本书分上、下两册，内容包括四部分：线性方程组 n 元数组的向量空间和矩阵理论；一元多项式环和多元多项式环的理论；任意域上的线性空间和线性映射的理论；具有度量的线性空间包括欧几里得空间、酉空间、正交空间和辛空间的理论。

本书注意渗透现代的观点，力图体现代数与几何、分析的联系，注重联系实际，努力培养学生的数学素质。本书写得深入浅出，特别注意阐述清楚所讨论问题的想法。每一节后面都配有一定数量的习题。

本书可作为大学数学系、概率统计系和应用数学系的高等代数课程的教材，它也是一部内容丰富的教学参考书，可供有关师生参考。

序　　言

为了把学生培养成为面向 21 世纪的高水平人才，作者积多年讲授高等代数、抽象代数和群表示论等课程的经验以及从事科研工作的体会，写了一套高等代数讲义，用这套讲义给北京大学数学系和概率统计系 94 级学生讲授高等代数课，取得了很好的教学效果. 接着又给这两个系的 95 级学生讲授此课，进一步修改这套讲义，现分上、下两册出版.

这套教材从我国的实际情况出发，面向 21 世纪，尝试对高等代数的教学内容进行一些改革，主要有以下几方面：

努力使教材现代化. 21 世纪的人才需要掌握现代数学的思想和方法. 为此，本书注意渗透现代数学的一些基本思想和观点. 例如，用等价关系把集合划分的思想；从代数结构着眼处理问题的思想；同构分类的思想；态射（保持运算的映射）的观点等. 用现代的观点组织和讲授传统的教学内容. 例如，通过讨论子空间的结构证明线性方程组有解判别定理；按照矩阵的相抵关系、相似关系、合同关系分别讨论矩阵的相抵分类、相似分类和合同分类，并且寻求每一种关系下的完全不变量；运用线性空间的同构分类思想证明域 F 上任一 n 维线性空间 V 与它的对偶空间 V^* 同构，以及 V 与它的双重对偶空间 V^{**} 同构；在讲一元多项式的概念时，用态射的观点阐述一元多项式环的通用性质；用环同构的观点讨论数域 K 上的多项式与多项式函数之间的关系；等等. 本书还注意渗透现代数学的一些基本概念. 例如，纤维、线性流形、商集等概念；结合高等代数的具体对象水到渠成地先后引进了抽象代数的一些基本概念：在一元多项式的概念之后引进环的概念；在讲完多项式环之后引进有理函数域、任意域和有限域的概念，以及域的特征的概念；在讲了线性变换的运算后引进域上的代数的概念；在最后一章

当学生已经熟悉了正交变换、酉变换和辛变换的性质后，引进群和子群的概念.

力图在教材中体现代数与几何、分析的联系. 21 世纪的数学，分析、代数、几何将会更加相互渗透和有机结合. 因此要使学生从大学一年级开始就逐步培养把代数与几何、分析联系起来的能力. 书中注意从几何直观或分析背景引出高等代数讨论的问题；在讲述高等代数的概念时列举几何或分析的例子；把高等代数的结论应用于解决几何或分析的问题. 例如，介绍了行列式的几何意义；从几何空间的结构引出向量空间的基的概念；运用线性方程组的理论解决一些几何问题；从平面旋转的合成引出矩阵乘法的定义；从二次曲面方程的化简引出实对称矩阵的对角化以及实二次型通过正交替换化成标准形的问题，并且运用所得到的代数结论解决二次曲面方程的化简问题；从函数极值问题引出正定(负定)二次型的概念，并且运用正定(负定)矩阵解决多元函数的极值问题；在讲线性相关性时，讲述了 n 个 $n-1$ 次可微函数线性无关的充分条件；从几何空间中的例子引出商空间的概念；等等. 为了将线性代数的理论应用到分析上，为泛函分析打下基础，本书讨论的线性空间可以是无限维的，尽量不加有限维的限制.

注重联系实际，加强应用. 面向 21 世纪，数学系不仅要培养从事数学科研和教学的人材，而且要培养在其他领域工作的人才. 因此要努力培养学生运用数学理论解决实际问题和其他领域中的问题的能力. 本书在讲完线性方程组的理论后，用它解决平板受热问题；讲了矩阵可对角化的条件后，用它解决色盲遗传问题；在讲了矩阵的运算之后，解决区组设计、图论、数论中的一些问题；在讲了一元多项式环的通用性质后，用它证明组合数的一些公式；等等. 考虑到矩阵在实际问题和许多领域中有广泛应用，本书不仅把矩阵贯穿始终，而且把矩阵的运算，矩阵的相抵分类、相似分类、合同分类集中在一起讲授，并且加强了矩阵的分块，矩阵的“打洞”以及巧用特殊矩阵的训练；讲述了 Binet-Cauchy 公式及其应用. 考虑

到有限域上的线性空间在计算机以及通讯编码中有重要应用，书中讨论的线性空间是任意域上的，不局限于数域.考虑到现代物理以及一些数学分支中的需要，加强了酉空间的内容，介绍了作为爱因斯坦相对论基础的 Minkowski 空间，并且讨论了一般的正交空间，以及辛空间.

提高数学素质，加强能力培养. 21 世纪所需要的人才应当有较高的数学素质和较强的分析问题、解决问题的能力.数学素质包括提出数学问题、理解力、逻辑思维、抽象思维、创造性等几个方面.为了从大学一年级开始就着力培养学生的数学素质，本书在每一单元的开头都要提出问题，然后阐述解决问题的想法，经过抽象思维和逻辑思维一步一步地去解决这些问题.书中特别注意讲清楚想法(idea).例如，在研究线性方程组有无解的判定时为什么会想到去研究 n 元数组的向量空间的结构？在讨论有理系数多项式的因式分解时怎么会想到引进本原多项式的概念？在研究不可以对角化的线性变换的结构时如何想到最小多项式的因式分解并且讨论相应线性变换的多项式核之间的关系？复线性空间上内积的定义为什么与实线性空间上内积的定义不同？在任意域上的线性空间中如何引进度量？为什么只有两种度量：对称双线性函数或者斜对称双线性函数，而不用一般的线性函数？等等，书中都给予了清晰的回答.为了培养学生的阅读、理解能力和扩大知识面，书中有较多的例题，并且有密切配合正文的加“*”号的章节内容和一些阅读材料.有些例题不必在课堂上讲授，留给学生自己看.加“*”号的内容和阅读材料不作为教学要求，供有兴趣的学生自学.本书配备了相当丰富的习题(每一节后面配有习题，每一章后面还有补充题)，有的是为了使学生理解正文的概念，掌握正文中的定理和方法，学会重要的解题方法和技巧；有的是正文内容的补充和拓宽；有的是应用.通过做题可以培养学生分析问题和解决问题的能力.习题中加“*”号的题以及补充题不作为教学要求，供学生选做.

本书内容的安排力求符合人们认识事物的客观规律.学生从中学进入大学,有一个学习方法的适应过程.为了帮助学生树立信心适应大学的学习,我们把高等代数研究的具体对象放在前半部分,而把抽象对象放在后半部分.全书分为四部分.第一部分是线性方程组;n 元数组的向量空间和矩阵的理论,包括线性方程组的解法、行列式、n 元数组的向量空间、线性方程组的理论、矩阵的运算、矩阵的相抵分类与相似分类、二次型与矩阵的合同分类.第二部分是一元多项式环与多元多项式环的理论.第三部分是线性空间和线性映射的理论,包括任意域上的线性空间、线性映射和线性变换、线性变换的 Jordan 标准形、线性函数和双线性函数.第四部分是具有度量的线性空间的理论,包括欧几里得空间、酉空间、正交空间、辛空间.

本书可作为大学数学系、概率统计系、应用数学系的高等代数教材,上、下册共讲授两个学期;还可作为大专院校有关教师和学生的参考书.

作者衷心感谢刘旭峰博士,他通读了全书,并提出了一些宝贵的建议.

特别要感谢本书的责任编辑胡乃冏,他为本书的编辑出版付出了辛勤劳动.

书中可能会有考虑不周和疏漏之处,热诚欢迎同行和读者批评指正.

丘维声

1996 年 *2* 月于北京大学燕北园

责任编辑 胡乃罔
封面设计 王 睢
责任绘图 朱 静
版式设计 胡乃罔
责任校对 胡乃罔
责任印制 孔 源

目　　录

第一章　线性方程组的解法

读者在中学数学里已经学过二元一次方程组和三元一次方程组. 由若干个一次方程组成的方程组称为**线性方程组**,它的一般形式是

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = b_s \end{cases} \tag{1}$$

其中 $x_1, x_2, \cdots, x_n$ 代表 n 个未知量,s 是方程的个数,a_{ij}是第 i 个方程中 x_j 的系数,b_i 是第 i 个方程的常数项,$i=1,2,\cdots,s$;　$j=1,2,\cdots,n$. s 可以等于 n,也可以大于 n 或者小于 n.

在数学的各个分支,在自然科学、工程技术以及生产实际中都经常遇到线性方程组,需要求出它的解,或者判断它有没有解,当有解时需要了解它的解的结构.

代数学的一个特点是:不满足于研究个别的具体问题,而是要研究一般的问题. 对于线性方程组来说,就是要研究一般的线性方程组(方程个数 s 和未知量个数 n 都是任意的自然数,并且系数 a_{ij} 和常数项 b_i,　$i=1,2,\cdots,s$;　$j=1,2,\cdots,n$,都是任意的数)怎么解?解的情况如何?本章就来讨论这些问题.

§1　高斯(Gauss)消去法

线性方程组(1)的一个**解**是指由 n 个数 $c_1, c_2, \cdots, c_n$ 组成的有

序数组$(c_1,c_2,\cdots,c_n)$，当 $x_1,x_2,\cdots,x_n$ 分别用 $c_1,c_2,\cdots,c_n$ 代入后，(1)中每一个方程都变成恒等式.

如何解一般的线性方程组？从中学数学里知道解二元(或三元)一次方程组的基本方法是**消元法**，即把方程组中一部分方程变成未知量较少的方程，直至得到一个一元一次方程，进而可求出方程组的解. 这种方法也适用于解一般的线性方程组. 在解未知量较多的线性方程组时，需要使这种解法有规律可遵循，尽可能简便. 下面通过一个具体例子来说明消元法的做法.

例 1 解线性方程组

$$\begin{cases} x_1+3x_2+x_3+2x_4=4 \\ 3x_1+4x_2+2x_3-3x_4=6 \\ -x_1-5x_2+4x_3+x_4=11 \\ 2x_1+7x_2+x_3-6x_4=-5 \end{cases} \tag{2}$$

解 把第一个方程的(-3)倍、1 倍、(-2)倍分别加到第二、三、四个方程上，使得在第二、三、四个方程中消去未知量 x_1：

$$\begin{array}{l} \\ ②+①\cdot(-3) \\ ③+① \\ ④+①\cdot(-2) \end{array}\begin{cases} x_1+3x_2+x_3+2x_4=4 \\ -5x_2-x_3-9x_4=-6 \\ -2x_2+5x_3+3x_4=15 \\ x_2-x_3-10x_4=-13 \end{cases}$$

把第二、四两个方程互换位置(目的是在下一步避免分数运算)：

$$(②,④)\quad\begin{cases} x_1+3x_2+x_3+2x_4=4 \\ x_2-x_3-10x_4=-13 \\ -2x_2+5x_3+3x_4=15 \\ -5x_2-x_3-9x_4=-6 \end{cases}$$

这里符号(②,④)表示二、四两个方程互换位置.

同理

$$
\begin{array}{r}
\\
\\
③+②\cdot 2\\
④+②\cdot 5
\end{array}
\left\{\begin{aligned}
x_1+3x_2+x_3+2x_4&=4\\
x_2-x_3-10x_4&=-13\\
3x_3-17x_4&=-11\\
-6x_3-59x_4&=-71
\end{aligned}\right.
$$

$$
\begin{array}{r}
\\
\\
\\
④+③\cdot 2
\end{array}
\left\{\begin{aligned}
x_1+3x_2+x_3+2x_4&=4\\
x_2-x_3-10x_4&=-13\\
3x_3-17x_4&=-11\\
-93x_4&=-93
\end{aligned}\right.
\tag{3}
$$

方程组(3)的最后一个方程是一元一次方程,解得 $x_4=1$. 再逐次代入(3)的第三、第二、第一个方程,可求得 $x_3=2$, $x_2=-1$, $x_1=3$. 于是$(3,-1,2,1)$是原方程组(2)的解.

像(3)这样形状的方程组称为**阶梯形方程组**.

从上述解题过程看出,用消元法解线性方程组的具体做法是:对方程组反复施行以下三种变换:

1° 把一个方程的倍数加到另一个方程上;

2° 互换两个方程的位置;

3° 用一个非零数乘某一个方程.

这三种变换称为线性方程组的**初等变换**. 经过初等变换,把原方程组变成阶梯形方程组,然后去解阶梯形方程组(从最后一个方程开始,逐次往上解),求得的解就是原方程组的解.

从例 1 的解题过程中看到,在对方程组作初等变换时,只是对方程组的系数和常数项进行运算,而未知量 $x_1,x_2,\cdots,x_n$ 并没有参加运算,因此在用消元法解线性方程组时,可以只写出方程组的全部系数和常数项. 即线性方程组(1)可以用下面的一张表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} & b_s \end{pmatrix} \tag{4}$$

来表示.

定义 1 由 $s\times m$ 个数 $a_{ij}(i=1,2,\cdots,s;\quad j=1,2,\cdots,m)$ 排成的 s 行、m 列的一张表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sm} \end{pmatrix} \tag{5}$$

称为一个 $s\times m$ **矩阵**. 这张表里的任何一个数称为这个矩阵的一个元素.

矩阵通常用一个大写拉丁字母表示,譬如,矩阵(5) 可记作 A,为了指明它是 s 行、m 列的矩阵,可以记作 $A_{s\times m}$. 若矩阵 A 的行数 s 与列数 m 相等,则称 A 为 m **级方阵**或者 m **级矩阵**. 元素全为零的矩阵称为**零矩阵**,就记作 0,为了指明它是 s 行、m 列的零矩阵,可记作 $0_{s\times m}$. 矩阵 A 的位于第 i 行与第 j 列交叉位置的元素称为 A 的 (i,j) 元,可记作 $A(i;j)$. i 称为 $A(i;j)$ 的行指标(或行标),j 称为 $A(i;j)$ 的列指标(或列标). 例如,若用 A 表示矩阵(5),则 $A(2;m)=a_{2m}$,$A(s;1)=a_{s1}$,等等. 如果矩阵 A 的 (i,j) 元为 $a_{ij}(i=1,\cdots,s;\quad j=1,\cdots,m)$,则可以写 $A=(a_{ij})$.

线性方程组(1)的全部系数和常数项按原来顺序组成的矩阵(4)称为方程组(1)的**增广矩阵**. 而只由全部系数组成的矩阵

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} \end{pmatrix} \tag{6}$$

称为方程组(1)的**系数矩阵**. 线性方程组(1)可以用它的增广矩阵

(4) 来表示.

线性方程组(2)用它的增广矩阵来表示,则例 1 的解题过程可以写为:

$$\begin{pmatrix} 1 & 3 & 1 & 2 & 4 \\ 3 & 4 & 2 & -3 & 6 \\ -1 & -5 & 4 & 1 & 11 \\ 2 & 7 & 1 & -6 & -5 \end{pmatrix}$$

$$\xrightarrow[\text{④}+\text{①}\cdot(-2)]{\substack{\text{②}+\text{①}\cdot(-3)\\ \text{③}+\text{①}}} \begin{pmatrix} 1 & 3 & 1 & 2 & 4 \\ 0 & -5 & -1 & -9 & -6 \\ 0 & -2 & 5 & 3 & 15 \\ 0 & 1 & -1 & -10 & -13 \end{pmatrix}$$

$$\xrightarrow{(\text{②},\ \text{④})} \begin{pmatrix} 1 & 3 & 1 & 2 & 4 \\ 0 & 1 & -1 & -10 & -13 \\ 0 & -2 & 5 & 3 & 15 \\ 0 & -5 & -1 & -9 & -6 \end{pmatrix}$$

$$\xrightarrow{\substack{\text{③}+\text{②}\cdot 2\\ \text{④}+\text{②}\cdot 5}} \begin{pmatrix} 1 & 3 & 1 & 2 & 4 \\ 0 & 1 & -1 & -10 & -13 \\ 0 & 0 & 3 & -17 & -11 \\ 0 & 0 & -6 & -59 & -71 \end{pmatrix}$$

$$\xrightarrow{\text{④}+\text{③}\cdot 2} \begin{pmatrix} 1 & 3 & 1 & 2 & 4 \\ 0 & 1 & -1 & -10 & -13 \\ 0 & 0 & 3 & -17 & -11 \\ 0 & 0 & 0 & -93 & -93 \end{pmatrix} \tag{7}$$

最后这个矩阵(7)称为行阶梯形矩阵,它对应的线性方程组是阶梯形方程组(3).解阶梯形方程组(3)得:

$$x_4 = 1, \quad x_3 = 2, \quad x_2 = -1, \quad x_1 = 3$$

从而(3,－1,2,1)是原方程组的解.

现在我们给出行阶梯形矩阵一个精确的定义.

定义 2 一个 $s \times m$ 矩阵 J 称为**行阶梯形矩阵**(或阶梯形矩阵),如果

1) J 的零行(即,元素全为 0 的行,如果有的话)在下方;

2) J 的每个非零行的第一个不为零的元素称为 J 的**主元**. 主元的列指标随着行指标的递增而严格增大. 即设 J 有 r 个非零行,第 i 行的主元的列指标记作 j_i,则有

$$1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant m$$

例如,

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

是一个行阶梯形矩阵,其中 $j_1 = 1$, $j_2 = 2$, $j_3 = 4$.

从上面所说的看出,利用线性方程组(2)的增广矩阵求(2)的解时,对矩阵反复施行了以下三种变换:

1° 把一行的倍数加到另一行上;

2° 互换两行的位置;

3° 用一个非零数乘某一行.

这三种变换称为**矩阵的初等行变换**.

运用矩阵的初等行变换把线性方程组(2)的增广矩阵化成行阶梯形矩阵(7)后,可以写出(7)表示的阶梯形方程组,进而求解;也可以对矩阵(7)继续作初等行变换,化成一种特殊的行阶梯形矩阵,以至于能直接从这种矩阵"读出"原方程组的解:

$$\begin{pmatrix} 1 & 3 & 1 & 2 & 4 \\ 0 & 1 & -1 & -10 & -13 \\ 0 & 0 & 3 & -17 & -11 \\ 0 & 0 & 0 & -93 & -93 \end{pmatrix}$$

$$\xrightarrow{④\cdot\left(-\frac{1}{93}\right)}\begin{pmatrix}1 & 3 & 1 & 2 & 4\\ \text{[illegible]} & & & & \\ 0 & 0 & 3 & -17 & -11\\ 0 & 0 & 0 & 1 & 1\end{pmatrix}$$

$$\xrightarrow[③+④\cdot 17]{\substack{①+④\cdot(-2)\\ ②+④\cdot 10}}\begin{pmatrix}1 & 3 & 1 & 0 & 2\\ 0 & 1 & -1 & 0 & -3\\ 0 & 0 & 3 & 0 & 6\\ 0 & 0 & 0 & 1 & 1\end{pmatrix}$$

$$\xrightarrow{③\cdot\frac{1}{3}}\begin{pmatrix}1 & 3 & 1 & 0 & 2\\ 0 & 1 & -1 & 0 & -3\\ 0 & 0 & 1 & 0 & 2\\ 0 & 0 & 0 & 1 & 1\end{pmatrix}$$

$$\xrightarrow{\substack{①+③\cdot(-1)\\ ②+③}}\begin{pmatrix}1 & 3 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & -1\\ 0 & 0 & 1 & 0 & 2\\ 0 & 0 & 0 & 1 & 1\end{pmatrix}$$

$$\xrightarrow{①+②\cdot(-3)}\begin{pmatrix}1 & 0 & 0 & 0 & 3\\ 0 & 1 & 0 & 0 & -1\\ 0 & 0 & 1 & 0 & 2\\ 0 & 0 & 0 & 1 & 1\end{pmatrix}$$

最后这个行阶梯形矩阵称为简化行阶梯形矩阵，它表示的线性方程组是：

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

从而立即得到原方程组的解是$(3, -1, 2, 1)$.

定义 3　一个行阶梯形矩阵 J 称为**简化行阶梯形矩阵**，如果 J 的主元都是 1，并且每个主元所在的列的其余元素都是零.

例如，

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

是简化行阶梯形矩阵.

综上所述，用消元法解线性方程组的具体做法是：第一步，写出线性方程组的增广矩阵，运用矩阵的初等行变换把它化成行阶梯形矩阵；第二步，写出行阶梯形矩阵表示的阶梯形方程组，对它自下而上求解；或者运用初等行变换把行阶梯形矩阵进一步化成简化行阶梯形矩阵，从而可立即写出解，这种解线性方程组的方法称为**高斯**(Gauss)**消去法**.

高斯消去法的理论根据是什么？即，为什么阶梯形方程组的解就是原方程组的解？任意一个矩阵是否一定能够通过矩阵的初等行变换化成行阶梯形矩阵？现在来讨论这两个问题.

定义 4　如果线性方程组 I 的每一个解都是线性方程组 II 的解，并且 II 的每一个解都是 I 的解，那么称 I 与 II **同解**.

同解是线性方程组之间的一种关系，它满足：

1)　**反身性**，即任何一个线性方程组与自身同解；

2)　**对称性**，即，若 I 与 II 同解，则 II 与 I 同解；

3)　**传递性**，即，若 I 与 II 同解，II 与 III 同解，则 I 与 III 同解.

命题 1.1.1　初等变换把线性方程组变成与它同解的方程组.

证明 设线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1 \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{i1}x_1+\cdots+a_{in}x_n=b_i \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{k1}x_1+\cdots+a_{kn}x_n=b_k \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{s1}x_1+\cdots+a_{sn}x_n=b_s \end{cases} \tag{8}$$

作初等变换:ⓚ+ ⓘ $\cdot l$,得

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1 \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{i1}x_1+\cdots+a_{in}x_n=b_i \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ (a_{k1}+la_{i1})x_1+\cdots+(a_{kn}+la_{in})x_n=b_k+lb_i \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{s1}x_1+\cdots+a_{sn}x_n=b_s \end{cases} \tag{9}$$

设$(c_1,\cdots,c_n)$是方程组(8)的一个解,则有一组恒等式:

$$\begin{cases} a_{11}c_1+\cdots+a_{1n}c_n=b_1 \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{i1}c_1+\cdots+a_{in}c_n=b_i \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{k1}c_1+\cdots+a_{kn}c_n=b_k \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{s1}c_1+\cdots+a_{sn}c_n=b_s \end{cases} \tag{10}$$

把(10)中第 i 个恒等式的 l 倍加到第 k 个恒等式上,得到又一组恒等式:

$$\begin{cases} a_{11}c_1+\cdots+ & a_{1n}c_n= & b_1 \\ \cdots\ \cdots\ \cdots & \cdots\ \cdots & \cdots \\ a_{i1}c_1+\cdots+ & a_{in}c_n= & b_i \\ \cdots\ \cdots\ \cdots & \cdots\ \cdots & \cdots \\ (a_{k1}+la_{i1})c_1+\cdots+ & (a_{kn}+la_{in})c_n= & b_k+lb_i \\ \cdots\ \cdots\ \cdots & \cdots\ \cdots & \cdots \\ a_{s1}c_1+\cdots+ & a_{sn}c_n= & b_s \end{cases} \tag{11}$$

(11) 表明 $(c_1,\cdots,c_n)$ 是方程组(9) 的一个解. 这证明了经过 1° 型初等变换,原方程组的每个解是新方程组的解.

方程组(9) 经过初等变换 ⓚ+ ⓘ·$(-l)$ 得到方程组(8),由刚才证明的结论知道,方程组(9) 的每个解是方程组(8) 的解. 因此方程组(8) 与(9) 同解.

对于 2° 型和 3° 型初等变换,作为练习留给读者. ▎

由命题 1.1.1 和同解的传递性得到,原方程组与阶梯形方程组同解.

命题 1.1.2　任意一个矩阵都可以经过一系列初等行变换化成行阶梯形矩阵.

证明　零矩阵按定义是行阶梯形矩阵. 下面考虑非零矩阵. 对非零矩阵的行数 s 作数学归纳法.

$s=1$ 时,矩阵只有一行,这是行阶梯形矩阵.

假设 $s-1$ 行的矩阵都能经过初等行变换化成行阶梯形. 下面看 s 行的矩阵 $A=(a_{ij})$. 如果 A 的第 1 列元素不全为 0,则互换两行位置可以使矩阵的左上角元素,即(1,1) 元不为零. 因此不妨就设 A 的(1,1) 元 $a_{11}\neq 0$. 把 A 的第 1 行的 $-\dfrac{a_{21}}{a_{11}}$ 倍加到第 2 行,把 A 的第 1 行的 $-\dfrac{a_{31}}{a_{11}}$ 倍加到第 3 行,…,把 A 的第 1 行的 $-\dfrac{a_{s1}}{a_{11}}$

倍加到第 s 行，A 变成下述矩阵 B

$$B=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}-\dfrac{a_{21}}{a_{11}}a_{12} & \cdots & a_{2n}-\dfrac{a_{21}}{a_{11}}a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a_{s2}-\dfrac{a_{s1}}{a_{11}}a_{12} & \cdots & a_{sn}-\dfrac{a_{s1}}{a_{11}}a_{1n} \end{pmatrix}$$

我们把矩阵 B 的右下方的 $(s-1)\times(n-1)$ 矩阵记作 B_1.

如果 A 的第 1 列元素全为零，则考虑第 2 列. 若 A 的第 2 列元素不全为零，不妨设 $a_{12}\neq 0$，同理，可以经过初等行变换把 A 变成下述矩阵 C

$$C=\begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23}-\dfrac{a_{22}}{a_{12}}a_{13} & \cdots & a_{2n}-\dfrac{a_{22}}{a_{12}}a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & a_{s3}-\dfrac{a_{s2}}{a_{12}}a_{13} & \cdots & a_{sn}-\dfrac{a_{s2}}{a_{12}}a_{1n} \end{pmatrix}$$

把矩阵 C 的右下方 $(s-1)\times(n-2)$ 矩阵记作 C_1.

如果 A 的第 1,2 列元素全都为零，则考虑 A 的第 3 列，依次类推.

由于 $B_1,C_1,\cdots$ 都是 $s-1$ 行矩阵，根据归纳假设，它们可以经过初等行变换分别化成行阶梯形 $J_1,J_2,\cdots$. 因此 A 可以经过初等行变换化成下述形式的矩阵之一：

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & J_1 & \\ 0 & & & \end{pmatrix},\qquad \begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & & & \\ \vdots & \vdots & & J_2 & \\ 0 & 0 & & & \end{pmatrix},\qquad \cdots$$

这些都是行阶梯形矩阵.

根据数学归纳法原理,对于任意自然数 s, s 行非零矩阵都可以经过初等行变换化成行阶梯形. ▌

推论 1.1.1 任意一个矩阵都可以经过一系列初等行变换化成简化行阶梯形矩阵.

证明 由命题 1.1.2 知,任一矩阵 A 可以经过初等行变换化成行阶梯形矩阵 J . 运用 3° 型初等行变换可以把 J 的每个主元变成 1,再从倒数第一个非零行起,用 1° 型初等行变换可以把每个主元所在的列的其余元素变成零,从而得到简化行阶梯形矩阵. ▌

习 题 1.1

解下列线性方程组:

(1) $$\begin{cases} x_1-3x_2-2x_3-x_4=6 \\ 3x_1-8x_2+x_3+5x_4=0 \\ -2x_1+x_2-4x_3+x_4=-12 \\ -x_1+4x_2-x_3-3x_4=2 \end{cases}$$

(2) $$\begin{cases} x_1+3x_2-7x_3=-8 \\ 2x_1+5x_2+4x_3=4 \\ -3x_1-7x_2-2x_3=-3 \\ x_1+4x_2-12x_3=-15 \end{cases}$$

(3) $$\begin{cases} x_1+3x_2-2x_3+2x_4-x_5=4 \\ 5x_1+8x_2+4x_3+3x_4-2x_5=-4 \\ -2x_1-5x_2+x_3-5x_4+4x_5=-12 \\ 3x_1+7x_2-x_3+x_4-3x_5=-4 \\ -x_1-4x_2+5x_3-x_4+5x_5=-6 \end{cases}$$

(4) $\begin{cases} x_1-2x_2+3x_3-4x_4= 4 \\ x_1+ x_2- x_3+ x_4=-11 \\ x_1+3x_2 \qquad + x_4= 1 \\ \quad -7x_2+3x_3+ x_4=- 3 \end{cases}$

§2　线性方程组的解的情况

这一节我们讨论线性方程组的解的可能情形.读者在中学平面解析几何中已经知道,在平面上取定一个直角坐标系后,一个二元一次方程表示平面上的一条直线. 因此两个方程的二元一次方程组

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases} \tag{1}$$

的解就是它们表示的两条直线的交点的坐标.由于平面上的两条直线的位置关系只有三种可能:相交,平行,重合,因此二元一次方程组(1)的解的情况只有三种可能:有唯一解,无解,有无穷多个解.

一般线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = b_s \end{cases} \tag{2}$$

的解的情况如何?由于方程组(2)与对它进行初等变换后得到的阶

梯形方程组同解，所以我们只要看阶梯形方程组的解有几种可能性.

情形 1. 阶梯形方程组中出现“$0=d$”，其中 d 是非零数. 这个方程无解，从而阶梯形方程组无解.

情形 2. 阶梯形方程组中不出现形如“$0=d$”这样的方程，其中 d 是非零数. 设阶梯形方程组中不包括“$0=0$”的方程的个数为 r.

情形 2.1 $r=n$. 这时阶梯形方程组(不包括“$0=0$”)为

$$\left\{\begin{aligned} c_{11}x_1+c_{12}x_2+\cdots+c_{1n}x_n&=d_1\\ c_{22}x_2+\cdots+c_{2n}x_n&=d_2\\ \cdots\quad\cdots\quad\cdots\quad&\cdots\\ c_{nn}x_n&=d_n \end{aligned}\right. \tag{3}$$

其中 $c_{11},c_{22},\cdots,c_{nn}$ 全不为零. 从最后一个方程可求出 x_n 的值，代入倒数第二个方程可求出 x_{n-1} 的值，依次往上代入，可求出 $x_{n-2},\cdots,x_2,x_1$ 的值. 从而(3) 有唯一解.

情形 2.2 $r<n$. 这时阶梯形方程组(不包括“$0=0$”)形如

$$\left\{\begin{aligned} c_{11}x_1+\cdots\quad\cdots\quad\cdots\quad\cdots+c_{1n}x_n&=d_1\\ c_{2j_2}x_{j_2}+\cdots\quad\cdots\quad\cdots+c_{2n}x_n&=d_2\\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad&\cdots\\ c_{rj_r}x_{j_r}+\cdots+c_{rn}x_n&=d_r \end{aligned}\right. \tag{4}$$

其中 $c_{11},c_{2j_2},\cdots,c_{rj_r}$ 全不为零，它们是(4) 的增广矩阵的主元. 以主元 $c_{11},c_{2j_2},\cdots,c_{rj_r}$ 为系数的未知量 $x_1,x_{j_2},\cdots,x_{j_r}$ 称为阶梯形方程组(4) 的**主变量**，也称为原方程组(2) 的一组主变量；其余的未知量称为(4) 的**自由未知量**，也称为原方程组(2) 的一组自由未知

量. 在(4) 中任给自由未知量一组值,则最后一个方程是 x_{j_r} 的一元一次方程,可唯一解出 x_{j_r};依次往上代入,可求出 $x_{j_{r-1}},\cdots,x_{j_2},x_1$ 的值. 这说明:任给自由未知量一组值,可唯一求出主变量 $x_1,x_{j_2},\cdots,x_{j_r}$ 的值,从而得到(4) 的一个解. 由于自由未知量可以取无穷多组值,因此(4) 有无穷多个解. 这无穷多个解可以通过用自由未知量表达主变量的下述表达式来表示:

$$\begin{cases} x_1 = c'_{1k}x_k + \cdots + c'_{1l}x_l + d'_1 \\ x_{j_2} = c'_{2k}x_k + \cdots + c'_{2l}x_l + d'_2 \\ \cdots\ \cdots\ \ \cdots\ \ \ \cdots\ \ \ \cdots\ \ \ \cdots\ \ \ \cdots \\ x_{j_r} = c'_{rk}x_k + \cdots + c'_{rl}x_l + d'_r \end{cases} \tag{5}$$

其中 $x_k,\cdots,x_l$ 是自由未知量. 表达式(5) 称为阶梯形方程组(4) 的一般解,也称为原方程组(2) 的**一般解**.

情形 2.3　$r > n$. 这时阶梯形方程组的增广矩阵有 r 个非零行,从而有 r 个主元. 由于阶梯形方程组的增广矩阵是行阶梯形矩阵,因此这 r 个主元应当位于不同的列,且不能位于最后一列(否则,会出现"$0 = d$"这样的方程,其中 $d \neq 0$). 但是,行阶梯形矩阵除去最后一列后,只有 n 列,因此 $r > n$ 是不可能的.

综上述,我们得到下面的结论:

定理 1.2.1　n 元线性方程组的解的情况只有三种可能:无解,有唯一解,有无穷多个解. 把 n 元线性方程组经过初等变换化成阶梯形方程组,如果出现"$0 = d$"(其中 d 是非零数) 这样的方程,那么原方程组无解;否则,原方程组有解. 有解时,设阶梯形方程组中不包括"$0 = 0$" 的方程的个数是 r,如果 $r = n$,则原方程组有唯一解;如果 $r < n$,则原方程组有无穷多个解,这无穷多个解可用一般解表达出,其中自由未知量有 $n - r$ 个,主变量有 r 个. ▌

例 1　解线性方程组

$$\begin{cases} x_1 - x_2 - 3x_3 + x_4 = 1 \\ x_1 - x_2 + 2x_3 - x_4 = 3 \\ 4x_1 - 4x_2 + 3x_3 - 2x_4 = 6 \\ 2x_1 - 2x_2 - 11x_3 + 4x_4 = 0 \end{cases} \tag{6}$$

解 把方程组(6)的增广矩阵经过初等行变换化成行阶梯形

$$\begin{pmatrix} 1 & -1 & -3 & 1 & 1 \\ 1 & -1 & 2 & -1 & 3 \\ 4 & -4 & 3 & -2 & 6 \\ 2 & -2 & -11 & 4 & 0 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & -1 & -3 & 1 & 1 \\ 0 & 0 & 5 & -2 & 2 \\ 0 & 0 & 15 & -6 & 2 \\ 0 & 0 & -5 & 2 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & -3 & 1 & 1 \\ 0 & 0 & 5 & -2 & 2 \\ 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

这个行阶梯形矩阵表示的阶梯形方程组出现方程“0 =－ 4”，因此，方程组(6) 无解.

例 2 解线性方程组

$$\begin{cases} x_1 - x_2 - 3x_3 + x_4 = 1 \\ x_1 - x_2 + 2x_3 - x_4 = 3 \\ 4x_1 - 4x_2 + 3x_3 - 2x_4 = 10 \\ 2x_1 - 2x_2 - 11x_3 + 4x_4 = 0 \end{cases} \tag{7}$$

解

$$\begin{pmatrix} 1 & -1 & -3 & 1 & 1 \\ 1 & -1 & 2 & -1 & 3 \\ 4 & -4 & 3 & -2 & 10 \\ 2 & -2 & -11 & 4 & 0 \end{pmatrix}$$

$$\rightarrow\begin{pmatrix}1 & -1 & -3 & 1 & 1\\0 & 0 & 5 & -2 & 2\\0 & 0 & 15 & -6 & 6\\0 & 0 & -5 & 2 & -2\end{pmatrix}\rightarrow\begin{pmatrix}1 & -1 & -3 & 1 & 1\\0 & 0 & 5 & -2 & 2\\0 & 0 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0\end{pmatrix}$$

$$\rightarrow\begin{pmatrix}1 & -1 & -3 & 1 & 1\\0 & 0 & 1 & -\frac{2}{5} & \frac{2}{5}\\0 & 0 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0\end{pmatrix}\rightarrow\begin{pmatrix}1 & -1 & 0 & -\frac{1}{5} & \frac{11}{5}\\0 & 0 & 1 & -\frac{2}{5} & \frac{2}{5}\\0 & 0 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0\end{pmatrix}$$

从最后一个简化行阶梯形矩阵可以看出原方程组(7)有无穷多个解,并且可以立即写出方程组(7)的一般解:

$$\begin{cases}x_1 = x_2 + \frac{1}{5}x_4 + \frac{11}{5}\\x_3 = \qquad \frac{2}{5}x_4 + \frac{2}{5}\end{cases}$$

其中 x_2, x_4 是自由未知量.

例 3　a 为何值时,下述线性方程组有解?当有解时,求它的一般解.

$$\begin{cases}3x_1 + x_2 - x_3 - 2x_4 = 2\\x_1 - 5x_2 + 2x_3 + x_4 = -1\\2x_1 + 6x_2 - 3x_3 - 3x_4 = a + 1\\-x_1 - 11x_2 + 5x_3 + 4x_4 = -4\end{cases}\tag{8}$$

解

$$\begin{pmatrix}3 & 1 & -1 & -2 & 2\\1 & -5 & 2 & 1 & -1\\2 & 6 & -3 & -3 & a+1\\-1 & -11 & 5 & 4 & -4\end{pmatrix}\rightarrow\begin{pmatrix}1 & -5 & 2 & 1 & -1\\3 & 1 & -1 & -2 & 2\\2 & 6 & -3 & -3 & a+1\\-1 & -11 & 5 & 4 & -4\end{pmatrix}$$

$$\to\begin{pmatrix}1&-5&2&1&-1\\0&16&-7&-5&5\\0&16&-7&-5&a+3\\0&-16&7&5&-5\end{pmatrix}\to\begin{pmatrix}1&-5&2&1&-1\\0&16&-7&-5&5\\0&0&0&0&a-2\\0&0&0&0&0\end{pmatrix}$$

当 $a-2\neq 0$ 时，方程组(8) 无解. 当 $a-2=0$，即 $a=2$ 时，方程组(8) 有解，并且有无穷多个解. 因为

$$\begin{pmatrix}1&-5&2&1&-1\\0&16&-7&-5&5\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&-5&2&1&-1\\0&1&-\frac{7}{16}&-\frac{5}{16}&\frac{5}{16}\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&0&-\frac{3}{16}&-\frac{9}{16}&\frac{9}{16}\\0&1&-\frac{7}{16}&-\frac{5}{16}&\frac{5}{16}\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}$$

所以 $a=2$ 时，方程组(8) 的一般解为

$$\begin{cases}x_1=\frac{3}{16}x_3+\frac{9}{16}x_4+\frac{9}{16}\\x_2=\frac{7}{16}x_3+\frac{5}{16}x_4+\frac{5}{16}\end{cases}\tag{9}$$

其中 x_3,x_4 是自由未知量.

定义 1 常数项全为零的线性方程组称为**齐次线性方程组**，它的一般形式是

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = 0 \end{cases} \tag{10}$$

齐次线性方程组(10)一定有解,因为$(0,0,\cdots,0)$就是方程组(10)的一个解,这个解称为**零解**.其它的解(如果有的话)称为**非零解**.把定理1.2.1应用到齐次线性方程组上,我们便得到

推论 1.2.1 n 元齐次线性方程组经过初等变换化成阶梯形方程组后,如果阶梯形方程组中不包括"$0=0$"的方程的个数 $r=n$,则齐次线性方程组只有零解;如果 $r<n$,则有非零解(此时有无穷多个解),自由未知量有 $n-r$ 个,主变量有 r 个. ▌

从推论1.2.1我们又可得到

推论 1.2.2 s 个方程的 n 元齐次线性方程组(10),如果 $s<n$,则它一定有非零解.

证明 我们把方程组(10)经过初等变换化成阶梯形方程组,其中不包括"$0=0$"的方程个数 $r\leqslant s<n$,所以方程组(10)有非零解. ▌

例 4 判断下述齐次线性方程组有无非零解,如果有非零解,写出它的一般解.

$$\begin{cases} x_1 + 3x_2 - 4x_3 + 2x_4 = 0 \\ 3x_1 - x_2 + 2x_3 - x_4 = 0 \\ -2x_1 + 4x_2 - x_3 + 3x_4 = 0 \\ 3x_1 + 9x_2 - 7x_3 + 6x_4 = 0 \end{cases} \tag{11}$$

解 因为齐次线性方程组的增广矩阵的最后一列元素全为零,在对它作初等行变换时,所得到的矩阵的最后一列元素也总是全为零,因此我们只要写出齐次线性方程组的系数矩阵,对系数

矩阵施行初等行变换化成行阶梯形.

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

$$\longrightarrow \begin{pmatrix} 1 & 3 & -4 & 2 \\ 0 & -10 & 14 & -7 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{10} \\ 0 & 1 & 0 & \frac{7}{10} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

从最后的简化行阶梯形矩阵看出,方程组(11)有非零解,并且可立即写出它的一般解:

$$\begin{cases} x_1 = \frac{1}{10}x_4 \\ x_2 = -\frac{7}{10}x_4 \\ x_3 = 0 \end{cases} \tag{12}$$

其中 x_4 是自由未知量.

习 题 1.2

1. 解下列线性方程组:

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

(2) $\begin{cases} 2x_1 - 3x_2 + x_3 + 5x_4 = 6 \\ -3x_1 + x_2 + 2x_3 - 4x_4 = 5 \\ -x_1 - 2x_2 + 3x_3 + x_4 = 11 \end{cases}$

(3) $\begin{cases} 3x_1-\ 9x_2+\ 6x_3+15x_4=-3 \\ x_1-\ 3x_2+\ 2x_3+\ 5x_4=-1 \\ -\ 6x_1+18x_2-12x_3-30x_4=\ 6 \\ 4x_1-12x_2+\ 8x_3+20x_4=-4 \\ 5x_1-15x_2+10x_3+25x_4=-5 \end{cases}$

(4) $\begin{cases} 2x_1+5x_2 \qquad -\ 7x_4 \qquad =\ 2 \\ x_1+3x_2 \qquad -\ 2x_4+x_5=-1 \\ 5x_1-2x_2+\ x_3-\ x_4+x_5=\ 3 \\ x_1-4x_2+\ x_3+\ x_4-x_5=\ 3 \\ x_1+2x_2+\ x_3-\ x_4+x_5=-1 \end{cases}$

2. 下列齐次线性方程组有无非零解？若有非零解，求出它的一般解.

(1) $\begin{cases} 3x_1-\ 5x_2+\ x_3-\ 2x_4=0 \\ 2x_1+\ 3x_2-\ 5x_3+\ x_4=0 \\ -\ x_1+\ 7x_2-\ 4x_3+\ 3x_4=0 \\ 4x_1+15x_2-\ 7x_3+\ 9x_4=0 \end{cases}$

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

3. a 为何值时，下述线性方程组有解？当有解时，求出它的一般解.

$$\begin{cases} x_1+\ x_2+\ x_3+\ x_4=-7 \\ x_1 \qquad +\ 3x_3-\ 4x_4=\ 8 \\ x_1+\ 2x_2-\ x_3+\ x_4=\ 2a+2 \\ 3x_1+\ 3x_2+\ 3x_3+\ 2x_4=-11 \\ 2x_1+\ 2x_2+\ 2x_3+\ x_4=\ 2a \end{cases}$$

4. 当 c 与 d 取什么值时，线性方程组

$$
\begin{cases}
x_1+\ \ x_2+\ \ x_3+\ \ x_4+x_5=1\\
3x_1+2x_2+\ \ x_3+\ \ x_4-3x_5=c\\
2x_1+2x_2+2x_3+2x_4+2x_5=2\\
3x_1+3x_2+3x_3+3x_4+3x_5=d+1
\end{cases}
$$

有解？在有解的情形，求它的一般解.

§3 数　　域

讨论线性方程组有没有解时，需要明确所考虑的数集是什么. 例如，一元一次方程

$$2x=1$$

在有理数集内有解：$x=\frac{1}{2}$；但是在整数集内没有解.

我们在用高斯消去法解线性方程组时，用到了数的加法、减法、乘法和除法四种运算. 如果所取的数集 K 具有下述性质：$\forall\ a,b\in K$，有 $a\pm b,ab,\frac{a}{b}\in K$(其中分母 $b\neq 0$)，那么线性方程组就不会因为数集取得太小而引起无解.（注：符号"$\forall$ "表示"对于一切".）事实上，前面两节讨论线性方程组的解时，已经假定了所取的数集具有这些性质.

定义 1　复数集的子集 K 称为一个**数域**，如果它满足：

1)　$0,1\in K$；

2)　$\forall\ a,b\in K$，都有 $a\pm b,ab\in K$，并且当 $b\neq 0$ 时，有 $\frac{a}{b}\in K$.

性质 2)称为 K 对于加、减、乘、除四种运算**封闭**.

有理数集 Q、实数集 R、复数集 C 都是数域. 整数集 Z 不是数域. 除了 Q,R,C 外，还有许多数域. 例如，令

$$Q(\sqrt{2}):=\{a+b\sqrt{2}\mid a,b\in Q\}$$

我们来证明 $Q(\sqrt{2})$ 是一个数域：首先，

$$0=0+0\sqrt{2}\in Q(\sqrt{2}),\quad 1=1+0\sqrt{2}\in Q(\sqrt{2})$$

其次，设 $\alpha,\beta\in Q(\sqrt{2})$，即

$$\alpha=a+b\sqrt{2},\quad \beta=c+d\sqrt{2}$$

其中 $a,b,c,d\in Q$. 我们有

$$\begin{aligned}\alpha\pm\beta&=(a+b\sqrt{2})\pm(c+d\sqrt{2})\\&=(a\pm c)+(b\pm d)\sqrt{2}\in Q(\sqrt{2})\\ \alpha\beta&=(a+b\sqrt{2})(c+d\sqrt{2})\\&=(ac+2bd)+(ad+bc)\sqrt{2}\in Q(\sqrt{2})\end{aligned}$$

设 $\beta\neq 0$，则 c,d 不全为零. 我们断言 $c-d\sqrt{2}\neq 0$. 否则，$c=d\sqrt{2}$，于是 d 不为零，从而得 $\frac{c}{d}=\sqrt{2}$，矛盾. 因此我们有

$$\begin{aligned}\frac{\alpha}{\beta}&=\frac{a+b\sqrt{2}}{c+d\sqrt{2}}\\&=\frac{(a+b\sqrt{2})(c-d\sqrt{2})}{(c+d\sqrt{2})(c-d\sqrt{2})}\\&=\frac{ac-2bd}{c^2-2d^2}+\frac{bc-ad}{c^2-2d^2}\sqrt{2}\in Q(\sqrt{2})\end{aligned}$$

综上述，$Q(\sqrt{2})$ 是一个数域.

命题 1.3.1　任一数域都包含有理数域 Q.

证明　设 K 是一个数域，则 $0,1\in K$. 从而有 $2=1+1\in K$，$3=2+1\in K,\cdots,n=(n-1)+1\in K$，即任一自然数 $n\in K$. 由于 $-n=0-n\in K$，所以任一负整数属于 K. 从而 $Z\subset K$. 于是任一有理数 $\frac{a}{b}\in K$，其中 $b\neq 0$. 所以 $Q\subset K$. ▌

注：　当 A 是集合 B 的子集时，记作 $A\subset B$. 如果已经知道 A 是 B 的真子集，则记作 $A\subsetneq B$. 这些记号与中学数学课本采用的记号有点不同，请读者注意.

从现在起,我们取定一个数域 K. 所讨论的线性方程组都是数域 K 上的,即它的全部系数和常数项都属于 K,并且在数域 K 里求它的解. 所讨论的矩阵都是数域 K 上的矩阵,即矩阵的全部元素都属于 K. 做矩阵的初等行变换时,"倍数"、"非零数"都属于 K.

习　题　1.3

1. 令 $Q(i) := \{a + bi \mid a, b \in Q\}$,证明 $Q(i)$ 是一个数域.

2. 令数集

$$K = \left\{ \frac{a_0 + a_1 e + \cdots + a_n e^n}{b_0 + b_1 e + \cdots + b_m e^m} \;\middle|\; \begin{array}{l} n, m \text{ 为任意非负整数}, \\ a_i,\quad b_j \in Z, \\ 0 \leqslant i \leqslant n, 0 \leqslant j \leqslant m. \end{array} \right\}$$

证明 K 是一个数域(注:e 是自然对数的底).

补 充 题 一

1. 解下列线性方程组:

$$(1)\quad \begin{cases} (1 + a_1)x_1 + x_2 + x_3 + \cdots + x_n = b_1 \\ x_1 + (1 + a_2)x_2 + x_3 + \cdots + x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_1 + x_2 + x_3 + \cdots + (1 + a_n)x_n = b_n \end{cases}$$

其中 $a_i \neq 0$, $i = 1, \cdots, n$,并且$\dfrac{1}{a_1} + \cdots + \dfrac{1}{a_n} \neq -1$.

(2) $$\begin{cases} x_1 + 2x_2 + 3x_3 + \cdots + (n-1)x_{n-1} + nx_n = b_1 \\ nx_1 + x_2 + 2x_3 + \cdots + (n-2)x_{n-1} + (n-1)x_n = b_2 \\ \cdots\cdots \\ 3x_1 + 4x_2 + 5x_3 + \cdots + x_{n-1} + 2x_n = b_{n-1} \\ 2x_1 + 3x_2 + 4x_3 + \cdots + nx_{n-1} + x_n = b_n \end{cases}$$

(3) $$\begin{cases} x_1 + x_2 + \cdots + x_n = 1 \\ x_2 + x_3 + \cdots + x_{n+1} = 2 \\ \cdots\cdots \\ x_{n+1} + x_{n+2} + \cdots + x_{2n} = n + 1 \end{cases}$$

2. 设有 n 个实数,其中任意 $n-1$ 个数之和都为零,求这 n 个数.

第二章 方阵的行列式

§1 引 言

第一章定理 1.2.1 给出的方法可用来判断线性方程组的解的情况，这需要首先把所给的线性方程组经过初等变换化成阶梯形方程组. 能不能直接从原来的线性方程组判断它的解的情况呢？我们在这一章和第三章将来讨论这个问题. 线性方程组可以用它的增广矩阵来表示，因此我们将从增广矩阵和系数矩阵中发掘出信息来判断线性方程组的解的情况.

这一章首先讨论方程个数与未知量个数相等的线性方程组，它的系数矩阵是方阵. 我们要利用的第一个信息是系数矩阵的行列式，用它来讨论线性方程组的解. 所以这一章我们主要讲方阵的行列式.

什么是行列式？如何利用线性方程组的系数矩阵的行列式讨论它的解？让我们先看二元一次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{1}$$

用加减消元法，得

$$\begin{cases} (a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12} \\ (a_{11}a_{22} - a_{12}a_{21})x_2 = b_2a_{11} - b_1a_{21} \end{cases} \tag{2}$$

于是，当 $a_{11}a_{22}-a_{12}a_{21}\neq 0$ 时，方程组(1) 有唯一解：

$$\left(\frac{b_1a_{22}-b_2a_{12}}{a_{11}a_{22}-a_{12}a_{21}},\frac{b_2a_{11}-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}}\right)$$

$a_{11}a_{22}-a_{12}a_{21}$ 是方程组(1) 的系数矩阵

$$A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$$

的元素组成的一个表达式，我们称它是2**级方阵** A **的行列式**，简称为 2 级(阶) 行列式，记作 $|A|$ 或者 $det(A)$，或者 $\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}$.
即

$$\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}:=a_{11}a_{22}-a_{12}a_{21} \qquad (4)$$

或

$$|A|:=a_{11}a_{22}-a_{12}a_{21} \qquad (5)$$

也就是，2 级矩阵 A 的行列式等于 A 的**主对角线**(由左上至右下的对角线) 上元素的乘积减去 A 的**次对角线**(由右上至左下的对角线) 上元素的乘积.

如果我们把方程组(1) 的系数矩阵 A 的第 1 列换成(1) 的常数 项，得到的矩阵记作 B_1；把 A 的第 2 列换成(1) 的常数项，得到的矩阵记作 B_2，则

$$B_1=\begin{pmatrix}b_1 & a_{12}\\ b_2 & a_{22}\end{pmatrix},\qquad B_2=\begin{pmatrix}a_{11} & b_1\\ a_{21} & b_2\end{pmatrix} \qquad (6)$$

按照 2 级矩阵的行列式的定义，我们有

$$|B_1| = b_1a_{22} - b_2a_{12}, \qquad |B_2| = b_2a_{11} - b_1a_{21} \tag{7}$$

利用行列式的概念，我们可以把上面得到的关于二元一次方程组(1)的结果叙述成：

如果二元一次方程组(1)的系数矩阵 A 的行列式 $|A| \neq 0$，则方程组(1)有唯一解，这个解是

$$\left(\frac{|B_1|}{|A|}, \quad \frac{|B_2|}{|A|}\right)$$

其中 B_1, B_2 如(6)所示.

这个结论能否推广到 n 个方程的 n 元线性方程组上?这首先需要 n 级矩阵的行列式的概念，并且讨论行列式的性质，以及会计算行列式.这些就是本章的主要内容.

习　题　2.1

1. 计算下列 2 级行列式.

(1) $\begin{vmatrix} 3 & -1 \\ 5 & 2 \end{vmatrix}$　　(2) $\begin{vmatrix} 0 & 0 \\ 1 & 4 \end{vmatrix}$　　(3) $\begin{vmatrix} -2 & 5 \\ 4 & -10 \end{vmatrix}$

2. 利用行列式解二元一次方程组：

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

§2 n 元排列

为了引进 n 级矩阵的行列式的概念，需要用到 n 元排列的一些知识.

定义 1 给定 n 个不同的自然数，它们的一个全排列称为一个 n **元排列**.

例如，自然数 1，2，3 形成的 3 元排列有：

$$123,\quad 132,\quad 213,\quad 231,\quad 312,\quad 321$$

给定 n 个不同的自然数，它们形成的全排列有 $n!$ 个，因此，对于给定的 n 个不同的自然数，n 元排列的总数是 $n!$.

我们在大多数情形下，考虑的是自然数 $1,2,\cdots,n$ 的 n 元排列，在某些情形下也需要考虑某 n 个不同的自然数的 n 元排列. 下面讨论的 n 元排列的性质. 如果没有特别声明，考虑的是前 n 个自然数的 n 元排列，但对任意 n 个不同的自然数的 n 元排列也都成立.

定义 2 在一个 n 元排列中，一对数如果前面的数小于后面的数，则称这对数构成一个**顺序**；如果前面的数大于后面的数，则称这对数构成一个**逆序**. 一个 n 元排列中出现的逆序的总数称为这个 n 元排列的**逆序数**.

n 元排列 $j_1j_2\cdots j_n$ 的逆序数记作 $\tau(j_1j_2\cdots j_n)$.

例如，5 元排列 24531 中构成逆序的数对有：21，43，41，53，51，31. 因此 $\tau(24531)=6$.

定义 3 逆序数为偶数的 n 元排列称为**偶排列**；逆序数为奇数的 n 元排列称为**奇排列**.

例如，24531 是偶排列. 因为 $\tau(23541)=5$，所以 23541 是奇排列.

定义 4 把一个 n 元排列中某两个数 i,j 的位置互换，而其余

的数不动，就得到另一个 n 元排列，这样一个变换称为一个**对换**，记作 (i,j).

给定 n 个不同的自然数，把所有 n 元排列组成的集合记作 S_n. 一个对换 (i,j) 是集合 S_n 到自身的一个映射，并且这个映射是一一对应，进一步，对换 (i,j) 把 S_n 中的 n 元排列两两配对，使每两个配成对的 n 元排列在 (i,j) 下互变. 例如，考虑 1,2,3 的 3 元排列组成的集合 S_3，对换 $(2,3)$ 是 S_3 到 S_3 的一个一一对应：

$$123 \xleftrightarrow{(2,3)} 132,\quad 213 \xleftrightarrow{(2,3)} 312,\quad 231 \xleftrightarrow{(2,3)} 321$$

5 元排列 24531 经对换 $(3,4)$ 变成 5 元排列 23541，前者是偶排列，后者是奇排列. 一般地有

定理 2.2.1　对换改变 n 元排列的奇偶性.

证明　先看对换的两个数在 n 元排列中相邻的情形：

$$\cdots\cdots \quad ij \quad \cdots\cdots \tag{Ⅰ}$$
$$\downarrow (i,j)$$
$$\cdots\cdots \quad ji \quad \cdots\cdots \tag{Ⅱ}$$

i,j 以外的数彼此间的序性在 Ⅰ 与 Ⅱ 中相同；i,j 以外的数与 i（或 j）的序性在 Ⅰ 与 Ⅱ 中也相同；若 $i<j$，则 Ⅰ 中 ij 是顺序，Ⅱ 中 ji 是逆序，于是 $\tau(\text{Ⅱ})=\tau(\text{Ⅰ})+1$；若 $i>j$，则 $\tau(\text{Ⅱ})=\tau(\text{Ⅰ})-1$. 总之，Ⅰ 与 Ⅱ 的逆序数相差 1，所以 Ⅰ 与 Ⅱ 的奇偶性相反.

再看一般情形：

$$\cdots\cdots \quad i \quad k_1\cdots\cdots k_s \quad j \quad \cdots\cdots \tag{Ⅲ}$$
$$\downarrow (i,j)$$
$$\cdots\cdots \quad j \quad k_1\cdots\cdots k_s \quad i \quad \cdots\cdots \tag{Ⅳ}$$

从 Ⅲ 变成 Ⅳ 也可以经过下列相邻两数的对换实现：

$$(i,k_1),(i,k_2),\cdots,(i,k_s),(i,j),(j,k_s),(j,k_{s-1}),\cdots,(j,k_1)$$

这一共作了 $(2s+1)$ 次相邻两数的对换. 由于 1 次相邻两数的对换

改变 n 元排列的奇偶性，所以奇数次相邻两数的对换会改变排列的奇偶性. 从而 Ⅲ 与 Ⅳ 的奇偶性相反. ▌

定理 2.2.2　$1,2,\cdots,n$ 的任一 n 元排列与排列 $12\cdots n$ 可以经过一系列对换互变，并且所作对换的次数与这个 n 元排列有相同的奇偶性.

证明　我们对 n 作归纳法证明 $1,2,\cdots,n$ 的任一 n 元排列可以经过一系列对换变成排列 $12\cdots n$.

$n=1$ 时，1 元排列只有一个，结论显然成立.

假设 $1,2,\cdots,n-1$ 的每一个 $n-1$ 元排列都能经过一系列对换变成排列 $12\cdots(n-1)$. 我们来看 $1,2,\cdots,n$ 的 n 元排列 $j_1j_2\cdots j_n$.

如果 $j_n=n$，则根据归纳假设，$j_1j_2\cdots j_{n-1}$ 可以经过一系列对换变成 $12\cdots(n-1)$，这些对换也就把 $j_1j_2\cdots j_{n-1}j_n$ 变成了 $12\cdots(n-1)n$.

如果 $j_n\neq n$，先作对换 (n,j_n) 把 $j_1j_2\cdots j_{n-1}j_n$ 变成 $j_1'\ j_2'\cdots j_{n-1}'n$，这归结为上一情形，从而 $j_1j_2\cdots j_{n-1}j_n$ 可以经过一系列对换变成 $12\cdots n$.

根据数学归纳法原理，对于任意自然数 n，$1,2,\cdots,n$ 的任一 n 元排列都可以经过一系列对换变成 $12\cdots n$.

设对换 (i_1,i_2)，(i_3,i_4)，$\cdots$，(i_{2s-1},i_{2s}) 把 $j_1j_2\cdots j_n$ 变成 $12\cdots n$（$i_1,i_2,\cdots,i_{2s-1},i_{2s}$ 可以有相同的数），则对换 (i_{2s-1},i_{2s})，$\cdots$，(i_3,i_4)，(i_1,i_2) 把 $12\cdots n$ 就变成 $j_1j_2\cdots j_n$.

设 $12\cdots n$ 经过 s 次对换变成 $j_1j_2\cdots j_n$. 若 s 为奇数，则 $j_1j_2\cdots j_n$ 与 $12\cdots n$ 奇偶性相反. 由于 $12\cdots n$ 是偶排列，因此 $j_1j_2\cdots j_n$ 是奇排列. 若 s 为偶数，则 $j_1j_2\cdots j_n$ 与 $12\cdots n$ 的奇偶性相同，于是 $j_1j_2\cdots j_n$ 是偶排列. ▌

习　题　2.2

1. 求下列各个排列的逆序数，并且指出它们的奇偶性：

(1) 315462　　　(2) 365412

(3) 654321　　　(4) 518394267

(5) 518694237　　　(6) 87654321

2. 求下列 n 元排列的逆序数：

(1) $(n-1)(n-2)\cdots 21n$　　　(2) $23\cdots(n-1)n1$

3. 在 $1,2,\cdots,n$ 的 n 元排列中

(1) 位于第 k 个位置的数 1 作成多少个逆序？

(2) 位于第 k 个位置的数 n 作成多少个逆序？

4. 写出把排列 315462 变成排列 123456 的那些对换.

5. 求 n 元排列 $n(n-1)\cdots 321$ 的逆序数，并且讨论它的奇偶性.

6. 如果 n 元排列 $j_1j_2\cdots j_{n-1}j_n$ 的逆序数为 r，求排列 $j_nj_{n-1}\cdots j_2j_1$ 的逆序数.

7. 证明在全部 n 元排列($n>1$)中，偶排列和奇排列各占一半.

§3　n 级矩阵的行列式的定义

为了给出 n 级矩阵的行列式的定义，我们来仔细分析一下 2 级矩阵的行列式的定义. 设 $A=(a_{ij})$ 是 2 级矩阵，则 $|A|=a_{11}a_{22}-a_{12}a_{21}$. 我们看出，$|A|$ 是两项的代数和；每一项是 A 的两个元素的乘积，这两个元素的行指标不同，列指标也不同，这说明它们位于 A 的不同行、不同列；第一项 $a_{11}a_{22}$ 带正号，其行指标成自然序排列，列指标的排列 12 是偶排列；第二项 $-a_{12}a_{21}$ 带负号，其行指标仍成自然序排列，但是列指标的排列 21 是奇排列. 根据这个分析，我们可以把 2 级矩阵 $A=(a_{ij})$ 的行列式写成

$$\begin{aligned}|A| &= (-1)^{\tau(12)}a_{11}a_{22}+(-1)^{\tau(21)}a_{12}a_{21}\\ &= \sum_{j_1j_2}(-1)^{\tau(j_1j_2)}a_{1j_1}a_{2j_2}\end{aligned}$$

$\sum$ 是连加号，上式表示把形如 $(-1)^{\tau(j_1j_2)}a_{1j_1}a_{2j_2}$ 的项都加起来，其中 j_1j_2 取遍 1，2 的所有 2 元排列.

这一分析启发我们如何给出 n 级矩阵的行列式的定义.

定义 1 数域 K 上的 **n 级矩阵 $A=(a_{ij})$ 的行列式** $|A|$ 是由 A 的元素组成的 $n!$ 项的代数和；其中每一项是 A 的位于不同行、不同列的 n 个元素的乘积；把每一项的 n 个元素按照行指标成自然序排好位置，当列指标的排列是偶排列时，该项带正号，当列指标的排列是奇排列时，该项带负号. 即

$$|A|:=\sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n} \tag{1}$$

其中 $j_1j_2\cdots j_n$ 是 $1,2,\cdots,n$ 的任一 n 元排列，$\sum\limits_{j_1j_2\cdots j_n}$ 表示对所有 n 元排列求和.

n 级矩阵的行列式简称为 n 级(阶)行列式. 定义 1 告诉我们，n 级行列式是指(1)式右端的表达式，它的简洁记号是 $|A|$ 或者

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

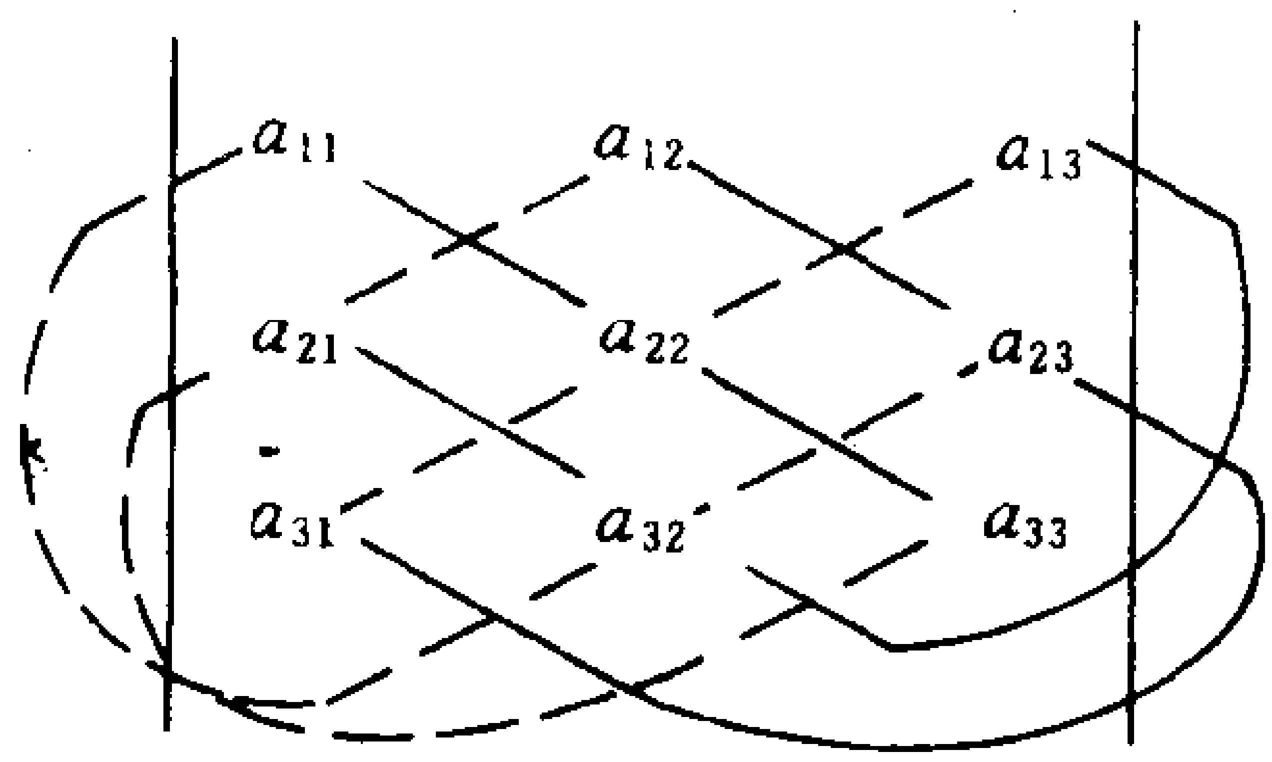
1 级矩阵 (a) 的行列式为 $|a|:=a$.

3 级矩阵 $A=(a_{ij})$ 的行列式为

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

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

3 级行列式的记忆方法如下图所示：



我们把数域 K 上的所有 n 级矩阵组成的集合记作 $M_n(K)$，则定义 1 给出了集合 $M_n(K)$ 到集合 K 的一个映射 det：

$$det(A)=\sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$$

这个映射 det 称为**行列式函数**.

主对角线下方的元素全为 0 的 n 级矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n-1} & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n-1} & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix} \qquad (2)$$

称为**上三角矩阵**.

命题 2.3.1 上三角矩阵的行列式等于它的主对角线上元素的乘积.

证明 设 $A=(a_{ij})$ 是 n 级上三角矩阵. 考虑 $|A|$ 的任意一项 $(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$. 如果 $j_n\neq n$，则 $a_{nj_n}=0$，从而该项等于零. 因此取 $j_n=n$. 如果 $j_{n-1}\neq n-1,n$，则 $a_{n-1j_{n-1}}=0$，从而该项等于零. 因此 j_{n-1} 取 $n-1$ 或 n. 但是由于 $j_n=n$，所以 j_{n-1} 不能取 n，于是取 $j_{n-1}=n-1$. 依次分析 $j_{n-2},j_{n-3},\cdots,j_2,j_1$ 可得只有取

$j_{n-2}=n-2, j_{n-3}=n-3, \cdots, j_2=2, j_1=1$ 时，$a_{1j_1}a_{2j_2}\cdots a_{nj_n}$ 才可能不等于零，而其他取法都会使 $a_{1j_1}a_{2j_2}\cdots a_{nj_n}=0$. 因此

$$|A|=(-1)^{\tau(12\cdots n)}a_{11}a_{22}\cdots a_{nn}=a_{11}a_{22}\cdots a_{nn}$$

上三角矩阵的行列式简称为上三角形行列式.

主对角线以外的元素全为 0 的 n 级矩阵

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \tag{3}$$

称为**对角矩阵**. 根据命题 2.3.1，对角矩阵的行列式等于它的主对角线上元素的乘积.

主对角线上元素全为 1 的对角矩阵称为**单位矩阵**，记作 I. 如果需要指明它是 n 级单位矩阵，则可以记成 I_n. 我们有 $|I|=1$.

在 n 级矩阵的行列式的定义中，每一项的 n 个元素是按照它们的行指标成自然序排好位置：$a_{1j_1}a_{2j_2}\cdots a_{nj_n}$，该项符号由列指标的排列 $j_1j_2\cdots j_n$ 的奇偶性决定. 由于数的乘法有交换律，因此也可以把这 n 个元素按任意次序排好位置：$a_{i_1k_1}a_{i_2k_2}\cdots a_{i_nk_n}$，这时如何从行指标的排列 $i_1i_2\cdots i_n$ 和列指标的排列 $k_1k_2\cdots k_n$ 来确定这一项所带的符号呢？设 $12\cdots n$ 经过 s 次对换变成 $i_1i_2\cdots i_n$，则 $a_{1j_1}a_{2j_2}\cdots a_{nj_n}$ 经过 s 次互换两个元素的位置变成 $a_{i_1k_1}a_{i_2k_2}\cdots a_{i_nk_n}$. 相应地，列指标的排列 $j_1j_2\cdots j_n$ 经过 s 次对换变成 $k_1k_2\cdots k_n$.

$i_1i_2\cdots i_n$ 为偶排列 $\Rightarrow s$ 为偶数 $\Rightarrow j_1j_2\cdots j_n$ 与 $k_1k_2\cdots k_n$ 有相同的奇偶性 $\Rightarrow (-1)^{\tau(j_1j_2\cdots j_n)}=(-1)^{\tau(k_1k_2\cdots k_n)}=(-1)^{\tau(k_1k_2\cdots k_n)}\cdot(-1)^{\tau(i_1i_2\cdots i_n)}=(-1)^{\tau(k_1k_2\cdots k_n)+\tau(i_1i_2\cdots i_n)}$；

$i_1i_2\cdots i_n$ 为奇排列 $\Rightarrow s$ 为奇数 $\Rightarrow j_1j_2\cdots j_n$ 与 $k_1k_2\cdots k_n$ 的奇偶性相反 $\Rightarrow (-1)^{\tau(j_1j_2\cdots j_n)}=-(-1)^{\tau(k_1k_2\cdots k_n)}=(-1)^{\tau(i_1i_2\cdots i_n)+\tau(k_1k_2\cdots k_n)}$.

因此 $$(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$$

$$=(-1)^{\tau(i_1i_2\cdots i_n)+\tau(k_1k_2\cdots k_n)}a_{i_1k_1}a_{i_2k_2}\cdots a_{i_nk_n}$$

这说明项 $a_{i_1k_1}a_{i_2k_2}\cdots a_{i_nk_n}$ 所带符号由行指标排列的逆序数与列指标排列的逆序数之和决定.

特别地，我们可以把 $|A|$ 中每一项的 n 个元素按照列指标成自然序排好位置，则这一项的符号由行指标的排列的奇偶性决定. 于是我们得到 n 级矩阵 $A=(a_{ij})$ 的行列式

$$|A|=\sum_{i_1i_2\cdots i_n}(-1)^{\tau(i_1i_2\cdots i_n)}a_{i_11}a_{i_22}\cdots a_{i_nn} \qquad (4)$$

习　题　2.3

1. 按定义计算行列式：

(1) $\begin{vmatrix} 0 & 0 & 0 & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$

(2) $\begin{vmatrix} 0 & 0 & \cdots & 0 & a_1 \\ 0 & 0 & \cdots & a_2 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n-1} & \cdots & 0 & 0 \\ a_n & 0 & \cdots & 0 & 0 \end{vmatrix}$

(3) $\begin{vmatrix} 0 & b_1 & 0 & \cdots & 0 \\ 0 & 0 & b_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & b_{n-1} \\ b_n & 0 & 0 & \cdots & 0 \end{vmatrix}$

(4) $\begin{vmatrix} 0 & \cdots & 0 & a_1 & 0 \\ 0 & \cdots & a_2 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n-1} & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & a_n \end{vmatrix}$.

2. 计算行列式

(1) $\begin{vmatrix} 2 & -1 & 5 \\ 3 & 1 & -2 \\ 1 & 4 & 6 \end{vmatrix}$　　(2) $\begin{vmatrix} c & 0 & 0 \\ 0 & a_1 & a_2 \\ 0 & b_1 & b_2 \end{vmatrix}$

3. n 阶行列式主对角线上 n 个元素的乘积带什么符号?次对角线上 n 个元素的乘积带什么符号?

4. 用行列式定义计算:

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ b_1 & b_2 & b_3 & b_4 & b_5 \\ c_1 & c_2 & 0 & 0 & 0 \\ d_1 & d_2 & 0 & 0 & 0 \\ e_1 & e_2 & 0 & 0 & 0 \end{vmatrix}$$

5. 证明:如果在 n 阶行列式中,处于某 k 行和某 l 列交叉处的各元素等于零,并且 $k+l>n$,则这个行列式等于零.

6. 求出行列式

$$\begin{vmatrix} 5x & x & 1 & x \\ 1 & x & 1 & -1 \\ 3 & 2 & x & 1 \\ 3 & 1 & 1 & x \end{vmatrix}$$

中分别包含 x^4 与 x^3 的项.

§ 4 行列式的性质

本节讨论的矩阵都是数域 K 上的矩阵.

n 级矩阵的行列式是 $n!$ 项的代数和，每一项又是 n 个元素的乘积. 如果直接用行列式的定义计算一个 n 级行列式，其计算量是相当大的. 因此我们必须研究行列式的性质，利用行列式的性质来简化行列式的计算；并且利用行列式的性质来研究线性方程组的解的情况.

设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} \end{pmatrix}$$

把 A 的行与列互换得到的矩阵

$$\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{s1} \\ a_{12} & a_{22} & \cdots & a_{s2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{sn} \end{pmatrix}$$

称为 A 的**转置**，记作 A'（有的文献记作 A^T 或 A^t）.

由定义，我们有

$$A' \text{ 的}(i,j)\text{ 元} = A \text{ 的}(j,i)\text{ 元}$$

即，$A'(i;j)=A(j;i)$.

例如，设

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

则

$$A' = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$$

此时，$A'(3;2) = 6 = A(2;3)$，等等.

命题 2.4.1　设 A 是 n 级矩阵，则 $|A'| = |A|$.

证明

$$\begin{aligned}|A'| &= \sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}A'(j_1;1)A'(j_2;2)\cdots A'(j_n;n)\\ &= \sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}A(1;j_1)A(2;j_2)A(n;j_n)\\ &= |A|\end{aligned}$$

▌

命题 2.4.1 表明，在行列式中行与列的地位是对称的，因此，行列式的有关行的性质，对于列也同样成立. 例如，主对角线上方元素全为 0 的 n 级矩阵称为下三角矩阵：

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdot & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

下三角矩阵 A 的转置 A' 是上三角矩阵. 利用命题 2.4.1 和命题 2.3.1，我们得到：$|A| = |A'| = a_{11}a_{22}\cdots a_{nn}$. 这表明，下三角矩阵的行列式也等于它的主对角元(即，主对角线上元素) 的乘积.

命题 2.4.2　行列式函数 det 对于 n 级矩阵的每一行是**线性的**.

(1) $$det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ ka_{i1} & \cdots & ka_{in} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = k \cdot det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

其中 $k \in K$， $i = 1,2,\cdots,n$;

(2) $$det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ b_1 + c_1 & \cdots & b_n + c_n \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

$$= det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ b_1 & \cdots & b_n \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} + det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ c_1 & \cdots & c_n \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

证明

(1)

$$\begin{aligned} & det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ ka_{i1} & \cdots & ka_{in} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \\ &= \sum_{j_1\cdots j_n} (-1)^{\tau(j_1\cdots j_n)} a_{1j_1}\cdots(ka_{ij_i})\cdots a_{nj_n} \\ &= k\sum_{j_1\cdots j_n} (-1)^{\tau(j_1\cdots j_n)} a_{1j_1}\cdots a_{ij_i}\cdots a_{nj_n} \\ &= k \cdot det(A) \end{aligned}$$

其中 $A = (a_{ij})$.

(2)

$$det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ b_1+c_1 & \cdots & b_n+c_n \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

$$= \sum_{j_1\cdots j_n}(-1)^{\tau(j_1\cdots j_n)}a_{1j_1}\cdots(b_{j_i}+c_{j_i})\cdots a_{nj_n}$$

$$= \sum_{j_1\cdots j_n}(-1)^{\tau(j_1\cdots j_n)}a_{1j_1}\cdots b_{j_i}\cdots a_{nj_n} + \sum_{j_1\cdots j_n}(-1)^{\tau(j_1\cdots j_n)}a_{1j_1}\cdots c_{j_i}\cdots a_{nj_n}$$

$$= det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ b_1 & \cdots & b_n \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} + det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ c_1 & \cdots & c_n \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

命题 2.4.2 的(1) 也可叙述成:行列式一行的公因子可以提出去.这条性质也表明:如果用一个非零数 k 乘 n 级矩阵 A 的某一行,则得到的矩阵的行列式等于 $|A|$ 的 k 倍.

命题 2.4.3　行列式函数 det **是斜对称的**,即如果把 n 级矩阵 $A=(a_{ij})$ 的第 i 行与第 k 行互换得到矩阵 B,则

$$det(B) = -det(A)$$

证明　因为

$$B = \begin{array}{c} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \text{第 } i \text{ 行} \\ \\ \text{第 } k \text{ 行} \\ \\ \\ \end{array}$$

所以

$$
\begin{aligned}
det(B) &= \sum_{j_1\cdots j_i\cdots j_k\cdots j_n} (-1)^{\tau(j_1\cdots j_i\cdots j_k\cdots j_n)} B(1;j_1) \\
&\quad \cdots B(i;j_i)\cdots B(k;j_k)\cdots B(n;j_n) \\
&= \sum_{j_1\cdots j_i\cdots j_k\cdots j_n} (-1)^{\tau(j_1\cdots j_i\cdots j_k\cdots j_n)} a_{1j_1}\cdots a_{kj_i}\cdots a_{ij_k}\cdots a_{nj_n} \\
&= \sum_{j_1\cdots j_k\cdots j_i\cdots j_n} -(-1)^{\tau(j_1\cdots j_k\cdots j_i\cdots j_n)} a_{1j_1}\cdots a_{ij_k}\cdots a_{kj_i}\cdots a_{nj_n} \\
&= -\sum_{j_1\cdots j_k\cdots j_i\cdots j_n} (-1)^{\tau(j_1\cdots j_k\cdots j_i\cdots j_n)} a_{1j_1}\cdots a_{ij_k}\cdots a_{kj_i}\cdots a_{nj_n} \\
&= -det(A)
\end{aligned}
$$

推论 2.4.1　如果 n 级矩阵 A 有零行，则 $|A| = 0$.

证明　在命题 2.4.2 的(1) 中取 $k = 0$ 即得.

推论 2.4.2　如果 A 有两行相同，则 $|A| = 0$.

证明　设 A 的第 i 行与第 k 行相同，则 $A \xrightarrow{(ⓘ,ⓚ)} A$. 由命题 2.4.3 得，$det(A) = -det(A)$. 由此得 $det(A) = 0$.

推论 2.4.3　如果 A 有两行成比例，则 $|A| = 0$.

证明　没

$$
A = \begin{pmatrix} \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ la_{i1} & \cdots & la_{in} \\ \cdots & \cdots & \cdots \end{pmatrix}
$$

据命题 2.4.2 的(1)和推论 2.4.2 得

$$
det(A) = l \cdot det \begin{pmatrix} \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \end{pmatrix} = 0
$$

推论 2.4.4　设 $A \xrightarrow{Ⓚ+Ⓘ\cdot l} B$，则 $|B| = |A|$.

证明　设 $A = (a_{ij})$，则

$$
|B| = \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ a_{k1} + la_{i1} & \cdots & a_{kn} + la_{in} \\ \cdots & \cdots & \cdots \end{vmatrix}
$$

$$
= \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots \end{vmatrix} + \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ la_{i1} & \cdots & la_{in} \\ \cdots & \cdots & \cdots \end{vmatrix}
$$

$$
= |A| + 0
$$

$$
= |A|
$$

推论 2.4.4 表明：1° 型初等行变换不改变行列式的值. 从推论 2.4.4、命题 2.4.3 和命题 2.4.2 的(1) 知道，如果 n 级矩阵 A 经过一系列初等行变换变成矩阵 B，则

$$|B| = l|a|，其中 l 是 K 中某个非零数$$

我们已经知道，n 级矩阵 A 一定可以经过一系列初等行变换化成行阶梯形矩阵，而行阶梯形方阵一定是上三角矩阵，并且上三角矩阵的行列式很容易计算，因此计算行列式的方法之一是把 n 级矩阵经过初等行变换化成上三角矩阵，要注意所作的初等行变换会不会引起行列式的变化. 此外，在计算行列式时还常常利用命题 2.4.2 的(2).

例 1　计算行列式

$$
\begin{vmatrix} -2 & 3 & 1 \\ 503 & 201 & 298 \\ 5 & 2 & 3 \end{vmatrix}
$$

解

$$\begin{vmatrix} -2 & 3 & 1 \\ 503 & 201 & 298 \\ 5 & 2 & 3 \end{vmatrix} = \begin{vmatrix} -2 & 3 & 1 \\ 500+3 & 200+1 & 300-2 \\ 5 & 2 & 3 \end{vmatrix}$$

$$= \begin{vmatrix} -2 & 3 & 1 \\ 500 & 200 & 300 \\ 5 & 2 & 3 \end{vmatrix} + \begin{vmatrix} -2 & 3 & 1 \\ 3 & 1 & -2 \\ 5 & 2 & 3 \end{vmatrix}$$

$$\underset{(①,③)}{=} 0 - \begin{vmatrix} 1 & 3 & -2 \\ -2 & 1 & 3 \\ 3 & 2 & 5 \end{vmatrix} = - \begin{vmatrix} 1 & 3 & -2 \\ 0 & 7 & -1 \\ 0 & -7 & 11 \end{vmatrix}$$

$$= - \begin{vmatrix} 1 & 3 & -2 \\ 0 & 7 & -1 \\ 0 & 0 & 10 \end{vmatrix} = -70$$

在例 1 的第三步计算中，(①,③) 表示把矩阵的第一列与第三列对换(我们约定，列的变换的记号写在等号或箭号下面，而行的变换的记号写在上面)，根据行与列的地位对称可知，两列互换，行列式要反号. 从这个例子看出，有时需要对矩阵的列作变化. 类似于矩阵的初等行变换，有矩阵的**初等列变换**：

1° 把矩阵的一列的倍数加到另一列上；

2° 互换两列的位置；

3° 用数域 K 中一个非零数乘矩阵的某一列.

矩阵的初等列变换如何引起行列式的变化，与矩阵的初等行变换的情形是一样的.

例 2 计算 n 级行列式

$$\begin{vmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b & b & b & \cdots & a \end{vmatrix}$$

解 这个行列式的特点是:每一行的元素之和等于常数 $a+(n-1)b$. 把第2列、第3列、…、第 n 列分别加到第一列上,便可以使第一列有公因子 $a+(n-1)b$,把它提出去,第一列便全变成1,再用矩阵的初等行变换便容易得到一个上三角矩阵的行列式.

$$\begin{vmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b & b & b & \cdots & a \end{vmatrix}$$

$$=\begin{vmatrix} a+(n-1)b & b & b & \cdots & b \\ a+(n-1)b & a & b & \cdots & b \\ a+(n-1)b & b & a & \cdots & b \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a+(n-1)b & b & b & \cdots & a \end{vmatrix}$$

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

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

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

当 $n=1$ 时,上述公式也成立.

