线　性　代　数

(成人高等教育试用教材)

李永乐　编

清 华 大 学 出 版 社

(京)新登字 158 号

内 容 提 要

本书是根据作者在清华大学继续教育学院多年来的教学实践编写的.符合成人教育对本科“线性代数”教学的基本要求,具有成人教学的特点.

本书概念清楚,对基本要求部分叙述详细充实,重点突出,层次清晰,说理浅显,各种类型的例题丰富,坡度较小,适于自学.内容包括:行列式,矩阵,线性方程组,向量空间,特征值与特征向量,二次型及线性代数应用举例的附录.每章之后备有适量的练习题,书末有习题答案及提示.

本书可作为成人高等院校,继续教育学院专升本,夜大学教材,也可作为大学专科的教材及工程技术人员自学用书.

图书在版编目(CIP)数据

线性代数/李永乐编.—北京:清华大学出版社,1996
成人高等教育试用教材
ISBN 7-302-02385-9

Ⅰ.线… Ⅱ.李… Ⅲ.线性代数-成人教育:高等教育-教材 Ⅳ.0151.2

中国版本图书馆 CIP 数据核字(96)第 23307 号

出版者:清华大学出版社(北京清华大学校内,邮编 100084)
印刷者:密云胶印厂
发行者:新华书店总店北京科技发行所
开　本:850×1168 1/32　印张:8 1/8　字数:212 千字
版　次:1997 年 2 月 第 1 版　1997 年 2 月 第 1 次印刷
书　号:ISBN 7-302-02385-9/O·174
印　数:0001—5000
定　价:8.90 元

前　　言

线性代数是一门重要的基础课,它所涉及到的处理问题的思想、方法和技巧被广泛地应用于各学科,尤其是随着计算机的发展,这种离散化解决问题的手法更显重要.

考虑到成人教学的特点和规律,为了适应今后成人教育发展的需要,在清华大学继续教育学院的领导下,特别是康静安先生的鼓励和帮助,编写了这本“线性代数”教材.在编写本书时,注意了以下几点:

根据教学基本要求,突出重点及基本方法,对矩阵、线性方程组及特征值阐述详细透彻,力求学后能熟悉现代科技中常用的矩阵方法.

对难度较大的某些基础理论,例如“线性相关”、“秩”、“矩阵对角化”、“惯性定理”等,未作过份严密的论证和推导,有的以二、三维为例介绍基本思想方法,有的用打 * 号形式给出一个论证供有兴趣有余力的读者自学参考,也有的只是罗列出结果.

教材中注重例题的选择匹配,有些是介绍基本概念和基本运算的,以帮助读者正确理解概念及运算法则,有些是澄清初学者易犯错误之处,有些是介绍线性代数中常用技巧,一些题目给出一题多解,以求开阔思路,活跃思维,还有一些超出要求的提高题,这些灵活的综合题打上 * 号供读者选用.

为了便于自学,本书力求条理清晰,深入浅出,循序渐近,例题丰富,利于理解和掌握.对于打 * 号的加深内容,略去之后不会影响后面的学习.

本书可作为成人高等院校及大学专科的教材,也可作为少学

时工科院校的本科教材,并可供工程技术人员自学参考．

在编写本书的过程中,胡冠章教授详细审阅了全稿,提出了许多宝贵的建议,附录线性代数应用举例就是在他提议下增写的．

编写本书时,主要参考书有：栾汝书编著的《线性代数》,居余马、胡金德等编的《线性代数》,G·Strang 著,侯自新等译的《线性代数及其应用》,И.В.普罗斯库烈柯夫著,周晓钟译的《线性代数习题集》.

感谢清华大学继续教育学院,清华大学应用数学系对编者的关心信任,感谢康静安先生,及两年来支持帮助我的所有同仁．

由于水平所限,疏漏错误难免,悬请读者批评指正．

编　者

1996 年 6 月清华园

目　　录

第 1 章　行　列　式

行列式是一个重要的数学工具,在微积分的学习中已经看到:向量的叉积与混合积可以用二、三阶行列式来表示,直线及平面的一些问题如果运用行列式是较简捷的,在重积分的计算中,也出现过 Jacobi 行列式 .

在线性代数中,行列式是一个不可缺少的工具,它在方程组、矩阵、特征值及二次型中有许多重要应用 .

1.1　二、三阶行列式

对二元一次方程组

$$\begin{aligned} a_1 x + b_1 y &= c_1, \\ a_2 x + b_2 y &= c_2 . \end{aligned} \tag{1.1}$$

用加减消元法,可以得到

$$\begin{aligned} (a_1 b_2 - a_2 b_1) x &= c_1 b_2 - c_2 b_1, \\ (a_1 b_2 - a_2 b_1) y &= a_1 c_2 - a_2 c_1 . \end{aligned}$$

当 $a_1 b_2 - a_2 b_1 \neq 0$ 时,方程组(1.1)有唯一解

$$x = \frac{c_1 b_2 - c_2 b_1}{a_1 b_2 - a_2 b_1},$$

$$y = \frac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1} .$$

为了方便,如果引入记号

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

我们称其是二阶行列式 .a, b 是行列式的第 1 行; c, d 是行列式的第 2 行; a, c 是行列式的第 1 列; b, d 是行列式的第 2 列,这个行列式的值是 $ad - bc$.

例如,由行列式定义我们有

$$\begin{vmatrix} 2 & -1 \\ 3 & 6 \end{vmatrix} = 2 \times 6 - (-1) \times 3 = 15,$$

$$\begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} = \cos^2\theta + \sin^2\theta = 1 .$$

把行列式用到方程组(1.1),我们称

$$\Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

为方程组的系数行列式,当其值 $\Delta \neq 0$ 时,方程组的解就是:

$$x = \frac{\Delta_1}{\Delta},$$

$$y = \frac{\Delta_2}{\Delta},$$

其中 x 的分子 Δ_1 是把系数行列式中 x 的系数用常数项替换后所得到的行列式, y 的分子 Δ_2 是把系数行列式中 y 的系数换成常数项后所得到的行列式,而 x 和 y 的分母都是系数行列式 .

例 1 用行列式解方程组

$$3x + 2y = 5,$$
$$2x + 3y = 0 .$$

解 由系数行列式

$$\Delta = \begin{vmatrix} 3 & 2 \\ 2 & 3 \end{vmatrix} = 5 \neq 0$$

知方程组有唯一解 .又因

$$\Delta_1 = \begin{vmatrix} 5 & 2 \\ 0 & 3 \end{vmatrix} = 15, \quad \Delta_2 = \begin{vmatrix} 3 & 5 \\ 2 & 0 \end{vmatrix} = -10$$

得到方程组的解是:

$$x=\frac{\Delta_1}{\Delta}=3,$$

$$y=\frac{\Delta_2}{\Delta}=-2.$$

■

完全类似地,对三元一次方程组也有相应的结果,为了今后便于推广到更复杂的情形,未知数现在用 x_1, x_2, x_3 来表示,未知数的系数用带有两个下标的 a_{ij}表示,其中第 1 个下标 i 表示该项在第 i 个方程,第 2 个下标 j 表示它是未知数 x_j 的系数.这样三元一次方程组可写成

$$\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3&=b_1,\\ a_{21}x_1+a_{22}x_2+a_{23}x_3&=b_2,\\ a_{31}x_1+a_{32}x_2+a_{33}x_3&=b_3.\end{aligned} \tag{1.2}$$

利用加减消元法可以得到

$$(a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32})x_1$$

$$=b_1a_{22}a_{33}+b_2a_{32}a_{13}+b_3a_{12}a_{23}-b_3a_{22}a_{13}-b_2a_{12}a_{33}-b_1a_{32}a_{23}$$

这个结果很难记忆,为此引进三阶行列式的定义,我们称

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

是一个三阶行列式,其值规定为:

$$a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32},$$

它是 6 项的代数和,而每一项都是 3 个元素的乘积.这 3 个元素取自不同的行不同的列,其中有 3 项前面带正号,另 3 项前面带负号.

为了便于计算三阶行列式的值,这 6 项可以这样来记忆:在图 1.1 中,由左上至右下的实线上 3 个元素的乘积所构成的 3 项都带正号,由右上至左下的虚线上 3 个元素的乘积所构成的 3 项都带负号.

图 1.1

例如,根据行列式定义,可计算出

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1\times5\times9+4\times8\times3+2\times6\times7-3\times5\times7-2\times4\times9-6\times8\times1$$

$$=45+96+84-105-72-48=0.$$

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

$$=1+3a^2.$$

利用三阶行列式,方程组(1.2)的解就容易记忆了,当系数行列式

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0$$

时,方程组(1.2)有唯一解

$$x_1 = \frac{\Delta_1}{\Delta},$$

$$x_2 = \frac{\Delta_2}{\Delta},$$

$$x_3 = \frac{\Delta_3}{\Delta}.$$

其中 Δ_j 就是把常数项替代系数行列式 Δ 中第 j 列系数所得到的行列式 .

例 2 解方程组

$$\begin{cases} x_1 - 2x_2 - x_3 = 1, \\ 3x_2 + 2x_3 = 1, \\ 3x_1 + x_2 - x_3 = 0 . \end{cases}$$

解 先计算系数行列式

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

$= 1\times3\times(-1) + 0 + (-2)\times2\times3 - (-1)\times3\times3 - 0 - 2\times1\times 1$

$= -8,$

由于 $\Delta\neq0$,方程组有唯一解 .

再计算未知数的分子行列式 Δ_j .据已知条件,可计算出

$$\Delta_1 = \begin{vmatrix} 1 & -2 & -1 \\ 1 & 3 & 2 \\ 0 & 1 & -1 \end{vmatrix} = -8, \qquad \Delta_2 = \begin{vmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 3 & 0 & -1 \end{vmatrix} = 8,$$

$$\Delta_3 = \begin{vmatrix} 1 & -2 & 1 \\ 0 & 3 & 1 \\ 3 & 1 & 0 \end{vmatrix} = -16.$$

所以,方程组的解是

$$x_1 = \frac{\Delta_1}{\Delta} = 1,$$

$$x_2 = \frac{\Delta_2}{\Delta} = -1,$$ ■

$$x_3 = \frac{\Delta_3}{\Delta} = 2.$$

对于三阶行列式,我们虽然已会计算,但是数字的计算还是很繁琐的,因此要分析研究行列式的基本性质,以便利用性质来简化我们的计算.

性质 1 行列互换,行列式的值不变.即

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = D^{T}$$

通常把后一个行列式称为是原行列式的转置,用记号 D^{T} 或 D'表示.这条性质也可叙述为:经转置行列式的值不变.

这条性质表明在行列式中,行与列的地位是对称的,如果行有某个性质,那么列也就有同样的相关性质.为了简洁,下面只叙述与讨论行的性质.关于列所具有的性质就不重复了.

至于性质 1 的证明,可以利用行列式的定义把每个行列式展开成六项后经比较得到,在这里把证明略去.

观察三阶行列式定义中的六项,我们发现可把这六项分成三个小组,每组按行提取公因式后,剩余部分正巧可用二阶行列式描述.例如按第 1 行的元素来分组,即

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

$$=a_{11}(a_{22}a_{33}-a_{23}a_{32})-a_{12}(a_{21}a_{33}-a_{23}a_{31})+a_{13}(a_{21}a_{32}-a_{22}a_{31}).$$

如果把每组中第 1 行的元素 a_{1j}作为公因子提出后,剩余部分记作 A_{1j},则 A_{1j}中不含行列式 D 中第 1 行及第 j 列的元素,它可以用 D 中去掉第 1 行及第 j 列的二阶行列式来表示.也就是

$$A_{11}=a_{22}a_{33}-a_{23}a_{32}=\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix},$$

$$A_{12}=-(a_{21}a_{33}-a_{23}a_{31})=-\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix},$$

$$A_{13}=a_{21}a_{32}-a_{22}a_{31}=\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix},$$

这时行列式 D 的值可写成

$$D=a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}.$$

完全类似地,D 中的六项也可按第 2 行或第 3 行的元素来分组,这就是

性质 2 $\quad D=a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}$

$$=a_{21}A_{21}+a_{22}A_{22}+a_{23}A_{23}$$

$$=a_{31}A_{31}+a_{32}A_{32}+a_{33}A_{33}.$$

这称为行列式按行展开公式,其中 A_{ij}叫做 a_{ij}的代数余子式,它是把行列式 D 去掉 a_{ij}所在的 i 行和 j 列后所得到的一个二阶行列式,并带有正负号$(-1)^{i+j}$.

例如,a_{31}的代数余子式 A_{31}是

$$A_{31}=(-1)^{3+1}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}=\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

例 3　计算下面行列式的值,

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

解　注意到 D **中第 2 行有两个元素** $a_{22} = a_{23} = 0$, **根据性质 2 按第 2 行展开,有**

$$\begin{aligned} D &= a_{21} A_{21} \\ &= 4 \times (-1)^{2+1} \begin{vmatrix} 2 & 3 \\ 6 & 7 \end{vmatrix} = 16. \end{aligned}$$ ■

利用行列式按行展开公式,立即可有行列式的另外三条性质.

性质 3　如某一行的元素全为 0,则行列式的值为 0.

性质 4　如果某一行的元素有公因数 k, **则** k **可以提到行列式记号之外.**

例如第 3 行有公因数,可写成

$$\begin{vmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{vmatrix} = 3 \begin{vmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 1 & 2 & 3 \end{vmatrix}.$$

性质 5　如某一行的元素各为两个数的和,那么这个行列式就等于两个行列式的和,而这两个行列式除这一行以外全与原来行列式的对应的行一样.

例如第 1 行是两个数的和,就有

$$\begin{vmatrix} a_1 + a_2 & b_1 + b_2 & c_1 + c_2 \\ m & n & p \\ x & y & z \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ m & n & p \\ x & y & z \end{vmatrix} + \begin{vmatrix} a_2 & b_2 & c_2 \\ m & n & p \\ x & y & z \end{vmatrix}.$$

性质 6　将行列式的两行互相调换,行列式的值只改变正负号.

例如 1,3 两行互换,有

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} a_3 & b_3 & c_3 \\ a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \end{vmatrix}.$$

关于性质 6,可以用行列式定义把其展开成六项的代数和之后来检验其正确.利用性质 6,又可推导出行列式的 3 条性质.

性质 7 如行列式中有两行元素对应相同,那么行列式的值为 0.

假若 D 中第 1 行与第 2 行元素对应相等,互换 1,2 两行得到行列式 D_1,一方面由性质 6 应有: $D = -D_1$;另一方面由于 1,2 两行相同,互换后的 D_1 其实就是原先的 D.因此

$$D = -D_1 = -D.$$

从而有: $D = 0$.

性质 8 如行列式中有两行元素对应成比例,则行列式的值为 0.

例如, D 中 2,3 两行成比例,利用性质 4 和 7 就有

$$\begin{vmatrix} a & b & c \\ m & n & p \\ km & kn & kp \end{vmatrix} = k \begin{vmatrix} a & b & c \\ m & n & p \\ m & n & p \end{vmatrix} = 0.$$

性质 9 若把行列式某行的 k 倍加至另一行,行列式的值不变.

假如把 D 中第 1 行的 k 倍加至第 2 行,利用性质 5 和 8 就可得到

$$\begin{vmatrix} a & b & c \\ ka+m & kb+n & kc+p \\ x & y & z \end{vmatrix} = \begin{vmatrix} a & b & c \\ ka & kb & kc \\ x & y & z \end{vmatrix} + \begin{vmatrix} a & b & c \\ m & n & p \\ x & y & z \end{vmatrix}$$

$$= \begin{vmatrix} a & b & c \\ m & n & p \\ x & y & z \end{vmatrix}.$$

这条性质在行列式计算中非常有用,当我们用按行展开公式计算行列式的值时,为了减少计算工作量.在展开之前一般先用性质 9,把某 i 行的 k 倍加至第 j 行,使得某一列有较多的零,然后再展开,这样计算比较简捷,参看下面的例 7、例 8 等.

性质 10 某行的每个元素与另一行相应元素的代数余子式的乘积之和等于 **0**.

例 4 以二阶行列式为例验证性质 **10**,如

$$D = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}.$$

按代数余子式的定义,有

$$A_{11} = 4,\ A_{12} = -3,\ A_{21} = -2,\ A_{22} = 1.$$

利用行列式按行展开公式,显然有

$$\begin{aligned} D &= a_{11}A_{11} + a_{12}A_{12} = 1\times 4 + 2\times(-3) = -2, \\ &= a_{21}A_{21} + a_{22}A_{22} = 3\times(-2) + 4\times 1 = -2. \end{aligned}$$

也容易验证

$$a_{11}A_{21} + a_{12}A_{22} = 1\cdot(-2) + 2\cdot 1 = 0,$$

$$a_{21}A_{11} + a_{22}A_{12} = 3\cdot 4 + 4\cdot(-3) = 0.$$

即第 1 行元素与第 2 行的代数余子式乘积之和,第 2 行元素与第 1 行的代数余子式乘积之和均为 0. ■

* **例 5** 对三阶行列式,证明:第 1 行的元素与第 2 行的代数余子式相应乘积之和为 0.

即
$$a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23} = 0$$

证 构造一个第 1 行与第 2 行一样的行列式 D(即 $a_{11} = a_{21}$, $a_{12} = a_{22}$, $a_{13} = a_{23}$).

一方面,由性质 7 知 $D = 0$;

另一方面,用性质 2 按第 2 行展开有

$$D = a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23},$$

把 $a_{11}=a_{21},\cdots$及 $D=0$ 代入上式,得

$$a_{11}A_{21}+a_{12}A_{22}+a_{13}A_{23}=a_{21}A_{21}+a_{22}A_{22}+a_{23}A_{23}=D=0,$$

即有第 1 行元素与第 2 行代数余子式乘积之和等于 0 . ■

现在综合运用行列式的这些性质来处理行列式问题 .

例 6 计算行列式的值

$$\begin{vmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{vmatrix}$$

解 注意到行列式中第 1 列有两个 0,可按第 1 列展开来计算 .这就有

$$D=a_{11}A_{11}+a_{21}A_{21}+a_{31}A_{31}=a_{11}A_{11},$$

现在

$$a_{11}=a,\quad A_{11}=(-1)^{1+1}\begin{vmatrix} d & e \\ 0 & f \end{vmatrix}=df,$$

所以,
$$D=adf.$$
■

从左上至右下的对角线称为行列式的主对角线,主对角线下(上)方元素全是 0 的行列式叫上(下)三角行列式 .由例 6 可看出:上三角行列式的值等于其主对角线元素的乘积 .

例 7 计算行列式的值

$$D=\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}.$$

解 在前面讲行列式定义时,我们已经用直接计算的方法算过此题,现在利用性质来解,希望读者通过比较能体会到在行列式计算中运用行列式性质的重要性,要学会用性质来解题 .

法 1 把第 1 列的 - 2 倍, - 3 倍分别加至第 2 列及第 3 列,即有

$$\begin{vmatrix}1&2&3\\4&5&6\\7&8&9\end{vmatrix}=\begin{vmatrix}1&0&0\\4&-3&-6\\7&-6&-12\end{vmatrix},$$

由于 2、3 **两列成比例,故** $D=0$. ■

法 2　把第 2 行的 - 1 倍加至第 3 行,再把第 1 **行的** - 1 **倍加至第** 2 **行,即有**

$$\begin{vmatrix}1&2&3\\4&5&6\\7&8&9\end{vmatrix}=\begin{vmatrix}1&2&3\\3&3&3\\3&3&3\end{vmatrix},$$

由于 2、3 **两行相同,故** $D=0$. ■

* **法 3　把 7**,8,9 **分别改写成** 8 - 1,10 - 2,12 - 3,**然后把行列式拆开为两个行列式,即有**

$$\begin{vmatrix}1&2&3\\4&5&6\\7&8&9\end{vmatrix}=\begin{vmatrix}1&2&3\\4&5&6\\8-1&10-2&12-3\end{vmatrix}$$

$$=\begin{vmatrix}1&2&3\\4&5&6\\8&10&12\end{vmatrix}+\begin{vmatrix}1&2&3\\4&5&6\\-1&-2&-3\end{vmatrix},$$

由于这两个行列式中分别有两行成比例其值都为 0,**而有** $D=0$.

■

例 8　计算行列式

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

解　利用 $a_{21}=0$,**把第** 1 **行的** - 3 **倍加至第** 3 **行,再按第** 1 **列展开 就有**

$$D=\begin{vmatrix}1&-2&-1\\0&3&2\\0&7&2\end{vmatrix}=\begin{vmatrix}3&2\\7&2\end{vmatrix}=-8.$$

我们也可以把第 1 列的 2 倍加至第 2 列,再把第 1 列加至第 3 列,然后按第 1 行展开,就有

$$D=\begin{vmatrix}1&0&0\\0&3&2\\3&7&2\end{vmatrix}=\begin{vmatrix}3&2\\7&2\end{vmatrix}=-8.$$ ■

当然,亦可利用性质 9 按其它的行或列展开计算,但是选择行列的原则应以计算方便简捷为主.显然,本题的解法比起例题 2 中的解法来要方便多了.

例 9 计算行列式

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

解 把各行都加至第 1 行,并提出公因式 $a+3$,然后把第 1 行的 -1 倍分别加至 2、3 行,得到上三角行列式,再用例 6 即可.也就是

$$D=\begin{vmatrix}a+3&a+3&a+3\\1&1+a&1\\1&1&1+a\end{vmatrix}=(a+3)\begin{vmatrix}1&1&1\\1&1+a&1\\1&1&1+a\end{vmatrix}$$

$$=(a+3)\begin{vmatrix}1&1&1\\0&a&0\\0&0&a\end{vmatrix}=a^2(a+3).$$ ■

* **解法 2** 利用性质 5,把 D 的每一项都看成是两个数的和,那么 D 可拆成 8 个行列式的和.在这 8 个行列式中,其中有 4 个"两行一样,其值为零",不为零的 4 个行列式或是上(下)三角行列式,或是易计算的.这样把 D 拆开以后立即可得到 D 的值.即

$$D=\begin{vmatrix}1+a & 1+0 & 1+0\\ 1+0 & 1+a & 1+0\\ 1+0 & 1+0 & 1+a\end{vmatrix}$$

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

$$=a^2+a^2+a^2+a^3=a^2(a+3).$$ ■

例 10 **解方程求** x **的值**

$$\begin{vmatrix}3-x & -1 & 1\\ -1 & 5-x & -1\\ 1 & -1 & 3-x\end{vmatrix}=0.$$

解 **把** 2、3 **行均加至第** 1 **行并提出公因式** $3-x$,**即**

$$\textbf{原式}=\begin{vmatrix}3-x & 3-x & 3-x\\ -1 & 5-x & -1\\ 1 & -1 & 3-x\end{vmatrix}$$

$$=(3-x)\begin{vmatrix}1 & 1 & 1\\ -1 & 5-x & -1\\ 1 & -1 & 3-x\end{vmatrix}$$

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

$$=(3-x)(6-x)(2-x)=0.$$

所以 $x_1=2,\ x_2=3,\ x_3=6.$ ■

例 11 **解方程求** x

$$\begin{vmatrix}2-x & 2 & -2\\ 2 & 5-x & -4\\ -2 & -4 & 5-x\end{vmatrix}=0.$$

解 **把第** 3 **行加至第** 2 **行并提出公因式** $1-x$,**即**

$$\text{原式}=\begin{vmatrix}2-x & 2 & -2\\ 0 & 1-x & 1-x\\ -2 & -4 & 5-x\end{vmatrix}$$

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

$$=(1-x)\begin{vmatrix}2-x & 2 & -4\\ 0 & 1 & 0\\ -2 & -4 & 9-x\end{vmatrix}$$

$$=(1-x)\begin{vmatrix}2-x & -4\\ -2 & 9-x\end{vmatrix}$$

$$=(1-x)[(2-x)(9-x)-8]=(1-x)^2(10-x)=0.$$

故 $x=1$(二重根), $x=10$. ■

[*]**例 12　证明**

$$\begin{vmatrix}a_{11}-\lambda & a_{12} & a_{13}\\ a_{21} & a_{22}-\lambda & a_{23}\\ a_{31} & a_{32} & a_{33}-\lambda\end{vmatrix}=-\lambda^3+(a_{11}+a_{22}+a_{33})\lambda^2$$

$$-\left(\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}+\begin{vmatrix}a_{11} & a_{13}\\ a_{31} & a_{33}\end{vmatrix}+\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}\right)\lambda+$$

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

证　把行列式的每一项都看成两个数的和,利用性质 5 把原行列式可拆成 8 个行列式的和.即

$$\text{原式}=\begin{vmatrix}a_{11}-\lambda & a_{12}+0 & a_{13}+0\\ a_{21}+0 & a_{22}-\lambda & a_{23}+0\\ a_{31}+0 & a_{32}+0 & a_{33}-\lambda\end{vmatrix}$$

$$
= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & -\lambda \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & a_{13} \\ a_{21} & -\lambda & a_{23} \\ a_{31} & 0 & a_{33} \end{vmatrix}
$$

$$
+ \begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & -\lambda & 0 \\ a_{31} & 0 & -\lambda \end{vmatrix} + \begin{vmatrix} -\lambda & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} +
$$

$$
\begin{vmatrix} -\lambda & a_{12} & 0 \\ 0 & a_{22} & 0 \\ 0 & a_{32} & -\lambda \end{vmatrix}
$$

$$
+ \begin{vmatrix} -\lambda & 0 & a_{13} \\ 0 & -\lambda & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix}
$$

$$
= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \lambda - \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \lambda + a_{11} \lambda^2
$$

$$
- \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \lambda + a_{22} \lambda^2 + a_{33} \lambda^2 - \lambda^3
$$

$= $右端 . ■

1.2　n 阶行列式

对二、三阶的行列式,可以有多种方式把它拓广为高阶行列式,我们现在采用的方法是数学归纳法,用低阶来定义高阶的行列式 .

根据三阶行列式按行展开的公式(性质 2),引入四阶行列式 D 的定义 .

我们称

$$D=\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$

是一个四阶行列式,它由 4^2 **个数组成,其值规定为**

$$a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}+a_{14}A_{14},$$

其中 A_{1j}**叫做** a_{1j}**的代数余子式,它是** D **中去掉第** 1 **行及第** j **列后的三阶行列式再带上正负号**$(-1)^{1+j}$.

例如, a_{11}**的代数余子式** A_{11}**是**

$$(-1)^{1+1}\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix},$$

而 a_{14}**的代数余子式** A_{14}**是**

$$(-1)^{1+4}\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix}.$$

例 1　求行列式

$$\begin{vmatrix} 1 & 0 & -3 & 7 \\ 0 & 1 & 2 & 1 \\ -3 & 4 & 0 & 3 \\ 1 & -2 & 2 & -1 \end{vmatrix}$$

的代数余子式 A_{12}**和** A_{13}**.**

解　按代数余子式定义,有

$$A_{12}=(-1)^{1+2}\begin{vmatrix} 0 & 2 & 1 \\ -3 & 0 & 3 \\ 1 & 2 & -1 \end{vmatrix}=-\begin{vmatrix} 0 & 2 & 1 \\ 0 & 6 & 0 \\ 1 & 2 & -1 \end{vmatrix}=6,$$

$$A_{13}=(-1)^{1+3}\begin{vmatrix}0&1&1\\-3&4&3\\1&-2&-1\end{vmatrix}=\begin{vmatrix}0&0&1\\-3&1&3\\1&-1&-1\end{vmatrix}=2.\blacksquare$$

需要指出的是,a_{1j}的代数余子式 A_{1j}只与 a_{1j}的位置有关,而与 a_{1j}的大小或正负无关.也就是说,如果只改变行列式中 a_{1j}的数值而其它元素不变,那么代数余子式 A_{1j}的值是不变的.

例如,在例 1 中 $A_{12}=6$,如果把 a_{12}由 0 改变为 3 或 -5 或任一数 x,而其它元素不变,那么新行列式的 a_{12}的代数余子式 A_{12}仍然是 6.

例 2 计算四阶行列式

$$\begin{vmatrix}a&0&0&0\\b&c&0&0\\d&e&f&0\\g&h&i&j\end{vmatrix}.$$

解 根据四阶行列式定义

$$D=a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}+a_{14}A_{14}.$$

现在 $a_{12}=a_{13}=a_{14}=0$,而 $a_{11}=a$.

$$A_{11}=\begin{vmatrix}c&0&0\\e&f&0\\h&i&j\end{vmatrix}=cfj.$$

从而有

$$D=acfj.\qquad\blacksquare$$

由例 2 知,下三角行列式(当 $i<j$ 时,$a_{ij}=0$)的值等于主对角线元素的乘积.

用三阶行列式定义了四阶行列式,类似地,可用四阶行列式定义五阶行列式,一般地用 $n-1$ 阶行列式来定义 n 阶行列式.

我们称

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

为 n 阶行列式,其值规定为

$$
a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}. \qquad (1.3)
$$

其中 A_{1j}**是带有正负号的** $n-1$ **阶行列式**

$$
A_{1j} = (-1)^{1+j}
\begin{vmatrix}
a_{21} & \cdots & a_{2\,j-1} & a_{2\,j+1} & \cdots & a_{2n} \\
a_{31} & \cdots & a_{3\,j-1} & a_{3\,j+1} & \cdots & a_{3n} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn}
\end{vmatrix} \qquad (1.4)
$$

称为 a_{1j}**的代数余子式 .**

n **阶行列式有着与三阶行列式完全一致的** 10 **条性质(证明略去),除去少数题目可以用行列式定义直接计算外,一般地讲,对** n **阶行列式也是要先利用行列式性质把其恒等变形化简后再计算 .**

例 3　证明: n **阶上三角行列式(当** $i>j$ **时,** $a_{ij}=0$**)的值等于主对角线元素的乘积 .即**

$$
D_n =
\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
 & a_{22} & \cdots & a_{2n} \\
0 & & \ddots & \cdots \\
 & & & a_{nn}
\end{vmatrix}
= a_{11}a_{22}\cdots a_{nn}. \qquad (1.5)
$$

证　当 $n=2$ **时,结论显然成立 .假设** $n-1$ **时命题成立,对** D_n **按第一列展开有**

$$
D_n = a_{11}\cdot(-1)^{1+1}
\begin{vmatrix}
a_{22} & \cdots & a_{2n} \\
 & \ddots & \cdots \\
0 & & a_{nn}
\end{vmatrix}.
$$

右端行列式是 $n-1$ 阶上三角行列式,由归纳假设

$$D_n = a_{11}a_{22}\cdots a_{nn}. \qquad \blacksquare$$

特别地,

$$\begin{vmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{vmatrix} = a_1 a_2 \cdots a_n .$$

由于行列式转置后其值不变,因此,下三角行列式的值也等于其主对角线元素的乘积 .

例 4　证明

$$\begin{vmatrix} & & & & a_{1n} \\ & & & a_{2\,n-1} & \\ & & ⋰ & & \\ & a_{n-1\,2} & & & \\ a_{n1} & & & & \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} a_{1n}a_{2\,n-1}\cdots a_{n1} .$$

证　$n=2$ 时,结论显然成立 .假设 $n-1$ 时命题正确,对于 n 阶行列式按第一列展开有

$$D_n = a_{n1}\cdot(-1)^{n+1}\begin{vmatrix} & & & a_{1n} \\ & & a_{2\,n-1} & \\ & ⋰ & & \\ a_{n-1\,2} & & & \end{vmatrix},$$

右端行列式是与 D_n 同类型的 $n-1$ 阶行列式,按归纳假设其值是 $(-1)^{\frac{(n-1)(n-2)}{2}} a_{1n}a_{2\,n-1}\cdots a_{n-1\,2}$.由于

$$(-1)^{n+1}\cdot(-1)^{\frac{(n-1)(n-2)}{2}} = (-1)^{\frac{n^2-n+4}{2}} = (-1)^{\frac{n(n-1)}{2}},$$

可见对于 n 命题成立 . $\blacksquare$

下面利用行列式的性质再给出计算行列式的一些方法和技

巧.

例 5　计算四阶行列式

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

解　观察行列式发现第 3 行数字比较简单且有 0,我们可先用性质 9 对列作恒等变形使第 3 行再出现两个 0.(这只须把第 4 列的 -1 倍加至第 2 列,再把第 3 列的 2 倍加至第 4 列),然后按第 3 行展开.即是

$$D=\begin{vmatrix}2&0&-5&-9\\1&3&0&-6\\0&0&-1&0\\1&-1&-7&-9\end{vmatrix}=(-1)\cdot(-1)^{3+3}\begin{vmatrix}2&0&-9\\1&3&-6\\1&-1&-9\end{vmatrix}$$

$$=3\begin{vmatrix}2&0&3\\1&3&2\\1&-1&3\end{vmatrix}=3\begin{vmatrix}2&0&3\\4&0&11\\1&-1&3\end{vmatrix}$$

$$=3\cdot(-1)\cdot(-1)^{3+2}\begin{vmatrix}2&3\\4&11\end{vmatrix}=30.$$

■

例 6　计算四阶行列式

$$D=\begin{vmatrix}1&2&3&4\\2&3&4&1\\3&4&1&2\\4&1&2&3\end{vmatrix}.$$

解 1　把各行均加至第 1 行,提出公因数 10 之后,再把各列减去第 1 列,按第 1 行展开就可把四阶行列式降阶为三阶行列式.即

$$D=\begin{vmatrix}10&10&10&10\\2&3&4&1\\3&4&1&2\\4&1&2&3\end{vmatrix}=10\begin{vmatrix}1&1&1&1\\2&3&4&1\\3&4&1&2\\4&1&2&3\end{vmatrix}$$

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

$$=10\begin{vmatrix}1&2&-1\\0&-4&0\\-3&-2&-1\end{vmatrix}=-40\begin{vmatrix}1&-1\\-3&-1\end{vmatrix}=160.\qquad\blacksquare$$

解 2　**把** 2,3 **列的** - 1 **倍及第** 4 **列都加至第** 1 **列,第** 1 **列可提出公因数** 4,**再把第** 4 **行加至第** 2 **行,行列式可降为三阶.即**

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

$$=-4\begin{vmatrix}2&3&4\\4&6&4\\4&1&2\end{vmatrix}=-8\begin{vmatrix}2&3&4\\2&3&2\\4&1&2\end{vmatrix}$$

$$=-8\begin{vmatrix}0&0&2\\2&3&2\\4&1&2\end{vmatrix}=-16\begin{vmatrix}2&3\\4&1\end{vmatrix}=160.\qquad\blacksquare$$

例 7　**计算** Vandermonde **行列式**

$$D=\begin{vmatrix}1&1&1\\x_1&x_2&x_3\\x_1^2&x_2^2&x_3^2\end{vmatrix}.$$

解 把第 2 行的 $-x_1$ 倍加至第 3 行,再把第 1 行的 $-x_1$ 倍加至第 2 行,然后按第 1 列展开.

$$D=\begin{vmatrix}1 & 1 & 1\\ 0 & x_2-x_1 & x_3-x_1\\ 0 & x_2^2-x_1x_2 & x_3^2-x_1x_3\end{vmatrix}=\begin{vmatrix}x_2-x_1 & x_3-x_1\\ x_2^2-x_1x_2 & x_3^2-x_1x_3\end{vmatrix}$$

$$=(x_2-x_1)(x_3-x_1)\begin{vmatrix}1 & 1\\ x_2 & x_3\end{vmatrix}$$

$$=(x_2-x_1)(x_3-x_1)(x_3-x_2)=\prod_{1\leq j<i\leq 3}(x_i-x_j).\quad \blacksquare$$

用数学归纳法可以写出这个行列式的一般形式

$$\begin{vmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\\ x_1^2 & x_2^2 & \cdots & x_n^2\\ \cdots & \cdots & & \cdots\\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1}\end{vmatrix}=\prod_{1\leq j<i\leq n}(x_i-x_j). \tag{1.6}$$

如果能观察出某些行列式是 Vandermonde 行列式,那么这个行列式的值就很容易算出.例如,

$$\begin{vmatrix}1 & 1 & 1 & 1\\ 2 & 3 & -1 & 4\\ 4 & 9 & 1 & 16\\ 8 & 27 & -1 & 64\end{vmatrix}$$

$$=(3-2)(-1-2)(4-2)(-1-3)(4-3)(4-(-1))$$

$$=120.$$

例 8 计算五阶行列式(其空白处全是 0)

$$D_5 = \begin{vmatrix} 2 & 1 & & & \\ 1 & 2 & 1 & & \\ & 1 & 2 & 1 & \\ & & 1 & 2 & 1 \\ & & & 1 & 2 \end{vmatrix}.$$

解 1　利用行列式性质把 D_5 化成上三角行列式,即

$$D_5 = \begin{vmatrix} 2 & 1 & & & \\ & \frac{3}{2} & 1 & & \\ & 1 & 2 & 1 & \\ & & 1 & 2 & 1 \\ & & & 1 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 1 & & & \\ & \frac{3}{2} & 1 & & \\ & & \frac{4}{3} & 1 & \\ & & 1 & 2 & 1 \\ & & & 1 & 2 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 1 & & & \\ & \frac{3}{2} & 1 & & \\ & & \frac{4}{3} & 1 & \\ & & & \frac{5}{4} & 1 \\ & & & 1 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 1 & & & \\ & \frac{3}{2} & 1 & & \\ & & \frac{4}{3} & 1 & \\ & & & \frac{5}{4} & 1 \\ & & & & \frac{6}{5} \end{vmatrix}$$

$$= 2 \cdot \frac{3}{2} \cdot \frac{4}{3} \cdot \frac{5}{4} \cdot \frac{6}{5} = 6. \qquad \blacksquare$$

*__解 2　把 D_5 按第 1 行展开,建立递推关系__

$$D_5 = 2D_4 - \begin{vmatrix} 1 & 1 & & \\ & 2 & 1 & \\ & 1 & 2 & 1 \\ & & 1 & 2 \end{vmatrix} = 2D_4 - D_3,$$

把上述关系整理成

$$D_5 - D_4 = D_4 - D_3,$$

那么,递推地运用上述关系式就有

$$D_5 - D_4 = D_4 - D_3 = D_3 - D_2 = D_2 - D_1 = 3 - 2 = 1,$$

因此,

$$D_5 = D_4 + 1$$

再一次递推,得到:

$$D_5 = D_4 + 1 = (D_3 + 1) + 1 = D_2 + 3 = D_1 + 4 = 6. \blacksquare$$

***例 9　证明**

$$\begin{vmatrix} a_1 & -1 & 0 & \cdots & 0 \\ a_2 & x & -1 & \ddots & \cdots \\ a_3 & 0 & x & \ddots & 0 \\ \cdots & \cdots & \ddots & \ddots & -1 \\ a_n & 0 & \cdots & 0 & x \end{vmatrix} = a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n$$

证 1　把第 i 行的 x 倍加至第 $i+1$ 行, i 由 1 至 $n-1$ 依次进行,即有

$$D = \begin{vmatrix} a_1 & -1 & 0 & \cdots & 0 \\ a_1 x + a_2 & 0 & -1 & \ddots & \cdots \\ a_3 & 0 & x & \ddots & 0 \\ \cdots & \cdots & \ddots & \ddots & -1 \\ a_n & 0 & \cdots & 0 & x \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & -1 & 0 & \cdots & 0 \\ a_1 x + a_2 & 0 & -1 & \ddots & \cdots \\ a_1 x^2 + a_2 x + a_3 & 0 & 0 & \ddots & 0 \\ \cdots & \cdots & \ddots & \ddots & -1 \\ a_n & 0 & \cdots & 0 & x \end{vmatrix}$$

$$
= \cdots = \begin{vmatrix} a_1 & -1 & 0 & \cdots & 0 \\ a_1 x + a_2 & 0 & -1 & \ddots & \cdots \\ a_1 x^2 + a_2 x + a_3 & 0 & 0 & \ddots & 0 \\ \cdots & \cdots & \ddots & \ddots & -1 \\ a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n & 0 & \cdots & 0 & 0 \end{vmatrix}
$$

$$
= (a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n) \cdot (-1)^{n+1} \begin{vmatrix} -1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{vmatrix}_{n-1}
$$

$$
= (a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n) \cdot (-1)^{n+1} \cdot (-1)^{n-1}
$$

$$
= a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n . \qquad \blacksquare
$$

证 2　按第 1 行展开,建立递推关系

$$
D_n = a_1 \begin{vmatrix} x & -1 & & & \\ & x & -1 & 0 & \\ & \ddots & & \ddots & \\ & 0 & \ddots & & -1 \\ & & & & x \end{vmatrix} + (-1) \cdot (-1)^{1+2} \begin{vmatrix} a_2 & -1 & 0 & \\ a_3 & x & \ddots & \\ \cdots & & \ddots & -1 \\ a_n & 0 & & x \end{vmatrix}
$$

$$
\begin{aligned}
&= a_1 x^{n-1} + D_{n-1} \\
&= a_1 x^{n-1} + a_2 x^{n-2} + D_{n-2} \\
&= \cdots\cdots \\
&= a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n .
\end{aligned} \qquad \blacksquare
$$

***例 10　证明**

$$\begin{vmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 2 & & \text{0} & \\ 1 & & 3 & & \\ \vdots & \text{0} & & \ddots & \\ 1 & & & & n \end{vmatrix} = - \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \cdot n\,!$$

证　根据数字的特点,可以第 j 列提出公因数 j,然后把每列的 -1 倍加至第 1 列,把行列式恒等变形为上三角行列式.就是

$$D = (n\,!)\begin{vmatrix} 0 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ 1 & 1 & & \text{0} & \\ 1 & & 1 & & \\ \vdots & \text{0} & & \ddots & \\ 1 & & & & 1 \end{vmatrix}$$

$$= (n\,!)\begin{vmatrix} -\frac{1}{2}-\frac{1}{3}-\cdots-\frac{1}{n} & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ 0 & 1 & & \text{0} & \\ 0 & & 1 & & \\ \vdots & \text{0} & & \ddots & \\ 0 & & & & 1 \end{vmatrix}$$

$$= - \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \cdot n\,! \qquad \blacksquare$$

在本节的最后,我们把行列式按行展开公式推广到按几行来展开,下面给出的是两种特殊情况下的公式(通常称为拉普拉斯(Laplace)展开式).

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} & 0 & \cdots & 0 \\ \cdots & & \cdots & \cdots & & \cdots \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1n} & b_{11} & \cdots & b_{1m} \\ \cdots & & \cdots & \cdots & & \cdots \\ c_{m1} & \cdots & c_{mn} & b_{m1} & \cdots & b_{mm} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1m} \\ \cdots & & \cdots \\ b_{m1} & \cdots & b_{mm} \end{vmatrix} . \tag{1.7}$$

$$\begin{vmatrix} 0 & \cdots & 0 & a_{11} & \cdots & a_{1n} \\ \cdots & & \cdots & \cdots & & \cdots \\ 0 & \cdots & 0 & a_{n1} & \cdots & a_{nn} \\ b_{11} & \cdots & b_{1m} & c_{11} & \cdots & c_{1n} \\ \cdots & & \cdots & \cdots & & \cdots \\ b_{m1} & \cdots & b_{mm} & c_{m1} & \cdots & c_{mn} \end{vmatrix} = (-1)^{mn} \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1m} \\ \cdots & & \cdots \\ b_{m1} & \cdots & b_{mm} \end{vmatrix} . \tag{1.8}$$

例如

$$\begin{vmatrix} 1 & 2 & a & b & c \\ 3 & 4 & d & e & f \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \cdot \begin{vmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{vmatrix} = -2,$$

$$\begin{vmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 & 4 \\ 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 & 4 \\ 0 & 0 & 1 & 5 & 6 \end{vmatrix} = (-1)^{2\times 3} \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \cdot \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = -2 .$$

1.3 克莱姆(Cramer)法则

利用行列式的性质我们可以得到方程组解的一些重要定理.

对于含有 n 个未知数 $x_1, x_2, \cdots, x_n$ 的 n 个线性方程所构成的非齐次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots, \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \tag{1.9}$$

及相应的齐次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots, \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0. \end{cases} \tag{1.10}$$

我们称 $a_{ij}(i=1,2,\cdots,n;\ j=1,2,\cdots,n)$为方程组的系数, $b_j(j=1,2,\cdots,n)$是常数项, 并称

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \hline \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{1.11}$$

是方程组的系数行列式.

如果有 n 个数 $c_1, c_2, \cdots, c_n$, 当把 $x_1 = c_1, x_2 = c_2, \cdots, x_n = c_n$ 代入到(1.9)的每个方程使其每个都成为恒等式, 我们就称这组数是线性方程组(1.9)的一个解. 所谓解方程组就是要找出这个方程组的所有的解.

克莱姆法则　如果线性方程组(1.9)的系数行列式 $\Delta \neq 0$, 则

方程组有唯一解,且解是

$$x_1=\frac{\Delta_1}{\Delta},\quad x_2=\frac{\Delta_2}{\Delta},\quad\cdots,\quad x_n=\frac{\Delta_n}{\Delta}. \tag{1.12}$$

其中 $\Delta_j(j=1,2,\cdots,n)$是把 Δ 中第 j 列系数换成常数项 $b_1,b_2,\cdots,b_n$ 后所构成的行列式.

克莱姆法则不仅给出了方程组(1.9)有唯一解的条件,而且还给出了求解的方法.关于这个法则的证明我们将在下一章中用矩阵运算的形式给出一个较简捷的论述.

例 1 用克莱姆法则解方程组

$$\begin{aligned} x_1+x_2+x_3&=1,\\ 2x_1-x_2+3x_3&=-2,\\ 4x_1+x_2+9x_3&=4.\end{aligned}$$

解 系数行列式是 Vandermonde 行列式,其值

$$\Delta=\begin{vmatrix}1&1&1\\2&-1&3\\4&1&9\end{vmatrix}=(-1-2)(3-2)(3-(-1))=-12,$$

不为 0,因此方程组有唯一解.由

$$\Delta_1=\begin{vmatrix}1&1&1\\-2&-1&3\\4&1&9\end{vmatrix}$$

$$=(-1-(-2))(3-(-2))(3-(-1))=20,$$

$$\Delta_2=\begin{vmatrix}1&1&1\\2&-2&3\\4&4&9\end{vmatrix}=-20,\quad \Delta_3=\begin{vmatrix}1&1&1\\2&-1&-2\\4&1&4\end{vmatrix}=-12.$$

所以,方程组的唯一解是

$$x_1=-\frac{5}{3},\quad x_2=\frac{5}{3},\quad x_3=1.$$ ■

当方程组(1.9)的系数行列式 $\Delta=0$ 时,方程组肯定没有唯一

解,否则将与克莱姆法则相矛盾,此时的方程组可能是无解的,也可能是有无穷多个解.例如

$$\begin{cases} x+y=1, \\ x+y=3, \end{cases} \text{和} \quad \begin{cases} x+y=1, \\ 2x+2y=2, \end{cases}$$

前者无解而后者有无穷多组解.

对于齐次方程组(1.10),$x_1=x_2=\cdots=x_n=0$ 一定是它的解,这个解叫做齐次方程组的零解,对齐次方程组我们感兴趣的是它是否有一组不全为 0 的数 $k_1,k_2,\cdots,k_n$ 能够成为它的解?这样的解叫做齐次方程组的非零解.

不一定每一个齐次方程组都有它的非零解,例如

$\begin{cases} x+y=0, \\ x-y=0, \end{cases}$ 只有零解,并没有非零解.

而 $\begin{cases} x+y=0, \\ 2x+2y=0, \end{cases}$ 有许多非零解,像 $\begin{cases} x=1, \\ y=-1, \end{cases}$ $\begin{cases} x=2, \\ y=-2, \end{cases}$ $\cdots$.

把克莱姆法则用到齐次方程组,就有

定理 1.1 如齐次方程组(1.10)的系数行列式 $\Delta\neq 0$,则方程组只有零解

$$x_1=x_2=\cdots=x_n=0. \qquad \blacksquare$$

定理 1.2 如齐次方程组(1.10)有非零解,则系数行列式

$$\Delta=0. \qquad \blacksquare$$

例 2 已知方程组

$$\begin{cases} x_1-x_2=\lambda x_1, \\ -x_1+2x_2-x_3=\lambda x_2, \\ -x_2+x_3=\lambda x_3, \end{cases}$$

有非零解,求参数 λ.

解 方程组是一个齐次方程组

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

因为它有非零解,其系数行列式为 0, **即**

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

把 2、3 **行均加至第** 1 **行,即有**

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

$$=-\lambda\begin{vmatrix}1 & 1 & 1\\ 0 & 3-\lambda & 0\\ 0 & -1 & 1-\lambda\end{vmatrix}=-\lambda(3-\lambda)(1-\lambda)=0$$

$$\therefore\qquad \lambda_1=0,\ \lambda_2=1,\ \lambda_3=3.$$ ■

* **例 3　求平面上经过两点** $A(3,-2)$, $B(-1,4)$ **的直线方程**.

解　设直线方程是: $ax+by+c=0$, **点** $P(x,y)$ **是直线上任一点,且点** A, B **也都在直线上,因此有**

$$\begin{aligned}ax+by+c&=0,\\ 3a-2b+c&=0,\\ -a+4b+c&=0,\end{aligned}$$

作为直线方程 a, b, c **不可能全是** 0, **所以上述关于** a, b, c **的齐次方程组必有非零解,从而系数行列式为** 0, **即**

$$\begin{vmatrix}x & y & 1\\ 3 & -2 & 1\\ -1 & 4 & 1\end{vmatrix}=0,$$

按第 1 行展开,得到: $-6x-4y+10=0$,
由于点 $P(x,y)$的坐标要满足关系: $3x+2y-5=0$
这就是所求的直线方程 . ■

习　题　1

1.1　计算行列式

(1) $\begin{vmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{vmatrix}$;　(2) $\begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix}$;

(3) $\begin{vmatrix} 4 & -3 & 5 \\ 2 & -3 & 6 \\ 5 & 2 & -1 \end{vmatrix}$;　(4) $\begin{vmatrix} 1 & 5 & 25 \\ 1 & 7 & 49 \\ 1 & 8 & 64 \end{vmatrix}$.

1.2　证明行列式的性质

(1) $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \\ a_2 & b_2 & c_2 \end{vmatrix} = -\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$.

(2) $\begin{vmatrix} a_1+x & b_1+y & c_1+z \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} x & y & z \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$.

1.3　求 x 的值

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

(3) $\begin{vmatrix} 1-x & 1 & 1 \\ 1 & 1-x & 1 \\ 1 & 1 & 1-x \end{vmatrix}=0$, (4) $\begin{vmatrix} 1+x & 2 & -2 \\ -1 & x & 1 \\ -2 & 2 & 1+x \end{vmatrix}=0$.

1.4 求第 2 列元素的代数余子式

(1) $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix}$, (2) $\begin{vmatrix} 1 & -1 & 2 & -3 \\ -1 & 0 & 0 & -1 \\ 0 & 3 & 2 & -3 \\ 1 & 2 & -1 & 2 \end{vmatrix}$.

1.5 计算行列式的值

(1) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{vmatrix}$, (2) $\begin{vmatrix} 2 & 1 & 4 & 1 \\ 3 & -1 & 2 & 1 \\ 1 & 2 & 3 & 2 \\ 5 & 0 & 6 & 2 \end{vmatrix}$,

(3) $\begin{vmatrix} 0 & 0 & 1 & 2 \\ 1 & 2 & 3 & 0 \\ 3 & 4 & 0 & 0 \\ 5 & 0 & 0 & 0 \end{vmatrix}$, (4) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 3 & -2 \\ 1 & 1 & 9 & 4 \\ 1 & -1 & 27 & -8 \end{vmatrix}$.

1.6 用行列式解方程组

(1) $\begin{cases} 2x+3y=5, \\ 3x-4y=-18. \end{cases}$ (2) $\begin{cases} x\cos\theta-y\sin\theta=1, \\ x\sin\theta+y\cos\theta=0. \end{cases}$

(3) $\begin{cases} x_1+2x_2-x_3=0, \\ 2x_1-3x_2+x_3=0, \\ 4x_1-x_2+x_3=0. \end{cases}$ (4) $\begin{cases} x_1+4x_2+x_3=1, \\ x_1+x_2+x_3=0, \\ 2x_1+3x_3=0. \end{cases}$

1.7 已知齐次方程组有非零解,求参数 λ

$$(1)\quad \begin{cases} (5-\lambda)x_1 - 4x_2 - 7x_3 = 0, \\ -6x_1 + (7-\lambda)x_2 + 11x_3 = 0, \\ 6x_1 - 6x_2 - (10+\lambda)x_3 = 0 . \end{cases}$$

$$^{*}(2)\quad \begin{cases} 2x_1 - x_2 + 2x_3 = \lambda x_1, \\ 5x_1 - 3x_2 + 3x_3 = \lambda x_2, \\ -x_1 - 2x_3 = \lambda x_3 . \end{cases}$$

1.8　证明

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

$$(2)\quad \begin{vmatrix} b_1+c_1 & c_1+a_1 & a_1+b_1 \\ b_2+c_2 & c_2+a_2 & a_2+b_2 \\ b_3+c_3 & c_3+a_3 & a_3+b_3 \end{vmatrix} = 2\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} .$$

1.9　计算行列式的值

$$(1)\quad \begin{vmatrix} 2 & 3 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{vmatrix}, \qquad (2)\quad \begin{vmatrix} -1 & -2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ * & * & -1 & 0 & 0 \\ * & * & 2 & -1 & 0 \\ * & * & 3 & 2 & -1 \end{vmatrix},$$

(3) $\begin{vmatrix} 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ a & a & 0 & * & * \\ 0 & a & a & * & * \\ 0 & 0 & a & * & * \end{vmatrix}$, (4) $\begin{vmatrix} * & * & * & a & a & 0 \\ * & * & * & 0 & 1 & a \\ * & * & * & 0 & 0 & a \\ b & 0 & 0 & 0 & 0 & 0 \\ 1 & b & 0 & 0 & 0 & 0 \\ 2 & 1 & b & 0 & 0 & 0 \end{vmatrix}$

*1.10 **如果行列式** D **的元素满足**: $a_{ij} = -a_{ji}$, $i, j = 1, 2, \cdots, n$.**则称** D **是反对称行列式** .**证明当** n **为奇数时**, $D = 0$.

*1.11 **计算下列** n **阶行列式**

(1) $\begin{vmatrix} a & b & & & \\ & a & b & & \\ & & \ddots & \ddots & \\ & & & \ddots & b \\ b & & & & a \end{vmatrix}$, (2) $\begin{vmatrix} 1 & 2 & 3 & \cdots & n \\ -1 & 0 & 3 & \cdots & n \\ -1 & -2 & 0 & \cdots & n \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ -1 & -2 & -3 & \cdots & 0 \end{vmatrix}$,

(3) $\begin{vmatrix} 1 & 2 & 3 & \cdots & n \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{vmatrix}$.

第 2 章　矩　　阵

矩阵在高等数学中是一个极重要且应用广泛的概念,它是线性代数的核心,从本章起以后的每一章都要运用矩阵的概念、运算及理论 .

我们把 $m\times n$ 个数排成如下 m 行 n 列的一个表格

$$\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \hline a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}$$

称为是一个 $m\times n$ 矩阵 .

例如

$$\begin{matrix} 1 & -1 & 0 \\ 2 & 3 & 5 \end{matrix}$$

是一个 2×3 矩阵,

$$\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

是一个 3×3 矩阵 .

数 a_{ij}, $i=1,2,\cdots,m$, $j=1,2,\cdots,n$,称为矩阵的元素,其中 i 叫元素 a_{ij}的行指标,它表示这个元素在矩阵中第 i 行, j 称为列指标它表示这个元素在矩阵中第 j 列 .习惯上用大写字母 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 等表示矩阵,有时也记成

$$\boldsymbol{A} = (a_{ij})_{m\times n}$$

我们要注意,矩阵与行列式是不同的,行列式要求行数与列数相同,而矩阵不要求行数 m 等于列数 n;行列式表示的是一个数值,而矩阵仅是由 mn 个数所排成的一个表格.在下面的学习中,我们应注意它们之间的联系与区别不要相混淆.

2.1 矩阵的概念及运算

假如某种动物的生长发育对蛋白质,矿物质,维生素的摄入特别敏感,对 5 种饲料 c_1, c_2, c_3, c_4, c_5 每 500 g 所含营养成份的调查分析有如下表:

	蛋白质/ g	矿物质/ g	维生素/ mg
c_1	0.3	0.1	0.05
c_2	2.0	0.05	0.1
c_3	1.0	0.02	0.02
c_4	0.6	0.2	0.2
c_5	1.8	0.05	0.08

对这份饲料营养成份表,我们可写成如下形式

$$\begin{matrix} 0.3 & 0.1 & 0.05 \\ 2.0 & 0.05 & 0.1 \\ 1.0 & 0.02 & 0.02 \\ 0.6 & 0.2 & 0.2 \\ 1.8 & 0.05 & 0.08 \end{matrix},$$

其中第 2 行第 1 列的数 2.0 就是表示第 2 种饲料 c_2 每 500g 所含蛋白质(第 1 种营养成份)的数量,其它数字的含义是类似地.

数学上把这种形式叫做矩阵,矩阵中的数字排成了 5 行 3 列,通常称其为 5×3 矩阵.

又如某工厂 A 可用 m 种原料加工成 s 种产品,我们用 a_{ij} 表

示生产一件第 j 种产品时第 i 种原料的用量,那么 $m\times s$ 矩阵

$$
\begin{array}{cccccc}
a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1s} \\
a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2s} \\
\hline
a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{is} \\
\hline
a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{ms}
\end{array}
$$

中的第 j 列就是该厂第 j 种产品的配方,而第 i 行是每件产品所需第 i 种原料的用量 .

再如在通讯领域,电报局用长短两个音节的各种组合发出讯号,我们可用 0 和 1 表示这两个基本音节,倘若规定用 0001 表示数字 1,0010 表示数字 2, …,1001 表示数字 9,那么,数字 1 到 9 可用 9×4 矩阵

$$
\begin{array}{cccc}
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
\hline
1 & 0 & 0 & 1
\end{array}
$$

表示,其中第 i 行的 4 个数字就是数字 i 的电报号码 .这样,要发电报就可先把电文翻译成电报号码,然后按号码(长短音)一组一组发送 .

在数学上,在实际应用中,矩阵都是一个很重要的概念,如果矩阵仅仅作为一个表格,其作用就发挥的不充分 .重要的是,对矩阵可以定义一些运算,正是这些运算使矩阵成为一个重要且实用的工具 .

如果 $\boldsymbol{A}=(a_{ij})$与 $\boldsymbol{B}=(b_{ij})$都是 $m\times n$ 矩阵,则称 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是同型的矩阵 .例如,

$$
\begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{matrix} \quad 与 \quad \begin{matrix} a & b & c \\ d & e & f \end{matrix}
$$

都是 2×3 矩阵,它们是同型的 .

两个同型的矩阵 $\boldsymbol{A}=(a_{ij})$与 $\boldsymbol{B}=(b_{ij})$,如 " i, j 都有 $a_{ij}=b_{ij}$,则称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相等,记作 $\boldsymbol{A}=\boldsymbol{B}$.

假如上例的两个矩阵相等,也就等于说:

$$a=1, b=2, c=3, d=4, e=5, f=6 .$$

又若某方程组的解是: $x_1=1, x_2=3, x_3=5$.那么可用矩阵 $\boldsymbol{X}=\boldsymbol{B}$ 来表示这个解,其中

$$\boldsymbol{X}=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \boldsymbol{B}=\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}.$$

如果一个矩阵的所有元素都是 0,

$$\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

则称这个矩阵是零矩阵,可简记为 $\boldsymbol{O}$.

当 $m=n$ 时,矩阵 $\boldsymbol{A}$ 称为 $\boldsymbol{n}$ 阶矩阵或叫 $\boldsymbol{n}$ 阶方阵 .例如

$$\begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}, \quad \begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}, \quad \begin{pmatrix} 2 & -1 & 3 \\ -1 & 0 & 5 \\ 3 & 5 & 6 \end{pmatrix},$$

等都是 3 阶矩阵 .

对 1×1 矩阵(a)看作与数 a 相同 .

n 阶方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$的元素所构成的行列式

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

称为 n 阶矩阵 A 的行列式,记成$|A|$或 $\det A$.

方阵 $A=O$ 及矩阵 A 的行列式$|A|=0$ 是一个易于混淆的概念,前者表示矩阵 A 的所有元素都是 0,而$|A|=0$ 并不能得到 A 的元素为 0 .例如,可构造

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

都有$|A|=0$,但 A 不是零矩阵,即 $A\neq O$.怎样来证明 $A=O$ 或证明$|A|=0$?希望读者在下面的学习中能细心揣摩 .

现在来学习矩阵的运算及其性质:

1) 加法

定义 2.1 设 $A=(a_{ij})$, $B=(b_{ij})$是两个 $m\times n$ 矩阵,则矩阵

$$C=(c_{ij})_{m\times n} > (a_{ij}+b_{ij})$$

称为矩阵 A 与 B 的和,记为 $C=A+B$.

例如,

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

则 $$A+B=\begin{pmatrix}1&2&-1\\-2&5&6\end{pmatrix}.$$

由定义,同型的矩阵可以相加,且矩阵的加法就是对应元素相加 .由于数的加法有交换律和结合律,因此矩阵的加法满足:

(1) $A+B=B+A$; (交换律)

(2) $(A+B)+C=A+(B+C)$; (结合律)

(3) $A+O=O+A=A$.

2) 数乘

定义 2.2 设 $A=(a_{ij})_{m\times n}$, k 是一个常数,则矩阵$(ka_{ij})_{m\times n}$ 称为数 k 与矩阵 A 的数乘,记为 kA .

例如,$A=\begin{pmatrix}2&-3&1\\0&4&-2\end{pmatrix}$ 则 $3A=\begin{pmatrix}6&-9&3\\0&12&-6\end{pmatrix}$.

由定义,数 k 乘矩阵 $\boldsymbol{A}$ 就是将 $\boldsymbol{A}$ 的每个元素都乘以 k, 作为数字乘法有结合律,对加法是可以分配的,因此数乘矩阵满足:

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

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

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

(4) $1\boldsymbol{A}=\boldsymbol{A}$, $0\boldsymbol{A}=\boldsymbol{O}$.

利用数乘运算, 我们可定义一个矩阵的负矩阵,从而可规定同型矩阵的减法 .

设 $\boldsymbol{A}=(a_{ij})_{m\times n}$,记

$$-\boldsymbol{A}=(-a_{ij})=(-1)\boldsymbol{A}$$

称 $-\boldsymbol{A}$ 是 $\boldsymbol{A}$ 的负矩阵 .显然

$$\boldsymbol{A}+(-\boldsymbol{A})=0$$

由此规定: $\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{A}+(-\boldsymbol{B})$

例 1 如 $2\boldsymbol{A}+\boldsymbol{X}=\boldsymbol{B}-2\boldsymbol{X}$, 其中

$$\boldsymbol{A}=\begin{pmatrix}1&-2&0\\4&3&5\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}8&2&6\\5&3&4\end{pmatrix}$$

求矩阵 $\boldsymbol{X}$.

解 移项,整理矩阵方程为:

$$\boldsymbol{X}=\frac{1}{3}(\boldsymbol{B}-2\boldsymbol{A})$$

把 $\boldsymbol{A}$, $\boldsymbol{B}$ 代入,得

$$\boldsymbol{X}=\frac{1}{3}\left[\begin{pmatrix}8&2&6\\5&3&4\end{pmatrix}-\begin{pmatrix}2&-4&0\\8&6&10\end{pmatrix}\right]$$

$$=\frac{1}{3}\begin{pmatrix}6&6&6\\-3&-3&-6\end{pmatrix}=\begin{pmatrix}2&2&2\\-1&-1&-2\end{pmatrix}. \qquad \blacksquare$$

例 2 设 $\boldsymbol{A}=\begin{pmatrix}1&3\\2&4\end{pmatrix}$, 求 $|5\boldsymbol{A}|$.

解 由数乘矩阵定义

$$5\boldsymbol{A}=\begin{pmatrix}5 & 15\\10 & 20\end{pmatrix},$$

那么

$$|5\boldsymbol{A}|=\begin{vmatrix}5 & 15\\10 & 20\end{vmatrix}=5^2\begin{vmatrix}1 & 3\\2 & 4\end{vmatrix}$$

$$=5^2|\boldsymbol{A}|=-50. \qquad \blacksquare$$

一般地,如 $\boldsymbol{A}$ 是 n 阶矩阵,则

$$|k\boldsymbol{A}|=k^n|\boldsymbol{A}|. \tag{2.1}$$

这是因为 $k\boldsymbol{A}$ 中每一行都有公因数 k,作为行列式 $|kA|$ 每行可提出一个公因数 k,故一共可提出 n 个.

作为矩阵,当所有的元素都有公因数 k 时才可把 k 提到矩阵符号之外,这是矩阵与行列式容易混淆的又一点.

例如

$$\begin{pmatrix}2 & -6\\4 & 4\end{pmatrix}=2\begin{pmatrix}1 & -3\\2 & 2\end{pmatrix}\neq 4\begin{pmatrix}1 & -3\\1 & 1\end{pmatrix},$$

而

$$\begin{vmatrix}2 & -6\\4 & 4\end{vmatrix}=2\begin{vmatrix}1 & -3\\4 & 4\end{vmatrix}=8\begin{vmatrix}1 & -3\\1 & 1\end{vmatrix}.$$

例 3 如三阶行列式 $|\boldsymbol{A}|=2$,二阶行列式 $|\boldsymbol{B}|=-1$,求行列式

$$\begin{vmatrix}-\boldsymbol{A} & O\\ * & 3\boldsymbol{B}\end{vmatrix}$$

的值.

解 由拉普拉斯展开式(1.7),

$$\text{原式}=|-\boldsymbol{A}|\cdot|3\boldsymbol{B}|$$

$$=(-1)^3|\boldsymbol{A}|\cdot 3^2|\boldsymbol{B}|=18. \qquad \blacksquare$$

3) 乘法

作为矩阵乘法,我们先来看一个例子,大家知道:如果从甲地

到乙地有 3 条路,从乙地到丙地有 2 条路,那么由甲经乙到丙就有 3×2 条路线 .现看下列图表

	y_1	y_2
x_1	a_{11}	a_{12}
x_2	a_{21}	a_{22}
x_3	a_{31}	a_{32}

	z_1	z_2	z_3	z_4
y_1	b_{11}	b_{12}	b_{13}	b_{14}
y_2	b_{21}	b_{22}	b_{23}	b_{24}

其中 a_{ij}表示由 x_i 到 y_i 路的条数, b_{jk}表示由 y_j 到 z_k 的路的条数 .显然,由 x_1 经 y 到 z_1 的路共有: $a_{11}b_{11}+a_{12}b_{21}$(条),它是由 x_1 经 y_1 到 z_1 及 x_1 经 y_2 到 z_1 路的总和 .一般地,由 x_i 经 y 到 z_k 的路共有: $a_{i1}b_{1k}+a_{i2}b_{2k}$(条) .

我们用矩阵描述这种情况, $\boldsymbol{A}$, $\boldsymbol{B}$ 分别表示由 x 到 y,由 y 到 z 的道路情况 .即

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{pmatrix},$$

用矩阵 $\boldsymbol{C}$ 表示由 x 经 y 到 z 的道路情况

$$\boldsymbol{C}=\begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \end{pmatrix},$$

那么,由 x_i 到 z_j(经 y)的路共有 c_{ij}条,而

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}$$

这样的矩阵 $\boldsymbol{C}$ 就称为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积,记为 $\boldsymbol{AB}=\boldsymbol{C}$.

定义 2.3 设 $\boldsymbol{A}=(a_{ij})_{m\times s}$, $\boldsymbol{B}=(b_{ij})_{s\times n}$,那么矩阵 $\boldsymbol{C}=(c_{ij})_{m\times n}$,其中

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{is}b_{sj}=\sum_{k=1}^{s}a_{ik}b_{kj},\qquad(2.2)$$

称为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积,记为 $\boldsymbol{C}=\boldsymbol{AB}$.

由矩阵乘法定义可以看出, $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积 $\boldsymbol{C}$ 的第 i 行第 j 列的元素等于前一个矩阵的第 i 行与后一个矩阵第 j 列的对应元素乘积的和,矩阵 $\boldsymbol{C}$ 的行数就是前一个矩阵的行数,而 $\boldsymbol{C}$ 的列数就是后一个矩阵的列数,为了便于记忆矩阵乘法,可形象地看图 2.1 .

$$
\begin{array}{c} \\ \\ i \\ \\ \end{array}
\underset{m\times s}{\begin{array}{c} \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ \boxed{a_{i1}\quad a_{i2}\quad \cdots\quad a_{is}} \\ \cdots\cdots\cdots\cdots\cdots \end{array}}
\quad
\underset{s\times n}{\underset{j}{\begin{array}{ccc} \cdots & \boxed{\begin{array}{c} b_{1j} \\ b_{2j} \\ \cdots \\ b_{sj} \end{array}} & \cdots \end{array}}}
=
\underset{m\times n}{\underset{j}{\begin{array}{ccc} & \cdots & \\ \cdots & \boxed{c_{ij}} & \cdots \\ & \cdots & \end{array}}}
\; i
$$

图　2.1

例如,设由 x 经 y 到 z 的道路如图 2.2 .

图　2.2

用矩阵表示即有

$$
\begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 1 & 1 \end{pmatrix}
\begin{pmatrix} 3 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix}
=
\begin{pmatrix} 3 & 0 & 1 \\ 7 & 2 & 3 \\ 4 & 2 & 2 \end{pmatrix},
$$

其中 $c_{12}=0$ 表示由 x_1 到 z_2 无路, $c_{21}=2\times3+1\times1=7$ 表示由 x_2 到 z_1 有 7 条路,其中 6 条经 y_1 ,1 条经 y_2 .

例 4　设

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

求 $\boldsymbol{AB}$

解　$\boldsymbol{A}$ 是 2×3 矩阵，$\boldsymbol{B}$ 是 3×4 矩阵，$\boldsymbol{A}$ 的列数等于 $\boldsymbol{B}$ 的行数，故 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可以相乘且乘积 $\boldsymbol{C}$ 是 2×4 矩阵，即有

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

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

其中 $c_{11}=1\cdot1+(-1)\cdot1+1\cdot0=0,$

$c_{23}=2\cdot(-2)+0\cdot4+3\cdot3=5,\cdots.$ ■

例 5　如 $\boldsymbol{A}=(a_1\quad a_2\quad a_3)$，$\boldsymbol{B}=\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}$，求 $\boldsymbol{AB}$ 及 $\boldsymbol{BA}$.

解 $$\boldsymbol{AB}=(a_1\quad a_2\quad a_3)\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}=a_1b_1+a_2b_2+a_3b_3,$$

$$\boldsymbol{BA}=\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}(a_1\quad a_2\quad a_3)=\begin{pmatrix}a_1b_1&a_2b_1&a_3b_1\\a_1b_2&a_2b_2&a_3b_2\\a_1b_3&a_2b_3&a_3b_3\end{pmatrix}.$$ ■

例 6　如 $\boldsymbol{A}=\begin{pmatrix}1&1\\1&1\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}1&1\\-1&-1\end{pmatrix}$，求 $\boldsymbol{AB}$ 及 $\boldsymbol{BA}$.

解　$$\boldsymbol{AB}=\begin{pmatrix}1&1\\1&1\end{pmatrix}\begin{pmatrix}1&1\\-1&-1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}=\boldsymbol{O},$$

$$BA = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix} \neq O.$$ ■

从上面的两个例题我们看到,矩阵的乘法一般没有交换律,即 $AB \neq BA$.因此,对矩阵乘法一定要注意先后次序.

从例 6 还看到,虽 $A \neq O$, $B \neq O$,但确有 $AB = O$.因此,矩阵的乘法没有消去律.即若 $AB = AC$,不一定有 $B = C$.

矩阵 $I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}_{n\times n}$ 称为 n 阶单位矩阵,它在矩阵乘法中的作用,就像数字乘法中的 1.

例如

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

关于矩阵的乘法有下述性质(假设运算都可进行):

(1) $(AB)C = A(BC)$,

(2) $A(B + C) = AB + AC$,

$(A + B)C = AC + BC$,

(3) $(kA)(lB) = (kl)AB$,

(4) $AI = A$, $IA = A$,

(5) $AO = O$, $OA = O$.

如 A 是 n 阶方阵,利用矩阵乘法有结合律,可以定义 n 阶方阵的方幂.我们规定:

$$A^1 = A,\ A^2 = AA, \cdots, A^{k+1} = A^k A.$$

利用乘法可结合,因此

$$\boldsymbol{A}^k \cdot \boldsymbol{A}^l = \boldsymbol{A}^{k+l},\ (\boldsymbol{A}^k)^l = \boldsymbol{A}^{kl}.$$

但因矩阵乘法没有交换律,所以$(\boldsymbol{AB})^k \neq \boldsymbol{A}^k\boldsymbol{B}^k$.

***例 7** $\boldsymbol{A}, \boldsymbol{B}, I$ **都是** n **阶矩阵,如** $\boldsymbol{AB}+\boldsymbol{BA}=\boldsymbol{I}$**,证明:** $\boldsymbol{A}^3\boldsymbol{B}+\boldsymbol{BA}^3=\boldsymbol{A}^2$**.**

证 对 $\boldsymbol{AB}+\boldsymbol{BA}=\boldsymbol{I}$,用 $\boldsymbol{A}^2$ 分别左乘和右乘就有

$$\boldsymbol{A}^3\boldsymbol{B}+\boldsymbol{A}^2\boldsymbol{BA}=\boldsymbol{A}^2,$$

$$\boldsymbol{ABA}^2+\boldsymbol{BA}^3=\boldsymbol{A}^2,$$

两式相加

$$\boldsymbol{A}^3\boldsymbol{B}+\boldsymbol{BA}^3+\boldsymbol{A}^2\boldsymbol{BA}+\boldsymbol{ABA}^2=2\boldsymbol{A}^2,$$

又 $$\boldsymbol{A}^2\boldsymbol{BA}+\boldsymbol{ABA}^2=\boldsymbol{A}(\boldsymbol{AB}+\boldsymbol{BA})\boldsymbol{A}=\boldsymbol{AIA}=\boldsymbol{A}^2,$$

故 $$\boldsymbol{A}^3\boldsymbol{B}+\boldsymbol{BA}^3=\boldsymbol{A}^2.$$ ■

例 8 **已知** $\boldsymbol{A}=\begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{pmatrix}$,**证明** $\boldsymbol{A}^2=l\boldsymbol{A}$ **并求** l.

证

$$\boldsymbol{A}=\begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{pmatrix}=\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}(b_1\ \ b_2\ \ b_3),$$

$$\boldsymbol{A}^2=\left[\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}(b_1\ \ b_2\ \ b_3)\right]\left[\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}(b_1\ \ b_2\ \ b_3)\right]$$

$$=\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}\left[(b_1\ \ b_2\ \ b_3)\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}\right](b_1\ \ b_2\ \ b_3)$$

$$=(a_1b_1+a_2b_2+a_3b_3)\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}(b_1\ \ b_2\ \ b_3)$$

$$= (a_1 b_1 + a_2 b_2 + a_3 b_3) \boldsymbol{A},$$

可见 $\boldsymbol{A}^2 = l\boldsymbol{A}$，其中 $l = a_1 b_1 + a_2 b_2 + a_3 b_3$. ■

例 9 用矩阵表示方程组

$$\begin{cases} 2x_1 - 3x_2 + 2x_3 = 2, \\ x_1 + 4x_2 - 3x_3 = 0, \\ 3x_1 + x_2 - x_3 = 2. \end{cases}$$

解 根据矩阵乘法的定义，方程组的左端可表示成

$$\begin{pmatrix} 2 & -3 & 2 \\ 1 & 4 & -3 \\ 3 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

若令

$$\boldsymbol{A} = \begin{pmatrix} 2 & -3 & 2 \\ 1 & 4 & -3 \\ 3 & 1 & -1 \end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix},$$

则方程组可表示成：

$$\boldsymbol{AX} = \boldsymbol{B}.$$ ■

4) 转置

定义 2.4 一个 $m \times n$ 矩阵 $\boldsymbol{A}$，若将它的行改为列，列改为行，得到一个 $n \times m$ 矩阵，叫做 $\boldsymbol{A}$ 的转置矩阵，记为 $\boldsymbol{A}'$ 或 $\boldsymbol{A}^{\mathrm{T}}$.

例如 $\boldsymbol{A} = \begin{pmatrix} 3 & 2 & 4 \\ 1 & 7 & 5 \end{pmatrix}$ 则 $\boldsymbol{A}^{\mathrm{T}} = \begin{pmatrix} 3 & 1 \\ 2 & 7 \\ 4 & 5 \end{pmatrix}$.

矩阵的转置满足下列运算规律：

(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) $(k\boldsymbol{A})^{\mathrm{T}} = k\boldsymbol{A}^{\mathrm{T}}$,

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

例 10 如 $A=(2\quad 1\quad -1)$，$B=\begin{pmatrix}0&3\\4&5\\2&7\end{pmatrix}$，求 $(AB)^{\mathrm{T}}$.

解 1 由于

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

故

$$(AB)^{\mathrm{T}}=\begin{pmatrix}2\\4\end{pmatrix}.\qquad \blacksquare$$

解 2 由 $A^{\mathrm{T}}=\begin{pmatrix}2\\1\\-1\end{pmatrix}$，$B^{\mathrm{T}}=\begin{pmatrix}0&4&2\\3&5&7\end{pmatrix}$,

$$B^{\mathrm{T}}A^{\mathrm{T}}=\begin{pmatrix}0&4&2\\3&5&7\end{pmatrix}\begin{pmatrix}2\\1\\-1\end{pmatrix}=\begin{pmatrix}2\\4\end{pmatrix},$$

所以

$$(AB)^{\mathrm{T}}=\begin{pmatrix}2\\4\end{pmatrix}.\qquad \blacksquare$$

例 11 如 $A=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$，$B=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$，求 $(AB)^2$，A^2B^2.

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

得

$$(AB)^2=(AB)(AB)=\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}=I,$$

由

$$A^2=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}1&0\\0&-1\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}=I,$$

$$B^2=\begin{pmatrix}0&1\\-1&0\end{pmatrix}\begin{pmatrix}0&1\\-1&0\end{pmatrix}=\begin{pmatrix}-1&0\\0&-1\end{pmatrix}=-I$$

得 $$A^2B^2 = I\cdot(-I) = -I.$$ ■

在本节的最后,我们介绍行列式的乘法公式.

定理 2.1 设 A,B 都是 n 阶矩阵,则

$$|AB| = |A|\cdot|B| \tag{2.3}$$

即矩阵乘积的行列式等于它的因子的行列式的乘积.

证 (为了简明,只看 $n=2$ 的情形,对任意 n 的证明其本质是一样的.)

设 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$,我们构造四阶行列式 D

$$D = \begin{vmatrix} A & 0 \\ -I & B \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ -1 & 0 & b_{11} & b_{12} \\ 0 & -1 & b_{21} & b_{22} \end{vmatrix},$$

按拉普拉斯展开公式,一方面有

$$D = |A|\cdot|B|,$$

另一方面,把第 1 列的 b_{11} 倍加至第 3 列,第 2 列的 b_{21} 倍也加至第 3 列,就可把 D 中原来 b_{11}, b_{21} 位置上的数变成 0.类似地再把 b_{12}, b_{22} 位置上的数变成 0.即

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{11}b_{11} & 0 \\ a_{21} & a_{22} & a_{21}b_{11} & 0 \\ -1 & 0 & 0 & b_{12} \\ 0 & -1 & b_{21} & b_{22} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{11}b_{11}+a_{12}b_{21} & 0 \\ a_{21} & a_{22} & a_{21}b_{11}+a_{22}b_{21} & 0 \\ -1 & 0 & 0 & b_{12} \\ 0 & -1 & 0 & b_{22} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21} & a_{22} & a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{vmatrix}$$

$$= \begin{vmatrix} \boldsymbol{A} & \boldsymbol{AB} \\ -\boldsymbol{I} & \boldsymbol{O} \end{vmatrix}$$

$$= |-\boldsymbol{I}|\cdot(-1)^{2\times 2}|\boldsymbol{AB}| = |\boldsymbol{AB}|,$$

所以，$$|\boldsymbol{AB}| = |\boldsymbol{A}|\cdot|\boldsymbol{B}|.$$ ■

一般地，如 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m$ 都是 n 阶矩阵，则

$$|\boldsymbol{A}_1||\boldsymbol{A}_2|\cdots|\boldsymbol{A}_m| = |\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_m|. \tag{2.4}$$

例 12 设 $\boldsymbol{A} = \begin{pmatrix} 3 & 2 \\ 5 & 6 \end{pmatrix}$，$\boldsymbol{B} = \begin{pmatrix} 4 & -3 \\ -2 & 1 \end{pmatrix}$，验证行列式乘法公式.

