荣获国家教委优秀教材二等奖

工科线性代数

崔荣泉　杨泮池　编

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

西安交通大学出版社
XI'AN JIAOTONG UNIVERSITY PRESS

工 科 线 性 代 数

崔荣泉　杨泮池　编

西安交通大学出版社

图书在版编目(CIP)数据

工科线性代数/崔荣泉,杨泮池编.—西安:西安交通大学出版社,2006.8
ISBN 7-5605-2295-5

Ⅰ.工...　Ⅱ.①崔...②杨...　Ⅲ.线性代数-高等学校-教材　Ⅳ.0151.2

中国版本图书馆 CIP 数据核字(2006)第 097226 号

书　　名　工科线性代数
编　　者　崔荣泉　杨泮池
出版发行　西安交通大学出版社
地　　址　西安市兴庆南路 25 号(邮编:710049)
电　　话　(029)82668357　82667874(发行部)
　　　　　(029)82668315　82669096(总编办)
印　　刷　西安建筑科技大学印刷厂
字　　数　129 千字
开　　本　850mm×1 168mm　1/32
印　　张　5.125
版　　次　2006 年 8 月第 1 版　2006 年 8 月第 1 次印刷
印　　数　0 001～3 500
书　　号　ISBN 7-5605-2295-5/O·251
定　　价　7.50 元

前　　言

随着计算机技术的迅速发展,解大型的线性方程组、求矩阵的特征值与特征向量等已经成为工程技术人员及其他领域科学工作者经常遇到的课题.而线性代数正是阐述这些问题的有关理论和方法的一门课程,它已成为高等工科院校大学生的一门必修课.

本书是依据1995年国家教委高教司修订的《线性代数课程教学基本要求》,参照国内外有关教材,并结合我们多年教学体会写成的讲义基础上编写而成的.本书具有以下特点:(1)紧密结合工科院校的情况,重视理论联系实际;(2)突出矩阵方法,注重学生能力的培养;(3)提供在计算机上实现线性代数计算所必需的数值方法(这些内容都标有*号);(4)问题引入直观,叙述简明易懂,条理清楚,例题丰富,便于自学.

本书中配有一定数量的习题,其中有些是全国硕士研究生入学试题。通过这些习题,可以加深对各章内容的理解,并掌握一定的解题方法和技巧.

在此,我们感谢西安建筑科技大学潘鼎坤教授,他详细审阅了原稿并提出许多宝贵的建设性意见.同时感谢刘林教授,任学明教授和黄泽民副教授,黄长钧副教授.

由于我们水平有限,错误和不妥之处在所难免,敬请读者批评指正,不胜感激.

编　者

2006年6月

目　　录

第 1 章　矩阵代数基础

矩阵是现代科学技术不可缺少的工具，特别是电子计算机出现之后，矩阵方法得到了更广泛的应用. 本章介绍矩阵的有关概念和矩阵代数基础.

1.1　矩阵概念

工程技术中的许多问题都归结为求解线性代数方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.1}$$

其中 x_i 为未知量，系数 a_{ij} 和右端 b_i 都是常数. 研究这类方程组何时无解，何时有解，以及有解时解的构造和求解的方法等构成了线性代数的重要内容.

方程组(1.1)的解取决于系数 a_{ij} 和右端的常数，若将方程组(1.1)左端的系数按其相对位置排成数表，并且括起来成为

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{1.2}$$

把它看作一个整体，这个矩形数表就是我们要研究的矩阵.

看一个关于矩阵的实例.

图 1.1 表示 A 国的两个机场 A_1、A_2 与 B 国的三个机场 B_1、B_2、B_3 间的通航关系. 图中的连线表示两机场间通航，连线上的数字表示航班数目. 图 1.1 表示的关系可用如下矩阵表示

$$\begin{array}{c} \\ A_1 \\ A_2 \end{array}\begin{array}{c} \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ \begin{pmatrix} 4 & 2 & 2 \\ 0 & 8 & 2 \end{pmatrix} \end{array} \qquad (1.3)$$

这个矩阵表达了 A 国与 B 国间的通航关系.

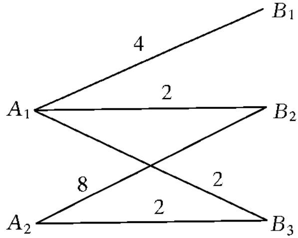


图 1.1　A 国与 B 国间航空网络

现在给出矩阵的一般定义.

定义 1.1　由 $m\times n$ 个元素 $a_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 排成矩形元素表(1.2),称这个元素表为 $m\times n$ **维矩阵**. 横的各排叫做矩阵的**行**,纵的各列叫做矩阵的**列**,a_{ij} 叫做矩阵第 i 行第 j 列的**元素**.

元素可以是数,也可以是函数,等等.

通常以大写字母 $\boldsymbol{A}$、$\boldsymbol{B}$、…等表示矩阵,矩阵(1.2)可以简记为 $\boldsymbol{A}=(a_{ij})_{m\times n}$,或 $\boldsymbol{A}=(a_{ij})$.

当 $m=n$ 即行数等于列数时,矩阵 $\boldsymbol{A}$ 称为 $\boldsymbol{n}$ **阶方阵**. 方阵的左上角到右下角的直线称为**主对角线**,主对角线上的元素 a_{ii} 叫做**主对角元素**.

只有一列的矩阵 $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$ 叫做 $\boldsymbol{m}$ **维列矩阵**或**列向量**,列向量常以小写字母 $\boldsymbol{a}$、$\boldsymbol{b}$ 或 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 等表示. 类似地称只有一行的矩阵$(b_1\ b_2 \cdots b_n)$为 $\boldsymbol{n}$ **维行矩阵**或**行向量**,它也常以小写字母表示. 向量中的元素称为**坐标**或**分量**. 行向量元素间可用逗号分开,以避免混淆,比如 $\boldsymbol{b}=(b_1,b_2,\cdots,b_n)$. 在这里指出了向量是特殊形式的矩阵,反过来也可以认为 $m\times n$ 维矩阵是由 n 个 m 维列向量组成,或是由 m 个 n 维行向量组成.

元素全为实数的矩阵称为**实矩阵**,元素中有复数的矩阵称为**复矩阵**. 特别地称元素全为零的矩阵为**零矩阵**,记为 $\boldsymbol{O}$.

今后还常常用到如下一些特殊方阵:

对角方阵

$$\boldsymbol{D}=\begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}$$

这是主对角线外元素全为零的方阵，简记为

$$\boldsymbol{D}=\mathrm{diag}(a_{11},a_{22},\cdots,a_{nn}).$$

若 $a_{11}=a_{22}=\cdots=a_{nn}$，则这种对角矩阵称为**数量矩阵**.

单位方阵

$$\boldsymbol{I}=\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

这是主对角元素全为 1 的对角方阵，简记为 $\boldsymbol{I}$（也常记为 $\boldsymbol{E}$）.

上三角矩阵 $\boldsymbol{U}$ 和下三角矩阵 $\boldsymbol{L}$

$$\boldsymbol{U}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix},\qquad \boldsymbol{L}=\begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

上三角矩阵 $\boldsymbol{U}$ 是主对角线以下元素全为零的方阵，下三角矩阵 $\boldsymbol{L}$ 是主对角线以上元素全为零的方阵.

本节介绍了矩阵的基本概念，矩阵这个词是英国数学家西尔威斯特(Sylvester)在 1850 年首先使用的，而矩阵记号则是英国数学家凯莱(Cayley)于 1855 年引进的.

1.2　矩阵基本运算

若矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 具有相同的行数与相同的列数，则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为**同型矩阵或同维矩阵**. 当同型矩阵 $\boldsymbol{A}=(a_{ij})$ 与 $\boldsymbol{B}=(b_{ij})$ 的所有对应元素相等，即 $a_{ij}=b_{ij}$ 时，称两**矩阵相等**，记为 $\boldsymbol{A}=\boldsymbol{B}$.

1.2.1 线性运算

1. 矩阵加法

同型矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 与 $\boldsymbol{B}=(b_{ij})_{m\times n}$ 的和 $\boldsymbol{A}+\boldsymbol{B}$ 规定为

$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

简记为 $\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij})$.

对于矩阵加法显然有

$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A},\quad (\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$$

2. 数与矩阵相乘

常数 k 与矩阵 $\boldsymbol{A}$ 的乘积 $k\boldsymbol{A}$(或 $\boldsymbol{A}k$)规定为

$$k\boldsymbol{A}=\boldsymbol{A}k=\begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}$$

简记为 $k\boldsymbol{A}=(ka_{ij})$.

$\boldsymbol{A}$ 的**负矩阵**(或**反矩阵**)记为 $-\boldsymbol{A}$,规定 $-\boldsymbol{A}=(-1)\boldsymbol{A}=(-a_{ij})$. 由此可规定同型矩阵 $\boldsymbol{A}=(a_{ij})$ 与 $\boldsymbol{B}=(b_{ij})$ 的减法为

$$\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{A}+(-\boldsymbol{B})=(a_{ij}-b_{ij})$$

数与矩阵相乘满足以下运算律(其中 k、l 为常数):

(1) $(kl)\boldsymbol{A}=k(l\boldsymbol{A})$;

(2) $(k+l)\boldsymbol{A}=k\boldsymbol{A}+l\boldsymbol{A}$;

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

例 1.1 设

$$\boldsymbol{A}=\begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

(1)求 $3\boldsymbol{A}+5\boldsymbol{B}$;

(2)解矩阵方程 $\boldsymbol{A}+\boldsymbol{X}=-\boldsymbol{B}$.

解

(1) $$3\boldsymbol{A}+5\boldsymbol{B}=3\begin{pmatrix}1 & -1 & 1\\ 0 & 2 & 0\end{pmatrix}+5\begin{pmatrix}0 & 1 & 0\\ 1 & -1 & 1\end{pmatrix}$$

$$=\begin{pmatrix}3 & -3 & 3\\ 0 & 6 & 0\end{pmatrix}+\begin{pmatrix}0 & 5 & 0\\ 5 & -5 & 5\end{pmatrix}=\begin{pmatrix}3 & 2 & 3\\ 5 & 1 & 5\end{pmatrix}$$

(2)由于 $\boldsymbol{A}+\boldsymbol{X}=-\boldsymbol{B}$,所以

$$\boldsymbol{X}=-\boldsymbol{B}-\boldsymbol{A}$$

$$=\begin{pmatrix}0 & -1 & 0\\ -1 & 1 & -1\end{pmatrix}+\begin{pmatrix}-1 & 1 & -1\\ 0 & -2 & 0\end{pmatrix}$$

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

1.2.2 矩阵的乘法和方阵的幂

1. 矩阵乘法

先看一个实例. 图 1.2 的航空网络图表明 A 国与 B 国的航班矩阵 $\boldsymbol{A}$ 和 B 国与 C 国的航班矩阵 $\boldsymbol{B}$ 分别为

$$\boldsymbol{A}=\begin{array}{c}\begin{array}{ccc}B_1 & B_2 & B_3\end{array}\\ \begin{pmatrix}4 & 2 & 6\\ 0 & 2 & 2\end{pmatrix}\end{array}\begin{array}{c}\\ A_1\\ A_2\end{array},\qquad \boldsymbol{B}=\begin{array}{c}\begin{array}{cc}C_1 & C_2\end{array}\\ \begin{pmatrix}2 & 2\\ 2 & 0\\ 0 & 8\end{pmatrix}\end{array}\begin{array}{c}\\ B_1\\ B_2\\ B_3\end{array}$$

试求由 A 国经 B 国到 C 国的航班矩阵 $\boldsymbol{C}$.

先求由 A_1 到 C_1 的航班数:由 A_1 经 B_1 到 C_1 的航班有 4×2 个;由 A_1 经 B_2 到 C_1 的航班有 2×2 个;由 A_1 经 B_3 到 C_1 的航班有 6×0 个;因此由 A_1 到达 C_1 的航班总数为

$$4\times2+2\times2+6\times0(\text{个})$$

同样可算出 A_1 到达 C_2 的航班总数为

$$4\times2+2\times0+6\times8(\text{个})$$

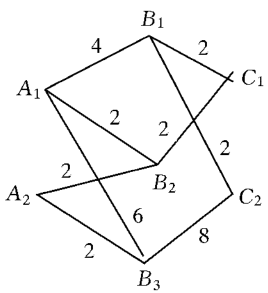


图 1.2 A 国经 B 国到 C 国的航空网络

由 A_2 到达 C_1 和 C_2 的航班总数的计算与上述算法相同. 总起来,航班矩阵 $\boldsymbol{C}$ 为

$$\boldsymbol{C} = \begin{array}{cc} C_1 & C_2 \\ \left(\begin{array}{cc} 4\times 2+2\times 2+6\times 0 & 4\times 2+2\times 0+6\times 8 \\ 0\times 2+2\times 2+2\times 0 & 0\times 2+2\times 0+2\times 8 \end{array}\right) & \begin{array}{c} A_1 \\ A_2 \end{array} \end{array}$$

$$= \begin{pmatrix} 12 & 56 \\ 4 & 16 \end{pmatrix} \begin{array}{c} A_1 \\ A_2 \end{array}$$

再看一个例子. 设变量 y 能用变量 x 线性地表示为

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{cases} \tag{1.4}$$

变量 x 能用变量 z 线性地表示为

$$\begin{cases} x_1 = b_{11}z_1 + b_{12}z_2 \\ x_2 = b_{21}z_1 + b_{22}z_2 \\ x_3 = b_{31}z_1 + b_{32}z_2 \end{cases} \tag{1.5}$$

其中系数 a_{ij} 和 b_{ij} 都是常数.

式(1.4)所表示的变量 x 到变量 y 的变换称为**线性变换**,同样式(1.5)表示了变量 z 到变量 x 的线性变换.

将式(1.5)代入式(1.4)得到变量 z 到变量 y 的线性变换

$$\begin{cases} y_1 = (a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31})z_1 + (a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32})z_2 \\ y_2 = (a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31})z_1 + (a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32})z_2 \end{cases} \tag{1.6}$$

若设式(1.4),式(1.5)的系数矩阵分别为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$,即

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$$

设式(1.6)的系数矩阵为 $\boldsymbol{C}$,则有

$$\boldsymbol{C} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}$$

由以上两例不难看出它们的共同点是,两例中矩阵 $\boldsymbol{C}$ 的元素

都是由矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的元素按同样计算规则得到. 我们把按这种计算规则得到的矩阵 $\boldsymbol{C}$,规定为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积.

定义 1.2 设矩阵 $\boldsymbol{A}=(a_{ij})_{m\times s}$, $\boldsymbol{B}=(b_{ij})_{s\times n}$, 如果矩阵 $\boldsymbol{C}=(c_{ij})$ 的元素规定为

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{is}b_{sj} = \sum_{k=1}^{s} a_{ik}b_{kj} \tag{1.7}$$

$$(i = 1,2,\cdots,m; j = 1,2,\cdots,n)$$

则称矩阵 $\boldsymbol{C}$ 为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的**乘积**,记为 $\boldsymbol{C}=\boldsymbol{AB}$.

注意:只有在矩阵 $\boldsymbol{A}$ 的列数与矩阵 $\boldsymbol{B}$ 的行数相同时,乘积 $\boldsymbol{AB}$ 才有意义,这时称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为可乘矩阵,乘积 $\boldsymbol{AB}=\boldsymbol{C}$ 的维数关系是

$$(a_{ij})_{m\times s}(b_{ij})_{s\times n}=(c_{ij})_{m\times n}$$

根据定义 1.2,航行例中航班矩阵 $\boldsymbol{C}$ 即为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积

$$\boldsymbol{C} = \boldsymbol{AB} = \begin{pmatrix} 4 & 2 & 6 \\ 0 & 2 & 2 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 2 & 0 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 12 & 56 \\ 4 & 16 \end{pmatrix}$$

线性变换例中连续两次线性变换的计算可用矩阵乘法表示为

$$\boldsymbol{C} = \boldsymbol{AB} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} \\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} \end{pmatrix}$$

由矩阵乘法规则可将线性方程组(1.1)表示为

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{1.8}$$

若以 $\boldsymbol{A}$ 表示系数矩阵, $\boldsymbol{x}$ 表示未知数列向量, $\boldsymbol{b}$ 表示常数列向量,那么线性方程组(1.1)或式(1.8)可简记为

$$\boldsymbol{Ax}=\boldsymbol{b} \tag{1.9}$$

这是线性方程组(1.1)的矩阵形式,这种简捷的表示法为以后的讨论带来许多方便.

矩阵乘法可以用来表示像连续的两次飞行、连续的两次线性变换这类问题,使得矩阵在工程技术中得到了广泛的应用.

例 1.2 (1)将微分方程组

$$\frac{\mathrm{d}x}{\mathrm{d}t}=y+z,\quad \frac{\mathrm{d}y}{\mathrm{d}t}=z+x,\quad \frac{\mathrm{d}z}{\mathrm{d}t}=x+y$$

写成矩阵形式;

(2)求出下式中待定的 a_{11}, a_{21}, a_{22}

$$x^2+2xy-2y^2=(x,y)\begin{pmatrix}a_{11} & 1\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}.$$

解 (1) 由于

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}= \quad y+z\\[2mm] \dfrac{\mathrm{d}y}{\mathrm{d}t}=x \quad +z\\[2mm] \dfrac{\mathrm{d}z}{\mathrm{d}t}=x+y\end{cases}$$

那么

$$\begin{pmatrix}\dfrac{\mathrm{d}x}{\mathrm{d}t}\\[2mm] \dfrac{\mathrm{d}y}{\mathrm{d}t}\\[2mm] \dfrac{\mathrm{d}z}{\mathrm{d}t}\end{pmatrix}=\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}$$

(2) $x^2+2xy-2y^2$

$$=(x,y)\begin{pmatrix}a_{11} & 1\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=(a_{11}x+a_{21}y, x+a_{22}y)\begin{pmatrix}x\\ y\end{pmatrix}$$

$$=a_{11}x^2+(a_{21}+1)xy+a_{22}y^2$$

比较两端的系数,得

$$a_{11}=1,\quad a_{21}=1,\quad a_{22}=-2.$$

例 1.3 下面各题中矩阵 $\boldsymbol{A}$、$\boldsymbol{B}$ 是否可作运算 $\boldsymbol{A}+\boldsymbol{B}$, $\boldsymbol{AB}$ 与

$\boldsymbol{BA}$，对可作的运算求出结果：

(1) $\boldsymbol{A}=\begin{pmatrix}1 & 1\\ -1 & -1\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}-1 & 1\\ 1 & -1\end{pmatrix}$；

(2) $\boldsymbol{A}=\begin{pmatrix}1 & 2 & 3\\ 1 & 3 & 6\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}1 & 1 & 1\\ 1 & 2 & 3\end{pmatrix}$；

(3) $\boldsymbol{A}=\begin{bmatrix}2 & -1\\ 1 & 0\\ -3 & 4\end{bmatrix}$，$\boldsymbol{B}=\begin{pmatrix}1 & -2\\ 3 & 4\end{pmatrix}$.

解 (1)可作运算 $\boldsymbol{A}+\boldsymbol{B}$，$\boldsymbol{AB}$ 与 $\boldsymbol{BA}$

$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}0 & 2\\ 0 & -2\end{pmatrix},\quad \boldsymbol{AB}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},\quad \boldsymbol{BA}=\begin{pmatrix}-2 & -2\\ 2 & 2\end{pmatrix}$$

(2)仅可作运算 $\boldsymbol{A}+\boldsymbol{B}$

$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}2 & 3 & 4\\ 2 & 5 & 9\end{pmatrix}$$

(3)仅可作运算 $\boldsymbol{AB}$

$$\boldsymbol{AB}=\begin{bmatrix}-1 & -8\\ 1 & -2\\ 9 & 22\end{bmatrix}$$

由(1)看出 $\boldsymbol{AB}\neq\boldsymbol{BA}$，即矩阵乘法不像数的乘法那样一定满足交换律．因此在谈到 $\boldsymbol{A}$ 乘 $\boldsymbol{B}$ 时，必须指明 $\boldsymbol{A}$ 左乘 $\boldsymbol{B}$ 还是右乘 $\boldsymbol{B}$．若矩阵 $\boldsymbol{A}$、$\boldsymbol{B}$ 满足 $\boldsymbol{AB}=\boldsymbol{BA}$，则称它们是关于乘法**可换**的．任何可换矩阵都是方阵．单位方阵与任何同阶的方阵关于乘法都可换，并且 $\boldsymbol{IA}=\boldsymbol{AI}=\boldsymbol{A}$.

由(1)还看出，虽然 $\boldsymbol{A}\neq\boldsymbol{O}$ 且 $\boldsymbol{B}\neq\boldsymbol{O}$，却有 $\boldsymbol{AB}=\boldsymbol{O}$．换句话说，$\boldsymbol{AB}=\boldsymbol{O}$，未必一定有 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{B}=\boldsymbol{O}$.

以上两点是矩阵乘法与数的乘法不同之处.

对于矩阵的加法与乘法运算有下列运算律：

结合律　　$(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC})$

左分配律　$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC}$

右分配律　$(\boldsymbol{B}+\boldsymbol{C})\boldsymbol{A}=\boldsymbol{BA}+\boldsymbol{CA}$

2. 方阵的幂

根据矩阵的乘法可以定义**方阵的幂**为

$$\boldsymbol{A}^1=\boldsymbol{A},\quad \boldsymbol{A}^2=\boldsymbol{A}\boldsymbol{A},\quad \cdots,\quad \boldsymbol{A}^{k+1}=\boldsymbol{A}^k\boldsymbol{A}\quad (k\text{ 为正整数}).$$

由此可以推出

$$\boldsymbol{A}^k\boldsymbol{A}^l=\boldsymbol{A}^{k+l},\quad (\boldsymbol{A}^k)^l=\boldsymbol{A}^{kl}\qquad (k,l\text{ 为正整数}).$$

注意，一般地 $(\boldsymbol{A}\boldsymbol{B})^k\neq\boldsymbol{A}^k\boldsymbol{B}^k$，例如

$$\boldsymbol{A}=\begin{pmatrix}1&0\\1&1\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}1&1\\0&1\end{pmatrix}$$

那么

$$(\boldsymbol{A}\boldsymbol{B})^2=\begin{pmatrix}2&3\\3&5\end{pmatrix},\qquad \boldsymbol{A}^2\boldsymbol{B}^2=\begin{pmatrix}1&2\\2&5\end{pmatrix},$$

可见 $$(\boldsymbol{A}\boldsymbol{B})^2\neq\boldsymbol{A}^2\boldsymbol{B}^2.$$

例 1.4 求证

$$\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}^n=\begin{pmatrix}\cos n\theta&-\sin n\theta\\\sin n\theta&\cos n\theta\end{pmatrix}$$

证 应用数学归纳法，$n=1$ 时等式显然成立，设 $n=k$ 时等式也成立，即有

$$\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}^k=\begin{pmatrix}\cos k\theta&-\sin k\theta\\\sin k\theta&\cos k\theta\end{pmatrix}$$

那么

$$\begin{aligned}\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}^{k+1}&=\begin{pmatrix}\cos k\theta&-\sin k\theta\\\sin k\theta&\cos k\theta\end{pmatrix}\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}\\&=\begin{pmatrix}\cos k\theta\cos\theta-\sin k\theta\sin\theta&-\cos k\theta\sin\theta-\sin k\theta\cos\theta\\\sin k\theta\cos\theta+\cos k\theta\sin\theta&-\sin k\theta\sin\theta+\cos k\theta\cos\theta\end{pmatrix}\\&=\begin{pmatrix}\cos(k+1)\theta&-\sin(k+1)\theta\\\sin(k+1)\theta&\cos(k+1)\theta\end{pmatrix}\end{aligned}$$

根据数学归纳法原理，等式得证.

本题结果与几何学中坐标轴旋转变换的结论一致，在几何学中只考虑坐标轴按逆时针旋转 θ 角，其转轴公式为

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

本例结果表明连续旋转 n 次 θ 角与一次旋转 $n\theta$ 角的变换一致.

1.2.3 逆矩阵

规定了矩阵的加法、减法和乘法之后，自然想到如何规定矩阵的除法. 还是回到数的除法，我们知道当 $a\neq 0$ 时，$b\div a=b\cdot 1/a=b\cdot a^{-1}$，这是把 b 除以 a 化为 b 与 a 的逆元素 a^{-1} 的乘法. 所谓 a 的逆元素 a^{-1} 是指满足 $aa^{-1}=a^{-1}a=1$ 的元素. 对于矩阵的"除法"可类似地处理，这就是用一个矩阵与另一个矩阵的逆矩阵相乘来代替两矩阵相"除"，因此应当像数的逆元素那样引出逆矩阵的概念.

定义 1.3 对于 n 阶方阵 $\boldsymbol{A}$，如果存在 n 阶方阵 $\boldsymbol{B}$ 使

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$$

则称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的**逆矩阵**，并说 $\boldsymbol{A}$ 是**可逆的**.

由定义知 $\boldsymbol{A}$ 也是 $\boldsymbol{B}$ 的逆矩阵，即 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 互为逆矩阵. 若 $\boldsymbol{A}$ 可逆，可以证明其逆矩阵唯一(见习题一 14(1))，于是 $\boldsymbol{A}$ 的逆矩阵记为 $\boldsymbol{A}^{-1}$，据定义 $\boldsymbol{AA}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$.

对于线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$，若 $\boldsymbol{A}$ 是可逆方阵，并能求出 $\boldsymbol{A}^{-1}$，那么对 $\boldsymbol{Ax}=\boldsymbol{b}$ 左乘 $\boldsymbol{A}^{-1}$，得 $\boldsymbol{A}^{-1}\boldsymbol{Ax}=\boldsymbol{A}^{-1}\boldsymbol{b}$，即

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

这就是线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解.

若 $\boldsymbol{A}$ 可逆，显然有性质：

(1) $(\boldsymbol{A}^{-1})^{-1}=\boldsymbol{A}$

(2) $(k\boldsymbol{A})^{-1}=\dfrac{1}{k}\cdot\boldsymbol{A}^{-1}(k\neq 0)$

(3) 若 $\boldsymbol{A},\boldsymbol{B}$ 皆为可逆矩阵，并且 $\boldsymbol{A},\boldsymbol{B}$ 同阶，则 $\boldsymbol{AB}$ 可逆，并且

$$(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

事实上 $\quad(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1})(\boldsymbol{AB})=\boldsymbol{B}^{-1}(\boldsymbol{A}^{-1}\boldsymbol{A})\boldsymbol{B}=\boldsymbol{B}^{-1}\boldsymbol{B}=\boldsymbol{I}$

及 $$(\boldsymbol{AB})(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1})=\boldsymbol{A}(\boldsymbol{BB}^{-1})\boldsymbol{A}^{-1}=\boldsymbol{AA}^{-1}=\boldsymbol{I}$$

根据逆矩阵定义

$$(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

一般地,有

$$(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_n)^{-1}=\boldsymbol{A}_n^{-1}\cdots\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}$$

对于可逆方阵 $\boldsymbol{A}$ 规定

$$\boldsymbol{A}^0=\boldsymbol{I},\qquad \boldsymbol{A}^{-k}=(\boldsymbol{A}^{-1})^k\qquad (k\text{ 为正整数})$$

由此

$$\boldsymbol{A}^k\boldsymbol{A}^l=\boldsymbol{A}^{k+l},\qquad (\boldsymbol{A}^k)^l=\boldsymbol{A}^{kl}\qquad (k,l\text{ 为整数})$$

例 1.5 验证

$$\boldsymbol{A}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\quad\text{与}\quad\boldsymbol{B}=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$$

互为逆矩阵.

解 由 $$\boldsymbol{AB}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

与

$$\boldsymbol{BA}=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

那么 $\boldsymbol{A},\boldsymbol{B}$ 互为逆矩阵.

注: 事实上,判定矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 互逆时,只需验证 $\boldsymbol{AB}=\boldsymbol{I}$ 或 $\boldsymbol{BA}=\boldsymbol{I}$之一成立即可,第 2 章例 2.10 给出这个结论的证明.

例 1.6 设方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 满足 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{AB}$,试证 $\boldsymbol{A}-\boldsymbol{I}$ 可逆.

证 由 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{AB}$,有

$$\boldsymbol{AB}-\boldsymbol{A}=\boldsymbol{B},\quad \text{即 }\boldsymbol{A}(\boldsymbol{B}-\boldsymbol{I})=\boldsymbol{B}$$

$$\boldsymbol{A}(\boldsymbol{B}-\boldsymbol{I})-\boldsymbol{I}=\boldsymbol{B}-\boldsymbol{I}$$

或 $$\boldsymbol{A}(\boldsymbol{B}-\boldsymbol{I})-(\boldsymbol{B}-\boldsymbol{I})=\boldsymbol{I}$$

于是 $$(\boldsymbol{A}-\boldsymbol{I})(\boldsymbol{B}-\boldsymbol{I})=\boldsymbol{I}$$

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

例 1.7 求解方程组 $\boldsymbol{Ax}=\boldsymbol{b}$，其中

$$\boldsymbol{A}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix},\quad \boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix}1\\1\\0\end{pmatrix}$$

解 由例 1.5 知

$$\boldsymbol{A}^{-1}=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$$

于是

$$\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}\begin{pmatrix}1\\1\\0\end{pmatrix}=\begin{pmatrix}0\\1\\1\end{pmatrix}$$

即方程组的解为

$$x_1=0,\quad x_2=x_3=1.$$

例 1.8 证明当 $ad-bc\neq0$ 时 $\boldsymbol{A}=\begin{pmatrix}a&b\\c&d\end{pmatrix}$可逆，并且

$$\boldsymbol{A}^{-1}=\frac{1}{ad-bc}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}$$

证 因 $ad-bc\neq0$，令

$$\boldsymbol{B}=\frac{1}{ad-bc}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}$$

那么

$$\begin{aligned}\boldsymbol{BA}&=\frac{1}{ad-bc}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}\\&=\frac{1}{ad-bc}\begin{pmatrix}ad-bc&0\\0&ad-bc\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}=\boldsymbol{I}\end{aligned}$$

同理可证 $\boldsymbol{AB}=\boldsymbol{I}$，因此 $\boldsymbol{A}$ 可逆，并且

$$\boldsymbol{A}^{-1}=\boldsymbol{B}=\frac{1}{ad-bc}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}.$$

这一段在不能定义矩阵除法的情况下给出了方阵的逆矩阵的概念，应当再次强调在矩阵运算中不能定义除法，这是因为若能定义矩阵 $\boldsymbol{A}$ 除矩阵 $\boldsymbol{B}$，就意味着可以找到满足关系式

$$\boldsymbol{AX}=\boldsymbol{B} \quad \text{或} \quad \boldsymbol{YA}=\boldsymbol{B}$$

的矩阵 $\boldsymbol{X}$ 或 $\boldsymbol{Y}$，一般来说这样的矩阵 $\boldsymbol{X}$、$\boldsymbol{Y}$ 未必存在，既便存在也未必唯一，因此这时谈论“矩阵 $\boldsymbol{A}$ 除矩阵 $\boldsymbol{B}$”就毫无意义. 并且我们还知道在实数或者复数集合中，$a\neq 0$ 时 $ax=b$ 有且仅有一个解 $x=b/a$，特别当 $b=0$ 时必有 $x=0$. 然而这种对于数的除法成立的结论对于矩阵并不成立. 即若 $\boldsymbol{A}$ 为非零矩阵而 $\boldsymbol{B}$ 是零矩阵时，满足 $\boldsymbol{AX}=\boldsymbol{B}$ 的 $\boldsymbol{X}$ 可以不是零矩阵，如

$$\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}\begin{pmatrix}0 & 0\\ 1 & 1\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

出现这种情形的原因在于，数 $a\neq 0$ 时，a 的逆元素 a^{-1} 一定存在，而非零方阵 $\boldsymbol{A}$ 的逆矩阵未必存在. 关于方阵满足什么条件时它的逆矩阵存在以及如何求出这个逆矩阵将在第 2 章叙述.

1.3 矩阵的转置及对称矩阵

1.3.1 矩阵的转置

定义 1.4 把矩阵 $\boldsymbol{A}$ 的每一行（或列）换成同序数的列（或行）得到的矩阵称为 $\boldsymbol{A}$ 的**转置矩阵**，记作 $\boldsymbol{A}^{\mathrm{T}}$.

例如 $$\boldsymbol{A}=\begin{pmatrix}-1 & 2\\ 0 & -1\\ 3 & 1\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}$$

则 $$\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix}-1 & 0 & 3\\ 2 & -1 & 1\end{pmatrix},\quad \boldsymbol{B}^{\mathrm{T}}=(b_1,b_2,\cdots,b_n)$$

矩阵的转置运算有以下性质：

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

(3) $(\lambda \boldsymbol{A})^{\mathrm{T}}=\lambda \boldsymbol{A}^{\mathrm{T}}$($\lambda$ 为数)　　　(4) $(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$

性质(1)、(2)、(3)容易证明,这里仅证性质(4).

设 $\boldsymbol{A}=(a_{ij})_{m\times n}$,$\boldsymbol{B}=(b_{ij})_{n\times s}$,欲证 $(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$,只须证明 $(\boldsymbol{AB})^{\mathrm{T}}$ 与 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 对应的第 i 行第 j 列元素相等.

$(\boldsymbol{AB})^{\mathrm{T}}$ 的第 i 行第 j 列元素等于 $\boldsymbol{AB}$ 的第 j 行第 i 列元素,即 $\boldsymbol{A}$ 的第 j 行与 $\boldsymbol{B}$ 的第 i 列对应元素乘积之和,就是

$$(a_{j1},a_{j2},\cdots,a_{jn})\begin{pmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{ni} \end{pmatrix}=\sum_{k=1}^{n} a_{jk}b_{ki}$$

而 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 的第 i 行第 j 列的元素是 $\boldsymbol{B}^{\mathrm{T}}$ 的第 i 行(即 $\boldsymbol{B}$ 的第 i 列)与 $\boldsymbol{A}^{\mathrm{T}}$ 的第 j 列(即 A 的第 j 行)对应元素乘积之和,就是

$$(b_{1i},b_{2i},\cdots,b_{ni})\begin{pmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jn} \end{pmatrix}=\sum_{k=1}^{n} b_{ki}a_{jk}=\sum_{k=1}^{n} a_{jk}b_{ki}$$

由此证得

$$(\boldsymbol{AB})^{\mathrm{T}}=\left(\sum_{k=1}^{n} a_{jk}b_{ki}\right)_{s\times m}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$$

例 1.9　设

$$\boldsymbol{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

求 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}$ 与 $\boldsymbol{a}\boldsymbol{b}^{\mathrm{T}}$.

解

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}=(a_1,a_2,\cdots,a_n)\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}=\sum_{k=1}^{n} a_k b_k \tag{1.10}$$

$$\boldsymbol{ab}^{\mathrm{T}}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1,b_2,\cdots,b_n)$$

$$=\begin{pmatrix}a_1b_1&a_1b_2&\cdots&a_1b_n\\a_2b_1&a_2b_2&\cdots&a_2b_n\\\vdots&\vdots&&\vdots\\a_nb_1&a_nb_2&\cdots&a_nb_n\end{pmatrix}\tag{1.11}$$

通常称式(1.10)为向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的**内积**.

例 1.10 证明,若 $\boldsymbol{A}$ 可逆,则 $\boldsymbol{A}^{\mathrm{T}}$ 可逆,且$(\boldsymbol{A}^{\mathrm{T}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}}$.

证 $\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}^{-1})^{\mathrm{T}}=(\boldsymbol{A}^{-1}\boldsymbol{A})^{\mathrm{T}}=\boldsymbol{I}^{\mathrm{T}}=\boldsymbol{I}$

$(\boldsymbol{A}^{-1})^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}=(\boldsymbol{A}\boldsymbol{A}^{-1})^{\mathrm{T}}=\boldsymbol{I}^{\mathrm{T}}=\boldsymbol{I}$

根据定义 $\boldsymbol{A}^{\mathrm{T}}$ 可逆,并且$(\boldsymbol{A}^{\mathrm{T}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}}$.

1.3.2 对称矩阵

定义 1.5 若矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$,称矩阵 $\boldsymbol{A}$ 为**对称矩阵**,若满足 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$,称 $\boldsymbol{A}$ 为**反对称矩阵**.

根据定义,若 $\boldsymbol{A}$ 为对称矩阵,则 $\boldsymbol{A}$ 的元素以其主对角线为对称,即 $a_{ij}=a_{ji}$(对一切 i,j). 若 $\boldsymbol{A}$ 为反对称矩阵,则 $a_{ij}=-a_{ji}$(对一切 i,j). 例如

$$\begin{pmatrix}1&4&3\\4&6&-5\\3&-5&2\end{pmatrix}\text{为对称矩阵},\begin{pmatrix}0&2&-3\\-2&0&-1\\3&1&0\end{pmatrix}\text{为反对称矩阵.}$$

对称矩阵和反对称矩阵都是方阵. 一个有趣的事实是:任何方阵 $\boldsymbol{A}$ 都可以表示为一个对称矩阵 $\boldsymbol{M}$ 与一个反对称矩阵 $\boldsymbol{S}$ 的和,即

$$\boldsymbol{A}=\boldsymbol{M}+\boldsymbol{S}$$

事实上,取 $\boldsymbol{M}=\frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})$,$\boldsymbol{S}=\frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}})$,显然有 $\boldsymbol{A}=\boldsymbol{M}+\boldsymbol{S}$,并且 $\boldsymbol{M}^{\mathrm{T}}=\frac{1}{2}(\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{A})=\boldsymbol{M}$,即 $\boldsymbol{M}$ 为对称矩阵,又 $\boldsymbol{S}^{\mathrm{T}}=\frac{1}{2}(\boldsymbol{A}^{\mathrm{T}}-$

$\boldsymbol{A})=-\boldsymbol{S}$,即 $\boldsymbol{S}$ 为反对称矩阵.

在空间解析几何中规定向量 $\boldsymbol{a}=a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k}$ 与向量 $\boldsymbol{b}=b_x\boldsymbol{i}+b_y\boldsymbol{j}+b_z\boldsymbol{k}$ 的向量积为

$$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ a_x & a_y & a_z\\ b_x & b_y & b_z\end{vmatrix}$$

$$=(a_yb_z-a_zb_y)\boldsymbol{i}+(a_zb_x-a_xb_z)\boldsymbol{j}+(a_xb_y-a_yb_x)\boldsymbol{k}$$

若记 $\boldsymbol{a}=(a_x,a_y,a_z)^{\mathrm{T}}$,$\boldsymbol{b}=(b_x,b_y,b_z)^{\mathrm{T}}$,则向量积 $\boldsymbol{a}\times\boldsymbol{b}$ 的矩阵表示为

$$\boldsymbol{a}\times\boldsymbol{b}=\begin{bmatrix}0 & -a_z & a_y\\ a_z & 0 & -a_x\\ -a_y & a_x & 0\end{bmatrix}\begin{bmatrix}b_x\\ b_y\\ b_z\end{bmatrix}$$

即向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的向量积 $\boldsymbol{a}\times\boldsymbol{b}$ 是由 $\boldsymbol{a}$ 的分量构成的一个反对称矩阵与向量 $\boldsymbol{b}$ 的乘积.

1.3.3 厄尔米特矩阵

定义 1.6 若 $\boldsymbol{A}=(a_{ij})$ 为复矩阵,则称 $\overline{\boldsymbol{A}}=(\overline{a_{ij}})$ 为 $\boldsymbol{A}$ 的**复共轭矩阵**. $\overline{a_{ij}}$ 是 a_{ij} 的共轭复数.

例如,

$$\text{若}\quad \boldsymbol{A}=\begin{pmatrix}1 & 1-\mathrm{i}\\ \mathrm{i} & 1+\mathrm{i}\end{pmatrix},\quad \text{则}\quad \overline{\boldsymbol{A}}=\begin{pmatrix}1 & 1+\mathrm{i}\\ -\mathrm{i} & 1-\mathrm{i}\end{pmatrix}$$

易知复共轭矩阵有性质:

(1)$\overline{\boldsymbol{A}+\boldsymbol{B}}=\overline{\boldsymbol{A}}+\overline{\boldsymbol{B}}$;　　(2)$\overline{k\boldsymbol{A}}=\overline{k}\ \overline{\boldsymbol{A}}$($k$ 为复数);

(3)$\overline{\boldsymbol{A}\boldsymbol{B}}=\overline{\boldsymbol{A}}\ \overline{\boldsymbol{B}}$;　　(4)$(\overline{\boldsymbol{A}^{\mathrm{T}}})=(\overline{\boldsymbol{A}})^{\mathrm{T}}$.

性质(4)表明 $\boldsymbol{A}$ 的**转置共轭矩阵**($\overline{\boldsymbol{A}^{\mathrm{T}}}$)等于 $\boldsymbol{A}$ 的**共轭转置矩阵**$(\overline{\boldsymbol{A}})^{\mathrm{T}}$. 记 $\boldsymbol{A}^*=(\overline{\boldsymbol{A}})^{\mathrm{T}}$,一般地 $\boldsymbol{A}^*\neq\boldsymbol{A}$,当 $\boldsymbol{A}^*=\boldsymbol{A}$ 时有如下定义.

定义 1.7 $\boldsymbol{A}$ 为复矩阵,若 $\boldsymbol{A}^*=\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为**厄尔米特(Hermite)矩阵**.

厄尔米特矩阵一定是方阵,而且 $\overline{a_{ij}}=a_{ji}$(对一切 i,j),若 $\boldsymbol{A}$ 为实的厄尔米特矩阵,它一定是实对称矩阵.

厄尔米特矩阵同样有共轭矩阵的四条性质，特别地，对于厄尔米特矩阵 $\boldsymbol{A}$ 有 $\overline{(\boldsymbol{A}^{\mathrm{T}})}=(\overline{\boldsymbol{A}})^{\mathrm{T}}=\boldsymbol{A}$.

类似地有定义：若 $\boldsymbol{A}^{*}=-A$，即对所有的 i,j 有 $\overline{a}_{ij}=-a_{ji}$，称 $\boldsymbol{A}$ 为**反厄尔米特矩阵**. 例如

$$\begin{pmatrix} 2 & 1+\mathrm{i} & 5-\mathrm{i} \\ 1-\mathrm{i} & 7 & \mathrm{i} \\ 5+\mathrm{i} & -\mathrm{i} & -1 \end{pmatrix} \quad 和 \quad \begin{pmatrix} 2\mathrm{i} & 1+\mathrm{i} & 5-\mathrm{i} \\ -1+\mathrm{i} & 7\mathrm{i} & \mathrm{i} \\ -5-\mathrm{i} & \mathrm{i} & -\mathrm{i} \end{pmatrix}$$

分别为厄尔米特矩阵和反厄尔米特矩阵. 由这个例子可以看出厄尔米特矩阵主对角线元素 a_{jj} 全为实数，反厄尔米特矩阵主对角线元素均为虚数，其实这不难由 $a_{jj}=\overline{a}_{jj}$ 与 $a_{jj}=-\overline{a}_{jj}$ 证得.

*1.3.4 矩阵分解

若

$$\boldsymbol{A}=\begin{pmatrix} 0 & 6 & 1 \\ 1 & 3 & 0 \\ 0 & 9 & 2 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 2 & 0 & 3 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$$

则称矩阵 $\boldsymbol{A}$ 分解为右端两个矩阵的积. 这种分解在线性方程组的求解中相当有用，这就是解线性方程组的 $\boldsymbol{LU}$ 分解法，它要求将方阵 $\boldsymbol{A}$ 分解为**单位下三角矩阵** $\boldsymbol{L}$（主对角线元素全为 1 的下三角矩阵）和上三角矩阵 $\boldsymbol{U}$ 的积，即 $\boldsymbol{A}=\boldsymbol{LU}$，设 $\boldsymbol{A}=(a_{ij})_{n\times n}$，那么 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 的形式为

$$\boldsymbol{L}=\begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \vdots & \vdots & & \ddots & \\ l_{n1} & l_{n2} & \cdots & \cdots & 1 \end{pmatrix}, \quad \boldsymbol{U}=\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix}$$

对 $\boldsymbol{A}$ 的这种分解称为 **$\boldsymbol{LU}$ 分解**或**三角分解**，当 $\boldsymbol{A}$ 满足一定条件[①]时，分解是唯一存在的.

① 这个条件是：方阵 $\boldsymbol{A}$ 的前主子式都不等于零. 前主子式的概念见 5.6 节.

由下面例子讨论如何进行矩阵的 $\boldsymbol{LU}$ 分解.

例 1.11 将矩阵

$$\boldsymbol{A}=\begin{pmatrix}1 & -3 & 7\\ 2 & 4 & -3\\ 3 & 7 & 2\end{pmatrix}$$

作 $\boldsymbol{LU}$ 分解.

解 设 $\boldsymbol{A}=\boldsymbol{LU}$,即

$$\begin{pmatrix}1 & -3 & 7\\ 2 & 4 & -3\\ -3 & 7 & 2\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ l_{21} & 1 & 0\\ l_{31} & l_{32} & 1\end{pmatrix}\begin{pmatrix}u_{11} & u_{12} & u_{13}\\ 0 & u_{22} & u_{23}\\ 0 & 0 & u_{33}\end{pmatrix}$$

$$=\begin{pmatrix}u_{11} & u_{12} & u_{13}\\ l_{21}u_{11} & l_{21}u_{12}+u_{22} & l_{21}u_{13}+u_{23}\\ l_{31}u_{11} & l_{31}u_{12}+l_{32}u_{22} & l_{31}u_{13}+l_{32}u_{23}+u_{33}\end{pmatrix}$$

对比上式两端的对应元素

第一行:$u_{11}=1,u_{12}=-3,u_{13}=7$;

第一列:$l_{21}u_{11}=2,l_{31}u_{11}=-3$;

将 $u_{11}=1$ 代入,求得 $l_{21}=2,l_{31}=-3$.