习 题 2.4

1. 计算下列行列式：

(1) $\begin{vmatrix} 5 & -1 & 3 \\ 2 & 2 & 2 \\ 196 & 203 & 199 \end{vmatrix}$ (2) $\begin{vmatrix} -1 & 203 & \frac{1}{3} \\ 3 & 298 & \frac{1}{2} \\ 5 & 399 & \frac{2}{3} \end{vmatrix}$

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

(5) $\begin{vmatrix} a_1+b & a_2 & \cdots & a_n \\ a_1 & a_2+b & \cdots & a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_1 & a_2 & \cdots & a_n+b \end{vmatrix}$

(6) $\begin{vmatrix} a^2 & (a-1)^2 & (a-2)^2 & (a-3)^2 \\ b^2 & (b-1)^2 & (b-2)^2 & (b-3)^2 \\ c^2 & (c-1)^2 & (c-2)^2 & (c-3)^2 \\ d^2 & (d-1)^2 & (d-2)^2 & (d-3)^2 \end{vmatrix}$

(7) $\begin{vmatrix} a_1+b_1 & a_1+b_2 & \cdots & a_1+b_n \\ a_2+b_1 & a_2+b_2 & \cdots & a_2+b_n \\ \cdots & \cdots & \cdots & \cdots \\ a_n+b_1 & a_n+b_2 & \cdots & a_n+b_n \end{vmatrix}$

2. 设 $i=\sqrt{-1}$，证明

(1) $$\begin{vmatrix} a_1+b_1i & a_1i-b_1 & c_1 \\ a_2+b_2i & a_2i-b_2 & c_2 \\ a_3+b_3i & a_3i-b_3 & c_3 \end{vmatrix}=0$$

(2) $$\begin{vmatrix} a_1+b_1i & a_1i+b_1 & c_1 \\ a_2+b_2i & a_2i+b_2 & c_2 \\ a_3+b_3i & a_3i+b_3 & c_3 \end{vmatrix}=2\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

§5 行列式按一行(列)展开

这一节我们进一步讨论行列式的性质. 我们先以 3 级矩阵

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

的行列式为例来说明行列式的又一性质.

$$\begin{aligned} |A| &= a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32} \\ &\quad -a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32} \\ &= a_{11}(a_{22}a_{33}-a_{23}a_{32})-a_{12}(a_{21}a_{33}-a_{23}a_{31}) \\ &\quad +a_{13}(a_{21}a_{32}-a_{22}a_{31}) \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}-a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}+a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \end{aligned} \tag{2}$$

公式(2)的右端的 2 级行列式

$$\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

是 A 中划去 A 的(1,1)元 a_{11} 所在的第 1 行、第 1 列，剩下的元素按原来的排法组成的 2 级矩阵的行列式. 公式(2) 右端的其他两个 2

级行列式可以用类似的方法得到. 公式(2) 的右端出现的 a_{11}, a_{12}, a_{13} 正好是 A 的第 1 行的元素. 公式(2) 称为 3 级矩阵 A 的行列式按第 1 行展开. 本节我们把这一性质加以推广.

定义 1 设 $A=(a_{ij})$ 是数域 K 上的 n 级矩阵. 划去 A 的 (i,j) 元 a_{ij} 所在的第 i 行、第 j 列，剩下的 $(n-1)^2$ 个元素按原来排法组成的 $n-1$ 级矩阵的行列式称为 **A 的 (i,j) 元的余子式**，记作 M_{ij}. 令

$$A_{ij} := (-1)^{i+j} M_{ij}$$

称 A_{ij} 是矩阵 **A 的 (i,j) 元的代数余子式**.

例如，在上面的 3 级矩阵 $A=(a_{ij})$ 中，公式(2) 右端出现的 3 个 2 级行列式依次是 A 的 (1,1) 元，(1,2) 元，(1,3) 元的余子式，它们的代数余子式则分别为

$$A_{11}=\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \qquad A_{12}=-\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \qquad A_{13}=\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

于是公式(2)可以写成

$$|A| = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \tag{2$'$}$$

运用代数余子式的记号，可以使公式(2)$'$ 的右端的每一项带正号，比较整齐. 又如，3 级矩阵 $A=(a_{ij})$ 的 (2,3) 元的代数余子式为

$$A_{23}=-\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}.$$

一般地，n 级矩阵 $A=(a_{ij})$ 的 (i,j) 元的余子式为

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

$$\overset{*}{=}\sum_{k_1\cdots k_{i-1}k_{i+1}\cdots k_n}(-1)^{\tau(k_1\cdots k_{i-1}k_{i+1}\cdots k_n)}a_{1k_1}\cdots a_{i-1,k_{i-1}}a_{i+1,k_{i+1}}\cdots a_{nk_n} \quad (3)$$

其中 $k_1\cdots k_{i-1}k_{i+1}\cdots k_n$ 是 $1,2,\cdots,j-1,j+1,\cdots,n$ 的 $n-1$ 元排列.

不难看出,如果作映射 $\sigma:1\mapsto 1,\cdots,j-1\mapsto j-1,j+1\mapsto j,\cdots,n\mapsto n-1$,那么 $\sigma(k_1)\cdots\sigma(k_{i-1})\sigma(k_{i+1})\cdots\sigma(k_n)$ 是 $1,\cdots,j-1,j,\cdots,n-1$ 的 $n-1$ 元排列,并且它的逆序数与 $k_1\cdots k_{i-1}k_{i+1}\cdots k_n$ 的逆序数相同.因此加 $*$ 的等号的确成立.

定理 2.5.1　设 $A=(a_{ij})$ 是数域 K 上的 n 级矩阵,则 $|A|$ 等于 A 的第 i 行元素与它自己的代数余子式的乘积之和,即

$$|A|=a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in} \quad (4)$$

其中 $i\in\{1,2,\cdots,n\}$.

证明

$$\begin{aligned}|A| &= \sum_{k_1\cdots k_{i-1}jk_{i+1}\cdots k_n}(-1)^{\tau(k_1\cdots k_{i-1}jk_{i+1}\cdots k_n)} \\ &\quad\cdot a_{1k_1}\cdots a_{i-1,k_{i-1}}a_{ij}a_{i+1,k_{i+1}}\cdots a_{nk_n} \\ &= \sum_{jk_1\cdots k_{i-1}k_{i+1}\cdots k_n}(-1)^{\tau(i1\cdots i-1i+1\cdots n)+\tau(jk_1\cdots k_{i-1}k_{i+1}\cdots k_n)} \\ &\quad\cdot a_{ij}a_{1k_1}\cdots a_{i-1,k_{i-1}}a_{i+1,k_{i+1}}\cdots a_{nk_n}\end{aligned}$$

$$= \sum_{jk_1\cdots k_{i-1}k_{i+1}\cdots k_n} (-1)^{i-1} \cdot (-1)^{(j-1)+\tau(k_1\cdots k_{i-1}k_{i+1}\cdots k_n)}$$

$$\cdot a_{ij}(a_{1k_1}\cdots a_{i-1,k_{i-1}}a_{i+1,k_{i+1}}\cdots a_{nk_n})$$

$$= \sum_{j=1}^{n} \sum_{k_1\cdots k_{i-1}k_{i+1}\cdots k_n} (-1)^{i+j} \cdot (-1)^{\tau(k_1\cdots k_{i-1}k_{i+1}\cdots k_n)}$$

$$\cdot a_{ij}(a_{1k_1}\cdots a_{i-1,k_{i-1}}a_{i+1,k_{i+1}}\cdots a_{nk_n})$$

$$= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \Big(\sum_{k_1\cdots k_{i-1}k_{i+1}\cdots k_n} (-1)^{\tau(k_1\cdots k_{i-1}k_{i+1}\cdots k_n)}$$

$$\cdot a_{1k_1}\cdots a_{i-1,k_{i-1}}a_{i+1,k_{i+1}}\cdots a_{nk_n}\Big)$$

$$= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}$$

$$= \sum_{j=1}^{n} a_{ij} A_{ij}$$

公式(4) 称为 n 级矩阵 A 的**行列式按第 i 行展开**.

根据定理 2.5.1,我们也可以把 3 级矩阵 $A=(a_{ij})$ 的行列式按第 2 行展开,得

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

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

利用行列式中行与列的地位的对称性,从定理 2.5.1 可以得到

定理 2.5.2　n 级矩阵 $A=(a_{ij})$ 的行列式等于 A 的第 j 列元素与它自己的代数余子式的乘积之和,即

$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \tag{5}$$

其中 $j \in \{1,2,\cdots,n\}$.

证明　将 A' 按第 j 行展开,由于 A' 的 (j,l) 元等于 A 的 (l,j) 元,并且 A' 的 (j,l) 元的代数余子式等于 A 的 (l,j) 元的代数余子式 A_{lj}, $l=1,2,\cdots,n$,因此

$$|A| = |A'| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}$$

公式(5) 称为 n 级矩阵的**行列式按第 j 列展开**.

例如，3 级矩阵 $A=(a_{ij})$ 的行列式按第 1 列展开得

$$\begin{aligned}|A| &= a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31} \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33}\end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33}\end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23}\end{vmatrix}\end{aligned}$$

又如,设 3 级矩阵 B_1 为

$$B_1 = \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33}\end{pmatrix}$$

则 $|B_1|$ 按第 1 列展开为

$$\begin{aligned}|B_1| &= b_1A_{11} + b_2A_{21} + b_3A_{31} \\ &= b_1\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33}\end{vmatrix} - b_2\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33}\end{vmatrix} + b_3\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23}\end{vmatrix}\end{aligned}$$

定理 2.5.3　n 级矩阵 $A=(a_{ij})$ 的第 i 行元素与第 k 行($k \neq i$) 相应元素的代数余子式的乘积之和等于零. 即

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} = 0, \quad 当\ k \neq i \tag{6}$$

证明　考虑矩阵

$$B = \begin{array}{l} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn}\end{pmatrix} \begin{array}{l} \\ \\ 第\ i\ 行 \\ \\ 第\ k\ 行 \\ \\ \\ \end{array}\end{array}$$

把 $|B|$ 按第 k 行展开，由于 B 的 (k,j) 元的代数余子式等于矩阵 A

的(k,j)元的代数余子式A_{kj}，其中$j=1,2,\cdots,n$，所以

$$|B| = a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn}$$

由于B有两行相同，所以$|B|=0$. 由此即得公式(6). ▎

由于行列式中行与列的地位对称，因此也有

定理 2.5.4　n级矩阵$A=(a_{ij})$的第j列元素与第l列$(l\neq j)$相应元素的代数余子式的乘积之和等于零：即

$$a_{1j}A_{1l} + a_{2j}A_{2l} + \cdots + a_{nj}A_{nl} = 0, \quad \text{当 } l \neq j \tag{7}$$

▎

公式(4)、(5)、(6)、(7)可以用连加号写成

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \begin{cases} |A| & \text{当 } k = i \\ 0 & \text{当 } k \neq i \end{cases}$$

$$\sum_{i=1}^{n} a_{ij}A_{il} = \begin{cases} |A| & \text{当 } l = j \\ 0 & \text{当 } l \neq j \end{cases}$$

利用行列式按一行(列)展开可以简化行列式的计算. 当然，如果直接用公式(4)或(5)去计算n级行列式，则一般来说需要计算n个$n-1$级行列式，这并没有减少计算量. 但是我们先作初等行(列)变换使行列式的某一行或某一列有许多元素为0，则按这一行或这一列展开就可减少计算量.

例1　计算4级矩阵

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

的行列式.

解　为了尽量避免分数运算，选择A中元素1所在的第2行(或第2列)展开$|A|$.

$$|A| \xlongequal[\substack{①+②\cdot 4\\ ③+②\cdot 2\\ ④+②\cdot(-3)}]{} \begin{vmatrix} -10 & -3 & 1 & 14 \\ 0 & 1 & 0 & 0 \\ 19 & 4 & 14 & -19 \\ 0 & -2 & -1 & 1 \end{vmatrix}$$

$$\xlongequal{\quad} 1\cdot(-1)^{2+2}\begin{vmatrix} -10 & 1 & 14 \\ 19 & 14 & -19 \\ 0 & -1 & 1 \end{vmatrix}$$

$$\xlongequal[②+③]{} \begin{vmatrix} -10 & 15 & 14 \\ 19 & -5 & -19 \\ 0 & 0 & 1 \end{vmatrix}$$

$$\xlongequal{\quad} 1\cdot(-1)^{3+3}\begin{vmatrix} -10 & 15 \\ 19 & -5 \end{vmatrix}$$

$$\xlongequal{\quad} 5\begin{vmatrix} -10 & 3 \\ 19 & -1 \end{vmatrix}$$

$$\xlongequal{\quad} -235$$

例 2　计算行列式

$$\begin{vmatrix} a-6 & 2 & -2 \\ 2 & a-3 & -4 \\ -2 & -4 & a-3 \end{vmatrix}$$

解

$$\begin{vmatrix} a-6 & 2 & -2 \\ 2 & a-3 & -4 \\ -2 & -4 & a-3 \end{vmatrix}$$

$$\xlongequal{③+②\cdot 1} \begin{vmatrix} a-6 & 2 & -2 \\ 2 & a-3 & -4 \\ 0 & a-7 & a-7 \end{vmatrix}$$

$$\xlongequal{②+③\cdot(-1)}\begin{vmatrix} a-6 & 4 & -2 \\ 2 & a+1 & -4 \\ 0 & 0 & a-7 \end{vmatrix}$$

$$=(a-7)(-1)^{3+3}\begin{vmatrix} a-6 & 4 \\ 2 & a+1 \end{vmatrix}$$

$$=(a-7)(a^2-5a-14)$$

$$=(a-7)^2(a+2)$$

例 3　计算 n 级矩阵

$$A=\begin{pmatrix} a & b & 0 & \cdots & 0 & 0 \\ 0 & a & b & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a & b \\ b & 0 & 0 & \cdots & 0 & a \end{pmatrix}.$$

的行列式($n>1$).

解　把 $|A|$ 按第 1 列展开得

$$|A|=a\begin{vmatrix} a & b & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a & b \\ 0 & 0 & \cdots & 0 & a \end{vmatrix}$$

$$+b\cdot(-1)^{n+1}\begin{vmatrix} b & 0 & \cdots & 0 & 0 \\ a & b & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a & b \end{vmatrix}$$

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

$$=a^n+(-1)^{n+1}b^n$$

习 题 2.5

1. 计算下列行列式

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

(2) $\begin{vmatrix} 2 & -4 & -3 & 5 \\ -3 & 1 & 4 & -2 \\ 7 & 2 & 5 & 3 \\ 4 & -3 & -2 & 6 \end{vmatrix}$

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

(4) $\begin{vmatrix} a-2 & -2 & 2 \\ -2 & a-5 & 4 \\ 2 & 4 & a-5 \end{vmatrix}$

(5) $\begin{vmatrix} a-2 & -3 & -2 \\ -1 & a-8 & -2 \\ 2 & 14 & a+3 \end{vmatrix}$

(6) $\begin{vmatrix} a-3 & -1 & 0 & 1 \\ -1 & a-3 & 1 & 0 \\ 0 & 1 & a-3 & -1 \\ 1 & 0 & -1 & a-3 \end{vmatrix}$

(7) $\begin{vmatrix} 1 & 0 & 1+i \\ 0 & 1 & i \\ 1-i & -i & 1 \end{vmatrix}$

(8) $\begin{vmatrix} a & b & c & 0 & d \\ 0 & e & 0 & 0 & f \\ 0 & g & h & 0 & k \\ l & p & q & r & s \\ 0 & 0 & 0 & 0 & t \end{vmatrix}$

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

2. 解方程

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1-x & 1 & \cdots & 1 \\ 1 & 1 & 2-x & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & n-x \end{vmatrix} = 0$$

3. 计算下列 n 级行列式

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

3

(2) $\begin{vmatrix} a_1 & a_2 & a_3 & \cdots & a_{n-1} & a_n \\ 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 2 & -2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & n-1 & 1-n \end{vmatrix}$.

§6 n 级行列式的计算

计算行列式的基本方法是利用行列式的性质，运用矩阵的初等行(列)变换把 n 级矩阵的行列式转化成上三角矩阵的行列式；或者是把行列式的某一行(列)的很多元素化成零，然后按这一行(列)展开. 此外，还需要针对所给行列式的具体特点运用一些技巧. 本节举一些例子.

例 1 计算 n 级矩阵

$$A=\begin{pmatrix} x_1-a_1 & x_2 & x_3 & \cdots & x_n \\ x_1 & x_2-a_2 & x_3 & \cdots & x_n \\ x_1 & x_2 & x_3-a_3 & \cdots & x_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_1 & x_2 & x_3 & \cdots & x_n-a_n \end{pmatrix},$$

$$a_i\neq 0,\quad i=1,2,\cdots,n$$

的行列式.

解 先把 A 的第 1 行的 (-1) 倍分别加到第 2 行，…，第 n 行上，然后各列分别提出公因子 $a_1,\cdots,a_n$：

$$|A| \xlongequal{\quad} \begin{vmatrix} x_1-a_1 & x_2 & x_3 & \cdots & x_n \\ a_1 & -a_2 & 0 & \cdots & 0 \\ a_1 & 0 & -a_3 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_1 & 0 & 0 & \cdots & -a_n \end{vmatrix}$$

$$
= a_1a_2\cdots a_n\begin{vmatrix} \frac{x_1}{a_1}-1 & \frac{x_2}{a_2} & \frac{x_3}{a_3} & \cdots & \frac{x_n}{a_n} \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & 0 & \cdots & -1 \end{vmatrix}
$$

$$
\xlongequal[\substack{①+② \\ \cdots \\ ①+ⓝ}]{} a_1a_2\cdots a_n\begin{vmatrix} \sum_{i=1}^{n}\frac{x_i}{a_i}-1 & \frac{x_2}{a_2} & \frac{x_3}{a_3} & \cdots & \frac{x_n}{a_n} \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -1 \end{vmatrix}
$$

$$
= (-1)^{n-1}a_1a_2\cdots a_n\left(\sum_{i=1}^{n}\frac{x_i}{a_i}-1\right)
$$

当 $n=1$ 时，上述公式也成立.

例 2 n 级矩阵

$$
A=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_1^{n-2} & a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2} \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \end{pmatrix} \tag{1}
$$

称为 n 级范德蒙(Vandermonde)矩阵，它的行列式称为 n 级范德蒙行列式. 证明：对于 $n\geqslant 2$，有

$$
\begin{aligned} |A| = &(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1) \\ &\cdot(a_3-a_2)\cdots(a_n-a_2) \\ &\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ &\cdot(a_n-a_{n-1}) \end{aligned}
$$

$$= \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j) \tag{2}$$

证明 对 n 作归纳法.

当 $n = 2$ 时,

$$\begin{vmatrix} 1 & 1 \\ a_1 & a_2 \end{vmatrix} = a_2 - a_1$$

命题为真.

假设对于 $n-1$ 级范德蒙行列式,命题为真. 现在来看 n 级范德蒙行列式 $|A|$. 把 A 的第 $n-1$ 行的 $(-a_1)$ 倍加到第 n 行上,然后把第 $n-2$ 行的 $(-a_1)$ 倍加到第 $n-1$ 行上,依次类推,得

$$|A| = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ 0 & a_2^2 - a_1 a_2 & a_3^2 - a_1 a_3 & \cdots & a_n^2 - a_1 a_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_2^{n-2} - a_1 a_2^{n-3} & a_3^{n-2} - a_1 a_3^{n-3} & \cdots & a_n^{n-2} - a_1 a_n^{n-3} \\ 0 & a_2^{n-1} - a_1 a_2^{n-2} & a_3^{n-1} - a_1 a_3^{n-2} & \cdots & a_n^{n-1} - a_1 a_n^{n-2} \end{vmatrix}$$

$$= \begin{vmatrix} a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ a_2(a_2 - a_1) & a_3(a_3 - a_1) & \cdots & a_n(a_n - a_1) \\ \cdots & \cdots & \cdots & \cdots \\ a_2^{n-3}(a_2 - a_1) & a_3^{n-3}(a_3 - a_1) & \cdots & a_n^{n-3}(a_n - a_1) \\ a_2^{n-2}(a_2 - a_1) & a_3^{n-2}(a_3 - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{vmatrix}$$

$$= (a_2 - a_1)(a_3 - a_1)\cdots(a_n - a_1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_2 & a_3 & \cdots & a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_2^{n-3} & a_3^{n-3} & \cdots & a_n^{n-3} \\ a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2} \end{vmatrix}$$

$$\xlongequal{\text{用归纳假设}}(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)(a_3-a_2)\cdots(a_n-a_2)$$
$$\cdot(a_4-a_3)\cdots(a_n-a_3)\cdots(a_n-a_{n-1})$$
$$=\prod_{1\leqslant j<i\leqslant n}(a_i-a_j)$$

由归纳法原理，对一切 $n\geqslant 2$，命题为真. ▌

从例 2 得出：n 级范德蒙行列式等于零的充分必要条件是 a_1，$a_2,\cdots,a_n$ 这 n 个数中至少有两个相等.

例 3 计算 n 级行列式

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

解 $D_1=|2|=2$. 当 $n>1$ 时，将第 2 列至第 n 列都加到第 1 列上，再按第 1 列展开得

$$D_n=\begin{vmatrix} 1 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{vmatrix}$$
$$=1\cdot D_{n-1}+1\cdot(-1)^{n+1}\cdot(-1)^{n-1}=D_{n-1}+1$$

于是 $D_1,D_2,\cdots,D_n$ 是首项为 2，公差为 1 的等差数列，因此

$$D_n=2+(n-1)\cdot 1=n+1$$

例 4 计算 n 级行列式

$$D_n=\begin{vmatrix} x & y & y & \cdots & y & y \\ z & x & y & \cdots & y & y \\ z & z & x & \cdots & y & y \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z & z & z & \cdots & x & y \\ z & z & z & \cdots & z & x \end{vmatrix}, \qquad y \neq z$$

解

$$D_n=\begin{vmatrix} x & y & y & \cdots & y & 0+y \\ z & x & y & \cdots & y & 0+y \\ z & z & x & \cdots & y & 0+y \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z & z & z & \cdots & x & 0+y \\ z & z & z & \cdots & z & (x-y)+y \end{vmatrix}$$

$$=\begin{vmatrix} x & y & y & \cdots & y & 0 \\ z & x & y & \cdots & y & 0 \\ z & z & x & \cdots & y & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z & z & z & \cdots & x & 0 \\ z & z & z & \cdots & z & x-y \end{vmatrix}$$

$$+\begin{vmatrix} x & y & y & \cdots & y & y \\ z & x & y & \cdots & y & y \\ z & z & x & \cdots & y & y \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z & z & z & \cdots & x & y \\ z & z & z & \cdots & z & y \end{vmatrix}$$

$$
=(x-y)\begin{vmatrix} x & y & y & \cdots & y \\ z & x & y & \cdots & y \\ z & z & x & \cdots & y \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ z & z & z & \cdots & x \end{vmatrix}
$$

$$
+\begin{vmatrix} x & y & y & \cdots & y & y \\ z-x & x-y & 0 & \cdots & 0 & 0 \\ 0 & z-x & x-y & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & x-y & 0 \\ 0 & 0 & 0 & \cdots & z-x & 0 \end{vmatrix}
$$

$$
=(x-y)D_{n-1}+y\cdot(-1)^{1+n}\cdot(z-x)^{n-1}
$$

$$
=(x-y)D_{n-1}+y(x-z)^{n-1}\qquad(n\geqslant 2)
$$

设 D_n 是 n 级矩阵 A 的行列式，则 $|A'|=|A|=D_n$. 对 $|A'|$ 运用刚刚证得的结果，便得到

$$
D_n=|A'|=(x-z)D_{n-1}+z(x-y)^{n-1}\qquad(n\geqslant 2)
$$

于是得到

$$
\begin{cases} D_n=(x-y)D_{n-1}+y(x-z)^{n-1} \\ D_n=(x-z)D_{n-1}+z(x-y)^{n-1} \end{cases}\qquad(n\geqslant 2)
$$

解之得

$$
D_n=\frac{y(x-z)^n-z(x-y)^n}{y-z}\qquad(n\geqslant 2)
$$

易验证上式对于 $n=1$ 时也成立.

习　题　2.6

1. 解方程

$$\begin{vmatrix} 1 & x & x^2 & \cdots & x^{n-1} \\ 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^{n-1} \end{vmatrix} = 0$$

其中 $a_1, a_2, \cdots, a_{n-1}$ 是互不相同的数.

2. 计算下列 n 级行列式

(1) $$\begin{vmatrix} a_1 - y & a_2 & \cdots & a_n \\ a_1 & a_2 - y & \cdots & a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_1 & a_2 & \cdots & a_n - y \end{vmatrix}$$

*(2) $$\begin{vmatrix} 1-a_1 & a_2 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1-a_2 & a_3 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1-a_3 & a_4 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1-a_n \end{vmatrix}$$

3. 用数学归纳法证明

$$\begin{vmatrix} x & 0 & \cdots & 0 & a_0 \\ -1 & x & \cdots & 0 & a_1 \\ 0 & -1 & \cdots & 0 & a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & x & a_{n-2} \\ 0 & 0 & \cdots & -1 & x + a_{n-1} \end{vmatrix} = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \qquad (n \geqslant 2)$$

4. 证明

(1) $$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & a_1 & 0 & \cdots & 0 \\ 1 & 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & 0 & \cdots & a_n \end{vmatrix} = a_1a_2\cdots a_n\left(1 - \sum_{i=1}^{n}\frac{1}{a_i}\right)$$

其中 $a_1, a_2, \cdots, a_n$ 全不为零；

(2) $$D_n = \begin{vmatrix} 5 & 3 & 0 & 0 & \cdots & 0 & 0 \\ 2 & 5 & 3 & 0 & \cdots & 0 & 0 \\ 0 & 2 & 5 & 3 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 2 & 5 \end{vmatrix} = 3^{n+1} - 2^{n+1}$$

(3) $$\begin{vmatrix} \alpha+\beta & \alpha\beta & 0 & \cdots & 0 & 0 \\ 2 & \alpha+\beta & \alpha\beta & \cdots & 0 & 0 \\ 0 & 1 & \alpha+\beta & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & \alpha+\beta \end{vmatrix} = \alpha^n + \beta^n$$

(4) $$D_n = \begin{vmatrix} 1+x^2 & x & 0 & 0 & \cdots & 0 & 0 \\ x & 1+x^2 & x & 0 & \cdots & 0 & 0 \\ 0 & x & 1+x^2 & x & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & x & 1+x^2 \end{vmatrix}$$
$$= 1 + x^2 + x^4 + \cdots + x^{2n}$$

5. 设 $a_1, a_2, \cdots, a_n$ 是数域 K 中互不相同的数，$b_1, b_2, \cdots, b_n$ 是 K 中任意一组给定的数，证明：存在唯一的数域 K 上的多项式 $f(x) = c_0x^{n-1} + c_1x^{n-2} + \cdots + c_{n-1}$ 使得

$$f(a_i) = b_i, \qquad i = 1, 2, \cdots, n$$

6. 计算 n 级行列式

$$\begin{vmatrix} 2a & a^2 & 0 & \cdots & 0 & 0 \\ 1 & 2a & a^2 & \cdots & 0 & 0 \\ 0 & 1 & 2a & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 2a & a^2 \\ 0 & 0 & 0 & \cdots & 1 & 2a \end{vmatrix}$$

§7 用行列式判断线性方程组的解的情况 · Cramer 法则

我们在 §4 已指出，如果 n 级矩阵 A 经过一系列初等行变换变成矩阵 B，则 $|B| = l|A|$，其中 l 是 K 中某个非零数. 根据这个结论，我们可以利用行列式来判断线性方程组的解的情况.

设数域 K 上的 n 元线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{cases} \tag{1}$$

的系数矩阵为 $A = (a_{ij})$，增广矩阵为 $\widetilde{A}$. 对 $\widetilde{A}$ 施行初等行变换化成行阶梯形矩阵 $\widetilde{J}$，则相应地，A 经过这些初等行变换被化成行阶梯形矩阵 J，其中 J 比 $\widetilde{J}$ 少最后一列. 我们有 $|J| = l|A|$，其中 l 是 K 中某个非零数. 于是

$|A| \neq 0 \quad \Rightarrow |J| \neq 0 \quad \Rightarrow J$ 没有零行 $\quad \Rightarrow J$ 有 n 个主元

设 J 的 n 个主元的列指标依次为 $j_1, j_2, \cdots, j_n$，则 $1 \leqslant j_1 < j_2 < \cdots < j_n \leqslant n$. 由此得 $j_1 = 1, j_2 = 2, \cdots, j_n = n$. 因此

$$J = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ 0 & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & c_{nn} \end{pmatrix}$$

其中 $c_{11} \neq 0, c_{22} \neq 0, \cdots, c_{nn} \neq 0$. 从而

$$\tilde{J} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} & d_1 \\ 0 & c_{22} & \cdots & c_{2n} & d_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & c_{nn} & d_n \end{pmatrix}$$

于是 $\tilde{J}$ 表示的阶梯形方程组不会出现方程“$0 = d$”(其中 d 是非零数),因此当 $|A| \neq 0$ 时,方程组(1) 有解;由于阶梯形方程组的方程个数等于 n,因此(1) 有唯一解.

若 $|A| = 0$,则 $|J| = 0$. 假如 J 没有零行,则从上面一段知道 $|J| = c_{11}c_{22}\cdots c_{nn} \neq 0$,矛盾. 因此 J 必有零行. 从而

$$\tilde{J} = \begin{pmatrix} c_{11} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & c_{1n} & d_1 \\ 0 & \cdots & 0 & c_{2j_2} & \cdots & \cdots & \cdots & \cdots & c_{2n} & d_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & c_{rj_r} & \cdots & c_{rn} & d_r \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & d_{r+1} \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{pmatrix}$$

其中 $c_{11} \neq 0, c_{2j_2} \neq 0, \cdots, c_{rj_r} \neq 0; r < n$.

此时,有两种可能:1) $d_{r+1} \neq 0$,这时方程组(1) 无解;

2） $d_{r+1}=0$，这时方程组(1)有无穷多个解.

这样我们便证明了下面的

定理 2.7.1 数域 K 上的 n 元线性方程组(1)，如果它的系数矩阵 A 的行列式 $|A|\neq 0$，则(1)有唯一解；如果 $|A|=0$，则(1)或者无解，或者有无穷多个解. ▌

推论 2.7.1 数域 K 上的 n 元线性方程组(1)有唯一解的充分必要条件是(1)的系数矩阵的行列式不等于零. ▌

把定理 2.7.1 应用到 n 个方程的 n 元齐次线性方程组上便得到

推论 2.7.2 数域 K 上的 n 个方程的 n 元齐次线性方程组有非零解的充分必要条件是它的系数矩阵的行列式等于零. ▌

例 1 当 λ 取什么值时，方程组

$$\begin{cases}\lambda x_1+\ \ x_2=0\\ \ \ x_1+\lambda x_2=0\end{cases}$$

有非零解？

解 据推论 2.7.2，上述方程组有非零解的充分必要条件是

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

即，$\lambda^2-1=0$. 由此得 $\lambda=\pm 1$.

利用行列式按一行(列)展开可以得到 n 个方程的 n 元线性方程组当其系数矩阵的行列式不等于零时解的公式.

定理 2.7.2(Cramer 法则) 如果线性方程组(1)的系数矩阵 $A=(a_{ij})$ 的行列式 $|A|\neq 0$，则方程组(1)有唯一解，这个解是

$$\left(\frac{|B_1|}{|A|},\frac{|B_2|}{|A|},\cdots,\frac{|B_n|}{|A|}\right)\tag{2}$$

其中

$$B_j = \begin{bmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{i,j-1} & b_i & a_{i,j+1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{bmatrix} \tag{3}$$

$j = 1, 2, \cdots, n$.

证明 定理 2.7.1 已证明了 $|A| \neq 0$ 时，方程组(1) 有唯一解. 现在证本定理的后半部分. 为了证有序数组(2) 是方程组(1) 的解，只要把它们代入(1) 的每个方程，看是否变成恒等式. 代入第 i 个方程，计算它的左边的值：

$$\sum_{j=1}^{n} a_{ij} \frac{|B_j|}{|A|} = \frac{1}{|A|} \sum_{j=1}^{n} a_{ij} |B_j| \tag{4}$$

由于矩阵 B_j 与矩阵 A 只有第 j 列不同，其余列相同，因此，B_j 的 (k,j) 元的代数余子式等于 A 的(k,j) 元的代数余子式 A_{kj}，从而将 $|B_j|$ 按第 j 列展开得

$$|B_j| = b_1 A_{1j} + b_2 A_{2j} + \cdots + b_n A_{nj} = \sum_{k=1}^{n} b_k A_{kj} \tag{5}$$

把(5)式代入(4)式得

$$\begin{aligned} \sum_{j=1}^{n} a_{ij} \frac{|B_j|}{|A|} &= \frac{1}{|A|} \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{n} b_k A_{kj} \\ &= \frac{1}{|A|} \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ij} b_k A_{kj} = \frac{1}{|A|} \sum_{k=1}^{n} \sum_{j=1}^{n} a_{ij} b_k A_{kj} \\ &= \frac{1}{|A|} \sum_{k=1}^{n} b_k \sum_{j=1}^{n} a_{ij} A_{kj} = \frac{1}{|A|} b_i |A| = b_i \end{aligned}$$

这与第 i 个方程的右边一致. 因此(2)是方程组(1)的解. ▎

例 2 解线性方程组

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

解 先计算方程组(6)的系数矩阵 A 的行列式:

$$|A|=\begin{vmatrix} 1 & 2 & 1 \\ -2 & 1 & -1 \\ 1 & -4 & 2 \end{vmatrix}=11$$

由于 $|A|\neq 0$,因此方程组(6)有唯一解. 因为

$$\begin{vmatrix} 3 & 2 & 1 \\ -3 & 1 & -1 \\ -5 & -4 & 2 \end{vmatrix}=33 \qquad \begin{vmatrix} 1 & 3 & 1 \\ -2 & -3 & -1 \\ 1 & -5 & 2 \end{vmatrix}=11$$

$$\begin{vmatrix} 1 & 2 & 3 \\ -2 & 1 & -3 \\ 1 & -4 & -5 \end{vmatrix}=-22$$

所以方程组(6)的解是$(3,1,-2)$.

习 题 2.7

1. a,b 取什么值时,方程组

$$\begin{cases} ax_1+x_2+x_3=0 \\ x_1+bx_2+x_3=0 \\ x_1+2bx_2+x_3=0 \end{cases}$$

有非零解?

2. a,b 取什么值时,方程组

$$\begin{cases} ax_1 + x_2 + x_3 = 2 \\ x_1 + bx_2 + x_3 = 1 \\ x_1 + 2bx_2 + x_3 = 2 \end{cases}$$

有唯一解?

3. 用 Cramer 法则解下列线性方程组:

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

(2)
$$\begin{cases} x_1 + 2x_2 + 3x_3 - 3x_4 = 8 \\ 3x_1 - x_2 - 2x_3 - 3x_4 = 9 \\ 2x_1 + 2x_2 - x_3 + 2x_4 = 3 \\ 3x_1 - 3x_2 + 2x_3 + x_4 = -7 \end{cases}$$

§8 拉普拉斯(Laplace)定理

本节介绍行列式的拉普拉斯定理,这是将行列式按 k 行(列)展开,它是行列式按一行(列)展开的推广.

定义 1 在数域 K 上的 n 级矩阵 A 中任意取定 k 行、k 列($k \leqslant n$),位于这些行和列的交叉处的 k^2 个元素按原来的排法组成的 k 级矩阵称为 A 的一个 **k 级子矩阵**,它的行列式称为 A 的一个 **k 级子式**. 设从 A 中取第 $i_1, i_2, \cdots, i_k$ 行($i_1 < i_2 < \cdots < i_k$),取第 $j_1, j_2, \cdots, j_k$ 列($j_1 < j_2 < \cdots < j_k$),所得到的 k 级子矩阵记作 $A(i_1, i_2, \cdots, i_k; j_1, j_2, \cdots, j_k)$,相应的 k 级子式记作 $|A(i_1, i_2, \cdots, i_k; j_1, j_2, \cdots, j_k)|$.

例如,设

$$A = \begin{pmatrix} b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{pmatrix}$$

则

$$A(2,3;1,4) = \begin{pmatrix} b_1 & b_4 \\ c_1 & c_4 \end{pmatrix}$$

$$|A(2,3;1,4)| = \begin{vmatrix} b_1 & b_4 \\ c_1 & c_4 \end{vmatrix}$$

如果划去 A 的第 2,3 行，并且划去第 1,4 列，剩下的元素按原来的排法组成的矩阵的行列式

$$\begin{vmatrix} a_2 & a_3 \\ d_2 & d_3 \end{vmatrix}$$

称为子式 $|A(2,3;1,4)|$ 的余子式，它恰好是 A 取定第 1,4 行，第 2,3 列所组成的 2 级子式 $|A(1,4;2,3)|$. 反之 A 的 2 级子式 $|A(1,4;2,3)|$ 的余子式是 $|A(2,3;1,4)|$. 一般地，我们有

定义 2 在 n 级矩阵 A 中划去第 $i_1,\cdots,i_k$ 行($i_1<\cdots<i_k$)，划去第 $j_1,\cdots,j_k$ 列($j_1<\cdots<j_k$)，剩下的元素按照原来的排法组成的 $n-k$ 级矩阵的行列式称为 k 级子式 $|A(i_1,\cdots,i_k;j_1,\cdots,j_k)|$ 的**余子式**，它恰好是 $|A(i'_1,\cdots,i'_{n-k};j'_1,\cdots,j'_{n-k})|$，其中

$$\{i'_1,\cdots,i'_{n-k}\} = \{1,\cdots,n\}\backslash\{i_1,\cdots,i_k\}，并且\ i'_1<\cdots<i'_{n-k}$$

$$\{j'_1,\cdots,j'_{n-k}\} = \{1,\cdots,n\}\backslash\{j_1,\cdots,j_k\}，并且\ j'_1<\cdots<j'_{n-k}$$

$(-1)^{(i_1+\cdots+i_k)+(j_1+\cdots+j_k)}|A(i'_1,\cdots,i'_{n-k};j'_1,\cdots,j'_{n-k})|$ 称为 k 级子式 $|A(i_1,\cdots,i_k;j_1,\cdots,j_k)|$ 的**代数余子式**.

定理 2.8.1(Laplace) 在 n 级矩阵 $A=(a_{ij})$ 中任意取定 k 行($1\leqslant k<n$)，则这 k 行元素所组成的一切 k 级子式与它们自己的代数余子式的乘积之和等于 $|A|$. 设取定的是 A 的第 $i_1,\cdots,i_k$ 行(其中 $i_1<i_2<\cdots<i_k$)，则

$$|A| = \sum_{1\leqslant j_1<\cdots<j_n\leqslant n} |A(i_1,\cdots,i_k;j_1,\cdots,j_k)|$$
$$\cdot(-1)^{(i_1+\cdots+i_k)+(j_1+\cdots+j_k)}\cdot|A(i'_1,\cdots,i'_{n-k};j'_1,\cdots,j'_{n-k})| \quad (1)$$

其中

$$\{i'_1,\cdots,i'_{n-k}\}=\{1,\cdots,n\}\backslash\{i_1,\cdots,i_k\},\text{且 } i'_1<\cdots<i'_{n-k}$$
$$\{j'_1,\cdots,j'_{n-k}\}=\{1,\cdots,n\}\backslash\{j_1,\cdots,j_k\},\text{且 } j'_1<\cdots<j'_{n-k}$$

证明 A 中取定的 k 行元素组成的 k 级子式一共有 C_n^k 个(从 A 的 n 列中每次取出 k 列的组合数),因此(1) 的右端有 C_n^k 个乘积项. 在每个乘积项中,k 级子式 $|A(i_1,\cdots,i_k;j_1,\cdots,j_k)|$ 有 $k!$ 项;$|A(i'_1,\cdots,i'_{n-k};j'_1,\cdots,j'_{n-k})|$ 有$(n-k)!$项,这两个行列式的展开式相乘共得$k!(n-k)!$项. 因此(1) 式右端总共有$C_n^k\cdot k!(n-k)!=n!$项. 这$n!$项两两不同. 如果我们能证明(1) 式右端的每一项都是 $|A|$ 的一项,那么(1) 的右端的$n!$项的和就正好是 $|A|$. 现在任取(1) 的右端的一项

$$(-1)^{\tau(\mu_1\cdots\mu_k)}a_{i_1\mu_1}\cdots a_{i_k\mu_k}\cdot(-1)^{(i_1+\cdots+i_k)+(j_1+\cdots+j_k)}$$
$$\cdot(-1)^{\tau(\nu_1\cdots\nu_{n-k})}a_{i'_1\nu_1}\cdots a_{i'_{n-k}\nu_{n-k}} \quad (2)$$

其中 $\mu_1\cdots\mu_k$ 是 $j_1,\cdots,j_k$ 的一个 k 元排列,$\nu_1\cdots\nu_{n-k}$ 是 $j'_1,\cdots,j'_{n-k}$的一个 $n-k$ 元排列. 我们来证明(2) 是 $|A|$ 的一项. 显然,$a_{i_1\mu_1}\cdots a_{i_k\mu_k}a_{i'_1\nu_1}\cdots a_{i'_{n-k}\nu_{n-k}}$是$A$中不同行、不同列的$n$个元素的乘积,它是 $|A|$ 的一项当且仅当它带的符号为

$$(-1)^{\tau(i_1\cdots i_k i'_1\cdots i'_{n-k})+\tau(\mu_1\cdots\mu_k\nu_1\cdots\nu_{n-k})}$$
$$\overset{*}{=}(-1)^{(i_1-1)+[(i_2-1)-1]+\cdots+[(i_k-1)-(k-1)]}$$
$$\cdot(-1)^{\tau(\mu_1\cdots\mu_k)+\tau(\nu_1\cdots\nu_{n-k})+(j_1-1)+[(j_2-1)-1]+\cdots+[(j_k-1)-(k-1)]}$$
$$=(-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)-2(1+2+\cdots+k)+\tau(\mu_1\cdots\mu_k)+\tau(\nu_1\cdots\nu_{n-k})}$$
$$=(-1)^{\tau(\mu_1\cdots\mu_k)}\cdot(-1)^{(i_1+\cdots+i_k)+(j_1+\cdots+j_k)}\cdot(-1)^{\tau(\nu_1\cdots\nu_{n-k})}$$

因此(2) 是 $|A|$ 的一项. ▌

注 *: 设排列 $\mu_1\cdots\mu_k$ 经过 s 次对换变成 $j_1\cdots j_k$. 由于 $j_1\cdots j_k$

是偶排列，所以 s 与 $\mu_1\cdots\mu_k$ 有相同奇偶性. 由于在上述 s 次对换下，$\mu_1\cdots\mu_k v_1\cdots v_{n-k}$ 变成 $j_1\cdots j_k v_1\cdots v_{n-k}$，所以

$$(-1)^{\tau(\mu_1\cdots\mu_k v_1\cdots v_{n-k})}=(-1)^s\cdot(-1)^{\tau(j_1\cdots j_k v_1\cdots v_{n-k})}$$
$$=(-1)^{\tau(\mu_1\cdots\mu_k)}\cdot(-1)^{(j_1-1)+[(j_2-1)-1]+\cdots+[(j_k-1)-(k-1)]+\tau(v_1\cdots v_{n-k})}$$

例 1 证明

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{1r} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{r1} & \cdots & c_{rk} & b_{r1} & \cdots & b_{rr} \end{vmatrix}$$
$$=\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}\cdot\begin{vmatrix} b_{11} & \cdots & b_{1r} \\ \cdots & \cdots & \cdots \\ b_{r1} & \cdots & b_{rr} \end{vmatrix} \tag{4}$$

证明 把(4)式左边的行列式按前 k 行展开，这 k 行元素组成的 k 级子式只有左上角的 k 级子式可能不为零，其余 k 级子式一定包含零列，从而为零. 左上角的 k 级子式的余子式恰好是右下角的 r 级子式，并且

$$(1+\cdots+k)+(1+\cdots+k)=2(1+\cdots+k)$$

由定理 2.8.1 即得(4) 式. ▌

设

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1k} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{1r} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{r1} & \cdots & c_{rk} & b_{r1} & \cdots & b_{rr} \end{pmatrix}$$

令

$$A_1 = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} \qquad C = \begin{pmatrix} c_{11} & \cdots & c_{1k} \\ \cdots & \cdots & \cdots \\ c_{r1} & \cdots & c_{rk} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1r} \\ \cdots & \cdots & \cdots \\ b_{r1} & \cdots & b_{rr} \end{pmatrix}$$

我们可以把 A 写成

$$A = \begin{pmatrix} A_1 & 0 \\ C & B \end{pmatrix}$$

于是公式(4) 可写成

$$\begin{vmatrix} A_1 & 0 \\ C & B \end{vmatrix} = |A_1|\,|B| \tag{4$'$}$$

注意公式(4′) 中 A_1, B 都必须是方阵.

习　题　2.8

1. 设

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

写出 $|A(2,4;2,3)|$ 以及它的余子式和代数余子式.

2. 计算行列式:

(1)
$$\begin{vmatrix} 2 & 3 & 0 & 0 & 0 \\ -1 & 4 & 0 & 0 & 0 \\ 180 & 79 & 1 & 2 & 0 \\ 87 & 43 & 0 & 3 & 4 \\ 15 & 67 & 1 & 0 & 2 \end{vmatrix}$$

$$
(2)\quad \begin{vmatrix} 0 & \cdots & 0 & a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & a_{k1} & \cdots & a_{kk} \\ b_{11} & \cdots & b_{1r} & c_{11} & \cdots & c_{1k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{r1} & \cdots & b_{rr} & c_{r1} & \cdots & c_{rk} \end{vmatrix}
$$

§9 行列式的几何意义

行列式有许多应用. §7 我们利用行列式讨论了 n 个方程的 n 元线性方程组的解的情况. 在解析几何中，我们在好几处运用了二阶行列式和三阶行列式. 行列式在线性代数中有很多应用，这些在本书的以后各章中会陆续看到. 此外，本书还在第四章 §8 及习题 4.8 中介绍了如何运用行列式证明一些著名的恒等式；并且指出了行列式在图论中的一个重要应用. 在第四章补充题四的第 5,9 题中还可看到行列式在整数理论中的应用. 读者在其他课程中也会看到行列式是一个有用的工具.

这一节我们着重介绍一下行列式的几何意义. 任给一个实数域上的 3 级矩阵 A 的行列式

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

在几何空间中取定一个直角坐标系 $[O;\vec{e}_1,\vec{e}_2,\vec{e}_3]$，设向量 $\vec{a},\vec{b},\vec{c}$ 的坐标分别为 (a_1,a_2,a_3)，(b_1,b_2,b_3)，(c_1,c_2,c_3). 以 $\vec{a},\vec{b},\vec{c}$ 为棱的平行六面体的体积为

$$
V = |\vec{a}\times\vec{b}|\cdot|\vec{c}|\,|\cos\langle\vec{c},\vec{a}\times\vec{b}\rangle| = |\vec{a}\times\vec{b}\cdot\vec{c}| \qquad (1)
$$

我们把$\vec{a}\times\vec{b}\cdot\vec{c}$叫做以$\vec{a},\vec{b},\vec{c}$为棱的平行六面体的定向体积. 现在我们来计算$\vec{a}\times\vec{b}\cdot\vec{c}$. 因为

$$\vec{a}\times\vec{b}=\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}\vec{e}_1-\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}\vec{e}_2+\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}\vec{e}_3$$

所以

$$\begin{aligned}\vec{a}\times\vec{b}\cdot\vec{c}&=\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}c_1-\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}c_2+\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}c_3\\ &=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}\end{aligned} \tag{2}$$

从(2)式看到,实数域上的3级矩阵A的行列式等于以$\vec{a},\vec{b},\vec{c}$为棱的平行六面体的定向体积,其中$\vec{a},\vec{b},\vec{c}$的直角坐标分别为A的第1,2,3列. 这就是3阶行列式的几何意义.

一般地,n阶行列式的几何意义是什么?本书第十二章的习题12.2中将作介绍.

补 充 题 二

1. 设$n\geqslant 2$,证明:如果n级矩阵A的元素为1或-1(这种矩阵称为$(1,-1)$矩阵),则$|A|$一定是偶数.

2. 求3级$(1,-1)$矩阵的行列式可取到的最大值.

3. 元素只取0或1的矩阵称为$(0,1)$矩阵. 求3级$(0,1)$矩阵的行列式的最大值.

4. 计算下列n级行列式:

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

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

(3)
$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1+1 & x_2+1 & \cdots & x_n+1 \\ x_1^2+x_1 & x_2^2+x_2 & \cdots & x_n^2+x_n \\ x_1^3+x_1^2 & x_2^3+x_2^2 & \cdots & x_n^3+x_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ x_1^{n-1}+x_1^{n-2} & x_2^{n-1}+x_2^{n-2} & \cdots & x_n^{n-1}+x_n^{n-2} \end{vmatrix}$$

5. 设数域 K 上的 n 级矩阵 $A=(a_{ij})$，它的 (i,j) 元的代数余子式记作 A_{ij}. 把 A 的每个元素都加上同一个数 t，得到的矩阵记作 $A(t)=(a_{ij}+t)$. 证明

$$|A(t)| = |A| + t\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}$$

6. 用第 5 题给出的计算行列式 $|A(t)|$ 的公式，计算下列行列式：

(1)
$$\begin{vmatrix} 1+x_1y_1 & 1+x_1y_2 & \cdots & 1+x_1y_n \\ 1+x_2y_1 & 1+x_2y_2 & \cdots & 1+x_2y_n \\ \cdots & \cdots & \cdots & \cdots \\ 1+x_ny_1 & 1+x_ny_2 & \cdots & 1+x_ny_n \end{vmatrix}$$

(2)

$$\begin{vmatrix} 1+t & t & t & \cdots & t \\ t & 2+t & t & \cdots & t \\ t & t & 3+t & \cdots & t \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ t & t & t & \cdots & n+t \end{vmatrix}$$

第三章　n 维向量空间·线性方程组的理论

§1　引　　言

第一章中我们把线性方程组通过初等变换化成阶梯形方程组，从阶梯形方程组去判断原方程组的解的情况. 在许多问题中希望直接从原方程组判断其解的情况. 为此，在第二章我们介绍了 n 级矩阵的行列式的概念和性质，利用线性方程组(方程个数与未知量个数相等的方程组) 的系数矩阵的行列式讨论其解的情况，这是直接从原方程组的系数去讨论方程组的解. 但是，即便是对于 n 个方程的 n 元线性方程组，用系数矩阵 A 的行列式去讨论其解，只是在 $|A|\neq 0$ 时，给出了确定的答案:原方程组有唯一解；而在 $|A|=0$时，只知道原方程组或者无解,或者有无穷多个解,并没有回答什么时候无解? 什么时候有无穷多个解?至于方程个数与未知量个数不相等的线性方程组，其系数矩阵没有行列式的概念. 此外，阶梯形方程组中不包括“$0=0$” 的方程的个数是不是由原线性方程组唯一决定呢?这个问题也没有解决. 这些问题促使我们应当学习新的理论来判断线性方程组的解的情况，这正是本章所要讲的内容.

在解线性方程组时，我们要对其增广矩阵作初等行变换. 1°型的初等行变换是把矩阵的一行的倍数加到另一行，譬如说，把矩阵$A=(a_{ij})$ 的第i行的l倍加到第k行，这里包括两个步骤：第一步,“第i行的l倍”,这意思是第i行的每个元素乘以l. 如果我们把 A 的第 i 行写成$(a_{i1},a_{i2},\cdots,a_{in})$,则“第 i 行的 l 倍” 可以写成

$$l \cdot (a_{i1}, a_{i2}, \cdots, a_{in}) = (la_{i1}, la_{i2}, \cdots, la_{in})$$

这实际上规定了数与 n 元有序数组的乘法，称它为数量乘法，即数 l 与一个 n 元有序数组相乘等于 l 乘该有序数组的每个元素后组成的 n 元有序数组. 第二步，第 i 行乘以 l 以后，再加到第 k 行，这时是把这两个有序数组的对应元素相加，即

$$\begin{aligned}&(la_{i1}, la_{i2}, \cdots, la_{in}) + (a_{k1}, a_{k2}, \cdots, a_{kn})\\ =&(la_{i1} + a_{k1}, la_{i2} + a_{k2}, \cdots, la_{in} + a_{kn})\end{aligned}$$

这实际上规定了两个 n 元有序数组的加法：对应元素相加得到一个 n 元有序数组. 这些启发我们可以在数域 K 上的所有 n 元有序数组组成的集合中规定加法运算，还可以规定数域 K 与上述有序数组集合之间的数量乘法运算. 这两种运算的引进，将给 n 元有序数组的集合带来崭新的面貌，进而为讨论线性方程组的解的情况开辟一条新的途径.

§2　n 维向量空间 K^n 及其线性子空间

设 n 是任意给定的一个自然数. 取定一个数域 K.

定义 1　数域 K 上的两个 n 元有序数组 $(a_1, a_2, \cdots, a_n)$ 与 $(b_1, b_2, \cdots, b_n)$ 称为**相等**，如果

$$a_i = b_i, \qquad i = 1, 2, \cdots, n$$

数域 K 上所有 n 元有序数组组成的集合记作 K^n，它的元素用小写希腊字母 $\alpha, \beta, \gamma, \cdots$ 表示.

在 K^n 中规定**加法**运算如下：

定义 2　设 $\alpha = (a_1, a_2, \cdots, a_n)$，$\beta = (b_1, b_2, \cdots, b_n)$，则 n 元有序数组 $(a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n)$ 称为 α 与 β 的**和**，记作 $\alpha + \beta$. 即

$$\alpha + \beta := (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n) \tag{1}$$

在数域 K 与集合 K^n 之间规定**数量乘法**运算如下：

定义 3　设 $k \in K, \alpha = (a_1, a_2, \cdots, a_n)$，则 n 元有序数组 $(ka_1, ka_2, \cdots, ka_n)$ 称为 k 与 α 的**数量乘积**，记作 $k\alpha$. 即

$$k\alpha := (ka_1, ka_2, \cdots, ka_n) \tag{2}$$

容易直接验证加法和数量乘法满足下述八条规则：

1°　加法**交换律**，即 $\forall\ \alpha, \beta \in K^n$，有

$$\alpha + \beta = \beta + \alpha$$

2°　加法**结合律**，即 $\forall\ \alpha, \beta, \gamma \in K^n$，有

$$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$$

3°　K^n 中有一个元素 $(0, 0, \cdots, 0)$，记作 0，它使得

$$\alpha + 0 = 0 + \alpha = \alpha$$

$\forall\ \alpha \in K^n$. 这个元素 0 称为**零元素**；

4°　对于 K^n 中每个元素 $\alpha = (a_1, a_2, \cdots, a_n)$，在 K^n 中有一个元素 $(-a_1, -a_2, \cdots, -a_n)$，使得

$$(a_1, \cdots, a_n) + (-a_1, \cdots, -a_n) = 0$$

把 $(-a_1, -a_2, \cdots, -a_n)$ 记成 $-\alpha$，上式即为

$$\alpha + (-\alpha) = 0$$

称 $-\alpha$ 是 α 的**负元素**；

5°　$\forall\ \alpha \in K^n$，有

$$1\alpha = \alpha$$

6°　$\forall\ k, l \in K, \alpha \in K^n$，有

$$(kl)\alpha = k(l\alpha)$$

7°　$\forall\ k, l \in K, \alpha \in K^n$，有

$$(k + l)\alpha = k\alpha + l\alpha$$

8°　$\forall\ k \in K, \alpha, \beta \in K^n$，有

$$k(\alpha + \beta) = k\alpha + k\beta$$

定义 4　数域 K 上的所有 n 元有序数组组成的集合 K^n，连同定义在它们上面的加法运算和数量乘法运算，称为数域 K 上的一个 n **维向量空间**. K^n 的元素称为 n 维向量，简称为**向量**；设向量 $\alpha = (a_1, a_2, \cdots, a_n)$，称 a_i 是 α 的第 i 个**分量**.

上述定义中“n 维向量空间”的**维数**的精确含义在本章 §4 将给出.

从向量空间关于加法和数量乘法的上述八条基本运算法则还可以推导出其他一些性质，例如，我们可以定义向量的减法如下：

$$\alpha-\beta := \alpha+(-\beta)$$

又如，

$$\begin{aligned} 0\alpha &= 0 \\ (-1)\alpha &= -\alpha \\ k0 &= 0 \\ k\alpha &= 0 \quad \Rightarrow \quad k=0 \text{ 或者 } \alpha=0 \end{aligned}$$

这四条性质从加法和数量乘法的定义是很容易得到的. 但是它们也可以不用加法和数量乘法的定义而从前面的八条基本运算法则推导出，建议有兴趣的读者不妨试一试.

上面考虑的 n 维向量空间 K^n，其向量写成一行，称为**行向量**. 我们也可以把每一个 n 元有序数组写成一列，并且同样定义有序数组之间的加法运算，以及数与有序数组之间的数量乘法运算，这样也得到数域 K 上的一个 n 维向量空间，其元素称为**列向量**. 这两个向量空间之间没有本质的区别，只是元素的写法不同，因此都记成 K^n. 在具体场合下，我们谈的究竟是行向量还是列向量，可以从上下文中看出.

实数域上的 n 维向量空间简称为 n 维实向量空间.

从解析几何知道，几何空间中取定了一个坐标系后，全体向量的集合与全体有序三元实数组的集合之间有一个一一对应：每个向量对应于它的坐标；并且 $\vec{a}+\vec{b}$ 的坐标等于 $\vec{a}$ 的坐标与 $\vec{b}$ 的坐标之和，$k\vec{a}$ 的坐标等于 k 乘以 $\vec{a}$ 的坐标. 因此 3 维实向量空间 R^3 可以看成是所有以原点为起点的向量组成的几何空间. 这使我们常常可以从几何直观中受到启发. 注意：几何空间既可以看成是所有点组成的集合，又可以看成是以原点为起点的所有向量组成的集合（在这两个集合之间有一个一一对应）. 今后我们说几何空间时，

通常指的是后者.同理,一个平面π(或一条直线L)既可以看成是π上(或L上)所有点组成的集合,又可以看成是以原点为起点,终点在π上(或L上)的所有向量组成的集合.今后当我们说平面π(或直线L)是几何空间R^3的子集时,指的是后者.

设π是通过原点的一个平面,它可以看成是R^3的一个子集.从向量的加法的平行四边形法则,以及实数与向量的数量乘法的定义,可以看出:若$\alpha,\beta\in\pi$,则$\alpha+\beta\in\pi$;若$\lambda\in R$,$\alpha\in\pi$,,则$\lambda\alpha\in\pi$.于是向量的加法,以及实数与向量的数量乘法,可分别看成是集合π的加法运算,以及R与π的数量乘法运算.由于这个原因,我们称π是R^3的一个线性子空间.从这个几何直观受到启发,我们引进下述概念:

定义 5　设W是数域K上n维向量空间K^n的一个非空子集,如果它满足下面两条性质:

1)　$\alpha,\beta\in W \quad\Rightarrow\quad \alpha+\beta\in W$,

2)　$k\in K,\alpha\in W \quad\Rightarrow\quad k\alpha\in W$,

则称W是K^n的一个**线性子空间**,简称为子空间.

通常把性质1)称为W对于加法封闭,把性质2)称为W对于数量乘法封闭.

由性质2)可以推出,零向量属于K^n的任一线性子空间.

例 1　$\{0\}$是K^n的一个线性子空间,这是因为$0+0=0\in\{0\}$,$k0=0\in\{0\}$.$\{0\}$称为K^n的零子空间,记成0.

例 2　K^n本身是K^n的一个线性子空间.

例 3　$\{(a_1,\cdots,a_{n-1},0)\mid a_i\in K,\quad i=1,\cdots,n-1\}$是$K^n$的一个线性子空间,请读者利用定义5验证它.

例 4　设π是不经过原点的一个平面,则它是R^3的一个子集,但是它不是R^3的一个线性子空间,这是因为它对于加法和数量乘法都不封闭.　?

设W是K^n的一个非零线性子空间,则W含有无限多个向量.理由是:在W中取一个非零向量α,则对于K中任一数k,有$k\alpha$

$\in W$. 于是 W 含有无限多个向量.

从解析几何知道，几何空间中取定三个不共面的向量 $\vec{e}_1,\vec{e}_2,\vec{e}_3$，则每一个向量 $\vec{a}$ 可以写成 $\vec{a}=k_1\vec{e}_1+k_2\vec{e}_2+k_3\vec{e}_3$ 的形式，因此

$$R^3=\{k_1\vec{e}_1+k_2\vec{e}_2+k_3\vec{e}_3 \mid k_i\in R,\quad i=1,2,3\}$$

设 π 是通过原点的一个平面，在 π 中取定两个不共线的向量 $\vec{a},\vec{b}$，则 π 中每一个向量可以写成 $k_1\vec{a}+k_2\vec{b}$ 的形式，因此

$$\pi=\{k_1\vec{a}+k_2\vec{b} \mid k_i\in R,\quad i=1,2\}$$

现在我们把上述形成线性子空间的方法加以推广：

设 $\alpha_1,\cdots,\alpha_s\in K^n$，其中 s 是任意给定的一个自然数. 对于 K 中一组数 $k_1,k_2,\cdots,k_s$，K^n 中的向量 $k_1\alpha_1+\cdots+k_s\alpha_s$ 称为向量组 $\alpha_1,\cdots,\alpha_s$ 的一个**线性组合**，$k_1,\cdots,k_s$ 称为这个线性组合的**系数**. 考虑向量组 $\alpha_1,\cdots,\alpha_s$ 的所有线性组合组成的集合

$$W=\{k_1\alpha_1+\cdots+k_s\alpha_s \mid k_i\in K,\quad i=1,\cdots,s\}$$

因为 $0=0\alpha_1+\cdots+0\alpha_s\in W$，所以 W 是 K^n 的一个非空子集. 因为

$$\begin{aligned}&(k_1\alpha_1+\cdots+k_s\alpha_s)+(l_1\alpha_1+\cdots+l_s\alpha_s)\\ =&(k_1+l_1)\alpha_1+\cdots+(k_s+l_s)\alpha_s\in W\end{aligned}$$

$$l(k_1\alpha_1+\cdots+k_s\alpha_s)=(lk_1)\alpha_1+\cdots+(lk_s)\alpha_s\in W$$

所以 W 是 K^n 的一个线性子空间，称它是**由** $\alpha_1,\cdots,\alpha_s$ **生成**（或**张成**）**的线性子空间**，记成 $\langle\alpha_1,\cdots,\alpha_s\rangle$. 即

$$\langle\alpha_1,\cdots,\alpha_s\rangle=\{k_1\alpha_1+\cdots+k_s\alpha_s \mid k_i\in K,\quad i=1,\cdots,s\} \quad (3)$$

定义 6 对于 K^n 中的向量组 $\alpha_1,\cdots,\alpha_s$，以及向量 β，如果有 K 中一组数 $k_1,\cdots,k_s$ 使得

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \tag{4}$$

则称 β 可以由向量组 $\alpha_1,\cdots,\alpha_s$ **线性表出**.

因为向量组 $\alpha_1,\cdots,\alpha_s$ 的所有线性组合都在 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 中,所以

$$\beta \text{ 可以由向量组 } \alpha_1,\cdots,\alpha_s \text{ 线性表出} \Leftrightarrow \beta \in \langle\alpha_1,\cdots,\alpha_s\rangle \tag{5}$$

设 $A=(a_{ij})$ 是数域 K 上的一个 $s\times n$ 矩阵. A 的每一行是一个 n 元有序数组,因此它是 K^n 中的一个向量,称 A 的每一行是 A 的一个**行向量**. A 的每一列是一个 s 元有序数组,因此它是 K^s 中的一个向量. 称 A 的每一列是 A 的一个**列向量**. A 的行向量组生成的线性子空间(K^n 的线性子空间)称为 A 的**行空间**. A 的列向量组生成的线性子空间(K^s 的线性子空间)称为 A 的**列空间**.

运用上述概念,我们来看线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = b_s \end{cases} \tag{6}$$

有解的充要条件. 设方程组(6)的系数矩阵 A 的列向量组是 $\alpha_1,\alpha_2,\cdots,\alpha_n$;增广矩阵 $\bar{A}$ 的列向量组是 $\alpha_1,\cdots,\alpha_n,\beta$,则线性方程组(6)可以写成

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{s1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{s2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{sn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_s \end{pmatrix}$$

即

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta \tag{6$'$}$$

由此得出

$$
\begin{aligned}
& n\text{ 元线性方程组(6) 有解} \\
\Leftrightarrow & \text{存在 } K \text{ 中一组数 } c_1, c_2, \cdots, c_n\text{，使得} \\
& c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = \beta \\
\Leftrightarrow & \beta \text{ 可以由 } \alpha_1, \alpha_2, \cdots, \alpha_n \text{ 线性表出} \qquad (7) \\
\Leftrightarrow & \beta \in \langle \alpha_1, \alpha_2, \cdots, \alpha_n \rangle \qquad (8)
\end{aligned}
$$

当方程组(6) 有解时，若它有唯一解，则 β 表示成 $\alpha_1, \cdots, \alpha_n$ 的线性组合只有一种方式，简称为表法唯一. 若(6) 有无穷多个解，则 β 表示成 $\alpha_1, \cdots, \alpha_n$ 的线性组合有无穷多种方式(以方程组(6) 的每一个解 $(c_1, \cdots, c_n)$ 作为 $\alpha_1, \cdots, \alpha_n$ 的线性组合的系数就是 β 的一种表示方式). 因此我们得到

$$
\begin{aligned}
& \text{线性方程组(6) 有唯一解} \\
\Leftrightarrow & \beta \text{ 可以由 } \alpha_1, \cdots, \alpha_n \text{ 线性表出并且表法唯一} \\
& \text{线性方程组(6) 有无穷多个解} \\
\Leftrightarrow & \beta \text{ 可以由 } \alpha_1, \cdots, \alpha_n \text{ 线性表出且表法不唯一}
\end{aligned}
$$

这样我们把 n 元线性方程组有没有解的问题归结为 n 维向量空间中向量之间的关系问题. 这是我们第一次遇到需要研究 $n > 3$ 的 n 维向量空间. 现实的几何空间只是 3 维实向量空间. 研究线性方程组的有无解问题和解的结构问题促使我们去研究 n 为任意自然数的 n 维向量空间. 我们将在 §3 和 §4 研究 n 维向量空间及其线性子空间的结构.

如果给定了 K^s 中一个向量组 $\alpha_1, \cdots, \alpha_n$ 以及一个向量 β，问 β 能否由 $\alpha_1, \cdots, \alpha_n$ 线性表出，那么我们可以利用(7)，即

$$
\begin{aligned}
& \beta \text{ 可以由向量组 } \alpha_1, \cdots, \alpha_n \text{ 线性表出} \\
\Leftrightarrow & \text{线性方程组 } x_1\alpha_1 + \cdots + x_n\alpha_n = \beta \text{ 有解} \qquad (9)
\end{aligned}
$$

例 5　K^4 中，判断向量 β 能否由向量组 $\alpha_1, \alpha_2, \alpha_3$ 线性表出，若能够，写出它的一种表达方式，其中

$$\beta=(2,-1,3,4),\qquad \alpha_1=(1,2,-3,1)$$
$$\alpha_2=(5,-5,12,11),\qquad \alpha_3=(1,-3,6,3)$$

解 据(9),应该考虑线性方程组

$$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=\beta \tag{10}$$

它的增广矩阵是把 $\alpha_1,\alpha_2,\alpha_3,\beta$ 都写成列向量以后组成的,对增广矩阵作初等行变换化成简化行阶梯形矩阵:

$$\begin{pmatrix} 1 & 5 & 1 & 2 \\ 2 & -5 & -3 & -1 \\ -3 & 12 & 6 & 3 \\ 1 & 11 & 3 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -\frac{2}{3} & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

于是方程组(10)有无穷多个解,它的一般解是

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

其中 x_3 是自由未知量.因此 β 可以由 $\alpha_1,\alpha_2,\alpha_3$ 线性表出,而且表法不唯一.取 $x_3=1$,得 $x_1=1,x_2=0$.于是得到一种表出方式:

$$\beta=\alpha_1+\alpha_3$$

若取 $x_3=4$,则得 $x_1=3,x_2=-1$ 于是又得到另一种表出方式:

$$\beta=3\alpha_1-\alpha_2+4\alpha_3$$

由于方程组(10)有无穷多个解,因此 β 由 $\alpha_1,\alpha_2,\alpha_3$ 线性表出的方式有无穷多种.

习 题 3.2

1. 在 K^4 中,设 $\alpha_1=(-2,0,1,3)$,$\alpha_2=(5,-7,4,6)$,$\alpha_3=(3,-1,7,2)$,求 $3\alpha_1+\alpha_2+2\alpha_3$.

2. 在 K^4 中,设 $\alpha=(6,-2,0,4)$,$\beta=(-3,1,5,7)$.求向量 γ,使得 $2\alpha+\gamma=3\beta$.

3. 在 K^4 中,判断向量 β 能否由 $\alpha_1,\alpha_2,\alpha_3$ 线性表出.若能,写出它的一种表出方式.

(1)

$$\beta=(8,3,-1,-25)\qquad \alpha_1=(-1,3,0,-5)$$
$$\alpha_2=(2,0,7,-3)\qquad \alpha_3=(-4,1,-2,6)$$

(2)

$$\beta=(-8,-3,7,-10)\qquad \alpha_1=(-2,7,1,3)$$
$$\alpha_2=(3,-5,0,-2)\qquad \alpha_3=(-5,-6,3,-1)$$

(3)

$$\beta=(2,-30,13,-26)\qquad \alpha_1=(3,-5,2,-4)$$
$$\alpha_2=(-1,7,-3,6)\qquad \alpha_3=(3,11,-5,10)$$

4. 在 K^n 中,设 $\varepsilon_1=(1,0,\cdots,0)$,$\varepsilon_2=(0,1,0,\cdots,0)$,$\cdots$,$\varepsilon_n=(0,0,\cdots,0,1)$.试证:$K^n$ 中任一向量都可以由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性表出,并且表法唯一,写出这种表出方式.

5. 试证:K^4 中任一向量都可以由向量组

$$\alpha_1=(1,0,0,0)\qquad \alpha_2=(1,1,0,0)$$
$$\alpha_3=(1,1,1,0)\qquad \alpha_4=(1,1,1,1)$$

线性表出,并且表法唯一,写出这种表出方式.

6. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s\in K^n$,说明 $\alpha_i\in\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$, $i=1,2,\cdots,s$.

7. 设 L_1,L_2 分别是通过原点,不通过原点的一条直线,问:L_1,L_2 是不是 R^3 的线性子空间?

8. 设 $r<n$,证明 K^n 的下述子集

$$W = \{(a_1, \cdots, a_r, 0, \cdots, 0) \mid a_i \in K, \quad i = 1, \cdots, r\}$$

是一个线性子空间.

9. 证明:如果线性方程组(I)的增广矩阵的第i个行向量可以由其余行向量线性表出,那么把方程组(I)的第i个方程去掉以后得到的方程组(II)与方程组(I)同解.

§3 线性相关的向量组与线性无关的向量组

上一节我们指出,研究线性方程组有无解的问题以及解的结构问题促使我们去研究n维向量空间及其线性子空间的结构.

先分析几何空间R^3及其线性子空间的结构,以便从中受到启发. R^3中只要取定三个不共面的向量$\vec{e}_1, \vec{e}_2, \vec{e}_3$,则

$$R^3 = \{x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3 \mid x, y, z \in R\}$$

于是R^3的结构就清楚了. 其次考虑R^3的一个线性子空间,譬如,通过原点的一个平面π. 在π上取定两个不共线的向量$\vec{a}, \vec{b}$,则

$$\pi = \{k_1\vec{a} + k_2\vec{b} \mid k_1, k_2 \in R\}$$

这样π的结构也就清楚了. 为了把这些推广到研究n维向量空间及其线性子空间的结构,就需要弄清楚几何空间中"不共面"、"不共线"向量的代数本质是什么?我们在解析几何中已讲了:

三个向量$\vec{a}, \vec{b}, \vec{c}$不共面

$\Leftrightarrow$从$k_1\vec{a} + k_2\vec{b} + k_3\vec{c} = \vec{0}$可以推出$k_1 = k_2 = k_3 = 0$

两个向量$\vec{a}, \vec{b}$不共线

$\Leftrightarrow$从$k_1\vec{a} + k_2\vec{b} = \vec{0}$可以推出$k_1 = k_2 = 0$

换句话说,

三个向量$\vec{a}, \vec{b}, \vec{c}$共面

$\Leftrightarrow$有不全为零的实数k_1, k_2, k_3,使

$$k_1\vec{a}+k_2\vec{b}+k_3\vec{c}=\vec{0}$$

类似地可写出两个向量共线的充分必要条件.

抓住了"不共面"、"不共线"、"共面"、"共线"向量的本质，就可以在K^n中引入相应的概念，这些概念在研究向量空间的结构中起着关键的作用.

定义 1 K^n 中的向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 1)$ 称为是**线性相关**的，如果有 K 中不全为零的数 $k_1,k_2,\cdots,k_s$，使

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0 \tag{1}$$

定义 2 K^n 中的向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 1)$ 如果不是线性相关的，则称为**线性无关**的. 换句话说，向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 1)$ 称为线性无关，如果从

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0$$

可以推出

$$k_1=k_2=\cdots=k_s=0$$

从定义 1 和定义 2 可以看出线性相关的向量组与线性无关的向量组是性质相反的两类向量组：

向量组 $\alpha_1,\cdots,\alpha_s(s\geqslant 1)$ 线性相关

$\Leftrightarrow\alpha_1,\cdots,\alpha_s$ 有系数不全为零的线性组合等于 0

$\Leftrightarrow$有 K 中不全为 0 的数 $k_1,\cdots,k_s$ 使(1)式成立

这时(1)式称为向量组 $\alpha_1,\cdots,\alpha_s$ 的**一个线性关系**.

向量组 $\alpha_1,\cdots,\alpha_s(s\geqslant 1)$ 线性无关

$\Leftrightarrow\alpha_1,\cdots,\alpha_s$ 没有系数不全为 0 的线性组合等于 0

$\Leftrightarrow\alpha_1,\cdots,\alpha_s$ 只有系数全为 0 的线性组合才等于 0

$\Leftrightarrow$从 $k_1\alpha_1+\cdots+k_s\alpha_s=0$ 可以推出 $k_1=\cdots=k_s=0$

例 1 包含零向量的向量组一定线性相关.

证明 考虑向量组 $0,\alpha_2,\cdots,\alpha_s$. 因为

$$1\cdot 0+0\alpha_2+\cdots+0\alpha_s=0$$

这里 $1,0,\cdots,0$ 不全为零，所以 $0,\alpha_2,\cdots,\alpha_s$ 线性相关. ▍

例 2 一个向量 α 组成的向量组线性无关的充分必要条件是 $\alpha\neq 0$.

证明 必要性. 用反证法，假如 $\alpha=0$，则据例 1 知向量组 α 线性相关.

充分性. 设 $\alpha\neq 0$，要证向量组 α 线性无关. 根据定义 2，要从 $k\alpha=0$ 去推出 $k=0$. 为此，设 $k\alpha=0$. 因为 $\alpha\neq 0$，并且根据 §2 列举的 n 维向量空间的一条性质："$k\alpha=0\Rightarrow k=0$ 或者 $\alpha=0$，"，所以得到 $k=0$. 因此向量组 α 线性无关. ▍

例 3 K^n 中，设 $\varepsilon_1=(1,0,0,\cdots,0)$，$\varepsilon_2=(0,1,0,\cdots,0)$，$\cdots$，$\varepsilon_n=(0,0,0,\cdots,1)$. 则向量组 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性无关.

证明 设 $k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n=0$，则得到

$$k_1\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}+k_2\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}+\cdots+k_n\begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\\\vdots\\0\end{pmatrix}$$

由上式得到：$k_1=0,k_2=0,\cdots,k_n=0$. 因此向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关. ▍

根据本节开头所说的可以看出，几何空间 R^3 中，三个向量 $\vec{a}$，$\vec{b}$，$\vec{c}$ 线性相关，就是它们共面；线性无关就是它们不共面. 两个向量 $\vec{a}$，$\vec{b}$ 线性相关就是它们共线；线性无关就是它们不共线.

下面我们从另一个角度来看线性相关的向量组与线性无关的向量组的本质区别.

命题 3.3.1 K^n 中，向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 2)$ 线性相关的充分必要条件是：其中至少有一个向量可以由其余向量线性表出.

证明 必要性. 设向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 2)$ 线性相关，则据定义 1，有数域 K 中不全为零的数 $k_1,k_2,\cdots,k_s$，使得 $k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0$. 设 $k_i\neq 0$，则得

$$\alpha_i = -\frac{k_1}{k_i}\alpha_1 - \cdots - \frac{k_{i-1}}{k_i}\alpha_{i-1} - \frac{k_{i+1}}{k_i}\alpha_{i+1} - \cdots - \frac{k_s}{k_i}\alpha_s$$

充分性. 设向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 2)$ 中有一个向量 α_j 可以由其余向量线性表出,即

$$\alpha_j = l_1\alpha_1 + \cdots + l_{j-1}\alpha_{j-1} + l_{j+1}\alpha_{j+1} + \cdots + l_s\alpha_s$$

则

$$l_1\alpha_1 + \cdots + l_{j-1}\alpha_{j-1} - \alpha_j + l_{j+1}\alpha_{j+1} + \cdots + l_s\alpha_s = 0$$

系数中有 -1,因此 $\alpha_1,\cdots,\alpha_{j-1},\alpha_j,\alpha_{j+1},\cdots\alpha_s$ 线性相关. ▌

推论 3.3.1　K^n 中,向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 2)$ 线性无关的充分必要条件是:其中每一个向量都不能由其余向量线性表出. ▌

命题 3.3.1 和推论 3.3.1 使我们对线性相关与线性无关的含义有进一步的理解.

下面的命题 3.3.2 给出了线性相关的向量组与线性无关的向量组的第三个区别:

命题 3.3.2　K^n 中,假设向量 β 可以由向量组 $\alpha_1,\cdots,\alpha_s$ 线性表出. 如果 $\alpha_1,\cdots,\alpha_s$ 线性无关,则 β 由 $\alpha_1,\cdots,\alpha_s$ 线性表出的方式唯一;如果 $\alpha_1,\cdots,\alpha_s$ 线性相关,则 β 由 $\alpha_1,\cdots,\alpha_s$ 线性表出的方式不唯一. 换句话说,β 由 $\alpha_1,\cdots,\alpha_s$ 线性表出的方式唯一的充分必要条件是 $\alpha_1,\cdots,\alpha_s$ 线性无关.

证明　首先设 $\alpha_1,\cdots,\alpha_s$ 线性无关. 设

$$\beta = k_1\alpha_1 + \cdots + k_s\alpha_s \qquad \beta = l_1\alpha_1 + \cdots + l_s\alpha_s$$

则得

$$k_1\alpha_1 + \cdots + k_s\alpha_s = l_1\alpha_1 + \cdots + l_s\alpha_s$$

从而有

$$(k_1 - l_1)\alpha_1 + \cdots + (k_s - l_s)\alpha_s = 0$$

因为 $\alpha_1,\cdots,\alpha_s$ 线性无关,由上式得

$$k_1 - l_1 = \cdots = k_s - l_s = 0$$

由此得

$$k_1 = l_1,\quad k_2 = l_2,\quad \cdots,\quad k_s = l_s$$

这说明 β 由 $\alpha_1,\cdots,\alpha_s$ 线性表出的方式唯一.

其次设 $\alpha_1,\cdots,\alpha_s$ 线性相关,则有 K 中不全为 0 的数 $k_1,\cdots,k_s$,使得

$$k_1\alpha_1+\cdots+k_s\alpha_s=0 \tag{2}$$

设

$$\beta=l_1\alpha_1+l_2\alpha_2+\cdots+l_s\alpha_s \tag{3}$$

从(2)和(3)得

$$\beta=(k_1+l_1)\alpha_1+\cdots+(k_s+l_s)\alpha_s \tag{4}$$

因为 $k_1,\cdots,k_s$ 不全为零,所以 $(k_1+l_1,\cdots,k_s+l_s)\neq(l_1,\cdots,l_s)$. 因此(3)与(4)给出了 β 由 $\alpha_1,\cdots,\alpha_s$ 线性表出的两种方式. ▌

往下我们总是取定一个数域 K,不再每次说明.

判断一个向量组是线性相关还是线性无关,最基本也是最重要的方法是根据线性相关的定义或线性无关的定义. 前面已经举了几个例子,现在再看两个例子.

例 4　若向量组的一个部分组线性相关,则整个向量组也线性相关.

证明　设向量组 $\alpha_1,\cdots,\alpha_t,\alpha_{t+1},\cdots,\alpha_s$ 的一个部分组,譬如说,$\alpha_1,\cdots,\alpha_t$ 线性相关,则有 K 中不全为 0 的数 $k_1,\cdots,k_t$,使得

$$k_1\alpha_1+\cdots+k_t\alpha_t=0$$

从而有

$$k_1\alpha_1+\cdots+k_t\alpha_t+0\alpha_{t+1}+\cdots+0\alpha_s=0$$

因为 $k_1,\cdots,k_t$ 不全为零,所以 $k_1,\cdots,k_t,0,\cdots,0$ 不全为零. 于是 $\alpha,\cdots,\alpha_t,\alpha_{t+1},\cdots,\alpha_s$ 线性相关. ▌

由例 4 立即得到:如果向量组线性无关,则它的任何一个部分组也线性无关.

例 5　试证:如果向量组 $\alpha_1,\alpha_2,\alpha_3$ 线性无关,则向量组 $2\alpha_1+\alpha_2,\alpha_2+5\alpha_3,4\alpha_3+3\alpha_1$ 也线性无关.

证明　设

$$k_1(2\alpha_1+\alpha_2)+k_2(\alpha_2+5\alpha_3)+k_3(4\alpha_3+3\alpha_1)=0$$

整理得

$$(2k_1 + 3k_3)\alpha_1 + (k_1 + k_2)\alpha_2 + (5k_2 + 4k_3)\alpha_3 = 0$$

因为 $\alpha_1,\alpha_2,\alpha_3$ 线性无关，由上式得

$$\begin{cases} 2k_1 \qquad\qquad + 3k_3 = 0 \\ \ \ k_1 + k_2 \qquad\qquad = 0 \\ \qquad\quad 5k_2 + 4k_3 = 0 \end{cases}$$

解这个线性方程组得

$$k_1 = 0,\quad k_2 = 0,\quad k_3 = 0$$

因此 $2\alpha_1 + \alpha_2,,\alpha_2 + 5\alpha_3,,4\alpha_3 + 3\alpha_1$ 线性无关.

利用 K^n 中的向量是 n 元有序数组这一特点，还可以采用下述方法判断一个向量组是否线性相关.

向量组 $\alpha_1,\cdots,\alpha_s$ 线性相关

$\Leftrightarrow$有 K 中不全为零的数 $k_1,\cdots,k_s$ 使得

$k_1\alpha_1 + \cdots + k_s\alpha_s = 0$

$\Leftrightarrow$0 可以由 $\alpha_1,\cdots,\alpha_s$ 的系数不全为 0 地线性表出

$\Leftrightarrow$齐次线性方程组

$x_1\alpha_1 + \cdots + x_s\alpha_s = 0$

有非零解

换句话说，

向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关

$\Leftrightarrow$齐次线性方程组

$x_1\alpha_1 + \cdots + x_s\alpha_s = 0$

只有零解

因为方程个数与未知量个数相等的齐次线性方程组有非零解的充分必要条件是它的系数矩阵的行列式等于零，所以从上面的结论可以得到：

在 K^n 中 n 个向量 $\alpha_1,\cdots,\alpha_n$ 线性相关

$\Leftrightarrow$以 $\alpha_1,\cdots,\alpha_n$ 为列向量组的矩阵的行列式等于零

换句话说，

在 K^n 中，n 个向量 $\alpha_1,\cdots,\alpha_n$ 线性无关

$\Leftrightarrow$以 $\alpha_1,\cdots,\alpha_n$ 为列向量组的矩阵的行列式不等于零.

由于 $|A'|=|A|$，所以我们也有

在 K^n 中，n 个向量 $\alpha_1,\cdots,\alpha_n$ 线性无关

$\Leftrightarrow$以 $\alpha_1,\cdots,\alpha_n$ 为行向量组的矩阵的行列式不等于零

例 6 判断下列向量组是线性相关还是线性无关？如果是线性相关，试找出其中的一个向量，使得这个向量能由其余向量线性表出，并且写出它的表达式.

(1) $\alpha_1=(1,-2,0,3)$ $\quad\alpha_2=(2,5,-1,0)$

$\alpha_3=(3,4,1,2)$

(2) $\alpha_1=(3,4,-2,5)$ $\quad\alpha_2=(2,-5,0,-3)$

$\alpha_3=(5,0,-1,2)$ $\quad\alpha_4=(3,3,-3,5)$

(3) $\alpha_1=(1,a,a^2,a^3)$ $\quad\alpha_2=(1,b,b^2,b^3)$

$\alpha_3=(1,c,c^2,c^3)$ $\quad\alpha_4=(1,d,d^2,d^3)$

其中 a,b,c,d 各不相同.

解 (1) 考虑齐次线性方程组

$$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3=0 \tag{5}$$

把它的增广矩阵经过初等行变换化成行阶梯形矩阵：

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

因为阶梯形方程组有 3 个方程，与未知量个数相等，所以齐次线性方程组(5)只有零解. 从而 $\alpha_1,\alpha_2,\alpha_3$ 线性无关.

