar{\boldsymbol{R}}(q^{-1}) \tag{5.9.13}$$

其中 $\boldsymbol{R}(q^{-1})$ 的阶次 $n_r = n_2 - 1$.

子系统(6.9.5)和(6.9.11)构成带 MA 有色观测噪声的常规系统.

【引理 5.9.1】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下,$(\boldsymbol{\Phi}_1, \boldsymbol{H}_1)$ 为完全可观对.

证明 对任意复数 z 有

$$\operatorname{rank}\begin{bmatrix} z\boldsymbol{M} - \boldsymbol{\Phi} \\ \boldsymbol{H} \end{bmatrix} = \operatorname{rank}\begin{bmatrix} \boldsymbol{P} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_m \end{bmatrix}\begin{bmatrix} z\boldsymbol{M} - \boldsymbol{\Phi} \\ \boldsymbol{H} \end{bmatrix}\boldsymbol{Q} =$$

$$\operatorname{rank}\begin{bmatrix} \boldsymbol{PMQ} - \boldsymbol{P\Phi Q} \\ \boldsymbol{HP} \end{bmatrix} = \operatorname{rank}\begin{bmatrix} z\boldsymbol{I}_{n_1} - \boldsymbol{\Phi}_1 & \boldsymbol{0} \\ \boldsymbol{0} & z\bar{\boldsymbol{N}} - \boldsymbol{I}_{n_2} \\ \boldsymbol{H}_1 & \boldsymbol{H}_2 \end{bmatrix} = n \tag{5.9.14}$$

其中应用了(5.9.1). 由矩阵秩的 Sylvester[101] 不等式引出

$$\operatorname{rank}\begin{bmatrix} z\boldsymbol{I}_{n_1} - \boldsymbol{\Phi}_1 \\ \boldsymbol{H}_1 \end{bmatrix} = n_1, \quad \operatorname{rank}\begin{bmatrix} z\bar{\boldsymbol{N}} - \boldsymbol{I}_{n_2} \\ \boldsymbol{H}_2 \end{bmatrix} = n_2 \tag{5.9.15}$$

上式第一式即 $(\boldsymbol{\Phi}_1, \boldsymbol{H}_1)$ 为完全可观对. □

应注意,不同于节 5.8 求 ARMA 新息模型的方法,这里可通过子系统(5.9.5)和(5.9.11)来求 ARMA 新息模型,因为处理低阶常规系统更为方便和简单. 由(5.9.5)和(5.9.11)有

$$\boldsymbol{y}(t) = \boldsymbol{H}_1(\boldsymbol{I}_{n_1} - q^{-1}\boldsymbol{\Phi}_1)^{-1}\boldsymbol{\Gamma}_1 q^{-1}\boldsymbol{w}(t) + \boldsymbol{\eta}(t) \tag{5.9.16}$$

引入左素分解

$$\boldsymbol{H}_1(\boldsymbol{I}_{n_1} - q^{-1}\boldsymbol{\Phi}_1)^{-1}\boldsymbol{\Gamma}_1 q^{-1} = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{C}(q^{-1}) \tag{5.9.17}$$

带 $\boldsymbol{A}_0 = \boldsymbol{I}_m, \boldsymbol{C}_0 = \boldsymbol{0}$,将上式代入(5.9.16)并由(5.9.12)有

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \bar{\boldsymbol{B}}(q^{-1})\boldsymbol{w}(t+n_2-1) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t),$$

$$\bar{\boldsymbol{B}}(q^{-1}) = \boldsymbol{C}(q^{-1})q^{-(n_2-1)} + \boldsymbol{A}(q^{-1})\boldsymbol{R}(q^{-1}) \tag{5.9.18}$$

将 $\bar{\boldsymbol{B}}(q^{-1})$ 表为

$$\bar{\boldsymbol{B}}(q^{-1}) = q^{-i_0}\boldsymbol{B}(q^{-1}), \quad \boldsymbol{B}_0 \neq \boldsymbol{0}, \quad i_0 \geqslant 0 \tag{5.9.19}$$

则(5.9.18)成为

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{B}(q^{-1})q^{\tau}\boldsymbol{w}(t) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t), \quad \tau = n_2 - 1 - i_0 \tag{5.9.20}$$

它们相同于(5.8.25).于是有 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.9.21}$$

其中 $\boldsymbol{D}(q^{-1})$ 是稳定的,$\boldsymbol{D}_0=\boldsymbol{I}_m$,新息 $\boldsymbol{\varepsilon}(t)\in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声,且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{B}(q^{-1})q^{\tau}\boldsymbol{w}(t)+\boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.9.22}$$

可用 Gevers - Wouters 算法求 $\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$.

应注意,ARMA 新息模型(5.9.21)完全相同于用节 5.8 方法求得的 ARMA 新息模型(5.8.26).

【引理 5.9.2】 MA 有色观测噪声 $\boldsymbol{\eta}(t)$ 有新息滤波器

$$\hat{\boldsymbol{\eta}}(t\mid t+N)=\boldsymbol{L}_N^{\eta}(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.9.23}$$

其中多项式矩阵 $\boldsymbol{L}_N^{\eta}(q^{-1})$ 定义为

$$\boldsymbol{L}_N^{\eta}(q^{-1})=\sum_{i=0}^{n_r}\boldsymbol{R}_i\boldsymbol{L}_{N+i-\tau_0}^{w}(q^{-1})+\boldsymbol{L}_N^{v}(q^{-1}) \tag{5.9.24}$$

其中 $\tau_0=n_2-1$,$\boldsymbol{R}_i$ 为 $\boldsymbol{R}(q^{-1})$ 的系数阵,$\boldsymbol{L}_N^{w}(q^{-1})$ 和 $\boldsymbol{L}_N^{v}(q^{-1})$ 由引理 5.8.3 的(5.8.45)定义.

证明 由(5.9.12)和 $\tau_0=n_2-1$,$n_r=n_2-1=\tau_0$ 有

$$\boldsymbol{\eta}(t)=\sum_{i=0}^{\tau_0}\boldsymbol{R}_i\boldsymbol{w}(t+\tau_0-i)+\boldsymbol{v}(t) \tag{5.9.25}$$

上式两边取射影运算有

$$\hat{\boldsymbol{\eta}}(t\mid t+N)=\sum_{i=0}^{\tau_0}\boldsymbol{R}_i\hat{\boldsymbol{w}}(t+\tau_0-i\mid t+N)+\hat{\boldsymbol{v}}(t\mid t+N) \tag{5.9.26}$$

应用白噪声新息滤波器(5.8.45)有

$$\hat{\boldsymbol{w}}(t+\tau_0-i\mid t+N)=\boldsymbol{L}_{N+i-\tau_0}^{w}(q^{-1})\boldsymbol{\varepsilon}(t+N)$$

$$\hat{\boldsymbol{v}}(t\mid t+N)=\boldsymbol{L}_N^{v}(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.9.27}$$

将它们代入(5.9.26)引出(5.9.23)和(5.9.24). □

【引理 5.9.3】 MA 有色观测噪声 $\boldsymbol{\eta}(t)$ 有 Wiener 滤波器

$$\hat{\boldsymbol{\eta}}(t\mid t+N)=\boldsymbol{L}_N^{\eta}(q^{-1})\tilde{\boldsymbol{A}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.9.28}$$

带伪交换

$$\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})=\tilde{\boldsymbol{A}}(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{5.9.29}$$

带 $\tilde{\boldsymbol{D}}_0=\boldsymbol{I}_m$,$\det\tilde{\boldsymbol{D}}(q^{-1})=\det\boldsymbol{D}(q^{-1})$.

证明 由(5.9.21)有

$$\boldsymbol{\varepsilon}(t+N)=\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{y}(t+N) \tag{5.9.30}$$

将它们代入(5.9.23)后,利用(5.9.29)得(5.9.28). □

【定理 5.9.1】 子系统(5.9.5)和(5.9.11)有非递推状态表达式

$$\boldsymbol{x}_1(t)=\sum_{i=0}^{\beta_1-1}\boldsymbol{\Omega}_i\left[\boldsymbol{y}(t+i)-\sum_{j=0}^{i-1}\boldsymbol{H}_1\boldsymbol{\Phi}_1^{i-1-j}\boldsymbol{\Gamma}_1\boldsymbol{w}(t+j)-\boldsymbol{\eta}(t+i)\right] \tag{5.9.31}$$

其中 β_1 为 $(\boldsymbol{\Phi}_1,\boldsymbol{H}_1)$ 的可观性指数,可观阵 $\boldsymbol{\Omega}$ 定义为

$$\boldsymbol{\Omega}=[\boldsymbol{H}_1^{\mathrm{T}},(\boldsymbol{H}_1\boldsymbol{\Phi}_1)^{\mathrm{T}},\cdots,(\boldsymbol{H}_1\boldsymbol{\Phi}_1^{\beta_1-1})^{\mathrm{T}}]^{\mathrm{T}} \tag{5.9.32}$$

且将伪逆 $\boldsymbol{\Omega}^{+}=(\boldsymbol{\Omega}^{\mathrm{T}}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathrm{T}}$ 分块表示为

$$\boldsymbol{\Omega}^{+}=[\boldsymbol{\Omega}_0,\boldsymbol{\Omega}_1,\cdots,\boldsymbol{\Omega}_{\beta_1-1}] \tag{5.9.33}$$

还规定 $\boldsymbol{\Phi}_1^i=\mathbf{0}(i<0)$ 且 $j\geqslant 0$，即规定当 $i=0$ 时(5.9.31)中括号内第二项为零.非递推稳态最优状态估值器为

$$\hat{\boldsymbol{x}}_1(t\mid t+N)=\sum_{i=0}^{\beta-1}\boldsymbol{\Omega}_i\Big[\hat{\boldsymbol{y}}(t+i\mid t+N)-\sum_{j=0}^{i-1}\boldsymbol{H}_1\boldsymbol{\Phi}_1^{i-1-j}\boldsymbol{\Gamma}_1\hat{\boldsymbol{w}}(t+j\mid t+N)-\hat{\boldsymbol{\eta}}(t+i\mid t+N)\Big] \tag{5.9.34}$$

状态 $\boldsymbol{x}_2(t)$ 有非递推表达式(5.9.9)，即

$$\boldsymbol{x}_2(t)=\sum_{i=0}^{\tau_0}\bar{\boldsymbol{R}}_i\boldsymbol{w}(t+\tau_0-i) \tag{5.9.35}$$

其中 $\tau_0=n_2-1$，$\bar{\boldsymbol{R}}_i$ 为由(5.9.10)定义的 $\bar{\boldsymbol{R}}(q^{-1})$ 的系数阵.非递推稳态最优状态估值器为

$$\hat{\boldsymbol{x}}_2(t\mid t+N)=\sum_{i=0}^{\tau_0}\bar{\boldsymbol{R}}_i\hat{\boldsymbol{w}}(t+\tau_0-i\mid t+N) \tag{5.9.36}$$

证明 因由引理5.9.1$(\boldsymbol{\Phi}_1,\boldsymbol{H}_1)$为完全可观对，故用定理5.8.1的证明方法得(5.9.31)～(5.9.34).由(5.9.9)引出(5.9.35)和(5.9.36). □

【引理5.9.4】 子系统(5.9.5)和(5.9.11)有状态估值误差 $\tilde{\boldsymbol{x}}_1(t\mid t+N)=\boldsymbol{x}_1(t)-\hat{\boldsymbol{x}}_1(t\mid t+N)$ 表达式

$$\tilde{\boldsymbol{x}}_1(t\mid t+N)=\sum_{i=0}^{\beta_1-1+\tau_0}\boldsymbol{\Omega}_i^w\boldsymbol{w}(t+i)-\sum_{i=0}^{\beta_1-1}\boldsymbol{\Omega}_i\boldsymbol{v}(t+i)+\sum_{i=0}^{\beta_1-1}\boldsymbol{\Omega}_i\tilde{\boldsymbol{y}}(t+i\mid t+N)-\sum_{i=0}^{\beta_1-1+\tau_0}\boldsymbol{\Omega}_i^w\hat{\boldsymbol{w}}(t+i\mid t+N)+\sum_{i=0}^{\beta_1-1}\boldsymbol{\Omega}_i\hat{\boldsymbol{v}}(t+i\mid t+N) \tag{5.9.37}$$

证明 由(5.9.31)减(5.9.34)，应用(5.9.12)，注意 $\tau_0=n_2-1\geqslant 0$，合并 $w(t+i)$ 的同类项得系数阵 $\boldsymbol{\Omega}_i^w$. □

【定理5.9.2】 子系统(5.9.5)和(5.9.11)有状态估值误差表达式

$$\tilde{\boldsymbol{x}}_1(t\mid t+N)=\sum_{i=0}^{\beta_1-1+\tau_0}\boldsymbol{\Omega}_i^w\boldsymbol{w}(t+i)-\sum_{i=0}^{\beta_1-1}\boldsymbol{\Omega}_i\boldsymbol{v}(t+i)+\sum_{i=0}^{t_\tau+\beta_1-1}\boldsymbol{\Omega}_i^c\boldsymbol{\varepsilon}(t-t_\tau+i),$$
$$\tau_0=n_2-1,\quad t_\tau=\max(\tau,0),\quad N<0 \tag{5.9.38}$$

$$\tilde{\boldsymbol{x}}_1(t\mid t+N)=\sum_{i=0}^{\beta_1-1+\tau_0}\boldsymbol{\Omega}_i^w\boldsymbol{w}(t+i)-\sum_{i=0}^{\beta_1-1}\boldsymbol{\Omega}_i\boldsymbol{v}(t+i)+\sum_{i=0}^{n_0+t_\tau}\boldsymbol{\Omega}_i^d\boldsymbol{\varepsilon}(t-t_\tau+i),$$
$$n_0=\max(\beta_1-1,N),\quad N\geqslant 0 \tag{5.9.39}$$

其中 $\boldsymbol{\Omega}_i^c,\boldsymbol{\Omega}_i^d$ 由合并同类项得到.合并上两式有

$$\tilde{\boldsymbol{x}}_1(t\mid t+N)=\sum_{i=0}^{\beta_1-1+\tau_0}\boldsymbol{\Omega}_i^w\boldsymbol{w}(t+i)-\sum_{i=0}^{\beta_1-1}\boldsymbol{\Omega}_i\boldsymbol{v}(t+i)+\sum_{i=0}^{n_0+t_\tau}\boldsymbol{\Omega}_i^e\boldsymbol{\varepsilon}(t-t_\tau+i) \tag{5.9.40}$$

上式可统一写为

$$\tilde{\boldsymbol{x}}_1(t\mid t+N)=\sum_{i=0}^{n_1}[\boldsymbol{\Omega}_i^w\boldsymbol{w}(t+i)+\boldsymbol{\Omega}_i^v\boldsymbol{v}(t+i)+\boldsymbol{\Omega}_i^\varepsilon\boldsymbol{\varepsilon}(t-t_\tau+i)] \tag{5.9.41}$$

$$n_1 = \max(\beta_1 - 1 + \tau_0, n_0 + t_\tau) \tag{5.9.42}$$

证明 平行于定理 5.8.2 的证明可得定理 5.9.2. 详细推导从略. □

【定理 5.9.3】 子系统(5.9.6)有状态估值误差 $\tilde{\boldsymbol{x}}_2(t \mid t+N) = \boldsymbol{x}_2(t) - \hat{\boldsymbol{x}}_2(t \mid t+N)$ 表达式

$$\tilde{\boldsymbol{x}}_2(t \mid t+N) = \sum_{i=0}^{\tau_0} \tilde{\boldsymbol{R}}_i \boldsymbol{w}(t+i) - \sum_{i=0}^{\tau_0} \tilde{\boldsymbol{R}}_i \hat{\boldsymbol{w}}(t+i \mid t+N) \tag{5.9.43}$$

其中 $\tilde{\boldsymbol{R}}_i = \bar{\boldsymbol{R}}_{\tau_0 - i}$. 它可表为

$$\tilde{\boldsymbol{x}}_2(t \mid t+N) = \sum_{i=0}^{\tau_0} \tilde{\boldsymbol{R}}_i \boldsymbol{w}(t+i) - \sum_{i=0}^{N+t_\tau} \boldsymbol{\Omega}_i^a \boldsymbol{\varepsilon}(t - t_\tau + i), \quad N \geqslant -t_\tau \tag{5.9.44}$$

$$\tilde{\boldsymbol{x}}_2(t \mid t+N) = \sum_{i=0}^{\tau_0} \tilde{\boldsymbol{R}}_i \boldsymbol{w}(t+i), \quad N < -t_\tau \tag{5.9.45}$$

其中 $\boldsymbol{\Omega}_i^a$ 由合并同类项得到.

证明 注意 $\hat{\boldsymbol{w}}(t+i \mid t+N) = \hat{\boldsymbol{w}}(t+i \mid t+i+N-i)$, 由(5.8.41)当 $N - i \geqslant -t_\tau$ 时它是$(\boldsymbol{\varepsilon}(t+i-t_\tau), \cdots, \boldsymbol{\varepsilon}(t+N))$的线性组合. 由(5.8.43)当 $N - i < -t_\tau$, $\hat{\boldsymbol{w}}(t+i \mid t+N) = \boldsymbol{0}$. 因此当 $N < -t_\tau$ 引出 $N - i < -t_\tau (i = 0,1,\cdots)$, 从而有(5.9.45)成立. 当 $N < -t_\tau$ 时有(5.9.43)右边第二项是$(\boldsymbol{\varepsilon}(t-t_\tau), \cdots, \boldsymbol{\varepsilon}(t+N))$的线性组合. 故(5.9.44)成立. □

【定理 5.9.4】 子系统(5.9.5)和(5.9.11)有估值误差方差阵 $\boldsymbol{P}_1(N) = \mathrm{E}[\tilde{\boldsymbol{x}}_1(t \mid t+N)\tilde{\boldsymbol{x}}_1^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}_1(N) = \sum_{i=0}^{n_1}\sum_{j=0}^{n_1} [\boldsymbol{\Omega}_i^w, \boldsymbol{\Omega}_i^v, \boldsymbol{\Omega}_i^\varepsilon] \begin{bmatrix} \boldsymbol{Q}_w \delta_{ij} & \boldsymbol{S}\delta_{ij} & \boldsymbol{\Lambda}_{j-i-t_\tau}^w \\ \boldsymbol{S}^{\mathrm{T}}\delta_{ij} & \boldsymbol{Q}_v \delta_{ij} & \boldsymbol{\Lambda}_{j-i-t_\tau}^v \\ \boldsymbol{\Lambda}_{i-j-t_\tau}^{w\mathrm{T}} & \boldsymbol{\Lambda}_{i-j-t_\tau}^{v\mathrm{T}} & \boldsymbol{Q}_\varepsilon \delta_{ij} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega}_j^{w\mathrm{T}} \\ \boldsymbol{\Omega}_j^{v\mathrm{T}} \\ \boldsymbol{\Omega}_j^{\varepsilon\mathrm{T}} \end{bmatrix} \tag{5.9.46}$$

其中 $\boldsymbol{\Lambda}_i^w, \boldsymbol{\Lambda}_i^v$ 由(5.8.28)计算.

证明 应用(5.9.41)和定理 5.8.3 得证. □

【定理 5.9.5】 子系统(5.9.6)有状态估值误差方差阵 $\boldsymbol{P}_1(N) = \mathrm{E}[\tilde{\boldsymbol{x}}_2(t \mid t+N)\tilde{\boldsymbol{x}}_2^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}_2(N) = \sum_{i=0}^{n_2}\sum_{j=0}^{n_2} [\tilde{\boldsymbol{R}}_i, \tilde{\boldsymbol{\Omega}}_i^a] \begin{bmatrix} \boldsymbol{Q}_w \delta_{ij} & \boldsymbol{\Lambda}_{j-i-t_\tau}^w \\ \boldsymbol{\Lambda}_{i-j-t_\tau}^{w\mathrm{T}} & \boldsymbol{Q}_\varepsilon \delta_{ij} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{R}}_j^{\mathrm{T}} \\ \tilde{\boldsymbol{\Omega}}_j^{a\mathrm{T}} \end{bmatrix} \tag{5.9.47}$$

$$n_2 = \max(\tau_0, N + t_\tau), \quad N \geqslant -t_\tau \tag{5.9.48}$$

其中规定 $\tilde{\boldsymbol{R}}_i = \boldsymbol{0}(i > \tau_0)$, $\boldsymbol{\Omega}_i^a = \boldsymbol{0}(i > N + t_\tau)$. 而且

$$\boldsymbol{P}_2(N) = \sum_{i=0}^{\tau_0} \tilde{\boldsymbol{R}}_i \boldsymbol{Q}_w \tilde{\boldsymbol{R}}_v^{\mathrm{T}}, \quad N < -t_\tau \tag{5.9.49}$$

证明 $N \geqslant -t_\tau$, 应用(5.8.28)由(5.9.44)得(5.9.47). 当 $N < -t_\tau$, 由(5.9.45)引出(5.9.49). □

【定理 5.9.6】 广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下有渐近稳定的降阶 Wiener 状态估值器

$$\hat{\boldsymbol{x}}_1(t \mid t+N) = \boldsymbol{K}_N^{(1)}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.9.50}$$

或

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{K}_N^{(1)}(q^{-1})\mathrm{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t+N) \tag{5.9.51}$$

其中多项式矩阵 $\boldsymbol{K}_N^{(1)}(q^{-1})$ 定义为

$$\boldsymbol{K}_N(q^{-1}) = \sum_{i=0}^{\beta_1-1}\boldsymbol{\Omega}_i\left[\boldsymbol{J}_{i-N}(q^{-1}) - \sum_{j=0}^{i-1}\boldsymbol{H}_1\boldsymbol{\Phi}_1^{i-1-j}\boldsymbol{\Gamma}_1\boldsymbol{L}_{N-j}^{w}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) - \boldsymbol{L}_{N-i}^{\eta}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})\right] \tag{5.9.52}$$

且有

$$\hat{\boldsymbol{x}}_2(t \mid t+N) = \boldsymbol{K}_N^{(2)}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{5.9.53}$$

或

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{x}}_2(t \mid t+N) = \boldsymbol{K}_N^{(2)}(q^{-1})\mathrm{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t+N) \tag{5.9.54}$$

其中定义

$$\boldsymbol{K}_N^{(2)}(q^{-1}) = \sum_{i=0}^{\tau_0}\overline{\boldsymbol{R}}_i\boldsymbol{L}_{N+i-\tau_0}^{w}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) \tag{5.9.55}$$

证明　将(5.8.50)、(5.8.53)和(5.9.28)代入(5.9.34)和(5.9.36)得证．由 $\det\widetilde{\boldsymbol{D}}(q^{-1})$ 是稳定的多项式，引出估值器的渐近稳定性． □

【推论 5.9.1】　对单输出广义系统(5.8.1)和(5.8.2)在假设 1 ~ 5 下，有渐近稳定降阶 Wiener 状态估值器

$$D(q^{-1})\hat{x}_1(t \mid t+N) = K_N^{(1)}(q^{-1})y(t+N) \tag{5.9.56}$$

$$D(q^{-1})\hat{x}_2(t \mid t+N) = K_N^{(2)}(q^{-1})y(t+N) \tag{5.9.57}$$

其中 $K_N^{(1)}(q^{-1})$ 和 $K_N^{(2)}(q^{-1})$ 由(5.9.52)和(5.9.55)置 $\widetilde{\boldsymbol{A}}(q^{-1}) = A(q^{-1})$ 来计算．

5.10　基于 ARMA 新息模型的稳态 Kalman 滤波器和预报器

文献[1,171,172]提出了基于 ARMA 新息模型的稳态 Kalman 滤波器和预报器增益的若干新算法．但算法的推导仅限于稳定系统(状态转置阵为稳定矩阵)，或限于单输出系统．本节利用 Diophantine 方程对系统输出进行有限展开，即将输出表为有限项之和形式，而不要求展为无穷级数，证明了这些增益算法对不稳定系统也成立．

考虑线性离散时不变随机系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}\boldsymbol{x}(t) + \boldsymbol{\Gamma}\boldsymbol{w}(t) \tag{5.10.1}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}(t) \tag{5.10.2}$$

其中 t 为离散时间，$\boldsymbol{x}(t) \in R^n$ 为状态，$\boldsymbol{y}(t) \in R^m$ 为观测，$\boldsymbol{\Phi}, \boldsymbol{\Gamma}, \boldsymbol{H}$ 为已知适当维数常阵．

【假设 1】　$\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}(t) \in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_v$ 的相关阵为 $\boldsymbol{S}$ 的相关白噪声：

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{w}^{\mathrm{T}}(j) & \boldsymbol{v}^{\mathrm{T}}(j)\end{bmatrix}\right\} = \begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix}\delta_{tj},\quad \begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}\\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v\end{bmatrix} > 0 \tag{5.10.3}$$

其中 E 为均值号，T 为转置号，$\delta_{tt} = 1, \delta_{tj} = 0 (t \neq j)$．

【假设 2】[173]　$(\boldsymbol{\Phi}, \boldsymbol{H})$ 为完全可观对，$(\overline{\boldsymbol{\Phi}}, \boldsymbol{\Gamma}\boldsymbol{Q}_0)$ 为完全能稳对，其中 $\overline{\boldsymbol{\Phi}} = \overline{\boldsymbol{\Phi}} - \boldsymbol{J}\boldsymbol{H}$，

$J = \Gamma S Q_v^{-1}, \bar{Q} = Q_w - S Q_v^{-1} S^{\mathrm{T}}, \bar{Q} = Q_0 Q_0^{\mathrm{T}}$.

【假设 3】 $\boldsymbol{\Phi}$ 是稳定或不稳定矩阵.

【假设 4】 初始观测时刻 $t_0 = -\infty$.

问题是求稳态 Kalman 滤波器和预报器增益.

5.10.1 ARMA 新息模型

由(5.10.1)和(5.10.2)有

$$y(t) = H(I_n - q^{-1}\Phi)^{-1}\Gamma q^{-1} w(t) + v(t) \tag{5.10.4}$$

其中 I_n 为 $n \times n$ 单位阵,q^{-1} 为单位滞后算子.引入左素分解

$$H(I_n - q^{-1}\Phi)^{-1}\Gamma q^{-1} = A^{-1}(q^{-1})B(q^{-1}) \tag{5.10.5}$$

其中 $A(q^{-1})$ 和 $B(q^{-1})$ 为形如 $X(q^{-1}) = X_0 + X_1 q^{-1} + \cdots + X_{n_x} q^{-n_x}$ 的多项式矩阵,且 $A_0 = I_m, B_0 = 0$.将上式代入(5.10.4)引出

$$A(q^{-1})y(t) = B(q^{-1})w(t) + A(q^{-1})v(t) \tag{5.10.6}$$

应用谱分解定理[1],上式右边的两个 MA 过程之和可用一个等价的可逆的 MA 过程 $D(q^{-1})\varepsilon(t)$ 表示为

$$D(q^{-1})\varepsilon(t) = B(q^{-1})w(t) + A(q^{-1})v(t) \tag{5.10.7}$$

其中 $D(q^{-1})$ 是稳定的,$D_0 = I_m$,新息 $\varepsilon(t) \in R^m$ 是零均值,方差阵为 Q_ε 的白噪声,且 $D(q^{-1})$ 和 Q_ε 可用 Gevers - Warters 算法求得.于是由上两式可得 ARMA 新息模型

$$A(q^{-1})y(t) = D(q^{-1})\varepsilon(t) \tag{5.10.8}$$

5.10.2 稳态 Kalman 滤波器

【定理 5.10.1】 系统(5.10.1)和(5.10.2)在假设 1 ~ 4 下有稳态 Kalman 滤波器

$$\hat{x}(t+1 \mid t++1) = \Psi_f \hat{x}(t \mid t) + (I_n - K_f H) J y(t) + K_f y(t+1) \tag{5.10.9}$$

$$\Psi_f = [I_n - K_f H]\bar{\Phi}, \quad \bar{\Phi} = \Phi - JH, \quad J = \Gamma S Q_v^{-1} \tag{5.10.10}$$

其中 Ψ_f 是稳定矩阵(即 Ψ_f 的所有特征值位于单位圆内),滤波增益 K 为

$$K_f = \begin{bmatrix} H \\ H\Phi \\ \vdots \\ H\Phi^{\beta-1} \end{bmatrix}^{+} \begin{bmatrix} I_m - Q_v Q_\varepsilon^{-1} \\ M_1 - H\Gamma S Q_\varepsilon^{-1} \\ \vdots \\ M_{\beta-1} - H\Phi^{\beta-2}\Gamma S Q_\varepsilon^{-1} \end{bmatrix} \tag{5.10.11}$$

其中矩阵 X 的伪逆 X^+ 定义为

$$X^+ = (X^{\mathrm{T}} X)^{-1} X^{\mathrm{T}} \tag{5.10.12}$$

β 为 (Φ, H) 的可观性指数,它是使可观阵 Ω 为列满秩的最小自然数,即

$$\operatorname{rank}\Omega = \operatorname{rank}[H^{\mathrm{T}}, (H\Phi)^{\mathrm{T}}, \cdots, (H\Phi^{\beta-1})^{\mathrm{T}}]^{\mathrm{T}} = n \tag{5.10.13}$$

矩阵 M_i 可递推计算为

$$M_i = -A_1 M_{i-1} - \cdots - A_{n_a} M_{i-n_a} + D_i, \quad i = 1, \cdots, \beta - 1 \tag{5.10.14}$$

其中规定 $M_0 = I_m, M_i = 0 (i < 0), D_i = 0 (i > n_d)$.

证明 由(5.10.2)有 $y(t) - Hx(t) - v(t) = 0$,则(5.10.1)等价于

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t)+\boldsymbol{J}[\boldsymbol{y}(t)-\boldsymbol{Hx}(t)-\boldsymbol{v}(t)] \tag{5.10.15}$$

其中 $\boldsymbol{J}$ 为待定矩阵.置

$$\bar{\boldsymbol{\Phi}}=\boldsymbol{\Phi}-\boldsymbol{JH},\quad \bar{\boldsymbol{w}}(t)=\boldsymbol{\Gamma w}(t)-\boldsymbol{Jv}(t) \tag{5.10.16}$$

则(5.10.15)成为

$$\boldsymbol{x}(t+1)=\bar{\boldsymbol{\Phi}}\boldsymbol{x}(t)+\boldsymbol{Jy}(t)+\bar{\boldsymbol{w}}(t) \tag{5.10.17}$$

且有

$$\mathrm{E}[\bar{\boldsymbol{w}}(t)\boldsymbol{v}^{\mathrm{T}}(k)]=(\boldsymbol{\Gamma S}-\boldsymbol{JQ}_v)\delta_{tk} \tag{5.10.18}$$

选取 $\boldsymbol{J}$ 为

$$\boldsymbol{J}=\boldsymbol{\Gamma S Q}_v^{-1} \tag{5.10.19}$$

则有

$$\mathrm{E}[\bar{\boldsymbol{w}}(t)\boldsymbol{v}^{\mathrm{T}}(k)]=\boldsymbol{0},\quad \forall\, t,k \tag{5.10.20}$$

即 $\bar{\boldsymbol{w}}(t)$ 与 $\boldsymbol{v}$ 是不相关的,且 $\boldsymbol{w}(t)$ 是白噪声,

$$\mathrm{E}\bar{\boldsymbol{w}}(t)=\boldsymbol{0},\quad \mathrm{E}[\bar{\boldsymbol{w}}(t)\boldsymbol{v}^{\mathrm{T}}(k)]=\boldsymbol{\Gamma}(\boldsymbol{Q}_w-\boldsymbol{SQ}_v^{-1}\boldsymbol{S}^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}}\delta_{tk} \tag{5.10.21}$$

对于带不相关噪声的等价系统(5.10.17)和(5.10.2)应用射影理论[193]有稳态Kalman滤波器

$$\hat{\boldsymbol{x}}(t+1\mid t+1)=\hat{\boldsymbol{x}}(t+1\mid t)+\boldsymbol{K}_f\boldsymbol{\varepsilon}(t+1) \tag{5.10.22}$$

$$\hat{\boldsymbol{x}}(t+1\mid t)=\bar{\boldsymbol{\Phi}}\hat{\boldsymbol{x}}(t\mid t)+\boldsymbol{Jy}(t) \tag{5.10.23}$$

$$\boldsymbol{\varepsilon}(t+1)=\boldsymbol{y}(t+1)-\boldsymbol{H}\hat{\boldsymbol{x}}(t+1\mid t) \tag{5.10.24}$$

其中 $\boldsymbol{K}_f$ 是待定的滤波增益阵.将(5.10.23)和(5.10.24)代入(5.10.22)引出(5.10.9).假设2引出 $\boldsymbol{\Psi}_f$ 是一个稳定矩阵[173].另一方面,应用射影性质由(5.10.1)和(5.10.2)有

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t\mid t)+\boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t\mid t) \tag{5.10.25}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t\mid t)+\hat{\boldsymbol{v}}(t\mid t) \tag{5.10.26}$$

其中由节5.1有白噪声滤波器

$$\hat{\boldsymbol{w}}(t\mid t)=\boldsymbol{SQ}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t),\quad \hat{\boldsymbol{v}}(t\mid t)=\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t) \tag{5.10.27}$$

将(5.10.25)和(5.10.27)代入(5.10.22)有

$$\hat{\boldsymbol{x}}(t+1\mid t+1)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t\mid t)+\boldsymbol{\Gamma SQ}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t)+K_f\boldsymbol{\varepsilon}(t+1) \tag{5.10.28}$$

再将上两式代入(5.10.26)引出

$$\boldsymbol{y}(t)=\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}[\boldsymbol{\Gamma SQ}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t-1)+\boldsymbol{K}_f\boldsymbol{\varepsilon}(t)]+\boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t) \tag{5.10.29}$$

因为 $\boldsymbol{\Phi}$ 可以是不稳定矩阵,因此可用引入如下Diophantine方程将上式展开为有限项之和:

$$\boldsymbol{I}_n=(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})\boldsymbol{\Phi}(q^{-1})+q^{-\beta}\boldsymbol{R}_0 \tag{5.10.30}$$

$$\boldsymbol{\Phi}(q^{-1})=\boldsymbol{\Phi}_0+\boldsymbol{\Phi}_1q^{-1}+\cdots+\boldsymbol{\Phi}_{\beta-1}q^{-(\beta-1)} \tag{5.10.31}$$

用比较(5.10.30)两边 q^{-j} 的系数阵的方法可得

$$\boldsymbol{\Phi}_i=\boldsymbol{\Phi}^i,\quad \boldsymbol{R}_0=\boldsymbol{\Phi}^\beta,\quad i=0,1,\cdots,\beta-1 \tag{5.10.32}$$

于是有有限展开式

$$(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}=\sum_{i=0}^{\beta-1}\boldsymbol{\Phi}^iq^{-i}+q^{-\beta}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\beta \tag{5.10.33}$$

这一展式对 $\boldsymbol{\Phi}$ 稳定或 $\boldsymbol{\Phi}$ 不稳定两种情形均成立.将(5.10.33)代入(5.10.29)中可得

$$
\begin{aligned}
\boldsymbol{y}(t) = & \sum_{i=0}^{\beta-1} \boldsymbol{H}\boldsymbol{\Phi}^i q^{-i}[\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t-1) + \boldsymbol{K}_f\boldsymbol{\varepsilon}(t)] + \boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t) + \\
& \boldsymbol{H}q^{-\beta}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\beta[\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{\varepsilon}(t-1) + \boldsymbol{K}_f\boldsymbol{\varepsilon}(t)]
\end{aligned} \tag{5.10.34}
$$

因为在 $\boldsymbol{\Phi}$ 为不稳定阵情形 $\boldsymbol{A}(q^{-1})$ 是不稳定的多项矩阵，因此(5.10.8)中 $\boldsymbol{y}(t)$ 不能展成关于 $\boldsymbol{\varepsilon}(t)$ 的收敛的无穷级数，但可将 $\boldsymbol{y}(t)$ 展成有限项和的形式. 为此，引入 Diophantine 方程

$$
\boldsymbol{D}(q^{-1}) = \boldsymbol{A}(q^{-1})\boldsymbol{M}(q^{-1}) + q^{-\beta}\boldsymbol{G}(q^{-1}) \tag{5.10.35}
$$

$$
\boldsymbol{M}_\beta(q^{-1}) = \boldsymbol{M}_0 + \boldsymbol{M}_1 q^{-1} + \cdots + \boldsymbol{M}_{\beta-1}q^{-(\beta-1)} \tag{5.10.36}
$$

其中用比较(5.10.35)中 q^{-i} 的系数阵方法可得(5.10.14)，多项式矩阵 $\boldsymbol{G}(q^{-1})$ 的阶次 $n_g = \max(n_a - 1, n_d - \beta)$. 于是由(5.10.8)和(5.10.35)有展式

$$
\boldsymbol{y}(t) = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \boldsymbol{M}_\beta(q^{-1})\boldsymbol{\varepsilon}(t) + q^{-\beta}\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{G}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.10.37}
$$

比较展式(5.10.34)和(5.10.37)右边 q^{-i} 的系数阵引出关系

$$
\begin{gathered}
\boldsymbol{M}_0 = \boldsymbol{H}\boldsymbol{K}_f + \boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}, \quad \boldsymbol{M}_0 = \boldsymbol{I}_m, \\
\boldsymbol{M}_i = \boldsymbol{H}\boldsymbol{\Phi}^{i-1}\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1} + \boldsymbol{H}\boldsymbol{\Phi}^i\boldsymbol{K}_f, \quad i = 1,\cdots,\beta-1
\end{gathered} \tag{5.10.38}
$$

(5.10.38)可写成

$$
\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{\Phi} \\ \vdots \\ \boldsymbol{H}\boldsymbol{\Phi}^{\beta-1} \end{bmatrix} \boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{I}_m - \boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1} \\ \boldsymbol{M}_1 - \boldsymbol{H}\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1} \\ \vdots \\ \boldsymbol{M}_{\beta-1} - \boldsymbol{H}\boldsymbol{\Phi}^{\beta-2}\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_\varepsilon^{-1} \end{bmatrix} \tag{5.10.39}
$$

对上式取伪逆运算得(5.10.11). □

【推论 5.10.1】 系统(5.10.1)和(5.10.2)在假设 1 ~ 4 下，若 $\boldsymbol{S} = \boldsymbol{0}$，即 $\boldsymbol{w}(t)$ 与 $\boldsymbol{v}(t)$ 不相关，则有稳态 Kalman 滤波器

$$
\hat{\boldsymbol{x}}(t+1 \mid t+1) = \boldsymbol{\Psi}_f\hat{\boldsymbol{x}}(t \mid t) + \boldsymbol{K}_f\boldsymbol{y}(t+1) \tag{5.10.40}
$$

$$
\boldsymbol{\Psi}_f = [\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Phi} \tag{5.10.41}
$$

$$
\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{\Phi} \\ \vdots \\ \boldsymbol{H}\boldsymbol{\Phi}^{\beta-1} \end{bmatrix}^+ \begin{bmatrix} \boldsymbol{I}_m - \boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1} \\ \boldsymbol{M}_1 \\ \vdots \\ \boldsymbol{M}_{\beta-1} \end{bmatrix} \tag{5.10.42}
$$

其中 $\boldsymbol{M}_i$ 由(5.10.14)计算.

【注 5.10.1】 定理 5.10.1 利用了输出 $\boldsymbol{y}(t)$ 的有限展式，或利用 $(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}$ 和 $\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{D}(q^{-1})$ 的有限展式，用比较系数法得到 $\boldsymbol{K}_f$ 所满足的 β 个关系式，进而求得了 $\boldsymbol{K}_f$. 这种有限项之和的展式适用于 $\boldsymbol{\Phi}$ 为稳定矩阵或 $\boldsymbol{\Phi}$ 为不稳定矩阵两种情形. 实际上我们只需要 $\boldsymbol{K}_f$ 所满足的 β 个关系式，因而并不要求无穷级数形式的展式. 而对 $\boldsymbol{\Phi}$ 不稳定情形，无穷级数展式是不收敛的，是无意义的. 只有当 $\boldsymbol{\Phi}$ 为稳定矩阵，这引出 $\boldsymbol{A}(q^{-1})$ 是稳定的多项式矩阵，我们才有收敛的无穷级数展式

$$
(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1} = \sum_{i=0}^{\infty} \boldsymbol{\Phi}^i q^{-i} \tag{5.10.43}
$$

$$\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{D}(q^{-1}) = \sum_{i=0}^{\infty}\boldsymbol{M}_i q^{-i} \tag{5.10.44}$$

其中 $\boldsymbol{M}_i$ 由(5.10.14)计算.但既使在这种情形下,展为无穷级数也是不必要的,因为 $\boldsymbol{K}_f$ 被展式的前 β 项惟一决定.注意,引入有限展式归结为解 Diophantine 方程,它相当于“综合除法”.

【定理 5.10.2】 系统(5.10.1)和(5.10.2)在假设 1 ~ 4 下,稳态滤波误差 $\tilde{\boldsymbol{x}}(t \mid t) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t \mid t)$ 方差阵 $\boldsymbol{P} = \mathrm{E}[\tilde{\boldsymbol{x}}(t \mid t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t \mid t)]$ 满足 Lyapunov 方程

$$\boldsymbol{P} = \boldsymbol{\Psi}_f \boldsymbol{P} \boldsymbol{\Psi}_f^{\mathrm{T}} + \boldsymbol{\Delta}_f \tag{5.10.45}$$

$$\boldsymbol{\Delta}_f = [\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Gamma}(\boldsymbol{Q}_w - \boldsymbol{S}\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}}[\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]^{\mathrm{T}} + \boldsymbol{K}_f\boldsymbol{Q}_v\boldsymbol{K}_f^{\mathrm{T}} \tag{5.10.46}$$

它可用迭代法求解

$$\boldsymbol{P}(t+1) = \boldsymbol{\Psi}_f \boldsymbol{P}(t) \boldsymbol{\Psi}_f^{\mathrm{T}} + \boldsymbol{\Delta}_f \tag{5.10.47}$$

带任意初值 $\boldsymbol{P}(0) = \alpha\boldsymbol{I}_n$,$\alpha > 0$,则有

$$\lim_{t\to\infty}\boldsymbol{P}(t) = \boldsymbol{P} \tag{5.10.48}$$

带指数收敛速度

$$\lim_{t\to\infty}\frac{\|\boldsymbol{\delta}(t)\|}{\rho^t} = 0, \quad 0 < \rho < 1 \tag{5.10.49}$$

其中迭代误差 $\boldsymbol{\delta}(t) = \boldsymbol{P} - \boldsymbol{P}(t)$,且 $\|\cdot\|$ 表示任意矩阵范数,ρ_f 为矩阵 $\boldsymbol{\Psi}_f$ 的谱半径,$0 < \rho_f < 1$,而任取 ρ_0 满足关系 $0 < \rho_f < \rho_0 < 1$,且定义 $\rho = \rho_0^2$.上式用符号记为

$$\|\boldsymbol{\delta}(t)\| = o(\rho^t) \tag{5.10.50}$$

因而当 t 充分大有

$$\boldsymbol{P} \approx \boldsymbol{P}(t) \tag{5.10.51}$$

证明 由(5.10.2)和(5.10.24)引出

$$\boldsymbol{\varepsilon}(t+1) = \boldsymbol{H}\tilde{\boldsymbol{x}}(t+1 \mid t) + \boldsymbol{v}(t+1) \tag{5.10.52}$$

其中定义 $\tilde{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{x}(t+1) - \hat{\boldsymbol{x}}(t+1 \mid t)$.由(5.10.17)减(5.10.23)引出

$$\tilde{\boldsymbol{x}}(t+1 \mid t) = \bar{\boldsymbol{\Phi}}\tilde{\boldsymbol{x}}(t \mid t) + \bar{\boldsymbol{w}}(t) \tag{5.10.53}$$

将(5.10.23)和上两式代入(5.10.22),且由(5.10.17)减(5.10.22)可得

$$\tilde{\boldsymbol{x}}(t+1 \mid t+1) = \boldsymbol{\Psi}_f\tilde{\boldsymbol{x}}(t \mid t) + [\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]\bar{\boldsymbol{w}}(t) - \boldsymbol{K}_f\boldsymbol{v}(t+1) \tag{5.10.54}$$

易知 $\tilde{\boldsymbol{x}}(t \mid t)$,$\bar{\boldsymbol{w}}(t)$ 和 $\boldsymbol{v}(t+1)$ 互不相关,由此引出(5.10.45)和(5.10.46).由(5.10.45)减(5.10.47)引出

$$\boldsymbol{\delta}(t+1) = \boldsymbol{\Psi}_f\boldsymbol{\delta}(t)\boldsymbol{\Psi}_f^{\mathrm{T}} \tag{5.10.55}$$

对上式迭代 t 次引出

$$\boldsymbol{\delta}(t) = \boldsymbol{\Psi}_f^t\boldsymbol{\delta}(0)\boldsymbol{\Psi}_f^{\mathrm{T}t} \tag{5.10.56}$$

由定理 5.10.1,$\boldsymbol{\Psi}_f$ 是一个稳定矩阵.设 $\boldsymbol{\Psi}$ 的特征值为 $\lambda_1,\cdots,\lambda_s$,$\lambda_i$ 的重数为 n_i,则有 $n_1 + \cdots + n_5 = n$,且 $\boldsymbol{\Psi}_f$ 的谱半径为 $\rho_f = \max(|\lambda_1|,\cdots,|\lambda_s|)$,$0 < \rho_f < 1$.由矩阵理论[107],存在非异矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{\Psi}_f$ 化为约当形,即

$$\boldsymbol{P}^{-1}\boldsymbol{\Psi}_f\boldsymbol{P} = \boldsymbol{J}_0, \quad \boldsymbol{J}_0 = \mathrm{diag}(\boldsymbol{J}_1,\cdots,\boldsymbol{J}_s)$$

$$
\boldsymbol{J}_i = \begin{bmatrix} \lambda_i & 1 & \cdots & 0 \\ & & \ddots & \vdots \\ & & \ddots & 1 \\ \boldsymbol{0} & & & \lambda_i \end{bmatrix}_{n_i \times n_i} \tag{5.10.57}
$$

于是有

$$
\boldsymbol{\Psi}_f^t = \boldsymbol{P}\mathrm{diag}(\boldsymbol{J}_1^t, \cdots, \boldsymbol{J}_s^t)\boldsymbol{P}^{-1}
$$

$$
\boldsymbol{J}_i^t = \begin{bmatrix} \lambda_i^t & c_i^1\lambda_i^{t-1} & \cdots & \cdots & c_i^{n_i-1}\lambda_i^{t-n_i+1} \\ & \lambda_i^t & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & c_i^1\lambda_i^{t-1} \\ \boldsymbol{0} & & & & \lambda_i^t \end{bmatrix} \tag{5.10.58}
$$

其中定义 $c_t^j = t(t-1)\cdots(t-j+1)/j!\,(j \leqslant t)$；$c_t^j = 0\,(j > t)$. $j! = 1 \times 2 \times \cdots \times j$.

记 $\boldsymbol{\Psi}_f^t = (\psi_{rj}^{(t)})$，它的元素为 $\psi_{rj}^{(t)}, r, j = 1, \cdots, n$，则 $\psi_{rj}^{(t)}$ 是形式为 $t^k\lambda_i^{t-\mu}$ 的项的线性组合，其中 $k = 0, 1, \cdots, n_m - 1$，$n_m = \max(n_1, \cdots, n_s)$，$i = 1, \cdots, s$，$\mu = 0, 1, \cdots, n_m - 1$. 应用公式：$t^k\lambda^t \to 0, t \to \infty, 0 < \lambda < 1$，注意 $|\lambda_i| \leqslant \rho_f < \rho_0 < 1$，则有

$$
(\psi_{rj}^{(t)}/\rho_0^t) \to 0, \quad t \to \infty \tag{5.10.59}
$$

记 $\delta(t) = (\delta_{\alpha\beta}^{(t)}), \alpha, \beta = 1, \cdots, n$，则 $\delta_{\alpha\beta}^{(t)}$ 是形式为 $\psi_{rj}^{(t)}\psi_{uv}^{(t)}$ 的项的线性组合，于是有

$$
(\delta_{\alpha\beta}^{(t)}/\rho^t) \to 0, \quad \rho = \rho_0^2, \quad t \to \infty \tag{5.10.60}
$$

定义 $n \times n$ 矩阵 $\boldsymbol{\delta}(t)$ 的范数 $\|\cdot\|_\delta$ 为

$$
\|\boldsymbol{\delta}(t)\|_\delta = \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} |\delta_{\alpha\beta}^{(t)}| \tag{5.10.61}
$$

则有

$$
(\|\boldsymbol{\delta}(t)\|_\delta/\rho_t) \to 0, \quad t \to \infty \tag{5.10.62}
$$

应用矩阵范数的等价性，对任意矩阵范数 $\|\cdot\|$，有关系

$$
c_1\|\boldsymbol{\delta}(t)\|_\delta \leqslant \|\boldsymbol{\delta}(t)\| \leqslant c_2\|\boldsymbol{\delta}(t)\|_\delta \tag{5.10.63}
$$

其中 $c_1 > 0, c_2 > 0$. 由上两式立刻得 $(\|\boldsymbol{\delta}(t)\|/\rho^t) \to 0, t \to \infty$. 证毕. □

【推论 5.10.2】 在定理 5.10.2 条件下，若 $\boldsymbol{S} = \boldsymbol{0}$，则 Lyapunov 方程(45)和(46)成为

$$
\boldsymbol{P} = \boldsymbol{\Psi}_f\boldsymbol{P}\boldsymbol{\Psi}_f^{\mathrm{T}} + \boldsymbol{\Delta}_f \tag{5.10.64}
$$

$$
\boldsymbol{\Delta}_f = [\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}[\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]^{\mathrm{T}} + \boldsymbol{K}_f\boldsymbol{Q}_v\boldsymbol{K}_f^{\mathrm{T}} \tag{5.10.65}
$$

其中 $\boldsymbol{K}_f$ 和 $\boldsymbol{\Psi}_f$ 由推论 5.10.1 计算.

【注 5.10.2】 在经典 Kalman 滤波理论中，滤波误差方差阵 $\boldsymbol{P}$ 是通过预报误差方差阵 $\boldsymbol{\Sigma}$ 来计算的，且有公式 $\boldsymbol{P} = [\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Sigma}$，而 $\boldsymbol{\Sigma}$ 是由 Riccati 方程

$$
\boldsymbol{\Sigma} = \overline{\boldsymbol{\Phi}}[\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\boldsymbol{H}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{Q}_v)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}]\overline{\boldsymbol{\Phi}}^{\mathrm{T}} + \boldsymbol{\Gamma}(\boldsymbol{Q}_w - \boldsymbol{S}\boldsymbol{Q}_v^{-1}\boldsymbol{S}^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}} \tag{5.10.66}
$$

来计算的. 而这里避免了 Riccati 方程，在由 ARMA 新息模型求出 $\boldsymbol{K}_f$ 之后，用解 Lyapunov 方程求 $\boldsymbol{P}$.

5.10.3 稳态 Kalman 预报器

【定理 5.10.3】 系统(5.10.1)和(5.10.2)在假设 1 ~ 4 下有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{K}_p\boldsymbol{y}(t) \tag{5.10.67}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H} \tag{5.10.68}$$

其中 $\boldsymbol{\Psi}_p$ 是一个稳定矩阵，预报增益阵 $\boldsymbol{K}_p$ 为

$$\boldsymbol{K}_p=\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{H\Phi}\\ \vdots\\ \boldsymbol{H\Phi}^{\beta-1}\end{bmatrix}^{+}\begin{bmatrix}\boldsymbol{M}_1\\ \boldsymbol{M}_2\\ \vdots\\ \boldsymbol{M}_{\beta-1}\end{bmatrix} \tag{5.10.69}$$

其中 $\boldsymbol{M}_i$ 由(5.10.14)递推计算.稳态预报误差方差阵 $\boldsymbol{\Sigma}=\mathrm{E}[\tilde{\boldsymbol{x}}(t+1\mid t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+1\mid t)]$满足 Lyapunov 方程

$$\boldsymbol{\Sigma}=\boldsymbol{\Psi}_p\boldsymbol{\Sigma}\boldsymbol{\Psi}_p^{\mathrm{T}}+\boldsymbol{\Delta}_p \tag{5.10.70}$$

$$\boldsymbol{\Delta}_p=\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{K}_p\boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{K}_p^{\mathrm{T}}+\boldsymbol{K}_p\boldsymbol{Q}_v\boldsymbol{K}_p^{\mathrm{T}} \tag{5.10.71}$$

证明　将(5.10.24)，(5.10.25)，(5.10.27)代入(5.10.22)有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{K}_p\boldsymbol{\varepsilon}(t) \tag{5.10.72}$$

$$\boldsymbol{y}(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{\varepsilon}(t) \tag{5.10.73}$$

$$\boldsymbol{K}_p=\boldsymbol{\Phi}\boldsymbol{K}_f+\boldsymbol{\Gamma}\boldsymbol{S}\boldsymbol{Q}_{\varepsilon}^{-1} \tag{5.10.74}$$

下面将推导与 $\boldsymbol{K}_f$ 无关的 $\boldsymbol{K}_p$ 的算法(5.10.69).将(5.10.72)代入(5.10.73)引出

$$\boldsymbol{y}(t)=\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{K}_p\boldsymbol{\varepsilon}(t-1)+\boldsymbol{\varepsilon}(t) \tag{5.10.75}$$

应用有限展式(5.10.33)有 $y(t)$ 的有限形式展式

$$\boldsymbol{y}(t)=\sum_{i=0}^{\beta-1}\boldsymbol{H\Phi}^i q^{-i}\boldsymbol{K}_p\boldsymbol{\varepsilon}(t-1)+\boldsymbol{H}q^{-\beta}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\beta}\boldsymbol{K}_p\boldsymbol{\varepsilon}(t-1)+\boldsymbol{\varepsilon}(t) \tag{5.10.76}$$

另一方面，类似于(5.10.35)和(5.10.36)引入 Diophantine 方程

$$\boldsymbol{D}(q^{-1})=\boldsymbol{A}(q^{-1})\boldsymbol{M}(q^{-1})+q^{-(\beta+1)}\boldsymbol{G}(q^{-1}) \tag{5.10.77}$$

$$\boldsymbol{M}_{\beta+1}(q^{-1})=\boldsymbol{M}_0+\boldsymbol{M}_1q^{-1}+\cdots+\boldsymbol{M}_{\beta}q^{-\beta} \tag{5.10.78}$$

用比较系数法可得 $\boldsymbol{M}_i$ 的递推公式为

$$\boldsymbol{M}_i=-\boldsymbol{A}_1\boldsymbol{M}_{i-1}-\cdots-\boldsymbol{A}_{n_a}\boldsymbol{M}_{i-n_a}+\boldsymbol{D}_i,\quad i=1,\cdots,\beta \tag{5.10.79}$$

其中规定 $\boldsymbol{M}_0=\boldsymbol{I}_m$，$\boldsymbol{M}_i=\boldsymbol{0}(i<0)$，$\boldsymbol{D}_i=\boldsymbol{0}(i>n_d)$，于是由(5.10.8)和(5.10.77)有展式

$$\boldsymbol{y}(t)=\boldsymbol{M}_{\beta+1}(q^{-1})\boldsymbol{\varepsilon}(t)+q^{-(\beta+1)}\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{G}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.10.80}$$

比较(5.10.76)与(5.10.80)右边 q^{-i} 的系数阵引出关系

$$\boldsymbol{M}_i=\boldsymbol{H\Phi}^{i-1}\boldsymbol{K}_p,\quad i=1,\cdots,\beta \tag{5.10.81}$$

于是有

$$\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{H\Phi}\\ \vdots\\ \boldsymbol{H\Phi}^{\beta-1}\end{bmatrix}\boldsymbol{K}_p=\begin{bmatrix}\boldsymbol{M}_1\\ \boldsymbol{M}_2\\ \vdots\\ \boldsymbol{M}_{\beta}\end{bmatrix} \tag{5.10.82}$$

对上式取伪逆运算得(5.10.69).

由(5.10.1)减(5.10.72)，并应用(5.10.52)可得

$$\tilde{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{\Psi}_p \tilde{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{\Gamma w}(t) - \boldsymbol{K}_p \boldsymbol{v}(t) \tag{5.10.83}$$

由此可得(5.10.70)和(5.10.71). □

【注 5.10.3】 可以证明由(5.10.35)和(5.10.36)定义的 $\boldsymbol{M}_\beta(q^{-1})$ 与由(5.10.77)和(5.10.78)定义的 $\boldsymbol{M}_{\beta+1}(q^{-1})$ 有相同的系数阵 $\boldsymbol{M}_0, \boldsymbol{M}_1, \cdots, \boldsymbol{M}_{\beta-1}$,而 $\boldsymbol{M}_{\beta+1}(q^{-1})$ 仅多了一个新的系数阵 $\boldsymbol{M}_\beta$.

【注 5.10.4】 不同于用解 Riccati 方程(5.10.66)求 $\boldsymbol{\Sigma}$ 的方法,这里用解 Lyapunov 方程(5.10.70)求 $\boldsymbol{\Sigma}$. 因 $\boldsymbol{K}_p$ 可单独由 ARMA 新息模型计算,故 $\boldsymbol{\Delta}_p$ 是已知的. 在经典 Kalman 滤波中,$\boldsymbol{K}_p$ 是由 $\boldsymbol{\Sigma}$ 来计算的,即 $\boldsymbol{K}_p = (\boldsymbol{\Phi\Sigma H}^{\mathrm{T}} + \boldsymbol{\Gamma S})[\boldsymbol{H\Sigma H}^{\mathrm{T}} + \boldsymbol{Q}_v]^{-1}$. 注意解 Lyapunov 方程(5.10.70)要比求解 Riccati 方程(5.10.66)简单. 因为在(5.10.66)中要求计算 $(\boldsymbol{H\Sigma H} + \boldsymbol{Q}_v)$ 的逆矩阵.

【例 5.10.1】 考虑跟踪系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi x}(t) + \boldsymbol{\Gamma} w(t) \tag{5.10.84}$$

$$y(t) = \boldsymbol{H x}(t) + v(t) \tag{5.10.85}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} 0.5T^2 \\ T_0 \end{bmatrix}, \quad \boldsymbol{H} = [1 \quad 0] \tag{5.10.86}$$

其中 T_0 为采样周期,状态 $\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}$, $x_1(t)$, $x_2(t)$ 和 $w(t)$ 分别为在采样时刻 tT_0 处运动目标的位置、速度和加速度,$y(t)$ 是对位置的观测信号,$v(t)$ 是观测噪声. 假设 $w(t)$ 和 $v(t)$ 是零均值、方差各为 σ_w^2 和 σ_v^2 的相互独立高斯白噪声. 问题是用基于 ARMA 新息模型和基于 Riccati 方程的两种方法求稳态 Kalman 滤波器 $\hat{\boldsymbol{x}}(t \mid t)$,并验证两种方法的功能等价性.

在仿真中取 $T_0 = 0.2$, $\sigma_w^2 = 1$, $\sigma_v^2 = 4$,可得 ARMA 新息模型为

$$(1 - q^{-1})^2 y(t) = (1 + d_1 q^{-1} + d_2 q^{-2})\varepsilon(t) \tag{5.10.87}$$

$$d_1 = -1.8007, \quad d_2 = 0.8188, \quad \sigma_\varepsilon^2 = 4.8852 \tag{5.10.88}$$

该系统可观性指数 $\beta = 2$,应用公式(5.10.11)有滤波增益 $\boldsymbol{K}_f$ 为

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \end{bmatrix}^{-1} \begin{bmatrix} 1 - \sigma_v^2/\sigma_w^2 \\ M_1 \end{bmatrix} = \begin{bmatrix} 0.1812 \\ 0.0905 \end{bmatrix} \tag{5.10.89}$$

其中由(5.10.14)有 $M_1 = 2 + d_1$. 应用公式(5.10.45)有滤波误差方差阵 P 满足 Lyapunov 方程

$$\boldsymbol{P} = \boldsymbol{\Psi}_f \boldsymbol{P} \boldsymbol{\Psi}_f^{\mathrm{T}} + \boldsymbol{\Delta}_f \tag{5.10.90}$$

$$\boldsymbol{\Delta}_f = [\boldsymbol{I}_2 - \boldsymbol{K}_f \boldsymbol{H}]\boldsymbol{\Gamma}\sigma_w^2\boldsymbol{\Gamma}^{\mathrm{T}}[\boldsymbol{I}_2 - \boldsymbol{K}_f \boldsymbol{H}]^{\mathrm{T}} + \sigma_v^2 \boldsymbol{K}_f \boldsymbol{K}_f^{\mathrm{T}} \tag{5.10.91}$$

$$\boldsymbol{\Psi}_f = [\boldsymbol{I}_2 - \boldsymbol{K}_f \boldsymbol{H}]\boldsymbol{\Phi} \tag{5.10.92}$$

用迭代法可求得

$$\boldsymbol{P} = \begin{bmatrix} 0.7248 & 0.3620 \\ 0.3620 & 0.3805 \end{bmatrix} \tag{5.10.93}$$

下面用经典 Kalman 滤波方法求 $\boldsymbol{K}_f$ 和 $\boldsymbol{P}$. 预报误差方差阵满足 Riccati 方程

$$\boldsymbol{\Sigma} = \boldsymbol{\Phi}[\boldsymbol{\Sigma} - \boldsymbol{\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}} + \sigma_v^2)^{-1}\boldsymbol{H\Sigma}] + \sigma_w^2 \boldsymbol{\Gamma\Gamma}^{\mathrm{T}} \tag{5.10.94}$$

用迭代法求得其解

$$\boldsymbol{\Sigma} = \begin{bmatrix} 0.885\,2 & 0.442\,1 \\ 0.442\,1 & 0.420\,5 \end{bmatrix} \tag{5.10.95}$$

增益 $\boldsymbol{K}_f$ 和滤波误差方差阵 $\boldsymbol{P}$ 各为

$$\boldsymbol{K}_f = \boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \sigma_v^2]^{-1} = \begin{bmatrix} 0.181\,2 \\ 0.090\,5 \end{bmatrix} \tag{5.10.96}$$

$$\boldsymbol{P} = [\boldsymbol{I}_2 - \boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Sigma} = \begin{bmatrix} 0.724\,8 & 0.362\,0 \\ 0.362\,0 & 0.380\,5 \end{bmatrix} \tag{5.10.97}$$

它们分别相同于(5.10.89)和(5.10.93).这验证了两种方法的功能等价性.仿真结果如图5.10.1和图5.10.2所示,其中实线为真实值 $x_1(t)$ 或 $x_2(t)$,虚线为 Kalman 滤波估值 $\hat{x}_1(t \mid t)$ 或 $\hat{x}_2(t \mid t)$.

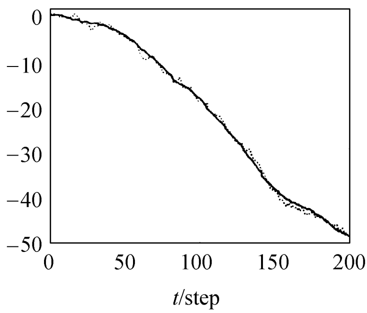


图 5.10.1 位置 $x_1(t)$ 和稳态 Kalman 跟踪滤波器 $\hat{x}_1(t \mid t)$

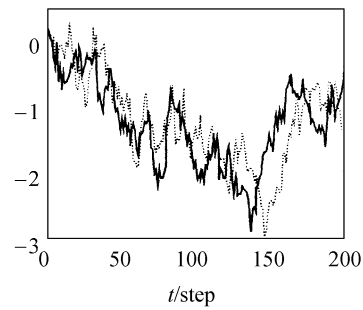


图 5.10.2 速度 $x_2(t)$ 和稳态 Kalman 跟踪滤波器 $\hat{x}_2(t \mid t)$

5.11 基于 ARMA 新息模型的稳态 Kalman 平滑器和多步 Kalman 预报器

考虑系统(5.10.1)和(5.10.2),在假设1 ~ 4下该系统存在稳态Kalman滤波器.现在考虑 Kalman 平滑 $\hat{\boldsymbol{x}}(t \mid t+N)(N>0)$.当初始观测时刻 t_0 为有限时刻时,应用射影理论有最优 Kalman 平滑器

$$\hat{\boldsymbol{x}}(t \mid t+N) = \hat{\boldsymbol{x}}(t \mid t+N-1) + \boldsymbol{K}_N(t)\boldsymbol{\varepsilon}(t+N) \tag{5.11.1}$$

$$\hat{\boldsymbol{x}}(t \mid t+N-1) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t-1 \mid t+N-1) + \boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t-1 \mid t-1+N) \tag{5.11.2}$$

这引出递推 Kalman 平滑器

$$\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t-1 \mid t+N-1) + \boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t-1 \mid t-1+N) + \boldsymbol{K}_N(t)\boldsymbol{\varepsilon}(t+N) \tag{5.11.3}$$

其中 $\boldsymbol{K}_N(t)$ 为最优 Kalman 平滑增益阵.令 $t_0 \to -\infty$,则存在稳态 Kalman 平滑器

$$\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t-1 \mid t+N-1) + \boldsymbol{\Gamma}\hat{\boldsymbol{w}}(t-1 \mid t-1+N) + \boldsymbol{K}_N\boldsymbol{\varepsilon}(t+N) \tag{5.11.4}$$

其中 $\boldsymbol{K}_N$ 为稳态平滑增益阵,且有

$$\boldsymbol{K}_N = \lim_{t_0 \to -\infty} \boldsymbol{K}_N(t) \tag{5.11.5}$$

在经典 Kalman 滤波理论中,稳态平滑增益 $\boldsymbol{K}_N$ 的计算公式为[171]

$$\boldsymbol{K}_N = \boldsymbol{\Sigma}\boldsymbol{\Psi}_p^{\mathrm{T}N}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{Q}_v]^{-1} \tag{5.11.6}$$

其中 $\boldsymbol{\Psi}_p$ 由(5.10.68)定义,$\boldsymbol{\Sigma}$ 由 Riccati 方程(5.10.66)计算,$\boldsymbol{K}_p = (\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Gamma}\boldsymbol{S})[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{Q}_v]^{-1}$. $\boldsymbol{K}_N$ 的计算完全建立在 Riccati 方程解 $\boldsymbol{\Sigma}$ 的基础上.

本节将提出基于 ARMA 新息模型的平滑增益 $\boldsymbol{K}_N$ 的算法.它适用于不稳定和稳定系统(5.10.1)和(5.10.2).推导方法首次在这里提出,不同于文献[1,171,172]的推导方法.文献[1,171,172]的推导严格地说仅适用于稳定系统.因为在推导中用到了数学期望运算 $\mathrm{E}[\boldsymbol{x}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(j)]$,$\mathrm{E}[\boldsymbol{y}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(j)]$,对不稳定系统当 $t_0 = -\infty$ 时,$\boldsymbol{x}(t)$ 和 $\boldsymbol{y}(t)$ 的分量均有无穷大方差[173],因而上述数学期望运算是无意义的.只有对稳定系统上述运算才有意义.由(5.11.4)有

$$\hat{\boldsymbol{x}}(t \mid t+N) = (\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}[\boldsymbol{\Gamma}q^{-1}\hat{\boldsymbol{w}}(t \mid t+N) + \boldsymbol{K}_N\boldsymbol{\varepsilon}(t+N)] \tag{5.11.7}$$

由节 5.6 的引理 5.6.2 和引理 5.6.3 有白噪声估值器

$$\hat{\boldsymbol{w}}(t \mid t+N) = \boldsymbol{L}_N^w(q^{-1})\boldsymbol{\varepsilon}(t+N) = \sum_{i=0}^{N}\boldsymbol{\Lambda}_i^w\boldsymbol{Q}_\varepsilon^{-1}q^{i-N}\boldsymbol{\varepsilon}(t+N) = \sum_{j=0}^{N}\boldsymbol{\Lambda}_{N-j}^w\boldsymbol{Q}_\varepsilon^{-1}q^{-j}\boldsymbol{\varepsilon}(t+N) \tag{5.11.8}$$

$$\hat{\boldsymbol{v}}(t \mid t+N) = \sum_{j=0}^{N}\boldsymbol{\Lambda}_{N-j}^v\boldsymbol{Q}_\varepsilon^{-1}q^{-j}\boldsymbol{\varepsilon}(t+N) \tag{5.11.9}$$

其中 $\boldsymbol{\Lambda}_i^w$,$\boldsymbol{\Lambda}_i^v$ 定义为

$$\boldsymbol{\Lambda}_i^w = \boldsymbol{Q}_w\boldsymbol{F}_i^{\mathrm{T}} + \boldsymbol{S}\boldsymbol{G}_i^{\mathrm{T}}, \quad \boldsymbol{\Lambda}_i^v = \boldsymbol{Q}_v\boldsymbol{G}_i^{\mathrm{T}} + \boldsymbol{S}^{\mathrm{T}}\boldsymbol{F}_i^{\mathrm{T}} \tag{5.11.10}$$

其中 $\boldsymbol{F}_i$ 和 $\boldsymbol{G}_i$ 可递推计算为

$$\begin{aligned}\boldsymbol{F}_i &= -\boldsymbol{D}_1\boldsymbol{F}_{i-1} - \cdots - \boldsymbol{D}_{n_d}\boldsymbol{F}_{i-n_d} + \boldsymbol{B}_i \\ \boldsymbol{G}_i &= -\boldsymbol{D}_1\boldsymbol{G}_{i-1} - \cdots - \boldsymbol{D}_{n_d}\boldsymbol{G}_{i-n_d} + \boldsymbol{A}_i\end{aligned} \tag{5.11.11}$$

其中规定 $\boldsymbol{F}_0 = \boldsymbol{B}_0 = \boldsymbol{0}$,$\boldsymbol{F}_i = \boldsymbol{0}(i<0)$,$\boldsymbol{G}_0 = \boldsymbol{A}_0 = \boldsymbol{I}_m$,$\boldsymbol{G}_i = \boldsymbol{0}(i<0)$,$\boldsymbol{B}_i = \boldsymbol{0}(i>n_b)$,$\boldsymbol{A}_i = \boldsymbol{0}(i>n_a)$.ARMA 新息模型为(5.10.8).

由射影理论,先取 t_0 为有限时刻,对(5.10.2)两边取射影运算,再令 $t_0 \to -\infty$ 有稳态关系

$$\boldsymbol{y}(t) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t+N) + \hat{\boldsymbol{v}}(t \mid t+N), \quad N \geqslant 0 \tag{5.11.12}$$

应用有限形式展式(5.10.33),将(5.11.7)~(5.11.9)代入上式

$$\begin{aligned}\boldsymbol{y}(t) = {}& \{\sum_{i=0}^{\beta-1}[\boldsymbol{H}\boldsymbol{\Phi}^i\boldsymbol{\Gamma}q^{-(i+1)}\sum_{j=0}^{N}\boldsymbol{\Lambda}_{N-j}^w\boldsymbol{Q}_\varepsilon^{-1}q^{-j} + \boldsymbol{H}\boldsymbol{\Phi}^iq^i\boldsymbol{K}_N] + \sum_{j=0}^{N}\boldsymbol{\Lambda}_{N-j}^w\boldsymbol{Q}_\varepsilon^{-1}q^{-j}\}\boldsymbol{\varepsilon}(t+N) + \\ & q^{-\beta}\boldsymbol{H}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\beta[\boldsymbol{\Gamma}q^{-1}\hat{\boldsymbol{w}}(t \mid t+N) + \boldsymbol{K}_N\boldsymbol{\varepsilon}(t+N)]\end{aligned} \tag{5.11.13}$$

由(5.10.37)有

$$\boldsymbol{y}(t) = \sum_{i=0}^{\beta-1}\boldsymbol{M}_iq^{-i-N}\boldsymbol{\varepsilon}(t+N) + q^{-\beta-N}\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{G}(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.11.14}$$

引入变换 $j = i+N$,并注意 $\boldsymbol{M}_i = \boldsymbol{0}(i<0)$,则上式可表为

$$\boldsymbol{y}(t) = \sum_{j=0}^{\beta-1+N}\boldsymbol{M}_{j-N}q^{-j}\boldsymbol{\varepsilon}(t+N) + q^{-\beta-N}\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{G}(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.11.15}$$

比较(5.11.13)和(5.11.15)右边第一项中含 q^{-i} 的项,并令它们相等可得关系

$$\boldsymbol{M}_{i-N}q^{-i} = \boldsymbol{H\Phi}^{i}\boldsymbol{K}_{N}q^{-i} + \sum_{j=0}^{i-1}\boldsymbol{H\Phi}^{i-1-j}\boldsymbol{\Gamma}q^{-(i-j)}\boldsymbol{\Lambda}_{N-j}^{w}\boldsymbol{Q}_{\varepsilon}^{-1}q^{-j} + \boldsymbol{\Lambda}_{N-i}^{v}q^{-i},\quad i = 1,2,\cdots,\beta-1 \tag{5.11.16}$$

且当 $i = 0$ 时有关系

$$\boldsymbol{M}_{-N}q^{0} = \boldsymbol{HK}_{N}q^{0} + \boldsymbol{\Lambda}_{N}^{v}\boldsymbol{Q}_{\varepsilon}^{-1}q^{0} \tag{5.11.17}$$

合并上述关系式有

$$\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{\beta-1} \end{bmatrix}\boldsymbol{K}_{N} = \begin{bmatrix} \boldsymbol{M}_{-N} - \boldsymbol{\Lambda}_{N}^{v}\boldsymbol{Q}_{\varepsilon}^{-1} \\ \boldsymbol{M}_{1-N} - \boldsymbol{H\Gamma\Lambda}_{N}^{w}\boldsymbol{Q}_{\varepsilon}^{-1} - \boldsymbol{\Lambda}_{N-1}^{v}\boldsymbol{Q}_{\varepsilon}^{-1} \\ \vdots \\ \boldsymbol{M}_{\beta-1-N} - \sum_{j=0}^{\beta-2}\boldsymbol{H\Phi}^{\beta-2-j}\boldsymbol{\Gamma\Lambda}_{N-j}^{w}\boldsymbol{Q}_{\varepsilon}^{-1} - \boldsymbol{\Lambda}_{N-\beta+1}^{v}\boldsymbol{Q}_{\varepsilon}^{-1} \end{bmatrix} \tag{5.11.18}$$

由此引出如下定理.

【定理 5.11.1】 系统(5.10.1)和(5.10.2)在假设 1 ~ 4 下,有非递推稳态最优 Kalman 平滑器 $\hat{\boldsymbol{x}}(t \mid t+N)(N \geqslant 0)$ 为

$$\hat{\boldsymbol{x}}(t \mid t+N) = \hat{\boldsymbol{x}}(t \mid t-1) + \sum_{i=0}^{N}\boldsymbol{K}_{i}\boldsymbol{\varepsilon}(t+i) \tag{5.11.19}$$

其中 $N = 0$ 时,$\boldsymbol{K}_{0} = \boldsymbol{K}_{f}$ 为稳态 Kalman 滤波器增益.有稳态最优递推 Kalman 平滑器为

$$\hat{\boldsymbol{x}}(t \mid t+N) = \hat{\boldsymbol{x}}(t \mid t+N-1) + \boldsymbol{K}_{N}\boldsymbol{\varepsilon}(t+N) \tag{5.11.20}$$

稳态 Kalman 平滑增益 $\boldsymbol{K}_{N}$ 为

$$\boldsymbol{K}_{N} = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{\beta-1} \end{bmatrix}^{+}\begin{bmatrix} \boldsymbol{M}_{-N} - \boldsymbol{\Lambda}_{N}^{v}\boldsymbol{Q}_{\varepsilon}^{-1} \\ \boldsymbol{M}_{1-N} - \boldsymbol{H\Gamma\Lambda}_{N}^{w}\boldsymbol{Q}_{\varepsilon}^{-1} - \boldsymbol{\Lambda}_{N-1}^{v}\boldsymbol{Q}_{\varepsilon}^{-1} \\ \vdots \\ \boldsymbol{M}_{\beta-1-N} - \sum_{j=0}^{\beta-2}\boldsymbol{H\Phi}^{\beta-2-j}\boldsymbol{\Gamma\Lambda}_{N-j}^{w}\boldsymbol{Q}_{\varepsilon}^{-1} - \boldsymbol{\Lambda}_{N-\beta+1}^{v}\boldsymbol{Q}_{\varepsilon}^{-1} \end{bmatrix} \tag{5.11.21}$$

特别对 $N = 0$ 有稳态 Kalman 滤波器增益 $\boldsymbol{K}_{0}$ 为(5.10.11),即 $\boldsymbol{K}_{0} = \boldsymbol{K}_{f}$.稳态平滑误差 $\tilde{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t \mid t+N)$ 方差阵 $\boldsymbol{P}_{N} = \mathrm{E}[\tilde{\boldsymbol{x}}(t \mid t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t \mid t+N)]$ 有递推公式

$$\boldsymbol{P}_{N} = \boldsymbol{P}_{N-1} - \boldsymbol{K}_{N}\boldsymbol{Q}_{\varepsilon}\boldsymbol{K}_{N}^{\mathrm{T}},\quad N = 0,1,2,\cdots \tag{5.11.22}$$

带初值 $\boldsymbol{P}_{-1} = \boldsymbol{\Sigma}$,$\boldsymbol{\Sigma}$ 为由(5.10.70)的 Lyapunov 方程决定的 Kalman 预报误差方差阵.非递推公式为

$$\boldsymbol{P}_{N} = \boldsymbol{\Sigma} - \sum_{i=0}^{N}\boldsymbol{K}_{i}\boldsymbol{Q}_{\varepsilon}\boldsymbol{K}_{i}^{\mathrm{T}} \tag{5.11.23}$$

证明 由射影理论有关系

$$\begin{aligned} &\hat{\boldsymbol{x}}(t \mid t) = \hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{K}_{0}\boldsymbol{\varepsilon}(t),\quad K_{0} = K_{f} \\ &\hat{\boldsymbol{x}}(t \mid t+1) = \hat{\boldsymbol{x}}(t \mid t) + \boldsymbol{K}_{1}\boldsymbol{\varepsilon}(t+1) \\ &\quad\vdots \\ &\hat{\boldsymbol{x}}(t \mid t+N) = \hat{\boldsymbol{x}}(t \mid t+N-1) + \boldsymbol{K}_{N}\boldsymbol{\varepsilon}(t+N) \end{aligned} \tag{5.11.24}$$

这引出(5.11.19)和(5.11.20).对(5.11.18)取伪逆运算得(5.11.21).当 $N = 0$ 时,注意

(5.11.8) 和 (5.11.11)，由 $\boldsymbol{B}_0 = \boldsymbol{0}$ 引出 $\boldsymbol{F}_0 = \boldsymbol{0}$，从而 $\boldsymbol{\Lambda}_0^w = \boldsymbol{S}$，并注意 $\boldsymbol{M}_i = \boldsymbol{0}(i < 0)$，在 (5.11.18) 中置 $N = 0$ 引出 (5.10.11)，即 $\boldsymbol{K}_0 = \boldsymbol{K}_f$，其中用到了事实 $\boldsymbol{\Lambda}_i^w = \boldsymbol{0}(i < 0)$，$\boldsymbol{\Lambda}_i^v = \boldsymbol{0}(i < 0)$.

由 (5.11.19) 有估值误差

$$\tilde{\boldsymbol{x}}(t \mid t + N) = \tilde{\boldsymbol{x}}(t \mid t - 1) - \sum_{i=0}^{N} \boldsymbol{K}_i \boldsymbol{\varepsilon}(t + i) \tag{5.11.25}$$

或

$$\tilde{\boldsymbol{x}}(t \mid t - 1) = \tilde{\boldsymbol{x}}(t \mid t + N) + \sum_{i=0}^{N} \boldsymbol{K}_i \boldsymbol{\varepsilon}(t + i) \tag{5.11.26}$$

由 $\tilde{\boldsymbol{x}}(t \mid t + N)$ 不相关于 $(\boldsymbol{\varepsilon}(t + N), \cdots, \boldsymbol{\varepsilon}(t + N))$，由上式立刻得 (5.11.23) 和 (5.11.22). 注意对 $\boldsymbol{\Phi}$ 是不稳定矩阵情形，公式 (5.11.22) 和 (5.11.23) 的证明是通过先取初始时刻 t_0 为有限的，然后令 $t_0 \to -\infty$ 取极限来完成的. □

【定理 5.11.2】 系统 (5.10.1) 和 (5.10.2) 在假设 1 ~ 4 下，有稳态 Kalman 平滑器 (5.11.19) 或 (5.11.20)，其中 Kalman 平滑增益 $\boldsymbol{K}_N$ 为

$$\boldsymbol{K}_N = \boldsymbol{\Sigma}(\boldsymbol{\Psi}_p^{\mathrm{T}})^N \boldsymbol{H}^{\mathrm{T}} \boldsymbol{Q}_\varepsilon^{-1}, \quad N \geqslant 0 \tag{5.11.27}$$

$$\boldsymbol{Q}_\varepsilon = \boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{Q}_v \tag{5.11.28}$$

$$\boldsymbol{\Psi}_p = \boldsymbol{\Phi} - \boldsymbol{K}_p \boldsymbol{H} \tag{5.11.29}$$

而 $\boldsymbol{K}_p$ 和 $\boldsymbol{\Sigma}$ 由定理 5.10.3 的 (5.10.69) 和 Lyapunov 方程 (5.10.70) 计算. 平滑误差方差阵 $\boldsymbol{P}_N$ 由 (5.11.23) 计算.

证明 由定理 3.8.5，定理 5.10.3 和定理 5.11.1 得证. □

下面给出不同于 (5.11.22) 的计算平滑误差方差阵 $\boldsymbol{P}_N$ 的另外的计算公式[172]. 注意到稳态关系

$$\boldsymbol{\varepsilon}(t + i) = \boldsymbol{H}\tilde{\boldsymbol{x}}(t + i \mid t + i - 1) + \boldsymbol{v}(t + i) \tag{5.11.30}$$

$$\tilde{\boldsymbol{x}}(t + 1 \mid t) = \boldsymbol{\Psi}_p \tilde{\boldsymbol{x}}(t \mid t - 1) + \boldsymbol{\Gamma}\boldsymbol{w}(t) - \boldsymbol{K}_p \boldsymbol{v}(t) \tag{5.11.31}$$

由 (5.11.31) 用迭代法有关系

$$\tilde{\boldsymbol{x}}(t + i \mid t + i - 1) = \boldsymbol{\Psi}_p^i \tilde{\boldsymbol{x}}(t \mid t - 1) + \sum_{s=t+1}^{t+i} \boldsymbol{\Psi}_p^{t+i-s} [\boldsymbol{\Gamma}\boldsymbol{w}(s - 1) - \boldsymbol{K}_p \boldsymbol{v}(s - 1)] \tag{5.11.32}$$

将 (5.11.30) 和 (5.11.32) 代入 (5.11.25)，合并同类项有

$$\tilde{\boldsymbol{x}}(t \mid t + N) = \boldsymbol{\Psi}_N \tilde{\boldsymbol{x}}(t \mid t - 1) + \sum_{i=0}^{N-1} \boldsymbol{K}_i^w \boldsymbol{w}(t + i) + \sum_{i=0}^{N} \boldsymbol{K}_i^v \boldsymbol{v}(t + i) \tag{5.11.33}$$

$$\boldsymbol{\Psi}_N = \boldsymbol{I}_n - \sum_{i=0}^{N} \boldsymbol{K}_i \boldsymbol{H} \boldsymbol{\Psi}_p^i$$

$$\boldsymbol{K}_i^w = -\sum_{j=i+1}^{N} \boldsymbol{K}_j \boldsymbol{H} \boldsymbol{\Psi}_p^{j-i-1} \boldsymbol{\Gamma}, \quad i = 0, \cdots, N - 1; \quad \boldsymbol{K}_N^w = \boldsymbol{0}$$

$$\boldsymbol{K}_i^v = -\sum_{j=i+1}^{N} \boldsymbol{K}_j \boldsymbol{H} \boldsymbol{\Psi}_p^{j-i-1} \boldsymbol{K}_p - \boldsymbol{K}_i, \quad i = 0, \cdots, N - 1; \quad \boldsymbol{K}_N^v = -\boldsymbol{K}_N \tag{5.11.34}$$

于是 (5.11.33) 可表为

$$\tilde{\boldsymbol{x}}(t \mid t + N) = \boldsymbol{\Psi}_N \tilde{\boldsymbol{x}}(t \mid t - 1) + \sum_{i=0}^{N} [\boldsymbol{K}_i^w, \boldsymbol{K}_i^v] \begin{bmatrix} \boldsymbol{w}(t + i) \\ \boldsymbol{v}(t + i) \end{bmatrix} \tag{5.11.35}$$

因 $\tilde{\boldsymbol{x}}(t \mid t-1)$ 不相关于 $\boldsymbol{w}(t+i)$ 和 $\boldsymbol{v}(t+i)(i \geqslant 0)$，于是引出如下定理.

【定理 5.11.3】 系统(5.10.1)和(5.10.2)在假设 1 ~ 4 下稳态 Kalman 平滑器(5.11.19)有稳态平滑误差方差阵 $\boldsymbol{P}_N$ 为

$$\boldsymbol{P}_N = \boldsymbol{\Psi}_N \boldsymbol{\Sigma} \boldsymbol{\Psi}_N^{\mathrm{T}} + \sum_{i=0}^{N} [\boldsymbol{K}_i^w, \boldsymbol{K}_i^v] \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v \end{bmatrix} \begin{bmatrix} \boldsymbol{K}_i^{w\mathrm{T}} \\ \boldsymbol{K}_i^{v\mathrm{T}} \end{bmatrix} \tag{5.11.36}$$

证明 暂时取初始观测时刻 t_0 是有限的，则应用射影理论有非稳态平滑误差

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \tilde{\boldsymbol{x}}(t \mid t-1) - \sum_{i=0}^{N} \boldsymbol{K}(t \mid t+i) \boldsymbol{\varepsilon}(t+i) \tag{5.11.37}$$

且有当 $t_0 \to -\infty$ 时

$$\lim_{t_0 \to -\infty} \boldsymbol{K}(t \mid t+i) = \boldsymbol{K}_i \tag{5.11.38}$$

且有非稳态预报误差

$$\tilde{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{\Psi}_p(t) \tilde{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{\Gamma} \boldsymbol{w}(t) - \boldsymbol{K}_p(t) \boldsymbol{v}(t) \tag{5.11.39}$$

其中 $\boldsymbol{\Psi}_p(t) = \boldsymbol{\Phi} - \boldsymbol{K}_p(t) \boldsymbol{H}$，且

$$\lim_{t_0 \to -\infty} \boldsymbol{\Psi}_p(t) = \boldsymbol{\Psi}_p, \quad \lim_{t_0 \to -\infty} \boldsymbol{K}_p(t) = \boldsymbol{K}_p \tag{5.11.40}$$

由此可进一步导出非稳态平滑误差

$$\tilde{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{\Psi}_N(t) \tilde{\boldsymbol{x}}(t \mid t-1) + \sum_{i=0}^{N} [\boldsymbol{K}_i^w(t), \boldsymbol{K}_i^v(t)] \begin{bmatrix} \boldsymbol{w}(t+i) \\ \boldsymbol{v}(t+i) \end{bmatrix} \tag{5.11.41}$$

且有关系

$$\lim_{t_0 \to -\infty} \boldsymbol{\Psi}_N(t) = \boldsymbol{\Psi}_N, \quad \lim_{t_0 \to -\infty} \boldsymbol{K}_i^\theta(t) = \boldsymbol{K}_i^\theta, \quad \theta = w, v \tag{5.11.42}$$

由(5.11.40)引出非稳态平滑误差方差阵

$$\boldsymbol{P}_N(t) = \boldsymbol{\Psi}_N(t) \boldsymbol{\Sigma}(t) \boldsymbol{\Psi}_N^{\mathrm{T}} + \sum_{i=0}^{N} [\boldsymbol{K}_i^w(t), \boldsymbol{K}_i^v(t)] \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{S} \\ \boldsymbol{S}^{\mathrm{T}} & \boldsymbol{Q}_v \end{bmatrix} \begin{bmatrix} \boldsymbol{K}_i^{w\mathrm{T}}(t) \\ \boldsymbol{K}_i^{w\mathrm{T}}(t) \end{bmatrix} \tag{5.11.43}$$

其中 $\boldsymbol{\Sigma}(t)$ 为非稳态预报误差方差阵，且有关系

$$\lim_{t_0 \to -\infty} \boldsymbol{\Sigma}(t) = \boldsymbol{\Sigma} \tag{5.11.44}$$

令 $t_0 \to \infty$，由(5.11.42)立刻引出(5.11.36). 证毕. □

【注 5.11.1】 上述证明是对 $\boldsymbol{\Phi}$ 为不稳定矩阵，即对系统(5.10.1)和(5.10.2)是不稳定的情形而言的. 对稳定系统则可直接由稳态关系(5.11.35)，通过取数学期望 $\mathrm{E}[\tilde{x}(t \mid t+N)\tilde{x}^{\mathrm{T}}(t \mid t+N)]$ 立刻得到(5.11.36). 因为此时这个数学期望存在.

基于稳态 Kalman 平滑器(5.11.19)，我们可以导出相应的渐近稳定的 Wiener 状态平滑器. 由(5.10.67)有

$$\hat{\boldsymbol{x}}(t \mid t-1) = (\boldsymbol{I}_n - q^{-1} \boldsymbol{\Psi}_p)^{-1} \boldsymbol{K}_p \boldsymbol{y}(t-1) \tag{5.11.45}$$

将上式代入(5.10.73)有

$$\boldsymbol{y}(t) = \boldsymbol{H}(\boldsymbol{I}_n - q^{-1} \boldsymbol{\Psi}_p)^{-1} \boldsymbol{K}_p \boldsymbol{y}(t-1) + \boldsymbol{\varepsilon}(t) \tag{5.11.46}$$

这引出另一种形式的 ARMA 新息模型

$$\boldsymbol{\Lambda}(q^{-1}) \boldsymbol{y}(t) = \psi(q^{-1}) \boldsymbol{\varepsilon}(t) \tag{5.11.47}$$

$$\psi(q^{-1}) = \det(\boldsymbol{I}_n - q^{-1} \boldsymbol{\Psi}_p) \tag{5.11.48}$$

$$\boldsymbol{\Lambda}(q^{-1}) = \psi(q^{-1})\boldsymbol{I}_m - \boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p q^{-1} \tag{5.11.49}$$

类似于定理 3.9.4 有如下 Wiener 状态平滑器.

【定理 5.11.4】 系统(5.10.1)和(5.10.2)在假设 1 ~ 4 下有渐近稳定的 Wiener 状态平滑器

$$\psi(q^{-1})\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\boldsymbol{y}(t+N) \tag{5.11.50}$$

$$\boldsymbol{K}_N(q^{-1}) = \mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p q^{-1-N} + \boldsymbol{K}_s(q^{-1})\boldsymbol{\Lambda}(q^{-1}) \tag{5.11.51}$$

$$\boldsymbol{K}_s(q^{-1}) = \sum_{i=0}^{N} \boldsymbol{K}_i q^{i-N} \tag{5.11.52}$$

稳态平滑误差方差阵 $\boldsymbol{P}_N$ 由(5.11.23)计算,$\boldsymbol{K}_p$,$\boldsymbol{\Psi}_p$,$\boldsymbol{\Sigma}$ 由定理 5.10.3 计算,且 $\boldsymbol{Q}_\varepsilon$ 由(5.11.28)定义,$\boldsymbol{K}_i$,$\boldsymbol{\Psi}_p$ 由定理 5.11.2 计算.

证明 (5.11.19)可表为

$$\hat{\boldsymbol{x}}(t \mid t+N) = \hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{K}_s(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{5.11.53}$$

其中 $\boldsymbol{K}_s(q^{-1})$ 由(5.11.52)定义.由(5.11.47)有

$$\boldsymbol{\varepsilon}(t+N) = \psi^{-1}(q^{-1})\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t+N) \tag{5.11.54}$$

将上式和(5.11.45)代入(5.11.53)得(5.11.50)和(5.11.51).由 $\boldsymbol{\Psi}_p$ 为稳定矩阵引出 $\psi(q^{-1})$ 是稳定多项式,故(5.11.50)是渐近稳定的. □

【注 5.11.2】 定理 5.11.4 与定理 3.9.4 不同之处在于这里基于 ARMA 新息模型(5.10.8)用(5.10.69)和(5.10.70)计算 $\boldsymbol{K}_p$ 和 $\boldsymbol{\Sigma}$,而定理 3.9.4 用 Riccati 方程计算 $\boldsymbol{\Sigma}$ 和 $\boldsymbol{K}_p$.

【定理 5.11.5】 系统(5.10.1)和(5.10.2)在假设 1 ~ 4 下有多步稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+N \mid t) = \boldsymbol{\Phi}^{N-1}\hat{\boldsymbol{x}}(t+1 \mid t), \quad N \geqslant 1 \tag{5.11.55}$$

$$\hat{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{K}_p\boldsymbol{y}(t) \tag{5.11.56}$$

其中 $\boldsymbol{K}_p$,$\boldsymbol{\Psi}_p$ 和 $\boldsymbol{\Sigma}$ 由定理 5.10.3 计算.预报误差方差阵 $\boldsymbol{P}_N = \mathrm{E}[\tilde{\boldsymbol{x}}(t+N \mid t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+N \mid t)]$,$\tilde{\boldsymbol{x}}(t+N \mid t) = \boldsymbol{x}(t+N) - \hat{\boldsymbol{x}}(t+N \mid t)$,为

$$\boldsymbol{P}_N = \boldsymbol{\Phi}^{N-1}\boldsymbol{\Sigma}(\boldsymbol{\Phi}^{N-1})^{\mathrm{T}} + \sum_{j=2}^{N} \boldsymbol{\Phi}^{N-j}\boldsymbol{\Gamma}\boldsymbol{Q}_w\boldsymbol{\Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^{N-j})^{\mathrm{T}}, \quad N \geqslant 2 \tag{5.11.57}$$

其中 $\boldsymbol{P}_1 = \boldsymbol{\Sigma}$.

证明 由定理 3.7.8 和定理 3.8.4 得证. □

【注 5.11.3】 定理 5.11.5 与定理 3.7.8 和定理 3.8.4 不同之处在于这里用定理 5.10.3 基于 ARMA 新息模型和 Lyapunov 方程计算 $\boldsymbol{K}_p$,$\boldsymbol{\Psi}_p$ 和 $\boldsymbol{\Sigma}$,而定理 3.7.8 和定理 3.8.4 基于 Riccati 方程计算 $\boldsymbol{K}_p$,$\boldsymbol{\Psi}_p$ 和 $\boldsymbol{\Sigma}$.

【定理 5.11.6】 系统(5.10.1)和(5.10.2)在假设 1 ~ 4 下有 Wiener 状态预报器

$$\psi(q^{-1})\hat{\boldsymbol{x}}(t+N \mid t) = \boldsymbol{K}_N(q^{-1})\boldsymbol{y}(t), \quad N \geqslant 2 \tag{5.11.58}$$

$$\boldsymbol{K}_N(q^{-1}) = \boldsymbol{\Phi}^{N-1}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p \tag{5.11.59}$$

其中 $\boldsymbol{K}_p$,$\boldsymbol{\Psi}_p$ 由定理 5.10.3 计算.相应的预报误差方差阵 $\boldsymbol{P}_N$ 由(5.11.57)计算.

证明 将(5.11.45)代入(5.11.55)且应用(5.11.48)得证.由 $\psi(q^{-1})$ 为稳定多项式引出(5.11.43)是渐近稳定的. □

【注 5.11.4】 在定理 5.11.1 中给出的平滑增益阵 $\boldsymbol{K}_N$ 可表为

$$\boldsymbol{K}_N = \sum_{i=0}^{\beta-1} \boldsymbol{\Omega}_i\left[\boldsymbol{M}_{i-N} - \sum_{j=0}^{i-1} \boldsymbol{H}\boldsymbol{\Phi}^{i-1-j}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_{N-j}^{w}\boldsymbol{Q}_\varepsilon^{-1} - \boldsymbol{\Lambda}_{N-i}^{v}\boldsymbol{Q}_\varepsilon^{-1}\right] \tag{5.11.60}$$

其中规定 $\boldsymbol{\Phi}^i = \boldsymbol{0}(i < 0)$，且 $j \geqslant 0$（即规定当 $i = 0$ 时上式中括号内第二项为零），且定义分块矩阵 $\boldsymbol{\Omega}_i$ 为

$$\boldsymbol{\Omega}^+ = (\boldsymbol{\Omega}^{\mathrm{T}}\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^{\mathrm{T}} = [\boldsymbol{\Omega}_0, \boldsymbol{\Omega}_1, \cdots, \boldsymbol{\Omega}_{\beta-1}] \tag{5.11.61}$$

$$\boldsymbol{\Omega} = [\boldsymbol{H}^{\mathrm{T}}, (\boldsymbol{H\Phi})^{\mathrm{T}}, \cdots, (\boldsymbol{H\Phi}^{\beta-1})^{\mathrm{T}}]^{\mathrm{T}} \tag{5.11.62}$$

公式(5.11.60)相同于文献[171]中提出的公式，但证明方法是完全不同的. 文献[171]的推导证明只适用于稳定系统.

5.12 单输出系统稳态 Kalman 滤波器和预报器增益算法

本节考虑对单输入单输出系统稳态 Kalman 滤波器和预报器增益的几个特殊算法[1,171,172]，算法的推导是基于矩阵求逆的 Fadeeva 公式.

考虑单输入单输出线性时不变离散随机系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi x}(t) + \boldsymbol{\Gamma} w(t) \tag{5.12.1}$$

$$y(t) = \boldsymbol{H x}(t) + v(t) \tag{5.12.2}$$

其中状态 $\boldsymbol{x}(t) \in R^n$，观测(输出) $y(t)$，白噪声输入 $w(t)$ 和观测白噪声 $v(t)$ 均为标量，$\boldsymbol{\Phi}$，$\boldsymbol{\Gamma}$ 和 $\boldsymbol{H}$ 分别为 $n \times n$，$n \times 1$ 和 $1 \times n$ 矩阵.

【假设 1】 $w(t)$ 和 $v(t)$ 是带零均值、方差各为 σ_w^2 和 σ_v^2 相关系数为 s 的相关白噪声：

$$\mathrm{E}\left\{\begin{bmatrix} w(t) \\ v(t) \end{bmatrix}[w(j), v(j)]\right\} = \begin{bmatrix} \sigma_w^2 & s \\ s & \sigma_v^2 \end{bmatrix}\delta_{tj} \tag{5.12.3}$$

【假设 2】 $(\boldsymbol{\Phi}, \boldsymbol{H})$ 为完全可观性，$(\boldsymbol{\Phi}, \boldsymbol{\Gamma})$ 为完全可控对，且 $(\bar{\boldsymbol{\Phi}}, \boldsymbol{\Gamma}\bar{\sigma}_w)$ 为完全能稳对[173]，其中 $\bar{\boldsymbol{\Phi}} = \boldsymbol{\Phi} - \boldsymbol{JH}$，$\boldsymbol{J} = \boldsymbol{\Gamma} s/\sigma_v^2$，$\bar{\sigma}_w^2 = \sigma_w^2 - s^2/\sigma_v^2$.

【假设 2】 初始观测时刻 $t_0 = -\infty$.

【引理 5.12.1】 (矩阵求逆 Fadeeva 公式) 矩阵求逆公式：

$$(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1} = \frac{\boldsymbol{F}(q^{-1})}{A(q^{-1})} \tag{5.12.4}$$

其中 $\boldsymbol{\Phi}$ 为 $n \times n$ 矩阵，q^{-1} 为单位滞后算子，$A(q^{-1}) = \det(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})$，$\boldsymbol{F}(q^{-1}) = \operatorname{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})$，记

$$A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}, a_{n_a} \neq 0, n_a \leqslant n; a_i = 0, i > n_a,$$

$$\boldsymbol{F}(q^{-1}) = \boldsymbol{I}_n + \boldsymbol{F}_1 q^{-1} + \cdots + \boldsymbol{F}_{n-1} q^{-(n-1)} \tag{5.12.5}$$

则系数 a_i 和系数阵 $\boldsymbol{F}_i$ 可递推计算为

$$a_i = -(1/i)\operatorname{trace}(\boldsymbol{\Phi F}_{i-1}), \quad i = 1, \cdots, n,$$

$$\boldsymbol{F}_i = \boldsymbol{\Phi F}_{i-1} + a_i \boldsymbol{I}_n, \quad i = 1, \cdots, n-1; \quad \boldsymbol{F}_0 = \boldsymbol{I}_n \tag{5.12.6}$$

由(5.12.1)和(5.12.2)有

$$y(t) = \boldsymbol{H}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma} q^{-1} w(t) + v(t) \tag{5.12.7}$$

将(5.12.4)代入(5.12.7)有

$$A(q^{-1})y(t) = \boldsymbol{HF}(q^{-1})\boldsymbol{\Gamma} q^{-1} w(t) + A(q^{-1})v(t) \tag{5.12.8}$$

由假设 2 引出 $(A(q^{-1}), \boldsymbol{HF}(q^{-1})\boldsymbol{\Gamma})$ 互质，于是由谱分解定理[1]有 ARMA 新息模型

$$A(q^{-1})y(t) = D(q^{-1})\varepsilon(t) \tag{5.12.9}$$

其中 $D(q^{-1}) = 1 + d_1 q^{-1} + \cdots + d_{n_d}$ 是稳定的，且(5.12.8)右端的随机过程的相关函数 $R(i)$ 在 $i = n_d$ 处截尾，新息 $\varepsilon(t)$ 是零均值、方差为 σ_ε^2 的白噪声，且有关系

$$D(q^{-1})\varepsilon(t) = \boldsymbol{HF}(q^{-1})\boldsymbol{\Gamma} q^{-1} w(t) + A(q^{-1})v(t) \tag{5.12.10}$$

另一方面由假设 2 和(5.10.72) ~ (5.10.74)存在稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{K}_p \varepsilon(t) \tag{5.12.11}$$

$$y(t) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t-1) + \varepsilon(t) \tag{5.12.12}$$

这引出

$$y(t) = \boldsymbol{H}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{K}_p q^{-1}\varepsilon(t) + \varepsilon(t) \tag{5.12.13}$$

将(5.12.4)代入上式有

$$A(q^{-1})y(t) = \boldsymbol{HF}(q^{-1})\boldsymbol{K}_p q^{-1}\varepsilon(t) + A(q^{-1})\varepsilon(t) \tag{5.12.14}$$

比较上式与(5.12.9)的右边 q^{-i} 系数引出关系

$$d_i = \boldsymbol{HF}_{i-1}\boldsymbol{K}_p + a_i, \quad i = 1, \cdots, n \tag{5.12.15}$$

其中规定 $a_i = 0(i > n_a)$，$d_i = 0(i > n_d)$. 上式可写成矩阵形式

$$\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{HF} \\ \vdots \\ \boldsymbol{HF}_{n-1} \end{bmatrix} \boldsymbol{K}_p = \begin{bmatrix} d_1 - a_1 \\ d_2 - a_2 \\ \vdots \\ d_n - a_n \end{bmatrix} \tag{5.12.16}$$

由假设 2 引出可观性指数 $\beta = n$，且由(5.12.6)有关系

$$\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{HF}_1 \\ \vdots \\ \boldsymbol{HF}_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ a_1 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & & & & & \vdots \\ a_{n-1} & a_{n-2} & 0 & \cdots & a_1 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{n-1} \end{bmatrix} \tag{5.12.17}$$

由$(\boldsymbol{\Phi}, \boldsymbol{H})$为完全可观时，引出

$$\operatorname{rank}\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{HF} \\ \vdots \\ \boldsymbol{HF}_{n-1} \end{bmatrix} = \operatorname{rank}\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{n-1} \end{bmatrix} = n \tag{5.12.18}$$

于是可得如下定理.

【定理 5.12.1】[2,171,172] 单输入单输出系统(5.12.1)和(5.12.2)在假设 1 ~ 3 下有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{\Psi}_p \hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{K}_p y(t) \tag{5.12.19}$$

$$\boldsymbol{\Psi}_p = \boldsymbol{\Phi} - \boldsymbol{K}_p \boldsymbol{H} \tag{5.12.20}$$

其中预报增益 $\boldsymbol{K}_p$ 为

$$\boldsymbol{K}_p = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{HF} \\ \vdots \\ \boldsymbol{HF}_{n-1} \end{bmatrix}^{-1} \begin{bmatrix} d_1 - a_1 \\ d_2 - a_2 \\ \vdots \\ d_n - a_n \end{bmatrix} \tag{5.12.21}$$

其中规定：$a_i = 0(i > n_a)$，$d_i = 0(i > n_d)$.

