华东师范大学 网络教育学院

数学与应用数学专业

线 性 代 数
讲 义

袁 海 荣

2009 年 12 月 18 日

目 录

说 明

这是基于教材[1]第一、二、三、四章内容，加以适当扩充后，为华东师范大学网络教育学院数学与应用数学专业《线性代数》课程准备的讲义. 考虑到网络教育学院学员多为中学一线教师的实际情况，本讲义仅完整地介绍了求解线性方程组的基本理论，而对这一理论的应用，如一般线性代数教材包括的矩阵对角化、二次型等内容予以省略.

除了[1]外，本讲义部分内容还取自、或参考了教材[2, 3]、习题集或辅导书[4, 5]，以及文[6].

[1] 线性代数，第二版，上海交通大学数学系编，科学出版社, 2007 年.

[2] 高等代数学，姚慕生编著，复旦大学出版社, 1999 年.

[3] 数学模型讲义，雷功炎编著，北京大学出版社, 1999 年.

[4] 高等代数，大学数学学习方法指导丛书，姚慕生编著，复旦大学出版社, 1999 年.

[5] 高等代数附册，北京大学数学系几何与代数教研室代数小组编，高等教育出版社, 1992 年.

[6] 例谈行列式在数学解题中的应用，安振平，数学教学, 2009 年第 8 期, 38—39.

欢迎各位学员就本讲义提出意见和建议！

袁 海 荣

2009 年 12 月 28 日

第 1 章　初识线性代数

本章介绍《线性代数》课程的主要内容、学习本课程的必要性、以及学习时需要注意的问题. 在第 2 节, 我们还将引入矩阵及其乘法. 这是线性代数的最核心概念和最重要工具之一, 也充分体现了代数学将具体问题转化为研究某一类抽象对象及其运算的特点.

第 1 节　线性代数初览

一. 线性代数的内容

1. 线性代数的核心是关于如何高效地求解线性代数方程组的理论.

本课程中我们只介绍线性代数方程组的求解理论, 以及该理论体现的数学思想和方法.

2. 什么是线性代数方程组?

线性代数方程组是中学数学里二元一次方程组、三元一次方程组等的自然推广. 它的最一般形式就是如下含有 n 个未知数 $x_1, x_2, \cdots, x_n$ 的由 m 个方程组成的方程组

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

这里 $a_{ij}, b_i\,(i = 1,2,\cdots,m, j = 1,2,\cdots,n)$ 都是已知的数字. a_{ij} 叫做系数, b_i 叫做非齐次项.

3. 求解线性代数方程组所需回答的具体问题.

我们必须解决以下问题, 才算是对线性代数方程组的求解给出了圆满的答案.

(a) 如何判断方程组 (1) 是否有解? 即是否存在一组数 $x_1, x_2, \cdots, x_n$ 使得 (1) 自动成立?

(b) 如果有解，那么有多少个解？特别地，何时只有一个解？如果解很多，怎样刻画这所有的解构成的集合？

(c) 如果没有解, 那么如何给出某种广义的解, 使得 (1) 在某种意义下 “有解 ”?

(d) 要给出 **高效** 地计算出 (1) 的解 (或广义解) 的方法.

本课程的内容，就是围绕这些问题展开的. 在解决这些问题的过程中，出现了许多卓有成效的思想和方法，我们会在学习中适时予以强调.

二. 为什么需要学习线性代数

1. *解决实际问题的需要.*

那么，为什么需要求解线性方程组，或者说，为什么需要提出并解决上述问题 (a)—(d) 呢？

这是因为，一方面，许多实际问题都可以通过建立适当的线性代数方程组得到解决 (参看计算机层析扫描成像数学原理的介绍), 而另一方面，几乎所有计算问题只有化为线性代数方程组的问题才可以得到解决. 例如，天气预报需要求解一个复杂的偏微分方程组，数学理论上没有办法解决，而且数据量巨大，只有采用计算机数值计算. 但是计算机不能直接做微分等运算，必须把偏微分方程离散化为线性代数方程组，利用线性代数方法求解，得到数值结果，作为天气预报的依据. 由于这里出现的线性代数方程组往往有上百万个未知数和方程，所以要求计算方法要高效，才能在短的时间内算出结果. 这也是我们强调问题 (d) 的原因.

由于实际测量的误差, 往往可能出现 $a_{ij}, b_i\,(i=1,2,\cdots,m, j=1,2,\cdots,n)$ 并不准确的情况. 这就可能导致本来应该有唯一解的方程组在数学上是没有解的，或者有多个解. 在无解的时候，就需要我们根据实际情况，给出方程组的某种 “广义解 ”, 这个广义解在某种意义下可看作方程组的近似解，给出实际问题的解答 (例如， CT 扫描的一个图像). 这就是提出问题 (c) 的原因. 对于出现多个解的情况，也需要我们对这些解的全体构成的集合有深入的了解，再根据某种标准选出最 “好 ” 的解. 这都需要我们搞清问题 (a)(b).

2. *可以提高我们的思维能力和数学修养.*

除去实用价值，线性代数作为一门典型的代数学课程，有着与微积分不尽相同的特点，例如，抽象程度较高，概念、定理很多，等等. 通过这门课

的学习，可以使我们得到与学习微积分颇不相同的体验，对高等数学有更全面的认识，也为学习某些后续课程打下基础.

三. 注意学习方法

1. 准确掌握概念、定理.

线性代数概念和定理比较多，比较抽象，而且某些概念和定理看上去好像差不多，但实际上却有着较大的差异. 为了度过这一关，就需要我们做到以下几点.

(A) 准确记忆概念、定理.

(B) 理解概念、定理的来龙去脉 —— 为什么提出这些概念、定理，它们有什么用？它们的优点是什么，不足是什么？

(C) 结合具体实例验证和理解概念、定理. 大家学习抽象的东西，心里必须有具体的例子，才能真正理解抽象内容的价值，提高自己的理解能力和直觉.

(B)(C) 两点也是本课程希望讲透的内容，希望大家认真学习相关的讲解.

2. 自己动手，完成适量习题.

线性代数的另一个特点是计算较多，而且有时需要一定的技巧和判断力(有的问题讲义介绍了有好几种方法，但针对具体问题各方法之间却有优劣之分，需要判断来选取恰当的方法). 技巧和判断力的养成需要一定的切身经验，这只能依靠做适量习题达到.

3. 注意反省总结，结合自身实际，形成自己的对线性代数的认识体系.

这其实是要真正掌握任何一门学科都必须的. 不同的教材或讲义，体现了不同的作者对这个学科的认识. 即使是最优秀的教材，相对于一个具体的学习者而言，也未必就是最适合的. 学习的最高境界，也绝不是能把这些书中的内容倒背如流，而是要根据自己对问题的认识，结合以前所学的基础，吸取这些书中的优点，形成自己的体系. 这也就是金庸小说《倚天屠龙记》中，当张无忌按自己的理解舞出太极剑时，虽然和张三丰教他的样式不同，张三丰却高兴地说他终于彻底领悟了太极剑法的原因.

习题 1. 请阅读计算机层析扫描成像数学原理的介绍.

第 2 节　矩阵及其乘法

大家阅读了计算机层析扫描成像数学原理的介绍，知道了 CT 扫描图像的获得依赖于线性方程组的求解．由此可归纳出一般的线性方程组 (1). 那么，线性代数是如何一步步解决问题 (a)—(d) 的呢？我们从引入一些记号对 (1) 加以简写开始.

一．求和符号 Σ

由于一般线性方程组 (1) 书写起来比较累赘，人们引入了求和符号 Σ 对之加以简写：

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \cdots, m. \tag{2}$$

这里，$\sum_{i\in I} f_i$ 表示把 i 在集合 I 中的每一个 f_i 加起来.（若 $I = \{m, m+1, m+2, \cdots, m+n\}$, 就把 $\sum_{i\in I} f_i$ 写为 $\sum_{i=m}^{m+n} f_i$, 它就是和 $f_m + f_{m+1} + \cdots + f_{m+n}$.) 例如：

$$\begin{aligned}
&\sum_{i=2}^{5}(5i+j) = \sum_{i\in\{2,3,4,5\}}(5i+j)\\
= &\ (5\times 2+j)+\cdots+(5\times 5+j)\\
= &\ 5\times[(2+5)\times 4/2]+4j = 70+4j.
\end{aligned}$$

关于求和符号的更多介绍，请看教材第 28 页．注意，求和符号可以使书写和运算更简便.

习题 1. 用和式写出并计算 $\sum_{n=5}^{10} n^2$.

[355]

二．矩阵及其乘法

我们还可以通过引入矩阵及其乘法，把 (2) 写得更加简单、更具有整体性、启发性.

1. 只有一个方程的情形：向量间的乘法

数学中的基本思想之一，就是从最简单的而又不失去本质的情形出发看待问题．所以我们先看一个方程的情形.

回忆空间解析几何中,我们曾定义两个三维向量 (a_{11},a_{12},a_{13}) 与 (x_1,x_2,x_3) 的内积 $\bullet$ 为对应坐标相乘后再相加

$$(a_{11},a_{12},a_{13})\bullet(x_1,x_2,x_3)=a_{11}x_1+a_{12}x_2+a_{13}x_3.$$

在线性代数里，我们把内积写为如下形式，即

$$(a_{11},a_{12},a_{13})\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}:=(a_{11},a_{12},a_{13})\bullet(x_1,x_2,x_3)=a_{11}x_1+a_{12}x_2+a_{13}x_3.$$

注意，我们把后一个向量 (x_1,x_2,x_3) 由横行的形式 (称为行向量) 写成了一竖列的形式，称为一个列向量.

一般地，我们定义一种乘法运算

$$\begin{aligned}(a_{i1},a_{i2},\cdots,a_{in})\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix} &= (a_{i1},a_{i2},\cdots,a_{in})\bullet(x_1,x_2,\cdots,x_n)\\ &= a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=\sum_{j=1}^{n}a_{ij}x_j. \qquad (3)\end{aligned}$$

从而方程组 (1) 就可以写为

$$(a_{i1},a_{i2},\cdots,a_{in})\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=b_i,\ i=1,2,\cdots,m. \qquad (4)$$

2. 方程组的情形：矩阵与列向量的乘法.

上面 (4) 中，仍然是把方程组作为若干个方程排在一起来看待的，而不是作为一个整体来看待的. 为此，我们引入如下的 m 行 n 列的由数字构成的

表 (我们总把水平的一排称为行，竖直的排称为列):

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \tag{5}$$

这种数表就叫做矩阵. 组成矩阵的数叫做矩阵的元素. 注意这个矩阵中在第 i 行第 j 列的元素就是 (1) 中第 i 个方程里 x_j 的系数 a_{ij}.

我们可把行 (列) 向量看成只有一行 (列) 的矩阵, 把一个数看成只有一行一列的矩阵. 从这个角度讲，矩阵也是数和向量的一种推广.

下面我们定义矩阵 (5) 与列向量 $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ 的一种运算, 我们称之为乘法:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}. \tag{6}$$

注意这个乘积的结果是得到另外一个列向量 (即上式的右端), 但它只有 m 个元素，其中第 i 个元素就是由矩阵 (5) 中第 i 行， $(a_{i1}, a_{i2}, \cdots, a_{in})$, 作为行向量，与列向量 $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ 作乘法得到的.

于是，按照这种乘法，线性方程组 (1) 就可写为

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \tag{7}$$

用 A 表示矩阵 (5), $\mathbf{x}$ 表示列向量 $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{b}$ 表示列向量 $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ (以后我们都用大写字母，如 A, B, Q 等表示矩阵，黑体小写字母表示列向量), 那么由 (7), 线性方程组 (1) 就可以很简单地、 **整体地** 写为

$$A\mathbf{x} = \mathbf{b}. \tag{8}$$

其中 A (称为 (1) 的系数矩阵), 以及非齐次项 $\mathbf{b}$, 都是已知的，而 $\mathbf{x}$ 是未知的.

3. 引入矩阵及乘法的意义.

我们可能会怀疑, 花这么大工夫把写起来较复杂的 (1) 改写为简单的 (8), 是否对解决问题 (即求解线性方程组) 有实质性地帮助 —- 否则就好像是玩文字游戏一样.

事实上， (8) 的形式并不能帮助我们立即解决问题．但是，非常非常重要的一点在于，它和如下我们熟知的一元一次方程

$$ax = b \tag{9}$$

(其中 a, b 为已知数， x 是未知数) 有某种相似性．这就把我们未知的东西 (线性方程组的求解) 与我们已知的东西 (9) 建立了某种联系．我们解决方程 (9) 的那些经验会指导我们如何去研究 (8). 例如，当 $a \neq 0$ 时，可在 (9) 两边同乘以 a^{-1} 而得到解 $x = a^{-1}b$. 这就启发我们尝试去寻找矩阵 A 的逆 A^{-1}, 通过在 (8) 两边同乘 A^{-1} 解出 $\mathbf{x}$. 将未知的东西与已知的东西建立某种联系，是研究解决问题的不二法门．同学们在今后进一步学习矩阵时，也要注意它和数加以联系，看有哪些相同之处，又有哪些不同之处.

此外，在上述引入 (8) 的过程中，也体现出了代数学的一个重要特点，即定义新的对象 (如矩阵), 新的运算 (如矩阵与列向量乘法) 等，而求解线性方程组的问题主要成了研究矩阵运算的问题 (例如如何求 A^{-1}?). 所以代数学主要研究各类带有某种运算的对象的性质.

4. 矩阵与矩阵间的乘法.

上面我们介绍了一个 m 行 n 列矩阵与一个有 n 行的列向量相乘．现在就

很容易定义一个 m 行 n 列矩阵 A 与 n 行 r 列矩阵 T 的乘法. 设

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1r} \\ t_{21} & t_{22} & \cdots & t_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nr} \end{pmatrix}. \tag{10}$$

我们把 T 的第 k 列看作列向量 $\mathbf{t}_k = \begin{pmatrix} t_{1k} \\ t_{2k} \\ \vdots \\ t_{nk} \end{pmatrix}$, 把 T 写为 $T = \begin{pmatrix} \mathbf{t}_1 & \mathbf{t}_2 & \cdots & \mathbf{t}_r \end{pmatrix}$.
那么我们定义 A 与 T 的积就是

$$AT = (A\mathbf{t}_1, A\mathbf{t}_2, \cdots, A\mathbf{t}_r). \tag{11}$$

如果具体把乘积 AT 的第 p 行第 q 列的元素 $(AT)_{pq}$ 写出来, 那么由 (6), 它就是

$$(AT)_{pq} = \sum_{i=1}^{n} a_{pi}\, t_{iq}, \tag{12}$$

即 A 的第 p 行与 T 的第 q 列作内积 (也就是行向量与列向量作乘法).

例 1. $\begin{pmatrix} 0 & 1 & 2 & 3 \\ -1 & 2 & 0 & 3 \\ 2 & -2 & 3 & -5 \end{pmatrix} \begin{pmatrix} 7 & 3 \\ 2 & -3 \\ -1 & 6 \\ -10 & 5 \end{pmatrix}$

$$= \begin{pmatrix} 0\times 7 + 1\times 2 + 2\times(-1) + 3\times(-10) & 0\times 3 + 1\times(-3) + 2\times 6 + 3\times 5 \\ -1\times 7 + 2\times 2 + 0\times(-1) + 3\times(-10) & -1\times 3 + 2\times(-3) + 0\times 6 + 3\times 5 \\ 2\times 7 + (-2)\times 2 + 3\times(-1) + (-5)\times(-10) & 2\times 3 + (-2)\times(-3) + 3\times 6 + (-5)\times 5 \end{pmatrix}$$

$$= \begin{pmatrix} -30 & 24 \\ -33 & 6 \\ 57 & 5 \end{pmatrix}.$$

例 2. $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & \vdots & & \vdots \\ x_ny_1 & x_ny_2 & \cdots & x_ny_n \end{pmatrix}.$

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

习题 3. 把如下线性方程组写为矩阵形式.

$$\begin{cases} 3x+7y-z+8=0, \\ x-w-4=0, \\ z+y=x-11. \end{cases}$$

在实际问题中，经常需要对几个不同的非齐次项 $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_r$ 求解方程组 (8). 设对应的解为 $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_r$:

$$A\mathbf{x}_k = \mathbf{b}_k, \qquad k = 1, 2, \cdots, r. \tag{13}$$

令 $X = (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_r)$ 是一个 n 行 r 列的矩阵，它的第 k 列就是 $\mathbf{x}_k$ $(k = 1, 2, \cdots, r)$; 令 $B = (\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_r)$ 是一个 m 行 r 列的矩阵，它的第 k 列就是 $\mathbf{b}_k$ $(k = 1, 2, \cdots, r)$. 那么我们就可以把 (13) 整体地写为

$$AX = B. \tag{14}$$

其中， AX 就是我们刚才定义的 m 行 n 列矩阵与 n 行 r 列矩阵的乘法. 以后我们看到，通过引入这种矩阵乘法，有助于整体地同时批量化求解多个线性方程组，极大地提高了计算效率.

我们强调一下，两个 m 行 n 列矩阵与 s 行 r 列矩阵，如果 $n \neq s$, 其乘法是没有定义的. 我们看到，引入矩阵及其乘法，是密切联系解决线性方程组这个重大问题的. 而脱离实际需要，为抽象而抽象，非要在任何两个矩阵之间都引入乘法，是没有必要的事情.

5. 矩阵的转置.

最后，我们再引入矩阵的另一种运算，叫转置. 设矩阵 A 按照 (5) 定义. A 的转置是如下 n 行 m 列矩阵 (记作 A^T)

$$\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1i} & a_{2i} & \cdots & a_{mi} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}. \tag{15}$$

它的特点是： A 中的第 k 行成了 A^T 的第 k 列 $(k=1,2,\cdots,m)$; 相应地 A 中的第 l 列成了 A^T 的第 l 行 $(l=1,2,\cdots,n)$. 例如,

$$\begin{pmatrix} 1 & 0 & -2 & 10 \\ -0.9 & 6 & 4.4 & -9 \\ 3.3 & -1 & -1 & 0 \end{pmatrix}^T = \begin{pmatrix} 1 & -0.9 & 3.3 \\ 0 & 6 & -1 \\ -2 & 4.4 & -1 \\ 10 & -9 & 0 \end{pmatrix}.$$

特别地，考虑一个列向量：

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

它可写为 $\mathbf{x}=(x_1,x_2,\cdots,x_n)^T$. 这样就可以节省书籍中排版所占的空间.

一般来讲，引入矩阵的转置，决不是为了书写上节省版面. 面对一个矩阵, 这个由若干个数排成的表, 我们会问: 它的各元素之间是平等的吗？行与行，列与列，行与列之间是平等的吗？转置矩阵将原矩阵中的行与列互换，所以架起了联系行与列之间关系的一座桥梁. 例如，对行所成立的性质，通过转置，就可能得到列具有的相应性质. 所以转置对于我们理解矩阵有着重要的意义.

习题 4. 求如下矩阵的转置： $A=\begin{pmatrix} a & b & c & d \\ b & e & f & g \\ c & f & h & i \\ d & g & i & j \end{pmatrix}$.

复习题

1. 设 $A=\begin{pmatrix} 3 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$, $B=\begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. 求 AB, BA.

2. 写出矩阵 $A=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$ 的转置 A^T, 并写出以 A^T 为系数矩阵的齐次线性方程组.

3. 证明: (1) $\sum_{i=1}^m(\sum_{j=1}^n a_{ij}) = \sum_{j=1}^n(\sum_{i=1}^m a_{ij})$; (2) $\sum_{i=1}^m(\sum_{j=1}^n a_i b_j) = (\sum_{j=1}^n b_j)(\sum_{i=1}^m a_i)$.

4. 用矩阵乘法表示如下线性方程组: $\begin{cases} 5x_1+6x_2=1, \\ x_1+5x_2+6x_3=0, \\ x_2+5x_3+7x_5=0, \\ x_2-x_4=-1, \\ x_1-x_3=8. \end{cases}$

5. 证明: 设 $a\neq 0$, 则三元一次方程 $ax+by+cz=d$ 表示三维空间中经过 $(d/a,0,0)$ 的一个平面. 于是, 方程组 $\begin{cases} a_1x+b_1y+c_1z=d_1, \\ a_2x+b_2y+c_2z=d_2, \end{cases}$ 表示求两个平面的公共部分 (交线).

6. 将矩阵 $\begin{pmatrix} a_1+b_1 & a_1+b_2 & \cdots & a_1+b_n \\ a_2+b_1 & a_2+b_2 & \cdots & a_2+b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_m+b_1 & a_m+b_2 & \cdots & a_m+b_n \end{pmatrix}$ 表示为两个矩阵的乘积.

第 2 章　行 列 式

这一章我们介绍行列式． n 阶行列式就是 n^2 个数按照某种规则作乘法和加法运算得到一个数．利用行列式可以给出由 n 个 n 元一次线性方程构成的方程组当系数行列式不为零时的求解公式 — Cramer 法则．它也是我们研究矩阵的性质以及高效求解一般线性方程组的一个重要的理论工具．

我们先介绍 2 阶和 3 阶行列式是如何通过实际计算引入数学的，以及相应的 Cramer 法则．我们希望能对一般的 n 也得到 Cramer 法则．这就必须给出 $n>3$ 阶行列式的定义．这是通过观察 2 阶和 3 阶行列式的特点，明确 n 阶行列式所应满足的要求，再利用排列的概念完成的．有了 n 阶行列式的概念，我们面对的问题却是直接用定义很难高效地计算一个行列式．这就使得我们要从定义推导出它的某些典型性质．利用这些性质，我们就可以较快地计算行列式，也可以推出一般的 Cramer 法则了．

这一章里，引入行列式及其记号，给出一般行列式定义，推导其性质直至证明一般的 Cramer 法则，都体现了许多重要的而且基本的数学思想．大家不但要掌握好行列式的相关知识，对整体思路和这些思想也要仔细体会．

第 1 节　n 阶行列式

一．二阶与三阶行列式

我们在中学里就已经会用消元法求如下二元一次方程组了：

$$\begin{cases} a_1x_1 & + & b_1x_2 & = & c_1, \\ a_2x_1 & + & b_2x_2 & = & c_2. \end{cases} \tag{1}$$

将第一个方程乘以 b_2，第二个方程乘以 b_1，两者相减，就会得到 $(a_1b_2-a_2b_1)x_1=c_1b_2-c_2b_1$. 如果

$$a_1b_2-a_2b_1\neq 0, \tag{2}$$

就会解出

$$x_1=\frac{c_1b_2-c_2b_1}{a_1b_2-a_2b_1}. \tag{3}$$

类似地，我们可以得到

$$x_2 = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}. \tag{4}$$

有了这个公式，若 (2) 成立，我们就可以直接写出 (1) 的解，而不必每次都重复消元法的过程了.

但是公式 (3)(4) 有两个缺点．首先，它有些复杂，不大好记；其次，由于复杂，很难看出规律，也就难以推广 — 比如，对 n 个 n 元一次方程组的解，是否有类似的公式？这就是人们引入行列式这个概念及其记号的原因.

我们定义四个数 a_1,a_2,b_1,b_2 的如下运算：

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} := a_1b_2 - a_2b_1, \tag{5}$$

称之为 2 阶行列式．注意，行列式本质代表若干个数进行某种运算，结果还是一个数，这和两个数作乘法运算得到乘积是类似的．例如，求行列式 $\begin{vmatrix} 3 & -10 \\ -4 & 2.7 \end{vmatrix}$ 的意思就是按定义计算 $3\times 2.7-(-4)\times(-10) = 8.1-40 = -31.9$, 最后要把这个数字结果算出来．在行列式里，我们把水平的一排数称作 **行**, 并按从上到下顺序依次称为第 1 行, 第 2 行; 我们把竖直的一排数称作 **列**, 并按从左到右顺序依次称为第 1 列, 第 2 列．例如 (5) 的第 2 行就是 $\begin{pmatrix} a_2 & b_2 \end{pmatrix}$, 其第 1 列就是 $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$. 这种叫法与矩阵中的行与列是一致的.

引入行列式有什么好处呢？我们发现，若记

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix},\ D_1 = \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix},\ D_2 = \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix},$$

那么 (3)(4) 就是

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}. \tag{6}$$

这里分母 D 就是 (1) 中系数按原来顺序组成的行列式，而求 x_1 时的分子 D_1 就是把 D 中第 1 列换为非齐次项 $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$，求 x_2 时的分子 D_2 就是把 D 中第 2 列换为 $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$. 这就容易看出一种规律性．这个规律 (这就是最简单的

Cramer 法则) 在表达式 (3) (4) 中是不容易看出的．这就体现出引入行列式这种记号的好处了：它是一种便于记忆复杂算式的符号，而且富有启发性；而一个好的符号 (语言) 就是一个强有力的 (思维) 工具.

习题 1. 求 $D = \begin{vmatrix} 3 & 5 \\ 2 & -7 \end{vmatrix}$.

我们再看三元一次方程组的情形：

$$\left\{\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & = & b_2, \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & = & b_3. \end{array}\right. \tag{7}$$

通过消元法，可以得到如下解的表达式 (要求分母 D 不为零)

$$x_1 = \frac{D_1}{D},\ x_2 = \frac{D_2}{D},\ x_3 = \frac{D_3}{D}. \tag{8}$$

这里

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

和

$$D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix},\ D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix},\ D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}$$

都是 3 阶行列式． 3 阶行列式是如下定义的：

$$\begin{aligned} D = & \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ = & \, a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \end{aligned} \tag{9}$$

例如,

$$\begin{vmatrix} 6 & -3 & 13 \\ 0 & -2 & 1 \\ 7 & 8 & -2 \end{vmatrix} = 137.$$

注意 3 阶行列式的定义是需要记住的．至于这个定义是如何得到的，则就完全和前面 2 阶行列式类似，是通过消元法把解的表达式 (类似 (3)(4)) 写出，然后根据表达式再给出上述定义的．由于过程比较繁琐，我们就略去了．

观察 (8), 我们发现它具有和 (6) 类似的规律．这就是三元一次方程组情形的 Cramer 法则．本章的目的之一，就是对一般 n 个 n 元一次方程组证明相应的 Cramer 法则.

下面看一个具体的算例.

例 1.

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

$\boldsymbol{x}=(x_1,x_2,x_3)^T$ *满足* $A\boldsymbol{x}=\boldsymbol{b}$, *求* $\boldsymbol{x}$.

解.

$$|A|=-1-6+2+1+3-4=-5,$$

$$D_1=\begin{vmatrix}0 & -2 & 1\\ 3 & 1 & -3\\ 1 & 1 & -1\end{vmatrix}=6+3-1-6=2,$$

$$D_2=\begin{vmatrix}1 & 0 & 1\\ 2 & 3 & -3\\ -1 & 1 & -1\end{vmatrix}=-3+2+3+3=5,$$

$$D_3=\begin{vmatrix}1 & -2 & 0\\ 2 & 1 & 3\\ -1 & 1 & 1\end{vmatrix}=1+6+4-3=8,$$

所以 $x_1=-\frac{2}{5}$, $x_2=-1$, $x_3=-\frac{8}{5}$.

验证. $\left\{\begin{array}{ccccccc}-\frac{2}{5} & + & 2 & - & \frac{8}{5} & = & 0,\\ 2\times(-\frac{2}{5}) & - & 1 & - & 3\times(-\frac{8}{5}) & = & 3,\\ \frac{2}{5} & - & 1 & + & \frac{8}{5} & = & 1.\end{array}\right.$

习题 2. 计算行列式 $\begin{pmatrix}a & b & c\\ b & c & a\\ c & a & b\end{pmatrix}$.

习题 3. 求解线性方程组

$$\begin{cases} x_1+2x_2+4x_3=31, \\ x_2+5x_1+2x_3=29, \\ x_3+3x_1-x_2=10. \end{cases}$$

二. 排列及其逆序数、 n 阶行列式的定义

要给出 n 个 n 元一次方程组的解的表达式 (Cramer 法则), 显然上面用消元法做实验一一计算的办法是行不通的. 从 $n=2,3$ 的情形我们看到，要是能把 n 阶行列式定义出来，那么我们就可以猜出解的表达式 ($x_i=D_i/D$, 其中 D 是系数行列式， D_i 是把 D 中第 i 列换为非齐次项 $\mathbf{b}$). 所以首要问题是给出 n 阶行列式的定义.

那么怎样定义 n 阶行列式呢？我们必须用数学中归纳综合的想法，即通过特例或试验发现一些规律 (性质), 然后把这些规律 (性质) 推广到一般情形. 我们首先考察 2 、 3 阶行列式，可以看出 n 阶行列式应有如下性质：

(a) n 阶行列式应当为 $n!$ 个乘积项的和；

(b) 每个乘积项由 n 个行列式中元素相乘而得. 这 n 个元素必须取自不同行、不同列；

(c) 每个乘积项的符号 (取正号还是负号) 由上述取法确定.

于是，我们写出如下 n 阶行列式的定义：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{(j_1\cdots j_n)} (-1)^{\tau(j_1\cdots j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n}, \tag{10}$$

其中 $(j_1\cdots j_n)$ 为 $(1\cdots n)$ 这 n 个不同的数的一个排列 (也就是按新的顺序排列，与中学数学中排列的概念是一样的), 而 $\sum\limits_{(j_1\cdots j_n)}$ 表示对所有的 $(1\cdots n)$ 的排列求和. 由于这样的排列共有 $n!$ 个，故共有 $n!$ 个项 $(-1)^{\tau(j_1\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$ 相加. 这与 (a) 相符.

再看每一项 $(-1)^{\tau(j_1\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$, 这就表示在第 1 行取位于第 j_1 列的那个元素, 然后乘以第2行的位于第 j_2 列的元素，$\cdots\cdots$, 最后乘以第 n 行的位于第 j_n 列的元素. 由于 $(j_1\cdots j_n)$ 是 $(12\cdots n)$ 的一个排列，故 $j_1,j_2,\cdots,j_n$ 互不相同. 这与 (b) 相符. 而 (c), 就表现为因子 $(-1)^{\tau(j_1\cdots j_n)}$. 我们称 $\tau(j_1\cdots j_n)$ 为排列 $(j_1\cdots j_n)$ 的 **逆序数**.

定义 1. 在排列 $(j_1\cdots j_n)$ 中，满足 $j_k > j_s$ 且 $k < s$ 的数对 (j_k, j_s) (这被叫做 **逆序对**) 的个数称为排列 $(j_1\cdots j_n)$ 的 **逆序数**, 记作 $\tau(j_1\cdots j_n)$. 逆序数为偶数的排列叫做偶排列，逆序数为奇数的排列叫奇排列.

例 2. 考虑 (12345) 的两个排列 (1,3,5,4,2) 和 (5,3,4,1,2), 求其逆序数.

解. (1,3,5,4,2) 的逆序对是 (32),(54),(52),(42). 故 $\tau(1,3,5,4,2) = 4$. (5,3,4,1,2)的逆序对是(53),(54),(51),(52),(31),(32),(41),(42),故$\tau(5,3,4,1,2)=8$.

例 3. 考虑 $(12\cdots n)$ 的排列 $(1,2,\cdots,n-1,n),(2,3,\cdots,n,1),(n,n-1,\cdots,2,1)$, 求其逆序数.

解. $\tau(1,2\cdots n-1,n) = 0$, 因为其中没有逆序对. $\tau(2,3\cdots n,1) = n-1$, 因为其中的逆序对就使 $(k1)$, $k = 2,3,\cdots,n$, 共有 $n-1$ 个. $(n,n-1,\cdots,2,1)$ 中任意一个数对都是逆序对，故有 $C_n^2 = \frac{n(n-1)}{2}$ 个，即得 $\tau(n,n-1,\cdots,2,1) = \frac{n(n-1)}{2}$.

习题 4. 求以下三个对 $1,2,\cdots,9$ 的排列的逆序数，并判定其奇偶性.

(1) (134782695);　(2) (217986354);　(3) (987654321)

这样就通过归纳猜想给出了 n 阶行列式的定义 (10). 下面我们用该定义计算一些特殊的 n 阶行列式.

例 4. 求如下上三角形式的 n 阶行列式

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix}.$$

解. $n!$ 个项 $(-1)^{\tau(j_1\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$ 中只有一个项 (即一种取法) 可能不为零，即第 n 行取 a_{nn}, 第 $n-1$ 行取 $a_{n-1,n-1}$ (不能取 $a_{n-1,n}$, 否则有两个元素取自同一列), $\cdots\cdots$, 最后第一行只能取 a_{11}. 故

$$D=(-1)^{\tau(12\cdots n)}a_{nn}a_{(n-1)(n-1)}\cdots a_{11}=a_{11}\cdots a_{nn}.$$

习题 5. 利用定义计算如下 n 阶行列式

$$\begin{vmatrix} 0 & 0 & \cdots & 0 & a_{1n} \\ 0 & 0 & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n(n-1)} & a_{nn} \end{vmatrix}.$$

习题 6. 按定义计算下列行列式：

$$(1)\begin{vmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n-1 \\ n & 0 & 0 & \cdots & 0 \end{vmatrix};\quad (2)\begin{vmatrix} 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ n-1 & \cdots & 0 & 0 & 0 \\ n & \cdots & 0 & 0 & n \end{vmatrix}.$$

我们看到，n 阶行列式定义最本质的是排列及其逆序数. 为了研究行列式的性质，我们给出以下结论.

命题 1. 交换一个给定排列中的两个数，得到的新排列的奇偶性与原排列的相反. 即 $(-1)^{\tau(j_1\cdots j_k\cdots j_s\cdots j_n)}=-(-1)^{\tau(j_1\cdots j_s\cdots j_k\cdots j_n)}$.

证明. 我们这里用的是简化的思想，即从简单的情形，即相邻两个数调换做起，把一般情形化为这种特殊情形得到证明.

首先考虑排列 $\sigma=(c_1\cdots c_kabd_1\cdots d_l)$, 调换 a,b, 我们得到排列

$$\sigma'=(c_1\cdots c_kbad_1\cdots d_l).$$

看 σ 与 σ' 逆序对的差异：在 σ 中若 (ab) 是逆序对，则 σ' 中 (ba) 就不是逆序对了 (相应的，若 σ 中若 (ab) 不是逆序对，则 σ' 中 (ba) 就是逆序对了); 而 σ 和 σ' 中其它的逆序对完全一样. 于是 σ' 的逆序对个数比 σ 多一个或少一个，其奇偶性与 σ 相反.

再考虑一般情形. 设 $\sigma=(c_1\cdots c_s be_1e_2\cdots e_r ad_1\cdots d_l)$, 调换 a,b 得到排列 $\sigma'=(c_1\cdots c_s ae_1e_2\cdots e_r bd_1\cdots d_l)$. 我们可以通过如下方式一步步地调换由 σ 得到 σ': 先交换相邻的 b,e_1, 得到 $\sigma_1=(c_1\cdots c_s e_1be_2\cdots e_r ad_1\cdots d_l)$, 然后在 σ_1 中交换 b,e_2, 得到 $\sigma_2=(c_1\cdots c_s e_1e_2b\cdots e_r ad_1\cdots d_l)$, ⋯⋯, 如此经过 r 次相邻数的交换, 我们就得到 $\sigma_r=(c_1\cdots c_s e_1e_2\cdots e_r bad_1\cdots d_l)$, 交换这里的 a,b, 得到 $\sigma_{r+1}=(c_1\cdots c_s e_1e_2\cdots e_r abd_1\cdots d_l)$. 然后 a 再依次与 $e_r,e_{r-1},\cdots,e_2,e_1$ 交换, 最后就得到 σ'. 这样, 从 σ 到 σ' 要经过 $r+1+r=2r+1$ 次相邻数的交换. 由于每次交换改变一下奇偶性, 故 σ 与 σ' 奇偶性相反. 证毕.