(2) 考虑齐次线性方程组

$$x_1\alpha_1+x_2\alpha_2+x_3\alpha_3+x_4\alpha_4=0 \tag{6}$$

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

$$\rightarrow \begin{pmatrix} 1 & 2 & 4 & 0 \\ 0 & 1 & -5 & 6 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

方程组(6)的一般解是

$$\begin{cases} x_1 = -2x_4 \\ x_2 = -x_4 \\ x_3 = x_4 \end{cases}$$

其中 x_4 是自由未知量. 由于方程组(6)有非零解,所以 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 线性相关. 令 $x_4=1$,得到方程组(6)的一个解是$(-2,-1,1,1)$. 于是有

$$-2\alpha_1-\alpha_2+\alpha_3+\alpha_4=0$$

由上式可得出

$$\alpha_1=-\frac{1}{2}\alpha_2+\frac{1}{2}\alpha_3+\frac{1}{2}\alpha_4 \qquad \alpha_2=-2\alpha_1+\alpha_3+\alpha_4$$

$$\alpha_3=2\alpha_1+\alpha_2-\alpha_4 \qquad \alpha_4=2\alpha_1+\alpha_2-\alpha_3$$

(3)以 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 为列向量组成的矩阵 A 的行列式为

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

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

因为 a,b,c,d 各不相同,所以 $|A|\neq 0$. 从而 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 线性无关.

例 7 设 $\alpha_1,\cdots,\alpha_s$ 是 K^n 中的向量组,每个 $\alpha_i(i=1,\cdots,s)$ 都添上 m 个分量(所添分量的位置对于 $\alpha_1,\cdots,\alpha_s$ 都一样),便得到 K^{n+m}

中的向量组 $\tilde{\alpha}_1,\cdots,\tilde{\alpha}_s$. 称 $\tilde{\alpha}_1,\cdots,\tilde{\alpha}_s$ 是向量组 $\alpha_1,\cdots,\alpha_s$ 的**延伸组**. 试证:如果向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关,则它的延伸组 $\tilde{\alpha}_1,\cdots,\tilde{\alpha}_s$ 也线性无关.

证明 因为 $\alpha_1,\cdots,\alpha_s$ 线性无关,所以 s 元齐次线性方程组

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_s\alpha_s=0 \tag{7}$$

只有零解. 由于方程组(7) 是下述 s 元齐次线性方程组

$$x_1\tilde{\alpha}_1+x_2\tilde{\alpha}_2+\cdots+x_s\tilde{\alpha}_s=0 \tag{8}$$

中的 n 个方程组成的,因此方程组(8) 也只有零解(假如(8) 有非零解 $(c_1,\cdots,c_s)$,则$(c_1,\cdots,c_s)$ 适合(8) 的每一个方程,从而适合方程组(7),因此 $(c_1,\cdots,c_s)$ 是(7) 的一个非零解,矛盾). 由此得出向量组 $\tilde{\alpha}_1,\cdots,\tilde{\alpha}_s$ 线性无关. ▌

在研究一个向量能否由向量组 $\alpha_1,\cdots,\alpha_s$ 线性表出时,下面的命题 3.3.3 是常用的.

命题 3.3.3 设向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关,而向量组 $\alpha_1,\cdots,\alpha_s,\beta$ 线性相关,则 β 可以由 $\alpha_1,\cdots,\alpha_s$ 线性表出.

证明 因为 $\alpha_1,\cdots,\alpha_s,\beta$ 线性相关,所以有 K 中不全为零的数 $k_1,\cdots,k_s,l$,使得

$$k_1\alpha_1+\cdots+k_s\alpha_s+l\beta=0 \tag{9}$$

假如 $l=0$,则 $k_1,\cdots,k_s$ 不全为零. 于是从(9) 式得出,$\alpha_1,\cdots,\alpha_s$ 线性相关,这与已知条件矛盾. 因此 $l\neq0$. 从而由(9) 式得

$$\beta=-\frac{k_1}{l}\alpha_1-\cdots-\frac{k_s}{l}\alpha_s$$ ▌

推论 3.3.2 设向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关,如果向量 β 不能由 $\alpha_1,\cdots,\alpha_s$ 线性表出,则 $\alpha_1,\cdots,\alpha_s,\beta$ 线性无关. ▌

习 题 3.3

1. 下述说法对吗? 为什么?

(1)“向量组 $\alpha_1,\cdots,\alpha_s$，如果有 K 中全为零的数 $k_1,\cdots,k_s$，使得

$$k_1\alpha_1+\cdots+k_s\alpha_s=0$$

则 $\alpha_1,\cdots,\alpha_s$ 线性无关.”

(2)“如果有 K 中不全为零的数 $k_1,\cdots,k_s$，使得

$$k_1\alpha_1+\cdots+k_s\alpha_s\neq 0$$

则 $\alpha_1,\cdots,\alpha_s$ 线性无关.”

(3)“若向量组 $\alpha_1,\cdots,\alpha_s$ 线性相关，则其中每一个向量都可以由其余向量线性表出.”

2. 试证：单个向量 α 线性相关的充分必要条件是 $\alpha=0$.

3. 试证：若向量组 $\alpha_1,\alpha_2,\alpha_3$ 线性无关，则

(1) $\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1$ 也线性无关；

(2) $2\alpha_1+3\alpha_2,\alpha_2+4\alpha_3,5\alpha_3+\alpha_1$ 也线性无关.

4. 试证：R^3 中任意四个向量都线性相关.

5. 试证：K^n 中，若 $s>n$，则任意 s 个向量都线性相关.

6. 试证：K^n 中，向量组 $\alpha_1=(1,0,0,\cdots,0),\alpha_2=(1,1,0,\cdots,0),\alpha_3=(1,1,1,\cdots,0),\cdots,\alpha_n=(1,1,1,\cdots,1)$ 线性无关.

7. 判断下列向量组是否线性相关？若线性相关，试找出其中的一个向量，使得这个向量可以由其余向量线性表出，并且写出它的一种表出方式.

(1) $\alpha_1=(3,1,2,-4)$　　$\alpha_2=(1,0,5,2)$

$\alpha_3=(-1,2,0,3)$

(2) $\alpha_1=(-2,1,0,3)$　　$\alpha_2=(1,-3,2,4)$

$\alpha_3=(3,0,2,-1)$　　$\alpha_4=(2,-2,4,6)$

(3) $\alpha_1=(3,-1,2)$　　$\alpha_2=(1,5,-7)$

$\alpha_3=(7,-13,20)$　　$\alpha_4=(-2,6,1)$

(4) $\alpha_1=(1,-2,4,-8)$　　$\alpha_2=(1,3,9,27)$

$\alpha_3=(1,4,16,64)$，　　$\alpha_4=(1,-1,1,-1)$

8. 设 $\alpha_1,\cdots,\alpha_s$ 是 K^n 中的向量组，每个 α_i 都去掉 m 个分量(去掉的分量的位置对于 $\alpha_1,\cdots,\alpha_s$ 都一样)，得到 K^{n-m} 中的向量组 $\alpha_1^*,\cdots,\alpha_s^*$，称 $\alpha_1^*,\cdots,\alpha_s^*$ 是 $\alpha_1,\cdots,\alpha_s$ 的**缩短组**. 试证：如果 $\alpha_1,\cdots,\alpha_s$ 线性相关，则它的缩短组也线性相关.

9. 设向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关，设 $\beta=b_1\alpha_1+\cdots+b_s\alpha_s$. 如果对于某个

$i(1 \leqslant i \leqslant s)$,$b_i \neq 0$,则用 β 替换 α_i 以后得到的向量组 $\alpha_1,\cdots,\alpha_{i-1},\beta,\alpha_{i+1},\cdots,\alpha_s$ 也线性无关.

10. 试证:由非零向量组成的向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关的充分必要条件是:每一个 $\alpha_i(1 < i \leqslant s)$ 都不能用它前面的向量线性表出.

11. 设 $a_1,a_2,\cdots,a_r$ 是互不相同的数,$r \leqslant n$. 试证:$\alpha_i = (1,a_i,\cdots,a_i^{n-1})$,$i = 1,2,\cdots,r$,是线性无关的.

12. 设 $\alpha_1,\cdots,\alpha_r$ 线性无关,设 $\beta_i = \sum\limits_{j=1}^{r} a_{ij}\alpha_j$, $i = 1,\cdots,r$. 证明:$\beta_1,\beta_2,\cdots,\beta_r$ 线性无关的充分必要条件是

$$\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{r1} \\ a_{12} & a_{22} & \cdots & a_{r2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1r} & a_{2r} & \cdots & a_{rr} \end{vmatrix} \neq 0$$

§4 基·维数·向量组的秩

有了向量组线性相关以及线性无关的概念之后,就可以来研究数域 K 上 n 维向量空间 K^n 及其线性子空间的结构了.

几何空间 R^3 中,取定三个不共面的向量 $\vec{e}_1,\vec{e}_2,\vec{e}_3$,则空间中任一向量都可以由 $\vec{e}_1,\vec{e}_2,\vec{e}_3$ 线性表出,并且表法唯一. 这样,几何空间 R^3 的结构就非常清楚了. 我们称 $\vec{e}_1,\vec{e}_2,\vec{e}_3$ 是几何空间 R^3 的一个基. 其次考虑过原点的一个平面 π,在 π 上取定两个不共线的向量 $\vec{a}$,$\vec{b}$,则 π 上每一个向量具有形式 $k\vec{a} + l\vec{b}$,并且如果 $k_1\vec{a} + l_1\vec{b}$ 与 $k_2\vec{a} + l_2\vec{b}$ 是 π 上相同的向量,则 $(k_1,l_1) = (k_2,l_2)$. 这样,π 的结构也很清楚. 类似地,我们称 $\vec{a},\vec{b}$ 是平面 π 的一个基. 注意到三个不共面的向量是线性无关的,两个不共线的向量也是线性无关的,那么我们容易从这几何直观中受到启发,把“基”的概念推广到 K^n 及其线性子空间上,从而了解它们的结构.

定义 1 设 W 是 K^n 的一个线性子空间，W 中的一个向量组 $\alpha_1,\cdots,\alpha_s$ 称为是 W 的一个**基**，如果它是线性无关的，并且 W 中每一个向量都可以由 $\alpha_1,\cdots,\alpha_s$ 线性表出.

定义 1 中，基的第二个条件等价于："$\alpha_1,\cdots,\alpha_s$ 生成的线性子空间等于 W，即 $\langle\alpha_1,\cdots,\alpha_s\rangle = W$". 于是有

命题 3.4.1 设 W 是 K^n 的一个线性子空间，W 中的一个向量组 $\alpha_1,\cdots,\alpha_s$ 是 W 的一个基当且仅当 $\alpha_1,\cdots,\alpha_s$ 线性无关，并且 $\langle\alpha_1,\cdots,\alpha_s\rangle = W$. ▌

根据命题 3.3.2，如果 $\alpha_1,\cdots,\alpha_s$ 是 W 的一个基，则 W 中每一个向量 ω 能唯一表示成形式 $\omega = k_1\alpha_1 + \cdots + k_s\alpha_s$，系数组成的 s 元有序数组 $(k_1,\cdots,k_s)$ 称为 ω 在基 $\alpha_1,\cdots,\alpha_s$ 下的**坐标**. 因此，如果找到了 W 的一个基，那么 W 的结构就清楚了. 所以基的概念是十分重要的.

例 1 K^n 中的向量组

$$\begin{aligned}
&\varepsilon_1 = (1,0,0,\cdots,0,0), \qquad &&\varepsilon_2 = (0,1,0,\cdots,0,0),\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots, &&\varepsilon_n = (0,0,0,\cdots,0,1)
\end{aligned}$$

是线性无关的，并且 K^n 中任一向量 $\alpha = (a_1,a_2,\cdots,a_n)$ 能表示成

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n$$

因此 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 K^n 的一个基. 这个基称为 K^n 的**标准基**. α 在标准基下的坐标恰好是 $(a_1,a_2,\cdots,a_n)$，这与 α 一致.

例 2 考虑 K^n 中的向量组

$$\begin{aligned}
&\alpha_1 = (1,0,0,0,\cdots,0), \qquad &&\alpha_2 = (1,1,0,0,\cdots,0),\\
&\alpha_3 = (1,1,1,0,\cdots,0), &&\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\
&\alpha_n = (1,1,1,1,\cdots,1)
\end{aligned}$$

由于

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{vmatrix} = 1$$

所以 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关，并且对于 K^n 中任一向量 $\alpha = (a_1, a_2, \cdots, a_n)$，线性方程组

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \alpha \tag{1}$$

有唯一解. 从而 α 可以由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表出. 所以 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 K^n 的一个基. 通过求出方程组(1)的解，得到 α 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标是

$$(a_1 - a_2, a_2 - a_3, \cdots, a_{n-1} - a_n, a_n)$$

从例1和例2中看出，K^n 有基，而且 K^n 的基不只一个. 现在要问：K^n 的每一个线性子空间是否都有基？如果有基，在各个基里基向量的数目是否相同？为了回答这两个问题，我们需要下面的引理3.4.1.

先给出一个术语：K^n 中，如果向量组(I)的每一个向量都能由向量组(II)线性表出，则称向量组(I)可以由向量组(II)线性表出. 如果向量组(I)和(II)可以互相线性表出，则称向量组(I)与(II) **等价**，记作$(I) \cong (II)$.

注意：本书中谈到向量组时，均指它只含有有限多个向量.

引理 3.4.1 K^n 中，设向量组 $\beta_1, \cdots, \beta_r$ 可以由向量组 $\alpha_1, \cdots, \alpha_s$ 线性表出. 如果 $r > s$，那么向量组 $\beta_1, \cdots, \beta_r$ 线性相关.

证明 设

$$\beta_j = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{sj}\alpha_s, \quad j = 1, \cdots, r \tag{2}$$

假设 $r > s$，我们想证明 $\beta_1, \cdots, \beta_r$ 一定线性相关. 为此做以 $x_j (j = 1, \cdots, r)$ 为系数的向量组 $\beta_1, \cdots, \beta_r$ 的线性组合：

$$x_1\beta_1 + x_2\beta_2 + \cdots + x_r\beta_r$$

$$
\begin{aligned}
&= x_1(a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{s1}\alpha_s) + \\
&\quad \cdots + x_r(a_{1r}\alpha_1 + a_{2r}\alpha_2 + \cdots + a_{sr}\alpha_s) \\
&= (a_{11}x_1 + \cdots + a_{1r}x_r)\alpha_1 + (a_{21}x_1 + \cdots + a_{2r}x_r)\alpha_2 + \\
&\quad \cdots + (a_{s1}x_1 + \cdots + a_{sr}x_r)\alpha_s
\end{aligned} \tag{3}
$$

考虑下述齐次线性方程组

$$
\begin{cases}
a_{11}x_1 + \cdots + a_{1r}x_r = 0 \\
a_{21}x_1 + \cdots + a_{2r}x_r = 0 \\
\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
a_{s1}x_1 + \cdots + a_{sr}x_r = 0
\end{cases} \tag{4}
$$

因为 $s < r$，所以齐次线性方程组(4)一定有非零解. 取它的一个非零解 $(c_1, \cdots, c_r)$，代入(4)和(3)中，得

$$
c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r = 0\alpha_1 + 0\alpha_2 + \cdots + 0\alpha_s = 0 \tag{5}
$$

所以 向量组 $\beta_1, \beta_2, \cdots, \beta_r$ 线性相关. ▌

推论 3.4.1　K^n 中任一线性无关的向量组所含向量的个数不超过 n.

证明　设 $\alpha_1, \cdots, \alpha_s$ 是线性无关的向量组. 因为 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 K^n 的一个基，所以向量组 $\alpha_1, \cdots, \alpha_s$ 可以由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 线性表出. 又由于 $\alpha_1, \cdots, \alpha_s$ 线性无关，据引理 3.4.1，得 $s \leqslant n$. ▌

现在我们来依次回答前面所提出的两个问题.

定理 3.4.1　K^n 的每一个非零线性子空间 W 都有一个基.

证明　因为 $W \neq 0$，所以可以在 W 中选取一个非零向量 α_1. 向量组 α_1 是线性无关的. 如果 $\langle \alpha_1 \rangle \neq W$，那么 W 中可选取一个向量 $\alpha_2 \overline{\in} \langle \alpha_1 \rangle$. 于是 α_2 不能由向量组 α_1 线性表出. 根据推论 3.3.2，向量组 α_1, α_2 是线性无关的. 如果 $\langle \alpha_1, \alpha_2 \rangle \neq W$，那么 W 中可选取向量 $\alpha_3 \overline{\in} \langle \alpha_1, \alpha_2 \rangle$. 于是 α_3 不能由向量组 α_1, α_2 线性表出，从而向量组 $\alpha_1, \alpha_2, \alpha_3$ 线性无关. 继续这样做. 但是这个过程不能无限进行下去，因为根据推论 3.4.1，K^n 中任一线性无关的向量组所含向量的个数不超过 n. 假设当我们得到了 W 中一个线性无关的向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 以后，不管把 W 中哪个向量 β 添进去，所得到的向量组

$\alpha_1,\alpha_2,\cdots,\alpha_s,\beta$ 都是线性相关的. 据命题 3.3.3,β 可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出. 这表明:W 中每一个向量都可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出. 于是由定义 1 即知,向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 W 的一个基. ▌

定理 3.4.2　设 W 是 K^n 的一个非零线性子空间,则 W 的所有基都含有相同数目的向量.

证明　设 $\alpha_1,\cdots,\alpha_s$ 与 $\beta_1,\cdots,\beta_r$ 是 W 的任意两个基. 因为 $\alpha_1,\cdots,\alpha_s$ 是 W 的一个基,所以向量组 $\beta_1,\cdots,\beta_r$ 可以由向量组 $\alpha_1,\cdots,\alpha_s$ 线性表出,又由于 $\beta_1,\cdots,\beta_r$ 是线性无关的,据引理 3.4.1 得,$r\leqslant s$. 因为 $\beta_1,\cdots,\beta_r$ 是 W 的一个基,同上面的道理可得 $s\leqslant r$. 因此 $s=r$. ▌

定义 2　设 W 是 K^n 的一个非零线性子空间,W 的一个基所含的向量的数目称为 W 的**维数**,记作 $\dim_K W$ 或简记作 $\dim W$. 零子空间的维数规定为 0.

据推论 3.4.1,K^n 的每一个非零线性子空间的维数不会超过 n. 显然,非零子空间的维数一定大于 0.

因为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 K^n 的一个基,所以 K^n 的维数是 n,这就是我们在 §2 中把 K^n 称为 n 维向量空间的根据.

几何空间 R^3 中,任意三个不共面的向量是它的一个基,因此几何空间是 3 维的. 通过原点的一个平面 π,它上面的任意两个不共线的向量是 π 的一个基,因此平面 π 是 2 维的. 通过原点的一条直线 L,它上面的任意一个非零向量是 L 的一个基,因此直线 L 是 1 维的.

维数也是一个十分重要的概念. 它在刻划 K^n 的线性子空间的结构上起着重要的作用.

命题 3.4.2　设 W 是 K^n 的一个非零线性子空间,如果 $\dim W=r$,那么 W 中任意 r 个线性无关的向量是 W 的一个基.

证明　设 $\alpha_1,\cdots,\alpha_r$ 是 W 中线性无关的向量组,设 $\delta_1,\cdots,\delta_r$ 是 W 的一个基. 任取 $\beta\in W$,由于 $\alpha_1,\cdots,\alpha_r,\beta$ 可以由 $\delta_1,\cdots,\delta_r$ 线性表

出，所以 $\alpha_1,\cdots,\alpha_r,\beta$ 线性相关（据引理 3.4.1）. 于是 β 可以由 $\alpha_1,\cdots,\alpha_r$ 线性表出（据命题 3.3.3）. 因此 $\alpha_1,\cdots,\alpha_r$ 是 W 的一个基. ▍

命题 3.4.3　设 W 和 U 是 K^n 的两个非零线性子空间，如果 $W\subset U$，那么 $\dim W\leqslant\dim U$.

证明　在 W 中取一个基 $\alpha_1,\cdots,\alpha_r$；在 U 中取一个基 $\beta_1,\cdots,\beta_s$. 因为 $W\subset U$，所以 $\alpha_i\in U$，$i=1,\cdots,r$. 于是 $\alpha_1,\cdots,\alpha_r$ 可以由 $\beta_1,\cdots,\beta_s$ 线性表出. 又由于 $\alpha_1,\cdots,\alpha_r$ 线性无关，据引理 3.4.1 得，$r\leqslant s$，即 $\dim W\leqslant\dim U$. ▍

命题 3.4.4　设 W 和 U 是 K^n 的两个非零线性子空间. 如果 $W\subset U$，且 $\dim W=\dim U$，则 $W=U$.

证明　在 W 中取一个基 $\alpha_1,\cdots,\alpha_r$. 由于 $W\subset U$，所以 $\alpha_i\in U$，$i=1,\cdots,r$. 由于 $\dim U=\dim W=r$，所以 $\alpha_1,\cdots,\alpha_r$ 是 U 的一个基（据命题 3.4.2）. 于是 $W=\langle\alpha_1,\cdots,\alpha_r\rangle=U$. ▍

设 $\alpha_1,\cdots,\alpha_s$ 是 K^n 的一个向量组，我们来讨论如何找 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的一个基以及求它的维数.

情形 1　若 $\alpha_1,\cdots,\alpha_s$ 线性无关，则 $\alpha_1,\cdots,\alpha_s$ 就是 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的一个基.

情形 2　若 $\alpha_1,\cdots,\alpha_s$ 线性相关，则 $\alpha_1,\cdots,\alpha_s$ 不是 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的一个基. 这时很自然地想在 $\alpha_1,\cdots,\alpha_s$ 中选出一部分向量是线性无关的，并且应选取最大的线性无关的部分组，这样的部分组才有可能成为 $\langle\alpha_1,\cdots,\alpha_r\rangle$ 的一个基. 为此引进下述概念：

定义 3　K^n 中向量组的一个部分组称为一个**极大线性无关组**，如果这个部分组本身是线性无关的，但是从这个向量组的其余向量（如果还有的话）中任取一个添进去，得到的新的部分组都线性相关.

例如，几何空间 R^3 中，向量组 $\vec{a},\vec{b},\vec{c}$ 如果共面，但 $\vec{a},\vec{b}$ 不共线，则 $\vec{a},\vec{b}$ 就是向量组 $\vec{a},\vec{b},\vec{c}$ 的一个极大线性无关组.

性质 1　K^n 中包含非零向量的任一向量组与它的极大线性无关组等价.

证明 任取 K^n 中包含非零向量的一个向量组(I),设 $\alpha_1,\cdots,\alpha_r$ 是向量组(I)的一个极大线性无关组.任取(I)中一个向量β,若$\beta=\alpha_i$,对于某个$i\in\{1,\cdots,r\}$,则$\beta=\alpha_i=1\cdot\alpha_i$,这说明$\beta$可以由$\alpha_1,\cdots,\alpha_r$线性表出.若$\beta\neq\alpha_i,i=1,\cdots,r$,据定义3,向量组$\alpha_1,\cdots,\alpha_r,\beta$线性相关.又由于$\alpha_1,\cdots,\alpha_r$线性无关,据命题3.3.3得,$\beta$可以由$\alpha_1,\cdots,\alpha_r$线性表出,因此向量组$(I)$可以由$\alpha_1,\cdots,\alpha_r$线性表出.

反之,因为$\alpha_i=1\cdot\alpha_i,i=1,\cdots,r$,所以$\alpha_1,\cdots,\alpha_r$可以由向量组$(I)$线性表出. ▍

前面指出,向量组$\alpha_1,\cdots,\alpha_s$的一个极大线性无关组有可能成为$\langle\alpha_1,\cdots,\alpha_s\rangle$的一个基,为此还需要下面的引理3.4.2:

引理 3.4.2(线性表出的传递性) K^n 中,如果向量组(I)可以由向量组(II)线性表出,并且(II)可以由向量组(III)线性表出,则(I)可以由(III)线性表出.

证明 任取(I)的一个向量α,由于(I)可以由(II):$\beta_1,\cdots,\beta_r$线性表出,所以可设$\alpha=\sum\limits_{i=1}^{r}k_i\beta_i$.又由于$(II)$可以由$(III)$:$\gamma_1,\cdots,\gamma_s$线性表出,所以

$$\beta_i=\sum_{j=1}^{s}l_{ij}\gamma_j,\quad i=1,\cdots,r$$

于是

$$\begin{aligned}\alpha&=\sum_{i=1}^{r}k_i\Big(\sum_{j=1}^{s}l_{ij}\gamma_j\Big)\\&=\sum_{i=1}^{r}\sum_{j=1}^{s}k_il_{ij}\gamma_j=\sum_{j=1}^{s}\Big(\sum_{i=1}^{r}k_il_{ij}\Big)\gamma_j\end{aligned}$$

这表明α可以由向量组(III)线性表出.因此(I)可以由(III)线性表出. ▍

有了性质1和引理3.4.2,我们就可以回答上面的问题了.

定理 3.4.3 设$\alpha_1,\cdots,\alpha_s$是K^n中包含非零向量的一个向量组,则$\alpha_1,\cdots,\alpha_s$的任意一个极大线性无关组是$\langle\alpha_1,\cdots,\alpha_s\rangle$的一个基.

证明 不妨设 $\alpha_1,\cdots,\alpha_r(r\leqslant s)$ 是 $\alpha_1,\cdots,\alpha_s$ 的一个极大线性无关组. 由性质 1, $\alpha_1,\cdots,\alpha_s$ 可以由 $\alpha_1,\cdots,\alpha_r$ 线性表出. 据引理 3.4.2, $\langle\alpha_1,\cdots,\alpha_s\rangle$ 中每个向量可以由 $\alpha_1,\cdots,\alpha_r$ 线性表出. 又 $\alpha_1,\cdots,\alpha_r$ 线性无关, 因此 $\alpha_1,\cdots,\alpha_r$ 是 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的一个基. ▌

反之, 容易看出, 如果 $\alpha_1,\cdots,\alpha_s$ 的一个部分组是子空间 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的一个基, 那么这个部分组是 $\alpha_1,\cdots,\alpha_s$ 的一个极大线性无关组.

由定理 3.4.3 立即得到

性质 2 设 $\alpha_1,\cdots,\alpha_s$ 是 K^n 中包含非零向量的一个向量组, 则它的所有极大线性无关组含有相同数目的向量.

证明 由定理 3.4.3, $\alpha_1,\cdots,\alpha_s$ 的任意一个极大线性无关组都是 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的一个基. 再据定理 3.4.2 即得结论. ▌

由于性质 2, 我们可引进下述概念:

定义 4 设 $\alpha_1,\cdots,\alpha_s$ 是 K^n 中包含非零向量的一个向量组, 它的一个极大线性无关组包含的向量的数目称为 $\alpha_1,\cdots,\alpha_s$ 的**秩**, 记作 $\operatorname{rank}\{\alpha_1,\cdots,\alpha_s\}$. 全由零向量组成的向量组的秩规定为 0.

由于 $\alpha_1,\cdots,\alpha_s$ 的任意一个极大线性无关组是 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的一个基, 以及全由零向量组成的向量组生成的线性子空间是零子空间, 所以我们得到

推论 3.4.2 K^n 中向量组 $\alpha_1,\cdots,\alpha_s$ 的秩等于 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的维数, 即

$$\operatorname{rank}\{\alpha_1,\cdots,\alpha_s\}=\dim\langle\alpha_1,\cdots,\alpha_s\rangle \tag{6}$$

▌

定理 3.4.3 和推论 3.4.2 回答了前面所提的问题, 即为了求 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的一个基, 就只要去找 $\alpha_1,\cdots,\alpha_s$ 的极大线性无关组, 为了求 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的维数, 就只要去求 $\alpha_1,\cdots,\alpha_s$ 的秩. 如果向量组 $\alpha_1,\cdots,\alpha_s$ 的每个向量的分量具体给出了, 那么在下一节我们将给出求 $\alpha_1,\cdots,\alpha_s$ 的极大线性无关组的方法以及求 $\alpha_1,\cdots,\alpha_s$ 的秩的方法. 现在先给出比较两个向量组的秩的一个常用方法:

命题 3.4.5　K^n 中，若向量组(I)：$\alpha_1,\cdots,\alpha_r$ 可以由向量组(II)：$\beta_1,\cdots,\beta_s$ 线性表出，则(I) 的秩 $\leqslant$ (II) 的秩.

证明　由已知条件得，$\alpha_i \in \langle\beta_1,\cdots,\beta_s\rangle, i=1,\cdots,r$. 从而$\langle\alpha_1,\cdots,\alpha_r\rangle \subset \langle\beta_1,\cdots,\beta_s\rangle$. 据命题 3.4.3 得

$$\dim\langle\alpha_1,\cdots,\alpha_r\rangle \leqslant \dim\langle\beta_1,\cdots,\beta_s\rangle$$

也就是

$$\operatorname{rank}\{\alpha_1,\cdots,\alpha_r\} \leqslant \operatorname{rank}\{\beta_1,\cdots,\beta_s\}$$ ▌

由命题 3.4.5 的结论及其证明过程立即得到下面两个推论：

推论 3.4.3　等价的向量组有相同的秩.　▌

推论 3.4.4

$$\{\alpha_1,\cdots,\alpha_r\} \cong \{\beta_1,\cdots,\beta_s\} \Leftrightarrow \langle\alpha_1,\cdots,\alpha_r\rangle = \langle\beta_1,\cdots,\beta_s\rangle$$

即两个向量组等价当且仅当它们分别生成的线性子空间相同. ▌

向量组的等价是 K^n 中向量组之间的一种关系，不难看出关系 $\cong$ 具有下述三条性质：

1°　反身性：每一个向量组与自身等价；

2°　对称性：若$(I) \cong (II)$，则$(II) \cong (I)$；

3°　传递性：若$(I) \cong (II)$，且$(II) \cong (III)$，则$(I) \cong (III)$.

利用向量组的秩可以得到向量组线性无关的又一个判定方法：

命题 3.4.6　向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关的充分必要条件是：它的秩等于 s.

证明　必要性. 若 $\alpha_1,\cdots,\alpha_s$ 线性无关，则 $\alpha_1,\cdots,\alpha_s$ 的极大线性无关组是它自身，所以 $\operatorname{rank}\{\alpha_1,\cdots,\alpha_s\}=s$.

充分性. 若$\operatorname{rank}\{\alpha_1,\cdots,\alpha_s\}=s$，则它的极大线性无关组含 s 个向量，从而就是 $\alpha_1,\cdots,\alpha_s$ 自身. 所以 $\alpha_1,\cdots,\alpha_s$ 线性无关.　▌

习 题 3.4

1. 找出 K^4 的两个基，并且求向量 $\alpha=(a_1,a_2,a_3,a_4)$ 分别在这两个基下的坐标.

2. 设 $r<n$，求 K^n 的下述子空间的一个基和维数：

$$W=\{(a_1,\cdots,a_r,0,\cdots,0)\,|\,a_i\in K,\quad i=1,\cdots,r\}$$

3. 证明：秩为 r 的向量组中任意 r 个线性无关的向量都构成它的一个极大线性无关组.

4. 证明：如果秩为 r 的向量组可以由它的 r 个向量线性表出，则这 r 个向量构成该向量组的一个极大线性无关组.

5. 试证：K^n 中，如果标准基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出，则 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关.

6. 试证：K^n 中，向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关的充分必要条件是 K^n 中任一向量都可由它们线性表出.

7. 设 $\alpha_1,\cdots,\alpha_n$ 是 K^n 的一个向量组，试证：如果 K^n 的每一个向量都可以由 $\alpha_1,\cdots,\alpha_n$ 线性表出，则 $\alpha_1,\cdots,\alpha_n$ 是 K^n 的一个基.

8. 设 W 是 K^n 的一个非零线性子空间，设 $\dim W=r$. 如果 W 中每一个向量都可以由 W 的一个向量组 $\alpha_1,\cdots,\alpha_r$ 线性表出，证明 $\alpha_1,\cdots,\alpha_r$ 是 W 的一个基.

9. 证明：在 K^n 中，一个向量组的任何一个线性无关组都可以扩充成一个极大线性无关组.

10. 设

$$\alpha_1=(3,-1,2,5),\quad \alpha_2=(4,3,7,2),\quad \alpha_3=(1,4,5,-3)$$
$$\alpha_4=(7,-1,2,0),\quad \alpha_5=(1,2,5,6)$$

(1) 证明：α_1,α_2 线性无关；

(2) 把 α_1,α_2 扩充成一个极大线性无关组.

*11. 证明：在 K^n 的一个线性子空间 W 里，每一个线性无关的向量组都能扩充成 W 的一个基.

12. 证明：数域 K 上的线性方程组

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta$$

对任何 $\beta \in K^n$ 都有解的充分必要条件是它的系数矩阵 A 的行列式 $|A| \neq 0$.

13. 试证：如果 $\mathrm{rank}\{\alpha_1,\cdots,\alpha_s\} = \mathrm{rank}\{\alpha_1,\cdots,\alpha_s,\beta\}$，则 β 可以由 $\alpha_1,\cdots,\alpha_s$ 线性表出.

*14. 试证：K^n 中如果两个向量组有相同的秩，并且其中一个向量组可以由另一个向量组线性表出，则这两个向量组等价.

*15. 在什么样的情形下，一个向量组具有唯一的极大线性无关组？

*16. s 个向量的向量组如果它的秩为 $s-1$，且包含成比例的非零向量，问：此向量组有多少个极大线性无关组？

§5 矩阵的秩

为了计算向量组的秩，以及研究线性方程组有无解的问题和解的结构问题，都需要利用“矩阵的秩”的概念. 矩阵的秩是线性代数中非常深刻、非常有用的概念.

定义 1 数域 K 上矩阵 A 的行向量组的秩称为 A 的**行秩**；A 的列向量组的秩称为 A 的**列秩**.

在 §2 中我们把矩阵 A 的行向量组生成的线性子空间称为 A 的行空间；A 的列向量组生成的线性子空间称为 A 的列空间. 于是据 §4 的推论 3.4.2 得，矩阵 A 的行秩等于 A 的行空间的维数，A 的列秩等于 A 的列空间的维数.

矩阵 A 的行秩与列秩有什么关系呢？让我们先看阶梯形矩阵.

命题 3.5.1 阶梯形矩阵 J 的行秩与列秩相等，都等于 J 的非零行的数目；J 的主元所在的列组成的向量组是 J 的列向量组的一个极大线性无关组.

证明 设 $s\times n$ 阶梯形矩阵 J 有 r 个非零行($r\leqslant s$),则 J 有 r 个主元,设它们分别位于第 $j_1,j_2,\cdots,j_r$ 列.于是 J 形如

$$\begin{pmatrix} 0 & \cdots & 0 & c_{1,j_1} & \cdots & c_{1,j_2} & \cdots & c_{1,j_r} & \cdots & c_{1n} \\ 0 & \cdots & 0 & 0 & \cdots & c_{2,j_2} & \cdots & c_{2,j_r} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & c_{r,j_r} & \cdots & c_{rn} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

其中 $c_{1,j_1}c_{2,j_2}\cdots c_{r,j_r}\neq 0$.

设 J 的列向量组为 $\beta_1,\cdots,\beta_n$,则

$$J\text{ 的列秩}=\dim\langle\beta_1,\cdots,\beta_n\rangle$$

考虑 K^s 的子空间:

$$W=\{(a_1,\cdots,a_r,0,\cdots,0)'\,|\,a_i\in K,\quad i=1,\cdots,r\}$$

据习题 3.4 的第 2 题,$\dim W=r$.由于 $\beta_i(i=1,\cdots,n)$ 的后面 $s-r$ 个分量全为零,因此有

$$\langle\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_r}\rangle\subset\langle\beta_1,\beta_2,\cdots,\beta_n\rangle\subset W \tag{1}$$

于是
$$\dim\langle\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_r}\rangle\leqslant\dim\langle\beta_1,\beta_2,\cdots,\beta_n\rangle\leqslant r \tag{2}$$

把 $\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_r}$ 的前 r 个分量组成的向量组记作 $\beta_{j_1}{}^*,\beta_{j_2}{}^*,\cdots,\beta_{j_r}{}^*$,以它们为列向量组的矩阵为

$$B_1=\begin{pmatrix} c_{1,j_1} & c_{1,j_2} & \cdots & c_{1,j_r} \\ 0 & c_{2,j_2} & \cdots & c_{2,j_r} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & c_{r,j_r} \end{pmatrix}$$

因为 $|B_1|=c_{1,j_1}c_{2,j_2}\cdots c_{r,j_r}\neq 0$,所以 $\beta_{j_1}{}^*,\beta_{j_2}{}^*,\cdots,\beta_{j_r}{}^*$ 线性无关.从而 $\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_r}$ 线性无关.于是 $\mathrm{rank}\{\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_r}\}=r$.再从(2)得 $\dim\langle\beta_1,\beta_2,\cdots,\beta_n\rangle=r$.于是 J 的列秩等于 r.又由于 β_{j_1},β_{j_2},

$\cdots,\beta_{j_r}$ 线性无关，因此它们是 J 的列向量组的一个极大线性无关组(据习题 3.4 第 3 题).

J 的行向量组记作 $\alpha_1,\cdots,\alpha_r,0,\cdots,0$. 把 $\alpha_1,\cdots,\alpha_r$ 的第 $j_1,j_2,\cdots,j_r$ 个分量组成的向量组记作$\alpha_1^*,\cdots,\alpha_r^*$，以它们为行向量组的矩阵为 B_1. 由于 $|B_1|\neq 0$，因此$\alpha_1^*,\cdots,\alpha_r^*$ 线性无关. 从而 $\alpha_1,\cdots,\alpha_r$ 线性无关. 所以 $\alpha_1,\cdots,\alpha_r$ 是 $\alpha_1,\cdots,\alpha_r,0,\cdots,0$ 的一个极大线性无关组. 于是 J 的行秩等于 r. ▌

一般的矩阵，其行秩与列秩是否也相等?它的列向量组的一个极大线性无关组如何求? 由于任何一个矩阵都可以经过一系列初等行变换化成阶梯形矩阵，因此解决上面两个问题的思路自然是去研究矩阵的初等行变换是否不改变矩阵的行秩? 是否不改变列向量组的线性相关性? 是否不改变矩阵的列秩?

定理 3.5.1 矩阵的初等行变换不改变矩阵的行秩.

证明 设矩阵 A 的行向量组是 $\alpha_1,\cdots,\alpha_s$. 设 A 经过 1° 型初等行变换 ⓙ + ⓘ $\cdot k$ 变成矩阵 B，则 B 的行向量组是 $\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j+k\alpha_i,\cdots,\alpha_s$. 这表明 B 的行向量组可以由 A 的行向量组线性表出. 由于 $\alpha_l=1\alpha_l$，当 $l\neq j$；并且 $\alpha_j=(\alpha_j+k\alpha_i)-k\alpha_i$，所以 A 的行向量组可以由 B 的行向量组线性表出. 因此 A 的行向量组与 B 的行向量组等价. 据 §4 的推论 3.4.3 得，A 的行秩等于 B 的行秩.

容易证明 2°，3° 型初等行变换使所得矩阵的行向量组与原矩阵的行向量组等价，从而不改变矩阵的行秩. ▌

定理 3.5.2 矩阵的初等行变换不改变矩阵的列向量组的线性相关性，从而不改变矩阵的列秩. 即

(1) 设矩阵 C 经过初等行变换变成矩阵 D，则 C 的列向量组线性相关的充分必要条件是 D 的列向量组线性相关；

(2) 设矩阵 A 经过初等行变换变成矩阵 B，并且设 B 的第 $j_1,\cdots,j_r$ 列组成 B 的列向量组的一个极大线性无关组，则 A 的第 $j_1,\cdots,j_r$ 列组成 A 的列向量组的一个极大线性无关组；从而 A 的列秩等于 B 的列秩.

证明 (1) 设 C 的列向量组是 $\gamma_1,\cdots,\gamma_n$;D 的列向量组是 $\beta_1,\cdots,\beta_n$. 我们知道

$$\gamma_1,\cdots,\gamma_n \text{ 线性相关}$$
$$\Leftrightarrow \text{齐次线性方程组 } x_1\gamma_1+\cdots+x_n\gamma_n=0 \text{ 有非零解}$$
$$\beta_1,\cdots,\beta_n \text{ 线性相关}$$
$$\Leftrightarrow \text{齐次线性方程组 } x_1\beta_1+\cdots+x_n\beta_n=0 \text{ 有非零解}$$

由于 $x_1\gamma_1+\cdots+x_n\gamma_n=0$ 的系数矩阵是 C,$x_1\beta_1+\cdots+x_n\beta_n=0$ 的系数矩阵是 D,并且 D 是 C 经过初等行变换得到的,因此 $x_1\beta_1+\cdots+x_n\beta_n=0$ 是 $x_1\gamma_1+\cdots+x_n\gamma_n=0$ 经过初等变换得到的方程组,从而这两个方程组同解. 所以 $x_1\gamma_1+\cdots+x_n\gamma_n=0$ 有非零解的充分必要条件是 $x_1\beta_1+\cdots+x_n\beta_n=0$ 有非零解. 由此得出,$\gamma_1,\cdots,\gamma_n$ 线性相关 $\Leftrightarrow\beta_1,\cdots,\beta_n$ 线性相关.

(2) 考虑 A 的第 $j_1,\cdots,j_r$ 列组成的子矩阵 A_1 与 B 的第 $j_1,\cdots,j_r$ 列组成的子矩阵 B_1. 当 A 经过初等行变换变成 B 时,矩阵 A_1 经过这些初等行变换便变成矩阵 B_1. 由于 B_1 的列向量组线性无关,据(1)的结论得,A_1 的列向量组也线性无关. 任取 $l\notin\{j_1,\cdots,j_r\}$. 考虑 A 的第 $j_1,\cdots,j_r,l$ 列组成的子矩阵 A_2 与 B 的第 $j_1,\cdots,j_r,l$ 列组成的子矩阵 B_2,后者是由前者经过初等行变换得到的. 因为 B_2 的列向量组线性相关,据(1)的结论得,A_2 的列向量组也线性相关. 所以,A 的第 $j_1,\cdots,j_r$ 列是 A 的列向量组的一个极大线性无关组. 由此即得,A 的列秩 $=r=B$ 的列秩. ▌

定理 3.5.3 任一矩阵的行秩等于它的列秩.

证明 任取矩阵 A,把它经过初等行变换化成阶梯形 J. 据定理 3.5.1,命题 3.5.1,和定理 3.5.2 得,A 的行秩 $=J$ 的行秩 $=J$ 的列秩 $=A$ 的列秩. ▌

定义 2 矩阵 A 的行秩与列秩统称为 A 的**秩**,记作 $\mathrm{rank}(A)$.

矩阵的秩是一个非常深刻的概念. 对于数域 K 上的任一 $s\times n$ 矩阵 A 来说,A 的行秩等于 A 的行空间的维数,A 的列秩等于 A 的列空间的维数. A 的秩的概念说明 A 的行空间的维数等于 A 的

列空间的维数. 注意 A 的行空间是 K^n 的一个线性子空间；而 A 的列空间是 K^s 的一个线性子空间. 它们的维数竟然一样！

推论 3.5.1　设矩阵 A 经过初等行变换变成阶梯形矩阵 J，则 A 的秩等于 J 的非零行的数目. 设 J 的主元所在的列是第 $j_1,\cdots,j_r$ 列，则 A 的第 $j_1,\cdots,j_r$ 列是 A 的列向量组的一个极大线性无关组.

证明　从命题 3.5.1，定理 3.5.2，和定理 3.5.3 即得. ▎

推论 3.5.1 也给出了求向量组 $\alpha_1,\cdots,\alpha_s$ 的一个极大线性无关组以及秩的一种方法. 即把 $\alpha_1,\cdots,\alpha_s$ 的每一个写成列向量，组成矩阵 A. 然后把 A 经过初等行变换化成阶梯形矩阵 J. 如果 J 的全部主元分别在第 $j_1,\cdots,j_r$ 列，那么 A 的第 $j_1,\cdots,j_r$ 列是 A 的列向量组的一个极大线性无关组，从而 $\alpha_{j_1},\cdots,\alpha_{j_r}$ 是向量组 $\alpha_1,\cdots,\alpha_s$ 的一个极大线性无关组. 向量组 $\alpha_1,\cdots,\alpha_s$ 的秩 $=A$ 的列秩 $=J$ 的非零行的数目. 这种方法也就给出了求向量组 $\alpha_1,\cdots,\alpha_s$ 生成的线性子空间 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 的一个基以及维数的方法.

例 1　求向量组

$$\alpha_1=(-1,5,3,-2),\qquad \alpha_2=(4,1,-2,9)$$

$$\alpha_3=(2,0,-1,4),\qquad \alpha_4=(0,3,4,-5)$$

的一个极大线性无关组以及它的秩.

解　把 $\alpha_1,\cdots,\alpha_4$ 写成列向量，组成矩阵 A，作初等行变换，把 A 化成行阶梯形 J：

$$A=\begin{pmatrix}-1 & 4 & 2 & 0\\ 5 & 1 & 0 & 3\\ 3 & -2 & -1 & 4\\ -2 & 9 & 4 & -5\end{pmatrix}\to\begin{pmatrix}-1 & 4 & 2 & 0\\ 0 & 1 & 0 & -5\\ 0 & 0 & 5 & 54\\ 0 & 0 & 0 & 0\end{pmatrix}=J$$

J 有 3 个非 0 行，所以 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的秩等于 3. J 的主元在第 1,2,3 列，所以 $\alpha_1,\alpha_2,\alpha_3$ 是向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的一个极大线性无关组.

例 2　求向量组

$$\alpha_1=(1,1,4,2),\qquad \alpha_2=(-1,-1,-4,-2)$$

$$\alpha_3=(-3,2,3,-11),\quad \alpha_4=(1,-1,-2,4)$$

$$\alpha_5=(1,3,10,0)$$

生成的线性子空间 $\langle\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5\rangle$ 的一个基和维数.

解 向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$ 的一个极大线性无关组就是 $\langle\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5\rangle$ 的一个基.

$$A=\begin{bmatrix}1&-1&-3&1&1\\1&-1&2&-1&3\\4&-4&3&-2&10\\2&-2&-11&4&0\end{bmatrix}$$

$$\rightarrow\begin{bmatrix}1&-1&-3&1&1\\0&0&5&-2&2\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}=J$$

因为阶梯形 J 的主元在第 1,3 列,所以 α_1,α_3 是 $\langle\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5\rangle$ 的一个基,其维数为 2.

由于 A' 的行(列)向量组是 A 的列(行)向量组,于是有

推论 3.5.2 $\operatorname{rank}(A')=\operatorname{rank}(A)$.

证明 $\operatorname{rank}(A')=A'$ 的行秩 $=A$ 的列秩 $=\operatorname{rank}(A)$. ▌

推论 3.5.3 矩阵的初等列变换不改变矩阵的秩.

证明 设矩阵 A 经过初等列变换变成矩阵 B. 由于一个矩阵的第 j 列是该矩阵的转置的第 j 行,因此 A' 经过相应的初等行变换变成 B'. 据定理 3.5.1 和推论 3.5.2 得

$$\operatorname{rank}(A)=\operatorname{rank}(A')=\operatorname{rank}(B')=\operatorname{rank}(B)$$ ▌

既然矩阵的初等行变换与初等列变换都不改变矩阵的秩,因此在求一个矩阵 A 的秩的时候,就可以对它既作初等行变换,又作初等列变换,更简便地把它化成阶梯形 J. 由于 $\operatorname{rank}(A)=\operatorname{rank}(J)$,并且 $\operatorname{rank}(J)=J$ 的非零行数目,因此我们得到下面的

推论 3.5.4 如果矩阵 A 经过初等行变换与初等列变换变成阶梯形矩阵 J,则

$$\operatorname{rank}(A)=J\text{ 的非零行的数目}$$ ▌

例 3 设

$$A=\begin{pmatrix}3 & -2 & 0 & 1 & -7\\ -1 & -3 & 2 & 0 & 4\\ 2 & 0 & -4 & 5 & 1\\ 4 & 1 & -2 & 1 & -11\end{pmatrix}$$

求 A 的秩.

解

$$A\to\begin{pmatrix}-1 & -3 & 2 & 0 & 4\\ 0 & -11 & 6 & 1 & 5\\ 0 & -6 & 0 & 5 & 9\\ 0 & -11 & 6 & 1 & 5\end{pmatrix}$$

$$\to\begin{pmatrix}-1 & -3 & 2 & 0 & 4\\ 0 & -11 & 6 & 1 & 5\\ 0 & -6 & 0 & 5 & 9\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}\to\begin{pmatrix}-1 & 2 & -3 & 0 & 4\\ 0 & 6 & -11 & 1 & 5\\ 0 & 0 & -6 & 5 & 9\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}$$

最后一个阶梯形矩阵有 3 个非零行,所以 $\mathrm{rank}(A)=3$.

现在我们从矩阵的子式(参见第二章 §8)的角度来刻划矩阵的秩. 先看一个特殊情形:

命题 3.5.2　一个 n 级方阵 A 的秩等于 n 的充分必要条件为 $|A|\neq 0$.

证明　设 A 的列向量组是 $\beta_1,\cdots,\beta_n$. 我们有

$$\begin{aligned}\mathrm{rank}(A)=n&\Leftrightarrow\mathrm{rank}\{\beta_1,\cdots,\beta_n\}=n\\ &\Leftrightarrow\beta_1,\cdots,\beta_n\text{ 线性无关}\Leftrightarrow|A|\neq 0\end{aligned}$$ ▍

定义 3　一个 n 级方阵的秩如果等于 n,则称它是**满秩矩阵**.

从命题 3.5.2 得,n 级方阵 A 为满秩矩阵 $\Leftrightarrow|A|\neq 0$.

现在我们来看一般的情形:

定理 3.5.4　一个 $s\times n$ 矩阵 A 的秩是 r 的充分必要条件为 A 有一个 r 级子式不为零,并且 A 的所有 $r+1$ 级子式(如有的话)全为零.

证明　先证必要性. 设 $A=(a_{ij})$ 的秩为 r. 于是 A 的行向量组

中有 r 个行向量线性无关. 设 A 的第 $i_1,\cdots,i_r$ 行线性无关,把它们取出来组成 A 的一个子矩阵 A_1. 由于 A_1 的行秩为 r,所以 A_1 的列秩也为 r. 于是 A_1 有 r 列线性无关,设 A_1 的第 $j_1,\cdots,j_r$ 列线性无关,把它们取出来组成 A_1 的一个子矩阵 A_2. 由于 A_2 的列向量组(它们是 A_1 的第 $j_1,\cdots,j_r$ 列)线性无关,并且 A_2 是 r 级方阵,因此,$|A_2| \neq 0$,即,$|A(i_1,\cdots,i_r;j_1,\cdots,j_r)| \neq 0$. 这证明了 A 至少有一个 r 级子式不为零.

任取 A 的一个 $r+1$ 级子式 $|A(k_1,\cdots,k_{r+1};l_1,\cdots,l_{r+1})|$. 因为 A 的秩为 r,所以 A 的任意 $r+1$ 列必线性相关(据 §4 的引理 3.4.1). 于是 A 的第 $l_1,\cdots,l_{r+1}$ 列线性相关. A 的子矩阵 $A(k_1,\cdots,k_{r+1};l_1,\cdots,l_{r+1})$ 的列向量组是 A 的第 $l_1,\cdots,l_{r+1}$ 列的缩短组,据习题 3.3 的第 8 题得,$A(k_1,\cdots,k_{r+1};l_1,\cdots,l_{r+1})$ 的列向量组线性相关. 从而 $|A(k_1,\cdots,k_{r+1};l_1,\cdots,l_{r+1})| = 0$.

再证充分性. 设矩阵 A 有一个 r 级子式不为零,而所有 $r+1$ 级子式全为零. 要证 A 的秩为 r.

设 A 的秩为 t. 假如 $t<r$,则据必要性,A 的所有 $t+1$ 级子式全为零. 由于 A 的任一 $t+2$ 级子式按一行展开等于 $t+2$ 个 $t+1$ 级代数余子式与相应元素的乘积之和,所以 A 的任一 $t+2$ 级子式也为零. 依次类推,A 的所有级数大于 t 的子式全为零,从而 A 的所有 r 级子式全为 0(因为 $r>t$). 这与已知条件矛盾. 假如 $t>r$,由已知条件和上述同样的理由可得,A 的所有 t 级子式全为零. 但是据必要性,A 应当至少有一个 t 级子式不为零,这又是矛盾. 因此 $t=r$,即 A 的秩为 r. ▌

定理 3.5.4 告诉我们,**矩阵 A 的秩等于它的不为零的子式的最高级数**. 这个结论不仅给出了求矩阵的秩的另一种方法,而且在理论上很有用.

推论 3.5.5　设 $s\times n$ 矩阵 A 的秩为 r,则 A 的不为零的 r 级子式所在的列是 A 的列向量组的一个极大线性无关组;不为零的 r 级子式所在的行是 A 的行向量组的一个极大线性无关组.

证明 设 $|A(i_1,\cdots,i_r;j_1,\cdots,j_r)| \neq 0$，于是 $A(i_1,\cdots,i_r;j_1,\cdots,j_r)$ 的列向量组线性无关. 从而 A 的第 $j_1,\cdots,j_r$ 列线性无关. 因此它们就是 A 的列向量组的一个极大线性无关组.

类似方法可证：A 的不为零的 r 级子式所在的行是 A 的行向量组的一个极大线性无关组. ▌

利用推论 3.5.5 以及推论 3.5.1、定理 3.5.2 可以求出一个向量组的许多极大线性无关组. 例如，在例 2 中，已经知道 J 的秩为 2，由于 J 的前两行和第 1,4 列组成的 2 级子式不为零，所以 J 的第 1,4 列是 J 的列向量组的一个极大线性无关组. 从而 α_1,α_4 是向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$ 的一个极大线性无关组. 请读者找出 $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$ 的其它极大线性无关组.

习 题 3.5

1. 求下列向量组的一个极大线性无关组与秩：

(1)

$\alpha_1 = (1,-1,2,3,4)$　　$\alpha_2 = (3,-7,8,9,13)$

$\alpha_3 = (-1,-3,0,-3,-3)$　　$\alpha_4 = (1,-9,6,3,6)$

(2)

$\alpha_1 = (1,-3,2,-1)$　　$\alpha_2 = (-2,1,5,3)$

$\alpha_3 = (4,-3,7,1)$　　$\alpha_4 = (-1,-11,8,-3)$

$\alpha_5 = (2,-12,30,6)$

(3)

$\alpha_1 = (3,0,1,-1,2)$　　$\alpha_2 = (5,0,2,3,-4)$

$\alpha_3 = (1,4,-9,-10,12)$　　$\alpha_4 = (2,1,0,-1,3)$

(4)

$\alpha_1 = (2,-1,1,3)$　　$\alpha_2 = (0,3,1,2)$

$\alpha_3 = (6,0,4,11)$　　$\alpha_4 = (4,-1,2,0)$

$\alpha_5 = (2,1,4,3)$

2. 计算下列矩阵的秩：

(1) $\begin{pmatrix} 3 & -7 & 6 & 1 & 5 \\ 1 & -2 & 4 & -1 & 3 \\ -1 & 1 & -10 & 5 & -7 \\ 4 & -11 & -2 & 8 & 0 \end{pmatrix}$

(2) $\begin{pmatrix} 25 & 31 & 17 & 43 \\ 75 & 94 & 53 & 132 \\ 75 & 94 & 54 & 134 \\ 25 & 32 & 20 & 48 \end{pmatrix}$

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

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

3. K^5 中，求下列向量组生成的线性子空间的一个基与维数：

(1) $\alpha_1 = (1,1,0,0,0)$　　$\alpha_2 = (0,1,1,0,1)$

$\alpha_3 = (1,0,1,1,0)$　　$\alpha_4 = (0,0,0,1,1)$

$\alpha_5 = (0,0,0,0,1)$

(2) $\alpha_1 = (1,-1,2,1,0)$　　$\alpha_2 = (2,-2,4,-2,0)$

$\alpha_3 = (3,0,6,-1,1)$　　$\alpha_4 = (2,1,4,2,1)$

(3) $\alpha_1 = (5,-1,3,1,-2)$　　$\alpha_2 = (1,4,-2,7,9)$

$\alpha_3 = (3,0,4,-1,-5)$　　$\alpha_4 = (0,2,-1,3,4)$

4. 证明：矩阵 A 的任一子矩阵的秩不大于 A 的秩.

5. 对各个不同的 λ 值，矩阵

$$\begin{pmatrix} 1 & \lambda & -1 & 2 \\ 2 & -1 & \lambda & 5 \\ 1 & 10 & -6 & 1 \end{pmatrix}$$

的秩等于多少?

6. 证明:如果 $m \times n$ 矩阵 A 的秩为 r,则它的任何 s 行组成的子矩阵的秩不小于 $r + s - m$.

7. 设 A,B 分别是数域 K 上的 $s \times n, s \times m$ 矩阵.用 (A,B) 表示在 A 的右边添写上 B 得到的 $s \times (n+m)$ 矩阵.如果 $\mathrm{rank}(A) = \mathrm{rank}(A,B)$,则 B 的列向量组可以由 A 的列向量组线性表出;反之也成立.

§6 用矩阵的秩判断线性方程组的解的情况

这一节我们来回答直接从线性方程组的系数和常数项判断线性方程组有无解的问题.

定理 3.6.1 (线性方程组有解判别定理) 线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = b_s \end{cases} \tag{1}$$

有解的充分必要条件为它的系数矩阵 A 与增广矩阵 $\widetilde{A}$ 有相同的秩.

证明 设 $\widetilde{A}$ 的列向量组为 $\alpha_1,\cdots,\alpha_n,\beta$,则 A 的列向量组为 $\alpha_1,\cdots,\alpha_n$. 于是线性方程组(1)可以写成

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta. \tag{1$'$}$$

先证必要性. 设线性方程组(1)有解,则 β 可以由向量组 $\alpha_1,\cdots,\alpha_n$ 线性表出. 于是向量组 $\alpha_1,\cdots,\alpha_n,\beta$ 与向量组 $\alpha_1,\cdots,\alpha_n$ 等价,

从而它们的秩相同. 由此即得

$$\operatorname{rank}(\widetilde{A}) = \operatorname{rank}(A)$$

再证充分性. 设 $\operatorname{rank}(A) = \operatorname{rank}(\widetilde{A})$，则

$$\dim\langle \alpha_1, \cdots, \alpha_n \rangle = \dim\langle \alpha_1, \cdots, \alpha_n, \beta \rangle$$

又由于 $\langle \alpha_1, \cdots, \alpha_n \rangle \subset \langle \alpha_1, \cdots, \alpha_n, \beta \rangle$，所以

$$\langle \alpha_1, \cdots, \alpha_n \rangle = \langle \alpha_1, \cdots, \alpha_n, \beta \rangle$$

因此 $\beta \in \langle \alpha_1, \cdots, \alpha_n \rangle$. 从而线性方程组(1) 有解. ▌

定理 3.6.2 线性方程组(1)有解时，如果它的系数矩阵 A 的秩等于未知量个数 n，则方程组(1) 有唯一解；如果 A 的秩小于 n，则方程组(1) 有无穷多个解.

证明 设方程组(1)的增广矩阵 $\widetilde{A}$ 的列向量组是 $\alpha_1, \cdots, \alpha_n, \beta$；于是 A 的列向量组是 $\alpha_1, \cdots, \alpha_n$. 因为方程组(1) 有解，所以 β 可以由 $\alpha_1, \cdots, \alpha_n$ 线性表出.

$\operatorname{rank}(A) = n \Rightarrow \alpha_1, \cdots, \alpha_n$ 线性无关

$\Rightarrow \beta$ 由 $\alpha_1, \cdots, \alpha_n$ 线性表出的方式唯一

$\Rightarrow$ 方程组(1) 有唯一解

$\operatorname{rank}(A) < n \Rightarrow \alpha_1, \cdots, \alpha_n$ 线性相关

$\Rightarrow \beta$ 由 $\alpha_1, \cdots, \alpha_n$ 线性表出的方式不唯一

$\Rightarrow$ 方程组(1) 有无穷多个解 ▌

这里顺便指出，当我们把方程组(1)的增广矩阵 $\widetilde{A}$ 经过初等行变换化成行阶梯形 $\widetilde{J}$ 时，$\widetilde{J}$ 表示的阶梯形方程组中不包含"$0 = 0$"的方程的数目等于 $\widetilde{J}$ 的非零行的数目，从而等于 $\widetilde{A}$ 的秩. 因此，虽然把一个线性方程组经过初等变换化成的阶梯形方程组不唯一，但是阶梯形方程组中不包含"$0 = 0$"的方程的数目是由原线性方程组唯一决定的.

利用线性方程组(1) 的系数矩阵和增广矩阵的秩去判断方程组 (1) 有没有解?有解时有多少个解?这是直接从原方程组的系数

和常数项去判断解的情况. 这种方法比用阶梯形方程组去判断解的情况优越. 当然用这两种方法去判断线性方程组的解的情况所得结果是一致的,这是因为阶梯形方程组中出现"$0=d$(d 是非零数)"意味着增广矩阵 $\widetilde{A}$ 的行阶梯形 $\widetilde{J}$ 比系数矩阵 A 的行阶梯形 J 多一个非零行,从而 $\widetilde{A}$ 的秩比 A 的秩多 1;而阶梯形方程组中不出现"$0=d$(d 是非零数)"意味着 $\widetilde{J}$ 的非零行数目与 J 的非零行数目一样,从而 $\operatorname{rank}(\widetilde{A})=\operatorname{rank}(A)$.

现在我们用矩阵的秩判断齐次线性方程组有没有非零解:

推论 3.6.1 齐次线性方程组有非零解的充分必要条件是它的系数矩阵的秩小于未知量的个数.

证明 从定理 3.6.2 立即得证. ▍

例 1 判断下述线性方程组有没有解?

$$\begin{cases} x_1+x_2+x_3=1 \\ a\,x_1+b\,x_2+c\,x_3=d \\ a^2x_1+b^2x_2+c^2x_3=d^2 \\ a^3x_1+b^3x_2+c^3x_3=d^3 \end{cases} \tag{2}$$

其中 a,b,c,d 各不相同.

解 线性方程组(2)的增广矩阵 $\widetilde{A}$ 的行列式是范德蒙行列式,由于 a,b,c,d 各不相同,因此 $|A|\neq 0$. 从而 $\widetilde{A}$ 是满秩矩阵,即 $\operatorname{rank}(\widetilde{A})=4$.

方程组(2)的系数矩阵 A 只有 3 列,因此 $\operatorname{rank}(A)\leqslant 3$. 所以 $\operatorname{rank}(A)\neq\operatorname{rank}(\widetilde{A})$. 从而方程组(2)无解.

例 2 讨论 a 取什么值时,下述线性方程组有解?

$$\begin{cases} ax_1+x_2+x_3=1 \\ x_1+ax_2+x_3=1 \\ x_1+x_2+ax_3=1 \end{cases} \tag{3}$$

解 方程组(3)的方程数目等于未知量数目. 先算它的系数矩阵 A 的行列式:

$$|A|=\begin{vmatrix}a&1&1\\1&a&1\\1&1&a\end{vmatrix}=\begin{vmatrix}a+2&1&1\\a+2&a&1\\a+2&1&a\end{vmatrix}$$

$$=(a+2)\begin{vmatrix}1&1&1\\1&a&1\\1&1&a\end{vmatrix}=(a+2)\begin{vmatrix}1&1&1\\0&a-1&0\\0&0&a-1\end{vmatrix}$$

$$=(a+2)(a-1)^2$$

1)当 $a\neq -2$,并且 $a\neq 1$ 时,$|A|\neq 0$,从而方程组(3)有唯一解.

2)当 $a=-2$ 时,$|A|=0$,从而 $\operatorname{rank}(A)<3$. 此时增广矩阵 $\widetilde{A}$ 为

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

$\widetilde{A}$ 的第 2,3,4 列组成的 3 级子式

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

由此得 $\operatorname{rank}(\widetilde{A})=3$. 因此 $\operatorname{rank}(A)\neq\operatorname{rank}(\widetilde{A})$. 从而方程组(3)当 $a=-2$ 时无解.

3)当 $a=1$ 时,系数矩阵 A 与增广矩阵 $\widetilde{A}$ 分别成为

$$A=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix},\qquad \widetilde{A}=\begin{pmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\end{pmatrix}$$

易看出,$\operatorname{rank}(A)=\operatorname{rank}(\widetilde{A})=1$. 所以当 $a=1$ 时,方程组(3)有无穷多个解.

例 3 当 λ 取何值时，齐次线性方程组

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

有非零解？这时求出它的一般解.

解 先计算方程组(4)的系数矩阵 A 的行列式

$$|A| = \begin{vmatrix}\lambda+3 & 1 & 2\\ \lambda & \lambda-1 & 1\\ 3(\lambda+1) & \lambda & \lambda+3\end{vmatrix} \xlongequal[\text{①}+\text{③}\cdot(-1)]{\text{①}+\text{②}\cdot(-1)} \begin{vmatrix}\lambda & 1 & 2\\ 0 & \lambda-1 & 1\\ \lambda & \lambda & \lambda+3\end{vmatrix} = \lambda^2(\lambda-1)$$

方程组 (4) 有非零解 $\Leftrightarrow |A| = 0 \Leftrightarrow \lambda = 0$ 或 $\lambda = 1$.

当 $\lambda = 0$ 时，易求出方程组(4) 的一般解为

$$\begin{cases}x_1 = -x_3\\ x_2 = x_3\end{cases}$$ 其中 x_3 为自由未知量

当 $\lambda = 1$ 时，易求出方程组(4) 的一般解为

$$\begin{cases}x_1 = -x_3\\ x_2 = 2x_3\end{cases}$$ 其中 x_3 为自由未知量

习 题 3.6

1. 判断下述线性方程组有无解？

$$\begin{cases} x_1+x_2+x_3=1 \\ 3x_1+5x_2+2x_3=4 \\ 9x_1+25x_2+4x_3=16 \\ 27x_1+125x_2+8x_3=64 \end{cases}$$

2. 讨论 λ 取什么值时，下述线性方程组有解？并且求出解.

$$\begin{cases} (\lambda+1)x_1+x_2+x_3=\lambda^2+3\lambda \\ x_1+(\lambda+1)x_2+x_3=\lambda^3+3\lambda^2 \\ x_1+x_2+(\lambda+1)x_3=\lambda^4+3\lambda^3 \end{cases}$$

3. 讨论 a,b 取什么值时，下述线性方程组有解？并且求解.

$$\begin{cases} ax_1+x_2+x_3=2 \\ x_1+bx_2+x_3=1 \\ x_1+2bx_2+x_3=2 \end{cases}$$

*4. 讨论 a,b,c 满足什么条件时，下述线性方程组有解？并且求解.

$$\begin{cases} ax+y+z=1 \\ x+by+z=1 \\ x+y+cz=1 \end{cases}$$

5. 当 λ 取何值时，下述齐次线性方程组有非零解？并且求出它的一般解：

$$\begin{cases} (\lambda-2)x_1-3x_2-2x_3=0 \\ -x_1+(\lambda-8)x_2-2x_3=0 \\ 2x_1+14x_2+(\lambda+3)x_3=0 \end{cases}$$

6. 已知线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases} \tag{5}$$

的系数矩阵 A 的秩等于矩阵

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \\ b_1 & b_2 & \cdots & b_n & 0 \end{pmatrix}.$$

的秩，求证方程组(5)有解.

§7 齐次线性方程组的解的结构·解空间

本节讨论齐次线性方程组的解的结构.

数域 K 上 n 元齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = 0 \end{cases} \tag{1}$$

的一个解 $(c_1,c_2,\cdots,c_n)$ 是 K^n 中的一个向量，称为方程组(1)的一个**解向量**. 方程组(1)的所有解组成的集合 W 是 K^n 的一个非空子集. W 具有什么样的性质？

让我们先看实数域 R 上三元齐次线性方程组的解集 W 的几何意义. 一个三元齐次线性方程表示过原点的一个平面，因此 W 是方程组中各个方程表示的平面的交. 于是 W 或者是原点，或者是过原点的一条直线，或者是过原点的一个平面. 即 W 或者是 R^3 的零子空间，或者是 R^3 的 1 维子空间，或者是 R^3 的 2 维子空间. 一般地，我们有

定理 3.7.1　数域 K 上 n 元齐次线性方程组(1)的解集 W 是 K^n 的一个线性子空间，称它为方程组(1)的解空间.

证明 设方程组(1)的系数矩阵的列向量组是 $\alpha_1,\cdots,\alpha_n$. 因为(1)必有零解,所以 W 非空.

设 $\gamma=(c_1,c_2,\cdots,c_n);\delta=(d_1,d_2,\cdots,d_n)\in W$,则有

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n=0$$
$$d_1\alpha_1+d_2\alpha_2+\cdots+d_n\alpha_n=0$$

把上两式相加得

$$(c_1+d_1)\alpha_1+(c_2+d_2)\alpha_2+\cdots+(c_n+d_n)\alpha_n=0$$

这说明 $\gamma+\delta=(c_1+d_1,c_2+d_2,\cdots,c_n+d_n)$ 是方程组(1)的一个解,因此 $\gamma+\delta\in W$.

设 $k\in K,\gamma=(c_1,\cdots,c_n)\in W$. 类似地可证:$k\gamma=(kc_1,\cdots,kc_n)$ 是方程组(1)的一个解. 因此 $k\gamma\in W$.

由上述得,W 是 K^n 的一个线性子空间. ▌

如果方程组(1)只有零解,则它的解空间是零子空间. 如果方程组(1)有非零解,则它的解空间是非零子空间,从而它有基.

定义 1 数域 K 上 n 元齐次线性方程组(1)有非零解时,它的解空间的一个基称为方程组(1)的一个**基础解系**.

用基的定义可把齐次线性方程组基础解系的定义叙述成:

定义 1′ 数域 K 上 n 元齐次线性方程组(1)有非零解时,它的有限多个解 $\eta_1,\cdots,\eta_t$ 称为方程组(1)的一个**基础解系**,如果

1) $\eta_1,\cdots,\eta_t$ 线性无关;

2) 方程组(1)的每一个解都可以由 $\eta_1,\cdots,\eta_t$ 线性表出.

如果我们找到了齐次线性方程组(1)的一个基础解系:$\eta_1,\cdots,\eta_t$,那么方程组(1)的每一个解都可以唯一地表示成 $\eta_1,\cdots,\eta_t$ 的一个线性组合:$k_1\eta_1+\cdots+k_t\eta_t$,其中 $k_i\in K;i=1,\cdots,t$. 于是齐次线性方程组(1)的全部解是集合

$$\{k_1\eta_1+\cdots+k_t\eta_t\,|\,k_i\in K,\quad i=1,\cdots,t\}$$

中的全部元素. 这说明:只要找到了齐次线性方程组(1)的一个基础解系,方程组(1)的解的结构就清楚了. 这时方程组(1)的解空间 $W=\langle\eta_1,\eta_2,\cdots,\eta_t\rangle$.

下面的定理 3.7.2 的证明过程给出了求齐次线性方程组的一个基础解系的方法，并且指出了每一个基础解系所含解向量的数目.

定理 3.7.2　若数域 K 上的 n 元齐次线性方程组(1) 的系数矩阵 A 的秩 $r < n$，则它一定有基础解系，并且它的每一个基础解系包含的解向量的数目都等于 $n-r$.

证明　此时方程组(1)有非零解，它的解空间 W 是 K^n 的一个非零线性子空间，从而 W 一定有基. 因此方程组(1) 一定有基础解系，并且每个基础解系包含的解向量的数目等于 $\dim W$. 只要具体找出一个基础解系便可知道 $\dim W$.

把 A 经过初等行变换化成简化行阶梯形 J. 因为 $\mathrm{rank}(A) = r$，所以 J 有 r 个非零行. 从而 J 有 r 个主元，不妨设它们分别在第 $1,2,\cdots,r$ 列，即

$$J=\begin{pmatrix} 1 & 0 & \cdots & 0 & b_{1,r+1} & \cdots & b_{1n} \\ 0 & 1 & \cdots & 0 & b_{2,r+1} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & b_{r,r+1} & \cdots & b_{rn} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

于是立即得到齐次线性方程组(1)的一般解为

$$\begin{cases} x_1 = -b_{1,r+1}x_{r+1} - \cdots - b_{1n}x_n \\ x_2 = -b_{2,r+1}x_{r+1} - \cdots - b_{2n}x_n \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ x_r = -b_{r,r+1}x_{r+1} - \cdots - b_{rn}x_n \end{cases} \tag{2}$$

其中 $x_{r+1},\cdots,x_n$ 为自由未知量.

让自由未知量 $(x_{r+1},\cdots,x_n)$ 分别取下述 $n-r$ 组数：

$$\begin{array}{l}(1,0,0,\cdots,0,0)\\(0,1,0,\cdots,0,0)\\\cdots\quad\cdots\quad\cdots\quad\cdots\\(0,0,0,\cdots,0,1)\end{array}\tag{3}$$

则得到方程组 (1) 的 $n-r$ 个解为：

$$\eta_1=\begin{pmatrix}-b_{1,r+1}\\-b_{2,r+1}\\\vdots\\-b_{r,r+1}\\1\\0\\\vdots\\0\end{pmatrix},\quad \eta_2=\begin{pmatrix}-b_{1,r+2}\\-b_{2,r+2}\\\vdots\\-b_{r,r+2}\\0\\1\\\vdots\\0\end{pmatrix},\cdots,\eta_{n-r}=\begin{pmatrix}-b_{1n}\\-b_{2n}\\\vdots\\-b_{rn}\\0\\0\\\vdots\\1\end{pmatrix}$$

因为

$$\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\quad\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\quad\cdots,\quad\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$$

线性无关，所以它们的延伸组 $\eta_1,\eta_2,\cdots,\eta_{n-r}$ 线性无关(据 §3 例 7). 为了证 $\eta_1,\eta_2,\cdots,\eta_{n-r}$ 是方程组(1) 的一个基础解系. 只还需要证方程组(1) 的每一个解 $\alpha=(c_1,\cdots,c_n)$ 可以由 $\eta_1,\eta_2,\cdots,\eta_{n-r}$ 线性表出. 因为 α 是(1) 的解，所以 α 满足方程组(1) 的一般解公式 (2)，即

$$\begin{cases}c_1=-b_{1,r+1}c_{r+1}-\cdots-b_{1n}c_n\\c_2=-b_{2,r+1}c_{r+1}-\cdots-b_{2n}c_n\\\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\c_r=-b_{r,r+1}c_{r+1}-\cdots-b_{rn}c_n\end{cases}\tag{4}$$

于是解向量 $\alpha=(c_1,c_2,\cdots,c_r,c_{r+1},\cdots,c_n)$ 可以写成下述形式：

$$
\alpha=\begin{pmatrix} c_1 \\ \vdots \\ c_r \\ c_{r+1} \\ \vdots \\ c_n \end{pmatrix}=\begin{pmatrix} -b_{1,r+1}c_{r+1}-\cdots-b_{1n}c_n \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ -b_{r,r+1}c_{r+1}-\cdots-b_{rn}c_n \\ 1\cdot c_{r+1}\quad+\cdots+0\cdot c_n \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \\ 0\cdot c_{r+1}\quad+\cdots+1\cdot c_n \end{pmatrix}
$$

$$
=\begin{pmatrix} -b_{1,r+1} \\ \vdots \\ -b_{r,r+1} \\ 1 \\ \vdots \\ 0 \end{pmatrix}c_{r+1}+\cdots+\begin{pmatrix} -b_{1n} \\ \vdots \\ -b_{rn} \\ 0 \\ \vdots \\ 1 \end{pmatrix}c_n
$$

$$
=c_{r+1}\eta_1+\cdots+c_n\eta_{n-r}
$$

这证明了方程组 (1) 的每一个解 α 可以由 $\eta_1,\cdots,\eta_{n-r}$ 线性表出. 因此 $\eta_1,\cdots,\eta_{n-r}$ 是方程组(1) 的一个基础解系. 从而 $\dim W=n-r$.

∎

注：当 n 元齐次线性方程组(1) 只有零解时,它的解空间 W 的维数是零,它的系数矩阵的秩 $r=n$. 因此公式 $\dim W=n-r$ 对于方程组(1) 只有零解时也成立.

从定理 3.7.2 的证明过程看出:求齐次线性方程组(1)的一个基础解系的步骤如下:

第一步：把方程组(1) 的系数矩阵 A 经过初等行变换化成简化行阶梯形 J;

第二步：从 J 直接可以写出方程组(1) 的一般解公式;

第三步：在(1) 的一般解公式中,每一次让一个自由未知量取值为 1,其余自由未知量取值为 0,求出方程组(1) 的一个解. 这样得到的 $n-r$ 个解就构成方程组(1) 的一个基础解系,其中 $r=\operatorname{rank}(A)$.

从定理 3.7.2 的证明过程还可以看出，齐次线性方程组(1)

的一般解公式(2)中出现的自由未知量的数目等于 $n-r$，其中 r 是方程组(1)的系数矩阵的秩.

在定理 3.7.2 的证明过程中，我们让 $n-r$ 个自由未知量 $(x_{r+1},\cdots,x_n)$ 取(3)所表示的 $n-r$ 组数，从而得到方程组(1)的 $n-r$ 个解 $\eta_1,\cdots,\eta_{n-r}$，它们构成方程组(1)的一个基础解系. 如果我们让自由未知量 $(x_{r+1},\cdots,x_n)$ 分别取

$$\begin{array}{c}(d_{11},\cdots,d_{1,n-r})\\ \cdots\quad\cdots\quad\cdots\quad\cdots\\ (d_{n-r,1},\cdots,d_{n-r,n-r})\end{array}\tag{5}$$

使得这 $n-r$ 个向量是线性无关的，则由它们代入一般解公式(3)所得到的 $n-r$ 个解 $\gamma_1,\cdots,\gamma_{n-r}$ 一定线性无关(仍然根据 §3 例 7). 由于方程组(1)的解空间 W 的维数是 $n-r$，所以 $\gamma_1,\cdots,\gamma_{n-r}$ 是 W 的一个基，从而 $\gamma_1,\cdots,\gamma_{n-r}$ 也是方程组(1)的一个基础解系. 为了计算简便，通常让自由未知量 $(x_{r+1},\cdots,x_n)$ 取如下的 $n-r$ 组数

$$(d_1,0,\cdots,0),(0,d_2,0,\cdots,0),\cdots,(0,\cdots,0,d_{n-r})$$

其中 $d_1,\cdots,d_{n-r}$ 是适当的非零数. 由于这 $n-r$ 个向量是线性无关的，因此由它们得到的方程组(1)的 $n-r$ 个解一定是一个基础解系.

例 1 求下述齐次线性方程组的一个基础解系：

$$\begin{cases}x_1-3x_2+5x_3-2x_4+x_5=0\\ -2x_1+x_2-3x_3+x_4-4x_5=0\\ -x_1-7x_2+9x_3-4x_4-5x_5=0\\ 3x_1-14x_2+22x_3-9x_4+x_5=0\end{cases}\tag{6}$$

解 把方程组(6)的系数矩阵经过初等行变换化成简化行阶梯形矩阵：

$$
\begin{pmatrix} 1 & -3 & 5 & -2 & 1 \\ -2 & 1 & -3 & 1 & -4 \\ -1 & -7 & 9 & -4 & -5 \\ 3 & -14 & 22 & -9 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & \frac{4}{5} & -\frac{1}{5} & \frac{11}{2} \\ 0 & 1 & -\frac{7}{5} & \frac{3}{5} & \frac{2}{5} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}
$$

于是方程组(6)的一般解为

$$
\begin{cases} x_1 = -\frac{4}{5}x_3 + \frac{1}{5}x_4 - \frac{11}{5}x_5 \\ x_2 = \quad \frac{7}{5}x_3 - \frac{3}{5}x_4 - \frac{2}{5}x_5 \end{cases} \tag{7}
$$

其中 x_3, x_4, x_5 是自由未知量.

让 (x_3, x_4, x_5) 分别取$(5,0,0)$, $(0,5,0)$, $(0,0,5)$ 代入一般解公式(7) 中可求出方程组(6) 的三个解:

$$
\eta_1 = \begin{pmatrix} -4 \\ 7 \\ 5 \\ 0 \\ 0 \end{pmatrix}, \quad \eta_2 = \begin{pmatrix} 1 \\ -3 \\ 0 \\ 5 \\ 0 \end{pmatrix}, \quad \eta_3 = \begin{pmatrix} -11 \\ -2 \\ 0 \\ 0 \\ 5 \end{pmatrix}
$$

η_1, η_2, η_3 就是方程组(6) 的一个基础解系.

习 题 3.7

1. 求下列齐次线性方程组的一个基础解系:

(1) $$
\begin{cases} x_1 - 3x_2 + x_3 - 2x_4 = 0 \\ -5x_1 + x_2 - 2x_3 + 3x_4 = 0 \\ -x_1 - 11x_2 + 2x_3 - 5x_4 = 0 \\ 3x_1 + 5x_2 + x_4 = 0 \end{cases}
$$

(2) $\begin{cases} 2x_1 - 5x_2 + x_3 - 3x_4 = 0 \\ -3x_1 + 4x_2 - 2x_3 + x_4 = 0 \\ x_1 + 2x_2 - x_3 + 3x_4 = 0 \\ -2x_1 + 15x_2 - 6x_3 + 13x_4 = 0 \end{cases}$

(3) $\begin{cases} x_1 - 3x_2 + x_3 - 2x_4 - x_5 = 0 \\ -3x_1 + 9x_2 - 3x_3 + 6x_4 + 3x_5 = 0 \\ 2x_1 - 6x_2 + 2x_3 - 4x_4 - 2x_5 = 0 \\ 5x_1 - 15x_2 + 5x_3 - 10x_4 - 5x_5 = 0 \end{cases}$

(4) $\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ 2x_1 + x_2 - 4x_5 = 0 \\ x_2 + 2x_3 + 2x_4 + 6x_5 = 0 \\ 4x_1 + 3x_2 + 2x_3 + 2x_4 - 2x_5 = 0 \end{cases}$

(5) $\begin{cases} x_1 - 3x_2 - 2x_3 + 2x_4 - 3x_5 = 0 \\ 2x_1 + x_2 - x_3 + 2x_4 - 3x_5 = 0 \\ 3x_1 - 2x_2 - x_3 + x_4 - 2x_5 = 0 \\ -6x_2 + 2x_3 - 4x_4 + 5x_5 = 0 \end{cases}$

(6) $\begin{cases} 3x_1 - x_2 + 2x_3 + x_4 = 0 \\ x_1 + 3x_2 - x_3 + 2x_4 = 0 \\ -2x_1 + 5x_2 + x_3 - x_4 = 0 \\ 3x_1 + 10x_2 + x_3 + 4x_4 = 0 \\ -2x_1 + 15x_2 - 4x_3 + 4x_4 = 0 \end{cases}$

2. 证明:与基础解系等价的线性无关向量组也是基础解系.

3. 设 n 元齐次线性方程组(1) 的系数矩阵的秩为 $r(< n)$,证明:方程组(1) 的任意 $n-r$ 个线性无关的解都是它的一个基础解系.

4. 证明:如果齐次线性方程组的秩比未知量个数少 1,则该方程组的任何两个解成比例.

5. 设 n 个方程的 n 元齐次线性方程组的系数矩阵 $A=(a_{ij})$ 的行列式等于零,并且 A 有一个元素 a_{kl} 的代数余子式 $A_{kl}\neq 0$. 证明:$\eta=(A_{k1},A_{k2},\cdots,$

$A_{kl},\cdots,A_{kn}$) 是这个方程组的一个基础解系.

*6.　证明：如果 $n(>1)$ 级矩阵 A 的行列式等于零，则 A 的任何两行(或两列)对应元素的代数余子式成比例.

7.　设 n 元齐次线性方程组的方程个数为 $n-1$，从它的系数矩阵中划掉第 j 列得到的子矩阵行列式记作 D_j，证明

(1)　$(D_1,-D_2,\cdots,(-1)^{n-1}D_n)$ 是这个方程组的一个解；

(2)　如果上述解不是零解，则这个方程组的每个解都是上述解的倍数.

8.　设 A_1 是 $s\times n$ 矩阵 $A=(a_{ij})$ 的前 $s-1$ 行组成的子矩阵. 证明：如果以 A_1 为系数矩阵的齐次线性方程组的解都是方程 $a_{s1}x_1+a_{s2}x_2+\cdots+a_{sn}x_n=0$ 的解，则 A 的第 s 行可以由 A 的前 $s-1$ 行线性表出.

§8　非齐次线性方程组的解的结构·线性流形

这一节我们来讨论非齐次线性方程组在有解时它的解的结构. 设数域 K 上 n 元非齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1+a_{s2}x_2+\cdots+a_{sn}x_n=b_s \end{cases} \tag{1}$$

的增广矩阵 $\widetilde{A}$ 的列向量组为 $\alpha_1,\cdots,\alpha_n,\beta$；于是它的系数矩阵 A 的列向量组是 $\alpha_1,\cdots,\alpha_n$. 把方程组(1)的常数项换成 0，就得到齐次线性方程组

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=0 \tag{2}$$

(2)称为(1)的**导出组**. 方程组(1)的解与它的导出组(2)的解之间有密切关系：

命题 3.8.1　线性方程组(1)的两个解的差是它的导出组(2)的一个解.