第二行:$l_{21}u_{12}+u_{22}=4,l_{21}u_{13}+u_{23}=-3$;

将已知的 l,u 值代入,求得 $u_{22}=10,u_{23}=-17$.

第二列:$l_{31}u_{12}+l_{32}u_{22}=7$,由此求得 $l_{32}=-1/5$;

第三行:$l_{31}u_{13}+l_{32}u_{23}+u_{33}=2$,求得 $u_{33}=98/5$.

于是求得 $\boldsymbol{A}$ 的 $\boldsymbol{LU}$ 分解

$$\begin{pmatrix}1 & -3 & 7\\ 2 & 4 & -3\\ -3 & 7 & 2\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ -3 & -1/5 & 1\end{pmatrix}\begin{pmatrix}1 & -3 & 7\\ 0 & 10 & -17\\ 0 & 0 & 98/5\end{pmatrix}$$

由本例知矩阵的 $\boldsymbol{LU}$ 分解的步骤是先求 $\boldsymbol{U}$ 的行再求 $\boldsymbol{L}$ 的列.

对 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 作 $\boldsymbol{LU}$ 分解的一般计算公式是

$$\begin{cases}u_{1j}=a_{1j} & j=1,2,\cdots,n\\ l_{i1}=a_{i1}/u_{11} & i=2,3,\cdots,n\end{cases}\tag{1.12}$$

当 $k=2,3,\cdots,n$ 时

$$
\begin{cases}
u_{kj} = a_{kj} - \sum\limits_{m=1}^{k-1} l_{km} u_{mj} & j = k, k+1, \cdots, n \\
l_{ik} = (a_{ik} - \sum\limits_{m=1}^{k-1} l_{im} u_{mk}) / u_{kk} & i = k+1, \cdots, n
\end{cases}
\tag{1.13}
$$

1.4 矩阵的分块

1.4.1 子矩阵

对于维数较高的矩阵，有时仅需考虑它的若干行与若干列相交处的元素（按相对位置）构成的矩阵，称它为原矩阵的**子矩阵**，如

$$
(1 \quad 3), \qquad \begin{pmatrix} 2 & 4 & 5 \\ 7 & 2 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 2 & 5 \\ 7 & 1 \\ -1 & 0 \end{pmatrix}
$$

都是矩阵

$$
\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 7 & 8 & 2 & 1 \\ 6 & -1 & 9 & 3 & 0 \end{pmatrix}
$$

的子矩阵.

当 $\boldsymbol{A}$ 为方阵时，一种重要的特殊子矩阵是由 $\boldsymbol{A}$ 的左上角元素开始，依次增加一行一列所构成的方阵，这些子矩阵称为方阵 $\boldsymbol{A}$ 的**前主子矩阵**. 前主子矩阵的构成是

$$
\boldsymbol{A} = \begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 0 & 2 \end{pmatrix}
$$

矩阵 $\boldsymbol{A}$ 的一切前主子矩阵是

$$
(4), \qquad \begin{pmatrix} 4 & 9 \\ 3 & 5 \end{pmatrix}, \qquad \begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 0 & 2 \end{pmatrix}
$$

1.4.2 矩阵的分块

为了简化高维矩阵的计算，常常采用分块方法，这就是用若干条纵线和横线将矩阵的行和列进行某种分划，使成一些矩形的子块（当然它们都是子矩阵），这时称以子块为元素的矩阵为**分块矩阵**. 显然分块矩阵的维数不超过原来矩阵的维数，每个子矩阵的维数也不超过原来矩阵的维数. 这种化高维矩阵为低维矩阵的分块方法常常是微型计算机用来克服容量不足的重要手段.

矩阵的分块形式多种多样，例如

$$\boldsymbol{A}=\left(\begin{array}{ccc:cc}1 & 0 & 2 & 3 & 5 \\ \hdashline 2 & 1 & 4 & 3 & 0 \\ 5 & 7 & 1 & 1 & 4\end{array}\right)=\begin{pmatrix}\boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{D} & \boldsymbol{E}\end{pmatrix}$$

这样分块的各子块

$$\boldsymbol{B}=(1\quad 0\quad 2),\qquad \boldsymbol{C}=(3\quad 5)$$

$$\boldsymbol{D}=\begin{pmatrix}2 & 1 & 4 \\ 5 & 7 & 1\end{pmatrix},\qquad \boldsymbol{E}=\begin{pmatrix}3 & 0 \\ 1 & 4\end{pmatrix}$$

还可分块为

$$\boldsymbol{A}=(\boldsymbol{A}_1,\boldsymbol{A}_2)$$

其中子块

$$\boldsymbol{A}_1=\begin{pmatrix}1 & 0 & 2 \\ 2 & 1 & 4 \\ 5 & 7 & 1\end{pmatrix},\qquad \boldsymbol{A}_2=\begin{pmatrix}3 & 5 \\ 3 & 0 \\ 1 & 4\end{pmatrix}$$

到底采用什么方式分块要根据矩阵的特点和需要而定，由于分块矩阵的计算法则与一般矩阵相同，还应当注意分块矩阵相加时子块应当同型，相乘时子块应当可乘.

分块矩阵的运算规则是

（1）对同维矩阵 $\boldsymbol{A},\boldsymbol{B}$ 作同样方式的分划，$\boldsymbol{A}=(\boldsymbol{A}_{rs})_{p\times q}$，$\boldsymbol{B}=(\boldsymbol{B}_{rs})_{p\times q}$，那么

$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_{11}+\boldsymbol{B}_{11} & \boldsymbol{A}_{12}+\boldsymbol{B}_{12} & \cdots & \boldsymbol{A}_{1q}+\boldsymbol{B}_{1q}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{p1}+\boldsymbol{B}_{p1} & \boldsymbol{A}_{p2}+\boldsymbol{B}_{p2} & \cdots & \boldsymbol{A}_{pq}+\boldsymbol{B}_{pq}\end{pmatrix}$$

(2) $k\boldsymbol{A}=\begin{pmatrix}k\boldsymbol{A}_{11} & k\boldsymbol{A}_{12} & \cdots & k\boldsymbol{A}_{1q}\\ \vdots & \vdots & & \vdots\\ k\boldsymbol{A}_{p1} & k\boldsymbol{A}_{p2} & \cdots & k\boldsymbol{A}_{pq}\end{pmatrix}$ （k 为常数）

(3)设 $\boldsymbol{A},\boldsymbol{B}$ 为可乘矩阵,分块成

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{p1} & \boldsymbol{A}_{p2} & \cdots & \boldsymbol{A}_{pt}\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1r}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{B}_{t1} & \boldsymbol{B}_{t2} & \cdots & \boldsymbol{B}_{tr}\end{pmatrix}$$

其中 $\boldsymbol{A}_{i1},\boldsymbol{A}_{i2},\cdots,\boldsymbol{A}_{it}$ 的列数分别等于 $\boldsymbol{B}_{1j},\boldsymbol{B}_{2j},\cdots,\boldsymbol{B}_{tj}$ 的行数,那么

$$\boldsymbol{AB}=\begin{pmatrix}\boldsymbol{C}_{11} & \boldsymbol{C}_{12} & \cdots & \boldsymbol{C}_{1r}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{C}_{p1} & \boldsymbol{C}_{p2} & \cdots & \boldsymbol{C}_{pr}\end{pmatrix}$$

其中子块

$$\boldsymbol{C}_{ij}=\sum_{k=1}^{t}\boldsymbol{A}_{ik}\boldsymbol{B}_{kj}\qquad (i=1,2,\cdots,p;j=1,2,\cdots,r)$$

注: 对矩阵乘积 $\boldsymbol{AB}$ 作分块计算时,若限定 $\boldsymbol{A}$ 不分块,则 $\boldsymbol{B}$ 只能按列分块为 $\boldsymbol{B}=(\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_k)$,作如下分块计算

$$\boldsymbol{AB}=\boldsymbol{A}(\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_k)=(\boldsymbol{AB}_1,\boldsymbol{AB}_2,\cdots,\boldsymbol{AB}_k),$$

若限定 $\boldsymbol{B}$ 不分块,则 $\boldsymbol{A}$ 只能按行分块并作如下分块计算

$$\boldsymbol{AB}=\begin{pmatrix}\boldsymbol{A}_1\\ \boldsymbol{A}_2\\ \vdots\\ \boldsymbol{A}_t\end{pmatrix}\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_1\boldsymbol{B}\\ \boldsymbol{A}_2\boldsymbol{B}\\ \vdots\\ \boldsymbol{A}_t\boldsymbol{B}\end{pmatrix}.$$

例 1.12 设 $\boldsymbol{A}$ 为 n 阶方阵,试证:若任意 n 维列向量 $\boldsymbol{x}$ 都满足 $\boldsymbol{Ax}=\boldsymbol{0}$,则 $\boldsymbol{A}=\boldsymbol{O}$.

证 取 $\boldsymbol{x}$ 为 n 维单位坐标向量

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \quad \boldsymbol{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

这里的 $\boldsymbol{e}_i$ 是第 i 个分量(元素)为 1,其余分量为 0 的向量. 根据题设 应有

$$\boldsymbol{A}\boldsymbol{e}_1 = \boldsymbol{0}, \quad \boldsymbol{A}\boldsymbol{e}_2 = \boldsymbol{0}, \cdots, \quad \boldsymbol{A}\boldsymbol{e}_n = \boldsymbol{0}$$

即 $\boldsymbol{A}(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n) = \boldsymbol{O}$ 或 $\boldsymbol{A}\boldsymbol{I} = \boldsymbol{O}$

由此立得 $\boldsymbol{A} = \boldsymbol{O}$.

注: n 维单位坐标向量也常以行向量形式表达.

(4)若

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{p1} & \boldsymbol{A}_{p2} & \cdots & \boldsymbol{A}_{pt} \end{pmatrix}$$

则

$$\boldsymbol{A}^{\mathrm{T}} = \begin{pmatrix} \boldsymbol{A}_{11}^{\mathrm{T}} & \boldsymbol{A}_{21}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{p1}^{\mathrm{T}} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{1t}^{\mathrm{T}} & \boldsymbol{A}_{2t}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{pt}^{\mathrm{T}} \end{pmatrix}$$

可见求分块矩阵 $\boldsymbol{A}$ 的转置 $\boldsymbol{A}^{\mathrm{T}}$ 是先将 $\boldsymbol{A}$ 的行列互换,同时将各子块转置.

例 1.13 求矩阵乘积 $\boldsymbol{AB}$,其中

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

解 把 $\boldsymbol{A}, \boldsymbol{B}$ 分块成

$$\boldsymbol{A} = \left(\begin{array}{cc:cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hdashline -1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{A}_{21} & \boldsymbol{I} \end{pmatrix}$$

$$\boldsymbol{B} = \left(\begin{array}{cc:cc} 1 & 0 & 1 & 0 \\ -1 & 2 & 0 & 1 \\ \hdashline 1 & 0 & 4 & 1 \\ -1 & -1 & 2 & 0 \end{array}\right) = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{I} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix}$$

那么

$$\begin{aligned} \boldsymbol{AB} &= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{A}_{21} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{I} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{I} \\ \boldsymbol{A}_{21}\boldsymbol{B}_{11} + \boldsymbol{B}_{21} & \boldsymbol{A}_{21} + \boldsymbol{B}_{22} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 2 & 0 & 1 \\ -2 & 4 & 3 & 3 \\ -1 & 1 & 3 & 1 \end{pmatrix} \end{aligned}$$

例 1.14 若

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

解 将 $\boldsymbol{A}$ 分块成分块对角阵

$$\boldsymbol{A} = \left(\begin{array}{cc:c:ccc} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 2 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 & -3 \end{array}\right) = \begin{pmatrix} \boldsymbol{A}_1 & & \\ & \boldsymbol{A}_2 & \\ & & \boldsymbol{A}_3 \end{pmatrix}$$

于是

$$\boldsymbol{A}^2 = \begin{pmatrix} \boldsymbol{A}_1^2 & & \\ & \boldsymbol{A}_2^2 & \\ & & \boldsymbol{A}_3^2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & & & & \\ 0 & 1 & & & & \\ & & 4 & & & \\ & & & 9 & -6 & 1 \\ & & & 0 & 9 & -6 \\ & & & 0 & 0 & 9 \end{pmatrix}.$$

由此可见，将矩阵划分成分块对角阵可简化计算，一般地对分块对角阵 $\boldsymbol{D}=\mathrm{diag}[\boldsymbol{D}_1,\boldsymbol{D}_2,\cdots,\boldsymbol{D}_n]$，有

(1) $\boldsymbol{D}^k=\mathrm{diag}[\boldsymbol{D}_1^k,\boldsymbol{D}_2^k,\cdots,\boldsymbol{D}_n^k]$；

(2) $\boldsymbol{D}^{\mathrm{T}}=\mathrm{diag}[\boldsymbol{D}_1^{\mathrm{T}},\boldsymbol{D}_2^{\mathrm{T}},\cdots,\boldsymbol{D}_n^{\mathrm{T}}]$；

(3) $\boldsymbol{D}^{-1}=\mathrm{diag}[\boldsymbol{D}_1^{-1},\boldsymbol{D}_2^{-1},\cdots,\boldsymbol{D}_n^{-1}]$.

这里要求(1)，(3)中的 $\boldsymbol{D}_i(i=1,2,\cdots,n)$均为方阵，并且(3)中的 $\boldsymbol{D}_i^{-1}(i=1,2,\cdots,n)$均存在.

*1.5 矩阵的微分与积分

若矩阵 $\boldsymbol{A}=(a_{ij}(t))$中各元素是 t 的可微函数，则定义**矩阵的导数**为

$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} = \left(\frac{\mathrm{d}a_{ij}(t)}{\mathrm{d}t}\right)$$

矩阵的微分

$$\mathrm{d}\boldsymbol{A} = (\mathrm{d}a_{ij}(t))$$

例如

$$\boldsymbol{A} = \begin{pmatrix} t^2 & t^3 \\ 2 & \sin t \end{pmatrix}$$

则

$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} = \begin{pmatrix} 2t & 3t^2 \\ 0 & \cos t \end{pmatrix}, \quad \mathrm{d}\boldsymbol{A} = \begin{pmatrix} 2t\mathrm{d}t & 3t^2\mathrm{d}t \\ 0 & \cos t\mathrm{d}t \end{pmatrix}$$

矩阵的和与乘积的导数法则是

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}+\boldsymbol{B}) = \frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}+\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}\boldsymbol{B}) = \frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\boldsymbol{B}+\boldsymbol{A}\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}$$

$\boldsymbol{A}^{-1}$的导数法则是

$$\frac{\mathrm{d}(\boldsymbol{A}^{-1})}{\mathrm{d}t}=-\boldsymbol{A}^{-1}\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\boldsymbol{A}^{-1}.$$

例 1.15 化三阶微分方程

$$\frac{\mathrm{d}^3 y}{\mathrm{d}x^3}+\sin x\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}+x\frac{\mathrm{d}y}{\mathrm{d}x}+2y=\cos x$$

为一阶方程组，并写为矩阵的形式.

解 设 $\dfrac{\mathrm{d}y}{\mathrm{d}x}=y_1$， $\dfrac{\mathrm{d}y_1}{\mathrm{d}x}=y_2$，

于是

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}=\frac{\mathrm{d}y_1}{\mathrm{d}x}=y_2,\quad \frac{\mathrm{d}^3 y}{\mathrm{d}x^3}=\frac{\mathrm{d}y_2}{\mathrm{d}x},$$

那么，这个三阶微分方程化为一阶方程组

$$\frac{\mathrm{d}y_2}{\mathrm{d}x}=(-\sin x)y_2-xy_1-2y+\cos x$$

$$\frac{\mathrm{d}y_1}{\mathrm{d}x}=y_2$$

$$\frac{\mathrm{d}y}{\mathrm{d}x}=y_1$$

即

$$\frac{\mathrm{d}}{\mathrm{d}x}\begin{pmatrix}y_2\\y_1\\y\end{pmatrix}=\begin{pmatrix}-\sin x & -x & -2\\1 & 0 & 0\\0 & 1 & 0\end{pmatrix}\begin{pmatrix}y_2\\y_1\\y\end{pmatrix}+\begin{pmatrix}\cos x\\0\\0\end{pmatrix}$$

或

$$\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}x}=\boldsymbol{A}\boldsymbol{y}+\boldsymbol{f}$$

其中

$$\boldsymbol{y}=\begin{pmatrix}y_2\\y_1\\y\end{pmatrix},\quad A=\begin{pmatrix}-\sin x & -x & -2\\1 & 0 & 0\\0 & 1 & 0\end{pmatrix},\quad \boldsymbol{f}=\begin{pmatrix}\cos x\\0\\0\end{pmatrix}$$

类似地可以定义**矩阵 $\boldsymbol{A}=(a_{ij}(t))_{m\times n}$ 的积分**，当 $a_{ij}(t)$ 是 t 的可积函数时，定义

$$\int \boldsymbol{A}\mathrm{d}t=\begin{pmatrix}\int a_{11}(t)\mathrm{d}t & \int a_{12}(t)\mathrm{d}t & \cdots & \int a_{1n}(t)\mathrm{d}t\\ \vdots & \vdots & & \vdots\\ \int a_{m1}(t)\mathrm{d}t & \int a_{m2}(t)\mathrm{d}t & \cdots & \int a_{mn}(t)\mathrm{d}t\end{pmatrix}$$

例如

$$\boldsymbol{A}=\begin{pmatrix}\sin t & 1\\ 2t & 0\end{pmatrix}$$

则

$$\int_0^{\pi}\boldsymbol{A}\mathrm{d}t=\begin{pmatrix}\int_0^{\pi}\sin t\mathrm{d}t & \int_0^{\pi}\mathrm{d}t\\ \int_0^{\pi}2t\mathrm{d}t & \int_0^{\pi}\mathrm{d}t\end{pmatrix}=\begin{pmatrix}2 & \pi\\ \pi^2 & 0\end{pmatrix}$$

习　题　一

1. 设函数 $z=f(u,v,w)$，$u=\varphi(x,y)$，$v=\psi(x,y)$，$w=w(x,y)$，试用矩阵形式表示连锁规则

$$\begin{cases}\dfrac{\partial z}{\partial x}=\dfrac{\partial f}{\partial u}\dfrac{\partial u}{\partial x}+\dfrac{\partial f}{\partial v}\dfrac{\partial v}{\partial x}+\dfrac{\partial f}{\partial w}\dfrac{\partial w}{\partial x}\\ \dfrac{\partial z}{\partial y}=\dfrac{\partial f}{\partial u}\dfrac{\partial u}{\partial y}+\dfrac{\partial f}{\partial v}\dfrac{\partial v}{\partial y}+\dfrac{\partial f}{\partial w}\dfrac{\partial w}{\partial y}\end{cases}$$

2. 某公司半成品分厂利用四种原材料生产三种半成品，它的原料消耗矩阵为 $\boldsymbol{A}=(a_{ij})_{4\times 3}$，其中 a_{ij} 表示生产一个单位第 j 种半成品需第 i 种原料的数量. 该公司装配厂利用分厂的三种半成品装配两种成品，它的半成品消耗矩阵为 $\boldsymbol{B}=(b_{ij})_{3\times 2}$，$b_{ij}$ 表示生产一个单位第 j 种成品需第 i 种半成品的数量. 若

$$\boldsymbol{A}=\begin{pmatrix}5 & 7 & 8\\ 1 & 0 & 2\\ 0 & 3 & 3\\ 4 & 1 & 0\end{pmatrix},\qquad \boldsymbol{B}=\begin{pmatrix}4 & 2\\ 5 & 3\\ 5 & 4\end{pmatrix}$$

试求成品对原料的消耗矩阵.

3. 已知两个线性变换

$$\begin{cases} y_1 = 2x_1 + x_2 \\ y_2 = 3x_1 - x_2 + 2x_3, \\ y_3 = 4x_1 + 2x_2 - x_3 \end{cases} \qquad \begin{cases} z_1 = y_1 + 2y_2 - y_3 \\ z_2 = 2y_1 - 2y_2 + 2y_3 \\ z_3 = y_1 - 3y_2 - 4y_3 \end{cases}$$

试用矩阵方法求出变量 z 与变量 x 间的线性关系式.

4. 设

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & -2 & 4 \\ 0 & 5 & 1 \end{pmatrix}$$

求(1)$3\boldsymbol{BA} - 2\boldsymbol{A}$；　(2)$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}$.

5. 若 $\boldsymbol{A} = \mathrm{diag}(a_{11}, a_{22}, \cdots, a_{nn})$, $\boldsymbol{B} = \mathrm{diag}(b_{11}, b_{22}, \cdots, b_{nn})$,求

(1)$\boldsymbol{A}^2$

(2)$\boldsymbol{AB}$

(3)当 $a_{ii} \neq 0 (i = 1, 2, \cdots, n)$时,验证

$$\boldsymbol{A}^{-1} = \mathrm{diag}(1/a_{11}, \cdots, 1/a_{nn})$$

(4)若 $\boldsymbol{C} = (c_{ij})_{n \times n}$,求 $\boldsymbol{AC}$ 与 $\boldsymbol{CA}$

6. (1)　设

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

求　$\boldsymbol{AB}, \boldsymbol{BA}, \boldsymbol{AC}, \boldsymbol{CA}$.

(2)　求

$$(x_1, x_2, x_3) \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

7. 已知 $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$,验证

(1)$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$　　(2)$(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$

(3)$(\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$　　(4)$(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$

8. 设

$$\boldsymbol{A}=\begin{pmatrix}2 & -2 & -4\\ -1 & 3 & 4\\ 1 & -2 & -3\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}2 & -3 & -5\\ -1 & 4 & 5\\ 1 & -3 & -4\end{pmatrix}$$

(1)验证 $\boldsymbol{AB}=\boldsymbol{A},\boldsymbol{BA}=\boldsymbol{B}$

(2)利用(1)证明 $\boldsymbol{A}^2=\boldsymbol{A},\boldsymbol{B}^2=\boldsymbol{B}$

一般地,若 $\boldsymbol{A}^2=\boldsymbol{A}$,称方阵 $\boldsymbol{A}$ 为**幂等矩阵**.

9. 求矩阵 $\boldsymbol{X},\boldsymbol{Y}$ 使

$$\begin{cases}\boldsymbol{X}+\boldsymbol{Y}=\boldsymbol{A}^2\\ \boldsymbol{X}-\boldsymbol{Y}=\boldsymbol{BA}\end{cases}$$

10. 举反例说明下列命题不成立:

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

(2)$(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})=\boldsymbol{A}^2-\boldsymbol{B}^2$

(3)若 $\boldsymbol{A}^2=\boldsymbol{A}$,则 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{A}=\boldsymbol{I}$

(4)若 $\boldsymbol{A}^2=\boldsymbol{O}$,则 $\boldsymbol{A}=\boldsymbol{O}$

(5)若 $\boldsymbol{Ax}=\boldsymbol{Ay}$,则 $\boldsymbol{x}=\boldsymbol{y}$

11. 设 $\boldsymbol{A}=\begin{pmatrix}1 & 2\\ -1 & 0\end{pmatrix},\boldsymbol{B}=\begin{pmatrix}2 & 1\\ 0 & 1\end{pmatrix}$,

(1)根据例 1.8 求出 $\boldsymbol{A}^{-1},\boldsymbol{B}^{-1},(\boldsymbol{AB})^{-1}$;

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

12. 设方阵 $\boldsymbol{A}$ 满足关系式 $\boldsymbol{A}^2-\boldsymbol{A}-2\boldsymbol{I}=\boldsymbol{O}$,试证 $\boldsymbol{A}$ 及 $\boldsymbol{A}+2\boldsymbol{I}$ 均可逆.

13$^{\triangle}$. 设 $\boldsymbol{A}$ 是如下 n 阶方阵:对于某个正整数 k 有 $\boldsymbol{A}^k=\boldsymbol{O}$,且 $\boldsymbol{A}^{k-1}\neq\boldsymbol{O}$,称 $\boldsymbol{A}$ 为**零幂矩阵**,试证

$$(\boldsymbol{I}-\boldsymbol{A})^{-1}=\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^{k-1}.$$

14. 证明下列命题:

(1)若方阵 $\boldsymbol{A}$ 可逆,则其逆矩阵唯一;

(2)若方阵 $\boldsymbol{A}$ 可逆,$\boldsymbol{B}$ 为同阶方阵,且 $\boldsymbol{AB}=\boldsymbol{O}$,证明 $\boldsymbol{B}=\boldsymbol{O}$;

△ 研究生入学试题

(3)若可逆方阵 $\boldsymbol{A}$ 是对称矩阵,则 $\boldsymbol{A}^{-1}$ 亦为对称矩阵;

(4)对于任意方阵 $\boldsymbol{A}$,证明 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 与 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 均为对称矩阵;

(5)如果 $\boldsymbol{A},\boldsymbol{B}$ 皆为对称矩阵,则 $\boldsymbol{AB}$ 也为对称矩阵的充要条件是 $\boldsymbol{AB}=\boldsymbol{BA}$;

(6)对于任意矩阵 $\boldsymbol{A}$,证明 $\boldsymbol{A}^*\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^*$ 均为厄尔米特矩阵;

(7)$\boldsymbol{A}$ 为反厄尔米特矩阵,证明 $\mathrm{i}\boldsymbol{A}$ 及 $-\mathrm{i}\boldsymbol{A}$ 均为厄尔米特矩阵.

15*. 作矩阵 $\boldsymbol{A}=\begin{pmatrix}2&2&3\\2&4&0\\1&0&2\end{pmatrix}$的 $\boldsymbol{LU}$ 分解.

16. 对下面矩阵乘积完成分划,使每个矩阵分为四个子块

$$\left(\begin{array}{cccc}*&*&*&*\\ *&*&*&*\\ *&*&*&*\\ \hdashline *&*&*&*\end{array}\right)\left(\begin{array}{cccc}*&*&*&*\\ \hdashline *&*&*&*\\ *&*&*&*\\ *&*&*&*\end{array}\right)=\left(\begin{array}{c:ccc}*&*&*&*\\ *&*&*&*\\ *&*&*&*\\ *&*&*&*\end{array}\right)$$

17. 若矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可乘,试证:

(1)若 $\boldsymbol{A}$ 有零行,则 $\boldsymbol{AB}$ 也有零行;

(2)若 $\boldsymbol{B}$ 有零列,则 $\boldsymbol{AB}$ 也有零列.

18. 证明两个同阶的上三角矩阵的乘积仍为上三角矩阵.

19. 计算

$$\boldsymbol{A}=\begin{pmatrix}1&2&1&0\\0&1&0&1\\0&0&2&1\\0&0&0&3\end{pmatrix}\begin{pmatrix}1&0&3&1\\0&1&2&-1\\0&0&-7&1\\0&0&-2&1\end{pmatrix}$$

及 $\boldsymbol{A}^2$.

20. 设分块矩阵 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{I}&\boldsymbol{O}\\\boldsymbol{A}_1&\boldsymbol{I}\end{pmatrix}$,求 $\boldsymbol{A}^n$.

21. 设 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_1&\boldsymbol{A}_2\\\boldsymbol{O}&\boldsymbol{A}_3\end{pmatrix}$,若 $\boldsymbol{A}_1,\boldsymbol{A}_3$ 可逆,验证

$$\boldsymbol{A}^{-1}=\begin{pmatrix}\boldsymbol{A}_1^{-1}&-\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\boldsymbol{A}_3^{-1}\\\boldsymbol{O}&\boldsymbol{A}_3^{-1}\end{pmatrix}.$$

22. n 阶方阵 $\boldsymbol{A}$ 的主对角元素之和称为 $\boldsymbol{A}$ 的**迹**，记为 $\mathrm{tr}(\boldsymbol{A})$，即 $\mathrm{tr}(\boldsymbol{A})=a_{11}+a_{22}+\cdots+a_{nn}$. 设 $\boldsymbol{A},\boldsymbol{B}$ 均为 n 阶方阵，试证：

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

(2) $\mathrm{tr}(\boldsymbol{AB})=\mathrm{tr}(\boldsymbol{BA})$.

23△. 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{A}+\boldsymbol{B},\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$ 均为 n 阶可逆方阵，则 $(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}$ 等于(　)

(a) $\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$　　(b) $\boldsymbol{A}+\boldsymbol{B}$

(c) $\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B}$　　(d) $(\boldsymbol{A}+\boldsymbol{B})^{-1}$

24△. 设 $\boldsymbol{A}$、$\boldsymbol{B}$ 均为 3 阶矩阵，$\boldsymbol{I}$ 为 3 阶单位矩阵，已知 $\boldsymbol{AB}=2\boldsymbol{A}+\boldsymbol{B}$，其中矩阵

$$\boldsymbol{B}=\begin{bmatrix}2&0&2\\0&4&0\\2&0&2\end{bmatrix}$$

则 $(\boldsymbol{A}-\boldsymbol{I})^{-1}=$________.

第 2 章　行列式　克莱姆法则　消元法

线性方程组的求解是线性代数的重要内容. 本章给出判定含 n 个未知数 n 个方程的 **n 阶线性方程组**有唯一解的克莱姆法则，它是中学里二、三阶线性方程组行列式解法的推广. 本章首先将二、三阶行列式推广到 n 阶行列式，并讨论它的性质及计算. 行列式在线性代数的其它问题中也有应用. 本章最后给出求解线性方程组的消元法，它不仅有着重要的理论意义，而且是目前计算机上常用的有效算法之一.

2.1　行列式的定义及性质

2.1.1　行列式定义

首先回顾高中代数关于二、三阶行列式的讨论，对于二阶线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{2.1}$$

它的矩阵形式是 $\boldsymbol{Ax}=\boldsymbol{b}$，系数矩阵 $\boldsymbol{A}$ 和它的行列式分别是

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad |\boldsymbol{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

常将矩阵 $\boldsymbol{A}$ 的行列式记作 $\det\boldsymbol{A}$，矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{A}$ 的行列式不同，矩阵仅是一个矩形元素表，而行列式是方阵中各元素的一个特殊函数，例如二阶方阵 $\boldsymbol{A}$ 的行列式

$$\det\boldsymbol{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{2.2}$$

三阶方阵 $\boldsymbol{A}$ 的行列式

$$\det\boldsymbol{A}=\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}$$

$$=a_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix}-a_{12}\begin{vmatrix}a_{21}&a_{23}\\a_{31}&a_{33}\end{vmatrix}+a_{13}\begin{vmatrix}a_{21}&a_{22}\\a_{31}&a_{32}\end{vmatrix}\tag{2.3}$$

这是把三阶行列式归结为二阶行列式计算,式(2.3)右端各项的第一个因子分别是原行列式第一行的各元素,另一因子是去掉第一因子所在行所在列的元素构成的低一阶的行列式,右端各项的符号正负相间.

这种由二阶行列式表示三阶行列式的方法启发我们就用这种方式递推地给出 n 阶行列式的定义.

定义 2.1 n 阶矩阵 $\boldsymbol{A}$ 的行列式 $\det\boldsymbol{A}$ 定义为

$n=1$ 时,$\det\boldsymbol{A}=|a_{11}|=a_{11}$,

n 为大于 1 的正整数时,

$$\det\boldsymbol{A}=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}=a_{11}\begin{vmatrix}a_{22}&a_{23}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n2}&a_{n3}&\cdots&a_{nn}\end{vmatrix}-$$

$$a_{12}\begin{vmatrix}a_{21}&a_{23}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n3}&\cdots&a_{nn}\end{vmatrix}+\cdots+(-1)^{1+n}a_{1n}\begin{vmatrix}a_{21}&a_{22}&\cdots&a_{2,n-1}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{n,n-1}\end{vmatrix}\tag{2.4}$$

n 阶行列式中去掉元素 a_{ij} 所在行所在列的元素后的 $n-1$ 阶行列式叫作 a_{ij} 的**余子式**,记作 M_{ij},即

$$M_{ij}=\begin{vmatrix}a_{11}&\cdots&a_{1,j-1}&a_{1,j+1}&\cdots&a_{1n}\\\vdots&&\vdots&\vdots&&\vdots\\a_{i-1,1}&\cdots&a_{i-1,j-1}&a_{i-1,j+1}&\cdots&a_{i-1,n}\\a_{i+1,1}&\cdots&a_{i+1,j-1}&a_{i+1,j+1}&\cdots&a_{i+1,n}\\\vdots&&\vdots&\vdots&&\vdots\\a_{n1}&\cdots&a_{n,j-1}&a_{n,j+1}&\cdots&a_{nn}\end{vmatrix}$$

并称 $D_{ij}=(-1)^{i+j}M_{ij}$ 为**代数余子式**. 引入这两个记号则可将(2.4)式简记为

$$\det\boldsymbol{A}=a_{11}M_{11}-a_{12}M_{12}+\cdots+(-1)^{1+n}a_{1n}M_{1n}$$
$$=\sum_{k=1}^{n}(-1)^{1+k}a_{1k}M_{1k} \tag{2.5}$$

或

$$\det\boldsymbol{A}=a_{11}D_{11}+a_{12}D_{12}+\cdots+a_{1n}D_{1n}=\sum_{k=1}^{n}a_{1k}D_{1k} \tag{2.6}$$

式(2.4),(2.5)和(2.6)统称为 **n 阶行列式按第一行的展开式.**

关于矩阵的行、列、主对角线、转置等术语均适用于行列式.

例 2.1 据行列式定义

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

$$=3\begin{vmatrix}1&0&0\\1&1&2\\1&2&0\end{vmatrix}+2\left\{2\begin{vmatrix}0&0&0\\0&1&2\\2&2&0\end{vmatrix}-3\begin{vmatrix}1&0&0\\0&1&2\\1&2&0\end{vmatrix}\right\}=12$$

例 2.2

$$\begin{vmatrix}a_{11}&&&\\a_{21}&a_{22}&&\\\vdots&&\ddots&\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}=a_{11}\begin{vmatrix}a_{22}&&\\\vdots&\ddots&\\a_{n2}&\cdots&a_{nn}\end{vmatrix}=\cdots=a_{11}a_{22}\cdots a_{nn}$$

本例表明任何下三角行列式都等于主对角元素的积.

2.1.2 行列式的性质

由行列式定义知道,计算二阶行列式要作 2 次乘法,而三阶行列式需作 9 次乘法,四阶行列式作 40 次乘法,这个数字增大得很快,到 10 阶行列式就要作 600 多万次乘法,这个计算量是相当可

观的，为了减少计算量也为了理论研究的需要，必须对行列式进行化简，为此应当研究行列式的性质.

性质 1　行列式转置后的值不变，即 $\det\boldsymbol{A}=\det(\boldsymbol{A}^{\mathrm{T}})$，或者

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix}$$

这个性质不难由二阶、三阶或四阶行列式验证，一般地采用数学归纳法证之.

有了这条性质可以证明行列式按第一行展开与按第一列展开相等，即 $\det\boldsymbol{A}=\sum\limits_{k=1}^{n}a_{1k}D_{1k}=\sum\limits_{k=1}^{n}a_{k1}D_{k1}$. 为了简单，仅以三阶行列式为例证之.

设
$$\det\boldsymbol{A}=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

根据性质 1，

$$\begin{aligned}\det\boldsymbol{A}=\det\boldsymbol{A}^{\mathrm{T}}&=\begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}\\ &=a_{11}\begin{vmatrix} a_{22} & a_{32} \\ a_{23} & a_{33} \end{vmatrix}-a_{21}\begin{vmatrix} a_{12} & a_{32} \\ a_{13} & a_{33} \end{vmatrix}+a_{31}\begin{vmatrix} a_{12} & a_{22} \\ a_{13} & a_{23} \end{vmatrix}\\ &=a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}-a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}+a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}\end{aligned}$$

最后的表达式恰好是行列式 $\det\boldsymbol{A}$ 按第一列展开的结果.

由于性质 1，以下凡是对行给出的性质同样适用于列.

例 2.3　对于上三角行列式始终按第一列展开，则有

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ & \ddots & \vdots \\ & & a_{nn} \end{vmatrix}$$

$$= \cdots = a_{11} \cdots a_{nn}$$

联系到例 2.2 知，任何三角形行列式等于主对角线元素之积，特别有 $\det \boldsymbol{I}=1$.

性质 2 对调行列式的任意两行（或两列），行列式仅改变符号.

例如$\det \boldsymbol{A}=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}$，调换两行后，

$$\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}=a_{21}a_{12}-a_{22}a_{11}=-\det \boldsymbol{A}.$$

推论 有两行（或两列）对应元素相同的行列式等于零.

性质 3 行列式中某一行（或列）的元素都乘以 k，等于用 k 乘原行列式，即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

例如

$$\begin{vmatrix} a_{11} & a_{12} \\ ka_{21} & ka_{22} \end{vmatrix} = a_{11}ka_{22} - a_{12}ka_{21} = k\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

推论 1 若行列式有一行（或列）元素全为零，则行列式为零.

推论 2 若行列式有两行（或两列）元素成比例，则此行列式为零.

性质 4 若行列式某行（或列）各元素均为两元素之和，则有

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1}+c_{i1} & b_{i2}+c_{i2} & \cdots & b_{in}+c_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ c_{i1} & c_{i2} & \cdots & c_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

以三阶行列式为例说明之,

$$\begin{vmatrix} b_{11}+c_{11} & b_{12}+c_{12} & b_{13}+c_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$=(b_{11}+c_{11})D_{11}+(b_{12}+c_{12})D_{12}+(b_{13}+c_{13})D_{13}$$

$$=(b_{11}D_{11}+b_{12}D_{12}+b_{13}D_{13})+(c_{11}D_{11}+c_{12}D_{12}+c_{13}D_{13})$$

$$= \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} c_{11} & c_{12} & c_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

性质 5 行列式某行(或列)各元素乘以同一个数分别加到另一行(或列)的对应元素上,行列式值不变,即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \xlongequal{r_j+kr_i} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1}+ka_{i1} & a_{j2}+ka_{i2} & \cdots & a_{jn}+ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

为叙述方便,这里用 r_i 表示第 i 行,以后还将用 c_j 表示第 j 列.

这个性质易由性质 3 和性质 4 推得.

根据行列式性质可以简化行列式计算,如对例 2.1 的行列式作如下简化计算

$$\begin{vmatrix} 1 & 0 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 1 & 2 & 1 & 2 & 0 \end{vmatrix} \xlongequal[r_4 \leftrightarrow r_5]{r_1+(-2)r_3} - \begin{vmatrix} -1 & 0 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 & 2 \end{vmatrix}$$

$$= -(-1)\times 3\times 1\times 2\times 2 = 12$$

例 2.4 计算

$$\det\boldsymbol{A} = \begin{vmatrix} a & b & c & d \\ a & a+b & a+b+c & a+b+c+d \\ a & 2a+b & 3a+2b+c & 4a+3b+2c+d \\ a & 3a+b & 6a+3b+c & 10a+6b+3c+d \end{vmatrix}$$

解

$$\det\boldsymbol{A} \xlongequal[r_2-r_1]{\substack{r_4-r_3 \\ r_3-r_2}} \begin{vmatrix} a & b & c & d \\ 0 & a & a+b & a+b+c \\ 0 & a & 2a+b & 3a+2b+c \\ 0 & a & 3a+b & 6a+3b+c \end{vmatrix}$$

$$\xlongequal[r_3-r_2]{r_4-r_3} \begin{vmatrix} a & b & c & d \\ 0 & a & a+b & a+b+c \\ 0 & 0 & a & 2a+b \\ 0 & 0 & a & 3a+b \end{vmatrix}$$

$$\xlongequal{r_4-r_3} \begin{vmatrix} a & b & c & d \\ 0 & a & a+b & a+b+c \\ 0 & 0 & a & 2a+b \\ 0 & 0 & 0 & a \end{vmatrix}$$

$$= a^4$$

由此可见,根据行列式性质将行列式三角化是简化行列式计算的一种重要方法.

2.2 行列式计算

2.2.1 行列式的展开

根据定义,行列式是按第一行展开计算的,其实按任一行(列)展开都可得到相同的值,这就是

定理 2.1 行列式 $\det\boldsymbol{A}$ 可以按任一行(列)展开,即

$$\det\boldsymbol{A} = \sum_{k=1}^{n} a_{ik} D_{ik} \quad (i = 1,2,\cdots,n) \tag{2.7}$$

证

$$\det\boldsymbol{A}=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$=(-1)^{i-1}\begin{vmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

这是将第 i 行与第 $i-1$ 行交换,然后再与第 $i-2$ 行交换,一直交换 $i-1$ 次使第 i 行位于第一行. 再根据定义按第一行展开右端

$$\begin{aligned} \det\boldsymbol{A} &= (-1)^{i-1} \sum_{k=1}^{n} (-1)^{1+k} a_{ik} M_{ik} \\ &= \sum_{k=1}^{n} (-1)^{i+k} a_{ik} M_{ik} = \sum_{k=1}^{n} a_{ik} D_{ik} \end{aligned}$$

此即所证.

若按第 j 列展开就是

$$\det\boldsymbol{A}=\sum_{k=1}^{n}a_{kj}D_{kj}\qquad(j=1,2,\cdots,n)\tag{2.8}$$

例 2.5

$$\begin{vmatrix}a&b&c&d\\2a&b&c&d\\a&2b&c&d\\0&b&c&0\end{vmatrix}\xlongequal{r_3-r_1}\begin{vmatrix}a&b&c&d\\2a&b&c&d\\0&b&0&0\\0&b&c&0\end{vmatrix}$$

$$\xlongequal{\text{按}r_3\text{展开}}(-1)^{3+2}b\begin{vmatrix}a&c&d\\2a&c&d\\0&c&0\end{vmatrix}$$

$$\xlongequal{\text{按}r_3\text{展开}}(-1)^{5}b\cdot(-1)^{3+2}c\begin{vmatrix}a&d\\2a&d\end{vmatrix}=-abcd$$

定理 2.2　行列式任一行(列)与另一行(列)对应元素的代数余子式的乘积之和等于零,即

$$\left.\begin{aligned}\sum_{k=1}^{n}a_{jk}D_{ik}=0\\\sum_{k=1}^{n}a_{kj}D_{ki}=0\end{aligned}\right.\qquad(i,j=1,2,\cdots,n;i\neq j)\tag{2.9}$$

证　因 $i\neq j$ 时

$$\det\boldsymbol{A}+\sum_{k=1}^{n}a_{jk}D_{ik}=\sum_{k=1}^{n}a_{ik}D_{ik}+\sum_{k=1}^{n}a_{jk}D_{ik}=\sum_{k=1}^{n}(a_{ik}+a_{jk})D_{ik}$$

$$\xlongequal{(i\neq j)}\begin{vmatrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{i1}+a_{j1}&\cdots&a_{in}+a_{jn}\\\vdots&&\vdots\\a_{j1}&\cdots&a_{jn}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nn}\end{vmatrix}\xlongequal{r_i-r_j}\begin{vmatrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{i1}&\cdots&a_{in}\\\vdots&&\vdots\\a_{j1}&\cdots&a_{jn}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nn}\end{vmatrix}=\det\boldsymbol{A}$$

于是
$$\sum_{k=1}^{n} a_{jk} D_{ik} = 0$$

综合定理 2.1 和定理 2.2,有

$$\sum_{k=1}^{n} a_{jk} D_{ik} = \begin{cases} \det\boldsymbol{A}, & i=j \\ 0, & i \neq j \end{cases} \tag{2.10}$$

最后给出一个应用广泛的**范德蒙行列式**[①].