证明 将(5.12.12)代入(5.12.11)引出(5.12.19)和(5.12.20).对(5.12.16)取逆矩阵运算得(5.12.21)，因(5.12.18)保证了逆矩阵存在. □

【定理 5.12.2】[172] Kalman 预报器(5.12.19)误差方差阵 $\boldsymbol{\Sigma}$ 满足 Lyapunov 方程

$$\boldsymbol{\Sigma} = \boldsymbol{\Psi}_p \boldsymbol{\Sigma} \boldsymbol{\Psi}_p^{\mathrm{T}} + \boldsymbol{\Delta}_p \tag{5.12.22}$$

$$\boldsymbol{\Delta}_p = \sigma_w^2 \boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}} - s\boldsymbol{K}_p\boldsymbol{\Gamma}^{\mathrm{T}} - s\boldsymbol{\Gamma}\boldsymbol{K}_p^{\mathrm{T}} + \sigma_v^2 \boldsymbol{K}_p\boldsymbol{K}_p^{\mathrm{T}} \tag{5.12.23}$$

证明 见定理 5.10.3. □

【定理 5.12.3】[172] 单输入单输出系统(5.12.1)和(5.12.2)在假设 1 ~ 3 下有稳态 Kalman 滤波器

$$\hat{\boldsymbol{x}}(t+1 \mid t+1) = \boldsymbol{\Psi}_f \hat{\boldsymbol{x}}(t \mid t) + [\boldsymbol{I}_n - \boldsymbol{K}_f \boldsymbol{H}]\boldsymbol{J} y(t) + \boldsymbol{K}_f y(t+1) \tag{5.12.24}$$

$$\boldsymbol{\Psi}_f = [\boldsymbol{I}_n - \boldsymbol{K}_f \boldsymbol{H}]\bar{\boldsymbol{\Phi}}, \quad \bar{\boldsymbol{\Phi}} = \boldsymbol{\Phi} - \boldsymbol{J}\boldsymbol{H}, \quad \boldsymbol{J} = \boldsymbol{\Gamma} s/\sigma_v^2 \tag{5.12.25}$$

稳态滤波增益为

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{F} \\ \vdots \\ \boldsymbol{H}\boldsymbol{F}_{n-1} \end{bmatrix}^{-1} \begin{bmatrix} 1 - \sigma_v^2/\sigma_\varepsilon^2 \\ d_1 - \boldsymbol{H}\boldsymbol{\Gamma} s/\sigma_\varepsilon^2 - a_1\sigma_v^2/\sigma_\varepsilon^2 \\ \vdots \\ d_{n-1} - \boldsymbol{H}\boldsymbol{F}_{n-2}\boldsymbol{\Gamma} s/\sigma_\varepsilon^2 - a_{n-1}\sigma_v^2/\sigma_\varepsilon^2 \end{bmatrix} \tag{5.12.26}$$

其中规定：$d_i = 0(i > n_d)$，$a_i = 0(i > n_a)$.滤波误差方差阵 $\boldsymbol{P}$ 满足 Lyapunov 方程

$$\boldsymbol{P} = \boldsymbol{\Psi}_f \boldsymbol{P} \boldsymbol{\Psi}_f^{\mathrm{T}} + \boldsymbol{\Delta}_f \tag{5.12.27}$$

$$\boldsymbol{\Delta}_f = (\sigma_w^2 - s^2/\sigma_v^2)[\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}[\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]^{\mathrm{T}} + \boldsymbol{K}_f\boldsymbol{K}_f^{\mathrm{T}}\sigma_v^2 \tag{5.12.28}$$

证明 由射影性质和(5.12.1)，(5.12.2)有

$$\hat{\boldsymbol{x}}(t \mid t) = \hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{K}_f \varepsilon(t) \tag{5.12.29}$$

$$\hat{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t \mid t) + \boldsymbol{\Gamma}\hat{w}(t \mid t) \tag{5.12.30}$$

$$y(t) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t) + \hat{v}(t \mid t) \tag{5.12.31}$$

又由(5.1.90)有稳态最优白噪声滤波器

$$\hat{w}(t \mid t) = (s/\sigma_\varepsilon^2)\varepsilon(t), \quad \hat{v}(t \mid t) = (\sigma_v^2/\sigma_\varepsilon^2)\varepsilon(t) \tag{5.12.32}$$

于是可得

$$\hat{\boldsymbol{x}}(t \mid t) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t-1 \mid t-1) + \boldsymbol{\Gamma}(s/\sigma_\varepsilon^2)\varepsilon(t-1) + \boldsymbol{K}_f\varepsilon(t) \tag{5.12.33}$$

从而有

$$y(t) = \boldsymbol{H}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}[\boldsymbol{\Gamma}(s/\sigma_\varepsilon^2)\varepsilon(t-1) + \boldsymbol{K}_f\varepsilon(t)] + (\sigma_v^2/\sigma_\varepsilon^2)\varepsilon(t) \tag{5.12.34}$$

应用(5.12.4)，引出 ARMA 新息模型

$$A(q^{-1})y(t) = \boldsymbol{H}\boldsymbol{F}(q^{-1})\boldsymbol{\Gamma}(s/\sigma_\varepsilon^2)\varepsilon(t-1) + \boldsymbol{H}\boldsymbol{F}(q^{-1})\boldsymbol{K}_f\varepsilon(t) + A(q^{-1})(\sigma_v^2/\sigma_\varepsilon^2)\varepsilon(t) \tag{5.12.35}$$

将它与 ARMA 新息模型(5.12.9)比较，比较两式右边 q^{-i} 的系数可得关系

$$d_i = \boldsymbol{H}\boldsymbol{F}_{i-1}\boldsymbol{\Gamma}(s/\sigma_\varepsilon^2) + \boldsymbol{H}\boldsymbol{F}_i\boldsymbol{K}_f + a_i\sigma_v^2/\sigma_\varepsilon^2,$$

$$i = 0,1,\cdots,n-1;\ \boldsymbol{F}_{-1} = \boldsymbol{0},\ \boldsymbol{F}_0 = \boldsymbol{I}_n,\ a_0 = 1 \tag{5.12.36}$$

因当 $i = 0$ 时(5.12.35)右边 q^0 的系数只有 $\boldsymbol{H}\boldsymbol{F}_0\boldsymbol{K}_f$ 和 $a_0\sigma_v^2/\sigma_\varepsilon^2$ 两项，故在(5.12.36)中应规定，$\boldsymbol{F}_{-1} = \boldsymbol{0}$.

上式可写成矩阵形式

$$\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{F}_1 \\ \vdots \\ \boldsymbol{H}\boldsymbol{F}_{n-1} \end{bmatrix} \boldsymbol{K}_f = \begin{bmatrix} 1 - \sigma_v^2/\sigma_\varepsilon^2 \\ d_1 - \boldsymbol{H}\boldsymbol{\Gamma} s/\sigma_\varepsilon^2 - a_1\sigma_v^2/\sigma_\varepsilon^2 \\ \vdots \\ d_{n-1} - \boldsymbol{H}\boldsymbol{F}_{n-2}\boldsymbol{\Gamma} s/\sigma_\varepsilon^2 - a_{n-1}\sigma_v^2/\sigma_\varepsilon^2 \end{bmatrix} \tag{5.12.37}$$

其中规定 $d_i = 0(i > n_d)$，$a_i = 0(i > n_a)$. 可证明[1]：$n_d \leqslant n$. 于是由(5.12.37)引出(5.12.26). 由(5.10.45)和(5.10.46)得(5.12.27)和(5.12.28). □

【推论 5.12.1】 单输入单输出系统(5.12.1)和(5.12.2)在假设 1 ~ 3 下，若 $s = 0$，则稳态 Kalman 滤波器为

$$\hat{\boldsymbol{x}}(t+1 \mid t+1) = \boldsymbol{\Psi}_f \hat{\boldsymbol{x}}(t \mid t) + \boldsymbol{K}_f y(t+1) \tag{5.12.38}$$

$$\boldsymbol{\Psi}_f = [\boldsymbol{I}_n - \boldsymbol{K}_f \boldsymbol{H}]\boldsymbol{\Phi} \tag{5.12.39}$$

稳态滤波增益为

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{F}_1 \\ \vdots \\ \boldsymbol{H}\boldsymbol{F}_{n-1} \end{bmatrix}^{-1} \begin{bmatrix} 1 - \sigma_v^2/\sigma_\varepsilon^2 \\ d_1 - \sigma_v^2/\sigma_\varepsilon^2 \\ \vdots \\ d_{n-1} - a_{n-1}\sigma_v^2/\sigma_\varepsilon^2 \end{bmatrix} \tag{5.12.40}$$

滤波误差方差阵 $\boldsymbol{P}$ 满足 Lyapunov 方程

$$\boldsymbol{P} = \boldsymbol{\Psi}_f \boldsymbol{P} \boldsymbol{\Psi}_f^{\mathrm{T}} + \boldsymbol{\Delta}_f \tag{5.12.41}$$

$$\boldsymbol{\Delta}_f = \sigma_w^2 [\boldsymbol{I}_n - \boldsymbol{K}_f \boldsymbol{H}]\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}[\boldsymbol{I}_n - \boldsymbol{K}_f \boldsymbol{H}]^{\mathrm{T}} + \boldsymbol{K}_f \boldsymbol{K}_f^{\mathrm{T}} \sigma_v^2 \tag{5.12.42}$$

对单输入单输出系统，可观性指数 $\beta = n$，因而在计算滤波或预报增益时可避免求伪逆运算，化为求 $n \times n$ 矩阵的逆矩阵. 应用定理 5.10.1 和定理 5.10.3 有如下定理.

【定理 5.12.4】 单输入单输出系统(5.12.1)和(5.12.2)在假设 1 ~ 3 下有稳态 Kalman 滤波器增益 K_f 为

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{\Phi} \\ \vdots \\ \boldsymbol{H}\boldsymbol{\Phi}^{n-1} \end{bmatrix}^{-1} \begin{bmatrix} 1 - \sigma_v^2/\sigma_\varepsilon^2 \\ m_1 - \boldsymbol{H}\boldsymbol{\Gamma} s/\sigma_\varepsilon^2 \\ \vdots \\ m_{n-1} - \boldsymbol{H}\boldsymbol{\Phi}^{n-2}\boldsymbol{\Gamma} s/\sigma_\varepsilon^2 \end{bmatrix} \tag{5.12.43}$$

其中系数 m_i 可递推计算为

$$m_i = -a_1 m_{i-1} - \cdots - a_{n_a} m_{i-n_a} + d_i \tag{5.12.44}$$

其中规定 $m_i = 0(i < 0)$，$d_i = 0(i > n_d)$. 稳态 Kalman 预报器增益 $\boldsymbol{K}_p$ 为

$$\boldsymbol{K}_p = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{\Phi} \\ \vdots \\ \boldsymbol{H}\boldsymbol{\Phi}^{n-1} \end{bmatrix}^{-1} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix} \tag{5.12.45}$$

滤波误差方差阵 $\boldsymbol{P}$ 满足 Lyapunov 方程(5.12.27)和(5.12.28). 预报误差方差阵 $\boldsymbol{\Sigma}$ 满足 Lyapunov 方程(5.12.22)和(5.12.23).

特别，若 $s = 0$，则滤波增益成为

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{n-1} \end{bmatrix}^{-1} \begin{bmatrix} 1 - \sigma_v^2/\sigma_\varepsilon^2 \\ m_1 \\ \vdots \\ m_{n-1} \end{bmatrix} \tag{5.12.46}$$

相应的滤波误差方差阵 $\boldsymbol{P}$ 由(5.12.41)和(5.12.42)计算.

5.13 ARMA 新息模型与状态空间新息模型关系

考虑系统(5.10.1)和(5.10.2),在假设 1 ~ 4 下由(5.10.72)和(5.10.73)我们有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{K}_p\boldsymbol{\varepsilon}(t) \tag{5.13.1}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{\varepsilon}(t) \tag{5.13.2}$$

它描写了观测 $\boldsymbol{y}(t)$ 与新息 $\boldsymbol{\varepsilon}(t)$ 之间的关系,称为状态空间新息模型.非稳态新息 $\boldsymbol{\varepsilon}(t)$ 可由上两式取初值 $\hat{\boldsymbol{x}}(0 \mid -1)$ 后递推计算.现在我们研究它与由(5.10.8)给出的 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.13.3}$$

之间有何关系?注意,因为 $\boldsymbol{\varepsilon}(t)$ 是 $\boldsymbol{y}(t)$ 的稳态新息过程,因而上述两种新息模型具有相同的稳态新息 $\boldsymbol{\varepsilon}(t)$.非稳态新息 $\boldsymbol{\varepsilon}(t)$ 也可由 ARMA 新息模型(5.13.3)取初值($\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(n_d-1)$)后递推计算.用上述两种模型计算的非稳态新息 $\boldsymbol{\varepsilon}(t)$ 有何对应关系?初值 $\hat{\boldsymbol{x}}(0 \mid -1)$ 和($\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(n_d-1)$)有何对应关系?为此,我们首先研究系统可观性指数 β 与 ARMA 新息模型的 MA 多项式矩阵 $\boldsymbol{D}(q^{-1})$ 的阶次 n_d 有何关系?

【定理 5.13.1】 系统(5.10.1)和(5.10.2)的可观性指数 $\beta(\beta \leqslant n)$ 与 ARMA 新息模型的 MA 多项式矩阵 $\boldsymbol{D}(q^{-1})$ 的阶次 n_d 有关系

$$n_d \leqslant \beta \tag{5.13.4}$$

证明 用反证法:设 $n_d > \beta$ 由(5.13.1)和(5.13.2)迭代可得关系

$$\begin{aligned} \boldsymbol{y}(0) &= \boldsymbol{H}\hat{\boldsymbol{x}}(0 \mid -1) + \boldsymbol{\varepsilon}(0), \\ \boldsymbol{y}(1) &= \boldsymbol{H\Phi}\hat{\boldsymbol{x}}(0 \mid -1) + \boldsymbol{HK}_p\boldsymbol{\varepsilon}(0) + \boldsymbol{\varepsilon}(1), \\ &\vdots \\ \boldsymbol{y}(n_d-1) &= \boldsymbol{H\Phi}^{n_d-1}\hat{\boldsymbol{x}}(0 \mid -1) + \sum_{i=0}^{n_d-2}\boldsymbol{H\Phi}^{n_d-2-i}\boldsymbol{K}_p\boldsymbol{\varepsilon}(i) + \boldsymbol{\varepsilon}(n_d-1) \end{aligned} \tag{5.13.5}$$

这引出初值 $\hat{\boldsymbol{x}}(0 \mid -1)$ 与初值($\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(n_d-1)$)有关系

$$\hat{\boldsymbol{x}}(0 \mid -1) = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{\beta-1} \\ \vdots \\ \boldsymbol{H\Phi}^{n_d-1} \end{bmatrix}^{+} \begin{bmatrix} \boldsymbol{y}(0) - \boldsymbol{\varepsilon}(0) \\ \boldsymbol{y}(1) - \boldsymbol{HK}_p\boldsymbol{\varepsilon}(0) - \boldsymbol{\varepsilon}(1) \\ \vdots \\ \boldsymbol{y}(\beta-1) - \sum_{i=0}^{\beta-2}\boldsymbol{H\Phi}^{\beta-2-i}\boldsymbol{K}_p\boldsymbol{\varepsilon}(i) - \boldsymbol{\varepsilon}(\beta-1) \\ \vdots \\ \boldsymbol{y}(n_d-1) - \sum_{i=0}^{n_d-2}\boldsymbol{H\Phi}^{n_d-2-i}\boldsymbol{K}_p\boldsymbol{\varepsilon}(i) - \boldsymbol{\varepsilon}(n_d-1) \end{bmatrix} \tag{5.13.6}$$

因$(\boldsymbol{\Phi},\boldsymbol{H})$为完全可观对，故上式中伪逆$\boldsymbol{M}^{+}=(\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M})^{-1}\boldsymbol{M}^{\mathrm{T}}$存在.因而(5.13.1)和(5.13.2)在初值$\hat{\boldsymbol{x}}(0\mid-1)$由(5.13.6)定义下所得非稳态新息$\boldsymbol{\varepsilon}(t)$恒同于(5.13.3)在初值$(\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(n_d-1))$下所得非稳态新息，即有恒同的$\boldsymbol{\varepsilon}(t)$，$t=0,\cdots,n_d-1,n_d,\cdots$.

任取(5.13.3)的两组不同的值

$$(\boldsymbol{\varepsilon}(0),\cdots,\boldsymbol{\varepsilon}(\beta-1);\boldsymbol{\varepsilon}^{(i)}(\beta),\cdots,\boldsymbol{\varepsilon}^{(i)}(n_d-1)),\quad i=1,2 \tag{5.13.7}$$

其中前β个分量相同，但后$(n_d-\beta)$个分量不全相同，即$(\boldsymbol{\varepsilon}^{(1)}(\beta),\cdots,\boldsymbol{\varepsilon}^{(1)}(n_d-1))\neq(\boldsymbol{\varepsilon}^{(2)}(\beta),\cdots,\boldsymbol{\varepsilon}^{(2)}(n_d-1))$，则在初值(5.13.6)下由(5.13.1)和(5.13.2)或在初值(5.13.7)下由(5.13.3)引出两个不同的非稳态新息序列$\boldsymbol{\varepsilon}^{(1)}(t)$和$\boldsymbol{\varepsilon}^{(2)}(t)$，即$\boldsymbol{\varepsilon}^{(1)}(t)\neq\boldsymbol{\varepsilon}^{(2)}(t)$.

另一方面，注意β为系统的可观性指数，因而在假设$n_d>\beta$下由(5.13.6)计算$\hat{\boldsymbol{x}}(0\mid-1)$与分量$(\boldsymbol{\varepsilon}(\beta),\cdots,\boldsymbol{\varepsilon}(n_d-1))$无关，即与初值(5.13.7)相应的初值$\hat{\boldsymbol{x}}^{(i)}(0\mid-1)$，$i=1,2$，为

$$\hat{\boldsymbol{x}}^{(1)}(0\mid-1)=\hat{\boldsymbol{x}}^{(2)}(0\mid-1)=\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{H\Phi}\\ \vdots\\ \boldsymbol{H\Phi}^{\beta-1}\end{bmatrix}^{+}\begin{bmatrix}\boldsymbol{y}(0)-\boldsymbol{\varepsilon}(0)\\ \boldsymbol{y}(1)-\boldsymbol{HK}_p\boldsymbol{\varepsilon}(0)-\boldsymbol{\varepsilon}(1)\\ \vdots\\ \boldsymbol{y}(\beta-1)-\sum\limits_{i=0}^{\beta-2}\boldsymbol{H\Phi}^{\beta-2-i}\boldsymbol{K}_p\boldsymbol{\varepsilon}(i)-\boldsymbol{\varepsilon}(\beta-1)\end{bmatrix} \tag{5.13.8}$$

这引出$\boldsymbol{\varepsilon}^{(1)}(t)\equiv\boldsymbol{\varepsilon}^{(2)}(t)$.这与上面推导的$\boldsymbol{\varepsilon}^{(1)}(t)\not\equiv\boldsymbol{\varepsilon}^{(2)}(t)$矛盾.故$n_d\leqslant\beta_0$.证毕. □

【注 5.13.1】 这个定理的重要意义在于：用谱分解基于 Gevers – Wouters 算法构造 ARMA 新息模型时，必须对有关多项式矩阵进行左素分解，否则可导致$n_d>\beta$，即构造了错误的 ARMA 新息模型.这将引出错误的或次优的滤波结果.

下面我们用稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1\mid t)=\boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t\mid t-1)+\boldsymbol{K}_p\boldsymbol{y}(t) \tag{5.13.9}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H} \tag{5.13.10}$$

的渐近稳定性，即$\boldsymbol{\Psi}_p$是一个稳定矩阵(即$\boldsymbol{\Psi}_p$的所有特征值位于单位圆内)，来证明(5.13.3)中$\boldsymbol{D}(q^{-1})$是稳定的多项式矩阵(即$\det\boldsymbol{D}(q^{-1})$是一个稳定的多项式，即$\det\boldsymbol{D}(x)$的零点全位于单位圆外).

【定理 5.13.2】 用谱分解和 Gevers – Wouters 算法构造的 ARMA 新息模型中的 MA 多项式矩阵$\boldsymbol{D}(q^{-1})$必须是稳定的.这个条件等价于稳态 Kalman 预报器(5.13.9)的渐近稳定性条件：$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H}$是一个稳定矩阵.

证明 任取(5.13.3)的两组初值$(\boldsymbol{\varepsilon}^{(i)}(0),\cdots,\boldsymbol{\varepsilon}^{(i)}(n_d-1))$，$i=1,2$，应用(5.13.6)可求得相应的初值$\hat{\boldsymbol{x}}^{(1)}(0\mid-1)$和$\hat{\boldsymbol{x}}^{(2)}(0\mid-1)$.因由定理 5.13.1 有$n_d\leqslant\beta$，故当$n_d<\beta$时，应由(5.13.3)由$(\boldsymbol{\varepsilon}^{(i)}(0),\cdots,\boldsymbol{\varepsilon}^{(i)}(n_d-1))$递推计算$(\boldsymbol{\varepsilon}^{(i)}(n_d),\cdots,\boldsymbol{\varepsilon}^{(i)}(\beta-1))$.于是相应的$\hat{\boldsymbol{x}}^{(1)}(t+1\mid t)$和$\hat{\boldsymbol{x}}^{(2)}(t+1\mid t)$满足(5.13.9).这引出$\boldsymbol{\delta}(t)=\hat{\boldsymbol{x}}^{(1)}(t\mid t-1)-\hat{\boldsymbol{x}}^{(2)}(t\mid t-1)$满足

$$\boldsymbol{\delta}(t+1)=\boldsymbol{\Psi}_p\boldsymbol{\delta}(t) \tag{5.13.11}$$

因 Kalman 滤波理论已证明[173]：$\boldsymbol{\Psi}$ 是一个稳定矩阵，故有 $\boldsymbol{\delta}(t)\to\boldsymbol{0}(t\to\infty)$. 从而由(5.13.2)引出相应的非稳态新息 $\boldsymbol{\varepsilon}^{(1)}(t)$ 与 $\boldsymbol{\varepsilon}^{(2)}(t)$ 之差 $\boldsymbol{e}(t)=\boldsymbol{\varepsilon}^{(1)}(t)-\boldsymbol{\varepsilon}^{(1)}(t)$ 有关系

$$\boldsymbol{e}(t)=-\boldsymbol{H}\boldsymbol{\delta}(t)\to\boldsymbol{0}\quad(t\to\infty)\tag{5.13.12}$$

注意关系

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}^{(i)}(t),\quad i=1,2\tag{5.13.13}$$

这引出差分方程

$$\boldsymbol{D}(q^{-1})\boldsymbol{e}(t)=\boldsymbol{0}\tag{5.13.14}$$

(5.13.14)有状态空间模型

$$\boldsymbol{x}(t+1)=\boldsymbol{P}\boldsymbol{x}(t)\tag{5.13.15}$$

其中定义

$$\boldsymbol{x}(t)=[\boldsymbol{e}^{\mathrm{T}}(t),\boldsymbol{e}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{e}^{\mathrm{T}}(t-n_d+1)]^{\mathrm{T}},$$

$$\boldsymbol{P}=\begin{bmatrix}-\boldsymbol{D}_1 & -\boldsymbol{D}_2 & \cdots & -\boldsymbol{D}_{n_d}\\ & & & \boldsymbol{0}\\ & \boldsymbol{I}_{(n_d-1)m} & & \vdots\\ & & & \boldsymbol{0}\end{bmatrix}\tag{5.13.16}$$

(5.13.15)由初值

$$\boldsymbol{x}(n_d-1)=[\boldsymbol{e}^{\mathrm{T}}(n_d-1),\cdots,\boldsymbol{e}^{\mathrm{T}}(0)]^{\mathrm{T}}\tag{5.13.17}$$

决定. 由初值$(\boldsymbol{\varepsilon}^{(i)}(0),\cdots,\boldsymbol{\varepsilon}^{(i)}(n_d-1))$的任意性引出初值 $\boldsymbol{x}(n_d-1)$ 的任意性. 注意对任意初值 $\boldsymbol{x}(n_d-1)$ 有

$$\boldsymbol{x}(t)\to\boldsymbol{0}\quad(t\to\infty)\tag{5.13.18}$$

这是因为由(5.13.12)有 $\boldsymbol{x}(t)$ 的每个分量 $\boldsymbol{e}^{\mathrm{T}}(t-i)\to\boldsymbol{0}(t\to\infty)$, $i=0,\cdots,n_d-1$. 因此系统(5.13.15)是全局渐近稳定的. 这引出[184]$\boldsymbol{P}$ 是一个稳定矩阵. 应用定理 1.6.2 的(1.6.110)有关系

$$\det(\boldsymbol{I}_{n_dm}-q^{-1}\boldsymbol{P}^{\mathrm{T}})=\det\boldsymbol{D}(q^{-1})\tag{5.13.19}$$

其中 $\boldsymbol{D}(q^{-1})=\boldsymbol{I}_m+\boldsymbol{D}_1q^{-1}+\cdots+\boldsymbol{D}_{n_d}q^{-n_d}$. 因 $\boldsymbol{P}$ 为稳定矩阵，则 $\boldsymbol{P}^{\mathrm{T}}$ 也为稳定矩阵，从而由(5.13.19)引出 $\boldsymbol{D}(q^{-1})$ 为稳定的多项式矩阵. 证毕. □

5.14 基于 ARMA 新息模型与基于 Riccati 方程的稳态 Kalman 滤波器的功能等价性

【定理 5.14.1】 基于 ARMA 新息模型的稳态 Kalman 滤波器或预报器与基于 Riccati 方程的稳态 Kalman 滤波器或预报器是功能等价的，即用两种方法计算的稳态 Kalman 滤波器增益是相同的，用两种方法计算的稳态 Kalman 预报器增益也是相同的.

证明 回忆定理 5.10.1 和定理 5.10.3 的推导过程，推导是基于状态空间 Kalman 滤波器和预报器与时间序列 ARMA 新息模型的有限展式关系来完成的. 若在稳态 Kalman 滤波器或预报器中的稳态 Kalman 滤波器增益或预报器增益是基于 Riccati 方程计算的，则定

理5.10.1和定理5.10.3的(5.10.11)和(5.10.69)恰好证明了它们在数值上与基于ARMA新息模型由(5.10.11)或(5.10.69)给出的公式的计算结果相同. □

【例5.14.1】 考虑二维随机系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}0 & 1.5\\0 & 1\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\1\end{bmatrix}w(t) \tag{5.14.1}$$

$$y(t)=\begin{bmatrix}1 & 0\end{bmatrix}\boldsymbol{x}(t)+v(t) \tag{5.14.2}$$

其中 $w(t)$ 和 $v(t)$ 是零均值、方差各为 $\sigma_w^2=1$ 和 $\sigma_v^2=1$ 的独立白噪声.下面用两种方法求稳态 Kalman 预报器增益.注意

$$\mathrm{rank}\begin{bmatrix}\boldsymbol{H}\\\boldsymbol{H\Phi}\end{bmatrix}=\mathrm{rank}\begin{bmatrix}1 & 0\\0 & 1.5\end{bmatrix}=2,$$

$$\mathrm{rank}[\boldsymbol{\Gamma},\boldsymbol{\Phi\Gamma}]=\mathrm{rank}\begin{bmatrix}1 & 1.5\\1 & 1\end{bmatrix}=2 \tag{5.14.3}$$

故系统是完全可观、完全可控的.因而存在稳态Kalman预报器.求稳态Kalman预报器增益 $\boldsymbol{K}_p$ 的一种方法是基于 Riccati 方程(3.6.36),即

$$\boldsymbol{\Sigma}=\boldsymbol{\Phi}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma H}(\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2)^{-1}\boldsymbol{H\Sigma}]\boldsymbol{\Phi}^{\mathrm{T}}+\sigma_w^2\boldsymbol{\Gamma\Gamma}^{\mathrm{T}} \tag{5.14.4}$$

由(3.6.39)和(3.6.45)有

$$\boldsymbol{K}_p=\boldsymbol{\Phi\Sigma H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2]^{-1} \tag{5.14.5}$$

置

$$\boldsymbol{\Sigma}=\begin{bmatrix}\sigma_{11} & \sigma_{12}\\\sigma_{12} & \sigma_{22}\end{bmatrix} \tag{5.14.6}$$

按分量写(5.14.4)引出非线性方程组

$$\begin{aligned}\sigma_{12}^2&=\sigma_{11}+1,\\\sigma_{11}&=2.25\sigma_{22}-1.25,\\\sigma_{12}&=1.5\sigma_{22}-0.5\end{aligned} \tag{5.14.7}$$

这引出一元二次方程

$$\sigma_{12}^2-1.5\sigma_{12}-0.5=0 \tag{5.14.8}$$

它有正根

$$\sigma_{12}=\frac{1.5+\sqrt{4.25}}{2}=1.780\,776\,4 \tag{5.14.9}$$

将它代入(5.14.7)可求得

$$\sigma_{22}=1.520\,517\,6,\quad \sigma_{11}=2.171\,164\,5 \tag{5.14.10}$$

于是得到 Riccati 方程解

$$\boldsymbol{\Sigma}=\begin{bmatrix}2.171\,164\,5 & 1.780\,776\,4\\1.780\,776\,4 & 1.520\,517\,6\end{bmatrix} \tag{5.14.11}$$

将它代入(5.14.5)得

$$\boldsymbol{K}_p=\begin{bmatrix}0.842\,329\,2\\0.561\,552\,8\end{bmatrix} \tag{5.14.12}$$

求增益 $\boldsymbol{K}_p$ 的另一种方法是基于 ARMA 新息模型.为此需要首先构造 ARMA 新息模

型.

由(5.14.1)和(5.14.2)有

$$y(t)=[1\quad 0]\begin{bmatrix}1 & 1.5q^{-1}\\ 0 & 1-q^{-1}\end{bmatrix}^{-1}\begin{bmatrix}1\\ 1\end{bmatrix}w(t-1)+v(t) \tag{5.14.13}$$

这引出关系

$$(1-q^{-1})y(t)=(1+0.5q^{-1})w(t-1)+(1-q^{-1})v(t) \tag{5.14.14}$$

令

$$(1+dq^{-1})\varepsilon(t)=(1+0.5q^{-1})w(t-1)+(1-q^{-1})v(t) \tag{5.14.15}$$

其中 $\varepsilon(t)$ 是零均值、方差为 σ_ε^2 的白噪声,且多项式 $(1+dq^{-1})$ 是稳定的.计算上式两边 MA 过程的相关函数,且代入 $\sigma_w^2=1$ 和 $\sigma_v^2=1$,可得关系

$$(1+d^2)\sigma_\varepsilon^2=1.25\sigma_w^2+2\sigma_v^2=3.25 \tag{5.14.16}$$

$$d\sigma_\varepsilon^2=0.5\sigma_w^2-2\sigma_v^2=0.5 \tag{5.14.17}$$

消去 σ_ε^2 引出一元二次方程

$$d^2+6.5d+1=0 \tag{5.14.18}$$

取 $|d|<1$ 的根可保证 $(1+dq^{-1})$ 是稳定的多项式,有

$$d=(-6.5+\sqrt{(6.5)^2-4})/2=-0.157\,670\,8 \tag{5.14.19}$$

于是有 ARMA 新息模型

$$(1-q^{-1})y(t)=(1-0.157\,670\,8q^{-1})\varepsilon(t) \tag{5.14.20}$$

且由(5.14.17)求得

$$\sigma_\varepsilon^2=3.1711645 \tag{5.14.21}$$

应用 Fedeeva 公式可求得

$$\boldsymbol{F}_1=\boldsymbol{\Phi}-\boldsymbol{I}_2=\begin{bmatrix}-1 & 1.5\\ 0 & 0\end{bmatrix} \tag{5.14.22}$$

应用公式(5.12.21)有稳态 Kalman 预报器增益 $\boldsymbol{K}_p$ 为

$$\boldsymbol{K}_p=\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{HF}_1\end{bmatrix}^{-1}\begin{bmatrix}d+1\\ 0\end{bmatrix} \tag{5.14.23}$$

这引出

$$\boldsymbol{K}_p=\begin{bmatrix}d+1\\ (d+1)/1.5\end{bmatrix}=\begin{bmatrix}0.842\,329\,2\\ 0.561\,552\,8\end{bmatrix} \tag{5.14.24}$$

它相同于用 Riccati 方程得到(5.14.12).

现在应用公式(5.12.45)求 $\boldsymbol{K}_p$.由(5.14.20)和(5.12.44)引出

$$m_1=1+d,\quad m_2=m_1 \tag{5.14.25}$$

从而有

$$\boldsymbol{K}_p=\begin{bmatrix}\boldsymbol{H}\\ \boldsymbol{H\Phi}\end{bmatrix}^{-1}\begin{bmatrix}m_1\\ m_2\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1/1.5\end{bmatrix}\begin{bmatrix}1+d\\ 1+d\end{bmatrix}=\begin{bmatrix}1+d\\ (1+d)/1.5\end{bmatrix}=\begin{bmatrix}0.842\,329\,2\\ 0.561\,552\,8\end{bmatrix} \tag{5.14.26}$$

它与(5.14.24)结果相同.

用两种方法得到相同的增益 $\boldsymbol{K}_p$ 的机理是用两种方法引出相同的 ARMA 新息模型.事实上由状态空间新息模型(稳态 Kalman 预报器)

$$\hat{\boldsymbol{x}}(t+1 \mid t)=\boldsymbol{\Phi}\hat{\boldsymbol{x}}(t \mid t-1)+\boldsymbol{K}_p\varepsilon(t) \tag{5.14.27}$$

$$y(t)=\boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t-1)+\varepsilon(t) \tag{5.14.28}$$

其中 $\boldsymbol{K}_p$ 由 Riccati 方程计算,即由(5.14.12)计算,我们有

$$y(t)=\boldsymbol{H}(\boldsymbol{I}_2-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{K}_p\varepsilon(t-1)+\varepsilon(t) \tag{5.14.29}$$

这引出 ARMA 新息模型为

$$(1-q^{-1})y(t)=(1-0.157\,670\,8q^{-1})\varepsilon(t) \tag{5.14.30}$$

另一方面,由谱分解[1],将两个 MA 过程用一个等价的稳定的 MA 过程表示有 ARMA 新息模型(5.14.20).可看到(5.14.30)与(5.14.20)是完全相同的.

【例 5.14.2】 考虑二维系统

$$\boldsymbol{x}(t+1)=\begin{bmatrix}0.3 & 2\\ 0 & 0.7\end{bmatrix}\boldsymbol{x}(t)+\begin{bmatrix}1\\ 0.1\end{bmatrix}w(t) \tag{5.14.31}$$

$$\boldsymbol{y}(t)=\begin{bmatrix}3 & 0\\ 2 & 5\end{bmatrix}\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{5.14.32}$$

其中 $w(t)$ 和 $\boldsymbol{v}(t)$ 是零均值、方差阵各为 σ_w^2 和 $\boldsymbol{Q}_v$ 的相互独立白噪声,且

$$\sigma_w^2=0.64,\quad \boldsymbol{Q}_v=\begin{bmatrix}1 & 0\\ 0 & 0.81\end{bmatrix} \tag{5.14.33}$$

下面用两种方法求稳态 Kalman 滤波器增益 $\boldsymbol{K}_f$.先用 Riccati 方程法求 $\boldsymbol{K}_f$.我们有

$$\boldsymbol{K}_f=\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v]^{-1} \tag{5.14.34}$$

其中 $\boldsymbol{\Sigma}$ 满足 Riccati 方程

$$\boldsymbol{\Sigma}=\boldsymbol{\Phi}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{Q}_v)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}]\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}\sigma_w^2 \tag{5.14.35}$$

可解得

$$\boldsymbol{K}_f=\begin{bmatrix}0.164\,2 & 0.169\,1\\ 0.016\,5 & 0.017\,2\end{bmatrix} \tag{5.14.36}$$

其次用 ARMA 新息模型求 $\boldsymbol{K}_f$.

对本例 $\boldsymbol{H}$ 为可逆矩阵,因而系统可观性指数 $\beta=1$.由定理5.13.1引出 ARMA 新息模型的 MA 多项式 $\boldsymbol{D}(q^{-1})$ 的阶次 $n_d=1$,而不是 $n_d=2$.由(5.14.31)和(5.14.32)有

$$\boldsymbol{y}(t)=\boldsymbol{H}(\boldsymbol{I}_2-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}w(t-1)+\boldsymbol{v}(t) \tag{5.14.37}$$

引入左素分解(伪交换)

$$\boldsymbol{H}(\boldsymbol{I}_2-q^{-1}\boldsymbol{\Phi})^{-1}=(\boldsymbol{I}_2-q^{-1}\tilde{\boldsymbol{\Phi}})^{-1}\tilde{\boldsymbol{H}} \tag{5.14.38}$$

其中 $\tilde{\boldsymbol{\Phi}}$ 和 $\tilde{\boldsymbol{H}}$ 为待定的 2×2 矩阵.这引出关系

$$(\boldsymbol{I}_2-q^{-1}\tilde{\boldsymbol{\Phi}})\boldsymbol{H}=\tilde{\boldsymbol{H}}(\boldsymbol{I}-q^{-1}\boldsymbol{\Phi}) \tag{5.14.39}$$

比较上式 q^0 和 q^{-1} 的系数阵,并注意 $\boldsymbol{H}$ 为可逆阵,可得

$$\tilde{\boldsymbol{H}}=\boldsymbol{H}=\begin{bmatrix}3 & 0\\ 2 & 5\end{bmatrix},\quad \tilde{\boldsymbol{\Phi}}=\boldsymbol{H}\boldsymbol{\Phi}\boldsymbol{H}^{-1}=\begin{bmatrix}-0.5 & 1.2\\ -0.8 & 1.5\end{bmatrix} \tag{5.14.40}$$

于是有

$$\boldsymbol{H}(\boldsymbol{I}_2 - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma} = (\boldsymbol{I}_2 - q^{-1}\widetilde{\boldsymbol{\Phi}})^{-1}\widetilde{\boldsymbol{H}}\boldsymbol{\Gamma} \tag{5.14.41}$$

将上式代入(5.14.37)可得

$$(\boldsymbol{I}_2 - q^{-1}\widetilde{\boldsymbol{\Phi}})\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{\Gamma}w(t-1) + (\boldsymbol{I}_2 - q^{-1}\widetilde{\boldsymbol{\Phi}})\boldsymbol{v}(t) \tag{5.14.42}$$

这引出 ARMA 新息模型

$$(\boldsymbol{I}_2 - q^{-1}\widetilde{\boldsymbol{\Phi}})\boldsymbol{y}(t) = (\boldsymbol{I}_2 + \boldsymbol{D}q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.14.43}$$

其中新息 $\boldsymbol{\varepsilon}(t)$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声,且有关系

$$(\boldsymbol{I}_2 + \boldsymbol{D}q^{-1})\boldsymbol{\varepsilon}(t) = \boldsymbol{H}\boldsymbol{\Gamma}w(t-1) + (\boldsymbol{I}_2 - q^{-1}\widetilde{\boldsymbol{\Phi}})\boldsymbol{v}(t) \tag{5.14.44}$$

用 Gevers – Wouters 算法可求得

$$\boldsymbol{D} = \begin{bmatrix} 0.746\,8 & -0.944\,8 \\ 1.022\,3 & -1.269\,8 \end{bmatrix}, \quad \boldsymbol{Q}_\varepsilon = \begin{bmatrix} 6.884\,7 & 4.912\,6 \\ 4.912\,6 & 4.911\,9 \end{bmatrix} \tag{5.14.45}$$

应用(5.10.11)有稳态 Kalman 滤波增益为

$$\boldsymbol{K}_f = \boldsymbol{H}^{-1}[\boldsymbol{I}_2 - \boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1}] = \begin{bmatrix} 0.164\,2 & 0.199\,1 \\ 0.016\,5 & 0.017\,2 \end{bmatrix} \tag{5.14.46}$$

它相同于用 Riccati 方程所得结果(5.14.36).

【注 5.14.1】 由(5.14.37)直接用矩阵求逆有关系

$$\begin{aligned}(1-0.3q^{-1})(1-0.7q^{-1})y(t) = {} & \boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_2 - q^{-1}\boldsymbol{\Phi})\boldsymbol{\Gamma}w(t-1) + \\ & (1-0.3q^{-1})(1-0.7q^{-1})v(t)\end{aligned} \tag{5.14.47}$$

用 Gevers – Wouters 算法可得 $n_d = 2$ 的 ARMA 新息模型

$$(1 - q^{-1} + 0.21q^{-2})\boldsymbol{y}(t) = (\boldsymbol{I}_2 + \boldsymbol{D}_1q^{-1} + \boldsymbol{D}_2q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.14.48}$$

其中

$$\boldsymbol{D}_1 = \begin{bmatrix} -0.753\,2 & 0.255\,2 \\ 0.222\,3 & -0.769\,8 \end{bmatrix}, \quad \boldsymbol{D}_2 = \begin{bmatrix} 0.106\,5 & -0.106\,5 \\ -0.086\,3 & 0.120\,9 \end{bmatrix},$$

$$\boldsymbol{Q}_\varepsilon = \begin{bmatrix} 6.884\,7 & 4.912\,6 \\ 4.912\,6 & 4.911\,9 \end{bmatrix} \tag{5.14.49}$$

由于这一新息模型没有进行左素分解,因而在(5.14.48)两边有稳定的一阶左因式,消去左因式后可化为一阶 ARMA 新息模型(5.14.43).但(5.14.48)两边有稳定的左因式并不影响 $\boldsymbol{Q}_\varepsilon$ 的值,因而基于(5.14.48)也引出相同的 $\boldsymbol{K}_f$.但理论上这一2阶 ARMA 新息模型是不正确的.但有稳定左因式的 ARMA 新息模型并不影响计算结果.因此这种 ARMA 新息模型也可以应用.

【例 5.14.3】 考虑跟踪系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}\boldsymbol{x}(t) + \boldsymbol{\Gamma}\boldsymbol{w}(t) \tag{5.14.50}$$

$$y(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}(t) \tag{5.14.51}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_0 & 0.5T_0^2 \\ 0 & 1 & T_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{H} = [1 \quad 0 \quad 0] \tag{5.14.52}$$

其中 T_0 为采样周期,状态 $\boldsymbol{x}(t) = [x_1(t), x_2(t), x_3(t)]^{\mathrm{T}}$,且 $x_1(t)$, $x_2(t)$ 和 $x_3(t)$ 分别为

在时刻 tT_0 处运动目标的位置、速度和加速度，$y(t)$ 为对位置的观测，$v(t)$ 为观测噪声. 假设 $w(t)$ 和 $v(t)$ 是零均值、方差各为 σ_w^2 和 σ_v^2 的相互独立高斯白噪声. 问题是求稳态 Kalman 滤波器和预报器 $\hat{\boldsymbol{x}}(t \mid t)$ 和 $\hat{\boldsymbol{x}}(t \mid t-1)$.

首先基于 ARMA 新息模型解决这个问题. 在仿真中取 $T_0 = 0.2$，$\sigma_w^2 = 0.8$，$\sigma_v^2 = 10$. 容易得到 ARMA 新息模型为

$$(1 - 3q^{-1} + 3q^{-2} - q^{-3})y(t) = (1 + d_1q^{-1} + d_2q^{-2} + d_3q^{-3})\varepsilon(t) \tag{5.14.53}$$

$$d_1 = -2.5529, \quad d_2 = 2.2002, \quad d_3 = -0.6383, \quad \sigma_\varepsilon^2 = 15.6672 \tag{5.14.54}$$

且由(5.12.44)可求得

$$m_1 = 0.4471, \quad m_2 = 0.5415, \quad m_3 = 0.6450 \tag{5.14.55}$$

稳态 Kalman 滤波器和预报器分别为

$$\hat{\boldsymbol{x}}(t \mid t) = \boldsymbol{\Psi}_f\hat{\boldsymbol{x}}(t-1 \mid t-1) + \boldsymbol{K}_f y(t) \tag{5.14.56}$$

$$\hat{\boldsymbol{x}}(t \mid t-1) = \boldsymbol{\Psi}_p\hat{\boldsymbol{x}}(t-1 \mid t-2) + \boldsymbol{K}_p y(t-1) \tag{5.14.57}$$

其中滤波增益 $\boldsymbol{K}_f$ 和预报增益 $\boldsymbol{K}_p$ 由(5.12.43)和(5.12.45)给出为

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \boldsymbol{H\Phi}^2 \end{bmatrix}^{-1} \begin{bmatrix} 1 - \sigma_v^2/\sigma_\varepsilon^2 \\ m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} 0.3617 \\ 0.4043 \\ 0.2260 \end{bmatrix} \tag{5.14.58}$$

$$\boldsymbol{K}_p = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \boldsymbol{H\Phi}^2 \end{bmatrix}^{-1} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} 0.4471 \\ 0.4495 \\ 0.2260 \end{bmatrix} \tag{5.14.59}$$

且可求得

$$\boldsymbol{\Psi}_f = [\boldsymbol{I}_3 - \boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Phi}, \quad \boldsymbol{\Psi}_p = \boldsymbol{\Phi}[\boldsymbol{I}_3 - \boldsymbol{K}_f\boldsymbol{H}] \tag{5.14.60}$$

其中基于 Riccati 方程求 $\boldsymbol{K}_f$ 和 $\boldsymbol{K}_p$.

$$\boldsymbol{K}_f = \boldsymbol{\Sigma H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}} + \sigma_v^2]^{-1} \tag{5.14.61}$$

其中 $\boldsymbol{\Sigma}$ 满足 Riccati 方程

$$\boldsymbol{\Sigma} = \boldsymbol{\Phi}[\boldsymbol{\Sigma} - \boldsymbol{\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}} + \sigma_v^2)^{-1}\boldsymbol{H\Sigma}]\boldsymbol{\Phi}^{\mathrm{T}} + \sigma_w^2\boldsymbol{\Gamma\Gamma}^{\mathrm{T}} \tag{5.14.62}$$

可求得

$$\boldsymbol{\Sigma} = \begin{bmatrix} 5.6672 & 6.3346 & 3.5403 \\ 6.3346 & 9.9090 & 7.1188 \\ 3.5403 & 7.1188 & 7.9571 \end{bmatrix} \tag{5.14.63}$$

$$\boldsymbol{K}_f = \begin{bmatrix} 0.3617 \\ 0.4043 \\ 0.2260 \end{bmatrix} \tag{5.14.64}$$

它相同于基于 ARMA 新息模型的结果(5.14.58)，且有

$$\boldsymbol{K}_p = \boldsymbol{\Phi K}_f = \begin{bmatrix} 0.4471 \\ 0.4495 \\ 0.2260 \end{bmatrix} \tag{5.14.65}$$

它相同于(5.14.59).上述计算说明两种方法功能是等价的.

仿真结果如图5.14.1至图5.14.3所示,其中实线代表真实值 $x_i(t)$,虚线代表预报值 $\hat{x}_i(t \mid t-1)$,长划线代表滤波值 $\hat{x}_i(t \mid t)$.

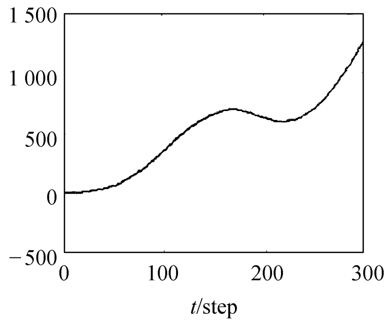


图5.14.1　位置 $x_1(t)$ 及其Kalman跟踪滤波器 $\hat{x}_1(t \mid t)$ 和预报器 $\hat{x}_1(t \mid t-1)$

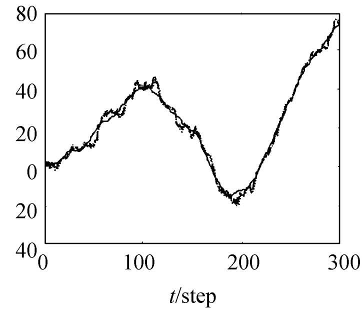


图5.14.2　速度 $x_2(t)$ 及其Kalman跟踪滤波器 $\hat{x}_2(t \mid t)$ 和预报器 $\hat{x}_2(t \mid t-1)$

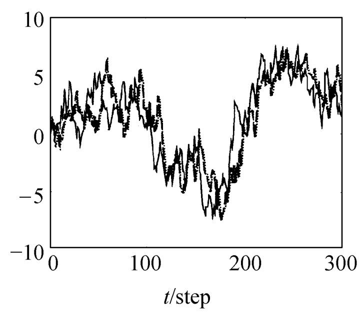


图 5.14.3　加速度 $x_3(t)$ 及其Kalman跟踪滤波器 $\hat{x}_3(t \mid t)$ 和预报器 $\hat{x}_3(t \mid t-1)$

5.15　多项式矩阵左素分解与ARMA新息模型

考虑一般随机系统(5.10.1)和(5.10.2)在假设1 ~ 4下构造ARMA新息模型时,要求对有理多项式矩阵实行左素分解

$$\boldsymbol{H}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma} = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1}) \tag{5.15.1}$$

其中 $(\boldsymbol{A}(q^{-1}), \boldsymbol{B}(q^{-1}))$ 左素,否则若 $\boldsymbol{A}(q^{-1})$ 与 $\boldsymbol{B}(q^{-1})$ 含有不稳定左因式时,将会引出错误的ARMA新息模型[185],进而将会引出错误的稳态Kalman滤波增益,导至错误的或次优的滤波结果.[39,185].

【引理5.15.1】[186]　具有相同行数的多项式矩阵 $\widetilde{\boldsymbol{A}}(q^{-1})$ 与 $\widetilde{\boldsymbol{B}}(q^{-1})$ 左素的充要条件是对一切复数 z 有多项式矩阵 $[\widetilde{\boldsymbol{A}}(z), \widetilde{\boldsymbol{B}}(z)]$ 行满秩.

【引理 5.15.2】[184] （求极大左因式方法）对多项式矩阵$[\widetilde{\boldsymbol{A}}(q^{-1}),\widetilde{\boldsymbol{B}}(q^{-1})]$实行一系列的列初等变换，即对其右乘以一个单位模阵 $\boldsymbol{U}(q^{-1})$，可将其化为

$$[\widetilde{\boldsymbol{A}}(q^{-1}),\widetilde{\boldsymbol{B}}(q^{-1})]\boldsymbol{U}(q^{-1}) = [\boldsymbol{M}(q^{-1}),\boldsymbol{0}] \tag{5.15.2}$$

则 $\boldsymbol{M}(q^{-1})$ 是 $\widetilde{\boldsymbol{A}}(q^{-1})$ 和 $\widetilde{\boldsymbol{B}}(q^{-1})$ 的极大左因式.

下面以三个不稳定系统 —— 目标跟踪系统为例，来说明如何进行左素分解及左素分解的重要性. 对不稳定系统不进行左素分解将导出错误的 ARMA 新息模型.

【例 5.15.1】 考虑带位置和速度观测的跟踪系统[185]

$$\boldsymbol{x}(t) = \boldsymbol{\Phi}\boldsymbol{x}(t) + \boldsymbol{\Gamma}w(t) \tag{5.15.3}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}(t) \tag{5.15.4}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_0 & 0.5T_0^2 \\ 0 & 1 & T_0 \\ 0 & 0 & 1 \end{bmatrix},\quad \boldsymbol{\Gamma} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},\quad \boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{5.15.5}$$

其中 T_0 为采样周期，状态 $\boldsymbol{x}(t) = [x_1(t),x_2(t),x_3(t)]^{\mathrm{T}}$，$x_1(t)$，$x_2(t)$ 和 $x_3(t)$ 表示在时刻 tT_0 处运动目标的位置、速度和加速度，观测 $\boldsymbol{y}(t) = [y_1(t),y_2(t)]^{\mathrm{T}}$，$y_1(t)$ 和 $y_2(t)$ 各为对位置和速度的观测信号，$\boldsymbol{v}(t) = [v_1(t),v_2(t)]^{\mathrm{T}}$ 为观测噪声. 设 $w(t)$ 和 $\boldsymbol{v}(t)$ 是带零均值、方差阵各为 σ_w^2 和 $\boldsymbol{Q}_v$ 的独立高斯白噪声. 问题是基于 ARMA 新息模型法和基于 Riccati 方程法分别求稳态 Kalman 滤波器，并验证它们是完全相同的.

由(5.15.3)和(5.15.4)有

$$\boldsymbol{y}(t) = \boldsymbol{H}(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}w(t-1) + \boldsymbol{v}(t) \tag{5.15.6}$$

由矩阵求逆公式有

$$(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi})^{-1} = \operatorname{adj}(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi})/\det(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi}) \tag{5.15.7}$$

注意

$$\det(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi}) = (1-q^{-1})^3,$$

$$\operatorname{adj}(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi}) = \begin{bmatrix} (1-q^{-1})^2 & T_0(1-q^{-1})q^{-1} & 0.5T_0^2(1-q^{-1})q^{-1} \\ 0 & (1-q^{-1})^2 & T_0(1-q^{-1})q^{-1} \\ 0 & 0 & (1-q^{-1})^2 \end{bmatrix} \tag{5.15.8}$$

这引出

$$\boldsymbol{H}\operatorname{adj}(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi})\boldsymbol{\Gamma} = \begin{bmatrix} 0.5T_0^2(1+q^{-1})q^{-1} \\ T_0(1-q^{-1})q^{-1} \end{bmatrix} \tag{5.15.9}$$

将(5.15.7)代入(5.15.6)并应用(5.15.9)有

$$\boldsymbol{y}(t) = \frac{1}{(1-q^{-1})^3}\begin{bmatrix} 0.5T_0^2(1+q^{-1}) \\ T_0(1-q^{-1}) \end{bmatrix}w(t-2) + \boldsymbol{v}(t) =$$

$$[(1-q^{-1})^3\boldsymbol{I}_2]^{-1}\begin{bmatrix} 0.5T_0^2(1+q^{-1}) \\ T_0(1-q^{-1}) \end{bmatrix}w(t-2) + \boldsymbol{v}(t) \tag{5.15.10}$$

注意

$$[(1-q^{-1})^3\boldsymbol{I}_2]^{-1}\begin{bmatrix} 0.5T_0^2(1+q^{-1}) \\ T_0(1-q^{-1}) \end{bmatrix} = \begin{bmatrix} (1-q^{-1})^3 & 0 \\ 0 & (1-q^{-1})^3 \end{bmatrix}^{-1}\begin{bmatrix} 0.5T_0^2(1+q^{-1}) \\ T_0(1-q^{-1}) \end{bmatrix} \tag{5.15.11}$$

为了对(5.15.11)实现左素分解，现在的问题是求如下两个多项式矩阵

$$\widetilde{\boldsymbol{A}}(q^{-1}) = \begin{bmatrix} (1-q^{-1})^3 & 0 \\ 0 & (1-q^{-1})^3 \end{bmatrix}, \quad \widetilde{\boldsymbol{B}}(q^{-1}) = \begin{bmatrix} 0.5T_0^2(1+q^{-1}) \\ T_0(1-q^{-1}) \end{bmatrix} \tag{5.15.12}$$

的极大左因式.消去极大左因式后才能实现左素分解.

由引理5.15.1，$\widetilde{\boldsymbol{A}}(q^{-1})$ 与 $\widetilde{\boldsymbol{B}}(q^{-1})$ 是非左素的，因为取 $q^{-1}=1$ 有

$$\operatorname{rank}[\widetilde{\boldsymbol{A}}(q^{-1}), \widetilde{\boldsymbol{B}}(q^{-1})] = \operatorname{rank}\begin{bmatrix} 0 & 0 & T_0^2 \\ 0 & 0 & 0 \end{bmatrix} = 1 \tag{5.15.13}$$

即 $[\widetilde{\boldsymbol{A}}(q^{-1}), \widetilde{\boldsymbol{B}}(q^{-1})]$ 非行满秩.

应用引理5.15.2，对 $[\widetilde{\boldsymbol{A}}(q^{-1}), \widetilde{\boldsymbol{B}}(q^{-1})]$ 实行如下列初等变换

$$\begin{bmatrix} (1-q^{-1})^3 & 0 & 0.5T_0^2(1+q^{-1}) \\ 0 & (1-q^{-1})^3 & T_0(1-q^{-1}) \end{bmatrix} \xrightarrow[E_a]{\text{第 3 列乘 } 1/T_0}$$

$$\begin{bmatrix} (1-q^{-1})^3 & 0 & 0.5T_0(1+q^{-1}) \\ 0 & (1-q^{-1})^3 & 1-q^{-1} \end{bmatrix} \xrightarrow[E_b]{\text{第 3 列乘 } -(1-q^{-1})^2 \text{ 加到第 2 列}}$$

$$\begin{bmatrix} (1-q^{-1})^3 & -0.5T_0(1+q^{-1})(1-q^{-1})^2 & 0.5T_0(1+q^{-1}) \\ 0 & 0 & 1-q^{-1} \end{bmatrix} \xrightarrow[E_c]{\text{第 2 列乘 } -2/T}$$

$$\begin{bmatrix} (1-q^{-1})^3 & (1-q^{-1})^2(1+q^{-1}) & 0.5T_0(1+q^{-1}) \\ 0 & 0 & 1-q^{-1} \end{bmatrix} \xrightarrow[E_d]{\text{第 1 列乘 } -1}$$

$$\begin{bmatrix} (q^{-1}-1)^3 & (1-q^{-1})^2(1+q^{-1}) & 0.5T_0(1+q^{-1}) \\ 0 & 0 & 1-q^{-1} \end{bmatrix} \xrightarrow[E_e]{\text{第 2 列乘 } -1 \text{ 加到第 1 列}}$$

$$\begin{bmatrix} -2+4q^{-1}-2q^{-2} & (1-q^{-1})^2(1+q^{-1}) & 0.5T_0(1+q^{-1}) \\ 0 & 0 & 1-q^{-1} \end{bmatrix} \xrightarrow[E_f]{\text{第 1 列乘 } -1/2}$$

$$\begin{bmatrix} (1-q^{-1})^2 & (1-q^{-1})^2(1+q^{-1}) & 0.5T_0(1+q^{-1}) \\ 0 & 0 & 1-q^{-1} \end{bmatrix} \xrightarrow[E_g]{\text{第 1 列乘 } -(1+q^{-1}) \text{ 加到第 2 列}}$$

$$\begin{bmatrix} (1-q^{-1})^2 & 0 & 0.5T_0(1+q^{-1}) \\ 0 & 0 & 1-q^{-1} \end{bmatrix} \xrightarrow[E_h]{\text{交换第 2 列和第 3 列}}$$

$$\begin{bmatrix} (1-q^{-1})^2 & 0.5T_0(1+q^{-1}) & 0 \\ 0 & 1-q^{-1} & 0 \end{bmatrix} \tag{5.15.14}$$

这引出 $\tilde{A}(q^{-1})$ 与 $\tilde{B}(q^{-1})$ 的极大左因式为

$$\boldsymbol{M}(q^{-1}) = \begin{bmatrix} (1-q^{-1})^2 & 0.5T_0(1+q^{-1}) \\ 0 & 1-q^{-1} \end{bmatrix} \tag{5.15.15}$$

相应的单位模阵 $\boldsymbol{U}(q^{-1})$ 为

$$\boldsymbol{U}(q^{-1}) = \boldsymbol{E}_a\boldsymbol{E}_b\boldsymbol{E}_c\boldsymbol{E}_d\boldsymbol{E}_e\boldsymbol{E}_f\boldsymbol{E}_g\boldsymbol{E}_h =$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/T_0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -(1-q^{-1})^2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2/T & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\times$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1-q^{-1} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 1/2 & 0 & -0.5(1+q^{-1}) \\ -1/T_0 & 0 & (-1+q^{-1})/T_0 \\ (1-q^{-1})^2/T_0^2 & 1/T_0 & (1-q^{-1})^3/T_0^2 \end{bmatrix} \tag{5.15.16}$$

即

$$[\widetilde{\boldsymbol{A}}(q^{-1}), \widetilde{\boldsymbol{B}}(q^{-1})]\boldsymbol{U}(q^{-1}) = \begin{bmatrix} (1-q^{-1})^2 & 0.5T_0(1+q^{-1}) \\ 0 & 1-q^{-1} \end{bmatrix} \tag{5.15.17}$$

于是我们有

$$\widetilde{\boldsymbol{A}}(q^{-1}) = \boldsymbol{M}(q^{-1})\overline{\boldsymbol{A}}(q^{-1}), \quad \widetilde{\boldsymbol{B}}(q^{-1}) = \boldsymbol{M}(q^{-1})\overline{\boldsymbol{B}}(q^{-1}) \tag{5.15.18}$$

注意

$$\boldsymbol{M}^{-1}(q^{-1}) = \frac{1}{(1-q^{-1})^3}\begin{bmatrix} 1-q^{-1} & -0.5T_0(1+q^{-1}) \\ 0 & (1-q^{-1})^2 \end{bmatrix} \tag{5.15.19}$$

这引出

$$\overline{\boldsymbol{A}}(q^{-1}) = \boldsymbol{M}^{-1}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) = \begin{bmatrix} 1-q^{-1} & -0.5T_0(1+q^{-1}) \\ 0 & (1-q^{-1})^2 \end{bmatrix},$$

$$\overline{\boldsymbol{B}}(q^{-1}) = \boldsymbol{M}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) = \begin{bmatrix} 0 \\ T_0 \end{bmatrix} \tag{5.15.20}$$

注意(5.15.20)$\overline{\boldsymbol{A}}(q^{-1})$ 可写为

$$\overline{\boldsymbol{A}}(q^{-1}) = \overline{\boldsymbol{A}}_0 + \overline{\boldsymbol{A}}_1 q^{-1} + \overline{\boldsymbol{A}}_2 q^{-2} =$$

$$\begin{bmatrix} 1 & -0.5T_0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & -0.5T_0 \\ 0 & -2 \end{bmatrix} q^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} q^{-2} \tag{5.15.21}$$

它的首系数阵 $\overline{A}_0$ 不是单位阵.可将其化为首系数阵为单位阵的多项式矩阵.为此引入

$$\boldsymbol{A}(q^{-1}) = \overline{\boldsymbol{A}}_0^{-1}\overline{\boldsymbol{A}}(q^{-1}), \quad \boldsymbol{B}(q^{-1}) = \overline{\boldsymbol{A}}_0^{-1}\overline{\boldsymbol{B}}(q^{-1}) \tag{5.15.22}$$

由(5.15.18)和(5.15.23),从 $\widetilde{\boldsymbol{A}}(q^{-1})$ 和 $\widetilde{\boldsymbol{B}}(q^{-1})$ 中消去极大左因式 $\boldsymbol{M}(q^{-1})$ 后有左素分解

$$\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) = \overline{\boldsymbol{A}}^{-1}(q^{-1})\overline{\boldsymbol{B}}(q^{-1}) = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1}) \tag{5.15.23}$$

其中 $\boldsymbol{A}(q^{-1})$ 的首系数阵 $\boldsymbol{A}_0 = \boldsymbol{I}_2$ 为单位阵.由上两式可得

$$\boldsymbol{A}(q^{-1}) = \boldsymbol{I}_2 + \begin{bmatrix} -1 & -1.5T_0 \\ 0 & -2 \end{bmatrix} q^{-1} + \begin{bmatrix} 0 & 0.5T_0 \\ 0 & 1 \end{bmatrix} q^{-2}, \quad \boldsymbol{B}(q^{-1}) = \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix} \tag{5.15.24}$$

由(5.15.9) ~ (5.15.11)和(5.15.24)有

$$\boldsymbol{y}(t) = \widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1})w(t-2) + \boldsymbol{v}(t) = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1})w(t-2) + \boldsymbol{v}(t) \tag{5.15.25}$$

这引出关系

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{B}(q^{-1})w(t-2) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.15.26}$$

应用 Gevers - Wouters 算法可得 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.15.27}$$

其中 $\boldsymbol{D}(q^{-1}) = \boldsymbol{I}_2 + \boldsymbol{D}_1 q^{-1} + \boldsymbol{D}_2 q^{-1}$ 是稳定的，新息 $\boldsymbol{\varepsilon}(t)$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \boldsymbol{B}(q^{-1})w(t-2) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.15.28}$$

在仿真中取

$$T_0 = 0.1,\quad \sigma_w^2 = 4,\quad \boldsymbol{Q}_v = \begin{bmatrix} 0.625 & 0 \\ 0 & 0.25 \end{bmatrix} \tag{5.15.29}$$

由 (5.15.24) 有

$$\boldsymbol{A}(q^{-1}) = \boldsymbol{I}_2 + \begin{bmatrix} -1 & -0.15 \\ 0 & -2 \end{bmatrix} q^{-1} + \begin{bmatrix} 0 & 0.05 \\ 0 & 1 \end{bmatrix} q^{-2} \tag{5.15.30}$$

且由 Gevers - Wouters 算法可得

$$\boldsymbol{D}_1 = \begin{bmatrix} -0.937\,5 & -0.011\,152 \\ 0.023\,767 & -1.151\,1 \end{bmatrix},\quad \boldsymbol{D}_2 = \begin{bmatrix} -0.001\,333\,4 & -0.020\,267 \\ -0.026\,668 & 0.405\,34 \end{bmatrix},$$

$$\boldsymbol{Q}_\varepsilon = \begin{bmatrix} 0.668 & 0.043\,948 \\ 0.043\,948 & 0.619\,65 \end{bmatrix} \tag{5.15.31}$$

注意对本例由 (5.15.4) 定义的 $\boldsymbol{H}$ 和 $\boldsymbol{\Phi}$，易验证

$$\mathrm{rank}\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \end{bmatrix} = 3 \tag{5.15.32}$$

因此系统的可观性指数 $\beta = 2$. 应用推论 5.10.1 有稳态 Kalman 跟踪滤波器

$$\hat{\boldsymbol{x}}(t+1 \mid t+1) = \boldsymbol{\Psi}_f \hat{\boldsymbol{x}}(t \mid t) + \boldsymbol{K}_f \boldsymbol{y}(t) \tag{5.15.33}$$

$$\boldsymbol{\Psi}_f = [\boldsymbol{I}_3 - \boldsymbol{K}_f \boldsymbol{H}]\boldsymbol{\Phi} \tag{5.15.34}$$

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \end{bmatrix}^{+} \begin{bmatrix} \boldsymbol{I}_2 - \boldsymbol{Q}_v \boldsymbol{Q}_\varepsilon^{-1} \\ \boldsymbol{M}_1 \end{bmatrix} \tag{5.15.35}$$

$$\boldsymbol{M}_1 = -\boldsymbol{A}_1 \boldsymbol{M}_0 + \boldsymbol{D}_1,\quad \boldsymbol{M}_0 = \boldsymbol{I}_2 \tag{5.15.36}$$

由 (5.15.35) 和 (5.15.36) 可得滤波增益

$$\boldsymbol{K}_f = \begin{bmatrix} 0.059\,981\,749\,5 & 0.066\,669\,727\,8 \\ 0.026\,667\,891\,1 & 0.594\,657\,991\,2 \\ -0.029\,013\,235\,6 & 2.542\,590\,489\,4 \end{bmatrix} \tag{5.15.37}$$

注意 (5.15.15) 极大左因式 $\boldsymbol{M}(q^{-1})$ 是不稳定的多项式矩阵，$\det\boldsymbol{M}(q^{-1})$ 有在单位圆上的零点 $q^{-1} = 1$. 若不消去极大左因式 $\boldsymbol{M}(q^{-1})$ 构造 ARMA 新息模型是错误的，将引出错误的稳态 Kalman 滤波增益和错误的或次优的滤波结果[39]. 此时有 $A(q^{-1}) = (1-q^{-1})^3$，可求得

$$\boldsymbol{K}_f = \begin{bmatrix} 0.289\,155\,828\,6 & -0.020\,759\,714\,2 \\ 0.038\,126\,595\,1 & 0.590\,286\,519\,1 \\ -0.266\,678\,909\,8 & 9.832\,100\,832\,0 \end{bmatrix} \tag{5.15.38}$$

它不同于 (5.15.37).

另一方面，由第三章有基于 Riccati 方程的稳态 Kalman 滤波器增益为

$$\boldsymbol{K}_f = \boldsymbol{\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}} + \boldsymbol{Q}_v)^{-1} \tag{5.15.39}$$

$\boldsymbol{\Sigma}$ 满足 Riccati 方程

$$\boldsymbol{\Sigma} = \boldsymbol{\Phi}[\boldsymbol{\Sigma\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}} + \boldsymbol{Q}_v)^{-1}\boldsymbol{H\Sigma}]\boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{\Gamma\Gamma}^{\mathrm{T}}\sigma_w^2 \tag{5.15.40}$$

可求得

$$\boldsymbol{K}_f = \begin{bmatrix} 0.059\,981\,749\,5 & 0.066\,669\,727\,8 \\ 0.026\,667\,891\,1 & 0.594\,657\,991\,2 \\ -0.029\,013\,235\,6 & 2.542\,590\,489\,4 \end{bmatrix} \tag{5.15.41}$$

它完全相同于基于 ARMA 新息模型的由(5.15.37)给出的 $\boldsymbol{K}_f$.因此两种求稳态 Kalman 滤波器增益方法是功能等价的.

【例 5.15.2】 两传感器集中式观测融合 Kalman 跟踪滤波器.考虑两传感器跟踪系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi x}(t) + \boldsymbol{\Gamma}w(t) \tag{5.15.42}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{H}_0\boldsymbol{x}(t) + v_i(t), \quad i = 1,2 \tag{5.15.43}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix}, \quad \boldsymbol{H}_0 = [1 \quad 0] \tag{5.15.44}$$

其中 T_0 为采样周期,状态 $\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}$,$x_1(t)$,$x_2(t)$ 和 $w(t)$ 表示在时刻 tT_0 处的位置、速度和加速度,$y_i(t)$ 为第 i 个传感器对目标位置的观测信号,$v_i(t)$ 为观测噪声,$i = 1,2$,$w(t)$,$v_1(t)$ 和 $v_2(t)$ 是零均值相互独立白噪声,方差各为 σ_w^2,σ_{v1}^2 和 σ_{v2}^2.仿真中取

$$T_0 = 0.2, \quad \sigma_w^2 = 1, \quad \sigma_{v1}^2 = 4, \quad \sigma_{v2}^2 = 0.25 \tag{5.15.45}$$

$$\boldsymbol{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}, \quad \boldsymbol{v}(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_0 \\ \boldsymbol{H}_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \tag{5.15.46}$$

$$\boldsymbol{y}(t) = \boldsymbol{Hx}(t) + \boldsymbol{v}(t) \tag{5.15.47}$$

则 $\boldsymbol{y}(t)$ 叫合成观测,合成观测噪声 $\boldsymbol{v}(t)$ 有方差阵 $\boldsymbol{Q}_v = \mathrm{diag}(\sigma_{v1}^2, \sigma_{v2}^2)$.

对系统(5.15.42)和(5.15.47)实行稳态 Kalman 滤波器叫集中式观测融合 Kalman 滤波器.下面我们验证用经典 Kalman 滤波方法,基于 Riccati 方程所求稳态 Kalman 滤波器增益相同于用现代时间序列分析方法,基于 ARMA 新息模型求得的稳态 Kalman 滤波器增益,但求 ARMA 新息模型时,必须对多项式矩阵实行左素分解.

应用经典 Kalman 滤波方法,由公式(5.15.39)和(5.15.40)可求得稳态 Kalman 滤波器增益为

$$\boldsymbol{K}_f = \begin{bmatrix} 0.019\,6 & 0.313\,7 \\ 0.019\,8 & 0.316\,9 \end{bmatrix} \tag{5.15.48}$$

下面基于 ARMA 新息模型求 $\boldsymbol{K}_f$.由(5.15.42),(5.15.44)和(5.15.47)有

$$\boldsymbol{y}(t) = \boldsymbol{H}(\boldsymbol{I}_2 - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}w(t-1) + \boldsymbol{v}(t) \tag{5.15.49}$$

$$\boldsymbol{H}(\boldsymbol{I}_2 - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma} = \begin{bmatrix} (1-q^{-1})^2 & 0 \\ 0 & (1-q^{-1})^2 \end{bmatrix}^{-1} \begin{bmatrix} 0.5T_0^2(1+q^{-1}) \\ 0.5T_0^2(1+q^{-1}) \end{bmatrix} = \widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) \tag{5.15.50}$$

其中定义 $\widetilde{\boldsymbol{A}}(q^{-1})$ 和 $\widetilde{\boldsymbol{B}}(q^{-1})$ 为

$$\widetilde{\boldsymbol{A}}(q^{-1}) = \begin{bmatrix} (1-q^{-1})^2 & 0 \\ 0 & (1-q^{-1})^2 \end{bmatrix}, \quad \widetilde{\boldsymbol{B}}(q^{-1}) = \begin{bmatrix} 0.5T_0^2(1+q^{-1}) \\ 0.5T_0^2(1+q^{-1}) \end{bmatrix} \tag{5.15.51}$$

因为当 $q^{-1}=1$ 时有 $\mathrm{rank}[\widetilde{\boldsymbol{A}}(1),\widetilde{\boldsymbol{B}}(1)]=1$，由引理 5.15.1 知 $\widetilde{\boldsymbol{A}}(q^{-1})$ 与 $\widetilde{\boldsymbol{B}}(q^{-1})$ 非左素. 故需消去它们的极大左因式，以实现左素分解. 这可用如下列初等变换方法实现：

$$\begin{aligned}
&\begin{bmatrix} (1-q^{-1})^2 & 0 & 0.5T_0^2(1+q^{-1}) \\ 0 & (1-q^{-1})^2 & 0.5T_0^2(1+q^{-1}) \end{bmatrix} \xrightarrow[E_a]{\text{第 3 列乘 } 1/0.5T_0^2} \\
&\begin{bmatrix} (1-q^{-1})^2 & 0 & (1+q^{-1}) \\ 0 & (1-q^{-1})^2 & (1+q^{-1}) \end{bmatrix} \xrightarrow[E_b]{\text{第 3 列乘 } (3-q^{-1}) \text{ 加到第 2 列}} \\
&\begin{bmatrix} (1-q^{-1})^2 & (3-q^{-1})(1+q^{-1}) & (1+q^{-1}) \\ 0 & 4 & (1+q^{-1}) \end{bmatrix} \xrightarrow[E_c]{\text{第 1 列加第 2 列}} \\
&\begin{bmatrix} (1-q^{-1})^2 & 4 & (1+q^{-1}) \\ 0 & 4 & (1+q^{-1}) \end{bmatrix} \xrightarrow[E_d]{\text{第 2 列乘 } 1/4} \\
&\begin{bmatrix} (1-q^{-1})^2 & 1 & (1+q^{-1}) \\ 0 & 1 & (1+q^{-1}) \end{bmatrix} \xrightarrow[E_e]{\text{第 2 列乘 } -(1+q^{-1}) \text{ 加到第 3 列}} \begin{bmatrix} (1-q^{-1})^2 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}
\end{aligned} \tag{5.15.52}$$

这引出 $\widetilde{\boldsymbol{A}}(q^{-1})$ 和 $\widetilde{\boldsymbol{B}}(q^{-1})$ 的极大左因式为

$$\boldsymbol{M}(q^{-1}) = \begin{bmatrix} (1-q^{-1})^2 & 1 \\ 0 & 1 \end{bmatrix} \tag{5.15.53}$$

且有关系

$$\widetilde{\boldsymbol{A}}(q^{-1}) = \boldsymbol{M}(q^{-1})\overline{\boldsymbol{A}}(q^{-1}), \quad \widetilde{\boldsymbol{B}}(q^{-1}) = \boldsymbol{M}(q^{-1})\overline{\boldsymbol{B}}(q^{-1}) \tag{5.15.54}$$

可求得

$$\overline{\boldsymbol{A}}(q^{-1}) = \boldsymbol{M}^{-1}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) = \begin{bmatrix} 1 & -1 \\ 0 & (1-q^{-1})^2 \end{bmatrix} \tag{5.15.55}$$

$$\overline{\boldsymbol{B}}(q^{-1}) = \boldsymbol{M}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) = \begin{bmatrix} 0 \\ 0.5T_0^2(1+q^{-1}) \end{bmatrix} \tag{5.15.56}$$

消去左极大公因式有左素分解

$$\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) = \overline{\boldsymbol{A}}^{-1}(q^{-1})\overline{\boldsymbol{B}}(q^{-1}) \tag{5.15.57}$$

由(5.15.55)，注意 $\overline{\boldsymbol{A}}(q^{-1}) = \overline{\boldsymbol{A}}_0 + \overline{\boldsymbol{A}}_1 q^{-1} + \overline{\boldsymbol{A}}_2 q^{-2}$ 为

$$\overline{\boldsymbol{A}}(q^{-1}) = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix} q^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} q^{-2} \tag{5.15.58}$$

它的首系数阵 $\overline{A}_0$ 不是单位阵. 可用如下方法化为首系数阵为单位阵的左素分解：定义 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{B}(q^{-1})$ 为

$$\boldsymbol{A}(q^{-1}) = \overline{\boldsymbol{A}}_0^{-1}\overline{\boldsymbol{A}}(q^{-1}), \quad \boldsymbol{B}(q^{-1}) = \overline{\boldsymbol{A}}_0^{-1}\overline{\boldsymbol{B}}(q^{-1}) \tag{5.15.59}$$

则有 $\boldsymbol{A}(q^{-1})$ 的首系数阵为单位阵，

$$\boldsymbol{A}(q^{-1}) = \boldsymbol{I}_2 + \begin{bmatrix} 0 & -2 \\ 0 & -2 \end{bmatrix} q^{-1} + \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} q^{-2} \tag{5.15.60}$$

$$\boldsymbol{B}(q^{-1}) = \begin{bmatrix} 0.5T_0^2(1+q^{-1}) \\ 0.5T_0^2(1+q^{-1}) \end{bmatrix} \tag{5.15.61}$$

则最终有左素分解

$$\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1}) \tag{5.15.62}$$

应用(5.15.49)~(5.15.51)和(5.15.62)引出关系

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{B}(q^{-1})w(t-1) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.15.63}$$

这引出 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.15.64}$$

且有关系 $\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \boldsymbol{B}(q^{-1})w(t-1) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t)$，其中新息 $\boldsymbol{\varepsilon}(t)$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，$\boldsymbol{D}(q^{-1}) = \boldsymbol{I}_2 + \boldsymbol{D}_1 q^{-1} + \boldsymbol{D}_2 q^{-2}$ 是稳定的.用 Gevers - Wouters 算法可得

$$\boldsymbol{D}_1 = \begin{bmatrix} 0.023\,567 & -1.622\,9 \\ 0.023\,567 & -1.622\,9 \end{bmatrix}, \quad \boldsymbol{D}_2 = \begin{bmatrix} -0.019\,606 & 0.686\,31 \\ -0.019\,606 & 0.686\,31 \end{bmatrix},$$

$$\boldsymbol{Q}_\varepsilon = \begin{bmatrix} 4.117\,6 & 0.117\,6 \\ 0.117\,6 & 0.367\,6 \end{bmatrix} \tag{5.15.65}$$

应用公式(5.15.35)和(5.15.36)可求得稳态 Kalman 滤波器增益

$$\boldsymbol{K}_f = \begin{bmatrix} 0.019\,6 & 0.313\,7 \\ 0.019\,8 & 0.316\,9 \end{bmatrix} \tag{5.15.66}$$

它相同于基于 Riccati 方程求得的由(5.15.48)给出的 $\boldsymbol{K}_f$，即两种方法引出相同结果，但必须完成上述左素分解.

【例 5.15.3】 考虑带位置和速度观测的跟踪系统

$$\boldsymbol{x}(t) = \boldsymbol{\Phi x}(t) + \boldsymbol{\Gamma w}(t) \tag{5.15.67}$$

$$\boldsymbol{y}(t) = \boldsymbol{Hx}(t) + \boldsymbol{v}(t) \tag{5.15.68}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_0 & 0.5T_0^2 \\ 0 & 1 & T_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} T_0^3/6 \\ T_0^2/2 \\ T_0 \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{5.15.69}$$

其中 T_0 为采样周期，状态 $\boldsymbol{x}(t) = [x_1(t), x_2(t), x_3(t)]^{\mathrm{T}}$，$x_1(t)$，$x_2(t)$ 和 $x_3(t)$ 分别为在时刻 tT_0 处运动目标的位置、速度和加速度，$\boldsymbol{y}(t) = [y_1(t), y_2(t)]^{\mathrm{T}}$ 是对位置和速度的观测信号，$\boldsymbol{v}(t)$ 为观测噪声.假设 $w(t)$ 和 $\boldsymbol{v}(t)$ 是零均值、方差阵各为 σ_w^2 和 $\boldsymbol{Q}_v = \mathrm{diag}(\sigma_{v1}^2, \sigma_{v2}^2)$ 的相互独立的高斯白噪声.问题是基于 ARMA 新息模型和基于 Riccati 方程两种方法，分别求稳态 Kalman 滤波器和预报器增益，并验证它们分别相等，即两种方法是功能等价的.

首先应用 ARMA 新息模型法，由(5.15.67)和(5.15.68)有

$$\boldsymbol{y}(t) = \boldsymbol{H}(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma} w(t-1) + \boldsymbol{v}(t) \tag{5.15.70}$$

应用(5.15.8)可得

$$\det(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi}) = (1 - q^{-1})^3 \tag{5.15.71}$$

$$\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi})\boldsymbol{\Gamma} = \begin{bmatrix} \dfrac{T_0^3}{6}(1 + 4q^{-1} + q^{-2}) \\ \dfrac{T_0^2}{6}(1 - q^{-2}) \end{bmatrix} \tag{5.15.72}$$

于是有

$$\boldsymbol{y}(t)=\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1})w(t-1)+\boldsymbol{v}(t) \tag{5.15.73}$$

其中定义多项式矩阵

$$\widetilde{\boldsymbol{A}}(q^{-1})=\begin{bmatrix}(1-q^{-1})^3 & 0\\ 0 & (1-q^{-1})^3\end{bmatrix},\quad \widetilde{\boldsymbol{B}}(q^{-1})=\begin{bmatrix}\dfrac{T_0^3}{6}(1+4q^{-1}+q^{-2})\\ \dfrac{T_0^2}{2}(1-q^{-2})\end{bmatrix} \tag{5.15.74}$$

应用引理 5.15.2 用列变换法求 $\widetilde{\boldsymbol{A}}(q^{-1})$ 和 $\widetilde{\boldsymbol{B}}(q^{-1})$ 的极大左因式：

$$[\widetilde{\boldsymbol{A}}(q^{-1}),\widetilde{\boldsymbol{B}}(q^{-1})]=\begin{bmatrix}(1-q^{-1})^3 & 0 & \dfrac{T_0^3}{6}(1+4q^{-1}+q^{-2})\\ 0 & (1-q^{-1})^3 & \dfrac{T_0^2}{2}(1-q^{-2})\end{bmatrix}\xrightarrow[E_a]{\text{第 3 列乘 } 2/T_0^2}$$

$$\begin{bmatrix}(1-q^{-1})^3 & 0 & \dfrac{T_0}{3}(1+4q^{-1}+q^{-2})\\ 0 & (1-q^{-1})^3 & 1-q^{-2}\end{bmatrix}\xrightarrow[E_b]{\text{第 3 列乘 }(3-q^{-1})\text{ 加到第 2 列}}$$

$$\begin{bmatrix}(1-q^{-1})^3 & \dfrac{T_0}{3}(1+4q^{-1}+q^{-2})(3-q^{-1}) & \dfrac{T_0}{3}(1+4q^{-1}+q^{-2})\\ 0 & 4(1-q^{-1}) & 1-q^{-2}\end{bmatrix}\xrightarrow[E_c]{\text{第 1 列乘}\frac{-T_0}{3}\text{ 加第 2 列}}$$

$$\begin{bmatrix}(1-q^{-1})^3 & (\dfrac{2T_0}{3}+\dfrac{14T_0}{3}q^{-1}-\dfrac{4T_0}{3}q^{-2}) & \dfrac{T_0}{3}(1+4q^{-1}+q^{-2})\\ 0 & 4(1-q^{-1}) & 1-q^{-2}\end{bmatrix}\xrightarrow[E_d]{\text{第 2 列乘 } 1/4}$$

$$\begin{bmatrix}(1-q^{-1})^3 & (\dfrac{T_0}{6}+\dfrac{7T_0}{6}q^{-1}-\dfrac{T_0}{3}q^{-2}) & \dfrac{T_0}{3}(1+4q^{-1}+q^{-2})\\ 0 & (1-q^{-1}) & 1-q^{-2}\end{bmatrix}\xrightarrow[E_e]{\text{第 2 列乘}-(1+q^{-1})\text{ 加到第 3 列}}$$

$$\begin{bmatrix}(1-q^{-1})^3 & (\dfrac{T_0}{6}+\dfrac{7T_0}{6}q^{-1}-\dfrac{T_0}{3}q^{-2}) & (\dfrac{T_0}{6}-\dfrac{T_0}{2}q^{-2}+\dfrac{T_0}{3}q^{-3})\\ 0 & (1-q^{-1}) & 0\end{bmatrix}\xrightarrow[E_f]{\text{第 3 列乘 } 3/T_0}$$

$$\begin{bmatrix}(1-q^{-1})^3 & (\dfrac{T_0}{6}+\dfrac{7T_0}{6}q^{-1}-\dfrac{T_0}{3}q^{-2}) & (\dfrac{1}{2}-\dfrac{3}{2}q^{-2}+q^{-3})\\ 0 & (1-q^{-1}) & 0\end{bmatrix}\xrightarrow[E_g]{\text{第 3 列加到第 1 列}}$$

$$\begin{bmatrix}\dfrac{3}{2}-3q^{-1}+\dfrac{3}{2}q^{-2} & (\dfrac{T_0}{6}+\dfrac{7T_0}{6}q^{-1}-\dfrac{T_0}{3}q^{-2}) & (\dfrac{1}{2}-\dfrac{3}{2}q^{-2}+q^{-3})\\ 0 & (1-q^{-1}) & 0\end{bmatrix}\xrightarrow[E_h]{\text{第 1 列乘 } 2/3}$$

$$\begin{bmatrix}(1-2q^{-1}+q^{-2}) & (\dfrac{T_0}{6}+\dfrac{7T_0}{6}q^{-1}-\dfrac{T_0}{3}q^{-2}) & (\dfrac{1}{2}-\dfrac{3}{2}q^{-2}+q^{-3})\\ 0 & (1-q^{-1}) & 0\end{bmatrix}\xrightarrow[E_i]{\text{第 1 列乘}\frac{T_0}{3}\text{ 加第 2 列}}$$

$$\begin{bmatrix}(1-2q^{-1}+q^{-2}) & (\dfrac{T_0}{2}+\dfrac{T_0}{2}q^{-1}) & (\dfrac{1}{2}-\dfrac{3}{2}q^{-2}+q^{-3})\\ 0 & (1-q^{-1}) & 0\end{bmatrix}\xrightarrow[E_j]{\text{第 1 列乘}-(0.5+q^{-1})\text{ 加第 3 列}}$$

$$\begin{bmatrix} (1-q^{-1})^2 & \frac{T_0}{2}(q^{-1}+1) & 0 \\ 0 & 1-q^{-1} & 0 \end{bmatrix} \tag{5.15.75}$$

上述初等列变换等价于用一个单位模阵 $\boldsymbol{U}(q^{-1})$ 右乘 $[\widetilde{\boldsymbol{A}}(q^{-1}), \widetilde{\boldsymbol{B}}(q^{-1})]$，即

$$[\widetilde{\boldsymbol{A}}(q^{-1}), \widetilde{\boldsymbol{B}}(q^{-1})]\boldsymbol{U}(q^{-1}) = [\boldsymbol{M}(q^{-1}), \boldsymbol{0}] \tag{5.15.76}$$

其中 $\boldsymbol{U}(q^{-1})$ 为

$$\boldsymbol{U}(q^{-1}) = \boldsymbol{E}_a\boldsymbol{E}_b\boldsymbol{E}_c\boldsymbol{E}_d\boldsymbol{E}_e\boldsymbol{E}_f\boldsymbol{E}_g\boldsymbol{E}_h\boldsymbol{E}_i\boldsymbol{E}_j =$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{2}{T_0^2} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3-q^{-1} & 1 \end{bmatrix}\begin{bmatrix} 1 & -\frac{T_0}{3} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1-q^{-1} \\ 0 & 0 & 1 \end{bmatrix}\times$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{3}{T_0} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & \frac{T_0}{3} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -\frac{1}{2}-q^{-1} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} =$$

$$\begin{bmatrix} \frac{5}{6}+\frac{1}{6}q^{-1} & \frac{7T_0}{36}+\frac{T_0}{18}q^{-1} & -\frac{1}{6}-\frac{2}{3}q^{-1}-\frac{1}{6}q^{-2} \\ -\frac{T_0}{2}(1+q^{-1}) & \frac{1}{12}-\frac{1}{6}q^{-1} & \frac{1+q^{-2}}{2T_0} \\ \frac{(1-q^{-1})^2}{T_0^3} & \frac{11-7q^{-1}+2q^{-2}}{6T_0^2} & \frac{(1-q^{-1})^3}{T_0^3} \end{bmatrix} \tag{5.15.77}$$

由 (5.15.2) 引出 $\widetilde{\boldsymbol{A}}(q^{-1})$ 与 $\widetilde{\boldsymbol{B}}(q^{-1})$ 的极大左因式为

$$\boldsymbol{M}(q^{-1}) = \begin{bmatrix} (1-q^{-1})^2 & 0.5T_0(1+q^{-1}) \\ 0 & 1-q^{-1} \end{bmatrix} \tag{5.15.78}$$

即

$$\widetilde{\boldsymbol{A}}(q^{-1}) = \boldsymbol{M}(q^{-1})\bar{\boldsymbol{A}}(q^{-1}), \widetilde{\boldsymbol{B}}(q^{-1}) = \boldsymbol{M}(q^{-1})\bar{\boldsymbol{B}}(q^{-1}) \tag{5.15.79}$$

$$\boldsymbol{M}^{-1}(q^{-1}) = \frac{1}{(1-q^{-1})^3}\begin{bmatrix} 1-q^{-1} & -0.5T_0(1+q^{-1}) \\ 0 & (1-q^{-1})^2 \end{bmatrix} \tag{5.15.80}$$

这引出

$$\bar{\boldsymbol{A}}(q^{-1}) = \boldsymbol{M}^{-1}(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) = \begin{bmatrix} 1-q^{-1} & -0.5T_0(1+q^{-1}) \\ 0 & (1-q^{-1})^2 \end{bmatrix} \tag{5.15.81}$$

$$\bar{\boldsymbol{B}}(q^{-1}) = \boldsymbol{M}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) = \begin{bmatrix} -T_0^3/12 \\ 0.5T_0^2(1+q^{-1}) \end{bmatrix} \tag{5.15.82}$$

于是消去极大左因式 $\boldsymbol{M}(q^{-1})$ 有左素分解

$$\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) = \bar{\boldsymbol{A}}^{-1}(q^{-1})\bar{\boldsymbol{B}}(q^{-1}) \tag{5.15.83}$$

注意 $\bar{\boldsymbol{A}}(q^{-1})$ 可写为矩阵多项式形式

$$\bar{\boldsymbol{A}}(q^{-1}) = \bar{\boldsymbol{A}}_0 + \bar{\boldsymbol{A}}_1q^{-1} + \bar{\boldsymbol{A}}_2q^{-2},$$

$$\bar{\boldsymbol{A}}_0 = \begin{bmatrix} 1 & -0.5T_0 \\ 0 & 1 \end{bmatrix}, \quad \bar{\boldsymbol{A}}_1 = \begin{bmatrix} -1 & -0.5T_0 \\ 0 & -2 \end{bmatrix}, \quad \bar{\boldsymbol{A}}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{5.15.84}$$

它的首系数阵 $\bar{A}_0$ 不是单位阵，引入

$$\boldsymbol{A}(q^{-1}) = \bar{\boldsymbol{A}}_0^{-1}\bar{\boldsymbol{A}}(q^{-1}), \quad \boldsymbol{B}(q^{-1}) = \bar{\boldsymbol{A}}_0^{-1}\bar{\boldsymbol{B}}(q^{-1}) \tag{5.15.85}$$

则有左素分解

$$\tilde{\boldsymbol{A}}^{-1}(q^{-1})\tilde{\boldsymbol{B}}(q^{-1}) = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1}) \tag{5.15.86}$$

其中 $\boldsymbol{A}(q^{-1})$ 的首系数阵 $\boldsymbol{A}_0 = \boldsymbol{I}_m$ 为单位阵. 易知

$$\boldsymbol{A}(q^{-1}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & -1.5T_0 \\ 0 & -2 \end{bmatrix} q^{-1} + \begin{bmatrix} 0 & 0.5T_0 \\ 0 & 1 \end{bmatrix} q^{-2} \tag{5.15.87}$$

$$\boldsymbol{B}(q^{-1}) = \begin{bmatrix} (-T_0^3/12) + 0.25T_0^2(1+q^{-1}) \\ 0.5T_0^2(1+q^{-1}) \end{bmatrix} \tag{5.15.88}$$

将 (5.15.86) 代入 (5.15.73) 有

$$\boldsymbol{A}(q^{-1})y(t) = \boldsymbol{B}(q^{-1})w(t-1) + \boldsymbol{A}(q^{-1})v(t) \tag{5.15.89}$$

这引出 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{5.15.90}$$

其中 $\boldsymbol{D}(q^{-1}) = \boldsymbol{I}_2 + \boldsymbol{D}_1 q^{-1} + \boldsymbol{D}_2 q^{-2}$ 是稳定的多项式矩阵，新息 $\boldsymbol{\varepsilon}(t)$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \boldsymbol{B}(q^{-1})w(t-1) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{5.15.91}$$

$\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers – Wouters 算法求得.

在仿真中取 $T_0 = 0.38$，$\sigma_w^2 = 1$，$\boldsymbol{Q}_v = \mathrm{diag}(0.1, 0.2)$ 可得 ARMA 新息模型为 (5.15.90)，其中

$$\boldsymbol{A}(q^{-1}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & -0.57 \\ 0 & -2 \end{bmatrix} + \begin{bmatrix} 0 & 0.19 \\ 0 & 1 \end{bmatrix} q^{-2},$$

$$\boldsymbol{D}(q^{-1}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -0.5097 & -0.2057 \\ 0.3510 & -1.3403 \end{bmatrix} q^{-1} + \begin{bmatrix} -0.0584 & 0.1047 \\ -0.3073 & 0.5510 \end{bmatrix} q^{-2},$$

$$\boldsymbol{Q}_\varepsilon = \begin{bmatrix} 0.1821 & 0.1016 \\ 0.1016 & 0.4196 \end{bmatrix} \tag{5.15.92}$$

应用 (5.10.14) 可求得

$$\boldsymbol{M}_1 = \begin{bmatrix} 0.4903 & 0.3643 \\ 0.3510 & 0.6597 \end{bmatrix}, \quad \boldsymbol{M}_2 = \begin{bmatrix} 0.6320 & 0.6550 \\ 0.3947 & 0.8704 \end{bmatrix} \tag{5.15.93}$$

易知该系统可观性指数 $\beta = 2$，应用 (5.10.42) 有稳态滤波和预报增益各为

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \end{bmatrix}^+ \begin{bmatrix} \boldsymbol{I}_2 - \boldsymbol{Q}_v\boldsymbol{Q}_\varepsilon^{-1} \\ \boldsymbol{M}_1 \end{bmatrix} = \begin{bmatrix} 0.3653 & 0.1536 \\ 0.3073 & 0.4490 \\ 0.1150 & 0.5546 \end{bmatrix} \tag{5.15.94}$$

$$\boldsymbol{K}_p = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \end{bmatrix}^+ \begin{bmatrix} \boldsymbol{M}_1 \\ \boldsymbol{M}_2 \end{bmatrix} = \begin{bmatrix} 0.4903 & 0.3643 \\ 0.3510 & 0.6597 \\ 0.1150 & 0.5546 \end{bmatrix} \tag{5.15.95}$$

其次基于 Riccati 方程 (5.15.40) 由 (5.15.39) 可求得

$$\boldsymbol{K}_f = \begin{bmatrix} 0.3653 & 0.1536 \\ 0.3073 & 0.4490 \\ 0.1150 & 0.5546 \end{bmatrix} \tag{5.15.96}$$

它相同于(5.15.94),且有

$$\boldsymbol{K}_p = \boldsymbol{\Phi}\boldsymbol{K}_f = \begin{bmatrix} 0.4903 & 0.3643 \\ 0.3510 & 0.6597 \\ 0.1150 & 0.5546 \end{bmatrix} \tag{5.15.97}$$

它相同于(5.15.95).这验证了两种方法的功能等价性.