证明 设 $\gamma=(c_1,c_2,\cdots,c_n)$ 与 $\delta=(d_1,d_2,\cdots,d_n)$ 是方程组(1) 的两个解. 则有

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n=\beta \tag{3}$$

$$d_1\alpha_1+d_2\alpha_2+\cdots+d_n\alpha_n=\beta \tag{4}$$

(3)式减去(4)式得

$$(c_1-d_1)\alpha_1+(c_2-d_2)\alpha_2+\cdots+(c_n-d_n)\alpha_n=0 \tag{5}$$

(5)式说明

$$\gamma-\delta=(c_1-d_1,c_2-d_2,\cdots,c_n-d_n)$$

是导出组(2)的一个解. ▌

命题 3.8.2 线性方程组(1)的一个解与它的导出组(2)的一个解之和仍是方程组(1)的一个解.

证明 设 $\gamma=(c_1,\cdots,c_n)$ 是方程组(1) 的一个解, $\eta=(e_1,\cdots,e_n)$ 是导出组(2) 的一个解. 则有

$$c_1\alpha_1+\cdots+c_n\alpha_n=\beta \tag{6}$$

$$e_1\alpha_1+\cdots+e_n\alpha_n=0 \tag{7}$$

(6)式和(7)式相加得

$$(c_1+e_1)\alpha_1+\cdots+(c_n+e_n)\alpha_n=\beta \tag{8}$$

(8)式说明

$$\gamma+\eta=(c_1+e_1,\cdots,c_n+e_n)$$

是方程组(1)的一个解. ▌

命题 3.8.1 与命题 3.8.2 说明线性方程组(1)的解与它的导出组(2)的解之间有密切联系. 我们已经知道导出组(2)的所有解组成 K^n 的一个线性子空间 W. 现在要问:线性方程组(1) 的所有解组成的集合有什么样的结构?

先看一个例子:考虑线性方程组

$$x-6y=12 \tag{9}$$

它的解集是平面上的一条直线 L, L 不经过原点,因此 L 不是 R^2 的子空间. (9) 的导出组的解集 W 是过原点且与 L 平行的直线. 取定 (9) 的一个解 α_0, 则 L 可以由 W 沿向量 α_0 作平移得到. 于是 L 中每

个向量可以表示成 $\alpha_0+\eta$ 的形式，其中 $\eta\in W$，如图所示. 所以

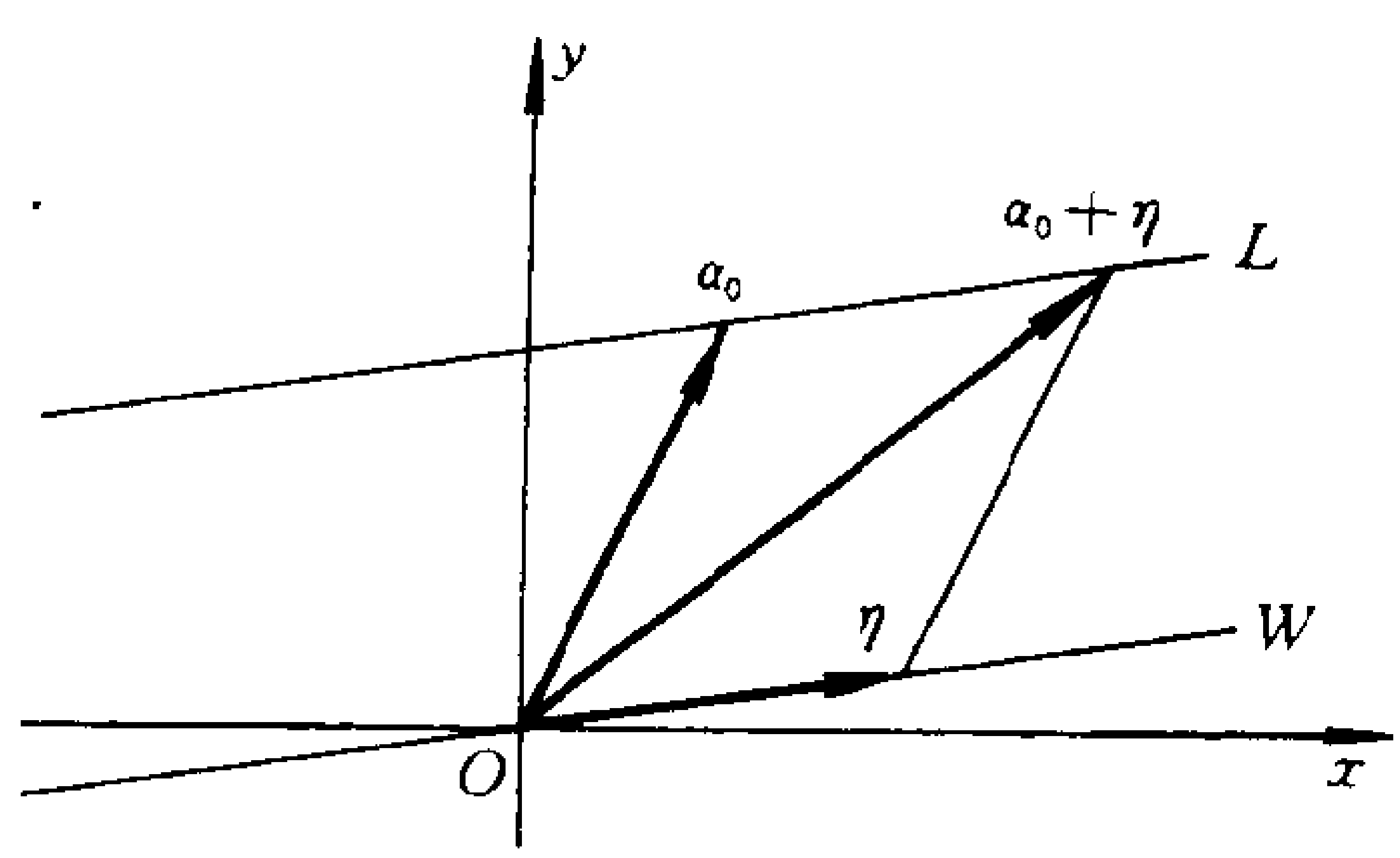
$$L=\{\alpha_0+\eta\mid\eta\in W\} \tag{10}$$

把(10)式右边的集合记作 α_0+W，于是线性方程组(9)的解集是 α_0+W. 由此受到启发引进一个概念：

定义 1 设 W 是 K^n 的一个线性子空间，设 α_0 是 K^n 的一个给定的向量. K^n 的一个子集

$$\alpha_0+W:=\{\alpha_0+\eta\mid\eta\in W\}$$

称为一个 W 型的**线性流形**. 把 $\dim W$ 称为线性流形 α_0+W 的维数.



由于 $W=0+W$，因此子空间 W 本身也是一个线性流形.

定理 3.8.1 如果数域 K 上的 n 元非齐次线性方程组(1)有解，则它的所有解组成的集合是一个 W 型的线性流形 γ_0+W，其中 γ_0 是方程组(1)的一个解(称它为特解)，W 是导出组(2)的解空间.

证明 把方程组(1)的所有解组成的集合记作 S. 据命题 3.8.2 得，$\gamma_0+W\subset S$. 现在任取 $\gamma\in S$，据命题 3.8.1 得，$\gamma-\gamma_0\in W$. 于是 $\gamma=\gamma_0+(\gamma-\gamma_0)\in\gamma_0+W$. 所以 $S\subset\gamma_0+W$. 因此 $S=\gamma_0+W$. ▌

推论 3.8.1 在线性方程组(1)有解的条件下，解是唯一的充

分必要条件为它的导出组(2)只有零解.

证明 设线性方程组(1)有解.据定理 3.8.1 得,它的解集合是 γ_0+W,其中 γ_0 是(1) 的一个特解,W 是导出组(2) 的解空间.于是

$$(1)\text{ 的解唯一} \Leftrightarrow \gamma_0+W=\{\gamma_0\} \Leftrightarrow W=0 \qquad \blacksquare$$

从定理 3.8.1 看出,为了求出非齐次线性方程组(1) 的全部解,就只要求它的一个特解 γ_0 以及它的导出组的解空间 W.而为了得到导出组(2) 的解空间 W,只要求出导出组(2) 的一个基础解系 $\eta_1,\cdots,\eta_{n-r}$,其中 r 是导出组(2) 的系数矩阵 A 的秩.于是方程组(1) 的全部解组成的集合是

$$\gamma_0+W=\{\gamma_0+k_1\eta_1+\cdots+k_{n-r}\eta_{n-r}\mid k_1,\cdots,k_{n-r}\in K\}$$

例 1 求下述数域 K 上线性方程组的全部解:

$$\begin{cases} x_1+3x_2-x_3+2x_4-x_5=-4 \\ -3x_1+x_2+2x_3-5x_4-4x_5=-1 \\ 2x_1-3x_2-x_3-x_4+x_5=4 \\ -4x_1+16x_2+x_3+3x_4-9x_5=-21 \end{cases} \tag{9}$$

解 第一步,求方程组(9)的一个特解 γ_0.为此先求出(9) 的一般解公式.把(9) 的增广矩阵 $\widetilde{A}$ 经过初等行变换化成简化行阶梯形 $\widetilde{J}$:

$$\widetilde{A}\longrightarrow\widetilde{J}=\begin{pmatrix} 1 & 0 & 0 & -27 & -22 & 2 \\ 0 & 1 & 0 & -4 & -4 & -1 \\ 0 & 0 & 1 & -41 & -33 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

于是方程组(9)的一般解是

$$\begin{cases} x_1=27x_4+22x_5+2 \\ x_2=4x_4+4x_5-1 \\ x_3=41x_4+33x_5+3 \end{cases} \tag{10}$$

其中 x_4,x_5 是自由未知量. 令 $x_4=0,x_5=0$，得到方程组(9)的一个特解 $\gamma_0=(2,-1,3,0,0)$.

第二步，求导出组的一个基础解系. 由于方程组(9) 与它的导出组的系数矩阵相同，因此 $\tilde{J}$ 去掉末列就是导出组的系数矩阵 A 的简化行阶梯形. 从而只要把方程组(9) 的一般解公式(10) 的常数项去掉就得到导出组的一般解公式：

$$\begin{cases} x_1=27x_4+22x_5 \\ x_2=\ \ 4x_4+\ \ 4x_5 \\ x_3=41x_4+33x_5 \end{cases} \tag{11}$$

令 $x_4=1,x_5=0$，得 $\eta_1=(27,4,41,1,0)$；令 $x_4=0,x_5=1$，得 $\eta_2=(22,4,33,0,1)$，所以导出组的一个基础解系为 η_1,η_2.

第三步，方程组(9)的全部解为

$$(2,-1,3,0,0)+k_1(27,4,41,1,0)+k_2(22,4,33,0,1)$$

其中 k_1,k_2 是数域 K 中任意数.

习　题　3.8

1. 求下列方程组的全部解

(1) $$\begin{cases} x_1-5x_2+2x_3-3x_4=11 \\ -3x_1+x_2-4x_3+2x_4=-5 \\ -x_1-9x_2\qquad\ -4x_4=17 \\ 5x_1+3x_2+6x_3-x_4=-1 \end{cases}$$

(2) $$\begin{cases} 2x_1-3x_2+x_3-5x_4=1 \\ -5x_1-10x_2-2x_3+x_4=-21 \\ x_1+4x_2+3x_3+2x_4=1 \\ 2x_1-4x_2+9x_3-3x_4=-16 \end{cases}$$

(3) $x_1 - 4x_2 + 2x_3 - 3x_4 + 6x_5 = 4$

2. 试证：n 个方程的 n 元线性方程组有唯一解的充分必要条件是它的导出组只有零解.

3. 试证：如果 $\gamma_1,\gamma_2,\cdots,\gamma_t$ 都是线性方程组(1)的解，并且一组数 $u_1,u_2,\cdots,u_t$ 满足

$$u_1 + u_2 + \cdots + u_t = 1$$

则 $u_1\gamma_1 + u_2\gamma_2 - \cdots + u_t\gamma_t$ 也是方程组(1)的一个解.

4. 试证：如果 γ_0 是线性方程组的一个特解，$\eta_1,\cdots,\eta_t$ 是它的导出组的一个基础解系. 令

$$\gamma_1 = \gamma_0 + \eta_1, \quad \gamma_2 = \gamma_0 + \eta_2, \quad \cdots, \quad \gamma_t = \gamma_0 + \eta_t$$

则线性方程组(1)的任意一个解 γ 可以表示成

$$\gamma = u_0\gamma_0 + u_1\gamma_1 + \cdots + u_t\gamma_t$$

其中 $u_0 + u_1 + \cdots + u_t = 1$.

*5. 设 W 是 K^n 的一个线性子空间，并且设 α_1 是线性流形 $\alpha_0 + W$ 里的任一向量，证明：$\alpha_1 + W = \alpha_0 + W$.（注：我们把线性流形 $\alpha_0 + W$ 中的 α_0 称为**流形的代表**. 此题表明：线性流形 $\alpha_0 + W$ 中的任一向量都可以作为这个流形的代表.）

*6. 设 W 是 K^n 的一个线性子空间，两个 W 型的线性流形 $\alpha_1 + W$ 与 $\alpha_2 + W$ 相等的充分必要条件是 $\alpha_1 - \alpha_2 \in W$.

*7. 在 K^5 中，设 $W = \langle \varepsilon_1,\varepsilon_2,\varepsilon_3 \rangle$. 判断线性流形 $\alpha_1 + W$ 与 $\alpha_2 + W$ 是否相等？

(1) $\alpha_1 = (0,0,1,1,0)$ $\qquad \alpha_2 = (0,0,0,1,0)$

(2) $\alpha_1 = (2,-1,0,5,1)$ $\qquad \alpha_2 = (-3,0,2,5,1)$

(3) $\alpha_1 = (2,-1,0,5,1)$ $\qquad \alpha_2 = (-3,0,2,5,0)$

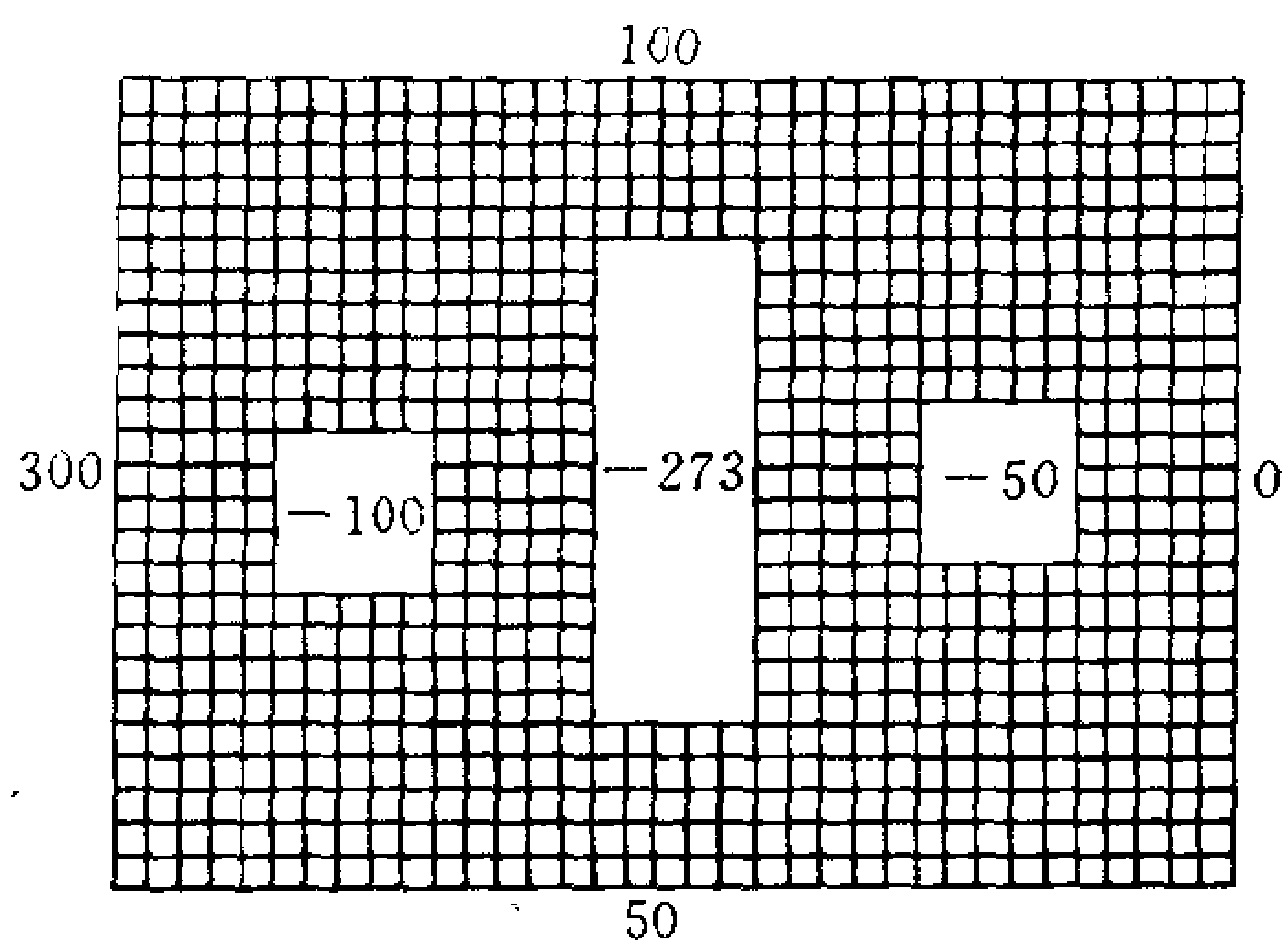
*8. 在 R^3 中，设 $\eta_1 = (1,3,2)$，$\eta_2 = (-2,1,3)$，$W = \langle \eta_1,\eta_2 \rangle$，$\alpha_0 = (0,1,2)$. 试问：线性流形 $\alpha_0 + W$ 是几何空间中的什么图形？

* §9 一个实际问题·线性方程组理论在几何上的应用

线性方程组的理论无论在数学的各个分支中，还是在自然科

学、工程技术以及生产实际中都有广泛的应用. 这一节首先我们举一个实际问题的例子.

平板的受热问题 一个带有三个孔的矩形平板作为阀门用于产生低温的一个理想装置中. 平板由正方形的方格组成的网复盖着. 网的周界上的顶点叫做边界顶点, 而所有其它的顶点叫做内部顶点. 内部顶点有 416 个, 边界顶点有 204 个. 实验表明, 当加热或冷却时, 任一内部顶点的温度是离它最近的四个顶点(内部顶点或者边界顶点)的温度的算术平均值. 我们希望沿着周界的顶点的温度取下图所示的值, 这是可能的吗? 如果是可能的, 试问内部顶点的温度分布是不是唯一确定的?



解 给平板的内部顶点编号, 从 1 标到 416; 接着给边界顶点编号, 从 417 标到 620. 顶点 $i(1 \leqslant i \leqslant 620)$ 的温度设为 t_i. 任取一个内部顶点 i, 设离它最近的四个顶点是 a, b, c, d. 由已知条件得

$$t_i = \frac{t_a + t_b + t_c + t_d}{4} \tag{1}$$

如果 a, b, c, d 都是内部顶点, 则把(1) 改写成

$$4t_i - t_a - t_b - t_c - t_d = 0 \tag{2}$$

如果 a, b, c 是内部顶点, 而 d 是边界顶点, 则把(1) 写成

$$4t_i - t_a - t_b - t_c = t_d \tag{3}$$

其中 $t_d \in \{300, 100, 50, 0, -273, -100, -50\}$.

如果 a,b 是内部顶点，而 c,d 是边界顶点，则把(1) 改写成

$$4t_i - t_a - t_b = t_c + t_d \tag{4}$$

其中 $t_c, t_d \in \{300, 100, 50, 0, -273, -100, -50\}$.

总之，任取一个内部顶点 i，可以列出一个一次方程，它形如(2)或(3)或(4)，等等. 其中未知量(内部顶点的温度)的系数或者是 4，或者是 -1 或者是 0. 这样我们可以列出 416 个方程，它们组成一个 416 元的线性方程组，记作 Ⅰ. 所问的问题就是：线性方程组 Ⅰ 有解吗？如果有解，解是唯一吗？直接讨论线性方程组 Ⅰ 的解的情况是困难的. 我们可以先讨论 Ⅰ 的导出组(记作 Ⅱ)的解的情况，然后利用线性方程组的解与它的导出组的解之间的关系来了解方程组 Ⅰ 的解的情况.

线性方程组 Ⅰ 的导出组 Ⅱ 就是把 Ⅰ 的常数项全换成 0. 换句话说，就是把所有边界顶点的温度都取为 0. 设

$$m = \max\{|t_i| \mid 1 \leqslant i \leqslant 416\}$$

设 e 是一个内部顶点，它的温度 t_e 满足 $|t_e| = m$. 由已知条件

$$t_e = \frac{t_a + t_b + t_c + t_d}{4}$$

得

$$\begin{aligned} 4m = 4|t_e| &= |t_a + t_b + t_c + t_d| \\ &\leqslant |t_a| + |t_b| + |t_c| + |t_d| \leqslant 4m \end{aligned}$$

所以 $|t_a| + |t_b| + |t_c| + |t_d| = 4m$. 由此得 $|t_a| = |t_b| = |t_c| = |t_d| = m$. 因此，若一个内部顶点的温度的绝对值为 m，则离它最近的四个顶点的温度的绝对值都是 m. 从内部顶点 e 出发，在四个方向的每一个方向上，一次移动一个顶点，据上述结论得，我们走过的每一个顶点的温度的绝对值都是 m. 最后我们走到一个边界顶点，由边界顶点的温度为 0，于是得出 $m = 0$. 由于上述移动可走遍所有内部顶点，因此我们得出对所有内部顶点 i 都有 $t_i = 0$. 这说明导出组 Ⅱ 只有零解. 因此导出组 Ⅱ 的系数矩阵 A 的行列式不等于零. 由于线性方程组 Ⅰ 与它的导出组 Ⅱ 有相同的系数矩阵

A,从 $|A| \neq 0$ 便得出方程组 I 有解,并且解是唯一的.这就解决了平板受热的问题:对于所给的边界顶点的温度要求,存在内部顶点的一个温度分布,并且这个温度分布是唯一确定的.

线性方程组的理论与几何有密切联系,现在我们举例说明如何利用线性方程组的理论解决一些几何问题.

例 1 求 n 个平面

$$a_ix + b_iy + c_iz + d_i = 0 \qquad (i = 1, \cdots, n)$$

通过一直线但不合并为一个平面的充分必要条件.

解 n 个平面

$$a_ix + b_iy + c_iz + d_i = 0 \qquad (i = 1, \cdots, n) \tag{5}$$

通过一直线但不合并为一个平面

$\Leftrightarrow$线性方程组(5)有解,并且解集是1维线性流形

$\Leftrightarrow$方程组(5)有解,且它的导出组的解空间是1维的

$\Leftrightarrow$方程组(5)有解,且导出组的系数矩阵 A 的秩是 2

$\Leftrightarrow$方程组(5)的系数矩阵 A 与增广矩阵 $\widetilde{A}$ 的秩都是 2

$\Leftrightarrow$方程组(5)的系数矩阵 A 与下述矩阵的秩都是 2

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ \cdots & \cdots & \cdots & \cdots \\ a_n & b_n & c_n & d_n \end{pmatrix}$$

例 2. 求平面上通过五点 $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), (x_5, y_5)$ 的二次曲线的方程.

解 设通过上述五点的二次曲线 C 的方程为

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{6}$$

于是有

$$ax_i^2 + bx_iy_i + cy_i^2 + dx_i + ey_i + f = 0, \quad i = 1, \cdots, 5 \tag{7}$$

点 $M(x, y)$ 在二次曲线 C 上

$\Leftrightarrow$点 M 的坐标(x, y)适合 C 的方程,即有

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{8}$$

$\Leftrightarrow a,b,c,d,e,f$ 为未知量的齐次线性方程组

$$\begin{cases} ax^2+bxy+cy^2+dx+ey+f=0 \\ ax_1^2+bx_1y_1+cy_1^2+dx_1+ey_1+f=0 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ ax_5^2+bx_5y_5+cy_5^2+dx_5+ey_5+f=0 \end{cases} \tag{9}$$

有非零解并且非零解的前三个分量不全为零

$\Leftrightarrow$方程组(9)的系数矩阵 A 的行列式等于零,即

$$\begin{vmatrix} x^2 & xy & y^2 & x & y & 1 \\ x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & 1 \end{vmatrix} = 0 \tag{10}$$

并且 A 的 $(1,j)$ 元的代数余子式 $A_{1j}(j=1,2,3)$ 中至少有一个不为零.(注意:当(10)式成立时,$(A_{11},A_{12},\cdots,A_{16})$ 是方程组(9)的一个解.)

因此,(10)式是所求的二次曲线方程.

*习　题　3.9

1. 求通过五点:$(3,0)$,$(-3,0)$,$(5,6\frac{2}{3})$,$(5,-6\frac{2}{3})$,$(-5,-6\frac{2}{3})$的二次曲线的方程并且确定其类型.

2. 求通过五点:$(0,1)$,$(\pm 2,0)$,$(\pm 1,-1)$,的二次曲线的方程并且确定其类型,位置和形状.

3. 利用矩阵的秩讨论几何空间中两个平面的位置关系;并且在两个平面的交不是空集时,指出这两个平面的交是多少维的线性流形?

4. 讨论几何空间中三个平面的相关位置的所有可能的情况,画出每种

情况的示意图；并且在三个平面的交不是空集时，指出它是几维的线性流形？

5. 在 xOy 平面上，下列四条直线是否相交于一点？

$$l_1: x - y = 3, \qquad l_2: 2x + y = 3$$

$$l_3: x - 4y = 6, \qquad l_4: \sqrt{2}\,x - y = 2\sqrt{2} + 1$$

6. 证明契维定理：若三角形的三边依次被分割成

$$\lambda : \mu, \quad \upsilon : \lambda, \quad \mu : \upsilon$$

其中 λ, μ, υ 均为正实数，则此三角形的顶点与对边分点的连线交于一点.

7. 证明：如果直线 l_1

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases}$$

与直线 l_2

$$\begin{cases} A_3x + B_3y + C_3z + D_3 = 0 \\ A_4x + B_4y + C_4z + D_4 = 0 \end{cases}$$

相交，则

$$\begin{vmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A_3 & B_3 & C_3 & D_3 \\ A_4 & B_4 & C_4 & D_4 \end{vmatrix} = 0$$

8. 求使平面上三点 $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ 位于一条直线上的充分必要条件.

9. 求使平面上 n 个点 $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ 位于一条直线上的充分必要条件.

10. 求一平面内不在一直线上的四点 $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)$ 位于一个圆周上的充分必要条件.

11. 写出通过不在一条直线上的三点 $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ 的圆周的方程.

12. 证明：通过具有有理坐标的三点的圆周，其圆心也有有理坐标.

13. 求四点 $(x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), (x_4, y_4, z_4)$ 位于一个平面内的充分必要条件.

14. 求通过不位于一个平面内的四点 $(x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), (x_4, y_4, z_4)$ 的球面的方程.

15. 求通过点$(1,1,1,)$,$(1,1,-1)$,$(1,-1,1)$,$(-1,0,0)$的球面的方程,并且求其中心和半径.

16. 怎样的线性方程组,给出平面上共点的三条不同直线?

17. 怎样的线性方程组,给出平面上组成三角形的三条直线?

18. 怎样的线性方程组,给出空间中三个没有公共点但两两相交的平面?

19. 怎样的线性方程组,给出空间中组成四面体的四个平面?

20. 给以下事实以几何解释:在三个未知量四个方程的线性方程组中,增广矩阵的秩为3,并且其中任意三个方程的系数矩阵的秩为3.

补 充 题 三

1. 证明:有理数域上的齐次线性方程组如果它的系数矩阵的秩小于未知量的个数,那么可构造出分量全为整数的基础解系.

2. 设向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$;$\beta_1,\beta_2,\cdots,\beta_t$;$\alpha_1,\alpha_2,\cdots,\alpha_s,\beta_1,\beta_2,\cdots,\beta_t$的秩分别为$r_1,r_2,r_3$,证明:

$$\max\{r_1,r_2\}\leqslant r_3\leqslant r_1+r_2$$

3. 一个n级方阵如果至少有n^2-n+1个元素为0,证明它的秩小于n,求这种n级方阵的秩的最大值.

4. 设$A=(a_{ij})$为一实数域上的矩阵,证明:

1) 如果$|a_{ii}|>\sum\limits_{j\neq i}|a_{ij}|$, $i=1,2,\cdots,n$,则$|A|\neq 0$;

2) 如果$a_{ii}>\sum\limits_{j\neq i}|a_{ij}|$, $i=1,2,\cdots,n$,则$|A|>0$.

5. 证明:方程个数比未知量个数大1的线性方程组有解的必要条件是它的增广矩阵的行列式等于零.如果系数矩阵的秩等于未知量的个数,则这一条件也是充分条件.

6. 给定一个非齐次线性方程组 Ⅰ,任取 Ⅰ 的有限多个解:$\gamma_1,\gamma_2,\cdots,\gamma_t$.求$c_1\gamma_1+c_2\gamma_2+\cdots c_t\gamma_t$仍是 Ⅰ 的解的充分必要条件.

第四章 矩阵的运算

§1 引 言

在前面三章中我们已经看到矩阵在线性方程组的理论中起着重要的作用.线性方程组的求解,解的情况的判定以及解的结构需要用到矩阵的初等行变换,方阵的行列式以及矩阵的秩.在数学的各个分支中,在自然科学,工程技术以及生产实际中,有大量的各种各样的问题都提出矩阵的概念,需要运用矩阵的理论.下面举几个典型例子.

例 1 平面上二次曲线的一般方程为

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_1x + 2a_2y + a_0 = 0 \qquad (1)$$

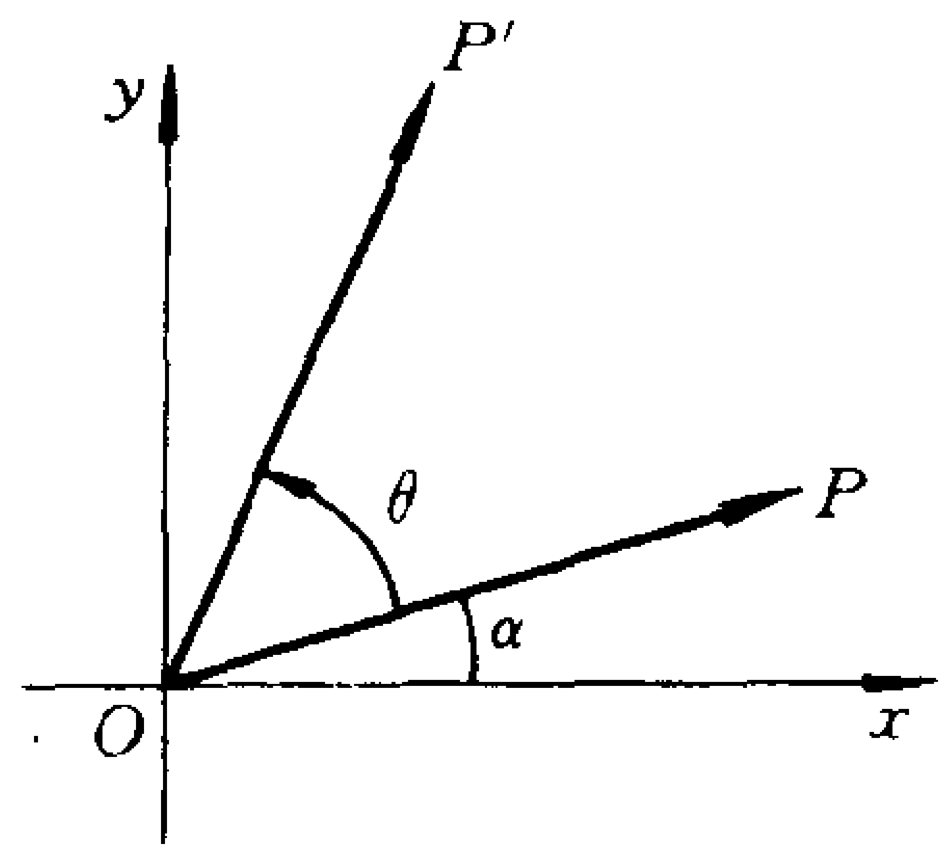
其中 a_{11}, a_{12}, a_{22} 不全为零.把方程(1)的系数和常数项排成如下的一张表:

$$\begin{bmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a_0 \end{bmatrix} \qquad (2)$$

称它为二次方程(1)的矩阵.这个矩阵在研究二次曲线方程的化简以及二次曲线的性质中起着重要作用.

例 2 平面上取定一个直角坐标系 Oxy,所有以原点为起点的向量组成的集合记作 V.让 V 中每个向量绕原点 O 旋转角度 θ,

这是集合 V 到自身的一个映射，记作 σ. 现在来求旋转 σ 的公式.



设$\overrightarrow{OP}$ 的坐标为(x,y)，它在 σ 下的象$\overrightarrow{OP'}$ 的坐标为(x',y'). 设以 x 轴的正半轴为始边，以射线OP 为终边的角为α. 设 $|\overrightarrow{OP}|=r$. 从三角函数的定义得到

$$x = r\cos\alpha \qquad y = r\sin\alpha$$
$$x' = r\cos(\alpha+\theta) \qquad y' = r\sin(\alpha+\theta)$$

由此得出

$$\begin{cases} x' = x\cos\theta - y\sin\theta \\ y' = x\sin\theta + y\cos\theta \end{cases} \tag{3}$$

(3) 式就是旋转 σ 的公式. 把公式中的系数排成如下一张表：

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{4}$$

这样旋转 σ 可以用矩阵(4) 来表示，其中的 θ 就是转角.

例 3　设某公司有四个商场销售电视机、电冰箱、洗衣机. 九月份的销售金额可以用一个矩阵表示：

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix} \tag{5}$$

其中 a_{i1}, a_{i2}, a_{i3} 分别表示商场 i 销售电视机、电冰箱、洗衣机的金额，$i = 1,2,3,4$.

＊**例** 4　设有七个水稻品种 P_i，$i = 1,\cdots,7$，要比较它们的优劣，应当如何安排试验方案？

如果将一大块田分成七小块分别种这七个品种的水稻，那么由于田块过大而难以保证七小块田之间的肥沃程度一样. 为此，一块田的面积不宜过大，以保证这块田里各处的肥沃程度一样. 这时，一块田里的水稻品种就要尽可能少些. 譬如，一块田里只种 3 个品种的水稻(即，把一块田均匀分成 3 小块，每小块种 1 个品种的水稻). 这时就需要好几块田. 由于田块与田块之间的肥沃程度不全一样，因此为了比较这七个品种的优劣，应当让每两个品种都能在同一块田里种. 我们把一块田叫做一个区组. 则上述安排试验方案的要求可叙述成：

1)每个区组种 3 个品种的水稻；

2)每两个品种相遇在恰好一个区组里.

如何求出满足这两个要求的试验方案？我们可用下述方法：

我们用 $B_1, B_2, B_3, \cdots$，表示各个区组. 现在我们构造一个矩阵 A，它的元素如下规定：

$$A(i;j) := \begin{cases} 1 & \text{当 } P_i \text{ 安排在 } B_j \text{ 里} \\ 0 & \text{当 } P_i \text{ 没有安排在 } B_j \text{ 里} \end{cases}$$

由于每个区组 B_j 都安排 3 个品种，并且每两个品种恰好相遇在一个区组里，因此可求出 A 如下：

$$\begin{array}{c} \\ P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \\ P_7 \end{array}\begin{array}{c} \begin{array}{ccccccc} B_1 & B_2 & B_3 & B_4 & B_5 & B_6 & B_7 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

A 只有 7 列，不可能有第 8 列. 因此安排这样的试验共需要 7 个区组. 从矩阵 A 还可看出，每个品种恰好安排在了 3 个区组里. 矩阵 A 就给出了一种试验方案.

从上述四个例子看到，不同领域中的问题都提出了矩阵的概念. 为了深入研究这些问题，需要使矩阵能进行运算. 这一章就来介绍矩阵的运算以及与运算有关的性质. 矩阵的乘法运算在矩阵理论及其应用中起着特别重要的作用. 为了理解矩阵乘法为什么那样规定，我们先介绍映射的乘法等概念. 本章中如果没有特别声明，所讲的矩阵都是任意给定的一个数域 K 上的矩阵.

§2 映　　射

定义 1　设 S 和 S' 是两个集合. 集合 S 到集合 S' 的一个**映射** f 是指一个法则，它使 S 中每一个元素 a 都有 S' 中唯一的元素 a' 与之对应. 此时，a' 称为 a 在映射 f 下的**象**，记作 $f(a)$ 或写成 fa；a 称为 a' 在 f 下的**一个原象**. 映射 f 记作

$$f: S \to S'$$
$$a \mapsto a'$$

或者记作 $f(a)=a'$. 集合 S 称为映射 f 的**定义域**(domain)，集合

S' 称为 f 的**伴域**(codomain).

定义 2　设 f 是集合 S 到集合 S' 的一个映射,S 的所有元素在 f 下的象组成的集合称为 f 的**象集**,或简称为 f 的**象**,记作 $\mathrm{Im}f$,或者记作 $f(S)$,即

$$\mathrm{Im}f := \{f(x) \mid x \in S\} =: f(S)$$

显然 $f(S) \subset S'$,也就是 $\mathrm{Im}f \subset S'$.

定义 3　如果映射 $f: S \to S'$ 使得 $\mathrm{Im}f = S'$,则称 f 是一个**满射**或一个**映上**的映射(或者称 f 是 S 到 S' 上的映射).

显然,f 为满射的充分必要条件是它的伴域中的每一个元素都有一个原象.

定义 4　如果映射 $f: S \to S'$ 使得不同元素的象也一定不同,即,如果对任意 $a_1, a_2 \in S$,且 $a_1 \neq a_2$,都有 $f(a_1) \neq f(a_2)$,那么称 f 是一个**单射**或 1—1 **的映射**.

显然,$f: S \to S'$ 为单射的充分必要条件是:从 $a_1, a_2 \in S$,且 $f(a_1) = f(a_2)$ 可以推出 $a_1 = a_2$.

定义 5　一个映射 $f: S \to S'$ 如果既是满射又是单射,则称它是**双射**或**一一对应**.

定义 6　设 f 是集合 S 到集合 S' 的一个映射,对于元素 $b \in S'$,b 在 f 下的所有原象组成的集合称为 b 在 f 下的**原象集**或**原象**,记作 $f^{-1}(b)$,即

$$f^{-1}(b) := \{x \in S \mid f(x) = b\}$$

如果 $b \in S' \backslash \mathrm{Im}f$,则 $f^{-1}(b) = \varnothing$.

如果 $b \in \mathrm{Im}f$,则 $f^{-1}(b)$ 称为 b 上的**纤维**.

定义 7　映射 f 和映射 g 称为**相等**,如果它们的定义域相同,伴域也相同,并且 $\forall\, x \in S$ 有 $f(x) = g(x)$.

例如,设 R_+ 表示非负实数的集合,那么由同一规则 $x \to x^2$ 定义的映射 $f: R \to R$,$g: R \to R_+$,$h: R_+ \to R_+$ 是互不相同的. 这里 f 既不是满射也不是单射;g 是满射但不是单射;h 是双射. 这个例子表明,在定义一个映射时一定要说清楚定义域和伴域,当然也要说

清楚对应法则.

定义 8　映射 $f:S\to S$ 如果使每一个元素对应到它自身，即 $f(x)=x, \forall\, x\in S$，则称 f 为**恒等映射**或**单位映射**，记作 1_S.

集合 S 到自身的一个映射通常称为 S 的一个**变换**.

集合 S 到数域 K 的一个映射通常称为 S 上的一个**函数**. 因此，函数可以认为是映射的特殊情形.

在许多问题中，我们需要相继作两次映射，为此引进下述概念：

定义 9　设 $g:S\to S'$，$f:S'\to S''$. 相继施行映射 g 和 f，得到一个 S 到 S'' 的映射，称为 f 与 g 的**乘积**(或**合成**)，记作 fg，即

$$(fg)(a):=f(g(a))\qquad \forall\, a\in S \tag{1}$$

用三角形图可以把这个定义直观地描绘成

$$\begin{tikzcd} S \arrow[rr, "fg"] \arrow[dr, "g"'] & & S'' \\ & S' \arrow[ur, "f"'] & \end{tikzcd}$$

映射的定义表明这个图“可交换”，即由 S 到 S'' 的结果，不依赖于我们是直接用 fg 得到的，还是用 g 经过 S' 再用 f 得到的.

注意：只有 g 的象集含于 f 的定义域才能作 f 与 g 的合成 fg. 当 g 的伴域与 f 的定义域一致时便能保证这点. 特别地，从集合 S 到自身的两个映射的合成总是有意义的.

在数学分析中，函数 $u=g(x)$ 与函数 $y=f(u)$ 的复合函数 $y=f(g(x))$ 就是映射 f 与映射 g 的合成 fg.

容易直接验证，对于任意一个映射 $f:S\to S'$，有

$$f1_S=f,\qquad 1_{S'}f=f \tag{2}$$

下面的定理给出了映射合成(乘法)的一个重要性质:

定理 4.2.1 映射的乘法(合成)适合结合律,即,如果 $h:S\to S'$, $g:S'\to S''$, $f:S''\to S'''$ 是三个映射,则

$$f(gh)=(fg)h \tag{3}$$

证明 $f(gh)$ 与 $(fg)h$ 都是 S 到 S''' 的映射. $\forall\, a\in S$,有

$$(f(gh))a=f((gh)a)=f(g(ha))$$

$$((fg)h)a=(fg)(ha)=f(g(ha))$$

因此, $(f(gh))a=((fg)h)a$, $\forall\, a\in S$. 所以 $f(gh)=(fg)h$. ▎

注意:映射的乘法不适合交换律. 可能 fg 有意义,但是 gf 没有意义. 即使 fg 与 gf 都有意义,也可能 $fg\neq gf$. 例如,设 f 与 g 都是集合 S 到 S 的**常值映射**,即 $f(x)$ 与 $g(x)$ 的值不依赖于 x. 设 $f(x)=b$, $\forall\, x\in S$;设 $g(x)=c$, $\forall\, x\in S$;并且设 $b\neq c$,则

$$(fg)x=f(gx)=f(c)=b$$

$$(gf)x=g(fx)=g(b)=c$$

因此 $fg\neq gf$.

与映射的乘法密切相关的一个概念是可逆映射的概念.

定义 10 设 $f:S\to S'$,如果存在一个映射 $g:S'\to S$,使得

$$fg=1_{S'},\qquad gf=1_S \tag{4}$$

则称映射 f 是**可逆的**,此时称 g 是 f 的一个**逆映射**.

如果 $f:S\to S'$ 是可逆的,则它的逆映射是唯一的. 证明如下:假如 g 和 g' 都是 f 的逆映射,则有

$$g'fg=(g'f)g=1_Sg=g$$

$$g'fg=g'(fg)=g'1_{S'}=g'$$

因此, $g=g'$. 当 f 是可逆映射时,我们用 f^{-1} 表示 f 的逆映射,于是有

$$ff^{-1}=1_{S'},\qquad f^{-1}f=1_S \tag{5}$$

(5) 式说明,当 f 是可逆映射时,它的逆映射 f^{-1} 也可逆,并且 f^{-1} 的逆映射就是 f,即有

$$(f^{-1})^{-1}=f \tag{6}$$

什么样的映射是可逆的？下面的定理回答了此问题.

定理 4.2.2　映射 $f:S\to S'$ 是可逆的充分必要条件为 f 是双射.

为了证明定理 4.2.2，先证下面的引理 4.2.1，引理 4.2.1 本身也是很有用的.

引理 4.2.1　如果两个映射 $f:S\to S'$ 与 $g:S'\to S$ 满足 $gf=1_s$，则 f 是一个单射，g 是一个满射.

证明　先证 f 是单射. 设 $a_1,a_2\in S$，并且 $f(a_1)=f(a_2)$，则

$$\begin{aligned} a_1 &= 1_s(a_1)=(gf)a_1 \\ &=g(fa_1)=g(fa_2)=(gf)a_2=1_s(a_2)=a_2 \end{aligned}$$

因此 f 是单射.

再证 g 是满射. 任取 $a\in S$，因为 $g(fa)=(gf)a=1_s(a)=a$，所以 $f(a)$ 是 a 在 g 下的一个原象. 这证明了 g 是满射. ▮

定理 4.2.2 的证明　先证必要性. 设 $f:S\to S'$ 可逆，则有

$$ff^{-1}=1_{s'},\qquad f^{-1}f=1_s$$

由引理 4.2.1 即得，f 是满射，并且 f 是单射. 因此 f 是双射.

再证充分性. 设 $f:S\to S'$ 是双射. 对于任给的 $a'\in S'$，因为 f 是满射，所以 a' 在 f 下有一个原象 a，即 $f(a)=a'$. 因为 f 是单射，所以 a' 的原象集只有一个元素 a. 令 $g(a')=a$，则 g 是 S' 到 S 的一个映射，并且有

$$(fg)a'=f(ga')=f(a)=a'$$

对任意 $x\in S$，记 $f(x)=x'$. 据上述议论得 $g(x')=x$，从而有

$$(gf)x=g(fx)=g(x')=x$$

因此，$fg=1_{s'}$，$gf=1_s$. 从而 f 是可逆的，且 $f^{-1}=g$. ▮

习　题　4.2

1.　判别下列对应法则是否为 R 到自身的映射？是否单射？是否满射？

(1) $x \mapsto x^3$ (2) $x \mapsto x^2 - x$

(3) $x \mapsto 2^x$ (4) $x \mapsto \ln x$

2. 设 $f:S \to S'$, $g:S' \to S''$. 证明:若 f 和 g 都是单(满)射,则 gf 也是单(满)射.

3. 设 $f:S \to S'$, $g:S' \to S''$. 证明:若 f 和 g 都是可逆的,则合成映射 gf 也是可逆的,并且有

$$(gf)^{-1} = f^{-1}g^{-1}$$

4. 证明:如果 S 是一个有限集合,并且映射 $f:S \to S$ 是单射,则 f 一定是双射.

5. 证明:一个有限集合到它自身的满射一定是双射.

6. 设 S 和 S' 是两个集合,如果存在一个双射 $f:S \to S'$,则称 S 和 S' 有相同的**基数**. 试证:整数集 Z 与偶数集有相同的基数. (注:与 Z 有相同基数的集合称为**可数集合**.)

7. 设 $f:S \to S'$. 设 $b, c \in \mathrm{Im} f$. 证明

(1) 若 $b \neq c$,则 b 上的纤维与 c 上的纤维不相交.

(2) S 是互不相交的纤维的并集.

§3 矩阵的运算

这一节介绍矩阵的运算及其满足的运算法则.

我们把数域 K 上所有 $s \times n$ 矩阵组成的集合记作 $M_{s\times n}(K)$.

定义 1 两个矩阵 $A_{s\times n}$, $B_{r\times m}$ 称为是**相等的**,如果 $s = r$, $n = m$,且 $A(i;j) = B(i;j)$, $i = 1, \cdots, s$; $j = 1, \cdots, n$.

3.1 矩阵的加法

在上一节的例 3 中,某公司的四个商场九月份销售电视机,电冰箱,洗衣机的金额用一个矩阵 $A = (a_{ij})$ 表示. 同样地,这四个商场十月份销售电视机,电冰箱,洗衣机的金额可以用一个矩阵 $B = (b_{ij})$ 表示. 于是九月份和十月份两个月销售这三种商品的总金额

可用下述矩阵 C 表示：

$$C=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \\ a_{31}+b_{31} & a_{32}+b_{32} & a_{33}+b_{33} \\ a_{41}+b_{41} & a_{42}+b_{42} & a_{43}+b_{43} \end{pmatrix}$$

从问题的实际意义很自然地应当把矩阵 C 称为矩阵 A 与矩阵 B 的和. 由此受到启发，规定矩阵的加法如下：

定义 2 设 $A=(a_{ij}), B=(b_{ij})\in M_{s\times n}(K)$. 令

$$C=(a_{ij}+b_{ij})_{s\times n} \tag{1}$$

把矩阵 C 称为矩阵 A 与 B 的**和**，记为 $C=A+B$.

从定义 2 看出，矩阵的加法就是把对应的元素相加，即

$$(A+B)(i;j)=A(i;j)+B(i;j), 1\leqslant i\leqslant s; 1\leqslant j\leqslant n \tag{2}$$

注意：只有行数相同，列数也相同的两个矩阵才能相加.

容易直接验证，矩阵的加法满足下述四条运算法则：$\forall A,B,C\in M_{s\times n}(K)$，有

1° 交换律：$A+B=B+A$；

2° 结合律：$A+(B+C)=(A+B)+C$；

3° $A+0=0+A=A$；

4° 设 $A=(a_{ij})$，矩阵

$$\begin{pmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{s1} & -a_{s2} & \cdots & -a_{sn} \end{pmatrix}$$

称为 A 的**负矩阵**，记作 $-A$. 显然有

$$A+(-A)=(-A)+A=0$$

矩阵的**减法**定义为

$$A-B:=A+(-B) \tag{3}$$

3.2 矩阵的数量乘法

如果某公司十月份每个商场销售每种家电产品的金额都是九月份销售金额的 k 倍(譬如,$k=1.05$),那么该公司的四个商场十月份的销售金额可用下述矩阵 D 表示:

$$D=\begin{pmatrix} ka_{11} & ka_{12} & ka_{13} \\ ka_{21} & ka_{22} & ka_{23} \\ ka_{31} & ka_{32} & ka_{33} \\ ka_{41} & ka_{42} & ka_{43} \end{pmatrix}$$

从问题的实际意义很自然地应当把矩阵 D 称为数 k 与矩阵 A 的数量乘积. 由此受到启发,规定矩阵的数量乘法如下:

定义 3　设 $A=(a_{ij})_{s\times n}$,$k\in K$,令

$$D=(ka_{ij})_{s\times n} \tag{4}$$

把矩阵 D 称为数 k 与矩阵 A 的**数量乘积**,记为 $D=kA$.

从定义 3 看出,用数 k 乘矩阵 A 就是把 A 的每一个元素都乘以 k,即

$$(kA)(i;j)=k[A(i;j)],1\leqslant i\leqslant s;1\leqslant j\leqslant n \tag{5}$$

容易直接验证,矩阵的数量乘法满足下述四条运算法则:$\forall\ A,B,C\in M_{s\times n}(K)$,$k,l\in K$,有

5°　$1A=A$;

6°　$k(lA)=(kl)A$;

7°　$(k+l)A=kA+lA$;

8°　$k(A+B)=kA+kB$.

3.3 矩阵的乘法

在上一节例 2 中,我们指出:平面上绕原点 O 的转角为 θ 的旋转 σ 可以用一个矩阵

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

来表示. 同理,绕原点 O 的转角为 γ 的旋转 τ 可以用矩阵

$$B = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$$

来表示.

现在考虑映射 σ 与 τ 的乘积 $\sigma\tau$,从映射的乘法定义知道,这是先作旋转 τ,接着作旋转 σ. 显然,其总的效果是作了一个转角为 $\theta+\gamma$ 的旋转 ψ,因此

$$\psi = \sigma\tau \tag{6}$$

同上理,旋转 ψ 可以用矩阵

$$C = \begin{pmatrix} \cos(\theta+\gamma) & -\sin(\theta+\gamma) \\ \sin(\theta+\gamma) & \cos(\theta+\gamma) \end{pmatrix}$$

来表示. 由于 $\psi = \sigma\tau$,因此很自然地可以把矩阵 C 称为矩阵 A 与 B 的乘积,即

$$C =: AB \tag{7}$$

现在我们来仔细看一下矩阵 C 的元素与矩阵 A,B 的元素之间有什么关系. 利用和角公式得

$$C = \begin{pmatrix} \cos\theta\cos\gamma - \sin\theta\sin\gamma & -\sin\theta\cos\gamma - \cos\theta\sin\gamma \\ \sin\theta\cos\gamma + \cos\theta\sin\gamma & \cos\theta\cos\gamma - \sin\theta\sin\gamma \end{pmatrix} \tag{8}$$

从(8)式看出,C 的(1,1)元是矩阵 A 的第 1 行与 B 的第 1 列的对应元素乘积之和;C 的(1,2)元是 A 的第 1 行与 B 的第 2 列的对应元素乘积之和;C 的(2,1)元是 A 的第 2 行与 B 的第 1 列的对应元素乘积之和;C 的(2,2)元是 A 的第 2 行与 B 的第 2 列的对应元素乘积之和.

由上述受到启发,规定矩阵的乘法如下:

定义 4 设 $A = (a_{ij})_{s\times n}, B = (b_{ij})_{n\times m}$,令

$$C = (c_{ij})_{s\times m}$$

其中

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \quad 1 \leqslant i \leqslant s, 1 \leqslant j \leqslant m \tag{9}$$

则矩阵 C 称为矩阵 A 与 B 的**乘积**，记作 $C = AB$.

矩阵的乘法有以下几个要点：

1） 只有左矩阵的列数与右矩阵的行数相同的两个矩阵才能相乘；

2） 乘积矩阵的 (i,j) 元等于左矩阵的第 i 行与右矩阵的第 j 列的对应元素的乘积之和，即

$$(AB)(i;j) = \sum_{k=1}^{n} [A(i;k)][B(k;j)] \tag{10}$$

3） 乘积矩阵的行数等于左矩阵的行数，乘积矩阵的列数等于右矩阵的列数.

例 1 设

$$A = \begin{pmatrix} 2 & -1 \\ -4 & 0 \\ 3 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 7 & -9 \\ -8 & 10 \end{pmatrix}$$

求 AB.

解

$$AB = \begin{pmatrix} 2 & -1 \\ -4 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 7 & -9 \\ -8 & 10 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \times 7 + (-1) \times (-8) & 2 \times (-9) + (-1) \times 10 \\ (-4) \times 7 + 0 \times (-8) & (-4) \times (-9) + 0 \times 10 \\ 3 \times 7 + 1 \times (-8) & 3 \times (-9) + 1 \times 10 \end{pmatrix}$$

$$= \begin{pmatrix} 22 & -28 \\ -28 & 36 \\ 13 & -17 \end{pmatrix}$$

现在来讨论矩阵乘法满足哪些运算法则.

1° 矩阵的乘法适合结合律：$\forall\, A \in M_{s \times n}(K), B \in$

$M_{n\times m}(K)$,$C \in M_{m\times r}(K)$,有

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

证明 显然,$(AB)C$ 与 $A(BC)$ 都是 $s \times r$ 矩阵. 为了证明它们相等,只要证它们的(i,j) 元对应相等. 设 $A = (a_{ij})$,$B = (b_{ij})$,$C = (c_{ij})$ 我们有

$$\begin{aligned}
[(AB)C](i;j) &= \sum_{l=1}^{m}[(AB)(i;l)][C(l;j)] \\
&= \sum_{l=1}^{m}\Big[\sum_{k=1}^{n}A(i;k)B(k;l)\Big]c_{lj} \\
&= \sum_{l=1}^{m}\sum_{k=1}^{n}a_{ik}b_{kl}c_{lj} \\
[A(BC)](i;j) &= \sum_{k=1}^{n}[A(i;k)][(BC)(k;j)] \\
&= \sum_{k=1}^{n}a_{ik}\Big[\sum_{l=1}^{m}b_{kl}c_{lj}\Big] = \sum_{k=1}^{n}\sum_{l=1}^{m}a_{ik}b_{kl}c_{lj} \\
&= \sum_{l=1}^{m}\sum_{k=1}^{n}a_{ik}b_{kl}c_{lj}
\end{aligned}$$

因此,$[(AB)C](i;j) = [A(BC)](i;j)$, $i = 1,\cdots,s$; $j = 1,\cdots,r$. 从而$(AB)C = A(BC)$. ▌

注意:矩阵的乘法不适合交换律. 我们举两个例子:

例 2 K^n 中的一个向量写成行向量,可以看成是一个 $1 \times n$ 矩阵,K^n 中的向量写成列向量可以看成是 $n \times 1$ 矩阵. 设

$$A = (a_1, a_2, \cdots, a_n)$$

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

求 AB,BA.

解

$$AB=(a_1,a_2,\cdots,a_n)\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

$$=(a_1b_1+a_2b_2+\cdots+a_nb_n)$$

如果运算的最后结果得到一个 1 级矩阵，那么我们可以把它写成一个数. 因此

$$AB=a_1b_1+a_2b_2+\cdots+a_nb_n$$

$$BA=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}(a_1,a_2,\cdots,a_n)$$

$$=\begin{pmatrix}b_1a_1 & b_1a_2 & \cdots & b_1a_n\\ b_2a_1 & b_2a_2 & \cdots & b_2a_n\\ \cdots & \cdots & \cdots & \cdots\\ b_na_1 & b_na_2 & \cdots & b_na_n\end{pmatrix}$$

从例 2 看出，$AB\neq BA$. 即使同级方阵的乘法一般也不适合交换律，看下面的例子.

例 3　设

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

求 AB,BA.

解

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

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

从例 3 看出，$AB\neq BA$. 例 3 还说明，$A\neq 0,B\neq 0$，但是 $AB=0$. 这与数的乘法是很不同的. 对于一个矩阵 A，如果存在一个矩

阵 $B\neq 0$ 使得 $AB=0$，则称 A 是一个**左零因子**；如果存在一个矩阵 $C\neq 0$ 使得 $CA=0$，则称 A 是一个**右零因子**. 显然，零矩阵既是左零因子，又是右零因子，称它是**平凡的零因子**. 上面的例子说明了存在非平凡的左零因子和右零因子：例 3 中的 A 是左零因子，B 是右零因子.

2° 矩阵的乘法适合左分配律

$$A(B+C)=AB+AC \tag{11}$$

也适合右分配律

$$(B+C)A=BA+CA \tag{12}$$

证明 设 $A=(a_{ij})\in M_{s\times n}(K)$；$B=(b_{ij})$，$C=(c_{ij})\in M_{n\times m}(K)$，则 $A(B+C)$，$AB+AC\in M_{s\times m}(K)$，并且有

$$\begin{aligned}
&[A(B+C)](i;j)\\
&=\sum_{k=1}^{n}A(i;k)[(B+C)(k;j)]\\
&=\sum_{k=1}^{n}a_{ik}(b_{kj}+c_{kj})=\sum_{k=1}^{n}a_{ik}b_{kj}+\sum_{k=1}^{n}a_{ik}c_{kj}\\
&=(AB)(i;j)+(AC)(i;j)\\
&=(AB+AC)(i;j),\quad 1\leqslant i\leqslant s,\quad 1\leqslant j\leqslant m
\end{aligned}$$

所以 $A(B+C)=AB+AC$. ▌

右分配律留给读者作为练习.

设 $C\neq 0$，如果 $AC=BC$，则 $(A-B)C=0$. 由于矩阵的乘法可能出现零因子，因此从 $(A-B)C=0$ 不能推出 $A-B=0$. 这说明从 $AC=BC$，且 $C\neq 0$，推不出 $A=B$. 即，矩阵的乘法不适合消去律. 例如，不难直接验证有

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

但是

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

3° 主对角线上元素都是 1，其余元素都是 0 的 n 级矩阵

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

称为 n 级**单位矩阵**，记作 I_n，或 I. 有

$$I_n A_{n\times m} = A_{n\times m}, A_{n\times m} I_m = A_{n\times m}$$

特别地，若 A 是 n 级方阵，则

$$IA = AI = A \tag{13}$$

证明 直接用矩阵的乘法进行验证. ▌

4° 矩阵的乘法与数量乘法满足下述关系式

$$k(AB) = (kA)B = A(kB) \tag{14}$$

证明 设 $A=(a_{ij})\in M_{s\times n}(K)$，$B=(b_{ij})\in M_{n\times m}(K)$，$k\in K$，则 $k(AB)$，$(kA)B$，$A(kB)\in M_{s\times m}(K)$，并且有

$$[k(AB)](i;j) = k(AB)(i;j) = k\sum_{l=1}^{n} a_{il}b_{lj}$$

$$[(kA)B](i;j) = \sum_{l=1}^{n}(kA)(i;l)B(l;j)$$

$$= \sum_{l=1}^{n} k a_{il}b_{lj} = k\sum_{l=1}^{n} a_{il}b_{lj}$$

$$[A(kB)](i;j) = \sum_{l=1}^{n} A(i;l)(kB)(l;j)$$

$$= \sum_{l=1}^{n} a_{il}kb_{lj} = k\sum_{l=1}^{n} a_{il}b_{lj}$$

$$1\leqslant i\leqslant s;\quad 1\leqslant j\leqslant m$$

因此(14)式成立. ▌

方阵

$$
\begin{pmatrix}
k & 0 & 0 & \cdots & 0 \\
0 & k & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & \cdots & k
\end{pmatrix} \tag{15}
$$

称为**数量矩阵**. 根据矩阵的数量乘法,数量矩阵(15)可以记成kI. 显然有

$$kI + lI = (k + l)I \tag{16}$$

$$(kI)(lI) = (kl)I \tag{17}$$

$$k(lI) = (kl)I \tag{18}$$

上述三个式子表明,n级数量矩阵组成的集合对于矩阵的加法、乘法与数量乘法三种运算都封闭.

据规则 4° 和 3°,我们可得到

$$(kI_n)A_{n\times m} = k(I_nA_{n\times m}) = kA_{n\times m} \tag{19}$$

$$A_{n\times m}(kI_m) = k(A_{n\times m}I_m) = kA_{n\times m} \tag{20}$$

(19)和(20)式表明,数量矩阵与矩阵的乘法相当于数与矩阵的数量乘法.

前面已指出,矩阵的乘法是非交换的. 但是对于具体的两个矩阵A与B,也有可能$AB = BA$. 如果A与B满足$AB = BA$,则称A与B**可交换**. 容易看出,两个矩阵可交换的一个必要条件是它们为同级方阵.

数量矩阵有一个重要性质,即

命题 4.3.1 任一数量矩阵与所有和它同级的方阵可交换.

证明 设kI为n级,任取一个n级矩阵A,据(19)和(20)式得

$$(kI)A = kA = A(kI)$$ ▌

设A是一个n级方阵. 因为矩阵的乘法适合结合律,所以$\underbrace{A\cdot A\cdot\cdots\cdot A}_{m个}$表示唯一的一个矩阵,于是可以定义

$$A^m := \underbrace{AA\cdots A}_{m个}, \quad 其中\ m\ 是正整数 \tag{21}$$

A^m 称为方阵 A 的 m 次**幂**. 换句话说，

$$A^1 := A, \quad A^{m+1} := A^m A, \quad 其中 m 是正整数 \tag{22}$$

我们还规定

$$A^0 := I \tag{23}$$

容易证明方阵的方幂适合下列规则：

$$A^k A^l = A^{k+l} \tag{24}$$

$$(A^k)^l = A^{kl} \tag{25}$$

其中 k,l 是任意非负整数.

注意：由于矩阵乘法不适合交换律，因此一般来说，$(AB)^k \neq A^k B^k$. 但是如果 A 与 B 可交换，则有 $(AB)^k = A^k B^k$.

例 4　证明：若矩阵 A 与 B 可交换，则有

$$(A+B)^2 = A^2 + 2AB + B^2$$

证明

$$\begin{aligned}(A+B)^2 &= (A+B)(A+B) = A^2 + AB + BA + B^2 \\ &= A^2 + AB + AB + B^2 = A^2 + 2AB + B^2\end{aligned} \qquad \blacksquare$$

类似地可以证明：若矩阵 A 与 B 可交换，则二项式定理成立，即有

$$(A+B)^m = A^m + C_m^1 A^{m-1} B + \cdots + C_m^i A^{m-i} B^i + \cdots + B^m$$

例 5　计算

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}^n, \quad n 是正整数$$

解　设

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

则

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = I + B, 其中 B = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$$

显然

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

因为 I 与 B 可交换，所以由二项式定理得

$$A^n=(I+B)^n=I^n+C_n^1I^{n-1}B=I+nIB=I+nB$$
$$=\begin{pmatrix}1&0\\0&1\end{pmatrix}+\begin{pmatrix}0&2n\\0&0\end{pmatrix}=\begin{pmatrix}1&2n\\0&1\end{pmatrix}$$

例 5 表明，应当善于利用运算法则去做矩阵运算.

3.4 矩阵的转置与矩阵运算的关系

这一小节我们来看矩阵的转置与矩阵运算的关系.

1° $(A')'=A$

2° $(A+B)'=A'+B'$

3° $(kA)'=kA'$

4° $(AB)'=B'A'$

1°—3°都容易验证.

4°的证明 设 $A=(a_{ij})_{s\times n}$，$B=(b_{ij})_{n\times m}$. 则 $A'\in M_{n\times s}(K)$，$B'\in M_{m\times n}(K)$，$AB\in M_{s\times m}(K)$，$(AB)'\in M_{m\times s}(K)$，$B'A'\in M_{m\times s}(K)$. 并且有

$$(AB)'(i;j)=(AB)(j;i)=\sum_{k=1}^{n}a_{jk}b_{ki}$$

$$(B'A')(i;j)=\sum_{k=1}^{n}B'(i;k)A'(k;j)=\sum_{k=1}^{n}B(k;i)A(j;k)$$
$$=\sum_{k=1}^{n}b_{ki}a_{jk}\qquad 1\leqslant i\leqslant m,\quad 1\leqslant j\leqslant s$$

因此 $(AB)'=B'A'$. ▌

3.5 矩阵乘法的另外两种表述方式

设 $A=(a_{ij})_{s\times n}$，$B=(b_{ij})_{n\times m}$，把 A 的列向量组记作 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 由矩阵乘法的定义得，$AB$ 的第 1 个列向量为

$$\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{s1}b_{11} + a_{s2}b_{21} + \cdots + a_{sn}b_{n1} \end{pmatrix}$$

$$= b_{11}\alpha_1 + b_{21}\alpha_2 + \cdots + b_{n1}\alpha_n$$

类似地可以得到，AB 的第 j 个列向量为

$$b_{1j}\alpha_1 + b_{2j}\alpha_2 + \cdots + b_{nj}\alpha_n \qquad j = 1, 2, \cdots, m$$

因此，AB 的列向量组为

$$\sum_{k=1}^{n} b_{k1}\alpha_k, \quad \sum_{k=1}^{n} b_{k2}\alpha_k, \quad \cdots, \quad \sum_{k=1}^{n} b_{km}\alpha_k$$

我们约定：若矩阵 A 的列向量组为 $\alpha_1, \alpha_2, \cdots, \alpha_n$，则可把 A 记成

$$A = (\alpha_1, \alpha_2, \cdots, \alpha_n)$$

同理，AB 可以记成

$$AB = \left(\sum_{k=1}^{n} b_{k1}\alpha_k, \sum_{k=1}^{n} b_{k2}\alpha_k, \cdots, \sum_{k=1}^{n} b_{km}\alpha_k\right)$$

于是有

$$AB = (\alpha_1, \alpha_2, \cdots, \alpha_n) \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix}$$

$$= \left(\sum_{k=1}^{n} b_{k1}\alpha_k, \sum_{k=1}^{n} b_{k2}\alpha_k, \cdots, \sum_{k=1}^{n} b_{km}\alpha_k\right) \tag{26}$$

(26)式说明，做矩阵乘法时，可以把左矩阵的列向量组分别与右矩阵的第 1 列，第 2 列，$\cdots$，第 m 列的对应元素乘积之和作为乘积矩阵的相应各列. 这是矩阵乘法的第二种表述方式.

按照矩阵乘法的第二种表述方式，我们可以把一个 $s \times n$ 矩阵 A 的列向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的一个线性组合 $x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$ 写成矩阵乘积的形式：

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$$

$$= (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \doteq AX$$

其中，$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. 于是线性方程组

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta \tag{27}$$

可以写成
$$AX = \beta \tag{28}$$

其中 X 是由未知量 $x_1, x_2, \cdots, x_n$ 组成的列向量，A 是方程组(27) 的系数矩阵.

含有未知矩阵的等式称为**矩阵方程**. 譬如，(28) 就是一个矩阵方程. 把线性方程组(27) 写成矩阵方程(28) 的形式不仅使线性方 程组获得了更加简洁的形式，而且可以把线性方程组的理论与矩阵的理论联系起来，这在以后会逐渐看到.

按照上面的议论，一个齐次线性方程组

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = 0 \tag{29}$$

可以写成下述矩阵方程

$$AX = 0 \tag{30}$$

其中 A 是方程组(29) 的系数矩阵. 于是一个列向量 $\eta \in K^n$ 是齐次线性方程组(30) 的解的充分必要条件是 $A\eta = 0$.

现在我们来看矩阵乘法的第三种表述方式. 设一个 $s \times n$ 矩阵的行向量组是 $\gamma_1, \gamma_2, \cdots, \gamma_s$. 我们可以把 A 记成

$$A = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_s \end{pmatrix} \tag{31}$$

设 $C=(c_{ij})$ 是一个 $m\times s$ 矩阵. 我们想利用 A 的行向量组的表出式(31) 来看 C 与 A 的乘积 CA 的表述式. 因为

$$A'C'=(\gamma_1',\gamma_2',\cdots,\gamma_s')\begin{pmatrix}c_{11}&c_{21}&\cdots&c_{m1}\\c_{12}&c_{22}&\cdots&c_{m2}\\\cdots&\cdots&\cdots&\cdots\\c_{1s}&c_{2s}&\cdots&c_{ms}\end{pmatrix}$$

$$=\left(\sum_{k=1}^{s}c_{1k}\gamma_k',\sum_{k=1}^{s}c_{2k}\gamma_k',\cdots,\sum_{k=1}^{s}c_{mk}\gamma_k'\right)$$

所以

$$CA=(A'C')'=\begin{pmatrix}\sum_{k=1}^{s}c_{1k}\gamma_k\\\sum_{k=1}^{s}c_{2k}\gamma_k\\\vdots\\\sum_{k=1}^{s}c_{mk}\gamma_k\end{pmatrix}$$

即

$$\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1s}\\c_{21}&c_{22}&\cdots&c_{2s}\\\cdots&\cdots&\cdots&\cdots\\c_{m1}&c_{m2}&\cdots&c_{ms}\end{pmatrix}\begin{pmatrix}\gamma_1\\\gamma_2\\\vdots\\\gamma_s\end{pmatrix}=\begin{pmatrix}\sum_{k=1}^{s}c_{1k}\gamma_k\\\sum_{k=1}^{s}c_{2k}\gamma_k\\\vdots\\\sum_{k=1}^{s}c_{mk}\gamma_k\end{pmatrix}\tag{32}$$

(32)式表明，如果把右矩阵 A 写成行向量组形式，则做矩阵乘法时，把左矩阵的每一行元素与右矩阵的行向量组的对应行向量的乘积之和作为乘积矩阵的相应的行向量. 这是矩阵乘法的第三种表述方式.

矩阵乘法的三种表达方式都有用，希望读者能逐渐掌握它们.

在下面两节，读者将会发现矩阵的第二、三种表达方式可以更清晰、更简洁地表述所讨论的问题.

习　题　4.3

1. 设

$$A=\begin{pmatrix}3&-2&7&5\\1&0&4&-3\\6&8&0&2\end{pmatrix},\qquad B=\begin{pmatrix}-2&0&1&4\\5&-1&7&6\\4&-2&1&-9\end{pmatrix}$$

求 $A+B, A-B, 3A-2B$.

2. 计算

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

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

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

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

3. 计算

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

4. 计算

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

5. 计算

(1) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2$ (2) $\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}^2$ (3) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$ n 是正整数

(4) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}^4$

(5) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^n$ n 是正整数 (6) $\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^n$ n 是正整数

(7) $\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}_{s\times s}^n$ n 是正整数

(8) $\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}_{s\times s}^n$ n 是正整数， $n \geqslant s$

(9) $\begin{pmatrix} a & c \\ 0 & b \end{pmatrix}^n$ n 是正整数

(10) $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}^n$ n 是正整数.

6. 证明：若 B_1, B_2 都与 A 可交换，则 B_1+B_2, B_1B_2 也都与 A 可交换.

7. 设 $A=\frac{1}{2}(B+I)$. 证明：$A^2=A$ 当且仅当 $B^2=I$.

8. 设

$$A=\begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 2 \\ 1 & 1 & 1 \end{bmatrix} \qquad B=\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 4 & 2 \end{bmatrix}$$

求 $AB-BA$.

9. 求与 A 可交换的所有矩阵，设

(1) $A=\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ (2) $A=\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$

10. 设 A 是 n 级方阵，如果对于 K^n 中的任一列向量 X，都有 $AX=0$，那么 $A=0$.

11. 设 A 是一个 n 级方阵，并且 $\operatorname{rank}(A)=1$. 证明：

(1) A 能表示成一个列向量与一个行向量的乘积；

(2) $A^2=kA$，其中 k 是某个数.

*12. 设 A,B 的元素都是非负实数. 如果 AB 中有一行的元素全为 0，那么 A 或者 B 中有一行元素全为 0.

*13. 求其平方等于零矩阵的所有 2 级矩阵.

*14. 求其平方等于单位矩阵的所有 2 级矩阵.

§4 几类常用的特殊矩阵

这一节介绍几类常用的特殊矩阵，着重讨论它们的乘法运算的特点.

（一） 基本矩阵

除了零矩阵以外，最简单的矩阵是只有一个元素是 1，其余所有元素都是 0 的矩阵，这种类型的矩阵称为**基本矩阵**. (i,j) 元为 1，其余元为 0 的基本矩阵记作 E_{ij}. 在 $M_{s\times n}(K)$ 里，基本矩阵共有 $s\cdot n$ 个. 例如，在 $M_{2\times 3}$ 里，基本矩阵共有 6 个，它们是：

$$E_{11}=\begin{pmatrix}1&0&0\\0&0&0\end{pmatrix},\quad E_{12}=\begin{pmatrix}0&1&0\\0&0&0\end{pmatrix},\quad E_{13}=\begin{pmatrix}0&0&1\\0&0&0\end{pmatrix}$$

$$E_{21}=\begin{pmatrix}0&0&0\\1&0&0\end{pmatrix},\quad E_{22}=\begin{pmatrix}0&0&0\\0&1&0\end{pmatrix},\quad E_{23}=\begin{pmatrix}0&0&0\\0&0&1\end{pmatrix}$$

在 $M_{2\times 3}(K)$ 里任一矩阵 $A=(a_{ij})$ 可以表示成基本矩阵的线性组合：

$$\begin{aligned}A&=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\end{pmatrix}\\&=a_{11}E_{11}+a_{12}E_{12}+a_{13}E_{13}+a_{21}E_{21}+a_{22}E_{22}+a_{23}E_{23}\end{aligned}$$

类似地，在 $M_{s\times n}(K)$ 里任一矩阵 $A=(a_{ij})$ 可以表示成基本矩阵的线性组合：

$$A=\sum_{i=1}^{s}\sum_{j=1}^{n}a_{ij}E_{ij} \tag{1}$$

这正是我们把 $E_{ij}(i=1,\cdots,s;j=1,\cdots,n)$ 称为基本矩阵的原因.

现在我们来看基本矩阵与矩阵相乘的特点.

设 A 是一个 $n\times m$ 矩阵，它的行向量组是 $\gamma_1,\gamma_2,\cdots,\gamma_n$. 利用矩

阵乘法的第三种表述方式，我们有

$$E_{ij}A=\begin{array}{c@{}c} i\left(\begin{array}{ccccc} & & \vdots & & \\ & & \vdots & & \\ \cdots & \cdots & 1 & \cdots & \cdots \\ & & \vdots & & \\ & & \vdots & & \\ & & \vdots & & \end{array}\right) & \begin{pmatrix}\gamma_1\\ \gamma_2\\ \vdots\\ \gamma_j\\ \vdots\\ \gamma_n\end{pmatrix} \\ \qquad\qquad j & \end{array}$$

$$=\begin{pmatrix}0\\ \vdots\\ 0\\ \gamma_j\\ 0\\ \vdots\\ 0\end{pmatrix}\text{第 } i \text{ 行} \tag{2}$$

在写出 E_{ij} 时，未标出元素处均为 0.（2）式表明，**在矩阵 A 的左边乘上 E_{ij}，就相当于把 A 的第 j 行搬到第 i 行的位置，而乘积矩阵的其余行全为零行.**

设 B 是一个 $r\times s$ 矩阵，它的列向量组是 $\beta_1,\beta_2,\cdots,\beta_s$. 利用矩阵乘法的第二种表述方式，我们得到

$$BE_{ij}=(\beta_1,\cdots,\beta_i,\cdots,\beta_s)\begin{array}{c} \left(\begin{array}{ccccc} & & \vdots & & \\ & & \vdots & & \\ \cdots & \cdots & 1 & \cdots & \cdots \\ & & \vdots & & \\ & & \vdots & & \end{array}\right) i \\ j \end{array}$$

$$=(0,\cdots,0,\underset{\text{第 } j \text{ 列}}{\beta_i},0,\cdots,0) \tag{3}$$

(3)式表明，**在矩阵 B 的右边乘上 E_{ij}，就相当于把 B 的第 i 列搬到**

第 j 列的位置上，而乘积矩阵的其余列全为零列.

基本矩阵 E_{kl} 的 (k,l) 元为 1，其余元为 0，因此 E_{kl} 的行向量组是

$$0,\cdots,0,\underbrace{\varepsilon_l}_{\text{第 } k \text{ 行}},0,\cdots,0$$

现在我们用 E_{ij} 乘以 E_{kl}，如果 $k=j$，则

$$E_{ij}E_{jl}=\begin{array}{c}\begin{pmatrix}0\\ \vdots\\ 0\\ \varepsilon_l\\ 0\\ \vdots\\ 0\end{pmatrix}\end{array}\begin{array}{l}\\ \\ \\ \text{第 } i \text{ 行}\\ \\ \\ \\ \end{array}=E_{il} \tag{4}$$

如果 $k\neq j$，则

$$E_{ij}E_{kl}=0 \tag{5}$$

把公式(4)和(5)合并起来写，得

$$E_{ij}E_{kl}=\begin{cases}E_{il} & \text{当 } k=j\\ 0 & \text{当 } k\neq j\end{cases} \tag{6}$$

作为基本矩阵的乘法特点的一个应用，我们来讲下面一个重要结论.

上一节我们指出，一个数量矩阵与所有跟它同级的方阵可交换. 现在我们要证明，这个命题的逆命题也成立，即

定理 4.4.1　与所有 n 级方阵可交换的矩阵一定是 n 级数量矩阵.

证明　设矩阵 A 与所有 n 级方阵可交换，则 A 必为 n 级方阵. 设 $A=(a_{ij})$. 由已知条件，A 与 n 级基本矩阵 $E_{1j}(j=1,2,\cdots,n)$ 可交换，即 $E_{1j}A=AE_{1j}$. 由此得

$$
\begin{pmatrix}
a_{j1} & \cdots & a_{jj} & \cdots & a_{jn} \\
0 & \cdots & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & \cdots & 0
\end{pmatrix}
$$

$$
= \begin{array}{c}
\begin{pmatrix}
0 & \vdots & 0 & a_{11} & 0 & \vdots & 0 \\
0 & \vdots & 0 & a_{21} & 0 & \vdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \vdots & 0 & a_{n1} & 0 & \vdots & 0
\end{pmatrix} \\
\text{第 } j \text{ 列}
\end{array}
\tag{7}
$$

从(7)式得，$a_{jj}=a_{11}$；$a_{jl}=0$，当 $l\neq j$. 由于 j 可取 $1,2,\cdots,n$，因此 $a_{11}=a_{22}=a_{33}=\cdots=a_{nn}$，并且 $a_{jl}=0$，当 $l\neq j$，$1\leqslant j,l\leqslant n$. 这说明

$$
A=a_{11}I
$$

▌

(二) 对角矩阵

主对角线以外的元素全为零的方阵称为**对角矩阵**，它形如

$$
\begin{pmatrix}
a_1 & 0 & \cdots & 0 \\
0 & a_2 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & a_n
\end{pmatrix}
$$

简记作 $\mathrm{diag}\{a_1,a_2,\cdots,a_n\}$.

我们来看对角矩阵与矩阵相乘的特点.

设 $D=\mathrm{diag}\{d_1,d_2,\cdots,d_n\}$，$A$ 是一个 $n\times m$ 矩阵，它的行向量组是 $\gamma_1,\gamma_2,\cdots,\gamma_n$. 则

$$
DA=\begin{pmatrix}
d_1 & 0 & 0 & \cdots & 0 \\
0 & d_2 & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & \cdots & d_n
\end{pmatrix}
\begin{pmatrix}
\gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n
\end{pmatrix}
=\begin{pmatrix}
d_1\gamma_1 \\ d_2\gamma_2 \\ \vdots \\ d_n\gamma_n
\end{pmatrix}
\tag{8}
$$

(8)式表明，**在矩阵 A 左边乘对角矩阵 D，就相当于把 A 的各行分别乘上 D 的相应的主对角元.**

设 B 是一个 $s\times n$ 矩阵，它的列向量组是 $\beta_1,\beta_2,\cdots,\beta_n$，则

$$
\begin{aligned}
BD &= (\beta_1,\beta_2,\cdots,\beta_n)\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} \\
&= (d_1\beta_1, d_2\beta_2,\cdots,d_n\beta_n) \qquad (9)
\end{aligned}
$$

(9)式表明，**在矩阵 B 右边乘上对角矩阵 D，就相当于把 B 的各列分别乘以 D 的相应的主对角元.**

特别地，有

$$
\begin{aligned}
&\operatorname{diag}\{a_1,a_2,\cdots,a_n\}\cdot\operatorname{diag}\{b_1,b_2,\cdots,b_n\} \\
&=\operatorname{diag}\{a_1b_1,a_2b_2,\cdots,a_nb_n\} \qquad (10)
\end{aligned}
$$

(10)式表明，两个 n 级对角矩阵的乘积仍是对角矩阵. 两个对角矩阵相乘时，只要把相应的主对角元相乘即可. 显然有

$$
\begin{aligned}
&\operatorname{diag}\{a_1,a_2,\cdots,a_n\} + \operatorname{diag}\{b_1,b_2,\cdots,b_n\} \\
&=\operatorname{diag}\{a_1+b_1,a_2+b_2,\cdots,a_n+b_n\} \qquad (11)
\end{aligned}
$$

$$
\begin{aligned}
&k\cdot\operatorname{diag}\{a_1,a_2,\cdots,a_n\} \\
&=\operatorname{diag}\{ka_1,ka_2,\cdots,ka_n\} \qquad (12)
\end{aligned}
$$

（三） 初等矩阵

在线性方程组的理论中，矩阵的初等行(列)变换起了重要作用. 现在我们来建立矩阵的初等变换与矩阵乘法的联系，为此需要引进初等矩阵的概念.

定义 1 由单位矩阵经过一次初等变换得到的矩阵称为**初等矩阵**.

我们先看 n 级单位矩阵 I 经过一次初等行变换得到的矩阵.

把 I 的第 j 行的 k 倍加到第 i 行得到的初等矩阵记作 $P(i,j(k))$，它是

$$P(i,j(k))=\begin{array}{c@{\quad}l}\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&\cdots&k&&\\&&&\ddots&\vdots&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}&\begin{matrix}\\ \\ \text{第 } i \text{ 行}\\ \\ \text{第 } j \text{ 行}\\ \\ \\ \end{matrix}\end{array}$$

把 I 的第 i 行与第 j 行互换位置得到的初等矩阵记作 $P(i,j)$，它是

$$P(i,j)=\begin{array}{c@{\quad}l}\begin{pmatrix}1&&&&&&&&&\\&\ddots&&&&&&&&\\&&0&\cdots&\cdots&\cdots&1&&&\\&&\vdots&1&&&\vdots&&&\\&&\vdots&&\ddots&&\vdots&&&\\&&\vdots&&&1&\vdots&&&\\&&1&\cdots&\cdots&\cdots&0&&&\\&&&&&&&1&&\\&&&&&&&&\ddots&\\&&&&&&&&&1\end{pmatrix}&\begin{matrix}\\ \\ \text{第 } i \text{ 行}\\ \\ \\ \\ \text{第 } j \text{ 行}\\ \\ \\ \\ \end{matrix}\end{array}$$

把 I 的第 i 行乘以一个非零数 c 得到的初等矩阵记作 $P(i(c))$，它是

$$P(i(c)) = \begin{array}{c@{}l} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \text{第}\, i\, \text{行} \\ \\ \\ \\ \end{array} \end{array}$$

对 I 作一次初等列变换得到的矩阵也包括在上面列举的三类矩阵中. 例如,把 I 的第 i 列的 k 倍加到第 j 列上得到的矩阵是 $P(i,j(k))$;把 I 的第 i 列与第 j 列互换位置得到的矩阵仍是 $P(i,j)$;把 I 的第 i 列乘以一个非零数 c 得到的矩阵仍是$P(i(c))$. 因此上述三类矩阵就是全部的初等矩阵.

利用初等矩阵可以把矩阵的初等变换与矩阵的乘法联系起来. 为此我们来看初等矩阵与矩阵相乘的特点.

定理 4.4.2 在一个 $s\times n$ 矩阵 A 的左边乘上一个 s 级初等矩阵,就相当于对 A 作一次相应的初等行变换;在 A 的右边乘上一个 n 级初等矩阵,就相当于对 A 作一次相应的初等列变换.