这个证明体现的思想在数学中是非常典型的. 以后我们还会遇到这种思想在其它问题中的应用. 它的要点, 就是先研究那些特殊的情形 (这也往往是比较简单的情形), 然后利用某些性质把一般问题转化为这些特殊情形处理.

习题 7. 证明：对 $1,2,\cdots,n$ 的所有排列中, 奇排列和偶排列各占一半.

第 2 节　行列式的性质与 Cramer 法则

我们已经给出了 n 阶行列式的定义, 所以从理论上讲, 我们可以计算任意 n 阶行列式了. 可是从定义出发计算 n 阶行列式的运算量是相当大的 (要把所有的排列写出来, 算出其逆序数, 然后作 $n(n!)$ 次乘法和 $n!-1$ 次加法). 所以除去一些特殊情形, 用定义计算行列式是很不明智的. 这就需要我们具体情况具体分析, 给出行列式的某些性质. 利用这些性质, 可以把复杂的行列式化为较简单的行列式, 极大地减少运算量, 而且简单地给出 Cramer 法则的证明. 这一节里, 大家要理解如何从行列式的定义出发证明行列式的若干性质, 以及这些性质的意义 (也就是说, 为什么它们是重要的), 并且要掌握 Cramer 法则. 下一节中我们将介绍如何利用这些性质来计算行列式.

一. 行列式的性质

下面依次介绍行列式的六条重要性质. 另外两个重要结果, 即 Laplace 展开定理和行列式的乘积, 将在第 4 节介绍.

设 $A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ 是一个 $n\times n$ 矩阵. 记 $|A|=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$,

称为矩阵 A 对应的行列式. $|A|$ 也常被写作 $\det A$.

性质 1. 一个矩阵的行列式与其转置矩阵的行列式是一样的，即 $|A^T| = |A|$.

注记. 这个性质就表明，在行列式中行与列地位是对称的：行列式中对行成立的性质对列也成立，反之亦然. 这是因为，矩阵转置是将行变为列，列变为行. 所以可通过转置将行列式的列变为行，应用行具有的性质之后，再通过转置回到列，就得到列的相应性质.

这个性质的证明涉及到排列的比较深刻的性质 (要将排列看作一种称为置换的映射，研究其逆映射的逆序数)，我们就省略了.

性质 2. 交换行列式中两行 (列) 对应元素的位置，行列式变号. 即若 (以行为例)

$$D = \left|\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{array}\right| \begin{array}{l} \\ \\ \leftarrow \text{第 } i \text{ 行} \\ \\ \leftarrow \text{第 } k \text{ 行} \\ \\ \\ \end{array},$$

$$D_1 = \left|\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{array}\right| \begin{array}{l} \\ \\ \leftarrow \text{第 } i \text{ 行} \\ \\ \leftarrow \text{第 } k \text{ 行} \\ \\ \\ \end{array}.$$

则 $D_1 = -D$.

证明. 我们只需对交换两行情形给出证明. 利用行列式定义和第 1 节命

题 1,

$$\begin{aligned}
D &= \sum_{(j_1\cdots j_n)} (-1)^{\tau(j_1\cdots j_i\cdots j_k\cdots j_n)} a_{1j_1}a_{2j_2}\cdots a_{ij_i}\cdots a_{kj_k}\cdots a_{nj_n} \\
&= -\sum_{(j_1\cdots j_n)} (-1)^{\tau(j_1\cdots j_k\cdots j_i\cdots j_n)} a_{1j_1}a_{2j_2}\cdots a_{kj_k}\cdots a_{ij_i}\cdots a_{nj_n} \\
&= -D_1.
\end{aligned}$$

这个性质表明，在行列式中，行与行之间 (对应地，列与列) 之间地位也具有对称性：第 1 行并不比第 n 行优越，完全可以把第 n 行放在第 1 行的位置，而把第 1 行放到第 n 行去，行列式只不过改变一下符号而已.

我们还可以看出，行列式中的 n^2 元素的地位也具有平等性. 例如，可以通过交换第 i 行与第 1 行，第 j 列与第 1 列，把本来位于第 i 行 j 列的元素 a_{ij} 放在新行列式第 1 行第 1 列，而行列式值却不变. 所以利用这些性质，可以把我们喜欢的行 (列或者元素) 放在我们喜欢的位置 (例如在这样的位置，我们就可能较简便地利用定义了), 而不改变行列式的绝对值.

推论. 若一个行列式中有两行 (列) 元素对应相同，则此行列式为 0.

习题 1. 证明此推论.

性质 3. 把行列式某一行 (列) 的所有元素同乘以常数 c , 等于用 c 乘以这个行列式. 即 (以行为例)

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{i1} & ca_{i2} & \cdots & ca_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = c\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \leftarrow \text{第 } i \text{ 行}.$$

证明. 利用行列式定义, 左 $= \sum_{(j_1\cdots j_n)}(-1)^{\tau(j_1\cdots j_i\cdots j_k\cdots j_n)}a_{1j_1}\cdots(ca_{ij_i})\cdots a_{nj_n} = c\sum_{(j_1\cdots j_n)}(-1)^{\tau(j_1\cdots j_i\cdots j_k\cdots j_n)}a_{1j_1}\cdots a_{ij_i}\cdots a_{nj_n} =$ 右.

推论 1. 行列式某行有公因子时可以将此公因子提到行列式符号外面.

推论 2. 如果行列式某两行元素对应成比例，则此行列式为 0 .

习题 2. 证明这两个推论.

性质 4: 如果行列式第 i 行 (列) 的各元素均为两元素的和，则此行列式为两行

列式之和. 即 (以行为例)

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

证明. 由行列式定义, 左 $=\sum_{(j_1\cdots j_n)}(-1)^{\tau(j_1\cdots j_i\cdots j_k\cdots j_n)}a_{1j_1}\cdots(a_{ij_i}+b_{jj_1})\cdots a_{nj_n}=\sum_{(j_1\cdots j_n)}(-1)^{\tau(j_1\cdots j_i\cdots j_k\cdots j_n)}a_{1j_1}\cdots a_{ij_i}\cdots a_{nj_n}+\sum_{(j_1\cdots j_n)}(-1)^{\tau(j_1\cdots j_i\cdots j_k\cdots j_n)}a_{1j_1}\cdots b_{ij_i}\cdots a_{nj_n}=$ 右. 证毕.

例 1. $\begin{vmatrix} 4 & 7 & 2 \\ 3 & 1 & 0 \\ 2 & 5 & 4 \end{vmatrix} = \begin{vmatrix} 4 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 5 & 4 \end{vmatrix} + \begin{vmatrix} 0 & 7 & 0 \\ 3 & 1 & 0 \\ 2 & 5 & 4 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 2 \\ 3 & 1 & 0 \\ 2 & 5 & 4 \end{vmatrix} = 16-84+26=-42.$

特别注意: 形如 $\begin{vmatrix} a+b & c+d \\ e+f & g+h \end{vmatrix} = \begin{vmatrix} a & c \\ e & g \end{vmatrix} + \begin{vmatrix} b & d \\ f & h \end{vmatrix}$ 的式子一般来讲都是 **不成立** 的. 大家不能犯这种 "想当然" 的错误.

性质 5. 将行列式某一行 (列) 的元素都乘以某个常数再对应加到另外一行 (列) 上去, 行列式值不变. 即 (以行为例)

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

证明. 利用性质 2, 3, 4 即可.

性质 3, 4, 5 的重要性体现在它们可以简化行列式, 比如使得行列式中某一行许多元素都是零. 下面介绍的性质 6 从另外一个角度来化简行列式: 它把 n 阶行列式的计算化为 $n-1$ 阶行列式的计算. 为此我们需要如下重要的定义.

定义 1. 在 n 阶行列式 $D=|a_{ij}|_n$ 中，划去元素 a_{ij} 所在的第 i 行及第 j 列，剩下的 $(n-1)^2$ 个元素按原来顺序位置构成的 $n-1$ 阶行列式

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

称为行列式 D 中元素 a_{ij} 的 **余子式**, 记作 M_{ij}; 称 $(-1)^{i+j}M_{ij}$ 为 a_{ij} 的 **代数余子式**, 记为 $\mathcal{A}_{ij}$, 即

$$\mathcal{A}_{ij}=(-1)^{i+j}M_{ij}.$$

例 2. $\begin{vmatrix} 3 & 2 & 7 \\ 0 & 1 & 8 \\ 0 & 2 & 5 \end{vmatrix}$ 中 *8* 的代数余子式为 $(-1)^{2+3}\times 6=-6$, *7* 的余子式为 *0*.

习题 3. 已知行列式 $D=\begin{vmatrix} 0 & 1 & 0 & 4 \\ 0 & 2 & -1 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & 5 & 2 & 7 \end{vmatrix}$. 求：(a) $2M_{43}+2M_{44}$; (b) $\mathcal{A}_{23}+\mathcal{A}_{33}$.

性质 6. 对 $D=|a_{ij}|_n$, 有

$$D=\sum_{j=1}^{n} a_{ij}\mathcal{A}_{ij} \tag{11}$$

对 $i=1,2,3\cdots,n$ 都成立. 这被称作行列式按第 i 行展开降阶. 类似地还成立行列式按第 k 列展开降阶，即

$$D=\sum_{j=1}^{n} a_{jk}\mathcal{A}_{jk}, \quad k=1,2,\cdots,n. \tag{12}$$

例 3. 按第 4 行展开，我们有

$$\begin{vmatrix} 0 & 1 & 0 & 4 \\ 0 & 2 & -1 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & 5 & 2 & 7 \end{vmatrix} = 0\times(-1)^{4+1}\begin{vmatrix} 1 & 0 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 1 \end{vmatrix} + 5\times(-1)^{4+2}\begin{vmatrix} 0 & 0 & 4 \\ 0 & -1 & 1 \\ 1 & 2 & 1 \end{vmatrix}$$
$$+2\times(-1)^{4+3}\begin{vmatrix} 0 & 1 & 4 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix} + 7\times(-1)^{4+4}\begin{vmatrix} 0 & 1 & 0 \\ 0 & 2 & -1 \\ 1 & 1 & 2 \end{vmatrix}.$$

但如果按第 1 列展开，

$$\begin{vmatrix} 0 & 1 & 0 & 4 \\ 0 & 2 & -1 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & 5 & 2 & 7 \end{vmatrix} = 1\times(-1)^{3+1}\begin{vmatrix} 1 & 0 & 4 \\ 2 & -1 & 1 \\ 5 & 2 & 7 \end{vmatrix}.$$

我们首先看这个结论的一个重要推论．由性质 6, 知

$$\sum_{m=1}^{n} k_m\mathcal{A}_{jm} = k_1\mathcal{A}_{j1} + k_2\mathcal{A}_{j2} + \cdots + k_n\mathcal{A}_{jn} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ k_1 & k_2 & \cdots & k_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \begin{matrix} \\ \\ \leftarrow \text{第 } j \text{ 行}. \\ \\ \\ \end{matrix}$$

取 $k_m = a_{im},\ i \neq j,\ m = 1,2,\cdots,n$，则由性质 2 推论，

$$\sum_{m=1}^{n} a_{im}\mathcal{A}_{jm} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \begin{matrix} \\ \\ \leftarrow \text{第 } i \text{ 行} \\ \\ \leftarrow \text{第 } j \text{ 行} \\ \\ \\ \end{matrix} = 0. \tag{13}$$

引入 Kronecker 记号

$$\delta_{ij}=\begin{cases}1 & i=j,\\ 0 & i\neq j,\end{cases}$$

(11)(13) 就可以简洁地写为

$$\sum_{m=1}^{n} a_{im}\mathcal{A}_{jm}=\delta_{ij}D,\qquad \text{其中}\quad D=|a_{ij}|_n. \tag{14}$$

类似地，采用按列展开，我们也有

$$\sum_{m=1}^{n} a_{mi}\mathcal{A}_{mj}=\delta_{ij}D,\qquad \text{其中}\quad D=|a_{ij}|_n. \tag{15}$$

这两个等式对于我们研究 Cramer 法则及矩阵可逆都有重要理论意义，大家务必要理解.

例 4. *我们利用性质 6 来解决习题 3.*

$$\begin{aligned}
2M_{43}+M_{44} &= 0\times\mathcal{A}_{41}+0\times\mathcal{A}_{42}-2\mathcal{A}_{43}+\mathcal{A}_{44}\\
&= \begin{vmatrix} 0 & 1 & 0 & 4\\ 0 & 2 & -1 & 1\\ 1 & 1 & 2 & 1\\ 0 & 0 & -2 & 1\end{vmatrix}\\
&= 1\times(-1)^{3+1}\times\begin{vmatrix} 1 & 0 & 4\\ 2 & -1 & 1\\ 0 & -2 & 1\end{vmatrix}\\
&= -15.
\end{aligned}$$

这里第三个等号用到了按第一列展开. 用类似办法可以算出 $\mathcal{A}_{23}+\mathcal{A}_{33}=1$.

这一小节的最后我们给出性质 6 的证明. 由性质 1, 我们仅需证明按行展开的情形，即 (11) 式.

首先考虑情形 1. 这也是最简单的情形，即行列式最后一行除最后一个

元素为 1, 其余元素均为零的情形. 此时，根据行列式定义，就有

$$\begin{aligned}
A &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ 0 & 0 & \cdots & 1 \end{vmatrix} \\
&= \sum_{(j_1 j_2 \cdots j_{n-1} n)} (-1)^{\tau(j_1 j_2 \cdots j_{n-1} n)} a_{1j_1} a_{2j_2} \cdots a_{n-1,j_{n-1}} \times 1 \\
&= \sum_{(j_1 j_2 \cdots j_{n-1})} (-1)^{\tau(j_1 j_2 \cdots j_{n-1})} a_{1j_1} a_{2j_2} \cdots a_{n-1,j_{n-1}} \\
&= \begin{vmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & \ddots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{vmatrix} \\
&= 1 \times \mathcal{A}_{nn}.
\end{aligned}$$

这里 $(j_1 j_2 \cdots j_n)$ 是 $(12, \cdots, n)$ 的排列，而 $\mathcal{A}_{ij}$ 就是 A 中位于第 i 行第 j 列的元素对应的代数余子式. 所以在这种情形下性质 6 成立.

其次考虑情形 2, 即行列式 A 最后一行除 $a_{nk} = 1 (1 \le k < n)$ 外其它元素均为零的情形. 通过第 k 列依次与第 $k+1, k+2, \cdots, n$ 列交换, 则所得行列式 B 就具有上面情形 1 的形式. 由于在 A 中 a_{nk} 对应的余子式 M_{nk} 与在 B 中第 n 行 n 列对应的余子式是一样的，但 $A = (-1)^{n-k} B = (-1)^{n-k} M_{nk} = (-1)^{n+k} M_{nk} = \mathcal{A}_{nk} = 1 \times \mathcal{A}_{nk}$. 这就是说对情形 2, 性质 6 依然成立.

现在我们证明行列式按最后一行展开成立. 利用行列式的加法和数乘,

我们可以把一个一般的行列式分解为若干个具有情形 1 、2 形式的行列式：

$$
\begin{aligned}
A &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \\
&= \sum_{k=1}^{n} a_{nk} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,k} & \cdots & a_{n-1,n} \\ 0 & 0 & \cdots & 1 & \cdots & 0 \end{vmatrix} \\
&= \sum_{k=1}^{n} a_{nk}\mathcal{A}_{nk}.
\end{aligned}
$$

最后考虑任意行列式 A 按任意一行的展开. 对于第 $l(1 \le l < n)$ 行，它总可以通过依次与 $l+1, l+2, \cdots, n$ 行的交换，换到第 n 行去，得到行列式 C. 则 A 中 a_{lk} 对应的余子式 M_{lk} 与行列式 C 中第 n 行第 k 列元素 a_{lk} 对应的余子式是一样的 (但后者的代数余子式为 $(-1)^{n+k}M_{lk}$). 利用上面已证明的对最后一行展开，有

$$
A = (-1)^{n-l}B = (-1)^{n-l}\sum_{k=1}^{n} a_{lk}(-1)^{n+k}M_{lk} = \sum_{k=1}^{n} a_{lk}(-1)^{k+l}M_{lk} = \sum_{k=1}^{n} a_{lk}\mathcal{A}_{lk}.
$$

这就完成了性质 6 的证明.

二. Cramer 法则

利用性质 6, 我们就可以证明如下 Cramer 法则了.

定理 1. 如果 n 元线性方程组

$$
A\boldsymbol{x} = \boldsymbol{b}, \qquad A = (a_{ij})_{n\times n} \tag{16}
$$

的系数行列式 $D=|A|\neq 0$, 则 (16) 有唯一解

$$x=\frac{1}{D}\begin{pmatrix} D_1 \\ \vdots \\ D_n \end{pmatrix}, \tag{17}$$

其中

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

即将 D 中第 j 列换为非齐次项 $\boldsymbol{b}=(b_1,b_2,\cdots,b_n)^T$.

证明．我们需要证明如下两点：

(a) (17) 为 (16) 的解;

(b) (16) 的解只能是 (17).

我们首先证明 (a). 考察 (16) 的第 i 个方程 ($i=1,2,\cdots,n$):

$$\sum_{j=1}^{n} a_{ij}x_j=b_i. \tag{18}$$

将 (17) 代入其左边，就得到

$$\begin{aligned} \frac{1}{D}\sum_{j=1}^{n} a_{ij}D_j &= \frac{1}{D}\sum_{j=1}^{n} a_{ij}(\sum_{k=1}^{n} b_k\mathcal{A}_{kj}) \\ &= \frac{1}{D}\sum_{k=1}^{n} b_k(\sum_{j=1}^{n} a_{ij}\mathcal{A}_{kj}) \\ &= \frac{1}{D}\sum_{k=1}^{n} b_k\delta_{ik}D \\ &= b_i. \end{aligned}$$

其中第一个等号用到了行列式 D_j 按其第 j 列展开：

$$D_j=b_1\mathcal{A}_{1j}+b_2\mathcal{A}_{2j}+\cdots+b_n\mathcal{A}_{nj}=\sum_{k=1}^{n} b_k\mathcal{A}_{kj}, j=1,2,\cdots,n, \tag{19}$$

而 $\mathcal{A}_{ij}$ 就是 $|A|$ 中第 i 行 j 列元素的代数余子式. 第二个等号用了数的乘法分配律和加法交换律. 第三个等号用了 (14) 式. 最后一个等号用了 δ_{ij} 的定义.

下面再证明 (b). 设 $\mathbf{c} = (c_1, c_2, \cdots, c_n)^T$ 是 (16) 的解, 即, 写为分量的形式, 第 i $(i = 1, 2, \cdots, n)$ 个方程成为如下恒等式:

$$\sum_{j=1}^{n} a_{ij} c_j = b_i.$$

它两边同乘以 $\mathcal{A}_{ik}$ (k 是 $1, 2, \cdots, n$ 中取定的一个数), 然后关于 i 从 1 到 n 相加, 得

$$\sum_{j=1}^{n} (c_j \delta_{jk}) |D| = \sum_{j=1}^{n} (\sum_{i=1}^{n} a_{ij} \mathcal{A}_{ik}) c_j = \sum_{i=1}^{n} b_i \mathcal{A}_{ik} = D_k.$$

这里第一个等号用了 (15), 而第三个等号用了 (19). 于是我们可以得到 $c_k |D| = D_k$. 由于 $D \neq 0$, 这就是 $c_k = \frac{D_k}{|D|}$, 即得 (17). 证毕.

作为 Cramer 法则的简单推论, 我们可知齐次线形方程组 $Ax = 0$ 当 $|A| \neq 0$ 时只有零解, 而 $Ax = 0$ 有非零解必然意味着 $|A| = 0$. 下面我们看几个例子. 这些例子体现了利用性质计算行列式的一些基本技巧.

例 5. 讨论 λ 为何值时, 线性方程组

$$\begin{cases} \lambda x_1 + x_2 + x_3 = 1, \\ x_1 + \lambda x_2 + x_3 = \lambda, \\ x_1 + x_2 + \lambda x_3 = \lambda^2 \end{cases}$$

有唯一解, 并求之.

解. 将系数行列式的第一行分别乘以 -1 和 $-\lambda$, 分别加到第 2 行和第 3 行, 利用行列式的性质, 得到

$$D = \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{vmatrix} = \begin{vmatrix} \lambda & 1 & 1 \\ 1-\lambda & \lambda-1 & 0 \\ 1-\lambda^2 & 1-\lambda & 0 \end{vmatrix}$$
$$= (1-\lambda)^2 + (1-\lambda)(1-\lambda^2)$$
$$= (1-\lambda)^2 (2+\lambda).$$

这里第三个等号可以由行列式按第 3 列展开得到．从而，由 Cramer 法则，当且仅当 $\lambda \neq 1$ 且 $\lambda \neq 2$ 时原方程组有唯一解.

下面再计算分子．用和前面类似的方式，我们可以让某一行或列除一个元素外全是零，然后再按行或列展开计算．我们有

$$\begin{aligned}
D_1 &= \begin{vmatrix} 1 & 1 & 1 \\ \lambda & \lambda & 1 \\ \lambda^2 & 1 & \lambda \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ \lambda-1 & \lambda-1 & 0 \\ \lambda^2-\lambda & 1-\lambda & 0 \end{vmatrix} \\
&= -(1-\lambda)^2-(\lambda-1)\lambda(\lambda-1) \\
&= -(1-\lambda)^2(1+\lambda),
\end{aligned}$$

$$\begin{aligned}
D_2 &= \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & \lambda^2 & \lambda \end{vmatrix} = \begin{vmatrix} \lambda & 1 & 1 \\ 1-\lambda & \lambda-1 & 0 \\ 1-\lambda^2 & \lambda^2-\lambda & 0 \end{vmatrix} \\
&= -\lambda(1-\lambda)^2+(1-\lambda)(1-\lambda^2) \\
&= (1-\lambda)^2,
\end{aligned}$$

$$\begin{aligned}
D_3 &= \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & \lambda \\ 1 & 1 & \lambda^2 \end{vmatrix} = \begin{vmatrix} 0 & 1-\lambda & 1-\lambda^3 \\ 0 & \lambda-1 & \lambda(1-\lambda) \\ 1 & \lambda & \lambda^2 \end{vmatrix} \\
&= \lambda(1-\lambda)^2+(1-\lambda)(1-\lambda^3) \\
&= (1-\lambda)^2(\lambda+1)^2.
\end{aligned}$$

所以

$$x_1 = -\frac{1+\lambda}{2+\lambda},\ x_2 = \frac{1}{\lambda+2},\ x_3 = \frac{(\lambda+1)^2}{\lambda+2}.$$

例 6. 在实际问题中常出现数据拟合的问题．例如通过测量得到某个函数 f 在若干点的值：$f(-1)=0, f(1)=4, f(2)=3, f(3)=16$. 我们希望得到 $f(0)$ 的一个近似值．此时一种办法就是假设 f 为一个三次多项式 $f(x)=ax^3+bx^2+cx+d$, 由上述四个函数值得到关于 a,b,c,d 的四元一次方程组

$$\begin{pmatrix} -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ 8 & 4 & 2 & 1 \\ 27 & 9 & 3 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 3 \\ 16 \end{pmatrix}.$$

用 *Cramer* 法则，按前面介绍的办法计算行列式，我们可以解得 $(a,b,c,d)^T = (2,-5,0,7)^T$. 所以可以近似拟合 $f(0)=d=7$. 当然，这样拟合的效果是否好，是要通过实际观测来检验的；如果好，就采用，如果不好，就要考虑其它的办法.

习题 4. 求解例 5 中的线性方程组.

注记. Cramer 法则的好处是对 $D\neq 0$ 的 n 元一次线性方程组的解利用 n 阶行列式给出了一个公式，这对于我们处理一些理论问题是非常重要的工具. 但它的缺点也很明显. 首先，没有解决 m 个方程 n 个未知数的一般情形方程组的可解问题；其次，因为需要计算行列式，一般来说计算量很大，效率不高；再次，当 D 接近于 0 时，由于它在分母上，故可能会产生很大的误差. 所以在下一章，我们会利用本章所学的知识，通过理论推导，发现一些更有效的办法.

习题 5. 计算行列式.

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

第 3 节　行列式的计算与应用举例

行列式的计算及应用具有较高技巧性. 这一节通过例子介绍一些基本技巧和思路，希望给大家以启发.

例 1. 计算 $n+1$ 阶行列式

$$D_n = \begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_n \\ b_1 & d_1 & 0 & \cdots & 0 \\ b_2 & 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_n & 0 & 0 & \cdots & d_n \end{vmatrix}.$$

一般来说，对于这样比较复杂的符号行列式，我们应通过看 $n=1,2,3$ 等特例，搞清它的结构. 然后通过保持这种结构的展开，得到一个递推关系式，利用数学归纳法或数列等的技巧求解. 这是一种比较常规的方法. 按照这种思路，解法如下.

首先看

$$n=1,\quad D_1=\begin{vmatrix} a_0 & a_1 \\ b_1 & d_1 \end{vmatrix};$$

$$n=2,\quad D_2=\begin{vmatrix} a_0 & a_1 & a_2 \\ b_1 & d_1 & 0 \\ b_2 & 0 & d_2 \end{vmatrix};$$

$$n=3,\quad D_3=\begin{vmatrix} a_0 & a_1 & a_2 & a_3 \\ b_1 & d_1 & 0 & 0 \\ b_2 & 0 & d_2 & 0 \\ b_3 & 0 & 0 & d_3 \end{vmatrix}.$$

这就表明，行列式最后一行、最后一列是“新生”的. 所以我们采取按最后一列展开的办法，

$$\begin{aligned} D_n &= d_n(-1)^{n+n+2}D_{n-1}+a_n(-1)^{1+n+1}\begin{vmatrix} b_1 & d_1 & 0 & \cdots & 0 \\ b_2 & 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_n & 0 & 0 & \cdots & 0 \end{vmatrix} \\ &= d_nD_{n-1}+(-1)^n a_n b_n(-1)^{n+1}\begin{vmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & d_{n-1} \end{vmatrix} \\ &= d_nD_{n-1}-a_nb_nd_1d_2\cdots d_{n-1}. \end{aligned}$$

如果按别的列展开，就会破坏这种结构，也就不能得到上述递推表达式了.

下面再介绍另一种较简洁的方法，即针对该行列式特点，利用消元法消去第一列中后 n 个元素，将行列式直接化为对角形式. 事实上，如果对任意

j, $d_j \neq 0$, 那么将第 j 列 $(j = 2, \cdots, n+1)$ 乘 $-\frac{b_{j-1}}{d_{j-1}}$ 再加到第 1 列就得到

$$D_n = \begin{vmatrix} a_0 - \sum\limits_{j=1}^{n} a_j \frac{b_j}{d_j} & a_1 & a_2 & \cdots & a_n \\ 0 & d_1 & 0 & \cdots & 0 \\ 0 & 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{vmatrix} = (a_0 - \sum_{k=1}^{n} \frac{a_k b_k}{d_k}) d_1 d_2 \cdots d_n.$$

若有某 $d_i = 0$, 则沿第 $i+1$ 行展开得

$$\begin{aligned} D_n = (-1)^{1+i+1} b_i & \begin{vmatrix} a_1 & a_2 & \cdots & a_{i-1} & a_i & a_{i+1} & \cdots & a_n \\ d_1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{i-1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & d_{i+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & d_n \end{vmatrix} \\ &= (-1)^i b_i (-1)^{i+1} a_i (d_1 \cdots d_{i-1} d_{i+1} \cdots d_n) \\ &= -a_i b_i d_1 \cdots d_{i-1} d_{i+1} \cdots d_n. \end{aligned}$$

例 2. 证明 *Vandermonde* 行列式

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

注. Π 是连乘符号, 表示满足要求的项相乘. 例如, $\Pi_{3 \le j < 5} j = 3 \times 4 = 12$. 我们已经在上一节例 5 种遇到过这种 Vandermonde 行列式, 它的用处之一就是多项式拟合.

证明方法一. 利用常规思路，先消元再展开得到递推关系式，再用归纳法. $n=2$ 时,

$$D_2=\begin{vmatrix}1 & 1\\ x_1 & x_2\end{vmatrix}=x_2-x_1=\prod_{1\le i<j\le 2}(x_j-x_i).$$

假设结论对 $n=k$ 成立，则将第 $n-1$ 行乘 $-x_n$ 加到 n 行，第 $n-2$ 行乘 $-x_n$ 加到第 $n-1$ 行，……, 第一行乘 $-x_n$ 加到第二行，得

$$\begin{aligned}
D_n&=\begin{vmatrix}1 & 1 & \cdots & 1\\ x_1-x_n & x_2-x_n & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ x_1^{n-1}-x_1^{n-2}x_n & x_2^{n-1}-x_2^{n-1}x_n & \cdots & 0\end{vmatrix}\\
&=(-1)^{n+1}\begin{vmatrix}x_1-x_n & x_2-x_n & \cdots & x_{n-1}-x_n\\ x_1^2-x_1x_n & x_2^2-x_2x_n & \cdots & x_{n-1}^2-x_{n-1}x_n\\ \vdots & \vdots & \ddots & \vdots\\ x_1^{n-2}(x_1-x_n) & x_2^{n-2}(x_2-x_n) & \cdots & x_{n-1}^{n-2}(x_{n-1}-x_n)\end{vmatrix}\\
&=(-1)^{n+1}(x_n-x_1)(x_n-x_2)\cdots(x_n-x_{n-1})(-1)^{n-1}\begin{vmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_{n-1}\\ \vdots & \vdots & \ddots & \vdots\\ x_1^{n-2} & x_2^{n-2} & \cdots & x_{n-1}^{n-2}\end{vmatrix}\\
&=(x_n-x_1)(x_n-x_2)\cdots(x_n-x_{n-1})D_{n-1}.
\end{aligned}$$

利用归纳假设，就有 $D_n=\prod\limits_{1\le i<j\le n}(x_j-x_i)$, 得证.

方法二. 令

$$f(x)=\begin{vmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x\\ x_1^2 & x_2^2 & \cdots & x^2\\ \vdots & \vdots & \ddots & \vdots\\ x_1^{n-1} & x_2^{n-1} & \cdots & x^{n-1}\end{vmatrix}.$$

按最后一列展开就发现这是一个 $n-1$ 次多项式且 $f(x_i)=0, i=1,\cdots,n-1$. 于是因式分解得到 $f(x)=(x-x_1)\cdots(x-x_{n-1})D_{n-1}(-1)^{n+n}$. 注意这里 D_{n-1} 就是最高次项的系数. 从而由归纳假设， $D_n=f(x_n)=\Pi_{1\le i<j\le n}(x_j-x_i)$.

例 3. 求

$$D_n=\begin{vmatrix} a_1+b_1 & a_2+b_1 & \cdots & a_n+b_1 \\ a_1+b_2 & a_2+b_2 & \cdots & a_n+b_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1+b_n & a_2+b_n & \cdots & a_n+b_n \end{vmatrix},\quad 即 a_{ij}=a_j+b_i.$$

解. 第 $2,3,\cdots,n$ 列依次减去第 1 列就得到

$$D_n=\begin{vmatrix} a_1+b_1 & a_2-a_1 & \cdots & a_n-a_1 \\ a_1+b_2 & a_2-a_1 & \cdots & a_n-a_1 \\ \vdots & \vdots & \ddots & \vdots \\ a_1+b_n & a_2-a_1 & \cdots & a_n-a_1 \end{vmatrix}=(a_2-a_1)\cdots(a_n-a_1)\begin{vmatrix} a_1+b_1 & 1 & \cdots & 1 \\ a_1+b_2 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_1+b_n & 1 & \cdots & 1 \end{vmatrix}=0,\ n\geq 3.$$

此外，容易算得

$$\begin{aligned} &D_1=a_1+b_1, \\ &D_2=a_1a_2+a_1b_2+b_1a_2+b_1b_2-a_1a_2-a_1b_1-a_2b_2-b_1b_2 \\ &\quad=a_1(b_2-b_1)+a_2(b_1-b_2) \\ &\quad=(a_2-a_1)(b_2-b_1). \end{aligned}$$

习题 1. 计算如下 n 阶行列式.

$$(1)\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 0 & \cdots & 0 \\ 1 & 0 & 3 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & n \end{vmatrix};\qquad (2)\begin{vmatrix} x & y & 0 & \cdots & 0 & 0 \\ 0 & x & y & \cdots & 0 & 0 \\ 0 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & x & y \\ y & 0 & 0 & \cdots & 0 & x \end{vmatrix};$$

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

下面介绍利用行列式解决问题的一些例子.

例 4. 已知 $(z-x)^2-4(x-y)(y-z)=0$, 求证： $x-y=y-z$.

证明. 本题虽然通过直接展开和因式分解也可较快得到答案，但利用行列式会更加快捷. 由于 $0=(z-x)^2-4(x-y)(y-z)=\begin{vmatrix} z-x & 2(x-y) \\ 2(y-z) & z-x \end{vmatrix}=\begin{vmatrix} 2y-z-x & z+x-2y \\ 2(y-z) & z-x \end{vmatrix}=(2y-x-z)\begin{vmatrix} 1 & -1 \\ 2(y-z) & z-x \end{vmatrix}=(2y-x-z)(z-x+2y-2z)=(2y-x-z)^2$. 于是 $2y-x-z=0$, 即得所证.

例 5. 若 $\log_{18}9=a, 18^b=5$, 求 $\log_{36}45$.

证明. 设 $\log_{36}45=c$. 将上面对数化为常用对数，就有

$$a=\frac{\lg 9}{\lg 18}=\frac{2\lg 3}{\lg 2+2\lg 3},\ b=\frac{\lg 5}{\lg 2+2\lg 3},\ c=\frac{\lg 5+2\lg 3}{2\lg 2+2\lg 3}.$$

这可改写为如下以 $\lg 2$, $\lg 3$, $\lg 5$ 为未知数的三元一次方程组：

$$\begin{cases} a\lg 2+2(a-1)\lg 3=0, \\ b\lg 2+2b\lg 3-\lg 5=0, \\ 2c\lg 2+2(c-1)\lg 3-\lg 5=0. \end{cases}$$

由 Cramer 法则，由于它有非零解，故系数行列式必为零，即

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

通过求行列式就得到 $ac+a+b-2c=0$, 从而 $\log_{36}45=c=\frac{a+b}{2-a}$.

例 6. 在 $\triangle ABC$ 中，证明： $\cos^2 A+\cos^2 B+\cos^2 C+2\cos A\cos B\cos C=1$.

证明．设角 A,B,C 对边边长分别为 a,b,c, 利用射影定理，就有

$$\begin{cases} a = b\cos C + c\cos B, \\ b = c\cos A + a\cos C, \\ c = a\cos B + b\cos A. \end{cases}$$

将它视作 a,b,c 的齐次线性方程组．由于 a,b,c 非零，利用 Cramer 法则，就得到

$$\begin{vmatrix} -1 & \cos C & \cos B \\ \cos C & -1 & \cos A \\ \cos B & \cos A & -1 \end{vmatrix} = 0.$$

可求得左边行列式是 $\cos^2 A + \cos^2 B + \cos^2 C + 2\cos A\cos B\cos C - 1$. 即得所证.

例 7. 平面上三点 $A(x_1,y_1),B(x_2,y_2),C(x_3,y_3)$ 构成的三角形 ABC 的面积是如下行列式 D 的绝对值的一半：

$$D = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

证明．首先假设 C 就是坐标原点，即 $(x_3,y_3) = (0,0)$. 回忆解析几何中向量内积定义为 $(x_1,y_1)\bullet(x_2,y_2) = x_1x_2 + y_1y_2$. 利用这个内积，我们可以计算 $\cos\angle AOB = \frac{(x_1,y_1)\bullet(x_2,y_2)}{\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2}}$. 那么相应三角形的面积就是

$$\begin{aligned} S &= \frac{1}{2}|OA||OB|\sin\angle AOB \\ &= \frac{1}{2}\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2}\sqrt{1-\cos^2\angle AOB} \\ &= \frac{1}{2}\sqrt{(x_1^2+y_1^2)(x_2^2+y_2^2)-(x_1x_2+y_1y_2)^2} \\ &= \frac{1}{2}\sqrt{(x_1y_2-x_2y_1)^2} = \frac{1}{2}|x_1y_2-x_2y_1|. \end{aligned}$$

注意这里最后一个式子中的 $|x_1y_2-x_2y_1|$ 表示数 $x_1y_2-x_2y_1$ 的绝对值，而不

是行列式. 注意到 $x_1y_2-x_2y_1=\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}=\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ 0 & 0 & 1 \end{vmatrix}$. 这就给出了 $C=O$ 情形下的证明.