解 一方面，由 $\boldsymbol{AB} = \begin{pmatrix} 3 & 2 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} 4 & -3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 8 & -7 \\ 8 & -9 \end{pmatrix}$

得
$$|\boldsymbol{AB}| = \begin{vmatrix} 8 & -7 \\ 8 & -9 \end{vmatrix} = -16.$$

另一方面，由 $|\boldsymbol{A}| = \begin{vmatrix} 3 & 2 \\ 5 & 6 \end{vmatrix} = 8$，$|\boldsymbol{B}| = \begin{vmatrix} 4 & -3 \\ -2 & 1 \end{vmatrix} = -2$，有 $|\boldsymbol{A}|\cdot|\boldsymbol{B}| = -16$.

可见
$$|\boldsymbol{AB}| = |\boldsymbol{A}|\cdot|\boldsymbol{B}|.$$ ■

2.2 可逆矩阵

对矩阵，已定义了加、减、乘运算，那么乘法是否会有逆运算呢？

由于 $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$，单位矩阵 $\boldsymbol{I}$ 在矩阵乘法中的作用类似于数字 1 在乘法中的作用，而一个数 $a \neq 0$ 时，其倒数 a^{-1} 可用 $a\cdot$

$a^{-1}=1$ 来刻划,这启发我们引出矩阵 $\boldsymbol{A}$ 的逆矩阵的概念.

定义 2.5 设 $\boldsymbol{A}$ 是 n 阶矩阵,如果存在 n 阶矩阵 $\boldsymbol{B}$ 使得 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}$ 成立,则称 $\boldsymbol{A}$ 是可逆矩阵 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆矩阵.

定理 2.2 如 $\boldsymbol{A}$ 是可逆矩阵,则 $\boldsymbol{A}$ 的逆矩阵是唯一的.

证 设 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 都是 $\boldsymbol{A}$ 的逆矩阵,即

$$\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I},\quad \boldsymbol{AC}=\boldsymbol{CA}=\boldsymbol{I}.$$

那么

$$\boldsymbol{B}=\boldsymbol{BI}=\boldsymbol{B}(\boldsymbol{AC})=(\boldsymbol{BA})\boldsymbol{C}=\boldsymbol{IC}=\boldsymbol{C},$$

所以 $\boldsymbol{A}$ 的逆矩阵是唯一的. ■

由于可逆矩阵的逆矩阵是唯一的,因此当可逆时,我们用 $\boldsymbol{A}^{-1}$ 表示 $\boldsymbol{A}$ 的逆矩阵.显然,单位矩阵是可逆的,且 $\boldsymbol{I}^{-1}=\boldsymbol{I}$.那么,一个矩阵何时可逆呢?如可逆其逆矩阵又怎样求呢?

定理 2.3 n 阶矩阵 $\boldsymbol{A}$ 可逆的充分必要条件是 $|\boldsymbol{A}|\neq 0$.

证)如 $\boldsymbol{A}$ 可逆,则 $\boldsymbol{AA}^{-1}=\boldsymbol{I}$,由行列式乘法公式(定理 2.1)有:

$$|\boldsymbol{A}|\cdot|\boldsymbol{A}^{-1}|=|\boldsymbol{AA}^{-1}|=|\boldsymbol{I}|=1,$$

所以 $|\boldsymbol{A}|\neq 0$;

) 如 $|\boldsymbol{A}|\neq 0$,我们构造矩阵 $\boldsymbol{A}^*$,它由 $|\boldsymbol{A}|$ 的代数余子式组成,为书写简单下面对三阶矩阵 $\boldsymbol{A}$ 写出它的 $\boldsymbol{A}^*$.即构造

$$\boldsymbol{A}^*=\begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}, \tag{2.5}$$

利用行列式性质 2 与性质 10,得到

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}=\begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$= \begin{pmatrix} |A| & 0 & 0 \\ 0 & |A| & 0 \\ 0 & 0 & |A| \end{pmatrix} = |A| I,$$

由于 $|A| \neq 0$,故有

$$A \cdot \frac{1}{|A|} A^* = \frac{1}{|A|} A^* \cdot A = I, \qquad (2.6)$$

根据可逆的定义知矩阵 A 可逆,且 $A^{-1} = \frac{1}{|A|} A^*$. ■

定理 2.3 不仅给出了矩阵可逆的充要条件,而且还构造性地给出了一个求逆矩阵的方法 .我们称 A^* 是矩阵 A 的伴随矩阵 .

由定理 2.3 我们还看到,如 n 阶矩阵 A, B 满足 $AB = I$,则 $|A| \cdot |B| = |I| = 1$,即 $|A| \neq 0$,于是 A 可逆 .又因逆矩阵是唯一的,得

$$BA = (A^{-1}A)BA = A^{-1}(AB)A = A^{-1}IA = I$$

因此, A 可逆,且 $A^{-1} = B$.所以,今后若用定义来证明 B 是 A 的逆矩阵时,我们只须检查 $AB = I$ 或 $BA = I$ 之一成立就可以了 .

例 1 求 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ 的伴随矩阵及逆矩阵 .

解 由于 $|A|$ 的代数余子式 $A_{11} = 4$, $A_{12} = -3$, $A_{21} = -2$, $A_{22} = 1$.有 A 的伴随矩阵

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

又因 $|A| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2 \neq 0$, 知 A 可逆,且

$$A^{-1} = \frac{1}{|A|} A^* = -\frac{1}{2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}. \quad ■$$

例 2 求 $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ 的逆矩阵 .

解 显然$|\boldsymbol{A}|=1\neq 0$，知 $\boldsymbol{A}$ 可逆，由于

$$A_{11}=\begin{vmatrix}1&2\\0&1\end{vmatrix}=1,\quad A_{12}=-\begin{vmatrix}0&2\\0&1\end{vmatrix}=0,$$

$$A_{13}=\begin{vmatrix}0&1\\0&0\end{vmatrix}=0,\quad A_{21}=-\begin{vmatrix}2&3\\0&1\end{vmatrix}=-2,$$

$$A_{22}=\begin{vmatrix}1&3\\0&1\end{vmatrix}=1,\quad A_{23}=-\begin{vmatrix}1&2\\0&0\end{vmatrix}=0,$$

$$A_{31}=\begin{vmatrix}2&3\\1&2\end{vmatrix}=1,\quad A_{32}=-\begin{vmatrix}1&3\\0&2\end{vmatrix}=-2,$$

$$A_{33}=\begin{vmatrix}1&2\\0&1\end{vmatrix}=1,$$

得到 $\boldsymbol{A}$ 的伴随矩阵 $\boldsymbol{A}^*$

$$\boldsymbol{A}^*=\begin{pmatrix}A_{11}&A_{21}&A_{31}\\A_{12}&A_{22}&A_{32}\\A_{13}&A_{23}&A_{33}\end{pmatrix}=\begin{pmatrix}1&-2&1\\0&1&-2\\0&0&1\end{pmatrix},$$

所以，$\boldsymbol{A}$ 的逆矩阵

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*=\begin{pmatrix}1&-2&1\\0&1&-2\\0&0&1\end{pmatrix}.$$ ■

例 3 $\boldsymbol{A}$ 是 n 阶矩阵满足关系：$\boldsymbol{A}^2-3\boldsymbol{A}+2\boldsymbol{I}=\boldsymbol{O}$，证明 $\boldsymbol{A}$ 可逆，并求其逆.

证 因为 $\boldsymbol{A}^2-3\boldsymbol{A}+2\boldsymbol{I}=\boldsymbol{O}$，移项后有

$$\boldsymbol{A}(\boldsymbol{A}-3\boldsymbol{I})=-2\boldsymbol{I},$$

得到

$$\boldsymbol{A}\left[-\frac{1}{2}(\boldsymbol{A}-3\boldsymbol{I})\right]=\boldsymbol{I},$$

因此 $\boldsymbol{A}$ 可逆，且 $\boldsymbol{A}^{-1}=-\frac{1}{2}(\boldsymbol{A}-3\boldsymbol{I})$. ■

例 4 设矩阵 A 可逆，证明 A^* 也可逆，并求 $(A^*)^{-1}$.

证 因为 A 可逆，知 $|A| \neq 0$. 又因

$$AA^* = A^* A = |A| I,$$

得到

$$A^* \cdot \frac{1}{|A|} A = I.$$

因此 A^* 可逆，且 $(A^*)^{-1} = \frac{1}{|A|} A$ ■

例 5 已知 $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$，如 $AX + I = A^2 + X$，求矩阵 X.

解 由 $AX + I = A^2 + X$，移项整理成

$$(A - I)X = A^2 - I = (A - I)(A + I),$$

因为 $A - I = \begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix}$ 可逆，用 $(A - I)^{-1}$ 左乘上式的两端得:

$$X = A + I = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$ ■

例 6 设

$$A = \begin{pmatrix} 2 & 1 & -1 \\ -1 & 4 & -1 \\ 1 & -1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -4 \\ 2 & 2 \\ 4 & 6 \end{pmatrix}$$

如 $AX = B + X$，求矩阵 X.

解 由 $AX = B + X$ 得 $(A - I)X = B$，因

$$A - I = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 3 & -1 \\ 1 & -1 & 1 \end{pmatrix},$$

可求出

$$|A - I| = 4 \neq 0,$$

$$(A - I)^{-1} = \frac{1}{4}\begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 2 \\ -2 & 2 & 4 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix},$$

所以

$$\begin{aligned} X &= (A - I)^{-1} B \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}\begin{pmatrix} 0 & -4 \\ 2 & 2 \\ 4 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ 5 & 9 \end{pmatrix}. \end{aligned}$$ ■

例 7 如 A, B 都是 n 阶可逆矩阵,证明 AB 可逆,且 $(AB)^{-1} = B^{-1}A^{-1}$.

证 ∵ A, B 都可逆,知 $|A| \neq 0$, $|B| \neq 0$.又

$$|AB| = |A| \cdot |B| \neq 0$$

∴ AB 可逆.

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

从而有
$$(AB)^{-1} = B^{-1}A^{-1}.$$ ■

例 8 如 A 可逆,证明$(A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1}$.

证 由 $A^{\mathrm{T}}(A^{-1})^{\mathrm{T}} = [(A^{-1})A]^{\mathrm{T}} = I^{\mathrm{T}} = I$

故
$$(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}.$$ ■

关于可逆矩阵的基本性质可归纳如下:

(1) 如 A 可逆,则 A^{-1} 也可逆且$(A^{-1})^{-1} = A$;

(2) 如 A 可逆, $k \neq 0$,则 kA 可逆,且$(kA)^{-1} = \frac{1}{k}A^{-1}$;

(3) 如 A_1, A_2 都是 n 阶可逆,则 A_1A_2 可逆且$(A_1A_2)^{-1} = A_2^{-1}A_1^{-1}$.更一般地,若 A_i $(i = 1, 2, \cdots, s)$都是 n 阶可逆,就有

$$(A_1A_2\cdots A_s)^{-1} = A_s^{-1}\cdots A_2^{-1}A_1^{-1};$$

(4) 如 A 可逆,则 $(A^{-1})^T = (A^T)^{-1}$.

要小心的是,可逆矩阵的和不一定是可逆的; 若 $A + B$ 可逆,也不一定有 A, B 可逆;即使 $A, B, A + B$ 都可逆,一般地 $(A + B)^{-1} \neq A^{-1} + B^{-1}$.例如:

$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,虽 A, B 都可逆,

但 $A + B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ 不可逆;

$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$,虽 A, B 都不可逆,

但 $A + B = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$ 可逆;

$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,有 $A + B = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$,虽 $A, B, A + B$ 都可逆,但

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

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

*例 9　已知 A 可逆, $(A + B)^2 = I$, 化简 $(I + BA^{-1})^{-1}$.

解

$$\begin{aligned}(I + BA^{-1})^{-1} &= (AA^{-1} + BA^{-1})^{-1} \\ &= [(A + B)A^{-1}]^{-1} \\ &= (A^{-1})^{-1}(A + B)^{-1} \\ &= A(A + B) .\end{aligned}$$

■

*例 10　已知 $A, B, A - B$ 都可逆,求 $(A^{-1} - B^{-1})^{-1}$.

解

$$(A^{-1} - B^{-1})^{-1} = (A^{-1}I - IB^{-1})^{-1}$$

$$= (A^{-1}BB^{-1} - A^{-1}AB^{-1})^{-1}$$
$$= [A^{-1}(B - A)B^{-1}]^{-1}$$
$$= (B^{-1})^{-1}(B - A)^{-1}(A^{-1})^{-1}$$
$$= - B(A - B)^{-1}A. \qquad ■$$

*例 11　A 是 3 阶矩阵, $AA^{T} = I$, $|A| = 1$, 证明 $I - A$ 不可逆.

证

$$|I - A| = |AA^{T} - A| = |A(A^{T} - I)|$$
$$= |A|\cdot|A^{T} - I| = |(A - I)^{T}|$$
$$= |A - I| = (-1)^{3}|I - A| = -|I - A|,$$

所以 $|I - A| = 0$ 即有 $I - A$ 不可逆. ■

最后,我们介绍用逆矩阵证明克莱姆法则的方法.

*例 12　对三元一次方程组

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1,$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2,$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3.$$

我们用矩阵表示成 $Ax = B$,其中

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},\quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},\quad B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

如系数行列式 $|A| \neq 0$,即矩阵 A 可逆,那么对方程组用 A^{-1} 左乘两端有:

$$x = A^{-1}B = \frac{1}{|A|}A^{*}B$$

$$= \frac{1}{|\boldsymbol{A}|}\begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$= \frac{1}{|\boldsymbol{A}|}\begin{pmatrix} b_1 A_{11} + b_2 A_{21} + b_3 A_{31} \\ b_1 A_{12} + b_2 A_{22} + b_3 A_{32} \\ b_1 A_{13} + b_2 A_{23} + b_3 A_{33} \end{pmatrix}$$

$$= \frac{1}{\boldsymbol{\Delta}}\begin{pmatrix} \boldsymbol{\Delta}_1 \\ \boldsymbol{\Delta}_2 \\ \boldsymbol{\Delta}_3 \end{pmatrix}.$$

于是得到方程组的解

$$x_1 = \frac{\boldsymbol{\Delta}_1}{\boldsymbol{\Delta}},\quad x_2 = \frac{\boldsymbol{\Delta}_2}{\boldsymbol{\Delta}},\quad x_3 = \frac{\boldsymbol{\Delta}_3}{\boldsymbol{\Delta}}.$$ ■

虽然我们已能通过伴随矩阵来求一个可逆矩阵的逆矩阵,但只要 n 稍大,这计算工作量是很大的.因此还要寻找简捷一些的方法,这就是初等变换,初等矩阵.

2.3 初等矩阵

本节先给出初等变换,初等矩阵的概念,建立起初等变换与矩阵乘法的联系,然后推导出求逆矩阵的方法.

对 $m \times n$ 矩阵,我们把

(1) 用非零常数 k 乘矩阵的某一行(列);

(2) 互换矩阵某两行(列)的位置;

(3) 把某行(列)的 k 倍加至另一行(列).

称为是矩阵的初等行(列)变换,统称为矩阵的初等变换.用记号

$A \to B$ 表示矩阵 A 经过初等变换得到矩阵 B .

定义 2.6 矩阵 A 经过一系列初等变换得到矩阵 B, 称矩阵 A 与矩阵 B 等价, 记作 AD B .

习惯上, 把形如

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2r} & \cdots & a_{2n} \\ & & \ddots & & & \\ & & & a_{rr} & \cdots & a_{rn} \end{pmatrix}$$

或

$$\begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1j} & \cdots\cdots & a_{1n} \\ & & a_{2i} & \cdots & a_{2j} & \cdots\cdots & a_{2n} \\ & & & & a_{3j} & \cdots\cdots & a_{3n} \\ & & & & & \cdots\cdots & \cdots \\ & & & & & a_{rs}\cdots & a_{rn} \end{pmatrix}$$

的矩阵叫做“阶梯形”矩阵 .阶梯形矩阵在解线性方程组中有重要的作用 .而任一矩阵都可经初等行变换化为阶梯形矩阵, 即任一矩阵都与阶梯形矩阵等价 .例如

$$A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 3 & 1 & 4 \\ 1 & 3 & 2 & 7 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 3 & -2 \\ 0 & 1 & 3 & 4 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 2 & -1 & 3 \\ & -1 & 3 & -2 \\ & & 6 & 2 \end{pmatrix} = B,$$

我们有 AD B, 而 B 就是一个阶梯形矩阵 .又如

$$A = \begin{pmatrix} 0 & 0 & 2 & 5 \\ 2 & 4 & 3 & 8 \\ 1 & 2 & 2 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 2 & 4 \\ 2 & 4 & 3 & 8 \\ 0 & 0 & 2 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 2 & 4 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 2 & 5 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 2 & 2 & 4 \\ & & -1 & 0 \\ & & & 5 \end{pmatrix} = \boldsymbol{B}.$$

这里 $\boldsymbol{B}$ 也是阶梯形矩阵,它与 A 等价 .

如果对矩阵既作初等行变换也作初等列变换,则与 $\boldsymbol{A}$ 等价的矩阵有更简单的形式 .

例如,对矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix}$ 先作行变换,把第 1 行的 - 2 倍加至第 2 行并把第 1 行加至第 3 行,经如此初等行变换,得到

$$\boldsymbol{A} \rightarrow \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 6 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \boldsymbol{B}_1,$$

再把第 2 行的 - 2 倍加至第 3 行,得阶梯形矩阵 $\boldsymbol{B}_1$,对 $\boldsymbol{B}_1$ 作初等列变换,把第 1 列的 - 3 倍分别加至 $\boldsymbol{B}_1$ 的第 2 和第 3 列,并把第 1 列的 - 2 倍加至第 4 列,就有

$$\boldsymbol{A} \rightarrow \cdots \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \boldsymbol{B},$$

再把 2、4 两列互换,最后把第 2 列的 - 3 倍加至第 3 列,就得到矩阵 $\boldsymbol{B}$,这样的矩阵 $\boldsymbol{B}$ 通常称为 $\boldsymbol{A}$ 的等价标准形 .更加一般地

对任一非零的 $m \times n$ 矩阵 $\boldsymbol{A}$,如 $a_{11} \neq 0$, 我们可把第 1 行的 $-\dfrac{a_{i1}}{a_{11}}$倍加至第 i 行$(i = 2, 3, \cdots, m)$,再把第 1 列的 $-\dfrac{a_{1j}}{a_{11}}$倍加至第 j 列$(j = 2, 3, \cdots, n)$,然后用$\dfrac{1}{a_{11}}$乘第一行, $\boldsymbol{A}$ 就变成

$$\begin{pmatrix} 1 & 0\cdots 0 \\ 0 & \\ \cdots & \boldsymbol{A}_1 \\ 0 & \end{pmatrix},$$

其中 $\boldsymbol{A}_1$ 是$(m-1)\times(n-1)$矩阵.如 $a_{11}=0$,由于 $\boldsymbol{A}$ 是非 0 的,不妨设 $a_{ij}\neq 0$,那么对 $\boldsymbol{A}$ 先把第 1 行与第 i 行对换,然后对换第 1 列与第 j 列,就可把 a_{ij}调用 $\boldsymbol{A}$ 中 a_{11}的位置上,再按前述方法变换 $\boldsymbol{A}$ 就等价于 $a_{11}=1$, $a_{1j}=0\ (j=2,3,\cdots,n)$, $a_{i1}=0\ (i=2,3,\cdots,m)$的矩阵.

对 $\boldsymbol{A}_1$ 再重复以上的步骤,就有

定理 2.4 任一个非 0 的 $m\times n$ 矩阵 $\boldsymbol{A}$ 都和一个形式为

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & & \cdots & & \cdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \hline \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \tag{2.7}$$

的矩阵等价,它称为矩阵 $\boldsymbol{A}$ 的等价标准形.

容易看出 $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ 的等价标准形是 $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}$ 的等价标准形是 $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

定义 2.7 单位矩阵经过一次初等变换所得到的矩阵叫做初等矩阵.

虽然有三种初等行变换,三种初等列变换,但反映到初等矩阵上却只有三种形式,也就是说每个初等矩阵既可看成经初等行变换也可看作由列变换所得到.

例如，二阶单位矩阵经过一次初等变换可有下列三种类型

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

它们分别表示把 I_2 的第 2 行(列)乘以 3；第 1、2 两行(列)互换；第 1 行的 -2 倍加至第 2 行或从列的角度来看是第 2 列的 -2 倍加至第 1 列．按定义，它们都是初等矩阵．

从行列式的角度来看，经初等变换行列式值的大小、正负有可能改变，但行列式的值是否等于 0 不会因初等变换而改变．所以，初等矩阵的行列式不会等于 0，因此初等矩阵都是可逆矩阵，并且初等矩阵的逆矩阵就是与它同类型的初等矩阵．例如

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

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

下面来看初等矩阵在矩阵乘法中所起的作用．容易算出

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 9 & 12 \end{pmatrix}; \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 6 \\ 3 & 12 \end{pmatrix};$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix};$$

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

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ -5 & 4 \end{pmatrix}.$$

其规律是：当用初等矩阵 P 左乘矩阵 A 时，PA 就是由 A 作了一个与 P 同样的初等行变换；当用初等矩阵 Q 右乘矩阵 A 时，AQ

就等于由 $\boldsymbol{A}$ 做了一个与 $\boldsymbol{Q}$ 同样的列变换 .

如 $\boldsymbol{A}$ 是 $m\times n$ 矩阵, $\boldsymbol{A}$ 经一次初等行(列)变换得到 $\boldsymbol{B}$, 也就是有 $m(n)$阶初等矩阵 $\boldsymbol{P}(\boldsymbol{Q})$使 $\boldsymbol{PA}=\boldsymbol{B}(\boldsymbol{AQ}=\boldsymbol{B})$成立 .那么, 两个 $m\times n$ 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价, 也就是 $\boldsymbol{A}$ 经过一系列初等行与列的变换得到 $\boldsymbol{B}$, 从而存在一串 m 阶初等矩阵 $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_s$ 及 n 阶初等矩阵 $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_t$ 使得

$$\boldsymbol{P}_s\cdots\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_t=\boldsymbol{B}$$

成立 .

因为初等矩阵是可逆矩阵,而可逆矩阵的乘积还是可逆矩阵 .所以 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价就相当于存在可逆矩阵 $\boldsymbol{P}$、$\boldsymbol{Q}$ 使 $\boldsymbol{PAQ}=\boldsymbol{B}$.

特别地, 对任一非 0 的 $m\times n$ 矩阵 $\boldsymbol{A}$, 存在可逆矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$, 使

$$\boldsymbol{PAQ}=\begin{matrix}1 & & & \\ & 1 & & \\ & & \mathrm{W} & \\ & & & 1\end{matrix} \qquad (2.8)$$

其右端的矩阵就是 $\boldsymbol{A}$ 的等价标准形 .

定理 2.5 n 阶矩阵 $\boldsymbol{A}$ 可逆的充分必要条件是 $\boldsymbol{A}$ 可表示成一串初等矩阵的乘积 .

*证 设 $\boldsymbol{A}$ 的等价标准形是 $\boldsymbol{B}$, 即有初等矩阵 $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots \boldsymbol{P}_s$ 及 $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_t$ 使

$$\boldsymbol{P}_s\cdots\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_t=\boldsymbol{B},$$

两边取行列式,得

$$|\boldsymbol{P}_s|\cdots|\boldsymbol{P}_2||\boldsymbol{P}_1||\boldsymbol{A}||\boldsymbol{Q}_1||\boldsymbol{Q}_2|\cdots|\boldsymbol{Q}_t|=|\boldsymbol{B}|.$$

由于 $\boldsymbol{P}_i$, $\boldsymbol{Q}_j$ 都是初等矩阵其行列式不为 0, 故

$$|\boldsymbol{A}|\neq 0 \quad |\boldsymbol{B}|\neq 0(\boldsymbol{B}\text{ 是等价标准形}) \quad \boldsymbol{B}=\boldsymbol{I}.$$

因此, $\boldsymbol{A}$ 可逆 $\quad \boldsymbol{A}=\boldsymbol{P}_1^{-1}\cdot\boldsymbol{P}_2^{-1}\cdot\cdots\boldsymbol{P}_s^{-1}\boldsymbol{Q}_t^{-1}\cdots\boldsymbol{Q}_2^{-1}\boldsymbol{Q}_1^{-1}$

$\boldsymbol{A}$ 是一些初等矩阵的乘积 . ■

根据定理 2.5,可得到下述用初等变换求逆矩阵的方法:

如 A 可逆,则 $A = P_1 P_2 \cdots P_t$ (P_i 是初等矩阵)

有 $$P_t^{-1} \cdots P_2^{-1} P_1^{-1} A = I, \tag{i}$$

那么 $$A^{-1} = P_t^{-1} \cdots P_2^{-1} P_1^{-1} = P_t^{-1} \cdots P_2^{-1} P_1^{-1} I. \tag{ii}$$

(i) 式说明可逆矩阵 A 经过初等行变换 $P_1^{-1}, P_2^{-1}, \cdots, P_t^{-1}$ 可一步步变成单位矩阵 I,而(ii)式就表示在同样的初等行变换下单位矩阵 I 变成了 A^{-1}.

这样,我们可把 A 和 I 凑在一起,作成一个 $n\times 2n$ 的矩阵

$$(A \mid I),$$

对这个矩阵作初等行变换,当把其左边的一半 A 变成单位矩阵 I 时,这时右边的一半 I 就已变成了 A^{-1}.

例 1 求 $A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ 的逆矩阵.

解

$$(A \mid I) = \left(\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 2 & 1 & -1 \\ 1 & 3 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{cc|cc} 1 & 2 & 1 & -1 \\ 0 & 1 & -1 & 2 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 0 & 3 & -5 \\ 0 & 1 & -1 & 2 \end{array}\right)$$

所以 $$A^{-1} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}.$$ ■

例 2 求 $A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ 的逆矩阵.

解

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

故

$$
\rightarrow\left(\begin{array}{ccc|ccc}1&0&2&1&0&0\\1&1&0&0&1&0\\0&0&1&0&-1&1\end{array}\right)
\rightarrow\left(\begin{array}{ccc|ccc}1&0&2&1&0&0\\0&1&-2&-1&1&0\\0&0&1&0&-1&1\end{array}\right)
$$

$$
\rightarrow\left(\begin{array}{ccc|ccc}1&0&2&1&0&0\\0&1&0&-1&-1&2\\0&0&1&0&-1&1\end{array}\right)
\rightarrow\left(\begin{array}{ccc|ccc}1&&&1&2&-2\\&1&&-1&-1&2\\&&1&0&-1&1\end{array}\right)
$$

$\therefore$
$$
A^{-1}=\begin{pmatrix}1&2&-2\\-1&-1&2\\0&-1&1\end{pmatrix}.
$$
■

例 3　**已知** $A=\begin{pmatrix}2&5\\3&8\end{pmatrix}$，$B=\begin{pmatrix}4&0&-1\\5&1&2\end{pmatrix}$，$AX=B$ **求矩阵** X.

解 1　**由** $|A|=\begin{vmatrix}2&5\\3&8\end{vmatrix}\neq 0$，**知** A **可逆，故**

$$
X=A^{-1}B.
$$

求出

$$
A^{-1}=\begin{pmatrix}8&-5\\-3&2\end{pmatrix},
$$

故
$$
X=\begin{pmatrix}8&-5\\-3&2\end{pmatrix}\begin{pmatrix}4&0&-1\\5&1&2\end{pmatrix}
=\begin{pmatrix}7&-5&-18\\-2&2&7\end{pmatrix}.
$$
■

由于 $(A \mid I) \to \cdots \to (I \mid A^{-1})$,

即 $P_t \cdots P_2 P_1 (A \mid I) = (P_t \cdots P_2 P_1 A \mid P_t \cdots P_2 P_1) = (I \mid A^{-1})$,

那么 $P_t \cdots P_2 P_1 (A \mid B) = (P_t \cdots P_2 P_1 A \mid P_t \cdots P_2 P_1 B) = (I \mid A^{-1} B)$,

因此,我们亦可不求出 A^{-1} 再作矩阵乘法来求 X,而是用初等变换直接求出 X .

*解 2

$$(A \mid B) = \left(\begin{array}{cc|ccc} 2 & 5 & 4 & 0 & -1 \\ 3 & 8 & 5 & 1 & 2 \end{array}\right) \to \left(\begin{array}{cc|ccc} -1 & -3 & -1 & -1 & -3 \\ 3 & 8 & 5 & 1 & 2 \end{array}\right)$$

$$\to \left(\begin{array}{cc|ccc} -1 & -3 & -1 & -1 & -3 \\ 0 & -1 & 2 & -2 & -7 \end{array}\right) \to \left(\begin{array}{cc|ccc} 1 & 3 & 1 & 1 & 3 \\ 0 & 1 & -2 & 2 & 7 \end{array}\right)$$

$$\to \left(\begin{array}{cc|ccc} 1 & 0 & 7 & -5 & -18 \\ 0 & 1 & -2 & 2 & 7 \end{array}\right).$$

$$\therefore \quad X = \begin{pmatrix} 7 & -5 & -18 \\ -2 & 2 & 7 \end{pmatrix}.$$ ■

2.4 特殊矩阵

在本节学习一些特殊类型的矩阵,后几章中它们都有较多的应用.

如果矩阵 A 满足条件: $A^{T} = A$,则称 A 是对称矩阵.显然,若 A 是对称矩阵,则 A 必然是 n 阶方阵,且 A 的元素要满足:

$$a_{ij} = a_{ji} \qquad i, j = 1, 2, \cdots, n .$$

例如,单位矩阵,零矩阵都是对称矩阵,而

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} -3 & 4 \\ 4 & 5 \end{pmatrix}, \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

等也都是对称矩阵 .

根据转置运算的性质,我们知道如果 $\boldsymbol{A}$, $\boldsymbol{B}$ 都是对称矩阵,那么

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

$$(k\boldsymbol{A})^{\mathrm{T}}=k\boldsymbol{A}^{\mathrm{T}}=k\boldsymbol{A},$$

即对称矩阵的和,以及对称矩阵与数相乘仍是对称矩阵 .

如果 $\boldsymbol{A}$ 是可逆的对称矩阵,那么

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

即可逆的对称矩阵其逆仍是对称矩阵 .

但是两个对称矩阵的乘积一般不一定还是对称矩阵 .例如

$$\begin{pmatrix}1&2\\2&4\end{pmatrix}\begin{pmatrix}1&1\\1&-1\end{pmatrix}=\begin{pmatrix}3&-1\\6&-2\end{pmatrix}.$$

例 1　设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵,则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 n 阶对称矩阵, $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 是 m 阶对称矩阵 .

证　$\boldsymbol{A}^{\mathrm{T}}$ 是 $n\times m$ 矩阵, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 n 阶矩阵,且

$$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},$$

故 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 n 阶对称矩阵,类似有 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 是 m 阶对称矩阵 . ■

例 2　$\boldsymbol{A}$ 是 n 阶矩阵,则 $\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}$ 是对称矩阵 .

证　由

$$(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{A}=\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}},$$

所以　$\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}$ 是对称矩阵 . ■

例 3　用对称矩阵表示二元二次函数

$$f(x_1,x_2)=2x_1^2-3x_2^2+8x_1x_2 .$$

解

$$\begin{aligned}f(x_1,x_2)&=2x_1^2+4x_1x_2+4x_1x_2-3x_2^2\\&=x_1(2x_1+4x_2)+x_2(4x_1-3x_2)\\&=(x_1\quad x_2)\begin{pmatrix}2x_1+4x_2\\4x_1-3x_2\end{pmatrix}\end{aligned}$$

$$= (x_1 \quad x_2) \begin{pmatrix} 2 & 4 \\ 4 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

若记 $\boldsymbol{A} = \begin{pmatrix} 2 & 4 \\ 4 & -3 \end{pmatrix}$, $\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$,

则 $f(x_1, x_2) = \boldsymbol{X}^{\mathrm{T}} \boldsymbol{AX}$, 其中 $\boldsymbol{A}$ 是对称矩阵 . ■

形如

$$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$$

的 n 阶矩阵称为对角矩阵 .

对角矩阵的和,数乘,乘积仍是对角矩阵,特别地,对角矩阵的方幂就是其对角线上元素方幂所构成的对角矩阵 .例如

$$\begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix}^n = \begin{pmatrix} a_1^n & & \\ & a_2^n & \\ & & a_3^n \end{pmatrix}, \tag{2.9}$$

而对角矩阵 $\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$ 可逆的充分必要条件是: $a_1 a_2 \cdots a_n \neq 0$, 且

$$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{a_1} & & & \\ & \frac{1}{a_2} & & \\ & & \ddots & \\ & & & \frac{1}{a_n} \end{pmatrix}. \quad (2.10)$$

例 4　已知 $A = PBQ$,其中

$$P = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & \\ & -2 \end{pmatrix}, \quad Q = \begin{pmatrix} -1 & 2 \\ 2 & -3 \end{pmatrix},$$

求 A 及 A^n.

解

$$\begin{aligned} A &= PBQ \\ &= \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & \\ & -2 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 2 & -3 \end{pmatrix} \\ &= \begin{pmatrix} -11 & 18 \\ -6 & 10 \end{pmatrix}, \end{aligned}$$

利用 A 来求 A^n 是困难的,如注意到 $PQ = I$,即 $Q = P^{-1}$,那么

$$\begin{aligned} A^n &= (PBQ)^n = (PBP^{-1})^n \\ &= (PBP^{-1})(PBP^{-1})\cdots(PBP^{-1}) \\ &= PB^nP^{-1} \\ &= \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & \\ & -2 \end{pmatrix}^n \begin{pmatrix} -1 & 2 \\ 2 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & \\ & (-2)^n \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 2 & -3 \end{pmatrix} \\ &= \begin{pmatrix} -3 + (-1)^n 2^{n+2} & 6 + (-1)^{n+1} 3 \cdot 2^{n+1} \\ -2 + (-1)^n 2^{n+1} & 4 + (-1)^{n+1} 3 \cdot 2^n \end{pmatrix}. \end{aligned}$$
■

满足条件

$$AA^{\mathrm{T}} = A^{\mathrm{T}}A = I$$

的 n 阶矩阵 $\boldsymbol{A}$ 称为正交矩阵.那么

$$\boldsymbol{A} \text{ 是正交矩阵} \qquad \boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}}.$$

例如,单位矩阵 $\boldsymbol{I}$ 是正交矩阵.

又如

$$\boldsymbol{A} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix},$$

有

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \cdot \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

$$= \frac{1}{5}\begin{pmatrix} 5 & \\ & 5 \end{pmatrix} = \boldsymbol{I}.$$

类似地 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$,所以

$$\begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{pmatrix}$$

是正交矩阵.

例 5 已知 $\boldsymbol{A} = \begin{pmatrix} x & \frac{1}{2} \\ \frac{1}{2} & y \end{pmatrix}$ 是正交矩阵,求 x 和 y 的值.

解 由于 $\boldsymbol{A}$ 是正交矩阵,$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$,即

$$\begin{pmatrix} x & \frac{1}{2} \\ \frac{1}{2} & y \end{pmatrix}\begin{pmatrix} x & \frac{1}{2} \\ \frac{1}{2} & y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

得到

$$x^2 + \frac{1}{4} = 1,$$

$$\frac{1}{2}x + \frac{1}{2}y = 0,$$

$$\frac{1}{4} + y^2 = 1,$$

所以

$$x = \frac{3}{2}, \qquad x = -\frac{3}{2},$$

或 ■

$$y = -\frac{3}{2}. \qquad y = \frac{3}{2}.$$

例 6　如 $\boldsymbol{A}$, $\boldsymbol{B}$ 都是 n 阶正交矩阵,证明 $\boldsymbol{AB}$ 是正交矩阵 .

证

$$(\boldsymbol{AB})^{\mathrm{T}}(\boldsymbol{AB}) = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AB} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{IB} = \boldsymbol{I},$$

$$(\boldsymbol{AB})\cdot(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{AB}\cdot\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{AIA}^{\mathrm{T}} = \boldsymbol{I}.$$

所以 $\boldsymbol{AB}$ 是正交矩阵,即有:正交矩阵的乘积是正交矩阵 . ■

例 7　如 $\boldsymbol{A}$ 是正交矩阵,则 $|\boldsymbol{A}| = \pm 1$.

证　由 $\boldsymbol{AA}^{\mathrm{T}} = \boldsymbol{I}$,两边取行列式(2.3)

$$|\boldsymbol{A}|^2 = |\boldsymbol{A}|\cdot|\boldsymbol{A}^{\mathrm{T}}| = |\boldsymbol{AA}^{\mathrm{T}}| = |\boldsymbol{I}| = 1,$$

所以,

$$|\boldsymbol{A}| = \pm 1.$$ ■

注意: 正交矩阵的行列式为 1 或 - 1,但行列式为 ±1 的矩阵不一定是正交矩阵 .例如

$\boldsymbol{A} =$ 2 $\frac{1}{2}$ **有** $|\boldsymbol{A}| = 1$**,但** $AA^{\mathrm{T}} =$ 4 $\frac{1}{4}$ $\neq \boldsymbol{I}$ **.因而 $\boldsymbol{A}$ 不是正交矩阵 .**

2.5 分 块 矩 阵

对某些矩阵问题,有时我们可把它适当分成一些小矩阵,对原矩阵的讨论就转化为对这些小矩阵的讨论,这样做或是为了计算简捷或是为了论证清楚明了.这种方法就是分块矩阵,它是矩阵中的一个重要技巧.其实,分块矩阵的方法在前面已出现过,例如2.3中用初等变换求逆矩阵的方法就采用了分块矩阵的写法.

下面通过一些例题来介绍分块矩阵的方法:

例 1 已知 $\boldsymbol{X}=\boldsymbol{ABA}$,其中

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

求矩阵 $\boldsymbol{X}$.

解 我们把 $\boldsymbol{A},\boldsymbol{B}$ 各分成四个 2×2 的小矩阵,即

$$\boldsymbol{A}=\left(\begin{array}{cc|cc}1&&&1\\&1&1&\\\hline&1&-1&\\1&&&-1\end{array}\right)=\begin{pmatrix}\boldsymbol{I}&\boldsymbol{J}\\\boldsymbol{J}&-\boldsymbol{I}\end{pmatrix},$$

$$\boldsymbol{B}=\left(\begin{array}{cc|cc}&&&1\\&&1&\\\hline&1&&\\1&&&\end{array}\right)=\begin{pmatrix}\boldsymbol{O}&\boldsymbol{J}\\\boldsymbol{J}&\boldsymbol{O}\end{pmatrix},$$

由于

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

就有

$$X = \begin{pmatrix} I & J \\ J & -I \end{pmatrix}\begin{pmatrix} O & J \\ J & O \end{pmatrix}\begin{pmatrix} I & J \\ J & -I \end{pmatrix}$$

$$= \begin{pmatrix} I & J \\ -J & I \end{pmatrix}\begin{pmatrix} I & J \\ J & -I \end{pmatrix}$$

$$= \begin{pmatrix} 2I & O \\ O & -2I \end{pmatrix},$$

所以
$$X = \begin{pmatrix} 2 & & & \\ & 2 & & \\ & & -2 & \\ & & & -2 \end{pmatrix}. \qquad \blacksquare$$

容易检验,这与直接按四阶矩阵乘法来作结果是一样的 .

例 2 求下三角矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积 .其中

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

解 把 $\boldsymbol{A}$, $\boldsymbol{B}$ 各分成 4 块,其中 $\boldsymbol{I}$ 与 $\boldsymbol{B}_1$ 是 2×2 矩阵, $\boldsymbol{A}_3$ 和 $\boldsymbol{B}_3$ 是 1×2 矩阵, $\boldsymbol{A}_4$ 和 $\boldsymbol{B}_4$ 是 1×1 矩阵, $\boldsymbol{O}$ 是 2×1 矩阵 .即

$$\boldsymbol{A} = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline -1 & 1 & 2 \end{array}\right) = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{pmatrix},$$

$$\boldsymbol{B} = \left(\begin{array}{cc|c} 3 & 0 & 0 \\ -1 & 2 & 0 \\ \hline 2 & -1 & 4 \end{array}\right) = \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{O} \\ \boldsymbol{B}_3 & \boldsymbol{B}_4 \end{pmatrix},$$

那么

$$\boldsymbol{AB} = \begin{pmatrix} I & \boldsymbol{O} \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{pmatrix}\begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{O} \\ \boldsymbol{B}_3 & \boldsymbol{B}_4 \end{pmatrix}$$

$$= \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{O} \\ \boldsymbol{A}_3\boldsymbol{B}_1 + \boldsymbol{A}_4\boldsymbol{B}_3 & \boldsymbol{A}_4\boldsymbol{B}_4 \end{pmatrix}.$$

由

$$\boldsymbol{A}_3\boldsymbol{B}_1 = (-1 \quad 1)\begin{pmatrix} 3 & 0 \\ -1 & 2 \end{pmatrix} = (-4, 2),$$

$$\boldsymbol{A}_4\boldsymbol{B}_3 = (4, \ -2)$$

知 $\boldsymbol{A}_3\boldsymbol{B}_1 + \boldsymbol{A}_4\boldsymbol{B}_3 = \boldsymbol{O}$,那么

$$\boldsymbol{AB} = \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_4\boldsymbol{B}_4 \end{pmatrix} = \begin{pmatrix} 3 & & \\ -1 & 2 & \\ 0 & 0 & 8 \end{pmatrix}. \qquad \blacksquare$$

用这种方法容易证出下(上)三角矩阵的乘积仍是下(上)三角矩阵.

例 3　求矩阵 $\boldsymbol{A}$ 的逆矩阵

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

解　把 $\boldsymbol{A}$ 分成四块,其中 $\boldsymbol{A}_1$ 是 2 阶矩阵,$\boldsymbol{A}_2$ 是 3 阶矩阵,两个 $\boldsymbol{O}$ 矩阵分别是 2×3 与 3×2 矩阵.即

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_2 \end{pmatrix},$$

设 $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的逆矩阵,并对 $\boldsymbol{X}$ 相应的分块

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \\ \boldsymbol{X}_3 & \boldsymbol{X}_4 \end{pmatrix},$$

那么

$$\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_2 \end{pmatrix}\begin{pmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \\ \boldsymbol{X}_3 & \boldsymbol{X}_4 \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_1\boldsymbol{X}_1 & \boldsymbol{A}_1\boldsymbol{X}_2 \\ \boldsymbol{A}_2\boldsymbol{X}_3 & \boldsymbol{A}_2\boldsymbol{X}_4 \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix},$$

比较两边得到

$$A_1X_1 = I,$$
$$A_1X_2 = O,$$
$$A_2X_3 = O,$$
$$A_2X_4 = I,$$

因为 A_1, A_2 都可逆,有

$$X_1 = A_1^{-1},\quad X_2 = O,\quad X_3 = O,\quad X_4 = A_2^{-1}.$$

由于

$$A_1^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix},\quad A_2^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

所以

$$A^{-1} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$ ■

例 4 已知 $A = \begin{pmatrix} O & B \\ C & O \end{pmatrix}$,其中 B, C 分别是 m 阶, n 阶可逆矩阵.证明 A 可逆并求 A^{-1}.

证 因为 B, C 都可逆,有 $|B| \neq 0$, $|C| \neq 0$,由行列式的拉普拉斯展开式(1.8)

$$|A| = \begin{vmatrix} O & B \\ C & O \end{vmatrix} = (-1)^{m\times n}|B|\cdot|C| \neq 0,$$

所以 A 可逆.设 A^{-1} 为

$$X = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}.$$

按分块矩阵乘法, X_1 是 $n\times m$ 矩阵, X_2 是 n 阶矩阵, X_3 是 m 阶

矩阵，X_4 是 $m\times n$ 矩阵.那么

$$\begin{pmatrix} O & B \\ C & O \end{pmatrix}\begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix} = \begin{pmatrix} I & O \\ O & I \end{pmatrix},$$

即有

$$BX_3 = I,$$

$$BX_4 = O,$$

$$CX_1 = O,$$

$$CX_2 = I.$$

据 B, C 都可逆，可得到

$$X_3 = B^{-1}, X_4 = O, \ X_1 = O, \ X_2 = C^{-1}.$$

所以

$$\begin{pmatrix} O & B \\ C & O \end{pmatrix}^{-1} = \begin{pmatrix} O & C^{-1} \\ B^{-1} & O \end{pmatrix}.$$ ■

下面来看把分块矩阵用到方程组的一些情况.

对三元一次方程组

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1,$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2,$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3.$$

用矩阵乘法可表示成

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

若对系数矩阵 A 按列分成三块，依次记为 $\alpha_1, \alpha_2, \alpha_3$，右端矩阵记为 β，则方程组又可写成：

$$(\alpha_1 \ \ \alpha_2 \ \ \alpha_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \beta$$

或

$$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=\beta,$$

方程组的这种种表示方法我们应当熟悉 .

又若 $\boldsymbol{AB}=\boldsymbol{O}$,其中 $\boldsymbol{A}$ 是 3 阶矩阵, $\boldsymbol{B}$ 是 $3\times n$ 矩阵,对 $\boldsymbol{B}$ 和 $\boldsymbol{O}$ 矩阵按列分块有

$$\begin{aligned}\boldsymbol{AB}&=\boldsymbol{A}(\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_n)=(\boldsymbol{AB}_1,\boldsymbol{AB}_2,\cdots,\boldsymbol{AB}_n)\\&=(\boldsymbol{O},\boldsymbol{O},\cdots,\boldsymbol{O}),\end{aligned}$$

那就有

$$\boldsymbol{AB}_i=\boldsymbol{O}\quad(i=1,2,\cdots,n).$$

这样我们可理解为矩阵 $\boldsymbol{B}$ 的每一列 $\boldsymbol{B}_i$ 都是齐次方程组

$$\boldsymbol{A}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

的解 .

例 5　已知 $\boldsymbol{A}^2=\boldsymbol{A}$ 且 $\boldsymbol{A}\neq\boldsymbol{I}$,证明: $|\boldsymbol{A}|=0$.

证 1　由 $\boldsymbol{A}^2=\boldsymbol{A}$,有

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

因为 $\boldsymbol{A}\neq\boldsymbol{I}$,故 $\boldsymbol{A}-\boldsymbol{I}\neq\boldsymbol{O}$.而 $\boldsymbol{A}-\boldsymbol{I}$ 的每一列都是齐次方程组 $\boldsymbol{AX}=\boldsymbol{O}$ 的解,因此 $\boldsymbol{Ax}=\boldsymbol{O}$ 有非零解 .

故
$$|\boldsymbol{A}|=0\qquad\blacksquare$$

证 2　(反证法)　如 $|\boldsymbol{A}|\neq0$,则 $\boldsymbol{A}$ 可逆,在等式 $\boldsymbol{A}^2=\boldsymbol{A}$ 的两端用 $\boldsymbol{A}^{-1}$ 左乘,得到

$$\boldsymbol{A}=\boldsymbol{A}^{-1}\boldsymbol{A}^2=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$$

这与 $\boldsymbol{A}\neq\boldsymbol{I}$ 相矛盾 .

故
$$|\boldsymbol{A}|=0.\qquad\blacksquare$$

例 6　$\boldsymbol{A}$ 是三阶矩阵,如果任一个三维向量都是齐次方程组 $\boldsymbol{Ax}=\boldsymbol{O}$ 的解 .证明 $\boldsymbol{A}=\boldsymbol{O}$.

证 1 由于任一个三维向量都是方程组的解,那么$(1,0,0)^T$也是 $Ax=O$ 的解,代入到方程组有

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

得

$$a_{11}=a_{21}=a_{31}=0.$$

类似地,把$(0,1,0)^T$,$(0,0,1)^T$ 是解,代入可得

$$a_{12}=a_{22}=a_{32}=0,\ a_{13}=a_{23}=a_{33}=0.$$

由于

$$a_{ij}\equiv 0,\ \forall\ i,j.$$

所以

$$A=O. \qquad \blacksquare$$

证 2 由于任一个三维向量都是方程组的解,那么 I 的三个列向量也应是解.于是

$$\begin{aligned} A &= AI = A(\varepsilon_1\ \ \varepsilon_2\ \ \varepsilon_3) \\ &= (A\varepsilon_1\ \ A\varepsilon_2\ \ A\varepsilon_3) = (0,0,0) = O \qquad \blacksquare \end{aligned}$$

习 题 2

2.1 若 $A=\begin{pmatrix} 3 & -2 \\ 5 & -4 \end{pmatrix}$, $B=\begin{pmatrix} 3 & 4 \\ 2 & 5 \end{pmatrix}$,求:

$$2A+B,\ A-B,\ AB,\ A^TB,\ A^2,\ |B-A|.$$

2.2 计算下列乘积

(1) $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}(-2\ \ 1)$, (2) $\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$,

(3) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2$, (4) $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}^5$,

(5) $\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$, (6) $\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$.

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

$\boldsymbol{C}=\begin{pmatrix}1& & \\ 0&1& \\ -1&0&1\end{pmatrix}$，求$(\boldsymbol{AB})\boldsymbol{C}$及$\boldsymbol{A}(\boldsymbol{BC})$.

2.4 求矩阵 $\boldsymbol{A}$ 的伴随矩阵 $\boldsymbol{A}^*$

(1) $\boldsymbol{A}=\begin{pmatrix}3&2\\ -1&4\end{pmatrix}$，(2) $\boldsymbol{A}=\begin{pmatrix}a&b\\ c&d\end{pmatrix}$,

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

2.5 求 $\boldsymbol{A}$ 的逆矩阵

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

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

2.6 求矩阵 $\boldsymbol{X}$

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

(2) $\begin{pmatrix}2&5\\ 1&3\end{pmatrix}\boldsymbol{X}=\begin{pmatrix}4&-6\\ 2&0\end{pmatrix}$，

(3) $\boldsymbol{X}\begin{pmatrix}2&3\\ 4&5\end{pmatrix}=\begin{pmatrix}-2&1\\ 0&4\\ 2&0\end{pmatrix}$，

(4) $\boldsymbol{AX}=\boldsymbol{A}+2\boldsymbol{X}$，其中 $\boldsymbol{A}=\begin{pmatrix}4&2\\ 1&1\end{pmatrix}$，

(5) $AX+I=A^2+X$, 其中 $A=\begin{pmatrix}3&4\\5&6\end{pmatrix}$.

2.7　举例说明下列命题不正确.

(1) 如 $A^2=O$, 则 $A=O$;

(2) 如 $A^2=A$, 则 $A=O$ 或 $A=I$;

(3) 如 $A^2=I$, 则 $A=I$ 或 $A=-I$;

(4) 如 $A\neq O$, 则 $|A|\neq O$.

2.8　如 $AB=BA$, 则称 A 与 B 可交换.求和 A 可交换的所有矩阵.

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

2.9　A 是任一 n 阶矩阵, 则下列命题中不正确的是____.

(1) $A^*A=AA^*$,　(2) $A^2\cdot A^3=A^3\cdot A^2$,

(3) $(A-I)(A+I)=(A+I)(A-I)$,　(4) $AA^T=A^TA$.

2.10　A 是 3 阶矩阵, $|A|=2$, A 的伴随矩阵是 A^*, 则 $|2A^*|=$________.

(1) 4,　(2) 8,　(3) 16,　(4) 32.

2.11　A, B 都是 n 阶矩阵, $(AB)^2=I$, 则下列各式中肯定正确的是________.

(1) $A^{-1}=B$,　(2) $(BA)^2=I$,

(3) $B^{-1}=ABA$,　(4) $BAB=A^{-1}$.

2.12　已知 $A=\frac{1}{2}\begin{pmatrix}2&0&0\\0&1&3\\0&2&5\end{pmatrix}$, 求 $|A|$, A^{-1}, $(A^*)^{-1}$.

2.13　用初等变换法求 2.5 题中矩阵 A 的逆矩阵.

2.14　求矩阵 $\begin{pmatrix}2&1&-1&1\\3&-2&1&-3\\1&4&-3&5\end{pmatrix}$ 的等价标准形.

2.15 证明：当 $\boldsymbol{A}$ 是对称矩阵时，$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 也是对称矩阵．

2.16 A，B 都是 n 阶对称矩阵，证明：$\boldsymbol{AB}$ 是对称矩阵的充要条件是 $\boldsymbol{AB}=\boldsymbol{BA}$．

2.17 如 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$，称 $\boldsymbol{A}$ 是反对称矩阵．证明 $\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}}$ 是反对称矩阵，且任一 n 阶矩阵都可表示成一个对称矩阵与一个反对称矩阵之和．

2.18 设 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{B} & \boldsymbol{D}\\ \boldsymbol{O} & \boldsymbol{C}\end{pmatrix}$，其中 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 都可逆，证明 $\boldsymbol{A}$ 可逆并求 $\boldsymbol{A}^{-1}$．

2.19 用分块矩阵求乘积 $\boldsymbol{AB}$．

$$\boldsymbol{A}=\begin{pmatrix}1 & 2 & 0 & 0\\ 2 & 8 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 2 & 3\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}1 & -1 & 0 & 0\\ -2 & 1 & 0 & 0\\ 2 & 0 & 2 & 3\\ 0 & 1 & 1 & 2\end{pmatrix}.$$

2.20 用分块矩阵求 $\boldsymbol{A}^{-1}$

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

2.21 已知 $\boldsymbol{A}^{k}=0$，证明 $\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^{2}+\cdots+\boldsymbol{A}^{k-1}$ 的逆矩阵是 $\boldsymbol{I}-\boldsymbol{A}$．

2.22 $\boldsymbol{A}$ 是 n 阶矩阵，$\boldsymbol{A}^{2}-\boldsymbol{A}-2\boldsymbol{I}=0$，证明 $\boldsymbol{A}$ 和 $\boldsymbol{A}+2\boldsymbol{I}$ 都可逆，并求它们的逆矩阵．

2.23 已知 $\boldsymbol{A}=\boldsymbol{I}+\boldsymbol{B}$，$\boldsymbol{B}^{2}=\boldsymbol{B}$，证明 $\boldsymbol{A}$ 可逆，并求 $\boldsymbol{A}^{-1}$．

（提示：$\boldsymbol{A}^{2}=(\boldsymbol{I}+\boldsymbol{B})^{2}$）

2.24 已知 $\boldsymbol{A}=\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}$，$\boldsymbol{B}$ 是二阶矩阵，$\boldsymbol{AB}=0$ 证明：$|\boldsymbol{B}|$

$=0$. (提示: 反证法)

*2.25 A 是三阶矩阵, $AA^{\mathrm{T}}=I$, $|A|=-1$, 证明: $|A+I|=0$. (提示: 参看 2.2 例 11)

*2.26 A 是 n 阶反对称矩阵, 证明 $(I-A)(I+A)^{-1}$ 是正交矩阵. (提示: 用定义, 注意 $I-A$ 与 $I+A$ 可交换)

*2.27 证明: 两个 n 阶上三角矩阵的乘积仍是 n 阶上三角矩阵.

*2.28 计算 A^n

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

*2.29 设 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$, 对 n 阶矩阵 A, 规定 $f(A)=a_0I+a_1A+a_2A^2+\cdots+a_nA^n$.

(1) 如 $A=\begin{pmatrix}1&0&0\\0&0&1\\0&0&0\end{pmatrix}$, $f(x)=1+x+x^2+\cdots+x^{100}$ 求 $f(A)$.

(2) 如 $f(x)=x^2+2x+3$, B 是 n 阶矩阵, 满足 $f(B)=0$, 求 B^{-1}

*2.30 已知 $A=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\&a_{22}&a_{23}\\&&a_{33}\end{pmatrix}$ 是正交矩阵, 证明 A 是对角矩阵, 且 $a_{ii}(i=1,2,3)$ 是 $+1$ 或 -1.

第 3 章　线 性 方 程 组

在前面已学过用克莱姆法则解线性方程组,但这种方法要求方程组中方程的个数要等于未知数的个数,而且还要求其系数行列式非零,一方面这种方法局限性较大,另一方面当未知数的个数 n 较大时,这 $n+1$ 个 n 阶行列式的计算工作量也太大 .为此,本章以矩阵为工具,利用向量组的理论来讨论一般的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{3.1}$$

何时有解 ? 如果有解,它共有多少组解 ? 怎样求出其所有解(或称一般解,或称通解) ?

3.1　高斯(Gauss)消元法

在中学代数里大家已学习过用加减消元法解三元一次方程组,现在通过例题重新分析这种解题方法 .例如,解方程组

$$\begin{cases} x_1 + 4x_2 - 3x_3 = 2, \\ 2x_1 + 5x_2 - x_3 = 11, \\ x_1 + 6x_2 + x_3 = 12 . \end{cases} \tag{Ⅰ}$$

把第 1 个方程的 - 2 倍加至第 2 个方程,并把第 1 个方程的 - 1 倍加至第 3 个方程,得到的这个新方程组与原方程组是同解的,而新方程组中的后两个方程的未知数 x_1 都被消去 .即是

$$\begin{cases} x_1 + 4x_2 - 3x_3 = 2 \\ -3x_2 + 5x_3 = 7 \\ 2x_2 + 4x_3 = 10 \end{cases} \qquad (\text{Ⅱ})$$

对方程组(Ⅱ),其第 3 个方程可化简(约去公因数 2),并与第 2 个方程互换位置,得到同解的方程组(Ⅲ),即

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

对方程组(Ⅲ),把第 2 个方程的 3 倍加至第 3 个方程消去未知数 x_2,得到同解的方程组(Ⅳ)

$$\begin{cases} x_1 + 4x_2 - 3x_3 = 2 \\ x_2 + 2x_3 = 5 \\ 11x_3 = 22 \end{cases} \qquad (\text{Ⅳ})$$

这个方程组的解已容易求出: 由最后一个方程求出 $x_3 = 2$,代入到第 2 个方程得 $x_2 = 1$,然后把 $x_2 = 1$, $x_3 = 2$ 代入到第 1 个方程得到 $x_1 = 4$.由于这些方程组有着完全相同的解,所以原方程组有唯一解,且其解是

$$x_1 = 4,$$
$$x_2 = 1,$$
$$x_3 = 2 .$$

加减消元法解方程组的步骤可分成“正向消元”及“反向代入”两大步 .所谓正向消元就是利用“第 1 个方程”的某些倍数依次把其它方程中 x_1 的系数化为零,再利用“第 2 个方程”中 x_2 的系数消去下面方程中的 x_2,依次类推,直至最后得到含有一个未知数的一个方程 .在前面的例题中由方程组(Ⅰ)同解变形为方程组(Ⅳ)的过程就叫做正向消元 .而“反向代入”是由下至上,从最下一个方程先求出最后一个未知数,再由倒数第 2 个方程求出倒数第

2 个未知数,最后由第 1 个方程求出 x_1 .上例中,依次求出 $x_3 = 2$, $x_2 = 1$, $x_1 = 4$ 的过程就叫反向代入 .

从这个例子我们看到,加减消元法的正向消元从本质上看是对方程组反复进行一些同解变换,而同解变换是指如下三种变换

(1) 用非 0 的常数乘某一方程;

(2) 互换两个方程的位置;

(3) 把某个方程的 k 倍加到另一个方程上 .

由于加减消元是对未知数的系数进行运算,为了书写简洁,消元时思路清晰,可以把方程组中的未知数,加号及等号全都省略,运用矩阵的理论来求方程组的解 .

我们把线性方程组(3.1)的系数及常数项写成矩阵的形式

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} . \tag{3.2}$$

这个 $m\times(n+1)$ 矩阵 $\boldsymbol{A}$ 比起方程组(3.1)的系数矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{3.3}$$

多了一列常数项 .称矩阵 $\boldsymbol{A}$ 是方程组(3.1)的增广矩阵 .这时,加减消元把方程组变成其同解方程组的正向消元就转化成为对增广矩阵作矩阵的初等行变换化其为阶梯形矩阵,而反向代入就是由阶梯形矩阵的最下面的非零行开始求未知数的值,依次往上求解 .这种用矩阵初等行变换解方程组的方法就是著名的高斯消元法 .

例 1 用高斯消元法解方程组

$$(1)\ \begin{cases} x + y = 1, \\ 2x - 3y = 12 . \end{cases} \qquad (2)\ \begin{cases} 3x - 2y = 4, \\ 4x - 5y = 3 . \end{cases}$$

解　对增广矩阵 $\boldsymbol{A}$ 作初等行变换化其为阶梯形,有

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & -3 & 12 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & -5 & 10 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \end{pmatrix}$$

从第二行求出 $x_2 = -2$,代入到第一行得 $x_1 = 3$.所以,方程组(1)的解是

$$\begin{cases} x_1 = 3, \\ x_2 = -2 . \end{cases}$$

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

在正向消元时,为回避分数计算的烦琐,先把第二行的 -1 倍加至第一行,使 x_1 的系数为 -1,然后再向下消元就简便了.经反向代入可得(2)的解

$$\begin{cases} x_1 = 2, \\ x_2 = 1 . \end{cases}$$ ■

下面再通过一些简单例题,我们来看在正向消元时矩阵都可能出现哪些形式?它们对方程组的解又有什么影响?

例 2　解方程组

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

解　对增广矩阵作初等行变换,有

$$\begin{pmatrix} 1 & 3 & -4 & 5 \\ 1 & -1 & 2 & 2 \\ 3 & 5 & -6 & 10 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -4 & 5 \\ 0 & -4 & 6 & -3 \\ 0 & -4 & 6 & -5 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 3 & -4 & 5 \\ 0 & -4 & 6 & -3 \\ 0 & 0 & 0 & -2 \end{pmatrix} .$$

由于第 3 行表示的方程是:

$$0x_1 + 0x_2 + 0x_3 = -2$$

这个方程是无解的,因此,原方程组无解 . ■

例 3 解方程组

$$\begin{cases} x_1 + 2x_2 - x_3 = 1, \\ 2x_1 + x_2 + x_3 = -4, \\ 4x_1 + 5x_2 - x_3 = -2. \end{cases}$$

解 对增广矩阵作初等行变换,有

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

$$\to \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & -3 & 3 & -6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

如对未知数 x_3 随意赋值 t,即令 $x_3 = t$,由第 2 个方程可求得 $x_2 = t + 2$,代入到第 1 个方程求出 $x_1 = -t - 3$.由于参数 t 的任意性,知方程组有无穷多组解

$$\begin{cases} x_1 = -t - 3, \\ x_2 = t + 2, \\ x_3 = t. \end{cases} \qquad t \text{ 为任意常数}.$$ ■

例 4 解方程组

$$\begin{cases} x_1 + 3x_2 + 2x_3 = 3, \\ 2x_1 + 6x_2 + 5x_3 = 8, \\ -x_1 - 3x_2 + x_3 = 3. \end{cases}$$

解 对增广矩阵作初等行变换,有

$$\begin{pmatrix} 1 & 3 & 2 & 3 \\ 2 & 6 & 5 & 8 \\ -1 & -3 & 1 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

由第 2 个方程得到: $x_3 = 2$,代入到第 1 个方程并令 $x_2 = t$,则

有 $x_1 = -3t - 1$ **.因此,方程组有无穷多组解**

$$\begin{cases} x_1 = -3t-1, \\ x_2 = t, \qquad t\text{为任意常数} \\ x_3 = 2. \end{cases}$$ ■

通过这些例题我们看到,高斯消元法的本质是通过矩阵的初等行变换将增广矩阵化为阶梯形,然后对阶梯形的方程组来研究其解的种种情况.下面给出一般情形,对线性方程(3.1),设其增广矩阵 $\boldsymbol{A}$ 经过一系列初等行变换化成阶梯形矩阵

$$\begin{pmatrix} c_{11} & c_{12} & \cdots & \cdots & \cdots & c_{1n} & d_1 \\ & c_{22} & \cdots & \cdots & \cdots & c_{2n} & d_2 \\ & & \ddots & & & \cdots & \cdots \\ & 0 & & c_{rr} & \cdots & c_{rn} & d_r \\ & & & & & & d_{r+1} \end{pmatrix} \tag{3.4a}$$

或阶梯形矩阵形如

$$\begin{pmatrix} c_{11} & \cdots & \cdots & \cdots & \cdots & \cdots & c_{1n} & d_1 \\ & c_{2i} & \cdots & \cdots & \cdots & \cdots & c_{2n} & d_2 \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & 0 & & c_{rs} & \cdots & c_{rn} & d_r \\ & & & & & & & d_{r+1} \end{pmatrix} \tag{3.4b}$$

显然,

如 $d_{r+1}\neq 0$,则同解方程组的第 $r+1$ 个方程矛盾无解,故原方程组无解;

如 $d_{r+1}=0$,方程组必然有解.并且

当阶梯形方程组中,方程的个数 r 等于未知数个数 n 时,方程组有唯一解;

当阶梯形方程组中,方程的个数 r 小于未知数的个数 n 时,方程组就有无穷多组解.这时,方程组中有些未知数可以任意赋

值,这样的未知数称为自由变量.(例如,例题 3 中的 x_3,例 4 中的 x_2)自由变量的个数为 $n-r$.

总之,高斯消元法的第一步是用初等行变换化增广矩阵为阶梯形,进而判断方程组是否有解.当方程组有解时,再利用阶梯形方程组由下往上依次求解.

例 5 对方程组

$$\begin{cases} x_1+x_2+ax_3=4, \\ x_1+2x_3=a, \\ -x_1+ax_2+x_3=0. \end{cases}$$

讨论 a 为何值时,方程组有解?当方程组有解时求出其所有解.

解 对增广矩阵化为阶梯形,有

$$\begin{pmatrix} 1 & 1 & a & 4 \\ 1 & 0 & 2 & a \\ -1 & a & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & a & 4 \\ 0 & -1 & 2-a & a-4 \\ 0 & a+1 & a+1 & 4 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 1 & a & 4 \\ 0 & -1 & 2-a & a-4 \\ 0 & 0 & (a+1)(3-a) & a(a-3) \end{pmatrix}.$$

如 $a\neq-1$ 且 $a\neq3$,方程组有唯一解

$$x_1=\frac{a^2+3a}{a+1},$$

$$x_2=\frac{a+4}{a+1},$$

$$x_3=-\frac{a}{a+1}.$$

如 $a=-1$,第 3 个方程矛盾,原方程组无解;如 $a=3$,方程组有无穷多组解.这时方程组是:

$$\begin{cases} x_1+x_2+3x_3=4, \\ -x_2-x_3=-1. \end{cases}$$

令 $x_3=t$,得 $x_2=1-t$,解出 $x_1=3-2t$
故方程组的解是

$$\begin{aligned} x_1 &= 3-2t, \\ x_2 &= 1-t, \\ x_3 &= t. \quad t \text{ 是任意常数}. \end{aligned}$$

■

对方程组(3.1)有关解的论述运用到齐次方程组,就有

定理 3.1 对齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0. \end{cases} \tag{3.5}$$

如 $m<n$,则(3.4)必有非零解.

证 把方程组化成阶梯形方程组之后,方程的个数 r 不会比原方程组中方程的个数多,故有: $r\leq m<n$
对齐次方程组总有 $d_{r+1}=0$,现又有 $r<n$,因此方程组的解不唯一.所以(3.5)必有非零解. ■

定理 3.2 齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=0. \end{cases} \tag{3.6}$$

有非零解的充分必要条件是系数行列式 $|A|=0$.

证 必要性由克莱姆法则给出(见第 1 章定理 1.3).下面证明充分性.

当把系数矩阵化成阶梯形矩阵时,虽然行列式的值可能会改变大小(假若用到了某行乘 k 这一初等行变换)或改变正负号(假若用到了两行互换或某行乘 -1 之类初等行变换),但是行列式的

值为 0 或不为 0 这一性质不会由初等行变换而改变 .因此,如系数行列式 $|\boldsymbol{A}| = 0$,则高斯消元把其变成阶梯形矩阵后其行列式也必为零 .那么,在阶梯形矩阵中必然有: $r < n$.

从而方程组有无穷多组解,即方程组有非零解 . ■

3.2 向量的线性相关

为了把线性方程组解的理论搞清楚,现在以向量为工具,把方程组改写成向量的形式,用向量的理论来研究方程组解的问题 .

首先,我们把空间解析几何中三维向量的概念、运算推广到更一般的情形 .

定义 3.1 n 个有顺序的数 $a_1, a_2, \cdots, a_n$ 所组成的有序数组

$\boldsymbol{\alpha} = (a_1, a_2, \cdots, a_n)^{\mathrm{T}}$ 或 $\boldsymbol{\alpha} = (a_1, a_2, \cdots, a_n)$.叫做 $\boldsymbol{n}$ 维向量 .其中 $a_1, a_2, \cdots, a_n$ 叫做向量 $\boldsymbol{\alpha}$ 的分量(或坐标) .

向量其实是一类特殊的矩阵, n 维行向量是 $1\times n$ 矩阵, n 维列向量是 $n\times 1$ 矩阵 .由此,矩阵的相关概念和运算都可平行地用到 n 维向量上 .例如,设

$\boldsymbol{\alpha} = (a_1, a_2, \cdots, a_n)^{\mathrm{T}}$ 和 $\boldsymbol{\beta} = (b_1, b_2, \cdots, b_n)^{\mathrm{T}}$ 都是 n 维(列)向量,那么

$$\boldsymbol{\alpha} = \boldsymbol{\beta} \qquad a_i = b_i,\ i = 1,2,\cdots,n .$$

$$\boldsymbol{\alpha} = \boldsymbol{0} \qquad a_i = 0,\ i = 1,2,\cdots,n .$$

$$\boldsymbol{\alpha} + \boldsymbol{\beta} \overset{\triangle}{\ } (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n)^{\mathrm{T}},$$

$$k\boldsymbol{\alpha} \overset{\triangle}{\ } (ka_1, ka_2, \cdots, ka_n)^{\mathrm{T}},$$

$$-\boldsymbol{\alpha} \overset{\triangle}{\ } (-a_1, -a_2, \cdots, -a_n)^{\mathrm{T}} .$$

并且向量的加法和数乘运算满足下列 8 条性质:

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

(2) $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$;

(3) $\alpha+0=0+\alpha=\alpha$;

(4) $\alpha+(-\alpha)=0$;

(5) $1\alpha=\alpha$;

(6) $k(l\alpha)=(kl)\alpha$;

(7) $k(\alpha+\beta)=k\alpha+k\beta$;

(8) $(k+l)\alpha=k\alpha+l\alpha$.

定义 3.2 对 n 维向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 和 β,如存在实数 $k_1,k_2,\cdots,k_s$ 使得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=\beta$$

则称 β 是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的线性组合,或者说 β 可由 $\alpha_1,\alpha_2,\cdots\alpha_s$ 线性表出.

例如,$\alpha_1=(1,0)^{\mathrm{T}}$,$\alpha_2=(0,1)^{\mathrm{T}}$,$\beta=(3,-2)^{\mathrm{T}}$,则由

$$3\alpha_1-2\alpha_2=\beta$$

知 β 是 α_1,α_2 的线性组合.

又如,$\alpha_1=(1,2,4)^{\mathrm{T}}$,$\alpha_2=(2,1,5)^{\mathrm{T}}$,$\alpha_3=(-1,1,-1)^{\mathrm{T}}$,$\beta=(1,-4,-2)^{\mathrm{T}}$,可以验证出

$$\begin{aligned}\beta&=-3\alpha_1+2\alpha_2+0\alpha_3=-\alpha_1+0\alpha_2-2\alpha_3\\&=-2\alpha_1+\alpha_2-\alpha_3=\cdots\end{aligned}$$

我们看出 β 可由 $\alpha_1,\alpha_2,\alpha_3$ 线性表出,而且表示方法是不唯一的.

一般地,怎样才能判断出 β 能不能由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出?如果能线性表出,那么表示方法是唯一还是不唯一呢?

设 $\alpha_1=(a_{11},a_{21},\cdots,a_{n1})^{\mathrm{T}}$,$\alpha_2=(a_{12},a_{22},\cdots,a_{n2})^{\mathrm{T}},\cdots$,$\alpha_s=(a_{1s},a_{2s},\cdots,a_{ns})^{\mathrm{T}}$,$\beta=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$.若

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_s\alpha_s=\beta,$$

按向量的分量写出来就是

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s &= b_1,\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s &= b_2,\\
&\cdots\cdots\\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{ns}x_s &= b_n .
\end{aligned}
$$

因此,β能否由$\alpha_1,\alpha_2,\cdots,\alpha_s$线性表出就等价于上述方程组是否有解?而表示法唯一就等价于上述方程组有唯一解.

例 1 设$\alpha_1=(1,2,3)^T,\alpha_2=(1,3,4)^T,\alpha_3=(2,-1,1)^T,\beta=(2,5,8)^T$,问$\beta$能否用$\alpha_1,\alpha_2,\alpha_3$线性表出.

解 设$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=\beta$,按分量写出就有

$$
\begin{aligned}
x_1 + x_2 + 2x_3 &= 2,\\
2x_1 + 3x_2 - x_3 &= 5,\\
3x_1 + 4x_2 + x_3 &= 8 .
\end{aligned}
$$

用高斯消元法

$$
\begin{pmatrix} 1 & 1 & 2 & 2\\ 2 & 3 & -1 & 5\\ 3 & 4 & 1 & 8 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 & 2\\ 0 & 1 & -5 & 1\\ 0 & 1 & -5 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 & 2\\ 0 & 1 & -5 & 1\\ 0 & 0 & 0 & 1 \end{pmatrix} .
$$

由于方程组无解,所以β不能由$\alpha_1,\alpha_2,\alpha_3$线性表出. ■

例 2 设$\alpha_1=(2,1,1)^T,\alpha_2=(-1,2,7)^T,\alpha_3=(1,-1,-4)^T,\beta=(1,2,t)^T$.问$t$取何值时$\beta$可由$\alpha_1,\alpha_2,\alpha_3$线性表出.并写出此时的线性表达式.

解 设$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=\beta$,按分量写出为:

$$
\begin{aligned}
2x_1 - x_2 + x_3 &= 1,\\
x_1 + 2x_2 - x_3 &= 2,\\
x_1 + 7x_2 - 4x_3 &= t .
\end{aligned}
$$

用高斯消元法

$$
\begin{pmatrix} 2 & -1 & 1 & 1\\ 1 & 2 & -1 & 2\\ 1 & 7 & -4 & t \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 & 2\\ 2 & -1 & 1 & 1\\ 1 & 7 & -4 & t \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 & 2\\ 0 & -5 & 3 & -3\\ 0 & 5 & -3 & t-2 \end{pmatrix}
$$

$$\rightarrow \begin{pmatrix} 1 & 2 & -1 & 2 \\ 0 & -5 & 3 & -3 \\ 0 & 0 & 0 & t-5 \end{pmatrix}.$$

仅当 $t=5$ 时，方程组有解.因此，仅当 $t=5$ 时 β 可由 $\alpha_1, \alpha_2, \alpha_3$ 线性表出.

解方程组，令 $x_3 = u$，有 $x_2 = \frac{3+3u}{5}$，得 $x_1 = \frac{4-u}{5}$.因此，$t=5$ 时 β 可由 $\alpha_1, \alpha_2, \alpha_3$ 线性表出.且表示法不唯一.即

$$\beta = \frac{4-u}{5}\alpha_1 + \frac{3+3u}{5}\alpha_2 + u\alpha_3,\ u \text{ 是任意常数}. \qquad \blacksquare$$

定义 3.3 对 n 维向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 如存在不全为零的数 $k_1, k_2, \cdots, k_s$ 使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0$$

则称向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关.否则，称向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关.(也就是说，仅在 $k_1 = k_2 = \cdots = k_s = 0$ 时，$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0$ 才成立.或者说，只要 $k_1, k_2, \cdots, k_s$ 不全为零，那么 $k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s$ 必不为零)

线性相关，线性无关是抽象难以掌握运用的概念，我们先从简单的例题开始熟悉这一概念，然后再来讨论向量组线性相关的理论.

例 3 设 $\alpha_1 = (1,2,3)^{\mathrm{T}}$，$\alpha_2 = (2,4,6)^{\mathrm{T}}$ 由于

$$2\alpha_1 - \alpha_2 = 0$$

所以 α_1, α_2 线性相关. $\blacksquare$

例 4 设 $\alpha_1 = (1,1,1)^{\mathrm{T}}$，$\alpha_2 = (1,1,0)^{\mathrm{T}}$，$\alpha_3 = (0,0,0)^{\mathrm{T}}$ 由于

$$0\alpha_1 + 0\alpha_2 + 1\alpha_3 = 0.$$

所以 $\alpha_1, \alpha_2, \alpha_3$ 线性相关. $\blacksquare$

例 5 设 $\alpha_1 = (1,2,1)^{\mathrm{T}}$，$\alpha_2 = (2,3,5)^{\mathrm{T}}$，$\alpha_3 = (3,5,6)^{\mathrm{T}}$.由于 $\alpha_1 + \alpha_2 = \alpha_3$，即有

$$\alpha_1+\alpha_2-\alpha_3=0,$$

所以 $\alpha_1,\alpha_2,\alpha_3$ 线性相关. ■

例 6 设 $\alpha_1=(1,2,3)^T,\alpha_2=(2,3,4)^T,\alpha_3=(1,2,3)^T$.由于 $\alpha_1=\alpha_3$,即有

$$\alpha_1+0\alpha_2-\alpha_3=0.$$

所以 $\alpha_1,\alpha_2,\alpha_3$ 线性相关. ■

以上例题告诉我们,如向量组中有零向量、有相等的向量、有能用其它向量线性表出的向量时,这个向量组是线性相关的.这些例题数字都比较简单,我们还能较容易地观察出向量间的一些联系.对于较复杂一些的向量如何判断是否线性相关呢?

例 7 判断 $\alpha_1=(3,4,2),\alpha_2=(2,1,-7),\alpha_3=(1,2,4)$的线性相关性.

解 1 考查是否有不全为 0 的 k_1,k_2,k_3 使 $k_1\alpha_1+k_2\alpha_2+k_3\alpha_3=0$ 成立,等价于问齐次方程组

$$\begin{aligned}3k_1+2k_2+k_3&=0,\\4k_1+k_2+2k_3&=0,\\2k_1-7k_2+4k_3&=0.\end{aligned}$$

是否有非零解.由于系数行列式

$$\begin{vmatrix}3&2&1\\4&1&2\\2&-7&4\end{vmatrix}=\begin{vmatrix}3&2&1\\-2&-3&0\\-10&-15&0\end{vmatrix}=\begin{vmatrix}-2&-3\\-10&-15\end{vmatrix}=0,$$

齐次方程组必有非零解(定理 3.2),所以 $\alpha_1,\alpha_2,\alpha_3$ 线性相关. ■

解 2 用高斯消元的方法,把向量拼成一矩阵,并把变换后的向量写在矩阵外相应的位置上.

$$\begin{matrix}3&4&2\\2&1&-7\\1&2&4\end{matrix}\begin{matrix}\alpha_1\\\alpha_2\\\alpha_3\end{matrix}\rightarrow\begin{matrix}1&2&4\\2&1&-7\\3&4&2\end{matrix}\begin{matrix}\alpha_3\\\alpha_2\\\alpha_1\end{matrix}\rightarrow$$

$$\begin{pmatrix} 1 & 2 & 4 \\ 0 & -3 & -15 \\ 0 & -2 & -10 \end{pmatrix} \begin{matrix} \alpha_3 \\ \alpha_2 - 2\alpha_3 \\ \alpha_1 - 3\alpha_3 \end{matrix}$$

$$\rightarrow \begin{pmatrix} 1 & 2 & 4 \\ 0 & -1 & -5 \\ 0 & -2 & -10 \end{pmatrix} \begin{matrix} \alpha_3 \\ \alpha_2 - 2\alpha_3 - (\alpha_1 - 3\alpha_3) \\ \alpha_1 - 3\alpha_3 \end{matrix}$$

$$\rightarrow \begin{pmatrix} 1 & 2 & 4 \\ 0 & -1 & -5 \\ 0 & 0 & 0 \end{pmatrix} \begin{matrix} \alpha_3 \\ \alpha_2 + \alpha_3 - \alpha_1 \\ \alpha_1 - 3\alpha_3 - 2(\alpha_2 + \alpha_3 - \alpha_1) \end{matrix}.$$

由于 $\alpha_1 - 3\alpha_3 - 2(\alpha_2 + \alpha_3 - \alpha_1) = 0$,

即 $3\alpha_1 - 2\alpha_2 - 5\alpha_3 = 0$,

所以, $\alpha_1, \alpha_2, \alpha_3$ 线性相关. ■

解法 2 不仅判断了向量组 α 的线性相关性,而且进一步看到其中有向量可用其余的向量线性表出,例如

$$\alpha_1 = \frac{1}{3}(2\alpha_2 + 5\alpha_3),$$

即 α_1 可由 α_2, α_3 线性表出.