【注 5.15.1】 例5.15.3的跟踪模型(5.15.67)~(5.15.69)要比例5.15.1的跟踪模型(5.15.3)~(5.15.5)更精确些.因为在例5.15.1中意味着对位置 $x_1(t)$ 用Taylor展式展开到 $x_1(t)$ 的二阶导数为止,对速度 $x_2(t)$ 展到 $x_2(t)$ 的一阶导数为止.而例5.15.3意味着用Taylor展式将 $x_1(t)$ 展到 $x_1(t)$ 的三阶导数项为止,对 $x_2(t)$ 展到 $x_2(t)$ 的二阶导数项为止.换言之,在例5.15.1中有

$$x_1(t+1) = x_1(t) + \dot{x}_1(t)T_0 + \frac{\ddot{x}_1(t)}{2!}T_0^2 \tag{5.15.98}$$

$$x_2(t+1) = x_2(t) + \dot{x}_2(t)T_0 \tag{5.15.99}$$

其中 $x_2(t) = \dot{x}_1(t)$,$x_3(t) = \ddot{x}_1(t)$,$\dot{x}_2(t) = x_3(t)$.在例5.15.3中有

$$x_1(t+1) = x_1(t) + \dot{x}_1(t)T_0 + \frac{\ddot{x}_1(t)}{2!}T_0^2 + \frac{\dddot{x}_1(t)}{3!}T_0^3 \tag{5.15.100}$$

$$x_2(t+1) = x_2(t) + \dot{x}_2(t)T_0 + \frac{\ddot{x}_2(t)}{2!}T_0^2 \tag{5.15.101}$$

其中 $x_2(t) = \dot{x}_1(t)$,$x_3(t) = \ddot{x}_1(t)$,$w(t) = \dddot{x}(t)$,$\dot{x}_2(t) = x_3(t)$

参 考 文 献

1 邓自立.最优滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2000

2 邓自立,郭一新.现代时间序列分析及其应用——建模、滤波、去卷、预报和控制.北京:知识出版社,1989

3 Deng Z L, Zhang H S, Liu S J, Zhou L. Optimal and Self-tuning White Noise Estimators with Applications to Deconvolution and Filtering Problems. Automatica, 1996,32(2):199~216

4 Deng Z L. White-Noise Filter and Smoother with Application to Seismic Data Deconvolution, Preprints of the 7th IFAC/IFORS Symposium on Identification and System Parameter Estimation, York UK, 1985,621~624

5 邓自立,张焕水.自校正Kalman滤波、预报、去卷、平滑新方法.控制理论与应用,1994,11(2):137~145

6 邓自立,刘叔军.单通道白噪声估值器.控制与决策,1994,9(6):426~430

7 邓自立.动态系统白噪声估计理论.控制理论与应用,1995,12(5):612~615

8 邓自立,周露,刘叔军.动态系统白噪声估值器.自动化学报,1995,21(5):592~596

9 邓自立,刘叔军.单通道自校正白噪声去卷平滑器.信息与控制,1995,24(3):129~134

10 邓自立.应用于地震数据去卷的自校正白噪声估值器.自动化学报,1996,12(2):156~161

11 邓自立,刘叔军.实时自校正白噪声去卷滤波器和平滑器.1994中国控制与决策学术年会论文集,厦门,1994,285~288

12 邓自立,张焕水.非最小相位和不稳定系统的自适应白噪声去卷平滑器.1992中国控制与决策学术年会论文集,哈尔滨,1992,67~71

13 Mendel J M. Optimal Seismic Deconvolution: An Estimation-Based Approach. New York: Academic Press, 1983

14 Mendel J M. White – Noise Estimators for Seismic Data Processing in oil Exploration. IEEE Trans. Automatic Control，1977，AC – 22 (5)：694 ~ 706
15 Mendel J M and Kormylo J. New Fast Optimal White – Noise Estimators for Deconvolution. IEEE Trans. Geoscience Electronics，1977，GE – 15 (1)：32 ~ 41
16 Mendel J M. Minimum Variance Deconvolution. IEEE Trans. Geoscience and Remote Sensing，1981，GE – 19 (3)：161 ~ 171
17 Mendel J M. Lessons in Estimation Theory for Signal Processing，Communications，and Control .Prentice Hall，Englewood Cliffs，New Tersey，1995
18 Deng Z L，Xu Y. Descriptor Wiener State Estimators. Automatica，2000，36 (9)：1761 ~ 1766
19 Deng Z L. Time – Domain Approaches to Multichannel Optimal Deconvolution. International Journal of Systems Science，2000，31 (6)：787 ~ 796
20 邓自立，郭金柱，孙书利.多通道 Wiener 滤波器设计新方法.控制与决策，2000，15 (3)：318 ~ 321
21 邓自立，刘玉梅.随机控制系统稳态 Kalman 滤波器新算法.自动化学报，2000，26 (1)：74 ~ 78
22 邓自立，郭金柱，许燕.广义系统 ARMA 最优递推状态估值器.自动化学报，2000，26 (2)：250 ~ 254
23 邓自立，王莅辉.固定区间 Kalman 平滑新算法.控制理论与应用，2000，17 (5)：777 ~ 780
24 邓自立，刘玉梅.一类新的固定点和固定区间 Kalman 平滑器.自动化学报，1999，25 (1)：32 ~ 37
25 邓自立，刘伟华，石莹.应用 Diophantine 方程的多通道最优去卷.自动化学报，1999，25 (3)：380 ~ 383
26 邓自立，刘玉梅.广义系统稳态 Kalman 估值器.自动化学报，1999，25 (4)：483 ~ 487
27 邓自立，许燕.广义系统 Wiener 滤波和 Kalman 滤波新方法.控制理论与应用，1999，16 (5)：634 ~ 638
28 邓自立，刘玉梅.稳态 Kalman 滤波的一种统一格式.控制与决策.1999，14 (1)：25 ~ 29
29 邓自立，郭金柱，孙书利.用时域方法设计多通道 Wiener 去卷滤波器.控制与决策，1999，14 (增刊)：491 ~ 495
30 Zi – Li Deng，Yu – Mei Liu. Descriptor Kalman Estimators. International Journal of Systems Science，1999，30 (11)：1205 ~ 1212
31 Zi – Li Deng. Multichannel Optimal Deconvolution Filters Using Modern Time Series Analysis Method. Proc. of 14th IFAC World Congress，Beijing，1999，Vol. I：1 ~ 6
32 邓自立，刘玉梅.一种统一的稳态 Kalman 估值器.信息与控制，1999，28 (4)：249 ~ 254
33 邓自立，刘玉梅.一类稳态 Kalman 滤波器及其渐近稳定性.信息与控制，1998，27 (1)：26 ~ 31
34 邓自立，许燕.一种统一的 Wiener 状态估值器.信息与控制，1998，27 (5)：336 ~ 341
35 邓自立，刘伟华，石莹.解决多通道 Wiener 滤波问题的 ARMA 新息滤波器.控制与决策，1998，13 (2)：169 ~ 172
36 邓自立，王莅辉.Wiener 滤波新方法.控制与决策，1998，13 (增刊)：396 ~ 401
37 邓自立.解耦 Wiener 状态滤波器.中国学术期刊文摘 (科技快报)，2000，6 (8)：981 ~ 982
38 邓自立.解耦稳态 Kalman 滤波器.中国学术期刊文摘 (科技快报)，2000，6 (8)：983 ~ 984
39 邓自立，金鹏.带位置和速度观测的稳态 Kalman 滤波器.中国学术期刊文摘 (科技快报)，2000，6 (8)：894 ~ 896
40 邓自立，金鹏.不带 Riccati 方程的 α-β 滤波器.中国学术期刊文摘 (科技快报)，2000，6 (8)：896 ~ 897
41 邓自立，祁荣宾.Wiener 去卷滤波器设计的一种新方法.中国学术期刊文摘 (科技快报)，2000，6 (7)：868 ~ 870
42 邓自立，罗秋滨.单通道 Wiener 去卷滤波器的一种新算法.中国学术期刊文摘 (科技快报)，2000，6 (7)：870 ~ 871
43 邓自立，齐国元.构造 ARMA 新息模型的 Gevevs – Wouters 算法.中国学术期刊文摘 (科技快报)，2000，6 (6)：727 ~ 729

44 邓自立,王玉成.广义 Kalman 滤波器新算法.中国学术期刊文摘(科技快报),2000,6(6):729~731
45 邓自立,罗秋滨.故障诊断的改进的加权残差平方和法.中国学术期刊文摘(科技快报),2000,6(5):602~604
46 邓自立,罗秋滨.偏差故障诊断的 U-检验法.中国学术期刊文摘(科技快报),2000,6(5):604~605
47 邓自立,祁荣宾.多传感器信息融合次优稳态 Kalman 滤波器.中国学术期刊文摘(科技快报),2000,6(2):183~184
48 邓自立,许燕.广义系统的正向和反向固定区间 Kalman 平滑器.中国学术期刊文摘(科技快报),2000,6(2):185~187
49 邓自立.白噪声估计理论的推广.中国学术期刊文摘(科技快报),2000,6(1):68~70
50 邓自立.基于 Markov 参数的稳态 Kalman 滤波器增益的一种新算法.中国学术期刊文摘(科技快报),2000,6(1):70~71
51 邓自立,王玉成.多通道最优去卷滤波器.中国学术期刊文摘(科技快报),1999,5(12):1499~1502
52 邓自立,郭金柱,王莅辉.用现代时间序列分析方法设计 Wiener 滤波器.中国学术期刊文摘(科技快报),1999,5(12):1502~1504
53 邓自立,王玉成,张明波.新的固定点和固定区间 Kalman 平滑算法.中国学术期刊文摘(科技快报),1999,5(10):1287~1289
54 邓自立,王玉成.多通道 ARMA 新息滤波器.中国学术期刊文摘(科技快报),1999,5(10):1289~1291
55 邓自立,郭金柱,张明波.带有色观测噪声系统 Wiener 状态滤波器.中国学术期刊文摘(科技快报),1999,5(7):893~895
56 邓自立,郭金柱,郭君红.统一的和通用的稳态 Kalman 估值器.中国学术期刊文摘(科技快报),1999,5(7):895~897
57 邓自立.带 MA 有色噪声系统 Wiener 滤波器.黑龙江大学自然科学学报,1999,16(4):41~44
58 邓自立,刘玉梅.Wiener 滤波器设计的一种新方法.黑龙江大学自然科学学报,1999,16(4):48~49
59 邓自立,许燕.广义 Kalman 估值器新算法.2000 中国控制与决策学术年会论文集,2000,181~184
60 张明波,邓自立,王玉成.一种简单的 Wiener 去卷滤波器.2000 中国控制与决策学术年会论文集,2000,189~192
61 孙书利,邓自立.带多重时滞系统 Wiener 状态滤波器.2000 中国控制与决策学术年会论文集,2000,193~196
62 郭君红,邓自立.随机控制系统稳态 Kalman 滤波器和预报器.1999 中国控制与决策学术年会论文集,258~261
63 王莅辉,邓自立.多变量 ARMA 信号的 Wiener 滤波器设计.1999 中国控制与决策学术年会论文集,266~269
64 邓自立,郭君红.统一的和通用的稳态 Kalman 估值器.1999 中国控制与决策学术年会论文集,274~278
65 邓自立,刘玉梅.一种统一的稳态 Kalman 滤波、平滑和预报算法.1998 中国控制与决策学术年会论文集,246~249
66 邓自立,胡萍,李国英.非递推最优状态估计的几种统一算法.1998 中国控制与决策学术年会论文集,250~253
67 刘玉梅,邓自立.应用 Kalman 滤波方法的 Wiener 滤波器设计.1998 中国控制与决策学术年会论文集,264~267
68 邓自立,许燕.广义系统非递推状态估计的几种统一算法.1998 中国控制会议论文集,1998,322~326
69 许燕,邓自立.广义系统最优递推状态估计.1998 中国控制会议论文集,1998,333~336
70 邓自立,石莹.最优和自校正去卷滤波器.1998 中国控制会议论文集,1998,337~341

71 邓自立，郭金柱，孙书利.Wiener 去卷滤波器设计新方法.自动化学报，2001，27(2)：241 ~ 246
72 陈建国，师自强，王培德.Kalman 滤波理论的推广.控制理论与应用，1990，7(3)：108 ~ 112
73 Lewis F L. Optimal Estimation. John Wiley & Sons，New York，1986
74 邓自立，李北新.自校正 α-β 跟踪滤波器，自动化学报，1992，18(6)：720 ~ 723
75 Chisci L，and Mosca E. Polynomial Equations for the linear MMSE State Estimation. IEEE Trans. Automatic Contvol，1992，37(5)：623 ~ 626
76 王莅辉，邓自立.多变量 ARMA 信号的 Wiener 滤波器设计.1999 中国控制与决策学术年会论文集.沈阳：东北大学出版社，1999，266 ~ 269
77 尤昌德.线性系统理论基础.北京：电子工业出版社，1985
78 Roberts A P and Newmann M M. Polynomial Approach to Wiener Filtering. Int，J. Control，1988，47(3)：681 ~ 696
79 Ahlen A and Sternad M. Optimal Deconvolution Using Polynomial Method. IEEE Trans. Acoustics，Speech，and Signal Processing，1989，37(2)：217 ~ 226
80 Grimble M J. Polynomial Approach to Optimal Linear Filterimg and Prediction. Int ，J. Control ，1985，41(6)：1545 ~ 1564
81 Kailath T. Linear Systems. Prentice – Hall，Inc，Englewood Cliffs，New Jersey，1980
82 邓自立，石莹 .Wiener 滤波问题的一种 Diophantine 方程解.控制与决策，1997，12(4)：289 ~ 294
83 Feinstein J and Bar – Ness Y. The Solution of the Matrix Polynomia1 Equation A(s)X(s) + B(s)Y(s) = C(s). IEEE Trans. Automatic Control ，1984，AC – 29(1)：75 ~ 77
84 Chisci L and Mosca E. A General Polynomial Solution to the MMSE Deconvolution Problem . IEEE Trans. Signal Processing，1991，39(4)：962 ~ 965
85 Chui C K and Chen G. Kalman Filtering with Real – Time Applications. Springer – Verlag，New York，1987
86 Ahlen A and Sternad M .Wiener Filter Design Using Polynomial Equations. IEEE Trans. Signal Processing，1991，39(11)：2387 ~ 2399
87 Chisci L and Mosca E. MMSE Deconvolution via Polynomial Method and Its Dual LQG Regulation. Automatica，1994，30(7)：1197 ~ 1201
88 Grimble M J. Multichannel Optimal Linear Deconvolution Filters and Strip Thickness Estimation From Gauge Measuremeuts，Trans. of the ASME，Journal of Dynamic Systems，Measurement，and Control，1995，117：165 ~ 173
89 Rogers S R. Alpha – Beta Filter with Correlated Measurement Noise. IEEE Trans. Aerospace and Electronic Systems，1987，AES – 23(4)：592 ~ 594
90 Guu J A and Wei C H. Tracking Technique for Manoeuvring Target with Correlated Measurement Noises and Unknown Parameters. IEE Proc – F，1991，138(3)：278 ~ 288
91 Gazit R. Digital Tracking Filters with High Order Correlated Measurement Noise. IEEE Trans. Aerospace and Electronic Systems，1997，AES – 33(1)：171 ~ 177
92 Guu J A and Wei C H. Tracking a Manoeuvring Target Using Input Etimation at High Measurement Frequency . Int. J. Systems Sci，1992，23(6)：871 ~ 883
93 Fung P T K and Grimble M J. Dynamic Ship Positioning Using a Self – Tuning Kalman Filter. IEEE Trans. Automatic Control，1983，AC ~ 28(3)：339 ~ 350
94 Friedland B. Treatment of Bias in Recursive Filtering. IEEE Trans. Automatic Control，1969，AC – 14(4)：359 ~ 367
95 Ignagni M B. Separate – Bias Kalman Filter with Bias State Noice. IEEE Trans. Automatic Control，1990，35(3)：338 ~ 341

96 Alouan A T et al. On the Optimality of Two – Stage State Estimation in the Presence of Random bias. IEEE Trans. Automatic Control，1993，38(8)：1279～1282
97 Keller J Y and Darouach M. Optimal Two – Stage Kalman Filter in the presence of Random Bias . Automatica，1997，33(9)：1745～1748
98 Nikoukhah R，Willsky A S and Bernard C L. Kalman Filtering and Riccati Equatiams for Descriptor Systems. IEEE Trans. Automatic Control ，1992，37(9)：1325～1341
99 Darouach M，Zasadzinski M，Mehdi D. State Estimation of Stochastic Singular Linear Systems. Int. J. Systems Sci.，1993，24(2)：345～354
100 甘特马赫尔 Φ P.矩阵论.北京：高等教育出版社，1955
101 须田信英等.自动控制中的矩阵理论.北京：科学出版社，1979
102 Kailath T. Linear Systems. Prentice – Hall，Inc.，Englewood Cliffs，N.J.，1980
103 Chisci L and Mosca E. Polynomial Equations for the Linear MMSE State Estimation. Automatic Control，1992 – 37(5)：623～626
104 Day L Y. Observers for Discrete Singular Systems. IEEE Trans. Automatic Control，1988，AC – 33(2)：187～191
105 秦超英，戴冠中.广义离散随机线性系统的状态估计.信息与控制，1992，21(5)：261～265
106 秦超英，戴冠中 . 广义离散随机线性系统的最优滤波.控制与决策.1993，8(1)：65～68
107 程云鹏.矩阵论.西安：西北工业大学出版社，1989
108 Syrmos V L and Lewis F L. Robust Eigenvalue Assignment for Generalized Systems. Automatica，1992，28(6)：1223～1228
109 邓自立，郭金柱，孙书利.Wiener 去卷滤波器设计新方法.自动化学报，2001，27(2)：241～246
110 邓自立，孙书利，郭金柱.Wiener 状态去卷滤波器.控制理论与应用，2001，18(4)：508～512
111 邓自立.估计 MA 参数的多维强 Gevers – Wouters 算法及其在构造 ARMA 新息模型中的应用.控制理论与应用，2001，18(5)：737～740
112 邓自立，罗秋滨.分离随机偏差两段触耦 Wiener 滤波器.控制理论与应用，2002，19(5)：755～758
113 邓自立，王玉成，刘伟华.不带 Diophantine 方程的多通道最优去卷滤波器. 自动化学报，28(1)：19～26
114 邓自立，张明波.统一的和通用的 Wiener 状态滤波器.自动化学报，2002，28(3)：427～430
115 邓自立，许燕.基于 Kalman 滤波的白噪声估计理论.自动化学报，2003，29(1)：23～31
116 许燕，邓自立.极点配置广义稳态 Kalman 估值器.自动化学报，2003，29(6)：835～841
117 邓自立，孙书利.极点配置固定滞后稳态 Kalman 平滑器.控制理论与应用，2003，20(5)：802～804
118 邓自立，孙书利.基于 Kalman 滤波的带相关噪声系统统一的 Wiener 状态估值器.控制理论与应用，2003.20(4)：573～576
119 邓自立.时变系统的统一和通用的最优白噪声估值器.控制理论与应用，2003，20(1)：143～146
120 邓自立，罗秋滨.分离随机偏差两段解耦 Wiener 滤波器.控制理论与应用，2002，19(5)：755～758
121 孙书利，邓自立.Wiener 状态滤波器设计新方法.控制与决策，2002，17(4)：503～505
122 许燕，邓自立.广义系统 Wiener 状态估值器.控制与决策，2003，18(3)：328～331
123 石莹，孙书利，许燕，邓自立.带有色观测噪声系统极点配置 Kalman 平滑器及其在跟踪系统中的应用.控制与决策，2002，17(增刊 2)：79～81
124 邓自立，石莹，孙书利，许燕.基于 Kalman 滤波的固定滞后 Wiener 状态平滑器.控制与决策，2002，17(增刊 2)：82～83
125 孙书利，石莹，邓自立.带有色观测噪声控制系统 Wiener 状态平滑器及其在跟踪系统中的应用.控制与决策，2002，17(增刊 2)：84～86

126 邓自立，孙书利．基于 Kalman 滤波的 Wiener 状态估值器．自动化学报，2004，30（1）：116～120
127 邓自立，许燕，石莹，孙书利．统一的快速次优固定区间白噪声 Wiener 平滑器．自动化学报，2004，30（2）：224～228
128 孙书利，邓自立．极点配置固定区间 Kalman 平滑器和 Wiener 平滑器．自动化学报，2004，30（2）：239～243
129 邓自立，石莹，孙书利，许燕．快速次优固定区间 Wiener 平滑器算法．控制理论与应用，2004，21（2）：275～278
130 邓自立．时域 Wiener 状态滤波新方法．控制理论与应用，2004，21（3）：367～372
131 邓自立．许燕．基于 Kalman 滤波的通用的和统一的白噪声估计方法．控制理论与应用，2004（4）：501～506
132 邓自立，孙书利，石莹．基于 Kalman 滤波的白噪声估计理论的推广．科学技术与工程，2003，3（6）：521～524
133 石莹，孟华，孙书利，邓自立．广义系统降阶极点配置 Kalman 平滑器．科学技术与工程，2003，3（6）：525～527
134 邓自立，石莹，孙书利．统一的和通用的 Kalman 滤波理论．科学技术与工程，2003，3（5）：400～404
135 孙书利，孟华，石莹，邓自立．广义系统降阶 Wiener 状态平滑器．科学技术与工程，2003，3（5）：405～407
136 邓自立，马建为，高媛．两传感器自校正信息融合 Kalman 滤波器．科学技术与工程，2003，3（4）：321～325
137 邓自立，马建为，高媛．两传感器自校正信息融合白噪声 Wiener 反卷积滤波器．科学技术与工程，2003，3（4）：325～327
138 邓自立，高媛，马建为．两传感器信息融合稳态最优 Kalman 滤波器．科学技术与工程，2003，3（3）：213～216
139 邓自立，高媛，马建为．两传感器信息融合最优白噪声反卷积 Wiener 滤波器．科学技术与工程，2003，3（3）：216～219
140 邓自立，孙书利，石莹．快速次优固定区间白噪声平滑器．科学技术与工程，2003，3（2）：106～109
141 石莹，孙书利，许燕，邓自立．带有色观测噪声系统统一的 Wiener 状态估值器．科学技术与工程，2003，3（2）：109～111
142 邓自立．线性离散随机系统最优和稳态 Kalman 平滑器．科学技术与工程，2002，2（6）：8～12
143 邓自立，孙书利，许燕，石莹．基于稳态 Riccati 方程迭代解的 ARMA 新息模型的构造．科学技术与工程，2002，2（6）：12～13
144 邓自立，杜洪越，马建为．改进的递推增广最小二乘参数估计方法．科学技术与工程，2002，2（5）：1～3
145 邓自立，马建为，杜洪越．ARMA 模型参数估计的两段最小二乘法．科学技术与工程，2002，2（5）：3～5
146 邓自立，王好谦，张明波．基于 Kalman 滤波的带相关噪声系统最优白噪声估值器．科学技术与工程，2002，2（3）：1～3
147 邓自立，王好谦．基于 Kalman 滤波的稳态白噪声估值器．科学技术与工程，2002，2（3）：3～5
148 邓自立，张明波，王玉成．统一的固定区间最优白噪声估值器．科学技术与工程，2002，2（3）：5～7
149 齐国元，邓自立．广义系统降阶极点配置 Kalman 估值器．黑龙江大学自然科学学报，2002，19（3）：40～43
150 齐国元，邓自立．Y 可观典范型广义系统降阶 Wiener 状态估值器．黑龙江大学自然科学学报，2002，19（2）：36～39

151 邓自立，高媛，李云，白敬，崔崇信．基于 Kalman 滤波的信息融合白噪声最优反卷积滤波器．科学技术与工程，2004，4（3）：169～171
152 邓自立，高媛，李云，崔崇信，白敬刚．信息融合稳态最优 Kalman 平滑器．科学技术与工程，2004，4（3）：172～175
153 许燕，邓自立．带 MA 有色观测噪声广义系统 Wiener 状态估值器．2002 第 17 届青年自动化学术年会论文集，83～87
154 邓自立，许燕．基于 Kalman 滤波的通用的和统一的白噪声估计方法．控制理论与应用，2004，21（4）：501～506
155 邓自立，石莹，孙书利，许燕．快速次优固定区间 Kalman 平滑算法．控制理论与应用，2004，21（2）：275～278
156 邓自立．时域 Wiener 状态滤波新方法．控制理论与应用，2004，21（3）：367～372
157 邓自立，石莹，孟华．基于 ARMA 新息模型与 Riccati 方程的两种 Kalman 跟踪滤波器的等价性．科学技术与工程，2004，4（11）：894～896
158 邓自立．ARMA 过程线性预报、滤波和平滑的新结果．全国控制理论及其应用学术交流会议论文集．北京：科学出版社，1981，164～169
159 Kailath T，Sayed A H，Hassibi B．Linear Estimation．Upper Saddle River，New Jevsey：Prentice－Hall，Inc．，2000
160 Hagander P，Wittenmark B．A Self－tuning Filter for Fixed－Lag Sinoothing IEEE Trans．Information Theory，1997，23（3）：377～384
161 Deng Zili．Multivariate Self－Tuning Filter and Smoother．Preprints of 6th IFAC Symp on Identification and System Parameter Estimation．Washington DC，1982，879～882
162 Deng Zili．White－Noise Filter and Smoothing with Aplication to Seismic Data Deconvolution．Preprints of the 7th IFAC/IFORS Symposium on Identification and System Parameter．York UK，1985，621～624
163 Moire T J．Optimal Decomvolution Smoother．Proc．IEE．Pt．D，1986，133（1）：13～18
164 Moire T J，Grimble M J．Optimal Self－Tuning Filtering，Prediction and Smoothing for Discrete Multivatiable Processes．IEEE Trans．Automatic Control，1984，AC－29（2）：128～135
165 Tajima K．Estimation of Seeady－State Kalman Filter Gain．IEEE Trans．Automatic Control，1978，AC－23（5）：944～945
166 Kalman RE．New Approach to Linear Filtering and Prediction Problems Trans．ASME，J Basic Eng．，1960，82D，34～35
167 Deng Zili．Estimation of Steady－State Optimeol Filter Gain with Colouried Observation Noise．Proceedings of the 1982 American Control Conference，Arlington，Virginia，1982，608～609
168 邓自立．稳态 Kalman 滤波器增益的估计．控制理论与应用，1985，2（1）：122～126
169 Gevers M，Wouters WRE．An Innovations Approach to the Discrete－Tine Stochastic Realization Problem．Quarterly Journal on Automatic Control，1978，19（2）：90～109
170 邓自立，郭一新．动态系统分析及其应用——建模、滤波、预报、控制的新方法和程序库．沈阳：辽宁科学技术出版社，1985
171 邓自立．卡尔曼滤波与维纳滤波——现代时间序列分析方法．哈尔滨：哈尔滨工业大学出版社，2001
172 邓自立．自校正滤波理论及其应用——现代时间序列分析方法．哈尔滨：哈尔滨工业大学出版社，2003
173 Anderson BDO，Moore JB．Optimal Filtering．Englewood Cliffs，New Jersey：Prentice－Hall，Inc．，1979
174 邓自立．估计 MA 参数的多维强 Gevers－Wouters 算法及其在构造 ARMA 新息模型中的应用．控制理论与应用，2001，18（5）：737～740

175 Jezek J, Kucera V. Efficient Algorithm for Matrix Spectral Factorization. Auto matica, 1985, 21, 663 ~ 669
176 Box G EP, Jenkins G M. Time Series Analysis, Forecasting and Control, San Francisco: Holden – Day, 1970
177 邓自立,孙书利,许燕,石莹.基于稳态 Riccati 方程迭代解的 ARMA 新息模型的构造.科学技术与工程,2002,2(6):12 ~ 13
178 Chisci L, Mosca E. Polynomial Equations for the Linear MMSE state Estimation. IEEE Trans. Automatic Control, 1992, 37(5):623 ~ 626
179 邓自立,石莹.应用 Diophantine 方程的最优去卷,1997 中国控制与决策学术年会论文集.沈阳:东北大学出版社,1997,316 ~ 320
180 邓自立,石莹.Wiener 滤波问题的一种 Diophantine 方程解.控制与决策,1994,6(2):143 ~ 148
181 邓自立,张焕水.应用 Diophantine 方程的最优估计的一种统一方法.1995 中国控制与决策学术年会论文集.沈阳:东北大学出版社,1995,281 ~ 286
182 邓自立,刘伟华,石莹.应用 Diphantime 方程的多通道最优去卷.自动化学报,1999,25(3):380 ~ 383
183 邓自立,李建国.稳态 Kalman 滤波器增益的一种新算法.自动化学报,1997,23(5):606 ~ 612
184 郑大钟.线性系统理论(第二版).北京:清华大学出版社,2002
185 邓自立,石莹,孟华.基于 ARMA 新息模型与 Riccati 方程的两种 Kalman 跟踪滤波器的等价性.科学技术与工程,2004,4(11):894 ~ 896
186 韩京清,何关钰,许可康.线性系统理论代数基础.沈阳:辽宁科学技术出版社,1985
187 Dai Y L. Observers for discrete singular systems. IEEE Trans. Automatic Control, 1988, 32(2):187 ~ 191
188 邓自立,孙书利.基于 Kalman 滤波的 Wiener 状态估值器.自动化学报,2004,30(1):126 ~ 130
189 孙书利,邓自立.极点配置固定区间 Kalman 平滑器和 Wiener 平滑器.自动化学报,2004,30(2):239 ~ 243
190 邓自立,许燕,石莹,孙书利.统一的快速次优固定区间白噪声 Wiener 平滑器.自动化学报,2004,30(2):224 ~ 248
191 邓自立,石莹,孙书利,许燕.快速次优固定区间 Wiener 平滑器算法.控制理论与应用,2004,21(2):275 ~ 278
192 邓自立,梁佐江.多传感器分布式信息融合 Wiener 信号滤波器.科学技术与工程,2005,5(9):539 ~ 542
193 邓自立,许燕.基于 Kalman 滤波的通用和统一的白噪声估计方法.控制理论与应用,2004,21(4):501 ~ 506
194 Sun Shuli, Deng Zili. Multi – sensor optimal information fusion Kalman filter. Automatica, 2004, 40:1017 ~ 1023
195 邓自立,高媛,王好谦.基于 Kalman 滤波的统一的 Wiener 状态滤波器.控制理论与应用,2004,21(6):1003 ~ 1006
196 石莹,沈永良,孙书利,邓自立.广义离散随机线性系统降阶 Wiener 滤波、平滑和预报器.控制理论与应用,2004,21(6):981 ~ 985
197 孙书利,邓自立.多传感器线性最小方差最优信息融合准则.科学技术与工程,2004,4(5):334 ~ 336
198 邓自立,高媛.两传感器信息融合超前 k 步稳态最优 Kalman 预报器.科学技术与工程,2004,4(5):334 ~ 340
199 邓自立,高媛,崔崇信.多传感器按对角阵加权信息融合 Kalman 滤波器.科学技术与工程,2004,4(7):518 ~ 521
200 高媛,白敬刚,邓自立.多传感器单通道信息融合 Wiener 滤波器.科学技术与工程,2004,4(7):522 ~ 525
201 邓自立,毛琳,高媛.多传感器最优信息融合稳态 Kalman 滤波器.科学技术与工程,2004,4(9):743 ~

747
202 邓自立，高媛，张明波.ARMA 信号自校正信息融合 Wiener 滤波器.科学技术与工程，2004，4(9)：749～752
203 邓自立，崔崇信，白敬刚.基于稳态 Kalman 滤波的两种观测融合方法的功能等价性.科学技术与工程，2004，4(11)：897～902
204 邓自立，高媛，毛琳，王欣.两传感器信息融合 Wiener 滤波器、平滑器和预报器.2004 中国控制与决策学术年会论文集.沈阳：东北大学出版社，2004，206～208
205 高媛，梁佐江，王欣，邓自立.两传感器信息融合 Wiener 反卷积滤波器.2004 中国控制与决策学术年会论文集.沈阳：东北大学出版社，2004，212～214
206 石莹，沈永良，孙书利，邓自立.广义离散随机线性系统降阶固定区间最优 Kalman 平滑器.黑龙江大学自然科学学报，2004，21(1)：47～50
207 孟华，孙书利，石莹，邓自立.广义离散系统降阶 Wiener 滤波、平滑和预报.黑龙江省大学自然科学学报，2004，21(2)：55～59
208 孙书利，邓自立.带有色观测噪声系统多传感标量加权最优信息融合稳态 Kalman 滤波器.控制理论与应用，2004，21(4)：635～638
209 高媛，毛琳，梁佐江，邓自立.两传感器按对角阵加权信息融合稳态 Kalman 滤波器.黑龙江大学自然科学学报，2004，21(2)：52～54
210 邓自立.Exponential Convergence of Iterative Solution to Lyapunov Equation. 科学技术与工程，2005，5(9)：543～546
211 邓自立.纯量谱分解的 Gevers－Woutors 算法收敛性分析.科学技术与工程，2005，5(1)：7～13

第六章　基于经典 Kalman 滤波的信息融合滤波理论

多传感器信息融合(Multisensor Information Fusion)也称多传感器数据融合(Multisensor Data Fusion)是 20 世纪 70 年代以来发展起来的一门新兴边缘学科,目前已成为倍受人们关注的热门领域.[1~7]

早在第二次世界大战期间人们就已把两传感器数据融合应用于火炮系统.当时在高炮火控雷达上加装了光学测距系统,对用雷达和光学传感器给出的两种信息进行融合测距,不仅提高了测距精度,而且增强了系统抗扰能力.但当时没有使用计算机进行融合处理,仅限于人工计算.20 世纪 70 年代初,美国海军就发现,对多个独立的声纳信号进行融合分析后,能够更准确地探测出敌方潜艇的位置.

20 世纪 70 年代以来,随着各种先进武器系统的出现和发展,例如精确制导、远程打击、导弹拦截等武器的出现,迫切要求提高对运动目标(导弹、飞机、卫星、坦克、车辆、船舰等)跟踪的精度或对系统状态与信号估计的精度.为了解决这类问题,随着电子技术和计算机应用技术的发展,出现了大量的具有不同应用背景的多传感器系统.例如对目标跟踪而言,有各种类型测量运动目标位置、速度或加速度的传感器.这是因为依靠传统的单传感器系统提供的信息很难满足目标跟踪或状态估计精度的要求.问题是如何对每个传感器提供的信息按某种最优融合准则进行融合,得出最优融合估计.显然,最优融合估计的基本性能应当是融合估计的精度应高于每个传感器估计的精度.

目前世界上许多科技发达国家都非常重视多传感器信息融合技术的开发和研究.早在 1988 年美国国防部就把信息融合技术列为 90 年代重点研究开发的二十项关键技术之一,且列为最优先发展的 A 类.我国起步较晚,在 20 世纪 90 年代后才逐渐形成研究热潮[5].

多传感器信息融合状态估计[4,5]是多传感器信息融合学科的一个重要分支或领域.目前基于经典 Kalman 滤波的多传感器信息融合理论尚在发展中,尚不完善,有许多研究课题有待进一步解决.目前有两种常用的信息融合方法[8]:一种方法是状态融合方法,另一种方法是观测融合方法.状态融合方法又分集中式 Kalman 滤波和分散化(分布式)Kalman 滤波.集中式 Kalman 滤波虽然在理论上可获得全局最优融合状态估计,但它的缺点是计算负担大,且容错性能差,而分散化 Kalman 滤波信息融合能克服这些缺点.在众多分散化 Kalman 滤波算法中[55,59,62],以 Carlson[9~11]的联邦 Kalman 滤波器最受重视,它已被美国空军容错导航系统"公共 Kalman 滤波器"计划选为基本算法.

Carlson 的联邦 Kalman 滤波器的基本原理[9~11]是:在按矩阵加权线性最小方差最优信息融合准则下,假设基于单个传感器的局部估计误差是不相关的,用加权最小二乘法导出了多传感器按矩阵加权最优融合算法,并提出用"方差上界技术"使局部估计误差不相

关，从而可实现全局最优信息融合估计.对实时应用而言，按矩阵加权最优融合算法要求较大的计算负担.为了克服这个缺点，文献[13]提出了在线性最小方差准则下按标量加权最优信息融合估计算法，用计算最优加权系数取代了最优加权阵，大大减小了计算负担，便于实时应用，但该算法的缺点是稍微损失了一点融合估计精度，且仍假设局部估计误差不相关.为了克服假设局部估计误差不相关的局限性，Kim[12]考虑了局部估计强差相关的情形，用极大似然法导出了按矩阵加权极大似然融合估计算法.但其缺点和局限性是假设局部估计误差服从联合正态分布.为了克服这个局限性，文献[14]针对局部估计误差相关情形，在线性最小方差准则下，用 Lagrange 乘数法导出了按矩阵加权最优融合公式，导出了相同于 Kim 的融合公式，但去掉了估计误差服从联合正态分布的假设.同时，文献[39]考虑了局部估计误差相关时的标量加权最优融合算法，用 Lagrange 乘数法导出了标量加权最优融合公式，推广了文献[13]的结果.文献[15]，提出了按对角阵加权最优信息融合估计算法，它的精度和计算量界于按矩阵加权融合算法和按标量加权融合算法之间.文献[64]给出了不同的按对角阵加权融合公式.文献[16]给出了上述三种加权融合准则的统一的简化的推导.上述三种加权分布式融合算法虽然计算量小，容错性强，但缺点是：它们是局部最优的，且它们是全局次优的.[35]

观测融合方法[8]有两种：一种是集中式观测融合方法，即将各传感器的观测方程合并成一个扩维的观测方程，然后与状态方程一起用于实现集中式全局最优 Kalman 滤波器.但它的缺点是计算负担大.另一种是加权观测融合方法.该方法按线性最小方差准则对每个传感器的观测方程加权得到一个融合的观测方程，其优点是不改变观测方程的维数，可明显减小计算负担，且可获得全局最优状态估计.但其局限性是要求每个传感器具有相同的观测阵.文献[8]证明了这两种观测融合方法是功能等价的，即用这两种观测融合方法所得 Kalman 滤波估值是恒同的.这叫部分功能等价.文献[20]进一步证明了完全功能等价性，即用这两种方法所得 Kalman 平滑估值、Kalman 预报估值、输入白噪声估值器，以及与状态有关的信号估值器在数值上是恒同的.

目前基于经典 Kalman 滤波的最优状态融合估计理论通常限于信息融合 Kalman 滤波器的设计.关于信息融合 Kalman 平滑器和信息融合多步 Kalman 预报器的报道甚少.没有形成可统一处理状态滤波、平滑和预报的统一的和通用的信息融合状态估计理论.关于这方面，文献[15，19，25，27，65]已有初步研究.

在许多应用问题中常常遇到信号估计问题.例如信号滤波、平滑和预报问题，信号反卷积估计问题，白噪声信号反卷积问题等.信号估计可作为状态估计的一种特殊形式，例如 ARMA 信号估计可转化为一个状态估计问题，其中信号是状态的分量.因此在对信号有多传感器进行观测情形，如何求最优信息融合信号估值器（滤波、平滑、预报）具有重要理论和应用意义.然而目前在国内外文献中关于这方面的报道和研究甚少.文献[15，18，22，23，24，29，31，32，33，37]在这方面做了一些初步工作.

针对现有基于经典 Kalman 滤波的信息融合估计理论的上述缺点、局限性和问题，本章将提出统一的和通用的信息融合状态和信号估计理论.可统一处理状态或信号信息融合滤波、平滑和预报问题，可在按矩阵加权、按对角阵加权和按标类加权三种加权方法下设计信息融合估值器.不仅提出了状态或信号的信息融合 Kalman 滤波器，而且还提出了它们的信息融合 Wiener 滤波器.难点或关键技术在于局部估计误差方差阵和局部估计误

差之间的协方差阵的计算.它们被用于计算最优加权.除了 Riccati 方程作为基本工具外,本章将用 Lyapunov 方程解决上述关键技术.还可统一处理时变系统和时不变系统最优信息融合问题,对时不变(定常)系统将提出相应的稳态信息融合滤波器.

综上所述,本章提出的基本经典 Kalman 滤波的信息融合估计理论包括:在线性最小方差意义下的按矩阵加权、按对角阵加权和按标量加权的三种最优信息融合准则和算法;时变和时不变系统的状态或信号的信息融合 Kalman 滤波器和 Wiener 滤波器、Kalman 平滑器和 Wiener 平滑器、Kalman 预报器和 Wiener 预报器;状态融合和观测融合;局部估计误差互协方差的计算.该理论具有重要的理论和应用意义,且在国防、军事、导航、机器人、GPS 定位、目标跟踪、通信、信号处理、控制等领域具有广泛的应用价值.

融合估计的两个主要问题是:融合规则和局部估计误差互协方差的计算.本章提出了三种加权融合规则,并基于 Lyapunov 方程解决了互协方差计算问题,从而解决了最优加权的计算问题.应指出,上述三种加权融合算法是局部最优的、全局次优的.[35]

6.1 三种加权多传感器最优信息融合准则

本节提出在线性最小方差意义下的多传感器系统的按矩阵加权、按标量加权和按对角阵加权的三种加权最优信息融合准则."最优"是相对的,是在指定的加权方式下,在线性最小方差意义下的"最优".最初有两传感器按矩阵加权融合[51],为了克服计算最优加权阵要求较大计算负担的缺点,文献[13]提出了按标量加权最优融合准则和算法,并且文献[15]进一步提出了按对角阵加权最优融合准则和算法.它有明显的物理意义,这种准则等价于对具有相同物理意义的分量按标量加权融合,实现了解耦信息融合.这三种加权最优融合准则和算法是本章的理论基础和特色.文献[14]推广了文献[51]的结果到多传感器情形.文献[37]推广文献[13]的结果到带非零互协方差情形.

6.1.1 按矩阵加权线性最小方差最优融合准则和算法

设随机向量 $\boldsymbol{x} \in R^n$,基于 L 个传感器观测已知它的 L 个无偏估计为 $\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \cdots, \hat{\boldsymbol{x}}_L$,即

$$\mathrm{E}\hat{\boldsymbol{x}}_i = \mathrm{E}\boldsymbol{x}, \quad i = 1, \cdots, L \tag{6.1.1}$$

设已知估计误差 $\tilde{\boldsymbol{x}}_i = \boldsymbol{x} - \hat{\boldsymbol{x}}_i$ 的方差阵 $\boldsymbol{P}_i = \mathrm{E}[\tilde{\boldsymbol{x}}_i \tilde{\boldsymbol{x}}_i^{\mathrm{T}}]$ 和协方差阵 $\boldsymbol{P}_{ij} = \mathrm{E}[\tilde{\boldsymbol{x}}_i \tilde{\boldsymbol{x}}_j^{\mathrm{T}}], i, j = 1, \cdots, L, \boldsymbol{P}_{ii} = \boldsymbol{P}_i, \boldsymbol{P}_{ij} = \boldsymbol{P}_{ji}^{\mathrm{T}}$,其中 E 为均值号,T 为转置号.问题是寻求 $\boldsymbol{x}$ 的按矩阵加权无偏融合估计 $\hat{\boldsymbol{x}}_0, \mathrm{E}\hat{\boldsymbol{x}}_0 = \mathrm{E}\boldsymbol{x}$,

$$\hat{\boldsymbol{x}}_0 = \sum_{i=1}^{L} \boldsymbol{A}_i \hat{\boldsymbol{x}}_i \tag{6.1.2}$$

其中加权阵 $\boldsymbol{A}_i$ 为 $n \times n$ 矩阵.在线性最小方差意义下,应选择加权阵 $\boldsymbol{A}_i$ 极小化融合估计误差的分量均方和 J,

$$J = \mathrm{E}[\tilde{\boldsymbol{x}}_0^{\mathrm{T}} \tilde{\boldsymbol{x}}_0], \quad \tilde{\boldsymbol{x}}_0 = \boldsymbol{x} - \hat{\boldsymbol{x}}_0 \tag{6.1.3}$$

它等价于

$$J = \mathrm{tr}\boldsymbol{P}_0, \quad \boldsymbol{P}_0 = \mathrm{E}[\tilde{\boldsymbol{x}}_0 \tilde{\boldsymbol{x}}_0^{\mathrm{T}}] \tag{6.1.4}$$

其中符号 tr 表示矩阵的迹.

由局部估计和融合估计的无偏性,由(6.1.2)引出约束条件

$$\sum_{i=1}^{L} \boldsymbol{A}_i = \boldsymbol{I}_n \tag{6.1.5}$$

其中 $\boldsymbol{I}_n$ 为 $n \times n$ 单位阵. 利用(6.1.5)有融合误差表达式

$$\tilde{\boldsymbol{x}}_0 = \sum_{i=1}^{L} \boldsymbol{A}_i \tilde{\boldsymbol{x}}_i \tag{6.1.6}$$

引入 $n \times nL$ 合成待定矩阵 $\boldsymbol{A}$,

$$\boldsymbol{A} = [\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_L] \tag{6.1.7}$$

定义 $nL \times nL$ 矩阵 $\boldsymbol{P}$ 为以 $\boldsymbol{P}_{ij}$ 为第 (i,j) 元素的分块矩阵,即

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}_{11} & \cdots & \boldsymbol{P}_{1L} \\ \vdots & & \vdots \\ \boldsymbol{P}_{L1} & \cdots & \boldsymbol{P}_{LL} \end{bmatrix} \tag{6.1.8}$$

则性能指标 J 可简写为

$$J = \mathrm{tr} \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{\mathrm{T}} \tag{6.1.9}$$

约束条件(6.1.5)可写为

$$\boldsymbol{A}\boldsymbol{e} = \boldsymbol{I}_n, \quad \boldsymbol{e} = \begin{bmatrix} \boldsymbol{I}_n \\ \vdots \\ \boldsymbol{I}_n \end{bmatrix} \tag{6.1.10}$$

因此问题归结为约束条件(6.1.10)下求矩阵 $\boldsymbol{A}$ 极小化性能指标(6.1.9).

因为约束(6.1.10)是一个矩阵等式,按矩阵元素它等价于 n^2 个标量约束条件. 为此应用 Lagrange 乘数法,引入辅助函数

$$F = J + \mathrm{tr} \sum_{i=1}^{n} \boldsymbol{\lambda}_i (\boldsymbol{A}\boldsymbol{e} - \boldsymbol{I}_n) \boldsymbol{e}_i \tag{6.1.11}$$

其中定义行向量

$$\boldsymbol{\lambda}_i = [\lambda_{i1}, \lambda_{i2}, \cdots, \lambda_{in}], \quad \boldsymbol{e}_i^{\mathrm{T}} = [0 \cdots 0\ 1\ 0 \cdots 0] \tag{6.1.12}$$

其中 $\boldsymbol{e}_i^{\mathrm{T}}$ 的第 i 列元素为 1,其余元素为零.

辅助函数(6.1.11)可解释为:矩阵约束 $\boldsymbol{A}\boldsymbol{e} = \boldsymbol{I}_n$ 按分量等价于 n^2 个标量约束条件,而 $(\boldsymbol{A}\boldsymbol{e} - \boldsymbol{I}_n)\boldsymbol{e}_i$ 为矩阵 $(\boldsymbol{A}\boldsymbol{e} - \boldsymbol{I}_n)$ 的第 i 列列向量,$\boldsymbol{\lambda}_i$ 为相应的 Lagrange 乘子行向量,这两个向量内积 $\boldsymbol{\lambda}_i(\boldsymbol{A}\boldsymbol{e} - \boldsymbol{I}_n)\boldsymbol{e}_i$ 相当于引入了 n 个标量约束条件的 n 个 Lagrange 乘子 $\boldsymbol{\lambda}_i = [\lambda_{i1}, \cdots, \lambda_{in}]$,于是(6.1.11)第二项相应于 n^2 个约束条件引入了 n^2 个 Lagrange 乘子 λ_{ij}, i, $j = 1, \cdots, n$.

应用矩阵迹微分公式

$$\frac{\partial}{\partial \boldsymbol{X}} \mathrm{tr}(\boldsymbol{X}\boldsymbol{B}\boldsymbol{X})^{\mathrm{T}} = 2\boldsymbol{X}\boldsymbol{B}\,(\boldsymbol{B} = \boldsymbol{B}^{\mathrm{T}}), \quad \frac{\partial}{\partial \boldsymbol{X}} \mathrm{tr}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}) = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}} \tag{6.1.13}$$

注意 $\boldsymbol{P} = \boldsymbol{P}^{\mathrm{T}}$ 和(6.1.9),置 $\partial F / \partial \boldsymbol{A} = \boldsymbol{0}$ 可得

$$2\boldsymbol{A}\boldsymbol{P} + \sum_{i=1}^{n} \boldsymbol{\lambda}_i^{\mathrm{T}} (\boldsymbol{e}\boldsymbol{e}_i)^{\mathrm{T}} = \boldsymbol{0} \tag{6.1.14}$$

上式左边第二项为

$$\sum_{i=1}^{n} \boldsymbol{\lambda}_i^{\mathrm{T}} \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{e}^{\mathrm{T}} = [\boldsymbol{\lambda}_i^{\mathrm{T}}, \cdots, \boldsymbol{\lambda}_n^{\mathrm{T}}] \begin{bmatrix} \boldsymbol{e}_i^{\mathrm{T}} \\ \vdots \\ \boldsymbol{e}_n^{\mathrm{T}} \end{bmatrix} \boldsymbol{e}^{\mathrm{T}} = \boldsymbol{\Lambda}^{\mathrm{T}} \boldsymbol{e}^{\mathrm{T}} \tag{6.1.15}$$

其中定义矩阵 $\boldsymbol{\Lambda}$ 为

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\lambda}_1 \\ \vdots \\ \boldsymbol{\lambda}_n \end{bmatrix} = (\lambda_{ij}) \tag{6.1.16}$$

它的第 i 行第 j 列元素为 λ_{ij}.

由(6.1.14)转置和(6.1.15)有关系

$$\boldsymbol{P}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{e}\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{0} \tag{6.1.17}$$

其中定义 $\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{\Lambda}/2$. 置 $\partial F/\partial \boldsymbol{\lambda}_i = \boldsymbol{0}$ 引出关系

$$\boldsymbol{e}_i^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{e} - \boldsymbol{I}_n)^{\mathrm{T}} = \boldsymbol{0}, \quad i = 1, \cdots, n \tag{6.1.18}$$

它等价于

$$\boldsymbol{e}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{I}_n = \boldsymbol{0} \tag{6.1.19}$$

合并(6.1.17)和(6.1.19)有矩阵方程

$$\begin{bmatrix} \boldsymbol{P} & \boldsymbol{e} \\ \boldsymbol{e}^{\mathrm{T}} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{U}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I}_n \end{bmatrix} \tag{6.1.20}$$

由分块矩阵求逆公式或直接解(6.1.17)和(6.1.19)可得

$$\begin{bmatrix} \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{U}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{P} & \boldsymbol{e} \\ \boldsymbol{e}^{\mathrm{T}} & \boldsymbol{0} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}^{-1}\boldsymbol{e}(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e})^{-1} \\ -(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e})^{-1} \end{bmatrix} \tag{6.1.21}$$

事实上只要注意关系 $\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{P}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}_n$ 即可. 这引出

$$\boldsymbol{A} = [\boldsymbol{A}_1, \cdots, \boldsymbol{A}_n] = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e})^{-1}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1} \tag{6.1.22}$$

引入定义

$$\boldsymbol{\mu} = \begin{bmatrix} \hat{\boldsymbol{x}}_1 \\ \vdots \\ \hat{\boldsymbol{x}}_L \end{bmatrix} \tag{6.1.23}$$

则有关系

$$\hat{\boldsymbol{x}}_0 = \boldsymbol{A}\boldsymbol{\mu}, \quad \mathrm{E}\boldsymbol{\mu} = \boldsymbol{e}\mathrm{E}\boldsymbol{x} \tag{6.1.24}$$

于是有

$$\mathrm{E}\hat{\boldsymbol{x}}_0 = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e})^{-1}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e}\mathrm{E}\boldsymbol{x} = \mathrm{E}\boldsymbol{x} \tag{6.1.25}$$

即融合估计 $\hat{x}_0$ 是无偏的. 注意关系

$$\boldsymbol{x} = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e})^{-1}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e}\boldsymbol{x} \tag{6.1.26}$$

$$\tilde{\boldsymbol{x}}_0 = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e})^{-1}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(\boldsymbol{e}\boldsymbol{x} - \boldsymbol{\mu}) \tag{6.1.27}$$

这引出最优融合误差方差阵 $\boldsymbol{P}_0 = \mathrm{E}[\tilde{\boldsymbol{x}}_0\tilde{\boldsymbol{x}}_0^{\mathrm{T}}]$ 为

$$\boldsymbol{P}_0 = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e})^{-1}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{P}\boldsymbol{P}^{-1}\boldsymbol{e}(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e})^{-1} = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{e})^{-1} \tag{6.1.28}$$

在(6.1.2)中置 $\boldsymbol{A}_i = \boldsymbol{I}_n, \boldsymbol{A}_j = \boldsymbol{0}(j \neq i)$,由(6.1.4)引出

$$\mathrm{tr}\boldsymbol{P}_0 \leqslant \mathrm{tr}\boldsymbol{P}_i, \quad i = 1, \cdots, L \tag{6.1.29}$$

这表明除了特殊情形(等号成立),融合估计精度高于每个局部估计精度.

用 Schwartz 不等式可以证明[14]:当且仅当 $\boldsymbol{P}_i = \boldsymbol{P}_{ji}(j = 1, \cdots, L)$ 时,有 $\boldsymbol{P}_0 = \boldsymbol{P}_i$,此时在不等式(6.1.29)中等号成立. 文献[14]用 Schwartz 还证明了比(6.1.29)更强的结果:$\boldsymbol{P}_0 \leqslant \boldsymbol{P}_i$.

上述推导可概括为如下定理.

【定理 6.1.1】[14] （按矩阵加权线性最小方差最优融合公式）已知基于 L 个传感器的随机向量 $\boldsymbol{x} \in R^n$ 的 L 个无偏估计 $\hat{\boldsymbol{x}}_i, i = 1, \cdots, L$，且已知估计误差协方差阵 $\boldsymbol{P}_{ij}, i, j = 1, \cdots, L$，则极小化性能指标(6.1.4)的按矩阵加权线性最小方差无偏融合估计 $\hat{\boldsymbol{x}}_0$ 为

$$\hat{\boldsymbol{x}}_0 = \sum_{i=1}^{L} \boldsymbol{A}_i \hat{\boldsymbol{x}}_i \tag{6.1.30}$$

最优加权阵为

$$[\boldsymbol{A}_1, \cdots, \boldsymbol{A}_L] = (\boldsymbol{e}^{\mathrm{T}} \boldsymbol{P}^{-1} \boldsymbol{e})^{-1} \boldsymbol{e}^{\mathrm{T}} \boldsymbol{P}^{-1} \tag{6.1.31}$$

最优融合估计误差方差阵 $\boldsymbol{P}_0$ 为

$$\boldsymbol{P}_0 = (\boldsymbol{e}^{\mathrm{T}} \boldsymbol{P}^{-1} \boldsymbol{e})^{-1} \tag{6.1.32}$$

且有关系

$$\boldsymbol{P}_0 \leqslant \boldsymbol{P}_i, \quad \mathrm{tr}\boldsymbol{P}_0 \leqslant \mathrm{tr}\boldsymbol{P}_i, \quad i = 1, \cdots, L \tag{6.1.33}$$

【推论 6.1.1】[9] 在定理 6.1.1 条件下，若 $\boldsymbol{P}_{ij} = \boldsymbol{0}(i \neq j)$，即局部估计误差不相关，则最优融合估计为(6.1.30)，其中最优加权阵 $\boldsymbol{A}_i$ 为

$$\boldsymbol{A}_i = \Big(\sum_{i=1}^{L} \boldsymbol{P}_i^{-1}\Big)^{-1} \boldsymbol{P}_i^{-1}, \quad i = 1, \cdots, L \tag{6.1.34}$$

其中 $\boldsymbol{P}_i = \boldsymbol{P}_{ii}$，且有最优融合误差方差阵 $\boldsymbol{P}_0$ 为

$$\boldsymbol{P}_0 = \Big(\sum_{i=1}^{L} \boldsymbol{P}_i^{-1}\Big)^{-1} \tag{6.1.35}$$

且融合估计的精度高于每个局部估计精度，即

$$\boldsymbol{P}_0 < \boldsymbol{P}_i, \quad i = 1, \cdots, L \tag{6.1.36}$$

$$\mathrm{tr}\boldsymbol{P}_0 < \mathrm{tr}\boldsymbol{P}_i \tag{6.1.37}$$

证明 由定理 6.1.1 直接得(6.1.34)和(6.1.35). 现在证明(6.1.36). 由(6.1.35)有

$$\boldsymbol{P}_0^{-1} = \boldsymbol{P}_1^{-1} + \boldsymbol{P}_2^{-1} + \cdots + \boldsymbol{P}_L^{-1} \tag{6.1.38}$$

设 $\boldsymbol{P}_i > 0, i = 1, \cdots, L$，则有

$$\boldsymbol{P}_0^{-1} > \boldsymbol{P}_i^{-1} \tag{6.1.39}$$

这引出[40]

$$\boldsymbol{P}_0 < \boldsymbol{P}_i \tag{6.1.40}$$

从而引出 $\mathrm{tr}\boldsymbol{P}_0 < \mathrm{tr}\boldsymbol{P}_i$. 证毕. □

【推论 6.1.2】[15] 对两传感器情形($L = 2$)，最优融合估计为

$$\hat{\boldsymbol{x}}_0 = \boldsymbol{A}_1 \hat{\boldsymbol{x}}_1 + \boldsymbol{A}_2 \hat{\boldsymbol{x}}_2 \tag{6.1.41}$$

最优加权阵为

$$\begin{aligned}\boldsymbol{A}_1 &= (\boldsymbol{P}_2 - \boldsymbol{P}_{21})(\boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_{12} - \boldsymbol{P}_{21})^{-1}, \\ \boldsymbol{A}_2 &= (\boldsymbol{P}_1 - \boldsymbol{P}_{12})(\boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_{12} - \boldsymbol{P}_{21})^{-1}\end{aligned} \tag{6.1.42}$$

最优融合误差方差阵为

$$\boldsymbol{P}_0 = \boldsymbol{P}_1 - (\boldsymbol{P}_1 - \boldsymbol{P}_{12})(\boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_{12} - \boldsymbol{P}_{21})^{-1}(\boldsymbol{P}_1 - \boldsymbol{P}_{12})^{\mathrm{T}} \tag{6.1.43}$$

或

$$\boldsymbol{P}_0 = \boldsymbol{P}_2 - (\boldsymbol{P}_2 - \boldsymbol{P}_{21})(\boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_{12} - \boldsymbol{P}_{21})^{-1}(\boldsymbol{P}_2 - \boldsymbol{P}_{21})^{\mathrm{T}} \tag{6.1.44}$$

且有关系 $\mathrm{tr}\boldsymbol{P}_0 \leqslant \mathrm{tr}\boldsymbol{P}_i, i = 1,2$.

【推论 6.1.3】 对两传感器情形，若 $\boldsymbol{P}_{12} = \boldsymbol{0}$，即局部估计误差不相关，则有按矩阵加权线性最小方差融合公式

$$\hat{\boldsymbol{x}}_0 = \boldsymbol{A}_1\hat{\boldsymbol{x}}_1 + \boldsymbol{A}_2\hat{\boldsymbol{x}}_2 \tag{6.1.45}$$

$$\boldsymbol{A}_1 = \boldsymbol{P}_2(\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1}, \quad \boldsymbol{A}_2 = \boldsymbol{P}_1(\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1} \tag{6.1.46}$$

最优融合误差方差阵为

$$\boldsymbol{P}_0 = \boldsymbol{P}_1 - \boldsymbol{P}_1(\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1}\boldsymbol{P}_1 \tag{6.1.47}$$

或

$$\boldsymbol{P}_0 = \boldsymbol{P}_2 - \boldsymbol{P}_2(\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1}\boldsymbol{P}_2 \tag{6.1.48}$$

且有关系 $\mathrm{tr}\boldsymbol{P}_0 \leqslant \mathrm{tr}\boldsymbol{P}_1, \mathrm{tr}\boldsymbol{P}_0 \leqslant \mathrm{tr}\boldsymbol{P}_2$. $\boldsymbol{P}_0$ 的更简单公式为

$$\boldsymbol{P}_0 = \boldsymbol{P}_1(\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1}\boldsymbol{P}_2 \tag{6.1.49}$$

或

$$\boldsymbol{P}_0 = \boldsymbol{P}_2(\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1}\boldsymbol{P}_1 \tag{6.1.50}$$

证明 由推论 6.1.2 引出(6.1.45) ~ (6.1.48).注意(6.1.47)可写为

$$\begin{aligned}\boldsymbol{P}_0 = \boldsymbol{P}_1[\boldsymbol{I}_n - (\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1}\boldsymbol{P}_1] = \boldsymbol{P}_1[(\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1}(\boldsymbol{P}_1 + \boldsymbol{P}_2) - \\ (\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1}\boldsymbol{P}_1] = \boldsymbol{P}_1(\boldsymbol{P}_1 + \boldsymbol{P}_2)^{-1}\boldsymbol{P}_2\end{aligned} \tag{6.1.51}$$

同理有(6.1.50)成立. □

【推论 6.1.4】 在推论 6.1.3 情形下，若 $\boldsymbol{P}_{12} = \boldsymbol{0}$，则最优融合估计为

$$\hat{\boldsymbol{x}}_0 = \boldsymbol{A}_1\hat{\boldsymbol{x}}_1 + \boldsymbol{A}_2\hat{\boldsymbol{x}}_2 \tag{6.1.52}$$

最优加权阵为 $$\boldsymbol{A}_1 = (\boldsymbol{P}_1^{-1} + \boldsymbol{P}_2^{-1})\boldsymbol{P}_1^{-1}, \quad \boldsymbol{A}_2 = (\boldsymbol{P}_1^{-1} + \boldsymbol{P}_2^{-1})\boldsymbol{P}_2^{-1} \tag{6.1.53}$$

最优融合误差方差阵为

$$\boldsymbol{P}_0 = (\boldsymbol{P}_1^{-1} + \boldsymbol{P}_2^{-1})^{-1} \tag{6.1.54}$$

且有精度关系 $\mathrm{tr}\boldsymbol{P}_0 < \mathrm{tr}\boldsymbol{P}_1, \mathrm{tr}\boldsymbol{P}_0 < \mathrm{tr}\boldsymbol{P}_2$.

证明 由推论 6.1.1 得证. □

【注 6.1.1】 按矩阵加权融合公式(6.1.31)和(6.1.32)是新近作者在文献[14]中提出的，它相同于在局部估计误差服从联合正态分布假设下的极大似然融合估计算法[12]，但定理 6.1.1 避免了正态分布的假设.这里的推导稍不同于文献[14]的推导.

【注 6.1.2】 带 Lagrange 乘子向量 $\boldsymbol{\lambda}_i = [\lambda_{i1}, \cdots, \lambda_{2n}]$ 的辅助函数(6.1.11)可简化为带 Lagrange 乘子矩阵 $\boldsymbol{\Lambda}$ 的辅助函数

$$F = J + \mathrm{tr}(\boldsymbol{A}\boldsymbol{e} - \boldsymbol{I}_n)\boldsymbol{\Lambda} \tag{6.1.55}$$

其中 Lagrange 乘子矩阵 $\boldsymbol{\Lambda}$ 由(6.1.16)定义，事实上，由(6.1.11)有

$$F = J + \mathrm{tr}\sum_{i=1}^{n}(\boldsymbol{A}\boldsymbol{e} - \boldsymbol{I}_n)\boldsymbol{e}_i\boldsymbol{\lambda}_i = J + \mathrm{tr}(\boldsymbol{A}\boldsymbol{e} - \boldsymbol{I}_n)[\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n]\begin{bmatrix}\boldsymbol{\lambda}_1 \\ \vdots \\ \boldsymbol{\lambda}_n\end{bmatrix} \tag{6.1.56}$$

注意 $[\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n] = \boldsymbol{I}_n$，于是有

$$F = J + \mathrm{tr}(\boldsymbol{A}\boldsymbol{e} - \boldsymbol{I}_n)\boldsymbol{\Lambda} \tag{6.1.57}$$

容易验证采用辅助函数(6.1.57)同样引出定理 6.1.1 的结果.

6.1.2 按标量加权线性最小方差最优信息融合准则和算法

由(6.1.31)看到,计算最优加权阵要求计算 $Ln \times Ln$ 协方差阵 $\boldsymbol{P}$ 的逆矩阵,它由(6.1.8)定义.当 n 和 L 较大时,计算负担大,对实时应用是不方便的.为了克服这个缺点,本节介绍由文献[13,15,16,39]提出的标量加权最优融合算法,它只需要计算加权系数,只要求计算一个 $L \times L$ 矩阵的逆矩阵,显著地减小了计算负担,适于实时应用.

【定理 6.1.2】 (按标量加权线性最小方差最优融合公式)设 $\hat{\boldsymbol{x}}_i, i = 1,\cdots,L$,为对随机向量 $\boldsymbol{x} \in R^n$ 的 L 个无偏估计,设估计误差 $\tilde{\boldsymbol{x}}_i = \boldsymbol{x} - \hat{\boldsymbol{x}}_i$ 与 $\tilde{\boldsymbol{x}}_j = \boldsymbol{x} - \hat{\boldsymbol{x}}_j$ 的协方差阵 $\boldsymbol{P}_{ij}$,$i,j = 1,\cdots,L$,是已知的,且记 $\boldsymbol{P}_i = \boldsymbol{P}_{ii}$,则在线性最小方差意义下按标量加权最优融合无偏估计为

$$\hat{\boldsymbol{x}}_0 = \sum_{i=1}^{L} a_i \hat{\boldsymbol{x}}_i \tag{6.1.58}$$

其中最优加权系数行向量 $\boldsymbol{a} = [a_1, a_2, \cdots, a_L]$ 为

$$\boldsymbol{a} = \frac{\boldsymbol{e}^{\mathrm{T}} \boldsymbol{P}_{\mathrm{tr}}^{-1}}{\boldsymbol{e}^{\mathrm{T}} \boldsymbol{P}_{\mathrm{tr}}^{-1} \boldsymbol{e}} \tag{6.1.59}$$

其中定义 $L \times L$ 矩阵 $\boldsymbol{P}_{\mathrm{tr}}$ 和 $L \times 1$ 列向量 $\boldsymbol{e}$ 为

$$\boldsymbol{P}_{\mathrm{tr}} = \begin{bmatrix} \mathrm{tr}\boldsymbol{P}_{11} & \cdots & \mathrm{tr}\boldsymbol{P}_{1L} \\ \vdots & & \vdots \\ \mathrm{tr}\boldsymbol{P}_{L1} & \cdots & \mathrm{tr}\boldsymbol{P}_{LL} \end{bmatrix}, \quad \boldsymbol{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \tag{6.1.60}$$

相应的最优融合估计误差方差阵 P_0 为

$$\boldsymbol{P}_0 = \sum_{i=1}^{L} \sum_{j=1}^{L} a_i a_j \boldsymbol{P}_{ij} \tag{6.1.61}$$

且有关系

$$\mathrm{tr}\boldsymbol{P}_0 \leqslant \mathrm{tr}\boldsymbol{P}_i, \quad i = 1,\cdots,L \tag{6.1.62}$$

证明 由无偏性假设有 $\mathrm{E}\hat{\boldsymbol{x}}_0 = \mathrm{E}\boldsymbol{x}, \mathrm{E}\hat{\boldsymbol{x}}_i = \mathrm{E}\boldsymbol{x}$,从而对(6.1.58)两边取数学期望运算引出约束关系

$$\sum_{i=1}^{L} a_i = 1 \tag{6.1.63}$$

于是有最优融合误差表达式

$$\tilde{\boldsymbol{x}}_0 = \boldsymbol{x} - \hat{\boldsymbol{x}}_0 = \sum_{i=1}^{L} a_i (\boldsymbol{x} - \hat{\boldsymbol{x}}_i) = \sum_{i=1}^{L} a_i \tilde{\boldsymbol{x}}_i \tag{6.1.64}$$

于是有融合估计误差方差阵 $\boldsymbol{P}_0$ 为

$$\boldsymbol{P}_0 = \mathrm{E}[\tilde{\boldsymbol{x}}_0 \tilde{\boldsymbol{x}}_0^{\mathrm{T}}] = \sum_{i=1}^{L} \sum_{j=1}^{L} a_i a_j \boldsymbol{P}_{ij} \tag{6.1.65}$$

最优加权系数向量应在线性最小方差意义下极小化性能指标

$$J = \mathrm{tr}\boldsymbol{P}_0 \tag{6.1.66}$$

由 $\boldsymbol{P}_{\mathrm{tr}}$ 的定义(6.1.60)和 $\boldsymbol{a}$ 的定义,由上两式它可表为

$$J = \boldsymbol{a}\boldsymbol{P}_{\mathrm{tr}}\boldsymbol{a}^{\mathrm{T}} \tag{6.1.67}$$

约束条件(6.1.63)成为

$$ae = 1 \tag{6.1.68}$$

应用 Lagrange 乘子法,引入辅助函数

$$F = J + \lambda(ae - 1) \tag{6.1.69}$$

其中 λ 为 Lagrange 乘数.置 $\partial F/\partial a = 0$ 得

$$2aP_{\text{tr}} + \lambda e^{\text{T}} = 0 \tag{6.1.70}$$

对上式转置有

$$P_{\text{tr}}a^{\text{T}} + e\lambda/2 = 0 \tag{6.1.71}$$

置 $\partial F/\partial\lambda = 0$ 得 $ae - 1 = 0$,置转后有

$$e^{\text{T}}a^{\text{T}} - 1 = 0 \tag{6.1.72}$$

上两式可写成矩阵方程形式

$$\begin{bmatrix} P_{\text{tr}} & e \\ e^{\text{T}} & 0 \end{bmatrix}\begin{bmatrix} a^{\text{T}} \\ \lambda/2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{6.1.73}$$

由分块矩阵求逆公式有

$$\begin{bmatrix} a^{\text{T}} \\ \lambda/2 \end{bmatrix} = \begin{bmatrix} P_{\text{tr}} & e \\ e^{\text{T}} & 0 \end{bmatrix}^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} P_{\text{tr}}^{-1}e(e^{\text{T}}P_{\text{tr}}^{-1}e)^{-1} \\ -(e^{\text{T}}P_{\text{tr}}e)^{-1} \end{bmatrix} \tag{6.1.74}$$

这引出

$$a^{\text{T}} = P_{\text{tr}}^{-1}e(e^{\text{T}}P_{\text{tr}}^{-1}e)^{-1} \tag{6.1.75}$$

它引出(6.1.59).在(6.1.58)中取 $a_i = 1, a_j = 0 (j \neq i)$,立刻有 $\text{tr}P_0 \leqslant \text{tr}P_i, i = 1, \cdots, L$.证毕. □

【推论 6.1.5】[13] 在定理 6.1.2 条件下,若局部估计误差是不相关的,即 $P_{ij} = 0 (i \neq j)$,则有按标量加权最优融合公式

$$\hat{x}_0 = \sum_{i=1}^{L} a_i\hat{x}_i \tag{6.1.76}$$

$$a_i = \left(\sum_{i=1}^{L}\frac{1}{\text{tr}P_i}\right)^{-1}\frac{1}{\text{tr}P_i}, \quad i = 1, \cdots, L \tag{6.1.77}$$

最优融合误差方差阵为

$$P_0 = \sum_{i=1}^{L} a_i^2 P_i \tag{6.1.78}$$

且有关系 $\text{tr}P_0 \leqslant \text{tr}P_i, i = 1, \cdots, L$.

【推论 6.1.6】[15] 对两传感器系统($L = 2$),在定理 6.1.2 条件下,按标量加权最优融合估计算法为

$$\hat{x}_0 = a_1\hat{x}_1 + a_2\hat{x}_2 \tag{6.1.79}$$

$$a_1 = \frac{\text{tr}P_2 - \text{tr}P_{12}}{\text{tr}P_1 + \text{tr}P_2 - 2\text{tr}P_{12}}, \quad a_2 = \frac{\text{tr}P_1 - \text{tr}P_{12}}{\text{tr}P_1 + \text{tr}P_2 - 2\text{tr}P_{12}} \tag{6.1.80}$$

最优融合估计误差方差阵为

$$P_0 = a_1^2P_1 + a_2^2P_2 + 2a_1a_2P_{12} \tag{6.1.81}$$

且有关系 $\text{tr}P_0 \leqslant \text{tr}P_i, i = 1, 2$.

6.1.3 按对角阵加权线性最小方差最优融合估计准则和算法

由标量加权融合公式(6.1.58)看到,它等价于各局部估计的分量按相同系数加权的融合估计.为了提高分量融合估计精度,各分量应采用不相同的加权系数.这引出按分量标量加权或按对角阵加权的最优融合估计准则和算法.它是对节 6.1.2 按标量加权准则的改进,同时在计算上又不增加很大负担,且比按矩阵加权计算负担小.

【定理 6.1.3】 (按对角阵加权或按分量标量加权线性最小方差最优融合公式)设已知对随机向量 $\boldsymbol{x} \in R^n$ 的 L 个无偏估计 $\hat{\boldsymbol{x}}_i$,且已知估计误差 $\tilde{\boldsymbol{x}}_i = \boldsymbol{x} - \hat{\boldsymbol{x}}_i$ 与 $\tilde{\boldsymbol{x}}_j = \boldsymbol{x} - \hat{\boldsymbol{x}}_j$ 的协方差阵 $\boldsymbol{P}_{ij} = \mathrm{E}[\tilde{\boldsymbol{x}}_i \tilde{\boldsymbol{x}}_j^{\mathrm{T}}]$, $i, j = 1, \cdots, L$, $\hat{\boldsymbol{x}}_0$ 为对 $\boldsymbol{x}$ 的融合估计.记 $\boldsymbol{x}$, $\hat{\boldsymbol{x}}_i$ 和 $\hat{\boldsymbol{x}}_0$ 的分量表示为

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \hat{\boldsymbol{x}}_j = \begin{bmatrix} \hat{x}_{j1} \\ \vdots \\ \hat{x}_{jn} \end{bmatrix}, \quad \hat{\boldsymbol{x}}_0 = \begin{bmatrix} \hat{x}_{01} \\ \vdots \\ \hat{x}_{0n} \end{bmatrix} \tag{6.1.82}$$

则极小化按对角阵加权融合估计性能指标

$$\min J = \mathrm{tr}\boldsymbol{P}_0, \quad \boldsymbol{P}_0 = \mathrm{E}[\tilde{\boldsymbol{x}}_0 \tilde{\boldsymbol{x}}_0^{\mathrm{T}}], \quad \tilde{\boldsymbol{x}}_0 = \boldsymbol{x} - \hat{\boldsymbol{x}}_0 \tag{6.1.83}$$

的最优融合估计 $\hat{\boldsymbol{x}}_0$,

$$\hat{\boldsymbol{x}}_0 = \sum_{j=1}^{L} \boldsymbol{A}_j \hat{\boldsymbol{x}}_j \tag{6.1.84}$$

其中加权阵 $\boldsymbol{A}_j$ 为对角阵

$$\boldsymbol{A}_j = \begin{bmatrix} a_{j1} & & 0 \\ & a_{j2} & \\ & & \ddots & \\ 0 & & & a_{jn} \end{bmatrix} \tag{6.1.85}$$

等价于 $\hat{\boldsymbol{x}}_j$ 的各分量按标量加权的最优融合估计

$$\hat{x}_{0i} = \sum_{j=1}^{L} a_{ji} \hat{x}_{ji} \tag{6.1.86}$$

其中最优加权系数向量

$$\boldsymbol{a}_i = [a_{1i}, a_{2i}, \cdots, a_{Li}], \quad i = 1, \cdots, n \tag{6.1.87}$$

为

$$\boldsymbol{a}_i = \frac{\boldsymbol{e}^{\mathrm{T}} (\boldsymbol{P}^{ii})^{-1}}{\boldsymbol{e}^{\mathrm{T}} (\boldsymbol{P}^{ii})^{-1} \boldsymbol{e}}, \quad i = 1, \cdots, n \tag{6.1.88}$$

其中定义 $\boldsymbol{e}^{\mathrm{T}} = [1, 1, \cdots, 1]$,且定义 $L \times L$ 矩阵

$$\boldsymbol{P}^{ii} = \begin{bmatrix} P_{11}^{(ii)} & \cdots & P_{1L}^{(ii)} \\ \vdots & & \vdots \\ P_{L1}^{(ii)} & \cdots & P_{LL}^{(ii)} \end{bmatrix} \tag{6.1.89}$$

其中 $P_{kj}^{(ii)}$ 为 $\boldsymbol{P}_{kj}$ 的第 (i, i) 对角元素.各分量的最小融合误差 $\tilde{x}_{0i} = x_i - \hat{x}_{0i}$ 的方差 $P_{0i} = \mathrm{E}[\tilde{x}_{0i}^2]$ 为

$$P_{0i} = [\boldsymbol{e}^{\mathrm{T}} (\boldsymbol{P}^{ii})^{-1} \boldsymbol{e}]^{-1}, \quad i = 1, \cdots, n \tag{6.1.90}$$

最小融合误差平方和为

$$\mathrm{tr}\boldsymbol{P}_0 = \sum_{i=1}^{n} P_{0i} \tag{6.1.91}$$

且有关系

$$P_{0i} \leqslant \boldsymbol{P}_{jj}^{ii}, \quad i = 1,\cdots,n; j = 1,\cdots,L \tag{6.1.92}$$

或

$$\mathrm{tr}\boldsymbol{P}_0 \leqslant \mathrm{tr}\boldsymbol{P}_{jj}, \quad j = 1,\cdots,L \tag{6.1.93}$$

证明 由定义(6.1.82),记分量估计误差

$$\tilde{x}_{0i} = x_i - \hat{x}_{0i}, \quad \tilde{x}_{ji} = x_i - \hat{x}_{ji} \tag{6.1.94}$$

并记对角阵加权参数向量 $\boldsymbol{a}_i$ 为(6.1.87).在线性最小方差意义下,我们的目的是选形如(6.1.85)的对角加权阵 $\boldsymbol{A}_j$ 或形如(6.1.87)的参数向量 $\boldsymbol{a}_i$,$i = 1,\cdots,n$,极小化全局性能指标

$$J = \mathrm{E}[\tilde{\boldsymbol{x}}_0^{\mathrm{T}}\tilde{\boldsymbol{x}}_0] = \sum_{i=1}^{n}\mathrm{E}(\tilde{x}_{0i}^2) = \sum_{i=1}^{n}\mathrm{E}(\sum_{j=1}^{L} a_{ji}\tilde{x}_{ji})^2 = \sum_{i=1}^{n}\boldsymbol{J}_i \tag{6.1.95}$$

其中定义局部性能指标

$$J_i = \mathrm{E}(\sum_{j=1}^{L} a_{ji}\tilde{x}_{ji})^2 = \mathrm{E}(\tilde{x}_{0i}^2) = \sum_{k,j=1}^{L} a_{ki}\boldsymbol{P}_{kj}^{ii}a_{ji} = \boldsymbol{a}_i\boldsymbol{P}^{ii}\boldsymbol{a}_i^{\mathrm{T}} \tag{6.1.96}$$

由估值无偏性引出关系

$$\sum_{i=1}^{L}\boldsymbol{A}_i = \boldsymbol{I}_n \tag{6.1.97}$$

它等价于 n 个标量约束关系

$$\sum_{j=1}^{L} a_{ji} = 1, \quad i = 1,\cdots,n \tag{6.1.98}$$

下面我们来证明,在约束(6.1.97)下极小化全局性能指标 J 等价于 n 个独立的分量极小化问题:在约束(6.1.98)下极小化分量性能指标 J_i,$i = 1,\cdots,n$. 事实上,应用Lagrange 乘子法,引入辅助函数

$$F = J + \sum_{i=1}^{n}\lambda_i(\sum_{j=1}^{L} a_{ji} - 1) \tag{6.1.99}$$

置

$$\frac{\partial J}{\partial \boldsymbol{a}_i} = \frac{\partial J_i}{\partial \boldsymbol{a}_i} = 0, \quad \frac{\partial J}{\partial \lambda_i} = \sum_{j=1}^{L} a_{ji} - 1 = 0, \quad i = 1,\cdots,n \tag{6.1.100}$$

这恰好等价于 n 个局部极小化问题:在约束(6.1.98)下极小化形如(6.1.96)的局部优化性能指标 J_i,$i = 1,\cdots,n$.而极小化 J_i 的问题又等价于分量按标量加权的最优融合估计问题.故应用定理 6.1.2 直接引出(6.1.88)~(6.1.93).证毕. □

【注 6.1.3】 按对角阵加权融合估计本质上是对分量而言按标量加权融合估计,将全局融合估计问题转化为局部(分量)按标量加权融合估计问题.在应用中,由于分量按标量加权是具有相同物理意义的分量按标量加权,因而有直观的物理意义.这种加权方式也叫解耦融合估计.

分别记定理 6.1.1 至定理 6.1.3 所述的三种加权融合估计误差方差阵为 $\boldsymbol{P}_0^m$,$\boldsymbol{P}_0^s$ 和

$\boldsymbol{P}_0^d$，则有如下定理.

【定理 6.1.4】 定理 6.1.1 至定理 6.1.3 所述的按矩阵加权、按标量加权和按对角阵加权最优融合估计误差方差阵有关系

$$\mathrm{tr}\boldsymbol{P}_0^m \leqslant \mathrm{tr}\boldsymbol{P}_0^d \leqslant \mathrm{tr}\boldsymbol{P}_0^s \leqslant \mathrm{tr}\boldsymbol{P}_i, \quad i = 1,\cdots,L \tag{6.1.101}$$

证明 按标量加权可看成是按特殊形式对角矩阵 $\boldsymbol{A}_i = a_i\boldsymbol{I}_n$ 加权，而按对角阵加权（即按分量标量加权）可看成是按特殊矩阵 $\boldsymbol{A}_i = \mathrm{diag}(a_{i1},\cdots,a_{in})$ 加权. 第 i 个传感器子系统的估计可看成是取加权阵 $\boldsymbol{A}_i = \boldsymbol{I}_n, \boldsymbol{A}_j = \boldsymbol{0}(j \neq i)$ 的融合估计. 于是在统一的按矩阵加权融合估计框架下有(6.1.101)成立. 证毕. □

6.2 时变系统多传感器信息融合 Kalman 滤波器和预报器

考虑多传感器线性离散时变随机控制系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)\boldsymbol{u}(t) + \boldsymbol{\Gamma}(t)\boldsymbol{w}(t) \tag{6.2.1}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{H}_i(t)\boldsymbol{x}(t) + \boldsymbol{v}_i(t), \quad i = 1,\cdots,L \tag{6.2.2}$$

其中 t 为离散时间，状态 $\boldsymbol{x}(t) \in R^n$，第 i 传感器的观测 $\boldsymbol{y}_i(t) \in R^{m_i}$，控制 $\boldsymbol{u}(t) \in R^p$，$\boldsymbol{\Phi}(t)$，$\boldsymbol{B}(t)$，$\boldsymbol{\Gamma}(t)$ 和 $\boldsymbol{H}_i(t)$ 是已知的适当维数时变矩阵. $\boldsymbol{w}(t)$ 为输入噪声，$\boldsymbol{v}_i(t)$ 为第 i 传感器的观测噪声.

【假设 1】 $\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}_i(t) \in R^{m_i}$ 是带零均值的相关白噪声：

$$\mathrm{E}\left\{\begin{bmatrix} \boldsymbol{w}(t) \\ \boldsymbol{v}_i(t) \end{bmatrix}\begin{bmatrix} \boldsymbol{w}^{\mathrm{T}}(k) & \boldsymbol{v}_i^{\mathrm{T}}(k) \end{bmatrix}\right\} = \begin{bmatrix} \boldsymbol{Q}(t) & \boldsymbol{S}_i(t) \\ \boldsymbol{S}_i^{\mathrm{T}}(t) & \boldsymbol{R}_i(t) \end{bmatrix}\delta_{tk}, \quad i = 1,\cdots,L \tag{6.2.3}$$

其中 E 为均值号，T 为转置号，$\delta_{tt} = 1, \delta_{tk} = 0(t \neq k)$，$\boldsymbol{w}(t)$ 与 $\boldsymbol{v}_i(t)$ 的方差阵和相关阵各为 $\boldsymbol{Q}(t)$，$\boldsymbol{R}_i(t)$ 和 $\boldsymbol{S}_i(t)$.

【假设 2】 观测噪声 $\boldsymbol{v}_i(t)$ 与 $\boldsymbol{v}_j(t)$ 是相关的，即

$$\mathrm{E}\left\{\begin{bmatrix} \boldsymbol{v}_i(t) \\ \boldsymbol{v}_j(t) \end{bmatrix}\begin{bmatrix} \boldsymbol{v}_i^{\mathrm{T}}(k) & \boldsymbol{v}_j^{\mathrm{T}}(k) \end{bmatrix}\right\} = \begin{bmatrix} \boldsymbol{R}_i(t) & \boldsymbol{R}_{ij}(t) \\ \boldsymbol{R}_{ij}^{\mathrm{T}}(t) & \boldsymbol{R}_j(t) \end{bmatrix}\delta_{tk}, \quad i = 1,\cdots,L \tag{6.2.4}$$

且定义 $\boldsymbol{R}_i(t) = \boldsymbol{R}_{ii}(t)$.

【假设 3】 $\boldsymbol{x}(0)$ 不相关于 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}_i(t)$，$i = 1,\cdots,L$，且 $\mathrm{E}\boldsymbol{x}(0) = \boldsymbol{\mu}$，$\mathrm{cov}\boldsymbol{x}(0) = \boldsymbol{P}_0$，其中 cov 为协方差号.

【假设 4】 $\boldsymbol{u}(t)$ 是已知的时间序列，或 $\boldsymbol{u}(t)$ 是 $(\boldsymbol{y}(t), \boldsymbol{y}(t-1), \cdots)$ 的线性函数（反馈控制）.

问题是基于观测 $(\boldsymbol{y}_i(t), \boldsymbol{y}_i(t-1), \cdots, \boldsymbol{y}_i(0))$ 和 $(\boldsymbol{u}(t-1), \boldsymbol{u}(t-2), \cdots, \boldsymbol{u}(0))$ 求相应于第 i 传感器的状态 $\boldsymbol{x}(t)$ 的局部 Kalman 滤波器 $\hat{\boldsymbol{x}}_i(t \mid t)$ 和状态 $\boldsymbol{x}(t+1)$ 局部 Kalman 预报器 $\hat{\boldsymbol{x}}_i(t+1 \mid t)$，$i = 1,\cdots,L$，并在三种加权融合估计准则下分别求相应的信息融合 Kalman 滤波器 $\hat{\boldsymbol{x}}_0(t \mid t)$ 和预报器 $\hat{\boldsymbol{x}}_0(t+1 \mid t)$.

【定理 6.2.1】 系统(6.2.1)和(6.2.2)在假设 1 ~ 4 下，第 i 传感器子系统有局部最优递推 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t+1 \mid t) = \boldsymbol{\Psi}_{pi}(t)\hat{\boldsymbol{x}}_i(t \mid t-1) + \boldsymbol{B}(t)\boldsymbol{u}(t) + \boldsymbol{K}_{pi}(t)\boldsymbol{y}_i(t) \tag{6.2.5}$$

$$\boldsymbol{\Psi}_{pi}(t) = \boldsymbol{\Phi}(t) - \boldsymbol{K}_{pi}(t)\boldsymbol{H}_i(t) \tag{6.2.6}$$

$$K_{pi}(t)=[\boldsymbol{\Phi}(t)P_i(t\mid t-1)H_i^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)S_i(t)]Q_{\varepsilon i}^{-1}(t) \tag{6.2.7}$$

$$Q_{\varepsilon i}(t)=H_i(t)P_i(t\mid t-1)H_i^{\mathrm{T}}(t)+R_i(t) \tag{6.2.8}$$

其中预报误差方差阵 $P_i(t+1\mid t)$ 满足 Riccati 方程

$$\begin{aligned}P_i(t+1\mid t)=&\boldsymbol{\Phi}(t)P_i(t\mid t-1)\boldsymbol{\Phi}^{\mathrm{T}}(t)-[\boldsymbol{\Phi}(t)P_i(t\mid t-1)H_i^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)S_i(t)]\times\\&[H_i(t)P_i(t\mid t-1)H_i^{\mathrm{T}}(t)+R_i(t)]^{-1}\times\\&[\boldsymbol{\Phi}(t)P_i(t\mid t-1)H_i^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)S_i(t)]^{\mathrm{T}}+\boldsymbol{\Gamma}(t)Q(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)\end{aligned} \tag{6.2.9}$$

带初值 $\hat{x}_i(0\mid-1)=\boldsymbol{\mu},P_i(0\mid-1)=P_0$.

$P_i(t+1\mid t)$ 的另一种等价算法为

$$P_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)P_i(t\mid t-1)\boldsymbol{\Psi}_{pi}^{\mathrm{T}}(t)+[\boldsymbol{\Gamma}(t),-K_{pi}(t)]\times\begin{bmatrix}Q(t)&S_i(t)\\S_i^{\mathrm{T}}(t)&R_i(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Gamma}^{\mathrm{T}}(t)\\-K_{pi}^{\mathrm{T}}(t)\end{bmatrix} \tag{6.2.10}$$

$$\begin{aligned}P_i(t+1\mid t)=&\boldsymbol{\Psi}_{pi}(t)P_i(t\mid t-1)\boldsymbol{\Psi}_{pi}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)Q(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\\&K_{pi}(t)S_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{\Gamma}(t)S_i(t)K_{pi}^{\mathrm{T}}(t)+K_{pi}(t)R_i(t)K_{pi}^{\mathrm{T}}(t)\end{aligned} \tag{6.2.11}$$

预报误差协方差阵 $P_{ij}(t+1\mid t)=\mathrm{E}[\tilde{x}_i(t+1\mid t)\tilde{x}_j^{\mathrm{T}}(t+1\mid t)]$ 为

$$P_{ij}(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)P_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t)+[\boldsymbol{\Gamma}(t),-K_{pi}(t)]\times\begin{bmatrix}Q(t)&S_j(t)\\S_i^{\mathrm{T}}(t)&R_{ij}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Gamma}^{\mathrm{T}}(t)\\-K_{pj}^{\mathrm{T}}(t)\end{bmatrix} \tag{6.2.12}$$

其中 $i,j=1,\cdots,L$,或

$$\begin{aligned}P_{ij}(t+1\mid t)=&\boldsymbol{\Psi}_{pi}(t)P_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)Q(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\\&K_{pi}(t)S_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{\Gamma}(t)S_j(t)K_{pj}^{\mathrm{T}}(t)+K_{pi}(t)R_{ij}(t)K_{pj}^{\mathrm{T}}(t)\end{aligned} \tag{6.2.13}$$

特别当 $i=j$ 时(6.2.12)化为(6.2.10),其中记 $P_{ii}(t+1\mid t)=P_i(t+1\mid t)$.初值为

$$P_{ij}(0\mid-1)=P_0,\quad i,j=1,\cdots,L \tag{6.2.14}$$

证明 由定理 3.7.1 有(6.2.5) ~ (6.2.9).由(3.7.12)有

$$\tilde{x}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)\tilde{x}_i(t\mid t-1)+\boldsymbol{\Gamma}(t)w(t)-K_{pi}(t)v_i(t) \tag{6.2.15}$$

注意 $\tilde{x}_i(t\mid t-1)$ 不相关于 $w(t)$ 和 $v_i(t),v_j(t)$,应用(6.2.3)和(6.2.4)得(6.2.10) ~ (6.2.13).由 $\hat{x}_i(0\mid-1)=\hat{x}_j(0\mid-1)=\boldsymbol{\mu}$ 引出 $P_{ij}(0\mid-1)=\mathrm{E}[(x(0)-\hat{x}_i(0\mid-1))(x(0)-\hat{x}_j(0\mid-1))^{\mathrm{T}}]=P_0$.证毕. □

引入节 3.7 的记号

$$\bar{\boldsymbol{\Phi}}_i(t)=\boldsymbol{\Phi}(t)-J_i(t)H_i(t) \tag{6.2.16}$$

$$J_i(t)=\boldsymbol{\Gamma}(t)S_i(t)R_i^{-1}(t) \tag{6.2.17}$$

$$\bar{w}_i(t)=\boldsymbol{\Gamma}(t)w(t)-J_i(t)v_i(t) \tag{6.2.18}$$

我们可以给出的不同于定理6.2.1的等价的局部 Kalman 预报器及误差方差阵和协方差阵的新算法.

【定理 6.2.2】 系统(6.2.1)和(6.2.2)在假设 1 ~ 4 下,第 i 传感器子系统有局部最优 Kalman 预报器

$$\hat{x}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)\hat{x}_i(t\mid t-1)+B(t)u(t)+K_{pi}(t)y_i(t) \tag{6.2.19}$$

$$\boldsymbol{\Psi}_{pi}(t)=\bar{\boldsymbol{\Phi}}_i(t)-\bar{\boldsymbol{K}}_{pi}(t)\boldsymbol{H}_i(t) \tag{6.2.20}$$

$$\bar{\boldsymbol{K}}_{pi}(t)=\bar{\boldsymbol{\Phi}}_i(t)\boldsymbol{K}_{fi}(t),\boldsymbol{K}_{pi}(t)=\bar{\boldsymbol{K}}_{pi}(t)+\boldsymbol{J}_i(t) \tag{6.2.21}$$

$$\boldsymbol{K}_{fi}(t)=\boldsymbol{P}(t\mid t-1)\boldsymbol{H}_i^{\mathrm{T}}(t)\boldsymbol{Q}_{\varepsilon i}^{-1}(t) \tag{6.2.22}$$

其中新息方差 $\boldsymbol{Q}_{\varepsilon i}(t)$ 由(6.2.8)定义.预报误差方差阵 $\boldsymbol{P}_i(t+1\mid t)$ 满足 Ricati 方程

$$\begin{aligned}\boldsymbol{P}_i(t+1\mid t)=&\bar{\boldsymbol{\Phi}}_i(t)[\boldsymbol{P}_i(t\mid t-1)-\boldsymbol{P}_i(t\mid t-1)\boldsymbol{H}_i^{\mathrm{T}}(t)(\boldsymbol{H}_i(t)\boldsymbol{P}_i(t\mid t-1)\times\\&\boldsymbol{H}_i^{\mathrm{T}}(t)+\boldsymbol{R}_i(t))^{-1}\boldsymbol{H}_i(t)\boldsymbol{P}_i(t\mid t-1)]\bar{\boldsymbol{\Phi}}_i^{\mathrm{T}}(t)+\\&\boldsymbol{\Gamma}(t)[\boldsymbol{Q}(t)-\boldsymbol{S}_i(t)\boldsymbol{R}_i^{-1}(t)\boldsymbol{S}_i^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t)\end{aligned} \tag{6.2.23}$$

带初值 $\boldsymbol{P}_i(0\mid-1)=\boldsymbol{P}_0$. $\boldsymbol{P}_i(t+1\mid t)$ 的另一种等价的递推公式为

$$\begin{aligned}\boldsymbol{P}_i(t+1\mid t)=&\boldsymbol{\Psi}_{pi}(t)\boldsymbol{P}_i(t\mid t-1)\boldsymbol{\Psi}_{pi}^{\mathrm{T}}(t)+\bar{\boldsymbol{K}}_{pi}(t)\boldsymbol{R}_i(t)\bar{\boldsymbol{K}}_{pi}^{\mathrm{T}}(t)+\\&\boldsymbol{\Gamma}(t)[\boldsymbol{Q}(t)-\boldsymbol{S}_i(t)\boldsymbol{R}_i^{-1}(t)\boldsymbol{S}_i^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t)\end{aligned} \tag{6.2.24}$$

带初值 $\boldsymbol{P}_i(0\mid-1)=\boldsymbol{P}_0$,其中 $\bar{\boldsymbol{K}}_{pi}(t)$ 和 $\boldsymbol{\Psi}_{pi}(t)$ 由(6.2.20) ~ (6.2.23)计算.

预报误差协方差阵 $\boldsymbol{P}_{ij}(t+1\mid t)(i\neq j)$ 可递推计算为

$$\begin{aligned}\boldsymbol{P}_{ij}(t+1\mid t)=&\boldsymbol{\Psi}_{pi}(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t)+[\boldsymbol{I}_n,-\bar{\boldsymbol{K}}_{pi}(t)]\times\\&\begin{bmatrix}\boldsymbol{S}_{ij}^{(1)}(t)&\boldsymbol{S}_{ij}^{(2)}(t)\\\boldsymbol{S}_{ij}^{(3)}(t)&\boldsymbol{S}_{ij}^{(4)}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_n\\-\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}}(t)\end{bmatrix}\end{aligned} \tag{6.2.25}$$

带初值 $\boldsymbol{P}_{ij}(0\mid-1)=\boldsymbol{P}_0,i,j=1,\cdots,L,i\neq j$,且定义

$$\begin{aligned}&\boldsymbol{S}_{ij}^{(1)}(t)=\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{J}_i(t)\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t)\boldsymbol{J}_j^{\mathrm{T}}(t)+\boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t),\\&\boldsymbol{S}_{ij}^{(2)}(t)=\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t)-\boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t),\\&\boldsymbol{S}_{ij}^{(3)}(t)=\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t),\\&\boldsymbol{S}_{ij}^{(4)}(t)=\boldsymbol{R}_{ij}(t)\end{aligned} \tag{6.2.26}$$

证明 (6.2.19) ~ (6.2.22)及(6.2.24)由定理3.7.3给出.(6.2.23)由(3.7.73)给出.下面证明(6.2.25)和(6.2.26).事实上由(3.7.51)有预报误差递推方程

$$\tilde{\boldsymbol{x}}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)\tilde{\boldsymbol{x}}_i(t\mid t-1)+\bar{\boldsymbol{w}}_i(t)-\bar{\boldsymbol{K}}_{pi}(t)\boldsymbol{v}_i(t) \tag{6.2.27}$$

它可写成

$$\tilde{\boldsymbol{x}}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)\tilde{\boldsymbol{x}}_i(t\mid t-1)+[\boldsymbol{I}_n,-\bar{\boldsymbol{K}}_{pi}(t)]\begin{bmatrix}\bar{\boldsymbol{w}}_i(t)\\\boldsymbol{v}_i(t)\end{bmatrix} \tag{6.2.28}$$

注意 $\bar{\boldsymbol{w}}_i(t)$ 和 $\boldsymbol{v}_i(t)$ 是不相关的,且不相关于 $\hat{\boldsymbol{x}}_i(t\mid t-1)$,由(6.2.27)可得(6.2.24).应用(6.2.28)可得(6.2.25),其中

$$\begin{aligned}&\boldsymbol{S}_{ij}^{(1)}(t)=\mathrm{E}[\bar{\boldsymbol{w}}_i(t)\bar{\boldsymbol{w}}_j^{\mathrm{T}}(t)],\\&\boldsymbol{S}_{ij}^{(2)}(t)=\mathrm{E}[\bar{\boldsymbol{w}}_i(t)\boldsymbol{v}_j^{\mathrm{T}}(t)],\\&\boldsymbol{S}_{ij}^{(3)}(t)=\mathrm{E}[\boldsymbol{v}_i(t)\bar{\boldsymbol{w}}_j^{\mathrm{T}}(t)],\\&\boldsymbol{S}_{ij}^{(4)}(t)=\mathrm{E}[\boldsymbol{v}_i(t)\boldsymbol{v}_j^{\mathrm{T}}(t)]\end{aligned} \tag{6.2.29}$$

由(6.2.3),(6.2.4)和(6.2.18)容易得(6.2.26).证毕. □

【定理6.2.3】 系统(6.2.1)和(6.2.2)在假设1 ~ 4下,第 i 传感器子系统有局部稳态最优 Kalman 滤波器

$$\hat{\boldsymbol{x}}_i(t+1\mid t+1)=\hat{\boldsymbol{x}}_i(t+1\mid t)+\boldsymbol{K}_{fi}(t+1)\boldsymbol{\varepsilon}_i(t+1) \tag{6.2.30}$$

$$\hat{\boldsymbol{x}}_i(t+1\mid t)=\bar{\boldsymbol{\Phi}}_i(t)\hat{\boldsymbol{x}}_i(t\mid t)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{J}_i(t)\boldsymbol{y}_i(t) \tag{6.2.31}$$

$$\boldsymbol{\varepsilon}_i(t+1)=\boldsymbol{y}_i(t+1)-\boldsymbol{H}_i(t+1)\hat{\boldsymbol{x}}_i(t+1\mid t) \tag{6.2.32}$$

$$\boldsymbol{K}_{fi}(t+1)=\boldsymbol{P}_i(t+1\mid t)\boldsymbol{H}_i^{\mathrm{T}}(t+1)\boldsymbol{Q}_{\varepsilon i}^{-1}(t+1) \tag{6.2.33}$$

$$\boldsymbol{Q}_{\varepsilon i}(t+1)=\boldsymbol{H}_i(t+1)\boldsymbol{P}_i(t+1\mid t)\boldsymbol{H}_i^{\mathrm{T}}(t+1)+\boldsymbol{R}_i(t+1) \tag{6.2.34}$$

$$\boldsymbol{P}_i(t+1\mid t)=\overline{\boldsymbol{\Phi}}_i(t)\boldsymbol{P}_i(t\mid t)\overline{\boldsymbol{\Phi}}_i^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)[\boldsymbol{Q}(t)-\boldsymbol{S}_i(t)\boldsymbol{R}_i^{-1}(t)\boldsymbol{S}_i^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{6.2.35}$$

$$\boldsymbol{P}_i(t+1\mid t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\boldsymbol{P}_i(t+1\mid t) \tag{6.2.36}$$

$$\hat{\boldsymbol{x}}_i(0\mid 0)=\boldsymbol{\mu},\quad \boldsymbol{P}_i(0\mid 0)=\boldsymbol{P}_0 \tag{6.2.37}$$

证明　见定理 3.7.4. □

【定理 6.2.4】　系统(6.2.1)和(6.2.2)在假设 1 ~ 4 下，局部滤波误差协方差阵 $\boldsymbol{P}_{ij}(t+1\mid t+1)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t+1\mid t+1)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t+1\mid t+1)]$ 有递推公式

$$\begin{aligned}\boldsymbol{P}_{ij}(t+1\mid t+1)=&\boldsymbol{\Psi}_{fi}(t+1)\boldsymbol{P}_{ij}(t\mid t)\boldsymbol{\Psi}_{fj}^{\mathrm{T}}(t+1)+\boldsymbol{\Psi}_{fi}(t+1)\boldsymbol{K}_{fi}(t)[\boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t)-\\&\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)][\boldsymbol{I}_n-\boldsymbol{K}_{fj}(t+1)\boldsymbol{H}_j(t+1)]^{\mathrm{T}}+[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\times\\&[\boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t)-\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t)]\boldsymbol{K}_{fj}^{\mathrm{T}}(t)\boldsymbol{\Psi}_{fj}^{\mathrm{T}}(t+1)+[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\times\\&[\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{J}_i(t)\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t)\boldsymbol{J}_j^{\mathrm{T}}(t)+\boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t)]\times\\&[\boldsymbol{I}_n-\boldsymbol{K}_{fj}(t+1)\boldsymbol{H}_j(t+1)]^{\mathrm{T}}+\boldsymbol{K}_{fi}(t+1)\boldsymbol{R}_{ij}(t+1)\boldsymbol{K}_{fj}^{\mathrm{T}}(t+1)\end{aligned} \tag{6.2.38}$$

其中 $i,j=1,\cdots,L,i\neq j$ 带初值 $P_{ij}(0\mid 0)=\boldsymbol{P}_0$，且定义

$$\boldsymbol{\Psi}_{fi}(t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\overline{\boldsymbol{\Phi}}_i(t) \tag{6.2.39}$$

(6.2.38)的另一种形式为

$$\begin{aligned}\boldsymbol{P}_{ij}(t+1\mid t+1)=&[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\{\overline{\boldsymbol{\Phi}}_i(t)\boldsymbol{P}_{ij}(t\mid t)\overline{\boldsymbol{\Phi}}_j^{\mathrm{T}}(t)+\\&\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{J}_j(t)\boldsymbol{R}_j(t)\boldsymbol{J}_j^{\mathrm{T}}(t)-\boldsymbol{J}_i(t)\boldsymbol{R}_i(t)\boldsymbol{J}_i^{\mathrm{T}}(t)+\boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t)+\\&\overline{\boldsymbol{\Phi}}_i(t)\boldsymbol{K}_{fi}(t)[\boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t)-\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)]+[\boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t)-\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t)]\boldsymbol{K}_j^{\mathrm{T}}(t)\overline{\boldsymbol{\Phi}}_j^{\mathrm{T}}(t)\}\times\\&[\boldsymbol{I}_n-\boldsymbol{K}_{fJ}(t+1)\boldsymbol{H}_J(t+1)]^{\mathrm{T}}+\boldsymbol{K}_{fi}(t+1)\boldsymbol{R}_{ij}(t+1)\boldsymbol{K}_{fj}^{\mathrm{T}}(t+1)\end{aligned} \tag{6.2.40}$$

证明　由(3.7.48)有局部滤波误差递推公式

$$\begin{aligned}\tilde{\boldsymbol{x}}_i(t+1\mid t+1)=&\boldsymbol{\Psi}_{fi}(t+1)\tilde{\boldsymbol{x}}_i(t\mid t)+[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\overline{\boldsymbol{w}}_i(t)-\\&\boldsymbol{K}_{fi}(t+1)\boldsymbol{v}_i(t+1)\end{aligned} \tag{6.2.41}$$

$$\begin{aligned}\tilde{\boldsymbol{x}}_j(t+1\mid t+1)=&\boldsymbol{\Psi}_{fj}(t+1)\tilde{\boldsymbol{x}}_j(t\mid t)+[\boldsymbol{I}_n-\boldsymbol{K}_{fj}(t+1)\boldsymbol{H}_j(t+1)]\overline{\boldsymbol{w}}_j(t)-\\&\boldsymbol{K}_{fi}(t+1)\boldsymbol{v}_j(t+1)\end{aligned} \tag{6.2.42}$$

于是有

$$\begin{aligned}\boldsymbol{P}_{ij}(t+1\mid t+1)=&\boldsymbol{\Psi}_{fi}(t+1)\boldsymbol{P}_{ij}(t\mid t)\boldsymbol{\Psi}_{fj}^{\mathrm{T}}(t+1)+\boldsymbol{\Psi}_{fi}(t+1)\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t)\overline{\boldsymbol{w}}_j^{\mathrm{T}}(t)]\times\\&[\boldsymbol{I}_n-\boldsymbol{K}_{fj}(t+1)\boldsymbol{H}_j(t+1)]^{\mathrm{T}}+[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\times\\&\mathrm{E}[\overline{\boldsymbol{w}}_i(t)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t\mid t)]\boldsymbol{\Psi}_{fj}^{\mathrm{T}}(t+1)+[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\times\\&\mathrm{E}[\overline{\boldsymbol{w}}_i(t)\overline{\boldsymbol{w}}_j^{\mathrm{T}}(t)][\boldsymbol{I}_n-\boldsymbol{K}_{fj}(t+1)\boldsymbol{H}_j(t+1)]^{\mathrm{T}}+\\&\boldsymbol{K}_{fi}(t+1)\boldsymbol{R}_{ij}(t+1)\boldsymbol{K}_{fj}^{\mathrm{T}}(t+1)\end{aligned} \tag{6.2.43}$$

应用(6.2.3)，(6.2.4)，(6.2.17)和(6.2.18)容易求得

$$\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t)\overline{\boldsymbol{w}}_j^{\mathrm{T}}(t)]=\mathrm{E}[-\boldsymbol{K}_{fi}(t)\boldsymbol{v}_i(t)(\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{J}_j(t)\boldsymbol{v}_j(t))^{\mathrm{T}}]=$$

$$-\boldsymbol{K}_{fi}(t)\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)+\boldsymbol{K}_{fi}(t)\boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t) \tag{6.2.44}$$

$$\mathrm{E}[\bar{\boldsymbol{w}}_i(t)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t\mid t)]=\mathrm{E}[(\boldsymbol{\Gamma}(t)\boldsymbol{w}(t)-\boldsymbol{J}_i(t)\boldsymbol{v}_i(t))(-\boldsymbol{K}_{fj}(t)\boldsymbol{v}_j(t))^{\mathrm{T}}]=$$
$$-\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t)\boldsymbol{K}_{fj}^{\mathrm{T}}(t)+\boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t)\boldsymbol{K}_{fj}^{\mathrm{T}}(t) \tag{6.2.45}$$

$$\mathrm{E}[\bar{\boldsymbol{w}}_i(t)\bar{\boldsymbol{w}}_j^{\mathrm{T}}(t)]=\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{J}_i(t)\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-$$
$$\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t)\boldsymbol{J}_j^{\mathrm{T}}(t)+\boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t) \tag{6.2.46}$$

将它们代入(6.2.43)得(6.2.38).应用定义(6.2.17)和(6.2.39)由(6.2.38)可得(6.2.40).证毕. □

【注 6.2.1】 在(6.2.38)中置 $i=j$ 或直接应用(6.2.41)可得不同于(6.2.36)的 $\boldsymbol{P}_i(t+1\mid t+1)$ 的另一种算法

$$\boldsymbol{P}_i(t+1\mid t+1)=\boldsymbol{\Psi}_{fi}(t+1)\boldsymbol{P}_i(t\mid t)\boldsymbol{\Psi}_{fi}^{\mathrm{T}}(t+1)+[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\times$$
$$\boldsymbol{\Gamma}(t)[\boldsymbol{Q}(t)-\boldsymbol{S}_i(t)\boldsymbol{R}_i^{-1}(t)\boldsymbol{S}_i^{\mathrm{T}}(t)]\boldsymbol{\Gamma}^{\mathrm{T}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\times$$
$$\boldsymbol{H}_i(t+1)]^{\mathrm{T}}+\boldsymbol{K}_{fi}(t+1)\boldsymbol{R}_i(t+1)\boldsymbol{K}_{fi}^{\mathrm{T}}(t+1) \tag{6.2.47}$$

【定理 6.2.5】 系统(6.2.1)和(6.2.2)在假设1～4下,有局部滤波误差协方差阵 $\boldsymbol{P}_{ij}(t\mid t)$ 为

$$\boldsymbol{P}_{ij}(t\mid t)=[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t)\boldsymbol{H}_i(t)]\boldsymbol{P}_{ij}(t\mid t-1)[\boldsymbol{I}_n-\boldsymbol{K}_{fj}(t)\boldsymbol{H}_j(t)]^{\mathrm{T}}+$$
$$\boldsymbol{K}_{fi}(t)\boldsymbol{R}_{ij}(t)\boldsymbol{K}_{fj}^{\mathrm{T}}(t) \tag{6.2.48}$$

证明 将(6.2.2)代入(6.2.32)有

$$\boldsymbol{\varepsilon}_i(t)=\boldsymbol{H}_i(t)\tilde{\boldsymbol{x}}_i(t\mid t-1)+\boldsymbol{v}_i(t) \tag{6.2.49}$$

将它代入(6.2.30)引出

$$\tilde{\boldsymbol{x}}_i(t\mid t)=[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t)\boldsymbol{H}_i(t)]\tilde{\boldsymbol{x}}_i(t\mid t-1)-\boldsymbol{K}_{fi}(t)\boldsymbol{v}_i(t) \tag{6.2.50}$$

由此引出(6.2.48).证毕. □

【推论 6.2.1】 时变系统(6.2.1)和(6.2.2)在假设1～3下,若噪声是不相关的,即 $\boldsymbol{S}_i(t)=\boldsymbol{0}$,且 $\boldsymbol{R}_{ij}(t)=\boldsymbol{0}(i\neq j)$,且 $\boldsymbol{u}(t)=\boldsymbol{0}$,则第 i 传感器子系统有局部最优 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)\hat{\boldsymbol{x}}_i(t\mid t-1)+\boldsymbol{K}_p(t)\boldsymbol{y}_i(t),$$
$$\boldsymbol{\Psi}_{pi}(t)=\boldsymbol{\Phi}(t)-\boldsymbol{K}_{pi}(t)\boldsymbol{H}_i(t),$$
$$\boldsymbol{K}_{pi}(t)=\boldsymbol{\Phi}(t)\boldsymbol{P}_i(t\mid t-1)\boldsymbol{H}_i^{\mathrm{T}}(t)\boldsymbol{Q}_{\varepsilon i}^{-1}(t),$$
$$\boldsymbol{Q}_{\varepsilon i}(t)=\boldsymbol{H}_i(t)\boldsymbol{P}_i(t\mid t-1)\boldsymbol{H}_i^{\mathrm{T}}(t)+\boldsymbol{R}_i(t) \tag{6.2.51}$$

其中预报误差方差阵 $\boldsymbol{P}_i(t+1\mid t)$ 满足 Riccati 方程

$$\boldsymbol{P}_i(t+1\mid t)=\boldsymbol{\Phi}(t)[\boldsymbol{P}_i(t\mid t-1)-\boldsymbol{P}_i(t\mid t-1)\boldsymbol{H}_i^{\mathrm{T}}(t)(\boldsymbol{H}_i(t)\boldsymbol{P}_i(t\mid t-1)\boldsymbol{H}_i^{\mathrm{T}}(t)+$$
$$\boldsymbol{R}_i(t))^{-1}\boldsymbol{H}_i(t)\boldsymbol{P}_i(t\mid t-1)]\boldsymbol{\Phi}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{6.2.52}$$

带初值 $\hat{\boldsymbol{x}}_i(0\mid-1)=\boldsymbol{\mu}$, $\boldsymbol{P}_i(0\mid-1)=\boldsymbol{P}_0$, $i=1,\cdots,L$.

$\boldsymbol{P}_i(t+1\mid t)$ 的另一种等价算法为

$$\boldsymbol{P}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)\boldsymbol{P}_i(t\mid t-1)\boldsymbol{\Psi}_{pi}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)+\boldsymbol{K}_{pi}(t)\boldsymbol{R}_i(t)\boldsymbol{K}_{pi}^{\mathrm{T}}(t) \tag{6.2.53}$$

预报误差协方差阵 $\boldsymbol{P}_{ij}(t+1\mid t)$ 有递推公式

$$\boldsymbol{P}_{ij}(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{6.2.54}$$

其中 $i,j=1,\cdots,L,i\neq j$，初值 $\boldsymbol{P}_{ij}(0\mid-1)=\boldsymbol{P}_0$.

【推论 6.2.2】 时变系统(6.2.1)和(6.2.2)在假设 1 ~ 3 下，若 $\boldsymbol{S}_i(t)=\boldsymbol{0},\boldsymbol{R}_{ij}(t)=\boldsymbol{0}(i\neq j)$，且 $\boldsymbol{u}(t)=\boldsymbol{0}$，则第 i 传感器子系统有局部最优 Kalman 滤波器

$$\hat{\boldsymbol{x}}_i(t+1\mid t+1)=\boldsymbol{\Psi}_{fi}(t+1)\hat{\boldsymbol{x}}_i(t\mid t)+\boldsymbol{K}_{fi}(t+1)\boldsymbol{y}_i(t+1) \tag{6.2.55}$$

$$\boldsymbol{\Psi}_{fi}(t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\boldsymbol{\Phi}(t) \tag{6.2.56}$$

$$\boldsymbol{K}_{fi}(t+1)=\boldsymbol{P}_i(t+1\mid t)\boldsymbol{H}_i^{\mathrm{T}}(t+1)\boldsymbol{Q}_{\varepsilon i}^{-1}(t+1) \tag{6.2.57}$$

$$\boldsymbol{Q}_{\varepsilon i}(t+1)=\boldsymbol{H}_i(t+1)\boldsymbol{P}_i(t+1\mid t)\boldsymbol{H}_i^{\mathrm{T}}(t+1)+\boldsymbol{R}_i(t+1) \tag{6.2.58}$$

$$\boldsymbol{P}_i(t+1\mid t)=\boldsymbol{\Phi}_i(t)\boldsymbol{P}_i(t\mid t)\boldsymbol{\Phi}_i^{\mathrm{T}}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) \tag{6.2.59}$$

$$\boldsymbol{P}_i(t+1\mid t+1)=[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\boldsymbol{P}_i(t+1\mid t) \tag{6.2.60}$$

$$\hat{\boldsymbol{x}}_i(0\mid 0)=\boldsymbol{\mu},\quad \boldsymbol{P}_i(0\mid 0)=\boldsymbol{P}_0 \tag{6.2.61}$$

且 $\boldsymbol{P}_i(t+1\mid t+1)$ 的另一种算法为

$$\begin{aligned}\boldsymbol{P}_i(t+1\mid t+1)=&\boldsymbol{\Psi}_{fi}(t+1)\boldsymbol{P}_i(t\mid t)\boldsymbol{\Psi}_{fi}^{\mathrm{T}}(t+1)+[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\times\\&\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]^{\mathrm{T}}+\\&\boldsymbol{K}_{fi}(t+1)\boldsymbol{R}_i(t+1)\boldsymbol{K}_{fi}^{\mathrm{T}}(t+1)\end{aligned} \tag{6.2.62}$$

且局部滤波误差协方差阵 $\boldsymbol{P}_{ij}(t+1\mid t+1)(i\neq j)$ 为

$$\begin{aligned}\boldsymbol{P}_{ij}(t+1\mid t+1)=&\boldsymbol{\Psi}_{fi}(t+1)\boldsymbol{P}_{ij}(t\mid t)\boldsymbol{\Psi}_{fj}^{\mathrm{T}}(t+1)+[\boldsymbol{I}_n-\boldsymbol{K}_{fi}(t+1)\boldsymbol{H}_i(t+1)]\times\\&\boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)[\boldsymbol{I}_n-\boldsymbol{K}_{fj}^{\mathrm{T}}(t+1)\boldsymbol{H}_j(t+1)]^{\mathrm{T}}\end{aligned} \tag{6.2.63}$$

带初值 $\boldsymbol{P}_{ij}(0\mid 0)=\boldsymbol{P}_0$.

根据节 6.1 的三种最优加权融合估计准则直接引出下述几个定理.

【定理 6.2.6】 多传感器时变系统(6.2.1)和(6.2.2)在假设 1 ~ 4 下，按矩阵加权最优信息融合 Kalman 滤波器 $\hat{x}_0(t\mid t)$ 为

$$\hat{\boldsymbol{x}}_0(t\mid t)=\sum_{i=1}^{L}\boldsymbol{A}_i(t)\hat{\boldsymbol{x}}_i(t\mid t) \tag{6.2.64}$$

其中 $\hat{\boldsymbol{x}}_i(t\mid t)$ 为第 i 传感器子系统局部 Kalman 滤波器，由定理 6.2.3 计算. 最优加权阵 $\boldsymbol{A}_i(t)$ 为

$$[\boldsymbol{A}_1(t),\cdots,\boldsymbol{A}_L(t)]=(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(t\mid t)\boldsymbol{e})^{-1}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(t\mid t) \tag{6.2.65}$$

其中 $\boldsymbol{e}^{\mathrm{T}}=[\boldsymbol{I}_n,\cdots,\boldsymbol{I}_n]$，且定义 $nL\times nL$ 矩阵 $\boldsymbol{P}(t\mid t)$ 为

$$\boldsymbol{P}(t\mid t)=\begin{bmatrix}\boldsymbol{P}_{11}(t\mid t) & \cdots & \boldsymbol{P}_{1L}(t\mid t)\\ \vdots & & \vdots\\ \boldsymbol{P}_{L1}(t\mid t) & \cdots & \boldsymbol{P}_{LL}(t\mid t)\end{bmatrix} \tag{6.2.66}$$

其中定义 $\boldsymbol{P}_{ii}(t\mid t)=\boldsymbol{P}_i(t\mid t)$，$\boldsymbol{P}_{ij}(t\mid t)$ 由定理 6.2.3 和定理 6.2.4 计算.