证明 设 A 的行向量组是 $\gamma_1,\gamma_2,\cdots,\gamma_s$,则

$$P(i,j(k))A$$

$$= \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & k & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_i \\ \vdots \\ \gamma_j \\ \vdots \\ \gamma_s \end{pmatrix} = \begin{array}{c@{}l} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_i + k\gamma_j \\ \vdots \\ \gamma_j \\ \vdots \\ \gamma_s \end{pmatrix} & \begin{array}{l} \\ \\ i\,\text{行} \\ \\ j\,\text{行} \\ \\ \\ \end{array} \end{array}$$

这相当于 把 A 的第 j 行的 k 倍加到第 i 行上.

$$P(i,j)A$$

$$=\begin{pmatrix}1&&&&&&&&\\&\ddots&&&&&&&\\&&0&\cdots&\cdots&\cdots&1&&\\&&\vdots&1&&&\vdots&&\\&&\vdots&&\ddots&&\vdots&&\\&&\vdots&&&1&\vdots&&\\&&1&\cdots&\cdots&\cdots&0&&\\&&&&&&&1&\\&&&&&&&&\ddots\\&&&&&&&&&1\end{pmatrix}\begin{pmatrix}\gamma_1\\\vdots\\\gamma_i\\\vdots\\\gamma_j\\\vdots\\\gamma_s\end{pmatrix}=\begin{array}{l}\begin{pmatrix}\gamma_1\\\vdots\\\gamma_j\\\vdots\\\gamma_i\\\vdots\\\gamma_s\end{pmatrix}\begin{array}{l}\\\\i\ 行\\\\j\ 行\\\\\\\end{array}\end{array}$$

这相当于把 A 的第 i 行与第 j 行互换位置.

$$P(i(c))A$$

$$=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&c&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{pmatrix}\gamma_1\\\vdots\\\gamma_i\\\vdots\\\gamma_s\end{pmatrix}=\begin{pmatrix}\gamma_1\\\vdots\\c\gamma_i\\\vdots\\\gamma_s\end{pmatrix}\begin{array}{l}\\\\i\ 行\\\\\\\end{array}$$

这相当于把 A 的第 i 行乘以一个非零数 c.

在 A 的右边乘上一个 n 级初等矩阵的情形可类似证明. 我们这里只证 $P(i,j(k))$ 的情形,其余两种情形留给读者作为练习. 设 A 的列向量组是 $\alpha_1,\alpha_2,\cdots,\alpha_n$,则

$$AP(i,j(k))$$

$$=(\alpha_1,\cdots,\alpha_i,\cdots,\alpha_j,\cdots,\alpha_n)\begin{array}{c}\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&\cdots&k&&\\&&&\ddots&\vdots&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\\ \qquad i\text{列}\quad j\text{列}\qquad\end{array}$$

$$=(\alpha_1,\cdots,\underset{i\text{列}}{\alpha_i},\cdots,\underset{j\text{列}}{k\alpha_i+\alpha_j},\cdots,\alpha_n)$$

这相当于把 A 的第 i 列的 k 倍加到第 j 列上. ▎

定理 4.4.2 也可以叙述成：**对一个 $s\times n$ 矩阵 A 作一次初等行变换，就相当于在 A 的左边乘上一个相应的 s 级初等矩阵；对 A 作一次初等列变换，就相当于在 A 的右边乘上一个相应的 n 级初等矩阵.**

定理 4.4.2 把矩阵的初等变换与矩阵的乘法相联系，这是非常有用的. 后面几节我们将会看到它的应用.

（四） 上(下)三角矩阵

我们已经知道，主对角线下方的元素全为零的方阵称为**上三角矩阵**. 容易看出，方阵 A 为上三角矩阵的充分必要条件是

$$A(i,j)=0,\qquad 当\ i>j$$

显然，两个 n 级上三角矩阵的和仍为上三角矩阵；用数 k 乘一个上三角矩阵仍得上三角矩阵.

命题 4.4.1 两个 n 级上三角矩阵 A 与 B 的乘积仍为上三角矩阵，并且 AB 的主对角元等于 A 与 B 的相应主对角元的乘积，即

$$(AB)(i;i)=A(i;i)B(i;i),\quad i=1,\cdots,n.$$

证明

$$(AB)(i;j)=\sum_{k=1}^{n}A(i;k)B(k;j)$$

$$=\sum_{k=1}^{j}A(i;k)B(k;j)+\sum_{k=j+1}^{n}A(i;k)B(k;j) \quad (13)$$

设 $i>j$. 当 $1\leqslant k\leqslant j$ 时，由于 $k\leqslant j<i$，以及 A 是上三角矩阵，因此 $A(i,k)=0$. 从而(13) 式右端第一部分为 0. 当 $j<k\leqslant n$ 时，因为 B 是上三角矩阵，所以 $B(k;j)=0$. 从而(13) 式右端的第二部分为 0. 因此，当 $i>j$ 时，有 $(AB)(i;j)=0$. 这证明了 AB 是上三角矩阵. 同理可得

$$\begin{aligned}
&(AB)(i;i)\\
&=\sum_{k=1}^{n}A(i;k)B(k;i)\\
&=\sum_{k=1}^{i-1}A(i;k)B(k;i)+\sum_{k=i+1}^{n}A(i;k)B(k;i)+A(i;i)B(i;i)\\
&=\sum_{k=1}^{i-1}0\cdot B(k;i)+\sum_{k=i+1}^{n}A(i;k)\cdot 0+A(i;i)B(i;i)\\
&=A(i;i)B(i;i)
\end{aligned}$$

∎

主对角线上方的元素全为零的方阵称为**下三角矩阵**. 方阵 A 为下三角矩阵当且仅当 $A(i;j)=0$，当 $i<j$. 下三角矩阵具有与上三角矩阵类似的性质，留给读者练习.

（五） 对称矩阵

定义 2　一个矩阵 A 如果满足

$$A'=A$$

则称 A 是**对称矩阵**.

容易从定义看出，对称矩阵一定是方阵，并且它的 (i,j) 元与 (j,i) 元相等. 于是对称矩阵必形如

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{12} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn} \end{pmatrix}. \tag{14}$$

两个 n 级对称矩阵的和仍然是对称矩阵. 这是因为如果 A,B 是 n 级对称矩阵,则有 $A' = A, B' = A$,从而有

$$(A + B)' = A' + B' = A + B$$

同样可证:若 A 是对称矩阵,则 kA 也是对称矩阵.

但是要注意:两个对称矩阵的乘积不一定是对称矩阵. 例如,设

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

显然, A 与 B 都是对称矩阵. 但是

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

AB 不是对称矩阵.

命题 4.4.2　设 A 与 B 都是 n 级对称矩阵,则 AB 为对称矩阵的充分必要条件是 A 与 B 可交换.

证明　因为 A 与 B 都是对称矩阵,所以

$$(AB)' = B'A' = BA$$

于是

$$AB \text{ 为对称矩阵} \Leftrightarrow (AB)' = AB \Leftrightarrow BA = AB \qquad \blacksquare$$

(六) 斜对称矩阵

定义 3　一个矩阵 A 如果满足

$$A' = -A$$

则称 A 是一个**斜对称矩阵**.

容易看出,斜对称矩阵一定是方阵,并且它的(j,i) 元等于它

的(i,j)元的相反数，它的主对角元全是零. 因此，斜对称矩阵必形如

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{pmatrix}$$

两个n级斜对称矩阵的和仍是斜对称矩阵；数乘斜对称矩阵仍得斜对称矩阵. 但是两个斜对称矩阵的乘积不一定是斜对称矩阵.

命题 4.4.3 数域K上奇数级斜对称矩阵的行列式等于零.

证明 设A是奇数n级的斜对称矩阵，则$A'=-A$. 从而$|A'|=|-A|$. 因为$|A'|=|A|$，并且$|-A|=(-1)^n|A|$，所以，$|A|=(-1)^n|A|$. 由于n是奇数，于是由上式得$|A|=-|A|$，从而$2|A|=0$. 因此$|A|=0$. ▌

习 题 4.4

1. 如果D是对角矩阵并且它的主对角元两两不同，则与D可交换的矩阵一定是对角矩阵.

2. 矩阵A与所有n级对角矩阵可交换的充分必要条件是A为n级对角矩阵.

3. 证明：两个n级下三角矩阵的乘积仍是下三角矩阵，并且乘积矩阵的主对角元等于因子矩阵的相应主对角元的乘积.

4. 证明：对任一$s\times n$矩阵A，都有AA'，$A'A$是对称矩阵.

5. 证明：对任一方阵A，都有$A+A'$是对称矩阵，$A-A'$是斜对称矩阵.

6. 证明：数域K上任一n级矩阵都可以表示成一个对称矩阵与一个斜对称矩阵之和，并且表法唯一.

7. 证明：如果A与B都是n级斜对称矩阵，则$AB-BA$也是斜对称矩

阵.

8. 证明：若 A 是一个实对称矩阵(即实数域上的对称矩阵)，并且 $A^2 = 0$，则 $A = 0$.

9. 如果 A 与 B 都是 n 级对称矩阵，则矩阵

$$C = ABAB\cdots ABA$$

是对称矩阵.

10. 证明：两个斜对称矩阵的乘积是对称矩阵当且仅当所给的两个矩阵是可交换的.

11. 证明：两个斜对称矩阵 A 与 B 的乘积是斜对称矩阵当且仅当 $AB = -BA$.

12. 证明：矩阵的 2° 型初等行变换(即，两行互换)可以通过一些 1° 型和 3° 型的初等行变换实现.

13. 设

$$A = \begin{pmatrix} \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$$B = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m2} & \cdots & b_{mm} \end{pmatrix}$$

证明：矩阵方程

$$AX = XB$$

有非零解.

§5 矩阵乘积的秩 · 方阵的迹

5.1 矩阵乘积的秩

关于矩阵乘积的秩，我们有下面的结果：

定理 4.5.1　设 $A=(a_{ij})\in M_{s\times n}(K)$，$B=(b_{ij})\in M_{n\times m}(K)$，则

$$\operatorname{rank}(AB)\leqslant\min\{\operatorname{rank}(A),\operatorname{rank}(B)\}$$

证明　设 A 的列向量组是 $\alpha_1,\alpha_2,\cdots,\alpha_n$，则

$$AB=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix}$$

$$=\left(\sum_{k=1}^{n}b_{k1}\alpha_k,\sum_{k=1}^{n}b_{k2}\alpha_k,\cdots,\sum_{k=1}^{n}b_{km}\alpha_k\right) \tag{1}$$

(1)式表明，AB 的列向量组可以由 A 的列向量组线性表出，因此 $\operatorname{rank}(AB)\leqslant\operatorname{rank}(A)$. 这表明乘积矩阵的秩不超过左矩阵的秩. 利用这个结论我们可以得到

$$\begin{aligned}\operatorname{rank}(AB)&=\operatorname{rank}[(AB)']=\operatorname{rank}(B'A')\leqslant\operatorname{rank}(B')\\&=\operatorname{rank}(B)\end{aligned}$$

因此 $\operatorname{rank}(AB)\leqslant\min\{\operatorname{rank}(A),\operatorname{rank}(B)\}$. ▌

用数学归纳法，可以把定理 4.5.1 推广到多个因子的情形，即我们有

推论 4.5.1

$$\operatorname{rank}(A_1,A_2,\cdots,A_r)\leqslant\min_{1\leqslant j\leqslant r}\operatorname{rank}(A_j)$$ ▌

对于实数域上的矩阵 A 与它的转置 A' 的乘积的秩有更进一步的结果：

命题 4.5.1　设 A 是实数域上一个 $s\times n$ 矩阵(简称为 $s\times n$ 实矩阵)，则

$$\operatorname{rank}(A'A)=\operatorname{rank}(AA')=\operatorname{rank}(A)$$

证明　如果我们能够证明 n 元齐次线性方程组 $(A'A)X=0$ 与 $AX=0$ 同解，那么它们的解空间一致，从解空间的维数公式得出

$$n-\operatorname{rank}(A'A)=n-\operatorname{rank}(A)$$

由此得，$\operatorname{rank}(A'A)=\operatorname{rank}(A)$.

现在来证明 $(A'A)X=0$ 与 $AX=0$ 同解. 设 η 是 $AX=0$ 的解，则有 $(A'A)\eta=A'(A\eta)=A'0=0$，这表明 η 也是方程组 $(A'A)X=0$ 的解. 反之，设 η 是方程组 $(A'A)X=0$ 的解，则有

$$(A'A)\eta=0 \tag{2}$$

(2)式 两边同时左乘 η' 得

$$\eta'(A'A)\eta=0$$

即

$$(A\eta)'(A\eta)=0 \tag{3}$$

设

$$A\eta=\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_s \end{pmatrix}$$

则由(3)式得

$$c_1^2+c_2^2+\cdots+c_s^2=0 \tag{4}$$

因为 $c_1,c_2,\cdots,c_s$ 都是实数，所以从(4) 式可以推出

$$c_1=c_2=\cdots=c_s=0$$

于是得 $A\eta=0$. 这表明 η 是方程组 $AX=0$ 的解. 因此方程组 $AX=0$ 与方程组 $(A'A)X=0$ 同解. 据前面所述，得

$$\operatorname{rank}(A'A)=\operatorname{rank}(A)$$

我们有

$$\operatorname{rank}(AA')=\operatorname{rank}[(A')'(A')]=\operatorname{rank}(A')=\operatorname{rank}(A) \quad \blacksquare$$

注意：对于复数域上的矩阵 A，$A'A$ 与 A 的秩不一定相等. 例如，设

$$A=\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$$

则

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

于是 $\operatorname{rank}(A'A) = 0$,但是容易看出 $\operatorname{rank}(A) = 1$.

对于复数域上的矩阵,有一个类似于命题 4.5.1 那样的结果,见本节习题第 7 题.

象从命题 4.5.1 的证明中看到的那样,齐次线性方程组的解空间的维数与系数矩阵的秩的关系常常可用来讨论有关矩阵的秩的问题.

5.2 方阵的迹

现在介绍方阵的迹的概念.

定义 1 n 级方阵 $A = (a_{ij})$ 的主对角元之和称为 A 的**迹**,记作 $T_r(A)$,即

$$T_r(A) := a_{11} + a_{12} + \cdots + a_{nn} \tag{5}$$

方阵的迹是集合 $M_n(K)$ 到数域 K 的一个映射,也就是说,方阵的迹是定义在集合 $M_n(K)$ 上的一个函数. 这个函数是相当有用的. 读者如果以后有机会学习群表示论的课程的话,会看到迹函数在群表示论中起着十分重要的作用. 方阵的迹在目前线性代数课中也有用.

迹函数具有以下性质:

1° $T_r(A + B) = T_r(A) + T_r(B)$;

2° $T_r(kA) = kT_r(A)$;

3° $T_r(A') = T_r(A)$;

4° $T_r(AB) = T_r(BA)$.

证明 1°—3°很容易直接验证. 现在证明 4°.

设 $A = (a_{ij}), B = (b_{ij}) \in M_n(K)$,则

$$T_r(AB) = \sum_{i=1}^{n} (AB)(i;i) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{ki}$$

$$T_r(BA)=\sum_{k=1}^{n}(BA)(k;k)=\sum_{k=1}^{n}\sum_{i=1}^{n}b_{ki}a_{ik}$$
$$=\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}b_{ki}$$
因此，$T_r(AB)=T_r(BA)$. ▌

性质 4°表明，虽然矩阵的乘法不适合交换律，但是 AB 与 BA 的迹却是相等的. 这个性质挺有用.

习 题 4.5

1. 证明：$\text{rank}(A+B)\leqslant\text{rank}(A)+\text{rank}(B)$.

2. 证明：对于实数域上的任一 $s\times n$ 矩阵 A，都有
$$\text{rank}(AA'A)=\text{rank}(A)$$

3. 设 A 是复数域上的矩阵，把 A 的每一个元素用共轭复数来代替所得到的矩阵记作 $\overline{A}$. 设 B 也是复数域上矩阵，证明：
$$\overline{AB}=\overline{A}\,\overline{B},\quad \overline{(A+B)}=\overline{A}+\overline{B},\quad \overline{kA}=\overline{k}\,\overline{A}$$

* 4. 设 A 是复矩阵，令 $A^*=\overline{A'}$. 证明：
$$(A+B)^*=A^*+B^*,\quad (AB)^*=B^*A^*,$$
$$(kA)^*=\overline{k}A^*,\quad (A^*)^*=A$$

* 5. 设 A 是复矩阵，如果 $A^*=A$，则称 A 为 Hermite **矩阵**. 证明：对任意复矩阵 B，都有 BB^*，B^*B 是 Hermite 矩阵.

* 6. 证明：对任意复矩阵 A，有 $\text{rank}(A^*)=\text{rank}(A)$.

* 7. 证明：对任意复矩阵 A，有
$$\text{rank}(AA^*)=\text{rank}(A^*A)=\text{rank}(A)$$

8. 一个矩阵称为**行(列)满秩矩阵**，如果它的行(列)向量组是线性无关的. 证明：如果一个 $m\times n$ 矩阵 A 的秩为 r，则有 $m\times r$ 的列满秩矩阵 B 和 $r\times n$ 的行满秩矩阵 C，使得 $A=BC$.

9. 举例说明 $T_r(AB)\neq T_r(A)\cdot T_r(B)$.

10. 证明：等式 $AB-BA=I$ 无论对于数域 K 上怎样的矩阵 A 和 B 都不成立.

11. 证明：对任一 2 级方阵 A，有

$$A^2 - T_r(A)A + |A|I_2 = 0$$

§6 矩阵的分块

这一节我们介绍在处理有关矩阵运算(特别是乘法)的问题中一个重要方法:矩阵的分块.

先看一个例子. 设

$$A = \left(\begin{array}{ccc:cc} 1 & 0 & 0 & 2 & 5 \\ 0 & 1 & 0 & 3 & -2 \\ 0 & 0 & 1 & -1 & 6 \\ \hdashline 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{array}\right)$$

把 A 的行分成两组,前三行为第一组,后两行为第二组;把 A 的列分成两组,前三列为第一组,后两列为第二组. 则 A 被分成 4 个矩形小块,其中每一小块是 A 的一个子矩阵,我们把位于第 i 个行组与第 j 个列组的交叉处的子矩阵记作 A_{ij},则

$$A_{11} = I_3 \quad A_{12} = \begin{pmatrix} 2 & 5 \\ 3 & -2 \\ -1 & 6 \end{pmatrix} \quad A_{21} = 0_{2\times 3} \quad A_{22} = 4I_2$$

于是 A 可以看成是由这 4 个子矩阵组成,即把 A 写成

$$A = \begin{pmatrix} I_3 & A_{12} \\ 0_{2\times 3} & 4I_2 \end{pmatrix}$$

这时 A 称为分块矩阵.

一般地,把一个矩阵 A 的行与列都分成若干组,从而 A 被分成若干个矩形小块,把 A 看成是由这些小矩阵组成的,这称为**矩阵的分块**,此时 A 称为**分块矩阵**.

矩阵分块的好处是使得矩阵的结构变得更明显清楚,譬如上

面的例子中，A 的左上角是一个 3 级单位矩阵，左下角是 2×3 零矩阵，右下角是 2 级数量矩阵 $4I$. 矩阵分块的更重要的用处是矩阵的运算可以通过它们的分块矩阵形式来进行，从而较为简便. 我们先通过一个例子来说明这点. 设

$$A=\left[\begin{array}{cc:c} a_{11} & a_{12} & a_{13} \\ \hdashline a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right],\qquad B=\left[\begin{array}{c:c:cc} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ \hdashline b_{31} & b_{32} & b_{33} & b_{34} \end{array}\right]$$

把 A,B 分块：A 的行分成两组，前一行为第一组，后两行为第二组；A 的列分成两组，前两列为第一组，后一列为第二组. B 的行分成两组，前两行为第一组，后一行为第二组；B 的列分成三组，第一，二组分别有一列，第三组有两列. 于是 A,B 分别可写成分块矩阵的形式：

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},\qquad B=\begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix}$$

其中 A_{ij},B_{ij} 是什么样的小矩阵，请读者写出. 这里我们只写出：

$$A_{21}=\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix},\qquad A_{22}=\begin{bmatrix} a_{23} \\ a_{33} \end{bmatrix}$$

$$B_{13}=\begin{bmatrix} b_{13} & b_{14} \\ b_{23} & b_{24} \end{bmatrix},\qquad B_{23}=(b_{33},b_{34})$$

现在我们试图模仿矩阵乘法的定义来规定分块矩阵的乘法运算，即令

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix}$$

$$:=\begin{bmatrix} A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22} & A_{11}B_{13}+A_{12}B_{23} \\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22} & A_{21}B_{13}+A_{22}B_{23} \end{bmatrix}$$

把上式右端的矩阵记作 C. 现在要问：把矩阵 A 与 B 按照它们的分

块矩阵形式相乘得到的矩阵C是否等于A与B直接相乘时得到的矩阵AB? 如果回答是肯定的,那么上述对于分块矩阵的乘法的规定就是合理的,从而在求乘积矩阵AB时,也就可以用A,B的分块矩阵形式相乘,这常常比较简便. 我们说,回答的确是肯定的. 理由如下:

用C_{pq}表示C的分块矩阵形式中位于第p行与第q列交叉处的小矩阵,即

$$C_{11}=A_{11}B_{11}+A_{12}B_{21},\cdots,C_{13}=A_{11}B_{13}+A_{12}B_{23}$$
$$C_{21}=A_{21}B_{11}+A_{22}B_{21},\cdots,C_{23}=A_{21}B_{13}+A_{22}B_{23}$$

则

$$C=\begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \end{bmatrix}$$

由于C_{11},C_{21}分别有1,2行,所以C有3行. 由于C_{11},C_{12},C_{13}分别有1,1,2列,所以C有4列. 因此C与AB都是3×4矩阵. 我们来看C与AB的对应元素是否相等. 看它们的(3,4)元. 有

$$AB(3;4)=a_{31}b_{14}+a_{32}b_{24}+a_{33}b_{34}$$

C的(3,4)元处于子矩阵C_{23}中. C_{23}是2×2矩阵,容易看出$C(3;4)$是C_{23}的(2,2)元. 因此

$$\begin{aligned} C(3;4) &= C_{23}(2;2)=(A_{21}B_{13}+A_{22}B_{23})(2;2) \\ &= A_{21}B_{13}(2;2)+A_{22}B_{23}(2;2) \\ &= a_{31}b_{14}+a_{32}b_{24}+a_{33}b_{34} \end{aligned}$$

从而$AB(3;4)=C(3;4)$. 类似可证AB与C的其它对应元素都相等. 所以$AB=C$.

一般地,我们有

定理4.6.1 设$A=(a_{ij})_{s\times n},B=(b_{ij})_{n\times m}$,把$A$的行分成$u$组,第$i$组有$s_i$行;$A$的列分成$t$组,第$j$组有$n_j$列. 把$B$的行分成$t$组,第$j$组有$n_j$行;列分成$v$组,第$l$组有$m_l$列. 即

$$
A = \begin{array}{c} \\ s_1 \\ s_2 \\ \\ s_u \end{array}
\begin{array}{c}
\begin{array}{cccc} n_1 & n_2 & & n_t \end{array} \\
\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \cdots & \cdots & \cdots & \cdots \\ A_{u1} & A_{u2} & \cdots & A_{ut} \end{pmatrix}
\end{array}
$$

$$
B = \begin{array}{c} \\ n_1 \\ n_2 \\ \\ n_t \end{array}
\begin{array}{c}
\begin{array}{cccc} m_1 & m_2 & & m_v \end{array} \\
\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1v} \\ B_{21} & B_{22} & \cdots & B_{2v} \\ \cdots & \cdots & \cdots & \cdots \\ B_{t1} & B_{t2} & \cdots & B_{tv} \end{pmatrix}
\end{array}
$$

其中每个 A_{ij} 是 $s_i \times n_j$ 小矩阵，每个 B_{ij} 是 $n_i \times m_j$ 小矩阵. 令 $C = (C_{pq})$，其中

$$
C_{pq} = \sum_{l=1}^{t} A_{pl}B_{lq}, \quad p = 1,\cdots,u; q = 1,\cdots,v \tag{1}
$$

则 $AB = C$.

***证明** 矩阵 C 的行数为 $s_1 + s_2 + \cdots + s_u = s$；$C$ 的列数为 $m_1 + m_2 + \cdots + m_v = m$. 于是 AB 与 C 都是 $s \times m$ 矩阵.

$$
(AB)(i;j) = \sum_{k=1}^{n} a_{ik}b_{kj}
$$

设 $i = s_1 + \cdots + s_{p-1} + f$，其中 $0 < f \leqslant s_p$；设 $j = m_1 + \cdots + m_{q-1} + g$，其中 $0 < g \leqslant m_q$. 这意思是 A 的第 i 行属于 A 的第 p 个行组，B 的第 j 列属于 B 的第 q 个列组. 于是

$$
\begin{aligned}
C(i;j) &= C_{pq}(f;g) \\
&= \left(\sum_{l=1}^{t} A_{pl}B_{lq}\right)(f;g) = \sum_{l=1}^{t} (A_{pl}B_{lq})(f;g) \\
&= \sum_{l=1}^{t} \sum_{r=1}^{n_l} A_{pl}(f;r)B_{lq}(r;g) \\
&= \sum_{r=1}^{n_1} A_{p1}(f;r)B_{1q}(r;g) + \cdots + \sum_{r=1}^{n_t} A_{pt}(f;r)B_{tq}(r;g)
\end{aligned}
$$

$$= \sum_{k=1}^{n_1} a_{ik}b_{kj} + \cdots + \sum_{k=n_1+\cdots+n_{l-1}+1}^{n} a_{ik}b_{kj}$$

$$= \sum_{k=1}^{n} a_{ik}b_{kj} = (AB)(i;j)$$

因此 $AB = C$. ▌

定理4.6.1告诉我们，我们模仿矩阵乘法的定义来规定分块矩阵的乘法是合理的，从而在求乘积矩阵 AB 时，就可以用它们的分块矩阵形式相乘. 注意：在把 A,B 分块时应满足下列两个条件：

1)左矩阵 A 的列组数等于右矩阵 B 的行组数；

2)左矩阵 A 的每个列组所含列数等于右矩阵 B 的相应行组所含的行数.

总而言之，左矩阵的列的分法应当与右矩阵的行的分法一致.

还要注意：小矩阵之间的相乘次序应当是把左分块矩阵的子矩阵写在左边，把右分块矩阵的子矩阵写在右边，不能颠倒次序.

例1 设 A 是 $s \times n$ 矩阵，B 是 $n \times m$ 矩阵，并且 B 的列向量组为 $\beta_1,\beta_2,\cdots,\beta_m$. 则

$$AB = (A\beta_1, A\beta_2, \cdots, A\beta_m) \tag{2}$$

证明 把 A 的所有行作为一组，A 的所有列作为一组；把 B 的所有行作为一组，B 的列分为 m 组，每组含 1 列. 则

$$AB = (A)(\beta_1,\beta_2,\cdots,\beta_m) = (A\beta_1,A\beta_2,\cdots,A\beta_m) \quad ▌$$

公式(2)是非常有用的. 譬如由公式(2)可以得到下面的例2的结论.

例2 设 A 是 $s \times n$ 矩阵，且 $A \neq 0$；$B_{n\times m}$ 的列向量组是 $\beta_1,\beta_2,\cdots,\beta_m$；$C_{s\times m}$ 的列向量组是 $\gamma_1,\gamma_2,\cdots,\gamma_m$. 则

1) $AB = 0 \Leftrightarrow \beta_1,\beta_2,\cdots,\beta_m$ 都是齐次线性方程组 $AX = 0$ 的解；

2) $AB = C \Leftrightarrow \beta_j$ 是线性方程组 $AX = \gamma_j$ 的解，　$j = 1,\cdots,m$.

证明 1) 因为 $AB = (A\beta_1,\cdots,A\beta_m)$，所以

$$AB = 0 \Leftrightarrow A\beta_1 = 0, A\beta_2 = 0, \cdots, A\beta_m = 0$$
$$\Leftrightarrow \beta_1, \beta_2, \cdots, \beta_m \text{ 都是方程组 } AX = 0 \text{ 的解}$$

2)同上理得

$$AB = C \Leftrightarrow (A\beta_1, \cdots, A\beta_m) = (\gamma_1, \cdots, \gamma_m)$$
$$\Leftrightarrow A\beta_j = \gamma_j, \quad j = 1, \cdots, m$$
$$\Leftrightarrow \beta_j \text{ 是方程组 } AX = \gamma_j \text{ 的解}, j = 1, \cdots, m \qquad \blacksquare$$

例2 的第1)个结论可以使我们利用齐次线性方程组的理论去解决矩阵理论中涉及到$AB=0$这样一类问题. 在本节的习题中有好几个题是属于这种类型.

设 A,B 都是 $s \times n$ 矩阵,并且用相同的分法把它们写成分块矩阵形式:

$$A = (A_{ij}), \quad B = (B_{ij})$$

其中 A_{ij}, B_{ij} 都是 $s_i \times n_j$ 矩阵,$s_1 + s_2 + \cdots + s_u = s, n_1 + n_2 + \cdots + n_t = n$,则容易看出

$$A + B = (A_{ij} + B_{ij})$$

同样容易看出

$$kA = (kA_{ij})$$

命题4.6.1　设 A 是 $s \times n$ 矩阵,把 A 分块写成

$$A = \begin{array}{c} \\ s_1 \\ s_2 \\ \\ s_u \end{array} \begin{array}{c} \begin{array}{cccc} n_1 & n_2 & & n_t \end{array} \\ \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \cdots & \cdots & \cdots & \cdots \\ A_{u1} & A_{u2} & \cdots & A_{ut} \end{pmatrix} \end{array}$$

则

$$A' = \begin{pmatrix} A'_{11} & A'_{21} & \cdots & A'_{u1} \\ A'_{12} & A'_{22} & \cdots & A'_{u2} \\ \cdots & \cdots & \cdots & \cdots \\ A'_{1t} & A'_{2t} & \cdots & A'_{ut} \end{pmatrix}$$

证明 显然上式左右两端都是 $n \times s$ 矩阵，现在来证它们的 (i,j) 元对应相等. 设

$$i = n_1 + \cdots + n_{p-1} + f, \quad j = s_1 + \cdots + s_{q-1} + g$$

其中 $0 < f \leqslant n_p$, $0 < g \leqslant s_q$. 则要证的等式右端的矩阵的 (i,j) 元为 $A'_{qp}(f;g)$. 我们有

$$A'_{qp}(f;g) = A_{qp}(g;f) = A(j;i) = A'(i;j)$$

因此要证的等式的确成立. ▍

注意：分块矩阵的转置不仅要把第 i 行组写成第 i 列组($i = 1, \cdots, u$)，而且要把每个子矩阵转置.

矩阵的分块还有很多应用，这在今后会逐渐看到 .

现在介绍两类特殊的分块矩阵.

主对角线上的所有小块都是方阵，其余小块全为零矩阵的分块矩阵称为**分块对角矩阵**(或**准对角矩阵**)，它形如

$$\begin{pmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & A_s \end{pmatrix} \tag{3}$$

其中 $A_1, A_2, \cdots, A_s$ 都是方阵. 分块对角矩阵(3) 可简记成 $\mathrm{diag}\{A_1, A_2, \cdots, A_s\}$，它一定是方阵.

注意：对角矩阵可看成是分块对角矩阵(每个小块都是1级矩阵). 但是一般地，分块对角矩阵不是对角矩阵.

设 $A = \mathrm{diag}\{A_1, A_2, \cdots, A_s\}$，其中 A_i 是 n_i 级方阵，$i = 1, \cdots, s$. 设 B 是具有与 A 相应分法的分块矩阵，即 B 的行分成 s 组，第 i 组含 n_i 行. 由于分块矩阵的乘法与矩阵的乘法遵循同样的规则，因此从对角矩阵与矩阵相乘的特点立即得出分块对角矩阵与分块矩阵相乘的特点：

在分块矩阵 B 的左边乘上分块对角矩阵 A，就相当于在 B 的每一个小块行里的子矩阵的左边乘上 A 的相应的主对角小块.

设 C 的列分成 s 组，第 j 组含 n_j 列，则在分块矩阵 C 的右边乘

上分块对角矩阵 A，就相当于在 C 的每一个小块列里的子矩阵的右边乘上 A 的相应的主对角小块.

特别地，若 $B=\text{diag}\{B_1,B_2,\cdots,B_s\}$，并且 B_i 是 n_i 级方阵，$i=1,\cdots,s$，则

$$\begin{aligned}&\text{diag}\{A_1,A_2,\cdots,A_s\}\cdot\text{diag}\{B_1,B_2,\cdots,B_s\}\\ =&\text{diag}\{A_1B_1,A_2B_2,\cdots,A_sB_s\}\end{aligned}\tag{4}$$

(4)式表明，两个具有相同分法的分块对角矩阵的乘积仍是一个分块对角矩阵.

显然有

$$\begin{aligned}&\text{diag}\{A_1,A_2,\cdots,A_s\}+\text{diag}\{B_1,B_2,\cdots,B_s\}\\ =&\text{diag}\{A_1+B_1,A_2+B_2,\cdots,A_s+B_s\}\end{aligned}\tag{5}$$

即两个具有相同分法的分块对角矩阵的和仍是分块对角矩阵.

也有

$$k\cdot\text{diag}\{A_1,\cdots,A_s\}=\text{diag}\{kA_1,\cdots,kA_s\}\tag{6}$$

主对角线上的所有小块都是方阵，而位于主对角线下方的所有小块都是零矩阵的分块矩阵称为**分块上三角矩阵**（或**准上三角矩阵**），它形如

$$\begin{pmatrix}A_{11}&A_{12}&\cdots&A_{1s}\\0&A_{22}&\cdots&A_{2s}\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&A_{ss}\end{pmatrix}\tag{7}$$

其中 $A_{ii}(i=1,\cdots,s)$ 都是方阵，于是分块上三角矩阵一定是方阵.（注：也可以讨论不是方阵的分块上三角矩阵，这时，它的主对角线上小块有的不是方阵. 但是本书只讨论主对角线上所有小块都是方阵的分块上三角矩阵.）

注意：上三角矩阵可看成是分块上三角矩阵，但是一般地，分块上三角矩阵不是上三角矩阵.

与上三角矩阵的乘法特点类似，两个具有相同分法的分块上三角矩阵的乘积仍是分块上三角矩阵，并且乘积矩阵的主对角小

块等于因子矩阵的相应主对角小块的乘积. 显然，两个具有相同分法的分块上三角矩阵的和仍是分块上三角矩阵. 数乘分块上三角矩阵仍为分块上三角矩阵.

类似地，可讨论**分块下三角矩阵**，它形如

$$\begin{pmatrix} A_{11} & 0 & \cdots & 0 \\ A_{21} & A_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ A_{s1} & A_{s2} & \cdots & A_{ss} \end{pmatrix} \tag{8}$$

其中 $A_{11},A_{22},\cdots,A_{ss}$ 都是方阵. 于是分块下三角矩阵一定是方阵.

第二章§8我们曾运用拉普拉斯(Laplace)定理证明了

$$\begin{vmatrix} A & 0 \\ C & B \end{vmatrix} = |A||B| \tag{9}$$

其中 A,B 都是方阵. 这个结论很容易推广到

$$\begin{vmatrix} A_{11} & 0 & \cdots & 0 \\ A_{21} & A_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ A_{s1} & A_{s2} & \cdots & A_{ss} \end{vmatrix} = |A_{11}| \cdot |A_{22}| \cdot \cdots \cdot |A_{ss}| \tag{10}$$

即分块下三角矩阵的行列式等于它的主对角线上所有小块的行列式的乘积. 利用矩阵的转置从(10)式可得出

$$\begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ 0 & A_{22} & \cdots & A_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & A_{ss} \end{vmatrix} = |A_{11}| \cdot |A_{22}| \cdot \cdots \cdot |A_{ss}| \tag{11}$$

即分块上三角矩阵的行列式等于它的主对角线上所有小块的行列式的乘积.

习 题 4.6

1. 设

$$A=\begin{pmatrix} I_3 & A_{12} \\ 0 & 4I_2 \end{pmatrix}, \qquad B=\begin{pmatrix} B_{11} & B_{12} \\ 0 & I_2 \end{pmatrix}$$

其中

$$A_{12}=\begin{bmatrix} 2 & 5 \\ 3 & -2 \\ -1 & 6 \end{bmatrix}, \quad B_{11}=\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}, \quad B_{12}=\begin{bmatrix} a_3 & a_4 \\ b_3 & b_4 \\ c_3 & c_4 \end{bmatrix}$$

计算 AB.

2. 设 A 是 n 级方阵,证明:存在一个 $n\times m$ 非零矩阵 B,使 $AB=0$ 的充分必要条件为 $|A|=0$.

3. 设 B 为 m 级方阵,C 为一个 $m\times n$ 行满秩矩阵. 证明:

1) 如果 $BC=0$,则 $B=0$;

2) 如果 $BC=C$,则 $B=I$.

4. 设 A 与 B 是 n 级方阵. 证明:如果 $AB=0$,则

$$\operatorname{rank}(A)-\operatorname{rank}(B)\leqslant n$$

5. 设 A 是 n 级方阵,证明:对于任给的一个满足下述条件的正整数 k, $\operatorname{rank}(A)\leqslant k\leqslant n$,存在一个 n 级方阵 B,使得 $AB=0$,并且

$$\operatorname{rank}(A)+\operatorname{rank}(B)=k$$

6. 设 A 是 n 级方阵,证明:如果 $A^2=I$,则

$$\operatorname{rank}(A+I)+\operatorname{rank}(A-I)=n$$

7. 设 A 是 n 级方阵,证明:如果 $A^2=A$,则

$$\operatorname{rank}(A)+\operatorname{rank}(A-I)=n$$

8. 设 A 是一个 $s\times n$ 矩阵,B 是一个 $s\times m$ 矩阵. 证明:矩阵方程 $AX=B$ 有解的充分必要条件是 $\operatorname{rank}(A)=\operatorname{rank}((A,\ B))$,这里 $(A,\ B)$ 表示在 A 的右边添写上 B 所得到的矩阵.

9. 设 A 是实数域上的一个 $s\times n$ 矩阵,β 是 R^s 的一个列向量. 证明:n 元线性方程组

$$A'AX = A'\beta$$

一定有解.

10. 设 $A = \mathrm{diag}\{a_1I_{n_1}, a_2I_{n_2}, \cdots, a_sI_{n_s}\}$，其中 $a_1, a_2, \cdots, a_s$ 是两两不同的数，$n_1 + n_2 + \cdots + n_s = n$. 证明：与 A 可交换的矩阵一定是分块对角矩阵 $\mathrm{diag}\{B_1, B_2, \cdots, B_s\}$，其中 B_i 是 n_i 级方阵，$i = 1, \cdots, s$.

§7 分块矩阵的初等变换

上一节我们指出，分块上三角矩阵的行列式等于主对角线上所有小块的行列式的乘积. 譬如

$$\begin{vmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{vmatrix} = |A_{11}| \cdot |A_{22}|$$

设一个分块矩阵 A 为

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

我们希望较容易地计算 $|A|$，自然想到能不能象第二章把一个行列式通过矩阵的初等行变换化成上三角形行列式那样，把 A 经过某种变换化成分块上三角矩阵，即把 A 的左下角那块变成零矩阵. 为此，我们仿照矩阵的初等行变换引进分块矩阵的初等行变换的概念：

定义1 下述三种变换称为**分块矩阵的初等行变换**：

1° 把一个分块矩阵的一个块行的 P(矩阵) 倍(即，这个块行里的每一个小矩阵都左边乘上或都右边乘上一个矩阵 P) 加到另一个块行上；

2° 把一个分块矩阵里的两个块行互换位置；

3° 用一个行列式不为 0 的方阵左乘(或右乘) 一个分块矩阵的某一个块行.

例如

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \xrightarrow{②+P\cdot①} \begin{pmatrix} A_{11} & A_{12} \\ PA_{11}+A_{21} & PA_{12}+A_{22} \end{pmatrix}$$

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \xrightarrow{(①,②)} \begin{pmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{pmatrix}$$

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \xrightarrow{C\cdot①} \begin{pmatrix} CA_{11} & CA_{12} \\ A_{21} & A_{22} \end{pmatrix}, \text{其中 } |C| \neq 0$$

类似地，可以定义**分块矩阵的初等列变换**.

注意：分块矩阵的初等行（列）变换不是矩阵的初等行（列）变换.

在本章§4我们把矩阵的初等变换与矩阵的乘法联系起来. 现在我们想把分块矩阵的初等变换与分块矩阵的乘法联系起来. 为此先引进下述概念：

定义2 分块单位矩阵（即把单位矩阵分块得到的分块矩阵）经过一次分块矩阵的初等变换得到的矩阵称为**分块初等矩阵**.

以具有两个块行和两个块列的分块单位矩阵为例：

$$\begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \xrightarrow{②+P\cdot①} \begin{pmatrix} I_n & 0 \\ P & I_m \end{pmatrix} \tag{1}$$

$$\begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \xrightarrow{①+Q\cdot②} \begin{pmatrix} I_n & Q \\ 0 & I_m \end{pmatrix} \tag{2}$$

$$\begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \xrightarrow{(①,②)} \begin{pmatrix} 0 & I_m \\ I_n & 0 \end{pmatrix} \tag{3}$$

$$\begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \xrightarrow{①\cdot C} \begin{pmatrix} C & 0 \\ 0 & I_m \end{pmatrix}, \text{其中 } |C| \neq 0 \tag{4}$$

$$\begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \xrightarrow{②\cdot D} \begin{pmatrix} I_n & 0 \\ 0 & D \end{pmatrix}, \text{其中 } |D| \neq 0 \tag{5}$$

不难看出，作一次分块矩阵的初等列变换仍得到上述这些分

块矩阵. 这些都是分块初等矩阵.

注意:分块初等矩阵不是初等矩阵.

现在以具有两个块行和两个块列的分块矩阵为例来说明分块矩阵的初等变换与分块矩阵的乘法之间的联系. 设

$$A = \begin{array}{c c} & \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \end{array} & \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \end{array} \tag{6}$$

其中 A_{11} 是 $s_1 \times n_1$ 矩阵,A_{12} 是 $s_1 \times n_2$ 矩阵,等等. 我们有

$$\begin{pmatrix} I_{s_1} & 0 \\ P & I_{s_2} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ PA_{11} + A_{21} & PA_{12} + A_{22} \end{pmatrix}$$

$$\begin{pmatrix} 0 & I_{s_2} \\ I_{s_1} & 0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{pmatrix}$$

$$\begin{pmatrix} C & 0 \\ 0 & I_{s_2} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} CA_{11} & CA_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

这些说明,**对一个分块矩阵 A 作一次分块矩阵的初等行变换,就相当于在 A 的左边乘上一个相应的分块初等矩阵.**

类似地可证,**对一个分块矩阵 A 作一次分块矩阵的初等列变换,就相当于在 A 的右边乘上一个相应的分块初等矩阵.**

给了一个形如(6)的分块矩阵 A,只要适当选择矩阵 P,使 $PA_{11} + A_{21} = 0$(具体选法在本章 §8 和 §9 将介绍),就可以把 A 化成分块上三角矩阵,而做这一次分块矩阵的初等行变换就相当于在 A 的左边乘上一个相应的分块初等矩阵.

在下一节我们将运用分块矩阵的初等变换及其与分块矩阵的乘法的联系来讨论矩阵乘积的行列式.

chemi gram

习　题　4.7

1.　设 A 是一个形如(6)的分块矩阵，证明：对 A 作一次分块矩阵的初等列变换就相当于在 A 的右边乘上一个相应的分块初等矩阵.

2.　计算2°型分块初等矩阵(3)的行列式.

§8　矩阵乘积的行列式 ·Binet-Cauchy 公式

本节介绍矩阵乘积的行列式的公式，它们有很多应用.

定理4.8.1　设 A 与 B 都是数域 K 上的 n 级方阵，则

$$|AB| = |A||B|$$

证明　考虑下述分块矩阵

$$\begin{pmatrix} A & 0 \\ -I & B \end{pmatrix} \tag{1}$$

它的行列式等于 $|A||B|$. 为了与 $|AB|$ 联系起来，我们把分块矩阵(1)的第2块行的 A 倍加到第一块行上，而这相当于在分块矩阵(1)的左边乘上一个相应的分块初等矩阵，即

$$\begin{pmatrix} I & A \\ 0 & I \end{pmatrix}\begin{pmatrix} A & 0 \\ -I & B \end{pmatrix} = \begin{pmatrix} 0 & AB \\ -I & B \end{pmatrix} \tag{2}$$

为了计算(2)式右端的分块矩阵的行列式，运用 Laplace 定理，按前 n 行展开得

$$\begin{aligned} \begin{vmatrix} 0 & AB \\ -I & B \end{vmatrix} &= |AB|(-1)^{(1+\cdots+n)+[(n+1)+\cdots+2n]}|-I| \\ &= |AB|(-1)^{n^2}(-1)^n|I| = |AB| \end{aligned} \tag{3}$$

现在来计算(2)式左端的两个分块矩阵乘积的行列式. 设 $A = (a_{ij})$. 因为

$$\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \xrightarrow[\text{①}+(\overline{2n})a_{1n}]{\text{①}+(\overline{n+1})a_{11} \ \cdots} \begin{pmatrix} & a_{11} & a_{12} & \cdots & a_{1n} \\ & 0 & 0 & \cdots & 0 \\ I & \cdots & \cdots & \cdots & \cdots \\ & 0 & 0 & \cdots & 0 \\ 0 & & I & & \end{pmatrix}$$

$$\xrightarrow[\text{②}+(\overline{2n})a_{2n}]{\text{②}+(\overline{n+1})a_{21} \ \cdots} \begin{pmatrix} & a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{21} & a_{22} & \cdots & a_{2n} \\ I & 0 & 0 & \cdots & 0 \\ & \cdots & \cdots & \cdots & \cdots \\ & 0 & 0 & \cdots & 0 \\ 0 & & I & & \end{pmatrix}$$

$$\xrightarrow{1^\circ\text{型行变换}} \cdots \xrightarrow{1^\circ\text{型行变换}} \begin{pmatrix} I & A \\ 0 & I \end{pmatrix}$$

所以据初等行变换与初等矩阵的乘法的关系得

$$\begin{pmatrix} I & A \\ 0 & I \end{pmatrix} = P_t \cdots P_2 P_1 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} = P_t \cdots P_2 P_1 \tag{4}$$

其中 $P_1, P_2, \cdots, P_t$ 是与上述 1° 型初等行变换相应的初等矩阵，并且 $t = n^2$. 于是(2) 式左端为

$$\begin{pmatrix} I & A \\ 0 & I \end{pmatrix} \begin{pmatrix} A & 0 \\ -I & B \end{pmatrix} = P_t \cdots P_2 P_1 \begin{pmatrix} A & 0 \\ -I & B \end{pmatrix} \tag{5}$$

因为1°型初等行变换不改变矩阵的行列式的值，所以

$$\left| \begin{pmatrix} I & A \\ 0 & I \end{pmatrix} \begin{pmatrix} A & 0 \\ -I & B \end{pmatrix} \right| = \begin{vmatrix} A & 0 \\ -I & B \end{vmatrix} = |A|\,|B| \tag{6}$$

由(2)，(3)，(6)式得

$$|AB| = |A|\,|B| \qquad \blacksquare$$

用数学归纳法，定理4.8.1可以推广到多个因子的情形：

推论4.8.1　设 $A_1, A_2, \cdots, A_s$ 都是数域 K 上的 n 级方阵，则

$$|A_1 A_2 \cdots A_s| = |A_1|\,|A_2| \cdots |A_s| \qquad \blacksquare$$

定义1　数域 K 上的 n 级方阵 A 称为**非退化的**，如果

$|A| \neq 0$；否则称为**退化的**.

显然，一个 n 级方阵是非退化的当且仅当它是满秩矩阵.

推论4.8.2 设 A,B 是数域 K 上的 n 级方阵，则 AB 为退化的充分必要条件是 A,B 中至少有一个是退化的. ▌

例1 设 A,B 分别是 $n\times m$ 与 $m\times n$ 矩阵，证明

$$\begin{vmatrix} I_m & B \\ A & I_n \end{vmatrix} = |I_n - AB| \tag{7}$$

分析 (7) 式左端是一个分块矩阵的行列式，如果能把它变成分块上三角矩阵，其行列式就容易计算了. 为此需作分块矩阵的初等行变换：②＋$(-A)$·①. 而这个变换就相当于在所给分块矩阵的左边乘上一个相应的分块初等矩阵.

证明 我们有

$$\begin{pmatrix} I_m & 0 \\ -A & I_n \end{pmatrix}\begin{pmatrix} I_m & B \\ A & I_n \end{pmatrix} = \begin{pmatrix} I_m & B \\ 0 & I_n - AB \end{pmatrix} \tag{8}$$

(8) 式右端的矩阵的行列式为 $|I_n - AB||I_m| = |I_n - AB|$. 关于 (8) 式左端的矩阵的行列式，利用定理 4.8.1，它等于

$$\begin{vmatrix} I_m & 0 \\ -A & I_n \end{vmatrix}\begin{vmatrix} I_m & B \\ A & I_n \end{vmatrix} = \begin{vmatrix} I_m & B \\ A & I_n \end{vmatrix}$$

所以(7)式成立. ▌

下面考虑 A 与 B 都不是方阵，但是 AB 为方阵，这种情形 $|AB|$ 等于什么？

先引进一个记号. 从 A 中取出第 $i_1,i_2,\cdots,i_p$ 行，第 $j_1,j_2,\cdots,j_p$ 列组成的子矩阵 $A(i_1,i_2,\cdots,i_p;j_1,j_2,\cdots,j_p)$ 的行列式记成

$$A\begin{pmatrix} i_1, & i_2, & \cdots, & i_p \\ j_1, & j_2, & \cdots, & j_p \end{pmatrix}$$

定理4.8.2 (Binet－Cauchy 公式) 设 A 是 $s\times n$ 矩阵，B 是 $n\times s$ 矩阵. 如果 $s>n$，则 $|AB|=0$；如果 $s\leqslant n$，则 $|AB|$ 等于 A 的所有 s 级子式与 B 的相应 s 级子式的乘积之和，即

$$|AB| = \sum_{1\leqslant v_1<\cdots<v_s\leqslant n} A\begin{pmatrix}1, 2, \cdots, s\\ v_1, v_2, \cdots, v_s\end{pmatrix} B\begin{pmatrix}v_1, v_2, \cdots, v_s\\ 1, 2, \cdots, s\end{pmatrix} \tag{9}$$

* **证明** 如果 $s>n$，则 $\operatorname{rank}(AB)\leqslant\operatorname{rank}(A)\leqslant n<s$. 所以 AB 不是满秩矩阵. 从而 $|AB|=0$.

下面 设 $s\leqslant n$. 考虑分块矩阵

$$\begin{pmatrix} A & 0\\ -I_n & B\end{pmatrix} \tag{10}$$

把它的第2块行的 A 倍加到第1块行上，即

$$\begin{pmatrix} I_s & A\\ 0 & I_n\end{pmatrix}\begin{pmatrix} A & 0\\ -I_n & B\end{pmatrix} = \begin{pmatrix} 0 & AB\\ -I_n & B\end{pmatrix} \tag{11}$$

利用定理4.8.1得，(11) 式左端的矩阵的行列式等于分块矩阵(10) 的行列式. 为了计算(11) 式右端矩阵的行列式，按前 s 行展开得

$$\begin{aligned}\begin{vmatrix} 0 & AB\\ -I_n & B\end{vmatrix} &= |AB|(-1)^{(1+\cdots+s)+[(n+1)+\cdots+(n+s)]}|-I_n| \\ &= (-1)^{n(s+1)}|AB| \end{aligned}\tag{12}$$

我们 来计算矩阵(10) 的行列式，按前 s 行展开得

$$\begin{vmatrix} A & 0\\ -I_n & B\end{vmatrix} = \sum_{1\leqslant v_1<\cdots<v_s\leqslant n} A\begin{pmatrix}1, & 2, & \cdots, & s\\ v_1, & v_2, & \cdots, & v_s\end{pmatrix}$$

n×s

$$\cdot(-1)^{(1+2+\cdots+s)+(v_1+v_2+\cdots+v_s)}\cdot|(-\varepsilon_{\mu_1},\cdots,-\varepsilon_{\mu_{n-s}},B)| \tag{13}$$

其中 $\{\mu_1,\mu_2,\cdots,\mu_{n-s}\}=\{1,2,\cdots,n\}\setminus\{v_1,v_2,\cdots,v_s\}$. 把行列式 $|(-\varepsilon_{\mu_1},\cdots,-\varepsilon_{\mu_{n-s}},B)|$ 按前 $n-s$ 列展开，注意前 $n-s$ 列只有一个 $n-s$ 级子式不为零，它是取第 $\mu_1,\mu_2,\cdots,\mu_{n-s}$ 行得到的那个 $n-s$ 级子式，这个子式的代数余子式是

$$(-1)^{(\mu_1+\mu_2+\cdots+\mu_{n-s})+[1+2+\cdots+(n-s)]}\cdot B\begin{pmatrix}v_1, & v_2, & \cdots, & v_s\\ 1, & 2, & \cdots, & s\end{pmatrix}$$

因此

$$|(-\varepsilon_{\mu_1},\cdots,-\varepsilon_{\mu n-s},B)|$$

$(-1)^{n-s}$

$$=|-I_{n-s}|\cdot B\begin{pmatrix}v_1, & v_2, & \cdots, & v_s\\ 1, & 2, & \cdots, & s\end{pmatrix}$$

$$\cdot(-1)^{(\mu_1+\cdots+\mu_{n-s})+[1+2-\cdots+(n-s)]} \tag{14}$$

代入（13）式，（－1）的指数经过计算化简，

$$\begin{vmatrix}A & 0\\ -I_n & B\end{vmatrix}=(-1)^{s^2+n^2-s(n+1)}$$

$$\cdot\sum_{1\leqslant v_1<\cdots<v_s\leqslant n}A\begin{pmatrix}1, 2, \cdots, s\\ v_1, v_2, \cdots, v_s\end{pmatrix}B\begin{pmatrix}v_1, v_2, \cdots, v_s\\ 1, 2, \cdots, s\end{pmatrix} \tag{15}$$

从(11),(12),(15)式得

$$|AB|=(-1)^{n(s+1)}(-1)^{s^2+n^2-s(n+1)}$$

$$\cdot\sum_{1\leqslant v_1<\cdots<v_s\leqslant n}A\begin{pmatrix}1, 2, \cdots, s\\ v_1, v_2, \cdots, v_s\end{pmatrix}B\begin{pmatrix}v_1, v_2, \cdots, v_s\\ 1, 2, \cdots, s\end{pmatrix}$$

$$=\sum_{1\leqslant v_1<\cdots<v_s\leqslant n}A\begin{pmatrix}1, 2, \cdots, s\\ v_1, v_2, \cdots, v_s\end{pmatrix}B\begin{pmatrix}v_1, v_2, \cdots, v_s\\ 1, 2, \cdots, s\end{pmatrix}$$ ▌

例如

$$\left|\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}\begin{pmatrix}7 & 8\\ 9 & 10\\ 11 & 12\end{pmatrix}\right|$$

$$=\begin{vmatrix}1 & 2\\ 4 & 5\end{vmatrix}\begin{vmatrix}7 & 8\\ 9 & 10\end{vmatrix}+\begin{vmatrix}1 & 3\\ 4 & 6\end{vmatrix}\begin{vmatrix}7 & 8\\ 11 & 12\end{vmatrix}+\begin{vmatrix}2 & 3\\ 5 & 6\end{vmatrix}\begin{vmatrix}9 & 10\\ 11 & 12\end{vmatrix}$$

当 $s=n$ 时，Binet－Cauchy 公式与定理 4.8.1 一致.

Binet－Cauchy 公式有很多应用，下面举几个例子.

该公式用处之一是可以用来计算乘积矩阵的各级子式.

命题4.8.1　设 A 是 $s\times n$ 矩阵，B 是 $n\times s$ 矩阵，设正整数 $r\leqslant s$. 如果 $r>n$，则 AB 的所有 r 级子式等于 0；如果 $r\leqslant n$，则 AB 的任一 r 级子式为

$$AB\begin{pmatrix}i_1, \cdots, i_r\\ j_1, \cdots, j_r\end{pmatrix}$$

$$= \sum_{1 \leqslant v_1 < \cdots < v_r \leqslant n} A\begin{pmatrix} i_1, \cdots, i_r \\ v_1, \cdots, v_r \end{pmatrix} B\begin{pmatrix} v_1, \cdots, v_r \\ j_1, \cdots, j_r \end{pmatrix} \tag{16}$$

证明 我们有

$$\begin{aligned} & AB\begin{pmatrix} i_1, & \cdots, & i_r \\ j_1, & \cdots, & j_r \end{pmatrix} \\ = & \left| \begin{pmatrix} AB(i_1;j_1) & \cdots & AB(i_1;j_r) \\ \cdots & \cdots & \cdots \\ AB(i_r;j_1) & \cdots & AB(i_r;j_r) \end{pmatrix} \right| \\ = & \left| \begin{pmatrix} a_{i_1,1} & \cdots & a_{i_1,n} \\ \cdots & \cdots & \cdots \\ a_{i_r,1} & \cdots & a_{i_r,n} \end{pmatrix} \begin{pmatrix} b_{1j_1} & \vdots & b_{1j_r} \\ \vdots & \vdots & \vdots \\ b_{nj_1} & \vdots & b_{nj_r} \end{pmatrix} \right| \end{aligned} \tag{17}$$

据 Binet — Cauchy 公式，如果 $r > n$，则(17) 式右端两个矩阵乘积的行列式等于 0，从而(17) 式左端 AB 的 r 级子式等于 0. 如果 $r \leqslant n$，则(17) 式右端两个矩阵(记作 A_1 与 B_1) 的乘积的行列式等于

$$\begin{aligned} & \sum_{1 \leqslant v_1 < \cdots < v_r \leqslant n} A_1\begin{pmatrix} 1, & 2, & \cdots, r \\ v_1, v_2, \cdots, v_r \end{pmatrix} B_1\begin{pmatrix} v_1, v_2, \cdots, v_r \\ 1, & 2, & \cdots, r \end{pmatrix} \\ = & \sum_{1 \leqslant v_1 < \cdots < v_r \leqslant n} A\begin{pmatrix} i_1, i_2, \cdots, i_r \\ v_1, v_2, \cdots, v_r \end{pmatrix} B\begin{pmatrix} v_1, v_2, \cdots, v_r \\ j_1, j_2, \cdots, j_r \end{pmatrix} \end{aligned} \tag{18}$$

从而 (17) 式左端 AB 的 r 级子式等于上述(18) 式，这证明了(16) 式成立. ▌

定义2 矩阵 A 的一个子式如果行指标与列指标相同，则称为 A 的**主子式**.

例2 设 A 是实数域上的一个 $s \times n$ 矩阵，则 AA' 的所有主子式是非负的.

证明 任给 正整数$r \leqslant s$，据命题 4. 8. 1，如果 $r > n$，则 AA' 的所有 r 级子式都为 0. 如果 $r \leqslant n$，则 AA' 的任一 r 级主子式为

$$AA'\begin{pmatrix} i_1, & \cdots, & i_r \\ i_1, & \cdots, & i_r \end{pmatrix}$$

$$=\sum_{1\leqslant v_1<\cdots<v_r\leqslant n} A\begin{pmatrix} i_1, & \cdots, & i_r \\ v_1, & \cdots, & v_r \end{pmatrix} A'\begin{pmatrix} v_1, & \cdots, & v_r \\ i_1, & \cdots, & i_r \end{pmatrix}$$

$$=\sum_{1\leqslant v_1<\cdots<v_r\leqslant n}\left[A\begin{pmatrix} i_1, & \cdots, & i_r \\ v_1, & \cdots, & v_r \end{pmatrix}\right]^2 \geqslant 0 \tag{19}$$

上述第二个等号成立是因为

$$[A(i_1,\cdots,i_r;v_1,\cdots,v_r)]' = A'(v_1,\cdots,v_r;i_1,\cdots,i_r) \quad \blacksquare$$

利用例2容易给出§5命题4.5.1的另一个证法，留给读者思考.

Binet－Cauchy 公式的应用之二是证明一些恒等式.

例3 证明 Cauchy 恒等式：当 $n \geqslant 2$ 时，有

$$\left(\sum_{i=1}^{n} a_i c_i\right)\left(\sum_{i=1}^{n} b_i d_i\right) - \left(\sum_{i=1}^{n} a_i d_i\right)\left(\sum_{i=1}^{n} b_i c_i\right)$$

$$=\sum_{1\leqslant j<k\leqslant n}(a_j b_k - a_k b_j)(c_j d_k - c_k d_j) \tag{20}$$

证明 (20)式左端为

$$\begin{vmatrix} \sum\limits_{i=1}^{n} a_i c_i & \sum\limits_{i=1}^{n} a_i d_i \\ \sum\limits_{i=1}^{n} b_i c_i & \sum\limits_{i=1}^{n} b_i d_i \end{vmatrix} = \left|\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}\begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \\ \vdots & \vdots \\ c_n & d_n \end{pmatrix}\right|$$

$$=\sum_{1\leqslant j<k\leqslant n}\begin{vmatrix} a_j & a_k \\ b_j & b_k \end{vmatrix}\begin{vmatrix} c_j & d_j \\ c_k & d_k \end{vmatrix}$$

$$=\sum_{1\leqslant j<k\leqslant n}(a_j b_k - a_k b_j)(c_j d_k - c_k d_j)$$

这证明了(20)式成立. $\blacksquare$

解析几何中讲的 Lagrange 恒等式(参见本节习题第 12 题的注)就是 Cauchy 恒等式在 $n=3$ 的情形.

应用之三：Binet－Cauchy 公式在证明图论里的矩阵-树定理中起了关键作用，有兴趣的读者可参看 F. Harary 著的《图论》176 页至177页.

矩阵乘积的行列式公式与 Binet－Cauchy 公式有很多应用，这里不一一列举了.

习　题　4.8

1. 设 A 是 n 级方阵,则 $|AA'| = |A|^2$.

2. 若 A 是 n 级方阵,并且满足 $AA' = I$,则 $|A|$ 等于 1 或 -1.

3. 证明:若 A 是数域 K 上 n 级方阵,且满足 $AA' = I$ 与 $|A| = -1$,则 $|I + A| = 0$.

4. 证明:若 A 是数域 K 上 n 级方阵,n 是奇数,并且 $AA' = I$,$|A| = 1$,则 $|I - A| = 0$.

5. 设 A 是 s 级方阵,B 是 r 级方阵.证明:

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

非退化的充分必要条件是:A 与 B 都非退化.

6. 设 $s_k = x_1^k + x_2^k + \cdots + x_n^k$,　$k = 0,1,2,\cdots$;设 $A = (a_{ij})$,其中 $a_{ij} = s_{i+j-2}$,　$i,j = 1,2,\cdots,n$.证明:

$$|A| = \prod_{1 \leqslant j < i \leqslant n} (x_i - x_j)^2$$

7. 设 A,B 分别是 $n \times m$,$m \times n$ 矩阵,证明:

$$\begin{vmatrix} I_m & B \\ A & I_n \end{vmatrix} = |I_m - BA|$$

8. 设 A,B 分别是 $n \times m$,$m \times n$ 矩阵,证明:

$$|I_n - AB| = |I_m - BA| \tag{21}$$

9. 设 A,B 分别是 $n \times m$,$m \times n$ 矩阵,且 $\lambda \neq 0$,证明:

$$|\lambda I_n - AB| = \lambda^{n-m}|\lambda I_m - BA| \tag{22}$$

10. 设 A,B 分别是 n,m 级方阵,求

$$\begin{vmatrix} 0 & A \\ B & C \end{vmatrix}$$

11. 设 A 是2级方阵,l 是大于2的整数.证明:$A^l = 0$ 当且仅当 $A^2 = 0$.

12. 证明拉格朗日(Lagrange)恒等式：

$$\left(\sum_{i=1}^{n}a_i^2\right)\left(\sum_{i=1}^{n}b_i^2\right)-\left(\sum_{i=1}^{n}a_ib_i\right)^2$$

$$=\sum_{1\leqslant j<k\leqslant n}(a_jb_k-a_kb_j)^2 \tag{23}$$

(注：解析几何中也讲了 Lagrange 恒等式：

$$(\vec{a}\times\vec{b})\cdot(\vec{c}\times\vec{d})=\begin{vmatrix}\vec{a}\cdot\vec{c} & \vec{a}\cdot\vec{d}\\ \vec{b}\cdot\vec{c} & \vec{b}\cdot\vec{d}\end{vmatrix} \tag{24}$$

当 $\vec{c}=\vec{a},\vec{d}=\vec{b}$ 时，(24) 式即为(23) 式在 $n=3$ 的情形；当 $\vec{a},\vec{b},\vec{c},\vec{d}$ 任意时，(24) 式即为 Cauchy 恒等式(20) 在 $n=3$ 的情形.)

13. 证明 Cauchy — Bunyakovsky 不等式：对任意实数 $a_1,a_2,\cdots,a_n$ 与 $b_1,b_2,\cdots,b_n$，有

$$(a_1^2+a_2^2+\cdots+a_n^2)(b_1^2+b_2^2+\cdots+b_n^2)$$

$$\geqslant(a_1b_1+a_2b_2+\cdots+a_nb_n)^2 \tag{25}$$

并且 等号成立当且仅当$(a_1,a_2,\cdots,a_n)$ 与$(b_1,b_2,\cdots,b_n)$ 线性相关.

14. 设 A 是 $s\times n$ 矩阵，$1\leqslant r\leqslant s$，且 $r\leqslant n$. 证明：AA' 的 r 级主子式之和等于 A 的所有 r 级子式的平方和.

*15. 设 A,B 都是 n 级方阵，证明：AB 与 BA 的 r 级$(1\leqslant r\leqslant n)$ 的所有主子式之和是相等的.

*16. 设 A 是一个 $n\times m$ 矩阵，其中 $m\geqslant n-1$，求 AA' 的$(1,1)$ 元的代数余子式.

*17. 设 A 是一个 $n\times m$ 矩阵，其中 $m\geqslant n-1$，并且 A 的每一列元素的和(简称为列和) 都为 0. 证明：AA' 的所有元素的代数余子式的值都相等.

(注：图论中的矩阵 - 树定理：设 G 为一个连通的标定图，其关联矩阵为 B. 把 B 的每一列中两个 1 的任何一个改为 -1，得到一个新的矩阵 E. 则 EE' 的所有元素的代数余子式的值都相等，并且这个公共值就是 G 的生成树的数目.)

*18. 形如

$$\begin{pmatrix}a_0 & a_1 & a_2 & \cdots & a_{n-1}\\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2}\\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ a_1 & a_2 & a_3 & \cdots & a_0\end{pmatrix}$$

的 n 级方阵 A 称为**循环矩阵**. 求复数域上循环矩阵 A 的行列式.

（提示：取一个 n 次本原单位根 ω，即 ω 是一个复数，它满足 $\omega^n=1$，但是 $\omega^i\neq 1$，对于 $1\leqslant i\leqslant n-1$. 令

$$B=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2} \end{pmatrix}$$

令 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1}$. 去验证

$$AB=B\mathrm{diag}\{f(1),f(\omega),f(\omega^2),\cdots,f(\omega^{n-1})\}$$

上式两端取行列式得 $|A||B|=|B|\prod\limits_{i=0}^{n-1}f(\omega^i)$.（注意 B 是 Vandermonde 行列式，并且 $1,\omega,\omega^2,\cdots,\omega^{n-1}$ 两两不同，因此 $|B|\neq 0$. 从而得 $|A|=\prod\limits_{i=0}^{n-1}f(\omega^i)$.）

19. 用 Binet—Cauchy 公式计算下述行列式

$$\begin{vmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \cdots & \cdots & \cdots & \cdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{vmatrix}$$

*20. 设

$$A=\begin{array}{c} \begin{array}{cc} \overbrace{\qquad\qquad\qquad}^{n-1} & \overbrace{\qquad\qquad\qquad\qquad\qquad}^{n} \end{array} \\ \begin{pmatrix} 1 & 0 & \cdots & 0 & a & b & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & a & b & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & a & b \\ n & 0 & \cdots & 0 & a & 0 & 0 & \cdots & 0 & 0 \\ 0 & n & \cdots & 0 & 0 & a & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & n & 0 & 0 & 0 & \cdots & a & 0 \\ 0 & 0 & \cdots & 0 & n & 0 & 0 & \cdots & 0 & a \end{pmatrix} \end{array} \begin{array}{l} \\ \left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\}n-1 \\ \left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\end{array}}\right\}n \end{array}$$

求 $|A|$.

§ 9 可逆矩阵·求逆矩阵的方法

9.1 可逆矩阵的定义、判定与性质

数的乘法有逆运算，矩阵的乘法有没有逆运算呢?数的乘法之所以有逆运算，关键是每个非零数 a 有“逆”a^{-1}，即$\frac{1}{a}$，于是 $b\div a=b\cdot\frac{1}{a}$. 非零数 a 的逆 a^{-1} 满足 $aa^{-1}=a^{-1}a=1$. 因此对于矩阵，自然要问：什么样的矩阵有类似于非零数那样的性质?

定义1 对于数域 K 上的一个矩阵 A，如果有数域 K 上的一个矩阵 B，使得

$$AB=BA=I \tag{1}$$

则称 A 为**可逆的**(或**非奇异的**).

从(1)式看出，A 与 B 可交换，因此可逆矩阵 A 一定是方阵. 换句话说，若一个矩阵不是方阵，则它一定不可逆. 适合(1)式的矩阵 B 也是方阵.

其次，适合(1)式的矩阵 B(如果有的话)是唯一的. 这是因为假如 B_1 与 B_2 都适合(1)式，则有

$$B_1AB_2=(B_1A)B_2=IB_2=B_2$$
$$B_1AB_2=B_1(AB_2)=B_1I=B_1$$

从而 $B_2=B_1$.

定义2 如果矩阵 B 适合(1)式，则 B 称为 A 的**逆矩阵**，记作 A^{-1}.

于是若 A 是可逆矩阵，则有逆矩阵 A^{-1}，使得下式成立

$$AA^{-1}=A^{-1}A=I \tag{2}$$

从(2)式可看出，A^{-1} 也是可逆的，并且有

$$(A^{-1})^{-1}=A \tag{3}$$

什么样的矩阵是可逆的？如果 A 可逆，如何求 A^{-1}？

首先看方阵 A 可逆的必要条件是什么？

命题4.9.1 若方阵 A 可逆，则 $|A| \neq 0$.

证明 因为 A 可逆，所以 $AA^{-1} = I$. 从而 $|AA^{-1}| = |I|$，由此得 $|A||A^{-1}| = 1$. 因此 $|A| \neq 0$. ▌

上述证明过程表明，若 A 可逆，则 $|A^{-1}| = |A|^{-1}$.

上述必要条件是否为充分条件？回答是肯定的. 为了证明这一点，需要引进一个概念.

定义3 设 $A = (a_{ij})$ 是 n 级方阵，用 A_{ij} 表示 A 的 (i,j) 元的代数余子式 $(i,j = 1,\cdots,n)$. 矩阵

$$\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \tag{4}$$

称为 A 的伴随矩阵，记作 $\boxed{A^*}$. (注：本书中，记号 A^* 有两种用法，一种是表示方阵的伴随矩阵(4)，另一种是用 A^* 表示复数域上矩阵 A 的转置共轭矩阵 $\overline{A}'$. 这可从上下文分辨出.)

利用行列式按一行展开的公式可以得出

$$AA^* = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & |A| \end{pmatrix} = |A|I \tag{5}$$

类似地，利用行列式按一列展开的公式可得出

$$A^*A = |A|I \tag{6}$$

于是我们可以得到

定理4.9.1　矩阵 A 可逆的充分必要条件为 $|A|\neq 0$(即 A 非退化),并且当 A 可逆时,有

$$A^{-1}=\frac{1}{|A|}A^{*} \tag{7}$$

证明　必要性已证.现在证充分性.若 $|A|\neq 0$,则由(5)与(6)式得

$$A\left(\frac{1}{|A|}A^{*}\right)=\left(\frac{1}{|A|}A^{*}\right)A=I$$

因此 A 可逆,并且 $A^{-1}=|A|^{-1}A^{*}$.　▌

定理4.9.1不仅给出了判断一个矩阵是否可逆的一种方法,并且给出了求逆矩阵的一种方法,称之为伴随矩阵法.当然,这种方法主要用在理论上以及2级或3级矩阵的情形,如果级数较大,那么应采用下面将介绍的其他方法,否则计算量太大.

例1　设

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

问:当 a,b,c,d 满足什么条件时,A 可逆?当 A 可逆时,求 A^{-1}.

解　A 可逆 $\Leftrightarrow |A|\neq 0 \Leftrightarrow ad-bc\neq 0$.

设 $ad-bc\neq 0$,则

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

推论4.9.1　n 级矩阵 A 可逆的充分必要条件是 $\mathrm{rank}(A)=n$(即 A 为满秩矩阵).　▌

推论4.9.2　n 级矩阵 A 可逆的充分必要条件是它的行(或列)向量组线性无关.　▌

推论4.9.3　数域 K 上 n 级矩阵 A 可逆的充分必要条件是 A 的行(或列)向量组是 K^{n} 的一个基.　▌

推论4.9.4　数域 K 上 n 级矩阵 A 可逆的充分必要条件是 A 的行空间(或列空间)等于 K^{n}.　▌

下面给出判别一个矩阵是否可逆的更简便的方法.

命题4.9.2 设 A 与 B 都是 n 级矩阵. 若 $AB=I$,则 A 与 B 都可逆,并且 $A^{-1}=B, B^{-1}=A$.

证明 因为 $AB=I$,所以 $|A||B|=1$. 从而 $|A|\neq 0, |B|\neq 0$. 因此 A 与 B 都可逆. 在 $AB=I$ 两边左乘 A^{-1},得 $A^{-1}(AB)=A^{-1}I$,由此得 $B=A^{-1}$. 于是 $B^{-1}=(A^{-1})^{-1}=A$. ▌

命题4.9.3 若 n 级矩阵 A,B 都可逆,则 AB 也可逆,并且

$$\boxed{(AB)^{-1}=B^{-1}A^{-1}} \tag{8}$$

证明 因为 A 与 B 都可逆,所以有

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

于是 AB 可逆,并且 $(AB)^{-1}=B^{-1}A^{-1}$. ▌

命题4.9.3 可以推广到多个 n 级可逆矩阵相乘的情形,即若 n 级矩阵 $A_1,A_2,\cdots,A_s$ 都可逆,则 $A_1A_2\cdots A_s$ 也可逆,并且有

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

命题4.9.4 若 A 可逆,则 A' 也可逆,并且

$$(A')^{-1}=(A^{-1})'$$

证明 因为 $A'(A^{-1})'=(A^{-1}A)'=I'=I$,所以 A' 可逆,并且 $(A')^{-1}=(A^{-1})'$. ▌

命题4.9.5 如果在一个矩阵 A 的左边(或右边)乘上一个可逆矩阵,则乘积矩阵的秩等于 A 的秩.

证明 设 A 是 $s\times n$ 矩阵,P 是 s 级可逆矩阵. 则

$$\operatorname{rank}(PA)\leqslant \operatorname{rank}(A) \tag{9}$$

设 $PA=B$,则 $A=P^{-1}B$. 于是有

$$\operatorname{rank}(A)=\operatorname{rank}(P^{-1}B)\leqslant \operatorname{rank}(B)=\operatorname{rank}(PA) \tag{10}$$

从(9)和(10)式得,

$$\operatorname{rank}(PA)=\operatorname{rank}(A)$$

若 Q 是 n 级可逆矩阵,则

$$\begin{aligned}\operatorname{rank}(AQ)&=\operatorname{rank}((AQ)')=\operatorname{rank}(Q'A')\\&=\operatorname{rank}(A')=\operatorname{rank}(A)\end{aligned}$$ ▌

* 利用定理4.9.1可以给出 Cramer 法则另一个证法：

证明 对于 n 个方程组成的 n 元线性方程组

$$AX=\beta \tag{11}$$

如果 $|A|\neq 0$，则 A 可逆. 于是我们有

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

这表明 $A^{-1}\beta$ 是线性方程组(11)的一个解. 把(7)式代入得

$$A^{-1}\beta=\frac{1}{|A|}A^*\beta=\frac{1}{|A|}\begin{pmatrix}A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \cdots & \cdots & \cdots & \cdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{pmatrix}\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}$$

$$=\frac{1}{|A|}(|B_1|,|B_2|,\cdots,|B_n|)$$

其中

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

如果 α 是方程组(11)的一个解，则 $A\alpha=\beta$. 两边左乘 A^{-1} 得 $\alpha=A^{-1}\beta$. 这证明了方程组(11)的解是唯一的. ∎

9.2 求逆矩阵的方法

方法一 伴随矩阵法，即若 $|A|\neq 0$，则 A 可逆，且

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

这方法在理论上很有用，在实际计算中常用于2级矩阵.

方法二 用命题4.9.2，即找一个矩阵 B，使 $AB=I$，则 A 可逆，并且 $A^{-1}=B$.

例2 设

$$B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix}$$

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

证明：$A = (I + B)$ 可逆，并且求 A^{-1}.

证明 因为 I 与 B 可交换，所以

$$(I + B)(I - B + B^2 - B^3) = I - B^4$$

易计算 $B^4 = 0$（参见习题 4.3 的第 5 题中第(7) 小题），于是得

$$(I + B)(I - B + B^2 - B^3) = I$$

据命题4.9.2得，$I + B$ 可逆，并且

$$(I + B)^{-1} = I - B + B^2 - B^3$$ ▌

例3 证明：初等矩阵都可逆，并且每个初等矩阵的逆矩阵是与它同型的初等矩阵.

证明 从初等矩阵与初等变换的联系可看出

$$P(i,j(-k))P(i,j(k)) = I$$

$$P(i,j)P(i,j) = I$$

$$P(i(c^{-1}))P(i(c)) = I$$

因此，$P(i,j(k))$，$P(i,j)$，$P(i(c))$ 都可逆，并且

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

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

$$P(i(c))^{-1} = P(i(c^{-1}))$$

这表明每个初等矩阵的逆矩阵是与它同型的初等矩阵. ▌

方法三 初等变换法，这是最常用的一种方法. 为了看出如何用初等变换求逆矩阵，先证一个引理：

引理4.9.1 可逆矩阵的简化行阶梯形一定是单位矩阵. 换句话说，可逆矩阵可以经过一系列初等行变换化成单位矩阵.

证明 设 A 是 n 级可逆矩阵，它的简化行阶梯形矩阵 J 的非零行的数目等于 $\mathrm{rank}(A) = n$. 于是 J 有 n 个主元，设它们的列指

标分别为 $j_1, j_2, \cdots, j_n$，则

$$1 \leqslant j_1 < j_2 < \cdots < j_n \leqslant n$$

由此得，$j_1 = 1, j_2 = 2, \cdots, j_n = n$. 因此 $J = I$. ▌

利用初等行变换与初等矩阵的乘法的联系，从引理 4.9.1 得出，有一系列初等矩阵 $P_1, P_2, \cdots, P_t$，使得

$$P_t \cdots P_2 P_1 A = I \tag{12}$$

在(12)式两边右乘 A^{-1} 得

$$P_t \cdots P_2 P_1 I = A^{-1} \tag{13}$$

比较(12)与(13)式得出，**如果用一系列初等行变换把 A 化成了单位矩阵，那么同样的这些初等行变换就把单位矩阵化成了 A^{-1}.**

因此，我们可以把 A 与 I 并排放在一起组成一个 $n \times 2n$ 矩阵 (A, I). **对矩阵 (A, I) 作一系列初等行变换，把它的左半部分化成 I，这时右半部分就是 A^{-1}.** 即

$$(A, I) \xrightarrow{\text{初等行变换}} (I, A^{-1})$$

例4 设

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

求 A^{-1}.

解

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

$$\xrightarrow{①+②\cdot(-1)} \begin{pmatrix} 1 & -1 & 1 & 1 & -1 & 0 \\ 3 & 2 & 1 & 0 & 1 & 0 \\ 5 & -3 & 2 & 0 & 0 & 1 \end{pmatrix}$$

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

$$\xrightarrow{②+①\cdot(-2)}\begin{pmatrix}1 & -1 & 1 & 1 & -1 & 0\\ 0 & 1 & 4 & 7 & -6 & -2\\ 0 & 2 & -3 & -5 & 5 & 1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & -1 & 1 & 1 & -1 & 0\\ 0 & 1 & 4 & 7 & -6 & -2\\ 0 & 0 & 1 & \frac{19}{11} & -\frac{17}{11} & -\frac{5}{11}\end{pmatrix}$$

−11 −19 17 5

$$\longrightarrow\begin{pmatrix}1 & 0 & 0 & -\frac{7}{11} & \frac{8}{11} & \frac{3}{11}\\ 0 & 1 & 0 & \frac{1}{11} & \frac{2}{11} & -\frac{2}{11}\\ 0 & 0 & 1 & \frac{19}{11} & -\frac{17}{11} & -\frac{5}{11}\end{pmatrix}$$

所以

$$A^{-1}=\begin{pmatrix}-\frac{7}{11} & \frac{8}{11} & \frac{3}{11}\\ \frac{1}{11} & \frac{2}{11} & -\frac{2}{11}\\ \frac{19}{11} & -\frac{17}{11} & -\frac{5}{11}\end{pmatrix}$$

从公式(12)我们可以得到可逆矩阵的又一个性质：

命题4.9.6 一个矩阵可逆的充分必要条件是它可以表示成一些初等矩阵的乘积.

证明 因为初等矩阵都可逆，所以它们的乘积仍可逆. 这证明了充分性. 现在证必要性. 设 n 级矩阵 A 可逆，从公式(12)得

$$A=(P_t\cdots P_2P_1)^{-1}=P_1^{-1}P_2^{-1}\cdots P_t^{-1} \tag{14}$$

由于初等矩阵的逆矩阵仍是初等矩阵，因此 A 是初等矩阵的乘积. ▌

方法四 利用解线性方程组来求逆矩阵. 若 n 级矩阵 A 可逆，则 $AA^{-1}=I$. 于是 A^{-1} 的第 j 列是线性方程组 $AX=\varepsilon_j$ 的解，$j=$

$1,2,\cdots,n$. 因此我们可以去解线性方程组 $AX=\beta$,其中 $\beta=(b_1,\cdots,b_n)'$,然后把所得的解的公式中的 $b_1,b_2,\cdots,b_n$ 分别用 $1,0,\cdots,0;0,1,0,\cdots,0;\cdots;0,\cdots,0,1$ 代替,便可求得 A^{-1} 的第 $1,2,\cdots,n$ 列. 这种方法在某些时候可能比用初等变换法求逆矩阵稍微简单一点.

方法五 分块求逆法.

当一个可逆矩阵的级数较大时,即使用初等变换法求它的逆矩阵仍然计算量较大. 如果把该矩阵分块,再对分块矩阵求逆矩阵,则有可能减少计算量.

如果把一个矩阵适当分块后能得到一个分块对角矩阵,那么很容易判断它是否可逆,并且当可逆时较易求它的逆矩阵.

例5 证明:分块对角矩阵 $A=\mathrm{diag}\{A_1,A_2,\cdots,A_s\}$ 可逆的充分必要条件是它的主对角线上每个小矩阵 A_i 可逆($i=1,\cdots,s$);并且 A 可逆时,

$$A^{-1}=\mathrm{diag}\{A_1^{-1},A_2^{-1},\cdots,A_s^{-1}\}$$ $|A|\neq 0 \Rightarrow |A_i|\neq 0$

证明 因为 $|A|=|A_1||A_2|\cdots|A_s|$,所以 A 可逆 $\Leftrightarrow$ 每个 A_i ($i=1,\cdots,s$) 可逆. 因为 A 可逆时,有

$$A\mathrm{diag}\{A_1^{-1},A_2^{-1},\cdots,A_s^{-1}\}=\mathrm{diag}\{A_1A_1^{-1},A_2A_2^{-1},\cdots,A_sA_s^{-1}\}=I$$

所以 $A^{-1}=\mathrm{diag}\{A_1^{-1},A_2^{-1},\cdots,A_s^{-1}\}$. ∎

例5表明,可逆的分块对角矩阵的逆矩阵仍是分块对角矩阵.

现在来看分块上三角形矩阵可逆时,怎样求逆?

例6 设

$$A=\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{pmatrix}$$

其中 A_{11},A_{22} 分别是 r 级、s 级方阵. 证明:A 可逆当且仅当 A_{11} 与 A_{22} 都可逆;此时求 A^{-1}.

证明 因为 $|A|=|A_{11}||A_{22}|$,所以 $|A|\neq 0 \Leftrightarrow |A_{11}|\neq 0$ 并且 $|A_{22}|\neq 0$. 于是 A 可逆 $\Leftrightarrow A_{11}$ 与 A_{22} 都可逆.

