硕士研究生入学考试

数学复习与解题指南

（下册）

同济大学应用数学系

徐建平 蒋福民 范麟馨 钱伟民 编

同济大学出版社

硕士研究生入学考试
数学复习与解题指南
（下册）

同 济 大 学 应 用 数 学 系

徐建平 蒋福民 范麟馨 钱伟民 编

同济大学出版社

图书在版编目(CIP)数据

硕士研究生入学考试数学复习与解题指南/徐建平编.
上海:同济大学出版社,2005.5
ISBN 7-5608-3011-0

Ⅰ.硕… Ⅱ.徐… Ⅲ.高等数学—研究生—入学考试—自学参考资料 Ⅳ.013

中国版本图书馆 CIP 数据核字(2005)第 008006 号

内容提要

本书由高等数学、线性代数和概率统计三部分组成。书中按教育部最新颁布的考研大纲对各部分的重要概念和基本定理(定理和公式)作了系统的归纳和提炼,着重讨论基本题型与解题方法,并对部分例题进行了详尽的分析和总结,以拓宽读者的思路,提高他们的分析能力。

本书从历届考题中筛选了近 700 道典型例题,选辑了 100 多道习题,并附有简答。书末录入近两年硕士入学考试数学试题及参考解答。本书可供报考硕士研究生的读者作为复习数学使用,也可作为高等院校师生的数学教学参考书。

硕士研究生入学考试
数学复习与解题指南(下册)
同济大学应用数学系
徐建平 蒋福民 范麟馨 钱伟民 编
责任编辑 李炳钊 责任校对 郁 峰 封面设计 李志云

出 版
发 行 同济大学出版社
(上海四平路 1239 号 邮编 200092 电话 021-65985622)
经 销 全国各地新华书店
印 刷 同济大学印刷厂印刷
开 本 889mm×1194mm 1/16
印 张 14
字 数 448000
印 数 1—4100
版 次 2005 年 5 月第 1 版 2005 年 5 月第 1 次印刷
书 号 ISBN 7-5608-3011-0/O·267
定 价 20.00 元

本书若有印装质量问题,请向本社发行部调换

前　言

随着我国高等教育事业的迅速发展，越来越多的人接受了高等教育，他们中的许多人都希望获得更高的学历，而我国四个现代化事业的蓬勃发展也需要更多更高层次的人才。正是因为个人的追求与国家的需要相一致，所以近年来报考硕士研究生的人数在逐年增加，2005 年已突破 100 万。数学是一门重要的基础课，广大考生最关心的问题便是怎样按照考试大纲进行复习，提高复习效果，从而取得理想的考试成绩。要达到这一点，首先必须要了解考试大纲中所规定的内容，分清哪些是主要的，哪些是次要的，这些内容在深度和广度上要求达到什么程度。其次要熟悉试卷中经常出现的题型，以及一些典型的甚至比较特殊的解题方法，这样，才能通过复习，深刻理解概念，牢固掌握解题方法，并在考试中付诸实践，从而取得好成绩。

为满足考生的需要，我们认真总结了近十年来所举办的考研复习辅导班积累的经验和资料，深入研究了历年来考试题型的结构、形式和特点，按照教育部新近颁布的数学考研大纲，编写了这本书。它不仅对报考研究生的人员是一本有用的辅导材料，对正在本科阶段学习高等数学、线性代数和概率统计的学生同样也是一本有价值的参考书。这不仅因为本书的内容与所选的例题、习题与在普通高校中使用很广的由同济大学应用数学系所编写、高等教育出版社出版的《高等数学》(第五版)和《线性代数》(第四版)有很好的衔接，而且因为本书所总结的题型和解题方法对本科生学好高等数学和线性代数以及概率统计也大有裨益。

本书的编写突出一个宗旨：应试针对性强，力求使考生通过较短时间的复习，能达到考试大纲所规定的要求，并掌握一整套常用的基本解题方法。

本书由“高等数学”、“线性代数”和“概率统计”三个部分组成，具备以下几个特色：(1) 吸收了历届试卷中经常出现的典型考题，并进行分析解答；(2) 许多章节对基本题型和解题方法作了总结归纳，便于学生总结提高；(3) 指出解题中容易出错的地方，了解命题意图或考生容易误入的“陷阱”；(4) 高等数学各章均包含一定数量的习题及简答，做好这些习题，有利于提高考生解题的基本功；线性代数和概率统计最后都安排了一些综合练习题，并附有简答，旨在提高考生的综合解题能力。

本书每章开头有复习要求、基本概念和理论，目的是帮助考生掌握好各章的内容。复习要求中对理论部分的要求分“了解”和“理解”两档，理解高于了解；对运算部分的要求分“会”、“掌握”和“熟练掌握”三档。本书录入 2005 年硕士研究生入学考试数学试题参考答案及评分标准，供考生参考。

本书由同济大学应用数学系徐建平、蒋福民、范麟馨、钱伟民等编写，由于编者水平有限，书中难免有错误和不妥之处，恳请同行批评指正。

编者

2005.1.28

目 录

线性代数

第一章 行列式 …… (2)
一、复习与考试要求 …… (2)
二、基本概念与理论 …… (2)
三、基本题型与解题方法 …… (3)
四、小结 …… (16)

第二章 矩阵及其运算 …… (18)
一、复习与考试要求 …… (18)
二、基本概念与理论 …… (18)
三、基本题型与解题方法 …… (23)
四、小结 …… (37)

第三章 向量组的线性相关性与矩阵的秩 …… (38)
一、复习与考试要求 …… (38)
二、基本概念与理论 …… (38)
三、基本题型与解题方法 …… (41)
四、小结 …… (55)

第四章 线性方程组 …… (56)
一、复习与考试要求 …… (56)
二、基本概念与理论 …… (56)
三、基本题型与解题方法 …… (57)
四、小结 …… (76)

第五章 矩阵的特征值与特征向量 …… (77)
一、复习与考试要求 …… (77)
二、基本概念与理论 …… (77)
三、基本题型与解题方法 …… (78)
四、小结 …… (96)

第六章 二次型 …… (98)
一、复习与考试要求 …… (98)
二、基本概念与理论 …… (98)
三、基本题型与解题方法 …… (99)
四、小结 …… (110)
综合练习题 …… (111)
简答与提示 …… (112)

概率统计

第一章 随机事件和概率 …… (116)
一、复习与考试要求 …… (116)
二、基本概念与理论 …… (116)
三、例题分析 …… (118)
四、小结 …… (123)

第二章 随机变量及其分布 …… (124)
一、复习与考试要求 …… (124)
二、基本概念与理论 …… (124)
三、例题分析 …… (126)
四、小结 …… (132)

第三章 二维随机变量及其分布 …… (133)
一、复习与考试要求 …… (133)
二、基本概念与理论 …… (133)
三、例题分析 …… (136)
四、小结 …… (143)

第四章 随机变量的数字特征 …… (144)
一、复习与考试要求 …… (144)
二、基本概念与理论 …… (144)
三、例题分析 …… (146)
四、小结 …… (152)

第五章 大数定律和中心极限定理 …… (153)
一、复习与考试要求 …… (153)
二、基本概念与理论 …… (153)
三、例题分析 …… (154)
四、小结 …… (156)

第六章 数理统计的基本概念 …… (157)
一、复习与考试要求 …… (157)
二、基本概念与理论 …… (157)
三、例题分析 …… (159)
四、小结 …… (162)

第七章 参数估计 …… (164)
一、复习与考试要求 …… (164)
二、基本概念与理论 …… (164)
三、例题分析 …… (166)
四、小结 …… (174)

第八章 假设检验 …… (175)
一、复习与考试要求 …… (175)
二、基本概念与理论 …… (175)
三、例题分析 …… (177)
四、小结 …… (179)

第九章 综合能力训练 …… (180)
一、例题分析 …… (180)
二、小结 …… (189)
综合练习题 …… (190)
习题简答 …… (192)

附录 全国硕士研究生入学统一考试答案和评分参考
2005 年数学(一)试卷 …… (195)
2005 年数学(二)试卷 …… (201)
2005 年数学(三)试卷 …… (206)
2005 年数学(四)试卷 …… (212)

线性代数

线性代数的复习，根据“考试大纲”的要求以及历年的考题情况，考生需掌握以下诸方面的内容：(1) 行列式的计算；(2) 矩阵的运算；(3) 线性方程组的求解；(4) 特征值和特征向量的计算；(5) 二次型的化简；(6) 有关的基本概念和基本理论.

有关线性代数内容的考题在数学试卷一、试卷二中约占 20%；在试卷三、试卷四中约占 26%.

考生如能按照本教材要求认真复习，掌握基本题型和解题方法，则不难达到考试大纲的要求，且能获得满意的成绩.

第一章　行列式

本章重点是利用行列式的性质计算行列式.

一、复习与考试要求

(1) 了解行列式的概念,掌握行列式的性质.

(2) 熟练应用行列式的性质和行列式按行(列)展开定理计算行列式.

(3) 掌握按某行(列)展开的方法建立特殊的 n 阶行列式的递推关系,再求出该行列式的值.

二、基本概念与理论

1. n 阶行列式的定义

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{(j_1,j_2,\cdots,j_n)} (-1)^{\tau(j_1,j_2,\cdots,j_n)} a_{1j_1} \cdot a_{2j_2} \cdot \cdots \cdot a_{nj_n}.$$

(1) 记号 $\sum\limits_{(j_1,j_2,\cdots,j_n)}$ 表明:对列指标所有的 n 元全排列求和,由此可知 n 阶行列式是有 n 项的代数和.

(2) 记号 $(-1)^{\tau(j_1,\cdots,j_n)}$ 确定了每一项前所带的正负号,其中 $\tau(j_1,j_2,\cdots,j_n)$ 是排列 $(j_1,j_2,\cdots,j_n)$ 的逆序数.

(3) 每一项 $a_{1j_1} \cdot a_{2j_2} \cdot \cdots \cdot a_{nj_n}$ 表明有 n 个元素相乘,而且这 n 个元素必须取自于不同行,不同列.

2. n 阶行列式的性质

(1) $|A| = |A^{\mathrm{T}}|$　(A^{T} 是 A 的转置矩阵)

此性质表明:在行列式中,"行"与"列"具有同等地位,即对"行"所具有的性质,对"列"也同样成立,反之亦然.

(2) 记 $A = (\alpha_1,\cdots,\alpha_n)$.这里 α_i 是 A 的第 i 列,则行列式关于"列"具有线性,即

$$|\alpha_1,\cdots,\alpha_i+\beta_i,\cdots,\alpha_n| = |\alpha_1,\cdots,\alpha_i,\cdots,\alpha_n| + |\alpha_1,\cdots,\beta_i,\cdots,\alpha_n|$$

$$|\alpha_1,\cdots,k\cdot\alpha_i,\cdots,\alpha_n| = k\cdot|\alpha_1,\cdots,\alpha_i,\cdots,\alpha_n|, \quad k \text{ 为常数}.$$

推论　行列式中若有一列为 0,则此行列式的值为 0.

(3) $|\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j,\cdots,\alpha_n| = -|\alpha_1,\cdots,\alpha_j,\cdots,\alpha_i,\cdots,\alpha_n|$,即行列式互换两列,则行列式反号.

推论 1　行列式中若有两列相同,则此行列式值为 0.

推论 2　行列式中若有两列对应成比例,则此行列式值为 0.

(4) $|\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j,\cdots,\alpha_n| = |\alpha_1,\cdots,\alpha_i,\cdots,k\cdot\alpha_i+\alpha_j,\cdots,\alpha_n|$.

(5) 行列式按行(列)展开定理

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} = \begin{cases} |A|, & k = i; \\ 0, & k \neq i. \end{cases}$$

$$a_{1j}A_{1s} + a_{2j}A_{2s} + \cdots + a_{nj}A_{ns} = \begin{cases} |A|, & s = j; \\ 0, & s \neq j. \end{cases}$$

3. 克莱默(Cramer)法则

线性方程组 $A_{n\times n}\cdot X_{n\times 1}=\boldsymbol{b}_{n\times 1}$,**若** $|A|\neq 0$,**则有惟一解**:

$$x_j=\frac{|A_j|}{|A|},\quad j=1,2,\cdots,n$$

其中,A_j **是用列向量** $\boldsymbol{b}$ **换掉** A **中的第** j **列而得到的矩阵**.

4. 范德蒙(Vandermende)行列式

$$D_n=\begin{vmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\\ x_1^2 & x_2^2 & \cdots & x_n^2\\ \vdots & \vdots & \ddots & \vdots\\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1}\end{vmatrix}=\prod_{n\geqslant i>j\geqslant 1}(x_i-x_j).$$

5. n **阶矩阵的行列式**

(1) $|\boldsymbol{A}\cdot\boldsymbol{B}|=|\boldsymbol{A}|\cdot|\boldsymbol{B}|=|\boldsymbol{B}\cdot\boldsymbol{A}|$,　$|\boldsymbol{A}^{\mathrm{T}}\cdot\boldsymbol{A}|=|\boldsymbol{A}|^2$

(2) $|k\boldsymbol{A}|=k^n\cdot|\boldsymbol{A}|$,　k **为任意常数**.

(3) $\begin{vmatrix}\boldsymbol{A}_{m\times m} & \boldsymbol{O}\\ \boldsymbol{C}_{n\times m} & \boldsymbol{B}_{n\times n}\end{vmatrix}=|\boldsymbol{A}_{m\times m}|\cdot|\boldsymbol{B}_{n\times n}|$

但是 $\begin{vmatrix}\boldsymbol{A} & \boldsymbol{D}\\ \boldsymbol{C} & \boldsymbol{B}\end{vmatrix}\neq|\boldsymbol{A}|\cdot|\boldsymbol{B}|-|\boldsymbol{C}|\cdot|\boldsymbol{D}|$,**其中**,$\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{D}$ **皆为方阵**.

三、基本题型与解题方法

例 1-1　计算

$$D_4=\begin{vmatrix}4 & 2 & 2 & 2\\ 2 & 4 & 2 & 2\\ 2 & 2 & 4 & 2\\ 2 & 2 & 2 & 4\end{vmatrix}.$$

分析　易见行列式的每行(列)都有公因子 2,而且公因子 2 提出后各行元素之和均为 5,因此可把各列元素都加到第一列上,再提取公因子 5,然后将改变后的行列式化成三角行列式,从而计算出值.

解

$$D_4=2^4\begin{vmatrix}2 & 1 & 1 & 1\\ 1 & 2 & 1 & 1\\ 1 & 1 & 2 & 1\\ 1 & 1 & 1 & 2\end{vmatrix}\xrightarrow[c_1\div 5]{c_1+c_2+c_3+c_4}2^4\times 5\begin{vmatrix}1 & 1 & 1 & 1\\ 1 & 2 & 1 & 1\\ 1 & 1 & 2 & 1\\ 1 & 1 & 1 & 2\end{vmatrix}\xrightarrow[r_4-r_1]{\substack{r_2-c_1\\ r_3-r_1}}80\begin{vmatrix}1 & 1 & 1 & 1\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{vmatrix}=80.$$

例 1-2　计算行列式

$$D_5=\begin{vmatrix}5 & 3 & -1 & -2 & 1\\ 1 & 1 & 4 & 2 & 0\\ 0 & -7 & -3 & 1 & 0\\ 0 & 0 & 3 & 2 & 0\\ 0 & 6 & -2 & -4 & 2\end{vmatrix}.$$

分析　此行列式的特点是第 1 列只有两个元素不等于 0(第 5 列,第 4 行也如此),这种情形是适合于按

列计算(行展开进行降阶计算). 此外第 5 行有公因子 2,且第 1 行与第 5 行有四个元素成比例.

解

$$D_5 \xrightarrow{r_1-\frac{1}{2}r_5} \begin{vmatrix} 5 & 0 & 0 & 0 & 0 \\ 1 & 1 & 4 & 2 & 0 \\ 0 & -7 & -3 & 1 & 0 \\ 0 & 0 & 3 & 2 & 0 \\ 0 & 6 & -2 & -4 & 2 \end{vmatrix} \xrightarrow{r_1\text{ 展开}} 5 \begin{vmatrix} 1 & 4 & 2 & 0 \\ -7 & -3 & 1 & 0 \\ 0 & 3 & 2 & 0 \\ 6 & -2 & -4 & 2 \end{vmatrix}$$

$$\xrightarrow{c_4\text{ 展开}} 5\times 2 \begin{vmatrix} 1 & 4 & 2 \\ -7 & -3 & 1 \\ 0 & 3 & 2 \end{vmatrix} \xrightarrow{r_2+7r_1} 10 \begin{vmatrix} 1 & 4 & 2 \\ 0 & 25 & 15 \\ 0 & 3 & 2 \end{vmatrix}$$

$$\xrightarrow{c_1\text{ 展开}} 10 \begin{vmatrix} 25 & 15 \\ 3 & 2 \end{vmatrix} \xrightarrow{r_1\div 5} 10\times 5 \begin{vmatrix} 5 & 3 \\ 3 & 2 \end{vmatrix} = 50.$$

例 1-3　计算

$$D_6 = \begin{vmatrix} 1 & 1 & 1 & & & a \\ 1 & 1 & 1 & & a & \\ 1 & 1 & 1 & a & & \\ & & b & 1 & & \\ & b & & & 1 & \\ b & & & & & 1 \end{vmatrix}, \qquad ab \neq 0,\text{未写出的元素为 } 0.$$

分析　消去第 1,2,3 行中的 a,将行列式化简为第 4,5,6 列都只有一个非零元素,再按此 3 列陆续展开降阶计算.

解

$$D_6 \xrightarrow[r_3-ar_4]{\substack{r_1-ar_6 \\ r_2-ar_5}} \begin{vmatrix} 1-ab & 1 & 1 & & & \\ 1 & 1-ab & 1 & & & \\ 1 & 1 & 1-ab & & & \\ & & & b & 1 & \\ & & b & & & 1 \\ b & & & & & \end{vmatrix}$$

$$\xrightarrow[\text{展开}]{\substack{\text{分别按第} \\ 6,5,4\text{ 列}}} \begin{vmatrix} 1-ab & 1 & 1 \\ 1 & 1-ab & 1 \\ 1 & 1 & 1-ab \end{vmatrix} \xrightarrow[\text{第一行提取 } 3-ab]{r_1+r_2+r_3} (3-ab) \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1-ab & 1 \\ 1 & 1 & 1-ab \end{vmatrix}$$

$$\xrightarrow[r_3-r_1]{r_2-r_1} (3-ab) \begin{vmatrix} 1 & 1 & 1 \\ 0 & -ab & 0 \\ 0 & 0 & -ab \end{vmatrix} = (3-ab)a^2b^2.$$

例 1-4　计算

$$D_{2n} = \begin{vmatrix} a_n & & & & & b_n \\ & \ddots & & & \iddots & \\ & & a_1 & b_1 & & \\ & & c_1 & d_1 & & \\ & \iddots & & & \ddots & \\ c_n & & & & & d_n \end{vmatrix}, \qquad a_ib_ic_id_i \neq 0, i = 1,\cdots,n.$$

分析　这个行列式的特点是不言自明的，适合于按行(列)展开降阶计算．也可以应用例 1-4 的方法将它化成下三角行列式，再计算出值．

$$D_{2n}\xrightarrow{\text{按 } r_1 \text{ 展开}} a_n(-1)^{1+1}\begin{vmatrix} a_{n-1} & & & & b_{n-1} & 0 \\ & \ddots & & \iddots & & \\ & & a_1 & b_1 & & \\ & & c_1 & d_1 & & \\ & \iddots & & & \ddots & \\ c_{n-1} & & & & d_{n-1} & 0 \\ 0 & & & & 0 & d_n \end{vmatrix}$$

$$+b_n(-1)^{1+2n}\begin{vmatrix} 0 & a_{n-1} & & & & b_{n-1} \\ & & \ddots & & \iddots & \\ & & & a_1 & b_1 & \\ & & & c_1 & d_1 & \\ & & 0 & \iddots & & \ddots \\ & & c_n-1 & & & d_{n-1} \\ c_n & & & & 0 & \end{vmatrix}$$

$$=(a_nd_n-b_nc_n)D_{2(n-1)}$$

所以

$$D_{2n}=(a_nd_n-b_nc_n)(a_{n-1}d_{n-1}-b_{n-1}c_{n-1})D_{2(n-2)}\cdots$$

$$=(a_nd_n-b_nc_n)\cdots(a_2d_2-b_2c_2)D_2=\prod_{i=1}^{n}(a_id_i-b_ic_i).$$

例 1-5　计算

$$D=\begin{vmatrix} o & a & b & o \\ c & o & o & d \\ e & o & f & o \\ o & g & o & h \end{vmatrix}.$$

分析　行列式元素中零元素较多，故可考虑按定义去求，也可考虑用按行展开定理去做．

解法 1　按定义，只要考虑非零项：$acfh$ 与 $bdeg$．$acfh$ 所带的符号为：$(-1)^{\tau(2134)}=(-1)^1=-1$．$bdeg$ 所带的符号为：$(-1)^{\tau(3412)}=(-1)^4=1$．所以，$D=bdeg-acfh$．

解法 2

$$D=-a\cdot\begin{vmatrix} c & o & d \\ e & f & o \\ o & o & h \end{vmatrix}+b\cdot\begin{vmatrix} c & o & d \\ e & o & o \\ o & g & h \end{vmatrix}$$

$$=-ah\cdot\begin{vmatrix} c & o \\ e & f \end{vmatrix}-be\cdot\begin{vmatrix} o & d \\ g & h \end{vmatrix}$$

$$=-ahcf+begd.$$

例 1-6　若 n 阶方阵 $\boldsymbol{A}$ 的元素皆为整数，试证 $2\boldsymbol{A}-\boldsymbol{E}$ 可逆．

分析　矩阵可逆的等价条件有好多种说法．本题可按定义论证 $|2\boldsymbol{A}-\boldsymbol{E}|\neq 0$，从而 $2\boldsymbol{A}-\boldsymbol{E}$ 可逆．

证　因 $\boldsymbol{A}$ 的元素皆为整数，故 $2\boldsymbol{A}-\boldsymbol{E}$ 的对角线元素皆为奇数，而非对角线元素皆为偶数．按定义考虑行列式 $|2\boldsymbol{A}-\boldsymbol{E}|$：只有所有对角线元素相乘这一项是奇数，而其余的 $(n!)-1$ 项均为偶数，因此，它们的代数和不等于零．

例 1-7　n 阶方阵 $\boldsymbol{A}$ 的元素全为 1 和 -1 组成，试证：$|\boldsymbol{A}|$ 可被 2^{n-1} 整除.

分析　行列式计算中，很基本的一个方法是通过消零，然后降阶. 在行列式中有很多 1 或 -1 时，消零是方便的.

证　不妨设 $a_{11}=1$，把第一列 a_{11} 以下的元素全消成零，即

$$|\boldsymbol{A}|=\begin{vmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & \boldsymbol{B}_{n-1} & \\ 0 & & & \end{vmatrix}=|\boldsymbol{B}_{n-1}|$$

而 $\boldsymbol{B}_{n-1}$ 中的元素全为 $2,-2,0$. 故可在每一行中提出公因子 2，则 $|\boldsymbol{B}_{n-1}|=2^{n-1}\cdot|\boldsymbol{C}_{n-1}|$. 即

$$|\boldsymbol{A}|=|\boldsymbol{B}_{n-1}|=2^{n-1}\cdot|\boldsymbol{C}_{n-1}|.$$

而 $\boldsymbol{C}_{n-1}$ 中的元素皆为整数，故 $|\boldsymbol{C}_{n-1}|$ 是整数，因此 $|\boldsymbol{A}|$ 可被 2^{n-1} 整除.

例 1-8　计算

$$D=\begin{vmatrix} 1 & -1 & 1 & x-1 \\ 1 & -1 & x+1 & -1 \\ 1 & x-1 & 1 & -1 \\ x+1 & -1 & 1 & -1 \end{vmatrix}.$$

分析　此行列式的特点是各行元素之和均为 x，故可把各列都加到第一列上，然后把公因子 x 提出，则第一列均为 1，再考虑消零.

解

$$D=x\cdot\begin{vmatrix} 1 & -1 & 1 & x-1 \\ 1 & -1 & x+1 & -1 \\ 1 & x-1 & 1 & -1 \\ 1 & -1 & 1 & -1 \end{vmatrix}=x\cdot\begin{vmatrix} 1 & 0 & 0 & x \\ 1 & 0 & x & 0 \\ 1 & x & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}=x^4$$

例 1-9　计算

$$D_n=\begin{vmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a \end{vmatrix}$$

分析　行列式的特点是各行元素之和是常数，则可把各列都加到第一列上去，把此常数提出，第一列全是 1，再考虑消零.

解

$$D_n=[a+(n-1)b]\cdot\begin{vmatrix} 1 & b & \cdots & b \\ 1 & a & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ 1 & b & \cdots & a \end{vmatrix}=[a+(n-1)b]\cdot\begin{vmatrix} 1 & b & \cdots & b \\ 0 & a-b & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a-b \end{vmatrix}$$

$$=[a+(n-1)b]\cdot(a-b)^{n-1}$$

例 1-10　设多项式

$$f(x)=\begin{vmatrix} a_{11}+x & a_{12}+x & a_{13}+x & a_{14}+x \\ a_{21}+x & a_{22}+x & a_{23}+x & a_{24}+x \\ a_{31}+x & a_{32}+x & a_{33}+x & a_{34}+x \\ a_{41}+x & a_{42}+x & a_{43}+x & a_{44}+x \end{vmatrix},$$

问：$f(x)$ 是几次多项式？

分析　要先把此行列式利用行列式的性质把它化简.

解　把第一行的(-1)倍分别加到第二、三、四行，得

$$f(x)=\begin{vmatrix} a_{11}+x & a_{12}+x & a_{13}+x & a_{14}+x \\ a_{21}-a_{11} & a_{22}-a_{12} & a_{23}-a_{13} & a_{24}-a_{14} \\ a_{31}-a_{11} & a_{32}-a_{12} & a_{33}-a_{13} & a_{34}-a_{14} \\ a_{41}-a_{11} & a_{42}-a_{12} & a_{43}-a_{13} & a_{44}-a_{14} \end{vmatrix},$$

再按第一行展开，可知 $f(x)$ 是 1 次多项式.

例 1-11　设

$$f(x)=\begin{vmatrix} x-2 & x-1 & x-2 & x-3 \\ 2x-2 & 2x-1 & 2x-2 & 2x-3 \\ 3x-3 & 3x-2 & 4x-5 & 3x-5 \\ 4x & 4x-3 & 5x-7 & 4x-3 \end{vmatrix},$$

问：方程 $f(x)=0$ 有几个根？

分析　本题实质是一个计算四阶行列式的问题. 通常采用行列式的性质，把它化简.

解　由行列式的性质知：

$$f(x)=\begin{vmatrix} x-2 & x-1 & x-2 & x-3 \\ 2 & 1 & 2 & 3 \\ 3x-3 & 3x-2 & 4x-5 & 3x-5 \\ 4x & 4x-3 & 5x-7 & 4x-3 \end{vmatrix}=x\cdot\begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 3 \\ 3x-3 & 3x-2 & 4x-5 & 3x-5 \\ 4x & 4x-3 & 5x-7 & 4x-3 \end{vmatrix}$$

$$=x\cdot\begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 3 \\ -3 & -2 & x-5 & -5 \\ 0 & -3 & x-7 & -3 \end{vmatrix}=x\cdot\begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & x-2 & -2 \\ 0 & -3 & x-7 & -3 \end{vmatrix}$$

$$=x\cdot\begin{vmatrix} -1 & 0 & 1 \\ 1 & x-2 & -2 \\ -3 & x-7 & -3 \end{vmatrix}=x\cdot\begin{vmatrix} -1 & 0 & 1 \\ 0 & x-2 & -1 \\ 0 & x-7 & -6 \end{vmatrix}$$

$=5x(x-1)=0$，　**故方程有 2 个根.**

例 1-12　方程求根：

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & x \\ 4 & 1 & 9 & x^2 \\ 8 & -1 & 27 & x^3 \end{vmatrix}=0.$$

分析　等号左边的行列式是范德蒙行列式，可直接利用结论.

解　左式 $=(x-2)(x+1)(x-3)(3-2)(3+1)(-1-2)$，可知此方程的三个根为：$x_1=2$，$x_2=-1$，$x_3=3$.

例 1-13　计算

$$D_4=\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{vmatrix}.$$

分析　此行列式的各列元素有类似的结构，可考虑列变换.

解　把第三列乘(-1)加到第四列上，再把第二列乘(-1)加到第三列上，再把第一列乘(-1)加到第二列上.

$$D_4=\begin{vmatrix} a^2 & 2a+1 & 2a+3 & 2a+5 \\ b^2 & 2b+1 & 2b+3 & 2b+5 \\ c^2 & 2c+1 & 2c+3 & 2c+5 \\ d^2 & 2d+1 & 2d+3 & 2d+5 \end{vmatrix}=\begin{vmatrix} a^2 & 2a+1 & 2 & 2 \\ b^2 & 2b+1 & 2 & 2 \\ c^2 & 2c+1 & 2 & 2 \\ d^2 & 2d+1 & 2 & 2 \end{vmatrix}=0.$$

例 1-14　计算

$$D_4=\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{vmatrix}.$$

其中，a,b,c,d **互不相同.**

解法 1　题给的行列式像范德蒙行列式，但不是范德蒙行列式，可考虑用加边法，即添加一行一列使之凑成范德蒙行列式，再利用它的结果.

$$V_5=\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ a & b & c & d & x \\ a^2 & b^2 & c^2 & d^2 & x^2 \\ a^3 & b^3 & c^3 & d^3 & x^3 \\ a^4 & b^4 & c^4 & d^4 & x^4 \end{vmatrix}$$

$$=(x-a)(x-b)(x-c)(x-d)\cdot[(d-a)(d-b)(d-c)(c-b)(c-a)(b-a)].$$

记方括号中的数为 k，**即**

$$V_5=k\cdot(x-a)(x-b)(x-c)(x-d) \tag{1}$$

另外，V_5 **按第五列展开，是** x **的四次多项式，而所要求的** D_4 **正是** x^3 **的余子式，即有**

$$V_5=1\cdot A_{51}+x\cdot A_{52}+x^2\cdot A_{53}+x^3A_{54}+x^4A_{51} \tag{2}$$

比较式(1) **与式**(2) **中** x^3 **的系数，可得**

$$-D_4=k(-a-b-c-d),$$

即　$D_4=k(a+b+c+d)=(d-a)(d-b)(d-c)(c-b)(c-a)(b-a)(a+b+c+d).$

解法 2　用降阶法，即把第三行乘$(-a^2)$加到第四行上，再把第二行乘$(-a)$加到第三行，再把第一行乘$(-a)$加到第二行上，即有

$$D_4=\begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & b-a & c-a & d-a \\ 0 & b(b-a) & c(c-a) & d(d-a) \\ 0 & b^2(b^2-a^2) & c^2(c^2-a^2) & d^2(d^2-a^2) \end{vmatrix}$$

$$=(b-a)(c-a)(d-a)\begin{vmatrix} 1 & 1 & 1 \\ b & c & d \\ b^2(b+a) & c^2(c+a) & d^2(d+a) \end{vmatrix}$$

$$\xlongequal[-b(b+a)r_2+r_3]{-b\cdot r_1+r_2}(b-a)(c-a)(d-a)\begin{vmatrix}1 & 1 & 1\\ 0 & c-b & d-b\\ 0 & (c-b)c(a+b+c) & (d-b)d(a+b+d)\end{vmatrix}$$

$$=(b-a)(c-a)(d-a)(c-b)(d-b)(d-c)(a+b+c+d).$$

解法 3　线性因式法.把 D_4 看成 a 的四次多项式 $D_4(a)$.由于 $D_4(b)=0,D_4(c)=0,D_4(d)=0$.故可记 $D_4(a)=h(a-b)(a-c)(a-d)(a-\beta)$,其中,$h,\beta$ 是待定常数.又因 b,c,d,β 是 $D_4(a)$ 的根,且 $D_4(a)$ 无 a^3 项,故由韦达定理,有 $b+c+d+\beta=0$.即 $\beta=-(b+c+d)$.而 h 是 a^4 的系数,故

$$h=-\begin{vmatrix}1 & 1 & 1\\ b & c & d\\ b^2 & c^2 & d^2\end{vmatrix}=(b-c)(d-b)(d-c),$$

把 h 及 β 代回 $D_4(a)$ 的表达式中,得

$$D_4=(b-c)(b-d)(c-d)(a-b)(a-c)(a-d)(a+b+c+d).$$

例 1-15　计算

$$D_n=\begin{vmatrix}1 & 2 & 3 & \cdots & n-1 & n\\ 2 & 1 & 2 & \cdots & n-2 & n-1\\ 3 & 2 & 1 & \cdots & n-3 & n-2\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ n-1 & n-2 & n-3 & \cdots & 1 & 2\\ n & n-1 & n-2 & \cdots & 2 & 1\end{vmatrix}$$

分析　此行列式的特点是每相邻两行对应位置的元素都差 1.这样,可利用行列式的性质使它出现很多 1 和 -1.再考虑消零就比较方便了.

解

$$D_n\xlongequal[(-1)r_1+r_2]{\substack{(-1)r_{n-1}+r_n\\(-1)r_{n-2}+r_{n-1}\\ \vdots\\ (-1)r_2+r_3}}\begin{vmatrix}1 & 2 & 3 & \cdots & n-1 & n\\ 1 & -1 & -1 & \cdots & -1 & -1\\ 1 & 1 & -1 & \cdots & -1 & -1\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 1 & 1 & 1 & \cdots & -1 & -1\\ 1 & 1 & 1 & \cdots & 1 & -1\end{vmatrix}$$

$$\xlongequal[c_n+c_{n-1}]{\substack{c_n+c_1\\c_n+c_2\\ \vdots}}\begin{vmatrix}n+1 & n+2 & n+3 & \cdots & 2n-1 & n\\ 0 & -2 & -2 & \cdots & -2 & -1\\ 0 & 0 & -2 & \cdots & -2 & -1\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & -2 & -1\\ 0 & 0 & 0 & \cdots & 0 & -1\end{vmatrix}$$

$$=(-1)(-2)^{n-2}\cdot(n+1).$$

例 1-16　计算

$$D_{n+1}=\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ b_1 & a_1 & a_1 & \cdots & \cdots & a_1 \\ b_1 & b_2 & a_2 & \cdots & \cdots & a_2 \\ b_1 & b_2 & b_3 & a_3 & \cdots & a_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & b_3 & \cdots & \cdots & a_{n-1} \\ b_1 & b_2 & b_3 & \cdots & \cdots & a_n \end{vmatrix}$$

解

$$D_{n+1}\xlongequal{\substack{\text{把第一列的 } b_1 \text{ 全消为 0 后,按第一列展开,并把展开式中第一行的公因子 } (a_1-b_1) \text{ 提出}}}(a_1-b_1)\cdot\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ b_2-b_1 & a_2-b_1 & a_2-b_1 & \cdots & a_2-b_1 & a_2-b_1 \\ b_2-b_1 & b_3-b_1 & a_3-b_1 & \cdots & a_3-b_1 & a_3-b_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b_2-b_1 & b_3-b_1 & b_4-b_1 & \cdots & a_{n-1}-b_1 & a_{n-1}-b_1 \\ b_2-b_1 & b_3-b_1 & b_4-b_1 & \cdots & b_n-b_1 & a_n-b_1 \end{vmatrix}$$

$$\xlongequal{\substack{\text{把第一列的 } b_2-b_1 \text{ 全消为 0 后,按第一列展开,并把展开式中第一行的公因子 } (a_2-b_2) \text{ 提出}}}(a_1-b_1)(a_2-b_2)\cdot\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ b_3-b_2 & a_3-b_2 & a_3-b_2 & \cdots & a_3-b_2 & a_3-b_2 \\ b_3-b_2 & b_4-b_2 & a_4-b_2 & \cdots & a_4-b_2 & a_4-b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b_3-b_2 & b_4-b_2 & b_5-b_2 & \cdots & a_{n-1}-b_2 & a_{n-1}-b_2 \\ b_3-b_2 & b_4-b_2 & b_5-b_2 & \cdots & b_n-b_2 & a_n-b_2 \end{vmatrix}$$

$$\xlongequal{\text{依次类推}}\cdots\cdots=\prod_{i=1}^{n}(a_i-b_i).$$

例 1-17　计算

$$D_n=\begin{vmatrix} 2 & 1 & & & & \\ 1 & 2 & 1 & & & \\ & 1 & 2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & & 2 & 1 \\ & & & & 1 & 2 \end{vmatrix}.$$

分析　此类行列式称为三对角行列式. 一般先经过按第一行展开,找出递推关系.

解

$$\begin{aligned} D_n &= 2\times D_{n-1}-D_{n-2}=2(2D_{n-2}-D_{n-3})-D_{n-2} \\ &= 3\times D_{n-2}-2D_{n-3}=3(2D_{n-3}-D_{n-4})-2D_{n-3} \\ &= 4D_{n-3}-3D_{n-4}=\cdots \\ &= (n-1)D_2-(n-2)D_1=(n-1)\cdot 3-(n-2)\cdot 2 \\ &= n+1. \end{aligned}$$

例 1-18　计算

$$D_n=\begin{vmatrix}\cos\alpha & 1 & & & & & \\ 1 & 2\cos\alpha & 1 & & & & \\ & 1 & 2\cos\alpha & 1 & & & \\ & \ddots & \ddots & \ddots & & & \\ & & & 1 & 2\cos\alpha & 1 & \\ & & & & 1 & 2\cos\alpha & \end{vmatrix}.$$

分析　这也是三对角行列式，但和例 1-17 题有所不同. 不能按第一行展开找逆推关系，但可以试算 D_1，D_2，D_3，看能否找到规律.

解

$$D_1=|\cos\alpha|=\cos\alpha,$$

$$D_2=\begin{vmatrix}\cos\alpha & 1\\ 1 & 2\cos\alpha\end{vmatrix}=2\cos^2\alpha-1=\cos 2\alpha,$$

$$D_3=\begin{vmatrix}\cos\alpha & 1 & 0\\ 1 & 2\cos\alpha & 1\\ 0 & 1 & 2\cos\alpha\end{vmatrix}=4\cos^3\alpha-3\cos\alpha=\cos 3\alpha.$$

试用归纳法证明：$D_n=\cos n\alpha$.

当 $n=1$ 时，结论成立. 归纳假设当 $n\leqslant k-1$ 时，结论成立；当 $n=k$ 时，把 D_k 按最后一行展开（注意：是按最后一行展开，而不是按第一行展开），得

$$\begin{aligned}D_k&=2\cos\alpha\cdot D_{k-1}-D_{k-2}=2\cos\alpha\cdot\cos(k-1)\alpha-\cos(k-2)\alpha\\&=2\cos\alpha[\cos(k-2)\alpha\cdot\cos\alpha-\sin(k-2)\alpha\sin\alpha]-\cos(k-2)\alpha\\&=(2\cos^2\alpha-1)\cos(k-2)\alpha-2\cos\alpha\sin(k-2)\alpha\sin\alpha\\&=\cos 2\alpha\cdot\cos(k-2)\alpha-\sin 2\alpha\cdot\sin(k-2)\alpha=\cos[2\alpha+(k-2)\alpha]=\cos k\alpha.\end{aligned}$$

例 1-19　计算

$$D_{n+1}=\begin{vmatrix}1 & 0 & 0 & \cdots & \cdots & 0 & 1\\ 1 & a_1 & 0 & \cdots & \cdots & 0 & 0\\ 1 & 1 & a_2 & \cdots & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots\\ 1 & 0 & 0 & \cdots & 1 & a_{n-1} & 0\\ 1 & 0 & 0 & \cdots & 0 & 1 & a_n\end{vmatrix}.$$

分析　这种行列式几乎是三条线的行列式. 即第一列一条，对角线一条，还有次对角线一条（次对角线的最后一个元素返到右上角）. 通常也是先按第一行展开.

解

$$D_{n+1}=a_1a_2\cdots a_n+(-1)^{1+(n+1)}\begin{vmatrix}1 & a_1 & 0 & \cdots & 0 & 0\\ 1 & 1 & a_2 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 1 & 0 & 0 & \cdots & 1 & a_{n-1}\\ 1 & 0 & 0 & \cdots & 0 & 1\end{vmatrix},$$

把第 n 行依次换到第一行，共换了 $(n-1)$ 次，即

$$= a_1a_2\cdots a_n + (-1)^{n+2}\cdot(-1)^{n-1}\cdot\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 1 & a_1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & a_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & a_{n-2} & 0 \\ 1 & 0 & 0 & \cdots & 1 & a_{n-1} \end{vmatrix}$$

$$= a_1a_2\cdots a_n - D_n = a_1a_2\cdots a_n - (a_1a_2\cdots a_{n-1} - D_{n-1})$$

$$= a_1a_2\cdots a_n - a_1a_2\cdots a_{n-1} + a_1a_2\cdots a_{n-2} - D_{n-2}$$

$$= \cdots$$

$$= a_1a_2\cdots a_n - a_1a_2\cdots a_{n-1} + a_1a_2\cdots a_{n-2} - \cdots + (-1)^{n-2}a_1a_2 + (-1)^{n-1}a_1 + (-1)^n 1.$$

例 1-20　计算

$$D_n = \begin{vmatrix} x_1 & a & a & \cdots & a \\ a & x_2 & a & \cdots & a \\ a & a & x_3 & \cdots & a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & a & a & \cdots & x_n \end{vmatrix}, \quad (x_i \neq a).$$

分析　加边后，使消零较多.

解

$$D_n = \begin{vmatrix} 1 & a & a & a & \cdots & a \\ 0 & x_1 & a & a & \cdots & a \\ 0 & a & x_2 & a & \cdots & a \\ 0 & a & a & x_3 & \cdots & a \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a & a & a & \cdots & x_n \end{vmatrix} = \begin{vmatrix} 1 & a & a & a & \cdots & a \\ -1 & x_1-a & 0 & 0 & \cdots & 0 \\ -1 & 0 & x_2-a & 0 & \cdots & 0 \\ -1 & 0 & 0 & x_3-a & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & 0 & \cdots & x_n-a \end{vmatrix}$$

把第 i 列$(i = 2,3,\cdots,n)$ 乘$\dfrac{1}{x_i - a}$，加到第一列上，

$$= \begin{vmatrix} 1+\sum\limits_{i=1}^{n}\dfrac{a}{x_i-a} & a & a & a & \cdots & a \\ 0 & x_1-a & 0 & 0 & \cdots & 0 \\ 0 & 0 & x_2-a & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & x_n-a \end{vmatrix} = \left(1+\sum_{i=1}^{n}\frac{a}{x_i-a}\right)\cdot\prod_{i=1}^{n}(x_i-a).$$

例 1-21　已知三阶行列式

$$D_3 = \begin{vmatrix} 2 & x & 1 \\ 1 & 0 & y \\ z & 1 & 2 \end{vmatrix}$$

且 $M_{11} - M_{13} + M_{33} = 0$，$M_{13} + M_{22} + M_{32} = 1$，$M_{31} - M_{21} + M_{13} = 2$. 其中，$M_{ij}$ 是 D_3 中元素 a_{ij} 的余子式，求 D_3 的值.

解　由余子式 M_{ij} 的定义得

$$M_{11}-M_{13}+M_{33}=-y-1-x=0 \tag{1}$$

$$M_{13}+M_{22}+M_{32}=1+(4-z)+(2y-1)=-1 \tag{2}$$

$$M_{31}-M_{21}+M_{13}=xy-(2x-1)+1=2 \tag{3}$$

解此三元方程组，由式(3)得 $x=0$ 或 $y=2$. 把 $x=0$ 代入式(1)与式(2)，得 $y=-1, z=3$. 把 $y=2$ 代入式(1)和式(2)，得 $x=-3, z=9$. 所以，当 $x=0, y=-1, z=3$ 时，

$$D_3=\begin{vmatrix}2&0&1\\1&0&-1\\3&1&2\end{vmatrix}=3.$$

当 $x=-3, y=2, z=9$ 时，

$$D_3=\begin{vmatrix}2&-3&1\\1&0&2\\9&1&2\end{vmatrix}=-51.$$

例 1-22 设

$$|\boldsymbol{A}|=\begin{vmatrix}-1&5&7&4\\2&2&2&2\\3&0&9&-6\\-3&4&1&5\end{vmatrix}.$$

试证：$A_{41}+A_{42}+A_{43}+A_{44}=0$. 其中，$A_{ij}$ 是 a_{ij} 的代数余子式.

证法 1 由行列式按行展开定理：

$$2\cdot A_{41}+2\cdot A_{42}+2\cdot A_{43}+2\cdot A_{44}=0,$$

即

$$A_{41}+A_{42}+A_{43}+A_{44}=0.$$

证法 2 构造

$$|\boldsymbol{B}|=\begin{vmatrix}-1&5&7&4\\2&2&2&2\\3&0&9&-6\\1&1&1&1\end{vmatrix}=1\cdot A_{41}+1\cdot A_{42}+1\cdot A_{43}+1\cdot A_{44},$$

而 $|\boldsymbol{B}|=0$ 是显然的，故有 $A_{41}+A_{42}+A_{43}+A_{44}=0$.

例 1-23 设 $\boldsymbol{A}=(a_{ij})$ 是 n 阶非零矩阵. 若 $|\boldsymbol{A}|$ 的每一元素 a_{ij} 都等于它自己的代数余子式 A_{ij}，试证 $|\boldsymbol{A}|\neq 0$.

证 因 $\boldsymbol{A}$ 非零，故不妨设 $a_{ij}\neq 0$. 则把 $|A|$ 按第 i 行展开，有

$$\begin{aligned}|\boldsymbol{A}|&=a_{i1}A_{i1}+\cdots+a_{ij}A_{ij}+\cdots+a_{in}A_{in}\\&=a_{i1}^2+\cdots+a_{ij}^2+\cdots+a_{in}^2>0.\end{aligned}$$

例 1-24 设 $\boldsymbol{A}=(a_{ij})$ 是 n 阶矩阵. 其中，$a_{1n}=a_{2n}=\cdots=a_{nn}=1$. 将 $\boldsymbol{A}$ 的第 i 行元素换成 $x_1, x_2, \cdots, x_{n-1}, 1$，而其余元素不变. 记这样的矩阵为 $\boldsymbol{A}_i(i=1,2,\cdots,n)$. 试证：$|\boldsymbol{A}|=|\boldsymbol{A}_1|+|\boldsymbol{A}_2|+\cdots+|\boldsymbol{A}_n|$.

证 把 $|\boldsymbol{A}_i|$ 按第 i 行展开：

$$|\boldsymbol{A}_i|=x_1\cdot A_{i1}+x_2\cdot A_{i2}+\cdots+x_{n-1}A_{in-1}+1\cdot A_{in},$$

则

$$\begin{aligned}|\boldsymbol{A}_1|+|\boldsymbol{A}_2|+\cdots+|\boldsymbol{A}_n|&=x_1(A_{11}+A_{21}+\cdots+A_{n1})+x_2(A_{12}+A_{22}+\cdots+A_{n2})+\cdots\\&\quad+x_{n-1}(A_{1n-1}+A_{2n-1}+\cdots+A_{nn-1})+1\cdot(A_{1n}+A_{2n}+\cdots+A_{nn}).\end{aligned}$$

由于 $|A|$ 的第 n 列元素与其余各列的对应元素的代数余子式乘积之和等于零. 即

$$A_{11}+A_{21}+\cdots+A_{n1}=0,$$

$$A_{12}+A_{22}+\cdots+A_{n2}=0,$$

$$\cdots$$

$$A_{1n-1}+A_{2n-1}+\cdots+A_{nn-1}=0.$$

故有

$$|\boldsymbol{A}_1|+|\boldsymbol{A}_2|+\cdots+|\boldsymbol{A}_n|=A_{1n}+A_{2n}+\cdots+A_{nn}=|\boldsymbol{A}|.$$

例 1-25　计算

$$D_n=\begin{vmatrix}1+x_1y_1 & 1+x_1y_2 & \cdots & 1+x_1y_n\\ 1+x_2y_1 & 1+x_2y_2 & \cdots & 1+x_2y_n\\ \vdots & \vdots & \ddots & \vdots\\ 1+x_ny_1 & 1+x_ny_2 & \cdots & 1+x_ny_n\end{vmatrix},\quad (n\geqslant 3).$$

分析　本题的关键在于把 D_n 中的矩阵分解为两个矩阵的乘积,称为积分解法.

解　因

$$\begin{pmatrix}1+x_1y_1 & 1+x_1y_2 & \cdots & 1+x_1y_n\\ 1+x_2y_1 & +x_2y_2 & \cdots & 1+x_2y_n\\ \vdots & \vdots & \ddots & \vdots\\ 1+x_ny_1 & 1+x_ny_2 & \cdots & 1+x_ny_n\end{pmatrix}=\begin{pmatrix}1 & x_1 & 0 & \cdots & 0\\ 1 & x_2 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & x_n & 0 & \cdots & 0\end{pmatrix}\cdot\begin{pmatrix}1 & 1 & \cdots & 1\\ y_1 & y_2 & \cdots & y_n\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix},$$

故有

$$D_n=\begin{vmatrix}1 & x_1 & 0 & \cdots & 0\\ 1 & x_2 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & x_n & 0 & \cdots & 0\end{vmatrix}\cdot\begin{vmatrix}1 & 1 & \cdots & 1\\ y_1 & y_2 & \cdots & y_n\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 0\end{vmatrix}=0\cdot 0=0.$$

例 1-26　计算

$$|\boldsymbol{A}|=\begin{vmatrix}a & b & c & d\\ -b & a & -d & c\\ -c & d & a & -b\\ -d & -c & b & a\end{vmatrix},\quad (a\neq 0).$$

分析　消零是不可取的. 根据 $\boldsymbol{A}$ 中元素排列的特点,可考虑利用性质:$|\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}|=|\boldsymbol{A}|^2$. 若先能把 $|\boldsymbol{A}|^2$ 求出,则求 $|\boldsymbol{A}|$ 就方便了.

解

$$\boldsymbol{A}\cdot\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix}a & b & c & d\\ -b & a & -d & c\\ -c & d & a & -b\\ -d & -c & b & a\end{pmatrix}\cdot\begin{vmatrix}a & -b & -c & -d\\ b & a & d & -c\\ c & -d & a & b\\ d & c & -b & a\end{vmatrix}$$

$$=\begin{pmatrix}a^2+b^2+c^2+d^2 & 0 & 0 & 0\\ 0 & a^2+b^2+c^2+d^2 & 0 & 0\\ 0 & 0 & a^2+b^2+c^2+d^2 & 0\\ 0 & 0 & 0 & a^2+b^2+c^2+d^2\end{pmatrix},$$

故 $$|\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}|=|\boldsymbol{A}|^2=(a^2+b^2+c^2+d^2)^4.$$

则 $|\boldsymbol{A}|$ 的值是 $\pm(a^2+b^2+c^2+d^2)^2$ 中之一. 由题设行列式知,必有 a^4 这一项,因此

$$|\boldsymbol{A}|=(a^2+b^2+c^2+d^2)^2.$$

例 1-27 解方程组:

$$\begin{cases} x+y+z=a+b+c \\ ax+by+cz=a^2+b^2+c^2 \\ bcx+cay+abz=3abc \end{cases}$$

其中,a,b,c 互异.

解

$$D=\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ bc & ca & ab \end{vmatrix}=\begin{vmatrix} 1 & 0 & 0 \\ a & b-a & c-a \\ bc & c(a-b) & b(a-c) \end{vmatrix}$$

$$=b(b-a)(a-c)-c(c-a)(a-b)=(a-c)(b-a)(b-c)\neq 0.$$

故由克莱默法则:

$$x=\frac{D_1}{D},\quad y=\frac{D_2}{D},\quad z=\frac{D_3}{D}.$$

$$D_1=\begin{vmatrix} a+b+c & 1 & 1 \\ a^2+b^2+c^2 & b & c \\ 3abc & ca & ab \end{vmatrix}=\begin{vmatrix} a & 1 & 1 \\ a^2 & b & c \\ abc & ca & ab \end{vmatrix}+\begin{vmatrix} b & 1 & 1 \\ b^2 & b & c \\ abc & ca & ab \end{vmatrix}+\begin{vmatrix} c & 1 & 1 \\ c^2 & b & c \\ abc & ca & ab \end{vmatrix}$$

$$=a\cdot D+0+0=a\cdot D$$

所以,$x=a$. 同理求得 $y=b,z=c$.

例 1-28 设 $\boldsymbol{A}$ 为 m 阶方阵,$\boldsymbol{B}$ 为 n 阶方阵,且 $|\boldsymbol{A}|=a$,$|\boldsymbol{B}|=b$,$\boldsymbol{C}=\begin{pmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix}$. 求 $|\boldsymbol{C}|$.

解 把 $\boldsymbol{C}$ 中的 $\boldsymbol{A}$ 的第一列向左依次换列,换了 n 次后,换到了 $\boldsymbol{C}_1$ 的第一列,再把 $\boldsymbol{A}$ 的第二列向左依次换列,换了 n 次后,换到了 $\boldsymbol{C}_1$ 的第二列,……,最后把 $\boldsymbol{A}$ 的第 m 列向左依次换列,换了 n 次后换到了 $\boldsymbol{C}_1$ 的第 m 列,此时

$$\boldsymbol{C}_1=\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix},$$

由行列式性质知:

$$|\boldsymbol{C}|=(-1)^{m\cdot n}\cdot|\boldsymbol{C}_1|=(-1)^{m\cdot n}\cdot|\boldsymbol{A}|\cdot|\boldsymbol{B}|=(-1)^{m\cdot n}\cdot a\cdot b.$$

例 1-29 $\boldsymbol{A}$ 是 4 阶矩阵,$|(2\boldsymbol{A})^*|=-64$,求 $|\boldsymbol{A}|$.

解法 1 因为 $$|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$$

所以 $$|(2\boldsymbol{A})^*|=|2\boldsymbol{A}|^3=(2^4\cdot|\boldsymbol{A}|)^3=(-4)^3,$$

则 $$2^4\cdot|\boldsymbol{A}|=-4,\quad \text{即}\quad |\boldsymbol{A}|=-\frac{1}{4}.$$

解法 2 因为 $$(k\boldsymbol{A})^*=k^{n-1}\boldsymbol{A}^*,$$

所以 $$|(2\boldsymbol{A})^*|=|2^3\cdot\boldsymbol{A}^*|=(2^3)^4\cdot|\boldsymbol{A}^*|=(64)^2\cdot|\boldsymbol{A}|^3=-64,$$

则 $$|\boldsymbol{A}|^3=-\frac{1}{64},\quad \text{即}\quad |\boldsymbol{A}|=-\frac{1}{4}.$$

例 1-30　已知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 是三维列向量，$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,2\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_4+\boldsymbol{\alpha}_2)$，$\boldsymbol{B}=(\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1)$，$\boldsymbol{C}=(\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2,2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_4+3\boldsymbol{\alpha}_1)$，$|\boldsymbol{B}|=-5$，$|\boldsymbol{C}|=40$，求 $|\boldsymbol{A}|$.

解

$$\begin{aligned}|\boldsymbol{A}| &= |\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,2\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_4+\boldsymbol{\alpha}_2| = |\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,2\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_4| \\ &= |\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,2\boldsymbol{\alpha}_3|-|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4| = -2\cdot|\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1|-|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4| \\ &= 10-|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4|.\end{aligned}$$

又因

$$\begin{aligned}\boldsymbol{C} &= (\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2,2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_4+3\boldsymbol{\alpha}_1) \\ &= (\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\boldsymbol{\alpha}_4)\begin{pmatrix}1&0&3\\2&2&0\\0&3&1\end{pmatrix},\end{aligned}$$

则有

$$|\boldsymbol{C}| = |\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4|\cdot\begin{vmatrix}1&0&3\\2&2&0\\0&3&1\end{vmatrix} = 20\cdot|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4| = 40.$$

知

$$|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4| = 2,$$

故

$$|\boldsymbol{A}| = 10-2 = 8.$$

例 1-31　设 $\boldsymbol{B}$ 为 4 阶方阵，它的第 i 行，第 j 列处的元素为 b_{ij}，且 $b_{ij}=i^j(i=1,2,3,4;j=1,2,3,4)$，试求 $|\boldsymbol{B}|$.

解

$$\begin{aligned}|\boldsymbol{B}| &= \begin{vmatrix}1&1&1&1\\2&4&8&16\\3&9&27&81\\4&16&64&256\end{vmatrix} = \begin{vmatrix}1&2&3&4\\1&4&9&16\\1&8&27&64\\1&16&81&256\end{vmatrix} \\ &= \begin{vmatrix}1&1&1&1&1\\0&1&2&3&4\\0&1&4&9&16\\0&1&8&27&64\\0&1&16&81&256\end{vmatrix} = (4\,!)\cdot(3\,!)\cdot(2\,!)\cdot 1 = 288.\end{aligned}$$

评注　本题解法是范德蒙行列式与加边法的综合运用.

例 1-32　设 4×4 矩阵 $\boldsymbol{A}=(\boldsymbol{\alpha},\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4)$，$\boldsymbol{B}=(\boldsymbol{\beta},\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4)$ 其中，$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4$ 均为 4 维列向量，已知 $|\boldsymbol{A}|=4$，$|\boldsymbol{B}|=1$，求 $|\boldsymbol{A}+\boldsymbol{B}|$.

分析　本题有很强的概念性，既涉及矩阵运算，又涉及行列式运算，因此要求概念很清楚. 此行列式的第 1 列每个元素都可以表示成两个元素和的形式. 此外，行列式的第 2,3,4 列都有公因子 2，可以提到行列式的外面.

解

$$\begin{aligned}|\boldsymbol{A}+\boldsymbol{B}| &= |\alpha+\beta,2\boldsymbol{\gamma}_2,2\boldsymbol{\gamma}_3,2\boldsymbol{\gamma}_4| = 2^3|\alpha+\beta,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4| \\ &= 8(|\alpha,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4|+|\beta,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\gamma}_4|) \\ &= 8(|\boldsymbol{A}|+|\boldsymbol{B}|) = 8(4+1) = 40.\end{aligned}$$

四、小　结

行列式是线性代数的重要工具，它不仅有自身的计算，还涉及计算矩阵的秩、逆阵、解线性方程组等，在

讨论向量组的线性相关性，方阵的特征值和特征向量时也会用到它，因此，对它的基本性质应非常熟悉.

n 阶行列式的计算是难点，我们要求对于一些较为常见的方法要比较熟悉. 例如：(1) 按定义法；(2) 降阶法；(3) 递推法(归纳法)；(4) 加边法；(5) 积分解法；(6) 用按行展开定理法等方法要会应用，对于一些有明显特点的行列式采用相应的方法. 例如，例 1-8 和例 1-9 的特点是每行元素相加之和为常数；例 1-15 的特点是每相邻两行元素均差 1；例 1-25 和例 1-26 的特点是矩阵的乘积与分解. 因此，计算行列式，一开始应先分析它的特点，然后考虑相应的方法.

第二章　矩阵与矩阵运算

线性代数离不开矩阵，矩阵概念及其运算是线性代数中最重要的工具.

一、复习与考试要求

(1) 理解矩阵的概念，了解单位矩阵、数量矩阵、对角矩阵、三角矩阵、对称矩阵和反对称矩阵，以及它们的性质.

(2) 掌握矩阵的线性运算，乘法、转置以及它们的运算规律，了解方阵的幂与方阵乘积的行列式.

(3) 理解逆矩阵的概念，掌握逆矩阵的性质，以及矩阵可逆的充分必要条件，理解伴随矩阵的概念，会用伴随矩阵求逆矩阵.

(4) 理解正交矩阵的概念及其性质.

(5) 掌握矩阵的初等变换，了解初等阵的性质和矩阵等价的概念，理解矩阵的秩的概念，掌握用初等变换求矩阵的秩和逆矩阵的方法.

(6) 了解分块矩阵及其运算，掌握分块矩阵的运算法则.

二、基本概念与理论

1. 矩阵加法

设矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$，$\boldsymbol{B}=(b_{ij})_{m\times n}$，则规定：$\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij})_{m\times n}$.

运算规律：

(1) $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$；

(2) $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$.

2. 数与矩阵相乘

设 $\boldsymbol{A}=(a_{ij})_{m\times n}$，$\lambda$ 是一个数，则规定：$\lambda\cdot\boldsymbol{A}=(\lambda\cdot a_{ij})_{m\times n}$.

运算规律：

(1) $(\lambda\cdot\mu)\cdot\boldsymbol{A}=\lambda\cdot(\mu\cdot\boldsymbol{A})$；

(2) $(\lambda+\mu)\cdot\boldsymbol{A}=\lambda\cdot\boldsymbol{A}+\mu\cdot\boldsymbol{A}$；

(3) $\lambda\cdot(\boldsymbol{A}+\boldsymbol{B})=\lambda\cdot\boldsymbol{A}+\lambda\cdot\boldsymbol{B}$.

(λ，μ 是数).

3. 矩阵与矩阵相乘

设 $\boldsymbol{A}=(a_{ij})_{m\times s}$，$\boldsymbol{B}=(b_{ij})_{s\times n}$，则规定：

$$\boldsymbol{A}\cdot\boldsymbol{B}=(c_{ij})_{m\times n},\text{其中 } c_{ij}=\sum_{k=1}^{s}a_{ik}b_{kj}\quad(i=1,\cdots,m;\ j=1,\cdots,n.)$$

运算规律：

(1) $(\boldsymbol{A}\cdot\boldsymbol{B})\cdot\boldsymbol{C}=\boldsymbol{A}\cdot(\boldsymbol{B}\cdot\boldsymbol{C})$；

(2) $\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{A}\boldsymbol{B}+\boldsymbol{A}\boldsymbol{C}$，$(\boldsymbol{B}+\boldsymbol{C})\boldsymbol{A}=\boldsymbol{B}\boldsymbol{A}+\boldsymbol{C}\boldsymbol{A}$；

(3) $\lambda\cdot(\boldsymbol{A}\cdot\boldsymbol{B})=(\lambda\cdot\boldsymbol{A})\cdot\boldsymbol{B}=\boldsymbol{A}\cdot(\lambda\boldsymbol{B})$，($\lambda$ 是数).

注意　(ⅰ) 一般 $\boldsymbol{A}\boldsymbol{B}\neq\boldsymbol{B}\boldsymbol{A}$. 但 $f(\boldsymbol{A})\cdot g(\boldsymbol{A})=g(\boldsymbol{A})\cdot f(\boldsymbol{A})$. 这里 $f(\boldsymbol{A})$，$g(\boldsymbol{A})$ 皆为方阵 $\boldsymbol{A}$ 的矩阵多项

式.

(ⅱ) 当 $AB = AC$,且 $A \neq O$ 时,$\not\Rightarrow B = C$;

当 $AB = A$,$\not\Rightarrow B = E$;

当 $AB = O$,$\not\Rightarrow A = O$ 或 $B = O$.

4. 矩阵的转置

把矩阵 A 的行换成同序数的列得到一个新的矩阵,称为 A 的转置矩阵,记为 A^{T}. 显见,A 的(i,j) 位置上的元素即为 A^{T} 的(j,i) 位置上的元素.

运算规律:

(1) $(A^{\mathrm{T}})^{\mathrm{T}} = A$;

(2) $(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$;

(3) $(\lambda \cdot A)^{\mathrm{T}} = \lambda \cdot A^{\mathrm{T}}$;

(4) $(A \cdot B)^{\mathrm{T}} = B^{\mathrm{T}} \cdot A^{\mathrm{T}}$.

5. 方阵的行列式

运算规律:

(1) $|A^{\mathrm{T}}| = |A|$;

(2) $|\lambda A| = \lambda^n \cdot |A|$;

(3) $|AB| = |A| \cdot |B| = |BA|$;

(4) $\begin{vmatrix} A & O \\ C & B \end{vmatrix} = |A| \cdot |B|$.

这里 A,B 皆为方阵,但阶数可以不同. 但是,

$$\begin{vmatrix} A & D \\ C & B \end{vmatrix} \neq |A| \cdot |B| - |C| \cdot |D| \quad (A,B,C,D \text{ 皆为方阵}).$$

6. 分块矩阵

矩阵分块是一种技巧,分块后应满足两个条件:一是要使分块后的矩阵运算能够进行;二是要便于所要说明(或证明)的问题,或便于计算.

常见矩阵分块的形式有:

(1) $A_{m\times n} \cdot B_{n\times p} = A \cdot (B_1, \cdots, B_p) = (AB_1, \cdots, AB_p)$. 这里,把 B 按列分块.

(2) $A_{m\times n} \cdot B_{n\times p} = \begin{pmatrix} \alpha_1^{\mathrm{T}} \\ \vdots \\ \alpha_m^{\mathrm{T}} \end{pmatrix} \cdot B = \begin{pmatrix} \alpha_1^{\mathrm{T}} \cdot B \\ \vdots \\ \alpha_m^{\mathrm{T}} \cdot B \end{pmatrix}$

这里,把 A 按行分块.

(3) $A_{m\times n} \cdot B_{n\times p} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} B_1^{\mathrm{T}} \\ \vdots \\ B_n^{\mathrm{T}} \end{pmatrix}$,

这里,把 B 按行分块.

$$A_{m\times n} \cdot B_{n\times p} = (A_1, \cdots, A_n) \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix},$$

这里,把 A 按列分块.

(4) $A_{n\times r} = B_{n\times s} \cdot K_{s\times r}$

$$\Leftrightarrow \quad (\boldsymbol{A}_1,\cdots,\boldsymbol{A}_r)=(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_s)\cdot\begin{pmatrix} k_{11} & \cdots & k_{1r} \\ \vdots & \vdots & \vdots \\ k_{s1} & \cdots & k_{sr} \end{pmatrix}$$

$\Leftrightarrow$ $\boldsymbol{A}$ 的列组可由 $\boldsymbol{B}$ 的列组线性表示.

(5) 线性方程组 $\boldsymbol{A}_{m\times n}\cdot\boldsymbol{X}_{n\times 1}=\boldsymbol{B}_{m\times 1}$

$$\Leftrightarrow \quad (\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)\begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix}=\boldsymbol{B} \quad (\boldsymbol{A}_i \text{ 是 } \boldsymbol{A} \text{ 的第 } i \text{ 列})$$

$$\Leftrightarrow \quad \boldsymbol{x}_1\boldsymbol{A}_1+\cdots+\boldsymbol{x}_n\boldsymbol{A}_n=\boldsymbol{B}$$

(6) 分块对角阵

$\boldsymbol{A}$ 是方阵,且分块后,对角线上均为方的子块 $\boldsymbol{A}_i(i=1,\cdots,s)$,其余全为零,即

$$\boldsymbol{A}=\begin{pmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_s \end{pmatrix}, \quad (\text{其中,} \boldsymbol{A}_i \text{ 的阶数可以不一样}).$$

分块对角阵的性质:

① $|\boldsymbol{A}|=|\boldsymbol{A}_1|\cdot|\boldsymbol{A}_2|\cdot\cdots\cdot|\boldsymbol{A}_s|$;

② $\boldsymbol{A}$ 可逆 $\Leftrightarrow\boldsymbol{A}_i$ 皆可逆 $\quad (i=1,2,\cdots,s)$

$$\text{且} \quad \boldsymbol{A}^{-1}=\begin{pmatrix} \boldsymbol{A}_1^{-1} & & & \\ & \boldsymbol{A}_2^{-1} & & \\ & & \ddots & \\ & & & \boldsymbol{A}_s^{-1} \end{pmatrix}$$

③ $\boldsymbol{A}^k=\begin{pmatrix} \boldsymbol{A}_1^k & & & \\ & \boldsymbol{A}_2^k & & \\ & & \ddots & \\ & & & \boldsymbol{A}_s^k \end{pmatrix} \quad k\in\mathbf{Z}^+$

7. 可逆矩阵

(1) $\boldsymbol{A}$ 是可逆阵的等价说法:

① 存在 n 阶阵 $\boldsymbol{B}$,使 $\boldsymbol{A}\cdot\boldsymbol{B}=\boldsymbol{E}$;

② $|\boldsymbol{A}|\neq 0$;

③ $R(\boldsymbol{A})=n$;

④ $\boldsymbol{A}$ 的 n 个列(行)向量线性无关;

⑤ $\boldsymbol{A}$ 可以分解为一些初等阵的乘积;

⑥ $\boldsymbol{A}$ 与 $\boldsymbol{E}$ 等价;

⑦ $\boldsymbol{A}$ 的所有特征值皆不为零.

(2) 可逆阵 $\boldsymbol{A}$ 的性质:

① 若 $\boldsymbol{A}$ 可逆,则 $\boldsymbol{A}^{-1}$ 也可逆,且$(\boldsymbol{A}^{-1})^{-1}=\boldsymbol{A}$;

② 若 $\boldsymbol{A},\boldsymbol{B}$ 皆可逆,则 $\boldsymbol{A}\cdot\boldsymbol{B}$ 也可逆,且$(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$;

③ 若 $\boldsymbol{A}$ 可逆,则 $\boldsymbol{A}^{\mathrm{T}}$ 也可逆,且$(\boldsymbol{A}^{\mathrm{T}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}}$;

④ 若 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^*$ 也可逆，且 $(\boldsymbol{A}^*)^{-1}=(\boldsymbol{A}^{-1})^*$；

⑤ 若 $\boldsymbol{A}$ 可逆，则 $|\boldsymbol{A}^{-1}|=|\boldsymbol{A}|^{-1}$；

⑥ 若 $\boldsymbol{A}$ 可逆，$\lambda\neq 0$，则 $\lambda\cdot\boldsymbol{A}$ 也可逆，且 $(\lambda\boldsymbol{A})^{-1}=\frac{1}{\lambda}\boldsymbol{A}^{-1}$；

⑦ 可逆阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 相乘后，不改变 $\boldsymbol{B}$ 的秩.

8. 伴随矩阵 $\boldsymbol{A}^*$

$$\boldsymbol{A}^*=\begin{pmatrix}A_{11}&A_{21}&\cdots&A_{n1}\\A_{12}&A_{22}&\cdots&A_{n2}\\\vdots&\vdots&\ddots&\vdots\\A_{1n}&A_{2n}&\cdots&A_{nn}\end{pmatrix}$$

其中，A_{ij} 是 $\boldsymbol{A}$ 中元素 a_{ij} 的代数余子式.

伴随矩阵 $\boldsymbol{A}^*$ 的性质：

① $\boldsymbol{A}\cdot\boldsymbol{A}^*=\boldsymbol{A}^*\cdot\boldsymbol{A}=|\boldsymbol{A}|\cdot\boldsymbol{E}$；

② 若 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\cdot\boldsymbol{A}^*$；

③ $R(\boldsymbol{A}^*)=\begin{cases}n, & \text{若 } R(\boldsymbol{A})=n;\\1, & \text{若 } R(\boldsymbol{A})=n-1;\\0, & \text{若 } R(\boldsymbol{A})<n-1.\end{cases}$

④ $|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$.

9. 初等阵

(1) **定义**　由单位阵经过一次初等变换后得到的矩阵，称为初等阵，即

$$\boldsymbol{P}(i,j)=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&0&&1&&i\\&&&\ddots&&&\\&&1&&0&&j\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}$$

$$\boldsymbol{P}(i(c))\underset{(c\neq 0)}{=}\begin{pmatrix}1&&&&\\&\ddots&&&\\&&c&&i\\&&&\ddots&\\&&&&1\end{pmatrix}$$

$$\boldsymbol{P}(i,j(k))=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&k&&i\\&&&\ddots&&&\\&&&&1&&j\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}$$

(2) 初等阵皆为可逆阵，且其逆阵还是它本身类型的初等阵，即有

$$\boldsymbol{P}(i,j)^{-1}=\boldsymbol{P}(i,j);$$

$$\boldsymbol{P}(i(c))^{-1}=\boldsymbol{P}\left(i\left(\frac{1}{c}\right)\right);$$

$$\boldsymbol{P}(i,j(k))^{-1}=\boldsymbol{P}(i,j(-k)).$$

(3) **定理**　对 $\boldsymbol{A}_{s\times n}$ 施行一次初等行(列)变换,相当于在 $\boldsymbol{A}$ 的左(右)边乘上一个相应的 $s(n)$ 阶初等阵.

(4) **矩阵 $\boldsymbol{A}_{m\times n}$ 与 $\boldsymbol{B}_{m\times n}$ 等价**

$\Leftrightarrow$　$\boldsymbol{P}_m\cdot\boldsymbol{A}\cdot\boldsymbol{Q}_n=\boldsymbol{B}$. 这里 $\boldsymbol{P},\boldsymbol{Q}$ 皆为可逆阵.

$\Leftrightarrow$　$R(\boldsymbol{A})=R(\boldsymbol{B})$.

$\Leftrightarrow$　$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的标准型.

(5) **求逆阵常用方法**:$(\boldsymbol{A},\boldsymbol{E})\xrightarrow{\text{行}}\cdots\to(\boldsymbol{E},\boldsymbol{A}^{-1})$.

(6) 注意"左行右列"规则,即用初等阵左乘 $\boldsymbol{A}$,就对 $\boldsymbol{A}$ 的行起作用,右乘 $\boldsymbol{A}$,就对 $\boldsymbol{A}$ 的列起作用,这里还要注意 $\boldsymbol{P}(i,j(k))$. 它左乘 $\boldsymbol{A}$,相当于把 $\boldsymbol{A}$ 的第 j 行的 k 倍加到第 i 行;而它右乘 $\boldsymbol{A}$,相当于把 $\boldsymbol{A}$ 的第 i 列的 k 倍加到第 j 列,即 $\boldsymbol{P}(i,j(k))$ 这个初等阵左乘,右乘时不仅是行换成列,还有指标 i,j 亦要变换.

10. 对称矩阵

(1) **定义**　$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$.

(2) **实对称阵的性质**:

① 实对称阵 $\boldsymbol{A}$ 的特征值全是实数;

② 实对称阵 $\boldsymbol{A}$ 的不同特征值对应的特征向量是正交的.

(3) **定理**　对于实对称阵 $\boldsymbol{A}$,一定存在正交阵 $\boldsymbol{Q}$,使 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{D}$,这里 $\boldsymbol{D}$ 是对角阵.

11. 反对称矩阵

(1) **定义**　$\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$.

(2) **性质**:

① 奇数阶的反对称阵一定不可逆;

② 对于反对称阵 $\boldsymbol{A}$,任给非零列向量 $\boldsymbol{X}$,恒有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=0$;

③ 实反对称阵 $\boldsymbol{A}$ 的特征值只能是零或纯虚数.

12. 正交矩阵

(1) **定义**　实方阵 $\boldsymbol{A}$,满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$,则称 $\boldsymbol{A}$ 为正交阵.

$\Leftrightarrow$　实方阵 $\boldsymbol{A}$,满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{-1}$.

$\Leftrightarrow$　实方阵 $\boldsymbol{A}$ 的列(行)向量组是两两正交的单位向量组.

(2) **性质**:

① $\boldsymbol{E}$ 是正交阵;

② 若 $\boldsymbol{A},\boldsymbol{B}$ 皆为正交阵,则 $\boldsymbol{A}\cdot\boldsymbol{B}$ 也是正交阵;

③ 若 $\boldsymbol{A}$ 是正交阵,则 $\boldsymbol{A}^{-1}$ 也是正交阵;

④ 若 $\boldsymbol{A}$ 是正交阵,则 $\boldsymbol{A}^{*}$ 也是正交阵;

⑤ 若 $\boldsymbol{A}$ 是正交阵,则 $|\boldsymbol{A}|=1$ 或 -1;

⑥ 若 $\boldsymbol{A}$ 是正交阵,且 $|\boldsymbol{A}|=1$,则 $\boldsymbol{A}_{ij}=a_{ij}$;

⑦ 若 $\boldsymbol{A}$ 是正交阵,则 $\boldsymbol{A}$ 的特征值的模必为 1;

⑧ 正交变换是保模变换,即若有 $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Y}$,这里 $\boldsymbol{Q}$ 是正交阵,$\boldsymbol{X},\boldsymbol{Y}$ 是 n 维列向量,则 $\|\boldsymbol{X}\|=\|\boldsymbol{Y}\|$.

13. 正定矩阵(在第六章详述)

14. 向量内积

(1) **定义**　$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\cdot\boldsymbol{\beta}$,　　$\boldsymbol{\alpha},\boldsymbol{\beta}$ 皆为 n 维实列向量.

(2) **性质**:

① $(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})$;

② $(k\boldsymbol{\alpha},\boldsymbol{\beta})=k\cdot(\boldsymbol{\alpha},\boldsymbol{\beta})$，　$k\in\mathbf{R}$；

③ $(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$；

④ $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$，当且仅当 $\boldsymbol{\alpha}=0$ 时，$(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$；

(3) 向量 $\boldsymbol{\alpha}$ 的长度(也称 $\boldsymbol{\alpha}$ 的模) 的定义：$\|\boldsymbol{\alpha}\|=\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}$. 它有如下性质：

① $\|\boldsymbol{\alpha}\|\geqslant 0$，　$\|\boldsymbol{\alpha}\|=0\Leftrightarrow\boldsymbol{\alpha}=0$；

② $\|k\cdot\boldsymbol{\alpha}\|=|k|\cdot\|\boldsymbol{\alpha}\|$；

③ $|(\boldsymbol{\alpha},\boldsymbol{\beta})|\leqslant\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|$；

④ $\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|\leqslant\|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\|$.

三、基本题型与解题方法

例 2-1　已知矩阵

$$\boldsymbol{A}=\begin{bmatrix}1/2 & 0 & 0\\ 0 & 1 & -3/4\\ 0 & -2/3 & 1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}0\\ 1\\ -1\end{bmatrix}$$

满足　$\boldsymbol{A}^{-1}(\boldsymbol{E}-\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{-1})^{-1}\boldsymbol{C}^{-1}=\boldsymbol{E}$，求 $\boldsymbol{C}$.

分析　利用矩阵的运算规则，先将所给的关系式化简成最简形式，再将矩阵代入，即可求出结果.

解　关系式两边右乘 $\boldsymbol{C}$，再将左边两个逆阵合并得

$$\left[(\boldsymbol{E}-\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{-1})\boldsymbol{A}\right]^{-1}=\boldsymbol{C},$$

从而得

$$\boldsymbol{C}=(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}})^{-1}=\begin{bmatrix}1/2 & 0 & 0\\ 0 & 0 & 1/4\\ 0 & 1/3 & 0\end{bmatrix}$$

$$=\begin{bmatrix}\left(\frac{1}{2}\right)^{-1} & & \\ & & \begin{pmatrix}0 & 1/4\\ 1/3 & 0\end{pmatrix}^{-1}\end{bmatrix}=\begin{bmatrix}2 & 0 & 0\\ 0 & 0 & 3\\ 0 & 4 & 0\end{bmatrix}.$$

例 2-2　已知三阶方阵 $\boldsymbol{A}$ 满足 $|\boldsymbol{A}|=\frac{1}{3}$，求 $|2\boldsymbol{A}^{*}-\boldsymbol{A}^{-1}|$.

分析　由公式 $\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}$，即有 $\boldsymbol{A}^{*}=|\boldsymbol{A}|\boldsymbol{A}^{-1}$，代入要求的行列式即可求出.

解　$|2\boldsymbol{A}^{*}-\boldsymbol{A}^{-1}|=\left|2\times\frac{1}{3}\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\right|=\left(-\frac{1}{3}\right)^{3}\frac{1}{|\boldsymbol{A}|}=-\frac{1}{9}.$

例 2-3　设

$$\boldsymbol{A}=\begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0\end{bmatrix},$$

(1) 求 $\boldsymbol{A}^{2},\boldsymbol{A}^{3},\boldsymbol{A}^{99}$；

(2) 设 $BP=PA$，其中，A 如上，而 $P=\begin{bmatrix}1&a&0&0&0\\0&1&0&0&0\\0&0&1&b&c\\0&0&0&1&b\\0&0&0&0&1\end{bmatrix}$，$abc\neq 0$，求 B^{10}.

分析　已知的 A 矩阵中有大量的零元素，因此进行分块可以简化运算.

解

(1) 记 $$A=\begin{bmatrix}A_1&0\\0&A_2\end{bmatrix},$$

其中，$A_1=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$，$A_2=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}$，

则 $$A^2=\begin{bmatrix}A_1^2&\\&A_2^2\end{bmatrix}=\begin{bmatrix}E&0\\0&A_2^2\end{bmatrix}=\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&0&0&1\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}.$$

$$A^3=\begin{bmatrix}A_1^3&0\\0&A_2^3\end{bmatrix}=\begin{bmatrix}A_1&0\\0&0\end{bmatrix}$$

$$=\begin{bmatrix}1&0&0&0&0\\0&-1&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}=\begin{bmatrix}1&&&\\&-1&&\\&&\mathbf{0}&\end{bmatrix},$$

$$A^{99}=(A^3)^{33}=\begin{bmatrix}1&&&\\&-1&&\\&&\mathbf{0}&\end{bmatrix}^{33}=\begin{bmatrix}1&&&\\&-1&&\\&&\mathbf{0}&\end{bmatrix}=A^3;$$

(2) 由已知得 $B=PAP^{-1}$，$B^{10}=PA^{10}P^{-1}$.

记 $P=\begin{bmatrix}P_1&0\\0&P_2\end{bmatrix}$，其中 $P_1=\begin{bmatrix}1&a\\0&1\end{bmatrix}$，$P_2=\begin{bmatrix}1&b&c\\0&1&b\\0&0&1\end{bmatrix}$.

则 $P^{-1}=\begin{bmatrix}P_1^{-1}&\\&P_2^{-1}\end{bmatrix}$，又 $A^{10}=\begin{bmatrix}A_1^{10}&0\\0&A_2^{10}\end{bmatrix}=\begin{bmatrix}E&0\\0&0\end{bmatrix}$，

所以 $$B^{10}=PA^{10}P^{-1}=\begin{bmatrix}P_1&\\&P_2\end{bmatrix}\begin{bmatrix}E&0\\0&0\end{bmatrix}\begin{bmatrix}P_1^{-1}&0\\0&P_2^{-1}\end{bmatrix}$$

$$=\begin{bmatrix}P_1P_1^{-1}&\\&0\end{bmatrix}=\begin{bmatrix}E&\\&0\end{bmatrix}.$$

例 2-4　设

$$\boldsymbol{A}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix},\quad \boldsymbol{C}=\begin{bmatrix}1&-4&3\\2&0&-1\\1&-2&0\end{bmatrix}$$

满足方程 $\boldsymbol{AXB}=\boldsymbol{C}$，求 $\boldsymbol{X}$.

分析　这是一类解矩阵方阵的问题，只要正确利用 $\boldsymbol{A}^{-1}$，$\boldsymbol{B}^{-1}$，及 $\boldsymbol{C}$ 的运算表达式来表示 $\boldsymbol{X}$，并熟练进行矩阵运算即可解出 $\boldsymbol{X}$.

解　易见 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是初等方阵，且 $\boldsymbol{A}^{-1}=\boldsymbol{A}$，$\boldsymbol{B}^{-1}=\boldsymbol{B}$，$\boldsymbol{A}$ 左乘 $\boldsymbol{C}$ 相当于对 $\boldsymbol{C}$ 进行行初等变换，$\boldsymbol{B}$ 右乘 $\boldsymbol{C}$ 相当于对 $\boldsymbol{C}$ 进行列初等变换，因此得

$$\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{C}\boldsymbol{B}^{-1}=\boldsymbol{ACB}\xlongequal[\text{再交换第 }2,3\text{ 列}]{\text{先交换 }\boldsymbol{C}\text{ 的第 }1,2\text{ 行}}\begin{bmatrix}2&-1&0\\1&3&-4\\1&0&-2\end{bmatrix}.$$

例 2-5　设实矩阵 $\boldsymbol{A}=(a_{ij})_{3\times4}$，且 $\boldsymbol{AA}^{\mathrm{T}}=\boldsymbol{0}$，求证 $\boldsymbol{A}=\boldsymbol{0}$.

分析　将 $\boldsymbol{AA}^{\mathrm{T}}$ 计算出，利用矩阵相等的充要条件得出 $\boldsymbol{AA}^{\mathrm{T}}$ 的每个元素为 0，从而得出 $\boldsymbol{A}=\boldsymbol{0}$.

证

$$\boldsymbol{AA}^{\mathrm{T}}=\begin{bmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\end{bmatrix}\begin{bmatrix}a_{11}&a_{21}&a_{31}\\a_{12}&a_{22}&a_{32}\\a_{13}&a_{23}&a_{33}\\a_{14}&a_{24}&a_{34}\end{bmatrix}$$

$$=\begin{bmatrix}\sum\limits_{i=1}^{4}a_{1i}^2&*&**&\sum\limits_{i=1}^{4}a_{2i}^2&**&*&\sum\limits_{i=1}^{4}a_{3i}^2\end{bmatrix}=\boldsymbol{0},$$

所以　$$\sum_{j=1}^{4}a_{1j}^2=\sum_{j=1}^{4}a_{2j}^2=\sum_{j=1}^{4}a_{3j}^2=0,$$

得到　$a_{ij}=0$，$i=1,2,3$，$j=1,2,3,4$，即 $\boldsymbol{A}=\boldsymbol{0}$.

本题的证明过程也可以采用下述方式表达：

将 $\boldsymbol{A}$ 按行分块，记为 $\boldsymbol{A}=\begin{bmatrix}\alpha_1^{\mathrm{T}}\\\alpha_2^{\mathrm{T}}\\\alpha_3^{\mathrm{T}}\end{bmatrix}$，因

$$\boldsymbol{AA}^{\mathrm{T}}=\begin{pmatrix}\alpha_1^{\mathrm{T}}\\\alpha_2^{\mathrm{T}}\\\alpha_3^{\mathrm{T}}\end{pmatrix}(\alpha_1,\alpha_2,\alpha_3)=\begin{pmatrix}\alpha_1^{\mathrm{T}}\alpha_1&*&**&\alpha_2^{\mathrm{T}}\alpha_2&**&*&\alpha_3^{\mathrm{T}}\alpha_3\end{pmatrix}=\boldsymbol{0},$$

所以　$$\alpha_i^{\mathrm{T}}\alpha_i=0,\ (i=1,2,3),\ \text{即}\ \sum_{j=1}^{4}a_{ij}^2=0,$$

因此　$$a_{ij}=0,\ (i=1,2,3;\ j=1,2,3,4),\ \text{即}\ \boldsymbol{A}=0.$$

评注　(1) 上面计算 $\boldsymbol{AA}^{\mathrm{T}}$ 的表达式中，打 $*$ 者未写出具体表达式，因为它对本题计算没有用，故略写成 $*$.

(2) 本题可以推广到 $m\times n$ 矩阵 $\boldsymbol{A}$，如果 $\boldsymbol{AA}^{\mathrm{T}}=0$ 或 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{0}$，结论同样成立.

例 2-6　选择题：设 n 阶方阵 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 满足关系 $\boldsymbol{ABC}=\boldsymbol{E}$，其中 $\boldsymbol{E}$ 是 n 阶单位阵，则必有　(　)

(A) $ACB=E$；　　(B) $CBA=E$；

(C) $BAC=E$；　　(D) $BCA=E$.

分析　这是 1991 年数学(试卷二)的 3 分题,目的是考查考生对矩阵乘法运算的掌握程度.

解法 1　$ABC=E$ 知,BC 及 A 都是可逆的,且 A 与 BC 互为逆矩阵,对照四个选项即可知选(D).

解法 2　考察四个选项中矩阵 A,B,C 的排列,知前三个选项为奇排列,只有选项(D)为偶排列,猜想应选(D).

例 2-7　设 A^* 是 A 的伴随阵,

$$A=\begin{pmatrix}1&0&0\\2&2&0\\3&4&5\end{pmatrix},\quad 求(A^*)^{-1}.$$

解　$$A\cdot A^*=|A|\cdot E,$$

由 A 可逆,　$\Rightarrow$　$A^*=|A|A^{-1}$.

$$\Rightarrow\quad (A^*)^{-1}=(|A|\cdot A^{-1})^{-1}=\frac{1}{|A|}A=\frac{1}{10}\begin{pmatrix}1&0&0\\2&2&0\\3&4&5\end{pmatrix}.$$

例 2-8　设矩阵 A,B 满足 $A^*BA=2BA-8E$,其中 $A=\begin{pmatrix}1&0&0\\0&-2&0\\0&0&1\end{pmatrix}$,$A^*$ 为伴随阵,E 为单位阵,求 B.

解　由　$$A^*BA=2BA-8E.$$

$$\Rightarrow\quad (A^*-2E)BA=-8E.$$

A 是可逆阵,$|A|=-2$,

$$A^{-1}=\begin{pmatrix}1&0&0\\0&-\frac{1}{2}&0\\0&0&1\end{pmatrix},\quad A^*=|A|\cdot A^{-1}=\begin{pmatrix}-2&0&0\\0&1&0\\0&0&-2\end{pmatrix}.$$

$$\Rightarrow\quad A^*-2E=\begin{pmatrix}-4&0&0\\0&-1&0\\0&0&-4\end{pmatrix},\quad (A^*-2E)^{-1}=\begin{pmatrix}-\frac{1}{4}&0&0\\0&-1&0\\0&0&-\frac{1}{4}\end{pmatrix}.$$

则　$$B=-8(A^*-2E)^{-1}\cdot A^{-1}=-8\cdot\begin{pmatrix}-\frac{1}{4}&0&0\\0&-1&0\\0&0&-\frac{1}{4}\end{pmatrix}\cdot\begin{pmatrix}1&0&0\\0&-\frac{1}{2}&0\\0&0&1\end{pmatrix}=\begin{pmatrix}2&0&0\\0&-4&0\\0&0&2\end{pmatrix}.$$

例 2-9　A,B 为三阶方阵,满足 $2A^{-1}B=B-4E$,

(1) 证明　$A-2E$ 可逆;

(2) 若 $B=\begin{pmatrix}1&-2&0\\1&2&0\\0&0&2\end{pmatrix}$,求 A.

解　(1) 由　$$2A^{-1}B=B-4E;$$

$$\Rightarrow\quad AB-2B-4A=0.$$

$$\Rightarrow \quad (\boldsymbol{A}-2\boldsymbol{E})(\boldsymbol{B}-4\boldsymbol{E})=8\boldsymbol{E}$$

$$\Rightarrow \quad (\boldsymbol{A}-2\boldsymbol{E})^{-1}=\frac{1}{8}(\boldsymbol{B}-4\boldsymbol{E}),$$

(2) $\boldsymbol{A}=2\boldsymbol{E}+8(\boldsymbol{B}-4\boldsymbol{E})^{-1}=\begin{pmatrix}0 & 2 & 0\\ -1 & -1 & 0\\ 0 & 0 & -2\end{pmatrix}$.

例 2-10 设 $\boldsymbol{A}$ 的伴随阵 $\boldsymbol{A}^*$.

$$\boldsymbol{A}^*=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 0 & -3 & 0 & 8\end{pmatrix},$$

且 $\boldsymbol{ABA}^{-1}=\boldsymbol{BA}^{-1}+3\boldsymbol{E}$,求 $\boldsymbol{B}$.

解 由 $|\boldsymbol{A}^*|=8$, $\Rightarrow$ $|\boldsymbol{A}|^{n-1}=|\boldsymbol{A}|^3=8$, $\Rightarrow$ $|\boldsymbol{A}|=2$.

又由

$$\boldsymbol{ABA}^{-1}=\boldsymbol{BA}^{-1}+3\boldsymbol{E}\Rightarrow(\boldsymbol{A}-\boldsymbol{E})\boldsymbol{BA}^{-1}=3\boldsymbol{E}$$

$$\Rightarrow \quad (\boldsymbol{A}-\boldsymbol{E})\boldsymbol{B}=3\boldsymbol{A}\Rightarrow\boldsymbol{A}^{-1}(\boldsymbol{A}-\boldsymbol{E})\boldsymbol{B}=3\boldsymbol{E}$$

$$\Rightarrow \quad (\boldsymbol{E}-\boldsymbol{A}^{-1})\boldsymbol{B}=3\boldsymbol{E}\Rightarrow\left(\boldsymbol{E}-\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*\right)\boldsymbol{B}=3\boldsymbol{E}$$

$$\Rightarrow \quad (2\boldsymbol{E}-\boldsymbol{A}^*)\boldsymbol{B}=6\boldsymbol{E}.\Rightarrow\boldsymbol{B}=6\cdot(2\boldsymbol{E}-\boldsymbol{A}^*)^{-1}$$

$$=\begin{pmatrix}6 & 0 & 0 & 0\\ 0 & 6 & 0 & 0\\ 6 & 0 & 6 & 0\\ 0 & 3 & 0 & -1\end{pmatrix}.$$

例 2-11 $\boldsymbol{A},\boldsymbol{B}$ 是 n 阶方阵,若 $\boldsymbol{E}+\boldsymbol{AB}$ 与 $\boldsymbol{E}+\boldsymbol{BA}$ 皆可逆,试证:$(\boldsymbol{E}+\boldsymbol{AB})^{-1}=\boldsymbol{E}-\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{BA})^{-1}\boldsymbol{B}$.

证 在等式两边左乘 $\boldsymbol{E}+\boldsymbol{AB}$,则

$$\text{右边}=(\boldsymbol{E}+\boldsymbol{AB})-(\boldsymbol{E}+\boldsymbol{AB})\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{BA})^{-1}\cdot\boldsymbol{B}$$

$$=\boldsymbol{E}+\boldsymbol{AB}-(\boldsymbol{A}+\boldsymbol{ABA})(\boldsymbol{E}+\boldsymbol{BA})^{-1}\cdot\boldsymbol{B}$$

$$=\boldsymbol{E}+\boldsymbol{AB}-\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{BA})(\boldsymbol{E}+\boldsymbol{BA})^{-1}\cdot\boldsymbol{B}$$

$$=\boldsymbol{E}+\boldsymbol{AB}-\boldsymbol{AB}=\boldsymbol{E}=\text{左边}.$$

评注 注意运算的第二步,第三步.

例 2-12 设 $\boldsymbol{A},\boldsymbol{B}$ 皆为对称阵,且 $\boldsymbol{AB}+\boldsymbol{E}$ 及 $\boldsymbol{A}$ 均可逆,试证:$(\boldsymbol{AB}+\boldsymbol{E})^{-1}\boldsymbol{A}$ 对称.

证

$$((\boldsymbol{AB}+\boldsymbol{E})^{-1}\boldsymbol{A})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{AB}+\boldsymbol{E})^{-1^{\mathrm{T}}}=\boldsymbol{A}(\boldsymbol{AB}+\boldsymbol{E})^{\mathrm{T}^{-1}}$$

$$=\boldsymbol{A}(\boldsymbol{BA}+\boldsymbol{E})^{-1}=\boldsymbol{A}(\boldsymbol{A}^{-1}\boldsymbol{ABA}+\boldsymbol{A}^{-1}\boldsymbol{A})^{-1}=\boldsymbol{A}(\boldsymbol{A}^{-1}(\boldsymbol{AB}+\boldsymbol{E})\boldsymbol{A})^{-1}$$

$$=\boldsymbol{A}\cdot\boldsymbol{A}^{-1}\cdot(\boldsymbol{AB}+\boldsymbol{E})^{-1}\cdot\boldsymbol{A}^{-1^{-1}}=(\boldsymbol{AB}+\boldsymbol{E})^{-1}\cdot\boldsymbol{A}.$$

评注 注意运算过程中的第四,第五步.

例 2-13 若 $\boldsymbol{A}^2=\boldsymbol{A}$,则 $\boldsymbol{A}+\boldsymbol{E}$ 可逆.

证法 1 设

$$(\boldsymbol{A}+\boldsymbol{E})(\lambda\boldsymbol{A}+\boldsymbol{E})=\boldsymbol{E}.$$

即

$$\lambda\boldsymbol{A}^2+\lambda\boldsymbol{A}+\boldsymbol{A}+\boldsymbol{E}=\boldsymbol{E}, \quad \Rightarrow \quad (2\lambda+1)\boldsymbol{A}=0, \quad \Rightarrow \quad \lambda=-\frac{1}{2}.$$

所以，$$(\boldsymbol{A}+\boldsymbol{E})^{-1}=-\frac{1}{2}\boldsymbol{A}+\boldsymbol{E}.$$

证法 2　由 $\boldsymbol{A}^2=\boldsymbol{A}$，设 $\boldsymbol{AX}=\lambda\boldsymbol{X}$，则 $\boldsymbol{A}^2\boldsymbol{X}=\lambda^2\boldsymbol{X}\Rightarrow(\lambda-\lambda^2)\boldsymbol{X}=0\Rightarrow\lambda=0$ 或 1.

$\Rightarrow$ $|\boldsymbol{E}+\boldsymbol{A}|\neq 0$（否则，$-1$ 是 $\boldsymbol{A}$ 的特征值，与证得结论矛盾），即 $\boldsymbol{A}+\boldsymbol{E}$ 可逆.

例 2-14　若 $\boldsymbol{A}$ 是实反对称阵，则 $\boldsymbol{E}\pm\boldsymbol{A}$ 均可逆，且 $(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}$ 是正交阵.

证　$\boldsymbol{A}$ 是实反对称阵，则 $\boldsymbol{E}\pm\boldsymbol{A}$ 均可逆，这在第二章线性方程组部分例题 2-13 已证. 下证是正交阵：

$$\begin{aligned}((\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1})^{\mathrm{T}}(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}&=(\boldsymbol{E}+\boldsymbol{A})^{-1^{\mathrm{T}}}\cdot(\boldsymbol{E}-\boldsymbol{A})^{\mathrm{T}}\cdot(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}\\&=(\boldsymbol{E}-\boldsymbol{A})^{-1}(\boldsymbol{E}+\boldsymbol{A})(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}=(\boldsymbol{E}-\boldsymbol{A})^{-1}(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}\\&=\boldsymbol{E}\cdot\boldsymbol{E}=\boldsymbol{E}.\end{aligned}$$

评注　本题证明中关键的一步是

$$(\boldsymbol{E}+\boldsymbol{A})(\boldsymbol{E}-\boldsymbol{A})=(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A}).$$

例 2-15　设 $\boldsymbol{X}=\begin{pmatrix}\boldsymbol{O}&\boldsymbol{A}\\\boldsymbol{B}&\boldsymbol{C}\end{pmatrix}$，其中，$\boldsymbol{A},\boldsymbol{B}$ 都是可逆矩阵，求 $\boldsymbol{X}^{-1}$.