最优融合滤波误差方差阵 $P_0(t\mid t)$ 为

$$\boldsymbol{P}_0(t\mid t)=(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(t\mid t)\boldsymbol{e})^{-1} \tag{6.2.67}$$

且有关系

$$\mathrm{tr}\boldsymbol{P}_0(t\mid t)\leqslant \mathrm{tr}\boldsymbol{P}_i(t\mid t),\quad i=1,\cdots,L \tag{6.2.68}$$

【定理 6.2.7】 多传感器时变系统(6.2.1)和(6.2.2)在假设 1 ~ 4 下，按标量加权最优信息融合 Kalman 滤波器为

$$\hat{\boldsymbol{x}}_0(t \mid t) = \sum_{i=1}^{L} a_i(t)\hat{\boldsymbol{x}}_i(t \mid t) \tag{6.2.69}$$

其中 $\hat{\boldsymbol{x}}_i(t \mid t)$ 为由定理 6.2.3 给出的局部 Kalman 滤波器.最优加权系数 $a_i(t)$ 为

$$[a_1(t), \cdots, a_L(t)] = \frac{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{tr}}^{-1}(t \mid t)}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{tr}}^{-1}(t \mid t)\boldsymbol{e}} \tag{6.2.70}$$

$$\boldsymbol{P}_{\mathrm{tr}}(t \mid t) = \begin{bmatrix} \mathrm{tr}\boldsymbol{P}_{11}(t \mid t) & \cdots & \mathrm{tr}\boldsymbol{P}_{1L}(t \mid t) \\ \vdots & & \vdots \\ \mathrm{tr}\boldsymbol{P}_{L1}(t \mid t) & \cdots & \mathrm{tr}\boldsymbol{P}_{LL}(t \mid t) \end{bmatrix} \tag{6.2.71}$$

其中 $\boldsymbol{e}^{\mathrm{T}} = [1, \cdots, 1]$,$\boldsymbol{P}_{ij}(t \mid t)$ 由定理 6.2.3 和定理 6.2.4 计算,且规定 $\boldsymbol{P}_i(t \mid t) = \boldsymbol{P}_{ii}(t \mid t)$.最优融合滤波误差方差阵为

$$\boldsymbol{P}_0(t \mid t) = \sum_{i=1}^{L}\sum_{j=1}^{L} a_i(t)a_j(t)\boldsymbol{P}_{ij}(t \mid t) \tag{6.2.72}$$

且有关系

$$\mathrm{tr}\boldsymbol{P}_0(t \mid t) \leqslant \mathrm{tr}\boldsymbol{P}_i(t \mid t), \quad i = 1, \cdots, L \tag{6.2.73}$$

【定理 6.2.8】 多传感器时变系统(6.2.1)和(6.2.2)在假设 1 ~ 4 下,按对角阵加权(按分量标量加权)最优信息融合 Kalman 滤波器为

$$\hat{\boldsymbol{x}}_0(t \mid t) = [\hat{x}_{01}(t \mid t), \cdots, \hat{x}_{0n}(t \mid t)]^{\mathrm{T}} \tag{6.2.74}$$

$$\hat{x}_{0i}(t \mid t) = \sum_{j=1}^{L} a_{ji}(t)\hat{x}_{ji}(t \mid t) \tag{6.2.75}$$

其中局部 Kalman 滤波器 $\hat{\boldsymbol{x}}_j(t \mid t) = [\hat{x}_{j1}(t \mid t), \cdots, \hat{x}_{jn}(t \mid t)]^{\mathrm{T}}$ 由定理 6.2.3 给出.最优加权系数 $a_{ji}(t)$ 为

$$[a_{1i}(t), \cdots, a_{Li}(t)] = \frac{\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}^{ii}(t \mid t))^{-1}}{\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}^{ii}(t \mid t))^{-1}\boldsymbol{e}}, \quad i = 1, \cdots, L \tag{6.2.76}$$

其中 $\boldsymbol{e}^{\mathrm{T}} = [1, \cdots, 1]$,且定义 $L \times L$ 矩阵 $P^{ii}(t \mid t)$ 为

$$\boldsymbol{P}^{ii}(t \mid t) = \begin{bmatrix} P_{11}^{ii}(t \mid t) & \cdots & P_{1L}^{ii}(t \mid t) \\ \vdots & & \vdots \\ P_{L1}^{ii}(t \mid t) & \cdots & P_{LL}^{ii}(t \mid t) \end{bmatrix} \tag{6.2.77}$$

其中 $P_{jk}^{ii}(t \mid t)$ 为 $\boldsymbol{P}_{jk}(t \mid t)$ 的第(i, i)对角元素.分量最优融合误差方差为

$$P_{0i}(t \mid t) = [\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}^{ii}(t \mid t))^{-1}\boldsymbol{e}]^{-1}, \quad i = 1, \cdots, n \tag{6.2.78}$$

且有最优融合误差方差阵 $\boldsymbol{P}_0(t \mid t)$ 的迹为

$$\mathrm{tr}\boldsymbol{P}_0(t \mid t) = \sum_{i=1}^{n} P_{0i}(t \mid t) \tag{6.2.79}$$

且有关系

$$\mathrm{tr}\boldsymbol{P}_0(t \mid t) \leqslant \mathrm{tr}\boldsymbol{P}_i(t \mid t), \quad i = 1, \cdots, L \tag{6.2.80}$$

上述三个定理的证明可直接由节 6.1 的三种加权最优融合准则和算法得出.

【定理 6.2.9】 多传感器时变系统(6.2.1)和(6.2.2)在假设 1 ~ 4 下有按矩阵加权最优融合 Kalman 预报器

$$\hat{\boldsymbol{x}}_0(t+1 \mid t) = \sum_{i=1}^{L} \boldsymbol{A}_i(t)\hat{\boldsymbol{x}}_i(t+1 \mid t) \tag{6.2.81}$$

$$[\boldsymbol{A}_1(t),\cdots,\boldsymbol{A}_L(t)] = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(t+1\mid t)\boldsymbol{e})^{-1}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(t+1\mid t) \tag{6.2.82}$$

$$\boldsymbol{P}(t+1\mid t) = (\boldsymbol{P}_{ij}(t+1\mid t)) \tag{6.2.83}$$

其中 $\boldsymbol{P}(t+1\mid t)$ 是以 $\boldsymbol{P}_{ij}(t+1\mid t)$ 为第 (i,j) 元素的 $nL\times nL$ 分块矩阵，$\hat{x}_i(t+1\mid t)$，$\boldsymbol{P}_{ij}(t+1\mid t)$ 由定理 6.2.1 计算. 最优融合预报误差方差阵为

$$\boldsymbol{P}_0(t+1\mid t) = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(t+1\mid t)\boldsymbol{e})^{-1} \tag{6.2.84}$$

按标量加权最优融合 Kalman 预报器为

$$\hat{\boldsymbol{x}}_0(t+1\mid t) = \sum_{i=1}^{L} a_i(t)\hat{\boldsymbol{x}}_1(t+1\mid t) \tag{6.2.85}$$

$$[a_1(t),\cdots,a_L(t)] = \boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{tr}}^{-1}(t+1\mid t)(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{tr}}^{-1}(t+1\mid t)\boldsymbol{e})^{-1} \tag{6.2.86}$$

$$\boldsymbol{P}_{\mathrm{tr}}(t+1\mid t) = (\mathrm{tr}\boldsymbol{P}_{ij}(t+1\mid t)) \tag{6.2.87}$$

其中 $\boldsymbol{P}_{\mathrm{tr}}(t+1\mid t)$ 是以 $\mathrm{tr}\boldsymbol{P}_{ij}(t+1\mid t)$ 为第 (i,j) 元素的 $L\times L$ 矩阵. 最优融合预报误差方差阵为

$$\boldsymbol{P}_0(t+1\mid t) = \sum_{i=1}^{L}\sum_{j=1}^{L} a_i(t)a_j(t)\boldsymbol{P}_{ij}(t+1\mid t) \tag{6.2.88}$$

按分量标量加权最优融合 Kalman 预报器为

$$\hat{\boldsymbol{x}}_0(t+1\mid t) = [\hat{x}_{01}(t+1\mid t),\cdots,\hat{x}_{0n}(t+1\mid t)]^{\mathrm{T}} \tag{6.2.89}$$

$$\hat{x}_{0i}(t+1\mid t) = \sum_{j=1}^{L} a_{ji}(t)\hat{x}_{ji}(t+1\mid t),\quad i=1,\cdots,n \tag{6.2.90}$$

其中局部 Kalman 预报器 $\hat{\boldsymbol{x}}_j(t+1\mid t) = [\hat{x}_{j1}(t+1\mid t),\cdots,\hat{x}_{jn}(t+1\mid t)]^{\mathrm{T}}$，且有

$$[a_{1i}(t),\cdots,a_{Li}(t)] = \frac{\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}^{ii}(t+1\mid t))^{-1}}{\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}^{ii}(t+1\mid t))^{-1}\boldsymbol{e}},\quad i=1,\cdots,L \tag{6.2.91}$$

其中 $\boldsymbol{P}^{ii}(t+1\mid t)$ 是以 $P_{kj}^{(ii)}(t+1\mid t)$ 为第 (k,j) 元素的 $L\times L$ 矩阵，而 $P_{kj}^{(ii)}(t+1\mid t)$ 是 $\boldsymbol{P}_{kj}(t+1\mid t)$ 的第 (i,i) 对角元素，即

$$\boldsymbol{P}^{ii}(t+1\mid t) = (P_{kj}^{(ii)}(t+1\mid t))_{L\times L} \tag{6.2.92}$$

按分量最优融合预报误差方差为

$$P_{0i}(t+1\mid t) = [\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}^{ii}(t+1\mid t))^{-1}\boldsymbol{e}]^{-1},\quad i=1,\cdots,n \tag{6.2.93}$$

且有最优融合预报误差方差阵 $\boldsymbol{P}_0$ 的迹为

$$\mathrm{tr}\boldsymbol{P}_0(t+1\mid t) = \sum_{i=1}^{n} P_{0i}(t+1\mid t) \tag{6.2.94}$$

在上述三种加权下，均有关系

$$\mathrm{tr}\boldsymbol{P}_0(t+1\mid t) \leqslant \mathrm{tr}\boldsymbol{P}_i(t+1\mid t),\quad i=1,\cdots,L \tag{6.2.95}$$

6.3 时变系统多传感器信息融合超前 *N* 步 Kalman 预报器

考虑 $\boldsymbol{u}(t)=\boldsymbol{0}$ 的多传感器时变系统 (6.2.1) 和 (6.2.2)

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}(t)\boldsymbol{x}(t) + \boldsymbol{\Gamma}(t)\boldsymbol{w}(t) \tag{6.3.1}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{H}_i(t)\boldsymbol{x}(t) + \boldsymbol{v}_i(t),\quad i=1,\cdots,L \tag{6.3.2}$$

假设 1 ~ 3 同节 6.2 的假设 1 ~ 3. 第 i 传感器子系统有局部超前一步 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}(t)\hat{\boldsymbol{x}}_i(t\mid t-1)+\boldsymbol{K}_{pi}(t)\boldsymbol{y}_i(t),\quad i=1,\cdots,L \tag{6.3.3}$$

其中 $\boldsymbol{\Psi}_{pi}(t)$，$\boldsymbol{K}_{pi}(t)$ 由定理 6.2.1 和定理 6.2.2 置 $\boldsymbol{u}(t)=\boldsymbol{0}$ 来计算.

【定理 6.3.1】 多传感器时变系统(6.3.1)和(6.3.2)在假设 1 ~ 3 下，第 i 传感器子系统有局部超前 N 步最优 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\hat{\boldsymbol{x}}_i(t+1\mid t),\quad i=1,\cdots,L,\quad N>1 \tag{6.3.4}$$

其中 $\hat{\boldsymbol{x}}_i(t+1\mid t)$ 由(6.3.3)计算，且定义

$$\boldsymbol{\Phi}(t,i)=\boldsymbol{\Phi}(t-1)\boldsymbol{\Phi}(t-2)\cdots\boldsymbol{\Phi}(i),$$
$$\boldsymbol{\Phi}(t,t)=\boldsymbol{I}_n \tag{6.3.5}$$

最优超前 N 步预报误差 $\tilde{\boldsymbol{x}}_i(t+N\mid t)=\boldsymbol{x}(t+N)-\hat{\boldsymbol{x}}_i(t+N\mid t)$ 方差阵 $\boldsymbol{P}_i(t+N\mid t)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t+N\mid t)\tilde{\boldsymbol{x}}_i^{\mathrm{T}}(t+N\mid t)]$ 为

$$\boldsymbol{P}_i(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{P}_i(t+1\mid t)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+1)+$$
$$\sum_{j=2}^{N}\boldsymbol{\Phi}(t+N,t+j)\boldsymbol{\Gamma}(t+j-1)\boldsymbol{Q}(t+j-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t+j-1)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+j) \tag{6.3.6}$$

其中一步预报误差方差阵 $\boldsymbol{P}_i(t+1\mid t)$ 由定理 6.2.1 和定理 6.2.2 计算. 超前 N 步预报误差协方差阵 $\boldsymbol{P}_{ij}(t+N\mid t)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t+N\mid t)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t+N\mid t)]$，$i,j=1,\cdots,L$，为

$$\boldsymbol{P}_{ij}(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{P}_{ij}(t+1\mid t)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+1)+$$
$$\sum_{j=2}^{N}\boldsymbol{\Phi}(t+N,t+j)\boldsymbol{\Gamma}(t+j-1)\boldsymbol{Q}(t+j-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t+j-1)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+j) \tag{6.3.7}$$

特别当 $i=j$ 时，$\boldsymbol{P}_{ii}(t+N\mid t)=\boldsymbol{P}_i(t+N\mid t)$ 同(6.3.6)一样.

证明 由(6.3.1)迭代有关系

$$\boldsymbol{x}(t+N)=\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{x}(t+1)+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}(t+N,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1) \tag{6.3.8}$$

其中 $\boldsymbol{\Phi}(t+N,i)$ 由(6.3.5)定义. 另一方面

$$\hat{\boldsymbol{x}}_i(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\hat{\boldsymbol{x}}_i(t+1\mid t),\quad N>1,i=1,\cdots,L \tag{6.3.9}$$

上两式相减有

$$\tilde{\boldsymbol{x}}_i(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\tilde{\boldsymbol{x}}_i(t+1\mid t)+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}(t+N,i)\times$$
$$\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1),\quad N>1 \tag{6.3.10}$$

在上式求和项引入变换 $i=t+j$，由(6.3.10)得

$$\tilde{\boldsymbol{x}}_i(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\tilde{\boldsymbol{x}}_i(t+1\mid t)+\sum_{j=2}^{N}\boldsymbol{\Phi}(t+N,t+j)\times$$
$$\boldsymbol{\Gamma}(t+j-1)\boldsymbol{w}(t+j-1) \tag{6.3.11}$$

由此利用 $\tilde{\boldsymbol{x}}_i(t+1\mid t)$ 不相关于 $\boldsymbol{w}(t+1),\cdots,\boldsymbol{w}(t+N+1)$，可得(6.3.6)和(6.3.7). 证毕. □

超前 N 步预报器也可表为如下定理给出的等价形式.

【定理 6.3.2】　多传感器时变系统(6.3.1)和(6.3.2)在假设 1 ~ 3 下,当 $N<0$ 时,第 i 子系统有局部超前($-N$)步最优 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{\Phi}(t,t+N+1)\hat{\boldsymbol{x}}_i(t+N+1\mid t+N),\quad N\leqslant -2 \tag{6.3.12}$$

其中 $\boldsymbol{\Phi}(t,i)$ 的定义同(6.3.5),相应的预报误差方差阵为

$$\begin{aligned}\boldsymbol{P}_i(t\mid t+N)=&\boldsymbol{\Phi}(t,t+N+1)\boldsymbol{P}_i(t+N+1\mid t+N)\boldsymbol{\Phi}^{\mathrm{T}}(t,t+N+1)+\\&\sum_{j=0}^{-N-2}\boldsymbol{\Phi}(t,t-j)\boldsymbol{\Gamma}(t-j-1)\boldsymbol{Q}(t-j-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-j-1)\boldsymbol{\Phi}^{\mathrm{T}}(t,t-j)\end{aligned} \tag{6.3.13}$$

而预报误差互协方差阵 $\boldsymbol{P}_{ij}(t\mid t+N)$

$$\begin{aligned}\boldsymbol{P}_{ij}(t\mid t+N)=&\boldsymbol{\Phi}(t,t+N+1)\boldsymbol{P}_{ij}(t+N+1\mid t+N)\boldsymbol{\Phi}^{\mathrm{T}}(t,t+N+1)+\\&\sum_{r=0}^{-N-2}\boldsymbol{\Phi}(t,t-r)\boldsymbol{\Gamma}(t-r-1)\boldsymbol{Q}(t-r-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-r-1)\boldsymbol{\Phi}^{\mathrm{T}}(t,t-r)\end{aligned} \tag{6.3.14}$$

证明　由(6.3.1)迭代有关系

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}(t,t+N+1)\boldsymbol{x}(t+N+1)+\sum_{i=t+N+2}^{t}\boldsymbol{\Phi}(t,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1) \tag{6.3.15}$$

由(6.3.1)和射影运算有

$$\hat{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{\Phi}(t,t+N+1)\hat{\boldsymbol{x}}_i(t+N+1\mid t+N),\quad N\leqslant -2 \tag{6.3.16}$$

上两式相减引出预报误差 $\tilde{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}_i(t\mid t+N)$ 为

$$\begin{aligned}\tilde{\boldsymbol{x}}_i(t\mid t+N)=&\boldsymbol{\Phi}(t,t+N+1)\tilde{\boldsymbol{x}}_i(t+N+1\mid t+N)-\\&\sum_{i=t+N+2}^{t}\boldsymbol{\Phi}(t,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1)\end{aligned} \tag{6.3.17}$$

在上式求和时引入变换 $r=t-i$,上式化为

$$\begin{aligned}\tilde{\boldsymbol{x}}_i(t\mid t+N)=&\boldsymbol{\Phi}(t,t+N+1)\tilde{\boldsymbol{x}}_i(t+N+1\mid t+N)-\\&\sum_{r=0}^{-N-2}\boldsymbol{\Phi}(t,t-r)\boldsymbol{\Gamma}(t-r-1)\boldsymbol{w}(t-r-1)\end{aligned} \tag{6.3.18}$$

由此引出(6.3.13)和(6.3.14). 证毕. □

【定理 6.3.2】　多传感器时变系统(6.3.1)和(6.3.2)在假设 1 ~ 3 下,有定理 6.2.7 中的(6.2.81) ~ (6.2.95)形式的三种加权多传感器最优信息融合超前 N 步 Kalman 预报器 $\hat{\boldsymbol{x}}_0(t+N\mid t)$,但应将其中的 $N=1$ 一步预报改为任意 N 步($N>1$)预报,即各公式相应地改动为

$$\begin{gathered}\hat{\boldsymbol{x}}_0(t+1\mid t)\to\hat{\boldsymbol{x}}_0(t+N\mid t)\\ \boldsymbol{P}_{ij}(t+1\mid t)\to\boldsymbol{P}_{ij}(t+N\mid t),\quad \boldsymbol{P}(t+1\mid t)\to\boldsymbol{P}(t+N\mid t)\\ \mathrm{tr}\boldsymbol{P}_{ij}(t+1\mid t)\to\mathrm{tr}\boldsymbol{P}_{ij}(t+N\mid t),\quad \boldsymbol{P}_{\mathrm{tr}}(t+1\mid t)\to\boldsymbol{P}_{\mathrm{tr}}(t+N\mid t)\\ \boldsymbol{P}_{kj}^{(ii)}(t+1\mid t)\to\boldsymbol{P}_{kj}^{(ii)}(t+N+1\mid t),\quad \boldsymbol{P}^{ii}(t+1\mid t)\to\boldsymbol{P}^{ii}(t+N\mid t)\\ \boldsymbol{P}_0(t+1\mid t)\to\boldsymbol{P}_0(t+N\mid t),\quad \boldsymbol{P}_{0i}(t+1\mid t)\to\boldsymbol{P}_{0i}(t+N\mid t)\end{gathered} \tag{6.3.19}$$

6.4　时变系统多传感器信息融合 Kalman 平滑器

考虑多传感器时变系统(6.2.1)和(6.2.2),在假设 1 ~ 4 下我们考虑信息融合平滑估

计问题.

【引理6.4.1】 时变系统(6.3.1)和(6.3.2)在假设1 ~ 4下,第 i 传感器子系统有局部最优 Kalman 平滑器

$$\hat{\boldsymbol{x}}_i(t \mid t+N) = \hat{\boldsymbol{x}}_i(t \mid t-1) + \sum_{j=0}^{N} \boldsymbol{K}_i(t \mid t+j)\boldsymbol{\varepsilon}_i(t+j), \quad N \geqslant 1 \tag{6.4.1}$$

其中 $i = 1,\cdots,L$. 平滑增益为

$$\boldsymbol{K}_i(t \mid t+j) = \boldsymbol{P}_i(t \mid t-1)\left\{\prod_{k=0}^{j-1} \boldsymbol{\Psi}_{pi}^{\mathrm{T}}(t+k)\right\}\boldsymbol{H}_i^{\mathrm{T}}(t+j)\boldsymbol{Q}_{\varepsilon i}^{-1}(t+j), \quad j \geqslant 1 \tag{6.4.2}$$

$$\boldsymbol{K}_i(t \mid t) = \boldsymbol{K}_{fi}(t) \tag{6.4.3}$$

其中 $\boldsymbol{\Psi}_{pi}(t)$, $\boldsymbol{P}_i(t \mid t-1)$, $\boldsymbol{Q}_{\varepsilon i}(t)$, $\boldsymbol{K}_{fi}(t)$, $\hat{\boldsymbol{x}}_i(t \mid t-1)$ 的定义见定理6.2.1 ~ 定理6.2.3,且有平滑误差 $\tilde{\boldsymbol{x}}_i(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}_i(t \mid t+N)$ 方差阵 $\boldsymbol{P}_i(t \mid t+N) = \mathrm{E}[\tilde{\boldsymbol{x}}_i(t \mid t+N)\tilde{\boldsymbol{x}}_i^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}_i(t \mid t+N) = \boldsymbol{P}_i(t \mid t-1) - \sum_{j=0}^{N} \boldsymbol{K}_i(t \mid t+j)\boldsymbol{Q}_{\varepsilon i}(t+j)\boldsymbol{K}_i^{\mathrm{T}}(t \mid t+j) \tag{6.4.4}$$

证明 见定理3.7.9和推论3.7.8. □

【定理6.4.1】 多传感器时变系统(6.2.1)和(6.2.2)在假设1 ~ 4下,第 i 传感器与第 j 传感器的平滑误差协方差阵 $\boldsymbol{P}_{ij}(t \mid t+N) = \mathrm{E}[\tilde{\boldsymbol{x}}_i(t \mid t+N)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t \mid t+N)](i \neq j)$ 为

$$\begin{aligned}\boldsymbol{P}_{ij}(t \mid t+N) = {} & \boldsymbol{P}_{ij}(t \mid t-1) - \sum_{r=0}^{N} \boldsymbol{K}_i(t \mid t+r)\boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}(t+r,t)\boldsymbol{P}_{ij}(t \mid t-1) - \\ & \sum_{s=0}^{N} \boldsymbol{P}_{ij}(t \mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+s,t)\boldsymbol{H}_j^{\mathrm{T}}(t+s)\boldsymbol{K}_j^{\mathrm{T}}(t \mid t+s) + \\ & \sum_{r=0}^{N}\sum_{s=0}^{N} \boldsymbol{K}_i(t \mid t+r)\boldsymbol{E}_{ij}(r,s,t)\boldsymbol{K}_j^{\mathrm{T}}(t \mid t+s)\end{aligned} \tag{6.4.5}$$

其中 $N > 0$ 且定义 $\boldsymbol{\Psi}_{pi}(t+1,t+1) = \boldsymbol{I}_n$,

$$\boldsymbol{\Psi}_{pi}(t+k,t) = \boldsymbol{\Psi}_{pi}(t+k-1)\cdots\boldsymbol{\Psi}_{pi}(t) \tag{6.4.6}$$

$$\boldsymbol{E}_{ij}(r,s,t) = \mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)] \tag{6.4.7}$$

而当 $\min(r,s) > 0$ 时 $\boldsymbol{E}_{ij}(r,s,t)$ 有公式

$$\begin{aligned}\boldsymbol{E}_{ij}(r,s,t) = {} & \boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}(t+r,t)\boldsymbol{P}_{ij}(t \mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+s,t)\boldsymbol{H}_j^{\mathrm{T}}(t+s) + \\ & \sum_{\nu=1}^{\min(r,s)} \boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}(t+r,t+\nu)[\boldsymbol{I}_n, -\bar{\boldsymbol{K}}_{pi}(t+\nu-1)] \times \\ & \begin{bmatrix} \boldsymbol{S}_{ij}^{11}(t+\nu-1) & \boldsymbol{S}_{ij}^{12}(t+\nu-1) \\ \boldsymbol{S}_{ij}^{21}(t+\nu-1) & \boldsymbol{S}_{ij}^{22}(t+\nu-1) \end{bmatrix}\begin{bmatrix} \boldsymbol{I}_n \\ -\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}}(t+\nu-1) \end{bmatrix} \times \\ & \boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+s,t+\nu)\boldsymbol{H}_j^{\mathrm{T}}(t+s) + \boldsymbol{R}_{ij}(t+r)\delta_{rs}\end{aligned} \tag{6.4.8}$$

其中

$$\boldsymbol{S}_{ij}^{11}(t) = \boldsymbol{\Gamma}(t)\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) - \boldsymbol{J}_i(t)\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) - \boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t)\boldsymbol{J}_j^{\mathrm{T}}(t) +$$

$$\boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t),$$
$$\boldsymbol{S}_{ij}^{12}(t) = \boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t) - \boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t),$$
$$\boldsymbol{S}_{ij}^{21}(t) = \boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) - \boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t),$$
$$\boldsymbol{S}_{ij}^{22}(t) = \boldsymbol{R}_{ij}(t) \tag{6.4.9}$$

当 $\min(r,s)=0$ 时有

$$\begin{aligned}
\boldsymbol{E}_{ij}(1,0,t) &= \boldsymbol{H}_i(t+1)\boldsymbol{\Psi}_{pi}(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t) + \\
&\quad \boldsymbol{H}_i(t+1)[\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t) - \boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t)] - \boldsymbol{H}_i(t+1)\bar{\boldsymbol{K}}_{pi}(t)\boldsymbol{R}_{ij}(t),\\
\boldsymbol{E}_{ij}(0,1,t) &= \boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t)\boldsymbol{H}_j^{\mathrm{T}}(t+1) + \\
&\quad [\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) - \boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t)]\boldsymbol{H}_j^{\mathrm{T}}(t+1) - \boldsymbol{R}_{ij}(t)\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}}(t)\boldsymbol{H}_j^{\mathrm{T}}(t+1),\\
\boldsymbol{E}_{ij}(0,0,t) &= \boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t) + \boldsymbol{R}_{ij}(t),\\
\boldsymbol{E}_{ij}(k,0,t) &= \boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t) + \\
&\quad \boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+1)[\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t) - \boldsymbol{J}_i(t)\boldsymbol{R}_{ij}(t) - \bar{\boldsymbol{K}}_{pi}(t)\boldsymbol{R}_{ij}(t)],\\
\boldsymbol{E}_{ij}(0,k,t) &= \boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+k,t)\boldsymbol{H}_j^{\mathrm{T}}(t+k) + \\
&\quad [\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) - \boldsymbol{R}_{ij}(t)\boldsymbol{J}_j^{\mathrm{T}}(t) - \boldsymbol{R}_{ij}(t)\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}}(t)]\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+k,t+1)\boldsymbol{H}_j^{\mathrm{T}}(t+k)
\end{aligned} \tag{6.4.10}$$

其中 $\boldsymbol{\Psi}_{pi}(t)$，$\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}}(t)$，$\boldsymbol{P}_{iJ}(t\mid t-1)$ 由定理 6.2.2 计算.

证明 由(6.4.1)有

$$\tilde{\boldsymbol{x}}_i(t\mid t+N) = \tilde{\boldsymbol{x}}_i(t\mid t-1) - \sum_{r=0}^{N}\boldsymbol{K}_i(t\mid t+r)\boldsymbol{\varepsilon}_i(t+r) \tag{6.4.11}$$

$$\tilde{\boldsymbol{x}}_j(t\mid t+N) = \tilde{\boldsymbol{x}}_j(t\mid t-1) - \sum_{s=0}^{N}\boldsymbol{K}_j(t\mid t+s)\boldsymbol{\varepsilon}_j(t+s) \tag{6.4.12}$$

注意

$$\boldsymbol{\varepsilon}_i(t+k) = \boldsymbol{H}_i(t+k)\tilde{\boldsymbol{x}}_i(t+k\mid t+k-1) + \boldsymbol{v}_i(t+k) \tag{6.4.13}$$

应用(6.2.27)有

$$\tilde{\boldsymbol{x}}_i(t+1\mid t) = \boldsymbol{\Psi}_{pi}(t)\tilde{\boldsymbol{x}}_i(t\mid t-1) - \bar{\boldsymbol{K}}_{pi}(t)\boldsymbol{v}_i(t) + \bar{\boldsymbol{w}}_i(t) \tag{6.4.14}$$

由它迭代引出关系

$$\begin{aligned}
\tilde{\boldsymbol{x}}_i(t+k\mid t+k-1) &= \boldsymbol{\Psi}_{pi}(t+k,t)\tilde{\boldsymbol{x}}_i(t\mid t-1) + \sum_{\nu=1}^{k}\boldsymbol{\Psi}_{pi}(t+k,t+\nu)\times\\
&\quad [\bar{\boldsymbol{w}}_i(t+\nu-1) - \bar{\boldsymbol{K}}_{pi}(t+\nu-1)\boldsymbol{v}_i(t+\nu-1)]
\end{aligned} \tag{6.4.15}$$

从而有

$$\begin{aligned}
\boldsymbol{\varepsilon}_i(t+k) &= \boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t)\tilde{\boldsymbol{x}}_i(t\mid t-1) + \sum_{\nu=1}^{k}\boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+\nu)\times\\
&\quad [\boldsymbol{I}_n, -\bar{\boldsymbol{K}}_{pi}(t+\nu-1)]\begin{bmatrix}\bar{\boldsymbol{w}}_i(t+\nu-1)\\ \boldsymbol{v}_i(t+\nu-1)\end{bmatrix} + \boldsymbol{v}_i(t+k)
\end{aligned} \tag{6.4.16}$$

于是有关系

$$\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t-1)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)] = \boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pi}^{\mathrm{T}}(t+s,t)\boldsymbol{H}_j(t+s) \tag{6.4.17}$$

$$\mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t\mid t-1)] = \boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}^{\mathrm{T}}(t+r,t)\boldsymbol{P}_{ij}(t\mid t-1) \tag{6.4.18}$$

应用(6.4.11)~(6.4.12)和(6.4.17)~(6.4.18)得(6.4.5).应用(6.4.16)和(6.4.8)，其

中

$$\begin{aligned}S_{ij}^{11}(t) &= \mathrm{E}[\bar{w}_i(t)\bar{w}_j^{\mathrm{T}}(t)], \quad S_{ij}^{12}(t) = \mathrm{E}[\bar{w}_i(t)v_j^{\mathrm{T}}(t)],\\ S_{ij}^{21}(t) &= \mathrm{E}[v_i(t)\bar{w}_j^{\mathrm{T}}(t)], \quad S_{ij}^{22}(t) = \mathrm{E}[v_i(t)v_j^{\mathrm{T}}(t)]\end{aligned} \tag{6.4.19}$$

由 $\bar{w}_i(t)$ 的定义和(6.2.3),(6.2.4) 可得(6.4.9).

当 $\min(r,s)=0$ 时,我们有

$$E_{ij}(1,0,t) = \mathrm{E}[\varepsilon_i(t+1)\varepsilon_j^{\mathrm{T}}(t)] = \mathrm{E}[(H_i(t+1)\tilde{x}_i(t+1\mid t) + v_i(t+1))\varepsilon_j^{\mathrm{T}}(t)] \tag{6.4.20}$$

将(6.4.13) 置 $k=0$ 和 $i=j$ 后将其和(6.4.14) 代入(6.4.20) 中可得(6.4.10) 的第1式.还有

$$\begin{aligned}E_{ij}(0,1,t) = \mathrm{E}[\varepsilon_i(t)\varepsilon_j^{\mathrm{T}}(t+1)] = \\ \mathrm{E}[\varepsilon_i(t)(H_j(t+1)\tilde{x}_j(t+1\mid t) + v_j(t+1))^{\mathrm{T}}]\end{aligned} \tag{6.4.21}$$

类似地可求得(6.4.10) 的第2式.最后有

$$\begin{aligned}E_{ij}(0,0,t) = \mathrm{E}[\varepsilon_i(t)\varepsilon_j^{\mathrm{T}}(t)] = \mathrm{E}[(H_i(t)\tilde{x}_i(t\mid t-1) + \\ v_i(t+1))(H_j(t)\tilde{x}_j(t\mid t-1) + v_j(t+1))^{\mathrm{T}}]\end{aligned} \tag{6.4.22}$$

由此可得(6.4.10) 的第3式.应用(6.4.15) 类似得(6.4.10) 的第4式和第5式. □

注意,推导定理6.4.1 的关键是利用预报误差表达式(6.4.14),它相应于定理6.2.2.现在利用定理6.2.1 中的预报误差表达式(6.2.15) 我们可得到等价于定理6.4.1 的等价新算法.

【定理6.4.2】 多传感器时变系统(6.2.1) 和(6.2.2) 在假设1 ~ 4下,局部平滑误差协方差阵 $P_{ij}(t\mid t+N)(i\neq j)$ 为

$$\begin{aligned}P_{ij}(t\mid t+N) = P_{ij}(t\mid t-1) - \sum_{r=0}^{N}K_i(t\mid t+r)H_i(t+r)\Psi_{pi}(t+r,t)P_{ij}(t\mid t-1) - \\ \sum_{s=0}^{N}P_{ij}(t\mid t-1)\Psi_{pj}^{\mathrm{T}}(t+s,t)H_j^{\mathrm{T}}(t+s)K_j^{\mathrm{T}}(t\mid t+s) + \\ \sum_{r=0}^{N}\sum_{s=0}^{N}K_i(t\mid t+r)E_{ij}(r,s,t)K_j^{\mathrm{T}}(t\mid t+s)\end{aligned} \tag{6.4.23}$$

其中 $P_{ij}(t\mid t-1)$,$\Psi_{pi}(t)$ 由定理6.2.1计算.当 $\min(r,s)>0$ 时,$E_{ij}(r,s,t) = \mathrm{E}[\varepsilon(t+r)\varepsilon_j^{\mathrm{T}}(t+s)]$,有公式

$$\begin{aligned}E_{ij}(r,s,t) = H_i(t+r)\Psi_{pi}(t+r,t)P_{ij}(t\mid t-1)\Psi_{pj}^{\mathrm{T}}(t+s,t)H_j^{\mathrm{T}}(t+s) + \\ \sum_{\nu=1}^{\min(r,s)}H_i(t+r)\Psi_{pi}(t+r,t+\nu)[\Gamma(t+\nu-1), -K_{pi}(t-\nu-1)]\times \\ \begin{bmatrix}Q(t+\nu-1) & S_j(t+\nu-1)\\ S_i^{\mathrm{T}}(t+\nu-1) & R_{ij}(t+\nu-1)\end{bmatrix}\begin{bmatrix}\Gamma^{\mathrm{T}}(t+\nu-1)\\ -K_{pj}^{\mathrm{T}}(t+\nu-1)\end{bmatrix}\times \\ \Psi_{pj}^{\mathrm{T}}(t+s,t+\nu)H_j^{\mathrm{T}}(t+s) + R_{ij}(t+r)\delta_{rs}\end{aligned} \tag{6.4.24}$$

若 $\min(r,s)=0$,则有

$$\begin{aligned}E_{ij}(1,0,t) = H_i(t+1)\Psi_{pi}(t)P_{ij}(t\mid t-1)H_j^{\mathrm{T}}(t) + H_i(t+1)\Gamma(t)S_j(t) - \\ H_i(t+1)K_{pi}(t)R_{ij}(t),\end{aligned}$$

$$E_{ij}(0,1,t) = H_i(t)P_{ij}(t\mid t-1)\Psi_{pj}^{\mathrm{T}}(t)H_j^{\mathrm{T}}(t+1) + S_i^{\mathrm{T}}(t)\Gamma^{\mathrm{T}}(t)H_j^{\mathrm{T}}(t+1) -$$

$$\boldsymbol{R}_{ij}(t)\boldsymbol{K}_{pi}^{\mathrm{T}}(t)\boldsymbol{H}_j^{\mathrm{T}}(t+1),$$

$$\begin{aligned}
\boldsymbol{E}_{ij}(0,0,t) &= \boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t \mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t) + \boldsymbol{R}_{ij}(t),\\
\boldsymbol{E}_{ij}(k,0,t) &= \boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t)\boldsymbol{P}_{ij}(t \mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t) + \\
&\quad \boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+1)[\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t) - \boldsymbol{K}_{pi}(t)\boldsymbol{R}_{ij}(t)],\\
\boldsymbol{E}_{ij}(0,k,t) &= \boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t \mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+k,t)\boldsymbol{H}_j^{\mathrm{T}}(t+k) + \\
&\quad [\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t) - \boldsymbol{R}_{ij}(t)\boldsymbol{K}_{pj}^{\mathrm{T}}(t)]\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+k,t+1)\boldsymbol{H}_j^{\mathrm{T}}(t+k)
\end{aligned} \tag{6.4.25}$$

其中 $\boldsymbol{\Psi}_{pi}(t)$，$\boldsymbol{K}_{pi}(t)$，$\boldsymbol{P}_{ij}(t \mid t-1)$ 均由定理 6.2.1 计算.

证明　(4.6.23) 的推导同 (6.4.5) 的推导. 应用 (6.2.15) 类似于 (6.4.16) 的推导有

$$\begin{aligned}
\boldsymbol{\varepsilon}_i(t+k) &= \boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t)\tilde{\boldsymbol{x}}_i(t \mid t-1) + \sum_{\nu=1}^{k}\boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+\nu) \times \\
&\quad [\boldsymbol{\Gamma}(t+\nu-1),\ -\boldsymbol{K}_{pi}(t+\nu-1)]\begin{bmatrix}\boldsymbol{w}(t+\nu-1)\\ \boldsymbol{v}_i(t+\nu-1)\end{bmatrix} + \boldsymbol{v}_i(t+k)
\end{aligned} \tag{6.4.26}$$

这引出 (4.6.24)，再应用 (6.2.15) 可得 (6.4.25). 证毕. □

【定理 6.4.3】　由定理 6.4.1 和定理 6.4.2 给出的两种 $\boldsymbol{P}_{ij}(t \mid t+N)$ 的算法是等价的，即它们的值是恒同的.

证明　注意 $\boldsymbol{K}_{pi}(t) = \bar{\boldsymbol{K}}_{pi}(t) + \boldsymbol{J}_i(t)$，应用 (6.2.15)，用直接展开方法可验证 (6.4.8) 恒同于 (6.4.24)，(6.4.10) 恒同于 (6.4.25). 详细验证如下：

$$\begin{aligned}
\boldsymbol{E}_{ij}(1,0,t) &= \mathrm{E}[\boldsymbol{\varepsilon}_i(t+1)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t)] = \mathrm{E}\{[\boldsymbol{H}_i(t+1)(\boldsymbol{\Psi}_{pi}(t)\tilde{\boldsymbol{x}}_i(t \mid t-1) + \\
&\quad \boldsymbol{\Gamma}(t)\boldsymbol{w}(t) - \boldsymbol{K}_{pi}(t)\boldsymbol{v}_i(t)) + \boldsymbol{v}_i(t+1)][\boldsymbol{H}_j(t)\tilde{\boldsymbol{x}}_j(t \mid t-1) + \boldsymbol{v}_j(t)]^{\mathrm{T}}\} = \\
&\quad \boldsymbol{H}_i(t+1)\boldsymbol{\Psi}_{pi}(t)\boldsymbol{P}_{ij}(t \mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t) + \boldsymbol{H}_i(t+1)\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t) - \\
&\quad \boldsymbol{H}_i(t+1)\boldsymbol{K}_{pi}(t)\boldsymbol{R}_{ij}(t)
\end{aligned} \tag{6.4.27}$$

它相同于 (6.4.10) 的第 1 式. 同理可得

$$\begin{aligned}
\boldsymbol{E}_{ij}(0,1,t) &= \boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t \mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t)\boldsymbol{H}_j^{\mathrm{T}}(t+1) + \boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)\boldsymbol{H}_j^{\mathrm{T}}(t+1) - \\
&\quad \boldsymbol{R}_{ij}(t)\boldsymbol{K}_{pj}^{\mathrm{T}}(t)\boldsymbol{H}_j^{\mathrm{T}}(t+1)
\end{aligned} \tag{6.4.28}$$

它相同于 (6.4.10) 的第 2 式，而 (6.4.25) 的第 3 式恰为 (6.4.10) 第 3 式. 应用 (6.4.26) 类似可得 (6.4.25) 的第4式和第5式. 最后利用 (6.4.9) 将 (6.4.8) 中的三个中括号内分块矩阵相乘展开为

$$\begin{aligned}
&[\boldsymbol{I}_n,\ -\bar{\boldsymbol{K}}_{pi}]\begin{bmatrix}\boldsymbol{S}_{ij}^{11} & \boldsymbol{S}_{ij}^{12}\\ \boldsymbol{S}_{ij}^{21} & \boldsymbol{S}_{ij}^{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_n\\ -\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}}\end{bmatrix} = \boldsymbol{S}_{ij}^{11} - \bar{\boldsymbol{K}}_{pi}\boldsymbol{S}_{ij}^{21} - \boldsymbol{S}_{ij}^{12}\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}} - \bar{\boldsymbol{K}}_{pi}\boldsymbol{S}_{ij}^{22}\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}} = \\
&\quad \boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{J}_i\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{\Gamma}\boldsymbol{S}_j\boldsymbol{J}_j^{\mathrm{T}} + \boldsymbol{J}_i\boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}} - \bar{\boldsymbol{K}}_{pi}(\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}}) - \\
&\quad (\boldsymbol{\Gamma}\boldsymbol{S}_j - \boldsymbol{J}_i\boldsymbol{R}_{ij})\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}} + \bar{\boldsymbol{K}}_{pi}\boldsymbol{R}_{ij}\bar{\boldsymbol{K}}_{pj}
\end{aligned} \tag{6.4.29}$$

其中为简单计，省略了时标.

另一方面利用关系 $\boldsymbol{K}_{pi}(t) = \bar{\boldsymbol{K}}_{pi}(t) + \boldsymbol{J}_i(t)$，将 (6.4.24) 中三个中括号相乘项展开为

$$\begin{aligned}
&[\boldsymbol{\Gamma},\ -\boldsymbol{K}_{pi}]\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}_j\\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Gamma}^{\mathrm{T}}\\ -\boldsymbol{K}_{pj}^{\mathrm{T}}\end{bmatrix} = (\boldsymbol{\Gamma}\boldsymbol{Q} - \boldsymbol{K}_{pi}\boldsymbol{S}_i^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}} - (\boldsymbol{\Gamma}\boldsymbol{S}_j - \boldsymbol{K}_{pi}\boldsymbol{R}_{ij})\boldsymbol{K}_{pj}^{\mathrm{T}} = \\
&\quad \boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} - \bar{\boldsymbol{K}}_{pi}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{J}_i\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{\Gamma}\boldsymbol{S}_j(\boldsymbol{K}_{pj}^{\mathrm{T}} + \boldsymbol{J}_j^{\mathrm{T}}) + \\
&\quad (\bar{\boldsymbol{K}}_{pi} + \boldsymbol{J}_i)\boldsymbol{R}_{ij}(\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}} + \boldsymbol{J}_j^{\mathrm{T}})
\end{aligned} \tag{6.4.30}$$

容易验证(6.4.29)相同于(6.4.30).这证明了(6.4.8)相同于(6.4.24).证毕. □

【推论 6.4.1】 多传感器时变系统(6.2.1)和(6.2.2)在假设 1 ~ 3 下,若 $\boldsymbol{S}_i(t)=\boldsymbol{0}$, $\boldsymbol{u}(t)=\boldsymbol{0}$, $\boldsymbol{R}_{ij}(t)=\boldsymbol{0}(i\neq j)$,则局部 Kalman 预报器 $\hat{\boldsymbol{x}}(t\mid t+1)$ 由推论 6.2.1 给出,局部 Kalman 平滑器由(6.4.1)和(6.4.2)给出,局部平滑误差方差阵 $\boldsymbol{P}_i(t\mid t+N)$ 由(6.4.4)给出,局部平滑误差协方差阵 $\boldsymbol{P}_{ij}(t\mid t+N)(i\neq j)$ 由(6.4.23)给出,其中 $\boldsymbol{E}_{ij}(r,s,t)$ 为

$$
\begin{aligned}
\boldsymbol{E}_{ij}(r,s,t)=&\ \boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}(t+r,t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+s,t)\boldsymbol{H}_j^{\mathrm{T}}(t+s)+\\
&\sum_{\nu=1}^{\min(r,s)}\boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}(t+r,t+\nu)\boldsymbol{\Gamma}(t+\nu-1)\boldsymbol{Q}(t+\nu-1)\times\\
&\boldsymbol{\Gamma}^{\mathrm{T}}(t+\nu-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+s,t+\nu)\boldsymbol{H}_j^{\mathrm{T}}(t+s)
\end{aligned}\tag{6.4.31}
$$

其中 $\min(r,s)>0$.若 $\min(r,s)=0$,则

$$
\begin{aligned}
&\boldsymbol{E}_{ij}(1,0,t)=\boldsymbol{H}_i(t+1)\boldsymbol{\Psi}_{pi}(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t),\\
&\boldsymbol{E}_{ij}(0,1,t)=\boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t)\boldsymbol{H}_j^{\mathrm{T}}(t+1),\\
&\boldsymbol{E}_{ij}(0,0,t)=\boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t),\\
&\boldsymbol{E}_{ij}(k,0,t)=\boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t),\\
&\boldsymbol{E}_{ij}(0,k,t)=\boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+k,t)\boldsymbol{H}_j^{\mathrm{T}}(t+k)
\end{aligned}\tag{6.4.32}
$$

下面我们将提出求 $\boldsymbol{P}_{ij}(t\mid t+N)$ 的另一种新算法.应用(6.4.11)有

$$
\tilde{\boldsymbol{x}}_i(t\mid t+N)=\tilde{\boldsymbol{x}}_i(t\mid t-1)+\sum_{k=0}^{N}\boldsymbol{K}_i(t\mid t+k)\boldsymbol{\varepsilon}_i(t+k)\tag{6.4.33}
$$

将(6.4.26)代入上式有

$$
\begin{aligned}
\tilde{\boldsymbol{x}}_i(t\mid t+N)=&\ \tilde{\boldsymbol{x}}_i(t\mid t-1)-\sum_{k=0}^{N}\boldsymbol{K}_i(t\mid t+k)\boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t)\tilde{\boldsymbol{x}}_i(t\mid t-1)-\\
&\sum_{k=1}^{N}\sum_{\nu=1}^{K}\boldsymbol{K}_i(t\mid t+k)\boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+\nu)\boldsymbol{\Gamma}(t+\nu-1)\boldsymbol{w}(t+\nu-1)+\\
&\sum_{k=1}^{N}\sum_{\nu=1}^{K}\boldsymbol{K}_i(t\mid t+k)\boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+\nu)\boldsymbol{K}_{pi}(t+\nu-1)\boldsymbol{v}_i(t+\nu-1)-\\
&\sum_{k=0}^{N}\boldsymbol{K}_i(t\mid t+k)\boldsymbol{v}_i(t+k)
\end{aligned}\tag{6.4.34}
$$

合并 $\tilde{\boldsymbol{x}}_i(t\mid t-1)$, $\boldsymbol{w}(t+\rho)$ 和 $\boldsymbol{v}_i(t+\rho)(\rho=0,1,\cdots,N)$ 的同类项,可将(6.4.34)化简为

$$
\tilde{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{\Psi}_{iN}(t)\tilde{\boldsymbol{x}}_i(t\mid t-1)+\sum_{\rho=0}^{N}\boldsymbol{K}_{i\rho}^{w}(t)\boldsymbol{w}(t+\rho)+\sum_{\rho=0}^{N}\boldsymbol{K}_{i\rho}^{v}(t)\boldsymbol{v}_i(t+\rho)\tag{6.4.35}
$$

其中定义

$$
\boldsymbol{\Psi}_{iN}(t)=\boldsymbol{I}_n-\sum_{k=0}^{N}\boldsymbol{K}_i(t\mid t+k)\boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t)\tag{6.4.36}
$$

$$
\begin{aligned}
\boldsymbol{K}_{i\rho}^{w}(t)=&-\sum_{k=\rho+1}^{N}\boldsymbol{K}_i(t\mid t+k)\boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+\rho+1)\boldsymbol{\Gamma}(t+\rho),\\
&\rho=0,\cdots,N-1;\quad \boldsymbol{K}_{iN}^{w}(t)=\boldsymbol{0}
\end{aligned}\tag{6.4.37}
$$

$$
\boldsymbol{K}_{i\rho}^{v}(t)=\sum_{k=\rho+1}^{N}\boldsymbol{K}_i(t\mid t+k)\boldsymbol{H}_i(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+\rho+1)\boldsymbol{K}_{pi}(t+\rho)+\boldsymbol{K}_i(t\mid t+\rho),
$$

$$\rho=0,\cdots,N-1;\boldsymbol{K}_{iN}^{v}=-\boldsymbol{K}_{i}(t\mid t+N) \tag{6.4.38}$$

【定理6.4.4】 多传感器时变系统(6.2.1)和(6.2.2)在假设1～4下,平滑误差协方差阵 $\boldsymbol{P}_{ij}(t\mid t+N)(N>0)(i\neq j)$,有公式

$$\boldsymbol{P}_{ij}(t\mid t+N)=\boldsymbol{\Psi}_{iN}(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{jN}^{\mathrm{T}}(t)+\sum_{\rho=0}^{N}[\boldsymbol{K}_{i\rho}^{w}(t),\boldsymbol{K}_{i\rho}^{v}(t)]\begin{bmatrix}\boldsymbol{Q}(t+\rho) & \boldsymbol{S}_{j}(t+\rho)\\ \boldsymbol{S}_{i}^{\mathrm{T}}(t+\rho) & \boldsymbol{R}_{ij}(t+\rho)\end{bmatrix}\begin{bmatrix}\boldsymbol{K}_{j\rho}^{w\mathrm{T}}(t)\\ \boldsymbol{K}_{j\rho}^{v\mathrm{T}}(t)\end{bmatrix} \tag{6.4.39}$$

特别对 $i=j$ 有 $\boldsymbol{P}_{i}(t\mid t+N)$ 的不同于(6.4.4)的公式

$$\boldsymbol{P}_{i}(t\mid t+N)=\boldsymbol{\Psi}_{iN}(t)\boldsymbol{P}_{i}(t\mid t-1)\boldsymbol{\Psi}_{iN}^{\mathrm{T}}(t)+\sum_{\rho=0}^{N}[\boldsymbol{K}_{i\rho}^{w}(t),\boldsymbol{K}_{i\rho}^{v}(t)]\begin{bmatrix}\boldsymbol{Q}(t+\rho) & \boldsymbol{S}_{j}(t+\rho)\\ \boldsymbol{S}_{i}^{\mathrm{T}}(t+\rho) & \boldsymbol{R}_{i}(t+\rho)\end{bmatrix}\begin{bmatrix}\boldsymbol{K}_{i\rho}^{w\mathrm{T}}(t)\\ \boldsymbol{K}_{i\rho}^{v\mathrm{T}}(t)\end{bmatrix} \tag{6.4.40}$$

证明 (6.4.35)可写为

$$\tilde{\boldsymbol{x}}_{i}(t\mid t+N)=\boldsymbol{\Psi}_{iN}(t)\tilde{\boldsymbol{x}}_{i}(t\mid t-1)+\sum_{\rho=0}^{N}[\boldsymbol{K}_{i\rho}^{w}(t),\boldsymbol{K}_{i\rho}^{v}(t)]\begin{bmatrix}\boldsymbol{w}(t+\rho)\\ \boldsymbol{v}_{i}(t+\rho)\end{bmatrix} \tag{6.4.41}$$

注意 $\tilde{\boldsymbol{x}}_{i}(t\mid t-1)$ 不相关于 $\boldsymbol{w}(t+\rho)$ 和 $\boldsymbol{v}_{j}(t+\rho)$,$\rho=0,\cdots,N$,于是应用(6.2.3)和(6.2.4)和上式可得(6.4.39)和(6.4.40). □

【注6.4.1】 从计算角度(6.4.4)比(6.4.40)简单.

【定理6.4.5】 多传感器时变系统(6.2.1)和(6.2.2)在假设1～4下,平滑误差协方差阵 $\boldsymbol{P}_{ij}(t\mid t+N)(i\neq j)$ 有另一种公式

$$\boldsymbol{P}_{ij}(t\mid t+N)=\boldsymbol{\Psi}_{iN}(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{jN}^{\mathrm{T}}(t)+\sum_{\rho=0}^{N}[\boldsymbol{M}_{i\rho}^{\bar{w}}(t),\boldsymbol{M}_{i\rho}^{v}(t)]\times\begin{bmatrix}\boldsymbol{S}_{ij}^{11}(t+\rho) & \boldsymbol{S}_{ij}^{12}(t+\rho)\\ \boldsymbol{S}_{ij}^{21}(t+\rho) & \boldsymbol{S}_{ij}^{22}(t+\rho)\end{bmatrix}\begin{bmatrix}\boldsymbol{M}_{j\rho}^{\bar{w}\mathrm{T}}(t)\\ \boldsymbol{M}_{j\rho}^{v\mathrm{T}}(t)\end{bmatrix} \tag{6.4.42}$$

其中定义

$$\boldsymbol{\Psi}_{iN}(t)=\boldsymbol{I}_{n}-\sum_{k=0}^{N}\boldsymbol{K}_{i}(t\mid t+k)\boldsymbol{H}_{i}(t+k)\boldsymbol{\Psi}_{pi}(t+k,t)$$

$$\boldsymbol{M}_{i\rho}^{\bar{w}\mathrm{T}}(t)=-\sum_{k=\rho+1}^{N}\boldsymbol{K}_{i}(t\mid t+k)\boldsymbol{H}_{i}(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+\rho+1),$$

$$\rho=0,\cdots,N-1,\quad \boldsymbol{M}_{iN}^{\bar{w}}(t)=\boldsymbol{0}$$

$$\boldsymbol{M}_{i\rho}^{v}(t)=\sum_{k=\rho+1}^{N}\boldsymbol{K}_{i}(t\mid t+k)\boldsymbol{H}_{i}(t+k)\boldsymbol{\Psi}_{pi}(t+k,t+\rho+1)\bar{\boldsymbol{K}}_{pi}(t+\rho)-\boldsymbol{K}_{i}(t\mid t+\rho),$$

$$\rho=0,\cdots,N-1,\quad \boldsymbol{M}_{iN}^{v}(t)=-\boldsymbol{K}_{i}(t\mid t+N) \tag{6.4.43}$$

其中 $\boldsymbol{S}_{ij}^{\alpha\beta}(t+\rho)$,$\alpha,\beta=1,2$,由(6.4.9)置 $t=t+\rho$ 得到.

特别对 $i=j$,有 $\boldsymbol{P}_{i}(t\mid t+N)$ 的另一种公式

$$\boldsymbol{P}_{i}(t\mid t+N)=\boldsymbol{\Psi}_{iN}(t)\boldsymbol{P}_{i}(t\mid t-1)\boldsymbol{\Psi}_{iN}^{\mathrm{T}}(t)+\sum_{\rho=0}^{N}[\boldsymbol{M}_{i\rho}^{\bar{w}}(t),\boldsymbol{M}_{i\rho}^{v}(t)]\begin{bmatrix}\boldsymbol{S}_{ii}^{11}(t+\rho) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{R}_{i}(t+\rho)\end{bmatrix}\begin{bmatrix}\boldsymbol{M}_{i\rho}^{\bar{w}\mathrm{T}}(t)\\ \boldsymbol{M}_{i\rho}^{v\mathrm{T}}(t)\end{bmatrix} \tag{6.4.44}$$

$$S_{ii}^{11}(t+\rho)=\boldsymbol{\Gamma}(t+\rho)(\boldsymbol{Q}(t+\rho)-\boldsymbol{S}_i(t+\rho)\boldsymbol{R}_i(t+\rho)\boldsymbol{S}_i^{\mathrm{T}}(t+\rho))\boldsymbol{\Gamma}^{\mathrm{T}}(t+\rho)\tag{6.4.45}$$

证明　利用公式(6.4.16)类似于定理6.4.1和定理6.4.2可得证.　□

【定理6.4.6】　平滑误差协方差阵 $\boldsymbol{P}_{ij}(t\mid t+N)$ 的两种算法(6.4.39)～(6.4.40)与(6.4.42)～(6.4.43)是等价(等值)的.

证明　详细证明留给读者,从略.　□

【定理6.4.7】　多传感器时变系统(6.2.1)和(6.2.2)在假设1～4下有最优信息融合Kalman平滑器 $\tilde{\boldsymbol{x}}_0(t\mid t+N)$ 为

$$\hat{\boldsymbol{x}}_0(t\mid t+N)=\sum_{i=1}^{L}\boldsymbol{A}_i(t)\hat{\boldsymbol{x}}_i(t\mid t+N)\tag{6.4.46}$$

其中 $\hat{\boldsymbol{x}}_i(t\mid t+N)$ 为由引理6.4.1计算的局部Kalman平滑器.在按矩阵加权准则下,$\boldsymbol{A}_i(t)$ 为加权阵;在按标量加权准则下,$\boldsymbol{A}_i(t)$ 为加权系数,即 $\boldsymbol{A}_i(t)$ 为特殊形式的加权阵 $\boldsymbol{A}_i(t)=a_i(t)\boldsymbol{I}_n$,$a_i(t)$ 为加权系数;在按分量标量加权准则下,$\boldsymbol{A}_i(t)$ 是对角阵.利用由引理6.4.1和定理6.4.1和定理6.4.2计算的平滑误差方差阵和协方差阵 $\boldsymbol{P}_i(t\mid t+N)$ 和 $\boldsymbol{P}_{ij}(t\mid t+N)$,按节6.1的三种最优加权融合准则和公式可求得加权 $\boldsymbol{A}_i(t)$,且最优融合平滑误差方差阵 $\boldsymbol{P}_0(t\mid t+N)$ 可用相应的公式求得,且有关系

$$\mathrm{tr}\boldsymbol{P}_0(t\mid t+N)\leqslant\mathrm{tr}\boldsymbol{P}_i(t\mid t+N),\quad i=1,\cdots,L\tag{6.4.47}$$

6.5　时变系统多传感器信息融合白噪声估值器

白噪声估计问题在石油地震勘探中有重要的应用背景[41,42].本节考虑时变系统多传感器信息融合白噪声估值器设计问题.时变系统按标量加权信息融合白噪声反卷积滤波器新近已由文献[37]给出,但它不能处理平滑估计问题.本节在统一框架下提出了按矩阵、对角阵和标量三种加权多传感器信息融合白噪声滤波器和平滑器,它包括文献[37]和[24]的结果作为特例.

考虑多传感器时变系统(6.2.1)和(6.2.2),在假设1～4下,问题是求相应于第 i 传感器子系统的白噪声 $\boldsymbol{w}(t)$ 的局部反卷积滤波器 $\hat{\boldsymbol{w}}_i(t\mid t+N)(N\geqslant 0)$,$i=1,\cdots,L$ 和信息融合白噪声反卷积滤波器和平滑器 $\hat{\boldsymbol{w}}_0(t\mid t+N)$,它由局部白噪声估值器 $\hat{\boldsymbol{w}}_i(t\mid t+N)(i=1,\cdots,L)$ 加权构成.注意,白噪声 $\boldsymbol{w}(t)$ 是系统(6.2.1)和(6.2.2)的输入,由输出估计系统的输入叫反卷积.解决问题的难点在于如何求各传感器白噪声估计误差的互协方差阵.

【引理6.5.1】　多传感器时变系统(6.2.1)和(6.2.2)在假设1～4下,第 i 传感器子系统有局部最优白噪声反卷积估值器

$$\hat{\boldsymbol{w}}_i(t\mid t+N)=\sum_{j=0}^{N}\boldsymbol{M}_i(t\mid t+j)\boldsymbol{\varepsilon}_i(t+j),\quad N\geqslant 0,i=1,\cdots,L\tag{6.5.1}$$

其中定义

$$\boldsymbol{M}(t\mid t)=\boldsymbol{S}_i(t)\boldsymbol{Q}_{\varepsilon i}^{-1}(t),$$

$$\boldsymbol{M}_i(t\mid t+1)=\boldsymbol{D}_i(t)\boldsymbol{H}_i^{\mathrm{T}}(t+1)\boldsymbol{Q}_{\varepsilon i}^{-1}(t+1),$$

$$\boldsymbol{M}_i(t\mid t+j)=\boldsymbol{D}_i(t)\{\prod_{k=1}^{j-1}\boldsymbol{\Psi}_{pi}^{\mathrm{T}}(t+k)\}\boldsymbol{H}_i^{\mathrm{T}}(t+j)\boldsymbol{Q}_{\varepsilon i}^{-1}(t+j),\quad j>1 \tag{6.5.2}$$

其中 $\boldsymbol{D}_i(t)$ 有两种算法:

$$\boldsymbol{D}_i(t)=\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{S}_i(t)\boldsymbol{K}_{pi}^{\mathrm{T}}(t) \tag{6.5.3}$$

或

$$\boldsymbol{D}_i(t)=-\boldsymbol{S}_i(t)\boldsymbol{K}_{fi}^{\mathrm{T}}(t)\overline{\boldsymbol{\Phi}}_i^{\mathrm{T}}(t)+\boldsymbol{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{S}_i(t)\boldsymbol{K}_{pi}^{\mathrm{T}}(t) \tag{6.5.4}$$

其中 $\boldsymbol{\Psi}_{pi}(t),\boldsymbol{K}_{pi}(t),\boldsymbol{K}_{fi}(t),\overline{\boldsymbol{\Phi}}_i(t),\boldsymbol{\varepsilon}_i(t),\boldsymbol{Q}_{\varepsilon i}(t)$ 由定理6.2.1 ~ 定理6.2.3计算,且应用 Riccati 方程(6.2.9)或(6.2.23).

局部估值误差 $\hat{\boldsymbol{w}}_i(t\mid t+N)=\boldsymbol{w}(t)-\hat{\boldsymbol{w}}_i(t\mid t+N)$ 方差阵 $\boldsymbol{P}_i^w(t\mid t+N)=\mathrm{E}[\tilde{\boldsymbol{w}}_i(t\mid t+N)\tilde{\boldsymbol{w}}_i^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}_i^w(t\mid t+N)=\boldsymbol{Q}(t)-\sum_{j=0}^{N}\boldsymbol{M}_i(t\mid t+j)\boldsymbol{Q}_{\varepsilon i}(t+j)\boldsymbol{M}_i^{\mathrm{T}}(t\mid t+j) \tag{6.5.5}$$

其中 $N\geqslant 0$.

证明 见定理3.7.12,定理3.7.13和推论3.7.9. □

【定理6.5.1】 多传感器时变系统(6.2.1)和(6.2.2)在假设1 ~ 4下,局部白噪声估计误差互协方差阵 $\boldsymbol{P}_{ij}^w(t\mid t+N)=\mathrm{E}[\tilde{\boldsymbol{w}}_i(t\mid t+N)\tilde{\boldsymbol{w}}_j^{\mathrm{T}}(t\mid t+N)](i\neq j)$ 有公式

$$\begin{aligned}\boldsymbol{P}_{ij}^w(t\mid t+N)=&\boldsymbol{Q}(t)-\sum_{r=0}^{N}\boldsymbol{M}_i(t\mid t+r)\boldsymbol{Q}_{\varepsilon i}(t+r)\boldsymbol{M}_i^{\mathrm{T}}(t\mid t+r)-\\&\sum_{s=0}^{N}\boldsymbol{M}_j(t\mid t+s)\boldsymbol{Q}_{\varepsilon j}(t+s)\boldsymbol{M}_j^{\mathrm{T}}(t\mid t+s)+\\&\sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{M}_i(t\mid t+r)\boldsymbol{E}_{ij}(r,s,t)\boldsymbol{M}_j^{\mathrm{T}}(t\mid t+s)\end{aligned} \tag{6.5.6}$$

其中 $N\geqslant 0$. 若 $\min(r,s)>0$,则有

$$\begin{aligned}\boldsymbol{E}_{ij}(r,s,t)=&\boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}(t+r,t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+s,t)\boldsymbol{H}_j^{\mathrm{T}}(t+s)+\\&\sum_{k=1}^{\min(r,s)}\boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}(t+r,t+k)[\boldsymbol{\Gamma}(t+k-1),-\boldsymbol{K}_{pi}(t-k-1)]\times\\&\begin{bmatrix}\boldsymbol{Q}(t+k-1) & \boldsymbol{S}_j(t+k-1)\\ \boldsymbol{S}_i^{\mathrm{T}}(t+k-1) & \boldsymbol{R}_{ij}(t+k-1)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Gamma}^{\mathrm{T}}(t+k-1)\\ -\boldsymbol{K}_{pi}^{\mathrm{T}}(t+k-1)\end{bmatrix}\times\\&\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+s,t+k)\boldsymbol{H}_j^{\mathrm{T}}(t+s)\end{aligned} \tag{6.5.7}$$

若 $\min(r,s)=0$,则有

$$\begin{aligned}\boldsymbol{E}_{ij}(0,0,t)=&\boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t)+\boldsymbol{R}_{ij}(t),\\\boldsymbol{E}_{ij}(0,s,t)=&\boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+s,t)\boldsymbol{H}_j^{\mathrm{T}}(t+s)+\\&[\boldsymbol{S}_i^{\mathrm{T}}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t)-\boldsymbol{R}_{ij}(t)\boldsymbol{K}_{pj}^{\mathrm{T}}(t)]\boldsymbol{\Psi}_{pj}^{\mathrm{T}}(t+s,t+1)\boldsymbol{H}_j^{\mathrm{T}}(t+s),\\\boldsymbol{E}_{ij}(r,0,t)=&\boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}(t+r,t)\boldsymbol{P}_{ij}(t\mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t)+\\&\boldsymbol{H}_i(t+r)\boldsymbol{\Psi}_{pi}(t+r,t+1)[\boldsymbol{\Gamma}(t)\boldsymbol{S}_j(t)-\boldsymbol{K}_{pi}(t)\boldsymbol{R}_{ij}(t)]\end{aligned} \tag{6.5.8}$$

其中 $\boldsymbol{P}_{ij}(t\mid t-1)$ 由定理6.2.1和定理6.2.2计算.

证明 由 $\boldsymbol{P}_{ij}^w(t\mid t+N)$ 的定义有

$$\boldsymbol{P}_{ij}^w(t\mid t+N)=\mathrm{E}[(\boldsymbol{w}(t)-\hat{\boldsymbol{w}}_i(t\mid t+N)(\boldsymbol{w}(t)-\hat{\boldsymbol{w}}_j(t\mid t+N))^{\mathrm{T}}]=$$

$$\boldsymbol{Q}(t) - \mathrm{E}[\hat{\boldsymbol{w}}_i(t \mid t+N)\boldsymbol{w}^{\mathrm{T}}(t)] - \mathrm{E}[\boldsymbol{w}(t)\hat{\boldsymbol{w}}_j^{\mathrm{T}}(t \mid t+N)^{\mathrm{T}}] + \mathrm{E}[\hat{\boldsymbol{w}}_i(t \mid t+N)\hat{\boldsymbol{w}}_j^{\mathrm{T}}(t \mid t+N)] \tag{6.5.9}$$

由射影公式有

$$\boldsymbol{M}_i(t \mid t+j) = \mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(t+j)]\boldsymbol{Q}_{\varepsilon i}^{-1}(t+j) \tag{6.5.10}$$

这引出

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(t+j)] = \boldsymbol{M}_i(t \mid t+j)\boldsymbol{Q}_{\varepsilon i}(t+j) \tag{6.5.11}$$

于是应用(6.5.1)有

$$\begin{aligned}\mathrm{E}[\hat{\boldsymbol{w}}_i(t \mid t+N)\boldsymbol{w}^{\mathrm{T}}(t)] &= \sum_{j=0}^{N}\boldsymbol{M}_i(t \mid t+j)\mathrm{E}[\boldsymbol{\varepsilon}_i(t+j)\boldsymbol{w}^{\mathrm{T}}(t)] = \\ &\sum_{j=0}^{N}\boldsymbol{M}_i(t \mid t+j)\boldsymbol{Q}_{\varepsilon i}(t+j)\boldsymbol{M}_i^{\mathrm{T}}(t \mid t+j)\end{aligned} \tag{6.5.12}$$

$$\begin{aligned}\mathrm{E}[\boldsymbol{w}(t)\hat{\boldsymbol{w}}_j^{\mathrm{T}}(t \mid t+N)] &= \sum_{s=0}^{N}\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)]\boldsymbol{M}_j^{\mathrm{T}}(t \mid t+s) = \\ &\sum_{s=0}^{N}\boldsymbol{M}_j(t \mid t+s)\boldsymbol{Q}_{\varepsilon j}(t+s)\boldsymbol{M}_j^{\mathrm{T}}(t \mid t+s)\end{aligned} \tag{6.5.13}$$

$$\mathrm{E}[\hat{\boldsymbol{w}}_i(t \mid t+N)\hat{\boldsymbol{w}}_j^{\mathrm{T}}(t \mid t+N)] = \sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{M}_i(t \mid t+r)\boldsymbol{E}_{ij}(r,s,t)\boldsymbol{M}_j^{\mathrm{T}}(t \mid t+s) \tag{6.5.14}$$

其中定义 $\boldsymbol{E}_{ij}(r,s,t) = \mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)]$. 由(6.5.9) ~ (6.5.14)得(6.5.6). 下面求 $\boldsymbol{E}_{ij}(r,s,t)$.

当 $\min(r,s) > 0$ 时,将(6.4.26)代入 $\boldsymbol{E}_{ij}(r,s,t) = \mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)]$ 中可得(6.5.7). 当 $\min(r,s) = 0$ 时,将(6.4.26)代入 $\boldsymbol{E}_{ij}(0,s,t) = \mathrm{E}[\boldsymbol{\varepsilon}_i(t)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)]$ 中,并将 $\boldsymbol{\varepsilon}_i(t) = \boldsymbol{H}_i(t)\tilde{\boldsymbol{x}}_i(t \mid t-1) + \boldsymbol{v}_i(t)$ 也代入其中可得(6.5.8)的第2式. 将(6.4.26)和 $\boldsymbol{\varepsilon}_j(t) = \boldsymbol{H}_j(t)\tilde{\boldsymbol{x}}_j(t \mid t-1) + \boldsymbol{v}_j(t)$ 代入 $\boldsymbol{E}_{ij}(r,0,t) = \mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t)]$ 中可得(6.5.8)的第3式. 将 $\boldsymbol{\varepsilon}_i(t)$ 和 $\boldsymbol{\varepsilon}_j(t)$ 的上述表达式代入 $\boldsymbol{E}_{ij}(0,0,t) = \mathrm{E}[\boldsymbol{\varepsilon}_i(t)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t)]$ 中可得(6.5.8)的第1式. 证毕. □

【推论 6.5.1】[37] 在定理 6.5.1 条件下,对 $N = 0$,有局部白噪声反卷积滤波器误差互协方差阵 $\boldsymbol{P}_{ij}^{w}(t \mid t)$ 为

$$\begin{aligned}\boldsymbol{P}_{ij}^{w}(t \mid t) = &\boldsymbol{Q}(t) - \boldsymbol{S}_i(t)\boldsymbol{Q}_{\varepsilon i}^{-1}(t)\boldsymbol{S}_i^{\mathrm{T}}(t) - \boldsymbol{S}_j(t)\boldsymbol{Q}_{\varepsilon j}^{-1}(t)\boldsymbol{S}_j^{\mathrm{T}}(t) + \\ &\boldsymbol{S}_i(t)\boldsymbol{Q}_{\varepsilon i}^{-1}(t)[\boldsymbol{H}_i(t)\boldsymbol{P}_{ij}(t \mid t-1)\boldsymbol{H}_j^{\mathrm{T}}(t) + \boldsymbol{R}_{ij}(t)]\boldsymbol{Q}_{\varepsilon j}^{-1}(t)\boldsymbol{S}_j^{\mathrm{T}}(t)\end{aligned} \tag{6.5.15}$$

【定理 6.5.2】 多传感器时变系统(6.5.1)和(6.2.2)在假设 1 ~ 4 下,有最优融合输入白噪声估值器

$$\hat{\boldsymbol{w}}_0(t \mid t+N) = \sum_{i=1}^{L}\boldsymbol{A}_i(t)\hat{\boldsymbol{w}}_i(t \mid t+N), \quad N \geqslant 0 \tag{6.5.16}$$

其中 $\hat{\boldsymbol{w}}_i(t \mid t+N)$ 由引理 6.5.1 计算. 在节 6.1 的三种加权最优融合准则下,加权 $A_i(t)$ 分别为矩阵、对角阵或标量,可由节 6.1 的三种加权融合公式基于由(6.5.5)和(6.5.6)给出的 $\boldsymbol{P}_i^{w}(t \mid t+N)$ 和 $\boldsymbol{P}_{ij}^{w}(t \mid t+N)$ 求得,且最优融合误差方差阵 $\boldsymbol{P}_0^{w}(t \mid t+N)$ 满足

$$\mathrm{tr}\boldsymbol{P}_0^{w}(t \mid t+N) \leqslant \mathrm{tr}\boldsymbol{P}_i^{w}(t \mid t+N), \quad i = 1,\cdots,L \tag{6.5.17}$$

【注 6.5.1】 当 $N<0$ 时有

$$\hat{\boldsymbol{w}}_i(t\mid t+N)=\boldsymbol{0},\quad N<0,i=1,\cdots,L \tag{6.5.18}$$

$$\boldsymbol{P}_i^w(t\mid t+N)=\boldsymbol{Q},\quad N<0 \tag{6.5.19}$$

因而有最优融合白噪声估值器

$$\hat{\boldsymbol{w}}_0(t\mid t+N)=\boldsymbol{0},\quad N<0 \tag{6.5.20}$$

$$\boldsymbol{P}_0^w(t\mid t+N)=\boldsymbol{Q},\quad N<0 \tag{6.5.21}$$

6.6 定常系统多传感器信息融合稳态 Kalman 估值器和白噪声估值器

考虑多传感器定常（时不变）线性离散随机系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{Bu}(t)+\boldsymbol{\Gamma w}(t) \tag{6.6.1}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{H}_i\boldsymbol{x}(t)+\boldsymbol{v}_i(t),\quad i=1,\cdots,L \tag{6.6.2}$$

其中 t 为离散时间，状态 $\boldsymbol{x}(t)\in R^n$，第 i 传感器的观测 $\boldsymbol{y}_i(t)\in R^{m_i}$，控制 $\boldsymbol{u}(t)\in R^p$，$\boldsymbol{\Phi}$，$\boldsymbol{B}$，$\boldsymbol{\Gamma}$ 和 $\boldsymbol{H}_i$ 是已知的适当维数的常阵. $\boldsymbol{w}(t)$ 为输入噪声，$\boldsymbol{v}_i(t)$ 为观测噪声.

【假设 1】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}_i(t)\in R^{m_i}$ 是带零均值的相关白噪声：

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}_i(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{w}^{\mathrm{T}}(j) & \boldsymbol{v}_i^{\mathrm{T}}(j)\end{bmatrix}\right\}=\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}_i\\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_i\end{bmatrix}\delta_{tj} \tag{6.6.3}$$

其中 E 为均值号，T 为转置号，$\delta_{tt}=1$，$\delta_{tj}=0(t\neq j)$，$\boldsymbol{w}(t)$ 与 $\boldsymbol{v}_i(t)$ 的方差阵和相关阵各为 $\boldsymbol{Q}$，$\boldsymbol{R}_i$ 和 $\boldsymbol{S}_i$.

【假设 2】 观测噪声 $\boldsymbol{v}_i(t)$ 和 $\boldsymbol{v}_j(t)$ 是相关的，即

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{v}_i(t)\\ \boldsymbol{v}_j(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{v}_i^{\mathrm{T}}(k) & \boldsymbol{v}_j^{\mathrm{T}}(k)\end{bmatrix}\right\}=\begin{bmatrix}\boldsymbol{R}i & \boldsymbol{R}_{ij}\\ \boldsymbol{R}_{ij}^{\mathrm{T}} & \boldsymbol{R}_j\end{bmatrix}\delta_{tk} \tag{6.6.4}$$

且定义 $\boldsymbol{R}_i=\boldsymbol{R}_{ii}$.

【假设 3】[43] $(\boldsymbol{\Phi},\boldsymbol{H}_i)$ 为完全可观对，$(\bar{\boldsymbol{\Phi}},\boldsymbol{\Gamma}\bar{\boldsymbol{Q}}_i^{1/2})$ 为完全能稳对，其中 $\bar{\boldsymbol{Q}}_i=\boldsymbol{Q}-\boldsymbol{S}_i\boldsymbol{R}_i^{-1}\boldsymbol{S}_i^{\mathrm{T}}$，$\bar{\boldsymbol{Q}}_i=\bar{\boldsymbol{Q}}_i^{1/2}(\bar{\boldsymbol{Q}}_i^{1/2})^{\mathrm{T}}$，$\bar{\boldsymbol{\Phi}}_i=\boldsymbol{\Phi}-\boldsymbol{J}_i\boldsymbol{H}_i$，$\boldsymbol{J}_i=\boldsymbol{\Gamma}\boldsymbol{S}_i\boldsymbol{R}_i^{-1}$.

【假设 4】 初始观测时刻 $t_0=-\infty$.

【假设 5】 $\boldsymbol{u}(t)$ 是已知的时间序列，或 $\boldsymbol{u}(t)$ 是 $(\boldsymbol{y}_i(t),\boldsymbol{y}_i(t-1),\cdots)$ 的线性函数（反馈控制）.

问题是基于无限观测历史 $(\boldsymbol{y}_i(t),\boldsymbol{y}_i(t-1),\cdots,\boldsymbol{u}(t-1),\boldsymbol{u}(t-2),\cdots)$，求局部稳态最优 Kalman 滤波器 $\hat{\boldsymbol{x}}_i(t\mid t)$ 和预报器 $\hat{\boldsymbol{x}}_i(t+1\mid t)$，$i=1,\cdots,L$，并在三种加权融合准则下求相应的信息融合稳态最优 Kalman 滤波器 $\hat{\boldsymbol{x}}_0(t\mid t)$ 和预报器 $\hat{\boldsymbol{x}}_0(t+1\mid t)$，并进一步求信息融合稳态 Kalman 平滑器和多步预报器，并求局部和信息融合稳态白噪声估值器 $\hat{\boldsymbol{w}}_i(t\mid t+N)$ 和 $\hat{\boldsymbol{w}}_0(t\mid t+N)$.

注意，假设 3 保证了各传感器子系统存在局部稳态 Kalman 滤波器和预报器.

应用节 6.2 的非稳态最优 Kalman 滤波器和预报器，置初始时刻 $t_0\to\infty$ 可得相应的稳态 Kalman 滤波器和预报器.

【定理 6.6.1】 定常系统(6.6.1)和(6.6.2)在假设 1 ~ 5 下，第 i 传感器子系统有局

部最优递推 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}\hat{\boldsymbol{x}}_i(t\mid t-1)+\boldsymbol{B}\boldsymbol{u}(t)+\boldsymbol{K}_{pi}\boldsymbol{y}_i(t) \tag{6.6.5}$$

$$\boldsymbol{\Psi}_{pi}=\boldsymbol{\Phi}-\boldsymbol{K}_{pi}\boldsymbol{H}_i,$$

$$\boldsymbol{K}_{pi}=[\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S}_i]\boldsymbol{Q}_{\varepsilon i}^{-1},$$

$$\boldsymbol{Q}_{\varepsilon i}=\boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{R}_i \tag{6.6.6}$$

其中稳态预报误差方差阵 $\boldsymbol{\Sigma}_i=\lim\boldsymbol{P}_i(t\mid t-1)(t_0\to-\infty$ 或 $t\to\infty)$ 满足 Riccati 方程

$$\boldsymbol{\Sigma}_i=\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{\Phi}^{\mathrm{T}}-[\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S}_i][\boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{R}_i]^{-1}\times$$

$$[\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S}_i]^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} \tag{6.6.7}$$

它可由(6.2.9)用迭代法求解.

稳态预报误差互协方差阵 $\boldsymbol{\Sigma}_{ij}=\lim\boldsymbol{P}_{ij}(t\mid t-1)(t_0\to-\infty$ 或 $t\to\infty)$ 满足 Lyapunov 方程

$$\boldsymbol{\Sigma}_{ij}=\boldsymbol{\Psi}_{pi}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}}+\boldsymbol{\Delta}_{pij} \tag{6.6.8}$$

其中 $i,j=1,\cdots,L$，且

$$\boldsymbol{\Delta}_{pij}=[\boldsymbol{\Gamma},-\boldsymbol{K}_{pi}]\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}_j\\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Gamma}^{\mathrm{T}}\\ -\boldsymbol{K}_{pj}^{\mathrm{T}}\end{bmatrix} \tag{6.6.9}$$

或等价地有

$$\boldsymbol{\Delta}_{pij}=\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{K}_{pi}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{\Gamma}\boldsymbol{S}_j\boldsymbol{K}_{pj}^{\mathrm{T}}+\boldsymbol{K}_{pi}\boldsymbol{R}_{ij}\boldsymbol{K}_{pj}^{\mathrm{T}} \tag{6.6.10}$$

定义 $\boldsymbol{\Sigma}_i=\boldsymbol{\Sigma}_{ii}$，在(6.6.8) ~ (6.6.10)中置 $i=j$ 有 $\boldsymbol{\Sigma}_i$ 满足 Lyapunov 方程

$$\boldsymbol{\Sigma}_i=\boldsymbol{\Psi}_{pi}\boldsymbol{\Sigma}_i\boldsymbol{\Psi}_{pi}^{\mathrm{T}}+\boldsymbol{\Delta}_{pii} \tag{6.6.11}$$

$$\boldsymbol{\Delta}_{pii}=\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}+\boldsymbol{K}_{pi}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{\Gamma}\boldsymbol{S}_i\boldsymbol{K}_{pi}^{\mathrm{T}}+\boldsymbol{K}_{pi}\boldsymbol{R}_i\boldsymbol{K}_{pi}^{\mathrm{T}} \tag{6.6.12}$$

证明　由定理 6.2.1 令 $t\to\infty$ 得证. □

【定理 6.6.2】　(Lyapunov 方程迭代解的指数收敛性)Lyapunov 方程(6.6.10)可用如下迭代法求解：

$$\boldsymbol{\Sigma}_{ij}(t+1)=\boldsymbol{\Psi}_{pi}\boldsymbol{\Sigma}_{ij}(t)\boldsymbol{\Psi}_{pj}^{\mathrm{T}}+\boldsymbol{\Delta}_{pij} \tag{6.6.13}$$

带初值

$$\boldsymbol{\Sigma}_{ij}(0)=\boldsymbol{P}_0 \quad 或 \quad \boldsymbol{\Sigma}_{ij}(0)=\boldsymbol{0} \tag{6.6.14}$$

我们有

$$\lim_{t\to\infty}\boldsymbol{\Sigma}_{ij}(t)=\boldsymbol{\Sigma}_{ij},\quad i,j=1,\cdots,L,i\neq j \tag{6.6.15}$$

且在任意矩阵范数 $\|\cdot\|$ 意义下有指数收敛速度为

$$\|\boldsymbol{\delta}_{ij}(t)\|\leqslant a\rho^t,\quad 0<\rho<1,a>0 \tag{6.6.16}$$

其中定义迭代误差阵 $\boldsymbol{\delta}_{ij}(t)=\boldsymbol{\Sigma}_{ij}-\boldsymbol{\Sigma}_{ij}(t)$，即 $\|\boldsymbol{\delta}_{ij}(t)\|$ 将以指数律 ρ^t 的速度衰减为零.

证明　类似于定理 5.10.2 的证明，由(6.6.10)减(6.6.13)引出

$$\boldsymbol{\delta}_{ij}(t+1)=\boldsymbol{\Psi}_{pi}\boldsymbol{\delta}_{ij}(t)\boldsymbol{\Psi}_{pj}^{\mathrm{T}} \tag{6.6.17}$$

将它迭代 t 次有

$$\boldsymbol{\delta}_{ij}(t)=\boldsymbol{\Psi}_{pi}^t\boldsymbol{\delta}_{ij}(0)\boldsymbol{\Psi}_{pj}^{\mathrm{T}t} \tag{6.6.18}$$

由稳态 Kalman 预报器性质[43]有 $\boldsymbol{\Psi}_{pi}$ 为稳定矩阵，即它的特征值均位于单位圆内. 设 $\boldsymbol{\Psi}_{pi}$ 的特征值为 $\lambda_{i1},\cdots,\lambda_{in}$，$|\lambda_{ir}|<1,r=1,\cdots,n$，则 $\boldsymbol{\Psi}_{pi}$ 的谱半径为 $\rho_{i0}=\max(|\lambda_{i1}|,\cdots,$

$|\lambda_{in}|$)，且有 $0<\rho_{i0}<1$. 应用矩阵理论[44]，取 $\varepsilon_i>0$ 使 $\rho_i=\rho_{i0}+\varepsilon_i=\rho_i<1$，则存在与 $\boldsymbol{\Psi}_{pi}$ 有关的矩阵范数 $\|\cdot\|_i$ 使

$$\|\boldsymbol{\Psi}_{pi}\|_i\leqslant\rho_{i0}+\varepsilon_i=\rho_i<1 \tag{6.6.19}$$

因 $\boldsymbol{\Psi}_{pj}$ 是一个稳态矩阵，而 $\boldsymbol{\Psi}_{pj}^{\mathrm{T}}$ 与 $\boldsymbol{\Psi}_{pj}$ 有相同特征值，故 $\boldsymbol{\Psi}_{pj}^{\mathrm{T}}$ 也是稳定矩阵. 因而

$$\lim_{t\to\infty}\boldsymbol{\Psi}_{pj}^{\mathrm{T}t}=\mathbf{0} \tag{6.6.20}$$

这引出

$$\lim_{t\to\infty}\|\boldsymbol{\Psi}_{pj}^{\mathrm{T}t}\|_i=\mathbf{0} \tag{6.6.21}$$

故有

$$\|\boldsymbol{\Psi}_{pj}^{\mathrm{T}t}\|_i<M_i,\quad i=1,\cdots,L \tag{6.6.22}$$

由矩阵范数性质，对(6.6.18)取范数运算有

$$\|\boldsymbol{\delta}_{ij}(t)\|_i\leqslant\|\boldsymbol{\Psi}_{pi}\|_i^t\|\boldsymbol{\delta}_{ij}(0)\|_i\|\boldsymbol{\Psi}_{pj}^{\mathrm{T}t}\|_i \tag{6.6.23}$$

应用(6.6.19)和(6.6.22)引出

$$\|\boldsymbol{\delta}_{ij}(t)\|_i\leqslant c\rho^t \tag{6.6.24}$$

其中 $\rho=\max(\rho_1,\cdots,\rho_L)$，$0<\rho<1$，$c=\max\limits_{i,j}(M_i\|\boldsymbol{\delta}_{ij}(0)\|_i)$. 对任意矩阵范数 $\|\cdot\|$ 有关系[44]

$$\alpha\|\boldsymbol{\delta}_{ij}(t)\|_i\leqslant\|\boldsymbol{\delta}_{ij}(t)\|\leqslant\beta\|\boldsymbol{\delta}_{ij}(t)\|_i \tag{6.6.25}$$

其中 $\alpha>0$，$\beta>0$. 于是有(6.6.16)成立，其中取 $a=c\beta>0$. 证毕. □

【定理 6.6.3】 定常系统(6.6.1)和(6.6.2)在假设 1 ~ 5 下，第 i 传感器子系统有局部最优 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t+1|t)=\boldsymbol{\Psi}_{pi}\hat{\boldsymbol{x}}_i(t|t-1)+\boldsymbol{B}\boldsymbol{u}(t)+\boldsymbol{K}_{pi}\boldsymbol{y}_i(t) \tag{6.6.26}$$

$$\boldsymbol{\Psi}_{pi}=\bar{\boldsymbol{\Phi}}_i-\bar{\boldsymbol{K}}_{pi}\boldsymbol{H}_i \tag{6.6.27}$$

$$\bar{\boldsymbol{\Phi}}_i=\boldsymbol{\Phi}-\boldsymbol{J}_i\boldsymbol{H}_i,\ \boldsymbol{J}_i=\boldsymbol{\Gamma}\boldsymbol{S}_i\boldsymbol{R}_i^{-1} \tag{6.6.28}$$

$$\begin{aligned}&\bar{\boldsymbol{K}}_{pi}=\bar{\boldsymbol{\Phi}}_i\boldsymbol{K}_{fi},\quad \boldsymbol{K}_{fi}=\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1},\\ &\boldsymbol{Q}_{\varepsilon i}=\boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{R}_i,\quad \boldsymbol{K}_{pi}=\bar{\boldsymbol{K}}_{pi}+\boldsymbol{J}_i\end{aligned} \tag{6.6.29}$$

其中稳态预报误差方差阵 $\boldsymbol{\Sigma}_i$ 满足 Riccati 方程

$$\begin{aligned}\boldsymbol{\Sigma}_i=&\bar{\boldsymbol{\Phi}}_i[\boldsymbol{\Sigma}_i-\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}_i\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}_i)^{-1}\boldsymbol{H}_i\boldsymbol{\Sigma}_i]\bar{\boldsymbol{\Phi}}_i+\\ &\boldsymbol{\Gamma}(\boldsymbol{Q}-\boldsymbol{S}_i\boldsymbol{R}_i^{-1}\boldsymbol{S}_i^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}}\end{aligned} \tag{6.6.30}$$

带初值 $\boldsymbol{\Sigma}_i(0)=\boldsymbol{P}_0$，可用迭代法求解.

稳态预报误差互协方差阵 $\boldsymbol{\Sigma}_{ij}$ 满足 Lyapunov 方程

$$\boldsymbol{\Sigma}_{ij}=\boldsymbol{\Psi}_{pi}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}}+\boldsymbol{\Delta}_{pij} \tag{6.6.31}$$

其中 $i,j=1,\cdots,L$，且

$$\boldsymbol{\Delta}_{pij}=[\boldsymbol{I}_n,\ -\bar{\boldsymbol{K}}_{pi}]\begin{bmatrix}\boldsymbol{S}_{ij}^{(1)} & \boldsymbol{S}_{ij}^{(2)}\\ \boldsymbol{S}_{ij}^{(3)} & \boldsymbol{S}_{ij}^{(4)}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_n\\ -\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}}\end{bmatrix} \tag{6.6.32}$$

(6.6.32)可用迭代法(6.6.13)求解，其中初值取为 $\boldsymbol{\Sigma}_{ij}(0)=\boldsymbol{P}_0$ 或任意取初值. 在(6.6.32)中定义

$$\boldsymbol{S}_{ij}^{(1)}=\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{J}_i\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{\Gamma}\boldsymbol{S}_j\boldsymbol{J}_j^{\mathrm{T}}+\boldsymbol{J}_i\boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}},$$

$$
\begin{aligned}
\boldsymbol{S}_{ij}^{(2)} &= \boldsymbol{\Gamma}\boldsymbol{S}_j - \boldsymbol{J}_i\boldsymbol{R}_{ij}, \\
\boldsymbol{S}_{ij}^{(3)} &= \boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}}, \\
\boldsymbol{S}_{ij}^{(4)} &= \boldsymbol{R}_{ij}
\end{aligned}
\tag{6.6.33}
$$

证明 在定理 6.2.2 中令 $t_0 \to \infty$ 或 $t \to \infty$，则引出定理 6.6.3. □

【定理 6.6.4】 定常系统(6.6.1)和(6.6.2)在假设 1 ~ 5 下，第 i 传感器子系统有局部稳态最优 Kalman 滤波器

$$\hat{\boldsymbol{x}}_i(t+1 \mid t+1) = \hat{\boldsymbol{x}}_i(t+1 \mid t) + \boldsymbol{K}_{fi}\boldsymbol{\varepsilon}(t+1) \tag{6.6.34}$$

$$\hat{\boldsymbol{x}}_i(t+1 \mid t) = \overline{\boldsymbol{\Phi}}_i\hat{\boldsymbol{x}}_i(t \mid t) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{J}_i\boldsymbol{y}_i(t) \tag{6.6.35}$$

$$\boldsymbol{\varepsilon}_i(t+1) = \boldsymbol{y}_i(t+1) - \boldsymbol{H}_i\hat{\boldsymbol{x}}_i(t+1 \mid t) \tag{6.6.36}$$

$$
\begin{aligned}
\boldsymbol{K}_{fi} &= \boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1}, \\
\boldsymbol{Q}_{\varepsilon i} &= \boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{R}_i, \\
\boldsymbol{P}_i &= [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i]\boldsymbol{\Sigma}_i
\end{aligned}
\tag{6.6.37}
$$

其中 $\boldsymbol{\Sigma}_i$ 由 Riccati 方程(6.6.9)或(6.6.31)计算. 初值 $\hat{\boldsymbol{x}}(0 \mid -1)$ 可任意选取. $\boldsymbol{P}_i$ 为稳态局部滤波误差方差阵. 稳态局部滤波误差互协方差阵 $\boldsymbol{P}_{ij} = \lim \boldsymbol{P}_{ij}(t \mid t)(t \to \infty)$ 满足 Lyapunov 方程

$$\boldsymbol{P}_{ij} = \boldsymbol{\Psi}_{fi}\boldsymbol{P}_{ij}\boldsymbol{\Psi}_{fj}^{\mathrm{T}} + \boldsymbol{\Delta}_{fij} \tag{6.6.38}$$

其中 $i,j = 1,\cdots,L,\ i \neq j$，且定义

$$
\begin{aligned}
\boldsymbol{\Delta}_{fij} = {} & \boldsymbol{\Psi}_{fi}\boldsymbol{K}_{fi}[\boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}} - \boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}][\boldsymbol{I}_n - \boldsymbol{K}_{fj}\boldsymbol{H}_j]^{\mathrm{T}} + \\
& [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i][\boldsymbol{J}_i\boldsymbol{R}_{ij} - \boldsymbol{\Gamma}\boldsymbol{S}_j]\boldsymbol{K}_{fj}^{\mathrm{T}}\boldsymbol{\Psi}_{fj}^{\mathrm{T}} + \\
& [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i][\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{J}_i\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{\Gamma}\boldsymbol{S}_j\boldsymbol{J}_j^{\mathrm{T}} + \boldsymbol{J}_i\boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}}] \times \\
& [\boldsymbol{I}_n - \boldsymbol{K}_{fj}\boldsymbol{H}_j]^{\mathrm{T}} + \boldsymbol{K}_{fi}\boldsymbol{R}_{ij}\boldsymbol{K}_{fj}^{\mathrm{T}}
\end{aligned}
\tag{6.6.39}
$$

其中定义

$$\boldsymbol{\Psi}_{fi} = [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i]\overline{\boldsymbol{\Phi}}_i \tag{6.6.40}$$

定义 $\boldsymbol{P}_i = \boldsymbol{P}_{ii}$，在(6.6.38)和(6.6.39)中置 $i = j$ 有 $\boldsymbol{P}_i$ 满足 Lyapunov 方程

$$\boldsymbol{P}_i = \boldsymbol{\Psi}_{fi}\boldsymbol{P}_i\boldsymbol{\Psi}_{fi}^{\mathrm{T}} + \boldsymbol{\Delta}_{fii} \tag{6.6.41}$$

$$\boldsymbol{\Delta}_{fii} = [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i]\boldsymbol{\Gamma}(\boldsymbol{Q} - \boldsymbol{S}_i\boldsymbol{R}_i^{-1}\boldsymbol{S}_i^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}}[\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i]^{\mathrm{T}} + \boldsymbol{K}_{fi}\boldsymbol{R}_i\boldsymbol{K}_{fi}^{\mathrm{T}} \tag{6.6.42}$$

证明 当 $t \to \infty$ 时由定理 6.2.4 得证. □

【注 6.6.1】 稳态 Kalman 滤波器(6.6.34) ~ (6.6.36)也可表为另一种形式

$$
\begin{aligned}
\hat{\boldsymbol{x}}_i(t+1 \mid t+1) = {} & \boldsymbol{\Psi}_{fi}\hat{\boldsymbol{x}}_i(t \mid t) + [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i]\boldsymbol{B}\boldsymbol{u}(t) + \\
& [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i]\boldsymbol{J}_i\boldsymbol{y}_i(t) + \boldsymbol{K}_{fi}\boldsymbol{y}_i(t+1)
\end{aligned}
\tag{6.6.43}
$$

【推论 6.6.1】 定常系统(6.6.1)和(6.6.2)在假设 1 ~ 4 下，若 $\boldsymbol{S}_i = \boldsymbol{0}$，$\boldsymbol{R}_{ij} = \boldsymbol{0}(i \neq j)$，$\boldsymbol{u}(t) = \boldsymbol{0}$，则第 i 传感器子系统有局部稳态最优 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t+1 \mid t) = \boldsymbol{\Psi}_{pi}\hat{\boldsymbol{x}}_i(t \mid t-1) + \boldsymbol{K}_{pi}\boldsymbol{y}_i(t) \tag{6.6.44}$$

$$\boldsymbol{\Psi}_{pi} = \boldsymbol{\Phi} - \boldsymbol{K}_{pi}\boldsymbol{H}_i \tag{6.6.45}$$

$$\boldsymbol{K}_{pi} = \boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{6.6.46}$$

$$\boldsymbol{Q}_{\varepsilon i} = \boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{R}_i \tag{6.6.47}$$

其中 $\boldsymbol{\Sigma}_i$ 满足 Riccati 方程

$$\boldsymbol{\Sigma}_i = \boldsymbol{\Phi}[\boldsymbol{\Sigma}_i - \boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}(\boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{R}_i)^{-1}\boldsymbol{H}_i\boldsymbol{\Sigma}_i]\boldsymbol{\Phi}^{\mathrm{T}} + \boldsymbol{\Gamma Q \Gamma}^{\mathrm{T}} \tag{6.6.48}$$

而稳态预报误差互协方差阵 $\boldsymbol{\Sigma}_{ij}$ 满足 Lyapunov 方程

$$\boldsymbol{\Sigma}_{ij} = \boldsymbol{\Psi}_{pi}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}} + \boldsymbol{\Gamma Q \Gamma}^{\mathrm{T}} \tag{6.6.49}$$

其中 $i,j = 1,\cdots,L$.

【推论 6.6.2】 定常系统(6.6.1)和(6.6.2)在假设 1 ~ 4 下,若 $\boldsymbol{S}_i = \boldsymbol{0}$,$\boldsymbol{R}_{ij} = \boldsymbol{0}(i \neq j)$,$\boldsymbol{u}(t) = \boldsymbol{0}$,则第 i 传感器子系统有局部稳态最优 Kalman 滤波器

$$\hat{\boldsymbol{x}}_i(t+1 \mid t+1) = \boldsymbol{\Psi}_{fi}\hat{\boldsymbol{x}}_i(t \mid t) + \boldsymbol{K}_{fi}\boldsymbol{y}(t+1) \tag{6.6.50}$$

$$\boldsymbol{\Psi}_{fi} = [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i]\boldsymbol{\Phi} \tag{6.6.51}$$

$$\boldsymbol{K}_{fi} = \boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1}, \quad \boldsymbol{Q}_{\varepsilon i} = \boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{R}_i \tag{6.6.52}$$

$$\boldsymbol{P}_i = [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i]\boldsymbol{\Sigma}_i \tag{6.6.53}$$

其中 $\boldsymbol{P}_i$ 为局部稳态滤波误差方差阵,$\boldsymbol{\Sigma}_i$ 由(6.6.48)计算.局部稳态滤波误差互协方差阵 $\boldsymbol{P}_{ij}(i \neq j)$ 满足 Lyapunov 方程

$$\boldsymbol{P}_{ij} = \boldsymbol{\Psi}_{fi}\boldsymbol{P}_{ij}\boldsymbol{\Psi}_{fj}^{\mathrm{T}} + \boldsymbol{\Delta}_{fij} \tag{6.6.54}$$

其中 $i,j = 1,\cdots,L$,且

$$\boldsymbol{\Delta}_{fij} = [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}_i]\boldsymbol{\Gamma Q \Gamma}^{\mathrm{T}}[\boldsymbol{I}_n - \boldsymbol{K}_{fj}\boldsymbol{H}_j]^{\mathrm{T}} \tag{6.6.55}$$

【定理 6.6.5】 定常系统(6.6.1)和(6.6.2)在假设 1 ~ 5 下,局部滤波误差互协方差阵 $\boldsymbol{P}_{ij}$ 与局部预报误差互协方差阵 $\boldsymbol{\Sigma}_{ij}$ 有关系

$$\boldsymbol{P}_{ij} = [\boldsymbol{I}_n - \boldsymbol{K}_{fi}\boldsymbol{H}]\boldsymbol{\Sigma}_{ij}[\boldsymbol{I}_n - \boldsymbol{K}_{fj}\boldsymbol{H}]^{\mathrm{T}} + \boldsymbol{K}_{fi}\boldsymbol{R}_{ij}\boldsymbol{K}_{fj}^{\mathrm{T}} \tag{6.6.56}$$

证明 由定理 6.2.5 得证. □

【定理 6.6.6】 定常系统(6.6.1)和(6.6.2)在假设 1 ~ 4 下,若 $\boldsymbol{u}(t) = \boldsymbol{0}$,则第 i 传感器子系统有超前$(-N)$步局部稳态最优 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t \mid t+N) = \boldsymbol{\Phi}^{-N-1}\hat{\boldsymbol{x}}_i(t+N+1 \mid t+N), \quad N \leqslant -2 \tag{6.6.57}$$

其中稳态预报误差 $\tilde{\boldsymbol{x}}_i(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}_i(t \mid t+N)$ 方差阵 $\boldsymbol{P}_i(N) = \mathrm{E}[\tilde{\boldsymbol{x}}_i(t \mid t+N)\tilde{\boldsymbol{x}}_i^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}_i(N) = \boldsymbol{\Phi}^{-N-1}\boldsymbol{\Sigma}_i(\boldsymbol{\Phi}^{-N-1})^{\mathrm{T}} + \sum_{j=0}^{-N-2}\boldsymbol{\Phi}^j\boldsymbol{\Gamma Q \Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^j)^{\mathrm{T}}, \quad N \leqslant -2 \tag{6.6.58}$$

其中 $i = 1,\cdots,L$,而稳态预报互协方差阵 $\boldsymbol{P}_{ij}(N) = \mathrm{E}[\tilde{\boldsymbol{x}}_i(t \mid t+N)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}_{ij}(N) = \boldsymbol{\Phi}^{-N-1}\boldsymbol{\Sigma}_{ij}(\boldsymbol{\Phi}^{-N-1})^{\mathrm{T}} + \sum_{j=0}^{-N-2}\boldsymbol{\Phi}^j\boldsymbol{\Gamma Q \Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^j)^{\mathrm{T}}, \quad N \leqslant -2 \tag{6.6.59}$$

其中 $i,j = 1,\cdots,L, i \neq j$,$\boldsymbol{\Sigma}_i$ 和 $\boldsymbol{\Sigma}_{ij}$ 由(6.6.9)和(6.6.10)计算,且 $\hat{\boldsymbol{x}}_i(t+N+1 \mid t+N)$ 由稳态 Kalman 预报器递推计算为

$$\hat{\boldsymbol{x}}_i(t+N+1 \mid t+N) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}_i(t+N \mid t+N-1) + \boldsymbol{K}_{pi}\boldsymbol{y}_i(t+N) \tag{6.6.60}$$

初始时刻取 $t_0 = -N$,且任取 $\hat{\boldsymbol{x}}_i(0 \mid -1)$.

证明 注意 $\boldsymbol{\Phi}(t) = \boldsymbol{\Phi}$,$\boldsymbol{\Gamma}(t) = \boldsymbol{\Gamma}$,$\boldsymbol{Q}(t) = \boldsymbol{Q}$,由定理 6.3.2 立刻得证. □

【定理 6.6.6】 定常系统(6.6.1)和(6.6.2)在假设 1 ~ 4 下,若 $\boldsymbol{u}(t) = \boldsymbol{0}$,则第 i 传感器子系统有超前 N 步局部稳态最优 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t+N \mid t) = \boldsymbol{\Phi}^{N-1}\hat{\boldsymbol{x}}_i(t+1 \mid t), \quad N \geqslant 2 \tag{6.6.61}$$

相应的稳态预报误差 $\tilde{\boldsymbol{x}}_i(t+N \mid t) = \boldsymbol{x}(t+N) - \hat{\boldsymbol{x}}_i(t+N \mid t)$ 方差阵 $\boldsymbol{\Sigma}_i(N) =$

$\mathrm{E}[\tilde{\boldsymbol{x}}_i(t+N\mid t)\tilde{\boldsymbol{x}}_i^{\mathrm{T}}(t+N\mid t)]$为

$$\boldsymbol{\Sigma}_i(N)=\boldsymbol{\Phi}^{N-1}\boldsymbol{\Sigma}_i(\boldsymbol{\Phi}^{N-1})^{\mathrm{T}}+\sum_{k=0}^{N-2}\boldsymbol{\Phi}^k\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^k)^{\mathrm{T}}\tag{6.6.62}$$

稳态预报误差互协方差阵 $\boldsymbol{\Sigma}_{ij}(N)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t+N\mid t)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t+N\mid t)]$为

$$\boldsymbol{\Sigma}_{ij}(N)=\boldsymbol{\Phi}^{N-1}\boldsymbol{\Sigma}_{ij}(\boldsymbol{\Phi}^{N-1})^{\mathrm{T}}+\sum_{k=0}^{N-2}\boldsymbol{\Phi}^k\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^k)^{\mathrm{T}}\tag{6.6.63}$$

证明　由定理6.3.1有

$$\boldsymbol{\Sigma}_i(N)=\boldsymbol{\Phi}^{N-1}\boldsymbol{\Sigma}_i(N)(\boldsymbol{\Phi}^{N-1})^{\mathrm{T}}+\sum_{j=2}^{N}\boldsymbol{\Phi}^{N-j}\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}(\boldsymbol{\Phi}^{N-j})^{\mathrm{T}},\quad N\geqslant 2\tag{6.6.64}$$

稳态预报误差互协方差阵 $\boldsymbol{\Sigma}_{ij}(N)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t+N\mid t)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t+N\mid t)]$为

$$\boldsymbol{\Sigma}_{ij}(N)=\boldsymbol{\Phi}^{N-1}\boldsymbol{\Sigma}_{ij}(N)(\boldsymbol{\Phi}^{N-1})^{\mathrm{T}}+\sum_{j=2}^{N}\boldsymbol{\Phi}^{N-j}\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}(\boldsymbol{\Phi}^{N-j})^{\mathrm{T}},\quad N\geqslant 2\tag{6.6.65}$$

其中 $i\neq j,i,j=1,\cdots,L$.引入变换 $N-j=k$ 上两式化为(6.6.62)和(6.6.63). □

注意(6.6.58)和(6.6.59)与(6.6.62)和(6.6.63)具有相同的形式.