再考虑一般情形. 将坐标平移使得原点就在 C 点，则 A 的新坐标就是 (x_1-x_3,y_1-y_3), B 的新坐标就是 (x_2-x_3,y_2-y_3). 利用上面已证结论，

$$2S_{\triangle ABC}=\left\|\begin{matrix} x_1-x_3 & y_1-y_3 & 1 \\ x_2-x_3 & y_2-y_3 & 1 \\ 0 & 0 & 1 \end{matrix}\right\|=\left\|\begin{matrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{matrix}\right\|.$$

注意这里内层的 $|\cdot|$ 是行列式的记号，外层的两条竖线表示取内层行列式值的绝对值. 最后一个等号是由第 3 列分别乘 x_3 和 y_3 加到第 1 列、第 2 列，利用行列式的性质得到的. 证毕.

习题 2. 平面上三点 $A(x_1,y_1),B(x_2,y_2),C(x_3,y_3)$ 共线的充要条件是

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}=0.$$

习题 3. 三次函数 $f(x)=ax^3+bx^2+cx+d\ (a\neq 0)$ 图像上有不同的三点 A,B,C, 其横坐标依次为 x_1,x_2,x_3, 则 A,B,C 三点共线的充分必要条件是 $x_1+x_2+x_3=-\frac{b}{a}$.

例 8. *解方程：* $\sqrt{4-2\sqrt{3}\sin x}+\sqrt{10-4\sqrt{3}\sin x-6\cos x}=2.$

解. 该方程可以改写为

$$\sqrt{(\sqrt{3}\cos x-0)^2+(\sqrt{3}\sin x-1)^2}+\sqrt{(\sqrt{3}\cos x-\sqrt{3})^2+(\sqrt{3}\sin x-2)}=2.$$

记 $A(0,1),B(\sqrt{3},2),C(\sqrt{3}\cos x,\sqrt{3}\sin x)$, 上面的方程就是

$$|AC|+|CB|=|AB|.$$

这就意味着 A,B,C 三点共线 (注意这一步是必要但不充分的，所以下面可能会出现增根). 从而

$$\begin{vmatrix} 0 & 1 & 1 \\ \sqrt{3} & 2 & 1 \\ \sqrt{3}\cos x & \sqrt{3}\sin x & 1 \end{vmatrix} = 0.$$

左边行列式也就是 $3\sin x - \sqrt{3}\cos x - \sqrt{3}$. 所以我们得到 $\sqrt{3}\sin x - \cos x = 1$, 或 $\sin(x - \frac{\pi}{6}) = \frac{1}{2}$. 于是 $x = 2k\pi + \pi/3$ 或 $x = (2k+1)\pi$ ($k \in \mathbf{Z}$). 经检验，原方程的解是 $x = 2k\pi + \pi/3$ ($k \in \mathbf{Z}$).

例 9. 设 $\alpha,\beta,\gamma \in \mathbf{R}$. 求 $u = \sin(\alpha-\beta) + \sin(\beta-\gamma) + \sin(\gamma-\alpha)$ 的最大值和最小值.

解. 利用三角公式及行列式定义有

$$\begin{aligned} u &= \sin\alpha\cos\beta + \sin\beta\cos\gamma + \sin\gamma\cos\alpha - \cos\alpha\sin\beta - \cos\beta\sin\gamma - \cos\gamma\sin\alpha \\ &= \begin{vmatrix} \sin\alpha & \cos\alpha & 1 \\ \sin\beta & \cos\beta & 1 \\ \sin\gamma & \cos\gamma & 1 \end{vmatrix}. \end{aligned}$$

构造三点 $A(\sin\alpha,\cos\alpha), B(\sin\beta,\cos\beta), C(\sin\gamma,\cos\gamma)$, 那么 u 就是三角形 ABC 面积 (带正号或负号) 的二倍. 注意到 $\triangle ABC$ 是内接在单位圆周 $x^2 + y^2 = 1$ 上的三角形，当其为正三角形时达到最大面积 $\frac{3\sqrt{3}}{4}$. 于是 u 最大值是 $\frac{3\sqrt{3}}{2}$, 最小值是 $-\frac{3\sqrt{3}}{2}$.

第 4 节　Laplace 展开定理与行列式乘积

这一节我们介绍行列式按行 (列) 展开的性质 6 的一个推广： Laplace 定理. 它说一个行列式也可以按照几行 (或几列) 同时展开. 这个定理提供了计算某些行列式的一个新方法，但更重要的是可以帮助我们证明，两个 n 阶矩阵按矩阵乘法所得的乘积矩阵的行列式就是两个矩阵行列式的积，即 $|AB| = |A||B|$. 这个结论是我们研究矩阵的一个非常重要的工具.

一. Laplace 展开定理 — 按行 (列) 逐项展开的推广

定义 1. 在 n 阶行列式 D 中，任取 k 行 k 列 $(k \le n)$，位于这些行列交叉处的 k^2 个元素按它们在原行列式中的相对位置组成的 k 阶行列式 N，称为 D 的一个 k **阶子式**. 在 D 中，划去 N 所在的行和列，由剩下的元素按它们在原行列式中的相对位置组成的 $n-k$ 阶行列式 M，称为 N 的 **余子式**. 如果 N 的行与列在 D 中的行标、列标分别为 $i_1,\cdots,i_k, j_1\cdots,j_k$，则

$$A = (-1)^{(i_1+i_2+\cdots+i_k)+(j_1+\cdots+j_k)}M$$

称为 N 的 **代数余子式**.

定理 1 (Laplace 定理)**.** 设在 n 阶行列式 D 中取定 k 行 $(1<k<n)$，则 D 等于位于这 k 行的所有 k 阶子式 $N_i(i=1,\cdots,t)$ 与它们各自对应的代数余子式 A_i 的乘积之和，即

$$D = \sum_{i=1}^{t} N_i A_i, \quad t = C_n^k.$$

这个定理的证明思想与性质 6 的类似，但是写起来要繁琐得多，我们就省略了. 下面看几个例子.

例 1. 求

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

解. 第 1 、3 行中 0 比较多，故按这两行展开计算量会比较小. 由于

$$N_1 = \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = 1, N_2 = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1, N_3 = \begin{vmatrix} 2 & 1 \\ 3 & 0 \end{vmatrix} = -3,$$

所对应代数余子式分别为

$$A_1=(-1)^{1+3+1+2}\begin{vmatrix}2&3&0\\2&1&0\\4&1&3\end{vmatrix}=-3\times(2-6)=12,$$

$$A_2=(-1)^{1+5}\begin{vmatrix}1&2&3\\2&2&1\\3&4&1\end{vmatrix}=\begin{vmatrix}-5&-4&0\\2&2&1\\1&2&0\end{vmatrix}=1\times 6=6,$$

$$A_3=(-1)^{2+5}\begin{vmatrix}0&2&3\\0&2&1\\0&4&1\end{vmatrix}=0,$$

故 $D=12-6=6$.

例 2. 求

$$D_n=\begin{vmatrix}a_1&&&&&&&b_1\\&a_2&&&&&b_2&\\&&\ddots&&&\iddots&&\\&&&a_{n-1}&b_{n-1}&&&\\&&&a_n\;\;b_n&&&&\\&&&b_n\;\;a_n&&&&\\&&&b_{n-1}&a_{n-1}&&&\\&&\iddots&&&\ddots&&\\&b_2&&&&&a_2&\\b_1&&&&&&&a_1\end{vmatrix},\text{ 空白处元素为 }0.$$

解. 按第 $n,n+1$ 行展开:

$$\begin{aligned}D_n&=\begin{vmatrix}a_n&b_n\\b_n&a_n\end{vmatrix}(-1)^{n+n+1+n+n+1}D_{(n-1)}\\&=\prod_{k=1}^{n}\begin{vmatrix}a_k&b_k\\b_k&a_k\end{vmatrix}=\prod_{k=1}^{n}(a_k^2-b_k^2).\end{aligned}$$

习题 1. 计算行列式 $\begin{vmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & x_1 & x_2 & x_3 & x_4 \\ 0 & x_1^2 & x_2^2 & x_3^2 & x_4^2 \end{vmatrix}$.

二. 行列式的乘积法则

定理 2. 设 A,B 为 n 阶矩阵， AB 为其乘积，则

$$|AB| = |A||B|.$$

证明. 设 $A = (a_{ij})_n, B = (b_{ij})_n$. 考察如下 $2n$ 阶行列式

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}. \tag{20}$$

利用 Laplace 展开定理，关于前 n 行展开，容易看出 $D = |A||B|$. 另一方面，为了使 (20) 第一行前 n 个元素全为零，将第 $n+1$ 行乘以 a_{11} 加到第一行，第 $n+2$ 行乘以 a_{12} 加到第一行， $\cdots\cdots$, 将第 $2n$ 行乘以 a_{1n} 加到第一行，就得到

$$D = \begin{vmatrix} 0 & 0 & \cdots & 0 & \sum_{k=1}^n a_{1k}b_{k1} & \sum_{k=1}^n a_{1k}b_{k2} & \cdots & \sum_{k=1}^n a_{1k}b_{kn} \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}. \tag{21}$$

类似地，可以用第 $n+1$ 行乘以 a_{21} 加到第 2 行，$\cdots\cdots$, 第 $n+l$ 行乘以 a_{2l} 加到第 2 行，$\cdots\cdots$, 第 $2n$ 行乘以 a_{2n} 加到第 2 行，使得 (21) 第 2 行前 n 个元素均为零：

$$D = \begin{vmatrix} 0 & 0 & \cdots & 0 & \sum_{k=1}^n a_{1k}b_{k1} & \sum_{k=1}^n a_{1k}b_{k2} & \cdots & \sum_{k=1}^n a_{1k}b_{kn} \\ 0 & 0 & \cdots & 0 & \sum_{k=1}^n a_{2k}b_{k1} & \sum_{k=1}^n a_{2k}b_{k2} & \cdots & \sum_{k=1}^n a_{2k}b_{kn} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}. \quad (22)$$

依次使得上述行列式前 n 行的前 n 个元素均为零，我们就得到

$$D = \begin{vmatrix} 0 & 0 & \cdots & 0 & \sum_{k=1}^n a_{1k}b_{k1} & \sum_{k=1}^n a_{1k}b_{k2} & \cdots & \sum_{k=1}^n a_{1k}b_{kn} \\ 0 & 0 & \cdots & 0 & \sum_{k=1}^n a_{2k}b_{k1} & \sum_{k=1}^n a_{2k}b_{k2} & \cdots & \sum_{k=1}^n a_{2k}b_{kn} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \sum_{k=1}^n a_{nk}b_{k1} & \sum_{k=1}^n a_{nk}b_{k2} & \cdots & \sum_{k=1}^n a_{nk}b_{kn} \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}. \quad (23)$$

对 (23) 的前 n 行用 Laplace 展开，并利用矩阵定义，就得到

$$D = |AB| \times (-1)^n \times (-1)^{1+2+\cdots n+(n+1)+(n+2)+\cdots+2n} = |AB|.$$

这就证明了定理 2.

注. 这个定理的证明是比较精巧的. 在下一章我们学习了分块矩阵之后，会对这里的证明更有清晰的认识.

例 3. 设 $D = \begin{vmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{vmatrix}$， 求 D^2.

解.

$$D^2 = DD^T = \begin{vmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{vmatrix} \begin{vmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{vmatrix}$$

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

$$= (a^2+b^2+c^2+d^2)^4.$$

习题 2. 设 a,b,c,d 是不全为零的常数. 证明如下齐次线性方程组只有零解.

$$\begin{cases} ax_1+bx_2+cx_3+dx_4=0, \\ bx_1-ax_2+dx_3-cx_4=0, \\ cx_1-dx\alpha_2-ax_3+bx_4=0, \\ dx_1+cx_2-bx_3-ax_4=0. \end{cases}$$

复习题

1. 用行列式定义证明:

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

2. 利用行列式定义计算

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

中 x^4 与 x^3 的系数.

3. 设 $\begin{vmatrix} a & b & c & d \\ c & b & d & a \\ d & b & c & a \\ a & b & d & c \end{vmatrix}$ 是四阶行列式，计算 $A_{14}+A_{24}+A_{34}+A_{44}$.

4. 利用行列式及 Laplace 展开定理证明如下恒等式：

$$(ab'-a'b)(cd'-c'd)-(ac'-a'c)(bd'-b'd)+(ad'-a'd)(bc'-b'c)=0.$$

5. 计算行列式.

(1) $\begin{vmatrix} y & \cdots & y & y & x \\ y & \cdots & y & x & y \\ y & \cdots & x & y & y \\ \vdots & \iddots & \cdots & \vdots & \vdots \\ x & \cdots & y & y & y \end{vmatrix}$; (2) $\begin{vmatrix} a_0 & -1 & 0 & \cdots & 0 & 0 \\ a_1 & x & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ a_{n-2} & 0 & 0 & \cdots & x & -1 \\ a_{n-1} & 0 & 0 & \cdots & 0 & x \end{vmatrix}$;

(3) $\begin{vmatrix} a_0 & 1 & 1 & \cdots & 1 \\ 1 & a_1 & 0 & \cdots & 0 \\ 1 & 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & a_n \end{vmatrix}$; (4) $\begin{vmatrix} \cos\alpha & 1 & 0 & \cdots & 0 & 0 \\ 1 & 2\cos\alpha & 1 & \cdots & 0 & 0 \\ 0 & 1 & 2\cos\alpha & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 2\cos\alpha \end{vmatrix}$.

第 3 章　矩　阵

我们在第 1 章已经介绍了矩阵的概念及其乘法和转置．本章第 1 、 2 节基本上是对这些重要概念的复习和简单延伸．在第 3 节，我们将介绍可逆矩阵．这一节理论性较强，从中也可初窥引入矩阵乘法对于求解线性方程组的意义．第 4 节介绍分块矩阵，利用矩阵的分块可以整体地批量地处理元素的运算，从而有助于高效地施行矩阵的乘法等运算．第 5 节是本章的重点和难点，可作为数学建模的一个典型例子，充分体现了如何用矩阵语言数学化地描述我们熟知的求解线性方程组的消元法．这一节概念多、想法多，需要多总结归纳．最后，在第 6 节，我们用第 5 节的理论和矩阵秩的概念，给出判断一个一般线性方程组是否有解，并在有解情形写出所有解的表达式的高效方法.

第 1 节　矩阵的概念

一．矩阵的定义

定义 1. 由 $m\times n$ 个数排成的 m 行 n 列的矩形 **数表**

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

称为一个 $m\times n$**矩阵**. 这里 a_{ij} 表示位于矩阵第 i 行第 j 列的数．构成矩阵的这些数也叫做矩阵的 **元素**.

实矩阵：元素都是实数的矩阵.

复矩阵：元素为复数的矩阵.

矩阵常用大写字母 A,B,C 等表示，或更具体地记作

$$[a_{ij}]_{m\times n}, \qquad (a_{ij})_{m\times n},$$

表示以 a_{ij} 为元素的 m 行 n 列的矩阵. 我们也常用 A_{ij} 表示矩阵 A 的位于第 i 行第 j 列的元素. 此外请注意区分行列式与矩阵. 行列式的行数与列数必须相同，表示一种对 $a_{ij}(1 \le i, j \le n)$ 这 n^2 个数的运算，实际上是一个数，而矩阵就是一个表格. 它是和数不同的一种新的数学对象，可看作数的一种推广.

下面是两个矩阵的例子：

$$\begin{pmatrix} 1 & -1 & a & \sqrt{7} \\ 0 & \pi & 0.23 & 3i \end{pmatrix}, \qquad \begin{pmatrix} 0 & 23 \\ d & 3+6.1i \\ f & 4 \\ 0 & 0 \end{pmatrix}.$$

这里 i 是虚数单位 $(i^2 = -1)$, a, d, f 表示数， π 是圆周率.

定义 2. 具有相同行数和列数的两个矩阵称作同类型矩阵. 对于同类型的两个矩阵 $[a_{ij}]_{m\times n}$ 和 $[b_{ij}]_{m\times n}$, 定义他们相等当且仅当对应元素相同： $a_{ij} = b_{ij}$.

例如, 若 $\begin{pmatrix} 1 & 0 & e \\ \frac{3}{4} & \pi & -0.2 \end{pmatrix} = \begin{pmatrix} x & y & z \\ u & v & w \end{pmatrix}$, 则可知 $x = 1, y = 0, z = e, u = 3/4, v = \pi, w = -0.2$.

注意，不同类型的矩阵不能定义 “相等 ”.

二. 常用特殊矩阵

下面这些矩阵，就像数字中的 0, 1, -1, i 一样，有着独特的作用，所以需要大家记住.

♦ <u>行矩阵 (行向量) 与列矩阵 (列向量)</u>: 即仅有一行 (一列) 的矩阵 A, 可等同于一个向量 $\mathbf{a}$:

$$A = (a_{11} \;\; a_{12} \;\; \cdots \;\; a_{1n}), \qquad \mathbf{a} = (a_{11}, a_{12}, \cdots, a_{1n})$$

♦ <u>零矩阵</u>: 所有元素均为零的矩阵，记作 O 或 0. m 行 n 列的零矩阵也记作 $O_{m\times n}$.

♦ <u>方阵</u>: 行数与列数相同的矩阵.

例. n 阶方阵

$$A=A_n=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{m2} & \cdots & a_{nn} \end{pmatrix}.$$

主对角线：左上到右下的对角线.

主对角元：在矩阵主对角线上的元素，即 $a_{11},a_{22},\cdots,a_{nn}$.

方阵的行列式：按方阵 A 元素的排列方式所构造的 n 阶行列式叫做方阵 A 的行列式，记作 $|A|$ 或 $\det A$. 即

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

♦ 对角阵: 除主对角元外其它元素均为零的方阵.

$$A=\begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}=\mathrm{diag}(a_{11},a_{22},\cdots,a_{nn}).$$

约定：未写出的元素均为零.

例. $\mathrm{diag}(0,1,3,0)=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$

♦ 单位矩阵: 主对角元全为 1 的对角矩阵. 记为 I 或 E.

$$I(=I_n)=E(=E_n)=\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}_{n\times n}$$

♦ <u>数量矩阵</u>: 主对角元全为相同常数的对角矩阵.

♦ <u>三角矩阵</u>: 主对角线下 (上) 元素全为零的方阵叫上 (下) 三角矩阵.

例. 上三角阵: $\begin{pmatrix} 0 & 3 & 0 & 0 & 0 \\ & 7 & 0 & -2 & 8 \\ & & 0 & 0 & 0 \\ & & & 0.9 & -1 \\ & & & & -11 \end{pmatrix}$, 下三角阵: $\begin{pmatrix} 1 & & \\ -1 & 0 & \\ 0 & -0.7 & 0 \end{pmatrix}$.

♦ <u>对称矩阵, 反对称矩阵</u>: 若方阵 $A=(a_{ij})_n$ 的元素满足

$$a_{ij}=a_{ji}, \quad (a_{ij}=-a_{ji}),$$

称 A 是对称 (反对称) 矩阵. 实的对称矩阵叫实对称阵.

例. 反对称阵: $\begin{pmatrix} 0 & 1 & e \\ -1 & 0 & -\pi \\ -e & \pi & 0 \end{pmatrix}$, 对称阵: $\begin{pmatrix} 1 & s & 0.7 & 0 \\ s & 2 & 0 & 8 \\ 0.7 & 0 & 4 & -31 \\ 0 & 8 & -31 & -7 \end{pmatrix}$.

习题 1. 证明反对称矩阵主对角元必为零; 矩阵 A 对称当且仅当 $A=A^T$, 反对称当且仅当 $A=-A^T$.

第 2 节 矩阵的运算

在第 1 章我们已经学习了矩阵与矩阵的乘法及矩阵的转置. 这一节里, 我们还将介绍矩阵的加法 (减法), 数乘以矩阵, 以及 方阵的幂. 引入这些运算有助于把复杂的矩阵分解化简为简单的矩阵.

在这一节的学习中, 大家要特别注意矩阵运算与数的运算的 **相同与不同** 之处, 避免以后 “想当然 ” 地使用某些对矩阵来说是错误的 “运算法则 ”.

一. 数乘和加法

定义 1. *(1)* 设 A,B 为同类型矩阵, 定义加法 $A+B$: $(A+B)_{ij}=A_{ij}+B_{ij}$.

(2) 定义数乘 kA: $(kA)_{ij}=kA_{ij}$, k 为一个数. 特别的, $-A=(-1)A$.

(3) 减法: $A-B=A+(-B)$.

例如，我们有

$$\begin{pmatrix} 2 & 0 \\ 4 & 6 \\ 2 & 3 \end{pmatrix} - 2\begin{pmatrix} 5 & 1 \\ 6 & 2 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 2-2\times 5 & 0-2\times 1 \\ 4-2\times 6 & 6-2\times 2 \\ 2-2\times 4 & 3-2\times 3 \end{pmatrix} = \begin{pmatrix} -8 & -2 \\ -8 & 2 \\ -6 & -3 \end{pmatrix}.$$

例 1. 设 $A = \begin{pmatrix} 3 & 0 \\ 4 & -3 \end{pmatrix}$, $B = \begin{pmatrix} -1 & -2 \\ 6 & 0 \end{pmatrix}$, 且 $2A - 3X = B$, 求矩阵 X.

解. $X = \frac{1}{3}(2A - B) = \begin{pmatrix} 7/3 & 2/3 \\ 2/3 & -2 \end{pmatrix}$.

习题 1. 证明： (1) $A+B = B+A$; (2) $k(A+B) = kA+kB$; (3) $(A+B)+C = A+(B+C)$.

二. 矩阵的乘法 (* 关键点 *)

定义 2. $A = [a_{ik}]_{m\times p}$, $B = [b_{kj}]_{p\times n}$. 定义 $C = [c_{ij}]_{m\times n}$ 为 A, B 的乘积，记作：$C = AB$. 其中

$$\boxed{c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.}$$

换句话说，乘积矩阵 AB 的第 i 行第 j 列的元素是由 A 的第 i 行与 B 的第 j 列作内积得到的：

$$\begin{pmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{pmatrix}_{m\times p} \begin{pmatrix} * & \cdots & b_{1j} & \cdots & * \\ * & \cdots & b_{2j} & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ * & \cdots & b_{pj} & \cdots & * \end{pmatrix}_{p\times n}$$

$$= \begin{pmatrix} * & \cdots & * & \cdots & * \\ * & \cdots & * & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ * & \cdots & c_{ij} & \cdots & * \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ * & \cdots & * & \cdots & * \end{pmatrix}_{m\times n}.$$

例. $\begin{pmatrix} 3 & 2 \\ -1 & 4 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} 1 & 8 & -1 \\ 2 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 7 & 24 & 3 \\ 7 & -8 & 13 \\ 7 & 40 & -2 \end{pmatrix}$.

特别注意: I. 矩阵乘法不满足交换律:

(a) 即使 A,B 可以相乘, B,A 未必可以相乘;

(b) 即使 AB, BA 都有定义, 它们也未必是相同类型的;

(c) 即使 AB, BA 同类型, 也未必相等.

例. $A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}$ 则 $AB = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $BA = \begin{pmatrix} -2 & -1 \\ 3 & 2 \end{pmatrix}$.

习题 2. 举例说明 (a)(b).

特别注意: II. 两个非零矩阵的乘积可以为零矩阵.

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

特别注意: III. 矩阵的乘积消去律不成立: $AB = CB$ 且 $B \neq O$ 未必可推出 $A = C$.

例. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}, B = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, C = \begin{pmatrix} 2 & 0 \\ 4 & 5 \end{pmatrix}$,

$AB = CB = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}$.

不过矩阵乘法仍然有许多好的性质.

运算法则: (1) 乘法零元: $OA = O, AO = O$.

(2) 乘法单位元: $EA = AE = A$.

(3) 乘法结合律: $(AB)C = A(BC)$, $(kA)B = k(AB) = A(kB)$.

(4) 乘法分配律: $A(B+C) = AB + AC$, $(B' + C')A' = B'A' + C'A'$.

(5) 行列式乘法: $|AB| = |A||B|$.

习题 2. 证明 (1)(2)(4).

下面给出 (3) 的证明. ((5) 在第 2 章已经学过.)

在证明这种最基本的关系式时唯一的办法就是利用矩阵乘法及矩阵相等的定义. 设 A 为 $m \times n$ 矩阵, B 为 $n \times p$ 矩阵, C 为 $p \times q$ 矩阵. 则 AB 为 $m \times p$ 矩阵, BC 为 $n \times q$ 矩阵. 我们考虑矩阵 $(AB)C$ 的第 i 行第 j 列的元素

$((AB)C)_{ij}$:

$$\begin{aligned}
((AB)C)_{ij} &= \sum_{k=1}^{p}(AB)_{ik}C_{kj} = \sum_{k=1}^{p}(\sum_{l=1}^{n}A_{il}B_{lk})C_{kj} \qquad \text{(利用了矩阵乘法定义)} \\
&= \sum_{k=1}^{p}(\sum_{l=1}^{n}A_{il}B_{lk}C_{kj}) \qquad \text{(利用了数的乘法分配律)} \\
&= \sum_{l=1}^{n}\sum_{k=1}^{p}A_{il}B_{lk}C_{kj} \qquad \text{(利用了数的加法交换律)} \\
&= \sum_{l=1}^{n}A_{il}(\sum_{k=1}^{p}B_{lk}C_{kj}) \qquad \text{(利用了数的乘法结合律)} \\
&= \sum_{l=1}^{n}A_{il}(BC)_{lj} \qquad \text{(利用了矩阵乘法定义)} \\
&= (A(BC))_{ij}. \qquad \text{(再用矩阵乘法定义)}
\end{aligned}$$

例 2. 证明两个 n 阶上三角方阵的乘积仍是上三角阵.

证明. 首先, 按定义, 一个矩阵 A 是上三角阵, 就相当于当 $i>j$ 时 $A_{ij}=0$. 现设 A,B 是两个上三角 n 阶方阵. 那么，当 $i>j$ 时,

$$(AB)_{ij} = \sum_{k=1}^{n}A_{ik}B_{kj} = \sum_{k=1}^{i-1}A_{ik}B_{kj} + \sum_{k=i}^{n}A_{ik}B_{kj} = \sum_{k=i}^{n}A_{ik}B_{kj}.$$

但由于 $i>j$, 故在上面最后一个和式中， $k=i,i+1,\cdots,n$, 均大于 j, 从而 $B_{kj}=0$. 于是上面和式为零，即得 $i>j$ 时 $(AB)_{ij}=0$. 所以 AB 仍是上三角阵.

习题 3. 设 A 是 n 阶方阵，且对于任意 $\mathbf{x}=(x_1,x_2,\cdots,x_n)^T$ 均成立 $A\mathbf{x}=\mathbf{0}=(0,0,\cdots,0)^T$. 证明 $A=O$.

三. 矩阵的幂

定义 3. 设 A 为 n 阶方阵， m 为自然数，定义 $A^m=AA\cdots A$ (m 个 A 相乘). 对多项式 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots a_2x^2+a_1x+a_0$, 定义

$$f(A)=a_nA^n+a_{n-1}A^{n-1}+\cdots a_2A^2+a_1A+a_0E.$$

运算法则: (1) $A^kA^l=A^{k+l}$, $(A^k)^l=A^{kl}$.

(2) $|A^m| = |A|^m$, $|kA| = k^n|A|$.

特别注意：未必成立 $(AB)^k = A^kB^k$.

习题 4. 证明 (1)(2).

矩阵函数有着广泛的应用. 例如, 我们知道常微分方程 $u'(t) = au(t)$ 的解是指数函数 $u(t) = ce^{at}$. 如果考察齐次常系数线性常微分方程组 $\mathbf{x}'(t) = A\mathbf{x}(t)$, (这里 $\mathbf{x}'(t)$ 表示对向量 $\mathbf{x}(t)$ 的每个分量 (还是函数) 求导数， A 是一个矩阵,) 它的解就可以写为 $\mathbf{x}(t) = e^{tA}\mathbf{b}$, 其中 $\mathbf{b}$ 是一个向量， tA 是时间 t 与 A 的数乘，而矩阵指数函数定义为

$$e^A = E + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^n}{n!} + \cdots.$$

这里利用了 Taylor 级数 $e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!}$. 当然，这样的无限和是否有意义，如何计算，都需要对矩阵深入的了解，比如其特征值以及是否可对角化等. 这些内容本课程就不涉及了. 下面只介绍一个典型的情形.

例 3. 设 $A = \begin{pmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}$.(空白处元素为零.) 求 A^n.

解. 记 $B = \begin{pmatrix} & 1 & \\ & & 1 \\ & & \end{pmatrix}$, 则 $A = \lambda E + B$, 其中 E 为三阶单位阵. 由于 $EB = BE$, 所以我们可以用二项式公式得到

$$A^n = \sum_{k=1}^{n} C_n^k (\lambda E)^{n-k} B^k = \sum_{k=1}^{n} C_n^k \lambda^{n-k} B^k.$$

另一方面直接计算得到 $B^2 = \begin{pmatrix} & & 1 \\ & & \\ & & \end{pmatrix}$, $B^3 = O$. 所以

$$A^n = \lambda^n E + n\lambda^{n-1} B + \frac{n(n-1)}{2}\lambda^{n-2} B^2 = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \frac{n(n-1)}{2}\lambda^{n-2} \\ & \lambda^n & n\lambda^{n-1} \\ & & \lambda^n \end{pmatrix}.$$

习题 5. 设

$$(1)\ f(\lambda)=\lambda^2-\lambda-1,\quad A=\begin{pmatrix}2&1&1\\3&1&2\\1&-1&0\end{pmatrix};$$

$$(2)\ f(\lambda)=\lambda^2-5\lambda+3,\quad A=\begin{pmatrix}2&-1\\-3&3\end{pmatrix}.$$

求 $f(A)$.

四. 矩阵的转置 (* 关键点 *)

定义 4. 对 $m\times n$ 阶矩阵 $A=[a_{ij}]$, 定义 $n\times m$ 阶矩阵 $B=[b_{ij}]$, 其中

$$b_{ij}=a_{ji},$$

称 B 为 A 的转置，记作 A^T. 所以，依定义，成立 $(A^T)_{ij}=A_{ji}$.

$$\text{例.}\quad \begin{pmatrix}1&0&3&t\\a_{21}&a_{22}&a_{23}&a_{24}\\2&1&-1&-0.9\end{pmatrix}^T=\begin{pmatrix}1&a_{21}&2\\0&a_{22}&1\\3&a_{23}&-1\\t&a_{24}&-0.9\end{pmatrix}.$$

运算法则: (1) $(A^T)^T=A$

(2) $(B+C)^T=B^T+C^T$

(3) $(kA)^T=kA^T$

(4) $(AB)^T=B^TA^T$, $(A_1A_2\cdots A_m)^T=A_m^TA_{m-1}^T\cdots A_2^TA_1^T$

(5) $(A^m)^T=(A^T)^m$

(6) $|A|=|A^T|$

(7) A 为对称阵当且仅当 $A^T=A$, A 为反对称阵当且仅当 $A^T=-A$.

习题 6. 证明 (2)(3)(5).

下面给出 (1)(4) 的证明. (6) 在第 2 章已经学过.

(1) 的证明. $((A^T)^T)_{ij}=(A^T)_{ji}=A_{ij}$.

(4) 的证明. 设 A 为 $m\times n$ 矩阵，B 为 $n\times p$ 矩阵. 则 $((AB)^T)_{ij}=(AB)_{ji}=\sum_{k=1}^n A_{jk}B_{ki}=\sum_{k=1}^n(B^T)_{ik}(A^T)_{kj}=(B^TA^T)_{ij}$. 后一部分用数学归纳法. $m=2$

是刚才所证明的．现设对 m 已成立，考虑 $m+1$ 个矩阵相乘．则

$$\begin{aligned}&(A_1A_2\cdots A_mA_{m+1})^T=((A_1A_2\cdots A_m)(A_{m+1}))^T\\=\ &A_{m+1}^T((A_1A_2\cdots A_m))^T=A_{m+1}^T(A_m^TA_{m-1}^T\cdots A_2^TA_1^T)\\=\ &A_{m+1}^TA_m^TA_{m-1}^T\cdots A_2^TA_1^T.\end{aligned}$$

得证！

在证明这些性质的时候，我们大多需要按照定义，用分量来计算．但是在以后证明中，要尽量利用这些整体的性质，而避免继续按照分量去证明.

例 4. 令 $\alpha=\begin{pmatrix} a_1 & a_2 & a_3\end{pmatrix}^T,\beta=\begin{pmatrix} b_1 & b_2 & b_3\end{pmatrix}$. 求 $(\alpha\beta)^n$.

解. $\beta\alpha=k:=\sum_{i=1}^3 a_ib_i$, 又 $(\alpha\beta)^n=\alpha(\beta\alpha)^{n-1}\beta=k^{n-1}\alpha\beta$.

习题 7. 设 A 为 $n\times 1$ 矩阵, 且 $A^TA=1,E_n$ 为 n 阶单位阵. 证明：$B=E_n-2AA^T$ 是对称阵且 $B^2=E_n$.

习题 8. 设 A 是实对称阵且 $A^2=O$. 证明： $A=O$.

第 3 节　可逆矩阵

在本章第 1 节我们指出， n 阶单位阵 E 与任何 n 阶方阵 A 相乘还是得到 A:

$$EA=AE=A.$$

现在考虑线性方程组

$$AX=B,$$

其中 A 为已知 n 阶方阵， X 为待求 $n\times p$ 矩阵， B 为已知 $n\times p$ 矩阵．如果我们可以找到一个 n 阶方阵 A^{-1}, 使得 $A^{-1}A=E$, 那么在上述方程两边同是左乘以 A^{-1}, 利用乘法结合律我们就得到

$$A^{-1}AX=A^{-1}B,\quad \text{或}\ X=A^{-1}B.$$

这样就可以一下子求出所有未知数了．本节的目的，就在于研究对什么样的矩阵 A, 可以存在 A^{-1} (这样的矩阵称为可逆矩阵), 并对可逆矩阵 A 找出求 A^{-1} 的办法.

一. 可逆矩阵的定义

定义 1. 设 A 是 n 阶矩阵. 若存在 n 阶矩阵 B 使得

$$AB = BA = E, \tag{1}$$

则称矩阵 A**可逆**, B 为 A 的 **逆矩阵**, 记作 $B = A^{-1}$.

如果不存在满足 (1) 的矩阵 B, 就称 A 是 **不可逆** 的，或 **奇异** 的.

很自然的，我们有如下 **问题**：什么样的矩阵是可逆的？可逆矩阵的逆是否唯一，如何求解？可逆矩阵有什么性质？下面将一一回答这些问题. 首先，从定义，我们有

定理 1. 如果 n 阶矩阵 A 可逆，则它的逆是唯一的.

证明. 设 B 和 C 都是 A 的逆. 那么

$$B = BE = B(AC) = (BA)C = EC = C.$$

所以 A 的逆唯一.

从这个证明可以看出，为保证逆的唯一性，定义 (1) 要求无论 A^{-1} 和 A 左乘还是右乘都是单位阵是必须的. 此外，"'1' 的替换" 也是一个值得注意的技巧.