解　设 $\boldsymbol{X}^{-1}=\begin{pmatrix}\boldsymbol{X}_1&\boldsymbol{X}_2\\\boldsymbol{X}_3&\boldsymbol{X}_4\end{pmatrix}$，则

$$\boldsymbol{X}\cdot\boldsymbol{X}^{-1}=\begin{pmatrix}\boldsymbol{O}&\boldsymbol{A}\\\boldsymbol{B}&\boldsymbol{C}\end{pmatrix}\cdot\begin{pmatrix}\boldsymbol{X}_1&\boldsymbol{X}_2\\\boldsymbol{X}_3&\boldsymbol{X}_4\end{pmatrix}=\begin{pmatrix}\boldsymbol{AX}_3&\boldsymbol{AX}_4\\\boldsymbol{BX}_1+\boldsymbol{CX}_3&\boldsymbol{BX}_2+\boldsymbol{CX}_4\end{pmatrix}=\begin{pmatrix}\boldsymbol{E}_k&\\&\boldsymbol{E}_{n-k}\end{pmatrix},$$

即 $$\boldsymbol{AX}_3=\boldsymbol{E}_k,\quad\boldsymbol{BX}_1+\boldsymbol{CX}_3=\boldsymbol{O},$$

$$\boldsymbol{AX}_4=\boldsymbol{O},\quad\boldsymbol{BX}_2+\boldsymbol{CX}_4=\boldsymbol{E}_{n-k}.$$

则 $$\boldsymbol{X}_3=\boldsymbol{A}^{-1},\quad\boldsymbol{X}_4=\boldsymbol{O},\quad\boldsymbol{X}_2=\boldsymbol{B}^{-1},\quad\boldsymbol{X}_1=-\boldsymbol{B}^{-1}\boldsymbol{CA}^{-1}.$$

所以 $$\boldsymbol{X}^{-1}=\begin{pmatrix}-\boldsymbol{B}^{-1}\boldsymbol{CA}^{-1}&\boldsymbol{B}^{-1}\\\boldsymbol{A}^{-1}&\boldsymbol{O}\end{pmatrix}.$$

评注　注意题中 $\boldsymbol{A},\boldsymbol{B}$ 可以是不同阶的方阵，而且在分块矩阵中，$\boldsymbol{X}^{-1}$ 中的 $\boldsymbol{X}_3$（即 $\boldsymbol{A}^{-1}$）的形状与 $\boldsymbol{X}$ 中的 $\boldsymbol{B}$ 是不一样的.

例 2-16　$\boldsymbol{A}$ 是 n 阶方阵，$\boldsymbol{A}+\boldsymbol{E}$ 可逆，且 $f(\boldsymbol{A})=(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}$. 证明：

(1) $(\boldsymbol{E}+f(\boldsymbol{A}))(\boldsymbol{E}+\boldsymbol{A})=2\boldsymbol{E}$；

(2) $f(f(\boldsymbol{A}))=\boldsymbol{A}$.

证　(1) 由 $$f(\boldsymbol{A})=(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}\Rightarrow f(\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})=\boldsymbol{E}-\boldsymbol{A}$$

$$\Rightarrow\quad\boldsymbol{E}+\boldsymbol{A}+f(\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})=2\boldsymbol{E}\Rightarrow(\boldsymbol{E}+f(\boldsymbol{A}))(\boldsymbol{E}+\boldsymbol{A})=2\boldsymbol{E}.$$

(2) $$\begin{aligned}f(f(\boldsymbol{A}))&=f((\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1})\\&=(\boldsymbol{E}-(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1})\cdot(\boldsymbol{E}+(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1})^{-1}\\&=((\boldsymbol{E}+\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}-(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1})\cdot((\boldsymbol{E}+\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}+(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1})^{-1}\\&=(2\boldsymbol{A}\cdot(\boldsymbol{E}+\boldsymbol{A})^{-1})\cdot(2\boldsymbol{E}(\boldsymbol{E}+\boldsymbol{A})^{-1})^{-1}=2\boldsymbol{A}\cdot(\boldsymbol{E}+\boldsymbol{A})^{-1}\cdot\frac{1}{2}\cdot(\boldsymbol{E}+\boldsymbol{A})=\boldsymbol{A}.\end{aligned}$$

评注　在矩阵运算中，$(\boldsymbol{A}+\boldsymbol{B})^{-1}\neq\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$，因此，在第(2)小题中的第二步后，要考虑恒等变形. 在不同题目中，可以有不同的恒等变形.

例 2-17　已知 n 阶方阵 $\boldsymbol{A},\boldsymbol{B}$ 满足 $\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}$，

(1) 求 $(\boldsymbol{A}-\boldsymbol{E})^{-1}$；

(2) 证明:$AB = BA$.

解 (1) 由已知 $AB = A + B \Rightarrow AB - A - B + E = E$

$$\Rightarrow \quad A(B-E) - (B-E) = E \Rightarrow (A-E)(B-E) = E$$

$$\Rightarrow \quad (A-E)^{-1} = B - E.$$

(2) 证 $E = (A-E)(B-E) = (B-E)(A-E) = BA - A - B + E$,

$$\Rightarrow \quad BA = A + B \Rightarrow \quad BA = AB.$$

评注 在方阵乘法中,若有 $AB = E$,则必有 $BA = E$.

例 2-18 证明:

$$R(A^*) = \begin{cases} n, & \text{若 } R(A) = n; \\ 1, & \text{若 } R(A) = n-1; \\ 0, & \text{若 } R(A) < n-1. \end{cases}$$

证 $$A \cdot A^* = |A| \cdot E$$

(1) 若 $R(A) = n \Rightarrow |A| \neq 0 \Rightarrow |A| \cdot |A^*| = |A|^n \Rightarrow |A^*| = |A|^{n-1} \neq 0 \Rightarrow A^*$ 可逆 $\Rightarrow R(A^*) = n$.(或者:A 可逆,$A \cdot A^*$ 不改变 A^* 的秩,而 $A \cdot A^* = |A| E$, $|A| \neq 0$,故 $|A| \cdot E$ 是可逆阵,秩为 n,故知 A^* 的秩为 n.)

(2) 若 $R(A) = n-1 \Rightarrow |A| = 0 \Rightarrow A \cdot A^* = O \Rightarrow R(A) + R(A^*) \leqslant n$

$\Rightarrow R(A^*) \leqslant n - (n-1) = 1$. 但由 $R(A) = n-1$,知 A^* 中至少有一个 $A_{ij} \neq 0$, $\Rightarrow A^* \neq O$

$\Rightarrow R(A^*) = 1$.

(3) 若 $R(A) < n-1$,A_{ij} 是 $n-1$ 阶代数余子式,故 A^* 中的全部 A_{ij} 皆为零,即 $A^* = O$

$\Rightarrow R(A^*) = 0$.

评注 这题应用了 A^* 的性质,同时,也易知,若 A^* 可逆,则 A 也可逆.

例 2-19 A 为 n 阶方阵,$n > 2$. 试证:$(A^*)^* = |A|^{n-2} \cdot A$.

证法 1 若 $|A| = 0 \Rightarrow R(A) \leqslant n-1$

$$\Rightarrow \quad R(A^*) \leqslant 1 < n-1 \quad \Rightarrow \quad R((A^*)^*) = 0$$

$$\Rightarrow \quad (A^*)^* = 0. \quad \text{得证.}$$

若 $|A| \neq 0$,则由 $$A^* = |A| \cdot A^{-1}$$

$$\Rightarrow \quad (A^*)^* = |A^*| \cdot (A^*)^{-1} = |A|^{n-1} \cdot (|A| A^{-1})^{-1} = |A|^{n-2} \cdot A.$$

证法 2 若 $|A| = 0$,证法同上.

若 $|A| \neq 0$,则由 $$A^* \cdot (A^*)^* = |A^*| \cdot E = |A|^{n-1} \cdot E.$$

在等式两边左乘 $\frac{1}{|A|} \cdot A$:

$$\frac{1}{|A|} \cdot A \cdot A^* \cdot (A^*)^* = |A|^{n-2} \cdot A,$$

即 $$(A^*)^* = |A|^{n-2} \cdot A.$$

例 2-20 试求满足 $(A^*)^* = A$ 的一切 n 阶矩阵 A.

解 由例 2-19 证得 $(A^*)^* = |A|^{n-2} \cdot A$,故即满足

$$|A|^{n-2} \cdot A = A, \text{即} (|A|^{n-2} - 1)A = 0.$$

所以,$A = O$,或者 $|A|^{n-2} = 1$,故所求矩阵 A 为

(1) $A = O$;

(2) $n=2$ 的 2 阶矩阵；

(3) 满足 $|\boldsymbol{A}|^{n-2}=1,(n>2)$ 的矩阵.

例 2-21　设 $\boldsymbol{A}=(a_{ij})_{3\times3}$ 是非零实方阵，若 $A_{ij}=a_{ij},(i=1,2,3;\ j=1,2,3)$.

(1) 试证 $|\boldsymbol{A}|\neq0$；

(2) 求 $|\boldsymbol{A}|,\boldsymbol{A}^{-1}$；

(3) 若 $a_{33}=1$，求解线性方程组：$\boldsymbol{AX}=\begin{pmatrix}0\\0\\1\end{pmatrix}$.

证　(1) 已知 $A_{ij}=a_{ij}$，即 $\boldsymbol{A}^{*}=\boldsymbol{A}^{\mathrm{T}}$.

由　　$\boldsymbol{AA}^{*}=|\boldsymbol{A}|\boldsymbol{E}$，　即有　$\boldsymbol{AA}^{\mathrm{T}}=|\boldsymbol{A}|\cdot\boldsymbol{E}$，

即　　$$\boldsymbol{AA}^{\mathrm{T}}=\begin{pmatrix}\sum\limits_{i=1}^{3}a_{1i}^{2}&&\\&\sum\limits_{i=1}^{3}a_{2i}^{2}&\\&&\sum\limits_{i=1}^{3}a_{3i}^{2}\end{pmatrix}=\begin{pmatrix}|\boldsymbol{A}|&&\\&|\boldsymbol{A}|&\\&&|\boldsymbol{A}|\end{pmatrix}.$$

由　　$$\boldsymbol{A}\neq\boldsymbol{O}\ \Rightarrow\ \sum_{i=1}^{3}a_{1i}^{2}+\sum_{i=1}^{3}a_{2i}^{2}+\sum_{i=1}^{3}a_{3i}^{2}=3\cdot|\boldsymbol{A}|\neq0,$$

即知　　$$|\boldsymbol{A}|\neq0.$$

解　(2) 由　　$$\boldsymbol{AA}^{\mathrm{T}}=|\boldsymbol{A}|\cdot\boldsymbol{E}\ \Rightarrow\ |\boldsymbol{A}|^{2}=|\boldsymbol{A}|^{3}$$

由　　$$|\boldsymbol{A}|\neq0\ \Rightarrow\ |\boldsymbol{A}|=1.$$

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\cdot\boldsymbol{A}^{*}=\boldsymbol{A}^{*}=\boldsymbol{A}^{\mathrm{T}},$$即知 $\boldsymbol{A}$ 是正交阵.

(3) 由　　$$\boldsymbol{AX}=\begin{pmatrix}0\\0\\1\end{pmatrix}\ \Rightarrow\ \boldsymbol{X}=\boldsymbol{A}^{-1}\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}=\boldsymbol{A}^{\mathrm{T}}\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}=\boldsymbol{A}^{\mathrm{T}}$$

的第三列，即 $\boldsymbol{A}$ 的第三行. 由于 $a_{33}=1$，且 $\boldsymbol{A}$ 是正交阵，故知 $\boldsymbol{A}$ 的第三行即为$(0,0,1)$.

所以，　　$$\boldsymbol{X}=\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

评注　(1) 由 $A_{ij}=a_{ij}$，即 $\boldsymbol{A}^{*}=\boldsymbol{A}^{\mathrm{T}}$，这个关系要理解.

(2) 对于矩阵乘法 $\boldsymbol{A}\cdot\boldsymbol{A}^{\mathrm{T}}$(或 $\boldsymbol{A}^{\mathrm{T}}\cdot\boldsymbol{A}$)，应知对角线元素就是 $\boldsymbol{A}$ 的每行元素的平方和.

例 2-22　设 $\boldsymbol{A}$ 为 n 阶可逆矩阵，$\boldsymbol{\alpha}$ 为 n 维列向量，b 为常数，记分块矩阵：

$$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{E}&\boldsymbol{O}\\-\boldsymbol{\alpha}^{\mathrm{T}}\cdot\boldsymbol{A}^{*}&|\boldsymbol{A}|\end{pmatrix},\quad\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{\alpha}\\\boldsymbol{\alpha}^{\mathrm{T}}&b\end{pmatrix}.$$

其中，$\boldsymbol{A}^{*}$ 是矩阵 $\boldsymbol{A}$ 的伴随阵，$\boldsymbol{E}$ 为 n 阶单位阵.

(1) 计算并化简 $\boldsymbol{P}\cdot\boldsymbol{Q}$；

(2) 证明：矩阵 $\boldsymbol{Q}$ 可逆的充要条件是 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{\alpha}\neq b$.

解　(1) 由 $\boldsymbol{AA}^{*}=\boldsymbol{A}^{*}\cdot\boldsymbol{A}=|\boldsymbol{A}|\cdot\boldsymbol{E}$，则

$$\boldsymbol{P}\cdot\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{E}&\boldsymbol{O}\\-\boldsymbol{\alpha}^{\mathrm{T}}\cdot\boldsymbol{A}^{*}&|\boldsymbol{A}|\end{pmatrix}\cdot\begin{pmatrix}\boldsymbol{A}&\boldsymbol{\alpha}\\\boldsymbol{\alpha}^{\mathrm{T}}&b\end{pmatrix}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{\alpha}\\-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{*}\boldsymbol{A}+|\boldsymbol{A}|\boldsymbol{\alpha}^{\mathrm{T}}&-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{*}\boldsymbol{\alpha}+b\cdot|\boldsymbol{A}|\end{pmatrix}$$

$$= \begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{O} & |\boldsymbol{A}|(b-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{\alpha}) \end{pmatrix}.$$

证 (2) 由(1)可得 $|\boldsymbol{P}\cdot\boldsymbol{Q}|=|\boldsymbol{P}|\cdot|\boldsymbol{Q}|=|\boldsymbol{A}|^2\cdot(b-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{\alpha})$,

又有 $$|\boldsymbol{P}|=|\boldsymbol{E}|\cdot|\boldsymbol{A}|=|\boldsymbol{A}|\neq 0,$$

故 $$|\boldsymbol{Q}|=|\boldsymbol{A}|\cdot(b-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{\alpha}),$$

即 $$|\boldsymbol{Q}|\neq 0 \quad\Leftrightarrow\quad b-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{\alpha}\neq 0.$$

例 2-23 (1) 若矩阵 $\boldsymbol{B}$ 是对调 $\boldsymbol{A}$ 的第三行与第五行得到的矩阵,则 $\boldsymbol{B}^{-1}$ 是对调 $\boldsymbol{A}^{-1}$ 的第三列与第五列得到的矩阵;

(2) 若 $\boldsymbol{A}$ 是对称阵,则任意对调 $\boldsymbol{A}$ 的两列及其相同的两行得到的矩阵仍对称.

证 (1) $\boldsymbol{B}=\boldsymbol{P}(3,5)\boldsymbol{A} \Rightarrow \boldsymbol{B}^{-1}=\boldsymbol{A}^{-1}\cdot\boldsymbol{P}(3,5)^{-1}=\boldsymbol{A}^{-1}\cdot\boldsymbol{P}(3,5)$.

(2) $\boldsymbol{B}=\boldsymbol{P}(i,j)\boldsymbol{A}\boldsymbol{P}(i,j) \Rightarrow \boldsymbol{B}^{\mathrm{T}}=\boldsymbol{P}(i,j)^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(i,j)^{\mathrm{T}}=\boldsymbol{P}(i,j)\boldsymbol{A}\boldsymbol{P}(i,j)=\boldsymbol{B}$.

例 2-24 设 $\boldsymbol{A}$ 是 n 阶可逆阵,将 $\boldsymbol{A}$ 的第 i 行和第 j 行对换后得到的矩阵记为 $\boldsymbol{B}$.

(1) 证明 $\boldsymbol{B}$ 可逆;

(2) 求 $\boldsymbol{A}\boldsymbol{B}^{-1}$.

解 (1) $\boldsymbol{B}=\boldsymbol{P}(i,j)\boldsymbol{A}$. 因为 $|\boldsymbol{P}(i,j)|=-1$, $|\boldsymbol{A}|\neq 0$,所以, $|\boldsymbol{B}|=-|\boldsymbol{A}|\neq 0$,即 $\boldsymbol{B}$ 可逆.

解 (2) $\boldsymbol{A}\boldsymbol{B}^{-1}=\boldsymbol{A}\cdot(\boldsymbol{P}(i,j)\boldsymbol{A})^{-1}=\boldsymbol{A}\cdot\boldsymbol{A}^{-1}\cdot\boldsymbol{P}(i,j)^{-1}=\boldsymbol{P}(i,j)$.

评注 对于初等变换,要会用初等阵.

例 2-25 设

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}, \quad \boldsymbol{B}=\begin{pmatrix} a_{14} & a_{13} & a_{12} & a_{11} \\ a_{24} & a_{23} & a_{22} & a_{21} \\ a_{34} & a_{33} & a_{32} & a_{31} \\ a_{44} & a_{43} & a_{42} & a_{41} \end{pmatrix},$$

$$\boldsymbol{P}_1=\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{P}_2=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

其中,$\boldsymbol{A}$ 可逆,则 $\boldsymbol{B}^{-1}=$ ( )

(A) $\boldsymbol{A}^{-1}\boldsymbol{P}_1\boldsymbol{P}_2$; (B) $\boldsymbol{P}_1\boldsymbol{A}^{-1}\boldsymbol{P}_2$;

(C) $\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{A}^{-1}$; (D) $\boldsymbol{P}_2\boldsymbol{A}^{-1}\boldsymbol{P}_1$.

解 逐个分析:

(1) 若 $\boldsymbol{B}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{P}_1\boldsymbol{P}_2 \Leftrightarrow \boldsymbol{B}=\boldsymbol{P}_2^{-1}\boldsymbol{P}_1^{-1}\boldsymbol{A} \Leftrightarrow \boldsymbol{A}=\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{B}$;

(2) 若 $\boldsymbol{B}^{-1}=\boldsymbol{P}_1\boldsymbol{A}^{-1}\boldsymbol{P}_2 \Leftrightarrow \boldsymbol{A}=\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{P}_1$;

(3) 若 $\boldsymbol{B}^{-1}=\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{A}^{-1} \Leftrightarrow \boldsymbol{A}=\boldsymbol{B}\boldsymbol{P}_1\boldsymbol{P}_2$;

(4) 若 $\boldsymbol{B}^{-1}=\boldsymbol{P}_2\boldsymbol{A}^{-1}\boldsymbol{P}_1 \Leftrightarrow \boldsymbol{A}=\boldsymbol{P}_1\boldsymbol{B}\boldsymbol{P}_2$.

由于 $\boldsymbol{P}_1=\boldsymbol{P}(1,4)$, $\boldsymbol{P}_2=\boldsymbol{P}(2,3)$, $\boldsymbol{A}$ 是 $\boldsymbol{B}$ 经过两次初等列变换而得到,易知(3)正确,故选(C).

例 2-26 设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶正交阵. 证明:

(1) 若 $|\boldsymbol{A}|+|\boldsymbol{B}|=0$,则 $|\boldsymbol{A}+\boldsymbol{B}|=0$;

(2) 若 $\boldsymbol{A}+\boldsymbol{B}$ 也是正交阵,则 $(\boldsymbol{A}+\boldsymbol{B})^{-1}=\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$;

(3) 若 n 为奇数,则 $|(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})|=0$.

证 (1) 因 $\boldsymbol{A},\boldsymbol{B}$ 是正交阵,故其行列式为 1 或 -1,则由

$$|\boldsymbol{A}|+|\boldsymbol{B}|=0 \quad \Rightarrow \quad |\boldsymbol{A}|\cdot|\boldsymbol{B}|=-1,$$

则
$$\begin{aligned}|\boldsymbol{A}+\boldsymbol{B}| &= |\boldsymbol{A}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}+\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}| = |\boldsymbol{A}|\cdot|\boldsymbol{B}^{\mathrm{T}}+\boldsymbol{A}^{\mathrm{T}}|\cdot|\boldsymbol{B}| \\ &= -|\boldsymbol{B}^{\mathrm{T}}+\boldsymbol{A}^{\mathrm{T}}| = -|\boldsymbol{B}+\boldsymbol{A}| = -|\boldsymbol{A}+\boldsymbol{B}|.\end{aligned}$$

移项即得
$$|\boldsymbol{A}+\boldsymbol{B}|=0.$$

(2) $(\boldsymbol{A}+\boldsymbol{B})^{-1}=(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$.

(3) $|(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})|=|\boldsymbol{A}+\boldsymbol{B}|\cdot|\boldsymbol{A}-\boldsymbol{B}|$.

若 $|\boldsymbol{A}|=-|\boldsymbol{B}|$，则由(1)已证，$|\boldsymbol{A}+\boldsymbol{B}|=0$.

若 $|\boldsymbol{A}|=|\boldsymbol{B}|$，则

$$\begin{aligned}|\boldsymbol{A}-\boldsymbol{B}| &= |\boldsymbol{A}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}-\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}| = |\boldsymbol{A}|\cdot|\boldsymbol{B}^{\mathrm{T}}-\boldsymbol{A}^{\mathrm{T}}||\boldsymbol{B}| \\ &= |\boldsymbol{B}^{\mathrm{T}}-\boldsymbol{A}^{\mathrm{T}}| = |\boldsymbol{B}-\boldsymbol{A}| = (-1)^{n}\cdot|\boldsymbol{A}-\boldsymbol{B}| \\ &= -|\boldsymbol{A}-\boldsymbol{B}|.\end{aligned}$$

移项便得
$$|\boldsymbol{A}-\boldsymbol{B}|=0.$$

综上，便有
$$|(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})|=0.$$

评注　(1) 行列式运算中，$|\boldsymbol{A}+\boldsymbol{B}|\neq|\boldsymbol{A}|+|\boldsymbol{B}|$；

(2) 在运用恒等变形时，有时用 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$，也有时用 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$. 这里主要为了左边可提出 $\boldsymbol{A}$，右边可提出 $\boldsymbol{B}$；

(3) 类似也可以证明："$\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶实对称阵，且 $\boldsymbol{A}^2=\boldsymbol{E},\boldsymbol{B}^2=\boldsymbol{E}$，$|\boldsymbol{A}|+|\boldsymbol{B}|=0$，则 $|\boldsymbol{A}+\boldsymbol{B}|=0$."

例 2-27　设
$$\boldsymbol{A}=\begin{pmatrix}1&0&0\\1&0&1\\0&1&0\end{pmatrix}.$$

试证：当 $n\geqslant 3$ 时，恒有 $\boldsymbol{A}^n=\boldsymbol{A}^{n-2}+\boldsymbol{A}^2-\boldsymbol{E}$，并计算 $\boldsymbol{A}^{100}$.

证　用归纳法，当 $n=3$ 时，

$$\boldsymbol{A}^2=\begin{pmatrix}1&0&0\\1&1&0\\1&0&1\end{pmatrix},\quad \boldsymbol{A}^3=\begin{pmatrix}1&0&0\\2&0&1\\1&1&0\end{pmatrix}.$$

易验得 $\boldsymbol{A}+\boldsymbol{A}^2-\boldsymbol{E}=\boldsymbol{A}^3$ 成立.

归纳假设结论对 $k\leqslant n-1$ 时正确，则 $k=n$ 时，

$$\begin{aligned}\boldsymbol{A}^n &= \boldsymbol{A}\cdot\boldsymbol{A}^{n-1}=\boldsymbol{A}\cdot(\boldsymbol{A}^{n-3}+\boldsymbol{A}^2-\boldsymbol{E}) \\ &= \boldsymbol{A}^{n-2}+\boldsymbol{A}^3-\boldsymbol{A}=\boldsymbol{A}^{n-2}+(\boldsymbol{A}+\boldsymbol{A}^2-\boldsymbol{E})-\boldsymbol{A} \\ &= \boldsymbol{A}^{n-2}+\boldsymbol{A}^2-\boldsymbol{E}. \quad \text{结论成立.}\end{aligned}$$

$$\begin{aligned}\boldsymbol{A}^{100} &= \boldsymbol{A}^{98}+\boldsymbol{A}^2-\boldsymbol{E}=\boldsymbol{A}^{96}+\boldsymbol{A}^2-\boldsymbol{E}+\boldsymbol{A}^2-\boldsymbol{E}=\boldsymbol{A}^{96}+2(\boldsymbol{A}^2-\boldsymbol{E}) \\ &= \cdots \\ &= \boldsymbol{A}^2+49(\boldsymbol{A}^2-\boldsymbol{E})=50\boldsymbol{A}^2-49\boldsymbol{E} \\ &= \begin{pmatrix}1&0&0\\50&1&0\\50&0&1\end{pmatrix}.\end{aligned}$$

评注　计算矩阵 $\boldsymbol{A}^n$，用归纳法是一种方法. 另外在下一章，若矩阵 $\boldsymbol{A}$ 可相似变换到对角阵，则计算 $\boldsymbol{A}^n$ 也方便. 还有，若 $R(\boldsymbol{A})=1$，则计算 $\boldsymbol{A}^n$ 也有简便的方法.

例 2-28 设 A 可逆，且 A 的各行元素之和都等于 k，试证：

(1) $k \neq 0$；

(2) A^{-1} 的各行元素之和都等于 $\frac{1}{k}$.

证 (1) 考虑 $|A|$. 把 A 的各列都加到第一列上去，则第一列全为 k，由 A 可逆，知 $|A| \neq 0$，则 $k \neq 0$.

(2) 把 A 和 A^{-1} 均按列分块，记为

$$A = (\alpha_1, \cdots, \alpha_n), \quad A^{-1} = (\beta_1, \cdots, \beta_n)$$

由

$$A^{-1}A = E \quad \Leftrightarrow \quad A^{-1}(\alpha_1, \cdots \alpha_n) = (\varepsilon_1, \cdots, \varepsilon_n)$$

$$\Leftrightarrow \quad A^{-1}\alpha_i = \varepsilon_i, \quad i = 1, \cdots, n,$$

则

$$A^{-1}(\alpha_1 + \alpha_2 + \cdots + \alpha_n) = A^{-1}\alpha_1 + \cdots + A^{-1}\alpha_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

又

$$A^{-1}(\alpha_1 + \alpha_2 + \cdots + \alpha_n) = A^{-1} \cdot \begin{pmatrix} k \\ k \\ \vdots \\ k \end{pmatrix} = (\beta_1, \cdots, \beta_n) \begin{pmatrix} k \\ k \\ \vdots \\ k \end{pmatrix}$$

$$= k\beta_1 + \cdots + k\beta_n = k(\beta_1 + \cdots + \beta_n),$$

则得

$$k(\beta_1 + \cdots + \beta_n) = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \text{即} \quad (\beta_1 + \cdots + \beta_n) = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k} \\ \vdots \\ \frac{1}{k} \end{pmatrix}.$$

证法 2 由 A 的各行元素之和皆为 k，可知：

$$A \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} k \\ k \\ \vdots \\ k \end{pmatrix} = k \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix},$$

即知 k 为 A 的特征值. 由于 A 可逆，故 $k \neq 0$.

而且，由上面关系式，可知

$$A^{-1} \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \frac{1}{k} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

即 A^{-1} 的各行元素之和皆为 $\frac{1}{k}$.

评注 本题和矩阵各行元素之和有关，因此把矩阵按列分块，考虑列向量之和是很自然的事.

证法 2 中，利用矩阵 A 和列向量 $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ 相乘，正好就表示了各行元素之和这样一个事实，使证明简洁得多

了.这种比较特殊的矩阵与列向量的乘法,应该熟练掌握.

例 2-29　已知 n 阶矩阵:

$$\boldsymbol{A}=\begin{pmatrix}0&1&0&\cdots&0\\0&0&2&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&n-1\\n&0&0&\cdots&0\end{pmatrix};$$

求 $\boldsymbol{A}$ 的全体 n^2 个元素的代数余子式 A_{ij} 的和.

解　记 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{O}&\boldsymbol{A}_1\\n&\boldsymbol{O}\end{pmatrix}$,　则 $\boldsymbol{A}^{-1}=\begin{pmatrix}\boldsymbol{O}&\frac{1}{n}\\\boldsymbol{A}_1^{-1}&\boldsymbol{O}\end{pmatrix}$.

即
$$\boldsymbol{A}^{-1}=\begin{pmatrix}0&0&\cdots&0&\frac{1}{n}\\1&0&\cdots&0&0\\0&\frac{1}{2}&\cdots&0&0\\\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&\cdots&\frac{1}{n-1}&0\end{pmatrix},$$

由
$$|\boldsymbol{A}|=(-1)^{n+1}\cdot n!$$

而
$$\boldsymbol{A}^*=|\boldsymbol{A}|\cdot\boldsymbol{A}^{-1}=(-1)^{n+1}\cdot n!\cdot\boldsymbol{A}^{-1},$$

所以,$\boldsymbol{A}$ 的全体元素的代数余子式的和即为

$$(-1)^{n+1}\cdot n!\cdot\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n-1}+\frac{1}{n}\right).$$

例 2-30　已知二阶矩阵 $\boldsymbol{M}$ 的伴随矩阵

$$\boldsymbol{M}^*=\begin{pmatrix}1&2\\2&4\end{pmatrix},\quad \text{求 } \boldsymbol{M}.$$

解　$\boldsymbol{M}^*=\begin{pmatrix}1&2\\2&4\end{pmatrix}=\begin{pmatrix}A_{11}&A_{21}\\A_{12}&A_{22}\end{pmatrix}$,　设　$\boldsymbol{M}=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}$.

则应有　$A_{11}=1=a_{22}$,

$A_{12}=2=-a_{21}$,　$\Rightarrow$　$a_{21}=-2$.

$A_{21}=2=-a_{12}$,　$\Rightarrow$　$a_{12}=-2$.

$A_{22}=4=a_{11}$.

所以,
$$\boldsymbol{M}=\begin{pmatrix}4&-2\\-2&1\end{pmatrix}.$$

例 2-31　已知 n 阶矩阵 $\boldsymbol{A}^2=\boldsymbol{A}$,$(\boldsymbol{A}+\boldsymbol{B})^2=\boldsymbol{A}^2+\boldsymbol{B}^2$.试证:$\boldsymbol{AB}=\boldsymbol{O}$.

证
$$(\boldsymbol{A}+\boldsymbol{B})^2=\boldsymbol{A}^2+\boldsymbol{AB}+\boldsymbol{BA}+\boldsymbol{B}^2=\boldsymbol{A}^2+\boldsymbol{B}^2$$

$$\Rightarrow\quad \boldsymbol{AB}+\boldsymbol{BA}=\boldsymbol{O}.$$

以 $\boldsymbol{A}$ 左乘上式两边:

$$A^2B+ABA=O,$$

即
$$AB+ABA=O.$$

$$\Rightarrow \quad AB(E+A)=O \tag{1}$$

由 $A^2=A \quad \Rightarrow \quad A^2-A=O$

$$\Rightarrow \quad (A+E)(A-2E)=-2E$$

$$\Rightarrow \quad A+E \text{ 可逆.}$$

则在式(1)中,易得 $AB=O$.

例 2-32 已知 n 阶矩阵 A 满足 $aA^2+bA+cE=O(a\neq 0)$. 若矩阵 A 对于任何实数 k,$A-kE$ 都可逆,问 a,b,c 需满足什么条件?(a,b,c 均为实数).

解 $aA^2+bA+cE=a\left(A^2+\frac{b}{a}A\right)+cE=a(A-kE)\cdot\left[A+\left(\frac{b}{a}+k\right)E\right]+ak\left(\frac{b}{a}+k\right)E+cE$

$$=a(A-kE)\cdot\left[A+\left(\frac{b}{a}+k\right)E\right]+(ak^2+bk+c)E=O.$$

故
$$a(A-kE)\left[A+\left(\frac{b}{a}+k\right)E\right]=-(ak^2+bk+c)E$$

即知:当 $ak^2+bk+c\neq 0$ 时,$A-kE$ 可逆. 而对任何实数 k,$ak^2+bk+c\neq 0$,即一元二次方程 $ax^2+bx+c=0$ 无实数解,即 $b^2-4ac<0$.

评注 由本题结果可知,若 A 满足 $3A^2+4A+5E=O$,则对任何实数 k,$A-kE$ 都可逆. 类似地,$2A^2-3A+2E=O$,$A^2+A+E=0$ 等都有此结论.

例 2-33 设 $Q=\begin{pmatrix}A & B\\ C & D\end{pmatrix}$,且 A 可逆,D 也是方阵. 证明:$|Q|=|A|\cdot|D-CA^{-1}B|$.

证 做初等行变换,把 Q 化成准对角阵. 令 $P_1=\begin{pmatrix}E_1 & O\\ X & E_2\end{pmatrix}$,其中,$E_1$ 与 A 同阶,E_2 与 D 为同阶的单位阵,使 $P_1\cdot Q=\begin{pmatrix}A & B\\ O & M\end{pmatrix}$,易得 $X=-CA^{-1}$,则

$$P_1Q=\begin{pmatrix}E_1 & O\\ -CA^{-1} & E_2\end{pmatrix}\cdot\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\begin{pmatrix}A & B\\ O & D-CA^{-1}B\end{pmatrix}.$$

则
$$|P_1|\cdot|Q|=|A|\cdot|D-CA^{-1}B|,$$

其中 $|P_1|=1$,即有

$$|Q|=|A|\cdot|D-CA^{-1}B|.$$

例 2-34 设 A,B 为 n 阶矩阵. 证明:

(1) $\begin{vmatrix}E & B\\ A & E\end{vmatrix}=|E-AB|$;

(2) $|E-AB|=|E-BA|$.

证 (1) 考虑初等行变换将矩阵化为准对角阵. 设 $P_1=\begin{pmatrix}E & O\\ X & E\end{pmatrix}$,其中,$X$ 为待定矩阵.

$$P_1\cdot\begin{pmatrix}E & B\\ A & E\end{pmatrix}=\begin{pmatrix}E & O\\ X & E\end{pmatrix}\begin{pmatrix}E & B\\ A & E\end{pmatrix}=\begin{pmatrix}E & B\\ X+A & XB+E\end{pmatrix}.$$

可取 $X=-A$，得 $P_1=\begin{pmatrix}E & O\\ -A & E\end{pmatrix}$，则上式即为

$$P_1\cdot\begin{pmatrix}E & B\\ A & E\end{pmatrix}=\begin{pmatrix}E & B\\ O & E-AB\end{pmatrix}.$$

等式两边取行列式，注意到 $|P_1|=1$，$|E|=1$，即有

$$\begin{vmatrix}E & B\\ A & E\end{vmatrix}=|E-AB|.$$

(2) 由(1)知：

$$|E-AB|=\begin{vmatrix}E & B\\ A & E\end{vmatrix}=(-1)^n\cdot\begin{vmatrix}A & E\\ E & B\end{vmatrix}=(-1)^{2n}\cdot\begin{vmatrix}E & A\\ B & E\end{vmatrix}$$

$$=\begin{vmatrix}E & A\\ B & E\end{vmatrix}=|E-BA|.$$

例 2-35　设 A,B 都是 n 阶矩阵. 试证：

$$|\lambda E-AB|=|\lambda E-BA|.$$

证法 1　构造矩阵乘法：

$$\begin{pmatrix}E & -A\\ O & \lambda E\end{pmatrix}\cdot\begin{pmatrix}A & \lambda E\\ E & B\end{pmatrix}=\begin{pmatrix}O & \lambda E-AB\\ \lambda E & \lambda B\end{pmatrix}.$$

两边取行列式：

$$|E|\cdot|\lambda E|\cdot\begin{vmatrix}A & \lambda E\\ E & B\end{vmatrix}=(-1)^n\cdot|\lambda E|\cdot|\lambda E-AB|,$$

即有

$$\begin{vmatrix}A & \lambda E\\ E & B\end{vmatrix}=(-1)^n\cdot|\lambda E-AB|. \tag{1}$$

类似，再构造矩阵乘法：

$$\begin{pmatrix}E & O\\ -B & \lambda E\end{pmatrix}\cdot\begin{pmatrix}A & \lambda E\\ E & B\end{pmatrix}=\begin{pmatrix}A & \lambda E\\ \lambda E-BA & O\end{pmatrix},$$

两边取行列式：

$$|E|\cdot|\lambda E|\cdot\begin{vmatrix}A & \lambda E\\ E & B\end{vmatrix}=(-1)^n\cdot|\lambda E|\cdot|\lambda E-BA|$$

即有

$$\begin{vmatrix}A & \lambda E\\ E & B\end{vmatrix}=|\lambda E-BA| \tag{2}$$

由式(1)、式(2)得

$$|\lambda E-AB|=|\lambda E-BA|.$$

证法 2　利用例 2-34 的结果：$|E-AB|=|E-BA|$. 则若 $\lambda\neq 0$，则

$$\left|E-\frac{A}{\lambda}B\right|=\left|E-B\cdot\frac{A}{\lambda}\right|,$$

即有

$$|\lambda E-AB|=|\lambda E-BA|.$$

若 $\lambda=0$，则 $|-AB|=(-1)^n|AB|=(-1)^n|BA|=|-BA|$. 故对任何 λ，均有

$$|\lambda E-AB|=|\lambda E-BA|.$$

评注 本题结论告诉我们，$A \cdot B$ 与 $B \cdot A$ 的特征多项式相同，因此，特征值也相同.

例 2-36 设 $\boldsymbol{\alpha}=(1,0,-1)^{\mathrm{T}}$，矩阵 $\boldsymbol{A}=\boldsymbol{\alpha} \cdot \boldsymbol{\alpha}^{\mathrm{T}}$，$n$ 为正整数，求 $|a\boldsymbol{E}-\boldsymbol{A}^n|$.

解法 1
$$\boldsymbol{A}^n=(\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}})^n=\boldsymbol{\alpha}(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha})\cdot\cdots\cdot(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha})\cdot\boldsymbol{\alpha}^{\mathrm{T}}.$$

而 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}=2$，故

$$\boldsymbol{A}^n=2^{n-1}\cdot\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}=2^{n-1}\cdot\begin{pmatrix}1&0&-1\\0&0&0\\-1&0&1\end{pmatrix}.$$

则
$$|a\boldsymbol{E}-\boldsymbol{A}^n|=\begin{vmatrix}a-2^{n-1}&0&2^{n-1}\\0&a&0\\2^{n-1}&0&a-2^{n-1}\end{vmatrix}=a^2(a-2^n).$$

解法 2 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}=\begin{pmatrix}1&0&-1\\0&0&0\\-1&0&1\end{pmatrix}$ 是实对称，且秩为 1，则易知 $\boldsymbol{A}$ 的三个特征值为 0，0，2，则 $\boldsymbol{A}^n$ 的三个特征值为 0，0，2^n. 故

$$|a\boldsymbol{E}-\boldsymbol{A}^n|=(a-0)(a-0)(a-2^n)=a^2\cdot(a-2^n).$$

评注 对于秩为 1 的对称矩阵，通常都能表为 $\boldsymbol{A}=\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}^{\mathrm{T}}$ 的形式. 这里 $\boldsymbol{\alpha}$ 是 n 维列向量. 若 $\boldsymbol{\alpha}$ 是实向量，则 $\boldsymbol{A}$ 又是实对称. 对于计算这种类型的 $\boldsymbol{A}^n$，则易知 $\boldsymbol{A}^n=\boldsymbol{K}^{n-1}\cdot\boldsymbol{A}$，其中 $\boldsymbol{K}=\boldsymbol{\alpha}^{\mathrm{T}}\cdot\boldsymbol{\alpha}$.

四、小　结

矩阵运算是历年试卷中总要考到的内容，矩阵运算涉及到的知识点很多，经常需要把矩阵的运算（如加法、数乘、乘法、转置、求逆等）与各类矩阵（如对称阵、反对称阵、可逆阵、伴随阵、正交阵、初等阵等）的基本性质加以综合运用. 一般，先要对给定的表达式进行化简，然后再进行计算.

涉及到初等变换的，要考虑初等矩阵这个工具. 例 2-33、例 2-34、例 2-35 都可视为初等矩阵的推广. 这类题目，左乘凑上的矩阵与乘积后的矩阵都是分块三角形矩阵，且这两个矩阵的行列式便于求出，这是原则，同时，对角线上的小块矩阵必须是方阵.

在等式两边同时作一种运算，这是在做题时经常会用到的方法. 这种运算，可以是两边取行列式，也可以是两边左乘（或右乘）一个矩阵（或向量），也可以是两边取逆，或两边转置等，这要视不同题目的条件而定.

分块矩阵的运用会使证明和计算简洁，应在学习例题时注意掌握这种技巧.

对于一些较为特殊的矩阵运算应该熟悉. 例如 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 的对角线元素；$\boldsymbol{A}\cdot\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$ 即表出 $\boldsymbol{A}$ 的各行元素之和；$\boldsymbol{A}\cdot\begin{pmatrix}0\\\vdots\\1\\\vdots\\0\end{pmatrix}i$ 即表示 $\boldsymbol{A}$ 的第 i 列等.

第三章　向量组的线性相关性与矩阵的秩

本章是线性代数中理论比较集中，而且是较为抽象的一章.

一、复习与考试要求

(1) 理解 n 维向量，向量的线性组合与线性表示的概念.

(2) 理解向量组线性相关，线性无关的定义，了解并会用向量组线性相关、线性无关的有关性质及判别法.

(3) 了解向量组的极大线性无关组和向量组的秩的概念，会求向量组的极大线性无关组及秩.

(4) 了解向量组等价的概念，以及向量组的秩与矩阵的秩的关系.

(5) 了解 n 维向量空间、子空间、基底、维数、坐标等概念.

(6) 了解基变换和坐标变换公式，会求过渡矩阵.

(7) 了解内积的概念，掌握线性无关向量组正交规范化的施密特(Schmidt) 方法.

(8) 了解规范正交基、正交矩阵的概念，以及它们的性质.

二、基本概念与理论

1. 线性方程组有解与向量的线性关系

$$\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times 1}=\boldsymbol{b}_{m\times 1}\text{ 有解}$$

$$\Leftrightarrow\quad (\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=b\text{ 有解，}(\boldsymbol{\alpha}_i\text{ 为 }\boldsymbol{A}\text{ 的列向量}).$$

$$\Leftrightarrow\quad x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n=b\text{ 有解.}$$

$$\Leftrightarrow\quad \text{向量 } b \text{ 可以由向量组 } \boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n \text{ 线性表示}$$

$$\Leftrightarrow\quad \text{向量组}\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n,b\}\text{ 与向量组}\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}\text{ 等价}$$

$$\Leftrightarrow\quad R(\overline{\boldsymbol{A}})=R(\boldsymbol{A})$$

注意　最后这个充分必要条件，对于一般的两个向量组，若它们的秩相等，并不能推出它们等价，但这里的这两个向量组中，还有“一个向量组可以由另一个向量组线性表示”这个条件.

2. 向量组的线性表示与矩阵乘法的关系

若向量组$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\}$ 可以由向量组$\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s\}$ 线性表示，即有

$$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\}=\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s\}\cdot\begin{pmatrix}k_{11}&k_{12}&\cdots&k_{1r}\\ k_{21}&k_{22}&\cdots&k_{2r}\\ \vdots&\vdots&\ddots&\vdots\\ k_{s1}&k_{s2}&\cdots&k_{sr}\end{pmatrix}$$

反之，若有矩阵乘法 $\boldsymbol{A}=\boldsymbol{B}\cdot\boldsymbol{K}$，并写成如上形式，即知 $\boldsymbol{A}$ 的列向量组可以由 $\boldsymbol{B}$ 的列向量组线性表示.

同理可知：$\boldsymbol{A}$ 的行向量组可以由 $\boldsymbol{K}$ 的行向量组线性表示 $\Leftrightarrow\boldsymbol{A}=\boldsymbol{BK}$.

3. 矩阵方程

对于矩阵方程 $\boldsymbol{A}_{m\times r}\boldsymbol{X}_{r\times s}=\boldsymbol{B}_{m\times s}$ 有解

$\Leftrightarrow$ $\boldsymbol{B}$ 的列向量组可以由 $\boldsymbol{A}$ 的列向量组线性表示

$\Leftrightarrow$ $(\boldsymbol{A},\boldsymbol{B})$ 的列组与 $\boldsymbol{A}$ 的列组等价

$\Leftrightarrow$ $R(\boldsymbol{A},\boldsymbol{B})=R(\boldsymbol{A})$.

这可视为线性方程组有解判别定理在矩阵方程的推广.

4. 向量组线性相关与线性无关

(1) **定义** 一组 n 维向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性相关,即存在一组不全为零的数 $\lambda_1,\cdots,\lambda_m$,使

$$\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_m\boldsymbol{\alpha}_m=0.$$

若 $k_1\boldsymbol{\alpha}_1+\cdots+k_m\boldsymbol{\alpha}_m=0$ 时,必须 $k_1=\cdots=k_m=0$ 时才成立,则称 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性无关.

(2) **等价定义** $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性相关 $\Leftrightarrow$ 组中有一个向量可以由其余向量线性表示.

$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性无关 $\Leftrightarrow$ 组中任一向量皆不能由其余向量线性表示.

(3) 按定义论证相关性时,还常用到如下的一些结论:

① 单个向量线性相关 $\Leftrightarrow$ 此向量是零向量;

② 含有零向量的向量组一定线性相关;

③ 若向量组中有一部分向量组成的部分组线性相关,则整组也线性相关;

④ 若向量组整组线性无关,则它的任一部分组也线性无关;

⑤ 若向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性无关,而 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}$ 线性相关,则 $\boldsymbol{\beta}$ 一定可由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性表示,且表法惟一;

⑥ 若向量组向量的个数大于向量的维数时,则此向量组必线性相关;

⑦ 若一向量组线性无关,则它的延长向量组(即维数增加后所成的向量组)也线性无关.

(4) 判别向量组的线性相关性时,有时还可用下述方法:

$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性相关 $\Leftrightarrow$ 存在一组不全为零的数 $X_1,\cdots,X_m$,使 $X_1\boldsymbol{\alpha}_1+\cdots+X_m\boldsymbol{\alpha}_m=0$

$$\Leftrightarrow \quad \text{齐次线性方程组}(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m)\begin{pmatrix}X_1\\ \vdots\\ X_m\end{pmatrix}=0,\text{有非零解}.$$

$\Leftrightarrow$ 秩$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m)<m$($m$ 即为未知量个数). 当$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m)$ 是方阵 $\boldsymbol{A}$ 时,还 $\Leftrightarrow |\boldsymbol{A}|=0$.

$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性无关 $\Leftrightarrow$ 当 $X_1\boldsymbol{\alpha}_1+\cdots+X_m\boldsymbol{\alpha}_m=0$ 时,系数 x_i 必须全等于零$(i=1,\cdots,m)$.

$$\Leftrightarrow \quad (\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m)\begin{pmatrix}X_1\\ \vdots\\ X_m\end{pmatrix}=0,\text{只有零解}$$

$\Leftrightarrow$ 秩$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m)=m$

$\Leftrightarrow$ $|\boldsymbol{A}|\neq 0$ (当$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m)$ 是方阵 $\boldsymbol{A}$ 时).

5. 向量组的秩

(1) **两个向量组等价的定义** 若向量组 Ⅰ 中的每个向量皆可由向量组 Ⅱ 线性表示,同时,向量组 Ⅱ 中的每个向量皆可由向量组 Ⅰ 线性表示,则称这两个向量组等价.

(2) **极大线性无关组的定义** 设向量组 $T:\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$,在 T 中选取 r 个向量 $\boldsymbol{\alpha}_{i1},\boldsymbol{\alpha}_{i2},\cdots,\boldsymbol{\alpha}_{ir}$,若满足:① $\boldsymbol{\alpha}_{i1},\boldsymbol{\alpha}_{i2},\cdots,\boldsymbol{\alpha}_{ir}$ 线性无关;② 任取 $\boldsymbol{\alpha}\in T$,则向量组 $\boldsymbol{\alpha}_{i1},\boldsymbol{\alpha}_{i2},\cdots,\boldsymbol{\alpha}_{ir},\boldsymbol{\alpha}$ 线性相关,则称向量组 $\boldsymbol{\alpha}_{i1},\boldsymbol{\alpha}_{i2},\cdots,\boldsymbol{\alpha}_{ir}$ 为向量组 T 的一个极大线性无关组.

(3) **向量组的秩的定义** 即向量组的极大线性无关组所含向量的个数.

(4) **定理** 若向量组$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\}$ 可由向量组$\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s\}$ 线性表示,且 $r>s$,则 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性相关(证明见本章例 3-23).

推论 1 若向量组$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\}$ 可由向量组$\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s\}$ 线性表示,且 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性无关,则 $r\leqslant s$.

推论 2 等价的线性无关的向量组含有相同个数的向量.

推论 3 向量组的任两个极大线性无关组含有相同个数的向量.

推论 4　任意 $n+1$ 个 n 维向量必线性相关.

命题　向量组与它的极大线性无关组等价.

推论 5　若向量组 Ⅰ 可以由向量组 Ⅱ 线性表示,则秩(Ⅰ)$\leqslant$秩(Ⅱ).

推论 6　若矩阵 $\boldsymbol{C}_{m\times n}=\boldsymbol{A}_{m\times s}\boldsymbol{B}_{s\times n}$,则 $R(\boldsymbol{C})\leqslant\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}$(见本章例 3-23).

推论 7　若向量组 Ⅰ 与向量组 Ⅱ 等价,则秩(Ⅰ)$=$秩(Ⅱ).

推论 8　若向量组 Ⅰ 与 Ⅱ,有秩(Ⅰ)$=$秩(Ⅱ),且 Ⅰ 可由 Ⅱ 线性表示,则 Ⅰ 与 Ⅱ 等价(见本章例 3-22).

6. 矩阵的秩

(1) **定义**　若在矩阵 $\boldsymbol{A}$ 中有一个 r 阶子式 $D_r\neq 0$,而所有的 $r+1$ 阶子式(若存在的话)全等于零,则称矩阵 $\boldsymbol{A}$ 的秩为 r,记为 $R(\boldsymbol{A})=r$.

由定义可知　$$R(\boldsymbol{A})=R(\boldsymbol{A}^{\mathrm{T}}).$$

(2) 矩阵 $\boldsymbol{A}$ 经过初等变换后化到阶梯型,阶梯型矩阵中非零行的行数即为矩阵的秩.

(3) **定理**　$R(\boldsymbol{A})=\boldsymbol{A}$ 的列秩 $=\boldsymbol{A}$ 的行秩.

(4) **定理**　初等变换不改变矩阵的秩.

(5) **求极大线性无关组**　把向量按列排成一个矩阵,然后用行初等变换化到阶梯型,若矩阵的秩为 r,则可在相应的 r 列中找到极大线性无关组.

(6) 常用的矩阵的秩的关系式:

① $R(\boldsymbol{A}+\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B})$;

② $R(\boldsymbol{A}\cdot\boldsymbol{B})\leqslant\min\{R(\boldsymbol{A})\cdot R(\boldsymbol{B})\}$;

③ $R(\boldsymbol{A})=R(\boldsymbol{PA})=R(\boldsymbol{AQ})=R(\boldsymbol{PAQ})$,这里,$\boldsymbol{P},\boldsymbol{Q}$ 皆为可逆阵.

7. 向量空间

(1) **向量空间的定义**　V 是一个非空集合,若 V 中元素对于所定义的加法和数乘这两种运算封闭,就称 V 为向量空间,也叫线性空间,简称空间.

(2) **空间的基的定义**　若空间 V 中有一组向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 满足两条:

① $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性无关;

② V 中任一向量 $\boldsymbol{\alpha}$ 皆可由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性表示,则称 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 为空间 V 的一个基.

(3) **向量空间 $\boldsymbol{V}$ 的维数的定义**　即 V 的基中所含向量的个数,记为 $\dim(V)$. 特别规定:只含有零向量的空间,其维数为零.

(4) **齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{O}$**,若有非零解,则它的所有解形成空间,也称为解空间,记为 $\boldsymbol{\eta}(\boldsymbol{A})$. 解空间的基即为基础解系,解空间的维数 $\dim(\boldsymbol{\eta}(\boldsymbol{A}))=n-R(\boldsymbol{A})$.

(5) 由向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 所生成的空间,记为

$$L(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m)=\{x\mid x=\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_m\boldsymbol{\alpha}_m,\ \forall\lambda_1,\cdots,\lambda_m\in F\}.$$

(6) **子空间的定义**　若有向量空间 V_1 及 V_2,若 $V_1\subset V_2$,就称 V_1 是 V_2 的子空间.

(7) **空间的基不惟一.** 若 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 是空间 V 的两组基,则必有 $(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)\boldsymbol{P}$,称 r 阶方阵 $\boldsymbol{P}$ 为从基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 到基 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 的过渡矩阵. 过渡矩阵一定是可逆矩阵.

(8) **坐标变换**　若 V 是 n 维线性空间,$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 与 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n$ 是 V 的两组基,任给 $\boldsymbol{\alpha}\in V$,$\boldsymbol{\alpha}$ 在这两组基下的坐标分别为

$$\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\quad\text{与}\quad\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},$$

即　$$\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

$$= y_1\boldsymbol{\beta}_1 + \cdots + y_n\boldsymbol{\beta}_n = (\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

又有 $$(\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n) = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{P}$$

则 $$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{P} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

三、基本题型与解题方法

例 3-1 若 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性无关，$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$，$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3$，$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_3 + \boldsymbol{\alpha}_1$，证明 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ 也线性无关.

分析 这是一组抽象向量，通常用定义证明.

证 设有一组数 k_1, k_2, k_3，使

$$k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2 + k_3\boldsymbol{\beta}_3 = 0,$$

得

$$k_1(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2) + k_2(\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3) + k_3(\boldsymbol{\alpha}_3 + \boldsymbol{\alpha}_1) = 0,$$

$$(k_1 + k_3)\boldsymbol{\alpha}_1 + (k_1 + k_2)\boldsymbol{\alpha}_2 + (k_2 + k_3)\boldsymbol{\alpha}_3 = 0,$$

因为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性无关，由定义得

$$k_1 + k_3 = k_1 + k_2 = k_2 + k_3 = 0,$$

推得 $k_1 = k_2 = k_3 = 0$，所以 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ 线性无关.

评注 该题可以推广到 m 个向量的情形，此时的结论是：

(1) 当 m 为奇数时，$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 线性无关；

(2) 当 m 为偶数时，$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 线性相关.

例 3-2 $\boldsymbol{\alpha}_i = (1, i+1, (i+1)^2), i = 1, 2, 3$，求证：$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性无关.

分析 本题所给的是具体的向量组，故可直接用公式证之.

证

$$\begin{vmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \boldsymbol{\alpha}_3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{vmatrix},$$

这是一个范德蒙行列式，它不为 0，从而 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_s$ 线性无关.

例 3-3 试判别向量组 $\boldsymbol{\alpha}_1 = (2, 1, 0, 0), \boldsymbol{\alpha}_2 = (3, 2, 1, 0), \boldsymbol{\alpha}_3 = (4, 2, 2, 1)$ 的线性相关性.

分析 这是一组具体的向量，通常用求秩的方法加以判别.

解

记 $$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \boldsymbol{\alpha}_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 4 & 2 & 2 & 1 \end{bmatrix},$$

显见 $\boldsymbol{A}$ 的后三列所成行列式不为 0，即

$$\begin{vmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{vmatrix} = 1 \neq 0,$$

可知 $R(\boldsymbol{A})=3$，从而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关.

例 3-4　已知向量组 $\boldsymbol{\alpha}_1=(1,2,3),\boldsymbol{\alpha}_2=(3,-1,2),\boldsymbol{\alpha}_3=(2,3,t)$，问

(1) 当 t 取何值时，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，并将 $\boldsymbol{\alpha}_3$ 表示成 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 的线性组合；

(2) 当 t 取何值时，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关.

分析　将 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 作为矩阵 $\boldsymbol{A}$ 的列向量组，对 $\boldsymbol{A}$ 进行行初等变换，则列向量组之间的线性关系不会改变. 若将 $\boldsymbol{A}$ 经行初等变换化成一个行最简形矩阵，则列向量组之间的线性关系即可一眼看出.

所谓行最简形矩阵，是指一个具有下述特性的行阶梯形矩阵：

(1) 每个非零行的第一个非零元素(也称首元素)均为 1；

(2) 首元素 1 所在列的其余元素全为零.

解

$$(\boldsymbol{\alpha}_1^{\mathrm{T}},\boldsymbol{\alpha}_2^{\mathrm{T}},\boldsymbol{\alpha}_3^{\mathrm{T}})=\begin{bmatrix}1&3&2\\2&-1&3\\3&2&t\end{bmatrix}\xrightarrow{\text{行初等变换}}\begin{bmatrix}1&0&11/7\\0&1&1/7\\0&0&t-5\end{bmatrix}=(\boldsymbol{\beta}_1^{\mathrm{T}},\boldsymbol{\beta}_2^{\mathrm{T}},\boldsymbol{\beta}_3^{\mathrm{T}}),$$

(1) 当 $t=5$ 时，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，且有

$$\boldsymbol{\beta}_3=\frac{11}{7}\boldsymbol{\beta}_1+\frac{1}{7}\boldsymbol{\beta}_2,$$

由于行初等变换并不改变列向量之间的线性关系，所以

$$\boldsymbol{\alpha}_3=\frac{11}{7}\boldsymbol{\alpha}_1+\frac{1}{7}\boldsymbol{\alpha}_2;$$

(2) 当 $t\neq 5$ 时，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关.

例 3-5　设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，向量组 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关，问

(1) $\boldsymbol{\alpha}_1$ 能否由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示？证明你的结论；

(2) $\boldsymbol{\alpha}_4$ 能否由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示？证明你的结论.

分析　这是 1992 年数学(试卷一)的 7 分考题，目的是考查考生对向量组线性相关性概念的理解和基本证明方法的掌握.

由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，可知其中必有一个向量可由其余向量线性表示，至于 $\boldsymbol{\alpha}_1$ 能否由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，还要看其他条件. 由于 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关，可以断言 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 也线性无关，从而得出 $\boldsymbol{\alpha}_1$ 能由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示.

证　(1) 因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，所以存在不全为零的数 k_1,k_2,k_3 使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3=0,$$

如果 $k_1=0$，则得出 $k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3=0$，表明 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，与题设向量组 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关矛盾，所以 $k_1\neq 0$，这时有

$$\boldsymbol{\alpha}_1=-\frac{1}{k_1}(k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3)=l_2\boldsymbol{\alpha}_2+l_3\boldsymbol{\alpha}_3,$$

即 $\boldsymbol{\alpha}_1$ 能由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示.

(2) 反证法　如果 $\boldsymbol{\alpha}_4$ 能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，即有

$$\boldsymbol{\alpha}_4=\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\lambda_3\boldsymbol{\alpha}_3,$$

则由上面(1)中已证得的结果知

$$\boldsymbol{\alpha}_4=\lambda_1(l_2\boldsymbol{\alpha}_2+l_3\boldsymbol{\alpha}_3)+\lambda_2\boldsymbol{\alpha}_2+\lambda_3\boldsymbol{\alpha}_3=(\lambda_1l_2+\lambda_2)\boldsymbol{\alpha}_2+(\lambda_1l_3+\lambda_3)\boldsymbol{\alpha}_3,$$

即 $\boldsymbol{\alpha}_4$ 可由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，这与 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关矛盾，因此 $\boldsymbol{\alpha}_4$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示.

例 3-6　设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，且可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示，证明存在一个 $\boldsymbol{\beta}_k(1\leqslant k\leqslant s)$，

使得 $\boldsymbol{\beta}_k,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关.

分析 证明存在或不存在的结论时常用反证法. 本题也宜用反证法,如果结论不成立,即对任何 k 有 $\boldsymbol{\beta}_k,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性相关,则 $\boldsymbol{\beta}_k$ 可由 $\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示,从而

$$\operatorname{rank}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m)\leqslant\operatorname{rank}(\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)=r-1,$$

这与条件矛盾.

证 因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关,且可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示,所以得 $r=\operatorname{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)\leqslant\operatorname{rank}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s)$. 如果求证的结论不成立,即对任何 k 有 $\boldsymbol{\beta}_k,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性相关,则 $\boldsymbol{\beta}_k$ 可由 $\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示,从而

$$\operatorname{rank}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s)\leqslant\operatorname{rank}(\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)=r-1,$$

这与上面已证得的 $\operatorname{rank}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s)\geqslant r$ 矛盾,所以结论成立.

例 3-7 设 $\boldsymbol{A}$ 为 n 阶方阵,求证 $\boldsymbol{A}$ 的秩 $R(\boldsymbol{A})=1$ 的充分必要条件为 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}$,其中 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 为 n 维非零列向量.

分析 充分性易证,由 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}$ 及 $\boldsymbol{\alpha}\neq0,\boldsymbol{\beta}\neq0$ 即可. 困难在于必要性的证明. 为了将 $\boldsymbol{A}$ 分解成两个向量 $\boldsymbol{\alpha},\boldsymbol{\beta}^{\mathrm{T}}$ 的乘积,需先用初等变换将 $\boldsymbol{A}$ 分解成两个向量乘积的简单形状.

证 充分性由题设得

$$R(\boldsymbol{A})\leqslant\min(R(\boldsymbol{\alpha}),R(\boldsymbol{\beta}^{\mathrm{T}}))\leqslant1,$$

而 $\boldsymbol{\alpha}\neq0,\boldsymbol{\beta}\neq0$,所以 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}$ 中至少有一个非零元,故得 $R(\boldsymbol{A})\geqslant1$,从而 $R(\boldsymbol{A})=1$.

必要性 因 $R(\boldsymbol{A})=1$,所以存在一系列初等方阵 $\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_t$,和 $\boldsymbol{Q}_1,\boldsymbol{Q}_2,\cdots,\boldsymbol{Q}_s$ 使

$$\boldsymbol{P}_t\cdots\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1\cdots\boldsymbol{Q}_s=\begin{bmatrix}1&&&\\&0&&\\&&\ddots&\\&&&0\end{bmatrix}=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}(1\quad0\cdots0),$$

得

$$\boldsymbol{A}=\boldsymbol{P}_1^{-1}\cdots\boldsymbol{P}_t^{-1}\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}(1\quad0\quad\cdots\quad0)\boldsymbol{Q}_s^{-1}\cdots\boldsymbol{Q}_1^{-1},$$

记

$$\boldsymbol{\alpha}=\boldsymbol{P}_1^{-1}\cdots\boldsymbol{P}_t^{-1}\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\boldsymbol{\beta}^{\mathrm{T}}=(1\quad0\quad\cdots\quad0)\boldsymbol{Q}_s^{-1}\cdots\boldsymbol{Q}_1^{-1},$$

则 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 都为 n 维非零列向量,证毕.

如果 $\boldsymbol{A}$ 为 $m\times n$ 矩阵,结论也成立,此时 $\boldsymbol{\alpha}$ 为 m 维非零列向量,而 $\boldsymbol{\beta}^{\mathrm{T}}$ 为 n 维非零行向量.

例 3-8 设实矩阵 $\boldsymbol{A}_{m\times n}$ 的秩为 $R(\boldsymbol{A})=m<n$,$\boldsymbol{I}_m$ 为 m 阶单位阵,下述结论中正确的是 ( )

(A) $\boldsymbol{A}$ 的任意 m 个列向量必线性无关;

(B) $\boldsymbol{A}$ 的任意一个 m 阶子式不等于零;

(C) 若矩阵 $\boldsymbol{B}$ 满足 $\boldsymbol{B}\boldsymbol{A}=\boldsymbol{O}$,则 $\boldsymbol{B}=\boldsymbol{O}$,或非齐次线性方程组 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$ 一定有无穷多组解;

(D) $\boldsymbol{A}$ 通过初等行变换,必可化为 $(\boldsymbol{I}_m,\boldsymbol{O})$ 的形式.

解 选(C).

由 $\boldsymbol{B}\boldsymbol{A}=\boldsymbol{O}\ \Rightarrow\ \boldsymbol{B}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{O}\cdot\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{O}$,而 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 是 m 阶方阵,且 $R(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})=R(\boldsymbol{A})=m$,故 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 是可逆阵,故

$$\boldsymbol{B}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\cdot(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}=\boldsymbol{O}\cdot(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}=\boldsymbol{O},\quad\text{即}\quad\boldsymbol{B}=\boldsymbol{O}.$$

或者：$\boldsymbol{AX}=\boldsymbol{b}$ 中，$R(\boldsymbol{A})=R(\boldsymbol{A},\boldsymbol{b})=m<n$，故必有无穷多组解.

评注　本题用到结论：$R(\boldsymbol{AA}^{\mathrm{T}})=R(\boldsymbol{A})$.

例 3-9　设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，则下列向量组中，线性无关的是　（　）

(A) $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$，$\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$，$\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_1$；

(B) $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$，$\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$，$\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$；

(C) $\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2$，$2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3$，$3\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1$；

(D) $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$，$2\boldsymbol{\alpha}_1-3\boldsymbol{\alpha}_2+22\boldsymbol{\alpha}_3$，$3\boldsymbol{\alpha}_1+5\boldsymbol{\alpha}_2-5\boldsymbol{\alpha}_3$.

解　选(C).

设有 k_1,k_2,k_3，使 $k_1(\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2)+k_2(2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3)+k_3(3\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1)=0$，即

$$(k_1+k_3)\boldsymbol{\alpha}_1+2(k_1+k_2)\boldsymbol{\alpha}_2+3(k_2+k_3)\boldsymbol{\alpha}_3=0.$$

因 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，故

$$\begin{cases}k_1+k_3=0\\2(k_1+k_2)=0\\3(k_2+k_3)=0\end{cases}\Rightarrow\quad k_1=k_2=k_3=0.$$

所以，$\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2,2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3,3\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1$ 线性无关.

(A) 不正确，是因为$(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)-(\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3)+(\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_1)=0$；

(B) 不正确，是因为$(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)+(\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3)-(\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3)=0$；

(D) 不正确：设有数 k_1,k_2,k_3，使

$$k_1(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3)+k_2(2\boldsymbol{\alpha}_1-3\boldsymbol{\alpha}_2+22\boldsymbol{\alpha}_3)+k_3(3\boldsymbol{\alpha}_1+5\boldsymbol{\alpha}_2-5\boldsymbol{\alpha}_3)=0$$

即

$$\begin{cases}k_1+2k_2+3k_3=0\\k_1-3k_2+5k_3=0\\k_1+22k_2-5k_3=0\end{cases}$$

此系数矩阵行列式：

$$\begin{vmatrix}1&2&3\\1&-3&5\\1&22&-5\end{vmatrix}=\begin{vmatrix}1&2&3\\0&-5&2\\0&20&-8\end{vmatrix}=0,$$

故 k_1,k_2,k_3 有非零解，即(D) 这一组线性相关.

例 3-10　设向量 $\boldsymbol{\beta}$ 可由向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性表示，但不能由向量组(Ⅰ)：$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{m-1}$ 线性表示. 记向量组(Ⅱ)：$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{m-1},\boldsymbol{\beta}$，则　（　）

(A) $\boldsymbol{\alpha}_m$ 不能由(Ⅰ)线性表示，也不能由(Ⅱ)线性表示；

(B) $\boldsymbol{\alpha}_m$ 不能由(Ⅰ)线性表示，但可由(Ⅱ)线性表示；

(C) $\boldsymbol{\alpha}_m$ 可由(Ⅰ)线性表示，也可由(Ⅱ)线性表示；

(D) $\boldsymbol{\alpha}_m$ 可由(Ⅰ)线性表示，但不可由(Ⅱ)线性表示.

解　选(B).

由已知：$\boldsymbol{\beta}=\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_{m-1}\boldsymbol{\alpha}_{m-1}+\lambda_m\boldsymbol{\alpha}_m$，这里 $\lambda_m\neq0$，故 $\boldsymbol{\alpha}_m=\dfrac{1}{\lambda_m}(\boldsymbol{\beta}-\lambda_1\boldsymbol{\alpha}_1-\cdots-\lambda_{m-1}\boldsymbol{\alpha}_{m-1})$，即 $\boldsymbol{\alpha}_m$ 可由(Ⅱ)线性表示，故知(A)，(D) 不对. 再看(C)，若 $\boldsymbol{\alpha}_m$ 可由(Ⅰ)线性表示，则可推得 $\boldsymbol{\beta}$ 可由(Ⅰ)线性表示，这与已知条件矛盾，故(C) 也不对.

例 3-11　设向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{O}$ 的一个基础解系，向量 $\boldsymbol{\beta}$ 不是方程组 $\boldsymbol{AX}=\boldsymbol{O}$ 的解，即 $\boldsymbol{A\beta}\neq\boldsymbol{O}$. 试证明：向量组 $\boldsymbol{\beta},\boldsymbol{\beta}+\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\beta}+\boldsymbol{\alpha}_t$ 线性无关.

证　设有一组数 $k,k_1,\cdots,k_t$，使

$$k\boldsymbol{\beta}+k_1(\boldsymbol{\beta}+\boldsymbol{\alpha}_1)+\cdots+k_t(\boldsymbol{\beta}+\boldsymbol{\alpha}_t)=0$$

即
$$(k+k_1+\cdots+k_t)\boldsymbol{\beta}=-(k_1\boldsymbol{\alpha}_1+\cdots+k_t\boldsymbol{\alpha}_t) \tag{1}$$

对式(1)两边左乘 $\boldsymbol{A}$,得
$$(k+k_1+\cdots+k_t)\boldsymbol{A\beta}=0$$

由 $\boldsymbol{A\beta}\neq\boldsymbol{O}$. $\Rightarrow$ $k+k_1+\cdots+k_t=0$ (2)

$\Rightarrow$ $-k_1\boldsymbol{\alpha}_1-\cdots-k_t\boldsymbol{\alpha}_t=0$

由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 线性无关, $\Rightarrow$ $k_1=\cdots=k_t=0$

再由式(2),可知 $k=0$.

所以,向量组 $\boldsymbol{\beta},\boldsymbol{\beta}+\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\beta}+\boldsymbol{\alpha}_t$ 线性无关.

例 3-12 设

$$\begin{aligned}\boldsymbol{\beta}_1&=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\cdots+\boldsymbol{\alpha}_t\\ \boldsymbol{\beta}_2&=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_3+\cdots+\boldsymbol{\alpha}_t\\ &\vdots\\ \boldsymbol{\beta}_t&=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_{t-1}\end{aligned}\qquad(t>1)$$

试证:秩$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t\}$ = 秩$\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t\}$.

证 已知 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t$ 可由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 线性表示,即

$$(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t)=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t)\begin{pmatrix}0&1&1&\cdots&1\\1&0&1&\cdots&1\\1&1&0&\cdots&1\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&1&1&\cdots&0\end{pmatrix}$$

$$|\boldsymbol{K}|=\begin{vmatrix}0&1&1&\cdots&1\\1&0&1&\cdots&1\\1&1&0&\cdots&1\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&1&1&\cdots&0\end{vmatrix}=(t-1)\cdot\begin{vmatrix}1&1&1&\cdots&1\\1&0&1&\cdots&1\\1&1&0&\cdots&1\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&1&1&\cdots&0\end{vmatrix}$$

$$=(t-1)\cdot\begin{vmatrix}1&1&1&\cdots&1\\0&-1&0&\cdots&0\\0&0&-1&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&-1\end{vmatrix}=(t-1)(-1)^{t-1}\neq0$$

故矩阵 $\boldsymbol{K}$ 可逆,则有

$$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t)=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t)\boldsymbol{K}^{-1}$$

即 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t$ 可由 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t$ 线性表示. 因此这两个向量组等价. 即有秩$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_t\}$ = 秩$\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t\}$.

例 3-13 试证:由

$$\boldsymbol{\alpha}_1=\begin{pmatrix}0\\1\\1\end{pmatrix},\quad\boldsymbol{\alpha}_2=\begin{pmatrix}1\\0\\1\end{pmatrix},\quad\boldsymbol{\alpha}_3=\begin{pmatrix}1\\1\\0\end{pmatrix}$$

所生成的空间 $L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$ 就是 $\boldsymbol{R}^3$.

证 $L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\subseteq\boldsymbol{R}^3$. 显然,再证明 $\boldsymbol{R}^3\subseteq L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$. 由于 $|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|\neq0$,知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关.

$\forall\, y \in \boldsymbol{R}^3$，由于 $y,\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，故 y 由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，即 $y \in L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$，则 $\boldsymbol{R}^3 \subseteq L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$.

例 3-14　若

$$\boldsymbol{\alpha}_1 = \begin{pmatrix}1\\1\\0\\0\end{pmatrix},\quad \boldsymbol{\alpha}_2 = \begin{pmatrix}1\\0\\1\\1\end{pmatrix},\quad \boldsymbol{\beta}_1 = \begin{pmatrix}2\\-1\\3\\3\end{pmatrix},\quad \boldsymbol{\beta}_2 = \begin{pmatrix}0\\1\\-1\\-1\end{pmatrix}.$$

试证：$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) = L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)$.

证　要证 $L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) = L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)$，只要证向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$ 与 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\}$ 等价.

向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\}$ 的秩为 2，这可从

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\} = \begin{pmatrix}1&1&2&0\\1&0&-1&1\\0&1&3&-1\\0&1&3&-1\end{pmatrix} \to \cdots \to \begin{pmatrix}1&1&2&0\\0&-1&-3&1\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

容易得知. 而 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$ 与 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\}$ 都是秩为 2 的向量组，因此它们都是 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\}$ 的极大线性无关组，因此等价. 即有 $L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) = L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)$.

评注　本题的方法比按定义证明 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$ 与 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\}$ 等价要方便，原因在于本题的向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\}$ 的秩必为 2（想一想，为什么？）.

例 3-15　设

$$\boldsymbol{B} = \begin{pmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_3\\ \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_3\\ \boldsymbol{\alpha}_3^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_3^{\mathrm{T}}\boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3^{\mathrm{T}}\boldsymbol{\alpha}_3\end{pmatrix}$$

其中，$\boldsymbol{\alpha}_i (i = 1,2,3)$ 是三维非零列向量. 试证：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关 $\Leftrightarrow |\boldsymbol{B}| = 0$.

证　因为

$$\boldsymbol{B} = \begin{pmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\ \boldsymbol{\alpha}_2^{\mathrm{T}}\\ \boldsymbol{\alpha}_3^{\mathrm{T}}\end{pmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$$

所以　　$$|\boldsymbol{B}| = |\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|^2$$

故　　$$|\boldsymbol{B}| = 0 \quad\Leftrightarrow\quad |\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3| = 0$$

$$\Leftrightarrow\quad \boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3 \text{ 线性相关.}$$

评注　本题关键是知道矩阵 $\boldsymbol{B}$ 是由两个矩阵乘积而得，也称为“积分解”. 对于一些比较简单的积分解，应该熟悉.

例 3-16　已知向量组(Ⅰ)：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$；(Ⅱ)：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$；(Ⅲ)：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_5$. 如果各向量组的秩分别为 $R(\text{Ⅰ}) = R(\text{Ⅱ}) = 3, R(\text{Ⅲ}) = 4$. 试证：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_5 - \boldsymbol{\alpha}_4$ 的秩为 4.

证　由 $R(\text{Ⅰ}) = R(\text{Ⅱ}) = 3$，知(Ⅰ)线性无关，(Ⅱ)线性相关，且 $\boldsymbol{\alpha}_4 = \lambda_1\boldsymbol{\alpha}_1 + \lambda_2\boldsymbol{\alpha}_2 + \lambda_3\boldsymbol{\alpha}_3$.

现考察　　$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + k_3\boldsymbol{\alpha}_3 + k_4(\boldsymbol{\alpha}_5 - \boldsymbol{\alpha}_4) = 0$$

即　　$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + k_3\boldsymbol{\alpha}_3 + k_4(\boldsymbol{\alpha}_5 - \lambda_1\boldsymbol{\alpha}_1 - \lambda_2\boldsymbol{\alpha}_2 - \lambda_3\boldsymbol{\alpha}_3) = 0$$

即　　$$(k_1 - \lambda_1 k_4)\boldsymbol{\alpha}_1 + (k_2 - \lambda_2 k_4)\boldsymbol{\alpha}_2 + (k_3 - \lambda_3 k_4)\boldsymbol{\alpha}_3 + k_4\boldsymbol{\alpha}_5 = 0$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_5$ 线性无关，故

$$\begin{cases} k_1 - \lambda_1 k_4 = 0 \\ k_2 - \lambda_2 k_4 = 0 \\ k_3 - \lambda_3 k_4 = 0 \\ k_4 = 0 \end{cases} \Rightarrow \quad k_1 = k_2 = k_3 = k_4 = 0,$$

所以 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_5 - \boldsymbol{\alpha}_4$ 线性无关，其秩为 4.

例 3-17 $\boldsymbol{A}$ 是 n 阶方阵，$\boldsymbol{X}$ 是 n 维非零列向量. 若 $\boldsymbol{A}^{k-1}\boldsymbol{X} \neq \boldsymbol{O}, \boldsymbol{A}^k\boldsymbol{X} = \boldsymbol{O}(k \in \mathbf{Z}^+)$. 试证：$\boldsymbol{X}, \boldsymbol{AX}, \boldsymbol{A}^2\boldsymbol{X}, \cdots, \boldsymbol{A}^{k-1}\boldsymbol{X}$ 线性无关.

证 考察 $\lambda_1\boldsymbol{X} + \lambda_2\boldsymbol{AX} + \lambda_3\boldsymbol{A}^2\boldsymbol{X} + \cdots + \lambda_k\boldsymbol{A}^{k-1}\boldsymbol{X} = \boldsymbol{O}$. 在等式两边左乘 $\boldsymbol{A}^{k-1}$，得 $\lambda_1\boldsymbol{A}^{k-1}\boldsymbol{X} = 0$. 由于 $\boldsymbol{A}^{k-1}\boldsymbol{X} \neq 0$，故 $\lambda_1 = 0$. 再左乘 $\boldsymbol{A}^{k-2}$，可得 $\lambda_2 = 0, \cdots$，类似可得 $\lambda_k = 0$. 所以，$\boldsymbol{X}, \boldsymbol{AX}, \boldsymbol{A}^2\boldsymbol{X}, \cdots, \boldsymbol{A}^{k-1}\boldsymbol{X}$ 线性无关.

评注 注意到 $\boldsymbol{AX}, \boldsymbol{A}^2\boldsymbol{X}, \cdots, \boldsymbol{A}^{k-1}\boldsymbol{X}$ 仍是 n 维列向量.

例 3-18 设 $s \times n$ 矩阵 $\boldsymbol{A}$ 是列线性无关的，$n \times s$ 矩阵 $\boldsymbol{B}, \boldsymbol{C}$ 满足 $\boldsymbol{BA} = \boldsymbol{E}, \boldsymbol{CA} = \boldsymbol{E}$，则矩阵 $\boldsymbol{B} - \boldsymbol{C}$ 的秩不大于 $s - n$.

证 已知

$$R(\boldsymbol{A}) = n$$

以及

$$(\boldsymbol{B} - \boldsymbol{C})\boldsymbol{A} = 0$$

推得

$$R(\boldsymbol{B} - \boldsymbol{C}) + R(\boldsymbol{A}) \leqslant s$$

即

$$R(\boldsymbol{B} - \boldsymbol{C}) \leqslant s - n.$$

例 3-19 设给定向量组

$$\boldsymbol{x}_i = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}, \quad (i = 1, 2, \cdots, s,\ s \leqslant n).$$

如果

$$|a_{ii}| > \sum_{\substack{i=1 \\ i \neq j}}^{s} |a_{ij}|, \quad (j = 1, 2, \cdots, s).$$

那么 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_s$ 一定是线性无关向量组.

证(反证法) 若 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_s$ 线性相关，则有

$$\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 + \cdots + \lambda_s\boldsymbol{x}_s = 0 \tag{1}$$

其中 λ_i 不全为零.

令

$$\max_{1 \leqslant i \leqslant s}\{|\lambda_i|\} = |\lambda_p|, \text{则 } |\lambda_p| \neq 0.$$

现看方程组(1)中第 p 个方程：

$$\lambda_1 a_{1p} + \lambda_2 \cdot a_{2p} + \cdots + \lambda_p a_{pp} + \cdots + \lambda_s a_{sp} = 0$$

移项

$$\lambda_p a_{pp} = -\lambda_1 a_{1p} - \lambda_2 a_{2p} - \cdots - \lambda_{p-1} a_{p-1\,p} - \lambda_{p+1} a_{p+1\,p} - \cdots - \lambda_s a_{sp},$$

即有

$$|\lambda_p a_{pp}| = |-\lambda_1 a_{1p} - \cdots - \lambda_s a_{sp}| \leqslant |\lambda_p| \cdot \sum_{\substack{i=1 \\ i \neq p}}^{s} |a_{ip}|.$$

即

$$|a_{pp}| \leqslant \sum_{\substack{i=1 \\ i \neq p}}^{s} |a_{ip}|,$$

这与已知 $|a_{pp}| > \sum\limits_{\substack{i=1 \\ i \neq p}}^{s} |a_{pp}|$ 矛盾.

例 3-20　设 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 是一组线性无关向量. $\boldsymbol{\beta}_i=\sum\limits_{j=1}^{r}a_{ij}\boldsymbol{\alpha}_j,i=1,2,\cdots,r$. 试证:$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 线性无关 $\Leftrightarrow\begin{pmatrix}a_{11}&\cdots&a_{1r}\\ \vdots&\ddots&\vdots\\ a_{r1}&\cdots&a_{rr}\end{pmatrix}\neq 0$.

证　考察
$$\sum_{i=1}^{r}k_i\boldsymbol{\beta}_i=0,$$
即
$$\sum_{i=1}^{r}k_i\sum_{j=1}^{r}a_{ij}\boldsymbol{\alpha}_j=0,\quad \text{即}\quad \sum_{j=1}^{r}\Big(\sum_{i=1}^{r}k_ia_{ij}\Big)\boldsymbol{\alpha}_j=0.$$
由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性无关,知 $\sum\limits_{i=1}^{r}k_ia_{ij}=0,j=1,\cdots,r$. 对于此齐次线性方程组,把 k_i 看成未知量,则
$$k_1,\cdots,k_r\text{ 只有零解}\quad\Leftrightarrow\quad\begin{vmatrix}a_{11}&\cdots&a_{1r}\\ \vdots&\ddots&\vdots\\ a_{r1}&\cdots&a_{rr}\end{vmatrix}\neq 0,$$
而 $k_1,\cdots,k_r$ 只有零解 $\Leftrightarrow\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 线性无关.

证法 2　已知条件即
$$(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)\cdot\boldsymbol{A}_{r\times r}$$

必要性　当 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 线性无关时,$R(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)=r$. 而 $R(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)\leqslant R(\boldsymbol{A})$. 即有 $R(\boldsymbol{A})\geqslant r$. 由于 $\boldsymbol{A}$ 是 r 阶方阵,故 $R(\boldsymbol{A})=r$,知 $\boldsymbol{A}$ 可逆,即 $|\boldsymbol{A}|\neq 0$.

充分性　当 $|\boldsymbol{A}|\neq 0$,即 $\boldsymbol{A}$ 可逆,则 $\boldsymbol{A}$ 乘了 $(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)$ 后,不改变 $(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)$ 的秩. 由于秩 $(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)=r$,故秩 $(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)=r$. 即知 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 线性无关.

评注　(1) 注意线性方程组也可用 $\sum$ 的形式表示:$\sum\limits_{i=1}^{r}k_ia_{ij}=0,j=1,\cdots,r$.

(2) 在证法 2 中,用到了向量组的线性表示与矩阵乘法的关系,也用到了矩阵乘法的秩的关系式:秩$(\boldsymbol{AB})\leqslant\min\{$秩$(\boldsymbol{A})$,秩$(\boldsymbol{B})\}$,以及可逆矩阵乘上另一矩阵后,不改变那个矩阵的秩,这些有用的结论应学会使用.

例 3-21　若向量组 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\}$ 可由 $\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s\}$ 线性表示,且 $r>s$,则 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性相关.

证　已知
$$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s)\cdot\boldsymbol{K}_{s\times r}\tag{1}$$
记 $\boldsymbol{K}_{s\times r}=(\boldsymbol{k}_1,\cdots,\boldsymbol{k}_r),\boldsymbol{k}_i(i=1,\cdots,r)$ 是 k 的列. 由于 $r>s$,即向量个数大于向量维数,则 $\boldsymbol{k}_1,\cdots,\boldsymbol{k}_r$ 线性相关,即存在一组不全为零的数 $\lambda_1,\cdots,\lambda_r$,使
$$(\boldsymbol{k}_1,\cdots,\boldsymbol{k}_r)\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_r\end{pmatrix}=0,$$
则在式(1) 两边右乘 $\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_r\end{pmatrix}$,得
$$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_r\end{pmatrix}=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)\cdot\boldsymbol{K}_{s\times r}\cdot\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_r\end{pmatrix}=0.$$
即 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性相关.

评注　(1) 例 3-23 的证明中,把矩阵 $\boldsymbol{K}_{s\times r}$ 按列分块后,使得证明简洁、方便,而把 $\boldsymbol{K}$ 按列分块是基于对命题:"向量个数大于向量维数时,必线性相关" 非常熟悉.

(2) 例 3-23 运用了"向量组的秩" 这一部分的重要定理,推论 1 是它的等价说法,并由此有一系列的推论. 推论 1 说明线性无关向量组决不能用向量个数比它少的向量组线性表示.

例 3-22 向量组 Ⅰ 与 Ⅱ，若 $R(\text{Ⅰ})=R(\text{Ⅱ})$，且 Ⅰ 可以由 Ⅱ 线性表示，则 Ⅰ 与 Ⅱ 等价.

证法 1 设 $R(\text{Ⅰ})=R(\text{Ⅱ})=r$，并设 Ⅰ_0，Ⅱ_0 分别是 Ⅰ，Ⅱ 的最大线性无关组：

$$\text{Ⅰ}_0:\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\},\quad \text{Ⅱ}_0:\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r\}.$$

已知 Ⅰ 可由 Ⅱ 线性表示，则有 Ⅰ_0 可由 Ⅱ_0 线性表示，则向量组 $\{\text{Ⅰ}_0,\boldsymbol{\beta}_j\}$ 可由 Ⅱ_0 线性表示（这里，$\boldsymbol{\beta}_j$ 是 Ⅱ_0 中的任一向量）. 由定理（即例 3-23）可知：向量组 $\{\text{Ⅰ}_0,\boldsymbol{\beta}_j\}$ 线性相关，而 Ⅰ_0 是最大无关组，则 $\boldsymbol{\beta}_j$ 可由 Ⅰ_0 线性表示，即有 Ⅱ_0 可由 Ⅰ_0 线性表示. 也即有 Ⅱ 可由 Ⅰ 线性表示，又已知 Ⅰ 可由 Ⅱ 线性表示，则 Ⅰ 与 Ⅱ 等价.

证法 2 由 Ⅰ_0 可由 Ⅱ_0 线性表示，即有矩阵乘法 $\text{Ⅰ}_0=\text{Ⅱ}_0\cdot\boldsymbol{K}_{r\times r}$，则 $R(\text{Ⅰ}_0)=r\leqslant R(\boldsymbol{K})$. 则 $R(\boldsymbol{K})=r$，则 $\boldsymbol{K}$ 可逆，则 $\text{Ⅱ}_0=\text{Ⅰ}_0\cdot\boldsymbol{K}^{-1}$. 即 Ⅱ_0 可由 Ⅰ_0 线性表示，即有 Ⅱ 可由 Ⅰ 线性表示，即得 Ⅱ 与 Ⅰ 等价.

评注 (1) 在证明中可看到最大线性无关组的作用. 一般，对于抽象的向量组，总要考虑它的最大线性无关组.

(2) 证法 1 中，关键是构造出向量组 $\{\text{Ⅰ}_0,\boldsymbol{\beta}_j\}$. 请想一想，怎么会想到构造这样的向量组.

证法 2 中用了向量组的线性表示与矩阵乘法的关系以及矩阵乘法的秩的关系式，比较简洁.

例 3-23 若 $\boldsymbol{A}_{m\times n}\cdot\boldsymbol{B}_{n\times s}=\boldsymbol{C}_{m\times s}$，则 $R(\boldsymbol{C})\leqslant\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}$.

证 由 $\boldsymbol{A}\cdot\boldsymbol{B}=\boldsymbol{C}$.

$\Rightarrow$ $\boldsymbol{C}$ 的列组可由 $\boldsymbol{A}$ 的列组线性表示

$$\Rightarrow\quad R(\boldsymbol{C})\leqslant R(\boldsymbol{A})\tag{1}$$

又由 $\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$，同理可得 $R(\boldsymbol{C}^{\mathrm{T}})\leqslant R(\boldsymbol{B}^{\mathrm{T}})$，即

$$R(\boldsymbol{C})\leqslant R(\boldsymbol{B})\tag{2}$$

由式(1)与式(2)，即得

$$R(\boldsymbol{C})\leqslant\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}.$$

例 3-24 若 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{C}$，则 $R(\boldsymbol{C})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B})$.

证法 1 记 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)$，$\boldsymbol{B}=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n)$，即把 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 按列分块. 则

$$\boldsymbol{C}=(\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1,\cdots,\boldsymbol{\alpha}_n+\boldsymbol{\beta}_n),$$

可知 $\boldsymbol{C}$ 的列组可由组 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n\}$ 线性表示.

则

$$\begin{aligned}\text{秩}(\boldsymbol{C})&\leqslant\text{秩}\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n\}\\&\leqslant\text{秩}\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}+\text{秩}\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n\}\\&=R(\boldsymbol{A})+R(\boldsymbol{B}).\end{aligned}$$

证法 2

$$R(\boldsymbol{A}+\boldsymbol{B})\leqslant R(\boldsymbol{A}+\boldsymbol{B},\boldsymbol{B})=R(\boldsymbol{A},\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B}).$$

例 3-25 设 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times s}=\boldsymbol{C}$，则 $R(\boldsymbol{A})+R(\boldsymbol{B})\leqslant n+R(\boldsymbol{AB})$.

证

$$R(\boldsymbol{A})+R(\boldsymbol{B})=R\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}\leqslant R\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\-\boldsymbol{E}&\boldsymbol{B}\end{pmatrix}=R\begin{pmatrix}\boldsymbol{O}&\boldsymbol{AB}\\-\boldsymbol{E}&\boldsymbol{B}\end{pmatrix}$$

$$=R\begin{pmatrix}\boldsymbol{O}&\boldsymbol{AB}\\\boldsymbol{E}&\boldsymbol{O}\end{pmatrix}=R(\boldsymbol{E})+R(\boldsymbol{AB})=n+R(\boldsymbol{AB}).$$

例 3-26 设 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times s}=\boldsymbol{C}$，则

(1) 若 $R(\boldsymbol{A})=n$，则 $R(\boldsymbol{B})=R(\boldsymbol{C})$；

(2) 若 $R(\boldsymbol{B})=n$,则 $R(\boldsymbol{A})=R(\boldsymbol{C})$.

证法 1　(1) 与(2) 证法类似,只证(2).

由
$$\boldsymbol{AB}=\boldsymbol{C}\quad\Rightarrow\quad R(\boldsymbol{C})\leqslant R(\boldsymbol{A}).$$

又由 $\boldsymbol{ABB}^{\mathrm{T}}=\boldsymbol{CB}^{\mathrm{T}}$,而 $R(\boldsymbol{BB}^{\mathrm{T}})=R(\boldsymbol{B})=n$,$\boldsymbol{BB}^{\mathrm{T}}$ 又是 n 阶方阵,故知 $\boldsymbol{BB}^{\mathrm{T}}$ 可逆,则有
$$\boldsymbol{A}=\boldsymbol{CB}^{\mathrm{T}}\cdot(\boldsymbol{BB}^{\mathrm{T}})^{-1},\quad\Rightarrow\quad R(\boldsymbol{A})\leqslant R(\boldsymbol{C}).$$

于是有
$$R(\boldsymbol{A})=R(\boldsymbol{C}).$$

证法 2　由
$$\boldsymbol{AB}=\boldsymbol{C}\quad\Rightarrow\quad R(\boldsymbol{C})\leqslant R(\boldsymbol{A}).$$

又由
$$R(\boldsymbol{B})=n\quad\Rightarrow\quad R(\boldsymbol{B},\boldsymbol{E}_n)=n=R(\boldsymbol{B})$$
$$\Rightarrow\quad \text{矩阵方程 } \boldsymbol{BX}=\boldsymbol{E}_n \text{ 有解. 设解为 } \boldsymbol{Q}_{s\times n},\text{即 } \boldsymbol{BQ}=\boldsymbol{E}_n.$$

在 $\boldsymbol{AB}=\boldsymbol{C}$ 等号两旁右乘 $\boldsymbol{Q}$,得 $\boldsymbol{ABQ}=\boldsymbol{CQ}$,即 $\boldsymbol{A}=\boldsymbol{CQ}\Rightarrow R(\boldsymbol{A})\leqslant R(\boldsymbol{C})$. 于是有 $R(\boldsymbol{A})=R(\boldsymbol{C})$.

评注　若利用例 3-27 的结论,可以很容易证明例 3-28,也可容易证得"若 $\boldsymbol{AB}=\boldsymbol{O}$,则 $R(\boldsymbol{A})+R(\boldsymbol{B})\leqslant n$"的结论(这里 n 是 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times s}$ 中的 n).

例 3-27　设 $\boldsymbol{A}_{n\times m}\boldsymbol{B}_{m\times n}=\boldsymbol{E}_n$,$(n<m)$. 试证:$\boldsymbol{B}$ 的列向量组线性无关.

证法 1　由 $\boldsymbol{AB}=\boldsymbol{E}_n\quad\Rightarrow\quad R(\boldsymbol{E})=n\leqslant R(\boldsymbol{B})\quad\Rightarrow\quad R(\boldsymbol{B})=n\quad\Rightarrow\quad\boldsymbol{B}$ 的 n 个列向量线性无关.

证法 2　记 $\boldsymbol{B}=(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)$,即把 $\boldsymbol{B}$ 按列分块. 考察 $x_1\boldsymbol{B}_1+\cdots+x_n\boldsymbol{B}_n=\boldsymbol{O}$,即
$$(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\boldsymbol{O}$$

在上式两边左乘 $\boldsymbol{A}$:
$$\boldsymbol{ABX}=\boldsymbol{A}\cdot\boldsymbol{O}$$

即 $\boldsymbol{X}=\boldsymbol{O}$,即 $x_1=x_2=\cdots=x_n=\boldsymbol{O}$,知 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n$ 线性无关.

证法 3　由 $\boldsymbol{AB}=\boldsymbol{E}_n$,可知,$\boldsymbol{E}_n$ 的行组可由 $\boldsymbol{B}$ 的行组线性表示. 又由 $\boldsymbol{BE}_n=\boldsymbol{B}$,可知:$\boldsymbol{B}$ 的行组可由 $\boldsymbol{E}_n$ 的行组线性表示. 就得
$$R(\boldsymbol{E}_n\text{ 的行组})=R(\boldsymbol{B}\text{ 的行组}),$$

即
$$R(\boldsymbol{E}_n)=R(\boldsymbol{B}).$$

则 $R(\boldsymbol{B})=n\quad\Rightarrow\quad\boldsymbol{B}$ 的 n 个列向量线性无关.

例 3-28　设 n 维列向量组 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s\}$ 与 $\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t\}$ 皆线性无关,且 $s>t$. 试证:必存在 $\boldsymbol{\alpha}_i(1\leqslant i\leqslant s)$,使得 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t,\boldsymbol{\alpha}_i$ 线性无关.

证　反证法　若不存在这样的 $\boldsymbol{\alpha}_i(1\leqslant i\leqslant s)$,即 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s\}$ 中的任一个 $\boldsymbol{\alpha}_i$,都有 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t,\boldsymbol{\alpha}_i$ 线性相关,则得 $\boldsymbol{\alpha}_i$ 可由 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t$ 线性表示,即有 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s\}$ 可由 $\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t\}$ 线性表示,又已知 $s>t$,则由定理(即例 3-23) 知:$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 线性相关,这与已知矛盾.

例 3-29　n 维列向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 线性无关$(m<n)$,矩阵 $\boldsymbol{K}=(k_{ij})_{m\times n}$,$\boldsymbol{\beta}_j=\sum\limits_{i=1}^{m}k_{ij}\boldsymbol{\alpha}_i$,$j=1,2,\cdots,n$. 试证:秩$\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n\}=$ 秩$(\boldsymbol{K})$.

证　已知
$$\boldsymbol{\beta}_j=\sum_{i=1}^{m}k_{ij}\boldsymbol{\alpha}_i,j=1,2,\cdots,n.$$

即
$$(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m)\boldsymbol{K}_{m\times n}$$

记为
$$\boldsymbol{B}_{n\times n}=\boldsymbol{A}_{n\times m}\boldsymbol{K}_{m\times n}$$

即知 $\boldsymbol{B}$ 的行组可由 $\boldsymbol{K}$ 的行组线性表示.

又由于 $R(\boldsymbol{A})=m$,而 $R(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=R(\boldsymbol{A})=m$,且 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 m 阶方阵,故 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 可逆,即有

$$\boldsymbol{K}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\cdot\boldsymbol{B}$$

此式表明：$\boldsymbol{K}$ 的行组可由 $\boldsymbol{B}$ 的行组线性表示. 这样就有 $\boldsymbol{K}$ 的行组与 $\boldsymbol{B}$ 的行组等价. 即有

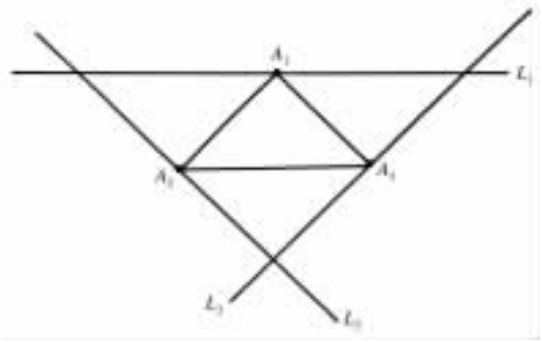
$$\text{秩}(\boldsymbol{K}\text{ 的行组})=\text{秩}(\boldsymbol{B}\text{ 的行组})$$

即
$$R(\boldsymbol{B})=R(\boldsymbol{K})$$

即有
$$\text{秩}\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n\}=\text{秩}(\boldsymbol{K}).$$

评注 本例与例 3-22 相像，结论更一般化了.