例 2.6 证明范德蒙行列式

$$V_n(x_1,\cdots,x_n) = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}$$
$$= \prod_{n \geqslant i > j \geqslant 1} (x_i - x_j) \tag{2.11}$$

其中记号 $\prod$ 表示连乘积,如 $\prod_{i=1}^{n} x_i = x_1 \cdots x_n$. 因此式(2.11)的右端是

$$\begin{aligned} \prod_{n \geqslant i > j \geqslant 1} (x_i - x_j) = {} & (x_n - x_{n-1})(x_n - x_{n-2})\cdots(x_n - x_1) \\ & \cdot (x_{n-1} - x_{n-2})(x_{n-1} - x_{n-3})\cdots(x_{n-1} - x_1) \\ & \cdots \\ & \cdot (x_3 - x_2) \cdot (x_3 - x_1) \\ & \cdot (x_2 - x_1) \end{aligned}$$

证
$$V_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} \xlongequal[\substack{r_{n-1} - x_1 r_{n-2} \\ \cdots \\ r_2 - x_1 r_1}]{r_n - x_1 r_{n-1}}$$

---

① 法国数学家范德蒙(Vandermonde)(1735～1796)第一个对行列式理论作出系统研究和阐述.

$$= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \\ 0 & x_2^2 - x_2x_1 & x_3^2 - x_3x_1 & \cdots & x_n^2 - x_nx_1 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_2^{n-1} - x_2^{n-2}x_1 & x_3^{n-1} - x_3^{n-2}x_1 & \cdots & x_n^{n-1} - x_n^{n-2}x_1 \end{vmatrix}$$

$$= \begin{vmatrix} x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \\ x_2(x_2 - x_1) & x_3(x_3 - x_1) & \cdots & x_n(x_n - x_1) \\ \vdots & \vdots & & \vdots \\ x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1)\cdots(x_n - x_1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2} \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1)\cdots(x_n - x_1)V_{n-1}(x_2, \cdots, x_n)$$

类似地可得

$$V_{n-1}(x_2, \cdots, x_n) = (x_3 - x_2)\cdots(x_n - x_2)V_{n-2}(x_3, \cdots, x_n)$$

直到 $V_2(x_{n-1}, x_n) = x_n - x_{n-1}$,

将 $V_2, V_3, \cdots$ 逐次回代,得

$$V_n(x_1, \cdots, x_n) = \prod_{n \geqslant i > j \geqslant 1} (x_i - x_j)$$

2.2.2 行列式的乘法法则

先看下例:

设 $\quad \boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$

那么 $\qquad \det\boldsymbol{A} = -2, \qquad \det\boldsymbol{B} = 10$

而 $\qquad \det(\boldsymbol{AB}) = \det\begin{pmatrix} 5 & 10 \\ 13 & 22 \end{pmatrix} = -20$

恰有等式 $\det\boldsymbol{A}\det\boldsymbol{B} = \det(\boldsymbol{AB})$,这不是偶然的.一般地,有如下行列式相乘定理 2.3,为证此定理,首先证明一个引理

引理 2.1 若 $\boldsymbol{A}=(a_{ij})_{m\times m}$，$\boldsymbol{B}=(b_{ij})_{n\times n}$，$\boldsymbol{C}=(c_{ij})_{n\times m}$，则

$$\det\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{C} & \boldsymbol{B}\end{pmatrix}=\det\boldsymbol{A}\det\boldsymbol{B} \tag{2.12}$$

证 应用行列式性质 2 和性质 5，将 $\det\boldsymbol{A}$ 和 $\det\boldsymbol{B}$ 三角化

$$\det\boldsymbol{A}=(-1)^k\begin{vmatrix}a'_{11} & & & \\ a'_{21} & a'_{22} & & \\ \vdots & \vdots & \ddots & \\ a'_{m1} & a'_{m2} & \cdots & a'_{mm}\end{vmatrix}\quad (k\text{ 为列交换次数})$$

$$=(-1)^k a'_{11}a'_{22}\cdots a'_{mm},$$

$$\det\boldsymbol{B}=(-1)^t\begin{vmatrix}b'_{11} & & & \\ b'_{21} & b'_{22} & & \\ \vdots & \vdots & \ddots & \\ b'_{n1} & b'_{n2} & \cdots & b'_{nn}\end{vmatrix}\quad (t\text{ 为列交换次数})$$

$$=(-1)^t b'_{11}b'_{22}\cdots b'_{nn},$$

如果将以上对 $\det\boldsymbol{A}$ 和 $\det\boldsymbol{B}$ 的三角化过程分别施用于行列式 $\det\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{C} & \boldsymbol{B}\end{pmatrix}$ 的前 m 列和后 n 列，有

$$\det\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{C} & \boldsymbol{B}\end{pmatrix}=(-1)^{k+t}\begin{vmatrix}a'_{11} & & & & & & & \\ a'_{21} & a'_{22} & & & & & & \\ \vdots & \vdots & \ddots & & & & & \\ a'_{m1} & a'_{m2} & \cdots & a'_{mm} & & & & \\ & & & & b'_{11} & & & \\ & \boldsymbol{C}' & & & b'_{21} & b'_{22} & & \\ & & & & \vdots & \vdots & \ddots & \\ & & & & b'_{n1} & b'_{n2} & \cdots & b'_{nn}\end{vmatrix}$$

$$=(-1)^{k+t}a'_{11}a'_{22}\cdots a'_{mm}\cdot b'_{11}b'_{22}\cdots b'_{nn}$$

$$=\det\boldsymbol{A}\det\boldsymbol{B}.$$

这就证明了引理 2.1.

推论 1 对于方阵 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_k$ 有

$$\det(\mathrm{diag}(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_k))=\det\boldsymbol{A}_1\det\boldsymbol{A}_2\cdots\det\boldsymbol{A}_k \tag{2.13}$$

定理 2.3　(行列式相乘定理)若 $\boldsymbol{A}=(a_{ij})_{n\times n}$,$\boldsymbol{B}=(b_{ij})_{n\times n}$,那么

$$\det(\boldsymbol{AB}) = \det\boldsymbol{A}\det\boldsymbol{B} \tag{2.14}$$

证　为简单计,仅就 $n=2$ 给出证明.

首先由给定的矩阵 $\boldsymbol{A}$,$\boldsymbol{B}$ 构造如下行列式

$$\det\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ -\boldsymbol{I}_2 & \boldsymbol{B}\end{pmatrix}=\begin{vmatrix}a_{11} & a_{12} & 0 & 0\\ a_{21} & a_{22} & 0 & 0\\ -1 & 0 & b_{11} & b_{12}\\ 0 & -1 & b_{21} & b_{22}\end{vmatrix}$$

对列施行计算 $c_3+b_{11}c_1$,$c_3+b_{21}c_2$,$c_4+b_{12}c_1$ 和 $c_4+b_{22}c_2$,有

$$\det\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ -\boldsymbol{I}_2 & \boldsymbol{B}\end{pmatrix}=\begin{vmatrix}a_{11} & a_{12} & a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22}\\ a_{21} & a_{22} & a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}\\ -1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\end{vmatrix}$$

$$\xlongequal[c_2\leftrightarrow c_4]{c_1\leftrightarrow c_3}\begin{vmatrix}a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} & a_{11} & a_{12}\\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} & a_{21} & a_{22}\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{vmatrix}$$

$$=\det\begin{pmatrix}\boldsymbol{AB} & \boldsymbol{A}\\ \boldsymbol{O} & -\boldsymbol{I}_2\end{pmatrix}$$

根据引理,上式左端

$$\det\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ -\boldsymbol{I}_2 & \boldsymbol{B}\end{pmatrix}= \det\boldsymbol{A}\det\boldsymbol{B}$$

上式右端

$$\det\begin{pmatrix}\boldsymbol{AB} & \boldsymbol{A}\\ \boldsymbol{O} & -\boldsymbol{I}_2\end{pmatrix}= \det(\boldsymbol{AB})\det(-\boldsymbol{I}_2) = \det(\boldsymbol{AB})$$

由此证得

$$\det(\boldsymbol{AB})=\det\boldsymbol{A}\det\boldsymbol{B}.$$

这里仅对 $n=2$ 给出证明,但对一般情形证明的方法一样.

推论 2　若 $\boldsymbol{A}$ 为方阵，则

$$\det\boldsymbol{A}^n = (\det\boldsymbol{A})^n \tag{2.15}$$

例 2.7　根据引理计算

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

$$= (-9) \times 5 = -45$$

例 2.8　由推论 2 知

$$\det\begin{pmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}^n = \left\{\det\begin{pmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}\right\}^n$$

$$= (\lambda^3)^n = \lambda^{3n}$$

关于行列式计算的方法，大致可归纳为：

(1)利用定义直接计算，即按某行(列)展开；

(2)利用行列式性质将行列式三角化，或使某行(列)出现更多个零元素，然后按此行(列)展开；

(3)利用行列式乘积定理及推论.

2.3　克莱姆法则

上节的行列式相乘定理无论在理论上还是计算上都是相当重要的，以致人们把它作为行列式的一条重要性质. 下面应用这个定理给出逆矩阵存在的充要条件，并由此证明著名的克莱姆(Cramer)法则.

2.3.1　逆矩阵存在的充要条件

现在讨论方阵 $\boldsymbol{A}$ 的逆矩阵何时存在，存在时求出 $\boldsymbol{A}^{-1}$.

定理 2.4　若 $\boldsymbol{A}$ 可逆，则 $\det\boldsymbol{A}\neq 0$.

证 A 可逆,则 $AA^{-1}=I$

根据行列式相乘定理,

$$\det(AA^{-1}) = \det A \det A^{-1} = 1$$

所以 $$\det A \neq 0.$$

称 $\det A\neq 0$ 的矩阵 A 为**非奇异矩阵**,$\det A=0$ 的矩阵 A 为**奇异矩阵**,本定理说明 A 非奇异是 A 可逆的必要条件,其实这个条件还是充分的,以下定理说明了这一点.

定理 2.5 (逆矩阵存在定理) 若 $\det A\neq 0$,**则** A **可逆,并且**

$$A^{-1} = \frac{\mathrm{adj} A}{\det A} \tag{2.16}$$

其中 $\mathrm{adj}A$ 称为 A 的**伴随矩阵,定义为**

$$\mathrm{adj} A = \begin{pmatrix} D_{11} & D_{21} & \cdots & D_{n1} \\ \vdots & \vdots & & \vdots \\ D_{1n} & D_{2n} & \cdots & D_{nn} \end{pmatrix} \tag{2.17}$$

其中 D_{ij} 为 a_{ij} **的代数余子式.**

证 因为

$$(\mathrm{adj} A)A = \begin{pmatrix} D_{11} & D_{21} & \cdots & D_{n1} \\ D_{12} & D_{22} & \cdots & D_{n2} \\ \vdots & \vdots & & \vdots \\ D_{1n} & D_{2n} & \cdots & D_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^{n} a_{k1} D_{k1} & & & \\ & \sum_{k=1}^{n} a_{k2} D_{k2} & & \\ & & \ddots & \\ & & & \sum_{k=1}^{n} a_{kn} D_{kn} \end{pmatrix}$$

$$
= \begin{pmatrix} \det \boldsymbol{A} & & & \\ & \det \boldsymbol{A} & & \\ & & \ddots & \\ & & & \det \boldsymbol{A} \end{pmatrix} = \det \boldsymbol{A} \cdot \boldsymbol{I}
$$

所以

$$
\frac{\operatorname{adj} \boldsymbol{A}}{\det \boldsymbol{A}} \boldsymbol{A} = \boldsymbol{I}
$$

同理可证

$$
\boldsymbol{A} \frac{\operatorname{adj} \boldsymbol{A}}{\det \boldsymbol{A}} = \boldsymbol{I}
$$

根据逆矩阵定义，$\boldsymbol{A}^{-1}$存在且为

$$
\boldsymbol{A}^{-1} = \frac{\operatorname{adj} \boldsymbol{A}}{\det \boldsymbol{A}}.
$$

若 $\boldsymbol{A}$ 为一阶非奇异矩阵，规定 $\operatorname{adj}\boldsymbol{A}=1$.

由以上两个定理可知，$\det\boldsymbol{A}\neq 0$（即 $\boldsymbol{A}$ 为非奇异）是 $\boldsymbol{A}$ 可逆的充要条件. 定理 2.5 同时给出了求逆矩阵的一种方法——伴随矩阵法，即按 $\boldsymbol{A}^{-1}=\dfrac{\operatorname{adj}\boldsymbol{A}}{\det\boldsymbol{A}}$求逆，第 1 章例 1.8 的 $\boldsymbol{A}^{-1}$就是根据此式写成的.

例 2.9 对于同阶方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，若 $\boldsymbol{A}\neq\boldsymbol{O}$，$\boldsymbol{B}\neq\boldsymbol{O}$，但 $\boldsymbol{AB}=\boldsymbol{O}$，试证 $\det\boldsymbol{A}=0$ 且 $\det\boldsymbol{B}=0$.

证 由 $\boldsymbol{AB}=\boldsymbol{O}$ 知 $\det(\boldsymbol{AB})=0$，由行列式相乘定理

$$
\det\boldsymbol{A}\det\boldsymbol{B}=\det(\boldsymbol{AB})=0
$$

于是

$$
\det\boldsymbol{A}=0 \quad 或 \quad \det\boldsymbol{B}=0,
$$

但 $\det\boldsymbol{A}$ 与 $\det\boldsymbol{B}$ 必须都等于零，若 $\det\boldsymbol{B}\neq 0$，那么 $\boldsymbol{B}^{-1}$存在，

$$
\boldsymbol{A} = \boldsymbol{ABB}^{-1} = (\boldsymbol{AB})\boldsymbol{B}^{-1} = \boldsymbol{O}
$$

这与题设 $\boldsymbol{A}\neq\boldsymbol{O}$ 矛盾，同样由 $\det\boldsymbol{A}\neq 0$ 将导致 $\boldsymbol{B}=\boldsymbol{O}$，亦与题设矛盾，由此证得本题结论.

例 2.10 设 $\boldsymbol{A}$ 为 n 阶方阵，若存在 n 阶方阵 $\boldsymbol{B}$ 使 $\boldsymbol{BA}=\boldsymbol{I}$，试证 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆矩阵.

证 对 $\boldsymbol{BA}=\boldsymbol{I}$ 两端取行列式,有

$$\det\boldsymbol{B}\det\boldsymbol{A}=1$$

由此 $\det\boldsymbol{A}\neq 0$,即 $\boldsymbol{A}^{-1}$存在,那么

$$\boldsymbol{AB}=\boldsymbol{AB}(\boldsymbol{AA}^{-1})=\boldsymbol{A}(\boldsymbol{BA})\boldsymbol{A}^{-1}=AA^{-1}=\boldsymbol{I}$$

由 $\boldsymbol{BA}=\boldsymbol{I}$ 及 $\boldsymbol{AB}=\boldsymbol{I}$ 知 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的逆矩阵.

由此可见,若有 $\boldsymbol{BA}=\boldsymbol{I}$ 定有 $\boldsymbol{AB}=\boldsymbol{I}$,反之亦然,那么今后判定逆矩阵时,只需验证 $\boldsymbol{BA}=\boldsymbol{I}$ 及 $\boldsymbol{AB}=\boldsymbol{I}$ 之一成立即可.

例 2.11 求 $\boldsymbol{A}$ 的逆矩阵,

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

解 由 $\det\boldsymbol{A}=-1\neq 0$ 知 $\boldsymbol{A}^{-1}$存在

求出
$$\begin{aligned}&D_{11}=-4, \quad D_{21}=-3, \quad D_{31}=-2\\ &D_{12}=-3, \quad D_{22}=-2, \quad D_{32}=-1\\ &D_{13}=-2, \quad D_{23}=-1, \quad D_{33}=-1\end{aligned}$$

于是

$$\boldsymbol{A}^{-1}=\frac{\mathrm{adj}\boldsymbol{A}}{\det\boldsymbol{A}}=\frac{1}{-1}\begin{pmatrix}-4 & -3 & -2\\ -3 & -2 & -1\\ -2 & -1 & -1\end{pmatrix}$$
$$=\begin{pmatrix}4 & 3 & 2\\ 3 & 2 & 1\\ 2 & 1 & 1\end{pmatrix}.$$

利用伴随矩阵法求高阶矩阵的逆矩阵是不可取的,后面将介绍更实用的矩阵求逆方法.

2.3.2 克莱姆法则

在中学里已经知道如何用行列式表示二阶或三阶线性方程组的解,现在将其推广到 n 阶线性方程组,这就是著名的克莱姆法则,它是克莱姆于 1750 年提出的.

定理 2.6(克莱姆法则) **若 $\boldsymbol{n}$ 阶线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的系数行**

列式 $D=\det\boldsymbol{A}\neq 0$，则方程组有唯一解

$$x_1=\frac{D_1}{D},\quad x_2=\frac{D_2}{D},\quad \cdots,\quad x_n=\frac{D_n}{D} \tag{2.18}$$

其中

$$D_j=\begin{vmatrix} a_{11} & \cdots & a_{1j-1} & b_1 & a_{1j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j-1} & b_2 & a_{2j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

$$(j=1,2,\cdots,n)$$

这个定理的结论包括两点，一是在 $D\neq 0$ 时方程组一定有解（解的存在性），二是这个解只能是 $x_i=\dfrac{D_i}{D}$（解的唯一性）. 以下按这两点证明.

证 首先证明唯一性.

设 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解 $\boldsymbol{x}$，据假设 $D=\det\boldsymbol{A}\neq 0$ 知 $\boldsymbol{A}^{-1}$ 存在，那么 $\boldsymbol{A}^{-1}\boldsymbol{Ax}=\boldsymbol{A}^{-1}\boldsymbol{b}$，即

$$\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$$

由(2.16)

$$\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}=\frac{\operatorname{adj}\boldsymbol{A}}{\det\boldsymbol{A}}\boldsymbol{b}=\frac{1}{D}\begin{pmatrix} D_{11} & D_{21} & \cdots & D_{n1} \\ D_{12} & D_{22} & \cdots & D_{n2} \\ \vdots & \vdots & & \vdots \\ D_{1n} & D_{2n} & \cdots & D_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$=\frac{1}{D}\begin{pmatrix} \sum\limits_{k=1}^{n} b_k D_{k1} \\ \vdots \\ \sum\limits_{k=1}^{n} b_k D_{kn} \end{pmatrix}=\begin{pmatrix} \dfrac{D_1}{D} \\ \vdots \\ \dfrac{D_n}{D} \end{pmatrix} \tag{2.19}$$

即

$$x_1=\frac{D_1}{D},\quad x_2=\frac{D_2}{D},\quad \cdots,\quad x_n=\frac{D_n}{D}$$

其次证明解的存在性，这只要将(2.18)确定的

$$\boldsymbol{x}=\left(\frac{D_1}{D},\frac{D_2}{D},\cdots,\frac{D_n}{D}\right)^{\mathrm{T}}$$

代入 $\boldsymbol{Ax}=\boldsymbol{b}$ 的左边，立刻验证满足这个方程组，事实上由式(2.19)知

$$\boldsymbol{A}\begin{pmatrix}\frac{D_1}{D}\\ \frac{D_2}{D}\\ \vdots\\ \frac{D_n}{D}\end{pmatrix}=\boldsymbol{AA}^{-1}\boldsymbol{b}=\boldsymbol{b}$$

至此，定理全部证毕.

例 2.12 解线性方程组

$$\begin{cases}2x_1-x_2+x_3-4x_4=3\\ x_1-4x_2+2x_3=-4\\ 3x_2-4x_4=5\\ 3x_1-3x_2+x_3+x_4=6\end{cases}$$

解

$$D=\begin{vmatrix}2&-1&1&-4\\1&-4&2&0\\0&3&0&-4\\3&-3&1&1\end{vmatrix}=-37$$

$$D_1=\begin{vmatrix}3&-1&1&-4\\-4&-4&2&0\\5&3&0&-4\\6&-3&1&1\end{vmatrix}=-148$$

$$D_2=-111,\quad D_3=-74,\quad D_4=-37$$

所以

$$x_1=\frac{-148}{-37}=4,\quad x_2=\frac{-111}{-37}=3$$

$$x_3=\frac{-74}{-37}=2,\quad x_4=\frac{-37}{-37}=1$$

2.4 解线性方程组的消元法

根据克莱姆法则，在 $\det\boldsymbol{A}\neq 0$ 的条件下可以求出任何 n 阶线性方程组的解，由此可以认为这种 n 阶线性方程组的求解问题已经完满的解决了. 但是对于高阶线性方程组来说，克莱姆法则只具有理论上的意义，实际求解时是行不通的. 比如求解一个 18 阶的线性方程组，按克莱姆法则需求 19 个 18 阶行列式的值，按定义求一个 18 阶的行列式大约作 1.1×10^{16} 次乘法，这个工作量多大呢? 用每秒 100 万次的计算机昼夜不停地计算，约需 350 年. 而工程技术中上百阶的线性方程组并不少见，克莱姆法则对它们已无能为力. 因此必须寻求更实用更有效的方法. 消元法就是其中的一种，它是一种古老的方法，但实践证明它仍是目前计算机上经常用到的有效方法之一.

2.4.1 高斯消元法

1. 引例

高斯(Gauss)消元法是求解线性方程组最简单的方法，我们由一个简单例子说明它.

例 2.13 解线性方程组

$$\begin{cases} 4x_1+2x_2+5x_3=4 & (1)\\ x_1+2x_2=7 & (2)\\ 2x_1-x_2+3x_3=1 & (3)\end{cases}$$

解 首先保留(1)，利用(1)消去(2)和(3)中含 x_1 的项

$$(2)-\frac{1}{4}(1),\quad \frac{3}{2}x_2-\frac{5}{4}x_3=6 \qquad (4)$$

$$(3)-\frac{2}{4}(1),\quad -2x_2+\frac{1}{2}x_3=-1 \tag{5}$$

在(4),(5)两方程中,保留(4),用(4)消去(5)中含 x_2 的项,

$$(5)-(-2)/(3/2)\times(4),\quad \frac{-7}{6}x_3=7 \tag{6}$$

得到由三个保留方程构成的**三角形方程组**,即其系数矩阵为三角形矩阵的方程组

$$\begin{cases} 4x_1+2x_2+5x_3=4 & (1)\\ \frac{3}{2}x_2-\frac{5}{4}x_3=6 & (4)\\ -\frac{7}{6}x_3=7 & (6)\end{cases}$$

由最后一个方程求得

$$x_3=-6$$

代入(4),得

$$x_2=-1$$

将 $x_3=-6,x_2=-1$ 代入(1),得

$$x_1=9$$

于是得到方程组的解

$$\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}=(9,-1,-6)^{\mathrm{T}}.$$

它不仅满足原方程组,也满足(1),(4)和(6)构成的三角形方程组,原方程组和化简后的三角形方程组为同解方程组.

由这个例子可以看出用消元法解线性方程组的全过程:首先保留一个方程(叫做**保留方程**),利用它消去其它方程的一个未知数(如 x_1),这叫**消元**,消元后的其它方程构成新方程组((4)和(5)),对新方程组继续以上作法又得到少一个未知数少一个方程的方程组,直至最后得到仅剩一个未知数的方程,将所有的保留方程合在一起构成一个三角形方程组,这个过程称为**消元过程**.另一过程是**回代过程**,就是由三角形方程组的最后一个方程求出一个未知数的值,代入上一方程求出又一个未知数的值,照这样依次向上代入直至求出全部未知数的值.消元过程和回代过程构成高斯

消元法的全过程.

2. 矩阵的初等变换

分析一下例 2.13 的消元过程不难看出，消元过程实际上是由原方程组出发，反复施行如下变换的过程：

(1)用一个非零的数乘一个方程；

(2)用一个数乘一个方程加到另一个方程；

(3)交换某两个方程的位置(有时用到).

我们称如上变换为**初等变换**，容易证明，初等变换把方程组变为与它同解的方程组.

对方程组施行初等变换涉及的只是方程组的系数和右端的常数项，也就是对系数矩阵 $\boldsymbol{A}$ 和常数列向量 $\boldsymbol{b}$ 拼起来的**增广矩阵** $(\boldsymbol{A},\boldsymbol{b})$施行初等变换，对矩阵的这种变换称为**矩阵的初等行变换**，并且规定：

(1)互换矩阵的两行(例如第 i 行与第 j 行，记作 $r_i \leftrightarrow r_j$)称为矩阵的第一种初等行变换；

(2)用数 $k\neq 0$ 乘矩阵的某行的所有元素(例如第 i 行乘 k，记作 kr_i)称为矩阵的第二种初等行变换；

(3)把矩阵某行的所有元素的 k 倍加到另一行的对应元素上去(例如第 j 行的 k 倍加到第 i 行上，记作 r_i+kr_j)称为矩阵的第三种初等行变换.

若把以上规定中的“行”换为“列”则有相应的初等列变换.

对矩阵施行初等变换是一种常用的矩阵方法，为此给出定义.

定义 2.2 如果矩阵 $\boldsymbol{A}$ 经过有限次初等变换变为矩阵 $\boldsymbol{B}$，则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **等价**，记作 $\boldsymbol{A}\sim\boldsymbol{B}$.

如果应用矩阵的初等变换表示例 2.13 的消元过程，就是

$$(\boldsymbol{A},\boldsymbol{b})=\left(\begin{array}{ccc:c}4 & 2 & 5 & 4\\ 1 & 2 & 0 & 7\\ 2 & -1 & 3 & 1\end{array}\right)$$

$$\xrightarrow[r_3-\frac{2}{4}r_1]{r_2-\frac{1}{4}r_1}\left(\begin{array}{ccc:c}4 & 2 & 5 & 4 \\ 0 & \frac{3}{2} & \frac{-5}{4} & 6 \\ 0 & -2 & \frac{1}{2} & -1\end{array}\right)$$

$$\xrightarrow{r_3+\frac{4}{3}r_2}\left(\begin{array}{ccc:c}4 & 2 & 5 & 4 \\ 0 & 3/2 & -5/4 & 6 \\ 0 & 0 & -7/6 & 7\end{array}\right).$$

3.高斯消元法

现在叙述一般 n 阶线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 & (1) \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 & (2) \\ \quad\vdots & \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n & (n)\end{cases} \tag{2.20}$$

的高斯消元法,设系数行列式 $\det\boldsymbol{A}\neq0$,则方程组有唯一解,消元过程分若干步完成.

第一步,设方程(1)中 x_1 的系数 $a_{11}\neq0$,这个假设不失一般性,否则总有方程(l)中 x_1 的系数 $a_{ll}\neq0$,这时可将方程(l)与(1)对调,使对调后的第一个方程 x_1 的系数不为零.

作$(i)-\dfrac{a_{i1}}{a_{11}}(1)(i=2,3,\cdots,n)$,得到同解方程组

$$\begin{cases} a_{11}^{(0)}x_1+a_{12}^{(0)}x_2+\cdots+a_{1n}^{(0)}x_n=b_1^{(0)} \\ \qquad\quad a_{22}^{(1)}x_2+\cdots+a_{2n}^{(1)}x_n=b_2^{(1)} \\ \qquad\qquad\qquad\vdots \\ \qquad\quad a_{n2}^{(1)}x_2+\cdots+a_{nn}^{(1)}x_n=b_n^{(1)}\end{cases} \tag{2.21}$$

这里引进上标(k)表明第 k 步得到的新方程的系数 $a_{ij}^{(k)}$ 和 $b_i^{(k)}$,方程组(2.21)中的 $a_{1j}^{(0)}=a_{1j}$,$b_1^{(0)}=b_1$,$(j=1,\cdots,n)$.

第二步,设 $a_{22}^{(1)}\neq0$,保留第二个方程,消去它以下方程中的含 x_2 的项,得

$$\begin{cases} a_{11}^{(0)}x_1 + a_{12}^{(0)}x_2 + a_{13}^{(0)}x_3 + \cdots + a_{1n}^{(0)}x_n = b_1^{(0)} \\ \qquad a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \cdots + a_{2n}^{(1)}x_n = b_2^{(1)} \\ \qquad\qquad a_{33}^{(2)}x_3 + \cdots + a_{3n}^{(2)}x_n = b_3^{(2)} \\ \qquad\qquad\qquad \vdots \\ \qquad\qquad a_{n3}^{(2)}x_3 + \cdots + a_{nn}^{(2)}x_n = b_n^{(2)} \end{cases} \tag{2.22}$$

照此消元,直至第 $n-1$ 步得到三角形方程组

$$\begin{cases} a_{11}^{(0)}x_1 + a_{12}^{(0)}x_2 + a_{13}^{(0)}x_3 + \cdots + a_{1n}^{(0)}x_n = b_1^{(0)} \\ \qquad a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \cdots + a_{2n}^{(1)}x_n = b_2^{(1)} \\ \qquad\qquad a_{33}^{(2)}x_3 + \cdots + a_{3n}^{(2)}x_n = b_3^{(2)} \\ \qquad\qquad\qquad \vdots \\ \qquad\qquad\qquad a_{nn}^{(n-1)}x_n = b_n^{(n-1)} \end{cases} \tag{2.23}$$

消元过程到此结束,其中系数 $a_{ij}^{(k)}$ 和 $b_{ij}^{(k)}$ 的计算公式是

$$\begin{cases} a_{ij}^{(k)} = a_{ij}^{(k-1)} - (a_{ik}^{(k-1)}/a_{kk}^{(k-1)})a_{kj}^{(k-1)} \\ b_i^{(k)} = b_i^{(k-1)} - (a_{ik}^{(k-1)}/a_{kk}^{(k-1)})b_k^{(k-1)} \end{cases} \tag{2.24}$$

$$(i,j=k+1,k+2,\cdots,n)$$

接下来的回代过程首先由(2.23)的最后方程求出 x_n,依次向上代入求出 $x_{n-1},x_{n-2},\cdots,x_1$,回代过程的计算公式是

$$\begin{cases} x_n = b_n^{(n-1)}/a_{nn}^{(n-1)} \\ x_i = \left(b_i^{(i-1)} - \sum\limits_{j=i+1}^{n} a_{ij}^{(i-1)}x_j\right)/a_{ii}^{(i-1)} \end{cases} \tag{2.25}$$

$$(i=n-1,n-2,\cdots,1)$$

这就是高斯消元法的全过程,公式(2.24)和(2.25)是编写计算机程序的依据.

由方程组(2.20)到(2.23)的消元过程用矩阵初等变换的方法表示就是

$$(\boldsymbol{A},\boldsymbol{b}) = \left(\begin{array}{cccc:c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right)$$

$$\xrightarrow[(i=2,\cdots,n)]{r_i-(a_{i1}/a_{11})r_1}\left(\begin{array}{cccc:c} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} & b_1^{(0)} \\ & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ & \vdots & & \vdots & \vdots \\ & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{array}\right)$$

$$\xrightarrow[(i=3,\cdots,n)]{r_i-(a_{i2}^{(1)}/a_{22}^{(1)})r_2}\left(\begin{array}{ccccc:c} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} & \cdots & a_{1n}^{(0)} & b_1^{(0)} \\ & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ & & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\ & & \vdots & & \vdots & \vdots \\ & & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right)$$

$$\rightarrow\cdots\rightarrow\left(\begin{array}{ccccc:c} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} & \cdots & a_{1n}^{(0)} & b_1^{(0)} \\ & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ & & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\ & & & \ddots & \vdots & \vdots \\ & & & & a_{nn}^{(n-1)} & b_n^{(n-1)} \end{array}\right)$$

2.4.2 按列选主元的高斯消元法

由于消元过程中位于系数矩阵主对角线上的元素 $a_{ii}^{(r)}$ 起着消元的主要作用,故称它们为**主元素**或**主元**. 上面指出消元过程中若出现 $a_{ii}^{(i-1)}=0$ 时,可在该列 $a_{ii}^{(i-1)}$ 之下选一非零元素,比如是 $a_{ji}^{(i-1)}(j>i)$作为主元,作行交换 $r_i\leftrightarrow r_j$,使 $a_{ji}^{(i-1)}$ 位于(i,i)位置,充当主元. 当 A 非奇异时这样的非零元素 $a_{ji}^{(i-1)}$ 一定存在,为什么?请读者思考.

上述这种按列另选主元的方法并不局限在 $a_{ii}^{(i-1)}=0$ 的时候,即便 $a_{ii}^{(i-1)}\neq0$,但 $a_{ii}^{(i-1)}$ 按绝对值是很小的量时亦应另选绝对值大的 $a_{ji}^{(i-1)}(j>i)$为主元,这是因为在数值计算中,以绝对值很小的数作除数将会带来较大的舍入误差,以致影响最后解的精确性. 因此,若按$\max\limits_{j\geqslant i}\{|a_{ji}^{(i-1)}|\}$选取的那个绝对值最大的元素作为主元则

可避免上述两种情形，这就是**按列选主元的高斯消元法**.

例 2.14 用按列选主元法解方程组

$$\begin{cases} -3x_1+2x_2+6x_3=4 \\ 10x_1-7x_2=7 \\ 5x_1-x_2+5x_3=6 \end{cases}$$

解

$$(\boldsymbol{A},\boldsymbol{b})=\left(\begin{array}{ccc:c} -3 & 2 & 6 & 4 \\ \boxed{10} & -7 & 0 & 7 \\ 5 & -1 & 5 & 6 \end{array}\right) \xrightarrow{r_1\leftrightarrow r_2} \left(\begin{array}{ccc:c} \boxed{10} & -7 & 0 & 7 \\ -3 & 2 & 6 & 4 \\ 5 & -1 & 5 & 6 \end{array}\right)$$

$$\xrightarrow[r_3+\frac{-5}{10}r_1]{r_2+\frac{3}{10}r_1} \left(\begin{array}{ccc:c} \boxed{10} & -7 & 0 & 7 \\ 0 & -0.1 & 6 & 6.1 \\ 0 & \boxed{2.5} & 5 & 2.5 \end{array}\right)$$

$$\xrightarrow[r_3+\frac{0.1}{2.5}r_2]{r_2\leftrightarrow r_3} \left(\begin{array}{ccc:c} \boxed{10} & -7 & 0 & 7 \\ 0 & \boxed{2.5} & 5 & 2.5 \\ 0 & 0 & \boxed{6.2} & 6.2 \end{array}\right)$$

回代得解

$$x_3=1,\quad x_2=-1,\quad x_1=0.$$

方框中的数为按列选的主元.

应用高斯消元法解 n 阶线性方程组要作 $\frac{n^3}{3}+n^2-\frac{n}{3}$ 次乘除法，解 18 阶方程组大约作 2.3×10^3 次乘除法，仍用每秒百万次计算机计算仅需 2.3×10^{-3} 秒，这与用克莱姆法则求解相比充分显示了消元法的威力，由此不难理解为什么有了现代化的计算机这种高速度的计算工具，仍要潜心研究数学方法.

上面针对 n 阶线性方程组叙述了高斯消元法，这种方法亦可用来讨论一般的线性方程组(1.1)，详见第 3 章.

2.4.3 高斯-若当消元法

高斯-若当(Jordan)消元法是将方程组(2.20)化为对角形方

程组

$$\begin{cases} x_1 & & & & = g_1 \\ & x_2 & & & = g_2 \\ & & \ddots & & \vdots \\ & & & x_n & = g_n \end{cases} \tag{2.26}$$

也就是对(2.18)的增广矩阵($\boldsymbol{A}$,$\boldsymbol{b}$)施行初等行变换化为

$$(\boldsymbol{I}_n, \boldsymbol{g}) = \left(\begin{array}{cccc:c} 1 & & & & g_1 \\ & 1 & & & g_2 \\ & & \ddots & & \vdots \\ & & & 1 & g_n \end{array}\right) \tag{2.27}$$

即

$$(\boldsymbol{A}, \boldsymbol{b}) \sim (\boldsymbol{I}_n, \boldsymbol{g}) \tag{2.28}$$

方程组(2.26)的解,也就是方程组(2.20)的解 $\boldsymbol{x}=\boldsymbol{g}$.

高斯-若当消元法与高斯消元法的区别在于,选主元后用它消去该列主元素上下的元素,而不仅是主元素以下的元素.另一区别是这种方法不需回代,因此也称之为无回代消元法.

高斯-若当消元法的计算量为$\frac{n^3}{2}+\frac{n^2}{2}-\frac{n}{2}$次乘除法,虽然它比高斯消元法的计算量约多$\frac{1}{6}$,但它在矩阵求逆中十分有用.

*2.4.4 利用矩阵的 *LU* 分解法解线性方程组

消元法是将方程组化为一个容易求解的三角形方程组,而矩阵的 $\boldsymbol{LU}$ 分解法是将方程组化为两个三角形方程组.

在第 1 章已经知道如何将矩阵作 $\boldsymbol{LU}$ 分解,若 $\boldsymbol{A}=\boldsymbol{LU}$,而 $\boldsymbol{Ax}=\boldsymbol{b}$,那么

$$\boldsymbol{LUx}=\boldsymbol{b}$$

设

$$\boldsymbol{Ux}=\boldsymbol{y} \tag{2.29}$$

则

$$\boldsymbol{Ly}=\boldsymbol{b} \tag{2.30}$$

这就将 $\boldsymbol{Ax}=\boldsymbol{b}$ 化成了两个三角形方程组(2.29)与(2.30).求解

时，先由(2.28)求出 $\boldsymbol{y}$，再由(2.29)求出 $\boldsymbol{x}$. 求解(2.30)与(2.29)都只是一个代入计算过程.

例 2.15 应用 $\boldsymbol{LU}$ 分解法解方程组

$$\begin{cases} 2x_1+2x_2+3x_3=9 \\ 2x_1+4x_2=10 \\ x_1+2x_3=3 \end{cases}$$

解

$$\boldsymbol{A}=\begin{pmatrix} 2 & 2 & 3 \\ 2 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/2 & -1/2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 2 & 3 \\ 0 & 2 & -3 \\ 0 & 0 & -1 \end{pmatrix}$$

(见习题一,15)

由方程组

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/2 & -1/2 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}=\begin{pmatrix} 9 \\ 10 \\ 3 \end{pmatrix}$$

解得
$$(y_1,y_2,y_3)^{\mathrm{T}}=(9,1,-1)^{\mathrm{T}}$$

再由方程组

$$\begin{pmatrix} 2 & 2 & 3 \\ 0 & 2 & -3 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 9 \\ 1 \\ -1 \end{pmatrix}$$

解得
$$(x_1,x_2,x_3)^{\mathrm{T}}=(1,2,1)^{\mathrm{T}}$$

由于 $\boldsymbol{LU}$ 分解法使系数矩阵的计算与右端向量 $\boldsymbol{b}$ 的计算分开，这就为仅右端项不同的一批方程组的求解减少了不少的计算量.

2.5 消元法的应用

2.5.1 化矩阵为阶梯形矩阵

利用初等行变换也就是利用消元法可以把矩阵化为等价的阶

梯形矩阵,这种形式的矩阵的应用见第 4 章.所谓**阶梯形矩阵**指的是,每一非零行的第一个非零元素前的零元素个数随行序数的增加而增加,如以下矩阵皆为阶梯形矩阵

$$\begin{pmatrix} 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

例 2.16 试用消元法化矩阵 $\boldsymbol{A}$ 为阶梯形矩阵,

$$\boldsymbol{A} = \begin{pmatrix} 1 & -2 & -1 & 0 & 2 \\ -2 & 4 & 2 & 6 & -6 \\ 2 & -1 & 0 & 2 & 3 \\ 3 & 3 & 3 & 3 & 4 \end{pmatrix}$$

解

$$\boldsymbol{A} \xrightarrow[r_4 - 3r_1]{\substack{r_2 + 2r_1 \\ r_3 - 2r_1}} \begin{pmatrix} 1 & -2 & -1 & 0 & -2 \\ 0 & 0 & 0 & 6 & -2 \\ 0 & 3 & 2 & 2 & -1 \\ 0 & 9 & 6 & 3 & -2 \end{pmatrix}$$

$$\xrightarrow{r_2 \leftrightarrow r_4} \begin{pmatrix} 1 & -2 & -1 & 0 & -2 \\ 0 & 9 & 6 & 3 & -2 \\ 0 & 3 & 2 & 2 & -1 \\ 0 & 0 & 0 & 6 & -2 \end{pmatrix}$$

$$\xrightarrow[r_4 - 6r_3]{r_3 - \frac{1}{3}r_2} \begin{pmatrix} 1 & -2 & -1 & 0 & 2 \\ 0 & 9 & 6 & 3 & 2 \\ 0 & 0 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \boldsymbol{B}$$

$\boldsymbol{B}$ 即为所求的与 $\boldsymbol{A}$ 等价的的阶梯形矩阵.

2.5.2 求逆矩阵

前面讲过,求逆矩阵的目的之一是为了解方程组 $\boldsymbol{Ax}=\boldsymbol{b}$,若能求得方阵 $\boldsymbol{A}$ 的逆 $\boldsymbol{A}^{-1}$,则方程组的解就是 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$. 实际上矩阵求

逆的目的并不局限于解方程组，它在线性代数的许多方面都是不可缺少的. 但是到现在为止我们仅知道一种求逆矩阵的方法，即伴随矩阵法，$\boldsymbol{A}^{-1}=\frac{\mathrm{adj}\boldsymbol{A}}{\det\boldsymbol{A}}$. 这种方法不适用于高阶矩阵，这是由于高阶行列式的计算量过大的缘故. 我们应当觅求新的矩阵求逆的方法，一种实用的方法的思想就是将先求 $\boldsymbol{A}^{-1}$ 再求 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解的过程倒过来，变成先求方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解 $\boldsymbol{x}$（这易由消元法实现）再求逆矩阵 $\boldsymbol{A}^{-1}$.

考虑矩阵方程

$$\boldsymbol{AX}=\boldsymbol{I}_n \tag{2.31}$$

$\boldsymbol{A}$ 为 n 阶非奇异方阵，$\boldsymbol{X}$ 为 n 阶未知量方阵，$\boldsymbol{A}^{-1}$左乘(2.31)式，得

$$\boldsymbol{X}=\boldsymbol{A}^{-1}$$

这表明 $\boldsymbol{A}^{-1}$实为矩阵方程(2.31)的解矩阵 X，因此求 $\boldsymbol{A}^{-1}$的问题转化为求解矩阵方程(2.31).

首先将(2.31)的矩阵 $\boldsymbol{X}$ 和 $\boldsymbol{I}_n$ 写成向量形式

$$\boldsymbol{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n),\qquad \boldsymbol{I}_n=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)$$

其中 $\boldsymbol{x}_i$ 为列向量，$\boldsymbol{e}_i(i=1,2,\cdots,n)$为单位坐标向量.

于是式(2.31)成为

$$\boldsymbol{A}(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)$$

这是 n 个 n 阶线性方程组

$$\boldsymbol{Ax}_i=\boldsymbol{e}_i\qquad (i=1,2,\cdots,n) \tag{2.32}$$

根据高斯-若当消元法求解，由(2.28)式知

$$(\boldsymbol{A},\boldsymbol{e}_i)\sim(\boldsymbol{I}_n,\boldsymbol{x}_i)\qquad (i=1,2,\cdots,n)$$

对(2.32)的 n 个方程组同时施行相同的初等行变换，就有

$$\begin{aligned}(\boldsymbol{A},\boldsymbol{I}_n)&=(\boldsymbol{A},\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)\sim(\boldsymbol{I}_n,\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\\&=(\boldsymbol{I}_n,\boldsymbol{X})=(\boldsymbol{I}_n,\boldsymbol{A}^{-1})\end{aligned} \tag{2.33}$$

这表明只需对$(\boldsymbol{A},\boldsymbol{I}_n)$施行初等行变换，当把 $\boldsymbol{A}$ 变换为 $\boldsymbol{I}_n$ 时，原来的 $\boldsymbol{I}_n$ 就变换为 $\boldsymbol{A}^{-1}$.

这种利用初等变换求逆矩阵的方法步骤归纳为：

(1)将矩阵 $\boldsymbol{A}$ 与同阶的单位方阵 $\boldsymbol{I}$ 拼成$(\boldsymbol{A},\boldsymbol{I})$；

(2)对$(\boldsymbol{A},\boldsymbol{I})$施行初等行变换，目标是将$\boldsymbol{A}$变换成$I$；

(3)当$\boldsymbol{A}$变换为$\boldsymbol{I}$时，原来的$\boldsymbol{I}$变换成$\boldsymbol{A}^{-1}$，即$(\boldsymbol{A},\boldsymbol{I})\rightarrow(\boldsymbol{I},\boldsymbol{A}^{-1})$.

注意：1)若$\boldsymbol{A}$不可逆，则无论如何变换也不会将$\boldsymbol{A}$变换为$\boldsymbol{I}$，因而不可能得到$\boldsymbol{A}^{-1}$；

2)若将$\boldsymbol{A}$、$\boldsymbol{I}$拼成$\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{I}\end{pmatrix}$，施行初等列变换，结果是$\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{I}\end{pmatrix}\rightarrow\begin{pmatrix}\boldsymbol{I}\\ \boldsymbol{A}^{-1}\end{pmatrix}$.

例 2.17 求例 2.11 矩阵$\boldsymbol{A}$的逆矩阵$\boldsymbol{A}^{-1}$.

解

$$(\boldsymbol{A},\boldsymbol{I})=\left(\begin{array}{ccc:ccc}-1 & 1 & 1 & 1 & 0 & 0\\ 1 & 0 & -2 & 0 & 1 & 0\\ 1 & -2 & 1 & 0 & 0 & 1\end{array}\right)$$

$$\xrightarrow[r_3+r_1]{r_2+r_1}\left(\begin{array}{ccc:ccc}-1 & 1 & 1 & 1 & 0 & 0\\ 0 & 1 & -1 & 1 & 1 & 0\\ 0 & -1 & 2 & 1 & 0 & 1\end{array}\right)$$

$$\xrightarrow[r_3+r_2]{(-1)r_1}\left(\begin{array}{ccc:ccc}1 & -1 & -1 & -1 & 0 & 0\\ 0 & 1 & -1 & 1 & 1 & 0\\ 0 & 0 & 1 & 2 & 1 & 1\end{array}\right)$$

$$\xrightarrow[r_1+r_2+r_3]{r_2+r_3}\left(\begin{array}{ccc:ccc}1 & 0 & 0 & 4 & 3 & 2\\ 0 & 1 & 0 & 3 & 2 & 1\\ 0 & 0 & 1 & 2 & 1 & 1\end{array}\right)$$

于是

$$\boldsymbol{A}^{-1}=\begin{pmatrix}4 & 3 & 2\\ 3 & 2 & 1\\ 2 & 1 & 1\end{pmatrix}$$

与例 2.11 的结果相同.

习 题 二

1. 设 $\det\boldsymbol{A}=\begin{vmatrix}1&2&3\\3&2&1\\0&0&2\end{vmatrix}$，写出 M_{21}，M_{22}，M_{23}，D_{21}，D_{22}，D_{23}.

2. 计算行列式

(1) $\begin{vmatrix}3&1&-1&2\\-5&1&3&-4\\2&0&1&-1\\1&-5&3&-3\end{vmatrix}$ (2) $\begin{vmatrix}3&1&1&1\\1&3&1&1\\1&1&3&1\\1&1&1&3\end{vmatrix}$

(3) $\begin{vmatrix}1&1&1&1\\4&3&7&-5\\16&9&49&25\\64&27&343&-125\end{vmatrix}$

(4) $\begin{vmatrix}x&y&0&\cdots&0&0\\0&x&y&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&x&y\\y&0&0&\cdots&0&x\end{vmatrix}$ (5) $\begin{vmatrix}&&&\lambda_1\\&&\lambda_2&\\&\iddots&&\\\lambda_n&&&\end{vmatrix}$

3. 证明以下命题

(1)证明行列式性质 3 的推论 2.