二. 矩阵可逆的充分必要条件和求可逆矩阵的逆的伴随矩阵法

首先我们引入如下重要定义.

定义 2. (* 伴随矩阵 *) 设 $A = [a_{ij}]_n$ 为 n 阶方阵，$\mathscr{A}_{ij}$ 为行列式 $|A|$ 中元素 a_{ij} 的 **代数余子式**, $i, j = 1, 2, \cdots, n$. 则称矩阵

$$\begin{pmatrix} \mathscr{A}_{11} & \mathscr{A}_{21} & \cdots & \mathscr{A}_{n1} \\ \mathscr{A}_{12} & \mathscr{A}_{22} & \cdots & \mathscr{A}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathscr{A}_{1n} & \mathscr{A}_{2n} & \cdots & \mathscr{A}_{nn} \end{pmatrix} \tag{2}$$

为 A 的 **伴随矩阵**, 记为 A^*.

请特别注意这里代数余子式 $\mathscr{A}_{ij}$(回忆第 2 章, 它是一个数) 的排列方法. 例如，若

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad 则\ A^*=\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

关于可逆矩阵的判定，我们有如下重要定理.

定理 2. *n 阶方阵 A 可逆的充分必要条件为 $|A|\neq 0$. 如果 A 可逆，则*

$$A^{-1}=\frac{1}{|A|}A^*. \tag{3}$$

证明. 必要性. 在 (1) 两边取行列式，得 $|A||B|=1$, 所以 $|A|\neq 0$.

充分性. 此时由 (3), 只需验证 $A^*A=AA^*=|A|E$. 这利用行列式按行 (列) 展开性质就可得到. 例如，考察矩阵 A^*A 的第 i 行第 j 列的元素 $(A^*A)_{ij}$: $(A^*A)_{ij}=\sum_{k=1}^n {A^*}_{ik}a_{kj}=\sum_{k=1}^n \mathscr{A}_{ki}a_{kj}=\delta_{ij}|A|$. 从而 $A^*A=|A|E$.

推论 1. 对 n 阶方阵 A, 若有 n 阶方阵 B 使得

$$AB=E \qquad (或\ \ BA=E),$$

则 A 可逆，且 $A^{-1}=B$.

证明. 由条件有 $|A|\neq 0$, 从而 A 可逆. 又 $B=(A^{-1}A)B=A^{-1}(AB)=A^{-1}$.

由此推论，以后只要 $AB=E$ 和 $BA=E$ 有一个成立，就可判定 $B=A^{-1}$. 这就比直接用定义 1 简单了一些.

推论 2. (Cramer 法则) 线性方程组 $A\mathbf{x}=\mathbf{b}$ 当 $|A|\neq 0$ 时有唯一解 $\mathbf{x}=A^{-1}\mathbf{b}$.

我们解释一下为什么说这就是第 2 章学习过的 Cramer 法则. 实际上，具体写出 x_i:

$$\begin{aligned} x_i &= \sum_{j=1}^n (A^{-1})_{ij}b_j=\sum_{j=1}^n \frac{1}{|A|}(A^*)_{ij}b_j \ \ ((A^*)_{ij}\ 表示伴随矩阵的第\ ij\ 个元素) \\ &= \sum_{j=1}^n \frac{1}{|A|}\mathscr{A}_{ji}b_j \ \ (\mathscr{A}_{ji}\ 表示\ |A|\ 中第\ ji\ 个元素的代数余子式) \\ &= \frac{1}{|A|}D_i. \quad (D_i\ 表示\ |A|\ 中第\ i\ 列元素代换为\ \mathbf{b}) \end{aligned}$$

上述定理 2 给出了 A^{-1} 的一个表达式. 直接按这个公式求 A^{-1} 的方法称为伴随矩阵法. 和 Cramer 法则类似，这个方法的价值主要在理论研究方

面．具体实际计算一个可逆矩阵的逆时，若用这个方法，则计算量很大．我们在第 5 节将介绍求逆矩阵的另一种高效的方法，即初等变换法．

习题 1. 利用伴随矩阵法求矩阵 $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 2 \\ -1 & 3 & 0 \end{pmatrix}$ 的逆．

习题 2. 求解矩阵方程 $\begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix} X = \begin{pmatrix} -77 & -70 & -99 \\ -45 & -41 & -58 \end{pmatrix}$．

例 1. *设可逆矩阵 A 的元素都是整数．证明：A^{-1} 的元素也都是整数，当且仅当 $|A| = \pm 1$.*

证明．充分性．若 A 的元素都是整数，则由定义不难看出 A^* 的元素也都是整数．所以根据 $A^{-1} = A^*/|A| = \pm A^*$, 可知 A^{-1} 的元素均为整数．

必要性．若 A, A^{-1} 的元素都是整数，则由行列式定义不难看出 $|A|, |A^{-1}|$ 都是整数．但是 $|A||A^{-1}| = 1$, 从而必有 $|A| = \pm 1$.

例 2. *设 n 阶方阵 A 可逆且每一行元素之和均为常数 c. 证明：(a) $c \neq 0$; (b) A^{-1} 的每一行元素之和均为 $\frac{1}{c}$.*

证明． (a) 若 $c = 0$, 那么我们就有 $A(1, \cdots, 1)^T = (c, \cdots, c)^T = \mathbf{0}$. 这里 $(1, \cdots, 1)^T$ 是由 n 个 1 构成的列向量，$(c, \cdots, c)^T$ 是由 n 个 c 构成的列向量．于是齐次线性方程组 $A\mathbf{x} = \mathbf{0}$ 有非零解，与 Cramer 法则矛盾 (注意 A 可逆的条件). 所以必有 $c \neq 0$.

(b) 由 $A(1, \cdots, 1)^T = (c, \cdots, c)^T$, 两边同时左乘 A^{-1}, 就得到 $A^{-1}(c, \cdots, c)^T = (1, \cdots, 1)^T$. 两边同时乘以 $\frac{1}{c}$, 就得到 $A^{-1}(1, \cdots, 1)^T = \frac{1}{c}(1, \cdots, 1)^T$. 证毕．

习题 3. 设 A 为 n 阶方阵且 $|A| = \frac{1}{3}$. 求 $|(\frac{1}{4}A)^{-1} - 15A^*|$.

三．可逆矩阵的性质

下面关于可逆矩阵的性质的定理要熟练掌握．

定理 3. *(1)* $(A^{-1})^{-1} = A$.

(2) $(kA)^{-1} = k^{-1}A^{-1}, \quad k \neq 0$.

(3) $(AB)^{-1} = B^{-1}A^{-1}$, $(A_1A_2 \cdots A_m)^{-1} = A_m^{-1} \cdots A_2^{-1}A_1^{-1}$.

(4) $(A^T)^{-1} = (A^{-1})^T$.

(5) $|A^{-1}| = \frac{1}{|A|}$.
(6) $(A^m)^{-1} = (A^{-1})^m$.

证明. (3) 由 $(AB)(B^{-1}A^{-1}) = E$ 即可看出. 后一部分用数学归纳法即可.

(4) 由 $A^T(A^{-1})^T = (A^{-1}A)^T = E^T = E$ 即知. 证毕.

习题 4. 证明 (1)(2)(5)(6).

例 3. 设 n 阶矩阵 A, B 满足 $A + B = AB$, 证明 $A - E$ 可逆并求出其逆.

证明. 由题意知 $AB - EB - (A - E) = E$, 从而 $(A - E)B - (A - E)E = E$, 由结合律得 $(A - E)(B - E) = E$. 所以 $A - E$ 可逆且 $(A - E)^{-1} = B - E$.

解决这类问题, 大家可以先考虑特殊情况, 即 A, B 都是数 $(n = 1)$, $E = 1$, 利用熟悉的代数关系式 $ab - a - b + 1 = (a - 1)(b - 1)$ 得到关系式 $(a - 1)(b - 1) = 1$, 从而可猜测 $(A - E)^{-1} = B - E$, 然后再严格地按照矩阵相应的运算律予以验证.

例 4. 设 n 阶矩阵 $A, B, A + B$ 均可逆, 证明 $A^{-1} + B^{-1}$ 可逆并求出其逆.

证明. 若 A, B 都是数字, 则通过分数通分可得 $A^{-1} + B^{-1} = A^{-1}(A + B)B^{-1}$. 由此可猜测 $(A^{-1} + B^{-1})^{-1} = B(A + B)^{-1}A$. 下面予以验证: $(A^{-1} + B^{-1})B(A + B)^{-1}A = (A^{-1}B + E)(A + B)^{-1}A = (A^{-1}B + A^{-1}A)(A + B)^{-1}A = A^{-1}(B + A)(A + B)^{-1}A = A^{-1}A = E$.

例 5. 设 A 为 $n(n \geq 2)$ 阶矩阵. 证明 $|A^*| = |A|^{n-1}$.

证明. $|A| \neq 0$ 的情形. 利用 (3), 我们有 $AA^* = |A|E$, 两边取行列式得 $|A||A^*| = |A|^n$. 两边同除 $|A|$ 得到 $|A^*| = |A|^{n-1}$.

$|A| = 0$ 但 A 不是零矩阵的情形. 此时成立 $AA^* = O$. 现要证 $|A^*| = 0$. 反证法. 否则, A^* 可逆, 从而有 $A = A(A^*(A^*)^{-1}) = (AA^*)(A^*)^{-1} = O$, 矛盾.

若 A 是零矩阵, 则显然成立. 证毕.

习题 5. 设 A 为 n 阶方阵且不是单位阵, $n \geq 2$. 若 $A^2 = A$, 证明: A^* 不可逆.

习题 6. 若 A, B 都是 n 阶矩阵, $B, A - E$ 可逆, 且 $(A - E)^{-1} = (B - E)^T$. 证明: A 可逆.

第 4 节　矩阵的分块

矩阵的分块, 就是将一个矩阵适当地分为若干个小矩阵. 这是可以实现集成高效的矩阵运算的一种方法.

例如, 对矩阵

$$A := \begin{pmatrix} 1 & 3 & -2 & 9 & 0 \\ 0 & -2 & -5 & -3 & -1 \\ 0 & 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & -2 & -3 \end{pmatrix},$$

令

$$A_{11} = \begin{pmatrix} 1 & 3 & -2 \\ 0 & -2 & -5 \end{pmatrix}, \qquad A_{21} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$A_{12} = \begin{pmatrix} 9 & 0 \\ -3 & -1 \end{pmatrix}, \qquad A_{22} = \begin{pmatrix} 3 & -1 \\ -2 & -3 \end{pmatrix}.$$

则 $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$.

将矩阵表为列向量 (或行向量) 的形式: $A = (\alpha_1, \cdots, \alpha_n)$, 也是一种分块矩阵.

本节重点: 熟悉分块矩阵的运算.

一. 分块矩阵的加法与数乘

对分块矩阵

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{pmatrix} =: (A_{ij})_{p\times q},$$

$$B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1q} \\ B_{21} & B_{22} & \cdots & B_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ B_{p1} & B_{p2} & \cdots & B_{pq} \end{pmatrix} =: (B_{ij})_{p\times q},$$

其中 $A_{ij}, B_{ij}(i=1,\cdots,p, j=1,\cdots q)$ 为同类型矩阵，则 **容易验证**

$$A+B=\begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} & \cdots & A_{1q}+B_{1q} \\ A_{21}+B_{21} & A_{22}+B_{22} & \cdots & A_{2q}+B_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1}+B_{p1} & A_{p2}+B_{p2} & \cdots & A_{pq}+B_{pq} \end{pmatrix} =: (A_{ij}+B_{ij})_{p\times q};$$

若 k 为数，

$$kA=\begin{pmatrix} kA_{11} & kA_{12} & \cdots & kA_{1q} \\ kA_{21} & kA_{22} & \cdots & kA_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ kA_{p1} & kA_{p2} & \cdots & kA_{pq} \end{pmatrix} =: (kA_{ij})_{p\times q}.$$

二. 分块矩阵的乘法 (* 重点 *)

设 A 为 $m\times p$ 矩阵， B 为 $p\times n$ 矩阵. 现对 A, B 分块，要求对 A **的列的分法与对** B **的行的分法一样** (但 A 的行与 B 的列的分法可各自随意), 则可以开展分块矩阵的乘法.

例.

$$A = \begin{array}{cc} \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix}, & \begin{array}{c} m_1\text{行} \\ m_2\text{行} \\ \vdots \\ m_r\text{行} \end{array} \\ \begin{array}{cccc} p_1\text{列} & p_2\text{列} & \cdots & p_s\text{列} \end{array} & \end{array}$$

$$B = \begin{array}{cc} \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1t} \\ B_{21} & B_{22} & \cdots & B_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{st} \end{pmatrix}, & \begin{array}{c} p_1\text{行} \\ p_2\text{行} \\ \vdots \\ p_s\text{行} \end{array} \\ \begin{array}{cccc} n_1\text{列} & n_2\text{列} & \cdots & n_t\text{列} \end{array} & \end{array}$$

这里 $m_1+m_2+\cdots+m_r=m,\ p_1+p_2+\cdots p_s=p,\ n_1+n_2+\cdots n_t=n$. 令分块矩阵

$$C=(C_{ij})_{r\times t},\qquad C_{ij}=\sum_{k=1}^{s}A_{ik}B_{kj},\ i=1,\cdots,r;\ j=1,\cdots,t.$$

则 **容易验证** $C=AB$.

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

解. 记 $A_1=\begin{pmatrix}1&-1\\3&-1\\0&1\end{pmatrix}$, $O=\begin{pmatrix}0&0\\0&0\\0&0\end{pmatrix}$, $O=\begin{pmatrix}0&0\end{pmatrix}$, $A_4=\begin{pmatrix}2&-1\end{pmatrix}$, $B_1=\begin{pmatrix}1\\-1\end{pmatrix}$, $O=\begin{pmatrix}0&0&0\\0&0&0\end{pmatrix}$, $O=\begin{pmatrix}0\\0\end{pmatrix}$, $B_4=\begin{pmatrix}1&3&-1\\2&1&4\end{pmatrix}$. 那么

$$\begin{pmatrix}A_1&O\\O&A_4\end{pmatrix}\begin{pmatrix}B_1&O\\O&B_4\end{pmatrix}=\begin{pmatrix}A_1B_1&O\\O&A_4B_4\end{pmatrix}=\begin{pmatrix}2&0&0&0\\4&0&0&0\\-1&0&0&0\\0&0&5&-6\end{pmatrix}.$$

我们看到，引入分块矩阵后可以直接看到某些块相乘之后为零，从而可避免做一些 “无用功 ”, 提高计算效率. 此外，分块矩阵也是处理一些理论问题的有效工具.

例 2. 设 A,B,C,D 都是 n 阶矩阵， $|A|\neq 0$, 且 $AC=CA$. 证明: $\begin{vmatrix}A&B\\C&D\end{vmatrix}=|AD-CB|$.

证明. 由分块矩阵乘法成立 $\begin{pmatrix}E&O\\-CA^{-1}&E\end{pmatrix}\begin{pmatrix}A&B\\C&D\end{pmatrix}=\begin{pmatrix}A&B\\O&-CA^{-1}B+D\end{pmatrix}$. 两边取行列式得到 $\begin{vmatrix}E&O\\-CA^{-1}&E\end{vmatrix}\begin{vmatrix}A&B\\C&D\end{vmatrix}=\begin{vmatrix}A&B\\O&-CA^{-1}B+D\end{vmatrix}$. 由 Laplace 展开定理, $\begin{vmatrix}E&O\\-CA^{-1}&E\end{vmatrix}=|E||E|=1$, $\begin{vmatrix}A&B\\O&-CA^{-1}B+D\end{vmatrix}=|A||D-CA^{-1}B|$.

所以我们得到 $\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A||D-CA^{-1}B| = |AD-ACA^{-1}B| = |AD-CAA^{-1}B| = |AD-CB|$. 证毕.

回忆第 2 章中行列式乘积定理的证明，本质也是用了如下分块矩阵乘法：$\begin{pmatrix} E & A \\ O & E \end{pmatrix}\begin{pmatrix} A & O \\ -E & B \end{pmatrix} = \begin{pmatrix} O & AB \\ -E & B \end{pmatrix}$. 两边取行列式，利用 Laplace 定理，右边就是 $|AB|$, 而左边乘积 $\begin{pmatrix} E & A \\ O & E \end{pmatrix}\begin{pmatrix} A & O \\ -E & B \end{pmatrix}$ 就相当于某一行乘适当常数加到另外一行 (参件第 2 章中相应证明，下一节学习初等矩阵之后会对此看得更清楚). 从而由行列式性质，$\left|\begin{pmatrix} E & A \\ O & E \end{pmatrix}\begin{pmatrix} A & O \\ -E & B \end{pmatrix}\right| = \begin{vmatrix} A & O \\ -E & B \end{vmatrix} = |A||B|$.

习题 1. 利用分块矩阵计算 $\begin{pmatrix} 1 & -1 & 0 & 0 \\ 2 & 7 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 2 & -1 \end{pmatrix}\begin{pmatrix} -2 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 2 & -1 & 4 \end{pmatrix}$.

三. 分块矩阵的转置

若 $A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{pmatrix}$, 则容易验证 $A^T = \begin{pmatrix} A_{11}^T & A_{21}^T & \cdots & A_{p1}^T \\ A_{12}^T & A_{22}^T & \cdots & A_{p2}^T \\ \vdots & \vdots & \ddots & \vdots \\ A_{1q}^T & A_{2q}^T & \cdots & A_{pq}^T \end{pmatrix}$.

注意：不但每个小矩阵要转置，而且每个小矩阵都要调到新的位置去！

四. 分块对角矩阵

称 $\mathrm{diag}(A_1, A_2, \cdots, A_n) := \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix}$ 为分块对角矩阵. 容易验证，若每个分块均可逆，则分块对角矩阵可逆，且 $\mathrm{diag}(A_1, A_2, \cdots, A_n)^{-1} = \mathrm{diag}((A_1)^{-1}, (A_2)^{-1}, \cdots, (A_n)^{-1})$.

习题 2. 验证上述论断.

习题 3. 设 $A_i(i=1,2,3)$ 可逆. 求分块矩阵 $\begin{pmatrix} & & A_1 \\ & A_2 & \\ A_3 & & \end{pmatrix}$ 的逆.

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

第 5 节 矩阵的初等变换与矩阵的秩

本节是线性代数课程中非常关键的一节，概念和难点较多，学习时要归纳总结，理清思路. 特别要注意以下几点：

1. 消元法求解线性方程组 → 线性方程组的初等变换 → 矩阵的初等行 (列) 变换 → 化矩阵为 “标准形 ”(按标准程度分为 阶梯形矩阵，简化的阶梯形矩阵，标准形矩阵 等).

方法：要求多练习，熟练掌握通过初等变换化矩阵为标准形.

2. 初等变换的数学严密化: 用三类 **初等矩阵** 左 (右) 乘矩阵 = 相应行 (列) 初等变换.

3. 必须熟练掌握通过矩阵的初等变换求矩阵的逆以及求解线性方程组的方法.

4. 矩阵秩的概念 (初等变换下矩阵的不变量；描述一般线性方程组可解与否的量); 秩的基本性质；会熟练地用初等变换计算矩阵的秩.

一. 矩阵的初等变换

1. 线性方程组的初等变换

利用消元法求解如下一般线性方程组

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

本质上是对方程组施行若干次如下 **线性方程组的初等变换**, 以使变换后的方程具有简单的形式 (标准形), 从而可以轻易地写出解:

(1) 交换某两个方程在方程组中的先后位置;

(2) 在一个方程的两端同时乘某个非零常数;

(3) 将一个方程的两端同时乘某个非零常数然后将之加到另一个方程上 (注意原来的那个方程不要变).

上述线性方程组本质上由其 **增广矩阵**

$$
\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} & b_1\\
a_{21} & a_{22} & \cdots & a_{2n} & b_2\\
\vdots & \vdots & \ddots & \vdots & \vdots\\
a_{m1} & a_{m2} & \cdots & a_{mn} & b_m
\end{pmatrix}
$$

确定. 所以上述线性方程组的初等变换可推广为如下对矩阵的行初等变换.

2. **矩阵的初等变换**

如下三种对矩阵的行 (列) 施行的操作称为 **矩阵的行 (列) 初等变换**:

(1) 交换矩阵中的某两行 (列);

(2) 在某行 (列) 乘某个非零常数;

(3) 将某行 (列) 乘某个非零常数再加到另一行 (列).

矩阵的初等变换的目的是将之化为某种具有简单形式的 "标准形 ". 根据 "简单 " 程度的不同, 可列出如下三种标准形.

定义 1. *(行) 阶梯矩阵, 就是满足以下条件的矩阵: (1) 元素不全为零的行 (非零行) 的标号小于元素全为零的行; 也就是说, 元素全为零的行都排在后面;*

(2) 设矩阵有 r 个非零行, 第 i 个非零行的第一个非零元素所在的列号为 t_i, 则 $t_1<t_2<\cdots<t_r$.

例. $\begin{pmatrix} 1 & 3 & 2 & 0 & 0 & 1 \\ 0 & -2 & 4 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ 是 (行) 阶梯矩阵，但 $\begin{pmatrix} 2 & 0 & -3 \\ 0 & 2 & -1 \\ 1 & 0 & 0 \end{pmatrix}$ 不是.

定义 2. *简化的 (行) 阶梯矩阵：满足如下条件的 (行) 阶梯矩阵：每个非零行的第一个非零元素为 1, 且其所在列的其它元素为零.*

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

定义 3. *标准形矩阵：左上角为单位阵，其它位置元素均为零的非零矩阵.*

例. $\begin{pmatrix} E & O \\ O & O \end{pmatrix}$, $\begin{pmatrix} E & O \end{pmatrix}$, $\begin{pmatrix} E \\ O \end{pmatrix}$, E.

3. **标准形定理**

定理 1. *任何矩阵都可以通过* **单纯的行初等变换** *化为简化的行阶梯矩阵；任何矩阵都可经* **初等变换** *(既包括行初等变换也包括列初等变换) 化为标准形矩阵.*

注意：(1) 仅通过行初等变换不能把任一矩阵化为标准形矩阵;

(2) 化矩阵为阶梯形的方法不是唯一的，结果也是不唯一的 (下面会看到，这与矩阵乘法不满足交换律有关);

(3) 但是一个矩阵的标准形矩阵是唯一的. (这与矩阵的 “秩” 的概念有关，将在本节后面介绍.)

我们通过下面的例子来讲这个定理的证明. 这个证明其实是给出了一个算法，可以编作程序由计算机完成.

例. 将如下矩阵通过行初等变换写出其简化的行阶梯标准形，再求出其标准形.

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

定理的证明思路和例题的解

(1) 找出矩阵中第一个非零列，并通过一二类行初等变换使该列第一个元素为 1;

在上述矩阵 A 中第一个元素不全为零的列是第 2 列. 我们可通过将第 3 行与第 1 行交换，得到

$$A \to A_1 = \begin{pmatrix} 0 & -1 & -2 & 6 & -7 \\ 0 & 2 & 4 & 3 & 5 \\ 0 & 4 & 2 & -1 & 4 \end{pmatrix}.$$

注意在作初等变换时不要写等号，而应写为 “→”. 然后在 A_1 的第 1 行乘 -1, 得到

$$A_1 \to A_2 = \begin{pmatrix} 0 & 1 & 2 & -6 & 7 \\ 0 & 2 & 4 & 3 & 5 \\ 0 & 4 & 2 & -1 & 4 \end{pmatrix}.$$

(2) 通过第三类行变换使该列其它元素为零;

将 A_2 的第 1 行分别乘 -2, -4 加到第 2 、 3 行，就得到

$$A_2 \to A_3 = \begin{pmatrix} 0 & 1 & 2 & -6 & 7 \\ 0 & 0 & 0 & 15 & -9 \\ 0 & 0 & -6 & 23 & -24 \end{pmatrix}.$$

(3) 考察除第 1 行外其它行组成的矩阵并重复上述过程；由此通过归纳法即可证明定理第一部分.

在 A_3 中，除去第 1 行后的矩阵为 $B_1 = \begin{pmatrix} 0 & 0 & 0 & 15 & -9 \\ 0 & 0 & -6 & 23 & -24 \end{pmatrix}$. 它的第一个非零列为第 3 列. 所以可在 A_3 中交换第 2 、 3 行，得到 $A_3 \to A_4 = \begin{pmatrix} 0 & 1 & 2 & -6 & 7 \\ 0 & 0 & -6 & 23 & -24 \\ 0 & 0 & 0 & 15 & -9 \end{pmatrix}$. 在 A_4 的第 2 行乘 $-\frac{1}{6}$, 就得到 $A_4 \to A_5 =$

$\begin{pmatrix} 0 & 1 & 2 & -6 & 7 \\ 0 & 0 & 1 & -23/6 & 4 \\ 0 & 0 & 0 & 15 & -9 \end{pmatrix}$. 这样一来，再用第 2 行乘 -2 到第 1 行，得到

$$A_5 \to A_6 = \begin{pmatrix} 0 & 1 & 0 & 5/3 & -1 \\ 0 & 0 & 1 & -23/6 & 4 \\ 0 & 0 & 0 & 15 & -9 \end{pmatrix}.$$

现在，前两行已经完成，余下的只有第 3 行．将之乘 1/15, 就得到

$$A_6 \to A_7 = \begin{pmatrix} 0 & 1 & 0 & 5/3 & -1 \\ 0 & 0 & 1 & -23/6 & 4 \\ 0 & 0 & 0 & 1 & -3/5 \end{pmatrix}.$$

在 A_7 中，第 3 行分别乘 23/6 、 $-5/3$ 到第 2 行、第 1 行，就得到

$$A_7 \to A_8 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -43/18 \\ 0 & 0 & 0 & 1 & -3/5 \end{pmatrix}.$$

这就得到了简化的行阶梯标准形.

(4) 在得到简化的行阶梯标准形后，通过第三类列初等变换将非零行除第一个非零元素 1 外其它均化为零；

A_8 中第 1 行已满足要求．将第 3 列乘 43/18 加到第 5 列，第 4 列乘 3/5 加到第 5 列，即得到

$$A_8 \to A_9 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

(5) 通过第一类列初等变换化为标准形.

在 A_9 中依次将第 1 列与第 2 、 3 、 4 交换，就得到标准形

$$A_9 \to A_{10} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

习题 1. 将下列矩阵通过行初等变换化为简化的行阶梯矩阵，再用初等变换化为标准形.

$$(1)\begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & 6 \\ 1 & 4 & 0 \end{pmatrix};\quad (2)\begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 3 & 0 \end{pmatrix}.$$

二. 初等矩阵

通过下面初等矩阵的概念，可以将上述矩阵初等变换的 **描述性定义** 转化为 **严格的数学语言**, 用矩阵的乘法运算来精确表达. 这相当于一种 “数学建模 ”, 其优点之一，就在于可以把上面的 “→” 写为精确的等号，为理论分析提供方便，从而可以发现一些新的性质、应用等.

1. **初等矩阵的定义**

单位矩阵经过一次初等变换所得到的矩阵称为初等矩阵.

(1) 第一类初等矩阵 $R_{i\leftrightarrow j}=C_{i\leftrightarrow j}$: 交换第 i,j 行 (列) 所得;

(2) 第二类初等矩阵 $R_{(k)i}=C_{(k)i}$: 第 i 行 (列) 乘以非零常数 k 所得;

(3) 第三类初等矩阵 $R_{i+(k)j}=C_{j+(k)i}$: 第 j 行 (i 列) 乘以非零常数 k 加到第 i 行 (j 列) 所得.

注意，我们用 R 表示由行初等变换所得的初等矩阵，C 表示由列初等变换所得的初等矩阵. 特别地，由定义看出，每一个初等矩阵既可以看成单位矩阵由行初等变换得来，也可以由同类列初等变换得来.

由行列式的性质，容易看出， **初等矩阵都是可逆矩阵**. 更进一步，可以验证, 初等矩阵的逆矩阵也都是初等矩阵. 例如, 我们有 $R_{i\leftrightarrow j}^{-1}=R_{j\leftrightarrow i}, R_{(k)i}^{-1}=R_{(\frac{1}{k})i}, R_{i+(k)j}^{-1}=R_{i-(k)j}$.

习题 2. 对三阶矩阵，写出 $C_{1\leftrightarrow 2}, R_{2+(k)3}, R_{2+(k)3}^{-1}$, 和 $R_{(2)2}$.

2. **初等矩阵与初等变换**

定理 2. *用一个 m 阶某类初等矩阵左乘一个 $m\times n$ 矩阵，相当于对该矩阵施行一次相应类的行初等变换；用一个 n 阶某类初等矩阵右乘一个 $m\times n$ 矩阵，相当于对该矩阵施行一次相应类的列初等变换.*

证明. 证明一般情形需要利用分块矩阵的乘法. 我们这里仅对 2 阶矩阵

的情形作一验证．证明的本质是一样的．例如，

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}=\begin{pmatrix} c & d \\ a & b \end{pmatrix},$$
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}=\begin{pmatrix} b & a \\ d & c \end{pmatrix};$$
$$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}=\begin{pmatrix} a & b \\ ka+c & kb+d \end{pmatrix},$$
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}=\begin{pmatrix} a+kb & b \\ c+kd & d \end{pmatrix}.$$

3. **标准形定理**

将定理 1 的结论改用初等矩阵来陈述，就是如下重要结论.

定理 3. *对任何 $m\times n$ 矩阵 A, 存在 m 阶初等矩阵 $R_1,\cdots,R_s$, 使得*

$$R_s\cdots R_1A$$

为简化的 (行) 阶梯形矩阵；存在 n 阶初等矩阵 $C_1,\cdots,C_t$, 使得

$$R_s\cdots R_1AC_1\cdots C_t=\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}.$$

其中 $r=r(A)$ 叫做矩阵 A 的 **秩**, *它是与上述初等矩阵选取无关的只依赖于矩阵本身的一个常数 (这一点将在本节后面证明).*

推论 1. 对任何 $m\times n$ 矩阵 A, 存在 m 阶可逆矩阵 P 与 n 阶可逆矩阵 Q 使得

$$PAQ=\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}.$$

证明：取 $P=R_s\cdots R_1$, $Q=C_1\cdots C_t$ 即可.

推论 2. 对任何 n 阶矩阵 A, 它可逆的充分必要条件为它的标准形是单位阵.

证明．若一个矩阵的标准形式单位阵，由推论 1, $PAQ=E$ 得到 $A=P^{-1}Q^{-1}$. 所以 A 可逆．反之，若 A 可逆，那么 PAQ 可逆，即标准形 (它现在是个方阵) 可逆，则必有 $r=n$, 即标准形是个单位阵．证毕.

推论 3. 矩阵 A 可逆的充要条件为 $A=P_1\cdots P_k$, 其中 P_i 为初等矩阵.

证明. 由推论 2 和推论 1 的证明，此时

$$A=P^{-1}Q^{-1}=R_1^{-1}R_2^{-1}\cdots R_s^{-1}C_t^{-1}\cdots C_2^{-1}C_1^{-1}.$$

由于初等矩阵的逆矩阵也是初等矩阵，所以此推论 3 得证.

推论 4. 任一可逆矩阵 A 可通过单纯的行初等变换化为单位阵，即：

$$A=P_k\cdots P_1E \qquad (\Leftrightarrow (P_1)^{-1}\cdots(P_k)^{-1}A=E)\ (\Leftrightarrow A^{-1}=(P_1)^{-1}\cdots(P_k)^{-1}E).$$

证明. 这只是推论 3 的另一种表述而已.

注意：由推论 4 后两式知可利用初等变换求矩阵逆 (当行初等变换把 A 化为单位阵时，相应的变换就把单位阵化为 A^{-1}). 这是矩阵求逆的一种高效方法.

三. 初等变换的应用：矩阵求逆及解线性方程组

1. **矩阵求逆算法**. 对 $(A\ \ E)$ 作行初等变换，将 A 化为标准形 (单位阵) 后，则 E 恰变换为 A^{-1}. (原理参见推论 4 中后两个表达式.) 若 A 不能化为单位阵，则 A 不可逆. 此算法比伴随矩阵法的优点：运算量小，效率高.

例 1. 求 $A=\begin{pmatrix}1&2&3\\2&1&2\\1&3&4\end{pmatrix}$ 的逆.

解.

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

所以 $A^{-1}=\begin{pmatrix}-2&1&1\\-6&1&4\\5&-1&-3\end{pmatrix}$.

2. **求解线性方程组 $AX=B$ 算法**. 对 $(A\ \ B)$ 作行初等变换将 A 化为单位阵, B 则变换为 $A^{-1}B$, 就是所要求的解.

例 2. 求解矩阵方程

$$\begin{pmatrix}1&1&1&\cdots&1&1\\&1&1&\cdots&1&1\\&&1&\cdots&1&1\\\cdots&&\cdots&&\cdots&\\&&&&1&1\\&&&&&1\end{pmatrix}X=\begin{pmatrix}2&1&&&&\\1&2&1&&&\\0&1&2&&&\\\cdots&&\cdots&&\cdots&\\&&&&2&1\\&&&&1&2\end{pmatrix}.$$

解. 依次将矩阵

$$\left(\begin{array}{cccccc|cccccc}1&1&1&\cdots&1&1&2&1&&&&\\&1&1&\cdots&1&1&1&2&1&&&\\&&1&\cdots&1&1&0&1&2&&&\\\cdots&&\cdots&&\cdots&&\cdots&&\cdots&&\cdots&\\&&&&1&1&&&&&2&1\\&&&&&1&&&&&1&2\end{array}\right)$$

第 k $(k=2,3,\cdots,n)$ 行乘以 -1 加到第 $k-1$ 行, 左边就得到单位阵, 而右边就是 $\begin{pmatrix}1&-1&-1&&\\1&1&-1&&\\0&1&1&&\\\cdots&&\cdots&\cdots&\\&&&1&-1\\&&&1&2\end{pmatrix}$, 这就是解 X.

上述矩阵求逆和线性方程组求解的初等变换法是线性代数非常重要的内容, 大家在理解原理的基础上要多做一些练习. 这里技巧性不强, 关键是要熟练.

习题 3. 设 $P=\begin{pmatrix}1&0&0\\2&-1&0\\2&1&1\end{pmatrix}, B=\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}$. 矩阵 X 满足方程 $XP=PB$. 求 X^5.

习题 4. 已知 $A=\begin{pmatrix}1&0&0\\1&1&0\\1&1&1\end{pmatrix}, B=\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}$, 且 X 满足 $AXA+BXB=AXB+BXA+E$. 求 X.

四. 矩阵的相抵及其不变量：矩阵的秩

1. 矩阵秩的定义及其性质

如果矩阵 A 经过有限次初等变换变为矩阵 B (或存在可逆阵 P,Q 使得 $PAQ=B$), 就称 A 与 B 相抵，记作 $A\sim B$. 特别地，一个矩阵和它的通过初等变换得到的标准形是相抵的.

习题 5. 证明：两个矩阵相抵，当且仅当它们可以通过初等变换化为同一个标准形.

那么，我们有什么办法可以判断两个矩阵相抵呢？将它们都通过初等变换化为标准形是一个办法. 但是还有更好的办法吗？如果这个问题有个更好的求解办法，那么我们就可以不必作详细的初等变换，也能一下子写出一个矩阵的标准形了. 为此我们引入矩阵的秩这个重要概念.

定义 4. *在矩阵 $A=(a_{ij})_{m\times n}$ 中，任取 k 行 k 列 $(1\le k\le \min\{m,n\})$, 由这 k 行 k 列交叉点上 k^2 个元素按原来位置构成的 k 阶行列式称为 A 的一个 k 阶子式. 若 A 中有一个 r 阶子式不为零，而所有的 $r+1$ 阶子式 (如果存在的话) 都为零，则称 r 为 A 的* **秩** *(rank), 记作 $r(A)$.*

即： **矩阵 A 的秩就是其非零子式的最高阶数**.