例 3-30 设矩阵 $\begin{pmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{pmatrix}$ 是满秩阵，$(a_i,b_i,c_i)\,i=1,2,3$，是空间的三点，则直线：



例 3-30 图

$$L_1:\frac{x-a_1}{a_2-a_3}=\frac{y-b_1}{b_2-b_3}=\frac{z-c_1}{c_2-c_3};$$

$$L_2:\frac{x-a_2}{a_3-a_1}=\frac{y-b_2}{b_3-b_1}=\frac{z-c_2}{c_3-c_1};$$

$$L_3:\frac{x-a_3}{a_1-a_2}=\frac{y-b_3}{b_1-b_2}=\frac{z-c_3}{c_1-c_2}.\qquad(\quad)$$

(A) 相交于一点；　(B) 两两相交；

(C) 平行但不重合；　(D) 重合.

解 由 $|\boldsymbol{A}|\neq 0\ \Rightarrow$ 三点 $(a_i,b_i,c_i),i=1,2,3$ 所成的矢径不共面，$\Rightarrow$ 此三点不共线. 由三条直线方程(点向式直线方程)，可知必为如例 3-30 图示，所以，(B) 正确.

例 3-31 设

$$\boldsymbol{\alpha}_1=\begin{pmatrix}a_1\\ a_2\\ a_3\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}b_1\\ b_2\\ b_3\end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix}c_1\\ c_2\\ c_3\end{pmatrix},$$

则三条 xOy 平面上的直线：

$$a_ix+b_iy+c_i=0\quad(i=1,2,3)$$

($\text{其中 }a_i^2+b_i^2\neq 0,i=1,2,3$) 相交于一点的充分必要条件是

(A) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关；　(B) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关；

(C) $R(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=R(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)$，而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 线性无关；　(D) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关.

解 三条直线

$$x\boldsymbol{\alpha}_1+y\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3=0\qquad(3)$$

相交于一点，即方程组(3) 有惟一解，而方程组(3) 有惟一解 $\Leftrightarrow R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})=$ 未知量个数(2)，可知(D) 是正确的.

例 3-32 设有向量组(Ⅰ)：

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\ 0\\ 2\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}1\\ 1\\ 3\end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix}1\\ -1\\ a+2\end{pmatrix}$$

和向量组(Ⅱ)：

$$\boldsymbol{\beta}_1=\begin{pmatrix}1\\2\\a+3\end{pmatrix},\quad \boldsymbol{\beta}_2=\begin{pmatrix}2\\1\\a+6\end{pmatrix},\quad \boldsymbol{\beta}_3=\begin{pmatrix}2\\1\\a+4\end{pmatrix}.$$

试问：当 a 为何值时，向量组(Ⅰ)与(Ⅱ)等价？当 a 为何值时，向量组(Ⅰ)与(Ⅱ)不等价？

解法 1　对向量组$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}$作初等行变换：

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}=\begin{pmatrix}1&1&1&1&2&2\\0&1&-1&2&1&1\\2&3&a+2&a+3&a+6&a+4\end{pmatrix}$$

$$\to\begin{pmatrix}1&0&2&-1&1&1\\0&1&-1&2&1&1\\0&0&a+1&a-1&a+1&a-1\end{pmatrix}.$$

(1) **当** $a\neq-1$ **时**，$|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|=a+1\neq0$. **故线性方程组** $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3=\boldsymbol{\beta}_i(i=1,2,3)$ **均有惟一解，即向量组(Ⅱ)可由(Ⅰ)线性表示. 同理，**$|\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3|=6\neq0$，**故(Ⅰ)也可由(Ⅱ)线性表示，即(Ⅰ)与(Ⅱ)等价.**

(2) **当** $a=-1$ **时，**

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)\to\begin{pmatrix}1&0&2&-1&1&1\\0&1&-1&2&1&1\\0&0&0&-2&0&-2\end{pmatrix}.$$

由于 $R(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\neq R(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)$. **即知线性方程组** $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3=\boldsymbol{\beta}_1$ **无解，即** $\boldsymbol{\beta}_1$ **不能由** $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ **线性表示，因此(Ⅰ)与(Ⅱ)不等价.**

解法 2

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)\xrightarrow{\text{行}}\begin{pmatrix}1&0&2&-1&1&1\\0&1&-1&2&1&1\\0&0&a+1&a-1&a+1&a-1\end{pmatrix}.$$

知向量组$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}$ **的秩为** 3. **当** $a\neq-1$ **时，易知** $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ **线性无关，即** $R(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=3$，**同时** $|\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3|=6$，**即** $R(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=3$. **因此它们都是向量组**$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}$ **的最大线性无关组，因此它们等价.**

$a=-1$ **时的讨论同解法** 1.

例 3-33　**已知** $\boldsymbol{R}^3$ **有两组基为**

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\\1\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}1\\0\\-1\end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix}1\\0\\1\end{pmatrix};\quad \boldsymbol{\beta}_1=\begin{pmatrix}1\\2\\1\end{pmatrix},\quad \boldsymbol{\beta}_2=\begin{pmatrix}2\\3\\4\end{pmatrix},\quad \boldsymbol{\beta}_3=\begin{pmatrix}3\\4\\3\end{pmatrix}.$$

求由基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ **过渡到基** $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ **的过渡矩阵.**

解　设　$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\cdot\boldsymbol{P}$$

则　$$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)^{-1}\cdot(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=\begin{pmatrix}1&1&1\\1&0&0\\1&-1&1\end{pmatrix}^{-1}\cdot\begin{pmatrix}1&2&3\\2&3&4\\1&4&3\end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 \\ 0 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix}.$$

例 3-34 已知 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 和 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_t$ 是两个线性无关的 n 维实向量组，并且每个 $\boldsymbol{\alpha}_i$ 与每个 $\boldsymbol{\beta}_j$ 都正交. 试证：$\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_t$ 线性无关.

证 考察

$$c_1\boldsymbol{\alpha}_1 + \cdots + c_s\boldsymbol{\alpha}_s + d_1\boldsymbol{\beta}_1 + \cdots + d_t\boldsymbol{\beta}_t = 0$$

令

$$\boldsymbol{\xi} = c_1\boldsymbol{\alpha}_1 + \cdots + c_s\boldsymbol{\alpha}_s = -d_1\boldsymbol{\beta}_1 - \cdots - d_t\boldsymbol{\beta}_t.$$

则内积

$$(\boldsymbol{\xi}, \boldsymbol{\xi}) = (c_1\boldsymbol{\alpha}_1 + \cdots + c_s\boldsymbol{\alpha}_s, -d_1\boldsymbol{\beta}_1 - \cdots - d_t\boldsymbol{\beta}_t)$$

$$= -\left(\sum_{i=1}^{s} c_i\boldsymbol{\alpha}_i, \sum_{j=1}^{t} d_j\boldsymbol{\beta}_j\right) = -\sum_{j=1}^{t}\sum_{i=1}^{s} c_i d_j(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_j)$$

由于

$$(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_j) = 0, (i = 1, 2, \cdots, s;\ j = 1, 2, \cdots, t).$$

所以

$$(\boldsymbol{\xi}, \boldsymbol{\xi}) = 0, \quad \text{则 } \boldsymbol{\xi} = 0.$$

则

$$c_1\boldsymbol{\alpha}_1 + \cdots + c_s\boldsymbol{\alpha}_s = 0, \quad d_1\boldsymbol{\beta}_1 + \cdots + d_t\boldsymbol{\beta}_t = 0.$$

由于这两组向量皆线性无关，故

$$c_1 = c_2 = \cdots = c_s = 0, \quad d_1 = d_2 = \cdots = d_t = 0.$$

即 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_t$ 线性无关.

例 3-35 设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 和 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_t$ 是两个线性无关的 n 维向量组. 试证：向量组 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_t$ 线性相关的充分必要条件是存在 n 维非零向量 $\boldsymbol{\xi}$，它既可由 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s$ 线性表示，又可由 $\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_t$ 线性表示.

证 “必要性”：设 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_t$ 线性相关，即存在不全为零的数 $k_1, k_2, \cdots, k_s, l_1, l_2, \cdots, l_t$，使得

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_s\boldsymbol{\alpha}_s + l_1\boldsymbol{\beta}_1 + l_2\boldsymbol{\beta}_2 + \cdots + l_t\boldsymbol{\beta}_t = 0.$$

令

$$\boldsymbol{\xi} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_s\boldsymbol{\alpha}_s = -l_1\boldsymbol{\beta}_1 - l_2\boldsymbol{\beta}_2 - \cdots - l_t\boldsymbol{\beta}_t,$$

可知 $\boldsymbol{\xi} \neq 0$，否则，若 $\boldsymbol{\xi} = 0$，则由这两组向量皆线性无关，可推知这些组合系数全为零，这与已知这些系数不全为零矛盾.

“充分性” 设非零向量 $\boldsymbol{\xi}$ 可以表示为

$$\boldsymbol{\xi} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_s\boldsymbol{\alpha}_s, \quad (k_i \text{ 不能全为零}).$$

及

$$\boldsymbol{\xi} = l_1\boldsymbol{\beta}_1 + l_2\boldsymbol{\beta}_2 + \cdots + l_t\boldsymbol{\alpha}_t, \quad (l_j \text{ 不能全为零}).$$

相减得

$$k_1\boldsymbol{\alpha}_1 + \cdots + k_s\boldsymbol{\alpha}_s - l_1\boldsymbol{\beta}_1 - \cdots - l_t\boldsymbol{\beta}_t = 0,$$

上式中的组合系数不全为零，故向量组 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_t$ 线性相关.

例 3-36 设 $\boldsymbol{A}$ 是 $s \times n$ 矩阵，$\boldsymbol{B}$ 是由 $\boldsymbol{A}$ 的前 m 行构成的 $m \times n$ 矩阵. 证明：若 $\boldsymbol{A}$ 的秩为 r，则 $R(\boldsymbol{B}) \geqslant r + m - s$.

证

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{C} \end{pmatrix} = \begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{O} \end{pmatrix} + \begin{pmatrix} \boldsymbol{O} \\ \boldsymbol{C} \end{pmatrix},$$

故有

$$R(\boldsymbol{A}) \leqslant R\begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{O} \end{pmatrix} + R\begin{pmatrix} \boldsymbol{O} \\ \boldsymbol{C} \end{pmatrix} = R(\boldsymbol{B}) + R(\boldsymbol{C}).$$

由于 $\boldsymbol{C}$ 是 $(s-m)\times n$ 型矩阵，故 $R(\boldsymbol{C})\leqslant s-m$

所以
$$R(\boldsymbol{B})\geqslant r+m-s.$$

例 3-37　在 $\boldsymbol{R}^4$ 中取两组基：

$$\boldsymbol{\varepsilon}_1=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad \boldsymbol{\varepsilon}_2=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\quad \boldsymbol{\varepsilon}_3=\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\quad \boldsymbol{\varepsilon}_4=\begin{pmatrix}0\\0\\0\\1\end{pmatrix};$$

$$\boldsymbol{\alpha}_1=\begin{pmatrix}2\\1\\-1\\1\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}0\\3\\1\\0\end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix}5\\3\\2\\1\end{pmatrix},\quad \boldsymbol{\alpha}_4=\begin{pmatrix}6\\6\\1\\3\end{pmatrix}.$$

（1）求由前一组基到后一组基的过渡矩阵；

（2）求向量 $(x_1,x_2,x_3,x_4)^{\mathrm{T}}$ 在后一组基下的坐标；

（3）求在两组基下有相同坐标的向量.

解　（1）
$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\cdot\boldsymbol{P}$$

则显然
$$\boldsymbol{P}=\begin{pmatrix}2&0&5&6\\1&3&3&6\\-1&1&2&1\\1&0&1&3\end{pmatrix}.$$

（2）
$$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=y_1\boldsymbol{\alpha}_1+y_2\boldsymbol{\alpha}_2+y_3\boldsymbol{\alpha}_3+y_4\boldsymbol{\alpha}_4=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}.$$

所以
$$\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)^{-1}\cdot\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\boldsymbol{P}^{-1}\cdot\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}$$

$$=\frac{1}{27}\begin{pmatrix}12&9&-27&-33\\1&12&-9&-23\\9&0&0&-18\\-7&-3&9&26\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}.$$

（3）设 $\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}$，即 $\boldsymbol{P}^{-1}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}$.

记 $\boldsymbol{X}=(x_1,x_2,x_3,x_4)^{\mathrm{T}}$. 即要求解 $\boldsymbol{PX}=\boldsymbol{X}$.

即 $(\boldsymbol{P}-\boldsymbol{E})\boldsymbol{X}=\boldsymbol{O}$. 求解此齐次线性方程组，得

$$\boldsymbol{X}=\begin{pmatrix} c \\ c \\ c \\ -c \end{pmatrix}=c\cdot\begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix},\quad c\text{ 为任意常数.}$$

例 3-38 设 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 是 R^n 的一组基,$\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n$ 是 R^n 的基本单位向量组(自然基). 又

$$(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\cdot\boldsymbol{A}.$$

试证:$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 是标准正交基的充要条件是 $\boldsymbol{A}$ 为正交矩阵.

证 设

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

由 $$(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\cdot\boldsymbol{A},$$

知 $$\boldsymbol{\eta}_i=a_{1i}\boldsymbol{\varepsilon}_1+a_{2i}\boldsymbol{\varepsilon}_2+\cdots+a_{ni}\boldsymbol{\varepsilon}_n.$$

于是
$$\begin{aligned}(\boldsymbol{\eta}_i,\boldsymbol{\eta}_j)&=(a_{1i}\boldsymbol{\varepsilon}_1+\cdots+a_{ni}\boldsymbol{\varepsilon}_n,\ a_{1j}\boldsymbol{\varepsilon}_1+\cdots+a_{nj}\boldsymbol{\varepsilon}_n)\\&=a_{1i}a_{1j}+a_{2i}a_{2j}+\cdots+a_{ni}a_{nj}\end{aligned}\tag{1}$$

若$(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n)$ 是标准正交基,则

$$(\boldsymbol{\eta}_i,\boldsymbol{\eta}_j)=\delta_{ij}=\begin{cases}1; & i=j,\\ 0; & i\neq j.\end{cases}\tag{2}$$

故 $$a_{1i}a_{1j}+a_{2i}a_{2j}+\cdots+a_{ni}a_{nj}=\delta_{ij}.\tag{3}$$

这就说明 $\boldsymbol{A}$ 的列向量组是两两正交的单位向量组,故 $\boldsymbol{A}$ 是正交矩阵.

反之,若 $\boldsymbol{A}$ 是正交矩阵,则有式(3),再有式(1) 与式(2),故 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 是标准正交基.

四、小 结

(1) 讨论向量组的线性相关性,若向量组是以具体数字形式给出的,则并不困难,难点在给出向量是抽象文字形式. 研究它们的相关性可直接按定义,考察 $\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_m\boldsymbol{\alpha}_m=0$ 时的组合系数 $\lambda_1,\cdots,\lambda_m$ 的情况,有时也用向量组的秩这个指标来判别相关性.

(2) 讨论向量组的相关性,不管用哪一种方法,一些命题及常用结论必须十分熟悉. 例如在按定义论证部分的七个命题及"向量组的秩" 部分的定理及 8 个推论.

(3) 两个向量组有线性表示关系,就可以有矩阵乘法,反之,若有矩阵乘法,也就有了向量组的线性表示关系,而有了矩阵乘法,就有矩阵秩之间的关系式,此概念也常用到.

(4) 最大线性无关组有重要地位,求最大线性无关组可有不同的方法. 一般,掌握一种方法即可.

(5) 有关"矩阵的秩" 的定理和结论都是经常用到的.

(6) 向量组的秩那一部分的定理(即基本概念与理论中的 5.(4)) 是向量组之间最基本的关系. 许多性质可由它导出,十分重要,也非常有用. 有关涉及到向量组之间的向量个数或秩之间的关系时,常可以试用此定理及推论.

第四章　线性方程组

线性方程组是线性代数的中心问题之一．许多问题的讨论都是围绕解线性方程组展开的，例如矩阵，向量组的线性相关性与秩的概念．其中许多问题已分散在前面各章解决，所以本章要讨论的问题是：(1) 齐次线性方程组有非零解的充要条件及解的结构；(2) 非齐次线性方程组有解的充要条件及解的结构；(3) 利用行初等变换解线性方程组．

一、复习与考试要求

(1) 理解齐次线性方程组有非零解的充要条件．

(2) 理解非齐次线性方程组有解的充要条件．

(3) 理解齐次线性方程组解的性质、基础解系、通解、解的结构及解空间的概念．

(4) 理解非齐次线性方程组解的性质，通解的概念及解的结构．

(5) 熟练掌握用行初等变换解线性方程组的方法．

二、基本概念与理论

1. 线性方程组的求解过程

解线性方程组 $\boldsymbol{A}_{m\times n}\cdot\boldsymbol{X}_{n\times 1}=\boldsymbol{b}_{m\times 1}$ 的过程，即把增广矩阵通过初等行变换化为阶梯型矩阵．这样做是基于两点：

(1) 任何矩阵都可经过初等行变换化为阶梯型；

(2) 行初等变换保持线性方程组同解．

因此，可从阶梯型矩阵来判断线性方程组解的情况．记

$$(\boldsymbol{A},\boldsymbol{b})=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1\\ a_{21} & a_{22} & \cdots & a_{2n} & b_2\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}\xrightarrow{\text{行}}\cdots\rightarrow\begin{pmatrix} c_{11} & \cdots & c_{1r} & \cdots & c_{1n} & d_1\\ & \ddots & \vdots & \ddots & \vdots & \vdots\\ & & c_{rr} & \cdots & c_{rn} & d_r\\ & & & & & d_{r+1} \end{pmatrix}$$

(这里，在讨论解的情况时，不妨设 $c_{ii}\neq 0,i=1,\cdots,r$)．可见：

(1) 若 $d_{r+1}\neq 0$，则线性方程组无解；

(2) 若 $d_{r+1}=0$，则 $\begin{cases}\text{若 } r<n, & \text{则有无穷多解；}\\ \text{若 } r=n, & \text{则有惟一解．}\end{cases}$

2. 线性方程组有解判别定理

线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times 1}=\boldsymbol{b}_{m\times 1}$ 有解判别定理：$\boldsymbol{AX}=\boldsymbol{b}$ 有解 $\Leftrightarrow R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})$．

$$\text{有解分两种情况：}\begin{cases}\text{无穷多解} & \Leftrightarrow R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})<n\\ \text{惟一解} & \Leftrightarrow R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})=n.\end{cases}$$

这里，n 是方程组未知量的个数．

对于齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times 1}=\boldsymbol{O}$，$R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})$ 总成立．因此，我们关心它是是否有惟一解(即只有零解)还是有无穷多解(即有非零解)．显见，

$$R(\boldsymbol{A})<n \quad\Leftrightarrow\quad \text{有非零解(若 } m=n\text{，则 } \Leftrightarrow |\boldsymbol{A}|=0).$$

$$R(\boldsymbol{A})=n \quad\Leftrightarrow\quad \text{只有零解(若 } m=n\text{，则 } \Leftrightarrow |\boldsymbol{A}|\neq 0).$$

注　从阶梯型矩阵去判断解的情况与用有解判断定理去判断解的情况都是基本而又重要的方法，两者都必须熟练掌握.

3. 线性方程组解的结构

记

$$\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times 1}=\boldsymbol{b}_{m\times 1} \tag{1}$$

它所对应的导出组为

$$\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times 1}=\boldsymbol{O} \tag{2}$$

定理　对于齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{O}$，有

$$\dim(\boldsymbol{\eta}(\boldsymbol{A}))=n-r$$

（这里，$\boldsymbol{\eta}(\boldsymbol{A})$ 是 $\boldsymbol{AX}=\boldsymbol{O}$ 的解空间，$\dim(\boldsymbol{\eta}(\boldsymbol{A}))$ 是解空间的维数，也即基础解系中所含向量的个数，n 是未知量个数，r 是 $R(\boldsymbol{A})$.）

一些有关性质：

（1）若 $\boldsymbol{x}=\xi_1,\boldsymbol{x}=\xi_2$ 是齐次方程(2)的解，则 $\boldsymbol{x}=c_1\xi_1+c_2\xi_2$ 也是方程(2)的解. c_1,c_2 为任意常数；

（2）若 $R(\boldsymbol{A})=r,\xi_1,\cdots,\xi_{n-r}$ 是方程(2)的基础解系，则 $\boldsymbol{x}=c_1\xi_1+\cdots+c_{n-r}\xi_{n-r}$ 是方程(2)的通解；

（3）若 $\boldsymbol{x}=\boldsymbol{\eta}^*$ 是非齐次方程(1)的特解，$\boldsymbol{x}=\xi$ 是方程(2)的通解，则 $\boldsymbol{x}=\boldsymbol{\eta}^*+\xi$ 是方程(1)的通解；

（4）若 $\boldsymbol{x}=\boldsymbol{\eta}_1,\boldsymbol{x}=\boldsymbol{\eta}_2$ 是方程(1)的两个解，则 $\boldsymbol{x}=\boldsymbol{\eta}_1-\boldsymbol{\eta}_2$ 是方程(2)的解；

（5）若 $R(\boldsymbol{A})=r,\xi_1,\cdots,\xi_{n-r}$ 是方程(2)的基础解系，$\boldsymbol{\eta}^*$ 是方程(1)的特解，则 $\boldsymbol{\eta}^*,\boldsymbol{\eta}^*+\xi_1,\cdots,\boldsymbol{\eta}^*+\xi_{n-r}$ 是方程(1)的 $n-r+1$ 个线性无关解. 由此可知，若 $R(\boldsymbol{A})=r$，则方程(2)可有 $n-r$ 个线性无关解，而方程(1)可有 $n-r+1$ 个线性无关的解；

（6）方程(2)的所有解形成线性空间，而方程(1)的所有解并不形成空间；

（7）从公式 $\dim(\boldsymbol{\eta}(\boldsymbol{A}))=n-r$，有时还可证明两个矩阵的秩相等（见例 4-21）.

三、基本题型与解题方法

用行初等变换将 $\boldsymbol{AX}=\boldsymbol{O}$ 的系数阵 $\boldsymbol{A}$ 化成行最简形，则可方便地得到通解表达式.

用行初等变换将 $\boldsymbol{AX}=\boldsymbol{b}$ 的增广矩阵 $\boldsymbol{B}$ 化成行阶梯形矩阵，则可以方便地讨论 $\boldsymbol{AX}=\boldsymbol{b}$ 的解的情况，即：① 有惟一解；② 有无穷多组解；③ 无解. 再进一步把增广矩阵 $\boldsymbol{B}$ 化为行最简形矩阵，就可以在有解时方便地得到通解表达式.

例 4-1　求解

$$\begin{cases} x_1+2x_2+2x_3+x_4=0 \\ 2x_1+x_2-2x_3-2x_4=0 \\ x_1-x_2-4x_3-3x_4=0 \end{cases}$$

分析　这是最基本的不含参数的线性方程组的求解问题，只要用行初等变换将系数矩阵化成行最简形即可求出其解.

解　将系数矩阵用行初等变换化成行最简形：

$$\begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & -2 & -2 \\ 1 & -1 & -4 & -3 \end{bmatrix} \xrightarrow{\text{行初等变换}} \begin{bmatrix} 1 & 0 & -2 & -5/3 \\ 0 & 1 & 2 & 4/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

由此得 $\begin{cases} x_1=2x_3+\dfrac{5}{3}x_4 \\ x_2=-2x_3-\dfrac{4}{3}x_4 \end{cases}$，$x_3,x_4$ 可以任意取值.

为把解表达得更清楚些，可以将

$$\begin{cases} x_1 = \quad 2x_3 + \dfrac{5}{3}x_4 \\ x_2 = -2x_3 - \dfrac{4}{3}x_4 \\ x_3 = \quad x_3 \\ x_4 = \quad\quad x_4 \end{cases}$$

写成向量形式得

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \end{bmatrix} x_3 + \begin{bmatrix} 5/3 \\ -4/3 \\ 0 \\ 1 \end{bmatrix} x_4 \triangleq c_1 \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \end{bmatrix} + c_4 \begin{bmatrix} 5 \\ -4 \\ 0 \\ 3 \end{bmatrix}$$

其中　$\xi_1 = \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \end{bmatrix}$，　$\xi_2 = \begin{bmatrix} 5 \\ -4 \\ 0 \\ 3 \end{bmatrix}$ 是原方程组的一个基础解系.

例 4-2　问 a,b 为何值时，线性方程组

$$\begin{cases} x_1 + x_2 \quad\quad + x_3 + x_4 = 0 \\ \quad x_2 \quad\quad + 2x_3 + 2x_4 = 1 \\ \quad -x_2 + (a-3) \quad x_3 - 2x_4 = b \\ 3x_1 + 2x_2 \quad\quad + x_3 + ax_4 = -1 \end{cases}$$

有(1) 惟一解；(2) 无解；(3) 有无穷多组解？并求出有无穷多组解时的通解.

分析　这是 1987 年数学(试卷一)的试题(计 8 分)，是典型的含有参数的非齐次线性方程组的讨论，要确定方程组有解或无解时的参数值，就要利用有解的充要条件 $R(\boldsymbol{A}) = R(\boldsymbol{B})$. 首先要对增广矩阵进行行初等变换，将它化为行阶梯形.

解

$$\boldsymbol{B} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 2 & 1 \\ 0 & -1 & a-3 & -2 & b \\ 3 & 2 & 1 & a & -1 \end{bmatrix} \xrightarrow{\text{行初等变换}} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & a-1 & 0 & b+1 \\ 0 & 0 & 0 & a-1 & 0 \end{bmatrix},$$

由此易见

(1) 惟一解，要求 $a \neq 1$；

(2) 无解，要求 $a = 1$ 且 $b \neq -1$；

(3) 无穷多组解，要求 $a = 1$，且 $b = -1$，此时将 $a = 1, b = -1$ 代入上面的阶梯形矩阵，再进一步化成行最简形矩阵

$$\boldsymbol{B} \longrightarrow \begin{bmatrix} 1 & 0 & -1 & -1 & -1 \\ 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

进而可以看出通解为

$$\boldsymbol{\eta}=k_1\begin{bmatrix}1\\-2\\1\\0\end{bmatrix}+k_2\begin{bmatrix}1\\-2\\0\\1\end{bmatrix}+\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}.$$

例 4-3 选择题：已知 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的两个不同的解，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是对应齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{O}$ 的基础解系，k_1,k_2 为任意常数，则方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的通解（一般解）必是 （ ）

(A) $k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)+\dfrac{\boldsymbol{\beta}_1-\boldsymbol{\beta}_2}{2}$；　(B) $k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)+\dfrac{\boldsymbol{\beta}_1+\boldsymbol{\beta}_2}{2}$；

(C) $k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\beta}_1+\boldsymbol{\beta}_2)+\dfrac{\boldsymbol{\beta}_1-\boldsymbol{\beta}_2}{2}$；　(D) $k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\beta}_1-\boldsymbol{\beta}_2)+\dfrac{\boldsymbol{\beta}_1+\boldsymbol{\beta}_2}{2}$.

分析 这是 1990 年数学（试卷一）的 3 分题，这类题通常采用排除法，由非齐次线性方程组的解的结构可知，$\boldsymbol{AX}=\boldsymbol{b}$ 的通解是它的一个特解及 $\boldsymbol{AX}=\boldsymbol{O}$ 的通解之和．而 $\dfrac{\boldsymbol{\beta}_1-\boldsymbol{\beta}_2}{2}$ 不是 $\boldsymbol{AX}=\boldsymbol{b}$ 的特解，所以排除(A)，(C)．在(B)，(D)中 $\dfrac{\boldsymbol{\beta}_1+\boldsymbol{\beta}_2}{2}$ 已是 $\boldsymbol{AX}=\boldsymbol{b}$ 的特解，还需检查齐次方程组的解向量是否线性无关．$\boldsymbol{\alpha}_1$ 与 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2$ 是线性无关的，而 $\boldsymbol{\alpha}_1$ 与 $\boldsymbol{\beta}_1-\boldsymbol{\beta}_2$ 是否线性无关不能肯定，所以选择(B)才是正确的答案．

例 4-4 设有三维列向量

$$\boldsymbol{\alpha}_1=\begin{bmatrix}1+\lambda\\1\\1\end{bmatrix},\ \boldsymbol{\alpha}_2=\begin{bmatrix}1\\\lambda+1\\1\end{bmatrix},\ \boldsymbol{\alpha}_3=\begin{bmatrix}1\\1\\1+\lambda\end{bmatrix},\ \boldsymbol{\beta}=\begin{bmatrix}0\\\lambda\\-\lambda^2\end{bmatrix},$$

问 λ 取何值时

(1) $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，且表达式惟一？

(2) $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，但表达式不惟一？

(3) $\boldsymbol{\beta}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示？

分析 这是 1991 年数学（试卷四）的 8 分题，本题形式上看是讨论参数 λ，以决定向量 $\boldsymbol{\beta}$ 能否用 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示的问题，其实质是决定一个线性方程组是否有解的问题，即讨论方程组

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3=\boldsymbol{\beta} \tag{1}$$

在 λ 取何值时有惟一解，无穷多组解及无解．因此本题是线性方程组问题的另一种表现形式．

$$\boldsymbol{B}=\begin{bmatrix}1+\lambda & 1 & 1 & 0\\1 & \lambda+1 & 1 & \lambda\\1 & 1 & 1+\lambda & -\lambda^2\end{bmatrix}$$

中有字母 λ 出现，给行初等变换带来很大麻烦，考虑到对应的线性方程组的系数矩阵是个方阵，因此，用克莱默法则求解更好．

解 由于 $\begin{vmatrix}1+\lambda & 1 & 1\\1 & \lambda+1 & 1\\1 & 1 & 1+\lambda\end{vmatrix}=\lambda^3+3\lambda^2=\lambda^2(\lambda+3)$，

(1) $\lambda\neq 0$ 且 $\lambda\neq -3$ 时，系数行列式不等于零，从而方程组(1)有惟一解，即 $\boldsymbol{\beta}$ 可惟一地由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示．

(2) $\lambda=0$ 时，有

$$\boldsymbol{B}=\begin{bmatrix}1 & 1 & 1 & 0\\1 & 1 & 1 & 0\\1 & 1 & 1 & 0\end{bmatrix},$$

$R(\boldsymbol{A})=R(\boldsymbol{B})=1<3$，从而方程组(1)有无穷多解，即 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，但表示式不惟一.

(3) $\lambda=-3$ 时，有

$$\boldsymbol{B}=\begin{bmatrix}-2 & 1 & 1 & 0\\ 1 & -2 & 1 & -3\\ 1 & 1 & -2 & -4\end{bmatrix}\xrightarrow{\text{行初等变换}}\begin{bmatrix}-2 & 1 & 1 & 0\\ 1 & -2 & 1 & -3\\ 0 & 0 & 0 & -7\end{bmatrix}$$

$R(\boldsymbol{A})<R(\boldsymbol{B})$，从而方程组(1)无解，即 $\boldsymbol{\beta}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示.

例 4-5　设 n 阶方阵 $\boldsymbol{A}\neq\boldsymbol{O}$，求证存在一个 n 阶方阵 $\boldsymbol{B}\neq\boldsymbol{O}$，使 $\boldsymbol{AB}=\boldsymbol{O}$ 的充要条件是 $|\boldsymbol{A}|=0$.

分析　题设中有 $\boldsymbol{AB}=\boldsymbol{O}$，因此仿照上题也可以与齐次线性方程组建立起联系，再用齐次线性方程组理论进行证明.

证　**必要性**　若存在 $\boldsymbol{B}\neq\boldsymbol{O}$，使 $\boldsymbol{AB}=\boldsymbol{O}$，则 $\boldsymbol{B}$ 的列向量都是 $\boldsymbol{AX}=\boldsymbol{O}$ 的解. 因为 $\boldsymbol{B}\neq\boldsymbol{O}$，所以 $\boldsymbol{B}$ 至少有一个非零列向量，也就是说明 $\boldsymbol{AX}=\boldsymbol{O}$ 有非零解，因此断言 $R(\boldsymbol{A})<n$，即 $|\boldsymbol{A}|=0$.

充分性　设 $|\boldsymbol{A}|=0$，则 $\boldsymbol{AX}=\boldsymbol{O}$ 有非零解，记为 X_0. 令 $\boldsymbol{B}=(X_0,0,\cdots,0)$，则 $\boldsymbol{B}\neq 0$，且满足 $\boldsymbol{AB}=0$.

例 4-6　已知：

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\0\\2\\3\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}1\\1\\3\\5\end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix}1\\-1\\a+2\\1\end{pmatrix},\quad \boldsymbol{\alpha}_4=\begin{pmatrix}1\\2\\4\\a+8\end{pmatrix},\quad \boldsymbol{\beta}=\begin{pmatrix}1\\1\\b+3\\5\end{pmatrix}.$$

(1) a,b 为何值时，$\boldsymbol{\beta}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性表示？

(2) a,b 为何值时，$\boldsymbol{\beta}$ 可惟一由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性表示？并写出该表达式.

解

$$\begin{aligned}\boldsymbol{\beta}&=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3+x_4\boldsymbol{\alpha}_4\\&=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\begin{pmatrix}x_1\\x_2\\x_2\\x_4\end{pmatrix}\end{aligned}\tag{1}$$

记线性方程组(1)为 $\boldsymbol{AX}=\boldsymbol{\beta}$. 可见本题即为判断 $\boldsymbol{AX}=\boldsymbol{\beta}$ 的解的情况.

$$(\bar{\boldsymbol{A}})=\begin{pmatrix}1&1&1&1&1\\0&1&-1&2&1\\2&3&a+2&4&b+3\\3&5&1&a+8&5\end{pmatrix}\xrightarrow{\text{行}}\cdots\rightarrow\begin{pmatrix}1&1&1&1&1\\0&1&-1&2&1\\0&0&a+1&0&b\\0&0&0&a+1&0\end{pmatrix}.$$

由阶梯型矩阵可知：

(1) 当 $a=-1$，且 $b\neq 0$ 时，$R(\boldsymbol{A})\neq R(\bar{\boldsymbol{A}})$，方程组无解，亦即 $\boldsymbol{\beta}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性表示；

(2) 当 $a\neq-1$ 时，$R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})=4=$ 未知量个数，故有惟一解：

$$x_4=0,\quad x_3=\frac{b}{a+1},\quad x_2=1+\frac{b}{a+1},\quad x_1=\frac{-2b}{a+1}.$$

即有

$$\boldsymbol{\beta}=\frac{-2b}{a+1}\cdot\begin{pmatrix}1\\0\\2\\3\end{pmatrix}+\left(1+\frac{b}{a+1}\right)\begin{pmatrix}1\\1\\3\\5\end{pmatrix}+\frac{b}{a+1}\begin{pmatrix}1\\-1\\a+2\\1\end{pmatrix}.$$

评注　(1) 一个向量由另外一组向量线性表示的问题实即线性方程组有解无解的问题.

(2) 把一个向量组的线性组合改写成矩阵乘法是经常用到的手法. 这里的写成线性方程组(1)式,以及第一章 1-30 题中矩阵 C 改写成两个矩阵相乘,都是此法的运用,请注意学会.

例 4-7 设

$$\begin{cases}(1+\lambda)x_1+ & x_2+ & x_3=0\\ x_1+(1+\lambda)x_2+ & x_3=3\\ x_1+ & x_2+(1+\lambda)x_3=\lambda\end{cases},$$

其中,λ 是参数. 试讨论解的情况.

解法 1

$$(\bar{\boldsymbol{A}})=\begin{pmatrix}1+\lambda & 1 & 1 & 0\\ 1 & 1+\lambda & 1 & 3\\ 1 & 1 & 1+\lambda & \lambda\end{pmatrix}\xrightarrow{\text{行}}\cdots\to\begin{pmatrix}1 & 1 & 1+\lambda & \lambda\\ 0 & \lambda & -\lambda & 3-\lambda\\ 0 & 0 & -\lambda(3+\lambda) & (1-\lambda)(3+\lambda)\end{pmatrix}.$$

由阶梯型矩阵可知:

(1) 当 $\lambda\neq 0$,且 $\lambda\neq -3$ 时,$R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})=3$(未知量个数),方程组有惟一解;

(2) 当 $\lambda=0$ 时,$R(\boldsymbol{A})=1<R(\bar{\boldsymbol{A}})=2$,方程组无解;

(3) 当 $\lambda=-3$ 时,$R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})=2<3$,方程组有无穷多解.

解法 2

$$|\boldsymbol{A}|=\begin{vmatrix}1+\lambda & 1 & 1\\ 1 & 1+\lambda & 1\\ 1 & 1 & 1+\lambda\end{vmatrix}=\lambda^2(\lambda+3).$$

(1) 当 $\lambda\neq 0$,且 $\lambda\neq -3$ 时,$|\boldsymbol{A}|\neq 0$,由克莱默法则知方程组有惟一解;

(2) 当 $\lambda=0$ 时,

$$(\bar{\boldsymbol{A}})=\begin{pmatrix}1 & 1 & 1 & 0\\ 1 & 1 & 1 & 3\\ 1 & 1 & 1 & 0\end{pmatrix}\to\begin{pmatrix}1 & 1 & 1 & 0\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0\end{pmatrix}.$$

$R(\boldsymbol{A})=1<R(\bar{\boldsymbol{A}})=2$,方程组无解.

(3) 当 $\lambda=-3$ 时,

$$(\bar{\boldsymbol{A}})=\begin{pmatrix}-2 & 1 & 1 & 0\\ 1 & -2 & 1 & 3\\ 1 & 1 & -2 & -3\end{pmatrix}\to\begin{pmatrix}1 & -2 & 1 & 3\\ 0 & 3 & -3 & -6\\ 0 & 0 & 0 & 0\end{pmatrix}.$$

$R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})=2<3$,方程组有无穷多解.

评注 (1) 判断方程组解的情况,用初等行变换化到阶梯型矩阵是一种基本的方法,但应指出,带有参数的矩阵要正确化到阶梯型有时也不是很方便的,像本题的方程组,系数矩阵是方阵,且参数也在系数矩阵中出现,则可考虑用行列式的方法(即克莱默法则),有时更为方便一些.

(2) 对于线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$,若 $|\boldsymbol{A}|=0$,则可能无解,也可能有无穷多解.

例 4-8

$$\begin{cases}x_1+ & x_2+ & x_3=1\\ a\,x_1+b\,x_2+c\,x_3=d\\ a^2x_1+b^2x_2+c^2x_3=d^2\\ a^3x_1+b^3x_2+c^3x_3=d^3\end{cases},$$

其中 a,b,c,d 互异. 试判断解的情况.

解　$|\bar{A}|$ 是范德蒙行列式，且已知 a,b,c,d 互异. 因此 $|\bar{A}|\neq 0$，可知 $R(\bar{A})=4$，且 $R(A)=3$. 因此，此线性方程组无解.

评注　本题不能考虑用初等行变换把 $\bar{A}$ 化为阶梯型. 因此，对于判断方程组解的情况，应该两种方法（即化为阶梯型的方法和有解判别定理）都熟练掌握.

例 4-9　要使

$$\xi_1=\begin{pmatrix}1\\0\\2\end{pmatrix},\quad \xi_2=\begin{pmatrix}0\\1\\-1\end{pmatrix}$$

都是 $AX=O$ 的解，只要系数矩阵 A 为　（　）

(A) $(-2,1,1)$；　(B) $\begin{pmatrix}2&0&-1\\0&1&1\end{pmatrix}$；

(C) $\begin{pmatrix}-1&0&2\\0&1&-1\end{pmatrix}$；　(D) $\begin{pmatrix}0&1&-1\\4&-2&-2\\0&1&1\end{pmatrix}$.

解　由已知条件知：未知量个数 $n=3$，解空间的维数 $\dim(\eta(A))=2$，故由 $\dim(\eta(A))=n-R(A)$，知 $R(A)=1$，所以选(A).

评注　题目涉及 $AX=O$ 的解的结构，总应先想到公式 $\dim(\eta(A))=n-R(A)$.

例 4-10　已知线性方程组：

$$(\text{Ⅰ})\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{12n}x_{2n}=0\\a_{21}x_1+a_{22}x_2+\cdots+a_{22n}x_{2n}=0\\\vdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{n2n}x_{2n}=0\end{cases}$$

的一个基础解系为

$$b_1=\begin{pmatrix}b_{11}\\b_{12}\\\vdots\\b_{12n}\end{pmatrix},\quad b_2=\begin{pmatrix}b_{21}\\b_{22}\\\vdots\\b_{22n}\end{pmatrix},\cdots,b_n=\begin{pmatrix}b_{n1}\\b_{n2}\\\vdots\\b_{n2n}\end{pmatrix}.$$

试写出方程组.

$$(\text{Ⅱ})\begin{cases}b_{11}y_1+b_{12}y_2+\cdots+b_{12n}y_{2n}=0\\b_{21}y_1+b_{22}y_2+\cdots+b_{22n}y_{2n}=0\\\vdots\\b_{n1}y_1+b_{n2}y_2+\cdots+b_{n2n}y_{2n}=0\end{cases}$$

的通解，并说明理由.

解　记方程组(Ⅰ)：$A_{n\times 2n}X_{2n\times 1}=O$，则由已知条件可有：

$$R(A)=2n-\dim(\eta(A))=2n-n=n,$$

且
$$Ab_i=O,\ (i=1,\cdots,n),\quad \text{即}\quad A(b_1,\cdots,b_n)=O.$$

记 $B=(b_1,b_2,\cdots,b_n)$，则上式为 $A_{n\times 2n}\cdot B_{2n\times n}=O$，即有

$$B^{\mathrm{T}}_{n\times 2n}\cdot A^{\mathrm{T}}_{2n\times n}=O.$$

而方程组(Ⅱ)即为 $B^{\mathrm{T}}_{n\times 2n}\cdot Y_{2n\times 1}=O$. 由于 $R(B)=n$，故 $R(B^{\mathrm{T}})=n$，故

$$\dim(\boldsymbol{\eta}(\boldsymbol{B}^{\mathrm{T}})) = 2n - R(\boldsymbol{B}^{T}) = 2n - n = n.$$

即知方程组(Ⅱ)的基础解系有 n 个线性无关的解向量. 现已有 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{O}$,即 $\boldsymbol{A}^{\mathrm{T}}$ 的 n 个线性无关列向量正可作为 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y} = \boldsymbol{O}$ 的基础解系. 所以,方程组(Ⅱ)的通解可表示为

$$\boldsymbol{Y} = c_1\boldsymbol{A}_1^{\mathrm{T}} + \cdots + c_n\boldsymbol{A}_n^{\mathrm{T}}.$$

这里,$\boldsymbol{A}_i^{\mathrm{T}}$ 是 $\boldsymbol{A}^{\mathrm{T}}$ 的第 i 列($i = 1, \cdots, n$),c_i 为任意常数.

评注 本题已知条件与所求通解,都与基础解系有关,则应考虑公式$\dim(\boldsymbol{\eta}(\boldsymbol{A})) = n - R(\boldsymbol{A})$. 另外,矩阵分块及矩阵转置等运算也要掌握.

例 4-11 已知

$$\boldsymbol{A}_{s\times 4} \cdot \boldsymbol{X}_{4\times 1} = \boldsymbol{b}_{s\times 1} \tag{1}$$

$R(\boldsymbol{A}) = 3$. $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 是方程组(1)的三个解向量,且

$$\boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 7 \end{pmatrix}, \quad \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix}.$$

求方程组(1)的通解.

解 非齐次方程组通解的结构是 $x = \boldsymbol{\eta}^{*} + \xi$. 其中,$\boldsymbol{\eta}^{*}$ 是方程组(1)的一个特解,ξ 是方程组(1)所对应的导出组的通解. 现取 $\boldsymbol{\eta}^{*} = \boldsymbol{\alpha}_1$.

由已知未知量个数 $n = 4$,$R(\boldsymbol{A}) = 3$,故基础解系中只有 1 个线性无关解,即只要找到齐次方程组的一个非零解即可,则可令:

$$\boldsymbol{\xi}_1 = (\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2) + (\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_3) = 2\boldsymbol{\alpha}_1 - (\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3) = \begin{pmatrix} 0 \\ 4 \\ 0 \\ 12 \end{pmatrix},$$

则方程组(1)的通解为

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 7 \end{pmatrix} + c_1 \cdot \begin{pmatrix} 0 \\ 4 \\ 0 \\ 12 \end{pmatrix}, \quad c_1 \text{ 为任意常数.}$$

评注 本题综合运用了线性方程组解的结构的一些性质,概念较强,值得回味.

例 4-12 若 $\boldsymbol{A}_{m\times n} \cdot \boldsymbol{B}_{n\times p} = \boldsymbol{O}$,则 $R(\boldsymbol{A}) + R(\boldsymbol{B}) \leqslant n$.

证 把 $\boldsymbol{B}$ 按列分块,记为 $\boldsymbol{B} = (\boldsymbol{B}_1, \cdots, \boldsymbol{B}_p)$.

则

$$\begin{aligned} \boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{O} &\Leftrightarrow \boldsymbol{A}(\boldsymbol{B}_1, \cdots, \boldsymbol{B}_p) = \boldsymbol{O} \\ &\Leftrightarrow (\boldsymbol{AB}_1, \cdots, \boldsymbol{AB}_p) = \boldsymbol{O} \\ &\Leftrightarrow \boldsymbol{AB}_i = \boldsymbol{O},\ i = 1, 2, \cdots, p. \end{aligned}$$

此条件表明:$\boldsymbol{B}_1, \cdots, \boldsymbol{B}_p$ 均为 $\boldsymbol{AX} = \boldsymbol{O}$ 的解. 则向量组$\{\boldsymbol{B}_1, \cdots, \boldsymbol{B}_p\}$ 可由 $\boldsymbol{AX} = \boldsymbol{O}$ 的基础解系线性表示. 则,

$$\text{秩}\{\boldsymbol{B}_1, \cdots, \boldsymbol{B}_p\} \leqslant \text{秩}\{\boldsymbol{AX} = \boldsymbol{O} \text{ 的基础解系}\}$$

即

$$R(\boldsymbol{B}) \leqslant n - R(\boldsymbol{A})$$

移项即得证.

评注　(1) 本证明中所用到的知识点都是经常用到的：① 把矩阵方程 $\boldsymbol{AB}=\boldsymbol{O}$ 写成等价的一组齐次线性方程组 $\boldsymbol{AB}_i=\boldsymbol{O}, i=1,\cdots,p$；② 向量组 Ⅰ 可由另一个向量组 Ⅱ 线性表示，则秩(Ⅰ)$\leqslant$秩(Ⅱ)；③ 基础解系的概念.

(2) 本题的结论在很多场合有用. 只要遇有二个矩阵相乘为零矩阵，则应知有本题的结论(见例 4-13).

例 4-13　试证：若 $\boldsymbol{A}^2=\boldsymbol{E}$，则 $R(\boldsymbol{A}+\boldsymbol{E})+R(\boldsymbol{A}-\boldsymbol{E})=n$.

证　由
$$\boldsymbol{A}^2=\boldsymbol{E} \quad \Rightarrow \quad (\boldsymbol{A}+\boldsymbol{E})(\boldsymbol{A}-\boldsymbol{E})=\boldsymbol{O}$$

则
$$R(\boldsymbol{A}+\boldsymbol{E})+R(\boldsymbol{A}-\boldsymbol{E})\leqslant n \tag{1}$$

又因
$$(\boldsymbol{A}+\boldsymbol{E})+(\boldsymbol{E}-\boldsymbol{A})=2\boldsymbol{E}$$

有
$$2(\boldsymbol{A}+\boldsymbol{E})+R(\boldsymbol{E}-\boldsymbol{A})\geqslant R(2\boldsymbol{E})=n \tag{2}$$

而
$$R(\boldsymbol{E}-\boldsymbol{A})=R(\boldsymbol{A}-\boldsymbol{E})$$

故由式(1)与式(2)，知
$$R(\boldsymbol{A}+\boldsymbol{E})+R(\boldsymbol{A}-\boldsymbol{E})=n.$$

评注　对于已知 $\boldsymbol{A}^2=\boldsymbol{A}$，求证 $R(\boldsymbol{A})+R(\boldsymbol{A}-\boldsymbol{E})=n$ 之类的题目，都可用类似的方法来证明.

例 4-14　设 $\boldsymbol{A},\boldsymbol{B}$ 是 n 阶方阵，满足 $\boldsymbol{ABA}=\boldsymbol{B}^{-1}$. 求证：
$$R(\boldsymbol{E}-\boldsymbol{AB})+R(\boldsymbol{E}+\boldsymbol{AB})=n$$

证　由
$$\boldsymbol{ABA}=\boldsymbol{B}^{-1} \quad \Rightarrow \quad \boldsymbol{ABAB}=\boldsymbol{E}$$

即
$$(\boldsymbol{AB})^2-\boldsymbol{E}=\boldsymbol{O} \quad \Rightarrow \quad (\boldsymbol{AB}+\boldsymbol{E})(\boldsymbol{AB}-\boldsymbol{E})=\boldsymbol{O}$$

则
$$R(\boldsymbol{E}+\boldsymbol{AB})+R(\boldsymbol{E}-\boldsymbol{AB})\leqslant n \tag{1}$$

又因
$$(\boldsymbol{E}-\boldsymbol{AB})+(\boldsymbol{E}+\boldsymbol{AB})=2\boldsymbol{E}$$

则有
$$R(\boldsymbol{E}-\boldsymbol{AB})+R(\boldsymbol{E}+\boldsymbol{AB})\geqslant R(2\boldsymbol{E})=n \tag{2}$$

由式(1)和式(2)，便有
$$R(\boldsymbol{E}-\boldsymbol{AB})+R(\boldsymbol{E}+\boldsymbol{AB})=n.$$

例 4-15　设 $\boldsymbol{A}$ 是 n 阶不可逆矩阵，$\boldsymbol{A}\neq\boldsymbol{O}$. 试证存在 n 阶非单位阵 $\boldsymbol{B}$，使 $\boldsymbol{AB}=\boldsymbol{A}$.

证　把 $\boldsymbol{A},\boldsymbol{B}$ 都按列分块，记为
$$\boldsymbol{A}=(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n), \quad \boldsymbol{B}=(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)$$

则
$$\begin{aligned}\boldsymbol{AB}=\boldsymbol{A} \quad &\Leftrightarrow \quad \boldsymbol{A}(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)=(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)\\ &\Leftrightarrow \quad (\boldsymbol{AB}_1,\cdots,\boldsymbol{AB}_n)=(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)\\ &\Leftrightarrow \quad \boldsymbol{AB}_i=\boldsymbol{A}_i, \quad i=1,\cdots,n.\end{aligned}$$

现要证矩阵 $\boldsymbol{B}$ 的存在性，即要证：非齐次线性方程组 $\boldsymbol{AB}_i=\boldsymbol{A}_i$ 中的 $\boldsymbol{B}_i$ 的存在性.

由于
$$R(\boldsymbol{A})=R(\boldsymbol{A},\boldsymbol{A}_i), \quad \text{且} \quad R(\boldsymbol{A})<n$$

故知方程组 $\boldsymbol{AB}_i=\boldsymbol{A}_i$ 有无穷多解. 则在每一个方程组中选取某 $\boldsymbol{B}_i(i=1,\cdots,n)$，使$(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)\neq\boldsymbol{E}$ 即可.

评注　首先把矩阵方程改写成等价的一组非齐次线性方程组，则本题就化为论证线性方程组 $\boldsymbol{AB}_i=\boldsymbol{A}_i$ 的解的存在性，则抓住有解判别定理 $R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})$ 去论证即可.

例 4-16　$\boldsymbol{A}$ 是 n 阶方阵，若有矩阵 $\boldsymbol{B}_{n\times p}\neq 0$，使 $\boldsymbol{AB}=\boldsymbol{O}$，则必有非零矩阵 $\boldsymbol{C}$，使 $\boldsymbol{CA}=\boldsymbol{O}$.

证法 1　由
$$\boldsymbol{AB}=\boldsymbol{O} \quad \Leftrightarrow \quad \boldsymbol{AB}_i=\boldsymbol{O} \quad (i=1,\cdots,p).$$

由 $\boldsymbol{B}\neq\boldsymbol{O} \quad \Rightarrow \quad \boldsymbol{AX}=\boldsymbol{O}$ 有非零解 $\quad \Rightarrow \quad |\boldsymbol{A}|=0 \quad \Rightarrow \quad \boldsymbol{A}$ 的行向量组线性相关，即存在不全为零的数 x_1，

…，x_n，使

$$(x_1,\cdots,x_n)\cdot\begin{pmatrix}A_1\\ \vdots\\ A_n\end{pmatrix}=\boldsymbol{O},\quad \text{这里 } A_i \text{ 表示 } \boldsymbol{A} \text{ 的第 } i \text{ 行.}$$

则把$(x_1,\cdots,x_n)$作为所要求的非零矩阵 $\boldsymbol{C}$ 即可.

证法 2 由

$$\boldsymbol{AB}=\boldsymbol{O}\ \Rightarrow\ R(\boldsymbol{A})+R(\boldsymbol{B})\leqslant n$$

又由 $\boldsymbol{B}\neq\boldsymbol{O}\ \Rightarrow\ R(\boldsymbol{A})<n\ \Rightarrow$ $\boldsymbol{A}$ 的行向量组线性相关，以下步骤同证法 1.

评注 本题要论证非零矩阵 $\boldsymbol{C}$，左乘 $\boldsymbol{A}$ 后等于零矩阵，因此要考虑 $\boldsymbol{A}$ 按行分块，这样才可以作分块矩阵的乘法.

例 4-17 设线性方程组

$$\begin{cases}x_1+a_1x_2+a_1^2x_3=a_1^3\\ x_1+a_2x_2+a_2^2x_3=a_2^3\\ x_1+a_3x_2+a_3^2x_3=a_3^3\\ x_1+a_4x_2+a_4^2x_3=a_4^3\end{cases}\tag{1}$$

(1) 证明：若 a_1,a_2,a_3,a_4 两两不相等，则此线性方程组无解；

(2) 设 $a_1=a_3=k,a_2=a_4=-k(k\neq0)$，且已知 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是该方程组的两个解，其中 $\boldsymbol{\beta}_1=(-1,1,1)^{\mathrm{T}}$，$\boldsymbol{\beta}_2=(1,1,-1)^{\mathrm{T}}$，写出此方程组的通解.

证 (1) 方程组(1)的增广矩阵行列式是范德蒙行列式，又已知 a_1,a_2,a_3,a_4 两两不相等，则

$$|\bar{\boldsymbol{A}}|=\begin{vmatrix}1&a_1&a_1^2&a_1^3\\ 1&a_2&a_2^2&a_2^3\\ 1&a_3&a_3^2&a_3^3\\ 1&a_4&a_4^2&a_4^3\end{vmatrix}\neq0$$

$\Rightarrow\ R(\bar{\boldsymbol{A}})=4\ \Rightarrow\ R(\boldsymbol{A})=3$

$\Rightarrow$ 此方程组无解.

解 (2) 当 $a_1=a_3=k,a_2=a_4=-k(k\neq0)$ 方程组(1)为

$$\begin{cases}x_1+kx_2+k^2x_3=k^3\\ x_1-kx_2+k^2x_3=-k^3\\ x_1+kx_2+k^2x_3=k^3\\ x_1-kx_2+k^2x_3=-k^3\end{cases},$$

即为

$$\begin{cases}x_1+kx_2+k^2x_3=k^3\\ x_1-kx_2+k^2x_3=-k^3\end{cases}.$$

由于

$$\begin{vmatrix}1&k\\ 1&-k\end{vmatrix}=-2\neq0\ \Rightarrow\ R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})=2.$$

未知量个数 $n=3$，故知有无穷多解，且其对应导出组的基础解系应含有 $3-2=1$ 个解向量，则令

$$\boldsymbol{\xi}=\boldsymbol{\beta}_1-\boldsymbol{\beta}_2=\begin{pmatrix}1\\1\\-1\end{pmatrix}-\begin{pmatrix}-1\\1\\1\end{pmatrix}=\begin{pmatrix}2\\0\\-2\end{pmatrix}.$$ 则(1)的通解为

$$\boldsymbol{x}=\boldsymbol{\beta}_1+c_1\boldsymbol{\xi}=\begin{pmatrix}-1\\1\\1\end{pmatrix}+c_1\cdot\begin{pmatrix}2\\0\\-2\end{pmatrix},\quad c_1\text{ 为任意常数.}$$

例 4-18　设 $\boldsymbol{A}_{m\times n}$ 为非零实矩阵. 试证：$R(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=R(\boldsymbol{A})$.

证　只要证明

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{O}\quad(\mathrm{I})\qquad\text{与}\qquad\boldsymbol{A}\boldsymbol{X}=\boldsymbol{O}\quad(\mathrm{II})$$

是同解方程组即可.

若 $\boldsymbol{Y}$ 是方程组(II)的解，则 $\boldsymbol{Y}$ 也是(I)的解，这是显然的. 再证：若 $\boldsymbol{Y}$ 是(I)的解，则 $\boldsymbol{Y}$ 也必是(II)的解，即由

$$\begin{aligned}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Y}=\boldsymbol{O}\quad&\Rightarrow\quad\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Y}=\boldsymbol{Y}^{\mathrm{T}}\cdot\boldsymbol{O}=\boldsymbol{O}\\&\Rightarrow\quad(\boldsymbol{A}\boldsymbol{Y})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{Y})=\boldsymbol{O}\\&\Rightarrow\quad\boldsymbol{A}\boldsymbol{Y}=\boldsymbol{O}\end{aligned}$$

这样　　$\dim(\boldsymbol{\eta}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}))=\dim(\boldsymbol{\eta}(\boldsymbol{A}))$.

由于　　$\dim(\boldsymbol{\eta}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}))=n-R(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$

$\dim(\boldsymbol{\eta}(\boldsymbol{A}))=n-R(\boldsymbol{A})$

故　　$R(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=R(\boldsymbol{A})$.

评注　(1) 从证明 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{O}$ 与 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{O}$ 是同解方程组. 再借助重要公式 $\dim(\boldsymbol{\eta}(\boldsymbol{A}))=n-R(\boldsymbol{A})$. 得到了重要结论 $R(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=R(\boldsymbol{A})$. 这个结论在不少地方有用(见例 4-19 ～ 例 4-21).

(2) 注意题目的条件：$\boldsymbol{A}$ 是实矩阵. 否则，从 $(\boldsymbol{A}\boldsymbol{Y})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{Y})=\boldsymbol{O}\not\Rightarrow\boldsymbol{A}\boldsymbol{Y}=\boldsymbol{O}$.

例 4-19　$\boldsymbol{A}_{m\times n}$ 为实矩阵. 试证 $R(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=R(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})$.

证　由　　$R(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=R(\boldsymbol{A})\quad\Rightarrow\quad R(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})=R(\boldsymbol{A}^{\mathrm{T}})$.

又由　　$R(\boldsymbol{A})=R(\boldsymbol{A}^{\mathrm{T}})$,

故有　　$R(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=R(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})$.

例 4-20　若 $\boldsymbol{A}_{m\times n}\cdot\boldsymbol{X}_{n\times 1}=\boldsymbol{b}_{m\times 1}$ 有惟一解，且 $\boldsymbol{A}$ 是实矩阵，求 $\boldsymbol{X}$.

解　因 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$ 有惟一解，故知

$$R(\boldsymbol{A})=R(\bar{\boldsymbol{A}})=n$$

又因 $R(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=R(\boldsymbol{A})=n$，且 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 n 阶方阵，故有 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 可逆.

所以　　$\boldsymbol{X}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$.

例 4-21　若 $\boldsymbol{A}$ 是 n 阶实矩阵，$\boldsymbol{B}_{n\times s}$ 是实矩阵，且 $R(\boldsymbol{B})=n$. 若 $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{O}$，则 $\boldsymbol{A}=\boldsymbol{O}$.

证法 1　由　　$\boldsymbol{A}\boldsymbol{B}=\boldsymbol{O}\quad\Rightarrow\quad\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{O}\cdot\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{O}$

因为 $R(\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}})=R(\boldsymbol{B})=n$，且 $\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}$ 为 n 阶方阵，故 $\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}$ 可逆. 所以

$$\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}})^{-1}=\boldsymbol{O}\cdot(\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}})^{-1}=\boldsymbol{O}$$

即　　$\boldsymbol{A}=\boldsymbol{O}$.

证法 2　由　　$\boldsymbol{A}\boldsymbol{B}=\boldsymbol{O}\quad\Rightarrow\quad R(\boldsymbol{A})+R(\boldsymbol{B})\leqslant n$,

又由 $R(\boldsymbol{B})=n\quad\Rightarrow\quad R(\boldsymbol{A})\leqslant 0\quad\Rightarrow\quad R(\boldsymbol{A})=0\quad\Rightarrow\quad\boldsymbol{A}=\boldsymbol{O}$.

评注　对于不是方阵 $\boldsymbol{A}_{m\times n}$，则可左乘或右乘它的转置阵，使之成为方阵. 若此矩阵秩已满(这里指秩不

能再大了,即秩为 m 或 n),则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 或 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 就是一个可逆方阵了.

例 4-22 $\boldsymbol{A}$ 是 n 阶方阵,$R(\boldsymbol{A})=n-1$. $\boldsymbol{A}$ 中元素 a_{rs} 的代数余子式 $\boldsymbol{A}_{rs}\neq\boldsymbol{O}$. 试证:$\boldsymbol{\alpha}=(\boldsymbol{A}_{r1},\cdots,\boldsymbol{A}_{rs},\cdots,\boldsymbol{A}_{rn})^{\mathrm{T}}$ 是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{O}$ 的基础解系.

证 由

$$\dim(\boldsymbol{\eta}(\boldsymbol{A}))=n-R(\boldsymbol{A})=n-(n-1)=1,$$

知此基础解系中只有一个解向量. 现 $\boldsymbol{\alpha}\neq\boldsymbol{O}$,故只要证明 $\boldsymbol{\alpha}$ 是 $\boldsymbol{AX}=\boldsymbol{O}$ 的解,则 $\boldsymbol{\alpha}$ 就是它的基础解系.

由于

$$a_{i1}A_{r1}+\cdots+a_{in}A_{rn}=0,\ (r\neq i).$$

$$a_{r1}A_{r1}+\cdots+a_{rn}A_{rn}=|A|=0,$$

表明 $\boldsymbol{\alpha}$ 适合 $\boldsymbol{AX}=\boldsymbol{O}$ 的所有 n 个方程,故 $\boldsymbol{\alpha}$ 是解.

例 4-23 设

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rn} \\ a_{r+11} & \cdots & a_{r+1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},\quad R(\boldsymbol{A})=n.$$

又设

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=0 \tag{1}$$

试证:

$$A_1=\begin{pmatrix} A_{r+11} \\ \vdots \\ A_{r+1n} \end{pmatrix},\cdots,A_{n-r}=\begin{pmatrix} A_{n1} \\ \vdots \\ A_{nn} \end{pmatrix}.$$

是方程组(1)的基础解系. 其中,A_{ij} 是 a_{ij} 的代数余子式.

证 记方程组(1)为 $\boldsymbol{BX}=\boldsymbol{O}$.

则由 $R(\boldsymbol{A})=n\ \Rightarrow\ R(\boldsymbol{B})=r$,则要证 $A_1,\cdots,A_{n-r}$ 是方程组(1)的基础解系,即要证两件事:一是证它们是方程组(1)的解;二是证它们线性无关.

先证 $A_1,\cdots,A_{n-r}$ 是方程组(1)的解:

$$a_{i1}A_{k1}+a_{i2}A_{k2}+\cdots+a_{in}A_{kn}=0\quad(i=1,\cdots,r,\ k=r+1,\cdots,n,\ i\neq k).$$

再证 $A_1,\cdots,A_{n-r}$ 线性无关:由于 $\boldsymbol{A}$ 可逆,故 $\boldsymbol{A}^*$ 可逆,则 $\boldsymbol{A}^*$ 的 n 个列向量线性无关,而 $A_1,\cdots,A_{n-r}$ 正是 $\boldsymbol{A}^*$ 的右边 $n-r$ 个列向量,故线性无关.

例 4-24 若 n 元线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 对任何 n 维列向量 $\boldsymbol{b}$ 均有解,则对于任何 n 维列向量 $\boldsymbol{\beta}$,方程组 $\boldsymbol{A}^*\boldsymbol{X}=\boldsymbol{\beta}$ 必有惟一解. 这里,$\boldsymbol{A}^*$ 是 $\boldsymbol{A}$ 的伴随矩阵.

证 要证 $\boldsymbol{A}^*\boldsymbol{X}=\boldsymbol{\beta}$ 有惟一解.

$$\Leftrightarrow\quad R(\boldsymbol{A}^*)=R(\boldsymbol{A}^*,\boldsymbol{\beta})=n$$

$$\Leftrightarrow\quad \boldsymbol{A}^*\text{ 可逆}.$$

则只要证 $\boldsymbol{A}$ 可逆即可.

由已知对任何列向量 b，$AX=b$ 均有解. 则可分别取

$$\boldsymbol{b}=\boldsymbol{\varepsilon}_i,\ (i=1,\cdots,n)$$

$$\boldsymbol{\varepsilon}_i=\begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix}i,\ i=1,2,\cdots,n.$$

则有　　$$A\boldsymbol{\alpha}_1=\boldsymbol{\varepsilon}_1,\cdots,A\boldsymbol{\alpha}_n=\boldsymbol{\varepsilon}_n.$$

即　　$$A(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)=(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)=E$$

则 A 可逆.

评注　(1) 对于题给条件中："对任何列向量 b，都有 ……" 这样的语句，我们应考虑选择性质最好，运用最熟练、最方便的向量.

(2) 把 n 个线性方程组 $A\boldsymbol{\alpha}_i=\boldsymbol{\varepsilon}_i,i=1,\cdots,n$. 写成一个矩阵方程：$A(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)=(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n)$，这种方法前面已经出现过. 前面是把矩阵方程写成等价的一组线性方程组，现是倒过来，应学会灵活运用.

例 4-25　若方阵 A 与任何矩阵乘法可交换，则 A 是数量阵（即 $A=aE$）.

证　由已知条件，对任何矩阵 B，皆有 $AB=BA$，则可选取：

$$B_1=\begin{pmatrix}1&&&\\&2&&\\&&\ddots&\\&&&n\end{pmatrix},\text{由 }AB_1=B_1A\ \Rightarrow\ a_{ij}=0\ (i\neq j).$$

再取

$$B_2=\begin{pmatrix}0&1&&&&\\1&0&&&&\\&&1&&&\\&&&\ddots&&\\&&&&1\end{pmatrix},\text{由 }AB_2=B_2A\ \Rightarrow\ a_{11}=a_{22}.$$

再取

$$B_3=\begin{pmatrix}1&&&&&\\&0&1&&&\\&1&0&&&\\&&&1&&\\&&&&\ddots&\\&&&&&1\end{pmatrix},\text{由 }AB_3=B_3A\ \Rightarrow\ a_{22}=a_{33},$$

…

类似可得　　$$a_{11}=a_{22}=\cdots=a_{nn},$$

所以　　$$A=aE.$$

例 4-26　设 A 是实反对称阵，即 $A^{\mathrm{T}}=-A$. 试证：$|E-A|\neq 0$.

证　$|E-A|\neq 0\ \Leftrightarrow$　线性方程组 $(E-A)X=O$ 只有零解.

现设 ξ 是 $(E-A)X=O$ 的解. 须证：$\xi=O$.

由　　$$(E-A)\xi=O\ \Rightarrow\ \xi^{\mathrm{T}}(E-A)\xi=\xi^{\mathrm{T}}\cdot O=O$$

$$\Rightarrow \quad \xi^{\mathrm{T}}\xi-\xi^{\mathrm{T}}A\xi=O$$

由于 A 反对称，故 $\xi^{\mathrm{T}}A\xi=O$，$\forall\,\xi\neq O$（这是因为：$\xi^{\mathrm{T}}A\xi=(\xi^{\mathrm{T}}A\xi)^{\mathrm{T}}=\xi^{\mathrm{T}}A^{\mathrm{T}}\xi=-\xi^{\mathrm{T}}A\xi$）.

所以

$$\xi^{\mathrm{T}}\xi=O \quad \Rightarrow \quad \xi=O.$$

评注 （1）把要证行列式 $|E-A|\neq 0$，等价地转化为考虑：齐次线性方程组 $(E-A)X=O$ 只有零解. 这是两方面的知识点，却是有关联的. 线性代数中，充分必要的条件要熟悉.

（2）论证某实向量是零向量时，常考虑它自己内积等于零，即 $X^{\mathrm{T}}X=O$.

（3）本题也可从特征值去论证，见第五章内容.

例 4-27 试证：线性方程组 $A_{s\times n}X_{n\times 1}=b_{s\times 1}$ 有解

$$\Leftrightarrow \quad A^{\mathrm{T}}_{n\times s}Z_{s\times 1}=O \quad \text{与} \quad \begin{pmatrix}A^{\mathrm{T}}\\ b^{\mathrm{T}}\end{pmatrix}_{(n+1)\times s}\cdot Z_{s\times 1}=O \text{ 是同解方程组.}$$

证 充分性 若 $A^{\mathrm{T}}Z=O$ 与 $\begin{pmatrix}A^{\mathrm{T}}\\ b^{\mathrm{T}}\end{pmatrix}Z=O$ 同解，则两个解空间的维数一样，且未知量个数（s）也一样，则有 $R(A^{\mathrm{T}})=R\begin{pmatrix}A^{\mathrm{T}}\\ b^{\mathrm{T}}\end{pmatrix}$，即有 $R(A)=R(A,b)$，则方程组 $AX=b$ 有解.

必要性 若 y 是 $\begin{pmatrix}A^{\mathrm{T}}\\ b^{\mathrm{T}}\end{pmatrix}Z=O$ 的解，即有 $\begin{pmatrix}A^{\mathrm{T}}y\\ b^{\mathrm{T}}y\end{pmatrix}=O$. 则显见 y 也是 $A^{\mathrm{T}}Z=O$ 的解. 另一方面，若 y 是 $A^{\mathrm{T}}Z=O$ 的解，经证 y 也是 $\begin{pmatrix}A^{\mathrm{T}}\\ b^{\mathrm{T}}\end{pmatrix}Z=O$ 的解. 也就是要证 $b^{\mathrm{T}}y=O$.

现已知 $AX=b$ 有解，即知 b 可由 A 的列向量组线性表示，也即 b^{T} 可由 A^{T} 的行向量组线性表示，记为

$$b^{\mathrm{T}}=(x_1,\cdots,x_n)\begin{pmatrix}A_1^{\mathrm{T}}\\ \vdots\\ A_n^{\mathrm{T}}\end{pmatrix} \quad A_i^{\mathrm{T}} \text{ 即为 } A^{\mathrm{T}} \text{ 的第 } i \text{ 行.}$$

则

$$b^{\mathrm{T}}y=(x_1,\cdots,x_n)\begin{pmatrix}A_1^{\mathrm{T}}\\ \vdots\\ A_n^{\mathrm{T}}\end{pmatrix}y=(x_1,\cdots,x_n)A^{\mathrm{T}}y$$

$$=(x_1,\cdots,x_n)\cdot O=O.$$

评注 （1）有解无解的问题，应当考虑有解判别定理.

（2）在充分性的证明中，要由 $A^{\mathrm{T}}y=O$ 去证 $b^{\mathrm{T}}y=O$，则在已知 $AX=b$ 中，转置即得 $b^{\mathrm{T}}=X^{\mathrm{T}}A^{\mathrm{T}}$，然后等号两边右乘 y 即得证.

例 4-28 设四元齐次线性方程组

$$(\mathrm{I}):\begin{cases}x_1+x_2=0\\ x_2-x_4=0\end{cases}$$

又已知某齐次线性方程组（Ⅱ）的通解为

$$k_1\begin{pmatrix}0\\1\\1\\0\end{pmatrix}+k_2\begin{pmatrix}-1\\2\\2\\1\end{pmatrix},$$

（1）求（Ⅰ）的基础解系；

（2）问（Ⅰ）与（Ⅱ）是否有非零公共解？若有，则求出所有的非零公共解. 若没有，请说明理由.

解　(1) 由(Ⅰ),可得

$$\begin{cases} x_1 = -x_2 \\ x_4 = x_2 \end{cases} \quad \text{分别令} \quad \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

则得(Ⅰ)的基础解系为

$$\boldsymbol{\xi}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\xi}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

(2) 有非零公共解.

(Ⅱ)的通解可表为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -k_2 \\ k_1 + 2k_2 \\ k_1 + 2k_2 \\ k_2 \end{pmatrix},$$

把它代入方程组(Ⅰ),得

$$\begin{cases} -k_2 + (k_1 + 2k_2) = 0 \\ (k_1 + 2k_2) - k_2 = 0 \end{cases} \quad \text{解得} \quad k_1 = -k_2.$$

故当 $k_1 = -k_2 \neq 0$ 时,(Ⅱ)的通解化为

$$k_1 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} -1 \\ 2 \\ 2 \\ 1 \end{pmatrix} = k_2 \left[\begin{pmatrix} 0 \\ -1 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \\ 2 \\ 1 \end{pmatrix} \right] = k_2 \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix};$$

即此向量是(Ⅰ)与(Ⅱ)的非零公共解(其中,k_2 是不等于零的任意常数).

评注　本题若(Ⅰ)的基础解系 ξ_1, ξ_2,与(Ⅱ)的基础解系 ξ_3, ξ_4,这四个向量还线性无关的话,则(Ⅰ)与(Ⅱ)就无非零公共解.

例 4-29　已知下列非齐次线性方程组(Ⅰ),(Ⅱ):

$$(\text{Ⅰ}): \begin{cases} x_1 + x_2 - 2x_4 = -6 \\ 4x_1 - x_2 - x_3 - x_4 = 1 \\ 3x_1 - x_2 - x_3 = 3 \end{cases}, \quad (\text{Ⅱ}): \begin{cases} x_1 - mx_2 - x_3 - x_4 = -5 \\ nx_2 - x_3 - 2x_4 = -11 \\ x_3 - 2x_4 = -t + 1 \end{cases}.$$

(1) 求解方程组(Ⅰ),用其导出组的基础解系表示通解;

(2) 当方程组(Ⅱ)中的参数 m, n, t 为何值时,方程组(Ⅰ)与(Ⅱ)同解?

解　(1)

$$(\bar{\boldsymbol{A}}_1) = \begin{pmatrix} 1 & 1 & 0 & -2 & -6 \\ 4 & -1 & -1 & -1 & 1 \\ 3 & -1 & -1 & 0 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & -1 & -2 \\ 0 & 1 & 0 & -1 & -4 \\ 0 & 0 & 1 & -2 & -5 \end{pmatrix},$$

$R(\boldsymbol{A}_1) = R(\bar{\boldsymbol{A}}_1) = 3 <$ 未知量个数 4,故有无穷多解. 其通解为

$$x = \begin{pmatrix} -2 \\ -4 \\ -5 \\ 0 \end{pmatrix} + k \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}, k \text{ 为任意常数}.$$

(2) 将上述通解 x 代入(Ⅱ)的第一个方程：

$$(-2+k)+m(-4+k)-(-5+2k)-k=-5,$$

解得

$$m=2.$$

再将 x 代入(Ⅱ)的第二个方程：

$$n(-4+k)-(-5+2k)-2k=-11,\ \text{解得}\ n=4.$$

再把 x 代入(Ⅱ)的第三个方程：

$$(-5+2k)-2k=-t+1,\ \text{得}\ t=6.$$

因此，方程组(Ⅱ)的参数 $m=2, n=4, t=6$ 时，方程组(Ⅰ)的全部解都是(Ⅱ)的解，此时，(Ⅱ)化为

$$(\text{Ⅱ})\quad \begin{cases} x_1+2x_2-x_3-x_4=-5 \\ 4x_2-x_3-2x_4=-11 \\ x_3-2x_4=-5 \end{cases}.$$

$$(\overline{\boldsymbol{A}}_2)=\begin{pmatrix} 1 & 2 & -1 & -1 & -5 \\ 0 & 4 & -1 & -2 & -11 \\ 0 & 0 & 1 & -2 & -5 \end{pmatrix}\to\begin{pmatrix} 1 & 0 & 0 & -1 & -2 \\ 0 & 1 & 0 & -1 & -4 \\ 0 & 0 & 1 & -2 & -5 \end{pmatrix}.$$

则(Ⅱ)的通解为

$$x=\begin{pmatrix} -2 \\ -4 \\ -5 \\ 0 \end{pmatrix}+k\begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix},\ k\ \text{为任意常数}.$$

可见(Ⅰ)与(Ⅱ)的解完全相同，即(Ⅰ)与(Ⅱ)同解.

例 4-30 设四元线性方程组为

$$(\text{Ⅰ})\quad \begin{cases} 2x_1+3x_2-x_3=0 \\ x_1+2x_2+x_3-x_4=0 \end{cases}$$

且已知另一个四元齐次线性方程组(Ⅱ)的一个基础解系为

$$\boldsymbol{\alpha}_1=\begin{pmatrix} 2 \\ -1 \\ a+2 \\ 1 \end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix} -1 \\ 2 \\ 4 \\ a+8 \end{pmatrix}.$$

(1) 求(Ⅰ)的一个基础解系；

(2) 当 a 为何值时，方程组(Ⅰ)与(Ⅱ)有非零的公共解？在有非零公共解时，求出全部非零公共解.

解法 1 (1)

$$\begin{pmatrix} 2 & 3 & -1 & 0 \\ 1 & 2 & 1 & -1 \end{pmatrix}\to\begin{pmatrix} 1 & 0 & -5 & 3 \\ 0 & 1 & 3 & -2 \end{pmatrix}.$$

解得：

$$\boldsymbol{\xi}_1=\begin{pmatrix} 5 \\ -3 \\ 1 \\ 0 \end{pmatrix},\quad \boldsymbol{\xi}_2=\begin{pmatrix} -3 \\ 2 \\ 0 \\ 1 \end{pmatrix}.$$

(2) 设有非零公共解,则应有数 k_1,k_2,k_3,k_4 使得

$$\boldsymbol{\eta}=k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2=k_3\boldsymbol{\alpha}_1+k_4\boldsymbol{\alpha}_2,$$

即

$$k_1\cdot\begin{pmatrix}5\\-3\\1\\0\end{pmatrix}+k_2\cdot\begin{pmatrix}-3\\2\\0\\1\end{pmatrix}=k_3\cdot\begin{pmatrix}2\\-1\\a+2\\1\end{pmatrix}+k_4\cdot\begin{pmatrix}-1\\2\\4\\a+8\end{pmatrix}.$$

即得方程组(Ⅲ):

$$(\text{Ⅲ})\begin{cases}5k_1-3k_2-2k_3+k_4=0\\-3k_1+2k_2+k_3-2k_4=0\\k_1-(a+2)k_3-4k_4=0\\k_2-k_3-(a+8)k_4=0\end{cases}.$$

$$\begin{pmatrix}5&-3&-2&1\\-3&2&1&-2\\1&0&-(a+2)&-4\\0&1&-1&-(a+8)\end{pmatrix}\to\begin{pmatrix}1&0&-1&-4\\0&1&-1&-7\\0&0&-a-1&0\\0&0&0&-a-1\end{pmatrix},$$

由此可知:当 $a\neq-1$ 时,方程组(Ⅲ)仅有零解,故(Ⅰ)与(Ⅱ)无非零公共解. 当 $a=-1$ 时,方程组(Ⅲ)的同解方程组为

$$\begin{cases}k_1=k_3+4k_4\\k_2=k_3+7k_4\end{cases}.$$

令 $k_3=c_1,k_4=c_2$,则得方程组(Ⅰ)与(Ⅱ)的非零公共解为

$$\boldsymbol{\eta}=c_1\cdot\begin{pmatrix}2\\-1\\1\\1\end{pmatrix}+c_2\cdot\begin{pmatrix}-1\\2\\4\\7\end{pmatrix},\quad c_1,c_2\text{ 为不全为零的任意常数.}$$

解法 2 (1)

$$A=\begin{pmatrix}2&3&-1&0\\1&2&1&-1\end{pmatrix}\to\begin{pmatrix}1&0&-5&3\\0&1&3&-2\end{pmatrix},$$

解得(Ⅰ)的一个基础解系:

$$\boldsymbol{\xi}_1=\begin{pmatrix}5\\-3\\1\\0\end{pmatrix},\quad\boldsymbol{\xi}_2=\begin{pmatrix}-3\\2\\0\\1\end{pmatrix}.$$

(2) 设方程组(Ⅱ)的通解为

$$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2=\begin{pmatrix}2k_1-k_2\\-k_1+2k_2\\(a+2)k_1+4k_2\\k_1+(a+8)k_2\end{pmatrix}$$

k_1,k_2 为任意常数.将此通解代入方程组(Ⅰ),得

$$\begin{cases}(a+1)k_1 & =0\\ (a+1)k_1-(a+1)k_2 & =0\end{cases} \tag{1}$$

要使方程组(Ⅰ)与(Ⅱ)有非零公共解,只要关于 k_1,k_2 的方程组(1)有非零解,因为

$$\begin{vmatrix}a+1 & 0\\ a+1 & -(a+1)\end{vmatrix}=-(a+1)^2.$$

所以,当 $a\neq-1$ 时,方程组(Ⅰ)与(Ⅱ)无非零公共解.当 $a=-1$ 时,方程组(1)有非零解,且 k_1,k_2 为不全为零的常数,此时,可得(Ⅰ)与(Ⅱ)的全部非零公共解为

$$\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\end{pmatrix}=k_1\cdot\begin{pmatrix}2\\ -1\\ 1\\ 1\end{pmatrix}+k_2\cdot\begin{pmatrix}-1\\ 2\\ 4\\ 7\end{pmatrix}.$$

评注　讨论两个方程组有非零公共解,常用的方法是把一个方程组的通解代入到另一个方程中去(如解法 2).有时,也可按公共解的定义,即 $\boldsymbol{\eta}=k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$,也有 $\boldsymbol{\eta}=k_3\boldsymbol{\alpha}_1+k_4\boldsymbol{\alpha}_2$(见解法 1),从此去判断这些 k_i(不全为零)的存在性.

例 4-31　已知四阶方阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)$,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 均为四维列向量,其中 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关.$\boldsymbol{\alpha}_1=2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$.若 $\boldsymbol{\beta}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4$,求 $\boldsymbol{AX}=\boldsymbol{\beta}$ 的通解.

解　由题给条件知:$R(\boldsymbol{A})=3$,未知量个数 $n=4$,故基础解系中只有一个线性无关向量.设 $\boldsymbol{AX}=\boldsymbol{\beta}$ 的通解为 $\boldsymbol{x}=\boldsymbol{\eta}^*+c\boldsymbol{\xi}$,其中 $\boldsymbol{\eta}^*$ 为 $\boldsymbol{AX}=\boldsymbol{\beta}$ 的一个特解,$\boldsymbol{\xi}$ 为 $\boldsymbol{AX}=\boldsymbol{O}$ 的一个非零解,c 为任意常数.

由　　$\boldsymbol{\alpha}_1=2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$　　即　　$\boldsymbol{\alpha}_1=2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3+0\cdot\boldsymbol{\alpha}_4$.

即有

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\begin{pmatrix}1\\ -2\\ 1\\ 0\end{pmatrix}=\boldsymbol{O},\text{知 }\boldsymbol{\xi}=\begin{pmatrix}1\\ -2\\ 1\\ 0\end{pmatrix}.$$

又由

$$\boldsymbol{\beta}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\begin{pmatrix}1\\ 1\\ 1\\ 1\end{pmatrix},$$

知

$$\boldsymbol{\eta}^*=\begin{pmatrix}1\\ 1\\ 1\\ 1\end{pmatrix}.$$

所以,$\boldsymbol{AX}=\boldsymbol{\beta}$ 的通解为

$$x=\begin{pmatrix}1\\ 1\\ 1\\ 1\end{pmatrix}+k\begin{pmatrix}1\\ -2\\ 1\\ 0\end{pmatrix},k\text{ 为任意常数.}$$

评注　注意把向量的线性组合写成矩阵乘法的这种方法的熟练运用.

例 4-32　已知平面上三条不同直线的方程分别为

$$L_1: ax + 2by + 3c = 0$$

$$L_2: bx + 2cy + 3a = 0$$

$$L_3: cx + 2ay + 3b = 0$$

试证：这三条直线交于一点的充分必要条件为 $a+b+c=0$.

证法 1　必要性　设三直线交于一点，则线性方程组：

$$\begin{cases} ax + 2by = -3c \\ bx + 2cy = -3a \\ cx + 2ay = -3b \end{cases}$$

有惟一解，即有　$R(\boldsymbol{A}) = R(\bar{\boldsymbol{A}}) = 2$,

所以　$|\bar{\boldsymbol{A}}| = 0$.

$$|\bar{\boldsymbol{A}}| = \begin{vmatrix} a & 2b & -3c \\ b & 2c & -3a \\ c & 2a & -3b \end{vmatrix} = 6(a+b+c)[a^2+b^2+c^2-ab-ac-bc]$$

$$= 3(a+b+c)\cdot[(a-b)^2+(b-c)^2+(c-a)^2]$$

但 $[(a-b)^2+(b-c)^2+(c-a)^2] \neq 0$，**故** $a+b+c=0$.

充分性　由 $a+b+c=0$，**则由上面** $|\bar{\boldsymbol{A}}|$ **的计算可知：** $|\bar{\boldsymbol{A}}| = 0$，**则** $R(\bar{\boldsymbol{A}}) < 3$. **但由于**

$$\begin{vmatrix} a & 2b \\ b & 2c \end{vmatrix} = 2(ac-b^2) = -2[a(a+b)+b^2]$$

$$= -2\left[\left(a+\frac{1}{2}b\right)^2 + \frac{3}{4}b^2\right] \neq 0.$$

知 $R(\boldsymbol{A}) = 2$，**于是** $R(\boldsymbol{A}) = R(\bar{\boldsymbol{A}}) = 2 =$ **未知量个数，即知方程组**(1) **有惟一解，即三直线交于一点.**

证法 2　必要性　设三直线交于一点(x_0, y_0)，**则知** $\begin{pmatrix} x_0 \\ y_0 \\ 1 \end{pmatrix}$ **是** $\boldsymbol{AX} = \boldsymbol{O}$ **的非零解.**

$$\boldsymbol{A} = \begin{pmatrix} a & 2b & 3c \\ b & 2c & 3a \\ c & 2a & 3b \end{pmatrix},$$

于是 $|\boldsymbol{A}| = 0$，**而**

$$|\bar{\boldsymbol{A}}| = \begin{vmatrix} a & 2b & 3c \\ b & 2c & 3a \\ c & 2a & 3b \end{vmatrix} = -3(a+b+c)[(a-b)^2+(b-c)^2+(c-a)^2]$$

可知　$a+b+c=0$.

充分性　考虑方程组：

$$\begin{cases} ax + 2by = -3c \\ bx + 2cy = -3a \\ cx + 2ay = -3b \end{cases} \tag{1}$$

将方程组(1) **的三个方程相加，并由** $a+b+c=0$，**可知方程组**(2) **等价于方程组：**

$$\begin{cases} ax + 2by = -3c \\ bx + 2cy = -3a \end{cases} \tag{2}$$

因为

$$\begin{vmatrix} a & 2b \\ b & 2c \end{vmatrix} = 2(ac-b^2) = -2[a(a+b)+b^2]$$

$$= -[a^2+b^2+(a+b)^2] \neq 0,$$

所以,方程组(2)有惟一解.即方程组(1)有惟一解,即三直线交于一点.

例 4-33 设 ξ_1,ξ_2,ξ_3 是三元线性方程组 $\boldsymbol{AX}=\boldsymbol{\beta}$ 的三个线性无关解,$R(\boldsymbol{A})=1$,则不正确的结论是 (　)

(A) $\boldsymbol{AX}=\boldsymbol{\beta}$ 的通解是 $k_1\xi_1+k_2\xi_2+k_3\xi_3$,其中,$k_1,k_2,k_3$ 是满足 $k_1+k_2+k_3=1$ 的任何数;

(B) $\boldsymbol{AX}=\boldsymbol{0}$ 的通解是 $k_1\xi_1+k_2\xi_2+k_3\xi_3$,其中 k_1,k_2,k_3 是满足 $k_1+k_2+k_3=0$ 的任何数;

(C) $\xi_1,\xi_2-\xi_3$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系;

(D) $\xi_1-\xi_2,\xi_2-\xi_3$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系.

解 选(C).因为 ξ_1 不是 $\boldsymbol{AX}=\boldsymbol{0}$ 的解.故 $\xi_1,\xi_2-\xi_3$ 不是 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系.

(A) 正确　因为 $k_1+k_2+k_3=1$,故 $k_3=1-k_1-k_2$,则 $k_1\xi_1+k_2\xi_2+k_3\xi_3=k_1(\xi_1-\xi_3)+k_2(\xi_2-\xi_3)+\xi_3$.$\xi_1-\xi_3$ 与 $\xi_2-\xi_3$ 都是 $\boldsymbol{AX}=\boldsymbol{0}$ 的解,且线性无关,所以(A)正确.

(B) 正确　因为 $k_1+k_2+k_3=0$,故 $k_3=-k_1-k_2$,则 $k_1\xi_1+k_2\xi_2+k_3\xi_3=k_1(\xi_1-\xi_3)+k_2(\xi_2-\xi_3)$,而 $\xi_1-\xi_3,\xi_2-\xi_3$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系.

(D) 正确　因为 $\xi_1-\xi_2,\xi_2-\xi_3$ 都是 $\boldsymbol{AX}=\boldsymbol{0}$ 的解.且线性无关,又因 $R(\boldsymbol{A})=1,n=3$,故它们是 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系.

例 2-34 设 $\boldsymbol{B}$ 是秩为 2 的 5×4 型矩阵.

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\\2\\3\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}-1\\1\\4\\-1\end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix}5\\-1\\-8\\9\end{pmatrix}$$

其齐次线性方程组 $\boldsymbol{BX}=\boldsymbol{0}$ 的三个解向量.求 $\boldsymbol{BX}=\boldsymbol{0}$ 的解空间的一个标准正交基.

解 已知 $R(\boldsymbol{B})=2$,未知量个数 $n=4$.故 $\boldsymbol{BX}=\boldsymbol{0}$ 的基础解系有 $4-2=2$ 个线性无关解向量.

现先求出 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的极大无关组:

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=\begin{pmatrix}1&-1&5\\1&1&-1\\2&4&-8\\3&-1&9\end{pmatrix}\xrightarrow{\text{行}\cdots}\begin{pmatrix}1&-1&5\\0&2&-6\\0&0&0\\0&0&0\end{pmatrix}$$

可知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的极大无关组,是 $\boldsymbol{BX}=\boldsymbol{0}$ 的基础解系.再用 Schmidt 正交化方法:

令

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\\2\\3\end{pmatrix},\quad \boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1=\begin{pmatrix}-\frac{4}{3}\\\frac{2}{3}\\\frac{10}{3}\\-2\end{pmatrix}.$$

再单位化:令

$$\boldsymbol{\gamma}_1=\frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|}=\begin{pmatrix}\frac{1}{\sqrt{15}}\\\frac{1}{\sqrt{15}}\\\frac{2}{\sqrt{15}}\\\frac{3}{\sqrt{15}}\end{pmatrix},\quad \boldsymbol{\gamma}_2=\frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|}=\begin{pmatrix}\frac{-2}{\sqrt{39}}\\\frac{1}{\sqrt{39}}\\\frac{5}{\sqrt{39}}\\\frac{-3}{\sqrt{39}}\end{pmatrix}$$

则 γ_1,γ_2 就是 $BX=0$ 的解空间的一个标准正交基.

例 4-35　设线性方程组 $A_{n\times n}\cdot X_{n\times 1}=b_{n\times 1}$. 若 $R(A)=R\begin{pmatrix}A & b\\ b^{\mathrm{T}} & 0\end{pmatrix}$,试证:这个方程组有解:

证　记$(\bar{A})=(A,b)$,

$$C=\begin{pmatrix}A & b\\ b^{\mathrm{T}} & 0\end{pmatrix}=\begin{pmatrix}\bar{A} & \\ b^{\mathrm{T}} & 0\end{pmatrix}.$$

显然有　　$R(\bar{A})\leqslant R(C)$,以及 $R(A)\leqslant R(\bar{A})$

则有　　$R(A)\leqslant R(\bar{A})\leqslant R(C)$

现已知　　$R(A)=R(C)$

故　　$R(A)=R(\bar{A})$,

即方程组 $AX=b$ 有解.

例 4-36　已知

$$B=\begin{pmatrix}1 & -2 & 1 & 0 & 0\\ 1 & -2 & 0 & 1 & 0\\ 1 & -2 & 3 & -2 & 0\\ 5 & -6 & 0 & 0 & 1\end{pmatrix}$$ 的行向量是方程组:

$$\begin{cases}x_1+x_2+x_3+x_4+x_5=0\\ 3x_1+2x_2+x_3+x_4-3x_5=0\\ x_2+2x_3+2x_4+6x_5=0\\ 5x_1+4x_2+3x_3+3x_4-x_5=0\end{cases}$$ 的解向量,试问:

(1) B 的 4 个行向量是否构成齐次方程组的基础解系;

(2) 若不是,不用解方程组的方法. 求一个基础解系.

解　(1) 先求 $R(B)$. 用初等变换化为阶梯型,易知 $R(B)=3$,故知 B 的四个行向量线性相关,它们不能作为齐次方程组的基础解系.

(2) 设此齐次方程组为 $AX=0$,容易算得 $R(A)=2$,故知基础解系中有 $n-r=5-2=3$ 个线性无关解.

可验证 B 的第一,第二,第四行这三个向量线性无关. 它们可作为 $AX=0$ 的一个基础解系.

例 4-37　设 A 是 $n\times s$ 型矩阵. 证明:存在 $s\times p$ 型矩阵 $B\neq 0$,使 $AB=0$ 的充分必要条件是 A 的 s 个列向量是线性相关的.

证　"⇒"

$AB=0\Leftrightarrow AB_i=0,\ i=1,\cdots,p$

这里 B_i 是 B 的第 i 列. 由 $B\neq 0$,知必有 $B_j\neq 0$. 即知 $AX=0$ 有非零解,则 $R(A)<s$,即 A 的 s 个列向量线性相关.

"⇐"

A 的 s 个列向量线性相关,可知 $R(A)<s$. 则 $AX=0$ 有非零解,选择 p 个非零解,记为 $B_1,\cdots,B_p$,把它们排成矩阵 B,即得 $AB=0$.

四、小　结

(1) 线性方程组这一部分主要是两个大定理:①线性方程组 $AX=b$ 有解判别定理. 即看 $R(A)\overset{?}{=}R(\bar{A})$. ②齐次方程组 $AX=0$ 的解空间的维数 $\dim(\eta(A))=n-R(A)$.

一般说,讨论解的情况,总从 $R(A)\overset{?}{=}R(\bar{A})$ 去考虑. 若涉及解的结构. 总应考虑 $\dim(\eta(A))=n-R(A)$ 以及相关的一些性质.

(2) 在论证 $R(A)\overset{?}{=}R(\bar{A})$ 时,一个常用的有效的方法是把增广矩阵 $\bar{A}$ 经过初等行变换化到阶梯型. 然后在阶梯型便可看清 $R(A)$ 与 $R(\bar{A})$ 的关系. 同时,在求解线性方程组时,也要用到这个方法. 因此,把矩阵经过初等行变换化到阶梯型是这一部分的重要的基本功,必须掌握得很好,特别是对带有参数的矩阵化为阶梯型不太容易做对,要重视.

第五章　矩阵的特征值与特征向量

矩阵的特征值与特征向量问题不仅是线性代数中的重要内容，也是历年硕士研究生入学考试的重点.显含或隐含的特征值和特征向量问题层出不穷.

一、复习与考试要求

(1) 理解矩阵的特征值和特征向量的概念及性质.

(2) 了解相似矩阵的概念、性质，矩阵可对角化的条件以及实对称阵相似于对角阵的理论.

(3) 熟练掌握求矩阵特征值和特征向量的方法.

(4) 熟练掌握实对称阵正交相似于对角阵的方法.

二、基本概念与理论

1. 矩阵的对角化

矩阵 A，若能相似变换到对角阵，就称 A 可以对角化，否则，就称 A 不能对角化.

2. 相似矩阵

若存在可逆阵 U，使 $U^{-1}AU=B$，则称 A 与 B 相似. 这里，A,B 皆为 n 阶方阵.

3. 相似矩阵的性质

(1) 相似矩阵有相同的行列式.

(2) 相似矩阵有相同的可逆性，且可逆时，它们的逆矩阵也相似.

(3) 相似矩阵有相同的秩.

(4) 相似矩阵有相同的特征多项式.

(5) 相似矩阵有相同的特征值.

(6) 相似矩阵有相同的迹.(矩阵 A 的迹，即为 $\mathrm{t}_r(\boldsymbol{A})=\sum\limits_{i=1}^{n}a_{ii}$).

(7) 若 $U^{-1}AU=B$，则 $U^{-1}A^kU=B^k$，$(k\in Z^+)$.

(8) 若 $U^{-1}AU=B$，则 $f(A)$ 与 $f(B)$ 也相似.

(9) 若 $U^{-1}AU=B$，则 $A+E$ 与 $B+E$ 也相似.

4. 特征值与特征向量

(1) 定义　n 阶方阵 A，$\lambda\in$ 数域 F，非零向量 α. 若 $A\alpha=\lambda\alpha$，则称数 λ 是 A 的特征值，非零向量 α 是 A 的属于 λ 的特征向量.

① 由定义可知，若 α 是 A 的属于 λ 的特征向量，则 $k\alpha(k\neq0,k\in F)$ 也是 A 的属于 λ 的特征向量.

② 特征向量 α 是属于某个特征值的，它不能同时分属于两个不同的特征值.

(2) 定理　数域 F 上的 n 阶矩阵 A 可以对角化$\Leftrightarrow A$ 有 n 个线性无关的特征向量.

这是因为

$$A\text{ 可以对角化}\Leftrightarrow U^{-1}AU=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$$

$$\Leftrightarrow AU=U\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$$

$$\Leftrightarrow \boldsymbol{A}(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$$

$\Leftrightarrow \boldsymbol{A}\boldsymbol{\alpha}_i=\lambda_i\boldsymbol{\alpha}_i \quad (i=1,2,\cdots,n)$.