【定理6.6.7】　定常系统(6.6.1)和(6.6.2)在假设1～5下,第 i 传感器子系统有局部稳态最优Kalman平滑器 $\hat{\boldsymbol{x}}_i(t\mid t+N)(N>0)$为

$$\hat{\boldsymbol{x}}_i(t\mid t+N)=\hat{\boldsymbol{x}}_i(t\mid t-1)+\sum_{k=0}^{N}\boldsymbol{K}_i(k)\boldsymbol{\varepsilon}_i(t+k),\quad i=1,\cdots,L\tag{6.6.66}$$

其中局部稳态Kalman预报器 $\hat{\boldsymbol{x}}_i(t\mid t-1)$由定理6.6.1计算.$\boldsymbol{\varepsilon}_i(t)=\boldsymbol{y}_i(t)-\boldsymbol{H}_i\hat{\boldsymbol{x}}_i(t\mid t-1)$,稳态平滑增益 $\boldsymbol{K}_i(k)$为

$$\boldsymbol{K}_i(k)=\boldsymbol{\Sigma}_i\boldsymbol{\Psi}_{pi}^{\mathrm{T}k}\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1},\quad k=0,\cdots,N\tag{6.6.67}$$

其中定义 $\boldsymbol{\Psi}_{pi}^{\mathrm{T}k}=(\boldsymbol{\Psi}_{pi}^{\mathrm{T}})^k$,且有

$$\begin{aligned}\boldsymbol{\Psi}_{pi}&=\boldsymbol{\Phi}-\boldsymbol{K}_{pi}\boldsymbol{H}_i,\\ \boldsymbol{K}_{pi}&=(\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S}_i)\boldsymbol{Q}_{\varepsilon i}^{-1},\\ \boldsymbol{Q}_{\varepsilon i}&=\boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{R}_i\end{aligned}\tag{6.6.68}$$

$\boldsymbol{\Sigma}_i$ 满足Riccati方程

$$\boldsymbol{\Sigma}_i=\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{\Phi}^{\mathrm{T}}-(\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S}_i)(\boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{R}_i)^{-1}(\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{S}_i)^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\tag{6.6.69}$$

稳态平滑误差方差阵 $\boldsymbol{P}_i(N)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t+N)\tilde{\boldsymbol{x}}_i^{\mathrm{T}}(t\mid t+N)]$为

$$\boldsymbol{P}_i(N)=\boldsymbol{\Sigma}_i-\sum_{k=0}^{N}\boldsymbol{K}_i(k)\boldsymbol{Q}_{\varepsilon i}\boldsymbol{K}_i^{\mathrm{T}}(k)\tag{6.6.70}$$

稳态平滑误差互协方差阵 $\boldsymbol{P}_{ij}(N)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t+N)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t\mid t+N)](i\neq j)$为

$$\begin{aligned}\boldsymbol{P}_{ij}(N)=\boldsymbol{\Sigma}_{ij}&-\sum_{r=0}^{N}\boldsymbol{K}_i(r)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij}-\sum_{s=0}^{N}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}s}\boldsymbol{H}_j^{\mathrm{T}}\boldsymbol{K}_j^{\mathrm{T}}(s)+\\ &\sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{K}_i(r)\boldsymbol{E}_{ij}(r,s)\boldsymbol{K}_j^{\mathrm{T}}(s)\end{aligned}\tag{6.6.71}$$

其中 $\boldsymbol{\Sigma}_{ij}$(6.6.10)计算,$\boldsymbol{E}_{ij}(r,s)=\mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)]$有公式:当 $\min(r,s)>0$ 时有

$$\boldsymbol{E}_{ij}(r,s)=\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}s}\boldsymbol{H}_j^{\mathrm{T}}+\sum_{k=1}^{\min(r,s)}\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r-k}[\boldsymbol{\Gamma},-\boldsymbol{K}_{pi}]\times$$

$$\begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S}_j \\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_{ij} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Gamma}^{\mathrm{T}} \\ -\boldsymbol{K}_{pj}^{\mathrm{T}} \end{bmatrix} \boldsymbol{\Psi}_{pj}^{\mathrm{T}(s-k)} \boldsymbol{H}_j^{\mathrm{T}} + \boldsymbol{R}_{ij}\delta_{rs} \tag{6.6.72}$$

若 $\min(r,s)=0$,则有

$$\boldsymbol{E}_{ij}(0,0) = \boldsymbol{H}_i\boldsymbol{\Sigma}_{ij}\boldsymbol{H}_j^{\mathrm{T}} + \boldsymbol{R}_{ij} \tag{6.6.73}$$

$$\boldsymbol{E}_{ij}(r,0) = \boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij}\boldsymbol{H}_j^{\mathrm{T}} + \boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r-1}[\boldsymbol{\Gamma}\boldsymbol{S}_j - \boldsymbol{K}_{pi}\boldsymbol{R}_{ij}] \tag{6.6.74}$$

$$\boldsymbol{E}_{ij}(0,s) = \boldsymbol{H}_i\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}s}\boldsymbol{H}_j^{\mathrm{T}} + [\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{R}_{ij}\boldsymbol{K}_{pj}^{\mathrm{T}}]\boldsymbol{\Psi}_{pj}^{\mathrm{T}(s-1)}\boldsymbol{H}_j^{\mathrm{T}} \tag{6.6.75}$$

证明 由引理 6.4.1 令 $t\to\infty$ 或 $t_0\to\infty$ 有(6.6.64),(6.6.65)和(6.6.70).由定理 6.4.2 令 $t\to\infty$ 或 $t_0\to\infty$ 引出(6.6.71)~(6.6.75). □

【定理 6.6.8】 定常系统(6.6.1)和(6.6.2)在假设 1 ~ 5 下,第 i 传感器子系统有局部稳态最优 Kalman 平滑器

$$\hat{\boldsymbol{x}}_i(t\mid t+N) = \hat{\boldsymbol{x}}_i(t\mid t-1) + \sum_{k=0}^{N}\boldsymbol{K}_i(k)\boldsymbol{\varepsilon}_i(t+k), \quad i=1,\cdots,L \tag{6.6.76}$$

其中局部稳态 Kalman 预报器 $\hat{\boldsymbol{x}}_i(t\mid t-1)$ 由定理 6.6.3 计算.且新息 $\boldsymbol{\varepsilon}_i(t) = \boldsymbol{y}_i(t) - \boldsymbol{H}_i\hat{\boldsymbol{x}}_i(t\mid t-1)$.稳态平滑增益为

$$\boldsymbol{K}_i(k) = \boldsymbol{\Sigma}_i\boldsymbol{\Psi}_{pi}^{\mathrm{T}k}\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{6.6.77}$$

$$\boldsymbol{\Psi}_{pi} = \bar{\boldsymbol{\Phi}}_i - \bar{\boldsymbol{K}}_{pi}\boldsymbol{H}_i \tag{6.6.78}$$

$$\bar{\boldsymbol{\Phi}}_i = \boldsymbol{\Phi} - \boldsymbol{J}_i\boldsymbol{H}_i, \quad \boldsymbol{J}_i = \boldsymbol{\Gamma}\boldsymbol{S}_i\boldsymbol{R}_i^{-1} \tag{6.6.79}$$

$$\boldsymbol{K}_{pi} = \bar{\boldsymbol{K}}_{pi} + \boldsymbol{J}_i \tag{6.6.80}$$

$$\bar{\boldsymbol{K}}_{pi} = \bar{\boldsymbol{\Phi}}_i\boldsymbol{K}_{fi}, \quad \boldsymbol{K}_{fi} = \boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{6.6.81}$$

$$\boldsymbol{Q}_{\varepsilon i} = \boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{R}_i \tag{6.6.82}$$

$\boldsymbol{\Sigma}_i$ 满足 Riccati 方程

$$\boldsymbol{\Sigma}_i = \bar{\boldsymbol{\Phi}}_i[\boldsymbol{\Sigma}_i - \boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}}(\boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{R}_i)^{-1}\boldsymbol{H}_i\boldsymbol{\Sigma}_i]\bar{\boldsymbol{\Phi}}_i^{\mathrm{T}} + \boldsymbol{\Gamma}(\boldsymbol{Q} - \boldsymbol{S}_i\boldsymbol{R}_i^{-1}\boldsymbol{S}_i^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}} \tag{6.6.83}$$

稳态局部平滑误差方差阵和互协方差阵各为

$$\boldsymbol{P}_i(N) = \boldsymbol{\Sigma}_i - \sum_{k=0}^{N}\boldsymbol{K}_i(k)\boldsymbol{Q}_{\varepsilon i}\boldsymbol{K}_i^{\mathrm{T}}(k) \tag{6.6.84}$$

$$\begin{aligned}\boldsymbol{P}_{ij}(N) = {} & \boldsymbol{\Sigma}_{ij} - \sum_{r=0}^{N}\boldsymbol{K}_i(r)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij} - \sum_{s=0}^{N}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}s}\boldsymbol{H}_j^{\mathrm{T}}\boldsymbol{K}_j^{\mathrm{T}}(s) + \\ & \sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{K}_i(r)\boldsymbol{E}_{ij}(r,s)\boldsymbol{K}_j^{\mathrm{T}}(s)\end{aligned} \tag{6.6.85}$$

其中 $\boldsymbol{\Sigma}_{ij}$(6.6.31)计算.当 $\min(r,s)>0$ 时有

$$\begin{aligned}\boldsymbol{E}_{ij}(r,s) = {} & \boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}s}\boldsymbol{H}_j^{\mathrm{T}} + \sum_{k=1}^{\min(r,s)}\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r-k}[\boldsymbol{I}_n, -\bar{\boldsymbol{K}}_{pi}] \times \\ & \begin{bmatrix} \boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{J}_i\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{\Gamma}\boldsymbol{S}_j\boldsymbol{J}_j^{\mathrm{T}} + \boldsymbol{J}_i\boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}}, & \boldsymbol{\Gamma}\boldsymbol{S}_j - \boldsymbol{J}_i\boldsymbol{R}_{ij} \\ \boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}} & \boldsymbol{R}_{ij} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_n \\ -\bar{\boldsymbol{K}}_{pj}^{\mathrm{T}} \end{bmatrix} \boldsymbol{\Psi}_{pj}^{\mathrm{T}(s-k)}\boldsymbol{H}_j^{\mathrm{T}} + \boldsymbol{R}_{ij}\delta_{rs}\end{aligned} \tag{6.6.86}$$

若 $\min(r,s)=0$,则有

$$\boldsymbol{E}_{ij}(0,0) = \boldsymbol{H}_i\boldsymbol{\Sigma}_{ij}\boldsymbol{H}_j^{\mathrm{T}} + \boldsymbol{R}_{ij} \tag{6.6.87}$$

$$\boldsymbol{E}_{ij}(r,0) = \boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij}\boldsymbol{H}_j^{\mathrm{T}} + \boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r-1}[\boldsymbol{\Gamma}\boldsymbol{S}_j - \boldsymbol{J}_i\boldsymbol{R}_{ij} - \bar{\boldsymbol{K}}_{pi}\boldsymbol{R}_{ij}] \tag{6.6.88}$$

$$\boldsymbol{E}_{ij}(0,s)=\boldsymbol{H}_i\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}s}\boldsymbol{H}_j^{\mathrm{T}}+[\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}}-\boldsymbol{R}_{ij}\bar{\boldsymbol{K}}_{Pj}^{\mathrm{T}}]\boldsymbol{\Psi}_{pj}^{\mathrm{T}(s-1)}\boldsymbol{H}_j^{\mathrm{T}} \tag{6.6.89}$$

证明　由引理6.4.1和定理6.4.1，令 $t\to\infty$ 得证.　□

【定理6.6.9】　由定理6.6.7和定理6.6.8给出的局部平滑增益 $\boldsymbol{K}_i(k)$ 和互协方差阵 $\boldsymbol{P}_{ij}(N)$ 的两种算法是等价(等值)的.

证明　由定理3.7.6引出两等 $\boldsymbol{K}_i(k)$ 算法是等价的. 由定理6.4.3的证明，令 $t\to\infty$ 立刻得两种 $\boldsymbol{P}_{ij}(N)$ 算法等价.　□

【推论6.6.3】　定常系统(6.6.1)和(6.6.2)在假设1～5下，若 $\boldsymbol{S}_i=\boldsymbol{0}$，$\boldsymbol{R}_{ij}=\boldsymbol{0}$，则局部稳态 Kalman 平滑器及平滑误差方差阵和互协方差阵为在定理6.6.7中有关公式中置 $\boldsymbol{S}_i=\boldsymbol{0}$，$\boldsymbol{S}_j=\boldsymbol{0}$，$\boldsymbol{R}_{ij}=\boldsymbol{0}$.

【定理6.6.10】　定常系统(6.6.1)和(6.6.2)在假设1～5下，第 i 传感器子系统有局部稳态 Kalman 平滑器(6.6.64)～(6.6.69)，平滑误差方差阵为(6.6.70)，而稳态平滑误差互协方差阵 $\boldsymbol{P}_{ij}(N)$ 有如下新公式：

$$\boldsymbol{P}_{ij}(N)=\boldsymbol{\Psi}_{iN}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{iN}^{\mathrm{T}}+\sum_{\rho=0}^{N}[\boldsymbol{K}_{i\rho}^{w},\boldsymbol{K}_{i\rho}^{v}]\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}_j\\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{K}_{j\rho}^{w\mathrm{T}}\\ \boldsymbol{K}_{j\rho}^{v\mathrm{T}}\end{bmatrix} \tag{6.6.90}$$

其中 $i,j=1,\cdots,L,i\neq j$，定义

$$\boldsymbol{\Psi}_{iN}=\boldsymbol{I}_n-\sum_{k=0}^{N}\boldsymbol{K}_i(k)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{k} \tag{6.6.91}$$

$$\boldsymbol{K}_{i\rho}^{w}=-\sum_{k=\rho+1}^{N}\boldsymbol{K}_i(k)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{k-\rho-1}\boldsymbol{\Gamma},\quad \rho=0,\cdots,N-1;\boldsymbol{K}_{iN}^{w}(t)=\boldsymbol{0}$$

$$\boldsymbol{K}_{i\rho}^{v}=\sum_{k=\rho+1}^{N}\boldsymbol{K}_i(k)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{k-\rho-1}\boldsymbol{K}_{pi}-\boldsymbol{K}_i(\rho),\quad \rho=0,\cdots,N-1;\boldsymbol{K}_{iN}^{v}=-\boldsymbol{K}_i(N) \tag{6.6.92}$$

证明　由定理6.4.4的稳态形式得证.　□

【定理6.6.11】　定常系统(6.6.1)和(6.6.2)在假设1～5下，第 i 传感器有局部稳态 Kalman 平滑器(6.6.76)～(6.6.83)，且平滑误差方差阵为(6.6.84)，而稳态平滑误差互协方差阵为

$$\boldsymbol{P}_{ij}(N)=\boldsymbol{\Psi}_{iN}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{jN}^{\mathrm{T}}+\sum_{\rho=0}^{N}[\boldsymbol{M}_{i\rho}^{\bar{w}},\boldsymbol{M}_{i\rho}^{v}]\begin{bmatrix}\boldsymbol{S}_{ij}^{11} & \boldsymbol{S}_{ij}^{12}\\ \boldsymbol{S}_{ij}^{21} & \boldsymbol{S}_{ij}^{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{M}_{j\rho}^{\bar{w}\mathrm{T}}\\ \boldsymbol{M}_{j\rho}^{v\mathrm{T}}\end{bmatrix} \tag{6.6.93}$$

其中 $i,j=1,\cdots,L,i\neq j$，且定义

$$\boldsymbol{\Psi}_{iN}=\boldsymbol{I}_n-\sum_{k=0}^{N}\boldsymbol{K}_i(k)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{k} \tag{6.6.94}$$

$$\boldsymbol{M}_{i\rho}^{\bar{w}}=-\sum_{k=\rho+1}^{N}\boldsymbol{K}_i(k)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{k-\rho-1},\quad \rho=0,\cdots,N-1,\boldsymbol{M}_{iN}^{\bar{w}}(t)=\boldsymbol{0}$$

$$\boldsymbol{M}_{i\rho}^{v}(t)=\sum_{k=\rho+1}^{N}\boldsymbol{K}_i(k)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{k-\rho-1}\bar{\boldsymbol{K}}_{pi}-\boldsymbol{K}_i(\rho),\quad \rho=0,\cdots,N-1,\boldsymbol{M}_{iN}^{v}=-\boldsymbol{K}_i(N) \tag{6.6.95}$$

$$\boldsymbol{S}_{ij}^{11}=\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{J}_i\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{\Gamma}\boldsymbol{S}_j\boldsymbol{J}_j^{\mathrm{T}}+\boldsymbol{J}_i\boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}},$$

$$\boldsymbol{S}_{ij}^{12}=\boldsymbol{\Gamma}\boldsymbol{S}_j-\boldsymbol{J}_i\boldsymbol{R}_{ij},$$

$$S_{ij}^{21} = S_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - R_{ij}J_j^{\mathrm{T}},$$

$$S_{ij}^{22} = R_{ij} \tag{6.6.96}$$

证明 由定理 6.4.5 和公式(6.4.9)令 $t \to \infty$ 引出(6.6.93) ~ (6.6.96). □

【定理 6.6.12】 多传感器定常系统(6.6.1)和(6.6.2)在假设 1 ~ 5 下,第 i 传感器子系统有局部稳态最优白噪声反卷积估值器

$$\hat{\boldsymbol{w}}_i(t \mid t+N) = \sum_{k=0}^{N} \boldsymbol{M}_i(k)\boldsymbol{\varepsilon}_i(t+k), \quad N \geqslant 0, i = 1,\cdots,L \tag{6.6.97}$$

$$\hat{\boldsymbol{w}}_i(t \mid t+N) = \boldsymbol{0}, \quad N < 0 \tag{6.6.98}$$

其中定义

$$\boldsymbol{M}_i(0) = \boldsymbol{S}_i\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{6.6.99}$$

$$\boldsymbol{M}_i(k) = \boldsymbol{D}_i\boldsymbol{\Psi}_{pi}^{\mathrm{T}(k-1)}\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1}, \quad k \geqslant 1 \tag{6.6.100}$$

其中规定 $\boldsymbol{\Psi}_{pi}^{\mathrm{T}(k-1)} = (\boldsymbol{\Psi}_{pi}^{\mathrm{T}})^{k-1}$,且定义

$$\boldsymbol{D}_i = \boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{S}_i\boldsymbol{K}_{pi}^{\mathrm{T}} \tag{6.6.101}$$

或等价地定义

$$\boldsymbol{D}_i = -\boldsymbol{S}_i\boldsymbol{K}_{fi}^{\mathrm{T}}\overline{\boldsymbol{\Phi}}_i^{\mathrm{T}} + \boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{S}_i\boldsymbol{J}_i^{\mathrm{T}} \tag{6.6.102}$$

稳态估值误差方差阵 $\boldsymbol{P}_i^w(N) = \mathrm{E}[\tilde{\boldsymbol{w}}_i(t \mid t+N)\tilde{\boldsymbol{w}}_i^{\mathrm{T}}(t \mid t+N)]$($\tilde{\boldsymbol{w}}_i(t \mid t+N) = \boldsymbol{w}(t) - \hat{\boldsymbol{w}}_i(t \mid t+N)$)为

$$\boldsymbol{P}_i^w(N) = \boldsymbol{Q} - \sum_{k=0}^{N} \boldsymbol{M}_i(k)\boldsymbol{Q}_{\varepsilon i}\boldsymbol{M}_i^{\mathrm{T}}(k), \quad N \geqslant 0 \tag{6.6.103}$$

$$\boldsymbol{P}_i^w(N) = \boldsymbol{Q}, \quad N < 0 \tag{6.6.104}$$

稳态局部估值误差互协方差阵 $\boldsymbol{P}_{ij}^w(N) = \mathrm{E}[\tilde{\boldsymbol{w}}_i(t \mid t+N)\tilde{\boldsymbol{w}}_j^{\mathrm{T}}(t \mid t+N)]$($i \neq j$)为

$$\begin{aligned}\boldsymbol{P}_{ij}^w(N) = \boldsymbol{Q} - \sum_{r=0}^{N} \boldsymbol{M}_i(r)\boldsymbol{Q}_{\varepsilon i}\boldsymbol{M}_i^{\mathrm{T}}(r) - \sum_{s=0}^{N} \boldsymbol{M}_j(s)\boldsymbol{Q}_{\varepsilon j}\boldsymbol{M}_j^{\mathrm{T}}(s) + \\ \sum_{r=0}^{N}\sum_{s=0}^{N} \boldsymbol{M}_i(r)\boldsymbol{E}_{ij}(r,s)\boldsymbol{M}_j^{\mathrm{T}}(s)\end{aligned} \tag{6.6.105}$$

其中 $i,j = 1,\cdots,L, i \neq j, N \geqslant 0$,$\boldsymbol{E}_{ij}(r,s)$ 由(6.6.72) ~ (6.6.75)计算. 当 $N < 0$ 时有

$$\boldsymbol{P}_{ij}^w(N) = \boldsymbol{Q}, \quad i,j = 1,\cdots,L, i \neq j, N < 0 \tag{6.6.106}$$

证明 在引理 6.5.1、定理 6.5.1 和定理 6.5.2 中令 $t \to \infty$ 或 $t_0 \to \infty$ 得证. □

统一记第 i 个传感器子系统 Kalman 估值器为 $\hat{\boldsymbol{x}}_i(t \mid t+N)$($N = 0, N > 0$ 或 $N < 0$),并记相应的估值误差互协方差阵为 $\boldsymbol{P}_{ij}(N)$,并定义

$$\boldsymbol{P}_i(N) = \boldsymbol{P}_{ii}(N), \quad \boldsymbol{P}_i(0) = \boldsymbol{P}_i, \quad \boldsymbol{P}_i(-1) = \boldsymbol{\Sigma}_i, \quad \boldsymbol{P}_{ij}(-1) = \boldsymbol{\Sigma}_{ij} \tag{6.6.107}$$

【定理 6.6.13】 多传感器系统(6.6.1)和(6.6.2)在假设 1 ~ 5 下,有按矩阵加权最优融合稳态 Kalman 估值器

$$\hat{\boldsymbol{x}}_0(t \mid t+N) = \sum_{i=1}^{L} \boldsymbol{A}_i\hat{\boldsymbol{x}}_i(t \mid t+N) \tag{6.6.108}$$

其中加权阵 $\boldsymbol{A}_i$ 由下式计算

$$[\boldsymbol{A}_i,\cdots,\boldsymbol{A}_L] = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(N)\boldsymbol{e})^{-1}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(N) \tag{6.6.109}$$

$$\boldsymbol{P}(N) = (\boldsymbol{P}_{ij}(N))_{nL \times nL} \tag{6.6.110}$$

其中 $\boldsymbol{e}^{\mathrm{T}}=[\boldsymbol{I}_n,\cdots,\boldsymbol{I}_n]$，$\boldsymbol{P}(N)$ 是以 $\boldsymbol{P}_{ij}(N)$ 为第 (i,j) 元素的 $nL\times nL$ 分块矩阵. 最优融合误差方差阵为

$$\boldsymbol{P}_0^m(N)=(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}^{-1}(N)\boldsymbol{e})^{-1} \tag{6.6.111}$$

按标量加权最优融合稳态 Kalman 估值器为

$$\hat{\boldsymbol{x}}_0(t\mid t+N)=\sum_{i=1}^{L}a_i\hat{\boldsymbol{x}}_i(t\mid t+N) \tag{6.6.112}$$

其中加权系数 a_i 由下式计算

$$[a_1,\cdots,a_L]=\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{tr}}^{-1}(N)(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{P}_{\mathrm{tr}}^{-1}(N)\boldsymbol{e})^{-1} \tag{6.6.113}$$

$$\boldsymbol{P}_{\mathrm{tr}}(N)=(\mathrm{tr}\boldsymbol{P}_{ij}(N))_{L\times L} \tag{6.6.114}$$

其中 $\boldsymbol{e}^{\mathrm{T}}=[1,\cdots,1]$，$\boldsymbol{P}_{\mathrm{tr}}(N)$ 是以 $\mathrm{tr}\boldsymbol{P}_{ij}(N)$ 为第 (i,j) 元素的 $L\times L$ 矩阵. 最优融合误差方差阵为

$$\boldsymbol{P}_0^s(N)=\sum_{i=1}^{L}\sum_{j=1}^{L}a_ia_j\boldsymbol{P}_{ij}(N) \tag{6.6.115}$$

按对角阵加权（按分量标量加权）最优融合稳态 Kalman 估值器为

$$\hat{\boldsymbol{x}}_0(t\mid t+N)=[\hat{x}_{01}(t\mid t+N),\cdots,\hat{x}_{0N}(t\mid t+N)]^{\mathrm{T}} \tag{6.6.116}$$

$$\hat{x}_{0i}(t\mid t+N)=\sum_{i=1}^{L}a_{ji}\hat{x}_{ji}(t\mid t+N) \tag{6.6.117}$$

$$\hat{\boldsymbol{x}}_j(t\mid t+N)=[\hat{x}_{j1}(t\mid t+N),\cdots,\hat{x}_{jn}(t\mid t+N)]^{\mathrm{T}} \tag{6.6.118}$$

且最优加权为

$$[a_{1i},\cdots,a_{Li}]=\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}^{ii}(N))^{-1}[\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}^{ii}(N))^{-1}\boldsymbol{e}]^{-1} \tag{6.6.119}$$

其中 $\boldsymbol{e}^{\mathrm{T}}=[1,\cdots,1]$，$\boldsymbol{P}^{ii}(N)$ 为以 $P_{kj}^{ii}(N)$ 为第 (k,j) 元素的 $L\times L$ 矩阵，而 $P_{kj}^{ii}(N)$ 为 $\boldsymbol{P}_{kj}(N)$ 的第 (i,i) 对角元素. 各分量最优融合估计误差方差为

$$P_{0i}(N)=[\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}^{ii}(N))^{-1}\boldsymbol{e}]^{-1} \tag{6.6.120}$$

且最优融合误差方差阵 $\boldsymbol{P}_0^c(N)$ 的迹为

$$\mathrm{tr}\boldsymbol{P}_0^c(N)=\sum_{i=1}^{n}P_{0i}(N) \tag{6.6.121}$$

在上述三种加权下，有精度关系

$$\mathrm{tr}\boldsymbol{P}_0^m(N)\leqslant\mathrm{tr}\boldsymbol{P}_0^d(N)\leqslant\mathrm{tr}\boldsymbol{P}_0^s(N)\leqslant\mathrm{tr}\boldsymbol{P}_i(N),\quad i=1,\cdots,L \tag{6.6.122}$$

且在上述公式中 $\boldsymbol{P}_{ij}(N)$，$\hat{\boldsymbol{x}}_i(t\mid t+N)$ 由定理 6.6.1 ~ 定理 6.6.8 有关公式计算. 注意上述加权阵或加权系数均与 N 有关.

证明　由节 6.1 的三种加权融合公式得证. □

【注 6.6.2】　同节 6.2 ~ 节 6.4 的非稳态最优信息融合 Kalman 估值器相比，信息融合稳态 Kalman 估值的优点是可减小计算负担，便于实际应用. 因为前者要求在每时刻计算 Kalman 估值器增益阵和最优融合加权，而后者增益阵和加权均为常的、非时变的，只需一次性离散计算即可，不要求每时刻更新.

【定理 6.6.14】　多传感器定常系统(6.6.1)和(6.6.2)在假设 1 ~ 5 下，有最优融合稳态白噪声反卷积估值器

$$\hat{\boldsymbol{w}}_0(t\mid t+N)=\sum_{i=1}^{L}\boldsymbol{A}_i\hat{\boldsymbol{w}}_i(t\mid t+N),\quad N\geqslant 0 \tag{6.6.123}$$

$$\hat{\boldsymbol{w}}_0(t \mid t+N)=\mathbf{0},\quad N<0 \tag{6.6.124}$$

其中 $\hat{\boldsymbol{w}}_i(t \mid t+N)$ 由定理6.6.12计算，在节6.1的三种加权融合准则下，$\boldsymbol{A}_i$ 分别为矩阵，对角阵或标量，可分别由节6.1的三种加权公式基于(6.6.103)～(6.6.106)给出的 $\boldsymbol{P}_i^w(N)$ 和 $\boldsymbol{P}_{ij}^w(N)$ 求得，且最优融合误差方差阵 $\boldsymbol{P}_0^w(N)$ 满足

$$\mathrm{tr}\boldsymbol{P}_0^w(N) \leqslant \mathrm{tr}\boldsymbol{P}_i^w(N),\quad i=1,\cdots,L \tag{6.6.125}$$

【例6.6.1】 考虑带有色观测噪声的三传感器雷达跟踪系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma}w(t) \tag{6.6.126}$$

$$z_i(t)=\boldsymbol{H}_0\boldsymbol{x}(t)+\eta_i(t) \tag{6.6.127}$$

$$\eta_i(t+1)=c_i\eta_i(t)+\xi_i(t),\quad i=1,2,3 \tag{6.6.128}$$

$$\boldsymbol{\Phi}=\begin{bmatrix}1 & T_0\\ 0 & 1\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}0.5T_0^2\\ T_0\end{bmatrix},\quad \boldsymbol{H}_0=\begin{bmatrix}1 & 0\end{bmatrix} \tag{6.6.129}$$

其中 T_0 为采样周期，状态 $\boldsymbol{x}(t)=[x_1(t),x_2(t)]^{\mathrm{T}}$，$x_1(t)$，$x_2(t)$ 和 $w(t)$ 分别为在时刻 tT_0 处运动目标(飞机、导弹、坦克、战舰等)的位置、速度和加速度，$z_i(t)$ 为第 i 传感器对位置的观测，$\eta_i(t)$ 为有色观测噪声.设 $w(t)$ 和 $\xi_i(t)$ 是零均值相互独立的、方差各为 σ_w^2 和 $\sigma_{\xi i}^2$ 的白噪声.仿真中取

$$T_0=0.2,\quad \sigma_w^2=4,\quad \sigma_{\xi1}^2=4,\quad \sigma_{\xi2}^2=8,\quad \sigma_{\xi3}^2=16,$$
$$c_1=0.5,\quad c_2=0.6,\quad c_3=0.8 \tag{6.6.130}$$

系统可化为带白色观测噪声系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t) \tag{6.6.131}$$

$$y_i(t)=\boldsymbol{H}_i\boldsymbol{x}(t)+v_i(t),\quad i=1,2,3 \tag{6.6.132}$$

其中定义

$$y_i(t)=z_i(t+1)-c_iz_i(t) \tag{6.6.133}$$

$$\boldsymbol{H}_i=\boldsymbol{H}_0\boldsymbol{\Phi}-c_i\boldsymbol{H}_0 \tag{6.6.134}$$

$$v_i(t)=\boldsymbol{H}_0\boldsymbol{\Gamma}w(t)+\xi_i(t) \tag{6.6.135}$$

显然 $v_i(t)$ 是带零均值、方差为 R_i 的相关于 $w(t)$ 的白噪声，且 $v_i(t)$ 与 $v_j(t)(i \neq j)$ 是相关的，且容易得到

$$R_i=\sigma_w^2\boldsymbol{H}_0\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}\boldsymbol{H}_0^{\mathrm{T}}+\sigma_{\xi i}^2 \tag{6.6.136}$$

$$S_i=\mathrm{E}[w(t)v_i^{\mathrm{T}}(t)]=\sigma_w^2\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{H}_0^{\mathrm{T}} \tag{6.6.137}$$

$$R_{ij}=\mathrm{E}[v_i(t)v_j^{\mathrm{T}}(t)]=\sigma_w^2\boldsymbol{H}_0\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}\boldsymbol{H}_0^{\mathrm{T}},\quad i \neq j \tag{6.6.138}$$

取 $N=2$，问题是求局部和融合的稳态 Kalman 平滑 $\hat{\boldsymbol{x}}_i(t \mid t+2)(i=1,2,3)$ 和 $\hat{\boldsymbol{x}}_0(t \mid t+2)$.

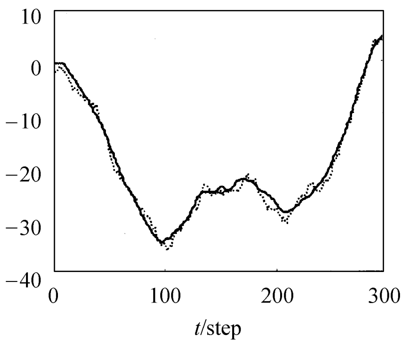
应用定理6.6.7，定理6.6.10和定理6.6.13可求得局部平滑误差方差阵的迹和按三种方式加权融合平滑误差方差阵的迹分别为

$$\mathrm{tr}\boldsymbol{P}_1(2)=2.355\,7,\quad \mathrm{tr}\boldsymbol{P}_2(2)=5.100\,7,\quad \mathrm{tr}\boldsymbol{P}_3(2)=21.488\,4,$$
$$\mathrm{tr}\boldsymbol{P}_0^m(2)=1.943\,0,\quad \mathrm{tr}\boldsymbol{P}_0^d(2)=1.952\,3,\quad \mathrm{tr}\boldsymbol{P}_0^s(2)=1.954\,7 \tag{6.6.139}$$

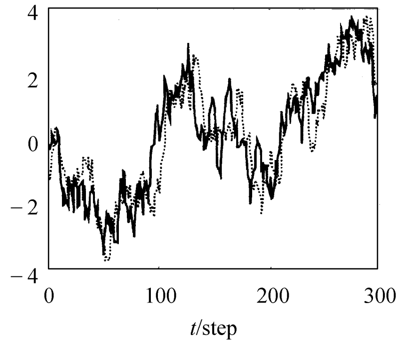
这引出精度关系

$$\mathrm{tr}\boldsymbol{P}_0^m(2)<\mathrm{tr}\boldsymbol{P}_0^d(2)<\mathrm{tr}\boldsymbol{P}_0^s(2)<\mathrm{tr}\boldsymbol{P}_i(2),\quad i=1,2,3 \tag{6.6.140}$$

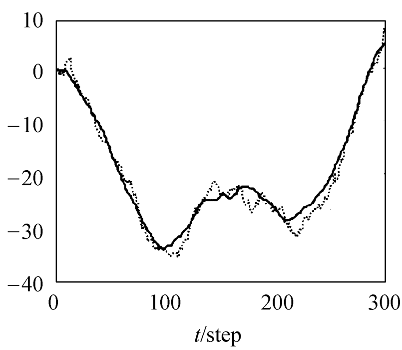
即精度关系(6.6.122)成立.仿真结果如图 6.6.1 ~ 图 6.6.2 所示,其中实线代表真实值 $x_i(t)$,虚线代表其平滑估值.可直观看到每种融合平滑估计精度均高于每个局部平滑估计精度,且三种加权融合平滑器的精度无显著区别.从理论上由(6.6.139)和(6.6.140)我们也得到相同的结论.因此从实时应用观点,应采用按标量加权融合估计.



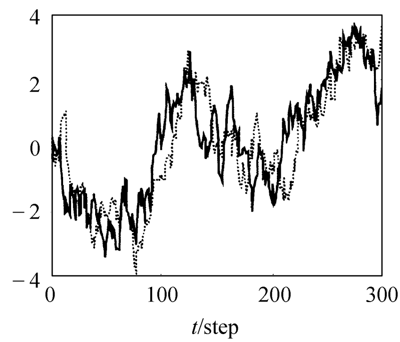
(a) 位置 $x_1(t)$ 与局部平滑器 $\hat{x}_{11}(t|t+2)$



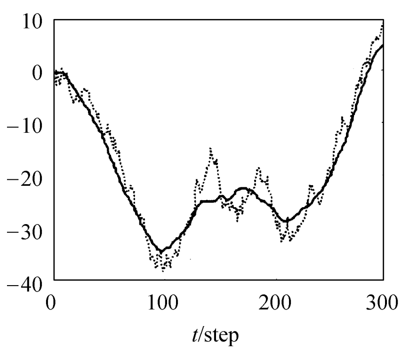
(b) 速度 $x_2(t)$ 与局部平滑器 $\hat{x}_{12}(t|t+2)$



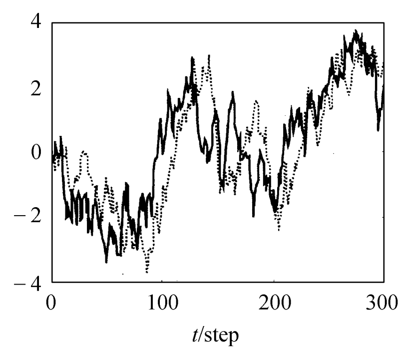
(c) 位置 $x_1(t)$ 与局部平滑器 $\hat{x}_{21}(t|t+2)$



(d) 速度 $x_2(t)$ 与局部平滑器 $\hat{x}_{22}(t|t+2)$



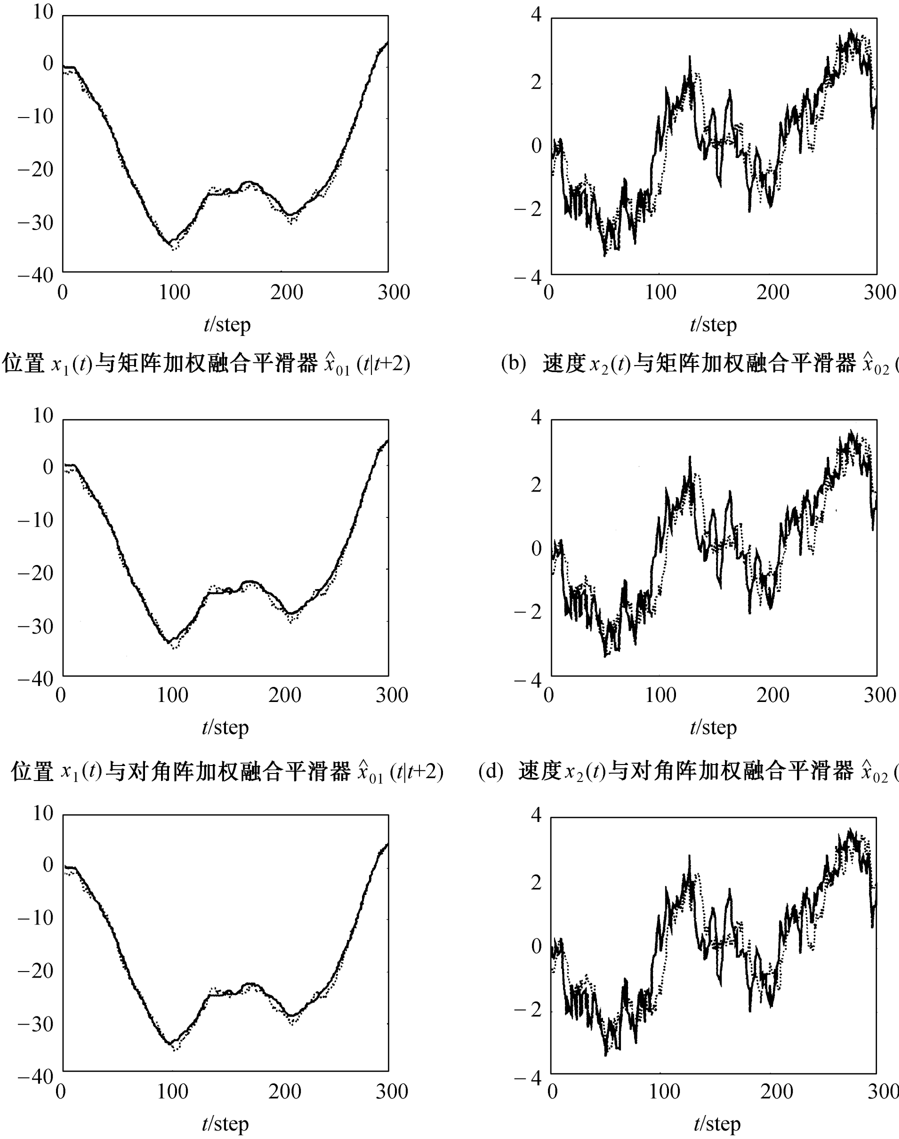
(e) 位置 $x_1(t)$ 与局部平滑器 $\hat{x}_{31}(t|t+2)$



(f) 速度 $x_2(t)$ 与局部平滑器 $\hat{x}_{32}(t|t+2)$

图 6.6.1 状态 $x(t)$ 和局部稳态 Kalman 跟踪平滑器 $\hat{x}_i(t\mid t+2)$

$(\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}, \hat{\boldsymbol{x}}_i(t\mid t+2) = [\hat{x}_{i1}(t\mid t+2), \hat{x}_{i2}(t\mid t+2)]^{\mathrm{T}}, i = 1,2,3)$



(a) 位置 $x_1(t)$ 与矩阵加权融合平滑器 $\hat{x}_{01}(t|t+2)$　(b) 速度 $x_2(t)$ 与矩阵加权融合平滑器 $\hat{x}_{02}(t|t+2)$

(c) 位置 $x_1(t)$ 与对角阵加权融合平滑器 $\hat{x}_{01}(t|t+2)$　(d) 速度 $x_2(t)$ 与对角阵加权融合平滑器 $\hat{x}_{02}(t|t+2)$

(e) 位置 $x_1(t)$ 与标量加权融合平滑器 $\hat{x}_{01}(t|t+2)$　(f) 速度 $x_2(t)$ 与标量加权融合平滑器 $\hat{x}_{02}(t|t+2)$

图 6.6.2　状态 $\boldsymbol{x}(t)$ 和按矩阵、对角和标量三种方式加权的信息融合稳态 Kalman 跟踪平滑器 $\hat{x}_0(t \mid t+2)$

6.7　基于 Kalman 滤波的两种观测融合方法的功能等价性

在节 6.2 ~ 节 6.4 及节 6.6 中提出了状态融合 Kalman 估值器，它由局部 Kalman 滤波器加权构成. 在节 6.1 的三种加权准则下，状态融合 Kalman 滤波器，也称分布式信息融合 Kalman 滤波器，并非全局最优（集中式）Kalman 滤波器，即三种加权信息融合 Kalman 滤波器是次优的. 本节介绍可得到全局最优估值的观测融合 Kalman 估值器[20]. 目前有两种最

优观测融合方法[8].一种方法(方法 Ⅰ)是集中式观测融合,即用增加观测向量维数的方法,合并各传感器的观测方程为一个增维的观测方程,这引出集中式 Kalman 滤波器,可获得全局最优(线性最小方差)状态估值.但缺点是由于观测方程维数的增加,增加了计算负担.另一种方法(方法 Ⅱ)是在各传感器具有相同的观测阵的条件下,用基于线性最小方差准则的加权方法合并各传感器的观测方程为一个新的观测方程,其中观测向量的维数不变,这也叫加权观测融合方法.这种观测融合方法不增加计算负担.问题是这两种观测融合方法是否功能等价?即用这两种观测融合方法所得 Kalman 估值器、信号和白噪声估值器在数值上是否相同?新近文献[8]证明了两种方法是部分功能等价的,即用两种方法得到的 Kalman 滤波器和一步 Kalman 预报器在数值上是分别相等的.但没有证明是否用两种方法得到的 Kalman 平滑器、多步 Kalman 预报器、信号和白噪声估值器也是数值上分别相同的?这叫两种方法的完全功能等价性.文献[20]对稳态估值情形证明了两种方法的完全功能等价性,推广了文献[8]的结果.

同方法 Ⅰ 相比,方法 Ⅱ 可明显减小计算负担,且可获得全局最优估计,具有全局最优性,但它的缺点是要求各传感器具有相同的观测阵,然而这一要求并不是苛刻的限制,例如在目标跟踪系统中各传感器均对目标位置或位置和速度进行观测就属这种情形.

本节对时变系统将提出不同于文献[8]的两种观测融合方法部分功能等价性的新的推导方法,并进一步证明两种方法的完全功能等价性,并包括文献[20]的结果作为特殊情形.

6.7.1 两种观测融合方法

考虑多传感器线性离散时变随机系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)+\boldsymbol{\Gamma}(t)\boldsymbol{w}(t) \tag{6.7.1}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{H}_i(t)\boldsymbol{x}(t)+\boldsymbol{v}_i(t) \tag{6.7.2}$$

$$\boldsymbol{s}(t)=\boldsymbol{F}(t)\boldsymbol{x}(t) \tag{6.7.3}$$

其中 t 为离散时间,$\boldsymbol{x}(t)\in R^n$ 为状态,$\boldsymbol{u}(t)\in R^p$ 为控制,$\boldsymbol{y}_i(t)\in R^{m_i}$ 为第 i 传感器的观测,$\boldsymbol{v}_i(t)\in R^{m_i}$ 为观测噪声,$\boldsymbol{s}(t)\in R^{n_s}$ 是有关于 $\boldsymbol{x}(t)$ 的信号,$\boldsymbol{\Phi}(t)$,$\boldsymbol{B}(t)$,$\boldsymbol{\Gamma}(t)$,$\boldsymbol{H}_i(t)$ 和 $\boldsymbol{F}(t)$ 是已知的适当维数的时变矩阵.

【假设 1】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}_i(t)\in R^{m_i}(i=1,\cdots,L)$ 是零均值、方差阵各为 $\boldsymbol{Q}(t)$ 和 $\boldsymbol{R}_i(t)>0$ 的相互独立的白噪声.

【假设 2】 $\boldsymbol{x}(0)$ 独立于 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}_i(t)(i=1,\cdots,L)$.

【假设 3】 $\boldsymbol{u}(t)$ 是已知的确定性时间序列.

观测融合方法 Ⅰ 是用增广观测向量方法合并各传感器观测方程为一个融合观测方程

$$\boldsymbol{y}^{(\mathrm{I})}(t)=\boldsymbol{H}^{(\mathrm{I})}(t)\boldsymbol{x}(t)+\boldsymbol{v}^{(\mathrm{I})}(t) \tag{6.7.4}$$

其中定义

$$\boldsymbol{y}^{(\mathrm{I})}(t)=\begin{bmatrix}\boldsymbol{y}_1(t)\\ \vdots\\ \boldsymbol{y}_L(t)\end{bmatrix},\quad \boldsymbol{H}^{(\mathrm{I})}(t)=\begin{bmatrix}\boldsymbol{H}_1(t)\\ \vdots\\ \boldsymbol{H}_L(t)\end{bmatrix},\quad \boldsymbol{v}^{(\mathrm{I})}(t)=\begin{bmatrix}\boldsymbol{v}_1(t)\\ \vdots\\ \boldsymbol{v}_L(t)\end{bmatrix} \tag{6.7.5}$$

显然合成观测白噪声 $\boldsymbol{v}^{(\mathrm{I})}(t)$ 的方差阵为

$$\boldsymbol{R}^{(\mathrm{I})}(t)=\operatorname{diag}(\boldsymbol{R}_1(t),\cdots,\boldsymbol{R}_L(t)) \tag{6.7.6}$$

对系统(6.7.1)和(6.7.4)应用 Kalman 滤波可求得全局最优观测融合估值器 $\hat{\boldsymbol{x}}^{(\mathrm{I})}(t\mid t+N)$，$\hat{\boldsymbol{w}}^{(\mathrm{I})}(t\mid t+N)$，$\hat{\boldsymbol{s}}^{(\mathrm{I})}(t\mid t+N)$，$N=0$，$N>0$ 或 $N<0$.

观测融合方法 Ⅱ 是用线性最小方差加权方法，即应用推论 6.1.1，假设各传感器有相同的观测阵，即

$$\boldsymbol{H}_1(t)=\boldsymbol{H}_2(t)=\cdots=\boldsymbol{H}_L(t)=\boldsymbol{H}_0(t) \tag{6.7.7}$$

可将 $\boldsymbol{y}_i(t)$ 视为对 $\boldsymbol{H}_0(t)\boldsymbol{x}(t)$ 的估计，$\boldsymbol{v}_i(t)$ 为估计误差，由推论 6.1.1，由 $\boldsymbol{v}_i(t)$ 的相互独立性，于是有最优加权观测融合方程为

$$\boldsymbol{y}^{(\mathrm{II})}(t)=\boldsymbol{H}^{(\mathrm{II})}(t)\boldsymbol{x}(t)+\boldsymbol{v}^{(\mathrm{II})}(t) \tag{6.7.8}$$

其中定义 $\boldsymbol{y}^{(\mathrm{II})}(t)$ 为对 $\boldsymbol{H}_0(t)\boldsymbol{x}(t)$ 的融合估计，$\boldsymbol{v}^{(\mathrm{II})}(t)$ 为融合估计误差，

$$\boldsymbol{y}^{(\mathrm{II})}(t)=\left[\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\right]^{-1}\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\boldsymbol{y}_i(t) \tag{6.7.9}$$

$$\boldsymbol{H}^{(\mathrm{II})}(t)=\left[\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\right]^{-1}\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\boldsymbol{H}_i(t)=\boldsymbol{H}_0(t) \tag{6.7.10}$$

$$\boldsymbol{v}^{(\mathrm{II})}(t)=\left[\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\right]^{-1}\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\boldsymbol{v}_i(t) \tag{6.7.11}$$

显然合成观测白噪声 $\boldsymbol{v}^{(\mathrm{II})}(t)$ 有方差阵

$$\boldsymbol{R}^{(\mathrm{II})}(t)=\left[\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\right]^{-1} \tag{6.7.12}$$

由假设 1，$\boldsymbol{R}_i(t)>0$ 引出

$$(\boldsymbol{R}^{(\mathrm{II})}(t))^{-1}=\boldsymbol{R}_1^{-1}(t)+\cdots+\boldsymbol{R}_L^{-1}(t)>\boldsymbol{R}_i^{-1}(t),\quad i=1,\cdots,L \tag{6.7.13}$$

从而有

$$\boldsymbol{R}^{(\mathrm{II})}(t)<\boldsymbol{R}_i(t),\quad i=1,\cdots,L \tag{6.7.14}$$

这引出

$$\operatorname{tr}\boldsymbol{R}^{(\mathrm{II})}(t)<\operatorname{tr}\boldsymbol{R}_i(t),\quad i=1,\cdots,L \tag{6.7.15}$$

这表明融合观测方程(6.7.8)改善了每个观测方程的精度.但融合观测 $\boldsymbol{y}^{(\mathrm{II})}(t)$ 的维数没增加.对系统(6.7.1)和(6.7.8)应用 Kalman 滤波可求得加权观测融合估值器 $\hat{\boldsymbol{x}}^{(\mathrm{II})}(t\mid t+N)$，$\hat{\boldsymbol{w}}^{(\mathrm{II})}(t\mid t+N)$，$\hat{\boldsymbol{s}}^{(\mathrm{II})}(t\mid t+N)$，$N=0$，$N>0$ 或 $N<0$.

问题是要证明两种方法是完全功能等价的，即 $\hat{\boldsymbol{x}}^{(\mathrm{I})}(t\mid t+N)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(t\mid t+N)$，$\hat{\boldsymbol{w}}^{(\mathrm{I})}(t\mid t+N)=\hat{\boldsymbol{w}}^{(\mathrm{II})}(t\mid t+N)$，$\hat{\boldsymbol{s}}^{(\mathrm{I})}(t\mid t+N)=\hat{\boldsymbol{s}}^{(\mathrm{II})}(t\mid t+N)$，$N=0$，$N>0$ 或 $N<0$.

6.7.2　两种观测融合方法的部分功能等价性

注意，观测融合方程(6.7.4)和(6.7.8)具有统一形式

$$\boldsymbol{y}(t)=\boldsymbol{H}(t)\boldsymbol{x}(t)+\boldsymbol{v}(t) \tag{6.7.16}$$

其中记 $\boldsymbol{v}(t)$ 的方差阵为 $\boldsymbol{R}(t)$.

【定理 6.7.1】　多传感器时变系统(6.7.1)和(6.7.16)在假设 1 ~ 3 下，有最优 Kalman 预报器 $\hat{\boldsymbol{x}}(t\mid t-1)$ 和最优 Kalman 滤波器 $\hat{\boldsymbol{x}}(t\mid t)$ 为

$$\hat{\boldsymbol{x}}(t \mid t-1) = \boldsymbol{\Phi}(t-1)\hat{\boldsymbol{x}}(t-1 \mid t-1) + \boldsymbol{B}(t-1)\boldsymbol{u}(t-1) \tag{6.7.17}$$

$$\hat{\boldsymbol{x}}(t \mid t) = \boldsymbol{\Psi}_f(t)\hat{\boldsymbol{x}}(t-1 \mid t-1) + [\boldsymbol{I}_n - \boldsymbol{K}_f(t)\boldsymbol{H}(t)]\boldsymbol{B}(t-1)\boldsymbol{u}(t-1) + \boldsymbol{K}_f(t)\boldsymbol{y}(t) \tag{6.7.18}$$

$$\boldsymbol{K}_f(t) = \boldsymbol{P}(t \mid t)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t) \tag{6.7.19}$$

$$\boldsymbol{\Psi}_f(t) = [\boldsymbol{I}_n - \boldsymbol{P}(t \mid t)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\boldsymbol{H}(t)]\boldsymbol{\Phi}(t-1) \tag{6.7.20}$$

$$[\boldsymbol{I}_n - \boldsymbol{K}_f(t)\boldsymbol{H}(t)] = \boldsymbol{I}_n - \boldsymbol{P}(t \mid t)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\boldsymbol{H}(t) \tag{6.7.21}$$

$$\boldsymbol{P}(t \mid t-1) = \boldsymbol{\Phi}(t-1)\boldsymbol{P}(t-1 \mid t-1)\boldsymbol{\Phi}^{\mathrm{T}}(t-1) + \boldsymbol{\Gamma}(t-1)\boldsymbol{Q}(t-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t-1) \tag{6.7.22}$$

$$\boldsymbol{P}^{-1}(t \mid t) = \boldsymbol{P}^{-1}(t \mid t-1) + \boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\boldsymbol{H}(t) \tag{6.7.23}$$

带初值 $\hat{\boldsymbol{x}}(0 \mid -1)$，$\boldsymbol{P}(0 \mid -1)$ 或 $\hat{\boldsymbol{x}}(0 \mid 0)$，$\boldsymbol{P}(0 \mid 0)$.

证明 由定理 3.3.4 有(6.7.17)、(6.7.18)和(6.7.22)，其中

$$\boldsymbol{K}_f(t) = \boldsymbol{P}(t \mid t-1)\boldsymbol{H}^{\mathrm{T}}(t)[\boldsymbol{H}(t)\boldsymbol{P}(t \mid t-1)\boldsymbol{H}^{\mathrm{T}}(t) + \boldsymbol{R}(t)]^{-1} \tag{6.7.24}$$

$$\boldsymbol{\Psi}_f(t) = [\boldsymbol{I}_n - \boldsymbol{K}_f(t)\boldsymbol{H}(t)]\boldsymbol{\Phi}(t-1) \tag{6.7.25}$$

$$\boldsymbol{P}(t \mid t) = [\boldsymbol{I}_n - \boldsymbol{K}_f(t)\boldsymbol{H}(t)]\boldsymbol{P}(t \mid t-1) \tag{6.7.26}$$

应用矩阵求逆引理

$$(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{C}^{\mathrm{T}})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I} + \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1} \tag{6.7.27}$$

可得(6.7.23). 事实上，应用(6.7.27)、(6.7.24)和(6.7.26)可引出(6.7.23). 对(6.7.24)应用矩阵求逆引理和(6.4.23)有

$$\begin{aligned}
\boldsymbol{K}_f(t) = {} & \boldsymbol{P}(t \mid t-1)\boldsymbol{H}^{\mathrm{T}}(t)[\boldsymbol{R}^{-1}(t) - \boldsymbol{R}^{-1}(t)\boldsymbol{H}(t)\boldsymbol{P}(t \mid t-1) \times \\
& (\boldsymbol{I} + \boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\boldsymbol{H}(t)\boldsymbol{P}(t \mid t-1))^{-1}\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)] = \\
& \boldsymbol{P}(t \mid t-1)\boldsymbol{H}^{\mathrm{T}}(t)\{\boldsymbol{R}^{-1}(t) + \boldsymbol{R}^{-1}(t)\boldsymbol{H}(t) \times \\
& [(\boldsymbol{I} + \boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\boldsymbol{H}(t)\boldsymbol{P}(t \mid t-1))\boldsymbol{P}^{-1}(t \mid t-1)]^{-1}\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\} = \\
& \boldsymbol{P}(t \mid t-1)[\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t) - \boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\boldsymbol{H}(t) \times \\
& (\boldsymbol{P}^{-1}(t \mid t) + \boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\boldsymbol{H}(t))^{-1}\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)] = \\
& \boldsymbol{P}(t \mid t-1)[\boldsymbol{I} - \boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\boldsymbol{H}(t)\boldsymbol{P}(t \mid t)]\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t) = \\
& \boldsymbol{P}(t \mid t-1)[\boldsymbol{I} - (\boldsymbol{P}^{-1}(t \mid t) - \boldsymbol{P}^{-1}(t \mid t-1))\boldsymbol{P}(t \mid t)]\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t) = \\
& \boldsymbol{P}(t \mid t)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)
\end{aligned} \tag{6.7.28}$$

这证明了(6.7.19). 将(6.7.19)代入(6.7.25)和(6.7.26)得(6.7.20)和(6.7.21). 证毕. □

由(6.7.22)和(6.7.23)看到，只要

$$\boldsymbol{H}^{(\mathrm{I})\mathrm{T}}(t)(\boldsymbol{R}^{(\mathrm{I})}(t))^{-1}\boldsymbol{H}^{(\mathrm{I})}(t) = \boldsymbol{H}^{(\mathrm{II})\mathrm{T}}(t)(\boldsymbol{R}^{(\mathrm{II})}(t))^{-1}\boldsymbol{H}^{(\mathrm{II})}(t) \tag{6.7.29}$$

且带相同初值 $\boldsymbol{P}^{(\mathrm{I})}(0 \mid -1) = \boldsymbol{P}^{(\mathrm{II})}(0 \mid -1)$ 或 $\boldsymbol{P}^{(\mathrm{I})}(0 \mid 0) = \boldsymbol{P}^{(\mathrm{II})}(0 \mid 0)$，则用两种观测融合方法引出相同的 $\boldsymbol{P}(t \mid t-1)$ 和 $\boldsymbol{P}(t \mid t)$，即

$$\boldsymbol{P}^{(\mathrm{I})}(t \mid t-1) = \boldsymbol{P}^{(\mathrm{II})}(t \mid t-1),\boldsymbol{P}^{(\mathrm{I})}(t \mid t) = \boldsymbol{P}^{(\mathrm{II})}(t \mid t) \tag{6.7.30}$$

进一步由(6.7.17) ~ (6.7.21)看到，只要

$$\boldsymbol{H}^{(\mathrm{I})\mathrm{T}}(t)(\boldsymbol{R}^{(\mathrm{I})}(t))^{-1}\boldsymbol{y}^{(\mathrm{I})}(t) = \boldsymbol{H}^{(\mathrm{II})\mathrm{T}}(t)(\boldsymbol{R}^{(\mathrm{II})}(t))^{-1}\boldsymbol{y}^{(\mathrm{II})}(t) \tag{6.7.31}$$

且带相同的初值 $\hat{\boldsymbol{x}}^{(\mathrm{I})}(0 \mid -1) = \hat{\boldsymbol{x}}^{(\mathrm{II})}(t \mid t)$ 或 $\hat{\boldsymbol{x}}^{(\mathrm{I})}(0 \mid 0) = \hat{\boldsymbol{x}}^{(\mathrm{II})}(t \mid 0)$，则两种观测融

合方法引出相同的 Kalman 预报器和滤波器，即

$$\hat{\boldsymbol{x}}^{(\mathrm{I})}(t\mid t-1)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(t\mid t-1),\quad \hat{\boldsymbol{x}}^{(\mathrm{I})}(t\mid t)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(t\mid t) \tag{6.7.32}$$

【定理 6.7.2】 （部分功能等价性）系统（6.7.1）和（6.7.2）在假设 1 ~ 3 下，再假设各传感器有相同的观测阵，即假设（6.7.7）成立，则用方法 Ⅰ 所得最优 Kalman 滤波器 $\hat{\boldsymbol{x}}^{(\mathrm{I})}(t\mid t)$ 和预报器 $\hat{\boldsymbol{x}}^{(\mathrm{I})}(t\mid t-1)$ 及相应的误差方差阵 $\boldsymbol{P}^{(\mathrm{I})}(t\mid t)$ 和 $\boldsymbol{P}^{(\mathrm{I})}(t\mid t-1)$ 在数值上恒同于用方法 Ⅱ 得到的最优 Kalman 滤波器 $\hat{\boldsymbol{x}}^{(\mathrm{II})}(t\mid t)$ 和预报器 $\hat{\boldsymbol{x}}^{(\mathrm{II})}(t\mid t-1)$ 及相应的误差方差阵 $\boldsymbol{P}^{(\mathrm{II})}(t\mid t)$ 和 $\boldsymbol{P}^{(\mathrm{II})}(t\mid t-1)$，即

$$\hat{\boldsymbol{x}}^{(\mathrm{I})}(t\mid t)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(t\mid t),\quad \boldsymbol{P}^{(\mathrm{I})}(t\mid t)=\boldsymbol{P}^{(\mathrm{II})}(t\mid t),\quad \forall t \tag{6.7.33}$$

$$\hat{\boldsymbol{x}}^{(\mathrm{I})}(t\mid t-1)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(t\mid t-1),$$
$$\boldsymbol{P}^{(\mathrm{I})}(t\mid t-1)=\boldsymbol{P}^{(\mathrm{II})}(t\mid t-1),\quad \forall t \tag{6.7.34}$$

只要它们有相同的初值

$$\hat{\boldsymbol{x}}^{(\mathrm{I})}(0\mid 0)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(0\mid 0),\quad \boldsymbol{P}^{(\mathrm{I})}(0\mid 0)=\boldsymbol{P}^{(\mathrm{II})}(0\mid 0) \tag{6.7.35}$$

或

$$\hat{\boldsymbol{x}}^{(\mathrm{I})}(0\mid -1)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(0\mid -1),\quad \boldsymbol{P}^{(\mathrm{I})}(0\mid -1)=\boldsymbol{P}^{(\mathrm{II})}(0\mid -1) \tag{6.7.36}$$

证明　只要证明（6.7.29）和（6.7.31）即可. 事实上，由（6.7.4）~（6.7.12）和（6.7.7）有

$$\boldsymbol{H}^{(\mathrm{I})\mathrm{T}}(t)(\boldsymbol{R}^{(\mathrm{I})}(t))^{-1}\boldsymbol{H}^{(\mathrm{I})}(t)=\sum_{i=1}^{L}\boldsymbol{H}_0^{\mathrm{T}}(t)\boldsymbol{R}_i^{-1}(t)\boldsymbol{H}_0(t) \tag{6.7.37}$$

$$\boldsymbol{H}^{(\mathrm{II})\mathrm{T}}(t)(\boldsymbol{R}^{(\mathrm{II})}(t))^{-1}\boldsymbol{H}^{(\mathrm{II})}(t)=\boldsymbol{H}_0^{\mathrm{T}}(t)\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\boldsymbol{H}_0(t)=\sum_{i=1}^{L}\boldsymbol{H}_0^{\mathrm{T}}(t)\boldsymbol{R}_i^{-1}(t)\boldsymbol{H}_0(t) \tag{6.7.38}$$

这引出（6.7.29）成立. 其次，注意

$$\boldsymbol{H}^{(\mathrm{I})\mathrm{T}}(t)(\boldsymbol{R}^{(\mathrm{I})}(t))^{-1}\boldsymbol{y}^{(\mathrm{I})}(t)=\sum_{i=1}^{L}\boldsymbol{H}_0^{\mathrm{T}}(t)\boldsymbol{R}_i^{-1}(t)\boldsymbol{y}_i(t) \tag{6.7.39}$$

$$\boldsymbol{H}^{(\mathrm{II})\mathrm{T}}(t)(\boldsymbol{R}^{(\mathrm{II})}(t))^{-1}\boldsymbol{y}^{(\mathrm{II})}(t)=\boldsymbol{H}_0^{\mathrm{T}}(t)\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\Big[\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\Big]^{-1}\sum_{i=1}^{L}\boldsymbol{R}_i^{-1}(t)\boldsymbol{y}_i(t) \tag{6.7.40}$$

这引出（6.7.30）成立. 证毕. □

以上证明了两种观测融合方法的部分功能等价性，即它们引出相同的 Kalman 滤波器和一步 Kalman 预报器.

6.7.3　两种观测融合方法的完全功能等价性

我们进一步证明两种观测融合方法引出相同的 Kalman 平滑器、多步 Kalman 预报器、输入白噪声估值器及信号估值器，这叫做完全功能等价性.

由定理 3.7.9 系统（6.7.1）和（6.7.16）有 Kalman 平滑器

$$\hat{\boldsymbol{x}}(t\mid t+N)=\hat{\boldsymbol{x}}(t\mid t)+\sum_{i=1}^{N}\boldsymbol{K}(t\mid t+i)\boldsymbol{\varepsilon}(t+i),\quad N\geqslant 1 \tag{6.7.41}$$

其中平滑增益 $\boldsymbol{K}(t\mid t+i)$ 为

$$\boldsymbol{K}(t \mid t+i)=\boldsymbol{P}(t \mid t-1)\{\prod_{j=0}^{i-1}\boldsymbol{\Psi}_p^{\mathrm{T}}(t+i)\}\boldsymbol{H}^{\mathrm{T}}(t+i)\boldsymbol{Q}_{\varepsilon}^{-1}(t+i), \quad i \geqslant 1 \tag{6.7.42}$$

$$\boldsymbol{\Psi}_p(t)=\boldsymbol{\Phi}(t)[\boldsymbol{I}_n-\boldsymbol{K}_f(t)\boldsymbol{H}(t)] \tag{6.7.43}$$

$$\boldsymbol{Q}_{\varepsilon}(t)=\boldsymbol{H}(t)\boldsymbol{P}(t \mid t-1)\boldsymbol{H}^{\mathrm{T}}(t)+\boldsymbol{R}(t) \tag{6.7.44}$$

且平滑误差方差阵 $\boldsymbol{P}(t \mid t+N)$ 为

$$\boldsymbol{P}(t \mid t+N)=\boldsymbol{P}(t \mid t-1)-\sum_{i=0}^{N}\boldsymbol{K}(t \mid t+i)\boldsymbol{Q}_{\varepsilon}(t+i)\boldsymbol{K}^{\mathrm{T}}(t \mid t+i) \tag{6.7.45}$$

由(6.7.24)有

$$\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{Q}_{\varepsilon}^{-1}(t)\boldsymbol{\varepsilon}(t)=\boldsymbol{P}^{-1}(t \mid t-1)\boldsymbol{K}_f(t)\boldsymbol{\varepsilon}(t) \tag{6.7.46}$$

而

$$\boldsymbol{K}_f(t)\boldsymbol{\varepsilon}(t)=\boldsymbol{K}_f(t)[\boldsymbol{y}(t)-\boldsymbol{H}(t)\hat{\boldsymbol{x}}(t \mid t-1)] \tag{6.7.47}$$

由(6.7.19)有

$$\boldsymbol{K}_f(t)\boldsymbol{y}(t)=\boldsymbol{P}(t \mid t)\boldsymbol{H}^{\mathrm{T}}(t)\boldsymbol{R}^{-1}(t)\boldsymbol{y}(t) \tag{6.7.48}$$

由(6.7.26)有

$$\boldsymbol{K}_f(t)\boldsymbol{H}(t)=[\boldsymbol{P}(t \mid t-1)-\boldsymbol{P}(t \mid t)]\boldsymbol{P}^{-1}(t \mid t-1) \tag{6.7.49}$$

由(6.7.31)和(6.7.30),由(6.7.46) ~ (6.7.49)有

$$\boldsymbol{\Psi}_p^{(\mathrm{I})}(t)=\boldsymbol{\Psi}_p^{(\mathrm{II})}(t) \tag{6.7.50}$$

$$\boldsymbol{H}^{(\mathrm{I})\mathrm{T}}(t)(\boldsymbol{Q}_{\varepsilon}^{(\mathrm{I})}(t))^{-1}\boldsymbol{\varepsilon}^{(\mathrm{I})}(t)=\boldsymbol{H}^{(\mathrm{II})\mathrm{T}}(t)(\boldsymbol{Q}_{\varepsilon}^{(\mathrm{II})}(t))^{-1}\boldsymbol{\varepsilon}^{(\mathrm{II})}(t) \tag{6.7.51}$$

进而有

$$\boldsymbol{K}^{(\mathrm{I})}(t \mid t+i)\boldsymbol{\varepsilon}^{(\mathrm{I})}(t+i)=\boldsymbol{K}^{(\mathrm{II})}(t \mid t+i)\boldsymbol{\varepsilon}^{(\mathrm{II})}(t+i) \tag{6.7.52}$$

于是由(6.7.41)引出

$$\hat{\boldsymbol{x}}^{(\mathrm{I})}(t \mid t+N)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(t \mid t+N) \tag{6.7.53}$$

且由(6.7.45)和(6.7.52)得

$$\boldsymbol{P}^{(\mathrm{I})}(t \mid t+N)=\boldsymbol{P}^{(\mathrm{II})}(t \mid t+N) \tag{6.7.54}$$

上述结果可概括为如下定理.

【定理 6.7.3】 多传感器观测融合系统(6.7.1)和(6.7.16)在假设 1 ~ 3 下,若(6.7.7)成立,用方法 Ⅰ 和方法 Ⅱ 所得观测融合 Kalman 平滑器及误差方差阵在数值上是恒同的,即

$$\hat{\boldsymbol{x}}^{(\mathrm{I})}(t \mid t+N)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(t \mid t+N), \quad \forall\, t, N>0 \tag{6.7.55}$$

$$\boldsymbol{P}^{(\mathrm{I})}(t \mid t+N)=\boldsymbol{P}^{(\mathrm{II})}(t \mid t+N) \tag{6.7.56}$$

它们具有相同形式(6.7.41) ~ (6.7.45).

下面研究超前 $N(>0)$ 步预报问题.由(6.7.1)迭代有关系(3.7.138),即

$$\boldsymbol{x}(t+N)=\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{x}(t+1)+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}(t+N,i)\boldsymbol{B}(i-1)\boldsymbol{u}(i-1)+\sum_{i=t+2}^{t+N}\boldsymbol{\Phi}(t+N,i)\boldsymbol{\Gamma}(i-1)\boldsymbol{w}(i-1), \quad N \geqslant 2 \tag{6.7.57}$$

$$\boldsymbol{\Phi}(t+N,i)=\boldsymbol{\Phi}(t+N-1)\cdots\boldsymbol{\Phi}(i),\quad \boldsymbol{\Phi}(t+N,t+N)=\boldsymbol{I}_n \tag{6.7.58}$$

引入变换 $i=t+j$,上式变为

$$\boldsymbol{x}(t+N)=\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{x}(t+1)+\sum_{j=2}^{N}\boldsymbol{\Phi}(t+N,t+j)\boldsymbol{B}(t+j-1)\boldsymbol{u}(i+j-1)+$$
$$\sum_{j=2}^{N}\boldsymbol{\Phi}(t+N,t+j)\boldsymbol{\Gamma}(t+j-1)\boldsymbol{w}(t+j-1) \tag{6.7.59}$$

取上式两边各项到线性流形 $L(\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots)$ 上的射影,由假设 1 ~ 3 有超前 N 步预报器

$$\hat{\boldsymbol{x}}(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\hat{\boldsymbol{x}}(t+1\mid t)+\sum_{j=2}^{N}\boldsymbol{\Phi}(t+N,t+j)\times$$
$$\boldsymbol{B}(t+j-1)\boldsymbol{u}(i+j-1) \tag{6.7.60}$$

上两式相减引出预报误差 $\tilde{\boldsymbol{x}}(t+N\mid t)=\boldsymbol{x}(t+N)-\hat{\boldsymbol{x}}(t+N\mid t)$ 为

$$\tilde{\boldsymbol{x}}(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\tilde{\boldsymbol{x}}(t+1\mid t)-\sum_{j=2}^{N}\boldsymbol{\Phi}(t+N,t+j)\times$$
$$\boldsymbol{\Gamma}(t+j-1)\boldsymbol{w}(i+j-1) \tag{6.7.61}$$

这引出预报误差方差阵 $\boldsymbol{P}(t+N\mid t)=\mathrm{E}[\tilde{\boldsymbol{x}}(t+N\mid t)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t+N\mid t)]$ 为

$$\boldsymbol{P}(t+N\mid t)=\boldsymbol{\Phi}(t+N,t+1)\boldsymbol{P}(t+1\mid t)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+1)+\sum_{j=2}^{N}\boldsymbol{\Phi}(t+N,t+j)\times$$
$$\boldsymbol{\Gamma}(t+j-1)\boldsymbol{Q}(t+j-1)\boldsymbol{\Gamma}^{\mathrm{T}}(t+j-1)\boldsymbol{\Phi}^{\mathrm{T}}(t+N,t+j) \tag{6.7.62}$$

上述结果和 $\hat{\boldsymbol{x}}^{(\mathrm{I})}(t+1\mid t)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(t+1\mid t)$ 引出如下定理.

【定理 6.7.4】 多传感器观测融合系统(6.7.1)和(6.7.16)在假设 1 ~ 3 下,若(6.7.7)成立,则用方法 Ⅰ 和方法 Ⅱ 所得观测融合超前 N 步 Kalman 预报器及误差方差阵在数值上是恒同的,即

$$\hat{\boldsymbol{x}}^{(\mathrm{I})}(t+N\mid t)=\hat{\boldsymbol{x}}^{(\mathrm{II})}(t+N\mid t),\quad \forall\, t,N>0 \tag{6.7.63}$$
$$\boldsymbol{P}^{(\mathrm{I})}(t+N\mid t)=\boldsymbol{P}^{(\mathrm{II})}(t+N\mid t) \tag{6.7.64}$$

它们具有相同形式(6.7.60)和(6.7.62).

【定理 6.7.5】 对多传感器观测融合系统(6.7.1),(6.7.3)和(6.7.16),在假设 1 ~ 3 下,若(6.7.7)成立,则用方法 Ⅰ 和方法 Ⅱ 所得信号估值器 $\hat{s}(t\mid t+N)(N=0,N>0,N<0)$ 在数值上是恒同的,即

$$\hat{\boldsymbol{s}}^{(\mathrm{I})}(t\mid t+N)=\hat{\boldsymbol{s}}^{(\mathrm{II})}(t\mid t+N) \tag{6.7.65}$$

且有相同的误差方差阵.

证明 由(6.7.3)有

$$\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{F}(t)\hat{\boldsymbol{x}}(t\mid t+N) \tag{6.7.66}$$

根据定理 6.7.2 ~ 定理 6.7.4,由 $\hat{\boldsymbol{x}}(t\mid t+N)(N=0,N>0,N<0)$ 在两种方法下的功能等价性引出(6.7.65). □

下面考虑白噪声估值器的功能等价性.对观测融合系统(6.7.1)和(6.7.16),由定理 3.7.12 有输入白噪声估值器

$$\hat{\boldsymbol{w}}(t\mid t+N)=\sum_{i=1}^{N}\boldsymbol{M}(t\mid t+i)\boldsymbol{\varepsilon}(t+i),\quad N\geqslant 1 \tag{6.7.67}$$
$$\hat{\boldsymbol{w}}(t\mid t+N)=0,\quad N\leqslant 0 \tag{6.7.68}$$

$$M(t \mid t+i) = Q(t)\Gamma^{\mathrm{T}}(t)\{\prod_{j=1}^{i-1}\Psi_p^{\mathrm{T}}(t+j)\}H^{\mathrm{T}}(t+i)Q_\varepsilon^{-1}(t+i), \quad N \geqslant 1 \tag{6.7.69}$$

相应的估值误差方差阵为

$$P_w(t \mid t+N) = Q(t) - \sum_{i=1}^{N} M(t \mid t+i)Q_\varepsilon(t+i)M^{\mathrm{T}}(t \mid t+i), \quad N \geqslant 1 \tag{6.7.70}$$

$$P_w(t \mid t+N) = Q(t), \quad N \leqslant 0 \tag{6.7.71}$$

由关系(6.7.50)和(6.7.51),由(6.7.67)~(6.7.71)引出用两种方法所得白噪声估值器的等价性.

【定理 6.7.6】 对观测融合系统(6.7.1)和(6.7.16)在假设1~3下,若(6.7.7)成立,则用两种方法所得白噪声估值器 $\hat{x}(t \mid t+N)$ 及其误差方差阵 $P_w(t \mid t+N)$ 在数值上是恒同的,即

$$\hat{w}^{(\mathrm{I})}(t \mid t+N) = \hat{w}^{(\mathrm{II})}(t \mid t+N) \tag{6.7.72}$$

$$P_w^{(\mathrm{I})}(t \mid t+N) = P_w^{(\mathrm{II})}(t \mid t+N) \tag{6.7.73}$$

它们有相同的形式(6.7.67)~(6.7.71).

【定理 6.7.7】 (两种观测融合方法的完全功能等价性)多传感器系统(6.7.1)~(6.7.3)在假设1~3下,若各传感器有相同的观测阵,即(6.7.7)成立,则观测融合方法Ⅰ和方法Ⅱ是完全功能等价的,即它们得到在数值上相同的最优融合Kalman估值器、白噪声估值器及信号估值器,并且具有相同的估值误差方差阵.

证明 由定理6.7.2~定理6.7.6得证. □

【注 6.7.1】 定理6.7.7推广了文献[8]的两种观测融合方法的部分功能等价性,并且证明部分功能等价性(两种方法分别引出相同的Kalman滤波器和一步Kalman预报器)的方法也不同于文献[8]的方法.文献[8]用信息滤波器来证明部分功能等价性,这里避免了信息滤波器.

6.7.4 基于稳态 Kalman 滤波的两种观测融合方法的完全功能等价性

考虑多传感器定常随机控制系统

$$x(t+1) = \Phi x(t) + Bu(t) + \Gamma w(t) \tag{6.7.74}$$

$$y_i(t) = Hx(t) + v_i(t), \quad i = 1,\cdots,L \tag{6.7.75}$$

$$s(t) = Fx(t) \tag{6.7.76}$$

其中 t 为离散时间,$x(t) \in R^n$ 为状态,$u(t) \in R^p$ 为控制,$y_i(t) \in R^m$ 为第 i 传感器的观测,$v_i(t) \in R^i$ 为观测噪声,L 为传感器个数,各传感器有相同的观测阵 H,$s(t) \in R^q$ 为有关于 $x(t)$ 的信号,Φ,B,Γ,H 和 F 为适当维数常阵.

【假设 1】 $w(t) \in R^r$ 和 $v_i(t) \in R^m (i = 1,\cdots,L)$ 是零均值、方差阵各为 Q 和 $R_i > 0$ 的相互独立的白噪声.

【假设 2】 (Φ, H) 为完全可观对,(Φ, Γ) 为完全可控对.

【假设 3】 $u(t)$ 是已知的确定性时间序列.

注意，假设 2 保证了局部稳态 Kalman 估值器存在.

【定理 6.7.7】 多传感器定常随机控制系统(6.7.74)～(6.7.76)在假设 1～3 下，用方法 Ⅰ 和方法 Ⅱ 所得稳态 Kalman 估值器、信号估值器及输入白噪声估值器在数值上分别相等，即它们是完全功能等价的.

证明 当 $t_0 = 0$ 时，由定理 6.7.7，只要两种方法的初值(6.7.34)或(6.7.35)相同，则所得非稳态最优状态信号和输入白噪声估值器在数值上是相同的. 当 $t_0 = -\infty$ 时，这等价于 $t_0 = 0$，但 $t \to \infty$. 在定理 6.7.1～6.7.7 中令 $t \to \infty$ 引出相应的稳态估值器也是恒同的. 证毕. □

6.7.5 应用于多传感器观测融合最优信号估计问题

考虑带多传感器的多通道 ARMA 信号

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{6.7.77}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{s}(t) + \boldsymbol{v}_i(t), \quad i = 1,\cdots,L \tag{6.7.78}$$

其中 $\boldsymbol{s}(t) \in R^m$ 为待估的 ARMA 信号，q^{-1} 为单位滞后算子，多项式矩阵 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{C}(q^{-1})$ 有形式

$$\boldsymbol{A}(q^{-1}) = \boldsymbol{I}_m + \boldsymbol{A}_1 q^{-1} + \cdots + \boldsymbol{A}_{n_a} q^{-n_a},$$

$$\boldsymbol{C}(q^{-1}) = \boldsymbol{C}_1 q^{-1} + \cdots + \boldsymbol{C}_{n_c} q^{-n_c}, \quad n_a \geqslant n_c \tag{6.7.79}$$

$\boldsymbol{y}_i(t) \in R^m$ 为第 i 个传感器对 $\boldsymbol{s}(t)$ 的观测信号，$\boldsymbol{v}_i(t) \in R^m$ 为观测噪声. 假设 $(\boldsymbol{A}(q^{-1}), \boldsymbol{C}(q^{-1}))$ 左素，且 $\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}_i(t) \in R^m$ 为零均值、方差阵各为 $\boldsymbol{Q}$ 和 $\boldsymbol{R}_i$ 的相互独立白噪声.

【定理 6.7.8】 带多传感器的多通道 ARMA 信号(6.7.77)和(6.7.78)，当 $t_0 = -\infty$ 时用观测融合方法 Ⅰ 得到的全局最优稳态估值器 $\hat{\boldsymbol{s}}(t \mid t+N)$ $(N = 0, N > 0$ 或 $N < 0)$ 在数值上恒同于基于用观测融合方法 Ⅱ 得到的融合观测方程

$$\boldsymbol{y}(t) = \boldsymbol{s}(t) + \boldsymbol{v}(t) \tag{6.7.80}$$

的 ARMA 信号(6.7.77)的稳态最优估值器 $\hat{\boldsymbol{s}}(t \mid t+N)$，其中 $\boldsymbol{y}(t)$ 为用方法 Ⅱ 的融合观测，$\boldsymbol{v}(t)$ 为用方法 Ⅱ 的融合观测噪声，$\boldsymbol{R}$ 为它的方差，即

$$\boldsymbol{y}(t) = \Big[\sum_{i=1}^{L} \boldsymbol{R}_i^{-1}\Big]^{-1} \sum_{i=1}^{L} \boldsymbol{R}_i^{-1} \boldsymbol{y}_i(t) \tag{6.7.81}$$

$$\boldsymbol{v}(t) = \Big[\sum_{i=1}^{L} \boldsymbol{R}_i^{-1}\Big]^{-1} \sum_{i=1}^{L} \boldsymbol{R}_i^{-1} \boldsymbol{v}_i(t) \tag{6.7.82}$$

$$\boldsymbol{R} = \Big[\sum_{i=1}^{L} \boldsymbol{R}_i^{-1}\Big]^{-1} \tag{6.7.83}$$

证明 (6.7.77)和(6.7.78)有状态空间模型

$$\boldsymbol{x}(t+1) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{C}\boldsymbol{w}(t) \tag{6.7.84}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}_i(t), \quad i = 1,\cdots,L \tag{6.7.85}$$

$$\boldsymbol{s}(t) = \boldsymbol{H}\boldsymbol{x}(t) \tag{6.7.86}$$

$$\boldsymbol{A} = \begin{bmatrix} -\boldsymbol{A}_1 & & \\ \vdots & \boldsymbol{I}_{m(n_a-1)} & \\ -\boldsymbol{A}_{n_a} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_1 \\ \vdots \\ \boldsymbol{C}_{n_a} \end{bmatrix}, \quad \boldsymbol{H} = [\boldsymbol{I}_m \quad \boldsymbol{0} \cdots \boldsymbol{0}] \tag{6.7.87}$$

其中规定 $\boldsymbol{C}_i = \boldsymbol{0}(i > n_c)$，由 $(\boldsymbol{A}(q^{-1}), \boldsymbol{C}(q^{-1}))$ 左素的假设引出系统(6.7.84)和(6.7.85)是完全可观、完全可控的[34]，因而稳态 Kalman 估值器存在. 因各传感器的观测方程(6.7.85)有相同的观测阵 $\boldsymbol{H}$，故系统(6.7.84)～(6.7.86)的用方法 Ⅰ 的全局最优估计问题等价于用方法 Ⅱ 的观测融合系统的最优估计问题：

$$\boldsymbol{x}(t+1) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{C}\boldsymbol{w}(t) \tag{6.7.88}$$

$$\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}(t) \tag{6.7.89}$$

$$\boldsymbol{s}(t) = \boldsymbol{H}\boldsymbol{x}(t) \tag{6.7.90}$$

其中 $\boldsymbol{A}, \boldsymbol{C}, \boldsymbol{H}, \boldsymbol{y}(t), \boldsymbol{v}(t)$ 和 $\boldsymbol{R}$ 由(6.7.87)和(6.7.81)～(6.7.83)定义，而系统(6.7.88)～(6.7.90)的最优估计问题等价于系统(6.7.77)和(6.7.80)的最优估计问题. 证毕. □

【注 6.7.2】 定理 6.7.8 的重要意义在于带多个观测方程的 ARMA 信号最优估计问题可转化为带一个观测融合方程的 ARMA 信号最优估计问题，因而可明显减小计算负担和计算复杂性，且还可获得全局最优估计.

【定理 6.7.9】 多传感器多通道 ARMA 信号(6.7.77)和(6.7.78)有全局最优加权观测融合 Wiener 滤波器 $\hat{\boldsymbol{s}}_0(t \mid t+N)$

$$\psi(q^{-1})\hat{\boldsymbol{s}}_0(t \mid t+N) = \boldsymbol{\Psi}_N(q^{-1})\boldsymbol{y}(t+N) \tag{6.7.91}$$

其中 $\boldsymbol{y}(t)$ 由(6.7.81)定义，且 $n = mn_a$，

$$\boldsymbol{\Psi}_N(q^{-1}) = \boldsymbol{H}[\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p q^{-1-N} + \boldsymbol{K}_N(q^{-1})\boldsymbol{\Lambda}(q^{-1})], \quad N \geqslant 0 \tag{6.7.92}$$

$$\boldsymbol{\Psi}_N(q^{-1}) = \boldsymbol{H}\boldsymbol{A}^{-N-1}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p, \quad N \leqslant -1 \tag{6.7.93}$$

$$\psi(q^{-1}) = \det(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p) \tag{6.7.94}$$

$$\boldsymbol{\Lambda}(q^{-1}) = \psi(q^{-1})\boldsymbol{I}_m - \boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p q^{-1} \tag{6.7.95}$$

$$\boldsymbol{\Psi}_p = \boldsymbol{A} - \boldsymbol{K}_p\boldsymbol{H} \tag{6.7.96}$$

$$\boldsymbol{K}_p = \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}]^{-1} \tag{6.7.97}$$

其中 $\boldsymbol{R}$ 由(6.7.83)定义，且 $\boldsymbol{\Sigma}$ 满足 Riccati 方程

$$\boldsymbol{\Sigma} = \boldsymbol{A}[\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\boldsymbol{H}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R})^{-1}\boldsymbol{H}\boldsymbol{\Sigma}]\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{C}\boldsymbol{Q}\boldsymbol{C}^{\mathrm{T}} \tag{6.7.98}$$

且由(6.7.87)定义 $\boldsymbol{A}, \boldsymbol{H}, \boldsymbol{C}$，且定义

$$\boldsymbol{K}_N(q^{-1}) = \sum_{j=0}^{N}\boldsymbol{K}_j q^{j-N}, \quad N \geqslant 0 \tag{6.7.99}$$

$$\boldsymbol{K}_j = \boldsymbol{\Sigma}\boldsymbol{\Psi}_p^{\mathrm{T}j}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R})^{-1} \tag{6.7.100}$$

且有最优融合误差方差阵 $\boldsymbol{P}_{s0}(N)$ 为

$$\boldsymbol{P}_{s0}(N) = \boldsymbol{H}\boldsymbol{\Sigma}(N)\boldsymbol{H}^{\mathrm{T}} \tag{6.7.101}$$

其中 $\boldsymbol{\Sigma}(N)$ 定义为

$$\boldsymbol{\Sigma}(N) = \boldsymbol{\Sigma} - \sum_{j=0}^{N}\boldsymbol{K}_j(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R})\boldsymbol{K}_j^{\mathrm{T}}, \quad N \geqslant 0 \tag{6.7.102}$$

$$\boldsymbol{\Sigma}(N) = \boldsymbol{A}^{-N-1}\boldsymbol{\Sigma}\boldsymbol{A}^{(-N-1)\mathrm{T}} + \sum_{j=2}^{N}\boldsymbol{A}^{-N-j}\boldsymbol{C}\boldsymbol{Q}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-(N-j)T}, \quad N < -1 \tag{6.7.103}$$

$$\boldsymbol{\Sigma}(-1) = \boldsymbol{\Sigma} \tag{6.7.104}$$

证明 最优加权观测融合系统(6.7.77)和(6.7.80)～(6.7.83)等价于状态空间模

型(6.7.88) ~ (6.7.90). 它有稳态最优 Kalman 预报器

$$\hat{\boldsymbol{x}}(t \mid t-1) = (\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)^{-1}\boldsymbol{K}_p \boldsymbol{y}(t-1) \tag{6.7.105}$$

将它代入

$$\boldsymbol{y}(t) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{\varepsilon}(t) \tag{6.7.106}$$

可得 ARMA 新息模型

$$\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t) = \psi(q^{-1})\boldsymbol{\varepsilon}(t) \tag{6.7.107}$$

由定理 3.8.5 有稳态最优 Kalman 平滑器

$$\hat{\boldsymbol{x}}(t \mid t+N) = \hat{\boldsymbol{x}}(t \mid t-1) + \boldsymbol{K}_N(q^{-1})\boldsymbol{\varepsilon}(t+N), \quad N \geqslant 0 \tag{6.7.108}$$

由(6.7.107)有

$$\boldsymbol{\varepsilon}(t+N) = \psi^{-1}(q^{-1})\boldsymbol{\Lambda}(q^{-1})\boldsymbol{y}(t+N) \tag{6.7.109}$$

将(6.7.105)和(6.7.109)代入(6.7.108),并应用

$$(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)^{-1} = \mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)/\psi(q^{-1}) \tag{6.7.110}$$

注意由(6.7.90)有关系 $\hat{\boldsymbol{s}}_0(t \mid t+N) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t+N)$,可得(6.7.91)和(6.7.92),当 $N \leqslant -1$ 时,由定理 3.8.4 有

$$\hat{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{A}^{-N-1}\hat{\boldsymbol{x}}(t+N+1 \mid t+N), \quad N \leqslant -1 \tag{6.7.111}$$

利用(6.7.105)和上式可得(6.7.91)和(6.7.93). 由(6.7.90)有 $\hat{\boldsymbol{s}}_0(t \mid t+N) = \boldsymbol{H}\hat{\boldsymbol{x}}(t \mid t+N)$,于是有估值误差 $\tilde{\boldsymbol{s}}_0(t \mid t+N) = \boldsymbol{s}(t) - \hat{\boldsymbol{s}}_0(t \mid t+N)$ 的方差阵 $\boldsymbol{P}_{s0}(N) = \mathrm{E}[\tilde{\boldsymbol{s}}_0(t \mid t+N)\tilde{\boldsymbol{s}}_0^{\mathrm{T}}(t \mid t+N)]$ 满足关系(6.7.101),其中 $\boldsymbol{\Sigma}(N)$ 为状态估计误差 $\tilde{\boldsymbol{x}}(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t \mid t+N)$ 的方差阵 $\boldsymbol{\Sigma}(N) = E[\tilde{\boldsymbol{x}}(t \mid t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t \mid t+N)]$,由定理 3.8.4 和定理 3.8.5 得(6.7.102) ~ (6.7.104). 证毕. □

【定理 6.7.10】 多传感器多通道 ARMA 信号(6.7.77)和(6.7.78)有渐近稳定的全局最优加权观测融合 Wiener 估值器 $\hat{\boldsymbol{s}}_0(t \mid t+N)$($N=0, N>0$ 或 $N<0$)为

$$\psi(q^{-1})\hat{\boldsymbol{s}}_0(t \mid t+N) = \boldsymbol{\Psi}_N(q^{-1})\boldsymbol{y}(t+N) \tag{6.7.112}$$

其中加权融合观测 $\boldsymbol{y}(t)$ 和融合噪声 $\boldsymbol{v}(t)$ 由(6.7.81) ~ (6.7.83)定义,$\psi(q^{-1})$ 由(6.7.94)定义,且

$$\boldsymbol{\Psi}_N(q^{-1}) = \boldsymbol{J}_{-N}(q^{-1}) - \boldsymbol{L}_N^v(q^{-1})\boldsymbol{\Lambda}(q^{-1}) \tag{6.7.113}$$

$$\boldsymbol{J}_i(q^{-1}) = \begin{cases} \boldsymbol{H}\boldsymbol{A}^{i-1}\mathrm{adj}(\boldsymbol{I}_{mn_a} - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p, & i > 0 \\ \psi(q^{-1})\boldsymbol{I}_m q^i, & i \leqslant 0 \end{cases} \tag{6.7.114}$$

$$\boldsymbol{L}_N^v(q^{-1}) = \sum_{i=1}^{N} \boldsymbol{M}_v(i) q^{i-N}, \quad N \geqslant 0 \tag{6.7.115}$$

$$\boldsymbol{L}_N^v(q^{-1}) = \boldsymbol{0}, \quad N < 0 \tag{6.7.116}$$

$$\boldsymbol{M}_v(i) = -\boldsymbol{R}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Psi}_p^{\mathrm{T}(i-1)}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}]^{-1} \tag{6.7.117}$$

$$\boldsymbol{M}_v(0) = \boldsymbol{R}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}]^{-1} \tag{6.7.118}$$

$$\boldsymbol{K} = \boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}]^{-1} \tag{6.7.119}$$

且 $\boldsymbol{\Lambda}(q^{-1}), \boldsymbol{K}_p, \boldsymbol{\Sigma}$ 由(6.7.95),(6.7.97)和(6.7.98)计算. 相应的估值误差 $\tilde{\boldsymbol{s}}(t \mid t+N) = \boldsymbol{s}(t) - \hat{\boldsymbol{s}}_0(t \mid t+N)$ 方差阵 $\boldsymbol{P}_{s0}(N) = \mathrm{E}[\tilde{\boldsymbol{s}}_0(t \mid t+N)\tilde{\boldsymbol{s}}_0^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}_{s0}(N) = \boldsymbol{R} - \sum_{i=0}^{N} \boldsymbol{M}_v(i)[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}]\boldsymbol{M}_v^{\mathrm{T}}(i), \quad N \geqslant 0 \tag{6.7.120}$$

$$P_{s0}(N) = H\Sigma(N)H^{\mathrm{T}}, \quad N < 0 \tag{6.7.121}$$

其中 $\Sigma(N)$ 由(6.7.103)和(6.7.104)定义.

证明 由(6.7.80)有

$$\hat{s}(t \mid t+N) = \hat{y}(t \mid t+N) - \hat{v}(t \mid t+N) \tag{6.7.122}$$

由定理 3.8.4 和定理 3.8.7 有

$$\hat{y}(t \mid t+N) = HA^{-N-1}\hat{x}(t+N+1 \mid t+N), \quad N \leqslant -1 \tag{6.7.123}$$

$$\hat{y}(t \mid t+N) = y(t), \quad N \geqslant 0 \tag{6.7.124}$$

$$\hat{v}(t \mid t+N) = L_N^v(q^{-1})\varepsilon(t+N) \tag{6.7.125}$$

将(6.7.105)代入(6.7.123),并将(6.7.109)代入(6.7.125)后,由(6.7.122)可得(6.7.112)~(6.7.114).而(6.7.115)~(6.7.119)由定理 3.8.6 给出.当 $N \geqslant 0$ 时易知有 $P_{s0}(N) = P_v(N)$,其中 $P_v(N)$ 是 $\hat{v}(t \mid t+N)$ 的估值误差方差阵,由定理 3.8.6 有(6.7.120).当 $N < 0$ 时,(6.7.121)由(6.7.101)、(6.7.103)和(6.7.104)给出.因 Ψ_p 为稳定矩阵,则 $\psi(q^{-1})$ 是稳定的多项式,故(6.7.112)又是渐近稳定的.证毕. □

6.7.6 应用于设计多传感器观测融合白噪声反卷积 Wiener 估值器

考虑多传感器白噪声反卷积系统

$$y_i(t) = \frac{B(q^{-1})}{A(q^{-1})}w(t) + v_i(t), \quad i = 1, \cdots, L \tag{6.7.126}$$

其中 $w(t)$ 是输入白噪声,观测 $y_i(t)$ 是第 i 个传感器的输出,$v_i(t)$ 是观测噪声,且 $w(t)$ 和 $v_i(t)$ 是零均值、方差各为 σ_w^2 和 σ_v^2 的相互独立白噪声.$A(q^{-1})$ 和 $B(q^{-1})$ 为形如 $X(q^{-1}) = x_0 + x_1 q^{-1} + \cdots + x_{n_x} q^{-n_x}$ 的单位滞后算子 q^{-1} 的多项式,$a_0 = 1$,$b_0 = 0$,$n_a \geqslant n_b$,且 $A(q^{-1})$ 与 $B(q^{-1})$ 互质.问题是求 $w(t)$ 的局部和全局最优加权观测融合白噪声 Wiener 反卷积估值器.

【定理 6.7.10】 多传感器白噪声反卷积系统(6.7.126)当 $t_0 = -\infty$ 时的全局最优观测融合白噪声反卷积 Wiener 估值器 $\hat{w}_0(t \mid t+N)$ $(N = 0, N > 0$ 或 $N < 0)$ 在数值上恒同于如下加权观测融合反卷积系统的稳态最优 Wiener 反卷积估值器 $\hat{w}_0(t \mid t+N)$,

$$y(t) = \frac{B(q^{-1})}{A(q^{-1})}w(t) + v(t) \tag{6.7.127}$$

其中 $y(t)$ 和 $v(t)$ 为用方法 Ⅱ 的加权融合观测和加权融合观测噪声,

$$y(t) = \left(\sum_{i=1}^{L} \frac{1}{\sigma_{vi}^2}\right)^{-1} \sum_{i=1}^{L} \frac{1}{\sigma_{vi}^2} y_i(t) \tag{6.7.128}$$

$$v(t) = \left(\sum_{i=1}^{L} \frac{1}{\sigma_{vi}^2}\right)^{-1} \sum_{i=1}^{L} \frac{1}{\sigma_{vi}^2} v_i(t) \tag{6.7.129}$$

且 $v(t)$ 为零均值、方差为 σ_v^2 的白噪声,

$$\sigma_v^2 = \left(\sum_{i=1}^{L} \frac{1}{\sigma_{vi}^2}\right)^{-1} \tag{6.7.130}$$

证明 类似于定理 6.7.8 的证明,从略. □

下面基于 Kalman 滤波方法解决问题.(6.7.127)有状态空间模型

$$x(t+1) = Ax(t) + Bw(t) \tag{6.7.131}$$

$$y(t)=\boldsymbol{H}\boldsymbol{x}(t)+v(t) \tag{6.7.132}$$

其中矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{H}$ 定义为

$$\boldsymbol{A}=\begin{bmatrix}-a_1 & & & \\ \vdots & \boldsymbol{I}_{n_a-1} & & \\ -a_{n_a} & 0 & \cdots & 0\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}b_1\\ \vdots\\ b_{n_a}\end{bmatrix},\quad \boldsymbol{H}=[1\quad 0\cdots 0] \tag{6.7.133}$$

其中规定 $b_i=0(i>n_b)$，易知系统(6.7.131)和(6.7.132)是完全可观、完全可控的.类似于定理6.7.9中(6.7.105)～(6.7.107)的推导可得ARMA新息模型

$$\Lambda(q^{-1})y(t)=\psi(q^{-1})\varepsilon(t) \tag{6.7.134}$$

$$\Lambda(q^{-1})=\psi(q^{-1})-\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_{n_a}-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_pq^{-1} \tag{6.7.135}$$

$$\psi(q^{-1})=\det(\boldsymbol{I}_{n_a}-q^{-1}\boldsymbol{\Psi}_p) \tag{6.7.136}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{A}-\boldsymbol{K}_p\boldsymbol{H} \tag{6.7.137}$$

$$\boldsymbol{K}_p=\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\sigma_v^2]^{-1} \tag{6.7.138}$$

其中 $\boldsymbol{\Sigma}$ 满足Riccati方程

$$\boldsymbol{\Sigma}=\boldsymbol{A}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\boldsymbol{H}(\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\sigma_v^2)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}]\boldsymbol{A}^{\mathrm{T}}+\sigma_w^2\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}} \tag{6.7.139}$$

它可用迭代法求解.

由定理6.6.12有稳态最优白噪声反卷积估值器

$$\hat{w}(t\mid t+N)=L_N^w(q^{-1})\varepsilon(t+N) \tag{6.7.140}$$

$$L_N^w(q^{-1})=\sum_{k=0}^{N}M(k)q^{k-N},\quad N\geqslant 0 \tag{6.7.141}$$

$$L_N^w(q^{-1})=0,\quad N<0 \tag{6.7.142}$$

$$M(0)=0 \tag{6.7.143}$$

$$M(k)=\sigma_w^2\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\Psi}_p^{\mathrm{T}(k-1)}\boldsymbol{H}^{\mathrm{T}}Q_\varepsilon^{-1},\quad k\geqslant 1 \tag{6.7.144}$$

$$Q_\varepsilon=\boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^{\mathrm{T}}+\sigma_v^2 \tag{6.7.145}$$

上述推导引出如下定理.

【定理6.7.11】 多传感器白噪声反卷积系统(6.7.126)有渐近稳定的全局最优加权观测融合白噪声Wiener反卷积估值器

$$\psi(q^{-1})\hat{w}_0(t\mid t+N)=K_N^w(q^{-1})y(t+N) \tag{6.7.146}$$

$$K_N^w(q^{-1})=L_N^w(q^{-1})\Lambda(q^{-1}) \tag{6.7.147}$$

且最优融合估值误差方差为

$$P_0^w(N)=\sigma_w^2-\sum_{k=0}^{N}M^2(k)Q_\varepsilon,\quad N\geqslant 0 \tag{6.7.148}$$

$$P_0^w(N)=\sigma_w^2,\quad N<0 \tag{6.7.149}$$

证明 由(6.7.134)有 $\varepsilon(t+N)=\psi^{-1}(q^{-1})\Lambda(q^{-1})y(t+N)$，由定理6.6.12有(6.7.140)～(6.7.145).将 $\varepsilon(t+N)$ 表达式代入(6.7.140)得(6.7.146)和(6.7.147).因 $\boldsymbol{\Psi}_p$ 为稳定矩阵，则 $\psi(q^{-1})$ 为稳定的多项式，故(6.7.146)是渐近稳定的.由(6.6.103)和(6.6.104)得(6.7.148)和(6.7.149). □

6.7.7 应用于设计多传感器观测融合 Wiener 反卷积滤波器

考虑多传感器反卷积系统

$$y_i(t) = \frac{B(q^{-1})}{P(q^{-1})}s(t) + v_i(t), \quad i = 1,\cdots,L \tag{6.7.150}$$

$$A(q^{-1})s(t) = C(q^{-1})w(t) \tag{6.7.151}$$

其中 $y_i(t)$ 是第 i 传感的输出(观测),$s(t)$ 是待估输入信号,$v_i(t)$ 是观测噪声,$A(q^{-1})$,$\cdots$,$P(q^{-1})$ 是形如 $X(q^{-1}) = x_0 + x_1q^{-1} + \cdots + x_{n_x}q^{-n_x}$ 的单位滞后算子 q^{-1} 的多项式,$a_0 = 1$,$b_0 = 0$,$c_0 = 0$,$n_a \geqslant n_b$,$w(t)$ 和 $v_i(t)$ 是零均值、方差各为 σ_w^2 和 σ_{vi}^2 的相互独立白噪声,$(A(q^{-1}),C(q^{-1}))$ 互质,$(P(q^{-1}),B(q^{-1}))$ 互质. 问题是基于 $y_i(t)$ 求 $s(t)$ 的局部反卷积滤波器 $\hat{s}_i(t \mid t+N)(i = 1,\cdots,L)$ 和全局最优加权观测融合反卷积滤波 $\hat{s}_0(t \mid t+N)(N = 0, N > 0$ 或 $N < 0)$.