设 A 为 $m\times n$ 矩阵. 若 $r(A)=m$ $(r(A)=n)$, 称 A 为行 (列) 满秩阵. 如果 A 为方阵，则 A 可逆 (非奇异) 与 A 满秩是等价概念.

根据定义，不难证明矩阵的秩具有如下基本性质.

(a) 一个矩阵的秩是唯一的.

(b) A 为 $m\times n$ 矩阵，则 $r(A)\le \min\{m,n\}$, 且 $r(A)=0 \Leftrightarrow A=O$.

(c) 若矩阵 A 中有一个 r 阶子式不为零，则 $r(A) \geq r$; 若所有 r 阶子式都为零，则 $r(A) < r$.

(d) 若 A_1 为 A 的子矩阵，则 $r(A_1) \leq r(A)$.

(e) $r(A) = r(A^T)$.

(f) 阶梯形矩阵的秩等于其非零行的个数.

习题 6. 证明上述性质.

秩的重要性质之一就是在初等变换下不变，即我们有如下重要定理.

定理 4. *初等变换不改变矩阵的秩.*

证明. 第一步. 简化问题的提法到最本质的地方.

(1) 只要证明行初等变换不改变矩阵的秩 (利用性质 (e));

(2) 只需要证明矩阵的秩在行初等变换下不增：设 A 经行初等变换化为 B, 证明 $r(B) \leq r(A)$ 即可.

事实上，由于初等变换的可逆性，A 也可由同类初等变换从 B 得到，于是 $r(A) \leq r(B)$.

两者对比，就有 $r(A) = r(B)$.

要证明 $r(B) \leq r = r(A)$, **只需说明**，B **的所有** s **阶子式** ($s > r$) **均为零**.

第二步. 实质性证明. 显然第二类行初等变换后矩阵的秩不增加.

课本上对第三类行初等变换给出了证明 (p.64). 我们在这里对第一类行初等变换的情形给出证明.

(1) 设交换 A 的 i, j 行得到 B. 考察 B 的任一 s 阶子式 D_s.

若 D_s 不包含 i, j 行，则它也是 A 的 $s(> r)$ 阶子式，由 $r = r(A)$ 的定义，$D_s = 0$.

(2) 设若 D_s 包含 i, j 行，则它由 A 的某个 s 阶子式交换两行得到，由行列式性质，$D_s = 0$.

(3) 若 D_s 只包含第 i 行而不包含第 j 行，则此 D_s 可由 A 的某个包含 j 行但不包含 i 行的 s 阶子式通过交换若干行得到. 由行列式性质知 $D_s = 0$.

(4) D_s 只包含第 j 行而不包含第 i 行，证明与 (3) 类似. 证毕.

推论. 设 A 为 $m \times n$ 矩阵，P 为 m 阶可逆矩阵，Q 为 n 阶可逆矩阵，则

$$r(A) = r(PA) = r(AQ) = r(PAQ).$$

即 **矩阵的秩是其相抵不变量**: 相抵的矩阵有相同的秩，有相同秩的同类型矩阵必相抵.

2. **利用初等变换求矩阵的秩**

定理 4 及其推论告诉我们，要判定同类型的两个矩阵是否相抵，只要计算它们的秩. 而在计算矩阵秩时，我们既可以使用初等行变换，也可以使用初等列变换，这就使得计算可以更快捷. (回忆在矩阵求逆、线性方程组求解时只能用初等行变换.)

例 3. 设 $A=\begin{pmatrix}1&1&0\\1&0&1\\0&1&1\end{pmatrix}$, $B=\begin{pmatrix}a&1&1\\2&1&a\\1&1&a\end{pmatrix}$. 若 AB 的秩为 2, 求 a.

解. $|A|\neq 0$. 从而 $r(B)=2$.

又 $B\sim\begin{pmatrix}1&1&a\\0&1&a\\0&0&1-a\end{pmatrix}\Rightarrow a=1$.

例 4. 求矩阵 $A=\begin{pmatrix}1&a&a&a\\a&1&a&a\\a&a&1&a\\a&a&a&1\end{pmatrix}$ 的秩.

解.

$$
\begin{aligned}
A \quad \to \quad & \begin{pmatrix} 1 & a & a & a \\ 0 & 1-a^2 & a-a^2 & a-a^2 \\ 0 & a-a^2 & 1-a^2 & a-a^2 \\ 0 & a-a^2 & a-a^2 & 1-a^2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-a^2 & a-a^2 & a-a^2 \\ 0 & a-a^2 & 1-a^2 & a-a^2 \\ 0 & a-a^2 & a-a^2 & 1-a^2 \end{pmatrix} \\
\to \quad & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-a & 0 & a-a^2 \\ 0 & a-1 & 1-a & a-a^2 \\ 0 & 0 & a-1 & 1-a^2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-a & 0 & a-a^2 \\ 0 & 0 & 1-a & 2(a-a^2) \\ 0 & 0 & a-1 & 1-a^2 \end{pmatrix} \\
\to \quad & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-a & 0 & a-a^2 \\ 0 & 0 & 1-a & 2(a-a^2) \\ 0 & 0 & 0 & 1+2a-3a^2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-a & 0 & 0 \\ 0 & 0 & 1-a & 0 \\ 0 & 0 & 0 & 1+2a-3a^2 \end{pmatrix}.
\end{aligned}
$$

注意到 $1+2a-3a^2=(1-a)(3a+1)$, 所以当 $a \neq 1$ 且 $a \neq -\frac{1}{3}$ 时 $r(A)=4$, 当 $a=-\frac{1}{3}$ 时 $r(A)=3$, 当 $a=1$ 时 $r(A)=1$.

例 5. 设 A 为 $m \times n$ 矩阵，证明：若 $r(A)=1$, 则存在 m 维非零列向量 α 及 n 维非零列向量 β 使得 $A=\alpha\beta^T$.

证明. 注意到特殊情形 $(1,0,\cdots,0)^T(1,0,\cdots,0)=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$, 而由于 $r(A)=1$, 这个矩阵就是 A 的标准形，故有可逆矩阵 P,Q 使得它等于 PAQ. 从而 $A=P^{-1}(1,0,\cdots,0)^T(1,0,\cdots,0)Q^{-1}$. 令 $\alpha=P^{-1}(1,0,\cdots,0)^T$, $\beta=(Q^{-1})^T(1,0,\cdots,0)^T$ 即可. 证毕.

这个证明体现了代数学的一个重要思想方法, 即对一类对象 (如矩阵) 按某种标准分类 (例如相抵的矩阵作为一类), 每一类找出一个比较简单的代表元 (这里就是初等变换下的标准形), 然后要证明某性质时，先对简单的代表

元作出证明 (这题中就是得出 $(1,0,\cdots,0)^T(1,0,\cdots,0)=\begin{pmatrix}1&0&\cdots&0\\0&0&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&0\end{pmatrix}$),
然后再推广到一般情形. 下一节将用这个思想研究一般线性方程组的可解性问题.

习题 7. 六阶方阵按矩阵相抵等价分类，一共可分几类？

习题 8. 计算下列矩阵的秩.

$$(1)\begin{pmatrix}0&1&1&-1&2\\0&2&-2&-2&0\\0&-1&-1&1&1\\1&1&0&1&-1\end{pmatrix};\ (2)\begin{pmatrix}14&12&6&8&2\\6&104&21&9&17\\7&6&3&4&1\\35&30&15&20&5\end{pmatrix}.$$

习题 9. 证明： (1) $r\begin{pmatrix}A&O\\O&B\end{pmatrix}=r(A)+r(B)$; (2) $r\begin{pmatrix}A&C\\O&B\end{pmatrix}\geq r(A)+r(B)$.

第 6 节　线性方程组的可解性判定定理及求解

作为第 5 节理论的应用，我们用矩阵秩的概念，给出判定一个一般线性方程组

$$A\mathbf{x}=\mathbf{b} \tag{4}$$

是否有解的判定定理，并利用初等变换写出有解情形下所有的解. 这里 A 是 $m\times n$ 的系数矩阵， $\mathbf{b}$ 为 m 维列向量. $\mathbf{x}$ 是由未知数构成的 n 维列向量. 分块矩阵 $\tilde{A}=(A,\mathbf{b})$ 称为方程组 (4) 的 **增广矩阵**.

定理 1. *(a)* (4) *有唯一解当且仅当* $r(A)=r(\tilde{A})=n$ *(其中* n *为未知数个数);*

(b) (4) *有无穷多解当且仅当* $r(A)=r(\tilde{A})<n$*;*

(c) (4) *无解当且仅当* $r(A)\neq r(\tilde{A})$.

证明. 由于秩在初等行变换下不变，而方程组 (4) 在初等行变换下解并不改变，故只需对 $\tilde{A}$ 的简化行阶梯标准形 C 进行讨论. 不失一般性，设 C 具

有如下形式：

$$\begin{pmatrix} 1 & c_{12} & \cdots & c_{1,k_2-1} & 0 & * & \cdots & * & 0 & * & \cdots & * & 0 & c_{1,k_r+1} & \cdots & c_{1n} & d_1 \\ 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & c_{2,k_r+1} & \cdots & c_{2n} & d_2 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & c_{3,k_r+1} & \cdots & c_{3n} & d_3 \\ \vdots & \vdots & & \vdots & \vdots & & & & \vdots & & & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & c_{r,k_r+1} & \cdots & c_{rn} & d_r \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & d_{r+1} \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}_{m\times(n+1)},$$

(一般情形可通过交换前 n 列化为这种情形，但这意味着相应的 $x_1,x_2,\cdots,x_n$ 也要作交换，) 其中 $*$ 表示的元素可能为零，也可能不为零．d_{r+1} 可能为零，也可能为 1.

讨论以 C 为增广矩阵的线性方程组．它的解和 (4) 的解是完全一样的．我们有如下几种情形：

(c) $d_{r+1}\neq 0$. 此时成立 $r(A)<r(\tilde{A})$. 由于会出现矛盾方程 $0=d_{r+1}$, 所以 (4) 无解.

(a) $d_{r+1}=0$ 且 $r=n$. 此时 $r(C)=r=n$, C 中由前 n 列构成的矩阵秩为 n, 即 $r(A)=n$. 于是有 $r(A)=r(C)=r(\tilde{A})=n$. 在这种情况下，我们有

$$C=\begin{pmatrix} 1 & & & & d_1 \\ & 1 & & & d_2 \\ & & \ddots & & \vdots \\ & & & 1 & d_n \\ & & & & \end{pmatrix},$$

其中空白处均表示矩阵相应元素为零. 于是我们可以写出 (4) 唯一的一组解：

$$x_1=d_1,\ x_2=d_2,\ \cdots,x_n=d_n.$$

(b) $d_{r+1}=0$ 且 $r<n$. 这也就是 $r(A)=r(\tilde{A})<n$ 的情形．不失一般性，设

C 具有如下形式：

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & c_{1,r+1} & \cdots & c_{1n} & d_1 \\ 0 & 1 & \cdots & 0 & c_{2,r+1} & \cdots & c_{2n} & d_2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_{r,r+1} & \cdots & c_{rn} & d_r \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

这就是说 (4) 可以等价地化简为

$$\begin{cases} x_1 + c_{1,r+1}x_{r+1} + \cdots + c_{1n}x_n = d_1, \\ x_2 + c_{2,r+1}x_{r+1} + \cdots + c_{2n}x_n = d_2, \\ \qquad \cdots\cdots \\ x_r + c_{r,r+1}x_{r+1} + \cdots + c_{rn}x_n = d_r. \end{cases}$$

所以如果以 $x_{r+1}, \cdots, x_n$ 作为 **自由未知量** (即 $x_{r+1}, \cdots, x_n$ 可以任意取值), 我们可以写出 (4) 的无穷多个解：

$$\begin{cases} x_1 = -c_{1,r+1}x_{r+1} - \cdots - c_{1n}x_n + d_1, \\ x_2 = -c_{2,r+1}x_{r+1} - \cdots - c_{2n}x_n + d_2, \\ \qquad \cdots\cdots \\ x_r = -c_{r,r+1}x_{r+1} - \cdots - c_{rn}x_n + d_r. \end{cases} \tag{5}$$

我们看到上述 (a)-(c) 是所有可能的三种情形. 所以我们证明的结论不但是充分的，而且也是必要的. 证毕.

上述定理回答了我们在第一章提出的问题 (a), 以及 (b) 的前一部分. 那里问题 (b) 的后一部分说，如何在解不唯一的情形给出解集的结构. 上面 (5) 已写出了所有的解，在下一章，我们将利用向量的理论讨论这个集合的大小等性质. 有解情形下问题 (d) 中高效求解的方法就是我们已学的初等变换法. 至于无解情形如何定义 “广义解 ”, 我们将在后面继续探讨.

例 1. 讨论 a, b 取何值时，线性方程组

$$\begin{cases} x_1+x_2-x_3=1, \\ 2x_1+(a+2)x_2-(b+2)x_3=3, \\ -3ax_2+(a+2b)x_3=-3 \end{cases}$$

有唯一解、无穷多解、以及无解．有解时写出所有的解．

解．将增广矩阵用初等行变换化简化到可以看出其秩的情形，然后分情况讨论．

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ 2 & a+2 & -(b+2) & 3 \\ 0 & -3a & a+2b & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & a & -b & 1 \\ 0 & -3a & a+2b & -3 \end{pmatrix}$$

$$\to \quad \tilde{A}_1 = \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & a & -b & 1 \\ 0 & 0 & a-b & 0 \end{pmatrix}.$$

(a) $a=b=0$ 时， $r(A)=1<r(\tilde{A})=2$, 无解．事实上，此时

$$\tilde{A}_1 \to \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(b) $a=b\neq 0$ 时， $r(A)=2=r(\tilde{A})<3$, 有无穷多解．事实上，此时

$$\tilde{A}_1 \to \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & a & -a & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -1 & 1/a \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1-\frac{1}{a} \\ 0 & 1 & -1 & \frac{1}{a} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

以 x_3 为自由未知量，那么解可写为 $x_1=1-\frac{1}{a}, x_2=x_3+\frac{1}{a}$.

(c) $a\neq b$ 且 $a\neq 0$ 时， $r(A)=r(\tilde{A})=3$, 从而有唯一解．进一步对 $\tilde{A}_1$ 作初等行变换得到

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -b/a & 1/a \\ 0 & 0 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1/a \\ 0 & 0 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1-1/a \\ 0 & 1 & 0 & 1/a \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

从而唯一解为 $x_1=1-\frac{1}{a}, x_2=\frac{1}{a}, x_3=0$.

(d) $a=0$ 且 $b\neq 0$. 对 $\tilde{A}_1$ 作初等行变换得到

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 0 & -b & 1 \\ 0 & 0 & -b & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 0 & 1 & -\frac{1}{b} \\ 0 & 0 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & -\frac{1}{b} \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

可以看出 $r(A)=2<r(\tilde{A})=3$. 所以无解.

例 2. 设有如下两个线性方程组：

$$(I)\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m; \end{cases} \qquad (II)\begin{cases} a_{11}x_1+a_{21}x_2+\cdots+a_{m1}x_m=0, \\ a_{12}x_1+a_{22}x_2+\cdots+a_{m2}x_m=0, \\ \cdots\cdots\cdots \\ a_{1n}x_1+a_{2n}x_2+\cdots+a_{mn}x_m=0, \\ b_1x_1+b_2x_2+\cdots+b_mx_m=1. \end{cases}$$

证明： *(I) 有解当且仅当 (II) 无解.*

证明. 记 $A=[a_{ij}]_{m\times n}$, $\mathbf{b}=(b_1,b_2,\cdots,b_m)^T$. 那么 (II) 的系数矩阵就是 $\begin{pmatrix} A^T \\ \mathbf{b}^T \end{pmatrix}$, 其增广矩阵就是分块矩阵 $\begin{pmatrix} A^T & \mathbf{0} \\ \mathbf{b}^T & 1 \end{pmatrix}$.

首先证明必要性. 由条件，成立 $r(A)=r(A,\mathbf{b})=r$. 从而由秩的定义，$r(A^T)=r$, 也就是说， A^T 中至少有一个 r 阶子式不为零. 设此子式取自第 $i_1,\cdots.i_r$ 行，$j_1,\cdots,j_r$ 列. 将此子式最后补充一行 $(b_{i_1},\cdots,b_{i_r})$、补充最后一列为 $(0,\cdots,0,1)^T$, 我们就得到 $\begin{pmatrix} A^T & \mathbf{0} \\ \mathbf{b}^T & 1 \end{pmatrix}$ 的一个 $r+1$ 阶子式, 将之按最后一列展开, 就知道该子式非零. 于是我们得到 $r\begin{pmatrix} A^T & \mathbf{0} \\ \mathbf{b}^T & 1 \end{pmatrix}\geq r+1>r=r\begin{pmatrix} A^T \\ \mathbf{b}^T \end{pmatrix}$. 这就表明 (II) 无解.

再证明充分性. 由条件和矩阵秩的性质，我们有不等式

$$r=r(A)=r(A^T)\leq r(A,\mathbf{b})=r\begin{pmatrix} A^T \\ \mathbf{b}^T \end{pmatrix}<r\begin{pmatrix} A^T & \mathbf{0} \\ \mathbf{b}^T & 1 \end{pmatrix}.$$

如果有 $r\begin{pmatrix} A^T & \mathbf{0} \\ \mathbf{b}^T & 1 \end{pmatrix} = r+1$, 那么由于秩只能是整数，就必有 $r(A,\mathbf{b}) = r$, 也即 (I) 有解.

我们已经在必要性部分证明了 $r\begin{pmatrix} A^T & \mathbf{0} \\ \mathbf{b}^T & 1 \end{pmatrix} \geq r+1$. 故只需再证 $r\begin{pmatrix} A^T & \mathbf{0} \\ \mathbf{b}^T & 1 \end{pmatrix} \leq r+1$.

事实上, 利用矩阵秩在初等列变化下不变, 我们有 $r\begin{pmatrix} A^T & \mathbf{0} \\ \mathbf{b}^T & 1 \end{pmatrix} = r\begin{pmatrix} A^T & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} = r\begin{pmatrix} E_r & O & \mathbf{0} \\ O & O & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1 \end{pmatrix} = r+1$. 其中 $\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$ 为 A^T 的的标准形.

我们在必要性中用了秩的定义证明，在充分性中用了标准形和初等变换来计算秩，目的是帮助大家复习秩.

习题 1. 求解如下线性方程组：

$$(1)\begin{cases} 2x_1 - x_2 + 3x_3 = 1, \\ 4x_1 - 2x_2 + 5x_3 = 4, \\ 2x_1 - x_2 + 4x_3 = -1, \\ 6x_1 - 3x_2 + 5x_3 = 11; \end{cases} \qquad (2)\begin{cases} 2x_1 + 2x_2 + 3x_3 = 1, \\ x_1 - x_2 = 2, \\ -x_1 + 2x_2 + x_3 = -2; \end{cases}$$

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

习题 2. 证明：线性方程组 $\begin{cases} x_1 - x_2 = a_1, \\ x_2 - x_3 = a_2, \\ \cdots, \\ x_{n-1} - x_n = a_{n-1}, \\ x_n - x_1 = a_n \end{cases}$ 有解当且仅当 $\sum_{i=1}^n a_i = 0$.

复习题

1. 若 A 是 n 阶方阵，我们定义 A 的迹为：

$$\operatorname{tr}(A)=\sum_{i=1}^{n}A_{ii},$$

即 A 的所有对角元之和．现 A,B 设都是 n 阶方阵，证明： $\operatorname{tr}(AB)=\operatorname{tr}(BA)$.

2. 解下列方程.

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

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

3. 求下列矩阵的逆.

(1) $\begin{pmatrix}0&0&1&-1\\0&3&1&4\\2&7&6&-1\\1&2&2&-1\end{pmatrix}$; (2) $\begin{pmatrix}2&1&&&\\&2&1&&\\&&2&1&\\&&&2&1\\&&&&2\end{pmatrix}$.

4. 设 A 是 n 阶方阵且 $A^k=O$. 证明：$(E-A)^{-1}=E+A+A^2+\cdots+A^{k-1}$.

5. 设 A 是 n 阶方阵．证明：存在 n 阶非零方阵 B 使得 $AB=O$ 的充分必要条件为 $|A|=0$.

6. 设 $A=\frac{1}{2}(B+E)$. 证明： $A^2=A$ 当且仅当 $B^2=E$.

7. 设 A 是 4 阶实方阵，记 $\mathscr{A}_{ij}$ 为 $|A|$ 中 A_{ij} 对应的代数余子式．若成立 $\mathscr{A}_{ij}=-A_{ij}$ 且 $A_{11}\neq 0$, 求 $|A|$.

8. 满足 $AA^T=E$ 的 n 阶方阵叫做正交阵．证明：(1) $|A|=\pm 1$; (2) 设 A,B 都是正交阵且 $|A|=1,|B|=-1$, 则 $A+B$ 不可逆.

9. 讨论如下线性方程组的解.

(1) $\begin{cases}\lambda x_1+x_2+x_3+x_4=1,\\ x_1+\lambda x_2+x_3+x_4=\lambda,\\ x_1+x_2+\lambda x_3+x_4=\lambda^2,\\ x_1+x_2+x_3+\lambda x_4=\lambda^3;\end{cases}$ (2) $\begin{cases}ax_1+x_2+x_3=4,\\ x_1+bx_2+x_3=3,\\ x_1+2bx_2+x_3=4.\end{cases}$

10. 设 A 是 n 阶方阵，A^* 是 A 的伴随矩阵．证明：$(A^*)^*=|A|^{n-2}A$.

11. 设 A,B 分别是 $n\times m,m\times n$ 矩阵．证明：$\begin{vmatrix}E_m & B\\ A & E_n\end{vmatrix}=|E_n-AB|=|E_m-BA|$.

12. 设 A,B 如上题．证明：$|\lambda E_n-AB|=\lambda^{n-m}|E_m-BA|$.

13. 考虑平面上三条不同的直线 $L_1:ax+by+c=0;L_2:bx+cy+a=0;L_3:cx+ay+b=0$. 证明：它们相交于一点的充分必要条件为 $a+b+c=0$.

14. 证明：通过平面内不共线的三点 $(x_1,y_1),(x_2,y_2),(x_3,y_3)$ 的圆的方程为

$$\begin{vmatrix}x^2+y^2 & x & y & 1\\ x_1^2+y_1^2 & x_1 & y_1 & 1\\ x_2^2+y_2^2 & x_2 & y_2 & 1\\ x_3^2+y_3^2 & x_3 & y_3 & 1\end{vmatrix}=0.$$

由此证明：通过三个有理点 (即横坐标与竖坐标都是有理数的点) 的圆的圆心也是有理点.

第 4 章　　线 性 空 间

通过引入矩阵及其乘法，以及初等变换，我们已经解决了判定给定线性代数方程组

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

受否有解的问题，并且在有解情形下可以给出所有的解组成的集合 $\mathscr{S}$. 但是，有如下问题还没有得到解决：

(A) 对 $\mathscr{S}$ 的细致了解；

(B) 对使得方程组 (1) 有解的非齐次项 $\mathbf{b}$ 构成的集合 $\mathscr{B}$ 的细致了解；

(C) 若 (1) 无解，如何定义某种 "广义解"?

我们来解释一下为什么要提出这些问题. 首先，我们不是都已经清清楚楚地把集合 $\mathscr{S}$ 写出来了吗？ (见第 3 章第 (5) 式.) 难道这还叫了解得不细致？

是的，从认识论角度讲，给出了你所关心的对象并不代表你对它就有了深刻的了解. 举个例子，作为教师自然非常关心学生考试成绩. 但是我们往往最关心的不是一个个具体分数，而是班级分数的平均分、最高分、最低分等一些数字指标. 如果知道了这些指标，再和别的班的考试成绩的相应指标对比，我们就会一下子看出到底哪个班相对更好一点.

同样的，对于集合 $\mathscr{S}$, 我们更希望了解它所具有的一些重要性质和数字指标. 比如，它到底有多 "大" (即其 "秩" 或 "维数" 的概念), 它是否可由某些简单的量完全确定 (即 "极大线性无关组" 或 "基" 的概念). 这样一来，我们才算真正搞清楚了这个集合. 这就是我们有问题 (A) 的原因.

我们已学的定理只是给出了对给定的 $\mathbf{b}$, 判断方程组 (1) 有解的方法，并没有指出使得它有解的集合 $\mathscr{B}$ 是什么. 而这对解决一些理论和实际问题 (比如，问题 (C)) 是很重要的. 例如，当 (1) 无解时，用最小二乘法给出这个无解方程组广义解的思想就是在 $\mathscr{B}$ 中找到一个元素 $\mathbf{b}_1$, 使得 $\mathbf{b}_1$ 与 $\mathbf{b}$ 尽可能 "接近". 如果不对 $\mathscr{B}$ 有个好的了解，又如何找到 $\mathbf{b}_1$? 这就是我们需要解决问题 (B) 的原因. 至于 (C) 的重要性，我们已经解释过，就不再赘述了.

那么，线性代数又是如何来解决这些问题的呢？

在代数学中，研究一个集合、一类对象的方法之一，是找出若干个代表元，然后通过这若干个代表元把集合中所有的元素表示出来．通过少数代表元就把所有元素都表示出来的关键，就是利用元素之间的运算，例如数乘和加法．所以代数学研究那些定义了元素之间的运算的集合．线性空间就是描述诸如 $\mathscr{S}, \mathscr{B}$ 这类对象的，具有加法和数乘运算的由向量组成的集合．这也是本章的主要内容．

第 1 节　n 维向量及线性空间的概念

一．n 维向量

由 n 个实数 (或复数) $a_1, a_2, \cdots, a_n$ 组成的有序数组 $\mathbf{a} = (a_1, a_2, \cdots, a_n)^T$ 称为 n 维实向量 (或复向量). 当然我们也可写之为行向量的形式 $(a_1, a_2, \cdots, a_n)$, 不过本书为了符号统一起见，总是写为列向量的形式．例如 (1) 的解 $\mathbf{x}$ 是 n 维向量，非齐次项 $\mathbf{b}$ 就是 m 维向量．n 维向量是空间解析几何中三维向量的一种推广．

将 n 维向量看作一种特殊的矩阵，可以按矩阵的数乘和加法定义 n 维实向量 (复向量) 中的数乘与加法．设 $\mathbf{a} = (a_1, a_2, \cdots, a_n)^T$, $\mathbf{b} = (b_1, \cdots, b_n)^T$, 则对 $k \in \mathbf{R}$ (或 $k \in \mathbf{C}$), 定义数乘 $k\mathbf{a}$ 与加法 $\mathbf{a}+\mathbf{b}$ 分别为

$$k\mathbf{a} = (ka_1, \cdots, ka_n)^T, \qquad \mathbf{a}+\mathbf{b} = (a_1+b_1, \cdots, a_n+b_n)^T. \tag{2}$$

我们把含有上述运算的 n 维实 (复) 向量的全体构成的集合记做 $\mathbf{R}^n(\mathbf{C}^n)$. 这是最简单，但也最典型的一种线性空间．

二．线性空间的抽象定义

定义 1. *设 V 是一个集合，K 是实数集 $\mathbf{R}$ 或复数集 $\mathbf{C}$. 在 V 的元素间定义了一种* **加法**: *对任意 $\boldsymbol{a}, \boldsymbol{b} \in V$, 有唯一的 $\boldsymbol{x} \in V$ 与它们对应，我们称 $\boldsymbol{x}$ 为 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的和，记作 $\boldsymbol{a}+\boldsymbol{b}$; 加法满足如下公理：*

(1)(加法交换律) $\boldsymbol{a}+\boldsymbol{b} = \boldsymbol{b}+\boldsymbol{a}$;

(2)(加法结合律) $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c} = \boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$;

(3) (加法零元) V 中有一个元素 $\mathbf{0}$, 对任意 $\boldsymbol{a} \in V$, 成立 $\boldsymbol{a}+\mathbf{0} = \mathbf{0}$;

(4) (加法负元) 对任意 $\boldsymbol{a} \in V$, 存在一个 $\boldsymbol{b} \in V$, 称为 $\boldsymbol{a}$ 的负元 (记作 $-\boldsymbol{a}$), 使得 $\boldsymbol{a}+\boldsymbol{b} = \mathbf{0}$.

另外，在 K 与 V 之间还定义了一种运算叫 **数乘**: 对任意 $k \in K, \boldsymbol{a} \in V$, V 中有唯一的元素 $\boldsymbol{y}$ 与它们对应，我们称 $\boldsymbol{y}$ 为 k 与 $\boldsymbol{a}$ 的数乘，记为 $k\boldsymbol{a}$. 数乘要满足以下公理：

(5) $1\boldsymbol{a} = \boldsymbol{a}$;

(6) $k(l\boldsymbol{a}) = (kl)\boldsymbol{a}$.

此外加法和数乘之间是有联系的，我们还要求成立如下公理 (分配律):

(7) $(k+l)\boldsymbol{a} = k\boldsymbol{a} + l\boldsymbol{a}$;

(8) $k(\boldsymbol{a}+\boldsymbol{b}) = k\boldsymbol{a} + k\boldsymbol{b}$.

以上公理 (1),(2),(5)—(8) 中，$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 是 V 中任意元素，k, l 是 K 中任意数. 我们称满足上述要求的集合 V 是一个 K **上的线性空间**. $\mathbf{R}$ 上的线性空间也叫作 **实线性空间**, $\mathbf{C}$ 上的线性空间也叫作 **复线性空间**.

线性空间的元素也常被叫做向量，但它们实际上很可能是与 n 维向量完全不同的东西. 学习、理解线性空间这类抽象的概念的途径之一就是看一些具体的例子 (既包括正面的例子，也包括反面的例子), 做到具体与抽象的结合.

例 1. $\mathbf{R}^n, \mathbf{C}^n$, 按照前面定义的加法与数乘，可以验证满足所有八条公理，所以都是线性空间.

例 2. 仅含 n 维零向量 $\mathbf{0} = (0, \cdots, 0)^T$ 的集合 $\{\mathbf{0}\}$ 按照 n 维向量的加法与数乘，也满足所有八条公理，所以也是线性空间.

例 3. 所有 $[a,b]$ 区间上的连续实函数，以函数的加法为加法：$(f+g)(x) = f(x)+g(x)$; 以函数与实数的乘法为数乘：$(kf)(x) = kf(x)$, 容易验证满足公理 (1)—(8), 所以构成一个实线性空间，我们记作 $C[a,b]$.

例 4. 所有 $m \times n$ 阶实矩阵，按照矩阵的加法与数乘，构成一个实线性空间.

例 5. 上半平面 $\mathbf{R}^2_+ = \{(x_1, x_2) \in \mathbf{R}^2 : x_2 \geq 0\}$ 按照向量的加法与数乘不构成线性空间，因为数乘定义不合要求.

例 6. 全体三维向量构成的集合，依照向量的加法与如下定义的数乘：

$$k \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{对任意 } k \in \mathbf{R},$$

不构成 $\mathbf{R}$ 上的线性空间．这是因为公理 *(5)* 不成立．

从例 1— 例 4 可以看出，线性空间确实是包含了相当广泛的对象．但是，我们的主要目的是研究关于线性方程组的问题 (A)—(C), 在那里 $\mathbf{b}, \mathbf{x}$ 就是普通的 $\mathbf{R}^m$ 或 $\mathbf{R}^n$ 中的向量，似乎我们只要知道 $\mathbf{R}^n$ 就够了，又何必弄出这么抽象的概念出来呢？

事实上，即使就我们研究 $\mathbf{R}^n$ 中线性方程组的目的而言，引入抽象的线性空间的概念也是必要的．这是因为，虽然 $\mathbf{R}^n$ 本身确实比较简单，但是它的子集却可能很复杂．例如，我们下面会看到，方程组 $A\mathbf{x}=\mathbf{0}$ 的解集是 $\mathbf{R}^n$ 的子集，它也是一个线性空间．但这个线性空间就不是简单地用我们描述 $\mathbf{R}^n$ 的类似方式就能说得清的．所以，引入线性空间的抽象概念有助于我们看清问题实质，予以整体统一地研究．这种抽象也就成了一种强有力的思维武器．

我们仅从定义 1 出发就可推出如下线性空间的性质．

性质 1. 设 V 是 K 上线性空间．对任意 $\mathbf{a}\in V$, 成立 $0\mathbf{a}=\mathbf{0}$.

注意，这里 $0\mathbf{a}$ 表示 $0\in K$ 与 $\mathbf{a}$ 的数乘，而 $\mathbf{0}$ 是 V 中的加法零元．

证明．

$$\begin{aligned} 0\mathbf{a} &= (0+0)\mathbf{a} \qquad \text{利用 } K \text{ 中 } 0+0=0 \text{ 的性质} \\ &= 0\mathbf{a}+0\mathbf{a}, \qquad \text{公理 (7)} \end{aligned}$$

由公理 (4), 两边加上 $0\mathbf{a}$ 的加法负元，就得到 $0\mathbf{a}+(-0\mathbf{a})=(0\mathbf{a}+0\mathbf{a})+(-0\mathbf{a})$, 左边就是 $\mathbf{0}$, 而右边用公理 (2), 就得到 $0\mathbf{a}+(0\mathbf{a}+(-0\mathbf{a}))=0\mathbf{a}+\mathbf{0}=0\mathbf{a}$. 其中最后一个等号用了公理 (3). 这就证明了 $0\mathbf{a}=\mathbf{0}$.

性质 2. 设 V 是 K 上线性空间．对任意 $\mathbf{a}\in V$, 成立 $(-1)\mathbf{a}=-\mathbf{a}$.

注意，这里 $(-1)\mathbf{a}$ 表示 $-1\in K$ 与 $\mathbf{a}$ 的数乘，而 $-\mathbf{a}$ 是 $\mathbf{a}$ 的加法负元．

证明．

$$\begin{aligned} \mathbf{a}+(-1)\mathbf{a} &= 1\mathbf{a}+(-1)\mathbf{a} \qquad \text{公理 (5)} \\ &= (1+(-1))\mathbf{a} \qquad \text{公理 (7)} \\ &= 0\mathbf{a}=\mathbf{0}, \qquad \text{性质 1} \end{aligned}$$

从而由公理 (4) 加法负元的定义，$(-1)\mathbf{a}=-\mathbf{a}$.

习题 1. 检验下列集合对于给定的运算是否构成实数集 $\mathbf{R}$ 上的线性空间．

(A) 实数对 $(a,b), a,b\in\mathbf{R}$ 全体，定义运算：

$$(a_1,b_1)\oplus(a_2,b_2)=(a_1+a_2,b_1+b_2+a_1a_2),$$
$$k\circ(a_1,b_1)=(ka_1,kb_1+\frac{k(k-1)}{2}a_1^2).$$

注意这里为了与普通的加法和数乘区别，我们分别用 “$\oplus$” 和 “$\circ$” 表示元素的加法和数乘.

(B) 全体正实数构成的集合，其中定义运算

$$a\oplus b=ab,$$
$$k\circ a=a^k.$$

三. 线性子空间

定义 2. *设 V 是一个 K 上的线性空间， W 是 V 的一个子集. 若 W 依照 V 中的加法与数乘也构成 K 上的线性空间，则称 W 是 V 的一个线性子空间.*

为什么要引入线性子空间的概念呢？这是因为，一方面，从下面的例 7, 例 8 可以看出，我们感兴趣的往往不是线性空间 $\mathbf{R}^n$ 本身，而是其中的某些线性子空间，如线性方程组的解空间；另一方面，研究线性空间的一个办法就是把它 (通过运算)“分解 ” 为若干个子空间，如果搞清楚了相对较简单的线性子空间，那么大空间也就搞清楚了.

我们经常需要判断一个线性空间的子集是否是线性子空间. 如果按照定义 2, 就需要一条条验证公理. 这有时是比较烦人的. 所以有如下命题可以帮助我们提高效率.