(2)证明行列式性质 5.

(3)证明 $\det(k\boldsymbol{A})=k^n\det\boldsymbol{A}$($\boldsymbol{A}$ 为 n 阶方阵，k 为任意常数).

(4)若 $\boldsymbol{A}$ 为非奇异矩阵，证明 $\det\boldsymbol{A}^{-1}=(\det\boldsymbol{A})^{-1}$.

(5)设 $\boldsymbol{A}$，$\boldsymbol{B}$ 为 n 阶方阵，则 $\det(\operatorname{adj}\boldsymbol{A})=(\det\boldsymbol{A})^{n-1}$.

(6) 试证：若 $\boldsymbol{A}$ 为 n 阶幂等矩阵，则 $\boldsymbol{A}=\boldsymbol{I}_n$，否则 $\boldsymbol{A}$ 必是奇异方阵.

(7) 设 $\boldsymbol{A}$，$\boldsymbol{B}$ 为同阶方阵，且均可逆，则 $\operatorname{adj}(\boldsymbol{AB})=\operatorname{adj}\boldsymbol{B}\operatorname{adj}\boldsymbol{A}$.

4. 解下列矩阵方程：

(1) $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \boldsymbol{X} = \begin{pmatrix} 4 & -6 \\ 2 & 1 \end{pmatrix}$

(2) $\begin{pmatrix} 1 & 4 \\ -1 & 2 \end{pmatrix} \boldsymbol{X} \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 0 & -1 \end{pmatrix}$

(3)$^{\triangle}$ $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{B}$, $\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 5 & -3 \end{pmatrix}$

5$^{\triangle}$**.** 设 $\det\boldsymbol{A} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{vmatrix}$

$$\det\boldsymbol{B} = \begin{vmatrix} a & -a & b & -b \\ a & a & -b & -b \\ -b & b & a & -a \\ b & b & a & a \end{vmatrix}$$

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

6. 根据克莱姆法则解方程组

(1) $\begin{cases} 2x+3y-z=-4 \\ x-y+z=5 \\ 7x-6y-4z=1 \end{cases}$

(2) $\begin{cases} 2x_1+x_2=1 \\ x_1+2x_2+x_3=0 \\ x_2+2x_3+x_4=0 \\ x_3+2x_4=1 \end{cases}$

7. 设方程组

$$\begin{cases} x_1+ax_2+a^2x_3=a^3 \\ x_1+bx_2+b^2x_3=b^3 \\ x_1+cx_2+c^2x_3=c^3 \end{cases}$$

试问 a、b、c 满足什么条件时方程组有唯一解，并求出这个条件下的唯一解.

8. 用按列选主元的高斯消元法解方程组

(1) $\begin{cases} 3x-2y+5z=-1 \\ x+2y-z=5 \\ 5x-y+2z=3; \end{cases}$

(2) $\begin{cases} x_1+4x_2-2x_3+3x_4=6 \\ 2x_1+2x_2+4x_4=2 \\ 3x_1-x_3+2x_4=1 \\ x_1+2x_2+2x_3-3x_4=8 \end{cases}$

9. 利用伴随矩阵法求下列矩阵的逆矩阵

(1) $\boldsymbol{A}=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$　(2) $\boldsymbol{B}=\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 1 & -1 & 0 \end{bmatrix}$

10. 利用初等行变换求下列矩阵的逆矩阵

(1) $\boldsymbol{A}=\begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix}$

(2)$^{\triangle}$设 $\boldsymbol{B}=\begin{bmatrix} 3 & 0 & 0 \\ 1 & 4 & 3 \\ 0 & 0 & 3 \end{bmatrix}$，求$(\boldsymbol{B}-2\boldsymbol{I}_3)^{-1}$

(3) $\boldsymbol{C}=\begin{bmatrix} 0 & c_1 & 0 & \cdots & 0 \\ 0 & 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & c_{n-1} \\ c_n & 0 & 0 & \cdots & 0 \end{bmatrix}$

$(c_i\neq 0, i=1,2,\cdots,n)$.

11$^{\triangle}$. 设 $\boldsymbol{A},\boldsymbol{B}$ 均为非奇异方阵，并且

$$\boldsymbol{X}=\begin{pmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix},\qquad 求\ \boldsymbol{X}^{-1}.$$

12. 已知 $\boldsymbol{AP}=\boldsymbol{PB}$，其中

$$\boldsymbol{B}=\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix},\quad \boldsymbol{P}=\begin{pmatrix}1&0&0\\2&-1&0\\2&1&1\end{pmatrix}$$

求 $\boldsymbol{A}$ 及 $\boldsymbol{A}^5$.

13$^{\triangle}$. 设 $\boldsymbol{A}$ 为 m 阶方阵，$\boldsymbol{B}$ 为 n 阶方阵，且 $|\boldsymbol{A}|=a$，$|\boldsymbol{B}|=b$，$\boldsymbol{C}=\begin{pmatrix}\boldsymbol{O}&\boldsymbol{A}\\\boldsymbol{B}&\boldsymbol{O}\end{pmatrix}$，则 $|\boldsymbol{C}|=$（　）.

（A）ab　（B）$(-1)^{m+n}ab$　（C）$-ab$　（D）$(-1)^{mn}ab$

第 3 章　矩阵的秩和线性方程组的相容性定理

解 n 阶线性代数方程组的克莱姆法则指出，对 n 阶线性方程组，当其系数矩阵为非奇异方阵时有唯一解，而当系数矩阵为奇异方阵，或更一般地为长方阵时，克莱姆法则就无法直接利用了. 因此，需要进一步讨论一般线性方程组何时有解，何时无解以及有解时，何时有唯一解，何时有无穷多个解的问题. 围绕这些问题，我们引进矩阵的秩的概念，得到线性方程组的相容性定理，从而对上述问题给出明确的答案.

3.1　矩阵的秩

克莱姆法则只是给出了线性方程组解的存在唯一性的一个充分条件. 对一般线性方程组，因常常不满足克莱姆法则的条件而不能直接使用，但能否经过某些改变便可用上克莱姆法则呢？若能，究竟经过怎样的改变才能用上克莱姆法则呢？先看下面的例子.

3.1.1　引例

对齐次线性方程组

$$\begin{cases} x+2y-z=0 \\ 2x-3y+z=0 \\ 4x+y-z=0 \end{cases} \tag{3.1}$$

由于其系数矩阵

$$\boldsymbol{A}=\begin{pmatrix} 1 & 2 & -1 \\ 2 & -3 & 1 \\ 4 & 1 & -1 \end{pmatrix}$$

是奇异方阵而不能直接用克莱姆法则，但可以看出，第 3 个方程（即 $\boldsymbol{A}$ 的第 3 行）可由第 1 个方程（$\boldsymbol{A}$ 的第 1 行）的 2 倍加上第 2 个方程（$\boldsymbol{A}$ 的第 2 行）得到. 这样，第 1、第 2 个方程的公共解自然是第 3 个方程的解，可见，第 3 个方程是多余的，称其为方程组的**多余方程**，应用高斯消元法，从方程组(3.1)中删去这种多余方程，得

$$\begin{cases} x+2y-z=0 \\ -7y+3z=0 \end{cases} \quad \text{或} \quad \begin{cases} x+2y-z=0 \\ 2x-3y+z=0 \end{cases} \tag{3.2}$$

(3.2)已不再含多余方程，且与(3.1)同解，(3.2)称为(3.1)的**保留方程组**，这样，求解(3.1)归结为求解同解的保留方程组(3.2)，若把(3.2)中的 z 看作可取任何值的参数，并把含 z 的项移至等号右边，则 x,y 的系数行列式

$$\begin{vmatrix} 1 & 2 \\ 0 & -7 \end{vmatrix} = -7 \neq 0 \quad \text{或} \quad \begin{vmatrix} 1 & 2 \\ 2 & -3 \end{vmatrix} = -7 \neq 0$$

因此可用克莱姆法则求解得 $x=\dfrac{1}{7}z, y=\dfrac{3}{7}z, z=z$，解中含有一个任意参数.

把上述删去多余方程求得保留方程组的过程用矩阵表示就是

$$\boldsymbol{A} \xrightarrow[r_3-4r_1]{r_2-2r_1} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -7 & 3 \\ 0 & -7 & 3 \end{pmatrix} \xrightarrow{r_3-r_2} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -7 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

很明显，最后变成的阶梯形矩阵中，元素全为零的行（简称为零行）对应的方程就是多余方程，而元素不全为零的行（简称为非零行）对应的方程组就是(3.2).

对一般齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \quad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{cases} \tag{3.3}$$

设法利用克莱姆法则的基本思想和步骤与求解(3.1)类似，即先用高斯消元法删去方程组中的多余方程（如果有多余方程的话）得到

有 $r(r\leqslant\min\{m,n\})$ 个方程的保留方程组，然后从保留方程组中选出 r 个未知数，使它们的系数行列式不为零，其余的未知数看作任意参数，这样就可以利用克莱姆法则了.

为表示方便，设方程组(3.3)的后面 $m-r$ 个方程是多余的，则可得到保留方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1r}x_r+a_{1r+1}x_{r+1}+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2r}x_r+a_{2r+1}x_{r+1}+\cdots+a_{2n}x_n=0 \\ \quad\vdots \\ a_{r1}x_1+a_{r2}x_2+\cdots+a_{rr}x_r+a_{r,r+1}x_{r+1}+\cdots+a_{rn}x_n=0 \end{cases}$$

仍为表示方便，设 $x_1,x_2,\cdots,x_r$ 的系数行列式

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}\neq 0$$

则可把 $x_{r+1},\cdots,x_n$ 看作可取任意值的参数，并把含它们的项移至等号右边，从而利用克莱姆法则知(3.3)存在含有 $n-r$ 个任意参数的解.

从上述讨论易见，对一般齐次线性方程组，欲用克莱姆法则的关键是从中删去多余方程得到保留方程组，并从保留方程组中找出 r 个未知数，使其系数行列式不为零. 这自然会提出下列问题：

(1) 如何判定方程组(3.3)是否有多余方程？

(2) 如何求方程组(3.3)的保留方程组？

由线性方程组与矩阵的对应关系可知对线性方程组的同解变形相当于对其相应的矩阵施行初等行变换. 因此，结合上述问题，我们引进矩阵的秩的概念.

3.1.2 矩阵的秩

定义 3.1 矩阵 $\boldsymbol{A}$ 中最大非奇异子方阵的阶数称为**矩阵 $\boldsymbol{A}$ 的秩**，记为 $R(\boldsymbol{A})$.

例 3.1 方程组(3.1)的系数矩阵 $\boldsymbol{A}$ 的秩 $R(\boldsymbol{A})=2$，这是由于

$|A|=0$,而 $\begin{vmatrix}1 & 2\\2 & -3\end{vmatrix}=-7\neq 0$.

例 3.2 求下列矩阵的秩

$$\boldsymbol{A}=\begin{pmatrix}1 & 2 & 4 & 1\\2 & 4 & 8 & 2\\3 & 6 & 2 & 0\end{pmatrix}$$

解 由于 A 的第 2 行是第 1 行的 2 倍,所以 $\boldsymbol{A}$ 的任一个三阶子方阵都是奇异的,因此有 $R(\boldsymbol{A})<3$,但因 $\boldsymbol{A}$ 有一个二阶子方阵的行列式

$$\begin{vmatrix}1 & 4\\3 & 2\end{vmatrix}=-10\neq 0$$

所以 $R(\boldsymbol{A})=2$.

由矩阵的秩的定义,容易得到下列的结论:

定理 3.1 **若矩阵 $\boldsymbol{A}$ 有一个 r 阶子方阵非奇异,且所有 $r+1$ 阶的子方阵都是奇异的,则 $R(\boldsymbol{A})=r$.**

定理 3.2 **若 $\boldsymbol{A}$ 是 $m\times n$ 维矩阵,则 $R(\boldsymbol{A})\leqslant \min(m,n)$.**

定理 3.3 **若 $\boldsymbol{A}$ 是 n 阶方阵,则 $R(\boldsymbol{A})=n$ 的充分必要条件是 $\boldsymbol{A}$ 为非奇异的.**

定理 3.4 **对任何矩阵 $\boldsymbol{A}$,有 $R(\boldsymbol{A})=R(\boldsymbol{A}^{\mathrm{T}})$.**

$R(\boldsymbol{A})=n$ 的 n 阶方阵 $\boldsymbol{A}$ 称为**满秩矩阵**,否则称为**降秩矩阵**.

一个明显的事实是,零矩阵是秩为零的唯一矩阵.

从上面的两个例子可以看出,按秩的定义求矩阵的秩时,要计算一些子方阵的行列式(称为**矩阵的子式**),当子式阶数较高时很不方便,因此需要给出一种有效的方法. 为此,我们在下节先介绍有关的概念和结论.

3.2 初等方阵

在第 2 章里,我们引进了矩阵的三种初等变换,并把它们用于消元法和矩阵求逆. 因此,矩阵的初等变换是矩阵的一种最基本、

最常用的运算.这些初等变换实际上都与称为初等方阵的一些特殊矩阵对应,而对矩阵施行初等变换相当于矩阵与初等方阵作乘法.本节介绍初等方阵的有关知识.

定义 3.2 单位阵经一次初等变换得到的矩阵称为**初等方阵**.

对应于三种初等变换,便有三种初等方阵.

1. **对调两行(或两列).** 把单位矩阵 $\boldsymbol{I}$ 中的第 i、j 两行对调(即 $r_i \leftrightarrow r_j$)得到的初等方阵记为 $\boldsymbol{E}_{ij}$,即

$$\boldsymbol{E}_{ij}=\left(\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 0 & & \cdots & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & \vdots & & \ddots & & \vdots & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & \cdots & & 0 & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right)\begin{array}{l}
\\ \\ \\ \text{第 } i \text{ 行} \\ \\ \\ \\ \text{第 } j \text{ 行} \\ \\ \\ \\
\end{array}$$

容易验证,用 m 阶初等方阵 $\boldsymbol{E}_{ij}$ 左乘 $m\times n$ 维矩阵 $\boldsymbol{A}$,即 $\boldsymbol{E}_{ij}\boldsymbol{A}$ 就是对 $\boldsymbol{A}$ 施行 $r_i \leftrightarrow r_j$ 变换,类似地,用 n 阶初等方阵 $\boldsymbol{E}_{ij}$ 右乘 $\boldsymbol{A}$ 即 $\boldsymbol{A}\boldsymbol{E}_{ij}$ 相当于对 $\boldsymbol{A}$ 施行 $c_i \leftrightarrow c_j$ 变换.

2. **用数 $k\neq 0$ 乘某行(或列).** 用不等于零的数 k 乘单位矩阵 $\boldsymbol{I}$ 的第 i 行得到的初等方阵记为 $\boldsymbol{E}_i(k)$,即

$$\boldsymbol{E}_i(k)=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&k&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{matrix}\\ \\ \\ \text{第 } i \text{ 行}\\ \\ \\ \\ \end{matrix}$$

容易验证，$\boldsymbol{A}\xrightarrow{kr_i}\boldsymbol{B}=\boldsymbol{E}_i(k)\boldsymbol{A}$，$\boldsymbol{A}\xrightarrow{kc_i}\boldsymbol{B}=\boldsymbol{A}\boldsymbol{E}_i(k)$.

3. **用数 k 乘某行(或列)的诸元素加到另一行(或列)的对应元素上去.** 用数 k 乘单位矩阵的第 j 行加到第 i 行上去得到的初等方阵记为 $\boldsymbol{E}_{ij}(k)$，即

$$\boldsymbol{E}_{ij}(k)=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&\cdots&k&&\\&&&\ddots&\vdots&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{matrix}\\ \\ \text{第 } i \text{ 行}\\ \\ \text{第 } j \text{ 行}\\ \\ \\ \end{matrix}$$

容易验证，$\boldsymbol{A}\xrightarrow{r_i+kr_j}\boldsymbol{B}=\boldsymbol{E}_{ij}(k)\boldsymbol{A}$，而 $\boldsymbol{A}\xrightarrow{c_j+kc_i}\boldsymbol{B}=\boldsymbol{A}\boldsymbol{E}_{ij}(k)$.

总之，对 $\boldsymbol{A}$ 施行初等行变换相当于用初等方阵左乘 $\boldsymbol{A}$，而对 $\boldsymbol{A}$ 施行初等列变换相当于用初等方阵右乘 $\boldsymbol{A}$.

显然，三种初等方阵都是可逆的，且逆矩阵也是初等方阵：

$$(\boldsymbol{E}_{ij})^{-1}=\boldsymbol{E}_{ij}，\ (\boldsymbol{E}_i(k))^{-1}=\boldsymbol{E}_i\left(\frac{1}{k}\right)，\ (\boldsymbol{E}_{ij}(k))^{-1}=\boldsymbol{E}_{ij}(-k)$$

由第 2 章知道，对非奇异方阵 $\boldsymbol{A}$ 用 $(\boldsymbol{A},\boldsymbol{I})\longrightarrow(\boldsymbol{I},\boldsymbol{A}^{-1})$ 求 $\boldsymbol{A}^{-1}$ 的方法实际上是对 $\boldsymbol{A}$ 和 $\boldsymbol{I}$ 同时施行一系列相同的初等行变换，当 $\boldsymbol{A}$ 变成 $\boldsymbol{I}$ 时，$\boldsymbol{I}$ 变得的就是 $\boldsymbol{A}^{-1}$. 用现在的术语就是用一系列初等方阵 $P_1,P_2,\cdots,P_t$ 依次左乘 $(\boldsymbol{A},\boldsymbol{I})$，便有

$$\boldsymbol{P}_t\cdots\boldsymbol{P}_1(\boldsymbol{A},\boldsymbol{I})=(\boldsymbol{P}_t\cdots\boldsymbol{P}_1\boldsymbol{A},\boldsymbol{P}_t\cdots\boldsymbol{P}_1\boldsymbol{I})=(\boldsymbol{I},\boldsymbol{A}^{-1})$$

即 $$\boldsymbol{A}^{-1}=\boldsymbol{P}_t\cdots\boldsymbol{P}_1$$

于是我们得到下面的结论.

定理 3.5 方阵可逆的充要条件是它可以表示成有限个初等方阵的乘积.

由于对矩阵 $\boldsymbol{A}$ 施行一系列初等行(列)变换相当于对 $\boldsymbol{A}$ 左(右)乘一系列相应的初等方阵。因此,若把左(右)乘的初等方阵的乘积记为 $\boldsymbol{P}(\boldsymbol{Q})$,由上述定理便得到下面的结论.

定理 3.6 $m\times n$ 维矩阵 $\boldsymbol{A}$、$\boldsymbol{B}$ 等价的充要条件是存在 m 阶可逆方阵 $\boldsymbol{P}$ 和 n 阶可逆方阵 $\boldsymbol{Q}$,使 $\boldsymbol{PAQ}=\boldsymbol{B}$.

3.3 矩阵的秩的求法和矩阵的标准形

3.3.1 等价矩阵具有相同的秩

在第 2 章里给出了等价矩阵的概念,即若矩阵 $\boldsymbol{A}$ 经有限次初等变换变到矩阵 $\boldsymbol{B}$,则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价,记作 $\boldsymbol{A}\sim\boldsymbol{B}$. 显然,等价关系具有:i)反身性,即 $\boldsymbol{A}\sim\boldsymbol{A}$;ii)对称性,即若 $\boldsymbol{A}\sim\boldsymbol{B}$,则 $\boldsymbol{B}\sim\boldsymbol{A}$;iii)传递性,即若 $\boldsymbol{A}\sim\boldsymbol{B}$,$\boldsymbol{B}\sim\boldsymbol{C}$,则 $\boldsymbol{A}\sim\boldsymbol{C}$.

对于等价的矩阵 $\boldsymbol{A}$、$\boldsymbol{B}$,有下列重要结论.

定理 3.7 若 $\boldsymbol{A}\sim\boldsymbol{B}$,则 $R(\boldsymbol{A})=R(\boldsymbol{B})$,即初等变换不改变矩阵的秩.

证 设 $R(\boldsymbol{A})=r$,首先考察第一、第二种初等行变换,即 $r_i\leftrightarrow r_j$ 和 $kr_i(k\neq 0)$.

对任何方阵,由行列式的性质,第一、二种初等行变换不改变它的奇异性,即奇异方阵仍保留奇异,非奇异方阵仍保留非奇异. 对矩阵 $\boldsymbol{A}$,第一、第二种初等行变换对 $\boldsymbol{A}$ 的任何子方阵具有完全相同的作用,因此第一、第二种初等行变换不改变矩阵 $\boldsymbol{A}$ 的秩.

其次考察第三种初等行变换,由于 $R(\boldsymbol{A})=r$,所以 $\boldsymbol{A}$ 的所有 $r+1$阶子方阵都是奇异的,设 $\boldsymbol{A}\overset{r_i+kr_j}{\sim}\boldsymbol{B}$. 显然,$\boldsymbol{B}$ 中只有含第 i 行的

子方阵不同于 $\boldsymbol{A}$ 中相应的子方阵，用 $\boldsymbol{S}_{r+1}(\boldsymbol{B})$ 表示 $\boldsymbol{B}$ 中含第 i 行的 $r+1$ 阶子方阵，$\boldsymbol{S}_{r+1}(\boldsymbol{A})$ 表示矩阵 $\boldsymbol{A}$ 中相应的子方阵，对这样的 $\boldsymbol{S}_{r+1}(\boldsymbol{B})$：

(1)若它含有第 j 行，由行列式性质，$|\boldsymbol{S}_{r+1}(\boldsymbol{B})|=|\boldsymbol{S}_{r+1}(\boldsymbol{A})|$，即经第三种初等行变换，$\boldsymbol{S}_{r+1}(\boldsymbol{A})$ 没改变奇异性；

(2)若它不含第 j 行，为书写方便，不妨取 $\boldsymbol{S}_{r+1}(\boldsymbol{B})$ 为由 $\boldsymbol{B}$ 的前 $r+1$ 列构成的子方阵，由行列式性质 4 有

$$
\begin{aligned}
|\boldsymbol{S}_{r+1}(\boldsymbol{B})| &= \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{i1}+ka_{j1} & \cdots & a_{ir+1}+ka_{jr+1} \\ \cdots & \cdots & \cdots \end{vmatrix} \\
&= \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{ir+1} \\ \cdots & \cdots & \cdots \end{vmatrix} + k \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{j1} & \cdots & a_{jr+1} \\ \cdots & \cdots & \cdots \end{vmatrix}
\end{aligned}
$$

其中未标出的元素均与 $\boldsymbol{A}$ 中相应位置上的元素相同. 很明显，最后的两个行列式是 $\boldsymbol{A}$ 的两个 $r+1$ 阶的行列式，它们都为零. 因此 $|\boldsymbol{S}_{r+1}(\boldsymbol{B})|=0$，即第三种初等行变换不改变 $\boldsymbol{S}_{r+1}(\boldsymbol{A})$ 的奇异性.

由(1)、(2)及定义 3.1 知 $R(\boldsymbol{B})<r+1$ 即 $R(\boldsymbol{B})\leqslant r$，这表明第三种初等行变换没增加 $\boldsymbol{A}$ 的秩，即 $R(\boldsymbol{B})\leqslant R(\boldsymbol{A})$，而 $\boldsymbol{B}$ 经逆变换(也是第三种初等行变换)可变换回到 $\boldsymbol{A}$，所以 $R(\boldsymbol{A})\leqslant R(\boldsymbol{B})$，总之有 $R(\boldsymbol{A})=R(\boldsymbol{B})$.

综合上述证明便得，$\boldsymbol{A}$ 经初等行变换不改变其秩，类似地可证初等列变换也不改变 A 的秩，总之，若 $A\sim B$，则 $R(\boldsymbol{A})=R(\boldsymbol{B})$.

对于同维矩阵，上述定理的逆定理也成立，即当矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 同维时，若 $R(\boldsymbol{A})=R(\boldsymbol{B})$，则 $\boldsymbol{A}\sim\boldsymbol{B}$.

3.3.2　秩的求法、矩阵的标准形

由于初等变换不改变矩阵的秩，所以可以通过初等变换把矩阵的许多元素变为零，从而可以直接看出矩阵的秩. 事实上，利用初等行变换可以把任一个 $m\times n$ 维的矩阵 $\boldsymbol{A}$ 变成与之等价的阶梯形矩阵，若这个阶梯形矩阵仅有 r 个非零行，那么可通过列交换变

成主元 $b_{ii}\neq 0(i=1,2,\cdots,r)$的阶梯形矩阵

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1r} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2r} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & b_{rr} & \cdots & b_{rn} \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

显然它有一个 r 阶的非奇异子方阵,且所有的 $r+1$ 阶子方阵都是奇异的,因此 $R(\boldsymbol{A})=r$.

由此可见,求矩阵的秩时,可以限定只用初等行变换把矩阵化为阶梯形,其中非零行数就是矩阵的秩.

例 3.3 求矩阵 $\boldsymbol{A}$ 的秩

$$\boldsymbol{A}=\begin{pmatrix} 1 & -1 & 2 & 1 & 0 \\ 2 & -2 & 4 & 2 & 0 \\ 3 & 0 & 6 & -1 & 1 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix}$$

解

$$\boldsymbol{A}\xrightarrow[r_3-3r_1]{r_2-2r_1}\begin{pmatrix} 1 & -1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -4 & 1 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{r_4-r_3}\begin{pmatrix} 1 & -1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -4 & 1 \\ 0 & 0 & 0 & 4 & 0 \end{pmatrix}$$

$$\xrightarrow[r_3\leftrightarrow r_4]{r_2\leftrightarrow r_3}\begin{pmatrix} 1 & -1 & 2 & 1 & 0 \\ 0 & 3 & 0 & -4 & 1 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

由最后的阶梯形矩阵即得 $R(\boldsymbol{A})=3$.

对任一个 $m\times n$ 维矩阵 $\boldsymbol{A}$ 经初等行变换变成的阶梯形矩阵若再施行初等列变换可以化为如下的最简形式

$$\boldsymbol{N}=\begin{pmatrix}1&0&\cdots&0&\cdots&0\\ \vdots&\vdots&&\vdots&&\vdots\\ 0&0&\cdots&1&\cdots&0\\ 0&0&\cdots&0&\cdots&0\\ \vdots&\vdots&&\vdots&&\vdots\\ 0&0&\cdots&0&\cdots&0\end{pmatrix}=\begin{array}{c}\begin{matrix}\phantom{\boldsymbol{I}_r}&n-r\end{matrix}\\ \begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{O}\end{pmatrix}_{m-r}\end{array}$$

$\boldsymbol{N}$ 称为 $\boldsymbol{A}$ 的**标准形**,其左上角是一个单位阵 $\boldsymbol{I}_r$,其余元素全为零,其中 r 就是矩阵 $\boldsymbol{A}$ 的秩.

由等价的性质易得下面的两个结论:

(1) 矩阵 $\boldsymbol{A}\sim\boldsymbol{B}$ 的充要条件是它们有相同的标准形.

(2) n 阶方阵 $\boldsymbol{A}$ 满秩的充要条件是 $\boldsymbol{A}\sim\boldsymbol{I}_n$.

例 3.4 对矩阵

$$\boldsymbol{A}=\begin{pmatrix}1&2&2\\ 2&2&2\end{pmatrix}$$

求可逆方阵 $\boldsymbol{P}$、$\boldsymbol{Q}$,使 $\boldsymbol{PAQ}$ 为 $\boldsymbol{A}$ 的标准形(显然,这种 $\boldsymbol{P}$、$\boldsymbol{Q}$ 是不唯一的).

解 先求 $\boldsymbol{P}$

$$(A,I)=\begin{pmatrix}1&2&2&1&0\\ 2&2&2&0&1\end{pmatrix}\xrightarrow{r_2-2r_1}\begin{pmatrix}1&2&2&1&0\\ 0&-2&-2&-2&1\end{pmatrix}$$

所以,$\boldsymbol{P}=\begin{pmatrix}1&0\\ -2&1\end{pmatrix}$,再求 $\boldsymbol{Q}$

$$\begin{pmatrix}1&2&2\\ 0&-2&-2\\ 1&0&0\\ 0&1&0\\ 0&0&1\end{pmatrix}\xrightarrow[c_2-2c_1]{c_3-2c_1}\begin{pmatrix}1&0&0\\ 0&-2&-2\\ 1&-2&-2\\ 0&1&0\\ 0&0&1\end{pmatrix}\xrightarrow[-\frac{1}{2}c_2]{c_3-c_2}\begin{pmatrix}1&0&0\\ 0&1&0\\ 1&1&0\\ 0&-1/2&-1\\ 0&0&1\end{pmatrix}$$

所以

$$
\boldsymbol{Q}=\begin{pmatrix} 1 & 1 & 0 \\ 0 & -1/2 & -1 \\ 0 & 0 & 1 \end{pmatrix}
$$

容易验证，$\boldsymbol{PAQ}$ 就是 $\boldsymbol{A}$ 的标准形 $\boldsymbol{N}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

3.4 线性方程组的相容性定理

至此，我们已经解决了本章 3.1 节提出的两个问题.

(1)方程组(3.3)含有多余方程就是其系数矩阵经初等行变换化为阶梯形后有零行出现，其充要条件是系数矩阵的秩小于方程个数 m.

(2)保留方程组中方程的个数就是系数矩阵的秩，保留方程组就是阶梯形矩阵中非零行对应的方程组.

这样，我们就有了足够的理论来讨论线性方程组何时无解(称为**不相容**)，何时有解(称为**相容**)，何时有唯一解，何时有无穷多个解的问题.

考察含有 n 个未知数 m 个方程的线性方程组

$$
\boldsymbol{Ax}=\boldsymbol{b} \tag{3.4}
$$

设 $R(\boldsymbol{A})=r$，记 $\boldsymbol{B}=(\boldsymbol{A},\boldsymbol{b})$，通常称 $\boldsymbol{B}$ 为方程组的**增广矩阵**，对 $\boldsymbol{B}$ 施行初等行变换化为阶梯形矩阵后，再将前 n 列作适当调整，记为 $\boldsymbol{B}_1=(\boldsymbol{A}_1,\boldsymbol{b}_1)$，则有 $R(\boldsymbol{B})=R(\boldsymbol{B}_1)$，$R(\boldsymbol{A})=R(\boldsymbol{A}_1)$，并且很明显地有 $R(\boldsymbol{B})\geqslant R(\boldsymbol{A})$，亦即 $R(\boldsymbol{B}_1)\geqslant R(\boldsymbol{A}_1)$.

$\boldsymbol{B}_1$ 只可能有下列两种情形：

$$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1r} & \cdots & c_{1n} & d_1 \\ 0 & c_{22} & \cdots & c_{2r} & \cdots & c_{2n} & d_2 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & c_{rr} & \cdots & c_{rn} & d_r \\ 0 & 0 & \cdots & 0 & \cdots & 0 & d_{r+1} \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \end{pmatrix}$$

或者

$$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1r} & \cdots & c_{1n} & d_1 \\ 0 & c_{22} & \cdots & c_{2r} & \cdots & c_{2n} & d_2 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & c_{rr} & \cdots & c_{rn} & d_r \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \end{pmatrix}$$

其中 $c_{ii}\neq 0, i=1,2\cdots,r, d_{r+1}\neq 0$，很明显，在前一种情形，原方程组无解，而在后一种情形原方程组有解. 由于上述矩阵的前 n 列是由式(3.4)的系数矩阵 $\boldsymbol{A}$ 经有限次初等变换得到的，因此，上述的结果表明，当 $R(\boldsymbol{B}_1)>R(\boldsymbol{A}_1)$，即 $R(\boldsymbol{B})>R(\boldsymbol{A})$ 时，(3.4)无解，当 $R(\boldsymbol{B}_1)=R(\boldsymbol{A}_1)=r$，即 $R(\boldsymbol{B})=R(\boldsymbol{A})=r$ 时，方程组(3.4)有解.

当 $r=n$ 时，由克莱姆法则知方程组(3.4)有唯一解.

当 $r<n$ 时，为方便，设 $x_1,\cdots,x_r$ 的系数矩阵是非奇异的，则方程组(3.4)可改写成

$$\begin{cases} a_{11}x_1+\cdots+a_{1r}x_r = b_1-a_{1r+1}x_{r+1}-\cdots-a_{1n}x_n \\ a_{21}x_1+\cdots+a_{2r}x_r = b_2-a_{2r+1}x_{r+1}-\cdots-a_{2n}x_n \\ \quad\vdots \\ a_{r1}x_1+\cdots+a_{rr}x_r = b_r-a_{rr+1}x_{r+1}-\cdots-a_{rn}x_n \end{cases} \tag{3.5}$$

对 $x_{r+1},\cdots,x_n$ 任意取定的一组值，由克莱姆法则，方程组(3.5)有唯一解 $x_1,\cdots,x_r$. 把它们同取定的 $x_{r+1},\cdots,x_n$ 合在一起，就得到

方程组(3.4)的一个解，因此，方程组(3.4)的解中有 $n-r$ 个可取任意值的参数，即方程组(3.4)有无穷多个解.

反之，当方程组(3.4)无解时，必然有 $R(\boldsymbol{B})>R(\boldsymbol{A})$；当方程组(3.4)有唯一解时，必然有 $R(\boldsymbol{B})=R(\boldsymbol{A})=n$；当方程组(3.4)有无穷多个解时，必然有 $R(\boldsymbol{B})=R(\boldsymbol{A})<n$.

综上所述，便得到线性方程组的**相容性定理**.

定理 3.8　对含有 $\boldsymbol{n}$ 个未知数的线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$，其增广矩阵 $\boldsymbol{B}=(\boldsymbol{A},\boldsymbol{b})$，有下列结论：

(1) $\boldsymbol{Ax}=\boldsymbol{b}$ **有唯一解**$\Leftrightarrow R(\boldsymbol{B})=R(\boldsymbol{A})=n$；

(2) $\boldsymbol{Ax}=\boldsymbol{b}$ **有无穷多个解**$\Leftrightarrow R(\boldsymbol{B})=R(\boldsymbol{A})<n$；

(3) $\boldsymbol{Ax}=\boldsymbol{b}$ **无解**$\Leftrightarrow R(\boldsymbol{B})>R(\boldsymbol{A})$.

在方程组(3.4)中，当 $\boldsymbol{b}=\boldsymbol{0}$ 时，$\boldsymbol{Ax}=\boldsymbol{0}$ 称为**齐次线性方程组**，对此种情况，由于总有 $R(\boldsymbol{B})=R(\boldsymbol{A})$，所以方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 总是有解，当 $R(\boldsymbol{A})=n$ 时，$\boldsymbol{Ax}=\boldsymbol{0}$ 有唯一解，即只有零解(称为**平凡解**)；当 $R(\boldsymbol{A})<n$ 时，方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有无穷多个解，即有非零解(称为**非平凡解**).

例 3.5　判断下列方程组是否有非零解

$$\begin{cases} x_1-2x_2+4x_3-7x_4=0 \\ 2x_1+3x_2-x_3+5x_4=0 \\ 3x_1+x_2+2x_3-7x_4=0 \\ 4x_1-x_2-3x_3+6x_4=0 \end{cases}$$

解　对其系数矩阵 $\boldsymbol{A}$ 施行初等行变换

$$\boldsymbol{A}=\begin{pmatrix} 1 & -2 & 4 & -7 \\ 2 & 3 & -1 & 5 \\ 3 & 1 & 2 & -7 \\ 4 & -1 & -3 & 6 \end{pmatrix} \xrightarrow[r_4-4r_1]{\substack{r_2-2r_1 \\ r_3-3r_1}} \begin{pmatrix} 1 & -2 & 4 & -7 \\ 0 & 7 & -9 & 19 \\ 0 & 7 & -10 & 14 \\ 0 & 7 & -19 & 34 \end{pmatrix}$$

$$\xrightarrow[r_4-r_2]{r_3-r_2} \begin{pmatrix} 1 & -2 & 4 & -7 \\ 0 & 7 & -9 & 19 \\ 0 & 0 & -1 & -5 \\ 0 & 0 & -10 & 15 \end{pmatrix} \xrightarrow{r_4-10r_3} \begin{pmatrix} 1 & -2 & 4 & -7 \\ 0 & 7 & -9 & 19 \\ 0 & 0 & -1 & -5 \\ 0 & 0 & 0 & 65 \end{pmatrix}$$

由于 $R(\boldsymbol{A})=4$（恰为未知数个数），所以原方程组只有零解.

例 3.6 判定下列方程组的相容性

$$\begin{cases} x_1-2x_2+3x_3-x_4=1 \\ 3x_1-x_2+5x_3-3x_4=2 \\ 2x_1+x_2+2x_3-2x_4=3 \end{cases}$$

解 对增广矩阵 $\boldsymbol{B}$ 施行初等行变换

$$\boldsymbol{B}=\begin{pmatrix} 1 & -2 & 3 & -1 & 1 \\ 3 & -1 & 5 & -3 & 2 \\ 2 & 1 & 2 & -2 & 3 \end{pmatrix} \xrightarrow[r_3-2r_1]{r_2-3r_1} \begin{pmatrix} 1 & -2 & 3 & -1 & 1 \\ 0 & 5 & -4 & 0 & -1 \\ 0 & 5 & -4 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{r_3-r_2} \begin{pmatrix} 1 & -2 & 3 & -1 & 1 \\ 0 & 5 & -4 & 0 & -1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

可见 $R(\boldsymbol{A})=2<R(\boldsymbol{B})=3$，故原方程组无解.

例 3.7 判定下列方程组是否有解

$$\begin{cases} x_1+x_2-3x_3-x_4=1 \\ 3x_1-x_2-3x_3+4x_4=4 \\ x_1+5x_2-9x_3-8x_4=0 \end{cases}$$

解

$$\boldsymbol{B}=\begin{pmatrix} 1 & 1 & -3 & -1 & 1 \\ 3 & -1 & -3 & 4 & 4 \\ 1 & 5 & -9 & -8 & 0 \end{pmatrix} \xrightarrow[r_3-r_1]{r_2-3r_1}$$

$$\begin{pmatrix} 1 & 1 & -3 & -1 & 1 \\ 0 & -4 & 6 & 7 & 1 \\ 0 & 4 & -6 & -7 & -1 \end{pmatrix} \xrightarrow{r_3+r_2} \begin{pmatrix} 1 & 1 & -3 & -1 & 1 \\ 0 & -4 & 6 & 7 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

可见 $R(\boldsymbol{A})=R(\boldsymbol{B})=2<4$（未知数个数），故方程组有无穷多个解.

例 3.8 问 k 取何值时，线性方程组

$$\begin{cases} x_1+x_2+kx_3=1 \\ x_1+kx_2+x_3=1 \\ kx_1+x_2+x_3=1 \end{cases}$$

(1)有唯一解；(2)有无穷多个解；(3)无解.

解 对增广矩阵 $\boldsymbol{B}$ 施行初等行变换

$$\boldsymbol{B}=\begin{pmatrix}1&1&k&1\\1&k&1&1\\k&1&1&1\end{pmatrix}\xrightarrow[r_3-kr_1]{r_2-r_1}\begin{pmatrix}1&1&k&1\\0&k-1&1-k&0\\0&1-k&1-k^2&1-k\end{pmatrix}$$

$$\xrightarrow{r_3+r_2}\begin{pmatrix}1&1&k&1\\0&k-1&1-k&0\\0&0&2-k-k^2&1-k\end{pmatrix}$$

(1)方程组有唯一解$\Leftrightarrow R(\boldsymbol{B})=R(\boldsymbol{A})=3\Leftrightarrow 2-k-k^2\neq 0$

$\Leftrightarrow k\neq -2$ 且 $k\neq 1$；

(2)方程组有无穷多个解$\Leftrightarrow R(\boldsymbol{B})=R(\boldsymbol{A})<3\Leftrightarrow\begin{cases}2-k-k^2=0\\1-k=0\end{cases}$

$\Leftrightarrow k=1$；

(3)方程组无解$\Leftrightarrow R(\boldsymbol{B})>R(\boldsymbol{A})\Leftrightarrow\begin{cases}2-k-k^2=0\\1-k\neq 0\end{cases}\Leftrightarrow k=-2.$

至于方程组有无穷多解时如何求它的通解，将在下章讨论.

习 题 三

1. 根据秩的定义求矩阵 $\boldsymbol{A}$ 的秩

$$\boldsymbol{A}=\begin{pmatrix}3&2&1\\3&1&5\\3&2&2\end{pmatrix}$$

2. 秩为 r 的矩阵中是否一定没有 r 阶奇异子方阵？

3. 矩阵 $\boldsymbol{A}$ 去掉一行得矩阵 $\boldsymbol{B}$，问 $R(\boldsymbol{B})$与 $R(\boldsymbol{A})$的关系.

4. 求作一秩为 4 的方阵，使其两行是

$$(1,0,1,0,0),(1,-1,0,0,0).$$

5. 对矩阵

$$\boldsymbol{A}=\begin{pmatrix}1&-1&2&-1\\3&1&0&2\\1&3&-4&4\end{pmatrix}$$

求可逆矩阵 $\boldsymbol{P}$，使

$$PA=\begin{pmatrix}1&-1&2&-1\\0&4&-6&5\\0&0&0&0\end{pmatrix}=B$$

再求可逆矩阵 Q,使

$$BQ=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\end{pmatrix}=N$$

从而得到 $PAQ=N$.

6. 设 A 与 B 为 $m\times n$ 维矩阵,试证 $R(B)=R(A)$ 时,$A\sim B$.

7△. 证明:$R(AB)\leqslant\min(R(A),R(B))$.

8. 证明:$R(A)=R(A_1)+R(A_2)$,其中 $A=\begin{pmatrix}A_1&O\\O&A_2\end{pmatrix}$.

9. 证明:$R(A+B)\leqslant R(A)+R(B)$

10. 试证,若 X 为幂等矩阵(即 $X^2=X$),则 $R(X)=\mathrm{tr}(X)$(提示:利用习题一,22(2)的结果).

11. 求下列矩阵的秩

(1)$A=\begin{pmatrix}1&2&-1&-3&-2\\2&-1&3&1&-3\\7&0&5&-1&8\end{pmatrix}$

(2)$B=\begin{pmatrix}25&31&17&4\\75&94&53&13\\75&94&54&12\\25&32&20&4\end{pmatrix}$

(3)$C=\begin{pmatrix}1&1&2&2&1\\0&2&1&5&-1\\2&0&3&-1&3\\1&1&0&4&-1\end{pmatrix}$

(4) $\boldsymbol{D}=\begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{pmatrix}$

$(a_ib_i\neq 0, i=1,2,\cdots,n)$

12. 对上题中的 $\boldsymbol{C}$,求可逆阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$,使 $\boldsymbol{PCQ}$ 为 $\boldsymbol{C}$ 的标准形.

13. 判别下列方程组的相容性(或有无非零解):

(1) $\begin{cases} x_1+x_2+2x_3-x_4=0 \\ 2x_1+x_2+x_3-x_4=0 \\ 2x_1+2x_2+x_3+2x_4=0 \end{cases}$　　(2) $\begin{cases} x_1+2x_2+x_3=0 \\ 3x_1+5x_2-x_3=0 \\ 5x_1+10x_2+x_3=0 \end{cases}$

(3) $\begin{cases} 4x_1+2x_2-x_3=2 \\ 3x_1-x_2+2x_3=10 \\ 11x_1+3x_2=8 \end{cases}$　　(4) $\begin{cases} 2x_1+x_2-x_3+x_4=1 \\ 4x_1+2x_2-2x_3+x_4=2 \\ 2x_1+x_2-x_3-x_4=1 \end{cases}$

(5) $\begin{cases} x_1-2x_2+4x_3=-5 \\ 2x_1+3x_2+x_3=4 \\ 3x_1+8x_2-2x_3=13 \\ 4x_1-x_2+9x_3=-6 \end{cases}$　　(6) $\begin{cases} x_1+4x_2-3x_3+5x_4=-2 \\ 2x_1+x_2-x_3+x_4=1 \\ 3x_1-2x_2+x_3-3x_4=4 \end{cases}$

14. λ 取何值时,方程组

$$\begin{cases} \lambda x_1+x_2+x_3=1 \\ x_1+\lambda x_2+x_3=\lambda \\ x_1+x_2+\lambda x_3=\lambda^2 \end{cases}$$

(1)有唯一解;(2)有无穷多个解;(3)无解.

15. λ 取何值时,下列方程组有解

$$\begin{cases} -2x_1+x_2+x_3=-2 \\ x_1-2x_2+x_3=\lambda \\ x_1+x_2-2x_3=\lambda^2 \end{cases}$$

16$^{\triangle}$. 4 阶方阵 $\boldsymbol{A}$ 的秩为 2,则 $\boldsymbol{A}$ 的伴随矩阵的秩=________.

第 4 章　向量组的线性相关性和线性方程组解的结构

向量可看成只有一行(列)的矩阵,矩阵可看成由一组同维向量组成.因此,可以用矩阵理论来研究向量,并把上章讨论过的关于线性方程组与矩阵的对应,看成方程组与一个向量组的对应.相应于线性方程组理论中讨论过的三个问题:(1)方程组是否有多余方程;(2)保留方程组中方程的个数;(3)保留方程组与同解方程组,本章引入这样一些概念:(1)向量组的线性相关性;(2)向量组的秩;(3)向量组的最大无关组和向量组的等价.并用这些概念得到线性方程组解的结构,最后解决求线性方程组的通解问题.