从例 7 的解法一,我们注意到向量组的线性相关等价于齐次方程组有非零解,根据定理 3.2,立即有

定理 3.3 n 个 n 维向量 $\alpha_i = (a_{i1}, a_{i2}, \cdots, a_{in})^{\mathrm{T}}$ $i = 1, 2, \cdots, n$,线性相关的充分必要条件是这些向量的分量所构成的行列式为零.即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0. \qquad (3.7)$$ ■

更一般地,对 m 个 n 维向量其线性相关性的判别方法是

定理 3.4 对 n 维向量

$$\alpha_1 = (a_{11}, a_{21}, \cdots, a_{n1})^T,$$
$$\alpha_2 = (a_{12}, a_{22}, \cdots, a_{n2})^T,$$
$$\cdots\cdots$$
$$\alpha_m = (a_{1m}, a_{2m}, \cdots, a_{nm})^T.$$

$\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性相关的充分必要条件是存在不全为 0 的数 $x_1, x_2, \cdots, x_m$ 使

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_m\alpha_m = 0$$

成立.即齐次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = 0, \\ \cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = 0. \end{cases}$$

有非零解. ■

特别地,如 $m > n$,根据定理 3.1,上述齐次方程组必有非零解.因此从向量来看,有

定理 3.5 任意 $n+1$ 个 n 维向量必定线性相关. ■

例 8 已知 $\alpha_1 = (1,0,2,3)^T$, $\alpha_2 = (1,1,3,5)^T$, $\alpha_3 = (1,-1,a,1)^T$, $\beta = (1,b,4,7)^T$.

(1) 判断 a 为何值时, $\alpha_1, \alpha_2, \alpha_3$ 线性相关?

(2) a, b 为何值时, β 可由 $\alpha_1, \alpha_2, \alpha_3$ 线性表出?

解 (1) 由定理 3.4, $\alpha_1, \alpha_2, \alpha_3$ 的线性相关性等价于齐次方程组

$$\begin{cases} x_1 + x_2 + x_3 = 0, \\ x_2 - x_3 = 0, \\ 2x_1 + 3x_2 + ax_3 = 0, \\ 3x_1 + 5x_2 + x_3 = 0. \end{cases}$$

有非零解,运用高斯消元法,有

$$\begin{pmatrix}1&1&1\\0&1&-1\\2&3&a\\3&5&1\end{pmatrix}\to\begin{pmatrix}1&1&1\\0&1&-1\\0&1&a-2\\0&2&-2\end{pmatrix}\to\begin{pmatrix}1&1&1\\0&1&-1\\0&0&a-1\\0&0&0\end{pmatrix}.$$

显然,仅当 $a=1$ 时,方程组才有非零解.所以,当 $a=1$ 时, $\alpha_1,\alpha_2,\alpha_3$ 线性相关.(例如, $x_1=2,\ x_2=x_3=-1$ 是方程组的一组解,可知 $2\alpha_1-\alpha_2-\alpha_3=0$,即 $\alpha_1,\alpha_2,\alpha_3$ 线性相关)

(2) 设 $x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=\beta$,按分量写出为

$$\begin{cases}x_1+x_2+x_3=1,\\ x_2-x_3=b,\\ 2x_1+3x_2+ax_3=4,\\ 3x_1+5x_2+x_3=7.\end{cases}$$

对增广矩阵作初等行变换化为阶梯形,有

$$\begin{pmatrix}1&1&1&1\\0&1&-1&b\\2&3&a&4\\3&5&1&7\end{pmatrix}\to\begin{pmatrix}1&1&1&1\\0&1&-1&b\\0&1&a-2&2\\0&2&-2&4\end{pmatrix}\to\begin{pmatrix}1&1&1&1\\0&1&-1&b\\0&0&a-1&2-b\\0&0&0&4-2b\end{pmatrix}.$$

当 $b\neq 2$ 时,方程组无解, β 不能由 α 线性表出;当 $b=2$ 且 $a\neq 1$ 时, β 可由 α 唯一的表出

$$\beta=-\alpha_1+2\alpha_2;$$

当 $b=2$ 且 $a=1$ 时, β 可由 α 线性表出,但表示方法不唯一.即

$$\beta=(-2t-1)\alpha_1+(t+2)\alpha_2+t\alpha_3,$$

其中 t 是任意常数. ■

例 9 证明 $\alpha_1=(1,0)^{\mathrm{T}},\ \alpha_2=(0,1)^{\mathrm{T}}$ 线性无关.

证 1 考查 $k_1\alpha_1+k_2\alpha_2=0$,按分量写出为

$$k_1=0,\ k_2=0,$$

根据线性无关定义, α_1, α_2 线性无关 . ■

证 2 由于向量的分量所构成的行列式

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

根据定理 3.3, α_1, α_2 线性无关 . ■

例 10 证明阶梯形向量组

$$\alpha_1 = (a_{11}, a_{12}, a_{13}, \cdots\cdots a_{1n})^T,$$
$$\alpha_2 = (0, a_{22}, a_{23}, \cdots\cdots a_{2n})^T,$$
$$\alpha_3 = (0, 0, a_{33}, \cdots\cdots a_{3n})^T,$$
$$\cdots\cdots$$
$$\alpha_r = (0, 0, 0, \cdots a_{rr} \cdots a_{rn})^T,$$

其中 $a_{ii} \neq 0$ $(i = 1, 2, \cdots, r)$, 线性无关 .

证 如果 $k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r = 0$, 写出分量就有齐次方程组

$$\begin{aligned}
a_{11}k_1 &= 0, \\
a_{12}k_1 + a_{22}k_2 &= 0, \\
a_{13}k_1 + a_{23}k_2 + a_{33}k_3 &= 0, \\
&\cdots\cdots \\
a_{1r}k_1 + a_{2r}k_2 + a_{3r}k_3 + \cdots + a_{rr}k_r &= 0 .
\end{aligned}$$

由 $a_{11} \neq 0$, 从第 1 个方程解出必有 $k_1 = 0$, 代入到第 2 个方程, 由 $a_{22} \neq 0$, 知必有 $k_2 = 0$, 如此一步步递推就得到所有 k_i 全为 0 .所以, α_1, α_2, $\cdots$, α_r 线性无关 . ■

例 11 证明: 如 α_1, α_2, α_3 线性相关, 则 α_1, α_2, α_3, α_4 线性相关 .

证 由于 α_1, α_2, α_3 线性相关, 故存在不全为 0 的 k_1, k_2, k_3 使

$$k_1\alpha_1 + k_2\alpha_2 + k_3\alpha_3 = 0$$

成立.那么, $k_1, k_2, k_3, 0$ 仍不全为零,而

$$k_1\alpha_1 + k_2\alpha_2 + k_3\alpha_3 + 0\alpha_4 = 0.$$

因此, $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 线性相关. ■

例 12 证明:如 $\alpha_1, \alpha_2, \alpha_3$ 线性无关,则 α_1, α_2 线性无关.

证(反证法) 如 α_1, α_2 线性相关,由例 11 知 $\alpha_1, \alpha_2, \alpha_3$ 线性相关,这与已知条件相矛盾.

因此, α_1, α_2 线性无关. ■

例 13 设向量组

$$\alpha_1 = (a_{11}, a_{21}, a_{31})^T, \alpha_2 = (a_{12}, a_{22}, a_{32})^T, \alpha_3 = (a_{13}, a_{23}, a_{33})^T$$

线性无关.证明向量组

$$\beta_1 = (a_{11}, a_{21}, a_{31}, a_{41})^T, \beta_2 = (a_{12}, a_{22}, a_{32}, a_{42})^T,$$

$$\beta_3 = (a_{13}, a_{23}, a_{33}, a_{43})^T$$

线性无关.

证 由于 $\alpha_1, \alpha_2, \alpha_3$ 线性无关,那么

$$k_1\alpha_1 + k_2\alpha_2 + k_3\alpha_3 = 0$$

仅在 $k_1 = k_2 = k_3 = 0$ 时才成立.即齐次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = 0. \end{cases} \tag{A}$$

只有零解.

考查 $x_1\beta_1 + x_2\beta_2 + x_3\beta_3 = 0$,按分量写成齐次方程组为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = 0, \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 = 0. \end{cases} \tag{B}$$

显然,方程组(B)的前 3 个方程就构成方程组(A),所以(B)的任何一个解必是(A)的解,现(A)只有零解,从而(B)也就只有零

解.也就是 β_1,β_2,β_3 线性无关. ■

通过这些例题我们看到,证明线性无关常用的方法有两个:一是先假定这些向量的一个线性组合为零,然后按分量写出来,得到一个齐次线性方程组,转而证明该方程组只有零解,进而得出线性无关的结论;另一方法是利用反证法,因为对向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$,或者是存在不全为零的 k_i 使 $k_1\alpha_1+\cdots+k_s\alpha_s=0$,或者是对任何不全为零的 k_i 都有 $k_1\alpha_1+\cdots+k_s\alpha_s\neq 0$,因此,$\alpha_1,\alpha_2,\cdots,\alpha_s$ 不是线性相关就是线性无关,二者必居其一,这样反证法是常用的.

下面讨论线性相关与线性表出之间的联系.

定理 3.6 如果 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geq 2)$ 线性相关,则其中必有一个向量可用其余的向量线性表出.反之,若有一个向量可用其余的 $s-1$ 个向量线性表出,则这 s 个向量必线性相关.

证 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关,则存在不全为零的 $k_1,k_2,\cdots,k_s$ 使得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0,$$

不妨设 $k_i\neq 0$,则

$$\alpha_i=\left(-\frac{k_1}{k_i}\right)\alpha_1+\cdots+\left(-\frac{k_{i-1}}{k_i}\right)\alpha_{i-1}+\left(-\frac{k_{i+1}}{k_i}\right)\alpha_{i+1}$$
$$+\cdots+\left(-\frac{k_s}{k_i}\right)\alpha_s,$$

即有 α_i 可用其余的 α 线性表出.

反之,若有一个向量可用其余的 $s-1$ 个向量线性表出.设 α_j 可由 $\alpha_1,\cdots,\alpha_{j-1},\alpha_{j+1},\cdots,\alpha_s$ 线性表出

$$\alpha_j=l_1\alpha_1+\cdots+l_{j-1}\alpha_{j-1}+l_{j+1}\alpha_{j+1}+\cdots+l_s\alpha_s,$$

那么移项,得

$$l_1\alpha_1+\cdots+l_{j-1}\alpha_{j-1}-\alpha_j+l_{j+1}\alpha_{j+1}+\cdots+l_s\alpha_s=0,$$

系数 $l_1,\cdots,l_{j-1},-1,l_{j+1},\cdots,l_s$ 不全为 0.从而 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关. ■

定理 3.7 如果 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关, $\alpha_1, \alpha_2, \cdots, \alpha_s, \beta$ 线性相关,则 β 可由 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性表出,且表示法唯一 .

证 因为 $\alpha_1, \alpha_2, \cdots, \alpha_s, \beta$ 线性相关,故存在不全为零的数 $k_1, k_2, \cdots, k_s, l$ 使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s + l\beta = 0,$$

如果其中 $l = 0$,那么必有 $k_1, k_2, \cdots, k_s$ 不全为 0,而

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0,$$

这与已知条件 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关相矛盾 .所以, $l \neq 0$.这样就有

$$\beta = \quad -\frac{k_1}{l}\ \alpha_1 + \quad -\frac{k_2}{l}\ \alpha_2 + \cdots + \quad -\frac{k_s}{l}\ \alpha_s,$$

即 β 可由 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性表出 .

关于表示法的唯一性 .若 β 有两种不同的表示法

$$\beta = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_s\alpha_s,$$

$$\beta = y_1\alpha_1 + y_2\alpha_2 + \cdots + y_s\alpha_s,$$

两式相减,得

$$(x_1 - y_1)\alpha_1 + (x_2 - y_2)\alpha_2 + \cdots + (x_s - y_s)\alpha_s = 0,$$

由于是两种不同的表示法, $x_i - y_i$ 不可能全为 0,那么 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关,与已知条件矛盾,所以 β 的表示法唯一 . ■

在本节的最后,我们以三维向量为例给出线性相关,线性无关在几何上的直观解释,这可帮助理解概念分辨是非 .

(1) α 线性相关 $\quad \alpha = 0$

证 如 α 线性相关,则存在 $k \neq 0$ 使 $k\alpha = 0$,两边同除以 k 即有 $\alpha = 0$.反之,若 $\alpha = 0$,则有 $1\alpha = 0$,按定义 α 线性相关 . ■

(2) α_1, α_2 线性相关 $\quad \alpha_1 /\!/ \alpha_2$.

证 如 α_1, α_2 线性相关,则存在不全为 0 的 k_1, k_2,使 $k_1\alpha_1 + k_2\alpha_2 = 0$.不妨设 $k_1 \neq 0$,那么

$$\alpha_1 = -\frac{k_2}{k_1}\alpha_2,$$

即有 $\alpha_1 \mathbin{/\!/} \alpha_2$.反之,若 $\alpha_1 \mathbin{/\!/} \alpha_2$.如 $\alpha_2 = 0$,则有

$$0\alpha_1 + 1\alpha_2 = \mathbf{0},$$

得 α_1, α_2 线性相关;如 $\alpha_2 \neq 0$,则有 $\alpha_1 = \lambda\alpha_2$,即

$$\alpha_1 - \lambda\alpha_2 = 0,$$

系数 1, $-\lambda$ 不全为 0,得 α_1, α_2 线性相关. ■

(3) $\alpha_1, \alpha_2, \alpha_3$ 线性相关　　$\alpha_1, \alpha_2, \alpha_3$ 共面.

证　如 $\alpha_1, \alpha_2, \alpha_3$ 线性相关,则存在不全为 0 的 k_1, k_2, k_3,使 $k_1\alpha_1 + k_2\alpha_2 + k_3\alpha_3 = 0$.不妨设 $k_3 \neq 0$,那么

$$\alpha_3 = -\frac{k_1}{k_3}\alpha_1 - \frac{k_2}{k_3}\alpha_2.$$

由向量加法的平行四边形法则,可知 α_3 在以 α_1, α_2 为边的平行四边形上.因此 $\alpha_1, \alpha_2, \alpha_3$ 共面.反之(略). ■

借助于几何直观,很容易判断出:

如 $\alpha_1, \alpha_2, \alpha_3$ 线性无关,则 $\alpha_1 - \alpha_2, \alpha_2 - \alpha_3, \alpha_3 - \alpha_1$ 线性相关, $\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3 + \alpha_1$ 线性无关.

我们可设想一个四面体 $O-ABC$(图 3.1),记 $\overrightarrow{OA} = \alpha_1$, $\overrightarrow{OB} = \alpha_2$, $\overrightarrow{OC} = \alpha_3$.显然, $\alpha_1 - \alpha_2 = \overrightarrow{BA}$, $\alpha_2 - \alpha_3 = \overrightarrow{CB}$, $\alpha_3 - \alpha_1 = \overrightarrow{AC}$,由于 $\alpha_1 - \alpha_2, \alpha_2 - \alpha_3, \alpha_3 - \alpha_1$ 构成四面体的底面 ABC,因此它们必然线性相关.而 $\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3 + \alpha_1$ 在四面体的 3 个侧面上,它们不共面.因此是线性无关的.请读者给出严格的代数证明

又如 $\alpha_1, \alpha_2, \alpha_3$ 线性无关,则 α_1, α_2 线性无关(例 12).从几何上看, $\alpha_1, \alpha_2, \alpha_3$ 不共面,那么 α_1 与 α_2 不可能平行,因此 α_1, α_2 线性无关.

图 3.1

3.3 向 量 组 的 秩

细心的读者也许会问,在高斯消元解方程组时,会不会由于选用了不同的消元次序而导致在最后的阶梯形方程组中方程的个数 r 不一样?会不会这样消元有 $d_{r+1}=0$.而换一个方式有 $d_{r+1}\neq 0$?

本节就是利用向量线性相关、线性无关的概念建立向量组秩的概念,为后面完整的讨论方程组解的理论作准备.

设有两个向量组

$$\alpha_1,\alpha_2,\cdots,\alpha_s \tag{Ⅰ}$$

$$\beta_1,\beta_2,\cdots,\beta_t \tag{Ⅱ}$$

如果(Ⅰ)中每个向量 $\alpha_i(i=1,2,\cdots,s)$都可用 $\beta_1,\beta_2,\cdots,\beta_t$ 线性表出,则称向量组(Ⅰ)可由向量组(Ⅱ)线性表出.习惯上用记号(Ⅰ)<(Ⅱ)表示.

如果两个向量组可以互相线性表出,则称这两个向量组等价.

若向量组(Ⅰ)与(Ⅱ)等价,常记为(Ⅰ)(Ⅱ).

例如, $\alpha_1=(1,0)$, $\alpha_2=(0,1)$与 $\beta_1=(1,1)$, $\beta_2=(1,-1)$

由于 $\beta_1=\alpha_1+\alpha_2$, $\beta_2=\alpha_1-\alpha_2$;

$$\alpha_1=\frac{1}{2}\beta_1+\frac{1}{2}\beta_2,\ \alpha_2=\frac{1}{2}\beta_1-\frac{1}{2}\beta_2;$$

知 α_1,α_2 与 β_1,β_2 可以互相线性表出.所以, α_1,α_2 与 β_1,β_2 是等价的向量组.

又如, $\alpha_1=(0,0)$, $\alpha_2=(0,3)$, $\alpha_3=(0,-1)$与 $\beta_1=(1,2)$, $\beta_2=(1,3)$.由于

$$\alpha_1=0\beta_1+0\beta_2,\alpha_2=-3\beta_1+3\beta_2,\alpha_3=\beta_1-\beta_2,$$

可见 $\alpha_1,\alpha_2,\alpha_3$ 可由 β_1,β_2 线性表出.但是 β_1,β_2 不能由 $\alpha_1,\alpha_2,\alpha_3$ 线性表出.

再如, $\alpha_1,\alpha_2,\alpha_3$ 总可由 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 线性表出.但 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 不一定能用 $\alpha_1,\alpha_2,\alpha_3$ 线性表出(请举例).

容易验证向量组的等价有下面三条性质:

1) (Ⅰ)(Ⅰ);(反身性)

2) 如(Ⅰ)～(Ⅱ),则(Ⅱ)～(Ⅰ);(对称性)

3) 如(Ⅰ)～(Ⅱ),(Ⅱ)～(Ⅲ),则(Ⅰ)～(Ⅲ).(传递性)

定理 3.8 如果 β_1,β_2,β_3 可由 α_1,α_2 线性表出,则 β_1,β_2,β_3 线性相关.

*证 因为 β_1,β_2,β_3 可由 α_1,α_2 线性表出,故可设

$$\beta_1=a_{11}\alpha_1+a_{21}\alpha_2,$$
$$\beta_2=a_{12}\alpha_1+a_{22}\alpha_2,$$
$$\beta_3=a_{13}\alpha_1+a_{23}\alpha_2.$$

利用矩阵表述上述关系就是

$$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2)\begin{matrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{matrix},$$

而要证 β_1,β_2,β_3 线性相关,就是要证明存在不全为零的 k_1,k_2,k_3 使

$$(\beta_1\ \ \beta_2\ \ \beta_3)\begin{pmatrix}k_1\\k_2\\k_3\end{pmatrix}=0,$$

成立.由于齐次方程组

$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

中,方程个数 $2<$ 未知数个数 3.根据定理 3.1,此方程组必有非零解,那么这非零解使

$$(\beta_1\ \beta_2\ \beta_3)\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=(\alpha_1\ \alpha_2)\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=(\alpha_1\ \alpha_2)\begin{pmatrix}0\\0\end{pmatrix}=0$$

成立.所以 β_1,β_2,β_3 线性相关. ■

从定理 3.8 的证明可以看出,只要多数向量能由少数向量线性表出,这多数向量一定线性相关.所以定理 3.8 的一般形式是:

如 $\beta_1,\beta_2,\cdots,\beta_t$ 可由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出,且 $t>s$,则 $\beta_1,\beta_2,\cdots,\beta_t$ 线性相关.

定理 3.8 的一个非常有用的等价说法是其逆否命题:

推论 1 如 $\beta_1,\beta_2,\cdots,\beta_t$ 线性无关,且它可由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出,则 $t\leq s$.

定义 3.4 在向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中,如存在 r 个向量 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线线无关,再加进任一个 α_j 都线性相关,则称 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个极大线性无关组.

例如,在 $\alpha_1=(0,0)$, $\alpha_2=(1,0)$, $\alpha_3=(2,0)$ 中, α_2 是线性无

关的,再加进 α_1 或 α_3 都线性相关,故 α_2 是 $\alpha_1, \alpha_2, \alpha_3$ 的一个极大线性无关组.类似地, α_3 也是 $\alpha_1, \alpha_2, \alpha_3$ 的一个极大线性无关组.

又如,向量组 $\alpha_1 = (1, 0)$, $\alpha_2 = (1, 1)$, $\alpha_3 = (2, 3)$, $\alpha_4 = (3, -5)$,其中任意两个向量都是线性无关的(它们不平行),而任意三个必定线性相关(定理 3.5)从而知 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的极大线性无关组由其中任意两个向量所构成.

一般地,我们可用高斯消元的方法把向量组化为阶梯形,从而找出其极大线性无关组.

例 1 求向量组 $\alpha_1 = (1, 3, 5, -1)$, $\alpha_2 = (2, -1, -3, 4)$, $\alpha_3 = (5, 1, -1, 7)$和 $\alpha_4 = (7, 7, 9, 1)$的极大线性无关组.

解 把这些向量拼成一个矩阵,然后化其为阶梯形.即

$$\begin{array}{rrrrl} 1 & 3 & 5 & -1 & \alpha_1 \\ 2 & -1 & -3 & 4 & \alpha_2 \\ 5 & 1 & -1 & 7 & \alpha_3 \\ 7 & 7 & 9 & 1 & \alpha_4 \end{array} \rightarrow \begin{array}{rrrrl} 1 & 3 & 5 & -1 & \alpha_1 \\ 0 & -7 & -13 & 6 & \alpha_2 - 2\alpha_1 \\ 0 & -14 & -26 & 12 & \alpha_3 - 5\alpha_1 \\ 0 & -14 & -26 & 8 & \alpha_4 - 7\alpha_1 \end{array}$$

$$\rightarrow \begin{array}{rrrrl} 1 & 3 & 5 & -1 & \alpha_1 \\ 0 & -7 & -13 & 6 & \alpha_2 - 2\alpha_1 \\ 0 & 0 & 0 & 0 & \alpha_3 - 5\alpha_1 - 2(\alpha_2 - 2\alpha_1) \\ 0 & 0 & 0 & -4 & \alpha_4 - 7\alpha_1 - 2(\alpha_2 - 2\alpha_1) \end{array}$$

从矩阵的第三行看到: $\alpha_3 - 5\alpha_1 - 2(\alpha_2 - 2\alpha_1) = 0$.

即 $\alpha_1 + 2\alpha_2 - \alpha_3 = 0$(这说明 $\alpha_1, \alpha_2, \alpha_3$ 线性相关)而一、二、四行的三个向量是线性无关的.所以 $\alpha_1, \alpha_2, \alpha_4$ 是 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的一个极大线性无关组.

如果改变高斯消元的先后顺序,例如

$$\begin{matrix}1&3&5&-1\\2&-1&-3&4\\5&1&-1&7\\7&7&9&1\end{matrix}\ \begin{array}{l}\alpha_1\\\alpha_2\\\alpha_3\\\alpha_4\end{array}\to\begin{matrix}5&1&-1&7\\2&-1&-3&4\\1&3&5&-1\\7&7&9&1\end{matrix}\ \begin{array}{l}\alpha_3\\\alpha_2\\\alpha_1\\\alpha_4\end{array}$$

$$\to\begin{matrix}1&3&5&-1\\2&-1&-3&4\\1&3&5&-1\\7&7&9&1\end{matrix}\ \begin{array}{l}\alpha_3-2\alpha_2\\\alpha_2\\\alpha_1\\\alpha_4\end{array}$$

$$\to\begin{matrix}1&3&5&-1\\0&-7&-13&6\\0&0&0&0\\0&-14&-26&8\end{matrix}\ \begin{array}{l}\alpha_3-2\alpha_2\\\alpha_2-2(\alpha_3-2\alpha_2)\\\alpha_1-(\alpha_3-2\alpha_2)\\\alpha_4-7(\alpha_3-2\alpha_2)\end{array}$$

$$\to\begin{matrix}1&3&5&-1\\0&-7&-13&6\\0&0&0&0\\0&0&0&-4\end{matrix}\ \begin{array}{l}\alpha_3-2\alpha_2\\5\alpha_2-2\alpha_3\\\alpha_1+2\alpha_2-\alpha_3\\\alpha_4-7(\alpha_3-2\alpha_2)-2(5\alpha_2-2\alpha_3)\end{array}$$

可看出 $\alpha_3,\alpha_2,\alpha_4$ 是线性无关的.而 $\alpha_1=-2\alpha_2+\alpha_3+0\alpha_4$,因此,$\alpha_2,\alpha_3,\alpha_4$ 是 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的一个极大线性无关组. ■

通过这些例子我们看到,一个向量组的极大线性无关组往往有多种选取方法,但是对每个向量组的极大线性无关组来说,其向量个数又是一定的,这究竟是一种偶然的巧合,还是有其必然性呢?

例 2 如果 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 与 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_t}$ 都是向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的极大线性无关组,证明: $r=t$.

证 因为 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是极大线性无关组,添加 α_{j_k} 后必然线性相关,根据定理 3.7, α_{j_k} 可由 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性表出.因此,

$\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_t}$可由$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$线性表出,又因 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_t}$线性无关,根据定理 3.8 的推论,得到 $t\leq r$.完全类似地又可有 $r\leq t$.所以,$t=r$. ■

例 2 告诉我们,虽然极大线性无关组的选取可以不唯一,但同一个向量组其极大线性无关组中有着相同的向量个数,由此引出向量组秩的概念.

定义 3.5 向量组的极大线性无关组中所含向量的个数称为这个向量组的秩.只含零向量的向量组其秩规定为 0.

例如,$\alpha_1=(1,0),\alpha_2=(2,0),\alpha_3=(3,0)$的极大线性无关组由 1 个向量构成,所以这个向量组的秩是 1.通常记作 $\mathrm{r}(\alpha_1,\alpha_2,\alpha_3)=1$.

又如,$\beta_1=(1,1,1),\beta_2=(0,1,1),\beta_3=(0,0,1)$是一个阶梯形向量组,它们线性无关.所以这个向量组的秩是 3,即 $\mathrm{r}(\beta_1,\beta_2,\beta_3)=3$.

定理 3.9 设 $\mathrm{r}(\alpha_1,\alpha_2,\cdots,\alpha_s)=r$, $\mathrm{r}(\beta_1,\beta_2,\cdots,\beta_t)=p$,如果 $\beta_1,\beta_2,\cdots,\beta_t$ 可由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出,则 $p\leq r$.即如果向量组$(\mathrm{I})<(\mathrm{II})$,则 $\mathrm{r}(\mathrm{I})\leq\mathrm{r}(\mathrm{II})$.

*证 设 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$是$\alpha_1,\alpha_2,\cdots,\alpha_s$ 的极大线性无关组,$\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_p}$是$\beta_1,\beta_2,\cdots,\beta_t$ 的极大线性无关组

由于 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 与 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$等价(习题 15)而 $\beta_1,\beta_2,\cdots,\beta_t$ 可由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出.那么 $\beta_1,\beta_2,\cdots,\beta_t$ 可由 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$线性表出,进而有 $\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_p}<\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$.
又因 $\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_p}$线性无关,据定理 3.8 推论,得 $p\leq r$. ■

推论 2 等价的向量组有相同的秩.即如果向量组$(\mathrm{I})(\mathrm{II})$,则 $\mathrm{r}(\mathrm{I})=\mathrm{r}(\mathrm{II})$.

证 因为向量组$(\mathrm{I})(\mathrm{II})$,所以向量组$(\mathrm{I})<(\mathrm{II})$,那么 $\mathrm{r}(\mathrm{I})\leq\mathrm{r}(\mathrm{II})$.类似有 $\mathrm{r}(\mathrm{II})\leq\mathrm{r}(\mathrm{I})$.故 $\mathrm{r}(\mathrm{I})=\mathrm{r}(\mathrm{II})$. ■

例 3 已知 $\alpha_1,\alpha_2,\alpha_3$ 及 $\beta_1=\alpha_1,\beta_2=\alpha_2+l_2\alpha_1,\beta_3=\alpha_3+l_3\alpha_1$.证明: $r(\alpha_1,\alpha_2,\alpha_3)=r(\beta_1,\beta_2,\beta_3)$

证 由于 β_1,β_2,β_3 可用 $\alpha_1,\alpha_2,\alpha_3$ 线性表出,而

$$\alpha_1=\beta_1,\alpha_2=\beta_2-l_2\beta_1,\alpha_3=\beta_3-l_3\beta_1 .$$

即 $\alpha_1,\alpha_2,\alpha_3$ 也可由 β_1,β_2,β_3 线性表出 .
那么, $\alpha_1,\alpha_2,\alpha_3$ 与 β_1,β_2,β_3 是等价的向量组 .它们有相同的秩(定理 3.9 推论 2) . ■

从矩阵的角度来看例 3,若构造矩阵

$$\boldsymbol{A}=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix}\qquad \boldsymbol{B}=\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix}=\begin{pmatrix}\alpha_1\\ \alpha_2+l_2\alpha_1\\ \alpha_3+l_3\alpha_1\end{pmatrix}$$

称 $\alpha_1,\alpha_2,\alpha_3$ 是矩阵 $\boldsymbol{A}$ 的行向量, β_1,β_2,β_3 是矩阵 $\boldsymbol{B}$ 的行向量 .我们可理解为 $\boldsymbol{A}$ 经初等行变换(某行的 k 倍加至另一行)得到 $\boldsymbol{B}$.例 3 说明在这种行变换下矩阵的行向量组的秩不变 .

容易看出,矩阵的另两种初等行变换(互换两行;某行乘上不为零的常数)所得到的行向量组仍是等价向量组,因而行向量组的秩仍不变 .这样就有

定理 3.10 初等行变换不改变矩阵行向量组的秩 . ■

利用定理 3.10 及阶梯形向量组是线性无关的,就能较方便的求出向量组的秩 .

例 4 已知 $\alpha_1=(2,-1,3,4),\alpha_2=(1,3,4,0),\alpha_3=(3,-5,2,8)$ 求 $r(\alpha_1,\alpha_2,\alpha_3)$

解 构造矩阵用高斯消元化为阶梯形

$$\begin{pmatrix}2&-1&3&4\\1&3&4&0\\3&-5&2&8\end{pmatrix}\begin{matrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{matrix}\to\begin{pmatrix}1&3&4&0\\2&-1&3&4\\3&-5&2&8\end{pmatrix}\begin{matrix}\alpha_2\\ \alpha_1\\ \alpha_3\end{matrix}$$

$$\to \begin{pmatrix} 1 & 3 & 4 & 0 \\ 0 & -7 & -5 & 4 \\ 0 & -14 & -10 & 8 \end{pmatrix} \begin{matrix} \alpha_2 \\ \alpha_1 - 2\alpha_2 \\ \alpha_3 - 3\alpha_2 \end{matrix}$$

$$\to \begin{pmatrix} 1 & 3 & 4 & 0 \\ 0 & -7 & -5 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \alpha_2 \\ \alpha_1 - 2\alpha_2 \\ \alpha_3 - 3\alpha_2 - 2(\alpha_1 - 2\alpha_2) \end{matrix}$$

最后一个矩阵有两个行向量线性无关,所以原向量组的秩 $r(\alpha_1, \alpha_2, \alpha_3) = 2$.

从消元我们也可得知: α_1, α_2 **是一极大线性无关组,** $\alpha_3 = 2\alpha_1 - \alpha_2$ **可用** α_1, α_2 **线性表出 .** ■

例 5　证明: $\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3 + \alpha_1$ **线性无关的充要条件是** $\alpha_1, \alpha_2, \alpha_3$ **线性无关 .**

证　记 $\beta_1 = \alpha_1 + \alpha_2, \beta_2 = \alpha_2 + \alpha_3, \beta_3 = \alpha_3 + \alpha_1$,**则有** $\alpha_1 = \frac{1}{2}(\beta_1 + \beta_3 - \beta_2), \alpha_2 = \frac{1}{2}(\beta_1 + \beta_2 - \beta_3), \alpha_3 = \frac{1}{2}(\beta_2 + \beta_3 - \beta_1)$.

这说明 $\alpha_1, \alpha_2, \alpha_3$ **与** $\beta_1, \beta_2, \beta_3$ **可以互相线性表出,它们是等价的向量组,因而有相同的秩**

$$r(\alpha_1, \alpha_2, \alpha_3) = r(\beta_1, \beta_2, \beta_3)$$

若 $\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3 + \alpha_1$ **线性无关,即** $r(\beta_1, \beta_2, \beta_3) = 3$ **.得** $r(\alpha_1, \alpha_2, \alpha_3) = 3$,**也就是** $\alpha_1, \alpha_2, \alpha_3$ **线性无关 .反之亦然,充要性成立 .** ■

3.4　矩 阵 的 秩

在 $m \times n$ **矩阵** $\boldsymbol{A}$ **中,任意取** k **行和** k **列,位于这些行和列的交叉点上的** k^2 **个元素按其原来的次序可组成一个** k **阶行列式,称其为矩阵** $\boldsymbol{A}$ **的一个** k **阶子式 .例如,在矩阵**

$$A = \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & 4 & 6 & -3 \\ 0 & 0 & 5 & 1 \end{pmatrix}$$

中,取第 1,2 行,第 2,4 列,得到 A 的一个二阶子式为

$$\begin{vmatrix} -1 & 2 \\ 4 & -3 \end{vmatrix}$$

若取第 2,3 行,第 1,2 列,得到的二阶子式为:

$$\begin{vmatrix} 0 & 4 \\ 0 & 0 \end{vmatrix}$$

定义 3.6 矩阵 A 的非 0 子式的最高阶数称为矩阵 A 的秩,记为 $\mathrm{r}(A)$.零矩阵的秩规定为 0 .

在前面的例题中,尽管有的 2 阶子式为 0,但存在着非 0 的 3 阶子式

$$\begin{vmatrix} 1 & -1 & 3 \\ 0 & 4 & 6 \\ 0 & 0 & 5 \end{vmatrix} = 20 \neq 0,$$

而矩阵 A 中没有 4 阶子式,按定义有 $\mathrm{r}(A) = 3$.

例 1 A 是 n 阶矩阵,则

$$\mathrm{r}(A) = n \qquad |A| \neq 0 \qquad A \text{ 可逆} .$$ ■

例 2 A 是 n 阶矩阵,$\mathrm{r}(A) = n$,证明:$\mathrm{r}(A^*) = n$.

证 因 $\mathrm{r}(A) = n$,有 $|A| \neq 0$.又因 $AA^* = |A| I$,两边取行列式,由行列式乘法公式有:

$$|A||A^*| = |AA^*| = ||A| I| = |A|^n$$

得到 $|A^*| = |A|^{n-1} \neq 0$,故有 $\mathrm{r}(A^*) = n$. ■

例 3 A 是 3 阶矩阵,$\mathrm{r}(A) = 1$,证明:$A^2 = kA$.

证 由 $\mathrm{r}(A) = 1$ 知,A 的所有二阶子式全为 0,那么 A 的任何两行,任何两列都成比例 .故可设

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}=\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}(b_1 \quad b_2 \quad b_3)$$

那么

$$A^2=\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}(b_1\ b_2\ b_3)\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}(b_1\ b_2\ b_3)=(a_1b_1+a_2b_2+a_3b_3)A.$$

■

*例 4　A 是 2 阶矩阵, $A^3=0$,证明: $A^2=0$.

证　$A^3=0$,由行列式乘法有

$$|A|^3=|A^3|=0 \quad 即 \quad |A|=0$$

那么, $r(A)=0$ 或 1　(A 是二阶矩阵).

如 $r(A)=0$,即 $A=0$,自然有 $A^2=0$;

如 $r(A)=1$,由例 3 知 $A^2=kA$,那么

$$A^3=kA^2=k^2A=0$$

而 $A\neq 0$,得 $k=0$.所以 $A^2=kA=0$. ■

虽然,用定义可求矩阵的秩,但在矩阵的行数、列数较大时,计算非 0 子式的工作量是很大的.我们设想一个秩是 2 的 4×6 的矩阵,按定义要证明其秩是 2 必须检验它的所有三阶子式为 0,而三阶子式共有 $C_4^3\cdot C_6^3=80$ 个,显然这种方法是麻烦的.下面给出求矩阵秩的方法.

定理 3.11　初等变换不改变矩阵的秩.即如 $A\to B$,有 $r(A)=r(B)$.

*证　由 $A\to B$ 一般要经过一系列初等变换,我们证明 A 经过一次初等变换化成 B_1 时,必有 $r(A)=r(B_1)$,那么也就有 $r(A)=r(B_1)=\cdots=r(B)$.又由于 $A\to B \Leftrightarrow B\to A$,我们只需证明对一次初等变换有 $r(B_1)\leq r(A)$.下面以列的初等变换为例进行论证

(行的情况完全是类似的):

设 $\mathrm{r}(\boldsymbol{A})=r$,我们要证 B_1 的任意 $r+1$ 阶子式 M 全为 0.对 $\boldsymbol{A}$ 按列分块,设 $\boldsymbol{A}=(\alpha_1,\alpha_2,\cdots,\alpha_n)$

(1) 如 $\boldsymbol{A}\to\boldsymbol{B}_1=(\alpha_1,\cdots,k\alpha_i,\cdots,\alpha_n)$

若 M 不含第 i 列,则 M 就是 $\boldsymbol{A}$ 的 $r+1$ 阶子式,从而 $M=0$ (为什么?);若 M 含第 i 列,则把第 i 列的公因数 k 提出知 M 等于 $\boldsymbol{A}$ 的一个 $r+1$ 阶子式乘以 k,仍有 $M=0$.

(2) 如 $\boldsymbol{A}\to\boldsymbol{B}_1=(\alpha_1\cdots\alpha_j\cdots\alpha_i\cdots\alpha_n)$

这时 M 或为 $\boldsymbol{A}$ 的一个 $r+1$ 阶子式或者可由 $\boldsymbol{A}$ 的某 $r+1$ 阶子式对换两列后所得到,因为 $r(\boldsymbol{A})=r$, $\boldsymbol{A}$ 的所有 $r+1$ 阶子式均为 0,故 $\boldsymbol{B}_1$ 的 $r+1$ 阶子式 M 也为 0.

(3) 如 $\boldsymbol{A}\to\boldsymbol{B}_1=(\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j+k\alpha_i,\cdots,\alpha_n)$

当 M 不含第 j 列时, M 就是 $\boldsymbol{A}$ 的 $r+1$ 阶子式,其值是 0,若 M 中有第 j 列,这时 M 可以表示成两个行列式之和,其中一个是 $\boldsymbol{A}$ 的 $r+1$ 阶子式,而另一个当第 i 列也在 M 中时,有两列成比例其值为 0,而当第 i 列不在 M 中时,提出公因数 k 后可化为 $\boldsymbol{A}$ 的 $r+1$ 阶子式,所以不论发生哪种情况都有 $M=0$.

这样就证明了 $\boldsymbol{A}$ 经过一次初等列变换得到矩阵 $\boldsymbol{B}_1$ 时, $\mathrm{r}(\boldsymbol{B}_1)\leq\mathrm{r}(\boldsymbol{A})$.进而可得 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})$. ■

推论 3 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵, $\boldsymbol{P},\boldsymbol{Q}$ 分别是 m 阶, n 阶可逆矩阵,则 $\mathrm{r}(\boldsymbol{PAQ})=\mathrm{r}(\boldsymbol{A})$.特别地, $\mathrm{r}(\boldsymbol{PA})=\mathrm{r}(\boldsymbol{A})$, $\mathrm{r}(\boldsymbol{AQ})=\mathrm{r}(\boldsymbol{A})$

证 可逆矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 可写成一些初等矩阵的乘积,设 $\boldsymbol{P}=\boldsymbol{P}_s\cdots\boldsymbol{P}_2\boldsymbol{P}_1$, $\boldsymbol{Q}=\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_t$,其中 $\boldsymbol{P}_i$, $\boldsymbol{Q}_j$ 都是初等矩阵.那么

$$\boldsymbol{PAQ}=\boldsymbol{P}_s\cdots\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_t$$

而用初等矩阵 $\boldsymbol{P}_i$ 左乘 $\boldsymbol{A}$ 相当于对 $\boldsymbol{A}$ 作初等行变换,用初等矩阵 $\boldsymbol{Q}_j$ 右乘 $\boldsymbol{A}$ 相当于对 $\boldsymbol{A}$ 作初等列变换.由于初等变换不改变矩阵的秩,这就有 $\mathrm{r}(\boldsymbol{PAQ})=\mathrm{r}(\boldsymbol{A})$. ■

定理 3.11 不仅告诉我们等价的矩阵有相同的秩,而且提供了

求矩阵秩的具体方法:对矩阵 $\boldsymbol{A}$ 作初等变换化其为阶梯形矩阵 $\boldsymbol{B}$,由 $\boldsymbol{B}$ 的秩而求出 $\boldsymbol{A}$ 的秩.

例 5 求矩阵 $\boldsymbol{A}$ 的秩

$$\boldsymbol{A}=\begin{pmatrix}1&1&2&5\\2&3&0&-1\\4&5&4&9\end{pmatrix}.$$

解 对 $\boldsymbol{A}$ 作初等变换化成阶梯形矩阵 $\boldsymbol{B}$

$$\boldsymbol{A}\to\begin{pmatrix}1&1&2&5\\0&1&-4&-11\\0&1&-4&-11\end{pmatrix}\to\begin{pmatrix}1&1&2&5\\0&1&-4&-11\\0&0&0&0\end{pmatrix}=\boldsymbol{B}$$

显然,$\boldsymbol{B}$ 中存在二阶子式非零而所有三阶子式全为零.即 $\mathrm{r}(\boldsymbol{B})=2$,从而得 $\mathrm{r}(\boldsymbol{A})=2$. ■

例 6 求矩阵 $\boldsymbol{A}$ 的秩,其中 a,b 是参数

$$\boldsymbol{A}=\begin{pmatrix}1&1&1&1\\0&1&-1&b\\2&3&a&4\\3&5&1&7\end{pmatrix}.$$

解 对 $\boldsymbol{A}$ 作初等变换化成阶梯形矩阵

$$\boldsymbol{A}\to\begin{pmatrix}1&1&1&1\\0&1&-1&b\\0&1&a-2&2\\0&2&-2&4\end{pmatrix}\to\begin{pmatrix}1&1&1&1\\0&1&-1&b\\0&0&a-1&2-b\\0&0&0&4-2b\end{pmatrix}$$

当 $a\neq1$ 且 $b\neq2$ 时,$\mathrm{r}(\boldsymbol{A})=4$;当 $a=1$ 且 $b=2$ 时,$\mathrm{r}(\boldsymbol{A})=2$;当 $a=1$, $b\neq2$ 或 $a\neq1$, $b=2$ 时,$\mathrm{r}(\boldsymbol{A})=3$. ■

现在来讨论矩阵的秩与向量组的秩之间的联系.设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵,对 $\boldsymbol{A}$ 分块如下:

$$A=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \cdots\\ \alpha_m\end{pmatrix}=(\beta_1\ \beta_2\cdots\beta_n),$$

其中 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 是 $1\times n$ 矩阵,称为矩阵 A 的行向量组,称行向量组的秩 $r(\alpha_1,\alpha_2,\cdots,\alpha_m)$是矩阵 A 的行秩.类似地,$\beta_1,\beta_2,\cdots,\beta_n$ 是 $m\times 1$ 矩阵,叫做矩阵 A 的列向量组,称列向量组的秩 $\mathrm{r}(\beta_1,\beta_2,\cdots,\beta_n)$是矩阵 A 的列秩.

定理 3.12 矩阵 A 的行秩等于其列秩也等于矩阵 A 的秩.即 $\mathrm{r}(\alpha_1,\alpha_2,\cdots,\alpha_m)=\mathrm{r}(\beta_1,\beta_2,\cdots,\beta_n)=\mathrm{r}(A)$. ■

关于定理 3.12 我们不给出其证明,仅通过一特例解释一下矩阵的三种秩的概念并验证其相等.例如阶梯形矩阵

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

矩阵 A 中有三阶子式

$$\begin{vmatrix}1&2&3\\0&3&2\\0&0&4\end{vmatrix}\neq 0,$$

而 A 的四阶子式只有一个,就是$|A|$,且$|A|=0$,所以,矩阵 A 的秩 $\mathrm{r}(A)=3$.

作为 A 的行向量组

$$\alpha_1=(1,2,1,3),\qquad \alpha_2=(0,3,-6,2),$$
$$\alpha_3=(0,0,0,4),\qquad \alpha_4=(0,0,0,0).$$

显然,$\alpha_1,\alpha_2,\alpha_3$ 是阶梯形向量组,它们线性无关,而 α_4 是零向量,故矩阵 A 的行秩 $\mathrm{r}(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=3$.

作为矩阵 $\boldsymbol{A}$ 的列向量组

$$\beta_1=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ \beta_2=\begin{pmatrix}2\\3\\0\\0\end{pmatrix},\ \beta_3=\begin{pmatrix}1\\-6\\0\\0\end{pmatrix},\ \beta_4=\begin{pmatrix}3\\2\\4\\0\end{pmatrix}.$$

易见 β_1,β_2,β_4 是阶梯形向量组,它们线性无关,而 $\beta_3=5\beta_1-2\beta_2+0\beta_4$,添加进 β_3 后就线性相关,所以矩阵 $\boldsymbol{A}$ 的列秩 $\mathrm{r}(\beta_1,\beta_2,\beta_3,\beta_4)=3$.我们看到 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=\mathrm{r}(\beta_1,\beta_2,\beta_3,\beta_4)$.

矩阵的秩是用行列式来刻划的,它与向量组的秩有其内在的联系,利用行列式的性质及向量组秩的概念、性质可推导出矩阵的秩的一些性质

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

(2) 当 $k\neq 0$ 时,$\mathrm{r}(k\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})$;

(3) $\mathrm{r}(\boldsymbol{A}+\boldsymbol{B})\leq\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$;

(4) $\mathrm{r}(\boldsymbol{AB})\leq\min(\mathrm{r}(\boldsymbol{A}),\mathrm{r}(\boldsymbol{B}))$;

(5) $\mathrm{r}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{B}\end{pmatrix}=\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$.

例 7 $\boldsymbol{A}$ 是 n 阶矩阵,$\boldsymbol{B}$ 是 $n\times m$ 矩阵,$r(\boldsymbol{B})=n$,若 $\boldsymbol{AB}=0$,证明:$\boldsymbol{A}=0$.

证 1 因为 $r(\boldsymbol{B})=n$,故 $\boldsymbol{B}$ 中存在 n 阶子式不为零,设其为 $|\boldsymbol{B}_1|$,那么 $\boldsymbol{B}_1$ 是 n 阶可逆矩阵.从 $\boldsymbol{AB}=0$,知 $\boldsymbol{AB}_1=0$.因此,$\boldsymbol{A}=(\boldsymbol{AB}_1)\boldsymbol{B}_1^{-1}=0\boldsymbol{B}_1^{-1}=0$. ■

证 2 对矩阵 $\boldsymbol{B}$ 按行分块,从行秩来看,由

$$\mathrm{r}(\boldsymbol{B})=r\begin{pmatrix}\alpha_1\\\alpha_2\\\cdots\\\alpha_n\end{pmatrix}=n,$$

得知 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关.对 $\boldsymbol{AB}$ 用分块矩阵乘法,有

$$\boldsymbol{AB}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}=\begin{pmatrix} a_{11}\alpha_1+a_{12}\alpha_2+\cdots+a_{1n}\alpha_n \\ a_{21}\alpha_1+a_{22}\alpha_2+\cdots+a_{2n}\alpha_n \\ \vdots \\ a_{n1}\alpha_1+a_{n2}\alpha_2+\cdots+a_{nn}\alpha_n \end{pmatrix}=0 .$$

按线性无关的定义,对

$$a_{11}\alpha_1+a_{12}\alpha_2+\cdots+a_{1n}\alpha_n=\boldsymbol{0},$$

有 $a_{11}=a_{12}=\cdots=a_{1n}=0$.同理可得

$$a_{21}=a_{22}=\cdots=a_{2n}=0,\cdots,a_{n1}=a_{n2}=\cdots=a_{nn}=0 .$$

由于 $a_{ij}\equiv 0$,所以 $\boldsymbol{A}=0$(证法三请看下节例 6) . ■

至此,已经讨论了向量组的线性相关、秩、矩阵的秩等重要概念及其基本性质,现在就可以运用它们来处理本章开头提出的有关线性方程组解的存在及解的结构的问题 .

3.5 齐次线性方程组

在 3.1 高斯消元法中已经看到,作为齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0, \\ \qquad\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 . \end{cases} \tag{3.8}$$

$x_1=0,x_2=0,\cdots,x_n=0$ 一定是它的解,也就是说齐次线性方程组一定有零解,因而我们关心的是它何时有非零解?如果它有非零解,那它就有无穷多个解,怎样表示这无穷多个解呢?

为了书写简洁明了,对方程组引入矩阵表示,设系数矩阵为 A,未知数矩阵是 $\boldsymbol{x}$,即

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \hline \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix},$$

则方程组(3.8)的矩阵表达形式为:

$$\boldsymbol{Ax}=0$$

定理 3.13　**如果** $\boldsymbol{X}_1, \boldsymbol{X}_2$ **都是齐次线性方程组**(3.8)**的解,则** $\boldsymbol{X}_1, \boldsymbol{X}_2$ **任意的线性组合** $k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2$ **仍是该齐次线性方程组**(3.8)**的解**.

(**注**: $\boldsymbol{A}(k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2)=\boldsymbol{A}(k_1\boldsymbol{X}_1)+\boldsymbol{A}(k_2\boldsymbol{X}_2)=k_1\boldsymbol{AX}_1+k_2\boldsymbol{AX}_2=0$) ■

定理 3.14　齐次线性方程组(3.8)**有非零解的充分必要条件是其系数矩阵的秩** $\mathrm{r}(\boldsymbol{A})<n$.

* **证**　**对方程组** $\boldsymbol{Ax}=0$,**将其系数矩阵按列分块** $\boldsymbol{A}=(\beta_1\ \beta_2\ \cdots\beta_n)$,**其中** $\beta_j=(a_{1j}, a_{2j}, \cdots, a_{mj})^{\mathrm{T}}$, $j=1,2,\cdots,n$ **是** m **维列向量**.**那么**

$$\boldsymbol{Ax}=(\beta_1\ \beta_2\cdots\beta_n)\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}=x_1\beta_1+x_2\beta_2+\cdots+x_n\beta_n=\boldsymbol{0},$$

因此

$$\boldsymbol{Ax}=\boldsymbol{0}\text{ 有非零解}$$

$$\beta_1,\beta_2,\cdots,\beta_n\text{ 线性相关}$$

$$\mathrm{r}(\beta_1,\beta_2,\cdots\beta_n)<n$$

$$\mathrm{r}(\boldsymbol{A})<n\quad(\text{定理 }3.12)$$ ■

现在给出本节最重要的一个结果

定理 3.15　如果齐次线性方程组(3.8)**系数矩阵** A **的秩**

$r(\boldsymbol{A})=r<n$,则方程组有 $n-r$ 个线性无关的解向量,而其它任意一个解都是这 $n-r$ 个线性无关的解向量的线性组合 .

*证　用高斯消元法把系数矩阵 A 化成阶梯形矩阵时,由于初等行变换不改变矩阵的秩,并且矩阵的秩等于其行秩,从 $r(A)=r$ 知阶梯形矩阵中有 r 个非零行,不妨设与(3.8)同解的阶梯形方程组是:

$$\begin{aligned}
c_{11}x_1+c_{12}x_2+\cdots+c_{1r}x_r+c_{1\,r+1}x_{r+1}+\cdots+c_{1n}x_n&=0,\\
c_{22}x_2+\cdots+c_{2r}x_r+c_{2\,r+1}x_{r+1}+\cdots+c_{2n}x_n&=0,\\
\hline
c_{rr}x_r+c_{r\,r+1}x_{r+1}+\cdots+c_{rn}x_n&=0,
\end{aligned}\tag{3.9}$$

其中 $c_{ii}\neq0,\ i=1,2,\cdots,r$.

把上述同解方程组(3.9)改写成与其同解的

$$\begin{aligned}
c_{11}x_1+c_{12}x_2+\cdots+c_{1r}x_r&=-c_{1\,r+1}x_{r+1}-\cdots-c_{1n}x_n\\
c_{22}x_2+\cdots+c_{2r}x_r&=-c_{2\,r+1}x_{r+1}-\cdots-c_{2n}x_n\\
\hline
c_{rr}x_r&=-c_{r\,r+1}x_{r+1}-\cdots-c_{rn}x_n
\end{aligned}\tag{3.10}$$

容易看出:任意给定 $x_{r+1},\cdots,x_n$ 一组数值,代入到方程组(3.10)中可以求出齐次方程组(3.8)的一个解 .

为了构造出 $n-r$ 个线性无关的解,可按下述方式给未知量 $x_{r+1},x_{r+2},\cdots,x_n$ 赋值,例如

令 $x_{r+1}=1,x_{x+2}=\cdots=x_n=0$,可求出(3.8)的一个解

$$\boldsymbol{\eta}_1=(x_1^{(1)},x_2^{(1)},\cdots,x_r^{(1)},1,0,0,\cdots,0)^{\mathrm{T}}$$

再令 $x_{r+1}=0,\ x_{r+2}=1,\ x_{r+3}=\cdots=x_n=0$,可求出(3.8)的另一个解

$$\boldsymbol{\eta}_2=(x_1^{(2)},x_2^{(2)},\cdots,x_r^{(2)},0,1,0,\cdots,0)^{\mathrm{T}}$$

类似地还可求出解

$$\eta_3 = (x_1^{(3)}, x_2^{(3)}, \cdots, x_r^{(3)}, 0, 0, 1, \cdots, 0)^{\mathrm{T}}$$

$$\eta_{n-r} = (x_1^{(n-r)}, x_2^{(n-r)}, \cdots, x_r^{(n-r)}, 0, 0, 0, \cdots, 1)^{\mathrm{T}}$$

根据 3.2 例 13,这样求出的 $n-r$ 个解向量

$$\eta_1, \eta_2, \cdots, \eta_{n-r},$$

是线性无关的

设 $\boldsymbol{X}_0 = (t_1, t_2, \cdots, t_n)^{\mathrm{T}}$ 是齐次方程组(3.8)的任一解,我们来证 $\boldsymbol{X}_0$ 可由 $\eta_1, \eta_2, \cdots, \eta_{n-r}$ 线性表出.由于 $\boldsymbol{X}_0, \eta_1, \eta_2, \cdots, \eta_{n-r}$ 都是齐次方程组(3.8)的解,那么这些解的线性组合

$$\boldsymbol{X}_0 - t_{r+1}\eta_1 - t_{r+2}\eta_2 - \cdots - t_n\eta_{n-r}$$

仍是齐次方程组的解(定理 3.13),但这个解的后 $n-r$ 个分量全为零,即这个解向量形如

$$(u_1, u_2, \cdots, u_r, 0, 0, \cdots, 0)^{\mathrm{T}}$$

把它代入到同解的齐次方程组(3.10),即有

$$\begin{aligned} c_{11}u_1 + c_{12}u_2 + \cdots + c_{1r}u_r &= 0, \\ c_{22}u_2 + \cdots + c_{2r}u_r &= 0, \\ &\cdots\cdots \\ c_{rr}u_r &= 0, \end{aligned}$$

由此可得: $u_r = 0, u_{r-1} = 0, \cdots, u_1 = 0$,因此

$$\boldsymbol{X}_0 - t_{r+1}\eta_1 - t_{r+2}\eta_2 - \cdots - t_n\eta_{n-r} = 0,$$

这表明,齐次方程组的任一解 $\boldsymbol{X}_0$ 都是这 $n-r$ 个线性无关解向量 $\eta_1, \eta_2, \cdots, \eta_{n-r}$ 的线性组合. ■

习惯上,我们称可以任意取值的这 $n-r$ 个变量 $x_{r+1}, x_{r+2}, \cdots, x_n$ 是自由变量,而线性无关的这 $n-r$ 个解向量 $\eta_1, \eta_2, \cdots, \eta_{n-r}$ 是齐次线性方程组的一个基础解系.也就是说,解向量组的一个极大线性无关组就是一个基础解系,它由 $n-r(A)$ 个线性无

关的解组成.

例 1 求齐次方程组

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

的基础解系及一般解.

解 用高斯消元法化系数矩阵为阶梯形

$$\begin{pmatrix} 1 & 2 & 4 & -3 \\ 3 & 5 & 6 & -4 \\ 4 & 5 & -2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 4 & -3 \\ 0 & -1 & -6 & 5 \\ 0 & -3 & -18 & 15 \end{pmatrix}$$

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

由于阶梯形矩阵有两个非零行,知 $r(\boldsymbol{A}) = 2$,那么基础解系由 $n - r(A) = 4 - 2 = 2$ 个解向量组成.从阶梯形矩阵的形状知,可选取 x_3, x_4 为自由变量,从而求出基础解系

$$\boldsymbol{\eta}_1 = (8, -6, 1, 0)^{\mathrm{T}},$$

$$\boldsymbol{\eta}_2 = (-7, 5, 0, 1)^{\mathrm{T}},$$

因此,方程组的一般解是: $k_1\boldsymbol{\eta}_1 + k_2\boldsymbol{\eta}_2$, k_1, k_2 是任意常数. ■

例 2 解齐次方程组

$$\begin{cases} x_1 + 3x_2 + 2x_3 + 3x_4 = 0, \\ 2x_1 + 6x_2 + 5x_3 + 9x_4 = 0, \\ 3x_1 + 9x_2 + 4x_3 + 3x_4 = 0. \end{cases}$$

解 用高斯消元化系数矩阵为阶梯形

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

由于阶梯形矩阵有两个非零行,知 $\mathrm{r}(\boldsymbol{A})=2$,那么基础解系应由 $n-\mathrm{r}(\boldsymbol{A})=4-2=2$ 个向量组成,从阶梯形矩阵看出可选取 x_2, x_4 为自由变量(能否选 x_3, x_4 为自由变量,为什么?能否选 x_1, x_4 为自由变量,为什么?选取自由变量的原则究竟是什么?)这样求出的基础解系是

$$\eta_1=(-3,1,0,0)^{\mathrm{T}},$$

$$\eta_2=(3,0,-3,1)^{\mathrm{T}},$$

得到齐次方程组的一般解是:$k_1\eta_1+k_2\eta_2$, k_1, k_2 是任意常数.■

例 3 解方程组

$$\begin{cases} x_1-2x_2+x_3+x_4-x_5=0,\\ 2x_1+x_2-x_3-x_4+x_5=0,\\ x_1+7x_2-5x_3-5x_4+5x_5=0,\\ 3x_1-x_2-2x_3+x_4-x_5=0. \end{cases}$$

解 用高斯消元法,有

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

$$\to \begin{pmatrix} 1 & -2 & 1 & 1 & -1\\ 0 & 5 & -3 & -3 & 3\\ 0 & -1 & 0 & 0 & 0\\ 0 & 0 & -2 & 1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 1 & 1 & -1\\ 0 & -1 & 0 & 0 & 0\\ 0 & 5 & -3 & -3 & 3\\ 0 & 0 & -2 & 1 & -1 \end{pmatrix}$$

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

$$\to\begin{pmatrix}1 & -2 & 1 & 1 & -1\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 1 & -1\\ 0 & 0 & 0 & 3 & -3\end{pmatrix}.$$

从阶梯形矩阵的四行均非零,得 $r(\boldsymbol{A})=4$,因此,基础解系由 $n-r(\boldsymbol{A})=5-4=1$ 个解向量组成.选取 x_5 作为自由变量,得到基础解系

$$\eta=(0,0,0,1,1)^{T}$$

所以方程组的一般解是: $k\eta$, k 为任意常数. ■

由于齐次方程组的解与矩阵的秩之间有内在的联系,所以一些有关秩的问题也可利用齐次方程组解的理论来处理,请看下面几个典型例题.

例 4 已知 $\boldsymbol{A}$ 是 $m\times n$ 矩阵, $\boldsymbol{B}$ 是 $n\times s$ 矩阵,若 $\boldsymbol{AB}=0$,证明: $r(\boldsymbol{A})+r(\boldsymbol{B})\leq n$.

证 对矩阵 B 按列分块,有

$$\boldsymbol{AB}=A(\boldsymbol{B}_1\ \boldsymbol{B}_2\cdots\boldsymbol{B}_s)=(\boldsymbol{AB}_1,\boldsymbol{AB}_2,\cdots,\boldsymbol{AB}_s)=\boldsymbol{0},$$

由于 $\boldsymbol{AB}_i=0\ (i=1,2,\cdots,s)$,知矩阵 $\boldsymbol{B}$ 的每一列 $\boldsymbol{B}_i$ 都是齐次方程组 $\boldsymbol{AX}=0$ 的解,那么 $\boldsymbol{B}$ 中列向量组的秩必小于等于 $\boldsymbol{AX}=0$ 的解向量组的秩.即有

$$r(\boldsymbol{B})=r(\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_s)\leq n-r(\boldsymbol{A})$$

所以 $r(\boldsymbol{A})+r(\boldsymbol{B})\leq n$. ■

例 5 设 $\boldsymbol{A}$ 是 $m\times s$ 矩阵, $\boldsymbol{B}$ 是 $s\times n$ 矩阵,证明 $r(\boldsymbol{AB})\leq r(\boldsymbol{B})$.

证 设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{T}$ 是未知数,构造两个齐次线性方程组

$$\boldsymbol{ABx}=\boldsymbol{0}, \tag{Ⅰ}$$

及

$$\boldsymbol{Bx}=0, \tag{Ⅱ}$$

显然,如果 $\boldsymbol{X}_0$ 是方程组(Ⅱ)的解,则 $\boldsymbol{X}_0$ 必是方程组(Ⅰ)的解,因

此

$$(\text{Ⅱ})\text{ 的解向量} \quad\quad (\text{Ⅰ})\text{ 的解向量},$$

从而有

$$n - \mathrm{r}(\boldsymbol{B}) = \mathrm{r}\{(\text{Ⅱ})\text{ 的解向量}\}$$

$$\leq \mathrm{r}\{(\text{Ⅰ})\text{ 的解向量}\} = n - \mathrm{r}(\boldsymbol{AB})$$

所以, $\mathrm{r}(\boldsymbol{AB}) \leq \mathrm{r}(\boldsymbol{B})$. ■

有兴趣的读者不妨利用本题的结果来证明 $\mathrm{r}(\boldsymbol{AB}) \leq \mathrm{r}(\boldsymbol{A})$, 这样也就证明了上一节给出的结果 $\mathrm{r}(\boldsymbol{AB}) \leq \min(\mathrm{r}(\boldsymbol{A}), \mathrm{r}(\boldsymbol{B}))$.(提示:注意 $\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\boldsymbol{A}^{\mathrm{T}})$)

*例 6　已知 $\boldsymbol{A}$ 是 $m \times n$ 矩阵, $\boldsymbol{B}$ 是 $n \times p$ 矩阵, 如果 $\boldsymbol{AB} = 0$ 且 $\mathrm{r}(\boldsymbol{B}) = n$, 证明: $\boldsymbol{A} = 0$.

证　构造齐次线性方程组 $\boldsymbol{Ax} = 0$, 其基础解系由 $n - \mathrm{r}(\boldsymbol{A})$ 个线性无关的解向量所构成 .现 $\boldsymbol{AB} = 0$, 知 $\boldsymbol{B}$ 的每一列都是 $\boldsymbol{Ax} = 0$ 的解, 而 $r(\boldsymbol{B}) = n$ 则说明 $\boldsymbol{Ax} = 0$ 的解中至少有 n 个线性无关的解向量 .因此有 $n - \mathrm{r}(\boldsymbol{A}) \geq n$ 这样就有 $\mathrm{r}(\boldsymbol{A}) \leq 0$, 但按照秩的定义, 对任何矩阵 $\boldsymbol{A}$ 都有 $\mathrm{r}(\boldsymbol{A}) \geq 0$.

所以 $\mathrm{r}(\boldsymbol{A}) = 0$, 即 $\boldsymbol{A} = 0$. ■

*例 7　$\boldsymbol{A}$ 是 n 阶矩阵, $\boldsymbol{A}^2 = \boldsymbol{I}$, 证明: $\mathrm{r}(\boldsymbol{A} - \boldsymbol{I}) + r(\boldsymbol{A} + \boldsymbol{I}) = n$.

证　从 $\boldsymbol{A}^2 = I$ 得 $(\boldsymbol{A} - \boldsymbol{I})(\boldsymbol{A} + \boldsymbol{I}) = 0$, 据例 4 有

$$\mathrm{r}(\boldsymbol{A} - \boldsymbol{I}) + \mathrm{r}(\boldsymbol{A} + \boldsymbol{I}) \leq n,$$

又因 $\mathrm{r}(\boldsymbol{A} - \boldsymbol{I}) = \mathrm{r}(\boldsymbol{I} - \boldsymbol{A})$,

$$\begin{aligned}\mathrm{r}(\boldsymbol{A} - \boldsymbol{I}) + \mathrm{r}(\boldsymbol{A} + \boldsymbol{I}) &= \mathrm{r}(\boldsymbol{I} - \boldsymbol{A}) + \mathrm{r}(\boldsymbol{A} + \boldsymbol{I}) \\ &\geq \mathrm{r}\ (\boldsymbol{I} - \boldsymbol{A}) + (\boldsymbol{A} + \boldsymbol{I}) \\ &= \mathrm{r}(2\boldsymbol{I}) = n\end{aligned}$$

所以, $\mathrm{r}(\boldsymbol{A} - \boldsymbol{I}) + \mathrm{r}(\boldsymbol{A} + \boldsymbol{I}) = n$. ■

3.6 非齐次线性方程组

有了前面 3.2—3.5 的种种准备,现在可以完满地回答在本章起始提出的问题:一个方程组何时有解?如果有解,那么有多少组解?怎样才能求出这所有的解?

设有 n 个未知数 m 个方程所构成的非齐次线性方程组,如(3.1)式

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1,\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2,\\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m .
\end{aligned}
$$

类似于 3.5 的表述,方程组(3.1)可用矩阵表示为

$$\boldsymbol{Ax} = \beta$$

其中 $\boldsymbol{A}$ 是系数矩阵, $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 是未知数, $\beta = (b_1, b_2, \cdots, b_m)^{\mathrm{T}}$ 是常数项,对系数矩阵 $\boldsymbol{A}$ 按列分块又有方程组(3.1)的向量表达形式

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta .$$

其中 $\alpha_j = (a_{1j}, a_{2j}, \cdots, a_{mj})^{\mathrm{T}}$ 是 A 中的第 j 列.

首先,研究非齐次方程组(3.1)有解的充分必要条件,即解的存在定理:

定理 3.16 非齐次线性方程组(3.1)有解的充分必要条件是系数矩阵 $\boldsymbol{A}$ 的秩等于增广矩阵 $\boldsymbol{A}$ 的秩.即 $\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\boldsymbol{A})$.

证 非齐次线性方程组(3.1)有解

β 可由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表出

向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 $\alpha_1, \alpha_2, \cdots, \alpha_n, \beta$ 等价

$$\mathrm{r}(\alpha_1, \alpha_2, \cdots, \alpha_n) = \mathrm{r}(\alpha_1, \alpha_2, \cdots, \alpha_n, \beta)$$

$$\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\boldsymbol{A}) \text{ .}$$ ■

例 1　已知

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0$$

证明: 方程组

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 &= a_{13} \\ a_{21} x_1 + a_{22} x_2 &= a_{23} \\ a_{31} x_1 + a_{32} x_2 &= a_{33} \end{aligned}$$

无解 .

证　方程组的系数矩阵

$$\boldsymbol{A} = \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

是 3×2 **的矩阵,它没有三阶子式,故其秩** $\mathrm{r}(\boldsymbol{A})\leq2$,**而增广矩阵**

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

据已知条件 $|\boldsymbol{A}|\neq0$,**从而知** $\mathrm{r}(\boldsymbol{A}) = 3$.

所以, $\mathrm{r}(\boldsymbol{A})\neq\mathrm{r}(\boldsymbol{A})$,**即方程组无解 .** ■

其次,我们来讨论当非齐次方程组(3.1)有解时,它共有多少解 ? 为此先来看方程组(3.1)与其相应的齐次方程组(3.8)的解之间的联系 .

(1) **设** $\boldsymbol{X}_1, \boldsymbol{X}_2$ **是** $\boldsymbol{Ax} = \beta$ **的任意两个解,则** $\boldsymbol{X}_1 - \boldsymbol{X}_2$ **是相应的**

齐次方程组 $\boldsymbol{Ax}=0$ 的解.(练习题)

(2) 如 X^* 是 $\boldsymbol{Ax}=\beta$ 的某个解,$\boldsymbol{X}_0$ 是 $A\boldsymbol{x}=0$ 的任一解,则 $\boldsymbol{X}^*+\boldsymbol{X}_0$ 仍是非齐次方程组 $\boldsymbol{Ax}=\beta$ 的解.(练习题)

根据方程组解向量间的这些联系,我们得到:

$$\{\boldsymbol{Ax}=\beta\text{的所有解}\}$$
$$=\{\boldsymbol{Ax}=\mathbf{0}\text{的所有解}\}+(\boldsymbol{Ax}=\beta\text{的某一个解})$$

这样,当 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})=n$ 时,即

$$\mathrm{r}(\alpha_1,\alpha_2,\cdots,\alpha_n)=\mathrm{r}(\alpha_1,\alpha_2,\cdots,\alpha_n,\beta)=n$$

则 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关,而 $\alpha_1,\alpha_2,\cdots,\alpha_n,\beta$ 线性相关,所以 β 可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出且表示方法唯一.也就是方程组 $\boldsymbol{Ax}=\beta$ 有唯一解.

当 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})<n$ 时,由于相应的齐次方程组 $\boldsymbol{Ax}=0$ 有无穷多个解,因而非齐次方程组 $\boldsymbol{Ax}=\beta$ 有无穷多解.

总之,对非齐次线性方程组(3.1)

当 $\mathrm{r}(\boldsymbol{A})\neq\mathrm{r}(\boldsymbol{A})$ 时,方程组无解;

当 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})=r$ 时,方程组有解,且

$r=n$ 时,有唯一解

$r<n$ 时,有无穷多解.

根据非齐次线性方程组(3.1)及其相应的齐次方程组(3.8)在其解之间的联系,为了求出(3.1)的所有解,我们只需设法找出它的某一个解(通常称其为特解)以及相应齐次方程组(3.8)的所有解.而后者在上一节已看到,这就是求基础解系.于是关于非齐次方程组(3.1)的解,有下述重要结果.

定理 3.17 如果非齐次线性方程组 $\boldsymbol{Ax}=\beta$ 中,A 是 $m\times n$ 矩阵,$r(\boldsymbol{A})=r(\boldsymbol{A})=r$,$\xi$ 是其一个特解,而 $\eta_1,\eta_2,\cdots,\eta_{n-r}$ 是相应齐次方程组 $\boldsymbol{Ax}=0$ 的基础解系.则方程组 $\boldsymbol{Ax}=\beta$ 的通解可表示为

$$\xi+k_1\eta_1+k_2\eta_2+\cdots+k_{n-r}\eta_{n-r}$$

其中 $k_1, k_2, \cdots, k_{n-r}$是任意常数 . ■

下面通过例题来介绍特解、通解的求法 .

例 2 解方程组

$$
\begin{cases}
x_1 + 2x_2 + 3x_3 + x_4 = 3, \\
3x_1 + 7x_2 + 7x_3 + 2x_4 = 12, \\
x_1 + 4x_2 + 5x_3 + 2x_4 = 6 .
\end{cases}
$$

解 对增广矩阵进行初等行变换化其为阶梯形

$$
\begin{pmatrix} 1 & 2 & 3 & 1 & 3 \\ 3 & 7 & 7 & 2 & 12 \\ 1 & 4 & 5 & 2 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 & 1 & 3 \\ 0 & 1 & -2 & -1 & 3 \\ 0 & 2 & 2 & 1 & 3 \end{pmatrix}
$$

$$
\rightarrow \begin{pmatrix} 1 & 2 & 3 & 1 & 3 \\ 0 & 1 & -2 & -1 & 3 \\ 0 & 0 & 6 & 3 & -3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 & 1 & 3 \\ 0 & 1 & -2 & -1 & 3 \\ 0 & 0 & 2 & 1 & -1 \end{pmatrix}
$$

由 $\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\boldsymbol{A}) = 3$,知方程组有解 .

要找出一组特解,可令 $x_3 = 0$,由下往上反向代入,可解得 $x_4 = -1$, $x_2 = 2$, $x_1 = 0$.于是有特解: $\boldsymbol{\xi} = (0, 2, 0, -1)^{\mathrm{T}}$

欲求相应齐次方程组的解,对上述消元所得系数矩阵

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

知其基础解系由 $n - \mathrm{r}(\boldsymbol{A}) = 4 - 3 = 1$ 个向量构成,选择 x_3 作为自由变量,令 $x_3 = 1$,经反向代入,可求出基础解系是: $\boldsymbol{\eta} = (-1, 0, 1, -2)^{\mathrm{T}}$ 所以,方程组的通解是

$$
\boldsymbol{\xi} + k\boldsymbol{\eta}
$$

其中 k 为任意常数 . ■

例 3 解方程组

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 2, \\ 2x_1 + 3x_2 + x_3 + x_4 - 3x_5 = 0, \\ x_1 + 2x_3 + 2x_4 + 6x_5 = 6, \\ 4x_1 + 5x_2 + 3x_3 + 3x_4 - x_5 = 4. \end{cases}$$

解　对增广矩阵作初等行变换化其为阶梯形

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

$$\rightarrow \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & -1 & -1 & -5 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

由于 $r(\boldsymbol{A}) = r(\bar{\boldsymbol{A}}) = 2$, **方程组有解**.

要找一组特解, **可令** $x_3 = x_4 = x_5 = 0$, **反向代入解出** $x_2 = -4$, $x_1 = 6$. **因而有特解**: $\boldsymbol{\xi} = (6, -4, 0, 0, 0)^{\mathrm{T}}$.

针对化为阶梯形的系数矩阵

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & -1 & -5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

其基础解系由 $n - r(\boldsymbol{A}) = 5 - 2 = 3$ **个解向量组成**. **可取** x_3, x_4, x_5 **为自由变量**, **求出基础解系是**

$$\boldsymbol{\eta}_1 = (-2, 1, 1, 0, 0)^{\mathrm{T}}, \quad \boldsymbol{\eta}_2 = (-2, 1, 0, 1, 0)^{\mathrm{T}},$$

$$\boldsymbol{\eta}_3 = (-6, 5, 0, 0, 1)^{\mathrm{T}}.$$

方程组通解是:

$$\boldsymbol{\xi} + k_1\boldsymbol{\eta}_1 + k_2\boldsymbol{\eta}_2 + k_3\boldsymbol{\eta}_3,\quad k_1, k_2, k_3 \text{ 是任意常数}.$$ ■

注 对 $\boldsymbol{Ax}=0$ 选取自由变量时,应保证同解方程组(3.1)的左端系数矩阵是 r 阶可逆矩阵.如只有一个自由变量,求基础解系时自由变量的赋值不能是 0.

$\boldsymbol{Ax}=\boldsymbol{b}$ 特解的求法是多种多样的,为了简捷,通常给自由变量赋值为 0.

对 $\boldsymbol{Ax}=\boldsymbol{b}$ 求相应 $\boldsymbol{Ax}=0$ 的基础解系时,齐次方程组的阶梯形矩阵可以不另写,而是借用求特解时的阶梯矩阵的前 n 列,但反向代入时千万不要把常数项混同进去运算.

例 4 已知线性方程组

$$\begin{cases} x_1 + x_2 + ax_3 = 4, \\ -x_1 + ax_2 + x_3 = a^2, \\ x_1 - x_2 + 2x_3 = -4, \end{cases}$$

有无穷多解,求 a 的取值并求这些解.

解 对增广矩阵作初等行变换化为阶梯形

$$\begin{pmatrix} 1 & 1 & a & 4 \\ -1 & a & 1 & a^2 \\ 1 & -1 & 2 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 2 & -4 \\ 1 & 1 & a & 4 \\ -1 & a & 1 & a^2 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & -1 & 2 & -4 \\ 0 & 2 & a-2 & 8 \\ 0 & a-1 & 3 & a^2-4 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & -1 & 2 & -4 \\ 0 & 2 & a-2 & 8 \\ 0 & 0 & \dfrac{(a+1)(4-a)}{2} & a(a-4) \end{pmatrix}.$$

由于方程组有无穷多解,故 $\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\bar{\boldsymbol{A}}) \leqslant 2$.易见仅当 $a=4$ 时,$\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\bar{\boldsymbol{A}}) = 2$,方程组才有无穷多解.这时增广矩阵可化为

$$\begin{matrix} 1 & -1 & 2 & -4 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{matrix}$$

令 $x_3 = 0$,得非齐次方程组的特解 $\xi = (0,4,0)^T$.

令 $x_3 = 1$,得相应齐次方程组基础解系 $\eta = (-3, -1,1)^T$.所以,方程组的所有解是 $\xi + k\eta$, k 是任意常数 . ■

*例 5　四元方程组 $\boldsymbol{Ax} = \beta$ 的三个解是:

$$\boldsymbol{X}_1 = (1,2,3,4)^T, \quad \boldsymbol{X}_2 = (2,3,1,4)^T, \quad \boldsymbol{X}_3 = (1,1,1,2)^T$$

又知 $r(\boldsymbol{A}) = 2$,求方程组的通解 .

解　对相应的齐次方程组 $\boldsymbol{Ax} = 0$,由 $r(\boldsymbol{A}) = 2$ 且未知数个数 $n = 4$,我们知其基础解系由两个线性无关的解向量组成 .根据方程组解之间的联系,知

$$\eta_1 = \boldsymbol{X}_2 - \boldsymbol{X}_1 = (1,1, -2,0)^T$$

$$\eta_2 = \boldsymbol{X}_2 - \boldsymbol{X}_3 = (1,2,0,2)^T$$

是齐次方程组 $\boldsymbol{Ax} = 0$ 的解,又因 η_1, η_2 线性无关,它们就是 $\boldsymbol{Ax} = 0$ 的基础解系 .从而

$$\boldsymbol{X}_1 + k_1 \eta_1 + k_2 \eta_2 \quad (k_1, k_2 \text{ 是任意常数})$$

是方程组 $\boldsymbol{Ax} = \beta$ 的通解 . ■

*例 6　已知两个线性方程组

$$\begin{cases} a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n = b_1, \\ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n = b_2, \\ \text{———————————} \\ a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n = b_m . \end{cases} \qquad (\text{I})$$

$$
\begin{cases}
a_{11}x_1 + a_{21}x_2 + \cdots + a_{m1}x_m = 0, \\
a_{12}x_1 + a_{22}x_2 + \cdots + a_{m2}x_m = 0, \\
\text{—————————————} \\
a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{mn}x_m = 0, \\
b_1x_1 + b_2x_2 + \cdots + b_mx_m = 1 .
\end{cases}
\qquad (\text{Ⅱ})
$$

证明: 方程组(Ⅰ)有解的充分必要条件是方程组(Ⅱ)无解 .

证 记方程组(Ⅰ)的系数矩阵是 $\boldsymbol{A}_1$, 增广矩阵是 $\boldsymbol{A}_1$, 方程组(Ⅱ)的系数矩阵是 $\boldsymbol{A}_2$, 增广矩阵是 $\boldsymbol{A}_2$.那么由题意有

$$\boldsymbol{A}_2 = \boldsymbol{A}_1^{\mathrm{T}} \quad \text{故} \quad \mathrm{r}(\boldsymbol{A}_2) = \mathrm{r}(\boldsymbol{A}_1^{\mathrm{T}}) = \mathrm{r}(\boldsymbol{A}_1),$$

又 $\boldsymbol{A}_2 = \begin{pmatrix} \boldsymbol{A}_1^{\mathrm{T}} & 0 \\ \boldsymbol{\beta}^{\mathrm{T}} & 1 \end{pmatrix}$, 其中 $\boldsymbol{\beta} = (b_1, b_2, \cdots, b_m)^{\mathrm{T}}$.从行向量的角度来看, $\boldsymbol{A}_2$ 的前 n 个行向量不可能表示其第 $n+1$ 个行向量 .因此

$$\mathrm{r}(\boldsymbol{A}_2) = \mathrm{r}(\boldsymbol{A}_1^{\mathrm{T}} \quad 0) + 1 = \mathrm{r}(\boldsymbol{A}_1^{\mathrm{T}}) + 1 = \mathrm{r}(\boldsymbol{A}_1) + 1 .$$

如果方程组(Ⅰ)有解, 则 $\mathrm{r}(\boldsymbol{A}_1) = \mathrm{r}(\boldsymbol{A}_1) = r$

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

$\mathrm{r}(\boldsymbol{A}_2) = \mathrm{r}(\boldsymbol{A}_1) + 1 = r + 1$

所以, $\mathrm{r}(\boldsymbol{A}_2) \neq \mathrm{r}(\boldsymbol{A}_2)$.即方程组(Ⅱ)无解 .必要性成立 .