命题 1. *若 W 是 K 上线性空间 V 的子集，且关于 V 中加法及数乘是封闭的 (即任意两个 W 中向量按照 V 的加法的和仍在 W 中， $k\in K$ 与 W 中任意向量的数乘也仍在 W 中), 则 W 就是 V 的一个线性子空间.*

证明. 只需逐一验证公理 (1)—(8) 对 W 成立.

(1) 对 $\mathbf{a},\mathbf{b}\in W\subset V$, 由于 V 是线性空间, 对 V 公理 (1) 成立：$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$, 从而得证.

(2)(5)(6)(7)(8) 类似于 (1) 可证.

(3) 首先说明 W 中加法零元的存在性．由性质 2, 对于任意 $\mathbf{a}\in W$, 成立 $0\mathbf{a}=\mathbf{0}$, 这里 $\mathbf{0}$ 是 V 中的加法零元. 由于 W 关于数乘封闭, 我们就知道 $\mathbf{0}\in W$. 下面验证 V 中的加法零元 $\mathbf{0}$ 也就是 W 中的加法零元．事实上，由 V 中的公理 (3), 任取 $\mathbf{a}\in W\subset V$, 都成立 $\mathbf{a}+\mathbf{0}=\mathbf{a}$. 所以在 W 中公理 (3) 成立.

(4) 对任意 $\mathbf{a}\in W$, 将 $\mathbf{a}$ 看作 V 中向量，则由性质 2, 成立 $(-1)\mathbf{a}=-\mathbf{a}$. 由于 W 关于数乘封闭，故 $(-1)\mathbf{a}\in W$, 从而， $-\mathbf{a}\in W$. 这就是说， $\mathbf{a}$ 的在 V 中的负元自然在 W 中．再由负元定义， $\mathbf{a}+(-\mathbf{a})=\mathbf{0}$. 这就表明， $\mathbf{a}$ 的在 V 中的负元也就是它在 W 中的负元．公理 (4) 也成立．证毕！

下面我们看一些关于线性子空间的例子.

例 7. *线性方程组 $A\mathbf{x}=\mathbf{0}$ 的解集是 $\mathbf{R}^n$ 的线性子空间．这里 A 是 $m\times n$ 矩阵, $\boldsymbol{x}\in\mathbf{R}^n$.*

证明． $A\mathbf{x}=\mathbf{0}$ 的解集，记作 $Ker(A)$, 显然是 $\mathbf{R}^n$ 的子集．现设 $\mathbf{x},\mathbf{y}$ 满足 $A\mathbf{x}=\mathbf{0}$, 则由矩阵乘法分配律， $A(\mathbf{x}+\mathbf{y})=A\mathbf{x}+A\mathbf{y}=\mathbf{0}+\mathbf{0}=\mathbf{0}$. 又对于任意 $k\in\mathbf{R}$, $A(k\mathbf{x})=kA\mathbf{x}=k\mathbf{0}=\mathbf{0}$. 所以 $Ker(A)$ 关于 $\mathbf{R}^n$ 的加法和数乘均封闭．由命题 1, $Ker(A)$ 是 $\mathbf{R}^n$ 的线性子空间.

例 8. *设 A 是 $m\times n$ 矩阵，考察集合 $R(A)=\{\boldsymbol{y}\in\mathbf{R}^m:$ 存在 $\boldsymbol{x}\in\mathbf{R}^n$使得 $\boldsymbol{y}=A\boldsymbol{x}\}$. 则 $R(A)$ 是 $\mathbf{R}^m$ 的线性子空间.*

证明．设 $\mathbf{y},\tilde{\mathbf{y}}\in R(A)$, 则存在 $\mathbf{x},\tilde{\mathbf{x}}\in\mathbf{R}^n$ 使得 $\mathbf{y}=A\mathbf{x},\tilde{\mathbf{y}}=A\tilde{\mathbf{x}}$. 于是成立 $A(\mathbf{x}+\tilde{\mathbf{x}})=\mathbf{y}+\tilde{\mathbf{y}}$. 这就是说 $\mathbf{y}+\tilde{\mathbf{y}}\in R(A)$. 类似的, 由于 $A(k\mathbf{x})=k\mathbf{y}$, 故 $k\mathbf{y}\in R(A)$. 所以由命题 1, $R(A)$ 是 $\mathbf{R}^m$ 的线性子空间.

例 9. *集合 $U_1=\{f\in C[-1,1]:f(1)-f(-1)=0\}$ 是 $C[-1,1]$ 的线性子空间 (回忆例 3).*

例 10. *集合 $U_2=\{f\in C[-1,1]:f(1)f(-1)=0\}$ 不是 $C[-1,1]$ 的线性子空间．这是因为它关于加法不封闭 (请自己举个例子).*

例 11. *集合 $W=\{\boldsymbol{x}=(x_1,x_2,x_3)^T\in\mathbf{R}^3:2x_1-7x_2+x_3=1\}$ 不是 $\mathbf{R}^3$ 的子空间．因为它关于加法和数乘都不封闭．例如， $(0,0,1)^T\in W$, 但 $(0,0,2)^T\notin W$.*

例 12. *设 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 是线性空间 V 中的一组向量．定义*

$$W=\{\boldsymbol{x}\in V:\boldsymbol{x}=k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m,\forall k_1,\cdots,k_m\in K\},$$

即 W 中的向量由 $\alpha_1,\cdots,\alpha_m$ 通过数乘和加法得到，则 W 为 V 的线性子空间，记为 $W=\mathrm{span}\{\alpha_1,\cdots,\alpha_m\}$, 称作由 $\alpha_1,\cdots,\alpha_m$ 张成的线性子空间.

习题 2. 设矩阵 $A=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 其中 α_i 是其列向量. 证明：线性方程组 $A\mathbf{x}=\mathbf{b}$ 有解当且仅当 $\mathbf{b}\in\mathrm{span}\{\alpha_1,\cdots,\alpha_m\}$.

习题 3. 证明： $\mathbf{R}^2$ 中在经过原点的一条直线上的所有向量构成 $\mathbf{R}^2$ 的一个线性子空间；如果某条直线不经过原点，则端点在这条直线上的向量全体不构成 $\mathbf{R}^2$ 的线性子空间.

第 2 节　向量间的线性关系

上一节我们引入了研究对象 — 线性空间 — 的公理化定义. 那么，我们怎么进一步了解一个线性空间的结构呢？代数学的一个基本思想，就是找到线性空间中若干个 “相互独立 ” 的向量，通过数乘和加法运算，由它们把所有的线性空间中的向量表示出来. 这组相互独立的向量，就称为该线性空间的一组基，而其中向量的个数，就叫做这个线性空间的维数. 一个 n 维的线性空间，在给定一组基之后，就可以和 $\mathbf{R}^n$(或 $\mathbf{C}^n$, 要看这个线性空间是 $\mathbf{R}$ 上的还是 $\mathbf{C}$ 上的) 等同起来. 这种想法与在平面上引入笛卡儿坐标类似，可把由公理定义的对抽象向量的研究转化为对表示坐标的数组 —n 维向量 — 的研究. 大家在学习时也要注意这种代数与几何的联系.

下面，我们就把上述想法严格化. 比如，什么叫向量之间 “相互独立 ”? 什么叫把一个向量用一组向量 “表示 ” 出来？我们都需要给出精确的数学定义. 这些就是以下几节的内容. 我们在这段学习中要特别注意哪些定义定理讲的是对抽象向量而言的，哪些是仅对 n 维向量 (即 $\mathbf{R}^n$ 中或 $\mathbf{C}^n$ 中向量) 而言的，切不可混淆. 此外还必须融会贯通，学会并熟练掌握将向量语言描述的概念、定理转化为线性方程组及矩阵的相应概念、定理.

一. 向量的线性组合

定义 1. 设 V 是 K 上线性空间，对于一组 (抽象) 向量 $\alpha_1,\alpha_2,\cdots,\alpha_m,\beta\in V$, 如果存在数 $k_1,k_2,\cdots,k_m\in K$ 使得

$$\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m,$$

则称向量 β 是向量 $\alpha_1,\cdots,\alpha_m$ 的一个**线性组合**, 或称向量 β 可由向量 $\alpha_1,\cdots,\alpha_m$**线性表示**. 称 $k_1,\cdots,k_m$ 为 **组合系数** 或 **表示系数**.

例 1. 在 $\mathbf{R}^n$ 中，称

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

为 n 维 **单位向量**. 任何 n 维向量 $\boldsymbol{b}=\begin{pmatrix} b_1 & b_2 & \cdots & b_n\end{pmatrix}^T$ 均可由单位向量线性表出： $\boldsymbol{b}=\sum_{i=1}^n b_i\boldsymbol{e}_i$.

例 2. 线性方程组的向量表示. 如下一般线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\ \qquad\cdots\qquad\cdots\qquad\cdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m,\end{cases} \tag{3}$$

可以写为矩阵形式 $Ax=\boldsymbol{b}$, 或向量形式

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots x_n\boldsymbol{a}_n=\boldsymbol{b},$$

其中

$$\boldsymbol{a}_i=\begin{pmatrix}a_{1i}\\a_{2i}\\\vdots\\a_{mi}\end{pmatrix},\ 1\le i\le n,\qquad \boldsymbol{b}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_m\end{pmatrix}.$$

于是，线性方程组的求解可以看作寻求 n 个 m 维向量 $\mathbf{a}_1,\cdots,\mathbf{a}_n$ 的线性组合 (的组合系数) 以得到 $\mathbf{b}$.

于是，由线性方程组的求解理论，我们可以得到如下判定定理.

定理 1. 设 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m,\boldsymbol{b}$ 为 n 维向量. 则 $\boldsymbol{b}$ 可由 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$ 线性表示的充分必要条件为线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 有解，其中 $A=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$ 是 $n\times m$ 矩阵.

推论. (1) $\mathbf{b}$ 可由 $\mathbf{a}_1,\cdots,\mathbf{a}_m$ 线性表示的充分必要条件为 $r(A)=r(A\,\mathbf{b})$;

(2) 进而，其表示系数唯一 $\Leftrightarrow r(A)=r(A\,\mathbf{b})=m$;

(3) $r(A)=r(A\,\mathbf{b})<m$ 时表示系数不唯一;

(4) $\mathbf{b}$ 不能由 $\mathbf{a}_1,\cdots,\mathbf{a}_m$ 线性表示的充分必要条件为 $r(A)\neq r(A\,\mathbf{b})$.

证明. 利用上一章线性方程组的可解性判定定理即可.

例 3. 判断下面四个向量中哪个不能由其余三个向量线性表示：

$$\boldsymbol{a}_1=\begin{pmatrix}1\\1\\1\\1\end{pmatrix},\ \boldsymbol{a}_2=\begin{pmatrix}0\\5\\2\\1\end{pmatrix},\ \boldsymbol{a}_3=\begin{pmatrix}1\\-1\\0\\0\end{pmatrix},\ \boldsymbol{a}_4=\begin{pmatrix}2\\-3\\0\\1\end{pmatrix}.$$

解. 对 $\begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4\end{pmatrix}$ 作行初等变换 $\Rightarrow \begin{pmatrix}1&0&1&2\\0&1&-2/5&-1\\0&0&1&0\\0&0&0&0\end{pmatrix}$. 所以可看出 $r(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4)=3, r(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4)=2$, 故 $\mathbf{a}_3$ 不能由 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_4$ 线性表示.

习题 1. 把向量 $\mathbf{b}$ 表为向量 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4$ 的线性组合：

$$\mathbf{b}=\begin{pmatrix}1\\2\\1\\1\end{pmatrix},\mathbf{a}_1=\begin{pmatrix}1\\1\\1\\1\end{pmatrix},\mathbf{a}_2=\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix},\mathbf{a}_3=\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix},\mathbf{a}_4=\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}$$

二. 向量组的线性相关、线性无关

定义 2. 对一组 (抽象) 向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$, 如果存在**不全为零的**数 $k_1,k_2,\cdots,k_m$ 使得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m=\mathbf{0},$$

则称向量 $\alpha_1,\cdots,\alpha_m$**线性相关**. 否则，若上式仅对 $k_1=\cdots=k_m=0$ 成立，则称向量 $\alpha_1,\cdots,\alpha_m$**线性无关**.

一组线性无关的向量可以认为是"相互独立"的向量. 线性相关在几何上，就相当于向量共线或共面等.

特别地，对于一组 n 维向量是否线性无关，利用线性方程组理论，我们有如下判定定理.

定理 2. n 维向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 线性无关的充要条件为相应齐次线性方程 $Ax=0$ 只有零解. 这里利用分块矩阵写法， $A=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)$.

推论 1. n 维向量 $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$ 线性相关 (无关)$\Leftrightarrow r(A)<m$ $(r(A)=m)$.
推论 2. n 个 n 维向量 $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n$ 线性相关 (无关)$\Leftrightarrow |A|=0$ $(|A|\neq 0)$.
推论 3. m 个 n 维向量 $(m>n)$ 必线性相关. (注意 m 是未知数个数， n 是方程个数.)

例 4. 单位向量线性无关. $\{\mathbf{0}\}$ 线性相关.

例 5. 讨论下列向量组的线性相关性：

$$\boldsymbol{a}=\begin{pmatrix}2\\-1\\1\\3\end{pmatrix},\ \boldsymbol{b}=\begin{pmatrix}1\\0\\4\\2\end{pmatrix},\ \boldsymbol{c}=\begin{pmatrix}-4\\2\\-2\\k\end{pmatrix}.$$

解. 作初等变换可得 $(\mathbf{a}\ \mathbf{b}\ \mathbf{c})\sim\begin{pmatrix}1&0&-2\\0&1&0\\0&0&k+6\\0&0&0\end{pmatrix}$. 从而 $k=-6$ 时秩为 $2<3\Leftrightarrow$ 线性相关.

例 6. 设向量 $\alpha_1,\cdots,\alpha_m$ 线性无关. 又 $(\beta_1\ \beta_2\cdots\ \beta_m)=(\alpha_1\ \alpha_2\cdots\alpha_m)A$, 其中 A 为 $m\times m$ 矩阵，则 $\beta_1,\cdots,\beta_m$ 线性无关 $\Leftrightarrow |A|\neq 0$.

证明. $\Leftarrow$. 如果 $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_m$ 都是 n 维向量, 那么利用矩阵的秩在可逆矩阵右乘下不变性， $r(\beta_1\ \beta_2\cdots\ \beta_m)=r(\alpha_1,\cdots,\alpha_m)$. 又 $\alpha_1,\cdots,\alpha_m$ 线性无关，利用推论 1, $r(\alpha_1,\cdots,\alpha_m)=m$. 于是 $r(\beta_1\ \beta_2\cdots\ \beta_m)=m$. 再用推论 1 知 $\beta_1,\cdots,\beta_m$ 线性无关.

但是这里题目并没有明确指出 $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_m$ 都是 n 维向量，从而可能都是抽象向量，上面的证明就行不通了．所以我们用定义和反证法．假设 $\beta_1,\cdots,\beta_m$ 线性相关，则存在不全为零的数 $k_1,\cdots,k_m$ 使得 $k_1\beta_1+\cdots+k_m\beta_m=\mathbf{0}$. 这就意味着 $(\alpha_1,\cdots,\alpha_m)A(k_1,\cdots,k_m)^T=\mathbf{0}$. 由于 $\alpha_1,\cdots,\alpha_m$ 线性无关，必有 $A(k_1,\cdots,k_m)^T=\mathbf{0}$. 但是由于 A 非奇异，由 Cramer 法则，应只有零解 $k_1=\cdots=k_m=0$, 与它们不全为零的假设矛盾！

$\Rightarrow$. 假若 $|A|=0$, 则存在非零向量 $\mathbf{c}$ 使得 $A\mathbf{c}=0$. 从而 $(\beta_1,\cdots,\beta_m)\mathbf{c}=\mathbf{0}$, 即知 $\beta_1,\cdots,\beta_m$ 线性相关．证毕.

例 7. 证明：若向量组 $\alpha_1,\cdots,\alpha_m$ 中有部分向量线性相关，则该向量组线性相关.

证明．不妨设 $\alpha_1,\cdots,\alpha_l(l\le m)$ 线性相关, 则存在不全为零的数 $k_1,\cdots,k_l$ 使得 $\sum_{i=1}^{l}k_i\alpha_i=0$. 于是存在不全为零的数 $k_1,k_2,\cdots,k_l,k_{l+1}=0,\cdots,k_m=0$ 使得 $\sum_{i=1}^{m}k_i\alpha_i=0$. 于是 $\alpha_1,\cdots,\alpha_m$ 线性相关.

习题 2. 证明：在 $C[0,1]$(见第 1 节例 3) 中，向量 $1,\cos^2 t,\cos 2t$ 线性相关.

三．线性组合与线性相关的关系

定理 3. 向量 $\alpha_1,\cdots,\alpha_m(m\ge 2)$ 线性相关 $\Leftrightarrow$ 其中至少有一个向量是其余向量的线性组合.

证明．必要性．由于 $\alpha_1,\cdots,\alpha_m(m\ge 2)$ 线性相关，故有不全为零的数 $k_1,\cdots,k_m$ 使得 $\sum_{i=1}^{m}k_i\alpha_i=\mathbf{0}$. 不妨设 $k_1\ne 0$, 那么我们就得到

$$\alpha_1=\sum_{i=2}^{m}(-\frac{k_i}{k_1})\alpha_i.$$

也就是说 α_1 可由其它向量线性表示.

充分性．不妨设 α_1 可由其它向量线性表示：

$$\alpha_1=\sum_{i=2}^{m}k_i\alpha_i.$$

则

$$-\alpha_1+\sum_{i=2}^{m}k_i\alpha_i=\mathbf{0}.$$

由于 α_1 前系数为 $-1\ne 0$, 故 $\alpha_1,\cdots,\alpha_m(m\ge 2)$ 线性相关．证毕.

定理 4. 向量 $\alpha_1,\cdots,\alpha_m$ 线性无关，但向量 $\alpha_1,\cdots,\alpha_m,\beta$ 线性相关，则 β 可唯一地由 $\alpha_1,\cdots,\alpha_m$ 线性表示.

证明. $\alpha_1,\cdots,\alpha_m,\beta$ 线性相关，故有不全为零的数 $k_1,\cdots,k_m,k$ 使得

$$\sum_{i=1}^{m} k_i\alpha_i + k\beta = \mathbf{0}.$$

我们断言必然成立 $k \neq 0$. 否则，如果 $k=0$, 那么 $k_1,\cdots,k_m$ 必不全为零且由上式成立 $\sum_{i=1}^{m} k_i\alpha_i = \mathbf{0}$. 这就与 $\alpha_1,\cdots,\alpha_m$ 线性无关矛盾．故 $k \neq 0$. 这样一来，就得到

$$\beta = \sum_{i=1}^{m}(-\frac{k_i}{k})\alpha_i,$$

即 β 可以由 $\alpha_1,\cdots,\alpha_m$ 线性表示.

再证上述线性表示的唯一性．设有两种线性表示：

$$\beta = \sum_{i=1}^{m} k_i\alpha_i, \quad \beta = \sum_{i=1}^{m} l_i\alpha_i.$$

两式相减，就得到

$$\mathbf{0} = \sum_{i=1}^{m}(k_i - l_i)\alpha_i.$$

由于 $\alpha_1,\cdots,\alpha_m$ 线性无关，必需成立 $k_i - l_i = 0, i = 1,\cdots,m$. 所以线性表示系数必然唯一．证毕.

如果仅需对 n 维向量证明定理 4, 那我们可以不从定义出发而给出如下证明：向量 $\alpha_1,\cdots,\alpha_m$ 线性无关，故由定理 2 的推论 1, $r(\alpha_1,\cdots,\alpha_m) = m$. $\alpha_1,\cdots,\alpha_m,\beta$ 线性相关，则 $r(\alpha_1,\cdots,\alpha_m,\beta) < m+1$. 由于 $r(\alpha_1,\cdots,\alpha_m) \leq r(\alpha_1,\cdots,\alpha_m,\beta)$, 可得 $r(\alpha_1,\cdots,\alpha_m) = r(\alpha_1,\cdots,\alpha_m,\beta) = m$. 这就是说线性方程组 $A\mathbf{x} = \beta$ 有唯一解，故而 β 可唯一地由 $\alpha_1,\cdots,\alpha_m$ 线性表示．注意这里 $A = (\alpha_1,\cdots,\alpha_m)$. 证毕.

例 8. $\alpha_1,\cdots,\alpha_n$ 是 n 个线性无关的向量，

$$\alpha_{n+1} = \sum_{i=1}^{n} k_i\alpha_i,$$

其中 k_i**全不为零**. 证明： $\alpha_1,\cdots,\alpha_n,\alpha_{n+1}$ 中任意 n 个向量均线性无关.

证明. 不失一般性, 我们证明 $\alpha_2,\cdots,\alpha_n,\alpha_{n+1}$ 线性无关. 用反证法. 假设存在不全为零的数 $l_2,\cdots,l_n,l_{n+1}$ 使得 $\sum_{i=2}^{n+1} l_i\alpha_i=\mathbf{0}$. 将 α_{n+1} 的表达式代入, 并合并同类项 (利用向量加法满足交换律和结合律), 我们得到

$$l_{n+1}k_1\alpha_1+\sum_{j=2}^{n}(l_j+l_{n+1}k_j)\alpha_j=\mathbf{0}.$$

由于 $\alpha_1,\cdots,\alpha_n$ 线性无关，故成立

$$l_{n+1}k_1=0,\qquad l_j+l_{n+1}k_j=0, j=2,\cdots,n.$$

由于 $k_i(i=1,\cdots,n)$ 全不为零，我们可解得 $l_{n+1}=0, l_j=0, j=2,\cdots,n$, 与前面假设它们不全为零矛盾．故 $\alpha_2,\cdots,\alpha_n,\alpha_{n+1}$ 线性无关．证毕.

例 9. *设 n 维单位向量 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$ 可由 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 线性表示．证明：$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 线性无关.*

证明．由题意，存在向量 $\mathbf{x}_1,\cdots,\mathbf{x}_n$ 使得 $\mathbf{e}_i=A\mathbf{x}_i$, 其中 A 就是由列向量构成的矩阵 $(\mathbf{a}_1,\cdots,\mathbf{a}_n)$. 置 $X=(\mathbf{x}_1,\cdots,\mathbf{x}_n)$, 我们就有

$$AX=(\mathbf{e}_1,\cdots,\mathbf{e}_n)=E_n.$$

由于单位矩阵 E_n 可逆，所以 $|A||X|=1\neq 0$. 所以 $|A|\neq 0$. 这就是说 $r(A)=n$. 从而由定理 2 推论 2, $\mathbf{a}_1,\cdots,\mathbf{a}_n$ 线性无关．证毕.

注记．从上述例题我们看出，研究 n 维向量线性关系的重要工具是线性方程组理论, 而对于抽象向量, 目前还只能依赖于定义. 以后我们会看到, 对于抽象向量的线性关系，也可化为 n 维向量的线性关系．所以，总的来说，线性方程组的理论是我们研究向量及线性空间的重要武器. 大家在学习中要特别注意把向量的语言与线性方程组的语言、矩阵的语言对应起来，这样才可以融会贯通，抓住实质，不至于被新概念牵着鼻子走.

四．向量组的等价

定义 3. *设有两个向量组：*

$$(\mathrm{I}):\alpha_1,\alpha_2,\cdots,\alpha_s;\qquad(\mathrm{II}):\beta_1,\beta_2,\cdots,\beta_t.$$

若 (I) 中的每个向量都可由 (II) 中的向量线性表示，就称向量组 (I) 可由向量组 (II) 线性表示．若 (I) 与 (II) 可以互相线性表示，就说它们是等价的.

例 10. *将向量组 (I) 可由向量组 (II) 线性表示用矩阵的语言表示.*

解. 设 $\alpha_m = \sum_{i=1}^{t} l_{im}\beta_i$, $m = 1, \cdots, s$. 那么利用矩阵乘法，这些等式可以合写为

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_s \end{pmatrix} = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_t \end{pmatrix} \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1s} \\ l_{21} & l_{22} & \cdots & l_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ l_{t1} & l_{t2} & \cdots & l_{ts} \end{pmatrix}_{t\times s}. \tag{4}$$

特别地，如果 (I) 和 (II) 均是由 n 维向量构成的向量组，那么若记

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_s \end{pmatrix},\ B = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_t \end{pmatrix},\ P = (l_{ij})_{t\times s},$$

(4) 就是

$$A = BP. \tag{5}$$

例 11. *设 (I) 和 (II) 均是由 n 维向量构成的向量组. 那么 (I) 可由 (II) 线性表示的充要条件为 $r(B) = r(B\quad A)$, 其中 $(B\quad A)$ 是 B, A 构成的分块矩阵.*

证明. 由 (5), (I) 可由 (II) 线性表示的充要条件是方程组 $BX = A$ 有解. 利用线性方程组可解性判定定理，这当且仅当 $r(B) = r(B\quad A)$. 证毕.

例 12. *设向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 可由向量组 $\beta_1, \beta_2, \cdots, \beta_t$ 线性表示, $\beta_1, \beta_2, \cdots, \beta_t$ 又可由 $\gamma_1, \gamma_2, \cdots, \gamma_m$ 线性表示, 则 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 可由向量组 $\gamma_1, \gamma_2, \cdots, \gamma_m$ 线性表示.*

证明. 注意此题所指向量未必就是 n 维向量, 所以我们只能使用 (4) 式. 不妨设

$$\begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_t \end{pmatrix} = \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_m \end{pmatrix} Q_{m\times t}.$$

则将该式代入 (4), 就得到

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_s \end{pmatrix} = \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_m \end{pmatrix} Q_{m\times t} P_{t\times s}.$$

这就证明了 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 可由向量组 $\gamma_1, \gamma_2, \cdots, \gamma_m$ 线性表示，其表示系数由矩阵 $Q_{m\times t}P_{t\times s}$ 给出.

最后，我们证明如下重要定理.

定理 5. 对给定向量组

$$(\mathrm{I}):\alpha_1,\alpha_2,\cdots,\alpha_s; \qquad (\mathrm{II}):\beta_1,\beta_2,\cdots,\beta_t.$$

若 *(I)* 可由 *(II)* 线性表示，且 $s>t$, 则 *(I)* 必线性相关.

证明. 我们仍沿用 (4) 的记号. 由于 $s>t$, 成立 $r(P)\leq t<s$, 故线性方程组 $P\mathbf{x}=\mathbf{0}$ 有非零解，其中 $\mathbf{x}$ 是 s 维向量. 于是 (4) 两边同时右乘 $\mathbf{x}$, 就得到

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_s \end{pmatrix}\mathbf{x}=\mathbf{0}.$$

从而 $\alpha_1,\cdots,\alpha_s$ 线性相关. 证毕.

推论. 若 (I) 可由 (II) 线性表示，且 (I) 线性无关，则 $s\leq t$.

习题 3. 证明此推论.

第 3 节　向量组的极大无关组和秩

我们这一节的目的，就是希望对给定的向量组 (I), 找到其中的一部分向量构成向量组 (II), 使得 (I) 可以由 (II) 线性表示. 我们当然希望向量组 (II) 越简单越好，也就是说，它所含向量的个数应该最少. 此时就称 (II) 是 (I) 的一个极大无关组，其所含向量的个数就称作向量组 (I) 的秩.

如果我们知道了一个向量组 (I) 的秩，我们不难利用向量间的线性相关 (无关) 性，找出它的一个极大无关组. 从这个极大无关组就可以通过线性表示生成 (I). 这时，我们对向量组 (I) 就算有了完全透彻的了解.

一. 极大线性无关组与秩

定义 1. 设 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 中部分向量构成的向量组，若满足：

(a) $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是线性无关的；

(b) $\alpha_1,\alpha_2,\cdots,\alpha_m$ 可以由 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性表示,

就说向量组 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的一个 **极大 (线性) 无关组**, 而 r, 即极大无关组所含向量的个数，被叫做向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的 **秩**, 记作 $r[\alpha_1,\alpha_2,\cdots,\alpha_m]$. 规定 $r[\mathbf{0}]=0$.

显然，一个向量组与其极大无关组是等价的．此外，注意一个给定向量组的极大无关组可能是不唯一的．不过利用上一节例 12 的结论，可知一个给定向量组的任意两个极大无关组是等价的.

习题 1. 证明上述论断.

下面的定理保证了极大无关组的存在性.

定理 1. 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 有极大无关组当且仅当 $\alpha_i(i=1,2,\cdots,m)$ 不全为零向量.

证明．必要性．用反证法．若 $\alpha_i(i=1,2,\cdots,m)$ 全为零向量，则由于零向量自己就是线性相关的，定义 1 的条件 (a) 不会成立，与 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 有极大无关组矛盾.

充分性. 由条件可取到 $\alpha_i\in\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$ 且 $\alpha_i\neq\mathbf{0}$. 若 α_i 就是 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的极大无关组，则定理已经得到证明．否则，必有

$$\alpha_j\in\{\alpha_1,\alpha_2,\cdots,\alpha_m\}\backslash\{\alpha_i\}$$

不能由 α_i 线性表示. 这就意味着，α_i,α_j 是线性无关的. 若 α_i,α_j 还不是极大无关组，则必有 α_k 不能由 α_i,α_j 线性表示，所以 $\alpha_i,\alpha_j,\alpha_k$ 是线性无关的．若 $\alpha_i,\alpha_j,\alpha_k$ 还不是极大无关组, 又可找到 α_l, ⋯ ⋯. 依次类推, 由于 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 向量个数是有限个，就可以在有限次后找到其极大无关组．证毕.

习题 2. 设 $\alpha_1,\cdots,\alpha_m$ 线性无关， β 不能由 $\alpha_1,\cdots,\alpha_m$ 线性表示．证明：$\alpha_1,\cdots,\alpha_m,\beta$ 线性无关.

一个向量组的秩是该向量组本身所具有的性质, 而与极大无关组的选取无关. 换句话说, 就是其所有极大无关组所含向量个数是一样的. 事实上, 设 $\alpha_1,\cdots,\alpha_s$ 和 $\beta_1,\cdots,\beta_t$ 都是某个给定向量组的极大无关组，则它们可以互相表示 (等价). 所以从上一节定理 5 的推论可知，必有 $s\leq t$ 且 $t\leq s$. 于是 $t=s$.

例 1. 向量组 $\alpha_1,\cdots,\alpha_s$ 线性无关, 则它就是自己的极大无关组, 且 $r[\alpha_1,\cdots,\alpha_s]=s$.

例 2. 等价的向量组必有相同的秩.

证明：分别取两个向量组的极大无关组，则这两个极大无关组之间也是等价的，从上一节定理 5 的推论可知，它们所含向量个数一样．证毕.

例 3. 设向量组 $\alpha_1,\cdots,\alpha_s$ 可由 $\beta_1,\cdots,\beta_t$ 线性表示，则

$$r[\alpha_1,\cdots,\alpha_s]\le r[\beta_1,\cdots,\beta_t].$$

证明. 设 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 为 $\alpha_1,\cdots,\alpha_s$ 的一个极大无关组， $\beta_{j_1},\cdots,\beta_{j_p}$ 为 $\beta_1,\cdots,\beta_t$ 的一个极大无关组. 则 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 可以由 $\beta_{j_1},\cdots,\beta_{j_p}$ 线性表示. 注意到 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 线性无关，由上一节定理 5 推论， $r\le p$. 即得证.

注记. 在这两个证明中我们可以初窥选取代表元后带来的方便：只要把极大无关组的情形证明了，一般情形也随之解决. 下面判定定理的证明也深刻体现了这个思想.

习题 3. 向量组的秩是为了能够线性表示所有该向量组向量所需要的部分向量的最少个数. 换句话说，设 $r(\alpha_1\cdots,\alpha_n)=r$, 则不可能有 $\alpha_{i_1},\cdots,\alpha_{i_s}$ $(s<r)$ 能够线性表示所有的 $\alpha_j(j=1,\cdots,n)$.

定理 2. 设给定向量组 $\alpha_1,\cdots,\alpha_m$ 的秩是 r. 则该向量组中任意 r 个线性无关的向量 $\beta_1,\cdots,\beta_r$ 都构成它的一个极大无关组.

证明. 只需证明 $\alpha_1,\cdots,\alpha_m$ 可由 $\beta_1,\cdots,\beta_r$ 线性表示. 为此, 设 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 是 $\alpha_1,\cdots,\alpha_m$ 的一个极大无关组，只需证明 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 可以由 $\beta_1,\cdots,\beta_r$ 线性表示.

首先，利用 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 是极大无关组，我们可以用它表示出 $\beta_1,\cdots,\beta_r$:

$$\begin{pmatrix}\beta_1 & \beta_2 & \cdots & \beta_r\end{pmatrix}=\begin{pmatrix}\alpha_{i_1} & \alpha_{i_2} & \cdots & \alpha_{i_r}\end{pmatrix}P_{r\times r}. \tag{6}$$

注意 P 是普通的 r 阶数字方阵 (元素为实数或复数). 我们断言 P 可逆.

假设 $|P|=0$, 则 $r(P)<r$. 于是线性方程组 $P\mathbf{x}=\mathbf{0}$ 有非零解 $\mathbf{x}\in K^r$ $(K=\mathbf{R}$ 或 $\mathbf{C})$. 这就意味着 $\begin{pmatrix}\beta_1 & \beta_2 & \cdots & \beta_r\end{pmatrix}\mathbf{x}=\mathbf{0}$, 即 $\beta_1,\beta_2,\cdots,\beta_r$ 线性相关，与题设矛盾. 这就证明了 P 可逆.

现在，在 (6) 两边同时右乘 P^{-1}, 我们得到

$$\begin{pmatrix}\alpha_{i_1} & \alpha_{i_2} & \cdots & \alpha_{i_r}\end{pmatrix}=\begin{pmatrix}\beta_1 & \beta_2 & \cdots & \beta_r\end{pmatrix}P^{-1}.$$

这就证明了 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 可以由 $\beta_1,\cdots,\beta_r$ 线性表示. 证毕.

二. m 维向量组的秩与矩阵的秩

我们已经知道了，要研究一个向量组，关键是要找到它的一个极大无关组. 定理 1 虽然保证了极大无关组的存在性，但是按照其证明方法把极大无关组找出来，却是非常低效的. 定理 2 说，我们可以先把向量组的秩 r 确定出来，然后随便找 r 个线性无关的向量就可以了. 这就在一定程度上提高了寻求极大无关组的效率. 可是，我们有什么办法快速地确定一个向量组的秩呢？此时我们显然不能依靠其定义，必须寻求新的思路.

另一方面，对由 m 维向量构成的向量组 $\mathbf{a}_1, \cdots, \mathbf{a}_n$, 我们可以考虑 $m \times n$ 分块矩阵 $(\mathbf{a}_1, \cdots, \mathbf{a}_n)$ 的秩 $r(\mathbf{a}_1, \cdots, \mathbf{a}_n)$. 那么这个矩阵的秩与向量组的秩 $r[\mathbf{a}_1, \cdots, \mathbf{a}_n]$ 是否一样？如果一样的话, 我们就可以利用上一章矩阵的初等变换法快速确定秩了. 这些，就是本小节的主要内容.

定理 3. 设 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ 为 m 维向量，则 $r(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n) = r[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$.

证明. 设 $r(\mathbf{a}_1, \cdots, \mathbf{a}_n) = r$. 我们首先证明 $\mathbf{a}_1, \cdots, \mathbf{a}_n$ 中必有 r 个向量线性无关.

由于 $r(\mathbf{a}_1, \cdots, \mathbf{a}_n) = r$, 由矩阵秩的定义，矩阵 $A = (\mathbf{a}_1, \cdots, \mathbf{a}_n)$ 必然有一个 r 阶子式 $D \neq 0$. 设 D 位于 A 的第 $i_1, i_2, \cdots, i_r$ 列. 考察由 $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_r}$ 按顺序构成的 $m \times r$ 矩阵 A_r. 则 $r(A_r) = r$. 于是线性方程组 $A_r\mathbf{x} = \mathbf{0}$ 在 $\mathbf{R}^r$ 中只有零解. 于是 $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_r}$ 线性无关.