4.1　向量组的线性相关性

4.1.1　线性相关和线性无关

在第 3 章 3.1 中曾指出,方程组

$$\begin{cases} x+2y-z=0 \\ 2x-3y+z=0 \\ 4x+y-z=0 \end{cases} \text{或}\quad x\begin{pmatrix}1\\2\\4\end{pmatrix}+y\begin{pmatrix}2\\-3\\1\end{pmatrix}+z\begin{pmatrix}-1\\1\\-1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix} \tag{4.1}$$

的第 3 个方程是多余方程,现在从向量的角度来讨论这个问题.

方程组(4.1)对应着三个向量

$$\boldsymbol{\alpha}_1=(1,2,4)^{\mathrm T},\quad \boldsymbol{\alpha}_2=(2,-3,1)^{\mathrm T},\quad \boldsymbol{\alpha}_3=(-1,1,-1)^{\mathrm T}$$

上面指出的 3 个方程有多余方程反映到它们对应的向量上就是 $\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+7\boldsymbol{\alpha}_3=\mathbf{0}$ 或 $\boldsymbol{\alpha}_3=-\dfrac{1}{7}\boldsymbol{\alpha}_1-\dfrac{3}{7}\boldsymbol{\alpha}_2$,即 $\boldsymbol{\alpha}_3$ 可由 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$ 线性运算得到,此时称 $\boldsymbol{\alpha}_3$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 的线性组合或称 $\boldsymbol{\alpha}_3$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 线性表示.

定义 4.1 对于向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\alpha}$，如果有数 $k_1,k_2,\cdots,k_m$，使 $\boldsymbol{\alpha}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m$，则称向量 $\boldsymbol{\alpha}$ 是向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的**线性组合**，或者 $\boldsymbol{\alpha}$ 可由向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ **线性表示**.

由此定义，有 n 个未知数的齐次线性方程组是否有多余方程就相当于相应的 n 个向量中是否有向量可由其余向量线性表示的问题. 这是向量组的一个重要性质，称为向量组的**线性相关性**. 为了使用方便，用另一种说法给出定义.

定义 4.2 设有 n 维向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$，如果存在不全为零的数 $k_1,k_2,\cdots,k_m$，使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\boldsymbol{0}$$

则称向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ **线性相关**，否则称为**线性无关**.

由此定义，(4.1)式对应的三个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，因为 $\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+7\boldsymbol{\alpha}_3=\boldsymbol{0}$. 而 $\boldsymbol{\beta}_1=(1,2,-1)$，$\boldsymbol{\beta}_2=(3,2,2)$，$\boldsymbol{\beta}_3=(3,3,3)$ 线性无关. 事实上，要考察 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 是否线性相关，就是看有没有不全为零的数 x_1,x_2,x_3 使 $x_1\boldsymbol{\beta}_1+x_2\boldsymbol{\beta}_2+x_3\boldsymbol{\beta}_3=\boldsymbol{0}$，亦即看方程组

$$\begin{cases}x_1+3x_2+3x_3=0\\2x_1+2x_2+3x_3=0\\-x_1+2x_2+3x_3=0\end{cases}$$

有无非零解，因其系数矩阵 $\boldsymbol{A}$ 非奇异，即 $R(\boldsymbol{A})=3$（它恰是向量的个数），因此方程组只有零解，亦即当且仅当 $x_1=x_2=x_3=0$ 时才有 $x_1\boldsymbol{\beta}_1+x_2\boldsymbol{\beta}_2+x_3\boldsymbol{\beta}_3=\boldsymbol{0}$，这表明 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性无关.

定理 4.1 **n 维向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关的充分必要条件是 $R(\boldsymbol{A})<m$，其中 $\boldsymbol{A}$ 是由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 组成的 $m\times n$（或 $n\times m$）维矩阵.**

证 设 $\boldsymbol{\alpha}_i$，$i=1,2,\cdots,m$，为 n 维列向量，下面证有 m 个不全为零的数 $x_1,\cdots,x_m$ 使 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_m\boldsymbol{\alpha}_m=\boldsymbol{0}$ 的充分且必要的条件是 $R(\boldsymbol{A})<m$.

$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_m\boldsymbol{\alpha}_m=\boldsymbol{0}$，即 $\boldsymbol{Ax}=\boldsymbol{0}$，其中 $\boldsymbol{x}=(x_1,x_2,\cdots,x_m)^{\mathrm{T}}$，由线性方程组的相容性定理知 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解的充分必要条件是 $R(\boldsymbol{A})<m$，这就证明了定理 4.1.

由此定理便得到判定一个 n 维向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关性的一般方法：以每一向量为列(或行)组成矩阵 $\boldsymbol{A}$ 并求矩阵 $\boldsymbol{A}$ 的秩 $R(\boldsymbol{A})$，若 $R(\boldsymbol{A})<m$，则向量组线性相关，否则线性无关.

推论　对 m 个 n 维向量，若 $m>n$，则该向量组线性相关.

证　由这些向量组成的矩阵记为 $\boldsymbol{A}$，则 $\boldsymbol{A}$ 是 $m\times n$(或 $n\times m$)维的，由于 $m>n$，所以

$$R(\boldsymbol{A})\leqslant \min(m,n)=n<m$$

由定理 4.1，该向量组线性相关.

该推论表明，当向量组中向量的个数大于向量的维数时，向量组一定线性相关，如，$\boldsymbol{\alpha}_1=(1,1,2)^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=(0,1,0)^{\mathrm{T}}$，$\boldsymbol{\alpha}_3=(1,0,1)^{\mathrm{T}}$，$\boldsymbol{\alpha}_4=(0,0,1)^{\mathrm{T}}$ 一定线性相关.

例 4.1　讨论向量组

$$\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(0,2,5)^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=(1,3,6)^{\mathrm{T}}$$

的线性相关性.

解　以所给向量为列组成矩阵 $\boldsymbol{A}$

$$\boldsymbol{A}=\begin{pmatrix}1&0&1\\1&2&3\\1&5&6\end{pmatrix}\xrightarrow[r_3-r_1]{r_2-r_1}\begin{pmatrix}1&0&1\\0&2&2\\0&5&5\end{pmatrix}\xrightarrow{r_3-\frac{5}{2}r_2}\begin{pmatrix}1&0&1\\0&2&2\\0&0&0\end{pmatrix}$$

因 $R(\boldsymbol{A})=2<3$(向量个数)，所以所给向量组线性相关.

例 4.2　讨论 n 维单位坐标向量 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 的线性相关性.

解　因为 $R(\boldsymbol{A})=R(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)=n$，所以向量组线性无关.

例 4.3　设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，试证 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$，$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$，$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1$ 线性无关.

证　无妨设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 均为列向量，则

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\begin{pmatrix}1&0&1\\1&1&0\\0&1&1\end{pmatrix}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\boldsymbol{C}$$

因为矩阵 $\boldsymbol{C}$ 可逆，所以 $\boldsymbol{C}$ 可表示为有限个初等矩阵的乘积，即矩阵$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$可认为由矩阵$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$经有限次初等列变换得到，从而矩阵$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$的秩等于矩阵$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$的秩，而$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$

的秩为3，所以$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$的秩为3，因此$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$线性无关.

例 4.4 设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性无关，若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}$线性相关，则$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性表示.

证 因$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}$线性相关，故有不全为零的数$k_1,k_2,\cdots,k_r,k$使$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r+k\boldsymbol{\beta}=\mathbf{0}$. 要证$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性表示，只需证$k\neq0$. 用反证法，设$k=0$，则$k_1,k_2,\cdots,k_r$不全为零且能使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r=\mathbf{0}$$

这与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性无关矛盾. 矛盾表明$k\neq0$，即$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$表示为$\boldsymbol{\beta}=\dfrac{-1}{k}(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r)$.

下面介绍向量组线性相关性的几个重要结论，用它们也可以判别向量组的线性相关性.

4.1.2 向量组线性相关性的几个重要结论

定理 4.2 若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性相关，则$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_m$也线性相关.

证 因$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性相关，所以存在不全为零的数$k_1,k_2,\cdots,k_r$使$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r=\mathbf{0}$，从而存在不全为零的$m$个数$k_1,k_2,\cdots,k_r,0,\cdots,0$，使

$$k_1\boldsymbol{\alpha}_1+\cdots+k_r\boldsymbol{\alpha}_r+0\boldsymbol{\alpha}_{r+1}+\cdots+0\boldsymbol{\alpha}_m=\mathbf{0}$$

因此，$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_m$线性相关.

由于一个零向量是线性相关的，所以任何含有零向量的向量组都线性相关.

推论 若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性无关，则由其中的部分向量构成的向量组线性无关.

定理4.3 设$\boldsymbol{\alpha}_i=(a_{i1},a_{i2},\cdots,a_{ir})$

$$\boldsymbol{\beta}_i=(a_{i1},a_{i2},\cdots,a_{ir},a_{ir+1})\quad(i=1,2,\cdots,m)$$

若r维向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性无关，则$r+1$维向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$线性无关.

证 显然 $\boldsymbol{\beta}_i=(\boldsymbol{\alpha}_i,\alpha_{ir+1})$，$i=1,2,\cdots,m$，设有 m 个数 $k_1,k_2,\cdots,k_m$ 使 $k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_m\boldsymbol{\beta}_m=\mathbf{0}$，即

$$(k_1\boldsymbol{\alpha}_1+\cdots+k_m\boldsymbol{\alpha}_m,k_1a_{1r+1}+\cdots+k_ma_{mr+1})=\mathbf{0}$$

因此有 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0}$. 由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关，所以当且仅当 $k_1=k_2=\cdots=k_m=0$ 时才成立，这就表明 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 线性无关.

推论 r **维向量组的每个向量加上** $n-r$ **个分量成为** n **维向量. 若** r **维向量组线性无关，则** n **维向量组线性无关.**

例 4.5 讨论向量组

$$\boldsymbol{\alpha}_1=(1,0,0,2,3),\quad \boldsymbol{\alpha}_2=(0,1,0,4,6),\quad \boldsymbol{\alpha}_3=(0,0,1,2,2)$$

的线性相关性.

解 因为向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 分别是由 $\boldsymbol{e}_1=(1,0,0)$，$\boldsymbol{e}_2=(0,1,0)$，$\boldsymbol{e}_3=(0,0,1)$加上两个分量得到的，而 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 线性无关，所以根据上面的结论知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关.

4.2 向量组的秩

由于一个向量可以看成只有一行或一列的矩阵，以向量组的每一向量为列(或行)可以组成矩阵，因此，可以用矩阵来研究向量组，用矩阵的秩来定义向量组的秩.

定义 4.3 向量组 A：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的秩定义为用它的每一向量为列(或行)所组成的矩阵 $\boldsymbol{A}$ 的秩，记为 $R(A)$.

有了这个概念，可以像求矩阵的秩一样求向量组的秩. 若以向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 为行组成矩阵 $\boldsymbol{A}$，则 $R(\boldsymbol{A})$称为矩阵 $\boldsymbol{A}$ 的**行秩**；若以向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 为列组成矩阵 $\boldsymbol{A}$，则 $R(\boldsymbol{A})$称为矩阵 $\boldsymbol{A}$ 的**列秩**. 由于矩阵的转置不改变矩阵的秩，所以矩阵的行秩等于列秩.

例 4.6 判定下列向量组的线性相关性，并求它的秩.

$$\boldsymbol{\alpha}_1=(1,2,3,-1),\quad \boldsymbol{\alpha}_2=(1,4,4,3)$$
$$\boldsymbol{\alpha}_3=(2,4,6,4),\quad \boldsymbol{\alpha}_4=(-1,-2,-3,4)$$

解 设 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3+x_4\boldsymbol{\alpha}_4=\mathbf{0}$，该方程组的系数矩阵 $\boldsymbol{A}$

是以向量为列组成的矩阵，对其进行初等行变换

$$\boldsymbol{A}=\begin{pmatrix}1&1&2&-1\\2&4&4&-2\\3&4&6&-3\\-1&3&4&4\end{pmatrix}\xrightarrow[r_4+r_1]{\substack{r_2-2r_1\\r_3-3r_1}}\begin{pmatrix}1&1&2&-1\\0&2&0&0\\0&1&0&0\\0&4&6&3\end{pmatrix}$$

$$\xrightarrow[r_4-2r_2]{r_3-\frac{1}{2}r_2}\begin{pmatrix}1&1&2&-1\\0&2&0&0\\0&0&0&0\\0&0&6&3\end{pmatrix}\xrightarrow{r_3\leftrightarrow r_4}\begin{pmatrix}1&1&2&-1\\0&2&0&0\\0&0&6&3\\0&0&0&0\end{pmatrix}$$

由于 $R(\boldsymbol{A})=3<4$（向量个数），所以向量组线性相关，且其秩为 3；又由于以向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 组成的矩阵（即 $\boldsymbol{A}$ 的前 3 列）的秩为 3（向量个数），因此向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，从而 $\boldsymbol{\alpha}_4$ 可由向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示：$\boldsymbol{\alpha}_4=-2\boldsymbol{\alpha}_1+0\boldsymbol{\alpha}_2+\frac{1}{2}\boldsymbol{\alpha}_3$.

前面已经指出，方程组中的每个未知数对应着一个向量，这样，第 3 章讨论过的方程组(3.3)有多余方程，用向量的术语来说就是向量组的秩小于方程的个数，方程组的保留方程组中有 r 个方程，用向量的术语来说则是对应的向量组的秩为 r，也就是说，向量组中有 r 个向量线性无关，且其余的向量可由这 r 个向量线性表示. 为了说明向量组的这种性质，引进如下定义.

定义 4.4 设 V 是 n 维向量组成的向量组. 若 V 中的 r 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 满足：(1) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关；(2) 任取 $\boldsymbol{\alpha}\in V$，总有 $\boldsymbol{\alpha},\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性相关. 那么称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 为向量组 V 的一个**最大线性无关向量组**，简称为**最大无关组**.

在例 4.6 中，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的一个最大无关组，方程组(4.1)对应的向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 中，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是它的一个最大无关组，同样 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 或 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3$ 也都是它的最大无关组. 由此可见，一个向量组的最大无关组一般不唯一. 但最大无关组中向量的个数唯一，而且它就是向量组的秩.

定理 4.4 **设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_m$**

的一个最大无关组，则向量组的秩为 r.

证 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 均为行向量，记矩阵

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_r\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}\boldsymbol{\alpha}_{r+1}\\ \boldsymbol{\alpha}_{r+2}\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix},\quad \boldsymbol{C}=\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}$$

由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是向量组的最大无关组知 $\boldsymbol{B}$ 中每一向量均可由 $\boldsymbol{A}$ 中的向量线性表示，且 $R(\boldsymbol{A})=r$，故可通过适当初等行变换使 $\boldsymbol{C}\sim\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{pmatrix}$，$R(\boldsymbol{C})=R\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{pmatrix}=R(\boldsymbol{A})=r$，即向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的秩为 r.

至此，如果以向量组的每一个向量为列组成矩阵 $\boldsymbol{A}$ ，并只用初等行变换将 $\boldsymbol{A}$ 化为阶梯形矩阵，因为阶梯形矩阵的非零行数就是矩阵 $\boldsymbol{A}$ 的秩，所以从中可以得到：(1)矩阵 $\boldsymbol{A}$ 的秩就是向量组的秩；(2)当秩小于向量个数时，向量组线性相关，当秩等于向量个数时，向量组线性无关；(3)阶梯形矩阵中所有非零行的第一个非零元素所在的列对应的列向量就是向量组的最大无关组.

例 4.7 全体 n 维向量构成的向量组记为 R^n，求 R^n 的一个最大无关组.

解 例 4.2 已证明 n 维单位坐标向量 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 线性无关，今对任一向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in R^n$，都有 $\boldsymbol{x}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n$，因此，$n$ 维单位坐标向量是 R^n 的一个最大无关组.

例 4.8 判断下列向量组的线性相关性并求它的秩和一个最大无关组：

$$\boldsymbol{\alpha}_1=(1,2,3),\quad \boldsymbol{\alpha}_2=(2,4,6),\quad \boldsymbol{\alpha}_3=(0,1,2)$$

解 以向量为列组成矩阵，并对其施行初等行变换

$$\begin{pmatrix}1&2&0\\2&4&1\\3&6&2\end{pmatrix}\xrightarrow[r_3-3r_1]{r_2-2r_1}\begin{pmatrix}1&2&0\\0&0&1\\0&0&2\end{pmatrix}\xrightarrow{r_3-2r_2}\begin{pmatrix}1&2&0\\0&0&1\\0&0&0\end{pmatrix}$$

因此，向量组线性相关，秩为 2，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3$ 是它的一个最大无关组.

最后，介绍与两个方程组同解相对应的向量组的等价概念.

定义 4.5 设有 n 维向量组 $A:\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r;B:\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$. 如果向量组 B 中的每一向量可由向量组 A 中的向量线性表示，则称**向量组 B 可由向量组 A 线性表示**，若向量组 B 可由向量组 A 线性表示，且向量组 A 也可由向量组 B 线性表示，则称**向量组 A 与向量组 B 等价**.

显然，向量组的等价具有反身性、对称性和传递性.

容易证明，一个向量组与它的最大无关组是等价的.

需要注意，向量组的等价与矩阵的等价是有区别的. 向量组等价不要求两向量组中向量个数相同，从而不能断定两向量组各自组成的矩阵等价，但当两向量组中向量个数相等时，两向量组等价能推出它们各自组成的矩阵等价；矩阵等价只是指两同维矩阵的秩相等，因此，从两个矩阵的等价不能推出它们对应的向量组等价. 但它们又是有联系的，这种联系反映在下面的定理中.

定理 4.5 **设有 n 维向量组 A，B，合起来的向量组记为 C(相同的向量不合并，下同)，那么向量组 B 可由向量组 A 线性表示的充要条件是** $R(A)=R(C)\geqslant R(B)$.

证 设以向量组 A 的向量为行组成的矩阵仍记为 $\boldsymbol{A}$，对向量组 B 类似地组成矩阵 $\boldsymbol{B}$，则向量组 C 组成的矩阵 $\boldsymbol{C}=\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}$.

必要性 由于向量组 B 可由向量组 A 线性表示，所以矩阵 $\boldsymbol{C}$ 中 $\boldsymbol{B}$ 块的每一行均可由 $\boldsymbol{A}$ 块的行线性表示，因此，可以通过适当的初等行变换可将矩阵 $\boldsymbol{C}$ 变为与之等价的矩阵 $\boldsymbol{C}_1=\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{pmatrix}$，故有 $R(C)=R(C_1)=R(A)$.

充分性 用反证法. 假设向量组 B 中至少有一向量不能用向量组 A 线性表示，即有向量不能用向量组 A 的最大无关组线性表示，根据例 4.4 的结果，则这个向量同向量组 A 的最大无关组组成的向量组线性无关，从而 $R(C)>R(A)$，这与 $R(C)=R(A)$ 矛盾，矛盾表明向量组 B 可由向量组 A 线性表示.

推论 1 **设有 n 维向量组 A、B，若向量组 B 可由向量组 A 线**

性表示,则 $R(A)\geqslant R(B)$.

推论 2 设有 n 维向量组 A、B,合起来的向量组记为 C,则向量组 A 与向量组 B 等价的充要条件是 $R(A)=R(B)=R(C)$.

例 4.9 讨论下列向量组 A 与向量组 B 的等价性

A:$\boldsymbol{\alpha}_1=(1,2,1)$, $\boldsymbol{\alpha}_2=(2,5,3)$, $\boldsymbol{\alpha}_3=(4,9,5)$

B:$\boldsymbol{\beta}_1=(2,4,2)$, $\boldsymbol{\beta}_2=(3,7,4)$

解 容易求得 $R(A)=R(B)=2$,又

$$\boldsymbol{C}=\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}=\begin{pmatrix}1&2&1\\2&5&3\\4&9&5\\2&4&2\\3&7&4\end{pmatrix}\xrightarrow[\substack{\boldsymbol{r}_4-2\boldsymbol{r}_1\\ \boldsymbol{r}_5-3\boldsymbol{r}_1}]{\substack{\boldsymbol{r}_2-2\boldsymbol{r}_1\\ \boldsymbol{r}_3-4\boldsymbol{r}_1}}\begin{pmatrix}1&2&1\\0&1&1\\0&1&1\\0&0&0\\0&1&1\end{pmatrix}\xrightarrow[\boldsymbol{r}_5-\boldsymbol{r}_2]{\boldsymbol{r}_3-\boldsymbol{r}_2}\begin{pmatrix}1&2&1\\0&1&1\\0&0&0\\0&0&0\\0&0&0\end{pmatrix}$$

因此 $R(C)=R(A)=R(B)$,这表明向量组 A 与 B 等价.

需注意,若向量组 A 与向量组 B 仅满足条件 $R(A)=R(B)$,则不能推出 A 与 B 等价,如 A:$\boldsymbol{\alpha}_1=(1,0,0,0)$, $\boldsymbol{\alpha}_2=(0,1,0,0)$ 与 B:$\boldsymbol{\beta}_1=(0,0,1,0)$, $\boldsymbol{\beta}_2=(0,0,0,1)$就是这种情形.

4.3 向量空间

由解析几何知道,平面上任何两个向量的和仍是该平面上的向量,任一向量同任一实数的乘积也是平面上的向量,对三维空间的向量,也有同样的结论.由前两节的讨论可以看到,n 维向量之间的基本关系是用向量的加法和数乘运算来表达的,因此,加法和数乘运算这两种线性运算对向量是至关重要的.对由其它元素构成的集合也是如此.事实上,在许多数学和物理系统中具有一种称为线性的性质,其实质就是系统中的任两元素之和及任一实数与系统的任一元素的乘积都仍为该系统的元素,即该系统对这两种线性运算是封闭的.数学上把这种具有对规定的加法和数乘运算封闭并满足相应的运算规律的元素的集合称为**向量空间**或**线性空间**.这里只介绍当集合的元素为向量时,有关向量空间的一些基本

概念.

定义 4.6 设 V 是 n 维向量组成的非空集合，如果 V 对向量加法和数乘运算封闭，即对任何向量 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$，任何 $k\in\mathbf{R}$(实数集)，有 $\boldsymbol{\alpha}+\boldsymbol{\beta}\in V,k\boldsymbol{\alpha}\in V$，则称 V 为**向量空间**.

例 4.10 n 维向量的全体 R^n 是一向量空间，因为任意两个 n 维向量的和仍是 n 维向量，数 k 乘 n 维向量仍为 n 维向量，即 R^n 对加法和数乘运算封闭.

例 4.11 集合 $V_1=\{\boldsymbol{x}=(0,x_2,\cdots,x_n)\mid x_i\in\mathbf{R},i=2,3,\cdots,n\}$ 是向量空间，因对任何 $\boldsymbol{\alpha}=(0,a_2,\cdots,a_n)\in V_1$，$\boldsymbol{\beta}=(0,b_2,\cdots,b_n)\in V_1$，有 $\boldsymbol{\alpha}+\boldsymbol{\beta}=(0,a_2+b_2,\cdots,a_n+b_n)\in V_1$，$k\boldsymbol{\alpha}=(0,ka_2,\cdots,ka_n)\in V_1$.

例 4.12 集合 $V_2=\{\boldsymbol{x}=(1,x_2,\cdots,x_n)\mid x_i\in\mathbf{R},i=2,3,\cdots,n\}$ 不是向量空间，因 $2\boldsymbol{x}=(2,2x_2,\cdots,2x_n)\notin V_2$.

例 4.13 设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 是两个已知的 n 维向量，则集合

$$V_3=\{\boldsymbol{x}=\lambda\boldsymbol{\alpha}+\mu\boldsymbol{\beta}\mid\lambda,\mu\in\mathbf{R}\}$$

是向量空间，因若 $\boldsymbol{x}_1=\lambda_1\boldsymbol{\alpha}+\mu_1\boldsymbol{\beta}\in V_3$，$\boldsymbol{x}_2=\lambda_2\boldsymbol{\alpha}+\mu_2\boldsymbol{\beta}\in V_3$，则有

$$\boldsymbol{x}_1+\boldsymbol{x}_2=(\lambda_1+\lambda_2)\boldsymbol{\alpha}+(\mu_1+\mu_2)\boldsymbol{\beta}\in V_3,$$

$$k\boldsymbol{x}_1=(k\lambda_1)\boldsymbol{\alpha}+(k\mu_1)\boldsymbol{\beta}\in V_3.$$

很明显，V_3 是由 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的所有可能的线性组合构成的，把它称为**由 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 所生成的向量空间**，一般地，由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 所生成向量空间为

$$V=\{\boldsymbol{x}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r\mid k_i\in\mathbf{R},i=1,2,\cdots,r\}$$

例 4.14 试证由 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 生成的向量空间是 R^n，即

$$R^n=\{\boldsymbol{x}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n\mid x_i\in\mathbf{R},i=1,2,\cdots,n\}.$$

证 设 $V=\{\boldsymbol{x}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n\mid x_i\in\mathbf{R},i=1,2,\cdots,n\}$，任取 $\boldsymbol{x}\in V$，则 $\boldsymbol{x}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in R^n$，这表明 $V\subset R^n$. 任取 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in R^n$，很明显，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n\in V$，这表明 $R^n\subset V$，于是 $R^n=V$，即

$$R^n=\{\boldsymbol{x}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n\mid x_i\in\mathbf{R},i=1,2,\cdots,n\}.$$

由本例可知，向量空间 R^n 是线性无关的 n 维单位坐标向量 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 生成的. 其实，任一向量空间都是由一组线性无关的向量生成，这组向量称为这个向量空间的一组基. 一般地有如下定义.

定义 4.7 设 V 是向量空间，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in V$ 且满足

(1) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关；

(2) 任取 $\boldsymbol{\alpha}\in V$，总有 $\boldsymbol{\alpha},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性相关.

那么向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 称为向量空间 V 的一组**基**，r 称为向量空间 V 的**维数**，并称 V 为 r **维**向量空间.

若把向量空间看成向量组，则向量空间的基就是向量组的最大无关组，向量空间的维数就是向量组的秩. 因此，一个向量空间的维数是确定的. 顺便指出，这里仅讨论有限维的向量空间.

应该注意，向量空间的维数与向量的维数是两个不同的概念，前者指基中向量的个数，后者指向量中分量的个数，如例 4.11 中向量空间 V_1 的维数是 $n-1$，而它所含的向量的维数是 n.

例 4.7 证明了 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是 R^n 的一组基，例 4.14 证明了 R^n 是一个向量空间，因此，R^n 是 n 维向量空间.

例 4.11 证明了 V_1 是向量空间，例 4.14 证明了 R^n 也是向量空间，很明显，$V_1\subset R^n$，就是说，向量空间 V_1 是向量空间 R^n 的部分空间，针对这种情况，给出如下定义.

定义 4.8 设有向量空间 V_1,V_2，若 $V_1\subset V_2$，则称 V_1 为 V_2 的**子空间**.

很明显，V_1 是 R^n 的子空间，其实任何由 n 维向量组生成的向量空间都是 R^n 的子空间，但例 4.12 中的 V_2 不是 R^n 的子空间，因为 V_2 不是向量空间，尽管 $V_2\subset R^n$.

若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是向量空间 V 的一组基，则对任何 $\boldsymbol{\alpha}\in V$ 均可由这组基线性表示，且表示式唯一（见习题四，3）. 设表示式为 $\boldsymbol{\alpha}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r$，则 $k_1,k_2,\cdots,k_r$ 这组有序数称为 $\boldsymbol{\alpha}$ 在这组基下的**坐标**. 若 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 也是 V 的一组基，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 等价，从而存在 r 阶可逆矩阵 $\boldsymbol{C}$ 使

$$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r)=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)\boldsymbol{C}$$

此式称为**基变换公式**，矩阵 $\boldsymbol{C}$ 称为由基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 到基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots\boldsymbol{\alpha}_r$ 的**过渡矩阵**，由此可得 $\boldsymbol{\alpha}$ 在基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 下的坐标为 $(l_1,l_2,\cdots,l_r)^{\mathrm{T}}=\boldsymbol{C}(k_1,k_2,\cdots,k_r)^{\mathrm{T}}$，它称为**坐标变换公式**.

由此定义可见，要求由基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 到基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 的过渡矩阵 $\boldsymbol{C}$ 时，只需解矩阵方程 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r)\boldsymbol{C}$ 即可. 若 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 为 r 维向量，则 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r)$ 为可逆矩阵，于是过渡矩阵

$$\boldsymbol{C}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r)^{-1}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r).$$

例 4.15 已知 R^3 的两组基：

$$\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(1,0,-1)^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=(1,0,1)^{\mathrm{T}};$$

$$\boldsymbol{\beta}_1=(1,2,1)^{\mathrm{T}},\quad \boldsymbol{\beta}_2=(2,3,4)^{\mathrm{T}},\quad \boldsymbol{\beta}_3=(3,4,3)^{\mathrm{T}}$$

求由基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 到基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 的过渡矩阵 $\boldsymbol{C}$.

解 因为 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\boldsymbol{C}$，所以

$$\boldsymbol{C}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)^{-1}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$$

而
$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)^{-1}=\begin{pmatrix}0 & 1 & 0\\ 1/2 & 0 & -1/2\\ 1/2 & -1 & 1/2\end{pmatrix}$$

因此
$$\boldsymbol{C}=\begin{pmatrix}2 & 3 & 4\\ 0 & -1 & 0\\ -1 & 0 & -1\end{pmatrix}$$

例 4.16 设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 等价

$$V_1=\{\boldsymbol{x}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r \mid k_i\in\mathbf{R},i=1,2,\cdots,r\}$$

$$V_2=\{\boldsymbol{x}=t_1\boldsymbol{\beta}_1+t_2\boldsymbol{\beta}_2+\cdots+t_m\boldsymbol{\beta}_m \mid t_i\in\mathbf{R},i=1,2,\cdots,m\}$$

试证 $V_1=V_2$.

证 V_1 中的任何 $\boldsymbol{x}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示，由等价性 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 线性表示，因此 $\boldsymbol{x}$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 线性表示，即 $\boldsymbol{x}\in V_2$，这说明 $V_1\subset V_2$，类似地可证 $V_2\subset V_1$. 综合起来有 $V_1=V_2$.

由上述证明可得到一般结论，**即任何两个等价的向量组所生**

成的是同一个向量空间.

4.4　线性方程组解的结构

线性方程组的相容性定理只是给出了判别方程组是否有解的准则,第 2 章介绍的是如何用高斯消元法求解有唯一解的 n 阶线性方程组.本节把高斯消元法应用到一般线性方程组,彻底解决它的求解问题.

4.4.1　齐次线性方程组的解空间和基础解系

设有 n 元齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \quad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{cases} \tag{4.2}$$

记 $\boldsymbol{A}=(a_{ij})_{m\times n}$,$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$,则(4.2)可写成

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0} \tag{4.3}$$

如果 $x_1=\xi_{11},x_2=\xi_{21},\cdots,x_n=\xi_{n1}$ 是(4.3)式的解,则

$$\boldsymbol{x}=\boldsymbol{\xi}_1=(\xi_{11},\xi_{21},\cdots,\xi_{n1})^{\mathrm{T}}$$

称为方程组(4.2)的**解向量**,自然它也是(4.3)的解.根据(4.3)式,容易验证解向量有下列性质:

(1)若 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2$ 是(4.3)的解,则 $\boldsymbol{\xi}_1+\boldsymbol{\xi}_2$ 也是(4.3)的解;

(2)若 $\boldsymbol{\xi}$ 是(4.3)的解,k 是任意实数,则 $k\boldsymbol{\xi}$ 是(4.3)的解.

如果用 S 表示(4.2)的全体解向量集合,则上述性质恰恰说明 S 对解向量的加法和数乘运算封闭,因此 S 是向量空间,称为方程(4.2)的**解空间**.如果能求出解空间的一组基,则解空间就是由这组基生成的向量空间,从而得到方程(4.2)的全部解.下面就来讨论如何求得解空间的一组基.

定理 4.6　在方程组(4.3)**中设** $R(\boldsymbol{A})=r<n$,**则方程组**(4.2)**的解空间有基,且其中含** $n-r$ **个解向量.**

证 无妨设 $\boldsymbol{A}$ 左上角的 r 阶子方阵非奇异，则方程组(4.2)可以改写成

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1r}x_r=-a_{1r+1}x_{r+1}-\cdots-a_{1n}x_n \\ \quad\vdots \\ a_{r1}x_1+a_{r2}x_2+\cdots+a_{rr}x_r=-a_{rr+1}x_{r+1}-\cdots-a_{rn}x_n \end{cases} \tag{4.4}$$

把 $x_{r+1},x_{r+2},\cdots,x_n$ 的任意一组值代入(4.4)式，根据克莱姆法则就唯一地确定了(4.4)式的一个解 $x_1,x_2,\cdots,x_r$. 把它同取定的 $x_{r+1},x_{r+2},\cdots,x_n$ 合在一起，就唯一地确定了方程组(4.2)的一个解向量. 换言之，对于方程组(4.2)的任意两个解向量，只要它们的后 $n-r$ 个分量相同，则前 r 个分量也相同，从而两个解向量就完全一样.

在(4.4)式中分别取

$$(x_{r+1},x_{r+2},\cdots,x_n)=(1,0,\cdots,0),(0,1,\cdots,0),\cdots,(0,0,\cdots,1) \tag{4.5}$$

这样 $n-r$ 组值，就得到(4.2)的 $n-r$ 个解向量

$$\begin{aligned} \boldsymbol{\eta}_1&=(c_{11},\cdots,c_{1r},1,0,\cdots 0)^{\mathrm{T}} \\ \boldsymbol{\eta}_2&=(c_{21},\cdots,c_{2r},0,1,\cdots,0)^{\mathrm{T}} \\ &\ \vdots \\ \boldsymbol{\eta}_{n-r}&=(c_{n-r1},\cdots,c_{n-rr},0,0,\cdots,1)^{\mathrm{T}} \end{aligned} \tag{4.6}$$

现在来证明(4.6)就是(4.2)解空间的一组基.

首先，由向量组(4.5)的线性无关及定理 4.3 的推论得知向量组(4.6)线性无关.

其次，证(4.2)的任一解向量可由向量组(4.6)线性表示. 设

$$\boldsymbol{\eta}=(c_1,\cdots,c_r,c_{r+1},\cdots,c_n)$$

是(4.2)的任一解向量，由于 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-r}$ 是(4.2)的解向量，所以它们的线性组合

$$c_{r+1}\boldsymbol{\eta}_1+c_{r+2}\boldsymbol{\eta}_2+\cdots+c_n\boldsymbol{\eta}_{n-r}$$

也是(4.2)的解向量，并且它的后 $n-r$ 个分量同 $\boldsymbol{\eta}$ 的后 $n-r$ 分量相同，由前面的分析知

$$\boldsymbol{\eta}=c_{r+1}\boldsymbol{\eta}_1+c_{r+2}\boldsymbol{\eta}_2+\cdots+c_n\boldsymbol{\eta}_{n-r}$$

这样就证明了(4.6)是(4.2)解空间的一组基.

一个齐次线性方程组解空间的基又称为方程组的**基础解系**. 当 $R(\boldsymbol{A})=r=n$ 时,方程组只有零解,即解空间只有一个零向量,此时,解空间没有基,方程组没有基础解系.

在求得基础解系 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-r}$后,解空间为

$$S=\{\boldsymbol{x}=k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+\cdots+k_{n-r}\boldsymbol{\eta}_{n-r}\mid k_i\in\mathbf{R},i=1,2,\cdots,n-r\}$$

而形如 $\boldsymbol{x}=k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+\cdots+k_{n-r}\boldsymbol{\eta}_{n-r}$ 的解称为方程组(4.2)的**通解**.

例 4.17 求方程组的基础解系和通解

$$\begin{cases}x_1-x_2-x_3+x_4=0\\x_1-x_2+x_3-3x_4=0\\x_1-x_2-2x_3+3x_4=0\end{cases}$$

解

$$\boldsymbol{A}=\begin{pmatrix}1&-1&-1&1\\1&-1&1&-3\\1&-1&-2&3\end{pmatrix}\xrightarrow[r_3-r_1]{r_2-r_1}\begin{pmatrix}1&-1&-1&1\\0&0&2&-4\\0&0&-1&2\end{pmatrix}$$

$$\xrightarrow[r_3+r_2]{\frac{1}{2}r_2}\begin{pmatrix}1&-1&-1&1\\0&0&1&-2\\0&0&0&0\end{pmatrix}\xrightarrow{r_1+r_2}\begin{pmatrix}1&-1&0&-1\\0&0&1&-2\\0&0&0&0\end{pmatrix}$$

因 $R(\boldsymbol{A})=2,n=4$,故基础解系有两个解. 保留方程组为

$$\begin{cases}x_1=x_2+x_4\\x_3=2x_4\end{cases}$$

取$(x_2,x_4)=(1,0),(0,1)$,解得$(x_1,x_3)=(1,0),(1,2)$,基础解系

$$\boldsymbol{\xi}_1=(1,1,0,0)^{\mathrm{T}},\quad \boldsymbol{\xi}_2=(1,0,2,1)^{\mathrm{T}}.$$

通解为 $\boldsymbol{x}=k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$ (k_1,k_2 为任意常数)

4.4.2 非齐次线性方程组解的结构及通解

设有非齐次线性方程组

$$
\begin{cases}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
\quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m
\end{cases} \tag{4.7}
$$

若记

$$
\boldsymbol{\alpha}_j = (a_{1j}, a_{2j}, \cdots, a_{mj})^{\mathrm{T}}, j=1,2,\cdots,n, \quad \boldsymbol{b} = (b_1, b_2, \cdots, b_m)^{\mathrm{T}}
$$

$$
\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}, \boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)
$$

则方程组(4.7)可写成

$$
\boldsymbol{Ax} = \boldsymbol{b} \tag{4.8}
$$

$$
x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n = \boldsymbol{b} \tag{4.9}
$$

容易验证，非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解具有下列性质：

(1)设 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解，则 $\boldsymbol{x}=\boldsymbol{\eta}_1-\boldsymbol{\eta}_2$ 是对应齐次方程组(也称为 $\boldsymbol{Ax}=\boldsymbol{b}$ 的导出组) $\boldsymbol{Ax}=\boldsymbol{0}$ 的解；

(2)若 $\boldsymbol{\eta}^*$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的一个解，$\boldsymbol{\xi}$ 是导出组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解，则称 $\boldsymbol{x}=\boldsymbol{\eta}^*+\boldsymbol{\xi}$ 为 $\boldsymbol{Ax}=\boldsymbol{b}$ 的通解.

性质(2)可由性质(1)直接得到. 性质(2)指明了非齐次线性方程组的解的结构，即任何非齐次线性方程组的通解由它的一个解与其导出组的通解的和构成.

至此，得到关于线性方程组(4.7)的相容性有下列四种等价说法：

(1)方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解；

(2) $R(\boldsymbol{A})=R(\boldsymbol{B})$，其中 $\boldsymbol{B}=(\boldsymbol{A},\boldsymbol{b})$；

(3)向量 $\boldsymbol{b}$ 可由向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性表示；

(4)向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 与 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n, \boldsymbol{b}$ 等价.

例 4.18 求方程组的通解

$$
\begin{cases}
x_1 + x_2 - 3x_3 - x_4 = 1 \\
3x_1 + 2x_2 - 3x_3 + 4x_4 = 4 \\
x_1 + 2x_2 - 9x_3 - 8x_4 = 0
\end{cases}
$$

解

$$\boldsymbol{B}=\begin{pmatrix}1&1&-3&-1&1\\3&2&-3&4&4\\1&2&-9&-8&0\end{pmatrix}\xrightarrow[r_2-3r_1]{r_3-r_1}\begin{pmatrix}1&1&-3&-1&1\\0&-1&6&7&1\\0&1&-6&-7&-1\end{pmatrix}$$

$$\xrightarrow{r_3+r_2}\begin{pmatrix}1&1&-3&-1&1\\0&-1&6&7&1\\0&0&0&0&0\end{pmatrix}\xrightarrow{r_1+r_2}\begin{pmatrix}1&0&3&6&2\\0&-1&6&7&1\\0&0&0&0&0\end{pmatrix}$$

于是有

$$\begin{cases}x_1=-3x_3-6x_4+2\\x_2=\quad 6x_3+7x_4-1\end{cases}\tag{4.10}$$

通解为 $\boldsymbol{x}=(2,-1,0,0)^{\mathrm{T}}+k_1(-3,6,1,0)^{\mathrm{T}}+k_2(-6,7,0,1)^{\mathrm{T}}$

$(k_1,k_2\in\mathbf{R})$

若在(4.10)中令 $x_3=k_1,x_4=k_2$,则

$$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}-3k_1-6k_2+2\\6k_1+7k_2-1\\k_1\\k_2\end{pmatrix}=k_1\begin{pmatrix}-3\\6\\1\\0\end{pmatrix}+k_2\begin{pmatrix}-6\\7\\0\\1\end{pmatrix}+\begin{pmatrix}2\\-1\\0\\0\end{pmatrix}$$

$(k_1,k_2\in\mathbf{R})$

*4.5 解线性方程组的迭代法

前面已完满地解决了线性方程组何时有解,何时无解以及有解时解的结构形式等理论问题,但还存在求解的技术问题,也就是对相容的方程组如何快速准确地求得它的解. 前面给出的消元法是在不考虑舍入误差时求精确解的直接方法,这种方法对高阶方程组由于占用计算机内存空间过多而难以在计算机上实现. 下面介绍一种求解线性方程组的近似方法——迭代法,其优点是占内存少且编制程序简单,因而是计算机上常用的算法之一.

4.5.1 迭代法的基本思想

设 n 阶线性方程组

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \tag{4.11}$$

的系数矩阵 $\boldsymbol{A}$ 非奇异,写出它的一个等价形式

$$\boldsymbol{x}=\boldsymbol{M}\boldsymbol{x}+\boldsymbol{g} \tag{4.12}$$

任给 $\boldsymbol{x}$ 的一个值 $\boldsymbol{x}^{(0)}$ 称为**初始向量**,代入(4.12)右边,算出的结果记为 $\boldsymbol{x}^{(1)}$,即 $\boldsymbol{x}^{(1)}=\boldsymbol{M}\boldsymbol{x}^{(0)}+\boldsymbol{g}$,再将 $\boldsymbol{x}^{(1)}$ 代入(4.12)的右边,算得的结果记为 $\boldsymbol{x}^{(2)}$,……一般地

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{M}\boldsymbol{x}^{(k)}+\boldsymbol{g} \tag{4.13}$$

它称为(4.11)的**迭代格式**,由 $\boldsymbol{x}^{(0)}$ 出发得到迭代序列 $\boldsymbol{x}^{(0)},\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(k)},\cdots$,如果这个序列有极限 $\boldsymbol{x}^*$,即

$$\lim_{k\to\infty}\boldsymbol{x}^{(k)}=\boldsymbol{x}^* \tag{4.14}$$

称迭代格式(4.13)收敛,这时 $\boldsymbol{x}^*$ 就是(4.11)的解.(4.14)是指

$$\lim_{k\to\infty}x_i^{(k)}=x_i^* \text{ 或 } |x_i^{(k)}-x_i^*|\to 0,i=1,2,\cdots,n.(k\to\infty)$$

这表示 n 维向量空间两向量 $\boldsymbol{x}^{(k)}$ 与 $\boldsymbol{x}^*$ 的“距离”趋于零,关于向量的“长度”及向量间的“距离”,下面将给出定义.

应用迭代法解线性方程组面临两个问题:

(1)如何构造一个收敛的迭代格式;

(2)由于迭代只能进行有限步,那么应在何时终止迭代以便能得到满意的近似解 $\boldsymbol{x}^{(k)}$.

4.5.2 向量的范数

n 维向量的范数是解析几何中向量长度的推广.

定义 4.9 n 维向量 $\boldsymbol{x}$ 的**范数** $\|\boldsymbol{x}\|$ 是满足下列关系的非负实数.

(1)对任何 $\boldsymbol{x}$,$\|\boldsymbol{x}\|\geqslant 0$,仅当 $\boldsymbol{x}=\boldsymbol{0}$ 时,$\|\boldsymbol{x}\|=0$(非负性);

(2)对任何实数 c,$\|c\boldsymbol{x}\|=|c|\,\|\boldsymbol{x}\|$(齐次性);

(3)对任何 n 维向量 $\boldsymbol{x},\boldsymbol{y}$,有 $\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$,(三角不等式).

设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$,由公式

$$\| \boldsymbol{x} \|_\infty = \max_{1 \leqslant i \leqslant n} \{ |x_i| \}$$

$$\| \boldsymbol{x} \|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

$$\| \boldsymbol{x} \|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

定义的实数 $\| \boldsymbol{x} \|_\infty$，$\| \boldsymbol{x} \|_1$，$\| \boldsymbol{x} \|_2$ 都满足范数的三个关系，因此它们都可作为 $\boldsymbol{x}$ 的范数，分别称为 $\boldsymbol{x}$ 的 l_∞（最大模）范数，l_1 范数和 l_2（欧几里德）范数.

有了范数的定义，应明确指出(4.14)式在什么范数下成立，如

对 l_1 范数，$\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^* \Leftrightarrow \| \boldsymbol{x}^{(k)} - \boldsymbol{x}^* \|_1 \to 0, (k \to \infty)$

对 l_2 范数，$\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^* \Leftrightarrow \| \boldsymbol{x}^{(k)} - \boldsymbol{x}^* \|_2 \to 0, (k \to \infty)$

对 l_∞ 范数，$\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^* \Leftrightarrow \| \boldsymbol{x}^{(k)} - \boldsymbol{x}^* \|_\infty \to 0, (k \to \infty)$

不过有定理保证，对某种范数 $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^*$ 成立时，则对任何范数都成立.