式中 $\boldsymbol{U}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)$. $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 是 n 个线性无关的特征向量. $\lambda_1,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值.

(3) 求特征值 λ 与特征向量 α.

$$\boldsymbol{A}\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha} \Leftrightarrow (\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{\alpha}=\boldsymbol{0}$$

$$\overset{\boldsymbol{\alpha}\neq\boldsymbol{0}}{\Leftrightarrow} |\lambda\boldsymbol{E}-\boldsymbol{A}|=0.$$

因此,特征值就是特征方程 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$ 的根. 属于某个特征值 λ_0 的特征向量,就是求解齐次线性方程组 $(\lambda_0\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系.

(4) 判断 n 阶矩阵 $\boldsymbol{A}$ 是否有 n 个线性无关的特征向量?

① 定理:$\boldsymbol{A}$ 的不同特征值所对应的特征向量是线性无关的.

② 推论 1:若 n 阶矩阵 $\boldsymbol{A}$ 有 n 个不同特征值,则 $\boldsymbol{A}$ 必可对角化.

③ 推论 2:若特征值 λ_i 对应 s 个线性无关的特征向量,特征值 λ_j 对应 t 个线性无关的特征向量,则这 $s+t$ 个特征向量还是线性无关的.

要注意,这里的定理和推论都只是充分条件,而不是充要条件.

5. 实对称矩阵必可对角化

(1) 实对称阵的性质:

① 实对称阵的特征值全是实数;

② 实对称阵的不同特征值对应的特征向量是正交的.

(2) 定理:对于实对称阵 $\boldsymbol{A}$,必存在正交阵 $\boldsymbol{T}$,使 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{D}$. 这里 $\boldsymbol{D}$ 为对角阵.

(3) 施密特正交化方法:把一组线性无关的向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$,通过下列公式化为一组与之等价的正交的向量组 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$.

$$\begin{cases}\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1 \\ \beta_2=\boldsymbol{\alpha}_2-\dfrac{(\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_r=\boldsymbol{\alpha}_r-\dfrac{(\boldsymbol{\beta}_1,\boldsymbol{\alpha}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\cdots-\dfrac{(\boldsymbol{\beta}_{r-1},\boldsymbol{\alpha}_r)}{(\boldsymbol{\beta}_{r-1},\boldsymbol{\beta}_{r-1})}\cdot\boldsymbol{\beta}_{r-1}.\end{cases}$$

6. n 阶矩阵 $\boldsymbol{A}$ 与特征值的关系

(1) $\lambda_1+\lambda_2+\cdots+\lambda_n=t_r(\boldsymbol{A})=\sum\limits_{i=1}^{n}a_{ii}$.

(2) $\lambda_1\cdot\lambda_2\cdot\cdots\cdot\lambda_n=|\boldsymbol{A}|$.

(3) 若可逆阵 $\boldsymbol{A}$ 有特征值 λ,则 $\boldsymbol{A}^{-1}$ 有特征值 $\dfrac{1}{\lambda}$,$\boldsymbol{A}^*$ 有特征值 $\dfrac{1}{\lambda}\cdot|\boldsymbol{A}|$.

(4) 若 $\boldsymbol{A}$ 有特征值 λ,则 $\boldsymbol{A}^n$ 有特征值 λ^n.

(5) 实对称阵 $\boldsymbol{A}$ 的特征值全是实数.

(6) 实反对称阵 $\boldsymbol{A}$ 的特征值只能是零或纯虚数.

(7) 正交阵的特征值的模必为 1.

(8) 正定阵的特征值全大于零.

(9) 上(下)三角阵的特征值即为其对角线元素.

三、基本题型与解题方法

例 5-1　求矩阵 $\boldsymbol{A}=\begin{bmatrix}-2 & 1 & 1\\ 0 & 2 & 0\\ -4 & 1 & 3\end{bmatrix}$ 的特征值和特征向量.

分析 这是最基本的特征值问题，只需按定义求解即可.

解 $|\boldsymbol{A}-\lambda\boldsymbol{E}|=\begin{bmatrix}-2-\lambda & 1 & 1\\ 0 & 2-\lambda & 0\\ -4 & 1 & 3-\lambda\end{bmatrix}=\cdots=-(\lambda+1)(\lambda-2)^2$，

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=-1,\lambda_2=\lambda_3=2$.

当 $x_1=-1$ 时，解齐次线性方程组 $(\boldsymbol{A}+\boldsymbol{E})\boldsymbol{X}=0$，由

$$\boldsymbol{A}+\boldsymbol{E}=\begin{bmatrix}-1 & 1 & 1\\ 0 & 3 & 0\\ -4 & 1 & 4\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}$$

得基础解系 $\xi_1=\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}$，

所以对应于 $\lambda_1=-1$ 的全部特征向量为 $k_1\xi_1(k\neq0)$.

当 $\lambda_2=\lambda_3=2$ 时，解齐次线性方程组 $(\boldsymbol{A}-2\boldsymbol{E})\boldsymbol{X}=\boldsymbol{0}$. 由

$$\boldsymbol{A}-2\boldsymbol{E}=\begin{bmatrix}-4 & 1 & 1\\ 0 & 0 & 0\\ -4 & 1 & 1\end{bmatrix}\longrightarrow\begin{bmatrix}1 & -1/4 & -1/4\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$

得基础解系 $\xi_2=\begin{bmatrix}0\\ 1\\ -1\end{bmatrix},\xi_3=\begin{bmatrix}1\\ 0\\ 4\end{bmatrix}$，

所以对应于 $\lambda_2=\lambda_3=1$ 的全部特征向量为

$$k_2\xi_2+k_3\xi_3(k_2,k_3\text{ 不同时为 }0),$$

因为 $\boldsymbol{A}$ 有三个线性无关的特征向量 ξ_1,ξ_2,ξ_3，所以 $\boldsymbol{A}$ 可以对角化.

例 5-2 求证 $\boldsymbol{A}$ 为奇异矩阵的充要条件为 0 是 $\boldsymbol{A}$ 的一个特征值.

分析 有关特征值的证明题基本上都是与特征值定义表达式 $\boldsymbol{AX}=\lambda\boldsymbol{X}(\boldsymbol{X}\neq\boldsymbol{0})$ 分不开的. 当 $\lambda=0$ 时，得 $\boldsymbol{AX}=\boldsymbol{0}$，问题便转化为讨论齐次线性方程组有非零解的条件.

证 必要性 设 $\boldsymbol{A}$ 奇异，则 $\boldsymbol{AX}=0$ 有非零解 $\boldsymbol{X}$ 即

$$\boldsymbol{AX}=0\boldsymbol{X}$$

从而 0 是 $\boldsymbol{A}$ 的特征值.

充分性 设 0 是 $\boldsymbol{A}$ 的一个特征值，则有 $\xi\neq0$，使 $\boldsymbol{A}\xi=0,\xi\neq\boldsymbol{0}$，即 $\boldsymbol{AX}=\boldsymbol{0}$ 有非零解，从而 $\boldsymbol{A}$ 是奇异的.

本题更为简单的证明方法是 $\boldsymbol{A}$ 奇异$\Leftrightarrow|\boldsymbol{A}|=\lambda_1\lambda_2\cdots\lambda_n=0\Leftrightarrow\lambda_1,\cdots,\lambda_n$ 中至少有一个特征值 $\lambda_i=0$.

例 5-3 若 $\lambda\neq0$ 是 $\boldsymbol{AB}$ 的特征值，则 λ 必是 $\boldsymbol{BA}$ 的特征值. $(\boldsymbol{A},\boldsymbol{B}$ 可以是奇异阵，也可以是长方矩阵$)$.

分析 若 λ 是 $\boldsymbol{AB}$ 的特征值，则由定义必存在特征向量 $\xi(\xi\neq\boldsymbol{0})$，使 $\boldsymbol{AB}\xi=\lambda\xi$，要证 λ 也是 $\boldsymbol{BA}$ 的特征值，则需证明存在向量 $\eta\neq\boldsymbol{0}$，使 $\boldsymbol{BA}\eta=\lambda\eta$，(未必 $\xi=\eta$). 由此可见将 $\boldsymbol{AB}\xi=\lambda\xi$ 两边左乘 $\boldsymbol{B}$ 才能过渡到 $\boldsymbol{BA}\eta=\lambda\eta$，这时有 $\eta=\boldsymbol{B}\xi$，剩下的问题是要证 $\eta\neq\boldsymbol{0}$，这时要用条件 $\lambda\neq0$.

证 因为 λ 是 $\boldsymbol{AB}$ 的特征值，所以存在列向量 $\xi\neq\boldsymbol{0}$，使

$$\boldsymbol{AB}\xi=\lambda\xi,$$

上式两边左乘 $\boldsymbol{B}$ 得

$$\boldsymbol{BAB}\xi=\lambda\boldsymbol{B}\xi,$$

令 $\eta=\boldsymbol{B}\xi$. 这时 $\eta\neq\boldsymbol{0}$. 否则有 $\lambda\xi=\boldsymbol{AB}\xi=\boldsymbol{0}$，而 $\xi\neq\boldsymbol{0}$，则得到 $\lambda=0$，这与已知条件矛盾，故 $\eta\neq\boldsymbol{0}$，这表明

$$\boldsymbol{BA}\eta=\lambda\eta$$

成立，所以 λ 也是 $\boldsymbol{BA}$ 的特征值.

例 5-4 已知矩阵 $\boldsymbol{A}=\begin{bmatrix}2 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & x\end{bmatrix}$ 与 $\boldsymbol{B}=\begin{bmatrix}2 & 0 & 0\\ 0 & y & 0\\ 0 & 0 & -1\end{bmatrix}$ 相似.

(1) 求 x 与 y；

(2) 求一个满足 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$ 的可逆阵 $\boldsymbol{P}$.

分析　这是 1988 年数学(试卷一)的 8 分题. 由于 $\boldsymbol{A}$ 已经相似于对角阵 $\boldsymbol{B}$，所以实际上已给出了 $\boldsymbol{A}$ 的两个特征值，这样就避免了需解三次方程来求特征值的困难，而将考试目的转向了其他方向：利用相似的概念求参数 x,y 以及求 $\boldsymbol{A}$ 的特征向量.

解

(1) 因为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似，所以有

$$|\boldsymbol{A}-\lambda\boldsymbol{E}|=|\boldsymbol{B}-\lambda\boldsymbol{E}|,$$

可以得到　$(2-\lambda)(x^2-x\lambda-1)=(2-\lambda)(\lambda-y)(\lambda+1)$,

比较等式两端的系数，得 $x=0,y=1$.

(2) 由(1)知矩阵 $\boldsymbol{A}$ 的特征值为 $2,1,-1$.

当 $\lambda_1=2$ 时，由

$$(\boldsymbol{A}-\lambda_1\boldsymbol{E})\boldsymbol{X}=\begin{bmatrix}0&0&0\\0&-2&1\\0&1&-2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=0,$$

解得　$\xi_1=(1,0,0)^{\mathrm{T}}$，这是对应于 $\lambda_1=2$ 的一个特征向量.

当 $\lambda_2=1$ 时，由

$$(\boldsymbol{A}-\lambda_2\boldsymbol{E})\boldsymbol{X}=\begin{bmatrix}1&0&0\\0&-1&1\\0&1&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\boldsymbol{0},$$

解得对应于 $\lambda_2=1$ 的特征向量 $\xi_2=(0,1,1)^{\mathrm{T}}$.

当 $\lambda_3=-1$ 时，由

$$(\boldsymbol{A}-\lambda_3\boldsymbol{E})\boldsymbol{X}=\begin{bmatrix}1&0&0\\0&1&1\\0&1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\boldsymbol{0},$$

解得对应于 $\lambda_3=-1$ 的一个特征向量 $\xi_3=(0,1,-1)^{\mathrm{T}}$.

令

$$\boldsymbol{P}=\begin{bmatrix}1&0&0\\0&1&1\\0&1&-1\end{bmatrix},$$

因为 $\boldsymbol{A}$ 的特征值互不相同，所以对应的特征向量 ξ_1,ξ_2,ξ_3 线性无关.(进一步，由于 $\boldsymbol{A}$ 是实对称阵，故 ξ_1,ξ_2,ξ_3 两两正交)，从而矩阵 $\boldsymbol{P}$ 可逆且满足　$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$.

例 5-5　设三阶实对称阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1=-1,\lambda_2=\lambda_3=1$，对应于 λ_1 的特征向量为 $\xi_1=(0,1,1)^{\mathrm{T}}$，求 $\boldsymbol{A}$.

分析　这是 1995 年数学(试卷一) 的 7 分题. 因为要求的三阶方阵 $\boldsymbol{A}$ 是实对称的，所以 $\boldsymbol{A}$ 一定有三个两两正交的单位特征向量. 只要求出这三个特征向量，再由给出的特征值即可求出 $\boldsymbol{A}$. 从已给出的 $\xi_1=(0,1,1)^{\mathrm{T}}$ 出发，可以求出与 ξ_1 正交的对应于 $\lambda_2=\lambda_3=1$ 的两个互相正交的特征向量 $\xi_2=(1,0,0)^{\mathrm{T}}$，$\xi_2=(0,1,-1)^{\mathrm{T}}$(这可通过解线性方程组 $(0,1,1)(x_1,x_2,x_3)^{\mathrm{T}}=\boldsymbol{0}$ 得出).

解　从上面的分析已经得到了 $\boldsymbol{A}$ 的三个互相正交的特征向量 ξ_1,ξ_2,ξ_3，再将它们单位化，故取

$$P=\begin{bmatrix}0&1&0\\1/\sqrt{2}&0&1/\sqrt{2}\\1/\sqrt{2}&0&-1/\sqrt{2}\end{bmatrix},$$

从而 $\boldsymbol{P}^{-1}=\boldsymbol{P}^{\mathrm{T}}$.

所以得

$$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{P}\Lambda\boldsymbol{P}^{-1}$$

$$=\begin{bmatrix}0&1&0\\1/\sqrt{2}&0&1/\sqrt{2}\\1/\sqrt{2}&0&-1/\sqrt{2}\end{bmatrix}\begin{bmatrix}-1&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}0&1/\sqrt{2}&1/\sqrt{2}\\1&0&0\\0&1/\sqrt{2}&-1/\sqrt{2}\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&-1\\0&-1&0\end{bmatrix}.$$

例 5-6 设 $\boldsymbol{A}$ 为 n 阶正交的对称正定阵,求证 $\boldsymbol{A}$ 必为单位矩阵.

分析 由 $\boldsymbol{A}$ 为对称阵的条件,可以断言 $\boldsymbol{A}$ 正交相似于对角阵,且对角线元素全为实数. 再由正定阵的条件得出 $\boldsymbol{A}$ 的特征值全为正数. 加上 $\boldsymbol{A}$ 是正交阵的条件,最后证明出 $\boldsymbol{A}$ 的特征值全为 1,即 $\boldsymbol{A}$ 正交相似于单位阵,从而 $\boldsymbol{A}$ 是单位阵.

证 因为 $\boldsymbol{A}$ 是实对称阵,所以 $\boldsymbol{A}$ 的特征值全为实数. 又 $\boldsymbol{A}$ 是正定阵,则 $\boldsymbol{A}$ 的特征值全为正数. 设有非零列向量 $\boldsymbol{X}$

使

$$\boldsymbol{A}\boldsymbol{X}=\lambda\boldsymbol{X},$$

由 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$,得到 $\boldsymbol{A}^2=\boldsymbol{E}$. 从 $\boldsymbol{A}\boldsymbol{X}=\lambda\boldsymbol{X}$,两边左乘 $\boldsymbol{A}$ 得 $\boldsymbol{A}^2\boldsymbol{X}=\lambda\boldsymbol{A}\boldsymbol{X}=\lambda^2\boldsymbol{X},\lambda^2\boldsymbol{X}=\boldsymbol{X}$,故 $\lambda^2=1$,从而 $\boldsymbol{A}$ 的特征值只能是 1.

由 $\boldsymbol{A}$ 对称,则存在正交阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{E}$,所以

$$\boldsymbol{A}=\boldsymbol{E}.$$

例 5-7 设 $\boldsymbol{A}=\begin{bmatrix}1&0&2\\0&1&0\\0&0&-1\end{bmatrix}$,求 $\boldsymbol{A}^9$.

分析 本题看上去似乎与特征值无关,但若硬求 $\boldsymbol{A}^9$,计算量必然很大. 这类问题的一般解法是先将 $\boldsymbol{A}$ 相似于对角形 Λ,再求 $\boldsymbol{P}$,使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\Lambda$,这样得 $\boldsymbol{A}=\boldsymbol{P}\Lambda\boldsymbol{P}^{-1},\boldsymbol{A}^9=\boldsymbol{P}\Lambda^9\boldsymbol{P}^{-1}$.

解法 1 先求出 $\boldsymbol{A}$ 的特征值 $\lambda_1=\lambda_2=1,\lambda_3=-1$.

当 $\lambda_1=\lambda_2=1$ 时,由 $(\boldsymbol{A}-\boldsymbol{E})\boldsymbol{X}=\boldsymbol{0}$ 解出对应于 $\lambda_1=\lambda_2=1$ 的两个线性无关的特征向量

$$\xi_1=(1,0,0)^{\mathrm{T}},\xi_2=(1,-1,0)^{\mathrm{T}}$$

当 $\lambda_3=-1$ 时,由 $(\boldsymbol{A}+\boldsymbol{E})\boldsymbol{X}=\boldsymbol{0}$ 解出对应于 $\lambda_3=-1$ 的特征向量

$$\xi_3=(1,0,-1)^{\mathrm{T}},$$

令 $\boldsymbol{P}=\begin{bmatrix}1&1&1\\0&-1&0\\0&0&-1\end{bmatrix}$,得 $\boldsymbol{P}^{-1}=\begin{bmatrix}1&1&1\\0&-1&0\\0&0&-1\end{bmatrix}$,

这时有 $\boldsymbol{A}=\boldsymbol{P}^{-1}\Lambda\boldsymbol{P}$

$$\boldsymbol{A}^9=\boldsymbol{P}^{-1}\Lambda^9\boldsymbol{P}=\begin{bmatrix}1&1&1\\0&-1&0\\0&0&-1\end{bmatrix}\begin{bmatrix}1&&\\&1&\\&&-1\end{bmatrix}^9\begin{bmatrix}1&1&1\\0&-1&0\\0&0&-1\end{bmatrix}$$

$$=\begin{bmatrix}\boldsymbol{A}_1&\boldsymbol{A}_2\\0&-1\end{bmatrix}\begin{bmatrix}\boldsymbol{E}&0\\0&-1\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_1&\boldsymbol{A}_2\\0&-1\end{bmatrix}$$

$$=\begin{bmatrix}\boldsymbol{A}_1&-\boldsymbol{A}_2\\0&1\end{bmatrix}\begin{bmatrix}-\boldsymbol{A}_2&\boldsymbol{A}_2\\0&-1\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_1^2&\boldsymbol{A}_1\boldsymbol{A}_2+\boldsymbol{A}_2\\0&-1\end{bmatrix}$$

$$=\begin{bmatrix}1&0&2\\0&1&0\\0&0&-1\end{bmatrix},$$

其中 $\boldsymbol{A}_1=\begin{bmatrix}1 & 1\\ 0 & -1\end{bmatrix}$，$\boldsymbol{A}_2=\begin{bmatrix}1\\ 0\end{bmatrix}$.

解法 2　因为 $\boldsymbol{A}^2=\boldsymbol{E}$，所以 $\boldsymbol{A}^9=\boldsymbol{A}$.

评注　这是试探法，先求 $\boldsymbol{A}^2$，$\boldsymbol{A}^3$，再从中找规律. 这种方法只在少数特殊情形才可以用得上.

例 5-8　已知 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 n 阶方阵 $\boldsymbol{a}=(a_{ij})$ 的 n 个特征值，试证 $\lambda_1^2+\lambda_2^2+\cdots+\lambda_n^2=\sum\limits_{i,j=1}^{n}a_{ij}a_{ji}$.

分析　本题初一看很难找到思路，但从特征值定义的表达式 $\boldsymbol{AX}=\lambda\boldsymbol{X}(\boldsymbol{X}\neq\boldsymbol{0})$，易得 $\boldsymbol{A}^2\boldsymbol{X}=\lambda^2\boldsymbol{X}$，这表明如果 λ 是 $\boldsymbol{A}$ 的特征值，则 λ^2 是 $\boldsymbol{A}^2$ 的特征值，另外 $\boldsymbol{A}=(a_{ij})$ 的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 与 $\boldsymbol{A}$ 的元素有以下关系，$\lambda_1+\lambda_2+\cdots+\lambda_n=\sum\limits_{i=1}^{n}a_{ii}$ 这两件事启发我们只要证明 $\sum\limits_{i,j=1}^{n}a_{ij}a_{ji}$ 就是 $\boldsymbol{A}^2$ 的对角元素之和即可完成本题的证明.

证　因为 $\lambda_1,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值，所以 $\lambda_1^2,\cdots,\lambda_n^2$ 是 $\boldsymbol{A}^2$ 的 n 个特征值.

又因为 $\lambda_1+\lambda_2+\cdots+\lambda_n=\sum\limits_{i=1}^{n}a_{ii}=\mathrm{tr}(\boldsymbol{A})$，

所以 $$\lambda_1^2+\lambda_2^2+\cdots+\lambda_n^2=\mathrm{tr}(\boldsymbol{A}^2),$$

而 $$\mathrm{tr}(\boldsymbol{A}^2)=\sum_{i,j=1}^{n}a_{ij}a_{ji}.$$

例 5-9　$\boldsymbol{A}$ 是实反对称阵. 证明：

(1) $\boldsymbol{A}$ 的特征值只能是零或纯虚数；

(2) $|\boldsymbol{E}-\boldsymbol{A}|\neq0$，　$|\boldsymbol{E}+\boldsymbol{A}|\neq0$.

证　(1)设 $\boldsymbol{AX}=\lambda\boldsymbol{X}$，$\lambda$ 可能复数，$\boldsymbol{X}$ 也可能是复向量. 两边共轭转置：$\overline{\boldsymbol{X}}^{\mathrm{T}}\overline{\boldsymbol{A}}^{\mathrm{T}}=\overline{\lambda}\,\overline{\boldsymbol{X}}^{\mathrm{T}}$，由于 $\overline{\boldsymbol{A}}=\boldsymbol{A}$，$\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$，故上式即为 $-\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{A}=\overline{\lambda}\,\overline{\boldsymbol{X}}^{\mathrm{T}}$，两边右乘 $\boldsymbol{X}$，得 $-\lambda\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}=\overline{\lambda}\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}$，

即 $$(\lambda+\overline{\lambda})\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{0},$$

记 $\boldsymbol{X}=\begin{pmatrix}X_1\\ \vdots\\ X_n\end{pmatrix}$，其中，$X_i$ 均为复数，则 $\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}=\sum\limits_{i=1}^{n}|X_i|^2\neq0$，

所以 $(\lambda+\overline{\lambda})=0$，即 $\lambda=0$. 或 λ 是纯虚数.

(2) 若 $|\boldsymbol{E}-\boldsymbol{A}|=0\Rightarrow1$ 是 $\boldsymbol{A}$ 的特征值. 这与(1)所证结果矛盾. 同样，若 $|E+\boldsymbol{A}|=0\Rightarrow-1$ 是 $\boldsymbol{A}$ 的特征值. 也与题设矛盾.

例 5-10　$\boldsymbol{A}$ 是正交阵，则 $\boldsymbol{A}$ 的特征值的模必为 1.

证　设 $\boldsymbol{AX}=\lambda\boldsymbol{X}$，$\lambda$ 与 $\boldsymbol{X}$ 都可能是复的. 两边共轭转置

$$\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}=\overline{\lambda}\,\overline{\boldsymbol{X}}^{\mathrm{T}}$$

两边右乘 $\boldsymbol{AX}$ $$\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\overline{\lambda}\,\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{AX}$$

即 $$\overline{X}^{\mathrm{T}}\boldsymbol{X}=\lambda\overline{\lambda}\,\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}$$

即有 $$(1-\lambda\overline{\lambda})\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{0}$$

由于 $\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}\neq0$，故 $1-\lambda\overline{\lambda}=0\Rightarrow|\lambda|^2=1\Rightarrow|\lambda|=1$.

评注　① 对于抽象矩阵讨论特征值，经常中从定义式 $\boldsymbol{AX}=\lambda\boldsymbol{X}$ 去考虑. 除了实对称阵(以及正定阵)，这里 λ 都有可能是复数.

② 在论证过程中，在等式两边同作一种运算. 对于复向量，常先考虑作共轭转置，然后运算. 在例 5-9 中，两边右乘 X；在例 5-10 题中，两边右乘了 $\boldsymbol{AX}$，这要视题目的条件不同而用不同的运算.

例 5-11　设 $\boldsymbol{A}$，$\boldsymbol{B}$ 为 n 阶矩阵，且 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，$\boldsymbol{E}$ 为 n 阶单位阵. 则　(　)

(A) $\lambda\boldsymbol{E}-\boldsymbol{A}=\lambda\boldsymbol{E}-\boldsymbol{B}$；

(B) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值和特征向量；

(C) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都相似于一个对角矩阵；

(D) 对任意常数 t，$t\boldsymbol{E}-\boldsymbol{A}$ 与 $t\boldsymbol{E}-\boldsymbol{B}$ 相似.

解　选(D).

(A)，(B)，(C)结论都明显不对.

由 $\boldsymbol{U}^{-1}\boldsymbol{AU}=\boldsymbol{B}\Rightarrow\boldsymbol{U}^{-1}(t\boldsymbol{E}-\boldsymbol{A})\boldsymbol{U}=\boldsymbol{U}^{-1}t\boldsymbol{E}\boldsymbol{U}-\boldsymbol{U}^{-1}\boldsymbol{AU}=t\boldsymbol{E}-\boldsymbol{B}$. 所以，$t\boldsymbol{E}-\boldsymbol{A}$ 与 $t\boldsymbol{E}-\boldsymbol{B}$ 相似.

例 5-12 设 A 是 n 阶下三角阵，且 $a_{11}=\cdots=a_{nn}=a$ 且至少有一元素 $a_{ij}\neq 0(i>j)$，则 A 不能对角化.

证 （反证法）不妨设 $a_{n1}=b\neq 0$，即

$$\boldsymbol{A}=\begin{pmatrix} a & & & \\ & a & & \\ & & \ddots & \\ b & & & a \end{pmatrix}$$

若 A 可以对角化，则存在可逆阵 U，使

$$\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\begin{pmatrix} a & & & \\ & a & & \\ & & \ddots & \\ & & & a \end{pmatrix}，即有\ \boldsymbol{U}^{-1}\boldsymbol{A}=\begin{pmatrix} a & & & \\ & a & & \\ & & \ddots & \\ & & & a \end{pmatrix}\boldsymbol{U}^{-1}$$

记 $\boldsymbol{U}^{-1}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)$，即有

$$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)\begin{pmatrix} a & & & \\ & a & & \\ & & \ddots & \\ b & & & a \end{pmatrix}=\begin{pmatrix} a & & & \\ & a & & \\ & & \ddots & \\ & & & a \end{pmatrix}(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n).$$

注意 等号右端的分块矩阵的乘法不能运算，但数量阵与任何矩阵乘法可交换.

$$=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)\begin{pmatrix} a & & & \\ & a & & \\ & & \ddots & \\ & & & a \end{pmatrix}$$

可得 $a\cdot\boldsymbol{\alpha}_1+b\cdot\boldsymbol{\alpha}_n=a\cdot\boldsymbol{\alpha}_1,\Rightarrow b\cdot\boldsymbol{\alpha}_n=0,\Rightarrow b=0$. 与已知矛盾.

例 5-13 设

$$\boldsymbol{A}=\begin{pmatrix} 1 & -2 & -4 \\ -2 & x & -2 \\ -4 & -2 & 1 \end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix} 5 & & \\ & y & \\ & & -4 \end{pmatrix}.$$

A 与 B 相似，求 x,y.

解 相似矩阵有相同的迹：$2+x=1+y$.（1）相似矩阵有相同的行列式：$|A|=|B|$.（2）相似矩阵有相同的特征值，故 -4 是 A 的特征值，因此有 $|-4\boldsymbol{E}-\boldsymbol{A}|=0$.

在上述三式中，只要任选 2 个，就可解出 x 和 y.

$$x=4,y=5.$$

例 5-14 用观察法判断：

（1）矩阵 A,B,C 能否对角化？

（2）矩阵 G 的特征值（四选一）；

（3）矩阵 H 的 4 个特征值.

其中

$$\boldsymbol{A}=\begin{pmatrix} 1 & 1 & -1 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 2 \end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix} -1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix},$$

$\boldsymbol{C}=\boldsymbol{E}_3-2\boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}}$（$\boldsymbol{E}_3$ 为三阶单位阵，$\boldsymbol{\beta}$ 为三维实列向量.）

$$\boldsymbol{G}=\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}\quad \begin{array}{ll} ①\ 1,0,1. & ②\ 1,1,2. \\ ③\ -1,1,2. & ④\ -1,1,1. \end{array}$$

$$\boldsymbol{H}=\begin{pmatrix}1&2&3&4\\2&4&6&8\\3&6&9&12\\4&8&12&16\end{pmatrix}.$$

解　$\boldsymbol{A}$ 可以. 因三阶矩阵有 3 个不同的特征值.

$\boldsymbol{B}$ 不可以,理由见例 5-15.

$\boldsymbol{C}$ 可以,因为 $\boldsymbol{C}$ 是实对称矩阵.

$\boldsymbol{G}$ 选③.

$\boldsymbol{H}$ 的 4 个特征值是:30,0,0,0. 因为 $\boldsymbol{H}$ 是实对称,必可相似变换到对角阵,又因 $\boldsymbol{H}$ 的秩为 1,故对角阵的对角线元素只有一个是非零的,其余 3 个为零. 又 4 个特征值之和应等于 $\boldsymbol{H}$ 的迹,故这个非零的特征值为 30.

例 5-15　设二阶实矩阵 $\boldsymbol{A}$,$|\boldsymbol{A}|<0$,试证:$\boldsymbol{A}$ 必可对角化.

证　$|\boldsymbol{A}|=\lambda_1\cdot\lambda_2<0$.

由于复根出现必成对共轭,故 λ_1 与 λ_2 必为实根,且 $\lambda_1\neq\lambda_2$. 二阶矩阵为 2 个不同特征值,因此必可对角化.

例 5-16　设 $\boldsymbol{A}$ 是 n 阶矩阵. 证明:

(1) 当 n 是偶数,且 $|\boldsymbol{A}|<0$ 时,$\boldsymbol{A}$ 既有正特征值,又有负特征值;

(2) 当 n 是奇数,且 $|\boldsymbol{A}|>0$ 时,$\boldsymbol{A}$ 必有正特征值. 当 n 是奇数;且 $|\boldsymbol{A}|<0$ 时,A 必有负特征值.

证　(1) $|\boldsymbol{A}|=\lambda_1\cdot\lambda_2\cdots\lambda_n<0$,当 n 是偶数时,特征值不可能全是复根. 由 $|\boldsymbol{A}|<0$,可知:在 n 个特征值中,于少有 2 个(或 4 个,6 个,…)实根,它们相乘小于零,则必有正的特征值和负的特征值.

(2) 类似(1)的讨论,可证.

例 5-17　已知三阶矩阵 A 的特征值为 $1,-1,2$,设 $\boldsymbol{B}=\boldsymbol{A}^3-5\boldsymbol{A}^2$. 试求:

(1) $\boldsymbol{B}$ 的特征值及其相似对角阵;

(2) 计算 $|\boldsymbol{B}|$ 及 $|\boldsymbol{A}-5\boldsymbol{E}|$.

解　(1) 因为三阶矩阵 $\boldsymbol{A}$ 有三个单特征值,故 $\boldsymbol{A}$ 必可对角化,即有

$$\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\begin{pmatrix}1&&\\&-1&\\&&2\end{pmatrix},\text{也即 }\boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}1&&\\&-1&\\&&2\end{pmatrix}\boldsymbol{U}^{-1},$$

则

$$\boldsymbol{B}=\boldsymbol{A}^3-5\boldsymbol{A}^2=\boldsymbol{U}\begin{pmatrix}1&&\\&-1&\\&&8\end{pmatrix}\boldsymbol{U}^{-1}-\boldsymbol{U}\begin{pmatrix}5&&\\&5&\\&&20\end{pmatrix}\boldsymbol{U}^{-1}$$

$$=\boldsymbol{U}\begin{pmatrix}-4&&\\&-6&\\&&-12\end{pmatrix}\boldsymbol{U}^{-1},$$

即 $\boldsymbol{U}^{-1}\boldsymbol{B}\boldsymbol{U}=\begin{pmatrix}-4&&\\&-6&\\&&-12\end{pmatrix}$,知 $\boldsymbol{B}$ 的特征值为 $-4,-6,-12$.

(2)　$$|\boldsymbol{B}|=-4\cdot(-6)\cdot(-12)=-288.$$

又　$$|\boldsymbol{B}|=|\boldsymbol{A}^2||\boldsymbol{A}-5\boldsymbol{E}|=|\boldsymbol{A}|^2\cdot|\boldsymbol{A}-5\boldsymbol{E}|,$$

所以　$$|\boldsymbol{A}-5\boldsymbol{E}|=-\frac{288}{4}=-72.$$

评注　一旦判断出 $\boldsymbol{A}$ 可以对角化,马上就写出 $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{D}$($\boldsymbol{D}$ 为对角阵). 这样,可以把 $\boldsymbol{A}$ 表示为:$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{-1}$. 再据题目的条件进行运算或论证.

例 5-18 若 A 为幂零阵(即 $A\neq 0$,但 $A^k=0$,$k\in Z^+$)证明:

(1) A 的特征值全为零;

(2) $|E+A|=1$;

(3) A 不能对角化.

证 (1) 设 $AX=\lambda X$,则 $A^kX=\lambda^kX$.

由 $A^k=0$,$\Rightarrow\lambda^kX=0$,$\Rightarrow\lambda^k=0$,$\Rightarrow\lambda=0$.

(2) $\lambda E-A\Leftrightarrow(\lambda+1)E-(E+A)$.

上面的这个恒等表述表明:

λ 是 A 的特征值$\Leftrightarrow\lambda+1$ 是 $E+A$ 的特征值. 由(1)知,A 的特征值全是零. 则 $E+A$ 的特征值全是 1,则 $|E+A|=\lambda_1\lambda_2\cdot\cdots\cdot\lambda_n=1\times1\times\cdots\times1=1$.

(3)(反证)若 A 可以对角化. 则 $U^{-1}AU=D$. 由于 A 的特征值全是零,故这里 $D=0$. 则有 $A=0$. 这与已知 $A\neq0$ 矛盾.

评注 (1) 第(2)题中,建立了 A 与 $E+A$ 这两个矩阵的一个恒等表述关系,从而建立了 A 与 $E+A$ 之间特征值的相互关系. 这种方法对于 A 与 $kE+A$ 都是适用的. 在求解某些题目时,是有用的. 但要注意,并不是任何两个矩阵之间都可建立这种关系的.

不妨可以思考一下:若 A 的特征值全是 2,B 的特征值全是 -1,能否断言 $A+B$ 的特征值全是 1?

(2) 用第(2)题的方法也可以证明:若 $(E+A)=0$ $(m\in Z^+)$,则 A 可逆.

例 5-19 设 A 是 n 阶方阵,$\lambda=2,4,6,\cdots,2n$ 是 A 的 n 个特征值,求 $|A-3E|$.

解法 1 由条件知 A 必可对角化,则

$$U^{-1}AU=\begin{pmatrix}2&&&\\&4&&\\&&\ddots&\\&&&2n\end{pmatrix},\quad\text{即}\quad A=U\begin{pmatrix}2&&&\\&4&&\\&&\ddots&\\&&&2n\end{pmatrix}U^{-1},$$

则

$$|A-3E|=\left|U\begin{pmatrix}2&&&\\&4&&\\&&\ddots&\\&&&2n\end{pmatrix}U^{-1}-3UU^{-1}\right|=|U|\cdot\left|\begin{pmatrix}2&&&\\&4&&\\&&\ddots&\\&&&2n\end{pmatrix}-\begin{pmatrix}3&&&\\&3&&\\&&\ddots&\\&&&3\end{pmatrix}\right||U^{-1}|$$

$$=\begin{vmatrix}-1&&&&\\&1&&&\\&&3&&\\&&&\ddots&\\&&&&2n-3\end{vmatrix}=-1\times1\times3\times5\times\cdots\times(2n-3).$$

解法 2 由于 $\lambda E-A\Leftrightarrow(\lambda-3)E-(A-3E)$ 已知:λ 是 A 的特征值$\Leftrightarrow\lambda-3$ 是 $A-3E$ 的特征值. 所以 $A-3E$ 的 n 个特征值为 $-1,1,3,\cdots,(2n-3)$. 即有 $|A-3E|=-1\times1\times3\times5\times\cdots\times(2n-3)$.

例 5-20 设

$$A=\begin{pmatrix}3&4&0&0\\4&-3&0&0\\0&0&2&4\\0&0&0&2\end{pmatrix},\text{求 }|A^{2k}|\text{ 与 }A^{2k}.$$

解 把矩阵 A 表为分块矩阵形式.

$$A=\begin{pmatrix}B&0\\0&C\end{pmatrix},\text{其中 }B=\begin{pmatrix}3&4\\4&-3\end{pmatrix},C=\begin{pmatrix}2&4\\0&2\end{pmatrix}.$$

$$|\boldsymbol{A}^{2k}|=|\boldsymbol{A}|^{2k}=(|\boldsymbol{B}|\cdot|\boldsymbol{C}|)^{2k}=(-25\times 4)^{2k}=100^{2k}.$$

$$\boldsymbol{A}^{2k}=\begin{pmatrix}\boldsymbol{B}^{2k} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{C}^{2k}\end{pmatrix}\quad \text{其中}$$

$$\boldsymbol{B}^{2k}=(\boldsymbol{B}^2)^k=\begin{pmatrix}25 & 0\\ 0 & 25\end{pmatrix}^k=\begin{pmatrix}25^k & 0\\ 0 & 25^k\end{pmatrix}.$$

$$\boldsymbol{C}^{2k}=\left(2\cdot\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}\right)^{2k}=2^{2k}\cdot\begin{pmatrix}1 & 4k\\ 0 & 1\end{pmatrix}=\begin{pmatrix}4^k & k\cdot 4^{k+1}\\ 0 & 4^k\end{pmatrix}.$$

所以

$$\boldsymbol{A}^{2k}=\begin{pmatrix}25^k & 0 & 0 & 0\\ 0 & 25^k & 0 & 0\\ 0 & 0 & 4^k & k\cdot 4^{k+1}\\ 0 & 0 & 0 & 4^k\end{pmatrix}.$$

评注　在计算 $\boldsymbol{C}^{2k}$ 的时候，用到了结论：$\begin{pmatrix}1 & \lambda\\ 0 & 1\end{pmatrix}^n=\begin{pmatrix}1 & n\lambda\\ 0 & 1\end{pmatrix}$. 这当然也可临时递推.

在计算 $\boldsymbol{B}^{2k}$ 时，由于 $\boldsymbol{B}^{2k}=(\boldsymbol{B}^2)^k$，而 $\boldsymbol{B}^2$ 正好是对角阵，使计算简化了，否则，还要考虑 $\boldsymbol{B}$ 能否对角化（求特征值，特征向量）. 若可以，则可计算 $\boldsymbol{B}$ 的高次幂.

例 5-21　设

$$\boldsymbol{A}=\begin{pmatrix}a & -1 & c\\ 5 & b & 3\\ 1-c & 0 & -a\end{pmatrix},\ |\boldsymbol{A}|=-1.$$

又 $\boldsymbol{A}$ 的伴随阵 $\boldsymbol{A}^*$ 有一个特征值 λ_0，属于 λ_0 的一个特征向量 $\boldsymbol{\alpha}=(-1,-1,1)^{\mathrm{T}}$. 求 a,b,c 和 λ_0 的值.

解　$\boldsymbol{A}^*\boldsymbol{\alpha}=\lambda_0\boldsymbol{\alpha}\ \Rightarrow\ \boldsymbol{A}\boldsymbol{A}^*\boldsymbol{\alpha}=\lambda_0\boldsymbol{A}\boldsymbol{\alpha}$

即　　$|\boldsymbol{A}|\boldsymbol{\alpha}=\lambda_0\boldsymbol{A}\boldsymbol{\alpha}$，即 $-\boldsymbol{\alpha}=\lambda_0\boldsymbol{A}\boldsymbol{\alpha}$

即

$$\lambda_0\cdot\begin{pmatrix}a & -1 & c\\ 5 & b & 3\\ 1-c & 0 & -a\end{pmatrix}\begin{pmatrix}-1\\ -1\\ 1\end{pmatrix}=-\begin{pmatrix}-1\\ -1\\ 1\end{pmatrix}.$$

即
$$\begin{cases}\lambda_0(-a+1+c)=1 & (1)\\ \lambda_0(-5-b+3)=1 & (2)\\ \lambda_0(-1+c-a)=-1 & (3)\end{cases}$$

式(1)—式(3)：　$\Rightarrow\ \lambda_0=1$

代入式(2)　$\Rightarrow\ b=-3$

代入式(1)　$\Rightarrow\ a=c$

又由 $|A|=-1$，可解出 $a=2$，

所以，$a=c=2,b=-3,\lambda_0=1$.

评注　和 $\boldsymbol{A}^*$ 有关的最基本的关系式就是 $\boldsymbol{A}\boldsymbol{A}^*=\boldsymbol{A}^*\boldsymbol{A}=|\boldsymbol{A}|\boldsymbol{E}$.

例 5-22　已知三阶矩阵 $\boldsymbol{A}$ 与三维列向量 $\boldsymbol{X}$，使得 $\boldsymbol{X},\boldsymbol{A}\boldsymbol{X},\boldsymbol{A}\boldsymbol{X}^2$ 线性无关，且满足 $\boldsymbol{A}^3\boldsymbol{X}=3\boldsymbol{A}\boldsymbol{X}-2\boldsymbol{A}^2\boldsymbol{X}$.

(1) 记 $\boldsymbol{P}=(\boldsymbol{X},\boldsymbol{A}\boldsymbol{X},\boldsymbol{A}\boldsymbol{X}^2)$，求三阶矩阵 $\boldsymbol{B}$，使 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}$.

(2) 计算 $|\boldsymbol{A}+\boldsymbol{E}|$.

解　(1) 设

$$\boldsymbol{B}=\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

由 $\boldsymbol{A}=\boldsymbol{PBP}^{-1}\Rightarrow\boldsymbol{AP}=\boldsymbol{PB}$

即

$$\boldsymbol{A}(\boldsymbol{X},\boldsymbol{AX},\boldsymbol{A}^2\boldsymbol{X})=(\boldsymbol{X},\boldsymbol{AX},\boldsymbol{A}^2\boldsymbol{X})\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

得

$$\boldsymbol{AX}=a_1\boldsymbol{X}+b_1\boldsymbol{AX}+c_1\boldsymbol{A}^2\boldsymbol{X}$$
$$\boldsymbol{A}^2\boldsymbol{X}=a_2\boldsymbol{X}+b_2\boldsymbol{AX}+c_2\boldsymbol{A}^2\boldsymbol{X}$$
$$\boldsymbol{A}^3\boldsymbol{X}=3\boldsymbol{AX}-2\boldsymbol{A}^2\boldsymbol{X}=a_3\boldsymbol{X}+b_3\boldsymbol{AX}+c_3\boldsymbol{A}^2\boldsymbol{X}$$

利用 $\boldsymbol{X},\boldsymbol{AX},\boldsymbol{A}^2\boldsymbol{X}$ 线性无关，可由上面三式解出：

$$\boldsymbol{B}=\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 3 \\ 0 & 1 & -2 \end{pmatrix}.$$

(2) 由 $\boldsymbol{A}=\boldsymbol{PBP}^{-1}$，知 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，则 $\boldsymbol{A}+\boldsymbol{E}$ 与 $\boldsymbol{B}+\boldsymbol{E}$ 也相似，则

$$|\boldsymbol{A}+\boldsymbol{E}|=|\boldsymbol{B}+\boldsymbol{E}|=\begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 3 \\ 0 & 1 & -1 \end{vmatrix}=-4.$$

例 5-23 设三阶实对称阵 $\boldsymbol{A}$ 的特征值是 1,2,3. 矩阵 $\boldsymbol{A}$ 属于特征值 1,2 的特征向量分别是：

$$\boldsymbol{\alpha}_1=(-1,-1,1)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(1,-2,-1)^{\mathrm{T}}.$$

(1) 求 $\boldsymbol{A}$ 的属于特征值 3 的特征向量；

(2) 求矩阵 A.

解 (1) 设 $\boldsymbol{A}$ 的属于特征值 3 的特征向量是 $\boldsymbol{\alpha}_3=(x_1,x_2,x_3)^{\mathrm{T}}$. 则因为 $\boldsymbol{A}$ 实对称，故属于不同特征值的特征向量必正交，则有

$$\begin{cases} \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_3=-\boldsymbol{X}_1-\boldsymbol{X}_2+\boldsymbol{X}_3=0 \\ \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_3=\boldsymbol{X}_1-2\boldsymbol{X}_2-\boldsymbol{X}_3=0 \end{cases} \tag{1}$$

方程组(1)的基础解系可解得为 $\xi=(1,0,1)^{\mathrm{T}}$。

所以，属于特征值 3 的特征向量为 $\boldsymbol{\alpha}_3=k\cdot(1,0,1)^{\mathrm{T}}$，$k$ 为任意非零常数.

(2) 若记

$$\boldsymbol{P}=(\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\ \boldsymbol{\alpha}_3)=\begin{pmatrix} -1 & 1 & 1 \\ -1 & -2 & 0 \\ 1 & -1 & 1 \end{pmatrix},$$

则有

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix},\text{即 }\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}\boldsymbol{P}^{-1}.$$

由 $\boldsymbol{P}$ 可求得 $\boldsymbol{P}^{-1}$：

$$\boldsymbol{P}^{-1}=\frac{1}{6}\begin{pmatrix} -2 & -2 & 2 \\ 1 & -2 & -1 \\ 3 & 0 & 3 \end{pmatrix},\text{则 }\boldsymbol{A}=\frac{1}{6}\begin{pmatrix} 13 & -2 & 5 \\ -2 & 10 & 2 \\ 5 & 2 & 13 \end{pmatrix}.$$

例 5-24　设矩阵

$$\boldsymbol{A}=\begin{pmatrix}1 & -1 & 1\\ x & 4 & y\\ -3 & -3 & 5\end{pmatrix},$$

已知 $\boldsymbol{A}$ 有 3 个线性无关的特征向量，$\lambda=\alpha$ 是 $\boldsymbol{A}$ 的二重特征值. 试求可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为对角阵.

解　因三阶矩阵 $\boldsymbol{A}$ 有三个线性无关的特征向量，故必可对角化. 又已知 $\lambda=2$ 是二重特征值，故对应于 $\lambda=2$ 的线性无关的特征向量有 2 个，因此在 $(2\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 中，秩 $(2\boldsymbol{E}-\boldsymbol{A})=1$.

$$(2\boldsymbol{E}-\boldsymbol{A})=\begin{pmatrix}1 & 1 & -1\\ -x & -2 & -y\\ 3 & 3 & -3\end{pmatrix}\to\begin{pmatrix}1 & 1 & -1\\ 0 & x-2 & -x-y\\ 0 & 0 & 0\end{pmatrix}.$$

由上式右端矩阵秩为 1，可知 $x=2$，$y=-2$.

所以，

$$\boldsymbol{A}=\begin{pmatrix}1 & -1 & 1\\ 2 & 4 & -2\\ -3 & -3 & 5\end{pmatrix},$$

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda-2)^2(\lambda-6).$$

对于 $\lambda_1=\lambda_2=2$，解 $(2\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得到对应的特征向量为 $\boldsymbol{\alpha}_1=(1,-1,0)^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=(1,0,1)^{\mathrm{T}}$；对于 $\lambda_3=6$，解 $(6\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=0$，得到对应的特征向量为

$$\boldsymbol{\alpha}_3=(1,-2,3)^{\mathrm{T}}.$$

令

$$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=\begin{pmatrix}1 & 1 & 1\\ -1 & 0 & -2\\ 0 & 1 & 3\end{pmatrix},$$

则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 6\end{pmatrix}.$$

评注　在矩阵对角化问题中，若可以对角化，则特征根 λ_0 的重数（也称代数重数）必与求解 $(\lambda_0\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系中线性无关向量的个数（也称几何重数）相同. 而在线性方程组理论中知道：此基础解系中线性无关向量的个数（即解空间的维数）等于 $n-R(\lambda_0\boldsymbol{E}-\boldsymbol{A})$.

例 5-25　设矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & 1 & a\\ 1 & a & 1\\ a & 1 & 1\end{pmatrix}$，向量 $\boldsymbol{\beta}=\begin{pmatrix}1\\ 1\\ -2\end{pmatrix}$，

已知线性方程组 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{\beta}$ 有解但不惟一，试求：

(1) a 的值；

(2) 正交矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}$ 为对角阵.

解　(1) 把增广阵作初等行变换

$$(\boldsymbol{A},\boldsymbol{\beta})=\begin{pmatrix}1 & 1 & a & 1\\ 1 & a & 1 & 1\\ a & 1 & 1 & -2\end{pmatrix}\to\cdots\to\begin{pmatrix}1 & 1 & a & 1\\ 0 & a-1 & 1-a & 0\\ 0 & 0 & (a-1)(a+2) & a+2\end{pmatrix},$$

由于方程组有解但不惟一，故 $R(\boldsymbol{A})=R(\boldsymbol{A},\boldsymbol{\beta})<3$，所以 $a=-2$.

(2) 由(1)，把 $a=-2$ 代入 $\boldsymbol{A}$ 中，计算 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=\lambda(\lambda-3)(\lambda+3)$. 知特征值为 $\lambda_1=3$，$\lambda_2=-3$，$\lambda_3=0$，并分别求出它们所对应的特征向量：

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\0\\-1\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}1\\-2\\1\end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

因为 A 实对称，故它们已正交，只需单位化：

$$\boldsymbol{\beta}_1=\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{-1}{\sqrt{2}}\end{pmatrix},\quad \boldsymbol{\beta}_2=\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{-2}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\end{pmatrix},\quad \boldsymbol{\beta}_3=\begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\end{pmatrix}.$$

令 $$Q=(\beta_1,\beta_2,\beta_3),\quad \text{即有}\quad Q^{-1}AQ=Q^{\mathrm{T}}AQ=\begin{pmatrix}3&0&0\\0&-3&0\\0&0&0\end{pmatrix}.$$

例 5-26

设矩阵 $A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&a\end{pmatrix}$ 可逆，向量 $\boldsymbol{\alpha}=(1,b,1)^{\mathrm{T}}$ 是矩阵 A^* 的一个特征向量，λ 是 $\boldsymbol{\alpha}$ 对应的特征值，其中 A^* 是 A 的伴随阵. 试求：a，b 和 λ 的值.

解 由 A 可逆$\Rightarrow A^*$ 可逆$\Rightarrow\lambda\neq 0$.

又由 $A^*\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}\Rightarrow AA^*\boldsymbol{\alpha}=\lambda A\boldsymbol{\alpha}$

$$\Rightarrow A\boldsymbol{\alpha}=\frac{1}{\lambda}|A|\boldsymbol{\alpha},$$

即

$$\begin{pmatrix}2&1&1\\1&2&1\\1&1&a\end{pmatrix}\begin{pmatrix}1\\b\\1\end{pmatrix}=\frac{|\boldsymbol{A}|}{\lambda}\cdot\begin{pmatrix}1\\b\\1\end{pmatrix}$$

即

$$\begin{cases}3+b=\dfrac{|A|}{\lambda} & (1)\\ 2+2b=\dfrac{|A|}{\lambda}\cdot b & (2)\\ a+b+1=\dfrac{|A|}{\lambda} & (3)\end{cases}$$

由式(1)，式(3) $\Rightarrow a=2$；

由式(1)，式(2) $\Rightarrow b=1$ 或 -2.

又由

$$|\boldsymbol{A}|=\begin{vmatrix}2&1&1\\1&2&1\\1&1&a\end{vmatrix}=3a-2=4,$$

故由式(1)得

$$\lambda=\frac{|\boldsymbol{A}|}{3+b}=\frac{4}{3+b}.$$

所以，当 $b=1$ 时，$\lambda=1$；

当 $b=-2$ 时，$\lambda=4$.

例 5-27

已知 $\xi=\begin{pmatrix}1\\1\\-1\end{pmatrix}$ 是矩阵 $A=\begin{pmatrix}2&-1&2\\5&a&3\\-1&b&-2\end{pmatrix}$ 的一个特征向量.

(1) 试确定参数 a,b 及特征向量 ξ 所对应的特征值 λ.

(2) 问 A 能否对角化？说明理由.

解　(1) $A\xi=\lambda\xi$，即 $(A-\lambda E)\xi=0$.

即
$$\begin{pmatrix}2-\lambda&-1&2\\5&a-\lambda&3\\-1&b&-2-\lambda\end{pmatrix}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

即
$$\begin{cases}(2-\lambda)-1-2=0\\5+(a-\lambda)-3=0\\-1+b+(2+\lambda)=0,\end{cases}\quad\text{解出}\begin{cases}\lambda=-1\\a=-3\\b=0.\end{cases}$$

(2) $A=\begin{pmatrix}2&-1&2\\5&-3&3\\-1&0&-2\end{pmatrix}$

$|\lambda E-A|=-(\lambda+1)^3$，$\lambda=-1$ 是三重根. 求解 $(-E-A)X=0$ 的基础解系，由于

$$(-E-A)=\begin{pmatrix}-3&1&-2\\-5&2&-3\\1&0&1\end{pmatrix}\to\begin{pmatrix}1&0&1\\0&1&1\\0&2&2\end{pmatrix}\to\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix},$$

知 $R(-E-A)=2$，即知 $(-E-A)X=0$ 的基础解系中只有一个线性无关的特征向量，因此，A 不能对角化.

例 5-28　设矩阵
$$A=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&y&1\\0&0&1&2\end{pmatrix}.$$

(1) 已知 A 的一个特征值为 3，试求 y；

(2) 求矩阵 P，使 $(AP)^{\mathrm{T}}(AP)$ 为对角矩阵.

解　(1)
$$|\lambda E-A|=\begin{vmatrix}\lambda&-1&0&0\\-1&\lambda&0&0\\0&0&\lambda-y&-1\\0&0&-1&\lambda-2\end{vmatrix}$$
$$=(\lambda^2-1)\cdot[\lambda^2-(y+2)\lambda+(2y-1)]=0,$$

令 $\lambda=3$ 代入上式，得 $y=2$. 则
$$A=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&2&1\\0&0&1&2\end{pmatrix}.$$

(2) 由于 $A^{\mathrm{T}}=A$，故 $(AP)^{\mathrm{T}}(AP)=P^{\mathrm{T}}A^{\mathrm{T}}AP=P^{\mathrm{T}}A^2P$，而

$$\boldsymbol{A}^2=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&5&4\\0&0&4&5\end{pmatrix},$$

由式(1)了解得 $\boldsymbol{A}$ 的特征值为:$\lambda_1=\lambda_2=1,\lambda_3=-1,\lambda_4=3$. 则 $\boldsymbol{A}^2$ 的特征值为 $\lambda_1=\lambda_2=\lambda_3=1,\lambda_4=9$.

分别求出 $\boldsymbol{A}^2$ 的对应于特征值的特征向量为

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix}0\\0\\-1\\1\end{pmatrix},\quad \boldsymbol{\alpha}_4=\begin{pmatrix}0\\0\\1\\1\end{pmatrix}.$$

它们已经正交,只需再单位化:

$$\boldsymbol{P}_1=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad \boldsymbol{P}_2=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\quad \boldsymbol{P}_3=\begin{pmatrix}0\\0\\0\\\frac{-1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix},\quad \boldsymbol{P}_4=\begin{pmatrix}0\\0\\0\\\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix}.$$

则

$$\boldsymbol{P}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&\frac{-1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\0&0&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix},$$

$$(\boldsymbol{AP})^{\mathrm{T}}(\boldsymbol{AP})=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}^2\boldsymbol{P}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&9\end{pmatrix}.$$

评注 实对称阵必可正交相似变换到对角阵,而正交相似变换正好是合同变换(即 $\boldsymbol{P}^{-1}=\boldsymbol{P}^{\mathrm{T}}$).

例 5-29 设矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似,

$$\boldsymbol{A}=\begin{pmatrix}1&-1&1\\2&4&-2\\-3&-3&a\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}2&0&0\\0&2&0\\0&0&b\end{pmatrix}.$$

(1) 求 a,b 的值;

(2) 求可逆阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$.

解 (1) 因为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似,所以

$$\begin{cases}|\boldsymbol{A}|=|\boldsymbol{B}|\\t_r(\boldsymbol{A})=t_r(\boldsymbol{B})\end{cases}\text{即}\begin{cases}6a-6=4b\\5+a=4+b\end{cases}\text{解出}\begin{cases}a=5\\b=6.\end{cases}$$

(2) 对于 $\lambda_1=\lambda_2=2$,求解 $(2\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系,即特征向量为

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\-1\\0\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}1\\0\\1\end{pmatrix}.$$

对于 $\lambda_3=6$，求解$(6\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系，得

$$\boldsymbol{\alpha}_3=\begin{pmatrix}1\\-2\\3\end{pmatrix}.\text{则令 }\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=\begin{pmatrix}1&1&1\\-1&0&-2\\0&1&3\end{pmatrix}.$$

就有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}.$$

例 5-30　设

$$\boldsymbol{A}=\begin{pmatrix}a&1&1\\1&a&-1\\1&-1&a\end{pmatrix}.$$

求可逆阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为对角阵，并计算$|\boldsymbol{A}-\boldsymbol{E}|$.

解

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-a&-1&-1\\-1&\lambda-a&1\\-1&1&\lambda-a\end{vmatrix}=(\lambda-a-1)^2(\lambda-a+2).$$

特征值

$$\lambda_1=\lambda_2=a+1,\quad \lambda_3=a-2.$$

对应于 $\lambda_1=\lambda_2=a+1$，求解$((a+1)\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系，得两个线性无关的特征向量.

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\\0\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}1\\0\\1\end{pmatrix}.$$

对应于 $\lambda_3=a-2$，求解$((a-2)\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系，求得特征向量：

$$\boldsymbol{\alpha}_3=\begin{pmatrix}-1\\1\\1\end{pmatrix}.$$

则令

$$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=\begin{pmatrix}1&1&-1\\1&0&1\\0&1&1\end{pmatrix}$$

可有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}a+1&0&0\\0&a+1&0\\0&0&a-2\end{pmatrix}.$$

$$|\boldsymbol{A}-\boldsymbol{E}|=\left|\boldsymbol{P}\cdot\begin{pmatrix}a+1&0&0\\0&a+1&0\\0&0&a-2\end{pmatrix}\boldsymbol{P}^{-1}-\boldsymbol{P}\boldsymbol{P}^{-1}\right|=|\boldsymbol{P}|\cdot\left|\begin{pmatrix}a+1&0&0\\0&a+1&0\\0&0&a-2\end{pmatrix}-\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\right|\cdot|\boldsymbol{P}^{-1}|$$

$$=\left|\begin{pmatrix}a&0&0\\0&a&0\\0&0&a-3\end{pmatrix}\right|=a^2(a-3).$$

评注　注意矩阵中带有参数的特征值的求解过程，以及求相应的特征向量的过程.

例 5-31　设矩阵

$$\boldsymbol{A}=\begin{pmatrix}3&2&2\\2&3&2\\2&2&3\end{pmatrix},\quad \boldsymbol{P}=\begin{pmatrix}0&1&0\\1&0&1\\0&0&1\end{pmatrix},\quad \boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}^*\boldsymbol{P}.$$

求 $\boldsymbol{B}+2\boldsymbol{E}$ 的特征值与特征向量.

解　由 $\boldsymbol{A}$ 可求得 $\boldsymbol{A}^*$，由 $\boldsymbol{P}$ 可求得 $\boldsymbol{P}^{-1}$：

$$\boldsymbol{A}^*=\begin{pmatrix}5&-2&-2\\-2&5&-2\\-2&-2&5\end{pmatrix},\quad \boldsymbol{P}^{-1}=\begin{pmatrix}0&1&-1\\1&0&0\\0&0&1\end{pmatrix}.$$

则

$$\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}^*\boldsymbol{P}=\begin{pmatrix}7&0&0\\-2&5&-4\\-2&-2&3\end{pmatrix},\quad \boldsymbol{B}+2\boldsymbol{E}=\begin{pmatrix}9&0&0\\-2&7&-4\\-2&-2&5\end{pmatrix}.$$

$$|\lambda\boldsymbol{E}-(\boldsymbol{B}+2\boldsymbol{E})|=(\lambda-9)^2(\lambda-3).$$

对应于特征值 $\lambda_1=\lambda_2=9$，求解 $(9\boldsymbol{E}-(\boldsymbol{B}+2\boldsymbol{E}))\boldsymbol{X}=\boldsymbol{0}$ 的基础解系. 求得特征向量为

$$\boldsymbol{\eta}_1=\begin{pmatrix}-1\\1\\0\end{pmatrix},\quad \boldsymbol{\eta}_2=\begin{pmatrix}-2\\0\\1\end{pmatrix}.$$

故对应于特征值 9 的全部特征向量为 $k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2$. 其中，$k_1$，$k_2$ 是任意的不全为零的常数.

对应于特征值 $\lambda_3=3$，求解 $(3\boldsymbol{E}-(\boldsymbol{B}+2\boldsymbol{E}))\boldsymbol{X}=\boldsymbol{0}$ 的基础解系. 求得特征向量为

$\boldsymbol{\eta}_3=\begin{pmatrix}0\\1\\1\end{pmatrix}$，则对应于特征值 3 的全部特征向量为 $k_3\boldsymbol{\eta}_3$. k_3 为任意的非零常数.

例 5-32　设

$$\boldsymbol{A}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(a_1\ a_2\cdots\ a_n),\quad a_i\in\boldsymbol{R},\quad a_1\neq0.$$

求 $\boldsymbol{A}$ 的特征值与特征向量.

解　记 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}$. $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$. 易知 $\boldsymbol{A}$ 是实对称.

$\boldsymbol{A}^2=\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}=\boldsymbol{\alpha}(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha})\boldsymbol{\alpha}^{\mathrm{T}}=K\cdot\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}=K\cdot\boldsymbol{A}$. 这里 $K=\sum\limits_{i=1}^{n}a_i^2$.

设 $\boldsymbol{A}\boldsymbol{X}=\lambda\boldsymbol{X}$，$\Rightarrow\boldsymbol{A}^2\boldsymbol{X}=\lambda^2\boldsymbol{X}\Rightarrow K\boldsymbol{A}\boldsymbol{X}=\lambda^2\boldsymbol{X}\Rightarrow K\lambda\boldsymbol{X}=\lambda^2\boldsymbol{X}\Rightarrow(K\lambda-\lambda^2)\boldsymbol{X}=\boldsymbol{0}\Rightarrow(K\lambda-\lambda^2)=0.$　$\Rightarrow\lambda=0$ 或 $\lambda=K$.

由于 $R(\boldsymbol{A})=1$，$\boldsymbol{A}$ 实对称必可对角化，故知 $\lambda_1=K$，$\lambda_2=\lambda_3=\cdots=\lambda_n=0$.

再求特征向量. 对于 $\lambda_1=K=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}$，$\boldsymbol{A}\boldsymbol{X}=\lambda_1\boldsymbol{X}$. 即 $\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}\boldsymbol{X}$，可见取 $\boldsymbol{X}=\boldsymbol{\alpha}$ 等式即成立，即属于特征值 λ_1 的特征向量为

$$\boldsymbol{P}_1=\boldsymbol{\alpha}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}.$$

对应于特征值 $\lambda=0$，$\boldsymbol{A}\boldsymbol{X}=O\cdot\boldsymbol{X}=\boldsymbol{O}$.

$$\boldsymbol{A}\xrightarrow{\text{行}}\cdots\longrightarrow\begin{pmatrix}a_1&a_2&\cdots&a_n\\0&0&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&0\end{pmatrix}$$

即求解 $a_1x_1+a_2x_2+\cdots+a_nx_n=0$ 的基础解系. 可得

$$\boldsymbol{P}_2=\begin{pmatrix}-a_2\\a_1\\0\\\vdots\\0\end{pmatrix},\quad \boldsymbol{P}_3=\begin{pmatrix}-a_3\\0\\a_1\\\vdots\\0\end{pmatrix},\quad \cdots\quad \boldsymbol{P}_n=\begin{pmatrix}-a_n\\0\\0\\\vdots\\a_1\end{pmatrix}.$$

例 5-33　$\boldsymbol{A},\boldsymbol{B}$ 皆为 n 阶方阵. 证明：$\boldsymbol{A}\cdot\boldsymbol{B}$ 与 $\boldsymbol{BA}$ 有相同的特征值.

证　设 $\boldsymbol{ABX}=\lambda\boldsymbol{X}$. λ 是 $\boldsymbol{AB}$ 的特征值. 若 $\lambda=0$. $\Rightarrow|\boldsymbol{AB}|=0$. $\Rightarrow|\boldsymbol{BA}|=0$. $\Rightarrow\lambda=0$ 也是 $\boldsymbol{BA}$ 的特征值.

若 $\lambda\neq0$，则 $\boldsymbol{BX}\neq\boldsymbol{O}$，在 $\boldsymbol{ABX}=\lambda\boldsymbol{X}$ 等式两边左乘矩阵 $\boldsymbol{B}$，得 $\boldsymbol{BABX}=\lambda\boldsymbol{BX}$，$\Rightarrow\lambda$ 也是 $\boldsymbol{BA}$ 的特征值.

评注　有本题的结论，就可以证明：若 $\boldsymbol{E}+\boldsymbol{AB}$ 可逆，则 $\boldsymbol{E}+\boldsymbol{BA}$ 也可逆.

例 5-34　设 $\boldsymbol{A},\boldsymbol{B}$ 均为实对称阵. 试证：存在正交阵 $\boldsymbol{T}$，使 $\boldsymbol{T}^{-1}\boldsymbol{AT}=\boldsymbol{B}\Leftrightarrow\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的全部特征值.

证　必要性：若存在正交阵 $\boldsymbol{T}$，使 $\boldsymbol{T}^{-1}\boldsymbol{AT}=\boldsymbol{B}$，则

$$\begin{aligned}|\lambda\boldsymbol{E}-\boldsymbol{B}|&=|\lambda\boldsymbol{E}-\boldsymbol{T}^{-1}\boldsymbol{AT}|=|\lambda\boldsymbol{T}^{-1}\boldsymbol{T}-\boldsymbol{T}^{-1}\boldsymbol{AT}|\\&=|\boldsymbol{T}^{-1}|\cdot|\lambda\boldsymbol{E}-\boldsymbol{A}|\cdot|\boldsymbol{T}|=|\lambda\boldsymbol{E}-\boldsymbol{A}|.\end{aligned}$$

即 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征多项式，即 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的全部特征值.

充分性：若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的全部特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 又 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 皆为实对称，故存在正交阵 $\boldsymbol{T}_1$ 与 $\boldsymbol{T}_2$，使

$$\boldsymbol{T}_1^{-1}\boldsymbol{AT}_1=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix},\quad \boldsymbol{T}_2^{-1}\boldsymbol{BT}_2=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}.$$

即有　$\boldsymbol{T}_1^{-1}\boldsymbol{AT}_1=\boldsymbol{T}_2^{-1}\boldsymbol{BT}_2\quad\Rightarrow\quad(\boldsymbol{T}_2\boldsymbol{T}_1^{-1})\boldsymbol{A}(\boldsymbol{T}_1\boldsymbol{T}_2^{-1})=\boldsymbol{B}.$

即　$$(\boldsymbol{T}_1\boldsymbol{T}_2^{-1})^{-1}\boldsymbol{A}(\boldsymbol{T}_1\boldsymbol{T}_2^{-1})=\boldsymbol{B}.$$

由于正交阵的逆阵还是正交阵，正交阵与正交阵相乘还是正交阵，故 $\boldsymbol{T}=\boldsymbol{T}_1\cdot\boldsymbol{T}_2^{-1}$ 也为正交阵.

例 5-35　设 $\boldsymbol{A},\boldsymbol{B}$ 是 n 阶实对称阵，λ 是 $\boldsymbol{AB}$ 的一个非实特征值，$\boldsymbol{X}$ 是 $\boldsymbol{AB}$ 的对应于 λ 的特征向量. 证明：$\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{BX}=\boldsymbol{O}$.

证　设 $\boldsymbol{ABX}=\lambda\boldsymbol{X}$，$\lambda$ 是复数，$\boldsymbol{X}$ 是复向量. 上式两边共轭转置得 $\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{BA}=\overline{\lambda}\,\overline{\boldsymbol{X}}^{\mathrm{T}}$.

再在等式两边右乘 BX 得　$\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{BABX}=\overline{\lambda}\,\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{BX}$.

即　$$\lambda\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{BX}=\overline{\lambda}\,\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{BX}.$$

即　$$(\lambda-\overline{\lambda})\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{BX}=0.$$

由于 λ 非实，故 $\lambda-\overline{\lambda}\neq0$，$\Rightarrow\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{BX}=0$.

例 5-36　设有 4 阶矩阵 $\boldsymbol{A}$ 满足条件 $|3\boldsymbol{E}+\boldsymbol{A}|=0$. $\boldsymbol{AA}^{\mathrm{T}}=2\boldsymbol{E}$，$|\boldsymbol{A}|<0$. 其中 $\boldsymbol{E}$ 为 4 阶单位阵. 求 $\boldsymbol{A}$ 的伴随阵 $\boldsymbol{A}^*$ 的一个特征值.

解　$|3\boldsymbol{E}+\boldsymbol{A}|=(-1)^4\cdot|-3\boldsymbol{E}-A|=|-3\boldsymbol{E}-\boldsymbol{A}|=0$ 可知 $\lambda=-3$ 是 $\boldsymbol{A}$ 的一个特征值.

由 $|\boldsymbol{AA}^{\mathrm{T}}|=|2\boldsymbol{E}|=2^4=16\quad\Rightarrow|\boldsymbol{A}|=4$ 或 -4. 又因已知 $|\boldsymbol{A}|<0$，故 $|\boldsymbol{A}|=-4$.

由 $\boldsymbol{A}$ 有特征值 -3，$\boldsymbol{A}$ 可逆(因 $|\boldsymbol{A}|\neq0$). 故 $\boldsymbol{A}^{-1}$ 有特征值 $-\dfrac{1}{3}$.

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*=-\frac{1}{4}\boldsymbol{A}^*\quad\Rightarrow\quad\boldsymbol{A}^{-1}\boldsymbol{\alpha}=-\frac{1}{4}\boldsymbol{A}^*\boldsymbol{\alpha}=-\frac{1}{3}\boldsymbol{\alpha},$$

所以 $\boldsymbol{A}^*\boldsymbol{\alpha}=\dfrac{3}{4}\boldsymbol{\alpha}$，即 $\boldsymbol{A}^*$ 有特征值 $\dfrac{4}{3}$.

例 5-37　已知 $\boldsymbol{A}$ 是 n 阶实对称阵，且 $\boldsymbol{A}^2=\boldsymbol{A}$，试证：存在正交阵 $\boldsymbol{T}$，使得

$$\boldsymbol{T}^{-1}\boldsymbol{AT}=\begin{pmatrix}\boldsymbol{E}_r&0\\0&0\end{pmatrix},$$

其中，$\boldsymbol{E}_r$ 是 r 阶单位阵，$r=R(\boldsymbol{A})$.

证　因 $\boldsymbol{A}$ 实对称，故存在正交阵 $\boldsymbol{T}$，使

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},\lambda_i(i=1,\cdots,n)\text{是 }\boldsymbol{A}\text{ 的 }n\text{ 个特征值，则有}$$

$$\boldsymbol{A}=\boldsymbol{T}\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}\boldsymbol{T}^{-1},\quad \boldsymbol{A}^2=\boldsymbol{T}\begin{pmatrix}\lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_n^2\end{pmatrix}\boldsymbol{T}^{-1}.$$

由于 $\boldsymbol{A}^2=\boldsymbol{A}$，可得 $\lambda_i^2=\lambda_i$，$(i=1,\cdots,n)$.

即 $\lambda_i=0$ 或 1. 由于 $R(\boldsymbol{A})=$对角阵中非零元素的个数，故知 $\boldsymbol{A}$ 的特征值中有 r 个 1，$n-r$ 个零，故

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\boldsymbol{E}_r & 0\\ 0 & 0\end{pmatrix}.$$

证法 2　由 $\boldsymbol{A}^2=\boldsymbol{A}$. 设 $\boldsymbol{A}\boldsymbol{X}=\lambda\boldsymbol{X}\Rightarrow\boldsymbol{A}^2\boldsymbol{X}=\lambda^2\boldsymbol{X}\Rightarrow(\lambda^2-\lambda)\boldsymbol{X}=0,\Rightarrow\lambda=0$ 或 1.

因 $\boldsymbol{A}$ 实对称，故必可正交相似变换到对角阵，对角阵的对角线元素即为 $\boldsymbol{A}$ 的特征值. $\boldsymbol{A}$ 的秩，即为特征值 1 的个数. 已知 $R(\boldsymbol{A})=r$，故

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\boldsymbol{E}_r & 0\\ 0 & 0\end{pmatrix}.$$

例 5-38　设 $\boldsymbol{A}$ 是 n 阶矩阵$(n>1)$. $R(\boldsymbol{A})=1$. 试证：$\boldsymbol{A}$ 可以对角化$\Leftrightarrow t_r(\boldsymbol{A})\neq 0$.

证　因 $R(\boldsymbol{A})=1$，故$|\boldsymbol{A}|=0,\Rightarrow\boldsymbol{A}$ 有特征值 0. 且对应于特征值 0 的特征向量是求解$(0\cdot\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系. 由于 $R(\boldsymbol{A})=1$，可知基础解系中线性无关的特征向量的个数为 $n-1$ 个. 即知特征值 0 的重数$\geqslant n-1$.

当 $\boldsymbol{A}$ 有 n 重 0 特征值时，此时 $t_r(\boldsymbol{A})=0$，但线性无关特征向量只有 $n-1$ 个，此时 $\boldsymbol{A}$ 不能对角化.

当 $\boldsymbol{A}$ 有 $n-1$ 重 0 特征值时，则另有一个特征值不等于 0，此非零特征值即为 $t_r(\boldsymbol{A})$. 特征值 $t_r(\boldsymbol{A})$对应一个特征向量. 此时，线性无关的特征向量总共有$(n-1)+1=n$ 个. 因此 $\boldsymbol{A}$ 可以对角化.

例 5-39　某试验性生产线每年一月份进行熟炼工与非熟炼工的人数统计. 然后将$\frac{1}{6}$熟炼工支援其他生产部门，其缺额由招收新的非熟炼工补齐. 新、老非熟炼工经过培训及实践至年终考核有$\frac{2}{5}$成为熟炼工. 设第 n 年一月份统计的熟炼工和非熟炼工所占百分比分别为 x_n 和 y_n，记成向量$\begin{pmatrix}x_n\\ y_n\end{pmatrix}$.

(1) 求$\begin{pmatrix}x_{n+1}\\ y_{n+1}\end{pmatrix}$与$\begin{pmatrix}x_n\\ y_n\end{pmatrix}$的关系式，并写成矩阵形成：$\begin{pmatrix}x_{n+1}\\ y_{n+1}\end{pmatrix}=\boldsymbol{A}\cdot\begin{pmatrix}x_n\\ y_n\end{pmatrix}$；

(2) 验证 $\boldsymbol{\eta}_1=\begin{pmatrix}4\\ 1\end{pmatrix},\boldsymbol{\eta}_2=\begin{pmatrix}-1\\ 1\end{pmatrix}$是 $\boldsymbol{A}$ 的两个线性无关的特征向量，并求出相应的特征值；

(3) 当$\begin{pmatrix}x_1\\ y_1\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}$时，求$\begin{pmatrix}x_{n+1}\\ y_{n+1}\end{pmatrix}$.

解　(1) $\begin{cases}x_{n+1}=\frac{5}{6}x_n+\frac{2}{5}\left(\frac{1}{6}x_n+y_n\right)\\ y_{n+1}=\frac{3}{5}\left(\frac{1}{6}x_n+y_n\right)\end{cases}$

化简得

$$\begin{cases} x_{n+1}=\dfrac{9}{10}x_n+\dfrac{2}{5}y_n \\ y_{n+1}=\dfrac{1}{10}x_n+\dfrac{3}{5}y_n \end{cases}$$

即 $\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix}=\begin{pmatrix} \dfrac{9}{10} & \dfrac{2}{5} \\ \dfrac{1}{10} & \dfrac{3}{5} \end{pmatrix}\begin{pmatrix} x_n \\ y_n \end{pmatrix}$，于是 $\boldsymbol{A}=\begin{pmatrix} \dfrac{9}{10} & \dfrac{2}{5} \\ \dfrac{1}{10} & \dfrac{3}{5} \end{pmatrix}$.

(2) 令 $\boldsymbol{P}=(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2)=\begin{pmatrix} 4 & -1 \\ 1 & 1 \end{pmatrix}$，则 $|\boldsymbol{P}|=5\neq 0$，

知 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2$ 线性无关.

因 $\boldsymbol{A}\boldsymbol{\eta}_1=\begin{pmatrix} 4 \\ 1 \end{pmatrix}=\boldsymbol{\eta}_1$ 故 $\boldsymbol{\eta}_1$ 为 $\boldsymbol{A}$ 的特征向量，且对应的特征值 $\lambda_1=1$.

固 $\boldsymbol{A}\boldsymbol{\eta}_2=\begin{pmatrix} -\dfrac{1}{2} \\ \dfrac{1}{2} \end{pmatrix}=\dfrac{1}{2}\boldsymbol{\eta}_2$，知 $\boldsymbol{\eta}_2$ 为 $\boldsymbol{A}$ 的特征向量，且对应的特征值为 $\lambda_2=\dfrac{1}{2}$.

(3) $\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix}=\boldsymbol{A}\cdot\begin{pmatrix} x_n \\ y_n \end{pmatrix}=\boldsymbol{A}^2\cdot\begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix}=\cdots=\boldsymbol{A}^n\begin{pmatrix} \dfrac{1}{2} \\ \dfrac{1}{2} \end{pmatrix}$.

由 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$，有 $\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\boldsymbol{P}^{-1}$

于是 $\boldsymbol{A}^n=\boldsymbol{P}\cdot\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^n\boldsymbol{P}^{-1}$，又 $\boldsymbol{P}^{-1}=\dfrac{1}{5}\begin{pmatrix} 1 & 1 \\ -1 & 4 \end{pmatrix}$.

故
$$\begin{aligned}\boldsymbol{A}^n&=\frac{1}{5}\begin{pmatrix} 4 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \left(\frac{1}{2}\right)^n \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 4 \end{pmatrix} \\ &=\frac{1}{5}\begin{pmatrix} 4+\left(\frac{1}{2}\right)^n & 4-4\cdot\left(\frac{1}{2}\right)^n \\ 1-\left(\frac{1}{2}\right)^n & 1+4\cdot\left(\frac{1}{2}\right)^n \end{pmatrix}\end{aligned}$$

因此
$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix}=\boldsymbol{A}^n\cdot\begin{pmatrix} \dfrac{1}{2} \\ \dfrac{1}{2} \end{pmatrix}=\frac{1}{10}\cdot\begin{pmatrix} 8-3\cdot\left(\frac{1}{2}\right)^n \\ 2+3\cdot\left(\frac{1}{2}\right)^n \end{pmatrix}.$$

四、小　结

(1) 特征值与特征向量问题经常与线性方程组的理论，与向量组的线性相关性，与下一章的二次型等内容有联系. 这里，首先是求解特征值与特征向量是基本功，必须熟炼掌握.

(2) 特征值与矩阵 $\boldsymbol{A}$ 的关系中，尤以(1)$\lambda_1+\lambda_2+\cdots+\lambda_n=\mathrm{f}_{\mathrm{t}}(A)$；(2) $\lambda_1\cdot\lambda_2\cdots\lambda_n=|\boldsymbol{A}|$ 更为重要. 另外，一些矩阵与特征值之间的关系也应熟悉.

(3) 相似矩阵的性质应熟悉.

(4) 矩阵能否对角化，是本章核心内容. “代数重数≥几何重数”. 这个结论，我们没有证明，但读者应该知道，当取得等号时，就可以对角化(对每个特征值皆如此)，否则就不能对角化.

(5) 实对称阵的性质和定理也是重要的,这与下一章二次型有很大关系.

(6) 这一章中有些是充分必要的结论,有些仅是充分而不必要,有些是必要而不充分,这些内容在复习时,要注意多思考.

第六章　二次型

二次型问题实质上是实对称阵的对角化和实对称阵的特征值问题的延伸. 所涉及的主要问题是二次型化标准形及实对称阵的正定性及判别法.

一、复习与考试要求

(1) 熟悉二次型的定义及其矩阵表示,了解二次型的秩.

(2) 了解二次型和矩阵的正定性及其判别法.

(3) 掌握用正交变换化二次型为标准形.

二、基本概念与理论

1. 二次型

含 n 个变量 $x_1,x_2,\cdots,x_n$ 的二次齐次函数

$$f(x_1,x_2,\cdots,x_n)=\sum_{i,j=1}^{n}a_{ij}x_ix_j$$

称为二次型,当 a_{ij} 为实数时,f 称为实二次型.

2. 二次型的矩阵及秩

在二次型的表达式中取 $a_{ij}=a_{ji}$,则

$$f(x_1,x_2,\cdots,x_n)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X},$$

其中

$$\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{mn}\end{bmatrix},\quad \boldsymbol{X}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}.$$

对称矩阵 $\boldsymbol{A}$ 称为二次型 f 的矩阵,$\boldsymbol{A}$ 的秩也称为二次型 f 的秩.

3. 正(负)定矩阵

若对任何列向量 $\boldsymbol{X}\neq 0$,有

$f=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}>0(<0)$,则称 f 为正定(负定)二次型,相应地称对称阵 $\boldsymbol{A}$ 为正定(负定)矩阵.

$f=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\geqslant 0(\leqslant 0)$,则称 f 为半正定(半负定)二次型,$\boldsymbol{A}$ 为半正定(半负定)矩阵.

4. 二次型 f 化为标准形的方法

(1) 对二次型 $f=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 存在正交变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$,化二次型 f 为标准形

$$\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{Y}=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2,$$

其中正交阵 $\boldsymbol{P}$ 使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{bmatrix},\boldsymbol{Y}=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}.$$

5. 正定二次型的判别法

二次型 $f=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 是正定的充要条件是下列条件之一成立:

① 对称阵 A 的特征值全大于零；

② 对称阵 A 的各阶顺序主子式全大于零，即

$$a_{11}>0,\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}>0,\cdots,|\boldsymbol{A}|>0.$$

6. 负定二次型的判别法

二次型 $f=\boldsymbol{X}^{\mathrm{T}}AX$ 是负定的充要条件是下列条件之一成立：

①对称阵 A 的特征值全小于零；

②对称阵 A 的各阶顺序主要式中，奇数阶的全小于零，偶数阶全大于零.

③二次型——$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 是正定的二次型.

三、基本题型与解题方法

例 6-1　求一个正交变换，化二次型

$$f=x_1^2+4x_2^2+4x_3^2-4x_1x_2+4x_1x_3-8x_2x_3$$

为标准形.

分析　这是 1990 年数学(试卷一)的 8 分题. 是一种最基本的常规计算题型，它的解题步骤如下：

(1) 写出二次型的矩阵 A，从而得矩阵表达形式 $f=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$，注意平方项系数 a_{ii} 为 A 的对角元素，非平方项系数的一半作为 A 的元素 a_{ij}，另一半作为 A 的元素 a_{ji}，这样得对称阵 A.

(2) 求二次型的矩阵 A 的特征值和特征向量；

(3) 将特征向量正交规范化，写出正交矩阵 P 及正交变换；

(4) 写出二次型的标准形.

解　(1) 写出 f 的矩阵

$$\boldsymbol{A}=\begin{bmatrix}1 & -2 & 2\\ -2 & 4 & -4\\ 2 & -4 & 4\end{bmatrix},\text{从而 } f=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}.$$

(2) 由 $|\boldsymbol{A}-\lambda\boldsymbol{E}|=-\lambda((1-\lambda)(8-\lambda)-8)=0$，

得矩阵 A 的特征值 $\lambda_1=\lambda_2=0,\lambda_3=9$.

对 $\lambda_1=\lambda_2=0$，解齐次线性方程组 $(\boldsymbol{A}-\lambda_1\boldsymbol{E})\boldsymbol{X}=0$ 得两个线性无关的向量 $\xi_1=(2,1,0)^{\mathrm{T}},\xi_2=(-2,0,1)^{\mathrm{T}}$. 为得到两个正交的特征向量，将 $\boldsymbol{A}-\lambda_1\boldsymbol{E}$ 化简得到的非零行向量 $(1,-2,2)$ 和 $\xi_1=(2,1,0)^{\mathrm{T}}$ 合在一起作系数阵，解线性方程组

$$\begin{pmatrix}1 & -2 & 2\\ 2 & 1 & 0\end{pmatrix}\boldsymbol{X}=0$$

得

$$\xi_2=(-2,4,5)^{\mathrm{T}}$$

对于 $\lambda_3=9$，解线性方程组　$(\boldsymbol{A}-\lambda_3\boldsymbol{E})\boldsymbol{X}=\boldsymbol{0}$ 得到

$$\xi_3=(1,-2,2)^{\mathrm{T}}.$$

(3) 将 ξ_1,ξ_2,ξ_3 单位化得

$$\boldsymbol{\beta}_1=\boldsymbol{\xi}_1/\|\boldsymbol{\xi}_1\|=\frac{1}{\sqrt{5}}(2,1,0)^{\mathrm{T}},$$

$$\boldsymbol{\beta}_2=\boldsymbol{\xi}_2/\|\boldsymbol{\xi}_2\|=\frac{1}{3\sqrt{5}}(-2,4,5)^{\mathrm{T}},$$

$$\boldsymbol{\beta}_3=\boldsymbol{\xi}_3/\|\boldsymbol{\xi}_3\|=\frac{1}{3}(1,-2,2)^{\mathrm{T}},$$

以 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 为列向量，构成正交阵

$$\boldsymbol{P}=\begin{bmatrix}\frac{2}{\sqrt{5}} & \frac{-2}{3\sqrt{5}} & \frac{1}{3}\\ \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} & \frac{-2}{3}\\ 0 & \frac{5}{3\sqrt{5}} & \frac{2}{3}\end{bmatrix},$$

再写出正交变换 $\boldsymbol{X}=\boldsymbol{PY}$.

(4) 写出二次型的标准形

$$f=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}\xlongequal{\boldsymbol{X}=\boldsymbol{PY}}9y_3^2$$

例 6-2　已知二次型

$$f(x_1,x_2,x_3)=x_1^2+x_2^2+5x_3^2+2\lambda x_1x_2-2x_1x_3-4x_2x_3.$$

问 λ 取何值时，该二次型是正定的？

分析　这类问题可用正定阵的判别法来确定参数 λ 的范围，应用正定阵的充要条件是它的各阶主子式大于零即可求出 λ 的取值范围.

解　二次型矩阵为

$$\boldsymbol{A}=\begin{bmatrix}1 & \lambda & -1\\ \lambda & 1 & -2\\ -1 & -2 & 5\end{bmatrix},$$

$$\boldsymbol{A}>0 \quad\Leftrightarrow\quad |\boldsymbol{A}_1|=1>0,$$

$$|\boldsymbol{A}_2|=\begin{vmatrix}1 & \lambda\\ \lambda & 1,\end{vmatrix}=1-\lambda^2>0, \qquad \text{得} \quad -1<\lambda<1,$$

$$|\boldsymbol{A}_3|=4\lambda-5\lambda^2>0, \qquad \text{得} \quad \begin{cases}\lambda>0\\ \lambda<\frac{4}{5}.\end{cases}$$

所以，要使所给二次型正定，其充要条件是 $0<\lambda<\frac{4}{5}$.

例 6-3　设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶正定对称阵，求证 $\boldsymbol{A}+\boldsymbol{B},\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$ 也都是对称正定阵.

分析　要证 $\boldsymbol{A}+\boldsymbol{B}$ 正定，必须用定义，要证 $\boldsymbol{A}^{-1},\boldsymbol{B}^{-1}$ 正定则要用到 $\boldsymbol{A}$ 正定的充要条件是其特征值全大于零.

证　因为 $\boldsymbol{A}$ 正定，$\boldsymbol{B}$ 正定，所以对任意 n 维非零列向量 $\boldsymbol{X}$，有

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}>0,\quad \boldsymbol{X}^{\mathrm{T}}\boldsymbol{BX}>0,$$

故得

$$\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{X}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}+\boldsymbol{X}^{\mathrm{T}}\boldsymbol{BX}>0,$$

此即表明 $\boldsymbol{A}+\boldsymbol{B}$ 正定.

设 λ 是 $\boldsymbol{A}$ 的特征值，则有非零列向量 $\boldsymbol{X}$ 使

$$\boldsymbol{AX}=\lambda\boldsymbol{X},$$

因为 $\boldsymbol{A}$ 正定，所以 $\boldsymbol{A}$ 可逆，$\lambda>0$. 从而

$$\boldsymbol{A}^{-1}\boldsymbol{X}=\frac{1}{\lambda}\boldsymbol{X},$$

因此，$\frac{1}{\lambda}$ 是 $\boldsymbol{A}^{-1}$ 的特征值，且 $\frac{1}{\lambda}>0$，故 $\boldsymbol{A}^{-1}$ 正定，同理 $\boldsymbol{B}^{-1}$ 也正定. 由前面得到的结果，则得 $\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$ 正定.

例 6-4　求证 n 阶对称阵 $\boldsymbol{A}$ 正定的充要条件是

(1) 存在可逆阵，使 $\boldsymbol{A}=\boldsymbol{QQ}^{\mathrm{T}}$；

(2) 存在正定阵 $\boldsymbol{S}$，使 $\boldsymbol{A}=\boldsymbol{S}^2$.

分析　对称正定阵的一个优良特性是可以对角化，而且正交相似的对角阵的对角线元素全大于零，从而可将此对角阵分解成一个正定对角阵的平方，再分别求出 $\boldsymbol{Q}$ 和 $\boldsymbol{S}$.

证 设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 A 的特征值，因为 A 正定，故 $\lambda_i>0,i=1,2,\cdots,n$，且存在正交阵 P 使

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}=\begin{bmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{bmatrix}\begin{bmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{bmatrix},$$

所以

$$A=P\begin{bmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{bmatrix}\begin{bmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{bmatrix}'P^{\mathrm{T}},$$

令

$$Q=P\begin{bmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{bmatrix}$$

则得 $$A=QQ^{\mathrm{T}}.$$

另外，A 可以表达成

$$A=P\begin{bmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{bmatrix}P^{\mathrm{T}}\cdot P\begin{bmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{bmatrix}P^{\mathrm{T}},$$

令

$$S=P\begin{bmatrix}\sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n}\end{bmatrix}P^{\mathrm{T}},$$

则得 $A=S^2$，且 S 正定.

例 6-5 已知方程 $5x^2+5y^2+cz^2-2xy+6xz-6yz=1$ 的图形是一个二次柱面，求常数 c，确定柱面的类型，并求柱面母线的方向向量.

分析 本例把二次型与解析几何相结合，超出考试大纲的要求，但二次型的几何意义，考生也应该有所了解，且设计线性代数与高等数学相结合的综合题，很可能是今后考试命题的一种趋势.

方程左端的二次型记为

$$f(x,y,z)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X},\boldsymbol{X}=(x,y,z)^{\mathrm{T}}$$

经正交变换 $X=PY$，二次型化为

$$f=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{Y}=\lambda_1u^2+\lambda_2v^2+\lambda_3\omega^2,(\boldsymbol{Y}=(u,v,w)^{\mathrm{T}})$$

由 $f=1$ 的图形为二次柱面，知 λ_i 中必有一个为 0，无妨设 $\lambda_3=0$，则柱面母线方向即为 w 轴的方向，亦即对应 $\lambda_3=0$ 的特征向量方向.

解 方程左端二次型对应的矩阵

$$\boldsymbol{A}=\begin{bmatrix}5&-1&3\\-1&5&-3\\3&-3&c\end{bmatrix}$$

由 $\boldsymbol{A}$ 的一个特征值为 0，知 $|\boldsymbol{A}|=0$，求得 $c=3$.

由　$\boldsymbol{AX}=\boldsymbol{0}$ 解得特征向量 $\boldsymbol{p}=(-1,1,2)^{\mathrm{T}}$，即为柱面母线的方向向量.

$$\boldsymbol{A}=\begin{bmatrix}5&-1&3\\-1&5&-3\\3&-3&c\end{bmatrix}\longrightarrow\begin{bmatrix}-1&5&-3\\0&2&-1\\0&0&c-3\end{bmatrix},$$

因 $R(\boldsymbol{A})=2$，解得 $c=3$.

这时，$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{bmatrix}\lambda-5&1&-3\\1&\lambda-5&3\\-3&3&\lambda-3\end{bmatrix}=\lambda(\lambda-4)(\lambda-9)=0$，

故所求特征值为 $\lambda_1=0,\lambda_2=4,\lambda_3=9$.

由上述特征值可知　$f(\boldsymbol{X})=1$ 所表示的二次曲面是椭圆柱面

$$4u^2+9v^2=1.$$

例 6-6　已知二次型 $f(x_1,x_2,x_3)=x_1^2-2x_2^2+x_3^2+2x_1x_3$.

(1) 试用配方法化二次型为标准形，并求变换矩阵；

(2) 试用正交变换化二次型为标准形，并求变换矩阵；

(3) 写出二次型的规范形，并求变换矩阵.

解　(1) $f(x_1,x_2,x_3)=x_1^2-2x_2^2+x_3^2+2x_1x_3=(x_1^2+2x_1x_3+x_3^2)-2x_2^2=(x_1+x_3)^2-2x_2^2$，

令

$$\begin{cases}y_1=x_1+x_3\\y_2=x_2\\y_3=x_3\end{cases}\quad\text{及解出}\quad\begin{cases}x_1=y_1-y_3\\x_2=y_2\\x_3=y_3,\end{cases}$$

即 $f(x)$ 在 $\boldsymbol{X}=\boldsymbol{Py}$ 后化为 $f(y)=y_1^2-2y_2^2$，其中

$$\boldsymbol{P}=\begin{pmatrix}1&0&-1\\0&1&0\\0&0&1\end{pmatrix}.$$

(2) 二次型 $f(x)$ 的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}1&0&1\\0&-2&0\\1&0&1\end{pmatrix}.$$

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\lambda(\lambda-2)(\lambda+2).$$

求出对应于 $\lambda_1=2$ 的特征向量：$\boldsymbol{\alpha}_1=(1,0,1)^{\mathrm{T}}$；对应于 $\lambda_2=-2$ 的特征向量：$\boldsymbol{\alpha}_2=(0,1,0)^{\mathrm{T}}$，对应于 $\lambda_3=0$ 的特征向量：$\boldsymbol{\alpha}_3=(1,0,-1)^{\mathrm{T}}$.

上述向量已经正交，只需单位化：

$$\boldsymbol{\xi}_1=\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix},\quad\boldsymbol{\xi}_2=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad\boldsymbol{\xi}_3=\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{-1}{\sqrt{2}}\end{pmatrix}.$$

则

$$\boldsymbol{T}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3)=\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ 0 & 1 & 0\\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}}\end{pmatrix}.$$

令 $\boldsymbol{X}=\boldsymbol{TY}$，则 $f(x)$ 在变换 $\boldsymbol{X}=\boldsymbol{TY}$ 下，化为标准形 $f(Y)=2y_1^2-2y_2^2$.

(3) 若由(2)往下做，再令 $\boldsymbol{Y}=\boldsymbol{RZ}$，这里，

$$\boldsymbol{R}=\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & 0\\ 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix},$$

则 $$f(Z)=\boldsymbol{Z}^{\mathrm{T}}(\boldsymbol{R}^{\mathrm{T}}\cdot(\boldsymbol{T}^{\mathrm{T}}\boldsymbol{AT})\boldsymbol{R})\boldsymbol{Z}=\boldsymbol{Z}^{\mathrm{T}}\cdot\begin{pmatrix}1 & & \\ & -1 & \\ & & 0\end{pmatrix}\cdot\boldsymbol{Z}=z_1^2-z_2^2.$$

即为规范形，此时所用的变换矩阵为

$$\boldsymbol{M}=\boldsymbol{T}\cdot\boldsymbol{R}=\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ 0 & 1 & 0\\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}}\end{pmatrix}\cdot\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & 0\\ 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}\frac{1}{2} & 0 & \frac{1}{\sqrt{2}}\\ 0 & \frac{1}{\sqrt{2}} & 0\\ 1 & 0 & \frac{-1}{\sqrt{2}}\end{pmatrix}.$$

若由(1)往下做，则可再令 $\boldsymbol{Y}=\boldsymbol{HZ}$，这里，

$$\boldsymbol{H}=\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix},$$

则 $$f(Z)=\boldsymbol{Z}^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP})\boldsymbol{H})\boldsymbol{Z}=\boldsymbol{Z}^{\mathrm{T}}\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}\cdot\boldsymbol{Z}=z_1^2-z_2^2.$$ 即得规范形.

这时，所用的变换矩阵为

$$\boldsymbol{N}=\boldsymbol{P}\cdot\boldsymbol{H}=\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}1 & 0 & -1\\ 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}.$$

例 6-7 已知二次曲面方程：$x^2+ay^2+z^2+2bxy+2xz+2yz=4$，可经过正交变换.

$$\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\boldsymbol{P}\cdot\begin{pmatrix}\boldsymbol{\xi}\\ \boldsymbol{\eta}\\ \boldsymbol{\zeta}\end{pmatrix}$$

化为椭圆柱面方程：$\boldsymbol{\eta}^2+4\boldsymbol{\zeta}^2=4$. 求 a,b 的值和正交阵 $\boldsymbol{P}$.

解

$$\boldsymbol{A}=\begin{pmatrix}1 & b & 1\\ b & a & 1\\ 1 & 1 & 1\end{pmatrix},$$

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 4\end{pmatrix}.$$

则由相似矩阵性质：$2+a=5$，$\Rightarrow a=3$.