再证明 $\mathbf{a}_1, \cdots, \mathbf{a}_n$ 可以由 $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_r}$ 线性表示.

显然 $\mathbf{a}_{i_k}$ $(k = 1, \cdots, r)$ 可由 $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_r}$ 线性表示. 现设 $i_1 < \cdots < i_j < k < i_{j+1} < \cdots < i_r$. 作 $m \times (r+1)$ 矩阵 $A' = (\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_j}, \mathbf{a}_k, \mathbf{a}_{i_{j+1}}, \cdots, \mathbf{a}_{i_r})$. 则由于 A' 是 A 的子阵，成立 $r(A') \leq r(A) = r < r+1$. 于是线性方程组 $A'\mathbf{y} = \mathbf{0}$ 有非零解 $\mathbf{y} = (k_1, \cdots, k_j, a, k_{j+1}, \cdots, k_r)^T \in \mathbf{R}^{r+1}$. 也就是说存在不全为零的系数使得成立

$$\sum_{l=1}^{j} k_l \mathbf{a}_{i_l} + a\mathbf{a}_k + \sum_{l=j+1}^{r} k_l \mathbf{a}_{i_l} = \mathbf{0}.$$

由于 $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_r}$ 线性无关, 必然成立 $a \neq 0$. 这就是说 $\mathbf{a}_k$ 可以由 $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_r}$ 线性表示.

这样, 由定义, $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_r}$ 就是 $\mathbf{a}_1, \cdots, \mathbf{a}_n$ 的一个极大无关组. 故 $r[\mathbf{a}_1, \cdots, \mathbf{a}_n] = r$. 证毕.

例 4. 一个矩阵的行 (列) 向量组成的向量组的秩叫做这个矩阵的行 (列) 秩. 由于 $r(A)=r(A^T)$, 而 A^T 的列秩就是 A 的行秩，由定理 3, 可知一个矩阵的秩与它的行秩、列秩是一样的.

例 5. 求如下向量组的秩和它的一个极大无关组，并且用该极大无关组表示整个向量组：

$$\boldsymbol{a}_1=\begin{pmatrix}1\\-2\\1\end{pmatrix},\boldsymbol{a}_2=\begin{pmatrix}2\\-4\\2\end{pmatrix},\boldsymbol{a}_3=\begin{pmatrix}1\\0\\3\end{pmatrix},\boldsymbol{a}_4=\begin{pmatrix}0\\-4\\-4\end{pmatrix}.$$

解. 作

$$A=\begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4\end{pmatrix}=\begin{pmatrix}1 & 2 & 1 & 0\\-2 & -4 & 0 & -4\\1 & 2 & 3 & -4\end{pmatrix}.$$

经过行初等变换，得到

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

将 B 的第 1 列至 4 列分别记作 $\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3,\mathbf{b}_4$. 从这一步就可算得 $r[\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4]=2$.

下面再求一个极大无关组. 注意到 $r(\mathbf{b}_1,\mathbf{b}_3)=2$. 所以 $\mathbf{b}_1,\mathbf{b}_3$ 就是 $\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3,\mathbf{b}_4$ 的一个极大无关组. 我们说，相应地， $\mathbf{a}_1,\mathbf{a}_3$ 也就是 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4$ 的一个极大无关组. 我们只需证明 $\mathbf{a}_1,\mathbf{a}_3$ 线性无关，或者等价地，$r(\mathbf{a}_1,\mathbf{a}_3)=2$. 事实上，矩阵 $(\mathbf{b}_1,\mathbf{b}_3)$ 就是由矩阵 $(\mathbf{a}_1,\mathbf{a}_3)$ 通过行初等变换得来的，而我们熟知行初等变换并不改变矩阵的秩.

下面我们用极大无关组 $\mathbf{a}_1,\mathbf{a}_3$ 表示向量组 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4$. 显然我们只需写出 $\mathbf{a}_2,\mathbf{a}_4$ 如何被表示. 这就是求解有如下增广矩阵的线性方程组

$$\begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_3 & \vdots & \mathbf{a}_2 & \mathbf{a}_4\end{pmatrix}=\begin{pmatrix}1 & 1 & \vdots & 2 & 0\\-2 & 0 & \vdots & -4 & -4\\1 & 3 & \vdots & 2 & -4\end{pmatrix}.$$

我们已经由 A 作了行初等变换得到 B, 故可直接写出上述矩阵经行初等变换化为

$$\left(\begin{array}{cc:cc} 1 & 0 & 2 & 2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

我们可直接写出解了，分别是 $(2,0)^T, (2,-2)^T$. 于是

$$\mathbf{a}_2 = 2\mathbf{a}_1,\ \ \mathbf{a}_4 = 2\mathbf{a}_1 - 2\mathbf{a}_3.$$

注记. 这个例题实际上给出了计算给定 m 维向量组的秩、其极大无关组、以及用极大无关组表示原向量组的高效算法. 它还告诉我们，对矩阵作初等行变换不改变其列向量间的线性关系.

下面我们给出计算一组抽象向量秩的办法.

定理 4. 设 $\alpha_1, \cdots, \alpha_m$ 线性无关， $\beta_j = \sum_{i=1}^m a_{ij}\alpha_i, j = 1, \cdots, s$. 令 $A = (a_{ij})_{m\times s}$. 则 $r[\beta_1, \cdots, \beta_s] = r(A)$.

证明. 写为矩阵语言，我们有

$$(\beta_1, \beta_2, \cdots, \beta_s) = (\alpha_1, \alpha_2, \cdots, \alpha_m)A. \tag{7}$$

记 $r[\beta_1, \cdots, \beta_s] = r'$.

若 $\beta_1, \cdots, \beta_s$ 全为零向量，则由于 $\alpha_1, \cdots, \alpha_m$ 线性无关，必成立 $A = O$. 于是 $r(A) = r' = 0$, 定理成立.

下面设 $\beta_1, \cdots, \beta_s$ 不全为零向量，于是 $A \neq O$. 设 $r(A) = r$, 则 A 的列秩也是 r. 记 $A = (\mathbf{a}_1, \cdots, \mathbf{a}_r, \mathbf{a}_{r+1}, \cdots, \mathbf{a}_s)$. 不失一般性，不妨设 A 的前面 r 列 $\mathbf{a}_1, \cdots, \mathbf{a}_r$ 就是 $\mathbf{a}_1, \cdots, \mathbf{a}_r, \mathbf{a}_{r+1}, \cdots, \mathbf{a}_s$ 的极大无关组 (一般情形下在 (7) 两边同时做交换列的列初等变换即可). 设 $\mathbf{a}_k (k = r+1, \cdots, s)$ 可表示为

$$\mathbf{a}_k = l_{1k}\mathbf{a}_1 + l_{2k}\mathbf{a}_2 + \cdots + l_{rk}\mathbf{a}_r.$$

如果记

$$A_1 = (\mathbf{a}_1, \cdots, \mathbf{a}_r),$$

那么由 (7), 利用矩阵乘法的性质，就有

$$(\beta_1,\cdots,\beta_r)=(\alpha_1,\cdots,\alpha_m)A_1.$$

我们断言 $\beta_1,\cdots,\beta_r$ 就是 $\beta_1,\beta_2,\cdots,\beta_s$ 的极大无关组．从而得证 $r'=r$.

首先说明 $\beta_1,\cdots,\beta_r$ 线性无关. 设有 $\mathbf{c}=(c_1,\cdots,c_r)^T\in K$ 使得 $(\beta_1,\cdots,\beta_r)\mathbf{c}=\mathbf{0}$, 那么 $(\alpha_1,\cdots,\alpha_m)(A_1\mathbf{c})=\mathbf{0}$. 由于 $\alpha_1,\cdots,\alpha_m$ 线性无关，必有 $A_1\mathbf{c}=\mathbf{0}$. 但是 $r(A_1)=r=$ 未知数个数，故线性方程组 $A_1\mathbf{x}=\mathbf{0}$ 只有零解，故 $\mathbf{c}=\mathbf{0}$.

另一方面，对 $k=r+1,\cdots,s$, 由 (7) 有

$$\begin{aligned}\beta_k &= (\alpha_1,\alpha_2,\cdots,\alpha_m)\begin{pmatrix}a_{1k}\\a_{2k}\\\vdots\\a_{mk}\end{pmatrix}\\ &= (\alpha_1,\alpha_2,\cdots,\alpha_m)\mathbf{a}_k\\ &= (\alpha_1,\alpha_2,\cdots,\alpha_m)(\mathbf{a}_1,\cdots,\mathbf{a}_r)\begin{pmatrix}l_{1k}\\l_{2k}\\\vdots\\l_{rk}\end{pmatrix}\\ &= (\beta_1,\cdots,\beta_r)\begin{pmatrix}l_{1k}\\l_{2k}\\\vdots\\l_{rk}\end{pmatrix}.\end{aligned}$$

这就是说， $\beta_1,\cdots,\beta_s$ 均可由 $\beta_1,\cdots,\beta_r$ 线性表示．证毕.

注记．这个定理表明，通过引入一组线性无关的向量 (相当于建立坐标系), 则抽象向量的很多性质可化为 m 维向量的性质来研究．此外，这个定理的证明也为求解类似例 5 的问题提供了另一种思路.

例 6. 设向量 α,β,γ 线性无关．又

$$\xi=\alpha,\ \ \eta=\alpha+\beta,\ \ \zeta=\alpha-\beta-\gamma.$$

求 ξ,η,ζ 的秩和一个极大无关组.

解. 我们有
$$(\xi,\eta,\zeta)=(\alpha,\beta,\gamma)\begin{pmatrix}1&1&1\\0&1&-1\\0&0&-1\end{pmatrix}.$$
由于系数矩阵的秩为 3, 利用定理 4, $r[\xi,\eta,\zeta]=3$, 它们线性无关. 它们就是自己的极大无关组.

习题 4. 求下列向量组的秩和极大无关组.

(1) $\mathbf{a}_1=(6,4,1,-1,2)^T$, $\mathbf{a}_2=(1,0,2,3,-4)^T$, $\mathbf{a}_3=(1,4,-9,-16,22)^T$, $\mathbf{a}_4=(7,1,0,-1,3)^T$;

(2) $\mathbf{a}_1=(1,-1,2,4)^T$, $\mathbf{a}_2=(0,3,1,2)^T$, $\mathbf{a}_3=(3,0,7,14)^T$, $\mathbf{a}_4=(1,-1,2,0)^T$, $\mathbf{a}_5=(2,1,5,6)^T$.

习题 5. 设 $r\geq 2, \beta_1=\alpha_2+\alpha_3+\cdots+\alpha_r$, $\beta_2=\alpha_1+\alpha_3+\cdots+\alpha_r,\cdots,\beta_r=\alpha_1+\alpha_2+\cdots+\alpha_{r-1}$. 证明: $\beta_1,\cdots,\beta_r$ 和 $\alpha_1,\cdots,\alpha_r$ 有相同的秩.

最后, 我们给出利用定理 4 反过来估计矩阵秩的一个例子.

例 7. 设 A,B 均为 $m\times n$ 矩阵. 证明: $r(A+B)\leq r(A)+r(B)$.

证明. 设 $A=(\mathbf{a}_1,\cdots,\mathbf{a}_n), B=(\mathbf{b}_1,\cdots,\mathbf{b}_n)$. 再设它们对应的列向量组的极大无关组分别为 $\mathbf{a}_{i_1},\cdots,\mathbf{a}_{i_s}$ 和 $\mathbf{b}_{j_1},\cdots,\mathbf{b}_{j_t}$. 那么 $\mathbf{a}_p,\mathbf{b}_p$ 分别可表示为
$$\mathbf{a}_p=\sum_{l=1}^{s}k_l\mathbf{a}_{i_l},\quad \mathbf{b}_p=\sum_{l=1}^{t}\tilde{k}_l\mathbf{b}_{j_l}.$$
于是
$$\mathbf{a}_p+\mathbf{b}_p=\sum_{l=1}^{s}k_l\mathbf{a}_{i_l}+\sum_{l=1}^{t}\tilde{k}_l\mathbf{b}_{j_l}.$$
从而 $A+B$ 的列向量均可用 $\mathbf{a}_{i_1},\cdots,\mathbf{a}_{i_s},\mathbf{b}_{j_1},\cdots,\mathbf{b}_{j_t}$ 这 $s+t$ 个向量线性表示, 从而其向量组的秩不大于 $s+t$. 于是, $r(A+B)\leq s+t=r(A)+r(B)$.

习题 6. 证明: $r(A-B)\geq r(A)-r(B)$.

第 4 节 线性空间的基、维数和向量的坐标

这一节将利用前面两节引入的概念、方法研究 K 上的线性空间 V. 将 V 看作一个向量组, 那么这个向量组的秩 n 就叫做线性空间 V 的维数, 而其

一个极大无关组就叫做 V 的一组基. V 中向量用这组基线性表示时的表示系数就称作该向量在这组基下的坐标. 正如平面几何中引入笛卡尔坐标系一样, 借助于基, 对于 V 中抽象向量的研究就完全转化为对于 $\mathbf{K}^n$ 中向量的研究, 所以也就可以使用矩阵、线性方程组等工具了.

这里需要注意的有两点. 首先, V 是含有无限个向量的向量组 (除非 $V=\{\mathbf{0}\}$), 所以第 3 节定理 1 的证明不再适用, 故基的存在性并不能保证. 所以在线性代数课程里, 只研究有限维空间, 即已存在一组基且其所含向量个数为有限的线性空间; 对于无限维的线性空间的研究是泛函分析的内容.

其次, 在线性空间中, 基的选取是带有人为色彩的, 从而一个向量的坐标不具有客观意义 (这里特别注意, 向量是客观的量, 不依赖于参考系的选取, 而坐标则不同). 从中学解析几何我们就知道, 选取一组好的基 (坐标系), 对于解决问题可能会带来极大方便. 这就需要我们研究在不同基下同一个向量的坐标之间变换的规律.

下面的内容, 就是把上述介绍严密化.

一. 线性空间的基、维数和向量的坐标

定义 1. *设 V 是 $K(K=\mathbf{R}$ 或 $\mathbf{C})$ 上的线性空间, $\alpha_1,\cdots,\alpha_n$ 为 V 中一组线性无关的向量. 若 V 中任意一个向量 β 均可由它们线性表示: $\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n$, $k_i\in K$, 就称 $\alpha_1,\cdots,\alpha_n$ 是 V 中的* **一组基**, *n 称作 V 的* **维数**, *记作 $\dim V$. n 维向量 $(k_1,k_2,\cdots,k_n)^T$ 称作 β 在基 $\alpha_1,\cdots,\alpha_n$ 下的* **坐标**.

注记. 由于基 $\alpha_1,\cdots,\alpha_n$ 线性无关, 故一个向量在给定基下的坐标是唯一的.

例 1. *(回忆本章第 1 节例 4) 考察由 2×2 实矩阵构成的线性空间, 记作 $M^{2\times2}$. 则*

$$E_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix},E_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix},E_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix},E_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

构成 $M^{2\times2}$ 的一组基, 故 $\dim M^{2\times2}=4$. 又 $\begin{pmatrix}a&b\\c&d\end{pmatrix}$ 在基 $E_{11},E_{12},E_{21},E_{22}$ 下的坐标就是 $(a,b,c,d)^T$.

例 2. 考察所有由次数不大于 5 的实系数一元多项式，按照普通加法及数乘构成的 $\mathbf{R}$ 上的线性空间 $P[x]_5$. 则它的一组基就是 $1, x, x^2, x^3, x^4, x^5$. 于是 $\dim P[x]_5 = 6$.

证明. 我们先证明 $1, x, x^2, x^3, x^4, x^5$ 线性无关. 若 $a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 = \mathbf{0}$ 成立，就意味着等号左边这个多项式恒为零，而由代数基本定理，若 $a_i (i = 0, \cdots, 5)$ 不全为零，则该方程最多只有五个复根，从而最多只有五个实根，不可能对所有 $x \in \mathbf{R}$ 成立. 故 $a_i (i = 0, \cdots, 5)$ 全为零，即得到 $1, x, x^2, x^3, x^4, x^5$ 线性无关.

又显然对任意 $f(x) \in P[x]_5$, 总可表示为 $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5$. 故在这组基下，$f$ 的坐标就是 $(a_0, a_1, a_2, a_3, a_4, a_5)^T$.

类似地，不难验证，例如，$1, x+1, (x+2)^2, (x-8)^3, (x+10.3)^4, (x-100)^5$ 也是 $P[x]_5$ 的一组基.

例 3. *(回忆第 1 节例 3)* 由于 $C[a,b]$ 包含所有次数的一元多项式，由上一例可以知道，$C[a,b]$ 是个无限维的线性空间.

例 4. *(回忆第 1 节例 12)* 对线性空间 $W = \mathrm{span}\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$, 若 $\alpha_1, \cdots, \alpha_n$ 的一个极大无关组是 $\alpha_{i_1}, \cdots, \alpha_{i_r}$, 则 $\alpha_{i_1}, \cdots, \alpha_{i_r}$ 就是 W 的一组基，$\dim W = r$.

例 5. 记 $W = \{(x_1, x_2, x_3)^T \in \mathbf{R}^3 : x_3 = 0\}$, 它是 $\mathbf{R}^3$ 的线性子空间. 可见 $(1,0,0)^T, (0,1,0)^T$ 是 W 的一组基，$\dim W = 2$.

习题 1. 在 $\mathbf{R}^4$ 中，求向量 $\mathbf{b} = (0,0,0,1)^T$ 在基 $\mathbf{a}_1 = (1,1,0,1)^T, \mathbf{a}_2 = (2,1,3,1)^T, \mathbf{a}_3 = (1,1,0,0)^T, \mathbf{a}_4 = (0,1,-1,-1)^T$ 下的坐标.

习题 2. 求本章第 1 节习题 1 中 (2) 所定义线性空间的维数与一组基.

二. 线性空间的基变换与向量的坐标变换

定理 1. 设 $\beta_1, \cdots, \beta_n$ 和 $\alpha_1, \cdots, \alpha_n$ 是 n 维线性空间 V 的两组基. 记 β_i 在基 $\alpha_1, \cdots, \alpha_n$ 下的坐标为 $\mathbf{c}_i = (c_{1i}, c_{2i}, \cdots, c_{ni})^T, i = 1, 2, \cdots, n$. 记 n 阶方阵

$$C = (\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_n), \tag{8}$$

称为从基 $\alpha_1, \cdots, \alpha_n$ 到基 $\beta_1, \cdots, \beta_n$ 的 **过渡矩阵**, 也就是说，写作矩阵语言，

$$(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)C. \tag{9}$$

又设向量 ξ 在基 $\alpha_1,\cdots,\alpha_n$ 下坐标为 $\boldsymbol{a}=(a_1,\cdots,a_n)^T$：

$$\xi=(\alpha_1,\cdots,\alpha_n)\boldsymbol{a}, \tag{10}$$

在基 $\beta_1,\cdots,\beta_n$ 下的坐标为 $\boldsymbol{b}=(b_1,\cdots,b_n)^T$：

$$\xi=(\beta_1,\cdots,\beta_n)\boldsymbol{b}. \tag{11}$$

则

$$\boldsymbol{b}=C^{-1}\boldsymbol{a}, \quad 或 \quad \boldsymbol{a}=C\boldsymbol{b}. \tag{12}$$

证明. 将 (9) 带入 (11), 得到

$$\xi=(\alpha_1,\cdots,\alpha_n)C\mathbf{b}.$$

将此式与 (10) 对比，利用在同一基下一个向量的坐标的唯一性，就得到 (12) 的第二式. 再注意到利用 (9) 及第 3 节定理 4, 就可知 $r(C)=n$. 所以 C 可逆. 于是 (12) 的第一式成立. 证毕.

注记. 利用 (8) 的意义，即它的第 i 列就是新基中第 i 个向量 β_i 在旧基 $\alpha_1,\cdots,\alpha_n$ 下的坐标，有时可以帮助我们快速写出过渡矩阵. 此外，如果我们把 (9) 写为列向量的形式：

$$\begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix}=C^T\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}, \tag{13}$$

那么我们就会看到一个向量的新旧坐标变换规律与新旧基的变换规律是不一样的：前者是后者对应矩阵的转置的逆. 所以在张量分析和微分几何里，也把向量空间里的向量称作一阶反变张量.

例 6. 给定 $\mathbf{R}^3$ 的基

$$\boldsymbol{a}_1=\begin{pmatrix}1\\2\\1\end{pmatrix},\boldsymbol{a}_2=\begin{pmatrix}2\\3\\3\end{pmatrix},\boldsymbol{a}_3=\begin{pmatrix}3\\7\\1\end{pmatrix}$$

和

$$\boldsymbol{b}_1=\begin{pmatrix}3\\1\\4\end{pmatrix},\boldsymbol{b}_2=\begin{pmatrix}5\\2\\1\end{pmatrix},\boldsymbol{b}_3=\begin{pmatrix}1\\1\\-6\end{pmatrix}.$$

(a) 求由 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 到 $\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3$ 的过渡矩阵.

(b) 已知向量 ξ 在 $\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3$ 下的坐标为 $(1,-1,0)^T$. 求 ξ 在 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 下的坐标 $\boldsymbol{a}$.

(c) 已知向量 δ 在 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 下的坐标为 $(1,-1,0)^T$. 求 δ 在 $\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3$ 下的坐标 $\boldsymbol{b}$.

解. (a) 记矩阵 $B=(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3)$, $A=(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3)$, 则 (9) 就是 $B=AC$. 我们要求出 C, 就需要求解线性方程组 $AX=B$. 为此，对增广矩阵 (AB) 作初等行变换：

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

$$\to\begin{pmatrix}1&0&5&\vdots&-7&-11&-1\\0&1&-1&\vdots&5&8&1\\0&0&-1&\vdots&-4&-12&-8\end{pmatrix}\to\begin{pmatrix}1&0&0&\vdots&-27&-71&-41\\0&1&0&\vdots&9&20&9\\0&0&1&\vdots&4&12&8\end{pmatrix}.$$

所以 $C=\begin{pmatrix}-27&-71&-41\\9&20&9\\4&12&8\end{pmatrix}$.

(b) 由 (12) 第二式 $\mathbf{a}=C\mathbf{b}=\begin{pmatrix}-27&-71&-41\\9&20&9\\4&12&8\end{pmatrix}\begin{pmatrix}1\\-1\\0\end{pmatrix}=\begin{pmatrix}44\\-11\\-8\end{pmatrix}$.

(c) 我们需计算 C^{-1}. 为此，对 $\begin{pmatrix}-27&-71&-41&\vdots&1&0&0\\9&20&9&\vdots&0&1&0\\4&12&8&\vdots&0&0&1\end{pmatrix}$ 作初

等行变换将左边化为单位阵，得到 $\left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 13 & 19 & \frac{181}{4} \\ 0 & 1 & 0 & -9 & -13 & -\frac{63}{2} \\ 0 & 0 & 1 & 7 & 10 & \frac{99}{4} \end{array}\right)$. 从而

$$\mathbf{b} = C^{-1}\mathbf{a} = \begin{pmatrix} 13 & 19 & \frac{181}{4} \\ -9 & -13 & -\frac{63}{2} \\ 7 & 10 & \frac{99}{4} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -6 \\ 4 \\ -3 \end{pmatrix}.$$

例 7. 考虑由所有二阶实上三角方阵 (即形如 $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ 的矩阵) 按矩阵加法和数乘组成的 $\mathbf{R}$ 上的线性空间 V. 显然

$$(\mathrm{I}):\ E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

是它的一组基.

(a) 证明：

$$(\mathrm{II}):\ M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, M_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, M_3 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

也是一组基，并写出由 *(I)* 到 *(II)* 的过渡矩阵.

(b) 求 $M = \begin{pmatrix} 2 & -1 \\ 0 & -3 \end{pmatrix}$ 在两组基下的坐标.

解. (a) 利用 (II) 向量在 (I) 下的坐标, 容易写出过渡矩阵 $C = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. 由于 $r(C) = 3$, 由第 3 节定理 4 可知 (II) 中向量线性无关. 又 $\dim V = 3$, 故由第 3 节定理 2 知 (II) 也是一组基.

(b) 显然 M 在 (I) 下的坐标为 $\mathbf{a} = (2, -1, -3)^T$. 故它在 (II) 下的坐标为

$$\mathbf{b} = C^{-1}(2, -1, 3)^T = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}.$$

习题 3. (1) 证明：$(1,1,0,)^T, (2,1,3,0)^T, (1,1,0,0)^T, (0,1,-1,-1)^T$ 和 $(1,0,0,1)^T, (0,0,1,-1)^T, (2,1,0,3)^T, (-1,0,1,2)^T$ 都是 $\mathbf{R}^4$ 的基;

(2) 求从前者到后者的过渡矩阵.

第 5 节　线性方程组解集的结构

在本章第 1 节我们已经证明了齐次线性方程组

$$A\mathbf{x}=\mathbf{0} \tag{14}$$

(A 是 $m\times n$ 矩阵，$\mathbf{x}\in\mathbf{K}^n$) 的解集 $Ker(A)$ 是 $\mathbf{K}^n$ 的线性子空间. 第 3 章第 6 节的 (5) 式给出了这个集合元素的一个表达式 (其中取 $d_1=d_2=\cdots=d_r=0$). 本节就是利用前面所学线性空间的理论对 $Ker(A)$ 加以研究，目的是确定其维数，并找到一组基. 然后，我们再研究非齐次方程组

$$A\mathbf{x}=\mathbf{b} \tag{15}$$

的解集 $\mathscr{A}$. 我们将发现，$\mathscr{A}$ 其实就是 $Ker(A)$ 再加一个平移.

定理 1. 设 $r(A)=r$, 则 $\dim Ker(A)=n-r$.

证明. 若 $r=n$, 则 (14) 只有零解，$Ker(A)=\{\mathbf{0}\}$, 故 $\dim Ker(A)=0$ 成立.

若 $r<n$, 让我们回到第 3 章第 6 节的 (5) 式 (由于现在只研究齐次方程组，故 $d_1=d_2=\cdots=d_r=0$):

$$\begin{cases} x_1=-c_{1,r+1}x_{r+1}-\cdots-c_{1n}x_n, \\ x_2=-c_{2,r+1}x_{r+1}-\cdots-c_{2n}x_n, \\ \cdots\cdots \\ x_r=-c_{r,r+1}x_{r+1}-\cdots-c_{rn}x_n. \end{cases} \tag{16}$$

我们只需要找出 $Ker(A)$ 的一组基即可.

为此，分别取 $(x_{r+1},\cdots,x_n)^T$ 为 $\mathbf{K}^{n-r}$ 的单位向量 (回忆第 2 节例 1):

$$\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\quad \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\quad \cdots,\quad \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}.$$

那么由 (16) 可写出如下 $n-r$ 个 (14) 的解，它们也被称作 (14) 的一个 **基础解系**:

$$\mathbf{x}_1=\begin{pmatrix}-c_{1,r+1}\\-c_{2,r+1}\\\vdots\\-c_{r,r+1}\\1\\0\\\vdots\\0\end{pmatrix},\mathbf{x}_2=\begin{pmatrix}-c_{1,r+2}\\-c_{2,r+2}\\\vdots\\-c_{r,r+2}\\0\\1\\\vdots\\0\end{pmatrix},\cdots,\mathbf{x}_{n-r}=\begin{pmatrix}-c_{1,n}\\-c_{2,n}\\\vdots\\-c_{r,n}\\0\\0\\\vdots\\1\end{pmatrix}.$$

我们证明 $\mathbf{x}_1,\cdots,\mathbf{x}_{n-r}$ 就是 $Ker(A)$ 的一组基.

首先说明它们是线性无关的. 由于矩阵 $(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_{n-r})$ 最后 $n-r$ 行其实就是一个 $n-r$ 阶单位阵, 故其秩就是 $n-r$. 于是齐次线性方程组 $\sum_{j=1}^{n-r}k_j\mathbf{x}_j=\mathbf{0}$ 只有零解 $k_j=0, j=1,\cdots,n-r$. 这就表明 $\mathbf{x}_1,\cdots,\mathbf{x}_{n-r}$ 线性无关.

再说明 (14) 的任意一个解均可由 $\mathbf{x}_1,\cdots,\mathbf{x}_{n-r}$ 线性表示. 事实上，由 (16),(14) 的任何一个解，写作列向量的形式，就是

$$\mathbf{x}=\begin{pmatrix}-c_{1,r+1}x_{r+1}-\cdots-c_{1n}x_n\\-c_{2,r+1}x_{r+1}-\cdots-c_{2n}x_n\\\vdots\\-c_{r,r+1}x_{r+1}-\cdots-c_{rn}x_n\\x_{r+1}\\x_{r+2}\\\vdots\\x_n\end{pmatrix},$$

其中 $x_{r+1},\cdots,x_n\in K$ 是自由未知量. 显然 $\mathbf{x}$ 可以分解为

$$\mathbf{x}=\sum_{j=1}^{n-r}x_{r+j}\mathbf{x}_j.$$

证毕.

推论 1. 设 A 为 $m\times n$ 矩阵，则 $r(A)+\dim Ker(A)=n$.

下面我们看一个具体的算例.

例 1. 求 $\begin{cases}3x_1+5x_2+6x_3-4x_4=0,\\ x_1+2x_2+4x_3-3x_4=0,\\ 4x_1+5x_2-2x_3+3x_4=0,\\ 3x_1+8x_2+29x_3-19x_4=0\end{cases}$ 的基础解系与通解.

解. 对系数矩阵 A 作行初等变换化为简化的行阶梯标准形得到 $\begin{pmatrix}1&0&-8&7\\0&1&6&-5\\0&0&0&0\\0&0&0&0\end{pmatrix}$. 从而可以写出 $\mathbf{x}_1=\begin{pmatrix}8\\-6\\1\\0\end{pmatrix}$, $\mathbf{x}_2=\begin{pmatrix}-7\\5\\0\\1\end{pmatrix}$. 于是通解就是 $k_1\mathbf{x}_1+k_2\mathbf{x}_2$.

例 2. 设有齐次方程组 $(\mathrm{I}):\begin{cases}x_1+x_2=0,\\ x_2-x_4=0,\end{cases}$ $(\mathrm{II}):\begin{cases}x_1-x_2+x_3=0,\\ x_2-x_3+x_4=0.\end{cases}$ 求它们的公共解.

解. 由于

$$\begin{pmatrix}1&1&0&0\\0&1&0&-1\end{pmatrix}\to\begin{pmatrix}1&0&0&1\\0&1&0&-1\end{pmatrix},$$

故 (I) 的一个基础解系就是 $\mathbf{x}_1=\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$, $\mathbf{x}_2=\begin{pmatrix}-1\\1\\0\\1\end{pmatrix}$. 现设 $\mathbf{x}=k_1\mathbf{x}_1+k_2\mathbf{x}_2$ 是公共解，将之代入 (II), 得到线性方程组

$$\begin{cases}k_1-2k_2=0,\\ -k_1+2k_2=0,\end{cases}$$

解得 $k_1=2k_2$. 从而公共解就是 $k(2\mathbf{x}_1+\mathbf{x}_2), k\in\mathbf{R}$.

注记. 另外一种求线性方程组 $A\mathbf{x}=\mathbf{0}$, $\mathbf{B}\mathbf{x}=\mathbf{0}$ 的公共解的办法就是考察方程组 $\begin{pmatrix} A \\ B \end{pmatrix}\mathbf{x}=\mathbf{0}$, 即通过对分块矩阵 $\begin{pmatrix} A \\ B \end{pmatrix}$ 作行初等变换求解.

例 3. 设 A 为 $m\times n$ 矩阵, B 为 $n\times p$ 矩阵, 且 $AB=O_{m\times p}$. 证明: $r(A)+r(B)\le n$.

证明. 将 B 按列向量写作分块矩阵 $B=(\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_p)$, 则成立 $A\mathbf{b}_j=\mathbf{0}$. 由于 $r(B)=r[\mathbf{b}_1,\mathbf{b}_2,\cdots,\mathbf{b}_p]\le \dim Ker(A)=n-r(A)$, 即得 $r(A)+r(B)\le n$.

例 4. A 为 n 阶方阵, $n\ge 2$, A^* 为 A 的伴随矩阵. 证明:

$$r(A^*)=\begin{cases} n, & r(A)=n, \\ 1, & r(A)=n-1, \\ 0, & r(A)\le n-2. \end{cases}$$

证明. 当 $r(A)=n$ 时 $|A|\ne 0$,A 可逆, 所以 $A^*=|A|A^{-1}$ 也可逆, 故 $r(A^*)=n$.

当 $r(A)\le n-2$ 时, 由矩阵秩的定义, A 的任意 $n-1$ 阶子式均为零. 故由伴随矩阵定义, $A^*=O$, 从而 $r(A^*)=0$.

当 $r(A)=n-1$ 时, 由于 $|A|=0$, 用第 2 章行列式按行列展开性质可得 $AA^*=O$. 于是利用例 3 结论, $r(A)+r(A^*)\le n$, 于是 $r(A^*)\le 1$. 若 $r(A^*)=0$, 则 $A^*=O$. 但是由矩阵秩定义, 此时 A 中至少有一个 $n-1$ 阶子式不为零. 故 A^* 中至少有一个元素不为零, 矛盾. 所以此时 $r(A^*)=1$. 证毕.

习题 1. 设 A 是 n 阶方阵且 $A^2=A$. 证明: $r(A)+r(A-E_{n\times n})=n$.

习题 2. 写出下列齐次线性方程组的一个基础解系.

$$(1)\begin{cases} x_1+x_2+x_3+x_4+x_5=0, \\ 3x_1+2x_2+x_3+x_4-3x_5=0, \\ x_2+2x_3+2x_4+6x_5=0, \\ 5x_1+4x_2+3x_3+3x_4-x_5=0; \end{cases}$$

$$(2)\begin{cases} x_1-2x_2+x_3+x_4-x_5=0, \\ 2x_1+x_2-x_3-x_4-x_5=0, \\ x_1+7x_2-5x_3-5x_4+5x_5=0, \\ 3x_1-x_2-2x_3+x_4-x_5=0. \end{cases}$$

下面我们再看非齐次方程组 (15) 的解集 $\mathscr{A}$, 此时 $\mathbf{b}\ne\mathbf{0}$.

定理 2. *设 $\boldsymbol{y}$ 为非齐次线性方程组 (15) 的一个解. 则它的任意解都可唯一地表示为 $\sum_{j=1}^{n-r} k_j \boldsymbol{x}_j + \boldsymbol{y}$, 其中 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{n-r}$ 是对应齐次线性方程组 (14) 的基础解系, $r = r(A)$.*

我们可把该结论简记为 $\mathscr{A} = Ker(A) + \mathbf{y}$. 所以，只要搞清楚了齐次方程的解集，再找到一个特解，那么非齐次方程组的解集也就清楚了. 我们特别注意，$\mathscr{A}$ 不是一个线性空间.

证明. 显然，形如 $\sum_{j=1}^{n-r} k_j \mathbf{x}_j + \mathbf{y}$ 的向量都是 (15) 的解. 反之，设 $\mathbf{z}$ 是 (15) 的解，则 $A(\mathbf{z}-\mathbf{y}) = A\mathbf{z} - A\mathbf{y} = \mathbf{b} - \mathbf{b} = \mathbf{0}$. 所以 $\mathbf{z}-\mathbf{y} \in Ker(A)$, 从而 $\mathbf{z}-\mathbf{y}$ 可以唯一地表为 $\sum_{j=1}^{n-r} k_j \mathbf{x}_j$, 这就是说，$\mathbf{z}$ 可唯一地写为 $\sum_{j=1}^{n-r} k_j \mathbf{x}_j + \mathbf{y}$ 的形式. 证毕.