关于充分性, 用反证法 .如(Ⅰ)无解, 设 $\mathrm{r}(\boldsymbol{A}_1) = r$, 则 $\mathrm{r}(\boldsymbol{A}_1) = r + 1$.这时

$$\mathrm{r}(\boldsymbol{A}_2) = \mathrm{r}(\boldsymbol{A}_1) = r + 1$$

$$\mathrm{r}(\boldsymbol{A}_2) = \mathrm{r}(\boldsymbol{A}_1) + 1 = r + 1$$

得到 $\mathrm{r}(\boldsymbol{A}_2) = \mathrm{r}(\boldsymbol{A}_2)$ 与已知条件方程组(Ⅱ)无解相矛盾 .所以, (Ⅰ)有解, 充分性成立 . ■

***例 7** 已知 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性相关, 证明对任何 $\boldsymbol{\beta}$ 都存在不全为零的 k_1, k_2, k_3 使得 $k_1\boldsymbol{\beta} + \boldsymbol{\alpha}_1, k_2\boldsymbol{\beta} + \boldsymbol{\alpha}_2, k_3\boldsymbol{\beta} + \boldsymbol{\alpha}_3$ 线性相关 .

证 因为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性相关, 故存在不全为零的 l_1, l_2, l_3 使 $l_1\boldsymbol{\alpha}_1 + l_2\boldsymbol{\alpha}_2 + l_3\boldsymbol{\alpha}_3 = 0$.对任意的 $\boldsymbol{\beta}$, 考查

$$l_1(x_1\beta+\alpha_1)+l_2(x_2\beta+\alpha_2)+l_3(x_3\beta+\alpha_3)=0$$

只需$(l_1x_1+l_2x_2+l_3x_3)\beta=0$ **总成立**

由于 l_1,l_2,l_3 **不全为零,方程** $l_1x_1+l_2x_2+l_3x_3=0$ **必有非零解** $(k_1,k_2,k_3)^{\mathrm{T}}$ **.因此,对任何** β **都存在不全为零的** k_1,k_2,k_3 **使** $k_1\beta+\alpha_1,k_2\beta+\alpha_2,k_3\beta+\alpha_3$ **线性相关** . ■

习 题 3

3.1 用高斯消元法解方程组

(1) $\begin{cases}2x_1-x_2+3x_3=3\\4x_1-x_2+x_3=3\\3x_1+x_2-5x_3=0\end{cases}$ (2) $\begin{cases}x_1-2x_2+x_3+x_4=1\\x_1-2x_2+x_3-x_4=-1\\x_1-2x_2+x_3+5x_4=5\end{cases}$

(3) $\begin{cases}2x_1+x_2-x_3+x_4=1\\3x_1-2x_2+2x_3-3x_4=2\\5x_1+x_2-x_3+2x_4=-1\\2x_1-x_2+x_3-3x_4=4\end{cases}$

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

3.2 **已知** $\alpha_1=(4,1,3),\alpha_2=(1,2,-1),\alpha_3=(7,9,2)$**求** $\alpha_1+2\alpha_2-\alpha_3$.

3.3 已知 $\alpha_1=(2,5,1,3),\alpha_2=(10,1,5,10),\alpha_3=(4,1,-1,1)$**由** $3(\alpha_1-\beta)+2(\alpha_2+\beta)=5(\alpha_3+\beta)$ **求** β .

3.4 已知 β **可由** $\alpha_1,\alpha_2,\alpha_3$ **线性表出,试求参数** t **的取值**

(1) $\alpha_1=(2,3,5),\alpha_2=(3,7,8),\alpha_3=(1,-6,1),\beta=(7,-2,t)$;

(2) $\alpha_1=(3,2,5),\alpha_2=(2,4,7),\alpha_3=(5,6,t),\beta=(1,3,$

5);

(3) $\alpha_1=(1,1,2)$, $\alpha_2=(2,t,4)$, $\alpha_3=(t,3,6)$, $\beta=(-1,5,5t)$

3.5 判断下列向量组是否线性相关,其中 α_3 能否用 α_1,α_2 线性表出.

(1) $\alpha_1=(1,3,6)$, $\alpha_2=(0,2,5)$, $\alpha_3=(1,-1,-4)$;

(2) $\alpha_1=(1,0,2)$, $\alpha_2=(2,0,4)$, $\alpha_3=(3,0,5)$;

(3) $\alpha_1=(1,3,4,-2)$, $\alpha_2=(2,1,3,-1)$,

$\alpha_3=(3,-1,2,0)$;

(4) $\alpha_1=(3,1,1,4)$, $\alpha_2=(2,2,4,3)$, $\alpha_3=(1,4,10,1)$.

3.6 判断下列命题是否正确.

(1) 对 $\alpha_1,\alpha_2,\alpha_3$,如其中任意两个向量都线性无关,则 $\alpha_1,\alpha_2,\alpha_3$ 线性无关;

(2) 如 α_1,α_2 线性相关,β_1,β_2 线性相关,则 $\alpha_1+\beta_1,\alpha_2+\beta_2$ 线性相关;

(3) 如 $\boldsymbol{A}$ 是 n 阶矩阵且 $|\boldsymbol{A}|=0$,则 $\boldsymbol{A}$ 的行向量中至少有一个是其余各行向量的线性组合;

(4) $\boldsymbol{A}$ 是 $m\times n$ 矩阵,齐次线性方程组 $\boldsymbol{Ax}=0$ 只有零解的充要条件是 $\boldsymbol{A}$ 的列向量线性无关.

3.7 如果 $\alpha_1,\alpha_2,\alpha_3$ 线性相关,试问 α_3 是否可由 α_1,α_2 线性表出?又若 α_3 不能由 α_1,α_2 线性表出,是否必有 $\alpha_1,\alpha_2,\alpha_3$ 线性无关?

3.8 已知 $\alpha_1,\alpha_2,\alpha_3$ 线性相关,$\alpha_2,\alpha_3,\alpha_4$ 线性无关,证明:α_1 可由 α_2,α_3 线性表出,而 α_4 不能由 $\alpha_1,\alpha_2,\alpha_3$ 线性表出.

3.9 已知 $\beta_1=\alpha_1+\alpha_2$, $\beta_2=\alpha_2+\alpha_3$, $\beta_3=\alpha_3+\alpha_4$, $\beta_4=\alpha_4+\alpha_1$,证明:向量组 $\beta_1,\beta_2,\beta_3,\beta_4$ 线性相关.

3.10 已知 $\alpha_1,\alpha_2,\alpha_3$ 线性无关,$\alpha_4=x_1\alpha_1+x_2\alpha_2+x_3\alpha_3$,其中 x_1,x_2,x_3 全不为零,证明:$\alpha_2,\alpha_3,\alpha_4$ 线性无关.

3.11　求下列向量组的秩，并求它的一个极大线性无关组

(1) $\alpha_1=(0,0)$，$\alpha_2=(0,3)$，$\alpha_3=(0,-5)$；

(2) $\alpha_1=(2,4,2)$，$\alpha_2=(-1,-2,-1)$，$\alpha_3=(3,5,2)$；

(3) $\alpha_1=(1,0,4,-1)$，$\alpha_2=(2,-1,5,-6)$，

$\alpha_3=(2,1,11,2)$.

3.12　求下列矩阵的秩

(1) $\begin{pmatrix}1&2&1&3\\2&4&0&5\\3&1&2&7\end{pmatrix}$，　(2) $\begin{pmatrix}1&2&3&4\\2&3&4&5\\3&4&5&6\\4&5&6&7\end{pmatrix}$，

(3) $\begin{pmatrix}2&0&2&0&2\\0&1&0&-1&0\\2&1&0&2&1\\0&-1&0&1&0\end{pmatrix}$，　(4) $\begin{pmatrix}1&2&3&-4\\2&3&-4&1\\2&-5&8&-3\\5&26&-9&-12\\3&-4&1&2\end{pmatrix}$.

3.13　**已知矩阵的秩是 3，求 a 的值**

$$\begin{pmatrix}1&1&2&a&3\\2&2&3&1&4\\1&0&1&1&5\\2&3&5&5&4\end{pmatrix}.$$

3.14　**如果** $r(\boldsymbol{A})=r$. $\boldsymbol{A}$ **中能否有等于零的** $r-1$ **阶子式？能否有等于零的** r **阶子式？能否有不为零的** $r+1$ **阶子式？**

3.15　已知 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ **是** $\alpha_1,\alpha_2,\cdots,\alpha_s$ **的一个极大线性无关组，证明这两个向量组等价.**

***3.16　如** $r(\alpha_1,\alpha_2,\alpha_3)=r(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$，**证明：**$\alpha_1,\alpha_2,\alpha_3$ **与** $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ **等价.**

3.17　试用矩阵，分块矩阵，向量等形式表示下列方程组

(1) $\begin{cases} 3x_1+2x_2=5, \\ 4x_1-x_2=6. \end{cases}$　　(2) $\begin{cases} x_1+x_2=1, \\ 2x_1-x_2=5, \\ x_1+2x_2=0. \end{cases}$

3.18　**判断下列命题是否正确**

(1) **如齐次方程组** $\boldsymbol{Ax}=0$ **只有零解,那么非齐次线性方程组** $\boldsymbol{Ax}=\boldsymbol{b}$ **有唯一解;**

(2) **如** $\boldsymbol{A}$ **是** $m\times n$ **矩阵,当** $m<n$ **时,方程组** $\boldsymbol{Ax}=\boldsymbol{b}$ **无唯一解;**

(3) **如** $\eta_1,\eta_2,\eta_3,\eta_4$ **是齐次方程组** $\boldsymbol{Ax}=0$ **的基础解系,则** $\eta_1+\eta_2,\eta_2+\eta_3,\eta_3+\eta_4,\eta_4+\eta_1$ **也是** $\boldsymbol{Ax}=0$ **的基础解系;**

(4) **如** n **元齐次方程组** $\boldsymbol{Ax}=0$ **有非零解,则** $|\boldsymbol{A}|=0$.

3.19　求齐次方程组的基础解系

(1) $\begin{cases} x_1+x_2-4x_3=0, \\ 2x_1+3x_2-6x_3=0. \end{cases}$

(2) $\begin{cases} x_1+2x_2+4x_3-3x_4=0, \\ 3x_1+5x_2+6x_3-4x_4=0, \\ 3x_1+8x_2+24x_3-19x_4=0. \end{cases}$

(3) $\begin{cases} x_1+x_2-3x_4-x_5=0 \\ x_1-x_2+2x_3-x_4=0, \\ 4x_1-2x_2+6x_3+3x_4-4x_5=0, \\ 2x_1+4x_2-2x_3+4x_4-7x_5=0. \end{cases}$

(4) $\begin{cases} x_1-x_3=0, \\ x_2-x_4=0, \\ -x_1+x_3-x_5=0. \end{cases}$

3.20　**求非齐次线性方程组的通解**

$$(1)\quad \begin{cases} 3x_1+4x_2+x_3+2x_4=3, \\ 6x_1+8x_2+2x_3+5x_4=7, \\ 9x_1+12x_2+3x_3+10x_4=13. \end{cases}$$

$$(2)\quad \begin{cases} 2x_1+7x_2+3x_3+x_4=6, \\ 3x_1+5x_2+2x_3+x_4=4, \\ 9x_1+4x_2+x_3+7x_4=2. \end{cases}$$

3.21 **当** a **取何值时方程组** $\boldsymbol{Ax}=\beta$ **有解,并求出其解**

(1) $\boldsymbol{A}=\begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & -1 \\ 3 & 4 & 1 \end{pmatrix}$, $\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $\beta=\begin{pmatrix} 2 \\ 5 \\ a \end{pmatrix}$;

(2) $\boldsymbol{A}=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & a \end{pmatrix}$, $\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $\beta=\begin{pmatrix} 2 \\ a \\ 4 \end{pmatrix}$;

(3) $\boldsymbol{A}=\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$, $\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $\beta=\begin{pmatrix} -2 \\ a \\ a^2 \end{pmatrix}$.

3.22 **当** a_1, a_2, b_1, b_2 **满足什么条件时,方程组**

$$\begin{cases} x_1+x_2=a_1, \\ x_3+x_4=a_2, \\ x_1+x_3=b_1, \\ x_2+x_4=b_2. \end{cases}$$

有解,当方程组有解时求出其通解.

3.23 设有三维向量

$\alpha_1=(1,1,1)^{\mathrm{T}}$, $\alpha_2=(a,1,1)^{\mathrm{T}}$, $\alpha_3=(1,2,b)^{\mathrm{T}}$, $\beta=(2,3,4)^{\mathrm{T}}$.

问当 a, b **取何值时**

(1) β 可由 $\alpha_1, \alpha_2, \alpha_3$ 线性表出,且表达式不唯一?

(2) β 不能由 $\alpha_1, \alpha_2, \alpha_3$ 线性表示?

* **3.24** $\boldsymbol{A}$ 是 $m\times n$ 矩阵,如对任何 m 维列向量 β,非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\beta$ 都有解,证明: $\mathrm{r}(\boldsymbol{A})=m$.

* **3.25** 设 $\boldsymbol{A}^*$ 是 n 阶矩阵($n\geq 2$) $\boldsymbol{A}$ 的伴随矩阵,证明

$$\mathrm{r}(\boldsymbol{A}^*)=\begin{cases} n & \text{如 } \mathrm{r}(\boldsymbol{A})=n, \\ 1 & \text{如 } \mathrm{r}(\boldsymbol{A})=n-1, \\ 0 & \text{如 } \mathrm{r}(\boldsymbol{A})<n-1. \end{cases}$$

第 4 章　向 量 空 间

在上一章,我们以向量为工具不仅讨论了向量的线性相关、线性无关等性质,并用它研究了线性方程组解的有关理论.由于矩阵,多项式,齐次线性方程组的解,……,这些完全不同的对象也都具有和向量运算相类似的性质.数学上抓住其共性引申抽象出空间的概念,它是线性代数中最重要的基本概念之一.

当把解析几何中向量点积的运算推广到向量空间,从而可计算 n 维向量的长度,两个 n 维向量的夹角,进而引出欧氏空间的有关理论.

我们还要介绍线性变换,它是函数概念的推广,线性变换是线性代数的一个重要研究对象.

4.1　向 量 空 间

在研究问题时,往往需要明确规定所考虑的数的范围.例如,对方程 $x^2+1=0$,如限定在实数范围内,则此方程无解,而若在复数范围内,则方程有解.为防止混淆,要事先规定数的取值范围,把全体实数所组成的数集叫实数域,全体复数所组成的数集叫复数域.在本教材,我们只讨论实数域内的代数问题.

定义 4.1　设 V 是 n 维向量的非空集合,R 是实数域,如果集合 V 中的向量对于向量的加法及数乘这两个运算封闭,即

$$\forall\ \alpha,\beta\in V,\text{有 }\alpha+\beta\in V;$$

$$\forall\ \alpha\in V,\ \forall\ k\in R,\text{有 }k\alpha\in V.$$

则称 V 是实数域 R 上的向量空间.

例 1　全体三维向量 $\mathbf{R}^3=\{(x,y,z)^{\mathrm{T}} \mid x,y,z\in\mathbf{R}\}$ 是实数域上的向量空间. ■

例 2　只由零向量一个向量构成的集合 $\{0\}$ 是实数域上的向量空间. ■

例 3　集合 $V=\{(0,0,x_3)^{\mathrm{T}} \mid x_3\in\mathbf{R}\}$ 是实数域上的向量空间. ■

例 4　证明向量集合 $V=\{(x_1,x_2,x_3)^{\mathrm{T}} \mid x_1+x_2=0,\ x_1,x_2,x_3\in\mathbf{R}\}$ 是实数域上的向量空间.

证　$\forall\ \alpha,\beta\in V$, 设 $\alpha=(a_1,a_2,a_3)^{\mathrm{T}}$, $\beta=(b_1,b_2,b_3)^{\mathrm{T}}$, 则 $a_1+a_2=0$, $b_1+b_2=0$. 由于

$$\alpha+\beta=(a_1+b_1,\ a_2+b_2,\ a_3+b_3)^{\mathrm{T}}.$$

其前两个分量之和

$$(a_1+b_1)+(a_2+b_2)=(a_1+a_2)+(b_1+b_2)=0$$

为零, 故知 $\alpha+\beta\in V$, 即 V 对向量加法封闭.

$$\forall\ \alpha\in V \text{ 及 } k\in\mathbf{R}, \text{由 } k\alpha=(ka_1,ka_2,ka_3)^{\mathrm{T}},$$

知其前两个分量之和

$$ka_1+ka_2=k(a_1+a_2)=0$$

为零. 即 $k\alpha\in V$, 也就有 V 对于数乘运算封闭. 所以, V 是实数域上的向量空间. ■

例 5　证明向量集合 $V=\{(x_1,x_2,x_3)^{\mathrm{T}} \mid x_1+x_2=1,\ x_1,x_2,x_3\in\mathbf{R}\}$ 不是实数域上的向量空间.

证　如 $\alpha=(a_1,a_2,a_3)^{\mathrm{T}}\in V$, 则 $a_1+a_2=1$. 但 $2\alpha=(2a_1,2a_2,2a_3)$ 而 $2a_1+2a_2=2\neq1$, 因此 $2\alpha\notin V$, 即集合 V 对于向量的数乘运算不封闭, 所以 V 不是向量空间. ■

对于实数域上的向量空间, 其向量个数一般有无穷多个, 如何来描述这些向量呢?

定义 4.2　如 V 是实数域上的向量空间, 若 V 中有 n 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关, 而 V 中任意 $n+1$ 个向量都线性相关, 则

称 V 是 n 维向量空间, 记作 $\dim V = n$. 而这 n 个线性无关的向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 就称为向量空间的一组基.

在例 1 中, $e_1 = (1, 0, 0)^T$, $e_2 = (0, 1, 0)^T$, $e_3 = (0, 0, 1)^T$ 线性无关, 它们可表示任一向量 $\alpha = (x_1, x_2, x_3)^T$, 所以称 R^3 是三维向量空间, 称 e_1, e_2, e_3 是自然基.

在例 3 中, $\alpha = (0, 0, 1)^T$ 是线性无关的, 它可表示任一 $\beta = (0, 0, x_3)^T$, 所以三维向量的集合 $\{(0, 0, x_3)^T | x_3 \in R\}$ 构成一维向量空间, $(0, 0, 1)^T$ 是空间的一组基.

在例 4 中, $\alpha_1 = (1, -1, 0)^T$ 与 $\alpha_2 = (0, 0, 1)^T$ 是线性无关的, 对任 $\beta = (x_1, x_2, x_3)^T \in V$, 由 $x_1 + x_2 = 0$ 知 $\beta = x_1\alpha_1 + x_3\alpha_2$, 即 V 中任何向量都可由 α_1, α_2 线性表出. 所以三维向量的集合 $\{(x_1, x_2, x_3)^T | x_1 + x_2 = 0\}$ 所形成的空间是二维向量空间. $(1, -1, 0)^T$ 与 $(0, 0, 1)^T$ 是其一组基.

设 V 是 n 维向量空间, $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是一组基, 按定义 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关, 而对任意 $\beta \in V$ 有 $\alpha_1, \alpha_2, \cdots \alpha_n, \beta$ 线性相关. 根据第三章定理 3.7, β 可由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表出且表示法唯一. 设

$$\begin{aligned} \beta &= x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \\ &= (\alpha_1, \alpha_2, \cdots, \alpha_n)x \end{aligned} \tag{4.1}$$

其中 $x = (x_1, x_2, \cdots, x_n)^T$. 我们称 x 是向量 β 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标.

例如, 在三维向量空间 R^3 中, 向量 $\beta = (1, 2, -1)^T$ 在自然基下有: $\beta = e_1 + 2e_2 - e_3$, 因此 β 在自然基 e_1, e_2, e_3 下的坐标是$(1, 2, -1)^T$. 显然, $\alpha_1 = (1, 1, 1)^T$, $\alpha_2 = (0, 1, 1)^T$, $\alpha_3 = (0, 0, 1)^T$ 是线性无关的, 它也构成 R^3 的一组基. 容易验证: $\beta = \alpha_1 + \alpha_2 - 3\alpha_3$, 因此 β 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的坐标是$(1, 1, -3)^T$.

在向量空间, 基向量的选取方法是多种多样的, 同一个向量在不同的基下一般坐标是不同的, 那么这些坐标之间有什么联系呢? 下面来讨论当基向量改变后坐标是如何变换的?

对三维向量空间,如果 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 与 η_1, η_2, η_3 是两组基,设

$$
\begin{aligned}
\eta_1 &= c_{11}\varepsilon_1 + c_{21}\varepsilon_2 + c_{31}\varepsilon_3,\\
\eta_2 &= c_{12}\varepsilon_1 + c_{22}\varepsilon_2 + c_{32}\varepsilon_3,\\
\eta_3 &= c_{13}\varepsilon_1 + c_{23}\varepsilon_2 + c_{33}\varepsilon_3.
\end{aligned} \tag{4.2}
$$

为了方便,现引用矩阵表示,记

$$
C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix},
$$

则两组基之间的联系可写成

$$
(\eta_1, \eta_2, \eta_3) = (\varepsilon_1, \varepsilon_2, \varepsilon_3)C. \tag{4.2.1}
$$

矩阵 C 叫做由基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 到基 η_1, η_2, η_3 的过渡矩阵,(注: C 中的第 i 列就是 η_i 在基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 下的坐标,可以证明 C 是可逆矩阵)

如果向量 α 在这两组基下的坐标分别是 x 和 y,也就是

$$
\alpha = (\varepsilon_1, \varepsilon_2, \varepsilon_3)x \quad 及 \quad \alpha = (\eta_1, \eta_2, \eta_3)y,
$$

那么,利用过渡矩阵我们有

$$
\begin{aligned}
\alpha &= (\eta_1, \eta_2, \eta_3)y \\
&= [(\varepsilon_1, \varepsilon_2, \varepsilon_3)C]y \\
&= (\varepsilon_1, \varepsilon_2, \varepsilon_3)Cy = (\varepsilon_1, \varepsilon_2, \varepsilon_3)x,
\end{aligned}
$$

由于 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 是基,它们是线性无关的,因而 α 在这组基下的表示方法是唯一的.这就有

$$
x = Cy \quad 或 \quad y = C^{-1}x, \tag{4.3}
$$

称其为坐标变换公式.

例 6 如图是两个正方形, $\varepsilon_1, \varepsilon_2$ 与 η_1, η_2 是 R^2 的两组基.求坐标变换公式.

解 按已知条件,有

$$
\eta_1 = \quad \varepsilon_1 + \quad \varepsilon_2, \quad \eta_2 = - \quad \varepsilon_1 + \quad \varepsilon_2,
$$

因此,

图 4.1

$$(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

如向量 $\boldsymbol{\alpha}$ 在基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2$ 下的坐标是 $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$，在基 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$ 下的坐标是 $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. 那么坐标变换公式是

$$x = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} y \quad 或 \quad y = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} x. \qquad \blacksquare$$

例 7 已知 $\boldsymbol{\varepsilon}_1 = (1, 1, 1)^T$, $\boldsymbol{\varepsilon}_2 = (0, 1, 1)^T$, $\boldsymbol{\varepsilon}_3 = (0, 0, 1)^T$ 与 $\boldsymbol{\eta}_1 = (1, 2, 1)^T$, $\boldsymbol{\eta}_2 = (2, 3, 3)^T$, $\boldsymbol{\eta}_3 = (3, 7, 1)^T$ 是三维向量空间的两组基.

(1) 求由基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3$ 到基 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3$ 的过渡矩阵 C.

(2) 求 $\boldsymbol{\beta} = (1, 0, 5)^T$ 在这两组基下的坐标.

解 由 $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3) = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3)C$ 得

$$C = (\boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_2 \boldsymbol{\varepsilon}_3)^{-1}(\boldsymbol{\eta}_1 \boldsymbol{\eta}_2 \boldsymbol{\eta}_3)$$

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

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

设 β在基 ε, ε, ε 下的坐标是 $x = (x_1, x_2, x_3)^T$, 即

$$\beta = x_1 \varepsilon + x_2 \varepsilon + x_3 \varepsilon = (\varepsilon, \varepsilon, \varepsilon) x.$$

那么

$$x = (\varepsilon \varepsilon \varepsilon)^{-1} \beta$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix}.$$

关于 β在基 η, η, η 下的坐标 y, 可类似于 x 的求法, 由 $y = (\eta, \eta, \eta)^{-1}\beta$求出. 现在用坐标变换公式来求 y.

$$y = C^{-1} x = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 4 \\ -1 & 0 & -6 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \\ -2 \end{pmatrix}. \quad ■$$

* 例 8　已知 $\varepsilon = (1, 1, -1)^T$, $\varepsilon = (2, -1, 2)^T$, $\varepsilon = (3, 0, 2)^T$ 与 $\eta = (1, -2, 1)^T$, $\eta = (1, 1, 1)^T$, $\eta = (1, 4, 0)^T$ 是两组基. 求非零向量 ξ使它在这两组基下有相同的坐标.

解法 1　设 ξ的坐标为$(x_1, x_2, x_3)^T$, 即

$$\xi = (\varepsilon \varepsilon \varepsilon) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (\eta \eta \eta) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

设由基 ε, ε, ε 到基 η, η, η 的过渡矩阵是 C,

即　$(\eta, \eta, \eta) = (\varepsilon \varepsilon \varepsilon) C$,

求出

$$C = (\varepsilon \varepsilon \varepsilon)^{-1} (\eta \eta \eta) = \begin{pmatrix} 1 & -1 & -2 \\ 3 & -2 & -6 \\ -2 & 2 & 5 \end{pmatrix}.$$

那么利用过渡矩阵, 有

$$(\varepsilon\ \varepsilon\ \varepsilon)\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=(\varepsilon\ \varepsilon\ \varepsilon)C\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}.$$

由于 ε, ε, ε 是一组基, 它们线性无关, 得

$$C\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},$$

即

$$(C-I)\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=0.$$

解这个齐次方程组

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

得基础解系$(0,2,-1)^T$.

所以 $\xi\ (\varepsilon\ \varepsilon\ \varepsilon)k\begin{pmatrix}0\\2\\-1\end{pmatrix}=k\begin{pmatrix}1\\-2\\2\end{pmatrix}$, 合于所求. ■

上述解法是麻烦的, 其实利用坐标的概念直接建立方程组, 可避两组基之间的过渡矩阵, 如

解法 2 设 ξ在这两组基下的坐标都是$(x_1, x_2, x_3)^T$. 即

$$\xi=x_1\varepsilon+x_2\varepsilon+x_3\varepsilon=x_1\eta+x_2\eta+x_3\eta$$

那么 $x_1(\varepsilon-\eta)+x_2(\varepsilon-\eta)+x_3(\varepsilon-\eta)=0.$

把已知条件代入上式, 得到齐次方程组

$$\begin{cases}x_2+2x_3=0,\\3x_1-2x_2-4x_3=0,\\-2x_1+x_2+2x_3=0,\end{cases}$$

它的基础解系是$(0,2,-1)^T$，下同解法 1，得到

$$\xi = (k, -2k, 2k)^T$$

其中 k 是任意非零实数. ■

现在介绍如何判断一组向量是否能构成向量空间一组基的方法.

定理 4.1 如果 $\alpha_1, \alpha_2, \alpha_3$ 是三维向量空间的一组基，对向量组 $\beta_1, \beta_2, \beta_3$，若

$$\beta_1 = c_{11}\alpha_1 + c_{21}\alpha_2 + c_{31}\alpha_3,$$
$$\beta_2 = c_{12}\alpha_1 + c_{22}\alpha_2 + c_{32}\alpha_3,$$
$$\beta_3 = c_{13}\alpha_1 + c_{23}\alpha_2 + c_{33}\alpha_3,$$

那么 $\beta_1, \beta_2, \beta_3$ 是一组基的充分必要条件是

$$\begin{vmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{vmatrix} \neq 0. \tag{4.4}$$

* **证** 考查 $k_1\beta_1 + k_2\beta_2 + k_3\beta_3 = 0$，往证 $k_1 = k_2 = k_3 = 0$.

若记 $C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$，则由已知条件有

$$(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3)C.$$

因此由$(\alpha_1, \alpha_2, \alpha_3)C\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (\beta_1\ \beta_2\ \beta_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$ 及 $\alpha_1, \alpha_2, \alpha_3$ 是一组基，它们是线性无关的，得到

$$C\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0. \qquad (*)$$

所以，$\beta_1, \beta_2, \beta_3$ 线性无关是一组基的充要条件是上述齐次线性方程组$(*)$只有零解，这等价于系数行列式$|C| \neq 0$. ■

例 9　证明 $\alpha_1=(2,1,-3)^{\mathrm T}$, $\alpha_2=(3,2,-5)^{\mathrm T}$, $\alpha_3=(1,-1,1)^{\mathrm T}$ **是三维向量空间的一组基, 并求** $\beta=(6,2,-7)^{\mathrm T}$ **在这组基下的坐标.**

证　设 ε_1, ε_2, ε_3 是三维向量空间的自然基, 根据已知条件, 有

$$(\alpha_1,\alpha_2,\alpha_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix}2&3&1\\1&2&-1\\-3&-5&1\end{pmatrix}.$$

因为

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

据定理 4.1, α_1, α_2, α_3 是一组基.

设 β 在 α_1, α_2, α_3 这组基下的坐标是 $(x_1,x_2,x_3)^{\mathrm T}$, 即 $\beta=x_1\alpha_1+x_2\alpha_2+x_3\alpha_3$, 按分量写出来得方程组

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

用高斯消元法, 有

$$\begin{pmatrix}2&3&1&6\\1&2&-1&2\\-3&-5&1&-7\end{pmatrix}\to\begin{pmatrix}1&2&-1&2\\2&3&1&6\\-3&-5&1&-7\end{pmatrix}$$

$$\to\begin{pmatrix}1&2&-1&2\\0&-1&3&2\\0&1&-2&-1\end{pmatrix}\to\begin{pmatrix}1&2&-1&2\\&-1&3&2\\&&1&1\end{pmatrix}.$$

解出 $x_1=1$, $x_2=1$, $x_3=1$. 所以 β 在基 α_1, α_2, α_3 下的坐标是 $(1,1,1)^{\mathrm T}$. ■

*4.2 线 性 空 间

我们已经知道向量的加法与数量乘法这两个运算有 8 条基本性质(见 3.2),如果向量集合 V 对于这两个运算封闭,那么 V 就是一个向量空间(定义 4.1).除此之外,还有许多类似的情况,例如.

对于 $m\times n$ 矩阵,也有加法与数量乘法两种运算,这两个运算也有着与向量运算相类似的 8 条性质(见 2.1),同时全体 $m\times n$ 矩阵所构的集合对这两个运算是封闭的;

对于多项式,亦有加法(合并同类项)及数量乘法(数与多项式乘法)这两种运算,运算的性质也是同样的八条,全体次数小于 n 的多项式的集合对这两个运算也是封闭的.

这样的例子还可举出一些,尽管所考虑的对象完全不同,运算的定义也是有差别的,但它们有共同点,抓住其共同点进行统一的研究是数学研究问题的基本方法,因此对向量空间作进一步的抽象归纳,引入线性空间的概念.

定义 4.3 设 S 是一个非空集合,R 是实数域,对 S 中的元素规定两种代数运算:

(1) 加法 $\forall\ \alpha,\beta\in S$,有 $\alpha+\beta\in S$;

(2) 数量乘法 $\forall\ \alpha\in S$,$\forall\ k\in R$ 有 $k\alpha\in S$(简称集合 S 对这两种运算封闭).如果这两种运算还满足以下八条性质

1) 加法交换律:$\alpha+\beta=\beta+\alpha$;

2) 加法结合律:$\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$;

3) 在 S 中有一个零元素 0,使得对任意的 $\alpha\in S$,都有 $\alpha+0=\alpha$;

4) 对于 S 中每一个元素 α,都有一个负元素 $-\alpha\in S$,使得 $\alpha+(-\alpha)=0$;

5) $1\alpha=\alpha$;

6) 数乘有结合律 $k(l\alpha)=(kl)\alpha$;

7) 数乘与加法可分配 $k(\alpha+\beta)=k\alpha+k\beta$;

8) 数乘与加法可分配 $(k+l)\alpha=k\alpha+l\alpha$

(在以上性质中, α,β,γ 是集合 S 中的任意元素, k,l 是实数域 R 中任意实数)就称 S 是实数域 R 上的线性空间.

例 1　所有元素是实数的 $m\times n$ 矩阵在矩阵加法及数乘矩阵两种运算下构成一个实数域 R 上的线性空间. ■

例 2　齐次线性方程组 $Ax=0$ 的全体解向量所构成的集合, 在向量的加法及数乘向量两种运算下构成实数域上的线性空间.

■

例 3　在闭区间 $[a,b]$ 上的全体连续函数, 对函数的加法及数与函数的乘法构成实数域上的线性空间. ■

今后, 线性空间的元素仍叫做向量. 当然, 这里所讲的向量比起向量空间中的向量涵义要广泛的多.

线性空间的简单性质是:

1) 线性空间中零元素是唯一的;

2) 线性空间中每个元素有唯一的负元素;

3) $0\alpha=0$, $k0=0$, $(-1)\alpha=-\alpha$;

4) 如 $k\alpha=0$, 则或者 $k=0$, 或者 $\alpha=0$.

由于线性空间中元素间的运算与向量之间的运算有着完全类同的性质, 与前面讨论向量问题时一样, 对线性空间的元素可照样引入线性相关、线性无关等概念, 进而有线性空间的维数、基、基变换, 元素的坐标, 坐标变换等与向量空间相平行的概念. 而且前面关于向量线性相关的理论中只要把其中的向量换成线性空间的元素其论断在线性空间都是成立的.

例 4　全体二阶实矩阵在矩阵加法和数乘矩阵两种运算下构成实数域上的线性空间, 记作 $M_{2\times 2}$.

任一 2 阶矩阵 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 可用 $E_{11}=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12}=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21}=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_{22}=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ 线性表出. 即

$$A = aE_{11} + bE_{12} + cE_{21} + dE_{22},$$

而 $k_1E_{11}+k_2E_{12}+k_3E_{21}+k_4E_{22}=0$ 的充分必要条件是 $k_1=k_2=k_3=k_4=0$.

所以, $E_{11}, E_{12}, E_{21}, E_{22}$ 是线性无关的, 它们构成线性空间 $M_{2\times 2}$ 的一组基. $M_{2\times 2}$ 是 4 维线性空间. 矩阵 A 在这组基下的坐标是 $(a, b, c, d)^T$. ■

例 5 全体次数小于 3 的实系数多项式, 记作 $P_3[x]$, 在通常多项式加法及实数与多项式相乘运算下构成线性空间.

对 $e_1=1, e_2=x, e_3=x^2\in P_3[x]$, 有

$$k_1e_1 + k_2e_2 + k_3e_3 = 0 \quad k_1 = k_2 = k_3 = 0$$

因此, e_1, e_2, e_3 线性无关. 而对任 $f=a_0+a_1x+a_2x^2\in P_3[x]$, 有

$$f = a_0e_1 + a_1e_2 + a_3e_3,$$

即任何 $f\in P_3[x]$, f 可由 $1, x, x^2$ 线性表出. 所以, $e_1=1, e_2=x, e_3=x^2$ 构成 $P_3[x]$ 的一组基. $P_3[x]$ 是三维线性空间. $f=a_0+a_1x+a_2x^2$ 在这组基下的坐标是 $(a_0, a_1, a_2)^T$.

容易验证

$$\alpha_1 = 1,\ \alpha_2 = 1 + x,\ \alpha_3 = 1 + x + x^2$$

也是 $P_3[x]$ 的一组基. 由于

$$(\alpha_1\alpha_2\alpha_3) = (e_1e_2e_3)\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

所以, 由基 e_1, e_2, e_3 到基 $\alpha_1, \alpha_2, \alpha_3$ 的过渡矩阵是

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

那么, f 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的坐标 y 可用坐标变换公式求出

$$y=C^{-1}x$$

$$=\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}^{-1}\begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix}=\begin{pmatrix}a_0-a_1\\a_1-a_2\\a_2\end{pmatrix}.$$ ■

4.3 向量的内积、欧氏(Euclid)空间

有解析几何中,向量的加法及数乘向量两种运算可用来讨论向量的共线、共面这些代数中的线性相关、线性无关问题.而向量的长度,两个向量的夹角等与物理、力学相关的问题,就需利用向量的点积(数量积)运算.现在把三维向量空间的向量点积推广到一般的 n 维向量空间.

定义 4.4 设有 n 维向量 $\alpha=(a_1,a_2,\cdots,a_n)^T$ 和 $\beta=(b_1,b_2,\cdots,b_n)^T$,令

$$(\alpha,\beta)=\alpha^T\beta=a_1b_1+a_2b_2+\cdots+a_nb_n.\tag{4.5}$$

称 (α,β) 为向量 α 与 β 的内积.

容易看出,向量的内积运算满足下列性质:

(1) $(\alpha,0)=0$;

(2) $(\alpha,\beta)=(\beta,\alpha)$;

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

(4) $(k\alpha,\beta)=k(\alpha,\beta)$;

(5) $(\alpha,\alpha)\geq 0$ 且 $(\alpha,\alpha)=0$ 的充要条件是 $\alpha=0$.

定义 4.5 规定了内积的实向量空间叫做欧氏空间(Euclid).

定义 4.6 在欧氏空间,对每一个向量 α,我们规定向量 α 的

长度 $\|\alpha\| = \sqrt{(\alpha,\alpha)}$. 如果 $\|\alpha\| = 1$, 则称 α 是单位向量.

例 1 对向量 $\alpha = (1, -1, 0, 2)^{T}$, 我们有

$$(\alpha, \alpha) = 1^2 + (-1)^2 + 0^2 + 2^2 = 6.$$

因此向量 α 的长度为 $\|\alpha\| = \sqrt{6}$. 那么和 α 平行的单位向量是 $\pm\frac{1}{\sqrt{6}}(1, -1, 0, 2)^{T}$. ■

为了引入向量的夹角, 先来证明柯西不等式.

定理 4.2 (柯西(Cauchy)不等式) 在欧氏空间, 对 n 维向量 α 和 β 我们总有

$$(\alpha, \beta)^2 \leq \|\alpha\|^2 \|\beta\|^2, \tag{4.6}$$

而等号成立的充分必要条件是 α, β 线性相关.

* **证** 如果 α, β 线性无关, 那么对任意实数 x, 恒有 $x\alpha + \beta \neq 0$, 由内积的性质 5, 有

$$(x\alpha + \beta, x\alpha + \beta) > 0,$$

再用内积的性质展开上式, 得到

$$(\alpha, \alpha)x^2 + 2(\alpha, \beta)x + (\beta, \beta) > 0,$$

其左端是 x 的二次函数, 对任意的 x 这个函数值恒大于零, 故它的判别式必小于零, 即

$$(\alpha, \beta)^2 - (\alpha, \alpha)(\beta, \beta) < 0,$$

也就是
$$(\alpha, \beta)^2 < \|\alpha\|^2 \|\beta\|^2.$$

如果 α, β 线性相关, 则或有 $\alpha = 0$, 或者 $\beta = k\alpha$

当 $\alpha = 0$ 时,

$$(\alpha, \beta)^2 = (0, \beta)^2 = 0 = \|0\|^2 \|\beta\|^2 = \|\alpha\|^2 \|\beta\|^2,$$

如 $\alpha \neq 0$, 则 $\beta = k\alpha$, 这时

$$\begin{aligned}(\alpha, \beta)^2 &= (\alpha, k\alpha)(\alpha, k\alpha) = k(\alpha, \alpha)(\alpha, k\alpha) \\ &= (\alpha, \alpha)(k\alpha, k\alpha) = (\alpha, \alpha)(\beta, \beta) = \|\alpha\|^2 \|\beta\|^2,\end{aligned}$$

所以, 不论哪种情况" α, β 总有

$$(\alpha, \beta)^2 \leq \|\alpha\|^2 \|\beta\|^2,$$

而等号仅在 α,β 线性相关时成立. ■

注: 若 $\alpha=(a_1,a_2,a_3)^T$, $\beta=(b_1,b_2,b_3)^T$ 柯西不等式用坐标写出来,就是

$$a_1b_1+a_2b_2+a_3b_3\leq(a_1^2+a_2^2+a_3^2)^{\frac{1}{2}}(b_1^2+b_2^2+b_3^2)^{\frac{1}{2}}.$$

这就是中学代数里常用的不等式.

利用柯西不等式,当 α,β 都不是零向量时,立即有:

$$\frac{|(\alpha,\beta)|}{\|\alpha\|\,\|\beta\|}\leq 1.$$

因而可定义两个 n 维向量的夹角.

定义 4.7 如果 α,β 是欧氏空间的任意两个非零向量,则规定 α 与 β 的夹角 θ 为

$$\theta=\arccos\frac{(\alpha,\beta)}{\|\alpha\|\,\|\beta\|},\quad 0\leq\theta\leq\pi \tag{4.7}$$

若 $(\alpha,\beta)=0$,则称向量 α 与 β 正交.

例如, $\alpha=(1,1,1,2)^T$, $\beta=(1,0,-2,-3)^T$, $\gamma=(1,2,3,-3)^T$ 则有

$$(\alpha,\beta)=-7,\quad(\alpha,\alpha)=7,\quad(\beta,\beta)=14.$$

于是

$$\cos\langle\alpha,\beta\rangle=\frac{-7}{\sqrt{7}\sqrt{14}}=-\frac{1}{\sqrt{2}},\text{得}\langle\alpha,\beta\rangle=\frac{3}{4}\pi$$

由于 $(\alpha,\gamma)=0$,故 α 与 γ 正交,即 $\langle\alpha,\gamma\rangle=\frac{\pi}{2}$.

按定义,零向量与任何向量都正交.

定理 4.3 如 n 维向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 s 个两两正交的非零向量,则 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关.

证 设 $k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0$,

两边用 α_i 作内积,由于 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 两两正交,我们有

$$\begin{aligned}&(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s,\alpha_i)\\&=k_1(\alpha_1,\alpha_i)+k_2(\alpha_2,\alpha_i)+\cdots+k_i(\alpha_i,\alpha_i)+\cdots+k_s(\alpha_s,\alpha_i)\end{aligned}$$

$$= k_i(\alpha_i, \alpha_i) = (0, \alpha_i) = 0,$$

又因 α_i 非零, $(\alpha_i, \alpha_i) > 0$. 从而必有 $k_i = 0$ $(i = 1, 2, \cdots, s)$.

所以, $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关. ■

在 n 维空间至多有 n 个向量线性无关, 根据定理 4.3, 那么 n 维欧氏空间最多有 n 个互相正交的非零向量. 而且也确实存在 n 个互相正交的非零向量(例如自然基), 这样就可把三维空间的直角坐标系推广到 n 维欧氏空间, 引入标准正交基的概念.

定义 4.8 n 维欧氏空间的一组基 $e_1, e_2, \cdots, e_n$ 如满足每两个基向量都互相正交, 且每个 e_i $(i = 1, 2, \cdots, n)$ 都是单位向量, 则称 $e_1, e_2, \cdots, e_n$ 是欧氏空间的一组标准正交基.

例如, $\alpha_1 = \frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\end{pmatrix}$, $\alpha_2 = \frac{1}{\sqrt{5}}\begin{pmatrix}-2\\1\end{pmatrix}$ 都是单位向量且相互正交. 因此, α_1, α_2 是二维欧氏空间的一组标准正交基.

例 2 设 $\alpha_1, \alpha_2, \alpha_3$ 是三维欧氏空间的一组标准正交基, 求向量 β 在这组基下的坐标.

解 设 $\beta = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3$, 用 α_1 作内积

$$\begin{aligned}(\beta, \alpha_1) &= (x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3, \alpha_1)\\ &= (x_1\alpha_1, \alpha_1) + (x_2\alpha_2, \alpha_1) + (x_3\alpha_3, \alpha_1)\\ &= x_1(\alpha_1, \alpha_1) = x_1.\end{aligned}$$

类似有 $x_2 = (\beta, \alpha_2)$, $x_3 = (\beta, \alpha_3)$.

所以, 向量 β 在标准正交基下的坐标就是 β 与每个基向量的内积. ■

在 2.4 中已经知道, 如果 n 阶矩阵 A 满足条件

$$AA^T = A^TA = I,$$

则 A 是正交矩阵. 现在从内积的角度来看正交矩阵. 设

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

是三阶正交矩阵，那么由 $AA^T=I$，即

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

可得到:

$$a_1^2+a_2^2+a_3^2=b_1^2+b_2^2+b_3^2=c_1^2+c_2^2+c_3^2=1,$$

$$a_1b_1+a_2b_2+a_3b_3=a_1c_1+a_2c_2+a_3c_3=b_1c_1+b_2c_2+b_3c_3=0$$

这说明矩阵 A 的三个行向量都是单位向量，且两两正交. 所以，正交矩阵的行向量构成标准正交基. 类似地，利用 $A^TA=I$ 可知 A 的列向量也是标准正交基.

例 3 已知 $\alpha_1=\frac{1}{3}(1,-2,-2)^T$，$\alpha_2=\frac{1}{3}(-2,1,-2)^T$，$\alpha_3=\frac{1}{3}(-2,-2,1)^T$ 是标准正交基. 求 $\beta=(2,0,1)^T$ 在这组基下的坐标.

解法 1 利用坐标的概念建立方程组，用高斯消元来求解. 设 $\beta=x_1\alpha_1+x_2\alpha_2+x_3\alpha_3$，按分量写出来，有

$$\begin{cases} \frac{1}{3}x_1-\frac{2}{3}x_2-\frac{2}{3}x_3=2, \\ -\frac{2}{3}x_1+\frac{1}{3}x_2-\frac{2}{3}x_3=0, \\ -\frac{2}{3}x_1-\frac{2}{3}x_2+\frac{1}{3}x_3=1, \end{cases}$$

解此方程组得 $x_1=0$，$x_2=-2$，$x_3=-1$，即 β 在 $\alpha_1,\alpha_2,\alpha_3$ 这组基下的坐标是 $(0,-2,-1)^T$. ■

解法 2 利用正交矩阵性质，通过矩阵运算来求解. 设 $\beta=x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=(\alpha_1\ \alpha_2\ \alpha_3)x$，其中 $x=(x_1,x_2,x_3)^T$.

由于 $\alpha_1,\alpha_2,\alpha_3$ 是标准正交基，$A=(\alpha_1,\alpha_2,\alpha_3)$ 是正交矩阵，$A^{-1}=A^T$，因此 $x=A^{-1}\beta=A^T\beta$ 可得

$$x=\frac{1}{3}\begin{pmatrix}1&-2&-2\\-2&1&-2\\-2&-2&1\end{pmatrix}\begin{pmatrix}2\\0\\1\end{pmatrix}=\begin{pmatrix}0\\-2\\-1\end{pmatrix}.$$ ■

解法 3 根据例 2,用内积来计算

$x_1=(\beta,\alpha_1)=0$, $x_2=(\beta,\alpha_2)=-2$, $x_3=(\beta,\alpha_3)=-1$. 即 $\beta=-2\alpha_2-\alpha_3$, 因此 β 的坐标是 $(0,-2,-1)^T$. ■

现设 $\alpha_1,\alpha_2,\alpha_3$ 和 β_1,β_2,β_3 是两组标准正交基, 由基 $\alpha_1,\alpha_2,\alpha_3$ 到基 β_1,β_2,β_3 的过渡矩阵是 C. 即

$$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2,\alpha_3)C$$

那么

$$\begin{pmatrix}\beta_1^T\\\beta_2^T\\\beta_3^T\end{pmatrix}=(\beta_1\ \beta_2\ \beta_3)^T=\left((\alpha_1\ \alpha_2\ \alpha_3)C\right)^T=C^T\begin{pmatrix}\alpha_1^T\\\alpha_2^T\\\alpha_3^T\end{pmatrix}.$$

利用 $\alpha_1,\alpha_2,\alpha_3$ 与 β_1,β_2,β_3 都是标准正交基, 作内积就有:

$$I=\begin{pmatrix}\beta_1^T\\\beta_2^T\\\beta_3^T\end{pmatrix}(\beta_1\ \beta_2\ \beta_3)=C^T\begin{pmatrix}\alpha_1^T\\\alpha_2^T\\\alpha_3^T\end{pmatrix}(\alpha_1\ \alpha_2\ \alpha_3)C=C^TIC=C^TC,$$

可见 C 是正交矩阵. 我们可以证明:

定理 4.4 由标准正交基到标准正交基的过渡矩阵是正交矩阵; 反过来, 如果两组基之间的过渡矩阵是正交矩阵, 有一组基是标准正交基, 那么另一组基一定也是标准正交的. ■

* **例 4** 已知 α_1,α_2 是二维欧氏空间的一组标准正交基, 证明

$$\beta_1=\frac{1}{\sqrt{2}}(\alpha_1+\alpha_2),\quad \beta_2=\frac{1}{\sqrt{2}}(-\alpha_1+\alpha_2)$$

也是一组标准正交基.

证 根据已知条件 $(\beta_1,\beta_2)=(\alpha_1,\alpha_2)C$, 其中

$$C=\begin{pmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix},$$

由于$|C|\neq 0$，按定理 4.1 知 β_1,β_2 是一组基. 又因 C 是正交矩阵，α_1,α_2 是标准正交基，据定理 4.4 知 β_1,β_2 是标准正交基. ■

在标准正交基下处理问题是方便的，那么如何把一组基改造成为标准正交基呢?

施密特(Schmidt)标准正交化方法：在欧氏空间，对任意给的一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 可以构造出一组标准正交基 $\gamma_1,\gamma_2,\cdots,\gamma_n$，使 γ_i 是 $\alpha_1,\alpha_2,\cdots,\alpha_i$ 的线性组合，$i=1,2,\cdots,n$.

下面以三维欧氏空间为例，说明施密特正交化方法.

设 $\alpha_1,\alpha_2,\alpha_3$ 是 R^3 的一组基，令 $\beta_1=\alpha_1$，

令 $\beta_2=k\beta_1+\alpha_2$，选取 k 使得$(\beta_1,\beta_2)=0$，这只需两边对 β_1 作内积，有

$$k(\beta_1,\beta_1)+(\beta_1,\alpha_2)=(\beta_1,\beta_2)=0$$

从而取 $k=-\dfrac{(\beta_1,\alpha_2)}{(\beta_1,\beta_1)}$ 即可.

令 $\beta_3=l\beta_1+m\beta_2+\alpha_3$，选取 l,m 使得

$$(\beta_1,\beta_3)=0,\ (\beta_2,\beta_3)=0$$

类似可得到

$$l=-\frac{(\beta_1,\alpha_3)}{(\beta_1,\beta_1)},\quad m=-\frac{(\beta_2,\alpha_3)}{(\beta_2,\beta_2)}.$$

这样构造出的 β_1,β_2,β_3 是两两正交的. 并且由 β_i 的构造方法，可看出 $\alpha_1,\alpha_2,\alpha_3$ 与 β_1,β_2,β_3 可以互相线性表出，它们是等价的向量组，因而有相同的秩，所以 β_1,β_2,β_3 必线性无关，是一组基.

再对 β_1,β_2,β_3 单位化，即令

$$\gamma_1=\frac{\beta_1}{\|\beta_1\|},\quad \gamma_2=\frac{\beta_2}{\|\beta_2\|},\quad \gamma_3=\frac{\beta_3}{\|\beta_3\|}.$$

这样得到的 $\gamma_1, \gamma_2, \gamma_3$ 就是一组标准正交基.

下面给出施密特正交化的一般公式:

设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关,构造

$$\beta_1 = \alpha_1,\quad \beta_2 = -\frac{(\beta_1, \alpha_2)}{(\beta_1, \beta_1)}\beta_1 + \alpha_2, \cdots$$

$$\beta_i = -\frac{(\beta_1, \alpha_i)}{(\beta_1, \beta_1)}\beta_1 - \frac{(\beta_2, \alpha_i)}{(\beta_2, \beta_2)}\beta_2 - \cdots - \frac{(\beta_{i-1}, \alpha_i)}{(\beta_{i-1}, \beta_{i-1})}\beta_{i-1} + \alpha_i. \tag{4.8}$$

令 $i = 2, 3, \cdots, s$,就得到两两正交的非零向量组 $\beta_1, \beta_2, \cdots, \beta_s$. 再把它们单位化为:

$$\gamma_1, \gamma_2, \cdots, \gamma_s$$

其中 $\gamma_j = \dfrac{\beta_j}{\|\beta_j\|}$, $j = 1, 2, \cdots, s$. 这时 $\gamma_1, \gamma_2, \cdots, \gamma_s$ 是标准正交的.

例 5　已知 $\alpha_1 = (1, 1, 1)^T$, $\alpha_2 = (1, 1, 2)^T$, $\alpha_3 = (0, 4, -1)^T$ 试用施密特方法把其标准正交化.

解　令 $\beta_1 = \alpha_1 = (1, 1, 1)^T$

再令 $\beta_2 = k\beta_1 + \alpha_2 = k(1, 1, 1)^T + (1, 1, 2)^T$,两边用 β_1 作内积,且令$(\beta_1, \beta_2) = 0$,有

$$3k + 4 = 0$$

所以 $\beta_2 = -\dfrac{4}{3}\beta_1 + \alpha_2 = -\dfrac{4}{3}(1, 1, 1)^T + (1, 1, 2)^T = \left(-\dfrac{1}{3}, -\dfrac{1}{3}, \dfrac{2}{3}\right)^T$

令 $\beta_3 = l\beta_1 + m\beta_2 + \alpha_3$,待定 l, m 使 β_3 与 β_1, β_2 都正交. 用 β_1 作内积,且令$(\beta_1, \beta_3) = 0$,有

$$3l + 3 = 0, \quad 得 \quad l = -1.$$

用 β_2 作内积,且令$(\beta_2, \beta_3) = 0$,有

$$\frac{2}{3}m - 2 = 0, \quad 得 \quad m = 3.$$

所以 $\beta_3=-\beta_1+3\beta_2+\alpha_3=(-2,2,0)^T$,

对 β_1,β_2,β_3 单位化,得

$$\gamma_1=\frac{\beta_1}{\|\beta_1\|}=\frac{1}{\sqrt{3}}(1,1,1)^T,$$

$$\gamma_2=\frac{\beta_2}{\|\beta_2\|}=\frac{1}{\sqrt{6}}(-1,-1,2)^T,$$

$$\gamma_3=\frac{\beta_3}{\|\beta_3\|}=\frac{1}{\sqrt{2}}(-1,1,0)^T.$$

以上 $\gamma_1,\gamma_2,\gamma_3$ 是一组标准正交基. ■

*4.4 子 空 间

本节介绍一个空间的子空间的概念及描述子空间的方法.

定义 4.9 线性空间 S 的一个非空子集 M,如果对线性空间的加法及数量乘法两个运算封闭,即 $\forall\alpha,\beta\in M$, $k\in R$ 有

$$\alpha+\beta\in M,\quad k\alpha\in M$$

对称 M 是 S 的线性子空间,简称子空间.

例 1 全体二阶矩阵在矩阵的加法及数乘运算下构成线性空间 $M_{2\times2}$,而二阶实对称矩阵的集合 M 就是 $M_{2\times2}$ 的子空间. ■

例 2 A 是 $m\times n$ 矩阵,齐次方程组 $Ax=0$ 的解的集合就是 R^n 的子空间,叫做 A 的解空间. ■

例 3 在线性空间 S 中,由其零元素所组成的集合是 S 的子空间,通常叫做零子空间. ■

设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是线性空间 S 中一组向量,这组向量所有可能的线性组合

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s$$

所构成的集合 $L(\alpha_1,\alpha_2,\cdots,\alpha_s)$ 是非空的,而且对两种运算封闭,因而它是 S 的子空间,叫做由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 生成的子空间.

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

A **的列向量** $\beta_1=\begin{pmatrix}1\\1\\2\end{pmatrix}$, $\beta_2=\begin{pmatrix}0\\2\\-1\end{pmatrix}$, $\beta_3=\begin{pmatrix}1\\3\\1\end{pmatrix}$ **所生成的空间** $L(\beta_1, \beta_2, \beta_3)$ **称为** A **的列空间, 记为** $\mathrm{R}(A)$

A **的行向量** $\alpha_1=(1,0,1)$, $\alpha_2=(1,2,3)$, $\alpha_3=(2,-1,1)$ **所生成的空间** $L(\alpha_1, \alpha_2, \alpha_3)$ **称为** A **的行空间. 记为** $\mathrm{R}(A^T)$.

利用空间解析几何的知识, 由于 $\beta_1+\beta_2=\beta_3$, A **的列空间也就是与** β_1, β_2 **共面的向量, 设** $(x,y,z)^T$ **与** β_1, β_2 **共面. 那么**

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

展开整理得: $5x-y-2z=0$

过原点的平面 $5x-y-2z=0$ **上的全体向量是** A **的列空间, 它是三维空间** R^3 **的子空间.** ■

子空间它本身就是一个线性空间, 因此线性空间的维数、基、坐标等概念对子空间也是完全适用的.

例如, 在例 4 中, A **的列空间** $\mathrm{R}(A)$ **就是一个二维的子空间,** β_1, β_2 **是其一组基.**

例 5　在三维向量空间, 前两个分量之和为零的向量构成三维空间的子空间. 其中

$$\alpha_1=(1,-1,0)^T, \quad \alpha_2=(0,0,1)^T$$

是线性无关的, 它能表示这子空间中的任一向量 $\beta=(a,-a,b)^T$. **所以这些三维向量构成二维子空间,** α_1, α_2 **是其一组基.** ■

关于子空间我们可以证明

定理 4.5　两个向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ **与** $\beta_1, \beta_2, \cdots, \beta_t$ **生成相同**

的子空间的充分必要条件是这两个向量组等价. 而子空间的维数等于向量组的秩. ■

例 6 已知

$$A=\begin{pmatrix}1 & 1 & 2 & 3\\ 1 & 3 & 6 & 1\\ 1 & -5 & -10 & 9\end{pmatrix}.$$

求 A 的列空间 $R(A)$ 及行空间 $R(A^T)$ 的基.

解 根据定理 4.5, 等价的向量组生成相同的子空间, 对 A 作初等行变换, 有

$$A=\begin{pmatrix}1 & 1 & 2 & 3\\ 1 & 3 & 6 & 1\\ 1 & -5 & -10 & 9\end{pmatrix}\rightarrow\cdots\rightarrow\begin{pmatrix}1 & 1 & 2 & 3\\ 0 & 1 & 2 & -1\\ 0 & 0 & 0 & 0\end{pmatrix}=U.$$

U 的行向量组与 A 的行向量组等价. 而秩 $r(U)=2$, 所以 A 的行空间 $R(A^T)$ 是二维子空间. $\alpha_1=(1,1,2,3)$, $\alpha_2=(1,3,6,1)$ 是行向量组的极大线性无关组, 也就是行空间的基.

由于矩阵的行秩等于其列秩, 所以 A 的列空间 $R(A)$ 是二维的, $\beta_1=(1,1,1)^T$, $\beta_2=(1,3,-5)^T$ 是它的一组基. ■

定义 4.10 设 W_1, W_2 是欧氏空间的两个子空间. 如果对于任意的 $\alpha\in W_1$, $\beta\in W_2$, 恒有

$$(\alpha,\beta)=0.$$

则称 W_1, W_2 是正交的, 记为 $W_1\perp W_2$.

定义 4.11 设 W 是欧氏空间的子空间, 一个向量 α 如果对于任意的 $\beta\in W$, 恒有

$$(\alpha,\beta)=0$$

则称 α 与子空间 W 正交, 记为 $\alpha\perp W$.

例 7 已知

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

则 A 的解空间的基就是齐次方程组

$$\begin{cases} x_1 + 2x_3 = 0, \\ x_1 + x_2 + 4x_3 = 0, \end{cases}$$

的基础解系 $\eta = (-2, -2, 1)^T$. 易见 η 与 A 的行向量 $\alpha_1 = (1, 0, 2)$, $\alpha_2 = (1, 1, 4)$ 都正交.

所以, A 的行空间⊥A 的解空间, 即可得 $\eta \perp$ A 的行空间. ■

例 8 求 $A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ 解空间的一组标准正交基.

解 对于齐次线性方程组

$$\begin{cases} x_1 + x_2 + x_4 = 0, \\ x_3 + x_4 = 0. \end{cases}$$

基础解系是: $\eta_1 = (-1, 1, 0, 0)^T$, $\eta_2 = (-1, 0, -1, 1)^T$, 将其标准正交化, 令 $\beta_1 = \eta_1$, 则

$$\beta_2 = \eta_2 - \frac{(\beta_1, \eta_2)}{(\beta_1, \beta_1)}\beta_1 = \frac{1}{2}(-1, -1, -2, -2)^T,$$

那么

$$\gamma_1 = \frac{\beta_1}{\|\beta_1\|} = \frac{1}{\sqrt{2}}(-1, 1, 0, 0)^T,$$

$$\gamma_2 = \frac{\beta_2}{\|\beta_2\|} = \frac{1}{\sqrt{10}}(-1, -1, -2, -2)^T$$

是解空间的一组标准正交基. ■

*4.5 线性变换

在前面学习了向量空间的一些基本概念, 那么向量空间的元素之间会有些什么样的联系呢?

函数 $y = f(x)$ 反映了两个变量 x 和 y 之间的种种联系, 作为函数概念的一个拓广, 我们可讨论向量空间、线性空间元素间的对

应关系 σ 在向量空间中的这种讨论可以完全平行的推广到线性空间，为了简洁下面的论述只针对向量空间，作为例题则向量空间、线性空间都涉及到.

设 V 是实数域 R 上的向量空间，如果对任意的向量 $\alpha\in V$，按照一定的规则，总有 V 中一个确定的向量 $\sigma(\alpha)$ 与它对应，那么称 σ 是向量空间 V 的一个变换. $\sigma(\alpha)$ 叫做 α 在变换 σ 下的象，α 叫做 $\sigma(\alpha)$ 在变换 σ 下的原象. 现在要研究的是向量空间的一类特殊的变换，即所谓的线性变换，它是线性代数中的一个重要课题.

定义 4.12 设 σ 是实数域 R 上的向量空间 V 的一个变换，如果对任意 $\alpha, \beta\in V$，及任意常数 $k\in R$，变换 σ 保持加法及数乘向量两种运算，即

$$\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta),$$
$$\sigma(k\alpha)=k\sigma(\alpha),$$

则称 σ 是向量空间 V 的一个线性变换.

例 1　设 V 是二维向量空间，$\alpha=\begin{pmatrix}x\\y\end{pmatrix}\in V$. 我们来看平面上的几种典型的线性变换.

(1) 如 $\sigma(\alpha)=c\alpha$，c 是一固定的常数. 则因

$$\sigma(\alpha+\beta)=c(\alpha+\beta)=c\alpha+c\beta=\sigma(\alpha)+\sigma(\beta),$$
$$\sigma(k\alpha)=c(k\alpha)=k(c\alpha)=k\sigma(\alpha),$$

σ 保持加法及数乘运算，它是向量空间 V 的一个线性变换. 通常称其为数乘变换. 从几何意义上看，这个变换是把向量放大 c 倍（如果 $c>1$）；把向量压缩 c 倍（如果 $0<c<1$）.

(2) 如 $\sigma(\alpha)=\begin{pmatrix}x\\0\end{pmatrix}$，容易验证 σ 也是 V 的线性变换. 通常称其为投影变换. 其几何意义是把向量 α 投影到 x 轴上. 类似有对 y 轴的投影变换.

(3) 如 $\sigma(\alpha)=\begin{pmatrix}-x\\ y\end{pmatrix}$, σ是 V 的线性变换, 通常叫做镜象变换. 若把 y 轴理解为一面镜子, 则 $\sigma(\alpha)$就是 α关于这镜面所成的象.

(4) 如 $\sigma(\alpha)=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$, σ是把向量 α围绕坐标原点按反时针方向旋转 θ角, 它是一个线性变换, 通常叫做旋转变换.

(5) $\sigma(\alpha)=\alpha$ 叫做恒等变换;

$\sigma(\alpha)=0$ 叫做零变换. ■

例 2 V 是三维向量空间, 给定一个单位向量 α_0, 对 V 中向量 α 令

$$\sigma(\alpha)=(\alpha,\alpha_0)\alpha_0.$$

则有

$$\sigma(\alpha+\beta)=(\alpha+\beta,\alpha_0)\alpha_0=[(\alpha,\alpha_0)+(\beta,\alpha_0)]\alpha_0$$
$$=(\alpha,\alpha_0)\alpha_0+(\beta,\alpha_0)\alpha_0=\sigma(\alpha)+\sigma(\beta),$$
$$\sigma(k\alpha)=(k\alpha,\alpha_0)\alpha_0=k(\alpha,\alpha_0)\alpha_0=k\sigma(\alpha),$$

所以 σ是三维空间 V 的线性变换.

下面我们来看这个线性变换的几何意义. 由于

$$(\alpha,\alpha_0)=\alpha\cdot\alpha_0=|\alpha||\alpha_0|\cos\theta$$
$$=|\alpha|\cos\theta=|\overrightarrow{OM}|,$$

知 $\sigma(\alpha)=(\alpha,\alpha_0)\alpha_0=\overrightarrow{OM}$.

所以, σ是把向量 α投影到 α_0 上所得到的投影向量. ■

图 4.2

例 3 设 A 是三阶正交矩阵, x 是三维列向量. 令

$$\sigma(x) = Ax.$$

由矩阵乘法的性质知道 σ保持加法和数乘两个运算. 所以 σ是三维向量空间上的一个线性变换. 通常叫做正交变换. ■

下面看一个线性空间上线性变换的例子.

例 4　设 $P_n[x]$是次数小于 n 的实系数多项式所构成的线性空间. 对任何 $f \in P_n[x]$

$$\sigma(f) = \frac{df}{dx} = f',$$

由导数的性质知道

$$\sigma(f+g) = (f+g)' = f' + g' = \sigma(f) + \sigma(g),$$

$$\sigma(kf) = (kf)' = kf' = k\sigma(f),$$

即求导运算是 $P_n[x]$上的一个线性变换. ■

线性变换的简单性质是:

定理 4.6　设 σ是向量空间 V 上的线性变换, 则有

(1) $\sigma(0) = 0$;

(2) $\sigma(-\alpha) = -\sigma(\alpha)$, $\forall \alpha \in V$;

(3) $\sigma(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s) = k_1\sigma(\alpha_1) + k_2\sigma(\alpha_2) + \cdots + k_s\sigma(\alpha_s)$, 其中 $\alpha_i \in V$, $k_i \in R$, $i = 1, 2, \cdots, s$. 特别地, 如果 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关, 则 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_s)$也线性相关.

证　(1) $\sigma(0) = \sigma(0\alpha) = 0\sigma(\alpha) = 0$;

(2) $\sigma(\alpha + (-\alpha)) = \sigma(\alpha) + \sigma(-\alpha)$, 又 $\sigma(\alpha + (-\alpha)) = \sigma(0) = 0$, 故有 $\sigma(-\alpha) = -\sigma(\alpha)$.

(3) 留作练习. ■

为了描述向量空间的元素, 对于向量空间引入了基的概念, 那么如何来描述线性变换呢?

设 V 是三维向量空间, $\alpha_1, \alpha_2, \alpha_3$ 是 V 的一组基, σ是 V 的线性变换. 如果对 $\alpha \in V$, α在基 $\alpha_1, \alpha_2, \alpha_3$ 下的坐标是 $x = (x_1, x_2, x_3)^T$, $\sigma(\alpha)$在这组基下的坐标是 $y = (y_1, y_2, y_3)^T$. 即

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = (\alpha_1\ \alpha_2\ \alpha_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= (\alpha_1\ \alpha_2\ \alpha_3)x$$

$$\sigma(\alpha) = y_1\alpha_1 + y_2\alpha_2 + y_3\alpha_3 = (\alpha_1\ \alpha_2\ \alpha_3)\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$= (\alpha_1\ \alpha_2\ \alpha_3)y.$$

假若我们知道了基向量在线性变换 σ下的象,设为

$$\begin{aligned} \sigma(\alpha_1) &= a_{11}\alpha_1 + a_{21}\alpha_2 + a_{31}\alpha_3, \\ \sigma(\alpha_2) &= a_{12}\alpha_1 + a_{22}\alpha_2 + a_{32}\alpha_3, \\ \sigma(\alpha_3) &= a_{13}\alpha_1 + a_{23}\alpha_2 + a_{33}\alpha_3, \end{aligned} \tag{4.9}$$

利用矩阵,上述关系可写成

$$\sigma(\alpha_1,\ \alpha_2,\ \alpha_3) \triangleq (\sigma(\alpha_1),\ \sigma(\alpha_2),\ \sigma(\alpha_3))$$

$$= (\alpha_1\ \alpha_2\ \alpha_3)\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \tag{4.9.1}$$

我们就很容易推导出向量 α的坐标 x 与象 $\sigma(\alpha)$的坐标 y 之间的关系 . 这是因为

$$\sigma(\alpha) = \sigma[x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3]$$

$$= x_1\sigma(\alpha_1) + x_2\sigma(\alpha_2) + x_3\sigma(\alpha_3)$$

$$= (\sigma(\alpha_1),\ \sigma(\alpha_2),\ \sigma(\alpha_3))\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= (\alpha_1\ \alpha_2\ \alpha_3)\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

又因 $\sigma(\alpha)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$，而 $\alpha_1,\alpha_2,\alpha_3$ 是一组基，$\sigma(\alpha)$ 在基下的表示法唯一，所以有

$$\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},\tag{4.10}$$

简记为 $y=Ax$.

从上面的分析讨论我们看出，要描述一个线性变换 σ 关键是搞清楚 σ 在一组基 $\alpha_1,\alpha_2,\alpha_3$ 下的象 $\sigma(\alpha_1),\sigma(\alpha_2),\sigma(\alpha_3)$ 变换成为什么样的向量．我们称矩阵 A 是线性变换 σ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的对应矩阵．以上讨论可总结为下述定理

定理 4.7 设 n 维向量空间 V 的线性变换 σ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的对应矩阵是 A，向量 α 及 $\sigma(\alpha)$ 在这组基下的坐标是 $x=(x_1,x_2,\cdots,x_n)^T$ 和 $y=(y_1,y_2,\cdots,y_n)^T$，则 $y=Ax$. ■

例 5 在三维向量空间，取基 $\varepsilon_1=(1,0,0)^T$，$\varepsilon_2=(0,1,0)^T$，$\varepsilon_3=(0,0,1)^T$．对向量 $\alpha=(x,y,z)^T$ 作投影变换 $\sigma(\alpha)=(x,y,0)^T$．那么，基向量的象是

$$\sigma(\varepsilon_1)=(1,0,0)^T=\varepsilon_1,$$
$$\sigma(\varepsilon_2)=(0,1,0)^T=\varepsilon_2,$$
$$\sigma(\varepsilon_3)=(0,0,0)^T=0,$$

因此，投影变换 σ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的对应矩阵是

$$A=\begin{pmatrix}1&&\\&1&\\&&0\end{pmatrix}.$$ ■

例 6 已知 $\alpha_0=(l,m,n)^T$ 是单位向量，σ 是到 α_0 的投影变换 $\sigma(\alpha)=(\alpha,\alpha_0)\alpha_0$（见例 2），那么基向量 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 的象是

$$\sigma(\varepsilon_1)=(\varepsilon_1,\alpha_0)\alpha_0=l\alpha_0=(l^2,lm,ln)^T,$$

$$\sigma(\varepsilon) = (\varepsilon, \alpha)\alpha = m\alpha = (ml, m^2, mn)^T,$$

$$\sigma(\varepsilon) = (\varepsilon, \alpha)\alpha = n\alpha = (nl, nm, n^2)^T.$$

因此,在 α 上的投影变换 σ在基 $\varepsilon, \varepsilon, \varepsilon$下的对应矩阵是

$$A = \begin{matrix} l^2 & ml & nl \\ lm & m^2 & nm \\ ln & mn & n^2 \end{matrix}. \qquad \blacksquare$$

例 7　**如果线性变换 σ在基 $\varepsilon, \varepsilon, \varepsilon$下的对应矩阵是**

$$A = \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix},$$

求 σ在基 $\varepsilon, \varepsilon, \varepsilon$下的对应矩阵.

解　**按对应矩阵的定义,有**

$$\sigma(\varepsilon) = \varepsilon + 4\varepsilon + 7\varepsilon = 7\varepsilon + 4\varepsilon + \varepsilon,$$

$$\sigma(\varepsilon) = 2\varepsilon + 5\varepsilon + 8\varepsilon = 8\varepsilon + 5\varepsilon + 2\varepsilon,$$

$$\sigma(\varepsilon) = 3\varepsilon + 6\varepsilon + 9\varepsilon = 9\varepsilon + 6\varepsilon + 3\varepsilon,$$

那么 $$(\sigma(\varepsilon), \sigma(\varepsilon), \sigma(\varepsilon)) = (\varepsilon, \varepsilon, \varepsilon)\begin{matrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{matrix},$$

所以,线性变换 σ在基 $\varepsilon, \varepsilon, \varepsilon$下的对应矩阵是

$$B = \begin{matrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{matrix}. \qquad \blacksquare$$

下面我们来讨论在两组基下的线性变换对应矩阵之间的联系:

设 α, α, α 与 β, β, β 是三维向量空间 V **的两组基,由基 α, α, α 到基 β, β, β 的过渡矩阵是** C,**即**

$$(\beta, \beta, \beta) = (\alpha, \alpha, \alpha)C.$$

如果线性变换 σ在这两组基下的对应矩阵分别是 A **和** B,**即**

$$\sigma(\alpha_1,\alpha_2,\alpha_3)=(\sigma(\alpha_1),\sigma(\alpha_2),\sigma(\alpha_3))=(\alpha_1,\alpha_2,\alpha_3)A,$$
$$\sigma(\beta_1,\beta_2,\beta_3)=(\sigma(\beta_1),\sigma(\beta_2),\sigma(\beta_3))=(\beta_1,\beta_2,\beta_3)B.$$

那么

$$\begin{aligned}\sigma(\beta_1,\beta_2,\beta_3)&=\sigma[(\alpha_1\ \alpha_2\ \alpha_3)C]\\&=[\sigma(\alpha_1,\alpha_2,\alpha_3)]C=(\alpha_1,\alpha_2,\alpha_3)AC.\end{aligned}$$

又有

$$\begin{aligned}\sigma(\beta_1,\beta_2,\beta_3)&=(\beta_1,\beta_2,\beta_3)B\\&=(\alpha_1,\alpha_2,\alpha_3)CB.\end{aligned}$$

由于 $\alpha_1,\alpha_2,\alpha_3$ 是一组基,它们线性无关 . 所以 $\sigma(\beta_1),\sigma(\beta_2),\sigma(\beta_3)$ 在这组基下的表示方法唯一,从而得到

$$AC=CB.$$

因为过渡矩阵 C 是可逆的,所以线性变换 σ在两组基下的对应矩阵 A, B 满足关系:

$$B=C^{-1}AC,\tag{4.11}$$

利用这一关系,例 7 的解法 2 是:

由于 $(\varepsilon_3,\varepsilon_2,\varepsilon_1)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$,

得知从基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 到基 $\varepsilon_3,\varepsilon_2,\varepsilon_1$ 的过渡矩阵是

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

那么,线性变换 σ在基 $\varepsilon_3,\varepsilon_2,\varepsilon_1$ 下的对应矩阵 B 是

$$B=C^{-1}AC=\begin{pmatrix}&&1\\&1&\\1&&\end{pmatrix}^{-1}\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}&&1\\&1&\\1&&\end{pmatrix}$$

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

例 8　在二维向量空间,有线性变换 σ,

$$\sigma\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}2x+y\\-x+2y\end{pmatrix}.$$

求 σ在自然基 $\varepsilon_1,\varepsilon_2$ 及基 $\eta_1=(1,1)^T,\eta_2=(1,-1)^T$ 下的对应矩阵.并当 $\alpha=(1,3)^T$ 时,求 $\sigma(\alpha)$在自然基下的坐标.

解　按线性变换 σ的定义

$$\sigma(\varepsilon_1)=\sigma\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}2\\-1\end{pmatrix}=2\varepsilon_1-\varepsilon_2,\quad \sigma(\varepsilon_2)=\sigma\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix}=\varepsilon_1+2\varepsilon_2,$$

得到 σ在自然基下的对应矩阵 A

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

由于 $(\eta_1,\eta_2)=(\varepsilon_1,\varepsilon_2)\begin{pmatrix}1&1\\1&-1\end{pmatrix}$,

故从基 $\varepsilon_1,\varepsilon_2$ 到基 η_1,η_2 的过渡矩阵是 $C=\begin{pmatrix}1&1\\1&-1\end{pmatrix}$,**那么,线性变换 σ在基 η_1,η_2 下的对应矩阵** B **是**:

$$B=C^{-1}AC=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&-\frac{1}{2}\end{pmatrix}\begin{pmatrix}2&1\\-1&2\end{pmatrix}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$$

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

设 $\sigma(\alpha)$在基 $\varepsilon_1,\varepsilon_2$ 下坐标为 $Y=(y_1,y_2)^T$,**由于**

$$y=Ax=\begin{pmatrix}2&1\\-1&2\end{pmatrix}\begin{pmatrix}1\\3\end{pmatrix}=\begin{pmatrix}5\\5\end{pmatrix},$$

所以,$\sigma(\alpha)$在自然基 $\varepsilon_1,\varepsilon_2$ 下的坐标是$(5,5)^T$.　■

例 9 设 σ 是向量空间 V 的正交变换(参看例 3),证明 $\|\sigma(\alpha)\| = \|\alpha\|$.

证 设 σ 在一组基下的对应矩阵是正交矩阵 T,对任意的 $\alpha, \beta \in V$,设 α 与 β 在这组基下的坐标分别是 x 和 y,那么

$$(Tx, Ty) = (Tx)^T Ty = x^T T^T Ty = x^T y = (x, y),$$

可见 $(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta)$ 即正交变换保持内积不变. 因此

$$\|Tx\|^2 = (Tx, Tx) = (x, x) = \|x\|^2,$$

也就是 $\|\sigma(\alpha)\| = \|\alpha\|$. 正交变换保持向量长度不变. ■

习 题 4

4.1 求向量 $\beta = (2, 3, 2)^T$ 在基 $\alpha_1 = (1, 1, 1)^T$, $\alpha_2 = (0, 1, 1)^T$, $\alpha_3 = (0, 0, 1)^T$ 下的坐标.

4.2 已知三维空间 R^3 的两组基为:

$\alpha_1 = (1, 1, 0)^T$, $\alpha_2 = (0, 1, 1)^T$, $\alpha_3 = (1, 0, 1)^T$,

$\beta_1 = (1, 0, 3)^T$, $\beta_2 = (1, -1, 0)^T$, $\beta_3 = (1, 2, 1)^T$,

求基 $\alpha_1, \alpha_2, \alpha_3$ 到基 $\beta_1, \beta_2, \beta_3$ 的过渡矩阵.

4.3 已知 R^3 的两组基为:

$\alpha_1 = (1, 0, -2)^T$, $\alpha_2 = (0, 1, 1)^T$, $\alpha_3 = (3, -1, 0)^T$.

$\beta_1 = (4, 0, -1)^T$, $\beta_2 = (3, 1, 2)^T$, $\beta_3 = (7, -2, -2)^T$.

求基 $\alpha_1, \alpha_2, \alpha_3$ 到基 $\beta_1, \beta_2, \beta_3$ 的过渡矩阵.

4.4 设 R^3 的两组基为:自然基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 和基 $\eta_1 = (1, 1, 1)^T$, $\eta_2 = (0, 1, 1)^T$, $\eta_3 = (0, 0, 1)^T$,求由自然基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 到基 η_1, η_2, η_3 的过渡矩阵,并求向量 $\alpha = (1, 3, 5)^T$ 在基 η_1, η_2, η_3 下的坐标.

4.5 求向量 $\alpha = (1, 2, 1, 1)^T$, $\beta = (3, -1, 0, 1)^T$ 的长度,并求 α 与 β 的内积和夹角.

4.6 求与向量 $\alpha = (1, 1, -1)$, $\beta = (1, -1, -1)$ 都正交的单位向量.

4.7 用 Schmidt 正交化方法,把下列 R^3 的基构造成为一组标准正交基

(1) $(1,1,1)^T$, $(0,1,1)^T$, $(1,0,1)^T$

(2) $(1,1,1)^T$, $(1,2,1)^T$, $(0,-1,1)^T$

(3) $(1,-1,1)^T$, $(-1,1,1)^T$, $(1,1,-1)^T$

并求向量 $\alpha=(1,-1,0)^T$ 在此标准正交基下的坐标 .

4.8 求出下面矩阵的行空间、列空间的维数和基

(1) $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix}$, (2) $\begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \end{pmatrix}$.

4.9 求出第 8 题矩阵的解空间的一组标准正交基,并验证解空间与行空间正交 .

4.10 证明全体二阶上三角矩阵构成 $M_{2\times 2}$ 的子空间,并求这子空间的维数和一组基 .

4.11 对二阶矩阵所构成的线性空间 $M_{2\times 2}$,求矩阵 $A=\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ 在基 $E_1=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_2=\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $E_3=\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $E_4=\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ 下的坐标 .