【定理 6.7.12】 多传感器反卷积系统(6.7.150)和(6.7.151)当 $t_0 = -\infty$ 时的全局最优观测融合 Wiener 反卷积滤波器 $\hat{s}(t \mid t+N)$ 在数值上恒同于如下加权观测融合反卷积系统的稳态最优 Wiener 反卷积滤波器:

$$y(t) = \frac{B(q^{-1})}{P(q^{-1})}s(t) + v(t) \tag{6.7.152}$$

$$A(q^{-1})s(t) = C(q^{-1})w(t) \tag{6.7.153}$$

其中 $y(t)$ 和 $v(t)$ 为用方法 Ⅱ 的加权融合观测和观测噪声,

$$y(t) = \left(\sum_{i=1}^{L}\frac{1}{\sigma_{vi}^2}\right)^{-1}\sum_{i=1}^{L}\frac{1}{\sigma_{vi}^2}y_i(t) \tag{6.7.154}$$

$$v(t) = \left(\sum_{i=1}^{L}\frac{1}{\sigma_{vi}^2}\right)^{-1}\sum_{i=1}^{L}\frac{1}{\sigma_{vi}^2}v_i(t) \tag{6.7.155}$$

且融合观测噪声 $v(t)$ 为零均值、方差为 σ_v^2 的白噪声,

$$\sigma_v^2 = \left(\sum_{i=1}^{L}\frac{1}{\sigma_{vi}^2}\right)^{-1} \tag{6.7.156}$$

证明 类似于定理 6.7.8 的证明,从略. □

下面应用经典 Kalman 滤波方法解决问题.(6.7.153)有状态空间模型

$$\boldsymbol{\alpha}(t+1) = \boldsymbol{A\alpha}(t) + \boldsymbol{C}w(t) \tag{6.7.157}$$

$$s(t) = \boldsymbol{H}_1\boldsymbol{\alpha}(t) \tag{6.7.158}$$

其中规定 $c_i = 0(i > n_c)$,且

$$\boldsymbol{A} = \begin{bmatrix} -a_1 & & & \\ \vdots & \boldsymbol{I}_{n_a-1} & & \\ -a_{n_a} & 0 & \cdots & 0 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} c_1 \\ \vdots \\ c_{n_a} \end{bmatrix}, \quad \boldsymbol{H}_1 = [1 \quad 0 \cdots 0] \tag{6.7.159}$$

而(6.7.152)有状态模型

$$\boldsymbol{\beta}(t+1) = \boldsymbol{P\beta}(t) + \boldsymbol{B}s(t) \tag{6.7.160}$$

$$y(t) = \boldsymbol{H}_2\boldsymbol{\beta}(t) + v(t) \tag{6.7.161}$$

其中规定 $b_i = 0(i > n_b)$,且

$$\boldsymbol{P}=\begin{bmatrix}-p_1 & & & \\ \vdots & \boldsymbol{I}_{n_p-1} & & \\ -p_{n_p} & 0 & \cdots & 0\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}b_1\\ \vdots\\ b_{n_a}\end{bmatrix},\quad \boldsymbol{H}_2=[1\quad 0\cdots 0] \tag{6.7.162}$$

于是有增广系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma}w(t) \tag{6.7.163}$$

$$y(t)=\boldsymbol{Hx}(t)+v(t) \tag{6.7.164}$$

其中增广状态 $\boldsymbol{x}(t)$ 的维数 $n=n_a+n_p$，且

$$\boldsymbol{x}(t)=\begin{bmatrix}\boldsymbol{\alpha}(t)\\ \boldsymbol{\beta}(t)\end{bmatrix},\quad \boldsymbol{\Phi}=\begin{bmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{BH}_1 & \boldsymbol{P}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\boldsymbol{C}\\ \boldsymbol{0}\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{0}\quad \boldsymbol{H}_2] \tag{6.7.165}$$

注意$(P(q^{-1}),B(q^{-1}))$互质和$(A(q^{-1}),C(q^{-1}))$互质的假设保证了增广系统完全可观、完全可控.由(6.7.158)有

$$\hat{s}_0(t\mid t+N)=\boldsymbol{H}_0\hat{\boldsymbol{x}}(t\mid t+N) \tag{6.7.166}$$

$$\boldsymbol{H}_0=[1\quad 0\cdots 0] \tag{6.7.167}$$

【引理6.7.1】[15,35] 增广系统(6.7.163)和(6.7.164)有渐近稳定的Wiener状态滤波器

$$\psi(q^{-1})\hat{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})y(t+N) \tag{6.7.168}$$

其中

$$\psi(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p) \tag{6.7.169}$$

$$\boldsymbol{\Psi}_p=\boldsymbol{\Phi}-\boldsymbol{K}_p\boldsymbol{H} \tag{6.7.170}$$

$$\boldsymbol{K}_p=\boldsymbol{\Phi\Sigma H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2]^{-1} \tag{6.7.171}$$

$\boldsymbol{\Sigma}$ 满足 Riccati 方程

$$\boldsymbol{\Sigma}=\boldsymbol{\Phi}[\boldsymbol{\Sigma}-\boldsymbol{\Sigma H}(\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2)^{-1}\boldsymbol{H\Sigma}]\boldsymbol{\Phi}^{\mathrm{T}}+\sigma_w^2\boldsymbol{\Gamma\Gamma}^{\mathrm{T}} \tag{6.7.172}$$

$$\boldsymbol{K}_N(q^{-1})=\sum_{i=0}^{n-1}\boldsymbol{\Omega}_i[J_{i-N}(q^{-1})-\Lambda_N^i(q^{-1})-\Delta_N^i(q^{-1})] \tag{6.7.173}$$

$$\Lambda_N^i(q^{-1})=\sum_{j=0}^{i-1}\boldsymbol{H\Phi}^{i-1-j}\boldsymbol{\Gamma}L_{N-j}^w(q^{-1})\Lambda(q^{-1});\ \Lambda_N^0(q^{-1})=0 \tag{6.7.174}$$

$$\Delta_N^i(q^{-1})=L_{N-i}^v(q^{-1})\Lambda(q^{-1}) \tag{6.7.175}$$

$$\Lambda(q^{-1})=\psi(q^{-1})-\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_pq^{-1}$$

$$J_i(q^{-1})=\begin{cases}\boldsymbol{H\Phi}^{i-1}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p, & i>0\\ \psi(q^{-1})q^i, & i\leqslant 0\end{cases} \tag{6.7.176}$$

$$\boldsymbol{\Omega}=[\boldsymbol{H}^{\mathrm{T}},(\boldsymbol{H\Phi})^{\mathrm{T}},\cdots,(\boldsymbol{H\Phi}^{n-1})^{\mathrm{T}}]^{\mathrm{T}} \tag{6.7.177}$$

$$\boldsymbol{\Omega}^{-1}=[\boldsymbol{\Omega}_0,\boldsymbol{\Omega}_1,\cdots,\boldsymbol{\Omega}_{n-1}] \tag{6.7.178}$$

其中 $n=n_a+n_p$ 为增广系统的状态维数.

$$L_N^\theta(q^{-1})=\sum_{i=1}^{N}M_\theta(i)q^{i-N},\quad \theta=w,v,N\geqslant 0 \tag{6.7.179}$$

$$L_N^\theta(q^{-1})=0,\quad N<0$$

$$M_w(i)=\sigma_w^2\boldsymbol{\Gamma\Psi}_p^{\mathrm{T}(i-1)}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2]^{-1},\quad i>0$$

$$M_v(i)=-\sigma_v^2\boldsymbol{K}^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Psi}_p^{\mathrm{T}(i-1)}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2]^{-1},\quad i>0$$

$$M_w(0)=0,\quad M_v(0)=\sigma_v^2[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2]^{-1} \tag{6.7.180}$$

且有估值误差 $\tilde{\boldsymbol{x}}(t\mid t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t\mid t+N)$ 稳态方差阵 $\boldsymbol{P}(N)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t+N)\tilde{\boldsymbol{x}}^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}(N)=\boldsymbol{\Sigma}-\sum_{i=0}^{N}\boldsymbol{K}_i[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2]\boldsymbol{K}_i^{\mathrm{T}},\quad N\geqslant 0 \tag{6.7.181}$$

$$\boldsymbol{K}_i=\boldsymbol{\Sigma}\boldsymbol{\Psi}_p^{\mathrm{T}i}\boldsymbol{H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2]^{-1},\quad i\geqslant 0 \tag{6.7.182}$$

其中定义 $\boldsymbol{K}=\boldsymbol{K}_0=\boldsymbol{\Sigma H}^{\mathrm{T}}[\boldsymbol{H\Sigma H}^{\mathrm{T}}+\sigma_v^2]^{-1}$. 当 $N=-1$ 时有 $\boldsymbol{P}(-1)=\boldsymbol{\Sigma}$，且当 $N<-1$ 时有

$$\boldsymbol{P}(N)=\boldsymbol{\Phi}^{-N-1}\boldsymbol{\Sigma}\boldsymbol{\Phi}^{\mathrm{T}(-N-1)}+\sigma_w^2\sum_{j=2}^{N}\boldsymbol{\Phi}^{-N-j}\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Phi}^{\mathrm{T}(-N-j)} \tag{6.7.183}$$

【定理 6.7.13】 多传感器反卷积系统 (6.7.150) 和 (6.7.151) 有全局最优加权观测融合 Wiener 反卷积滤波器 $\hat{s}_0(t\mid t+N)(N=0,N>0$ 或 $N<0)$ 为

$$\psi(q^{-1})\hat{s}_0(t\mid t+N)=K_N^0(q^{-1})y(t+N) \tag{6.7.184}$$

$$K_N^0(q^{-1})=\boldsymbol{H}_0\boldsymbol{K}_N(q^{-1}) \tag{6.7.185}$$

其中 $\boldsymbol{K}_N(q^{-1})$ 由 (6.7.173) ~ (6.7.180) 定义，且滤波误差 $\tilde{s}_0(t\mid t+N)=s(t)-\hat{s}_0(t\mid t+N)$ 稳态方差 $P_{s0}(N)=\mathrm{E}[\tilde{s}_0^2(t\mid t+N)]$ 为

$$P_{s0}(N)=H_0\boldsymbol{P}(N)H_0^{\mathrm{T}} \tag{6.7.186}$$

其中 $\boldsymbol{P}(N)$ 由 (6.7.181) 和 (6.7.183) 定义.

证明 由 (6.7.166) 和 (6.7.168) 得 (6.7.184) 和 (6.7.185). 由 (6.7.166) 有 $\tilde{s}_0(t\mid t+N)=H_0\tilde{\boldsymbol{x}}(t\mid t+N)$，于是有 (6.7.186). □

【例 6.7.1】 考虑三传感器系统

$$A(q^{-1})s(t)=C(q^{-1})w(t) \tag{6.7.187}$$

$$y_i(t)=s(t)+v_i(t),\quad i=1,2,3 \tag{6.7.188}$$

$$A(q^{-1})=1-1.8q^{-1}+1.05q^{-2}-0.2q^{-3},\quad C(q^{-1})=q^{-1}+0.3q^{-2} \tag{6.7.189}$$

其中 $w(t)$ 和 $v_i(t)$ 是零均值、方差各为 $\sigma_w^2=0.45$，$\sigma_{v1}^2=2$，$\sigma_{v2}^2=6$ 和 $\sigma_{v3}^2=13$ 的相互独立高斯白噪声. 问题是求 $s(t)$ 的局部滤波器 $\hat{s}_i(t\mid t)$ 和加权观测融合滤波器 $\hat{s}_0(t\mid t)$.

取 $N=0$，应用定理 6.7.10 可求得 $s(t)$ 的加权观测融合滤波器 $\hat{s}_0(t\mid t)$ 为

$$\psi(q^{-1})\hat{s}_0(t\mid t)=K_0^0(q^{-1})y(t) \tag{6.7.190}$$

$$\psi(q^{-1})=1-0.914\,8q^{-1}+0.433\,3q^{-2}-0.070\,6q^{-3} \tag{6.7.191}$$

$$K_0^0(q^{-1})=0.646\,8-0.279\,1q^{-1}+0.062\,5q^{-3} \tag{6.7.192}$$

其中 $y(t)$ 为由 (6.7.128) 定义的加权融合观测.

取 $N=0$，套用定理 6.7.10 可求得局部滤波器 $\hat{s}_i(t\mid t)$，$i=1,2,3$. 应用 (6.7.120) 可求得融合和局部滤波误差方差 $P_{si}(0)=\mathrm{E}[\hat{s}_i^2(t\mid t)]$，$\tilde{s}_i(t\mid t)=s(t)-\hat{s}_i(t\mid t)$，$i=1,2,3$，分别为

$$P_{s0}(0)=0.869\,9,\quad P_{s1}(0)=1.192\,9,\quad P_{s2}(0)=2.698\,5,\quad P_{s3}(0)=4.511\,3 \tag{6.7.193}$$

可从理论上看到加权观测融合滤波精度高于每个局部滤波精度，即

$$P_{s0}(0) < P_{si}(0), \quad i = 1,2,3 \tag{6.7.194}$$

仿真结果如图6.7.1 ~ 图6.7.4所示，其中实线代表真实值 $s(t)$，虚线代表滤波估值 $\hat{s}(t \mid t)$，也可直观上看到精度关系(6.7.194)成立.

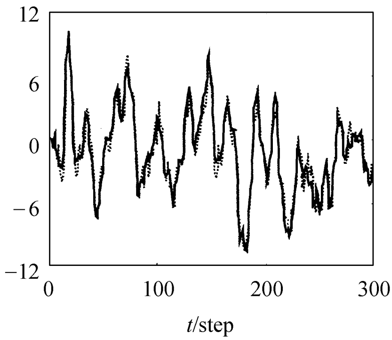


图6.7.1　信号 $s(t)$ 和观测融合Wiener滤波器 $\hat{s}_0(t \mid t)$

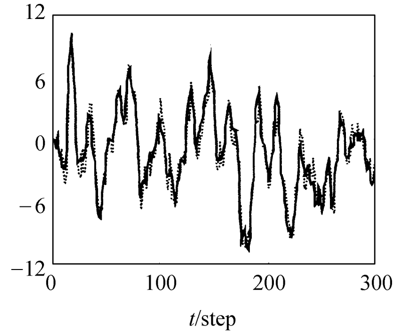


图6.7.2　信号 $s(t)$ 和传感器1的局部Wiener滤波器 $\hat{s}_1(t \mid t)$

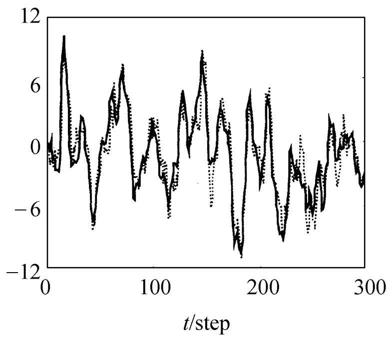


图6.7.3　信号 $s(t)$ 和传感器2的局部Wiener滤波器 $\hat{s}_2(t \mid t)$

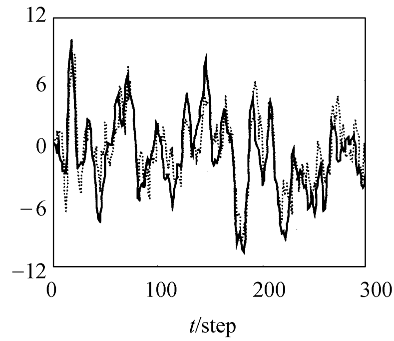


图6.7.4　信号 $s(t)$ 和传感器3的局部Wiener滤波器 $\hat{s}_3(t \mid t)$

上述加权观测融合滤波器 $\hat{s}_0(t \mid t)$ 的全局最优性可验证如下：

集中式观测融合系统有状态空间模型

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{C}w(t) \tag{6.7.195}$$

$$y_c(t) = \boldsymbol{H}_c x(t) + v_c(t) \tag{6.7.196}$$

$$s(t) = \boldsymbol{H}x(t) \tag{6.7.197}$$

$$\boldsymbol{y}_c(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix}, \quad \boldsymbol{H}_c = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H} \\ \boldsymbol{H} \end{bmatrix}, \quad \boldsymbol{v}_c(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \end{bmatrix}, \quad \boldsymbol{H} = [1 \quad 0 \quad 0],$$

$$\boldsymbol{A} = \begin{bmatrix} 1.8 & 1 & 0 \\ -1.05 & 0 & 1 \\ 0.2 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 1 \\ 0.3 \\ 0 \end{bmatrix} \tag{6.7.198}$$

用解Riccati方程的经典Kalman滤波方法可求得集中式稳态全局最优Kalman滤波器

$\hat{\boldsymbol{x}}_c(t \mid t)$，进而由(6.7.197)引出的关系 $\hat{s}_c(t \mid t) = \boldsymbol{H}\hat{\boldsymbol{x}}_c(t \mid t)$ 可求得集中式信号 Wiener 滤波器 $\hat{s}_c(t \mid t)$ 为

$$\psi_c(q^{-1})\hat{s}_c(t \mid t) = \boldsymbol{K}_0^c(q^{-1})\boldsymbol{y}_c(t) \tag{6.7.199}$$

其中

$$\psi_c(q^{-1}) = 1 - 0.914\,8q^{-1} + 0.433\,3q^{-2} - 0.070\,6q^{-3} \tag{6.7.200}$$

$$\boldsymbol{K}_0^c(q^{-1}) = \begin{bmatrix} 0.434\,9 \\ 0.145\,0 \\ 0.066\,9 \end{bmatrix}^{\mathrm{T}} - \begin{bmatrix} 0.187\,7 \\ 0.067\,6 \\ 0.028\,9 \end{bmatrix}^{\mathrm{T}} q^{-1} + \begin{bmatrix} 0.042\,0 \\ 0.014\,0 \\ 0.006\,5 \end{bmatrix}^{\mathrm{T}} q^{-2} \tag{6.7.201}$$

且可求得集中式全局最优滤波误差方差为

$$P_{sc}(0) = 0.869\,9 \tag{6.7.202}$$

由(6.7.193)有

$$P_{sc}(0) = P_{s0}(0) \tag{6.7.203}$$

即加权观测融合滤波 $\hat{s}_0(t \mid t)$ 具有全局最优性.

【例 6.7.2】 考虑四传感器白噪声反卷积系统

$$y_i(t) = \frac{q^{-1}(1 - 0.8q^{-1})}{1 - q^{-1} + 0.24q^{-2}}w(t) + v_i(t), \quad i = 1,2,3,4 \tag{6.7.204}$$

其中 $w(t) = b(t)g(t)$ 是 Bernoulli – Gaussian 白噪声，$b(t)$ 是取值 1 或 0 的 Bernoulli 白噪声，取值概率为 $P(b(t) = 1) = \lambda$，$P(b(t) = 0) = 1 - \lambda$，$g(t)$ 是零均值、方差为 σ_g^2 独立于 $b(t)$ 的 Gaussian 白噪声，于是 $w(t)$ 有零均值，且方差 $\sigma_w^2 = \lambda\sigma_g^2$. 设 $v_i(t)$ 是零均值、方差为 σ_{vi}^2 独立于 $w(t)$ 的 Gaussian 白噪声. 取 $\lambda = 0.25$，$\sigma_g^2 = 1$，$\sigma_{v1}^2 = \sigma_{v2}^2 = \sigma_{v3}^2 = \sigma_{v4}^2 = 0.05$，问题是求全局最优加权观测融合白噪声 Wiener 反卷积平滑器 $\hat{w}_0(t \mid t+1)$ 和局部平滑器 $\hat{w}_i(t \mid t+1)(i = 1,2,3,4)$.

应用定理 6.7.11 可求得最优加权观测融合和局部平滑误差方差各为

$$P_0^w(1) = 0.012\,5, \quad P_1^w(1) = P_2^w(1) = P_3^w(1) = P_4^w(1) = 0.043\,3 \tag{6.7.205}$$

这表明融合平滑估计精度高于每个局部平滑估计精度. 仿真结果如图 6.7.5 和图 6.7.6 所示. 在图 6.7.5 中实线端点纵坐标为真实值 $w(t)$，实圆点纵坐标为平滑估值 $\hat{w}_i(t \mid t+1)(i = 0,1,2,3,4)$. 可直观看到加权观测融合估计精度明显高于每个局部平滑估计精度. 在图 6.7.6 中给出了各平滑器的累积平滑误差平方比较曲线，其中实的曲线为观测融合平滑误差平方累积曲线，可进一步说明融合估计精度明显高于每个局部估计精度.

【例 6.7.3】 考虑四传感器反卷积系统

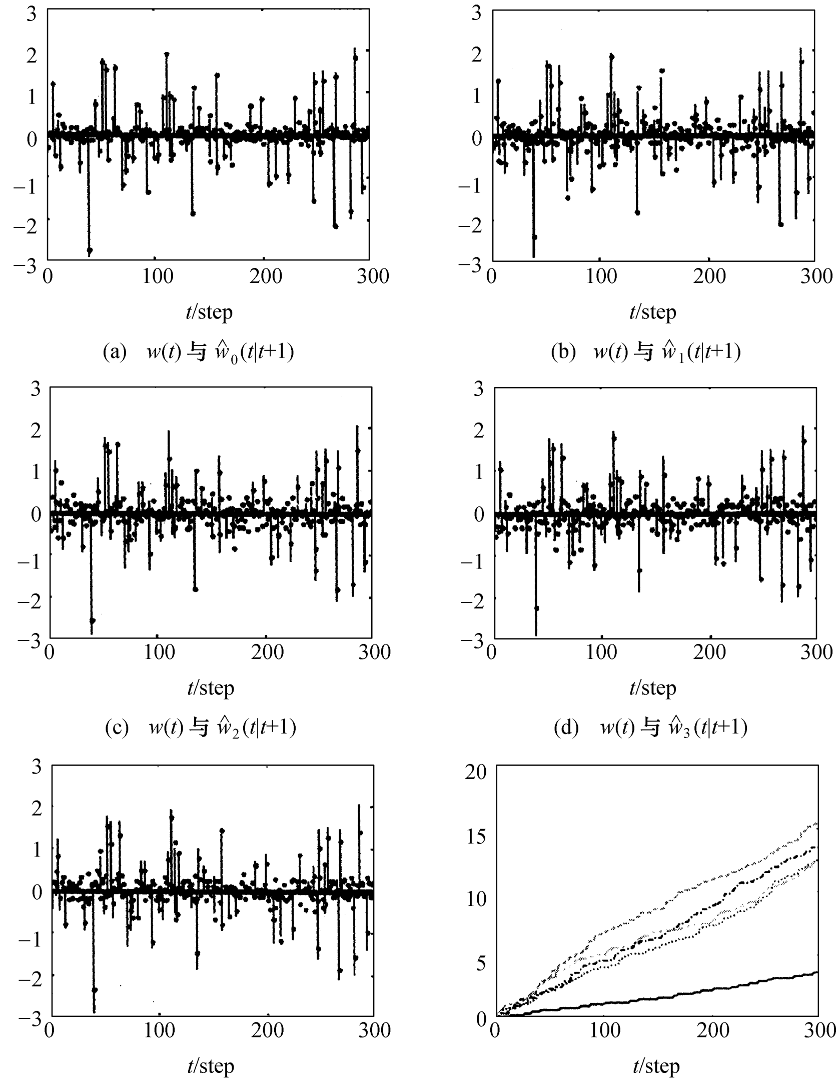
$$y_i(t) = \frac{q^{-1}}{1 - 1.3q^{-1} + 0.4q^{-2}}s(t) + v_i(t), \quad i = 1,2,3,4 \tag{6.7.206}$$

$$(1 - 0.9q^{-1})s(t) = w(t-1) \tag{6.7.207}$$

其中 $w(t)$ 和 $v_i(t)$ 是零均值、方差各为 $\sigma_w^2 = 0.45$，$\sigma_{v1}^2 = 2$，$\sigma_{v2}^2 = 6$，$\sigma_{v3}^2 = 13$，$\sigma_{v4}^2 = 10$ 的相互独立的高斯白噪声. 问题是求全局最优加权观测融合 Wiener 反卷积预报器 $\hat{s}_0(t \mid t-2)$.

可得加权观测融合系统增广状态空间模型

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}\boldsymbol{x}(t) + \boldsymbol{\Gamma}w(t) \tag{6.7.208}$$



(a) $w(t)$ 与 $\hat{w}_0(t|t+1)$

(b) $w(t)$ 与 $\hat{w}_1(t|t+1)$

(c) $w(t)$ 与 $\hat{w}_2(t|t+1)$

(d) $w(t)$ 与 $\hat{w}_3(t|t+1)$

(e) $w(t)$ 与 $\hat{w}_4(t|t+1)$

(f) **累积平滑误差平方曲线**

图 6.7.5　Bernoulli – Gaussian 白噪声 $w(t)$，局部 Wiener 平滑器 $\hat{w}_i(t \mid t+1)$ 和观测融合 Wiener 平滑器 $\hat{w}_0(t \mid t+1)$ 及它们的累积平滑误差平方曲线

$$y(t) = \boldsymbol{H}x(t) + v(t) \tag{6.7.209}$$

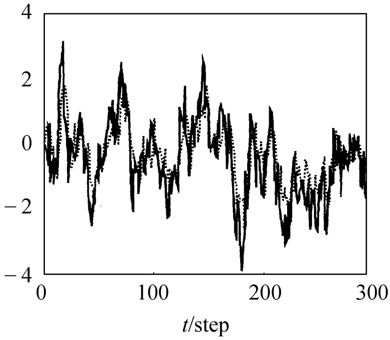
其中 $y(t)$ 和 $v(t)$ 是由(6.7.154)和(6.7.155)计算的融合观测和融合观测噪声，且

$$\boldsymbol{\Phi} = \begin{bmatrix} 0.9 & 0 & 0 \\ 1 & 1.3 & 1 \\ 0 & -0.4 & 0 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{H} = [0 \quad 1 \quad 0] \tag{6.7.210}$$

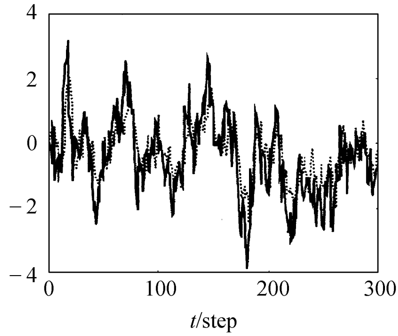
解 Riccati 方程(6.7.172)可求得

$$
\boldsymbol{\Sigma} = \begin{bmatrix} 1.0680 & 1.0891 & -0.1180 \\ 1.0891 & 2.7185 & -0.4966 \\ -0.1180 & -0.4966 & 0.1309 \end{bmatrix}, \boldsymbol{K}_p = \begin{bmatrix} 0.2521 \\ 1.0612 \\ -0.2796 \end{bmatrix}, \boldsymbol{\Psi}_p = \begin{bmatrix} 0.9 & -0.2521 & 0 \\ 1 & 0.2388 & 1 \\ 0 & -0.1204 & 0 \end{bmatrix} \tag{6.7.211}
$$

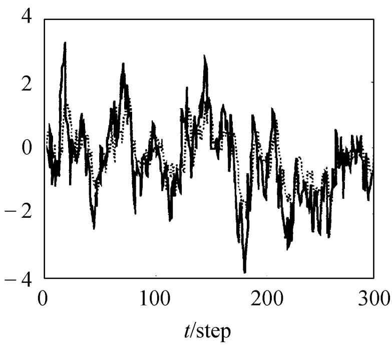
应用定理 6.7.13 可求得全局最优加权观测融合 Wiener 反卷积预报器 $\hat{s}_0(t \mid t-2)$ 为



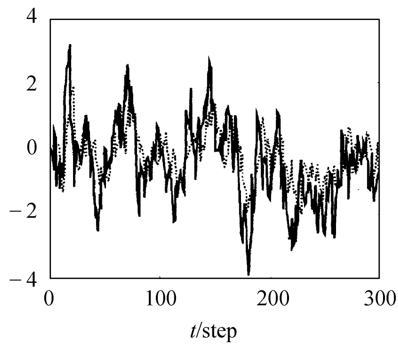
(a) 信号 $s(t)$ 和观测融合信号 Wiener 预报器 $\hat{s}_0(t|t-2)$



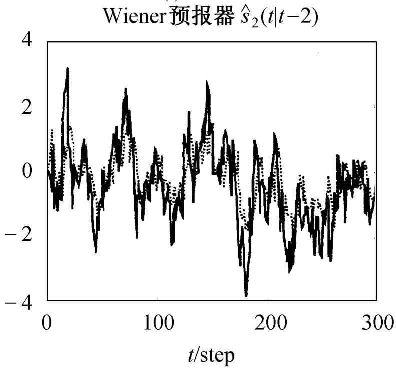
(b) 信号 $s(t)$ 和传感器 1 的局部 Wiener 预报器 $\hat{s}_1(t|t-2)$



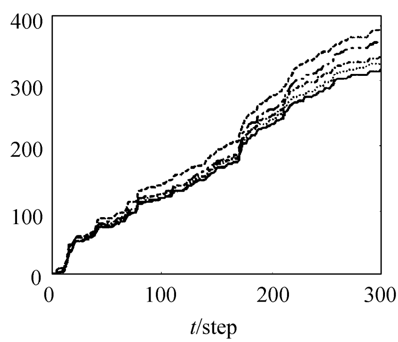
(c) 信号 $s(t)$ 和传感器 2 的局部 Wiener 预报器 $\hat{s}_2(t|t-2)$



(d) 信号 $s(t)$ 和传感器 3 的局部 Wiener 预报器 $\hat{s}_3(t|t-2)$



(e) 信号 $s(t)$ 和传感器 4 的局部 Wiener 预报器 $\hat{s}_4(t|t-2)$



(f) 预报误差平方和曲线

——— :观测融合预报误差平方和曲线
······ :第一个传感器预报误差平方和曲线
—·—·— :第二个传感器预报误差平方和曲线
- - - - - - :第三个传感器预报误差平方和曲线
—··—··- :第四个传感器预报误差平方和曲线

图 6.7.6 信号 $s(t)$,局部和加权观测融合 Wiener 反卷积预报器 $\hat{s}_0(t \mid t-2)$,$\hat{s}_i(t \mid t-2)$ 及累积预报误差平方曲线

$$\psi(q^{-1})\hat{s}_0(t \mid t-2) = K_{-2}^0(q^{-1})y(t-2) \tag{6.7.212}$$

$$\psi(q^{-1}) = 1 - 1.138\,8q^{-1} + 0.587\,3q^{-2} - 0.108\,3q^{-3} \tag{6.7.213}$$

$$K_{-2}^0(q^{-1}) = 0.226\,9 - 0.294\,9q^{-1} + 0.090\,7q^{-2} \tag{6.7.214}$$

类似可求得局部 Wiener 反卷积预报器 $\hat{s}_i(t \mid t-2)$，且可求得融合预报和局部预报误差方差各为

$$P_{s0}(-2) = 1.315\,1,\quad P_{s1}(-2) = 1.356\,4,\quad P_{s2}(-2) = 1.454\,1,$$
$$P_{s3}(-2) = 1.533\,5,\quad P_{s4}(-2) = 1.494\,7 \tag{6.7.215}$$

这引出精度关系

$$P_{s0}(-2) < P_{si}(-2),\quad i = 1,2,3,4 \tag{6.7.216}$$

即融合预报精度高于局部预报精度.仿真结果如图 6.7.6 所示，其中在图 6.7.6(a) ~ (e)中实线代表真实值 $s(t)$，虚线代表预报值 $\hat{s}_i(t \mid t-2)(i = 0,1,2,3,4)$.图 6.7.6(f) 中各曲线为累积预报误差平方比较曲线，也可看到观测融合预报精度最高.

6.8 多通道 ARMA 信号分布式信息融合 Wiener 滤波器

ARMA 信号滤波问题广泛出现在通讯、信号处理、目标跟踪、GPS 定位、语音增强等应用问题中.新近文献[32]用现代时间序列分析方法，基于 ARMA 新息模型提出了两传感器单通道 ARMA 信号信息融合 Wiener 滤波器.本节用 Kalman 滤波方法提出多通道 ARMA 信号多传感器分布式信息融合 Wiener 滤波器.它是次优的，即它是非全局最优的.

考虑带白色观测噪声的多传感器多通道 ARMA 信号 $s(t)$

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{6.8.1}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{s}(t) + \boldsymbol{v}_i(t),\quad i = 1,\cdots,L \tag{6.8.2}$$

其中 $\boldsymbol{y}_i(t) \in R^m$ 为第 i 个传感器的观测信号，待估 ARMA 信号 $\boldsymbol{s}(t) \in R^m$，白色观测噪声 $\boldsymbol{v}_i(t) \in R^m$，$\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}_i(t)(i = 1,\cdots,L)$ 为零均值、方差阵各为 $\boldsymbol{Q}$ 和 $\boldsymbol{R}_i$ 的相互独立白噪声：

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}_i(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{w}^{\mathrm{T}}(j) & \boldsymbol{v}_i^{\mathrm{T}}(j)\end{bmatrix}\right\} = \begin{bmatrix}\boldsymbol{Q} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{R}_i\end{bmatrix}\delta_{tj},$$
$$\mathrm{E}[\boldsymbol{v}_i(t)\boldsymbol{v}_j^{\mathrm{T}}(k)] = \boldsymbol{0},\quad \forall\, t,k,j \tag{6.8.3}$$

其中 E 为均值号，T 为转置号，$\delta_{tt} = 1$，$\delta_{tj} = 0(t \neq j)$. q^{-1} 为单位滞后算子，$\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{C}(q^{-1})$ 为多项式矩阵，

$$\boldsymbol{A}(q^{-1}) = \boldsymbol{I}_m + \boldsymbol{A}_1 q^{-1} + \cdots + \boldsymbol{A}_{n_a} q^{-n_a},$$
$$\boldsymbol{C}(q^{-1}) = \boldsymbol{C}_1 q^{-1} + \cdots + \boldsymbol{C}_{n_c} q^{-n_c},\quad n_a \geqslant n_c \tag{6.8.4}$$

且设 $(\boldsymbol{A}(q^{-1}),\boldsymbol{C}(q^{-1}))$ 左素.问题是基于观测 $(\boldsymbol{y}_i(t+N),\boldsymbol{y}_i(t+N-1),\cdots)$ 求信号 $\boldsymbol{s}(t)$ 的局部最优 Wiener 滤波器 $\hat{\boldsymbol{s}}_i(t \mid t+N)$，$i = 1,\cdots,L$，并求它们的分布式最优融合 Wiener 滤波器 $\hat{\boldsymbol{s}}_0(t \mid t+N)$.对 $N = 0$，$N > 0$，$N < 0$，各称其为滤波器、平滑器和预报器.

为了应用 Kalman 滤波方法，首先将(6.8.1)和(6.8.2)写成状态空间模型

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}\boldsymbol{x}(t) + \boldsymbol{\Gamma}\boldsymbol{w}(t) \tag{6.8.5}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}_i(t),\quad i = 1,\cdots,L \tag{6.8.6}$$

$$\boldsymbol{s}(t)=\boldsymbol{H}\boldsymbol{x}(t) \tag{6.8.7}$$

其中规定 $\boldsymbol{C}_i=\boldsymbol{0}(i>n_c)$，且定义

$$\boldsymbol{\Phi}=\begin{bmatrix}-\boldsymbol{A}_1 & & & \\ \vdots & & \boldsymbol{I}_{m(n_a-1)} & \\ -\boldsymbol{A}_{n_a} & \boldsymbol{0} & \cdots & \boldsymbol{0}\end{bmatrix},\quad \boldsymbol{\Gamma}=\begin{bmatrix}\boldsymbol{C}_1\\ \vdots \\ \boldsymbol{C}_{n_a}\end{bmatrix},\quad \boldsymbol{H}=[\boldsymbol{I}_m \quad \boldsymbol{0}\cdots\boldsymbol{0}] \tag{6.8.8}$$

【定理 6.8.1】 第 i 传感器子系统(6.8.5)～(6.8.7)有局部稳态最优 Wiener 滤波器 $\hat{s}_i(t\mid t+N)$ 为

$$\psi(q^{-1})\hat{\boldsymbol{s}}_i(t\mid t+N)=\boldsymbol{\Psi}_N^{(i)}(q^{-1})\boldsymbol{y}_i(t+N),\quad i=1,\cdots,L \tag{6.8.9}$$

其中定义 $n=mn_a$，且

$$\boldsymbol{\Psi}_N^{(i)}(q^{-1})=\boldsymbol{H}[\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_{pi})\boldsymbol{K}_{pi}q^{-1-N}+\boldsymbol{K}_N^{(i)}(q^{-1})\boldsymbol{\Lambda}_i(q^{-1})],\quad N\geqslant 0 \tag{6.8.10}$$

$$\boldsymbol{\Psi}_N^{(i)}(q^{-1})=\boldsymbol{H}\boldsymbol{\Phi}^{-N-1}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_{pi})\boldsymbol{K}_{pi},\quad N\leqslant -1 \tag{6.8.11}$$

$$\psi_i(q^{-1})=\det(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_{pi}) \tag{6.8.12}$$

$$\boldsymbol{\Lambda}_i(q^{-1})=\psi_i(q^{-1})\boldsymbol{I}_m-\boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_{pi})\boldsymbol{K}_{pi}q^{-1} \tag{6.8.13}$$

$$\boldsymbol{\Psi}_{pi}=\boldsymbol{\Phi}-\boldsymbol{K}_{pi}\boldsymbol{H} \tag{6.8.14}$$

$$\boldsymbol{K}_{pi}=\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{6.8.15}$$

$$\boldsymbol{Q}_{\varepsilon i}=\boldsymbol{H}\boldsymbol{\Sigma}_i\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}_i \tag{6.8.16}$$

$$\boldsymbol{K}_N^{(i)}(q^{-1})=\sum_{j=0}^{N}\boldsymbol{K}_j^{(i)}q^{j-N},\quad N\geqslant 0 \tag{6.8.17}$$

$$\boldsymbol{K}_j^{(i)}=\boldsymbol{\Sigma}_i\boldsymbol{\Psi}_{pi}^{\mathrm{T}j}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1},\quad i=1,\cdots,L \tag{6.8.18}$$

而 $\boldsymbol{\Sigma}$ 满足 Riccati 方程

$$\boldsymbol{\Sigma}_i=\boldsymbol{\Phi}[\boldsymbol{\Sigma}_i-\boldsymbol{\Sigma}_i\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Sigma}_i\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}_i)^{-1}\boldsymbol{H}\boldsymbol{\Sigma}_i]\boldsymbol{\Phi}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} \tag{6.8.19}$$

它可用迭代法求解.

估值误差 $\tilde{\boldsymbol{s}}_i(t\mid t+N)=\boldsymbol{s}(t)-\hat{\boldsymbol{s}}_i(t\mid t+N)$ 的稳态方差阵 $\boldsymbol{P}_{si}(N)=\mathrm{E}[\tilde{\boldsymbol{s}}_i(t\mid t+N)\tilde{\boldsymbol{s}}_i^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}_{si}(N)=\boldsymbol{H}\boldsymbol{\Sigma}_i(N)\boldsymbol{H}^{\mathrm{T}} \tag{6.8.20}$$

其中 $\boldsymbol{\Sigma}_i(N)$ 由下式计算

$$\boldsymbol{\Sigma}_i(N)=\boldsymbol{\Sigma}_i-\sum_{j=0}^{N}\boldsymbol{K}_j^{(i)}\boldsymbol{Q}_{\varepsilon i}\boldsymbol{K}_j^{(i)\mathrm{T}},\quad N\geqslant 0 \tag{6.8.21}$$

$$\boldsymbol{\Sigma}_i(N)=\boldsymbol{\Phi}^{-N-1}\boldsymbol{\Sigma}_i(\boldsymbol{\Phi}^{\mathrm{T}})^{(-N-1)}+\sum_{j=2}^{-N}\boldsymbol{\Phi}^{-N-j}\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Phi}^{-(N-j)T},\quad N<-1 \tag{6.8.22}$$

$$\boldsymbol{\Sigma}_i(-1)=\boldsymbol{\Sigma}_i \tag{6.8.23}$$

证明 见定理 6.7.9 的证明，从略. □

【定理 6.8.2】 多传感器系统(6.8.1)和(6.8.2)局部估计误差 $\tilde{\boldsymbol{s}}(t\mid t+N)$ 与 $\tilde{\boldsymbol{s}}_j(t\mid t+N)(i\neq j,i,j=1,\cdots,L)$ 的稳态协方差阵 $\boldsymbol{P}_{sij}(N)=\mathrm{E}[\tilde{\boldsymbol{s}}_i(t\mid t+N)\tilde{\boldsymbol{s}}_j^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}_{sij}(N)=\boldsymbol{H}\boldsymbol{P}_{ij}(N)\boldsymbol{H}^{\mathrm{T}} \tag{6.8.24}$$

其中估值误差 $\tilde{\boldsymbol{x}}_i(t\mid t+N)$ 与 $\tilde{\boldsymbol{x}}_j(t\mid t+N)$ 的协方差阵 $\boldsymbol{P}_{ij}(N)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t+N)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t$

$|t+N)]$ 当 $N\geqslant 0$ 时为

$$\boldsymbol{P}_{ij}(N)=\boldsymbol{\Sigma}_{ij}-\sum_{r=0}^{N}\boldsymbol{K}_r^{(i)}\boldsymbol{H}\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij}-\sum_{s=0}^{N}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{s\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{K}_s^{(j)\mathrm{T}}+$$
$$\sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{K}_r^{(i)}\boldsymbol{E}_{ij}(r,s)\boldsymbol{K}_s^{(j)\mathrm{T}},\quad N\geqslant 0 \tag{6.8.25}$$

其中 $\boldsymbol{E}_{ij}(r,s)=\mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)]$ 为

$$\boldsymbol{E}_{ij}(r,s)=\boldsymbol{H}\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{s\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}+\sum_{k=1}^{\min(r,s)}\boldsymbol{H}\boldsymbol{\Psi}_{pi}^{r-k}\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Psi}_{pj}^{(s-k)\mathrm{T}}\boldsymbol{H}^{\mathrm{T}} \tag{6.8.26}$$

其中当 $\min(r,s)=0$ 时应置上式中的第二项为零.

$\boldsymbol{\Sigma}_{ij}=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t|t-1)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t|t-1)]$ 满足 Lyapunov 方程

$$\boldsymbol{\Sigma}_{ij}=\boldsymbol{\Psi}_{pi}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}}+\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} \tag{6.8.27}$$

它可用迭代法求解.

当 $N=-1$ 时记

$$\boldsymbol{P}_{ij}(-1)=\boldsymbol{\Sigma}_{ij} \tag{6.8.28}$$

当 $N<-1$ 时有

$$\boldsymbol{P}_{ij}(N)=\boldsymbol{\Phi}^{-N-1}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Phi}^{(-N-1)\mathrm{T}}+\sum_{k=2}^{-N}\boldsymbol{\Phi}^{-N-k}\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Phi}^{-(N-k)\mathrm{T}} \tag{6.8.29}$$

证明　在定理6.6.7中的(6.6.71)~(6.6.75)中置 $\boldsymbol{S}_i=\boldsymbol{0}$，$\boldsymbol{R}_{ij}=\boldsymbol{0}$ 得(6.8.25)和(6.8.26)，由推论6.6.1的(6.6.49)得(6.8.27)，由 $\tilde{\boldsymbol{s}}_i(t|t+N)=\boldsymbol{H}\tilde{\boldsymbol{x}}_i(t|t+N)$，$\tilde{\boldsymbol{s}}_j(t|t+N)=\boldsymbol{H}\tilde{\boldsymbol{x}}_j(t|t+N)$ 得(6.8.24). 证毕. □

注意，定理6.8.1和定理6.8.2的基本原理是将信号估计问题转化为状态估计问题. 下面给出另一方法，它将信号估计问题转化为白噪声估计和观测预报问题.

【定理6.8.3】　第 i 传感器子系统(6.8.1)和(6.8.2)有渐近稳定的局部稳态最优 Wiener 滤波器

$$\psi(q^{-1})\hat{\boldsymbol{s}}_i(t|t+N)=\boldsymbol{\Psi}_N^{(i)}(q^{-1})\boldsymbol{y}_i(t+N),\quad i=1,\cdots,L \tag{6.8.30}$$

其中 $N=0$，$N>0$ 或 $N<0$，且定义

$$\boldsymbol{\Psi}_N^{(i)}(q^{-1})=\boldsymbol{J}_{-N}^{(i)}(q^{-1})-\boldsymbol{L}_N^{vi}(q^{-1})\boldsymbol{\Lambda}_i(q^{-1}) \tag{6.8.31}$$

其中定义

$$\boldsymbol{J}_N^{(i)}(q^{-1})=\begin{cases}\boldsymbol{H}\boldsymbol{\Phi}^{N-1}\mathrm{adj}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Psi}_{pi})\boldsymbol{K}_{pi}, & N>0\\ \psi_i(q^{-1})\boldsymbol{I}_m q^{N}, & N\leqslant 0\end{cases} \tag{6.8.32}$$

$$\boldsymbol{L}_N^{vi}(q^{-1})=\sum_{k=0}^{N}\boldsymbol{M}_k^{(i)}q^{k-N},\quad N\geqslant 0 \tag{6.8.33}$$

$$\boldsymbol{L}_N^{vi}(q^{-1})=\boldsymbol{0},\quad N<0 \tag{6.8.34}$$

$$\boldsymbol{M}_k^{(i)}=-\boldsymbol{R}_i\boldsymbol{K}_{pi}^{\mathrm{T}}\boldsymbol{\Psi}_{pi}^{\mathrm{T}(k-1)}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{6.8.35}$$

$$\boldsymbol{M}_0^{(i)}=\boldsymbol{R}_i\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{6.8.36}$$

且估值误差 $\tilde{\boldsymbol{s}}_i(t|t+N)=\boldsymbol{s}(t)-\hat{\boldsymbol{s}}_i(t|t+N)$ 的稳态方差阵 $\boldsymbol{P}_{si}(N)=\mathrm{E}[\tilde{\boldsymbol{s}}_i(t|t+N)\tilde{\boldsymbol{s}}_i^{\mathrm{T}}(t|t+N)]$ 为

$$\boldsymbol{P}_{si}(N)=\boldsymbol{R}_i-\sum_{i=0}^{N}\boldsymbol{M}_k^{(i)}\boldsymbol{Q}_{\varepsilon i}\boldsymbol{M}_k^{(i)},\quad N\geqslant 0 \tag{6.8.37}$$

$$\boldsymbol{P}_{si}(N)=\boldsymbol{H}\boldsymbol{\Sigma}_i(N)\boldsymbol{H}^{\mathrm{T}},\quad N<0 \tag{6.8.38}$$

其中 $\boldsymbol{\Sigma}(N)$ 由(6.8.22)和(6.8.23)定义.

证明 见定理6.7.10,从略.

【定理 6.8.4】 多传感器系统(6.8.1)和(6.8.2)的局部估计误差协方差阵 $\boldsymbol{P}_{sij}(N)=\mathrm{E}[\tilde{\boldsymbol{s}}_i(t\mid t+N)\tilde{\boldsymbol{s}}_j^{\mathrm{T}}(t\mid t+N)](i\neq j)$ 为

$$\boldsymbol{P}_{sij}(N)=\boldsymbol{\Psi}_{iN}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{jN}^{\mathrm{T}}+\sum_{\rho=1}^{N}\boldsymbol{\alpha}_{N\rho}^{(i)}\boldsymbol{Q}\boldsymbol{\alpha}_{N\rho}^{(j)\mathrm{T}},\quad N\geqslant 0 \tag{6.8.39}$$

其中 $\boldsymbol{\Sigma}_{ij}$ 由(6.8.37)决定,且定义

$$\boldsymbol{\Psi}_{iN}=\sum_{k=0}^{N}\boldsymbol{M}_k^{(i)}\boldsymbol{H}\boldsymbol{\Psi}_{pi}^{k} \tag{6.8.40}$$

$$\boldsymbol{\alpha}_{N\rho}^{(i)}=\sum_{k=\rho}^{N}\boldsymbol{M}_k^{(i)}\boldsymbol{H}\boldsymbol{\Psi}_{pi}^{k-\rho}\boldsymbol{\Gamma} \tag{6.8.41}$$

$$\boldsymbol{\beta}_{N\rho}^{(i)}=\sum_{k=\rho}^{N}\boldsymbol{M}_k^{(i)}\boldsymbol{H}\boldsymbol{\Psi}_{pi}^{k-\rho}\boldsymbol{K}_{pi},\quad \rho=2,\cdots,N \tag{6.8.42}$$

$$\boldsymbol{\beta}_{N1}^{(i)}=\sum_{k=\rho}^{N}\boldsymbol{M}_k^{(i)}\boldsymbol{H}\boldsymbol{\Psi}_{pi}^{k-1}\boldsymbol{K}_{pi}+\boldsymbol{M}_0^{(i)}-\boldsymbol{I}_m \tag{6.8.43}$$

且当 $N<0$ 时有

$$\boldsymbol{P}_{sij}(N)=\boldsymbol{H}\boldsymbol{P}_{ij}(N)\boldsymbol{H}^{\mathrm{T}},\quad N<0 \tag{6.8.44}$$

其中 $\boldsymbol{P}_{ij}(N)$ 由(6.8.29)计算.

证明 由(6.8.2)和射影性质有

$$\hat{\boldsymbol{s}}_i(t\mid t+N)=\hat{\boldsymbol{y}}_i(t\mid t+N)-\hat{\boldsymbol{v}}_i(t\mid t+N) \tag{6.8.45}$$

当 $N\geqslant 0$ 时 $\hat{\boldsymbol{y}}_i(t\mid t+N)=\boldsymbol{y}_i(t)$,于是有

$$\tilde{\boldsymbol{s}}_i(t\mid t+N)=-\tilde{\boldsymbol{v}}_i(t\mid t+N)=\sum_{k=0}^{N}\boldsymbol{M}_k^{(i)}\boldsymbol{\varepsilon}_i(t+k)-\boldsymbol{v}_i(t) \tag{6.8.46}$$

注意关系

$$\boldsymbol{\varepsilon}_i(t+k)=\boldsymbol{H}\tilde{\boldsymbol{x}}_i(t+k\mid t+k-1)+\boldsymbol{v}_i(t+k) \tag{6.8.47}$$

$$\tilde{\boldsymbol{x}}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}\tilde{\boldsymbol{x}}_i(t\mid t-1)+\boldsymbol{\Gamma}\boldsymbol{w}(t)-\boldsymbol{K}_{pi}\boldsymbol{v}_i(t) \tag{6.8.48}$$

将上式迭代 $(k-1)$ 次有

$$\tilde{\boldsymbol{x}}_i(t+k\mid t+k-1)=\boldsymbol{\Psi}_{pi}\tilde{\boldsymbol{x}}_i(t\mid t-1)+\sum_{r=1}^{k}\boldsymbol{\Psi}_{pi}^{k-r}[\boldsymbol{\Gamma}\boldsymbol{w}(t+r-1)-\boldsymbol{K}_{pi}\boldsymbol{v}_i(t+r-1)] \tag{6.8.49}$$

将(6.8.47)和(6.8.49)代入(6.8.46)并合并同类项可得

$$\tilde{\boldsymbol{s}}_i(t\mid t+N)=\boldsymbol{\Psi}_{iN}\tilde{\boldsymbol{x}}_i(t\mid t-1)+\sum_{\rho=1}^{N}[\boldsymbol{\alpha}_{N\rho}^{(i)}w(t+\rho-1)+\boldsymbol{\beta}_{N\rho}^{(i)}\boldsymbol{v}_i(t+\rho-1)] \tag{6.8.50}$$

其中 $\boldsymbol{\alpha}_{N\rho}^{(i)},\boldsymbol{\beta}_{N\rho}^{(i)},\boldsymbol{\Psi}_{iN}$ 由(6.8.40)~(6.8.43)定义.将上式代入 $\boldsymbol{P}_{sij}(N)$ 定义中便得(6.8.39),其中用到事实 $\tilde{\boldsymbol{x}}_i(t\mid t-1)$ 不相关于 $\boldsymbol{w}(t),\boldsymbol{w}(t+1),\cdots,\boldsymbol{v}_j(t),\boldsymbol{v}_j(t+1)\cdots$.

当 $N<0$ 时,由(6.8.45)有

$$\hat{\boldsymbol{s}}_i(t\mid t+N)=\hat{\boldsymbol{y}}_i(t\mid t+N),\quad N<0 \tag{6.8.51}$$

这引出

$$\tilde{s}_i(t \mid t+N) = \tilde{y}_i(t \mid t+N) - v_i(t) \tag{6.8.52}$$

从而有

$$P_{sij}(N) = \mathrm{E}[\tilde{y}_i(t \mid t+N)\tilde{y}_i^{\mathrm{T}}(t \mid t+N)] \tag{6.8.53}$$

由(6.8.6)有

$$\hat{y}_i(t \mid t+N) = H\hat{x}_i(t \mid t+N), \quad N < 0 \tag{6.8.54}$$

这引出

$$\tilde{y}_i(t \mid t+N) = H\tilde{x}_i(t \mid t+N) + v_i(t) \tag{6.8.55}$$

于是有

$$P_{sij}(N) = HP_{ij}(N)H^{\mathrm{T}}, \quad N < 0 \tag{6.8.56}$$

证毕. □

【定理 6.8.5】 多传感器系统(6.8.1)~(6.8.3)有稳态局部估计误差协方差阵 $P_{sij}(N) = \mathrm{E}[\tilde{s}_i(t \mid t+N)\tilde{s}_j^{\mathrm{T}}(t \mid t+N)](i \neq j)$ 为

$$P_{sij}(N) = \sum_{r=0}^{N}\sum_{s=0}^{N} M_r^{(i)} E_{ij}(r,s) M_s^{(j)\mathrm{T}}, \quad N \geqslant 0 \tag{6.8.57}$$

其中 $E_{ij}(r,s)$ 由(6.8.26)计算.

证明 由(6.8.2)有

$$\hat{s}_i(t \mid t+N) = \hat{y}_i(t \mid t+N) - \hat{v}_i(t \mid t+N) \tag{6.8.58}$$

当 $N \geqslant 0$ 有

$$\hat{s}_i(t \mid t+N) = y(t) - \hat{v}_i(t \mid t+N) \tag{6.8.59}$$

这引出估值误差 $\tilde{s}_i(t \mid t+N) = s(t) - \hat{s}_i(t \mid t+N)$ 与 $\tilde{v}_i(t \mid t+N) = v(t) - \hat{v}_i(t \mid t+N)$ 有关系

$$\tilde{s}_i(t \mid t+N) = -\tilde{v}_i(t \mid t+N) \tag{6.8.60}$$

由推论 3.5.2 或定理 3.8.6 有白噪声估值器

$$\tilde{v}_i(t \mid t+N) = \sum_{I=0}^{N} M_j^{(i)} \varepsilon_i(t+j), \quad i = 1,\cdots,L \tag{6.8.61}$$

这引出

$$P_{sij}(N) = \mathrm{E}\{[v_i(t) - \sum_{r=0}^{N} M_r^{(i)}\varepsilon_i(t+r)][v_j(t) - \sum_{s=0}^{N} M_s^{(j)}\varepsilon_j(t+s)]^{\mathrm{T}}\} \tag{6.8.62}$$

将上式展开,利用关系(6.8.47)~(6.8.49)可得(6.8.57)和(6.8.26),(6.8.27). □

【定理 6.8.6】 多传感器多通道系统(6.8.1)和(6.8.2)有最优分布式信息融合 Wiener 滤波器

$$\hat{s}_0(t \mid t+N) = \sum_{i=1}^{L} A_i \hat{s}_i(t \mid t+N) \tag{6.8.63}$$

在节 6.1 三种加权最优融合准则下,加权 A_i 分别为矩阵、对角阵或标量,它们可分别由三种加权融合公式通过 $P_{si}(N)$ 和 $P_{sij}(N)$ 求得,$\hat{s}_i(t \mid t+N)$ 可由定理 6.8.1 或定理 6.8.3 求得.相应的最优融合误差方差阵 $P_{s0}(N)$ 满足关系

$$\mathrm{tr}P_{s0}(N) \leqslant \mathrm{tr}P_{si}(N), \quad i = 1,\cdots,L \tag{6.8.64}$$

证明 由节 6.1 的三种加权公式得证. □

【例 6.8.1】 考虑两传感器二维跟踪系统

$$(\boldsymbol{I}_2 + \boldsymbol{A}_1 q^{-1}) \boldsymbol{s}(t) = \boldsymbol{C}_1 q^{-1} \boldsymbol{w}(t) \tag{6.8.65}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{s}(t) + \boldsymbol{v}_i(t), \quad i = 1,2 \tag{6.8.66}$$

$$\boldsymbol{A}_1 = \begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{C}_1 = \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix} \tag{6.8.67}$$

其中 $\boldsymbol{s}(t) = [s_1(t), s_2(t)]^{\mathrm{T}}$, T_0 为采样周期，$s_1(t)$，$s_2(t)$ 和 $w(t)$ 分别为在时刻 tT_0 处运动目标的位置、速度和加速度，$y_i(t)$ 为第 i 个传感器对 $\boldsymbol{s}(t)$ 的观测信号，$\boldsymbol{v}_i(t)$ 为观测噪声. 设 $w(t)$ 和 $\boldsymbol{v}_i(t)$ 是零均值、方差阵各为 Q（标量）和 $\boldsymbol{R}_i$ 的相互独立白噪声，且取

$$T_0 = 0.3, \quad Q = 1, \quad \boldsymbol{Q}_{v1} = \begin{bmatrix} 1 & 0 \\ 0 & 2.25 \end{bmatrix}, \quad \boldsymbol{Q}_{v2} = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \tag{6.8.68}$$

取 $N = 1$，问题是应用定理 6.8.2 和定理 6.8.3 求局部跟踪平滑器 $\hat{\boldsymbol{s}}_i(t \mid t+1)$ 和最优按矩阵加权融合跟踪平滑器 $\hat{\boldsymbol{s}}_0(t \mid t+1)$，可求得

$$\mathrm{tr}\boldsymbol{P}_{s1}(1) = 0.409\,0, \quad \mathrm{tr}\boldsymbol{P}_{s2}(1) = 1.083\,7, \quad \mathrm{tr}\boldsymbol{P}_{s0}(1) = 0.364\,0 \tag{6.8.69}$$

注意 $\mathrm{tr}\boldsymbol{P}_{s0}(1) < \mathrm{tr}\boldsymbol{P}_{si}(1)$，$i = 1,2$. 这表明融合平滑器的精度高于局部平滑器的精度. 仿真结果如图 6.8.1 ~ 6.8.6 所示，其中实线代表真实值 $s_i(t)$，虚线代表平滑估值 $\hat{s}_{ij}(t \mid t+1)$，即第 j 个传感器对 $s_i(t)$ 的平滑估值. 可看到融合估计精度高于局部估计精度. 图 6.8.7 和图 6.8.8 分别为位置和速度的累积平滑误差平方比较曲线，其中实曲线为融合平滑器的累积误差平方曲线. 可进一步看到融合平滑估计的精度高于局部平滑估计精度.

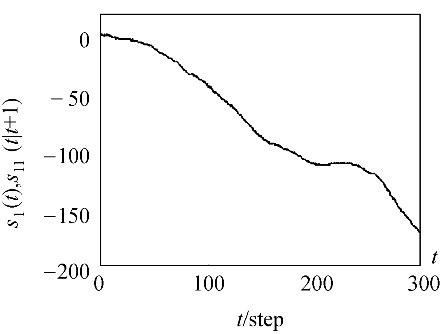


图 6.8.1　信号 $s_1(t)$ 与局部 Wiener 平滑器 $\hat{s}_{11}(t \mid t+1)$ 的比较图

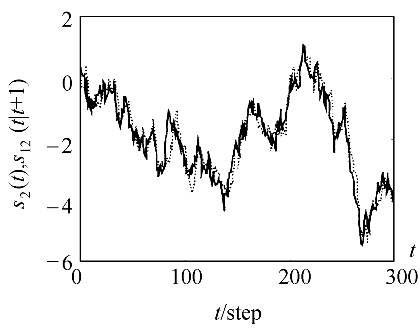


图 6.8.2　速度 $s_2(t)$ 与局部 Wiener 平滑器 $\hat{s}_{12}(t \mid t+1)$ 的比较图

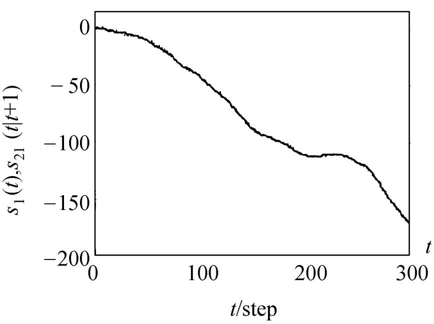


图 6.8.3　位置 $s_1(t)$ 与局部 Wiener 平滑器 $\hat{s}_{21}(t \mid t+1)$ 的比较图

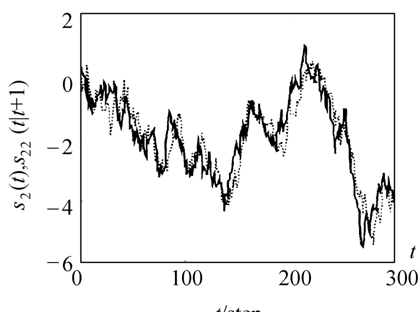


图 6.8.4　速度 $s_2(t)$ 与局部 Wiener 平滑器 $\hat{s}_{22}(t \mid t+1)$ 的比较图

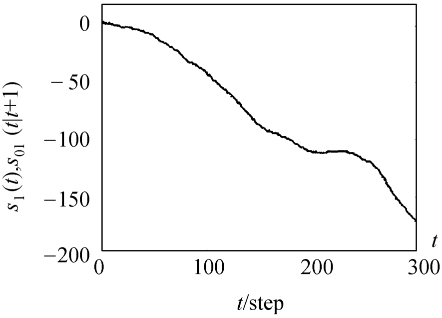


图 6.8.5 位置 $s_1(t)$ 与最优融合跟踪平滑器 $\hat{s}_{01}(t \mid t+1)$

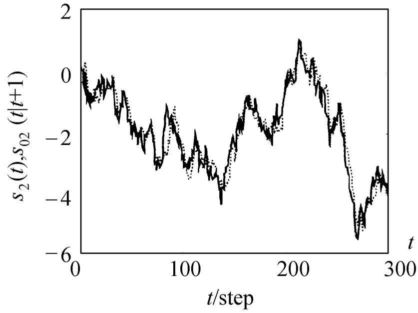


图 6.8.6 位置 $s_1(t)$ 与最优融合跟踪平滑器 $\hat{s}_{02}(t \mid t+1)$

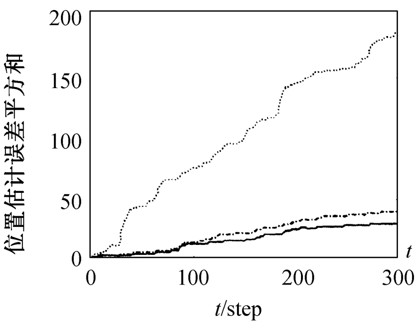


图 6.8.7 位置估计误差平方和比较曲线

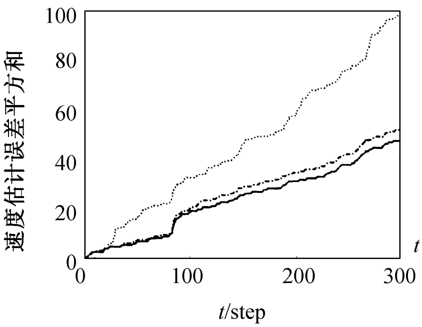


图 6.8.8 速度估计误差平方和比较曲线

6.9 广义系统多传感器信息融合降阶状态估值器

广义系统大量出现在电网络、机器人、经济系统、航空航天、复杂化工过程控制等领域，近年来尤为人们关注. 文献[34,35]用现代时间序列分析方法提出了广义系统降阶状态估值器. 原理是通过变换将广义系统化为典范型，它由两个互耦的子系统组成. 因而可将整个系统状态估计问题转化为子系统状态估计问题. 因为每个子系统的状态有较小维数，故称子系统状态估值器为降维状态估值器. 它具有计算量较小，便于实时应用的优点. 关于广义系统信息融合状态估值器目前尚未见国内外报道. 本节对两种广义系统典范型提出了相应的两种降阶状态估值器[67]. 它可转化为正常（非广义）系统的信息融合状态估计问题，因而可利用本章节 6.1 ~ 节 6.7 的有关结果解决问题.

6.9.1 在典范型 Ⅰ 下的广义系统分布式信息融合降阶状态估值器[67]

考虑多传感器线性离散时间定常广义随机系统

$$\boldsymbol{Mx}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t) \tag{6.9.1}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{H}_i\boldsymbol{x}(t)+\boldsymbol{v}_i(t),\quad i=1,\cdots,L \tag{6.9.2}$$

其中 t 为离散时间，状态 $\boldsymbol{x}(t)\in R^n$，第 i 个传感器的观测为 $\boldsymbol{y}_i(t)\in R^{m_i}$，$\boldsymbol{v}_i(t)\in R^{m_i}$ 为

观测噪声.

【假设 1】 $\boldsymbol{M}$ 为奇异方阵($\det\boldsymbol{M}=0$).

【假设 2】 系统是正则的,即 $\det(\boldsymbol{M}-q^{-1}\boldsymbol{\Phi})\not\equiv 0$.

【假设 3】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}_i(t)\in R^{m_i}$ 是零均值相关白噪声:

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}_i(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{w}^{\mathrm{T}}(k) & \boldsymbol{v}_i^{\mathrm{T}}(k)\end{bmatrix}\right\}=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}_i\\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_i\end{bmatrix}\delta_{tk},$$

$$\mathrm{E}[\boldsymbol{v}_i(t)\boldsymbol{v}_j^{\mathrm{T}}(t)]=\boldsymbol{0},\quad \forall\, i,j,t,k \tag{6.9.3}$$

其中 E 为均值号,T 为转置号,$\delta_{tt}=1$,$\delta_{tk}=0(t\neq k)$.

【假设 4】 系统是完全可观的,即对任意复数 z 有

$$\mathrm{rank}\begin{bmatrix}z\boldsymbol{M}-\boldsymbol{\Phi}\\ \boldsymbol{H}_i\end{bmatrix}=n,\quad \mathrm{rank}\begin{bmatrix}\boldsymbol{M}\\ \boldsymbol{H}_i\end{bmatrix}=n,\quad i=1,\cdots,L \tag{6.9.4}$$

问题是基于观测($\boldsymbol{y}_i(t+N),\boldsymbol{y}_i(t+N-1),\cdots$)求 $\boldsymbol{x}(t)$ 的局部稳态状态估值器 $\hat{\boldsymbol{x}}_i(t\mid t+N)$ 和信息融合稳态状态估值器 $\hat{\boldsymbol{x}}_0(t\mid t+N)$.对 $N=0$,$N>0$ 或 $N<0$,称其为滤波器、平滑器或预报器.

由假设 1 和假设 2,设 $\mathrm{rank}\boldsymbol{M}=n_1<n$,利用奇异值分解[44],存在非异阵 $\boldsymbol{P},\boldsymbol{Q}$ 使

$$\boldsymbol{PMQ}=\begin{bmatrix}\boldsymbol{\Sigma} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix},\quad \boldsymbol{\Sigma}=\begin{bmatrix}\sigma_1 & & \boldsymbol{0}\\ & \ddots & \\ \boldsymbol{0} & & \sigma_{n_1}\end{bmatrix},\quad \sigma_i>0,i=1,\cdots,n_1 \tag{6.9.5}$$

引入分块表示

$$\boldsymbol{P\Phi Q}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix},\quad \boldsymbol{P\Gamma}=\begin{bmatrix}\boldsymbol{\Gamma}_1\\ \boldsymbol{\Gamma}_2\end{bmatrix},\quad \boldsymbol{H}_i\boldsymbol{Q}=[\boldsymbol{H}_{i1},\boldsymbol{H}_{i2}] \tag{6.9.6}$$

且引入状态变换

$$\boldsymbol{x}(t)=\boldsymbol{Q}\begin{bmatrix}\boldsymbol{x}_1(t)\\ \boldsymbol{x}_2(t)\end{bmatrix} \tag{6.9.7}$$

则系统(6.9.1)和(6.9.2)化为典范型 Ⅰ 为

$$\begin{bmatrix}\boldsymbol{\Sigma} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_1(t+1)\\ \boldsymbol{x}_2(t+1)\end{bmatrix}=\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_1(t)\\ \boldsymbol{x}_2(t)\end{bmatrix}+\begin{bmatrix}\boldsymbol{\Gamma}_1\\ \boldsymbol{\Gamma}_2\end{bmatrix}\boldsymbol{w}(t) \tag{6.9.8}$$

$$\boldsymbol{y}_i(t)=[\boldsymbol{H}_{i1},\boldsymbol{H}_{i2}]\begin{bmatrix}\boldsymbol{x}_1(t)\\ \boldsymbol{x}_2(t)\end{bmatrix}+\boldsymbol{v}_i(t),\quad i=1,\cdots,L \tag{6.9.9}$$

其中 $n=n_1+n_2$,$\boldsymbol{x}_1(t)\in R^{n_1}$,$\boldsymbol{x}_2(t)\in R^{n_2}$.它可写成互耦的子系统

$$\boldsymbol{x}_1(t+1)=\boldsymbol{\Sigma}^{-1}\boldsymbol{A}_{11}\boldsymbol{x}_1(t)+\boldsymbol{\Sigma}^{-1}\boldsymbol{A}_{12}\boldsymbol{x}_2(t)+\boldsymbol{\Sigma}^{-1}\boldsymbol{\Gamma}_1\boldsymbol{w}(t) \tag{6.9.10}$$

$$\boldsymbol{0}=\boldsymbol{A}_{21}\boldsymbol{x}_1(t)+\boldsymbol{A}_{22}\boldsymbol{x}_2(t)+\boldsymbol{\Gamma}_2\boldsymbol{w}(t) \tag{6.9.11}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{H}_{i1}\boldsymbol{x}_1(t)+\boldsymbol{H}_{i2}\boldsymbol{x}_2(t)+\boldsymbol{v}_i(t),\quad i=1,\cdots,L \tag{6.9.12}$$

假设 $\boldsymbol{A}_{22}$ 非异,则由(6.9.11)可解出

$$\boldsymbol{x}_2(t)=-\boldsymbol{A}_{22}^{-1}[\boldsymbol{A}_{21}\boldsymbol{x}_1(t)+\boldsymbol{\Gamma}_2\boldsymbol{w}(t)] \tag{6.9.13}$$

将其代入(6.9.12)有多传感器降阶子系统

$$\boldsymbol{x}_1(t+1)=\boldsymbol{A}_0\boldsymbol{x}_1(t)+\boldsymbol{\Gamma}_0\boldsymbol{w}(t) \tag{6.9.14}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{H}_{i0}\boldsymbol{x}_1(t)+\boldsymbol{\eta}_i(t),\quad i=1,\cdots,L \tag{6.9.15}$$

其中定义

$$\begin{aligned}
\boldsymbol{A}_0&=\boldsymbol{\Sigma}^{-1}(\boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}),\\
\boldsymbol{\Gamma}_0&=\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Gamma}_1-\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2),\\
\boldsymbol{H}_{i0}&=\boldsymbol{H}_{i1}-\boldsymbol{H}_{i2}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21},\\
\boldsymbol{\eta}_i(t)&=\boldsymbol{v}_i(t)-\boldsymbol{H}_{i2}\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{w}(t)
\end{aligned} \tag{6.9.16}$$

注意 $\boldsymbol{w}(t)$ 与 $\boldsymbol{\eta}_i(t)$ 是相关白噪声：

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{\eta}_i(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{w}^{\mathrm{T}}(k) & \boldsymbol{\eta}_i^{\mathrm{T}}(k)\end{bmatrix}\right\}=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}_{i0}\\ \boldsymbol{S}_{i0}^{\mathrm{T}} & \boldsymbol{Q}_{\eta i}\end{bmatrix}\delta_{tk} \tag{6.9.17}$$

$$\boldsymbol{S}_{i0}=\boldsymbol{S}_i-\boldsymbol{Q}_w\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}\boldsymbol{H}_{i2}^{\mathrm{T}} \tag{6.9.18}$$

$$\boldsymbol{Q}_{\eta i}=\boldsymbol{R}_i-\boldsymbol{H}_{i2}\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{S}_i-\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-1}\boldsymbol{H}_{i2}^{\mathrm{T}}+\boldsymbol{H}_{i2}\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{Q}_w\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-1}\boldsymbol{H}_{i2}^{\mathrm{T}} \tag{6.9.19}$$

可证明 $(\boldsymbol{A}_0,\boldsymbol{H}_{i0})$ $(i=1,\cdots,L)$ 为完全可观对[35].因而子系统(6.9.14)和(6.9.15)存在稳态 Kalman 估值器.

对于多传感器常规降阶子系统(6.9.14)和(6.9.15)用本章方法可求得 $\boldsymbol{x}_1(t)$ 的局部稳态状态估值器 $\hat{\boldsymbol{x}}_i(t\mid t+N)$ $(i=1,\cdots,L)$ 和 $\boldsymbol{w}(t)$ 的局部稳态白噪声估值器 $\hat{\boldsymbol{w}}_i(t\mid t+N)$，进而由(6.9.13)可得 $\boldsymbol{x}_2(t)$ 的局部稳态估值器

$$\hat{\boldsymbol{x}}_{2i}(t\mid t+N)=-\boldsymbol{A}_{22}^{-1}[\boldsymbol{A}_{21}\hat{\boldsymbol{x}}_{1i}(t\mid t+N)+\boldsymbol{\Gamma}_2\hat{\boldsymbol{w}}_i(t\mid t+N)],\quad i=1,\cdots,L \tag{6.9.20}$$

定义 $\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)=\boldsymbol{x}_1(t)-\hat{\boldsymbol{x}}_{1i}(t\mid t+N)$，$\tilde{\boldsymbol{x}}_{2i}(t\mid t+N)=\boldsymbol{x}_2(t)-\hat{\boldsymbol{x}}_{2i}(t\mid t+N)$，

$$\begin{aligned}
\boldsymbol{P}_i^{x_1}(N)&=\mathrm{E}[\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)\tilde{\boldsymbol{x}}_{1i}^{\mathrm{T}}(t\mid t+N)],\\
\boldsymbol{P}_i^{x_2}(N)&=\mathrm{E}[\tilde{\boldsymbol{x}}_{2i}(t\mid t+N)\tilde{\boldsymbol{x}}_{2i}^{\mathrm{T}}(t\mid t+N)],\\
\boldsymbol{P}_{ij}^{x_1}(N)&=\mathrm{E}[\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)\tilde{\boldsymbol{x}}_{1j}^{\mathrm{T}}(t\mid t+N)],\\
\boldsymbol{P}_{ij}^{x_2}(N)&=\mathrm{E}[\tilde{\boldsymbol{x}}_{2i}(t\mid t+N)\tilde{\boldsymbol{x}}_{2j}^{\mathrm{T}}(t\mid t+N)]
\end{aligned} \tag{6.9.21}$$

且定义 $\tilde{\boldsymbol{w}}_i(t\mid t+N)=\boldsymbol{w}_1(t)-\hat{\boldsymbol{w}}_i(t\mid t+N)$，

$$\begin{aligned}
\boldsymbol{P}_i^{w}(N)&=\mathrm{E}[\tilde{\boldsymbol{w}}_i(t\mid t+N)\tilde{\boldsymbol{w}}_i^{\mathrm{T}}(t\mid t+N)],\\
\boldsymbol{P}_{ij}^{w}(N)&=\mathrm{E}[\tilde{\boldsymbol{w}}_i(t\mid t+N)\tilde{\boldsymbol{w}}_j^{\mathrm{T}}(t\mid t+N)],\\
\boldsymbol{P}_{ij}^{x_1w}(N)&=\mathrm{E}[\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)\tilde{\boldsymbol{w}}_j^{\mathrm{T}}(t\mid t+N)],\\
\boldsymbol{P}_{ij}^{wx_1}(N)&=\mathrm{E}[\tilde{\boldsymbol{w}}_i(t\mid t+N)\tilde{\boldsymbol{x}}_{1j}^{\mathrm{T}}(t\mid t+N)]
\end{aligned} \tag{6.9.22}$$

还定义 $\tilde{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}(t\mid t+N)$，

$$\begin{aligned}
\boldsymbol{P}_i^{x}(N)&=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t+N)\tilde{\boldsymbol{x}}_i^{\mathrm{T}}(t\mid t+N)],\\
\boldsymbol{P}_{ij}^{x}(N)&=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t+N)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t\mid t+N)]
\end{aligned} \tag{6.9.23}$$

$$\begin{aligned}
\boldsymbol{P}_{ij}^{x_1x_2}(N)&=\mathrm{E}[\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)\tilde{\boldsymbol{x}}_{2j}^{\mathrm{T}}(t\mid t+N)],\\
\boldsymbol{P}_{ij}^{x_2x_1}(N)&=\mathrm{E}[\tilde{\boldsymbol{x}}_{2i}(t\mid t+N)\tilde{\boldsymbol{x}}_{1j}^{\mathrm{T}}(t\mid t+N)]
\end{aligned} \tag{6.9.24}$$

由(6.9.7)有状态 $\boldsymbol{x}(t)$ 的局部稳态估值器为

$$\hat{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{Q}\begin{bmatrix}\hat{\boldsymbol{x}}_{1i}(t\mid t+N)\\ \hat{\boldsymbol{x}}_{2i}(t\mid t+N)\end{bmatrix},\quad i=1,\cdots,L \tag{6.9.25}$$

对第 i 个传感器子系统(6.9.14)和(6.9.15)由节 6.6 有稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}_{1i}(t+1\mid t)=\boldsymbol{\Psi}_{pi}\hat{\boldsymbol{x}}_{1i}(t\mid t-1)+\boldsymbol{K}_{pi}\boldsymbol{y}_i(t) \tag{6.9.26}$$

$$\boldsymbol{\varepsilon}_i(t)=\boldsymbol{y}_i(t)-\boldsymbol{H}_{i0}\hat{\boldsymbol{x}}_{1i}(t\mid t-1),\quad i=1,\cdots,L \tag{6.9.27}$$

且当 $N\geqslant 0$ 时有稳态 Kalman 滤波器和平滑器

$$\hat{\boldsymbol{x}}_{1i}(t\mid t+N)=\hat{\boldsymbol{x}}_{1i}(t\mid t-1)+\sum_{k=0}^{N}\boldsymbol{K}_k^{(i)}\boldsymbol{\varepsilon}_i(t+k) \tag{6.9.28}$$

当 $N<-1$ 有 Kalman 预报器

$$\hat{\boldsymbol{x}}_{1i}(t\mid t+N)=\boldsymbol{A}_0^{-N-1}\hat{\boldsymbol{x}}_{1i}(t+N+1\mid t+N) \tag{6.9.29}$$

且有 $w(t)$ 的局部白噪声估值器

$$\hat{\boldsymbol{w}}_i(t\mid t+N)=\sum_{k=0}^{N}\boldsymbol{M}_k^{(i)}\boldsymbol{\varepsilon}_i(t+k),\quad N\geqslant 0 \tag{6.9.30}$$

$$\hat{\boldsymbol{w}}_i(t\mid t+N)=\boldsymbol{0},\quad N<0 \tag{6.9.31}$$

其中 $\boldsymbol{\Psi}_{pi},\boldsymbol{K}_{pi},\boldsymbol{K}_k^{(i)},\boldsymbol{M}_k^{(i)}$ 及一步预报误差互协方差阵 $\boldsymbol{\Sigma}_{ij}$，新息互协方差阵 $\boldsymbol{E}_{ij}(r,s)$，有关估值误差方差阵和互协方差阵 $\boldsymbol{P}_i^{x_1}(N),\boldsymbol{P}_{ij}^{x_1}(N)$，均可由节 6.6 的有关公式得到.

由(6.9.7)和(6.9.25)有稳态局部估计误差为

$$\tilde{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{Q}\begin{bmatrix}\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)\\ \tilde{\boldsymbol{x}}_{2i}(t\mid t+N)\end{bmatrix} \tag{6.9.32}$$

由(6.9.21)～(6.9.24)，这引出稳态局部估值误差方差阵和互协方差阵为

$$\boldsymbol{P}_i^{x}(N)=\boldsymbol{Q}\begin{bmatrix}\boldsymbol{P}_i^{x_1}(N) & \boldsymbol{P}_{ii}^{x_1x_2}(N)\\ \boldsymbol{P}_{ii}^{x_2x_1}(N) & \boldsymbol{P}_i^{x_2}(N)\end{bmatrix}\boldsymbol{Q}^{\mathrm{T}} \tag{6.9.33}$$

$$\boldsymbol{P}_{ij}^{x}(N)=\boldsymbol{Q}\begin{bmatrix}\boldsymbol{P}_{ij}^{x_1}(N) & \boldsymbol{P}_{ij}^{x_1x_2}(N)\\ \boldsymbol{P}_{ij}^{x_2x_1}(N) & \boldsymbol{P}_{ij}^{x_2}(N)\end{bmatrix}\boldsymbol{Q}^{\mathrm{T}},\quad i,j=1,\cdots,L,i\neq j \tag{6.9.34}$$

为了求最优融合估计 $\hat{\boldsymbol{x}}_0(t\mid t+N)$，要求计算 $\boldsymbol{P}_i^{x}(N)$ 和 $\boldsymbol{P}_{ij}^{x}(N)$. 除了已由节 6.6 得到 $\boldsymbol{P}_i^{x_1}(N)$ 和 $\boldsymbol{P}_{ij}^{x_1}(N)$ 外，需要进一步求 $\boldsymbol{P}_i^{x_2}(N),\boldsymbol{P}_{ij}^{x_2}(N),\boldsymbol{P}_{ij}^{x_1x_2}(N),\boldsymbol{P}_{ij}^{x_2x_1}(N),i,j=1,\cdots,L.$

为了推导方便，以下推导假设子系统(6.9.14)和(6.9.15)是稳定的，即 $\boldsymbol{A}_0$ 是稳定矩阵，因而 $\boldsymbol{x}_1(t)$ 和 $\boldsymbol{x}_2(t)$ 均为平稳随机过程，故在推导中所有数学期望运算均有意义(即所有数学期望值存在). 对 $\boldsymbol{A}_0$ 不稳定情形，可设初始观测时刻为 $t_0=0$，平行于 $\boldsymbol{A}_0$ 稳定情形的推导，可求得时变最优状态估值器及时变估值误差方差阵和协方差阵，再令 $t\to\infty$ 取极限可得相应的稳态估值器及稳态估值误差方差阵和协方差阵. 有关公式同 $\boldsymbol{A}_0$ 稳定情形的公式完全一样. 取 $t_0=0$ 为有限的可保证推导中遇到的数学期望运算有意义. 在取 $t_0=0$ 时，还应假设 $\boldsymbol{x}_1(0)$ 与 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}_i(t),i=1,\cdots,L$，相互独立.

【定理 6.9.1】 广义系统(6.9.1)和(6.9.2)在假设 1 ～ 4 下，当 $N\geqslant 0$ 时，有

$$\begin{aligned}\boldsymbol{P}_i^{x_2}(N)=&\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_i^{x_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}+\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{P}_i^{w}(N)\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}+\\&\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_{ii}^{x_1w}(N)\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}+\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{P}_{ii}^{wx_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}\end{aligned} \tag{6.9.35}$$

$$\boldsymbol{P}_{ii}^{x_1w}(N)=-\sum_{k=0}^{N}\boldsymbol{K}_i^{(i)}\boldsymbol{Q}_{\varepsilon i}\boldsymbol{M}_k^{(i)\mathrm{T}},\quad \boldsymbol{P}_{ii}^{wx_1}(N)=-\sum_{k=0}^{N}\boldsymbol{M}_i^{(i)}\boldsymbol{Q}_{\varepsilon i}\boldsymbol{K}_k^{(i)\mathrm{T}} \tag{6.9.36}$$

证明 由(6.9.13)和(6.9.20)有

$$\tilde{\boldsymbol{x}}_{2i}(t\mid t+N)=-\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)-\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\tilde{\boldsymbol{w}}_i(t\mid t+N) \tag{6.9.37}$$

由 $\boldsymbol{P}_i^{x_2}(N)$ 的定义(6.9.21)得(6.9.35),其中应用了定义(6.9.22).注意 $\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)$ 正交(不相关)于由(6.9.30)给出的 $\hat{\boldsymbol{w}}_i(t\mid t+N)$,$N\geqslant 0$,并注意关系

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(t+k)]=\boldsymbol{M}_k^{(i)}\boldsymbol{Q}_{\varepsilon i}\tag{6.9.38}$$

并利用(6.9.28)及 $\boldsymbol{w}(t)$ 不相关于 $\boldsymbol{x}_1(t)$ 和 $\hat{\boldsymbol{x}}_{1i}(t\mid t-1)$,则当 $N\geqslant 0$ 时有

$$\boldsymbol{P}_{ii}^{x_1w}(N)=\mathrm{E}[\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)\tilde{\boldsymbol{w}}_i^{\mathrm{T}}(t\mid t+N)]=\mathrm{E}[\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)(\boldsymbol{w}(t)-\hat{\boldsymbol{w}}_i(t\mid t+N)^{\mathrm{T}}]=$$

$$\mathrm{E}[(\boldsymbol{x}_1(t)-\hat{\boldsymbol{x}}_{1i}(t\mid t+N)\boldsymbol{w}^{\mathrm{T}}(t)]=-\sum_{k=0}^{N}\boldsymbol{K}_k^{(i)}\boldsymbol{Q}_{\varepsilon i}\boldsymbol{M}_k^{(i)\mathrm{T}}\tag{6.9.39}$$

而 $\boldsymbol{P}_{ii}^{wx_1}(N)=(\boldsymbol{P}_{ii}^{x_1w}(N))^{\mathrm{T}}$.证毕. □

【定理 6.9.2】 广义系统(6.9.1)和(6.9.2)在假设 1 ~ 4 下,当 $N\geqslant 0$ 时,有

$$\begin{aligned}\boldsymbol{P}_{ij}^{x_2}(N)=&\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_{ij}^{x_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}+\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{P}_{ij}^{w}(N)\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}+\\&\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_{ij}^{x_1w}(N)\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}+\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{P}_{ij}^{wx_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}\end{aligned}\tag{6.9.40}$$

$$\begin{aligned}\boldsymbol{P}_{ij}^{x_1w}(N)=&-\sum_{k=0}^{N}\boldsymbol{K}_k^{(i)}\boldsymbol{Q}_{\varepsilon i}\boldsymbol{M}_k^{(i)\mathrm{T}}-\sum_{k=0}^{N}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{k\mathrm{T}}\boldsymbol{H}_{0j}^{\mathrm{T}}\boldsymbol{M}_k^{(j)\mathrm{T}}+\\&\sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{K}_r^{(i)}\boldsymbol{E}_{ij}(r,s)\boldsymbol{M}_s^{(j)\mathrm{T}}\end{aligned}\tag{6.9.41}$$

$$\begin{aligned}\boldsymbol{P}_{ij}^{wx_1}(N)=&-\sum_{k=0}^{N}\boldsymbol{M}_k^{(j)}\boldsymbol{Q}_{\varepsilon j}\boldsymbol{K}_k^{(j)\mathrm{T}}-\sum_{k=0}^{N}\boldsymbol{M}_k^{(i)}\boldsymbol{H}_{0i}\boldsymbol{\Psi}_{pi}^{k}\boldsymbol{\Sigma}_{ij}+\\&\sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{M}_r^{(i)}\boldsymbol{E}_{ij}(r,s)\boldsymbol{K}_s^{(j)\mathrm{T}}\end{aligned}\tag{6.9.42}$$

其中 $\boldsymbol{E}_{ij}(r,s)=\mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)]$ 由(6.6.72) ~ (6.6.75)计算.

证明 应用(6.9.28)和(6.9.30),当 $N\geqslant 0$ 且 $i\neq j$ 时有

$$\boldsymbol{P}_{ij}^{x_1w}(N)=\mathrm{E}[\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)\tilde{\boldsymbol{w}}_j^{\mathrm{T}}(t\mid t+N)]=$$

$$\mathrm{E}[(\tilde{\boldsymbol{x}}_{1i}(t\mid t-1)-\sum_{k=0}^{N}\boldsymbol{K}_k^{(i)}\boldsymbol{\varepsilon}_i(t+k))(\boldsymbol{w}(t)-\sum_{k=0}^{N}\boldsymbol{M}_k^{(j)}\boldsymbol{\varepsilon}_j(t+k))^{\mathrm{T}}]\tag{6.9.43}$$

将上式展开为四项,注意 $\boldsymbol{w}(t)$ 正交(不相关)于 $\hat{\boldsymbol{x}}_{1i}(t\mid t-1)$,应用(6.9.38),并应用由(6.4.26)引出的稳态表达式

$$\begin{aligned}\boldsymbol{\varepsilon}_i(t+k)=&\boldsymbol{H}_{i0}\boldsymbol{\Psi}_{pi}^{k}\tilde{\boldsymbol{x}}_{1i}(t\mid t-1)+\sum_{r=1}^{k}\boldsymbol{H}_{i0}\boldsymbol{\Psi}_{pi}^{k-r}[\boldsymbol{\Gamma}_0\boldsymbol{w}(t+r-1)-\\&\boldsymbol{K}_{pi}\boldsymbol{\eta}_i(t+r-1)]+\boldsymbol{\eta}_i(t+k)\end{aligned}\tag{6.9.44}$$

其中规定 $k=0$ 时上式第二项为零,则立刻得(6.9.41).同理得(6.9.42). □

【定理 6.9.3】 广义系统(6.9.1)和(6.9.2)在假设 1 ~ 4 下,当 $N<0$ 时有

$$\boldsymbol{P}_i^{x_2}(N)=\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_i^{x_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-1}+\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{Q}_w\boldsymbol{\Gamma}_2^{\mathrm{T}}\overline{\boldsymbol{A}}_{22}^{\mathrm{T}}\tag{6.9.45}$$

$$\boldsymbol{P}_{ij}^{x_2}(N)=\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_{ij}^{x_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}+\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{Q}_w\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}},\quad i\neq j\tag{6.9.46}$$

证明 注意当 $N<0$ 时有 $\hat{\boldsymbol{w}}_i(t\mid t+N)=\mathbf{0}$,故有

$$\tilde{\boldsymbol{x}}_{2i}(t\mid t+N)=-\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)-\boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{w}(t)\tag{6.9.47}$$

注意当 $N<0$ 时 $\boldsymbol{w}(t)$ 不相关于 $\tilde{\boldsymbol{x}}_{1i}(t\mid t+N)$,由定义(6.9.21)直接引出(6.9.45)和(6.9.46).证毕. □

【定理 6.9.4】 广义系统(6.9.1)和(6.9.2)在假设 1 ~ 4 下有

$$\boldsymbol{P}_{ii}^{x_1x_2}(N) = -\boldsymbol{P}_i^{x_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}} - \boldsymbol{P}_{ii}^{x_1w}(N)\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}, \quad N \geqslant 0 \tag{6.9.48}$$

$$\boldsymbol{P}_{ii}^{x_1x_2}(N) = -\boldsymbol{P}_i^{x_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}, \quad N < 0 \tag{6.9.49}$$

$$\boldsymbol{P}_{ii}^{x_2x_1}(N) = -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_i^{x_1}(N) - \boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{P}_{ii}^{wx_1}(N), \quad N \geqslant 0 \tag{6.9.50}$$

$$\boldsymbol{P}_{ii}^{x_2x_1}(N) = -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_i^{x_1}(N), \quad N < 0 \tag{6.9.51}$$

$$\boldsymbol{P}_i^x(N) = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{P}_i^{x_1}(N) & \boldsymbol{P}_{ii}^{x_1x_2}(N) \\ \boldsymbol{P}_{ii}^{x_2x_1}(N) & \boldsymbol{P}_i^{x_2}(N) \end{bmatrix}\boldsymbol{Q}^{\mathrm{T}} \tag{6.9.52}$$

证明 利用(6.9.37)和定义(6.9.24)立刻得(6.9.48)和(6.9.50). 当 $N < 0$ 时利用(6.9.47)和 $\boldsymbol{w}(t)$ 不相关于 $\tilde{\boldsymbol{x}}_{1i}(t \mid t+N)$ 引出(6.9.49)和(6.9.51). 由(6.9.33)得(6.9.52). 证毕. □

【定理 6.9.5】 广义系统(6.9.1)和(6.9.2)在假设 1 ~ 4 下有

$$\boldsymbol{P}_{ij}^{x_1x_2}(N) = -\boldsymbol{P}_{ij}^{x_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}} - \boldsymbol{P}_{ij}^{x_1w}(N)\boldsymbol{\Gamma}_2^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}, \quad N \geqslant 0 \tag{6.9.53}$$

$$\boldsymbol{P}_{ij}^{x_1x_2}(N) = -\boldsymbol{P}_{ij}^{x_1}(N)\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{A}_{22}^{-\mathrm{T}}, \quad N < 0 \tag{6.9.54}$$

$$\boldsymbol{P}_{ij}^{x_2x_1}(N) = -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_{ij}^{x_1}(N) - \boldsymbol{A}_{22}^{-1}\boldsymbol{\Gamma}_2\boldsymbol{P}_{ij}^{wx_1}(N), \quad N \geqslant 0 \tag{6.9.55}$$

$$\boldsymbol{P}_{ij}^{x_2x_1}(N) = -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{P}_{ij}^{x_1}(N), \quad N < 0 \tag{6.9.56}$$

$$\boldsymbol{P}_{ij}^x(N) = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{P}_{ij}^{x_1}(N) & \boldsymbol{P}_{ij}^{x_1x_2}(N) \\ \boldsymbol{P}_{ij}^{x_2x_1}(N) & \boldsymbol{P}_{ij}^{x_2}(N) \end{bmatrix}\boldsymbol{Q}^{\mathrm{T}}, \quad i \neq j, i, j = 1, \cdots, L \tag{6.9.57}$$

证明 利用(6.9.37)和定义(6.9.24),当 $N \geqslant 0$ 时立刻得(6.9.53)和(6.9.55). 当 $N < 0$ 时 $\hat{\boldsymbol{w}}_j(t \mid t+N) = \boldsymbol{0}$,于是 $\hat{\boldsymbol{w}}_j(t \mid t+N) = \boldsymbol{w}(t)$,$\tilde{\boldsymbol{w}}_i(t \mid t+N) = \boldsymbol{w}(t)$. 注意当 $N < 0$ 时 $\boldsymbol{w}(t)$ 不相关于 $\tilde{\boldsymbol{x}}_{1i}(t \mid t+N)$ 和 $\tilde{\boldsymbol{x}}_{1j}(t \mid t+N)$,于是 $\boldsymbol{P}_{ij}^{x_1w}(N) = \boldsymbol{0}$,$\boldsymbol{P}_{ij}^{wx_1}(N) = \boldsymbol{0}$. 这引出(6.9.54)和(6.9.56). 证毕. □

【定理 6.9.6】 广义系统(6.9.1)和(6.9.2)在假设 1 ~ 4 下有局部稳态估值器

$$\hat{\boldsymbol{x}}_i(t \mid t+N) = \boldsymbol{Q}\begin{bmatrix} \hat{\boldsymbol{x}}_{1i}(t \mid t+N) \\ \hat{\boldsymbol{x}}_{2i}(t \mid t+N) \end{bmatrix}, \quad i = 1, \cdots, L \tag{6.9.58}$$

在节 6.1 的三种线性最小方差加权准则下,有相应的三种分布式最优融合状态估值器 $\hat{\boldsymbol{x}}_0(t \mid t+N)$ 为

$$\hat{\boldsymbol{x}}_0(t \mid t+N) = \sum_{i=1}^{L}\boldsymbol{\beta}_i\hat{\boldsymbol{x}}_i(t \mid t+N) \tag{6.9.59}$$

其中 $\boldsymbol{\beta}_i$ 分别为加权阵、加权对角阵或加权系数,它们可用节 6.1 的最优融合公式根据由定理 6.9.1 ~ 定理 6.9.5 所求得的 $\boldsymbol{P}_i^x(N)$ 和 $\boldsymbol{P}_{ij}^x(N)$ 来计算,且最优融合误差方差阵 $\boldsymbol{P}_0^x(N)$ 满足精度关系

$$\mathrm{tr}\boldsymbol{P}_0^x(N) \leqslant \mathrm{tr}\boldsymbol{P}_i^x(N), \quad i = 1, \cdots, L \tag{6.9.60}$$

证明 由节 6.1 的融合公式得证. □

【注 6.9.1】 (6.9.26) ~ (6.9.29)系子系统 Kalman 估值器,也可将其化为相应的 Wiener 状态估值器[35]. 这样便可得到信息融合 Kalman 估值器或 Wiener 估值器.

6.9.2　在典范型 Ⅱ 下的广义系统分布式信息融合降阶状态估值器

考虑广义系统(6.9.1)和(6.9.2),在假设 1 ~ 4 下,设 $\mathrm{rank}M = n_1 < n$,则存在非异矩阵[45] $\boldsymbol{P},\boldsymbol{Q}$ 使

$$\boldsymbol{PMQ} = \begin{bmatrix} \boldsymbol{M}_1 & \boldsymbol{0} \\ \boldsymbol{M}_2 & \boldsymbol{0} \end{bmatrix},\quad \boldsymbol{P\Phi Q} = \begin{bmatrix} \boldsymbol{T}_1 & \boldsymbol{0} \\ \boldsymbol{T}_2 & \boldsymbol{T}_3 \end{bmatrix} \tag{6.9.61}$$

其中 $\boldsymbol{M}_1$ 为 $n_1 \times n_1$ 非异下三角阵,$\boldsymbol{T}_1$ 为 $n_1 \times n_1$ 拟下三角矩阵,$\boldsymbol{T}_3$ 为 $n_2 \times n_2$ 非异下三角阵,$n = n_1 + n_2$. 记

$$\boldsymbol{H}_i\boldsymbol{Q} = [\boldsymbol{H}_{i1},\boldsymbol{H}_{i2}],\quad \boldsymbol{P\Gamma} = \begin{bmatrix} \boldsymbol{\Gamma}_1 \\ \boldsymbol{\Gamma}_2 \end{bmatrix} \tag{6.9.62}$$

引入变换

$$\boldsymbol{x}(t) = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{x}_1(t) \\ \boldsymbol{x}_2(t) \end{bmatrix} \tag{6.9.63}$$

其中 $\boldsymbol{x}_1(t) \in R^{n_1}$,$\boldsymbol{x}_2(t) \in R^{n_2}$,则原系统为如下典范型 Ⅱ:

$$\begin{bmatrix} \boldsymbol{M}_1 & \boldsymbol{0} \\ \boldsymbol{M}_2 & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{x}_1(t+1) \\ \boldsymbol{x}_2(t+1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{T}_1 & \boldsymbol{0} \\ \boldsymbol{T}_2 & \boldsymbol{T}_3 \end{bmatrix}\begin{bmatrix} \boldsymbol{x}_1(t) \\ \boldsymbol{x}_2(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Gamma}_1 \\ \boldsymbol{\Gamma}_2 \end{bmatrix}\boldsymbol{w}(t) \tag{6.9.64}$$

$$\boldsymbol{y}(t) = [\boldsymbol{H}_{i1},\boldsymbol{H}_{i2}]\begin{bmatrix} \boldsymbol{x}_1(t) \\ \boldsymbol{x}_2(t) \end{bmatrix} + \boldsymbol{v}_i(t),\quad i = 1,\cdots,L \tag{6.9.65}$$

即

$$\boldsymbol{M}_1\boldsymbol{x}_1(t+1) = \boldsymbol{T}_1\boldsymbol{x}_1(t) + \boldsymbol{\Gamma}_1\boldsymbol{w}_1(t) \tag{6.9.66}$$

$$\boldsymbol{M}_2\boldsymbol{x}_1(t+1) = \boldsymbol{T}_2\boldsymbol{x}_1(t) + \boldsymbol{T}_3\boldsymbol{x}_2(t) + \boldsymbol{\Gamma}_2\boldsymbol{w}(t) \tag{6.9.67}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{H}_{i1}\boldsymbol{x}_1(t) + \boldsymbol{H}_{i2}\boldsymbol{x}_2(t) + \boldsymbol{v}_i(t),\quad i = 1,\cdots,L \tag{6.9.68}$$

由(6.9.66)有

$$\boldsymbol{x}_1(t+1) = \boldsymbol{M}_1^{-1}\boldsymbol{T}_1\boldsymbol{x}_1(t) + \boldsymbol{M}_1^{-1}\boldsymbol{\Gamma}_1\boldsymbol{w}(t) \tag{6.9.69}$$

由(6.9.67)有

$$\boldsymbol{x}_2(t) = \boldsymbol{T}_3^{-1}\boldsymbol{M}_2\boldsymbol{x}_1(t+1) - \boldsymbol{T}_3^{-1}\boldsymbol{T}_2\boldsymbol{x}_1(t) - \boldsymbol{T}_3^{-1}\boldsymbol{\Gamma}_2\boldsymbol{w}_1(t) \tag{6.9.70}$$

将它代入(6.9.68),有子系统

$$\boldsymbol{x}_1(t+1) = \boldsymbol{A}_0\boldsymbol{x}_1(t) + \boldsymbol{\Gamma}_0\boldsymbol{w}(t) \tag{6.9.71}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{H}_{i0}\boldsymbol{x}_1(t) + \boldsymbol{\eta}_i(t),\quad i = 1,\cdots,L \tag{6.9.72}$$

其中定义

$$\boldsymbol{A}_0 = \boldsymbol{M}_1^{-1}\boldsymbol{T}_1 \tag{6.9.73}$$

$$\boldsymbol{\Gamma}_0 = \boldsymbol{M}_1^{-1}\boldsymbol{\Gamma}_1 \tag{6.9.74}$$

$$\boldsymbol{H}_{i0} = \boldsymbol{H}_{i1} - \boldsymbol{H}_{i2}\boldsymbol{T}_3^{-1}\boldsymbol{M}_2\boldsymbol{M}_1^{-1}\boldsymbol{T}_1 - \boldsymbol{H}_{i2}\boldsymbol{T}_3^{-1}\boldsymbol{T}_2 \tag{6.9.75}$$

$$\boldsymbol{\eta}_i(t) = \boldsymbol{G}_i\boldsymbol{w}(t) + \boldsymbol{v}_i(t),\quad i = 1,\cdots,L \tag{6.9.76}$$

$$\boldsymbol{G}_i = \boldsymbol{H}_{i2}\boldsymbol{T}_3^{-1}\boldsymbol{M}_2\boldsymbol{M}_1^{-1}\boldsymbol{\Gamma}_1 - \boldsymbol{H}_{i2}\boldsymbol{T}_3^{-1}\boldsymbol{\Gamma}_2 \tag{6.9.77}$$

显然 $\boldsymbol{w}(t)$ 与 $\boldsymbol{\eta}_i(t)$ 是带零均值的相关白噪声:

$$\mathrm{E}\left\{\begin{bmatrix} \boldsymbol{w}(t) \\ \boldsymbol{\eta}_i(t) \end{bmatrix} \begin{bmatrix} \boldsymbol{w}^{\mathrm{T}}(k) & \boldsymbol{\eta}^{\mathrm{T}}(k) \end{bmatrix}\right\} = \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{S}_{\eta i} \\ \boldsymbol{S}_{\eta i}^{\mathrm{T}} & \boldsymbol{Q}_{\eta i} \end{bmatrix} \delta_{tk} \tag{6.9.78}$$

其中

$$\boldsymbol{S}_{\eta i} = \mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\eta}_i^{\mathrm{T}}(t)] = \boldsymbol{S}_i + \boldsymbol{Q}_w \boldsymbol{G}_i^{\mathrm{T}} \tag{6.9.79}$$

$$\boldsymbol{Q}_{\eta i} = \mathrm{E}[\boldsymbol{\eta}^i(t)\boldsymbol{\eta}_i^{\mathrm{T}}(t)] = \boldsymbol{G}_i \boldsymbol{Q}_w \boldsymbol{G}_i^{\mathrm{T}} + \boldsymbol{R}_i \tag{6.9.80}$$

可以证明在假设 1 ~ 4 下 $(\boldsymbol{A}_0, \boldsymbol{H}_{i0})$ $(i = 1, \cdots, L)$ 为完全可观对[34].

对于多传感器常规系统 (6.9.71) 和 (6.9.72)，用本章方法可求得 $\boldsymbol{x}_1(t)$ 的局部状态估值器 $\hat{\boldsymbol{x}}_{1i}(t \mid t + N)$ $(i = 1, \cdots, L)$ 和 $\boldsymbol{w}(t)$ 的局部白噪声估值器 $\hat{\boldsymbol{w}}_i(t \mid t + N)$，进而由 (6.9.70) 可求得 $\boldsymbol{x}_2(t)$ 的局部状态估值器

$$\begin{aligned} \hat{\boldsymbol{x}}_{2i}(t \mid t + N) = {} & \boldsymbol{T}_3^{-1}\boldsymbol{M}_2 \hat{\boldsymbol{x}}_{1i}(t \mid t + N) - \boldsymbol{T}_3^{-1}\boldsymbol{T}_2 \hat{\boldsymbol{x}}_{1i}(t \mid t + N) - \\ & \boldsymbol{T}_3^{-1}\boldsymbol{\Gamma}_2 \hat{\boldsymbol{w}}_i(t \mid t + N), \quad i = 1, \cdots, L \end{aligned} \tag{6.9.81}$$

于是由 (6.9.63) 有局部原始状态估值器

$$\hat{\boldsymbol{x}}_i(t \mid t + N) = \boldsymbol{Q} \begin{bmatrix} \hat{\boldsymbol{x}}_{1i}(t \mid t + N) \\ \hat{\boldsymbol{x}}_{2i}(t \mid t + N) \end{bmatrix}, \quad i = 1, \cdots, L \tag{6.9.82}$$

类似于在典范型 Ⅰ 下信息融合状态估值器的推导可得相应的三种分布式最优加权融合状态估值器. 详细推导留给读者.