又由 $|A|=0$，可解得 $b=1$.

再求属于特征值 $\lambda_1=0,\lambda_2=1,\lambda_3=4$ 的特征向量，并单位化，得

$$\boldsymbol{P}_1=\begin{pmatrix}\frac{1}{\sqrt{2}}\\ 0\\ \frac{-1}{\sqrt{2}}\end{pmatrix},\quad \boldsymbol{P}_2=\begin{pmatrix}\frac{1}{\sqrt{3}}\\ \frac{-1}{\sqrt{3}}\\ \frac{1}{\sqrt{3}}\end{pmatrix},\quad \boldsymbol{P}_3=\begin{pmatrix}\frac{1}{\sqrt{6}}\\ \frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{6}}\end{pmatrix}.$$

则 $\boldsymbol{P}=(\boldsymbol{P}_1,\boldsymbol{P}_2,\boldsymbol{P}_3)$ 即为所求的正交阵.

例 6-8　试证：

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\quad 与\quad \begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}$$

在实数域不是合同的，在复数域是合同的.

证　若合同，即存在可逆阵 $\boldsymbol{P}$，使

$$\boldsymbol{P}^{\mathrm{T}}\cdot\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\cdot\boldsymbol{P}=\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}.$$

两边取行列式，得 $|\boldsymbol{P}|^2=-1$. 这在实数域是不可能的. 而在复数域，可有

$$\begin{pmatrix}i & 0\\ 0 & 1\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}i & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix},$$

即 $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ 与 $\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}$ 是合同的.

例 6-9　证明：

$$\boldsymbol{A}=\begin{pmatrix}a_1 & 0 & 0\\ 0 & a_2 & 0\\ 0 & 0 & a_3\end{pmatrix}\quad 与\quad \boldsymbol{B}=\begin{pmatrix}a_2 & 0 & 0\\ 0 & a_3 & 0\\ 0 & 0 & a_1\end{pmatrix}是合同的.$$

证

$$\boldsymbol{P}(1,2)\cdot\boldsymbol{P}(1,3)\cdot\boldsymbol{A}\cdot\boldsymbol{P}(1,3)\cdot\boldsymbol{P}(1,2)=\boldsymbol{B}$$

而

$$\boldsymbol{P}(1,2)\cdot\boldsymbol{P}(1,3)=(\boldsymbol{P}(1,3)\cdot\boldsymbol{P}(1,2))^{\mathrm{T}},$$

记 $\boldsymbol{P}=\boldsymbol{P}(1,3)\cdot\boldsymbol{P}(1,2)$，即有 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$.

而 $\boldsymbol{P}$ 是两个初等阵的乘积，故还是可逆阵. 因此 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同.

例 6-10　设二次型 $f=x_1^2+x_2^2+x_3^2+2\alpha x_1x_2+2\beta x_2x_3+2x_1x_3$ 经正交变换 $x=\boldsymbol{P}\boldsymbol{Y}$ 化成 $f=y_2^2+2y_3^2$. 其中，$\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}$，$\boldsymbol{y}=(y_1,y_2,y_3)^{\mathrm{T}}$ 都是三维列向量，$\boldsymbol{P}$ 是三阶正交矩阵. 试求常数 α,β.

解　变换前后二次型的矩阵分别为

$$\boldsymbol{A}=\begin{pmatrix}1 & \alpha & 1\\ \alpha & 1 & \beta\\ 1 & \beta & 1\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 2\end{pmatrix}.$$

由于 P 是正交阵，故 $P^{\mathrm{T}}AP=B$ 即 $P^{-1}AP=B$，即 A 与 B 相似，故 $|\lambda E-A|=|\lambda E-B|$，展开即为

$$\lambda^3-3\lambda^2+(2-\boldsymbol{\alpha}^2-\boldsymbol{\beta}^2)\lambda+(\boldsymbol{\alpha}-\boldsymbol{\beta})^2\equiv\lambda^3-3\lambda^2+2\lambda.$$

比较两端同次幂的系数，得 $\boldsymbol{\alpha}=\boldsymbol{\beta}=\boldsymbol{O}$.

评注 注意二次型的矩阵变换，多为合同变换，而正交矩阵 $P^{\mathrm{T}}=P^{-1}$，故也正是相似变换，可用相似矩阵的一些性质.

例 6-11 若 $f(x)=x^{\mathrm{T}}Ax$ 经过可逆变换 $x=Py$ 后，化为 $f(Py)=y^{\mathrm{T}}(P^{\mathrm{T}}AP)y$，则在二次型 $f(x)$ 与 $f(Py)$ 中，只要有一个是正定二次型，则另一个也是正定二次型.

证 设 $f(x)$ 正定，往证 $f(Py)$ 也正定. 即证 $P^{\mathrm{T}}AP$ 正定.

$$\forall\quad Z\neq O,\quad Z^{\mathrm{T}}(P^{\mathrm{T}}AP)Z=(PZ)^{\mathrm{T}}A(PZ).$$

由于 P 可逆，故 PZ 还是任给的非零向量. 由 A 正定，故 $(PZ)^{\mathrm{T}}A(PZ)>0$，即得 $P^{\mathrm{T}}AP$ 正定. 即 $f(Py)$ 正定.

由于上述变换是可逆变换，故也有 $Y=P^{-1}X$. 故上述证明过程可逆. 证毕.

评注 从矩阵角度说，本题即证明了：合同变换保持矩阵的正定性不变.

例 6-12 正定阵 A 的等价说法：

(1) 对于实对称阵 A，$\forall X\neq0$，恒有 $X^{\mathrm{T}}AX>0$.

(2) 实对称阵 A 的所有特征值 $\lambda_1,\cdots,\lambda_n$ 全大于零.

(3) A 与 E 合同；

(4) 存在可逆阵 U，使 $A=U^{\mathrm{T}}U$.

证 “(1)⇒(2)”：已知 $f(x)=X^{\mathrm{T}}AX>0$，由 A 实对称，故存在正交阵 Q，使经过 $X=QY$ 后，

$$f(Qy)=Y^{\mathrm{T}}(Q^{\mathrm{T}}AQ)Y=Y^{\mathrm{T}}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}Y=\lambda_1y_1^2+\cdots+\lambda_ny_n^2,$$

由例 6-11 知，$f(Qy)$ 正定. 而 $f(Qy)$ 正定的充分必要条件是 $\lambda_1,\cdots,\lambda_n$ 全大于零. 这里的 $\lambda_1,\cdots,\lambda_n$ 是 A 的全部特征值.

$$(2)\Rightarrow(3):\quad f(x)=X^{\mathrm{T}}AX\xrightarrow[Q\text{ 正交}]{X=QY}f(Qy)=Y^{\mathrm{T}}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}Y,$$

再令 $Y=PZ$，其中 $P=\begin{pmatrix}\frac{1}{\sqrt{\lambda_1}}&&\\&\ddots&\\&&\frac{1}{\sqrt{\lambda_n}}\end{pmatrix}$. 则得

$f(QPZ)=Z^{\mathrm{T}}EZ$，其中的矩阵变换即 $P^{\mathrm{T}}(Q^{\mathrm{T}}AQ)P=E$.

即 A 与 E 合同.

(3)⇒(4)：已知存在可逆阵 P，$P^{\mathrm{T}}AP=E$，则有 $A=(P^{\mathrm{T}})^{-1}P^{-1}=(P^{-1})^{\mathrm{T}}P^{-1}=U^{\mathrm{T}}U$. 这里 $U=P^{-1}$ 是可逆阵.

(4)⇒(1)： 已知 $A=U^{\mathrm{T}}U$. U 可逆.

$\forall X\neq0$，$X^{\mathrm{T}}AX=X^{\mathrm{T}}U^{\mathrm{T}}UX=(UX)^{\mathrm{T}}(UX)>0$. 即 A 正定.

例 6-13 若 A,B 都是正定阵，则 $A+B$ 也是正定阵.

证 $(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}}=A+B$ 故对称.

$\forall X\neq0$，$X^{\mathrm{T}}(A+B)X=X^{\mathrm{T}}AX+X^{\mathrm{T}}BX>0$.

所以，$A+B$ 是正定阵.

评注 若已知正定阵，则它必是实对称. 若要证矩阵是正定阵. 则应先证对称.

例 6-14 若 A 是正定阵，则 A^{-1}，A^*，A^n 皆为正定阵.

证 $(A^{-1})^{\mathrm{T}}=(A^{\mathrm{T}})^{-1}=A^{-1}$，故 A^{-1} 对称.

设 A 的特征值 λ，则由 A 正定，知 $\lambda>0$，而 A^{-1} 的特征值为 $\frac{1}{\lambda}$，则可知 A^{-1} 的特征值也全大于零，故 A^{-1} 正定.

A^*，A^n 的正定性类似可证.

例 6-15　若 A 是正定阵，则 $a_{ii}>0$，$(i=1,\cdots,n)$.

证　因 A 正定，故 $A=U^{\mathrm{T}}U$.（U 为可逆阵）. 而 $U^{\mathrm{T}}U$ 的对角线元素正是 U 的每一列元素的平方和. 由于 U 可逆，故每一列皆非零向量，故平方和必大于零.

例 6-16　若 A，B 都是正定阵，且 $AB=BA$，则 $A\cdot B$ 也是正定阵.

证　$(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}=BA=AB$，故 $A\cdot B$ 对称.

由 A 正定，$\Rightarrow A=P^{\mathrm{T}}P$. P 为可逆阵，同理 $B=Q^{\mathrm{T}}Q$，Q 可逆. 故 $AB=P^{\mathrm{T}}PQ^{\mathrm{T}}Q$.

再作相似变换：$Q(AB)Q^{-1}=QP^{\mathrm{T}}PQ^{\mathrm{T}}QQ^{-1}$

$$=QP^{\mathrm{T}}PQ^{\mathrm{T}}=(PQ^{\mathrm{T}})^{\mathrm{T}}(PQ^{\mathrm{T}}).$$

即等号右端是正定阵. 即 $A\cdot B$ 与此正定阵相似. 从而有相同的特征值，从而知 $A\cdot B$ 的特征值全大于零. 因此 $A\cdot B$ 正定.

例 6-17　证明：二次型 $f=X^{\mathrm{T}}AX$ 在 $\|X\|=1$ 时的最大值为矩阵 A 的最大特征值.

证　因 A 实对称，故必存在正交阵 Q，使 $Q^{\mathrm{T}}AQ=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$，这里 λ_i 是 A 的特征值，且全为实数.

$$f=X^{\mathrm{T}}AX=X^{\mathrm{T}}QQ^{\mathrm{T}}AQQ^{\mathrm{T}}X=(Q^{T}X)^{\mathrm{T}}(Q^{\mathrm{T}}AQ)(Q^{T}X)$$

记向量 $y=Q^{\mathrm{T}}X$，则由于正交变换是保模变换，故知 $\|y\|=\|x\|=1$，则

$$f=y^{\mathrm{T}}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}y=\lambda_1 y_1^2+\cdots+\lambda_n y_n^2$$

$$\leqslant\lambda_{\max}(y_1^2+\cdots+y_n^2)=\lambda_{\max}\cdot\|y\|=\lambda_{\max}.$$

设 $\lambda_{\max}=\lambda_i$，只要取 $Y_0=(0\cdots1\cdots0)^{\mathrm{T}}$，此时

$$Y_0=Q^{\mathrm{T}}X_0,\ \text{即}\ X_0=QY_0,\ \text{且有}\ \|X_0\|=1.\ \text{则}$$

$$f(X_0)=X_0^{\mathrm{T}}AX_0=(QY_0)^{\mathrm{T}}A(QY_0)=Y_0^{\mathrm{T}}Q^{\mathrm{T}}AQY_0$$

$$=(0,\cdots,\underset{j}{1},\cdots,0)\begin{pmatrix}\lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_j & & \\ & & & \ddots & \\ & & & & \lambda_n\end{pmatrix}\begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix}=\lambda_j.$$

评注　证最大值，一般分为两步：第一步先证“$\leqslant$”；第二步再证可以取到等号. 与本题相仿，可以证明：二次型 f 在 n 维向量空间的单位球面上的最小值是二次型矩阵 A 的最小特征值.

例 6-18　A 实对称，试证：t 充分大后，$tE+A$ 是正定阵.

证　$(tE+A)^{\mathrm{T}}=tE+A^{\mathrm{T}}=tE+A$，对称. 因 A 实对称，故 A 的特征值全是实数，则由

$$\lambda E-A\Leftrightarrow(\lambda+t)E-(tE+A)$$

可知：λ 是 A 的特征值$\Leftrightarrow\lambda+t$ 是 $tE+A$ 的特征值. 当 t 充分大后，由于 λ 是实数，故可使 $\lambda+t>0$，即 t 充分大后，$tE+A$ 的特征值全大于零，故 $tE+A$ 正定.

例 6-19　若 A 实对称，$A^3-A^2+2A=2E$，证明，A 正定.

证　设 $AX=\lambda X$.

$$(A^3-A^2+2A)X=2X$$

$$\Rightarrow\quad(\lambda^3-\lambda^2+2\lambda-2)X=0$$

$$\Rightarrow\quad\lambda^3-\lambda^2+2\lambda-2=0$$

$\Rightarrow \quad (\lambda-1)(\lambda^2+2)=0$

$\Rightarrow \quad \lambda=1,(\lambda^2+2\neq 0$. 因 λ 是实数)

$\Rightarrow \quad A$ 的特征值大于零,故 A 正定.

例 6-20 证明:

$$A=\begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}$$ 是正定阵.

证 A 对称显然. 记 $|A|=D_n$.

则
$$\begin{aligned} D_n &= 2D_{n-1}-D_{n-2}=2(2D_{n-2}-D_{n-3})-D_{n-2} \\ &= 3D_{n-2}-2D_{n-3}=3(2D_{n-3}-D_{n-4})-2D_{n-3} \\ &= 4D_{n-3}-3D_{n-4}=\cdots=(n-1)D_2-(n-2)D_1 \\ &= (n-1)\cdot 3-(n-2)\cdot 2=n+1. \end{aligned}$$

可知 A 的所有顺序主子式全大于零. 故 A 正定.

评注 矩阵 A 正定,等价说法有好几种,有的按定义论证方便,有的用特征值论证方便,有的用所有顺序主子式大于零论证方便.

例 6-21 若 A 为 n 阶实对称阵,且 $|A|<0$. 试证:必存在 $0\neq x\in R^n$,使 $X^{\mathrm{T}}AX<0$.

证 因 A 实对称,故必存在正交阵 Q,使

$Q^{-1}AQ=\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$,$\lambda_i(i=1,\cdots,n)$是 A 的特征值,且均为实数.

由于 $|A|=\lambda_1\lambda_2\cdots\lambda_n<0$,知 λ_i 中必有负数,不妨设 $\lambda_1<0$,则可令 $Y_0=(1,0,\cdots,0)^{\mathrm{T}}$,可有

$$Y_0^{\mathrm{T}}Q^{\mathrm{T}}AQY_0=(1,0,\cdots,0)\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}=\lambda_1<0,$$

则只要取 $X=QY_0$ 即为所求.

例 6-22 设 $A=(a_{ij})_{n\times n}$ 正定. 证明:$\forall\, i\neq j$,都有 $|a_{ij}|<|a_{ii}\cdot a_{jj}|^{\frac{1}{2}}$.

证 把 A 的第 i 行换到第一行,第 j 行换到第二行,并把 A 的第 i 列换到第一列,第 j 列换到第二列,即

$$P(2,j)P(1,i)AP(1,i)P(2,j)=B,$$

即
$$(P(1,i)P(2,j))^{\mathrm{T}}A(P(1,i)P(2,j))=B.$$

即有 A 与 B 合同. 由于合同变换保持正定性不变,知 B 正定. 则 B 的二阶顺 序主子式大于零. 即

$$\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ij} & a_{jj} \end{vmatrix}>0. \quad \text{即}|a_{ij}|<|a_{ii}\quad a_{jj}|^{\frac{1}{2}}.$$

例 6-23 设 A,B 是两个 n 阶实对称阵,且 B 正定. 试证:存在一个 n 阶可逆阵 C,使 $C^{\mathrm{T}}AC$ 与 $C^{\mathrm{T}}BC$ 同时成为对角阵.

证 因 B 正定,所以,B 与 E 合同,即存在可逆阵 P,使 $P^{\mathrm{T}}BP=E$. 而 $P^{\mathrm{T}}AP$ 仍为实对称,则存在正交阵 Q,使 $Q^{\mathrm{T}}(P^{\mathrm{T}}AP)Q=D$. ($D$ 为对角阵). 而 $Q^{\mathrm{T}}(P^{\mathrm{T}}BP)Q=Q^{\mathrm{T}}EQ=E$.

则可令 $C=PQ$. 是可逆阵,有

$C^{\mathrm{T}}AC=D$,$C^{\mathrm{T}}BC=E$. 皆为对角阵.

评注 有两个矩阵 A,B,先从条件强的 B 开始考虑.

例 6-24 设 A,B 均为 n 阶实对称阵,且 A 为正定阵. 试证:$A\cdot B$ 的特证值都是实数.

证　因 A 正定，故 $A=U^{\mathrm{T}}U$，U 为可逆阵.

$$\begin{aligned}|\lambda E-AB| &= |\lambda E-U^{\mathrm{T}}UB| = |\lambda E-U^{\mathrm{T}}UBU^{\mathrm{T}}(U^{\mathrm{T}})^{-1}| \\ &= |\lambda U^{\mathrm{T}}(U^{\mathrm{T}})^{-1}-U^{\mathrm{T}}(UBU^{\mathrm{T}})(U^{\mathrm{T}})^{-1}| = |U^{\mathrm{T}}|\cdot|\lambda E-UBU^{\mathrm{T}}||(U^{\mathrm{T}})^{-1}| \\ &= |\lambda E-UBU^{\mathrm{T}}|.\end{aligned}$$

由于 B 实对称，$\Rightarrow UBU^{\mathrm{T}}$ 仍为实对称，故 UBU^{T} 的特征值皆为实数. 上式表明：$A\cdot B$ 与 UBU^{T} 有相同的特征多项式，即有相同的特征值. 故 $A\cdot B$ 的特征值都是实数.

例 6-25　设 A 为 m 阶实对称阵，且 A 正定. B 为 $m\times n$ 矩阵. 试证：$B^{\mathrm{T}}AB$ 为正定矩阵的充分必要条件是 B 的秩等于 n.

证　“必要性”. 设 A 与 $B^{\mathrm{T}}AB$ 是正定阵. 故 $\forall\,(0\neq x)\in R^n$，恒有 $X^{\mathrm{T}}(B^{\mathrm{T}}AB)X>0$，即

$$X^{\mathrm{T}}(B^{\mathrm{T}}AB)X=(BX)^{\mathrm{T}}A(BX)>0.$$

由于 A 正定，欲使上不等式成立，必须 $BX\neq 0$.（当 $X\neq 0$ 时）. 也即线性方程组 $BX=0$ 只有零解，即知 $R(B)=n$.

“充分性”设 $R(B)=n$，则 $BX=O$ 只有零解，即 $\forall\,X\neq O$，$BX\neq O$. 则由 A 正定可知，$\forall\,X\neq O$，恒有 $(BX)^{\mathrm{T}}A(BX)>O$，即 $X^{\mathrm{T}}(B^{\mathrm{T}}AB)X>O$，即 $B^{\mathrm{T}}AB$ 正定.

$(B^{\mathrm{T}}AB)^{\mathrm{T}}=B^{\mathrm{T}}AB$　对称显然.

评注　$R(B)=n \Leftrightarrow BX=O$ 只有零解

$\Leftrightarrow \forall\,X\neq O, BX\neq 0.$

例 6-26　设 A 为三阶实对称阵，且满足条件：

$A^2+2A=O$，　$R(A)=2$.

(1) 求 A 的全部特征值；

(2) 当 k 为何值时，矩阵 $A+kE$ 为正定阵.

解　(1) 设 $AX=\lambda X$，则 $A^2X=\lambda^2X$

$$(A^2+2A)X=(\lambda^2+2\lambda)X=O.$$

$$\Rightarrow \quad \lambda=0,\quad \lambda=-2.$$

由于 A 实对称，必可对角化，对角线上的元素即为 A 的特征值. 又已知 $R(A)=2$，故知

$$\lambda_1=\lambda_2=-2,\quad \lambda_3=0.$$

(2) $\lambda E-A \Leftrightarrow (\lambda+k)E-(A+kE)$，

即 λ 是 A 的特征值$\Leftrightarrow\lambda+k$ 是 $A+kE$ 的特征值. 即知 $A+kE$ 的特征值为 $k-2, k-2, k$.

由此可知，当 $k>2$ 时，$A+kE$ 的特征值全大于零. 即正定.

例 6-27　设矩阵

$$A=\begin{pmatrix}1 & 0 & 1\\ 0 & 2 & 0\\ 1 & 0 & 1\end{pmatrix}.$$

知阵 $B=(kE+A)^2$，其中，k 为实数. E 为单位阵. 求对角阵 Λ，使 B 与 Λ 相似. 并求 k 为何值时，B 为正定阵.

解法 1　$|\lambda E-A|=\lambda(\lambda-2)^2$.

知 A 的特征值为 $\lambda_1=\lambda_2=2, \lambda_3=0$.

由 A 实对称，故存在正交阵 P，使

$$P^{\mathrm{T}}AP=\begin{pmatrix}2 & & \\ & 2 & \\ & & 0\end{pmatrix}=D,\quad \text{即有 } A=PDP^{\mathrm{T}},$$

则

$$\begin{aligned}B &= (kE+A)^2=(kPP^{\mathrm{T}}+PDP^{\mathrm{T}})^2 \\ &= (kPP^{\mathrm{T}}+PDP^{\mathrm{T}})\cdot(kPP^{\mathrm{T}}+PDP^{\mathrm{T}})=P(kE+D)^2P^{\mathrm{T}} \\ &= P\cdot\begin{pmatrix}(k+2)^2 & 0 & 0\\ 0 & (k+2)^2 & 0\\ 0 & 0 & k^2\end{pmatrix}P^{\mathrm{T}}.\end{aligned}$$

则

$$\boldsymbol{\Lambda}=\begin{pmatrix}(k+2)^2 & 0 & 0\\ 0 & (k+2)^2 & 0\\ 0 & 0 & k^2\end{pmatrix},$$

只要 $k\neq -2$，且 $k\neq 0$ 时，$\boldsymbol{B}$ 的特征值全为正数，这时 $\boldsymbol{B}$ 正定.

解法 2 $\boldsymbol{A}$ 是实对称阵，$R(\boldsymbol{A})=2$ 是显然的.

由于 $\boldsymbol{A}$ 的每行元素之和皆为 2，知 A 有特征值 2. 又因 $R(\boldsymbol{A})=2$，知 A 有特征值 0. 又 $\boldsymbol{A}$ 必可对角化，对角阵的对角线元素就是 $\boldsymbol{A}$ 的特征值，因此，对角线上另一个元素也是 2，这可从 $\lambda_1+\lambda_2+\lambda_3=\mathrm{t}_r(\boldsymbol{A})=4$ 中求得.

因此，$\boldsymbol{A}$ 的特征值为 2，2，0. $k\boldsymbol{E}+\boldsymbol{A}$ 的特征值为 $k+2,k+2,k$. $(k\boldsymbol{E}+\boldsymbol{A})^2$ 的特征值为 $(k+2)^2,(k+2)^2,k^2$.

由于 $k\boldsymbol{E}+\boldsymbol{A}$ 对称. $(k\boldsymbol{E}+\boldsymbol{A})^2$ 也是实对称，因此，$\boldsymbol{B}=(k\boldsymbol{E}+\boldsymbol{A})^2$ 必可对角化. 而对角阵即为

$$\boldsymbol{\Lambda}=\begin{pmatrix}(k+2)^2 & 0 & 0\\ 0 & (k+2)^2 & 0\\ 0 & 0 & k^2\end{pmatrix}.$$

当 $k\neq -2$，且 $k\neq 0$ 时，$\boldsymbol{B}$ 的特征值全大于零，即 B 正定.

例 6-28 设有 n 元实二次型：

$$f(x_1,x_2,\cdots,x_n)=(x_1+a_1x_2)^2+(x_2+a_2x_3)^2+\cdots+(x_{n-1}+a_{n-1}x_n)^2+(x_n+a_nx_1)^2$$

其中，$a_i(i=1,\cdots,n)$ 为实数. 试问：当 $a_1,a_2,\cdots,a_n$ 满足何种条件时，$f(x_1,x_2,\cdots,x_n)$ 为正定二次型.

解 由题设知，对任意的 $x_1,x_2,\cdots,x_n$，有 $f(x_1,x_2,\cdots,x_n)\geqslant 0$，其中，等号成立当且仅当

$$\begin{cases}x_1+a_1x_2=0\\ x_2+a_2x_3=0\\ \vdots\\ x_{n-1}+a_{n-1}x_n=0\\ x_n+a_nx_1=0\end{cases}\tag{1}$$

方程组(1)仅有零解的充分必要条件是：

$$\begin{vmatrix}1 & a_1 & 0 & \cdots & \cdots & 0 & 0\\ 0 & 1 & a_2 & \cdots & \cdots & 0 & 0\\ 0 & 0 & 1 & a_3 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & \cdots & 1 & a_{n-1}\\ a_n & 0 & 0 & \cdots & \cdots & 0 & 1\end{vmatrix}=1+(-1)^{n+1}\cdot a_1a_2\cdots a_n\neq 0,$$

因此，当 $1+(-1)^{n+1}a_1a_2\cdots a_n\neq 0$ 时，对于任意的不全为零的 $x_1,x_2,\cdots,x_n$ 有 $f(x_1,x_2,\cdots,x_n)>0$，即当 $a_1a_2\cdots a_n\neq(-1)^n$ 时，二次型 $f(x_1,x_2,\cdots,x_n)$ 是正定二次型.

例 6-29 设 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots\boldsymbol{v}_n$ 是 n 个两两正交的 n 维单位列向量，σ 是实数，n 阶矩阵 $\boldsymbol{H}=\boldsymbol{E}-\sigma\boldsymbol{v}_1\boldsymbol{v}_1^{\mathrm{T}}$.

(1) 证明 $\boldsymbol{H}$ 是对称阵；

(2) 证明 $\boldsymbol{v}_1$ 是 $\boldsymbol{H}$ 的特征向量，并求相应特征值；

(3) 证明 $\boldsymbol{v}_2,\boldsymbol{v}_3,\cdots\boldsymbol{v}_n$ 也是 $\boldsymbol{H}$ 的特征向量，并求相应特征值；

(4) σ 为何值时，$\boldsymbol{H}$ 是正交阵；

(5) σ 在何范围内，$\boldsymbol{H}$ 是正定阵；

(6) σ 在何范围内，$\boldsymbol{H}$ 是可逆阵；

(7) 若 $\boldsymbol{H}$ 可逆，且 $\boldsymbol{H}^{-1}=\boldsymbol{E}-\mu\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}}$（$\mu$ 为实数），求 μ 与 σ 应满足的关系式.

证　(1) $\boldsymbol{H}^{\mathrm{T}}=(\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}}=\boldsymbol{H}$,即 $\boldsymbol{H}$ 对称.

(2) $\boldsymbol{H}\boldsymbol{v}_1=(\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}})\boldsymbol{v}_1=\boldsymbol{v}_1-\sigma\boldsymbol{v}_1=(1-\sigma)\boldsymbol{v}_1$. 且 $\lambda_1=1-\sigma$.

(3) $\boldsymbol{H}\boldsymbol{v}_2=(\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}})\boldsymbol{v}_2=\boldsymbol{v}_2-0=\boldsymbol{v}_2$,知 $\lambda_2=1$. 类似可证 $\boldsymbol{v}_3,\cdots,\boldsymbol{v}_n$ 是特征向量,且 $\lambda_3=\cdots=\lambda_n=1$.

(4) $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}=(\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}})(\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}})=(\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}})\cdot(\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}})$

$=\boldsymbol{E}-2\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}}+\sigma^2\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}}=\boldsymbol{E}-(2\sigma-\sigma^2)\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}}$.

当 $2\sigma-\sigma^2=0$ 时,即 $\sigma=0$ 或 $\sigma=2$ 时,$\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}=\boldsymbol{E}$,$\boldsymbol{H}$ 为正交阵.

(5) 由(2),(3)可知,当 $\sigma<1$ 时,$\boldsymbol{H}$ 的所有特征值全大于零,$\boldsymbol{H}$ 正定 。

(6) 由(2),(3)可知,当 $\sigma\neq1$ 时,所有特征值不等于零,$\boldsymbol{H}$ 可逆.

(7) $\boldsymbol{H}\cdot\boldsymbol{H}^{-1}=(\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}})\cdot(\boldsymbol{E}-\mu\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}})=\boldsymbol{E}-\sigma\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}}-\mu\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}}+\sigma\cdot\mu\boldsymbol{v}_1\cdot\boldsymbol{v}_1^{\mathrm{T}}$,

应有 $-\sigma-\mu+\sigma\mu=0$,即 $\mu=\dfrac{\sigma}{\sigma-1}$,$(\sigma\neq1)$.

例 6-30　设 $\boldsymbol{A},\boldsymbol{B}$ 均为实对称阵,若 $\boldsymbol{A}$ 的特征值全大于 a,$\boldsymbol{B}$ 的特征值全大于 b,试证:

(1) $\boldsymbol{A}-a\boldsymbol{E}$ 的特征值全大于零;

(2) $\boldsymbol{A}+\boldsymbol{B}-(a+b)\boldsymbol{E}$ 是正定矩阵;

(3) $\boldsymbol{A}+\boldsymbol{B}$ 的特征值全大于 $a+b$.

证　(1) $\lambda\boldsymbol{E}-\boldsymbol{A}\Leftrightarrow(\lambda-a)\boldsymbol{E}-(\boldsymbol{A}-a\boldsymbol{E})$

即 λ 是 $\boldsymbol{A}$ 的特征值$\Leftrightarrow\lambda-a$ 是 $\boldsymbol{A}-a\boldsymbol{E}$ 的特征值. 现已知 $\lambda>a$,故 $\boldsymbol{A}-a\boldsymbol{E}$ 的特征值 $\lambda-a>0$.

(2) 同理可证:$\boldsymbol{B}-b\boldsymbol{E}$ 的特征值全大于零. 由于 $\boldsymbol{A},\boldsymbol{B}$ 均为实对称阵,知 $\boldsymbol{A}-a\boldsymbol{E}$ 与 $\boldsymbol{B}-b\boldsymbol{E}$ 均为实对称阵,它们的特征值又全大于零,故 $\boldsymbol{A}-a\boldsymbol{E}$ 与 $\boldsymbol{B}-b\boldsymbol{E}$ 均为正定阵. 所以,$\forall\,\boldsymbol{X}\neq0$,

$$\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{B}-(a+b)\boldsymbol{E})\boldsymbol{X}=\boldsymbol{X}^{\mathrm{T}}((\boldsymbol{A}-a\boldsymbol{E})+(\boldsymbol{B}-b\boldsymbol{E}))\boldsymbol{X}=\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{A}-a\boldsymbol{E})\boldsymbol{X}+\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{B}-b\boldsymbol{E})\boldsymbol{X}>0.$$

所以,$\boldsymbol{A}+\boldsymbol{B}-(a+b)\boldsymbol{E}$ 是正定阵. (显然它也是对称阵).

(3) $\lambda\boldsymbol{E}-(\boldsymbol{A}+\boldsymbol{B})\Leftrightarrow(\lambda-(a+b))\boldsymbol{E}-(\boldsymbol{A}+\boldsymbol{B}-(a+b)\boldsymbol{E})$

即 λ 是 $\boldsymbol{A}+\boldsymbol{B}$ 的特征值$\Leftrightarrow\lambda-(a+b)$是 $\boldsymbol{A}+\boldsymbol{B}-(a+b)\boldsymbol{E}$ 的特征值. 由(2)知,$\boldsymbol{A}+\boldsymbol{B}-(a+b)\boldsymbol{E}$ 是正定阵,故 $\lambda-(a+b)>0$. 即 $\lambda>a+b$,所以,$\boldsymbol{A}+\boldsymbol{B}$ 的特征值全大于 $a+b$.

例 6-31　设 $\boldsymbol{A}$ 为 n 阶实对称矩阵 ,$R(\boldsymbol{A})=n$,A_{ij} 是 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 中元素 a_{ij} 的代数余子式$(i,j=1,2,\cdots,n)$二次型

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{A_{ij}}{|\boldsymbol{A}|}X_iX_j.$$

(1) 记 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$,把 $f(x_1,x_2,\cdots,x_n)$写成矩阵形式,并证明二次型 $f(x)$的矩阵为 $\boldsymbol{A}^{-1}$;

(2) 二次型 $g(x)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 与 $f(x)$的规范形是否相同?说明理由.

解　(1) 二次型 $f(x_1,x_2,\cdots,x_n)$的矩阵形式为

$$f(x)=(x_1,x_2,\cdots,x_n)\cdot\frac{1}{|\boldsymbol{A}|}\cdot\begin{pmatrix}A_{11}&A_{21}\cdots&A_{n1}\\A_{12}&A_{22}&\cdots&A_{n2}\\\vdots&\vdots&&\vdots\\A_{1n}&A_{2n}&\cdots&A_{mn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

由 $R(\boldsymbol{A})=n$,即 $\boldsymbol{A}$ 可逆,且 $\boldsymbol{A}^{-1}=\dfrac{1}{|\boldsymbol{A}|}\cdot\boldsymbol{A}^{*}$,又 $\boldsymbol{A}$ 实对称,$(\boldsymbol{A}^{-1})^{\mathrm{T}}=(\boldsymbol{A}^{\mathrm{T}})^{-1}=\boldsymbol{A}^{-1}$. 故 $\boldsymbol{A}^{-1}$ 也是实对称. 因此,$f(x)$的矩阵为 $\boldsymbol{A}^{-1}$.

(2) $g(x)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$,　$f(x)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{X}$,

由于 $(\boldsymbol{A}^{-1})^{\mathrm{T}}\boldsymbol{A}\cdot\boldsymbol{A}^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}}=(\boldsymbol{A}^{\mathrm{T}})^{-1}=A^{-1}$,所以,$\boldsymbol{A}$ 与 $\boldsymbol{A}^{-1}$ 合同. 于是 $g(x)$与 $f(x)$有相同的规范形.

四、小　结

本章的中心问题有两类:

(1) 二次型化标准形. 用正交变换化二次型为标准形与上一章的对称阵正交相似于对角阵是同一个问

题，而以两种形式出现. 需要注意的是特征多项式有重根的情形，所求的的特征向量可能不一定正交，这时一定要先正交规范化.

(2) 正定矩阵，常用的正定阵的充要条件是：① 其特征值全大于零；② 各阶顺序主子式全大于零.

综合练习题

1. 求证 $\begin{vmatrix} ax+by & ay+bz & az+bx \\ ay+bz & az+bx & ax+by \\ az+bx & ax+by & ay+bz \end{vmatrix} = (a^2+b^2)\begin{vmatrix} x & y & z \\ y & z & x \\ z & x & y \end{vmatrix}$.

2. 求证 $\begin{vmatrix} \cos\alpha & 1 & 0 & \cdots & 0 & 0 \\ 1 & 2\cos\alpha & 1 & \cdots & 0 & 0 \\ 0 & 1 & 2\cos\alpha & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 2\cos\alpha \end{vmatrix} = \cos n\alpha.$

3. 已知三阶矩阵 $\boldsymbol{A}$ 的特征值为 $1,2,-4$，又设 $\boldsymbol{B}=\boldsymbol{A}^2+2\boldsymbol{A}$，求：

(1) 矩阵 $\boldsymbol{B}$ 的特征值及其相似对角阵；

(2) $|\boldsymbol{B}|$ 及 $|\boldsymbol{A}-5\boldsymbol{E}|$.

4. (1989 年数学(试卷一)的 3 分题)设 A 是四阶矩阵，且 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|=0$，则 $\boldsymbol{A}$ 中 ( )

(A) 必有一列元素全为 0；

(B) 必有两列元素对应成比例；

(C) 必有一列向量是其余各列向量的线性组合；

(D) 任一列向量是其他各列向量的线性组合.

5. 设 $\boldsymbol{A}=\begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & -1 \end{bmatrix}$，试计算 $(\boldsymbol{A}+2\boldsymbol{E})^{-1}(\boldsymbol{A}^2-4\boldsymbol{E})$

6. 设 $\boldsymbol{A}^k=0$(k 为正整数)，证明

$$(\boldsymbol{E}-\boldsymbol{A})^{-1}=\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^{k-1}$$

7. 设 m 次多项式 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_mx^m$，记

$$f(\boldsymbol{A})=a_0\boldsymbol{E}+a_1\boldsymbol{A}+a_2\boldsymbol{A}^2+\cdots+a_m\boldsymbol{A}^m,$$

$f(\boldsymbol{A})$称为方阵 $\boldsymbol{A}$ 的 m 次多项式.

(1) 设 $\boldsymbol{\Lambda}=\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$，求证 $\boldsymbol{\Lambda}^k=\begin{bmatrix} {\lambda_1}^k & 0 \\ 0 & {\lambda_2}^k \end{bmatrix}$，$f(\boldsymbol{\Lambda})=\begin{bmatrix} f(\lambda_1) & 0 \\ 0 & f(\lambda_2) \end{bmatrix}$；

(2) 设 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$，求证 $\boldsymbol{A}^k=\boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1}$，$f(\boldsymbol{A})=\boldsymbol{P}f(\boldsymbol{\Lambda})\boldsymbol{P}^{-1}$.

8. 证明：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$(其中 $\boldsymbol{\alpha}_1\neq 0$)线性相关的充分必要条件是至少有一个 α_i($1\leqslant i\leqslant S$)可被 $\alpha_1,\alpha_2,\cdots,\alpha_{i-1}$ 线性表示.

9. 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是一组线性无关的向量，$\boldsymbol{\beta}_i=\sum\limits_{j=1}^{r}a_{ij}\boldsymbol{\alpha}_j$，$i=1,2,\cdots,r$，证明 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性无关的充分必要条件是

$$|\boldsymbol{A}|=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{bmatrix}\neq 0.$$

10. 当把向量 $\boldsymbol{\alpha}$ 加于向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 后，向量组的秩不变，那么向量 $\boldsymbol{\alpha}$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 线性表示.

11. 非零有序向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$，若任何一个向量都不能被它前面所有的向量线性表示，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,$

$\boldsymbol{\alpha}_k$ 线性无关.

12. 求下述线性方程组的通解

$$\begin{cases}(n-1)x_1=x_2+x_3+\cdots+x_n\\(n-1)x_2=x_1+x_3+\cdots+x_n\\\cdots\\(n-1)x_n=x_1+x_2+\cdots+x_{n-1}\end{cases}$$

13. 当 λ 取何值时，线性方程组

$$\begin{cases}2x_1-x_2+x_3+x_4=1\\x_1+2x_2-x_3+4x_4=2\\x_1+7x_2-4x_3+11x_4=\lambda\end{cases}$$

有解？在有解时，求出所有解.

14. 当 a,b,c 满足什么条件时，线性方程组

$$\begin{cases}-2x_1+x_2+x_3=a\\x_1-2x_2+x_3=b\\x_1+x_2-2x_3=c\end{cases}$$

有解？在有解时，求出所有解.

15. 设 $x_1-x_2=a_1,x_2-x_3=a_2,\cdots,x_5-x_1=a_5$，求证此方程组有解的充要条件为 $\sum\limits_{i=1}^{5}a_i=0$，在有解的情况下，求出它的一组解.

16. 如果方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2=A$，称 $\boldsymbol{A}$ 为幂等阵，求证幂等阵的特征值只能是 0 和 1.

17. 设 λ 是正交阵 $\boldsymbol{A}$ 的特征值，问 $1/\lambda$ 是不是 $\boldsymbol{A}$ 的特征值？为什么？

18. (1) 正交阵如果有实特征值，则它们只能是 1 和 -1.

(2) 正交阵不同的实特征值所对应的特征向量必定正交.

19. 设 λ 是方阵 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的对应于 λ 的特征向量，问 λ 是不是方阵 $\boldsymbol{A}^{\mathrm{T}}$ 的特征值？为什么？$\boldsymbol{X}$ 是不是 $\boldsymbol{A}^{\mathrm{t}}$ 的对应于 λ 的特征向量？

20. 设方阵 $\boldsymbol{A}$ 的各行元素之和为 a，则 a 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{X}=(1,1,\cdots,1)^{\mathrm{T}}$ 是 $\boldsymbol{A}$ 的对应于特征值 a 的特征向量.

21. 写出二次型 $f(x_1,x_2,x_3)=(ax_1+bx_2+cx_3)^2$ 的矩阵形式，并求出二次型的对称阵的特征值.

22. 设 $\boldsymbol{A}$ 为正交阵，$|\boldsymbol{A}|=-1$，则 -1 一定是 $\boldsymbol{A}$ 的特征值.

简答与提示

3. (1) $\boldsymbol{B}$ 的特征值为 3,8,8，$\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}=\mathrm{diag}\{3,8,8\}$；

(2) $|\boldsymbol{B}|=192$；　$|\boldsymbol{A}-5\boldsymbol{E}|=-108$.

4. (C).

5. $\begin{bmatrix}-3&0&0\\1&-3&0\\1&1&-3\end{bmatrix}$.

6. 在 $(\boldsymbol{E}-\boldsymbol{A})^{-1}$ 等式两边同乘 $\boldsymbol{E}-\boldsymbol{A}$.

7. (1) $f(\boldsymbol{\Lambda})=a_0\boldsymbol{E}+a_1\boldsymbol{\Lambda}+a_2\boldsymbol{\Lambda}^2+\cdots+a_m\boldsymbol{\Lambda}^m$

$$=\begin{bmatrix}a_0+a_1\lambda_1+\cdots+a_m{\lambda_1}^m&0\\0&a_0+a_1\lambda_1+\cdots+a_m{\lambda_1}^m\end{bmatrix}=\begin{bmatrix}f(\lambda_1)&0\\0&f(\lambda_2)\end{bmatrix}.$$

(2) $f(\boldsymbol{A})=a_0\boldsymbol{E}+a_1\boldsymbol{A}+a_2\boldsymbol{A}^2+\cdots+a_m\boldsymbol{A}^m$

$=a_0\boldsymbol{P}\boldsymbol{P}^{-1}+a_1\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}+a_2\boldsymbol{P}\boldsymbol{\Lambda}^2\boldsymbol{P}^{-1}+\cdots+a_m\boldsymbol{P}\boldsymbol{\Lambda}^m\boldsymbol{P}^{-1}$

$=\boldsymbol{P}(a_0\boldsymbol{E}+a_1\boldsymbol{\Lambda}+\cdots a_m\boldsymbol{\Lambda}^m)\boldsymbol{P}^{-1}=\boldsymbol{P}f(\Lambda)\boldsymbol{P}^{-1}$.

8. 存在一组不全为 0 的数 $k_1,k_2,\cdots,k_s$ 使

$k_1\alpha_1+\cdots+k_s\alpha_s=0$，从 $k_s=0$ 同前讨论，直到某个 $k_i\neq 0$ 为止.

9. 必要性 $r\geqslant R(\boldsymbol{A})\geqslant R(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)\geqslant R(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)=r$. 得 $|\boldsymbol{A}|\neq 0$；

充分性 $|\boldsymbol{A}|\neq 0$ 得 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 可由 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 线性表示.

从而 $R(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)\leqslant R(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)$，得出 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 线性无关.

10. 设 $\boldsymbol{\alpha}_{i1},\cdots,\boldsymbol{\alpha}_{ir}$ 是 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k$ 的极大线性无关组，从而这两组向量等价. 得 $R(\boldsymbol{\alpha}_{i1},\cdots,\boldsymbol{\alpha}_{ir})=R(\boldsymbol{\alpha}_{i1},\cdots,\boldsymbol{\alpha}_{ir},\boldsymbol{\alpha})$得出 $\boldsymbol{\alpha}$ 可由 $\boldsymbol{\alpha}_{i1},\cdots,\boldsymbol{\alpha}_{ir}$ 线性表示，从而 $\boldsymbol{\alpha}$ 可由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k$ 线性表示.

11. 设有一组数 $\lambda_1,\cdots,\lambda_k$ 使 $\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_k\boldsymbol{\alpha}_k=0$. 从 $\lambda_k\neq 0$ 开始向前推至 $\lambda_1\neq 0$，导出矛盾.

12. $\boldsymbol{X}=(1,1,\cdots,1)^{\mathrm{T}}$

13. $\lambda=5$ 时有解. 通解为 $\boldsymbol{X}=C_1\begin{bmatrix}-1\\3\\5\\0\end{bmatrix}+C_2\begin{bmatrix}-6\\7\\0\\5\end{bmatrix}+\begin{bmatrix}\frac{4}{5}\\\frac{3}{5}\\0\\0\end{bmatrix}$.

14. $a+b+c=0$ 时有解. 通解为

$$\boldsymbol{X}=C\begin{bmatrix}1\\1\\1\end{bmatrix}+\begin{bmatrix}\frac{1}{3}(2c-b)\\\frac{1}{3}(c-b)\\0\end{bmatrix}.$$

15. 必要性. 设 $x_1,\cdots,x_5$ 是解. 得出$(x_1-x_2)+\cdots+(x_5-x_1)=\sum_{i=1}^{5}a_i=0$；

充分性. 由 $\sum_{i=1}^{5}a_i=0$，得系数阵与增广阵秩相等，通解为

$$\begin{bmatrix}x_1\\\vdots\\x_5\end{bmatrix}=x_5\begin{bmatrix}1\\\vdots\\1\end{bmatrix}+\begin{bmatrix}a_1+\cdots+a_4\\a_2+a_3+a_4\\a_3+a_4\\a_4\\0\end{bmatrix}.$$

16. 由 $\boldsymbol{AX}=\lambda\boldsymbol{X}$，$\boldsymbol{A}^2=\boldsymbol{A}$ 得 $\lambda^2\boldsymbol{X}=\lambda\boldsymbol{X}$，又 $\boldsymbol{X}\neq 0$，得出 $\lambda=0$ 或 $\lambda=1$.

17. 由 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$，知 $|\boldsymbol{A}|\neq 0$，再由 $\boldsymbol{AX}=\lambda\boldsymbol{X}$ 得 $\frac{1}{\lambda}\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}$. 从而 $\frac{1}{\lambda}$ 是 $\boldsymbol{A}^{\mathrm{T}}$ 的特征值，又 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=|\lambda\boldsymbol{E}-\boldsymbol{A}^{\mathrm{T}}|$ 得 $\frac{1}{\lambda}$ 是 $\boldsymbol{A}$ 的特征值.

18. (1) 由 $\boldsymbol{AX}=\lambda\boldsymbol{X}$，$x\neq 0$，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$ 得 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\lambda\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$，从而 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}=\lambda^2\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$，若 λ 是实数，则 $\lambda=\pm 1$；

(2) 由 $\boldsymbol{AX}_1=\lambda_1\boldsymbol{X}_1$，$\boldsymbol{AX}_2=\lambda_2\boldsymbol{X}_2$ 得 $\lambda_1\lambda_2\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2=\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2$ 由 $\lambda_1\neq\lambda_2$，且 $\lambda_i=\pm 1$，得 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 正交.

19. 由 18 题已得若 λ 是 $\boldsymbol{A}$ 的特征值，则 λ 是 $\boldsymbol{A}^{\mathrm{T}}$ 的特征值. 但 $\boldsymbol{X}$ 未必是 $\boldsymbol{A}^{\mathrm{T}}$ 的特征向量，例如 $\boldsymbol{A}=\begin{bmatrix}0&1\\4&0\end{bmatrix}$；$\lambda_1=2$，$\boldsymbol{X}_1(1,2)^{\mathrm{T}}$；$\lambda_2=-2$，$\boldsymbol{X}_2=(1,-2)^{\mathrm{T}}$，但 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}_1\neq 2\boldsymbol{X}_1$.

20. $\boldsymbol{AX}=\boldsymbol{A}\begin{bmatrix}1\\\vdots\\1\end{bmatrix}=\begin{bmatrix}a\\\vdots\\a\end{bmatrix}=a\begin{bmatrix}1\\\vdots\\1\end{bmatrix}=a\boldsymbol{X}$. 结论成立.

21. $f=(x_1,x_2,x_3)\begin{bmatrix}a^2&ab&ac\\ab&b^2&bc\\ac&bc&c^2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$.

概率统计

概率统计是数学(试卷一、试卷三和试卷四)考试的内容之一.按照考试大纲中试卷结构的要求,“概率论与数理统计”在试卷一中约占20%,在试卷三和试卷四中约占25%.

根据考试大纲所规定的范围,把概率论分成五章加以叙述,即随机事件和概率,随机变量及其(概率)分布,二维随机变量及其(概率)分布,随机变量的数字特征,大数定律和中心极限定理;把数理统计分成三章加以叙述,即数理统计的基本概念,参数估计,假设检验。还单列了综合能力训练一章以提高考生综合运用概率统计知识的能力。在每一章中选讲了一些典型的例题,并在求解之后作了若干分析. 这些例题大部分取材于历年试题.

概率统计考试大纲的内容虽然庞杂,但究其实质,无非是六类问题:概率论中求概率、求分布和求数字特征、数理统计中求点估计、求区间估计与求检验的拒绝域,为此,最后有针对性地给出一些习题.从形式上看,这些习题未必与试题雷同,但内容是与试题紧扣的,认真完成这些习题,将有助于提高考生的应试能力.

第一章　随机事件和概率

随机事件及其概率是概率论中最基本的概念.

一、复习与考试要求

(1) 理解随机事件的概念,了解样本空间的概念,掌握事件之间的关系与运算.

(2) 了解概率的定义,掌握概率的基本性质与应用这些性质进行概率计算.

(3) 理解条件概率的概念,掌握概率的加法公式、乘法公式、全概率公式、贝叶斯公式以及应用这些公式进行概率计算.

(4) 理解事件的独立性概念,掌握应用事件独立性进行概率计算.

(5) 掌握伯努利概型及其二项概率.

二、基本概念与理论

1. 随机试验与样本空间

具有下列三个特性的试验称为随机试验(简称试验):

(1) 试验可以在相同的条件下重复进行;

(2) 每次试验的可能结果不止一个,并且事先明确知道试验的所有可能结果(称所有可能结果所组成的集合为试验的样本空间).

(3) 进行一次试验之前不能确定哪一个结果会出现.

2. 随机事件

在随机试验中,把在一次试验中可能出现也可能不出现,而在大量重复试验中具有某种规律性的事情称为随机事件(简称事件). 通常把必然事件(记作 Ω)与不可能的事件(记作 $\varnothing$)看作特殊的随机事件.

事件之间有下列关系:

(1) 包含　如果事件 A 发生必然导致事件 B 发生,那么称事件 B 包含事件 A,记作 $A\subset B$(或 $B\supset A$).

(2) 相等　如果 $A\supset B$ 且 $B\supset A$,那么,称事件 A 与事件 B 相等,记作 $A=B$.

(3) 互不相容　如果事件 A 与事件 B 不能同时发生,那么称事件 A 与 B 互不相容(或互斥),记作 $AB=\varnothing$.

事件之间有下列运算:

(1) 和事件　"事件 A 与事件 B 中至少有一个发生"是一个事件,称这个事件为事件 A 与事件 B 的和事件,记作 $A\cup B$.

和事件可推广到 $\bigcup_{i=1}^{n}A_i$ 与 $\bigcup_{i=1}^{\infty}A_i$. $\bigcup_{i=1}^{n}A_i$ 表示"事件 $A_1,A_2,\cdots,A_n$ 中至少有一个发生". 而 $\bigcup_{i=1}^{\infty}A_i$ 表示"事件 $A_1,A_2,\cdots$ 中至少有一个发生".

(2) 积事件　"事件 A 与事件 B 同时发生"是一个事件,称这个事件为事件 A 与事件 B 的积事件,记作 $A\cap B$(或 AB).

积事件可推广到 $\bigcap_{i=1}^{n}A_i$ 与 $\bigcap_{i=1}^{\infty}A_i$.

(3) 对立事件　"事件 A 不发生"是一个事件,称这个事件为事件 A 的对立事件(或逆事件),记作 $\overline{A}$.

(4) 差事件　"事件 A 发生且事件 B 不发生"是一个事件,称这个事件为事件 A 与事件 B 的差事件,记作 $A-B$(或 $A\overline{B}$).

事件之间的运算满足下述定律:

(ⅰ) **交换律** $A\cup B=B\cup A, A\cap B=B\cap A$;

(ⅱ) **结合律** $A\cup(B\cup C)=(A\cup B)\cup C$,

$A\cap(B\cap C)=(A\cap B)\cap C$;

(ⅲ) **分配律** $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$,

$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$;

(ⅳ) **德摩根法则** $\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i}$ $\overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A_i}$.

3. 概率的定义

直观上说,概率是随机事件发生的可能性大小的度量. 记事件 A **发生的概率为** $P(A)$.

(1) **古典概率的定义**

如果随机试验的样本空间是一个有限集,而且样本空间中每个试验结果的出现具有等可能性,那么,规定事件 A **发生的概率为**

$$P(A)=\frac{A\text{ 中所含元素的个数}}{\text{样本空间中所含元素的个数}}.$$

(2) **几何概率的定义**

如果随机试验的样本空间是一个区域(例如直线上的区间、平面或空间中的区域),而且样本空间中每个试验结果的出现具有等可能性,那么,规定事件 A **发生的概率为**

$$P(A)=\frac{A\text{ 的长度(或面积、体积)}}{\text{样本空间的长度(或面积、体积)}}$$

(3) **概率的统计定义**

在进行大量重复试验时,随机事件发生的频率具有稳定性。规定事件 A **发生的频率的稳定值** p **为事件** A **发生的概率,即** $P(A)=p$.

(4) **概率的公理化定义**

给定一个随机试验. 对于任意一个随机事件 A**,规定一个实数,记作** $P(A)$**. 如果函数** P **满足下列三条公理,就称** $P(A)$**为事件** A **的概率:**

公理 1 (非负性)**对任意一个事件** A,$P(A)\geqslant 0$;

公理 2 (规范性)$P(\Omega)=1$;

公理 3 (可列可加性)**如果事件** $A_1, A_2, \cdots$ **两两互不相容,那么**

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i).$$

上述三条公理蕴含了

$$0\leqslant P(A)\leqslant 1,$$

$$P(\varnothing)=0.$$

(5) **条件概率**

设 A **与** B **是两个事件. 在事件** B **发生的条件下事件** A **发生的概率称为条件概率,记作** $P(A|B)$**,当** $P(B)>0$**时,规定**

$$P(A|B)=\frac{P(AB)}{P(B)}.$$

在同一条件下的条件概率具有概率的一切性质. 例如 $P(\Omega|B)=1, P(\phi|B)=0, 0\leqslant P(A|B)\leqslant 1$ **等.**

4. 概率的计算公式

(1) **求逆公式** $P(\overline{A})=1-P(A)$.

(2) **加法公式** $P(A\cup B)=P(A)+P(B)-P(AB)$**. 如果** A **与** B **互不相容,那么**

$$P(A\cup B)=P(A)+P(B).$$

一般地,如果 $A_1, A_2, \cdots, A_n$ **两两互不相容,那么,**

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n).$$

(3) **求差公式**　$P(A-B)=P(A)-P(AB)$. **如果** $A\supset B$, **那么** $P(A)\geqslant P(B)$, **且**

$$P(A-B)=P(A)-P(B).$$

(4) **乘法公式**　$P(AB)=P(A|B)P(B)=P(B|A)P(A)$. **如果** A **与** B **相互独立, 那么**, $P(AB)=P(A)P(B)$. **一般地, 如果** $A_1,A_2,\cdots,A_n$ **相互独立, 那么**

$$P(A_1\cap A_2\cap\cdots\cap A_n)=P(A_1)P(A_2)\cdots P(A_n).$$

(5) **全概率公式**　**如果** $A_1,A_2,\cdots,A_n$ **构成一个完备事件组(即** $A_1,A_2,\cdots,A_n$ **两两互不相容, 且** $\bigcup_{i=1}^{n}A_i=\Omega$), **且** $P(A_i)>0,i=1,2,\cdots,n$, **那么**

$$P(B)=\sum_{i=1}^{n}P(B|A_i)P(A_i).$$

(6) **贝叶斯公式**　**如果** $A_1,A_2,\cdots,A_n$ **构成一个完备事件组, 且** $P(A_i)>0,i=1,2,\cdots,n$, **那么, 当** $P(B)>0$ **时,**

$$P(A_k|B)=\frac{P(B|A_k)P(A_k)}{P(B)}=\frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{n}P(B|A_i)P(A_i)},k=1,2,\cdots,n$$

5. 随机事件的相互独立性

如果事件 A **与** B **满足**

$$P(AB)=P(A)P(B)$$

那么, 称事件 A **与** B **相互独立.**

定理　**如果** $P(A)>0$(**或** $P(B)>0$), **那么, 事件** A **与** B **相互独立的充分必要条件是** $P(B|A)=P(B)$(**或** $P(A|B)=P(A)$).

这个定理表明: 事件 A **与** B **相互独立的直观意义是"事件** A **与** B **发生与否互不影响".**

定理　**下列四个命题是等价的:**

(1) **事件** A **与** B **相互独立;**

(2) **事件** A **与** $\overline{B}$ **相互独立;**

(3) **事件** $\overline{A}$ **与** B **相互独立;**

(4) **事件** $\overline{A}$ **与** $\overline{B}$ **相互独立.**

事件的相互独立性这一概念可以推广到多个随机事件上: 如果事件 $A_1,\cdots,A_n$ **满足**

$$P(A_{i_1}\cap\cdots\cap A_{i_k})=P(A_{i_1})\cdots P(A_{i_k}),$$

其中 $k=2,\cdots,n,i_1,i_2,\cdots,i_k$ **是** $1,2,\cdots,n$ **中任意** k **个不同的数, 那么, 称事件** $A_1,\cdots,A_n$ **相互独立. 注意, 这里实际上包含了** 2^n-n-1 **个等式.**

6. 伯努利概型与二项概率

一次试验中, 事件 A **发生的概率** $P(A)=p,(0<p<1)$, **独立重复地作** n **次试验, 在这** n **次试验中** A **恰发生** k **次的概率为**

$$P_n(k)=\binom{n}{k}p^k(1-p)^{n-k},k=0,1,\cdots,n,$$

称这组概率为二项概率.

三、例题分析

例 1-1　**已知** $P(A)=0.4,P(B)=0.3,P(A\cup B)=0.6$, **求** $P(A\overline{B})$.

解　**由** $P(A\cup B)=P(A)+P(B)-P(AB)=0.6$, **可得** $P(AB)=0.1$, **因此**

$$P(A\overline{B})=P(A-B)=P(A)-P(AB)=0.4-0.1=0.3.$$

分析　**本题中事件** A **与** B **不具有特殊的关系, 因此, 计算** $P(A-B)$**时要使用不附加前提的公式. 仔细观察求解过程, 可以发现条件"**$P(A)=0.4$**"是多余的, 因为由加法公式** $P(A\cup B)=P(A)+P(B)-P(AB)$ **可知**

$$P(A)-P(AB)=P(A\cup B)-P(B),$$

因此 $$P(A\bar{B})=P(A-B)=P(A)-P(AB)=P(A\cup B)-P(B)=0.6-0.3=0.3.$$

本题是 1990 年试题. 添加多余条件的后果是降低了难度.

例 1-2 设两两相互独立的三事件 A,B 和 C 满足条件：$ABC=\phi$，$P(A)=P(B)=P(C)<\frac{1}{2}$，且已知 $P(A\cup B\cup C)=\frac{9}{16}$，求 $P(A)$.

解 由加法公式可得

$$\begin{aligned}P(A\cup B\cup C)&=P(A\cup B)+P(C)-P((A\cup B)C)\\&=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC),\end{aligned}$$

由 A,B,C 两两相互独立，且 $P(A)=P(B)=P(C)$ 得

$$P(AB)=P(A)P(B)=(P(A))^2,P(AC)=(P(A))^2.$$

$$P(BC)=(P(A))^2$$

$$P(A\cup B\cup C)=3P(A)-3(P(A))^2$$

记 $P(A)=x$，则

$$\frac{9}{16}=3x-3x^2$$

解得 $x_1=\frac{1}{4},x_2=\frac{3}{4}$，由题意 $P(A)<\frac{1}{2}$，故应取 $x=\frac{1}{4}$ 即 $P(A)=\frac{1}{4}$.

分析 本题中 $P(A\cup B\cup C)$ 的计算公式是解题的关键，由于 A,B,C 只是两两相互独立，因此 A,B,C 并不相互独立. 建立方程求解也是求概率常用的技巧.

本题是 1999 年试题.

例 1-3 已知 $P(A)=0.6,P(B)=0.8,P(B|\bar{A})=0.2$，求 $P(A|B)$.

解 由于

$$P(\bar{A}B)=P(B|\bar{A})P(\bar{A})=0.2\times0.4=0.08$$

$$P(\bar{A}B)=P(B-A)=P(B)-P(AB),$$

因此 $$P(AB)=P(B)-P(\bar{A}B)=0.8-0.08=0.72,$$

从而

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{0.72}{0.8}=0.9.$$

分析 在事件 A 与 B 不具有特殊关系时，先设法求出 $P(AB)$，这是解题的关键.

例 1-4 已知 $P(A)=P(B)=P(C)=\frac{1}{4}$，$P(AB)=0$，$P(AC)=P(BC)=\frac{1}{6}$，求事件 A,B,C 全不发生的概率.

解 由德摩根法则及求逆公式，事件 A,B,C 全不发生的概率

$$P(\bar{A}\bar{B}\bar{C})=P(\overline{A\cup B\cup C})=1-P(A\cup B\cup C),$$

再由 $$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC),$$

及 $ABC\subset AB$ 推得

$$0\leqslant P(ABC)\leqslant P(AB)=0.$$

即 $P(ABC)=0$，因此

$$P(A\cup B\cup C)=\frac{1}{4}+\frac{1}{4}+\frac{1}{4}-0-\frac{1}{6}-\frac{1}{6}+0=\frac{5}{12},$$

从而

$$P(\bar{A}\bar{B}\bar{C})=1-\frac{5}{12}=\frac{7}{12}.$$

分析 本题表面上看缺条件“$P(ABC)$ 的值”，但利用概率的性质可以推得 $P(ABC)=0$. 这种方法不能误用，它只能用在概率为 0 的事件上. 这里选用了 AB，而不是 BC 与 CA，虽然 ABC 同样是它们的一个子集.

本题是 1992 年试题.细心的读者也许会发现:本题所给的已知条件中数据是有矛盾的.因为一方面由 $AC\subset C, BC\subset C$.推得 $(AC\cup BC)\subset C$,从而应有

$$P(AC\cup BC)\leqslant P(C)=\frac{1}{4};$$

另一方面,由加法公式推得

$$P(AC\cup BC)=P(AC)+P(BC)-P(ABC)=\frac{1}{6}+\frac{1}{6}-0=\frac{1}{3}>\frac{1}{4}.$$

考生不必理会这些问题,因为本题的目的是检查考生是否掌握前述技巧。

例 1-5　设三次独立试验中,事件 A 出现的概率相等,已知 A 至少出现一次的概率为 $\frac{19}{27}$,求事件 A 在一次试验中出现的概率.

解　设事件 A 在一次试验中出现的概率为 P.三次独立试验中,"A 至少出现一次"的逆事件为"A 出现 0 次"由求逆公式得到,"A 出现 0 次"的概率为 $1-\frac{19}{27}=\frac{8}{27}$;另一方面,由二项概率

$$P_3(0)=\binom{3}{0}p^0(1-p)^{3-0}=(1-p)^3$$

因此,由 $(1-p)^3=\frac{8}{27}$ 得到 $p=\frac{1}{3}$.

分析　本题使用了两个技巧:一是把事件转化成其逆事件,因为直接讨论将导致复杂的计算;二是建立方程求解,这个技巧已在例 1-2 中用过。上述两个技巧在概率论中是很有用的.

例 1-6　随机地向半圆 $0<y<\sqrt{2ax-x^2}$ (a 为正常数)内扔一个点,点落在半圆内任何区域内的概率与区域的面积成正比,求原点与该点的连线与 x 轴的夹角小于 $\frac{\pi}{4}$ 的概率.

解　设事件 A 为"原点与投点的连线与 x 轴的夹角小于 $\frac{\pi}{4}$".样本空间 $\Omega=\{(x,y)\mid 0<y<\sqrt{2ax-x^2}\}$,$\Omega$ 的面积为 $\frac{1}{2}\pi a^2$;$A=\{(x,y)\mid (x,y)\in\Omega, y<x\}$,$A$ 由 $\frac{1}{4}$ 个圆与一等腰直角三角形组成,A 的面积为 $S_A=\frac{1}{4}\pi a^2+\frac{1}{2}a^2$,由几何概率计算公式得到

$$P(A)=\frac{S_A}{S_\Omega}=\frac{\frac{1}{4}\pi a^2+\frac{1}{2}a^2}{\frac{1}{2}\pi a^2}=\frac{1}{2}+\frac{1}{\pi}.$$

分析　题目中"随机地"表示试验结果的出现具有等可能性,"点落在半圆内任何区域内的概率与区域的面积成正比"重复强调了试验结果等可能性出现,且暗示需用二维情形下的几何概率公式.解决本题涉及计算区域的面积.

例 1-7　假设 $P(A)=0.4, P(A\cup B)=0.7$.

(1) 若 A 与 B 互不相容,求 $P(B)$;

(2) 若 A 与 B 相互独立,求 $P(B)$;

解　(1) 由于 $AB=\phi$,因此

$$P(A\cup B)=P(A)+P(B),$$

$$P(B)=P(A\cup B)-P(A)=0.7-0.4=0.3.$$

(2) 由加法公式和 A, B 相互独立得到

$$P(A\cup B)=P(A)+P(B)-P(AB)=P(A)+P(B)-P(A)P(B).$$

因此

$$P(B)=\frac{P(A\cup B)-P(A)}{1-P(A)}=\frac{0.7-0.4}{1-0.4}=0.5.$$

分析　解本题要求分清两个事件互不相容和相互独立之间的差别.A, B 互不相容是指 A, B 不能同时发

生. 而 A,B 相互独立则是表示 $P(AB)=P(A)P(B)$. 这类题目在历年考试中很常见. 注意由 $P(AB)=0$ 无法推出 $AB=\phi$.

本题是 1988 年试题.

例 1-8 对于任意两事件 A 和 B,

(A) 若 $AB\neq\phi$,则 A,B 一定独立;

(B) 若 $AB\neq\phi$,则 A,B 有可能独立;

(C) 若 $AB=\phi$,则 A,B 一定独立;

(D) 若 $AB=\phi$,则 A,B 一定不独立.

应是(B).

分析 本题也是区分互不相容和相互独立这两个概念的题,这二者之间无因果关系,试举一例如下. 设 X 服从 $N(0,1)$,若记 $A=\{X\geqslant 1\}$,$B=\{X\leqslant 2\}$,则 $P(AB)=P(1\leqslant X\leqslant 2)=\Phi(2)-\Phi(1)$,$P(A)=1-\Phi(1)$,$P(B)=\Phi(2)$,由此推出 $AB\neq\phi$,且 A,B 不相互独立,即(A)不成立. 若取 $A=\{X=1\}$,$B=\{X\geqslant 1\}$,则 $P(AB)=P(X=1)=0$,$P(A)=0$,$P(B)=1-\Phi(1)\neq 0$,由于 $P(AB)=P(A)P(B)=0$,故 A,B 相互独立且 $AB\neq\phi$,即(B)成立. 考生可自行验证:若取 $A=\{X\geqslant 1\}$,$B=\{X<1\}$时,$AB=\phi$,A,B 不相互独立,即(C)不成立. 若取 $A=\{X=1\}$,$B=\{X=2\}$时,$AB=\phi$ 且 A,B 相互独立,即(D)不成立.

本题是 2003 年试题.

例 1-9 在区间(0,1)中随机地取出两个数,求两数之和小于 1.2 的概率.

解 设 x,y 为区间(0,1)中随机地取出的两个数,事件 A 表示"$x+y<1.2$". 样本空间 $\Omega=\{(x,y)\mid 0<x<1,0<y<1\}$,$\Omega$ 的面积为 1;$A=\{(x,y)\mid (x,y)\in\Omega,x+y<1.2\}$,$A$ 由 Ω 扣除一个等腰直角三角形组成,A 的面积为 $1-\frac{1}{2}\times 0.8^2=0.68$. 由几何概率计算公式得到

$$P(A)=\frac{0.68}{1}=0.68$$

分析 本题的难点是把所求问题归结为一个几何概率问题. 读者若有困难,不妨把它理解成这样一个问题:"X 与 Y 相互独立,且 X,Y 服从区间(0,1)上的均匀分布,求 $P(X+Y<1.2)$"。遗憾的是题目中没有明确给出"独立性"这一条件,而代之以概念比较模糊的字眼"随机地". 在本题中,"随机地"除了表示试验结果出现具有等可能性外,还表示取出的两个数相互独立. 本题是 1988 年试题.

例 1-10 有三个箱子,第一个箱子有四个黑球和一个白球,第二个箱子有三个黑球和三个白球,第三个箱子有三个黑球和五个白球. 现随机地取出一个箱子,再从这个箱子中取出一个球.

(1) 求取出的这个球是白球的概率;

(2) 已知取出的是白球,求此球属于第二个箱子的概率.

解 设 A_i 表示"取出第 i 个箱子",$i=1,2,3$.

$$P(A_1)=P(A_2)=P(A_3)=\frac{1}{3}.$$

设 B 表示"取出的这个球是白球",

$$P(B|A_1)=\frac{1}{5},P(B|A_2)=\frac{3}{6},P(B|A_3)=\frac{5}{8}.$$

(1) 由全概率公式得到

$$P(B)=\sum_{i=1}^{3}P(B|A_i)P(A_i)=\frac{1}{5}\times\frac{1}{3}+\frac{3}{6}\times\frac{1}{3}+\frac{5}{8}\times\frac{1}{3}=\frac{53}{120}.$$

(2) 由贝叶斯公式得到

$$P(A_2|B)=\frac{P(B|A_2)P(A_2)}{P(B)}=\frac{\frac{3}{6}\times\frac{1}{3}}{\frac{53}{120}}=\frac{20}{53}.$$

分析 本题是 1987 年试题. 本题是考查全概率公式与贝叶斯公式的典型试题. 识别用全概率公式的方

法是：随机试验分成两个阶段，且第一阶段的试验结果不明确给出，而最终结果是已知的，现在要追究最终试验结果是第一阶段的哪一个试验结果所造成的. 在复习时，要有意识地把这两类问题联系起来.

例 1-11　甲、乙两人对同一目标进行射击，命中率分别为 0.6，0.5. 在下列两种情形下，分别求事件“已知目标被命中，它是甲射中”的概率.

(1) 在甲、乙两人中随机地挑选一人，由他射击一次；

(2) 甲、乙两个独立地各射击一次.

解　(1) 设事件 A_1 表示“甲射击”，事件 A_2 表示“乙射击”，$P(A_1)=P(A_2)=\frac{1}{2}$.

设 B 表示“命中目标”，

$$P(B|A_1)=0.6,\ P(B|A_2)=0.5,$$

由全概率公式

$$P(B)=\sum_{i=1}^{2}P(B|A_i)P(A_i)=0.6\times\frac{1}{2}+0.5\times\frac{1}{2}=0.55;$$

由贝叶斯公式

$$P(A_1|B)=\frac{P(B|A_1)P(A_1)}{P(B)}=\frac{0.6\times\frac{1}{2}}{0.55}=\frac{6}{11}.$$

(2) 设事件 A 表示“甲射中”，事件 B 表示“乙射中”，$A\cup B$ 表示“命中目标”，所求概率为

$$P(A|A\cup B)=\frac{P(A\cap(A\cup B))}{P(A\cup B)}=\frac{P(A)}{P(A\cup B)}.$$

由于 A,B 相互独立，因此

$$\begin{aligned}P(A\cup B)&=P(A)+P(B)-P(AB)=P(A)+P(B)-P(A)P(B)\\&=0.6+0.5-0.6\times0.5=0.8.\end{aligned}$$

由此得到

$$P(A|A\cup B)=\frac{0.6}{0.8}=0.75.$$

分析　两种不同的试验导致不同的结果. 弄清随机试验的全过程是正确解题的前提. 在(2)中也可以用求逆公式

$$\begin{aligned}P(A\cup B)&=1-P(\overline{A\cup B})=1-P(\bar{A}\bar{B})=1-P(\bar{A})P(\bar{B})\\&=1-(1-P(A))(1-P(B))=1-0.4\times0.5=0.8.\end{aligned}$$

这种方法对多个射手射击的情形特别有效. 一般地，若 $A_1,A_2,\cdots,A_n$ 相互独立，则有

$$P(A_1\cup A_2\cup\cdots\cup A_n)=1-(1-P(A_1))(1-P(A_2))\cdots(1-P(A_n)).$$

例 1-12　袋中有 50 个乒乓球，其中 20 个是黄球，30 个是白球. 今有两人依次随机地从袋中各取一球，取后不放回. 求第二个人取得黄球的概率.

解　设 A_i 表示“第 i 个人取得黄球”，$i=1,2$.

$$P(A_1)=\frac{20}{50},\quad P(\bar{A}_1)=\frac{30}{50};$$

$$P(A_2|A_1)=\frac{19}{49},\quad P(A_2|\bar{A}_1)=\frac{20}{49},$$

由全概率公式得到

$$P(A_2)=P(A_2|A_1)P(A_1)+P(A_2|\bar{A}_1)P(\bar{A}_1)=\frac{19}{49}\times\frac{20}{50}+\frac{20}{49}\times\frac{30}{50}=0.4$$

分析　本题是 1997 年试题. 类似的试题也曾出现在 1993 年的试卷上. 题目本意是考查全概率公式. 由于这类题目在日常生活中的原型即是“抓阄”，因此有经验的读者应该不假思索地知道：先取与后取的概率是相等的，这对处理填空(或选择)题更具有特殊的意义.

四、小　结

本章的重点是概率计算.一要熟练常握各种概率计算公式,特别要注意在事件满足一定关系(包含、互斥、独立)时计算公式的异同之处;二要学会将普通语言表达的随机事件翻译成数学符号(包括事件运算符号),以便运用公式进行计算.

第二章　随机变量及其分布

随机变量是概率论中最重要的概念，本章主要讨论随机变量的(概率)分布.

一、复习与考试要求

(1) 理解随机变量的概念.

(2) 理解随机变量分布函数的概念及性质，理解离散型随机变量的分布律及其性质，理解连续型随机变量的概率密度及其性质，会应用概率分布计算有关事件的概率.

(3) 掌握二项分布、泊松分布、正态分布、均匀分布和指数分布.

(4) 会求简单随机变量函数的概率分布 。

二、基本概念与理论

1. 随机变量

在随机试验中，如果存在一个变量，它依试验结果的改变而取不同的实数值，那么，称这个变量为(一维)随机变量.

在随机试验进行之前无法确定随机变量的确切取值，随机变量取各个可能实数值的概率是概率论研究的重点. 即概率论主要研究随机变量取值的统计规律. 这种统计规律称为随机变量的分布.

随机事件可以通过随机变量来表示，例如$\{a<X\leqslant b\}$，$\{a\leqslant X<b\}$，…相应的概率表示为 $P(a<X\leqslant b)$，$P(a\leqslant X<b)$，…其中$-\infty<a\leqslant b<\infty$.

2. 分布函数

随机变量的分布可以用其分布函数来表示. 随机变量 X 的分布函数规定为

$$F(x)=P(X\leqslant x),\ -\infty<x<\infty.$$

分布函数具有以下性质：

(1) $0\leqslant F(x)\leqslant 1$；

(2) $\lim\limits_{x\to-\infty}F(x)=0$，$\lim\limits_{x\to+\infty}F(x)=1$；

(3) 单调非减　当 $x_1<x_2$ 时，$F(x_1)\leqslant F(x_2)$；

(4) 右连续　$\lim\limits_{x\to x_0^+}F(x)=F(x_0)$

由已知的分布函数可以算出概率：

$$P(a<X\leqslant b)=F(b)-F(a),\quad P(X=a)=F(a)-F(a^-).$$

其中　$F(a^-)=\lim\limits_{x\to a^-}F(x)$.

3. 离散型随机变量及其分布律

如果随机变量 X 仅可能取有限个或可列无限个值，那么称 X 为离散型随机变量.

离散型随机变量 X 的分布可以用下列分布律来表示

$$P(X=a_i)=p_i,\quad i=1,2,\cdots,$$

其中　$\sum\limits_i p_i=1$，　$p_i>0$，　$i=1,2,\cdots$.

由已知的分布律可以算得概率

$$P(x\in I)=\sum_{i=a_i\in I}p_i,$$

其中，I 是实数轴上的一个集合.

离散型随机变量的分布函数是非降的阶梯函数，它在 a_i 处的跳跃度为 $p_i=F(a_i)-F(a_i^-)$，$i=1,2,\cdots$.

下面给出常用的离散型随机变量的分布.

(1) 二项分布 $B(n,p)$,它的分布律为

$$p(X=i)=\binom{n}{i}p^i(1-p)^{n-i},i=0,1,\cdots,n$$

其中,$0<p<1$. 当 $n=1$ 时,$B(1,p)$称为 0-1 分布,它的分布律简化成

$$P(X=0)=1-p,P(X=1)=p.$$

(2) 泊松分布 $P(\lambda)$,它的分布律为

$$P(X=i)=\mathrm{e}^{-\lambda}\frac{\lambda^i}{i!},\quad i=0,1,2,\cdots,$$

其中,$\lambda>0$.

(3) 均匀分布,它的分布律为

$$P(X=a_i)=\frac{1}{n},i=1,2,\cdots,n.$$

4. 连续型随机变量及其密度函数

如果随机变量 X 的分布函数 $F(x)$可以表示成

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t,-\infty<x<\infty$$

那么,称 X 为连续型随机变量,其中,函数 $f(x)$称为 X 的(概率)密度函数,它必须满足.

(1) $f(x)\geqslant 0,-\infty<x<\infty$;

(2) $\int_{-\infty}^{\infty}f(x)\mathrm{d}x=1$.

密度函数刻画了连续型随机变量的分布,因为由已知的密度函数可以算得概率:

$$P(a<X\leqslant b)=\int_a^b f(x)\mathrm{d}x.$$

连续型随机变量具有下列性质:

(1) 分布函数 $F(x)$是连续函数;

(2) 对任意一个实数 c,$P(X=c)=0$;

(3) 在 $f(x)$的连续点处,$F'(x)=f(x)$.

下面给出常用的连续型随机变量的分布:

(1) 均匀分布 $R(a,b)$,它的密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b;\\ 0, & \text{其余},\end{cases}$$

其中,$-\infty<a<b<\infty$,分布函数为

$$F(x)=\begin{cases}0, & x<a;\\ \dfrac{x-a}{b-a}, & a\leqslant x<b;\\ 1, & x\geqslant b.\end{cases}$$

(2) 指数分布 $E(\lambda)$,它的密度函数为

$$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0;\\ 0, & \text{其余},\end{cases}$$

其中,$\lambda>0$;分布函数为

$$F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & x\geqslant 0;\\ 0, & x<0.\end{cases}$$

(3) 正态分布 $N(\mu,\sigma^2)$,它的密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},-\infty<x<\infty,$$

其中，$-\infty<\mu<\infty,\sigma^2>0$，称 $N(0,1)$ 为标准正态分布，它的分布函数记作 $\Phi(x)$. $\Phi(x)$ 具有下列性质：

(1) $\Phi(0)=\frac{1}{2}$；

(2) $\Phi(-x)=1-\Phi(x)$.

如果 $X\sim N(\mu,\sigma^2)$，那么，X 的分布函数为

$$F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right),$$

$$P(a<X\leqslant b)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right).$$

5. 随机变量函数的分布

设 X 是随机变量，$g(x)$ 是一个已知函数，$Y=g(X)$ 是随机变量 X 的函数，它也是一个随机变量. 现要在已知 X 的分布的前提下求出 $Y=g(X)$ 的分布.

(1) 离散型　如果 X 的分布律为

$$P(X=a_i)=p_i,i=1,2,\cdots$$

那么，$Y=g(X)$ 的分布律为

$$P(Y=g(a_i))=p_i,i=1,2,\cdots$$

但要注意：与取相同 $g(a_i)$ 值对应的那些概率应合并相加.

(2) 连续型　要求出 $Y=g(X)$ 的密度函数可以先求其分布函数. 如果 X 的密度函数为 $f(x)$，那么，$Y=g(X)$ 的分布函数为

$$F_Y(y)=P(Y\leqslant y)=P(g(X)\leqslant y)=P(X\in I_y)=\int_{I_y}f(x)\mathrm{d}x,$$

其中，$\{X\in I_y\}$ 是与 $\{g(X)\leqslant y\}$ 相等的随机事件，而 $I_y=\{x\mid g(x)\leqslant y\}$ 是实数轴上的一个集合（通常 I_y 是一个区间或若干个区间的并集）.

定理　如果 $y=g(x)$ 是单调函数，且具有一阶连续导数，$y=g(x)$ 的反函数为 $x=h(y)$，那么，当 X 的密度函数为 $f(x)$ 时，$Y=g(X)$ 的密度函数为

$$f_Y(y)=f(h(y))\cdot|h'(y)|.$$

定理　如果 $X\sim N(\mu,\sigma^2)$，那么，当 $k\neq0$ 时，$Y=kX+b\sim N(k\mu+b,k^2\sigma^2)$. 特别地，$\frac{X-\mu}{\sigma}\sim N(0,1)$.

三、例题分析

例 2-1　设 X 与 Y 是两个随机变量，且 $P(X\geqslant0,Y\geqslant0)=\frac{3}{7}$，$P(X\geqslant0)=P(Y\geqslant0)=\frac{4}{7}$，求 $P(\max(X,Y)\geqslant0)$.

解　事件 $\{\max(X,Y)\geqslant0\}=\{X\geqslant0\}\cup\{Y\geqslant0\}$. 由加法公式

$$\begin{aligned}P\{\max(X,Y)\geqslant0\}&=P(\{X\geqslant0\}\cup\{Y\geqslant0\})=P(X\geqslant0)+P(Y\geqslant0)-P(\{X\geqslant0\}\cap\{Y\geqslant0\})\\&=\frac{4}{7}+\frac{4}{7}-\frac{3}{7}=\frac{5}{7}.\end{aligned}$$

分析　从形式上看，本题涉及二维随机向量，但实质上仍是第一章的概率计算，只是随机事件通过随机变量来表示. 本题等价于“已知 $P(A)=P(B)=\frac{4}{7}$，$P(AB)=\frac{3}{7}$，求 $P(A\cup B)$”. 本题的另一难点是分解随机事件 $\{\max(X,Y)\geqslant0\}$. 一般地，对任意一个常数 c，

$$\{\max(X,Y)\geqslant c\}=\{X\geqslant c\}\cup\{Y\geqslant c\};$$

$$\{\max(X,Y)\leqslant c\}=\{X\leqslant c\}\cap\{Y\leqslant c\};$$

$$\{\min(X,Y)\geqslant c\}=\{X\geqslant c\}\cap\{Y\geqslant c\};$$

$$\{\min(X,Y)\leqslant c\}=\{X\leqslant c\}\cup\{Y\leqslant c\}.$$

这些分解公式在第三章求 $\max(X,Y)$ 与 $\min(X,Y)$ 分布时也很有用. 本题是 1995 年试题.

例 2-2 设随机变量 X 服从正态分布 $N(\mu,\sigma^2)(\sigma>0)$ 且二次方程 $y^2+4y+X=0$ 无实根的概率为 $\frac{1}{2}$,求 μ.

解 由题意

$$P(16-4X<0)=\frac{1}{2},$$

而

$$P(16-4X<0)=P(X>4)=1-P(X\leqslant 4)=1-\Phi\left(\frac{4-\mu}{\sigma}\right)=\frac{1}{2},$$

因此

$$\Phi\left(\frac{4-\mu}{\sigma}\right)=\frac{1}{2},$$

由 $\Phi(0)=\frac{1}{2}$,知

$$\frac{4-\mu}{\sigma}=0,$$

即 $\mu=4$.

分析 本题是 2002 年试题. 初看似乎 σ 未知时无法求出 μ. 但本题巧妙地给出恰当的数据,使得问题通过使用 $\Phi(0)=0.5$ 而得到解决. 遇到这类问题,不必考虑条件是否足够充分,只要正确使用正态分布的概率计算公式,问题总会迎刃而解.

例 2-3 在下列两种情形下,求方程 $t^2+Xt+1=0$ 有实根的概率,其中,X 是随机变量.

(1) X 服从 $\{1,2,\cdots,6\}$ 上的均匀分布;

(2) X 服从区间 $[1,6]$ 上的均匀分布.

解 方程 $t^2+Xt+1=0$ 有实根的充分必要条件是 $X^2-4\geqslant 0$. 于是,所求概率可表示成 $P(X^2-4\geqslant 0)=P(X\leqslant -2)+P(X\geqslant 2)$.

(1) $P(X\leqslant -2)=0$,

$$P(X\geqslant 2)=\sum_{i=2}^{6}P(X=i)=\sum_{i=2}^{6}\frac{1}{6}=\frac{5}{6},$$

因此所求概率为 $\frac{5}{6}$;

(2) $P(X\leqslant -2)=\int_{-\infty}^{-2}f(x)\mathrm{d}x=\int_{-\infty}^{-2}0\mathrm{d}x=0$,

$$P(X\geqslant 2)=\int_{2}^{6}\frac{1}{5}\mathrm{d}x+\int_{6}^{\infty}0\mathrm{d}x=\frac{4}{5},$$

因此所求概率为 $\frac{4}{5}$.

分析 在(1)中用到了离散型均匀分布的分布律

$$P(X=i)=\frac{1}{6},i=1,2,\cdots,6;$$

在(2)中用到了连续型均匀分布的密度函数

$$f(x)=\begin{cases}\frac{1}{5}, & 1\leqslant x\leqslant 6;\\ 0, & \text{其 余}.\end{cases}$$

在概率论中,被积函数常常是分段函数,因此计算积分时要仔细.

例 2-4 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\frac{1}{3}, & \text{若 } x\in[0,1];\\ \frac{2}{9}, & \text{若 } x\in[3,6];\\ 0, & \text{其余}.\end{cases}$$

若 k 使得 $P(X\geqslant k)=\frac{2}{3}$，求 k 的取值范围.

解　由 $P(X\geqslant k)=1-P(X<k)=\frac{2}{3}$，得到

$$P(X<k)=\frac{1}{3}.$$

而 $P(X<k)=\int_{-\infty}^{k}f(x)\mathrm{d}x$，

当 $k\leqslant 0$ 时，$P(X<k)=\int_{-\infty}^{k}0\mathrm{d}x=0$；

当 $0<k\leqslant 1$ 时，$P(X<k)=\int_{0}^{k}\frac{1}{3}\mathrm{d}x=\frac{k}{3}$；

当 $1<k<3$ 时，$P(X<k)=\int_{0}^{1}\frac{1}{3}\mathrm{d}x+\int_{1}^{k}0\mathrm{d}x=\frac{1}{3}$；

当 $3\leqslant k\leqslant 6$ 时，$P(X<k)=\int_{0}^{1}\frac{1}{3}\mathrm{d}x+\int_{1}^{3}0\mathrm{d}x+\int_{3}^{k}\frac{2}{9}\mathrm{d}x=\frac{1}{3}+\frac{2}{9}(k-3)\geqslant\frac{1}{3}$，

因此，当 $k\in[1,3]$ 时，$P(X<k)=\frac{1}{3}$. 即 k 的取值范围为 $[1,3]$.

分析　本题是 2000 年试题. 实质上本题是考分布函数的计算，由于密度函数 $f(x)$ 是一个分段函数，因此，要讨论分布函数在各段上的取值.

例 2-5 选择题　设随机随机变量 X 服从指数分布，则随机变量 $Y=\min(X,2)$ 的分布函数　　(　)

(A) 是连续函数；

(B) 至少有两个间断点；

(C) 是阶梯函数；

(D) 恰好有一个间断点.

解　应选(D).

分析　这是 1999 年的试题. 本题与例 2-4 一样也是考查分布函数的计算，只是 $Y=\min(X,2)$ 是 X 的函数，由例 2-1 的分析可知：

$$\overline{\{Y\leqslant y\}}=\{Y>y\}=\{\min(X,2)>y\}=\{X>y\}\cap\{2>y\}.$$

因此，Y 的分布函数为

$$F_Y(y)=P(Y\leqslant y)=1-P(Y>y)=1-P(\{X>y\}\cap\{2>y\}).$$

当 $y\geqslant 2$ 时，

$$F_Y(y)=1-P(\phi)=1;$$

当 $y<2$ 时，由 $Y\sim E(\lambda)$ 容易算出

$$F_Y(y)=1-P(X>y)=P(X\leqslant y)=\begin{cases}1-\mathrm{e}^{-\lambda y}, & 0<y<2;\\ 0, & y\leqslant 0.\end{cases}$$

即 Y 的分布函数为

$$F_Y(y)=\begin{cases}0, & y\leqslant 0;\\ 1-\mathrm{e}^{-\lambda y}, & 0<y<2;\\ 1, & y\geqslant 2.\end{cases}$$

$y=2$ 是 $F_Y(y)$ 的惟一间断点. 故选(D).

例 2-6　已知随机变量 X 的密度函数为

$$f(x)=\frac{1}{2}\mathrm{e}^{-|x|},\ -\infty<x<\infty.$$

求 X 的分布函数 $F(x)$.

解　$F(x)=P(X\leqslant x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\int_{-\infty}^{x}\frac{1}{2}\mathrm{e}^{-|t|}\mathrm{d}t.$

当 $x<0$ 时，

$$F(x)=\int_{-\infty}^{x}\frac{1}{2}e^{t}\,dt=\frac{1}{2}e^{x};$$

当 $x\geqslant 0$ 时，

$$F(x)=\int_{-\infty}^{x}\frac{1}{2}e^{-|t|}\,dt=\int_{-\infty}^{0}\frac{1}{2}e^{t}\,dt+\int_{0}^{x}\frac{1}{2}e^{-t}\,dt$$
$$=\frac{1}{2}+\left(\frac{1}{2}-\frac{1}{2}e^{-x}\right)=1-\frac{1}{2}e^{-x}.$$

于是，

$$F(x)=\begin{cases}\dfrac{1}{2}e^{x}, & x<0;\\ 1-\dfrac{1}{2}e^{-x}, & x\geqslant 0.\end{cases}$$

分析 本题考查分布函数的定义及其计算，除了积分计算要注意绝对值符号外，下面提出三个问题.

(1) 分布函数最终要用一个分段函数来完整表达；

(2) 当 $x\geqslant 0$ 时，也可以用求逆公式来计算：

$$F(x)=P(X\leqslant x)=1-P(X>x)=1-\int_{x}^{\infty}\frac{1}{2}e^{-|t|}\,dt=1-\int_{x}^{\infty}\frac{1}{2}e^{-t}dt=1-\frac{1}{2}e^{-x};$$

(3) 题目中的 $f(x)$ 可以改成

$$\varphi(x)=ke^{-|x|},\quad -\infty<x<\infty,$$

利用 $\int_{-\infty}^{\infty}\varphi(x)\,dx=1$ 可以求得 $k=\frac{1}{2}$，因为

$$k=\frac{1}{\int_{-\infty}^{\infty}e^{-|x|}\,dx}=\frac{1}{2\int_{0}^{\infty}e^{-x}\,dx}=\frac{1}{2}.$$

例 2-7 已知 X 的分布律为

X	-1	0	1	2
P_r	0.1	0.2	0.3	0.4

求 $Y=2X^2-1$ 的分布律与分布函数.

解 X 的取值范围为 $\{-1,0,1,2\}$，$Y=2X^2-1$ 的取值范围为 $\{1,-1,7\}$，且

$$P(Y=1)=P(X=1)+P(X=-1)=0.3+0.1=0.4,$$
$$P(Y=-1)=P(X=0)=0.2,$$
$$P(Y=7)=P(X=2)=0.4.$$

因此，Y 的分布律为

Y	-1	1	7
P_r	0.2	0.4	0.4

Y 的分布函数为

$$F_Y(y)=\begin{cases}0, & y<-1;\\ 0.2, & -1\leqslant y<1;\\ 0.2+0.4, & 1\leqslant y<7;\\ 0.2+0.4+0.4, & y\geqslant 7.\end{cases}=\begin{cases}0, & y<-1;\\ 0.2, & -1\leqslant y<1;\\ 0.6, & 1\leqslant y<7;\\ 1, & y\geqslant 7.\end{cases}$$

分析 求离散型随机变量函数的分布律的关键是求出 $x=a_i$ 时 $g(x)$ 的函数值 $g(a_i)$. 在本题中，$g(x)=2x^2-1$，当 $x=-1,0,1,2$ 时，$g(-1)=g(1)=1,g(0)=-1,g(2)=7$. 需要注意的是，事件 $\{Y=1\}=\{X=-1\}\cup\{X=1\}$. 在求分布律时，$P(Y=1)=P(X=-1)+P(X=1)$. 在求分布函数值时，按自变量 y 取值从

小到大排列，相应的概率累计相加.

例 2-8　设随机变量 X 的密度函数为

$$f(x)=\frac{1}{\pi(1+x^2)},\quad -\infty<x<\infty.$$

求随机变量 $Y=1-\sqrt[3]{X}$ 的密度函数 $f_Y(y)$.

解　$y=1-\sqrt[3]{X}$ 是单调函数，且反函数为 $x=(1-y)^3$，$\dfrac{\mathrm{d}x}{\mathrm{d}y}=3(1-y)^2$. 于是

$$f_Y(y)=\frac{1}{\pi(1+(1-y)^6)}\cdot 3(1-y)^2=\frac{3}{\pi}\cdot\frac{(1-y)^2}{1+(1-y)^6},\quad -\infty<y<\infty.$$

分析　本题使用公式求解，十分方便，但在使用公式时，一定要看清条件——$y=g(x)$ 是单调函数. 下面用一般的先求分布函数的方法求解：

$$F_Y(y)=P(Y\leqslant y)=P(1-\sqrt[3]{X}\leqslant y)=P(X\geqslant(1-y)^3)=\int_{(1-y)^3}^{\infty}\frac{1}{\pi(1+x^2)}\mathrm{d}x,$$

求导得到

$$f_Y(y)=-3(1-y)^2\cdot(-1)\cdot\frac{1}{\pi(1+(1-y)^6)}=\frac{3}{\pi}\cdot\frac{(1-y)^2}{1+(1-y)^6},\quad -\infty<y<\infty.$$

例 2-9　设随机变量 $X\sim R(0,2)$，求随机变量 $Y=X^2$ 的分布函数与密度函数.

解　X 的取值范围为 $(0,2)$，$Y=X^2$ 的取值范围为 $(0,4)$. 由于 X 的密度函数为

$$f_X(x)=\begin{cases}\dfrac{1}{2}, & 0<x<2;\\ 0, & \text{其余.}\end{cases}$$

因此，当 $0<y<4$ 时，

$$F_Y(y)=P(Y\leqslant y)=P(X^2\leqslant y)=P(-\sqrt{y}\leqslant X\leqslant\sqrt{y})=\int_0^{\sqrt{y}}\frac{1}{2}\mathrm{d}x=\frac{\sqrt{y}}{2}.$$

从而 Y 的分布函数为

$$F_Y(y)=\begin{cases}0, & y<0;\\ \dfrac{\sqrt{y}}{2}, & 0\leqslant y<4;\\ 1, & y\geqslant 4.\end{cases}$$

Y 的密度函数

$$f_Y(y)=\begin{cases}\dfrac{1}{4\sqrt{y}}, & 0<y<4;\\ 0, & \text{其　余.}\end{cases}$$

分析　本题若只需求密度函数，也可以仿照上题用公式直接求解，因为 $y=x^2(0<x<2)$ 依然是一个单调函数. 它的反函数是 $x=\sqrt{y}(0<y<4)$，$\dfrac{\mathrm{d}x}{\mathrm{d}y}=\dfrac{1}{2\sqrt{y}}$，于是，当 $0<y<4$ 时，有

$$f_Y(y)=\frac{1}{2}\cdot\frac{1}{2\sqrt{y}}=\frac{1}{4\sqrt{y}}.$$

例 2-10　设随机变量 X 的密度函数为

$$f_X(x)=\begin{cases}1-|x|, & |x|<1;\\ 0, & \text{其　余.}\end{cases}$$

求随机变量 $Y=X^2+1$ 的分布函数与密度函数.

解　X 的取值范围为 $(-1,1)$，Y 的取值范围为 $[1,2)$. 当 $1\leqslant y<2$ 时，有

$$F_Y(y)=P(Y\leqslant y)=P(X^2+1\leqslant y)=P(-\sqrt{y-1}\leqslant X\leqslant\sqrt{1-y})=\int_{-\sqrt{y-1}}^{\sqrt{y-1}}(1-|x|)\mathrm{d}x$$

$$=2\int_0^{\sqrt{y-1}}(1-x)\mathrm{d}x=1-(1-\sqrt{y-1})^2.$$

从而,Y **的分布函数**

$$F_Y(y)=\begin{cases}0, & y<1;\\ 1-(1-\sqrt{y-1})^2, & 1\leqslant y<2;\\ 1, & y\geqslant 2.\end{cases}$$

Y **的密度函数为**

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{y-1}}-1, & 1<y<2;\\ 0, & \text{其 余}.\end{cases}$$

分析 本题中 $y=x^2+1(-1<x<1)$ **不是单调函数,不能直接用公式求解. 这个例子表明,先考虑随机变量的取值范围,然后在此范围内求分布函数值可以简化运算. 至于取值范围之外的分布函数值,可以由分布函数的性质决定其为** 0 **或** 1.

例 2-11 假设随机变量 X 的概率密度为

$$f(x)=\begin{cases}2x, & 0<x<1;\\ 0, & \text{其余}.\end{cases}$$

现在对 X 进行 n 次独立重复观测,以 V_n 表示观测值不大于 0.1 的次数. 试求随机变量 V_n 的概率分布.