4.12 求出 4.5 例 1 中各线性变换在自然基下的对应矩阵 .

4.13 在三维向量空间,有线性变换 σ

$$\sigma\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} x+y \\ x-y \\ z \end{pmatrix}.$$

求 σ 在自然基下的对应矩阵,并求 σ 在基

$$\alpha_1=(1,1,1),\quad \alpha_2=(1,1,0),\quad \alpha_3=(1,0,0)$$

下的对应矩阵

4.14 设 σ 是向量空间 V 的正交变换,证明向量 $\sigma(\alpha)$,与

$\sigma(\beta)$的夹角等于向量 α与 β的夹角 .

4.15　设 $\alpha_1, \alpha_2, \alpha_3$ 是 R^3 的一组基, σ是线性变换, 且 $\sigma(\alpha_1) = \alpha_3$, $\sigma(\alpha_2) = \alpha_2$, $\sigma(\alpha_3) = \alpha_1$.

(1) 求 σ在基 $\alpha_1, \alpha_2, \alpha_3$ 下的对应矩阵;

(2) 如向量 α在基 $\alpha_1, \alpha_2, \alpha_3$ 下的坐标是$(2, -1, 1)^T$ 求 $\sigma(\alpha)$在这组基下的坐标;

(3) 验证 $\beta_1 = \alpha_1 + \alpha_3$, $\beta_2 = \alpha_2$, $\beta_3 = \alpha_1 - \alpha_3$ 也是 R^3 的一组基, 并求 σ在基 $\beta_1, \beta_2, \beta_3$ 下的对应矩阵 .

*4.16　已知 $\alpha_1, \alpha_2, \alpha_3$ 线性无关, β非零与每个 $\alpha_i (i = 1, 2, 3)$都正交, 证明: $\alpha_1, \alpha_2, \alpha_3, \beta$线性无关 .(提示: 反证法)

*4.17　四维向量 $\alpha_1, \alpha_2, \alpha_3$ 线性无关, 且与非零的 β_1, β_2 都正交, 证明: β_1, β_2 线性相关 .

*4.18　$\alpha_1, \alpha_2, \alpha_3$ 是欧氏空间的一组基, 且与 β都正交, 证明: $\beta = 0$.

第 5 章　特征值和特征向量

前几章的内容几乎都涉及到线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的求解，而基本技巧是高斯消元法与矩阵的初等行变换．现在新的问题是矩阵能否通过相似变换化为对角形的问题，虽然现在也还要用到行变换，但它只起辅助性的作用了．当前的关键是 n 阶矩阵的特征值与特征向量问题，这些在振动，稳定性、经济等问题上有着许多重要的应用．

5.1　特征值和特征向量

定义 5.1　设 $\boldsymbol{A}$ 是 n 阶矩阵，如果存在数 λ 及非零的 n 维向量 $\boldsymbol{X}$，使得

$$\boldsymbol{AX}=\lambda\boldsymbol{X} \tag{5.1}$$

成立，就称 λ 是矩阵 $\boldsymbol{A}$ 的特征值，$\boldsymbol{X}$ 是矩阵 $\boldsymbol{A}$ 属于特征值 λ 的一个特征向量．

对已给出的 n 阶矩阵 $\boldsymbol{A}$，怎样来求出 $\boldsymbol{A}$ 的特征值及其相对应的特征向量呢？

例如，已知二阶矩阵

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

按定义，若 λ 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{X}=(x_1,\ x_2)^{\mathrm{T}}$ 是属于 λ 的一个特征向量，那么，$\boldsymbol{AX}=\lambda\boldsymbol{X}$ 也就是

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\lambda\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

上述关系可以表示为:

$$a_{11}x_1 + a_{12}x_2 = \lambda x_1,$$
$$a_{21}x_1 + a_{22}x_2 = \lambda x_2,$$

作为未知数 λ, x_1, x_2 的方程已不是前面学过的线性方程组,因为它涉及到 λ 与 x_i 的乘积。但如果把上述方程改写成

$$\begin{aligned}(a_{11} - \lambda)x_1 + a_{12}x_2 = 0.\\ a_{21}x_1 + (a_{22} - \lambda)x_2 = 0.\end{aligned} \tag{5.1.1}$$

可以看出,当特征值 λ 已求出时,(5.1.1)就成为 x_1, x_2 的齐次线性方程组,而齐次方程组(5.1.1)的非零解 $\boldsymbol{X} = (x_1, x_2)^{\mathrm{T}}$ 正是矩阵 $\boldsymbol{A}$ 属于特征值 λ 的特征向量。由于(5.1.1)有非零解等价于它的系数行列式为零,即

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0, \tag{5.1.2}$$

这是 λ 的一个二次方程. 这样我们就可以先解方程(5.1.2),求出特征值 λ_i,然后把 λ_i 代入到齐次线性方程组(5.1.1)再求出其非零解,也就是 λ_i 所对应的特征向量.(为什么此时一定有非零解?)

对上面的分析,现在用矩阵来表示,并给出相关的术语、名词.

对 $\boldsymbol{AX} = \lambda\boldsymbol{X}$, $\boldsymbol{X} \neq 0$,即齐次线性方程组

$$(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0} \tag{5.2}$$

有非零解,因此,系数行列式为零,即

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = \boldsymbol{0} \tag{5.3}$$

称 $|\boldsymbol{A} - \lambda\boldsymbol{I}|$ 是矩阵 $\boldsymbol{A}$ 的特征多项式,对 n 阶矩阵 $\boldsymbol{A}$ 来说,其特征多项式是 λ 的 n 次多项式.称 $|\boldsymbol{A} - \lambda\boldsymbol{I}| = \boldsymbol{0}$ 是矩阵 $\boldsymbol{A}$ 的特征方程,它是 λ 的 n 次方程,因而有 n 个根,所以矩阵 $\boldsymbol{A}$ 有 n 个特征值. 对每个特征值 λ_i,解齐次方程组

$$(\boldsymbol{A} - \lambda_i\boldsymbol{I})\boldsymbol{x} = 0$$

其非零解就是 A 的特征值 λ_i 所对应的特征向量.

对于三阶矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

它的特征多项式是

$$|A-\lambda I|=\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix},$$

这是 λ 的三次多项式,从特征方程

$$|A-\lambda I|=0$$

可求出 A 的三个特征根 $\lambda_1,\lambda_2,\lambda_3$,而属于 λ_1 的特征向量就是齐次方程组 $(A-\lambda_1 I)x=\mathbf{0}$,即

$$\begin{aligned} (a_{11}-\lambda_1)x_1+a_{12}x_2+a_{13}x_3&=0,\\ a_{21}x_1+(a_{22}-\lambda_1)x_2+a_{23}x_3&=0,\\ a_{31}x_1+a_{32}x_2+(a_{33}-\lambda_1)x_3&=0, \end{aligned}$$

的非零解,类似地可求 λ_2,λ_3 所对应的特征向量.

根据齐次线性方程组解的性质,如果 X_1,X_2 都是方程组 $(A-\lambda I)x=0$ 的解,那么 X_1 与 X_2 的线性组合 $C_1X_1+C_2X_2$ 仍是这方程组的解.这样,关于矩阵的特征向量就有下述性质:

定理 5.1 如果 X_1,X_2 都是矩阵 A 的属于特征值 λ 的特征向量,那么当 $C_1X_1+C_2X_2$ 非零时,$C_1X_1+C_2X_2$ 仍是矩阵 A 属于特征值 λ 的特征向量. ■

关于特征向量,要注意的是:

(1) 如果 X_1 是矩阵 A 的属于特征值 λ_1 的特征向量,那么 X_1 就不可能再是矩阵 A 的属于特征值 $\lambda_2(\lambda_1\neq\lambda_2)$ 的特征向量。

这是因为,如果

$$\boldsymbol{A X}_1 = \lambda_1 \boldsymbol{X}_1 \quad 且 \quad \boldsymbol{A X}_1 = \lambda_2 \boldsymbol{X}_1,$$

那么 $\lambda_1 \boldsymbol{X}_1 = \lambda_2 \boldsymbol{X}_1$，得到 $(\lambda_1 - \lambda_2)\boldsymbol{X}_1 = 0$，

但由 $\lambda_1 \neq \lambda_2$ 有 $\boldsymbol{X}_1 = 0$,这与 $\boldsymbol{X}_1$ 是特征向量应当非零相矛盾 . 所以, $\boldsymbol{X}_1$ 不能再是 λ_2 所对应的特征向量 .

(2) 如果 $\boldsymbol{X}_1, \boldsymbol{X}_2$ 分别是矩阵 $\boldsymbol{A}$ 的不同特征值所对应的特征向量,那么 $\boldsymbol{X}_1 + \boldsymbol{X}_2$ 不再是 $\boldsymbol{A}$ 的特征向量(请看例 9) .

例 1 求矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ & 4 & 5 \\ & & 6 \end{pmatrix}$ 的特征值和特征向量 .

解 矩阵 $\boldsymbol{A}$ 的特征多项式为:

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} 1-\lambda & 2 & 3 \\ & 4-\lambda & 5 \\ & & 6-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda)(6-\lambda).$$

因此,由特征方程

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = (1-\lambda)(4-\lambda)(6-\lambda) = 0$$

得到矩阵 $\boldsymbol{A}$ 的特征值是:

$$\lambda_1 = 1, \ \lambda_2 = 4, \ \lambda_3 = 6.$$

当 $\lambda = 1$ 时,由齐次线性方程组 $(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{x} = 0$,即

$$\begin{pmatrix} 0 & 2 & 3 \\ 0 & 3 & 5 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

也就是

$$\begin{cases} 2x_2 + 3x_3 = 0, \\ 3x_2 + 5x_3 = 0, \\ 5x_3 = 0, \end{cases}$$

解得基础解系是: $\boldsymbol{X}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. 因此,属于 $\lambda_1 = 1$ 的全体特征向量就

是 $c_1 \boldsymbol{X}_1$**，** c_1 **取自所有非零常数**;

再将 $\lambda_2 = 4$ **代入到**$(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = 0$**,即有**$(\boldsymbol{A} - 4\boldsymbol{I})\boldsymbol{x} = 0$ **也就是**

$$\begin{cases} -3x_1+2x_2+3x_3=0, \\ 5x_3=0, \\ 2x_3=0, \end{cases}$$

求出基础解系是: $\boldsymbol{X}_2=\begin{pmatrix}2\\3\\0\end{pmatrix}$. 因此,属于 $\lambda_2=4$ 的全体特征向量是 $c_2\boldsymbol{X}_2$, c_2 为非零常数;最后,把 $\lambda_3=6$ 代入到 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=0$ 中,即 $(\boldsymbol{A}-6\boldsymbol{I})\boldsymbol{x}=0$, 也就是

$$\begin{cases} -5x_1+2x_2+3x_3=0, \\ -2x_2+5x_3=0, \\ 0=0, \end{cases}$$

其基础解系是 $X_3=\begin{pmatrix}16\\25\\10\end{pmatrix}$. 因此, $\lambda_3=6$ 所对应的特征向量是 $c_3\boldsymbol{X}_3$, c_3是非零常数. ■

从例 1 的推导求解不难看出,对于上(下)三角矩阵,对角矩阵来说,对角线上的元素就是其特征值.

例 2 求矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & -1\\2 & 4\end{pmatrix}$ 的特征值,特征向量.

解 矩阵 $\boldsymbol{A}$ 的特征多项式为

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}1-\lambda & -1\\2 & 4-\lambda\end{vmatrix}=\lambda^2-5\lambda+6,$$

解 $\boldsymbol{A}$ 的特征方程 $\lambda^2-5\lambda+6=0$, 得到矩阵 $\boldsymbol{A}$ 的特征值是 $\lambda_1=2$ 和 $\lambda_2=3$.

当 $\lambda_1=2$ 时,由 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=0$ 即

$$\begin{cases} -x_1-x_2=0 \\ 2x_1+2x_2=0 \end{cases}$$

解得 $\boldsymbol{X}_1=\begin{pmatrix}1\\-1\end{pmatrix}$. 故 $\boldsymbol{A}$ 对应于 $\lambda_1=2$ 的特征向量是 $c_1\boldsymbol{X}_1$, c_1 是任意的非零常数.

当 $\lambda_2=3$ 时, 由 $(\boldsymbol{A}-3\boldsymbol{I})\boldsymbol{x}=0$, 即

$$\begin{cases}-2x_1-x_2=0\\2x_1+x_2=0\end{cases}$$

解得 $\boldsymbol{X}_2=\begin{pmatrix}1\\-2\end{pmatrix}$. 因此, $\boldsymbol{A}$ 对应于 $\lambda_2=3$ 的特征向量是 $c_2\boldsymbol{X}_2$, c_2 是任意的非零常数. ■

例 3 求矩阵 $\boldsymbol{A}=\begin{pmatrix}1&1&0\\0&0&1\\0&0&0\end{pmatrix}$ 的特征值和特征向量.

解 矩阵 $\boldsymbol{A}$ 的特征多项式为

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}1-\lambda&1&0\\0&-\lambda&1\\0&0&-\lambda\end{vmatrix}=\lambda^2(1-\lambda),$$

由特征方程 $\lambda^2(1-\lambda)=0$, 得 $\boldsymbol{A}$ 的特征值是: $\lambda_1=1$, $\lambda_2=\lambda_3=0$.

当 $\lambda_1=1$ 时, 由 $(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{x}=0$, 即

$$\begin{pmatrix}0&1&0\\0&-1&1\\0&0&-1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

解得 $\boldsymbol{X}_1=(1,0,0)^{\mathrm{T}}$. 因此, $\boldsymbol{A}$ 对应于 $\lambda_1=1$ 的特征向量是 $c_1\boldsymbol{X}_1$, c_1 是任意的非零常数;

当 $\lambda_2=\lambda_3=0$ 时, 由 $(\boldsymbol{A}-0\boldsymbol{I})\boldsymbol{x}=0$, 即

$$\begin{pmatrix}1&1&0\\0&0&1\\0&0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

解得 $\boldsymbol{X}_2=(1,-1,0)^{\mathrm{T}}$. 因此, $\boldsymbol{A}$ 对应于 $\lambda=0$ 的特征向量是 $c_2\boldsymbol{X}_2$, c_2 是非零常数. ■

例 4 求矩阵 $\boldsymbol{A}=\begin{pmatrix}2&2&-2\\2&5&-4\\-2&-4&5\end{pmatrix}$ 的特征值及特征向量.

解 矩阵 $\boldsymbol{A}$ 的特征多项式为:

$$
\begin{aligned}
|\boldsymbol{A}-\lambda\boldsymbol{I}| &= \begin{vmatrix}2-\lambda&2&-2\\2&5-\lambda&-4\\-2&-4&5-\lambda\end{vmatrix}=\begin{vmatrix}2-\lambda&2&-2\\2&5-\lambda&-4\\0&1-\lambda&1-\lambda\end{vmatrix}\\
&=(1-\lambda)\begin{vmatrix}2-\lambda&2&-2\\2&5-\lambda&-4\\0&1&1\end{vmatrix}\\
&=(1-\lambda)\begin{vmatrix}2-\lambda&4&-2\\2&9-\lambda&-4\\0&0&1\end{vmatrix}\\
&=(1-\lambda)\begin{vmatrix}2-\lambda&4\\2&9-\lambda\end{vmatrix}\\
&=(1-\lambda)[(2-\lambda)(9-\lambda)-8]\\
&=(1-\lambda)^2(10-\lambda),
\end{aligned}
$$

由特征方程 $(1-\lambda)^2(10-\lambda)=0$, 得到 $\boldsymbol{A}$ 的特征值:

$$\lambda_1=\lambda_2=1,\ \lambda_3=10.$$

当 $\lambda_1=\lambda_2=1$ 时, 由 $(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$, 即

$$\begin{pmatrix}1&2&-2\\2&4&-4\\-2&-4&4\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

因为 $r(\boldsymbol{A}-\boldsymbol{I})=1$, 其基础解系由两个向量 $\boldsymbol{X}_1=(-2,1,0)^{\mathrm{T}}$ 和 $\boldsymbol{X}_2=(2,0,1)^{\mathrm{T}}$ 构成. 根据定理 5.1, $\boldsymbol{A}$ 对应于特征值 $\lambda_1=\lambda_2=1$

的特征向量是 $c_1\boldsymbol{X}_1+c_2\boldsymbol{X}_2$,其中 c_1, c_2 是不全为零的常数 .

当 $\lambda_3=10$ 时,由 $(\boldsymbol{A}-10\boldsymbol{I})\boldsymbol{x}=0$,即

$$\begin{pmatrix} -8 & 2 & -2 \\ 2 & -5 & -4 \\ -2 & -4 & -5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

经高斯消元,把系数矩阵化为阶梯形

$$\begin{pmatrix} 2 & -5 & -4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

可得基础解系 $\boldsymbol{X}_3=(1,2,-2)^{\mathrm{T}}$. 所以, $\boldsymbol{A}$ 关于特征值 $\lambda_3=10$ 所对应的特征向量是 $c_3\boldsymbol{X}_3$,其中 c_3 是不为零的常数 . ■

*例 5　求矩阵 $\boldsymbol{A}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ 的特征值和特征向量 .

解　矩阵 $\boldsymbol{A}$ 的特征多项式为:

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix}=\lambda^2+1,$$

由特征方程 $\lambda^2+1=0$ 得到 $\boldsymbol{A}$ 的特征值是 $\pm i$.

当 $\lambda=i$ 时,由 $(\boldsymbol{A}-i\boldsymbol{I})\boldsymbol{x}=0$,即

$$\begin{cases} -ix_1+x_2=0, \\ -x_1-ix_2=0, \end{cases}$$

解得 $\boldsymbol{X}_1=\begin{pmatrix} -i \\ 1 \end{pmatrix}$. 故 $\boldsymbol{A}$ 属于特征值 i 的特征向量是 $c_1\boldsymbol{X}_1$, c_1 是非零常数;

当 $\lambda=-i$ 时,由 $(\boldsymbol{A}+i\boldsymbol{I})\boldsymbol{x}=0$,即

$$\begin{cases} ix_1+x_2=0, \\ -x_1+ix_2=0, \end{cases}$$

解得 $\boldsymbol{X}_2=\begin{pmatrix} i \\ 1 \end{pmatrix}$. 故 $\boldsymbol{A}$ 对应于 $\lambda=-i$ 的特征向量是 $c_2\boldsymbol{X}_2$, c_2 是非

零常数 . ■

例 6 已知 $A^2 = A$, 求 A 的特征值 .

解 设 λ 是 A 的任一特征值, X 是 λ 所属的特征向量, 按定义有:

$$AX = \lambda X,$$

两边同用矩阵 A 左乘, 有

$$A^2 X = A(\lambda X) = \lambda AX = \lambda^2 X,$$

利用已知条件 $A^2 = A$, 有

$$A^2 X = AX = \lambda X,$$

所以
$$\lambda^2 X = \lambda X,$$

即有
$$(\lambda^2 - \lambda) X = 0,$$

因为 X 是特征向量, 按定义 $X \neq 0$, 故有

$$\lambda^2 - \lambda = 0.$$

所以, 矩阵 A 的特征值只能是 0 或 1. ■

注 满足 $A^2 = A$ 的矩阵 A 是不唯一的, 例如, 零矩阵, 单位矩阵都是这种类型的矩阵, 而零矩阵的特征值全是 0, 单位矩阵的特征值全是 1. 又如矩阵 $\begin{matrix}1 & \\ & 0\end{matrix}$ 亦满足 $A^2 = A$ 的条件而它的特征值是 $\lambda_1 = 1$, $\lambda_2 = 0$.

例 7 A 是三阶矩阵, 它的特征值是 $-1, 0, 4$. 又知 $A + B = 2I$, 求 B 的特征值 .

解 设矩阵 A 属于 -1 的特征向量是 X_1, 即 $AX_1 = -X_1$. 那么, 由 $A + B = 2I$ 有

$$BX_1 = (2I - A)X_1 = 2X_1 - AX_1 = 3X_1.$$

按定义, 3 是 B 的特征值, X_1 是其相应的特征向量 . 完全类似地,

$$BX_2 = (2I - A)X_2 = 2X_2 - AX_2 = 2X_2,$$

$$BX_3 = (2I - A)X_3 = 2X_3 - AX_3 = -2X_3,$$

可得到 B 的特征值是 $3, 2, -2$. ■

例 8 设 $\boldsymbol{A}$ 是可逆矩阵,证明 $\boldsymbol{A}$ 的所有特征值都不为 0.

证 设 λ 是 $\boldsymbol{A}$ 的任一特征值, $\boldsymbol{X}_0$ 是相应的特征向量,即 $\boldsymbol{A}\boldsymbol{X}_0=\lambda\boldsymbol{X}_0$.

如果 $\lambda=0$,则有 $\boldsymbol{A}\boldsymbol{X}_0=0$,这表明 $\boldsymbol{X}_0$ 是齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=0$ 的解. 因为 $\boldsymbol{X}_0$ 是特征向量,它是非零的. 于是 $\boldsymbol{A}\boldsymbol{x}=0$ 有非零解.

而 $\boldsymbol{A}\boldsymbol{x}=0$ 有非零解的充要条件是其系数行列式等于 0. 因此,$|\boldsymbol{A}|=0$. 这与 $\boldsymbol{A}$ 可逆相矛盾. 所以,$\boldsymbol{A}$ 的特征值 λ 都不为 0. ■

从定理 5.1 我们知道了矩阵 $\boldsymbol{A}$ 的同一个特征值所对应的特征向量之间的关系,现在进一步来看不同的特征值所对应的特征向量之间的关系.

定理 5.2 如果 λ_1,λ_2 是矩阵 $\boldsymbol{A}$ 的不同的特征值,其对应的特征向量分别是 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$,那么 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 线性无关.

证 考查 $k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2=0$,两边同左乘矩阵 $\boldsymbol{A}$,有:

$$\begin{aligned}\boldsymbol{A}(k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2)&=k_1\boldsymbol{A}\boldsymbol{X}_1+k_2\boldsymbol{A}\boldsymbol{X}_2\\&=k_1\lambda_1\boldsymbol{X}_1+k_2\lambda_2\boldsymbol{X}_2=0.\end{aligned}$$

若两边同乘以 λ_1 又有: $\lambda_1k_1\boldsymbol{X}_1+\lambda_1k_2\boldsymbol{X}_2=0$,

上、下两式相减得: $(\lambda_2-\lambda_1)k_2\boldsymbol{X}_2=0$,

由于 $\lambda_1\neq\lambda_2$, $\boldsymbol{X}_2$ 是特征向量,有 $\boldsymbol{X}_2\neq0$,从而有: $k_2=0$.

类似地, $k_1=0$.

因此,不同特征值所对应的特征向量 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 线性无关. ■

用数学归纳法很容易把定理 5.2 推广为:

如果 $\lambda_1,\lambda_2,\cdots,\lambda_t$ 是矩阵 $\boldsymbol{A}$ 的互不相同的特征值,其对应的特征向量分别是 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_t$,则 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_t$ 线性无关.

例 9 如果 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 分别是矩阵 $\boldsymbol{A}$ 属于不同特征值 λ_1,λ_2 的特征向量,则 $\boldsymbol{X}_1+\boldsymbol{X}_2$ 不是 $\boldsymbol{A}$ 的特征向量.

证 (用反证法) 如果 $\boldsymbol{X}_1+\boldsymbol{X}_2$ 是矩阵 $\boldsymbol{A}$ 对应于特征值 λ

的特征向量,即

$$A(X_1+X_2)=\lambda(X_1+X_2)=\lambda X_1+\lambda X_2.$$

又有 $$A(X_1+X_2)=AX_1+AX_2=\lambda_1X_1+\lambda_2X_2.$$

那么 $$(\lambda-\lambda_1)X_1+(\lambda-\lambda_2)X_2=0.$$

据定理 5.2, X_1, X_2 线性无关,所以

$$\lambda-\lambda_1=0,\ \lambda-\lambda_2=0.$$

这与 $\lambda_1\neq\lambda_2$ 相矛盾. 所以 X_1+X_2 不是 A 的特征向量. ■

在本节的最后,我们给出三阶矩阵的特征值与矩阵之间的一些联系.

***定理 5.3** 设 $\lambda_1,\lambda_2,\lambda_3$ 是矩阵

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

的特征值. 则有

$$\lambda_1+\lambda_2+\lambda_3=a_{11}+a_{22}+a_{33}\quad (\text{称为矩阵 } A \text{ 的迹})$$

$$\lambda_1\lambda_2\lambda_3=|A|. \tag{5.4}$$

证 利用行列式性质(参看 1.1 例 12)

$$|A-\lambda I|=\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13}\\ a_{21} & a_{22}-\lambda & a_{23}\\ a_{31} & a_{32} & a_{33}-\lambda\end{vmatrix}$$

$$=|A|-\left(\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}+\begin{vmatrix} a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}+\begin{vmatrix} a_{11} & a_{13}\\ a_{31} & a_{33}\end{vmatrix}\right)\lambda$$

$$+(a_{11}+a_{22}+a_{33})\lambda^2-\lambda^3,$$

又因 A 的特征值是 $\lambda_1,\lambda_2,\lambda_3$,应有

$$|A-\lambda I|=(\lambda_1-\lambda)(\lambda_2-\lambda)(\lambda_3-\lambda)$$

$$=\lambda_1\lambda_2\lambda_3-(\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_3\lambda_1)\lambda$$

$$+(\lambda_1+\lambda_2+\lambda_3)\lambda^2-\lambda^3,$$

比较特征多项式的系数,就有

$$\lambda_1+\lambda_2+\lambda_3=a_{11}+a_{22}+a_{33},$$

$$\lambda_1\lambda_2\lambda_3=|\boldsymbol{A}|. \qquad \blacksquare$$

从定理 5.3,**我们看到**

$|\boldsymbol{A}|\neq 0$　　　$\boldsymbol{A}$ **的特征值全不为** 0 .

这样,我们又有一个判断矩阵可逆的充分必要条件 .

推论 1　**矩阵 $\boldsymbol{A}$ 可逆的充分必要条件是 $\boldsymbol{A}$ 的特征值全不为** 0 .(**其证明也可如例** 8)

如 $\boldsymbol{A}$ 是 n 阶矩阵,则其特征多项式是 λ 的 n 次多项式,作为 n 次方程,当 $n\geq 3$ 时方程的根是很难求的 . 因此,对特征多项式最好像例 4 **那样,先利用行列式的性质化简 $|\boldsymbol{A}-\lambda\boldsymbol{I}|$,提取出 λ 的因子再来求根 . 而在实际问题中,一般要用近似计算的方法来求特征值 .**

5.2　相 似 矩 阵

在 5.1 **的例** 2 **中,对矩阵**

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

我们已知 $\boldsymbol{A}$ 的特征值是 2 **和** 3,**而** $\begin{pmatrix}1\\-1\end{pmatrix}$ **和** $\begin{pmatrix}1\\-2\end{pmatrix}$ **是相应的特征向量 . 即**

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

用分块矩阵可把上述结果写成

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

定理 5.2 保证了这两个特征向量线性无关,因此

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

必定是可逆矩阵 . 那么

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

易见

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

一般地,可以有

$$\begin{aligned} A^n &= \boldsymbol{P} \begin{pmatrix} 2^n & \\ & 3^n \end{pmatrix} \boldsymbol{P}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 2^n & \\ & 3^n \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2^{n+1} - 3^n & 2^n - 3^n \\ -2^{n+1} + 2.3^n & -2^n + 2.3^n \end{pmatrix}, \end{aligned}$$

从这里我们看到,如果能利用特征值和特征向量把矩阵 $\boldsymbol{A}$ 改写成 $\boldsymbol{P}^{-1}\boldsymbol{BP}$ 的形式,对矩阵 $\boldsymbol{A}$ 的计算有时是方便的,现引入相似矩阵的概念 .

定义 5.2 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 n 阶矩阵,如果存在可逆矩阵 $\boldsymbol{P}$,使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{B}, \tag{5.5}$$

则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似,记作 $\boldsymbol{A} \sim \boldsymbol{B}$.

矩阵的相似具有:

反身性: $\boldsymbol{A} \sim \boldsymbol{A}$;

对称性: 如 $\boldsymbol{A} \sim \boldsymbol{B}$,则 $\boldsymbol{B} \sim \boldsymbol{A}$;

传递性: 如 $\boldsymbol{A} \sim \boldsymbol{B}$, $\boldsymbol{B} \sim \boldsymbol{C}$,则 $\boldsymbol{A} \sim \boldsymbol{C}$.

数学上把具有反身性、对称性和传递性称为是一个等价关系,矩阵

的相似就是一个等价关系.

定理 5.4 如 $A\sim B$,则 A 与 B 有相同的特征值.

证 因为 $A\sim B$,故存在可逆矩阵 P 使得 $P^{-1}AP=B$,那么

$$
\begin{aligned}
|B-\lambda I| &= |P^{-1}AP-\lambda I| = |P^{-1}(A-\lambda I)P| \\
&= |P^{-1}|\,|A-\lambda I|\,|P| = |A-\lambda I|,
\end{aligned}
$$

这表明 A 和 B 有相同的特征多项式,所以它们有相同的特征值. ■

推论 2 如果 A 和对角矩阵 Λ

$$
\Lambda = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}
$$

相似,则 $a_1, a_2, \cdots, a_n$ 就是 A 的 n 个特征值.

证 由 $A\sim\Lambda$,按定理 5.4,有

$$
|A-\lambda I| = |\Lambda-\lambda I| = \begin{vmatrix} a_1-\lambda & & & \\ & a_2-\lambda & & \\ & & \ddots & \\ & & & a_n-\lambda \end{vmatrix}
$$

$$
= (a_1-\lambda)(a_2-\lambda)\cdots(a_n-\lambda).
$$

所以,对角元素 $a_1, a_2, \cdots, a_n$ 就是 A 的特征值. ■

注意 定理 5.4 的逆命题不成立,也就是说如 A 与 B 特征值相同,不一定有 A 与 B 相似. 例如, $A=\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$, $B=\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$, A 和 B 的特征值是相同的,但是 A 和 B 不相似. 因为

$$
\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix} = B = P^{-1}AP = P^{-1}\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}P = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}
$$

是不可能的.

例 1 如 A 可逆,证明: $AB\sim BA$.

证　由于 A 可逆,有

$$A^{-1}(AB)A=BA,$$

按相似定义知:$AB\sim BA$. ■

例 2　如 $A\sim B$,证明:$A^2\sim B^2$.

证　由于 $A\sim B$,有可逆矩阵 P 使 $P^{-1}AP=B$,那么

$$B^2=(P^{-1}AP)(P^{-1}AP)=P^{-1}A^2P,$$

所以, $A^2\sim B^2$. ■

例 3　如 $A\sim B$, $C\sim D$,证明 $\begin{pmatrix}A&\\&C\end{pmatrix}\sim\begin{pmatrix}B&\\&D\end{pmatrix}$.

证　由 $A\sim B$, $C\sim D$,知存在可逆矩阵 P 和 Q 使 $P^{-1}AP=B$, $Q^{-1}CQ=D$.

又因 $\begin{pmatrix}P&\\&Q\end{pmatrix}$ 可逆,且 $\begin{pmatrix}P&\\&Q\end{pmatrix}^{-1}=\begin{pmatrix}P^{-1}&\\&Q^{-1}\end{pmatrix}$. 于是

$$\begin{pmatrix}P&\\&Q\end{pmatrix}^{-1}\begin{pmatrix}A&\\&C\end{pmatrix}\begin{pmatrix}P&\\&Q\end{pmatrix}=\begin{pmatrix}P^{-1}&\\&Q^{-1}\end{pmatrix}\begin{pmatrix}A&\\&C\end{pmatrix}\begin{pmatrix}P&\\&Q\end{pmatrix}$$

$$=\begin{pmatrix}P^{-1}AP&\\&Q^{-1}CQ\end{pmatrix}=\begin{pmatrix}B&\\&D\end{pmatrix},$$

所以, $\begin{pmatrix}A&\\&C\end{pmatrix}\sim\begin{pmatrix}B&\\&D\end{pmatrix}$. ■

例 4　已知 $A\sim\Lambda=\begin{pmatrix}2&\\&3\end{pmatrix}$,求 $|A+I|$

解　由 $A\sim\Lambda$,故存在可逆矩阵 P 使 $P^{-1}AP=\Lambda$ 成立. 那么, $A=P\Lambda P^{-1}$,于是

$$|A+I|=|P\Lambda P^{-1}+I|=|P(\Lambda+I)P^{-1}|$$

$$=|P||\Lambda+I||P^{-1}|=|\Lambda+I|$$

$$= \begin{vmatrix} 3 & \\ & 4 \end{vmatrix} = 12 .$$ ■

*例 5 证明 $\boldsymbol{A} = \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix}$ 与 $B = \begin{pmatrix} a_2 & & \\ & a_1 & \\ & & a_3 \end{pmatrix}$ 相似 .

证 取 $\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,有 $\boldsymbol{P}^{-1} = \boldsymbol{P}$,则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & a_2 & 0 \\ a_1 & 0 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_2 & & \\ & a_1 & \\ & & a_3 \end{pmatrix} = \boldsymbol{B},$$

所以, $\boldsymbol{A} \sim \boldsymbol{B}$. ■

5.3 矩阵可对角化的条件

矩阵的相似是一个等价关系,对矩阵 $\boldsymbol{A}$ 可以有许多矩阵和它相似,现在我们关心的问题是在和 $\boldsymbol{A}$ 相似的那些矩阵中是否有对角矩阵?换句话讲 $\boldsymbol{A}$ 能否和一个对角矩阵相似?

如果矩阵 $\boldsymbol{A}$ 和一个对角矩阵相似,就称矩阵 $\boldsymbol{A}$ 可对角化 . 本节就是来讨论矩阵 $\boldsymbol{A}$ 满足什么条件时,它可对角化 .

设 $\boldsymbol{A}$ 是三阶矩阵 . 如果 $\boldsymbol{A} \sim \Lambda$,其中

$$\Lambda = \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix}$$

由相似定义,存在可逆矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\Lambda$,则有:$\boldsymbol{AP}=\boldsymbol{P}\Lambda$
对可逆矩阵 $\boldsymbol{P}$ 按列分块,我们有

$$\boldsymbol{A}(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)=\boldsymbol{AP}=\boldsymbol{P}\Lambda=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)\begin{pmatrix}a_1&&\\&a_2&\\&&a_3\end{pmatrix}$$

$$=(a_1\boldsymbol{X}_1,a_2\boldsymbol{X}_2,a_3\boldsymbol{X}_3),$$

因此,$\boldsymbol{AX}_1=a_1\boldsymbol{X}_1$, $\boldsymbol{AX}_2=a_2\boldsymbol{X}_2$, $\boldsymbol{AX}_3=a_3\boldsymbol{X}_3$.
由于 $\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)$是可逆矩阵,所以,$\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 都是非零向量,且它们是线性无关的.这样,$\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 分别是 a_1,a_2,a_3 所对应的特征向量.也就是说,如果 $\boldsymbol{A}\sim\Lambda$,则 $\boldsymbol{A}$ 有三个线性无关的特征向量.

反之,如 $\boldsymbol{A}$ 有三个线性无关的特征向量 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$,设相对应的特征值是 $\lambda_1,\lambda_2,\lambda_3$.即 $\boldsymbol{AX}_i=\lambda_i\boldsymbol{X}_i(i=1,2,3)$,那么

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)&=(\boldsymbol{AX}_1,\boldsymbol{AX}_2,\boldsymbol{AX}_3)\\&=(\lambda_1\boldsymbol{X}_1,\lambda_2\boldsymbol{X}_2,\lambda_3\boldsymbol{X}_3)\\&=(\boldsymbol{X}_1\ \boldsymbol{X}_2\ \boldsymbol{X}_3)\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix},\end{aligned}$$

因为 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 线性无关,故

$$\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)$$

是可逆矩阵,于是

$$\boldsymbol{AP}=\boldsymbol{P}\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix},\quad \text{即}\quad \boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix}.$$

所以,$\boldsymbol{A}\sim\Lambda$.

上述分析可推广到 n 阶矩阵,这就是

定理 5.5 n 阶矩阵 $\boldsymbol{A}$ 可对角化(即 $\boldsymbol{A}\sim\Lambda$)的充分必要条件

是 A 有 n 个线性无关的特征向量. ■

再利用定理5.2,可有

推论3 如 n 阶矩阵 A 有 n 个不同的特征值,则 A 可对角化,即 $A \sim \Lambda$. ■

对 n 阶矩阵 A,当 A 的特征值有重根时,A 的线性无关的特征向量的个数可能有 n 个,也可能是小于 n 个(如5.1中例4,例3),这表明不是任何 n 阶矩阵都可以对角化. 那么,当 A 的特征值有重根时,A 怎样才能对角化呢?

定理5.6 如 A 是 n 阶矩阵,λ_1 是 A 的 m 重特征值,则属于 λ_1 的线性无关的特征向量的个数不超过 m 个.(证明较难,略去) ■

我们设想 A 是一个5阶矩阵. λ_1 是其三重特征值,λ_2 是 A 的二重特征值. 如若 λ_1 只有两个线性无关的特征向量,按定理5.6,那么 λ_2 至多也只有两个线性无关的特征向量. 于是,A 最多有四个线性无关的特征向量,这样 A 就不可能对角化. 从这得到矩阵可对角化的充分必要条件:

推论4 矩阵 A 可对角化的充分必要条件是 A 的每个特征值中,线性无关的特征向量的个数恰好等于该特征值的重数. ■

例1 如 $A = \begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}$,问 A 是否与对角矩阵相似?若与对角矩阵相似,就求对角矩阵 Λ 及可逆矩阵 P,使得 $P^{-1}AP = \Lambda$.

解 先求 A 的特征多项式

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 \\ 5 & 4-\lambda \end{vmatrix} = \lambda^2 - 5\lambda - 6 = (\lambda - 6)(\lambda + 1).$$

所以,A 的特征值是 $\lambda_1 = 6$, $\lambda_2 = -1$.

由于 A 有两个不同的特征值,按推论3知 $A \sim \Lambda$.

下面求 A 的特征向量.

当 $\lambda_1 = 6$ 时,由 $(A - 6I)x = 0$,即

$$\begin{pmatrix} -5 & 2 \\ 5 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

得到特征向量 $\boldsymbol{X}_1 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$**,**

当 $\lambda_2 = -1$ **时,由**$(\boldsymbol{A} + \boldsymbol{I})\boldsymbol{x} = 0$**,即**

$$\begin{pmatrix} 2 & 2 \\ 5 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

得到特征向量 $\boldsymbol{X}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$**.**

因为 $\boldsymbol{AX}_1 = 6\boldsymbol{X}_1$**,** $\boldsymbol{AX}_2 = -\boldsymbol{X}_2$**且** $\boldsymbol{X}_1$**,** $\boldsymbol{X}_2$ **线性无关,故可令**

$$\boldsymbol{P} = (\boldsymbol{X}_1\ \ \boldsymbol{X}_2) = \begin{pmatrix} 2 & 1 \\ 5 & -1 \end{pmatrix},\ \text{有}\ \boldsymbol{P}^{-1} = \frac{1}{7}\begin{pmatrix} 1 & 1 \\ 5 & -2 \end{pmatrix},$$

这样

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \frac{1}{7}\begin{pmatrix} 1 & 1 \\ 5 & -2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 5 & -1 \end{pmatrix} = \begin{pmatrix} 6 & \\ & -1 \end{pmatrix},$$

即有

$$\boldsymbol{A} \sim \Lambda = \begin{pmatrix} 6 & \\ & -1 \end{pmatrix}. \qquad ■$$

例 2　证明:$\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$ **不可能对角化.**

证　**$\boldsymbol{A}$ **的特征多项式是

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = \begin{vmatrix} 1-\lambda & 1 \\ -1 & 3-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2,$$

所以,$\boldsymbol{A}$ **的特征值是** $\lambda = 2$**(二重根).**

当 $\lambda = 2$ **时,由**$(\boldsymbol{A} - 2\boldsymbol{I})\boldsymbol{x} = 0$**,即**

$$\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

得到 $\boldsymbol{A}$ **的线性无关的特征向量是** $\boldsymbol{X} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$**.**

由于 $\lambda=2$(二重根)只有一个线性无关的特征向量．据推论4, A 不可能对角化． ■

例 3 求可逆矩阵 P, 使

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

相似于对角矩阵 Λ.

解 $\boldsymbol{A}$ 的特征多项式为

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}-\lambda & 0 & 1\\ 0 & 1-\lambda & 0\\ 1 & 0 & -\lambda\end{vmatrix}=(1-\lambda)\begin{vmatrix}-\lambda & 1\\ 1 & -\lambda\end{vmatrix}$$

$$=(1-\lambda)(\lambda^2-1);$$

所以, $\boldsymbol{A}$ 的特征值是 $\lambda_1=\lambda_2=1$, $\lambda_3=-1$.

当 $\lambda_1=\lambda_2=1$ 时, 由 $(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{x}=0$, 即

$$\begin{pmatrix}-1 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix},$$

得到特征向量 $\boldsymbol{X}_1=\begin{pmatrix}1\\ 0\\ 1\end{pmatrix}$, $\boldsymbol{X}_2=\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}$,

当 $\lambda_3=-1$ 时, 由 $(\boldsymbol{A}+\boldsymbol{I})\boldsymbol{x}=0$, 即

$$\begin{pmatrix}1 & 0 & 1\\ 0 & 2 & 0\\ 1 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix},$$

得到特征向量 $\boldsymbol{X}_3=\begin{pmatrix}1\\ 0\\ -1\end{pmatrix}$.

$\boldsymbol{A}$ 有三个线性无关的特征向量, $\boldsymbol{A}$ 可对角化．

当取

$$\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)=\begin{pmatrix}1&0&1\\0&1&0\\1&0&-1\end{pmatrix}$$

时,我们就有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\frac{1}{2}&0&\frac{1}{2}\\0&1&0\\\frac{1}{2}&0&-\frac{1}{2}\end{pmatrix}\begin{pmatrix}&&1\\&1&\\1&&\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&0\\1&0&-1\end{pmatrix}=\begin{pmatrix}1&&\\&1&\\&&-1\end{pmatrix},$$

因此, $\boldsymbol{A}\sim\begin{pmatrix}1&&\\&1&\\&&-1\end{pmatrix}$. ■

例 4 已知 $\boldsymbol{A}=\begin{pmatrix}1&1&1\\0&0&x\\0&0&0\end{pmatrix}$ 有三个线性无关的特征向量,求 x 的值.

解 $\boldsymbol{A}$ 的特征多项式

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}1-\lambda&1&1\\0&-\lambda&x\\0&0&-\lambda\end{vmatrix}=\lambda^2(1-\lambda)$$

得到 $\boldsymbol{A}$ 的特征值是 $\lambda_1=1,\ \lambda_2=\lambda_3=0$.

由于 $\boldsymbol{A}$ 有三个线性无关的特征向量,故当 $\lambda_2=\lambda_3=0$ 时, $\boldsymbol{A}$ 应有两个线性无关的特征向量.即方程组 $(\boldsymbol{A}-0\boldsymbol{I})\boldsymbol{x}=0$ 的基础解系应由两个向量组成,那么系数矩阵的秩 $r(\boldsymbol{A}-0\boldsymbol{I})=1$,也就是 $r(\boldsymbol{A})=1$. 所以, $x=0$. ■

例 5 已知 $\boldsymbol{A}$ 的特征值是 2 和 3,相应的特征向量是 $\begin{pmatrix}1\\1\end{pmatrix}$ 和

$\begin{pmatrix}3\\2\end{pmatrix}$，求矩阵 $\boldsymbol{A}$.

解　由于特征向量是二维向量，知 $\boldsymbol{A}$ 是二阶矩阵，现在 $\boldsymbol{A}$ 有两个不同的特征值．故 A 可以对角化为

$$\boldsymbol{A} \sim \Lambda = \begin{pmatrix}2 & \\ & 3\end{pmatrix},$$

这时的可逆矩阵 $\boldsymbol{P}$ 由特征向量构成

$$\boldsymbol{P} = \begin{pmatrix}1 & 3\\ 1 & 2\end{pmatrix},$$

而 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \Lambda$，这就有

$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{P}\Lambda\boldsymbol{P}^{-1}\\ &= \begin{pmatrix}1 & 3\\ 1 & 2\end{pmatrix}\begin{pmatrix}2 & \\ & 3\end{pmatrix}\begin{pmatrix}-2 & 3\\ 1 & -1\end{pmatrix} = \begin{pmatrix}5 & -3\\ 2 & 0\end{pmatrix}.\end{aligned}$$ ■

*例 6　如果每个非零三维向量都是三阶矩阵 $\boldsymbol{A}$ 的特征向量，试证 $\boldsymbol{A}$ 是一个数量矩阵．

证　因为任何 3 维非零向量都是 $\boldsymbol{A}$ 的特征向量，所以 $\boldsymbol{A}$ 有三个线性无关的特征向量，那么 $\boldsymbol{A}$ 可化为对角形，设

$$\boldsymbol{A} \sim \begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{pmatrix},$$

其中 $\lambda_1, \lambda_2, \lambda_3$ 是 $\boldsymbol{A}$ 的特征值．

如果 $\lambda_1 \neq \lambda_2$，设 $\boldsymbol{X}_1$，$\boldsymbol{X}_2$ 分别是 $\boldsymbol{A}$ 属于 λ_1, λ_2 的特征向量，那么 $\boldsymbol{X}_1$，$\boldsymbol{X}_2$ 线性无关(定理 5.2)，所以 $\boldsymbol{X}_1 + \boldsymbol{X}_2 \neq 0$，一方面按已知条件 $\boldsymbol{X}_1 + \boldsymbol{X}_2$ 应是 $\boldsymbol{A}$ 的一个特征向量，另一方面，$\lambda_1 \neq \lambda_2$ 时，$\boldsymbol{X}_1 + \boldsymbol{X}_2$ 不能是 $\boldsymbol{A}$ 的特征向量(参看 5.1 中例 9)二者矛盾．所以，必有 $\lambda_1 = \lambda_2 = \lambda_3$．因此，$\boldsymbol{A} \sim \lambda_1 \boldsymbol{I}$．从而

$$\boldsymbol{A} = \boldsymbol{P}(\lambda_1 \boldsymbol{I})\boldsymbol{P}^{-1} = \lambda_1 \boldsymbol{I}$$

即 $\boldsymbol{A}$ 是一个数量矩阵. ■

这几个例题从不同的角度说明了矩阵对角化的判断方法,以及对能对角化的矩阵如何实现其对角化. 我们注意到:一个矩阵如果可以化为对角矩阵,那么对角矩阵的对角线上的元素就是该矩阵的特征值,而实现对角化的可逆矩阵 $\boldsymbol{P}$ 的每一列就是相对应的特征向量. 当然,在对角矩阵中,特征值排列的次序是任意的,可逆矩阵 $\boldsymbol{P}$ 中特征向量的选取方法也是不唯一的,这些是本节的核心,望读者仔细体会.

矩阵化为对角矩阵的问题在许多领域都有着广泛的应用,下面的例题介绍在计算矩阵的方幂方面的情况,在附录中还有求解线性常系数微分方程组的应用,供有兴趣有余力的读者参考.

*例 7 已知 $\boldsymbol{A} = \begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}$, 求 $\boldsymbol{A}^n$

解 先求 $\boldsymbol{A}$ 的特征值,由

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} -\lambda & -2 \\ 1 & 3-\lambda \end{vmatrix}$$

$$= \lambda^2 - 3\lambda + 2 = (\lambda - 1)(\lambda - 2),$$

得到 $\boldsymbol{A}$ 的特征值是 $\lambda_1 = 1, \lambda_2 = 2$.

再求 $\boldsymbol{A}$ 的特征向量:

当 $\lambda_1 = 1$ 时,由 $(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{x} = 0$, 得 $\boldsymbol{X}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$,

当 $\lambda_2 = 2$ 时,由 $(\boldsymbol{A} - 2\boldsymbol{I})\boldsymbol{x} = 0$, 得 $\boldsymbol{X}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

构造可逆矩阵

$$\boldsymbol{P} = (\boldsymbol{X}_1 \ \boldsymbol{X}_2) = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix},$$

则 $\boldsymbol{A} = \boldsymbol{P}\Lambda\boldsymbol{P}^{-1}$, 那么

$$\boldsymbol{A}^2 = (\boldsymbol{P}\Lambda\boldsymbol{P}^{-1})(\boldsymbol{P}\Lambda\boldsymbol{P}^{-1}) = \boldsymbol{P}\Lambda^2\boldsymbol{P}^{-1},$$

以此递推

$$
\begin{aligned}
\boldsymbol{A}^{n} &= (\boldsymbol{P}\Lambda\boldsymbol{P}^{-1})(\boldsymbol{P}\Lambda\boldsymbol{P}^{-1})\cdots(\boldsymbol{P}\Lambda\boldsymbol{P}^{-1}) \\
&= \boldsymbol{P}\Lambda^{n}\boldsymbol{P}^{-1} \\
&= \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & \\ & 2^{n} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\
&= \begin{pmatrix} 2-2^{n} & 2-2^{n+1} \\ -1+2^{n} & -1+2^{n+1} \end{pmatrix}.
\end{aligned}
$$

■

5.4　实对称矩阵的对角化

在上一节已看到有些 n 阶矩阵可以和对角矩阵相似，而也有些 n 阶矩阵不能对角化，本节专门讨论实对称矩阵，主要的结论是任一个实对称矩阵 $\boldsymbol{A}$ 一定可以和一个对角矩阵相似，并且可找到正交矩阵 $\boldsymbol{T}$，使 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 是对角矩阵．作为 n 阶实对称矩阵，通常 n 阶矩阵有关特征值、特征向量的性质它都具备，并且它还有一些特殊的性质．

定理 5.7　实对称矩阵 A 的特征值都是实数．

*证　设 λ 是 $\boldsymbol{A}$ 的任一特征值，$\boldsymbol{X}$ 是相应的特征向量．记 $\overline{\boldsymbol{X}}$ 是将 $\boldsymbol{X}$ 的所有分量都取共轭复数后所得到的向量，记 $\overline{\lambda}$ 是 λ 的共轭复数．

根据 $\boldsymbol{AX}=\lambda\boldsymbol{X}$ 及 $\overline{\boldsymbol{A}}=\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$(实对称)并利用复数运算的性质，我们取共轭

$$\boldsymbol{A}\overline{\boldsymbol{X}}=\overline{\boldsymbol{A}}\,\overline{\boldsymbol{X}}=\overline{\boldsymbol{AX}}=\overline{\lambda\boldsymbol{X}}=\overline{\lambda}\,\overline{\boldsymbol{X}},$$

对上式两端同时用 $\boldsymbol{X}^{\mathrm{T}}$ 左乘，就有

$$\overline{\lambda}\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\overline{\boldsymbol{X}}=(\boldsymbol{AX})^{\mathrm{T}}\overline{\boldsymbol{X}}=(\lambda\boldsymbol{X})^{\mathrm{T}}\overline{\boldsymbol{X}}=\lambda\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}},$$

从而得到

$$(\overline{\lambda}-\lambda)\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}}=0,$$

由于特征向量 $\boldsymbol{X}\neq 0$，故知

$$\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}} = x_1\overline{x_1} + x_2\overline{x_2} + \cdots + x_n\overline{x_n}$$
$$= \| x_1 \|^2 + \| x_2 \|^2 + \cdots + \| x_n \|^2 > 0,$$

因此必有 $\overline{\lambda} - \lambda = 0$, 即 $\lambda = \overline{\lambda}$,也就是 $\boldsymbol{A}$ 的特征值是实数 . ■

对一个实矩阵,虽然其特征多项式是实系数多项式,但它的特征值有可能是复数,相应的特征向量也就可能是复向量(如 5.1 中例 5)。但对于实对称矩阵来说,由定理 5.7 知其特征值一定是实数,那么$(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = 0$ 必是实系数方程组,就总可取实向量作为其特征向量 .

作为实对称矩阵,不同特征值所对应的特征向量不仅是线性无关的(定理 5.2),而且有进一步的结果:它们是正交的 .

定理 5.8 实对称矩阵 $\boldsymbol{A}$ 的不同特征值 λ_1, λ_2 所对应的特征向量 $\boldsymbol{X}_1, \boldsymbol{X}_2$ 是正交的 .

证 对 $\boldsymbol{A}\boldsymbol{X}_1 = \lambda_1\boldsymbol{X}_1$, $\boldsymbol{A}\boldsymbol{X}_2 = \lambda_2\boldsymbol{X}_2$,由

$$\lambda_1\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_1 = \boldsymbol{X}_2^{\mathrm{T}}(\lambda_1\boldsymbol{X}_1) = \boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}_1 = (\boldsymbol{A}\boldsymbol{X}_2)^{\mathrm{T}}\boldsymbol{X}_1$$
$$= (\lambda_2\boldsymbol{X}_2)^{\mathrm{T}}\boldsymbol{X}_1 = \lambda_2\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_1,$$

因此 $(\lambda_1 - \lambda_2)\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_1 = 0$.

由于 $\lambda_1 \neq \lambda_2$,故 $\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_1 = 0$,即$(\boldsymbol{X}_2, \boldsymbol{X}_1) = 0$,

所以,$\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 正交 . ■

例 1 求 $\boldsymbol{A} = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$ 的特征值、特征向量

解 $\boldsymbol{A}$ 的特征多项式是

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = \begin{vmatrix} 2-\lambda & 3 \\ 3 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda - 5 = (\lambda - 5)(\lambda + 1),$$

$\boldsymbol{A}$ 的特征值是 $\lambda_1 = 5$, $\lambda_2 = -1$.

当 $\lambda_1 = 5$ 时,由$(\boldsymbol{A} - 5\boldsymbol{I})\boldsymbol{x} = 0$. 即

$$\begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

得到 $\lambda_1 = 5$ 所对应的特征向量 $\boldsymbol{X}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

当 $\lambda_2 = -1$ 时,由 $(\boldsymbol{A} + \boldsymbol{I})\boldsymbol{x} = 0$,即

$$\begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

得到 $\lambda_2 = -1$ 所对应的特征向量 $\boldsymbol{X}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. ■

不难验证:$(\boldsymbol{X}_1, \boldsymbol{X}_2) = 0$,即不同特征值所对应的特征向量是正交的.

请注意,在例 1 中,如果把正交的特征向量 $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ 单位化为

$$\gamma_1 = \frac{\boldsymbol{X}_1}{\|\boldsymbol{X}_1\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \gamma_2 = \frac{\boldsymbol{X}_2}{\|\boldsymbol{X}_2\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

构造可逆矩阵

$$\boldsymbol{T} = (\gamma_1, \gamma_2) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix},$$

由于 γ_1, γ_2 仍是 $\boldsymbol{A}$ 的特征向量(定理 5.1),即

$$\boldsymbol{A}\gamma_1 = 5\gamma_1, \quad \boldsymbol{A}\gamma_2 = -\gamma_2,$$

我们就有

$$\boldsymbol{A}(\gamma_1, \gamma_2) = (\gamma_1, \gamma_2)\begin{pmatrix} 5 & \\ & -1 \end{pmatrix},$$

因此

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \Lambda = \begin{pmatrix} 5 & \\ & -1 \end{pmatrix},$$

显然,$\boldsymbol{T}\boldsymbol{T}^{\mathrm{T}} = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{T} = \boldsymbol{I}$. 所以,实对称矩阵 $\boldsymbol{A}$ 不仅能化为对角矩阵. 而且可选取可逆矩阵为正交矩阵来实现对角化.

更加一般地,对于 n 阶实对称矩阵 $\boldsymbol{A}$,如果它有 n 个不同的

特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$,根据定理 5.8,它们对应的特征向量 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ 是互相正交的 . 因此,只须把这些特征向量单位化为 $\gamma_1, \gamma_2, \cdots, \gamma_n$. 由此构造正交矩阵 $T = (\gamma_1, \gamma_2, \cdots, \gamma_n)$,就可得到

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

如果 $\boldsymbol{A}$ 的特征值 λ_i 是 k 重根,对方程组 $(\boldsymbol{A} - \lambda_i \boldsymbol{I})\boldsymbol{x} = 0$,必可求出 k 个线性无关的解 $\boldsymbol{X}_{i1}, \boldsymbol{X}_{i2}, \cdots, \boldsymbol{X}_{ik}$. 根据定理 5.1,这些特征向量的线性组合还是 λ_i 所对应的特征向量,故可用施密特正交化方法,做出互相正交的特征向量 $\beta_1, \beta_2, \cdots, \beta_k$,然后再单位化 . 把每个不同的特征值经如此处理的特征向量放在一起,按定理 5.8 它们可构造成一个正交矩阵 $\boldsymbol{T}$,同样可有

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

总之,我们可以证明如下结果:

定理 5.9 对任一 n 阶实对称矩阵 $\boldsymbol{A}$,存在正交矩阵 $\boldsymbol{T}$,使得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}, \tag{5.6}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的特征值 . ■

例 2 求正交矩阵 $\boldsymbol{T}$,化

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

为对角形 .

解　A 的特征多项式

$$|A-\lambda I|=\begin{vmatrix}1-\lambda & 1 & 1\\ 1 & 1-\lambda & 1\\ 1 & 1 & 1-\lambda\end{vmatrix}$$

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

$$=(3-\lambda)\begin{vmatrix}1 & 1 & 1\\ 1 & 1-\lambda & 1\\ 1 & 1 & 1-\lambda\end{vmatrix}=(3-\lambda)\begin{vmatrix}1 & 1 & 1\\ & -\lambda & \\ & & -\lambda\end{vmatrix}$$

$$=\lambda^2(3-\lambda),$$

得到 A 的特征值: $\lambda_1=3,\ \lambda_2=\lambda_3=0$.

当 $\lambda_1=3$ 时,由 $(A-3I)x=0$,即

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

得特征向量 $X_1=(1,1,1)^{\mathrm{T}}$.

当 $\lambda_2=\lambda_3=0$ 时,由 $(A-0I)x=0$,即

$$\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix},$$

得特征向量 $X_2=(-1,1,0)^{\mathrm{T}}$, $X_3=(-1,0,1)^{\mathrm{T}}$.

易见 X_1 与 X_2, X_3 都正交(不同特征值!),但 X_2 与 X_3 不正交,现对 X_2, X_3 用 Schmidt 正交化方法

令 $\beta_2=X_2=(-1,1,0)^{\mathrm{T}}$, $\beta_3=k\beta_2+X_3$.

由　$(\beta_2,\beta_3)=(\beta_2,k\beta_2+X_3)=k(\beta_2,\beta_2)+(\beta_2,X_3)=0$,

即 $\qquad 2k+1=0$

得 $\quad \beta_3 = -\frac{1}{2}\beta_2 + \boldsymbol{X}_3 = \frac{1}{2}(-1, -1, 2)^{\mathrm{T}}$,

再把 $\boldsymbol{X}_1, \beta_2, \beta_3$ 单位化,得 $\gamma_1, \gamma_2, \gamma_3$,即

$$\gamma_1 = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \gamma_2 = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \gamma_3 = \frac{1}{6}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix},$$

构造正交矩阵

$$\boldsymbol{T} = (\gamma_1 \ \ \gamma_2 \ \ \gamma_3) = \begin{pmatrix} \frac{1}{3} & \frac{-1}{2} & \frac{-1}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{-1}{6} \\ \frac{1}{3} & 0 & \frac{2}{6} \end{pmatrix},$$

我们就有

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} 3 & & \\ & 0 & \\ & & 0 \end{pmatrix}.$$ ■

*例 3 $\quad \boldsymbol{A}$ 是三阶实对称矩阵,2,1,1 是其特征值,对应于 $\lambda = 2$ 的特征向量为 $\boldsymbol{X}_1 = (2, 1, 2)^{\mathrm{T}}$,求 $\boldsymbol{A}$.

解法 1 $\quad A$ 是实对称矩阵它必可对角化,因而对应于 $\lambda = 1$, $\boldsymbol{A}$ 有两个线性无关的特征向量且它们与 $\boldsymbol{X}_1$ 都正交(定理 5.8).设$(x_1, x_2, x_3)^{\mathrm{T}}$ 是 $\lambda = 1$ 所对应的特征向量,那么

$$2x_1 + x_2 + 2x_3 = 0,$$

求出特征向量

$$\boldsymbol{X}_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \quad \boldsymbol{X}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},$$

构造可逆矩阵 $\boldsymbol{P}=(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3)$，则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\Lambda,$$

即

$$\boldsymbol{A}=\boldsymbol{P}\Lambda\boldsymbol{P}^{-1}=\begin{pmatrix}2&-1&-1\\1&2&0\\2&0&1\end{pmatrix}\begin{pmatrix}2&&\\&1&\\&&1\end{pmatrix}\cdot\frac{1}{9}\begin{pmatrix}2&1&2\\-1&4&-1\\-4&-2&5\end{pmatrix}$$

$$=\frac{1}{9}\begin{pmatrix}13&2&4\\2&10&2\\4&2&13\end{pmatrix}.$$ ■

解法 2 对 $\boldsymbol{A}$ 我们可用正交矩阵来实现对角化．只须对解法一中的 $\boldsymbol{X}_2, \boldsymbol{X}_3$ 用 Schmidt 正交化，令 $\beta_2=\boldsymbol{X}_2$，$\beta_3=k\beta_2+\boldsymbol{X}_3$

由

$$(\beta_2, \beta_3)=(\beta_2, k\beta_2+\boldsymbol{X}_3)=k(\beta_2, \beta_2)+(\beta_2, \boldsymbol{X}_3)=0$$

得

$$5k+1=0,$$

有

$$\beta_3=-\frac{1}{5}\beta_2+\boldsymbol{X}_3=\frac{1}{5}(-4, -2, 5)^{\mathrm{T}},$$

再单位化 $\boldsymbol{X}_1, \beta_2, \beta_3$，构造正交矩阵

$$\boldsymbol{T}=\begin{pmatrix}\frac{2}{3}&\frac{-1}{\sqrt{5}}&\frac{-4}{3\sqrt{5}}\\[2pt]\frac{1}{3}&\frac{2}{\sqrt{5}}&\frac{-2}{3\sqrt{5}}\\[2pt]\frac{2}{3}&0&\frac{5}{3\sqrt{5}}\end{pmatrix},$$

这时，$\boldsymbol{T}^{-1}=\boldsymbol{T}^{\mathrm{T}}$，

$$\boldsymbol{A}=\boldsymbol{T}\Lambda\boldsymbol{T}^{-1}$$

$$= \begin{pmatrix} \frac{2}{3} & \frac{-1}{5} & \frac{-4}{3\ 5} \\ \frac{1}{3} & \frac{2}{5} & \frac{-2}{3\ 5} \\ \frac{2}{3} & 0 & \frac{5}{3\ 5} \end{pmatrix} \begin{pmatrix} 2 & & \\ & 1 & \\ & & 1 \end{pmatrix}$$

$$\times \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{-1}{5} & \frac{2}{5} & 0 \\ \frac{-4}{3\ 5} & \frac{-2}{3\ 5} & \frac{5}{3\ 5} \end{pmatrix}$$

$$= \frac{1}{9} \begin{pmatrix} 13 & 2 & 4 \\ 2 & 10 & 2 \\ 4 & 2 & 13 \end{pmatrix} . \qquad ■$$

*例 4　A 是三阶实对称矩阵, $A^2 = A$, 秩 $r(\boldsymbol{A}) = 2$, 求 $|\boldsymbol{A} + 3\boldsymbol{I}|$

解　由 $\boldsymbol{A}^2 = \boldsymbol{A}$ 知 $\boldsymbol{A}$ 的特征值是 1 和 0(参看 5.1 例 6). 因 $\boldsymbol{A}$ 是实对称矩阵, $\boldsymbol{A}$ 可对角化, 故 $\boldsymbol{A}$ 有三个线性无关的特征向量.

对 $\lambda = 0$, 由 $r(\boldsymbol{A} - 0\boldsymbol{I}) = r(\boldsymbol{A}) = 2$, 知 $\lambda = 0$ 有一个线性无关的特征向量. 所以, $\lambda = 0$ 是 $\boldsymbol{A}$ 的特征值单根. 那么, $\lambda = 1$ 就必是 $\boldsymbol{A}$ 的二重特征根. 根据定理 5.9, 存在正交矩阵 $\boldsymbol{T}$ 使得:

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

那么

$$\boldsymbol{A} + 3\boldsymbol{I} = \boldsymbol{T} \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix} \boldsymbol{T}^{-1} + 3\boldsymbol{I}$$

$$= \boldsymbol{T}\begin{pmatrix}0 & & \\ & 1 & \\ & & 1\end{pmatrix}\boldsymbol{T}^{-1} + \boldsymbol{T}(3\boldsymbol{I})\boldsymbol{T}^{-1}$$

$$= \boldsymbol{T}\left(\begin{pmatrix}0 & & \\ & 1 & \\ & & 1\end{pmatrix} + \begin{pmatrix}3 & & \\ & 3 & \\ & & 3\end{pmatrix}\right)\boldsymbol{T}^{-1}$$

$$= \boldsymbol{T}\begin{pmatrix}3 & & \\ & 4 & \\ & & 4\end{pmatrix}\boldsymbol{T}^{-1},$$

所以

$$|\boldsymbol{A} + 3\boldsymbol{I}| = |\boldsymbol{T}|\begin{vmatrix}3 & & \\ & 4 & \\ & & 4\end{vmatrix}|\boldsymbol{T}^{-1}| = 48. \qquad \blacksquare$$

习　题　5

5.1　求矩阵的特征值、特征向量

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

(4) $\begin{pmatrix}1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1\end{pmatrix}$,(5) $\begin{pmatrix}0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0\end{pmatrix}$,　(6) $\begin{pmatrix}2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2\end{pmatrix}$.

5.2　求矩阵 $\boldsymbol{A}$ 的特征值

(1) 如 $\boldsymbol{A}^2 = \boldsymbol{I}$;

(2) 如 $\boldsymbol{A} + 2\boldsymbol{I} = \boldsymbol{B}$, 而 $\boldsymbol{B}$ 的特征值是 5;

(3) 如 $\boldsymbol{A}^2 + 2\boldsymbol{A} - 3\boldsymbol{I} = 0$.

5.3　如果 $\boldsymbol{A}$ 是可逆矩阵，试讨论 $\boldsymbol{A}$, $\boldsymbol{A}^*$, $\boldsymbol{A}^{-1}$ 三个矩阵特征

值之间的关系(提示:用定义,注意 $A^{-1}=\frac{1}{|A|}A^*$).

5.4 证明: A 和 A^T 有相同的特征值.

5.5 已知 $A\sim B$,证明 A 可逆的充分必要条件是 B 可逆.

5.6 选择题

(1) 三阶矩阵 A 的特征值是 1, 2, 3,相应的特征向量依次是 X_1, X_2, X_3,设 $P=(X_3, X_2, X_1)$. 则 $P^{-1}AP=$____

(A) $\begin{matrix}1&&\\&2&\\&&3\end{matrix}$, (B) $\begin{matrix}2&&\\&3&\\&&1\end{matrix}$, (C) $\begin{matrix}3&&\\&1&\\&&2\end{matrix}$, (D) $\begin{matrix}3&&\\&2&\\&&1\end{matrix}$.

(2) 已知 A 可逆, $\lambda=2$ 是 A 的一个特征值,则矩阵 $(\frac{1}{3}A^2)^{-1}$ 有一个特征值等于____

(A) $\frac{4}{3}$, (B) $\frac{3}{4}$, (C) $\frac{1}{2}$, (D) $\frac{1}{4}$.

(3) 已知三阶矩阵 A 的特征值是 $0, \pm 1$,则下列命题中不正确的是____

(A) 矩阵 A 是不可逆的,

(B) A 和对角矩阵相似,

(C) 1 和 -1 所对应的特征向量是正交的,

(D) $AX=0$ 的基础解系由一个向量组成.

(4) A、B 都是 n 阶实对称可逆矩阵,下列命题中不正确的是____

(A) 存在可逆矩阵 P,使 $P^{-1}(A+B)P=\Lambda$,

(B) 存在正交矩阵 T,使 $T^{-1}(A+B)T=\Lambda$,

(C) A^*+B^* 必可以对角化,

(D) AB 必可以对角化.

5.7 已知 $A\sim B$, $A^2=0$,证明: $B^2=0$.

5.8 已知 $A\sim B$,其中

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

求 x.

5.9　在 1 题中,哪些矩阵可对角化?对可对角化的矩阵求出与其相似的对角矩阵,并写出可逆矩阵 $\boldsymbol{P}$.

5.10　已知 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\Lambda_1$, $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}=\Lambda_2$,其中 Λ_1, Λ_2 是对角矩阵,证明:$\boldsymbol{A}\boldsymbol{B}=\boldsymbol{B}\boldsymbol{A}$.

5.11　$\boldsymbol{A}$ 是 2 阶矩阵,特征值是 2 和 -2,对应的特征向量是 $(1,0)^{\mathrm{T}}$ 和 $(2,1)^{\mathrm{T}}$,求矩阵 $\boldsymbol{A}$.

5.12　已知

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

问 $\boldsymbol{A}$、$\boldsymbol{B}$ 是否相似?为什么?

5.13　求正交矩阵 $\boldsymbol{T}$ 和对角矩阵 Λ,使 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\Lambda$

(1) $\boldsymbol{A}=\begin{pmatrix}1 & 2\\ 2 & -2\end{pmatrix}$, (2) $\boldsymbol{A}=\begin{pmatrix} & & 1\\ & 1 & \\ 1 & & \end{pmatrix}$,

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

5.14　已知 $\boldsymbol{A}\sim\begin{pmatrix}1 & \\ & 0\end{pmatrix}$,求 $|\boldsymbol{A}+2\boldsymbol{I}|$.

* **5.15　已知 $\boldsymbol{A}=\begin{pmatrix}2 & 2\\ 4 & 4\end{pmatrix}$,求 $\boldsymbol{A}^n$.(用对角化及 $\mathrm{r}(\boldsymbol{A})=1$ 分别求解,可参看 3.4 中例 3)**

5.16　已知

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

有三个线性无关的特征向量,求参数 t.

*5.17　已知三阶实对称矩阵 $\boldsymbol{A}$ 的特征值是 $\lambda_1=-1$, $\lambda_2=\lambda_3=1$,对应于 λ_1 的特征向量为 $\boldsymbol{x}_1=(0,1,1)^{\mathrm{T}}$,求 $\boldsymbol{A}$.

5.18　三阶矩阵 $\boldsymbol{A}$ 有二重特征值 λ_0,问向量 $\alpha_1=(1,0,1)^{\mathrm{T}}$, $\alpha_2=(1,1,1)^{\mathrm{T}}$, $\alpha_3=(2,-1,2)^{\mathrm{T}}$, $\alpha_4=(0,1,-1)^{\mathrm{T}}$ 能否都是矩阵 $\boldsymbol{A}$ 的属于特征值 λ_0 的特征向量,为什么?

*5.19　三阶矩阵 $\boldsymbol{A}$ 的特征值是 $1,2,-3$, $\boldsymbol{B}=\boldsymbol{A}^2+3\boldsymbol{A}-2\boldsymbol{I}$,求矩阵 $\boldsymbol{B}$ 的特征值, $\boldsymbol{B}$ 能否化为对角形,并说明理由.

*5.20　$\boldsymbol{A}$ 是三阶实对称矩阵, $\boldsymbol{A}^2=\boldsymbol{I}$, 秩 $\mathrm{r}(\boldsymbol{A}-\boldsymbol{I})=1$, 求行列式 $|\boldsymbol{A}-2\boldsymbol{I}|$ 的值.(提示:参看 5.4 例 4)

第 6 章　二　次　型

二次型就是二次齐次多项式．在解析几何中，对二次曲线

$$ax^2 + 2bxy + cy^2 = 1,$$

我们总可通过直角坐标变换把方程化为不含混合项 xy 的标准方程

$$a'x'^2 + b'y'^2 = 1$$

二次齐次式不仅在几何问题中出现，对于二元函数 $z = f(x, y)$ 极值问题的研究就涉及到形如

$$\frac{1}{2!}[f''_{xx}(\Delta x)^2 + 2f''_{xy}\Delta x\Delta y + f''_{yy}(\Delta y)^2]$$

的二次齐次式．类似的问题在数学的其它领域，在物理、力学等学科亦是经常会遇到的．

从代数的观点来看，二次型可以用一个实对称矩阵来表示，而实对称矩阵是可以用正交矩阵化为对角形的，其几何意义就是二次曲线的方程经适当的直角坐标变换可化为标准方程．在上一章，已经讨论了矩阵的特征值，但并未关心到特征值的正负号，现在我们要研究实对称矩阵特征值的符号，引入正定矩阵的概念，它的背景就是函数的极值．

6.1　二次型的矩阵表示

定义 6.1　含有 n 个变量 $x_1, x_2, \cdots, x_n$ 的二次齐次函数

$$\begin{aligned} &f(x_1, x_2, \cdots, x_n) \\ &= a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 \end{aligned}$$

$$+\cdots+2a_{1n}x_1x_n+2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n$$
$$+\cdots+2a_{n-1\,n}x_{n-1}x_n$$

如果系数 a_{ij}都是实数,就称 $f(x_1,x_2,\cdots,x_n)$是实数域上的 n 元二次型.

在二次型中,我们规定

$$a_{ij}=a_{ji},\quad \forall\ i,j=1,2,\cdots,n$$

则二次型可用实对称矩阵唯一地表出,例如,对于三元二次型

$$f(x_1,x_2,x_3)=a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2+2a_{12}x_1x_2$$
$$+2a_{13}x_1x_3+2a_{23}x_2x_3$$

可用代数恒等变形把 f(x_1,x_2,x_3)改写成

$$\begin{aligned}f(x_1,x_2,x_3)&=a_{11}x_1^2+a_{12}x_1x_2+a_{13}x_1x_3\\&\quad+a_{21}x_1x_2+a_{22}x_2^2+a_{23}x_2x_3\\&\quad+a_{31}x_1x_3+a_{32}x_2x_3+a_{33}x_3^2\\&=x_1(a_{11}x_1+a_{12}x_2+a_{13}x_3)\\&\quad+x_2(a_{21}x_1+a_{22}x_2+a_{23}x_3)\\&\quad+x_3(a_{31}x_1+a_{32}x_2+a_{33}x_3),\end{aligned}$$

再引进矩阵的运算,我们就有

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

如果记

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},\quad x=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

那么二次型 f 就可用矩阵简记为

$$f(x_1, x_2, x_3) = x^{T} A x, \tag{6.1}$$

其中矩阵 A 是一个实对称矩阵,叫做二次型 $f(x_1, x_2, x_3)$ 的对应矩阵.称(6.1)是二次型的矩阵表示.

要注意的是,给定二次型后,对应矩阵 A 的元素 a_{ii} 正是二次型中平方项 x_i^2 的系数,而 $a_{ij}(i\neq j)$ 恰是混合项 $x_i x_j$ 系数的一半.因此,二次型和它的对应矩阵是相互唯一决定的.

例如,二次型 $f(x_1, x_2, x_3) = x_1^2 - 3x_2^2 + 4x_3^2 + 2x_1x_2 + 6x_2x_3$ 的对应矩阵是

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -3 & 3 \\ 0 & 3 & 4 \end{pmatrix},$$

反之,若二次型的对应矩阵是

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 4 \\ -1 & 4 & 0 \end{pmatrix},$$

则这个二次型是

$$f(x_1, x_2, x_3) = x_1^2 + 3x_2^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3,$$

对二次型的矩阵表示 $x^{T}Ax$,也可以理解为:在 n 维向量空间有一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$,向量 α 在这组基下的坐标是 $x=(x_1, x_2, \cdots, x_n)^{T}$,那么二次型 $f(x_1, x_2, \cdots, x_n) = x^{T}Ax$ 就是向量 α 的 n 个坐标 $x_1, x_2, \cdots, x_n$ 的一个 n 元二次齐次函数.二次型的矩阵 A 是依赖于这组基的选取的.

下面以三元二次型为例,来研究当基向量改变后,二次型的对

应矩阵是如何改变的？对应矩阵之间又有什么样的联系？

设 $\alpha_1,\alpha_2,\alpha_3$ 和 β_1,β_2,β_3 是三维向量空间的两组基，由基 $\alpha_1,\alpha_2,\alpha_3$ 到基 β_1,β_2,β_3 的过渡矩阵是 $\boldsymbol{C}$（注：$\boldsymbol{C}$ 是可逆矩阵，见(4.1)）

$$(\beta_1\ \ \beta_2\ \ \beta_3)=(\alpha_1\ \ \alpha_2\ \ \alpha_3)\ \boldsymbol{C}$$

如果向量 α 在这两组基下的坐标分别是 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}$ 和 $\boldsymbol{y}=(y_1,y_2,y_3)^{\mathrm{T}}$，那么把坐标变换

$$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$$

（参看 4.1 中(4.3)）代入到变量 $\boldsymbol{x}$ 的二次型中，有

$$\begin{aligned} f(x_1\ \ x_2\ \ x_3) &= \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \\ &= (\boldsymbol{C}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{C}\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}, \end{aligned}$$

其中　$\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$.

由于 $\boldsymbol{A}$ 是对称矩阵，用转置的性质得到

$$\boldsymbol{B}^{\mathrm{T}}=(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})^{\mathrm{T}}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B},$$

$\boldsymbol{B}$ 是一个对称矩阵．因而 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 是二次型 f 的矩阵表示．

这说明，变量 x_1,x_2,x_3 的三元二次齐次式 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 后成为变量 y_1,y_2,y_3 的三元二次齐次式 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$，二次型的对应矩阵亦由基 $\alpha_1,\alpha_2,\alpha_3$ 下的 $\boldsymbol{A}$ 转化为基 β_1,β_2,β_3 下的矩阵 $\boldsymbol{B}$，且 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$，一般地

定理 6.1　变量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 的 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 后，成为变量 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$ 的 n 元二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$，其中 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$. ■

二次型经坐标变换后仍是二次型，现在把二次型矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 之间的这种关系抽象出来，引出矩阵间合同的概念．

定义 6.2　两个实对称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，如存在可逆矩阵 $\boldsymbol{C}$，使得

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B},\tag{6.2}$$

就称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 合同，记作 $\boldsymbol{A}\simeq\boldsymbol{B}$.

容易验证,矩阵的合同与矩阵的相似一样,合同也具有反身性,对称性和传递性,合同是实对称矩阵之间的一种等价关系 .

例 1 证明 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同,其中

$$\boldsymbol{A} = \begin{pmatrix} a_1 & \\ & a_2 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} a_2 & \\ & a_1 \end{pmatrix}.$$

证 取可逆矩阵 $\boldsymbol{C} = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$,则

$$\boldsymbol{C}^{\mathrm{T}} \boldsymbol{AC} = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} a_1 & \\ & a_2 \end{pmatrix} \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$$

$$= \begin{pmatrix} & a_2 \\ a_1 & \end{pmatrix} \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} = \begin{pmatrix} a_2 & \\ & a_1 \end{pmatrix} = \boldsymbol{B},$$

所以, $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同 . ■

* **例 2** 证明在实数范围内 $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ 与 $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ 不合同 .

证 (用反证法)如果这两个矩阵合同,则存在可逆矩阵 $\boldsymbol{C}$,使得

$$\boldsymbol{C}^{\mathrm{T}} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \boldsymbol{C} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix},$$

两边取行列式,并利用行列式乘法公式(定理 2.1),有

$$|\boldsymbol{C}^{\mathrm{T}}| \begin{vmatrix} 1 & \\ & -1 \end{vmatrix} |\boldsymbol{C}| = \begin{vmatrix} 1 & \\ & 1 \end{vmatrix},$$

由于 $|\boldsymbol{C}^{\mathrm{T}}| = |\boldsymbol{C}|$, 得 $|\boldsymbol{C}|^2 = -1$.
因为矩阵 $\boldsymbol{C}$ 中元素是实数,行列式 $|\boldsymbol{C}|$ 是实数,那么 $|\boldsymbol{C}|^2 > 0$,这与 $|\boldsymbol{C}|^2 = -1$ 相矛盾 . 所以, $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ 与 $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ 不合同 . ■

例 3 证明 $\begin{pmatrix} 2 & \\ & 3 \end{pmatrix}$ 与 $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ 合同但不相似 .

证　取可逆矩阵 $\boldsymbol{C}=\begin{pmatrix}\frac{1}{2} & \\ & \frac{1}{3}\end{pmatrix}$,由于

$$\boldsymbol{C}^{\mathrm{T}}\begin{pmatrix}2 & \\ & 3\end{pmatrix}\boldsymbol{C}=\begin{pmatrix}\frac{1}{2} & \\ & \frac{1}{3}\end{pmatrix}\begin{pmatrix}2 & \\ & 3\end{pmatrix}\begin{pmatrix}\frac{1}{2} & \\ & \frac{1}{3}\end{pmatrix}=\begin{pmatrix}1 & \\ & 1\end{pmatrix},$$

所以 $\begin{pmatrix}2 & \\ & 3\end{pmatrix}$ D $\begin{pmatrix}1 & \\ & 1\end{pmatrix}$.