A 可逆时,怎样求 A^{-1}?如果能把 A 化成分块对角矩阵,那么

求逆问题就变得容易解决了. 为此需要把右上角变成零矩阵(形象地说,在右上角打出一个洞). 办法是把 A 的第 2 块行的 $-A_{12}A_{22}^{-1}$ 倍加到第 1 块行上,则右上角变成 0. 这相当于在 A 的左边乘上一个相应的分块初等矩阵,因此我们有

$$\begin{pmatrix} I_r & -A_{12}A_{22}^{-1} \\ 0 & I_s \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} \tag{15}$$

从(15)式得　　$AB = C$　　$B = \frac{C}{A}$　　$B^{-1} = \frac{A}{C}$

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}^{-1}\begin{pmatrix} I_r & -A_{12}A_{22}^{-1} \\ 0 & I_s \end{pmatrix} = C^{-1}A$$

$$= \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{pmatrix} \tag{16}$$

■

例6说明:可逆的分块上三角矩阵的逆矩阵仍然是分块上三角矩阵.

现在再看一般的分块矩阵可逆时,怎样求逆?

例7 设

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

其中 A_{11} 是 r 级可逆矩阵,A_{22} 是 s 级方阵. 问:A 何时可逆?当 A 可逆时,求 A^{-1}.

解 如果能把 A 变成分块上三角矩阵,则问题易于解决. 为此我们把 A 的第 1 块行的 $-A_{21}A_{11}^{-1}$ 倍加到第 2 块行上,这相当于在 A 的左边乘上一个相应的分块初等矩阵:

$$\begin{pmatrix} I_r & 0 \\ -A_{21}A_{11}^{-1} & I_s \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix} \tag{17}$$

从(17)式看出,当 $D = A_{22} - A_{21}A_{11}^{-1}A_{12}$ 可逆时,A 可逆;并且

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & D \end{pmatrix}^{-1} \begin{pmatrix} I_r & 0 \\ -A_{21}A_{11}^{-1} & I_s \end{pmatrix}$$

$$= \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}D^{-1} \\ 0 & D^{-1} \end{pmatrix} \begin{pmatrix} I_r & 0 \\ -A_{21}A_{11}^{-1} & I_s \end{pmatrix}$$

$$= \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}D^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}D^{-1} \\ -D^{-1}A_{21}A_{11}^{-1} & D^{-1} \end{pmatrix}$$

9.3 矩阵方程

含有未知矩阵的等式称为**矩阵方程**.

线性 方程组 $AX = \beta$ 是最简单的一类矩阵方程.

本小节讨论两种简单的矩阵方程的解法.

(一) $AX = B$,其中 A,B 分别是 $s \times n, s \times m$ 矩阵.

易看出(见习题4.6的第8题),$AX = B$ 有解的充分必要条件是 $\mathrm{rank}(A) = \mathrm{rank}((A,B))$. 此时,我们来求解. 设 B 的列向量组是 $\beta_1, \cdots, \beta_m$,则 (A,B)表示在A的右边添写上B所得到的矩阵.

$n \times m$ 矩阵 C 是 $AX = B$ 的解 $\Leftrightarrow AC = B$

$\Leftrightarrow C$ 的第 j 列 γ_j 是线性方程组 $AY = \beta_j$ 的解,$j = 1, \cdots, m$

由于 解线性方程组 $AY = \beta_j$ 时,当我们把增广矩阵 (A, β_j) 经过初等行变换化成简化行阶梯形 $\tilde{J}$,则去掉 $\tilde{J}$ 的最后一列得到的矩阵 J 就是系数矩阵 A 的简化行阶梯形 J,反之也亦然;并且从 $\tilde{J}$ 易直接写出方程组 $AY = \beta_j$ 的解. 基于这个道理,我们可同时来解 m 个线性方程组 $AY = \beta_j, j = 1, \cdots, m$,即

$$(A,B) \xrightarrow{\text{一系列初等行变换}} (J,D)$$

然后 从 (J,D) 可以写出每个线性方程组 $AY = \beta_j$ 的一般解公式,进而可写出 $AX = B$ 的解.

特别地,如果 A 是 n 级可逆矩阵,则 $AX = B$ 有唯一解:$X = A^{-1}B$. 这时上述方法就成为

$$(A,B) \xrightarrow{\text{一系列初等行变换}} (I, A^{-1}B)$$

即当我们把(A,B)经过初等行变换使其左半部分变成 I 时，右半部分就是 $AX = B$ 的解 $A^{-1}B$.

（二） $XA = B$.

两边取转置得 $A'X' = B'$，归结为上一情形. 如果 A 可逆，则 $X = BA^{-1}$.

习 题 4.9

1. 写出数量矩阵 kI 可逆的充分必要条件，并且当 kI 可逆时，写出 $(kI)^{-1}$.

2. 证明：若矩阵 A 可逆，则 A 的伴随矩阵 A^* 也可逆，并求出$(A^*)^{-1}$.

3. 设

$$B = \begin{pmatrix} 0 & B_1 \\ B_2 & 0 \end{pmatrix}$$

其中 B_1, B_2 分别是 r 级，s 级方阵. 求 B 可逆的充分必要条件；当 B 可逆时，求 B^{-1}.

4. 求下述矩阵的逆矩阵：

$$A = \begin{pmatrix} 0 & a_1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & a_2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & a_{n-1} \\ a_n & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

其中 $a_1, a_2, \cdots, a_n$ 全不为零.

5. 证明：如果 $A^3 = 0$，则 $I - A$ 可逆；求$(I - A)^{-1}$.

6. 写出 n 级矩阵 kA 可逆的充分必要条件；并且当 kA 可逆时，写出 $(kA)^{-1}$.

7. 证明：可逆的对称（斜对称）矩阵的逆矩阵仍是对称（斜对称）矩阵.

8. 证明：如果 $A^k = 0$，则 $I - A$ 可逆；求$(I - A)^{-1}$.

9. 设 A 是 n 级矩阵($n \geqslant 2$)，证明：

$$|A^*| = |A|^{n-1}$$

10. 设 A 是 n 级矩阵($n \geqslant 2$)，证明：

$$\operatorname{rank}(A^*) = \begin{cases} n, & 当 \operatorname{rank}(A) = n \\ 1, & 当 \operatorname{rank}(A) = n-1 \\ 0, & 当 \operatorname{rank}(A) < n-1 \end{cases}$$

11. 设 A 是 n 级矩阵($n \geqslant 2$)，证明：

$$(A^*)^* = |A|^{n-2}A$$

*12. 设 A,B 都是 n 级矩阵($n \geqslant 2$)，证明：

$$(AB)^* = B^*A^*$$

13. 设 A,B,C,D 都是 n 级矩阵，并且 $|A| \neq 0, AC = CA$，证明

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD - CB|$$

14. 求 A^{-1}，设

(1)

$$A = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & 2 \\ 2 & -2 & 1 \end{pmatrix} \begin{pmatrix} \frac{2}{9} & \frac{1}{9} & \frac{2}{9} \\ \frac{1}{9} & \frac{2}{9} & -\frac{2}{9} \\ -\frac{2}{9} & \frac{2}{9} & \frac{1}{9} \end{pmatrix}$$

(2)

$$A = \begin{pmatrix} 2 & 5 & 7 \\ 5 & -2 & -3 \\ 6 & 3 & 4 \end{pmatrix}$$

(3)

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

(4)

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

(5)

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

(6)

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

(7)

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ 2 & 4 & 1 & 7 \\ -1 & -3 & -1 & -5 \end{pmatrix}$$

(8)

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

分块矩阵

(9)

$$A=\begin{pmatrix}3&1&0&0&0\\0&3&1&0&0\\0&0&3&1&0\\0&0&0&3&1\\0&0&0&0&3\end{pmatrix}$$

15. 求下列 n 级矩阵的逆矩阵：

(1) $\begin{pmatrix}0&1&1&\cdots&1&1\\1&0&1&\cdots&1&1\\1&1&0&\cdots&1&1\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\1&1&1&\cdots&1&0\end{pmatrix}$

* (2) $\begin{pmatrix}1&2&3&\cdots&n-1&n\\n&1&2&\cdots&n-2&n-1\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\2&3&4&\cdots&n&1\end{pmatrix}$

(3) $\begin{pmatrix}1&1&1&\cdots&1&1\\1&0&1&\cdots&1&1\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\1&1&1&\cdots&1&0\end{pmatrix}$

* (4) $\begin{pmatrix}1+a&1&1&\cdots&1\\1&1+a&1&\cdots&1\\\cdots&\cdots&\cdots&\cdots&\cdots\\1&1&1&\cdots&1+a\end{pmatrix}$， $a\neq 0$

(5) $\begin{pmatrix}1&1&1&\cdots&1&1\\0&1&1&\cdots&1&1\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&0&1\end{pmatrix}$

$$*(6)\quad \begin{pmatrix} 1 & b & b^2 & \cdots & b^{n-2} & b^{n-1} \\ 0 & 1 & b & \cdots & b^{n-3} & b^{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & b \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

$$*(7)\quad \begin{pmatrix} 1+a_1 & 1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1+a_n \end{pmatrix}, \quad a_1a_2\cdots,a_n \neq 0$$

$$*(8)\quad \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

$$*(9)\quad \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n-1 & n \\ 0 & 1 & 2 & 3 & \cdots & n-2 & n-1 \\ 0 & 0 & 1 & 2 & \cdots & n-3 & n-2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

*16. 设 A,B 分别是 $n\times m, m\times n$ 矩阵. 证明:如果 $I_n - AB$ 可逆,则 $I_m - BA$ 也可逆;并且求 $(I_m - BA)^{-1}$.

17. 求矩阵 X,设

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

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

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

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

* 18. 用分块求逆法求 A^{-1},设

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

19. 证明:可逆的上(下)三角矩阵的逆仍是上(下)三角矩阵.

* 20. 设 A 是一个 n 级矩阵,并且 $\operatorname{rank}(A) = r$. 证明:存在一个 n 级可逆矩阵 P,使 PAP^{-1} 的后 $n-r$ 行全为零.

21. 设 A 是 $m \times r$ 矩阵,证明

(1) A 是列满秩的充分必要条件为存在 m 级可逆矩阵 P,使

$$A = P\begin{pmatrix} I_r \\ 0 \end{pmatrix}$$

(2) A 行满秩的充分必要条件为存在 r 级可逆矩阵 Q,使

$$A = (I_m, 0)Q$$

* 22.

(1) 把矩阵

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$$

表示 成 1° 型初等矩阵 $P(i, j(k))$ 的乘积;

(2) 设

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

是复数域上一个矩阵，且 $|A|=1$. 证明：A 可以表示成 1° 型初等矩阵 $P(i,j(k))$ 的乘积.

*23. 设 A 是 n 级复矩阵，且 $|A|=1$. 证明：A 可以表示成 1° 型初等矩阵 $P(i,j(k))$ 的乘积.

*24. 证明：如果 n 级可逆矩阵 A 的每一行(或列)元素的和都是 a，则 $a\neq 0$，并且 A^{-1} 的每一行(或列)元素的和都是 a^{-1}.

(提示：用 1_n 表示 K^n 中分量全为 1 的列向量，则从 A 的每一行元素的和都是 a 可以得出 $A1_n=\begin{pmatrix}a\\ \vdots\\ a\end{pmatrix}=a1_n$.)

*25. 证明：任何方阵都可以表示成一些上三角形矩阵与下三角形矩阵的乘积.

*26. 证明：如果 n 级矩阵 A 的所有顺序主子式(即主子式 $A\begin{pmatrix}1, & 2, & \cdots, & k\\ 1, & 2, & \cdots, & k\end{pmatrix}$，$k=1,2,\cdots,n$) 不为 0，则有 n 级下三角矩阵 B，使 BA 为上三角矩阵.

§10 正交矩阵·R^n 的标准正交基

从§9知道，数域 K 上 n 级可逆矩阵 A 的行(或列)向量组是 K^n 的一个基. 例如，实数域 R 上的 3 级可逆矩阵 A 的行(或列)向量组是 R^3 的一个基. 我们还知道，几何空间 R^3 中有一种基，它的三个基向量是两两垂直的单位向量，这种基称为 R^3 的标准正交基. 一般地，R^n 中有没有标准正交基？什么样的 n 级矩阵其行(或列)向量组是 R^n 的一个标准正交基？这一节就来讨论这两个问题.

定义1 实数域上的方阵 A 如果满足

$$AA'=I \tag{1}$$

则称 A 是**正交矩阵**.

由(1)式知，正交矩阵 A 一定是可逆的，并且 $A^{-1}=A'$. 反之，若实方阵（即，实数域上的方阵）A 可逆，并且满足 $A^{-1}=A'$，则 $I=AA^{-1}=AA'$，从而 A 是正交矩阵. 这证明了下述命题：

命题4.10.1　实方阵 A 是正交矩阵的充分必要条件为 A 可逆，并且 $A^{-1}=A'$.　▌

从命题4.10.1立即得到

推论4.10.1　实方阵 A 是正交矩阵的充分必要条件为

$$A'A=I \tag{2}$$

▌

例1　判断下述矩阵是否正交矩阵：

$$A=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}，\text{其中 }\theta\text{ 是实数}$$

解　因为

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

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

所以 A 是正交矩阵.

正交矩阵具有下列性质：

1°　I 是正交矩阵；

2°　若 A 与 B 都是 n 级正交矩阵，则 AB 也是正交矩阵；

3°　若 A 是正交矩阵，则 A^{-1}（即 A'）也是正交矩阵；

4°　若 A 是正交矩阵，则 $|A|=1$ 或 -1.

证明　1°与3°的证明很容易. 现在证2°与4°.

若 A,B 都是正交矩阵，则

$$(AB)(AB)'=A(BB')A'=AIA'=I$$

因此 AB 也是正交矩阵.

若 A 是正交矩阵，则 $|AA'|=1$. 由此得 $|A||A'|=1$，即 $|A|^2=1$. 所以 $|A|=1$ 或 -1.　▌

命题4.10.2　n 级实矩阵 $A=(a_{ij})$ 是正交矩阵的充分必要条件为：A 的每一行元素的平方和等于 1，每两行对应元素的乘积之和等于零，即

$$\sum_{k=1}^{n} a_{ik}^{2}=1, \quad i=1,2,\cdots,n \tag{3}$$

$$\sum_{k=1}^{n} a_{ik}a_{jk}=0, \quad i\neq j \tag{4}$$

证明　对于任一 n 级矩阵 $A=(a_{ij})$ 都有

$$\begin{aligned}&(AA')(i;j)\\ &=\sum_{k=1}^{n} A(i;k)A'(k;j)=\sum_{k=1}^{n} a_{ik}a_{jk}\end{aligned} \tag{5}$$

于是

$$\begin{aligned}&n \text{ 级实矩阵 } A \text{ 是正交矩阵} \Leftrightarrow AA'=I\\ &\Leftrightarrow \begin{cases}(AA')(i;i)=1, & i=1,2,\cdots,n\\ (AA')(i;j)=0, & j\neq i\end{cases}\\ &\Leftrightarrow \begin{cases}\displaystyle\sum_{k=1}^{n} a_{ik}^{2}=1, & i=1,2,\cdots,n\\ \displaystyle\sum_{k=1}^{n} a_{ik}a_{jk}=0, & j\neq i\end{cases}\end{aligned} \tag{6}$$

▌

我们引用一个符号 δ_{ij}，它的含意是

$$\delta_{ij}=\begin{cases}1, & \text{当 } i=j\\ 0, & \text{当 } i\neq j\end{cases} \tag{7}$$

δ_{ij} 称为 Kronecker 记号. 采用这个符号，则(3)和(4)式可合并写成

$$\sum_{k=1}^{n} a_{ik}a_{jk}=\delta_{ij}, \quad 1\leqslant i,j\leqslant n \tag{8}$$

命题4.10.3　n 级实矩阵 $A=(a_{ij})$ 是正交矩阵的充分必要条件为：A 的每一列元素的平方和等于 1，每两列对应元素的乘积

之和等于零,即

$$\sum_{k=1}^{n} a_{ki}a_{kj} = \delta_{ij}, \quad 1 \leqslant i,j \leqslant n \tag{9}$$

证明 A 是正交矩阵 $\Leftrightarrow A'$ 是正交矩阵. 对 A' 用命题 4.10.2 即得(9)式. ▌

(8)式(或(9)式)的左端使我们联想到解析几何中两个向量 $\vec{a}$ 与 $\vec{b}$ 的内积 $\vec{a}\cdot\vec{b}$ 在直角坐标系中的计算公式:

$$\vec{a}\cdot\vec{b} = a_1b_1 + a_2b_2 + a_3b_3$$

其中 (a_1,a_2,a_3), (b_1,b_2,b_3) 分别为 $\vec{a}$, $\vec{b}$ 的直角坐标. 从这个几何背景受到启发,我们可以在实数域上的 n 维向量空间 R^n 中引进内积的概念:

定义2 在 R^n 中,任给 $\alpha=(a_1,a_2,\cdots,a_n)$, $\beta=(b_1,b_2,\cdots,b_n)$, 规定

$$(\alpha,\beta) := a_1b_1 + a_2b_2 + \cdots + a_nb_n \tag{10}$$

这个二元实值函数 (α,β) 称为 R^n 的一个**内积**,通常称这个内积为 R^n 的**标准内积**.

容易直接按照定义验证 R^n 的标准内积具有下列基本性质: $\forall\ \alpha,\beta,\gamma \in R^n, k \in R$, 有

1° $(\alpha,\beta) = (\beta,\alpha)$;

2° $(\alpha+\gamma,\beta) = (\alpha,\beta) + (\gamma,\beta)$;

3° $(k\alpha,\beta) = k(\alpha,\beta)$;

4° $(\alpha,\alpha) \geqslant 0$, 并且 $(\alpha,\alpha) = 0$ 当且仅当 $\alpha = 0$.

从以上四条性质看出, R^n 的标准内积具有对称性,对每个变量都是线性,以及正定性(即 4°).

由1°,2°,3°可以得出

$$(k_1\alpha_1 + k_2\alpha_2,\beta) = k_1(\alpha_1,\beta) + k_2(\alpha_2,\beta)$$

$$(\alpha,k_1\beta_1 + k_2\beta_2) = k_1(\alpha,\beta_1) + k_2(\alpha,\beta_2)$$

定义3 向量空间 R^n 如果定义了一个内积,则称 R^n 为一个**欧几里得空间**,简称为**欧氏空间**.

本章和第五、六章提到欧氏空间 R^n 时，其内积均指由(10) 式定义的内积，即标准内积.

解析几何中向量 $\vec{a}$ 的长度可用内积计算：$|\vec{a}| = \sqrt{(\vec{a} \cdot \vec{a})}$. 以这事实为背景，我们在欧氏空间 R^n 中引进向量的长度的概念：

定义4　欧氏空间 R^n 中向量 α 的**长度** $|\alpha|$ 规定为

$$|\alpha| := \sqrt{(\alpha,\alpha)} \tag{11}$$

长度为 1 的向量称为**单位向量**. α 是单位向量 $\Leftrightarrow (\alpha,\alpha) = 1$.

容易验证，$|k\alpha| = |k||\alpha|$. 于是对于 $\alpha \neq 0$，$|\alpha|^{-1}\alpha$ 一定是单位向量. 把非零向量 α 乘以 $|\alpha|^{-1}$ 称为把 α **单位化**.

解析几何中有 $\vec{a} \perp \vec{b} \Leftrightarrow \vec{a} \cdot \vec{b} = 0$. 以这为背景我们在欧氏空间 R^n 中引进正交的概念：

定义5　欧氏空间 R^n 中，如果 $(\alpha,\beta) = 0$，则称 α 与 β 是**正交的或互相垂直的**，记作 $\alpha \perp \beta$.

因为 $(0,\beta) = 0, \forall \beta \in R^n$，所以 $0 \perp \beta, \forall \beta \in R^n$.

定义6　欧氏空间 R^n 中由非零向量组成的向量组称为**正交向量组**，如果其中每两个不同的向量都正交(即它们两两正交).

据定义6，由一个非零向量组成的向量组也是正交向量组.

如果一个正交向量组的每个向量都是单位向量，则称它为**正交单位向量组**.

命题4.10.4　欧氏空间 R^n 中，正交向量组一定是线性无关的.

证明　设 $\alpha_1, \cdots, \alpha_s$ 是 R^n 的一个正交向量组. 设

$$k_1\alpha_1 + \cdots + k_s\alpha_s = 0 \tag{12}$$

(12)式两端的向量都与 α_i 作内积，得

$$(k_1\alpha_1 + \cdots + k_s\alpha_s,\ \alpha_i) = (0,\alpha_i) \tag{13}$$

由于 $\alpha_1, \cdots, \alpha_s$ 彼此正交，因此(13) 式成为

$$k_i(\alpha_i,\alpha_i) = 0 \tag{14}$$

因为 $\alpha_i \neq 0$，所以 $(\alpha_i,\alpha_i) \neq 0$. 于是由(14) 式得

$$k_i = 0, \quad i = 1,2,\cdots,s$$

这证明了 $\alpha_1,\cdots,\alpha_s$ 线性无关. ▌

由命题 4.10.4 得出，欧氏空间 R^n 中，n 个向量组成的正交向量组是 R^n 的一个基，称它为**正交基**. 而 n 个向量组成的正交单位向量组称为 R^n 的一个**标准正交基**.

例如，容易验证，向量空间 R^n 的标准基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是欧氏空间 R^n 的一个标准正交基.

命题4.10.5 $A=(a_{ij})$ 是 n 级正交矩阵的充分必要条件是，A 的行(或列)向量组是欧氏空间 R^n 的一个标准正交基.

证明 A 的行向量组记作 $\gamma_1,\cdots,\gamma_n$. 据命题 4.10.2 得

n 级实矩阵 A 是正交矩阵

$$\Leftrightarrow(\gamma_i,\gamma_j)=\sum_{k=1}^{n}a_{ik}a_{jk}=\delta_{ij},\quad 1\leqslant i,j\leqslant n$$

$\Leftrightarrow\gamma_1,\cdots,\gamma_n$ 是正交的单位向量组

$\Leftrightarrow\gamma_1,\cdots,\gamma_n$ 是 R^n 的一个标准正交基

同理，A 是正交矩阵的充分必要条件是，A 的列向量组 $\alpha_1,\cdots,\alpha_n$ 是 R^n 的一个标准正交基. ▌

命题4.10.5 使我们对正交矩阵有了更深一层的认识. 找一个 n 级正交矩阵等价于在 R^n 中找一个标准正交基. 给了 R^n 的一个基，如何从它出发构造出 R^n 的一个标准正交基？我们可以把问题提得更一般化：给了 R^n 的一个线性无关的向量组 $\alpha_1,\cdots,\alpha_s$，如何从它出发构造出一个正交向量组 $\beta_1,\cdots,\beta_s$，并且使得 $\langle\alpha_1,\cdots,\alpha_s\rangle=\langle\beta_1,\cdots,\beta_s\rangle$？让我们先看几何空间中的例子，以便受到启发.

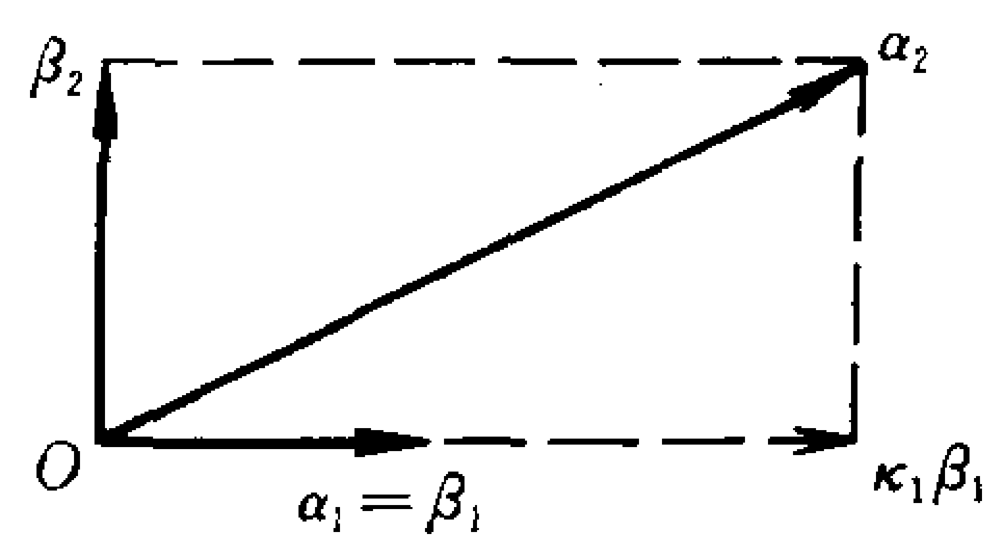
几何空间中，设 α_1 与 α_2 不共线，要找向量 β_1,β_2，使得 $\beta_1\perp\beta_2$，并且 $\langle\alpha_1,\alpha_2\rangle=\langle\beta_1,\beta_2\rangle$. 取 $\beta_1=\alpha_1$，把 α_2 表示成两个向量的和，其中一个与 β_1 平行，记作 $k_1\beta_1$，另一个与 β_1 垂直，记作 β_2(β_2 称为 α_2 沿方向 β_1 下的外射影)，即

$$\alpha_2=k_1\beta_1+\beta_2 \tag{15}$$

(15)式两边用 β_1 作内积得

$$(\alpha_2,\beta_1)=k_1(\beta_1,\beta_1)$$

由此得出
$$k_1 = \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)} \tag{16}$$



代入(15)式得
$$\beta_2 = \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1 \tag{17}$$
(17)式是α_2沿方向β_1下的外射影的公式. 由前述知,$\beta_1 \perp \beta_2$,并且显然有$\langle \alpha_1, \alpha_2 \rangle = \langle \beta_1, \beta_2 \rangle$.

定理4.10.1　设 $\alpha_1, \cdots, \alpha_s$ 是欧氏空间 R^n 的一个线性无关的向量组. 令
$$\begin{aligned}
\beta_1 &= \alpha_1 \\
\beta_2 &= \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1 \\
&\cdots \quad \cdots \quad \cdots \quad \cdots \\
\beta_s &= \alpha_s - \sum_{i=1}^{s-1} \frac{(\alpha_s, \beta_i)}{(\beta_i, \beta_i)}\beta_i
\end{aligned} \tag{18}$$
则 $\beta_1, \beta_2, \cdots, \beta_s$ 是正交向量组,并且
$$\langle \alpha_1, \alpha_2, \cdots, \alpha_s \rangle = \langle \beta_1, \beta_2, \cdots, \beta_s \rangle$$

证明　对 s 用归纳法. $s = 1$ 时,令 $\beta_1 = \alpha_1$,则 β_1 是正交向量组(因为 α_1 线性无关,所以 $\beta_1 = \alpha_1 \neq 0$),并且显然$\langle \alpha_1 \rangle = \langle \beta_1 \rangle$.

假设 $s = k$ 时命题为真,即 $\beta_1, \cdots, \beta_k$ 是正交向量组,并且$\langle \alpha_1, \cdots, \alpha_k \rangle = \langle \beta_1, \cdots, \beta_k \rangle$. 现在看 $s = k + 1$ 情形: 令

$$\beta_{k+1} = \alpha_{k+1} - \sum_{i=1}^{k} \frac{(\alpha_{k+1},\beta_i)}{(\beta_i,\beta_i)}\beta_i \tag{19}$$

对于 $1 \leqslant j \leqslant k$,(19) 式两边用 β_j 去作内积,得

$$\begin{aligned}(\beta_{k+1},\beta_j) &= (\alpha_{k+1},\beta_j) - \sum_{i=1}^{k} \frac{(\alpha_{k+1},\beta_i)}{(\beta_i,\beta_i)}(\beta_i,\beta_j) \\ &= (\alpha_{k+1},\beta_j) - \frac{(\alpha_{k+1},\beta_j)}{(\beta_j,\beta_j)}(\beta_j,\beta_j) = 0\end{aligned}$$

因此 $\beta_{k+1} \perp \beta_j (1 \leqslant j \leqslant k)$. 从(19)式知,$\beta_{k+1}$ 可由 $\alpha_1,\cdots,\alpha_k,\alpha_{k+1}$ 线性表出,并且表出式中 α_{k+1} 的系数为 1. 由于 $\alpha_1,\cdots,\alpha_k,\alpha_{k+1}$ 线性无关,所以 $\beta_{k+1} \neq 0$. 因此 $\beta_1,\cdots,\beta_k,\beta_{k+1}$ 是正交向量组. 易看出 $\{\alpha_1,\cdots,\alpha_k,\alpha_{k+1}\} \cong \{\beta_1,\cdots,\beta_k,\beta_{k+1}\}$,因此 $\langle\alpha_1,\cdots,\alpha_k,\alpha_{k+1}\rangle = \langle\beta_1,\cdots,\beta_k,\beta_{k+1}\rangle$.

根据数学归纳法原理,对一切 $s \geqslant 1$,命题为真. ▌

公式 (18) 给出了由 R^n 中一个线性无关的向量组 $\alpha_1,\cdots,\alpha_s$ 出发,构造出与它等价的一个正交向量组 $\beta_1,\cdots,\beta_s$ 的方法,这种方法称为 Gram — Schmidt **正交化过程**. 只要再将 $\beta_1,\cdots,\beta_s$ 中每个向量 β_j 单位化,即,令 $\eta_j = |\beta_j|^{-1}\beta_j, j = 1,\cdots,s$,则 $\eta_1,\cdots,\eta_s$ 是与 $\alpha_1,\cdots,\alpha_s$ 等价的正交单位向量组.

例2 设向量组

$$\alpha_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \qquad \alpha_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

求与 α_1,α_2 等价的一个正交单位向量组.

解 先正交化,令

$$\beta_1 = \alpha_1$$

$$\beta_2 = \alpha_2 - \frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} - \frac{4}{5}\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} \\ \frac{4}{5} \\ 1 \end{pmatrix}$$

然后单位化,令

$$\eta_1 = \frac{1}{|\beta_1|}\beta_1 = \frac{1}{\sqrt{5}}\beta_1 = \begin{bmatrix} \frac{2}{5}\sqrt{5} \\ -\frac{1}{5}\sqrt{5} \\ 0 \end{bmatrix}$$

$$\eta_2 = \frac{1}{|\beta_2|}\beta_2 = \frac{\sqrt{5}}{3}\beta_2 = \begin{bmatrix} \frac{2}{15}\sqrt{5} \\ \frac{4}{15}\sqrt{5} \\ \frac{1}{3}\sqrt{5} \end{bmatrix}.$$

则 η_1,η_2 是与 α_1,α_2 等价的一个正交单位向量组.

推论4.10.1　欧氏空间 R^n 中,任给一个线性无关的向量组 $\alpha_1,\cdots,\alpha_s$,都能从它出发构造出 R^n 的一个标准正交基.

证明　先把 $\alpha_1,\cdots,\alpha_s$ 扩充成向量空间 R^n 的一个基:$\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$. 然后用 Gram – Schmidt 正交化过程并且把所得向量单位化,得到一个正交的单位向量组:$\eta_1,\cdots,\eta_n$. 这就是欧氏空间 R^n 的一个标准正交基. ▌

欧氏空间 R^n 中,两个列向量 α 与 β 的内积也可以用矩阵乘法的形式表达出来:设

$$\alpha = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \qquad \beta = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

则

$$(\alpha,\beta) = a_1b_1 + a_2b_2 + \cdots + a_nb_n = \alpha'\beta \tag{20}$$

习 题 4.10

1. 证明：若 A 是 n 级实对称矩阵，T 是 n 级正交矩阵，则 $T^{-1}AT$ 是实对称矩阵.

2. 证明：如果正交矩阵 A 是上三角矩阵，则 A 是对角矩阵，并且 A 的主对角元是 1 或 -1.

3. 证明：如果正交矩阵 A 是分块上三角矩阵，则 A 是分块对角矩阵，并且 A 的主对角线上的所有子矩阵都是正交矩阵.

4. 设 A 是正交矩阵，证明：如果 $|A|=1$，则 A 的每一个元素等于它自己的代数余子式；如果 $|A|=-1$，则 A 的每一个元素等于它自己的代数余子式乘以 -1.

5. 证明：(1) 若 $|A|=1$，并且 A 的每一个元素等于它自己的代数余子式，则 A 是正交矩阵；(2) 若 $|A|=-1$，并且 A 的每一个元素等于它自己的代数余子式乘以 -1，则 A 是正交矩阵.

*6. 设 A 是 n 级实方阵，$n\geqslant 3$. 证明：(1) 如果 A 的每一个元素等于它自己的代数余子式，并且 A 至少有一个元素不为零，则 A 是正交矩阵；(2) 如果 A 的每一个元素等于它自己的代数余子式乘以 -1，并且 A 至少有一个元素不为零，则 A 是正交矩阵.

*7. 证明：设 A 是正交矩阵，任意取定 A 的两行(或两列)，证明：位于这两行(或两列)的所有 2 级子式的平方和等于 1.(提示：利用乘积矩阵的各级子式的计算公式，参见本章 §8 命题 4.8.1.)

*8. 证明：位于正交矩阵的任何 k 行(或 k 列)的所有 k 级子式的平方和等于 1.(提示：类似于第 7 题.)

9. 在怎样的条件下，对角矩阵是正交矩阵？

10. 方阵 A 称为**对合矩阵**，如果 $A^2=I$. 证明：一个实方阵如果有下列三个性质中的任两个性质，则必有第三个性质：对称矩阵，正交矩阵，对合矩阵.

11. 证明：在欧氏空间 R^n 中，若 α 与 β 正交，则对任意实数 k,l，有 $k\alpha$ 与 $l\beta$ 正交.

12. 证明：在欧氏空间 R^n 中，若 β 与 $\alpha_1,\cdots,\alpha_s$ 都正交，则 β 与 $\langle\alpha_1,\cdots,\alpha_s\rangle$ 中每一个向量正交.

13. 证明：在欧氏空间 R^n 中，若 α 与任意向量都正交，则 α 必定是零向量.

14. 给了下列线性无关的向量组，分别求与它们等价的正交单位向量组.

(1) $\alpha_1=(1,-2,0)$, $\alpha_2=(1,0,-1)$

(2) $\alpha_1=(1,1,1,1)$, $\alpha_2=(1,0,1,0)$

(3) $\alpha_1=(1,1,0,0)$, $\alpha_2=(1,0,1,0)$

$\alpha_3=(1,0,0,1)$

(4) $\alpha_1=(1,1,0,0)$, $\alpha_2=(1,0,1,0)$

$\alpha_3=(-1,0,0,1)$, $\alpha_4=(1,-1,-1,1)$

15. 设 A 是 n 级正交矩阵，α 是欧氏空间 R^n 中的列向量，证明：$|A\alpha|=|\alpha|$.

16. 设 A 是一个 n 级可逆实矩阵，证明：A 可以分解成

$$A=TB$$

其中 T 是正交矩阵，B 是上三角矩阵，并且 B 的主对角元都为正数；证明这个分解是唯一的.

*17. 证明：齐次线性方程组

$$\begin{cases} ax_1+bx_2+cx_3+dx_4=0 \\ bx_1-ax_2+dx_3-cx_4=0 \\ cx_1-dx_2-ax_3+bx_4=0 \\ dx_1+cx_2-bx_3-ax_4=0 \end{cases}$$

只有零解.

*18. 设 A 是复数域上矩阵，用 A^* 表示 $\overline{A'}$. 复方阵 A 称为**酉矩阵**，如果 $AA^*=I$. 证明：下列每一个条件都是 n 级复方阵 $A=(a_{ij})$ 为酉矩阵的充分必要条件：

(1) A 可逆，并且 $A^{-1}=A^*$；

(2) $A^*A=I$；

(3) $\sum_{k=1}^{n} a_{ik}\bar{a}_{jk} = \delta_{ij}$;

(4) $\sum_{k=1}^{n} a_{ki}\bar{a}_{kj} = \delta_{ij}$.

* 19. 证明：酉矩阵的行列式的模等于 1.

* 20. 证明：酉矩阵的乘积是酉矩阵；酉矩阵的逆矩阵是酉矩阵.

* 21. 设 W 是欧氏空间 R^n 的一个子空间，如果向量 α 与 W 中每一个向量正交，则称 α 与 W **正交**，或者称 α **垂直于** W. 设 φ 是 R^n 到 R^n 的一个映射，如果对一切 $X \in R^n$，都有 $\varphi(X) \in W$，并且 $X - \varphi(X)$ 垂直于 W，则称 φ 是 R^n 在 W 上的**正投影**（简称**投影**），并且把向量 $\varphi(X)$ 称为向量 X 在 W 上的**正投影**（简称**投影**，或**射影**）. 现在设 A 是 $n \times m$ 列满秩实矩阵，$n > m$，它的列向量组 $\alpha_1, \cdots, \alpha_m$ 生成的子空间记作 W. 设 $P_A = A(A'A)^{-1}A'$. 令

$$\underline{P_A}(X) = P_A X, \quad \forall X \in R^n$$

证明：$\underline{P_A}$ 是 R^n 在 W 上的投影.

补 充 题 四

1. 方阵 P 称为**幂等矩阵**，如果 $P^2 = P$. 证明：如果 P 是幂等矩阵，则 $2P - I$ 是对合矩阵；反之，如果 A 是对合矩阵，则 $\frac{1}{2}(A + I)$ 是幂等矩阵.

2. 方阵 A 称为**幂零矩阵**，如果 A 的某个正整数次幂等于零矩阵. 使 $A^k = 0$ 成立的最小正整数 k 称为 A 的**幂零指数**. 证明：上（下）三角形矩阵是幂零阵当且仅当所有的主对角元等于零，并且它的幂零指数不超过矩阵的级数.（提示：用归纳法.）

3. 证明：与所有行列式为 1 的 n 级矩阵可交换的矩阵一定是数量矩阵.（提示：$I + E_{1j}$ 的行列式为 1，当 $j \neq 1$；$|I + E_{21}| = 1$.）

4. 证明：与所有 n 级可逆矩阵可交换的矩阵一定是数量矩阵.

5. 证明：如果整数 a, b 都能表示成两个整数的平方和，则 ab 也能表示成两个整数的平方和.

6. 把 533 表示成两个整数的平方和.

7. 证明：如果整数 a, b 都能表示成四个整数的平方和，则 ab 也能表示

成四个整数的平方和.（提示：设 $a = m_1^2 + m_2^2 + m_3^2 + m_4^2, b = n_1^2 + n_2^2 + n_3^2 + n_4^2$. 令

$$A = \begin{pmatrix} m_1 & m_2 & m_3 & m_4 \\ m_2 & -m_1 & m_4 & -m_3 \\ m_3 & -m_4 & -m_1 & m_2 \\ m_4 & m_3 & -m_2 & -m_1 \end{pmatrix}$$

则 $AA' = aI$. 于是 $ab = a\beta\beta' = \beta(aI)\beta' = \beta AA'\beta'$，其中 $\beta = (n_1, n_2, n_3, n_4)$.）

8. 把 1457 表示成四个整数的平方和.

9. 证明：如果整数 a, b 都能表示成形式为 $x^3 + y^3 + z^3 - 3xyz$ 的数，那么 ab 也能表示成这种形式的数.

10. 用 1_n 表示 K^n 中分量全为 1 的列向量. 实数域 R 上每行（列）的元素之和都等于 1 的非负矩阵（即元素都是非负数的矩阵）称为**行（列）随机矩阵**. 证明：

(1) 非负矩阵 $A_{s\times n}$ 是行随机矩阵 $\Leftrightarrow A1_n = 1_s$；

(2) 非负矩阵 $A_{s\times n}$ 是列随机矩阵 $\Leftrightarrow 1'_s A = 1'_n$；

(3) 若 $A_{s\times n}, B_{n\times s}$ 都是行（列）随机矩阵，则 AB 也是行（列）随机矩阵；

(4) 可逆的行（列）随机矩阵的逆仍是行（列）随机矩阵.

11. 设 $A = (a_{ij})$ 为一个 n 级实矩阵，已知

$$a_{ii} > 0, \qquad i = 1,2,\cdots,n$$
$$a_{ij} < 0, \qquad i \neq j;\quad i,j = 1,2,\cdots,n$$
$$\sum_{j=1}^{n} a_{ij} = 0, \qquad i = 1,2,\cdots,n$$

证明：$\mathrm{rank}(A) = n - 1$.（提示：去证 $A\begin{pmatrix} 1, & 2, & \cdots, & n-1 \\ 1, & 2, & \cdots, & n-1 \end{pmatrix} \neq 0$.）

12. 一个**区组设计**是把 v 个不同的对象编进 b 个区组里的一种排法，使得

1）每个区组恰好包含 k 个不同的对象（$2 \leqslant k < v$）；

2）每两个不同的对象一起恰好出现在 λ 个区组里.

一个参数为 (v,b,k,λ) 区组设计可以用一个 $v \times b$ 矩阵 A 来表示，其中

$$A(i;j)=\begin{cases}1, & \text{当对象 } P_i \text{ 出现在区组 } B_j \text{ 里}\\ 0, & \text{其他}\end{cases}$$

这个 矩阵 A 称为该区组设计的**关联矩阵**. 证明：A 的每列元素的和(简称为列和) 都是 k；A 的每两行的内积等于 λ.

13. 设 A 是参数为 (v,b,k,λ) 的区组设计的关联矩阵，证明：A 的每一行元素的和(简称为行和) 是个常数，它等于 $\frac{\lambda(v-1)}{k-1}$. 把这个数记作 r. (提示：设 A 的行向量组是 $\alpha_1,\cdots,\alpha_v$，任意给定 $i(1\leqslant i\leqslant v)$，用两种方法计算 $\sum\limits_{j\neq i}(\alpha_i,\alpha_j)$.)

14. 证明：参数为 (v,b,k,λ) 的区组设计必满足：

$$\lambda(v-1)=r(k-1),\qquad rv=kb.$$

15. 设 A 是参数为 (v,b,k,λ) 的区组设计的关联矩阵，计算 $|AA'|$；并且求 $\operatorname{rank}(AA')$.

16. 证明：区组设计的参数满足：

$$v\leqslant b$$

17. $v=b$ 的区组设计称为**对称设计**. 设 A 是对称设计的关联矩阵. 证明：

(1) $r=k$；

(2) $A'A=AA'$；

(3) 若 v 是偶数，则 $k-\lambda$ 一定是平方数.

18. 元素 全为整数的矩阵称为**整数矩阵**. 行列式等于 ± 1 的整数矩阵称为**么模矩阵**. 对于一个整数矩阵 A，如果存在一个整数矩阵 B，使得 $AB=BA=I$，则称 A **有整数逆矩阵**. 证明：整数矩阵 A 有整数逆矩阵当且仅当 A 是么模矩阵.

第五章　矩阵的相抵分类与相似分类

§1　引　　言

这一章我们要讨论矩阵的相抵分类与方阵的相似分类.

把一个集合的元素按照一定的关系进行分类，在每一类中选取一个元素研究，以便了解这一类元素的共同性质. 这种方法无论在日常工作中或者数学中经常被使用. 这一章我们将分别对数域 K 上所有 $s\times n$ 矩阵组成的集合 $M_{s\times n}(K)$ 以及所有 n 级矩阵组成的集合 $M_n(K)$ 按照某种关系进行分类，讨论每一类中具有最简单形式的矩阵是什么样，即所谓标准形的问题；并且研究同一类中的矩阵的共同性质以及这种共同性质在量上的反映，即所谓不变量的问题，而且希望找到的不变量足以区分不同的类，即所谓完全不变量的问题. 寻求矩阵在某种关系下的标准形和完全不变量是研究矩阵分类的两个基本问题.

在讨论方阵的相似分类时，将涉及方阵的特征值和特征向量这两个重要的概念，它们在数学的各个分支以及自然科学，工程技术中都是十分有用的.

为了把集合中元素的分类说得严密，我们在下一节将介绍数学中的两个基本概念：等价关系以及集合的划分.

§2 等价关系·集合的划分

(一) 二元关系

给定 两个集合 A 与 B,下述集合

$$\{(a,b)|a \in A, \quad b \in B\}$$

称为 A 与 B 的**笛卡儿积**,记作 $A \times B$.

我们 常常谈论一个集合 S 中元素之间的各种关系.譬如,实数集 R 中元素之间的小于关系,大于关系,等于关系.又如,设 S 表示北京大学数学系和概率统计系 95 级全体同学组成的集合,则 S 中 元素之间有同系关系,同班关系,同乡关系(从同一个省或市来到北大),等等.如何用数学语言描述一个关系?以 S 中同系关系为例.设 S_1,S_2 分别表示数学系 95 级全体同学、概率系 95 级全体同学组成的集合.对于 $a,b \in S$,我们有

$$\begin{aligned} a \text{ 与 } b \text{ 同系} &\Leftrightarrow (a,b) \in S_1 \times S_1, \text{或} (a,b) \in S_2 \times S_2 \\ &\Leftrightarrow (a,b) \in (S_1 \times S_1) \cup (S_2 \times S_2) \end{aligned}$$

记 $W_1 = (S_1 \times S_1) \cup (S_2 \times S_2)$,显然 W_1 是 $S \times S$ 的一个子集.于是

$$a \text{ 与 } b \text{ 同系} \Leftrightarrow (a,b) \in W_1$$

因此,我们把 $S \times S$ 的这个子集 W_1 就叫做 S 上的同系关系.如果 $(a,b) \in W_1$,就说 a 与 b 有同系关系,记作 aW_1b;若 $(a,b) \bar{\in} W_1$,就说 a 与 b 没有同系关系.一般地,我们有

定义 1 设 A 与 B 是两个集合,$A \times B$ 的任一子集 W 称为 A 与 B 之间的一个**二元关系**.若 $B = A$,则 $A \times A$ 的子集 W 称为 A 上的一个二元关系.如果 $(a,b) \in W$,则称 a 对 b 有关系 W,并且用符号 aWb(或 $a \sim b$)来表示.如果 $(a,b) \bar{\in} W$,则称 a 对 b 没有关系 W.

例如，实数集 R 上的顺序关系“$<$”就是 R 上的一个二元关系，这时 $W=\{(x,y)\mid x<y\}$，即 W 是由平面内位于直线 $y=x$ 的上侧的点的坐标组成的. 这时记号 xWy 用 $x<y$ 代替.

(二)　等价关系

定义 2　集合 S 上的一个二元关系 $\sim$ 称为**等价关系**，如果对所有的 $a,b,c\in S$，下列条件成立：

1°　$a\sim a$(反身性)；

2°　$a\sim b\Rightarrow b\sim a$(对称性)；

3°　$a\sim b$ 且 $b\sim c\Rightarrow a\sim c$(传递性).

前面讲的 S 上的同系关系是一个等价关系. R 上的小于关系不是等价关系，因为它不满足反身性和对称性.

定义 3　设 $\sim$ 是集合 S 上的一个等价关系，$a\in S$. 所有与 a 有关系 $\sim$ 的元素组成的子集

$$\{x\in S\mid x\sim a\}$$

称为由 a 确定的**等价类**($\sim$ 等价类)，记作 $\bar{a}$.

由定义 3 知

$$x\in\bar{a}\Leftrightarrow x\sim a$$

因为 $a\sim a$，所以 $a\in\bar{a}$. 从而 $\bar{a}$ 也称为含 a 的等价类.

因为

$$x,y\in\bar{a}\Rightarrow x\sim a \text{ 且 } y\sim a\Rightarrow x\sim y$$

所以同一个等价类里的任意两个元素都有关系 $\sim$.

例如，在(一)中讲的 S 上的同系关系下，设 $a\in S_1$，则 $\bar{a}=S_1$；设 $b\in S_2$，则 $\bar{b}=S_2$.

又如，设 S 是平面上所有点组成的集合. 定义

$$a\sim b\Leftrightarrow a \text{ 与 } b \text{ 位于同一条水平线上}$$

易看出，$\sim$ 是 S 上的一个等价关系. a 确定的等价类 $\bar{a}$ 是通过点 a 的一条水平线. 还可看出，若 $a\sim b$，则 $\bar{a}$ 与 $\bar{b}$ 是同一条水平线；反之，若 $\bar{a}$ 与 $\bar{b}$ 是同一条水平线，显然有 $a\sim b$. 一般地，我们有

命题 5.2.1　设 $\sim$ 是集合 S 上的一个等价关系，$a,b \in S$，则

$$\bar{a} = \bar{b} \Leftrightarrow a \sim b$$

证明　设 $a \sim b$. 任取 $c \in \bar{a}$，则 $c \sim a$. 由传递性得，$c \sim b$. 因此 $c \in \bar{b}$. 这证明了 $\bar{a} \subset \bar{b}$. 由于对称性，有 $b \sim a$，于是据刚才证得的结论，我们有 $\bar{b} \subset \bar{a}$. 所以 $\bar{a} = \bar{b}$.

反之，设 $\bar{a} = \bar{b}$. 因为 $a \in \bar{a} = \bar{b}$，所以 $a \sim b$. ▌

$\bar{a}$ 中的任一元素都叫做 $\bar{a}$ 的一个**代表**.

（三）　集合的划分

（一）中讲的集合 S 是它的两个子集 S_1 和 S_2 的并集，并且 S_1 与 S_2 不相交（即 $S_1 \cap S_2 = \varnothing$）. 我们称集合 $\{S_1, S_2\}$ 是 S 的一个划分. 一般地，我们有

定义 4　如果集合 S 是一些非空子集 S_i（$i \in I$，这里 I 表示指标集）的并集，并且其中不相等的子集一定不相交，则称集合 $\{S_i | i \in I\}$ 是 S 的一个**划分**，记作 $\pi(S)$.

例如，设 $S = \{1,2,3,4,5,6\}$，令 $S_1 = \{1,2\}$，$S_2 = \{3,4,5\}$，$S_3 = \{6\}$. 则 $S = S_1 \cup S_2 \cup S_3$，并且 S_1, S_2, S_3 彼此不相交. 于是集合 $\{S_1, S_2, S_3\}$ 是 S 的一个划分. 显然，集合 $\{\{1,2,3\},\{4,5,6\}\}$ 是 S 的另一个划分. 但是集合 $\{\{1,2,3\},\{3,4\},\{5,6\}\}$ 不是 S 的划分.

集合的划分与等价关系这两个概念有密切的关系. 例如，（一）中的 S 在同系关系下，恰有两个等价类：S_1, S_2. 它们组成的集合 $\{S_1, S_2\}$ 正好是 S 的一个划分. 一般地，我们有

命题 5.2.2　设 $\sim$ 是集合 S 上的一个等价关系，则所有 $\sim$ 等价类组成的集合是 S 的一个划分，称为 $\sim$ 伴随的划分，记作 $\pi_{\sim}(S)$.

证明　因为对于 $a \in S$，有 $a \in \bar{a}$，所以 $S = \bigcup_{a \in S} \bar{a}$.

现在来证明不相等的等价类是不相交的：设 $\bar{a} \neq \bar{b}$. 假如 $c \in$

$\bar{a} \cap \bar{b}$,则 $c \sim a$,并且 $c \sim b$. 由对称性和传递性得 $a \sim b$. 于是 $\bar{a} = \bar{b}$,矛盾. 所以 $\bar{a} \cap \bar{b} = \varnothing$.

综上述,所有等价类组成的集合是 S 的一个划分. ▎

例如,在上面我们曾提出,设 S 是平面的所有点组成的集合,$\sim$ 是如上定义的一个等价关系,则等价类 $\bar{a}$ 是通过 a 的一条水平线. 显然,所有等价类(即所有水平线)组成的集合是平面 S 的一个划分.

命题 5.2.2 说明,给了集合 S 上的一个等价关系,就可以得到 S 的一个划分. 反之,给了集合 S 的一个划分,是否可得到 S 上的一个等价关系? 例如,(一) 中的 S 有一个划分:S_1,S_2. 若规定

$$a \sim b \Leftrightarrow (a,b) \in \bigcup_{i=1}^{2} S_i \times S_i$$

则 $\sim$ 就是 S 上的同系关系,它是一个等价关系,并且 S_1,S_2 都是等价类. 一般地,有

命题 5.2.3 设 $\pi(S) = \{ S_i | i \in I\}$ 是集合 S 的一个划分,则我们可以得到 S 上的一个等价关系 $\sim$,使得每个 S_i 是一个等价类,从而 $\pi_{\sim}(S) = \pi(S)$.

* **证明** 因为 $\pi(S) = \{ S_i | i \in I\}$ 是 S 的一个划分,所以每个 $a \in S$ 恰好被包含在一个子集 S_i 中. 我们定义 S 上的一个二元关系 $\sim$ 如下:

$$a \sim b \Leftrightarrow (a,b) \in \bigcup_{i \in I} S_i \times S_i$$

显然,这个关系 $\sim$ 是反身的,对称的,传递的. 因此它是 S 上的一个等价关系. 显然,a 确定的等价类 $\bar{a}$ 是包含 a 的子集 S_i. 因此这个等价关系伴随的 S 的划分 $\pi_{\sim}(S)$ 与所给的 S 的划分 $\pi(S)$ 一致. ▎

从命题 5.2.2 和从命题 5.2.3 得出,在集合 S 的等价关系与集合 S 的划分之间存在着一一对应.

因此,为了把集合 S 的元素分类,用现在精确的语言说,为了给出集合 S 的一个划分,只要给出 S 上的一个等价关系就可以了.

(四) 商集

定义 5 设 $\sim$ 是集合 S 上的一个等价关系，则 $\sim$ 伴随的 S 的划分 $\pi_\sim(S)=\{\bar{a}\mid a\in S\}$ 称为 S 对于关系 $\sim$ 的**商集**，记作 $S/\sim$.

注意 商集 $S/\sim$ 里的元素是等价类，因此商集 $S/\sim$ 不是 S 的子集.

例如，(一) 中 S 对于同系关系的商集是 $\{S_1,S_2\}$.

从命题 5.2.3 和定义 5 得出，集合 S 的任意一个划分 $\pi(S)$ 都是 S 的一个商集 $S/\sim$，其中 $\sim$ 是命题 5.2.3 的证明中所定义的.

譬如，集合 $S=\{1,2,3,4\}$，它的一个划分 $\{\{1\},\{2,3,4\}\}$ 是 S 的一个商集；S 的另一个划分 $\{\{1,2\},\{3,4\}\}$ 是 S 的又一个商集.

例 1 在整数集 Z 上定义一个二元关系 $\sim$ 如下：

$$a\sim b\Leftrightarrow a \text{ 与 } b \text{ 被 } 7 \text{ 除所得余数相同}$$

显然，$\sim$ 具有反身性，对称性，传递性. 因此 $\sim$ 是 Z 上的一个等价关系. Z 对于 $\sim$ 的商集 $Z/\sim$ 由 7 个等价类：$\bar{0},\bar{1},\bar{2},\bar{3},\bar{4},\bar{5},\bar{6}$ 组成，这是因为任意一整数被 7 除所得余数只有 7 种可能：0，1，2，3，4，5，6.

在数论里，"a 与 b 被 7 除所得余数相同" 称为 a 与 b 模 7 同余，记作 $a\equiv b(\mathrm{mod}7)$. 于是例 1 中的等价关系 $\sim$ 就是模 7 同余，等价类就是同余类. Z 对于模 7 同余的商集

$$Z/\sim=\{\bar{0},\bar{1},\bar{2},\bar{3},\bar{4},\bar{5},\bar{6}\}$$

有许多应用. 譬如，可以用 $\bar{0}$ 表示星期日，$\bar{1}$ 表示星期一，…，$\bar{6}$ 表示星期六. 这样我们排一个学期的课程表时，不用详细写出 9 月 1 日至 1 月 14 日的每一天的课程表，而只要写出星期一至星期五的课程表就够了.

习 题 5.2

1. 在平面 S 上定义一个二元关系 $\sim$ 如下：

$$P(x,y) \sim P'(x',y') \Leftrightarrow x - x' \in Z \text{ 且 } y - y' \in Z$$

证明：$\sim$ 是 S 上的一个等价关系.

2. 设 $S = \{a,b,c\}$，问：S 有多少种划分？S 有多少个不同的商集？

§3 矩阵的相抵分类

这一节我们来讨论数域 K 上的所有 $s \times n$ 矩阵组成的集合 $M_{s\times n}(K)$ 的一种划分. 为此我们先定义 $M_{s\times n}(K)$ 上的一个二元关系如下：

定义 1 设 $A,B \in M_{s\times n}(K)$，如果 A 可以经过有限次初等变换变成 B，则称 A 与 B 是**相抵的**(或**等价的**).

容易验证，相抵关系具有反身性，对称性与传递性. 因此，它是 $M_{s\times n}(K)$ 上的一个等价关系. 从而相抵等价类组成的集合是 $M_{s\times n}(K)$ 的一个划分.

由于矩阵的初等变换可以通过初等矩阵与矩阵的乘法实现，并且一个矩阵可逆的充分必要条件是它能表示成一些初等矩阵的乘积，因此我们得到

命题 5.3.1 数域 K 上两个 $s \times n$ 矩阵 A 与 B 相抵的充分必要条件为，存在数域 K 上 s 级可逆矩阵 P 与 n 级可逆矩阵 Q，使得

$$PAQ = B \tag{1}$$

▌

在每一个相抵的等价类里，我们希望选取一个代表具有最简单的形式，这个简单的矩阵是什麼样子？下面的定理 5.3.1 回答

了这个问题.

定理 5.3.1 设一个 $s\times n$ 矩阵 A 的秩为 $r(\neq 0)$,则 A 相抵于下述矩阵

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \tag{1}$$

它称为矩阵 A 的**相抵标准形**. 如果 $\operatorname{rank}(A)=0$,则 A 的相抵标准形是零矩阵.

证明 如果 $\operatorname{rank}(A)=0$,则 $A=0$,从而,它确定的相抵等价类里只有一个元素:0. 因此 A 的标准形是它自身.

现在设 $\operatorname{rank}(A)=r\neq 0$. 我们已经知道,$A$ 可以经过有限次初等行变换变成简化行阶梯形矩阵 J,它有 r 个非零行,从而有 r 个主元,并且主元都为 1,主元所在列的其余元素都为 0. 把 J 经过有限次两列互换可以变成形如下述的矩阵 J_1:

简化行阶梯形矩阵.

$$J_1=\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & c_{1,r+1} & \cdots & c_{1n} \\ 0 & 1 & 0 & \cdots & 0 & c_{2,r+1} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & c_{r,r+1} & \cdots & c_{rn} \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

把 J_1 的第1列的 $-c_{1,r+1}$ 倍加到第 $r+1$ 列,可以使所得矩阵的$(1,r+1)$元为 0,并且 J_1 的其余元素没有变化. 因此看出,只要把 J_1 的第 $1,2,\cdots,r$ 列的适当倍数分别加到第 $r+1,\cdots,n$ 列上,就可以变成形如下述的矩阵

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \tag{2}$$

所以 A 相抵于矩阵(2). ▌

定理 5.3.2 两个 $s \times n$ 矩阵 A 与 B 相抵的充分必要条件是它们有相同的秩.

证明 必要性. 设 A 与 B 相抵. 由于初等变换不改变矩阵的秩,因此 $\operatorname{rank}(A) = \operatorname{rank}(B)$.

充分性. 设 $\operatorname{rank}(A) = \operatorname{rank}(B) = r$. 若 $r = 0$,则 $A = B = 0$,它们相抵. 若 $r \neq 0$,则据定理 5.3.1 得,A 与 B 都相抵于矩阵(2),由对称性和传递性即得 A 与 B 相抵. ▌

定理 5.3.2 告诉我们,在相抵关系下,同一个等价类里的矩阵由于它们彼此相抵,从而有相同的秩; 秩相同的矩阵(由于它们相抵,从而)在同一个等价类里. 由此即得

推论 5.3.1 集合 $M_{s\times n}(K)$ 中,秩为 r 的所有矩阵恰好组成一个相抵等价类,其中 $0 \leqslant r \leqslant \min\{s,n\}$;从而 $M_{s\times n}(K)$ 一共有 $\min\{s,n\}+1$ 个相抵等价类. ▌

定义 2 设 $\sim$ 是集合 S 的一个等价关系,一种量或一种表达式如果对于同一等价类里的元素是相同的,则称这种量或这种表达式为集合 S 在等价关系 $\sim$ 下的一个**不变量**. 集合 S 在等价关系 $\sim$ 下的一组不变量,如果对于不同的等价类的元素是不完全相同的,而去掉其中任何一个,剩下的不变量就有可能对于不同的等价类的元素是相同的,那么称这组不变量是集合 S 在关系 $\sim$ 下的一组**完全不变量**.

由于同一个相抵等价类里的矩阵有相同的秩,因此矩阵的秩是集合 $M_{s\times n}(K)$ 在相抵关系下的一个不变量. 又由于秩相同的矩阵在同一个相抵等价类里,因此矩阵的秩是 $M_{s\times n}(K)$ 在相抵关系下的完全不变量.

从命题 5.3.1 和定理 5.3.1 可得到

推论 5.3.2 设 $s \times n$ 矩阵 A 的秩为 $r(\neq 0)$,则存在 s 级可逆矩阵 P 与 n 级可逆矩阵 Q,使得

$$A = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q \tag{3}$$

把矩阵 A 表示成(3)式是非常有用的.

习 题 5.3

1. 设 $s\times n$ 矩阵 A 的秩为 $r(\neq 0)$,证明:存在 $s\times r$ 列满秩矩阵 P_1 与 $r\times n$ 行满秩矩阵 Q_1,使得 $A=P_1Q_1$.

2. 设 A,B 分别是 $s\times n,n\times m$ 矩阵,证明:

$$\mathrm{rank}(AB)\geqslant \mathrm{rank}(A)+\mathrm{rank}(B)-n.$$

3. 证明:任意一个秩为 $r(\neq 0)$ 的矩阵都可以表示成 r 个秩为 1 的矩阵之和.

§4 广义逆矩阵

本节介绍矩阵的相抵标准形的一个应用.

我们知道,线性方程组 $AX=\beta$,如果它的系数矩阵 A 可逆(也就是,$|A|\neq 0$),那么它有唯一解,并且这个解可以用公式

$$X=A^{-1}\beta \tag{1}$$

表示. 现在要问:如果 A 不可逆,但是 $AX=\beta$ 有解,那么它的解能否也用类似于(1)的简洁公式表达?为此首先要把逆矩阵的概念推广,使得每个矩阵都有广义逆矩阵. 广义逆矩阵有许多应用,特别是在数理统计和计算数学等方面.

为了把逆矩阵的概念加以推广,我们先看一看逆矩阵的性质:若 A 可逆,则有 $AA^{-1}=I$. 两边右乘 A 得,$AA^{-1}A=A$. 这表明 A 可逆时,矩阵方程 $AXA=A$ 有解:$X=A^{-1}$. 我们进一步指出,A 可逆时,矩阵方程 $AXA=A$ 的解是唯一的,这是因为如果 $X=C$ 是解,则 $ACA=A$. 两边左乘 A^{-1},右乘 A^{-1} 得 $C=A^{-1}$. 综上述,当 A 可逆时,矩阵方程 $AXA=A$ 有唯一解,这个解就是 A^{-1}.

现在我们来看对任意 $s\times n$ 矩阵 A，矩阵方程 $AXA=A$ 是否有解？如果它有解的话，我们就可以把它的解称为 A 的广义逆矩阵.

定理 5.4.1 设 A 是数域 K 上一个 $s\times n$ 矩阵，则矩阵方程

$$AXA=A \tag{2}$$

总是有解. 如果 $\operatorname{rank}(A)=r$，并且

$$A=P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q \tag{3}$$

其中 P 与 Q 分别是 s 级、n 级可逆矩阵，则矩阵方程(2) 的一般解(通解) 为

$$X=Q^{-1}\begin{pmatrix} I_r & B \\ C & D \end{pmatrix}P^{-1} \tag{4}$$

其中 B,C,D 分别是任意 $r\times(s-r),(n-r)\times r,(n-r)\times(s-r)$ 矩阵.

证明 把形如(4)的矩阵以及(3)式代入矩阵方程(2)得

$$\begin{aligned}
\text{左边}&=P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}QQ^{-1}\begin{pmatrix} I_r & B \\ C & D \end{pmatrix}P^{-1}P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q \\
&=P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & B \\ C & D \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q \\
&=P\begin{pmatrix} I_r & B \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q \\
&=P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q \\
&=A=\text{右边}
\end{aligned}$$

所以形如(4)的每一个矩阵都是矩阵方程(2)的解.

为了说明(4)是矩阵方程(2)的通解，现在任取(2)的一个解 $X=G$，则由(2) 和(3) 得

$$P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}QGP\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q=P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q$$

因为 P,Q 可逆，所以从上式得

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} QGP \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \tag{5}$$

把矩阵 QGP 分块，设

$$QGP = \begin{pmatrix} H & B \\ C & D \end{pmatrix} \tag{6}$$

代入(5)式得

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} H & B \\ C & D \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

即

$$\begin{pmatrix} H & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

由此得出，$H = I_r$. 代入(6) 式便得出

$$G = Q^{-1} \begin{pmatrix} I_r & B \\ C & D \end{pmatrix} P^{-1}$$

这证明了矩阵方程(2)的任意一个解都能表示成(4)的形式，所以公式(4)是矩阵方程(2)的通解. ▌

定义 1 设 A 是一个 $s \times n$ 矩阵，矩阵方程 $AXA = A$ 的通解称为 A 的**广义逆矩阵**，简称为 A 的**广义逆**. 我们用记号 A^- 表示 A 的一个广义逆.

从定义 1 得出

$$AA^- A = A$$

应当注意的是，从公式(4)看出，矩阵方程 $AXA = A$ 的解一般来说是不唯一的，因为当 $r < \min\{s,n\}$ 时，公式(4) 中的 B,C,D 可任意取. 只有当 $r = s = n$ 时，也就是 A 可逆时，矩阵方程(2) 的解才是唯一的. 因此当 A 不是可逆矩阵时，A 的广义逆不唯一.

现在用矩阵的广义逆来讨论线性方程组的解.

一个线性方程组有解时，称它是**相容的**；否则称它是**不相容的**.

定理 5.4.2 （非齐次线性方程组的相容性定理） 非齐次线性方程组 $AX=\beta$ 有解的充分必要条件是 $\beta=AA^{-}\beta$.

证明 必要性. 设 $AX=\beta$ 有解 $X=\alpha$, 则 $A\alpha=\beta$. 因为 $A=AA^{-}A$, 所以

$$\beta=A\alpha=AA^{-}A\alpha=AA^{-}\beta$$

充分性. 设 $\beta=AA^{-}\beta$, 则取 $\alpha=A^{-}\beta$ 得

$$A\alpha=A(A^{-}\beta)=\beta$$

所以 $\alpha=A^{-}\beta$ 是 $AX=\beta$ 的解. ▌

定理 5.4.3 （非齐次线性方程组解的结构定理） 设非齐次线性方程组 $AX=\beta$ 有解, 则它的一般解(通解)为

$$\boxed{X=A^{-}\beta} \tag{7}$$

其中 A^{-} 是 A 的任意一个广义逆.

证明 任取 A 的一个广义逆 A^{-}, 我们来证 $X=A^{-}\beta$ 是方程组 $AX=\beta$ 的解: 已知 $AX=\beta$ 有解, 据定理 5.4.2 得, $\beta=AA^{-}\beta=A(A^{-}\beta)$, 这表明 $A^{-}\beta$ 是 $AX=\beta$ 的一个解.

反之, 对于 $AX=\beta$ 的任意一个解 γ, 我们要证存在 A 的一个广义逆 A^{-}, 使得 $\gamma=A^{-}\beta$. 设 A 是 $s\times n$ 矩阵, 它的秩为 r, 且

$$A=P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q \tag{8}$$

其中 P,Q 分别是 s 级、n 级可逆矩阵. 由于 A 的广义逆具有形式 (4), 因此我们要找矩阵 B,C,D, 使

$$\gamma=Q^{-1}\begin{pmatrix} I_r & B \\ C & D \end{pmatrix}P^{-1}\beta \tag{9}$$

即

$$Q\gamma=\begin{pmatrix} I_r & B \\ C & D \end{pmatrix}P^{-1}\beta \tag{10}$$

先分析 $Q\gamma$ 与 $P^{-1}\beta$ 之间的关系. 由已知, $A\gamma=\beta$, 因此我们有

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q\gamma=P^{-1}\beta \tag{11}$$

分别把 $Q\gamma, P^{-1}\beta$ 分块，设

$$Q\gamma = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \begin{matrix} \}r \quad 行 \\ \}n-r行 \end{matrix}$$

$$P^{-1}\beta = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \begin{matrix} \}r \quad 行 \\ \}s-r行 \end{matrix}$$

则(11)式成为

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \tag{12}$$

由此得

$$\begin{pmatrix} Y_1 \\ 0 \end{pmatrix} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$$

所以 $Y_1 = Z_1, Z_2 = 0$，因为 $\beta \neq 0$，所以 $P^{-1}\beta \neq 0$，从而 $Z_1 \neq 0$. 设 $Z'_1 = (k_1, \cdots, k_r)$，且设 $k_i \neq 0$.

取 $B = 0, D = 0, C = (0, \cdots, 0, k_i^{-1}Y_2, 0, \cdots, 0)$，则

$$\begin{pmatrix} I_r & 0 \\ C & 0 \end{pmatrix} P^{-1}\beta = \begin{pmatrix} I_r & 0 \\ C & 0 \end{pmatrix} \begin{pmatrix} Z_1 \\ 0 \end{pmatrix} = \begin{bmatrix} Z_1 \\ CZ_1 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = Q\gamma \tag{13}$$

于是

$$\gamma = Q^{-1} \begin{pmatrix} I_r & 0 \\ C & 0 \end{pmatrix} P^{-1}\beta$$

从而只要取

$$A^- = Q^{-1} \begin{pmatrix} I_r & 0 \\ C & 0 \end{pmatrix} P^{-1}$$

则 $\gamma = A^-\beta$. ▌

从定理 5.4.3 看出，任意非齐次线性方程组 $AX = \beta$ 有解时，其通解的形式为 $X = A^-\beta$，这与 A 可逆时方程组 $AX = \beta$ 的解的形式 $X = A^{-1}\beta$ 很相似.

下面讨论齐次线性方程组的解的结构：

定理 5.4.4(齐次线性方程组解的结构定理) 数域 K 上 n 元齐次线性方程组 $AX=0$ 的通解为

$$X=(I_n-A^-A)Z \tag{14}$$

其中 A^- 是 A 的任意给定的一个广义逆,Z 取遍 K^n 中任意列向量.

证明 任取 $Z\in K^n$,我们有

$$A[(I_n-A^-A)Z]=(A-AA^-A)Z=0Z=0$$

所以 $X=(I_n-A^-A)Z$ 是方程组 $AX=0$ 的解.

反之,设 η 是方程组 $AX=0$ 的解,要证存在 $Z\in K^n$,使得 $\eta=(I_n-A^-A)Z$. 取 $Z=\eta$,我们有

$$(I_n-A^-A)\eta=\eta-A^-A\eta=\eta-A^-(A\eta)=\eta-0=\eta$$

所以(14)是方程组 $AX=0$ 的通解. ▌

利用定理 5.4.4,我们可以得到非齐次线性方程组的另一种形式的通解.

推论 5.4.1 设数域 K 上 n 元非齐次线性方程组 $AX=\beta$ 有解,则它的通解为

$$X=A^-\beta+(I_n-A^-A)Z \tag{15}$$

其中 A^- 是 A 的任意给定的一个广义逆,Z 取遍 K^n 中任意列向量.

证明 我们已经知道 $A^-\beta$ 是非齐次线性方程组 $AX=\beta$ 的一个解,又知道 $(I_n-A^-A)Z$ 是导出组 $AX=0$ 的通解,所以 $X=A^-\beta+(I_n-A^-A)Z$ 是 $AX=\beta$ 的通解. ▌

我们已指出,一个矩阵 A 的广义逆一般来说不唯一,这是因为矩阵方程 $AXA=A$ 的解一般来说不唯一. 有时我们希望矩阵 A 的某种类型的广义逆是唯一的,这就需要再添上一些矩阵方程,然后规定 A 的某种类型的广义逆是这若干个矩阵方程组成的方程组

的解. 下面就是这样一种做法：

定义 2　设 A 是一个 $s\times n$ 复矩阵，下述矩阵方程组

$$\begin{cases} AXA = A & (16) \\ XAX = X & (17) \\ (\overline{AX})' = AX & (18) \\ (\overline{XA})' = XA & (19) \end{cases}$$

称为关于 A 的 Penrose **方程组**，它的解称为 A 的 Moore-Penrose **广义逆**，记作 A^{+}.

下面以及习题 5.4 第 3 题，我们将证明，对任意复矩阵 A，Penrose 方程组总是有解，并且解唯一. 因此，A 的 Moore-Penrose 广义逆存在且唯一.

定理 5.4.5　设 A 是一个 $s\times n$ 非零复矩阵，关于 A 的 Penrose 方程组总是有解，并且它的解唯一. 设矩阵 $A = BC$，其中 B，C 分别是列满秩与行满秩矩阵，则 Penrose 方程组的唯一解是

$$X = \overline{C}'(C\overline{C}')^{-1}(\overline{B}'B)^{-1}\overline{B}' \tag{20}$$

* **证明**　把(20)式代入 Penrose 方程组的每一个方程，验证每一个方程都变成恒等式：

$$AXA = (BC)\overline{C}'(C\overline{C}')^{-1}(\overline{B}'B)^{-1}\overline{B}'(BC) = BC = A$$

$$XAX = \overline{C}'(C\overline{C}')^{-1}(\overline{B}'B)^{-1}\overline{B}'(BC)\overline{C}'(C\overline{C}')^{-1}(\overline{B}'B)^{-1}\overline{B}'$$

$$= \overline{C}'(C\overline{C}')^{-1}(\overline{B}'B)^{-1}\overline{B}' = X$$

$$(\overline{AX})' = (\overline{B}\,\overline{C}C'(\overline{C}C')^{-1}(B'\overline{B})^{-1}B')'$$

$$= B(\overline{B}'B)^{-1}\overline{B}' = B(C\overline{C}')(C\overline{C}')^{-1}(\overline{B}'B)^{-1}\overline{B}' = AX$$

$$(\overline{XA})' = (C'(\overline{C}C')^{-1}(B'\overline{B})^{-1}B'\overline{B}\,\overline{C})'$$

$$= \overline{C}'(C\overline{C}')^{-1}C = \overline{C}'(C\overline{C}')^{-1}(\overline{B}'B)^{-1}(\overline{B}'B)C = XA$$

所以(20)式的确是 Penrose 方程组的解.

下面证解的唯一性. 设矩阵 X_1 和 X_2 都是 Penrose 方程组的解，则

$$
\begin{aligned}
X_1 &\overset{(17)}{=} X_1AX_1 \overset{(16)}{=} X_1(AX_2A)X_1 = X_1(AX_2)(AX_1) \\
&\overset{(18)}{=} X_1(\overline{AX_2})'(\overline{AX_1})' = X_1(\overline{AX_1AX_2})' \\
&= X_1(\overline{AX_1A}\,\overline{X_2})' \overset{(16)}{=} X_1(\overline{A}\,\overline{X_2})' = X_1(\overline{AX_2})' \\
&\overset{(18)}{=} X_1AX_2 \overset{(16)}{=} X_1(AX_2A)X_2 = (X_1A)(X_2A)X_2 \\
&\overset{(19)}{=} (\overline{X_1A})'(\overline{X_2A})'X_2 = (\overline{X_2AX_1A})'X_2 \overset{(16)}{=} (\overline{X_2A})'X_2 \\
&\overset{(19)}{=} (X_2A)X_2 \overset{(17)}{=} X_2
\end{aligned}
$$

这证明了 Penrose 方程组的解的唯一性. ▌

请读者思考：公式(20)中的矩阵 $C\overline{C}'$ 与 $\overline{B}'B$ 为什么是可逆矩阵?(提示:用 Binet-Cauchy 公式.)

习 题 5.4

1. 设 A 与 B 分别是 $s\times n, s\times m$ 矩阵,证明:矩阵方程 $AX=B$ 有解的充分必要条件是

$$B = AA^{-}B$$

在有解时,它的通解为

$$X = A^{-}B + (I_n - A^{-}A)W$$

其中 A^{-} 是 A 的任意取定的一个广义逆,W 是任意 $n\times m$ 矩阵.

*2. 设 A 是 $s\times n$ 复矩阵,证明:

$$(A^{+})^{+} = A$$

$$(\overline{A^{+}})' = (\overline{A})'^{+}$$

$$(kA)^{+} = k^{+}A^{+}, \text{其中 } k \text{ 是复数}$$

*3. 说明零矩阵的 Moore-Penrose 广义逆是零矩阵自身. 从而数 0 的 Moore-Penrose 广义逆为数 0.

*4. 说明非零复数 k 的 Moore-Penrose 广义逆等于 k^{-1}.

*5. 说明可逆复矩阵 A 的 Moore-Penrose 广义逆等于 A^{-1}.

§5 矩阵的相似分类导引

在§3中我们讨论了集合 $M_{s\times n}(K)$ 在相抵关系下的分类. 设 $A,B\in M_{s\times n}(K)$,则 A 与 B 相抵的充分必要条件是存在数域 K 上的 s 级可逆矩阵 P 与 n 级可逆矩阵 Q,使得 $PAQ=B$.

这一节我们讨论集合 $M_n(K)$ 上的另一个二元关系:相似关系. 讨论这种关系的动力来自向量空间的同一个线性变换在不同基下的矩阵之间的关系,这在第九章中将详细阐述. 方阵的相似关系是非常有用的.

定义1 设 A 与 B 都是数域 K 上的 n 级矩阵,如果存在数域 K 上的一个 n 级可逆矩阵 U,使得

$$U^{-1}AU=B \tag{1}$$

则称 A 与 B 是**相似的**(或者称 A 相似于 B),记作 $A\sim B$.

从定义看出,相似比相抵的条件要强. 若 A 与 B 相似,则 A 与 B 一定相抵;反之不对.

相似是集合 $M_n(K)$ 的一个二元关系,容易验证它具有反身性,对称性与传递性,因此它是一个等价关系. 相似关系具有传递性的证明如下:

设 $A\sim B$ 且 $B\sim C$. 据定义,分别有可逆矩阵 U_1 与 U_2 使得: $U_1{}^{-1}AU_1=B, U_2{}^{-1}AU_2=C$. 由于 U_1U_2 仍为可逆矩阵,并且有

$$(U_1U_2)^{-1}A(U_1U_2)=U_2{}^{-1}(U_1{}^{-1}AU_1)U_2=U_2{}^{-1}BU_2=C$$

因此 $A\sim C$.

集合 $M_n(K)$ 在相似关系下的等价类组成 $M_n(K)$ 的一个划分. 我们需要讨论两个基本问题:1° 每一个相似等价类里具有最简单形式的矩阵是什么?这个最简单形式的矩阵称为这一个相似等价类的标准形,简称为相似标准形;2° 哪些量是相似关系下的不变量?哪些量组成相似关系的一组完全不变量?现在我们将开始讨论

这两个问题，但是它们的完全解决需要留待第十章进行.

矩阵的相似对于矩阵的运算有下面的性质：

命题5.5.1　如果 $B_1=U^{-1}A_1U$ 并且 $B_2=U^{-1}A_2U$，那么

$$B_1+B_2=U^{-1}(A_1+A_2)U \tag{2}$$

$$B_1B_2=U^{-1}(A_1A_2)U \tag{3}$$

$$B_1^m=U^{-1}A_1^mU,\text{其中 } m \text{ 是非负整数} \tag{4}$$

证明　(2)式是显然的. 下面证(3)式：

$$\begin{aligned}B_1B_2&=(U^{-1}A_1U)(U^{-1}A_2U)=U^{-1}A_1(UU^{-1})A_2U\\&=U^{-1}A_1IA_2U=U^{-1}A_1A_2U\end{aligned}$$

利用数学归纳法，从(3)式可以推出(4)式成立. ▍

值得注意的是，命题5.5.1的条件比"$A_1\sim B_1$ 且 $A_2\sim B_2$"更强.

同一个相似等价类里的矩阵有许多共同的性质，现在先列举一些，以后还会继续指出.

1° 相似的矩阵其行列式的值相同.

证明　设 $A\sim B$，则存在可逆矩阵 U，使得 $U^{-1}AU=B$. 从而 $|B|=|U^{-1}AU|=|U|^{-1}|A||U|=|A|$. ▍

2° 相似的矩阵或者都可逆，或者都不可逆；并且当它们可逆时，它们的逆矩阵也相似.

证明　由性质1°即得结论的前半部分. 现在设 $A\sim B$ 且 A 可逆，则有 $B=U^{-1}AU$. 因此即得

$$B^{-1}=(U^{-1}AU)^{-1}=U^{-1}A^{-1}U$$

所以 $A^{-1}\sim B^{-1}$. ▍

3° 相似的矩阵有相同的秩.

证明　设 $A\sim B$，则有可逆矩阵 U，使得 $U^{-1}AU=B$. 这表明 A 与 B 相抵. 从而 $\mathrm{rank}(A)=\mathrm{rank}(B)$. ▍

注意：由于相抵的矩阵不一定相似，因此秩相同的矩阵不一定相似.

4° 相似的矩阵有相同的迹.

证明 设 $A \sim B$,则 $U^{-1}AU = B$. 于是

$$Tr(B) = Tr(U^{-1}AU) = Tr(U^{-1}(AU)) = Tr((AU)U^{-1})$$
$$= Tr(A(UU^{-1})) = Tr(A) \qquad \blacksquare$$

在上述第三个等式中利用了第四章§5讲的迹函数的性质4°.

以上性质说明:矩阵的行列式,矩阵的秩,矩阵的迹都是相似关系下的不变量. 请读者举例说明行列式相同的 n 级矩阵不一定相似;秩相同的 n 级矩阵不一定相似;迹相同的 n 级矩阵也不一定相似.(提示:考虑 $n = 2$,易于举例).

现在来讨论相似标准形的问题.

由于对一切 n 级可逆矩阵 U 都有 $U^{-1}IU = I$,因此单位矩阵 I 确定的相似等价类里只有一个元素:I.

由于对一切可逆矩阵 U 都有 $U^{-1}(kI)U = kI$,因此给定一个数量矩阵 kI,它确定的相似等价类里只有一个元素:kI.

于是一个非数量矩阵不可能相似于一个数量矩阵. 非数量矩阵中最简单的矩阵是对角矩阵,自然要问:什么样的矩阵能够相似于对角矩阵呢?设 A 是数域 K 上一个 n 级矩阵,设 $D = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$,其中 $\lambda_i \in K$, $i = 1, \cdots, n$.

$$A \sim D = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$$

$\Leftrightarrow$有数域 K 上可逆矩阵 $U = (\alpha_1, \alpha_2, \cdots, \alpha_n)$,使得

$$U^{-1}AU = D$$

即 $\quad AU = UD$

即 $\quad A(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)D$

即 $\quad (A\alpha_1, A\alpha_2, \cdots, A\alpha_n) = (\lambda_1\alpha_1, \lambda_2\alpha_2, \cdots, \lambda_n\alpha_n)$

$\Leftrightarrow$有 K^n 中 n 个线性无关的列向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$,使得

$$A\alpha_i = \lambda_i\alpha_i, \quad i = 1, 2, \cdots, n$$

于是我们证明了

定理5.5.1 数域 K 上 n 级矩阵 A 能够相似于对角矩阵的充分必要条件是:K^n中有 n 个线性无关的列向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$,以及 K 中有 n 个数 $\lambda_1, \lambda_2, \cdots, \lambda_n$(可以相同) 满足

$$A\alpha_i = \lambda_i\alpha_i, \quad i = 1,2,\cdots,n \tag{5}$$

这时，令 $U = (\alpha_1,\alpha_2,\cdots,\alpha_n)$，则

$$U^{-1}AU = \mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\} \tag{6}$$

∎

如果一个 n 级矩阵 A 能够相似于对角矩阵 D，则称 A **可对角化**，把对角矩阵 D 称为 A 的**相似标准形**. 上述定理 5.5.1 给出了 n 级矩阵 A 可对角化的充分必要条件，从中看出满足(5) 式的数 λ_i 以及列向量 α_i 起着关键作用. 我们在下一节就来研究这样的数 λ_i 和这样的列向量 α_i.

习　题　5.5

1. 证明：若 $A \sim B$，则 $kA \sim kB$，$A' \sim B'$.

2. 证明：若 A 可逆，则 $AB \sim BA$，其中 B 是任意与 A 同级的矩阵.

3. 证明：若 $A_1 \sim B_1$ 且 $A_2 \sim B_2$，则 $\mathrm{diag}\{A_1,A_2\} \sim \mathrm{diag}\{B_1,B_2\}$.

4. 证明：若 A 与 B 可交换，则对任一可逆矩阵 U，有 $U^{-1}AU$ 与 $U^{-1}BU$ 也可交换.

5. 设 $f(x) = a_0 + a_1x + \cdots + a_mx^m$ 是数域 K 上的一元多项式，设 A 是数域 K 上的 n 级矩阵，定义

$$f(A) := a_0I + a_1A + \cdots + a_mA^m$$

$f(A)$ 是数域 K 上一个 n 级矩阵，称 $f(A)$ 是**矩阵 A 的多项式**. 证明：若 $A \sim B$，则 $f(A) \sim f(B)$.

6. 方阵 A 称为**幂零矩阵**，如果 A 的某个正整数次幂等于零矩阵，使 $A^l = 0$ 成立的最小正整数 l 称为 A 的**幂零指数**. 证明：与幂零矩阵相似的矩阵仍然是幂零矩阵，并且它们的幂零指数相同.