解 $p=P\{X\leqslant 0.1\}=\int_{-\infty}^{0.1}f(x)\mathrm{d}x=\int_0^{0.1}2x\mathrm{d}x=0.01.$

V_n **服从参数为**(n,p)**的二项分布**:

$$P(V_n=k)=\binom{n}{k}p^k(1-p)^{n-k}=\binom{n}{k}\times 0.01^k\times 0.99^{n-k},k=0,1,\cdots,n.$$

分析 本题是 1994 **年的试题. 解本题要求对二项分布的直观背景有深入的了解. 这题中** X **的作用在于定义事件** $A=\{X\leqslant 0.1\}$,**这样** V_n **表示在** n **重伯努利试验中** A **出现的次数,因此,**$V_n\sim B(n,p)$,**其中** $p=P(A)=P(X\leqslant 0.1)$.

例 2-12 假设随机变量 X 的绝对值不大于 1;$P\{X=-1\}=\frac{1}{8}$, $P\{X=1\}=\frac{1}{4}$;**在事件**$\{-1<x<1\}$**出现的条件下,**X **在**$(-1,1)$**内的任一子区间上取值的条件概率与该子区间长度成正比. 试求**

(1) X **的分布函数** $F(x)=P\{X\leqslant x\}$;(2) X **取负值的概率** p.

解 (1) **由于** X **的取值范围为**$[-1,1]$,**因此,**

$x<-1$ **时,**$F(x)=0$;

$x\geqslant 1$ **时,**$F(x)=P(X\leqslant 1)=1$;

以下考虑$-1\leqslant x<1$ **时的情形.**

$x=-1$ **时,**$F(x)=P(X=-1)=\frac{1}{8}$;

由于

$$1=P(|X|\leqslant 1)=P(X=-1)+P(-1<X<1)+P(X=1)=\frac{1}{8}+P(-1<X<1)+\frac{1}{4},$$

从而

$$P(-1<X<1)=1-\frac{1}{8}-\frac{1}{4}=\frac{5}{8}。$$

由题意,在事件$\{-1<X<1\}$**出现的条件下,**X **服从区间**$(-1,1)$**上的均匀分布,即对于**$-1<x<1$,**有**

$$P\{-1<X\leqslant x\mid -1<X<1\}=\frac{1}{2}(x+1).$$

即

$$P\{-1<X\leqslant x,-1<X<1\}=P\{-1<X<1\}\frac{1}{2}(x+1),$$

从而

$$P\{-1<X\leqslant x\}=P\{-1<X<1\}\frac{1}{2}(x+1)=\frac{5}{8}\cdot\frac{1}{2}(x+1)=\frac{5}{16}(x+1).$$

$$F(x)=P(X\leqslant x)=P(X\leqslant -1)+P(-1<X\leqslant x)=\frac{1}{8}+\frac{5}{16}(x+1)=\frac{5x+7}{16}.$$

综上，有

$$F(x)=\begin{cases}0, & x<-1;\\ \dfrac{5x+7}{16}, & -1\leqslant x<1;\\ 1, & x\geqslant 1.\end{cases}$$

分析　本题是 1997 年试题. 本题中 X 既不是连续型随机变量也不是离散型随机变量. 对这类随机变量求分布函数就不能简单套用离散型随机变量和连续型随机变量的分布函数的求法，而需要从题目条件中找到突破口. 本题中“在条件$\{-1<X<1\}$下，X 服从$(-1,1)$上的均匀分布”就是一个突破口，由此可算出$-1<x<1$ 时 $P(-1<X\leqslant x)$的值.

四、小　结

本章重点要掌握两种方法：一是承接第一章，会计算用随机变量表示的随机事件的概率；二是会求随机变量函数的分布，特别是应该掌握先求分布函数的方法，这一方法在第三章中还要经常使用.

第三章　二维随机变量及其分布

在随机试验中，如果存在二维变量(X,Y)，它依试验结果的改变而取不同的向量值，那么称(X,Y)为二维随机变量(或随机向量)．本章主要讨论二维随机变量取值的统计规律，即二维随机变量的(概率)分布．

随机事件可以通过二维随机变量来表示，例如$\{(X,Y)\in D\}$，其中，D是二维平面上的某个区域，相应的概率表示为$P((X,Y)\in D)$．有时候，随机事件$\{(X,Y)\in D\}$通过一个或若干个不等式来表示．

一、复习与考试要求

(1) 了解二维随机变量的概念．

(2) 了解二维随机变量的联合分布函数及其性质、二维离散型随机变量的联合分布律及其性质和二维连续型随机变量的联合概率密度及其性质，并会用它们计算有关事件的概率．

(3) 了解二维随机变量的边缘分布和条件分布．

(4) 理解随机变量独立性的概念，掌握应用随机变量的独立性进行概率计算．

(5) 会求两个独立随机变量的简单函数的分布．

二、基本概念与理论

1. 联合分布函数与边缘分布函数

二维随机变量的分布可以用其联合分布函数来表示．随机变量X与Y的联合分布函数规定为二元函数

$$F(x,y)=P(X\leqslant x,Y\leqslant y),\quad -\infty<x,y<\infty.$$

联合分布函数具有以下性质：

(1) $0\leqslant F(x,y)\leqslant 1$；

(2) $F(-\infty,-\infty)=F(x,-\infty)=F(-\infty,y)=0$，

$F(\infty,\infty)=1$；

(3) 固定一个自变量时，对于另一个自变量单调非减；

(4) 固定一个自变量时，对于另一个自变量右连续．

由已知的联合分布函数可以算得概率：

$$P(x_1<X\leqslant x_2,\quad y_1<Y\leqslant y_2)=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1).$$

设X与Y的联合分布函数为$F(x,y)$，分别称

$$F_X(x)=P(X\leqslant x)=F(x,\infty),\quad -\infty<x<\infty,$$

$$F_Y(y)=P(Y\leqslant y)=F(\infty,y),\quad -\infty<y<\infty.$$

为X,Y的边缘分布函数．

2. 离散型二维随机变量及其分布律

如果二维随机变量(X,Y)仅可能取有限个或可列无限个值，那么称(X,Y)为离散型二维随机变量．

离散型二维随机变量(X,Y)的分布可以用下列联合分布律来表示：

$$P(X=a_i,Y=b_j)=p_{ij},\quad i,j=1,2,\cdots.$$

其中

$$\sum_i\sum_j p_{ij}=1,\quad p_{ij}\geqslant 0,\quad i,j=1,2,\cdots.$$

由已知的联合分布律可以算得概率：

$$P((X,Y)\in D)=\sum_{(i,j):(a_i,b_j)\in D}\sum p_{ij}$$

其中，D是二维平面上的一个集合

X的边缘分布律是

$$P(X=a_i)=p_{i\cdot},\quad i=1,2,\cdots.$$

其中，$0<p_{i\cdot}=\sum_j p_{ij}$，$i=1,2,\cdots$；$Y$ 的边缘分布律是

$$P(Y=b_j)=p_{\cdot j},\quad j=1,2,\cdots.$$

其中，$0<p_{\cdot j}=\sum_i p_{ij}$，$\quad j=1,2,\cdots.$

3. 连续型二维随机变量及其分布律

如果 X 与 Y 的联合分布函数 $F(x,y)$ 可以表示成

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\mathrm{d}u\mathrm{d}v,\quad -\infty<x,y<\infty,$$

那么称 (X,Y) 为连续型二维随机变量，其中，二元函数 $f(x,y)$ 称为 X 与 Y 的联合（概率）密度函数，它必须满足

(1) $f(x,y)\geqslant 0$，$-\infty<x,y<\infty$；

(2) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\mathrm{d}x\mathrm{d}y=1.$

联合密度函数刻画了连续型二维随机变量的分布，因为由已知的联合密度函数可以算得概率

$$P((X,Y)\in D)=\iint_D f(x,y)\mathrm{d}x\mathrm{d}y.$$

二维连续型随机变量具有下列性质：

(1) 对任意一条平面曲线 L，$P((X,Y)\in L)=0$；

(2) 在 $f(x,y)$ 的连续点处，$\dfrac{\partial^2}{\partial x\partial y}F(x,y)=f(x,y)$.

X 的边缘密度函数是

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\mathrm{d}y,\quad -\infty<x<\infty;$$

Y 的边缘密度函数是

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\mathrm{d}x,\quad -\infty<y<\infty.$$

下面给出常用的连续型二维随机变量的分布：

(1) 均匀分布　如果 (X,Y) 在二维平面上某个区域 G 中服从均匀分布，那么它的密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{G\text{的面积}}, & (x,y)\in G;\\ 0, & \text{其　余}.\end{cases}$$

(2) 二维正态分布 $N(\mu_1,\mu_2,\sigma_1^2\sigma_2^2,\rho)$，它的密度函数为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}\right.\right.$$

$$\left.\left.-2\rho\cdot\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\},\quad -\infty<x,y<\infty,$$

其中　$-\infty<\mu_1,\mu_2<\infty$，$\sigma_1^2$，$\sigma_2^2>0$，$|\rho|<1$.

定理　如果 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，那么 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$.

4. 随机变量的相互独立性

给定二维随机变量 (X,Y). 如果 X 与 Y 的联合分布函数等于 X，Y 的边缘分布函数之积，即

$$F(x,y)=F_X(x)F_Y(y)\text{，对一切}-\infty<x,y<\infty,$$

那么称随机变量 X 与 Y 相互独立. 随机变量 X 与 Y 相互独立的直观意义是“X 与 Y 各自的取值互不影响”，因为对实数轴上任意两个集合 I_1，I_2，总有

$$P(X\in I_1,Y\in I_2)=P(X\in I_1)P(Y\in I_2).$$

当 (X,Y) 为离散型二维随机变量时，X 与 Y 相互独立的充分必要条件为

$$p_{ij}=p_{i\cdot}\,p_{\cdot j}\text{，对一切 }i,j=1,2,\cdots.$$

当 (X,Y) 为连续型二维随机变量时，X 与 Y 相互独立的充分必要条件为

$$f(x,y)=f_X(x)f_Y(y)\text{，至少在一切连续点处成立.}$$

一般地，由 X 与 Y 的边缘分布不能惟一确定它们的联合分布. 但是，当 X 与 Y 相互独立时，由 X 与 Y 的边缘分布惟一确定了 X 与 Y 的联合分布.

定理 如果$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$,那么 X 与 Y 相互独立的充分必要条件是$\rho=0$.

定理 如果 X 与 Y 相互独立,那么 $g(X)$与 $h(Y)$相互独立.

5. 条件分布

给定二维随机变量(X,Y)当已知 Y 取值 y 时,X 仍是随机变量。这时,X 的分布会受到 $Y=y$ 的影响。X 在 $Y=y$ 的条件下的分布称为 X 的条件分布。类似地有 Y 在 $X=x$ 的条件下的条件分布.

当(X,Y)为离散型二维随机变量时,X 在给定 $Y=b_j$ 的条件下的条件分布律为

$X\mid Y=b_j$	a_1	a_2	…
P_r	$\frac{p_{1j}}{p_{\cdot j}}$	$\frac{p_{2j}}{p_{\cdot j}}$	…

$j=1,2,\cdots$;Y 在给定 $X=a_i$ 的条件下的条件分布律为

$Y\mid X=a_i$	b_1	b_2	…
P_r	$\frac{p_{i1}}{p_{i\cdot}}$	$\frac{p_{i2}}{p_{i\cdot}}$	…

当(X,Y)为连续型二维随机变量时,X 在给定 $Y=y$ 的条件下的条件(概率)密度函数为

$$f(x\mid y)=\frac{f(x,y)}{f_Y(y)},\quad -\infty<x<\infty,$$

其中,$f_Y(y)>0$;Y 在给定 $X=x$ 的条件下的条件(概率)密度函数为

$$f(y\mid x)=\frac{f(x,y)}{f_X(x)},-\infty<y<\infty,$$

其中,$f_X(x)>0$.

当 X 与 Y 相互独立时,X(或 Y)的一切条件分布都与其边缘分布相同.

6. 两个独立随机变量的简单函数的分布

设随机变量 X 与 Y 相互独立,$g(x,y)$是一个已知的简单函数,$Z=g(X,Y)$是随机变量 X 与 Y 的函数,它也是一个随机变量. 如何在已知 X,Y 的分布的前提下求出 Z 的分布?

(1) 离散型 如果 X 的分布律为

$$P(X=a_i)=p_i,\quad i=1,2,\cdots,$$

Y 的分布律为

$$P(Y=b_j)=q_j,\quad j=1,2,\cdots,$$

那么 $Z=g(X,Y)$的分布律为

$$P(Z=g(a_i,b_j))=p_iq_j,\quad i,j=1,2,\cdots.$$

但要注意,与取相同 $g(a_i,b_j)$值对应的那些概率应合并相加.

(2) 连续型 要求 $Z=g(X,Y)$的密度函数可以先求其分布函数. 如果 X 的密度函数为 $f_X(x)$,Y 的密度函数为 $f_Y(y)$,那么$Z=g(X,Y)$的分布函数为

$$F_Z(z)=P(Z\leqslant z)=P(g(X,Y)\leqslant z)=P((X,Y)\in D_z)=\iint_{D_z}f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y,$$

其中,$\{(X,Y)\in D_z\}$是与$\{g(X,Y)\leqslant z\}$相等的随机事件,而$D_z=\{(x,y)\mid g(x,y)\leqslant z\}$是二维平面上的某个集合(通常是一个区域或若干区域的并集).

定理(卷积公式) 当 X 与 Y 相互独立时,$Z=X+Y$ 的密度函数为

$$f_Z(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)\mathrm{d}x.\ (\text{或 }f_Z(z)=\int_{-\infty}^{\infty}f_X(z-y)f_Y(y)\mathrm{d}y)$$

定理(分布的可加性) 设 X 与 Y 相互独立

(1) 当 $X\sim B(m,p)$,$Y\sim B(n,p)$时,$X+Y\sim B(m+n,p)$;

(2) 当 $X\sim P(\lambda_1)$,$Y\sim P(\lambda_2)$时,$X+Y\sim P(\lambda_1+\lambda_2)$;

(3) 当 $X\sim N(\mu_1,\sigma_1^2)$,$Y\sim N(\mu_2,\sigma_2^2)$时,$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$,更一般地,

$$kX+lY+b\sim N(k\mu_1+l\mu_2+b,\ k^2\sigma_1^2+l^2\sigma_2^2),$$

其中,k,l 不同时为 0.

设 X 与 Y 相互独立，X 的分布函数为 $F_X(x)$，Y 的分布函数为 $F_Y(y)$，那么，$U=\max(X,Y)$ 的分布函数为

$$F_U(u)=F_X(u)F_Y(u),$$

$V=\min(X,Y)$ 的分布函数为

$$F_V(v)=1-(1-F_X(v))(1-F_Y(v)).$$

通过求导，可以求得 U,V 的密度函数.

三、例题分析

例 3-1　设随机变量 X 与 Y 相互独立. 在下列两种情况下，求方程 $t^2+2Xt+Y=0$ 有不同的实根、有相同实根与无实根的概率.

(1) $X\sim B(1,\frac{1}{2})$，$Y\sim B(1,\frac{1}{3})$；

(2) $X\sim R(-1,1)$，$Y\sim R(0,1)$.

解　方程 $t^2+2Xt+Y=0$ 的判别式 $\Delta=4(X^2-Y)$

(1) 所给方程有不同实根的概率为

$$P(\Delta>0)=P(X^2>Y)=P(X=1,Y=0)=P(X=1)P(Y=0)=\frac{1}{2}\times\left(1-\frac{1}{3}\right)=\frac{1}{3};$$

有相同实根的概率为

$$\begin{aligned}P(\Delta=0)&=P(X^2=Y)=P(X=0,Y=0)+P(X=1,Y=1)\\&=P(X=0)P(Y=0)+P(X=1)P(Y=1)\\&=\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)+\frac{1}{2}\times\frac{1}{3}=\frac{1}{2};\end{aligned}$$

无实根的概率为

$$P(\Delta<0)=P(X^2<Y)=P(X=0,Y=1)=P(X=0)P(Y=1)=\left(1-\frac{1}{2}\right)\times\frac{1}{3}=\frac{1}{6}.$$

(2) 所给方程有不同实根的概率为

$$P(\Delta>0)=P(X^2>Y)=\iint_{x^2>y}f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y=\int_{-1}^{1}\mathrm{d}x\int_{0}^{x^2}\frac{1}{2}\times 1\mathrm{d}y=\frac{1}{3};$$

有相同实根的概率为

$$P(\Delta=0)=P(X^2=Y)=0;$$

无实根的概率为

$$P(\Delta<0)=1-P(\Delta>0)-P(\Delta=0)=\frac{2}{3}.$$

分析　对于 $P(X^2=Y)$，离散型与连续型二维随机变量的性质是不一样的. $P(X^2=Y)=P((X,Y)\in L)$，L 是抛物线 $y=x^2$ 在 $[-1,1]$ 之间的一段曲线弧，因此当 (X,Y) 是连续型二维随机变量时，这个概率永远是 0，另外，在(2)中，运用了求逆公式以简化运算. 这个方法也可以用在(1)中. 例如，求得了 $P(\Delta>0)=\frac{1}{3}$，$P(\Delta<0)=\frac{1}{6}$ 之后，立即可推得 $P(\Delta=0)=1-\frac{1}{3}-\frac{1}{6}=\frac{1}{2}$.

例 3-2　设二维随机变量 (X,Y) 在区域 $G=\{(x,y)\mid 0<x<1,|y|<x\}$ 内服从均匀分布，求 X 与 Y 的边缘密度函数及条件密度函数.

解　G 的面积为 1，X 与 Y 的联合密度函数为

$$f(x,y)=\begin{cases}1, & 0<x<1,|y|<x;\\ 0, & \text{其　余.}\end{cases}$$

X 的取值范围为 $(0,1)$. 当 $0<x<1$ 时，

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y=\int_{-x}^{x}1\mathrm{d}y=2x.$$

因此，X 的边缘密度函数是

$$f_X(x)=\begin{cases}2x, & 0<x<1;\\ 0, & \text{其 余.}\end{cases}$$

Y 的取值范围为$(-1,1)$，当 $0<|y|<1$ 时，

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x=\int_{|y|}^{1}1\mathrm{d}x=1-|y|.$$

因此，Y 的边缘密度函数是

$$f_Y(y)=\begin{cases}1-|y|, & -1<y<1;\\ 0, & \text{其 余.}\end{cases}$$

在 $Y=y$ 的条件下，当 $y\notin(-1,1)$ 时，$f_Y(y)=0$，X 的条件分布不存在；当 $y\in(-1,1)$ 时，$f_Y(y)=1-|y|$，X 的条件密度函数为

$$f(x|y)=\begin{cases}\dfrac{1}{1-|y|}, & |y|<x<1;\\ 0, & \text{其 余.}\end{cases}$$

在 $X=x$ 的条件下，当 $x\notin(0,1)$ 时，$f_X(x)=0$，X 的条件分布不存在；当 $x\in(0,1)$ 时，$f_X(x)=2x$，Y 的条件密度函数为

$$f(y|x)=\begin{cases}\dfrac{1}{2x}, & -x<y<x;\\ 0, & \text{其 余.}\end{cases}$$

分析 求边缘密度函数的公式貌似简单，但很容易出错，因为 $f(x,y)$ 常常是二元分段函数. 本例用先求随机变量取值范围的方法. 随机变量 X(或 Y)的取值范围是区域 G 在 x 轴(或 y 轴)上的投影. 在取值范围之外，密度函数值是 0.

条件密度函数是概率论中的一个难点，但不是重点. 求条件密度函数的公式很简单，但要认清自变量，$f(x|y)$ 与 $f(y|x)$ 都不是二元函数，而是一元函数，自变量分别为 x 与 y，另一个变量在所给定的条件下要视作一个常量. 这相当于高等数学中，当二元函数一个自变量固定时，它只是另一个变量的一元函数. 另外，随着所给条件的变化，条件密度函数也不同. 例如，在本题中，$f(x|\frac{1}{2})$ 与 $f(x|\frac{1}{3})$ 便是不同的条件密度函数，因为它们所给条件分别是“$Y=\frac{1}{2}$”与“$Y=\frac{1}{3}$”. 读者不妨作为练习进行计算.

例 3-3 设随机变量 X 的密度函数为

$$f(x)=\frac{1}{2}\mathrm{e}^{-|x|},\quad -\infty<x<\infty.$$

问 X 与 $|X|$ 是否相互独立？为什么？

解 X 与 $|X|$ 不独立，其理由如下：

$$P(X>1)=\int_{1}^{\infty}\frac{1}{2}\mathrm{e}^{-|x|}\mathrm{d}x=\frac{1}{2\mathrm{e}},$$

$$P(|X|>1)=P(X>1)+P(X<-1)=\int_{1}^{\infty}\frac{1}{2}\mathrm{e}^{-|x|}\mathrm{d}x+\int_{-\infty}^{-1}\frac{1}{2}\mathrm{e}^{-|x|}\mathrm{d}x=\frac{1}{\mathrm{e}},$$

$$P(X>1,|X|>1)=P(X>1)=\frac{1}{2\mathrm{e}},$$

易见

$$P(X>1,|X|>1)\neq P(X>1)P(|X|>1).$$

分析 本题的难点在于事先不知道结论. 但从随机变量相互独立性的直观意义上去理解，X 与 $|X|$ 的取值是有关系的，因此它们一般不会相互独立. 另外，说明 X 与 $|X|$ 不相互独立容易进入一个误区，即去证

明 X 与 $|X|$ 的联合密度函数不等于 X 与 $|X|$ 的边缘密度函数之积. 在本题中,这种解题方法将带来较繁的计算. 本题是 1993 年试题,说明理由部分仅占 1 分.

例 3-4　设随机变量 X 与 Y 相互独立,$X\sim N(1,2)$, $Y\sim N(0,1)$,求 $Z=2X-Y+3$ 的密度函数.

解　由所给条件,按正态分布的可加性定理

$$Z\sim N(2\times1-1\times0+3,2^2\times2+(-1)^2\times1)=N(5,9),$$

因此,Z 的密度函数是

$$f_Z(z)=\frac{1}{3\sqrt{2\pi}}\exp\left\{-\frac{(z-5)^2}{18}\right\},\quad -\infty<z<\infty.$$

分析　本题是 1989 年试题. 它虽然是考查二维随机变量的简单函数的分布,但由于仅涉及两个正态分布,因此可以直接用可加性定理得到.

例 3-5　设随机变量 X 与 Y 相互独立,$X\sim N(\mu,\sigma^2)$,$Y\sim R(-\pi,\pi)$,试求 $Z=X+Y$ 的密度函数(计算结果用 $N(0,1)$ 的分布函数 Φ 表示,其中,$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t$).

解　X 的密度函数为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{\frac{(x-\mu)^2}{2\sigma^2}\right\},\quad -\infty<x<\infty,$$

Y 的密度函数为

$$f_Y(y)=\begin{cases}\dfrac{1}{2\pi}, & -\pi<y<\pi;\\ 0, & \text{其　余.}\end{cases}$$

按卷积公式,$Z=X+Y$ 的密度函数为

$$\begin{aligned}f_Z(z)&=\int_{-\infty}^{\infty}f_X(z-y)f_Y(y)\mathrm{d}y=\int_{-\pi}^{\pi}f_X(z-y)\frac{1}{2\pi}\mathrm{d}y\\&=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(z-y-\mu)^2}{2\sigma^2}\right\}\mathrm{d}y\\&=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{[y-(z-\mu)]^2}{2\sigma^2}\right\}\mathrm{d}y\\&=\frac{1}{2\pi}\left[\Phi\left(\frac{\pi-(z-\mu)}{\sigma}\right)-\Phi\left(\frac{-\pi-(z-\mu)}{\sigma}\right)\right],\quad -\infty<z<\infty.\end{aligned}$$

分析　上述计算最后一步中利用概率来计算积分. 当随机变量 $U\sim N(z-\mu,\sigma^2)$ 时,则

$$P(-\pi<U<\pi)=\int_{-\pi}^{\pi}f_U(u)\mathrm{d}u=\int_{-\pi}^{\pi}\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{[u-(z-\mu)]^2}{2\sigma^2}\right\}\mathrm{d}u,$$

另一方面,按照正态分布的概率计算公式,有

$$P(-\pi<U<\pi)=\Phi\left(\frac{\pi-(z-\mu)}{\sigma}\right)-\Phi\left(\frac{-\pi-(z-\mu)}{\sigma}\right).$$

使用卷积公式要仔细,因为 $f_X(x)$,$f_Y(y)$ 常常是分段函数,这样便会涉及分段函数的复合. 在上述求解过程中,没有采用另一个公式 $f_X(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)\mathrm{d}x$,就是为了避开 $f_Y(z-x)$.

例 3-6　设二维随机变量 (X,Y) 的密度函数为

$$f(x,y)=\begin{cases}2\mathrm{e}^{-(x+2y)}, & x>0,\quad y>0;\\ 0, & \text{其　余.}\end{cases}$$

求随机变量 $Z=X+2Y$ 的分布函数.

解　Z 的取值范围为 $(0,\infty)$　当 $z>0$ 时,

$$F_Z(z)=P(Z\leqslant z)=P(X+2Y\leqslant z)$$

$$=\iint_{x+2y\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^z \mathrm{d}x\int_0^{\frac{1}{2}(z-x)} 2\mathrm{e}^{-(x+2y)}\mathrm{d}y$$
$$=\int_0^z \mathrm{e}^{-x}(1-\mathrm{e}^{x-z})\mathrm{d}x=\int_0^z (\mathrm{e}^{-x}-\mathrm{e}^{-z})\mathrm{d}x$$
$$=1-\mathrm{e}^{-z}-z\mathrm{e}^{-z}.$$

于是，Z 的分布函数为

$$F_Z(z)=\begin{cases}1-\mathrm{e}^{-z}-z\mathrm{e}^{-z}, & z>0;\\ 0, & z\leqslant 0.\end{cases}$$

分析　这类题目还是要注意被积函数 $f(x,y)$ 是二元分段函数. 正确写出二次积分式是解题的关键，由 $Z=X+2Y$ 的分布函数 $F_Z(z)$，通过求导便可得到 Z 的密度函数：

$$f_Z(z)=\begin{cases}z\mathrm{e}^{-z}, & z>0;\\ 0, & \text{其　余.}\end{cases}$$

本题是 1991 年试题. 实际上本题中的 X 与 Y 是相互独立，且 $X\sim E(1)$，$Y\sim E(2)$. 读者不妨作为练习加以证明.

例 3-7　设随机变量 X 与 Y 相互独立，其密度函数分别是

$$f_X(x)=\begin{cases}1, & 0<x<1;\\ 0, & \text{其　余,}\end{cases}\qquad f_Y(y)=\begin{cases}\mathrm{e}^{-y}, & y>0;\\ 0, & \text{其　余.}\end{cases}$$

求随机变量 $Z=2X+Y$ 的密度函数.

解　X 与 Y 的联合密度函数为

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}\mathrm{e}^{-y}, & 0<x<1,y>0;\\ 0, & \text{其　余.}\end{cases}$$

Z 的取值范围为 $(0,\infty)$. 当 $z>0$ 时，有

$$F_Z(z)=P(Z\leqslant z)=P(2X+Y\leqslant z)=\iint_{2x+y\leqslant z} f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y.=\iint_{D_z}\mathrm{e}^{-y}\mathrm{d}x\mathrm{d}y,$$

其中，$D_z=\{(x,y)\mid 2x+y\leqslant z,0<x<1,y>0\}$.

当 $0<z<2$ 时，$F_Z(z)=\int_0^{\frac{z}{2}}\mathrm{d}x\int_0^{z-2x}\mathrm{e}^{-y}\mathrm{d}y=\int_0^{\frac{z}{2}}(1-\mathrm{e}^{2x-z})\mathrm{d}x=\frac{z}{2}+\frac{1}{2}\mathrm{e}^{-z}-\frac{1}{2}$，

$$f_Z(z)=\frac{1}{2}(1-\mathrm{e}^{-z});$$

当 $z>2$ 时，$F_Z(z)=\int_0^1\mathrm{d}x\int_0^{z-2x}\mathrm{e}^{-y}\mathrm{d}y=\int_0^1(1-\mathrm{e}^{2x-z})\mathrm{d}x=1-\frac{\mathrm{e}^2-1}{2}\mathrm{e}^{-z}$，

$$f_Z(z)=\frac{1}{2}(\mathrm{e}^2-1)\mathrm{e}^{-z}.$$

因此，Z 的密度函数为

$$f_Z(z)=\begin{cases}\frac{1}{2}(1-\mathrm{e}^{-z}), & 0<z<2;\\ \frac{1}{2}(\mathrm{e}^2-1)\mathrm{e}^{-z}, & z>2;\\ 0, & \text{其　余.}\end{cases}$$

分析　本题采用先求分布函数的方法. 在求解过程中，把 $z>0$ 分成两个部分：$0<z<2$ 与 $z>2$，这是因为在两种情形下积分区域 D_z 的形状不同，前者是三角形，后者是梯形. 在 $z=2$ 处不必特别关注，因为由分布函数的性质——$F_Z(z)$ 连续，立即可以得出它的值，而对密度函数而言，不妨让该处值为 0.

例 3-8　设随机变量 X 与 Y 相互独立.

(1) $X\sim B(1,\frac{1}{2})$，$Y\sim B(1,\frac{1}{2})$，**求** $U=\max(X,Y)$ **的分布律；**

(2) $X\sim E(\lambda_1)$,$Y\sim E(\lambda_2)$,**求** $V=\min(X,Y)$**的分布函数与密度函数.**

解　(1) X **与** Y **的取值范围都是**$\{0,1\}$,**因此** $U=\max(X,Y)$**的取值范围也是**$\{0,1\}$,U **的分布律为**

$$P(U=0)=P(\max(X,Y)=0)=P(X=0,Y=0)=P(X=0)P(Y=0)$$
$$=\left(1-\frac{1}{2}\right)\times\left(1-\frac{1}{2}\right)=\frac{1}{4}.$$

$$P(U=1)=1-P(U=0)=1-\frac{1}{4}=\frac{3}{4}.$$

(2) X **的分布函数为**

$$F_X(x)=\begin{cases}1-\mathrm{e}^{-\lambda_1 x}, & x>0;\\ 0, & x\leqslant 0.\end{cases}$$

Y **的分布函数为**

$$F_Y(y)=\begin{cases}1-\mathrm{e}^{-\lambda_2 y}, & y>0;\\ 0, & y\leqslant 0.\end{cases}$$

因此,$V=\min(X,Y)$**的分布函数为**

$$F_V(v)=1-(1-F_X(v))(1-F_Y(v))=\begin{cases}1-\mathrm{e}^{-(\lambda_1+\lambda_2)v}, & v>0;\\ 0, & \text{其余}.\end{cases}$$

V **的密度函数为**

$$f_V(v)=\begin{cases}(\lambda_1+\lambda_2)\mathrm{e}^{-(\lambda_1+\lambda_2)v}, & v>0;\\ 0, & \text{其余}.\end{cases}$$

分析　本题是 1994 年试题. 同样是求最大(小)值的分布,本题对离散型与连续型二维随机变量采用了不同的方法. 如果对(1)采用先求分布函数的方法,将导致易犯错误的分段函数运算. 顺便指出,$U\sim B\left(1,\frac{3}{4}\right)$,$V\sim E(\lambda_1+\lambda_2)$.

例 3-9　**设** X,Y **是相互独立且服从同一分布的两个随机变量,已知** X **的分布律为** $P(X=i)=\frac{1}{3}$,$i=1,2,3$. **令** $U=\max(X,Y)$,$V=\min(X,Y)$. **试求**:

(1) (U,V)**的分布律**;

(2) U **与** V **的边缘分布律**;

(3) **给定** $V=2$ **时** U **的条件分布律与给定** $U=3$ **时** V **的条件分布律.**

解　(1) X **与** Y **的取值范围都是**$\{1,2,3\}$,**因此**,U **与** V **的取值范围也都是**$\{1,2,3]$,(U,V)**的分布律为**

U \ V	1	2	3
1	$\frac{1}{9}$	0	0
2	$\frac{2}{9}$	$\frac{1}{9}$	0
3	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

这里,九个概率分别计算如下:

$$P(U=i,V=i)=P(X=i,Y=i)=P(X=i)P(Y=i)=\frac{1}{3}\times\frac{1}{3}=\frac{1}{9},i=1,2,3.$$

当 $i<j$ **时**,$\{U=i,V=j\}=\phi$,**因此**,$P(U=i,V=j)=0$.

当 $i>j$ **时**,$\{U=i,V=j\}=\{X=i,Y=j\}\cup\{X=j,Y=i\}$,**因此**,

$$P(U=i,V=j)=P(X=i,Y=j)+P(X=j,Y=i)=P(X=i)P(Y=j)+P(X=j)P(Y=i)$$
$$=\frac{1}{3}\times\frac{1}{3}+\frac{1}{3}\times\frac{1}{3}=\frac{2}{9}.$$

(2) U 的边缘分布律为

U	1	2	3
P_r	$\frac{1}{9}+0+0=\frac{1}{9}$	$\frac{2}{9}+\frac{1}{9}+0=\frac{3}{9}$	$\frac{2}{9}+\frac{2}{9}+\frac{1}{9}=\frac{5}{9}$

V 的边缘分布律为

V	1	2	3
P_r	$\frac{1}{9}+\frac{2}{9}+\frac{2}{9}=\frac{5}{9}$	$0+\frac{1}{9}+\frac{2}{9}=\frac{3}{9}$	$0+0+\frac{1}{9}=\frac{1}{9}$

(3) 给定 $V=2$ 时 U 的条件分布律为

$U\|V=2$	1	2	3
P_r	$\frac{0}{3/9}=0$	$\frac{1/9}{3/9}=\frac{1}{3}$	$\frac{2/9}{3/9}=\frac{2}{3}$

给定 $U=3$ 时 V 的条件分布律为

$V\|U=3$	1	2	3
P_r	$\frac{2/9}{5/9}=\frac{2}{5}$	$\frac{2/9}{5/9}=\frac{2}{5}$	$\frac{1/9}{5/9}=\frac{1}{5}$

分析 本题(1)是 1996 年试题，解决本题所用的方法仅仅是基本要求的综合运用. 望读者仔细体会。在(2)中，由(1)得到的联合分布律推出边缘分布律，也可以仿照例 3-8(1)，直接由 (X,Y) 的分布律推出 U，V 的分布律.

例 3-10(选择题) 设 X_1 和 X_2 是任意两个相互独立的连续型随机变量，它们的概率密度分别为 $f_1(x)$ 和 $f_2(x)$，分布函数分别为 $F_1(x)$ 和 $F_2(x)$，则

(A) $f_1(x)+f_2(x)$ 必为某一随机变量的概率密度；

(B) $f_1(x)f_2(x)$ 必为某一随机变量的概率密度；

(C) $F_1(x)+F_2(x)$ 必为某一随机变量的分布函数；

(D) $F_1(x)F_2(x)$ 必为某一随机变量的分布函数.

解 应选(D).

分析 本题是 2002 年试题. 解本题要求对分布函数和密度函数的性质十分熟悉. 由于 $\int_{-\infty}^{\infty}(f_1(x)+f_2(x))\mathrm{d}x=\int_{-\infty}^{\infty}f_1(x)\mathrm{d}x+\int_{-\infty}^{\infty}f_2(x)\mathrm{d}x=2$，因此，(A)不成立. 至于(B)，若取 $X_1\sim R(0,1)$，$X_2\sim R(1,2)$ 时，$f_1(x)f_2(x)\equiv 0$，因此 $\int_{-\infty}^{\infty}f_1(x)f_2(x)\mathrm{d}x=0$. 从而(B)不成立. 由于 $F_1(+\infty)=1$，$F_2(+\infty)=1$，故 $F_1(+\infty)+F_2(+\infty)=2$，因而(C)不成立，(D)是正确的，实际上 $F_1(x)F_2(x)$ 是 $\max(X_1,X_2)$ 的分布函数.

例 3-11 设随机变量 X 与 Y 相互独立，下面列出了二维随机变量 (X,Y) 的联合分布律及关于 X 和关于 Y 的边缘分布律中的部分数值，试将其余数值填入表中的空白处.

X \ Y	y_1	y_2	y_3	$P\{X=x_i\}=p_{i\cdot}$
x_1		$\frac{1}{8}$		
x_2	$\frac{1}{8}$			
$P\{Y=y_j\}=p_{\cdot j}$	$\frac{1}{6}$			1

解

X \ Y	y_1	y_2	y_3	$P\{X=x_i\}=p_{i\cdot}$
x_1	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{4}$
x_2	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{4}$
$P\{Y=y_j\}=p_{\cdot j}$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$	1

分析　本题是1999年试题.解本题要求对联合分布律和边缘分布律的概念十分熟悉. 先由 $P\{X=x_1,Y=y_1\}+\frac{1}{8}=\frac{1}{6}$,解出 $P\{X=x_1,Y=y_1\}=\frac{1}{6}-\frac{1}{8}=\frac{1}{24}$. 由 X,Y 相互独立得到 $P\{X=x_1,Y=y\}=p_{11}=p_1.p_{\cdot 1}$,因此

$$P\{X=x_1\}=p_1.=\frac{p_{11}}{p_{\cdot 1}}=\frac{\frac{1}{24}}{\frac{1}{6}}=\frac{1}{4}.$$

同样由独立性可得　$P\{X=x_2\}=p_2.=\frac{p_{21}}{p_{\cdot 1}}=\frac{\frac{1}{8}}{\frac{1}{6}}=\frac{3}{4}.$

$$P\{Y=y_2\}=p_{\cdot 2}=\frac{p_{12}}{p_1.}=\frac{\frac{1}{8}}{\frac{1}{4}}=\frac{1}{2}.$$

$$P\{Y=y_3\}=1-P\{Y=y_1\}-P\{Y=y_2\}=1-\frac{1}{6}-\frac{1}{2}=\frac{1}{3}.$$

$$P\{X=x_2,Y=y_2\}=p_{22}=p_2.\,p_{\cdot 2}=\frac{3}{4}\times\frac{1}{2}=\frac{3}{8},$$

$$P\{X=x_2,Y=y_3\}=p_{23}=p_2.\,p_{\cdot 3}=\frac{3}{4}\times\frac{1}{3}=\frac{1}{4}.$$

例 3-12　设随机变量 $X_i(i=1,2)$的分布律为

X_i	-1	0	1
概率	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

且满足 $P\{X_1X_2=0\}=1$,求 $P\{X_1=X_2\}$.

解　设(X_1,X_2)的联合分布律为

X_1 \ X_2	-1	0	1
-1	p_{11}	p_{12}	p_{13}
0	p_{21}	p_{22}	p_{23}
1	p_{31}	p_{32}	p_{33}

由于 $P\{X_1X_2=0\}=1$,因此

$$P\{X_1X_2=0\}=p_{12}+p_{22}+p_{32}+p_{21}+p_{23}=1$$

另一方面　$\sum\limits_{i=1}^{3}\sum\limits_{j=1}^{3}p_{ij}=1$. 因此

$$p_{11}=p_{13}=p_{31}=p_{33}=0.$$

即(X_1,X_2)的联合分布律为

X_1 \ X_2	-1	0	1
-1	0	p_{12}	0
0	p_{21}	p_{22}	p_{23}
1	0	p_{32}	0

由于已知 $X_i(i=1,2)$的边缘分布律，因此，有

$$p_{21}=P\{X_2=-1\}=\frac{1}{4},\ p_{23}=P\{X_2=1\}=\frac{1}{4};$$

$$p_{12}=P\{X_1=-1\}=\frac{1}{4},\ p_{32}=P\{X_1=1\}=\frac{1}{4}.$$

因而

$$p_{22}=1-p_{12}-p_{21}-p_{23}-p_{32}=1-\frac{1}{4}-\frac{1}{4}-\frac{1}{4}-\frac{1}{4}=0.$$

从而

$$P\{X_1=X_2\}=p_{11}+p_{22}+p_{33}=0+0+0=0.$$

分析 本题是 1999 年的考题. 本题中 X_1,X_2 并不相互独立，此时，一般的教材都强调由边缘分布律得不出联合分布律，但适当补充条件后，可以由边缘分布律推出联合分布律. 本例中通过添加条件 $P\{X_1X_2=0\}=1$，使得 $p_{11}=p_{13}=p_{31}=p_{33}=0$，从而使得未知量个数由 9 个减少至 5 个. 再结合边缘分布律求出联合分布律.

四、小 结

本章重点是求分布——联合分布、边缘分布(条件分布)与二维随机变量简单函数的分布，即使是讨论独立性，也离不开关于分布的讨论. 当然，求分布还是建立在求概率的基础上. 另外，对正态分布要给予足够的重视，因为它具有一些特殊的性质，运用这些性质可以简化计算.

第四章　随机变量的数字特征

随机变量的分布一般不易得到，退而求其次，本章讨论随机变量的数字特征.

一、复习与考试要求

(1) 理解数学期望和方差的概念，掌握它们的性质与计算.

(2) 掌握二项分布、泊松分布和正态分布的数学期望和方差，了解均匀分布和指数分布的数学期望和方差.

(3) 会计算随机变量函数的数学期望.

(4) 了解矩、协方差和相关系数的概念和性质，并会计算.

二、基本概念与理论

1. 数学期望

设 X 是离散型随机变量，X 的分布律为

$$P(X=a_i)=p_i,\quad i=1,2,\cdots,$$

如果级数 $\sum_i a_i p_i$ 绝对收敛①，那么，规定 X 的数学期望(或均值)为

$$E(X)=\sum_i a_i p_i;$$

规定随机变量函数 $g(X)$ 的数学期望为

$$E[g(X)]=\sum_i g(a_i)p_i.$$

设 X 是连续型随机变量，X 的密度函数为 $f(x)$. 规定 X 的数学期望为

$$E(X)=\int_{-\infty}^{\infty} xf(x)\mathrm{d}x;$$

规定随机变量函数 $g(X)$ 的数学期望为

$$E[g(X)]=\int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x.$$

设 (X,Y) 为二维离散型随机变量，X 与 Y 的联合分布律为

$$P(X=a_i,Y=b_j)=p_{ij},\quad i,j=1,2,\cdots.$$

规定二维随机变量的函数 $g(X,Y)$ 的数学期望为

$$E[g(X,Y)]=\sum_i\sum_j g(a_i,b_j)p_{ij}.$$

特殊地

$$E(X)=\sum_i\sum_j a_i\ p_{ij}=\sum_i a_i\ p_{i\cdot},\ E(Y)=\sum_i\sum_j b_i\ p_{ij}=\sum_i b_i\ p_{\cdot j}.$$

设 (X,Y) 为二维连续型随机变量，X 与 Y 的联合密度函数为 $f(x,y)$，规定二维随机变量的函数 $g(X,Y)$ 的数学期望为

① 如果级数不绝对收敛，那么就认为 X 的数学期望不存在. 以下类似的情形(包括连续型情形下的广义积分)不再一一指明.

$$E[g(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y.$$

特殊地

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x\,f(x,y)\mathrm{d}x\mathrm{d}y=\int_{-\infty}^{\infty}x\,f_X(x)\mathrm{d}x,$$

$$E(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}y\,f(x,y)\mathrm{d}x\mathrm{d}y=\int_{-\infty}^{\infty}y\,f_Y(y)\mathrm{d}y.$$

数学期望具有下列性质：

(1) $E(c)=c$ （c 是常数）；

(2) $E(kX+b)=kE(X)+b$（k,b 是常数）；

(3) $E(X\pm Y)=E(X)\pm E(Y)$；

(4) 如果 X 与 Y 相互独立，那么 $E(XY)=E(X)E(Y)$.

2. 方差与标准差

随机变量 X 的函数 $g(X)=[X-E(X)]^2$ 的数学期望称为随机变量 X 的方差，记作 $D(X)$，即

$$D(X)=E([X-E(X)]^2).$$

计算方差常用下列公式：

$$D(X)=E(X^2)-[E(X)]^2.$$

方差具有下列性质：

(1) $D(c)=0$（c 是常数）；

(2) $D(kX+b)=k^2D(X)$（k,b 是常数）；

(3) 如果 X 与 Y 相互独立，那么，$D(X\pm Y)=D(X)+D(Y)$.

方差 $D(X)$ 的算术平方根 $\sqrt{D(X)}$ 称为随机变量 X 的标准差.

3. 协方差

设 (X,Y) 为二维随机变量，称

$$\mathrm{cov}(X,Y)=E([X-E(X)][Y-E(Y)])$$

为随机变量 X 与 Y 的协方差. 计算协方差常用下列公式：

$$\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y).$$

当 $X=Y$ 时，$\mathrm{cov}(X,Y)=\mathrm{cov}(X,X)$ 表示 $D(X)$.

协方差具有下列性质：

(1) $\mathrm{cov}(X,c)=0$（c 是常数）；

(2) $\mathrm{cov}(kX,lY)=kl\mathrm{cov}(X,Y)$（$k,l$ 是常数）；

(3) $\mathrm{cov}(X,Y\pm Z)=\mathrm{cov}(X,Y)\pm\mathrm{cov}(X,Z)$.

4. 相关系数

设 (X,Y) 为二维随机变量，称

$$\rho(X,Y)=\frac{\mathrm{cov}(X,Y)}{\sqrt{D(X)D(Y)}}$$

为随机变量 X 与 Y 的相关系数. 相关系数 $\rho(X,Y)$ 反映了随机变量 X 与 Y 之间线性联系的紧密程度. 当 $\rho(X,Y)=0$ 时，称 X 与 Y 不相关.

相关系数具有下列性质：

(1) $|\rho(X,Y)|\leqslant 1$；

(2) $|\rho(X,Y)|=1$ 的充分必要条件是 X 与 Y 之间存在线性关系的概率为 1.

定理 下列 5 个命题是等价的：

(1) $\rho(X,Y)=0$；

(2) $\mathrm{cov}(X,Y)=0$；

(3) $E(XY)=E(X)E(Y)$；

(4) $D(X+Y)=D(X)+D(Y)$；

(5) $D(X-Y)=D(X)+D(Y)$.

定理　如果随机变量 X 与 Y 相互独立，那么 X 与 Y 不相关.

定理　如果 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，那么 X 与 Y 相互独立等价于 X 与 Y 不相关.

利用协方差或相关系数可以计算

$$D(X\pm Y)=D(X)+D(Y)\pm 2\mathrm{cov}(X,Y)$$
$$=D(X)+D(Y)\pm 2\rho(X,Y)\sqrt{D(X)D(Y)}.$$

5. 常用分布的数字特征

(1) 当 $X\sim B(n,p)$ 时，

$E(X)=np$，　　$D(X)=np(1-p)$.

(2) 当 $X\sim P(\lambda)$ 时，

$E(X)=\lambda$，　　$D(X)=\lambda$.

(3) 当 $X\sim R(a,b)$ 时，

$E(X)=\dfrac{1}{2}(a+b)$，　　$D(X)=\dfrac{1}{12}(b-a)^2$.

(4) 当 $X\sim E(\lambda)$ 时，

$E(X)=\dfrac{1}{\lambda}$，　　$D(X)=\dfrac{1}{\lambda^2}$.

(5) 当 $X\sim N(\mu,\sigma^2)$ 时，

$E(X)=\mu$，　　$D(X)=\sigma^2$.

(6) 当 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ 时，

$E(X)=\mu_1$，　　$D(X)=\sigma_1^2$.

$E(Y)=\mu_2$，　　$D(Y)=\sigma_2^2$.

$\mathrm{cov}(X,Y)=\rho\sigma_1\sigma_2$，　　$\rho(X,Y)=\rho$.

6. 矩

设 X 是随机变量，称 $E(X^k)$ 为随机变量 X 的 k 阶原点矩；称 $E([X-E(X)]^k)$ 为随机变量 X 的 k 阶中心矩.

设 (X,Y) 是二维随机变量，称 $E(X^kY^l)$ 为随机变量 X 与 Y 的 (k,l) 阶联合原点矩；称 $E([X-E(X)]^k[Y-E(Y)]^l)$ 为随机变量 X 与 Y 的 (k,l) 阶联合中心矩.

数学期望是一阶原点矩；方差是二阶中心矩；协方差是 $(1,1)$ 阶联合中心矩.

三、例题分析

例 4-1　已知 $X\sim P(2)$，即

$$P(X=k)=\mathrm{e}^{-2}\frac{2^k}{k!},\quad k=0,1,2,\cdots.$$

求随机变量 $Z=3X-2$ 的数学期望 $E(Z)$.

解　由 $X\sim P(2)$ 推知 $E(X)=2$. 于是，

$$E(Z)=E(3X-2)=3E(X)-2=3\times 2-2=4.$$

分析　如果不知道泊松分布的数学期望，那就只好先求级数的和：

$$E(X)=\mathrm{e}^{-2}\sum_{k=0}^{\infty}k\cdot\frac{2^k}{k!}=\mathrm{e}^{-2}\sum_{k=1}^{\infty}\frac{2^k}{(k-1)!}=2\mathrm{e}^{-2}\sum_{k-1=0}^{\infty}\frac{2^{k-1}}{(k-1)!}=2.$$

例 4-2　已知 $X\sim E(1)$，求 $E(X+\mathrm{e}^{-2X})$

解　由 $X\sim E(1)$ 推知 $E(X)=1$，且

$$E(\mathrm{e}^{-2X})=\int_0^{\infty}\mathrm{e}^{-2x}\mathrm{e}^{-x}\mathrm{d}x=\int_0^{\infty}\mathrm{e}^{-3x}\mathrm{d}x=\frac{1}{3}.$$

因此，

$$E(X+e^{-2X})=E(X)+E(e^{-2X})=1+\frac{1}{3}=\frac{4}{3}.$$

分析 本题切忌错用公式

$$E(X+e^{-2X})=E(X)+e^{-2E(X)}.$$

因为 e^{-2X} 不是 X 的线性函数，所以上式不成立.

例 4-3 已知连续型随机变量 X 的密度函数为

$$f(x)=\frac{1}{\sqrt{\pi}}e^{-x^2+2x-1},\quad -\infty<x<\infty,$$

求 $E(X)$ 与 $D(X)$.

解 由于

$$f(x)=\frac{1}{\sqrt{2\pi}\sqrt{\frac{1}{2}}}\exp\left\{-\frac{(x-1)^2}{2\times\frac{1}{2}}\right\}$$

是 $N(1,\frac{1}{2})$ 的密度函数，因此

$$E(X)=1,\qquad D(X)=\frac{1}{2}.$$

分析 本题是 1987 年试题. 解决本题的关键是要善于识别常用分布的密度函数. 不然的话，直接计算将会带来较大的工作量. 本题中的 $f(x)$ 也可以改成

$$\varphi(x)=ke^{-x^2+2x},\quad -\infty<x<\infty.$$

最终结果不变. 理由请参阅例 2-3 的分析.

例 4-4 设 X 表示 10 次独立重复射击命中目标的次数，每次射中目标的概率为 0.4，求 $E(X^2)$.

解 由题意得到 $X\sim B(10,0.4)$，于是

$$E(X)=10\times0.4=4,$$

$$D(X)=10\times0.4\times(1-0.4)=2.4.$$

由 $D(X)=E(X^2)-[E(X)]^2$

推得 $E(X^2)=D(X)+[E(X)]^2=2.4+4^2=18.4$

分析 本题是 1995 年试题，它考查了两个内容. 一是由题意归结出随机变量 X 的分布，二是灵活应用方差计算公式，如果直接求解，那么

$$E(X^2)=\sum_{k=0}^{10}k^2\binom{10}{k}0.4^k(1-0.4)^{10-k}$$

的计算是较繁的.

例 4-5 设二维随机变量 (X,Y) 在区域 $G=\{(x,y)\}|0<x<1,|y|<x\}$ 内服从均匀分布，求随机变量 $Z=2X+1$ 的方差 $D(Z)$.

解 由方差的性质得到

$$D(Z)=D(2X+1)=4D(X).$$

在例 3-2 中已经求得 X 的边缘密度函数为

$$f_X(x)=\begin{cases}2x, & 0<x<1;\\ 0, & \text{其余}.\end{cases}$$

于是，

$$E(X)=\int_0^1 x\times 2x\mathrm{d}x=\frac{2}{3},$$

$$E(X^2)=\int_0^1 x^2\times 2x\mathrm{d}x=\frac{1}{2},$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{1}{2}-\left(\frac{2}{3}\right)^2=\frac{1}{18}.$$

因此，$D(Z)=4\times\frac{1}{18}=\frac{2}{9}$.

分析　尽管本题给出的是二维随机变量，但在求 X 的数学期望与方差时可以从 X 的边缘密度函数出发，而不必从 X 与 Y 的联合密度函数开始. 一般情形下，采用边缘密度函数较方便.

例 4-6　已知 $X\sim N(1,3^2)$，$Y\sim N(0,4^2)$，$\rho(X,Y)=-\frac{1}{2}$. 设随机变量 $Z=\frac{X}{3}+\frac{Y}{2}$.

(1) 求 $E(Z)$ 与 $D(Z)$；

(2) 求 X 与 Z 的相关系数 $\rho(X,Z)$.

解　(1) 由数学期望与方差的性质得到

$$E(Z)=E\left(\frac{X}{3}+\frac{Y}{2}\right)=\frac{1}{3}E(X)+\frac{1}{2}E(Y)=\frac{1}{3}\times 1+\frac{1}{2}\times 0=\frac{1}{3},$$

$$D(Z)=D\left(\frac{X}{3}+\frac{Y}{2}\right)=D\left(\frac{X}{3}\right)+D\left(\frac{Y}{2}\right)+2\mathrm{cov}\left(\frac{X}{3},\frac{Y}{2}\right)$$

$$=\frac{1}{9}D(X)+\frac{1}{4}D(Y)+2\times\frac{1}{3}\times\frac{1}{2}\mathrm{cov}(X,Y)$$

$$=\frac{1}{9}\times 3^2+\frac{1}{4}\times 4^2+\frac{1}{3}\rho(X,Y)\sqrt{D(X)D(Y)}$$

$$=1+4+\frac{1}{3}\times\left(-\frac{1}{2}\right)\times 3\times 4=3.$$

(2) 由于

$$\mathrm{cov}(X,Z)=\mathrm{cov}(X,\frac{X}{3}+\frac{Y}{2})=\frac{1}{3}\mathrm{cov}(X,X)+\frac{1}{2}\mathrm{cov}(X,Y)$$

$$=\frac{1}{3}D(X)+\frac{1}{2}\rho(X,Y)\sqrt{D(X)D(Y)}$$

$$=\frac{1}{3}\times 3^2+\frac{1}{2}\times(-\frac{1}{2})\times 3\times 4=0.$$

因此 $\rho(X,Z)=0$.

分析　本题是 1994 年试题. 主要考查方差、协方差与相关系数的性质. 即使 $X\sim N(1,3^2)$，$Y\sim N(0,4^2)$，$\rho(X,Y)=0$，但不能由此推出 X 与 Y 相互独立，因为 (X,Z) 未必一定服从二维正态分布. 这方面内容的深入讨论超出了考试大纲与教学大纲的要求. 当然，如果添上条件“已知 (X,Y) 服从二维正态分布”，那么由 $\rho(X,Y)=0$，立刻推得 X 与 Z 相互独立，因为这时 (X,Z) 必定服从二维正态分布.

例 4-7　设随机变量 X 的密度函数为

$$f(x)=\frac{1}{2}\mathrm{e}^{-|x|},\quad -\infty<x<\infty.$$

(1) 求 $E(X)$ 与 $D(X)$；

(2) 求 X 与 $|X|$ 的协方差 $\mathrm{cov}(X,|X|)$，并问 X 与 $|X|$ 是否不相关？

解　(1) 广义积分 $\int_{-\infty}^{\infty}x\cdot\frac{1}{2}\mathrm{e}^{-|x|}\mathrm{d}x$ 绝对收敛，因此

$$E(X)=\int_{-\infty}^{\infty}x\cdot\frac{1}{2}\mathrm{e}^{-|x|}\mathrm{d}x=0,$$

$$E(X^2)=\int_{-\infty}^{\infty}x^2\cdot\frac{1}{2}e^{-|x|}dx=\int_0^{\infty}x^2e^{-|x|}dx=\int_0^{\infty}x^2e^{-x}dx=2.$$

$$D(X)=E(X^2)-[E(X)]^2=2;$$

(2) **应用协方差计算公式**

$$\operatorname{cov}(X,|X|)=E(X\cdot|X|)-E(X)\cdot E(|X|),$$

由于随机变量 X **的函数** $g(X)=X\cdot|X|$ **的数学期望**

$$E(X\cdot|X|)=\int_{-\infty}^{\infty}x\cdot|x|\times\frac{1}{2}e^{-|x|}dx=0,$$

因此

$$\operatorname{cov}(X,|X|)=0-0\times E(|X|)=0.$$

因此推得 $\rho(X,|X|)=0$,**即** X **与** $|X|$ **不相关.**

分析　在例 3-3 **中曾经说明了** X **与** $|X|$ **不相互独立. 在本题计算积分时,多次用到奇(偶)函数在对称区间上积分的性质. 但要注意,现在是广义积分,遇到奇函数时特别要当心. 严格地说,在**(2)**中计算** $\int_{-\infty}^{\infty}x\cdot|x|\times\frac{1}{2}e^{-|x|}dx$ **时也应先说明它绝对收敛.**

例 4-8　一零件的横截面是圆,对截面直径进行测量,设其直径服从区间$[0,2]$**上的均匀分布,求截面面积的数学期望与方差.**

解　设截面直径为 X,$X\sim R(0,2)$,**即** X **的密度函数为**

$$f_X(x)=\begin{cases}\frac{1}{2}, & 0\leqslant x\leqslant 2;\\ 0, & \text{其　余.}\end{cases}$$

设横截面面积为 Y,$Y=\frac{\pi}{4}X^2$. **易见,**

$$E(Y)=E\left(\frac{\pi}{4}X^2\right)=\int_{-\infty}^{\infty}\frac{\pi}{4}x^2f_X(x)dx=\int_0^2\frac{\pi}{4}x^2\times\frac{1}{2}dx=\frac{\pi}{3}.$$

另外,

$$E(Y^2)=E\left(\frac{\pi^2}{16}X^4\right)=\int_{-\infty}^{\infty}\frac{\pi^2}{16}x^4f_X(x)dx=\int_0^2\frac{\pi^2}{16}x^4\times\frac{1}{2}dx=\frac{\pi^2}{5},$$

因此

$$D(Y)=E(Y^2)-[E(Y)]^2=\frac{\pi^2}{5}-\left(\frac{\pi}{3}\right)^2=\frac{4}{45}\pi^2.$$

分析　本题以应用题的形式出现,要求首先建立随机变量之间的函数关系. 此题本质上提出了如何求随机变量的函数$Y=g(X)$**的方差?**

由上述求解过程可以看出,一般地,有

$$D[g(X)]=D(Y)=E(Y^2)-[E(Y)]^2=E[g^2(X)]-(E[g(X)])^2.$$

本题是 1987 **年备用试题.**

例 4-9　设 X **与** Y **是相互独立且均服从** $N\left(0,\frac{1}{2}\right)$ **的随机变量,求**(1)$E(|X-Y|)$,(2) $D(|X-Y|)$.

解　记 $U=X-Y$. **由所给条件,按正态分布的可加性定理,有**

$$U\sim N\left(0-0,\frac{1}{2}+(-1)^2\times\frac{1}{2}\right),$$

即

$$U\sim N(0,1).$$

于是

$$E(|X-Y|)=E(|U|)=\int_{-\infty}^{\infty}|u|\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du=2\int_0^{\infty}u\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du$$

$$=\sqrt{\frac{2}{\pi}}\int_0^{\infty} u\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u=\sqrt{\frac{2}{\pi}}$$

$$E(|X-Y|^2)=EU^2=D(U)+(EU)^2=1,$$

因此

$$D(|X-Y|)=E(|X-Y|^2)-(E(|X-Y|))^2=1-\left(\sqrt{\frac{2}{\pi}}\right)^2=1-\frac{2}{\pi}.$$

分析　本题(1)是 1996 年试题. 本题(2)是 1998 年试题. 本题所用的技巧在例 3-4 中已经有所体现，关键是引进一个中间随机变量 U，一般它是独立正态随机变量的线性函数. 如果直接用求 $E(g(X,Y))$ 的公式来求解本题，计算便会相当繁琐.

例 4-10　设随机变量 X 的密度为

$$f(x)=\begin{cases}\frac{1}{2}\cos\left(\frac{x}{2}\right), & 0\leqslant x\leqslant\pi;\\ 0, & \text{其余}.\end{cases}$$

对 X 独立地重复观察 4 次，用 Y 表示观察值大于 $\frac{\pi}{3}$ 的次数，求 Y^2 的数学期望.

解　由于

$$P\left(X>\frac{\pi}{3}\right)=\int_{\frac{\pi}{3}}^{\pi}\frac{1}{2}\cos\left(\frac{x}{2}\right)\mathrm{d}x=\frac{1}{2},$$

因此

$$Y\sim B\left(4,\frac{1}{2}\right)$$

因此，$EY=4\times\frac{1}{2}=2$，$D(Y)=4\times\frac{1}{2}\times\left(1-\frac{1}{2}\right)=1$，于是

$$E(Y^2)=D(Y)+(E(Y))^2=1+2^2=5.$$

分析　本题是 2002 年试题. 利用二项概率使其成为求常用分布的数学特征问题. 这种综合运用概率知识的能力是考生必须具备的.

例 4-11　设随机变量 X 服从参数为 λ 的泊松(Poisson)分布，且已知 $E[(X-1)(X-2)]=1$，求 λ.

解　由于 X 服从 $P(\lambda)$，因此

$$EX=D(X)=\lambda,E(X^2)=D(X)+(EX)^2=\lambda+\lambda^2.$$

从而

$$E[(X-1)(X-2)]=E[X^2-3X+2]=EX^2-3EX+2=\lambda+\lambda^2-3\lambda+2=1,$$

即

$$(\lambda-1)^2=\lambda^2-2\lambda+1=0.$$

由此得

$$\lambda=1.$$

分析　本题是 1999 年试题. 解本题的关键是理解数字特征 $E[(X-1)(X-2)]$ 是参数 λ 的函数，可记这个函数为 $h(\lambda)$. 由方程 $h(\lambda)=1$ 解出 λ. 一般地，若随机变量 X 的密度函数(或分布律)为 $f(x;\theta)$，其中，θ 为参数，那么，由条件 $Eg(X)=0$ 可建立方程

$$u(\theta)\equiv Eg(x)=0,$$

即

$$u(\theta)=0.$$

解此方程可得到 θ.

本题中 $g(x)=(X-1)(X-2)-1$，$\theta=\lambda$，$u(\lambda)=\lambda^2-2\lambda+1$，由 $u(\lambda)=0$ 解出 $\lambda=1$.

例 4-12　一商店经销某种商品，每周进货的数量 X 与顾客对该种商品的需求量 Y 是相互独立的随机变量，且都服从区间[10,20]上的均匀分布. 商店每售出一单位商品可得利润 1000 元；若需求量超过了进货量，

商店可从其他商店调剂供应，这时每单位商品获利润为 500 元，试计算此商店经销该种商品每周所得到利润的期望值.

解 记 Z 表示商店一周所得的利润(单位：元). 由题意

$$Z=\begin{cases}1000Y, & Y\leqslant X\\ 1000X+500(Y-X), & Y>X\end{cases}$$

由于 X 与 Y 相互独立，所以 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{100}, & 10\leqslant x\leqslant 20 \text{ 且 } 10\leqslant y\leqslant 20;\\ 0, & \text{其余.}\end{cases}$$

由 Z 的表达式可知，Z 是 X,Y 的函数，即 $Z=g(X,Y)$. 从而，

$$\begin{aligned}EZ=Eg(X,Y)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y\\ &=\iint\limits_{\{y\leqslant x\}}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y+\iint\limits_{\{y>x\}}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y\\ &=\iint\limits_{\{y\leqslant x\}}1000y\,f(x,y)\mathrm{d}x\mathrm{d}y+\iint\limits_{\{y>x\}}(500(y-x)+1000x)f(x,y)\mathrm{d}x\mathrm{d}y\\ &=\int_{10}^{20}\mathrm{d}y\int_{y}^{20}1000y\cdot\frac{1}{100}\mathrm{d}x+\int_{10}^{20}\mathrm{d}y\int_{10}^{y}500(x+y)\frac{1}{100}\mathrm{d}x\\ &=10\int_{10}^{20}y(20-y)\mathrm{d}y+5\int_{10}^{20}\left(\frac{3}{2}y^2-10y-50\right)\mathrm{d}y\\ &=\frac{20000}{3}+5\times 1500\approx 14166.7(\text{元}).\end{aligned}$$

分析 本题是 1998 年试题. 随机变量的数字特征在经济学和管理科学中有许多应用. 如本例中的利润的期望值(也称为期望利润)，投资学中的平均收益用收益的期望表示，投资的风险则用方差来度量等. 本例中一周的利润是一个随机变量 Z，Z 可用进货量 X 和需求量 Y 表示，只是这种表示式是一个分段的函数，这样求 EZ 就转化为对 X,Y 的函数 $g(X,Y)$ 求期望. $Eg(X,Y)$ 的计算公式为：若 (X,Y) 为二维连续型随机变量，(X,Y) 的联合密度为 $f(x,y)$，则 $Eg(X,Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$. 若 (X,Y) 为二维离散型随机变量，X 与 Y 的联合分布律为

$$P(X=a_i,Y=b_j)=p_{ij},\ i,j=1,2,\cdots.$$

则

$$Eg(X,Y)=\sum_i\sum_j g(a_i,b_j)p_{ij}.$$

利用上述公式可容易算出协方差 $\mathrm{cov}(X,Y)=E(XY)-E(x)E(y)$，事实上只需分别取 $g(X,Y)$ 为 XY，X 和 Y 即可. 读者可做下面的习题来练习并熟悉计算方法：

设随机变量 X 和 Y 的联合概率分布为

X \ Y	-1	0	1
0	0.07	0.18	0.15
1	0.08	0.32	0.20

求 $\mathrm{cov}(X,Y)$ 和 $\mathrm{cov}(X^2,Y^2)$. (这是 2002 年试题)答案：$0,-0.02$.

四、小　结

本章重点是求随机变量的数字特征. 一般有两条途径:一是利用分布律(或密度函数)直接计算,二是利用性质来计算. 不论采用哪一种方法,熟记常用分布的数字特征是有益的.

第五章　大数定律和中心极限定理

大数定律和中心极限定理属于随机变量序列的极限性质，它们涉及的理论较深.

一、复习与考试要求

(1) 了解切比雪夫不等式.

(2) 了解切比雪夫大数定律和伯努利大数定律.

(3) 了解独立同分布的中心极限定理和棣莫弗-拉普拉斯中心极限定理。

二、基本概念与理论

1. 切比雪夫不等式

设 X 是随机变量，对任意一个 $\varepsilon>0$，有

$$P(|X-E(X)|\geqslant\varepsilon)\leqslant\frac{D(X)}{\varepsilon^2}.\ \left(\text{或 } P(|X-E(X)|<\varepsilon)\geqslant 1-\frac{D(X)}{\varepsilon^2}\right)$$

切比雪夫不等式的直观意义是，不论 X 服从何种分布，X 取值于 $E(X)$ 的 ε 的邻域内的概率不小于 $1-\frac{D(X)}{\varepsilon^2}$.

2. 大数定律

(1) 切比雪夫大数定律　设 $X_1,X_2,\cdots$ 是一列两两不相关的随机变量，$E(X_i)=\mu_i$，$D(X_i)=\sigma_i^2$，且存在常数 c，使得 $\sigma_i^2\leqslant c$，$i=1,2,\cdots$，那么，对任意一个 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}\mu_i\right|<\varepsilon\right)=1.$$

(2) 伯努利大数定律　设随机变量 $Y_n\sim B(n,p)$，那么，对任意一个 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left(\left|\frac{Y_n}{n}-p\right|<\varepsilon\right)=1.$$

伯努利大数定律的直观意义是，在大量独立重复试验中可以用某个事件 A 发生的频率来近似每次试验中事件 A 发生的概率.

3. 中心极限定理

(1) 独立同分布情形下的中心极限定理　设 $X_1,X_2,\cdots$ 是一列独立同分布的随机变量，$E(X_i)=\mu$，$D(X_i)=\sigma^2$，$i=1,2,\cdots$，那么，对任意一个实数 x，有

$$\lim_{n\to\infty}P\left(\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\leqslant x\right)=\Phi(x).$$

这个定理的直观意义是，当 n 足够大时，可以近似地认为 $\sum_{i=1}^{n}X_i\sim N(n\mu,n\sigma^2)$. 这个定理也称为列维-林德贝格(Levy-Lindberg)中心极限定理.

(2) 棣莫弗-拉普拉斯中心极限定理　设随机变量 $Y_n\sim B(n,p)$，那么，对任意一个实数 x，有

$$\lim_{n\to\infty}P\left(\frac{Y_n-np}{\sqrt{np(1-p)}}\leqslant x\right)=\Phi(x).$$

4. 二项概率的近似计算

(1) 当 n 足够大且 p 较小时，二项概率

$$P_n(k)=\binom{n}{k}p^k(1-p)^{n-k}\approx e^{-\lambda}\cdot\frac{\lambda^k}{k!}.$$

其中，$\lambda=np$.

(2) 当 n 足够大时，有

$$\sum_{i=0}^{k}P_n(i)=\sum_{i=0}^{k}\binom{n}{i}p^i(1-p)^{n-i}\approx\Phi\left(\frac{k-np}{\sqrt{np(1-p)}}\right).$$

5. 依概率收敛性

设 $Y_1,Y_2,\cdots$ 是一个随机变量序列，a 是一个常数. 如果对任意一个 $\varepsilon>0$，有

$$\lim_{n\to\infty}P(|Y_n-a|\geqslant\varepsilon)=0,\ (\text{或}\lim_{n\to\infty}P(|Y_n-a|<\varepsilon)=1).$$

那么称这个随机变量序列 $Y_1,Y_2,\cdots$ 依概率收敛于 a，记作 $Y_n\xrightarrow{P}a$.

切比雪夫大数定律表明 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_i)\xrightarrow{P}0$；伯努利大数定理表明 $\frac{Y_n}{n}\xrightarrow{P}p$，这里 $Y_n\sim B(n,P)$.

当 $X_1,X_2,\cdots$ 是独立同分布随机变量时，有

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\mu,$$

其中，$\mu=E(X_i)$，$i=1,2,\cdots$.

三、例题分析

例 5-1　设连续型随机变量 X 具有数学期望 $E(X)=a$ 和方差 $D(X)=\sigma^2$. 求证，对任意的正数 ε，有

$$P(|X-a|<\varepsilon)\geqslant1-\frac{\sigma^2}{\varepsilon^2}.$$

证　设 X 的密度函数为 $f(x)$. 由于在集合 $\{x\,||x-a|\geqslant\varepsilon\}$ 上 $\frac{(x-a)^2}{\varepsilon^2}\geqslant1$，因此

$$\begin{aligned}P(|X-a|\geqslant\varepsilon)&=\int_{|x-a|\geqslant\varepsilon}f(x)\mathrm{d}x\\&\leqslant\int_{|x-a|\geqslant\varepsilon}\frac{(x-a)^2}{\varepsilon^2}f(x)\mathrm{d}x\\&\leqslant\int_{|x-a|\geqslant\varepsilon}\frac{(x-a)^2}{\varepsilon^2}f(x)\mathrm{d}x+\int_{|x-a|<\varepsilon}\frac{(x-a)^2}{\varepsilon^2}f(x)\mathrm{d}x\\&=\int_{-\infty}^{\infty}\frac{(x-a)^2}{\varepsilon^2}f(x)\mathrm{d}x=\frac{1}{\varepsilon^2}\int_{-\infty}^{\infty}(x-a)^2f(x)\mathrm{d}x\\&=\frac{1}{\varepsilon^2}\cdot D(X)=\frac{\sigma^2}{\varepsilon^2}.\end{aligned}$$

从而由求逆公式推得

$$P(|X-a|<\varepsilon)=1-P(|X-a|\geqslant\varepsilon)\geqslant1-\frac{\sigma^2}{\varepsilon^2}.$$

分析　本题要求证明切比雪夫不等式，必须写出证明过程. 需要用到的技巧必须记住. 本题是 1988 年备用试题.

例 5-2　某人要测量甲、乙两地之间的距离，限于测量工具，他把总距离分成 1200 段进行测量，每段测量误差(单位：cm)相互独立，且都服从区间 $(-0.5,0.5)$ 上的均匀分布. 试求总距离测量误差的绝对值不超过 20cm 的概率(已知 $\Phi(2)=0.95$).

解　设第 i 段测量误差为 X_i，$X_i\sim R(-0.5,0.5)$，

$$E(X_i)=0,\quad D(X_i)=\frac{1}{12},$$

$i=1,\cdots,1200$. 由独立同分布情形下的中心极限定理推得，可以近似地认为 $\sum_{i=1}^{1200}X_i\sim N\left(1200\times 0,1200\times\frac{1}{12}\right)=N(0,100)$. 于是，所求概率为

$$P\left(\left|\sum_{i=1}^{1200}X_i\right|\leqslant 20\right)=P\left(-20\leqslant\sum_{i=1}^{1200}X_i\leqslant 20\right)=\Phi\left(\frac{20-0}{10}\right)-\Phi\left(\frac{-20-0}{10}\right)$$

$$=\Phi(2)-\Phi(-2)=\Phi(2)-[1-\Phi(2)]=2\Phi(2)-1=2\times 0.95-1=0.90.$$

分析 本题的难点在于构造适当的随机变量. 这里要避免犯一种错误：设每段测量误差为 X，总误差为 $1200X$. 这种错误在于假定每段测量误差全相等. 这是不符合实际情况的，也与题目中的相互独立性违背.

例 5-3 设随机变量 X 具有数学期望 $E(X)=\mu$ 与方差 $D(X)=\sigma^2$. 对任意一个正常数 $k>1$，求证：

$$P(|X-\mu|\geqslant k\sigma)\leqslant\frac{1}{k^2}.$$

证 在切比雪夫不等式中，令 $\varepsilon=k\sigma$ 便得

$$P(|X-\mu|\geqslant k\sigma)\leqslant\frac{\sigma^2}{k^2\sigma^2}=\frac{1}{k^2}.$$

分析 本题是切比雪夫不等式的简单应用. 当 $k\leqslant 1$ 时，$P(|X-\mu|\geqslant k\sigma)\leqslant\frac{1}{k^2}$ 总是成立的，因为 $\frac{1}{k^2}\geqslant 1$. 这表明切比雪夫不等式只是给出了概率的上界，不能把它当作一个近似计算公式.

例 5-4 设 $X_1,X_2,\cdots,X_n$ 是独立同分布的随机变量. 已知 $E(X^k)=\alpha_k$，$k=1,2,3,4$. 试证，当 n 充分大时，随机变量 $Z_n=\frac{1}{n}\cdot\sum_{i=1}^{n}X_i^2$ 近似地服从正态分布，并指出其分布参数.

证 记 $Y_i=X_i^2$，$i=1,\cdots,n$.

$$E(Y_i)=E(X_i^2)=\alpha_2,$$

$$D(Y_i)=E(Y_i^2)-[D(Y_i)]^2=E(X_i^4)-\alpha_2^2=\alpha_4-\alpha_2^2.$$

$Y_1,\cdots,Y_n$ 是独立同分布的随机变量. 由独立同分布情形下的中心极限定理推得 $nZ=\sum_{i=1}^{n}X_i^2=\sum_{i=1}^{n}Y_i$ 近似地服从正态分布 $N(n\alpha_2,n(\alpha_4-\alpha_2^2))$，从而 Z 近似地服从正态分布 $N(\alpha_2,\frac{1}{n}(\alpha_4-\alpha_2^2))$.

例 5-5 设随机变量 X 和 Y 的数学期望都是 2，方差分别为 1 和 4，而相关系数为 0.5，试根据切比雪夫不等式给出 $P\{|X-Y|\geqslant 6\}$ 的上界.

解 记 $U=X-Y$，则

$$EU=EX-EY=2-2=0,$$

$$D(U)=D(X-Y)=D(X)+D(Y)-2\mathrm{cov}(X,Y)=D(X)+D(Y)-2\rho(X,Y)\sqrt{D(X)D(Y)}$$

$$=1+4-2\times 0.5\times\sqrt{1\times 4}=5-2=3.$$

从而，由切比雪夫不等式并注意到 $U-EU=U$，

$$P\{|X-Y|\geqslant 6\}=P\{|U|\geqslant 6\}\leqslant\frac{D(U)}{6^2}=\frac{3}{36}=\frac{1}{12}.$$

分析 本题是 2001 年试题. 解本题用的技巧仍是引入中间随机变量. 本题中引入 $U=X-Y$，然后对随机变量 U 用切比雪夫不等式就得到所要的上界.

例 5-6 设总体 X 服从参数为 2 的指数分布，$X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本，则当 $n\to\infty$ 时，$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 依概率收敛于_____.

解 本题是填空题. 应填“$\frac{1}{2}$”.

分析　本题是 2003 年试题. 由于 $X_1, X_2, \cdots, X_n$ 是独立且服从相同分布的随机变量序列,因此, $X_1^2, X_2^2, \cdots, X_n^2$ 也是独立同分布的随机变量序列,从而由切比雪夫大数定律知 $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} EX_1^2$. 由题设 $X_1 \sim E(2)$, 因此

$$EX_1 = \frac{1}{2}, D(X_1) = \frac{1}{4}.$$

$$E(X_1^2) = D(X_1) + (EX_1)^2 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

因而　$\frac{1}{n}\sum_{i=1}^{n} X_i^2$ 依概率收敛于 $\frac{1}{2}$.

例 5-7　一生产线生产的产品成箱包装,每箱的重量是随机的. 假设每箱平均重 50kg,标准差为 5kg. 若用最大载重量为 5t 的汽车承运,试利用中心极限定理说明每辆车最多可以装多少箱,才能保障不超载的概率大于 0.977. ($\Phi(2) = 0.977$,其中, $\Phi(x)$ 是标准正态分布函数.)

解　设最多装 n 箱. $X_i (i = 1, 2, \cdots, n)$ 为第 i 箱的重量. 由题意 $EX_i = 50$, $\sqrt{D(X_i)} = 5$. $i = 1, 2, \cdots, n$. 并视 $X_1, X_2, \cdots, X_n$ 相互独立且服从相同的分布. 货物的总重量为 $X_1 + \cdots + X_n$,不超载即指"$X_1 + \cdots X_n \leqslant 5000$". 由题意,要求满足

$$P\{X_1 + \cdots + X_n \leqslant 5000) > 0.977.$$

由列维-林德贝格中心极限定理

$$P\{X_1 + \cdots + X_n \leqslant 5000\} = P\left(\frac{X_1 + \cdots + X_n - nX50}{\sqrt{n}5} \leqslant \frac{5000 - 50n}{5\sqrt{n}}\right)$$

$$\approx \Phi\left(\frac{1000 - 10n}{\sqrt{n}}\right) > 0.977 = \Phi(2),$$

由此得到

$$\frac{1000 - 10n}{\sqrt{n}} > 2.$$

即

$$10n + 2\sqrt{n} - 1000 < 0,$$

记 $\sqrt{n} = x$,解方程 $10x^2 + 2x - 1000 = 0$ 得两个根

$$x_1 = \frac{-2 + 2\sqrt{10001}}{2 \times 10} \approx 9.9, \quad x_2 = \frac{-2 - 2\sqrt{10001}}{2 \times 10} \approx -10.1.$$

而使 $10x^2 + 2x - 1000 < 0$, x 满足 $-10.1 < x < 9.9$. 注意到 $x > 0$,因此,只需 $x = \sqrt{n} < 9.9$. 即 $n < 98.01$ 取 $n = 98$ 即可. 即最多可以装 98 箱.

分析　本题是 2001 年试题. 本题中没有指明各箱重量相互独立且服从相同分布. 遇到这类题时只能作上述假设,否则无法用中心极限定理解题. 中心极限定理的应用也是经常考的知识点. 这些题涉及范围很广,例如电力的分配,电话线的配制安排等. 考生对此应加以熟悉.

四、小　结

本章的理论性较强. 大数定律和中心极限定理的应用是要重点掌握的内容,对切比雪夫不等式也要有一定的了解. 这些年对这方面内容的要求有所提高,因此读者还须认真对待.

第六章　数理统计的基本概念

本章内容既是由概率论向数理统计过渡的桥梁，也是今后讨论统计推断（估计与检验）的必要准备.

一、复习与考试要求

（1）理解总体、个体、简单随机样本和统计量的概念，掌握样本均值、样本方差及样本矩的计算.

（2）了解 χ^2 分布、t 分布和 F 分布的定义及性质，了解分位数的概念并会查表计算.

（3）了解正态总体的某些常用统计量的分布.

二、基本概念与理论

1. 总体与样本

研究对象的某个数值指标的全体称为总体；从总体中抽取若干个成员，称这些成员为样本，称成员个数为样本大小（或样本容量）。

设总体 X 是具有未知的分布函数 $F(x)$（或未知的分布律，或未知的密度函数）的随机变量，一般有下列两种情形：

（1）X 的分布类型已知，但含有若干个未知参数；

（2）X 的分布类型未知.

简单随机样本（以下简称样本）$(X_1,\cdots,X_n)$ 是一个 n 维随机变量，它表示 $X_1,\cdots,X_n$ 是独立同分布的随机变量，且每一个 X_i 的分布都与总体 X 的分布相同，$i=1,\cdots,n$. 在抽样后，记样本观测值为 $(x_1,\cdots,x_n)$. 在实际问题中，$(x_1,\cdots,x_n)$ 即是数据.

在区间估计与假设检验中，只讨论正态总体，即已知总体 $X\sim N(\mu,\sigma^2)$. 正态总体有以下三种类型：

（1）μ 未知，但 σ^2 已知；

（2）σ^2 未知，但 μ 已知；

（3）μ 与 σ^2 均未知.

2. 统计量

样本 $(X_1,\cdots,X_n)$ 的函数称为统计量，要求它不直接包含总体分布中任何未知参数。在抽样前，统计量是一个随机变量；在抽样后，统计量的观测值是一个可以由数据算得的实数.

样本均值定义为 $\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$；样本方差定义为 $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$，称 $S=\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}$ 为样本标准差. 记 $S_n^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-X)^2$，$S_n=\sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-X)^2}$. 这些都是常用统计量.

对任意一个正整数 k，称

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^{\,k}$$

为样本的 k 阶原点矩；称

$$M_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k$$

为样本的 k 阶中心矩，这些都是统计量.

3. 统计量的数字特征

设 $(X_1,\cdots,X_n)$ 是取自总体 X 的一个样本，$E(X)=\mu$，$D(X)=\sigma^2$. 那么，有

（1）$E(\overline{X})=\mu$，$D(\overline{X})=\dfrac{\sigma^2}{n}$；

(2) $E(S^2)=\sigma^2$，　$E(S_n^2)=\dfrac{n-1}{n}\sigma^2$.

4. χ^2 分布、t 分布与 F 分布

(1) 设 $X_1,\cdots,X_n$ 是独立同分布的随机变量，且都服从 $N(0,1)$，那么，$\chi^2=\sum\limits_{i=1}^{n}X_i^2$ 服从自由度为 n 的 χ^2 分布 $\chi^2(n)$，且

$$E(\chi^2)=n,\quad D(\chi^2)=2n.$$

当 $\chi_1^2\sim\chi^2(m)$，$\chi_2^2\sim\chi^2(n)$，且 χ_1^2 与 χ_2^2 相互独立时，有

$$\chi_1^2+\chi_2^2\sim\chi^2(m+n).$$

$\chi^2(n)$分布的 p 分位数记作 $\chi_p^2(n)$. 它表示：当 $\chi^2\sim\chi^2(n)$时，有

$$P(\chi^2\leqslant\chi_p^2(n))=p.$$

$\chi_p^2(n)$的值可以查表得到.

(2) 设随机变量 X 与 Y 相互独立，且 $X\sim N(0,1)$，$Y\sim\chi^2(n)$，那么，$T=\dfrac{X}{\sqrt{Y/n}}$服从自由度为 n 的 t 分布 $t(n)$.

$t(n)$分布的 p 分位数记作 $t_p(n)$. 它表示：当 $T\sim t(n)$时，有

$$P(T\leqslant t_p(n))=p.$$

$t_p(n)$的值可以查表得到，它满足关系式

$$t_p(n)=-t_{1-p}(n).$$

(3) 设随机变量 X 与 Y 相互独立，且 $X\sim\chi^2(m)$，$Y\sim\chi^2(n)$，那么，$F=\dfrac{X/m}{Y/n}$服从自由度为(m,n)的 F 分布 $F(m,n)$.

$F(m,n)$分布的 p 分位数记作 $F_p(m,n)$. 它表示：当$F\sim F(m,n)$时，有

$$P(F\leqslant F_p(m,n))=p.$$

$F_p(m,n)$的值可以查表得到，它满足关系式

$$F_p(m,n)=\frac{1}{F_{1-p}(n,m)}.$$

(4) $N(0,1)$分布的 p 分位数记作 u_p. 它表示：当 $X\sim N(0,1)$时，有

$$P(X\leqslant u_p)=\Phi(u_p)=p,$$

即 $u_p=\Phi^{-1}(p)$. u_p 的值可以查表得到，它满足关系式

$$u_p=-u_{1-p}.$$

5. 样本 $X_1,\cdots,X_n$ 的联合分布

设总体 X 是一个连续型随机变量，密度函数为 $f(x)$. 样本 $X_1,\cdots,X_n$ 的联合密度函数为

$$f^*(x_1,\cdots,x_n)=\prod_{i=1}^{n}f(x_i).$$

设总体 X 是一个离散型随机变量，分布律为 $P(X=a_i)=p_i$，$i=1,2,\cdots$. 样本 $X_1,\cdots,X_n$ 的联合分布律为

$$f^*(x_1,\cdots,x_n)=\prod_{i=1}^{n}P(X_i=x_i),\quad x_i\in\{a_1,a_2,\cdots\}.$$

6. 正态总体下常用统计量的分布

定理　设$(X_1,\cdots,X_n)$是取自正态总体 $N(\mu,\sigma^2)$的一个样本. 那么，有

(1) $\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$，$\sqrt{n}\,\dfrac{\overline{X}-\mu}{\sigma}\sim N(0,1)$；

(2) $\dfrac{nS_n^2}{\sigma^2}=\dfrac{(n-1)S^2}{\sigma^2}=\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$;

(3) $\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n)$;

(4) $\overline{X}$ **与** S^2(**或** S_n^2)**相互独立**;

(5) $\sqrt{n}\dfrac{\overline{X}-\mu}{S}\sim t(n-1)$.

定理　设$(X_1,\cdots,X_m)$**是取自正态总体** $N(\mu_1,\sigma_1^2)$**的一个样本**,$(Y_1,\cdots,Y_n)$**是取自正态总体** $N(\mu_2,\sigma_2^2)$**的一个样本.那么,有**

(1) $\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}\sim N(0,1)$;

(2) $\dfrac{\sum\limits_{i=1}^{m}(X_i-\mu_1)^2/m\sigma_1^2}{\sum\limits_{i=1}^{n}(Y_i-\mu_2)^2/n\sigma_2^2}\sim F(m,n)$;

(3) $\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(m-1,n-1)$;

(4) **当** $\sigma_1^2=\sigma_2^2$ **时,有**

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}\sim t(m+n-2),$$

其中 $S_w=\sqrt{\dfrac{1}{(m+n-2)}\left[\sum\limits_{i=1}^{m}(X_i-\overline{X})^2+\sum\limits_{i=1}^{n}(Y_i-\overline{Y})^2\right]}$;(3)**成为**

$$\frac{S_1^2}{S_2^2}\sim F(m-1,n-1).$$

其中 $S_1^2=\dfrac{1}{m-1}\sum\limits_{i=1}^{m}(X_i-\overline{X})^2$, $S_2^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(Y_i-\overline{Y})^2$,$\overline{X}=\dfrac{1}{m}\sum\limits_{i=1}^{m}X_i$, $\overline{Y}=\dfrac{1}{n}\sum\limits_{i=1}^{n}Y_i$.

三、例题分析

例 6-1　已知数据 $x_1,\cdots,x_n$ **满足** $\sum\limits_{i=1}^{n}x_i=a$, $\sum\limits_{i=1}^{n}x_i^2=b$. **试求** $\bar{x},s^2,s_n^2$.

解　由所给条件得到 $\bar{x}=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i=\dfrac{a}{n}$;**另外,由**

$$\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}x_i^2-n\bar{x}^2=b-\frac{a^2}{n}$$

推得 $s_n^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i-\bar{x})^2=\dfrac{1}{n^2}(nb-a^2)$,$s^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\bar{x})^2=\dfrac{1}{n(n-1)}(nb-a^2)$.

分析　本题的意义在于:可以利用中间数据算得常用统计量的观测值.这些公式在数据计算题中常常很有用。

例 6-2　设 $X_1,\cdots,X_8$ **是取自正态总体** $N(\mu,\sigma^2)$**的一个样本,其中** $\mu=1$,σ^2 **未知.在** $\max\limits_{1\leqslant i\leqslant 8}X_i$,$\dfrac{1}{8}\sum\limits_{i=1}^{8}(X_i-\mu)^2$,$k\overline{X}+b(k,b$ **为已知常数**),$\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{8}(X_i-1)^2$ **中,哪个不是统计量?**

解　$\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{8}(X_i-1)^2$ **不是统计量,因为它包含未知参数** σ^2.

例 6-3　设 $X_1,\cdots,X_n$ 是取自正态总体 $N(0,\sigma^2)$ 的一个样本. 试求样本二阶原点矩 $A_2=\frac{1}{n}\sum\limits_{i=1}^{n}X_i^2$ 的数学期望与方差.

解　由数学期望与方差的性质推得

$$E(A_2)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)=\frac{1}{n}\sum_{i=1}^{n}E(X^2)=E(X^2),$$

$$D(A_2)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i^2)=\frac{1}{n^2}\sum_{i=1}^{n}D(X^2)=\frac{1}{n}D(X^2).$$

因为 $X\sim N(0,\sigma^2)$, $\frac{X}{\sigma}\sim N(0,1)$, 所以 $\frac{X^2}{\sigma^2}\sim\chi^2(1)$. 由此推得

$$E\left(\frac{X^2}{\sigma^2}\right)=1,\quad D\left(\frac{X^2}{\sigma^2}\right)=2,$$

即

$$E(X^2)=\sigma^2,\quad D(X^2)=2\sigma^4.$$

从而

$$E(A_2)=\sigma^2,\quad D(A_2)=\frac{2\sigma^4}{n}.$$

分析　一般要求 A_2 的方差涉及四阶矩. 但是, 在正态分布情形下, 可以利用 χ^2 分布的方差来回避四阶矩的计算.

例 6-4　设 $X_1,\cdots,X_6$ 是取自正态总体 $N(1,\sigma^2)$ 的一个样本. 已知 $P(\overline{X}>1,S<c\sigma)=p$. 试用分位数表示常数 c, 其中 $0<p<0.5$.

解　由 $\overline{X}$ 与 S 的相互独立性推得

$$p=P(\overline{X}>1,\quad S<c\sigma)=P(\overline{X}>1)P(S<c\sigma).$$

由 $\overline{X}\sim N\left(1,\frac{\sigma^2}{6}\right)$ 知道

$$P(\overline{X}>1)=1-P(\overline{X}\leqslant 1)=1-\Phi\left(\frac{1-1}{\sigma/\sqrt{6}}\right)$$

$$=1-\Phi(0)=1-\frac{1}{2}=\frac{1}{2},$$

从而, $P(S<c\sigma)=2p$. 另一方面, 由 $\frac{5S^2}{\sigma^2}\sim\chi^2(5)$ 推得

$$P(S<c\sigma)=P(S^2<c^2\sigma^2)=P(5S^2<5c^2\sigma^2)=P\left(\frac{5S^2}{\sigma^2}<5c^2\right)=2p,$$

因此, $5c^2=\chi^2_{2p}(5)$, 即 $c=\sqrt{\frac{1}{5}\chi^2_{2p}(5)}$.

分析　本题的关键是利用 $\overline{X}$ 与 S^2 的相互独立性.

例 6-5　设随机变量 $T\sim t(n)$. 试证 $T^2\sim F(1,n)$.

证　由 T 分布的定义知道, $T=\frac{X}{\sqrt{Y/n}}$, 其中, X 与 Y 相互独立, 且 $X\sim N(0,1)$, $Y\sim\chi^2(n)$, 现在

$$T^2=\frac{X^2}{Y/n}=\frac{X^2/1}{Y/n},$$

其中, X^2 与 Y 相互独立, $Y\sim\chi^2(n)$, 由 χ^2 分布的定义推得 $X^2\sim\chi^2(1)$, 于是, 由 F 分布的定义得到 $T^2\sim F(1,n)$.

分析　服从 $N(0,1)$ 的随机变量的平方服从 $\chi^2(1)$ 分布. 这相当于 χ^2 分布定义中取 $n=1$. 这一事实我们在例 6-3 中已经使用过一次了.

例 6-6　设 $X_1,\cdots,X_n$ 是取自正态总体 $N(0,\sigma^2)$ 的一个样本, 试问, 当 c 取何值时, 随机变量 $Y=c$

· $(\sum_{i=1}^{n} X_i)^2$ 服从 χ^2 分布？指出这个 χ^2 分布的自由度.

解 由正态分布的可加性定理知道，$\sum_{i=1}^{n} X_i \sim N(0, n\sigma^2)$，即 $\frac{1}{\sqrt{n\sigma^2}} \sum_{i=1}^{n} X_i \sim N(0,1)$，于是，有

$$\left(\frac{1}{\sqrt{n\sigma^2}} \sum_{i=1}^{n} X_i\right)^2 = \frac{1}{n\sigma^2} (\sum_{i=1}^{n} X_i)^2 \sim \chi^2(1),$$

这表明 $c=\frac{1}{n\sigma^2}$，且自由度为 1.

例 6-7 （选择题）设随机变量 $X \sim t(n)(n>1)$，$Y=\frac{1}{X^2}$，则 （ ）

(A) $Y \sim \chi^2(n)$；　　(B) $Y \sim \chi^2(n-1)$；

(C) $Y \sim F(n,1)$；　　(D) $Y \sim F(1,n)$.

解 应选(C).

分析 本题是 2003 年试题. 解本题的关键是要对 χ^2 分布、t 分布和 F 分布的定义十分熟悉. 由于 $X \sim t(n)$，因此 $X=\frac{U}{\sqrt{V/n}}$，其中，U,V 相互独立且 $U \sim N(0,1)$，$V \sim \chi^2(n)$. 从而 $Y=\frac{1}{X^2}=\frac{V/n}{U^2}$，而由于 U,V 独立，且 $U^2 \sim \chi^2(1)$，由 F 分布的定义得到 $Y \sim F(n,1)$.

例 6-8 设 $X_1, X_2, \cdots, X_9$ 是来自正态总体 X 的简单随机样本，

$$Y_1=\frac{1}{6}(X_1+\cdots+X_6),\ Y_2=\frac{1}{3}(X_7+X_8+X_9),$$

$$S^2=\frac{1}{2}\sum_{i=7}^{9}(X_i-Y_2)^2,\ Z=\frac{\sqrt{2}(Y_1-Y_2)}{S}.$$

证明统计量 Z 服从自由度为 2 的 t 分布.

证 设 $X \sim N(\mu,\sigma^2)$，由于

$$Y_1-Y_2=\frac{1}{6}X_1+\cdots+\frac{1}{6}X_6-\frac{1}{3}X_7-\frac{1}{3}X_8-\frac{1}{3}X_9$$

即 Y_1-Y_2 是 $X_1, X_2, \cdots, X_9$ 的线性组合，因此 Y_1-Y_2 服从正态分布. 又因为 Y_1 与 Y_2 相互独立，所以

$$E(Y_1-Y_2)=EY_1-EY_2=\mu-\mu=0,$$

$$D(Y_1-Y_2)=D(Y_1)+D(Y_2)=\frac{1}{6}\sigma^2+\frac{1}{3}\sigma^2=\frac{1}{2}\sigma^2.$$

从而

$$\frac{\sqrt{2}(Y_1-Y_2)}{\sigma} \sim N(0,1).$$

由 S^2 的定义可知：S^2 是由 X_7, X_8, X_9 建立的样本方差. 由正态总体的定理可知：$\frac{2S^2}{\sigma^2} \sim \chi^2(2)$ 且 S^2 与 Y_2 相互独立. 再由 Y_1 是 $X_1, \cdots, X_6$ 的函数，S^2 是 X_7, X_8, X_9 的函数，所以有 S^2 与 Y_1 相互独立. 从而，S^2 与 Y_1-Y_2 相互独立. 由 t 分布的定义

$$Z=\frac{\frac{\sqrt{2}(Y_1-Y_2)}{\sigma}}{\sqrt{\frac{\frac{2S^2}{\sigma^2}}{2}}} \sim t(2).$$

分析 本题是 1999 年试题. 解本题的关键是要对正态总体下常用统计量的分布十分熟悉，并能灵活运用.

例 6-9 （选择题）设随机变量 X 和 Y 都服从标准正态分布，则（ ）

(A) $X+Y$ 服从正态分布；　　(B) X^2+Y^2 服从 χ^2 分布；

(C) X^2 和 Y^2 都服从 χ^2 分布；　　(D) X^2/Y^2 服从 F 分布.

解　应选(C).

分析　本题是 2002 年试题. 由题意，$X^2 \sim \chi^2(1)$，$Y^2 \sim \chi^2(1)$. 由于 X，Y 并不一定相互独立，因此(A)，(B)，(D)均不一定成立，例如取 $Y=-X$，那么，$X+Y=0$，$X^2+Y^2=2X^2$，$X^2/Y^2=1$，此时(A)，(B)和(D)均不成立.

例 6-10　设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0,2^2)$ 的简单样本，$X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$，则当 a，b 取何值时 X 服从 χ^2 分布？并指出其自由度.

解　由于 X_1-2X_2，$3X_3-4X_4$ 均服从正态分布，且

$$E(X_1-2X_2)=0, D(X_1-2X_2)=D(X_1)+4D(X_2)=5\times 2^2=20.$$

$$E(3X_3-4X_4)=0, D(3X_3-4X_4)=9D(X_3)+16D(X_4)=25\times 2^2=100.$$

由此得

$$\frac{X_1-2X_2}{\sqrt{20}} \sim N(0,1),\ \frac{3X_3-4X_4}{10} \sim N(0,1)$$

由 X_1-2X_2 与 $3X_3-4X_4$ 相互独立，所以，由 χ^2 分布定义

$$\frac{(X_1-2X_2)^2}{20}+\frac{(3X_3-4X_4)^2}{100} \sim \chi^2(2).$$

从而，当 $a=\frac{1}{20}$，$b=\frac{1}{100}$ 时 $X \sim \chi^2(2)$.