但矩阵 $\begin{pmatrix}2 & \\ & 3\end{pmatrix}$ 的特征值是 2 和 3,矩阵 $\begin{pmatrix}1 & \\ & 1\end{pmatrix}$ 的特征值是 1 (二重根),两个矩阵的特征值不同,根据第 5 章定理 5.4,矩阵相似的必要条件是有相同的特征值,所以这两个矩阵不可能相似 .■

例 4　$\boldsymbol{A}$ 是三阶实对称矩阵,如果对任何三维列向量 $\boldsymbol{X}$ 都有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=0$,证明:$\boldsymbol{A}=0$.

证　令 $\boldsymbol{x}=(1,0,0)^{\mathrm{T}}$,我们有

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(1\ \ 0\ \ 0)\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}=a_{11},$$

又因," $\boldsymbol{x}$ 有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=0$,所以 $a_{11}=0$.

类似取 $\boldsymbol{x}$ 为 $(0,1,0)^{\mathrm{T}}$ 和 $(0,0,1)^{\mathrm{T}}$ 可得到 $a_{22}=a_{33}=0$;

再取 $\boldsymbol{x}=(1,1,0)^{\mathrm{T}}$. 由于

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(1\ \ 1\ \ 0)\begin{pmatrix}0 & a_{12} & a_{13}\\ a_{21} & 0 & a_{23}\\ a_{31} & a_{32} & 0\end{pmatrix}\begin{pmatrix}1\\1\\0\end{pmatrix}=a_{21}+a_{12}=0.$$

又因 $a_{12}=a_{21}$,所以 $a_{12}=a_{21}=0$.

同样地,$\boldsymbol{x}$ 为 $(1,0,1)^{\mathrm{T}}$ 及 $(0,1,1)^{\mathrm{T}}$ 可得 $a_{13}=a_{31}=0$, $a_{23}=a_{32}=0$.

因为 $a_{ij}\equiv 0\ (i,j=1,2,3)$，故 $\boldsymbol{A}=0$. ■

如果二次型中只含有变量的平方项,所有混合项 $x_i x_j(i\neq j)$ 的系数全是零,即

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2. \tag{6.3}$$

这样的二次型称为标准形.

一个二次型如是标准形,它的二次型矩阵就是对角形.对(6.3)可用对角矩阵表示成

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\begin{pmatrix}d_1&&&\\&d_2&&\\&&\ddots&\\&&&d_n\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.$$

定理6.1说明,经坐标变换后二次型的两个矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是合同的.那么,现在感兴趣的问题是:对任给的二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是否存在好的坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$,使变换后的二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 为标准形呢?换句话说,任给一个实对称矩阵 $\boldsymbol{A}$,它能否合同于一个对角矩阵呢?

答案是肯定的.并且可以用多种方法把二次型化成标准形.

6.2 用配方法化二次型为标准形

本节的化标准形方法是用初等代数中的配平方法.下面通过例题介绍这种方法:

例1 求坐标变换化二次型

$$f(x_1,x_2,x_3)=2x_1^2+3x_2^2+8x_3^2+4x_1x_2-8x_2x_3-4x_3x_1,$$

为标准形.

解 先把所有含 x_1 的项配成一个完全平方,即有

$$f=2(x_1+x_2-x_3)^2+x_2^2+6x_3^2-4x_2x_3,$$

再把所有含 x_2 的项配成一个完全平方.即

$$f=2(x_1+x_2-x_3)^2+(x_2-2x_3)^2+2x_3^2,$$

引入新变量,令

$$\begin{aligned} y_1 &= x_1+x_2-x_3,\\ y_2 &= x_2-2x_3,\\ y_3 &= x_3,\end{aligned}$$

则有 f 的标准形

$$f=2y_1^2+y_2^2+2y_3^2$$

相应的坐标变换公式,很容易由 $\boldsymbol{y}$ 与 $\boldsymbol{x}$ 的关系中求出

$$\begin{aligned} x_1 &= y_1-y_2-y_3,\\ x_2 &= y_2+2y_3,\\ x_3 &= y_3,\end{aligned}$$

用矩阵可表示为 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$,其中

$$\boldsymbol{C}=\begin{pmatrix}1&-1&-1\\0&1&2\\0&0&1\end{pmatrix},$$

$\boldsymbol{C}$ 是可逆的上三角矩阵 .

注:(1) 从矩阵的角度来说,例 1 表示实对称矩阵

$$\boldsymbol{A}=\begin{pmatrix}2&2&-2\\2&3&-4\\-2&-4&8\end{pmatrix}$$ 合同于对角矩阵 $$\boldsymbol{B}=\begin{pmatrix}2&&\\&1&\\&&2\end{pmatrix}.$$

(2) 配方时,我们强调一次要把一种字母配完,这必能保证矩阵 C 是上三角矩阵且是可逆的,确保 $x=Cy$ 是一坐标变换 .

二次型不含平方项时,可先作一个坐标变换构造出平方项,然后再用例 1 的方法 .

例 2 求坐标变换 $x=Cy$,化二次型

$$f(x_1,x_2,x_3)=2x_1x_2-4x_2x_3$$

为标准形 .

解　我们先令

$$\begin{aligned} x_1 &= y_1 + y_2, \\ x_2 &= y_1 - y_2, \\ x_3 &= y_3, \end{aligned}$$

代入到二次型中,得

$$\begin{aligned} f(x_1, x_2, x_3) &= 2(y_1 + y_2)(y_1 - y_2) - 4(y_1 - y_2) y_3 \\ &= 2y_1^2 - 2y_2^2 - 4y_1 y_3 + 4y_2 y_3 \\ &= 2(y_1 - y_3)^2 - 2y_2^2 + 4y_2 y_3 - 2y_3^2 \\ &= 2(y_1 - y_3)^2 - 2(y_2 - y_3)^2, \end{aligned}$$

再令

$$\begin{aligned} z_1 &= y_1 - y_3, \\ z_2 &= y_2 - y_3, \\ z_3 &= y_3, \end{aligned}$$

就得到二次型的标准形: $f = 2z_1^2 - 2z_2^2$.

而所用的坐标变换是:

$$\begin{aligned} x_1 &= y_1 + y_2 &= z_1 + z_2 + 2z_3, \\ x_2 &= y_1 - y_2 &= z_1 - z_2, \\ x_3 &= y_3 &= z_3, \end{aligned}$$

用矩阵可表示为,由

$$\boldsymbol{x} = \boldsymbol{C}_1 \boldsymbol{y}, \quad \boldsymbol{y} = \boldsymbol{C}_2 \boldsymbol{z}$$

得坐标变换: $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{z}$, $\boldsymbol{C} = \boldsymbol{C}_1 \boldsymbol{C}_2$. **其中**

$$\boldsymbol{C}_1 = \begin{pmatrix} 1 & 1 & \\ 1 & -1 & \\ & & 1 \end{pmatrix}, \quad \boldsymbol{C}_2 = \begin{pmatrix} 1 & & 1 \\ & 1 & 1 \\ & & 1 \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad ■$$

通过上述例题可以看出,对任何二次型都可用这些方法化其为平方项的代数和. 这就是:

定理 6.2 任意的 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 都可以通过坐标变换化成标准形

$$d_1 y_1^2 + d_2 y_2^2 + \cdots + d_n y_n^2, \tag{6.3}$$

其中 $d_i(i=1,2,\cdots,n)$ 是实数.(证略) ■

定理 6.2 的等价说法是:

定理 6.3 任一 n 阶实对称矩阵 $\boldsymbol{A}$,总可以合同于一个对角矩阵

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C} = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}. \tag{6.4}$$ ■

一个二次型可以用多种方法化其为标准形,在标准形中正平方项的个数及负平方项的个数与所选用的坐标变换是无关的,这就是著名的惯性定理,其证明从略.

定理 6.4 (惯性定理)对于一个二次型,不论选取怎样的坐标变换使它化为仅含平方项的标准形,其中正平方项的个数 p,负平方项的个数 q 都是由所给二次型唯一确定的. ■

通常,称 p 为二次型的正惯性指数,q 是负惯性指数,$r=p+q$ 为二次型的秩.

例如,例 1 的二次型经坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 化成 $2y_1^2+y_2^2+2y_3^2$,所以二次型 $2x_1^2+3x_2^2+8x_3^2+4x_1x_2-8x_2x_3-4x_3x_1$ 的正惯性指数 $p=3$,负惯性指数 $q=0$,二次型的秩 $p+q=3$. 而例 2 中二次型 $2x_1x_2-4x_2x_3$ 的正、负惯性指数 $p=q=1$,二次型的秩 $p+q=2$.

在二次型的标准形 $d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2$ 中,如果系数 $d_i(i=1,2,\cdots,n)$ 的值为 ±1 或 0,这样的标准形称为二次型的规范形.

一个二次型总可用坐标变换化其为规范形. 例如,在例 1 中

再令

$$\begin{aligned} z_1 &= 2y_1, \\ z_2 &= \quad y_2, \\ z_3 &= \qquad 2y_3, \end{aligned}$$

就有 $f = z_1^2 + z_2^2 + z_3^2$,它就是 f 的规范形 . 又如,在例 2 中再令

$$\begin{aligned} u_1 &= 2z_1, \\ u_2 &= \quad 2z_2, \\ u_3 &= \qquad z_3, \end{aligned}$$

就是 $f = u_1^2 - u_2^2$,它是 $2x_1x_2 - 4x_2x_3$ 的规范形 .

设 $\boldsymbol{A}$ 是 n 阶实对称矩阵,秩 $\mathrm{r}(\boldsymbol{A}) = r$,若可逆矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{A}$ 合同于对角矩阵 Λ,即 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC} = \Lambda$. 那么根据秩的性质我们有

$$\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}) = \mathrm{r}(\Lambda),$$

而对角矩阵 Λ 的秩 $\mathrm{r}(\Lambda)$就是对角线上非 0 元素的和 . 所以二次型的秩 $p + q = \mathrm{r}(\Lambda)$就是矩阵 $\boldsymbol{A}$ 的秩 $\mathrm{r}(\boldsymbol{A})$. 即 $p + q = \mathrm{r}$.

根据惯性定理,如果 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$ 的正负惯性指数分别为 p 和 q,如同前面例题的讨论,我们可先选坐标变换化其为标准形

$$\begin{aligned} & d_1y_1^2 + d_2y_2^2 + \cdots + d_py_p^2 - d_{p+1}y_{p+1}^2 \\ & - d_{p+2}y_{p+2}^2 - \cdots - d_{p+q}y_{p+q}^2 \end{aligned}$$

其中系数 $d_i > 0\ (i = 1, 2, \cdots, p + q)$.

再选坐标变换化其为规范形

$$z_1^2 + z_2^2 + \cdots + z_p^2 - z_{p+1}^2 - z_{p+2}^2 - \cdots - z_{p+q}^2 .$$

用矩阵的话来说,上述二次型矩阵 $\boldsymbol{A}$ 合同于

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}. \tag{6.5}$$

其中 1 有 p 个, -1 有 q 个, 0 有 $n-(p+q)$ 个. 这样的对角矩阵称为 $\boldsymbol{A}$ 的合同规范形.

如果二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经过坐标变换化为二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$, 那么称这两个二次型是合同的. 显然, 这充要条件是它们的正惯性指数相同, 负惯性指数也相同, 也即矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是合同的.

6.3 用正交变换化二次型为标准形

二次型的矩阵是实对称矩阵, 而实对称矩阵是可以经过正交变换化为对角形的(5.4), 如 $\boldsymbol{T}$ 是正交矩阵, 则 $\boldsymbol{T}^{-1}=\boldsymbol{T}^{\mathrm{T}}$, 那么

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{T}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{T}.$$

因此, 实对称矩阵化为对角形(相似)与实对称矩阵合同于一个对角矩阵就统一起来了. 所以二次型除用配方法外亦可用正交变换化为标准形. 下面通过例题介绍这种方法.

例 1 求正交变换 $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$, 化二次型

$$f(x_1, x_2) = 2x_1^2 + 2x_2^2 - 2x_1x_2$$

为标准形.

解 二次型的矩阵为 $\boldsymbol{A}=\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$,

由此求出 $\boldsymbol{A}$ 的特征多项式

$$|\boldsymbol{A}-\lambda\boldsymbol{I}| = \begin{vmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3,$$

得到 $\boldsymbol{A}$ 的特征值是 $\lambda_1=1,\lambda_2=3$.

当 $\lambda=1$ 时,由特征方程 $(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{x}=0$,即

$$\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix},$$

得到基础解系(即特征向量):$\boldsymbol{X}_1=\begin{pmatrix}1\\ 1\end{pmatrix}$.

当 $\lambda=3$ 时,由特征方程 $(\boldsymbol{A}-3\boldsymbol{I})\boldsymbol{x}=0$,即

$$\begin{pmatrix}-1 & -1\\ -1 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix},$$

得基础解系(即特征向量) $\boldsymbol{X}_2=\begin{pmatrix}1\\ -1\end{pmatrix}$.

由于 $\lambda_1\neq\lambda_2$,特征向量 $\boldsymbol{X}_1$、$\boldsymbol{X}_2$ 是正交的(定理 5.8),故只须把其单位化,有

$$\gamma_1=\frac{\boldsymbol{X}_1}{\|\boldsymbol{X}_1\|}=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix},\quad \gamma_2=\frac{\boldsymbol{X}_2}{\|\boldsymbol{X}_2\|}=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -1\end{pmatrix}$$

构造正交矩阵 T

$$\boldsymbol{T}=(\gamma_1,\gamma_2)=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{pmatrix}$$

令 $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$,则有

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{T}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{T}\boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{T}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{T}\boldsymbol{y}$$

$$=\boldsymbol{y}^{\mathrm{T}}\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{pmatrix}\boldsymbol{y}$$

$$=\boldsymbol{y}^{\mathrm{T}}\begin{pmatrix}1 & \\ & 3\end{pmatrix}\boldsymbol{y}=y_1^2+3y_2^2.$$

所以,经正交变换 $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$,二次型 $2x_1^2+2x_2^2-2x_1x_2$ 的标准形是 $y_1^2+3y_2^2$. ■

例 2 求正交变换 $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$,化二次型

$$3x_1^2+6x_2^2+3x_3^2-4x_1x_2-8x_1x_3-4x_2x_3$$

为标准形.

解 二次型的矩阵是

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

矩阵 $\boldsymbol{A}$ 的特征多项式

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}3-\lambda&-2&-4\\-2&6-\lambda&-2\\-4&-2&3-\lambda\end{vmatrix}=\begin{vmatrix}7-\lambda&0&\lambda-7\\-2&6-\lambda&-2\\-4&-2&3-\lambda\end{vmatrix}$$

$$=(7-\lambda)\begin{vmatrix}1&0&-1\\-2&6-\lambda&-2\\-4&-2&3-\lambda\end{vmatrix}$$

$$=(7-\lambda)\begin{vmatrix}1&0&0\\-2&6-\lambda&-4\\-4&-2&-1-\lambda\end{vmatrix}$$

$$=(7-\lambda)(\lambda-7)(\lambda+2).$$

得到 $\boldsymbol{A}$ 的特征值:$\lambda_1=\lambda_2=7,\lambda_3=-2$.

当 $\lambda=7$ 时,由 $(\boldsymbol{A}-7\boldsymbol{I})\boldsymbol{x}=0$,即

$$\begin{pmatrix}-4&-2&-4\\-2&-1&-2\\-4&-2&-4\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

系数矩阵的秩为 1,得基础解系:

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

当 $\lambda=-2$ 时,由 $(\boldsymbol{A}+2\boldsymbol{I})\boldsymbol{x}=0$,即

$$\begin{pmatrix}5&-2&-4\\-2&8&-2\\-4&-2&5\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

对系数矩阵高斯消元,有

$$\begin{pmatrix}1&-4&1\\0&2&-1\\0&0&0\end{pmatrix},$$

得到基础解系(即 $\lambda=-2$ 所对应的特征向量)

$$\boldsymbol{X}_3=\begin{pmatrix}2\\1\\2\end{pmatrix}.$$

显然, $\boldsymbol{X}_3$ 与 $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ 都正交(它们是不同特征值所对应的特征向量),但 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 并不正交,对 $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ 运用 Schmidt 标准正交化,有

$$\beta_1=\boldsymbol{X}_1=(-1,2,0)^{\mathrm{T}},$$

$$\beta_2=-\frac{(\beta_1,\boldsymbol{X}_2)}{(\beta_1,\beta_1)}\beta_1+\boldsymbol{X}_2=\left(-\frac{4}{5},-\frac{2}{5},1\right)^{\mathrm{T}},$$

再单位化,得

$$\gamma_1=\frac{\beta_1}{\|\beta_1\|}=\frac{1}{5}\begin{pmatrix}-1\\2\\0\end{pmatrix},\quad \gamma_2=\frac{\beta_2}{\|\beta_2\|}=\frac{1}{3\;5}\begin{pmatrix}-4\\-2\\5\end{pmatrix},$$

$$\gamma_3=\frac{\boldsymbol{X}_3}{\|\boldsymbol{X}_3\|}=\frac{1}{3}\begin{pmatrix}2\\1\\2\end{pmatrix}.$$

构造正交矩阵

$$\boldsymbol{T} = (\gamma_1 \ \ \gamma_2 \ \ \gamma_3) = \begin{pmatrix} \frac{-1}{\sqrt{5}} & \frac{-4}{3\sqrt{5}} & \frac{2}{3} \\ \frac{2}{\sqrt{5}} & \frac{-2}{3\sqrt{5}} & \frac{1}{3} \\ 0 & \frac{5}{3\sqrt{5}} & \frac{2}{3} \end{pmatrix},$$

容易验证

$$\boldsymbol{T}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}^{-1} \boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} 7 & & \\ & 7 & \\ & & -2 \end{pmatrix}.$$

即经过 $\boldsymbol{x} = \boldsymbol{T}\boldsymbol{y}$,二次型化为标准形

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}} (\boldsymbol{T}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{T}) \boldsymbol{y} = 7y_1^2 + 7y_2^2 - 2y_3^2 . \qquad \blacksquare$$

通过以上例题我们看到,二次型经正交变换可以化成标准形,并且在标准形中平方项的系数就是该二次型矩阵的特征值,这不是偶然的巧合 . 关键在于二次型的矩阵是实对称矩阵,实对称矩阵的特征值及特征向量的性质保证我们可以构造正交矩阵 T,由于 $T^{-1} = T^{\mathrm{T}}$,那么 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{T}$,因此 $\boldsymbol{A}$ 相似于对角阵与 $\boldsymbol{A}$ 合同于对角阵是一致的 . 这样,我们就可用二次型的语言来叙述上一章的定理 5.9.

定理 6.5 对任一个 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,其中 $\boldsymbol{A}$ 是 n 阶实对称矩阵,必存在正交变换 $\boldsymbol{x} = \boldsymbol{T}\boldsymbol{y}$($\boldsymbol{T}$ 是正交矩阵),使得 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 化成标准形

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 . \tag{6.6}$$

这里 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值 . $\blacksquare$

二次型通过正交变换化为标准形的问题来源于平面的二次曲线及空间的二次曲面方程的化简 . 也就是通过对直角坐标系的平移及旋转把二次方程化为最简单的形式(不含混合项). 下面以二次曲线方程的化简为例介绍这一方法 .

*例 3　求直角坐标变换化简二次方程

$$5x^2 - 6xy + 5y^2 - 6\;\;2x + 2\;\;2y - 4 = 0$$

解　先把二次项化为标准形 . 由对应矩阵

$$\boldsymbol{A} = \begin{matrix} 5 & -3 \\ -3 & 5 \end{matrix}$$

求出 $\boldsymbol{A}$ 的特征值: $\lambda_1 = 2, \lambda_2 = 8$

由 $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = 0$ 求出相应的特征向量, 再单位化(现特征值不同, 特征向量必正交; 若特征值是等根就用 Schmidt 正交化)就有

$$\gamma_1 = \frac{1}{2}\begin{matrix} 1 \\ 1 \end{matrix}, \quad \gamma_2 = \frac{1}{2}\begin{matrix} -1 \\ 1 \end{matrix}.$$

构造正交矩阵 $\boldsymbol{T} = (\gamma_1, \gamma_2)$, 那么经正交变换(即直角坐标系的旋转)

$$\begin{matrix} x \\ y \end{matrix} = \begin{matrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{matrix} \begin{matrix} x' \\ y' \end{matrix},$$

也就是把

$$x = \frac{1}{2}x' - \frac{1}{2}y'$$

$$y = \frac{1}{2}x' + \frac{1}{2}y'$$

代入到原方程, 方程化简为

$$2x'^2 + 8y'^2 - 6\;\;2 \cdot \frac{1}{2}(x' - y') + 2\;\;2$$

$$\cdot \frac{1}{2}(x' + y') - 4 = 0$$

整理后即有

$$x'^2 + 4y'^2 - 2x' + 4y' - 2 = 0.$$

再用配平方法,得

$$(x'-1)^2+4\left(y'+\frac{1}{2}\right)^2=4.$$

因此,经平移坐标轴

$$x''=x'-1,$$

$$y''=y'+\frac{1}{2}.$$

得到二次曲线的标准方程

$$\frac{1}{4}x''^2+y''^2=1.$$

其图形是椭圆,而所用的直角坐标变换是

$$x''=\frac{1}{2}x+\frac{1}{2}y-1,$$

$$y''=-\frac{1}{2}x+\frac{1}{2}y+\frac{1}{2}.$$ ■

6.4 正定二次型

观察 6.2 中的例 1,由于二次型

$$\begin{aligned}f(x_1,x_2,x_3)&=2x_1^2+3x_2^2+8x_3^2+4x_1x_2-8x_2x_3-4x_3x_1\\&=2(x_1+x_2-x_3)^2+(x_2-2x_3)^2+2x_3^2\end{aligned}$$

那么对任何 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}\neq 0$,我们恒有

$$f(x_1,x_2,x_3)>0$$

且仅有 $f(0,0,0)=0$.

说明这个三元二次齐次式恒正,仅在 $\boldsymbol{x}=0$ 处函数有极小值 0. 而该节例 2 所给出的二次型 $f(x_1,x_2,x_3)=2x_1x_2-4x_2x_3$ 就不是恒正的. 例如,$f(1,0,0)=0$, $f(0,1,1)=-4$.

如何判断一个二次型 f 是否恒正,仅在 $\boldsymbol{x}=0$ 才有 $f=0$?这

就是本节所要学习的正定二次型、正定矩阵.

定义 6.3 对二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,如对任何 $\boldsymbol{x}\neq 0$,恒有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$,则称二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是正定二次型.并称实对称矩阵 $\boldsymbol{A}$ 是正定矩阵.

例如,$f(x_1,x_2)=x_1^2+3x_2^2$,对任何 $\boldsymbol{x}=\begin{pmatrix} t \\ u \end{pmatrix}\neq 0$ 总有 $f(t,u)=t^2+3u^2>0$,因此,这二次型是正定二次型,$\boldsymbol{A}=\begin{pmatrix} 1 & \\ & 3 \end{pmatrix}$ 是正定矩阵.

二次型 $f(x_1,x_2,x_3)=x_1^2+3x_2^2-5x_3^2+4x_1x_2-6x_2x_3$ 中,平方项 x_3^2 的系数 $a_{33}=-5<0$,如果选 $\boldsymbol{x}=(0,0,1)^{\mathrm{T}}\neq 0$,代入有 $f(0,0,1)=-5<0$,所以 f 不是正定的.

二次型 $f(x_1,x_2,x_3)=4x_2^2+x_3^2+2x_1x_2$ 中,平方项 x_1^2 的系数 $a_{11}=0$,如果取 $\boldsymbol{x}=(1,0,0)^{\mathrm{T}}\neq 0$,代入得 $\mathrm{f}(1,0,0)=0$,这个二次型也不是正定的.

从上面的讨论分析,容易得到:

若 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定,则其平方项的系数 $a_{ii}(i=1,2,\cdots,n)$ 必须全大于零.从正定矩阵的角度来说就是主对角线元素恒为正.

这是一个必要条件,对二次型

$$\mathrm{f}(x_1,x_2,x_3)=x_1^2+x_2^2+3x_3^2-10x_1x_2,$$

虽然 a_{ii} 全大于零,但 $\mathrm{f}(1,1,1)=-5<0$,f 并不是正定的二次型.

下面来研究二次型正定的充分必要条件

引理 6.6 设 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 是坐标变换,则 $\boldsymbol{x}\neq 0$ 的充要条件是 $\boldsymbol{y}\neq 0$.

证 必要性(用反证法)

如 $\boldsymbol{y}=0$,则 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}=\boldsymbol{C}0=0$,与已知矛盾.

充分性(用反证法).

如 $\boldsymbol{x}=\boldsymbol{0}$,由 $\boldsymbol{C}\boldsymbol{y}=0$ 而 $\boldsymbol{y}\neq 0$,知齐次方程组

$$\boldsymbol{C}\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}$$

有非零解,那么方程组的系数行列式 $|\boldsymbol{C}|=0$(定理 3.2),这与 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 是坐标变换, $\boldsymbol{C}$ 为可逆矩阵(参看(4.2.1))相矛盾. ■

引理 6.6 表明,二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 化成二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 时,二次型的正定性是不会因此而改变的. 这是因为:

对任意的 $\boldsymbol{y}_0\neq 0$, 令 $\boldsymbol{x}_0=\boldsymbol{C}\boldsymbol{y}_0$, 则 $\boldsymbol{x}_0\neq 0$,

那么,由 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定,就可得到

$$\boldsymbol{y}_0^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}_0=\boldsymbol{y}_0^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y}_0=\boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_0>0,$$

即二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 正定.

反之,若二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 正定,那么经坐标变换 $\boldsymbol{y}=\boldsymbol{C}^{-1}\boldsymbol{x}$,可知 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是正定二次型.

根据二次型正定的这一性质,要判断 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的正定性,就可以用它的标准形的正定性来决定,而标准形的正定性是容易看出的.

定理 6.7 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定的充分必要条件有:

(1) $\boldsymbol{A}$ 的正惯性指数是 n;

(2) $\boldsymbol{A}$ 与 $\boldsymbol{I}$ 合同,即存在可逆矩阵 $\boldsymbol{C}$,使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{I}$;

(3) $\boldsymbol{A}$ 的所有特征值 $\lambda_i\ (i=1,2,\cdots,n)$均为正数.

证 (1) 设坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 化二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为标准形

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y},$$

由引理 6.6 知: $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 正定

而 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 正定 $d_i>0\ (i=1,2,\cdots,n)$

正惯性指数 $p=n$.

(2) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定 正惯性指数 $p=n$

A 的合同规范形是 I(参看(6.5))

A 与 I 合同 .

(3) **对 $x^{T}Ax$ 有正交变换 $x = Ty$ 化其为标准形**

$$x^{T}Ax = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \qquad (6.6)$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ **是 $x^{T}Ax$ 的特征值,由**(1)**知**

$$x^{T}Ax \text{ 正定} \qquad \lambda_1, \lambda_2, \cdots, \lambda_n \text{ 全大于零 .}$$ ■

我们还可以用行列式来判断正定性,即

定理 6.8 n **元二次型 $x^{T}Ax$ 正定的充分必要条件是 A 的各阶顺序主子式** $\Delta_k(k = 1, 2, \cdots, n)$ **均大于零**,**即**

$$\Delta_1 = a_{11} > 0,\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0,\ \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0,$$

$$\cdots,\ \Delta_n = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} > 0.$$ ■

推论　实对称矩阵 A 正定的必要条件是 $|A| > 0$ **.** ■

例 1　说明下列矩阵都不是正定的 .

$$A = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 0 & 4 \\ -2 & 4 & 3 \end{pmatrix},\quad B = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ -3 & -7 & 10 \end{pmatrix}$$

解　对矩阵 A,由 $a_{22} = 0$,**不满足 A 正定的必要条件** a_{ii} **全大于** 0**. 所以 A 不是正定的 .**

对矩阵 B,由于 $|B| = 0$ **(把一、二两行全加至第三行 !),与正定的必要条件行列式大于** 0 **相矛盾(推论),所以 B 不是正定的 .**

■

例 2　判断二次型

$$f(x_1, x_2, x_3) = 6x_1^2 + 5x_2^2 + x_3^2 - 4x_1x_2 + 4x_1x_3$$

是否是正定二次型?

解 二次型的对应矩阵是

$$\boldsymbol{A} = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

其顺序主子式

$$\Delta_1 = a_{11} = 6 > 0,$$

$$\Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 6 & -2 \\ -2 & 5 \end{vmatrix} = 26 > 0,$$

$$\Delta_3 = |\boldsymbol{A}| = \begin{vmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 2 & -2 & 0 \\ -2 & 5 & 0 \\ 2 & 0 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & -2 \\ -2 & 5 \end{vmatrix} = 6 > 0$$

所以二次型是正定的. ■

例 3 二次型 $x_1^2 + 2x_2^2 + 4x_3^2 - 2x_1x_2 + 2tx_2x_3$ 正定,求参数 t 的取值范围.

解 二次型的对应矩阵是

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

其顺序主子式

$$\Delta_1 = 1,\quad \Delta_2 = \begin{vmatrix} 1 & -1 \\ -1 & 2 \end{vmatrix} = 1,\quad \Delta_3 = |\boldsymbol{A}| = 4 - t^2.$$

由于二次型正定,故 $4 - t^2 > 0$ (定理 6.8),解得当 $t \in (-2, 2)$时,二次型正定. ■

例 4 已知 $\boldsymbol{A}$ 是正定矩阵,证明 $\boldsymbol{A}^{-1}$是正定矩阵.

证 由于 $\boldsymbol{A}$ 正定,一方面知 $|\boldsymbol{A}|>0$,得 $\boldsymbol{A}^{-1}$存在;另一方面知 $\boldsymbol{A}$ 是对称的,即 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$. 那么

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

得 $\boldsymbol{A}^{-1}$是对称矩阵.

根据 $\boldsymbol{AX}=\lambda\boldsymbol{X}$ 及 $\boldsymbol{A}$ 正定时 $\lambda>0$(定理 6.7),有 $\boldsymbol{A}^{-1}\boldsymbol{X}=\dfrac{1}{\lambda}X$,所以 $\boldsymbol{A}^{-1}$的特征值$\dfrac{1}{\lambda}$全大于零,故 $\boldsymbol{A}^{-1}$正定. ■

注 证明 $\boldsymbol{A}^{-1}$正定的方法很多,例如

由 $\boldsymbol{A}$ 正定知存在可逆矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\boldsymbol{I}$,那么两边取逆,有

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

记 $(\boldsymbol{C}^{-1})^{\mathrm{T}}=\boldsymbol{D}$,则 $\boldsymbol{C}^{-1}=\boldsymbol{D}^{\mathrm{T}}$,故

$$\boldsymbol{D}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{D}=\boldsymbol{I}.$$

所以 $\boldsymbol{A}^{-1}$正定.

例 5 $\boldsymbol{A}$ 是 3 阶实对称矩阵,证明存在实数 t 使 $\boldsymbol{A}+t\boldsymbol{I}$ 是正定矩阵.

证 设 $\boldsymbol{A}$ 的特征值是 $\lambda_1,\lambda_2,\lambda_3$. 那么 $\boldsymbol{A}+t\boldsymbol{I}$ 的特征值是 $\lambda_1+t,\lambda_2+t,\lambda_3+t$. 总可选取实数 t 使 $\lambda_1+t>0,\lambda_2+t>0,\lambda_3+t>0$. 由定理 6.7,$\boldsymbol{A}+t\boldsymbol{I}$ 的特征值全大于零知其正定. ■

***例 6** $\boldsymbol{A}$ 是 3 阶正定矩阵,证明 $|\boldsymbol{A}+\boldsymbol{I}|>1$.

证 因为 $\boldsymbol{A}$ 是实对称的,故存在正交矩阵 $\boldsymbol{T}$ 使

$$\boldsymbol{T}^{-1}\boldsymbol{AT}=\boldsymbol{T}^{\mathrm{T}}\boldsymbol{AT}=\begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{pmatrix},$$

其中 $\lambda_1,\lambda_2,\lambda_3$ 是 $\boldsymbol{A}$ 的特征值. 由 $\boldsymbol{A}$ 正定知 $\lambda_i(i=1,2,3)$全大于零,这样

$$A+I=T\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix}T^{-1}+I=T\begin{pmatrix}\lambda_1+1&&\\&\lambda_2+1&\\&&\lambda_3+1\end{pmatrix}T^{-1},$$

所以

$$|A+I|=|T|\begin{vmatrix}\lambda_1+1&&\\&\lambda_2+1&\\&&\lambda_3+1\end{vmatrix}|T^{-1}|=\begin{vmatrix}\lambda_1+1&&\\&\lambda_2+1&\\&&\lambda_3+1\end{vmatrix}$$

$$=(\lambda_1+1)(\lambda_2+1)(\lambda_3+1)>1.$$ ■

* **例 7** 已知 A 是三阶正定矩阵,证明存在三阶正定矩阵 B,使 $A=B^2$.

证 设 A 的特征值是 $\lambda_1,\lambda_2,\lambda_3$,由于 A 是正定的,这些特征值均大于零. 又因 A 是实对称矩阵,故有正交矩阵 T 使

$$T^{-1}AT=\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix},$$

那么

$$A=T\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix}T^{-1}=T\begin{pmatrix}\sqrt{\lambda_1}&&\\&\sqrt{\lambda_2}&\\&&\sqrt{\lambda_3}\end{pmatrix}T^{-1}T\begin{pmatrix}\sqrt{\lambda_1}&&\\&\sqrt{\lambda_2}&\\&&\sqrt{\lambda_3}\end{pmatrix}T^{-1},$$

令

$$B=T\begin{pmatrix}\sqrt{\lambda_1}&&\\&\sqrt{\lambda_2}&\\&&\sqrt{\lambda_3}\end{pmatrix}T^{-1},$$

则

$$\boldsymbol{B} \sim \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}.$$

$\boldsymbol{B}$ 的特征值 λ_1，λ_2，λ_2 全大于零，故 $\boldsymbol{B}$ 是正定矩阵，且 $\boldsymbol{A} = \boldsymbol{B}^2$. ■

与正定性相平行，还有半正定，负定，半负定等概念.

*定义 6.4　对 n 元二次型 $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x}$，如果对任意的非零向量 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}} \neq \boldsymbol{0}$，恒有

(1) $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x} \geq 0$，且存在非零的 $\boldsymbol{X}_0$ 使得 $\boldsymbol{X}_0^{\mathrm{T}} \boldsymbol{A}\boldsymbol{X}_0 = 0$，就称 $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x}$ 是半正定二次型，$\boldsymbol{A}$ 叫半正定矩阵；

(2) $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x} < 0$，则称 $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x}$ 是负定二次型，$\boldsymbol{A}$ 叫做负定矩阵；

(3) $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x} \leq 0$，且存在非零的 $\boldsymbol{X}_0$ 使 $\boldsymbol{X}_0^{\mathrm{T}} \boldsymbol{A}\boldsymbol{X}_0 = 0$，就称 $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x}$ 是半负定二次型，$\boldsymbol{A}$ 叫半负定矩阵.

除此之外，其它的二次型称为不定二次型.

*定理 6.9　n 元二次型 $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x}$ 负定的充分必要条件有：

(1) $\boldsymbol{x}^{\mathrm{T}} (-\boldsymbol{A}) \boldsymbol{x}$ 是正定二次型；

(2) $\boldsymbol{A}$ 的负惯性指数是 n，即 $\boldsymbol{A}$ 与 $-\boldsymbol{I}$ 合同；

(3) $\boldsymbol{A}$ 的顺序主子式负正相间，即奇数阶的主子式均小于零，偶数阶主子式均大于零.

(4) $\boldsymbol{A}$ 的特征值全是负数. ■

*例 8　证明三元二次型

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x} = -\boldsymbol{x}_1^2 - 2\boldsymbol{x}_2^2 - 3\boldsymbol{x}_3^2 + 2\boldsymbol{x}_1\boldsymbol{x}_2 + 2\boldsymbol{x}_2\boldsymbol{x}_3$$

是负定的.

证法 1　经配平方，有

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x} = -(x_1 - x_2)^2 - (x_2 - x_3)^2 - 2x_3^2$$

对任何 $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathrm{T}} \neq 0$，恒有 $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x} \leq 0$，而等式仅在

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

成立,即 $x_1 = x_2 = x_3 = 0$ 时才有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = 0$

所以,$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是负定二次型. ■

证法 2 二次型的对应矩阵

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

其顺序主子式

$$\Delta_1 = -1 < 0,\quad \Delta_2 = \begin{vmatrix} -1 & 1 \\ 1 & -2 \end{vmatrix} = 1 > 0,$$

$$\Delta_3 = |\boldsymbol{A}| = \begin{vmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{vmatrix} = \begin{vmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -3 \end{vmatrix} = -2 < 0$$

所以,$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是负定二次型. ■

习 题 6

6.1 写出下列二次型 $f(x_1, x_2, x_3)$ 的对应矩阵

(1) $x_1^2 + 4_3^2 - 2x_1x_2 + 6x_2x_3$,

(2) $2x_1^2 + x_2^2 + 5x_3^2 + 6x_1x_2 - 8x_1x_3 + 2x_2x_3$.

6.2 用配方法化二次型 $f(x_1, x_2, x_3)$ 为标准形,并写出所用坐标变换

(1) $x_1^2 + 2x_2^2 + 5x_3^2 + 2x_1x_2 + 6x_2x_3 + 2x_1x_3$,

(2) $2x_1^2 + 2x_2^2 + 7x_3^2 + 2x_1x_3 + 4x_2x_3$.

6.3 求下列实对称矩阵 $\boldsymbol{A}$ 的合同规范形,并求它所对应的二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的正、负惯性指数及二次型的秩.

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

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

6.4 求正交变换 $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$ 化二次型为标准形

(1) $x_1^2+2x_1x_2+x_2^2$,

(2) $x_1^2-x_2^2+2x_3^2+4x_2x_3$,

(3) $3x_1^2+6x_2^2+3x_3^2-4x_1x_2-8x_1x_3-4x_2x_3$,

(4) $3x_1^2+3x_2^2+5x_3^2+4x_1x_3-4x_2x_3$.

6.5 判断下列二次型 $f(x_1, x_2, x_3)$ 是否正定?(要求用特征值及顺序主子式两种方法)

(1) $3x_1^2+3x_2^2+3x_3^2+2x_1x_2+2x_1x_3+2x_2x_3$,

(2) $x_1^2-2x_1x_2+2x_1x_3+2x_2x_3$,

(3) $x_1^2+x_2^2+2x_3^2+2x_1x_3+2x_2x_3$.

6.6 下列二次型是正定的,求参数 t 的取值范围

(1) $x_1^2+x_2^2+x_3^2+2tx_1x_3$,

(2) $x_1^2+3x_2^2+tx_3^2+2x_1x_2+2tx_2x_3$,

(3) $2x_1^2+2x_2^2+x_3^2+2tx_1x_2+6x_1x_3+2x_2x_3$.

6.7 选择题

(1) **如 $\boldsymbol{A}$ 是任一正定矩阵,则下列命题中正确的是( )**

A. **存在可逆矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{I}$,**

B. **存在正交矩阵 $\boldsymbol{T}$,使 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{I}$,**

C. **$\boldsymbol{A}$ 必是对环可逆矩阵,**

D. **$\boldsymbol{A}$ 的负惯性指数为零是其充要条件.**

(2) **与矩阵 $\boldsymbol{A}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 2 \\ 0 & 2 & 2 \end{pmatrix}$ 合同的矩阵是( )**

A. $\begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}$, B. $\begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}$,

$C.\begin{pmatrix}1 & & \\ & -1 & \\ & & -1\end{pmatrix}$，　$D.\begin{pmatrix}-1 & & \\ & -1 & \\ & & 0\end{pmatrix}$.

(3) 二次型 $f(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2$ 在条件 $x_1^2 + x_2^2 + x_3^2 = 1$ 下的最大值是( ).

A.1，　B.3，　C.5，　D.7.

(4) 与矩阵 $\boldsymbol{A} = \begin{pmatrix}2 & 0 & 2\\ 0 & -3 & 0\\ 2 & 0 & 2\end{pmatrix}$ 既相似又合同的矩阵是( )

$A.\begin{pmatrix}1 & & \\ & -1 & \\ & & 0\end{pmatrix}$，　$B.\begin{pmatrix}2 & & \\ & -3 & \\ & & 0\end{pmatrix}$，

$C.\begin{pmatrix}3 & & \\ & -4 & \\ & & 0\end{pmatrix}$，　$D.\begin{pmatrix}4 & & \\ & -3 & \\ & & 0\end{pmatrix}$.

6.8　$\boldsymbol{A}$、$\boldsymbol{B}$ 都是 3 阶正定矩阵，证明 $\boldsymbol{A} + 2\boldsymbol{B}$ 是正定矩阵.(提示:用正定的定义)

6.9　$\boldsymbol{A}$ 是 n 阶正定矩阵，证明 $\boldsymbol{A}$ 的伴随矩阵 $\boldsymbol{A}^*$ 是正定矩阵.(参看习题五第 3 题)

6.10　已知 $\lambda_1, \lambda_2, \lambda_3$ 是二次型 $f(x_1, x_2, x_3) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的三个特征值，若 $\lambda_1 < \lambda_2 < \lambda_3$. 证明:对任意的 $x = (x_1, x_2, x_3)^{\mathrm{T}}$ 恒有

$$\lambda_1 \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} \leq \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \leq \lambda_3 \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}.$$

6.11　$\boldsymbol{A}$ 是 n 阶矩阵，证明 $\boldsymbol{A}$ 可逆的充分必要条件是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 为正定矩阵.

6.12　已知二次型 $f(x_1, x_2, x_3) = 2x_1^2 + 3x_2^2 + 3x_3^2 + 2ax_2x_3$ $(a > 0)$，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.

附　录　线性代数应用举例

1. 把连续问题转化为离散问题

为描述函数 $y = f(x)$,我们应当知道在每一点 x 处其函数值 $f(x)$,把每一点的函数值 $f(x)$看成未知数,这就成为一个未知数的个数是无穷的连续问题。数字计算机不能求连续问题的解。因此要设法转化成为离散问题用近似解来逼近。

例如　一根棒上的温度分布由微分方程

$$-\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = \mathrm{f}(x), \qquad 0 \leq x \leq 1 .$$

$$u(0) = 0, \qquad u(1) = 1 .$$

来描述,为了求出 $u(x)$,我们用差商来近似二阶导数,即

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} \approx \frac{u(x+h) - 2u(x) + u(x-h)}{h^2},$$

把区间$[0,1]$ $n+1$ 等分,得到分点 $x_1 = h, x_2 = 2h, \cdots, x_n = nh$.记 $u_i = u(ih), u_0 = u(0), u_{n+1} = u(1)$,那么方程的解 $u(x)$在各分点的近似值 u_i 应满足:

$$-u_{i+1} + 2u_i - u_{i-1} = h^2\mathrm{f}(ih), \qquad (i = 1,2,\cdots,n)$$

对每个 i 有一个这样的方程.于是得到含有 n 个未知量 u_i, n 个方程所组成的一个线性方程组.由于 $u_0 = u_{n+1} = 0$,这个方程组是:

$$\begin{aligned}
2u_1 - u_2 \qquad\qquad\qquad\qquad &= h^2\mathrm{f}(h),\\
-u_1 + 2u_2 - u_3 \qquad\qquad\quad &= h^2\mathrm{f}(2h),\\
-u_2 + 2u_3 - u_4 \qquad &= h^2\mathrm{f}(3h),\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
-u_{n-2} + 2u_{n-1} - u_n &= h^2\mathrm{f}((n-1)h),\\
-u_{n-1} + 2u_n &= h^2\mathrm{f}(nh) .
\end{aligned}$$

容易看出,此方程组的系数行列式不为0(类似于1.2中例5),由克莱姆法则方程组有唯一解.利用高斯消元,就可求出$u(x)$在各分点的近似值u_i,因而可得到方程的近似解.显然,如果缩小差分的步长h,它将提高用差商代替导数的精度,同时对$u(x)$的描述也更细致一些,当然,由于未知数u_i个数的增多,计算工作量也就随之增大.

总之,一个连续问题可离散化,未知数的个数越多逼近连续问题的精度就越高,大线性方程组的求解问题是常会遇到的。

2. 矩阵对角化解微分方程组

在电路网络、溶液扩散振动理论等许多方面常会遇到线性常系数微分方程组的求解问题.如果利用矩阵特征值、特征向量的理论,转化为可分离变量型微分方程,问题的求解就方便许多.下面通过一个例题介绍这一方法.

例　解微分方程组

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -2x_1 + x_2,$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = x_1 - 2x_2 .$$

解　引入矩阵记号,记

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix},$$

则方程组可表示为

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \boldsymbol{AX}$$

求出$\boldsymbol{A}$的特征值$\lambda_1 = -1$, $\lambda_2 = -3$. 相应的特征向量取为$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$和$\begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

由于 $\boldsymbol{A}$ 有两个不同的特征值,故 $\boldsymbol{A}$ 可以对角化 . 构造可逆矩阵

$$\boldsymbol{C} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

则有

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

这时微分方程可改写成

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \boldsymbol{C}\begin{pmatrix} -1 & \\ & -3 \end{pmatrix}\boldsymbol{C}^{-1}\boldsymbol{X},$$

引入新的待求函数 $\boldsymbol{Y}$,令

$$\boldsymbol{Y} = \boldsymbol{C}^{-1}\boldsymbol{X}, \qquad \boldsymbol{X} = \boldsymbol{C}\boldsymbol{Y},$$

则有

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \boldsymbol{C}\frac{\mathrm{d}\boldsymbol{Y}}{\mathrm{d}t},$$

代入到原方程组,消去 $\boldsymbol{X}$,可得 $\boldsymbol{Y}$ 的方程

$$\frac{\mathrm{d}\boldsymbol{Y}}{\mathrm{d}t} = \begin{pmatrix} -1 & \\ & -3 \end{pmatrix}\boldsymbol{Y},$$

也就是

$$\frac{\mathrm{d}y_1}{\mathrm{d}t} = -y_1,$$

$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = -3y_2 .$$

可解出

$$y_1 = k_1\mathrm{e}^{-t}, \qquad y_2 = k_2\mathrm{e}^{-3t}$$

从而有

$$\boldsymbol{X} = \boldsymbol{C}\boldsymbol{Y} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} k_1\mathrm{e}^{-t} \\ k_2\mathrm{e}^{-3t} \end{pmatrix},$$

也就是说,原方程组的解是:

$$x_1 = k_1 e^{-t} - k_2 e^{-3t},$$

$$x_2 = k_1 e^{-t} + k_2 e^{-3t}.$$

3．最小二乘法

最小二乘法是一种在工程技术、商业与经济等方面常用的求经验公式的方法.

例如,经实验得到一批数据资料

$$(x_i, y_i) \qquad i = 1,2,\cdots,n.$$

通过把(x_i, y_i)当成点的坐标,把这些点描在坐标纸上,观察其图形,然后选用某函数来拟合、近似.

假若我们要用一次函数作为经验公式来刻化变量 x 与 y 之间的关系.把这批数据代入有:

$$kx_1 + b = y_1,$$

$$kx_2 + b = y_2,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$kx_n + b = y_n,$$

写成矩阵的形式,为

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \cdots & \cdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} k \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}.$$

一般说来,由于公式本来就是近似的,而且实验数据测量的本身也会有各种误差,所以上述方程组的系数矩阵及增广矩阵的秩不大可能会相等,也就是说方程组无解.

如果确实需要有一个近似公式,那么是任取两组实验数据来确定 k 和 b 合理,还是用这一批 n 组数据来求,并选用使对各组误差的平方和最小的 k 和 b 更合理、更可靠呢?

对线性方程组

$$\boldsymbol{AX} = \beta \qquad (\boldsymbol{A}\text{ 是 } m \times n \text{ 矩阵})$$

如果方程组无解,那么

$$\forall \boldsymbol{X}, \qquad \boldsymbol{AX} - \beta \neq 0 .$$

这样, $\boldsymbol{AX} - \beta$ 作为 n 维向量(非 0)其内积大于 0,即

$$\| \boldsymbol{AX} - \beta \|^2 = (\boldsymbol{AX} - \beta, \boldsymbol{AX} - \beta) > 0$$

若 $\boldsymbol{X}_0$ 使得 $\| \boldsymbol{AX} - \beta \|^2$ 取最小值,则称 $\boldsymbol{X}_0$ 是方程组 $\boldsymbol{AX} = \beta$ 的最小二乘解 .

我们通过例题介绍如何来求最小二乘解:

例 1　对 x 测试三次,有关系式

$$3x = 10,$$

$$4x = 7,$$

$$5x = 8 .$$

求在最小二乘意义下 x 的值 .

解　设

$$\boldsymbol{A} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}, \qquad \boldsymbol{X} = (x), \qquad \beta = \begin{pmatrix} 10 \\ 7 \\ 8 \end{pmatrix} .$$

所测关系式可用矩阵表示为

$$\boldsymbol{AX} = \beta .$$

由于

$$\begin{aligned} \| \boldsymbol{AX} - \beta \|^2 &= (\boldsymbol{AX} - \beta, \boldsymbol{AX} - \beta) = (\boldsymbol{AX} - \beta)^{\mathrm{T}} (\boldsymbol{AX} - \beta) \\ &= \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A} x^2 - 2 \boldsymbol{A}^{\mathrm{T}} \beta x + \beta^{\mathrm{T}} \beta \qquad (\text{注 } \boldsymbol{A}^{\mathrm{T}} \beta = \beta^{\mathrm{T}} \boldsymbol{A}) \\ &= 50 x^2 - 196 x + 213 \end{aligned}$$

二次函数的最小值在

$$x = \frac{196}{2 \times 50} = 1.96$$

时达到,所以在最小二乘意义下, x 的最佳取值是 $x = 1.96$.

一般地,关于一个变量的方程组 $\boldsymbol{AX}=\boldsymbol{\beta}$ 的最小二乘解是

$$x=\frac{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\beta}}{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}}.$$

例 2 求拟合(1,13),(2,18),(3,22),(4,29)四个点的最小二乘直线的方程.

解 设所求最小二乘直线的方程是 $y=kx+b$. 把四个点的坐标代入,得到方程组

$$\begin{pmatrix}1&1\\2&1\\3&1\\4&1\end{pmatrix}\begin{pmatrix}k\\b\end{pmatrix}=\begin{pmatrix}13\\18\\22\\29\end{pmatrix}.$$

令
$$\boldsymbol{A}=\begin{pmatrix}1&1\\2&1\\3&1\\4&1\end{pmatrix},\quad \boldsymbol{X}=\begin{pmatrix}k\\b\end{pmatrix},\quad \boldsymbol{\beta}=\begin{pmatrix}13\\18\\22\\29\end{pmatrix}$$

由于
$$\|\boldsymbol{AX}-\boldsymbol{\beta}\|^2=(k+b-13)^2+(2k+b-18)^2+(3k+b-22)^2+(4k+b-29)^2$$

二元函数 $f(k,b)=\|\boldsymbol{AX}-\boldsymbol{\beta}\|^2$ 取最小值的必要条件是

$$\frac{\partial f}{\partial k}=0,\qquad \frac{\partial f}{\partial b}=0$$

即

$$(k+b-13)+2(2k+b-18)+3(3k+b-22)+4(4k+b-29)=0,$$

$$(k+b-13)+(2k+b-18)+(3k+b-22)+(4k+b-29)=0,$$

整理后,为

$$30k+10b=231,$$
$$10k+4b=82.$$

解出:　　$k = 5.2$,　$b = 7.5$

所以,最小二乘直线的方程是 $y = 5.2x + 7.5$.

利用向量空间的理论,可以证明

方程组 $\boldsymbol{AX} = \beta$ 的最小二乘解也就是方程组 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX} = \boldsymbol{A}^{\mathrm{T}}\beta$ 的解.

注　请验证在例 2 中

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \begin{pmatrix} 30 & 10 \\ 10 & 4 \end{pmatrix}, \qquad \boldsymbol{A}^{\mathrm{T}}\beta = \begin{pmatrix} 231 \\ 82 \end{pmatrix}.$$

4. 编码问题

在引入矩阵概念时,我们已提到现代通讯可以是数字通讯.但数字信息在传递过程中可能会遇到种种干扰,难免出现差错,那么如何提高传递的可靠性呢?

假若我们用 16 个符号 $\alpha_1 — \alpha_{16}$ 可构成各种信息,为表示这 16 个符号,在二进制下只须选用四维向量:

$$(0,0,0,0)^{\mathrm{T}},\quad (0,0,0,1)^{\mathrm{T}},\quad (0,0,1,0)^{\mathrm{T}},\cdots,(1,1,1,1)^{\mathrm{T}}$$

分别表示每一个符号就可以了.每个这样的向量叫做信息码,向量的维数叫码长.

如果只有信息码:那么传递时是否有失误就不好判断了,为此要增加校验位,以确定传递是否正确,如有差错,是错在哪一个位置?

为了纠正信息码上一个错误,要用 3 个校验位*,这样构成码长为 7 的编码

$$\alpha = (p_1\ \ p_2\ \ m_1\ \ p_3\ \ m_2\ \ m_3\ \ m_4)^{\mathrm{T}}$$

其中 $m_1 \sim m_4$ 是信息码,$p_1 \sim p_3$ 是校验位.

纠错的原理是:构造齐次方程组 $\boldsymbol{HX} = 0$.使这 16 个 α_i 是方程组的解.如果接到编码 β 不是方程的解,假如第 i 位有错,在二进制下,我们的运算应在模 2 数域上进行,其运算规律是:

加法　$0 + 0 = 1 + 1 = 0$,　$0 + 1 = 1 + 0 = 1$;

乘法 $1\times0=0\times1=0\times0=0,1\times1=1$.

因此,必有

$$\beta=\alpha_j+e_i,$$

其中

$$e_i=(0,\cdots,0,\underset{\text{第 } i \text{ 位}}{1},0,\cdots,0)^{\mathrm{T}}$$

由于

$$H\beta=H(\alpha_j+e_i)=H\alpha_j+He_i=He_i,$$

而我们要求

$$He_i=H\begin{pmatrix}0\\ \cdots\\ 0\\ 1\\ 0\\ \cdots\\ 0\end{pmatrix}$$

正好是矩阵 H 中的第 i 列 . 因此,只要检查 $H\beta$ 与 H 中第几列一样,就知道 β 中那一个位置上的信息量错了 . 而每个位置上信息只有 0 与 1 两种,也就很容易改正传递上的差错了 .

例如,给出校验矩阵 H

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

如收到信号 $\beta=(1\ 0\ 1\ 0\ 0\ 1\ 1)^{\mathrm{T}}$,则由

$$H\beta=\begin{pmatrix}0\\1\\1\end{pmatrix}\neq0$$

知 β 有差错,从 $H\beta$ 与 H 的第 3 列一样,得知 β 中第 3 个位置的

信号有错．那么,正确的信号就应当是

$$(1\ 0\ 0\ 0\ 0\ 1\ 1)^{T},$$

对照信息码的位置,知传递的符号是

$$(0\ 0\ 1\ 1)^{T}.$$

*　从方程组 $\boldsymbol{HX}=0$,由系数矩阵的秩为 3,故要有 4 个自由变量,才能得到 $2^4=16$ 个解,因此方程组中应有 $3+4=7$ 个未知数,现在信息码长为 4,所以要添加 3 个校验位．

部分习题答案及提示

习题 1

1.1 (1) 18; (2) 2; (3) -37; (4) 6 .

1.3 (1) $\pm 1, 3$; (2) $2, 5, -6$; (3) 0(二重根), 3;

(4) $0, 1, -3$. 把第 3 列加至第 1 列 .

1.4 (1) $A_{12}=0$, $A_{22}=6$, $A_{32}=-5$. (2) $A_{12}=-1$, $A_{22}=1$, $A_{32}=2$, $A_{42}=2$.

1.5 (1) -8; (2) 0; (3) 120; (4) 240 .

1.6 (1) $\begin{cases} x=-2, \\ y=3, \end{cases}$ (2) $\begin{cases} x=\cos\theta, \\ y=-\sin\theta, \end{cases}$ (3) $\begin{cases} x_1=0, \\ x_2=0, \\ x_3=0, \end{cases}$

(4) $\begin{cases} x_1=3, \\ x_2=-1, \\ x_3=-2 . \end{cases}$

1.7 (1) $\pm 1, 2$, (2) -1(三重根) . 1 列 + 2 列 - 3 列 .

1.8 (1) 把第一列变换出 0, (2) 把 2、3 列加至第 1 列 .

1.9 (1) -7, (2) -2, (3) a^3, (4) $-a^2b^3$.

1.10 设法证明 $D=-D$, 注意 $D=D^{\mathrm{T}}$.

1.11 (1) $a^n+(-1)^{n+1}b^n$, (2) $n!$ (把 D 化成上三角行列式), (3) $\frac{1}{2}n(n+1)$, 把各列均加至第 1 列 .

习题 2

2.1 $\begin{pmatrix} 9 & 0 \\ 12 & -3 \end{pmatrix}$; $\begin{pmatrix} 0 & -6 \\ 3 & -9 \end{pmatrix}$; $\begin{pmatrix} 5 & 2 \\ 7 & 0 \end{pmatrix}$; $\begin{pmatrix} 19 & 37 \\ -14 & -28 \end{pmatrix}$;

$\begin{pmatrix}-1 & 2\\ -5 & 6\end{pmatrix}$; 18.

2.5 (1) $-\frac{1}{5}\begin{pmatrix}-2 & 3\\ -1 & 4\end{pmatrix}$, (2) $\frac{1}{4}\begin{pmatrix}3 & 1\\ -2 & 0\end{pmatrix}$,

(3) $-\frac{1}{5}\begin{pmatrix}1 & 0 & -2\\ -2 & -5 & 4\\ -3 & 0 & 1\end{pmatrix}$, (4) $\begin{pmatrix}1 & -1 & -1\\ & 1 & -1\\ & & 1\end{pmatrix}$.

2.6 (1) $\begin{pmatrix}3 & -2 & 8\\ 5 & -4 & 15\\ -7 & 4 & -2\end{pmatrix}$, (2) $\begin{pmatrix}2 & -18\\ 0 & 6\end{pmatrix}$,

(3) $\begin{pmatrix}7 & -4\\ 8 & -4\\ -5 & 3\end{pmatrix}$, (4) $\frac{1}{2}\begin{pmatrix}3 & 2\\ 1 & 0\end{pmatrix}$, (5) $\begin{pmatrix}4 & 4\\ 5 & 7\end{pmatrix}$.

2.8 (1) $\begin{pmatrix}t & u\\ 0 & t\end{pmatrix}$, (2) $\begin{pmatrix}2t+u & 2t\\ t & u\end{pmatrix}$. t, u **是任意实数.**

2.10 **注意** $\boldsymbol{A}\cdot\boldsymbol{A}^{*}=|\boldsymbol{A}|\boldsymbol{I}$,**现** $\boldsymbol{A}$ **是** 3 **阶矩阵,两边取行列式后**, $|\boldsymbol{A}^{*}|=|\boldsymbol{A}|^{2}$,**应选择**(4).

2.11 **选**(2)、(3)、(4). **注**: $(\boldsymbol{AB})^{2}=(\boldsymbol{AB})(\boldsymbol{AB})=\boldsymbol{A}(\boldsymbol{BAB})$.

2.12 $|\boldsymbol{A}|=-\frac{1}{4}$, $\boldsymbol{A}^{-1}=\begin{pmatrix}1 & & \\ & -10 & 6\\ & 4 & -2\end{pmatrix}$, $(\boldsymbol{A}^{*})^{-1}$

$=\begin{pmatrix}-4 & & \\ & -2 & -6\\ & -4 & -10\end{pmatrix}$. **参看** 2.2 **例** 4.

2.17 **联系 2.4 例 1**

$$\frac{1}{3}\qquad\qquad \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$$

2.20 (1) $\begin{pmatrix} 1 & \\ & \frac{1}{2} \end{pmatrix}$, (2) $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 1 & -1 & -1 \end{pmatrix}$.

2.22 $\boldsymbol{A}^{-1}=\frac{1}{2}(\boldsymbol{A}-\boldsymbol{I})$, $(\boldsymbol{A}+2\boldsymbol{I})^{-1}=\frac{1}{4}(3\boldsymbol{I}-\boldsymbol{A})$.

注:$(\boldsymbol{A}+2\boldsymbol{I})\cdot(\boldsymbol{A}-3\boldsymbol{I})=\boldsymbol{A}^2-\boldsymbol{A}-6\boldsymbol{I}$.

2.23 $\boldsymbol{A}^{-1}=\frac{1}{2}(3\boldsymbol{I}-\boldsymbol{A})$.注:$\boldsymbol{A}^2=(\boldsymbol{I}+\boldsymbol{B})^2=\boldsymbol{I}+3\boldsymbol{B}=$

2.26
$$\begin{aligned} & [(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}+\boldsymbol{A})^{-1}]^{\mathrm{T}}[(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}+\boldsymbol{A})^{-1}] \\ = & [(\boldsymbol{I}+\boldsymbol{A})^{-1}]^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{A})^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}+\boldsymbol{A})^{-1} \\ = & [(\boldsymbol{I}+\boldsymbol{A})^{\mathrm{T}}]^{-1}(\boldsymbol{I}^{\mathrm{T}}-\boldsymbol{A}^{\mathrm{T}})(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}+\boldsymbol{A})^{-1}=\cdots \end{aligned}$$

2.27 用归纳法、分块技巧:

$$\begin{pmatrix} a_{11} & \alpha \\ 0 & \boldsymbol{A}_{n-1} \end{pmatrix}\begin{pmatrix} b_{11} & \beta \\ 0 & \boldsymbol{B}_{n-1} \end{pmatrix}=\begin{pmatrix} a_{11}b_{11} & a_{11}\beta+\alpha\boldsymbol{B}_{n-1} \\ 0 & A_{n-1}\boldsymbol{B}_{n-1} \end{pmatrix}$$

这里 $\boldsymbol{A}_{n-1}$, $\boldsymbol{B}_{n-1}$是 $n-1$ 阶上三角矩阵, α、β是 $1\times(n-1)$矩阵,**0** 是$(n-1)\times 1$ 矩阵

2.28 (1) $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$. 注: $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n=\left(\boldsymbol{I}+\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right)^n$

$=\boldsymbol{I}+n\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 一般地,$(\boldsymbol{A}+\boldsymbol{B})^2\neq\boldsymbol{A}^2+2\boldsymbol{AB}+\boldsymbol{B}^2$,两者不要混淆.

(2) $3^{n-1}\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$. 因 $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}=\begin{pmatrix} 1 \\ 2 \end{pmatrix}(1\quad 1)$,参看 2.1 中例 8.

2.29 (1) $f(\boldsymbol{A})=\begin{pmatrix} |0| & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

注: $\boldsymbol{A}^2=\boldsymbol{A}^3=\cdots=\boldsymbol{A}^{100}=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}$

(2) $\boldsymbol{B}^{-1}=-\frac{1}{3}(\boldsymbol{B}+2\boldsymbol{I})$. 注: $f(\boldsymbol{B})=\boldsymbol{B}^2+2\boldsymbol{B}+3\boldsymbol{I}$.

2.30 $\boldsymbol{A}$ 是上三角阵,可证 $\boldsymbol{A}^{-1}$ 是上三角矩阵,再设 $\boldsymbol{A}^{-1}=\begin{pmatrix}b_{11}&b_{12}&b_{13}\\&b_{22}&b_{23}\\&&b_{33}\end{pmatrix}$, 由 $\boldsymbol{A}$ 是正交的, $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{T}}$ 可证 $a_{ij}=0\ (i\neq j)$. 利用 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{I}$ 来证 a_{ii} 为 $+1$ 或 -1.

习题 3

3.1 (1) $x_1=1$, $x_2=2$, $x_3=1$; (2) $x_1=2t-u$, $x_2=t$, $x_3=u$, $x_4=1$; (3) 无解; (4) $x_1=-8$, $x_2=t+3$, $x_3=2t+6$, $x_4=t$.

3.4 (1) $t=15$, (2) $t\neq12$, (3) $t\neq2$ 且 $t\neq3$ 时, β 表示式唯一; $t=2$ 时, β 表示法不唯一.

3.5 (1) 线性相关, $\alpha_3=\alpha_1-2\alpha_2$; (2) 线性相关, α_3 不能由 α_1, α_2 线性表出; (3) 线性相关, $\alpha_3=-\alpha_1+2\alpha_2$, (4) 线性无关.

3.6 (1) 错. 例如(1,0), (0,1)和(1,1); (2) 错. 例如 $\alpha_1=(1,0)$, $\alpha_2=(2,0)$ 与 $\beta_1=(0,2)$, $\beta_2=(0,3)$; (3) 正确. 定理 3.3 与 3.6; (4) 正确. 定理 3.14.

3.7 不一定(参看习题中 5(2))

3.8 参看 3.2 例 12.

3.10 α_1, α_2, α_3 与 α_2, α_3, α_4 是等价向量组. 或用定义, 对 $k_2\alpha_2+k_3\alpha_3+k_4\alpha_4=0$, 把 α_4 代入, 再用 α_1, α_2, α_3 线性无关.

3.11 (1) $r=1$, α_2 是极大线性无关组. (2) $r=2$, α_1, α_3 是极大线性无关组. (3) $r=2$, α_1, α_2 是极大线性无关组.

3.13 $a=2$

3.18 (1) 错. 可能无解, 如 $\begin{cases} x+y=3, \\ x+y=2, \\ x-y=1. \end{cases}$ (2) 正确. $\boldsymbol{A}$ 的列向量必线性相关, 表示法不可能唯一; (3) 错. 参看第 9 题; (4) 错. $\boldsymbol{A}$ 不一定是 n 阶矩阵.

3.19 (1) $\eta=(6, -2, 1)^{\mathrm{T}}$; (2) $\eta_1=(8, -6, 1, 0)^{\mathrm{T}}$, $\eta_2=(-7, 5, 0, 1)^{\mathrm{T}}$; (3) $\eta_1=\left(\frac{7}{6}, \frac{5}{6}, 0, \frac{1}{3}, 1\right)^{\mathrm{T}}$, $\eta_2=(-1, 1, 1, 0, 0)^{\mathrm{T}}$; (4) $\eta_1=(1, 0, 1, 0, 0)^{\mathrm{T}}$, $\eta_2=(0, 1, 0, 1, 0)^{\mathrm{T}}$.

3.20 (1) $\left(\frac{1}{3}, 0, 0, 1\right)^{\mathrm{T}}+k_1(4, -3, 0, 0)^{\mathrm{T}}+k_2(1, 0, -3, 0)^{\mathrm{T}}$; (2) $(0, 0, 2, 0)^{\mathrm{T}}+k(1, -5, 11, 0)^{\mathrm{T}}$

3.21 (1) $a=7$ 时有解, $(1, 1, 0)^{\mathrm{T}}+k(-7, 5, 1)^{\mathrm{T}}$; (2) $a=0$ 或 3 时有解. 当 $a=0$ 时, $(4, 0, -2)^{\mathrm{T}}+k(1, -1, 0)^{\mathrm{T}}$, 当 $a=3$ 时, $(1, 0, 1)^{\mathrm{T}}+k(1, -1, 0)^{\mathrm{T}}$; (3) $a=1$ 或 -2 时有解, 当 $a=1$ 时, $(1, 0, 0)^{\mathrm{T}}+k(0, 1, 1)^{\mathrm{T}}$, 当 $a=-2$ 时, $(2, 2, 0)^{\mathrm{T}}+k(1, 1, 1)^{\mathrm{T}}$.

3.22 $a_1+a_2=b_1+b_2$ 时有解, $(a_1-b_2, b_2, a_2, 0)^{\mathrm{T}}+k(1, -1, -1, 1)^{\mathrm{T}}$.

3.23 (1) 当 $a=1$, $b=3$ 时, $\beta=(1-t)\alpha_1+t\alpha_2+\alpha_3$ 表示法不唯一; (2) 当 $b=2$ 或 $a=1$ 但 $b\neq 3$ 时 β 不能由 α_1, α_2, α_3 线性表出.

3.24 对 $\boldsymbol{A}=(\alpha_1, \alpha_2, \cdots, \alpha_n)$ 按列分块, $\boldsymbol{Ax}=\beta$ 即 $x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=\beta$,

" β, $\boldsymbol{Ax}=\beta$ 有解

$\alpha_1, \alpha_2, \cdots, \alpha_n$ 可表示任一 m 维向量 β

$\alpha_1, \alpha_2, \cdots, \alpha_n$ 可表示 $\varepsilon_1 = (1, 0, \cdots, 0)^T$，$\varepsilon_2 = (0, 1, \cdots, 0)^T, \cdots$，$\varepsilon_m = (0, 0, \cdots, 1)^T$

$\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m$ 是等价向量组

$r(\alpha_1, \alpha_2, \cdots, \alpha_n) = r(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m)$

$r(\boldsymbol{A}) = m$.

3.25 $\boldsymbol{A} \cdot \boldsymbol{A}^* = |\boldsymbol{A}| \boldsymbol{I}$，如 $r(\boldsymbol{A}) < n$，则 $|\boldsymbol{A}| = 0$. 若 $r(\boldsymbol{A}) < n - 1$，$\boldsymbol{A}$ 的 $n-1$ 阶行列式？

习题 4

4.1 $(2,\ 1,\ -1)^{\mathrm{T}}$.

4.2 $\begin{matrix} -1 & 0 & 1 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{matrix}$, 3 . $\begin{matrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \end{matrix}$, 4 . $\begin{matrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{matrix}$, $\begin{matrix} 1 \\ 2 \\ 2 \end{matrix}$.

4.5 7 , 11 , 2 , $\arccos \dfrac{2}{77}$;

4.6 $\pm \dfrac{1}{2}(1,\ 0,\ 1)^{\mathrm{T}}$

4.7 (1) $\dfrac{1}{3}(1,\ 1,\ 1)^{\mathrm{T}}$, $\dfrac{1}{6}(-2,\ 1,\ 1)^{\mathrm{T}}$, $\dfrac{1}{2}(0,\ -1,\ 1)^{\mathrm{T}}$;

$\alpha = \left(0,\ -\dfrac{3}{6},\ -\dfrac{1}{2}\right)^{\mathrm{T}}$,

(2) $\dfrac{1}{3}(1,\ 1,\ 1)^{\mathrm{T}}$, $\dfrac{1}{6}(-1,\ 2,\ -1)^{\mathrm{T}}$, $\dfrac{1}{2}(-1,\ 0,\ 1)^{\mathrm{T}}$; $\alpha =$

$\left(0,\ -\dfrac{3}{6},\dfrac{1}{2}\right)^{\mathrm{T}}$,

(3) $\dfrac{1}{3}(1,\ -1,\ 1)^{\mathrm{T}}$, $\dfrac{1}{6}(-1,\ 1,\ 2)^{\mathrm{T}}$, $\dfrac{1}{2}(1,\ 1,\ 0)^{\mathrm{T}}$; $\alpha =$

$\left(\dfrac{2}{3},\ -\dfrac{2}{6},0\right)^{\mathrm{T}}$.

4.8 (1) **行空间** $R(\boldsymbol{A}^{\mathrm{T}})$**的基** $\alpha_1 = (1,\ 2,\ 0)$, $\alpha_2 = (0,\ 1,\ 1)$ **列空间** $R(\boldsymbol{A})$**的基** $\beta_1 = (1,\ 0,\ 1)^{\mathrm{T}}$, $\beta_2 = (0,\ 1,\ 0)^{\mathrm{T}}$.

(2) **行空间一维**,**基**: $(1,\ 1,\ 2)$; **列空间一维**, **基是** $\begin{matrix} 1 \\ 2 \end{matrix}$ **.**

4.9 (1) **解空间一维**,**标准正交基**: $\dfrac{1}{6}(2,\ -1,\ 1)^{\mathrm{T}}$

(2) **解空间二维**, $\dfrac{1}{2}(-1,\ 1,\ 0)^{\mathrm{T}}$, $\dfrac{1}{3}(1,\ 1,\ -1)^{\mathrm{T}}$ **是标准正交基 .**

4.10 3维. $\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $\begin{pmatrix}0&1\\0&0\end{pmatrix}$, $\begin{pmatrix}0&0\\0&1\end{pmatrix}$ 是一组基.

4.11 $(-1, -1, -1, 4)^{\mathrm{T}}$.

4.13 $\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix}$, $\begin{pmatrix}1&0&0\\-1&0&1\\2&2&0\end{pmatrix}$.

4.15 (1) $\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$, (2) $\begin{pmatrix}1\\-1\\2\end{pmatrix}$, (3) $\begin{pmatrix}1&&\\&1&\\&&-1\end{pmatrix}$.

4.16 如 $k_1\alpha_1 + k_2\alpha_2 + k_3\alpha_3 + l\beta = 0$, 用 β 作内积可得 $l(\beta, \beta) = 0$.

4.17 构造齐次线性方程组 $\boldsymbol{Ax} = 0$, 其中

$$\boldsymbol{A} = \begin{pmatrix}\alpha_1^{\mathrm{T}}\\\alpha_2^{\mathrm{T}}\\\alpha_3^{\mathrm{T}}\end{pmatrix}$$

是 3×4 矩阵, 且 $r(\boldsymbol{A}) = r(\alpha_1, \alpha_2, \alpha_3) = 3$. 基础解系?

4.18 $\alpha_1, \alpha_2, \alpha_3$ 是一组基, 可设 $\beta = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3$, 考查 $(\beta, \beta) = ?$

习题 5

5.1 (1) $\lambda_1 = 2$, $\boldsymbol{X}_1 = (5, 2)^{\mathrm{T}}$; $\lambda_2 = -1$, $\boldsymbol{X}_2 = (1, 1)^{\mathrm{T}}$; (2) $\lambda_1 = 2$, $\boldsymbol{X}_1 = (1, -1)^{\mathrm{T}}$; $\lambda_2 = 9$, $\boldsymbol{X}_2 = (2, 5)^{\mathrm{T}}$; (3) $\lambda_1 = 1$, $\boldsymbol{X}_1 = (1, 0, 0)^{\mathrm{T}}$, $\lambda_2 = \lambda_3 = 0$, $\boldsymbol{X}_2 = (1, -1, 0)^{\mathrm{T}}$, $\boldsymbol{X}_3 = (1, 0, -1)^{\mathrm{T}}$; (4) $\lambda_1 = 1$, $\boldsymbol{X}_1 = (1, 0, -1)^{\mathrm{T}}$; $\lambda_2 = 3$, $\boldsymbol{X}_2 = (1, -2, 1)^{\mathrm{T}}$; $\lambda_3 = 0$, $\boldsymbol{X}_3 = (1, 1, 1)^{\mathrm{T}}$; (5) $\lambda_1 = \lambda_2 = \lambda_3 = 0$, $\boldsymbol{X}_1 = (1, 0, 0)^{\mathrm{T}}$; (6) $\lambda_1 = \lambda_2 = \lambda_3 = -1$ (参看习题 1 1.7(2)), $\boldsymbol{X}_1 = (-1, -1, 1)^{\mathrm{T}}$.

5.2 (1) ± 1; (2) 2; (3) 1, -3.

5.3 如 λ 是 $\boldsymbol{A}$ 的特征值, 则 $\boldsymbol{A}^{-1}$ 的特征值是 $\frac{1}{\lambda}$, $\boldsymbol{A}^*$ 的特征值是 $\frac{1}{\lambda}|\boldsymbol{A}|$.

5.6 (1) D; (2) B; (3) C; (4) D.

5.8 **0**. **提示**: **定理** 5.4 **与** 5.3.

5.9 (1) $\boldsymbol{\Lambda}=\begin{pmatrix}2&\\&-1\end{pmatrix}$, $\boldsymbol{P}=\begin{pmatrix}5&1\\2&1\end{pmatrix}$; (2) $\boldsymbol{\Lambda}=\begin{pmatrix}2&\\&9\end{pmatrix}$, $\boldsymbol{P}=\begin{pmatrix}1&2\\-1&5\end{pmatrix}$; (3) $\boldsymbol{\Lambda}=\begin{pmatrix}1&&\\&0&\\&&0\end{pmatrix}$, $\boldsymbol{P}=\begin{pmatrix}1&1&1\\0&-1&0\\0&0&-11\end{pmatrix}$; (4) $\boldsymbol{\Lambda}=\begin{pmatrix}1&&\\&3&\\&&0\end{pmatrix}$, $\boldsymbol{P}=\begin{pmatrix}1&1&1\\0&-2&1\\-1&1&1\end{pmatrix}$; (5), (6) **不能对角化**.

5.11 $\boldsymbol{A}=\begin{pmatrix}2&-8\\0&-2\end{pmatrix}$

5.12 **不相似**. $\boldsymbol{A}\sim\Lambda$ **而** $\boldsymbol{B}$ **不能对角化**.

5.13 (1) $\boldsymbol{\Lambda}=\begin{pmatrix}2&\\&-3\end{pmatrix}$, $\boldsymbol{T}=\begin{pmatrix}\frac{2}{5}&\frac{1}{5}\\\frac{1}{5}&-\frac{2}{5}\end{pmatrix}$;

(2) $\boldsymbol{\Lambda}=\begin{pmatrix}1&&\\&1&\\&&-1\end{pmatrix}$, $\boldsymbol{T}=\begin{pmatrix}\frac{1}{2}&0&\frac{1}{2}\\0&1&0\\\frac{1}{2}&0&-\frac{1}{2}\end{pmatrix}$;

(3) $\boldsymbol{\Lambda}=\begin{pmatrix}3&&\\&1&\\&&-3\end{pmatrix}$,

$$\boldsymbol{T}=\begin{pmatrix}\frac{1}{3} & -\frac{1}{2} & \frac{1}{6}\\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6}\\ \frac{1}{3} & 0 & -\frac{2}{6}\end{pmatrix};$$

(4) $\boldsymbol{\Lambda}=\begin{pmatrix}3 & & \\ & 3 & \\ & & 0\end{pmatrix}$，$\boldsymbol{T}=\begin{pmatrix}\frac{1}{2} & \frac{1}{6} & \frac{1}{3}\\ 0 & -\frac{2}{6} & \frac{1}{3}\\ -\frac{1}{2} & \frac{1}{6} & \frac{1}{3}\end{pmatrix}$.

5.14　**6 .**

5.15　$\boldsymbol{6}^{n-1}\begin{pmatrix}2 & 2\\ 4 & 4\end{pmatrix}$,

5.17　$\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & -1\\ 0 & -1 & 0\end{pmatrix}$

5.18　**不能 . 因为** $r(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=3$,**而** λ_0 **是二重根**,**至多有** 2 **个线性无关的特征向量 .**

5.19　$\pm$**2**, 8. $\boldsymbol{B}$ **能对角化**(5.3 **推论** 3).

5.20　**- 3. 注意** $\boldsymbol{A}\sim\begin{pmatrix}1 & & \\ & 1 & \\ & & -1\end{pmatrix}$

习题 6

6.2　(1) $y_1^2+y_2^2$, $\begin{cases}x_1=y_1-y_2+y_3,\\ x_2=y_2-2y_3,\\ x_3=y_3,\end{cases}$

(2) $2y_1^2+2y_2^2+\frac{9}{2}y_3^2$, $\begin{cases} x_1 = y_1 - \frac{1}{2}y_3, \\ x_2 = y_2 - y_3, \\ x_3 = y_3. \end{cases}$

6.3 (1) $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$, $p=1$, $q=0$, $r=1$,

(2) $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$, $p=2$, $q=0$, $r=2$,

(3) $\begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}$, $p=2$, $q=1$, $r=3$,

(4) $\begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix}$, $p=1$, $q=2$, $r=3$.

(提示:可用配方法化为标准形,或看特征值的正负号)

6.4 (1) $2y_1^2$, $\boldsymbol{T}=\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$,

(2) $y_1^2+3y_2^2-2y_3^2$, $\boldsymbol{T}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{5} & -\frac{2}{5} \\ 0 & \frac{2}{5} & \frac{1}{5} \end{pmatrix}$,

(3) $7y_1^2+7y_2^2-2y_3^2$, $\boldsymbol{T}=\begin{pmatrix} \frac{1}{2} & \frac{1}{3\ 2} & \frac{2}{3} \\ 0 & -\frac{4}{3\ 2} & \frac{1}{3} \\ -\frac{1}{2} & \frac{1}{3\ 2} & \frac{2}{3} \end{pmatrix}$,

(4) $y_1^2+3y_2^2+7y_3^2$, $\boldsymbol{T}=\begin{pmatrix} -\frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & -\frac{1}{6} \\ \frac{1}{3} & 0 & \frac{2}{6} \end{pmatrix}$.

6.5 (1) **正定.(特征值是:**2(**二重**),5),

(2) **不定.(特征值是:**1, ± 3),

(3) **半正定.(求特征值时可**(1)**行** + (2)**行** - (3)**行)**

6.6 (1) $|t|<1$, (2) $t<0$ **或** $t>2$, (3) **不存在(对任何** t, f **均不可能正定)**.