7. 方阵 A 称为**幂等矩阵**，如果 $A^2 = A$. 证明：与幂等矩阵相似的矩阵仍然是幂等矩阵.

8. 方阵 A 称为**对合矩阵**，如果 $A^2 = I$. 证明：与对合矩阵相似的矩阵仍是对合矩阵.

9. 方阵 A 称为**周期矩阵**,如果有正整数 m 使得 $A^m = I$. 使 $A^m = I$ 成立的最小正整数 m 称为 A 的**周期**. 证明:与周期矩阵相似的矩阵仍是周期矩阵,并且它们的周期相同.

10. 证明:若 A 可对角化,则 $A \sim A'$.

11. 设 n 级矩阵 B 是从 n 级矩阵 A 交换第 i 行与第 j 行,并且交换第 i 列与第 j 列得到的,证明:$A \sim B$,并且求可逆矩阵 U,使得 $B = U^{-1}AU$.

*12. 证明:

$$\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\} \sim \mathrm{diag}\{\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_n}\}$$

其中 $i_1 i_2 \cdots i_n$ 是一个 n 元排列.(提示:用归纳法,并且利用第11题的结果.)

*13. 设 $i_1 i_2 \cdots i_n$ 是自然数 $1,2,\cdots,n$ 的一个排列,把 n 级单位矩阵 I_n 的第 $1,2,\cdots,n$ 行分别调到第 $i_1, i_2, \cdots, i_n$ 行得到的矩阵称为**置换矩阵**,它可以写成 $(\varepsilon_{i_1}, \varepsilon_{i_2} \cdots, \varepsilon_{i_n})$,从而它也可看成把 I_n 的第 $i_1, i_2, \cdots, i_n$ 列分别调到第 $1,2,\cdots,n$ 列得到的矩阵. 证明:置换矩阵 $(\varepsilon_{i_1}, \varepsilon_{i_2} \cdots, \varepsilon_{i_n})$ 可逆,并且它的逆矩阵是把 I_n 的第 $i_1, i_2, \cdots, i_n$ 行分别调到第 $1,2,\cdots,n$ 行得到的矩阵.

*14. 证明:在一个 n 级矩阵 A 的右边乘上置换矩阵 $(\varepsilon_{i_1}, \varepsilon_{i_2}, \cdots, \varepsilon_{i_n})$ 相当于把 A 的第 $i_1, i_2, \cdots, i_n$ 列分别调到第 $1,2,\cdots,n$ 列;在 A 的左边乘上置换矩阵 $(\varepsilon_{i_1}, \varepsilon_{i_2} \cdots, \varepsilon_{i_n})$ 相当于把 A 的第 $1,2,\cdots,n$ 行分别调到第 $i_1, i_2, \cdots, i_n$ 行.

*15. 设 $i_1 i_2 \cdots i_n$ 是 $1,2,\cdots,n$ 的一个排列,证明:矩阵 $A \sim B$,其中

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

$$B = \begin{pmatrix} a_{i_1, i_1} & a_{i_1, i_2} & \cdots & a_{i_1, i_n} \\ a_{i_2, i_1} & a_{i_2, i_2} & \cdots & a_{i_2, i_n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_n, i_1} & a_{i_n, i_2} & \cdots & a_{i_n, i_n} \end{pmatrix}$$

*16. 证明:如果方程 A 确定的相似等价类里只有一个元素,则 A 一定是数量矩阵.(提示:先证与所有 n 级可逆矩阵可交换的矩阵一定是数量矩阵.)

17. 设

$$J_0 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}_{n\times n}$$

证明：$J_0 \sim J_0'$.（提示：考虑矩阵 $U = (\varepsilon_n, \varepsilon_{n-1}, \cdots, \varepsilon_2, \varepsilon_1)$）.

*18. 设 $A \sim B$. 证明：使得 $B = U^{-1}AU$ 的所有可逆矩阵 U 的集合，可以用下述方法得到：将与 A 可交换的所有可逆矩阵的集合中的矩阵，右乘以一个任何具有性质 $B = U_0^{-1}AU_0$ 的矩阵 U_0 而得到.

19. 证明：如果数域 K 上的 n 级矩阵 A,B 满足：$AB - BA = A$，则 A 不可逆.

*20. 证明：如果数域 K 上的 2 级矩阵 A,B 满足：$AB - BA = A$，则 $A^2 = 0$.

§6 矩阵的特征值和特征向量

定义 1 设 A 是数域 K 上的一个 n 级矩阵，如果对于数域 K 中一个数 λ_0，存在 K^n 中的非零列向量 α，使得

$$A\alpha = \lambda_0\alpha \tag{1}$$

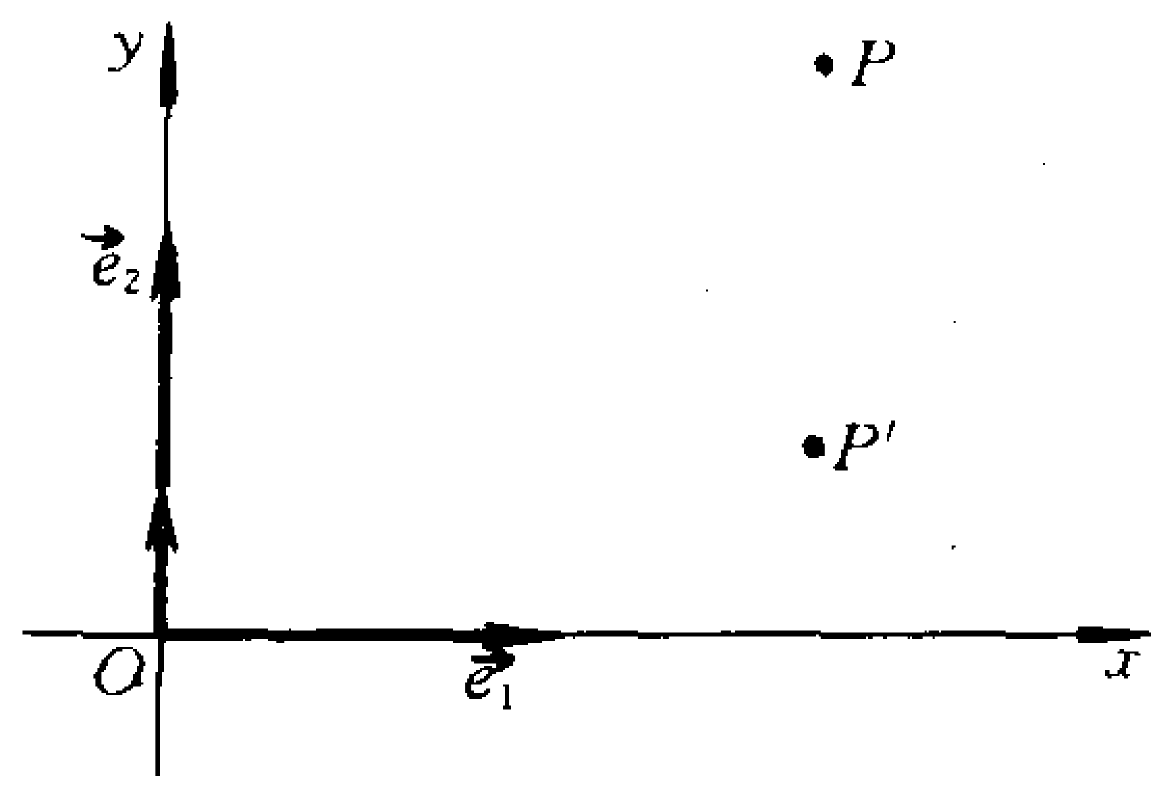
则 λ_0 称为矩阵 A 的一个**特征值**，α 称为 A 的属于特征值 λ_0 的一个**特征向量**.

如果 α 是矩阵 A 的属于特征值 λ_0 的特征向量，则 α 的任一非零数倍 $k\alpha$ 也是 A 的属于 λ_0 的特征向量，这是因为 $k\alpha \neq 0$，且

$$A(k\alpha) = k(A\alpha) = k(\lambda_0\alpha) = \lambda_0(k\alpha) \tag{2}$$

我们通过例子来看矩阵的特征值和特征向量的几何意义. 在平面上取一个直角坐标系 $[O;\vec{e}_1,\vec{e}_2]$. 考虑平面沿方向 $\vec{e}_2$ 或 $-\vec{e}_2$ 向

着 x 轴的压缩 τ，压缩系数为 $k(>0)$.



点 $P(x,y)$ 在 τ 下的象 P' 的坐标 (x',y') 为

$$\begin{cases} x' = x \\ y' = ky \end{cases} \tag{3}$$

即

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{4}$$

我们把(4)式右端的系数矩阵称为压缩 τ 的矩阵，记作 A. 从(4)式立即看出，原点 O 在 τ 下的象是点 O 自身. 从而向量 $\overrightarrow{OP}$ 在 τ 下的象是 $\overrightarrow{OP'}$. 由于 (x,y)，(x',y') 分别可看成是向量 $\overrightarrow{OP}$，$\overrightarrow{OP'}$ 的坐标，所以公式(4)也可看成是向量 $\overrightarrow{OP}$ 的坐标 (x,y) 与它在 τ 下的象 $\overrightarrow{OP'}$ 的坐标 (x',y') 之间的关系式. 从而有

$$\tau(\overrightarrow{OP}) = \overrightarrow{OP'} \Leftrightarrow A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{5}$$

从(5)得出，设 $\begin{pmatrix} x \\ y \end{pmatrix}$ 是向量 $\overrightarrow{OP}$ 的坐标，则

$$\begin{pmatrix} x \\ y \end{pmatrix} \text{是 } A \text{ 的属于特征值 } \lambda_0 \text{ 的特征向量}$$

$$\Leftrightarrow A\begin{pmatrix} x \\ y \end{pmatrix} = \lambda_0 \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \tau(\overrightarrow{OP}) = \lambda_0 \overrightarrow{OP}, \tag{6}$$

这说明，$\begin{pmatrix}x\\y\end{pmatrix}$是矩阵 A 的属于特征值 λ_0 的特征向量当且仅当以 $\begin{pmatrix}x\\y\end{pmatrix}$为坐标的向量$\overrightarrow{OP}$ 在压缩 τ 下的象 $\tau(\overrightarrow{OP})$ 与$\overrightarrow{OP}$ 共线，且其系数就是 A 的特征值 λ_0，即，$\tau(\overrightarrow{OP}) = \lambda_0 \overrightarrow{OP}$. 从这几何意义立即得出，由于 $\tau(\vec{e}_1) = \vec{e}_1, \tau(\vec{e}_2) = k\vec{e}_2$，因此 1 和 k 都是 A 的特征值，$\begin{pmatrix}1\\0\end{pmatrix}$是 A 的属于特征值 1 的一个特征向量，$\begin{pmatrix}0\\1\end{pmatrix}$是 A 的属于特征值 k 的一个特征向量.

现在考虑平面绕原点 O 转角为 θ 的旋转 σ，设点 P 的坐标为 (x,y)，点 P 在 τ 下的象 P' 的坐标为(x',y')，则

$$\begin{pmatrix}x'\\y'\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} \tag{7}$$

把(7)式右端的系数矩阵称为旋转 σ 的矩阵，记作 B. 由于原点 O 在 σ 下的象是点 O 自身，所以向量$\overrightarrow{OP}$ 在 σ 下的象是$\overrightarrow{OP'}$. 因此公式(7) 也可看成是向量$\overrightarrow{OP}$ 的坐标(x,y) 与它在 σ 下的象$\overrightarrow{OP'}$ 的坐标(x',y') 之间的关系式. 从而有

$$\sigma(\overrightarrow{OP}) = \overrightarrow{OP'} \Leftrightarrow B\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x'\\y'\end{pmatrix} \tag{8}$$

从(8)得出，

$$\begin{pmatrix}x\\y\end{pmatrix}\text{是 } B \text{ 的属于特征值 } \lambda_0 \text{ 的特征向量}$$

$$\Leftrightarrow B\begin{pmatrix}x\\y\end{pmatrix} = \lambda_0\begin{pmatrix}x\\y\end{pmatrix} \Leftrightarrow \sigma(\overrightarrow{OP}) = \lambda_0 \overrightarrow{OP}$$

这说明，$\begin{pmatrix}x\\y\end{pmatrix}$是矩阵 B 的属于特征值 λ_0 的特征向量当且仅当以 $\begin{pmatrix}x\\y\end{pmatrix}$为坐标的向量$\overrightarrow{OP}$ 在旋转 σ 下的象 $\sigma(\overrightarrow{OP})$ 与$\overrightarrow{OP}$ 共线，且其系数就是 B 的特征值 λ_0，即 $\sigma(\overrightarrow{OP}) = \lambda_0 \overrightarrow{OP}$. 从这几何意义立即得出，如果 $\theta \neq k\pi (k \in Z)$，则矩阵 B 没有特征向量(因为当 $\theta \neq k\pi$ 时，

任何一个向量$\overrightarrow{OP}$在旋转σ下的象不可能与$\overrightarrow{OP}$共线),从而B也就没有特征值.

任给数域K上的一个n级矩阵A,如何知道它是否有特征值和特征向量?如果有的话,怎样求出A的全部特征值和特征向量?为了回答这个问题,先引进一个概念:

定义 2 设$A=(a_{ij})$是数域K上一个n级矩阵,λ是一个变量.矩阵$\lambda I-A$的行列式

$$|\lambda I-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & -a_{23} & \cdots & -a_{2n}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & \lambda-a_{nn}\end{vmatrix}$$

称为A的**特征多项式**.

注:λ的精确含义将在第七章 §1 给出.矩阵$\lambda I-A$的主对角线上的元素不是数域K中的数,而是λ的一次多项式$\lambda-a_{ii}$,$(i=1,\cdots,n)$.对于这种类型的矩阵可以象数域K上的矩阵那样定义它的行列式,并且具有第二章讲的行列式的一切性质.由行列式的定义容易看出,$|\lambda I-A|$是λ的n次多项式.由行列式的定义还可以看出,对于$\lambda_0\in K$,矩阵$\lambda_0 I-A$的行列式$|\lambda_0 I-A|$正好等于多项式$|\lambda I-A|$在λ_0处的值.

现在我们来讨论如何求数域K上一个n级矩阵A的特征值和特征向量:

λ_0是A的一个特征值,α是A的属于λ_0的一个特征向量

$\Leftrightarrow A\alpha=\lambda_0\alpha,\lambda_0\in K,\alpha\in K^n,\alpha\neq 0$

$\Leftrightarrow(\lambda_0 I-A)\alpha=0,\lambda_0\in K,\alpha\in K^n,\alpha\neq 0$

$\Leftrightarrow\alpha$是齐次线性方程组$(\lambda_0 I-A)X=0$的一个非零解,$\lambda_0\in K$

$\Leftrightarrow|\lambda_0 I-A|=0,\lambda_0\in K,\alpha$是$(\lambda_0 I-A)X=0$的一个非零解

$\Leftrightarrow\lambda_0$是A的特征多项式$|\lambda I-A|$在数域K中的一个根,α是齐次线性方程组$(\lambda_0 I-A)X=0$的一个非零解

以上推导过程证明了

定理 5.6.1　设 A 是数域 K 上一个 n 级矩阵,则

(1) λ_0 是 A 的一个特征值的充分必要条件为 λ_0 是 A 的特征多项式在数域 K 中的一个根;

(2) α 是 A 的属于特征值 λ_0 的一个特征向量的充分必要条件为 α 是齐次线性方程组 $(\lambda_0 I - A)X = 0$ 的一个非零解,并且写成列向量的形式. ▌

因此,如果数域 K 上的 n 级矩阵 A 的特征多项式 $|\lambda I - A|$ 在数域 K 中没有根,则 A 没有特征值. 否则,A 有特征值,此时,A 的特征多项式 $|\lambda I - A|$ 在数域 K 中的全部根就是 A 的全部特征值. 若 λ_0 是 A 的一个特征值,则齐次线性方程组 $(\lambda_0 I - A)X = 0$ 的解空间中的所有非零列向量就是 A 的属于特征值 λ_0 的全部特征向量. 这也表明:A 的属于特征值 λ_0 的全部特征向量添上零向量所成的集合是 K^n 的一个线性子空间,称它是 A 的属于特征值 λ_0 的**特征子空间**,它就是齐次线性方程组 $(\lambda_0 I - A)X = 0$ 的解空间.

于是求数域 K 上的一个 n 级矩阵 A 的全部特征值和全部特征向量的方法是:

第一步,计算行列式 $|\lambda I - A|$;

第二步,求出多项式 $f(\lambda) = |\lambda I - A|$ 在数域 K 中的全部根,它们就是 A 的全部特征值;

第三步,对于 A 的每一个特征值 λ_i,解齐次线性方程组 $(\lambda_i I - A)X = 0$,求出它的一个基础解系 $\eta_1, \cdots, \eta_t$,并且写成列向量的形式,它们就是 A 的属于特征值 λ_0 的极大线性无关特征向量组. 于是 A 的属于 λ_0 的全部特征向量是

$$k_1\eta_1 + k_2\eta_2 + \cdots + k_t\eta_t$$

其中 $k_1, k_2, \cdots, k_t$ 是数域 K 中的不全为零的任意数.

例 1　设

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

是一个实矩阵，求 A 的全部特征值和全部特征向量.

解 A 的特征多项式为

$$|\lambda I - A| = \begin{vmatrix} \lambda-2 & 2 & -2 \\ 2 & \lambda+1 & -4 \\ -2 & -4 & \lambda+1 \end{vmatrix} = \begin{vmatrix} \lambda-2 & 2 & -2 \\ 2 & \lambda+1 & -4 \\ 0 & \lambda-3 & \lambda-3 \end{vmatrix}$$

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

$$= (\lambda-3)(\lambda^2+3\lambda-18) = (\lambda-3)^2(\lambda+6)$$

所以 A 的特征值是 3(二重) 与 -6.

对于特征值 3，解齐次线性方程组 $(3I-A)X=0$：

$$\begin{pmatrix} 1 & 2 & -2 \\ 2 & 4 & -4 \\ -2 & -4 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

从而该方程组的一般解为：$x_1 = -2x_2 + 2x_3$，于是它的一个基础解系是

$$\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

所以 A 的属于特征值 3 的全部特征向量是

$$k_1\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + k_2\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

其中 k_1, k_2 是不全为零的任意实数.

对于特征值 -6，解齐次线性方程组 $(-6I-A)X=0$，求得

它的一个基础解系为

$$\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$$

所以 A 的属于特征值 -6 的全部特征向量是

$$k\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$$

其中 k 是任意非零实数.

例 2　设

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

如果把 A 看成实数域上的矩阵,A 有没有特征值?如果把 A 看成复数域上的矩阵,求 A 的全部特征值和全部特征向量.

解　A 的特征多项式为

$$|\lambda I-A|=\begin{vmatrix} \lambda-1 & 1 \\ -1 & \lambda-1 \end{vmatrix}=\lambda^2-2\lambda+2$$

如果 A 是实矩阵,由于实系数多项式 $\lambda^2-2\lambda+2$ 没有实根,因此 A 没有特征值.如果 A 是复矩阵,由于多项式 $\lambda^2-2\lambda+2$ 有两个复根:$1+i$ 与 $1-i$,因此 A 的特征值是 $1+i$ 与 $1-i$.

对于特征值 $1+i$,解齐次线性方程组$((1+i)I-A)X=0$,求得它的一个基础解系为

$$\begin{pmatrix} i \\ 1 \end{pmatrix}$$

所以 A 的属于特征值 $1+i$ 的全部特征向量是

$$k\begin{pmatrix} i \\ 1 \end{pmatrix}$$

其中 k 是任意非零复数.

同理可求得 A 的属于特征值 $1-i$ 的全部特征向量是

$$l\begin{pmatrix} -i \\ 1 \end{pmatrix}$$

其中 l 是任意非零复数.

从例 2 看出，一个方阵 A 有没有特征值，是与把它看成哪个数域上的矩阵有关系的. 根据代数基本定理可知，复数域上方阵总有特征值.

方阵 A 称为**幂零矩阵**，如果 A 的某个正整数次幂等于零矩阵；使 $A^l = 0$ 成立的最小正整数 l 称为 A 的**幂零指数**.

例 3 证明：任意数域 K 上的幂零矩阵一定有特征值，并且它的特征值一定是零.

证明 设 A 是数域 K 上的 n 级幂零矩阵，其幂零指数为 l，则 $A^l = 0$. 于是 $|A|^l = 0$，因此 $|A| = 0$. 由此得出

$$|0I - A| = |-A| = (-1)^n |A| = 0$$

这表明 0 是 A 的特征多项式 $|\lambda I - A|$ 的一个根，因此 0 是 A 的一个特征值.

设 λ_0 是幂零矩阵 A 的一个特征值，则有 K^n 中非零列向量 α 使得 $A\alpha = \lambda_0\alpha$. 两边左乘 A 得，$A^2\alpha = A(\lambda_0\alpha) = \lambda_0(A\alpha) = \lambda_0(\lambda_0\alpha) = \lambda^2\alpha$. 继续这个过程，可得到 $A^l\alpha = \lambda_0{}^l\alpha$. 由于 $A^l = 0$，因此 $\lambda_0{}^l\alpha = 0$. 因为 $\alpha \neq 0$ 所以 $\lambda_0 = 0$. 这说明 A 的特征值一定是零. ▌

现在我们继续讨论同一个相似等价类里的矩阵的共同性质.

5° 相似的矩阵有相同的特征多项式.

证明 设 $A \sim B$，则有可逆矩阵 U，使得 $B = U^{-1}AU$. 于是

$$\begin{aligned}|\lambda I - B| &= |\lambda I - U^{-1}AU| = |U^{-1}(\lambda I)U - U^{-1}AU| \\ &= |U^{-1}(\lambda I - A)U| = |U^{-1}||\lambda I - A||U| \\ &= |\lambda I - A|\end{aligned}$$ ▌

性质 5°说明，矩阵的特征多项式是一个相似不变量. 但是反之，有相同特征多项式的 n 级矩阵不一定相似. 例如，

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

A 与 I 有相同的特征多项式$(\lambda-1)^2$，但是 A 与 I 不相似.

由性质 5°立即得到

6° 相似的矩阵有相同的特征值(包括重数相同). ▍

因此，矩阵的特征值也是相似不变量. 但是，由上面的例子可知，有相同特征值(包括重数相同)的 n 级矩阵不一定相似.

下面我们来求一个 n 级矩阵的特征多项式的各项系数.

命题 5.6.1 设 A 是数域 K 上的一个 n 级矩阵，则 A 的特征多项式 $|\lambda I-A|$ 是一个 n 次多项式，它的首项系数为 1，$n-1$ 次项的系数等于 $-Tr(A)$，常数项为$(-1)^n|A|$，$n-k$ 次项的系数为 A 的所有 k 级主子式的和乘以$(-1)^k$，$1\leqslant k<n$.

* **证明**

$$|\lambda I-A|=\begin{vmatrix} \lambda-a_{11} & 0-a_{12} & \cdots & 0-a_{1n} \\ 0-a_{21} & \lambda-a_{22} & \cdots & 0-a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0-a_{n1} & 0-a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}$$

利用行列式函数的线性性质，$|\lambda I-A|$ 可以拆成下列行列式的和：

$$\begin{vmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda \end{vmatrix}, \qquad \begin{vmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & -a_{nn} \end{vmatrix}$$

$$|(-\alpha_1,\cdots,-\alpha_{j_1-1},\lambda\varepsilon_{j_1},-\alpha_{j_1+1},\cdots,-\alpha_{j_{n-k}-1},\lambda\varepsilon_{j_{n-k}},\cdots,-\alpha_n)|$$

其中

$$(\alpha_1,\cdots,\alpha_n)=A,\quad 1\leqslant j_1<j_2<\cdots<j_{n-k}\leqslant n,\quad 1\leqslant k<n$$

上述第 1 个行列式等于 λ^n，第 2 个行列式等于$(-1)^n|A|$. 对于上述第 3 种类型的行列式，按第 $j_1,j_2,\cdots,j_{n-k}$ 列展开，这 $n-k$ 列元素组成的 $n-k$ 级子式只有一个不为 0，其余全为 0，这个不为 0 的 $n-k$ 级子式等于 λ^{n-k}，它的代数余子式为

$$(-1)^{(j_1+\cdots+j_{n-k})+(j_1+\cdots+j_{n-k})}(-A)\begin{pmatrix} j_1' & , & \cdots, & j_k' \\ j_1' & , & \cdots, & j_k' \end{pmatrix}$$

$$=(-1)^k A\begin{pmatrix} j_1' & , & \cdots, & j_k' \\ j_1' & , & \cdots, & j_k' \end{pmatrix}$$

其中 $\{j_1',\cdots,j_k'\}=\{1,2,\cdots,n\}\backslash\{j_1,\cdots,j_{n-k}\}$，并且 $j_1'<\cdots<j_k'$. 所以上述第 3 种类型的行列式的值等于

$$(-1)^k A\begin{pmatrix} j_1' & , & \cdots, & j_k' \\ j_1' & , & \cdots, & j_k' \end{pmatrix}\lambda^{n-k}$$

由于 $1\leqslant j_1'<j_2'<\cdots<j_k'\leqslant n$，因此 $|\lambda I-A|$ 的 $n-k$ 次项的系数为

$$(-1)^k \sum_{1\leqslant j_1'<\cdots<j_k'\leqslant n} A\begin{pmatrix} j_1' & , & \cdots, & j_k' \\ j_1' & , & \cdots, & j_k' \end{pmatrix}$$

其中 $1\leqslant k<n$. 特别地，当 $k=1$ 时，得到 $|\lambda I-A|$ 的 $n-1$ 次项的系数为

$$-(a_{11}+a_{22}+\cdots+a_{nn})=-Tr(A) \qquad \blacksquare$$

习　题　5.6

1. 求下列复数域上的矩阵的全部特征值和全部特征向量：

(1) $A=\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$　　(2) $A=\begin{pmatrix} 2 & 3 & 2 \\ 1 & 8 & 2 \\ -2 & -14 & -3 \end{pmatrix}$

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

(5) $A=\begin{pmatrix}1 & -1 & 2\\ 5 & -4 & 3\\ -1 & 0 & -3\end{pmatrix}$ (6) $A=\begin{pmatrix}4 & 7 & -3\\ -2 & -4 & 2\\ -4 & -10 & 4\end{pmatrix}$

(7) $A=\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\end{pmatrix}$.

(8) $A=\begin{pmatrix}3 & -2 & -4 & 3\\ 3 & -3 & -3 & 2\\ -3 & \frac{1}{2} & \frac{5}{2} & -\frac{5}{2}\\ -10 & 3 & 11 & -9\end{pmatrix}$.

2. 设 A 是复数域上的一个 n 级矩阵，并且 A 的元素全是实数. 若复数 λ_0 是 A 的一个特征值，α 是 A 的属于 λ_0 的一个特征向量，则 $\overline{\lambda_0}$ 也是 A 的特征值，并且 $\overline{\alpha}$ 是 A 的属于 $\overline{\lambda_0}$ 的一个特征向量.

3. 下列矩阵 A 如果看成实数域上的矩阵，它有没有特征值？如果看成复数域上矩阵，求 A 的特征值和特征向量.

(1) $A=\begin{pmatrix}1 & -\sqrt{3}\\ \sqrt{3} & 1\end{pmatrix}$ (2) $A=\begin{pmatrix}0 & a\\ -a & 0\end{pmatrix}$，$a$ 是实数

4. 证明：任一数域 K 上的幂等矩阵一定有特征值，并且它的特征值是 1 或者 0.

5. 证明：复数域上周期为 m 的周期矩阵的特征值都是 m 次单位根. （注：如果一个复数 z 满足 $z^m=1$，则称 z 是一个 m 次单位根.）

6. 证明：方阵 A 与 A' 有相同的特征多项式，从而它们有相同的特征值.

7. 设 A 是数域 K 上的一个可逆矩阵，证明：

(1) 如果 A 有特征值，则 A 的特征值不等于零；

(2) 若 λ_0 是 A 的一个特征值，则 λ_0^{-1} 是 A^{-1} 的一个特征值.

8. 证明：方阵 A 的行列式为 0 的充分必要条件是 A 有特征值 0.

9. 设 A 是一个 n 级正交矩阵，证明：

(1) 若 A 有特征值,则它的特征值是 1 或 -1;

(2) 若 n 是奇数,并且 $|A|=1$,则 1 是 A 的一个特征值;

(3) 若 $|A|=-1$,则 -1 是 A 的一个特征值.

10. 举例说明:一个正交矩阵可以没有特征值.

11. 设 λ_0 是数域 K 上的一个 n 级矩阵 A 的一个特征值,证明:

(1) $k\lambda_0$ 是矩阵 kA 的一个特征值,

(2) λ_0^m 是矩阵 A^m 的一个特征值,m 是正整数.

12. 设 $f(x)=a_0+a_1x+\cdots+a_mx^m$ 是数域 K 上的一个一元多项式,设 λ_0 是 K 上的一个 n 级矩阵 A 的一个特征值.证明:$f(\lambda_0)$ 是矩阵 $f(A)$ 的一个特征值.

13. 设 A 是数域 K 上一个 n 级矩阵,证明:如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 A 的属于同一个特征值 λ_0 的线性无关的特征向量,则对于数域 K 中任意一组不全为零的数 $k_1,k_2,\cdots,k_s$,有:$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s$ 也是 A 的属于特征值 λ_0 的特征向量.

*14. 证明:设 A 与 B 都是数域 K 上的 n 级矩阵,则 AB 与 BA 的特征多项式相同.(提示:用第四章 §8 命题 4.8.1.)

*15. 求复数域上 n 级循环矩阵

$$A=\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}$$

的全部特征值和全部特征向量.(提示:设 $\omega_k=\cos\frac{2k\pi}{n}+i\sin\frac{2k\pi}{n}$, $k=0,1,\cdots,n-1$. 设 $f(x)=a_1+a_2x+\cdots+a_nx^{n-1}$. 去计算

$$A\begin{pmatrix} 1 \\ w_k \\ \vdots \\ w_k^{n-1} \end{pmatrix},\quad k=0,1,\cdots,n-1$$

由此可得 A 的全部特征值和全部特征向量.)

§7 n 级矩阵可对角化的条件

在 §5 的定理 5.5.1 我们指出了一个 n 级矩阵可对角化的充

分必要条件，现在我们用特征值和特征向量的术语把这个条件重新叙述一遍：

定理 5.7.1 数域 K 上的一个 n 级矩阵 A 可对角化的充分必要条件是 A 有 n 个线性无关的特征向量.

证明

数域 K 上 n 级矩阵 A 可对角化

$\Leftrightarrow$有 K 上可逆矩阵 $U=(\alpha_1,\cdots,\alpha_n)$ 与 K 中数 $\lambda_1,\cdots,\lambda_n$ 使得 $U^{-1}AU=\mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}$，即

$$A(\alpha_1,\cdots,\alpha_n)=(\lambda_1\alpha_1,\cdots,\lambda_n\alpha_n)$$

$\Leftrightarrow$有 K^n 中 n 个线性无关的列向量 $\alpha_1,\cdots,\alpha_n$ 与 K 中数 $\lambda_1,\cdots,\lambda_n$，使得

$$A\alpha_i=\lambda_i\alpha_i,\quad i=1,\cdots,n$$

$\Leftrightarrow A$ 有 n 个线性无关的特征向量 $\alpha_1,\cdots,\alpha_n$ ∎

从定理 5.7.1 及其证明过程看出，如果一个 n 级矩阵 A 有 n 个线性无关的特征向量 $\alpha_1,\cdots,\alpha_n$，那么 A 可对角化，并且令 $U=(\alpha_1,\cdots,\alpha_n)$，则 $U^{-1}AU=\mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}$，其中 λ_i 是 α_i 所属的特征值，即 $A\alpha_i=\lambda_i\alpha_i,\quad i=1,\cdots,n.$

给了一个 n 级矩阵 A，怎样判别它有没有 n 个线性无关的特征向量呢？首先我们要求出 A 的所有不同的特征值：$\lambda_1,\cdots,\lambda_m$；然后对每个特征值 λ_i，求出齐次线性方程组 $(\lambda_iI-A)X=0$ 的一个基础解系：$\alpha_{i1},\cdots,\alpha_{ir_i}$；最后把它们合在一起：$\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{m1},\cdots,\alpha_{mr_m}$. 如果这个向量组仍线性无关，则它就是 A 的最大的线性无关特征向量组，这意思是：把 A 的任一特征向量 β 添进去得到的向量组 $\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{m1},\cdots,\alpha_{mr_m},\beta$ 一定线性相关，这是因为：设 β 是 A 的属于 λ_j 的特征向量，则 $\beta=k_1\alpha_{j1}+\cdots+k_{r_j}\alpha_{jr_j}$. 下面我们来证明：$\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{m1},\cdots,\alpha_{mr_m}$ 的确是线性无关的.

定理 5.7.2 设 λ_1,λ_2 是数域 K 上 n 级矩阵 A 的不同的特征值，并且 $\alpha_1,\cdots,\alpha_s$ 是 A 的属于特征值 λ_1 的线性无关的特征向量，

$\beta_1,\cdots,\beta_r$ 是 A 的属于特征值 λ_2 的线性无关的特征向量. 则 $\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_r$ 线性无关.

证明 设

$$k_1\alpha_1+\cdots+k_s\alpha_s+l_1\beta_1+\cdots+l_r\beta_r=0 \tag{1}$$

(1)式两边左乘 A 得

$$k_1A\alpha_1+\cdots+k_sA\alpha_s+l_1A\beta_1+\cdots+l_rA\beta_r=0$$

从而有

$$k_1\lambda_1\alpha_1+\cdots+k_s\lambda_1\alpha_s+l_1\lambda_2\beta_1+\cdots+l_r\lambda_2\beta_r=0 \tag{2}$$

(1)式两边乘以 λ_2 得

$$k_1\lambda_2\alpha_1+\cdots+k_s\lambda_2\alpha_s+l_1\lambda_2\beta_1+\cdots+l_r\lambda_2\beta_r=0 \tag{3}$$

(2)式减去(3)式得

$$k_1(\lambda_1-\lambda_2)\alpha_1+\cdots+k_s(\lambda_1-\lambda_2)\alpha_s=0 \tag{4}$$

因为 $\alpha_1,\cdots,\alpha_s$ 线性无关,从(4) 式得

$$k_1(\lambda_1-\lambda_2)=0,\cdots,k_s(\lambda_1-\lambda_2)=0$$

由于 $\lambda_1\neq\lambda_2$,因此从上述式子得 $k_1=0,\cdots,k_s=0$. 再代入(1)式得

$$l_1\beta_1+\cdots+l_r\beta_r=0$$

因为 $\beta_1,\cdots,\beta_r$ 线性无关,所以 $l_1=\cdots=l_r=0$. 这样我们证明了: $\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_r$ 线性无关. ▎

对 A 的特征值的个数作数学归纳法,我们可得到:

定理 5.7.3 设 $\lambda_1,\cdots,\lambda_m$ 是数域 K 上 n 级矩阵 A 的不同的特征值,并且 $\alpha_{i1},\cdots,\alpha_{ir_i}$ 是 A 的属于特征值 λ_i 的线性无关的特征向量,$i=1,\cdots,m$. 则向量组

$$\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{m1},\cdots,\alpha_{mr_m}$$

是线性无关的.

证明的细节留给读者.

推论 5.7.1 设 A 是数域 K 上的一个 n 级矩阵,则 A 的属于不同特征值的特征向量是线性无关的. ▎

从定理 5.7.3 得出，设 $\lambda_1,\cdots,\lambda_m$ 是数域 K 上 n 级矩阵 A 的全部不同的特征值，并且 $\alpha_{i1},\cdots,\alpha_{ir_i}$ 是齐次线性方程组 $(\lambda_i I - A)X = 0$ 的一个基础解系，$i = 1,\cdots,m$. 则 $\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{m1},\cdots,\alpha_{mr_m}$ 是 A 的最大的线性无关特征向量组. 由于 K^n 中任一线性无关向量组所含向量的数目不超过 n，因此 $r_1+\cdots+r_m \leqslant n$. 据定理 5.7.1，如果 $r_1+\cdots+r_m < n$，则 A 不能对角化；如果 $r_1+\cdots+r_m = n$，则 A 可对角化. 当 A 可对角化时，令

$$U = (\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{m1},\cdots,\alpha_{mr_m})$$

则

$$U^{-1}AU = \mathrm{diag}\{\underbrace{\lambda_1,\cdots,\lambda_1}_{r_1},\cdots,\underbrace{\lambda_m,\cdots,\lambda_m}_{r_m}\}$$

即，此时 A 的相似标准形是以 A 的全部特征值为主对角元的对角矩阵，每一个特征值 λ_i 在主对角线上出现的次数等于齐次线性方程组 $(\lambda_i I - A) = 0$ 的解空间的维数 r_i，也就是属于 λ_i 的特征子空间的维数. 因此这时 A 的相似标准形除了主对角线上元素的排列次序外是被矩阵 A 唯一决定的. 注意：由于齐次线性方程组的基础解系有无穷多个，因此可逆矩阵 U 的取法有无穷多种.

从上一段的议论我们立即得到

推论 5.7.2 数域 K 上的 n 级矩阵 A 可对角化的充分必要条件是 A 的属于不同特征值的特征子空间的维数之和等于 n. ▌

从定理 5.7.3 和定理 5.7.1 我们还可得到

推论 5.7.3 设 A 是数域 K 上的一个 n 级矩阵，如果 A 的特征多项式在 K 中有 n 个不同的根，那么 A 可对角化. ▌

例 1 判断 §6 的例 1 的实矩阵 A 能否对角化？如果可对角化，找出一个可逆矩阵 U，使 $U^{-1}AU$ 为对角矩阵.

解 §6 例 1 中 A 的特征值为 2(二重) 与 -7. 已经求出，对于特征值 2，齐次线性方程组 $(2I - A)X = 0$ 的一个基础解系是

$$\alpha_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \qquad \alpha_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

对于特征值 -7，$(-7I-A)X=0$ 的一个基础解系是

$$\alpha_3=\begin{pmatrix}1\\2\\-2\end{pmatrix}$$

于是 3 级矩阵 A 有 3 个线性无关的特征向量，所以 A 可对角化. 令

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

则 $U^{-1}AU=\mathrm{diag}\{2,2,-7\}$.

例 2 §6 中例 2 的矩阵 A 如果看成实数域上的矩阵，它能否对角化？如果把 A 看成复数域上的矩阵，它能否对角化？若可对角化，求出一个可逆矩阵 U，使 $U^{-1}AU$ 为对角矩阵.

解 §6 例 2 中已经指出，A 看成是实矩阵时，它没有特征值，从而 A 没有特征向量，所以实矩阵 A 不能对角化.

如果把 A 看成复矩阵，则 A 有两个不同的特征值：$1+i$ 与 $1-i$，因此复矩阵 A 可对角化. 从 §6 例 2 的结果知道，令

$$U=\begin{pmatrix}i&-i\\1&1\end{pmatrix}$$

则 $U^{-1}AU=\mathrm{diag}\{1+i,1-i\}$.

从例 2 看到，一个矩阵能否对角化与所考虑的数域有关. 但是，有的矩阵不能对角化并不是数域的原因，即使把它看成复数域上的矩阵，如果它的线性无关的特征向量的数目小于该矩阵的级数，则它不能对角化.

习　题　5.7

1. 习题 5.6 的第 1 题与第 3 题中哪些矩阵可对角化？哪些矩阵不能对

角化?如果 A 可对角化,求出一个可逆矩阵 U 使得 $U^{-1}AU$ 为对角矩阵.

2. 在有理数域、实数域、复数域中,判断下列矩阵是否可对角化? 如果可对角化,写出对角矩阵.

(1) $\begin{pmatrix} 5 & 2 & -3 \\ 4 & 5 & -4 \\ 6 & 4 & -4 \end{pmatrix}$ (2) $\begin{pmatrix} 8 & 15 & -36 \\ 8 & 21 & -46 \\ 5 & 12 & -27 \end{pmatrix}$

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

3. 证明:若 α 与 β 是 n 级矩阵 A 的属于不同特征值的特征向量,则 $\alpha+\beta$ 不是 A 的特征向量.

4. 设 A 是数域 K 上一个 n 级矩阵,证明:如果 K^n 中任意非零列向量都是 A 的特征向量,则 A 一定是数量矩阵.

5. 设 $A=(a_{ij})$ 是数域 K 上一个 n 级上三角矩阵,证明:

(1) 如果 $a_{11},a_{22},\cdots,a_{nn}$ 两两不同,则 A 可对角化;

(2) 如果 $a_{11}=a_{22}=\cdots=a_{nn}$,并且至少有一个 $a_{kl}\neq 0(k<l)$,则 A 一定不能对角化.

*6. 复数域上的 n 级循环矩阵 A(参见习题 5.6 第 15 题)是否可对角化?如果可以,求出一个可逆矩阵 U,使得 $U^{-1}AU$ 为对角矩阵.由此说明:复数域上所有 n 级循环矩阵可以用同一个可逆矩阵 U 相似于对角矩阵.

§8 矩阵的相似标准形的一些应用

这一节我们来介绍矩阵的相似标准形的一些应用.

矩阵的相似标准形可以用来讨论同一个相似等价类里的矩阵的共同性质.其想法是:如果我们想了解矩阵 A 的某一性质(在相似关系下不变的性质),那么我们可以先求出 A 的相似标准形 D.

由于矩阵 D 比较简单，它的性质容易研究，由此可了解矩阵 A 的性质. 下面举一个例子.

方阵 A 称为**幂等矩阵**，如果 $A^2 = A$. 我们想了解数域 K 上的幂等矩阵的秩与迹的关系. 由于矩阵的秩与迹都是相似关系下的不变量，因此我们可以先求出幂等矩阵 A 的相似标准形 D，从 D 容易看出它的秩与迹有什么关系，进而了解 A 的秩与迹的关系.

例 1 证明：数域 K 上的幂等矩阵一定可对角化，并且它的相似标准形是 $\mathrm{diag}\{I_r, 0\}$，其中 r 是该幂等矩阵的秩.

证明 设 A 是数域 K 上的一个 n 级幂等矩阵，它的秩为 r. 如果 $r = 0$，则 $A = 0$，结论显然成立. 如果 $r = n$，则 A 可逆，从而由 $A^2 = A$ 得出 $A = I$，结论也成立. 下面设 $0 < r < n$.

设 λ_0 是 A 的一个特征值，则有 K^n 中非零列向量 α，使得 $A\alpha = \lambda_0\alpha$. 两边左乘 A 得，$A^2\alpha = \lambda_0 A\alpha$. 由此得出，$A\alpha = \lambda_0\lambda_0\alpha$，即 $\lambda_0\alpha = \lambda_0^2\alpha$，亦即，$\lambda_0(\lambda_0 - 1)\alpha = 0$. 由于 $\alpha \neq 0$，因此 $\lambda_0 = 0$ 或 $\lambda_0 = 1$.

因为 $|0I - A| = |-A| = 0$，所以 0 是 A 的一个特征值. 齐次线性方程组 $(0I - A)X = 0$ 的解空间的维数为

$$n - \mathrm{rank}(-A) = n - \mathrm{rank}(A) = n - r$$

因为 $A^2 = A$，所以 $A(I - A) = 0$. 于是

$$\mathrm{rank}(A) + \mathrm{rank}(I - A) \leqslant n.$$

又有

$$\begin{aligned} n &= \mathrm{rank}(I) = \mathrm{rank}(A + (I - A)) \\ &\leqslant \mathrm{rank}(A) + \mathrm{rank}(I - A) \end{aligned}$$

因此得到

$$\mathrm{rank}(A) + \mathrm{rank}(I - A) = n$$

从而得到

$$\mathrm{rank}(I - A) = n - \mathrm{rank}(A) = n - r$$

由于 $r > 0$，所以 $\mathrm{rank}(I - A) = n - r < n$. 从而 $|I - A| = 0$. 由此得出，1 是 A 的一个特征值. 齐次线性方程组 $(I - A)X = 0$ 的解空间的维数等于

$$n - \mathrm{rank}(I - A) = n - (n - r) = r$$

综上述，A 的属于不同特征值的特征子空间的维数之和等于 $(n-r)+r=n$,因此 A 可对角化. A 的相似标准形中，特征值 1 在主对角线上出现的次数等于相应的特征子空间的维数 r;特征值 0 在主对角线上出现的次数等于相应的特征子空间的维数 $n-r$. 于是 A 的相似标准形为 $\mathrm{diag}\{I_r,0\}$. ▌

例 2 证明：数域 K 上幂等矩阵的秩等于它的迹.

证明 设 A 是数域 K 上的一个 n 级幂等矩阵，它的秩为 r. 如果 $r=0$,则 $\mathrm{rank}(A)=0=Tr(A)$. 如果 $r=n$,则 $A=I$. 从而 $\mathrm{rank}(A)=n=Tr(A)$. 下面设 $0<r<n$. 据例 1,A 相似于 $\mathrm{diag}\{I_r,0\}$. 于是

$$\mathrm{rank}(A) = r = Tr(\mathrm{diag}\{I_r,0\}) = Tr(A) \qquad ▌$$

例 2 给出了求数域 K 上幂等矩阵 A 的秩的简便方法：只要把 A 的主对角元相加即得 $\mathrm{rank}(A)$.

矩阵的相似标准形还可以用来简化矩阵的方幂的计算. 其理论根据是：如果 $A \sim B$,则存在可逆矩阵 U,使

$$U^{-1}AU = B$$

由此得出

$$U^{-1}A^mU = B^m$$

从而有

$$A^m = U\,B^m\,U^{-1} \tag{1}$$

因此，如果 B 的形式比较简单，那么 B^m 就比较容易计算，从(1)式就可以算出 A^m,这比直接计算 A^m 要简便得多.

在许多实际问题中都会遇到矩阵的方幂，需要计算它们. 我们来举两个例子.

例 3 著名的**斐波那契**(Fibonacci)**数列**是

$$0,1,1,2,3,5,8,13,\cdots$$

它满足下述递归公式：

$$f_{k+2} = f_{k+1} + f_k, \quad k = 0,1,2,\cdots \tag{2}$$

以及初始条件：$f_0=0, f_1=1$.（Fibonacci 数列与叫做叶序（叶子生在茎上的次序）的植物现象有关.）试求 Fibonacci 数列的通项公式，并且求出$\lim\limits_{k\to\infty}\dfrac{f_k}{f_{k+1}}$.

解 令

$$U_k=\begin{pmatrix} f_{k+1} \\ f_k \end{pmatrix}, \quad k=0,1,2,\cdots$$

因为 $f_{k+2}=f_{k+1}+f_k$，所以

$$\begin{pmatrix} f_{k+2} \\ f_{k+1} \end{pmatrix}=\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} f_{k+1} \\ f_k \end{pmatrix} \tag{3}$$

令

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

则(3)式成为

$$U_{k+1}=AU_k \tag{4}$$

由(4)式得出

$$U_k=A^kU_0 \tag{5}$$

于是为了求 Fibonacci 数列的通项公式就只要去计算 A^k，我们利用 A 的相似标准形来简化 A^k 的计算. 我们把 A 看成实数域上的矩阵.

A 的特征多项式为 $|\lambda I-A|=\lambda^2-\lambda-1$，它的两个根是：$\lambda_1=\frac{1}{2}(1+\sqrt{5})$，$\lambda_2=\frac{1}{2}(1-\sqrt{5})$，这些是 A 的特征值. 由此看出 A 可对角化. 解齐次线性方程组

$$\left(\frac{1}{2}(1+\sqrt{5})I-A\right)X=0$$

得到它的一个基础解系：

$$\alpha_1=\begin{pmatrix} \frac{1}{2}(1+\sqrt{5}) \\ 1 \end{pmatrix}=\begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}$$

同理可得

$$\left(\frac{1}{2}(1-\sqrt{5})I-A\right)X=0$$

的一个基础解系是

$$\alpha_2=\begin{pmatrix}\frac{1}{2}(1-\sqrt{5})\\ 1\end{pmatrix}=\begin{pmatrix}\lambda_2\\ 1\end{pmatrix}$$

令

$$U=\begin{pmatrix}\lambda_1 & \lambda_2\\ 1 & 1\end{pmatrix}$$

则

$$U^{-1}AU=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$$

从而

$$A^k=U\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}^k U^{-1}=\begin{pmatrix}\lambda_1 & \lambda_2\\ 1 & 1\end{pmatrix}\begin{pmatrix}\lambda_1^k & 0\\ 0 & \lambda_2^k\end{pmatrix}\frac{1}{\sqrt{5}}\begin{pmatrix}1 & -\lambda_2\\ -1 & \lambda_1\end{pmatrix}$$

$$=\frac{1}{\sqrt{5}}\begin{pmatrix}\lambda_1^{k+1} & \lambda_2^{k+1}\\ \lambda_1^k & \lambda_2^k\end{pmatrix}\begin{pmatrix}1 & -\lambda_2\\ -1 & \lambda_1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} \tag{6}$$

从(5)式及初始条件得

$$\begin{pmatrix}f_{k+1}\\ f_k\end{pmatrix}=A^k\begin{pmatrix}1\\ 0\end{pmatrix} \tag{7}$$

$$\begin{pmatrix}1\\ -1\end{pmatrix}$$

$(\lambda$

比较(7)式两边的第 2 个分量得

$$f_k=\frac{1}{\sqrt{5}}(\lambda_1^k-\lambda_2^k)=\frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^k-\left(\frac{1-\sqrt{5}}{2}\right)^k\right] \tag{8}$$

这就是 Fibonacci 数列的通项公式. 容易算出:

$$\lim_{k\to\infty}\frac{f_k}{f_{k+1}}=\frac{1}{\lambda_1}=\frac{\sqrt{5}-1}{2}\approx 0.618 \tag{9}$$

* 注: $\frac{f_k}{f_{k+1}}$ 当 $k\to\infty$ 时的极限, 即 $\frac{\sqrt{5}-1}{2}\approx 0.618$, 在最优化

方法中有重要应用. 一些实际问题常常可归结为求目标函数 $y=f(x)$ 在区间$[a,b]$上的最大值(或最小值),其中 $y=f(x)$ 的解析表达式并不知道. 假定 $y=f(x)$ 在$[a,b]$上只有一个极值点(否则可将区间$[a,b]$划分),这时称 $y=f(x)$ 是单峰函数. 为了求单峰函数 $y=f(x)$ 在$[a,b]$上的最大值点?可以在区间$[a,b]$的若干点上做试验求出函数值,再比较函数值的大小. 如何选取这些试验点,使得所做试验次数比较少,却能迅速找出最大值点,可采用下述方法选取试验点:

第一个试验点 $t_1=a+0.618(b-a)$,第二个试验点 $t_1'=a+0.382(b-a)$,即 t_1' 是点 t_1 关于区间$[a,b]$中点的对称点,比较 $f(t_1)$ 与 $f(t_1')$,如果 $f(t_1)>f(t_1')$,则由于 $y=f(x)$ 是单峰函数,因此,最大值点不可能出现在区间$[a,t_1']$里,从而可以去掉$[a,t_1']$,剩下区间$[t_1',b]$. 第三个试验点 $t_2=t_1'+0.618(b-t_1')$,第四个试验点 $t_2'=t_1'+0.382(b-t_1')$. 比较 $f(t_2)$ 与 $f(t_2')$,如果 $f(t_2)<f(t_2')$,则去掉区间$[t_2,b]$,剩下区间$[t_1',t_2]$. 依次进行下去,当剩下的区间长度比指定的正数 ε 小时,就取剩下区间的中点作为所要求的点,称它为最优点(与真正的最大值点很接近的点).

上述这种方法称为 0.618 **法**,也称为**黄金分割法**. 它的优点是可以迅速缩短搜索区间,以便找出最优点.

* **例** 4　考察某地区居民的色盲遗传情况.

每一个人都有 46 个染色体. 染色体是成对的,有 22 对是常染色体,一对是性染色体. 男性的一对性染色体是(X, Y);女性的一对性染色体是(X, X). 基因位于染色体上,因此基因也是成对的. 在一对染色体的某一点位上的一对基因称为两个等位基因. 显性的基因用 A 表示,隐性的基因用 a 表示. 色盲基因是隐性的,且只位于 X 染色体上. 一个女性居民如果她的一对性染色体的某一点位 P 上的两个等位基因是 X^aX^A(包括 X^AX^a 这一情形,以下同)或 X^aX^a,则她患色盲,其中 X^a 表示色盲基因. 如果她的点位 P 上的

两个等位基因是 X^AX^A，则她不患色盲. 设 N 个女性居民中有 N_1 个人的点位 P 上的两个等位基因是 X^AX^A，N_2 个人的点位 P 上的两个等位基因是 X^AX^a，N_3 个人的点位 P 上的两个等位基因是 X^aX^a. 则这 N 个女性居民中色盲基因的频率为

$$\frac{N_2+2N_3}{2N}=\frac{N_2}{2N}+\frac{N_3}{N} \tag{10}$$

令

$$r=\frac{N_1}{N},\qquad 2s=\frac{N_2}{N},\qquad t=\frac{N_3}{N} \tag{11}$$

则 $r,2s,t$ 为这 N 个女性居民中点位 P 上的等位基因分别为 X^AX^A,X^AX^a,X^aX^a 的人所占的比例，这些比例记成 $(r,2s,t)$. 显然有 $r+2s+t=1$. 用这些记号，则这 N 个女性居民中色盲基因的频率为 $s+t$.

类似地，一个男性居民如果他的一对性染色体的某一点位 P 上的两个等位基因是 X^aY，则他患色盲；如果他的点位 P 上的两个等位基因是 X^AY，则他不患色盲. 设 M 个男性居民中有 M_1 个人的点位 P 上的两个等位基因是 X^AY，M_2 个人的点位 P 上的两个等位基因是 X^aY，则这 M 个男性居民中色盲基因的频率为 $\frac{M_2}{M}$.

令

$$p=\frac{M_1}{M},\qquad q=\frac{M_2}{M} \tag{12}$$

则这 M 个男性居民中色盲基因的频率为 q. 这里 p,q 为这 M 个男性居民中点位 P 上的等位基因分别为 X^AY,X^aY 的人所占的比例，这些比例记成 (p,q). 显然有 $p+q=1$. 由此可见，男性居民的色盲基因频率等于男性色盲者的比例 q.

现在设某地区第一代男性居民中，点位 P 上的等位基因分别为 X^AY,X^aY 的人所占的比例为 (p,q)；女性居民中点位 P 上的等位基因分别为 X^AX^A,X^AX^a,X^aX^a 的人所占的比例为 $(r,2s,t)$. 则

第一代男性居民，女性居民的色盲基因频率分别为 $q, s+t$. 我们来求该地区第二代男性居民，女性居民的色盲基因频率. 这里假设第一代男性居民与女性居民的结合是随机的. 设第二代男性居民共有 L 人，其中具有等位基因 X^AY 的人，由于他的基因 X^A 来自母亲，而第一代女性居民中，基因 X^A 的频率为

$$\frac{2N_1+N_2}{2N}=r+s \tag{13}$$

因此具有等位基因 X^AY 的人的数目为 $L(r+s)$. 同理，具有等位基因 X^aY 的人的数目为 $L(s+t)$. 因此第二代男性居民中色盲基因的频率（它等于男性色盲者的比例）为

$$\frac{L(s+t)}{L}=s+t \tag{14}$$

由此看出，第二代男性居民中色盲基因的频率等于第一代女性居民中色盲基因的频率.

设第二代女性居民共有 W 人，其中具有等位基因 X^AX^A 的人的数目为 $Wp(r+s)$，具有等位基因 X^AX^a 的人的数目为 $W[p(s+t)+(r+s)q]$，具有等位基因 X^aX^a 的人的数目为 $Wq(s+t)$. 由此得出，第二代女性居民色盲基因的频率为

$$\begin{aligned}
&\frac{W[p(s+t)+(r+s)q]+2Wq(s+t)}{2W}\\
&=\frac{1}{2}[p(s+t)+(r+s)q]+q(s+t)\\
&=\frac{1}{2}[(s+t)+q(r+s)+q(s+t)]\\
&=\frac{1}{2}[(s+t)+q(r+2s+t)]=\frac{1}{2}(s+t+q)
\end{aligned} \tag{15}$$

由(15)式看出，第二代女性居民中色盲基因的频率等于第一代男性居民和女性居民的色盲基因频率的算术平均数.

我们用 b_i, c_i 分别表示该地区第 i 代男性居民和女性居民的色盲基因频率，由上述知

$$\begin{cases} b_i = c_{i-1} \\ c_i = \dfrac{1}{2}(b_{i-1} + c_{i-1}) \end{cases} \tag{16}$$

其中 $i = 2,3,\cdots$. 如果知道了 b_1,c_1,我们来求 b_n,c_n. 从(16)式得

$$\begin{pmatrix} b_i \\ c_i \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix} \begin{pmatrix} b_{i-1} \\ c_{i-1} \end{pmatrix} \tag{17}$$

把(17)式右端的系数矩阵记作 B. 从(17)式容易得出

$$\begin{pmatrix} b_n \\ c_n \end{pmatrix} = B^{n-1} \begin{pmatrix} b_1 \\ c_1 \end{pmatrix} \tag{18}$$

由此可见,求 b_n,c_n 归结为求出 B^{n-1}. 为此我们来求 B 的相似标准形. B 的特征多项式为

$$|\lambda I - B| = (\lambda - 1)(\lambda + \frac{1}{2})$$

因此 B 的特征值是 $1,-\frac{1}{2}$. 由此看出,B 可对角化. 解齐次线性方程组 $(1 \cdot I - B)X = 0$,得到它的一个基础解系:$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

解齐次线性方程组 $(-\frac{1}{2}I - B)X = 0$,得到它的一个基础解系:$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$. 令

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

则

$$U^{-1}BU = \begin{pmatrix} 1 & 0 \\ 0 & -\dfrac{1}{2} \end{pmatrix}$$

从而

$$B^{n-1} = U \begin{pmatrix} 1 & 0 \\ 0 & -\dfrac{1}{2} \end{pmatrix}^{n-1} U^{-1}$$

$$=\begin{pmatrix}1 & -2\\ 1 & 1\end{pmatrix}\begin{bmatrix}1 & 0\\ 0 & (-\frac{1}{2})^{n-1}\end{bmatrix}\frac{1}{3}\begin{pmatrix}1 & 2\\ -1 & 1\end{pmatrix}$$

$$=\frac{1}{3}\begin{bmatrix}1-(-\frac{1}{2})^{n-2} & 2+(-\frac{1}{2})^{n-2}\\ 1-(-\frac{1}{2})^{n-1} & 2+(-\frac{1}{2})^{n-1}\end{bmatrix} \tag{19}$$

因此

$$\begin{cases} b_n=\frac{1}{3}\left[1-(-\frac{1}{2})^{n-2}\right]b_1+\frac{1}{3}\left[2+(-\frac{1}{2})^{n-2}\right]c_1\\ c_n=\frac{1}{3}\left[1-(-\frac{1}{2})^{n-1}\right]b_1+\frac{1}{3}\left[2+(-\frac{1}{2})^{n-1}\right]c_1\end{cases} \tag{20}$$

从(20)式得

$$\lim_{n\to\infty}b_n=\lim_{n\to\infty}c_n=\frac{1}{3}b_1+\frac{2}{3}c_1 \tag{21}$$

这说明,尽管第一代男性居民、女性居民的色盲基因频率可能不相同,但是经过好几代(每一代都是随机结合)之后,两个性别的居民的色盲基因频率将接近相等.

矩阵的相似标准形在几何上也有重要应用.我们举一个例子.

设二次曲面在直角坐标系 I 中的方程为

$$x^2+4y^2+z^2-4xy-8xz-4yz-1=0 \tag{22}$$

试问:这是什么样的二次曲面?

解决这个问题的方法是:作直角坐标变换,使得在直角坐标系 II 中,这个二次曲面的新方程不含交叉项(xy 项,xz 项,yz 项),那么就可看出它是什么二次曲面.设直角坐标变换公式为

$$\underset{\mathrm{I}}{\begin{pmatrix}x\\ y\\ z\end{pmatrix}}=T\underset{\mathrm{II}}{\begin{pmatrix}x^*\\ y^*\\ z^*\end{pmatrix}} \tag{23}$$

其中 T 是直角坐标系 I 到直角坐标系 II 的过渡矩阵,在解析几何中我们已证明了 T 一定是正交矩阵.(22) 左端的二次项部分可以

写成

$$x^2+4y^2+z^2-4xy-8xz-4yz$$

$$=(x,y,z)\underbrace{\begin{bmatrix}1&-2&-4\\-2&4&-2\\-4&-2&1\end{bmatrix}}_{A}\begin{bmatrix}x\\y\\z\end{bmatrix} \tag{24}$$

把(24)式右端的 3 级矩阵记作 A. 用公式(23) 代入(24) 式，则(24) 变成

$$(x^*,y^*,z^*)T'AT\begin{bmatrix}x^*\\y^*\\z^*\end{bmatrix} \tag{25}$$

T′ 转置.

为了使(25)式中不出现交叉项，只要使矩阵 $T'AT$ 为对角矩阵，也就是使矩阵 $T^{-1}AT$ 为对角矩阵(因为 $T'=T^{-1}$). 由此可见，为了判断方程(22)表示什么二次曲面，关键是要找一个正交矩阵 T，使 $T^{-1}AT$ 为对角矩阵. 注意到 A 是实对称矩阵，因此问题在于：实对称矩阵 A 能不能对角化？并且能不能找到正交矩阵 T，使 $T^{-1}AT$ 为对角矩阵？我们将在下一节回答这个问题.

本节的后半部分，我们通过例题来介绍求矩阵的特征值和特征向量的一些方法和技巧，这些例题的结论也是有用的.

例 5　设 A,B 分别是数域 K 上的 $n\times m, m\times n$ 矩阵，证明：AB 与 BA 有相同的非零特征值.

证法一　设 $\lambda_0\neq 0$ 是 AB 的一个特征值，则有 K^n 中的一个非零列向量 α，使得

$$(AB)\alpha=\lambda_0\alpha \tag{26}$$

(26)式两边左乘 B 得

$$(BA)(B\alpha)=\lambda_0(B\alpha) \tag{27}$$

假如 $B\alpha=0$，则由(26)式得 $\lambda_0\alpha=A(B\alpha)=0$. 由于 $\alpha\neq 0$，于是推出 $\lambda_0=0$，矛盾. 所以 $B\alpha\neq 0$. (27)式说明 λ_0 是 BA 的一个特征值，并且 $B\alpha$ 是 BA 的属于 λ_0 的一个特征向量.

同理，BA 的任一非零特征值也是 AB 的特征值. ∎

证法二 设 $\lambda_0 \neq 0$ 是 AB 的一个特征值，则 λ_0 是 AB 的特征多项式 $|\lambda I_n - AB|$ 在 K 中的一个根. 据习题 4.8 的第 9 题得

$$|\lambda_0 I_n - AB| = \lambda_0^{n-m} |\lambda_0 I_m - BA|$$

于是由 $|\lambda_0 I_n - AB| = 0$ 得出 $|\lambda_0 I_m - BA| = 0$，这说明 λ_0 是 BA 的特征多项式 $|\lambda I_m - BA|$ 在 K 中的一个根. 因此 λ_0 是 BA 的一个特征值.

同理，BA 的任一非零特征值也是 AB 的特征值. ∎

注 1 例 5 的证法一的好处是：不仅知道 λ_0 也是 BA 的特征值，而且知道 $B\alpha$ 是 BA 的属于 λ_0 的一个特征向量，其中 α 是 AB 的属于特征值 λ_0 的一个特征向量.

注 2 在第七章我们将指出，λ_0 是多项式 $f(\lambda)$ 的一个根的充分必要条件为 $\lambda - \lambda_0$ 是 $f(\lambda)$ 的一个因式. 如果 $f(\lambda)$ 含因式 $(\lambda - \lambda_0)$ 共 l 个，则称 $\lambda - \lambda_0$ 是 $f(\lambda)$ 的 l 重因式，此时称 λ_0 是 $f(\lambda)$ 的 l 重根. 我们指出，用 Cramer 法则可证明，数域 K 上两个 n 次多项式 $f(\lambda)$ 和 $g(\lambda)$，如果它们对 K 中 $n+1$ 个不同的数 $a_1, a_2, \cdots, a_{n+1}$ 有相同的值，即 $f(a_i) = g(a_i)$，$\quad i = 1, 2, \cdots, n+1$，则 $f(\lambda) = g(\lambda)$. 运用这些结论，我们从例 1 的证法二所依据的习题 4.8 的第 9 题出发，可得到更强的结论：AB 与 BA 不仅有相同的非零特征值，而且 $\lambda_0 \neq 0$ 作为 AB 的特征值的重数与 λ_0 作为 BA 的特征值的重数也相同. 理由如下：

不妨设 $n \geqslant m$，据习题 4.8 的第 9 题得，对于 $a \in K$，只要 $a \neq 0$，就有 $|aI_n - AB| = a^{n-m}|aI_m - BA|$. 因此多项式 $|\lambda I_n - AB|$ 与多项式 $\lambda^{n-m}|\lambda I_m - BA|$ 相等. 设 $\lambda_0 \neq 0$ 是 AB 的 l 重特征值，则 $|\lambda I_n - AB|$ 含因式 $(\lambda - \lambda_0)$ 恰好 l 个，由于 $|\lambda I_n - AB| = \lambda^{n-m}|\lambda I_m - BA|$，所以 $|\lambda I_m - BA|$ 也含因式 $(\lambda - \lambda_0)$ 恰好 l 个，这证明了 λ_0 是 BA 的 l 重特征值.

对于 $n < m$ 的情形可类似证得结论.

例 6 求有理数域上 n 级矩阵 J 的特征值和特征向量；J 是否

可对角化?如果可对角化,求出一个可逆矩阵U,使$U^{-1}JU$为对角矩阵.

$$J=\begin{pmatrix}1 & 1 & \cdots & 1\\ 1 & 1 & \cdots & 1\\ \cdots & \cdots & \cdots & \cdots\\ 1 & 1 & \cdots & 1\end{pmatrix}$$

解法一 用§6介绍的求矩阵的特征值和特征向量的基本方法. 留给读者完成.

解法二 我们用1_n表示K^n中分量全为1的列向量,则$J=1_n1_n'$. 利用例5的结论,J与$1_n'1_n=(n)$有相同的非零特征值. 1级矩阵(n)的特征值只有一个:n,并且是1重特征值;属于特征值n的线性无关的特征向量只有一个:$\alpha=(1)$. 因此J的非零特征值只有一个:n,并且它是1重特征值;J的属于特征值n的线性无关的特征向量只有一个:$1_n\alpha=1_n(1)=1_n$.

因为$|J|=0$,所以0是J的一个特征值. 显然J的秩是1,所以齐次线性方程组$(0I-J)X=0$的解空间的维数为$n-1$. 容易求出该方程组的一般解公式为$x_1=-x_2-\cdots-x_n$,其中$x_2,\cdots,x_n$是自由未知量. 从而该方程组的一个基础解系为

$$\begin{pmatrix}1\\ -1\\ 0\\ 0\\ \vdots\\ 0\end{pmatrix},\quad \begin{pmatrix}1\\ 0\\ -1\\ 0\\ \vdots\\ 0\end{pmatrix},\quad \cdots,\quad \begin{pmatrix}1\\ 0\\ 0\\ \vdots\\ 0\\ -1\end{pmatrix}$$

这些是J的属于特征值0的线性无关的特征向量.

由于J有n个线性无关的特征向量,所以J可对角化. 令

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

则 $U^{-1}JU=\mathrm{diag}\{n,0,\cdots,0\}$.

* **例** 7　设 A 是数域 K 上的一个 n 级矩阵. 证明：A 可对角化的充分必要条件是：A 的特征多项式 $|\lambda I-A|$ 在复数域中的全部根都属于 K（从而每个根都是 A 的特征值），并且每一个根 λ_i 的重数等于 A 的属于特征值 λ_i 的特征子空间的维数.

证明　必要性. 设 A 可对角化，则

$$A\sim\mathrm{diag}\{\underbrace{\lambda_1,\cdots,\lambda_1}_{r_1},\cdots,\underbrace{\lambda_m,\cdots,\lambda_m}_{r_m}\}$$

其中 $\lambda_1,\cdots,\lambda_m$ 是 A 的全部不同的特征值，r_i 是 A 的属于特征值 λ_i 的特征子空间的维数. 因为相似矩阵有相同的特征多项式，所以

$$|\lambda I-A|=(\lambda-\lambda_1)^{r_1}\cdots(\lambda-\lambda_m)^{r_m}$$

这表明 A 的特征多项式的全部根都属于 K，并且每一个根 λ_i 的重数等于 r_i.

充分性. 设 A 的特征多项式 $|\lambda I-A|$ 在复数域中的全部不同的根 $\lambda_1,\cdots,\lambda_m$ 都属于 K，并且每个根 λ_i 的重数等于 A 的属于特征值 λ_i 的特征子空间的维数 r_i. 则

$$|\lambda I-A|=(\lambda-\lambda_1)^{r_1}\cdots(\lambda-\lambda_m)^{r_m}$$

从而 $r_1+\cdots+r_m=n$. 这表明 A 的属于不同特征值的特征子空间的维数之和等于 n，所以 A 可对角化. ▌

* **例** 8　设 A 是数域 K 上的一个 n 级矩阵，设 λ_1 是 A 的一个特征值. 证明：A 的属于特征值 λ_1 的特征子空间的维数不超过 λ_1 作为 A 的特征多项式的根的重数.

证明　设 A 的属于特征值 λ_1 的特征子空间 W_1 的维数为 r_1. 在

W_1 中取一个基：$\alpha_1,\cdots,\alpha_{r_1}$. 把它扩充成 K^n 的一个基：$\alpha_1,\cdots,\alpha_{r_1},\beta_1,\cdots,\beta_m$，其中 $r_1+m=n$. 令

$$U=(\alpha_1,\cdots,\alpha_{r_1},\beta_1,\cdots,\beta_m)$$

则 U 是可逆矩阵，并且有

$$\begin{aligned}U^{-1}AU&=U^{-1}(A\alpha_1,\cdots,A\alpha_{r_1},A\beta_1,\cdots,A\beta_m)\\&=(\lambda_1U^{-1}\alpha_1,\cdots,\lambda_1U^{-1}\alpha_{r_1},U^{-1}A\beta_1,\cdots,U^{-1}A\beta_m)\end{aligned}$$

因为

$$I=U^{-1}U=(U^{-1}\alpha_1,\cdots,U^{-1}\alpha_{r_1},U^{-1}\beta_1,\cdots,U^{-1}\beta_m)$$

所以

$$U^{-1}\alpha_1=\varepsilon_1\quad,\cdots\quad,U^{-1}\alpha_{r_1}=\varepsilon_{r_1}$$

从而得到

$$U^{-1}AU=\begin{pmatrix}\lambda_1I_{r_1} & B\\ 0 & C\end{pmatrix}$$

由于相似的矩阵有相同的特征多项式，所以

$$|\lambda I-A|=|\lambda I_{r_1}-\lambda_1I_{r_1}|\cdot|\lambda I_m-C|=(\lambda-\lambda_1)^{r_1}|\lambda I_m-C|$$

由此得出，λ_1 作为 A 的特征多项式的根的重数至少是 r_1. 换句话说，r_1 不超过 λ_1 作为 $|\lambda I-A|$ 的根的重数. ▍

注：设 λ_1 是数域 K 上 n 级矩阵 A 的一个特征值，我们把 A 的属于 λ_1 的特征子空间的维数叫做特征值 λ_1 的**几何重数**；而把 λ_1 作为 A 的特征多项式的根的重数叫做特征值 λ_1 的**代数重数**. 例 8 说明：A 的任一特征值 λ_1 的几何重数不超过它的代数重数.

习　题　5.8

1. 设 $a_1,a_2,\cdots,a_n$ 是不全为零的实数，$n>1$. 令 $A'=(a_1,a_2,\cdots,a_n)$，求 n 级矩阵 AA' 的特征值和特征向量；矩阵 AA' 是否可对角化？如果可对角化，求出一个可逆矩阵 U，使得 $U^{-1}(AA')U$ 为对角矩阵.

2. 设复数域上的矩阵

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

问：AB是否可对角化？如果可以，求出一个可逆矩阵U，使$U^{-1}(AB)U$为对角矩阵.

3. 用J表示元素全为1的n级矩阵，设有理数域上的矩阵$A=rI+mJ$，其中r,m是给定的正整数. 求A的特征值和特征向量；问A是否可对角化？如果可对角化，求出一个可逆矩阵U，使$U^{-1}AU$为对角矩阵.（提示：利用J的特征值和特征向量去求A的特征值和特征向量.）

4. 证明：不为零矩阵的幂零矩阵不能对角化.

5. 证明：数域K上的对合矩阵一定可对角化，并且写出这种对角矩阵的形式.

*6. 设$b_1,\cdots,b_n$是正实数，并且$\sum\limits_{i=1}^{n}b_i=1$. 设$A=(a_{ij})$，其中

$$a_{ij}=\begin{cases}1-b_i, & \text{当 } j=i\\-\sqrt{b_ib_j}, & \text{当 } j\neq i\end{cases}$$

求矩阵A的秩.（提示：去证A是幂等矩阵.）

7. 设

$$A=\begin{pmatrix}0 & 10 & 6\\1 & -3 & -3\\-2 & 10 & 8\end{pmatrix}$$

求A^{100}.

8. 设数列$\{a_k\}$满足下述递归公式

$$a_{k+2}=\frac{1}{2}(a_{k+1}+a_k),\ k=0,1,2,\cdots$$

以及初始条件：$a_0=0,a_1=\frac{1}{2}$. 求这个数列的通项公式，并且求$\lim\limits_{k\to\infty}a_k$.

§9 实对称矩阵的对角化

上一节我们从二次曲面方程的化简中提出了一个问题：实对

称矩阵 A 能不能对角化，能不能找到一个正交矩阵 T，使得 $T^{-1}AT$ 为对角矩阵？这一节就来回答这个问题.

定义 1 设 A,B 都是 n 级实矩阵，如果有一个 n 级正交矩阵 T，使得 $B = T^{-1}AT$，则称 **A 正交相似于 B.** 正交相似 $B=T^{-1}AT$

容易验证，正交相似是一个等价关系.

显然，如果 A 正交相似于 B，则 A 一定相似于 B. 反之，如果实矩阵 A 相似于 B，则 A 不一定能正交相似于 B，例如，习题 5.6 第 1 题的(4) 中的 3 级实矩阵 A，它的特征值是 $0,1,-1$，因此 A 可对角化，即 $A \sim D = \text{ding}\{0,1,-1\}$. 但是 A 不能正交相似于 D，这是因为假如有正交矩阵 T 使得 $D = T^{-1}AT$，则 $A = TDT^{-1}$. 从而 $A' = (TDT^{-1})' = TDT^{-1} = A$，由此推出 A 是对称矩阵，可是已知的 A 不是对称矩阵，矛盾. 从这个例子的论证过程，我们可看出有下面的结论：

定理 5.9.1 如果实矩阵 A 正交相似于一个对角矩阵 D，则 A 一定是对称矩阵.

证明 由已知条件，有正交矩阵 T 使得 $D = T^{-1}AT$. 于是 $A = TDT^{-1}$. 从而

$$A' = (TDT^{-1})' = (T^{-1})'D'T' = TDT^{-1} = A$$

所以 A 是对称矩阵. ▌

下面我们要证明定理 5.9.1 的逆命题也成立，即实对称矩阵一定能正交相似于一个对角矩阵. 为此首先来研究实对称矩阵的特征多项式的根的性质和特征向量的性质.

定理 5.9.2 实对称矩阵的特征多项式在复数域中的每一个根都是实数.

证明 设 A 是 n 级实对称矩阵，设 λ_0 是 A 的特征多项式 $|\lambda I - A|$ 在复数域中的任意一个根. 于是有 $|\lambda_0 I - A| = 0$. 从而齐次线性方程组$(\lambda_0 I - A)X = 0$ 有非零解. 设它的一个非零解是

$$\alpha = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

则 $(\lambda_0 I - A)\alpha = 0$,从而得

$$A\alpha = \lambda_0 \alpha \tag{1}$$

在(1)式两边左乘 $\overline{\alpha}'$ 得

$$\overline{\alpha}' A\alpha = \lambda_0 \overline{\alpha}' \alpha \tag{2}$$

在(2)式两边取复数共轭(参见习题 4.5 的第 3 题),注意 A 是实矩阵,于是得

$\Downarrow$ $A = \overline{A}$

$$\alpha' A \overline{\alpha} = \overline{\lambda}_0 \alpha' \overline{\alpha} \tag{3}$$

(3)式两边取转置,注意 A 是对称矩阵,于是得

$$\overline{\alpha}' A\alpha = \overline{\lambda}_0 \overline{\alpha}' \alpha \tag{4}$$

比较(2)和(4)式得

$$\lambda_0 \overline{\alpha}' \alpha = \overline{\lambda}_0 \overline{\alpha}' \alpha \tag{5}$$

从(5)式得 $(\lambda_0 - \overline{\lambda}_0)\overline{\alpha}' \alpha = 0$. 因为 $\alpha \neq 0$,所以

$$\begin{aligned} \overline{\alpha}' \alpha &= \overline{c}_1 c_1 + \overline{c}_2 c_2 + \cdots + \overline{c}_n c_n \\ &= |c_1|^2 + |c_2|^2 + \cdots + |c_n|^2 \neq 0 \end{aligned}$$

于是 $\lambda_0 = \overline{\lambda}_0$. 这表明 λ_0 是实数. ▌

从定理 5.9.2 得出,实对称矩阵的特征多项式在复数域内的每一个根都是这个矩阵的特征值.

定理 5.9.3 实对称矩阵 A 的属于不同特征值的特征向量一定是正交的.

证明 设 λ_1 与 λ_2 是 A 的不同的特征值,α_i 是 A 的属于 λ_i 的特征向量,$i = 1,2$. 因为

$$\lambda_1(\alpha_1,\alpha_2) = (\lambda_1\alpha_1,\alpha_2) = (A\alpha_1,\alpha_2) = (A\alpha_1)'\alpha_2 = \alpha_1' A\alpha_2$$

$$\lambda_2(\alpha_1,\alpha_2) = (\alpha_1,\lambda_2\alpha_2) = (\alpha_1,A\alpha_2) = \alpha_1' A\alpha_2$$

所以 $\lambda_1(\alpha_1,\alpha_2) = \lambda_2(\alpha_1,\alpha_2)$. 于是得 $(\lambda_1 - \lambda_2)(\alpha_1,\alpha_2) = 0$. 因为

$\lambda_1 \neq \lambda_2$，所以 $(\alpha_1,\alpha_2)=0$，即 α_1 与 α_2 正交. ▌

定理 5.9.4 实对称矩阵一定正交相似于一个对角矩阵.

证明 对实对称矩阵的级数 n 作数学归纳法.

$n=1$ 时，1 级实对称矩阵 (a) 已经是对角矩阵，取 I_1 这个正交矩阵，即有 $I_1^{-1}(a)I_1=(a)$.

假设任意一个 $n-1$ 级实对称矩阵都能正交相似于一个对角矩阵. 现在来看 n 级实对称矩阵 A.

据定理 5.9.2，A 一定有特征值. 设 λ_1 是 A 的一个特征值，设 η_1 是 A 的属于 λ_1 的一个特征向量，并且 $|\eta_1|=1$. 据第四章 §10 的推论 4.10.1，η_1 可扩充成 R^n 的一个标准正交基：$\eta_1,\eta_2,\cdots,\eta_n$. 于是矩阵 $T_1=(\eta_1,\eta_2,\cdots,\eta_n)$ 是正交矩阵. 我们有

$$\begin{aligned} T_1^{-1}AT_1 &= T_1^{-1}(A\eta_1,A\eta_2,\cdots A\eta_n) \\ &= (T_1^{-1}\lambda_1\eta_1, T_1^{-1}A\eta_2,\cdots,T_1^{-1}A\eta_n) \end{aligned}$$

因为 $T_1^{-1}T_1=I$，所以

$$T_1^{-1}(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$$

于是得，$T_1^{-1}\eta_1=\varepsilon_1$. 所以 $T_1^{-1}AT_1$ 的第 1 列是 $\lambda_1\varepsilon_1$. 从而可以设

$$T_1^{-1}AT_1=\begin{pmatrix}\lambda_1 & \alpha \\ 0 & B\end{pmatrix}$$

因为 T_1 是正交矩阵，并且 A 是对称矩阵，所以

$$(T_1^{-1}AT_1)'=T_1'A'(T_1^{-1})'=T_1^{-1}AT_1$$

于是 $T_1^{-1}AT_1$ 是对称矩阵. 从而得，$\alpha=0$，并且 B 是 $n-1$ 级实对称矩阵. 据归纳假设，有 $n-1$ 级正交矩阵 T_2，使

$$T_2^{-1}BT_2=\mathrm{diag}\{\lambda_2,\cdots,\lambda_n\}$$

令
$$T=T_1\begin{pmatrix}1 & 0 \\ 0 & T_2\end{pmatrix}$$

由于上式右端的两个矩阵都是正交矩阵，从而 T 是正交矩阵，并

且有

$$
\begin{aligned}
T^{-1}AT &= \begin{pmatrix} 1 & 0 \\ 0 & T_2 \end{pmatrix}^{-1} T_1{}^{-1}AT_1 \begin{pmatrix} 1 & 0 \\ 0 & T_2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & T_2{}^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & T_2 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & 0 \\ 0 & T_2{}^{-1}BT_2 \end{pmatrix} \\
&= \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}
\end{aligned}
$$

据数学归纳法原理，对任意自然数 n，任一 n 级实对称矩阵都正交相似于一个对角矩阵. ▌

从定理 5.9.4 得出，实对称矩阵一定可对角化. 于是 n 级实对称矩阵 A 的属于不同特征值的特征子空间的维数之和一定等于 n；并且 A 的相似标准形是对角矩阵，其主对角线上元素是 A 的全部特征值，每个特征值 λ_i 在主对角线上出现的次数等于 A 的属于 λ_i 的特征子空间的维数. 那么如何找一个正交矩阵 T，使 $T^{-1}AT$ 为对角矩阵呢？方法如下：

第一步，求出 A 的特征多项式 $|\lambda I - A|$ 在复数域内的全部不同的根 $\lambda_1, \cdots, \lambda_m$，它们都是实数，从而它们都是 A 的特征值.

第二步，对于每一个特征值 λ_i，求出齐次线性方程组 $(\lambda_i I - A)X = 0$ 的一个基础解系：$\alpha_{i1}, \cdots, \alpha_{ir_i}$. 然后把 $\alpha_{i1}, \cdots, \alpha_{ir_i}$ 正交化（运用 Gram－Schmidt 正交化过程），再单位化得 $\eta_{i1}, \cdots, \eta_{ir_i}$. 它们与 $\alpha_{i1}, \cdots, \alpha_{ir_i}$ 等价，因此它们是 A 的属于 λ_i 的特征向量，并且它们是正交单位向量组.

第三步，令

$$
T = (\eta_{11}, \cdots, \eta_{1r_1}, \cdots, \eta_{m1}, \cdots, \eta_{mr_m}).
$$

据定理 5.9.3 得，T 的列向量组是正交单位向量组. 从而 T 是 n 级正交矩阵，并且有

$$
T^{-1}AT = \mathrm{diag}\{\underbrace{\lambda_1, \cdots, \lambda_1}_{r_1}, \cdots, \underbrace{\lambda_m, \cdots, \lambda_m}_{r_m}\}
$$

例 1　设

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

求正交矩阵 T，使 $T^{-1}AT$ 为对角矩阵.

解

$$|\lambda I-A|=\begin{vmatrix}\lambda-4 & 1 & 1 & -1\\ 1 & \lambda-4 & -1 & 1\\ 1 & -1 & \lambda-4 & 1\\ -1 & 1 & 1 & \lambda-4\end{vmatrix}$$

$$=(\lambda-3)^3(\lambda-7)$$

所以 A 的特征值是 3(三重)，7.