分析　本题是 1998 年试题. 本题的解法与例 6-6 是一样的.

例 6-11　从正态总体 $N(3.4,6^2)$ 中抽取容量为 n 的样本，如果要求其样本均值位于区间(1.4，5.4)内的概率不小于 0.95，问样本容量 n 至少应取多大？

解　设 $X_1, X_2, \cdots, X_n$ 是取自总体 $N(3.4,6^2)$ 的样本，样本均值为 $\overline{X}$. 由于 $\frac{\sqrt{n}(\overline{X}-3.4)}{6} \sim N(0,1)$，因此

$$P(1.4<\overline{X}<5.4)=P\left(-\frac{2\sqrt{n}}{6}<\sqrt{n}\frac{(\overline{X}-3.4)}{6}<\frac{2\sqrt{n}}{6}\right)$$

$$=\Phi\left(\frac{2\sqrt{n}}{6}\right)-\Phi\left(-\frac{2\sqrt{n}}{6}\right)=2\Phi\left(\frac{2\sqrt{n}}{6}\right)-1.$$

由题意

$$2\Phi\left(\frac{2\sqrt{n}}{6}\right)-1 \geqslant 0.95,$$

即

$$\Phi\left(\frac{2\sqrt{n}}{6}\right) \geqslant 0.975.$$

由于 $\Phi(1.96)=0.975$，所以

$$\frac{2\sqrt{n}}{6} \geqslant 1.96.$$

解得　$n \geqslant 34.57$，所以，n 至少应取 35.

分析　本题是 1998 年试题. 本题中样本容量 n 为正整数，因此，最后答案应取满足条件 $n \geqslant 34.57$ 的最小整数，即取 n 为 35.

四、小　结

在本章中，数理统计的名词仅仅是表象，实质仍是概率论中随机变量的分布与数字特征. “设 $X_1, X_2,$

…，X_n 是取自某个总体的一个样本”蕴含了“X_1，…，X_n 是独立同(总体)分布的随机变量”，由此推得 $E(X_i)=E(X)$，$D(X_i)=D(X)$，$i=1,2,\cdots,n$. 对正态总体下常用统计量的分布的有关结论在以后的内容中常会出现，读者应熟练掌握这些结论.

第七章　参数估计

参数估计就是利用样本来估计总体分布中的未知参数. 参数估计有两种形式:点估计(用某个实数作估计)与区间估计(用某个区间作估计).

一、复习与考试要求

(1) 理解点估计的概念.
(2) 掌握矩估计法(一阶、二阶)和极大似然估计法.
(3) 了解估计量的评选标准(无偏性、有效性和一致性).
(4) 理解区间估计的概念.
(5) 会求单个正态总体的均值和方差的置信区间.
(6) 会求两个正态总体的均值差和方差比的置信区间.

二、基本概念与理论

1. 点估计的统计意义

假定要利用样本 $X_1,\cdots,X_n$ 来估计总体分布中的未知参数 θ. 要求 θ 的点估计为 $\hat{\theta}$,即是构造一个统计量 $h(X_1,\cdots,X_n)=\hat{\theta}$,称 $\hat{\theta}$ 为 θ 的估计量. 在抽样后,称 $\hat{\theta}$ 为 θ 的估计值. 两者都可以简称为 θ 的估计.

设 $\hat{\theta}$ 是 θ 的估计量,那么 $g(\hat{\theta})$ 便是 $g(\theta)$ 的估计量.

2. 矩估计法与极大似然估计法

(1) **矩估计量**　设总体 X 的 k 阶原点矩 $E(X^k)=\alpha_k,k=1,2,\cdots$. 如果未知参数 $\theta=\varphi(\alpha_1,\cdots,\alpha_m)$,那么 θ 的矩估计量 $\hat{\theta}=\varphi(A_1,\cdots,A_m)$,其中 $A_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^k$ 是样本的 k 阶原点矩. 只要掌握 $m=1,2$ 的情形.

定理　设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本,$E(X)=\mu,D(X)=\sigma^2$,其中,μ 与 σ^2 均未知,那么,$\overline{X}$ 是 μ 的矩估计量,S_n^2 是 σ^2 的矩估计量,S_n 是 σ 的矩估计量.

(2) **极大似然估计量**　设样本 $X_1,\cdots,X_n$ 的联合密度函数(或联合分布律)为 $f^*(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k)$,其中 $\theta_1,\cdots,\theta_k$ 是总体分布的 k 个未知参数. 记

$$L(\theta_1,\cdots,\theta_k)=f^*(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k),$$

并称它为似然函数. $(\theta_1,\cdots,\theta_k)$ 的极大似然估计量 $(\hat{\theta}_1,\cdots,\hat{\theta}_k)$ 满足

$$L(\hat{\theta}_1,\cdots,\hat{\theta}_k)=\max_{(\theta_1,\cdots,\theta_k)}L(\theta_1,\cdots,\theta_k).$$

只要掌握 $k=1,2$ 的情形.

求极大似然估计量的常用方法是求解似然方程组

$$\begin{cases}\dfrac{\partial}{\partial\theta_1}\ln L(\theta_1,\cdots,\theta_k)=0\\ \cdots\\ \dfrac{\partial}{\partial\theta_k}\ln L(\theta_1,\cdots,\theta_k)=0.\end{cases}$$

上述方程组的解是样本 $(x_1,\cdots,x_n)$ 的函数,即得 $\theta_1,\cdots,\theta_k$ 的极大似然估计量.

3. 估计量的评选标准

(1) **无偏性**　$\hat{\theta}$ 是 θ 的无偏估计量的充分必要条件为

$$E(\hat{\theta})=\theta.$$

定理 设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本，$E(X)=\mu$，$D(X)=\sigma^2$. 那么，$\overline{X}$ 是 μ 的无偏估计量，S^2 是 σ^2 的无偏估计量；一般地 S_n^2 不是 σ^2 的无偏估计量，S_n 与 S 不是 σ 的无偏估计量.

(2) **有效性** 如果 $\hat{\theta}^*$ 与 $\hat{\theta}$ 都是 θ 的无偏估计量，且

$$D(\hat{\theta}^*)<D(\hat{\theta}),$$

那么称 $\hat{\theta}^*$ 比 $\hat{\theta}$ 有效.

(3) **一致性(相合性)** 如果 $\hat{\theta}$ 是 θ 的估计量，且

$$\hat{\theta}\xrightarrow{P}\theta,\quad n\to\infty,$$

那么称 $\hat{\theta}$ 是 θ 的一致(或相合)估计量.

定理 设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本，$E(X)=\mu$，$D(X)=\sigma^2$. 那么，$\overline{X}$ 是 μ 的一致估计量；S^2 与 S_n^2 都是 σ^2 的一致估计量；S 与 S_n 都是 σ 的一致估计量.

定理 设 $\hat{\theta}$ 是 θ 的无偏估计量. 那么，当 $D(\hat{\theta})\to 0$ 时，$\hat{\theta}$ 是 θ 的一致估计量.

4. 区间估计的统计意义

假定要利用样本 $X_1,\cdots,X_n$ 来估计总体分布中的未知参数 θ. 要求 θ 的区间估计为 $[\underline{\theta},\overline{\theta}]$，即是构造两个统计量 $h_1(X_1,\cdots,X_n)=\underline{\theta}$，$h_2(X_1,\cdots,X_n)=\overline{\theta}$.

置信区间是区间估计的一种常用形式. 置信水平(或置信度)$1-\alpha$ 下的置信区间 $[\underline{\theta},\overline{\theta}]$ 必定满足

$$P(\underline{\theta}\leqslant\theta\leqslant\overline{\theta})\geqslant 1-\alpha$$

其中，$\underline{\theta}$ 称为置信下限，$\overline{\theta}$ 称为置信上限.

在置信水平 $1-\alpha$ 下求未知参数 θ 的置信区间的一般步骤如下：

(1) 求出未知参数 θ 的一个较优的点估计 $\hat{\theta}$. 我们建议尽可能使用 θ 的极大似然估计；

(2) 以 $\hat{\theta}$ 为基础，构造一个随机变量

$$J=J(X_1,\cdots,X_n;\theta),$$

它必须包含，也只能包含所要估计的未知参数 θ. 要求 J 的分位数能通过查表或计算得到具体数值；

(3) 设 J 的 $\frac{\alpha}{2}$ 分位数为 a，$1-\frac{\alpha}{2}$ 分位数为 b，即

$$P(a\leqslant J\leqslant b)=1-\alpha;$$

(4) 把不等式“$a\leqslant J\leqslant b$”作等价变形，使它成为“$\underline{\theta}\leqslant\theta\leqslant\overline{\theta}$”，这个 $[\underline{\theta},\overline{\theta}]$ 便为所求的置信区间.

如果在置信水平 $1-\alpha$ 下 θ 的置信区间为 $[\underline{\theta},\overline{\theta}]$，那么，当 $g(\theta)$ 是 θ 的单调增加函数时，$[g(\underline{\theta}),g(\overline{\theta})]$ 便是在置信水平 $1-\alpha$ 下 $g(\theta)$ 的置信区间.

5. 单个正态总体

设 $X_1,\cdots,X_n$ 是取自正态总体 $N(\mu,\sigma^2)$ 的一个样本；取置信水平为 $1-\alpha$.

(1) 当 σ^2 已知时，μ 的置信区间的置信上、下限分别是

$$\overline{X}\pm u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}.$$

(2) 当 μ 已知时，σ^2 的置信区间是

$$\left[\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)},\quad \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}\right].$$

(3) 当 μ 与 σ^2 均未知时，μ 的置信区间的置信上、下限分别是

$$\overline{X}\pm t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}};$$

σ^2 的置信区间是

$$\left[\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)},\quad \frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right].$$

6. 两个正态总体

设 $X_1,\cdots,X_m$ 是取自正态总体 $N(\mu_1,\sigma_1^2)$ 的一个样本，$Y_1,\cdots,Y_n$ 是取自正态总体 $N(\mu_2,\sigma_2^2)$ 的一个样本；取置信水平为 $1-\alpha$.

(1) 当 σ_1^2,σ_2^2 已知时，$\mu_1-\mu_2$ 的置信区间的置信上、下限分别是

$$(\overline{X}-\overline{Y})\pm u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}};$$

(2) 当 σ_1,σ_2 未知，但 $\sigma_1^2=\sigma_2^2$ 时，$\mu_1-\mu_2$ 的置信上、下限分别是

$$(\overline{X}-\overline{Y})\pm t_{1-\frac{\alpha}{2}}(m+n-2)S_w\sqrt{\frac{1}{m}+\frac{1}{n}};$$

(3) 当 μ_1,μ_2 已知时，$\frac{\sigma_1^2}{\sigma_2^2}$ 的置信区间是

$$\left[\frac{n\sum_{i=1}^{m}(X_i-\mu_1)^2/m\sum_{i=1}^{n}(Y_i-\mu_2)^2}{F_{1-\frac{\alpha}{2}}(m,n)},\frac{n\sum_{i=1}^{m}(X_i-\mu_1)^2/m\sum_{i=1}^{n}(Y_i-\mu_2)^2}{F_{\frac{\alpha}{2}}(m,n)}\right];$$

(4) 当 μ_1,μ_2 均未知时，$\frac{\sigma_1^2}{\sigma_2^2}$ 的置信区间是

$$\left[\frac{S_1^2/S_2^2}{F_{1-\frac{\alpha}{2}}(m-1,n-1)},\quad \frac{S_1^2/S_2^2}{F_{\frac{\alpha}{2}}(m-1,n-1)}\right].$$

三、例题分析

例 7-1　设 $X_1,\cdots,X_n$ 是取自正态总体 $N(\mu,\sigma^2)$ 的一个样本. 在正态总体的三种类型下，分别求出未知参数 μ,σ^2 的矩估计量与极大似然估计量，并讨论它们的无偏性.

解　由前述定理知道，在正态总体的三种类型下，μ 的矩估计量都是 $\overline{X}$，且是无偏估计量；当 μ 未知时，σ^2 的矩估计量是 S_n^2，且不是无偏估计量.

样本 $X_1,\cdots,X_n$ 的联合密度函数

$$f^*(x_1,\cdots,x_n)=\prod_{i=1}^{n}\left[\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}\right]=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}.$$

当 μ 未知但 σ^2 已知时，似然函数

$$L(\mu)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\},$$

$$\ln L(\mu)=-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.$$

由似然方程 $\frac{\mathrm{d}}{\mathrm{d}\mu}\ln L(\mu)=0$ 解出 $\mu=\frac{1}{n}\sum_{i=1}^{n}x_i=\overline{x}$. 因此，$\mu$ 的极大似然估计量 $\hat{\mu}=\overline{X}$，且是 μ 的无偏估计量. 当 σ^2

未知但 μ 已知时，似然函数

$$L(\sigma^2)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\},$$

$$\ln L(\sigma^2)=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln(\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.$$

由似然方程 $\frac{\mathrm{d}}{\mathrm{d}\sigma^2}\ln L(\sigma^2)=0$ 解出 $\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$. 因此，$\sigma^2$ 的极大似然估计量

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2.$$

由于

$$E(\hat{\sigma}^2)=E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i-\mu)^2$$

$$=\frac{1}{n}\sum_{i=1}^{n}D(X_i)=\frac{1}{n}\sum_{i=1}^{n}\sigma^2=\frac{1}{n}\cdot n\sigma^2=\sigma^2,$$

因此，$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu^2)$ 是 σ^2 的无偏估计量. 当 μ 与 σ^2 均未知时，似然函数

$$L(\mu,\sigma^2)=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\},$$

$$\ln L(\mu,\sigma^2)=-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.$$

由似然方程组

$$\begin{cases}\dfrac{\partial}{\partial\mu}\ln L(\mu,\sigma^2)=\dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(x_i-\mu)=0\\[2ex]\dfrac{\partial}{\partial\sigma^2}\ln L(\mu,\sigma^2)=-\dfrac{n}{2\sigma^2}+\dfrac{\displaystyle\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4}=0\end{cases}$$

解出 $\mu=\bar{x}$，$\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$. 因此，$\mu$ 的极大似然估计量 $\hat{\mu}=\overline{X}$，且是 μ 的无偏估计量；σ^2 的极大似然估计量 $\hat{\sigma}^2=S_n^2$，且不是 σ^2 的无偏估计量.

分析 本题的结论要熟记，它们是讨论置信区间与假设检验的基础. 另外，三种不同类型下的似然函数的本质区别在于自变量的不同，应当加以注意.

例 7-2 设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本，$X\sim B(1,p)$，其中 p 未知. 试求 p 的矩估计量与极大似然估计量，并讨论它们的无偏性.

解 由于 $E(X)=p$，因此 p 的矩估计量是 $\overline{X}$，且是 p 的无偏估计量.

总体 $X\sim B(1,p)$，它的分布律可以表达成

$$P(X=x)=p^x(1-p)^{1-x},\quad x=0,1.$$

似然函数

$$L(p)=\prod_{i=1}^{n}P(X_i=x_i)=\prod_{i=1}^{n}\left[p^{x_i}(1-p)^{1-x_i}\right]=p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i},$$

$$\ln L(p)=\left(\sum_{i=1}^{n}x_i\right)\ln p+\left(n-\sum_{i=1}^{n}x_i\right)\ln(1-p).$$

由似然方程 $\frac{\mathrm{d}}{\mathrm{d}p}\ln L(p)=0$ 解出 $p=\bar{x}$. 因此，p 的极大似然估计量仍是 $\overline{X}$，且是 p 的无偏估计量.

分析 本题以解析形式写出 $B(1,p)$ 的分布律，这是求极大似然估计量的关键.

例 7-3 设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本，X 的密度函数为

$$f(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1;\\ 0, & \text{其余}.\end{cases}$$

其中 θ 未知. 试求 θ 的矩估计量与极大似然估计量.

解　由于

$$E(X)=\int_0^1 x\cdot(\theta+1)x^{\theta}\mathrm{d}x=\frac{\theta+1}{\theta+2}.$$

因此,从方程

$$\overline{X}=\frac{\hat{\theta}+1}{\hat{\theta}+2}$$

解得 θ 的矩估计量 $\hat{\theta}=\dfrac{2\overline{X}-1}{1-\overline{X}}$.

$0<x_i<1, i=1,\cdots,n$ 时,似然函数

$$L(\theta)=\prod_{i=1}^{n}[(\theta+1)x_i^{\theta}]=(\theta+1)^n\left(\prod_{i=1}^{n}x_i\right)^{\theta},$$

$$\ln L(\theta)=n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln x_i,$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=\frac{n}{\theta+1}+\sum_{i=1}^{n}\ln x_i.$$

由似然方程 $\dfrac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=0$ 解出 $\theta=-1-\dfrac{n}{\sum\limits_{i=1}^{n}\ln x_i}$. 因此,$\theta$ 的极大似然估计量是 $\hat{\theta}=-1-\dfrac{n}{\sum\limits_{i=1}^{n}\ln X_i}$.

分析　本题是 1997 年试题. 由于这是第一次出现数理统计试题,因此考查的要求较低,没有涉及无偏性之类的概念.

例 7-4　设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本,$X\sim P(\lambda)$,其中 λ 未知. 试证:

(1) 当常数 $c_1,\cdots,c_n$ 满足 $\sum\limits_{i=1}^{n}c_i=1$ 时,$\sum\limits_{i=1}^{n}c_iX_i$ 是 λ 的无偏估计量;

(2) 在 λ 的所有形如 $\sum\limits_{i=1}^{n}c_iX_i$ 的无偏估计量中,$\overline{X}$ 的方差最小(即 $\overline{X}$ 最有效),并求出最小值 $D(\overline{X})$.

证　(1) 由于 $\sum\limits_{i=1}^{n}c_i=1$, $E(X)=\lambda$,因此

$$E\left(\sum_{i=1}^{n}c_iX_i\right)=\sum_{i=1}^{n}c_iE(X_i)=\sum_{i=1}^{n}c_i\lambda=\lambda,$$

即 $\sum\limits_{i=1}^{n}c_iX_i$ 是 λ 的无偏估计量;

(2) 由于 $D(X)=\lambda$,因此

$$D\left(\sum_{i=1}^{n}c_iX_i\right)=\sum_{i=1}^{n}c_i^2D(X_i)=\sum_{i=1}^{n}c_i^2\lambda.$$

用高等数学中的拉格朗日乘数法可以得到:在约束条件 $\sum\limits_{i=1}^{n}c_i=1$ 下,n 元函数

$$f(c_1,\cdots,c_n)=\sum_{i=1}^{n}c_i^2$$

在 $c_1=\cdots=c_n=\dfrac{1}{n}$ 时达到最小值. 这表明 $\overline{X}=\dfrac{1}{n}X_1+\cdots+\dfrac{1}{n}X_n$ 的方差最小,且 $D(\overline{X})=\dfrac{\lambda}{n}$.

分析　比较无偏估计量的有效性等价于找出方差最小. 有时候,这类问题与高等数学中求最小值可以联系起来.

例 7-5　设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本,$X\sim R(0,\theta)$,其中 θ 未知,$\theta>0$. 试求 θ 的矩估计量与极大似然估计量,并讨论它们的无偏性与一致性.

解　由于

$$E(X)=\int_0^{\theta}x\cdot\frac{1}{\theta}\mathrm{d}x=\frac{\theta}{2}.$$

因此，从方程

$$\overline{X}=\frac{\hat{\theta}}{2}$$

解得 θ 的矩估计量 $\hat{\theta}=2\overline{X}$. 由 $E(\hat{\theta})=2E(\overline{X})=2\times\frac{\theta}{2}=\theta$ 知道，$\hat{\theta}$ 是 θ 的无偏估计量. 由

$$D(\hat{\theta})=4D(\overline{X})=4\ \frac{1}{n}D(X)=\frac{4}{n}\cdot\frac{(\theta-0)^2}{12}=\frac{\theta^2}{3n}\to 0$$

知道，$\hat{\theta}$ 是 θ 的一致估计量。

似然函数

$$L(\theta)=\begin{cases}\theta^{-n}, & 0\leqslant x_1\leqslant\cdots\leqslant x_n\leqslant\theta;\\ 0, & \text{其余.}\end{cases}$$

一方面 θ 越小，$L(\theta)$ 越大；另一方面，θ 不能太小，它要满足 $\theta\geqslant x_i$，$i=1,\cdots,n$. 因此，当 $\theta=\max\limits_{1\leqslant i\leqslant n}x_i$ 时，$L(\theta)$ 达最大值，即 θ 的极大似然估计量是 $\hat{\theta}^*=\max\limits_{1\leqslant i\leqslant n}X_i$.

当 $0<x<\theta$ 时，总体 X 的分布函数 $f(x)=\frac{x}{\theta}$. 由第三章给出的随机变量函数的分布知道，$\hat{\theta}^*=\max\limits_{1\leqslant i\leqslant n}X_i$ 的分布函数

$$F_{\hat{\theta}^*}(u)=\begin{cases}0, & u<0;\\ \frac{u^n}{\theta^n}, & 0\leqslant u<\theta;\\ 1, & u\geqslant\theta.\end{cases}$$

于是，$\hat{\theta}^*$ 的密度函数

$$f_{\hat{\theta}^*}(u)=\begin{cases}\frac{nu^{n-1}}{\theta^n}, & 0<u<\theta;\\ 0, & \text{其余.}\end{cases}$$

由此算得

$$E(\hat{\theta}^*)=\int_{-\infty}^{\infty}uf_{\hat{\theta}^*}(u)\mathrm{d}u=\int_0^{\theta}u\cdot\frac{nu^{n-1}}{\theta^n}\mathrm{d}u=\int_0^{\theta}\frac{n}{\theta^n}\cdot u^n\mathrm{d}u=\frac{n}{n+1}\theta\neq\theta,$$

这表明 θ 的极大似然估计 $\hat{\theta}^*$ 不是 θ 的无偏估计量. 另外，

$$D(\hat{\theta}^*)=E(\hat{\theta}^{*2})-[E(\hat{\theta}^*)]^2=\int_0^{\theta}u^2\ \frac{nu^{n-1}}{\theta^n}\mathrm{d}u-\left(\frac{n}{n+1}\theta\right)^2$$

$$=\frac{n}{n+2}\ \theta^2-\frac{n^2}{(n+1)^2}\theta^2=\frac{n}{(n+1)^2(n+2)}\theta^2.$$

现在来证明 $\hat{\theta}^*$ 是 θ 的一致估计量，即证明 $\hat{\theta}^*\xrightarrow{P}\theta$. 由 $E(\hat{\theta}^*)=\frac{n}{n+1}\theta$ 知道，$\frac{n+1}{n}\hat{\theta}^*$ 是 θ 的无偏估计量，且

$$D\left(\frac{n+1}{n}\hat{\theta}^*\right)=\frac{(n+1)^2}{n^2}D(\hat{\theta}^*)=\frac{(n+1)^2}{n^2}\cdot\frac{n\theta^2}{(n+1)^2(n+2)}=\frac{\theta^2}{n(n+2)}\to 0,$$

这表明 $\frac{n+1}{n}\hat{\theta}^*\xrightarrow{P}\theta$. 由 $\frac{n+1}{n}\to 1$ 推得 $\hat{\theta}^*\xrightarrow{P}\theta$.

分析　本题中极大似然估计量的部分有一定的难度，其方法也有一定的特殊性. 读者若感困难较大，不必纠缠于此.

例 7-6　从正态总体 $N(\mu,\sigma^2)$ 中获取四个数据 x_1,x_2,x_3,x_4，其中，μ 与 σ^2 均未知。已知

$$\sum_{i=1}^{4} x_i = 24, \quad \sum_{i=1}^{4} x_i^2 = 147.$$

在置信水平 90% 下，试求 μ 与 σ 的置信区间.

解　本题中 $n=4, \alpha=0.10$. 由例 6-1 所得公式算得

$$\bar{x} = \frac{1}{4} \times 24 = 6, \quad \sum_{i=1}^{4} (x_i - \bar{x})^2 = 147 - \frac{1}{4} \times 24^2 = 3,$$

$$S^2 = \frac{1}{3} \sum_{i=1}^{4} (x_i - \bar{x})^2 = 1, \quad S = 1.$$

于是，μ 的 90% 置信区间的上、下限分别为

$$\bar{x} \pm t_{1-\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}} = 6 \pm t_{0.95}(3) \frac{1}{\sqrt{4}} = 6 \pm \frac{1}{2} t_{0.95}(3) = 6 \pm \frac{1}{2} \times 2.3534.$$

由此算出 μ 的 90% 双侧置信区间为 $[4.8233, 7.1767]$. 另外，由 σ^2 的 90% 置信区间 $[3/\chi^2_{0.95}(3), \quad 3/\chi^2_{0.05}(3)]$ 推得 σ 的 90% 置信区间为

$$\left[\sqrt{\frac{3}{\chi^2_{0.95}(3)}}, \ \sqrt{\frac{3}{\chi^2_{0.05}(3)}}\right] = \left[\sqrt{\frac{3}{6.815}}, \ \sqrt{\frac{3}{0.352}}\right] = [0.66, \ 2.92]$$

分析　数理统计问题可以分成抽样前和抽样后两大类，本题属于后者，即样本通过具体数据来表达. 解这类题目的特点是用数据（抽样前）代公式作具体计算.

例 7-7　设 $X_1, \cdots, X_{2n}$ 是取自正态总体 $N(\mu_1, 18)$ 的一个样本，$Y_1, \cdots, Y_n$ 是取自正态总体 $N(\mu_2, 16)$ 的一个样本. 要使 $\mu_1 - \mu_2$ 的 95% 置信区间长度不超过 l，问 n 至少要取多大？

解　本题中 $m=2n, \alpha=0.05$. $\mu_1 - \mu_2$ 的 95% 置信区间的上、下限分别是

$$(\bar{X} - \bar{Y}) \pm u_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} = (\bar{X} - \bar{Y}) \pm u_{0.975} \sqrt{\frac{18}{2n} + \frac{16}{n}}$$

由置信区间的长度

$$\bar{\theta} - \underline{\theta} = 2u_{0.975} \sqrt{\frac{18}{2n} + \frac{16}{n}} = \frac{10}{\sqrt{n}} u_{0.975} \leqslant l$$

解得 $n \geqslant \frac{100}{l^2} u_{0.975}^2$（取最小整数）.

分析　置信区间的长度是一个很有意义的概念. 从直观上看，长度太大的置信区间是没有意义的.

例 7-8　设 $X_1, \cdots, X_n$ 是取自正态总体 $N(\mu, \sigma^2)$ 的一个样本，其中，μ 与 σ^2 均未知，给定置信水平 $1-\alpha$.

(1) 试求 $\underline{\mu}$，使得 $P(\mu \geqslant \underline{\mu}) = 1-\alpha$;

(2) 试求统计量 $\overline{\sigma^2}$，使得 $P(\sigma^2 \leqslant \overline{\sigma^2}) = 1-\alpha$.

解　仿照求置信区间的一般步骤求解.

(1) μ 的极大似然估计是 $\bar{X}$，构造随机变量

$$J = \sqrt{n} \frac{\bar{X} - \mu}{S} \sim t(n-1).$$

由

$$P(J \leqslant t_{1-\alpha}(n-1)) = P\left(\sqrt{n} \frac{\bar{X} - \mu}{S} \leqslant t_{1-\alpha}(n-1)\right) = 1-\alpha$$

得到

$$P\left(\mu \geqslant \bar{X} - t_{1-\alpha}(n-1) \frac{S}{\sqrt{n}}\right) = 1-\alpha,$$

即所求的 $\underline{\mu} = \bar{X} - t_{1-\alpha}(n-1) \frac{S}{\sqrt{n}}$.

(2) σ^2 的极大似然估计是 $S_n^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$. 构造随机变量：

$$J=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{nS_n^2}{\sigma^2}\sim\chi^2(n-1).$$

由

$$P(J\geqslant\chi_\alpha^2(n-1))=P\left(\frac{nS_n^2}{\sigma^2}\geqslant\chi_\alpha^2(n-1)\right)=1-\alpha$$

得到

$$P\left(\sigma^2\leqslant\frac{nS_n^2}{\chi_\alpha^2(n-1)}\right)=1-\alpha,$$

即所求的$\overline{\sigma^2}=nS_n^2/\chi_\alpha^2(n-1)$.

分析 本题是两种特殊类型的置信区间：$[\underline{\theta},+\infty)$与$(-\infty,\overline{\theta}]$. 但解题方法基本不变，仅在确定分位数时要注意区别.

例 7-9 设总体 X 的概率密度为

$$f(x)=\begin{cases}\frac{6x}{\theta^3}(\theta-x); & 0<x<\theta;\\ 0, & \text{其余}.\end{cases}$$

$X_1,X_2,\cdots,X_n$ 是取自总体 X 的简单随机样本.

(1) 求 θ 的矩估计量$\hat{\theta}$；(2) 求 $\hat{\theta}$ 的方差 $D(\hat{\theta})$.

解 (1) $E(X)=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x=\int_0^\theta\frac{6x^2}{\theta^3}(\theta-x)\mathrm{d}x=\frac{\theta}{2}$.

由此得到

$$\theta=2E(X).$$

所以 θ 的矩估计量为

$$\hat{\theta}=2\overline{X}.$$

(2) 由于

$$E(X^2)=\int_{-\infty}^{\infty}x^2f(x)\mathrm{d}x=\int_0^\theta\frac{6x^3}{\theta^3}(\theta-x)\mathrm{d}x=\frac{6\theta^2}{20}.$$

$$D(X)=E(X^2)-(EX)^2=\frac{6\theta^2}{20}-\left(\frac{\theta}{2}\right)^2=\frac{\theta^2}{20},$$

所以，$\hat{\theta}$ 的方差为

$$D(\hat{\theta})=D(2\overline{X})=4D(\overline{X})=\frac{4}{n}D(X)=\frac{\theta^2}{5n}.$$

分析 本题是 1999 年试题. 通常参数的矩估计量并不惟一. 如本题中 $\theta^2=\frac{10}{3}E(X^2)$，所以 $\hat{\theta}=\sqrt{\frac{10}{3}\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)}$也是 θ 的一个矩估计量. 一般情况下先考虑从 EX 的表达式建立 θ 的矩估计量.

例 7-10 设某种元件的使用寿命 X 的概率密度为

$$f(x;\theta)=\begin{cases}2\mathrm{e}^{-2(x-\theta)}, & x>\theta;\\ 0, & x\leqslant\theta.\end{cases}$$

其中，$\theta>0$ 为未知参数. 又设 $x_1,x_2,\cdots,x_n$ 是 X 的一组样本观测值，求参数 θ 的最大似然估计值.

解 似然函数为

$$L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\begin{cases}2^n\mathrm{e}^{-2\sum_{i=1}^{n}(x_i-\theta)}, & x_i>\theta,i=1,\cdots,n;\\ 0, & \text{其余}.\end{cases}$$

$$=\begin{cases}2^n \mathrm{e}^{-2\sum\limits_{i=1}^{n} x_i} \cdot \mathrm{e}^{2n\theta}, & x_{(1)}>\theta;\\ 0, & \text{其余}.\end{cases}$$

这里 $x_{(1)}=\min(x_1,\cdots,x_n)$.

要使 $L(\theta)$达到最大，应在条件“$x_{(1)}>\theta$”下使 θ 尽可能大. 因此，当 $\hat{\theta}=x_{(1)}$ 时 $L(\hat{\theta})=\sup\limits_{\theta} L(\theta)$. 即 $x_{(1)}$ 是 θ 的最大似然估计值.

分析　本题是 2000 年试题. 由于 $L(\theta)$在 $\theta=x_{(1)}$ 处不连续，因此不能对 $L(\theta)$求导. 本题中求 θ 的极大似然估计的方法与例 7-5 中的方法类似.

例 7-11　设总体 X 的概率分布为

X	0	1	2	3
p_r	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中，$\theta\left(0<\theta<\dfrac{1}{2}\right)$是未知参数，利用总体 X 的如下样本值：

$$3,1,3,0,3,1,2,3,$$

求 θ 的矩估计值和最大似然估计值.

解　由于 $EX=0\times\theta^2+1\times2\theta(1-\theta)+2\theta^2+3(1-2\theta)$

$=3-4\theta$,

因此，由 $\theta=\dfrac{1}{4}(3-EX)$得到 θ 的矩估计量为

$$\hat{\theta}=\frac{1}{4}(3-\overline{X})$$

θ 的矩估计值为

$$\hat{\theta}=\frac{1}{4}(3-\bar{x})=\frac{1}{4}\left[3-\frac{1}{8}(3+1+3+0+3+1+2+3)\right]=\frac{1}{4}.$$

再求 θ 的最大似然估计值. 先写出似然函数

$$\begin{aligned}L(\theta)&=P(X_1=3,X_2=1,X_3=3,X_4=0,X_5=3,X_6=1,X_7=2,X_8=3)\\&=P(X_1=3)P(X_2=1)P(X_3=3)P(X_4=0)P(X_5=3)P(X_6=1)\cdot P(X_7=2)P(X_8=3)\\&=(1-2\theta)^4\cdot(2\theta(1-\theta))^2\cdot(\theta^2)\cdot\theta^2=4\theta^6(1-\theta)^2(1-2\theta)^4.\end{aligned}$$

$$\ln L(\theta)=\ln4+6\ln\theta+2\ln(1-\theta)+4\ln(1-2\theta)$$

$$\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=\frac{6}{\theta}-\frac{2}{1-\theta}-\frac{8}{1-2\theta}=\frac{24\theta^2-28\theta+6}{\theta(1-\theta)(1-2\theta)}$$

令 $\dfrac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=0$，解得　　$\theta=\dfrac{7\pm\sqrt{13}}{12}$.

由于 $\theta=\dfrac{7+\sqrt{13}}{12}>\dfrac{1}{2}$，因此，取 $\hat{\theta}=\dfrac{7-\sqrt{13}}{12}$，即 θ 的极大似然估计值为$\dfrac{7-\sqrt{13}}{12}$.

分析　本题是 2002 年试题. 本题中求极大似然估计值部分很有新意. 本题的关键在于写出似然函数. 在写出似然函数过程中用到了独立性和同分布性，即样本为简单随机样本. $L(\theta)=P(X_1=3)P(X_2=1)P(X_3=3)P(X_4=0)\cdot P(X_5=3)P(X_6=1)\cdot P(X_7=2)P(X_8=3)=(P(X=3))^4\cdot(P(X=1))^2\cdot P(X=0)\cdot P(X=2)=(1-2\theta)^4(2\theta(1-\theta))^2\cdot\theta^2\cdot\theta^2$. 通过解本题读者对最大似然估计方法的原理一定有了进一步的了解.

例 7-12　设总体 X 的概率密度为

$$f(x)=\begin{cases}2\mathrm{e}^{-2(x-\theta)}, & x>\theta;\\ 0, & x\leqslant\theta.\end{cases}$$

其中，$\theta>0$ 是未知参数，从总体 X 中抽取简单随机样本 $X_1,X_2,\cdots,X_n$，记 $\hat{\theta}=\min(X_1,X_2,\cdots,X_n)$.

(1) 求总体 X 的分布函数 $F(x)$；

(2) 求统计量 $\hat{\theta}$ 的分布函数 $F_{\hat{\theta}}(x)$；

(3) 如果用 $\hat{\theta}$ 作为 θ 的估计量，讨论它是否具有无偏性.

解 (1) $x\leqslant\theta$ 时，$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\int_{-\infty}^{x}0\mathrm{d}t=0$；

$$x>\theta\text{ 时，}F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\int_{\theta}^{x}2\mathrm{e}^{-2(t-\theta)}\mathrm{d}t=1-\mathrm{e}^{-2(x-\theta)}.$$

即
$$F(x)=\begin{cases}1-\mathrm{e}^{-2(x-\theta)}, & x>\theta;\\ 0, & x\leqslant\theta.\end{cases}$$

(2) $F_{\hat{\theta}}(x)=P(\hat{\theta}\leqslant x)=P(X_{(1)}\leqslant x)$

$$=1-P(X_{(1)}>x)=1-P\{X_1>x,X_2>x,\cdots,X_n>x\}$$

$$=1-P\{X_1>x\}P\{X_2>x\}\cdots P\{X_n>x\}=1-(1-F(x))^n$$

$$=\begin{cases}1-\mathrm{e}^{-2n(x-\theta)}, & x>\theta;\\ 0, & \text{其余.}\end{cases}$$

(3) $\hat{\theta}$ 的概率密度为

$$f_{\hat{\theta}}(x)=F'_{\hat{\theta}}(x)=\begin{cases}2n\mathrm{e}^{-2n(x-\theta)}, & x>\theta;\\ 0, & x\leqslant\theta.\end{cases}$$

$$E\hat{\theta}=\int_{-\infty}^{\infty}xf_{\hat{\theta}}(x)\mathrm{d}x=\int_{\theta}^{+\infty}2nx\mathrm{e}^{-2n(x-\theta)}\mathrm{d}x=\theta+\frac{1}{2n}\neq0.$$

因此，$\hat{\theta}$ 不是 θ 的无偏估计量. 即 $\hat{\theta}$ 作为 θ 的估计量不具有无偏性.

分析 本题是 2003 年试题. 由例 7-10 可知 $\hat{\theta}=X_{(1)}$ 是 θ 的最大似然估计量. 本题表明：θ 的最大似然估计量不是 θ 的无偏估计量. 但 $\hat{\theta}$ 经修正后可以成为 θ 的无偏估计量. 只需记 $\hat{\theta}^*=X_{(1)}-\frac{1}{2n}$，则 $\hat{\theta}^*$ 是 θ 的无偏估计量. 一般地，若 $E\hat{\theta}=a\theta+b(a\neq0,a,b\text{ 为常数})$，则 $\hat{\theta}^*=\frac{\hat{\theta}-b}{a}$ 是 θ 的无偏估计量.

例 7-13 假设 0.50，1.25，0.80，2.00 是来自总体 X 的简单随机样本值. 已知 $Y=\ln X$ 服从正态分布 $N(\mu,1)$.

(1) 求 X 的数学期望 EX(记 EX 为 b)；

(2) 求 μ 的置信度为 0.95 的置信区间；

(3) 利用上述结果求 b 的置信度为 0.95 的置信区间.

解 由于 $Y=\ln X$，因此，ln0.50，ln1.25，ln0.8，ln2.00 可视为取自总体 Y 的样本值，而 $r\sim N(\mu,1)$，从而 $\bar{y}=\frac{1}{4}(\ln0.50+\ln1.25+\ln0.80+\ln2.00)=0$

(1) $EX=E\mathrm{e}^{Y}=\int_{-\infty}^{\infty}\mathrm{e}^{y}f_Y(y)\mathrm{d}y=\int_{-\infty}^{\infty}\mathrm{e}^{y}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(y-\mu)^2}{2}}\mathrm{d}y$

$$=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2-(2+2\mu)y+\mu^2}{2}}=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(y-\mu-1)^2}{2}}\mathrm{e}^{\frac{1+2\mu}{2}}\mathrm{d}y$$

$$=\mathrm{e}^{\frac{1}{2}+\mu}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(y-\mu-1)^2}{2}}\mathrm{d}y=\mathrm{e}^{\frac{1}{2}+\mu}.$$

这里用到 $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(y-\mu-1)^2}{2}}$ 是对应于正态分布 $N(\mu+1,1)$ 的密度函数，所以它在 $(-\infty,+\infty)$ 上积分为 1.

(2) 由于 $Y\sim N(\mu,1)$,故 μ 的置信度为 0.95 的置信区间的上、下限为

$$\bar{y}\pm\frac{1}{\sqrt{n}}u_{0.975}=0\pm\frac{1}{\sqrt{4}}\times1.96=\pm0.98,$$

所以 μ 的置信区间为$[-0.98,0.98]$.

(3) 由于 $e^{\frac{1}{2}+\mu}$ 的严格单增函数,所以

$$P\left(\bar{Y}-\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}\leqslant\mu\leqslant\bar{Y}+\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}\right)=P\left(e^{\frac{1}{2}+\bar{Y}-\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}}\leqslant e^{\frac{1}{2}+\mu}\leqslant e^{\frac{1}{2}+\bar{Y}+\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}}\right)=1-\alpha.$$

从而 $b=e^{\frac{1}{2}+\mu}$ 的置信度为 0.95 的置信区间为

$$[e^{\frac{1}{2}-0.98},e^{\frac{1}{2}+0.98}]=[e^{-0.48},e^{1.48}]=[0.6188,4.3929].$$

分析　本题是 2000 年试题. 本题涉及两个总体间的变换.

四、小　结

本章是数理统计部分的重点,读者尤其要熟练地掌握求点估计的两种方法以及讨论所求出的点估计的无偏性. 对正态总体参数的置信区间的求法也应熟练地掌握.

第八章　假设检验

假设检验就是利用样本来检验关于总体分布的某个假设.假设有两种形式:涉及总体分布中的未知参数与未知的总体分布本身.读者应主要掌握正态总体分布中的未知参数的假设检验方法.

一、复习与考试要求

(1) 理解显著性检验的基本思想,掌握假设检验的基本步骤,了解假设检验可能产生的两类错误.

(2) 了解单个及两个正态总体的均值和方差的假设检验.

二、基本概念与理论

1. 假设检验的统计意义

样本$(X_1,\cdots,X_n)$的取值范围称为样本空间.对于取自正态总体的样本,样本空间是 n 维欧氏空间 R^n.

原假设(或零假设)H_0 是针对未知总体分布的某个陈述.它可以是关于总体分布中的未知参数的某个陈述,例如 $\theta=\theta_0$;也可以是关于未知的总体分布本身的某个陈述,例如总体 X 服从正态分布.备择假设(或对立假设)H_1 常常是原假设 H_0 的否定命题.

求解一个假设检验问题,其结果分成两种类型.

(1) 抽样前　求出一个拒绝域 W_1.它是样本空间的一个子集,通常用一个或两个不等式来表示.拒绝域 W_1 的含义是:当$(x_1,\cdots,x_n)\in W_1$ 时拒绝 H_0(即认为命题 H_0 不成立);当$(x_1,\cdots,x_n)\notin W_1$ 时接受 H_0(即可以认为命题 H_0 成立).

(2) 抽样后　在求出拒绝域 W_1 的基础上,用数据检查样本是否落在 W_1 中,以得出最终结论究竟是"拒绝 H_0"还是"接受 H_0".

2. 假设检验的两类错误

利用样本 $X_1,\cdots,X_n$ 检验原假设 H_0 是否成立最终会面临犯两类错误的风险,具体内容见表 8-1:

表 8-1

检验带来的后果		根据样本观测值所得的结论	
		当$(x_1,\cdots,x_n)\notin W_1$ 接受 H_0	当$(x_1,\cdots,x_n)\in W_1$ 拒绝 H_0
总体分布的实际情况(未知)	H_0 成立	判断正确	犯第Ⅰ类错误
	H_0 不成立	犯第Ⅱ类错误	判断正确

3. 显著性检验的统计思想

显著性检验要求控制犯第Ⅰ类错误的概率不超过显著性水平 α,即要求所求出的拒绝域 W_1 满足:

$$P((X_1,\cdots,X_n)\in W_1)\leqslant\alpha,\quad \text{当 } H_0 \text{ 成立时.}$$

"拒绝 H_0"也称为"有显著性结果";"接受 H_0"也称为"无显著性结果".

在显著性水平 α 下求拒绝域 W_1 的一般步骤如下:

(1) 求出 H_0 所涉及的未知参数 θ 的一个较优的点估计 $\hat{\theta}$.我们建议尽可能使用 θ 的极大似然估计;

(2) 以 $\hat{\theta}$ 为基础,构造一个检验统计量

$$T=t(X_1,\cdots,X_n).$$

要求当 H_0 成立$(\theta=\theta_0)$时 T 的分布已知,从而 T 的分位数(称为临界值)能通过查表或计算得到具体数值;

(3) 利用检验统计量 T 与临界值构造一个或两个不等式，以确定拒绝域 W_1 的形状，W_1 应满足显著性水平 α 的要求.

4. 单个正态总体

设 $X_1,\cdots,X_n$ 是取自正态总体 $N(\mu,\sigma^2)$ 的一个样本；取显著性水平为 α.

(1) $H_0:\mu=\mu_0$　($H_1:\mu\neq\mu_0$)

当 σ^2 已知时，拒绝域

$$W_1=\left\{(x_1,\cdots,x_n)\,\middle|\,|u|=\sqrt{n}\,\frac{|\bar{x}-\mu_0|}{\sigma}>u_{1-\frac{\alpha}{2}}\right\},$$

称它为 u 检验法；

当 σ^2 未知时，拒绝域

$$W_1=\left\{(x_1,\cdots,x_n)\,\middle|\,|t|=\sqrt{n}\,\frac{|\bar{x}-\mu_0|}{S}>t_{1-\frac{\alpha}{2}}(n-1)\right\},$$

称它为 t 检验法；

(2) $H_0:\sigma^2=\sigma_0^2$　($H_1:\sigma^2\neq\sigma_0^2$)

当 μ 已知时，拒绝域

$$W_1=\{(x_1,\cdots,x_n)\,|\,\chi^2<\chi^2_{\frac{\alpha}{2}}(n)\text{ 或 }\chi^2>\chi^2_{1-\frac{\alpha}{2}}(n)\},$$

其中 $\chi^2=\dfrac{1}{\sigma_0^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2$；

当 μ 未知时，拒绝域

$$W_1=\{(x_1,\cdots,x_n)\,|\,\chi^2<\chi^2_{\frac{\alpha}{2}}(n-1)\text{ 或 }\chi^2>\chi^2_{1-\frac{\alpha}{2}}(n-1)\},$$

其中 $\chi^2=\dfrac{1}{\sigma_0^2}\sum\limits_{i=1}^{n}(x_i-\bar{x})^2=\dfrac{ns_n^2}{\sigma_0^2}=\dfrac{(n-1)s^2}{\sigma_0^2}$. 这两个检验都称为 χ^2 检验法.

5. 两个正态总体

设 $X_1,\cdots,X_m$ 是取自正态总体 $N(\mu_1,\sigma_1^2)$ 的一个样本，$Y_1\cdots,Y_n$ 是取自正态总体 $N(\mu_2,\sigma_2^2)$ 的一个样本；取显著性水平为 α.

(1) $H_0:\mu_1=\mu_2$　($H_1:\mu_1\neq\mu_2$)

当 σ_1^2,σ_2^2 已知时，拒绝域

$$W_1=\{(x_1,\cdots,x_m;y_1,\cdots,y_n)\,|\,|u|=\frac{|\bar{x}-\bar{y}|}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}>u_{1-\frac{\alpha}{2}}\},$$

称它为 u 检验法；

当 σ_1^2,σ_2^2 未知，但 $\sigma_1^2=\sigma_2^2$ 时，拒绝域

$$W_1=\{(x_1,\cdots,x_m;y_1,\cdots,y_n)\,|\,|t|=\frac{|\bar{x}+\bar{y}|}{S_W\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}>t_{1-\frac{\alpha}{2}}(m+n-2)\},$$

称它为 t 检验法.

(2) $H_0:\sigma_1^2=\sigma_2^2$ ($H_1:\sigma_1^2\neq\sigma_2^2$)

当 μ_1,μ_2 已知时，拒绝域

$$W_1=\{(x_1,\cdots,x_m;y_1,\cdots,y_n)\,|\,f<F_{\frac{\alpha}{2}}(m,n)\text{ 或 }f>F_{\frac{\alpha}{2}}(m,n)\},$$

其中 $f=n\sum\limits_{i=1}^{m}(x_i-\mu_1)^2/m\sum\limits_{i=1}^{n}(y_i-\mu_2)^2$；

当 μ_1,μ_2 均未知时，拒绝域

$$W_1=\{(x_1,\cdots,x_m;y_1,\cdots,y_n)\,|\,f<F_{\frac{\alpha}{2}}(m-1,n-1)\text{ 或 }f>F_{\frac{\alpha}{2}}(m-1,n-1)\},$$

其中，$f=s_1^2/s_2^2$，这两个检验都称为 F 检验法.

三、例题分析

例 8-1 在假设检验问题中，在样本观测值的基础上试判断：

(1) 如果最终结果是拒绝原假设 H_0，可能犯何类错误？

(2) 如果最终结果是接受原假设 H_0，可能犯何类错误？

解 (1) 当实际情况为 H_0 成立时，要犯第 Ⅰ 类错误；当实际情况为 H_0 不成立时，判断正确，即不犯错误.

(2) 当实际情况为 H_0 成立时，判断正确，即不犯错误；当实际情况为 H_0 不成立时，要犯第 Ⅱ 类错误.

分析 显著性检验仅仅控制了犯第 Ⅰ 类错误的概率不超过显著性水平 α，因而检验结果为“拒绝 H_0”时比较可靠. 当检验结果为“接受 H_0”时，由于不知道犯第 Ⅱ 类错误的概率究竟有多大，因此结论很不可信.

例 8-2 从两个正态总体 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$ 中分别抽取样本 $(X_1,\cdots,X_m)$，$(Y_1,\cdots,Y_n)$，其中，μ_1，μ_2，σ_1^2，σ_2^2 均未知. 假定 $\sigma_1^2=\sigma_2^2$. 在显著性水平 α 下，要检验

$$H_0:\mu_1=\mu_2+\delta \quad (H_1:\mu_1\neq\mu_2+\delta),$$

其中，δ 是已知常数. 试求拒绝域 W_1.

解 记 $\theta=\mu_1-\mu_2$. 现在要检验

$$H_0:\theta=\delta \quad (H_1:\theta\neq\delta).$$

θ 的极大似然估计是 $\overline{X}-\overline{Y}$. 构造检验统计量

$$T=\frac{(\overline{X}-\overline{Y})-\delta}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}}$$

当 H_0 成立时，$T\sim t(m+n-2)$. 因此拒绝域

$$W_1=\left\{(x_1,\cdots,x_m;y_1,\cdots,y_n)\,\middle|\,|t|=\frac{|(\overline{x}-\overline{y})-\delta|}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}}>t_{1-\frac{\alpha}{2}}(m+n-2)\right\}.$$

分析 本题中取 $\delta=0$ 便得检验 $H_0:\mu_1=\mu_2$ 的拒绝域. 类似地，在取消本题关于“等方差”的假定时，在显著性水平 α 下，检验

$$H_0:\sigma_1^2=\delta\sigma_2^2 \quad (H_1:\sigma_1^2\neq\delta\sigma_2^2)$$

的拒绝域为

$$W_1=\{(x_1,\cdots,x_m;y_1,\cdots,y_n)\,|\,f<F_{\frac{\alpha}{2}}(m-1,n-1)$$

$$\text{或 } f>F_{1-\frac{\alpha}{2}}(m-1,n-1)\},$$

其中，$f=\delta\dfrac{s_1^2}{s_2^2}$.

例 8-3 设 $X_1,\cdots,X_n$ 是取自正态总体 $N(\mu,\sigma^2)$ 的一个样本，其中，μ 与 σ^2 均未知，在显著性水平 α 下，试求下列两个假设检验问题的拒绝域.

(1) $H_0:\mu=\mu_0 \quad (H_1:\mu>\mu_0)$；

(2) $H_0:\sigma^2=\sigma_0^2 \quad (H_1:\sigma^2<\sigma_0^2)$.

解 仿照求显著性检验的拒绝域的一般步骤求解.

(1) μ 的极大似然估计是 $\overline{X}$. 构造检验统计量

$$T=\sqrt{n}\,\frac{\overline{X}-\mu_0}{S}$$

当 H_0 成立时，$T\sim t(n-1)$，由

$$P(T>t_{1-\alpha}(n-1))=\alpha$$

推得拒绝域

$$W_1=\left\{(x_1,\cdots,x_n)\,|\,t=\sqrt{n}\,\frac{\bar{x}-\mu_0}{s}>t_{1-\alpha}(n-1)\right\};$$

(2) σ^2 的极大似然估计是 $S_n^2=\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2$. 构造检验统计量

$$\chi^2=\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\bar{X})^2.$$

当 H_0 成立时，$\chi^2\sim\chi^2(n-1)$. 由

$$P(\chi^2<\chi_\alpha^2(n-1))=\alpha$$

推得拒绝域

$$W_1=\{(x_1,\cdots,x_n)\,|\,\chi^2<\chi_\alpha^2(n-1)\},$$

其中 $\chi^2=\frac{1}{\sigma_0^2}\sum\limits_{i=1}^{n}(x_i-\bar{x})^2=\frac{ns_n^2}{\sigma_0^2}=\frac{(n-1)s^2}{\sigma_0^2}$.

分析　要注意本题中拒绝域的形状(从而临界值)已有改变，这是因为备择假设 H_1 有了变动。另外，如果本题中原假设 H_0 改成“$\mu\leqslant\mu_0,\sigma^2\geqslant\sigma_0^2$”，那么拒绝域仍保持不变.

例 8-4　设 $X_1,\cdots,X_n$ 是取自正态总体 $N(\mu,4)$ 的一个样本. 在显著性水平 α 下检验

$$H_0:\mu=0\quad(H_1:\mu\neq0),$$

我们取拒绝域 $W_1=\left\{(x_1,\cdots,x_n)\,|\sqrt{n}\,\frac{|\bar{x}|}{2}>u_{1-\frac{\alpha}{2}}\right\}$. 当实际情况为 $\mu=1$(即总体 $X\sim N(1,4)$)时，试求犯第 Ⅱ 类错误的概率.

解　当实际情况为 H_0 不成立，而最终结果为接受 H_0 时要犯第 Ⅱ 类错误，其概率为 $P((X_1,\cdots,X_n)\notin W_1)$. 现在实际情况 $\mu=1$，表示 H_0 不成立. 当 $X\sim N(1,4)$ 时，$\bar{X}\sim N\left(1,\frac{4}{n}\right)$. 由正态分布的概率计算公式得到，犯第 Ⅱ 类错误的概率为

$$P((X_1,\cdots,X_n)\notin W_1)=P\left(\sqrt{n}\,\frac{|\bar{X}|}{2}\leqslant u_{1-\frac{\alpha}{2}}\right)=P\left(-\frac{2u_{1-\frac{\alpha}{2}}}{\sqrt{n}}\leqslant\bar{X}\leqslant\frac{2u_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right)$$

$$=\Phi\left(\frac{\frac{2u_{1-\frac{\alpha}{2}}}{\sqrt{n}}-1}{2/\sqrt{n}}\right)-\Phi\left(\frac{-\frac{2u_{1-\frac{\alpha}{2}}}{\sqrt{n}}-1}{2/\sqrt{n}}\right)=\Phi\left(u_{1-\frac{\alpha}{2}}-\frac{\sqrt{n}}{2}\right)-\Phi\left(-u_{1-\frac{\alpha}{2}}-\frac{\sqrt{n}}{2}\right)$$

$$=\left[1-\Phi\left(\frac{\sqrt{n}}{2}-u_{1-\frac{\alpha}{2}}\right)\right]-\left[1-\Phi\left(\frac{\sqrt{n}}{2}+u_{1-\frac{\alpha}{2}}\right)\right]=\Phi\left(\frac{\sqrt{n}}{2}+u_{1-\frac{\alpha}{2}}\right)-\Phi\left(\frac{\sqrt{n}}{2}-u_{1-\frac{\alpha}{2}}\right).$$

分析　犯第 Ⅱ 类错误的概率一般不易计算. 本题目的是让读者通过一个具体例子来加深对显著性检验的理解，因为在显著性检验中我们没有考虑犯第 Ⅱ 类错误的概率. 例如，在本题中，取 $n=4,\alpha=0.05$. 通过查表知道，$u_{0.975}=1.96$，犯第 Ⅱ 类错误的概率为

$$\Phi(1+1.96)-\Phi(1-1.96)=\Phi(2.96)-\Phi(-0.96)=0.9985-0.1685=0.83.$$

这表明尽管犯第 Ⅰ 类错误的概率不超过 5%，但犯第 Ⅱ 类错误的概率高达 83%.

例 8-5　设某次考试的考生成绩服从正态分布，从中随机地抽取 36 位考生的成绩，算得平均成绩为 66.5 分，标准差为 15 分. 问在显著性水平 0.05 下，是否可以认为这次考试全体考生的平均成绩为 70 分？并给出检验过程.

解　设这次考试考生的成绩为 X，$X\sim N(\mu,\sigma^2)$. 其中，μ,σ^2 均未知. 现要在显著性水平 0.05 下，检验假设：

$$H_0:\mu=70,\ (H_1:\mu\neq 70).$$

$$|t|=\sqrt{n}\frac{|\bar{x}-70|}{s}=\sqrt{36}\frac{|66.5-70|}{15}=1.4,$$

$$t_{1-\frac{\alpha}{2}}(n-1)=t_{0.975}(35)=2.0301.$$

由于$|t|=1.4<t_{1-\frac{\alpha}{2}}(n-1)=2.0301$,因此,拒绝域条件不成立.从而,在显著性水平 0.05 下,可以认为这次考试考生的平均成绩为 70 分.

分析　本题是 1998 年试题.考生只要能记住拒绝域条件,会代公式来判断拒绝域条件是否成立即可.

四、小　结

假设检验虽然也是数理统计的一个重要内容,但在本科阶段的教学中要求较弱.考生除了记住各种拒绝域的公式外,只需对假设检验的基本概念有一般性的了解.

第九章　综合能力训练

随着概率统计知识的普及，现在对考生综合运用概率统计知识的能力的要求越来越高. 综合运用概率统计知识的能力主要体现在以下几个方面：

(1) 综合运用概率论知识的能力；

(2) 综合运用统计知识的能力；

(3) 综合运用概率论和数理统计知识的能力；

(4) 综合运用概率统计、线性代数和高等数学的能力.

本章结合一些例题来说明如何提高综合能力，希望对读者有所启发和帮助.

一、例题分析

例 9-1　已知随机变量 X 服从参数为 1 的泊松分布，并记事件 $A=\{X\geqslant 2\}$，$B=\{X<1\}$. 求 $P(A\cup B)$，$P(A-B)$，$P(B|\overline{A})$.

解　由于 $X\sim P(1)$，因此

$$P(A)=P(X\geqslant 2)=1-P(X=0)-P(X=1)=1-\mathrm{e}^{-1}-\mathrm{e}^{-1}=1-2\mathrm{e}^{-1}\approx 0.26,$$

$$P(B)=P(X<1)=P(X=0)=\mathrm{e}^{-1}\approx 0.37.$$

又由于 $AB=\phi$，即 A，B 互不相容，因而

$$P(A\cup B)=P(A)+P(B)=0.26+0.37=0.63,$$

$$P(A-B)=P(A)-P(AB)=P(A)=0.26,$$

$$P(B|\overline{A})=\frac{P(\overline{A}B)}{P(\overline{A})}=\frac{P(X\leqslant 1,X<1)}{1-P(A)}=\frac{P(X<1)}{1-P(A)}=\frac{\mathrm{e}^{-1}}{1-(1-2\mathrm{e}^{-1})}=\frac{1}{2}.$$

分析　本题主要考虑综合运用概率论知识的能力. 通过随机变量 X 定义的两个随机事件 A，B 的概率要考生利用 X 的分布来求出，同时要判断出 A，B 互不相容和 $\overline{A}\supset B$. 例 2-1 与本题有些相似之处，可谓异曲同工.

例 9-2　设某班车起点站上客人数 X 服从参数为 $\lambda(\lambda>0)$ 的泊松分布，每位乘客在中途下车的概率为 $p(0<p<1)$，且中途下车与否相互独立. 以 Y 表示在中途下车的人数，求：

(1) 在发车时有 n 个乘客的条件下，中途有 m 人下车的概率；

(2) 二维随机变量 (X,Y) 的概率分布.

解　(1) 记 $Z_i=\begin{cases}1, & \text{第 } i \text{ 位乘客中途下车;}\\ 0, & \text{第 } i \text{ 位乘客中途不下车.}\end{cases}$

则　$Y=Z_1+Z_2+\cdots+Z_X$. 由题意，$Z_1,Z_2,\cdots$ 相互独立且服从相同的分布，$Z_i\sim B(1,p)$，$i=1,2\cdots$

$$P(Y=m|X=n)=P(Z_1+Z_2+\cdots+Z_X=m|X=n)=P(Z_1+Z_2+\cdots+Z_n=m|X=n)$$

$$=P(Z_1+Z_2+\cdots+Z_n=m).$$

由二项分布的可加性定理得到 $Z_1+Z_2+\cdots+Z_n\sim B(n,p)$. 因此

$$P(Y=m|X=n)=C_n^m p^m(1-p)^{n-m},m=0,1,2,\cdots,n.$$

(2) 对于 $m=0,1,2,\cdots,n$，有

$$P(X=n,Y=m)=P(Y=m|X=n)P(X=n)=C_n^m p^m(1-p)^{n-m}\cdot\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}$$

分析　本题是 2001 年试题. 题目中“中途下车与否相互独立”的含义比较模糊，事实上是指各位乘客中

途下车与否相互独立且各位乘客中途下车与否与起点站上客人数 X 相互独立，即 $Z_1, Z_2, \cdots$ 相互独立，且 $X, Z_1, Z_2, \cdots$，相互独立. 在解题过程中 $P(Z_1+\cdots+Z_n=m \mid X=n)=P(Z_1+\cdots+Z_n=m)$，用到了“$Z_1+\cdots Z_n$ 与 X 相互独立”. 生活经验表明：若起点站上客人数太多，有些人不得不考虑中途下车. 但遗憾的是题目中没有写清楚，而用了比较模糊的说法. 考生对此可以不加理会，毕竟本题只是考查考生综合运用概率论知识的能力.

例 9-3 已知甲、乙两箱中装有同种产品，其中，甲箱中装有 3 件合格品和 3 件次品，乙箱中仅装有 3 件合格品. 从甲箱中任取 3 件产品放入乙箱后，求：

(1) 乙箱中次品件数 X 的数学期望；

(2) 从乙箱中任取一件产品是次品的概率.

解 (1) 记

$$X_i=\begin{cases}1, & \text{从甲箱中取出的第 } i \text{ 件产品是次品；}\\ 0, & \text{从甲箱中取出的第 } i \text{ 件产品是合格品.}\end{cases}, i=1,2,3$$

由例 1-12 可知，X_1, X_2, X_3 服从相同的分布. X_1 的概率分布律为

X_1	0	1
概率	$\frac{1}{2}$	$\frac{1}{2}$

从甲箱中任取 3 件产品放入乙箱后，乙箱中次品数为

$$X=X_1+X_2+X_3$$

$$EX=EX_1+EX_2+EX_3=\frac{1}{2}+\frac{1}{2}+\frac{1}{2}=\frac{3}{2}.$$

(2) 设 A 表示事件“从乙箱中任取一件产品是次品”，由于 $\{X=0\}, \{X=1\}, \{X=2\}$ 和 $\{X=3\}$ 构成完全事件组，因此，由全概率公式

$$P(A)=\sum_{K=0}^{3}P(A \mid X=k)P(X=k)=\sum_{k=0}^{3}\frac{k}{6}P(X=k)=\frac{1}{6}\sum_{k=0}^{3}kP(X=k)$$

$$=\frac{1}{6}EX=\frac{1}{6}\times\frac{3}{2}=\frac{1}{4}.$$

分析 本题是 2003 年试题. 在 $P(A)$ 的计算中巧妙地利用了(1)中已求得的 EX. 当然，本题也可以用常规的方法求解. 先确定 X 的可能取值为 0,1,2,3. 求出 X 的分布律为

$$P(X=k)=\frac{C_3^kC_3^{3-k}}{C_6^3}, k=0,1,2,3.$$

即算出分布律

X	0	1	2	3
P_r	$\frac{1}{20}$	$\frac{9}{20}$	$\frac{9}{20}$	$\frac{1}{20}$

因此

$$EX=0\times\frac{1}{20}+1\times\frac{9}{20}+2\times\frac{9}{20}+3\times\frac{1}{20}=\frac{3}{2}.$$

$$P(A)=\sum_{k=0}^{3}P(X=k)P(A \mid X=k)=\frac{1}{20}\times0+\frac{9}{20}\times\frac{1}{6}+\frac{9}{20}\times\frac{2}{6}+\frac{1}{20}\times\frac{3}{6}=\frac{1}{4}.$$

例 9-4 假设随机变量 X_1, X_2, X_3, X_4 相互独立且同分布 $P(X_i=0)=0.6, P(X_i=1)=0.4, i=1,2,3,4$. 求行列式 $X=\begin{vmatrix}X_1 & X_2\\ X_3 & X_4\end{vmatrix}$ 的概率分布.

解　记 $Y_1=X_1X_4$，$Y_2=X_2X_3$，则

$$X=Y_1-Y_2.$$

由于 X_1,X_2,X_3,X_4 独立同分布，因此，Y_1 与 Y_2 也独立同分布．Y_1 可能取值为 0，1.

$$P(Y_1=1)=P(X_1=1,X_4=1)=P(X_1=1)P(X_4=1)=0.4\times0.4=0.16,$$

$$P(Y_1=0)=1-P(Y_1=1)=1-0.16=0.84.$$

$X=Y_1-Y_2$ 可能取值为 0，1，-1

$$P(X=-1)=P(Y_1=0,Y_2=1)=P(Y_1=0)P(Y_2=1)=0.84\times0.16=0.1344,$$

$$P(X=1)=P(Y_1=1,Y_2=0)=P(Y_1=1)P(Y_2=0)=0.16\times0.84=0.1344.$$

$$P(X=0)=1-P(X=-1)-P(X=1)=1-2\times0.1344=0.7312.$$

于是 X 的概率分布律为

X	-1	0	1
$p_{r.}$	0.1344	0.7312	0.1344

分析　本题是 1994 年试题．本题要求考生能综合运用概率统计和线性代数的知识．其中 $Y_1=X_1X_4$ 与 $Y_2=X_2X_3$ 独立同分布是解题的关键．

例 9-5　设随机变量 $X_{ij}(i,j=1,2,\cdots,n,n\geqslant2)$ 独立同分布，$EX_{ij}=2$，求行列式

$$Y=\begin{vmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{vmatrix}$$

的数学期望 EY.

解　由 n 阶行列式的定义

$$Y=\sum_{(i_1,\cdots,i_n)\in S_n}(-1)^{\tau(i_1\cdots i_n)}X_{1i_1}X_{2i_2}\cdot\cdots\cdot X_{ni_n},$$ 其中 $S_n=\{$所有 n 元排列$\}$.

因此

$$\begin{aligned} EY&=\sum_{(i_1,\cdots,i_n)\in S_n}(-1)^{\tau(i_1\cdots i_n)}E(X_{1i_1}X_{2i_2}\cdot\cdots\cdot X_{ni_n}) \\ &=\sum_{(i_1,\cdots,i_n)\in S_n}(-1)^{\tau(i_1\cdots i_n)}EX_{1i_1}EX_{2i_2}\cdot\cdots\cdot EX_{ni_n} \\ &=\begin{vmatrix} EX_{11} & EX_{12} & \cdots & EX_{1n} \\ EX_{21} & EX_{22} & \cdots & EX_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ EX_{n1} & EX_{n2} & \cdots & EX_{nn} \end{vmatrix}=\begin{vmatrix} 2 & 2 & \cdots & 2 \\ 2 & 2 & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 2 & 2 & \cdots & 2 \end{vmatrix}=0. \end{aligned}$$

分析　本题是 1999 年考题．本题考查考生综合运用线性代数和概率论知识的能力．其中，$E(X_{i_1}\cdot\cdots\cdot X_{i_n})=EX_{i_1}\cdot\cdots\cdot EX_{i_n}$ 用到 $X_{i_1},X_{i_2},\cdots,X_{i_n}$ 相互独立．

例 9-6　设随机变量 X 在区间 $[-1,2]$ 上服从均匀分布；随机变量

$$Y=\begin{cases} 1, & \text{若 } X>0; \\ 0, & \text{若 } X=0; \\ -1, & \text{若 } X<0. \end{cases}$$

求 Y 的方差 $D(Y)$.

解　先求出 Y 的分布律．注意到 $\{Y=1\}=\{X>0\}$ 而 $X\sim R[-1,2]$，X 的密度函数 $f(x)=$

$\begin{cases}\frac{1}{3}, & -1\leqslant x\leqslant 2\\ 0, & \text{其余.}\end{cases}$ 因此 $P(Y=1)=P(X>0)=\int_0^{+\infty} f(x)\mathrm{d}x=\int_0^2 \frac{1}{3}\mathrm{d}x=\frac{2}{3}$,

$$P(Y=0)=P(X=0)=0,$$

$$P(Y=-1)=P(X<0)=\int_{-1}^{0}\frac{1}{3}\mathrm{d}x=\frac{1}{3}.$$

由此得到 Y 的分布律为

Y	-1	1
$p_{r.}$	$\frac{1}{3}$	$\frac{2}{3}$

$$EY=(-1)\times\frac{1}{3}+1\times\frac{2}{3}=\frac{1}{3},$$

$$EY^2=(-1)^2\times\frac{1}{3}+1^2\times\frac{2}{3}=1.$$

从而 $$D(Y)=E(Y^2)-(EY)^2=1-\left(\frac{1}{3}\right)^2=\frac{8}{9}.$$

分析 本题是 2000 年试题. 本题中 X 是连续型随机变量,Y 是离散型随机变量. Y 由 X 所定义,因此 Y 的分布律可由 X 的分布求得. 本题仍是考查考生综合运用概率论知识的能力.

例 9-7 假设测量的随机误差 $X\sim N(0,10^2)$,试求在 100 次独立重复测量中,至少有三次测量误差的绝对值大于 19.6 的概率 α,并利用泊松分布求出 α 的近似值.(要求小数点后取两位有效数字).

附表

λ	1	2	3	4	5	6	…
$\mathrm{e}^{-\lambda}$	0.368	0.135	0.050	0.018	0.007	0.002	…

解 先求出

$$p=P(|X|>19.6)=P\left(\frac{|X|}{10}>1.96\right)=2(1-\Phi(1.96))=2\times 0.025=0.05.$$

记 Y 表示在 100 次重复测量中测量误差的绝对值超过 19.6 的次数,则 $Y\sim B(100,p)$.

因此

$$\alpha=P(Y\geqslant 3)=1-P(Y=0)-P(r=1)-P(r=2)$$

$$=1-0.95^{100}-100\times 0.05\times 0.95^{99}-\frac{100\times 99}{2}\times 0.05^2\times 0.95^{98}$$

由泊松定理,Y 近似服从参数为 $\lambda=np=100\times 0.05=5$ 的泊松分布

$$\alpha=1-P(Y=0)-P(Y=1)-P(Y=2)$$

$$\approx 1-\mathrm{e}^{-\lambda}-\lambda\mathrm{e}^{-\lambda}-\frac{\lambda^2}{2}\mathrm{e}^{-\lambda}=1-\mathrm{e}^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2}\right)$$

$$=1-\mathrm{e}^{-5}(1+5+12.5)=1-0.007\times 18.5\approx 0.87.$$

分析 本题是 1992 年试题. 在本题中出现了三个常用分布. 且二项分布中的参数 p 由 X 的分布算出. 这些都要求考生具有一定的综合能力.

例 9-8 设随机变量 X 和 Y 同分布,X 的概率密度为

$$f(x)=\begin{cases}\frac{3}{8}x^2, & 0<x<2;\\ 0,\text{其余.}\end{cases}$$

(1) 已知事件 $A=\{X>a\}$ 和 $B=\{Y>a\}$ 独立,且 $P(A\cup B)=\frac{3}{4}$ 求常数 a;

(2) 求$\frac{1}{X^2}$的数学期望.

解　(1) 由于 X,Y 同分布,因此,$P(A)=P(B)$. 又由于 A,B 独立,所以 $P(AB)=P(A)P(B)$.

$$P(A\cup B)=P(A)+P(B)-P(AB)=P(A)+P(A)-P(A)P(B)=2P(A)-(P(A))^2.$$

由条件知

$$2P(A)-(P(A))^2=\frac{3}{4},$$

即

$$4(P(A))^2-8P(A)+3=0,$$

解得

$$P(A)=\frac{1}{2}\text{或 }P(A)=\frac{3}{2}.$$

由于 $0\leqslant P(A)\leqslant 1$,因此取 $P(A)=\frac{1}{2}$.

$$P(A)=P(X>a)=\int_a^{+\infty}f(x)\mathrm{d}x=\int_a^2\frac{3}{8}x^2\mathrm{d}x=1-\frac{1}{8}a^3.$$

所以

$$1-\frac{1}{8}a^3=\frac{1}{2},$$

由此解得

$$a=\sqrt[3]{4}.$$

(2) $E\left(\frac{1}{X^2}\right)=\int_{-\infty}^{\infty}\frac{1}{x^2}f(x)\mathrm{d}x=\int_0^2\frac{1}{x^2}\cdot\frac{3}{8}x^2\mathrm{d}x=\frac{3}{8}\times 2=\frac{3}{4}$.

分析　本题是 1993 年试题. 本题考查的知识点有:事件的独立性、概率的加法公式、概率的性质($0\leqslant P(A)\leqslant 1$)、连续型随机变量的密度函数的应用和数学期望的计算. 只有能够综合运用学过的知识,才能取得理想的成绩.

例 9-9　假设由自动线加工的某种零件的内径 X(μm)服从正态分布 $N(\mu,1)$,内径小于 10 或大于 12 为不合格品,其余为合格品. 销售每件合格品获利,销售每件不合格品亏损,已知销售利润 T(单位:百元)与销售零件的内径 X 有如下关系:

$$T=\begin{cases}-1, & \text{若 } X<10;\\ 20, & \text{若 } 10\leqslant X\leqslant 12;\\ -5, & \text{若 } X>12.\end{cases}$$

问平均内径 μ 取何值时,销售一个零件的平均利润最大?

解　T 可能取值为 $-1,20,-5$,因此,T 为离散型随机变量,以下求 T 的分布律:

$$P(T=-1)=P(X<10)=P(X-\mu<10-\mu)=\Phi(10-\mu),$$

$$P(T=20)=P(10\leqslant X\leqslant 12)=P(10-\mu\leqslant X-\mu\leqslant 12-\mu)=\Phi(12-\mu)-\Phi(10-\mu),$$

$$P(T=-5)=P(X>12)=1-\Phi(12-\mu).$$

所以

$$\begin{aligned}E(T)&=(-1)\times\Phi(10-\mu)+20\times(\Phi(12-\mu)-\Phi(10-\mu))+(-5)\times(1-\Phi(12-\mu))\\&=25\Phi(12-\mu)-21\Phi(10-\mu)-5.\end{aligned}$$

ET 是 μ 的函数,为使 ET 最大,令$\frac{\mathrm{d}E(T)}{\mathrm{d}\mu}=0$. 即

$$\frac{\mathrm{d}E(T)}{\mathrm{d}\mu}=-25\varphi(12-\mu)+21\varphi(10-\mu)=0.$$

其中 $\varphi(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$ 是 $N(0,1)$ 的密度函数.

从而

$$25\mathrm{e}^{-\frac{(12-\mu)^2}{2}}=21\mathrm{e}^{-\frac{(10-\mu)^2}{2}},$$

$$\ln 25-\frac{(12-\mu)^2}{2}=\ln 21-\frac{(10-\mu)^2}{2}.$$

由此解得

$$\mu=11-\frac{1}{2}\ln\frac{25}{21}\approx 10.9.$$

所以当取 $\mu=10.9\mu m$ **时,平均利润最大.**

分析　本题是 1994 **年试题.本题考查考生综合运用高等数学知识和概率论知识的能力.**

例 9-10　假设二维随机变量(X,Y)**在矩形** $G=\{(x,y)\mid 0\leqslant x\leqslant 2,0\leqslant y\leqslant 1\}$**上服从均匀分布,记**

$$U=\begin{cases}0, & \text{若 } X\leqslant Y;\\ 1, & \text{若 } X>Y.\end{cases}\qquad V=\begin{cases}0, & \text{若 } X\leqslant 2Y;\\ 1, & \text{若 } X>2Y.\end{cases}$$

(1) **求** U **和** V **的联合分布;**

(2) **求** U **和** V **的相关系数** r.

解　(1) U **可能取值为** 0,1,V **可能取值也是** 0,1.**因此,**(U,V)**是二维离散型随机变量,可知取值**(0,0),(0,1),(1,0)**和**(1,1).

$$P(U=0,V=0)=P(X\leqslant Y,X\leqslant 2Y)=P(X\leqslant Y)=\iint\limits_{\{x\leqslant y\}}f(x,y)\mathrm{d}x\mathrm{d}y=\frac{1}{2}\int_0^1\mathrm{d}y\int_0^y\mathrm{d}x=\frac{1}{2}\cdot\frac{y^2}{2}\Big|_0^1=\frac{1}{4}.$$

这里用到(X,Y)**的联合密度函数**

$$f(x,y)=\begin{cases}\frac{1}{2}, & 0\leqslant x\leqslant 2 \text{ 且 } 0\leqslant y\leqslant 1;\\ 0, & \text{其余.}\end{cases}$$

$$P(U=0,V=1)=P(X\leqslant Y,X>2Y)=\iint\limits_{\left\{\substack{x\leqslant y\\ x>2y}\right\}}f(x,y)\mathrm{d}x\mathrm{d}y=0.$$

$$P(U=1,V=0)=P(X>Y,X\leqslant 2Y)=\iint\limits_{\{y<x\leqslant 2y\}}f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_0^1\mathrm{d}y\int_y^{2y}\frac{1}{2}\mathrm{d}x=\frac{1}{2}\int_0^1 y\mathrm{d}y=\frac{1}{4}.$$

$$P(U=1,V=1)=1-\frac{1}{4}-0-\frac{1}{4}=\frac{1}{2}.$$

因此,(U,V)**的联合分布律为**

U \ V	0	1
0	$\frac{1}{4}$	0
1	$\frac{1}{4}$	$\frac{1}{2}$

(2) U,V **的边缘分布律分别为**

U	0	1
$p_{r\cdot}$	$\frac{1}{4}$	$\frac{3}{4}$

V	0	1
$p_{r\cdot}$	$\frac{1}{2}$	$\frac{1}{2}$

$$EU=0\times\frac{1}{4}+1\times\frac{3}{4}=\frac{3}{4},\ EU^2=O^2\times\frac{1}{4}+1^2\times\frac{3}{4}=\frac{3}{4},$$

$$EV=0\times\frac{1}{2}+1\times\frac{1}{2}=\frac{1}{2},EV^2=O^2\times\frac{1}{2}+1^2\times\frac{1}{2}=\frac{1}{2},$$

$$EUV=0\times 0\times\frac{1}{4}+0\times 1\times 0+1\times 0\times\frac{1}{4}+1\times 1\times\frac{1}{2}=\frac{1}{2},$$

$$\operatorname{cov}(U,V)=EUV-EUEV=\frac{1}{2}-\frac{3}{4}\times\frac{1}{2}=\frac{1}{8},$$

$$D(U)=\frac{3}{4}\times\left(1-\frac{3}{4}\right)=\frac{3}{16},D(V)=\frac{1}{2}\times\left(1-\frac{1}{2}\right)=\frac{1}{4},$$

因此，U 和 V 的相关系数为

$$r=\frac{\operatorname{cov}(U,V)}{\sqrt{D(U)D(V)}}=\frac{\frac{1}{8}}{\sqrt{\frac{3}{16}\times\frac{1}{4}}}=\frac{1}{\sqrt{3}}=\frac{\sqrt{3}}{3}.$$

分析　本题是 1999 年试题. 本题中(X,Y)是二维连续型随机变量. 而由(X,Y)定义的(U,V)是二维离散型随机变量. 由(X,Y)的联合密度函数来求(U,V)的联合分布律，这类题目在教材中较少出现. 但这类题目应引起读者的重视. 读者可自行练习下面的习题(此题是 1997 年的试题)：

假设随机变量 Y 服从参数为 $\lambda=1$ 的指数分布，随机变量

$$X_k=\begin{cases}0, & \text{若 } Y\leqslant k;\\ 1, & \text{若 } Y>k.\end{cases}\quad(k=1,2)$$

(1) 求 X_1 和 X_2 的联合概率分布；

(2) 求 $E(X_1+X_2)$.

答案：

(1)

X_2 \ X_1	0	1
0	$1-\mathrm{e}^{-1}$	$\mathrm{e}^{-1}-\mathrm{e}^{-2}$
1	0	e^{-2}

(2) $\mathrm{e}^{-1}+\mathrm{e}^{-2}$

例 9-11　设随机变量 X 与 Y 独立，其中，X 的概率分布律为

X	1	2
$p_{r.}$	0.3	0.7

而 Y 的概率密度为 $f(y)$，求随机变量 $U=X+Y$ 的概率密度 $g(u)$.

解　先求 U 的分布函数. 记 Y 的分布函数为 $F(y)$，则 U 的分布函数为

$$G(u)=P(U\leqslant u)=P(X+Y\leqslant u),$$

由于 X 可能取值为 1 和 2，因此，$\{X=1\}$和$\{X=2\}$是完全事件组. 由全概率公式

$$G(u)=P\{X+Y\leqslant u\mid X=1\}P(X=1)+P\{X+Y\leqslant u\mid X=2\}P(X=2)$$

$$=P\{Y\leqslant u-1\mid X=1\}\times0.3+P\{Y\leqslant u-2\mid X=2\}\times0.7,$$

由于 X 与 Y 相互独立，因而

$$G(u)=0.3P(Y\leqslant u-1)+0.7P(Y\leqslant u-2)=0.3F(u-1)+0.7F(u-2).$$

U 的密度函数为

$$g(u)=\frac{\mathrm{d}}{\mathrm{d}u}G(u)=0.3f(u-1)+0.7f(u-2).$$

分析　本题是 2003 年试题. 在本题中 X 是离散型随机变量，Y 是连续型随机变量，$U=X+Y$ 是连续型

随机变量. 由 X 的概率分布和 Y 的密度函数可求出 U 的密度函数.

例 9-12 设 $X_1, X_2, \cdots, X_n$ 是取自总体 $N(0,\sigma^2)$ 的样本,其中 σ^2 未知,$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$,$\hat{\sigma}^2=\frac{1}{n}\cdot\sum_{i=1}^{n}X_i^2\ (n\geqslant 2)$

(1) 证明:s^2 和 $\hat{\sigma}^2$ 均是 σ^2 的无偏估计量;

(2) s^2 与 $\hat{\sigma}^2$ 哪个更有效?

(3) 证明:s^2 与 $\hat{\sigma}^2$ 均是 σ^2 的相合估计量.

证 (1) $Es^2=D(X)=\sigma^2$

$$E\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}EX_i^2=EX_1^2=\sigma^2,$$

因此,s^2 和 $\hat{\sigma}^2$ 均是 σ^2 的无偏估计量.

(2) 由于 $\frac{(n-1)}{\sigma^2}s^2\sim\chi^2(n-1)$,因此

$$D(s^2)=D\left(\frac{\sigma^2}{n-1}\cdot\frac{n-1}{\sigma^2}s^2\right)=\frac{\sigma^4}{(n-1)^2}D\left(\frac{n-1}{\sigma^2}s^2\right)=\frac{\sigma^4}{(n-1)^2}\cdot 2(n-1)=\frac{2}{n-1}\sigma^4,$$

由于 $\frac{1}{\sigma^2}\sum_{i=1}^{n}X_i^2\sim\chi^2(n)$,因此

$$D(\hat{\sigma}^2)=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)=D\left(\frac{\sigma^2}{n}\frac{\sum_{i=1}^{n}X_i^2}{\sigma^2}\right)=\frac{\sigma^4}{n^2}D\left(\frac{\sum_{i=1}^{n}X_i^2}{\sigma^2}\right)=\frac{\sigma^4}{n^2}\cdot 2n=\frac{2}{n}\sigma^4.$$

由于 $D(\hat{\sigma}^2)<D(s^2)$,因此,$\hat{\sigma}^2$ 比 s^2 更有效.

(3) 对任意 $\varepsilon>0$,由切比雪夫不等式

$$0\leqslant P(|s^2-\sigma^2|\geqslant\varepsilon)\leqslant\frac{D(s^2)}{\varepsilon^2}=\frac{2}{(n-1)}\frac{\sigma^4}{\varepsilon^2},$$

所以 $\lim\limits_{n\to\infty}P(|s^2-\sigma^2|\geqslant\varepsilon)=0$,即 s^2 是 σ^2 的相合估计.

$$0\leqslant P(|\hat{\sigma}^2-\sigma^2|\geqslant\varepsilon)\leqslant\frac{D(\hat{\sigma}^2)}{\varepsilon^2}=\frac{2\sigma^4}{n\varepsilon^2},$$

从而 $\lim\limits_{n\to\infty}P(|\hat{\sigma}^2-\sigma^2|\geqslant\varepsilon)=0$ 即 $\hat{\sigma}^2$ 是 σ^2 的相合估计.

分析 本题在 $D(s^2)$ 和 $D(\hat{\sigma}^2)$ 的计算中利用了正态总体下常用统计量分布的结果. 本题是考查考生综合运用数理统计知识的能力. 不用这些常用结果则计算十分繁琐.

例 9-13 设 A,B 是两个随机事件;随机变量

$$X=\begin{cases}1, & \text{若 } A \text{ 出现;}\\ 0, & \text{若 } A \text{ 不出现.}\end{cases}\qquad Y=\begin{cases}1, & \text{若 } B \text{ 出现;}\\ 0, & \text{若 } B \text{ 不出现.}\end{cases}$$

证明下列三个命题等价:

(1) 随机变量 X,Y 不相关;

(2) 随机变量 X,Y 独立;

(3) 随机事件 A,B 相互独立.

证 由 X,Y 的定义得到

$P(X=1)=P(A),P(X=0)=1-P(A)$,

$P(Y=1)=P(B),P(Y=0)=1-P(B)$,

$P(X=1,Y=1)=P(AB),P(X=1,Y=0)=P(A\overline{B}),P(X=0,Y=0)=P(\overline{A}\overline{B}),P(X=0,Y=1)=P(\overline{A}B)$.

由此得到 (X,Y) 的联合分布律为

X \ Y	0	1
0	$P(\overline{A}\overline{B})$	$P(\overline{A}B)$
1	$P(A\overline{B})$	$P(AB)$

X 和 Y 的边缘分布律分别为

X	0	1
$p_{r\cdot}$	$1-P(A)$	$P(A)$

Y	0	1
$p_{r\cdot}$	$1-P(B)$	$P(B)$

先证：(1)**与**(3)**等价**.

$$E(XY)=0\times0\times P(\overline{A}\overline{B})+0\times1\times P(\overline{A}B)+1\times0\times P(A\overline{B}+1\times1\times P(AB)$$

$$=P(AB),$$

$$EX=P(A),EY=P(B)$$

$$\operatorname{cov}(X,Y)=E(XY)-EXEY=P(AB)-P(A)P(B).$$

所以 $\operatorname{cov}(X,Y)=0$ **的充要充件是** $P(AB)=P(A)P(B)$，**即** A,B **相互独立**.

再证(2)**与**(3)**等价**.

若 X **与** Y **独立**，**则** $P(X=1,Y=1)=P(X=1)P(Y=1)$**即** $P(AB)=P(A)P(B)$，**因而** A,B **相互独立**.

若 A,B **相互独立**，**则** $P(AB)=P(A)P(B)$，**且** $\overline{A},B$ **独立**，$\overline{B},A$ **独立**，$\overline{A},\overline{B}$ **独立**，**所以**

$$P(X=0,Y=1)=P(\overline{A}B)=P(\overline{A})P(B)=(1-P(A))P(B)=P(X=0)P(Y=1),$$

$$P(X=1,Y=0)=P(A\overline{B})=P(A)P(\overline{B})=P(A)(1-P(B)=P(X=1)P(Y=0),$$

$$P(X=0,Y=0)=P(\overline{A}\overline{B})=P(\overline{A})P(\overline{B})=P(X=0)P(Y=0),$$

$$P(X=1,Y=1)=P(AB)=P(A)P(B)=P(X=1)P(Y=1).$$

因此，X **与** Y **相互独立**.

分析　**本题中** X,Y **分别由随机事件** A,B **定义**，**因此**，X **与** Y **的分布律可由** $P(A),P(B)$**定出**，**而**(X,Y)**的联合分布律由** $P(\overline{A}\overline{B}),P(\overline{A}B),P(A\overline{B})$**和** $P(AB)$**定出**. **当** X **与** Y **独立时**，(X,Y)**的联合分布律也可由** $P(A),P(B)$**定出**.

例 9-14　**对于任意两个事件** A **和** B，$0<P(A)<1,0<P(B)<1$.

$$\rho=\frac{P(AB)-P(A)P(B)}{\sqrt{P(A)P(B)P(\overline{A})P(\overline{B})}}$$

称作事件 A **和** B **的相关系数**.

(1) **证明**：**事件** A **和** B **独立的充分必要条件是其相关系数等于零**；

(2) **利用随机变量相关系数的基本性质**，**证明**：$|\rho|\leqslant1$.

证　(1) **由于** $0<P(A)<1,0<P(B)<1$，**所以**

$\rho=0$ **的充要条件为** $P(AB)=P(A)P(B)$，**即** $\rho=0$ **的充要条件是** A,B **相互独立**.

(2) **引入随机变量**

$$X=\begin{cases}1, & \text{若 } A \text{ 出现};\\ 0, & \text{若 } A \text{ 不出现}.\end{cases}\qquad Y=\begin{cases}1, & \text{若 } B \text{ 出现};\\ 0, & \text{若 } B \text{ 不出现}.\end{cases}$$

由例 9-13 **可知**，$\operatorname{cov}(X,Y)=P(AB)-P(A)P(B)$，

$$X\sim B(1,P(A)),\ Y\sim B(1,P(B)),$$

从而

$$D(X)=P(A)(1-P(A))=P(A)P(\overline{A}),$$

$$D(Y)=P(B)(1-P(B))=P(B)P(\overline{B}).$$

所以(X,Y)**的相关系数**

$$\rho(X,Y)=\frac{P(AB)-P(A)P(B)}{\sqrt{P(A)P(\overline{A}P(B)P(\overline{B})}}=\rho.$$

由相关系数的性质有

$$|\rho| = |\rho(X,Y)| \leqslant |.$$

分析 本题是 2003 年试题. 本题(2)的证明中引入随机变量 X,Y 的做法对读者来说有一定的难度. 但如果做过例 9-13 后,也许会比较容易想到.

二、小 结

概率统计的综合题的解法灵活多变,但要做到应付自如、得心应手离不开对基本知识的理解和掌握. 读者应着重把握概率论知识的理解和应用. 综合题主要考查考生对概率论的相关知识点的灵活运用. 由于课时的关系,对统计部分的要求相对较低.

综合练习题

1. 设 $P(A)=\frac{1}{3}, P(B)=\frac{1}{2}$. 在下列三种情形下,分别求出 $P(A\cup B), P(A-B), P(B-A), P(B\mid A), P(\overline{B}\mid\overline{A})$.

(1) $A=\{X\geqslant 2\}, B=\{X\geqslant 1\}$,其中,$X$ 是随机变量;

(2) $A=\{X\geqslant 2\}, B=\{X\leqslant 1\}$;

(3) $A=\{X\geqslant 2\}, B=\{Y\geqslant 1\}$,$X$ 与 Y 相互独立.

2. 设 $P(B)=0.2, P(A\cup B)=0.3, P(B\mid A)=0.1$,求 $P(A)$.

3. 在一次试验中事件 A 发生的概率为 p,把这个试验独立重复地做两次. 在下列两种情形下分别求 p 的值.

(1) 已知事件 A 至少发生一次的概率与事件 A 至多发生一次的概率相等;

(2) 已知事件 A 至多发生一次的条件下事件 A 至少发生一次的概率为 $\frac{1}{2}$.

4. 设随机变量 $X\sim N(\mu,1), \mu>0$. 已知 $P(X<0)=0.3$,求 $P(\mu\leqslant X\leqslant 2\mu)$.

5. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases} cx^3, & 0<x<1;\\ 0, & \text{其 余}. \end{cases}$$

(1) 求常数 c 的值;

(2) 求常数 a 的值,使得 $P(X\leqslant a)=P(X\geqslant a)$;

(3) 求 X 的 k 阶原点矩.

6. 设随机变量 X 与 Y 相互独立,$X\sim B(2,0.5), Y\sim P(1)$,试求关于 t 的一元二次方程 $t^2+2Xt+Y=0$ 有不同实根、无实根、有相同实根的概率.

7. 设二维随机变量 (X,Y) 的分布律为

X \ Y	-0.1	0	1
-1	0.1	0	0.3
0	0	0.2	0.4

求随机变量 $Z=X^2+2Y$ 的分布律.

8. 在第 5 题中,求随机变量 $Y=\ln X$ 的分布函数与密度函数.

9. 设随机变量 $X\sim N(-1,2)$. 试求:

(1) $U=2X+2$ 的密度函数;

(2) $V=16(X+1)^4$ 的密度函数.

10. 设随机变量 X 与 Y 相互独立,且 $X\sim N(2,1)$, $Y\sim N(0,2)$. 试求:

(1) $U=3X-2Y-6$ 的密度函数;

(2) $V=|3X-2Y-6|$ 的密度函数.

11. 设随机变量 X 与 Y 相互独立,X 与 Y 都服从正态分布 $N(0,2)$. 试求随机变量 $Z=X^2+Y^2$ 的分布函数与密度函数.

12. 设随机变量 X 的分布律为

$$P(X=-1)=0.1, P(X=1)=0.5, P(X=3)=0.4.$$

已知 X 的分布函数为 $F(x)$,求 $F(\sqrt{e})$.

13. 已知随机变量 X 的密度函数为

$$f(x)=c\exp\{-x^2-4x-1\}, -\infty<x<\infty.$$

试求 $E(X), D(X), E(X^2)$,X 的 k 阶中心矩与常数 c.

14. 设随机变量 $X\sim B(3,\frac{1}{3}), Y\sim P(2)$. 设 $Z=2X-3Y+1$. 在下列两种情形下,分别求 $E(Z), D(Z)$

与 $E(Z^2)$.

(1) X 与 Y 相互独立；

(2) $\rho(X,Y)=-0.5$.

15. 已知 $E(X)=E(Y)=1, E(Z)=-1, D(X)=D(Y)=D(Z)=2, \rho(X,Y)=0, \rho(Y,Z)=\frac{1}{2}$, $\rho(Z,X)=-\frac{1}{2}$. 设随机变量 $W=X-Y+Z$. 试求 $E(W)$, $D(W)$ 与 $E(W^2)$.

16. 设随机变量 X 的密度函数为

$$f(x)=\frac{1}{2}e^{-|x|}, \quad -\infty<x<\infty.$$

(1) 求 $E(|X|)$ 与 $D(|X|)$；

(2) 求随机变量 $Y=|X|$ 的分布函数.

17. 设二维随机变量 (X,Y) 服从区域 G 上的均匀分布，区域 G 由四条直线 $x+y=1, x+y=-1, x-y=1, x-y=-1$ 所围成.

(1) 求 X 的边缘密度函数与边缘分布函数；

(2) X 与 Y 是否不相关？

(3) X 与 Y 是否独立？

(4) 求给定 $Y=y$ 时 X 的条件密度函数与给定 $X=0$ 时 Y 的条件密度函数.

18. 设二维随机变量 (X,Y) 服从区域 G 上的均匀分布，区域 G 由三条直线 $y=x, y=1, x=0$ 所围成. 试求：

(1) $\mathrm{cov}(X,Y)$ 与 $\rho(X,Y)$；

(2) $E(X^2Y)$；

(3) $D(-2X+1)$ 与 $D(-2X+3Y-1)$.

19. 假设生产线上组装每件成品的时间服从指数分布，统计资料表明：该生产线上组装一件成品的平均组装时间为 10min，各件产品的组装时间相互独立.

(1) 试求在该生产线组装 100 件产品需要 15h 至 20h 的概率；

(2) 以 95% 的概率在 16h 之内最多可以组装多少件成品？

20. 设 $X_1,\cdots,X_5$ 是取自正态总体 $N(0,\sigma^2)$ 的一个样本. 试问，当常数 c,d 取何值时，随机变量 $Y=\frac{1}{\sigma^2}[c(X_1+X_2)^2+d(X_3+X_4+X_5)^2]$ 服从自由度为 2 的 χ^2 分布 $\chi^2(2)$？

21. 设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本，$X\sim E\left(\frac{1}{\theta}\right)$，其中 θ 未知，$\theta>0$. 试求未知参数 θ 的矩估计量与极大似然估计量，并讨论它们的无偏性与求出它们的方差.

22. 设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本，$X\sim P(\lambda)$，其中 λ 未知. 试问，$\hat{\lambda}=\alpha\overline{X}+(1-\alpha)S^2$ 是否为未知参数 λ 的无偏估计？其中，α 是常数.

23. 设 $X_1,\cdots,X_n$ 是取自正态总体 $N(\mu,1)$ 的一个样本，其中 μ 未知. 在 μ 的 n 个无偏估计量

$$\overline{X}_k=\frac{1}{k}\sum_{i=1}^{k}X_i, \quad k=1,\cdots,n$$

中，找出方差最小的一个，并写出它的方差.

24. 设 $X_1,\cdots,X_n$ 是取自正态总体 $N(\mu,\sigma^2)$ 的一个样本，其中 σ^2 未知但 μ 已知. 试问，σ^2 的两个无偏估计 S^2 与 $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$ 中，哪一个有效？

25. 设 $X_1,\cdots,X_n$ 是取自总体 X 的一个样本，$X\sim R(-\theta,\theta)$，其中 θ 未知，$\theta>0$.

(1) 试求 θ 与 θ^2 的矩估计量；

(2) θ^2 的矩估计量是否具有无偏性？

(3) θ^2 的矩估计量是否具有一致性？

26. 设 $X_1,\cdots,X_n$ 是取自正总体 $N(\mu,1)$ 的一个样本，其中 μ 未知. 在置信水平 $1-\alpha$ 下，试求未知参数 $\theta=k\mu+c$ 的置信区间，其中，k,c 是常数，$k\neq 0$.

27. 设 $X_1,\cdots,X_n$ 是取自正态总体 $N(\mu,\sigma^2)$ 的一个样本，其中 μ 未知但 σ^2 已知. 在置信水平 $1-\alpha$ 下，要使 μ 的置信区间的长度不超过 l_0，问 n 至少要取多大？

28. 从两个正态总体 $N(\mu_1,\sigma^2)$，$N(\mu_2,\sigma^2)$ 中分别抽取样本 $(X_1,\cdots,X_n)$，$(Y_1,\cdots,Y_n)$，其中 μ_1,μ_2,σ^2 均未知，在置信水平 $1-\alpha$ 下，试求未知参数 $\theta=\mu_1-\mu_2$ 的形如 $(-\infty,\overline{\theta}]$ 的置信区间.