参 考 文 献

1 Bar - Shalom Y，Fortmann T E. Tracking and Data Association. Academic Press，1988

2 Hall D L. Mathematical Techniques in. Multisensor Data Fusion. Artech House，Boston，London，1992

3 Bar - Shalom Y，Li X R. Multitarget - Multisensor Tracking: Principles and Techniques. Stors，CT: YBS Publishing，1995

4 Bar - Shalom Y，Li X R. Estimation and Tracking: Principle，Techniques and Software. Boston，MA: Artech House，INC，1993

5 何友，王国宏，陆大俭，彭应宁. 多传感器信息融合及应用. 北京：电子工业出版社，2000

6 杨万海. 多传感器数据融合及其应用. 西安：西安电子科技大学出版社，2004

7 Klein L A. 多传感顺数据融合理论及应用 (第二版). 北京：北京理工大学出版社，2004

8 Gan Q，Harris C J. Comparison of Two Measurement Fusion Methods for Kalman Filter - Based Multisensor Data Fusion. IEEE Trans. Aerospace and Electronic Systems，2001，37 (1)：273 ~ 280

9 Carlson N A. Federated Square Root Filter for Decentralized Parallel Processes. IEEE Trans. Aerospace Electronic Systems，1990，AES - 26，517 ~ 525

10 Carlson N A. Federated Filter for Fault - Tolerant Integated Navigation Systems. Proc. of IEEE PLANS' 88，Orlando，FL，1988. 110 ~ 119

11 Loomis PVW，Carlson N A. Berarducci M P. Common Kalman Filter: Fault - Tolerant Navigation for Next Generation Aircraft. Proc. of the Inst. of Navigation Conf.，Santa Barbara，CA，1988. 38 ~ 45

12 Kim K H. Development of Track to Track Fusion Algorithm Proceedings of the American Control Conference，Maryland，June 1994. 1037 ~ 1041

13 邓自立，祁荣宾. 多传感器信息融合次优稳态 Kalman 滤波器. 中国学术期刊文摘 (科技快报)，2000，6 (2)：183 ~ 184

14 Sun Shuli，Deng Zili. Muliti - sensor Optimal Information Fusion Kalman Filter. Automatica，2004，40，

1017 ~ 1023
15 邓自立.自校正滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2003
16 孙书利,邓自立.多传感器线性最小方差最大信息融合准则.科学技术与工程,2004,4(5):334 ~ 336
17 邓自立,高媛,崔崇信.多传感器按对角阵加权信息融合 Kalman 滤波器.科学技术与工程,2004,4(7):519 ~ 521
18 高缓,白敬刚,邓自立.多传感器单通道信息融合 Wiener 滤波器.科学技术与工程,2004,4(7):522 ~ 525
19 邓自立,高媛.两传感器信息融合超前 k 步稳态最优 Kalman 预报器.科学技术与工程,2004,4(5):337 ~ 340
20 邓自立,崔崇信,白敬刚.基于稳态 Kalman 滤波的两种观测融合方法的功能等价性.科学技术与工程,2004,4(11):897 ~ 902
21 邓自立,高媛.快速信息融合 Kalman 滤波器.控制与决策,2005,20(1):27 ~ 31
22 邓自立,崔崇信.多传感器全局最优观测融合白噪声反卷积滤波器.科学技术与工程,2005,5(5):267 ~ 270
23 邓自立,李云,王欣.带 MA 有色观测噪声的单通道多传感器信息融合 Wiener 滤波器.科学技术与工程,2005,5(5):271 ~ 275
24 邓自立,高媛,李云等.基于 Kalman 滤波的信息融合白噪声最优反卷积滤波器.科学技术与工程,2004,4(3):161 ~ 171
25 邓自立,高媛,李云等.信息融合稳态最优 Kalman 平滑器.科学技术与工程,204,4(3):172 ~ 175
26 高媛,王琳,梁佐江,邓自立.两传感器按对角阵加权信息融合稳态 Kalman 滤波器.黑龙江大学自然科学学报,2004,21(2):52 ~ 54
27 邓自立,高媛.基于 Kalman 滤波的 ARMA 信号信息融合 Wiener 滤波器.控制理论与应用,2005,22(4)
28 邓自立,高媛,马建为.两传感器信息融合稳态最优 Kalman 滤波器.科学技术与工程,2003,3(3):213 ~ 215
29 邓自立,高媛,马建为.两传感器信息融合最优白噪声反卷积 Wiener 滤波器.科学技术与工程,2003,3(3):216 ~ 218
30 邓自立,马建为,高媛.两传感器自校正信息融合 Kalman 滤波器.科学技术与工程,2003,3(4):321 ~ 324
31 邓自立,马建为,高媛,两传感器自校正信息融合白噪声 Wiener 反卷积滤波器.科学技术与工程,2003.3(4):325 ~ 327
32 邓自立,高媛,王琳,王欣.两传感器信息融合 Wiener 滤波器、平滑器和预报器.2004 中国控制与决策学术年会论文集.沈阳:东北大学出版社,2004.206 ~ 208
33 高媛,梁佐江,王欣,邓自立.两传感器信息融合 Wiener 反卷积滤波器.2004 中国控制与决策学术年会论文集.沈阳:东北大学出版社,2004.212 ~ 214
34 邓自立.最优滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2000
35 邓自立.卡尔曼滤波与维纳滤波——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2001
36 Gao J B. Harris C J. Some Remarks on Kalman Filters for the Multisensor Fusion. Information Fusion,2002,3:191 ~ 201
37 Sun S L. Multi – sensor Information Fusion Waite Noise Filter weighted by scalars based on Kalman Predictor. Automatica,2004,40,1447 ~ 1453
38 Sun S L. Multi – sensor Optimal Information Kalman Filters with Applications. Aerospace Science and Fechnology,2004,8,57 ~ 62

39 孙书利，崔平远.多传感器标量加权最优信息融合稳态 Kalman 滤波器.控制与决策，2004，19（2）：208～211
40 戴华.矩阵论.北京：科学出版社，2001
41 Mendel. J M. White Noise Estimaors for Seismic Data Processing in Oil exploration. IEEE Trans. Automatic Control，1977，22（5）：694～706
42 Deng Z L，Zhang H S，Liu S J，Zhou L. Optimal and Self－tuning White Noise Estimators with Applications to Deconvolution and Filtering Problems. Automatica，1996，32（2）：199～216
43 Anderson B D O，Moore J B. Optimal Filtering. Englewood Cliffs，New Jersey：Prentice－Hall，1979
44 程云鹏.矩阵论（第二版）.西安：西北工业大学出版社，2001
45 Syrmos V L，Lewis F L. Robust Eigenvalue Assignment for Generalized Systems. Automatica，1992，28（6）：1223～1228
46 孟华，石莹，邓自立.带位置和速度观测的信息融合 Kalman 跟踪滤波器.科学技术与工程，2005，5（1）：14～18
47 邓自立，王琳，高媛.多传感器最优信息融合稳态 Kalman 滤波器.科学技术与工程，2004，4（9）：743～748
48 邓自立，高媛，张明波.ARMA 信号自校正信息融合 Wiener 滤波器.科学技术与工程，2004，4（9）：749～752
49 邓自立，石莹，孟华.基于 ARMA 新息模型与 Riccati 方程的两种 Kalman 跟踪滤波器的等价性.科学技术与工程，2004，4（11）：894～896
50 邓自立.Lyapunov 方程迭代解的指数收敛性.科学技术与工程，2005，5（9）：543～546
51 Bar－Shalom，Y. and L. Campo（1986）. The Effect of Common Process Noise on the Two－Sensor Fused－Track Covariance. IEEE Transactions on Aerospace and Electronic Systems，22（6）：803～805
52 Chang，K C，Zhi，T and R K Saha（2002）. Performance Evaluation of Track Fusion with Information Matrix Filter，IEEE Transactions on Aerospace and Electronic Systems，38（2）：455～465
53 Chang K C，Saha R，and Y Bar－Shalom（1997）. On Optimal Track－to－Track Fusion. IEEE Transactions on Aerospace and Electronic Systems. 33（4）：1271～1276
54 Chen H，Kirubarajan，T and Y Bar－Shalom（2003）. Performance Limits of Track－to－Tack Fusion Versus Centralized Estimation：Theory and Application. IEEE Transactions on Aerospace and Electronic Systems，39（2）：386～398
55 Hashemipour H R，Roy S and A J Laub（1998）. Decentralized Structures for Parallel Kalman Filtering. IEEE Transactions on Automatic Control，33（1）：88～93
56 Kailath T，A H Sayed and B Hassibi（2000），Linear Estimation. Upper Saddle River，New Jersey：Prentice－Hall
57 Li X R，Y M Zhu，J Wang，and C Z Han（2003）. Optimal Linear Estimation Fusion－Part. I：Unified Fusion Rules. IEEE Tansactions on Information Theory，49（9）：2192～2208
58 Mori S，W H Barker，C Y Chong，and K C Chang（2002）. Track Association and Track Fusion with Nondeterministic Arget Dynamics. IEEE Transactions on Aerospace and Electronic Systems，38（2）：659～668
59 Paik B S，and J H Oh（2000）. Gain Fusion Algorithm for Decentralised Parallel Kalman Filters. IEEE Proc. Control theory Appl. 147（1）：97～103
60 Roecker J A，and C D McGillem（1988）. Comparison of Two－Sensor Tracking Method Based on State Vector Fusion and Measurement Fusion. IEEE Transaction on Aerospace and Electronic Systems，24（4）：447～449
61 Saha R K and K C Chang（1998）. Efficient Algorithm for Multisensor Track Fusion. IEEE Transactions on Aerospace and Electronic Systems，34（1）：200～210

62 Zhu Y M，Z H You J Zhao，K S Zhang，and X R Li（2001）. The Optimality for the Distributed Kalman Filtering Fusion with feedback. Automatica，37，1489～1493
63 邓自立，梁佐江.多传感器分布式信息融合 Wiener 滤波器.科学技术与工程，2005，5（9）：539～542
64 Hong Z Q，Hong Y Z，Hong J. Fusion Algorithm of Correlated Local Estimates. Aerospace Science and Technology，2004，8：619～626
65 邓自立，孙小君.多传感器分布式协方差信息融合 Kalman 滤波理论.科学技术与工程，2005，5（11）
66 邓自立，郝钢，吴孝慧.两种加权观测融合算法的全局最优性和完全功能等价性.科学技术与工程，2005，5（13）
67 邓自立，陶贵丽.广义系统多传感器信息融合降阶状态估值器.科学技术与工程，2005，5（13）
68 邓自立，王欣，李云.多传感器最优信息融合白噪声反卷积滤波器.电子学报，2005，33（5）：104～107

第七章　基于现代时间序列分析方法的协方差信息融合滤波理论

在第六章我们提出了基于经典 Kalman 滤波的统一的、通用的和实用的最优信息融合滤波理论.它的基本工具是经典 Kalman 滤波方法和 Riccati 方程.所谓统一性是指该理论既可处理时变系统最优信息融合状态估计问题,又可处理时不变(定常)系统的稳态最优融合状态估计问题;既可处理状态融合估计问题,又可处理信号融合估计问题;既可处理常规系统最优融合状态估计问题,又可处理广义系统融合状态估计问题;不仅可处理信息融合状态或信号滤波问题,而且可处理状态或信号的融合预报和平滑问题,所谓通用性是指该理论可带输入噪声和观测噪声相关和带相关的各传感器观测噪声的多传感器系统信息融合最优估计问题,而且提出了适用于不同具体情况的三种最优信息融合准则及相应的最优融合公式.所谓实用性是指该理论可直接应用于解决目标跟踪,石油地震勘探等许多实际最优融合估计问题.该理论填补了在信息融合状态和信号估计领域中的许多空白.许多在文献中未被研究和未被解决的问题首次在该理论被解决.例如信息融合 Kalman 平滑估计问题,信息融合白噪声反卷积问题,信息融合信号估计问题,信息融合广义系统状态估计问题等.最优融合估计的关键问题和难点问题在于各传感器局部估计误差方差阵和互协方差阵的计算,该理论成功地解决了这个难题.

本章我们将提出基于第五章介绍的现代时间序列分析方法的新的最优信息融合滤波理论.它的基本工具是 ARMA 新息模型和 Lyapunov 方程,它避免了 Riccati 方程.该理论仅适用于处理定常(时不变)系统的稳态最优信息融合状态或信号估计问题,它不能处理时变系统.但该理论的优点是不仅可处理用经典 Kalman 滤波方法能解决的最优融合估计问题,可得到完全不同形式的新结果,而且还可处理用经典 Kalman 滤波方法不容易解决和未解决的问题,可发现许多新结果.该理论的特色是基于 ARMA 新息模型计算稳态 Kalman 估值器增益阵,避免了 Riccati 方程,且通过 Lyapunov 方程来计算各传感器局部状态或信号估计误差方差阵和协方差阵或直接通过计算输入白噪声、观测白噪声和新息三者之间的互协方差阵来计算局部状态或信号估计误差方差阵和互协方差阵.该理论另一优点是可处理含未知模型参数和噪声统计系统的自校正信息融合状态或信号估值器[1,19,21,22].

该理论的基本框架首先在文献[1]中被提出,有关最新结果见文献[8～16,18～22,25～35].特别,文献[35]系统地提出了信息融合稳态 Kalman 滤波理论.

由于在现代时间序列分析方法中白噪声估计的重要性,它是信号和状态估计的基础,本章提出的信息融合估计理论从白噪声融合估计开始,进而推广到信号融合估计,最后推广到状态融合估计.本章在统一框架下解决多传感器系统信号或状态的局部估计误差方差阵和互协方差阵计算问题,将该问题归结为简单地、统一地计算输入白噪声、观测白噪

声、新息过程三者间的互协方差阵，即最终转化为三种白噪声的互协方差阵的计算.这个问题可基于 ARMA 新息模型来解决.在处理 Kalman 估值器时，本章还提出了一种基于 ARMA 新息模型计算稳态 Kalman 估值器增益阵和基于 Lyapunov 方程求多传感系统局部估计误差互协方差阵的新方法.该方法完全不同于经典 Kalman 滤波方法基于 Riccati 方程求稳态 Kalman 估值器和计算局部估计误差互协方差阵.由于 Lyapunov 方程可用迭代法简单地求解，且迭代解具有指数收敛速度[33]，因而用迭代法可快速求解.由于本章研究信息融合稳态估计问题，在节 6.1 的三种加权准则下，最优加权均为常量(常矩阵、常对角阵或常系数)，因而相应的最优信息融合估值器的计算较为简单，便于实时应用.因而本章理论具有较大实用价值.我们将提出本章结果在目标跟踪系统中的仿真应用，说明其有效性.

综上所述，本章用现代时间序列分析方法首次系统提出了多传感器协方差信息融合理论.该理论包括节 6.1 提出基于估计误差协方差的三种最优信息融合加权估计准则及对白噪声、信号及状态估计误差协方差计算的方法和公式，以及信息融合估计算法.所谓估计误差协方差包括局部估计误差的自协方差和局部估计误差之间的互协方差.协方差信息融合理论的基本原理是根据局部估计误差协方差信息按三种最优加权准则可求得最优加权，进而可求得由局部估计进行最优加权所得到的最优信息融合估计，也叫分布式融合估计.应用现代时间序列分析方法本章提出了计算局部估计误差协方差的统一框架，即将问题归结为计算输入白噪声、观测白噪声和新息三者之间的协方差.这个问题容易通过 ARMA 新息模型来解决.应指出，三种加权 Kalman 融合器是全局次优的.[35]

7.1 多传感器信息融合白噪声估值器

考虑多传感器线性离散随机控制系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{Bu}(t)+\boldsymbol{\Gamma w}(t) \tag{7.1.1}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{H}_i\boldsymbol{x}(t)+\boldsymbol{v}_i(t),\quad i=1,\cdots,L \tag{7.1.2}$$

其中 t 为离散时间，状态 $\boldsymbol{x}(t)\in R^n$，第 i 个传感器的观测 $\boldsymbol{y}_i(t)\in R^{m_i}$，控制 $\boldsymbol{u}(t)\in R^p$，$\boldsymbol{v}_i(t)$ 为观测白噪声，$\boldsymbol{w}(t)$ 为输入白噪声，$\boldsymbol{\Phi},\boldsymbol{B},\boldsymbol{\Gamma},\boldsymbol{H}_i$ 为已知适当维数常数.

【假设 1】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}_i(t)\in R^{m_i}$ 是带零均值的相关白噪声

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}_i(t)\end{bmatrix}\left[\boldsymbol{w}^{\mathrm{T}}(k),\boldsymbol{v}_i^{\mathrm{T}}(k)\right]\right\}=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}_i\\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_i\end{bmatrix}\delta_{tk} \tag{7.1.3}$$

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{v}_i(t)\\ \boldsymbol{v}_j(t)\end{bmatrix}\left[\boldsymbol{v}^{\mathrm{T}}(k),\boldsymbol{v}_j^{\mathrm{T}}(k)\right]\right\}=\begin{bmatrix}\boldsymbol{R}_i & \boldsymbol{R}_{ij}\\ \boldsymbol{R}_{ij}^{\mathrm{T}} & \boldsymbol{R}_j\end{bmatrix}\delta_{tk} \tag{7.1.4}$$

且定义 $\boldsymbol{R}_{ii}=\boldsymbol{R}_i$.

【假设 2】 $\boldsymbol{u}(t)$ 是已知的时间序列，或 $\boldsymbol{u}(t)$ 是 $(\boldsymbol{y}_i(t),\boldsymbol{y}_i(t-1),\cdots)$ 的函数(反馈控制).

【假设 3】 初始观测时刻 $t_0=-\infty$.

该状态空间模型的输出 $\boldsymbol{y}_i(t)$ 可写为

$$\boldsymbol{y}_i(t)=\left[\boldsymbol{H}_i(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{B}q^{-1},\boldsymbol{H}_i(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}\right]\begin{bmatrix}\boldsymbol{u}(t)\\ \boldsymbol{w}(t)\end{bmatrix}+\boldsymbol{v}_i(t) \tag{7.1.5}$$

其中 q^{-1} 为单位滞后算子.引入多项矩阵左素分解

$$[\boldsymbol{H}_i(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{B}q^{-1}, \boldsymbol{H}_i(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}] = \boldsymbol{A}_i^{-1}(q^{-1})[\boldsymbol{B}_i(q^{-1}), \boldsymbol{P}_i(q^{-1})] \tag{7.1.6}$$

其中 $\boldsymbol{X}_i(q^{-1}) = \boldsymbol{X}_{i0} + \boldsymbol{X}_{i1}q^{-1} + \cdots + \boldsymbol{X}_{in_x}q^{-n_{xi}}$,且有 $\boldsymbol{A}_{i0} = \boldsymbol{I}_{m_i}, \boldsymbol{B}_{i0} = \boldsymbol{0}, \boldsymbol{P}_{i0} = \boldsymbol{0}$.

【假设 4】[3] $(\boldsymbol{A}_i(q^{-1}), \boldsymbol{P}_i(q^{-1}))$ 无零点位于单位圆上的左因式.

将(7.1.6)代入(7.1.5)引出第 i 个传感器的 MA 新息模型为

$$\boldsymbol{z}_i(t) = \boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t), \quad i = 1,\cdots,L \tag{7.1.7}$$

其中定义新的观测为

$$\boldsymbol{z}_i(t) = \boldsymbol{A}_i(q^{-1})\boldsymbol{y}_i(t) - \boldsymbol{B}_i(q^{-1})\boldsymbol{u}(t)$$

而 $\boldsymbol{D}_i(q^{-1})$ 是稳定的,且 $\boldsymbol{D}_{i0} = \boldsymbol{I}_{m_i}$,新息 $\boldsymbol{\varepsilon}_i(t) \in R^{m_i}$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声,且有关系

$$\boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t) = \boldsymbol{P}_i(q^{-1})\boldsymbol{w}_i(t) + \boldsymbol{A}_i(q^{-1})\boldsymbol{v}_i(t), \quad i = 1,\cdots,L \tag{7.1.8}$$

其中 $\boldsymbol{D}_i(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers – Wouters 算法求得.

问题是基于局部 MA 新息模型(7.1.7)求局部白噪声反卷积估值器 $\hat{\boldsymbol{w}}_i(t \mid t + N)(N = 0, N > 0$ 或 $N < 0)$,并求最优信息融合白噪声反卷积估值器 $\hat{\boldsymbol{w}}_0(t \mid t + N)$.

注意,当 $\boldsymbol{u}(t) = \boldsymbol{0}$ 时,这一问题在石油地震勘探[23]有重要应用背景.当 $\boldsymbol{B} = \boldsymbol{\Gamma}$,$\boldsymbol{u}(t)$ 是已知的运动目标的机动加速度,$\boldsymbol{w}(t)$ 是它的随机加速度,这一问题相当于求随机加速度估计.

【引理 7.1.1】 多传感器系统(7.1.1)和(7.1.2)在假设 1 ~ 4 下,第 i 个传感器有局部稳态最优白噪声反卷积估值器

$$\hat{\boldsymbol{w}}_i(t \mid t + N) = \sum_{k=0}^{N}\boldsymbol{M}_k^{iw}\boldsymbol{\varepsilon}_i(t + k), \quad N \geqslant 0, i = 1,\cdots,L \tag{7.1.9}$$

$$\hat{\boldsymbol{w}}_i(t \mid t + N) = \boldsymbol{0}, \quad N < 0, i = 1,\cdots,L \tag{7.1.10}$$

其中定义

$$\boldsymbol{M}_k^{iw} = \boldsymbol{\Lambda}_k^{iw}\boldsymbol{Q}_{\varepsilon i}^{-1}, \quad i = 1,\cdots,L \tag{7.1.11}$$

$$\boldsymbol{\Lambda}_k^{iw} = \boldsymbol{Q}_w\boldsymbol{F}_k^{(i)\mathrm{T}} + \boldsymbol{S}_i\boldsymbol{G}_k^{(i)\mathrm{T}}, \quad i = 1,\cdots,L \tag{7.1.12}$$

其中上标 i 为传感器序号,而系数阵 $\boldsymbol{F}_k^{(i)}$ 和 $\boldsymbol{G}_k^{(i)}$ 可递推计算为

$$\boldsymbol{F}_k^{(i)} = -\boldsymbol{D}_{i1}\boldsymbol{F}_{k-1}^{(i)} - \cdots - \boldsymbol{D}_{in_{di}}\boldsymbol{F}_{k-n_{di}}^{(i)} + \boldsymbol{P}_{ik} \tag{7.1.13}$$

$$\boldsymbol{G}_k^{(i)} = -\boldsymbol{D}_{i1}\boldsymbol{G}_{k-1}^{(i)} - \cdots - \boldsymbol{D}_{in_{di}}\boldsymbol{G}_{k-n_{di}}^{(i)} + \boldsymbol{A}_{ik} \tag{7.1.14}$$

其中定义 $\boldsymbol{F}_k^{(i)} = \boldsymbol{0}(k < 0), \boldsymbol{G}_k^{(i)} = \boldsymbol{0}(k < 0), \boldsymbol{P}_{ik} = \boldsymbol{0}(k > n_{pi}), \boldsymbol{A}_{ik} = \boldsymbol{0}(k > n_{ai})$.定义

$$\boldsymbol{L}_N^{iw}(q^{-1}) = \sum_{k=0}^{N}\boldsymbol{M}_k^{iw}q^{k-N}, \quad N \geqslant 0,$$

$$\boldsymbol{L}_N^{iw}(q^{-1}) = \boldsymbol{0}, \quad N < 0 \tag{7.1.15}$$

则(7.1.9)和(7.1.10)可写为新息滤波器形式

$$\hat{\boldsymbol{w}}_i(t \mid t + N) = \boldsymbol{L}_N^{iw}(q^{-1})\boldsymbol{\varepsilon}_i(t + N) \tag{7.1.16}$$

相应的稳态估值误差方差阵为

$$P_i^w(N) = Q_w - \sum_{k=0}^{N} M_k^{iw} Q_{\varepsilon i} M_k^{iw\mathrm{T}}, \quad N \geqslant 0, \quad i = 1, \cdots, L \tag{7.1.17}$$

$$P_i^w(N) = Q_w, \quad N < 0, i = 1, \cdots, L \tag{7.1.18}$$

证明　见定理 5.1.2. □

【定理 7.1.1】　多传感器系统(7.1.1)和(7.1.2)在假设 1 ~ 4 下有局部稳态白噪声估值误差互协方差阵 $P_{ij}^w(N) = \mathrm{E}[\tilde{w}_i(t \mid t+N)\tilde{w}_j^{\mathrm{T}}(t \mid t+N)](i \neq j)$ 为

$$P_{ij}^w(N) = Q_w - \sum_{k=0}^{N} M_k^{iw} Q_{\varepsilon i} M_k^{iw\mathrm{T}} - \sum_{k=0}^{N} M_k^{jw} Q_{\varepsilon j} M_k^{jw\mathrm{T}} + \sum_{r=0}^{N}\sum_{s=0}^{N} M_r^{iw} E_{ij}(r,s) M_s^{jw\mathrm{T}}, \quad i \neq j, N \geqslant 0 \tag{7.1.19}$$

其中 $E_{ij}(r,s) = \mathrm{E}[\varepsilon_i(t+r)\varepsilon_j^{\mathrm{T}}(t+s)](i \neq j)$ 为

$$E_{ij}(r,s) = \sum_{k=0}^{\infty} F_k^{(i)} Q_w F_{k+s-r}^{(j)\mathrm{T}} + \sum_{k=0}^{\infty} F_k^{(i)} S_j G_{k+s-r}^{(j)\mathrm{T}} + \sum_{k=0}^{\infty} G_k^{(i)} S_i^{\mathrm{T}} F_{k+s-r}^{(j)\mathrm{T}} + \sum_{k=0}^{\infty} G_k^{(i)} R_{ij} G_{k+s-r}^{(j)\mathrm{T}} \tag{7.1.20}$$

且当 $N < 0$ 时有

$$P_{ij}^w(N) = Q_w, \quad i \neq j, N < 0 \tag{7.1.21}$$

证明　由定义有

$$P_{ij}^w(N) = \mathrm{E}[(w(t) - \hat{w}_i(t \mid t+N))(w(t) - \hat{w}_j(t \mid t+N))^{\mathrm{T}}] \tag{7.1.22}$$

当 $N \geqslant 0$ 时将 $\hat{w}_i(t \mid t+N)$ 和 $\hat{w}_j(t \mid t+N)$ 的表达式(7.1.9)代入上式后展开为四项.由(5.1.39)和(5.1.40)有关系

$$\mathrm{E}[w(t)\varepsilon_j^{\mathrm{T}}(t+k)] = M_k^{wj} Q_{\varepsilon j} \tag{7.1.23}$$

由此式和(7.1.9)可引出(7.1.19).由(5.1.35),(7.1.8)有展式

$$\varepsilon_i(t) = D_i^{-1}(q^{-1})P_i(q^{-1})w(t) + D_i^{-1}(q^{-1})A_i(q^{-1})v_i(t) = \sum_{k=0}^{\infty} F_k^{(i)} w(t-k) + \sum_{k=0}^{\infty} G_k^{(i)} v_i(t-k) \tag{7.1.24}$$

其中 $F_k^{(i)}$ 和 $G_k^{(i)}$ 由(7.1.13)和(7.1.14)计算.利用此式和(7.1.3)、(7.1.4)可得(7.1.20).注意当 $N < 0$ 时有 $\hat{w}(t \mid t+N) = \mathbf{0}$,故得(7.1.21).证毕. □

【定理 7.1.2】　多传感器系统(7.1.1)和(7.1.2)在假设 1 ~ 4 下有 $w(t)$ 的稳态最优融合白噪声估值器

$$\hat{w}_0(t \mid t+N) = \sum_{i=1}^{L} \alpha_i \hat{w}_i(t \mid t+N) \tag{7.1.25}$$

其中在节 6.1 的三种最优加权融合准则下,α_i 分别为与 N 有关的加权阵、加权对角阵和加权系数,它们可用相应的最优加权公式基于 $P_i^w(N)$ 和 $P_{ij}^w(N)(i \neq j)$ 求得,且最优融合估值误差方差阵 $P_0^w(N)$ 满足关系

$$\mathrm{tr} P_0^w(N) \leqslant \mathrm{tr} P_i^w(N), \quad i = 1, \cdots, L \tag{7.1.26}$$

证明　见节 6.1. □

7.2 多传感器多通道 ARMA 信号信息融合 Wiener 滤波器

带白色观测噪声的多通道 ARMA 信号滤波问题在信号处理、通信、目标跟踪等领域经常会遇到，节 5.3 已给出了它的 Wiener 滤波器. 本节进一步考虑多传感器信息融合估计问题. 考虑带多传感器的多通道 ARMA 信号最优融合滤波问题：

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{7.2.1}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{s}(t) + \boldsymbol{v}_i(t), \quad i = 1,\cdots,L \tag{7.2.2}$$

其中第 i 个传感器的观测 $\boldsymbol{y}_i(t) \in R^m$，待估信号 $\boldsymbol{s}(t) \in R^m$，$\boldsymbol{v}_i(t) \in R^m$ 是观测噪声，且 $\boldsymbol{w}(t)$ 和 $\boldsymbol{v}(t)$ 是零均值、方差各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_{vi}$ 的相互独立白噪声：

$$\mathrm{E}\left\{\begin{bmatrix} \boldsymbol{w}(t) \\ \boldsymbol{v}_i(t) \end{bmatrix}[\boldsymbol{w}^{\mathrm{T}}(k), \boldsymbol{v}_i^{\mathrm{T}}(k)]\right\} = \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_{vi} \end{bmatrix}\delta_{tk} \tag{7.2.3}$$

$$\mathrm{E}[\boldsymbol{v}_i(t)\boldsymbol{v}_j^T(k)] = \boldsymbol{0}, \quad \forall\, i,j,t,k$$

其中 E 为均值号，T 为转置号，$\delta_{tt} = 1$，$\delta_{tk} = 0(t \neq k)$. q^{-1} 为单位滞后算子，多项式矩阵 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{C}(q^{-1})$ 有形式 $\boldsymbol{X}(q^{-1}) = \boldsymbol{X}_0 + \boldsymbol{X}_1 q^{-1} + \cdots + \boldsymbol{X}_{n_x} q^{-n_x}$，$\boldsymbol{A}_0 = \boldsymbol{I}_m$，$\boldsymbol{C}_0 = \boldsymbol{0}$. 假设 $(\boldsymbol{A}(q^{-1}), \boldsymbol{C}(q^{-1}))$ 左素，初始时刻 $t_0 = -\infty$，问题是求 $\boldsymbol{s}(t)$ 的局部滤波器 $\hat{\boldsymbol{s}}_i(t \mid t + N)(i = 1,\cdots,L)$ 和最优融合滤波器 $\hat{\boldsymbol{s}}_0(t \mid t + N)(N = 0, N > 0$ 或 $N < 0)$.

将(7.2.2)代入(7.2.1)可得基于第 i 传感器的局部 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}_i(t) = \boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t), \quad i = 1,\cdots,L \tag{7.2.4}$$

其中 $\boldsymbol{D}_i(q^{-1})$ 是稳定的，$\boldsymbol{D}_i(q^{-1}) = \boldsymbol{D}_{i0} + \boldsymbol{D}_{i1}q^{-1} + \cdots + \boldsymbol{D}_{in_{di}}q^{-n_{di}}$，$\boldsymbol{D}_{i0} = \boldsymbol{I}_m$ 新息 $\boldsymbol{\varepsilon}_i(t) \in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_{\varepsilon i}$ 的白噪声，且

$$\boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) + \boldsymbol{A}(q^{-1})\boldsymbol{v}_i(t) \tag{7.2.5}$$

而 $\boldsymbol{D}_i(q^{-1})$ 和 $\boldsymbol{Q}_{\varepsilon i}$ 可用 Gevers – Wouters 算法求得.

【引理 7.2.1】 第 i 传感器有稳态最优观测预报器

$$\hat{\boldsymbol{y}}_i(t \mid t + N) = \boldsymbol{J}_{-N}^{(i)}(q^{-1})\widetilde{\boldsymbol{D}}_i^{-1}(q^{-1})\boldsymbol{y}_i(t + N), \quad i = 1,\cdots,L \tag{7.2.6}$$

其中 $\boldsymbol{J}_{-N}^{(i)}(q^{-1})$ 定义为

$$\begin{aligned} &\widetilde{\boldsymbol{D}}_i(q^{-1}) = \boldsymbol{E}_{-N}^{(i)}(q^{-1})\widetilde{\boldsymbol{A}}_i(q^{-1}) + q^N\boldsymbol{J}_{-N}^{(i)}(q^{-1}), \quad N < 0, \\ &\boldsymbol{J}_{-N}^{(i)}(q^{-1}) = \widetilde{\boldsymbol{D}}_i(q^{-1})q^{-N}, \quad N \geqslant 0 \end{aligned} \tag{7.2.7}$$

$$\boldsymbol{E}_{-N}^{(i)}(q^{-1}) = \boldsymbol{E}_{-N0}^{(i)} + \boldsymbol{E}_{-N1}^{(i)}q^{-1} + \cdots + \boldsymbol{E}_{-N,-N-1}^{(i)}q^{N+1}, E_{-N0}^{(i)} = \boldsymbol{I}_m, \quad N < 0 \tag{7.2.8}$$

$$\boldsymbol{J}_{-N}^{(i)}(q^{-1}) = \boldsymbol{J}_{-N0}^{(i)} + \boldsymbol{J}_{-N1}^{(i)}q^{-1} + \cdots + \boldsymbol{J}_{-Nn_j}^{(i)}q^{-n_j}, \quad N < 0 \tag{7.2.9}$$

其中 $n_j = \max(n_a - 1, n_{di} + N)$，而

$$\boldsymbol{D}_i^{-1}(q^{-1})\boldsymbol{A}(q^{-1}) = \widetilde{\boldsymbol{A}}_i(q^{-1})\widetilde{\boldsymbol{D}}_i^{-1}(q^{-1}) \tag{7.2.10}$$

且 $\widetilde{\boldsymbol{A}}_{i0} = \boldsymbol{I}_m$，$\widetilde{\boldsymbol{D}}_{i0} = \boldsymbol{I}_m$，且有

$$\det\widetilde{\boldsymbol{D}}_i(q^{-1}) = \det\boldsymbol{D}_i(q^{-1}) \tag{7.2.11}$$

预报误差为

$$\begin{aligned} &\tilde{\boldsymbol{y}}_i(t \mid t + N) = \boldsymbol{y}_i(t) - \hat{\boldsymbol{y}}(t \mid t + N) = \boldsymbol{E}_{-N}^{(i)}(q^{-1})\boldsymbol{\varepsilon}_i(t), \quad N < 0 \\ &\tilde{\boldsymbol{y}}_i(t \mid t + N) = \boldsymbol{0}, \quad N \geqslant 0 \end{aligned} \tag{7.2.12}$$

预报误差方差阵为

$$\boldsymbol{P}_i^{\gamma}(N) = \sum_{k=0}^{-N-1} \boldsymbol{E}_{-Nk}^{(i)} \boldsymbol{Q}_{\varepsilon i} \boldsymbol{E}_{-Nk}^{(i)\mathrm{T}}, \quad N < 0,$$

$$\boldsymbol{P}_i^{\gamma}(N) = \boldsymbol{0}, \quad N \geqslant 0 \tag{7.2.13}$$

证明 见推论 5.3.2. □

【引理 7.2.2】 多传感器系统(7.2.1)~(7.2.3)的第 i 个传感器有局部稳态最优观测白噪声估值器

$$\hat{\boldsymbol{v}}_i(t \mid t+N) = \sum_{k=0}^{N} \boldsymbol{M}_k^{iv} \boldsymbol{\varepsilon}_i(t+k), \quad N \geqslant 0, i = 1, \cdots, L$$

$$\hat{\boldsymbol{v}}_i(t \mid t+N) = \boldsymbol{0}, \quad N < 0, i = 1, \cdots, L \tag{7.2.14}$$

其中系数阵 $\boldsymbol{M}_k^{iv}$ 为

$$\boldsymbol{M}_k^{iv} = \boldsymbol{Q}_{vi} \boldsymbol{G}_k^{(i)\mathrm{T}} \boldsymbol{Q}_{\varepsilon i}^{-1} \tag{7.2.15}$$

而系数阵 $\boldsymbol{G}_k^{(i)}$ 可递推计算为

$$\boldsymbol{G}_k^{(i)} = -\boldsymbol{D}_{i1} \boldsymbol{G}_{k-1}^{(i)} - \cdots - \boldsymbol{D}_{in_{di}} \boldsymbol{G}_{k-n_{di}}^{(i)} + \boldsymbol{A}_k \tag{7.2.16}$$

其中规定 $\boldsymbol{G}_k^{(i)} = \boldsymbol{0}(k < 0)$, $\boldsymbol{A}_i = \boldsymbol{0}(k > n_a)$.

它可表为新息滤波器形式

$$\hat{\boldsymbol{v}}_i(t \mid t+N) = \boldsymbol{L}_N^{iv}(q^{-1}) \boldsymbol{\varepsilon}_i(t+N) \tag{7.2.17}$$

其中上、下标 i 为传感器序号,且定义

$$\boldsymbol{L}_N^{iv}(q^{-1}) = \sum_{k=0}^{N} \boldsymbol{M}_k^{iv} q^{k-N}, \quad N \geqslant 0, i = 1, \cdots, L \tag{7.2.18}$$

$$\boldsymbol{L}_N^{iv}(q^{-1}) = \boldsymbol{0}, \quad N < 0, i = 1, \cdots, L \tag{7.2.19}$$

相应的估值误差方差阵为

$$\boldsymbol{P}_i^{v}(N) = \boldsymbol{Q}_{vi} - \sum_{k=0}^{N} \boldsymbol{M}_k^{iv} \boldsymbol{Q}_{\varepsilon i} \boldsymbol{M}_k^{iv\mathrm{T}}, \quad N \geqslant 0 \tag{7.2.20}$$

$$\boldsymbol{P}_i^{v}(N) = \boldsymbol{Q}_{vi}, \quad N < 0 \tag{7.2.21}$$

其中上标或下标 i 为传感器序号.

证明 见定理 5.1.2. □

【引理 7.2.3】 多传感器系统(7.2.1)~(7.2.3)有 $s(t)$ 的基于第 i 个传感器的局部 Wiener 滤波器

$$\hat{\boldsymbol{s}}_i(t \mid t+N) = \boldsymbol{K}_N^{(i)}(q^{-1}) \widetilde{\boldsymbol{D}}_i^{-1}(q^{-1}) \boldsymbol{y}_i(t+N) \tag{7.2.22}$$

$$\boldsymbol{K}_N^{(i)}(q^{-1}) = \boldsymbol{J}_{-N}^{(i)}(q^{-1}) - \boldsymbol{L}_N^{iv}(q^{-1}) \widetilde{\boldsymbol{A}}_i(q^{-1}) \tag{7.2.23}$$

它有 ARMA 递推形式

$$\det \widetilde{\boldsymbol{D}}_i(q^{-1}) \hat{\boldsymbol{s}}_i(t \mid t+N) = \boldsymbol{K}_N^{(i)}(q^{-1}) \operatorname{adj} \widetilde{\boldsymbol{D}}_i(q^{-1}) \boldsymbol{y}_i(t+N) \tag{7.2.24}$$

且有滤波误差方差阵

$$\boldsymbol{P}_i^{s}(N) = \boldsymbol{Q}_{vi} - \sum_{k=0}^{N} \boldsymbol{Q}_{vi} \boldsymbol{G}_k^{(i)\mathrm{T}} \boldsymbol{Q}_{\varepsilon i}^{-1} \boldsymbol{G}_k^{(i)} \boldsymbol{Q}_{vi}, \quad N \geqslant 0 \tag{7.2.25}$$

$$\boldsymbol{P}_i^{s}(N) = \sum_{k=0}^{-N-1} \boldsymbol{E}_{-Nk}^{(i)} \boldsymbol{Q}_{\varepsilon i} \boldsymbol{E}_{-Nk}^{(i)\mathrm{T}} - \boldsymbol{Q}_{vi}, \quad N < 0 \tag{7.2.26}$$

证明 见定理 5.3.1. □

【定理 7.2.1】 多传感器系统(7.2.1) ~ (7.2.3)有局部稳态估值误差 $\tilde{\boldsymbol{s}}_i(t \mid t + N) = \boldsymbol{s}(t) - \hat{\boldsymbol{s}}_i(t \mid t + N)$ 的互协方差阵 $\boldsymbol{P}_{ij}^s(N) = \mathrm{E}[\tilde{\boldsymbol{s}}_i(t \mid t + N)\tilde{\boldsymbol{s}}_j^{\mathrm{T}}(t \mid t + N)](i \neq j)$ 为

$$\boldsymbol{P}_{ij}^s(N) = \sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{M}_r^{iv}\boldsymbol{E}_{ij}(r,s)\boldsymbol{M}_s^{jv\mathrm{T}}, \quad N \geqslant 0, i \neq j \tag{7.2.27}$$

其中 $\boldsymbol{E}_{ij}(r,s) = \mathrm{E}[\boldsymbol{\varepsilon}_i(t + r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t + s)]$ 为

$$\boldsymbol{E}_{ij}(r,s) = \sum_{k=0}^{\infty}\boldsymbol{F}_k^{(i)}\boldsymbol{Q}_w\boldsymbol{F}_{k+s-r}^{(j)\mathrm{T}}, \quad i \neq j \tag{7.2.28}$$

$$\boldsymbol{P}_{ij}^s(N) = \sum_{r=0}^{-N-1}\sum_{s=0}^{-N-1}\boldsymbol{E}_{-Nr}^{(i)}\boldsymbol{\Sigma}_{ij}(r,s)\boldsymbol{E}_{-Ns}^{(j)\mathrm{T}}, \quad N < 0, i \neq j \tag{7.2.29}$$

其中 $\boldsymbol{\Sigma}_{ij}(r,s) = \mathrm{E}[\boldsymbol{\varepsilon}_i(t - r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t - s)]$ 为

$$\boldsymbol{\Sigma}_{ij}(r,s) = \sum_{k=0}^{\infty}\boldsymbol{F}_k^{(i)}\boldsymbol{Q}_w\boldsymbol{F}_{k+r-s}^{(j)\mathrm{T}}, \quad i \neq j \tag{7.2.30}$$

证明 当 $N \geqslant 0$ 时,由(7.2.2)有 $\hat{\boldsymbol{s}}_i(t \mid t + N) = \boldsymbol{y}_i(t) - \hat{\boldsymbol{v}}(t \mid t + N)$,

$$\tilde{\boldsymbol{s}}_i(t \mid t + N) = -(\boldsymbol{v}_i(t) - \hat{\boldsymbol{v}}_i(t \mid t + N)) = -\tilde{\boldsymbol{v}}_i(t \mid t + N) \tag{7.2.31}$$

$$\boldsymbol{P}_{ij}^s(N) = \mathrm{E}[(\boldsymbol{v}_i(t) - \hat{\boldsymbol{v}}_i(t \mid t + N))(\boldsymbol{v}_j(t) - \hat{\boldsymbol{v}}_j(t \mid t + N))^{\mathrm{T}}], \quad i \neq j \tag{7.2.32}$$

将(7.2.32)展开为四项并应用(7.2.14)和(7.2.3)得(7.2.27). 由(7.2.5)有展式

$$\boldsymbol{\varepsilon}_i(t) = \boldsymbol{D}_i^{-1}(q^{-1})\boldsymbol{C}(q^{-1})\boldsymbol{w}(t) + \boldsymbol{D}_i^{-1}(q^{-1})\boldsymbol{A}(q^{-1})\boldsymbol{w}_i(t) =$$

$$\sum_{k=0}^{\infty}\boldsymbol{F}_k^{(i)}\boldsymbol{w}(t - k) + \sum_{k=0}^{\infty}\boldsymbol{G}_k^{(i)}\boldsymbol{v}_i(t - k) \tag{7.2.33}$$

其中 $\boldsymbol{F}_k^{(i)}$ 和 $\boldsymbol{G}_k^{(i)}$ 由(7.1.13)和(7.1.14)计算,但应将 $\boldsymbol{P}_{ik}$ 改为 $\boldsymbol{C}_k$. 应用(7.2.3)和(7.2.33)得(7.2.28).

当 $N < 0$ 时,由(7.2.2)有 $\hat{\boldsymbol{s}}_i(t \mid t + N) = \hat{\boldsymbol{y}}_i(t \mid t + N)$,故

$$\tilde{\boldsymbol{s}}_i(t \mid t + N) = \tilde{\boldsymbol{y}}_i(t \mid t + N) - \boldsymbol{v}_i(t) \tag{7.2.34}$$

$$\boldsymbol{P}_{ij}^s(N) = \mathrm{E}[(\tilde{\boldsymbol{y}}_i(t \mid t + N) - \boldsymbol{v}_i(t)(\tilde{\boldsymbol{y}}_j(t \mid t + N) - \boldsymbol{v}_j(t))^{\mathrm{T}}] \tag{7.2.35}$$

应用(7.2.3)和(7.2.12)得(7.2.29),且应用(7.2.33)得(7.2.30). 证毕. □

【注 7.2.1】 注意无穷级数(7.2.28)和(7.2.30)在计算中可取其前有限项之和作为级数和伪近似值. 可证明系数阵 $\boldsymbol{F}_k^{(i)}$ 当 $k \to \infty$ 时以指数律衰减至零,因此只要取有限项的数目充分大,用级数的前有限项之和就可任意逼近级数和的真实值.

【定理 7.2.2】 多传感器多通道 ARMA 信号(7.2.1) ~ (7.2.3)有最优信息融合 Wiener 滤波器

$$\hat{\boldsymbol{s}}_0(t \mid t + N) = \sum_{I=1}^{L}\boldsymbol{\alpha}_i\hat{\boldsymbol{s}}_i(t \mid t + N) \tag{7.2.36}$$

其中局部 Wiener 滤波器 $\hat{\boldsymbol{s}}_i(t \mid t + N)$ 由引理 7.2.3 计算. 在节 6.1 的三种最优加权融合准则下,$\boldsymbol{\alpha}_i$ 分别为与 N 有关的加权阵、加权对角阵和加权系数,分别可用相应的最优加权公式,由引理 7.2.3 和定理 7.2.1 给出的 $\boldsymbol{P}_i^s(N)$ 和 $\boldsymbol{P}_{ij}^s(N)$ 求得,且最优融合估值误差方差阵 $\boldsymbol{P}_0^s(N)$ 有性质

$$\mathrm{tr}\boldsymbol{P}_0^s(N) \leqslant \mathrm{tr}\boldsymbol{P}_i^s(N), \quad i = 1,\cdots,L \tag{7.2.37}$$

证明 见节 6.1. □

7.3 多传感器信息融合 Wiener 状态估值器

节 5.6 用现代时间序列分析方法提出了统一的 Wiener 状态估值器.方法原理是将状态估计问题转化为白噪声估计和观测预报问题,且将状态估值误差方差阵的计算问题转化为计算白噪声与新息之间的协方差阵问题.本节在此基础上进一步提出计算局部状态估值误差互协方差阵公式.这是解决信息融合状态估计问题的关键技术.方法原理仍为将问题归结为计算各传感器系统白噪声与新息之间的互协方差阵问题.这可基于 ARMA 新息模型来解决.因为 ARMA 新息模型建立了新息与模型白噪声和观测白噪声之间的关系.

考虑多传感器线性离散定常随机系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t) \tag{7.3.1}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{H}_i\boldsymbol{x}(t)+\boldsymbol{v}_i(t),\quad i=1,\cdots,L \tag{7.3.2}$$

其中状态 $\boldsymbol{x}(t)\in R^n$,第 i 个传感器的观测 $\boldsymbol{y}_i(t)\in R^{m_i}$,$\boldsymbol{\Phi},\boldsymbol{\Gamma},\boldsymbol{H}_i$ 为适当维数常阵.

【假设 1】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}_i(t)\in R^{m_i}$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_{vi}$、相关阵各为 $\boldsymbol{S}_i$ 和 $\boldsymbol{R}_{ij}$ 的相关白噪声:

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}_i(t)\end{bmatrix}\left[\boldsymbol{w}^{\mathrm{T}}(k),\boldsymbol{v}_i^{\mathrm{T}}(k)\right]\right\}=\begin{bmatrix}\boldsymbol{Q}_w & \boldsymbol{S}_i\\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{Q}_{vi}\end{bmatrix}\delta_{tk},\quad i=1,\cdots,L$$

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{v}_i(t)\\ \boldsymbol{v}_j(t)\end{bmatrix}\left[\boldsymbol{v}^{\mathrm{T}}(k),\boldsymbol{v}_j^{\mathrm{T}}(k)\right]\right\}=\begin{bmatrix}\boldsymbol{Q}_{vi} & \boldsymbol{R}_{ij}\\ \boldsymbol{R}_{ij}^{\mathrm{T}} & \boldsymbol{Q}_{vj}\end{bmatrix}\delta_{tk},\quad i,j=1,\cdots,L \tag{7.3.3}$$

其中 E 为均值号,T 为转置号,$\delta_{tt}=1,\delta_{tk}=0(t\neq k)$.

【假设 2】 $(\boldsymbol{\Phi},\boldsymbol{H}_i)$ 为完全可观对,即

$$\mathrm{rank}\boldsymbol{\Omega}_i=n,\quad \boldsymbol{\Omega}_i=[\boldsymbol{H}_i^{\mathrm{T}},(\boldsymbol{H}_i\boldsymbol{\Phi})^{\mathrm{T}},\cdots,(\boldsymbol{H}_i\boldsymbol{\Phi}^{\beta_i-1})^{\mathrm{T}}]^{\mathrm{T}},\quad i=1,\cdots,L \tag{7.3.4}$$

其中 β_i 为第 i 个传感器子系统的可观性指数.

【假设 3】 初始观测时刻 $t_0=-\infty$.

问题是求状态 $\boldsymbol{x}(t)$ 的局部稳态最优 Wiener 状态估值器 $\hat{\boldsymbol{x}}_i(t\mid t+N)(N=0,N>0$ 或 $N<0)$ 和最优信息融合状态估值器 $\hat{\boldsymbol{x}}_0(t\mid t+N)$.

由(7.3.1)和(7.3.2)有

$$\boldsymbol{y}_i(t)=\boldsymbol{H}_i(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}\boldsymbol{w}(t)+\boldsymbol{v}_i(t) \tag{7.3.5}$$

其中 q^{-1} 为单位滞后算子.引入左素分解

$$\boldsymbol{H}_i(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}=\boldsymbol{A}_i^{-1}(q^{-1})\boldsymbol{B}_i(q^{-1}),\quad i=1,\cdots,L \tag{7.3.6}$$

其中 $\boldsymbol{A}_i(q^{-1})$ 和 $\boldsymbol{B}_i(q^{-1})$ 是形如 $\boldsymbol{X}_i(q^{-1})=\boldsymbol{X}_{i0}+\boldsymbol{X}_{i1}q^{-1}+\cdots+\boldsymbol{X}_{in_{xi}}q^{-n_{xi}}$ 的多项式矩阵,且 $\boldsymbol{A}_{i0}=\boldsymbol{I}_{m_i},\boldsymbol{B}_{i0}=\boldsymbol{0}$.将(7.3.6)代入(7.3.5)引出局部 ARMA 新息模型

$$\boldsymbol{A}_i(q^{-1})\boldsymbol{y}_i(t)=\boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t),\quad i=1,\cdots,L \tag{7.3.7}$$

其中 $\boldsymbol{D}_i(q^{-1})$ 是稳定的,$\boldsymbol{D}_{i0}=\boldsymbol{I}_{m_i}$,新息 $\boldsymbol{\varepsilon}(t)\in R^{m_i}$ 是零均值、方差阵为 $\boldsymbol{Q}_{\varepsilon i}$ 的白噪声,且有关系

$$\boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t)=\boldsymbol{B}_i(q^{-1})\boldsymbol{w}(t)+\boldsymbol{A}_i(q^{-1})\boldsymbol{v}_i(t) \tag{7.3.8}$$

而 $\boldsymbol{D}_i(q^{-1})$ 和 $\boldsymbol{Q}_{\varepsilon i}$ 可用 Gevers - Wouters 算法求得.

【引理 7.3.1】 多传感器系统(7.3.1)和(7.3.2)在假设 1 ~ 3 下,基于第 i 个传感器状态 $\boldsymbol{x}(t)$ 的渐近稳定的局部 Wiener 状态估值器为

$$\hat{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{K}_N^{(i)}(q^{-1})\tilde{\boldsymbol{D}}_i^{-1}(q^{-1})\boldsymbol{y}_i(t+N),\quad i=1,\cdots,L \tag{7.3.9}$$

它有 ARMA 递推形式

$$\det\widetilde{\boldsymbol{D}}_i(q^{-1})\hat{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{K}_N^{(i)}(q^{-1})\mathrm{adj}\widetilde{\boldsymbol{D}}_i(q^{-1})\boldsymbol{y}_i(t+N) \tag{7.3.10}$$

其中定义多项式矩阵 $\boldsymbol{K}_N^{(i)}(q^{-1})$ 为

$$\boldsymbol{K}_N^{(i)}(q^{-1})=\sum_{k=0}^{\beta_i-1}\boldsymbol{\Omega}_k^{(i)}\left[\boldsymbol{J}_{k-N}^{(i)}(q^{-1})-\sum_{r=0}^{k-1}\boldsymbol{H}_i\boldsymbol{\Phi}^{k-1-r}\boldsymbol{T}\boldsymbol{L}_{N-r}^{iw}(q^{-1})\widetilde{\boldsymbol{A}}_i(q^{-1})-\boldsymbol{L}_{N-k}^{iv}(q^{-1})\widetilde{\boldsymbol{A}}_i(q^{-1})\right] \tag{7.3.11}$$

其中规定 $\boldsymbol{\Phi}^i=\boldsymbol{0}(i<0)$，$r\geqslant 0$，且

$$\boldsymbol{D}_i^{-1}(q^{-1})\boldsymbol{A}_i(q^{-1})=\widetilde{\boldsymbol{A}}_i(q^{-1})\widetilde{\boldsymbol{D}}_i^{-1}(q^{-1}) \tag{7.3.12}$$

其中 $\boldsymbol{A}_{i0}=\boldsymbol{I}_{m_i}$，$\boldsymbol{D}_{i0}=\boldsymbol{I}_{m_i}$，$\det\widetilde{\boldsymbol{D}}_i(q^{-1})=\det\boldsymbol{D}_i(q^{-1})$. 定义

$$\boldsymbol{L}_N^{i\theta}(q^{-1})=\sum_{k=0}^{N}\boldsymbol{M}_k^{i\theta}q^{k-N},\quad N\geqslant 0,\theta=w,v \tag{7.3.13}$$

$$\boldsymbol{L}_N^{i\theta}(q^{-1})=\boldsymbol{0},\quad N<0,\theta=w,v \tag{7.3.14}$$

$$\boldsymbol{M}_k^{i\theta}=\boldsymbol{\Lambda}_k^{i\theta}\boldsymbol{Q}_{\varepsilon i}^{-1},\quad \theta=w,v \tag{7.3.15}$$

$$\boldsymbol{\Lambda}_k^{iw}=\boldsymbol{Q}_w\boldsymbol{F}_k^{(i)\mathrm{T}}+\boldsymbol{S}_i\boldsymbol{G}_k^{(i)\mathrm{T}} \tag{7.3.16}$$

$$\boldsymbol{\Lambda}_k^{iv}=\boldsymbol{Q}_{vi}\boldsymbol{G}_k^{(i)\mathrm{T}}+\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{F}_k^{(i)\mathrm{T}} \tag{7.3.17}$$

其中上标 i 为传感器序号，系数阵 $\boldsymbol{F}_k^{(i)\mathrm{T}}$ 和 $\boldsymbol{G}_k^{(i)\mathrm{T}}$ 可由(7.1.3)和(7.1.14)递推计算，但应将 $\boldsymbol{P}_{ik}$ 用 $\boldsymbol{B}_{ik}$ 代替. $\boldsymbol{J}_N^{(i)}(q^{-1})$ 由(7.2.7) ~ (7.2.9)定义. $n\times m_i$ 矩阵 $\boldsymbol{\Omega}_k^{(i)}$ 定义为

$$\boldsymbol{\Omega}_i^{+}=(\boldsymbol{\Omega}_i^{\mathrm{T}}\boldsymbol{\Omega}_i)^{-1}\boldsymbol{\Omega}_i^{\mathrm{T}}=[\boldsymbol{\Omega}_0^{(i)},\boldsymbol{\Omega}_1^{(i)},\cdots,\boldsymbol{\Omega}_{\beta_i-1}^{(i)}] \tag{7.3.18}$$

其中 $\boldsymbol{\Omega}_i$ 由(7.3.4)定义.

证明　见定理 5.6.1.　□

【引理 7.3.2】　系统(7.3.1)和(7.3.2)在假设 1 ~ 3 下有局部协方差公式

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(k)]=\boldsymbol{\Lambda}_{k-t}^{iw},\quad i=1,\cdots,L \tag{7.3.19}$$

$$\mathrm{E}[\boldsymbol{v}_i(t)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(k)]=\boldsymbol{\Lambda}_{k-t}^{iv},\quad i=1,\cdots,L \tag{7.3.20}$$

其中 $\boldsymbol{\Lambda}_k^{iw}$ 和 $\boldsymbol{\Lambda}_k^{iv}$ 由(7.3.16)和(7.3.17)定义.

证明　见定理 5.1.1.　□

【引理 7.3.3】　多传感器系统(7.3.1)和(7.3.2)在假设 1 ~ 3 下，有局部状态估计误差 $\tilde{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}_i(t\mid t+N)$ 表达式

$$\tilde{\boldsymbol{x}}_i(t\mid t+N)=\sum_{k=0}^{n_{i0}}[\boldsymbol{\Omega}_k^{iw}\boldsymbol{w}(t+k)-\boldsymbol{\Omega}_k^{(i)}\boldsymbol{v}_i(t+k)+\boldsymbol{\Omega}_k^{i\varepsilon}\boldsymbol{\varepsilon}_i(t+k)],\quad N\geqslant 0 \tag{7.3.21}$$

$$\tilde{\boldsymbol{x}}_i(t\mid t+N)=\sum_{k=0}^{n_{i1}}[\boldsymbol{\Omega}_k^{iw}\boldsymbol{w}(t+k)-\boldsymbol{\Omega}_k^{(i)}\boldsymbol{v}_i(t+k)+\boldsymbol{\Omega}_k^{iy}\boldsymbol{\varepsilon}_i(t+N+1+k)],\quad N<0 \tag{7.3.22}$$

其中上标和下标 i 为传感器序号 $n_{i0}=\max(\beta_i-1,N)$，$n_{i1}=\beta_i-2-N$，系数阵 $\boldsymbol{\Omega}_k^{iw}$，$\boldsymbol{\Omega}_k^{i\varepsilon}$ 和 $\boldsymbol{\Omega}_k^{iy}$ 由合并同类项得到. 相应的稳态估值误差方差阵 $\boldsymbol{P}_i(N)=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t+N)\tilde{\boldsymbol{x}}_i^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}_i(N)=\sum_{k=0}^{n_{i0}}\sum_{r=0}^{n_{i0}}[\boldsymbol{\Omega}_k^{iw},-\boldsymbol{\Omega}_k^{(i)},\boldsymbol{\Omega}_k^{i\varepsilon}]\begin{bmatrix}\boldsymbol{Q}_w\delta_{kr} & \boldsymbol{S}_i\delta_{kr} & \boldsymbol{\Lambda}_{r-k}^{iw}\\ \boldsymbol{S}_i^{\mathrm{T}}\delta_{kr} & \boldsymbol{Q}_{vi}\delta_{kr} & \boldsymbol{\Lambda}_{r-k}^{iv}\\ \boldsymbol{\Lambda}_{k-r}^{iw\mathrm{T}} & \boldsymbol{\Lambda}_{k-r}^{iv\mathrm{T}} & \boldsymbol{Q}_{\varepsilon i}\delta_{kr}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}_r^{iw\mathrm{T}}\\ -\boldsymbol{\Omega}_r^{(i)\mathrm{T}}\\ \boldsymbol{\Omega}_r^{i\varepsilon\mathrm{T}}\end{bmatrix},\quad N\geqslant 0 \tag{7.3.23}$$

$$P_i(N) = \sum_{k=0}^{n_{i1}} \sum_{r=0}^{n_{i1}} [\boldsymbol{\Omega}_k^{iw}, -\boldsymbol{\Omega}_k^{(i)}, \boldsymbol{\Omega}_k^{iy}] \begin{bmatrix} Q_w\delta_{kr} & S_i\delta_{kr} & \boldsymbol{\Lambda}_{N+1+r-k}^{iw} \\ S_i^{\mathrm{T}}\delta_{kr} & Q_{vi}\delta_{kr} & \boldsymbol{\Lambda}_{N+1+r-k}^{iv} \\ \boldsymbol{\Lambda}_{N+1+k-r}^{iw\mathrm{T}} & \boldsymbol{\Lambda}_{N+1k-r}^{iv\mathrm{T}} & Q_{\varepsilon i}\delta_{kr} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega}_r^{iw\mathrm{T}} \\ -\boldsymbol{\Omega}_r^{(i)\mathrm{T}} \\ \boldsymbol{\Omega}_r^{iy\mathrm{T}} \end{bmatrix}, \quad N < 0 \tag{7.3.24}$$

证明 见定理 5.6.3 和定理 5.6.4. □

【定理 7.3.1】 多传感器系统(7.3.1)和(7.3.2)在假设 1 ~ 3 下,对 $i \neq j, \forall t, k$,有公式

$$E_{ji}^{v\varepsilon}(k,t) = \mathrm{E}[\boldsymbol{v}_j(k)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(t)] = S_j^{\mathrm{T}}F_{t-k}^{(i)\mathrm{T}} + R_{ij}^{\mathrm{T}}G_{t-k}^{(i)\mathrm{T}} \tag{7.3.25}$$

$$E_{ij}^{v\varepsilon}(k,t) = \mathrm{E}[\boldsymbol{v}_i(k)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t)] = S_i^{\mathrm{T}}F_{t-k}^{(j)\mathrm{T}} + R_{ij}G_{t-k}^{(j)\mathrm{T}} \tag{7.3.26}$$

$$\begin{aligned} E_{ij}^{\varepsilon\varepsilon}(k,t) = \mathrm{E}[\boldsymbol{\varepsilon}_i(k)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t)] &= \sum_{r=0}^{\infty} F_r^{(i)}Q_wF_{t-k+r}^{(j)\mathrm{T}} + \sum_{r=0}^{\infty} F_r^{(i)}S_jG_{t-k+r}^{(j)\mathrm{T}} + \\ &\sum_{r=0}^{\infty} G_r^{(i)}S_i^{\mathrm{T}}F_{t-k+r}^{(j)\mathrm{T}} + \sum_{r=0}^{\infty} G_r^{(i)}R_{ij}G_{t-k+r}^{(j)\mathrm{T}} \end{aligned} \tag{7.3.27}$$

证明 由(7.3.8)$\boldsymbol{\varepsilon}_i(t)$有展式

$$\begin{aligned} \boldsymbol{\varepsilon}_i(t) &= D_i^{-1}(q^{-1})B_i(q^{-1})\boldsymbol{w}(t) + D_i^{-1}(q^{-1})A_i(q^{-1})\boldsymbol{v}_i(t) = \\ &\sum_{r=0}^{\infty} F_r^{(i)}\boldsymbol{w}(t-r) + \sum_{r=0}^{\infty} G_r^{(i)}\boldsymbol{v}_i(t-r) \end{aligned} \tag{7.3.28}$$

其中 $F_r^{(i)}$ 和 $G_r^{(i)}$ 由(7.1.13)和(7.1.14)计算,但应将 P_{ik} 用 B_{ik} 代替.应用上式和(7.3.3)容易得到(7.3.25) ~ (7.3.27). □

【定理 7.3.2】 多传感器系统(7.3.1)和(7.3.2)在假设 1 ~ 3 下,有稳态局部估值误差互协方差阵 $P_{ij}(N) = \mathrm{E}[\tilde{\boldsymbol{x}}_i(t \mid t+N)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t \mid t+N)](i \neq j)$ 为

$$P_{ij}(N) = \sum_{k=0}^{n_{i0}} \sum_{r=0}^{n_{j0}} [\boldsymbol{\Omega}_{ik}^{w}, -\boldsymbol{\Omega}_{ik}, \boldsymbol{\Omega}_{ik}^{c}] Q_{ij}(k,r) \begin{bmatrix} \boldsymbol{\Omega}_{jr}^{w\mathrm{T}} \\ -\boldsymbol{\Omega}_{jr}^{\mathrm{T}} \\ \boldsymbol{\Omega}_{jr}^{c\mathrm{T}} \end{bmatrix}, \quad N \geqslant 0 \tag{7.3.29}$$

$$Q_{ij}(k,r) = \begin{bmatrix} Q_w\delta_{kr} & S_j\delta_{kr} & \boldsymbol{\Lambda}_{r-k}^{jw} \\ S_i^{\mathrm{T}}\delta_{kr} & R_{ij}\delta_{kr} & E_{ij}^{v\varepsilon}(t+k, t+r) \\ \boldsymbol{\Lambda}_{k-r}^{iw\mathrm{T}} & E_{ji}^{v\varepsilon\mathrm{T}}(t+r, t+k) & E_{ij}^{\varepsilon\varepsilon}(t+k, t+r) \end{bmatrix} \tag{7.3.30}$$

且

$$P_{ij}(N) = \sum_{k=0}^{n_{i1}} \sum_{r=0}^{n_{j1}} [\boldsymbol{\Omega}_{ik}^{w}, -\boldsymbol{\Omega}_{ik}, \boldsymbol{\Omega}_{ik}^{y}] R_{ij}(k,r) \begin{bmatrix} \boldsymbol{\Omega}_{jr}^{w\mathrm{T}} \\ -\boldsymbol{\Omega}_{jr}^{\mathrm{T}} \\ \boldsymbol{\Omega}_{jr}^{y\mathrm{T}} \end{bmatrix}, \quad N < 0 \tag{7.3.31}$$

$$R_{ij}(k,r) = \begin{bmatrix} Q_w\delta_{kr} & S_j\delta_{kr} & \boldsymbol{\Lambda}_{N+1+r-k}^{jw} \\ S_i^{\mathrm{T}}\delta_{kr} & R_{ij}\delta_{kr} & E_{ij}^{v\varepsilon}(t+k, t+N+1+r) \\ \boldsymbol{\Lambda}_{N+1+k-r}^{iw\mathrm{T}} & E_{ji}^{v\varepsilon\mathrm{T}}(t+r, t+N+1+k) & E_{ij}^{\varepsilon\varepsilon}(t+N+1+k, t+N+1+r) \end{bmatrix} \tag{7.3.32}$$

证明 应用(7.3.21),(7.3.22),(7.3.25) ~ (7.3.27),(7.3.19)和(7.3.20)得证. □

【定理 7.3.3】 多传感器系统(7.3.1)和(7.3.2)在假设 1 ~ 3 下有最优信息融合 Wiener 滤波器

$$\hat{\boldsymbol{x}}_0(t \mid t+N) = \sum_{i=1}^{L} \boldsymbol{\alpha}_i \hat{\boldsymbol{x}}_i(t \mid t+N) \tag{7.3.33}$$

其中 $\hat{\boldsymbol{x}}_i(t \mid t+N)$ 由引理7.3.1给出.在节6.1的三种最优加权准则下,加权 $\boldsymbol{\alpha}_i$ 为加权阵、加权对角阵或加权系数,它们分别可由节6.1的三种最优融合加权公式根据由引理7.3.2和定理7.3.2给出的 $\boldsymbol{P}_i(N)$ 和 $\boldsymbol{P}_{ij}(N)$ 来计算,且最优融合估值误差方差阵 $\boldsymbol{P}_0(N)$ 满足关系

$$\mathrm{tr}\boldsymbol{P}_0(N) \leqslant \mathrm{tr}\boldsymbol{P}_i(N), \quad i = 1,\cdots,L \tag{7.3.34}$$

【例7.3.1】 考虑三传感器跟踪系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi x}(t) + \boldsymbol{\Gamma w}(t) \tag{7.3.35}$$

$$z_i(t) = \boldsymbol{H}_0 x(t) + \eta_i(t), \quad i = 1,2,3 \tag{7.3.36}$$

$$\eta_i(t+1) = b_i\eta_i(t) + \xi_i(t) \tag{7.3.37}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{H}_0 = [1 \quad 0], \quad \boldsymbol{\Gamma} = \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix} \tag{7.3.38}$$

其中 T_0 为采样周期,状态 $\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}$, $x_1(t)$, $x_2(t)$ 和 $w(t)$ 各为在时刻 tT_0 处运动目标的位置、速度和加速度,$z_i(t)$ 为第 i 个传感器对位置的观测信号,$\eta_i(t)$ 为有色观测噪声,$w(t)$ 和 $\xi_i(t)$ 是零均值、方差各为 σ_w^2 和 $\sigma_{\xi i}^2$ 的相互独立白噪声.问题是求局部Wiener状态滤波器 $\hat{\boldsymbol{x}}_i(t \mid t)$ 和信息融合Wiener状态滤波器 $\hat{\boldsymbol{x}}_0(t \mid t)$.

引入变换 $y_i(t) = z_i(t+1) - b_i z_i(t)$ 可将观测方差(7.3.36)化为带白色观测噪声 $v_i(t)$ 的观测方程

$$y_i(t) = \boldsymbol{H}_i x(t) + v_i(t), \quad i = 1,2,3 \tag{7.3.39}$$

$$\boldsymbol{H}_i = \boldsymbol{H}_0\boldsymbol{\Phi} - b_i\boldsymbol{H}_0 \tag{7.3.40}$$

$$v_i(t) = \boldsymbol{H}_0\boldsymbol{\Gamma} w(t) + \xi_i(t) \tag{7.3.41}$$

显然 $v_i(t)$ 是带零均值、方差为 σ_{vi}^2 的白噪声,

$$\sigma_{vi}^2 = \sigma_w^2\boldsymbol{H}_0\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}\boldsymbol{H}_0^{\mathrm{T}} + \sigma_{\xi i}^2 \tag{7.3.42}$$

且 $w(t)$ 与 $v_i(t)$ 相关,即 $S_i \neq 0$, $v_i(t)$ 与 $v_j(t)$, $(i \neq j)$ 相关,即 $R_{ij} \neq 0$.在仿真中取

$$\begin{aligned} &T_0 = 0.2, \quad \sigma_w^2 = 1, \quad b_1 = 0.3, \quad b_2 = 0.5, \quad b_3 = 0.4, \\ &\sigma_{\xi 1}^2 = 0.81, \quad \sigma_{\xi 2}^2 = 3, \quad \sigma_{\xi 3}^2 = 6 \end{aligned} \tag{7.3.43}$$

对系统(7.3.35)和(7.3.39)应用引理7.3.1,引理7.3.3,定理7.3.2和定理7.3.3可得局部Winener滤波器 $\hat{\boldsymbol{x}}_i(t \mid t)(i = 1,2,3)$ 和按矩阵加权融合Wiener滤波器 $\hat{x}_0(t \mid t)$,相应的滤波误差方差阵 $\boldsymbol{P}_i(0)$ 和 $\boldsymbol{P}_0(0)$ 各为

$$\begin{aligned} \boldsymbol{P}_1(0) &= \begin{bmatrix} 0.258\,633\,425\,358\,99 & 0.151\,903\,898\,814\,48 \\ 0.151\,903\,898\,814\,48 & 0.247\,267\,534\,465\,07 \end{bmatrix}, \\ \boldsymbol{P}_2(0) &= \begin{bmatrix} 1.259\,458\,687\,860\,65 & 0.462\,821\,544\,417\,58 \\ 0.462\,821\,544\,417\,58 & 0.430\,689\,094\,168\,66 \end{bmatrix}, \\ \boldsymbol{P}_3(0) &= \begin{bmatrix} 1.731\,578\,318\,014\,57 & 0.592\,990\,881\,189\,91 \\ 0.592\,990\,881\,189\,91 & 0.487\,330\,235\,384\,23 \end{bmatrix}, \\ \boldsymbol{P}_0(0) &= \begin{bmatrix} 0.224\,213\,030\,446\,96 & 0.140\,547\,982\,426\,75 \\ 0.140\,547\,982\,426\,75 & 0.241\,228\,749\,607\,33 \end{bmatrix} \end{aligned} \tag{7.3.44}$$

相应的迹各为

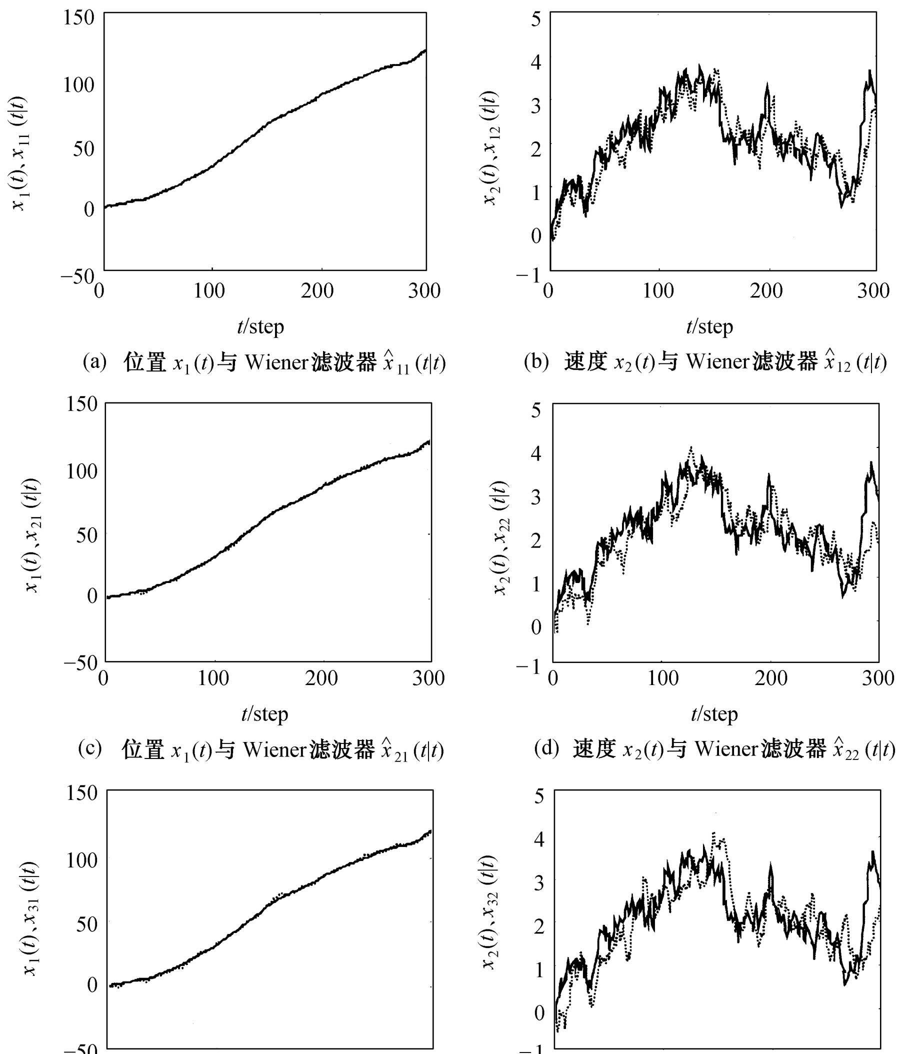
$$\begin{aligned} &\mathrm{tr}\boldsymbol{P}_1(0) = 0.505\,9, \quad \mathrm{tr}\boldsymbol{P}_2(0) = 1.690\,1, \\ &\mathrm{tr}\boldsymbol{P}_3(0) = 2.218\,9, \quad \mathrm{tr}\boldsymbol{P}_0(0) = 0.465\,4 \end{aligned} \tag{7.3.45}$$

于是有精度关系 $\boldsymbol{P}_0(0) < \boldsymbol{P}_i(0)$, $i = 1,2,3$,即融合估计精度高于每个局部估计精度.

为了验证上述方法的正确性,用经典Kalman滤波方法可得

$$P_1(0)=\begin{bmatrix}0.258\,633\,425\,358\,94 & 0.151\,903\,898\,814\,48\\ 0.151\,903\,898\,814\,43 & 0.247\,267\,534\,464\,77\end{bmatrix},$$
$$P_2(0)=\begin{bmatrix}1.259\,458\,687\,859\,77 & 0.462\,821\,544\,416\,88\\ 0.462\,821\,544\,416\,88 & 0.430\,689\,094\,166\,96\end{bmatrix},$$
$$P_3(0)=\begin{bmatrix}1.731\,578\,318\,016\,26 & 0.592\,990\,881\,189\,93\\ 0.592\,990\,881\,189\,93 & 0.487\,330\,235\,388\,18\end{bmatrix}\quad(7.3.46)$$

比较(7.3.44)与(7.3.46),可看到相应方差阵的每个分量仅小数点末尾三位数有所不同,这是由计算机计算误差所引起.因而这证明了两种方法是功能等价的.仿真结果如图7.3.1和图7.3.2所示,其中实线为真实值 $x_i(t)$,虚线为估值 $\hat{x}_i(t\mid t)$ $(i=0,1,2,3)$.图7.3.3为局部



(a) 位置 $x_1(t)$ 与 Wiener滤波器 $\hat{x}_{11}(t|t)$　(b) 速度 $x_2(t)$ 与 Wiener滤波器 $\hat{x}_{12}(t|t)$

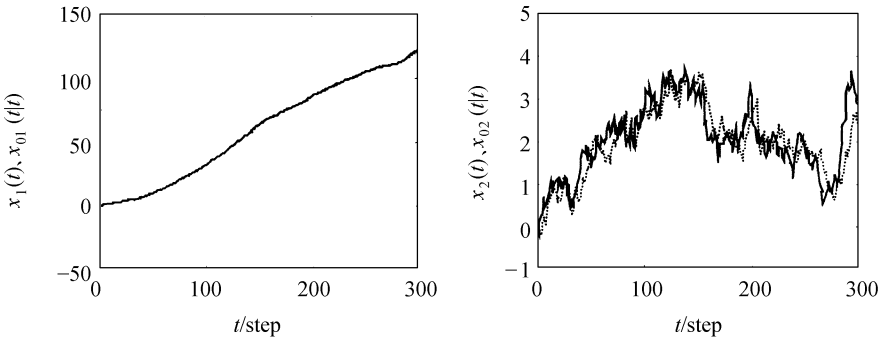
(c) 位置 $x_1(t)$ 与 Wiener滤波器 $\hat{x}_{21}(t|t)$　(d) 速度 $x_2(t)$ 与 Wiener滤波器 $\hat{x}_{22}(t|t)$

(e) 位置 $x_1(t)$ 与 Wiener滤波器 $\hat{x}_{31}(t|t)$　(f) 速度 $x_2(t)$ 与 Wiener滤波器 $\hat{x}_{32}(t|t)$

图7.3.1　状态 $\boldsymbol{x}(t)$ 与局部 Wiener 跟踪滤波器 $\hat{\boldsymbol{x}}_i(t\mid t)$

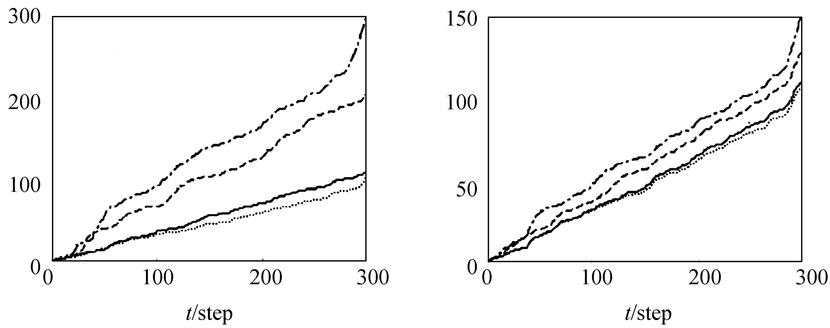
$(\boldsymbol{x}(t)=[x_1(t),x_2(t)]^{\mathrm{T}},\hat{\boldsymbol{x}}_i(t\mid t)=[\hat{x}_{i1}(t\mid t),\hat{x}_{i2}(t\mid t)]^{\mathrm{T}},i=1,2,3)$

和融合滤波绝对误差累积比较曲线,也可看到融合估计精度高于每个局部估计精度.



(a) 位置 $x_1(t)$ 与融合 Wiener 滤波器 $\hat{x}_{01}(t|t)$ (b) 速度 $x_2(t)$ 与融合 Wiener 滤波器 $\hat{x}_{02}(t|t)$

图 7.3.2 状态 $\boldsymbol{x}(t)$ 与按矩阵加权融合 Wiener 跟踪滤波器 $\hat{\boldsymbol{x}}_0(t \mid t)$



(a) 位置滤波累积绝对误差曲线 (b) 速度滤波累积绝对误差曲线

——:单传感器 1 的累积绝对误差曲线 —·—:单传感器 3 的累积绝对误差曲线

- - -:单传感器 2 的累积绝对误差曲线 ……:最优信息融合的累积绝对误差曲线

图 7.3.3 局部和融合 Wiener 滤波器累积绝对误差曲线

7.4 广义系统多传感器信息融合 Wiener 状态估值器

考虑带多传感器的广义随机系统

$$\boldsymbol{M}\boldsymbol{x}(t+1) = \boldsymbol{\Phi}\boldsymbol{x}(t) + \boldsymbol{\Gamma}\boldsymbol{w}(t) \tag{7.4.1}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{H}_i\boldsymbol{x}(t) + \boldsymbol{v}_i(t), \quad i = 1, \cdots, L \tag{7.4.2}$$

其中 t 为离散时间,状态 $\boldsymbol{x}(t) \in R^n$,第 i 个传感器的观测(输出)$\boldsymbol{y}_i(t) \in R^{m_i}$,$\boldsymbol{v}_i(t)$ 为观测噪声,$\boldsymbol{M}, \boldsymbol{\Phi}, \boldsymbol{\Gamma}, \boldsymbol{H}$ 为已知适当维数的常阵.

【假设 1】 $\boldsymbol{M}$ 为奇异方阵($\det\boldsymbol{M} = 0$).

【假设 2】 系统是正则的,即 $\det(\boldsymbol{M} - q^{-1}\boldsymbol{\Phi}) \neq 0$.

【假设 3】 $\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}_i(t) \in R^{m_i}$ 是零均值相关白噪声:

$$\mathrm{E}\left\{\begin{bmatrix} \boldsymbol{w}(t) \\ \boldsymbol{v}_i(t) \end{bmatrix} [\boldsymbol{w}^{\mathrm{T}}(k), \boldsymbol{v}_i^{\mathrm{T}}(k)]\right\} = \begin{bmatrix} \boldsymbol{Q}_w & \boldsymbol{S}_i \\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_i \end{bmatrix}\delta_{tk}, \quad i = 1, \cdots, L \tag{7.4.3}$$

但 $\boldsymbol{v}_i(t)(i = 1, \cdots, L)$ 互不相关:

$$\mathrm{E}[\boldsymbol{v}_i(t)\boldsymbol{v}_j^{\mathrm{T}}(k)] = \boldsymbol{0}, \quad \forall t, k, i, j = 1, \cdots, L \tag{7.4.4}$$

其中 E 为均值号,T 为转置号,$\delta_{tt} = 1$,$\delta_{tk} = 0(t \neq k)$.

【假设 4】 系统是完全可观的，即对任意复数 z 有

$$\operatorname{rank}\begin{bmatrix} z\boldsymbol{M}-\boldsymbol{\Phi} \\ \boldsymbol{H}_i \end{bmatrix}=n,\quad \operatorname{rank}\begin{bmatrix} \boldsymbol{M} \\ \boldsymbol{H}_i \end{bmatrix}=n,\quad i=1,\cdots,L \tag{7.4.5}$$

【假设 5】 初始时刻 $t_0=-\infty$.

【引理 7.4.1】 第 i 传感器子系统(7.4.1)和(7.4.2)在假设 1 ~ 5 下有 ARMA 新息模型

$$\boldsymbol{A}_i(q^{-1})\boldsymbol{y}_i(t)=\boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t),\quad i=1,\cdots,L \tag{7.4.6}$$

其中 $\boldsymbol{A}_i(q^{-1})$ 和 $\boldsymbol{D}_i(q^{-1})$ 是单位滞后算子 q^{-1} 的多项式矩阵，有形式 $\boldsymbol{X}_i(q^{-1})=\boldsymbol{X}_{i0}+\boldsymbol{X}_{i1}q^{-1}+\cdots+\boldsymbol{X}_{in_x}q^{-n_{xi}}$，$\boldsymbol{A}_{i0}=\boldsymbol{I}_{m_i}$，$\boldsymbol{D}_{i0}=\boldsymbol{I}_{m_i}$，$\boldsymbol{I}_{m_i}$ 为 $m_i\times m_i$ 单位阵，$\boldsymbol{D}_i(q^{-1})$ 是稳定的，新息 $\boldsymbol{\varepsilon}_i(t)\in R^{m_i}$ 是零均值、方差阵为 $\boldsymbol{Q}_{\varepsilon i}$ 的白噪声，且有关系

$$\boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t)=\boldsymbol{B}_i(q^{-1})q^{\tau_i}\boldsymbol{w}_i(t)+\boldsymbol{A}_i(q^{-1})\boldsymbol{v}_i(t) \tag{7.4.7}$$

其中 $B_{i0}\neq 0$，$\tau_i=0$，$\tau_i>0$ 或 $\tau_i<0$，q 为单位前进算子. $\boldsymbol{A}_i(q^{-1})$ 和 $\boldsymbol{B}_i(q^{-1})$ 由左素分解

$$\boldsymbol{H}_i(\boldsymbol{M}-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}=\boldsymbol{A}_i^{-1}(q^{-1})\boldsymbol{B}_i(q^{-1})q^{\tau_i} \tag{7.4.8}$$

决定，$\boldsymbol{D}_i(q^{-1})$ 和 $\boldsymbol{Q}_{\varepsilon i}$ 可由 Gevers – Wouters 算法求得.

证明 见节 5.8. □

【引理 7.4.2】 广义系统(7.4.1)和(7.4.2)在假设 1 ~ 5 下，对第 i 个传感器，对任意 t,k 有关系

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(k)]=\boldsymbol{\Lambda}_{k-t}^{iw} \tag{7.4.9}$$

$$\mathrm{E}[\boldsymbol{v}_i(t)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(k)]=\boldsymbol{\Lambda}_{k-t}^{iv} \tag{7.4.10}$$

$$\boldsymbol{\Lambda}_k^{iw}=\boldsymbol{Q}_w\boldsymbol{F}_{k+t_{\tau i}}^{iw\mathrm{T}}+\boldsymbol{S}_i\boldsymbol{F}_{k+t_{\tau i}}^{iv\mathrm{T}} \tag{7.4.11}$$

$$\boldsymbol{\Lambda}_k^{iv}=\boldsymbol{Q}_{vi}\boldsymbol{F}_{k+t_{\tau i}}^{iv\mathrm{T}}+\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{F}_{k+t_{\tau i}}^{iw\mathrm{T}} \tag{7.4.12}$$

其中上、下标 i 为传感器序号，且定义

$$t_{\tau i}=\max(\tau_i,0)=(\tau_i\vee 0),\quad i=1,\cdots,L \tag{7.4.13}$$

而系数阵 $\boldsymbol{F}_k^{iw}$ 和 $\boldsymbol{F}_k^{iv}$ 可递推计算为

$$\boldsymbol{F}_k^{i\theta}=-\boldsymbol{D}_{i1}\boldsymbol{F}_{k-1}^{i\theta}-\cdots-\boldsymbol{D}_{in_{di}}\boldsymbol{F}_{k-n_{di}}^{i\theta}+\boldsymbol{X}_k^{i\theta},\quad \theta=w,v,i=1,\cdots,L \tag{7.4.14}$$

其中上标 i 为传感器序号，规定 $\boldsymbol{F}_k^{i\theta}=\boldsymbol{0}(k<0)$，$\boldsymbol{X}_k^{i\theta}=\boldsymbol{0}(k>n_{xi\theta})$，且定义

$$\boldsymbol{X}^{iw}(q^{-1})=\boldsymbol{B}_i(q^{-1})q^{(\tau_i\wedge 0)},\boldsymbol{X}^{iv}(q^{-1})=\boldsymbol{A}_i(q^{-1})q^{(-\tau_i\wedge 0)} \tag{7.4.15}$$

它们可统一记为

$$\boldsymbol{X}^{i\theta}(q^{-1})=\boldsymbol{X}_0^{i\theta}+\boldsymbol{X}_1^{i\theta}q^{-1}+\cdots+\boldsymbol{X}_{n_{xi\theta}}^{i\theta}q^{-n_{xi\theta}},\quad \theta=w,v \tag{7.4.16}$$

证明 见引理 5.8.2. □

【引理 7.4.3】 广义系统(7.4.1)和(7.4.2)在假设 1 ~ 5 下，对第 i 传感器有渐近稳定的局部 Wiener 状态估值器

$$\hat{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{K}_N^{(i)}(q^{-1})\widetilde{\boldsymbol{D}}_i^{-1}(q^{-1})\boldsymbol{y}_i(t+N),\quad i=1,\cdots,L \tag{7.4.17}$$

它有 ARMA 递推形式

$$\det\widetilde{\boldsymbol{D}}_i(q^{-1})\hat{\boldsymbol{x}}_i(t\mid t+N)=\boldsymbol{K}_N^{(i)}(q^{-1})\operatorname{adj}\widetilde{\boldsymbol{D}}_i(q^{-1})\boldsymbol{y}_i(t+N) \tag{7.4.18}$$

其中 $\widetilde{\boldsymbol{A}}_i(q^{-1})$ 和 $\widetilde{\boldsymbol{D}}_i(q^{-1})$ 用如下伪交换定义：

$$\boldsymbol{D}_i^{-1}(q^{-1})\boldsymbol{A}_i(q^{-1})=\widetilde{\boldsymbol{A}}_i(q^{-1})\widetilde{\boldsymbol{D}}_i^{-1}(q^{-1}) \tag{7.4.19}$$

带 $\boldsymbol{D}_{i0}=\boldsymbol{I}_{m_i}$，$\boldsymbol{A}_{i0}=\boldsymbol{I}_{m_i}$，$\det\widetilde{\boldsymbol{D}}_i^{-1}(q^{-1})=\det\boldsymbol{D}_i(q^{-1})$，且定义

$$\boldsymbol{K}_N^{(i)}(q^{-1}) = \sum_{k=0}^{\beta_i-2}\boldsymbol{\Omega}_k^{i1}\boldsymbol{\Gamma}\boldsymbol{L}_{N-k}^{iw}(q^{-1})\widetilde{\boldsymbol{A}}_i(q^{-1}) + \sum_{k=0}^{\beta_i-1}\boldsymbol{\Omega}_k^{i2}[\boldsymbol{J}_{k-N}^{(i)}(q^{-1}) - \boldsymbol{L}_{N-k}^{iv}(q^{-1})\widetilde{\boldsymbol{A}}_i(q^{-1})] \tag{7.4.20}$$

其中 β_i 为第 i 个传感器相应广义系统的可观性指数，且定义 $\boldsymbol{\Omega}_i$，$\boldsymbol{\Omega}_k^{i1}$，$\boldsymbol{\Omega}_k^{i2}$ 为

$$\boldsymbol{\Omega}_i = \begin{bmatrix} -\boldsymbol{\Phi} & \boldsymbol{M} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & -\boldsymbol{\Phi} & \boldsymbol{M} \\ \boldsymbol{H}_i & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{H}_i \end{bmatrix}, \quad i = 1,\cdots,L \tag{7.4.21}$$

$$\boldsymbol{\Omega}_i^{+} = (\boldsymbol{\Omega}_i^{\mathrm{T}}\boldsymbol{\Omega}_i)^{-1}\boldsymbol{\Omega}_i^{\mathrm{T}},$$

$$[\boldsymbol{I}_n \quad \boldsymbol{0}\cdots\boldsymbol{0}]\boldsymbol{\Omega}_i^{+} = [\boldsymbol{\Omega}_0^{i1}\cdots\boldsymbol{\Omega}_{\beta_i-2}^{i1},\boldsymbol{\Omega}_0^{i2}\cdots\boldsymbol{\Omega}_{\beta_i-1}^{i2}] \tag{7.4.22}$$

且定义

$$\boldsymbol{L}_N^{i\theta}(q^{-1}) = \sum_{k=-t_{\tau i}}^{N}\boldsymbol{\Lambda}_k^{i\theta}\boldsymbol{Q}_{\varepsilon i}^{-1}q^{k-N}, \quad N \geqslant -t_{\tau i},$$

$$\boldsymbol{L}_N^{i\theta}(q^{-1}) = \boldsymbol{0}, \quad N < -t_{\tau i}, \theta = w, v \tag{7.4.23}$$

其中 $i = 1,\cdots,L$ 为传感器序号. $\boldsymbol{J}_N^{(i)}(q^{-1})$ 由下式定义：

$$\widetilde{\boldsymbol{D}}_i(q^{-1}) = \boldsymbol{E}_N^{(i)}(q^{-1})\widetilde{\boldsymbol{A}}_i(q^{-1}) + q^{-N}\boldsymbol{J}_N^{(i)}(q^{-1}), \quad N > 0 \tag{7.4.24}$$

$$\boldsymbol{J}_N^{(i)}(q^{-1}) = \widetilde{\boldsymbol{D}}_i(q^{-1})q^N, \quad N \leqslant 0 \tag{7.4.25}$$

$$\boldsymbol{E}_N^{(i)}(q^{-1}) = \boldsymbol{E}_{N0}^{(i)} + \boldsymbol{E}_{N1}^{(i)}q^{-1} + \cdots + \boldsymbol{E}_{N,N-1}^{(i)}q^{-(N-1)}, \quad N > 0 \tag{7.4.26}$$

$$\boldsymbol{J}_N^{(i)}(q^{-1}) = \boldsymbol{J}_{N0}^{(i)} + \boldsymbol{J}_{N1}^{(i)}q^{-1} + \cdots + \boldsymbol{J}_{Nn_j}^{(i)}q^{-n_j} \tag{7.4.27}$$

$$n_j = \max(n_{ai} - 1, n_{di} - N) \tag{7.4.28}$$

证明　见定理 5.8.4. □

【引理 7.4.4】　广义系统(7.4.1)和(7.4.2)在假设 1 ~ 5 下，Wiener 状态估值器(7.4.17)有状态估值误差 $\tilde{\boldsymbol{x}}_i(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}_i(t \mid t+N)$ 表达式

$$\tilde{\boldsymbol{x}}_i(t \mid t+N) = \sum_{k=1}^{n_{1i}}[\boldsymbol{\Omega}_k^{iw}\boldsymbol{w}(t+k) + \boldsymbol{\Omega}_k^{iv}\boldsymbol{v}_i(t+k) + \boldsymbol{\Omega}_k^{i\varepsilon}\boldsymbol{\varepsilon}_i(t - t_{\tau i} + k)] \tag{7.4.29}$$

其中上、下标 i 为传感器序号，n_{1i}，$t_{\tau i}$，$\boldsymbol{\Omega}_k^{iw}$，$\boldsymbol{\Omega}_k^{i\varepsilon}$ 用定理 5.8.2 方法和推论 5.8.1 求得，且估值误差方差阵 $\boldsymbol{P}_i(N) = \mathrm{E}[\tilde{\boldsymbol{x}}_i(t \mid t+N)\tilde{\boldsymbol{x}}_i^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}_i(N) = \sum_{k=0}^{n_{1i}}\sum_{r=0}^{n_{1i}}[\boldsymbol{\Omega}_k^{iw}, -\boldsymbol{\Omega}_k^{iv}, \boldsymbol{\Omega}_k^{i\varepsilon}]\begin{bmatrix} \boldsymbol{Q}_w\delta_{kr} & \boldsymbol{S}_i\delta_{kr} & \boldsymbol{\Lambda}_{r-k-t_{\tau i}}^{iw} \\ \boldsymbol{S}_i^{\mathrm{T}}\delta_{kr} & \boldsymbol{Q}_{vi}\delta_{kr} & \boldsymbol{\Lambda}_{r-k-t_{\tau i}}^{iv} \\ \boldsymbol{\Lambda}_{k-r-t_{\tau i}}^{iw\mathrm{T}} & \boldsymbol{\Lambda}_{k-r-t_{\tau i}}^{iv\mathrm{T}} & \boldsymbol{Q}_{\varepsilon i}\delta_{kr} \end{bmatrix}\begin{bmatrix} \boldsymbol{\Omega}_r^{iw\mathrm{T}} \\ -\boldsymbol{\Omega}_r^{iv\mathrm{T}} \\ \boldsymbol{\Omega}_r^{i\varepsilon\mathrm{T}} \end{bmatrix},$$

$$i = 1,\cdots,L \tag{7.4.30}$$

证明　见定理 5.8.2，定理 5.8.3 和定理 5.8.4. □

【定理 7.4.1】　广义系统(7.4.1)和(7.4.2)在假设 1 ~ 5 下，有局部估值误差协方差阵 $P_{ij}(N) = \mathrm{E}[\tilde{x}_i(t \mid t+N)\tilde{x}_j^{\mathrm{T}}(t \mid t+N)]$ 为

$$\boldsymbol{P}_{ij}(N) = \sum_{k=0}^{n_{i1}}\sum_{r=0}^{n_{1j}}[\boldsymbol{\Omega}_k^{iw},\boldsymbol{\Omega}_k^{iv},\boldsymbol{\Omega}_k^{i\varepsilon}]\begin{bmatrix} \boldsymbol{Q}_w\delta_{kr} & \boldsymbol{S}_j\delta_{kr} & \boldsymbol{\Lambda}_{r-k-t_{\tau j}}^{jw} \\ \boldsymbol{S}_i^{\mathrm{T}}\delta_{kr} & \boldsymbol{0} & \boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{F}_{r-k}^{jw\mathrm{T}} \\ \boldsymbol{\Lambda}_{k-r-t_{\tau i}}^{iw\mathrm{T}} & \boldsymbol{F}_{k-r}^{iw}\boldsymbol{S}_j & \boldsymbol{E}_{ij}(t_{\tau 0}-t_{\tau i}+k,t_{\tau 0}-t_{\tau j}+r)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}_r^{jw\mathrm{T}}\\ \boldsymbol{\Omega}_r^{jv\mathrm{T}}\\ \boldsymbol{\Omega}_r^{j\varepsilon\mathrm{T}}\end{bmatrix},$$

$$i,j=1,\cdots,L,\ i\neq j$$

$$t_{\tau 0}=\max(t_{\tau i},t_{\tau j}) \tag{7.4.31}$$

其中定义 $\boldsymbol{E}_{ij}(\alpha,\beta)=\mathrm{E}[\boldsymbol{\varepsilon}_i(t+\alpha)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+\beta)]$，且有

$$\boldsymbol{E}_{ij}(\alpha,\beta)=\sum_{s=0}^{\infty}\boldsymbol{F}_s^{iw}\boldsymbol{Q}_w\boldsymbol{F}_{\beta+s+t_{\tau j}-t_{\tau i}-\alpha}^{jw\mathrm{T}}+\sum_{s=0}^{\infty}\boldsymbol{F}_s^{iv}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{F}_{\beta+s+t_{\tau j}-t_{\tau i}-\alpha}^{jw\mathrm{T}}+\sum_{s=0}^{\infty}\boldsymbol{F}_s^{iw}\boldsymbol{S}_j\boldsymbol{F}_{\beta+s+t_{\tau j}-t_{\tau i}-\alpha}^{jv\mathrm{T}} \tag{7.4.32}$$

$$\boldsymbol{E}_{ij}(t_{\tau 0}-t_{\tau i}+k,t_{\tau 0}-t_{\tau j}+r)=\sum_{s=0}^{\infty}\boldsymbol{F}_s^{iw}\boldsymbol{Q}_w\boldsymbol{F}_{r+s-k}^{jw\mathrm{T}}+\sum_{s=0}^{\infty}\boldsymbol{F}^{iw}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{F}_{r+s-k}^{jw\mathrm{T}}+\sum_{s=0}^{\infty}\boldsymbol{F}_s^{iw}\boldsymbol{S}_j\boldsymbol{F}_{r+s-k}^{jw\mathrm{T}} \tag{7.4.33}$$

证明　由(7.4.29)有

$$\tilde{\boldsymbol{x}}_i(t\mid t+N)=\sum_{k=0}^{n_{1i}}[\boldsymbol{\Omega}_k^{iw},\boldsymbol{\Omega}_k^{iv},\boldsymbol{\Omega}_k^{i\varepsilon}]\begin{bmatrix}\boldsymbol{w}(t+k)\\ \boldsymbol{v}_i(t+k)\\ \boldsymbol{\varepsilon}_i(t-t_{\tau i}+k)\end{bmatrix} \tag{7.4.34}$$

$$\tilde{\boldsymbol{x}}_j(t\mid t+N)=\sum_{r=0}^{n_{1i}}[\boldsymbol{\Omega}_r^{jw},\boldsymbol{\Omega}_r^{jv},\boldsymbol{\Omega}_r^{j\varepsilon}]\begin{bmatrix}\boldsymbol{w}(t+r)\\ \boldsymbol{v}_j(t+r)\\ \boldsymbol{\varepsilon}_j(t-t_{\tau j}+r)\end{bmatrix} \tag{7.4.35}$$

且由(7.4.7)，(7.4.13)和(7.4.15)有展式

$$\boldsymbol{\varepsilon}_i(t)=\boldsymbol{D}_i^{-1}(q^{-1})\boldsymbol{X}^{iw}(q^{-1})\boldsymbol{w}(t+t_{\tau i})+\boldsymbol{D}_i^{-1}(q^{-1})\boldsymbol{X}^{iv}(q^{-1})\boldsymbol{v}_i(t+t_{\tau i})=\sum_{k=0}^{\infty}\boldsymbol{F}_k^{iw}\boldsymbol{w}(t+t_{\tau i}-k)+\sum_{k=0}^{\infty}\boldsymbol{F}_k^{iv}\boldsymbol{v}_i(t+t_{\tau i}-k) \tag{7.4.36}$$

应用(7.4.34) ~ (7.4.36)，(7.4.3)，(7.4.9)和(7.4.14)容易得到(7.4.31)、(7.4.33).证毕. □

【定理 7.4.2】　多传感器系统(7.4.1)和(7.4.2)在假设 1 ~ 5 下，有最优信息融合 Wiener 滤波器

$$\hat{\boldsymbol{x}}_0(t\mid t+N)=\sum_{i=1}^{L}\boldsymbol{\alpha}_i\hat{\boldsymbol{x}}_i(t\mid t+N) \tag{7.4.37}$$

其中 $\hat{\boldsymbol{x}}_i(t\mid t+N)$ 由引理7.4.3给出.在节6.1的三种最优加权准则下，加权 $\boldsymbol{\alpha}_i$ 为与 N 有关的加权阵、加权对角阵或加权系数，可由节 6.1 的三种最优加权公式，根据由(7.4.30)和(7.4.31)给出的 $\boldsymbol{P}_i(N)$ 和 $\boldsymbol{P}_{ij}(N)$ 来计算.最优融合估值误差方差阵 $\boldsymbol{P}_0(N)$ 满足关系

$$\mathrm{tr}\boldsymbol{P}_0(N)\leqslant \mathrm{tr}\boldsymbol{P}_i(N),\quad i=1,\cdots,L \tag{7.4.38}$$

【注 7.4.1】　类似地，由定理 5.8.6 ~ 定理 5.8.13 可导出另外两种信息融合广义 Wiener 状态估值器.详细推导留给读者.

【注 7.4.2】　应用定理7.4.1的方法，且应用节5.9广义系统降阶 Wiener 状态估值器设计方法，容易设计广义系统信息融合降阶 Wiener 状态估值器.详细推导留给读者.

7.5　多传感器信息融合稳态 Kalman 估值器

在节 6.6 提出了基于经典 Kalman 滤波方法的多传感器信息融合 Kalman 估值器(滤波、预报和平滑).方法特点是基于求解稳态 Riccati 方程来求稳态 Kalman 估值器增益阵和求估值误差方差阵和互协方差阵.本节提出基于现代时间序列分析方法的多传感器信息融合 Kalman 估值器.[35] 方法特点是基于 ARMA 新息模型来求稳态 Kalman 估值器增益,且用 Lyapunov 方程求估值误差方差阵和互协方差阵.该方法完全避免了 Riccati 方程,用构造 ARMA 新息模型取代了求解 Riccati 方程.ARMA 新息模型可用 Gevers - Wouters 迭代算法得到.对于单输出系统该方法可明显减小计算负担,且该方法可用于设计含有未知噪声统计系统的多传感器信息融合自校正 Kalman 估值器[1].

考虑多传感器线性离散定常随机系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t) \tag{7.5.1}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{H}_i\boldsymbol{x}(t)+\boldsymbol{v}_i(t),\quad i=1,\cdots,L \tag{7.5.2}$$

其中 t 为离散时间,状态 $\boldsymbol{x}(t)\in R^n$,$\boldsymbol{y}_i(t)\in R^{m_i}$ 为第 i 传感器的输出(观测),$\boldsymbol{\Phi},\boldsymbol{\Gamma},\boldsymbol{H}_i$ 为已知适当维数常阵,$\boldsymbol{v}_i(t)$ 为观测噪声,$\boldsymbol{w}(t)$ 为输入噪声.

【假设 1】　$\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}_i(t)\in R^{m_i}$ 是带零均值的相关白噪声:

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{v}_i(t)\end{bmatrix}\left[\boldsymbol{w}^{\mathrm{T}}(k),\boldsymbol{v}_i^{\mathrm{T}}(k)\right]\right\}=\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}_i\\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_i\end{bmatrix}\delta_{tk} \tag{7.5.3}$$

【假设 2】　观测噪声 $\boldsymbol{v}_i(t)$ 和 $\boldsymbol{v}_j(t)$ 是相关的,即

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{v}_i(t)\\ \boldsymbol{v}_j(t)\end{bmatrix}\left[\boldsymbol{v}_i^{\mathrm{T}}(k),\boldsymbol{v}_j^{\mathrm{T}}(k)\right]\right\}=\begin{bmatrix}\boldsymbol{R}_i & \boldsymbol{R}_{ij}\\ \boldsymbol{R}_{ij}^{\mathrm{T}} & \boldsymbol{R}_j\end{bmatrix}\delta_{tk} \tag{7.5.4}$$

且定义 $\boldsymbol{R}_{ii}=\boldsymbol{R}_i$.

【假设 3】　$(\boldsymbol{\Phi},\boldsymbol{H}_i)$ 为完全可观对,$(\bar{\boldsymbol{\Phi}}_i,\boldsymbol{\Gamma Q}_{i0})$ 为完全能稳对,其中 $\bar{\boldsymbol{Q}}_i=\boldsymbol{Q}-\boldsymbol{S}_i\boldsymbol{R}_i^{-1}\boldsymbol{S}_i^{\mathrm{T}}$,$\bar{\boldsymbol{Q}}_i=\boldsymbol{Q}_{i0}\boldsymbol{Q}_{i0}^{\mathrm{T}}$,$\bar{\boldsymbol{\Phi}}_i=\boldsymbol{\Phi}-\boldsymbol{J}_i\boldsymbol{H}_i$,$\boldsymbol{J}_i=\boldsymbol{\Gamma S}_i\boldsymbol{R}_i^{-1}$.

【假设 4】　初始观测时刻 $t_0=-\infty$.

问题是基于第 i 个传感器求局部稳态 Kalman 估值器 $\hat{\boldsymbol{x}}_i(t\mid t+N)$,$i=1,\cdots,L$,并求由它们加权构成的最优信息融合 Kalman 估值器 $\hat{\boldsymbol{x}}_0(t\mid t+N)$,$N=0$,$N>0$ 或 $N<0$.

引入左素分解

$$\boldsymbol{H}_i(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}=\boldsymbol{A}_i^{-1}(q^{-1})\boldsymbol{B}_i(q^{-1}) \tag{7.5.5}$$

其中 q^{-1} 为单位滞后算子,$\boldsymbol{A}_i(q^{-1})$ 和 $\boldsymbol{B}_i(q^{-1})$ 有形式 $\boldsymbol{X}_i(q^{-1})=\boldsymbol{X}_{i0}+\boldsymbol{X}_{i1}q^{-1}+\cdots+\boldsymbol{X}_{in_x}q^{-n_{xi}}$,且有 $\boldsymbol{A}_{i0}=\boldsymbol{I}_{m_i}$,$\boldsymbol{B}_{i0}=\boldsymbol{0}$.由节 5.10,第 i 传感器有 ARMA 新息模型

$$\boldsymbol{A}_i(q^{-1})\boldsymbol{y}_i(t)=\boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t),\quad i=1,\cdots,L \tag{7.5.6}$$

其中 $\boldsymbol{D}_i(q^{-1})$ 是稳定的,$\boldsymbol{D}_{i0}=\boldsymbol{I}_{m_i}$,新息 $\boldsymbol{\varepsilon}_i(t)\in R^{m_i}$ 是零均值、方差阵为 $\boldsymbol{Q}_{\varepsilon i}$ 的白噪声,$\boldsymbol{D}_i(q^{-1})$ 和 $\boldsymbol{Q}_{\varepsilon i}$ 可用 Gevers - Wouters 算法求得.