6.7 (1) C, (2) B, (3) B, (4) D.

提示: **在**(3)**中,用正交变换** $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$ **化二次型为标准形,注意** $\boldsymbol{y}^{\mathrm T}\boldsymbol{y}=\boldsymbol{y}^{\mathrm T}\boldsymbol{T}^{\mathrm T}\boldsymbol{T}\boldsymbol{y}=\boldsymbol{x}^{\mathrm T}\boldsymbol{x}=1$.

6.9 **如** λ **是** $\boldsymbol{A}$ **的特征值,则** $\boldsymbol{A}^*$ **的特征值是** $\frac{|\boldsymbol{A}|}{\lambda}$, **由** $\boldsymbol{A}$ **正定可证** $\boldsymbol{A}^*$ **的特征值全大于零.**

或者, $\boldsymbol{A}^*=|\boldsymbol{A}|\boldsymbol{A}^{-1}$**,对任** $\boldsymbol{x}\neq 0$**,由**

$$\boldsymbol{x}^{\mathrm T}\boldsymbol{A}^*\boldsymbol{x}=|\boldsymbol{A}|\boldsymbol{x}^{\mathrm T}\boldsymbol{A}^{-1}\boldsymbol{x}$$

用 $\boldsymbol{A}$、$\boldsymbol{A}^{-1}$ **正定可证.**

6.10 **用正交变换化二次型为** $\boldsymbol{x}^{\mathrm T}\boldsymbol{A}\boldsymbol{x}=\lambda_1 y_1^2+\lambda_2 y_2^2+\lambda_3 y_3^2$, **看** 6.7(3)**提示.**

6.11 **如** $\boldsymbol{A}$ **可逆,则** $\boldsymbol{A}^{\mathrm T}\boldsymbol{A}=\boldsymbol{A}^{\mathrm T}\boldsymbol{I}\boldsymbol{A}$**,有** $\boldsymbol{A}^{\mathrm T}\boldsymbol{A}$D $\boldsymbol{I}$

如 $\boldsymbol{A}^{\mathrm T}\boldsymbol{A}$ **正定,则** $|\boldsymbol{A}^{\mathrm T}\boldsymbol{A}|>0$

6.12 $a=2$, $\boldsymbol{T}=\begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$.

注意： $\boldsymbol{A} = \begin{pmatrix} 2 & & \\ & 3 & a \\ & a & 3 \end{pmatrix}$ **的特征值是** 1, 2, 5**. 代到特征方程可求出** a**.**

第5章　特征值和特征向量

前几章的内容几乎都涉及到线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的求解,而基本技巧是高斯消元法与矩阵的初等行变换.现在新的问题是矩阵能否通过相似变换化为对角形的问题,虽然现在也还要用到行变换,但它只起辅助性的作用了.当前的关键是 n 阶矩阵的特征值与特征向量问题,这些在振动,稳定性、经济等问题上有着许多重要的应用.

5.1　特征值和特征向量

定义 5.1　设 $\boldsymbol{A}$ 是 n 阶矩阵,如果存在数 λ 及非零的 n 维向量 $\boldsymbol{X}$,使得

$$\boldsymbol{AX}=\lambda\boldsymbol{X} \tag{5.1}$$

成立,就称 λ 是矩阵 $\boldsymbol{A}$ 的特征值,$\boldsymbol{X}$ 是矩阵 $\boldsymbol{A}$ 属于特征值 λ 的一个特征向量.

对已给出的 n 阶矩阵 $\boldsymbol{A}$,怎样来求出 $\boldsymbol{A}$ 的特征值及其相对应的特征向量呢?

例如,已知二阶矩阵

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

按定义,若 λ 是 $\boldsymbol{A}$ 的特征值,$\boldsymbol{X}=(x_1,\ x_2)^{\mathrm{T}}$ 是属于 λ 的一个特征向量,那么,$\boldsymbol{AX}=\lambda\boldsymbol{X}$ 也就是

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\lambda\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

上述关系可以表示为:

$$a_{11}x_1 + a_{12}x_2 = \lambda x_1,$$
$$a_{21}x_1 + a_{22}x_2 = \lambda x_2,$$

作为未知数 λ, x_1, x_2 的方程已不是前面学过的线性方程组,因为它涉及到 λ 与 x_i 的乘积。但如果把上述方程改写成

$$\begin{aligned}(a_{11} - \lambda)x_1 + a_{12}x_2 = 0.\\ a_{21}x_1 + (a_{22} - \lambda)x_2 = 0.\end{aligned} \tag{5.1.1}$$

可以看出,当特征值 λ 已求出时,(5.1.1)就成为 x_1, x_2 的齐次线性方程组,而齐次方程组(5.1.1)的非零解 $\boldsymbol{X} = (x_1, x_2)^{\mathrm{T}}$ 正是矩阵 $\boldsymbol{A}$ 属于特征值 λ 的特征向量。由于(5.1.1)有非零解等价于它的系数行列式为零,即

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0, \tag{5.1.2}$$

这是 λ 的一个二次方程 . 这样我们就可以先解方程(5.1.2),求出特征值 λ_i,然后把 λ_i 代入到齐次线性方程组(5.1.1)再求出其非零解,也就是 λ_i 所对应的特征向量 .(为什么此时一定有非零解?)

对上面的分析,现在用矩阵来表示,并给出相关的术语、名词 .

对 $\boldsymbol{AX} = \lambda\boldsymbol{X}$, $\boldsymbol{X} \neq 0$,即齐次线性方程组

$$(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0} \tag{5.2}$$

有非零解,因此,系数行列式为零,即

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = \boldsymbol{0} \tag{5.3}$$

称 $|\boldsymbol{A} - \lambda\boldsymbol{I}|$ 是矩阵 $\boldsymbol{A}$ 的特征多项式,对 n 阶矩阵 $\boldsymbol{A}$ 来说,其特征多项式是 λ 的 n 次多项式 .称 $|\boldsymbol{A} - \lambda\boldsymbol{I}| = \boldsymbol{0}$ 是矩阵 $\boldsymbol{A}$ 的特征方程,它是 λ 的 n 次方程,因而有 n 个根,所以矩阵 $\boldsymbol{A}$ 有 n 个特征值 . 对每个特征值 λ_i,解齐次方程组

$$(\boldsymbol{A} - \lambda_i\boldsymbol{I})\boldsymbol{x} = 0$$

其非零解就是 A 的特征值 λ_i 所对应的特征向量.

对于三阶矩阵

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

它的特征多项式是

$$|A - \lambda I| = \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix},$$

这是 λ 的三次多项式,从特征方程

$$|A - \lambda I| = 0$$

可求出 A 的三个特征根 $\lambda_1, \lambda_2, \lambda_3$,而属于 λ_1 的特征向量就是齐次方程组 $(A - \lambda_1 I)x = \mathbf{0}$,即

$$\begin{aligned} (a_{11}-\lambda_1)x_1 + a_{12}x_2 + a_{13}x_3 &= 0, \\ a_{21}x_1 + (a_{22}-\lambda_1)x_2 + a_{23}x_3 &= 0, \\ a_{31}x_1 + a_{32}x_2 + (a_{33}-\lambda_1)x_3 &= 0, \end{aligned}$$

的非零解,类似地可求 λ_2, λ_3 所对应的特征向量.

根据齐次线性方程组解的性质,如果 X_1, X_2 都是方程组 $(A - \lambda I)x = 0$ 的解,那么 X_1 与 X_2 的线性组合 $C_1X_1 + C_2X_2$ 仍是这方程组的解.这样,关于矩阵的特征向量就有下述性质:

定理 5.1 如果 X_1, X_2 都是矩阵 A 的属于特征值 λ 的特征向量,那么当 $C_1X_1 + C_2X_2$ 非零时,$C_1X_1 + C_2X_2$ 仍是矩阵 A 属于特征值 λ 的特征向量. ■

关于特征向量,要注意的是:

(1) 如果 X_1 是矩阵 A 的属于特征值 λ_1 的特征向量,那么 X_1 就不可能再是矩阵 A 的属于特征值 $\lambda_2(\lambda_1 \neq \lambda_2)$ 的特征向量。

这是因为,如果

$$\boldsymbol{A X}_1 = \lambda_1 \boldsymbol{X}_1 \quad \text{且} \quad \boldsymbol{A X}_1 = \lambda_2 \boldsymbol{X}_1,$$

那么 $\lambda_1 \boldsymbol{X}_1 = \lambda_2 \boldsymbol{X}_1$ **,得到** $(\lambda_1 - \lambda_2)\boldsymbol{X}_1 = 0$,

但由 $\lambda_1 \neq \lambda_2$ **有** $\boldsymbol{X}_1 = 0$,**这与** $\boldsymbol{X}_1$ **是特征向量应当非零相矛盾.所以,** $\boldsymbol{X}_1$ **不能再是** λ_2 **所对应的特征向量.**

(2) **如果** $\boldsymbol{X}_1, \boldsymbol{X}_2$ **分别是矩阵** $\boldsymbol{A}$ **的不同特征值所对应的特征向量,那么** $\boldsymbol{X}_1 + \boldsymbol{X}_2$ **不再是** $\boldsymbol{A}$ **的特征向量(请看例 9).**

例 1　求矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ & 4 & 5 \\ & & 6 \end{pmatrix}$ **的特征值和特征向量.**

解　矩阵 $\boldsymbol{A}$ **的特征多项式为:**

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} 1-\lambda & 2 & 3 \\ & 4-\lambda & 5 \\ & & 6-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda)(6-\lambda).$$

因此,由特征方程

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = (1-\lambda)(4-\lambda)(6-\lambda) = 0$$

得到矩阵 $\boldsymbol{A}$ **的特征值是:**

$$\lambda_1 = 1,\ \lambda_2 = 4,\ \lambda_3 = 6.$$

当 $\lambda = 1$ **时,由齐次线性方程组** $(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{x} = 0$,**即**

$$\begin{pmatrix} 0 & 2 & 3 \\ 0 & 3 & 5 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

也就是

$$\begin{aligned} 2x_2 + 3x_3 &= 0, \\ 3x_2 + 5x_3 &= 0, \\ 5x_3 &= 0, \end{aligned}$$

解得基础解系是: $\boldsymbol{X}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ **.因此,属于** $\lambda_1 = 1$ **的全体特征向量就**

是 $c_1\boldsymbol{X}_1$，c_1 取自所有非零常数;

再将 $\lambda_2=4$ 代入到$(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=0$,即有$(\boldsymbol{A}-4\boldsymbol{I})\boldsymbol{x}=0$ 也就是

$$-3x_1+2x_2+3x_3=0,$$
$$5x_3=0,$$
$$2x_3=0,$$

求出基础解系是：$\boldsymbol{X}_2=\begin{pmatrix}2\\3\\0\end{pmatrix}$. 因此，属于 $\lambda_2=4$ 的全体特征向量是 $c_2\boldsymbol{X}_2$，c_2 为非零常数；最后，把 $\lambda_3=6$ 代入到 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=0$ 中，即 $(\boldsymbol{A}-6\boldsymbol{I})\boldsymbol{x}=0$，也就是

$$-5x_1+2x_2+3x_3=0,$$
$$-2x_2+5x_3=0,$$
$$0=0,$$

其基础解系是 $X_3=\begin{pmatrix}16\\25\\10\end{pmatrix}$. 因此，$\lambda_3=6$ 所对应的特征向量是 $c_3\boldsymbol{X}_3$，c_3 是非零常数. ■

从例 1 的推导求解不难看出，对于上(下)三角矩阵，对角矩阵来说，对角线上的元素就是其特征值.

例 2 求矩阵 $\boldsymbol{A}=\begin{pmatrix}1&-1\\2&4\end{pmatrix}$ 的特征值，特征向量.

解 矩阵 $\boldsymbol{A}$ 的特征多项式为

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}1-\lambda&-1\\2&4-\lambda\end{vmatrix}=\lambda^2-5\lambda+6,$$

解 $\boldsymbol{A}$ 的特征方程 $\lambda^2-5\lambda+6=0$，得到矩阵 $\boldsymbol{A}$ 的特征值是 $\lambda_1=2$ 和 $\lambda_2=3$.

当 $\lambda_1=2$ 时，由 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=0$ 即

$$-x_1-x_2=0$$
$$2x_1+2x_2=0$$

解得 $X_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. 故 A 对应于 $\lambda_1 = 2$ 的特征向量是 $c_1 X_1$, c_1 是任意的非零常数.

当 $\lambda_2 = 3$ 时, 由 $(A - 3I)x = 0$, 即

$$\begin{cases} -2x_1 - x_2 = 0 \\ 2x_1 + x_2 = 0 \end{cases}$$

解得 $X_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. 因此, A 对应于 $\lambda_2 = 3$ 的特征向量是 $c_2 X_2$, c_2 是任意的非零常数. ■

例 3 求矩阵 $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ 的特征值和特征向量.

解 矩阵 A 的特征多项式为

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 0 & 0 & -\lambda \end{vmatrix} = \lambda^2(1-\lambda),$$

由特征方程 $\lambda^2(1-\lambda) = 0$, 得 A 的特征值是: $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 0$.

当 $\lambda_1 = 1$ 时, 由 $(A - I)x = 0$, 即

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

解得 $X_1 = (1, 0, 0)^{\mathrm{T}}$. 因此, A 对应于 $\lambda_1 = 1$ 的特征向量是 $c_1 X_1$, c_1 是任意的非零常数;

当 $\lambda_2 = \lambda_3 = 0$ 时, 由 $(A - 0I)x = 0$, 即

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

解得 $X_2=(1,-1,0)^{\mathrm{T}}$. 因此, A 对应于 $\lambda=0$ 的特征向量是 c_2X_2, c_2 是非零常数. ■

例 4 求矩阵 $A=\begin{pmatrix}2&2&-2\\2&5&-4\\-2&-4&5\end{pmatrix}$ 的特征值及特征向量.

解 矩阵 A 的特征多项式为:

$$
\begin{aligned}
|A-\lambda I| &= \begin{vmatrix}2-\lambda&2&-2\\2&5-\lambda&-4\\-2&-4&5-\lambda\end{vmatrix}=\begin{vmatrix}2-\lambda&2&-2\\2&5-\lambda&-4\\0&1-\lambda&1-\lambda\end{vmatrix}\\
&=(1-\lambda)\begin{vmatrix}2-\lambda&2&-2\\2&5-\lambda&-4\\0&1&1\end{vmatrix}\\
&=(1-\lambda)\begin{vmatrix}2-\lambda&4&-2\\2&9-\lambda&-4\\0&0&1\end{vmatrix}\\
&=(1-\lambda)\begin{vmatrix}2-\lambda&4\\2&9-\lambda\end{vmatrix}\\
&=(1-\lambda)[(2-\lambda)(9-\lambda)-8]\\
&=(1-\lambda)^2(10-\lambda),
\end{aligned}
$$

由特征方程 $(1-\lambda)^2(10-\lambda)=0$, 得到 A 的特征值:

$$\lambda_1=\lambda_2=1,\ \lambda_3=10.$$

当 $\lambda_1=\lambda_2=1$ 时, 由 $(A-I)x=0$, 即

$$\begin{pmatrix}1&2&-2\\2&4&-4\\-2&-4&4\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

因为 $r(A-I)=1$, 其基础解系由两个向量 $X_1=(-2,1,0)^{\mathrm{T}}$ 和 $X_2=(2,0,1)^{\mathrm{T}}$ 构成. 根据定理 5.1, A 对应于特征值 $\lambda_1=\lambda_2=1$

的特征向量是 $c_1 \boldsymbol{X}_1 + c_2 \boldsymbol{X}_2$,其中 c_1, c_2 是不全为零的常数 .

当 $\lambda_3 = 10$ 时,由 $(\boldsymbol{A} - 10\boldsymbol{I})\boldsymbol{x} = 0$,即

$$\begin{pmatrix} -8 & 2 & -2 \\ 2 & -5 & -4 \\ -2 & -4 & -5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

经高斯消元,把系数矩阵化为阶梯形

$$\begin{pmatrix} 2 & -5 & -4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

可得基础解系 $\boldsymbol{X}_3 = (1, 2, -2)^{\mathrm{T}}$. 所以, $\boldsymbol{A}$ 关于特征值 $\lambda_3 = 10$ 所对应的特征向量是 $c_3 \boldsymbol{X}_3$,其中 c_3 是不为零的常数 . ■

*例 5　求矩阵 $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ 的特征值和特征向量 .

解　矩阵 $\boldsymbol{A}$ 的特征多项式为:

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1,$$

由特征方程 $\lambda^2 + 1 = 0$ 得到 $\boldsymbol{A}$ 的特征值是 $\pm i$.

当 $\lambda = i$ 时,由 $(\boldsymbol{A} - i\boldsymbol{I})\boldsymbol{x} = 0$,即

$$\begin{cases} -ix_1 + x_2 = 0, \\ -x_1 - ix_2 = 0, \end{cases}$$

解得 $\boldsymbol{X}_1 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$. 故 $\boldsymbol{A}$ 属于特征值 i 的特征向量是 $c_1 \boldsymbol{X}_1$, c_1 是非零常数;

当 $\lambda = -i$ 时,由 $(\boldsymbol{A} + i\boldsymbol{I})\boldsymbol{x} = 0$,即

$$\begin{cases} ix_1 + x_2 = 0, \\ -x_1 + ix_2 = 0, \end{cases}$$

解得 $\boldsymbol{X}_2 = \begin{pmatrix} i \\ 1 \end{pmatrix}$. 故 $\boldsymbol{A}$ 对应于 $\lambda = -i$ 的特征向量是 $c_2 \boldsymbol{X}_2$, c_2 是非

零常数． ■

例 6 已知 $A^2=A$，求 A 的特征值．

解 设 λ 是 A 的任一特征值，X 是 λ 所属的特征向量，按定义有：

$$AX=\lambda X,$$

两边同用矩阵 A 左乘，有

$$A^2X=A(\lambda X)=\lambda AX=\lambda^2X,$$

利用已知条件 $A^2=A$，有

$$A^2X=AX=\lambda X,$$

所以
$$\lambda^2X=\lambda X,$$

即有
$$(\lambda^2-\lambda)X=0,$$

因为 X 是特征向量，按定义 $X\neq0$，故有

$$\lambda^2-\lambda=0.$$

所以，矩阵 A 的特征值只能是 0 或 1． ■

注 满足 $A^2=A$ 的矩阵 A 是不唯一的，例如，零矩阵，单位矩阵都是这种类型的矩阵，而零矩阵的特征值全是 0，单位矩阵的特征值全是 1．又如矩阵 $\begin{matrix}1&\\&0\end{matrix}$ 亦满足 $A^2=A$ 的条件而它的特征值是 $\lambda_1=1$，$\lambda_2=0$．

例 7 A 是三阶矩阵，它的特征值是 $-1,0,4$．又知 $A+B=2I$，求 B 的特征值．

解 设矩阵 A 属于 -1 的特征向量是 X_1，即 $AX_1=-X_1$．那么，由 $A+B=2I$ 有

$$BX_1=(2I-A)X_1=2X_1-AX_1=3X_1.$$

按定义，3 是 B 的特征值，X_1 是其相应的特征向量．完全类似地，

$$BX_2=(2I-A)X_2=2X_2-AX_2=2X_2,$$

$$BX_3=(2I-A)X_3=2X_3-AX_3=-2X_3,$$

可得到 B 的特征值是 $3,2,-2$． ■

例 8 设 $\boldsymbol{A}$ 是可逆矩阵,证明 $\boldsymbol{A}$ 的所有特征值都不为 0.

证 设 λ 是 $\boldsymbol{A}$ 的任一特征值, $\boldsymbol{X}_0$ 是相应的特征向量,即 $\boldsymbol{A}\boldsymbol{X}_0=\lambda\boldsymbol{X}_0$.

如果 $\lambda=0$,则有 $\boldsymbol{A}\boldsymbol{X}_0=0$,这表明 $\boldsymbol{X}_0$ 是齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=0$ 的解.因为 $\boldsymbol{X}_0$ 是特征向量,它是非零的.于是 $\boldsymbol{A}\boldsymbol{x}=0$ 有非零解.

而 $\boldsymbol{A}\boldsymbol{x}=0$ 有非零解的充要条件是其系数行列式等于 0.因此,$|\boldsymbol{A}|=0$.这与 $\boldsymbol{A}$ 可逆相矛盾.所以,$\boldsymbol{A}$ 的特征值 λ 都不为 0.

■

从定理 5.1 我们知道了矩阵 $\boldsymbol{A}$ 的同一个特征值所对应的特征向量之间的关系,现在进一步来看不同的特征值所对应的特征向量之间的关系.

定理 5.2 如果 λ_1,λ_2 是矩阵 $\boldsymbol{A}$ 的不同的特征值,其对应的特征向量分别是 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$,那么 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 线性无关.

证 考查 $k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2=0$,两边同左乘矩阵 $\boldsymbol{A}$,有:

$$\begin{aligned}\boldsymbol{A}(k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2)&=k_1\boldsymbol{A}\boldsymbol{X}_1+k_2\boldsymbol{A}\boldsymbol{X}_2\\&=k_1\lambda_1\boldsymbol{X}_1+k_2\lambda_2\boldsymbol{X}_2=0.\end{aligned}$$

若两边同乘以 λ_1 又有: $\lambda_1k_1\boldsymbol{X}_1+\lambda_1k_2\boldsymbol{X}_2=0$,

上、下两式相减得: $(\lambda_2-\lambda_1)k_2\boldsymbol{X}_2=0$,

由于 $\lambda_1\neq\lambda_2$, $\boldsymbol{X}_2$ 是特征向量,有 $\boldsymbol{X}_2\neq 0$,从而有: $k_2=0$.

类似地, $k_1=0$.

因此,不同特征值所对应的特征向量 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 线性无关. ■

用数学归纳法很容易把定理 5.2 推广为:

如果 $\lambda_1,\lambda_2,\cdots,\lambda_t$ 是矩阵 $\boldsymbol{A}$ 的互不相同的特征值,其对应的特征向量分别是 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_t$,则 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_t$ 线性无关.

例 9 如果 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 分别是矩阵 $\boldsymbol{A}$ 属于不同特征值 λ_1,λ_2 的特征向量,则 $\boldsymbol{X}_1+\boldsymbol{X}_2$ 不是 $\boldsymbol{A}$ 的特征向量.

证 (用反证法) 如果 $\boldsymbol{X}_1+\boldsymbol{X}_2$ 是矩阵 $\boldsymbol{A}$ 对应于特征值 λ

的特征向量,即

$$A(X_1+X_2)=\lambda(X_1+X_2)=\lambda X_1+\lambda X_2.$$

又有 $$A(X_1+X_2)=AX_1+AX_2=\lambda_1X_1+\lambda_2X_2.$$

那么 $$(\lambda-\lambda_1)X_1+(\lambda-\lambda_2)X_2=0.$$

据定理 5.2, X_1, X_2 线性无关,所以

$$\lambda-\lambda_1=0,\ \lambda-\lambda_2=0.$$

这与 $\lambda_1\neq\lambda_2$ 相矛盾. 所以 X_1+X_2 不是 A 的特征向量. ■

在本节的最后,我们给出三阶矩阵的特征值与矩阵之间的一些联系.

***定理 5.3** 设 $\lambda_1,\lambda_2,\lambda_3$ 是矩阵

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

的特征值. 则有

$$\lambda_1+\lambda_2+\lambda_3=a_{11}+a_{22}+a_{33}\quad(\text{称为矩阵 }A\text{ 的迹})$$

$$\lambda_1\lambda_2\lambda_3=|A|. \tag{5.4}$$

证 利用行列式性质(参看 1.1 例 12)

$$|A-\lambda I|=\begin{vmatrix}a_{11}-\lambda&a_{12}&a_{13}\\a_{21}&a_{22}-\lambda&a_{23}\\a_{31}&a_{32}&a_{33}-\lambda\end{vmatrix}$$

$$=|A|-\left(\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}+\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix}+\begin{vmatrix}a_{11}&a_{13}\\a_{31}&a_{33}\end{vmatrix}\right)\lambda$$

$$+(a_{11}+a_{22}+a_{33})\lambda^2-\lambda^3,$$

又因 A 的特征值是 $\lambda_1,\lambda_2,\lambda_3$,应有

$$|A-\lambda I|=(\lambda_1-\lambda)(\lambda_2-\lambda)(\lambda_3-\lambda)$$

$$=\lambda_1\lambda_2\lambda_3-(\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_3\lambda_1)\lambda$$

$$+ (\lambda_1 + \lambda_2 + \lambda_3)\lambda^2 - \lambda^3,$$

比较特征多项式的系数,就有

$$\lambda_1 + \lambda_2 + \lambda_3 = a_{11} + a_{22} + a_{33},$$

$$\lambda_1\lambda_2\lambda_3 = |\boldsymbol{A}|. \qquad \blacksquare$$

从定理 5.3,**我们看到**

$|\boldsymbol{A}| \neq 0 \qquad\quad$ $\boldsymbol{A}$ **的特征值全不为** 0 .

这样,我们又有一个判断矩阵可逆的充分必要条件 .

推论 1　矩阵 $\boldsymbol{A}$ 可逆的充分必要条件是 $\boldsymbol{A}$ 的特征值全不为 0 .(**其证明也可如例** 8)

如 $\boldsymbol{A}$ 是 n 阶矩阵,则其特征多项式是 λ 的 n 次多项式,作为 n 次方程,当 $n \geq 3$ 时方程的根是很难求的 . 因此,对特征多项式最好像例 4 **那样,先利用行列式的性质化简 $|\boldsymbol{A} - \lambda\boldsymbol{I}|$,提取出 λ 的因子再来求根 . 而在实际问题中,一般要用近似计算的方法来求特征值 .**

5.2　相 似 矩 阵

在 5.1 **的例** 2 **中,对矩阵**

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

我们已知 $\boldsymbol{A}$ 的特征值是 2 **和** 3,**而** $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ **和** $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ **是相应的特征向量 . 即**

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

用分块矩阵可把上述结果写成

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

定理 5.2 保证了这两个特征向量线性无关,因此

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

必定是可逆矩阵 . 那么

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

易见

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

一般地,可以有

$$\begin{aligned} A^n &= \boldsymbol{P}\begin{pmatrix} 2^n & \\ & 3^n \end{pmatrix}\boldsymbol{P}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 2^n & \\ & 3^n \end{pmatrix}\begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2^{n+1} - 3^n & 2^n - 3^n \\ -2^{n+1} + 2 \cdot 3^n & -2^n + 2 \cdot 3^n \end{pmatrix}, \end{aligned}$$

从这里我们看到,如果能利用特征值和特征向量把矩阵 $\boldsymbol{A}$ 改写成 $\boldsymbol{P}^{-1}\boldsymbol{BP}$ 的形式,对矩阵 $\boldsymbol{A}$ 的计算有时是方便的,现引入相似矩阵的概念 .

定义 5.2 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 n 阶矩阵,如果存在可逆矩阵 $\boldsymbol{P}$,使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{B}, \tag{5.5}$$

则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似,记作 $\boldsymbol{A} \sim \boldsymbol{B}$.

矩阵的相似具有:

反身性: $\boldsymbol{A} \sim \boldsymbol{A}$;

对称性: 如 $\boldsymbol{A} \sim \boldsymbol{B}$,则 $\boldsymbol{B} \sim \boldsymbol{A}$;

传递性: 如 $\boldsymbol{A} \sim \boldsymbol{B}$, $\boldsymbol{B} \sim \boldsymbol{C}$,则 $\boldsymbol{A} \sim \boldsymbol{C}$.

数学上把具有反身性、对称性和传递性称为是一个等价关系,矩阵

的相似就是一个等价关系 .

定理 5.4 如 $A \sim B$,则 A 与 B 有相同的特征值 .

证 因为 $A \sim B$,故存在可逆矩阵 P 使得 $P^{-1}AP = B$,那么

$$
\begin{aligned}
|B - \lambda I| &= |P^{-1}AP - \lambda I| = |P^{-1}(A - \lambda I)P| \\
&= |P^{-1}||A - \lambda I||P| = |A - \lambda I|,
\end{aligned}
$$

这表明 A 和 B 有相同的特征多项式,所以它们有相同的特征值 . ■

推论 2 如果 A 和对角矩阵 Λ

$$
\Lambda = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}
$$

相似,则 $a_1, a_2, \cdots, a_n$ 就是 A 的 n 个特征值 .

证 由 $A \sim \Lambda$,按定理 5.4,有

$$
|A - \lambda I| = |\Lambda - \lambda I| = \begin{vmatrix} a_1 - \lambda & & & \\ & a_2 - \lambda & & \\ & & \ddots & \\ & & & a_n - \lambda \end{vmatrix}
$$

$$
= (a_1 - \lambda)(a_2 - \lambda)\cdots(a_n - \lambda).
$$

所以,对角元素 $a_1, a_2, \cdots, a_n$ 就是 A 的特征值 . ■

注意 定理 5.4 的逆命题不成立,也就是说如 A 与 B 特征值相同,不一定有 A 与 B 相似 . 例如, $A = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$, A 和 B 的特征值是相同的,但是 A 和 B 不相似 . 因为

$$
\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix} = B = P^{-1}AP = P^{-1}\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}P = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}
$$

是不可能的 .

例 1 如 A 可逆,证明: $AB \sim BA$.

证 由于 A 可逆,有

$$A^{-1}(AB)A=BA,$$

按相似定义知:$AB\sim BA$. ■

例 2 如 $A\sim B$,证明:$A^2\sim B^2$.

证 由于 $A\sim B$,有可逆矩阵 P 使 $P^{-1}AP=B$,那么

$$B^2=(P^{-1}AP)(P^{-1}AP)=P^{-1}A^2P,$$

所以, $A^2\sim B^2$. ■

例 3 如 $A\sim B$, $C\sim D$,证明 $\begin{pmatrix}A & \\ & C\end{pmatrix}\sim\begin{pmatrix}B & \\ & D\end{pmatrix}$.

证 由 $A\sim B$, $C\sim D$,知存在可逆矩阵 P 和 Q 使 $P^{-1}AP=B$, $Q^{-1}CQ=D$.

又因 $\begin{pmatrix}P & \\ & Q\end{pmatrix}$ 可逆,且 $\begin{pmatrix}P & \\ & Q\end{pmatrix}^{-1}=\begin{pmatrix}P^{-1} & \\ & Q^{-1}\end{pmatrix}$. 于是

$$\begin{pmatrix}P & \\ & Q\end{pmatrix}^{-1}\begin{pmatrix}A & \\ & C\end{pmatrix}\begin{pmatrix}P & \\ & Q\end{pmatrix}=\begin{pmatrix}P^{-1} & \\ & Q^{-1}\end{pmatrix}\begin{pmatrix}A & \\ & C\end{pmatrix}\begin{pmatrix}P & \\ & Q\end{pmatrix}$$

$$=\begin{pmatrix}P^{-1}AP & \\ & Q^{-1}CQ\end{pmatrix}=\begin{pmatrix}B & \\ & D\end{pmatrix},$$

所以, $\begin{pmatrix}A & \\ & C\end{pmatrix}\sim\begin{pmatrix}B & \\ & D\end{pmatrix}$. ■

例 4 已知 $A\sim\Lambda=\begin{pmatrix}2 & \\ & 3\end{pmatrix}$,求 $|A+I|$

解 由 $A\sim\Lambda$,故存在可逆矩阵 P 使 $P^{-1}AP=\Lambda$ 成立. 那么,$A=P\Lambda P^{-1}$,于是

$$|A+I|=|P\Lambda P^{-1}+I|=|P(\Lambda+I)P^{-1}|$$

$$=|P||\Lambda+I||P^{-1}|=|\Lambda+I|$$

$$= \begin{vmatrix} 3 & \\ & 4 \end{vmatrix} = 12.$$ ■

*例 5　证明 $A = \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix}$ 与 $B = \begin{pmatrix} a_2 & & \\ & a_1 & \\ & & a_3 \end{pmatrix}$ 相似 .

证　取 $P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, 有 $P^{-1} = P$, 则

$$P^{-1}AP = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & a_2 & 0 \\ a_1 & 0 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_2 & & \\ & a_1 & \\ & & a_3 \end{pmatrix} = B,$$

所以, $A \sim B$. ■

5.3　矩阵可对角化的条件

矩阵的相似是一个等价关系, 对矩阵 A 可以有许多矩阵和它相似, 现在我们关心的问题是在和 A 相似的那些矩阵中是否有对角矩阵? 换句话讲 A 能否和一个对角矩阵相似?

如果矩阵 A 和一个对角矩阵相似, 就称矩阵 A 可对角化 . 本节就是来讨论矩阵 A 满足什么条件时, 它可对角化 .

设 A 是三阶矩阵 . 如果 $A \sim \Lambda$, 其中

$$\Lambda = \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix}$$

由相似定义,存在可逆矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\Lambda$,则有:$\boldsymbol{AP}=\boldsymbol{P}\Lambda$
对可逆矩阵 $\boldsymbol{P}$ 按列分块,我们有

$$\boldsymbol{A}(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)=\boldsymbol{AP}=\boldsymbol{P}\Lambda=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)\begin{pmatrix}a_1 & & \\ & a_2 & \\ & & a_3\end{pmatrix}$$

$$=(a_1\boldsymbol{X}_1,a_2\boldsymbol{X}_2,a_3\boldsymbol{X}_3),$$

因此,$\boldsymbol{AX}_1=a_1\boldsymbol{X}_1$, $\boldsymbol{AX}_2=a_2\boldsymbol{X}_2$, $\boldsymbol{AX}_3=a_3\boldsymbol{X}_3$.
由于 $\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)$是可逆矩阵,所以,$\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 都是非零向量,且它们是线性无关的.这样,$\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 分别是 a_1,a_2,a_3 所对应的特征向量.也就是说,如果 $\boldsymbol{A}\sim\Lambda$,则 $\boldsymbol{A}$ 有三个线性无关的特征向量.

反之,如 $\boldsymbol{A}$ 有三个线性无关的特征向量 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$,设相对应的特征值是 $\lambda_1,\lambda_2,\lambda_3$.即 $\boldsymbol{AX}_i=\lambda_i\boldsymbol{X}_i\ (i=1,2,3)$,那么

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)&=(\boldsymbol{AX}_1,\boldsymbol{AX}_2,\boldsymbol{AX}_3)\\&=(\lambda_1\boldsymbol{X}_1,\lambda_2\boldsymbol{X}_2,\lambda_3\boldsymbol{X}_3)\\&=(\boldsymbol{X}_1\ \boldsymbol{X}_2\ \boldsymbol{X}_3)\begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{pmatrix},\end{aligned}$$

因为 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 线性无关,故

$$\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)$$

是可逆矩阵,于是

$$\boldsymbol{AP}=\boldsymbol{P}\begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{pmatrix},\quad 即\ \boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{pmatrix}.$$

所以,$\boldsymbol{A}\sim\Lambda$.

上述分析可推广到 n 阶矩阵,这就是

定理 5.5 n 阶矩阵 $\boldsymbol{A}$ 可对角化(即 $\boldsymbol{A}\sim\Lambda$)的充分必要条件

是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量 . ■

再利用定理 5.2,可有

推论 3 如 n 阶矩阵 $\boldsymbol{A}$ 有 n 个不同的特征值,则 $\boldsymbol{A}$ 可对角化,即 $\boldsymbol{A}\sim\Lambda$. ■

对 n 阶矩阵 $\boldsymbol{A}$,当 $\boldsymbol{A}$ 的特征值有重根时,$\boldsymbol{A}$ 的线性无关的特征向量的个数可能有 n 个,也可能是小于 n 个(如 5.1 中例 4,例 3),这表明不是任何 n 阶矩阵都可以对角化 . 那么,当 $\boldsymbol{A}$ 的特征值有重根时,$\boldsymbol{A}$ 怎样才能对角化呢?

定理 5.6 如 $\boldsymbol{A}$ 是 n 阶矩阵,λ_1 是 $\boldsymbol{A}$ 的 m 重特征值,则属于 λ_1 的线性无关的特征向量的个数不超过 m 个 .(证明较难,略去) ■

我们设想 $\boldsymbol{A}$ 是一个 5 阶矩阵 . λ_1 是其三重特征值,λ_2 是 $\boldsymbol{A}$ 的二重特征值 . 如若 λ_1 只有两个线性无关的特征向量,按定理 5.6,那么 λ_2 至多也只有两个线性无关的特征向量 . 于是,$\boldsymbol{A}$ 最多有四个线性无关的特征向量,这样 $\boldsymbol{A}$ 就不可能对角化 . 从这得到矩阵可对角化的充分必要条件:

推论 4 矩阵 $\boldsymbol{A}$ 可对角化的充分必要条件是 $\boldsymbol{A}$ 的每个特征值中,线性无关的特征向量的个数恰好等于该特征值的重数 . ■

例 1 如 $\boldsymbol{A}=\begin{pmatrix}1&2\\5&4\end{pmatrix}$,问 $\boldsymbol{A}$ 是否与对角矩阵相似?若与对角矩阵相似,就求对角矩阵 Λ 及可逆矩阵 $\boldsymbol{P}$,使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\Lambda$.

解 先求 $\boldsymbol{A}$ 的特征多项式

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}1-\lambda&2\\5&4-\lambda\end{vmatrix}=\lambda^2-5\lambda-6=(\lambda-6)(\lambda+1).$$

所以,$\boldsymbol{A}$ 的特征值是 $\lambda_1=6$, $\lambda_2=-1$.

由于 $\boldsymbol{A}$ 有两个不同的特征值,按推论 3 知 $\boldsymbol{A}\sim\Lambda$.

下面求 $\boldsymbol{A}$ 的特征向量 .

当 $\lambda_1=6$ 时,由 $(\boldsymbol{A}-6\boldsymbol{I})\boldsymbol{x}=0$,即

$$\begin{pmatrix} -5 & 2 \\ 5 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

得到特征向量 $X_1 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$，

当 $\lambda_2 = -1$ 时，由 $(A + I)x = 0$，即

$$\begin{pmatrix} 2 & 2 \\ 5 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

得到特征向量 $X_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

因为 $AX_1 = 6X_1$，$AX_2 = -X_2$ 且 X_1，X_2 线性无关，故可令

$$P = (X_1\ X_2) = \begin{pmatrix} 2 & 1 \\ 5 & -1 \end{pmatrix}，有\ P^{-1} = \frac{1}{7}\begin{pmatrix} 1 & 1 \\ 5 & -2 \end{pmatrix},$$

这样

$$P^{-1}AP = \frac{1}{7}\begin{pmatrix} 1 & 1 \\ 5 & -2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 5 & -1 \end{pmatrix} = \begin{pmatrix} 6 & \\ & -1 \end{pmatrix},$$

即有
$$A \sim \Lambda = \begin{pmatrix} 6 & \\ & -1 \end{pmatrix}.$$
■

例 2　证明：$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$ 不可能对角化.

证　A 的特征多项式是

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 \\ -1 & 3-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2,$$

所以，A 的特征值是 $\lambda = 2$(二重根).

当 $\lambda = 2$ 时，由 $(A - 2I)x = 0$，即

$$\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

得到 A 的线性无关的特征向量是 $X = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

由于 $\lambda = 2$(二重根)只有一个线性无关的特征向量．据推论 4, A 不可能对角化． ■

例 3 求可逆矩阵 P,使

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

相似于对角矩阵 Λ.

解 $\boldsymbol{A}$ 的特征多项式为

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} -\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix}$$

$$= (1-\lambda)(\lambda^2 - 1);$$

所以, $\boldsymbol{A}$ 的特征值是 $\lambda_1 = \lambda_2 = 1, \lambda_3 = -1$.

当 $\lambda_1 = \lambda_2 = 1$ 时,由 $(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{x} = 0$,即

$$\begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

得到特征向量 $\boldsymbol{X}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\boldsymbol{X}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$,

当 $\lambda_3 = -1$ 时,由 $(\boldsymbol{A} + \boldsymbol{I})\boldsymbol{x} = 0$,即

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

得到特征向量 $\boldsymbol{X}_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

$\boldsymbol{A}$ 有三个线性无关的特征向量, $\boldsymbol{A}$ 可对角化．

当取

$$\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)=\begin{pmatrix}1&0&1\\0&1&0\\1&0&-1\end{pmatrix}$$

时,我们就有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\frac{1}{2}&0&\frac{1}{2}\\0&1&0\\\frac{1}{2}&0&-\frac{1}{2}\end{pmatrix}\begin{pmatrix}&&1\\&1&\\1&&\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&0\\1&0&-1\end{pmatrix}=\begin{pmatrix}1&&\\&1&\\&&-1\end{pmatrix},$$

因此, $\boldsymbol{A}\sim\begin{pmatrix}1&&\\&1&\\&&-1\end{pmatrix}$. ■

例 4 已知 $\boldsymbol{A}=\begin{pmatrix}1&1&1\\0&0&x\\0&0&0\end{pmatrix}$ 有三个线性无关的特征向量,求 x 的值.

解 $\boldsymbol{A}$ 的特征多项式

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}1-\lambda&1&1\\0&-\lambda&x\\0&0&-\lambda\end{vmatrix}=\lambda^2(1-\lambda)$$

得到 $\boldsymbol{A}$ 的特征值是 $\lambda_1=1$, $\lambda_2=\lambda_3=0$.

由于 $\boldsymbol{A}$ 有三个线性无关的特征向量,故当 $\lambda_2=\lambda_3=0$ 时, $\boldsymbol{A}$ 应有两个线性无关的特征向量.即方程组 $(\boldsymbol{A}-0\boldsymbol{I})\boldsymbol{x}=0$ 的基础解系应由两个向量组成,那么系数矩阵的秩 $r(\boldsymbol{A}-0\boldsymbol{I})=1$,也就是 $r(\boldsymbol{A})=1$. 所以, $x=0$. ■

例 5 已知 $\boldsymbol{A}$ 的特征值是 2 和 3,相应的特征向量是 $\begin{pmatrix}1\\1\end{pmatrix}$ 和

$\begin{pmatrix}3\\2\end{pmatrix}$，求矩阵 $\boldsymbol{A}$.

解　由于特征向量是二维向量，知 $\boldsymbol{A}$ 是二阶矩阵，现在 $\boldsymbol{A}$ 有两个不同的特征值. 故 A 可以对角化为

$$\boldsymbol{A}\sim \Lambda = \begin{pmatrix}2 & \\ & 3\end{pmatrix},$$

这时的可逆矩阵 $\boldsymbol{P}$ 由特征向量构成

$$\boldsymbol{P} = \begin{pmatrix}1 & 3\\ 1 & 2\end{pmatrix},$$

而 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \Lambda$，这就有

$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{P}\Lambda\boldsymbol{P}^{-1}\\ &= \begin{pmatrix}1 & 3\\ 1 & 2\end{pmatrix}\begin{pmatrix}2 & \\ & 3\end{pmatrix}\begin{pmatrix}-2 & 3\\ 1 & -1\end{pmatrix} = \begin{pmatrix}5 & -3\\ 2 & 0\end{pmatrix}.\end{aligned}$$ ■

*例 6　如果每个非零三维向量都是三阶矩阵 $\boldsymbol{A}$ 的特征向量，试证 $\boldsymbol{A}$ 是一个数量矩阵.

证　因为任何 3 维非零向量都是 $\boldsymbol{A}$ 的特征向量，所以 $\boldsymbol{A}$ 有三个线性无关的特征向量，那么 $\boldsymbol{A}$ 可化为对角形，设

$$\boldsymbol{A}\sim \begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{pmatrix},$$

其中 $\lambda_1,\lambda_2,\lambda_3$ 是 $\boldsymbol{A}$ 的特征值.

如果 $\lambda_1\neq\lambda_2$，设 $\boldsymbol{X}_1$，$\boldsymbol{X}_2$ 分别是 $\boldsymbol{A}$ 属于 λ_1,λ_2 的特征向量，那么 $\boldsymbol{X}_1$，$\boldsymbol{X}_2$ 线性无关(定理 5.2)，所以 $\boldsymbol{X}_1+\boldsymbol{X}_2\neq 0$，一方面按已知条件 $\boldsymbol{X}_1+\boldsymbol{X}_2$ 应是 $\boldsymbol{A}$ 的一个特征向量，另一方面，$\lambda_1\neq\lambda_2$ 时，$\boldsymbol{X}_1+\boldsymbol{X}_2$ 不能是 $\boldsymbol{A}$ 的特征向量(参看 5.1 中例 9)二者矛盾. 所以，必有 $\lambda_1=\lambda_2=\lambda_3$. 因此，$\boldsymbol{A}\sim\lambda_1\boldsymbol{I}$. 从而

$$\boldsymbol{A} = \boldsymbol{P}(\lambda_1\boldsymbol{I})\boldsymbol{P}^{-1} = \lambda_1\boldsymbol{I}$$

即 A 是一个数量矩阵 . ■

这几个例题从不同的角度说明了矩阵对角化的判断方法,以及对能对角化的矩阵如何实现其对角化 . 我们注意到:一个矩阵如果可以化为对角矩阵,那么对角矩阵的对角线上的元素就是该矩阵的特征值,而实现对角化的可逆矩阵 P 的每一列就是相对应的特征向量 . 当然,在对角矩阵中,特征值排列的次序是任意的,可逆矩阵 P 中特征向量的选取方法也是不唯一的,这些是本节的核心,望读者仔细体会 .

矩阵化为对角矩阵的问题在许多领域都有着广泛的应用,下面的例题介绍在计算矩阵的方幂方面的情况,在附录中还有求解线性常系数微分方程组的应用,供有兴趣有余力的读者参考 .

*例 7 已知 $A=\begin{pmatrix}0 & -2\\ 1 & 3\end{pmatrix}$,求 A^n

解 先求 A 的特征值,由

$$|A-\lambda I|=\begin{vmatrix}-\lambda & -2\\ 1 & 3-\lambda\end{vmatrix}$$

$$=\lambda^2-3\lambda+2=(\lambda-1)(\lambda-2),$$

得到 A 的特征值是 $\lambda_1=1,\lambda_2=2$.

再求 A 的特征向量:

当 $\lambda_1=1$ 时,由 $(A-I)x=0$, 得 $X_1=\begin{pmatrix}2\\ -1\end{pmatrix}$,

当 $\lambda_2=2$ 时,由 $(A-2I)x=0$, 得 $X_2=\begin{pmatrix}-1\\ 1\end{pmatrix}$.

构造可逆矩阵

$$P=(X_1\ X_2)=\begin{pmatrix}2 & -1\\ -1 & 1\end{pmatrix},$$

则 $A=P\Lambda P^{-1}$,那么

$$A^2=(P\Lambda P^{-1})(P\Lambda P^{-1})=P\Lambda^2P^{-1},$$

以此递推

$$\begin{aligned} \boldsymbol{A}^n &= (\boldsymbol{P\Lambda P}^{-1})(\boldsymbol{P\Lambda P}^{-1})\cdots(\boldsymbol{P\Lambda P}^{-1}) \\ &= \boldsymbol{P\Lambda}^n\boldsymbol{P}^{-1} \\ &= \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & \\ & 2^n \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2-2^n & 2-2^{n+1} \\ -1+2^n & -1+2^{n+1} \end{pmatrix}. \end{aligned}$$

■

5.4　实对称矩阵的对角化

在上一节已看到有些 n 阶矩阵可以和对角矩阵相似,而也有些 n 阶矩阵不能对角化,本节专门讨论实对称矩阵,主要的结论是任一个实对称矩阵 $\boldsymbol{A}$ 一定可以和一个对角矩阵相似,并且可找到正交矩阵 $\boldsymbol{T}$,使 $\boldsymbol{T}^{-1}\boldsymbol{AT}$ 是对角矩阵. 作为 n 阶实对称矩阵,通常 n 阶矩阵有关特征值、特征向量的性质它都具备,并且它还有一些特殊的性质.

定理 5.7　实对称矩阵 A 的特征值都是实数.

*证　设 λ 是 $\boldsymbol{A}$ 的任一特征值,$\boldsymbol{X}$ 是相应的特征向量. 记 $\overline{\boldsymbol{X}}$ 是将 $\boldsymbol{X}$ 的所有分量都取共轭复数后所得到的向量,记 $\overline{\lambda}$ 是 λ 的共轭复数.

根据 $\boldsymbol{AX}=\lambda\boldsymbol{X}$ 及 $\overline{\boldsymbol{A}}=\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$(实对称)并利用复数运算的性质,我们取共轭

$$\boldsymbol{A}\overline{\boldsymbol{X}}=\overline{\boldsymbol{A}}\,\overline{\boldsymbol{X}}=\overline{\boldsymbol{AX}}=\overline{\lambda\boldsymbol{X}}=\overline{\lambda}\,\overline{\boldsymbol{X}},$$

对上式两端同时用 $\boldsymbol{X}^{\mathrm{T}}$ 左乘,就有

$$\overline{\lambda}\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\overline{\boldsymbol{X}}=(\boldsymbol{AX})^{\mathrm{T}}\overline{\boldsymbol{X}}=(\lambda\boldsymbol{X})^{\mathrm{T}}\overline{\boldsymbol{X}}=\lambda\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}},$$

从而得到

$$(\overline{\lambda}-\lambda)\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}}=0,$$

由于特征向量 $\boldsymbol{X}\neq 0$,故知

$$\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}} = x_1\overline{x_1} + x_2\overline{x_2} + \cdots + x_n\overline{x_n}$$
$$= \| x_1 \|^2 + \| x_2 \|^2 + \cdots + \| x_n \|^2 > 0,$$

因此必有 $\overline{\lambda} - \lambda = 0$, 即 $\lambda = \overline{\lambda}$, 也就是 $\boldsymbol{A}$ 的特征值是实数. ■

对一个实矩阵,虽然其特征多项式是实系数多项式,但它的特征值有可能是复数,相应的特征向量也就可能是复向量(如 5.1 中例 5)。但对于实对称矩阵来说,由定理 5.7 知其特征值一定是实数,那么 $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = 0$ 必是实系数方程组,就总可取实向量作为其特征向量.

作为实对称矩阵,不同特征值所对应的特征向量不仅是线性无关的(定理 5.2),而且有进一步的结果:它们是正交的.

定理 5.8 实对称矩阵 $\boldsymbol{A}$ 的不同特征值 λ_1, λ_2 所对应的特征向量 $\boldsymbol{X}_1, \boldsymbol{X}_2$ 是正交的.

证 对 $\boldsymbol{A}\boldsymbol{X}_1 = \lambda_1\boldsymbol{X}_1$, $\boldsymbol{A}\boldsymbol{X}_2 = \lambda_2\boldsymbol{X}_2$, 由

$$\lambda_1\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_1 = \boldsymbol{X}_2^{\mathrm{T}}(\lambda_1\boldsymbol{X}_1) = \boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}_1 = (\boldsymbol{A}\boldsymbol{X}_2)^{\mathrm{T}}\boldsymbol{X}_1$$
$$= (\lambda_2\boldsymbol{X}_2)^{\mathrm{T}}\boldsymbol{X}_1 = \lambda_2\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_1,$$

因此 $(\lambda_1 - \lambda_2)\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_1 = 0$.

由于 $\lambda_1 \neq \lambda_2$, 故 $\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_1 = 0$, 即 $(\boldsymbol{X}_2, \boldsymbol{X}_1) = 0$,

所以, $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 正交. ■

例 1 求 $\boldsymbol{A} = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$ 的特征值、特征向量

解 $\boldsymbol{A}$ 的特征多项式是

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = \begin{vmatrix} 2-\lambda & 3 \\ 3 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda - 5 = (\lambda - 5)(\lambda + 1),$$

$\boldsymbol{A}$ 的特征值是 $\lambda_1 = 5$, $\lambda_2 = -1$.

当 $\lambda_1 = 5$ 时,由 $(\boldsymbol{A} - 5\boldsymbol{I})\boldsymbol{x} = 0$. 即

$$\begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

得到 $\lambda_1 = 5$ 所对应的特征向量 $\boldsymbol{X}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

当 $\lambda_2 = -1$ 时,由 $(\boldsymbol{A} + \boldsymbol{I})\boldsymbol{x} = 0$,即

$$\begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

得到 $\lambda_2 = -1$ 所对应的特征向量 $\boldsymbol{X}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. ■

不难验证:$(\boldsymbol{X}_1, \boldsymbol{X}_2) = 0$,即不同特征值所对应的特征向量是正交的.

请注意,在例 1 中,如果把正交的特征向量 $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ 单位化为

$$\gamma_1 = \frac{\boldsymbol{X}_1}{\|\boldsymbol{X}_1\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \gamma_2 = \frac{\boldsymbol{X}_2}{\|\boldsymbol{X}_2\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

构造可逆矩阵

$$\boldsymbol{T} = (\gamma_1, \gamma_2) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix},$$

由于 γ_1, γ_2 仍是 $\boldsymbol{A}$ 的特征向量(定理 5.1),即

$$\boldsymbol{A}\gamma_1 = 5\gamma_1, \quad \boldsymbol{A}\gamma_2 = -\gamma_2,$$

我们就有

$$\boldsymbol{A}(\gamma_1, \gamma_2) = (\gamma_1, \gamma_2)\begin{pmatrix} 5 & \\ & -1 \end{pmatrix},$$

因此

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \Lambda = \begin{pmatrix} 5 & \\ & -1 \end{pmatrix},$$

显然,$\boldsymbol{T}\boldsymbol{T}^{\mathrm{T}} = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{T} = \boldsymbol{I}$. 所以,实对称矩阵 $\boldsymbol{A}$ 不仅能化为对角矩阵. 而且可选取可逆矩阵为正交矩阵来实现对角化.

更加一般地,对于 n 阶实对称矩阵 $\boldsymbol{A}$,如果它有 n 个不同的

特征值$\lambda_1,\lambda_2,\cdots,\lambda_n$,根据定理5.8,它们对应的特征向量$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$是互相正交的.因此,只须把这些特征向量单位化为$\gamma_1,\gamma_2,\cdots,\gamma_n$.由此构造正交矩阵$T=(\gamma_1,\gamma_2,\cdots,\gamma_n)$,就可得到

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}.$$

如果$\boldsymbol{A}$的特征值λ_i是k重根,对方程组$(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{x}=0$,必可求出$k$个线性无关的解$\boldsymbol{X}_{i1},\boldsymbol{X}_{i2},\cdots,\boldsymbol{X}_{ik}$.根据定理5.1,这些特征向量的线性组合还是$\lambda_i$所对应的特征向量,故可用施密特正交化方法,做出互相正交的特征向量$\beta_1,\beta_2,\cdots,\beta_k$,然后再单位化.把每个不同的特征值经如此处理的特征向量放在一起,按定理5.8它们可构造成一个正交矩阵$\boldsymbol{T}$,同样可有

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}.$$

总之,我们可以证明如下结果:

定理5.9　对任一n阶实对称矩阵$\boldsymbol{A}$,存在正交矩阵$\boldsymbol{T}$,使得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{T}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix},\tag{5.6}$$

其中$\lambda_1,\lambda_2,\cdots,\lambda_n$是$\boldsymbol{A}$的特征值.　■

例2　求正交矩阵$\boldsymbol{T}$,化

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

为对角形.

解 A 的特征多项式

$$|A-\lambda I|=\begin{vmatrix}1-\lambda & 1 & 1\\ 1 & 1-\lambda & 1\\ 1 & 1 & 1-\lambda\end{vmatrix}$$

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

$$=(3-\lambda)\begin{vmatrix}1 & 1 & 1\\ 1 & 1-\lambda & 1\\ 1 & 1 & 1-\lambda\end{vmatrix}=(3-\lambda)\begin{vmatrix}1 & 1 & 1\\ & -\lambda & \\ & & -\lambda\end{vmatrix}$$

$$=\lambda^2(3-\lambda),$$

得到 A 的特征值: $\lambda_1=3$, $\lambda_2=\lambda_3=0$.

当 $\lambda_1=3$ 时,由 $(A-3I)x=0$,即

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

得特征向量 $X_1=(1,1,1)^{\mathrm{T}}$.

当 $\lambda_2=\lambda_3=0$ 时,由 $(A-0I)x=0$,即

$$\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix},$$

得特征向量 $X_2=(-1,1,0)^{\mathrm{T}}$, $X_3=(-1,0,1)^{\mathrm{T}}$.

易见 X_1 与 X_2, X_3 都正交(不同特征值!),但 X_2 与 X_3 不正交,现对 X_2, X_3 用 Schmidt 正交化方法

令 $\beta_2=X_2=(-1,1,0)^{\mathrm{T}}$, $\beta_3=k\beta_2+X_3$.

由 $(\beta_2,\beta_3)=(\beta_2,k\beta_2+X_3)=k(\beta_2,\beta_2)+(\beta_2,X_3)=0$,

即 $\qquad 2k + 1 = 0$

得 $\quad \beta_3 = -\frac{1}{2}\beta_2 + \boldsymbol{X}_3 = \frac{1}{2}(-1, -1, 2)^{\mathrm{T}}$,

再把 $\boldsymbol{X}_1, \beta_2, \beta_3$ 单位化,得 $\gamma_1, \gamma_2, \gamma_3$,即

$$\gamma_1 = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \gamma_2 = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \gamma_3 = \frac{1}{6}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix},$$

构造正交矩阵

$$\boldsymbol{T} = (\gamma_1 \ \ \gamma_2 \ \ \gamma_3) = \begin{pmatrix} \frac{1}{3} & \frac{-1}{2} & \frac{-1}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{-1}{6} \\ \frac{1}{3} & 0 & \frac{2}{6} \end{pmatrix},$$

我们就有

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} 3 & & \\ & 0 & \\ & & 0 \end{pmatrix}.$$ ■

*例 3 $\quad \boldsymbol{A}$ 是三阶实对称矩阵,2,1,1 是其特征值,对应于 $\lambda = 2$ 的特征向量为 $\boldsymbol{X}_1 = (2, 1, 2)^{\mathrm{T}}$,求 $\boldsymbol{A}$.

解法 1 $\quad A$ 是实对称矩阵它必可对角化,因而对应于 $\lambda = 1$,$\boldsymbol{A}$ 有两个线性无关的特征向量且它们与 $\boldsymbol{X}_1$ 都正交(定理 5.8).设 $(x_1, x_2, x_3)^{\mathrm{T}}$ 是 $\lambda = 1$ 所对应的特征向量,那么

$$2x_1 + x_2 + 2x_3 = 0,$$

求出特征向量

$$\boldsymbol{X}_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \quad \boldsymbol{X}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},$$

构造可逆矩阵 $\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)$，则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\Lambda,$$

即

$$\boldsymbol{A}=\boldsymbol{P}\Lambda\boldsymbol{P}^{-1}=\begin{pmatrix}2&-1&-1\\1&2&0\\2&0&1\end{pmatrix}\begin{pmatrix}2&&\\&1&\\&&1\end{pmatrix}\cdot\frac{1}{9}\begin{pmatrix}2&1&2\\-1&4&-1\\-4&-2&5\end{pmatrix}$$

$$=\frac{1}{9}\begin{pmatrix}13&2&4\\2&10&2\\4&2&13\end{pmatrix}.$$ ■

解法 2　对 $\boldsymbol{A}$ 我们可用正交矩阵来实现对角化．只须对解法一中的 $\boldsymbol{X}_2,\boldsymbol{X}_3$ 用 Schmidt 正交化，令 $\beta_2=\boldsymbol{X}_2,\ \beta_3=k\beta_2+\boldsymbol{X}_3$

由

$$(\beta_2,\beta_3)=(\beta_2,k\beta_2+\boldsymbol{X}_3)=k(\beta_2,\beta_2)+(\beta_2,\boldsymbol{X}_3)=0$$

得

$$5k+1=0,$$

有

$$\beta_3=-\frac{1}{5}\beta_2+\boldsymbol{X}_3=\frac{1}{5}(-4,-2,5)^{\mathrm{T}},$$

再单位化 $\boldsymbol{X}_1,\beta_2,\beta_3$，构造正交矩阵

$$\boldsymbol{T}=\begin{pmatrix}\frac{2}{3}&\frac{-1}{\sqrt{5}}&\frac{-4}{3\sqrt{5}}\\\frac{1}{3}&\frac{2}{\sqrt{5}}&\frac{-2}{3\sqrt{5}}\\\frac{2}{3}&0&\frac{5}{3\sqrt{5}}\end{pmatrix},$$

这时，$\boldsymbol{T}^{-1}=\boldsymbol{T}^{\mathrm{T}}$，

$$\boldsymbol{A}=\boldsymbol{T}\Lambda\boldsymbol{T}^{-1}$$

$$
= \begin{pmatrix} \frac{2}{3} & \frac{-1}{\sqrt{5}} & \frac{-4}{3\sqrt{5}} \\ \frac{1}{3} & \frac{2}{\sqrt{5}} & \frac{-2}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{pmatrix} \begin{pmatrix} 2 & & \\ & 1 & \\ & & 1 \end{pmatrix} \times \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ \frac{-4}{3\sqrt{5}} & \frac{-2}{3\sqrt{5}} & \frac{5}{3\sqrt{5}} \end{pmatrix}
$$

$$
= \frac{1}{9}\begin{pmatrix} 13 & 2 & 4 \\ 2 & 10 & 2 \\ 4 & 2 & 13 \end{pmatrix}. \qquad \blacksquare
$$

* **例 4** A **是三阶实对称矩阵**, $A^2 = A$, **秩** $r(\boldsymbol{A}) = 2$, **求** $|\boldsymbol{A} + 3\boldsymbol{I}|$

解 **由** $\boldsymbol{A}^2 = \boldsymbol{A}$ **知** $\boldsymbol{A}$ **的特征值是** 1 **和** 0 **(参看** 5.1 **例** 6**). 因** $\boldsymbol{A}$ **是实对称矩阵,** $\boldsymbol{A}$ **可对角化, 故** $\boldsymbol{A}$ **有三个线性无关的特征向量.**

对 $\lambda = 0$, **由** $r(\boldsymbol{A} - 0\boldsymbol{I}) = r(\boldsymbol{A}) = 2$, **知** $\lambda = 0$ **有一个线性无关的特征向量. 所以,** $\lambda = 0$ **是** $\boldsymbol{A}$ **的特征值单根. 那么,** $\lambda = 1$ **就必是** $\boldsymbol{A}$ **的二重特征根. 根据定理** 5.9**, 存在正交矩阵** $\boldsymbol{T}$ **使得:**

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

那么

$$
\boldsymbol{A} + 3\boldsymbol{I} = \boldsymbol{T}\begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix}\boldsymbol{T}^{-1} + 3\boldsymbol{I}
$$

$$= T\begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix} T^{-1} + T(3I)T^{-1}$$

$$= T\left[\begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix} + \begin{pmatrix} 3 & & \\ & 3 & \\ & & 3 \end{pmatrix}\right] T^{-1}$$

$$= T\begin{pmatrix} 3 & & \\ & 4 & \\ & & 4 \end{pmatrix} T^{-1},$$

所以

$$|A + 3I| = |T| \begin{vmatrix} 3 & & \\ & 4 & \\ & & 4 \end{vmatrix} |T^{-1}| = 48 .$$ ■

习　题　5

5.1　求矩阵的特征值、特征向量

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

(4) $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$,(5) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$,　(6) $\begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix}$.

5.2　求矩阵 A 的特征值

(1) 如 $A^2 = I$;

(2) 如 $A + 2I = B$, 而 B 的特征值是 5;

(3) 如 $A^2 + 2A - 3I = 0$.

5.3　如果 A 是可逆矩阵,试讨论 A, A^*, A^{-1}三个矩阵特征

值之间的关系(提示:用定义,注意 $\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}$).

5.4 证明: $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 有相同的特征值.

5.5 已知 $\boldsymbol{A}\sim\boldsymbol{B}$,证明 $\boldsymbol{A}$ 可逆的充分必要条件是 $\boldsymbol{B}$ 可逆.

5.6 选择题

(1) 三阶矩阵 $\boldsymbol{A}$ 的特征值是 1, 2, 3, 相应的特征向量依次是 $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3$, 设 $\boldsymbol{P}=(\boldsymbol{X}_3, \boldsymbol{X}_2, \boldsymbol{X}_1)$. 则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=$____

(A) $\begin{pmatrix}1&&\\&2&\\&&3\end{pmatrix}$, (B) $\begin{pmatrix}2&&\\&3&\\&&1\end{pmatrix}$, (C) $\begin{pmatrix}3&&\\&1&\\&&2\end{pmatrix}$, (D) $\begin{pmatrix}3&&\\&2&\\&&1\end{pmatrix}$.

(2) 已知 $\boldsymbol{A}$ 可逆, $\lambda=2$ 是 $\boldsymbol{A}$ 的一个特征值, 则矩阵 $\left(\frac{1}{3}\boldsymbol{A}^2\right)^{-1}$有一个特征值等于____

(A) $\frac{4}{3}$, (B) $\frac{3}{4}$, (C) $\frac{1}{2}$, (D) $\frac{1}{4}$.

(3) 已知三阶矩阵 $\boldsymbol{A}$ 的特征值是 $0, \pm 1$,则下列命题中不正确的是____

(A) 矩阵 $\boldsymbol{A}$ 是不可逆的,

(B) $\boldsymbol{A}$ 和对角矩阵相似,

(C) 1 和 -1 所对应的特征向量是正交的,

(D) $\boldsymbol{A}\boldsymbol{X}=0$ 的基础解系由一个向量组成.

(4) $\boldsymbol{A}$、$\boldsymbol{B}$ 都是 n 阶实对称可逆矩阵,下列命题中不正确的是____

(A) 存在可逆矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{-1}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{P}=\Lambda$,

(B) 存在正交矩阵 $\boldsymbol{T}$,使 $\boldsymbol{T}^{-1}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{T}=\Lambda$,

(C) $\boldsymbol{A}^{*}+\boldsymbol{B}^{*}$ 必可以对角化,

(D) $\boldsymbol{A}\boldsymbol{B}$ 必可以对角化.

5.7 已知 $\boldsymbol{A}\sim\boldsymbol{B}$, $\boldsymbol{A}^2=0$,证明: $\boldsymbol{B}^2=0$.

5.8 已知 $\boldsymbol{A}\sim\boldsymbol{B}$,其中

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

求 x.

5.9　在 1 题中,哪些矩阵可对角化?对可对角化的矩阵求出与其相似的对角矩阵,并写出可逆矩阵 $\boldsymbol{P}$.

5.10　已知 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\Lambda_1$, $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}=\Lambda_2$,其中 Λ_1, Λ_2 是对角矩阵,证明:$\boldsymbol{AB}=\boldsymbol{BA}$.

5.11　$\boldsymbol{A}$ 是 2 阶矩阵,特征值是 2 和 -2,对应的特征向量是 $(1,0)^{\mathrm{T}}$ 和 $(2,1)^{\mathrm{T}}$,求矩阵 $\boldsymbol{A}$.

5.12　已知

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

问 $\boldsymbol{A}$、$\boldsymbol{B}$ 是否相似?为什么?

5.13　求正交矩阵 $\boldsymbol{T}$ 和对角矩阵 Λ,使 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\Lambda$

(1) $\boldsymbol{A}=\begin{pmatrix}1&2\\2&-2\end{pmatrix}$, (2) $\boldsymbol{A}=\begin{pmatrix}&&1\\&1&\\1&&\end{pmatrix}$,

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

5.14　已知 $\boldsymbol{A}\sim\begin{pmatrix}1&\\&0\end{pmatrix}$,求 $|\boldsymbol{A}+2\boldsymbol{I}|$.

***5.15　已知 $\boldsymbol{A}=\begin{pmatrix}2&2\\4&4\end{pmatrix}$,求 $\boldsymbol{A}^n$.(用对角化及 $\mathrm{r}(\boldsymbol{A})=1$ 分别求解,可参看 3.4 中例 3)**

5.16　已知

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

有三个线性无关的特征向量,求参数 t.

*5.17　已知三阶实对称矩阵 $\boldsymbol{A}$ 的特征值是 $\lambda_1=-1$, $\lambda_2=\lambda_3=1$,对应于 λ_1 的特征向量为 $\boldsymbol{x}_1=(0,1,1)^{\mathrm{T}}$,求 $\boldsymbol{A}$.

5.18　三阶矩阵 $\boldsymbol{A}$ 有二重特征值 λ_0,问向量 $\alpha_1=(1,0,1)^{\mathrm{T}}$, $\alpha_2=(1,1,1)^{\mathrm{T}}$, $\alpha_3=(2,-1,2)^{\mathrm{T}}$, $\alpha_4=(0,1,-1)^{\mathrm{T}}$ 能否都是矩阵 $\boldsymbol{A}$ 的属于特征值 λ_0 的特征向量,为什么?

*5.19　三阶矩阵 $\boldsymbol{A}$ 的特征值是 $1,2,-3$, $\boldsymbol{B}=\boldsymbol{A}^2+3\boldsymbol{A}-2\boldsymbol{I}$,求矩阵 $\boldsymbol{B}$ 的特征值, $\boldsymbol{B}$ 能否化为对角形,并说明理由.

*5.20　$\boldsymbol{A}$ 是三阶实对称矩阵, $\boldsymbol{A}^2=\boldsymbol{I}$, 秩 $\mathrm{r}(\boldsymbol{A}-\boldsymbol{I})=1$, 求行列式 $|\boldsymbol{A}-2\boldsymbol{I}|$ 的值.(提示:参看 5.4 例 4)

第 6 章　二　次　型

二次型就是二次齐次多项式．在解析几何中,对二次曲线

$$ax^2 + 2bxy + cy^2 = 1,$$

我们总可通过直角坐标变换把方程化为不含混合项 xy 的标准方程

$$a'x'^2 + b'y'^2 = 1$$

二次齐次式不仅在几何问题中出现,对于二元函数 $z = f(x, y)$ 极值问题的研究就涉及到形如

$$\frac{1}{2!}[f''_{xx}(\Delta x)^2 + 2f''_{xy}\Delta x\Delta y + f''_{yy}(\Delta y)^2]$$

的二次齐次式．类似的问题在数学的其它领域,在物理、力学等学科亦是经常会遇到的．

从代数的观点来看,二次型可以用一个实对称矩阵来表示,而实对称矩阵是可以用正交矩阵化为对角形的,其几何意义就是二次曲线的方程经适当的直角坐标变换可化为标准方程．在上一章,已经讨论了矩阵的特征值,但并未关心到特征值的正负号,现在我们要研究实对称矩阵特征值的符号,引入正定矩阵的概念,它的背景就是函数的极值．

6.1　二次型的矩阵表示

定义 6.1　含有 n 个变量 $x_1, x_2, \cdots, x_n$ 的二次齐次函数

$$f(x_1, x_2, \cdots, x_n)$$
$$= a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3$$

$$+ \cdots + 2a_{1n}x_1x_n + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n$$
$$+ \cdots + 2a_{n-1\,n}x_{n-1}x_n$$

如果系数 a_{ij} 都是实数,就称 $f(x_1, x_2, \cdots, x_n)$是实数域上的 n 元二次型 .

在二次型中,我们规定

$$a_{ij} = a_{ji}, \quad \forall \; i, j = 1, 2, \cdots, n$$

则二次型可用实对称矩阵唯一地表出,例如,对于三元二次型

$$f(x_1, x_2, x_3) = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3$$

可用代数恒等变形把 $f(x_1, x_2, x_3)$改写成

$$\begin{aligned} f(x_1, x_2, x_3) &= a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 \\ &\quad + a_{21}x_1x_2 + a_{22}x_2^2 + a_{23}x_2x_3 \\ &\quad + a_{31}x_1x_3 + a_{32}x_2x_3 + a_{33}x_3^2 \\ &= x_1(a_{11}x_1 + a_{12}x_2 + a_{13}x_3) \\ &\quad + x_2(a_{21}x_1 + a_{22}x_2 + a_{23}x_3) \\ &\quad + x_3(a_{31}x_1 + a_{32}x_2 + a_{33}x_3), \end{aligned}$$

再引进矩阵的运算,我们就有

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

如果记

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

那么二次型 f 就可用矩阵简记为

$$f(x_1, x_2, x_3) = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}, \tag{6.1}$$

其中矩阵 A 是一个实对称矩阵,叫做二次型 $f(x_1, x_2, x_3)$ 的对应矩阵.称(6.1)是二次型的矩阵表示.

要注意的是,给定二次型后,对应矩阵 A 的元素 a_{ii} 正是二次型中平方项 x_i^2 的系数,而 $a_{ij}(i \neq j)$ 恰是混合项 $x_i x_j$ 系数的一半.因此,二次型和它的对应矩阵是相互唯一决定的.

例如,二次型 $f(x_1, x_2, x_3) = x_1^2 - 3x_2^2 + 4x_3^2 + 2x_1x_2 + 6x_2x_3$ 的对应矩阵是

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -3 & 3 \\ 0 & 3 & 4 \end{pmatrix},$$

反之,若二次型的对应矩阵是

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 4 \\ -1 & 4 & 0 \end{pmatrix},$$

则这个二次型是

$$f(x_1, x_2, x_3) = x_1^2 + 3x_2^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3,$$

对二次型的矩阵表示 $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}$,也可以理解为:在 n 维向量空间有一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$,向量 α 在这组基下的坐标是 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$,那么二次型 $f(x_1, x_2, \cdots, x_n) = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}$ 就是向量 α 的 n 个坐标 $x_1, x_2, \cdots, x_n$ 的一个 n 元二次齐次函数.二次型的矩阵 A 是依赖于这组基的选取的.

下面以三元二次型为例,来研究当基向量改变后,二次型的对

应矩阵是如何改变的?对应矩阵之间又有什么样的联系?

设 $\alpha_1,\alpha_2,\alpha_3$ 和 β_1,β_2,β_3 是三维向量空间的两组基,由基 $\alpha_1,\alpha_2,\alpha_3$ 到基 β_1,β_2,β_3 的过渡矩阵是 $\boldsymbol{C}$(注:$\boldsymbol{C}$ 是可逆矩阵,见(4.1))

$$(\beta_1\ \beta_2\ \beta_3)=(\alpha_1\ \alpha_2\ \alpha_3)\ \boldsymbol{C}$$

如果向量 α 在这两组基下的坐标分别是 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}$ 和 $\boldsymbol{y}=(y_1,y_2,y_3)^{\mathrm{T}}$,那么把坐标变换

$$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$$

(参看 4.1 中(4.3))代入到变量 $\boldsymbol{x}$ 的二次型中,有

$$\begin{aligned} f(x_1\ x_2\ x_3) &= \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \\ &= (\boldsymbol{C}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{C}\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}, \end{aligned}$$

其中 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$.

由于 $\boldsymbol{A}$ 是对称矩阵,用转置的性质得到

$$\boldsymbol{B}^{\mathrm{T}}=(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})^{\mathrm{T}}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B},$$

$\boldsymbol{B}$ 是一个对称矩阵. 因而 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 是二次型 f 的矩阵表示.

这说明,变量 x_1,x_2,x_3 的三元二次齐次式 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 后成为变量 y_1,y_2,y_3 的三元二次齐次式 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$,二次型的对应矩阵亦由基 $\alpha_1,\alpha_2,\alpha_3$ 下的 $\boldsymbol{A}$ 转化为基 β_1,β_2,β_3 下的矩阵 $\boldsymbol{B}$,且 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$,一般地

定理 6.1 变量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 的 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 后,成为变量 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$ 的 n 元二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$,其中 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$. ■

二次型经坐标变换后仍是二次型,现在把二次型矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 之间的这种关系抽象出来,引出矩阵间合同的概念.

定义 6.2 两个实对称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$,如存在可逆矩阵 $\boldsymbol{C}$,使得

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B}, \tag{6.2}$$

就称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 合同,记作 $\boldsymbol{A}\simeq\boldsymbol{B}$.

容易验证,矩阵的合同与矩阵的相似一样,合同也具有反身性,对称性和传递性,合同是实对称矩阵之间的一种等价关系 .

例 1 证明 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同,其中

$$\boldsymbol{A}=\begin{pmatrix} a_1 & \\ & a_2 \end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix} a_2 & \\ & a_1 \end{pmatrix}.$$

证 取可逆矩阵 $\boldsymbol{C}=\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$,则

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}\begin{pmatrix} a_1 & \\ & a_2 \end{pmatrix}\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$$

$$=\begin{pmatrix} & a_2 \\ a_1 & \end{pmatrix}\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}=\begin{pmatrix} a_2 & \\ & a_1 \end{pmatrix}=\boldsymbol{B},$$

所以,$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同 . ■

*例 2 证明在实数范围内 $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ 与 $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ 不合同 .

证 (用反证法)如果这两个矩阵合同,则存在可逆矩阵 $\boldsymbol{C}$,使得

$$\boldsymbol{C}^{\mathrm{T}}\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}\boldsymbol{C}=\begin{pmatrix} 1 & \\ & 1 \end{pmatrix},$$

两边取行列式,并利用行列式乘法公式(定理 2.1),有

$$|\boldsymbol{C}^{\mathrm{T}}|\begin{vmatrix} 1 & \\ & -1 \end{vmatrix}|\boldsymbol{C}|=\begin{vmatrix} 1 & \\ & 1 \end{vmatrix},$$

由于 $|\boldsymbol{C}^{\mathrm{T}}|=|\boldsymbol{C}|$,得 $|\boldsymbol{C}|^2=-1$.
因为矩阵 $\boldsymbol{C}$ 中元素是实数,行列式 $|\boldsymbol{C}|$ 是实数,那么 $|\boldsymbol{C}|^2>0$,这与 $|\boldsymbol{C}|^2=-1$ 相矛盾 . 所以,$\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ 与 $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ 不合同 . ■

例 3 证明 $\begin{pmatrix} 2 & \\ & 3 \end{pmatrix}$ 与 $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ 合同但不相似 .

证　取可逆矩阵 $\boldsymbol{C}=\begin{pmatrix}\frac{1}{2} & \\ & \frac{1}{3}\end{pmatrix}$,由于

$$\boldsymbol{C}^{\mathrm{T}}\begin{pmatrix}2 & \\ & 3\end{pmatrix}\boldsymbol{C}=\begin{pmatrix}\frac{1}{2} & \\ & \frac{1}{3}\end{pmatrix}\begin{pmatrix}2 & \\ & 3\end{pmatrix}\begin{pmatrix}\frac{1}{2} & \\ & \frac{1}{3}\end{pmatrix}=\begin{pmatrix}1 & \\ & 1\end{pmatrix},$$

所以 $\begin{pmatrix}2 & \\ & 3\end{pmatrix}$ D $\begin{pmatrix}1 & \\ & 1\end{pmatrix}$.

但矩阵 $\begin{pmatrix}2 & \\ & 3\end{pmatrix}$ 的特征值是 2 和 3,矩阵 $\begin{pmatrix}1 & \\ & 1\end{pmatrix}$ 的特征值是 1 (二重根),两个矩阵的特征值不同,根据第 5 章定理 5.4,矩阵相似的必要条件是有相同的特征值,所以这两个矩阵不可能相似.■

例 4　$\boldsymbol{A}$ 是三阶实对称矩阵,如果对任何三维列向量 $\boldsymbol{X}$ 都有 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=0$,证明:$\boldsymbol{A}=0$.

证　令 $\boldsymbol{x}=(1,0,0)^{\mathrm{T}}$,我们有

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(1\ 0\ 0)\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}=a_{11},$$

又因," $\boldsymbol{x}$ 有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=0$,所以 $a_{11}=0$.

类似取 $\boldsymbol{x}$ 为$(0,1,0)^{\mathrm{T}}$ 和$(0,0,1)^{\mathrm{T}}$ 可得到 $a_{22}=a_{33}=0$;

再取 $\boldsymbol{x}=(1,1,0)^{\mathrm{T}}$. 由于

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(1\ 1\ 0)\begin{pmatrix}0 & a_{12} & a_{13}\\ a_{21} & 0 & a_{23}\\ a_{31} & a_{32} & 0\end{pmatrix}\begin{pmatrix}1\\1\\0\end{pmatrix}=a_{21}+a_{12}=0.$$

又因 $a_{12}=a_{21}$,所以 $a_{12}=a_{21}=0$.

同样地,$\boldsymbol{x}$ 为$(1,0,1)^{\mathrm{T}}$ 及$(0,1,1)^{\mathrm{T}}$ 可得 $a_{13}=a_{31}=0$, $a_{23}=a_{32}=0$.

因为 $a_{ij}\equiv 0\ (i,j=1,2,3)$，故 $\boldsymbol{A}=0$. ■

如果二次型中只含有变量的平方项,所有混合项 $x_i x_j (i\neq j)$ 的系数全是零,即

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2. \tag{6.3}$$

这样的二次型称为标准形.

一个二次型如是标准形,它的二次型矩阵就是对角形. 对(6.3)可用对角矩阵表示成

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

定理 6.1 说明,经坐标变换后二次型的两个矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是合同的. 那么,现在感兴趣的问题是:对任给的二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是否存在好的坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$,使变换后的二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 为标准形呢? 换句话说,任给一个实对称矩阵 $\boldsymbol{A}$,它能否合同于一个对角矩阵呢?

答案是肯定的. 并且可以用多种方法把二次型化成标准形.