29. 设 $X_1,\cdots,X_9$ 是取自正态总体 $N(\mu,1)$ 的一个样本. 现在要检验

$$H_0:\mu=\mu_0 \quad (H_1:\mu\neq 0).$$

取拒绝域 $W_1=\{(x_1,\cdots,x_9)\mid|\bar{x}|>d\}$. 试确定常数 d，使得犯第 Ⅰ 类错误的概率为 0.10.

30. 设 $X_1,\cdots,X_n$ 是取自态总体 $N(\mu,\sigma^2)$ 的一个样本，其中 μ 与 σ^2 均未知. 在显著性水平 α 下，要检验

$$H_0:\mu=\mu_0 \quad (H_1:\mu<\mu_0).$$

试求拒绝域 W_1.

习题简答

1. (1) $\frac{1}{2}$，0，$\frac{1}{6}$，1，$\frac{3}{4}$；(2) $\frac{5}{6}$，$\frac{1}{3}$，$\frac{1}{2}$，0，$\frac{1}{4}$；(3) $\frac{2}{3}$，$\frac{1}{6}$，$\frac{1}{3}$，$\frac{1}{2}$，$\frac{1}{2}$.

2. $\frac{1}{9}$.

3. (1) $\frac{1}{2}$；(2) $\frac{1}{3}$.

4. 0.2.

5. (1) 4；(2) $1/\sqrt[4]{2}$；(3) $\frac{4}{k+4}$.

6. $\frac{7}{6\mathrm{e}}$，$1-\frac{185}{96\mathrm{e}}$，$\frac{73}{96\mathrm{e}}$.

7.

Z	0	0.8	2	3
P_r	0.2	0.1	0.4	0.3

8. $F_Y(y)=\begin{cases}\mathrm{e}^{4y}, & y<0;\\ 1, & y\geqslant 0,\end{cases}$ $f_Y(y)=\begin{cases}4\mathrm{e}^{4y}, & y<0;\\ 0, & \text{其余}.\end{cases}$

9. (1) $f_U(u)=\frac{1}{4\sqrt{\pi}}\mathrm{e}^{-\frac{u^2}{16}}$，$-\infty<u<\infty$；(2) $f_V(v)=\begin{cases}\frac{1}{8\sqrt{\pi}\sqrt[4]{v^3}}\mathrm{e}^{-\frac{\sqrt{v}}{16}}, & v>0;\\ 0, & \text{其余}.\end{cases}$

10. (1) $f_U(u)=\frac{1}{\sqrt{34\pi}}\mathrm{e}^{-\frac{u^2}{34}}$，$-\infty<u<\infty$；(2) $f_V(v)=\begin{cases}\sqrt{\frac{2}{17\pi}}\mathrm{e}^{-\frac{v^2}{34}}, & v>0;\\ 0, & \text{其余}.\end{cases}$

11. $F_Z(z)=\begin{cases}1-\mathrm{e}^{-\frac{z}{4}}, & z>0;\\ 0, & z\leqslant 0,\end{cases}$ $f_Z(z)=\begin{cases}\frac{1}{4}\mathrm{e}^{-\frac{z}{4}}, & z>0;\\ 0, & \text{其余}.\end{cases}$

12. 0.6.

13. -2；$\frac{1}{2}$；$\frac{9}{2}$；当 k 为奇数时是 0；当 k 为偶数时是 $2^{-\frac{k}{2}}(k-1)!!$；$\frac{1}{\mathrm{e}^3\sqrt{\pi}}$.

14. (1) -3，$\frac{62}{3}$，$\frac{89}{3}$；(2) -3，$\frac{62}{3}+4\sqrt{3}$，$\frac{89}{3}+4\sqrt{3}$.

15. -1，2，3.

16. (1) 1，1；(2) $F_Y(y)=\begin{cases}1-\mathrm{e}^{-y}, & y>0;\\ 0, & \text{其余}.\end{cases}$

17. (1) $f_X(x)=\begin{cases}1-x, & 0<x<1;\\ 1+x, & -1<x<0;\\ 0, & \text{其余}.\end{cases}$ 或 $f_X(x)=\begin{cases}1-|x|, & -1<x<1;\\ 0, & \text{其余}.\end{cases}$

$$F_X(x)=\begin{cases}0, & x<-1;\\ \frac{1}{2}(1+x)^2, & -1\leqslant x<0;\\ 1-\frac{1}{2}(1-x)^2, & 0\leqslant x<1;\\ 1, & x\geqslant 1.\end{cases}$$

(2) **不相关**； (3) **不独立**；

(4) **当** $-1<y<1$ **时**，

$$f(x\mid y)=\begin{cases}\frac{1}{2(1-|y|)}, & |y|-1<x<1-|y|;\\ 0, & \text{**其余**.}\end{cases}$$

当 $y\geqslant 1$ **或** $y\leqslant -1$ **时，条件密度函数不存在.**

$$f(y\mid 0)=\begin{cases}\frac{1}{2}, & -1<y<1;\\ 0, & \text{**其余**.}\end{cases}$$

18. (1) $\frac{1}{36}$, $\frac{1}{2}$; (2) $\frac{1}{5}$; (3) $\frac{2}{9}$, $\frac{7}{18}$.

19. (1) 0.8185; (2) 81

20. $c=\frac{1}{2}, d=\frac{1}{3}$.

21. **矩估计量与极大似然估计量都是** $\overline{X}$，**且是无偏估计，方差为** $\frac{\theta^2}{n}$.

22. **是无偏估计.**

23. $\overline{X}_n=\frac{1}{n}\sum\limits_{i=1}^{n}X_i, D(\overline{X}_n)=\frac{1}{n}$.

24. $\hat{\sigma}^2$ **有效.**

25. (1) $\sqrt{\frac{3}{n}\sum\limits_{i=1}^{n}X_i^2}$, $\frac{3}{n}\sum\limits_{i=1}^{n}X_i^2$; (2) **是无偏估计**; (3) **是一致估计.**

26. **当** $k>0$ **时**，$\left[k\overline{X}+c-u_{1-\frac{\alpha}{2}}\frac{k}{\sqrt{n}}, \quad k\overline{X}+c+u_{1-\frac{\alpha}{2}}\frac{k}{\sqrt{n}}\right]$；**当** $k<0$ **时**，$\left[k\overline{X}+c+u_{1-\frac{\alpha}{2}}\frac{k}{\sqrt{n}}, \quad k\overline{X}+c-u_{\frac{\alpha}{2}}\frac{k}{\sqrt{n}}\right]$.

27. $n\geqslant 4\sigma^2u_{1-\frac{\alpha}{2}}^2/l_0^2$ **的最小整数.**

28. $\overline{\theta}=(\overline{X}-\overline{Y})+t_{1-\alpha}(2n-2)\sqrt{\left[\sum\limits_{i=1}^{n}(X_i-\overline{X})^2+\sum\limits_{i=1}^{n}(Y_i-\overline{Y})^2\right]/n(n-1)}$.

29. $d=\frac{1}{3}u_{0.95}$.

30. $W_1=\{(x_1,\cdots,x_n)\mid t=\sqrt{n}\frac{\overline{x}-\mu_0}{S}<-t_{1-\alpha}(n-1)\}$.

附录　全国硕士研究生入学统一考试答案和评分参考

2005 年数学(一) 试卷

一、填空题(本题共 6 小题,每小题 4 分,满分 24 分.)

(1) 曲线 $y=\dfrac{x^2}{2x+1}$ 的斜渐近线方程为 $\underline{y=\dfrac{1}{2}x-\dfrac{1}{4}}$.

(2) 微分方程 $xy'+2y=x\ln x$ 满足 $y(1)=-\dfrac{1}{9}$ 的解为 $\underline{y=\dfrac{x}{3}\left(\ln x-\dfrac{1}{3}\right)}$.

(3) 设函数 $u(x,y,z)=1+\dfrac{x^2}{6}+\dfrac{y^2}{12}+\dfrac{z^2}{18}$,单位向量 $\boldsymbol{n}=\dfrac{1}{\sqrt{3}}\{1,1,1\}$,则 $\left.\dfrac{\partial u}{\partial \boldsymbol{n}}\right|_{(1,2,3)}=\underline{\dfrac{\sqrt{3}}{3}}$.

(4) 设 Ω 是由锥面 $z=\sqrt{x^2+y^2}$ 与半球面 $z=\sqrt{R^2-x^2-y^2}$ 围成的空间区域,Σ 是 Ω 的整个边界的外侧,则 $\iint\limits_{\Sigma} x\mathrm{d}y\mathrm{d}z+y\mathrm{d}z\mathrm{d}x+z\mathrm{d}x\mathrm{d}y=\underline{(2-\sqrt{2})\pi R^3}$.

(5) 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 均为 3 维列向量,记矩阵

$$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3),\boldsymbol{B}=(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+4\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+9\boldsymbol{\alpha}_3).$$

如果 $|\boldsymbol{A}|=1$,那么 $|\boldsymbol{B}|=\underline{2}$.

(6) 从数 1,2,3,4 中任取一个数,记为 X,再从 $1,\cdots,X$ 中任取一个数,记为 Y,则 $P\{Y=2\}=\underline{\dfrac{13}{48}}$.

二、选择题(本题共 8 小题,每小题 4 分,满分 32 分.)

(7) 设函数 $f(x)=\lim\limits_{n\to\infty}\sqrt[n]{1+|x|^{3n}}$,则 $f(x)$ 在 $(-\infty,+\infty)$ 内 【C】

(A) 处处可导;　　(B) 恰有一个不可导点;

(C) 恰有两个不可导点;　　(D) 至少有三个不可导点.

(8) 设 $F(x)$ 是连续函数 $f(x)$ 的一个原函数,"$M\Leftrightarrow N$" 表示"M 的充分必要条件是 N",则必有 【A】

(A) $F(x)$ 是偶函数 $\Leftrightarrow f(x)$ 是奇函数;　　(B) $F(x)$ 是奇函数 $\Leftrightarrow f(x)$ 是偶函数;

(C) $F(x)$ 是周期函数 $\Leftrightarrow f(x)$ 是周期函数;　　(D) $F(x)$ 是单调函数 $\Leftrightarrow f(x)$ 是单调函数.

(9) 设函数 $u(x,y)=\varphi(x+y)+\varphi(x-y)+\int_{x-y}^{x+y}\psi(t)\mathrm{d}t$,其中函数 φ 具有二阶导数,ψ 具有一阶导数,则必有 【B】

(A) $\dfrac{\partial^2 u}{\partial x^2}=-\dfrac{\partial^2 u}{\partial y^2}$;　　(B) $\dfrac{\partial^2 u}{\partial x^2}=\dfrac{\partial^2 u}{\partial y^2}$;

(C) $\dfrac{\partial^2 u}{\partial x\partial y}=\dfrac{\partial^2 u}{\partial y^2}$;　　(D) $\dfrac{\partial^2 u}{\partial x\partial y}=\dfrac{\partial^2 u}{\partial x^2}$.

(10) 设有三元方程 $xy-z\ln y+\mathrm{e}^{xz}=1$,根据隐函数存在定理,存在点 $(0,1,1)$ 的一个邻域,在此邻域内该方程 【D】

(A) 只能确定一个具有连续偏导数的隐函数 $z=z(x,y)$;

(B) 可确定两个具有连续偏导数的隐函数 $y=y(x,z)$ 和 $z=z(x,y)$;

(C) 可确定两个具有连续偏导数的隐函数 $x=x(y,z)$ 和 $z=z(x,y)$;

(D) 可确定两个具有连续偏导数的隐函数 $x=x(y,z)$ 和 $y=y(x,z)$.

(11) 设 λ_1,λ_2 是矩阵 $\boldsymbol{A}$ 的两个不同的特征值,对应的特征向量分别为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$,则 $\boldsymbol{\alpha}_1,\boldsymbol{A}(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)$ 线性无关的充分必要条件是 【B】

(A) $\lambda_1\neq 0$　　(B) $\lambda_2\neq 0$　　(C) $\lambda_1=0$　　(D) $\lambda_2=0$.

(12) 设 $\boldsymbol{A}$ 为 $n(n\geqslant 2)$ 阶可逆矩阵,交换 $\boldsymbol{A}$ 的第 1 行与第 2 行得矩阵 $\boldsymbol{B},\boldsymbol{A}^*,\boldsymbol{B}^*$ 分别为 $\boldsymbol{A},\boldsymbol{B}$ 的伴随矩阵,

则 【C】

(A) 交换 $\boldsymbol{A}^*$ 的第 1 列与第 2 列得 $\boldsymbol{B}^*$；　(B) 交换 $\boldsymbol{A}^*$ 的第 1 行与第 2 行得 $\boldsymbol{B}^*$；

(C) 交换 $\boldsymbol{A}^*$ 的第 1 列与第 2 列得 $-\boldsymbol{B}^*$；　(D) 交换 $\boldsymbol{A}^*$ 的第 1 行与第 2 行得 $-\boldsymbol{B}^*$.

(13) 设二维随机变量(X,Y) 的概率分布为 【B】

X \ Y	0	1
0	0.4	a
1	b	0.1

已知随机事件$\{X=0\}$ 与$\{X+Y=1\}$ 相互独立，则

(A) $a=0.2, b=0.3$；　(B) $a=0.4, b=0.1$；

(C) $a=0.3, b=0.2$；　(D) $a=0.1, b=0.4$.

(14) 设 $X_1, X_2, \cdots, X_n (n \geqslant 2)$ 为来自总体 $N(0,1)$ 的简单随机样本，$\overline{X}$ 为样本均值，S^2 为样本方差，则 【D】

(A) $n\overline{X} \sim N(0,1)$；　(B) $nS^2 \sim \chi^2(n)$；

(C) $\dfrac{(n-1)\overline{X}}{S} \sim t(n-1)$；　(D) $\dfrac{(n-1)X_1^2}{\sum\limits_{i=2}^{n} X_i^2} \sim F(1,n-1)$.

三、解答题(本题共 9 小题，满分 94 分).

(15)(本题满分 11 分)

设 $D=\{(x,y) \mid x^2+y^2 \leqslant \sqrt{2}, x \geqslant 0, y \geqslant 0\}$，$[1+x^2+y^2]$ 表示不超过 $1+x^2+y^2$ 的最大整数. 计算二重积分 $\iint\limits_D xy[1+x^2+y^2]\mathrm{d}x\mathrm{d}y$.

解法 1

$$\iint_D xy[1+x^2+y^2]\mathrm{d}x\mathrm{d}y = \int_0^{\frac{\pi}{2}} \mathrm{d}\theta \int_0^{\sqrt{2}} r^3 \sin\theta\cos\theta[1+r^2]\mathrm{d}r \quad (3\text{ 分})$$

$$= \int_0^{\frac{\pi}{2}} \sin\theta\cos\theta \mathrm{d}\theta \int_0^{\sqrt{2}} r^3[1+r^2]\mathrm{d}r \quad (5\text{ 分})$$

$$= \frac{1}{2}\left(\int_0^1 r^3\mathrm{d}r + \int_1^{\sqrt{2}} 2r^3\mathrm{d}r\right) \quad (10\text{ 分})$$

$$= \frac{3}{8}. \quad (11\text{ 分})$$

解法 2 记 $D_1=\{(x,y) \mid x^2+y^2<1, x \geqslant 0, y \geqslant 0\}$,

$D_2=\{(x,y) \mid 1 \leqslant x^2+y^2 \leqslant \sqrt{2}, x \geqslant 0, y \geqslant 0\}$,

则 $[1+x^2+y^2]=1, (x,y) \in D_1, [1+x^2+y^2]=2, (x,y) \in D_2$. (4 分)

于是

$$\iint_D xy[1+x^2+y^2]\mathrm{d}x\mathrm{d}y = \iint_{D_1} xy\mathrm{d}x\mathrm{d}y + \iint_{D_2} 2xy\mathrm{d}x\mathrm{d}y \quad (6\text{ 分})$$

$$= \int_0^{\frac{\pi}{2}} \mathrm{d}\theta \int_0^1 r^3\sin\theta\cos\theta\mathrm{d}r + \int_0^{\frac{\pi}{2}} \mathrm{d}\theta \int_1^{\sqrt{2}} 2r^3\sin\theta\cos\theta\mathrm{d}r \quad (9\text{ 分})$$

$$= \frac{1}{8} + \frac{1}{4} = \frac{3}{8}. \quad (11\text{ 分})$$

(16)(本题满分 12 分)

求幂级数 $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\left(1+\dfrac{1}{n(2n-1)}x^{2n}\right)$ 的收敛区间与和函数 $f(x)$.

解 因为 $\lim\limits_{n\to\infty}\dfrac{(n+1)(2n+1)+1}{(n+1)(2n+1)}\cdot\dfrac{n(2n-1)}{n(2n-1)+1}=1$，所以，当 $x^2<1$ 时，原级数绝对收敛，当 $x^2>1$ 时，原级数发散，因此，原级数的收敛半径为 1，收敛区间为 $(-1,1)$. （3 分）

记 $$S(x)=\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n(2n-1)}x^{2n},x\in(-1,1),$$

则 $S'(x)=\sum\limits_{n=1}^{\infty}\dfrac{(-1)^{n-1}}{2n-1}x^{2n-1},x\in(-1,1),S''(x)=\sum\limits_{n=1}^{\infty}(-1)^{n-1}x^{2n-2}=\dfrac{1}{1+x^2},x\in(-1,1).$ （6 分）

由于 $$S(0)=0,\quad S'(0)=0,$$

所以 $$S'(x)=\int_0^x S''(t)\mathrm{d}t=\int_0^x\frac{1}{1+t^2}\mathrm{d}t=\arctan x,$$

$$S(x)=\int_0^x S'(t)\mathrm{d}t=\int_0^x\arctan t\mathrm{d}t=x\arctan x-\frac{1}{2}\ln(1+x^2).$$ （9 分）

又 $$\sum_{n=1}^{\infty}(-1)^{n-1}x^{2n}=\frac{x^2}{1+x^2},x\in(-1,1),$$ （11 分）

从而 $$f(x)=2S(x)+\frac{x^2}{1+x^2}=2x\arctan x-\ln(1+x^2)+\frac{x^2}{1+x^2},x\in(-1,1).$$ （12 分）

(17)（本题满分 11 分）

如图，曲线 C 的方程为 $y=f(x)$，点 $(3,2)$ 是它的一个拐点，直线 l_1 与 l_2 分别是曲线 C 在点 $(0,0)$ 与 $(3,2)$ 处的切线，其交点为 $(2,4)$. 设函数 $f(x)$ 具有三阶连续导数，计算定积分 $\int_0^3(x^2+x)f'''(x)\mathrm{d}x$.

解 $\int_0^3(x^2+x)f'''(x)\mathrm{d}x$

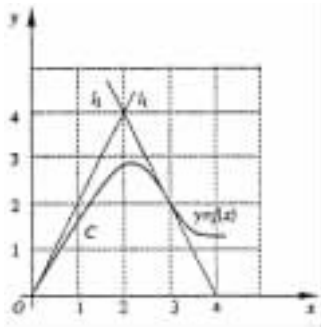
$=(x^2+x)f''(x)\Big|_0^3-\int_0^3(2x+1)f''(x)\mathrm{d}x$ （2 分）

$=-\int_0^3(2x+1)f''(x)\mathrm{d}x$ （4 分）

$=-(2x+1)f'(x)\Big|_0^3+2\int_0^3 f'(x)\mathrm{d}x$ （6 分）

$=-[7\times(-2)-2]+2\int_0^3 f'(x)\mathrm{d}x$ （8 分）

$=16+2f(x)\Big|_0^3=16+4=20.$ （11 分）



(18)（本题满分 12 分）

已知函数 $f(x)$ 在 $[0,1]$ 上连续，在 $(0,1)$ 内可导，且 $f(0)=0,f(1)=1$. 证明：

（Ⅰ）存在 $\xi\in(0,1)$，使得 $f(\xi)=1-\xi$；

（Ⅱ）存在两个不同的点 $\eta,\zeta\in(0,1)$，使得 $f'(\eta)f'(\zeta)=1$.

证 （Ⅰ）令 $g(x)=f(x)+x-1$，则 $g(x)$ 在 $[0,1]$ 上连续，且

$$g(0)=-1<0,\qquad g(1)=1>0,$$ （3 分）

所以存在 $\xi\in(0,1)$，使得

$$g(\xi)=f(\xi)+\xi-1=0,$$

即 $$f(\xi)=1-\xi$$ （5 分）

（Ⅱ）根据拉格朗日中值得定理，存在 $\eta\in(0,\xi),\zeta\in(\xi,1)$，使得

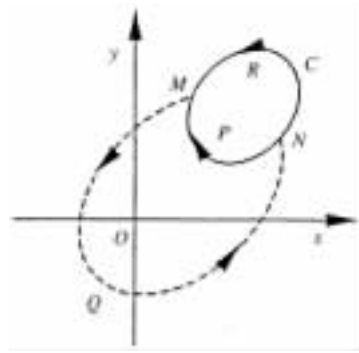
$$f'(\eta)=\frac{f(\xi)-f(0)}{\xi}=\frac{1-\xi}{\xi},$$ （8 分）

$$f'(\zeta)=\frac{f(1)-f(\xi)}{1-\xi}=\frac{1-(1-\xi)}{1-\xi}=\frac{\xi}{1-\xi},\qquad (11\text{ 分})$$

从而
$$f'(\eta)f'(\zeta)=\frac{1-\xi}{\xi}\cdot\frac{\xi}{1-\xi}=1.\qquad (12\text{ 分})$$

(19)(本题满分 12 分)

设函数 $\varphi(y)$ 具有连续导数,在围绕原点的任意分段光滑简单闭曲线 L 上,曲线积分 $\oint_L \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}$ 的值恒为同一常数.



(Ⅰ)证明:对右半平面 $x>0$ 内的任意分段光滑简单闭曲线 C,有 $\oint_C \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}=0$;

(Ⅱ)求函数 $\varphi(y)$ 的表达式.

(Ⅰ)证　如图,设 C 是半平面 $x>0$ 内的任一分段光滑简单闭曲线,在 C 上任意取定两点 M,N,作围绕原点的闭曲线 $\overset{\frown}{MQNRM}$,同时得到另一转绕原点的闭曲线 $\overset{\frown}{MQNPM}$.

根据题设可知

$$\oint_{\overset{\frown}{MQNRM}} \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}-\oint_{\overset{\frown}{MQNPM}} \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}=0.\qquad (2\text{ 分})$$

根据第二类曲线积分的性质,利用上式可得

$$\begin{aligned}&\oint_C \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}\\&=\int_{\overset{\frown}{NRM}} \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}+\int_{\overset{\frown}{MPN}} \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}=\int_{\overset{\frown}{NRM}} \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}-\int_{\overset{\frown}{NPM}} \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}\\&=\oint_{\overset{\frown}{MQNRM}} \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}-\oint_{\overset{\frown}{MQNPM}} \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}=0.\end{aligned}\qquad (4\text{ 分})$$

(Ⅱ)解　设 $P=\frac{\varphi(y)}{2x^2+y^4}$,$Q=\frac{2xy}{2x^2+y^4}$,$P,Q$ 在单连通区域 $x>0$ 内具有一阶连续偏导数.由(Ⅰ)知,曲线积分 $\int_L \frac{\varphi(y)\mathrm{d}x+2xy\mathrm{d}y}{2x^2+y^4}$ 在该区域内与路径无关,故当 $x>0$ 时,总有 $\frac{\partial Q}{\partial x}=\frac{\partial P}{\partial y}$.　(6 分)

$$\begin{aligned}&\frac{\partial Q}{\partial x}=\frac{2y(2x^2+y^4)-4x\cdot 2xy}{(2x^2+y^4)^2}=\frac{-4x^2y+2y^5}{(2x^2+y^4)^2},&①\\&\frac{\partial P}{\partial y}=\frac{\varphi'(y)(2x^2+y^4)-4\varphi(y)y^3}{(2x^2+y^4)^2}=\frac{2x^2\varphi'(y)+\varphi'(y)y^4-4\varphi(y)y^3}{(2x^2+y^4)^2}.&②\end{aligned}\qquad (9\text{ 分})$$

比较①、②两式的右端,得

$$\begin{cases}\varphi'(y)=-2y,&③\\ \varphi'(y)y^4-4\varphi(y)y^3=2y^5.&④\end{cases}$$

由式③得　$\varphi(y)=-y^2+c$,

将 $\varphi(y)$ 代入④得　$2y^5-4cy^3=2y^5$,　(11 分)

所以 $c=0$,从而 $\varphi(y)=-y^2$.　(12 分)

(20)(本题满分 9 分)

已知二次型 $f(x_1,x_2,x_3)=(1-a)x_1^2+(1-a)x_2^2+2x_3^2+2(1+a)x_1x_2$ 的秩为 2.

(Ⅰ)求 a 的值;

(Ⅱ)求正交变换 $x=Qy$ 把 $f(x_1,x_2,x_3)$ 化成标准形;

(Ⅲ)求方程 $f(x_1,x_2,x_3)=0$ 的解.

解 （Ⅰ）由于二次型 f 的秩为 2，对应的矩阵 $\boldsymbol{A}=\begin{pmatrix}1-a & 1+a & 0\\ 1+a & 1-a & 0\\ 0 & 0 & 2\end{pmatrix}$ 的秩为 2，所以有 $\begin{vmatrix}1-a & 1+a\\ 1+a & 1-a\end{vmatrix}=-4a=0$，得 $a=0$.

（Ⅱ）当 $a=0$ 时，$\boldsymbol{A}=\begin{pmatrix}1 & 1 & 0\\ 1 & 1 & 0\\ 0 & 0 & 2\end{pmatrix}$，

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1 & -1 & 0\\ -1 & \lambda-1 & 0\\ 0 & 0 & \lambda-2\end{vmatrix}=(\lambda-2)^2\lambda,$$

可知 $\boldsymbol{A}$ 的特征值为 $\lambda_1=\lambda_2=2,\lambda_3=0$. （4 分）

$\boldsymbol{A}$ 的属于 $\lambda_1=2$ 的线性无关的特征向量为

$$\boldsymbol{\eta}_1=(1,1,0)^{\mathrm{T}},\ \boldsymbol{\eta}_2=(0,0,1)^{\mathrm{T}};$$

$\boldsymbol{A}$ 的属于 $\lambda_3=0$ 的线性无关的特征向量为

$$\boldsymbol{\eta}_3=(-1,1,0)^{\mathrm{T}}.$$

易见 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3$ 两两正交.

将 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3$ 单位化，得

$$\mathrm{e}_1=\frac{1}{\sqrt{2}}(1,1,0)^{\mathrm{T}},\ \boldsymbol{e}_2,=(0,0,1)^{\mathrm{T}},\ \boldsymbol{e}_3,=\frac{1}{\sqrt{2}}(-1,1,0)^{\mathrm{T}},$$

取 $\boldsymbol{Q}=(e_1,e_2,e_3)$，则 Q 为正交矩阵. 令 $\boldsymbol{x}=\boldsymbol{Q}y$，得

$$f(x_1,x_2,x_3)=\lambda_1y_1^2+\lambda_2y_2^2+\lambda_3y_3^2=2y_1^2+2y_2^2.$$

（Ⅲ）**解法 1** 在正交变换 $\boldsymbol{x}=\boldsymbol{Q}y$ 下，$f(x_1,x_2,x_3)=0$ 化成 $2y_1^2+2y_2^2=0$，解之得 $y_1=y_2=0$，从而

$$\boldsymbol{x}=\boldsymbol{Q}\begin{pmatrix}0\\0\\y_3\end{pmatrix}=(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)\begin{pmatrix}0\\0\\y_3\end{pmatrix}=y_3\boldsymbol{e}_3=k(-1,1,0)^{\mathrm{T}},$$

其中 k 为任意常数. （9 分）

解法 2 由于 $f(x_1,x_2,x_3)=x_1^2+x_2^2+2x_3^2+2x_1x_2=(x_1+x_2)^2+2x_3^2=0$，所以

$$\begin{cases}x_1+x_2=0,\\ x_3=0,\end{cases}$$

其通解为 $\boldsymbol{x}=k(-1,1,0)^{\mathrm{T}}$，其中 k 为任意常数. （9 分）

(21)（本题满分 9 分）

已知 3 阶矩阵 $\boldsymbol{A}$ 的第一行是 (a,b,c)，a,b,c 不全为零，矩阵 $\boldsymbol{B}=\begin{pmatrix}1 & 2 & 3\\ 2 & 4 & 6\\ 3 & 6 & k\end{pmatrix}$（$k$ 为常数），且 $\boldsymbol{AB}=\boldsymbol{O}$，求线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解.

解 由于 $\boldsymbol{AB}=\boldsymbol{O}$，故 $r(\boldsymbol{A})+r(\boldsymbol{B})\leqslant 3$，又由 a,b,c 不全为零，可知 $r(A)\geqslant 1$.

当 $k\neq 9$ 时，$r(\boldsymbol{B})=2$，于是 $r(\boldsymbol{A})=1$；

当 $k=9$ 时，$r(\boldsymbol{B})=1$，于是 $r(\boldsymbol{A})=1$ 或 $r(\boldsymbol{A})=2$. （2 分）

对于 $k\neq 9$，由 $\boldsymbol{AB}=\boldsymbol{O}$ 可得

$$\boldsymbol{A}\begin{pmatrix}1\\2\\3\end{pmatrix}=\boldsymbol{0}\quad\text{和}\quad\boldsymbol{A}\begin{pmatrix}3\\6\\k\end{pmatrix}=\boldsymbol{0}.$$

由于 $\boldsymbol{\eta}_1=(1,2,3)^{\mathrm{T}}$，$\boldsymbol{\eta}_2=(3,6,k)^{\mathrm{T}}$ 线性无关，故 $\boldsymbol{\eta}_1$，$\boldsymbol{\eta}_2$ 为 $\boldsymbol{Ax}=\boldsymbol{0}$ 的一个基础解系，于是，$\boldsymbol{Ax}=\boldsymbol{0}$ 的通解为

$$\boldsymbol{x}=c_1\boldsymbol{\eta}_1+c_2\boldsymbol{\eta}_2\text{，其中 }c_1,c_2\text{ 为任意常数.}\qquad(5\text{ 分})$$

对于 $k=9$，分别就 $r(\boldsymbol{A})=2$ 和 $r(\boldsymbol{A})=1$ 进行讨论.

如果 $r(\boldsymbol{A})=2$，则 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系由一个向量构成. 又因为 $\boldsymbol{A}\begin{pmatrix}1\\2\\3\end{pmatrix}=\boldsymbol{0}$，所以，$\boldsymbol{Ax}=\boldsymbol{0}$ 的通解为 $\boldsymbol{x}=c_1(1,2,3)^{\mathrm{T}}$，其中 c_1 为任意常数.

如果 $r(\boldsymbol{A})=1$，则 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系由两个向量构成. 又因为 $\boldsymbol{A}$ 的第一行为 (a,b,c) 且 a,b,c 不全为零，所以 $\boldsymbol{Ax}=\boldsymbol{0}$ 等价于 $ax_1+bx_2+cx_3=0$. 不妨设 $a\neq0$，$\boldsymbol{\eta}_1=(-b,a,0)^{\mathrm{T}}$，$\boldsymbol{\eta}_2=(-c,0,a)^{\mathrm{T}}$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的两个线性无关的解，故 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解为

$$\boldsymbol{x}=c_1\boldsymbol{\eta}_1+c_2,c_2\text{ 为任意常数.}\qquad(9\text{ 分})$$

(22)（本题满分 9 分）

设二维随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}1, & 0<x<,0<y<2x,\\ 0 & \text{其他.}\end{cases}$$

求：(Ⅰ) (X,Y) 的边缘概率密度 $f_X(x)$，$f_Y(y)$；(Ⅱ) $Z=2X-Y$ 的概率密度 $f_z(z)$.

解　(Ⅰ) 当 $0<x<1$ 时，$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_0^{2x}\mathrm{d}y=2x$；

当 $x\leqslant0$ 或 $x\geqslant1$ 时，$f_X(x)=0$，即

$$f_X(x)=\begin{cases}2x & 0<x<1,\\ 0, & \text{其他.}\end{cases}\qquad(2\text{ 分})$$

当 $0<y<2$ 时，$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_{\frac{y}{2}}^{1}\mathrm{d}x=1-\frac{y}{2}$；

当 $y\leqslant0$ 或 $y\geqslant1$ 时，$f_Y(y)=0$，即

$$f_Y(y)=\begin{cases}1-\frac{z}{2}, & 0<z<2,\\ 0, & \text{其他.}\end{cases}\qquad(4\text{ 分})$$

(Ⅱ) 解法 1

当 $z\leqslant0$ 时，$F_Z(z)=0$；

当 $0<z<2$ 时，　$F_Z(z)=P\{2X-Y\leqslant z\}=\iint\limits_{2x-y\leqslant z}f(x,y)\mathrm{d}x\mathrm{d}y=z-\frac{z^2}{4}$；　(5 分)

当 $z\geqslant2$ 时，$F_Z(z)=1$，　(7 分)

所以

$$f_Z(z)=\begin{cases}1-\frac{z}{2}, & 0<z<2,\\ 0. & \text{其他.}\end{cases}\qquad(9\text{ 分})$$

解法 2

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,2x)-z)\mathrm{d}x,\qquad(6\text{ 分})$$

其中

$$f(x,2x-z)=\begin{cases}1, & 0<x<1,0<z<2x,\\ 0, & \text{其他.}\end{cases}$$

当 $z\leqslant0$ 或 $z\geqslant2$ 时，$f_Z(z)=0$；　(7 分)

当 $0<z<2$ 时，$F_Z(z)=\int_{\frac{z}{2}}^{1}\mathrm{d}x=1-\frac{z}{2}$，

即 $$F_Z(z)=\begin{cases}1-\frac{z}{2}, & 0<z<2,\\ 0, & \text{其他}.\end{cases}$$ (9 分)

(23)(本题满分 9 分)

设 $X_1,X_2,\cdots,X_n(n>2)$ 为来自总体 $N(0,1)$ 的简单随机样本,$\overline{X}$ 为样本均值,记 $Y_i=X_i-\overline{X},i=1,2,\cdots,n$.

求:(Ⅰ) Y_i 的方差 $DY_i,i=1,2,\cdots,n$;

(Ⅱ) Y_1 与 Y_n 的协方差 $\operatorname{cov}(Y_1,Y_n)$.

解 (Ⅰ) $DY_i=D(X_i-\overline{X})=D\left[\left(1-\frac{1}{n}\right)X_i-\frac{1}{n}\sum\limits_{\substack{k=1\\k\neq i}}^{n}X_k\right]$ (2 分)

$$=\frac{n-1}{n}(i=1,2,\cdots,n).$$ (4 分)

(Ⅱ) $\operatorname{cov}(Y_1,Y_n)=E(Y_1-EY_1)(Y_n-EY_n)$

$$=E(X_1-\overline{X})(X_n-\overline{X})=E(X_1X_n)+E(\overline{X}^2)-E(X_1\overline{X})-E(X_n\overline{X})$$

$$=EX_1EX_n+D\overline{X}-\frac{1}{n}E(X_1^2)-\frac{1}{n}\sum_{k=2}^{n}E(X_1X_k)-\frac{1}{n}E(X_n^2)-\frac{1}{n}\sum_{k=1}^{n-1}-\frac{1}{n}\sum_{k=1}^{n-1}E(X_kX_n)$$

$$=-\frac{1}{n}.$$ (9 分)

2005 年数学(二)试卷

一、填空题(本题共 6 小题,每小题 4 分,满分 24 分.)

(1) 设 $y=(1+\sin x)^x$,则 $\mathrm{d}y\big|_{x=\pi}=\underline{-\pi\mathrm{d}x}$.

(2) 曲线 $y=\frac{(1+x)^{\frac{3}{2}}}{\sqrt{x}}$ 的斜渐近线方程为$\underline{y=x+\frac{3}{2}}$.

(3) $\int_0^1\frac{x\mathrm{d}x}{(2-x^2)\sqrt{1-x^2}}=\underline{\frac{\pi}{4}}$.

(4) 微分方程 $xy'+2y=x\ln x$ 满足 $y(1)=-\frac{1}{9}$ 的解为$\underline{y=\frac{x}{3}\left(\ln x-\frac{1}{3}\right)}$.

(5) 当 $x\to 0$ 时,$\alpha(x)=kx^2$ 与 $\beta(x)=\sqrt{1+x\arcsin x}-\sqrt{\cos x}$ 是等价无穷小,则 $k=\underline{\frac{3}{4}}$.

(6) 同数学(一)第(5)题.

二、选择题(本题共 8 小题,每小题 4 分,满分 32 分.)

(7) 同数学(一)第(7)题

(8) 同数学(一)第(8)题.

(9) 设函数 $y=y(x)$ 由参数方程 $\begin{cases}x=t^2+2t,\\ y=\ln(1+t)\end{cases}$ 确定,则曲线 $y=y(x)$ 在 $x=3$ 处的法线与 x 轴交点的横坐标是 【A】

(A) $\frac{1}{8}\ln 2+3$; (B) $-\frac{1}{8}\ln 2+3$;

(C) $-8\ln 2+3$; (D) $8\ln 2+3$.

(10) 设区域 $D=\{(x,y)\mid x^2+y^2\leqslant 4,x\geqslant 0\}$,$f(x)$ 为 D 上的正值连续函数,a,b 为常数,则 $\iint\limits_D\frac{a\sqrt{f(x)}+b\sqrt{f(y)}}{\sqrt{f(x)}+\sqrt{f(y)}}\mathrm{d}\sigma=$ 【D】

(A) $ab\pi$; (B) $\frac{ab}{2}\pi$ (C) $(a+b)\pi$ (D) $\frac{a+b}{2}\pi$.

(11) 设函数 $u(x,y)=\varphi(x+y)+\varphi(x-y)+\int_{x-y}^{x+y}\psi\mathrm{d}t$,其中函数 φ 具有二阶导数,ψ 具有一阶导数,则必有 【B】

(A) $\frac{\partial^2 u}{\partial x^2} = -\frac{\partial^2 u}{\partial y^2}$；　　(B) $\frac{\partial u}{\partial x^2} = \frac{\partial^2 u}{\partial y^2}$；

(C) $\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y^2}$；　　(D) $\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial x^2}$.

(12) 设函数 $f(x) = \frac{1}{e^{\frac{x}{x-1}} - 1}$，则　　【D】

(A) $x = 0, x = 1$ 都是 $f(x)$ 的第一类间断点.

(B) $x = 0, x = 1$ 都是 $f(x)$ 的第二类间断点.

(C) $x = 0$ 是 $f(x)$ 的第一类间断点，$x = 1$ 是 $f(x)$ 的第二类间断点.

(D) $x = 0$ 是 $f(x)$ 的第二类间断点，$x = 1$ 是 $f(x)$ 的第一类间断点.

(13) 设 λ_1, λ_2 是矩阵 $\boldsymbol{A}$ 的两个不同的特征值，对应的特征向量分别为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$，则 $\boldsymbol{\alpha}_1, \boldsymbol{A}(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2)$ 线性无关的充分必要条件是　　【B】

(A) $\lambda_1 \neq 0$　　(B) $\lambda_2 \neq 0$　　(C) $\lambda_1 = 0$　　(D) $\lambda_2 = 0$.

(14) 设 $\boldsymbol{A}$ 为 $n(n \geqslant 2)$ 阶可逆矩阵，交换 $\boldsymbol{A}$ 的第 1 行与第 2 行，得矩阵 $\boldsymbol{B}_1, \boldsymbol{A}^*, \boldsymbol{B}^*$ 分别为 $\boldsymbol{A}, \boldsymbol{B}$ 的伴随矩阵，则　　【C】

(A) 交换 $\boldsymbol{A}^*$ 的第 1 列与第 2 列，得 $\boldsymbol{B}^*$；　　(B) 交换 $\boldsymbol{A}^*$ 的第 1 行与第 2 行得 $\boldsymbol{B}^*$；

(C) 交换 $\boldsymbol{A}^*$ 的第 1 列与第 2 列，得 $-\boldsymbol{B}^*$；　　(D) 交换 $\boldsymbol{A}^*$ 的第 1 行与第 2 行，得 $-\boldsymbol{B}$.

三、解答题(本题共 9 小题，满分 94 分.)

(15)(本题满分 11 分)

设函数 $f(x)$ 连续，且 $f(0) \neq 0$，求极限 $\lim\limits_{x \to 0} \frac{\int_0^x (x-t) f(t) \mathrm{d}t}{x \int_0^x f(x-t) \mathrm{d}t}$.

解　原式 $= \lim\limits_{x \to 0} \frac{x \int_0^x f(t) \mathrm{d}t - \int_0^x t f(t) \mathrm{d}t}{x \int_0^x f(x-t) \mathrm{d}t}$　　(2 分)

$\xrightarrow{\text{设 } x-t=u} \lim\limits_{x \to 0} \frac{x \int_0^x f(t) \mathrm{d}t - \int_0^x t f(t) \mathrm{d}t}{x \int_0^x f(u) \mathrm{d}u}$　　(4 分)

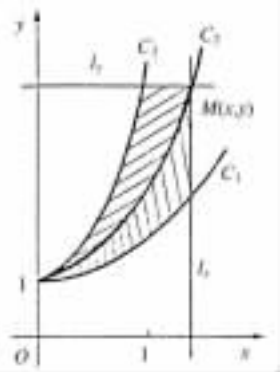
$\xrightarrow{\text{洛必达法则}} \lim\limits_{x \to 0} \frac{\int_0^x f(t) \mathrm{d}t + x f(x) - x f(x)}{\int_0^x f(u) \mathrm{d}u + x f(x)}$　　(7 分)

$\xrightarrow[(\xi \text{在 } 0 \text{ 和 } x \text{ 之间})]{\text{积分中值定理}} \lim\limits_{\substack{x \to 0 \\ (\xi \to 0)}} \frac{x f(\xi)}{x f(\xi) + x f(x)}$　　(10 分)

$= \frac{f(0)}{f(0) + f(0)} = \frac{1}{2}$.　　(11 分)

(16)(本题满分 11 分)

如图，C_1 和 C_2 分别是 $y = \frac{1}{2}(1 + e^x)$ 和 $y = e^x$ 的图像，过点 $(0,1)$ 的曲线 C_3 是一单调增函数的图像. 过 C_2 上任一点 $M(x,y)$ 分别作垂直于 x 轴和 y 轴的直线 l_x 和 l_y. 记 C_1, C_2 与 l_x 所围图形的面积为 $S_1(x)$；C_2, C_3 与 l_y 所围图形的面积为 $S_2(y)$. 如果总有 $S_1(x) = S_2(y)$，求曲线 C_3 的方程 $x = \varphi(y)$.



解　由题设 $S_1(x) = S_2(y)$，知

$$\int_0^x \left[e^x - \frac{1}{2}(1 + e^x) \right] \mathrm{d}x = \int_1^y [\ln y - \varphi(y)] \mathrm{d}y,$$

即
$$\int_0^x \left(\frac{1}{2} e^x - \frac{1}{2} \right) \mathrm{d}x = \int_1^y [\ln y - \varphi(y)] \mathrm{d}y,$$　　(5 分)

两边对 x 求导，得

$$\frac{1}{2}e^x-\frac{1}{2}=[\ln y-\varphi(y)]\frac{dy}{dx},\qquad (8\text{ 分})$$

由 $y=e^x$ 得

$$\frac{1}{2}e^x-\frac{1}{2}=[x-\varphi(e^x)]e^x,$$

于是
$$\varphi(e^x)=x+\frac{1}{2e^x}-\frac{1}{2},\qquad (10\text{ 分})$$

从而
$$\varphi(y)=\ln y+\frac{1}{2y}-\frac{1}{2},$$

故曲线 C_3 的方程为
$$x=\ln y+\frac{1}{2y}-\frac{1}{2}.\qquad (11\text{ 分})$$

(17) 同数学(一)第(17)题

(18)(本题满分 12 分)

用变量代换 $x=\cos t(0<t<\pi)$ 化简微分方程 $(1-x^2)y''-xy'+y=0$,并求其满足 $y|_{x=0}=1,y'|_{x=0}=2$ 的特解.

解
$$y'=\frac{dy}{dx}=\frac{dy}{dt}\cdot\frac{1}{\frac{dx}{dt}}=-\frac{1}{\sin t}\frac{dy}{dt},$$

$$y''=\frac{d^2y}{dx^2}=\frac{d}{dt}\left(\frac{dy}{dx}\right)\cdot\frac{1}{\frac{dx}{dt}}$$

$$=\left[\frac{\cos t}{\sin^2 t}\frac{dy}{dt}-\frac{1}{\sin t}\frac{d^2y}{dt^2}\right]\left(-\frac{1}{\sin t}\right)=\frac{1}{\sin^2 t}\frac{d^2y}{dt^2}-\frac{\cos t}{\sin^3 t}\frac{dy}{dt}\qquad (3\text{ 分})$$

将 y',y'' 代入原方程,得

$$(1-\cos^2 t)\left(\frac{1}{\sin^2 t}\frac{d^2y}{dt^2}-\frac{\cos t}{\sin^3 t}\frac{dy}{dt}\right)+\frac{\cos t}{\sin t}\frac{dy}{dt}+y=0,$$

即
$$\frac{d^2y}{dt^2}+y=0,\qquad (5\text{ 分})$$

其特征方程为 $\lambda^2+1=0$,解得 $\lambda=\pm i$,于是,此方程的通解为

$$y=C_1\cos t+C_2\sin t,\qquad (8\text{ 分})$$

从而原方程的通解为
$$y=C_1x+C_2\sqrt{1-x^2}.\qquad (10\text{ 分})$$

由 $y|_{x=0}=1,y'|_{x=0}=2$,得 $C_1=2,C_2=1$. 故所求方程的特解为

$$y=2x+\sqrt{1-x^2}.\qquad (12\text{ 分})$$

(19)(本题满分 12 分)

已知函数 $f(x)=[0,1]$ 上连续,在 $(0,1)$ 内可导,且 $f(0)=0,f(1)=1$. 证明:

(Ⅰ)存在 $\xi\in(0,1)$,使得 $f(\xi)=1-\xi$;

(Ⅱ)存在两个不同的点 $\eta,\zeta\in(0,1)$,使得 $f'(\eta)f'(\zeta)=1$.

证 (Ⅰ)令 $g(x)=f(x)+x-1$,则 $g(x)$ 在 $[0,1]$ 上连续,且

$$g(0)=-1<0,\ g(1)=1>0,\qquad (3\text{ 分})$$

所以,存在 $\xi\in(0,1)$,使得

$$g(\xi)=f(\xi)+\xi-1)=0,$$

即
$$f(\xi)=1-\xi.\qquad (5\text{ 分})$$

(Ⅱ)根据拉格朗日中值定理,存在 $\eta\in(0,\xi),\zeta\in(\xi,1)$,使得

$$f'(\eta)=\frac{f(\xi)-f(0)}{\xi}=\frac{1-\xi}{\xi},\qquad (8\text{ 分})$$

$$f'(\zeta)=\frac{f(1)-f(\xi)}{1-\xi}=\frac{1-(1-\xi)}{1-\xi}=\frac{\xi}{1-\xi},\qquad(11\text{ 分})$$

从而
$$f'(\eta)f'(\zeta)=\frac{1-\xi}{\xi}\cdot\frac{\xi}{1-\xi}=1.\qquad(12\text{ 分})$$

(20)(**本题满分** 10 **分**)

已知函数 $z=f(x,y)$ **的全微分** $\mathrm{d}z=2x\mathrm{d}x-2y\mathrm{d}y$,**并且** $f(1,1)=2$. **求** $f(x,y)$ **在椭圆域** $D=\left\{(x,y)\middle|x^2+\frac{y^2}{4}\leqslant 1\right\}$ **上的最大值和最小值**.

解法 1 **由** $\mathrm{d}z=2x\mathrm{d}x-2y\mathrm{d}y$ **可知**

$$z=f(x,y)=x^2-y^2+C,$$

再由 $f(1,1)=2$,**得** $C=2$,**故**

$$z=f(x,y)=x^2-y^2+2.\qquad(3\text{ 分})$$

令$\frac{\partial f}{\partial x}=2x=0$,$\frac{\partial f}{\partial y}=-2y=0$,**解得驻点**$(0,0)$.(5 **分**)

在椭圆 $x^2+\frac{y^2}{4}=1$ **上**,$z=x^2-(4-4x^2)+2$,**即**

$$z=5x^2-2\ (-1\leqslant x\leqslant 1),$$

其最大值为 $z|_{x=\pm1}=3$,**最小值为** $z|_{x=0}=-2$,(8 **分**)

再与 $f(0,0)=2$ **比较,可知** $f(x,y)$ **在椭圆域** D **上的最大值为** 3,**最小值为** -2.(10 **分**)

解法 2 **同解法** 1,**得驻点**$(0,0)$.(5 **分**)

用拉格朗日乘数法求此函数在椭圆 $x^2+\frac{y^2}{4}=1$ **上的极值**.

设
$$L=x^2-y^2+2+\lambda\left(x^2+\frac{y^2}{4}-1\right),\qquad(6\text{ 分})$$

令
$$\begin{cases}L'_x=2x+2\lambda x=0,\\ L'_y=-2y+\frac{\lambda}{2}y=0,\\ L'_\lambda=x^2+\frac{y^2}{4}-1=0,\end{cases}$$

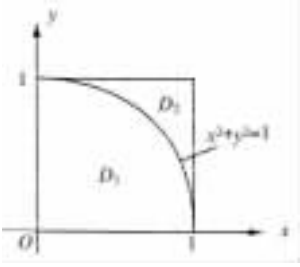
解得 4 **个可能的极值点**$(0,2),(0,-2),(1,0)$ **和**$(-1,0)$.(8 **分**)

又 $f(0,2)=-2,f(0,-2)=-2,f(1,0)=3,f(-1,0)=3$,**再与** $f(0,0)=2$ **比较,得** $f(x,y)$ **在** D **上的最大值为** 3,**最小值为** -2.(10 **分**)

(21)(**本题满分** 9 **分**)

计算二重积分$\iint\limits_D|x^2+y^2-1|\mathrm{d}\sigma$,**其中** $D=\{(x,y)\mid 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\}$.

解 **如图,将** D **分成** D_1 **与** D_2 **两部分**.



$$\iint\limits_D|x^2+y^2-1|\mathrm{d}\sigma=\iint\limits_{D_1}(1-x^2-y^2)\mathrm{d}\sigma+\iint\limits_{D_2}(x^2+y^2-1)\mathrm{d}\sigma,(2\text{ 分})$$

由于
$$\iint\limits_{D_1}(1-x^2-y^2)\mathrm{d}\sigma=\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\int_0^1(1-r^2)r\mathrm{d}r=\frac{\pi}{8},\qquad(4\text{ 分})$$

$$\iint\limits_{D_2}(x^2+y^2-1)\mathrm{d}\sigma=\int_0^1\mathrm{d}x\int_{\sqrt{1-x^2}}^1(x^2+y^2-1)\mathrm{d}y\qquad(6\text{ 分})$$

$$=\int_0^1\left[x^2y+\frac{y^3}{3}-y\right]_{\sqrt{1-x^2}}^1\mathrm{d}x$$

$$=\int_0^1\left[x^2-\frac{2}{3}+\frac{2}{3}(1-x^2)^{\frac{3}{2}}\right]\mathrm{d}x$$

$$=\int_0^1\left(x^2-\frac{2}{3}\right)\mathrm{d}x+\frac{2}{3}\int_0^1(1-x^2)^{\frac{3}{2}}\mathrm{d}x=-\frac{1}{3}+\frac{2}{3}I,$$

其中

$$I=\int_0^1(1-x^2)^{\frac{3}{2}}\mathrm{d}x\xlongequal{x=\sin t}\int_0^{\frac{\pi}{2}}\cos^4 t\mathrm{d}t=\frac{3}{4}\times\frac{1}{2}\times\frac{\pi}{2}=\frac{3\pi}{16},$$

$$\iint\limits_{D_2}(x^2+y^2-1)\mathrm{d}\sigma=-\frac{1}{3}+\frac{2}{3}\times\frac{3\pi}{16}=\frac{\pi}{8}-\frac{1}{3}.\quad(8\text{ 分})$$

因此

$$\iint\limits_{D}|x^2+y^2-1|\mathrm{d}\sigma=\frac{\pi}{8}+\frac{\pi}{8}-\frac{1}{3}=\frac{\pi}{4}-\frac{1}{3}.\quad(9\text{ 分})$$

(22)(**本题满分 9 分**)

确定常数 a,使向量组 $\boldsymbol{\alpha}_1=(1,1,a)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,a,1)^{\mathrm{T}},\boldsymbol{\alpha}_3=(a,1,1)^{\mathrm{T}}$ 可由向量组 $\boldsymbol{\beta}_1=(1,1,a)^{\mathrm{T}},\boldsymbol{\beta}_2=(-2,a,4)^{\mathrm{T}},\boldsymbol{\beta}_3=(-2,a,a)^{\mathrm{T}}$ 线性表示,但向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 不能由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示.

解法 1 记 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3),\boldsymbol{B}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$,由于 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示,故秩 $r(\boldsymbol{A})<3$,从而 $|\boldsymbol{A}|=-(a-1)^2(a+2)=0$,所以 $a=1$ 或 $a=-2$.(4 分)

当 $a=1$ 时,$\boldsymbol{\alpha}_1=\boldsymbol{\alpha}_2=\boldsymbol{\alpha}_3=\boldsymbol{\beta}_1(1,1,1)^{\mathrm{T}}$,故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性表示,但 $\boldsymbol{\beta}_2=(-2,1,4)^{\mathrm{T}}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示,所以 $a=1$ 符合题意.

当 $a=-2$ 时,由于

$$(\boldsymbol{B}\,|\,\boldsymbol{A})=\left(\begin{array}{ccc:ccc}1&-2&-2&1&1&-2\\1&-2&-2&1&-2&1\\-2&4&-2&-2&1&1\end{array}\right)\to\left(\begin{array}{ccc:ccc}1&-2&-2&1&1&-2\\0&0&-6&0&3&-3\\0&0&0&0&-3&3\end{array}\right),$$

考虑线性方程组 $\boldsymbol{Bx}=\boldsymbol{\alpha}_2$,因为秩 $r(\boldsymbol{B})=2$,秩 $r(\boldsymbol{B}\,|\,\boldsymbol{\alpha}_2)=3$,所以方程组 $\boldsymbol{Bx}=\boldsymbol{\alpha}_2$ 无解,即 $\boldsymbol{\alpha}_2$ 不能由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性表示,与题设矛盾.因此 $a=1$.

解法 2 记 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3),\boldsymbol{B}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$,对矩阵 $(\boldsymbol{A}\,|\,\boldsymbol{B})$ 施行初等行变换:

$$(A\,|\,B)=\left(\begin{array}{ccc:ccc}1&1&a&1&-2&-2\\1&a&1&1&a&a\\a&1&1&a&4&a\end{array}\right)\to\left(\begin{array}{ccc:ccc}1&1&a&1&-2&-2\\0&a-1&1-a&0&a+2&a+2\\0&1-a&1-a^2&0&4+2a&3a\end{array}\right)$$

$$\to\left(\begin{array}{ccc:ccc}1&1&a&1&-2&-2\\0&a-1&1-a&0&a+2&a+2\\0&0&-(a-1)(a+2)&0&3a+6&4a+2\end{array}\right),$$

由于 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示,故 $r(\boldsymbol{A})<3$,因此 $a=1$ 或 $a=-2$.(4 分)

当 $a=1$ 时,由于

$$(\boldsymbol{A}\,|\,\boldsymbol{B})=\left(\begin{array}{ccc:ccc}1&1&1&1&-2&-2\\1&1&1&1&1&1\\1&1&1&1&4&1\end{array}\right)\to\left(\begin{array}{ccc:ccc}1&1&1&1&-2&-2\\0&0&0&0&3&3\\0&0&0&0&0&-3\end{array}\right),$$

考虑线性方程组 $\boldsymbol{Ax}=\boldsymbol{\beta}_2$.由于秩 $r(\boldsymbol{A})=1$,秩 $r(\boldsymbol{A}\,|\,\boldsymbol{\beta}_2)=2$,故方程组 $\boldsymbol{Ax}=\boldsymbol{\beta}_2$ 无解,所以 $\boldsymbol{\beta}_2$ 不能由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示.另一方面,由于 $|\boldsymbol{B}|=-9\neq0$,故 $\boldsymbol{Bx}=\boldsymbol{\alpha}_i(i=1,2,3)$ 有惟一解,即 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性表示,所以 $a=1$ 符合题意.

当 $a=-2$ 时,

$$(\boldsymbol{A}\,|\,\boldsymbol{B})=\left(\begin{array}{ccc:ccc}1&1&-2&1&-2&-2\\1&-2&1&1&-2&-2\\-2&1&1&-2&4&-2\end{array}\right)\to\left(\begin{array}{ccc:ccc}1&1&-2&1&-2&-2\\0&-3&3&0&0&0\\0&0&0&0&0&-6\end{array}\right),$$

考虑线性方程组 $\boldsymbol{Bx}=\boldsymbol{\alpha}_2$.

$$(\boldsymbol{B}\,|\,\boldsymbol{\alpha}_2)=\left(\begin{array}{ccc:c}1&-2&-2&1\\1&-2&-2&-2\\-2&4&-2&1\end{array}\right)\to\left(\begin{array}{ccc:c}1&-2&-2&1\\0&0&0&-3\\0&0&-6&0\end{array}\right),$$

由于秩 $r(\boldsymbol{B})=2$，秩 $r(\boldsymbol{B}\mid\boldsymbol{\alpha}_2)=3$，故方程组 $\boldsymbol{Bx}=\boldsymbol{\alpha}_2$ 无解，即 $\boldsymbol{\alpha}_2$ 不能由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性表示，与题设矛盾. 因此 $a=1$. （9 分）

（23）（本题满分 9 分）

已知 3 阶矩阵 $\boldsymbol{A}$ 的第一行是 (a,b,c)，a,b,c 不全为零，矩阵 $\boldsymbol{B}=\begin{pmatrix}1&2&3\\2&4&6\\3&6&k\end{pmatrix}$（$k$ 为常数），且 $\boldsymbol{AB}=\boldsymbol{O}$，求线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解.

解　由于 $\boldsymbol{AB}=\boldsymbol{O}$，故 $r(\boldsymbol{A})+r(\boldsymbol{B})\leqslant 3$，又由 a,b,c 不全为零，可知 $r(\boldsymbol{A})\geqslant 1$.

当 $k\neq 0$ 时，$r(\boldsymbol{B})=2$，于是 $r(\boldsymbol{A})=1$；

当 $r=9$ 时，$r(\boldsymbol{B})=1$，于是 $r(\boldsymbol{A})=1$ 或 $r(A)=2$. （2 分）

对于 $k\neq 9$，由 $\boldsymbol{AB}=\boldsymbol{O}$ 可得

$$\boldsymbol{A}\begin{pmatrix}1\\2\\3\end{pmatrix}=\boldsymbol{0}\quad\text{和}\quad\boldsymbol{A}\begin{pmatrix}3\\6\\k\end{pmatrix}=\boldsymbol{0}.$$

由于 $\boldsymbol{\eta}_1=(1,2,3)^{\mathrm{T}}$，$\boldsymbol{\eta}_2=(3,6,k)^{\mathrm{T}}$ 线性无关，故 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2$ 为 $\boldsymbol{Ax}=\boldsymbol{0}$ 的一个基础解系，于是，$\boldsymbol{Ax}=\boldsymbol{0}$ 的通解为

$$\boldsymbol{x}=c_1\boldsymbol{\eta}_1+c_2\boldsymbol{\eta}_2,\text{其中 }c_1,c_2\text{ 为任意常数}.\quad（5\text{ 分}）$$

对于 $k=0$，分别就 $r(\boldsymbol{A})=2$ 和 $r(\boldsymbol{A})=1$ 进行讨论.

如果 $r(\boldsymbol{A})=2$，则 $\boldsymbol{Ax}=0$ 的基础解系由一个向量构成. 又因为 $\boldsymbol{A}\begin{pmatrix}1\\2\\3\end{pmatrix}=\boldsymbol{0}$，所以 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解为

$$\boldsymbol{x}=c_1(1,2,3)^{\mathrm{T}},\text{其中 }c_1\text{ 为任意常数}.\quad（7\text{ 分}）$$

如果 $r(\boldsymbol{A})=1$，则 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系由两个向量构成. 又因为 $\boldsymbol{A}$ 的第一行为 (a,b,c) 且 a,b,c 不全为零，所以 $\boldsymbol{Ax}=0$ 等价于 $ax_1+bx_2+cx_3=0$. 不妨设 $a\neq 0$，$\boldsymbol{\eta}_1=(-b,a,0)^{\mathrm{T}}$，$\boldsymbol{\eta}_2=(-c,0,a)^{\mathrm{T}}$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的两个线性无关的解，故 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解为

$$\boldsymbol{x}=c_1\boldsymbol{\eta}_1+c_2\boldsymbol{\eta}_2,\text{其中 }c_1,c_2\text{ 为任意常数}.\quad（9\text{ 分}）$$

2005 年数学（三）试卷

一、填空题（本题共 6 小题，每小题 4 分，满分 24 分.）

（1）极限 $\lim\limits_{x\to\infty}x\sin\dfrac{2x}{x^2+1}=$ $\underline{\quad 2\quad}$.

（2）微分方程 $xy'+y=0$ 满足初始条件 $y(1)=2$ 的特解为 $\underline{xy=2}$.

（3）设二元函数 $z=xe^{x+y}+(x+1)\ln(1+y)$，则 $\mathrm{d}z\mid_{(1,0)}=$ $\underline{2\mathrm{e}\mathrm{d}x+(\mathrm{e}+2)\mathrm{d}y}$.

（4）设行向量组 $(2,1,1,1),(2,1,a,a),(3,2,1,a),(4,3,2,1)$ 线性相关，且 $a\neq 1$，则 $a=$ $\underline{\dfrac{1}{2}}$.

（5）从数 1,2,3,4 中任取一个数，记为 X，再从 $1,\cdots,X$ 中任取一个数，记为 Y，则 $P\{Y=2\}=$ $\underline{\dfrac{13}{48}}$.

（6）设二维随机变量 (X,Y) 的概率分布为

X \ Y	0	1
0	0.4	a
1	b	0.1

若随机事件 $\{X=0\}$ 与 $\{X+Y=1\}$ 相互独立，则 $a=\underline{0.4}$，$b=\underline{0.1}$.

二、选择题（本题共 8 小题，每小题 4 分，满分 32 分.）

（7）当 a 取下列哪个值时，函数 $f(x)=2x^3-9x^2+12x-a$ 恰有两个不同的零点. 【B】

（A）2　（B）4　（C）6　（D）8

(8) 设 $I_1=\iint\limits_D\cos\sqrt{x^2+y^2}\,d\sigma$, $I_2=\iint\limits_D\cos(x^2+y^2)\,d\sigma$, $I_3=\iint\limits_D\cos(x^2+y^2)^2\,d\sigma$, 其中 $D=\{(x,y)\mid x^2+y^2\leqslant 1\}$, 则 【A】

(A) $I_3>I_2>I_1$; (B) $I_1>I_2>I_3$;

(C) $I_2>I_1>I_3$; (D) $I_3>I_1>I_2$.

(9) 设 $a_n>0$, $n=1,2,\cdots$, 若 $\sum\limits_{n=1}^{\infty}a_n$ 发散, $\sum\limits_{n=1}^{\infty}(-1)^{n-1}a_n$ 收敛, 则下列结论正确的是 【D】

(A) $\sum\limits_{n=1}^{\infty}a_{2n-1}$ 收敛, $\sum\limits_{n=1}^{\infty}a_{2n}$ 发散; (B) $\sum\limits_{n=1}^{n}a_{2n}$ 收敛, $\sum\limits_{n=1}^{\infty}a_{2n-1}$ 发散;

(C) $\sum\limits_{n=1}^{\infty}(a_{2n-1}+a_{2n})$ 收敛; (D) $\sum\limits_{n=1}^{\infty}(a_{2n-1}-a_{2n})$ 收敛.

(10) 设 $f(x)=x\sin x+\cos x$, 下列命题中正确的是 【B】

(A) $f(0)$ 是极大值, $f\left(\frac{\pi}{2}\right)$ 是极小值; (B) $f(0)$ 是极小值, $f\left(\frac{\pi}{2}\right)$ 是极大值;

(C) $f(0)$ 是极大值, $f\left(\frac{\pi}{2}\right)$ 也是极大值; (D) $f(0)$ 是极小值, $f\left(\frac{\pi}{2}\right)$ 也是极小值.

(11) 以下四个命题中, 正确的是 【C】

(A) 若 $f'(x)$ 在 $(0,1)$ 内连续, 则 $f(x)$ 在 $(0,1)$ 内有界;

(B) 若 $f(x)$ 在 $(0,1)$ 内连续, 则 $f(x)$ 在 $(0,1)$ 内有界;

(C) 若 $f'(x)$ 在 $(0,1)$ 内有界, 则 $f(x)$ 在 $(0,1)$ 内有界;

(D) 若 $f(x)$ 在 $(0,1)$ 内有界, 则 $f'(x)$ 在 $(0,1)$ 内有界.

(12) 设矩阵 $\boldsymbol{A}=(a_{ij})_{3\times 3}$ 满足 $\boldsymbol{A}^*=\boldsymbol{A}^{\mathrm{T}}$, 其中 $\boldsymbol{A}^*$ 为 $\boldsymbol{A}$ 的伴随矩阵, $\boldsymbol{A}^{\mathrm{T}}$ 为 $\boldsymbol{A}$ 的转置矩阵. 若 a_{11},a_{12},a_{13} 为三个相等的正数, 则 a_{11} 为 【A】

(A) $\frac{\sqrt{3}}{3}$ (B) 3 (C) $\frac{1}{3}$ (D) $\sqrt{3}$

(13) 设 λ_1,λ_2 是矩阵 $\boldsymbol{A}$ 的两个不同的特征值, 对应的特征向量分别为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$, 则 $\boldsymbol{\alpha}_1,\boldsymbol{A}(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)$ 线性无关的充分必要条件是 【D】

(A) $\lambda_1=0$ (B) $\lambda_2=0$ (C) $\lambda_1\neq 0$ (D) $\lambda_2\neq 0$

(14) 设一批零件的长度服从正态分布 $N(\mu,\sigma^2)$, 其中 μ,σ^2 均未知. 现从中随机抽取 16 个零件, 测得样本均值 $\bar{x}=20(\mathrm{cm})$, 样本标准差 $s=1(\mathrm{cm})$, 则 μ 的置信度为 0.90 的置信区间是 【C】【D】

(A) $\left(20-\frac{1}{4}t_{0.05}(16),20+\frac{1}{4}t_{0.05}(16)\right)$; (B) $\left(20-\frac{1}{4}t_{0.1}(16),20+\frac{1}{4}t_{0.1}(16)\right)$;

(C) $\left(20-\frac{1}{4}t_{0.05}(15),20+\frac{1}{4}t_{0.05}(15)\right)$; (D) $\left(20-\frac{1}{4}t_{0.1}(15),20+\frac{1}{4}t_{0.1}(15)\right)$.

三、解答题(本题共 9 小题, 满分 94 分.)

(15) (本题满分 8 分)

求 $\lim\limits_{x\to 0}\left(\frac{1+x}{1-e^{-x}}-\frac{1}{x}\right)$.

解 $\lim\limits_{x\to 0}\left(\frac{1+x}{1-e^{-x}}-\frac{1}{x}\right)=\lim\limits_{x\to 0}\frac{x+x^2-1+e^{-x}}{x(1-e^{-x})}$ (2 分)

$=\lim\limits_{x\to 0}\frac{x+x^2-1+e^{-x}}{x^2}$ (3 分)

$=\lim\limits_{x\to 0}\frac{1+2x-e^{-x}}{2x}$ (5 分)

$=\lim\limits_{x\to 0}\frac{2+e^{-x}}{2}$ (7 分)

$=\frac{3}{2}$.

(16) (本题满分 8 分)

设 $f(u)$ 具有二阶连续导数, 且 $g(x,y)=f\left(\frac{y}{x}\right)+yf\left(\frac{x}{y}\right)$, 求 $x^2\frac{\partial^2 g}{\partial x^2}-y^2\frac{\partial^2 g}{\partial y^2}$.

解 由已知条件可得

$$\frac{\partial g}{\partial x}=-\frac{y}{x^2}f'\left(\frac{y}{x}\right)+f'\left(\frac{x}{y}\right),\qquad(1\text{ 分})$$

$$\frac{\partial^2 g}{\partial x^2}=\frac{2y}{x^3}f'\left(\frac{y}{x}\right)+\frac{y^2}{x^4}f''\left(\frac{y}{x}\right)+\frac{1}{y}f''\left(\frac{x}{y}\right),\qquad(3\text{ 分})$$

$$\frac{\partial g}{\partial y}=\frac{1}{x}f'\left(\frac{y}{x}\right)+f\left(\frac{x}{y}\right)-\frac{x}{y}f'\left(\frac{x}{y}\right),\qquad(4\text{ 分})$$

$$\frac{\partial^2 g}{\partial y^2}=\frac{1}{x^2}f''\left(\frac{y}{x}\right)-\frac{x}{y^2}f'\left(\frac{x}{y}\right)+\frac{x}{y^2}f'\left(\frac{x}{y}\right)+\frac{x^2}{y^3}f''\left(\frac{x}{y}\right)=\frac{1}{x^2}f''\left(\frac{y}{x}\right)+\frac{x^2}{y^3}f''\left(\frac{x}{y}\right),\qquad(6\text{ 分})$$

所以

$$x^2\frac{\partial^2 g}{\partial x^2}-y^2\frac{\partial^2 g}{\partial y^2}=\frac{2y}{x}f'\left(\frac{y}{x}\right)+\frac{y^2}{x^2}f''\left(\frac{y}{x}\right)+\frac{x^2}{y}f''\left(\frac{x}{y}\right)-\frac{y^2}{x^2}f''\left(\frac{y}{x}\right)-\frac{x^2}{y}f''\left(\frac{x}{y}\right)=\frac{2y}{x}f'\left(\frac{y}{x}\right).$$

(8 分)

(17) **同数学(二)第(21)题.**

(18) **(本题满分 9 分)**

求幂级数 $\sum\limits_{n=1}^{\infty}\left(\frac{1}{2n+1}-1\right)x^{2n}$ **在区间**$(-1,1)$ **内的和函数** $S(x)$.

解　设

$$S(x)=\sum_{n=1}^{\infty}\left(\frac{1}{2n+1}-1\right)x^{2n},$$

$$S_1(x)=\sum_{n=1}^{\infty}\frac{x^{2n}}{2n+1},\ S_2(x)=\sum_{n=1}^{\infty}x^{2n},$$

则　$S(x)=S_1(x)-S_2(x),\ x\in(-1,1).$

由于

$$S_2(x)=\sum_{n=1}^{\infty}x^{2n}=\frac{x^2}{1-x^2},\qquad(3\text{ 分})$$

$$(xS_1(x))'=\sum_{n=1}^{\infty}x^{2n}=\frac{x^2}{1-x^2},x\in(-1,1),\qquad(5\text{ 分})$$

因此

$$xS_1(x)=\int_0^x\frac{t^2}{1-t^2}\mathrm{d}t=-x+\frac{1}{2}\ln\frac{1+x}{1-x}.\qquad(6\text{ 分})$$

又由于

$$S_1(0)=0,$$

故

$$S_1(x)=\begin{cases}-1+\dfrac{1}{2x}\ln\dfrac{1+x}{1-x}, & |x|\in(0,1),\\ 0, & x=0,\end{cases}\qquad(8\text{ 分})$$

所以

$$S(x)=S_1(x)-S_2(x)=\begin{cases}\dfrac{1}{2x}\ln\dfrac{1+x}{1-x}-\dfrac{1}{1-x^2}, & |x|\in(0,1),\\ 0, & x=0.\end{cases}\qquad(9\text{ 分})$$

(19) **(本题满分 8 分)**

设 $f(x),g(x)$ **在**$[0,1]$ **上的导数连续,且** $f(0)=0,f'(x)\geqslant 0$. **证明:对任何** $a\in[0,1]$**,有**

$$\int_0^a g(x)f'(x)\mathrm{d}x+\int_0^1 f(x)g'(x)\mathrm{d}x\geqslant f(a)g(1).$$

证法一　设

$$F(x)=\int_0^x g(t)f'(t)\mathrm{d}t+\int_0^1 f(t)g'(t)\mathrm{d}t-f(x)g(1),x\in[0,1],\qquad(2\text{ 分})$$

则 $F(x)$ **在**$[0,1]$ **上的导数连续,并且**

$$F'(x)=g(x)f'(x)-f'(x)g(1)=f'(x)[g(x)-g(1)],$$

由于 $x\in[0,1]$ **时,**$f'(x)\geqslant 0,g'(x)\geqslant 0$**,因此,**$F'(x)\leqslant 0$**,即** $F(x)$ **在**$[0,1]$ **上单调递减.**　(4 分)

注意到

$$F(1)=\int_0^1 g(t)f'(t)\mathrm{d}t+\int_0^1 f(t)g'(t)\mathrm{d}t-f(1)g(1),$$

而

$$\int_0^1 g(t)f'(t)\mathrm{d}t=\int_0^1 g(t)\mathrm{d}f(t)=g(t)f(t)\Big|_0^1-\int_0^1 f(t)g'(t)\mathrm{d}t=f(1)g(1)-\int_0^1 f(t)g'(t)\mathrm{d}t,\quad(7\text{ 分})$$

故 $F(1)=0$.

因此,$x\in[0,1]$ **时**,$F(x)\geqslant 0$,**由此可得对任何** $a\in[0,1]$,**有**

$$\int_0^a g(x)f'(x)\mathrm{d}x+\int_0^1 f(x)g'(x)\mathrm{d}x\geqslant f(a)g(1).\quad(8\text{ 分})$$

证法二 $\int_0^a g(x)f'(x)\mathrm{d}x=g(x)f(x)\Big|_0^a-\int_0^a f(x)g'(x)\mathrm{d}x=f(a)g(a)-\int_0^a f(x)g'(x)\mathrm{d}x,\quad(2\text{ 分})$

$$\int_0^a g(x)f'(x)\mathrm{d}x+\int_0^1 f(x)g'(x)\mathrm{d}x=f(a)g(a)-\int_0^a f(x)g'(x)\mathrm{d}x+\int_0^1 f(x)g'(x)\mathrm{d}x$$

$$=f(x)g(a)+\int_a^1 f(x)g'(x)\mathrm{d}x.\quad(4\text{ 分})$$

由于 $x\in[0,1]$ **时**,$f'(x)\geqslant 0$,**因此**,$f(x)$ **在**$[0,1]$ **上单调递增**,

$$f(x)\geqslant f(a),x\in[a,1],$$

又由于 $x\in[0,1]$ **时**,$g'(x)\geqslant 0$,**因此**

$$f(x)g'(x)\geqslant f(a)g'(x),x\in[a,1],$$

$$\int_a^1 f(x)g'(x)\mathrm{d}x\geqslant\int_a^1 f(a)g'(x)\mathrm{d}x=f(a)[g(1)-g(a)],\quad(7\text{ 分})$$

从而

$$\int_0^a g(x)f'(x)\mathrm{d}x+\int_0^1 f(x)g'(x)\mathrm{d}x\geqslant f(a)g(a)+f(a)[g(1)-g(a)]=f(a)g(1).\quad(8\text{ 分})$$

(20)(**本题满分** 13 **分**)

已知齐次线性方程组

$$(\text{ⅰ})\begin{cases}x_1+2x_2+3x_3=0,\\2x_1+3x_2+5x_3=0,\\x_1+x_2+ax_3=0,\end{cases}$$

和

$$(\text{ⅱ})\begin{cases}x_1+bx_2+cx_3=0,\\2x_1+b^2x_2+(c+1)x_3=0,\end{cases}$$

同解,求 a,b,c **的值.**

解　方程组(ⅱ)**的未知量个数大于方程的个数,故方程组**(ⅱ)**有无穷多个解.因为方程组**(ⅰ)**与**(ⅱ)**同解,所以方程组**(ⅰ)**的系数矩阵的秩小于** 3. (2 **分**)

对方程组(ⅰ)**的系数矩阵施以初等行变换:**

$$\begin{pmatrix}1&2&3\\2&3&5\\1&1&a\end{pmatrix}\to\begin{pmatrix}1&0&1\\0&1&1\\0&0&a-2\end{pmatrix},$$

从而 $a=2$. (5 **分**)

此时,方程组(ⅰ)**的系数矩阵可化为**

$$\begin{pmatrix}1&2&3\\2&3&5\\1&1&2\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix},$$

故$(-1,-1,1)^{\mathrm{T}}$是方程组(ⅰ)的一个基础解系. (7分)

将$x_1=-1,x_2=-1,x_3=1$代入方程组(ⅱ),可得

$$b=1,c=2 \text{ 或 } b=0,c=1. \tag{9分}$$

当$b=1,c=2$时,对方程组(ⅱ)的系数矩阵施以初等行变换,有

$$\begin{pmatrix}1&1&2\\2&1&3\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&1\\0&1&1\end{pmatrix},$$

故方程组(ⅰ)与(ⅱ)同解. (11分)

当$b=0,c=1$时,方程组(ⅱ)的系数矩阵可化为

$$\begin{pmatrix}1&0&1\\2&0&2\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&1\\0&0&0\end{pmatrix},$$

故方程组(ⅰ)与(ⅱ)的解不相同.

综上所述,当$a=2,b=1,c=2$时,方程组(ⅰ)与(ⅱ)同解. (13分)

(21)(本题满分13分)

设$\boldsymbol{D}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{C}\\\boldsymbol{C}^{\mathrm{T}}&\boldsymbol{B}\end{pmatrix}$为正定矩阵,其中$\boldsymbol{A},\boldsymbol{B}$分别为$m$阶,$n$阶对称矩阵,$\boldsymbol{C}$为$m\times n$矩阵.

(Ⅰ)计算$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{P}$,其中$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{E}_m&-\boldsymbol{A}^{-1}\boldsymbol{C}\\\boldsymbol{O}&\boldsymbol{E}_n\end{pmatrix}$;

(Ⅱ)利用(Ⅰ)的结果判断矩阵$B-C^{T}A^{-1}C$是否为正定矩阵,并证明你的结论.

解 (Ⅰ)因$\boldsymbol{P}^{\mathrm{T}}=\begin{pmatrix}\boldsymbol{E}_m&\boldsymbol{O}\\-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}&\boldsymbol{E}_n\end{pmatrix}$, (2分)

有

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{P}=\begin{pmatrix}\boldsymbol{E}_m&\boldsymbol{O}\\-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}&\boldsymbol{E}_n\end{pmatrix}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{C}\\\boldsymbol{C}^{\mathrm{T}}&\boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{E}_m&-A^{-1}\boldsymbol{C}\\\boldsymbol{O}&\boldsymbol{E}_n\end{pmatrix}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{C}\\\boldsymbol{O}&\boldsymbol{B}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{C}\end{pmatrix}\begin{pmatrix}\boldsymbol{E}_m&-\boldsymbol{A}^{-1}\boldsymbol{C}\\\boldsymbol{O}&\boldsymbol{E}_n\end{pmatrix} \tag{4分}$$

$$=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{C}\end{pmatrix}. \tag{6分}$$

(Ⅱ)矩阵$\boldsymbol{B}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{C}$是正定矩阵. (8分)

由(Ⅰ)的结果可知,矩阵D合同于矩阵

$$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{C}\end{pmatrix}.$$

又,$\boldsymbol{D}$为正定矩阵,可知矩阵$\boldsymbol{M}$为正定矩阵.

因矩阵$\boldsymbol{M}$为对称矩阵,故$\boldsymbol{B}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{C}$为对称矩阵. 对$X=(\underbrace{0,0,\cdots,0}_{m})^{\mathrm{T}}$及任意的$\boldsymbol{Y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}\neq\boldsymbol{0}$,有

$$(\boldsymbol{X}^{\mathrm{T}},\boldsymbol{Y}^{\mathrm{T}})\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{C}\end{pmatrix}\begin{pmatrix}\boldsymbol{X}\\\boldsymbol{Y}\end{pmatrix}>0, \tag{11分}$$

即$\boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{B}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{C})\boldsymbol{Y}>0$. 故$\boldsymbol{B}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{C}$为正定矩阵. (13分)

(22)(本题满分13分)

设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}1, & 0<x<1,0<y<2x,\\0, & \text{其他.}\end{cases}$$

求:(Ⅰ)(X,Y)的边缘概率密度$f_X(x),f_Y(y)$;

(Ⅱ)$Z=2X-Y$的概率密度$f_Z(z)$;

(Ⅲ)$P\left\{Y\leqslant\frac{1}{2}\middle|X\leqslant\frac{1}{2}\right\}$.

解 （Ⅰ）**当** $0<x<1$ **时**，

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_0^{2x}\mathrm{d}y=2x;$$

当 $x\leqslant 0$ **或** $x\geqslant 1$ **时**，$f_X(x)=0$，**故**

$$f_X(x)=\begin{cases}2x, & 0<x<1,\\ 0, & \text{其他}.\end{cases}\qquad (2\text{ 分})$$

当 $0<y<2$ **时**，

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\int_{\frac{y}{2}}^{1}\mathrm{d}x=1-\frac{y}{2};$$

当 $y\leqslant 0$ **或** $y\leqslant 2$ **时**，$f_Y(y)=0$，**故**

$$f_Y(y)=\begin{cases}1-\dfrac{y}{2} & 0<y<2,\\ 0, & \text{其他}.\end{cases}\qquad (4\text{ 分})$$

（Ⅱ）**解法 1**

当 $z\leqslant 0$ **时**，$F_Z(z)=0$；（5 **分**）

当 $z\geqslant 2$ **时**，$F_Z(z)=1$；（6 **分**）

当 $0<z<2$ **时**，

$$F_Z(z)=P\{2X-Y\leqslant z\}=\iint\limits_{2x-y\leqslant z}f(x,y)\mathrm{d}x\mathrm{d}y=z-\frac{z^2}{4};\qquad (8\text{ 分})$$

故

$$f_Z(z)=F_Z'=\begin{cases}1-\dfrac{z}{2}, & 0<z<2,\\ 0, & \text{其他}.\end{cases}\qquad (9\text{ 分})$$

解法 2

由于 $f_Z(z)=\int_{-\infty}^{+\infty}f(x,2x-z)\mathrm{d}x$，**其中**

$$f(x,2x-z)=\begin{cases}1, & 0<x<1,0<z<2x,\\ 0, & \text{其他},\end{cases}\qquad (6\text{ 分})$$

当 $z\leqslant 0$ **或** $z\geqslant 2$ **时**，$f_Z(z)=0$；（8 **分**）

当 $0<z<2$ **时**，$f_Z(z)=\int_{\frac{z}{2}}^{1}\mathrm{d}x=1-\frac{z}{2}$，

故

$$f_Z(z)=\begin{cases}1-\dfrac{z}{2}, & 0<z<2,\\ 0, & \text{其他}.\end{cases}\qquad (9\text{ 分})$$

（Ⅲ）
$$P\left\{Y\leqslant\frac{1}{2}\,\middle|\,X\leqslant\frac{1}{2}\right\}=\frac{P\left\{X\leqslant\frac{1}{2},Y\leqslant\frac{1}{2}\right\}}{P\left\{X\leqslant\frac{1}{2}\right\}}\qquad (11\text{ 分})$$

$$=\frac{\frac{3}{16}}{\frac{1}{4}}=\frac{3}{4}.\qquad (13\text{ 分})$$

（23）（**本题满分** 13 **分**）

设 $X_1,X_2,\cdots,X_n(n>2)$ **为来自总体** $N(0,\sigma^2)$ **的简单随机样本，其样本均值为** $\overline{X}$. **记** $Y_i=X_i-\overline{X}$，$i=1,2,\cdots,n$.

（Ⅰ）**求** Y_i **的方差** DY_i，$i=1,2,\cdots,n$；

(Ⅱ) 求 Y_1 与 Y_n 的协方差 $\mathrm{cov}(Y_1,Y_n)$；

(Ⅲ) 若 $c(Y_1+Y_n)^2$ 是 σ^2 的无偏估计量，求常数 c.

解 (Ⅰ) $DY_i = D(X_i-\overline{X}) = D\left[\left(1-\frac{1}{n}\right)X_i - \frac{1}{n}\sum_{k\neq i}X_k\right] = \frac{n-1}{n}\sigma^2, i=1,2,\cdots,n.$ (6 分)

(Ⅱ) $\mathrm{cov}(Y_1,Y_n) = E(Y_1-EY_1)(Y_n-EY_n)$ (5 分)

$= E(X_1-\overline{X})(X_n-\overline{X}) = E(X_1X_n)+E(\overline{X}^2)-E(X_1\overline{X})-E(X_n\overline{X})$ (7 分)

$= EX_1EX_n + D\overline{X} - \frac{1}{n}E(X_1^2) - \frac{1}{n}\sum_{i=2}^{n}E(X_1X_i) - \frac{1}{n}E(X_n^2) - \frac{1}{n}\sum_{i=1}^{n-1}E(X_iX_n)$

$= -\frac{1}{n}\sigma^2.$ (9 分)

(Ⅲ) $E[c(Y_1+Y_n)^2] = cD(Y_1+Y_n) = c[DY_1+DY_n+2\mathrm{cov}(Y_1,Y_n)]$ (11 分)

$= c\left[\frac{n-1}{n}+\frac{n-1}{n}-\frac{2}{n}\right]\sigma^2 = \frac{2(n-2)}{n}c\sigma^2 = \sigma^2,$

故 $c = \frac{n}{2(n-2)}.$ (13 分)

2005 年数学(四) 试卷

一、填空题(本题共 6 小题，每小题 4 分，满分 24 分.)

(1) 同数学(三) 第(1) 题.

(2) 同数学(三) 第(2) 题.

(3) 同数学(三) 第(3) 题.

(4) 同数学(三) 第(4) 题.

(5) 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 均为三维列向量，记三阶矩阵

$$\boldsymbol{A} = (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3),\ \boldsymbol{B} = (\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+4\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+9\boldsymbol{\alpha}_3).$$

如果 $|\boldsymbol{A}|=1$，那么，$|\boldsymbol{B}| =$ __2__.

(6) 同数学(三) 第(5) 题.

二、选择题(本题共 8 小题，每小题 4 分，满分 32 分.)

(7) 同数学(三) 第(7) 题.

(8) 同数学(三) 第(8) 题.

(9) 下列结论中正确的是 【D】

(A) $\int_1^{+\infty}\frac{\mathrm{d}x}{x(x+1)}$ 与 $\int_0^1\frac{\mathrm{d}x}{x(x+1)}$ 都收敛；　(B) $\int_1^{+\infty}\frac{\mathrm{d}x}{x(x+1)}$ 与 $\int_0^1\frac{\mathrm{d}x}{x(x+1)}$ 都发散；

(C) $\int_1^{+\infty}\frac{\mathrm{d}x}{x(x+1)}$ 发散，$\int_0^1\frac{\mathrm{d}x}{x(x+1)}$ 收敛；　(D) $\int_1^{+\infty}\frac{\mathrm{d}x}{x(x+1)}$ 收敛，$\int_0^1\frac{\mathrm{d}x}{x(x+1)}$ 发散；

(10) 同数学(三) 第(10) 题.

(11) 同数学(三) 第(11) 题.

(12) 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 均为 n 阶矩阵，$\boldsymbol{E}$ 为 n 阶单位矩阵，若 $\boldsymbol{B}=\boldsymbol{E}+\boldsymbol{AB},\boldsymbol{C}=\boldsymbol{A}+\boldsymbol{CA}$，则 $\boldsymbol{B}-\boldsymbol{C}$ 为 【A】

(A) $\boldsymbol{E}$　(B) $-\boldsymbol{E}$　(C) $\boldsymbol{A}$　(D) $-\boldsymbol{A}$.

(13) 设二维随机变量 (X,Y) 的概率分布为

X \ Y	0	1
0	0.4	a
1	b	0.1

若随机事件 $\{X=0\}$ 与 $\{X+Y=1\}$ 相互独立，则 【D】

(A) $a=0.2,b=0.3$；　(B) $a=0.1,b=0.4$；

(C) $a=0.3,b=0.2$；　(D) $a=0.4,b=0.1$.

(14) 设 $X_1, X_2, \cdots, X_n, \cdots$ 为独立同分布的随机变量列，且均服从参数为 $\lambda(\lambda > 1)$ 的指数分布，记 $\Phi(x)$ 为标准正态分布函数，则 【C】

(A) $\lim\limits_{n\to\infty} P\left\{\dfrac{\sum\limits_{i=1}^{n} X_i - n\lambda}{\lambda\sqrt{n}} \leqslant x\right\} = \Phi(x)$；

(B) $\lim\limits_{n\to\infty} P\left\{\dfrac{\sum\limits_{i=1}^{n} X_i - n\lambda}{\sqrt{n\lambda}} \leqslant x\right\} = \Phi(x)$；

(C) $\lim\limits_{n\to\infty} P\left\{\dfrac{\lambda\sum\limits_{i=1}^{n} X_i - n}{\sqrt{n}} \leqslant x\right\} = \Phi(x)$；

(D) $\lim\limits_{n\to\infty} P\left\{\dfrac{\sum\limits_{i=1}^{n} X_i - \lambda}{\sqrt{n\lambda}} \leqslant x\right\} = \Phi(x)$.

三、解答题(本题共 9 小题，满分 94 分.)

(15) 同数学(三)第(15)题.

(16) 同数学(三)第(16)题.

(17) 同数学(二)第(21)题.

(18) (本题满分 9 分)

求 $f(x,y) = x^2 - y^2 + 2$ 在椭圆域 $D = \left\{(x,y) \middle| x^2 + \dfrac{y^2}{4} \leqslant 1\right\}$ 上的最大值和最小值.

解法 1 由

$$\begin{cases} f'_x(x,y) = 2x = 0, \\ f'_y(x,y) = -2y = 0, \end{cases}$$

可求得 $f(x,y) = x^2 - y^2 + 2$ 在区域 $x^2 + \dfrac{y^2}{4} < 1$ 内的惟一驻点$(0,0)$. (2 分)

在椭圆 $x^2 + \dfrac{y^2}{4} = 1$ 上，

$$f(x,y) = x^2 - (4 - 4x^2) + 2 = 5x^2 - 2 \quad (-1 \leqslant x \leqslant 1), \tag{4 分}$$

求得椭圆域 D 上可能的最值

$$f(\pm 1, 0) = 3,\ f(0, \pm 2) = -2. \tag{6 分}$$

又由于 $f(0,0) = 2$，因此，$f(x,y)$ 在椭圆域 D 上的最大值为 3，最小值为 -2. (9 分)

解法 2 由

$$\begin{cases} f'_x(x,y) = 2x = 0, \\ f'_y(x,y) = -2y = 0, \end{cases}$$

可求得 $f(x,y) = x^2 - y^2 + 2$ 在区域 $x^2 + \dfrac{y^2}{4} < 1$ 内的惟一驻点$(0,0)$. (2 分)

又由于

$$A = f''_{xx}(0,0) = 2,\ B = f''_{xy}(0,0) = 0,\ C = f''_{yy}(0,0) = -2,\ B^2 - AC = 4 > 0,$$

因此，$(0,0)$ 不是极值点，$f(x,y)$ 在 D 上的最大值与最小值在椭圆 $x^2 + \dfrac{y^2}{4} = 1$ 上取得. (5 分)

在椭圆 $x^2 + \dfrac{y^2}{4} = 1$ 上，

$$f(x,y) = x^2 - (4 - 4x^2) + 2 = 5x^2 - 2\ (-1 \leqslant x \leqslant 1), \tag{7 分}$$

记

$$g(x) = 5x^2 - 2, (-1 \leqslant x \leqslant 1),$$

其最大值为 $g(\pm 1) = 3$，最小值为 $g(0) = -2$.

因此，$f(x,y)$ 在椭圆域 D 上的最大值为 3，最小值为 -2. (9 分)

解法 3 (在区域 $x^2 + \dfrac{y^2}{4} < 1$ 内同解法 2) 在椭圆 $x^2 + \dfrac{y^2}{4} = 1$ 上，用拉格朗日乘数法求最大值与最小值.

设 $$L = x^2 - y^2 + 2 + \lambda\left(x^2 + \frac{y^2}{4} - 1\right), \tag{6 分}$$

由
$$\begin{cases} L'_x = 2x + 2\lambda x = 0, \\ L'_y = -2y + \dfrac{\lambda}{2} y = 0, \\ L'_\lambda = x^2 + \dfrac{y^2}{4} - 1 = 0, \end{cases}$$

求得 4 个可能的极值点 $M_1(0,2), M_2(0,-2), M_3(1,0)$ 和 $M_4(-1,0)$.

又，$f(M_1) = -2, f(M_2) = -2, f(M_3) = 3, f(M_4) = 3$，得 $f(x,y)$ 在 D 上的最大值为 3，最小值为 -2.

(19) 同数学(三)第(19)题.

(20) 同数学(三)第(20)题.

(21)（本题满分 13 分）.

设 $\boldsymbol{A}$ 为三阶矩阵，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 是线性无关的三维列向量，且满足

$$\boldsymbol{A}\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3, \ \boldsymbol{A}\boldsymbol{\alpha}_1 = 2\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3, \ \boldsymbol{A}\boldsymbol{\alpha}_3 = 2\boldsymbol{\alpha}_2 + 3\boldsymbol{\alpha}_3.$$

(Ⅰ) 求矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)\boldsymbol{B}$；

(Ⅱ) 求矩阵 $\boldsymbol{A}$ 的特征值；

(Ⅲ) 求可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为对角矩阵.

解 (Ⅰ) 由题设条件，有

$$\boldsymbol{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)\begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 2 \\ 1 & 1 & 3 \end{pmatrix},$$

可知

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 2 \\ 1 & 1 & 3 \end{pmatrix}. \qquad (3\text{ 分})$$

(Ⅱ) 因为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 是线性无关的三维列向量，可知矩阵 $\boldsymbol{C} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$ 可逆，所以，$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{B}$，即矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似. 由此可得矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值. (5 分)

由

$$|\lambda\boldsymbol{E} - \boldsymbol{B}| = \begin{vmatrix} \lambda - 1 & 0 & 0 \\ -1 & \lambda - 2 & -2 \\ -1 & -1 & \lambda - 3 \end{vmatrix} = (\lambda - 1)^2(\lambda - 4) = 0,$$

得矩阵 B 的特征值，也即矩阵 $\boldsymbol{A}$ 的特征值：

$$\lambda_1 = \lambda_2 = 1, \ \lambda_3 = 4. \qquad (7\text{ 分})$$

(Ⅲ) 对应于 $\lambda_1 = \lambda_2 = 1$，解齐次线性方程组 $(\boldsymbol{E} - \boldsymbol{B})\boldsymbol{X} = \boldsymbol{0}$，得基础解系：

$$\boldsymbol{\xi}_1 = (-1,1,0)^{\mathrm{T}}, \ \boldsymbol{\xi}_2 = (-2,0,1)^{\mathrm{T}};$$

对应于 $\lambda_3 = 4$，解齐次线性方程组 $(4\boldsymbol{E} - \boldsymbol{B})\boldsymbol{X} = \boldsymbol{0}$，得基础解系：

$$\boldsymbol{\xi}_3 = (0,1,1)^{\mathrm{T}} \qquad (10\text{ 分})$$

令矩阵

$$\boldsymbol{Q} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) = \begin{pmatrix} -1 & -2 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

则

$$\boldsymbol{Q}^{-1}B\boldsymbol{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

因 $\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q} = \boldsymbol{Q}^{-1}\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}\boldsymbol{Q} = (\boldsymbol{C}\boldsymbol{Q})^{-1}A(\boldsymbol{C}\boldsymbol{Q})$，记矩阵

$$\boldsymbol{P}=\boldsymbol{CQ}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\begin{pmatrix}-1&-2&0\\1&0&1\\0&1&1\end{pmatrix}=(-\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,-2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3),\quad(13\text{ 分})$$

故 P 即为所求的可逆矩阵.

(22) **同数学(三)第(22)题.**

(23) **(本题满分 13 分)**

设 $X_1,X_2,\cdots,X_n(n>2)$ 为独立同分布的随机变量,且均服从 $N(0,1)$,记 $\overline{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$,$Y_i=X_i-\overline{X}$,$i=1,2,\cdots,n$. 求:

(Ⅰ) Y_i 的方差 DY_i,$i=1,2,\cdots,n$;

(Ⅱ) Y_1 与 Y_n 的协方差 $\mathrm{cov}(Y_1,Y_n)$;

(Ⅲ) $P\{Y_1+Y_n\leqslant 0\}$.

解 (Ⅰ) $DY_i=D(X_i-\overline{X})=D\left[\left(1-\frac{1}{n}\right)X_i-\frac{1}{n}\sum\limits_{k\neq i}X_k\right]$ (2 分)

$$=\frac{n-1}{n},i=1,2,\cdots,n.\quad(4\text{ 分})$$

(Ⅱ) $\mathrm{cov}(Y_1,Y_n)=E(Y_1-EY_1)(Y_n-EY_n)$ (5 分)

$$=E(X_1-\overline{X})(X_n-\overline{X})=E(X_1X_n)+E(\overline{X}^2)-E(X_1\overline{X})-E(X_n\overline{X})\quad(7\text{ 分})$$

$$=EX_1EX_n+D\overline{X}-\frac{1}{n}E(X_1^2)-\frac{1}{n}\sum_{i=2}^{n}E(X_1X_i)-\frac{1}{n}E(X_n^2)$$

$$-\frac{1}{n}\sum_{i=1}^{n-1}E(X_iX_n)$$

$$=-\frac{1}{n}.\quad(9\text{ 分})$$

(Ⅲ) $Y_1+Y_n=X_1-\overline{X}+X_n-\overline{X}=\frac{n-2}{n}X_1-\frac{2}{n}\sum\limits_{i=2}^{n-1}X_i+\frac{n-2}{n}X_n$,

上式是相互独立的正态随机变量的线性组合,所以,Y_1+Y_n 服从正态分布. (11 分)

由于

$$E(Y_1+Y_n)=0,$$

故

$$P\{Y_1+Y_n\leqslant 0\}=\frac{1}{2}.\quad(13\text{ 分})$$