【引理 7.5.1】　多传感器系统(7.5.1)和(7.5.2)在假设 1 ~ 4 下,第 i 个传感器有 $\boldsymbol{x}(t)$ 的局部稳态 Kalman 滤波器

$$\hat{\boldsymbol{x}}_i(t+1\mid t+1)=\boldsymbol{\Psi}_{fi}\hat{\boldsymbol{x}}_i(t\mid t)+(\boldsymbol{I}_n-\boldsymbol{K}_{fi}\boldsymbol{H}_i)\boldsymbol{J}_i\boldsymbol{y}_i(t)+\boldsymbol{K}_{fi}\boldsymbol{y}_i(t+1) \tag{7.5.7}$$

$$\boldsymbol{\Psi}_{fi}=[\boldsymbol{I}_n-\boldsymbol{K}_{fi}\boldsymbol{H}_i]\bar{\boldsymbol{\Phi}}_i \tag{7.5.8}$$

其中 $\boldsymbol{\Psi}_{fi}$ 为稳定矩阵(即 $\boldsymbol{\Psi}_{fi}$ 的所有特征值位于单位圆内),滤波增益 $\boldsymbol{K}_{fi}$ 为

$$\boldsymbol{K}_{fi}=\begin{bmatrix}\boldsymbol{H}_i\\ \boldsymbol{H}_i\boldsymbol{\Phi}\\ \vdots\\ \boldsymbol{H}_i\boldsymbol{\Phi}^{\beta_i-1}\end{bmatrix}^{+}\begin{bmatrix}\boldsymbol{I}_{m_i}-\boldsymbol{R}_i\boldsymbol{Q}_{\varepsilon i}^{-1}\\ \boldsymbol{M}_1^{(i)}-\boldsymbol{H}_i\boldsymbol{\Gamma}\boldsymbol{S}_i\boldsymbol{Q}_{\varepsilon i}^{-1}\\ \vdots\\ \boldsymbol{M}_{\beta_i-1}^{(i)}-\boldsymbol{H}_i\boldsymbol{\Phi}^{\beta_i-2}\boldsymbol{\Gamma}\boldsymbol{S}_i\boldsymbol{Q}_{\varepsilon i}^{-1}\end{bmatrix}\tag{7.5.9}$$

其中矩阵 $\boldsymbol{X}$ 的伪逆 $\boldsymbol{X}^{+}$ 定义为 $\boldsymbol{X}^{+}=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}$,$\beta_i$ 为 $(\boldsymbol{\Phi},\boldsymbol{H}_i)$ 的可观性指数,系数阵 $\boldsymbol{M}_k^{(i)}$ 可递推计算为

$$\boldsymbol{M}_k^{(i)}=-\boldsymbol{A}_{i1}\boldsymbol{M}_{k-1}^{(i)}-\cdots-\boldsymbol{A}_{in_{ai}}\boldsymbol{M}_{k-n_{ai}}^{(i)}+\boldsymbol{D}_{ik}\tag{7.5.10}$$

其中规定 $\boldsymbol{M}_0^{(i)}=I_{m_i}$,$\boldsymbol{M}_k^{(i)}=\boldsymbol{0}(k<0)$,$\boldsymbol{D}_{ik}=\boldsymbol{0}(k>n_{di})$.

稳态滤波误差 $\tilde{\boldsymbol{x}}_i(t\mid t)=\boldsymbol{x}(t)-\hat{\boldsymbol{x}}_i(t\mid t)$ 方差阵 $\boldsymbol{P}_i$ 满足 Lyapunov 方程

$$\boldsymbol{P}_i=\boldsymbol{\Psi}_{fi}\boldsymbol{P}_i\boldsymbol{\Psi}_{fi}^{\mathrm{T}}+\boldsymbol{\Delta}_{fi},\quad i=1,\cdots,L\tag{7.5.11}$$

$$\boldsymbol{\Delta}_{fi}=[\boldsymbol{I}_n-\boldsymbol{K}_{fi}\boldsymbol{H}_i]\boldsymbol{\Gamma}(\boldsymbol{Q}-\boldsymbol{S}_i\boldsymbol{R}_i^{-1}\boldsymbol{S}_i^{\mathrm{T}})\boldsymbol{\Gamma}^{\mathrm{T}}[\boldsymbol{I}_n-\boldsymbol{K}_{fi}\boldsymbol{H}_i]^{\mathrm{T}}+\boldsymbol{K}_{fi}\boldsymbol{R}_i\boldsymbol{K}_{fi}^{\mathrm{T}}\tag{7.5.12}$$

而稳态局部滤波误差互协方差阵 $\boldsymbol{P}_{ij}=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t\mid t)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t\mid t)](i\neq j)$ 满足 Lyapunov 方程

$$\boldsymbol{P}_{ij}=\boldsymbol{\Psi}_{fi}\boldsymbol{P}_{ij}\boldsymbol{\Psi}_{fj}^{\mathrm{T}}+\boldsymbol{\Delta}_{fij},\quad i\neq j,i,j=1,\cdots,L\tag{7.5.13}$$

其中 $\boldsymbol{\Delta}_{fij}$ 定义为

$$\begin{aligned}\boldsymbol{\Delta}_{fij}=&\boldsymbol{\Psi}_{fi}\boldsymbol{K}_{fi}[\boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}}-\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}][\boldsymbol{I}_n-\boldsymbol{K}_{fj}\boldsymbol{H}_j]^{\mathrm{T}}+[\boldsymbol{I}_n-\boldsymbol{K}_{fi}\boldsymbol{H}_i][\boldsymbol{J}_i\boldsymbol{R}_{ij}-\boldsymbol{\Gamma}\boldsymbol{S}_j]\boldsymbol{K}_{fj}^{\mathrm{T}}\boldsymbol{\Psi}_{fj}^{\mathrm{T}}+\\&[\boldsymbol{I}_n-\boldsymbol{K}_{fi}\boldsymbol{H}_i][\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{J}_i\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{\Gamma}\boldsymbol{S}_j\boldsymbol{J}_j^{\mathrm{T}}+\boldsymbol{J}_i\boldsymbol{R}_{ij}\boldsymbol{J}_j^{\mathrm{T}}]\times\\&[\boldsymbol{I}_n-\boldsymbol{K}_{fj}\boldsymbol{H}_j]^{\mathrm{T}}+\boldsymbol{K}_{fi}\boldsymbol{R}_{ij}\boldsymbol{K}_{fj}^{\mathrm{T}}\end{aligned}\tag{7.5.14}$$

证明 见定理 5.10.1,定理 5.10.2,定理 6.2.4 和定理 6.6.4. □

【定理 7.5.2】 多传感器系统(7.5.1)和(7.5.2)在假设 1 ~ 4 下,第 i 传感器子系统有局部稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t+1\mid t)=\boldsymbol{\Psi}_{pi}\hat{\boldsymbol{x}}_i(t\mid t-1)+\boldsymbol{K}_{pi}\boldsymbol{y}_i(t)\tag{7.5.15}$$

$$\boldsymbol{\Psi}_{pi}=\boldsymbol{\Phi}-\boldsymbol{K}_{pi}\boldsymbol{H}_i\tag{7.5.16}$$

其中 $\boldsymbol{\Psi}_{pi}$ 是一个稳定矩阵,预报增益阵 $\boldsymbol{K}_{pi}$ 为

$$\boldsymbol{K}_{pi}=\begin{bmatrix}\boldsymbol{H}_i\\ \boldsymbol{H}_i\boldsymbol{\Phi}\\ \vdots\\ \boldsymbol{H}_i\boldsymbol{\Phi}^{\beta_i-1}\end{bmatrix}^{+}\begin{bmatrix}\boldsymbol{M}_1^{(i)}\\ \boldsymbol{M}_2^{(i)}\\ \vdots\\ \boldsymbol{M}_{\beta_i}^{(i)}\end{bmatrix}\tag{7.5.17}$$

其中 $\boldsymbol{M}_k^{(i)}$ 由(7.5.10)计算,且局部稳态预报误差 $\tilde{\boldsymbol{x}}_i(t+1\mid t)=\boldsymbol{x}(t+1)-\hat{\boldsymbol{x}}_i(t+1\mid t)$ 方差阵 $\boldsymbol{\Sigma}_i$ 满足 Lyapunov 方程

$$\boldsymbol{\Sigma}_i=\boldsymbol{\Psi}_{pi}\boldsymbol{\Sigma}_i\boldsymbol{\Psi}_{pi}^{\mathrm{T}}+\boldsymbol{\Delta}_{pi},\quad i=1,\cdots,L\tag{7.5.18}$$

其中定义

$$\boldsymbol{\Delta}_{pi}=\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{K}_{pi}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{\Gamma}\boldsymbol{S}_i\boldsymbol{K}_{pi}^{\mathrm{T}}+\boldsymbol{K}_{pi}\boldsymbol{R}_i\boldsymbol{K}_{pi}^{\mathrm{T}}\tag{7.5.19}$$

而局部稳态预报误差互协方差阵 $\boldsymbol{\Sigma}_{ij}=\mathrm{E}[\tilde{\boldsymbol{x}}_i(t+1\mid t)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t+1\mid t)](i\neq j)$ 满足 Lyapunov 方程

$$\boldsymbol{\Sigma}_{ij}=\boldsymbol{\Psi}_{pi}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}}+\boldsymbol{\Delta}_{pij},\quad i,j=1,\cdots,L,i\neq j\tag{7.5.20}$$

其中定义

$$\boldsymbol{\Delta}_{pij}=\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{K}_{pi}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}-\boldsymbol{\Gamma}\boldsymbol{S}_j\boldsymbol{K}_{pj}^{\mathrm{T}}+\boldsymbol{K}_{pi}\boldsymbol{R}_{ij}\boldsymbol{K}_{pi}^{\mathrm{T}}\tag{7.5.21}$$

证明 见定理 5.10.3 和定理 6.6.1. □

【注 7.5.1】 上述信息融合 Kalman 滤波器和预报器与定理 6.6.1 和定理 6.6.3 给出的信息融合 Kalman 滤波器和预报器的根本区别是在这里我们基于 ARMA 新息模型求滤波和预报增益，而定理 6.6.1 和定理 6.6.3 是基于 Riccati 方程求滤波和预报增益. 相同点是计算局部估计误差互协方差阵的 Lyapunov 方程是一样的.

【定理 7.5.3】 多传感器系统(7.5.1)和(7.5.2)在假设 1 ~ 4 下，第 i 传感器子系统有超前($-N$)步局部稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}_i(t \mid t+N) = \boldsymbol{\Phi}^{-N-1}\hat{\boldsymbol{x}}_i(t+N+1 \mid t+N), \quad N \leqslant -2, i = 1,\cdots,L \tag{7.5.22}$$

相应的局部稳态预报误差 $\tilde{\boldsymbol{x}}_i(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}_i(t \mid t+N)$ 方差阵为

$$\boldsymbol{P}_i(N) = \boldsymbol{\Phi}^{-N-1}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Phi}^{-(N-1)\mathrm{T}} + \sum_{j=0}^{-N-2}\boldsymbol{\Phi}^j\boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Phi}^{j\mathrm{T}}, \quad N \leqslant -2 \tag{7.5.23}$$

其中上标 T 为转置号，$\boldsymbol{\Sigma}_i$，$\boldsymbol{\Sigma}_{ij}$ 由定理 7.5.1 和定理 7.5.2 计算. $\hat{\boldsymbol{x}}_i(t+N+1 \mid t+N)$ 由定理 7.5.2 计算.

证明 见定理 6.6.5. □

【定理 7.5.4】 多传感器系统(7.5.1)和(7.5.2)在假设 1 ~ 4 下，第 i 传感器子系统有局部稳态 Kalman 平滑器 $\hat{\boldsymbol{x}}_i(t \mid t+N)(N>0)$ 为

$$\hat{\boldsymbol{x}}_i(t \mid t+N) = \hat{\boldsymbol{x}}_i(t \mid t-1) + \sum_{k=0}^{N}\boldsymbol{K}_i(k)\boldsymbol{\varepsilon}_i(t+k), \quad i = 1,\cdots,L \tag{7.5.24}$$

其中 $\hat{\boldsymbol{x}}_i(t \mid t-1)$ 和 $\boldsymbol{\varepsilon}_i(t) = \boldsymbol{y}_i(t) - \boldsymbol{H}_i\hat{\boldsymbol{x}}_i(t \mid t-1)$ 由定理 7.5.2 计算. 平滑增益 $\boldsymbol{K}_i(k)$ 为

$$\boldsymbol{K}_i(k) = \boldsymbol{\Sigma}_i\boldsymbol{\Psi}_{pi}^{\mathrm{T}k}\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{7.5.25}$$

$$\boldsymbol{Q}_{\varepsilon i} = \boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{R}_i \tag{7.5.26}$$

$$\boldsymbol{K}_{pi} = (\boldsymbol{\Phi}\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{\Gamma}\boldsymbol{S}_i)\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{7.5.27}$$

$$\boldsymbol{\Psi}_{pi} = \boldsymbol{\Phi} - \boldsymbol{K}_{pi}\boldsymbol{H}_i \tag{7.5.28}$$

其中 $\boldsymbol{\Sigma}_i$ 由定理 7.5.2 计算. 局部稳态平滑误差 $\tilde{\boldsymbol{x}}_i(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}_i(t \mid t+N)$ 方差阵为

$$\boldsymbol{P}_i(N) = \boldsymbol{\Sigma}_i - \sum_{k=0}^{N}\boldsymbol{K}_i(k)\boldsymbol{Q}_{\varepsilon i}\boldsymbol{K}_i^{\mathrm{T}}(k), \quad i = 1,\cdots,L \tag{7.5.29}$$

而局部稳态平滑误差互协方差阵 $\boldsymbol{P}_{ij}(N) = \mathrm{E}[\tilde{\boldsymbol{x}}_i(t \mid t+N)\tilde{\boldsymbol{x}}_j^{\mathrm{T}}(t \mid t+N)](i \neq j)$ 为

$$\begin{aligned}\boldsymbol{P}_{ij}(N) = {} & \boldsymbol{\Sigma}_{ij} - \sum_{r=0}^{N}\boldsymbol{K}_i(r)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij} - \sum_{s=0}^{N}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}s}\boldsymbol{H}_j\boldsymbol{K}_j^{\mathrm{T}}(s) + \\ & \sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{K}_i(r)\boldsymbol{E}_{ij}(r,s)\boldsymbol{K}_j^{\mathrm{T}}(s), \quad i \neq j\end{aligned} \tag{7.5.30}$$

其中 $\boldsymbol{\Sigma}_{ij}$ 由(7.5.20)计算，$\boldsymbol{E}_{ij}(r,s) = \mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)]$ 为

$$\begin{aligned}\boldsymbol{E}_{ij}(r,s) = {} & \boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}s}\boldsymbol{H}_j^{\mathrm{T}} + \sum_{k=1}^{\min(r,s)}\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r-k}[\boldsymbol{\Gamma}, -\boldsymbol{K}_{pi}] \times \\ & \begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}_j \\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Gamma}^{\mathrm{T}} \\ -\boldsymbol{K}_{pj}^{\mathrm{T}}\end{bmatrix}\boldsymbol{\Psi}_{pj}^{\mathrm{T}(s-k)}\boldsymbol{H}_j^{\mathrm{T}} + \boldsymbol{R}_{ij}\delta_{rs}, \quad i \neq j, \min(r,s) > 0\end{aligned} \tag{7.5.31}$$

若 $\min(r,s) = 0$，则有

$$\boldsymbol{E}_{ij}(0,0) = \boldsymbol{H}_i\boldsymbol{\Sigma}_{ij}\boldsymbol{H}_j^{\mathrm{T}} + \boldsymbol{R}_{ij} \tag{7.5.32}$$

$$\boldsymbol{E}_{ij}(r,0) = \boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^r\boldsymbol{\Sigma}_{ij}\boldsymbol{H}_j^{\mathrm{T}} + \boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{r-1}[\boldsymbol{\Gamma}\boldsymbol{S}_j - \boldsymbol{K}_{pi}\boldsymbol{R}_{ij}] \tag{7.5.33}$$

$$\boldsymbol{E}_{ij}(0,s) = \boldsymbol{H}_i\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{pj}^{\mathrm{T}s}\boldsymbol{H}_j^{\mathrm{T}} + [\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{R}_{ij}\boldsymbol{K}_{pj}^{\mathrm{T}}]\boldsymbol{\Psi}^{\mathrm{T}(s-1)}\boldsymbol{H}_j^{\mathrm{T}} \tag{7.5.34}$$

证明　见定理 6.6.7. □

【注 7.5.2】　在定理 7.5.4 中 $\boldsymbol{P}_{ij}(N)(i \neq j)$ 也可用如下公式计算：

$$\boldsymbol{P}_{ij}(N) = \boldsymbol{\Psi}_{iN}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Psi}_{iN}^{\mathrm{T}} + \sum_{\rho=0}^{N}[\boldsymbol{K}_{i\rho}^{w},\boldsymbol{K}_{i\rho}^{v}]\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}_j\\ \boldsymbol{S}_i^{\mathrm{T}} & \boldsymbol{R}_{ij}\end{bmatrix}\begin{bmatrix}\boldsymbol{K}_{j\rho}^{w\mathrm{T}}\\ \boldsymbol{K}_{j\rho}^{v\mathrm{T}}\end{bmatrix} \tag{7.5.35}$$

其中 $i,j = 1,\cdots,L,i \neq j$,且定义

$$\boldsymbol{\Psi}_{iN} = \boldsymbol{I}_n - \sum_{k=0}^{N}\boldsymbol{K}_i(k)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{k}$$

$$\boldsymbol{K}_{i\rho}^{w} = -\sum_{k=\rho+1}^{N}\boldsymbol{K}_i(k)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{k-\rho-1}\boldsymbol{\Gamma},\quad \rho = 0,\cdots,N-1;\quad K_{iN}^{w} = \boldsymbol{0},$$

$$\boldsymbol{K}_{i\rho}^{v} = \sum_{k=\rho+1}^{N}\boldsymbol{K}_i(k)\boldsymbol{H}_i\boldsymbol{\Psi}_{pi}^{k-\rho-1}\boldsymbol{K}_{pi} - \boldsymbol{K}_i(\rho),\quad \rho = 0,\cdots,N-1,$$

$$\boldsymbol{K}_{iN}^{v} = -\boldsymbol{K}_i(N) \tag{7.5.36}$$

这个公式的证明见定理 6.4.4 和定理 6.6.10.

在节 7.1 我们提出了一种信息融合白噪声反卷积估值器,在这里我们提出一种新的信息融合白噪声反卷积估值器.

【定理 7.5.5】　多传感器系统(7.5.1)和(7.5.2)在假设 1 ~ 4 下,第 i 传感器子系统有局部稳态最优白噪声反卷积估值器

$$\hat{\boldsymbol{w}}_i(t \mid t+N) = \sum_{k=0}^{N}\boldsymbol{M}_i(k)\boldsymbol{\varepsilon}_i(t+k),\quad N \geqslant 0,i = 1,\cdots,L \tag{7.5.37}$$

$$\hat{\boldsymbol{w}}_i(t \mid t+N) = \boldsymbol{0},\quad N < 0 \tag{7.5.38}$$

其中定义

$$\boldsymbol{M}_i(0) = \boldsymbol{S}_i\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{7.5.39}$$

$$\boldsymbol{M}_i(k) = \boldsymbol{D}_i\boldsymbol{\Psi}_{pi}^{\mathrm{T}(k-1)}\boldsymbol{H}_i^{\mathrm{T}}\boldsymbol{Q}_{\varepsilon i}^{-1} \tag{7.5.40}$$

$$\boldsymbol{D}_i = \boldsymbol{Q}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{S}_i\boldsymbol{K}_{pi}^{\mathrm{T}} \tag{7.5.41}$$

$$\boldsymbol{Q}_{\varepsilon i} = \boldsymbol{H}_i\boldsymbol{\Sigma}_i\boldsymbol{H}_i^{\mathrm{T}} + \boldsymbol{R}_i \tag{7.5.42}$$

其中 $\boldsymbol{\Sigma}_i$ 由定理 7.5.2 计算.局部稳态估值误差 $\tilde{\boldsymbol{w}}_i(t \mid t+N) = \boldsymbol{w}(t) - \hat{\boldsymbol{w}}_i(t \mid t+N)$ 方差阵为

$$\boldsymbol{P}_i^{w} = \boldsymbol{Q} - \sum_{k=0}^{N}\boldsymbol{M}_i(k)\boldsymbol{Q}_{\varepsilon i}\boldsymbol{M}_i^{\mathrm{T}}(k),\quad N \geqslant 0 \tag{7.5.43}$$

$$\boldsymbol{P}_i^{w} = \boldsymbol{Q},\quad N < 0 \tag{7.5.44}$$

而局部稳态估值误差互协方差阵 $\boldsymbol{P}_{ij}^{w}(N) = \mathrm{E}[\tilde{\boldsymbol{w}}_i(t \mid t+N)\tilde{\boldsymbol{w}}_j^{\mathrm{T}}(t \mid t+N)](i \neq j)$ 为

$$\boldsymbol{P}_{ij}^{w}(N) = \boldsymbol{Q} - \sum_{r=0}^{N}\boldsymbol{M}_i(r)\boldsymbol{Q}_{\varepsilon i}\boldsymbol{M}_i^{\mathrm{T}}(r) - \sum_{s=0}^{N}\boldsymbol{M}_j(s)\boldsymbol{Q}_{\varepsilon j}\boldsymbol{M}_j^{\mathrm{T}}(s) + \sum_{r=0}^{N}\sum_{s=0}^{N}\boldsymbol{M}_i(r)\boldsymbol{E}_{ij}(r,s)\boldsymbol{M}_j^{\mathrm{T}}(s),\quad N \geqslant 0 \tag{7.5.45}$$

其中 $i,j = 1,\cdots,L,i \neq j,N \geqslant 0$. $\boldsymbol{E}_{ij}(r,s)$ 由(7.5.31) ~ (7.5.34)计算,$\boldsymbol{\Sigma}_{ij}$ 由定理 7.5.2 计算.当 $N < 0$ 时有

$$\boldsymbol{P}_{ij}^{w}(N) = \boldsymbol{Q},\quad i,j = 1,\cdots,L,i \neq j,N < 0 \tag{7.5.46}$$

信息融合白噪声反卷积估值器为

$$\hat{\boldsymbol{w}}_0(t \mid t+N) = \sum_{i=1}^{L} \boldsymbol{\alpha}_i \hat{\boldsymbol{w}}_i(t \mid t+N) \tag{7.5.47}$$

其中 $\boldsymbol{\alpha}_i$ 为与 N 有关的加权阵、加权对角阵或加权系数，可由 $\boldsymbol{P}_i^w(N)$ 和 $\boldsymbol{P}_{ij}^w(N)$ 根据节 6.1 的三种最优加权公式计算. 最优融合估值误差方差阵 $\boldsymbol{P}_0^w(N)$ 满足关系

$$\mathrm{tr}\boldsymbol{P}_0^w(N) \leqslant \mathrm{tr}\boldsymbol{P}_i^w(N), \quad i = 1, \cdots, L \tag{7.5.48}$$

证明 见定理 6.6.12 和节 6.1. □

【定理 7.5.6】 多传感器系统(7.5.1)和(7.5.2)在假设 1 ~ 4 下，最优信息融合稳态 Kalman 估值器为

$$\hat{\boldsymbol{x}}_0(t \mid t+N) = \sum_{i=1}^{L} \boldsymbol{\alpha}_i \hat{\boldsymbol{x}}_i(t \mid t+N) \tag{7.5.49}$$

其中 $\hat{\boldsymbol{x}}_i(t \mid t+N)$ 由定理 7.5.1 ~ 定理 7.5.4 计算，$\boldsymbol{\alpha}_i$ 为与 N 有关的加权阵、对角加权阵或加权系数，可分别用节 6.1 的三种最优加权公式由定理 7.5.1 ~ 定理 7.5.4 给出的 $\boldsymbol{P}_i(N)$ 和 $\boldsymbol{P}_{ij}(N)$ 计算. 最优加权融合估计误差方差阵 $\boldsymbol{P}_0(N)$ 满足关系

$$\mathrm{tr}\boldsymbol{P}_0(N) \leqslant \mathrm{tr}\boldsymbol{P}_i(N), \quad i = 1, \cdots, L \tag{7.5.50}$$

【例 7.5.1】 考虑带有色观测噪声的两传感器跟踪系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi x}(t) + \boldsymbol{\Gamma} w(t) \tag{7.5.51}$$

$$z_i(t) = \boldsymbol{H}_0 \boldsymbol{x}(t) + \eta_i(t), \quad i = 1, 2 \tag{7.5.52}$$

$$\eta_i(t+1) = b_i \eta_i(t) + \xi_i(t) \tag{7.5.53}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix}, \quad \boldsymbol{H}_0 = [1 \quad 0] \tag{7.5.54}$$

其中 T_0 为采样周期，状态 $x(t) = [x_1(t), x_2(t)]^{\mathrm{T}}$，$x_1(t)$，$x_2(t)$ 和 $w(t)$ 各为时刻 tT_0 处运动目标(飞机，导弹，坦克等)的位置、速度和加速度. $z_i(t)$ 为第 i 传感器对位置的观测，$\eta_i(t)$ 为有色观测噪声，$w(t)$ 和 $\xi_i(t)(i = 1, 2)$ 是零均值、方差各为 σ_w^2 和 $\sigma_{\xi i}^2$ 的相互独立高斯白噪声. 问题是求 $\boldsymbol{x}(t)$ 的局部稳态 Kalman 跟踪平滑器 $\hat{\boldsymbol{x}}_i(t \mid t+2)(i = 1, 2)$ 和加权信息融合 Kalman 跟踪平滑器 $\hat{\boldsymbol{x}}_0(t \mid t+2)$.

引入新的观测过程 $y_i(t) = z_i(t+1) - b_i z_i(t)$，有观测方程

$$y_i(t) = \boldsymbol{H}_i \boldsymbol{x}(t) + v_i(t) \tag{7.5.55}$$

$$\boldsymbol{H}_i = \boldsymbol{H}_0 \boldsymbol{\Phi} - b_i \boldsymbol{H}_0 \tag{7.5.56}$$

$$v_i(t) = \boldsymbol{H}_0 \boldsymbol{\Gamma} w(t) + \xi_i(t) \tag{7.5.57}$$

显然 $v_i(t)$ 是带零均值、方差为 σ_{vi}^2 的白噪声，且

$$\sigma_{vi}^2 = \sigma_w^2 \boldsymbol{H}_0 \boldsymbol{\Gamma} \boldsymbol{\Gamma}^{\mathrm{T}} \boldsymbol{H}_0^{\mathrm{T}} + \sigma_{\xi i}^2, \quad i = 1, 2 \tag{7.5.58}$$

且 $v_i(t)$ 与 $w(t)$ 是相关的，即 $S_i \neq 0$，$v_1(t)$ 与 $v_2(t)$ 是相关的，即 $R_{ij} \neq 0$. 因此对系统(7.5.51)和(7.5.55)应用定理 7.5.4 和定理 7.5.6 可得 $\boldsymbol{x}(t)$ 的局部 Kalman 平滑器 $\hat{\boldsymbol{x}}_i(t \mid t+2)$ 和三种加权融合 Kalman 平滑器 $\hat{\boldsymbol{x}}_0(t \mid t+2)$. 在仿真中取 $T_0 = 0.1$，$\sigma_w^2 = 5$，$\sigma_{\xi 1}^2 = 1$，$\sigma_{\xi 2}^2 = 10$，$b_1 = 0.1$，$b_1 = 0.3$，$N = 2$. 可求得局部平滑误差方差阵的迹为

$$\mathrm{tr}\boldsymbol{P}_1(2) = 0.455\,5, \quad \mathrm{tr}\boldsymbol{P}_2(2) = 2.195\,8 \tag{7.5.59}$$

且可求得三种加权融合 Kalman 平滑器为

$$\hat{\boldsymbol{x}}_0(t \mid t+2) = \boldsymbol{A}_1 \hat{\boldsymbol{x}}_1(t \mid t+2) + \boldsymbol{A}_2 \hat{\boldsymbol{x}}_1(t \mid t+2) \tag{7.5.60}$$

按矩阵加权，可求得最优加权阵

$$A_1 = \begin{bmatrix} 0.940\,9 & -0.002\,4 \\ 0.251\,5 & 0.432\,4 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.059\,1 & 0.002\,4 \\ -0.251\,5 & 0.567\,6 \end{bmatrix} \tag{7.5.61}$$

且有融合平滑误差方差阵的迹为

$$\mathrm{tr}\boldsymbol{P}_0^m(2) = 0.265\,5 \tag{7.5.62}$$

按对角阵加权,可求得最优对角加权阵

$$\boldsymbol{A}_1 = \mathrm{diag}(a_{11}, a_{12}), \quad \boldsymbol{A}_2 = \mathrm{diag}(a_{21}, a_{22}),$$
$$a_{11} = 0.940\,3, \quad a_{21} = 0.059\,7, \quad a_{12} = 0.547\,8, \quad a_{22} = 0.451\,2 \tag{7.5.63}$$

且有融合平滑误差方差阵的迹为

$$\mathrm{tr}\boldsymbol{P}_0^d(2) = 0.296\,7 \tag{7.5.64}$$

按标量加权,可求得最优加权系数为

$$A_1 = a_1 = 0.745\,3, \quad A_2 = a_2 = 0.254\,7 \tag{7.5.65}$$

且有融合平滑误差方差阵的迹为

$$\mathrm{tr}\boldsymbol{P}_0^s(2) = 0.378\,3 \tag{7.5.66}$$

于是有精度关系

$$\mathrm{tr}\boldsymbol{P}_0^m(2) < \mathrm{tr}\boldsymbol{P}_0^d(2) < \mathrm{tr}\boldsymbol{P}_0^s(2) < \mathrm{tr}\boldsymbol{P}_i(2), \quad i = 1,2 \tag{7.5.67}$$

仿真结果如图7.5.1和图7.5.2所示,其中实线代表真实值 $x_i(t)$,虚线代表Kalman平滑估值 $\hat{x}_i(t \mid t+2)$ $(i = 0,1,2)$. 图7.5.3为三种加权融合平滑误差平方累积比较曲线,可进一步看到精度关系(7.5.67)成立.

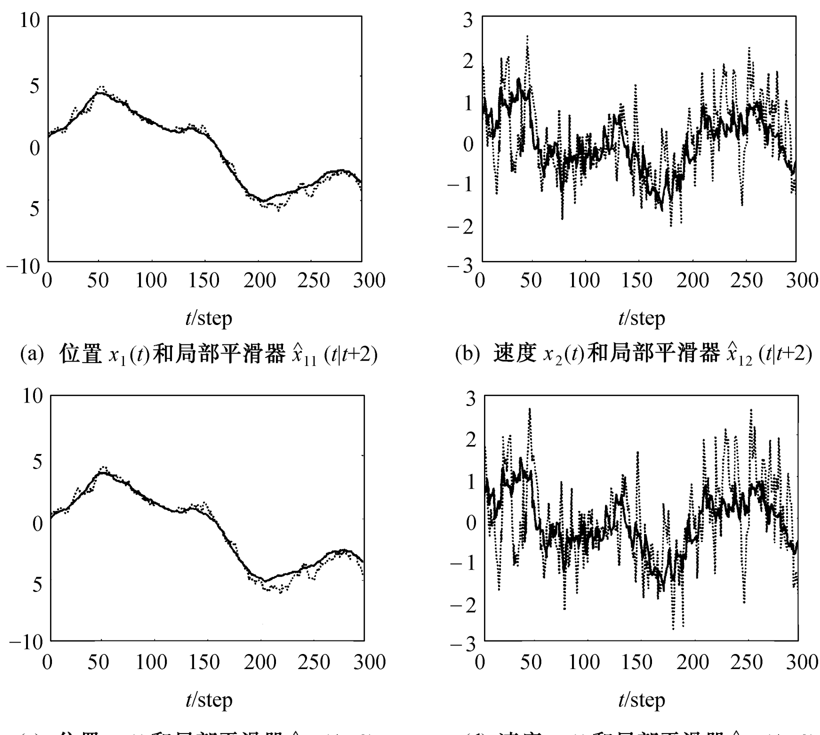


图7.5.1　状态 $\boldsymbol{x}(t)$ 与局部稳态 Kalman 平滑器 $\hat{\boldsymbol{x}}_i(t \mid t+1)$

$(\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}, \hat{\boldsymbol{x}}_i(t \mid t) = [\hat{x}_{i1}(t \mid t), \hat{x}_{i2}(t \mid t)]^{\mathrm{T}}, i = 1,2)$

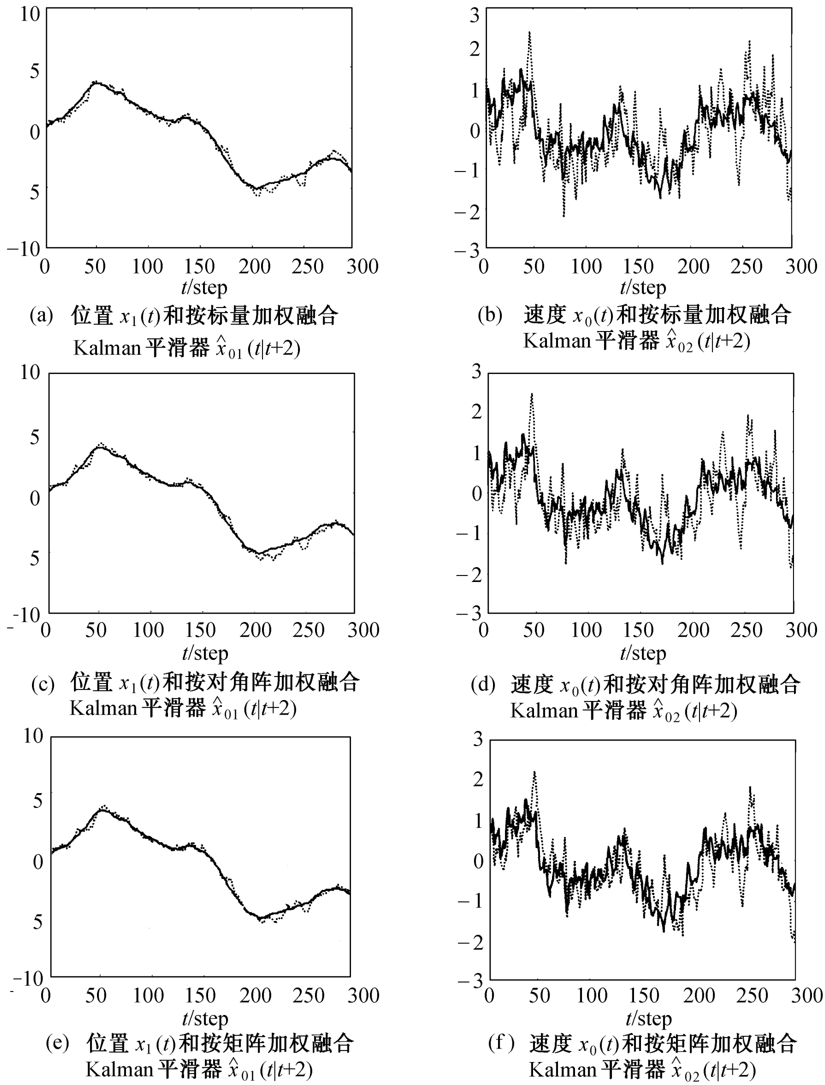


图 7.5.2　状态 $\boldsymbol{x}(t)$ 与标量，对角阵和矩阵三种方式加权的融合 Kalman 平滑器 $\hat{\boldsymbol{x}}_0(t \mid t+1)$

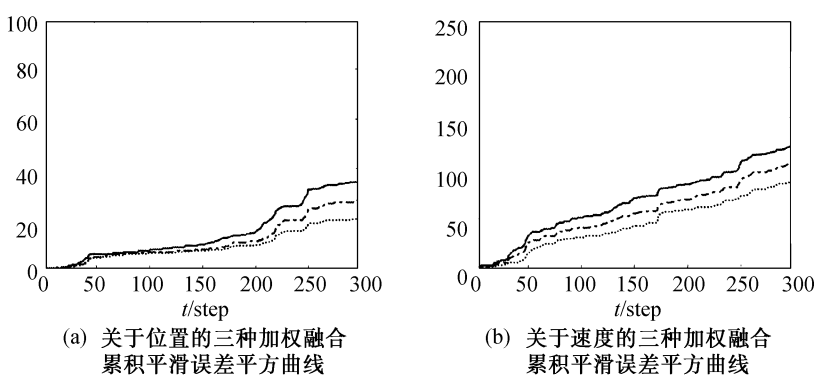


图 7.5.3　三种加权融合累积平滑误差平方比较曲线

（虚线为按矩阵加权融合情形，点划线为按对角阵加权融合情形，实线为按标量加权融合情形）

7.6　多传感器多通道 ARMA 信号全局最优加权观测融合 Wiener 滤波器

考虑多通道多传感器 ARMA 信号

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{7.6.1}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{s}(t) + \boldsymbol{v}_i(t), \quad i = 1,\cdots,L \tag{7.6.2}$$

其中 $\boldsymbol{s}(t) \in R^m$ 为待估的信号，$\boldsymbol{y}_i(t) \in R^m$ 为第 i 个传感器的观测信号，$\boldsymbol{v}_i(t) \in R^m$ 为白色观测噪声. 假设 $\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}_i(t)$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_{vi}$ 的相互独立白噪声，且 $\boldsymbol{v}_i(t)(i = 1,\cdots,L)$ 也相互独立. $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{C}(q^{-1})$ 为单位滞后算子 q^{-1} 的多项式矩阵，有形式 $\boldsymbol{X}(q^{-1}) = \boldsymbol{X}_0 + \boldsymbol{X}_1 q^{-1} + \cdots + \boldsymbol{X}_{n_x} q^{-n_x}$，$\boldsymbol{A}_0 = \boldsymbol{I}_m$，$\boldsymbol{C}_0 = \boldsymbol{0}$.

问题是当 $t_0 = -\infty$ 时求 $\boldsymbol{s}(t)$ 的全局最优加权观测融合 Wiener 滤波器 $\hat{\boldsymbol{s}}_0(t \mid t+N)$，$N = 0$，$N > 0$ 或 $N < 0$.

在节 7.2 的定理 7.2.2 中，我们已提出了分布式加权融合 Wiener 滤波器 $\hat{\boldsymbol{s}}_0(t \mid t+N)$. 它的缺点是对三种加权准则的每种加权准则，相应的融合 Wiener 滤波器均不是全局最优的，即分布式融合滤波误差方差阵的迹大于集中式最优融合滤波误差方差阵的迹，而且分布式融合要求计算每个局部 Wiener 滤波器 $\hat{\boldsymbol{s}}_i(t \mid t+N)(i = 1,\cdots,L)$，计算负担较大.

本节利用定理 6.7.8 证明的多通道多传感器 ARMA 信号的加权观测融合稳态估值器具有全局最优性的功能等价性定理，提出了不同于定理 6.7.9 的新的加权观测融合 Wiener 滤波器. 它的特点是基于 ARMA 新息模型来解决融合估计问题，避免了 Riccati 方程. 而且同节 7.2 的分布式融合滤波器相比，不要求计算局部滤波器，显著地减小了计算负担，且可获得全局最优融合估计.

由节 6.7 的加权观测融合方法，引入等价于(7.6.2)的加权融合观测方程

$$\boldsymbol{y}(t) = \boldsymbol{s}(t) + \boldsymbol{v}(t) \tag{7.6.3}$$

其中融合观测 $\boldsymbol{y}(t)$ 和融合观测白噪声 $\boldsymbol{v}(t)$ 及其方差阵 $\boldsymbol{Q}_v$ 分别为

$$\boldsymbol{y}(t) = \left[\sum_{i=1}^{L} \boldsymbol{Q}_{vi}^{-1}\right]^{-1} \sum_{i=1}^{L} \boldsymbol{Q}_{vi}^{-1} \boldsymbol{y}_i(t) \tag{7.6.4}$$

$$\boldsymbol{v}(t) = \left[\sum_{i=1}^{L} \boldsymbol{Q}_{vi}^{-1}\right]^{-1} \sum_{i=1}^{L} \boldsymbol{Q}_{vi}^{-1} \boldsymbol{v}_i(t) \tag{7.6.5}$$

$$\boldsymbol{Q}_v = \left[\sum_{i=1}^{L} \boldsymbol{Q}_{vi}^{-1}\right]^{-1} \tag{7.6.6}$$

由节 6.7 加权观测融合功能等价性，带融合观测方程(7.6.3)的 ARMA 信号(7.6.1)的稳态最优 Wiener 滤波器 $\hat{\boldsymbol{s}}_0(t \mid t+N)$ 就是带多传感器(7.6.2)的 ARMA 信号(7.6.1)的全局最优 Wiener 滤波器.

将(7.6.3)代入(7.6.1)可得融合观测的 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{7.6.7}$$

其中 $\boldsymbol{D}(q^{-1})$ 是稳定的，$\boldsymbol{D}_0 = \boldsymbol{I}_m$，新息 $\boldsymbol{\varepsilon}(t) \in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{C}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{7.6.8}$$

$\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers - Wouters 算法求得.

引入伪交换

$$\boldsymbol{D}^{-1}(q^{-1})\boldsymbol{A}(q^{-1})=\widetilde{\boldsymbol{A}}(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1}) \tag{7.6.9}$$

带 $\widetilde{\boldsymbol{D}}=\boldsymbol{I}_m,\widetilde{\boldsymbol{A}}=\boldsymbol{I}_m,\det\widetilde{\boldsymbol{D}}(q^{-1})=\det\boldsymbol{D}(q^{-1})$.

引入 Diophantine 方程

$$\widetilde{\boldsymbol{D}}(q^{-1})=\boldsymbol{E}_N(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1})+q^{-N}\boldsymbol{J}_N(q^{-1}),\quad N>0$$

可解得

$$\boldsymbol{E}_N(q^{-1})=\boldsymbol{E}_{N0}+\boldsymbol{E}_{N1}q^{-1}+\cdots+\boldsymbol{E}_{N,N-1}q^{-(N-1)} \tag{7.6.10}$$

$$\boldsymbol{J}_N(q^{-1})=\boldsymbol{J}_{N0}+\boldsymbol{J}_{N1}q^{-1}+\cdots+\boldsymbol{J}_{Nn_j}q^{-n_j} \tag{7.6.11}$$

其中 $n_j=\max(n_a-1,n_d-N)$. 当 $N\leqslant 0$ 时定义

$$\boldsymbol{J}_N(q^{-1})=\widetilde{\boldsymbol{D}}(q^{-1})q^N,\quad N\leqslant 0 \tag{7.6.12}$$

应用引理 7.2.3 有如下定理.

【定理7.6.1】 多传感器多通道系统(7.6.1)和(7.6.2)有 $s(t)$ 的全局稳态最优加权观测融合 Wiener 滤波器

$$\hat{\boldsymbol{s}}_0(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\widetilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{7.6.13}$$

$$\boldsymbol{K}_N(q^{-1})=\boldsymbol{J}_{-N}(q^{-1})-\boldsymbol{L}_N^v(q^{-1})\widetilde{\boldsymbol{A}}(q^{-1}) \tag{7.6.14}$$

它有 ARMA 递推形式

$$\det\widetilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{s}}(t\mid t+N)=\boldsymbol{K}_N(q^{-1})\mathrm{adj}\widetilde{\boldsymbol{D}}(q^{-1})\boldsymbol{y}(t+N) \tag{7.6.15}$$

其中定义

$$\boldsymbol{L}_N^v(q^{-1})=\sum_{k=0}^{N}\boldsymbol{M}_k^v\boldsymbol{Q}_\varepsilon\boldsymbol{M}_k^{v\mathrm{T}},\quad N\geqslant 0,$$

$$\boldsymbol{L}_N^v(q^{-1})=\boldsymbol{0},\quad N<0 \tag{7.6.16}$$

$$\boldsymbol{M}_k^v=\boldsymbol{Q}_v\boldsymbol{G}_k^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1} \tag{7.6.17}$$

而系数阵 $\boldsymbol{G}_k$ 可递推计算为

$$\boldsymbol{G}_k=-\boldsymbol{D}_1\boldsymbol{G}_{k-1}-\cdots-\boldsymbol{D}_{n_d}\boldsymbol{G}_{k-n_d}+\boldsymbol{A}_k \tag{7.6.18}$$

其中定义 $\boldsymbol{G}_k=\boldsymbol{0}(k<0),\boldsymbol{A}_k=\boldsymbol{0}(k>n_a)$. 全局最优融合滤波误差方差阵为

$$\boldsymbol{P}_0^s(N)=\boldsymbol{Q}_v-\sum_{k=0}^{N}\boldsymbol{Q}_v\boldsymbol{G}_k^{\mathrm{T}}\boldsymbol{Q}_\varepsilon^{-1}\boldsymbol{G}_k\boldsymbol{G}_k,\quad N\geqslant 0 \tag{7.6.19}$$

$$\boldsymbol{P}_0^s(N)=\sum_{k=0}^{-N-1}\boldsymbol{E}_{-Nk}\boldsymbol{Q}_\varepsilon\boldsymbol{E}_{-Nk}^{\mathrm{T}}-\boldsymbol{Q}_v,\quad N<0 \tag{7.6.20}$$

且(7.6.13)和(7.6.15)是渐近稳定的.

7.7　多传感器全局最优加权观测融合状态估值器

节 6.7 证明了对于各传感器带有相同观测阵的多传感器系统,用加权观测融合方法可获得全局最优融合估计,即融合估计精度相同于集中式观测融合估计精度,这叫做完全

功能等价性.本节利用加权观测融合方法基于 ARMA 新息模型提出新的加权观测融合 Kalman 估值器(滤波、平滑和预报)和 Wiener 状态估值器.文献[4]曾应用经典 Kalman 滤波方法基于 Riccati 方程提出了非稳态加权观测融合 Kalman 滤波器和一步 Kalman 预报器.而本节采用的方法是现代时间序列分析方法,所提出的最优融合估值器是稳态估值器,且可处理平滑和多步预报,并且除了 Kalman 估值器外,还提出了最优融合 Wiener 状态估值器并给出了一个加权观测融合 Kalman 跟踪滤波器的仿真例子,并验证了其全局最优性.

考虑带相同观测阵的多传感器系统

$$\boldsymbol{x}(t+1)=\boldsymbol{\Phi x}(t)+\boldsymbol{\Gamma w}(t) \tag{7.7.1}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}_i(t),\quad i=1,\cdots,L \tag{7.7.2}$$

其中状态 $\boldsymbol{x}(t)\in R^n$,第 i 个传感器的观测 $\boldsymbol{y}_i(t)\in R^m$,$v_i(t)$ 为观测噪声,各传感器观测方程有相同的观测 $\boldsymbol{H}$,$\boldsymbol{\Phi}$,$\boldsymbol{\Gamma}$ 和 $\boldsymbol{H}$ 为已知常阵.

【假设 1】 $\boldsymbol{w}(t)\in R^r$ 和 $\boldsymbol{v}_i(t)\in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_{vi}$ 的相互独立白噪声,且 $\boldsymbol{v}_i(t)$,$i=1,\cdots,L$,也相互独立.

【假设 2】 $(\boldsymbol{\Phi},\boldsymbol{H})$ 为完全可观对,$(\boldsymbol{\Phi},\boldsymbol{\Gamma})$ 为完全可控对.

【假设 3】 初始观测时刻 $t_0=-\infty$.

问题是求加权观测融合稳态 Kalman 估值器 $\hat{\boldsymbol{x}}_0(t\mid t+N)$ 和 Wiener 状态估值器 $\hat{\boldsymbol{x}}_0(t\mid t+N)$.

应用节 6.7 的加权观测融合方法,引入等价于(7.7.2)的加权融合观测方程

$$\boldsymbol{y}(t)=\boldsymbol{Hx}(t)+\boldsymbol{v}(t) \tag{7.7.3}$$

其中融合观测 $\boldsymbol{y}(t)$,融合观测白噪声 $\boldsymbol{v}(t)$ 及其方差阵 $\boldsymbol{Q}_v$ 分别为

$$\boldsymbol{y}(t)=\left[\sum_{i=1}^{L}\boldsymbol{Q}_{vi}^{-1}\right]^{-1}\sum_{i=1}^{L}\boldsymbol{Q}_{vi}^{-1}\boldsymbol{y}_i(t),$$

$$\boldsymbol{v}(t)=\left[\sum_{i=1}^{L}\boldsymbol{Q}_{vi}^{-1}\right]^{-1}\sum_{i=1}^{L}\boldsymbol{Q}_{vi}^{-1}\boldsymbol{v}_i(t) \tag{7.7.4}$$

$$\boldsymbol{Q}_v=\left[\sum_{i=1}^{L}\boldsymbol{Q}_{vi}^{-1}\right]^{-1} \tag{7.7.5}$$

由(7.7.1)和(7.7.3)有

$$\boldsymbol{y}(t)=\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}\boldsymbol{w}(t)+\boldsymbol{v}(t) \tag{7.7.6}$$

其中 q^{-1} 为单位滞后算子,引入左素分解

$$\boldsymbol{H}(\boldsymbol{I}_n-q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}=\boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1}) \tag{7.7.7}$$

其中 $\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{B}(q^{-1})$ 是形如 $\boldsymbol{X}(q^{-1})=\boldsymbol{X}_0+\boldsymbol{X}_1q^{-1}+\cdots+\boldsymbol{X}_{n_x}q^{-n_x}$ 的多项式矩阵,且 $\boldsymbol{A}_0=\boldsymbol{I}_m$,$\boldsymbol{B}_0=\boldsymbol{0}$.将上式代入(7.7.6)引出 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t)=\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{7.7.8}$$

其中 $\boldsymbol{D}(q^{-1})$ 是稳定的,$\boldsymbol{D}_0=\boldsymbol{I}_m$,新息 $\boldsymbol{\varepsilon}(t)\in R^m$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声,且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t)=\boldsymbol{B}(q^{-1})\boldsymbol{w}(t)+\boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{7.7.9}$$

$\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers - Wouters 算法求得.

【定理7.7.1】 多传感器系统(7.7.1)和(7.7.2)在假设1 ~ 3下,有全局最优加权观测融合稳态 Kalman 滤波器

$$\hat{\boldsymbol{x}}_0(t+1 \mid t+1) = \boldsymbol{\Psi}_f \hat{\boldsymbol{x}}_0(t \mid t) + \boldsymbol{K}_f \boldsymbol{y}(t) \tag{7.7.10}$$

$$\boldsymbol{\Psi}_f = [\boldsymbol{I}_n - \boldsymbol{K}_f \boldsymbol{H}]\boldsymbol{\Phi} \tag{7.7.11}$$

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{\beta-1} \end{bmatrix}^+ \begin{bmatrix} \boldsymbol{I}_m - \boldsymbol{Q}_v \boldsymbol{Q}_\varepsilon^{-1} \\ \boldsymbol{M}_1 \\ \vdots \\ \boldsymbol{M}_{\beta-1} \end{bmatrix} \tag{7.7.12}$$

其中 $\boldsymbol{\Psi}_f$ 为稳定阵,而 $\boldsymbol{X}^+$ 表示矩阵 $\boldsymbol{X}$ 的伪逆,定义为 $\boldsymbol{X}^+ = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}$. $\boldsymbol{M}_i$ 可递推计算为

$$\boldsymbol{M}_i = -\boldsymbol{A}_1 \boldsymbol{M}_{i-1} - \cdots - \boldsymbol{A}_{n_a} \boldsymbol{M}_{i-n_a} + \boldsymbol{D}_i \tag{7.7.13}$$

其中规定 $\boldsymbol{M}_i = \boldsymbol{0}(i<0)$, $\boldsymbol{D}_i = \boldsymbol{0}(i>n_d)$. β 为可观性指数. 最优加权融合稳态滤波误差 $\tilde{\boldsymbol{x}}_0(t \mid t) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}_0(t \mid t)$ 方差阵 $\boldsymbol{P}_0 = \mathrm{E}[\tilde{\boldsymbol{x}}_0(t \mid t)\tilde{\boldsymbol{x}}_0^{\mathrm{T}}(t \mid t)]$ 满足 Lyapunov 方程

$$\boldsymbol{P}_0 = \boldsymbol{\Psi}_f \boldsymbol{P}_0 \boldsymbol{\Psi}_f^{\mathrm{T}} + \boldsymbol{\Delta}_f \tag{7.7.14}$$

其中定义

$$\boldsymbol{\Delta}_f = [\boldsymbol{I}_n - \boldsymbol{K}_f \boldsymbol{H}]\boldsymbol{\Gamma} \boldsymbol{Q}_w \boldsymbol{\Gamma}^{\mathrm{T}} [\boldsymbol{I}_n - \boldsymbol{K}_f \boldsymbol{H}]^{\mathrm{T}} + \boldsymbol{K}_f \boldsymbol{Q}_v \boldsymbol{K}_f^{\mathrm{T}} \tag{7.7.15}$$

证明 见推论5.10.1和推论5.10.2,及定理6.7.7. □

【定理7.7.2】 多传感器系统(7.7.1)和(7.7.2)在假设1 ~ 3下,有全局最优加权观测融合稳态一步 Kalman 预报器

$$\hat{\boldsymbol{x}}(t+1 \mid t) = \boldsymbol{\Psi}_p \hat{\boldsymbol{x}}_0(t \mid t-1) + \boldsymbol{K}_p \boldsymbol{y}(t) \tag{7.7.16}$$

$$\boldsymbol{\Psi}_p = \boldsymbol{\Phi} - \boldsymbol{K}_p \boldsymbol{H} \tag{7.7.17}$$

$$\boldsymbol{K}_p = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \\ \vdots \\ \boldsymbol{H\Phi}^{\beta-1} \end{bmatrix}^+ \begin{bmatrix} \boldsymbol{M}_1 \\ \boldsymbol{M}_2 \\ \vdots \\ \boldsymbol{M}_\beta \end{bmatrix} \tag{7.7.18}$$

其中 $\boldsymbol{\Psi}_p$ 为稳定阵,$\boldsymbol{M}_i$ 由(7.7.13)计算,且稳态预报误差 $\tilde{\boldsymbol{x}}_0(t+1 \mid t) = \boldsymbol{x}(t+1) - \hat{\boldsymbol{x}}_0(t+1 \mid t)$ 方差阵 $\boldsymbol{\Sigma}_0 = \mathrm{E}[\tilde{\boldsymbol{x}}_0(t+1 \mid t)\tilde{\boldsymbol{x}}_0^{\mathrm{T}}(t+1 \mid t)]$ 满足 Lyapunov 方程

$$\boldsymbol{\Sigma}_0 = \boldsymbol{\Psi}_p \boldsymbol{\Sigma}_0 \boldsymbol{\Psi}_p^{\mathrm{T}} + \boldsymbol{\Delta}_p \tag{7.7.19}$$

其中定义

$$\boldsymbol{\Delta}_p = \boldsymbol{\Gamma} \boldsymbol{Q}_w \boldsymbol{\Gamma}^{\mathrm{T}} + \boldsymbol{K}_p \boldsymbol{Q}_v \boldsymbol{K}_p^{\mathrm{T}} \tag{7.7.20}$$

且有全局最优加权观测融合超前 N 步稳态 Kalman 预报器

$$\hat{\boldsymbol{x}}_0(t \mid t+N) = \boldsymbol{\Phi}^{N-1} \hat{\boldsymbol{x}}_0(t+1 \mid t), \quad N \geqslant 2 \tag{7.7.21}$$

相应的预报误差 $\tilde{\boldsymbol{x}}_0(t+N \mid t) = \boldsymbol{x}(t+N) - \hat{\boldsymbol{x}}_0(t \mid t+N)$ 方差阵 $\boldsymbol{\Sigma}_0(N)$ 为

$$\boldsymbol{\Sigma}_0(N) = \boldsymbol{\Phi}^{N-1} \boldsymbol{\Sigma}_0 \boldsymbol{\Phi}^{(N-1)\mathrm{T}} + \sum_{j=0}^{N-2} \boldsymbol{\Phi}^j \boldsymbol{\Gamma} \boldsymbol{Q}_w \boldsymbol{\Gamma}^{\mathrm{T}} \boldsymbol{\Phi}^{j\mathrm{T}}, \quad N \geqslant 2 \tag{7.7.22}$$

证明 见定理7.5.3. □

【定理 7.7.3】 多传感器系统(7.7.1)和(7.7.2)在假设 1 ~ 3 下,有全局最优加权观测融合稳态 Kalman 平滑器 $\hat{\boldsymbol{x}}(t \mid t+N)(N>0)$ 为

$$\hat{\boldsymbol{x}}_0(t \mid t+N) = \hat{\boldsymbol{x}}_0(t \mid t-1) + \sum_{k=0}^{N} \boldsymbol{K}(k)\boldsymbol{\varepsilon}(t+k) \tag{7.7.23}$$

$$\boldsymbol{K}(k) = \boldsymbol{\Sigma}_0 \boldsymbol{\Psi}_p^{\mathrm{T}k} \boldsymbol{H}^{\mathrm{T}} \boldsymbol{Q}_{\varepsilon}^{-1} \tag{7.7.24}$$

$$\boldsymbol{Q}_{\varepsilon} = \boldsymbol{H}\boldsymbol{\Sigma}_0 \boldsymbol{H}^{\mathrm{T}} + \boldsymbol{Q}_v \tag{7.7.25}$$

$$\boldsymbol{\Psi}_p = \boldsymbol{\Phi} - \boldsymbol{K}_p \boldsymbol{H} \tag{7.7.26}$$

其中 $\boldsymbol{K}_p$ 由(7.7.18)计算,$\hat{\boldsymbol{x}}_0(t \mid t-1)$ 和 $\boldsymbol{\Sigma}_0$ 由定理 7.7.2 计算,且相应的平滑误差 $\tilde{\boldsymbol{x}}_0(t \mid t+N) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}_0(t \mid t+N)$ 方差阵为

$$\boldsymbol{P}_0(N) = \boldsymbol{\Sigma}_0 - \sum_{k=0}^{N} \boldsymbol{K}(k)\boldsymbol{Q}_{\varepsilon}\boldsymbol{K}^{\mathrm{T}}(k) \tag{7.7.27}$$

证明 见定理 7.5.4. □

【定理 7.7.4】 多传感器系统(7.7.1)和(7.7.2)在假设 1 ~ 3 下,有全局最优加权观测融合 Wiener 状态估值器 $\hat{\boldsymbol{x}}_0(t \mid t+N)(N=0, N>0$ 或 $N<0)$ 为

$$\hat{\boldsymbol{x}}_0(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{7.7.28}$$

它有 ARMA 递推形式

$$\det\tilde{\boldsymbol{D}}(q^{-1})\hat{\boldsymbol{x}}_0(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\mathrm{adj}\tilde{\boldsymbol{D}}^{-1}(q^{-1})\boldsymbol{y}(t+N) \tag{7.7.29}$$

其中 $\boldsymbol{K}_N(q^{-1})$ 和 $\tilde{\boldsymbol{D}}(q^{-1})$ 由定理 5.6.1 定义,且相应的估值误差方差阵 $\boldsymbol{P}_0(N)$ 由定理 5.6.4 给出.

证明 见定理 5.6.1 和定理 5.6.4. □

【定理 7.7.5】 多传感器系统(7.7.1)和(7.7.2)在假设 1 ~ 3 下,有全局最优加权观测融合 Wiener 状态平滑器 $\hat{\boldsymbol{x}}_0(t \mid t+N)(N>0)$ 为

$$\psi(q^{-1})\hat{\boldsymbol{x}}_0(t \mid t+N) = \boldsymbol{K}_N(q^{-1})\boldsymbol{y}(t+N) \tag{7.7.30}$$

其中定义

$$\boldsymbol{K}_N(q^{-1}) = \mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p q^{-1-N} + \boldsymbol{K}_N^s(q^{-1})\boldsymbol{\Lambda}(q^{-1}) \tag{7.7.31}$$

$$\boldsymbol{\Lambda}(q^{-1}) = \psi(q^{-1})\boldsymbol{I}_m - \boldsymbol{H}\mathrm{adj}(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p)\boldsymbol{K}_p q^{-1} \tag{7.7.32}$$

$$\psi(q^{-1}) = \det(\boldsymbol{I}_n - q^{-1}\boldsymbol{\Psi}_p) \tag{7.7.33}$$

$$\boldsymbol{K}_N^s(q^{-1}) = \sum_{k=0}^{N} \boldsymbol{K}(k) q^{k-N} \tag{7.7.34}$$

其中 $\boldsymbol{K}_p$,$\boldsymbol{K}(k)$ 和 $\boldsymbol{\Psi}_p$ 由(7.7.18),(7.7.24)和(7.7.26)计算,且平滑误差方差阵由(7.7.27)计算.

证明 见定理 5.11.3 和定理 5.11.4. □

【例 7.7.1】 全局最优加权观测融合 Kalman 跟踪滤波器.

考虑带两传感器的目标跟踪系统

$$\boldsymbol{x}(t+1) = \boldsymbol{\Phi}x(t) + \boldsymbol{\Gamma}w(t) \tag{7.7.35}$$

$$\boldsymbol{y}_1(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}_1(t) \tag{7.7.36}$$

$$\boldsymbol{y}_2(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}_2(t) \tag{7.7.37}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_0 & 0.5T_0^2 \\ & 1 & T_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{7.7.38}$$

其中 T_0 为采样周期，状态 $\boldsymbol{x}(t) = [x_1(t), x_2(t), x_3(t)]^{\mathrm{T}}$，$x_1(t)$，$x_2(t)$ 和 $x_3(t)$ 各为在时刻 tT_0 处运动目标（例如导弹，飞机，坦克等）的位置、速度和加速度，$\boldsymbol{y}_i(t) = [y_{i1}(t), y_{i2}(t)]^{\mathrm{T}}$，$i = 1,2$，$y_{i1}(t)$ 和 $y_{i2}(t)$ 各为第 i 个传感器对位置和速度的观测. 输入噪声 $w(t)$ 和观测噪声 $\boldsymbol{v}_i(t)$ 为零均值、方差阵各为 σ_w^2 和 $\boldsymbol{Q}_{vi}$ 的相互独立的高斯白噪声. 问题是基于观测 $(\boldsymbol{y}_i(t), \boldsymbol{y}_i(t-1), \cdots)$，$i = 1,2$，求全局最优加权观测融合跟踪滤波器 $\hat{\boldsymbol{x}}_0(t \mid t)$.

由(7.7.3) ~ (7.7.6)有最优加权观测融合方程

$$\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{v}(t) \tag{7.7.39}$$

$$\boldsymbol{y}(t) = (\boldsymbol{Q}_{v1}^{-1} + \boldsymbol{Q}_{v2}^{-1})^{-1}\boldsymbol{Q}_{v1}^{-1}\boldsymbol{y}_1(t) + (\boldsymbol{Q}_{v1}^{-1} + \boldsymbol{Q}_{v2}^{-1})^{-1}\boldsymbol{Q}_{v2}^{-1}\boldsymbol{y}_2(t) \tag{7.7.40}$$

$$\boldsymbol{v}(t) = (\boldsymbol{Q}_{v1}^{-1} + \boldsymbol{Q}_{v2}^{-1})^{-1}\boldsymbol{Q}_{v1}^{-1}\boldsymbol{v}_1(t) + (\boldsymbol{Q}_{v1}^{-1} + \boldsymbol{Q}_{v2}^{-1})^{-1}\boldsymbol{Q}_{v2}^{-1}\boldsymbol{v}_2(t) \tag{7.7.41}$$

而融合观测噪声 $v(t)$ 的方差阵 $\boldsymbol{Q}_v$ 为

$$\boldsymbol{Q}_v = (\boldsymbol{Q}_{v1}^{-1} + \boldsymbol{Q}_{v2}^{-1})^{-1} \tag{7.7.42}$$

对系统(7.7.35)和(7.7.39)有全局最优加权融合稳态 Kalman 跟踪滤波器

$$\hat{\boldsymbol{x}}_0(t \mid t) = \boldsymbol{\Psi}_f\hat{\boldsymbol{x}}_0(t-1 \mid t-1) + \boldsymbol{K}_f\boldsymbol{y}(t) \tag{7.7.43}$$

$$\boldsymbol{\Psi}_f = [\boldsymbol{I}_n - \boldsymbol{K}_f\boldsymbol{H}]\boldsymbol{\Phi} \tag{7.7.44}$$

其中 $\boldsymbol{K}_f$ 为稳态 Kalman 跟踪滤波器增益.

下面我们用基于 ARMA 新息模型和基于 Riccati 方程两种方法求增益 $\boldsymbol{K}_f$，并验证两种方法引出相同的 $\boldsymbol{K}_f$.

首先用 ARMA 新息模型法. 由(7.7.35)和(7.7.39)有

$$\boldsymbol{y}(t) = \boldsymbol{H}(\boldsymbol{I}_3 - q^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Gamma}q^{-1}w(t) + \boldsymbol{v}(t) = \widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1})w(t-2) + \boldsymbol{v}(t) \tag{7.7.45}$$

其中 q^{-1} 为单位滞后算子，$\boldsymbol{I}_3$ 为 3×3 单位阵，且定义

$$\widetilde{\boldsymbol{A}}(q^{-1}) = \begin{bmatrix} (1-q^{-1})^3 & 0 \\ 0 & (1-q^{-1})^3 \end{bmatrix}, \quad \widetilde{\boldsymbol{B}}(q^{-1}) = \begin{bmatrix} 0.5T_0^2(1+q^{-1}) \\ T_0(1-q^{-1}) \end{bmatrix} \tag{7.7.46}$$

因为当 $q^{-1} = 1$ 时有 $\operatorname{rank}[\widetilde{\boldsymbol{A}}(q^{-1}), \widetilde{\boldsymbol{B}}(q^{-1})] = 1$，由多项式矩阵理论[24]，可知 $\widetilde{\boldsymbol{A}}(q^{-1})$ 和 $\widetilde{\boldsymbol{B}}(q^{-1})$ 非左素，因此必须进行左素分解. 根据求极大左因式方法[24]，用例 5.15.1 的列初等变换方法，可求得 $\widetilde{\boldsymbol{A}}(q^{-1})$ 和 $\widetilde{\boldsymbol{B}}(q^{-1})$ 的极大左因式为

$$\boldsymbol{M}(q^{-1}) = \begin{bmatrix} (1-q^{-1})^2 & 0.5T_0^2(1+q^{-1}) \\ 0 & 1-q^{-1} \end{bmatrix} \tag{7.7.47}$$

消去极大左因式，可得左素分解

$$\widetilde{\boldsymbol{A}}^{-1}(q^{-1})\widetilde{\boldsymbol{B}}(q^{-1}) = \boldsymbol{A}^{-1}(q^{-1})\boldsymbol{B}(q^{-1}) \tag{7.7.48}$$

其中可化 $\boldsymbol{A}(q^{-1})$ 的首系数阵为单位阵，可得

$$\boldsymbol{A}(q^{-1}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & -1.5T_0 \\ 0 & -2 \end{bmatrix}q^{-1} + \begin{bmatrix} 0 & 0.5T_0 \\ 0 & 1 \end{bmatrix}q^{-2} \tag{7.7.49}$$

即 $\boldsymbol{A}(q^{-1}) = \boldsymbol{I}_2 + \boldsymbol{A}_1 q^{-1} + \boldsymbol{A}_2 q^{-2}$，且可得

$$\boldsymbol{B}(q^{-1}) = \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix} \tag{7.7.50}$$

将(7.7.48)代入(7.7.45)可得 ARMA 新息模型

$$\boldsymbol{A}(q^{-1})\boldsymbol{y}(t) = \boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) \tag{7.7.51}$$

其中 $\boldsymbol{D}(q^{-1}) = \boldsymbol{I}_2 + \boldsymbol{D}_1 q^{-1} + \boldsymbol{D}_2 q^{-2}$ 是稳定的，新息 $\boldsymbol{\varepsilon}(t) \in R^2$ 是零均值、方差阵为 $\boldsymbol{Q}_\varepsilon$ 的白噪声，且有关系

$$\boldsymbol{D}(q^{-1})\boldsymbol{\varepsilon}(t) = \boldsymbol{B}(q^{-1})w(t-2) + \boldsymbol{A}(q^{-1})\boldsymbol{v}(t) \tag{7.7.52}$$

而 $\boldsymbol{D}(q^{-1})$ 和 $\boldsymbol{Q}_\varepsilon$ 可用 Gevers – Wouters[2] 算法求得.

容易验证系统(7.7.35)和(7.7.39)是完全可观的，可观性指数为 $\beta = 2$，于是由(7.1.12)有增益 K_f 为

$$\boldsymbol{K}_f = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H\Phi} \end{bmatrix}^{+} \begin{bmatrix} \boldsymbol{I}_2 - \boldsymbol{Q}_v \boldsymbol{Q}_\varepsilon^{-1} \\ \boldsymbol{M}_1 \end{bmatrix} \tag{7.7.53}$$

且由(7.7.13)和 $\boldsymbol{M}_0 = \boldsymbol{D}_0 = \boldsymbol{I}_2$ 有

$$\boldsymbol{M}_1 = -\boldsymbol{A}_1 + \boldsymbol{D}_1 \tag{7.7.54}$$

其次用 Riccati 方程法有

$$\boldsymbol{K}_f = \boldsymbol{\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}} + \boldsymbol{Q}_v)^{-1} \tag{7.7.55}$$

其中 $\boldsymbol{\Sigma}$ 满足稳态 Riccati 方程

$$\boldsymbol{\Sigma} = \boldsymbol{\Phi}[\boldsymbol{\Sigma} - \boldsymbol{\Sigma H}^{\mathrm{T}}(\boldsymbol{H\Sigma H}^{\mathrm{T}} + \boldsymbol{Q}_v)^{-1}\boldsymbol{H\Sigma}]\boldsymbol{\Phi}^{\mathrm{T}} + \sigma_w^2 \boldsymbol{\Gamma\Gamma}^{\mathrm{T}} \tag{7.7.56}$$

滤波误差方差阵 $\boldsymbol{P}$ 为

$$\boldsymbol{P} = [\boldsymbol{I}_3 - \boldsymbol{K}_f \boldsymbol{H}]\boldsymbol{\Sigma} \tag{7.7.57}$$

为了验证全局最优性，考虑集中式观测融合观测方程

$$\boldsymbol{y}_0(t) = \boldsymbol{H}_0 \boldsymbol{x}(t) + \boldsymbol{v}_0(t) \tag{7.7.58}$$

其中定义

$$\boldsymbol{y}_0(t) = \begin{bmatrix} \boldsymbol{y}_1(t) \\ \boldsymbol{y}_2(t) \end{bmatrix}, \quad \boldsymbol{H}_0 = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H} \end{bmatrix}, \quad \boldsymbol{v}_0(t) = \begin{bmatrix} \boldsymbol{v}_1(t) \\ \boldsymbol{v}_2(t) \end{bmatrix} \tag{7.7.59}$$

则有融合观测噪声 $\boldsymbol{v}_0(t)$ 的方差阵为

$$\boldsymbol{Q}_{v0} = \begin{bmatrix} \boldsymbol{Q}_{v1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_{v2} \end{bmatrix} \tag{7.7.60}$$

于是可得全局稳态最优 Kalman 滤波器增益为

$$\boldsymbol{K}_{f0} = \boldsymbol{\Sigma}_0 \boldsymbol{H}_0(\boldsymbol{H}_0 \boldsymbol{\Sigma}_0 \boldsymbol{H}_0^{\mathrm{T}} + \boldsymbol{Q}_{v0})^{-1} \tag{7.7.61}$$

其中 $\boldsymbol{\Sigma}_0$ 满足稳态 Riccati 方程

$$\boldsymbol{\Sigma}_0 = \boldsymbol{\Phi}[\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_0 \boldsymbol{H}_0^{\mathrm{T}}(\boldsymbol{H}_0 \boldsymbol{\Sigma}_0 \boldsymbol{H}_0^{\mathrm{T}} + \boldsymbol{Q}_{v0})^{-1}\boldsymbol{H}_0 \boldsymbol{\Sigma}_0]\boldsymbol{\Phi}^{\mathrm{T}} + \sigma_w^2 \boldsymbol{\Gamma\Gamma}^{\mathrm{T}} \tag{7.7.62}$$

全局最优滤波误差方差阵 $\boldsymbol{P}_0$ 为

$$\boldsymbol{P}_0 = [\boldsymbol{I}_3 - \boldsymbol{K}_{f0}\boldsymbol{H}_0]\boldsymbol{\Sigma}_0 \tag{7.7.63}$$

基于 ARMA 新息模型由(7.7.53)可求得增益

$$\boldsymbol{K}_f = \begin{bmatrix} 0.1249 & 0.1110 \\ 0.1942 & 0.5136 \\ 0.0609 & 1.0404 \end{bmatrix} \tag{7.7.64}$$

基于 Riccati 新息模型，由(7.7.55)可求得增益

$$\boldsymbol{K}_f = \begin{bmatrix} 0.1249 & 0.1110 \\ 0.1942 & 0.5136 \\ 0.0609 & 1.0404 \end{bmatrix} \tag{7.7.65}$$

这验证了用两种方法引出相同的增益 K_f.

由(7.7.57)可求得加权观测融合滤波误差方差阵为

$$\boldsymbol{P} = \begin{bmatrix} 0.2046 & 0.1850 & 0.0580 \\ 0.1850 & 0.8559 & 1.7339 \\ 0.0580 & 1.7339 & 10.3753 \end{bmatrix} \tag{7.7.66}$$

而应用(7.7.61)～(7.7.63)可求得全局最优滤波增益和误差方差阵各为

$$\boldsymbol{K}_{f0} = \begin{bmatrix} 0.2046 & 0.0925 & 0.0102 & 0.0185 \\ 0.1850 & 0.4280 & 0.0092 & 0.0856 \\ 0.0580 & 0.8670 & 0.0029 & 0.1734 \end{bmatrix} \tag{7.7.67}$$

$$\boldsymbol{P}_0 = \begin{bmatrix} 0.2046 & 0.1850 & 0.0580 \\ 0.1850 & 0.8559 & 1.7339 \\ 0.0580 & 1.7339 & 10.3753 \end{bmatrix} \tag{7.7.68}$$

于是有

$$\boldsymbol{P} = \boldsymbol{P}_0 \tag{7.7.69}$$

这证明了加权观测融合稳态 Kalman 滤波器的全局最优性.

仿真结果见图 7.7.1 ～ 7.7.3 所示，其中实线代表真实值，虚线代表 Kalman 滤波估值.

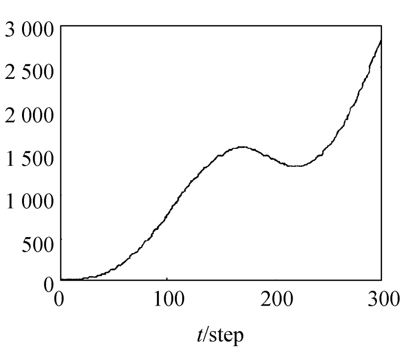


图 7.7.1　位置 $x_1(t)$ 和加权观测融合 Kalman 滤波器 $\hat{\boldsymbol{x}}_{01}(t \mid t)$

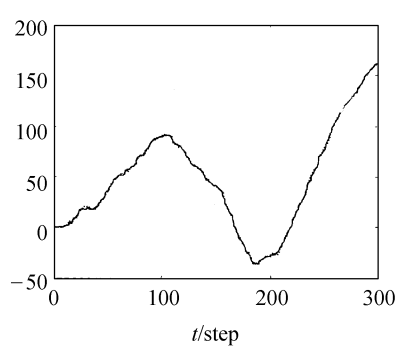


图 7.7.2　速度 $x_2(t)$ 和加权观测融合 Kalman 滤波器 $\hat{\boldsymbol{x}}_{02}(t \mid t)$

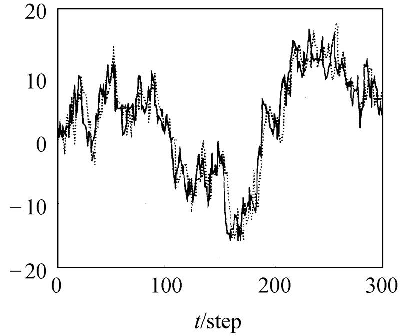


图 7.7.3　加速度 $x_3(t)$ 和加权观测融合 Kalman 滤波器 $\hat{\boldsymbol{x}}_{03}(t \mid t)$

7.8　多传感器分布式信息融合 Wiener 反卷积滤波器

考虑多通道多传感器信息融合反卷积问题

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{7.8.1}$$

$$\boldsymbol{P}_i(q^{-1})\boldsymbol{y}_i(t) = \boldsymbol{B}_i(q^{-1})\boldsymbol{s}(t) + \boldsymbol{R}_i(q^{-1})\boldsymbol{v}_i(t),\quad i = 1,\cdots,L \tag{7.8.2}$$

其中 $\boldsymbol{s}(t) \in R^m$ 为待估的多通道 ARMA 信号，它作为观测系统的输入被间接观测，$\boldsymbol{y}_i(t)$ 是第 i 个传感器的输出(观测)，$\boldsymbol{y}_i(t) \in R^m$，$\boldsymbol{v}_i(t) \in R^m$ 是观测噪声. $\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}_i(t)$ 是零均值、方差各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_{vi}$ 的相互独立的噪声. $\boldsymbol{P}_i(q^{-1})$，$\boldsymbol{B}_i(q^{-1})$，$\boldsymbol{R}_i(q^{-1})$ 是单位滞后算子 q^{-1} 的多项式矩阵，形如 $\boldsymbol{X}_i(q^{-1}) = \boldsymbol{X}_{i0} + \boldsymbol{X}_{i1}q^{-1} + \cdots + \boldsymbol{X}_{in_{xi}}q^{-n_{xi}}$，$\boldsymbol{A}(q^{-1})$ 和 $\boldsymbol{C}(q^{-1})$ 形如 $\boldsymbol{X}(q^{-1}) = \boldsymbol{X}_0 + \boldsymbol{X}_1 q^{-1} + \cdots + \boldsymbol{X}_{n_x}q^{-n_x}$，且 $\boldsymbol{A}_0 = \boldsymbol{I}_m$，$\boldsymbol{P}_0 = \boldsymbol{I}_m$. 问题是求 $\boldsymbol{s}(t)$ 的局部反卷积滤波器 $\hat{\boldsymbol{s}}_i(t \mid t+N)$ 和分布式融合反卷积滤波器 $\hat{\boldsymbol{s}}(t \mid t+N)$，$N = 0$，$N > 0$ 或 $N < 0$.

由节 5.5 可求得各传感器子系统的 ARMA 新息模型

$$\boldsymbol{\Lambda}_i(q^{-1})\boldsymbol{y}_i(t) = \boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t),\quad i = 1,\cdots,L \tag{7.8.3}$$

且由定理 5.5.1 可求得局部反卷积 Wiener 滤波器为

$$\hat{\boldsymbol{s}}_i(t \mid t+N) = \boldsymbol{K}_{Ni}(q^{-1})\widetilde{\boldsymbol{D}}_i^{-1}(q^{-1})\boldsymbol{y}_i(t+N) \tag{7.8.4}$$

由定理 5.5.2 可求得局部滤波误差为

$$\tilde{\boldsymbol{s}}_i(t \mid t+N) = \sum_{s=0}^{n_{0i}+N}\left[\boldsymbol{\Omega}_s^{3i}\boldsymbol{w}(t-s) + \boldsymbol{\Omega}_s^{4i}\boldsymbol{v}_i(t-s) + \boldsymbol{\Omega}_s^{5i}\boldsymbol{\varepsilon}_i(t-n_{0i}+s)\right],\quad N \geqslant 0 \tag{7.8.5}$$

$$\tilde{\boldsymbol{s}}_i(t \mid t+N) = \sum_{s=0}^{n_{3i}}\left[\boldsymbol{\Omega}_s^{3i}\boldsymbol{w}(t-s) + \boldsymbol{\Omega}_s^{4i}\boldsymbol{v}_i(t-s) + \boldsymbol{\Omega}_s^{6i}\boldsymbol{\varepsilon}_i(t-s)\right],\quad N < 0 \tag{7.8.6}$$

由左素分解(5.5.4)有关系

$$\boldsymbol{D}_i(q^{-1})\boldsymbol{\varepsilon}_i(t) = \overline{\boldsymbol{B}}_i(q^{-1})\boldsymbol{w}(t) + \overline{\boldsymbol{R}}_i(q^{-1})\boldsymbol{v}_i(t) \tag{7.8.7}$$

于是有展式

$$\boldsymbol{\varepsilon}_i(t)=\sum_{k=0}^{\infty}\boldsymbol{F}_k^{(i)}\boldsymbol{w}(t-k)+\sum_{k=0}^{\infty}\boldsymbol{G}_k^{(i)}\boldsymbol{v}_i(t-k) \tag{7.8.8}$$

其中 $\boldsymbol{F}_k^{(i)}$ 和 $\boldsymbol{G}_k^{(i)}$ 用公式(5.5.14)计算.由引理5.5.4有关系

$$\mathrm{E}[\boldsymbol{v}_i(t)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(k)]=\boldsymbol{Q}_{vi}\boldsymbol{G}_{k-t}^{(i)\mathrm{T}} \tag{7.8.9}$$

$$\mathrm{E}[\boldsymbol{w}(t)\boldsymbol{\varepsilon}_i^{\mathrm{T}}(k)]=\boldsymbol{Q}_{w}\boldsymbol{F}_{k-t}^{(i)\mathrm{T}} \tag{7.8.10}$$

上述 $n_{0i},n_{3i},\Omega^{3i},\Omega^{4i},\Omega^{5i},\Omega^{6i},\boldsymbol{\Lambda}_i(q^{-1}),\boldsymbol{G}_k^{(i)},\boldsymbol{F}_k^{(i)},\boldsymbol{K}_{Ni}(q^{-1}),\tilde{\boldsymbol{D}}_i(q^{-1}),\bar{\boldsymbol{B}}_i(q^{-1})$，$\bar{\boldsymbol{R}}_i(q^{-1})$ 均由节5.5有关公式计算，并附上标或下标 i，以表示第 i 个传感器子系统.

【定理7.8.1】 多传感器系统(7.8.1)和(7.8.2)有局部估计误差方差阵 $\boldsymbol{P}_i(N)=\mathrm{E}[\tilde{\boldsymbol{s}}_i(t\mid t+N)\tilde{\boldsymbol{s}}_i^{\mathrm{T}}(t\mid t+N)]$ 为

$$\boldsymbol{P}_i(N)=\sum_{r=0}^{n_{0i}+N}\sum_{s=0}^{n_{0i}+N}[\boldsymbol{\Omega}_r^{3i},\boldsymbol{\Omega}_i^{4i},\boldsymbol{\Omega}_r^{5i}]\begin{bmatrix}\boldsymbol{Q}_w\delta_{rs} & \boldsymbol{0} & \boldsymbol{Q}_w\boldsymbol{F}_{r+s-n_{0i}}^{(i)\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{Q}_{vi}\delta_{rs} & \boldsymbol{Q}_{vi}\boldsymbol{G}_{r+s-n_{0i}}^{(i)\mathrm{T}}\\ \boldsymbol{F}_{r+s-n_{0i}}^{(i)}\boldsymbol{Q}_w & \boldsymbol{G}_{r+s-n_{0i}}^{(i)}\boldsymbol{Q}_{vi} & \boldsymbol{Q}_{\varepsilon i}\delta_{rs}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}_s^{3i\mathrm{T}}\\ \boldsymbol{\Omega}_s^{4i\mathrm{T}}\\ \boldsymbol{\Omega}_s^{5i\mathrm{T}}\end{bmatrix},\quad N\geqslant 0 \tag{7.8.11}$$

$$\boldsymbol{P}_i(N)=\sum_{r=0}^{n_{3i}}\sum_{s=0}^{n_{3i}}[\boldsymbol{\Omega}_r^{3i},\boldsymbol{\Omega}_i^{4i},\boldsymbol{\Omega}_r^{6i}]\begin{bmatrix}\boldsymbol{Q}_w\delta_{rs} & \boldsymbol{0} & \boldsymbol{Q}_w\boldsymbol{F}_{r-s}^{(i)\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{Q}_{vi}\delta_{rs} & \boldsymbol{Q}_{vi}\boldsymbol{G}_{r-s}^{(i)\mathrm{T}}\\ \boldsymbol{F}_{s-r}^{(i)}\boldsymbol{Q}_w & \boldsymbol{G}_{s-r}^{(i)}\boldsymbol{Q}_{vi} & \boldsymbol{Q}_{\varepsilon i}\delta_{rs}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}_s^{3i\mathrm{T}}\\ \boldsymbol{\Omega}_s^{4i\mathrm{T}}\\ \boldsymbol{\Omega}_s^{6i\mathrm{T}}\end{bmatrix},\quad N<0 \tag{7.8.12}$$

证明 见定理5.5.3. □

【定理7.8.2】 多传感器系统(7.8.1)和(7.8.2)有局部估计误差互协方差阵 $\boldsymbol{P}_{ij}(N)=\mathrm{E}[\tilde{\boldsymbol{s}}_i(t\mid t+N)\tilde{\boldsymbol{s}}_j^{\mathrm{T}}(t\mid t+N)](i\neq j)$ 为

$$\boldsymbol{P}_{ij}(N)=\sum_{r=0}^{n_{0i}+N}\sum_{s=0}^{n_{0j}+N}[\boldsymbol{\Omega}_r^{3i},\boldsymbol{\Omega}_i^{4i},\boldsymbol{\Omega}_r^{5i}]\begin{bmatrix}\boldsymbol{Q}_w\delta_{rs} & \boldsymbol{0} & \boldsymbol{Q}_w\boldsymbol{F}_{r+s-n_{0j}}^{(j)\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{F}_{r+s-n_{0i}}^{(i)}\boldsymbol{Q}_w & \boldsymbol{0} & \boldsymbol{E}_{ij}(n_0-n_{0i}+r,n_0-n_{0j}+s)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}_s^{3j\mathrm{T}}\\ \boldsymbol{\Omega}_s^{4j\mathrm{T}}\\ \boldsymbol{\Omega}_s^{5j\mathrm{T}}\end{bmatrix},$$

$$n_0=\max(n_{0i},n_{0j}),N\geqslant 0 \tag{7.8.13}$$

$$\boldsymbol{P}_{ij}(N)=\sum_{r=0}^{n_{3i}}\sum_{s=0}^{n_{3j}}[\boldsymbol{\Omega}_r^{3i},\boldsymbol{\Omega}_i^{4i},\boldsymbol{\Omega}_r^{6i}]\begin{bmatrix}\boldsymbol{Q}_w\delta_{rs} & \boldsymbol{0} & \boldsymbol{Q}_w\boldsymbol{F}_{r-s}^{(j)\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{F}_{s-r}^{(i)}\boldsymbol{Q}_w & \boldsymbol{0} & \boldsymbol{E}_{ij}(-r,-s)\end{bmatrix}\begin{bmatrix}\boldsymbol{\Omega}_s^{3j\mathrm{T}}\\ \boldsymbol{\Omega}_s^{4j\mathrm{T}}\\ \boldsymbol{\Omega}_s^{6j\mathrm{T}}\end{bmatrix},\quad N<0 \tag{7.8.14}$$

其中 $\boldsymbol{E}_{ij}(r,s)=\mathrm{E}[\boldsymbol{\varepsilon}_i(t+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t+s)]$ 为

$$\boldsymbol{E}_{ij}(r,s)=\sum_{k=0}^{\infty}\boldsymbol{F}_k^{(i)}\boldsymbol{Q}_w\boldsymbol{F}_{s-r+k}^{(j)\mathrm{T}} \tag{7.8.15}$$

证明 由 $\boldsymbol{w}(t)$ 与 $\boldsymbol{v}_i(t)(i=1,\cdots,L)$ 的相互独立性，应用(7.8.8)可得(7.8.15).应用(7.8.5)，(7.8.6)和(7.8.15)容易得(7.8.13)和(7.8.14)，其中用到了事实

$$\mathrm{E}[\boldsymbol{\varepsilon}_i(t-n_{0i}+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t-n_{0j}+s)]=\mathrm{E}[\boldsymbol{\varepsilon}_i(t-n_0+n_0-n_{0i}+r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t-n_0+n_0-n_{0j}+s)]=$$

$$\boldsymbol{E}_{ij}(n_0-n_{0i}+r,n_0-n_{0j}+s) \tag{7.8.16}$$

$$\mathrm{E}[\boldsymbol{\varepsilon}_i(t-r)\boldsymbol{\varepsilon}_j^{\mathrm{T}}(t-r)] = \boldsymbol{E}_{ij}(-r,-s) \tag{7.8.17}$$

【注 7.8.1】 考虑多传感器多通道 ARMA 信号融合 Wiener 滤波问题:

$$\boldsymbol{A}(q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}(q^{-1})\boldsymbol{w}(t) \tag{7.8.18}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{s}(t) + \boldsymbol{\eta}_i(t), \quad i = 1,\cdots,L \tag{7.8.19}$$

$$\boldsymbol{P}_i(q^{-1})\boldsymbol{\eta}_i(t) = \boldsymbol{R}_i(q^{-1})\boldsymbol{v}_i(t) \tag{7.8.20}$$

其中 $\boldsymbol{s}(t) \in R^m$ 为待估的 ARMA 信号,$\boldsymbol{\eta}_i(t) \in R^m$ 为 ARMA 有色观测噪声,观测 $\boldsymbol{y}_i(t) \in R^m$,$\boldsymbol{w}(t) \in R^r$ 和 $\boldsymbol{v}_i(t) \in R^m$ 是零均值、方差阵各为 $\boldsymbol{Q}_w$ 和 $\boldsymbol{Q}_{vi}$ 的相互独立白噪声.

问题可化为特殊的反卷积问题.事实上,由(7.8.19)和(7.8.20)有观测系统

$$\boldsymbol{P}_i(q^{-1})\boldsymbol{y}_i(t) = \boldsymbol{P}_i(q^{-1})\boldsymbol{s}(t) + \boldsymbol{R}_i(q^{-1})\boldsymbol{v}_i(t) \tag{7.8.21}$$

它是(7.8.2)取 $\boldsymbol{B}_i(q^{-1}) = \boldsymbol{P}_i(q^{-1})$ 的特殊情形.于是有如下定理.

【定理7.8.3】 多传感器Wiener滤波问题(7.8.18)~(7.8.20)有局部Wiener滤波器 $\hat{\boldsymbol{s}}_i(t \mid t+N)$ 为(7.8.4),局部滤波误差方差阵 $\boldsymbol{P}_i(N)$ 为(7.8.11)和(7.8.12),局部滤波误差互协方差阵 $\boldsymbol{P}_{ij}(N)(i \neq j)$ 为(7.8.13)和(7.8.14),但在所有公式中应置 $\boldsymbol{B}_i(q^{-1}) = \boldsymbol{P}_i(q^{-1})$ 来计算.

【例 7.8.1】 考虑带三传感器的二维跟踪系统的信息融合 Wiener 滤波问题:

$$(\boldsymbol{I}_2 + \boldsymbol{A}_1 q^{-1})\boldsymbol{s}(t) = \boldsymbol{C}_1 w(t-1) \tag{7.8.22}$$

$$\boldsymbol{y}_i(t) = \boldsymbol{s}(t) + \boldsymbol{v}_i(t), \quad i = 1,2,3 \tag{7.8.23}$$

$$(\boldsymbol{I}_2 + \boldsymbol{P}_{i1}q^{-1})\boldsymbol{v}_i(t) = (\boldsymbol{I}_2 + \boldsymbol{R}_{i1}q^{-1})\xi_i(t) \tag{7.8.24}$$

$$\boldsymbol{A}_1 = -\begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{C}_1 = \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix} \tag{7.8.25}$$

其中 T_0 为采样周期,$\boldsymbol{s}(t) = [s_1(t), s_2(t)]^{\mathrm{T}}$ 为待估信号,$s_1(t)$,$s_2(t)$ 和 $w(t)$ 分别为在时刻 tT_0 处运动目标的位置、速度和加速度,$\boldsymbol{y}_i(t)$ 为第 i 个传感器对位置和速度的观测信号,$\boldsymbol{v}_i(t)$ 为有色观测噪声,$w(t)$ 和 $\boldsymbol{\xi}_i(t)$ 是零均值、方差阵各为 σ_w^2 和 $\boldsymbol{Q}_{\xi i}$ 的相互独立高斯白噪声.在仿真中取 $T_0 = 0.25$,$\sigma_w^2 = 0.5$,$\boldsymbol{Q}_{\xi 1} = \mathrm{diag}(0.03, 0.05)$,$\boldsymbol{Q}_{\xi 2} = \mathrm{diag}(0.09, 0.15)$,$\boldsymbol{Q}_{\xi 3} = \mathrm{diag}(0.18, 0.3)$,$\boldsymbol{P}_{11} = \mathrm{diag}(0.12, 0.15)$,$\boldsymbol{P}_{21} = \mathrm{diag}(0.12, 0.1)$,$\boldsymbol{P}_{31} = \mathrm{diag}(0.15, 0.1)$,$\boldsymbol{R}_{\xi 1} = \mathrm{diag}(0.1, 0.2)$,$\boldsymbol{R}_{21} = \mathrm{diag}(0.15, 0.12)$,$\boldsymbol{R}_{31} = \mathrm{diag}(0.1, 0.12)$.问题是求 $\boldsymbol{s}(t)$ 的局部和融合滤波器 $\hat{\boldsymbol{s}}_i(t \mid t)(i = 1,2,3)$ 和 $\hat{\boldsymbol{s}}_0(t \mid t)$.

由定理7.8.3,可求得各子系统局部Wiener滤波器 $\hat{\boldsymbol{s}}_i(t \mid t+N)$ 及相应的误差方差阵 $\boldsymbol{P}_i(0)$ 的迹为

$$\mathrm{tr}\boldsymbol{P}_1(0) = 0.033\,3, \quad \mathrm{tr}\boldsymbol{P}_2(0) = 0.071\,7, \quad \mathrm{tr}\boldsymbol{P}_3(0) = 0.110\,5 \tag{7.8.26}$$

可求得按矩阵、对角阵和标量加权融合滤波器 $\hat{s}_0(t \mid t)$ 及相应误差方差阵 $\boldsymbol{P}_0^m(0)$,$\boldsymbol{P}_0^d(0)$ 和 $\boldsymbol{P}_0^s(0)$ 的迹为

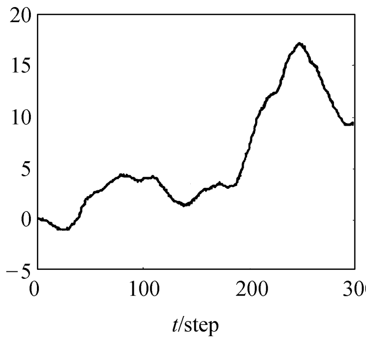
$$\mathrm{tr}\boldsymbol{P}_1^m(0) = 0.027\,67, \quad \mathrm{tr}\boldsymbol{P}_0^d(0) = 0.027\,71, \quad \mathrm{tr}\boldsymbol{P}_0^s(0) = 0.027\,77 \tag{7.8.27}$$

这引出精度关系

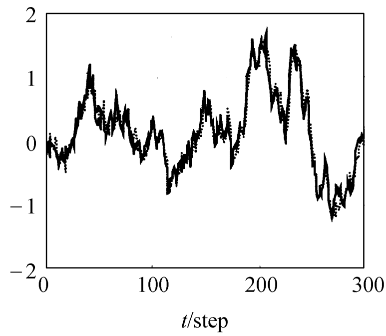
$$\mathrm{tr}\boldsymbol{P}_0^m(0) < \mathrm{tr}\boldsymbol{P}_0^d(0) < \mathrm{tr}\boldsymbol{P}_0^s(0) < \mathrm{tr}\boldsymbol{P}_i(0), \quad i = 1,2,3 \tag{7.8.28}$$

即融合估计精度高于每个局部估计精度,且按矩阵加权融合估计精度相对较高,而按标量融合估计精度相对较低,且按对角阵加权融合估计精度介于两者之间.但由仿真结果图

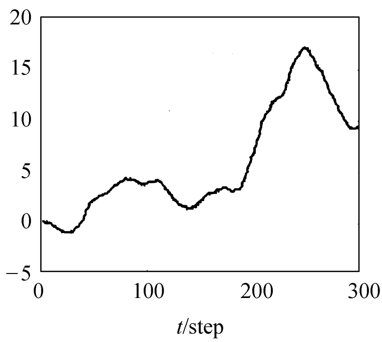
7.8.1 和图 7.8.2 可看到，三种加权融合估计精度并无显著差异. 因此从工程实时应用角度，采用按标量加权融合估计只稍微损失了一点精度，但却可显著减小计算负担，便于实时应用. 在图 7.8.1 中，实线代表真实值 $s_i(t)$，虚线代表在三种加权下的融合滤波估值 $\hat{s}_{0i}(t \mid t)$，$i = 1,2$. 图 7.8.2 为局部滤波和三种加权融合滤波绝对误差累积曲线，从中可看到三种加权融合滤波绝对误差累积曲线均在每个局部滤波绝对误差累积曲线下方，即它们的精度高于每个局部估计精度，但三条融合估计绝对误差累积曲线几乎相重合，这说明对本例三种加权融合估计精度无显著差别.



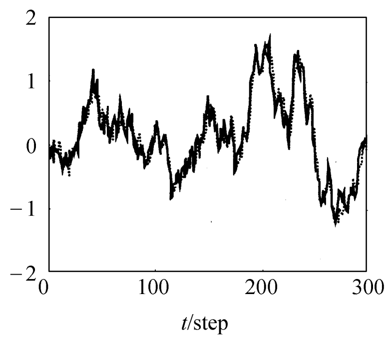
(a) 位置 $s_1(t)$ 和按矩阵加权融合 Wiener 滤波器 $\hat{s}_{01}(t|t)$



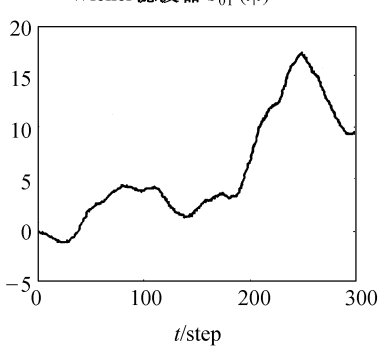
(b) 速度 $s_2(t)$ 和按矩阵加权融合 Wiener 滤波器 $\hat{s}_{02}(t|t)$



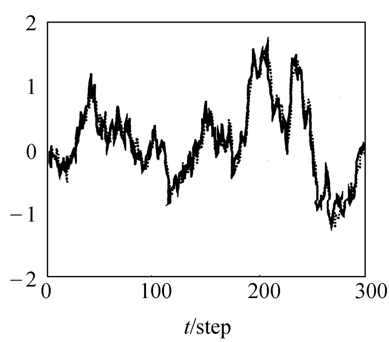
(c) 位置 $s_1(t)$ 和按对角阵加权融合 Wiener 滤波器 $\hat{s}_{01}(t|t)$



(d) 速度 $s_2(t)$ 和按对角阵加权融合 Wiener 滤波器 $\hat{s}_{02}(t|t)$



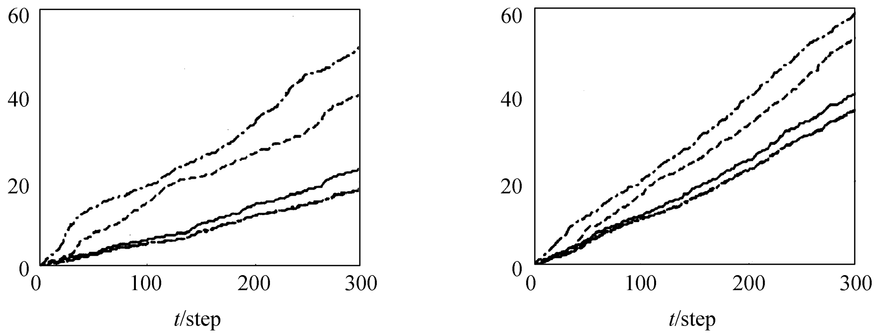
(e) 位置 $s_1(t)$ 和按标量加权融合 Wiener 滤波器 $\hat{s}_{01}(t|t)$



(f) 速度 $s_2(t)$ 和按标量加权融合 Wiener 滤波器 $\hat{s}_{02}(t|t)$

图 7.8.1　信号 $\boldsymbol{s}(t)$ 和按三种方式加权融合 Wiener 滤波器 $\hat{\boldsymbol{s}}_0(t \mid t)$

$(\boldsymbol{s}(t) = [s_1(t), s_2(t)]^{\mathrm{T}}, \hat{\boldsymbol{s}}_0(t \mid t) = [\hat{s}_{01}(t \mid t), \hat{s}_{02}(t \mid t)]^{\mathrm{T}})$



(a) 位置滤波绝对误差累积曲线　　(b) 速度滤波绝对误差累积曲线

—:单传感器 1 的累积滤波绝对误差曲线
- -:单传感器 2 的累积滤波绝对误差曲线
—·:单传感器 3 的累积滤波绝对误差曲线
……: 按矩阵加权融合的累积滤波绝对误差曲线
—·: 按对角阵加权融合的累积滤波绝对误差曲线
-·-: 按标量加权融合的累积滤波绝对误差曲线

图 7.8.2　局部和融合滤波绝对误差累积曲线

参　考　文　献

1 邓自立.自校正滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2003

2 邓自立.卡尔曼滤波与维纳滤波——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2001

3 邓自立.最优滤波理论及其应用——现代时间序列分析方法.哈尔滨:哈尔滨工业大学出版社,2000

4 Gan Q, Harris C J. Comparison of Two Measurement Fusion Methods for Kalman Filter – Based Measurement Data Fusion IEEE Trans. Aerospace and Electronic Systems, 2001,37(1):273 ~ 280

5 Sun Shuli, Deng Zili. Multi – sensor Optimal Information Fusion Kalman Filter. Automatica, 2004, 40, 1017 ~ 1023

6 邓自立,祁荣宾.多传感器信息融合次优稳态 Kalman 滤波器.中国学术期刊文摘(科技快报),2000,6(2):183 ~ 184

7 孙书利,邓自立.多传感器线性最小方差最优信息融合准则.科学技术与工程,2004,4(5):334 ~ 336

8 邓自立,高媛,崔崇信.多传感器按对角阵加权信息融合 Kalman 滤波器.科学技术与工程,2004,4(7):519 ~ 521

9 高媛,白敬刚,邓自立.多传感器单通道信息融合 Wiener 滤波器.科学技术与工程,2004,4(7):522 ~ 525

10 邓自立,高媛.两传感器信息融合超前 k 步稳态最优 Kalman 预报器.科学技术与工程,2004,4(5):337 ~ 340

11 邓自立,崔崇信,白敬刚.基于稳态 Kalman 滤波的两种观测融合方法的功能等价性.科学技术与工程,2004,4(11):897 ~ 902

12 邓自立,高媛.快速信息融合 Kalman 滤波器.控制与决策,2005,20(1)

13 邓自立,高媛,马建为.两传感器信息融合稳态最优 Kalman 滤波器.科学技术与工程,2003,3(3):213 ~ 215

14 邓自立,高媛,马建为.两传感器信息融合最优白噪声反卷积 Wiener 滤波器.科学技术与工程,2003,3(3):216 ~ 218

15 邓自立,高媛,王琳等.两传感器信息融合 Wiener 滤波器、平滑器和预报器.2004 中国控制与决策学术年会论文集.沈阳:东北大学出版社,2004.206 ~ 208

16 高媛，梁佐江，王欣等．两传感器信息融合 Wiener 反卷积滤波器．2004 中国控制与决策学术年会论文集．沈阳：东北大学出版社，2004．212 ~ 214
17 Deng Z L，Zhang H S，Liu S J，Zhou L. Optimal and Self－Tuning White Noise Estimators with Applications to Deconvolution and Filtering Problems．Automatica，1996，32（2）：199 ~ 216
18 孟华，石莹，邓自立．带位置和速度观测的信息融合 Kalman 跟踪滤波器．科学技术与工程，2005，5（1）：14 ~ 18
19 邓自立，高媛，张明波．ARMA 信号自校正信息融合 Wiener 滤波器．科学技术与工程，2004，4（9）：749 ~ 752
20 邓自立，石莹，孟华．基于 ARMA 新息模型与 Riccati 方程的两种 Kalman 跟踪滤波器的等价性．科学技术与工程，2004，4（11）：894 ~ 896
21 邓自立，马建为，高媛．两传感器自校正信息融合 Kalman 滤波器．科学技术与工程，2003，3（4）：321 ~ 324
22 邓自立，马建为，高媛．两传感器自校正信息融合白噪声反卷积滤波器．科学技术与工程，2003，3（4）：325 ~ 327
23 Mendel J．White Noise Estimators for Seismic Data Processing in Oil Exploration．IEEE Trans．Automatic Control，1977，22（5）：694 ~ 706
24 郑大钟．线性系统理论．第 2 版．北京：清华大学出版社，2002
25 邓自立，李云，王欣．带 MA 有色观测噪声的单通道多传感器信息融合 Wiener 滤波器．科学技术与工程，2005，5（5）：271 ~ 275
26 毛琳，邓自立．多传感器信息融合 ARMA 信号 Wiener 预报器．2005 中国控制与决策学术年会论文集．沈阳：东北大学出版社，2005
27 邓自立，崔崇信．多传感器全局最优观测融合白噪声反卷积滤波器．科学技术与工程，2005，5（5）：267 ~ 270
28 邓自立，李云，王欣．多传感器最优信息融合白噪声反卷积滤波器．2005 中国控制与决策学术年会论文集．沈阳：东北大学出版社，2005
29 孟华，石莹，邓自立．加权观测融合 Kalman 跟踪滤波器．2005 中国控制与决策学术年会论文集．沈阳：东北大学出版社，2005
30 石莹，孟华，邓自立．多传感器信息融合稳态 Kalman 跟踪滤波器．2005 中国控制与决策学术年会论文集．沈阳：东北大学出版社，2005
31 邓自立．Convergence Analysis of Gevers－Wouters Algorithm for Scalar Spectral Facorization．科学技术与工程，2005，5（1）：7 ~ 13
32 邓自立，梁佐江．多传感器分布式信息融合 Wiener 信号滤波器．科学技术与工程，2005，5（9）：539 ~ 542
33 邓自立．Lyapunov 方程迭代解的指数收敛性．科学技术与工程，2005，5（9）：543 ~ 546
34 邓自立，王欣，李云．多传感器最优信息融合白噪声反卷积滤波器．电子学报，2005，33（5）：104 ~ 106
35 Deng Z L，Gao Y，Mao L，Li Y，Hao G．New approach to information fusion steady-state Kalman filtering．Automatica，2005，41（10）