4.5.3 雅可比迭代法

对同一个方程组 $\boldsymbol{Ax} = \boldsymbol{b}$，可以构造各种迭代格式，从而有各种迭代方法，其中较简单的是雅可比(Jacobi)迭代法也称简单迭代法，用下例说明这种方法.

例 4.19 解方程组

$$\begin{cases} 10x_1 - x_2 - 2x_3 = 7.2 \\ -x_1 + 10x_2 - 2x_3 = 8.3 \\ -x_1 - x_2 + 5x_3 = 4.2 \end{cases} \tag{4.15}$$

解 将方程组化为等价形式

$$\begin{cases} x_1 = 0.1x_2 + 0.2x_3 + 0.72 \\ x_2 = 0.1x_1 + 0.2x_3 + 0.83 \\ x_3 = 0.2x_1 + 0.2x_2 + 0.84 \end{cases} \tag{4.16}$$

构造迭代格式

$$\begin{cases} x_1^{(k+1)} = 0.1x_2^{(k)} + 0.2x_3^{(k)} + 0.72 \\ x_2^{(k+1)} = 0.1x_1^{(k)} + 0.2x_3^{(k)} + 0.83 \\ x^{(k+1)} = 0.2x_1^{(k)} + 0.2x_2^{(k)} + 0.84 \end{cases} \tag{4.17}$$

(4.17)就是(4.15)的**雅可比迭代格式**,任取迭代初值 $\boldsymbol{x}^{(0)}=(0,0,0)$,依次迭代得一系列近似值 $\boldsymbol{x}^{(k)}$,表 4.1 是迭代 9 次的结果,由于继续迭代不会有较大变化,所以取

$$\boldsymbol{x}^* \approx \boldsymbol{x}^{(9)}=(1.09994,1.19994,1.29992)^{\mathrm{T}}$$

而精确解 $\boldsymbol{x}^*=(1.1,1.2,1.3)^{\mathrm{T}}$,因此误差量 $\boldsymbol{e}=\boldsymbol{x}^*-\boldsymbol{x}^{(9)}$. 可以取任何范数衡量误差大小,如取 $\|\cdot\|_\infty$,则

$$\|\boldsymbol{e}\|_\infty=8.0\times 10^{-5}$$

一般地,由于 $\boldsymbol{x}^*$ 未知,常以 $\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^{(k-1)}\|$ 作为误差估计,即当它很小时停止迭代,取 $\boldsymbol{x}=\boldsymbol{x}^{(k)}$ 为近似解,在本例中 $\|\boldsymbol{x}^{(9)}-\boldsymbol{x}^{(8)}\|_\infty=1.4\times 10^{-4}$,略大于实际误差.

综上所述,雅可比迭代法的解题步骤是:

(1)写出等价方程并构造收敛的迭代格式,如(4.17). 若由(4.15)的第二个方程分离出 x_1,第三个方程分离出 x_2 及第一个方程分离出 x_3,有

$$\begin{cases} x_1^{(k+1)}=10x_2^{(k)}-2x_3^{(k)}-8.3 \\ x_2^{(k+1)}=-x_1^{(k)}+5x_3^{(k)}-4.2 \\ x_3^{(k+1)}=5x_1^{(k)}-0.5x_2^{(k)}-3.6 \end{cases} \tag{4.18}$$

这个迭代格式不收敛.

(2)按收敛的迭代格式计算,直至 $\|\boldsymbol{x}^{(k-1)}-\boldsymbol{x}^{(k)}\|<\varepsilon$ 时,取 $\boldsymbol{x}^{(k)}$ 为近似解,这里 ε 是一较小的正数,$\|\cdot\|$ 是任一指定的范数.

4.5.4 塞得尔迭代法

仍以上例说明塞得尔(Seidel)迭代过程,由(4.16)构造迭代格式

$$\begin{cases} x_1^{(k+1)}= \qquad 0.1x_2^{(k)}+0.2x_3^{(k)}+0.72 \\ x_2^{(k+1)}=0.1x_1^{(k+1)} \qquad +0.2x_3^{(k)}+0.83 \\ x_3^{(k+1)}=0.2x_1^{(k+1)}+0.2x_2^{(k+1)} \qquad +0.84 \end{cases} \tag{4.19}$$

迭代初始向量仍取 $\boldsymbol{x}^{(0)}=\mathbf{0}$,迭代 6 次结果见表 4.2.

表 4.1

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
1	0.720 00	0.830 00	0.840 00
2	0.971 00	1.070 00	1.150 00
3	1.057 00	1.157 10	1.248 20
4	1.085 35	1.185 34	1.282 82
5	1.095 10	1.195 10	1.294 14
6	1.098 34	1.198 34	1.298 04
7	1.099 44	1.199 44	1.299 34
8	1.099 81	1.199 81	1.299 78
9	1.099 94	1.199 94	1.299 92

表 4.2

$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0.720 00	0.902 00	1.164 40
1.043 08	1.167 19	1.282 05
1.093 13	1.195 72	1.297 77
1.099 13	1.199 47	1.299 72
1.099 89	1.199 93	1.299 96
1.099 99	1.199 99	1.300 00

对比两个表可知，塞得尔迭代 6 次的结果已优于雅可比迭代 9 次的结果，可见塞得尔法的收敛速度高于雅可比法，这是由于塞得尔法每次都及时地利用了新值，对于收敛的迭代格式，靠后的 $\boldsymbol{x}^{(k)}$ 通常更接近于 $\boldsymbol{x}^*$.

塞得尔法除收敛快外，编制程序更简单且节省内存.

4.5.5 收敛定理

对一般 n 阶线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$，若能改写成等价方程组

$$\begin{cases} x_1 = c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n + g_1 \\ x_2 = c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n + g_2 \\ \quad\vdots \\ x_n = c_{n1}x_1 + c_{n2}x_2 + \cdots + c_{nn}x_n + g_n \end{cases} \tag{4.20}$$

或
$$x_i = \sum_{j=1}^{n} c_{ij}x_j + g_i, \quad i = 1,2,\cdots,n$$

（这里 x_i 未必是由原来的第 i 个方程分离出来的）. 雅可比迭代格式为

$$x_i^{(k+1)}=\sum_{j=1}^{n}c_{ij}x_j^{(k)}+g_i,\quad i=1,2,\cdots,n \tag{4.21}$$

塞得尔迭代格式为

$$x_i^{(k+1)}=\sum_{j=1}^{i-1}c_{ij}x_j^{(k+1)}+\sum_{j=i}^{n}c_{ij}x_j^{(k)}+g_i,\quad i=1,2,\cdots,n \tag{4.22}$$

定理 4.8　在等价方程组(4.20)中,若

$$\boldsymbol{L}=\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|c_{ij}|<1$$

则(4.21)和(4.22)这两种迭代格式对任意给定的迭代初值均收敛.

仅就(4.21)给出证明.

证　设精确解 $\boldsymbol{x}^*=(x_1^*,x_2^*,\cdots,x_n^*)^{\mathrm{T}}$,(4.21)第 k 步迭代为 $\boldsymbol{x}^{(k)}=(x_1^{(k)},x_2^{(k)}\cdots,x_n^{(k)})^{\mathrm{T}}$,即

$$x_i^*=\sum_{j=1}^{n}c_{ij}x_j^*+g_i,\quad x_i^{(k)}=\sum_{j=1}^{n}c_{ij}x_j^{(k-1)}+g_i,\quad i=1,2,\cdots,n$$

误差为

$$\begin{aligned}\|\boldsymbol{e}^{(k)}\|_\infty&=\max_{1\leqslant i\leqslant n}\{|x_i^*-x_i^{(k)}|\}\\&\leqslant\max_{1\leqslant i\leqslant n}\{\sum_{j=1}^{n}|c_{ij}(x_j^*-x_j^{(k-1)})|\}\\&\leqslant\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|c_{ij}|\max_{1\leqslant j\leqslant n}|x_j^*-x_j^{(k-1)}|\\&=L\|\boldsymbol{x}^*-\boldsymbol{x}^{(k-1)}\|_\infty=L\|e^{(k-1)}\|_\infty\\&=\cdots=L^k\|\boldsymbol{e}^{(0)}\|_\infty\to 0,\quad k\to\infty\end{aligned}$$

这是由等价形式(或迭代格式)右端未知数的系数来判断迭代格式的收敛性,其实由该定理可以得到一个直接由原方程组判断迭代格式收敛的充分条件,即下列定理.

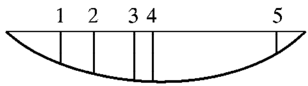
定理 4.9　设 n 阶线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的系数矩阵 $\boldsymbol{A}=(a_{ij})$ 按行或按列严格对角占优,即满足

$$a_{ii}>\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|,\text{或者}\ |a_{ii}|>\sum_{\substack{i=1\\i\neq j}}^{n}|a_{ij}|\quad,i=1,2,\cdots,n$$

那么 $Ax=b$ 有唯一解，且雅可比和塞得尔迭代均收敛.

证明略.

例 4.20 如图表示一水平放置的梁在垂直向下作用力下的弯曲形变.



在不同点每千牛力使各点下降的挠度(单位为 cm)列于表中：

挠度 / 作用点 / 下垂点	1	2	3	4	5
1	5.0	1.5	1.3	0.9	0.0
2	1.5	4.5	1.4	1.0	0.5
3	1.3	1.4	5.5	1.5	0.7
4	0.9	1.0	1.5	5.5	1.2
5	0.0	0.5	0.7	1.2	2.5

假定在第 i 点加力 F_i 时下降挠度为 h_i，试求在各点容许挠度为

$$h_1=21.1,\ h_2=21.9,\ h_3=23.1,\ h_4=26.4,\ h_5=10.0$$

时，各点所能承受的力.

解 设第 i 点所承受的力为 F_i，由力的线性迭加原理

$$\begin{cases}5.0F_1+1.5F_2+1.3F_3+0.9F_4 & =21.1\\ 1.5F_1+4.5F_2+1.4F_3+1.0F_4+0.5F_5 & =21.9\\ 1.3F_1+1.4F_2+5.5F_3+1.5F_4+0.7F_5 & =23.1\\ 0.9F_1+1.0F_2+1.5F_3+5.5F_4+1.2F_5 & =26.4\\ \qquad\qquad 0.5F_2+0.7F_3+1.2F_4+2.5F_5 & =10.0\end{cases}$$

可以验证其系数矩阵严格对角占优，因此有收敛的塞得尔迭代

$$\begin{cases} F_1^{(k+1)}=-0.300\,0F_2^{(k)}-0.260\,0F_3^{(k)}-0.180\,0F_4^{(k)}+4.220\,0 \\ F_2^{(k+1)}=-0.333\,3F_1^{(k+1)}-0.311\,1F_3^{(k)}-0.222\,2F_4^{(k)} \\ \qquad\qquad -0.111\,1F_5^{(k)}+4.866\,7 \\ F_3^{(k+1)}=-0.236\,4F_1^{(k+1)}-0.254\,5F_2^{(k+1)}-0.272\,7F_4^{(k)} \\ \qquad\qquad -0.127\,3F_5^{(k)}+4.200\,0 \\ F_4^{(k+1)}=-0.163\,6F_1^{(k+1)}-0.181\,8F_2^{(k+1)}-0.272\,7F_3^{(k+1)} \\ \qquad\qquad -0.218\,2F_5^{(k)}+4.800\,0 \\ F_5^{(k+1)}=-0.200\,0F_2^{(k+1)}-0.280\,0F_3^{(k+1)}-0.480\,0F_4^{(k+1)}+4.00\,0 \end{cases}$$

取迭代初值 $\boldsymbol{F}^{(0)}=\boldsymbol{0}$，迭代 6 次的结果列于表中

k	$F_1^{(k)}$	$F_2^{(k)}$	$F_3^{(k)}$	$F_4^{(k)}$	$F_5^{(K)}$
0	0	0	0	0	0
1	4.220 0	3.460 2	2.321 8	2.847 4	1.291 1
2	2.065 8	2.679 7	2.088 8	3.123 5	1.379 9
3	2.310 8	2.599 3	1.964 8	3.112 5	1.436 0
4	2.369 1	2.614 7	1.942 9	3.093 9	1.448 0
5	2.373 5	2.622 8	1.943 4	3.089 0	1.448 6
6	2.371 9	2.624 3	1.944 6	3.088 5	1.448 2

$$\| \boldsymbol{F}^{(6)}-\boldsymbol{F}^{(5)} \|_\infty=1.6\times10^{-3}$$

$\boldsymbol{F}$ 的近似值为

$$F_1=2.371\,9,\quad F_2=2.624\,3,\quad F_3=1.944\,6,$$
$$F_4=3.088\,5,\quad F_5=1.448\,2.$$

习　题　四

1. 如果把定义 4.2 中的“不全为零”改为“全不为零”，是否还与定义相同？

2. 举例说明下列命题是错误的：

(1)若向量 $\boldsymbol{\alpha}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示，则表示式是唯一的.

(2)若向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性相关,则其中任一向量均可由其余向量线性表示.

(3)若有不全为零的数 $k_1,k_2,\cdots,k_r$ 使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r+k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_r\boldsymbol{\beta}_r=\mathbf{0}$$

且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性相关,则 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性相关.

3. 证明:若 $\boldsymbol{\alpha}$ 可由线性无关的向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示,则表示式唯一.

4. 证明:若 n 维单位坐标向量 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 可由 n 维向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示,则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关.

5. 证明:n 维向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关的充要条件是任一 n 维向量均可由它们线性表示.

6. 判断下列向量组的线性相关性,求它们的秩和最大无关组:

(1) $\boldsymbol{\alpha}_1=(1,2,-1,4)$, $\boldsymbol{\alpha}_2=(9,100,10,4)$,
$\boldsymbol{\alpha}_3=(-2,-4,2,-8)$

(2) $\boldsymbol{\alpha}_1=(1,1,0)$, $\boldsymbol{\alpha}_2=(0,2,0)$, $\boldsymbol{\alpha}_3=(0,0,3)$

(3) $\boldsymbol{\alpha}_1=(3,-2,0,1)$, $\boldsymbol{\alpha}_2=(0,2,2,1)$
$\boldsymbol{\alpha}_3=(1,-2,-3,-2)$, $\boldsymbol{\alpha}_4=(0,1,2,1)$

(4) $\boldsymbol{\alpha}_1=(3,1,0,2)$,$\boldsymbol{\alpha}_2=(1,-1,2,-1)$,$\boldsymbol{\alpha}_3=(1,3,-4,4)$

7. 设 $V_1=\{\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\mid x_i\in\mathbf{R},i=1,2,\cdots,n,\sum_{i=1}^{n}x_i=0\}$

$V_2=\{\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\mid x_i\in\mathbf{R},i=1,2\cdots,n,\sum_{i=1}^{n}x_i=1\}$

问 V_1、V_2 是否是向量空间,若是,指出与 R^n 的关系.

8. 证明:$R^3=\{\boldsymbol{x}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3\mid k_i\in\mathbf{R}\}$,其中

$\boldsymbol{\alpha}_1=(0,1,1)$,$\boldsymbol{\alpha}_2=(1,0,1)$,$\boldsymbol{\alpha}_3=(1,1,0)$.

9. 验证 $\boldsymbol{\alpha}_1=(1,-1,0)$,$\boldsymbol{\alpha}_2=(2,1,3)$,$\boldsymbol{\alpha}_3=(3,1,2)$是 R^3 的一组基,并分别把 $\boldsymbol{\beta}_1=(5,0,7)$,$\boldsymbol{\beta}_2=(-9,-8,-13)$用这组基表示出来.

10. 证明 $\boldsymbol{\alpha}_1=(1,1,0,0)$，$\boldsymbol{\alpha}_2=(1,0,1,1)$ 和 $\boldsymbol{\beta}_1=(2,-1,3,3)$，$\boldsymbol{\beta}_2=(0,1,-1,-1)$ 是同一向量空间的基，并求由基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 到基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 的过渡矩阵.

11. 求下列齐次线性方程组的基础解系和通解

(1) $\begin{cases} 3x_1+x_2+8x_3+2x_4+x_5=0 \\ 2x_1-2x_2-3x_3-7x_4+2x_5=0 \\ x_1+11x_2-12x_3+34x_4-5x_5=0 \\ x_1-5x_2+2x_3-16x_4+3x_5=0 \end{cases}$

(2) $\begin{cases} x_1-2x_2+x_3+x_4-x_5=0 \\ 2x_1+x_2-x_3-x_4+x_5=0 \\ x_1+7x_2-5x_3-5x_4+5x_5=0 \\ 3x_1-x_2-2x_3+x_4-x_5=0 \end{cases}$

12. 求下列非齐次线性方程组的通解

(1) $\begin{cases} 2x_1-x_2+3x_3-x_4=1 \\ 3x_1-2x_2-2x_3+3x_4=3 \\ x_1-x_2-5x_3+4x_4=2 \\ 7x_1-5x_2-9x_3+10x_4=8 \end{cases}$

(2) $\begin{cases} x_1-2x_2+3x_3-4x_4=4 \\ x_2-x_3+x_4=-3 \\ x_1-3x_2-3x_4=1 \\ -7x_2+3x_3+x_4=-3 \end{cases}$

(3) $\begin{cases} x_1-4x_2-3x_3=1 \\ x_1-5x_2-3x_3=0 \\ x_1+6x_2+4x_3=4 \end{cases}$

13. 讨论并解线性方程组

$$\begin{cases} \lambda x_1+x_2+x_3=1 \\ x_1+\lambda x_2+x_3=\lambda \\ x_1+x_2+\lambda x_3=\lambda^2 \end{cases}$$

14. 方程组 $\begin{cases} a_1x+b_1y=c_1 \\ a_2x+b_2y=c_2 \end{cases}$ 的每一方程在平面上表示一条直

线，试说明方程组无解、有唯一解及有无穷多个解的几何意义. 类似地，说明

$$\begin{cases} a_1x+b_1y+c_1z=d_1 \\ a_2x+b_2y+c_2z=d_2 \\ a_3x+b_3y+c_3z=d_3 \end{cases}$$

无解、有唯一解、有无穷多个解的几何意义.

15△. 设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关，讨论向量组 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\beta}_s=\boldsymbol{\alpha}_s+\boldsymbol{\alpha}_1$ 的线性相关性.

16△. 设 $\boldsymbol{A},\boldsymbol{B}$ 均为 n 阶方阵，$R(\boldsymbol{A})=r,R(\boldsymbol{B})=s$，且 $\boldsymbol{AB}=\boldsymbol{O}$，证明 $r+s\leqslant n$.

17△. 已知 $\boldsymbol{\alpha}_1=(1,0,2,3)$，　$\boldsymbol{\alpha}_2=(1,1,3,5)$，

$\boldsymbol{\alpha}_3=(1,-1,a+2,1)$，　$\boldsymbol{\alpha}_4=(1,2,4,a+8)$，

$\boldsymbol{\beta}=(1,1,b+3,5)$

(1)a,b 为何值时，$\boldsymbol{\beta}$ 不能表示成 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的线性组合？

(2)a,b 为何值时，$\boldsymbol{\beta}$ 有 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的唯一的线性表示式？并写出该表示式.

18△. 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 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.

19. 设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 中两两线性无关，问能否由此推出该向量组线性无关？若能，证明你的结论；若不能，举出反例.

20. 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，且可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性表示，证明 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性无关.

21. 已知向量组 A 与向量组 B 有相同的秩，且组 A 可由组 B 线性表示，证明向量组 A 与向量组 B 等价.

22. 证明 n 维向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关的充要条件是

$$\boldsymbol{G}=\begin{vmatrix} \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_n \\ \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_n \\ \vdots & \vdots & & \vdots \\ \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_n \end{vmatrix}\neq 0$$

23$^{\triangle}$. 设 $\boldsymbol{A}$ 是 $n\times m$ 维矩阵,$\boldsymbol{B}$ 为 $m\times n$ 维矩阵,其中 $n<m$,$\boldsymbol{I}$ 是 n 维单位矩阵,若 $\boldsymbol{AB}=\boldsymbol{I}$,证明 $\boldsymbol{B}$ 的列向量线性无关.

24$^{\triangle}$. 已知 $\boldsymbol{Q}=\begin{pmatrix}1 & 2 & 3\\ 2 & 4 & t\\ 3 & 6 & 9\end{pmatrix}$,$\boldsymbol{P}$ 为三阶非零矩阵,$\boldsymbol{PQ}=\boldsymbol{O}$,则

(A)$t=6$ 时,$\boldsymbol{P}$ 的秩必为 1, (B)$t=6$ 时,$\boldsymbol{P}$ 的秩必为 2,

(C)$t\neq 6$ 时,$\boldsymbol{P}$ 的秩必为 1, (D)$t\neq 6$ 时,$\boldsymbol{P}$ 的秩必为 2.

25*. 用两种迭代法解线性方程组

$$\begin{cases} 0.78x_1-0.02x_2-0.12x_3-0.14x_4=0.76\\ -0.02x_1+0.86x_2-0.04x_3-0.06x_4=0.08\\ -0.12x_1-0.04x_2+0.72x_3-0.08x_4=1.12\\ -0.14x_1-0.06x_2-0.08x_3+0.74x_4=0.68\end{cases}$$

第5章　特征值　特征向量　二次型

以上所讨论的基本上都是围绕着解线性方程组的问题，但为了研究其它的应用，例如线性微分方程组的解法，弹性系统的振动问题，预测某种生物群体的增殖情况以及可以化为函数最优化的很多实际问题，都需要这一章的知识. 事实上，对于矩阵理论的所有进一步的发展，本章介绍的概念都是很关键的.

5.1　正交向量组与正交矩阵

5.1.1　向量的内积与许瓦兹不等式

在第1章里，曾引进向量内积的概念，即对 n 维列向量（今后除特别指明，凡说到向量时，均指列向量，且其分量均为实数）

$$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}},\qquad \boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$$

称 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=x_1y_1+x_2y_2+\cdots+x_ny_n$ 为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的**内积**，记为 $[\boldsymbol{x},\boldsymbol{y}]$.

不难看出，n 维向量的内积是解析几何中向量的数量积的推广. 类似于向量的数量积，n 维向量的内积满足下列运算规律（其中 $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}$ 均为 n 维向量，k 为实数）：

（1）　$[\boldsymbol{x},\boldsymbol{y}]=[\boldsymbol{y},\boldsymbol{x}]$

（2）　$[k\boldsymbol{x},\boldsymbol{y}]=k[\boldsymbol{x},\boldsymbol{y}]$

（3）　$[\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z}]=[\boldsymbol{x},\boldsymbol{z}]+[\boldsymbol{y},\boldsymbol{z}]$

在解析几何中，曾定义过向量的长度和夹角，但对 n 维向量，没有那样直观的长度和夹角概念，因此，通常按数量积的直角坐标计算公式来推广，并反过来利用内积来定义 n 维向量的长度和夹角.

在第4章4.5节中已定义过 n 维向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 的

长度(或称范数)为 $\sqrt{[\boldsymbol{x},\boldsymbol{x}]}=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}$，记为 $\|\boldsymbol{x}\|$. 如同解析几何中一样，$\|\boldsymbol{x}\|$ 具有下列性质：

(1)非负性：对于任何向量 $\boldsymbol{x}$，有 $\|\boldsymbol{x}\|\geqslant 0$，当且仅当 $\boldsymbol{x}=\boldsymbol{0}$ 时，$\|\boldsymbol{x}\|=0$；

(2)齐次性：$\|k\boldsymbol{x}\|=|k|\ \|\boldsymbol{x}\|$；

(3)三角不等式：$\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$.

当 $\|\boldsymbol{x}\|=1$ 时，称 $\boldsymbol{x}$ 为**单位向量**.

对任何实数 k 和 n 维向量 $\boldsymbol{x},\boldsymbol{y}$，有

$$[k\boldsymbol{x}+\boldsymbol{y},k\boldsymbol{x}+\boldsymbol{y}]\geqslant 0$$

即
$$k^2[\boldsymbol{x},\boldsymbol{x}]+2k[\boldsymbol{x},\boldsymbol{y}]+[\boldsymbol{y},\boldsymbol{y}]\geqslant 0$$

其左端是关于 k 的二次三项式，由于它不改变符号，所以其判别式

$$(2[\boldsymbol{x},\boldsymbol{y}])^2-4[\boldsymbol{x},\boldsymbol{x}][\boldsymbol{y},\boldsymbol{y}]\leqslant 0$$

即
$$[\boldsymbol{x},\boldsymbol{y}]^2\leqslant[\boldsymbol{x},\boldsymbol{x}][\boldsymbol{y},\boldsymbol{y}] \tag{5.1}$$

(5.1)称为**许瓦兹**(Schwarz)不等式，该不等式在数学物理中有着广泛应用.

5.1.2 正交向量组与正交化方法

由许瓦兹不等式可知，当 $\|\boldsymbol{x}\|\ \|\boldsymbol{y}\|\neq 0$ 时可得

$$\left|\frac{[\boldsymbol{x},\boldsymbol{y}]}{\|\boldsymbol{x}\|\ \|\boldsymbol{y}\|}\right|\leqslant 1$$

于是可以定义 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 夹角 $\boldsymbol{\theta}$ 的余弦为

$$\cos\theta=\frac{[\boldsymbol{x},\boldsymbol{y}]}{\|\boldsymbol{x}\|\ \|\boldsymbol{y}\|}$$

作为解析几何中向量垂直概念的自然推广，当 $[\boldsymbol{x},\boldsymbol{y}]=0$ 时，称向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ **正交**. 显然，零向量与任何向量正交.

对不含零向量的向量组，若其中的向量两两正交，则称该向量组为**正交向量组**. 如 n 维单位坐标向量组就是一个正交向量组.

定理 5.1　若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 为正交向量组，则它线性无关.

证　设有数 $k_1,k_2,\cdots,k_r$ 使 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r=\boldsymbol{0}$，用 $\boldsymbol{\alpha}_i^{\mathrm{T}}$ 左乘两边得 $k_i\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\alpha}_i=0$，因 $\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\alpha}_i\neq 0$ 所以 $k_i=0$，此结果对 $i=1,2,$

$\cdots,r$ 均成立. 因此 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关.

如果向量空间的一组基是正交向量组,则称它为向量空间的**正交基**. 在讨论向量空间时,经常采用正交基,正像解析几何中采用直角坐标一样.

那么,对一个向量空间,能否找出它的一组正交基呢? 下面的定理给出了肯定的回答.

定理 5.2　若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 为 n 维正交向量组,且 $r<n$,则必有非零 n 维向量 $\boldsymbol{x}$,使 $\boldsymbol{x}$ 与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 两两正交.

证　要求的 n 维非零向量 $\boldsymbol{x}$ 应满足

$$\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{x}=0 \qquad i=1,2,\cdots,r$$

即

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)^{\mathrm{T}}\boldsymbol{x}=\mathbf{0} \tag{5.2}$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关,且 $r<n$,所以 $R((\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)^{\mathrm{T}})=r<n$,由线性方程组的相容性定理知齐次线性方程组(5.2)一定有非零解向量 $\boldsymbol{x}$,非零解向量即为所求的 $\boldsymbol{x}$.

推论　对 $r(r<n)$ 个两两正交的 n 维非零向量,总可以添上 $n-r$ 个 n 维非零向量,使 n 个向量两两正交,从而这 n 个向量就构成了向量空间 R^n 的一组正交基.

这样,对任何一个 r 维向量空间,都可以按上述结论得到它的一组正交基. 作法是,先从向量空间中任取一非零向量 $\boldsymbol{\alpha}_1$,接下来按证明定理 5.2 的方法在该向量空间中找非零向量 $\boldsymbol{\alpha}_2$,使 $\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_2=0$,再找 $\boldsymbol{\alpha}_3$,使 $\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\alpha}_3=0$,$i=1,2$,如此下去,便得到 r 个两两正交的向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$,它就是要求的一组正交基.

例 5.1　已知 R^3 的一个向量 $\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}}$,求 R^3 的一组正交基.

解　求 $\boldsymbol{\alpha}_2=(x_{21},x_{22},x_{23})^{\mathrm{T}}$,使 $\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_2=0$,即

$$x_{21}+x_{22}+x_{23}=0$$

取 $x_{22}=0$,$x_{23}=-1$,得 $x_{21}=1$,则 $\boldsymbol{\alpha}_2=(1,0,-1)^{\mathrm{T}}$ 与 $\boldsymbol{\alpha}_1$ 正交.

再求 $\boldsymbol{\alpha}_3=(x_{31},x_{32},x_{33})^{\mathrm{T}}$,使 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)^{\mathrm{T}}\boldsymbol{\alpha}_3=\mathbf{0}$,

$$\text{由 } \boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\ \boldsymbol{\alpha}_2^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}1 & 1 & 1\\ 1 & 0 & -1\end{pmatrix}\text{得 }\begin{cases}x_{31}+x_{32}=-x_{33}\\ x_{31}=x_{33}\end{cases}$$

取 $x_{33}=1$，得 $(x_{31},x_{32})=(1,-2)$，于是 $\boldsymbol{\alpha}_3=(1,-2,1)^{\mathrm{T}}$ 即为所求，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 就构成了 R^3 的一组正交基.

把解析几何中直角坐标系的单位坐标向量推广到一般向量空间便有

定义 5.1 设 n 维向量 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_r$ 是向量空间 $V(V\subset R^n)$ 的一组正交基. 如果它们均为单位向量，则称 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_r$ 为 V 的一组**正交规范基**或**标准正交基**.

例如 $\boldsymbol{e}_1=(1,0,0,0)^{\mathrm{T}}$，$\quad \boldsymbol{e}_2=(0,1,0,0)^{\mathrm{T}}$

$\boldsymbol{e}_3=(0,0,1,0)^{\mathrm{T}}$，$\quad \boldsymbol{e}_4=(0,0,0,1)^{\mathrm{T}}$

与
$$\boldsymbol{\varepsilon}_1=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0,0\right)^{\mathrm{T}},\quad \boldsymbol{\varepsilon}_2=\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},0,0\right)^{\mathrm{T}}$$
$$\boldsymbol{\varepsilon}_3=\left(0,0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^{\mathrm{T}},\quad \boldsymbol{\varepsilon}_4=\left(0,0,\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$$

都是 R^4 的正交规范基.

如果 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_r$ 是向量空间 V 的一组正交规范基，则 V 中的任一向量 $\boldsymbol{\alpha}$ 可由这组基线性表示，设表示式为 $\boldsymbol{\alpha}=k_1\boldsymbol{\varepsilon}_1+k_2\boldsymbol{\varepsilon}_2+\cdots+k_r\boldsymbol{\varepsilon}_r$，这些 $k_i(i=1,2\cdots,r)$ 便是向量 $\boldsymbol{\alpha}$ 在这组正交规范基下的坐标，为求坐标 k_i，用 $\boldsymbol{\varepsilon}_i^T$ 左乘两边便得
$$k_i=[\boldsymbol{\varepsilon}_i,\boldsymbol{\alpha}],i=1,2,\cdots,r.$$

定理 5.2 及推论给出了求向量空间的正交基的方法，但现在有这样一个问题，若已知向量空间的一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$，能否从它导出一组正交规范基呢？或更一般地，若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，能否找到两两正交的单位向量 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_r$，使两个向量组等价呢？回答是肯定的. 这样的问题称为把基(或线性无关向量组)**正交规范化的施密特**(Schimidt)**方法**.

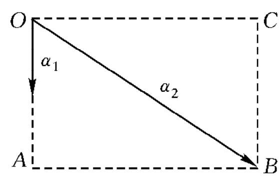


图 5.1

设不共线的两非零向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$(如图 5.1 所示)，先考虑正交化，记 $\boldsymbol{b}_1=\boldsymbol{\alpha}_1$，设 $\boldsymbol{\alpha}_1^0$ 为 $\boldsymbol{\alpha}_1$ 的单位向量，即 $\boldsymbol{\alpha}_1^0=\dfrac{\boldsymbol{\alpha}_1}{\|\boldsymbol{\alpha}_1\|}$，以 $\boldsymbol{\alpha}_2$ 为对

角线，$\boldsymbol{\alpha}_1$ 所在直线为边作矩形 $OABC$. 显然

$$\boldsymbol{\alpha}_2=\overrightarrow{OA}+\overrightarrow{OC}, \qquad \overrightarrow{OC}\perp\overrightarrow{OA}$$

而

$$\overrightarrow{OA}=(\mathrm{Prj}_{\boldsymbol{\alpha}_1}\boldsymbol{\alpha}_2)\boldsymbol{\alpha}_1^0=\frac{[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2]}{\|\boldsymbol{\alpha}_1\|}\boldsymbol{\alpha}_1^0=\frac{[\boldsymbol{b}_1,\boldsymbol{\alpha}_2]}{\|\boldsymbol{b}_1\|}\boldsymbol{b}_1^0=\frac{[\boldsymbol{b}_1,\boldsymbol{\alpha}_2]}{[\boldsymbol{b}_1,\boldsymbol{b}_1]}\boldsymbol{b}_1$$

因此 $$\overrightarrow{OC}=\boldsymbol{\alpha}_2-\overrightarrow{OA}=\boldsymbol{\alpha}_2-\frac{[\boldsymbol{b}_1,\boldsymbol{\alpha}_2]}{[\boldsymbol{b}_1,\boldsymbol{b}_1]}\boldsymbol{b}_1$$

记 $\boldsymbol{b}_2=\overrightarrow{OC}$，则 $[\boldsymbol{b}_1,\boldsymbol{b}_2]=0$，把这种方法推广到一般的线性无关向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$，得到下列正交化公式：

$$\boldsymbol{b}_1=\boldsymbol{\alpha}_1$$

$$\boldsymbol{b}_2=\boldsymbol{\alpha}_2-\frac{[\boldsymbol{b}_1,\boldsymbol{\alpha}_2]}{[\boldsymbol{b}_1,\boldsymbol{b}_1]}\boldsymbol{b}_1$$

$$\boldsymbol{b}_3=\boldsymbol{\alpha}_3-\frac{[\boldsymbol{b}_1,\boldsymbol{\alpha}_3]}{[\boldsymbol{b}_1,\boldsymbol{b}_1]}\boldsymbol{b}_1-\frac{[\boldsymbol{b}_2,\boldsymbol{\alpha}_3]}{[\boldsymbol{b}_2,\boldsymbol{b}_2]}\boldsymbol{b}_2$$

$$\vdots$$

$$\boldsymbol{b}_r=\boldsymbol{\alpha}_r-\frac{[\boldsymbol{b}_1,\boldsymbol{\alpha}_r]}{[\boldsymbol{b}_1,\boldsymbol{b}_1]}\boldsymbol{b}_1-\frac{[\boldsymbol{b}_2,\boldsymbol{\alpha}_r]}{[\boldsymbol{b}_2,\boldsymbol{b}_2]}\boldsymbol{b}_2-\cdots-\frac{[\boldsymbol{b}_{r-1},\boldsymbol{\alpha}_r]}{[\boldsymbol{b}_{r-1},\boldsymbol{b}_{r-1}]}\boldsymbol{b}_{r-1}$$

然后取

$$\boldsymbol{\varepsilon}_i=\frac{\boldsymbol{b}_i}{\|\boldsymbol{b}_i\|},(i=1,2,\cdots,r)$$

显然 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_r$ 为正交规范化的向量组，且与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 等价.

上述由线性无关向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 导出正交向量组 $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_r$ 的方法称为施密特(Schmidt)正交化方法. 这种方法导出的正交向量组满足这样的关系，即对任何 $k(1\leqslant k\leqslant r)$，向量组 $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_k$ 与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 等价.

例 5.2 已知 $\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=(1,2,1)^{\mathrm{T}}$，$\boldsymbol{\alpha}_3=(1,1,2)^{\mathrm{T}}$ 线性无关，试将它们正交规范化.

解 取 $\boldsymbol{b}_1=\boldsymbol{\alpha}_1$

$$\boldsymbol{b}_2=\boldsymbol{\alpha}_2-\frac{[\boldsymbol{b}_1,\boldsymbol{\alpha}_2]}{[\boldsymbol{b}_1,\boldsymbol{b}_1]}\boldsymbol{b}_1=\begin{pmatrix}1\\2\\1\end{pmatrix}-\frac{4}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}=\frac{1}{3}\begin{pmatrix}-1\\2\\-1\end{pmatrix}$$

$$\boldsymbol{b}_3=\boldsymbol{\alpha}_3-\frac{[\boldsymbol{b}_1,\boldsymbol{\alpha}_3]}{[\boldsymbol{b}_1,\boldsymbol{b}_1]}\boldsymbol{b}_1-\frac{[\boldsymbol{b}_2,\boldsymbol{\alpha}_3]}{[\boldsymbol{b}_2,\boldsymbol{b}_2]}\boldsymbol{b}_2$$

$$=\begin{pmatrix}1\\1\\2\end{pmatrix}-\frac{4}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}+\frac{1}{2}\times\frac{1}{3}\begin{pmatrix}-1\\2\\-1\end{pmatrix}=\frac{1}{2}\begin{pmatrix}-1\\0\\1\end{pmatrix}$$

则 $\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3$ 两两正交. 令

$$\boldsymbol{\varepsilon}_1=\frac{\boldsymbol{b}_1}{\|\boldsymbol{b}_1\|}=\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)^{\mathrm{T}}$$

$$\boldsymbol{\varepsilon}_2=\frac{\boldsymbol{b}_2}{\|\boldsymbol{b}_2\|}=\left(-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},-\frac{1}{\sqrt{6}}\right)^{\mathrm{T}}$$

$$\boldsymbol{\varepsilon}_3=\frac{\boldsymbol{b}_3}{\|\boldsymbol{b}_3\|}=\left(-\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$$

则 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 两两正交且均为单位向量.

5.1.3 正交矩阵与正交变换

由正交规范化的 n 维向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为列所组成的矩阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 具有性质

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1^{T}\\\boldsymbol{\alpha}_2^{\mathrm{T}}\\\vdots\\\boldsymbol{\alpha}_n^{\mathrm{T}}\end{pmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=\begin{pmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_n\\\boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_n\\\vdots & \vdots & & \vdots\\\boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_n\end{pmatrix}$$

$$=\begin{pmatrix}1 & 0 & \cdots & 0\\0 & 1 & \cdots & 0\\\vdots & \vdots & & \vdots\\0 & 0 & \cdots & 1\end{pmatrix}=\boldsymbol{I}$$

根据这个特性,给出所谓正交矩阵的定义.

定义 5.2 如果 n 阶方阵 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$$

则称 $\boldsymbol{A}$ 为**正交矩阵**.

显然,若 $\boldsymbol{A}$ 是正交矩阵,则它一定是非奇异的,即 $\boldsymbol{A}^{-1}$ 存在. 关

于正交矩阵，还有下列的结论.

定理 5.3 如果 $\boldsymbol{A},\boldsymbol{B}$ 均为 n 阶正交矩阵，那么

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

(2) $\boldsymbol{A}^{\mathrm{T}}$ 即 $\boldsymbol{A}^{-1}$ 为正交矩阵；

(3) $\dfrac{1}{\sqrt{2}}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{A}\\ -\boldsymbol{A} & \boldsymbol{A}\end{pmatrix}$ 为 $2n$ 阶正交矩阵；

(4) $\boldsymbol{AB}$、$\boldsymbol{BA}$ 都是正交矩阵.

其证明留作练习. 下面的定理给出了正交矩阵的具体构造.

定理 5.4 n 阶方阵 $\boldsymbol{A}$ 为正交矩阵的充要条件是 $\boldsymbol{A}$ 的列(行)向量两两正交且均为单位向量.

证 用 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 表示 $\boldsymbol{A}$ 的列向量，即 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 则

$$\boldsymbol{A}\text{ 为正交矩阵}\Leftrightarrow\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}\Leftrightarrow\begin{pmatrix}\boldsymbol{\alpha}_1^{T}\\ \boldsymbol{\alpha}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\alpha}_n^{\mathrm{T}}\end{pmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=\boldsymbol{I}$$

$$\Leftrightarrow\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\alpha}_j=\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}\quad i,j=1,2,\cdots,n$$

由定理 5.3 的(2)知上述结论对行向量也成立.

由此定理可见，n 阶正交矩阵的行(列)向量就构成了向量空间 R^n 的一组正交规范基.

定义 5.3 设 $\boldsymbol{P}$ 为 n 阶正交矩阵，$\boldsymbol{x}$ 为 n 维列向量，则线性变换 $\boldsymbol{y}=\boldsymbol{Px}$ 称为**正交变换.**

正交变换具有良好的特性，那就是任何向量 $\boldsymbol{x}$ 在正交变换下长度不变. 事实上，设 $\boldsymbol{y}=\boldsymbol{Px}$ 为正交变换，则有

$$\|\boldsymbol{y}\|=\sqrt{[\boldsymbol{y},\boldsymbol{y}]}=\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{Px}}=\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}=\|\boldsymbol{x}\|.$$

它意味着正交变换不改变向量的长度，这正是线性代数中常常使用正交变换的理由. 实际上，正交变换相当于解析几何中的旋转和反射，例如

$$\boldsymbol{P}=\begin{pmatrix}-\cos\theta & \sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}=\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

为正交矩阵,这个正交变换 $\boldsymbol{y}=\boldsymbol{P}\boldsymbol{x}$ 相当于将向量 $\boldsymbol{x}$ 在平面直角坐标系中绕原点逆时针旋转 θ 角,再关于纵轴对称反射.

5.2 方阵的特征值和特征向量

5.2.1 引例

例 5.3 分析图 5.2 所示的两个自由度系统的无阻尼自由振动的特性. m_1,m_2 表示质量,k_1,k_2 表示弹簧的弹性系数,x_1,x_2 表示物体离开各自平衡位置的位移. 根据牛顿第二定律,系统运动的微分方程为

$$\begin{cases}m_1\ddot{x}_1+k_1(x_1-x_2)=0\\ m_2\ddot{x}_2-k_1x_1+(k_1+k_2)x_2=0\end{cases}\tag{5.3}$$

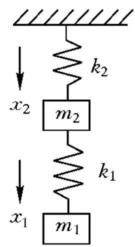


图 5.2

其中 $\ddot{x}$ 表示$\dfrac{\mathrm{d}^2x}{\mathrm{d}t^2}$,由振动理论知,(5.3)有下列形式的解

$$x_1=y_1\sin\omega t,\qquad x_2=y_2\sin\omega t$$

其中待求的 y_1,y_2 表示振幅,ω 表示角频率. 将它们代入式(5.3)得

$$\begin{cases}\left(\omega^2-\dfrac{k_1}{m_1}\right)y_1+\dfrac{k_1}{m_1}y_2=0\\ \dfrac{k_1}{m_2}y_1+\left(\omega^2-\dfrac{k_1+k_2}{m_2}\right)y_2=0\end{cases}\tag{5.4}$$

若记
$$\boldsymbol{A}=\begin{pmatrix}\dfrac{k_1}{m_1} & -\dfrac{k_1}{m_1}\\ \dfrac{-k_1}{m_2} & \dfrac{k_1+k_2}{m_2}\end{pmatrix},\quad \boldsymbol{y}=\begin{pmatrix}y_1\\ y_2\end{pmatrix},\quad \lambda=\omega^2$$

则式(5.4)的矩阵形式就是

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{y} = \boldsymbol{0} \tag{5.5}$$

其中 $\boldsymbol{I}$ 为 2×2 单位矩阵.

实际上,式(5.5)是一个齐次线性代数方程组,它有非零解的充要条件是 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$. 这样,求角频率 ω 及振幅向量 $\boldsymbol{y}$ 的问题就归结为求满足方程

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0 \tag{5.6}$$

的根 λ 及对应于每个 λ,求满足

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{y} = \boldsymbol{0} \tag{5.7}$$

的非零向量 $\boldsymbol{y}$ 的问题.

在其它机械振动以及电网和控制系统等许多应用领域经常出现类似于上面的数学模型和求解问题. 满足(5.6)的 λ 称为方阵 $\boldsymbol{A}$ 的特征值,满足(5.7)的非零向量 $\boldsymbol{y}$ 称为方阵 $\boldsymbol{A}$ 的对应于特征值 λ 的特征向量.

方阵的特征值和特征向量本身在工程技术中常具有实际的物理意义,因此,求方阵的特征值问题,不论在实用上还是在理论上都是很重要的.

5.2.2 方阵的特征值和特征向量

定义 5.4 设 $\boldsymbol{A}$ 为 n 阶方阵,如果数 λ 和 n 维非零向量 $\boldsymbol{x}$ 使

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \tag{5.8}$$

成立,则称数 λ 为方阵 $\boldsymbol{A}$ 的**特征值**,非零向量 $\boldsymbol{x}$ 称为方阵 $\boldsymbol{A}$ 的对应于特征值 λ 的**特征向量**. (5.8)又可写成

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \tag{5.9}$$

这是一个 n 阶齐次线性方程组,它有非零解的充要条件是

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0 \tag{5.10}$$

显然(5.10)是一个以 λ 为未知数的一元 n 次方程,称为方阵

$\boldsymbol{A}$ 的**特征方程**，方程的左端称为方阵 $\boldsymbol{A}$ 的**特征多项式**. 由于一元 n 次方程在复数范围内有 n 个根，即矩阵 $\boldsymbol{A}$ 的 n 个特征值（重根按重数计算），并设它们为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，于是有

$$\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix} = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n) \tag{5.11}$$

称集合 $\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$ 为方阵 $\boldsymbol{A}$ 的**谱**，记为 $\lambda(\boldsymbol{A})$.