对于特征值 3，求得 $(3I-A)X=0$ 的一个基础解系：

$$\alpha_1=\begin{pmatrix}1\\1\\0\\0\end{pmatrix},\quad \alpha_2=\begin{pmatrix}1\\0\\1\\0\end{pmatrix},\quad \alpha_3=\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}$$

把 $\alpha_1,\alpha_2,\alpha_3$ 正交化，令 $\beta_1=\alpha_1$，

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1$$

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

$$\beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2$$

$$=\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}-\frac{1}{2}\begin{pmatrix}1\\1\\0\\0\end{pmatrix}-\frac{\frac{1}{2}}{\frac{3}{2}}\begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\\0\end{pmatrix}=\begin{pmatrix}\frac{1}{3}\\-\frac{1}{3}\\-\frac{1}{3}\\-1\end{pmatrix}$$

把 β_1,β_2,β_3 分别单位化，得

$$\eta_1=\begin{pmatrix}\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\\0\\0\end{pmatrix},\quad \eta_2=\begin{pmatrix}\frac{\sqrt{6}}{6}\\-\frac{\sqrt{6}}{6}\\\frac{\sqrt{6}}{3}\\0\end{pmatrix},\quad \eta_3=\begin{pmatrix}\frac{\sqrt{3}}{6}\\-\frac{\sqrt{3}}{6}\\-\frac{\sqrt{3}}{6}\\-\frac{\sqrt{3}}{2}\end{pmatrix}$$

对于特征值 7，求得$(7I-A)X=0$ 的一个基础解系：

$$\alpha_4=\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}$$

把 α_4 单位化，得

$$\eta_4=\begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\end{pmatrix}$$

令

$$
T=\begin{pmatrix}
\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\
\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & -\frac{1}{2} \\
0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{6} & -\frac{1}{2} \\
0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2}
\end{pmatrix},
$$

则 T 是正交矩阵，并且有

$$T^{-1}AT=\mathrm{diag}\{3,3,3,7\}$$

前面我们已指出，两个实矩阵相似，但不一定能正交相似. 然而对于实对称矩阵，我们却有肯定的结论，即

定理 5.9.5 两个 n 级实对称矩阵正交相似的充分必要条件是它们相似. 相似 $\Longleftrightarrow$ 正交相似

证明 必要性是显然的. 现在证充分性. 设 A 与 B 是 n 级实对称矩阵，并且 $A\sim B$. 于是 A 与 B 有相同的特征多项式. 又由于 A 与 B 都是实对称矩阵，它们的特征多项式的根都是特征值，因此 A 与 B 的特征值完全相同(包括重数也相同). 设 A 与 B 的特征值是 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 据定理 5.9.4，$A$ 正交相似于对角矩阵 $\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$；$B$ 也正交相似于 $\mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}$. 由于正交相似具有对称性和传递性，从而 A 正交相似于 B. ▌

定理 5.9.5 说明，实对称矩阵 A 的正交相似等价类与它的相似等价类是相等的. 而对于一般的实矩阵 A 来说，A 的正交相似等价类是它的相似等价类的子集.

从定理 5.9.5 的充分性的证明过程中可以看出，如果两个实对称矩阵 A 与 B 的特征值完全相同(包括重数也相同)，则它们正交相似，从而相似. 因此对于所有 n 级实对称矩阵组成的集合来说，特征值(包括重数)是相似关系的完全不变量.

习 题 5.9

1. 求正交矩阵 T,使 $T^{-1}AT$ 为对角矩阵:

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

(2)
$$A=\begin{pmatrix}6&2&4\\2&3&2\\4&2&6\end{pmatrix}$$

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

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

2. 证明:如果 A 是 n 级实对称矩阵,且 $A^2=A$,则存在正交矩阵 T,使得 $T^{-1}AT=\mathrm{diag}\{I_r,0\}$,其中 r 是 A 的秩.

3. 证明:如果 A 是 n 级实对称矩阵,并且 $A^2=I$,则存在正交矩阵 T,使得 $T^{-1}AT=\mathrm{diag}\{I_r,-I_{n-r}\}$,其中 $r=\mathrm{rank}(I+A)$.

4. 证明:如果 A 是 n 级实对称矩阵,并且 A 是幂零矩阵,则 $A=0$.

5. 证明:如果 A 与 B 都是 n 级实对称矩阵,则 A 与 B 相似的充分必要条件是它们有相同的特征多项式.

6. 证明:如果 A 是一个 $s\times n$ 实矩阵,则 $A'A$ 的特征值都是非负实数.

7. 证明:如果 n 级实矩阵 A 的特征多项式的根都是实数,则 A 一定正交相似于一个上三角矩阵.

8. 证明:如果 n 级实矩阵 A 正交相似于一个上三角矩阵,则 A 的特征多项式的根全是实数.

*9. 证明:任一 n 级复矩阵一定相似于一个上三角矩阵.

*10. 证明:如果 n 级实矩阵 A 的特征多项式的根都是实数,并且 $AA'=$

$A'A$,则 A 一定是对称矩阵.(提示:利用本节定理 5.9.1 以及第 7 题结论.)

补 充 题 五

1 设 A 是复数域上 n 级可逆矩阵,并且 $A \sim A^k$,其中 k 是某个正整数.证明:A 的特征值都是单位根.(提示:设 λ_0 是 A 的特征值,由 $A \sim A^k$,去证 λ_0^k 是 A 的特征值,然后运用这个结论.)

2. 证明:正交矩阵的特征多项式的根的模等于 1.

3. 设 A 是 2 级正交矩阵.证明:

(1) 如果 $|A| = 1$,则 A 正交相似于下述形式的矩阵

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

其中 θ 是实数.

(2) 如果 $|A| = -1$,则 A 正交相似于对角矩阵

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

4. 设 A 是 3 级正交矩阵,证明:存在正交矩阵 T,使得

$$T^{-1}AT = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & a \end{pmatrix}$$

其中,当 $|A| = 1$ 时,$a = 1$;当 $|A| = -1$ 时,$a = -1$;θ 是实数.

5. 在复数域上的 n 维向量空间 C^n 中,任给两个列向量 α,β,规定

$$(\alpha,\beta) := \alpha'\overline{\beta}$$

这个二元复值函数 (α,β) 称为 C^n 的一个**内积**.证明:$\forall\ \alpha,\beta,\gamma \in C^n, k \in C$,有

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

2° $(\alpha+\gamma,\beta) = (\alpha,\beta) + (\gamma,\beta)$;

3° $(k\alpha,\beta) = k(\alpha,\beta)$;

4° (α,α) 是非负实数;$(\alpha,\alpha) = 0$ 当且仅当 $\alpha = 0$.

6. 在 C^n 中,定义了第 5 题所述的内积后,称 C^n 为**酉空间**。如果$(\alpha,\beta) =$

0，则称 α 与 β 正交。C^n 中由非零向量组成的向量组称为**正交向量组**，如果它们两两正交. 证明：C^n 中，正交向量组一定是线性无关的.

7. 在酉空间 C^n 中，把 $\sqrt{(\alpha,\alpha)}$ 称为 α 的**长度**，记作 $|\alpha|$. 长度为 1 的向量称为**单位向量**. C^n 中 n 个单位向量组成的正交向量组是 C^n 的一个基，称它为**标准正交基**。设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 C^n 的一个基，令

$$\beta_1=\alpha_1,\quad \beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1,\cdots$$

$$\beta_n=\alpha_n-\sum_{i=1}^{n-1}\frac{(\alpha_n,\beta_i)}{(\beta_i,\beta_i)}\beta_i$$

证明：$\beta_1,\beta_2,\cdots,\beta_n$ 是正交向量组，并且它与 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 等价.

8. 复数域上的方阵 A 称为**酉矩阵**，如果 $AA^*=I$。证明：酉矩阵的特征值的模为 1.

9. 证明：n 级复矩阵 A 是酉矩阵当且仅当 A 的行（或列）向量组是酉空间 C^n 的一个标准正交基.

10. 证明：对于任一 n 级酉矩阵 A，存在一个 n 级酉矩阵 T，使得 $T^{-1}AT$ 为对角矩阵. 这时称 A **酉相似**于一个对角矩阵.（提示：类似于定理 9.4 的证明.）

11. 证明：酉矩阵 A 的属于不同特征值的特征向量一定正交.

12. 设 $A=(a_{ij})$ 是 n 级复矩阵，A 的所有特征值组成的 n 元数组 $(\lambda_1,\lambda_2,\cdots,\lambda_n)$ 称为 A 的**谱**. A 的特征值的模的最大值称为 A 的**谱半径**，记作 $Sr(A)$. 证明：

$$Sr(A)\leqslant\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}|,\quad Sr(A)\leqslant\max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|$$

13. 设 $A=(a_{ij})$ 是 n 级复矩阵，令

$$D_i(A)=\{z\in C\mid |z-a_{ii}|\leqslant\sum_{k\neq i}|a_{ik}|\},\quad i=1,\cdots,n$$

称 $D_i(A)$，$i=1,\cdots,n$，是 A 的 Gersgorin **圆盘**. 证明下述的 Gersgorin 圆盘定理：n 级复矩阵 A 的每一个特征值都处于 A 的某个 Gersgorin 圆盘里.

14. 设 $A=(a_{ii})$ 是 n 级复矩阵，用 Gersgorin 圆盘定理证明：如果 $|a_{ii}|>\sum_{k\neq i}|a_{ik}|$，$i=1,2,\cdots,n$，则 A 可逆.

15. 设 $A=(a_{ij})$ 是 n 级复矩阵，证明：如果 $(n-1)|a_{ij}|<|a_{ii}|$，$j\neq i$，$i,j=1,2,\cdots,n$，则 A 可逆.

16. 设 A 是 n 级复矩阵，如果 A 的每一个特征值的实部都是负数，则 A 称为**稳定矩阵**. 稳定矩阵在微分方程理论中有重要应用. 判断下述矩阵 A 是否为稳定矩阵.

$$A=\begin{pmatrix} -8 & 2 & 0 & 2 \\ 0 & -5 & 1 & 1 \\ 1 & -1 & -8 & -2 \\ 1 & 1 & 1 & -6 \end{pmatrix}$$

* 17. 设 $B(t)=(f_{ij}(t))$ 是由可微函数 $f_{ij}(t)$，$i,j=1,2,\cdots,n$，组成的 n 级矩阵，规定：

$$\frac{\mathrm{d}B(t)}{\mathrm{d}t}:=(f'_{ij}(t))$$

由导数性质得，对于 n 级实矩阵 C 有

$$\frac{\mathrm{d}(CB(t))}{\mathrm{d}t}=C\frac{\mathrm{d}B(t)}{\mathrm{d}t}$$

求下述线性微分方程组的通解：

$$\begin{cases} \dfrac{\mathrm{d}y_1}{\mathrm{d}t}=y_1-2y_2+2y_3 \\ \dfrac{\mathrm{d}y_2}{\mathrm{d}t}=-2y_1-2y_2+4y_3 \\ \dfrac{\mathrm{d}y_3}{\mathrm{d}t}=2y_1+4y_2-2y_3 \end{cases}$$

其中 y_1,y_2,y_3 都是 t 的函数，它们是未知的.

第六章　二次型·矩阵的合同分类

§1　引　　言

在第五章§8我们曾指出，为了判别二次曲面方程

$$x^2+4y^2+z^2-4xy-8xz-4yz-1=0 \tag{1}$$

表示什么样的二次曲面，应当作直角坐标变换

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}=T\begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix} \tag{2}$$

其中T是正交矩阵，使得在新的直角坐标系中，(1)的二次项部分

$$\begin{aligned} &x^2+4y^2+z^2-4xy-8xz-4yz \\ &=(x,y,z)\underbrace{\begin{pmatrix} 1 & -2 & -4 \\ -2 & 4 & -2 \\ -4 & -2 & 1 \end{pmatrix}}_{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned} \tag{3}$$

变成

$$(x^*,y^*,z^*)T'AT\begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix} \tag{4}$$

其中A是(3)式右端的3级矩阵. 当矩阵$T'AT$为对角矩阵时，(4)式就只含x^*,y^*,z^*的平方项，从而可判别方程(1)表示什么二次曲面.

(3)式左端是x,y,z的二次齐次多项式(各项的次数相同称为齐次)，称它为x,y,z的二次型. 公式(2)是x,y,z到x^*,y^*,z^*的

一个线性替换，它的系数矩阵 T 是正交矩阵，称公式(2)是正交的线性替换. 如果系数矩阵是可逆矩阵，则称之为非退化(或非奇异，或可逆)的线性替换.

从上述二次曲面方程的化简可抽象出一个问题：把一个二次型经过非退化的线性替换化成只含平方项的形式. 用矩阵的术语来说就是，对于一个对称矩阵 A，要找一个可逆矩阵 C，使得 $C'AC$ 为对角矩阵. 由此引出需要讨论矩阵的合同关系、合同分类以及合同标准形等问题. 这一章就来讨论这些问题. 二次型在数学的其他分支以及物理、力学和工程技术中都很有用.

§2 二次型和它的标准形·矩阵的合同关系

定义 1 **数域 K 上的一个 n 元二次型**是系数在 K 中的 n 个变量的二次齐次多项式. 它的一般形式是

$$\begin{aligned}&f(x_1,x_2,\cdots,x_n)\\&=a_{11}x_1^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n+a_{22}x_2^2\\&\quad+2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n+\cdots+a_{nn}x_n^2\end{aligned}\tag{1}$$

其中 $a_{ij}\in K$， $i<j$， $i,j,=1,2,\cdots,n$. 这里我们把 x_ix_j 的系数写成 $2a_{ij}$(当 $i\neq j$)是为了使系数具有对称性.

如果令 $a_{ji}=a_{ij}$，当 $i<j$，则二次型(1)也可写成

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\tag{2}$$

我们也可以用矩阵的术语来表示二次型(1). 为此我们用二次型(1)的系数按下述规则组成一个矩阵 A：A 的 (i,i) 元等于 x_i^2 项的系数，$i=1,2,\cdots,n$；A 的 (i,j) 元等于 x_ix_j 项的系数的一半，$i<j$，$i,j=1,2,\cdots,n$；A 的 (j,i) 元等于 A 的 (i,j) 元，$i<j$， $i,j=1,2,\cdots,n$. 这样作出的矩阵 A 称为**二次型(1)的矩阵**，它是对称矩阵. 显然，给定一个二次型，它的矩阵是唯一的. 把变量

$x_1,x_2,\cdots,x_n$ 组成的列向量记作 X，则我们有

$$X'AX=(x_1,x_2,\cdots,x_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{12}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\vdots&\vdots\\a_{1n}&a_{2n}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

$$=(x_1,x_2,\cdots,x_n)\begin{pmatrix}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\\a_{12}x_1+a_{22}x_2+\cdots+a_{2n}x_n\\\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\a_{1n}x_1+a_{2n}x_2+\cdots+a_{nn}x_n\end{pmatrix}$$

$$=a_{11}x_1^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n+a_{22}x_2^2$$
$$+2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n+\cdots+a_{nn}x_n^2$$

因此二次型(1)可以写成 $X'AX$，其中 A 是这个二次型的矩阵，它是对称矩阵.

从上面的计算过程看出，任给一个数域 K 上的 n 级对称矩阵 A，可以得到数域 K 上的一个 n 元二次型 $X'AX$，并且这个二次型的矩阵就是 A.

如果二次型 $X'AX=X'BX$，并且 A,B 都是对称矩阵，则 A，B 分别是二次型 $X'AX,X'BX$ 的矩阵. 由于 $X'AX=X'BX$，所以 $A=B$. 由此得出，一个二次型 $f(x_1,x_2,\cdots,x_n)$ 表示成 $X'AX$(其中 A 是对称矩阵)的形式的表法唯一.

综上述，数域 K 上所有 n 元二次型组成的集合与 K 上所有 n 级对称矩阵组成的集合之间有一个一一对应，这个对应法则是让二次型 $X'AX$(A 是对称矩阵)对应到它的矩阵 A.

今后我们一提到二次型 $X'AX$ 时，总是要求 A 是对称矩阵，不再每次声明.

我们在引言中指出，许多问题希望通过变量的线性替换把二次型化成只含平方项的形式. 下面我们给出变量的线性替换的确切含意：

定义 2　设 $x_1,\cdots,x_n;y_1,\cdots y_n$ 是两组变量，令

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$

设 $C=(c_{ij})$ 是数域 K 上的一个 n 级矩阵. 下述关系式

$$X=CY \tag{3}$$

即

$$\begin{cases} x_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n\\ x_2=c_{21}y_1+c_{22}y_2+\cdots+c_{2n}y_n\\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\ x_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n\end{cases}$$

称为由 $x_1,\cdots,x_n$ 到 $y_1,\cdots,y_n$ 的一个**线性替换**,简称为**线性替换**. 如果 C 是可逆矩阵,则称线性替换(3) 是**非退化的**(或**非奇异的**,或**可逆的**).

对于二次型 $X'AX$,如果作一个线性替换(3),则

$$X'AX=(CY)'A(CY)=Y'(C'AC)Y \tag{4}$$

A—对称矩阵.

显然 $Y'(C'AC)Y$ 仍为二次型,并且 $C'AC$ 仍为对称矩阵,从而 $C'AC$ 是二次型 $Y'(C'AC)Y$ 的矩阵. 这说明:线性替换把二次型变成二次型,并且替换后的二次型 $Y'(C'AC)Y$ 的矩阵为 $C'AC$,其中 A 是原二次型 $X'AX$ 的矩阵. 如果所作的线性替换(3) 是非退化的,则 C 是可逆矩阵. 由此我们引入

定义 3 数域 K 上的 n 级矩阵 A 与 B 称为**合同的**,如果有数域 K 上的 n 级可逆矩阵 C,使

$$B=C'AC \tag{5}$$

合同是矩阵之间的一个关系,记作 $\simeq$,它具有反身性(因为 $A\simeq I'AI$),对称性(因为 $A\simeq B\Rightarrow$ 有可逆矩阵 C,使得 $B=C'AC\Rightarrow A=(C^{-1})'BC^{-1}\Rightarrow B\simeq A$),传递性(因为 $A\simeq B$ 且 $B\simeq D\Rightarrow$ 有可逆矩阵 C_1,C_2,使得 $B=C_1'AC_1,D=C_2'BC_2\Rightarrow D=C_2'(C_1'AC_1)C_2=(C_1C_2)'A(C_1C_2)\Rightarrow A\simeq D$). 因此,合同是一个等价关系. 集合 $M_n(K)$ 的所有合同等价类组成该集合的一个划分.

显然，如果 A 与 B 合同，则 A 与 B 一定相抵. 反之，如果 A 与 B 相抵，则 A 与 B 不一定合同. 于是，合同的矩阵有相同的秩，但是秩相同的矩阵不一定合同.

从前三段的结论知道，经过非退化的线性替换，新二次型的矩阵与原二次型的矩阵是合同的.

定义 4 数域 K 上的两个 n 元二次型称为**等价的**，如果用一个非退化的线性替换可将其中一个化为另一个.

容易验证二次型的等价是一个等价关系.

由上述知，如果两个 n 元二次型等价，则它们的矩阵是合同的. 反之，如果两个 n 元二次型 $X'AX$ 与 $Y'BY$ 的矩阵 A 与 B 合同，则这两个二次型等价. 这是因为从 A 与 B 合同推出，有可逆矩阵 C，使得 $B = C'AC$. 从而作非退化的线性替换：$X = CY$ 就得到

$$X'AX = (CY)'A(CY) = Y'C'ACY = Y'BY$$

这说明：研究二次型的等价分类问题与研究对称矩阵的合同分类问题本质上是同一个问题. 在二次型 $X'AX$ 的等价类中找一个只含平方项的二次型，就相当于在对称矩阵 A 的合同类中找一个对角矩阵，这是因为二次型只含平方项当且仅当它的矩阵是对角矩阵. 现在我们来讨论对于任意一个数域 K 上的 n 元二次型 $X'AX$，能不能在它的等价类中找到一个只含平方项的二次型？也就是说，能不能通过非退化的线性替换，把 $X'AX$ 化成只含平方项的形式？用矩阵的术语来说，对于任意数域 K 上的一个对称矩阵 A，能不能在它的合同类中找到一个对角矩阵？也就是说，数域 K 上的任一对称矩阵能不能合同于一个对角矩阵？回答是肯定的.

二次型 $X'AX$ 经过非退化的线性替换化成的只含平方项的二次型称为 $X'AX$ 的一个**标准形**.

对称矩阵 A 经过合同变换变成的对角矩阵称为 A 的一个**合同标准形**.

在证明一般结论之前，让我们先看两个具体例子.

例 1 作非退化的线性替换把下述二次型化成标准形，并且

写出所作的非退化线性替换.

$$f(x_1,x_2,x_3)=x_1^2+2x_2^2-x_3^2+4x_1x_2-4x_1x_3-4x_2x_3$$

解 因为标准形只含变量的平方项,所以我们可用配方法把变量 x_1,x_2,x_3 逐个地配成完全平方的形式.

$$\begin{aligned}
&f(x_1,x_2,x_3)\\
&=x_1^2+4x_1(x_2-x_3)+[2(x_2-x_3)]^2-[2(x_2-x_3)]^2\\
&\quad+2x_2^2-x_3^2-4x_2x_3\\
&=[x_1+2(x_2-x_3)]^2-4(x_2^2-2x_2x_3+x_3^2)\\
&\quad+2x_2^2-x_3^2-4x_2x_3\\
&=(x_1+2x_2-2x_3)^2-2x_2^2+4x_2x_3-5x_3^2\\
&=(x_1+2x_2-2x_3)^2-2(x_2^2-2x_2x_3+x_3^2-x_3^2)-5x_3^2\\
&=(x_1+2x_2-2x_3)^2-2(x_2-x_3)^2-3x_3^2
\end{aligned}$$

令

$$\begin{cases}y_1=x_1+2x_2-2x_3\\ y_2=x_2-x_3\\ y_3=x_3\end{cases}$$

则 $f(x_1,x_2,x_3)=y_1^2-2y_2^2-3y_3^2$.

所作的线性替换是

$$\begin{cases}x_1=y_1-2y_2\\ x_2=y_2+y_3\\ x_3=y_3\end{cases}\tag{6}$$

其系数矩阵的行列式

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

因此这个线性替换是非退化的.

例 2 同例 1 的要求,所给二次型为

$$f(x_1,x_2,x_3)=x_1x_2+x_1x_3-3x_2x_3$$

解　这个二次型中没有平方项，为了能够进行配方，首先要通过非退化的线性替换变成含有平方项的形式. 令

$$\begin{cases} x_1 = y_1 - y_2 \\ x_2 = y_1 + y_2 \\ x_3 = y_3 \end{cases} \tag{7}$$

则

$$\begin{aligned} & f(x_1, x_2, x_3) \\ =& (y_1 - y_2)(y_1 + y_2) + (y_1 - y_2)y_3 - 3(y_1 + y_2)y_3 \\ =& y_1^2 - y_2^2 - 2y_1y_3 - 4y_2y_3 \\ =& y_1^2 - 2y_1y_3 + y_3^2 - y_3^2 - y_2^2 - 4y_2y_3 \\ =& (y_1 - y_3)^2 - y_3^2 - [y_2^2 + 4y_2y_3 + (2y_3)^2 - (2y_3)^2] \\ =& (y_1 - y_3)^2 - y_3^2 - (y_2 + 2y_3)^2 + 4y_3^2 \\ =& (y_1 - y_3)^2 - (y_2 + 2y_3)^2 + 3y_3^2 \end{aligned}$$

令

$$\begin{cases} z_1 = y_1 - y_3 \\ z_2 = y_2 + 2y_3 \\ z_3 = y_3 \end{cases} \tag{8}$$

则 $f(x_1, x_2, x_3) = z_1^2 - z_2^2 + 3z_3^2$.

为了写出所作的线性替换，先从(8)解出 y_1, y_2, y_3，得

$$\begin{cases} y_1 = z_1 + z_3 \\ y_2 = z_2 - 2z_3 \\ y_3 = z_3 \end{cases} \tag{9}$$

把(9)代入(7)得

$$\begin{cases} x_1 = z_1 - z_2 + 3z_3 \\ x_2 = z_1 + z_2 - z_3 \\ x_3 = z_3 \end{cases} \tag{10}$$

容易看出线性替换(7)和(9)的系数矩阵都是可逆矩阵，而线性替换(10)的系数矩阵是上述两个可逆矩阵的乘积（设 $X = C_1Y, Y =$

C_2Z,则 $X=(C_1C_2)Z)$,所以它也是可逆矩阵,从而(10)是非退化的线性替换.

例 1 和例 2 中所用的配方法能够把数域 K 上任何一个二次型经过非退化线性替换变成只含平方项的二次型.下面就来证明这一结论.

定理 6.2.1 数域 K 上任意一个二次型都能经过非退化的线性替换变成只含平方项的二次型.

证明 对二次型的变量个数 n 作数学归纳法.

$n=1$ 时,二次型就是 $a_{11}x_1^2$,这只含平方项.

假设对于数域 K 上任意一个 $n-1$ 元二次型命题为真.现在来看 n 元二次型(1),分三种情形讨论:

1)平方项系数 $a_{11},a_{22},\cdots,a_{nn}$ 中至少有一个不为零.不失普遍性,设 $a_{11}\neq 0$.我们有

$$\begin{aligned}
&f(x_1,x_2,\cdots,x_n)\\
=&a_{11}[x_1^2+2a_{11}{}^{-1}x_1(a_{12}x_2+\cdots+a_{1n}x_n)]\\
&+a_{22}x_2^2+2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n+\cdots+a_{nn}x_n^2\\
=&a_{11}[x_1+a_{11}{}^{-1}(a_{12}x_2+\cdots+a_{1n}x_n)]^2\\
&-a_{11}{}^{-1}(a_{12}x_2+\cdots+a_{1n}x_n)^2\\
&+a_{22}x_2^2+2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n+\cdots+a_{nn}x_n^2
\end{aligned}$$

令

$$\begin{cases}
y_1=x_1+a_{11}{}^{-1}(a_{12}x_2+\cdots+a_{1n}x_n)\\
y_2=x_2\\
\cdots\quad\cdots\quad\cdots\\
y_n=x_n
\end{cases}$$

即

$$\begin{cases}
x_1=y_1-a_{11}{}^{-1}(a_{12}y_2+\cdots+a_{1n}y_n)\\
x_2=y_2\\
\cdots\quad\cdots\quad\cdots\\
x_n=y_n
\end{cases}\tag{11}$$

这是非退化的线性替换，它使得

$$\begin{aligned} &f(x_1,x_2,\cdots,x_n) \\ &=a_{11}y_1^2-a_{11}{}^{-1}(a_{12}y_2+\cdots+a_{1n}y_n)^2+a_{22}y_2^2+2a_{23}y_2y_3 \\ &\quad+\cdots+2a_{2n}y_2y_n+\cdots+a_{nn}y_n^2 \end{aligned} \tag{12}$$

(12)式右端除去第一项 $a_{11}y_1^2$ 外，其余项是 $y_2,\cdots,y_n$ 的 $n-1$ 元二次型.由归纳假设，有非退化线性替换

$$\begin{pmatrix} y_2 \\ \vdots \\ y_n \end{pmatrix} = C_2 \begin{pmatrix} z_2 \\ \vdots \\ z_n \end{pmatrix} \tag{13}$$

使它变成只含平方项的二次型

$$d_2z_2^2+d_3z_3^2+\cdots+d_nz_n^2 \tag{14}$$

于是作非退化线性替换

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & C_2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \tag{15}$$

就使 $f(x_1,x_2,\cdots,x_n)$ 变成了

$$f(x_1,x_2,\cdots,x_n)=a_{11}z_1^2+d_2z_2^2+d_3z_3^2+\cdots+d_nz_n^2 \tag{16}$$

由于线性替换(11)和(15)都是非退化的，所以从 $x_1,\cdots,x_n$ 到 $z_1,\cdots,z_n$ 的线性替换也是非退化的.

2) $a_{11}=a_{22}=\cdots=a_{nn}=0$，但是至少有一个 $a_{1j}\neq 0(j>1)$.不妨设 $a_{12}\neq 0$.令

$$\begin{cases} x_1=y_1-y_2 \\ x_2=y_1+y_2 \\ x_3=y_3 \\ \cdots \quad \cdots \quad \cdots \\ x_n=y_n \end{cases} \tag{17}$$

显然这是非退化线性替换，并且使得

$$
\begin{aligned}
& f(x_1, x_2, \cdots, x_n) \\
={} & 2a_{12}(y_1 - y_2)(y_1 + y_2) + 2a_{13}(y_1 - y_2)y_3 \\
& + \cdots + 2a_{1n}(y_1 - y_2)y_n + 2a_{23}(y_1 + y_2)y_3 \\
& + \cdots + 2a_{2n}(y_1 + y_2)y_n + \cdots + 2a_{n-1,n}y_{n-1}y_n \\
={} & 2a_{12}y_1^2 - 2a_{12}y_2^2 + 2a_{13}y_1y_3 - 2a_{13}y_2y_3 \\
& + \cdots + 2a_{1n}y_1y_n - 2a_{1n}y_2y_n + 2a_{23}y_1y_3 + 2a_{23}y_2y_3 \\
& + \cdots + 2a_{2n}y_1y_n + 2a_{2n}y_2y_n + \cdots + 2a_{n-1,n}y_{n-1}y_n \qquad (18)
\end{aligned}
$$

(18)式右端是 $y_1, y_2, \cdots, y_n$ 的二次型，并且 y_1^2 的系数不为零. 由情形 1) 知，它可以经过非退化线性替换

$$Y = CZ \qquad (19)$$

变成只含平方项的二次型

$$d_1z_1^2 + d_2z_2^2 + \cdots + d_nz_n^2 \qquad (20)$$

从而在非退化线性替换(17)与(19)的相继作用下(这相当于作了一个总的线性替换)，二次型(1)变成了二次型(20).

3) $a_{11} = a_{22} = \cdots = a_{nn} = 0$，并且 $a_{12} = a_{13} = \cdots = a_{1n} = 0$. 这时

$$f(x_1, x_2, \cdots, x_n) = 2a_{23}x_2x_3 + \cdots + 2a_{n-1,n}x_{n-1}x_n$$

它是 $n - 1$ 元二次型，据归纳假设，它能用非退化的线性替换化成只含平方项的形式.

上述表明，在归纳假设下，数域 K 上任意一个 n 元二次型都能经过非退化线性替换化成只含平方项的形式. 根据数学归纳法原理，定理 6.2.1 得证. ▌

用矩阵的术语，定理 6.2.1 可以叙述为

定理 6.2.2　数域 K 上任一对称矩阵都合同于一个对角矩阵. ▌

前面讲的配方法是把二次型化成标准形的基本方法. 现在我们再介绍一种方法.

由于把一个二次型 $X'AX$ 经过非退化线性替换化成标准形的问题等价于把一个对称矩阵 A 经过合同变换化成对角矩阵的问

题，因此我们来讨论如何把一个对称矩阵 A 经过合同变换化成对角形.

设 A 是数域 K 上一个 n 级对称矩阵，据定理 6.2.2，存在一个可逆矩阵 C，使得

$$C'AC = \mathrm{diag}\{d_1, d_2, \cdots, d_n\} =: D \tag{21}$$

因为 C 可逆，所以 C 可以表示成一些初等矩阵 $P_1, \cdots, P_s$ 的乘积，即

$$C = P_1 P_2 \cdots P_s \tag{22}$$

于是从(21)和(22)式得

$$\begin{aligned} D &= (P_1 P_2 \cdots P_s)' A (P_1 P_2 \cdots P_s) \\ &= P_s' \cdots P_2' P_1' A P_1 P_2 \cdots P_s \end{aligned} \tag{23}$$

我们知道，初等矩阵有三种类型：$P(i, j(k))$，$P(i, j)$，$P(i(b))$. 显然，$P(i,j)' = P(i,j)$，$P(i(b))' = P(i(b))$. 易直接验证：$P(i,j(k))' = P(j,i(k))$. 于是我们有

$$P(i,j)'AP(i,j) = P(i,j)AP(i,j)$$

这相当于把 A 的第 i, j 行互换，接着把所得矩阵的第 i, j 列互换. 我们也有

$$P(i(b))'AP(i(b)) = P(i(b))AP(i(b))$$

这相当于把 A 的第 i 行乘以非零数 b，接着把所得矩阵的第 i 列乘以 b. 我们还有

$$P(i,j(k))'AP(i,j(k)) = P(j,i(k))AP(i,j(k))$$

这相当于把 A 的第 i 行的 k 倍加到第 j 行上，接着把所得矩阵的第 i 列的 k 倍加到第 j 列上.

综上述得，若 P_i 是一个初等矩阵，则 $P_i'AP_i$ 就相当于先对 A 作一次初等行变换，接着对所得矩阵作一次同样类型的初等列变换. 我们称这是矩阵的**成对初等行、列变换**.

(23)式表明，只要对 A 施行一系列成对的初等行、列变换就可以把 A 变成对角矩阵 D. 为了同时求出所用的可逆矩阵 C，注意到(22)式也可写成

$$C = IP_1P_2\cdots P_s \tag{24}$$

把(24)与(23)式比较便知道：当我们对 A 施行成对的初等行、列变换将 A 变成了对角矩阵 D 时，对单位矩阵 I 只作其中的初等列变换就得到可逆矩阵 C. 为此我们在矩阵 A 的下方写上 I，组成一个 $2n \times n$ 矩阵，对这个矩阵的上半部分作成对的初等行、列变换，与此同时，让下半部分作其中的初等列变换，当上半部分变成对角矩阵 D 时，下半部分变成的矩阵就是我们要求的可逆矩阵 C：

$$\begin{pmatrix} A \\ I \end{pmatrix} \xrightarrow[\text{对 } I \text{ 只作其中的初等列变换}]{\text{对 } A \text{ 作成对的初等行、列变换}} \begin{pmatrix} D \\ C \end{pmatrix}$$

这种把对称矩阵经过合同变换化成对角矩阵的方法称为**矩阵的成对初等行、列变换法**.

利用矩阵的成对初等行、列变换可以把一个二次型化成标准形，这只要把二次型 $X'AX$ 的矩阵 A 按上述方法化成对角矩阵 $D = \mathrm{diag}\{d_1, d_2, \cdots, d_n\}$，并且求出一个可逆矩阵 C，则 $C'AC = D$. 令 $X = CY$，则

$$\begin{aligned} X'AX &= Y'(C'AC)Y = Y'DY \\ &= d_1y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2 \end{aligned}$$

例 3 用矩阵的成对初等行、列变换法把例 2 中所给出的二次型经过非退化线性替换化成标准形，并且写出所作的非退化线性替换：

$$f(x_1, x_2, x_3) = x_1x_2 + x_1x_3 - 3x_2x_3$$

解 $f(x_1, x_2, x_3)$ 的矩阵是

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

用矩阵的成对初等行、列变换法，我们有

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{3}{2} \\ \frac{1}{2} & -\frac{3}{2} & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{①+②\cdot 1} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -1 \\ \frac{1}{2} & 0 & -\frac{3}{2} \\ \frac{1}{2} & -\frac{3}{2} & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[①+②\cdot 1]{} \begin{pmatrix} 1 & \frac{1}{2} & -1 \\ \frac{1}{2} & 0 & -\frac{3}{2} \\ -1 & -\frac{3}{2} & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{②+①\cdot(-\frac{1}{2})} \begin{pmatrix} 1 & \frac{1}{2} & -1 \\ 0 & -\frac{1}{4} & -1 \\ -1 & -\frac{3}{2} & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[②+①\cdot(-\frac{1}{2})]{} \begin{pmatrix} 1 & 0 & -1 \\ 0 & -\frac{1}{4} & -1 \\ -1 & -1 & 0 \\ 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{③+①\cdot 1} \begin{pmatrix} 1 & 0 & -1 \\ 0 & -\frac{1}{4} & -1 \\ 0 & -1 & -1 \\ 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[③+①\cdot 1]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & -1 \\ 0 & -1 & -1 \\ 1 & -\frac{1}{2} & 1 \\ 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{③+②\cdot(-4)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & -1 \\ 0 & 0 & 3 \\ 1 & -\frac{1}{2} & 1 \\ 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[③+②\cdot(-4)]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & 3 \\ 1 & -\frac{1}{2} & 3 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

因此

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

令 $X = CY$,得

$$f(x_1, x_2, x_3) = y_1^2 - \frac{1}{4}y_2^2 + 3y_3^2$$

所作的非退化线性替换 $X = CY$ 详细写出就是

$$\begin{cases} x_1 = y_1 - \frac{1}{2}y_2 + 3y_3 \\ x_2 = y_1 + \frac{1}{2}y_2 - y_3 \\ x_3 = y_3 \end{cases}$$

比较例 2 和例 3 的结果可看出,同一个二次型,其标准形不唯一. 但是我们看到,这个例子中同一个二次型的不同的标准形,它们的系数不为零的平方项的数目相同,并且系数为正的平方项的数目是一样的,系数为负的平方项的数目也是一样的. 在下一节我们将详细讨论这个问题.

习 题 6.2

1. 用非退化线性替换把下列二次型化成标准形,并且写出所作的非退化线性替换:

(1) $f(x_1,x_2,x_3) = x_1^2 + 2x_2^2 - x_3^2 + 2x_1x_2 - 2x_1x_3$;

(2) $f(x_1,x_2,x_3) = x_1^2 - x_3^2 + 2x_1x_2 + 2x_2x_3 - 2x_1x_3$;

(3) $f(x_1,x_2,x_3) = x_1^2 - 2x_2^2 + x_3^2 - 2x_1x_2 + 4x_2x_3$;

(4) $f(x_1,x_2,x_3) = 2x_1x_2 + 2x_1x_3 + 2x_2x_3$;

(5) $f(x_1,x_2,x_3,x_4) = 2x_1x_2 - 2x_3x_4$;

(6) $f(x_1,x_2,x_3,x_4) = 4x_1x_4 + 2x_3x_4 + 2x_2x_3 + 4x_2x_4$;

2. 指出下列实二次型中哪些是等价的,要求说明理由:

$$f_1(x_1,x_2,x_3) = x_1^2 - x_2x_3$$
$$f_2(y_1,y_2,y_3) = y_1y_2 - y_3^2$$
$$f_3(z_1,z_2,z_3) = z_1z_2 + z_3^2$$

3. 证明:

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \quad 与 \quad \begin{pmatrix} \lambda_{i_1} & & & 0 \\ & \lambda_{i_2} & & \\ & & \ddots & \\ 0 & & & \lambda_{i_n} \end{pmatrix}$$

合同，其中 $i_1 i_2 \cdots i_n$ 是 $1,2,\cdots,n$ 的一个排列.

4. 证明：秩等于 r 的对称矩阵可以表示成 r 个秩等于 1 的对称矩阵之和.

5. 设 A 是数域 K 上的一个 n 级矩阵，证明：A 是斜对称矩阵的充分必要条件是对于 K^n 中任一列向量 α，有 $\alpha' A\alpha = 0$.

6. 证明：设 A 是数域 K 上的一个 n 级对称矩阵，如果对于 K^n 中任一列向量 α，有 $\alpha' A\alpha = 0$，则 $A = 0$.

7. 证明：(1) 在实数域上，$-I$ 与 I 不是合同的；(2) 在复数域上，$-I$ 与 I 合同.

*8. 用非退化线性替换化下列二次型为标准形，并且写出所作的非退化线性替换：

(1) $f(x_1,x_2,x_3) = \sum\limits_{i=1}^{3} x_i^2 + \sum\limits_{1\leqslant i<j\leqslant 3} x_i x_j$；

(2) $f(x_1,x_2,x_3,x_4) = \sum\limits_{i=1}^{4} x_i^2 + \sum\limits_{1\leqslant i<j\leqslant 4} x_i x_j$；

(3) $f(x_1,x_2,\cdots,x_n) = \sum\limits_{i=1}^{n} x_i^2 + \sum\limits_{1\leqslant i<j\leqslant n} x_i x_j$；

(4) $f(x_1,x_2,\cdots,x_n) = \sum\limits_{i=1}^{n} (x_i - s)^2$，其中

$$s = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

*9. 设

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

是一个对称矩阵，并且 A_{11} 可逆. 证明：存在可逆矩阵 C 使得，$C'AC = \mathrm{diag}\{A_{11}, B\}$，其中 B 是某个级数与 A_{22} 的级数相同的矩阵.

10. 设 A 是数域 K 上的一个斜对称矩阵；证明：A 合同于下述分块对角矩阵

$$\mathrm{diag}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, (0), \cdots, (0)\right\}$$

（提示：对斜对称矩阵的级数作第二数学归纳法.）

* 11. 证明:实数域上任一斜对称矩阵的行列式必为非负实数.

* 12. 证明:元素全为整数的斜对称矩阵的行列式一定是一个整数的平方.

§3 规范形·实(复)对称矩阵的合同分类

我们已经指出,一个二次型的标准形是不唯一的,与所作的非退化线性替换有关.这一节我们来讨论二次型的哪些量是与所作的非退化线性替换无关.换句话说,同一个等价类里的二次型有哪些共同的性质?哪些量有相同的值?在一个二次型的等价类里最简单形式的二次型是什么样?它是否唯一?

今后我们把一个二次型 $X'AX$ 的矩阵的秩就称为这个**二次型的秩**.由于合同的矩阵有相同的秩,因此等价的二次型有相同的秩.由于二次型的标准形的矩阵是对角矩阵,它的秩等于主对角线上不为零的元素的数目,因此二次型的标准形中系数不为零的平方项的个数等于这个二次型的秩,从而是唯一确定的,与所作的非退化线性替换无关.

从§2的例2与例3还看到,同一个二次型的不同的标准形中,系数为正的平方项的个数也相同,系数为负的平方项的个数也相同.这个结论对于实数域上的二次型是对的,对于复数域上的二次型却是不对的.下面我们就分别对于实数域上的二次型与复数域上的二次型进行讨论.

设 $f(x_1,x_2,\cdots,x_n)$ 是复数域上的一个二次型.设它经过一个非退化线性替换 $X=CY$ 变成的标准形是

$$d_1y_1^2+d_2y_2^2+\cdots+d_ry_r^2,\quad d_i\neq 0,\quad i=1,\cdots,r$$

其中 r 是这个二次型的秩. 因为任一复数都可以开平方,所以可以再作一个非退化线性替换:

$$\begin{cases} y_1 = \dfrac{1}{\sqrt{d_1}} z_1 \\ \cdots \quad \cdots \quad \cdots \\ y_r = \dfrac{1}{\sqrt{d_r}} z_r \\ y_{r+1} = z_{r+1} \\ \cdots \quad \cdots \quad \cdots \\ y_n = z_n \end{cases}$$

于是得

$$f(x_1, x_2, \cdots, x_n) = z_1^2 + z_2^2 + \cdots + z_r^2 \tag{1}$$

(1)式的右端称为复二次型 $f(x_1, x_2, \cdots x_n)$ 的**规范形**,它的系数为 1 或 0,并且系数为 1 的平方项的个数等于这个二次型的秩. 因此,一个复二次型的规范形完全被这个二次型的秩所决定. 于是我们证明了下面的定理:

定理 6.3.1　任意一个复二次型都可以经过一个适当的非退化线性替换化成规范形,并且规范形是唯一的. ▌

定理 6.3.1 表明,一个复二次型的等价类里具有最简单形式的二次型是规范形. 两个 n 元复二次型等价的充分必要条件是它们有相同的秩.

用矩阵的术语就是,任一复对称矩阵合同于一个形式为 $\mathrm{diag}\{1, \cdots, 1, 0, \cdots, 0\}$ 的对角矩阵. 两个 n 级复对称矩阵合同的充分必要条件是它们的秩相等. 从而秩是 n 级复对称矩阵组成的集合在合同关系下的完全不变量.

现在讨论实数域上的二次型.

设 $f(x_1, x_2, \cdots, x_n)$ 是一个实数域上的二次型,经过一个适当

非退化线性替换 $X = CY$ 可以化成如下的标准形

$$d_1y_1^2 + \cdots + d_py_p^2 - d_{p+1}y_{p+1}^2 - \cdots - d_ry_r^2 \tag{2}$$

其中 $d_i > 0$, $i = 1,\cdots,r$;并且 r 是这个二次型的秩.因为正实数总可以开平方,所以可以再作一个非退化线性替换

$$\begin{cases} y_1 = \dfrac{1}{\sqrt{d_1}}z_1 \\ \cdots\quad\cdots\quad\cdots \\ y_r = \dfrac{1}{\sqrt{d_r}}z_r \\ y_{r+1} = z_{r+1} \\ \cdots\quad\cdots\quad\cdots \\ y_n = z_n \end{cases}$$

(2)就变成

$$z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2 \tag{3}$$

(3)式称为实二次型 $f(x_1,x_2,\cdots,x_n)$ 的**规范形**,它的系数为 1,-1 或 0.显然,规范形完全被 r,p 这两个数决定.

定理 6.3.2(**惯性定理**) 实数域上任意一个二次型都可以经过一个适当的非退化线性替换化成规范形,并且规范形是唯一的.

证明 定理 6.3.2 的前半部分已在上面证明.现在证唯一性.设 $f(x_1,x_2,\cdots,x_n)$ 经过非退化线性替换 $X = CY$ 化成规范形

$$f(x_1,x_2,\cdots,x_n) = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 \tag{4}$$

又设 $f(x_1,x_2,\cdots,x_n)$ 经过非退化线性替换 $X = BZ$ 化成规范形

$$f(x_1,x_2,\cdots,x_n) = z_1^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2 \tag{5}$$

现在来证 $p = q$.用反证法.假如 $p > q$.从(4)和(5)式得

$$\begin{aligned} &y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 \\ =&z_1^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2 \end{aligned} \tag{6}$$

其中 $Z = B^{-1}CY$.设 $B^{-1}C = (g_{ij})$.为了推出矛盾,我们想取 Y 的一组值,使得(6)式的左端大于零,而右端小于或等于零.为此取

$$y_{p+1}=\cdots=y_n=0$$

此时有

$$\begin{cases} z_1=g_{11}y_1+\cdots+g_{1p}y_p \\ \cdots\quad\cdots\quad\cdots\quad\cdots \\ z_q=g_{q1}y_1+\cdots+g_{qp}y_p \\ z_{q+1}=g_{q+1,1}y_1+\cdots+g_{q+1,p}y_p \\ \cdots\quad\cdots\quad\cdots\quad\cdots \\ z_n=g_{n1}y_1+\cdots+g_{np}y_p \end{cases} \tag{7}$$

为了使(6)式右端不大于零，只要让 $y_1,\cdots,y_p$ 取一组适当的值，使得 $z_1=\cdots=z_q=0$. 为此考虑齐次线性方程组

$$\begin{cases} g_{11}y_1+\cdots+g_{1p}y_p=0 \\ \cdots\quad\cdots\quad\cdots\quad\cdots \\ g_{q1}y_1+\cdots+g_{qp}y_p=0 \end{cases} \tag{8}$$

因为 $q<p$，所以方程组(8)有非零解. 取一个非零解 $(k_1,\cdots,k_p)$，则当 $y_1=k_1,\cdots,y_p=k_p,y_{p+1}=y_{p+2}=\cdots=y_n=0$ 时，(6)式左端 $=k_1^2+\cdots+k_p^2>0$，而右端 $=-z_{q+1}^2-\cdots-z_r^2\leqslant 0$，矛盾. 因此，$p>q$ 是不对的. 这证明了 $p\leqslant q$.

同理可证 $q\leqslant p$. 因此 $p=q$. 这证明了规范形的唯一性. ▌

定义 1 在实二次型 $f(x_1,x_2,\cdots,x_n)$ 的规范形中，系数为 $+1$ 的平方项的个数 p 称为 $f(x_1,x_2,\cdots,x_n)$ 的**正惯性指数**；系数为 -1 的平方项的个数 $r-p$ 称为 $f(x_1,x_2,\cdots,x_n)$ 的**负惯性指数**；正惯性指数减去负惯性指数所得的差 $p-(r-p)=2p-r$ 称为 $f(x_1,x_2,\cdots,x_n)$ 的**符号差**.

从上面把实二次型 $f(x_1,x_2,\cdots,x_n)$ 的一个标准形化成规范形的过程可看出，标准形中系数为正的平方项个数与规范形中系数为 $+1$ 的平方项的个数是一致的；标准形中系数为负的平方项的个数与规范形中系数为 -1 的平方项个数是一致的. 因此，实二次型的标准形中系数为正的平方项的个数是唯一确定的，它等于该二次型的正惯性指数；而系数为负的平方项的个数等于负惯性

指数.

推论 6.3.1 两个 n 元实二次型等价的充分必要条件是它们的秩相等，并且它们的正惯性指数也相等. ▌

用矩阵的术语叙述上述结果就是

推论 6.3.2 任一实对称矩阵 A 合同于一个形式为 $\mathrm{diag}\{1,\cdots,1,-1,\cdots,-1,0,\cdots,0\}$ 的对角矩阵，其中 $+1$ 的个数由 A 唯一确定，称它为 A 的正惯性指数；-1 的个数也由 A 唯一确定，称它为 A 的负惯性指数. ▌

推论 6.3.3 两个 n 级实对称矩阵合同的充分必要条件是它们的秩相等，并且它们的正惯性指数也相等. ▌

推论 6.3.3 说明，秩和正惯性指数是 n 级实对称矩阵组成的集合在合同关系下的一组完全不变量.

习 题 6.3

1. n 级复对称矩阵组成的集合有多少个合同类?

2. n 级实对称矩阵组成的集合有多少个合同类?

3. n 级实对称矩阵组成的集合中，如果一个合同类既含有 A 又含有 $-A$，那么这个合同类的秩与符号差有什么特点?

*4. n 级实对称矩阵组成的集合中，符号差为给定数 s 的合同类有多少个?

5. 数域 K 上的一个 n 元**线性型**是系数在 K 中的 n 个变量的一次齐次多项式. 证明：一个实二次型可以分解成两个实线性型的乘积的充分必要条件是，它的秩等于 2 并且符号差等于 0，或者秩等于 1.

6. 设 A 为一个 n 级实对称矩阵，且 $|A|<0$. 证明：在 R^n 中有非零列向量 α，使得 $\alpha' A\alpha<0$.

7. 把习题 6.2 的第 1 题的(1)，(3)，(4)中的二次型化为规范形(分别考虑复数域和实数域的情形)，并且写出所作的非退化线性替换.

8. 设 $f(x_1,x_2,\cdots,x_n)=X'AX$ 是一个实二次型. 证明：如果 R^n 中有列

向量 α_1,α_2,使得

$$\alpha_1'A\alpha_1 > 0,\qquad \alpha_2'A\alpha_2 < 0$$

则 R^n 中一定有非零列向量 α_3,使得 $\alpha_3'A\alpha_3 = 0$.

*9. 设实二次型

$$f(x_1,x_2,\cdots,x_n) = l_1^2 + \cdots + l_s^2 - l_{s+1}^2 - \cdots - l_{s+u}^2$$

其中 $l_i(i=1,2,\cdots,s+u)$ 是 $x_1,x_2,\cdots,x_n$ 的实线性型. 证明:$f(x_1,x_2,\cdots,x_n)$ 的正惯性指数 $p \leqslant s$,负惯性指数 $q \leqslant u$. (提示:类似于惯性定理的证明方法.)

§4 用正交替换化实二次型为标准形

在本章引言中我们曾指出,为了把一个二次曲面方程化成标准方程,首先需要作直角坐标变换把方程中的二次项部分(x,y,z 的二次型)化成只含平方项的形式. 由于一个直角坐标系到另一个直角坐标系的过渡矩阵是正交矩阵,因此要求所作的线性替换的系数矩阵是正交矩阵.

如果线性替换的系数矩阵是正交矩阵,则称它为**正交的线性替换**,简称为**正交替换**.

把一个 n 元实二次型 $X'AX$ 通过正交替换 $X=TY$ 化成标准形 $d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2$,这相当于对实对称矩阵 A 找一个正交矩阵 T,使得 $T'AT=\mathrm{diag}\{d_1,d_2,\cdots,d_n\}$. 注意到正交矩阵 T 的转置 $T'=T^{-1}$,因此 $T'AT=\mathrm{diag}\{d_1,d_2,\cdots,d_n\}$ 也就是 $T^{-1}AT=\mathrm{diag}\{d_1,d_2,\cdots,d_n\}$. 而后者意味着 A 正交相似于对角矩阵 $\mathrm{diag}\{d_1,d_2,\cdots,d_n\}$. 在上一章 §9 已指出,实对称矩阵一定能正交相似于一个对角矩阵,并且对角矩阵的主对角线上元素是 A 的全部特征值;还给出了正交矩阵 T 的求法. 因此,我们一定能用正交替换把实二次型化成标准形. 即我们有下面的结果:

定理 6.4.1 对于实二次型 $f(x_1,x_2,\cdots,x_n)=X'AX$,一定

能找到一个正交矩阵 T，使得经过正交替换 $X = TY$ 把它化成标准形

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \tag{1}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是该二次型的矩阵 A 的全部特征值. ▌

用正交替换把实二次型化成标准形，这在理论上和实际应用上都是非常重要的.

例 1 作直角坐标变换，把下述二次曲面方程化成标准方程，并且指出它是什么曲面？

$$x^2 + 4y^2 + z^2 - 4xy - 8xz - 4yz + 2x + y + 2z - \frac{25}{16} = 0 \tag{2}$$

解 首先把二次曲面方程(2)的二次项部分通过正交替换化成只含平方项的形式. 设

$$f(x,y,z) = x^2 + 4y^2 + z^2 - 4xy - 8xz - 4yz \tag{3}$$

二次型(3)的矩阵是

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

因为

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & 2 & 4 \\ 2 & \lambda - 4 & 2 \\ 4 & 2 & \lambda - 1 \end{vmatrix} = (\lambda - 5)^2(\lambda + 4)$$

所以 A 的特征值是 5(二重)，-4.

对于特征值 5，求得齐次线性方程组 $(5I - A)X = 0$ 的一个基础解系为：

$$\alpha_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

把 α_1,α_2 正交化：令 $\beta_1=\alpha_1$，

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1=\begin{pmatrix}\frac{4}{5}\\ \frac{2}{5}\\ -1\end{pmatrix}$$

把 β_1,β_2 单位化：

$$\eta_1=\frac{1}{|\beta_1|}\beta_1=\begin{pmatrix}\frac{1}{5}\sqrt{5}\\ -\frac{2}{5}\sqrt{5}\\ 0\end{pmatrix}、$$

$$\eta_2=\frac{1}{|\beta_2|}\beta_2=\begin{pmatrix}\frac{4}{15}\sqrt{5}\\ \frac{2}{15}\sqrt{5}\\ -\frac{1}{3}\sqrt{5}\end{pmatrix}$$

对于特征值 -4，求得齐次线性方程组 $(-4I-A)X=0$ 的一个基础解系为

$$\alpha_3=\begin{pmatrix}2\\1\\2\end{pmatrix}$$

把 α_3 单位化：

$$\eta_3=\frac{1}{|\alpha_3|}\alpha_3=\begin{pmatrix}\frac{2}{3}\\ \frac{1}{3}\\ \frac{2}{3}\end{pmatrix}$$

令

$$T = \begin{bmatrix} \frac{1}{5}\sqrt{5} & \frac{4}{15}\sqrt{5} & \frac{2}{3} \\ -\frac{2}{5}\sqrt{5} & \frac{2}{15}\sqrt{5} & \frac{1}{3} \\ 0 & -\frac{1}{3}\sqrt{5} & \frac{2}{3} \end{bmatrix}$$

则 T 是正交矩阵，并且

$$T^{-1}AT = \mathrm{diag}\{5,5,-4\}$$

于是作正交替换

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = T \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \tag{4}$$

可以把二次型(3)化成下述标准形：

$$f(x,y,z) = 5x'^2 + 5y'^2 - 4z'^2 \tag{5}$$

因此，作直角坐标变换(4)，二次曲面(2)的新方程为

$$5x'^2 + 5y'^2 - 4z'^2 + 3z' - \frac{25}{16} = 0 \tag{6}$$

将(6)的左端配方得

$$5x'^2 + 5y'^2 - 4(z' - \frac{3}{8})^2 - 1 = 0 \tag{7}$$

作移轴

$$\begin{cases} x' = x^* \\ y' = y^* \\ z' = z^* + \frac{3}{8} \end{cases} \tag{8}$$

得

$$5x^{*2} + 5y^{*2} - 4z^{*2} = 1 \tag{9}$$

方程(9)是二次曲面(2)的标准方程，它是单叶双曲面. 从方程(2)

到方程(9)的直角坐标变换公式为

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{5}\sqrt{5} & \frac{4}{15}\sqrt{5} & \frac{2}{3} \\ -\frac{2}{5}\sqrt{5} & \frac{2}{15}\sqrt{5} & \frac{1}{3} \\ 0 & -\frac{1}{3}\sqrt{5} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix} + \begin{pmatrix} \frac{1}{4} \\ \frac{1}{8} \\ \frac{1}{4} \end{pmatrix}$$

* **例** 2　设 n 级实对称矩阵 A 的全部特征值按大小顺序排成：$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. 证明：对于 R^n 中任一列向量 $\alpha \neq 0$，都有

$$\lambda_n \leqslant \frac{\alpha' A\alpha}{|\alpha|^2} \leqslant \lambda_1$$

证明　因为 A 是实对称矩阵，所以有正交矩阵 T，使得 $T^{-1}AT = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$. 任取 R^n 中一个非零列向量 α，设 $(T'\alpha)' = (b_1, b_2, \cdots, b_n)$. 我们有

$$\begin{aligned} \alpha' A\alpha &= \alpha'(T\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}T')\alpha \\ &= (T'\alpha)'\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}(T'\alpha) \\ &= \lambda_1 b_1^2 + \lambda_2 b_2^2 + \cdots + \lambda_n b_n^2 \\ &\leqslant \lambda_1(b_1^2 + b_2^2 + \cdots + b_n^2) \\ &= \lambda_1 |T'\alpha|^2 \end{aligned}$$

同理，我们有 $\alpha' A\alpha \geqslant \lambda_n |T'\alpha|^2$. 因为

$$|T'\alpha|^2 = (T'\alpha, T'\alpha) = (T'\alpha)'(T'\alpha) = \alpha' TT'\alpha = \alpha'\alpha = |\alpha|^2$$

所以

$$\lambda_n \leqslant \frac{\alpha' A\alpha}{|\alpha|^2} \leqslant \lambda_1 \qquad \blacksquare$$

习　题　6.4

1. 用正交替换把下列实二次型化成标准形，并且写出所作的正交替换：

(1) $f(x_1,x_2,x_3)=3x_2^2+3x_3^2+4x_1x_2-4x_1x_3-8x_2x_3$;

(2) $f(x_1,x_2,x_3,x_4)=x_1^2+x_2^2+x_3^2+x_4^2+2x_1x_2-2x_3x_4$;

2. 作直角坐标变换，把下列二次曲面方程化为标准方程，并且指出它是什么曲面？

(1) $x^2+2z^2+3y^2-4xy-4yz=1$;

(2) $2x^2+6y^2+2z^2+8xz=1$.

3. 设 A 是一个 n 级实对称矩阵，证明：存在一个正实数 c，使得对于 R^n 中任一列向量 α，都有 $|\alpha' A\alpha|\leqslant c\alpha'\alpha$.

*4. 设 B 是 n 级实矩阵，设 $B'B$ 的全部特征值按大小顺序排成 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$. 证明：如果 B 有特征值，则 B 的任一特征值 μ 满足：

$$\sqrt{\lambda_n}\leqslant|\mu|\leqslant\sqrt{\lambda_1}$$

5. 设实二次型 $f(x_1,x_2,\cdots,x_n)$ 的矩阵为 A，设 λ_i 是 A 的一个特征值. 证明：存在 R^n 中的非零向量 $\alpha=(a_1,a_2,\cdots,a_n)$，使得

$$f(a_1,a_2,\cdots,a_n)=\lambda_i(a_1^2+a_2^2+\cdots+a_n^2)$$

6. 设 A,B 都是 n 级实对称矩阵，并且 $AB=BA$. 证明：存在一个 n 级正交矩阵，使得 $T'AT$ 与 $T'BT$ 都为对角矩阵.（提示：因为 A 实对称，所以有正交矩阵 T_1，使

$$T_1^{-1}AT_1=\operatorname{diag}\{\lambda_1I_{r_1},\lambda_2I_{r_2},\cdots,\lambda_mI_{r_m}\}$$

其中 $\lambda_1,\cdots,\lambda_m$ 是 A 的全部不同的特征值. 由已知条件：$AB=BA$ 可得出，

$$(T_1^{-1}AT_1)(T_1^{-1}BT_1)=(T_1^{-1}BT_1)(T_1^{-1}AT_1)$$

从而可推出

$$T_1^{-1}BT_1=\operatorname{diag}\{B_1,B_2,\cdots,B_m\}$$

其中每个 B_i 是实对称矩阵.）

§5 正定二次型与正定矩阵

函数的极值问题有广泛的实际应用. 读者在中学学习了二次函数的极值，在数学分析课程中又进一步学习了一元函数的极值. 譬如，二次函数 $f(x)=x^2+1$ 在 $x=0$ 处达到最小值 1，而 $g(x)$

$=-x^2+1$ 在 $x=0$ 处达到最大值1,这些从它们的图象可以立即看出.一般地,设 $f(x)$ 在 x_0 的邻域 $U(x_0;\delta)$ 上可导,且 $f'(x_0)=0$,$f''(x_0)$ 存在,则据泰勒公式得

$$f(x)-f(x_0)=\left[\frac{1}{2}f''(x_0)+o(1)\right](x-x_0)^2 \qquad (1)$$

因为 $x\to x_0$ 时,$o(1)$ 是无穷小量,所以存在 $\delta_1\in(0,\delta)$,使得当 $x\in U(x_0;\delta_1)$ 时,$\frac{1}{2}f''(x_0)+o(1)$ 的符号与 $f''(x_0)$ 的符号相同.令 $t=x-x_0$,考虑变量 t 的二次型

$$f''(x_0)t^2 \qquad (2)$$

当 $f''(x_0)>0$ 时,对于任一非零实数 c,有

$$f''(x_0)c^2>0$$

从而对一切 $x\in U(x_0;\delta_1)$ 且 $x\neq x_0$,有

$$f(x)-f(x_0)>0$$

因此,当 $f''(x_0)>0$时,函数 $f(x)$ 在 $x=x_0$ 处达到严格极小值.类似地可以推出,当 $f''(x_0)<0$时,$f(x)$ 在 x_0 处达到严格极大值.上述推导过程有两步是关键的:一步是利用泰勒公式,把 $f(x)-f(x_0)$ 表示成(1)式右端的形式;另一步是研究实二次型 $f''(x_0)t^2$ 是否具有性质:"对任一非零实数 c,有 $f''(x_0)c^2$ 恒大于零(或者恒小于零)".由此受到启发,为了研究多元函数 $f(x_1,x_2,\cdots,x_n)$ 的极值问题,也有关键的两步:一步是需要有多元函数的泰勒公式,读者在数学分析课程中将会学到.另一步是需要研究实二次型 $g(t_1,t_2,\cdots,t_n)$ 是否具有性质:"对于任意一组不全为零的实数 $c_1,c_2,\cdots,c_n$,有 $g(c_1,c_2,\cdots,c_n)$ 恒大于零(或者恒小于零)".我们现在就来讨论这第二步提出的问题.

定义 1 实二次型 $f(x_1,x_2,\cdots,x_n)=X'AX$ 称为**正定的**,如果对于任意一组不全为零的实数 $c_1,c_2,\cdots,c_n$ 都有

$$f(c_1,c_2,\cdots,c_n)>0$$

换句话说,如果对于一切 $\alpha\in R^n$ 且 $\alpha\neq 0$,有

$$\alpha' A\alpha > 0$$

例 1 判断下列 3 元实二次型是否正定：

(1) $f(y_1,y_2,y_3) = y_1^2 + y_2^2 + y_3^2$；

(2) $g(y_1,y_2,y_3) = y_1^2 + y_2^2$；

(3) $h(y_1,y_2,y_3) = y_1^2 + y_2^2 - y_3^2$.

解 (1) 因为对于任意一组不全为零的实数 c_1,c_2,c_3，都有 $f(c_1,c_2,c_3) = c_1^2 + c_2^2 + c_3^2 > 0$，所以 $f(y_1,y_2,y_3)$ 是正定的.

(2) 因为 $g(0,0,1) = 0$，所以 $g(y_1,y_2,y_3)$ 不是正定的.

(3) 因为 $h(0,0,1) = -1$，所以 $h(y_1,y_2,y_3)$ 不是正定的.

例 1 中的三个二次型都是规范形，从中容易看出，当正惯性指数等于变量个数时，该二次型是正定的；否则，就不是正定的. 这个结论对于任意一个实二次型都成立. 下面我们来证明它.

定理 6.5.1 n 元实二次型 $X'AX$ 是正定的充分必要条件为，它的正惯性指数等于 n.

证明 充分性. 设 $X'AX$ 的正惯性指数等于变量个数 n，则可以作非退化线性替换 $X = CY$，把 $X'AX$ 化成如下的规范形

$$y_1^2 + y_2^2 + \cdots + y_n^2$$

任取 $\alpha \in R^n$ 且 $\alpha \neq 0$，记 $C^{-1}\alpha = \beta = (b_1,b_2,\cdots,b_n)'$，则 $\alpha = C\beta$ 且 $\beta \neq 0$. 于是有

$$\alpha' A\alpha = b_1^2 + b_2^2 + \cdots + b_n^2 > 0$$

所以 $X'AX$ 是正定的.

必要性. 设 $X'AX$ 是正定的. 作非退化线性替换 $X = CY$，把 $X'AX$ 化成规范形 $g(y_1,y_2,\cdots,y_n)$. 假如 $X'AX$ 的正惯性指数不等于 n，则 $g(y_1,y_2,\cdots,y_n)$ 中的 y_n^2 的系数为 0 或 -1. 取 $\beta = (0,\cdots,0,1)'$，令 $\alpha = C\beta$. 显然，$\alpha \neq 0$，并且有

$$\alpha' A\alpha = g(0,\cdots,0,1) = 0 \text{ 或 } -1$$

这与 $X'AX$ 是正定的矛盾. ▌

推论 6.5.1

n 元实二次型 $X'AX$ 是正定的

$\Leftrightarrow$它的规范形为 $y_1^2+y_2^2+\cdots+y_n^2$

$\Leftrightarrow$它的标准形中 n 个系数全大于零 ▌

下面我们来讨论正定二次型的矩阵的特性.

定义 2 实对称矩阵 A 称为**正定的**,如果二次型 $X'AX$ 是正定的.

正定的实对称矩阵简称为正定矩阵.

从定义 2 和定义 1 立即得出:

命题 6.5.1

n 级实对称矩阵 A 是正定的

$\Leftrightarrow \forall\ \alpha\in R^n$ 且 $\alpha\neq 0$、有 $\alpha' A\alpha>0$ ▌

由于两个 n 元二次型等价当且仅当它们的矩阵合同,因此从定理 6.5.1 和推论 6.5.1,以及定理 6.4.1 立即得出:

定理 6.5.2

n 级实对称矩阵 A 是正定的

$\Leftrightarrow A$ 的正惯性指数等于 n

$\Leftrightarrow A\simeq I$

$\Leftrightarrow A$ 合同于主对角元全大于零的对角矩阵

$\Leftrightarrow A$ 的特征值全大于零 ▌

从定理 6.5.2 看出,n 级实对称矩阵是正定的当且仅当它的正惯性指数等于 n 并且秩等于 n. 由于在所有 n 级实对称矩阵组成的集合 S 中,由正惯性指数等于 n 且秩等于 n 的所有矩阵组成的子集是 S 的一个合同类. 因此所有 n 级正定矩阵组成 S 的一个合同类. 在这个合同类中具有最简单形式的矩阵是单位矩阵. 由上述立即得到:

推论 6.5.2 与正定矩阵合同的实对称矩阵也是正定矩阵. ▌

从推论 6.5.2 立即得出:

推论 6.5.3 与正定二次型等价的实二次型也是正定的. 从而非退化实线性替换不改变实二次型的正定性. ▌

现在我们从子式的角度来研究正定矩阵. 先给出两个概念:

定义 3 设 A 是一个 n 级矩阵. A 的一个子式称为**主子式**,如果它的行指标与列指标相同,即它形如

$$A\begin{pmatrix} i_1, & i_2, & \cdots, & i_k \\ i_1, & i_2, & \cdots, & i_k \end{pmatrix}$$

其中 $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$.

对于 $1 \leqslant k \leqslant n$, A 的下述 k 级主子式

$$A\begin{pmatrix} 1, & 2, & \cdots, & k \\ 1, & 2, & \cdots, & k \end{pmatrix}$$

称为 A 的 k 级**顺序主子式**.

例如,设

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

则 A 的顺序主子式共有 3 个,如下述:

$$|-1|, \quad \begin{vmatrix} -1 & 0 \\ 7 & 2 \end{vmatrix}, \quad \begin{vmatrix} -1 & 0 & 3 \\ 7 & 2 & -4 \\ 5 & 8 & 6 \end{vmatrix}$$

单位矩阵 I 是最简单的正定矩阵,显然, I 的所有顺序主子式全大于零. 一般地,我们有

定理 6.5.3 实对称矩阵 A 是正定的充分必要条件为 A 的所有顺序主子式全大于零.

证明 必要性. 设 A 是 n 级正定矩阵,则 $A \simeq I$. 于是存在实可逆矩阵 C,使得

$$A = C'IC = C'C$$

对于 $1 \leqslant k \leqslant n$,据命题 4.8.1 得

$$A\begin{pmatrix} 1, & \cdots, & k \\ 1, & \cdots, & k \end{pmatrix}$$

$$=(C'C)\begin{pmatrix}1, & \cdots, & k\\ 1, & \cdots, & k\end{pmatrix}$$

$$=\sum_{1\leqslant v_1<\cdots<v_k\leqslant n} C'\begin{pmatrix}1, & \cdots, & k\\ v_1, & \cdots, & v_k\end{pmatrix} C\begin{pmatrix}v_1, & \cdots, & v_k\\ 1, & \cdots, & k\end{pmatrix}$$

$$=\sum_{1\leqslant v_1<\cdots<v_k\leqslant n}\left[C\begin{pmatrix}v_1, & \cdots, & v_k\\ 1, & \cdots, & k\end{pmatrix}\right]^2 \qquad (3)$$

因为 C 可逆，所以 C 的列向量组线性无关. 从而 C 的前 k 列组成的子矩阵 C_1 的秩为 k. 因此 C_1 至少有一个 k 级子式不为 0. 而 C_1 的 k 级子式是 C 的前 k 列元素组成的 k 级子式，因此(3) 式右端中至少有一项不为 0. 于是(3) 式右端大于 0. 这证明了 A 的 k 级顺序主子式大于 0.

充分性. 对于实对称矩阵的级数 n 作归纳法.

当 $n=1$ 时，1 级实对称矩阵为(a). 由已知条件 $a>0$，因此(a) 是正定矩阵.

假设对于 $n-1$ 级实对称矩阵，定理 6.5.3 的充分性成立. 现在来看 n 级实对称矩阵 A. 假设 A 的顺序主子式全大于零，要证 A 正定. 把 A 分块写成

$$A=\begin{pmatrix}A_{n-1} & \alpha\\ \alpha' & a_{nn}\end{pmatrix}$$

其中 A_{n-1} 是 $n-1$ 级实对称矩阵. 显然，A_{n-1} 的所有顺序主子式是 A 的 1 级到 $n-1$ 级的顺序主子式，由已知条件得，A_{n-1} 的顺序主子式全大于零. 据归纳假设得，A_{n-1} 正定. 于是有 $n-1$ 级实可逆矩阵 C_1，使得

$$C_1{}'A_{n-1}C_1=I_{n-1} \qquad (4)$$

由于

$$\begin{pmatrix}A_{n-1} & \alpha\\ \alpha' & a_{nn}\end{pmatrix}\xrightarrow{②+(-\alpha' A_{n-1}^{-1})\cdot ①}\begin{pmatrix}A_{n-1} & \alpha\\ 0 & a_{nn}-\alpha' A_{n-1}^{-1}\alpha\end{pmatrix}$$

$$\xrightarrow[②+①\cdot(-A_{n-1}^{-1}\alpha)]{}\begin{pmatrix}A_{n-1} & 0\\ 0 & a_{nn}-\alpha' A_{n-1}^{-1}\alpha\end{pmatrix}$$

所以

$$\begin{pmatrix} I_{n-1} & -A_{n-1}^{-1}\alpha \\ 0 & 1 \end{pmatrix}' \begin{pmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{pmatrix} \begin{pmatrix} I_{n-1} & -A_{n-1}^{-1}\alpha \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} A_{n-1} & 0 \\ 0 & a_{nn} - \alpha' A_{n-1}^{-1}\alpha \end{pmatrix} \tag{5}$$

记 $b = a_{nn} - \alpha' A^{-1}\alpha$. 于是由(5)得

$$A \simeq \begin{pmatrix} A_{n-1} & 0 \\ 0 & b \end{pmatrix} \tag{6}$$

由(5)式还可得到

$$|A| = |A_{n-1}| b$$

由于 $|A| > 0$,并且 $|A_{n-1}| > 0$,所以 $b > 0$. 因为

$$\begin{pmatrix} C_1 & 0 \\ 0 & 1 \end{pmatrix}' \begin{pmatrix} A_{n-1} & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} C_1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} C_1' A_{n-1} C_1 & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} I_{n-1} & 0 \\ 0 & b \end{pmatrix}$$

所以

$$\begin{pmatrix} A_{n-1} & 0 \\ 0 & b \end{pmatrix} \simeq \begin{pmatrix} I_{n-1} & 0 \\ 0 & b \end{pmatrix}$$

由合同的传递性得

$$A \simeq \begin{pmatrix} I_{n-1} & 0 \\ 0 & b \end{pmatrix} \tag{7}$$

(7)式右端是正定矩阵,从而 A 也正定.

根据归纳法原理,充分性得证. ▌

从定理 6.5.3 立即得出:

推论 6.5.4 正定矩阵的行列式大于零. ▌

例 2 判别下列实二次型是否正定:

(1) $f(x_1, x_2, x_3) = 4x_1^2 + 5x_2^2 + 6x_3^2 + 4x_1x_2 - 4x_2x_3$;

(2) $f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 - 3x_3^2 + 4x_1x_2 + 2x_2x_3$.

解 (1) $f(x_1, x_2, x_3)$ 的矩阵是

$$A = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 5 & -2 \\ 0 & -2 & 6 \end{pmatrix}$$

因为

$$|4| = 4 > 0, \quad \begin{vmatrix} 4 & 2 \\ 2 & 5 \end{vmatrix} = 16 > 0, \quad \begin{vmatrix} 4 & 2 & 0 \\ 2 & 5 & -2 \\ 0 & -2 & 6 \end{vmatrix} = 80 > 0$$

所以 A 正定. 从而二次型 $X'AX$ 正定.

(2)　$f(x_1, x_2, x_3)$ 的矩阵是

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

因为

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

所以 A 不是正定的. 从而二次型 $X'AX$ 不是正定的.

上面我们介绍了实对称矩阵 A 是正定的六个充分必要条件，分别从二次型、正惯性指数、合同规范形、合同标准形、特征值、顺序主子式六个角度看. 希望读者能熟练掌握它们，并且能灵活运用. 从这六个充分必要条件还可以推导出实对称矩阵是正定的其它一些充分必要条件. 下面的例 3 是从矩阵的分解的角度给出一个充分必要条件：

例 3　证明：实对称矩阵 A 是正定的充分必要条件为有实可逆矩阵 C，使得

$$A = C'C$$

证明　实对称矩阵 A 正定 $\Leftrightarrow A \simeq I \Leftrightarrow$ 有实可逆矩阵 C，使得 $A = C'IC = C'C$. ▌

本节习题中我们还将给出正定矩阵的一些充分必要条件.

实二次型中除了正定二次型这一类型外，还有其它一些类型：

定义 4　设 $f(x_1,x_2,\cdots,x_n)$ 是一个实二次型，对于任意一组不全为零的实数 $c_1,c_2,\cdots,c_n$，如果都有 $f(c_1,c_2,\cdots,c_n)\geqslant 0$，那么 $f(x_1,x_2,\cdots,x_n)$ 称为**半正定的**；如果都有 $f(c_1,c_2,\cdots,c_n)<0$，则 $f(x_1,x_2,\cdots,x_n)$ 称为**负定的**；如果都有 $f(c_1,c_2,\cdots,c_n)\leqslant 0$，则 $f(x_1,x_2,\cdots,x_n)$ 称为**半负定的**；如果它既不是半正定又不是半负定的，则 $f(x_1,x_2,\cdots,x_n)$ 称为**不定的**.

定义 5　实对称矩阵 A 称为**半正定的**（**负定的**、**半负定的**、**不定的**）如果实二次型 $X'AX$ 是半正定的（负定的、半负定的、不定的）.

例 4　判别下列 3 元实二次型属于哪种类型：

(1)　$y_1^2+y_2^2$；

(2)　y_1^2；

(3)　$y_1^2+y_2^2-y_3^2$；

(4)　$-y_1^2-y_2^2-y_3^2$；

(5)　$-y_1^2-y_2^2$.

解　(1)　半正定；(2)　半正定；(3)　不定；(4)　负定；(5)　半负定.

从例 4 受到启发，一般地，有

定理 6.5.4

n 元实二次型 $X'AX$ 是半正定的

$\Leftrightarrow$它的正惯性指数等于它的秩

$\Leftrightarrow$它的规范形是 $y_1^2+\cdots+y_r^2$，其中 $1\leqslant r\leqslant n$

$\Leftrightarrow$它的标准形中 n 个系数全非负

证明类似于定理 6.5.1 的证明，请读者自己写出.

从定义 5 和定理 6.5.4 立即得出：

定理 6.5.5

n 级实对称矩阵 A 是半正定的

$\Leftrightarrow\forall\ \alpha\in R^n$ 且 $\alpha\neq 0$，有 $\alpha'A\alpha\geqslant 0$

$\Leftrightarrow A$ 的正惯性指数等于 $\mathrm{rank}A$

$\Leftrightarrow A \simeq \mathrm{diag}\{I_r, 0\}$，其中 $r = \mathrm{rank} A$

$\Leftrightarrow A$ 合同于主对角元全非负的对角矩阵

$\Leftrightarrow A$ 的特征值全非负 ▌

从子式的角度如何判定实对称矩阵是不是半正定的？我们有

定理 6.5.6　实对称矩阵 A 是半正定的充分必要条件为 A 的所有主子式全非负.

证明　必要性的证明类似于定理 6.5.3 的证法. 充分性的证明要利用命题 5.6.1，去证 A 的特征值全非负. 细节留给读者. ▌

注意，实对称矩阵 A 如果只有顺序主子式全非负，则 A 不一定是半正定的. 例如

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

显然 A 的顺序主子式都为零，但是 A 不是半正定的.

由于实二次型 $f(x_1, \cdots, x_n)$ 是负定的（或半负定的）充分必要条件为 $-f(x_1, \cdots, x_n)$ 是正定的（或半正定的），因此容易得出负定（或半负定）二次型以及负定（或半负定）矩阵的判别条件.

习　题　6.5

1.　判断下列实二次型是否正定？

（1）　$7x_1^2 + 8x_2^2 + 6x_3^2 - 4x_1x_2 - 4x_2x_3$；

（2）　$2x_1^2 + 8x_1x_2 + 4x_1x_3 + 2x_2^2 - 8x_2x_3 + x_3^2$；

（3）　$X'(I + J)X$，其中 J 是元素全为 1 的 n 级矩阵.

2.　证明：n 元实二次型 $f(x_1, x_2, \cdots, x_n)$ 为正定的必要条件是，它的 n 个平方项的系数全是正的. 举例说明这个条件不是 $f(x_1, x_2, \cdots, x_n)$ 为正定的充分条件.

3.　证明：若 A 是正定矩阵，则 A^{-1} 也是正定矩阵.

4.　证明：若 A 是正定矩阵，则 A 的伴随矩阵 A^* 也是正定矩阵.

5.　证明：若 A 与 B 都是 n 级正定矩阵，则 $A + B$ 也是正定矩阵.

6. 设 A 是实对称矩阵，它的 n 个特征值的绝对值中最大者记作 $S_r(A)$. 证明：当实数 $t > S_r(A)$ 时，$tI + A$ 是正定矩阵.

7. 证明：实对称矩阵 A 是正定矩阵的充分必要条件为 A 的所有主子式都大于零.

*8. 证明：实对称矩阵 A 正定的充分必要条件为：有实上三角矩阵 B 并且 B 的主对角元全大于零，使得 $A = B'B$.

*9. 证明：n 级实对称矩阵 A 正定的充分必要条件为：有 $m \times n$ 列满秩实矩阵 P 使得 $A = P'P$.

10. 证明：实对称矩阵 A 正定的充分必要条件为：有可逆实对称矩阵 C 使得 $A = C^2$.

11. 证明：若 A 正定，则存在一个正定矩阵 C 使得 $A = C^2$.

12. 设 A 是 n 级实对称矩阵，B 是 n 级正定矩阵. 证明：存在一个 n 级实可逆矩阵 C 使得 $C'AC$ 与 $C'BC$ 都是对角矩阵.

13. 证明本节定理 6.5.4.

14. 证明本节定理 6.5.6.

15. 证明：$n\sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{i=1}^{n} x_i)^2$ 是半正定的.

*16. 证明：n 级实对称矩阵 A 是半正定的充分必要条件为：有 $r \times n$ 行满秩实矩阵 Q 使得 $A = Q'Q$.

*17. 证明：实对称矩阵 A 半正定的充分必要条件为：有实对称矩阵 C 使得 $A = C^2$.

18. 证明：实二次型 $f(x_1, x_2, \cdots, x_n)$ 负定的充分必要条件为：它的矩阵 A 的奇数级顺序主子式全小于零，偶数级顺序主子式全大于零.

*19. 证明：若 A 与 B 都是 n 级正定矩阵，则 AB 是正定矩阵的充分必要条件为 A 与 B 可交换.（提示：充分性利用习题 6.4 第 6 题结论）.

*20. 证明：若 $A \neq 0$ 是 n 级半正定矩阵，S 是 n 级正定矩阵，则 $|A + S| > |S|$，且 $|A + S| > |A|$.（提示：利用第 12 题结果）.

21. 设

$$M = \begin{pmatrix} A & B \\ B' & D \end{pmatrix}$$

是 n 级正定矩阵，其中 A 是 r 级方阵（$r < n$）. 证明 $A, D, D - B'A^{-1}B$ 都是正定矩阵.

* 22. 设 $\begin{pmatrix} A & B \\ B' & D \end{pmatrix}$ 是正定矩阵,其中 A,D 均为方阵.证明:

$$\begin{vmatrix} A & B \\ B' & D \end{vmatrix} \leqslant |A||D|$$

并且等号成立当且仅当 $B=0$.(提示:利用第 21 题、第 20 题结果).

补 充 题 六

1. 证明:若 A 正定,则存在唯一的正定矩阵 C 使得 $A=C^2$.(注:存在性是习题 6.5 的第 11 题).

2. 证明 **极分解定理**:对于任一实可逆矩阵 A,一定存在一个正交矩阵 T 和两个正定矩阵 S_1 与 S_2,使得

$$A=TS_1=S_2T$$

并且这种分解是唯一的.

3. 证明:对于任一 n 级实可逆矩阵 A,都存在正交矩阵 T_1 和 T_2,使得

$$A=T_1\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}T_2$$

并且 $\lambda_1^2,\lambda_2^2,\cdots,\lambda_n^2$ 是 $A'A$ 的特征值.(提示:用第 1 题和第 2 题结论).

4. 证明:几何空间中任一仿射变换可以分解成一些正交变换与沿着三个互相垂直的方向的压缩的乘积.

5. 证明:如果对称矩阵 A 的顺序主子式全不为零,则一定有主对角元全为 1 的上三角矩阵 T 与主对角元全不为零的对角矩阵 D,使得

$$A=T'DT$$

并且 A 的这种分解式是唯一的.(提示:存在性类似于定理 6.5.3 的证法).

6. 证明:

(1) 如果 A 是 n 级正定矩阵,则对于 Rn 中任一非零列向量 α,有

$$\begin{vmatrix} A & \alpha \\ \alpha' & 0 \end{vmatrix} < 0$$

(2) 如果 $A=(a_{ij})$ 是 n 级正定矩阵,则

$$|A| \leqslant a_{nn}A\begin{pmatrix} 1, & 2, & \cdots, & n-1 \\ 1, & 2, & \cdots, & n-1 \end{pmatrix}$$

(3) 如果 $C=(c_{ij})$ 是 n 级实可逆矩阵，则

$$|C|^2 \leqslant \prod_{j=1}^{n}(c_{1j}^2+c_{2j}^2+\cdots c_{nj}^2)$$

7. 设 $A=(a_{ij})$ 是 n 级正定矩阵，$b_1,b_2,\cdots,b_n$ 是任意 n 个非零实数，证明：$C=(a_{ij}b_ib_j)$ 是正定矩阵.

* 8. 证明：若 n 级矩阵 $A=(a_{ij}),B=(b_{ij})$ 都是正定的，则矩阵 $C=(a_{ij}b_{ij})$ 也是正定的.

* 9. 正定(负定)矩阵在多元函数极值问题中的应用：

设 $F(X)=F(x_1,x_2,\cdots,x_n)$ 有三阶连续偏导数，设 $\alpha=(a_1,a_2,\cdots,a_n)$，如果 $F(X)$ 在 α 处的一阶偏导数全为 0，则称 α 是 $F(X)$ 的一个稳定点. 令

$$H=(F_{x_ix_j}(\alpha))$$

其中 α 是 $F(X)$ 的一个稳定点，称 H 是函数 $F(X)$ 在 α 处的 Hesse 矩阵. 如果 H 是正定的，则 $F(X)$ 在 α 处达到极小值；如果 H 是负定的，则 $F(X)$ 在 α 处达到极大值；如果 H 是不定的，则 $F(X)$ 在 α 处既不是极大，也不是极小，这时称 α 是 $F(X)$ 的鞍点. 试以二元函数 $F(x,y)$ 为例证明上述结论.（提示：用泰勒公式.）

* 10. 求 $F(x,y)=6xy-x^3-y^3$ 的极值.

［Ｇｅｎｅｒａｌ Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝高等代数（上册）
作者＝丘维声　　编著
页数＝３４４
ＳＳ号＝１０６５２８９０
出版日期＝１９９６年０６月第１版