6.2 用配方法化二次型为标准形

本节的化标准形方法是用初等代数中的配平方法. 下面通过例题介绍这种方法:

例 1 求坐标变换化二次型

$$f(x_1,x_2,x_3)=2x_1^2+3x_2^2+8x_3^2+4x_1x_2-8x_2x_3-4x_3x_1,$$

为标准形.

解 先把所有含 x_1 的项配成一个完全平方,即有

$$f=2(x_1+x_2-x_3)^2+x_2^2+6x_3^2-4x_2x_3,$$

再把所有含 x_2 的项配成一个完全平方. 即

$$f = 2(x_1 + x_2 - x_3)^2 + (x_2 - 2x_3)^2 + 2x_3^2,$$

引入新变量,令

$$\begin{cases} y_1 = x_1 + x_2 - x_3, \\ y_2 = x_2 - 2x_3, \\ y_3 = x_3, \end{cases}$$

则有 f 的标准形

$$f = 2y_1^2 + y_2^2 + 2y_3^2$$

相应的坐标变换公式,很容易由 $\boldsymbol{y}$ 与 $\boldsymbol{x}$ 的关系中求出

$$\begin{cases} x_1 = y_1 - y_2 - y_3, \\ x_2 = y_2 + 2y_3, \\ x_3 = y_3, \end{cases}$$

用矩阵可表示为 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$, 其中

$$\boldsymbol{C} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix},$$

$\boldsymbol{C}$ 是可逆的上三角矩阵 .

注:(1) 从矩阵的角度来说, 例 1 表示实对称矩阵

$\boldsymbol{A} = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 3 & -4 \\ -2 & -4 & 8 \end{pmatrix}$ 合同于对角矩阵 $\boldsymbol{B} = \begin{pmatrix} 2 & & \\ & 1 & \\ & & 2 \end{pmatrix}$.

(2) 配方时,我们强调一次要把一种字母配完,这必能保证矩阵 C 是上三角矩阵且是可逆的,确保 $x = Cy$ 是一坐标变换 .

二次型不含平方项时,可先作一个坐标变换构造出平方项,然后再用例 1 的方法 .

例 2 求坐标变换 $x = Cy$, 化二次型

$$f(x_1, x_2, x_3) = 2x_1x_2 - 4x_2x_3$$

为标准形 .

解　我们先令

$$\begin{aligned} x_1 &= y_1 + y_2, \\ x_2 &= y_1 - y_2, \\ x_3 &= y_3, \end{aligned}$$

代入到二次型中,得

$$\begin{aligned} f(x_1, x_2, x_3) &= 2(y_1 + y_2)(y_1 - y_2) - 4(y_1 - y_2)y_3 \\ &= 2y_1^2 - 2y_2^2 - 4y_1y_3 + 4y_2y_3 \\ &= 2(y_1 - y_3)^2 - 2y_2^2 + 4y_2y_3 - 2y_3^2 \\ &= 2(y_1 - y_3)^2 - 2(y_2 - y_3)^2, \end{aligned}$$

再令

$$\begin{aligned} z_1 &= y_1 - y_3, \\ z_2 &= y_2 - y_3, \\ z_3 &= y_3, \end{aligned}$$

就得到二次型的标准形: $f = 2z_1^2 - 2z_2^2$.

而所用的坐标变换是:

$$\begin{aligned} x_1 &= y_1 + y_2 &&= z_1 + z_2 + 2z_3, \\ x_2 &= y_1 - y_2 &&= z_1 - z_2, \\ x_3 &= y_3 &&= z_3, \end{aligned}$$

用矩阵可表示为,由

$$\boldsymbol{x} = \boldsymbol{C}_1\boldsymbol{y},\quad \boldsymbol{y} = \boldsymbol{C}_2\boldsymbol{z}$$

得坐标变换: $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{z}$, $\boldsymbol{C} = \boldsymbol{C}_1\boldsymbol{C}_2$. **其中**

$$\boldsymbol{C}_1 = \begin{pmatrix} 1 & 1 & \\ 1 & -1 & \\ & & 1 \end{pmatrix},\quad \boldsymbol{C}_2 = \begin{pmatrix} 1 & & 1 \\ & 1 & 1 \\ & & 1 \end{pmatrix},\quad \boldsymbol{C} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.\ \blacksquare$$

通过上述例题可以看出,对任何二次型都可用这些方法化其为平方项的代数和. 这就是:

定理 6.2 任意的 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 都可以通过坐标变换化成标准形

$$d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2,\tag{6.3}$$

其中 $d_i(i=1,2,\cdots,n)$是实数.(证略) ■

定理 6.2 的等价说法是:

定理 6.3 任一 n 阶实对称矩阵 $\boldsymbol{A}$,总可以合同于一个对角矩阵

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix}d_1&&&\\&d_2&&\\&&\ddots&\\&&&d_n\end{pmatrix}.\tag{6.4}$$ ■

一个二次型可以用多种方法化其为标准形,在标准形中正平方项的个数及负平方项的个数与所选用的坐标变换是无关的,这就是著名的惯性定理,其证明从略.

定理 6.4 (惯性定理)对于一个二次型,不论选取怎样的坐标变换使它化为仅含平方项的标准形,其中正平方项的个数 p,负平方项的个数 q 都是由所给二次型唯一确定的. ■

通常,称 p 为二次型的正惯性指数,q 是负惯性指数,$r=p+q$ 为二次型的秩.

例如,例 1 的二次型经坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 化成 $2y_1^2+y_2^2+2y_3^2$,所以二次型 $2x_1^2+3x_2^2+8x_3^2+4x_1x_2-8x_2x_3-4x_3x_1$ 的正惯性指数 $p=3$,负惯性指数 $q=0$,二次型的秩 $p+q=3$.而例 2 中二次型 $2x_1x_2-4x_2x_3$ 的正、负惯性指数 $p=q=1$,二次型的秩 $p+q=2$.

在二次型的标准形 $d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2$ 中,如果系数 $d_i(i=1,2,\cdots,n)$的值为±1 或 0,这样的标准形称为二次型的规范形.

一个二次型总可用坐标变换化其为规范形.例如,在例 1 中

再令

$$\begin{aligned} z_1 &= 2y_1, \\ z_2 &= y_2, \\ z_3 &= 2y_3, \end{aligned}$$

就有 $f = z_1^2 + z_2^2 + z_3^2$,它就是 f 的规范形 . 又如,在例 2 中再令

$$\begin{aligned} u_1 &= 2z_1, \\ u_2 &= 2z_2, \\ u_3 &= z_3, \end{aligned}$$

就是 $f = u_1^2 - u_2^2$,它是 $2x_1x_2 - 4x_2x_3$ 的规范形 .

设 $\boldsymbol{A}$ 是 n 阶实对称矩阵,秩 $\mathrm{r}(\boldsymbol{A}) = r$,若可逆矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{A}$ 合同于对角矩阵 Λ,即 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC} = \Lambda$. 那么根据秩的性质我们有

$$\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}) = \mathrm{r}(\Lambda),$$

而对角矩阵 Λ 的秩 $\mathrm{r}(\Lambda)$就是对角线上非 0 元素的和 . 所以二次型的秩 $p + q = \mathrm{r}(\Lambda)$就是矩阵 $\boldsymbol{A}$ 的秩 $\mathrm{r}(\boldsymbol{A})$. 即 $p + q = \mathrm{r}$.

根据惯性定理,如果 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$ 的正负惯性指数分别为 p 和 q,如同前面例题的讨论,我们可先选坐标变换化其为标准形

$$\begin{aligned} & d_1y_1^2 + d_2y_2^2 + \cdots + d_py_p^2 - d_{p+1}y_{p+1}^2 \\ & - d_{p+2}y_{p+2}^2 - \cdots - d_{p+q}y_{p+q}^2 \end{aligned}$$

其中系数 $d_i > 0\ (i = 1, 2, \cdots, p + q)$.

再选坐标变换化其为规范形

$$z_1^2 + z_2^2 + \cdots + z_p^2 - z_{p+1}^2 - z_{p+2}^2 - \cdots - z_{p+q}^2 .$$

用矩阵的话来说,上述二次型矩阵 $\boldsymbol{A}$ 合同于

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} . \qquad (6.5)$$

其中 1 有 p 个, -1 有 q 个, 0 有 $n-(p+q)$ 个 . 这样的对角矩阵称为 $\boldsymbol{A}$ 的合同规范形 .

如果二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经过坐标变换化为二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$, 那么称这两个二次型是合同的 . 显然, 这充要条件是它们的正惯性指数相同, 负惯性指数也相同, 也即矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是合同的 .

6.3　用正交变换化二次型为标准形

二次型的矩阵是实对称矩阵, 而实对称矩阵是可以经过正交变换化为对角形的(5.4), 如 $\boldsymbol{T}$ 是正交矩阵, 则 $\boldsymbol{T}^{-1}=\boldsymbol{T}^{\mathrm{T}}$, 那么

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{T}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{T}.$$

因此, 实对称矩阵化为对角形(相似)与实对称矩阵合同于一个对角矩阵就统一起来了 . 所以二次型除用配方法外亦可用正交变换化为标准形 . 下面通过例题介绍这种方法 .

例 1　求正交变换 $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$, 化二次型

$$f(x_1, x_2) = 2x_1^2 + 2x_2^2 - 2x_1x_2$$

为标准形 .

解　二次型的矩阵为 $\boldsymbol{A}=\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$,

由此求出 $\boldsymbol{A}$ 的特征多项式

$$|\boldsymbol{A}-\lambda\boldsymbol{I}| = \begin{vmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} = \lambda^2-4\lambda+3,$$

得到 A 的特征值是 $\lambda_1 = 1, \lambda_2 = 3$.

当 $\lambda = 1$ 时,由特征方程 $(A - I)x = 0$,即

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

得到基础解系(即特征向量): $X_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

当 $\lambda = 3$ 时,由特征方程 $(A - 3I)x = 0$,即

$$\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

得基础解系(即特征向量) $X_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

由于 $\lambda_1 \neq \lambda_2$,特征向量 X_1、X_2 是正交的(定理 5.8),故只须把其单位化,有

$$\gamma_1 = \frac{X_1}{\| X_1 \|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \gamma_2 = \frac{X_2}{\| X_2 \|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

构造正交矩阵 T

$$T = (\gamma_1, \gamma_2) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

令 $x = Ty$,则有

$$x^{\mathrm{T}} A x = (Ty)^{\mathrm{T}} A (Ty) = y^{\mathrm{T}} T^{\mathrm{T}} A T y$$

$$= y^{\mathrm{T}} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} y$$

$$= y^{\mathrm{T}} \begin{pmatrix} 1 & \\ & 3 \end{pmatrix} y = y_1^2 + 3y_2^2.$$

所以,经正交变换 $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$,二次型 $2x_1^2+2x_2^2-2x_1x_2$ 的标准形是 $y_1^2+3y_2^2$. ■

例 2 求正交变换 $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$,化二次型

$$3x_1^2+6x_2^2+3x_3^2-4x_1x_2-8x_1x_3-4x_2x_3$$

为标准形.

解 二次型的矩阵是

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

矩阵 $\boldsymbol{A}$ 的特征多项式

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}3-\lambda&-2&-4\\-2&6-\lambda&-2\\-4&-2&3-\lambda\end{vmatrix}=\begin{vmatrix}7-\lambda&0&\lambda-7\\-2&6-\lambda&-2\\-4&-2&3-\lambda\end{vmatrix}$$

$$=(7-\lambda)\begin{vmatrix}1&0&-1\\-2&6-\lambda&-2\\-4&-2&3-\lambda\end{vmatrix}$$

$$=(7-\lambda)\begin{vmatrix}1&0&0\\-2&6-\lambda&-4\\-4&-2&-1-\lambda\end{vmatrix}$$

$$=(7-\lambda)(\lambda-7)(\lambda+2).$$

得到 $\boldsymbol{A}$ 的特征值:$\lambda_1=\lambda_2=7,\ \lambda_3=-2$.

当 $\lambda=7$ 时,由 $(\boldsymbol{A}-7\boldsymbol{I})\boldsymbol{x}=0$,即

$$\begin{pmatrix}-4&-2&-4\\-2&-1&-2\\-4&-2&-4\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

系数矩阵的秩为 1,得基础解系:

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

当 $\lambda = -2$ 时,由$(\boldsymbol{A} + 2\boldsymbol{I})\boldsymbol{x} = 0$,即

$$\begin{pmatrix} 5 & -2 & -4 \\ -2 & 8 & -2 \\ -4 & -2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

对系数矩阵高斯消元,有

$$\begin{pmatrix} 1 & -4 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

得到基础解系(即 $\lambda = -2$ 所对应的特征向量)

$$\boldsymbol{X}_3 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

显然,$\boldsymbol{X}_3$ 与 $\boldsymbol{X}_1$,$\boldsymbol{X}_2$ 都正交(它们是不同特征值所对应的特征向量),但 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 并不正交,对 $\boldsymbol{X}_1$,$\boldsymbol{X}_2$ 运用 Schmidt 标准正交化,有

$$\beta_1 = \boldsymbol{X}_1 = (-1, 2, 0)^{\mathrm{T}},$$

$$\beta_2 = -\frac{(\beta_1, \boldsymbol{X}_2)}{(\beta_1, \beta_1)}\beta_1 + \boldsymbol{X}_2 = \left(-\frac{4}{5}, -\frac{2}{5}, 1\right)^{\mathrm{T}},$$

再单位化,得

$$\gamma_1 = \frac{\beta_1}{\|\beta_1\|} = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \quad \gamma_2 = \frac{\beta_2}{\|\beta_2\|} = \frac{1}{3\sqrt{5}}\begin{pmatrix} -4 \\ -2 \\ 5 \end{pmatrix},$$

$$\gamma_3 = \frac{\boldsymbol{X}_3}{\|\boldsymbol{X}_3\|} = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

构造正交矩阵

$$
\boldsymbol{T}=(\gamma_1\ \ \gamma_2\ \ \gamma_3)=\begin{pmatrix}\frac{-1}{\sqrt{5}} & \frac{-4}{3\sqrt{5}} & \frac{2}{3}\\ \frac{2}{\sqrt{5}} & \frac{-2}{3\sqrt{5}} & \frac{1}{3}\\ 0 & \frac{5}{3\sqrt{5}} & \frac{2}{3}\end{pmatrix},
$$

容易验证

$$
\boldsymbol{T}^{\mathrm{T}}\boldsymbol{AT}=\boldsymbol{T}^{-1}\boldsymbol{AT}=\begin{pmatrix}7 & & \\ & 7 & \\ & & -2\end{pmatrix}.
$$

即经过 $\boldsymbol{x}=\boldsymbol{Ty}$，二次型化为标准形

$$
\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{T}^{\mathrm{T}}\boldsymbol{AT})\boldsymbol{y}=7y_1^2+7y_2^2-2y_3^2.
$$ ■

通过以上例题我们看到，二次型经正交变换可以化成标准形，并且在标准形中平方项的系数就是该二次型矩阵的特征值，这不是偶然的巧合．关键在于二次型的矩阵是实对称矩阵，实对称矩阵的特征值及特征向量的性质保证我们可以构造正交矩阵 T，由于 $T^{-1}=T^{\mathrm{T}}$，那么 $\boldsymbol{T}^{-1}\boldsymbol{AT}=\boldsymbol{T}^{\mathrm{T}}\boldsymbol{AT}$，因此 $\boldsymbol{A}$ 相似于对角阵与 $\boldsymbol{A}$ 合同于对角阵是一致的．这样，我们就可用二次型的语言来叙述上一章的定理 5.9.

定理 6.5 对任一个 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$，其中 $\boldsymbol{A}$ 是 n 阶实对称矩阵，必存在正交变换 $\boldsymbol{x}=\boldsymbol{Ty}$（$\boldsymbol{T}$ 是正交矩阵），使得 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$ 化成标准形

$$
\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2. \tag{6.6}
$$

这里 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值． ■

二次型通过正交变换化为标准形的问题来源于平面的二次曲线及空间的二次曲面方程的化简．也就是通过对直角坐标系的平移及旋转把二次方程化为最简单的形式（不含混合项）．下面以二次曲线方程的化简为例介绍这一方法．

*例 3　求直角坐标变换化简二次方程

$$5x^2 - 6xy + 5y^2 - 6\ 2x + 2\ 2y - 4 = 0$$

解　先把二次项化为标准形．由对应矩阵

$$\boldsymbol{A} = \begin{matrix} 5 & -3 \\ -3 & 5 \end{matrix}$$

求出 $\boldsymbol{A}$ 的特征值：$\lambda_1 = 2, \lambda_2 = 8$

由$(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = 0$ 求出相应的特征向量，再单位化(现特征值不同，特征向量必正交；若特征值是等根就用 Schmidt 正交化)就有

$$\gamma_1 = \frac{1}{2}\begin{matrix} 1 \\ 1 \end{matrix}, \quad \gamma_2 = \frac{1}{2}\begin{matrix} -1 \\ 1 \end{matrix}.$$

构造正交矩阵 $\boldsymbol{T} = (\gamma_1, \gamma_2)$，那么经正交变换(即直角坐标系的旋转)

$$\begin{matrix} x \\ y \end{matrix} = \begin{matrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{matrix} \begin{matrix} x' \\ y' \end{matrix},$$

也就是把

$$x = \frac{1}{2}x' - \frac{1}{2}y'$$

$$y = \frac{1}{2}x' + \frac{1}{2}y'$$

代入到原方程，方程化简为

$$2x'^2 + 8y'^2 - 6\ 2 \cdot \frac{1}{2}(x' - y') + 2\ 2$$

$$\cdot \frac{1}{2}(x' + y') - 4 = 0$$

整理后即有

$$x'^2 + 4y'^2 - 2x' + 4y' - 2 = 0.$$

再用配平方法,得

$$(x'-1)^2+4\left(y'+\frac{1}{2}\right)^2=4.$$

因此,经平移坐标轴

$$x''=x'-1,$$

$$y''=y'+\frac{1}{2}.$$

得到二次曲线的标准方程

$$\frac{1}{4}x''^2+y''^2=1.$$

其图形是椭圆,而所用的直角坐标变换是

$$x''=\frac{1}{2}x+\frac{1}{2}y-1,$$

$$y''=-\frac{1}{2}x+\frac{1}{2}y+\frac{1}{2}.$$ ■

6.4 正定二次型

观察 6.2 中的例 1,由于二次型

$$\begin{aligned}f(x_1,x_2,x_3)&=2x_1^2+3x_2^2+8x_3^2+4x_1x_2-8x_2x_3-4x_3x_1\\&=2(x_1+x_2-x_3)^2+(x_2-2x_3)^2+2x_3^2\end{aligned}$$

那么对任何 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}\neq 0$,我们恒有

$$f(x_1,x_2,x_3)>0$$

且仅有 $f(0,0,0)=0$.

说明这个三元二次齐次式恒正,仅在 $\boldsymbol{x}=0$ 处函数有极小值 0.而该节例 2 所给出的二次型 $f(x_1,x_2,x_3)=2x_1x_2-4x_2x_3$ 就不是恒正的.例如,$f(1,0,0)=0$, $f(0,1,1)=-4$.

如何判断一个二次型 f 是否恒正,仅在 $\boldsymbol{x}=0$ 才有 $f=0$?这

就是本节所要学习的正定二次型、正定矩阵 .

定义 6.3 对二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,如对任何 $\boldsymbol{x}\neq 0$,恒有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$,则称二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是正定二次型 . 并称实对称矩阵 $\boldsymbol{A}$ 是正定矩阵 .

例如,$f(x_1, x_2) = x_1^2 + 3x_2^2$,对任何 $\boldsymbol{x} = \begin{pmatrix} t \\ u \end{pmatrix} \neq 0$ 总有 $f(t, u) = t^2 + 3u^2 > 0$,因此,这二次型是正定二次型, $\boldsymbol{A} = \begin{pmatrix} 1 & \\ & 3 \end{pmatrix}$ 是正定矩阵 .

二次型 $f(x_1, x_2, x_3) = x_1^2 + 3x_2^2 - 5x_3^2 + 4x_1x_2 - 6x_2x_3$ 中,平方项 x_3^2 的系数 $a_{33} = -5 < 0$,如果选 $\boldsymbol{x} = (0, 0, 1)^{\mathrm{T}} \neq 0$,代入有 $f(0, 0, 1) = -5 < 0$,所以 f 不是正定的 .

二次型 $f(x_1, x_2, x_3) = 4x_2^2 + x_3^2 + 2x_1x_2$ 中,平方项 x_1^2 的系数 $a_{11} = 0$,如果取 $\boldsymbol{x} = (1, 0, 0)^{\mathrm{T}} \neq 0$,代入得 $\mathrm{f}(1, 0, 0) = 0$,这个二次型也不是正定的 .

从上面的讨论分析,容易得到:

若 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定,则其平方项的系数 $a_{ii}(i = 1, 2, \cdots, n)$ 必须全大于零 . 从正定矩阵的角度来说就是主对角线元素恒为正 .

这是一个必要条件,对二次型

$$\mathrm{f}(x_1, x_2, x_3) = x_1^2 + x_2^2 + 3x_3^2 - 10x_1x_2,$$

虽然 a_{ii} 全大于零,但 $\mathrm{f}(1, 1, 1) = -5 < 0$, f 并不是正定的二次型 .

下面来研究二次型正定的充分必要条件

引理 6.6 设 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$ 是坐标变换,则 $\boldsymbol{x}\neq 0$ 的充要条件是 $\boldsymbol{y}\neq 0$.

证 必要性(用反证法)

如 $\boldsymbol{y} = 0$,则 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y} = \boldsymbol{C}0 = 0$,与已知矛盾 .

充分性(用反证法) .

如 $\boldsymbol{x}=\boldsymbol{0}$,由 $\boldsymbol{C}\boldsymbol{y}=0$ 而 $\boldsymbol{y}\neq 0$,知齐次方程组

$$\boldsymbol{C}\begin{matrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{matrix} = \begin{matrix} 0 \\ 0 \\ \cdots \\ 0 \end{matrix}$$

有非零解,那么方程组的系数行列式 $|\boldsymbol{C}|=0$(定理 3.2),这与 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 是坐标变换,$\boldsymbol{C}$ 为可逆矩阵(参看(4.2.1))相矛盾. ■

引理 6.6 表明,二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 化成二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 时,二次型的正定性是不会因此而改变的.这是因为:

对任意的 $\boldsymbol{y}_0\neq 0$,令 $\boldsymbol{x}_0=\boldsymbol{C}\boldsymbol{y}_0$,则 $\boldsymbol{x}_0\neq 0$,

那么,由 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定,就可得到

$$\boldsymbol{y}_0^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}_0=\boldsymbol{y}_0^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y}_0=\boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_0>0,$$

即二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 正定.

反之,若二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 正定,那么经坐标变换 $\boldsymbol{y}=\boldsymbol{C}^{-1}\boldsymbol{x}$,可知 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是正定二次型.

根据二次型正定的这一性质,要判断 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的正定性,就可以用它的标准形的正定性来决定,而标准形的正定性是容易看出的.

定理 6.7 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定的充分必要条件有:

(1) $\boldsymbol{A}$ 的正惯性指数是 n;

(2) $\boldsymbol{A}$ 与 $\boldsymbol{I}$ 合同,即存在可逆矩阵 $\boldsymbol{C}$,使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{I}$;

(3) $\boldsymbol{A}$ 的所有特征值 $\lambda_i(i=1,2,\cdots,n)$ 均为正数.

证 (1) 设坐标变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 化二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为标准形

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y},$$

由引理 6.6 知:$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定 $\quad$ $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 正定

而 $\quad\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}$ 正定 $\quad d_i>0(i=1,2,\cdots,n)$

正惯性指数 $p=n$.

(2) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定 $\quad$ 正惯性指数 $p=n$

$\boldsymbol{A}$ 的合同规范形是 $\boldsymbol{I}$(参看(6.5))

$\boldsymbol{A}$ 与 $\boldsymbol{I}$ 合同.

(3) 对 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 有正交变换 $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$ 化其为标准形

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2 \tag{6.6}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的特征值,由(1)知

$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 全大于零. ■

我们还可以用行列式来判断正定性,即

定理 6.8 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定的充分必要条件是 $\boldsymbol{A}$ 的各阶顺序主子式 $\Delta_k(k=1,2,\cdots,n)$ 均大于零,即

$$\Delta_1=a_{11}>0,\quad \Delta_2=\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}>0,\quad \Delta_3=\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}>0,$$

$$\cdots,\quad \Delta_n=|\boldsymbol{A}|=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}>0. \qquad ■$$

推论 实对称矩阵 $\boldsymbol{A}$ 正定的必要条件是 $|\boldsymbol{A}|>0$. ■

例 1 说明下列矩阵都不是正定的.

$$\boldsymbol{A}=\begin{pmatrix}2&1&-2\\1&0&4\\-2&4&3\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}1&2&-3\\2&5&-7\\-3&-7&10\end{pmatrix}$$

解 对矩阵 $\boldsymbol{A}$,由 $a_{22}=0$,不满足 $\boldsymbol{A}$ 正定的必要条件 a_{ii} 全大于 0. 所以 $\boldsymbol{A}$ 不是正定的.

对矩阵 $\boldsymbol{B}$,由于 $|\boldsymbol{B}|=0$(把一、二两行全加至第三行!),与正定的必要条件行列式大于 0 相矛盾(推论),所以 $\boldsymbol{B}$ 不是正定的.

■

例 2 判断二次型

$$f(x_1,x_2,x_3)=6x_1^2+5x_2^2+x_3^2-4x_1x_2+4x_1x_3$$

是否是正定二次型?

解 二次型的对应矩阵是

$$\boldsymbol{A}=\begin{pmatrix}6&-2&2\\-2&5&0\\2&0&1\end{pmatrix}$$

其顺序主子式

$$\Delta_1=a_{11}=6>0,$$

$$\Delta_2=\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}=\begin{vmatrix}6&-2\\-2&5\end{vmatrix}=26>0,$$

$$\Delta_3=|\boldsymbol{A}|=\begin{vmatrix}6&-2&2\\-2&5&0\\2&0&1\end{vmatrix}=\begin{vmatrix}2&-2&0\\-2&5&0\\2&0&1\end{vmatrix}$$

$$=\begin{vmatrix}2&-2\\-2&5\end{vmatrix}=6>0$$

所以二次型是正定的. ■

例 3 二次型 $x_1^2+2x_2^2+4x_3^2-2x_1x_2+2tx_2x_3$ **正定,求参数** t **的取值范围.**

解 二次型的对应矩阵是

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

其顺序主子式

$$\Delta_1=1,\ \Delta_2=\begin{vmatrix}1&-1\\-1&2\end{vmatrix}=1,\ \Delta_3=|\boldsymbol{A}|=4-t^2.$$

由于二次型正定,故 $4-t^2>0$ (**定理** 6.8),**解得当** $t\in(-2,2)$**时,二次型正定.** ■

例 4 已知 $\boldsymbol{A}$ 是正定矩阵,证明 $\boldsymbol{A}^{-1}$是正定矩阵.

证 由于 $\boldsymbol{A}$ 正定,一方面知 $|\boldsymbol{A}|>0$,得 $\boldsymbol{A}^{-1}$存在;另一方面知 $\boldsymbol{A}$ 是对称的,即 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$. 那么

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

得 $\boldsymbol{A}^{-1}$是对称矩阵.

根据 $\boldsymbol{AX}=\lambda\boldsymbol{X}$ 及 $\boldsymbol{A}$ 正定时 $\lambda>0$(定理 6.7),有 $\boldsymbol{A}^{-1}\boldsymbol{X}=\dfrac{1}{\lambda}X$,所以 $\boldsymbol{A}^{-1}$的特征值 $\dfrac{1}{\lambda}$ 全大于零,故 $\boldsymbol{A}^{-1}$正定. ■

注 证明 $\boldsymbol{A}^{-1}$正定的方法很多,例如

由 $\boldsymbol{A}$ 正定知存在可逆矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\boldsymbol{I}$,那么两边取逆,有

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

记 $(\boldsymbol{C}^{-1})^{\mathrm{T}}=\boldsymbol{D}$,则 $\boldsymbol{C}^{-1}=\boldsymbol{D}^{\mathrm{T}}$,故

$$\boldsymbol{D}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{D}=\boldsymbol{I}.$$

所以 $\boldsymbol{A}^{-1}$正定.

例 5 $\boldsymbol{A}$ 是 3 阶实对称矩阵,证明存在实数 t 使 $\boldsymbol{A}+t\boldsymbol{I}$ 是正定矩阵.

证 设 $\boldsymbol{A}$ 的特征值是 $\lambda_1,\lambda_2,\lambda_3$. 那么 $\boldsymbol{A}+t\boldsymbol{I}$ 的特征值是 $\lambda_1+t,\lambda_2+t,\lambda_3+t$. 总可选取实数 t 使 $\lambda_1+t>0,\lambda_2+t>0,\lambda_3+t>0$. 由定理 6.7, $\boldsymbol{A}+t\boldsymbol{I}$ 的特征值全大于零知其正定. ■

*例 6 $\boldsymbol{A}$ 是 3 阶正定矩阵,证明 $|\boldsymbol{A}+\boldsymbol{I}|>1$.

证 因为 $\boldsymbol{A}$ 是实对称的,故存在正交矩阵 $\boldsymbol{T}$ 使

$$\boldsymbol{T}^{-1}\boldsymbol{AT}=\boldsymbol{T}^{\mathrm{T}}\boldsymbol{AT}=\begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{pmatrix},$$

其中 $\lambda_1,\lambda_2,\lambda_3$ 是 $\boldsymbol{A}$ 的特征值. 由 $\boldsymbol{A}$ 正定知 $\lambda_i(i=1,2,3)$全大于零,这样

$$\boldsymbol{A}+\boldsymbol{I}=\boldsymbol{T}\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix}\boldsymbol{T}^{-1}+\boldsymbol{I}=\boldsymbol{T}\begin{pmatrix}\lambda_1+1&&\\&\lambda_2+1&\\&&\lambda_3+1\end{pmatrix}\boldsymbol{T}^{-1},$$

所以

$$|\boldsymbol{A}+\boldsymbol{I}|=|\boldsymbol{T}|\begin{vmatrix}\lambda_1+1&&\\&\lambda_2+1&\\&&\lambda_3+1\end{vmatrix}|\boldsymbol{T}^{-1}|=\begin{vmatrix}\lambda_1+1&&\\&\lambda_2+1&\\&&\lambda_3+1\end{vmatrix}$$

$$=(\lambda_1+1)(\lambda_2+1)(\lambda_3+1)>1.$$ ■

***例 7** 已知 $\boldsymbol{A}$ 是三阶正定矩阵，证明存在三阶正定矩阵 $\boldsymbol{B}$，使 $\boldsymbol{A}=\boldsymbol{B}^2$.

证 设 $\boldsymbol{A}$ 的特征值是 $\lambda_1,\lambda_2,\lambda_3$，由于 $\boldsymbol{A}$ 是正定的，这些特征值均大于零．又因 $\boldsymbol{A}$ 是实对称矩阵，故有正交矩阵 $\boldsymbol{T}$ 使

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix},$$

那么

$$\boldsymbol{A}=\boldsymbol{T}\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix}\boldsymbol{T}^{-1}=\boldsymbol{T}\begin{pmatrix}\sqrt{\lambda_1}&&\\&\sqrt{\lambda_2}&\\&&\sqrt{\lambda_3}\end{pmatrix}\boldsymbol{T}^{-1}\boldsymbol{T}\begin{pmatrix}\sqrt{\lambda_1}&&\\&\sqrt{\lambda_2}&\\&&\sqrt{\lambda_3}\end{pmatrix}\boldsymbol{T}^{-1},$$

令

$$\boldsymbol{B}=\boldsymbol{T}\begin{pmatrix}\sqrt{\lambda_1}&&\\&\sqrt{\lambda_2}&\\&&\sqrt{\lambda_3}\end{pmatrix}\boldsymbol{T}^{-1},$$

则

$$\boldsymbol{B}\sim\begin{matrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{matrix}.$$

$\boldsymbol{B}$ 的特征值 λ_1, λ_2, λ_2 全大于零,故 $\boldsymbol{B}$ 是正定矩阵,且 $\boldsymbol{A}=\boldsymbol{B}^2$. ■

与正定性相平行,还有半正定,负定,半负定等概念.

*定义 6.4　对 n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,如果对任意的非零向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\neq\boldsymbol{0}$,恒有

(1) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\geq 0$,且存在非零的 $\boldsymbol{X}_0$ 使得 $\boldsymbol{X}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}_0=0$,就称 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是半正定二次型,$\boldsymbol{A}$ 叫半正定矩阵;

(2) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}<0$,则称 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是负定二次型,$\boldsymbol{A}$ 叫做负定矩阵;

(3) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\leq 0$,且存在非零的 $\boldsymbol{X}_0$ 使 $\boldsymbol{X}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}_0=0$,就称 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是半负定二次型,$\boldsymbol{A}$ 叫半负定矩阵.

除此之外,其它的二次型称为不定二次型.

*定理 6.9　n 元二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 负定的充分必要条件有:

(1) $\boldsymbol{x}^{\mathrm{T}}(-\boldsymbol{A})\boldsymbol{x}$ 是正定二次型;

(2) $\boldsymbol{A}$ 的负惯性指数是 n,即 $\boldsymbol{A}$ 与 $-\boldsymbol{I}$ 合同;

(3) $\boldsymbol{A}$ 的顺序主子式负正相间,即奇数阶的主子式均小于零,偶数阶主子式均大于零.

(4) $\boldsymbol{A}$ 的特征值全是负数. ■

*例 8　证明三元二次型

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=-\boldsymbol{x}_1^2-2\boldsymbol{x}_2^2-3\boldsymbol{x}_3^2+2\boldsymbol{x}_1\boldsymbol{x}_2+2\boldsymbol{x}_2\boldsymbol{x}_3$$

是负定的.

证法 1　经配平方,有

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=-(x_1-x_2)^2-(x_2-x_3)^2-2x_3^2$$

对任何 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}\neq 0$,恒有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\leq 0$,而等式仅在

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

成立,即 $x_1 = x_2 = x_3 = 0$ 时才有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = 0$

所以, $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是负定二次型 . ■

证法 2 二次型的对应矩阵

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

其顺序主子式

$$\Delta_1 = -1 < 0, \quad \Delta_2 = \begin{vmatrix} -1 & 1 \\ 1 & -2 \end{vmatrix} = 1 > 0,$$

$$\Delta_3 = |\boldsymbol{A}| = \begin{vmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{vmatrix} = \begin{vmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -3 \end{vmatrix} = -2 < 0$$

所以, $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是负定二次型 . ■

习 题 6

6.1 写出下列二次型 $f(x_1, x_2, x_3)$ 的对应矩阵

(1) $x_1^2 + 4_3^2 - 2x_1x_2 + 6x_2x_3$,

(2) $2x_1^2 + x_2^2 + 5x_3^2 + 6x_1x_2 - 8x_1x_3 + 2x_2x_3$.

6.2 用配方法化二次型 $f(x_1, x_2, x_3)$ 为标准形,并写出所用坐标变换

(1) $x_1^2 + 2x_2^2 + 5x_3^2 + 2x_1x_2 + 6x_2x_3 + 2x_1x_3$,

(2) $2x_1^2 + 2x_2^2 + 7x_3^2 + 2x_1x_3 + 4x_2x_3$.

6.3 求下列实对称矩阵 $\boldsymbol{A}$ 的合同规范形,并求它所对应的二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的正、负惯性指数及二次型的秩 .

(1) $\boldsymbol{A} = \begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix}$, (2) $\boldsymbol{A} = \begin{matrix} 1 & 2 \\ 2 & 5 \end{matrix}$,

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

6.4　求正交变换 $\boldsymbol{x}=\boldsymbol{Ty}$ 化二次型为标准形

(1) $x_1^2+2x_1x_2+x_2^2$,

(2) $x_1^2-x_2^2+2x_3^2+4x_2x_3$,

(3) $3x_1^2+6x_2^2+3x_3^2-4x_1x_2-8x_1x_3-4x_2x_3$,

(4) $3x_1^2+3x_2^2+5x_3^2+4x_1x_3-4x_2x_3$.

6.5　判断下列二次型 $f(x_1,x_2,x_3)$ 是否正定？(要求用特征值及顺序主子式两种方法)

(1) $3x_1^2+3x_2^2+3x_3^2+2x_1x_2+2x_1x_3+2x_2x_3$,

(2) $x_1^2-2x_1x_2+2x_1x_3+2x_2x_3$,

(3) $x_1^2+x_2^2+2x_3^2+2x_1x_3+2x_2x_3$.

6.6　下列二次型是正定的，求参数 t 的取值范围

(1) $x_1^2+x_2^2+x_3^2+2tx_1x_3$,

(2) $x_1^2+3x_2^2+tx_3^2+2x_1x_2+2tx_2x_3$,

(3) $2x_1^2+2x_2^2+x_3^2+2tx_1x_2+6x_1x_3+2x_2x_3$.

6.7　选择题

(1) **如 $\boldsymbol{A}$ 是任一正定矩阵，则下列命题中正确的是( )**

A. **存在可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{I}$,**

B. **存在正交矩阵 $\boldsymbol{T}$，使 $\boldsymbol{T}^{-1}\boldsymbol{AT}=\boldsymbol{I}$,**

C. **$\boldsymbol{A}$ 必是对环可逆矩阵,**

D. **$\boldsymbol{A}$ 的负惯性指数为零是其充要条件 .**

(2) **与矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 2\\ 0 & 2 & 2\end{pmatrix}$ 合同的矩阵是( )**

A. $\begin{pmatrix}1 & & \\ & -1 & \\ & & 0\end{pmatrix}$,　*B*. $\begin{pmatrix}1 & & \\ & 1 & \\ & & -1\end{pmatrix}$,

$C.\begin{pmatrix}1&&\\&-1&\\&&-1\end{pmatrix}$, $D.\begin{pmatrix}-1&&\\&-1&\\&&0\end{pmatrix}$.

(3) 二次型 $f(x_1,x_2,x_3)=2x_1^2+2x_2^2+2x_3^2-2x_1x_2$ 在条件 $x_1^2+x_2^2+x_3^2=1$ 下的最大值是( ).

$A.1$, $B.3$, $C.5$, $D.7$.

(4) 与矩阵 $\boldsymbol{A}=\begin{pmatrix}2&0&2\\0&-3&0\\2&0&2\end{pmatrix}$ 既相似又合同的矩阵是( )

$A.\begin{pmatrix}1&&\\&-1&\\&&0\end{pmatrix}$, $B.\begin{pmatrix}2&&\\&-3&\\&&0\end{pmatrix}$,

$C.\begin{pmatrix}3&&\\&-4&\\&&0\end{pmatrix}$, $D.\begin{pmatrix}4&&\\&-3&\\&&0\end{pmatrix}$.

6.8 $\boldsymbol{A}$、$\boldsymbol{B}$ 都是 3 阶正定矩阵,证明 $\boldsymbol{A}+2\boldsymbol{B}$ 是正定矩阵.(提示:用正定的定义)

6.9 $\boldsymbol{A}$ 是 n 阶正定矩阵,证明 $\boldsymbol{A}$ 的伴随矩阵 $\boldsymbol{A}^*$ 是正定矩阵.(参看习题五第 3 题)

6.10 已知 $\lambda_1,\lambda_2,\lambda_3$ 是二次型 $f(x_1,x_2,x_3)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的三个特征值,若 $\lambda_1<\lambda_2<\lambda_3$.证明:对任意的 $x=(x_1,x_2,x_3)^{\mathrm{T}}$ 恒有

$$\lambda_1\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}\le\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\le\lambda_3\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}.$$

6.11 $\boldsymbol{A}$ 是 n 阶矩阵,证明 $\boldsymbol{A}$ 可逆的充分必要条件是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 为正定矩阵.

6.12 已知二次型 $f(x_1,x_2,x_3)=2x_1^2+3x_2^2+3x_3^2+2ax_2x_3$ $(a>0)$,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.

附　录　线性代数应用举例

1. 把连续问题转化为离散问题

为描述函数 $y = f(x)$,我们应当知道在每一点 x 处其函数值 $f(x)$,把每一点的函数值 $f(x)$看成未知数,这就成为一个未知数的个数是无穷的连续问题。数字计算机不能求连续问题的解。因此要设法转化成为离散问题用近似解来逼近。

例如　一根棒上的温度分布由微分方程

$$-\frac{\mathrm{d}^2 u}{\mathrm{d} x^2} = \mathrm{f}(x), \qquad 0 \leq x \leq 1 .$$

$$u(0) = 0, \qquad u(1) = 1 .$$

来描述,为了求出 $u(x)$,我们用差商来近似二阶导数,即

$$\frac{\mathrm{d}^2 u}{\mathrm{d} x^2} \approx \frac{u(x+h) - 2u(x) + u(x-h)}{h^2},$$

把区间[0,1] $n+1$ 等分,得到分点 $x_1 = h, x_2 = 2h, \cdots, x_n = nh$. 记 $u_i = u(ih)$, $u_0 = u(0)$, $u_{n+1} = u(1)$,那么方程的解 $u(x)$在各分点的近似值 u_i 应满足:

$$-u_{i+1} + 2u_i - u_{i-1} = h^2 \mathrm{f}(ih), \qquad (i = 1, 2, \cdots, n)$$

对每个 i 有一个这样的方程 .于是得到含有 n 个未知量 u_i, n 个方程所组成的一个线性方程组 .由于 $u_0 = u_{n+1} = 0$,这个方程组是:

$$\begin{array}{rcl}
2u_1 - u_2 & = & h^2 \mathrm{f}(h), \\
-u_1 + 2u_2 - u_3 & = & h^2 \mathrm{f}(2h), \\
-u_2 + 2u_3 - u_4 & = & h^2 \mathrm{f}(3h), \\
\cdots\cdots\cdots\cdots\cdots\cdots & & \\
-u_{n-2} + 2u_{n-1} - u_n & = & h^2 \mathrm{f}((n-1)h), \\
-u_{n-1} + 2u_n & = & h^2 \mathrm{f}(nh) .
\end{array}$$

容易看出,此方程组的系数行列式不为0(类似于1.2中例5),由克莱姆法则方程组有唯一解.利用高斯消元,就可求出$u(x)$在各分点的近似值u_i,因而可得到方程的近似解.显然,如果缩小差分的步长h,它将提高用差商代替导数的精度,同时对$u(x)$的描述也更细致一些,当然,由于未知数u_i个数的增多,计算工作量也就随之增大.

总之,一个连续问题可离散化,未知数的个数越多逼近连续问题的精度就越高,大线性方程组的求解问题是常会遇到的。

2. 矩阵对角化解微分方程组

在电路网络、溶液扩散振动理论等许多方面常会遇到线性常系数微分方程组的求解问题.如果利用矩阵特征值、特征向量的理论,转化为可分离变量型微分方程,问题的求解就方便许多.下面通过一个例题介绍这一方法.

例　解微分方程组

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -2x_1 + x_2,$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = x_1 - 2x_2.$$

解　引入矩阵记号,记

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix},$$

则方程组可表示为

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \boldsymbol{AX}$$

求出$\boldsymbol{A}$的特征值$\lambda_1 = -1, \lambda_2 = -3$. 相应的特征向量取为$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$和$\begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

由于 $\boldsymbol{A}$ 有两个不同的特征值，故 $\boldsymbol{A}$ 可以对角化．构造可逆矩阵

$$\boldsymbol{C} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

则有

$$\boldsymbol{C}^{-1}\boldsymbol{AC} = \begin{pmatrix} -1 & \\ & -3 \end{pmatrix},$$

这时微分方程可改写成

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \boldsymbol{C}\begin{pmatrix} -1 & \\ & -3 \end{pmatrix}\boldsymbol{C}^{-1}\boldsymbol{X},$$

引入新的待求函数 $\boldsymbol{Y}$，令

$$\boldsymbol{Y} = \boldsymbol{C}^{-1}\boldsymbol{X}, \qquad \boldsymbol{X} = \boldsymbol{CY},$$

则有

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \boldsymbol{C}\frac{\mathrm{d}\boldsymbol{Y}}{\mathrm{d}t},$$

代入到原方程组，消去 $\boldsymbol{X}$，可得 $\boldsymbol{Y}$ 的方程

$$\frac{\mathrm{d}\boldsymbol{Y}}{\mathrm{d}t} = \begin{pmatrix} -1 & \\ & -3 \end{pmatrix}\boldsymbol{Y},$$

也就是

$$\frac{\mathrm{d}y_1}{\mathrm{d}t} = -y_1,$$

$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = -3y_2.$$

可解出

$$y_1 = k_1\mathrm{e}^{-t}, \qquad y_2 = k_2\mathrm{e}^{-3t}$$

从而有

$$\boldsymbol{X} = \boldsymbol{CY} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} k_1\mathrm{e}^{-t} \\ k_2\mathrm{e}^{-3t} \end{pmatrix},$$

也就是说,原方程组的解是:

$$x_1 = k_1 e^{-t} - k_2 e^{-3t},$$

$$x_2 = k_1 e^{-t} + k_2 e^{-3t}.$$

3. 最小二乘法

最小二乘法是一种在工程技术、商业与经济等方面常用的求经验公式的方法.

例如,经实验得到一批数据资料

$$(x_i, y_i) \qquad i = 1, 2, \cdots, n.$$

通过把(x_i, y_i)当成点的坐标,把这些点描在坐标纸上,观察其图形,然后选用某函数来拟合、近似.

假若我们要用一次函数作为经验公式来刻化变量 x 与 y 之间的关系.把这批数据代入有:

$$\begin{aligned} kx_1 + b &= y_1, \\ kx_2 + b &= y_2, \\ &\cdots\cdots \\ kx_n + b &= y_n, \end{aligned}$$

写成矩阵的形式,为

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \cdots & \cdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} k \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}.$$

一般说来,由于公式本来就是近似的,而且实验数据测量的本身也会有各种误差,所以上述方程组的系数矩阵及增广矩阵的秩不大可能会相等,也就是说方程组无解.

如果确实需要有一个近似公式,那么是任取两组实验数据来确定 k 和 b 合理,还是用这一批 n 组数据来求,并选用使对各组误差的平方和最小的 k 和 b 更合理、更可靠呢?

对线性方程组

$$\boldsymbol{AX} = \beta \qquad (\boldsymbol{A} \text{ 是 } m \times n \text{ 矩阵})$$

如果方程组无解,那么

$$" \boldsymbol{X}, \qquad \boldsymbol{AX} - \beta \neq 0 .$$

这样,$\boldsymbol{AX} - \beta$ 作为 n 维向量(非 0)其内积大于 0,即

$$\| \boldsymbol{AX} - \beta \|^2 = (\boldsymbol{AX} - \beta, \boldsymbol{AX} - \beta) > 0$$

若 $\boldsymbol{X}_0$ 使得 $\| \boldsymbol{AX} - \beta \|^2$ 取最小值,则称 $\boldsymbol{X}_0$ 是方程组 $\boldsymbol{AX} = \beta$ 的最小二乘解 .

我们通过例题介绍如何来求最小二乘解:

例 1　对 x 测试三次,有关系式

$$3x = 10,$$

$$4x = 7,$$

$$5x = 8 .$$

求在最小二乘意义下 x 的值 .

解　设

$$\boldsymbol{A} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}, \qquad \boldsymbol{X} = (x), \qquad \beta = \begin{pmatrix} 10 \\ 7 \\ 8 \end{pmatrix} .$$

所测关系式可用矩阵表示为

$$\boldsymbol{AX} = \beta .$$

由于

$$\begin{aligned} \| \boldsymbol{AX} - \beta \|^2 &= (\boldsymbol{AX} - \beta, \boldsymbol{AX} - \beta) = (\boldsymbol{AX} - \beta)^{\mathrm{T}} (\boldsymbol{AX} - \beta) \\ &= \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A} x^2 - 2 \boldsymbol{A}^{\mathrm{T}} \beta x + \beta^{\mathrm{T}} \beta \qquad (\text{注 } \boldsymbol{A}^{\mathrm{T}} \beta = \beta^{\mathrm{T}} \boldsymbol{A}) \\ &= 50 x^2 - 196 x + 213 \end{aligned}$$

二次函数的最小值在

$$x = \frac{196}{2 \times 50} = 1.96$$

时达到,所以在最小二乘意义下,x 的最佳取值是 $x = 1.96$.

一般地,关于一个变量的方程组 $\boldsymbol{AX}=\boldsymbol{\beta}$ 的最小二乘解是

$$x=\frac{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\beta}}{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}}.$$

例 2 求拟合(1,13),(2,18),(3,22),(4,29)四个点的最小二乘直线的方程.

解 设所求最小二乘直线的方程是 $y=kx+b$. 把四个点的坐标代入,得到方程组

$$\begin{pmatrix}1 & 1\\ 2 & 1\\ 3 & 1\\ 4 & 1\end{pmatrix}\begin{pmatrix}k\\ b\end{pmatrix}=\begin{pmatrix}13\\ 18\\ 22\\ 29\end{pmatrix}.$$

令

$$\boldsymbol{A}=\begin{pmatrix}1 & 1\\ 2 & 1\\ 3 & 1\\ 4 & 1\end{pmatrix},\quad \boldsymbol{X}=\begin{pmatrix}k\\ b\end{pmatrix},\quad \boldsymbol{\beta}=\begin{pmatrix}13\\ 18\\ 22\\ 29\end{pmatrix}$$

由于

$$\|\boldsymbol{AX}-\boldsymbol{\beta}\|^2=(k+b-13)^2+(2k+b-18)^2+(3k+b-22)^2+(4k+b-29)^2$$

二元函数 $f(k,b)=\|\boldsymbol{AX}-\boldsymbol{\beta}\|^2$ 取最小值的必要条件是

$$\frac{\partial f}{\partial k}=0,\qquad \frac{\partial f}{\partial b}=0$$

即

$$(k+b-13)+2(2k+b-18)+3(3k+b-22)+4(4k+b-29)=0,$$

$$(k+b-13)+(2k+b-18)+(3k+b-22)+(4k+b-29)=0,$$

整理后,为

$$30k+10b=231,$$

$$10k+4b=82.$$

解出：　　$k=5.2,\quad b=7.5$

所以，最小二乘直线的方程是 $y=5.2x+7.5$.

利用向量空间的理论，可以证明

方程组 $\boldsymbol{AX}=\beta$ 的最小二乘解也就是方程组 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{A}^{\mathrm{T}}\beta$ 的解.

注　请验证在例 2 中

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}30 & 10\\ 10 & 4\end{pmatrix},\qquad \boldsymbol{A}^{\mathrm{T}}\beta=\begin{pmatrix}231\\ 82\end{pmatrix}.$$

4．编码问题

在引入矩阵概念时，我们已提到现代通讯可以是数字通讯．但数字信息在传递过程中可能会遇到种种干扰，难免出现差错，那么如何提高传递的可靠性呢？

假若我们用 16 个符号 $\alpha_1-\alpha_{16}$ 可构成各种信息，为表示这 16 个符号，在二进制下只须选用四维向量：

$$(0,0,0,0)^{\mathrm{T}},\quad (0,0,0,1)^{\mathrm{T}},\quad (0,0,1,0)^{\mathrm{T}},\cdots,(1,1,1,1)^{\mathrm{T}}$$

分别表示每一个符号就可以了．每个这样的向量叫做信息码，向量的维数叫码长．

如果只有信息码：那么传递时是否有失误就不好判断了，为此要增加校验位，以确定传递是否正确，如有差错，是错在哪一个位置？

为了纠正信息码上一个错误，要用 3 个校验位*，这样构成码长为 7 的编码

$$\alpha=(p_1\ \ p_2\ \ m_1\ \ p_3\ \ m_2\ \ m_3\ \ m_4)^{\mathrm{T}}$$

其中 $m_1\sim m_4$ 是信息码，$p_1\sim p_3$ 是校验位．

纠错的原理是：构造齐次方程组 $\boldsymbol{HX}=0$．使这 16 个 α_i 是方程组的解．如果接到编码 β 不是方程的解，假如第 i 位有错，在二进制下，我们的运算应在模 2 数域上进行，其运算规律是：

加法　$0+0=1+1=0,\quad 0+1=1+0=1;$

乘法 $1\times0=0\times1=0\times0=0, 1\times1=1$.

因此,必有

$$\beta=\alpha_j+e_i,$$

其中

$$e_i=(0,\cdots,0,\ 1,0,\cdots,0)^{\mathrm{T}} \qquad \text{第 } i \text{ 位}$$

由于

$$H\beta=H(\alpha_j+e_i)=H\alpha_j+He_i=He_i,$$

而我们要求

$$He_i=H\begin{pmatrix}0\\ \cdots\\ 0\\ 1\\ 0\\ \cdots\\ 0\end{pmatrix}$$

正好是矩阵 H 中的第 i 列 . 因此,只要检查 $H\beta$ 与 H 中第几列一样,就知道 β 中那一个位置上的信息量错了 . 而每个位置上信息只有 0 与 1 两种,也就很容易改正传递上的差错了 .

例如,给出校验矩阵 H

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

如收到信号 $\beta=(1\ 0\ 1\ 0\ 0\ 1\ 1)^{\mathrm{T}}$,则由

$$H\beta=\begin{pmatrix}0\\ 1\\ 1\end{pmatrix}\neq 0$$

知 β 有差错,从 $H\beta$ 与 H 的第 3 列一样,得知 β 中第 3 个位置的

信号有错．那么，正确的信号就应当是

$$(1\ 0\ 0\ 0\ 0\ 1\ 1)^{\mathrm{T}},$$

对照信息码的位置，知传递的符号是

$$(0\ 0\ 1\ 1)^{\mathrm{T}}.$$

* 从方程组 $\boldsymbol{HX}=0$，由系数矩阵的秩为 3，故要有 4 个自由变量，才能得到 $2^4=16$ 个解，因此方程组中应有 $3+4=7$ 个未知数，现在信息码长为 4，所以要添加 3 个校验位．

部分习题答案及提示

习题 1

1.1 (1) 18; (2) 2; (3) −37; (4) 6.

1.3 (1) ±1, 3; (2) 2, 5, −6; (3) 0(二重根), 3; (4) 0, 1, −3. 把第 3 列加至第 1 列.

1.4 (1) $A_{12}=0$, $A_{22}=6$, $A_{32}=-5$. (2) $A_{12}=-1$, $A_{22}=1$, $A_{32}=2$, $A_{42}=2$.

1.5 (1) −8; (2) 0; (3) 120; (4) 240.

1.6 (1) $x=-2,\ y=3,$ (2) $x=\cos\theta,\ y=-\sin\theta,$ (3) $x_1=0,\ x_2=0,\ x_3=0,$ (4) $x_1=3,\ x_2=-1,\ x_3=-2.$

1.7 (1) ±1, 2, (2) −1(三重根). 1 列 + 2 列 − 3 列.

1.8 (1) 把第一列变换出 0, (2) 把 2、3 列加至第 1 列.

1.9 (1) −7, (2) −2, (3) a^3, (4) $-a^2b^3$.

1.10 设法证明 $D=-D$, 注意 $D=D^{\mathrm{T}}$.

1.11 (1) $a^n+(-1)^{n+1}b^n$, (2) $n!$ (把 D 化成上三角行列式), (3) $\frac{1}{2}n(n+1)$, 把各列均加至第 1 列.

习题 2

2.1 $\begin{pmatrix}9 & 0\\ 12 & -3\end{pmatrix}$; $\begin{pmatrix}0 & -6\\ 3 & -9\end{pmatrix}$; $\begin{pmatrix}5 & 2\\ 7 & 0\end{pmatrix}$; $\begin{pmatrix}19 & 37\\ -14 & -28\end{pmatrix}$;

$\begin{pmatrix} -1 & 2 \\ -5 & 6 \end{pmatrix}$; 18 .

2.5 (1) $-\frac{1}{5}\begin{pmatrix} -2 & 3 \\ -1 & 4 \end{pmatrix}$, (2) $\frac{1}{4}\begin{pmatrix} 3 & 1 \\ -2 & 0 \end{pmatrix}$,

(3) $-\frac{1}{5}\begin{pmatrix} 1 & 0 & -2 \\ -2 & -5 & 4 \\ -3 & 0 & 1 \end{pmatrix}$, (4) $\begin{pmatrix} 1 & -1 & -1 \\ & 1 & -1 \\ & & 1 \end{pmatrix}$.

2.6 (1) $\begin{pmatrix} 3 & -2 & 8 \\ 5 & -4 & 15 \\ -7 & 4 & -2 \end{pmatrix}$, (2) $\begin{pmatrix} 2 & -18 \\ 0 & 6 \end{pmatrix}$,

(3) $\begin{pmatrix} 7 & -4 \\ 8 & -4 \\ -5 & 3 \end{pmatrix}$, (4) $\frac{1}{2}\begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}$, (5) $\begin{pmatrix} 4 & 4 \\ 5 & 7 \end{pmatrix}$.

2.8 (1) $\begin{pmatrix} t & u \\ 0 & t \end{pmatrix}$, (2) $\begin{pmatrix} 2t+u & 2t \\ t & u \end{pmatrix}$. t, u **是任意实数.**

2.10 **注意** $\boldsymbol{A}\cdot\boldsymbol{A}^* = |\boldsymbol{A}|\boldsymbol{I}$,**现** $\boldsymbol{A}$ **是** 3 **阶矩阵,两边取行列式后,** $|\boldsymbol{A}^*| = |\boldsymbol{A}|^2$,**应选择**(4)**.**

2.11 **选**(2)、(3)、(4)**. 注:** $(\boldsymbol{AB})^2 = (\boldsymbol{AB})(\boldsymbol{AB}) = \boldsymbol{A}(\boldsymbol{BAB})$.

2.12 $|\boldsymbol{A}| = -\frac{1}{4}$, $\boldsymbol{A}^{-1} = \begin{pmatrix} 1 & & \\ & -10 & 6 \\ & 4 & -2 \end{pmatrix}$, $(\boldsymbol{A}^*)^{-1}$

$= \begin{pmatrix} -4 & & \\ & -2 & -6 \\ & -4 & -10 \end{pmatrix}$. **参看** 2.2 **例** 4**.**

2.17 **联系 2.4 例 1**

$\frac{1}{3}$ $\quad\quad$ $\begin{matrix}2 & -3\\ -1 & 2\end{matrix}$

2.20 (1) 1 $\quad\quad$, $\quad$ (2) $\begin{pmatrix}1 & -1 & 0\\ -1 & 2 & 1\\ 1 & -1 & -1\end{pmatrix}$.

$\frac{1}{2}$

2.22 $\boldsymbol{A}^{-1}=\frac{1}{2}(\boldsymbol{A}-\boldsymbol{I})$, $(\boldsymbol{A}+2\boldsymbol{I})^{-1}=\frac{1}{4}(3\boldsymbol{I}-\boldsymbol{A})$.

注: $(\boldsymbol{A}+2\boldsymbol{I})\cdot(\boldsymbol{A}-3\boldsymbol{I})=\boldsymbol{A}^2-\boldsymbol{A}-6\boldsymbol{I}$.

2.23 $\boldsymbol{A}^{-1}=\frac{1}{2}(3\boldsymbol{I}-\boldsymbol{A})$. **注:** $\boldsymbol{A}^2=(\boldsymbol{I}+\boldsymbol{B})^2=\boldsymbol{I}+3\boldsymbol{B}=$

2.26 $[(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}+\boldsymbol{A})^{-1}]^{\mathrm{T}}[(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}+\boldsymbol{A})^{-1}]$

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

$=[(\boldsymbol{I}+\boldsymbol{A})^{\mathrm{T}}]^{-1}(\boldsymbol{I}^{\mathrm{T}}-\boldsymbol{A}^{\mathrm{T}})(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}+\boldsymbol{A})^{-1}=\cdots$

2.27 **用归纳法、分块技巧**:

$$\begin{pmatrix}a_{11} & \alpha\\ 0 & \boldsymbol{A}_{n-1}\end{pmatrix}\begin{pmatrix}b_{11} & \beta\\ 0 & \boldsymbol{B}_{n-1}\end{pmatrix}=\begin{pmatrix}a_{11}b_{11} & a_{11}\beta+\alpha\boldsymbol{B}_{n-1}\\ 0 & A_{n-1}\boldsymbol{B}_{n-1}\end{pmatrix}$$

这里 $\boldsymbol{A}_{n-1}$, $\boldsymbol{B}_{n-1}$ **是** $n-1$ **阶上三角矩阵**, α、β **是** $1\times(n-1)$ **矩阵**, **0 是** $(n-1)\times1$ **矩阵**

2.28 (1) $\begin{pmatrix}1 & n\\ 0 & 1\end{pmatrix}$. **注**: $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}^n=\left(\boldsymbol{I}+\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\right)^n$

$=\boldsymbol{I}+n\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$ **一般地**, $(\boldsymbol{A}+\boldsymbol{B})^2\neq\boldsymbol{A}^2+2\boldsymbol{A}\boldsymbol{B}+\boldsymbol{B}^2$, **两者不要混淆**

.

(2) $3^{n-1}\begin{pmatrix}1 & 1\\ 2 & 2\end{pmatrix}$. **因** $\begin{pmatrix}1 & 1\\ 2 & 2\end{pmatrix}=\begin{pmatrix}1\\ 2\end{pmatrix}(1 \quad 1)$, **参看** 2.1 **中例** 8.

2.29 (1) $f(\boldsymbol{A})=\begin{pmatrix}|0| & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}$

注: $\boldsymbol{A}^2 = \boldsymbol{A}^3 = \cdots = \boldsymbol{A}^{100} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

(2) $\boldsymbol{B}^{-1} = -\frac{1}{3}(\boldsymbol{B} + 2\boldsymbol{I})$. 注: $f(\boldsymbol{B}) = \boldsymbol{B}^2 + 2\boldsymbol{B} + 3\boldsymbol{I}$.

2.30 **$\boldsymbol{A}$ 是上三角阵, 可证 $\boldsymbol{A}^{-1}$ 是上三角矩阵, 再设** $\boldsymbol{A}^{-1} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ & b_{22} & b_{23} \\ & & b_{33} \end{pmatrix}$, **由 $\boldsymbol{A}$ 是正交的, $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}}$ 可证** $a_{ij} = 0\ (i \neq j)$. **利用 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$ 来证** a_{ii}**为** $+1$ **或** -1.

习题 3

3.1 (1) $x_1 = 1$, $x_2 = 2$, $x_3 = 1$; (2) $x_1 = 2t - u$, $x_2 = t$, $x_3 = u$, $x_4 = 1$; (3) **无解**; (4) $x_1 = -8$, $x_2 = t + 3$, $x_3 = 2t + 6$, $x_4 = t$.

3.4 (1) $t = 15$, (2) $t \neq 12$, (3) $t \neq 2$ **且** $t \neq 3$ **时**, β **表示式唯一**; $t = 2$ **时**, β **表示法不唯一** .

3.5 (1) **线性相关**, $\alpha_3 = \alpha_1 - 2\alpha_2$; (2) **线性相关**, α_3 **不能由** α_1, α_2 **线性表出**; (3) **线性相关**, $\alpha_3 = -\alpha_1 + 2\alpha_2$, (4) **线性无关** .

3.6 (1) **错 . 例如** $(1,0)$, $(0,1)$ **和** $(1,1)$; (2) **错 . 例如** $\alpha_1 = (1,0)$, $\alpha_2 = (2,0)$ **与** $\beta_1 = (0,2)$, $\beta_2 = (0,3)$; (3) **正确 . 定理** 3.3 **与** 3.6; (4) **正确 . 定理** 3.14.

3.7 **不一定(参看习题中** 5(2))

3.8 **参看 3.2 例 12.**

3.10 α_1, α_2, α_3 **与** α_2, α_3, α_4 **是等价向量组 . 或用定义, 对** $k_2\alpha_2 + k_3\alpha_3 + k_4\alpha_4 = 0$, **把** α_4 **代入, 再用** α_1, α_2, α_3 **线性无关** .

3.11 (1) $r=1$, α_2 是极大线性无关组 . (2) $r=2$, α_1, α_3 是极大线性无关组 .(3) $r=2$, α_1, α_2 是极大线性无关组 .

3.13 $a=2$

3.18 (1) 错 . 可能无解,如 $\begin{cases} x+y=3, \\ x+y=2, \\ x-y=1 . \end{cases}$ (2) 正确 . $\boldsymbol{A}$ 的列向量必线性相关,表示法不可能唯一; (3) 错 . 参看第 9 题; (4) 错 . $\boldsymbol{A}$ 不一定是 n 阶矩阵 .

3.19 (1) $\eta=(6,\ -2,\ 1)^{\mathrm{T}}$; (2) $\eta_1=(8,\ -6,1,0)^{\mathrm{T}}$, $\eta_2=(-7,\ 5,\ 0,\ 1)^{\mathrm{T}}$; (3) $\eta_1=\left(\frac{7}{6},\ \frac{5}{6},\ 0,\ \frac{1}{3},\ 1\right)^{\mathrm{T}}$, $\eta_2=(-1,\ 1,\ 1,\ 0,\ 0)^{\mathrm{T}}$; (4) $\eta_1=(1,\ 0,\ 1,\ 0,\ 0)^{\mathrm{T}}$, $\eta_2=(0,\ 1,\ 0,\ 1,\ 0)^{\mathrm{T}}$.

3.20 (1) $\left(\frac{1}{3},\ 0,\ 0,\ 1\right)^{\mathrm{T}}+k_1(4,\ -3,\ 0,\ 0)^{\mathrm{T}}+k_2(1,\ 0,\ -3,\ 0)^{\mathrm{T}}$; (2) $(0,\ 0,\ 2,\ 0)^{\mathrm{T}}+k(1,\ -5,\ 11,\ 0)^{\mathrm{T}}$

3.21 (1) $a=7$ 时有解,$(1,1,0)^{\mathrm{T}}+k(-7,\ 5,\ 1)^{\mathrm{T}}$; (2) $a=0$或 3 时有解 . 当 $a=0$ 时,$(4,0,\ -2)^{\mathrm{T}}+k(1,\ -1,\ 0)^{\mathrm{T}}$, 当 $a=3$ 时,$(1,0,1)^{\mathrm{T}}+k(1,\ -1,\ 0)^{\mathrm{T}}$; (3) $a=1$ 或 -2 时有解,当 $a=1$ 时,$(1,\ 0,\ 0)^{\mathrm{T}}+k(0,\ 1,\ 1)^{\mathrm{T}}$, 当 $a=-2$ 时,$(2,\ 2,\ 0)^{\mathrm{T}}+k(1,\ 1,\ 1)^{\mathrm{T}}$.

3.22 $a_1+a_2=b_1+b_2$ 时有解,$(a_1-b_2,\ b_2,\ a_2,\ 0)^{\mathrm{T}}+k(1,\ -1,\ -1,\ 1)^{\mathrm{T}}$.

3.23 (1) 当 $a=1$, $b=3$ 时,$\beta=(1-t)\alpha_1+t\alpha_2+\alpha_3$ 表示法不唯一; (2) 当 $b=2$ 或 $a=1$ 但 $b\neq 3$ 时 β 不能由 α_1, α_2, α_3 线性表出 .

3.24 对 $\boldsymbol{A}=(\alpha_1,\alpha_2,\cdots,\alpha_n)$按列分块, $\boldsymbol{Ax}=\beta$ 即 $x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=\beta$,

" β, $\boldsymbol{Ax}=\beta$ 有解

$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 可表示任一 m 维向量 $\boldsymbol{\beta}$

$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 可表示 $\boldsymbol{\varepsilon}_1 = (1, 0, \cdots, 0)^{\mathrm{T}}$, $\boldsymbol{\varepsilon}_2 = (0, 1, \cdots, 0)^{\mathrm{T}}, \cdots$, $\boldsymbol{\varepsilon}_m = (0, 0, \cdots, 1)^{\mathrm{T}}$

$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 与 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_m$ 是等价向量组

$r(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = r(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_m)$

$r(\boldsymbol{A}) = m$.

3.25　$\boldsymbol{A} \cdot \boldsymbol{A}^* = |\boldsymbol{A}| \boldsymbol{I}$, 如 $r(\boldsymbol{A}) < n$, 则 $|\boldsymbol{A}| = 0$. 若 $r(\boldsymbol{A}) < n - 1$, $\boldsymbol{A}$ 的 $n-1$ 阶行列式?

习题 4

4.1 $(2, 1, -1)^{\mathrm{T}}$.

4.2 $\begin{matrix} -1 & 0 & 1 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{matrix}$, 3. $\begin{matrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \end{matrix}$, 4. $\begin{matrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{matrix}$, $\begin{matrix} 1 \\ 2 \\ 2 \end{matrix}$.

4.5 7, 11, 2, $\arccos \frac{2}{77}$;

4.6 $\pm \frac{1}{2}(1, 0, 1)^{\mathrm{T}}$

4.7 (1) $\frac{1}{3}(1, 1, 1)^{\mathrm{T}}$, $\frac{1}{6}(-2, 1, 1)^{\mathrm{T}}$, $\frac{1}{2}(0, -1, 1)^{\mathrm{T}}$;

$\alpha = \left(0, -\frac{3}{6}, -\frac{1}{2}\right)^{\mathrm{T}}$,

(2) $\frac{1}{3}(1, 1, 1)^{\mathrm{T}}$, $\frac{1}{6}(-1, 2, -1)^{\mathrm{T}}$, $\frac{1}{2}(-1, 0, 1)^{\mathrm{T}}$; $\alpha =$

$\left(0, -\frac{3}{6}, \frac{1}{2}\right)^{\mathrm{T}}$,

(3) $\frac{1}{3}(1, -1, 1)^{\mathrm{T}}$, $\frac{1}{6}(-1, 1, 2)^{\mathrm{T}}$, $\frac{1}{2}(1, 1, 0)^{\mathrm{T}}$; $\alpha =$

$\left(\frac{2}{3}, -\frac{2}{6}, 0\right)^{\mathrm{T}}$.

4.8 (1) **行空间** $R(\boldsymbol{A}^{\mathrm{T}})$**的基** $\alpha_1 = (1, 2, 0)$, $\alpha_2 = (0, 1, 1)$ **列空间** $R(\boldsymbol{A})$**的基** $\beta_1 = (1, 0, 1)^{\mathrm{T}}$, $\beta_2 = (0, 1, 0)^{\mathrm{T}}$.

(2) **行空间一维,基**: $(1, 1, 2)$; **列空间一维, 基是** $\begin{matrix} 1 \\ 2 \end{matrix}$.

4.9 (1) **解空间一维,标准正交基**: $\frac{1}{6}(2, -1, 1)^{\mathrm{T}}$

(2) **解空间二维**, $\frac{1}{2}(-1, 1, 0)^{\mathrm{T}}$, $\frac{1}{3}(1, 1, -1)^{\mathrm{T}}$ **是标准正交基**.

4.10 3维. $\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $\begin{pmatrix}0&1\\0&0\end{pmatrix}$, $\begin{pmatrix}0&0\\0&1\end{pmatrix}$ 是一组基.

4.11 $(-1,-1,-1,4)^{\mathrm T}$.

4.13 $\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix}$, $\begin{pmatrix}1&0&0\\-1&0&1\\2&2&0\end{pmatrix}$.

4.15 (1) $\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$, (2) $\begin{pmatrix}1\\-1\\2\end{pmatrix}$, (3) $\begin{pmatrix}1\\1\\-1\end{pmatrix}$.

4.16 如 $k_1\alpha_1+k_2\alpha_2+k_3\alpha_3+l\beta=0$, 用 β 作内积可得 $l(\beta,\beta)=0$.

4.17 构造齐次线性方程组 $\boldsymbol{Ax}=0$, 其中

$$\boldsymbol{A}=\begin{pmatrix}\alpha_1^{\mathrm T}\\\alpha_2^{\mathrm T}\\\alpha_3^{\mathrm T}\end{pmatrix}$$

是 3×4 矩阵, 且 $r(\boldsymbol{A})=r(\alpha_1,\alpha_2,\alpha_3)=3$. 基础解系?

4.18 $\alpha_1,\alpha_2,\alpha_3$ 是一组基, 可设 $\beta=x_1\alpha_1+x_2\alpha_2+x_3\alpha_3$, 考查 $(\beta,\beta)=$?

习题 5

5.1 (1) $\lambda_1=2$, $\boldsymbol{X}_1=(5,2)^{\mathrm T}$; $\lambda_2=-1$, $\boldsymbol{X}_2=(1,1)^{\mathrm T}$; (2) $\lambda_1=2$, $\boldsymbol{X}_1=(1,-1)^{\mathrm T}$; $\lambda_2=9$, $\boldsymbol{X}_2=(2,5)^{\mathrm T}$; (3) $\lambda_1=1$, $\boldsymbol{X}_1=(1,0,0)^{\mathrm T}$, $\lambda_2=\lambda_3=0$, $\boldsymbol{X}_2=(1,-1,0)^{\mathrm T}$, $\boldsymbol{X}_3=(1,0,-1)^{\mathrm T}$; (4) $\lambda_1=1$, $\boldsymbol{X}_1=(1,0,-1)^{\mathrm T}$; $\lambda_2=3$, $\boldsymbol{X}_2=(1,-2,1)^{\mathrm T}$; $\lambda_3=0$, $\boldsymbol{X}_3=(1,1,1)^{\mathrm T}$; (5) $\lambda_1=\lambda_2=\lambda_3=0$, $\boldsymbol{X}_1=(1,0,0)^{\mathrm T}$; (6) $\lambda_1=\lambda_2=\lambda_3=-1$(参看习题1 1.7(2)), $\boldsymbol{X}_1=(-1,-1,1)^{\mathrm T}$.

5.2 (1) ±1; (2) 2; (3) 1, -3.

5.3 如 λ 是 $\boldsymbol{A}$ 的特征值, 则 $\boldsymbol{A}^{-1}$ 的特征值是 $\dfrac{1}{\lambda}$, $\boldsymbol{A}^*$ 的特征值是 $\dfrac{1}{\lambda}|\boldsymbol{A}|$.

5.6 (1) D; (2) B; (3) C; (4) D.

5.8 **0**. **提示**: **定理** 5.4 **与** 5.3.

5.9 (1) $\boldsymbol{\Lambda}=\begin{pmatrix}2&\\&-1\end{pmatrix}$, $\boldsymbol{P}=\begin{pmatrix}5&1\\2&1\end{pmatrix}$; (2) $\boldsymbol{\Lambda}=\begin{pmatrix}2&\\&9\end{pmatrix}$, $\boldsymbol{P}=\begin{pmatrix}1&2\\-1&5\end{pmatrix}$; (3) $\boldsymbol{\Lambda}=\begin{pmatrix}1&&\\&0&\\&&0\end{pmatrix}$, $\boldsymbol{P}=\begin{pmatrix}1&1&1\\0&-1&0\\0&0&-11\end{pmatrix}$; (4) $\boldsymbol{\Lambda}=\begin{pmatrix}1&&\\&3&\\&&0\end{pmatrix}$, $\boldsymbol{P}=\begin{pmatrix}1&1&1\\0&-2&1\\-1&1&1\end{pmatrix}$; (5), (6) **不能对角化**.

5.11 $\boldsymbol{A}=\begin{pmatrix}2&-8\\0&-2\end{pmatrix}$

5.12 **不相似**. $\boldsymbol{A}\sim\Lambda$ **而** $\boldsymbol{B}$ **不能对角化**.

5.13 (1) $\boldsymbol{\Lambda}=\begin{pmatrix}2&\\&-3\end{pmatrix}$, $\boldsymbol{T}=\begin{pmatrix}\frac{2}{5}&\frac{1}{5}\\\frac{1}{5}&-\frac{2}{5}\end{pmatrix}$;

(2) $\boldsymbol{\Lambda}=\begin{pmatrix}1&&\\&1&\\&&-1\end{pmatrix}$, $\boldsymbol{T}=\begin{pmatrix}\frac{1}{2}&0&\frac{1}{2}\\0&1&0\\\frac{1}{2}&0&-\frac{1}{2}\end{pmatrix}$;

(3) $\boldsymbol{\Lambda}=\begin{pmatrix}3&&\\&1&\\&&-3\end{pmatrix}$,

$$\boldsymbol{T}=\begin{pmatrix} \frac{1}{3} & -\frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & 0 & -\frac{2}{6} \end{pmatrix};$$

(4) $\Lambda=\begin{pmatrix} 3 & & \\ & 3 & \\ & & 0 \end{pmatrix}$, $\boldsymbol{T}=\begin{pmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ 0 & -\frac{2}{6} & \frac{1}{3} \\ -\frac{1}{2} & \frac{1}{6} & \frac{1}{3} \end{pmatrix}$.

5.14 **6 .**

5.15 $6^{n-1}\begin{pmatrix} 2 & 2 \\ 4 & 4 \end{pmatrix}$,

5.17 $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$

5.18 **不能．因为** $r(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=3$,**而** λ_0 **是二重根**,**至多有** 2 **个线性无关的特征向量．**

5.19 $\pm$**2**, 8. $\boldsymbol{B}$ **能对角化**(5.3 **推论** 3).

5.20 **-3. 注意** $\boldsymbol{A}\sim\begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}$

习题 6

6.2 (1) $y_1^2+y_2^2$, $\begin{cases} x_1=y_1-y_2+y_3, \\ x_2=y_2-2y_3, \\ x_3=y_3, \end{cases}$

(2) $2y_1^2+2y_2^2+\frac{9}{2}y_3^2$, $\begin{cases} x_1 = y_1 - \frac{1}{2}y_3, \\ x_2 = y_2 - y_3, \\ x_3 = y_3. \end{cases}$

6.3 (1) $\begin{matrix} 1 & \\ & 0 \end{matrix}$, $p=1$, $q=0$, $r=1$,

(2) $\begin{matrix} 1 & \\ & 1 \end{matrix}$, $p=2$, $q=0$, $r=2$,

(3) $\begin{matrix} 1 & & \\ & 1 & \\ & & -1 \end{matrix}$, $p=2$, $q=1$, $r=3$,

(4) $\begin{matrix} 1 & & \\ & -1 & \\ & & -1 \end{matrix}$, $p=1$, $q=2$, $r=3$.

(提示:可用配方法化为标准形,或看特征值的正负号)

6.4 (1) $2y_1^2$, $\boldsymbol{T}=\begin{matrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{matrix}$,

(2) $y_1^2+3y_2^2-2y_3^2$, $\boldsymbol{T}=\begin{matrix} 1 & 0 & 0 \\ 0 & \frac{1}{5} & -\frac{2}{5} \\ 0 & \frac{2}{5} & \frac{1}{5} \end{matrix}$,

(3) $7y_1^2+7y_2^2-2y_3^2$, $\boldsymbol{T}=\begin{matrix} \frac{1}{2} & \frac{1}{3\ 2} & \frac{2}{3} \\ 0 & -\frac{4}{3\ 2} & \frac{1}{3} \\ -\frac{1}{2} & \frac{1}{3\ 2} & \frac{2}{3} \end{matrix}$,

(4) $y_1^2+3y_2^2+7y_3^2$, $\boldsymbol{T}=\begin{pmatrix} -\frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & -\frac{1}{6} \\ \frac{1}{3} & 0 & \frac{2}{6} \end{pmatrix}$.

6.5 (1) **正定.(特征值是:**2**(**二重**),**5**),**

(2) **不定.(特征值是:**1**,** $\pm$ 3**),**

(3) **半正定.(求特征值时可**(1)**行**+(2)**行**-(3)**行)**

6.6 (1) $|t|<1$, (2) $t<0$ **或** $t>2$, (3) **不存在(对任何** t, f **均不可能正定)**.

6.7 (1) C, (2) B, (3) B, (4) D.

提示: **在**(3)**中,用正交变换** $\boldsymbol{x}=\boldsymbol{T}\boldsymbol{y}$ **化二次型为标准形,注意** $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{T}^{\mathrm{T}}\boldsymbol{T}\boldsymbol{y}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$.

6.9 **如** λ **是** $\boldsymbol{A}$ **的特征值,则** $\boldsymbol{A}^*$ **的特征值是** $\frac{|\boldsymbol{A}|}{\lambda}$**,由** $\boldsymbol{A}$ **正定可证** $\boldsymbol{A}^*$ **的特征值全大于零.**

或者, $\boldsymbol{A}^*=|\boldsymbol{A}|\boldsymbol{A}^{-1}$**,对任** $\boldsymbol{x}\neq 0$**,由**

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^*\boldsymbol{x}=|\boldsymbol{A}|\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{x}$$

用 $\boldsymbol{A}$、$\boldsymbol{A}^{-1}$ **正定可证.**

6.10 **用正交变换化二次型为** $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\lambda_1 y_1^2+\lambda_2 y_2^2+\lambda_3 y_3^2$, **看** 6.7(3)**提示.**

6.11 **如** $\boldsymbol{A}$ **可逆,则** $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{I}\boldsymbol{A}$**,有** $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$D $\boldsymbol{I}$

如 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ **正定,则** $|\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}|>0$

6.12 $a=2$, $\boldsymbol{T}=\begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$.

注意： $\boldsymbol{A} = \begin{matrix} 2 & & \\ & 3 & a \\ & a & 3 \end{matrix}$ **的特征值是** $1, 2, 5$**. 代到特征方程可求出** a**.**