设 $\boldsymbol{A}$ 为 n 阶矩阵，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 为特征方程 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$ 的根，即 $\boldsymbol{A}$ 的 n 个特征值，那么

(1) $\det\boldsymbol{A}=\lambda_1\lambda_2\cdots\lambda_n$.

事实上，只要令 $\lambda=0$ 代入(5.11)式即得

$$\begin{aligned} \det(-\boldsymbol{A}) &= (-\lambda_1)(-\lambda_2)\cdots(-\lambda_n) \\ &= (-1)^n\det\boldsymbol{A} = (-1)^n\lambda_1\lambda_2\cdots\lambda_n \end{aligned}$$

即
$$\det\boldsymbol{A}=\lambda_1\lambda_2\cdots\lambda_n$$

这表明，方阵 $\boldsymbol{A}$ 的行列式等于它的特征值的乘积，于是有结论：**$\boldsymbol{A}$ 为奇异矩阵的充要条件是它至少有一个零特征值.**

(2) $\mathrm{tr}(\boldsymbol{A})=\lambda_1+\lambda_2+\cdots+\lambda_n$（其中 $\mathrm{tr}(\boldsymbol{A})$ 表示方阵 $\boldsymbol{A}$ 的迹）.

事实上，在等式(5.11)中，考察两边 λ^{n-1} 的系数：在等式左边，只有主对角线元素的代数余子式中才出现 λ^{n-1}，比如，$\lambda-a_{11}$ 与其代数余子式的乘积中 λ^{n-1} 的系数是 $-a_{11}$，$\lambda-a_{22}$ 与其代数余子式的乘积中 λ^{n-1} 的系数是 $-a_{22}$，如此等等，因此，等式左边 λ^{n-1} 的系数为 $-(a_{11}+a_{22}+\cdots+a_{nn})$，而等式右边 λ^{n-1} 的系数为 $-(\lambda_1+\lambda_2+\cdots+\lambda_n)$，由此得

$$a_{11}+a_{22}+\cdots+a_{nn}=\lambda_1+\lambda_2+\cdots+\lambda_n.$$

例 5.4 求 $\boldsymbol{A}=\begin{pmatrix} 2 & 0 \\ -1 & 4 \end{pmatrix}$ 的特征值和特征向量.

解 $\boldsymbol{A}$ 的特征方程为

$$|\lambda \boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2 & 0\\ 1 & \lambda-4\end{vmatrix}=(\lambda-2)(\lambda-4)=0$$

所以 $\boldsymbol{A}$ 的特征值 $\lambda_1=4,\lambda_2=2$.

对 $\lambda_1=2$,由 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$,所以对应的特征向量可取为

$$\boldsymbol{p}_1=(2,1)^{\mathrm{T}}$$

对 $\lambda_2=4$,对应的特征向量应满足 $(4\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$,所以对应的特征向量可取为

$$\boldsymbol{p}_2=(0,1)^{\mathrm{T}}$$

根据齐次线性方程组解的性质,若 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$,那么由于 $\boldsymbol{A}(k\boldsymbol{x})=\lambda(k\boldsymbol{x})$,所以 $k\boldsymbol{x}(k\neq 0)$ 也是 $\boldsymbol{A}$ 的对应于特征值 λ 的特征向量. 这表明,特征向量不能由特征值唯一确定,反过来,对不同的特征值,我们有下列结论.

定理 5.5　方阵 $\boldsymbol{A}$ 的对应于不同特征值的特征向量是线性无关的.

证　设 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_k$ 为 $\boldsymbol{A}$ 的依次对应于互不相同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 的特征向量. 假定有 k 个常数 $x_1,x_2,\cdots,x_k$ 使

$$x_1\boldsymbol{p}_1+x_2\boldsymbol{p}_2+\cdots+x_k\boldsymbol{p}_k=\boldsymbol{0}$$

则 $\boldsymbol{A}(x_1\boldsymbol{p}_1+\cdots+x_k\boldsymbol{p}_k)=0$,由 $\boldsymbol{A}\boldsymbol{p}_i=\lambda_i\boldsymbol{p}_i$ 得

$$\lambda_1x_1\boldsymbol{p}_1+\lambda_2x_2\boldsymbol{p}_2+\cdots+\lambda_kx_k\boldsymbol{p}_k=\boldsymbol{0}$$

以此类推便有

$$\lambda_1^m x_1\boldsymbol{p}_1+\lambda_2^m x_2\boldsymbol{p}_2+\cdots+\lambda_k^m x_k\boldsymbol{p}_k=\boldsymbol{0},\quad m=0,1,\cdots,k-1$$

将其写成矩阵形式有

$$(x_1\boldsymbol{p}_1,x_1\boldsymbol{p}_2,\cdots,x_k\boldsymbol{p}_k)\begin{pmatrix}1 & \lambda_1 & \cdots & \lambda_1^{k-1}\\ 1 & \lambda_2 & \cdots & \lambda_2^{k-1}\\ \vdots & \vdots & & \vdots\\ 1 & \lambda_k & \cdots & \lambda_k^{k-1}\end{pmatrix}=(\boldsymbol{0},\boldsymbol{0},\cdots,\boldsymbol{0})$$

上式等号左端第二个矩阵的行列式为范德蒙行列式 V,当 λ_i 互不相同时,$V\neq 0$,因此有

$$(x_1\boldsymbol{p}_1,x_2\boldsymbol{p}_2,\cdots,x_k\boldsymbol{p}_k)=(\boldsymbol{0},\cdots,\boldsymbol{0})$$

即 $x_i\boldsymbol{p}_i=\boldsymbol{0}(i=1,2,\cdots,k)$,而由 $\boldsymbol{p}_i\neq\boldsymbol{0}$ 得知 $x_i=0,i=1,2,\cdots,k$,

从而 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_k$ 线性无关.

当方阵 $\boldsymbol{A}$ 具有重特征值时，对应于这些重根的特征向量会怎样呢？请看下面的两个例子.

例 5.5 求下列矩阵 $\boldsymbol{A}$ 的特征值和特征向量

$$\boldsymbol{A} = \begin{pmatrix} -1 & -2 & 0 \\ 2 & 3 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

解 $\boldsymbol{A}$ 的特征方程为

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda+1 & 2 & 0 \\ -2 & \lambda-3 & 0 \\ -2 & 0 & \lambda-2 \end{vmatrix} = (\lambda-1)^2(\lambda-2) = 0$$

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = \lambda_2 = 1, \lambda_3 = 2$.

对 $\lambda_1 = \lambda_2 = 1$，解方程 $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 得基础解系

$$\boldsymbol{p}_1 = (-1, 1, 2)^{\mathrm{T}}$$

它就是对应于 $\lambda_1 = \lambda_2 = 1$ 的全部线性无关的特征向量.

对 $\lambda_3 = 2$，解方程 $(2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 得基础解系

$$\boldsymbol{p}_2 = (0, 0, 1)^{\mathrm{T}}$$

它是对应于 $\boldsymbol{\lambda}_3 = 2$ 的特征向量.

例 5.6 求下列矩阵 $\boldsymbol{A}$ 的特征值和特征向量

$$\boldsymbol{A} = \begin{pmatrix} -2 & 2 & 2 \\ 0 & 2 & 0 \\ -2 & 1 & 3 \end{pmatrix}$$

解

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda+2 & -2 & -2 \\ 0 & \lambda-2 & 0 \\ 2 & -1 & \lambda-3 \end{vmatrix} = (\lambda+1)(\lambda-2)^2$$

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = -1, \lambda_2 = \lambda_3 = 2$.

对 $\lambda_1 = -1$，解方程 $(-\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 得基础解系

$$\boldsymbol{p}_1 = (2, 0, 1)^{\mathrm{T}}$$

它是对应于 $\lambda_1 = -1$ 的一个特征向量.

对 $\lambda_2=\lambda_3=2$，解方程 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 得基础解系

$$\boldsymbol{p}_2=(1,2,0)^{\mathrm{T}},\boldsymbol{p}_3=(1,0,2)^{\mathrm{T}}$$

它们是对应于 $\boldsymbol{\lambda}_2=\boldsymbol{\lambda}_3=2$ 的两个线性无关的特征向量.

由这两个例子可见，方阵对应于它的重特征值的线性无关的特征向量的最大个数不多于重特征值的重数，这对任何方阵都是对的.

定理 5.6　设 λ_0 是 n 阶矩阵 $\boldsymbol{A}$ 的 k 重特征值，则对应于 λ_0 的线性无关的特征向量最大个数 $l\leqslant k$.

证　用反证法. 假设 $l>k$，并设 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_l$ 为 $\boldsymbol{A}$ 的对应于 λ_0 的线性无关的特征向量，即 $\boldsymbol{A}\boldsymbol{p}_i=\lambda_0\boldsymbol{p}_i$，$i=1,\cdots,l$，将 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_l$ 扩充为 n 维向量空间 R^n 的一组基：$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_l,\boldsymbol{p}_{l+1},\cdots,\boldsymbol{p}_n$，由于 $\boldsymbol{A}\boldsymbol{p}_m\in R^n$，所以

$$\boldsymbol{A}\boldsymbol{p}_m=c_{1m}\boldsymbol{p}_1+\cdots+c_{lm}\boldsymbol{p}_l+c_{l+1,m}\boldsymbol{p}_{l+1}+\cdots+c_{nm}\boldsymbol{p}_n$$
$$(m=l+1,l+2,\cdots,n)$$

令 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_l,\boldsymbol{p}_{l+1},\cdots,\boldsymbol{p}_n)$，则矩阵 $\boldsymbol{P}$ 可逆，记

$$\boldsymbol{A}_1=\begin{pmatrix}c_{1,l+1} & \cdots & c_{1n}\\ \vdots & & \vdots\\ c_{l,l+1} & \cdots & c_{ln}\end{pmatrix},\quad \boldsymbol{A}_2=\begin{pmatrix}c_{l+1,l+1} & \cdots & c_{l+1,n}\\ \vdots & & \vdots\\ c_{n,l+1} & \cdots & c_{nn}\end{pmatrix}$$

则 $\boldsymbol{AP}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{\lambda}_0\boldsymbol{I}_l & \boldsymbol{A}_1\\ \boldsymbol{O} & \boldsymbol{A}_2\end{pmatrix}$，$\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{\lambda}_0\boldsymbol{I}_l & \boldsymbol{A}_1\\ \boldsymbol{O} & \boldsymbol{A}_2\end{pmatrix}\boldsymbol{P}^{-1}$，从而

$$|\lambda\boldsymbol{I}_n-\boldsymbol{A}|=\begin{vmatrix}(\lambda-\lambda_0)\boldsymbol{I}_l & -\boldsymbol{A}_1\\ \boldsymbol{O} & \lambda\boldsymbol{I}_{n-l}-\boldsymbol{A}_2\end{vmatrix}=(\lambda-\lambda_0)^l|\lambda\boldsymbol{I}_{n-l}-\boldsymbol{A}_2|$$

这表明 λ_0 至少为矩阵 $\boldsymbol{A}$ 的 l 重特征值，与 λ_0 为 $\boldsymbol{A}$ 的 k 重特征值矛盾，因此 $l\leqslant k$.

这样，对有些 n 阶矩阵，对应于每个重特征值的线性无关的特征向量个数等于重特征值的重数，从而它就有 n 个线性无关的特征向量，这种矩阵称为**非亏损矩阵**；而对有些 n 阶矩阵，对应于某个重特征值的线性无关的特征向量的个数小于重数，从而它就没有 n 个线性无关的特征向量，这种矩阵称为**亏损矩阵**. 如例 5.6 中的 $\boldsymbol{A}$ 就是非亏损矩阵，而例 5.5 中的 $\boldsymbol{A}$ 就是亏损矩阵.

*5.3 求矩阵特征值的数值方法

求矩阵的特征值就是求特征方程的根，而高次方程的求解是相当困难的，为此，本节介绍两种求矩阵特征值的数值方法—**幂法**和**反幂法**.

5.3.1 幂法

幂法是通过求矩阵的特征向量来求矩阵的按模最大的特征值的一种迭代方法.

为叙述简便，以下凡提到最大或最小特征值都指按模最大或最小，如果有 n 个特征值，那么规定由大到小的顺序是

$$|\lambda_1|\geqslant|\lambda_2|\geqslant\cdots\geqslant|\lambda_n|.$$

幂法的基本思想是，欲求 n 阶方阵 $\boldsymbol{A}$ 的最大特征值 λ_1，从任意 n 维非零向量 $\boldsymbol{x}^{(0)}$ 出发，构造迭代序列

$$\boldsymbol{x}^{(0)},\boldsymbol{x}^{(1)}=\boldsymbol{A}\boldsymbol{x}^{(0)},\boldsymbol{x}^{(2)}=\boldsymbol{A}\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(k)}=\boldsymbol{A}\boldsymbol{x}^{(k-1)},\cdots$$

可以证明，当 k 充分大时，$\boldsymbol{x}^{(k+1)}\approx\lambda_1\boldsymbol{x}^{(k)}$，即相邻的两个迭代向量成正比例，比例系数 λ_1 即为所求之最大特征值.

事实上，若设 $\boldsymbol{A}$ 是 n 阶非亏损矩阵，并假定最大特征值 λ_1 为单实根，即 $|\lambda_1|>|\lambda_i|$，$i=2,3,\cdots,n$，$\boldsymbol{A}$ 的 n 个特征值对应的 n 个线性无关的特征向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$ 构成 n 维向量空间的一组基，于是对任何非零初始向量 $\boldsymbol{x}^{(0)}$，有

$$\boldsymbol{x}^{(0)}=a_1\boldsymbol{v}_1+a_2\boldsymbol{v}_2+\cdots+a_n\boldsymbol{v}_n$$

而

$$\begin{aligned}
\boldsymbol{x}^{(1)}=\boldsymbol{A}\boldsymbol{x}^{(0)}&=a_1\boldsymbol{A}\boldsymbol{v}_1+a_2\boldsymbol{A}\boldsymbol{v}_2+\cdots+a_n\boldsymbol{A}\boldsymbol{v}_n\\
&=a_1\lambda_1\boldsymbol{v}_1+a_2\lambda_2\boldsymbol{v}_2+\cdots+a_n\lambda_n\boldsymbol{v}_n\\
\boldsymbol{x}^{(2)}=\boldsymbol{A}\boldsymbol{x}^{(1)}&=a_1\lambda_1\boldsymbol{A}\boldsymbol{v}_1+a_2\lambda_2\boldsymbol{A}\boldsymbol{v}_2+\cdots+a_n\lambda_n\boldsymbol{A}\boldsymbol{v}_n\\
&=a_1\lambda_1^2\boldsymbol{v}_1+a_2\lambda_2^2\boldsymbol{v}_2+\cdots+a_n\lambda_n^2\boldsymbol{v}_n\\
&\ \ \vdots\\
\boldsymbol{x}^{(k)}&=a_1\lambda_1^k\boldsymbol{v}_1+a_2\lambda_2^k\boldsymbol{v}_2+\cdots+a_n\lambda_n^k\boldsymbol{v}_n
\end{aligned}$$

$$\vdots$$

设 $a_1 \neq 0$，由于 $\left|\dfrac{\lambda_i}{\lambda_1}\right| < 1, i=2,3,\cdots,n$，故当 k 充分大时，

$$\boldsymbol{x}^{(k)} = \lambda_1^k \left(a_1 \boldsymbol{v}_1 + a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \boldsymbol{v}_2 + \cdots + a_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \boldsymbol{v}_n\right) \approx \lambda_1^k a_1 \boldsymbol{v}_1$$

同理 $\boldsymbol{x}^{(k+1)} \approx \lambda_1^{k+1} a_1 \boldsymbol{v}_1$

由此 $\boldsymbol{x}^{(k+1)} \approx \lambda_1 \boldsymbol{x}^{(k)}$

即 $\lambda_1 = \dfrac{x_i^{(k+1)}}{x_i^{(k)}}, \quad i=1,2,\cdots,n.$

应当说明，在上述推证中假定了 $a_1 \neq 0$，但在实际计算中，这个假定不是必要的. 因为即便选取的 $\boldsymbol{x}^{(0)}$ 的第一个坐标 $a_1 = 0$，由于舍入误差的影响，经若干步后，向量的第一个坐标就不再是零，继续迭代就可认为是以这个向量为初始向量，它的 $a_1 \neq 0$. 因此，在数值计算中不必考虑 $\boldsymbol{x}^{(0)}$ 的第一个坐标 a_1 是否为零.

在应用幂法计算时，为了避免计算过程中出现绝对值过大或过小的数 x_i，以防止在计算机中出现溢出现象，通常在每步迭代时首先将向量 $\boldsymbol{x}^{(k)}$ 规范化为 $\boldsymbol{y}^{(k)} = \boldsymbol{x}^{(k)} / \|\boldsymbol{x}^{(k)}\|_\infty$，用 $\boldsymbol{y}^{(k)}$ 作为新的迭代向量. 因此，实际计算的迭代公式是

$$\begin{cases} \boldsymbol{y}^{(k)} = \boldsymbol{x}^{(k)} / \|\boldsymbol{x}^{(k)}\|_\infty \\ \boldsymbol{x}^{(k+1)} = \boldsymbol{A}\boldsymbol{y}^{(k)}, \quad k = 0,1,2,\cdots \end{cases} \tag{5.12}$$

当 k 充分大时，

$$\begin{cases} \boldsymbol{y}^{(k)} \approx \boldsymbol{v}_1 \\ \|\boldsymbol{x}^{(k+1)}\|_\infty \approx |\lambda_1| \end{cases} \tag{5.13}$$

例 5.7 求矩阵 $\boldsymbol{A}$ 按模最大的特征值 λ_1 和特征向量 $\boldsymbol{v}_1$.

$$\boldsymbol{A} = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}$$

解 取 $\boldsymbol{x}^{(0)} = (1,0,0,0)^{\mathrm{T}}$，按公式(5.13)迭代

$$x^{(1)}=Ax^{(0)}=\begin{pmatrix}10&7&8&7\\7&5&6&5\\8&6&10&9\\7&5&9&10\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}=\begin{pmatrix}10\\7\\8\\7\end{pmatrix}=10\begin{pmatrix}1\\0.7\\0.8\\0.7\end{pmatrix}=10y^{(1)}$$

$$x^{(2)}=Ay^{(1)}=\begin{pmatrix}10&7&8&7\\7&5&6&5\\8&6&10&9\\7&5&9&10\end{pmatrix}\begin{pmatrix}1\\0.7\\0.8\\0.7\end{pmatrix}=\begin{pmatrix}26.2\\18.8\\26.5\\24.7\end{pmatrix}=26.5y^{(2)}$$

连续迭代 5 次的结果列于表 5.1.

矩阵 A 的按模最大的特征值 $\lambda=30.292\,68$,对应的特征向量 $v_1=(0.95769,0.68898,1.00000,0.94376)^{\mathrm{T}}$.

表 5.1

k	λ_1	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
0		1.000 00	0.000 00	0.000 00	0.000 00
1	10.000 00	1.000 00	0.700 00	0.800 00	0.700 00
2	26.500 00	0.988 68	0.709 43	1.000 00	0.932 08
3	30.554 72	0.961 47	0.691 49	1.000 00	0.942 20
4	30.320 49	0.958 11	0.689 26	1.000 00	0.943 58
5	30.292 68	0.957 69	0.688 98	1.000 00	0.943 76

5.3.2 反幂法

反幂法是求矩阵按模最小的特征值和特征向量的迭代法.

设 A 是非奇异矩阵,即 A 的一切特征值 $\lambda_i\neq 0$,则 A^{-1} 的特征值为$\dfrac{1}{\lambda_i}$,$i=1,2,\cdots,n$,(见习题五,18),由于 $|\lambda_1|\geqslant|\lambda_2|\geqslant\cdots\geqslant|\lambda_n|$,所以 $\left|\dfrac{1}{\lambda_1}\right|\leqslant\left|\dfrac{1}{\lambda_2}\right|\leqslant\cdots\leqslant\left|\dfrac{1}{\lambda_n}\right|$,那么 A^{-1} 的最大特征值$\dfrac{1}{\lambda_n}$的倒数即为 A 的最小特征值 λ_n,因此,将幂法用于 A^{-1} 求出它的最大特征值 即得 A 的最小特征值,这就是反幂法.

取非零初始向量 $\boldsymbol{x}^{(0)}$，按迭代公式

$$\begin{cases}\boldsymbol{y}^{(k)} = \boldsymbol{x}^{(k)} / \| \boldsymbol{x}^{(k)} \|_\infty \\ \boldsymbol{x}^{(k+1)} = \boldsymbol{A}^{-1} \boldsymbol{y}^{(k)}\end{cases}$$

$$k = 0,1,2,\cdots \tag{5.14}$$

算出 $\boldsymbol{x}^{(k+1)}$，最后按公式（k 充分大时）

$$\begin{cases}\boldsymbol{y}^{(k+1)} \approx \boldsymbol{v}_n \\ \| \boldsymbol{x}^{(k+1)} \|_\infty \approx | \mu_n |\end{cases} \tag{5.15}$$

得到 $\boldsymbol{A}$ 的最小特征值 $|\lambda_n| = \dfrac{1}{|\mu_n|}$ 及对应的特征向量 $\boldsymbol{v}_n = \boldsymbol{y}^{(k+1)}$.

应用反幂法必须首先求出 $\boldsymbol{A}^{-1}$，而矩阵求逆是很困难的，为此，可将迭代式(5.14)改写为

$$\boldsymbol{A}\boldsymbol{x}^{(k+1)} = \boldsymbol{y}^{(k+1)} \tag{5.16}$$

求解这个 n 阶方程组即得 $\boldsymbol{x}^{(k+1)}$，这里避免了求 $\boldsymbol{A}^{-1}$. 由于在整个迭代过程中不断求解(5.16)，这个方程组的右端 $\boldsymbol{y}^{(k)}$ 不断变化，而左端系数矩阵 $\boldsymbol{A}$ 始终不变，所以采用 $\boldsymbol{LU}$ 分解法解(5.16)是最方便的解法. 于是首先作分解 $\boldsymbol{A}=\boldsymbol{LU}$，这样，反幂法的迭代公式为

$$\begin{cases}\boldsymbol{y}^{(k)} = \boldsymbol{x}^{(k)} / \| \boldsymbol{x}^{(k)} \|_\infty \\ \boldsymbol{L}\boldsymbol{z} = \boldsymbol{y}^{(k)} \\ \boldsymbol{U}\boldsymbol{x}^{(k+1)} = \boldsymbol{z}\end{cases} \quad k = 0,1,2,\cdots \tag{5.17}$$

例 5.8 求矩阵

$$\boldsymbol{A} = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}$$

的所有特征值.

解 由例 5.7 知 $\lambda_1 = 30.2927$. 下面由反幂法求 λ_4.

首先作 $\boldsymbol{A}$ 的 $\boldsymbol{LU}$ 分解

$$\boldsymbol{A} = \boldsymbol{LU} = \begin{pmatrix} 1 & & & \\ 0.7 & 1 & & \\ 0.8 & 4 & 1 & \\ 0.7 & 1 & 1.5 & 1 \end{pmatrix} \begin{pmatrix} 10 & 7 & 8 & 7 \\ & 0.1 & 0.4 & 0.1 \\ & & 2 & 3 \\ & & & 0.5 \end{pmatrix}$$

取 $\boldsymbol{x}^{(0)}=(1,0,0,0)^{\mathrm{T}}$，按(5.17)迭代 3 次，结果列于表 5.2.

矩阵 $\boldsymbol{A}$ 的最小特征值 $\lambda_4=\dfrac{1}{\mu_4}=\dfrac{1}{98.52309}=0.01015$.

据特征多项式的性质，有

$$\begin{cases}\lambda_1+\lambda_2+\lambda_3+\lambda_4=\mathrm{tr}(\boldsymbol{A})=35\\ \lambda_1\lambda_2\lambda_3\lambda_4=\det\boldsymbol{A}=1\end{cases}$$

将求得的 λ_1 和 λ_2 代入并解方程组，得

$$\lambda_2=3.85308\qquad \lambda_3=0.84409$$

于是得到了矩阵 $\boldsymbol{A}$ 的全部特征值

$$\lambda_1=30.29268\qquad \lambda_2=3.85308,$$
$$\lambda_3=0.84409\qquad \lambda_4=0.01015$$

表 5.2

k	μ_4	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
0		1.000 00	0	0	0
1	20.000 30	−0.600 00	1.000 00	−0.250 00	0.150 00
2	98.351 49	−0.603 97	1.000 00	−0.251 14	0.148 96
3	98.523 09	−0.603 97	1.000 00	−0.251 14	0.148 95

5.4 相似矩阵与实对称矩阵的对角化

前面我们讨论了求方阵 $\boldsymbol{A}$ 的特征值和特征向量的方法，如果已经求得 n 阶方阵 $\boldsymbol{A}$ 的 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$，试问 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$ 有什么关系？是否所有的方阵 $\boldsymbol{A}$ 与其特征值构成的对角方阵 $\boldsymbol{\Lambda}$ 都有这样的关系？围绕这些问题并为了后面讨论二次型化标准形的需要，我们引进相似矩阵和合同矩阵的概念，并进而讨论实对称矩阵的对角化.

5.4.1 相似矩阵与合同矩阵

定义 5.5 对 n 阶方阵 $\boldsymbol{A}$、$\boldsymbol{B}$，如果存在可逆矩阵 $\boldsymbol{P}$，使

$$P^{-1}AP = B$$

则称 B 是 A 的**相似矩阵**,或称矩阵 A 与 B **相似**,而对 A 进行运算 $P^{-1}AP$ 称为对 A 进行**相似变换**,可逆矩阵 P 称为把 A 变成 B 的**相似变换矩阵.**

显然,若 A 与 B 相似,则 A 与 B 等价,反之不然;矩阵相似具有反身性、对称性和传递性.

定理 5.7 **若 n 阶方阵 A 与 B 相似,则 A 与 B 有相同的特征多项式,从而有相同的特征值.**

证 因 A 与 B 相似,即有 P 使 $P^{-1}AP=B$,所以

$$|\lambda I - B| = |\lambda I - P^{-1}AP| = |P^{-1}||\lambda I - A||P| = |\lambda I - A|.$$

这表明 A,B 有相同的特征多项式,从而有相同的特征值.

推论 **若 n 阶方阵 A 与对角矩阵 $\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$ 相似,则 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 A 的 n 个特征值.**

设 n 阶方阵 A 的 n 个特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$,$\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$,根据上述讨论,自然会提出下列问题:

(1)是否任何方阵 A 都与某个对角矩阵 Λ 相似?(2)若 A 与某个对角矩阵 Λ 相似,即若存在 P,使 $P^{-1}AP=\Lambda$,P 的构造如何?

我们先来讨论第二个问题.设存在可逆矩阵 P,使

$$P^{-1}AP = \Lambda = \mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

将 P 用其列向量表示为

$$P = (p_1,p_2,\cdots,p_n)$$

由 $P^{-1}AP=\Lambda$ 得 $AP=P\Lambda$,即

$$A(p_1,p_2,\cdots,p_n) = (p_1,p_2,\cdots,p_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$$

$$= (\lambda_1 p_1,\lambda_2 p_2,\cdots,\lambda_n p_n)$$

于是 $Ap_i=\lambda_i p_i$,$i=1,2,\cdots,n$.这说明,λ_i 是 A 的特征值,p_i 是 A 的对应于特征值 λ_i 的特征向量.这就是相似变换矩阵的具体构造.

现在讨论第一个问题. 由 5.2 节知道，n 阶方阵恰有 n 个特征值(可以有重根)，并可对应地求出 n 个特征向量(但未必线性无关)，这 n 个特征向量组成的矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{AP}=\boldsymbol{P\Lambda}$($\boldsymbol{P}$ 可能为复矩阵).

余下的问题是矩阵 $\boldsymbol{P}$ 是否可逆，即 $\boldsymbol{p}_1,\cdots,\boldsymbol{p}_n$ 是否线性无关. 因若 $\boldsymbol{P}$ 可逆，则 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}$，即 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似.

由上节定理 5.5 可知，若 $\boldsymbol{A}$ 的 n 个特征值均不相同，则它们对应的 n 个特征向量线性无关，因此，由 n 个特征向量组成的矩阵 $\boldsymbol{P}$ 可逆；若 $\boldsymbol{A}$ 有重特征值，且 $\boldsymbol{A}$ 是非亏损矩阵，则存在 n 个线性无关的特征向量，以它们为列可以组成可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}$；若 $\boldsymbol{A}$ 是亏损矩阵，则不存在 n 个线性无关的特征向量，所以不能使 $\boldsymbol{A}$ 与对角矩阵相似. 因此我们有

定理 5.8　方阵 $\boldsymbol{A}$ 与对角矩阵相似(或方阵 $\boldsymbol{A}$ 可对角化)的充要条件是 $\boldsymbol{A}$ 为非亏损矩阵.

不过，我们不对非亏损矩阵进行一般性的讨论，而仅讨论 $\boldsymbol{A}$ 为实对称矩阵(即 $\boldsymbol{A}$ 的元素全是实数，且 $\boldsymbol{A}$ 为对称矩阵)的情形. 这种情形比较简单，而且实际应用上较为常见.

定义 5.6　对 n 阶矩阵 $\boldsymbol{A}$、$\boldsymbol{B}$，若存在 n 阶可逆矩阵 $\boldsymbol{C}$，使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\boldsymbol{B}$，则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同.

显然，合同关系也具有反身性、对称性和传递性. $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同时，$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 一定等价.

5.4.2　实对称矩阵的对角化

定理 5.9　实对称矩阵的特征值为实数.

证　设实对称矩阵 $\boldsymbol{A}$ 的特征值为复数 λ，对应的特征向量为复向量 $\boldsymbol{x}$，则 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$，两端取共轭，则有 $\boldsymbol{A}\bar{\boldsymbol{x}}=\overline{\lambda\boldsymbol{x}}$，于是

$$\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{Ax}=\bar{\boldsymbol{x}}^{\mathrm{T}}(\boldsymbol{Ax})=\bar{\boldsymbol{x}}^{\mathrm{T}}\lambda\boldsymbol{x}=\lambda\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}$$

$$\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{Ax}=(\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x}=(\boldsymbol{A}\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}=\overline{\lambda\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}$$

两式相减得

$$(\lambda-\bar{\lambda})\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}=0$$

因 $\boldsymbol{x}\neq\boldsymbol{0}$,所以

$$\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}=\sum_{i=1}^{n}\bar{x}_i x_i=\sum_{i=1}^{n}|x_i|^2\neq 0$$

故有 $\lambda-\bar{\lambda}=0$. 这表明 λ 为实数.

显然,对实特征值 λ,$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 为实系数的齐次线性方程组,由 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$ 知其必有实的基础解系,从而相应地,可取实的特征向量.

对于实对称矩阵,属于不同的特征值的特征向量不仅线性无关,而且还有更强的结论,这就是

定理 5.10 设 λ_1,λ_2 是实对称矩阵 $\boldsymbol{A}$ 的两个特征值,$\boldsymbol{p}_1,\boldsymbol{p}_2$ 为对应的特征向量. 若 $\lambda_1\neq\lambda_2$ 则 $\boldsymbol{p}_1,\boldsymbol{p}_2$ 正交.

证 $\lambda_1\boldsymbol{p}_1=\boldsymbol{A}\boldsymbol{p}_1,\lambda_2\boldsymbol{p}_2=\boldsymbol{A}\boldsymbol{p}_2,\lambda_1\neq\lambda_2$

由于 $\boldsymbol{A}$ 对称,所以

$$\lambda_1\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{p}_2=(\lambda_1\boldsymbol{p}_1)^{\mathrm{T}}\boldsymbol{p}_2=(\boldsymbol{A}\boldsymbol{p}_1)^{\mathrm{T}}\boldsymbol{p}_2=\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_2=\boldsymbol{p}_1^{\mathrm{T}}\lambda_2\boldsymbol{p}_2=\lambda_2\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{p}_2$$

于是$(\lambda_1-\lambda_2)\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{p}_2=0$. 但 $\lambda_1\neq\lambda_2$,故有 $\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{p}_2=0$,即 $\boldsymbol{p}_1,\boldsymbol{p}_2$ 正交.

定理 5.11 设 $\boldsymbol{A}$ 为 n 阶实对称方阵,则必有正交矩阵 $\boldsymbol{P}$,使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值.

证 任取 $\boldsymbol{A}$ 的一个特征值 λ_1 及对应的单位特征向量 $\boldsymbol{p}_1$,由定理 5.2 的推论知,可找到 $n-1$ 个单位列向量 $\boldsymbol{p}_2,\boldsymbol{p}_3,\cdots,\boldsymbol{p}_n$,使 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 为正交向量组. 因此

$$\boldsymbol{P}_1=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)$$

为正交矩阵,因为

$$\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_1=\boldsymbol{p}_1^{\mathrm{T}}\lambda_1\boldsymbol{p}_1=\lambda_1\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{p}_1=\lambda_1$$

$$\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_i=(\boldsymbol{p}_i^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_1)^{\mathrm{T}}=(\boldsymbol{p}_i^{\mathrm{T}}\lambda_1\boldsymbol{p}_1)^{\mathrm{T}}=(\lambda_1\boldsymbol{p}_i^{\mathrm{T}}\boldsymbol{p}_1)^{\mathrm{T}}=0$$

$$(i=2,3,\cdots,n)$$

所以

$$\boldsymbol{P}_1^{-1}\boldsymbol{A}\boldsymbol{P}_1=\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1=\begin{pmatrix}\boldsymbol{p}_1^{\mathrm{T}}\\\boldsymbol{p}_2^{\mathrm{T}}\\\vdots\\\boldsymbol{p}_n^{\mathrm{T}}\end{pmatrix}\boldsymbol{A}(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)$$

$$=\begin{pmatrix} \boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_1 & \boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_2 & \cdots & \boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_n \\ \boldsymbol{p}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_1 & \boldsymbol{p}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_2 & \cdots & \boldsymbol{p}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_n \\ \vdots & \vdots & & \vdots \\ \boldsymbol{p}_n^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_1 & \boldsymbol{p}_n^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_2 & \cdots & \boldsymbol{p}_n^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_n \end{pmatrix}$$

$$=\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \boldsymbol{p}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_2 & \cdots & \boldsymbol{p}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_n \\ \vdots & \vdots & & \vdots \\ 0 & \boldsymbol{p}_n^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_2 & \cdots & \boldsymbol{p}_n^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_1 \end{pmatrix}$$

其中 $\boldsymbol{A}_1$ 为上述矩阵中删去第一行、第一列所余下的子块. 显然，$\boldsymbol{A}_1$ 是 $n-1$ 阶的实对称矩阵.

对 $\boldsymbol{A}_1$ 重复上面的作法，可找到 $n-1$ 阶正交矩阵 $\boldsymbol{Q}$，使

$$\boldsymbol{Q}^{-1}\boldsymbol{A}_1\boldsymbol{Q} = \begin{pmatrix} \lambda_2 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_2 \end{pmatrix}$$

其中 λ_2 是 $\boldsymbol{A}_1$ 的特征值，$\boldsymbol{A}_2$ 是 $n-2$ 阶实对称矩阵. 令

$$\boldsymbol{P}_2 = \begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Q} \end{pmatrix}$$

显然 $\boldsymbol{P}_2$ 也是正交矩阵，由分块矩阵乘法得

$$\boldsymbol{P}_2^{-1}(\boldsymbol{P}_1^{-1}\boldsymbol{A}\boldsymbol{P}_1)\boldsymbol{P}_2 = \begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Q}^{-1} \end{pmatrix}\begin{pmatrix} \lambda_1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_1 \end{pmatrix}\begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Q} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Q}^{-1}\boldsymbol{A}_1\boldsymbol{Q} \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \boldsymbol{A}_2 \end{pmatrix}$$

如此继续下去，可得到正交矩阵 $\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_{n-1}$ 使

$$\boldsymbol{P}_{n-1}^{-1}\cdots\boldsymbol{P}_1^{-1}\boldsymbol{A}\boldsymbol{P}_1\cdots\boldsymbol{P}_{n-1} = \mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

令 $\boldsymbol{P}=\boldsymbol{P}_1\cdots\boldsymbol{P}_{n-1}$，由定理 5.3(4)知 P 仍为正交矩阵，并有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

再由定理 5.7 的推论知 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值.

这个定理说明，任何实对称矩阵 $\boldsymbol{A}$ 都能对角化为对角矩阵，且对角矩阵的主对角线元素就是 $\boldsymbol{A}$ 特征值，同时说明实对称矩阵

$\boldsymbol{A}$ 是非亏损矩阵.

例 5.9 设

$$\boldsymbol{A}=\begin{pmatrix}4&0&0\\0&3&1\\0&1&3\end{pmatrix}$$

求一正交矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$.

解

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-4&0&0\\0&\lambda-3&-1\\0&-1&\lambda-3\end{vmatrix}=(\lambda-2)(\lambda-4)^2$$

A 的特征值为 $\lambda_1=2,\lambda_2=\lambda_3=4$.

对 $\lambda_1=2$,解 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 得基础解系 $\boldsymbol{\xi}_1=(0,1,-1)^{\mathrm{T}}$,规范化得

$$\boldsymbol{p}_1=\left(0,\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$$

当 $\lambda_2=\lambda_3=4$ 时,解方程 $(4\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 得基础解系

$$\boldsymbol{\xi}_2=(1,0,0)^{\mathrm{T}},\quad \boldsymbol{\xi}_3=(0,1,1)^{\mathrm{T}}$$

由于 $\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 恰好正交,所以只要规范化为

$$\boldsymbol{p}_2=(1,0,0)^{\mathrm{T}},\quad \boldsymbol{p}_3=\left(0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$$

因此

$$\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3)=\begin{pmatrix}0&1&0\\\frac{1}{\sqrt{2}}&0&\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}&0&\frac{1}{\sqrt{2}}\end{pmatrix}$$

并且

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(2,4,4)$$

由这个例子可见,对实对称矩阵 $\boldsymbol{A}$,求一个正交矩阵 $\boldsymbol{P}$,使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ 的步骤如下:

第一步　求 $\boldsymbol{A}$ 的特征值；

第二步　求对应于每个特征值的特征向量．对单特征值，只需将属于它的特征向量规范化；对 r 重特征值，需要求出属于它的 r 个线性无关的特征向量，并对这 r 个特征向量进行正交规范化，这样就得到 n 个两两正交的单位特征向量；

第三步　以正交规范化的特征向量为列组成矩阵，它就是要求的正交矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$，这时 $\boldsymbol{\Lambda}$ 的主对角线元素只需按组成 $\boldsymbol{P}$ 时特征向量的顺序依次将它们所属的特征值排列即可．

最后顺便指出，要求的矩阵 $\boldsymbol{P}$ 不是唯一的，比如在例 5.9 中，对应于 $\lambda_1=2$ 的单位特征向量可取成

$$\boldsymbol{p}_1=\left(0,-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$$

而对应于重特征值 $\lambda_2=\lambda_3=4$ 的两个线性无关的特征向量可取为

$$\boldsymbol{\xi}_2=(1,1,1)^{\mathrm{T}},\quad \boldsymbol{\xi}_3=(-1,1,1)^{\mathrm{T}}$$

由于 $\boldsymbol{\xi}_2$ 与 $\boldsymbol{\xi}_3$ 不正交，所以需先正交化，取

$$\boldsymbol{\eta}_2=\boldsymbol{\xi}_2$$

$$\boldsymbol{\eta}_3=\boldsymbol{\xi}_3-\frac{[\boldsymbol{\eta}_2,\boldsymbol{\xi}_3]}{[\boldsymbol{\eta}_2,\boldsymbol{\eta}_2]}\boldsymbol{\eta}_2=\begin{pmatrix}-1\\1\\1\end{pmatrix}-\frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}=\frac{1}{3}\begin{pmatrix}-4\\2\\2\end{pmatrix}$$

再将 $\boldsymbol{\eta}_2,\boldsymbol{\eta}_3$ 规范化得

$$\boldsymbol{p}_2=\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)^{\mathrm{T}},\boldsymbol{p}_3=\left(-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}}\right)^{\mathrm{T}}$$

于是

$$\boldsymbol{P}=\begin{pmatrix}0&\frac{1}{\sqrt{3}}&-\frac{2}{\sqrt{6}}\\-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{6}}\end{pmatrix}$$

5.5　二次型及其标准形

5.5.1　二次型

在解析几何中，平面二次曲线

$$ax^2+bxy+cy^2=1$$

的左端是变量 x,y 的二次齐次多项式，而二次曲面

$$ax^2+bxy+cyz+dzx+ey^2+fz^2=1$$

的左端是变量 x,y,z 的二次齐次多项式. 这类二次齐次多项式通称为二次型或二次齐式.

为了便于研究二次曲线或二次曲面的几何性质，常常将坐标轴经过适当的旋转，把二次曲线、二次曲面化为只含有坐标平方项的所谓标准形

$$mx^2+ny^2=1$$

$$px^2+qy^2+rz^2=1$$

从代数学的观点看，这种化标准形的过程，就是通过变量的适当的线性变换化简一个二次型，使它只含变量的平方项. 这种问题不仅在理论上，而且在实际应用上，比如在气象统计、大气动力学、求多元函数的极值以及多元统计分析中经常遇到. 因此，需要对这个问题进行一般化的讨论.

定义 5.7　含有 n 个变量 $x_1,x_2,\cdots,x_n$ 的二次齐次函数

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)=&a_{11}x_1^2+a_{22}x_2^2+\cdots+a_{nn}x_n^2\\&+2a_{12}x_1x_2+2a_{13}x_1x_3+\cdots+2a_{(n-1)n}x_{n-1}x_n\end{aligned}\tag{5.18}$$

称为**二次型**.

对 $i\neq j$，若取 $a_{ji}=a_{ij}$，则

$$2a_{ij}x_ix_j=a_{ij}x_ix_j+a_{ji}x_jx_i$$

于是(5.18)式可写成

$$\begin{aligned}f=&x_1(a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n)\\&+x_2(a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n)\end{aligned}$$

$$+\cdots+x_n(a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n)$$

$$=(x_1,x_2,\cdots,x_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

$$=\sum_{i,j=1}^{n}a_{ij}x_ix_j$$

若记

$$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix},\quad \boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

则 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$,(5.18)式可写成

$$f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \tag{5.19}$$

(5.19)式称为二次型 f 的矩阵表示式.

例如,二次型 $f=x^2-4xy-3z^2+yz$ 用矩阵表示就是

$$f=(x,y,z)\begin{pmatrix}1&-2&0\\-2&0&\frac{1}{2}\\0&\frac{1}{2}&-3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$$

不难看出,任给一个二次型 f,就唯一地确定了一个对称矩阵 $\boldsymbol{A}$;反之,任给一个对称矩阵 $\boldsymbol{A}$,也由(5.19)式唯一地确定了一个二次型 f. 因此二次型与对称矩阵之间建立了一一对应关系. 这样,把对称矩阵 $\boldsymbol{A}$ 称为二次型 f 的矩阵,把 f 叫做对称矩阵 $\boldsymbol{A}$ 的二次型;对称矩阵 $\boldsymbol{A}$ 的秩就叫做**二次型的秩**. 如果一个二次型的矩阵是实对称矩阵,则称其为**实二次型**. 今后我们仅讨论实二次型.

对二次型,要讨论的主要问题是,寻求可逆的线性变换

$$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y} \tag{5.20}$$

其中 $\boldsymbol{C}=(c_{ij})_{n\times n}$，$\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$，$\boldsymbol{C}$ 可逆，使二次型只含平方项，即将(5.20)代入(5.19)式后，二次型变为

$$\begin{aligned} f &= \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{C}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{C}\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})\boldsymbol{y} \\ &= k_1y_1^2+k_2y_2^2+\cdots+k_ny_n^2 \end{aligned} \tag{5.21}$$

这种只含平方项的二次型称为二次型的**标准形**. 若 $k_i>0$，$(i=1,\cdots,r)$，$k_j<0$，$(j=r+1,\cdots,m)$，$k_t=0$，$(t=m+1,\cdots,n)$，如果记矩阵 $\boldsymbol{B}=\mathrm{diag}\left(\dfrac{1}{\sqrt{|k_1|}},\cdots,\dfrac{1}{\sqrt{|k_m|}},1,\cdots,1\right)$，再经变换 $\boldsymbol{y}=\boldsymbol{B}\boldsymbol{z}$，标准形变为 $f=z_1^2+\cdots+z_r^2-z_{r+1}^2\cdots-z_m^2$，它称为二次型的**规范形**.

由(5.21)可见，为使二次型 f 化为标准形，要寻求的可逆变换(5.20)中的 $\boldsymbol{C}$ 必须满足 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵，即

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C} = \mathrm{diag}(k_1,k_2,\cdots,k_n)$$

因此，寻求可逆变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 化二次型为标准形的问题就归结为寻求对称矩阵 $\boldsymbol{A}$ 的可逆变换矩阵 $\boldsymbol{C}$，使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵的问题.

5.5.2 用正交变换化二次型为标准形

由定理 5.10 知道，对任何实对称矩阵 $\boldsymbol{A}$，总有正交矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$，即 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ 为对角阵. 把这个结论用于二次型就有

定理 5.12　任给实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，总有正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$，使 f 化为标准形

$$f = \lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值.

可以证明，正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 不改变二次型的秩. 事实上，由于 $\boldsymbol{P}$ 可逆，所以它可以表示成一系初等方阵的乘积，即 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ 可以看成 $\boldsymbol{A}$ 经一系列初等变换得到，亦即 $\boldsymbol{A}$ 与 $\boldsymbol{\Lambda}$ 等价，因此 $R(\boldsymbol{A})=R(\boldsymbol{\Lambda})$. 其实，这个结论对任何可逆变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 也是成立的，即若 $\boldsymbol{C}$ 可逆，$\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$，则 $R(\boldsymbol{A})=R(\boldsymbol{B})$.

例 5.10　求正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$，把二次型

$$f = 2x_1x_3+x_2^2$$

化为标准形.

解 二次型的矩阵

$$\boldsymbol{A}=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$$

它的特征多项式为

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda&0&-1\\0&\lambda-1&0\\-1&0&\lambda\end{vmatrix}=(\lambda+1)(\lambda-1)^2$$

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=-1,\lambda_2=\lambda_3=1.$

对 $\lambda_1=-1$，解 $(-\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 得基础解系 $\boldsymbol{\xi}_1=(1,0,-1)^{\mathrm{T}}$，规范化得 $\boldsymbol{p}_1=\left(\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$.

当 $\lambda_2=\lambda_3=1$ 时，解方程 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 得基础解系

$$\boldsymbol{\xi}_2=(0,1,0)^{\mathrm{T}},\quad \boldsymbol{\xi}_3=(1,0,1)^{\mathrm{T}}$$

由于它们已正交，只需规范化

$$\boldsymbol{p}_2=(0,1,0)^{\mathrm{T}},\quad \boldsymbol{p}_3=\left(\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$$

所以正交变换为

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}}&0&\frac{1}{\sqrt{2}}\\0&1&0\\-\frac{1}{\sqrt{2}}&0&\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$$

f 的标准形为

$$f=-y_1^2+y_2^2+y_3^2.$$

由此例，我们可以得到求正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 化二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为标准形的步骤：

(1)写出二次型 f 的矩阵 $\boldsymbol{A}$；

(2)求 $\boldsymbol{A}$ 的特征值；

(3)求对应于特征值的特征向量. 对单特征值，需将属于它的

特征向量规范化；对 r 重的特征值，需将属于它的 r 个线性无关的特征向量正交规范化；

(4)以正交规范化的特征向量为列组成正交矩阵 $\boldsymbol{P}$，写出正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$；

(5)按组成 $\boldsymbol{P}$ 时特征向量的次序，以其所属特征值为系数写出标准形.

5.5.3 用配方法化二次型为标准形

把一个二次型化成标准形，如果不限于用正交变换，可以用多种方法，即有多个可逆变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 为对角阵，即使实对称矩阵 $\boldsymbol{A}$ 与对角矩阵合同. 在这一节里，只介绍化二次型为标准形的拉格朗日配方法.

这种方法的基本想法是：若二次型 f 中含 x_1^2 和 $x_1x_j(j\neq 1)$，则可先把含 x_1 的项归并起来配平方，使余下的项中不再含 x_1；若余下的项中含 x_2^2 和 $x_2x_j(j\neq 1,2)$，则类似地处理. 如此继续直至全配成平方项为止.

例 5.11 化二次型

$$f=x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3$$

为标准形，并求相应的变换矩阵.

解 由于 f 中含 x_1^2, x_1x_2, x_1x_3，所以先把它们归并起来配平方得

$$\begin{aligned} f &= x_1^2+2x_1x_2+2x_1x_3+2x_2^2+2x_3^2 \\ &= (x_1+x_2+x_3)^2-x_2^2-x_3^2-2x_2x_3+2x_2^2+2x_3^2 \\ &= (x_1+x_2+x_3)^2+x_2^2-2x_2x_3+x_3^2 \end{aligned}$$

显然除第一项外，不再含 x_1. 将含 x_2^2 和 x_2x_3 的项再归并配平方得

$$f=(x_1+x_2+x_3)^2+(x_2-x_3)^2$$

此时，f 中只有平方项，因此，令

$$\begin{cases} y_1=x_1+x_2+x_3 \\ y_2=x_2-x_3 \\ y_3=x_3 \end{cases} \quad \text{即} \quad \begin{cases} x_1=y_1-y_2-2y_3 \\ x_2=y_2+y_3 \\ x_3=y_3 \end{cases}$$

便有二次型的标准形 $f=y_1^2+y_2^2$，所用合同变换变换矩阵为

$$C=\begin{pmatrix}1 & -1 & -2\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}\qquad (\det C\neq 0)$$

当二次型 f 中不含任何变量的平方项时，上述方法不能直接应用，这时可先作适当的变量替换使二次型中出现平方项，比如，若 f 中含 $x_ix_j(i\neq j)$，可作变换 $x_i=y_i-y_j$，$x_j=y_i+y_j$，$x_k=y_k$，$k\neq i,j$ 使二次型中出现 y_j 的平方项．这样就可像例 5.11 一样进行配方了．

*5.5.4　用初等变换法化二次型为标准形

只需考虑将实对称矩阵 $\boldsymbol{A}$ 用初等变换化为对角矩阵即可．方法是：在 $\boldsymbol{A}$ 的下方放一个同阶单位矩阵，记其为矩阵 $\boldsymbol{B}$，(1) 若 $\boldsymbol{A}$ 中的 $a_{11}\neq 0$，则可用初等行变换将 $\boldsymbol{A}$（只对 $\boldsymbol{A}$）的第一列各行的元素变为 0，然后再对 $\boldsymbol{B}$（一定对整个矩阵 $\boldsymbol{B}$）进行相同的列变换，这时 $\boldsymbol{A}$ 的第一行和第一列除 a_{11} 外的元素全变成了零，若在 $\boldsymbol{A}$ 的原 a_{22} 位置上的元素不为零，则重复上述步骤，如此下去直至将 $\boldsymbol{A}$ 化为对角矩阵，此时单位矩阵变成的矩阵 $\boldsymbol{C}$ 就是使 $\boldsymbol{A}$ 与对角矩阵合同的合同变换矩阵，即 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵；(2) 若 $a_{11}=0$，但有 $a_{ii}\neq 0(2\leqslant i\leqslant n)$，则对 $\boldsymbol{A}$（只对 $\boldsymbol{A}$）施行行交换 $r_i\leftrightarrow r_1$，同时对 $\boldsymbol{B}$（对整个矩阵 $\boldsymbol{B}$）施行列交换 $c_i\leftrightarrow c_1$，接下来再按 (1) 的方法进行；(3) 若 $a_{ii}=0(i=1,2,\cdots,n)$，则必有 $a_{ij}=a_{ji}\neq 0(i<j)$，这时对 $\boldsymbol{A}$（只对 $\boldsymbol{A}$）施行行变换 r_i+r_j，同时对 $\boldsymbol{B}$（对整个矩阵 $\boldsymbol{B}$）施行 c_i+c_j，然后再按 (1) 的方法进行．例如

$$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{I}\end{pmatrix}=\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 3\\ 1 & 3 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\xrightarrow[\substack{\text{对 }\boldsymbol{B}:\\ c_2-c_1\\ c_3-c_1}]{\substack{\text{对 }\boldsymbol{A}:\\ r_2-r_1\\ r_3-r_1}}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 2\\ 0 & 2 & 0\\ 1 & -1 & -1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$\xrightarrow[\substack{\text{对 } \boldsymbol{B}: \\ c_2+c_3}]{\substack{\text{对 } \boldsymbol{A}: \\ r_2+r_3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 2 \\ 0 & 2 & 0 \\ 1 & -2 & -1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\xrightarrow[\substack{\text{对 } \boldsymbol{B}: \\ 2c_3-c_2}]{\substack{\text{对 } \boldsymbol{A}: \\ 2\boldsymbol{r}_3-\boldsymbol{r}_2}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -4 \\ 1 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

这里使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\mathrm{diag}(1,4,-4)$ 的矩阵 $\boldsymbol{C}=\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}$.

5.6 惯性定理和正定二次型

二次型化为标准形,由于采用的可逆变换不同而使标准形不同,也就是说,二次型的标准形不是唯一的.但因任何可逆变换不改变二次型的秩,所以标准形中所含非零系数的项数(即二次型的秩)是确定的.不仅如此,在限定变换为实变换时,还可以证明(这里不证),标准形中正系数的个数(从而负系数的个数)是不变的.这就是下面的**惯性定理**.

定理 5.13　设实二次型 $\boldsymbol{f}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的秩为 r,有两个实可逆变换

$$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y} \quad \text{及} \quad \boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$$

使
$$f=k_1y_1^2+k_2y_2^2+\cdots+k_ry_r^2,\quad (k_i\neq 0,i=1,2,\cdots,r)$$
及
$$f=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ry_r^2,\quad (\lambda_i\neq 0,i=1,2,\cdots,r)$$
则 $k_1,\cdots,k_r$ 中的正数个数(称为二次型的正惯性指数)与 $\lambda_1,\cdots,\lambda_r$ 中的正数个数相等.

经常用到的二次型是其标准形中的 n 个系数全为正或全为负的情形,据此,有下列定义.

定义 5.8　设有实二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,如果对任何 $\boldsymbol{x}\neq\boldsymbol{0}$ 有 $f(\boldsymbol{x})>0$(或<0),则称 f 为**正定(或负定)二次型**,并称对称矩阵 $\boldsymbol{A}$ **是正定(或负定)的**;如果对任何 $\boldsymbol{x}\neq\boldsymbol{0}$ 有 $f(\boldsymbol{x})\geqslant 0$(或$\leqslant 0$),且至

少有一个向量 $\boldsymbol{x}_0\neq\mathbf{0}$ 使 $f(\boldsymbol{x}_0)=0$ 则称 f 为**半正定(或半负定)二次型**,并称对称矩阵 $\boldsymbol{A}$ 是**半正定(或半负定)的.**

定理 5.14　实二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定(负定)的充要条件是它的标准形的 n 个系数全为正(负).

证　设可逆变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 使

$$f(\boldsymbol{x})=f(\boldsymbol{C}\boldsymbol{y})=\sum_{i=1}^{n}k_i y_i^2$$

充分性,设 $k_j>0,j=1,2,\cdots,n$,任给 $\boldsymbol{x}\neq\mathbf{0}$,则 $\boldsymbol{y}=\boldsymbol{C}^{-1}\boldsymbol{x}\neq 0$,因此 $f(\boldsymbol{x})=\sum_{i=1}^{n}k_i y_i^2>0$.

必要性,用反证法.假设有 $k_s\leqslant 0$,则当 $\boldsymbol{y}$ 为第 s 个单位坐标向量 $\boldsymbol{e}_s$ 时,显然 $\boldsymbol{C}\boldsymbol{e}_s\neq\mathbf{0}$,但 $f(\boldsymbol{C}\boldsymbol{e}_s)=k_s\leqslant 0$,这与 f 为正定矛盾,矛盾表明 $k_i>0,i=1,2,\cdots,n$.

对负定的充要条件可类似地证明.

例 5.12　判定二次型

$$f=-x_1^2-2x_2^2-3x_3^2+2x_1x_2+2x_1x_3$$

的正定性.

解　应用配方法得

$$f=-(x_1-x_2-x_3)^2-(x_2-x_3)^2-x_3^2=-y_1^2-y_2^2-y_3^2$$

由于标准形中的系数全为负,所以 f 为负定二次型.

例 5.13　设 $\boldsymbol{A}$、$\boldsymbol{B}$ 分别为 m 阶、n 阶正定矩阵,试判定矩阵

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

的正定性.

解　设 $\boldsymbol{x}^{\mathrm{T}}=(\boldsymbol{x}_1^{\mathrm{T}},\boldsymbol{x}_2^{\mathrm{T}})$,其中,$\boldsymbol{x}_1$ 为 m 维列向量,$\boldsymbol{x}_2$ 为 n 维列向量,则

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}=(\boldsymbol{x}_1^{\mathrm{T}},\boldsymbol{x}_2^{\mathrm{T}})\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{pmatrix}=\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_1+\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}_2^{\mathrm{T}}$$

由于 $\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}_2$ 均为正定二次型,故 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}$ 为正定二次型,从而 $\boldsymbol{C}$ 为正定矩阵.

推论　实对称矩阵 $\boldsymbol{A}$ 为正定(负定)的充要条件是 $\boldsymbol{A}$ 的特征值全为正数(负数).

定理 5.15　n 阶实对称矩阵 $\boldsymbol{A}$ 为正定的充要条件是 $\boldsymbol{A}$ 的 n 个前主子式(即前主子矩阵的行列式)均为正,即

$$a_{11}>0,\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0,\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}>0,\cdots,\det\boldsymbol{A}>0$$

实对称矩阵 $\boldsymbol{A}$ 为负定的充要条件是它的奇数阶前主子式为负,偶数阶主子式为正,即

$$(-1)^r\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{vmatrix}>0,r=1,2,\cdots,n.$$

这个定理称为霍尔维兹定理,这里不证.

例 5.14　判别矩阵

$$\boldsymbol{A}=\begin{pmatrix} -5 & 2 & 2 \\ 2 & -6 & 0 \\ 2 & 0 & -4 \end{pmatrix}$$

的正定性.

解　因为

$$a_{11}=-5<0,\quad \begin{vmatrix} -5 & 2 \\ 2 & -6 \end{vmatrix}>0,\det\boldsymbol{A}=-80<0$$

所以 $\boldsymbol{A}$ 是负定的,从而它的二次型

$$f=-5x_1^2+4x_1x_2+4x_1x_3-6x_2^2-4x_3^2$$

是负定的.

* 5.7 一些应用

5.7.1 主轴问题

在空间解析几何中，方程

$$ax^2+by^2+cz^2+2dyz+2exz+2fxy+2gx+2hy+2kz=1 \tag{5.22}$$

一般表示曲面，其中 a,b,c,d,e,f,g,h,k 均为常数. 为了识别曲面的形状，方程必须利用适当的几何变换化简为标准形.

方程(5.22)中的一次项 $2gx+2hy+2kz$ 可以通过平移变换 $x=X+\alpha, y=Y+\beta, z=Z+\gamma$ 消去，其中 α,β,γ 必须满足方程组

$$\begin{cases} a\alpha+f\beta+e\gamma=-g \\ f\alpha+b\beta+d\gamma=-h \\ e\alpha+d\beta+c\gamma=-k \end{cases} \tag{5.23}$$

如果方程(5.23)无解，则(5.22)表示无心二次曲面，称为抛物面；如果(5.23)有解，从中解出 α,β,γ，并作平移变换，则(5.22)式变为

$$aX^2+bY^2+cZ^2+2dYZ+2eXZ+2fXY=K \tag{5.24}$$

(5.24)式一般是有心(中心在原点的)二次曲面方程. 如果 $K\neq 0$，用 K 除两边，则(5.24)式可变为下列形式

$$a_{11}X^2+a_{22}Y^2+a_{33}Z^2+2a_{23}YZ+2a_{13}XZ+2a_{12}XY=1 \tag{5.25}$$

上式左边是一个二次型，若记

$$\boldsymbol{x}=(X,Y,Z)^{\mathrm{T}}, \boldsymbol{A}=(a_{ij})_{3\times 3}, (a_{ij}=a_{ji}, i,j=1,2,3)$$

则式(5.25)可写成

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=1$$

因为 $\boldsymbol{A}$ 是实对称矩阵，所以 $\boldsymbol{A}$ 有三个实特征值 $\lambda_1,\lambda_2,\lambda_3$，且对它们有三个相互正交的特征向量，将它们规范化后分别记为 $\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3$；则矩阵 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3)$ 为正交矩阵，正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 对应着坐

标系的旋转，经此变换，(5.25)式变成

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 = 1 \tag{5.26}$$

$\boldsymbol{A}$ 的特征向量称为二次曲面 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=1$ 的**主轴**，(5.26)式就是关于其主轴的有心曲面的方程.

容易看出，若 $\lambda_1,\lambda_2,\lambda_3$ 均为正数，则曲面是半轴为 $\sqrt{\frac{1}{\lambda_1}}$，$\sqrt{\frac{1}{\lambda_2}}$，$\sqrt{\frac{1}{\lambda_3}}$ 的椭球面；若 $\lambda_1,\lambda_2,\lambda_3$ 中有两个是正数，另一个是负数，则曲面为单叶双曲面；若 $\lambda_1,\lambda_2,\lambda_3$ 中有两个为负数，另一个为正数，则曲面为双叶双曲面；若 $\lambda_1,\lambda_2,\lambda_3$ 均为负数，则方程不表示任何曲面.

5.7.2　线性微分方程组的解法

工程技术中的很多问题其数学模型是形如

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} \tag{5.27}$$

的线性微分方程组，其中 $\boldsymbol{x}=\boldsymbol{x}(t)=(x_1(t),x_2(t),\cdots,x_n(t))^{\mathrm{T}}$，$\boldsymbol{A}$ 为 n 阶实对称矩阵. 下面讨论它的求解问题.

根据定理 5.10 存在正交矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{A}$ 对角化，即 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值. 因此，通过正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$，式(5.27)变为

$$\dot{\boldsymbol{y}} = \boldsymbol{\Lambda}\boldsymbol{y}$$

即

$$\dot{y}_1=\lambda_1 y_1,\quad \dot{y}_2=\lambda_2 y_2,\quad \cdots,\quad \dot{y}_n=\lambda_n y_n \tag{5.28}$$

显然(5.28)中的每一个都是可分离变量的一阶微分方程，它们的解是

$$y_i = a_i \mathrm{e}^{\lambda_i t},\ i = 1,2,\cdots,n$$

其中 $a_1,\cdots,a_n$ 为任意积分常数，于是得(5.28)的通解为

$$\boldsymbol{x}(t) = \boldsymbol{P}(a_1\mathrm{e}^{\lambda_1 t},a_2\mathrm{e}^{\lambda_2 t},\cdots,a_n\mathrm{e}^{\lambda_n t})^{\mathrm{T}}$$

对于工程技术中经常遇到的形如

$$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k),\ k = 0,1,2,\cdots$$

的线性差分方程组，其中 $\boldsymbol{x}(k)=(x_1(k),x_2(k),\cdots,x_n(k))^{\mathrm{T}}$，$\boldsymbol{A}$ 为

n 阶实对称矩阵,可以用同上面十分类似的方法来解.

5.7.3 函数最优化

从微积分中我们已知,求函数的极值是工程技术中经常遇到的问题.对于二元函数 $f(x,y)$,如果它具有一、二阶连续偏导数,则向量

$$\left(\frac{\partial f}{\partial x}\boldsymbol{i}+\frac{\partial f}{\partial y}\boldsymbol{j}\right)_{P_0}$$

称为函数 $f(x,y)$在 $P_0(x_0,y_0)$的**梯度**,记为 $\mathbf{grad}f$,并且 $f(x,y)$在 $P_0(x_0,y_0)$取得极值的必要条件是

$$\mathbf{grad}f\,\big|_{P_0}=\mathbf{0}$$

若记 $A=\left.\frac{\partial^2 f}{\partial x^2}\right|_{P_0}$, $B=\left.\frac{\partial^2 f}{\partial x\partial y}\right|_{P_0}$, $C=\left.\frac{\partial^2 f}{\partial y^2}\right|_{P_0}$,从微积分中知道,当 $A>0$,且 $AC-B^2>0$ 时,$f(x,y)$在 P_0 取极小值;而当 $A<0$,且 $AC-B^2>0$ 时,$f(x,y)$在 P_0 取极大值.现在若引入矩阵

$$\boldsymbol{H}=\begin{pmatrix}A & B\\ B & C\end{pmatrix}$$

显然 $\boldsymbol{H}$ 是实对称矩阵,而 $A>0$,且 $AC-B^2>0$ 意味着 $\boldsymbol{H}$ 是正定的;$A<0$,且 $AC-B^2>0$ 意味着 $\boldsymbol{H}$ 是负定的.这样,微积分中有关二元函数极值的充分条件可叙述为,当 $\boldsymbol{H}$ 为正定矩阵时,$f(x,y)$在 P_0 取极小值;当 $\boldsymbol{H}$ 为负定矩阵时,$f(x,y)$在 P_0 取极大值.

把有关二元函数 $f(x,y)$的上述结论推广到 n 元函数 $f(\boldsymbol{x})$,其中 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$,有:

(1)必要条件:设 $f(x_1,x_2,\cdots,x_n)$具有一阶偏导数.若 $f(\boldsymbol{x})$在 $\boldsymbol{x}=\boldsymbol{a}$ 取得极值,则

$$\mathbf{grad}f\,\big|_{\boldsymbol{x}=\boldsymbol{a}}=\mathbf{0}$$

其中 $\mathbf{grad}f\big|_{\boldsymbol{x}=\boldsymbol{a}}=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_n}\right)^{\mathrm{T}}_{\boldsymbol{x}=\boldsymbol{a}}$, $\boldsymbol{a}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$.

(2)充分条件:假设 $f(x_1,x_2,\cdots,x_n)$具有一、二阶连续偏导数,$\boldsymbol{x}=\boldsymbol{a}$ 是 $f(\boldsymbol{x})$的极值点.若记

$$h_{ij}=\frac{\partial^2 f}{\partial x_i \partial x_j}, i,j=1,2,\cdots,n, \boldsymbol{H}=(h_{ij})_{n\times n}$$

则

（Ⅰ）当 $\boldsymbol{H}$ 为正定矩阵时，$f(\boldsymbol{x})$ 在 $\boldsymbol{x}=\boldsymbol{a}$ 取极小值；

（Ⅱ）当 $\boldsymbol{H}$ 为负定矩阵时，$f(\boldsymbol{x})$ 在 $\boldsymbol{x}=\boldsymbol{a}$ 取极大值.

显然 $\boldsymbol{H}$ 为一实对称矩阵. 通常称 $\boldsymbol{H}$ 为**海赛(Hesse)矩阵**.

习 题 五

1. 试将线性无关的向量组

$$\boldsymbol{\alpha}_1=(1,2,0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(2,0,2)^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=(2,1,2)^{\mathrm{T}}$$

正交规范化.

2. 检验下列矩阵是否为正交矩阵

(1) $\begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{3} \\ -\frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & -1 \end{pmatrix}$　　(2) $\begin{pmatrix} \frac{1}{9} & -\frac{8}{9} & -\frac{4}{9} \\ -\frac{8}{9} & \frac{1}{9} & -\frac{4}{9} \\ -\frac{4}{9} & -\frac{4}{9} & \frac{7}{9} \end{pmatrix}$

3. 证明定理 5.3.

4. 设 $\boldsymbol{A}$ 是正交矩阵，证明 $\det\boldsymbol{A}=\pm 1$.

5. 设 $\boldsymbol{x}$ 为 n 维列向量，且 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$，证明 $\boldsymbol{H}=\boldsymbol{I}-2\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}$ 是对称的正交矩阵.

6. 求

$$\boldsymbol{A}=\begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

的特征值和特征向量，并将特征向量正交规范化.

7$^{\triangle}$. 设 $\boldsymbol{A},\boldsymbol{B}$ 均为 n 阶方阵，且 $\boldsymbol{A}$ 是满秩的，证明 $\boldsymbol{BA}$ 与 $\boldsymbol{AB}$ 有相同的特征多项式.

8. 设三阶方阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1=1,\lambda_2=0,\lambda_3=-1$，对应的特征向量依次是

$$\boldsymbol{p}_1=(1,2,2)^{\mathrm{T}},\quad \boldsymbol{p}_2=(2,-2,1)^{\mathrm{T}},\quad \boldsymbol{p}_3=(-2,-1,2)^{\mathrm{T}}$$

求矩阵 $\boldsymbol{A}$.

9. 求一个正交相似变换矩阵，将下列矩阵化为对角矩阵

(1) $\begin{pmatrix} 1 & 0 & \sqrt{3} \\ 0 & 3 & 0 \\ \sqrt{3} & 0 & -1 \end{pmatrix}$ (2) $\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$.

10. 用矩阵表示下列二次型：

(1) $f=2x_1^2+x_2^2-4x_1x_2-4x_2x_3$

(2) $f=2x_1x_3+x_2^2$

(3) $f=x_1^2+x_2^2+x_3^2+x_4^2-2x_1x_2+4x_1x_3-2x_1x_4+6x_2x_3+4x_2x_4$

11. 通过正交变换，化下列二次型为标准形并写出所用的正交变换矩阵 $\boldsymbol{P}$.

(1) $f=2x_1^2+3x_2^2+3x_3^2+4x_2x_3$

(2) $f=x_1^2+x_2^2+x_3^2+x_4^2+2x_1x_2-2x_1x_4-2x_2x_3+2x_3x_4$

(3)$^{\triangle}$ 已知二次型 $f=2x_1^2+3x_2^2+3x_3^2+2ax_2x_3\ (a>0)$，通过正交变换化成标准形 $f=y_1^2+2y_2^2+5y_3^2$，求参数 a 及所用的正交变换矩阵.

12. 判别下列矩阵的正定性

(1) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}$ (2) $\begin{pmatrix} -5 & 2 & 2 \\ 2 & -2 & 0 \\ 2 & 0 & -4 \end{pmatrix}$

13. 判别下列二次型的正定性

(1) $f=-2x_1^2-6x_2^2-4x_3^2+2x_1x_2+2x_1x_3$

(2) $f=x_1^2+3x_2^2+9x_3^2+19x_4^2-2x_1x_2+4x_1x_3+2x_1x_4-6x_2x_4-12x_3x_4$

14. 设二次型 $f_1=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定的，$f_2=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}$ 为半正定的，其

中 $\boldsymbol{A}$,$\boldsymbol{B}$ 均为 n 阶实对称矩阵. 证明二次型 f_1+f_2 是正定的.

15. 设 $\boldsymbol{A}$ 为实对称矩阵,证明 $\boldsymbol{A}$ 为正定的充要条件是存在满秩方阵 $\boldsymbol{U}$ 使 $\boldsymbol{A}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}$.

16. 设 $\boldsymbol{A}$ 为正定矩阵,$\boldsymbol{B}$ 为同阶实对称矩阵,证明存在可逆矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{I}$ 且 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}$ 为对角阵.

17. 正定矩阵 $\boldsymbol{A}$ 为正交矩阵的充要条件是 $\boldsymbol{A}=\boldsymbol{I}$.

18. (1)证明 n 阶方阵为非奇异阵的充要条件是其特征值 $\lambda_i\neq 0, i=1,2,\cdots,n$.

(2)设 λ 为非奇异矩阵 A 的特征值,证明 $1/\lambda$ 是 $\boldsymbol{A}^{-1}$ 的特征值.

* **19.** 用幂法求矩阵

$$\boldsymbol{A}=\begin{pmatrix}1 & -3 & 2\\ 4 & 4 & -1\\ 6 & 3 & 5\end{pmatrix}$$

按模最大的特征值和对应的特征向量.

20$^{\triangle}$. 设 $\boldsymbol{A}$ 是正定矩阵,试证 $\det(\boldsymbol{A}+\boldsymbol{I})>1$,其中 $\boldsymbol{I}$ 为单位矩阵.

21. 设 $\boldsymbol{A}$ 为实对称矩阵,证明存在实数 $t_0>0$,使当 $t>t_0$ 时,$\boldsymbol{A}+t\boldsymbol{I}$ 为正定矩阵.

22. 设矩阵

$$\boldsymbol{A}=\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\\ 1 & 1 & 1 & 1\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}4 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0\end{pmatrix}$$

则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$(　　).

(A) 合同且相似　(B) 合同但不相似　(C) 不合同但相似

(D) 不合同且不相似

习 题 答 案

习 题 一

1. $$\begin{pmatrix} \dfrac{\partial z}{\partial x} \\ \dfrac{\partial z}{\partial y} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x} & \dfrac{\partial w}{\partial x} \\ \dfrac{\partial u}{\partial y} & \dfrac{\partial v}{\partial y} & \dfrac{\partial w}{\partial y} \end{pmatrix} \begin{pmatrix} \dfrac{\partial f}{\partial u} \\ \dfrac{\partial f}{\partial v} \\ \dfrac{\partial f}{\partial w} \end{pmatrix}$$

2. $$\begin{array}{cc} \begin{array}{cc} \text{成 1} & \text{成 2} \end{array} & \\ \begin{pmatrix} 95 & 63 \\ 14 & 10 \\ 30 & 21 \\ 21 & 11 \end{pmatrix} & \begin{array}{l} \text{原 1} \\ \text{原 2} \\ \text{原 3} \\ \text{原 4} \end{array} \end{array}$$

3. $$\begin{cases} z_1 = 4x_1 - 3x_2 + 5x_3 \\ z_2 = 6x_1 + 8x_2 - 6x_3 \\ z_3 = -23x_1 - 4x_2 - 2x_3 \end{cases}$$

4. (1) $\begin{pmatrix} -2 & 13 & 22 \\ -2 & -17 & 20 \\ 4 & 29 & -2 \end{pmatrix}$ (2) $\begin{pmatrix} 0 & 0 & 2 \\ 5 & -5 & 9 \\ 8 & 6 & 0 \end{pmatrix}$

5. (1) $\boldsymbol{A}^2 = \mathrm{diag}(a_{11}^2, a_{22}^2, \cdots, a_{nn}^2)$

(2) $\boldsymbol{AB} = \mathrm{diag}(a_{11}b_{11}, a_{22}b_{22}, \cdots, a_{nn}b_{nn})$

(4) $$\boldsymbol{AC} = \begin{pmatrix} a_{11}c_{11} & a_{11}c_{12} & \cdots & a_{11}c_{1n} \\ a_{22}c_{21} & a_{22}c_{22} & \cdots & a_{22}c_{2n} \\ \vdots & \vdots & & \vdots \\ a_{nn}c_{n1} & a_{nn}c_{n2} & \cdots & a_{nn}c_{nn} \end{pmatrix}$$

$$\boldsymbol{CA} = \begin{pmatrix} c_{11}a_{11} & c_{12}a_{22} & \cdots & c_{1n}a_{nn} \\ c_{21}a_{11} & c_{22}a_{22} & \cdots & c_{2n}a_{nn} \\ \vdots & \vdots & & \vdots \\ c_{n1}a_{11} & c_{n2}a_{22} & \cdots & c_{nn}a_{nn} \end{pmatrix}$$

6. (1) $\boldsymbol{AB} = \begin{pmatrix} a_{11} & a_{12} & ka_{13} \\ a_{21} & a_{22} & ka_{23} \\ a_{31} & a_{32} & ka_{33} \end{pmatrix}$, $\boldsymbol{BA} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ ka_{31} & ka_{32} & ka_{33} \end{pmatrix}$

$$\boldsymbol{AC} = \begin{pmatrix} a_{11} + ka_{12} & a_{12} & a_{13} \\ a_{21} + ka_{22} & a_{22} & a_{23} \\ a_{31} + ka_{32} & a_{32} & a_{33} \end{pmatrix}$$

$$CA=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ ka_{11}+a_{21} & ka_{12}+a_{22} & ka_{13}+a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

(2) $a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2$

$+(a_{12}+a_{21})x_1x_2+(a_{13}+a_{31})x_1x_3+(a_{23}+a_{32})x_2x_3$

9. $X=\dfrac{1}{2}(A^2+BA)$, $\quad Y=\dfrac{1}{2}(A^2-BA)$.

10. (1) $A=\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, $\quad B=\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$, $\qquad$ (2) 同(1)

(3) $A=\begin{pmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{pmatrix}$ $\qquad$ (4) $A=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

(5) $A=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, $\quad X=\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\quad Y=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

11. (1) $A^{-1}=\begin{pmatrix} 0 & -1 \\ \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix}$, $\quad B^{-1}=\begin{pmatrix} \dfrac{1}{2} & -\dfrac{1}{2} \\ 0 & 1 \end{pmatrix}$

$(AB)^{-1}=\begin{pmatrix} -\dfrac{1}{4} & -\dfrac{3}{4} \\ \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix}$, $\quad$ (2) $C^{-1}=\begin{pmatrix} 0 & -1 & 0 & 0 \\ \dfrac{1}{2} & \dfrac{1}{2} & 0 & 0 \\ 0 & 0 & \dfrac{1}{2} & -\dfrac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$

15*. $A=\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \dfrac{1}{2} & -\dfrac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} 2 & 2 & 3 \\ 0 & 2 & -3 \\ 0 & 0 & -1 \end{pmatrix}$

16. $$\left(\begin{array}{c:ccc} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ \hdashline * & * & * & * \end{array}\right)\left(\begin{array}{c:ccc} * & * & * & * \\ \hdashline * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{array}\right)=\left(\begin{array}{c:ccc} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ \hdashline * & * & * & * \end{array}\right)$$

19. $\boldsymbol{A}=\begin{pmatrix}1&2&0&0\\0&1&0&0\\0&0&-16&3\\0&0&-6&3\end{pmatrix}$, $\boldsymbol{A}^2=\begin{pmatrix}1&4&0&0\\0&1&0&0\\0&0&238&-39\\0&0&78&-9\end{pmatrix}$

20. $\boldsymbol{A}^n=\begin{pmatrix}\boldsymbol{I}&\boldsymbol{O}\\n\boldsymbol{A}_1&\boldsymbol{I}\end{pmatrix}$

23. (C)

24. $\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$

习　题　二

1. $M_{21}=\begin{vmatrix}2&3\\0&2\end{vmatrix}=4$　$M_{22}=\begin{vmatrix}1&3\\0&2\end{vmatrix}=2$　$M_{23}=\begin{vmatrix}1&2\\0&0\end{vmatrix}=0$

$D_{21}=(-1)^{2+1}M_{21}=-4$　$D_{22}=(-1)^{2+2}M_{22}=2$

$D_{23}=(-1)^{2+3}M_{23}=0$

2. (1)40，（2）48，（3）10368，

（4）$x^n+(-1)^{n+1}y^n$，（5）$(-1)^{n(n-1)/2}\lambda_1\lambda_2\cdots\lambda_n$.

4. (1) $\boldsymbol{X}=\begin{pmatrix}2&-23\\0&8\end{pmatrix}$　(2) $\boldsymbol{X}=\begin{pmatrix}1&1\\\frac{1}{4}&0\end{pmatrix}$　(3) $\boldsymbol{X}=\begin{pmatrix}3&-1\\2&0\\1&-1\end{pmatrix}$

5. $2^{12}(a^2+b^2)^4$

6. (1)　$x=1,y=-1,z=3$

(2) $x_1=\frac{3}{5}$，　$x_2=-\frac{1}{5}$，　$x_3=-\frac{1}{5}$，　$x_4=\frac{3}{5}$

7. 当 $a\neq b,c,b\neq c$ 时，方程组有唯一解，它为

$$x_1=abc,\quad x_2=-ca-ab-bc,x_3=a+b+c$$

8. (1) $z=0,y=2,x=1$，（2）$x_4=-1,x_3=0,x_2=2,x_1=1$

9. (1) $\boldsymbol{A}^{-1}=\begin{pmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{pmatrix}$　(2) $\boldsymbol{B}^{-1}=\begin{pmatrix}-\frac{1}{3}&1&\frac{4}{3}\\-\frac{1}{3}&1&\frac{1}{3}\\\frac{2}{3}&-1&-\frac{2}{3}\end{pmatrix}$

10. (1) $\boldsymbol{A}^{-1}=\begin{pmatrix} 2 & 0 & -1 \\ -\dfrac{3}{2} & \dfrac{1}{2} & \dfrac{1}{2} \\ -1 & 0 & 1 \end{pmatrix}$

(2) $C^{-1}=\begin{pmatrix} 0 & 0 & \cdots & 0 & \dfrac{1}{c_n} \\ \dfrac{1}{c_1} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \dfrac{1}{c_{n-1}} & 0 \end{pmatrix}$

(3) $(\boldsymbol{B}-2\boldsymbol{I}_3)^{-1}=\begin{pmatrix} 1 & 0 & 0 \\ -\dfrac{1}{2} & \dfrac{1}{2} & -\dfrac{3}{2} \\ 0 & 0 & 1 \end{pmatrix}$

11. $\boldsymbol{X}^{-1}=\begin{pmatrix} \boldsymbol{O} & \boldsymbol{B}^{-1} \\ \boldsymbol{A}^{-1} & \boldsymbol{O} \end{pmatrix}$ **12.** $\boldsymbol{A}=\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 6 & -1 & -1 \end{pmatrix}$, $\boldsymbol{A}^5=\boldsymbol{A}$.

13. (D).

习 题 三

1. $R(\boldsymbol{A})=3$.

2. 可以有 r 阶奇异子方阵，如 $\boldsymbol{A}_{4\times 4}=\begin{pmatrix} \boldsymbol{I}_2 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$

3. $R(\boldsymbol{B})\leqslant R(\boldsymbol{A})$.

4. $\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

5. $P=\begin{pmatrix}1&0&0\\-3&1&0\\2&-1&1\end{pmatrix}$, $Q=\begin{pmatrix}1&\frac{1}{4}&-\frac{1}{2}&-\frac{1}{4}\\0&\frac{1}{4}&\frac{3}{2}&-\frac{5}{4}\\0&0&1&0\\0&0&0&1\end{pmatrix}$

11. (1)秩为 3,(2)秩为 3,(3)秩为 3,(4)秩为 1.

12. $P=\begin{pmatrix}1&0&0&0\\0&1&0&0\\-1&0&1&1\\-2&1&1&0\end{pmatrix}$, $Q=\begin{pmatrix}1&-\frac{1}{2}&\frac{3}{4}&-\frac{1}{2}&0\\0&\frac{1}{2}&\frac{1}{4}&-\frac{3}{2}&-\frac{1}{2}\\0&0&-\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\0&0&0&\frac{1}{2}&0\\0&0&0&0&-\frac{1}{2}\end{pmatrix}$

13. (1) 有非零解; (2) 仅有零解; (3) 无解

(4) 有无穷多个解; (5) 有无穷多个解; (6) 有无穷多个解.

14. $\lambda\neq-2,1$ 时有唯一解,$\lambda=1$ 时有无穷多个解,$\lambda=-2$ 时无解.

15. $\lambda=-2$ 或 1 时有无穷多个解.

16. 0

习 题 四

1. 不同.

2. (1) 如 $\boldsymbol{\alpha}=(4,4)$,$\boldsymbol{\alpha}_1=(2,2)$,$\boldsymbol{\alpha}_2=(1,1)$线性相关,$\boldsymbol{\alpha}$ 可由 $\boldsymbol{\alpha}_1$,$\boldsymbol{\alpha}_2$ 线性表示,但表示式不唯一.

(2) 如 $\boldsymbol{\alpha}_1=(1,0,0)$,$\boldsymbol{\alpha}_2=(0,1,0)$,$\boldsymbol{\alpha}_3=(0,2,0)$,显然它们线性相关,但 $\boldsymbol{\alpha}_1$ 不能由 $\boldsymbol{\alpha}_2$,$\boldsymbol{\alpha}_3$ 线性表示.

(3) 如 $\boldsymbol{\alpha}_1=(2,2)$,$\boldsymbol{\alpha}_2=(1,1)$,$\boldsymbol{\beta}_1=(-1,0)$,$\boldsymbol{\beta}_2=(0,1)$,取

$$\lambda_1=1,\lambda_2=-1$$

6. (1) 线性相关,秩为 2,$\boldsymbol{\alpha}_1$,$\boldsymbol{\alpha}_2$ 是一个最大无关组.

(2) 线性无关,秩为 3. (3) 线性无关,秩为 4.

(4) 线性相关,秩为 2,$\boldsymbol{\alpha}_1$,$\boldsymbol{\alpha}_2$ 是一个最大无关组.

7. V_1 是向量空间，是 R^n 的子空间，V_2 不是向量空间.

10. $\boldsymbol{C}=\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & \frac{1}{2} \end{pmatrix}$

11. (1) 基础解系 $\boldsymbol{\xi}_1=(\frac{3}{8},-\frac{25}{8},0,1,0)^{\mathrm{T}}$，$\boldsymbol{\xi}_2=(-\frac{1}{2},\frac{1}{2},0,0,1)^{\mathrm{T}}$，

通解 $\boldsymbol{x}=k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$.

(2) 基础解系 $\boldsymbol{\xi}=(0,0,0,1,1)^{\mathrm{T}}$，通解为 $x=k\boldsymbol{\xi}$.

12. (1) 通解 $\boldsymbol{x}=\boldsymbol{\eta}^*+k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$，其中

$\boldsymbol{\eta}^*=(-1,-3,0,0)^{\mathrm{T}}$，$\boldsymbol{\xi}_1=(-8,-13,1,0)^{\mathrm{T}}$，$\boldsymbol{\xi}_2=(5,9,0,1)^{\mathrm{T}}$

(2) 唯一解 $\boldsymbol{x}=(-8,0,0,-3)^{\mathrm{T}}$

(3) 唯一解 $\boldsymbol{x}=(2,1,-1)^{\mathrm{T}}$

13. $\lambda\neq1,-2$ 时，有唯一解

$$x_3=\frac{(\lambda+1)^2}{(\lambda+2)},\quad x_2=\frac{1}{(\lambda+2)},\quad x_1=-\frac{\lambda+1}{\lambda+2}$$

$\lambda=-2$ 时无解；$\lambda=1$ 时有无穷多个解，通解为

$$(x_1,x_2,x_3)^{\mathrm{T}}=(1,0,0)^{\mathrm{T}}+k_1(-1,1,0)^{\mathrm{T}}+k_2(-1,0,1)^{\mathrm{T}}$$

14. $n=2$ 时，无解：两直线平行；有唯一解：两直线相交；有无穷多个解：两直线重合.

$n=3$ 时，无解：三平面平行或两平面平行，另一平面与其中一平面重合或三平面中两两的交线中有两条交线平行；有唯一解：三平面只有一个交点；无穷多个解：解中含有一个任意参数时，三平面交于一直线，或两平面交于一直线，另一平面与其中一平面重合；解中有两个任意参数时，三平面重合.

15. s 为奇数时线性无关，s 为偶数时线性相关.

17. (1) $a=-1,b\neq0$ 时，$\boldsymbol{\beta}$ 不能表示成 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的线性组合.

(2) $\boldsymbol{a}\neq-1$ 时，$\boldsymbol{\beta}=\frac{-2b}{a+1}\boldsymbol{\alpha}_1+\frac{a+b+1}{a+1}\boldsymbol{\alpha}_2+\frac{b}{a+1}\boldsymbol{\alpha}_3+0\boldsymbol{\alpha}_4$

24. (C)

25*. 取初始向量

$$\boldsymbol{x}^{(0)}=(0.7600000,\ 0.0800000,\ 1.1200000,\ 0.6800000)^{\mathrm{T}}$$

时，雅可比方法迭代 12 次的结果是

$$\boldsymbol{x}^{(11)}=(1.5347196,\ 0.1220096,\ 1.9749109,\ 1.4127100)^{\mathrm{T}}$$

$$\boldsymbol{x}^{(12)}=(1.5348472,\ 0.1220096,\ 1.9750386,\ 1.4128376)^{\mathrm{T}}$$

塞得尔迭代 10 次的结果是

$$x^{(9)} = (1.5346912, 0.1220044, 1.9749370, 1.4127954)^{T}$$

$$x^{(10)} = (1.5348559, 0.1220073, 1.9750690, 1.4128917)^{T}$$

习　题　五

1. $\boldsymbol{\eta}_1 = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0\right)^{T}$,

$\boldsymbol{\eta}_2 = \left(\frac{4}{3\sqrt{5}}, -\frac{2}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}\right)^{T}$,

$\boldsymbol{\eta}_3 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)^{T}$

2. (1) 不是,因第一列不是单位向量　(2) 是.

6. 特征值为 $\lambda_1=-1, \lambda_2=1, \lambda_3=\lambda_4=2$,其对应的特征向量依次为

$\boldsymbol{\xi}_1 = (-3,2,0,0)^{T}, \boldsymbol{\xi}_2 = (1,0,0,0)^{T}, \boldsymbol{\xi}_3 = (6,1,3,0)^{T}$.

正交规范化后为

$$\boldsymbol{p}_1 = \left(-\frac{3}{\sqrt{13}}, \frac{2}{\sqrt{13}}, 0, 0\right)^{T}, \boldsymbol{p}_2 = \left(\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}, 0, 0\right)^{T}, \boldsymbol{p}_3 = (0,0,1,0)^{T}$$

8. $\boldsymbol{A} = \frac{1}{3}\begin{pmatrix} -1 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 0 \end{pmatrix}$

9. (1) $\boldsymbol{P} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \end{pmatrix}$

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \mathrm{diag}(-2,2,3)$$

(2) $\boldsymbol{P} = \begin{pmatrix} 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \end{pmatrix}$,

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \mathrm{diag}(1,1,10)$$

10. (1) $(x_1,x_2,x_3)\begin{pmatrix}2 & -2 & 0\\ -2 & 1 & -2\\ 0 & -2 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$

(2) $(x_1,x_2,x_3)\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$

(3) $(x_1,x_2,x_3,x_4)\begin{pmatrix}1 & -1 & 2 & -1\\ -1 & 1 & 3 & 2\\ 2 & 3 & 1 & 0\\ -1 & 2 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\end{pmatrix}$

11. (1) $\boldsymbol{P}=\begin{pmatrix}0 & 1 & 0\\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{pmatrix}$, $\quad f=y_1^2+2y_2^2+5y_3^2$

(2) $\boldsymbol{P}=\begin{pmatrix}\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2}\\ -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2}\\ -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{2}\\ \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{2}\end{pmatrix}$, $\quad f=-y_1^2+y_2^2+y_3^2+3y_4^2$

(3) $a=2$, $\quad \boldsymbol{P}=\begin{pmatrix}0 & 1 & 0\\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{pmatrix}$

12. (1) 正定矩阵,(2) 负定矩阵.

13. (1) 负定的,(2) 正定的.

19. 取 $\boldsymbol{x}^{(0)}=(0,0,1)^{\mathrm{T}}$ 迭代 10 次结果得特征值 $\lambda_1=7.14352$ 特征向量为$(0.29563, 0.06006, 1)^{\mathrm{T}}$.

22. (A)