那么，怎么寻找特解呢？回忆第 3 章第 6 节关于通解的 (5) 式：

$$\begin{cases} x_1 = -c_{1,r+1}x_{r+1} - \cdots - c_{1n}x_n + d_1, \\ x_2 = -c_{2,r+1}x_{r+1} - \cdots - c_{2n}x_n + d_2, \\ \cdots\cdots \\ x_r = -c_{r,r+1}x_{r+1} - \cdots - c_{rn}x_n + d_r. \end{cases}$$

取 $x_{r+1} = \cdots = x_n = 0$, 就可以得到一个特解 $\mathbf{y} = (d_1, \cdots, d_r, 0, \cdots, 0)^T$. 这里 $d_1, \cdots, d_r$ 也就依次是用行初等变换把增广矩阵 $(A \quad b)$ 化为简化的行阶梯标准形矩阵后的最后一列的非零元素.

例 5. *求解* $\begin{cases} x_1 + x_2 + 2x_3 + x_4 + 2x_5 = 7, \\ x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 15, \\ 2x_1 + 3x_2 + 5x_3 + 5x_4 + 7x_5 = 22. \end{cases}$

解. 对增广矩阵作行初等变换化为简化的行阶梯标准形，得到

$$\begin{pmatrix} 1 & 0 & 1 & -2 & -1 & \vdots & -1 \\ 0 & 1 & 1 & 3 & 3 & \vdots & 8 \\ 0 & 0 & 0 & 0 & 0 & \vdots & 0 \end{pmatrix}.$$

最后一列就是一个特解： $\mathbf{y} = (-1, 8, 0, 0, 0)^T$. 只看前面 5 列，就可写出齐次

方程组的基础解系：

$$\mathbf{x}_1=\begin{pmatrix}-1\\-1\\1\\0\\0\end{pmatrix},\mathbf{x}_2=\begin{pmatrix}2\\-3\\0\\1\\0\end{pmatrix},\mathbf{x}_3=\begin{pmatrix}1\\-3\\0\\0\\1\end{pmatrix}.$$

于是可以写出通解 $\mathbf{x}=k_1\mathbf{x}_1+k_2\mathbf{x}_2+k_3\mathbf{x}_3+\mathbf{y}$, $k_1,k_2,k_3\in\mathbf{R}$.

复习题

1. 求向量组 $(1,0,-1,3,-2)^T,(2,1,0,-1,0)^T,(3,1,-1,2,-2)^T$ 的秩.

2. 设 V 是所有次数小于 n 的实系数一元多项式构成的实数集上的线性空间. 又设 $a_1,a_2,\cdots,a_n$ 是 n 个互不相同的实数. 记 $f(x)=(x-a_1)(x-a_2)\cdots(x-a_n)$, $f_i(x)=\frac{f(x)}{x-a_i}$. 证明： $f_1,\cdots,f_n$ 构成 V 的一组基.

3. 设 n 阶方阵 A 满足 $A^2=E$. 证明： $r(A+E)+r(A-E)=n$.

4. 设 A,B 是 n 阶方阵且方程组 $A\mathbf{x}=\mathbf{0}$ 和 $B\mathbf{x}=\mathbf{0}$ 同解且基础解系含有 m 个向量. 证明： $r(A-B)\le n-m$.

5. 设 A 是 $m\times n$ 矩阵， $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_m$ 依次是 A 的行向量， $\mathbf{b}=(b_1,b_2,\cdots,b_n)$. 求证：若线性方程组 $A\mathbf{x}=\mathbf{0}$ 的解全是方程 $b_1x_1+b_2x_2+\cdots+b_nx_n=0$ 的解，则 $\mathbf{b}$ 是 $\mathbf{a}_1,\cdots,\mathbf{a}_m$ 的线性组合.

6. 设 $A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}(m<n.)$ 若 $A\mathbf{x}=\mathbf{0}$ 的基础解系是 $\mathbf{b}_i=(b_{i1},b_{i2},\cdots,b_{in})^T$, $(i=1,2,\cdots,n-m.)$ 试求齐次线性方程组

$$\sum_{j=1}^{n}b_{ij}y_j=0,\quad i=1,\cdots,n-m$$

的基础解系.

7. 设 W 是 $\mathbf{R}^n$ 的真子空间. 证明：存在 $n\times n$ 矩阵 A 使得 $W=Ker(A)$.

8. 设 A 是一个 n 阶方阵，A 的第 $i_1,\cdots,i_r$ 行和第 $i_1,\cdots,i_r$ 列的交点上的元素按原来顺序组成的子式称为 A 的一个 r 阶主子式. 若 A 是对称阵或反对称阵，且秩为 r, 证明：A 必有一个 r 阶子式不为零.

9. 证明：(1) 奇数阶反对称方阵必不可逆；(2) 反对称矩阵的秩必为偶数.

10. 设 A 是 $m\times n$ 矩阵且 $r(A)=n$. 证明：存在秩为 n 的 $n\times m$ 矩阵 B 使得 $BA=E_n$.

11. 考察方程组 $A\mathbf{x}=\mathbf{0}$, 其中 A 是 n 阶方阵，$|A|=0$. 若 A 中有某个元素 a_{ij} 的代数余子式 $\mathscr{A}_{ij}\neq 0$. 证明：该方程组的所有解都可表为

$$\mathbf{x}=k(\mathscr{A}_{i1},\mathscr{A}_{i2},\cdots,\mathscr{A}_{in})^T.$$

第 5 章　欧 氏 空 间

在这一章里, 我们研究关于线性方程组的最后两个问题, 即 (a) 使得 $A\mathbf{x}=\mathbf{b}$ 有解的 $\mathbf{b}$ 构成的集合 $R(A)$ 的刻画，以及 (b) 当 $A\mathbf{x}=\mathbf{b}$ 无解时如何定义 “广义解 ”— 我们将介绍最小二乘法．解决这两个问题都要用到欧式空间的概念.

第 1 节　欧式空间

距离和方向是我们日常生活常用的概念．在线性代数中，这对应于向量的长度和向量间的夹角．可是在线性空间里，向量间只有线性关系，其长度和夹角都是没有定义的．为了对现实世界建立更精确的模型，人们引入了欧式空间的概念：它就是两个向量之间又定义了一种叫做内积的运算并且满足若干公理的线性空间．这里的内积，也就是解析几何里两个 3 维向量内积概念的推广．有了内积就可以定义任一个向量的长度及任意两个向量间的夹角，也就有了两个向量是否垂直 (正交) 的概念．于是我们就可以如解析几何那样，建立直角坐标系 — 在线性代数中，被称作标准正交基 — 了.

有了距离，就可以比较大小远近．这就为介绍用最小二乘法定义无解方程的广义解 — 某种意义下的最优解 — 作了必要的准备．特别地，这使得我们对 “最优 ” 有了一个数学上的可以量化的标准．利用垂直的概念，还可以帮我们方便地刻画使得给定线性方程组有解的非齐次项构成的集合.

一．内积与欧式空间

定义 1. 设 V 是一个实线性空间．在 V 中定义了一种运算，称为 **内积**: 对任意 $\alpha,\beta\in V$, 有唯一的实数与它们对应，这个实数也被叫做 α 与 β 的内积, 记作 (α,β), 它要满足如下公理：

(a) $(\alpha,\beta)=(\beta,\alpha)$;

(b) $(k\alpha,\beta)=k(\alpha,\beta)$;

(c) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$;

(d) $(\alpha,\alpha)\geq 0$, 且 $(\alpha,\alpha)=0$ 当且仅当 $\alpha=\mathbf{0}$.

以上 α,β,γ 为 V 中任意向量，k 为任意实数．定义了内积的实线性空间叫做 **欧几里得空间**, 简称欧式空间.

在欧式空间 V 里，定义向量 α 的 **长度**$|\alpha|$ 为 $\sqrt{(\alpha,\alpha)}$. 长度为 1 的向量称作 **单位向量**. 定义向量 α,β 的 **距离** 就是 $|\alpha-\beta|$. 定义两个非零向量 α,β 的 **夹角**$<\alpha,\beta>$ 为

$$<\alpha,\beta>=\arccos\frac{(\alpha,\beta)}{|\alpha||\beta|}.$$

若 $(\alpha,\beta)=0$, 就称 α 与 β**正交**. 特别地，零向量与任何向量均正交.

可以看出内积的定义是由解析几何的内积抽象化用公理方法定义的，长度与夹角的定义也与解析几何相关公式一样．但是由于这里内积定义的抽象性，我们必须说明上述长度及夹角定义的合理性．由内积的公理 (d), 向量长度是可以定义的 (可以开方). 此外不难看出 $|\alpha|=0$ 当且仅当 $\alpha=\mathbf{0}$, 且 $|k\alpha|=|k||\alpha|$ $(k\in\mathbf{R})$. 但是由于总是成立 $|\cos<\alpha,\beta>|\leq 1$, 所以为了使夹角定义合理，必须成立如下不等式

$$|(\alpha,\beta)|\leq|\alpha||\beta|, \tag{1}$$

其中等号成立当且仅当 α,β 线性相关.

(1) 就是 **柯西 — 施瓦茨不等式**．它可由内积的公理推出．事实上，若 α,β 线性相关, 则成立 $\alpha=k\beta$ 或 $\beta=l\alpha$. 不妨考虑第一种情形, 则 $|(\alpha,\beta)|=|(k\beta,\beta)|=|k||\beta|^2=(|k||\beta|)(|\beta|)=|\alpha||\beta|$.

再考虑 α,β 线性无关的情形. 此时 $\alpha\neq\mathbf{0}$. 对任意实数 t, 则向量 $t\alpha-\beta\neq\mathbf{0}$. 于是由公理 (a)—(d), 成立

$$0<(t\alpha-\beta,t\alpha-\beta)=t^2(\alpha,\alpha)-2t(\alpha,\beta)+(\beta,\beta).$$

由于上式右边关于 t 的一元二次函数恒正, 其判别式必严格小于零, 即 $(\alpha,\beta)^2<(\alpha,\alpha)(\beta,\beta)$, 两边开方就得到 (1). 所以向量夹角的定义也是合理的.

下面看一些例子.

例 1. 在 $\mathbf{R}^n$ 中，对 $\boldsymbol{x}=(x_1,\cdots,x_n)^T,\boldsymbol{y}=(y_1,\cdots,y_n)^T$, 定义内积

$$(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_iy_i=\boldsymbol{y}^T\boldsymbol{x}.$$

注意 $\boldsymbol{y}^T\boldsymbol{x}$ 是矩阵乘法. 很容易验证公理 (a)—(d) 成立, 所以这确实是内积. 以后, 在 $\mathbf{R}^n$ 中, 内积总是约定如此定义. 显然这就是我们解析几何中所熟悉的内积. 于是向量 $\boldsymbol{x}$ 的长度就是 $|\boldsymbol{x}| = \sqrt{\sum_{i=1}^n x_i^2}$, 而这种情形下的柯西 — 施瓦茨不等式就是

$$|\sum_{i=1}^n x_i y_i| \le \sqrt{(\sum_{i=1}^n x_i^2)(\sum_{i=1}^n y_i^2)}. \tag{2}$$

例 2. 在实线性空间 $C[a,b]$ 中, 对任意函数 $f(x), g(x) \in C[a,b]$, 定义

$$(f,g) = \int_a^b f(x)g(x)\,dx.$$

则由积分的性质, 不难验证公理 (a)—(d) 成立. 于是这就给出了 $C[a,b]$ 上的一个内积. 相应的向量 $f(x)$ 的长度就是 $|f| = \sqrt{\int_a^b |f(x)|^2\,dx}$, 柯西 — 施瓦茨不等式就是

$$|\int_a^b f(x)g(x)\,dx| \le \sqrt{(\int_a^b |f(x)|^2\,dx)(\int_a^b |g(x)|^2\,dx)}. \tag{3}$$

看着形式颇不相同的不等式 (2)(3), 我们是不是得佩服抽象的力量?

例 3. 在 $\mathbf{R}^4$ 中, 寻求一单位向量 $\boldsymbol{x}$ 使之与以下向量均正交:

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix},\ \boldsymbol{a}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix},\ \boldsymbol{a}_3 = \begin{pmatrix} 2 \\ 1 \\ 3 \\ 1 \end{pmatrix}.$$

解. 首先求解齐次线性方程组 $A\mathbf{x} = \mathbf{0}$, 其中

$$A = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 2 & 1 & 3 & 1 \end{pmatrix}.$$

利用行初等变换可得 $\mathbf{x} = (4x_4, 0, -3x_4, x_4)^T$, 其中 x_4 为自由未知量. 由于 $|\mathbf{x}| = 1$, 所以要求 $26x_4^2 = 1$, 从而 $k_4 = \pm\sqrt{\frac{1}{26}}$, 得到 $\mathbf{x} = (\pm 4\sqrt{\frac{1}{26}}, 0, \mp 3\sqrt{\frac{1}{26}}, \pm\sqrt{\frac{1}{26}})^T$ 即所求.

习题 1. 设线性空间 V 上可以定义两个不同的内积 $(\mathbf{x},\mathbf{y})_1,(\mathbf{x},\mathbf{y})_2$，证明：$(\mathbf{x},\mathbf{y})=(\mathbf{x},\mathbf{y})_1+(\mathbf{x},\mathbf{y})_2$ 也是 V 上的内积，并写出相应的柯西 — 施瓦茨不等式.

有了内积之后，我们就可以定义欧式空间的标准正交基，即直角坐标系了.

定义 2. 若欧式空间 V 中一组非零向量两两互相正交，就称这是一个 **正交向量组**. 若 $\varepsilon_1,\cdots,\varepsilon_n$ 是 V 的一组基，而且也是一个正交向量组，就称之为 V 的一组 **正交基**. 若进一步 $\varepsilon_i(i=1,\cdots,n)$ 还是单位向量，就称 $\varepsilon_1,\cdots,\varepsilon_n$ 是 V 的一组 **标准正交基**.

例 4. 设 $\alpha_1,\cdots,\alpha_m$ 是欧式空间的一组正交向量，则它们必线性无关.

证明. 设 $\sum_{i=1}^m l_i\alpha_i=\mathbf{0}$, 两边与 α_k (k 分别取 $1,2,\cdots,m$) 作内积，有

$$\sum_{i=1}^m l_i(\alpha_i,\alpha_k)=0.$$

利用 $\alpha_1,\cdots,\alpha_m$ 两两互相正交，则 $(\alpha_i,\alpha_k)=\delta_{ik}|\alpha_i|^2$, 其中 δ_{ij} 是第 2 章引入的 Kronecker 记号. 由于 $\alpha_i\neq\mathbf{0}$, 于是得到 $l_k=0$. 这就证明了 $\alpha_1,\cdots,\alpha_m$ 线性无关.

例 5. 设 $\alpha_1,\cdots,\alpha_n$ 是欧式空间 V 的一组标准正交基. 设 V 中向量 ξ,η 在此基下的坐标分别为 $(x_1,\cdots,x_n)^T$ 和 $(y_1,\cdots,y_n)^T$. 则 $(\xi,\eta)=\sum_{i=1}^n x_iy_i$. 特别地，$|\xi|=\sqrt{\sum_{i=1}^n x_i^2}$. 又坐标分量可用内积表示： $x_i=(\xi,\alpha_i), i=1,\cdots,n$.

证明. 设 $\xi=\sum_{i=1}^n x_i\alpha_i$, 两边与 α_l ($l=1,\cdots,n$) 作内积，就得到 $(\xi,\alpha_l)=\sum_{i=1}^n x_i\delta_{il}=x_l$, 于是 $x_l=(\xi,\alpha_l)$. 又

$$\begin{aligned}(\xi,\eta) &= (\sum_{i=1}^n x_i\alpha_i,\sum_{j=1}^n y_j\alpha_j)\\ &= \sum_{i=1}^n\sum_{j=1}^n x_iy_j(\alpha_i,\alpha_j)\\ &= \sum_{i=1}^n\sum_{j=1}^n x_iy_j\delta_{ij}\\ &= \sum_{i=1}^n x_iy_i.\end{aligned}$$

证毕.

这个结果表明，在欧式空间中，无论是怎样定义的内积，如果选定了一组关于这个内积的标准正交基，则两个抽象向量的内积的计算就可以化为其坐标作为 n 维向量所定义的标准内积 (回忆例 1) 的计算. 所以在欧式空间中找到一组标准正交基对于内积、向量长度及向量夹角的计算都会带来很大方便.

二. 向量组的标准正交化

那么一个自然的问题就是欧式空间中标准正交基的存在性及其构造方法. 下面介绍的格莱姆 — 施密特正交化法解决了这个问题.

定理 1. *设 V 是内积空间， $\alpha_1,\cdots,\alpha_m \in V$ 是 m 个线性无关的向量. 则在 V 中存在正交向量组 $\beta_1,\cdots,\beta_m$, 它与 $\alpha_1,\cdots,\alpha_m$ 是等价的 (即可以互相表示).*

证明. 下面构造 $\beta_1,\cdots,\beta_m$ 的算法叫做格莱姆 — 施密特正交化法.

首先取 $\beta_1=\alpha_1$.

现假设 $\beta_1,\beta_2,\cdots,\beta_k (1\le k\le m-1)$ 已经定义, 它们互相正交, 且与 $\alpha_1,\cdots,\alpha_k$ 等价. 我们定义 β_{k+1} 如下:

$$\beta_{k+1}=\alpha_{k+1}-\sum_{j=1}^{k}\frac{(\alpha_{k+1},\beta_j)}{|\beta_j|^2}\beta_j, \tag{4}$$

则由归纳假设，显然 β_{k+1} 可由 $\alpha_1,\cdots,\alpha_{k+1}$ 线性表示；反之， α_{k+1} 也可以由 $\beta_1,\beta_2,\cdots,\beta_{k+1}$ 线性表示. 再对 $l=1,2,\cdots,k$ 计算

$$\begin{aligned}(\beta_{k+1},\beta_l) &= (\alpha_{k+1},\beta_l)-\sum_{j=1}^{k}\frac{(\alpha_{k+1},\beta_j)}{|\beta_j|^2}(\beta_j,\beta_l)\\ &= (\alpha_{k+1},\beta_l)-\sum_{j=1}^{k}\frac{(\alpha_{k+1},\beta_j)}{|\beta_j|^2}(\delta_{jl}|\beta_l|^2)\\ &= (\alpha_{k+1},\beta_l)-(\alpha_{k+1},\beta_l)=0.\end{aligned}$$

所以 $\beta_1,\cdots,\beta_k,\beta_{k+1}$ 仍然互相正交. 这样就用归纳法证明了定理.

注记. 显然, 得到了正交向量组 $\beta_1,\cdots,\beta_m$ 之后, 令 $\varepsilon_i=\beta_i/|\beta_i|, i=1,\cdots,m$, 就得到标准正交向量组. 如果 $\alpha_1,\cdots,\alpha_m$ 是一组基，那么 $\varepsilon_1,\cdots,\varepsilon_m$ 就是一组标准正交基. 下面我们看一个具体的算例. 大家可以进一步通过这个例子理解格莱姆 — 施密特正交化法.

例 6. 将 $\boldsymbol{x}_1=\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \quad \boldsymbol{x}_2=\begin{pmatrix}0\\1\\1\\0\end{pmatrix}, \quad \boldsymbol{x}_3=\begin{pmatrix}1\\0\\1\\1\end{pmatrix}$ 标准正交化.

解. 首先检验 $r(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)=3$, 故它们线性无关.

令 $\mathbf{y}_1=\mathbf{x}_1$. 则 $|\mathbf{y}_1|=\sqrt{2}$. 作

$$\begin{aligned}\mathbf{y}_2 &= \mathbf{x}_2-\frac{(\mathbf{x}_2,\mathbf{y}_1)}{|\mathbf{y}_1|^2}\mathbf{y}_1\\ &= \begin{pmatrix}0\\1\\1\\0\end{pmatrix}-\frac{1}{2}\begin{pmatrix}1\\1\\0\\0\end{pmatrix}=\begin{pmatrix}-\frac{1}{2}\\\frac{1}{2}\\1\\0\end{pmatrix},\end{aligned}$$

所以 $|\mathbf{y}_2|=\sqrt{\frac{3}{2}}$. 再作

$$\begin{aligned}\mathbf{y}_3 &= \mathbf{x}_3-\frac{(\mathbf{x}_3,\mathbf{y}_1)}{|\mathbf{y}_1|^2}\mathbf{y}_1-\frac{(\mathbf{x}_3,\mathbf{y}_2)}{|\mathbf{y}_2|^2}\mathbf{y}_2\\ &= \begin{pmatrix}1\\0\\1\\1\end{pmatrix}-\frac{1}{2}\begin{pmatrix}1\\1\\0\\0\end{pmatrix}-\frac{\frac{1}{2}}{\frac{3}{2}}\begin{pmatrix}-\frac{1}{2}\\\frac{1}{2}\\1\\0\end{pmatrix}=\begin{pmatrix}\frac{2}{3}\\-\frac{2}{3}\\\frac{2}{3}\\1\end{pmatrix},\end{aligned}$$

于是 $|\mathbf{y}_3|=\sqrt{\frac{7}{3}}$.

故所求标准正交向量组为 $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \quad \frac{1}{\sqrt{6}}\begin{pmatrix}-1\\1\\2\\0\end{pmatrix}, \quad \frac{1}{\sqrt{21}}\begin{pmatrix}2\\-2\\2\\3\end{pmatrix}$.

习题 2. 将如下向量标准正交化： $(1,1,0,0)^T,(1,0,1,0)^T,(-1,0,0,1)^T$.

例 7. 设 V 是一个欧式空间， $\xi\in V$, S 是 V 的一个子集. 若 ξ 和 S 中任意一个向量均正交，就称 ξ 与 S 正交，记作 $\xi\perp S$. 证明：若 ξ 与 $\alpha_1,\cdots,\alpha_m\in V$ 正交，则 $\xi\perp W=\mathrm{span}\{\alpha_1,\cdots,\alpha_m\}$. *(回忆第 4 章第 1 节例 12.)* 特别地，若还有 $\xi\in W$, 则 $\xi=\mathbf{0}$.

证明. 设 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 是 $\alpha_1,\cdots,\alpha_m$ 的一个极大无关组. 它也就是 W 的一组基. 由定理 1, 可以由之构造出与其等价的 W 的标准正交基 $\beta_1,\cdots,\beta_r$. 由于 β_l 可以由 $\alpha_{i_1},\cdots,\alpha_{i_r}$ 线性表示, 且 $\xi\perp\alpha_i, i=1,\cdots,m$, 故成立 $\xi\perp\beta_j, j=1,\cdots,r$. 于是, 对于任意 $\eta\in W$, η 可表为 $\eta=\sum_{i=1}^r y_i\beta_i$, 从而 $(\xi,\eta)=\sum_{i=1}^r y_i(\xi,\beta_i)=0$.

若 $\xi\in W$, 视 W 为一个 r 维欧式空间，则 ξ 在 $\beta_1,\cdots,\beta_r$ 下坐标的第 l 个坐标分量 $y_l=(\xi,\beta_l)=0, l=1,\cdots,r$. 所以 ξ 的坐标为 $(0,0,\cdots,0)^T\in\mathbf{R}^r$. 于是 $\xi=\mathbf{0}$. 证毕.

在本节最后，作为例 8 中概念的一个推广，我们给出如下定义.

定义 3. 设 S 是欧式空间 V 的一个非空子集. V 中所有与 S 垂直的向量构成的集合，叫做 S 的正交补空间，记作 $S^\perp$.

注意 $S^\perp$ 不是空集，因为总有 $\mathbf{0}\in\mathbf{S}^\perp$.

习题 3. 证明 $S^\perp$ 是 V 的一个线性子空间.

例 8. 设 S 是 n 维欧式空间 V 的线性子空间，且 $\dim S=r$, 则 $\dim S^\perp=n-r$.

证明. 设 $\alpha_1,\cdots,\alpha_r$ 是 S 的一组基. 由于 $S^\perp$ 也是 V 的线性子空间，且 $\dim V=n<\infty$, 可以找到 $S^\perp$ 的一组基 $\beta_1,\cdots,\beta_p$. (参考第 4 章第 3 节定理 1 的证明.)

对 $\alpha_1,\cdots,\alpha_r$ 作格莱姆 – 施密特正交化，就可得到 S 的一组标准正交基 $\varepsilon_1,\cdots,\varepsilon_r$. 对 $\beta_1,\cdots,\beta_p$ 格莱姆 – 施密特正交化，就可得到 $S^\perp$ 的一组标准正交基 $\zeta_1,\cdots,\zeta_p$. 下面证明 $\varepsilon_1,\cdots,\varepsilon_r,\zeta_1,\cdots,\zeta_p$ 就是 V 的一组标准正交基.

首先证明 $\varepsilon_1,\cdots,\varepsilon_r,\zeta_1,\cdots,\zeta_p$ 线性无关. 设

$$k_1\varepsilon_1+\cdots+k_r\varepsilon_r+l_1\zeta_1+\cdots+l_p\zeta_p=\mathbf{0}. \tag{5}$$

该等式两边同时与 $k_1\varepsilon_1+\cdots+k_r\varepsilon_r$ 作内积，由于 $\zeta_1,\cdots,\zeta_p$ 均与 S 正交. 就得到

$$(k_1\varepsilon_1+\cdots+k_r\varepsilon_r,\ k_1\varepsilon_1+\cdots+k_r\varepsilon_r)=0.$$

由内积公理，可得 $k_1\varepsilon_1+\cdots+k_r\varepsilon_r=0$. 由于 $\varepsilon_1,\cdots,\varepsilon_r$ 是基，从而 $k_1=\cdots=k_r=0$. 类似地，通过在 (5) 两边同时与 $l_1\zeta_1+\cdots+l_p\zeta_p$ 作内积就可以证明 $l_1=\cdots=l_p=0$. 这就证明了 $\varepsilon_1,\cdots,\varepsilon_r,\zeta_1,\cdots,\zeta_p$ 线性无关.

再证明 V 中任一向量 ξ 均可用 $\varepsilon_1,\cdots,\varepsilon_r,\zeta_1,\cdots,\zeta_p$ 表示. 作向量 $\eta=\sum_{i=1}^r(\xi,\varepsilon_i)\varepsilon_i+\sum_{j=1}^p(\xi,\zeta_j)\zeta_j\in V$. 那么由于 $(\xi-\eta,\varepsilon_i)=0,(\xi-\eta,\zeta_j)=0$ 对

$i=1,\cdots,r, j=1,\cdots,p$ 都成立．所以 $\xi-\eta\in S^{\perp}$ 且 $(\xi-\eta)\perp S^{\perp}$. 由例 7, 知 $\xi-\eta=\mathbf{0}$. 所以 $\xi=\eta$. 这就证明了 $\varepsilon_1,\cdots,\varepsilon_r,\zeta_1,\cdots,\zeta_p$ 是 V 的一组基．不难看出这还是 V 的一组标准正交基.

所以 $\dim V=r+p$. 又已知 $\dim V=n$. 所以 $\dim S^{\perp}=p=n-r$. 证毕.

习题 4. 若 S 是欧式空间 V 的一个线性子空间，证明 $(S^{\perp})^{\perp}=S$.

第 **2** 节　$R(A)$ 的刻画

考察线性方程组

$$A\mathbf{x}=\mathbf{b}, \tag{6}$$

其中 A 是 $m\times n$ 矩阵，$\mathbf{x}\in\mathbf{R}^n, \mathbf{b}\in\mathbf{R}^m$. 记 $R(A)=\{\mathbf{y}\in\mathbf{R}^m:$ 存在 $\mathbf{x}\in\mathbf{R}^n$ 使得 $\mathbf{y}=A\mathbf{x}\}$. 在第 4 章第 1 节例 8 我们已经知道了 $R(A)$ 是 $\mathbf{R}^m$ 的一个线性子空间. 此外回忆 $Ker(A)=\{\mathbf{x}\in\mathbf{R}^n:A\mathbf{x}=\mathbf{0}\}$. 作为欧式空间理论的一个应用，我们还有如下对 $R(A)$ 更深入的刻画.

定理 1. 设 $r(A)=r$, 则 $\dim R(A)=r$ 且 $R(A)=Ker(A^T)^{\perp}$. 矩阵 A 的列向量组的一个极大无关组就是 $R(A)$ 的一组基.

证明．1. 首先证明 $\dim R(A)=r$. 将矩阵 A 写为 $A=\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}$ 的形式，其中 $\alpha_i(i=1,2,\cdots,n)$ 是 m 维向量．由定义，我们就有 $R(A)=\operatorname{span}\{\alpha_1,\cdots,\alpha_n\}$, 从而 $\dim R(A)=r[\alpha_1,\cdots,\alpha_n]=r(A)=r$. $\alpha_1,\cdots,\alpha_n$ 的一个极大无关组显然就是 $R(A)$ 的一组基.

2. 由于 $r(A^T)=r(A)=r$, A^T 是 $n\times m$ 矩阵，所以 $Ker(A^T)$ 是 $\mathbf{R}^m$ 的子空间，且 $\dim Ker(A^T)=m-r$. 于是由上一节例 8, 知 $\dim Ker(A^T)^{\perp}=m-(m-r)=r$.

3. 再证明 $R(A)\subset Ker(A^T)^{\perp}$. 按照定义，也就是要证明，设 $\mathbf{y}\in R(A)$, 则 $\mathbf{y}\perp Ker(A^T)$, 即对于任意的 $\xi\in Ker(A^T)$, 要证明 $(\mathbf{y},\xi)=0$. 这里 $(\cdot,\cdot)$ 是 $\mathbf{R}^m$ 中的标准内积.

事实上，假设对 $\mathbf{x}\in\mathbf{R}^n$ 成立 $A\mathbf{x}=\mathbf{y}$, 那么

$$(\mathbf{y},\xi)=(A\mathbf{x},\xi)=\xi^T(A\mathbf{x})=(\xi^T(A\mathbf{x}))^T=\mathbf{x}^TA^T\xi=(A^T\xi,\mathbf{x})=(\mathbf{0},\mathbf{x})=0.$$

其中第三个等号用到了 $\xi^T(A\mathbf{x})$ 是一个数，从而它和它的转置是一样的．第四个等号用了矩阵乘积与转置之间的关系：$(A_1A_2\cdots A_n)^T = A_n^T\cdots A_2^TA_1^T$.

4. 这样我们证明了 $R(A)$ 是 $Ker(A^T)^{\perp}$ 的线性子空间，且具有相同的维数 r. 由下面命题 1, 可知它们必然相同．证毕.

命题 1.设 W 是 n 维线性空间 V 的子空间且 $\dim W = n$, 则 $W = V$.

证明．设 $\alpha_1,\cdots,\alpha_n$ 是 W 的一组基，又设 $\beta_1,\cdots\beta_n$ 是 V 的一组基．由于 $W \subset V$, 那么 $\alpha_1,\cdots,\alpha_n$ 可以由 $\beta_1,\cdots\beta_n$ 线性表示．由第 4 章第 3 节定理 2 的证明，知 $\beta_1,\cdots\beta_n$ 也可以由 $\alpha_1,\cdots,\alpha_n$ 线性表示．于是得到 $V \subset W$. 故成立 $W = V$. 证毕.

定理 1 的好处是抛开了秩和初等变换，所以它可以被推广到无限维线性空间中的线性方程，比如偏微分方程．下面是定理 1 的几个简单推论.

推论 1. 若 $r(A) = m$, 则 (6) 对任意 $\mathbf{b} \in \mathbf{R}^m$ 都有解.

推论 2. 当齐次方程 $A^T\mathbf{y} = \mathbf{0}$ 只有零解时，(6) 对任意 $\mathbf{b} \in \mathbf{R}^m$ 都有解.

推论 3. 若齐次方程 $A^T\mathbf{y} = \mathbf{0}$ 有 s 个线性无关的解，则只有当 $\mathbf{b}$ 满足 s 个齐次线性方程组时 (6) 才有解.

习题 1. 证明上述三个推论.

第 3 节　最小二乘法

在本课程的最后一节，我们介绍定义无解情形的方程 (6) 的广义解的一种方法：“最小二乘法”. 它的想法就是在 $\mathbf{R}^m$ 的线性子空间 $R(A)$ 中找到一个向量 $\mathbf{b}_0$, 使得

$$A\mathbf{x} = \mathbf{b}_0 \tag{7}$$

有解 $\mathbf{x}_0$, 而且 $\mathbf{b}_0$ 与 $\mathbf{b}$ 最 “接近”. 用欧式空间中向量长度的概念刻画 “接近” 与否，也就是要求

$$|\mathbf{b}_0 - \mathbf{b}| = \sqrt{(\mathbf{b}_0 - \mathbf{b}, \mathbf{b}_0 - \mathbf{b})} \tag{8}$$

最小．$\mathbf{x}_0$ 就是在最小二乘法意义下的解．换句话说，我们定义当 $r(A) \neq r(A \quad \mathbf{b})$ 时无解方程组 $A\mathbf{x} = \mathbf{b}$ 的 “广义解” 为使得定义在 $\mathbf{R}^n$ 上的如下函数

$$f(\mathbf{x}) = (A\mathbf{x} - \mathbf{b}, A\mathbf{x} - \mathbf{b})$$

取到最小值的那个 (些) 向量 $\mathbf{x}_0$, 而 $A\mathbf{x}_0=\mathbf{b}_0$.

本节我们将首先证明 $\mathbf{b}_0$ 的存在唯一性 (定理 1). 然后给出求广义解 $\mathbf{x}_0$ 的一个公式. 下面我们总假设 A 是 $m\times n$ 矩阵, $m>n$, 且 $r(A)=n$, 也就是说, 若记 $A=\begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}$, 那么 m 维列向量 $\mathbf{a}_1,\cdots,\mathbf{a}_n$ 线性无关.

定理 1. *设 W 是 n 维欧式空间 V 的真子空间, $\boldsymbol{v}\in V$ 但 $\boldsymbol{v}\notin W$. 则存在唯一的向量 $\boldsymbol{w}\in W$ 使得定义在 $\boldsymbol{x}\in W$ 上的函数 $|\boldsymbol{v}-\boldsymbol{x}|$ 取到最小值. 此外, $\boldsymbol{w}$ 还满足性质: $\boldsymbol{v}-\boldsymbol{w}\perp W$. 称 $\boldsymbol{w}$ 为 $\boldsymbol{v}$ 在 W 上的* **正交投影**. *反之, 若 $\boldsymbol{w}\in W$ 使得 $\boldsymbol{v}-\boldsymbol{w}\perp W$, 则 $\boldsymbol{w}$ 就是使得关于 $\boldsymbol{x}$ 的函数 $|\boldsymbol{v}-\boldsymbol{x}|$ 在 W 上取到最小值的向量.*

证明. 依本章第 1 节例 8 的证明, 我们设 $\varepsilon_1,\cdots,\varepsilon_r$ 是 W 的一组标准正交基, $\varepsilon_{r+1},\cdots,\varepsilon_n$ 是 $W^{\perp}$ 的一组标准正交基, 那么 $\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n$ 就是 V 的一组标准正交基. 设 $\mathbf{v}=\sum_{i=1}^{r}k_i\varepsilon_i+\sum_{j=r+1}^{n}k_j\varepsilon_j$, 我们就取 $\mathbf{w}=\sum_{i=1}^{r}k_i\varepsilon_i$. 显然 $\mathbf{v}-\mathbf{w}=\sum_{j=r+1}^{n}k_j\varepsilon_j\perp W$. 下面证明 $\mathbf{w}\in W$ 使得函数 $|\mathbf{v}-\mathbf{x}|(\mathbf{x}\in W)$ 取到最小值.

事实上, 对于任意 $\mathbf{x}\in W$ 可表示为 $\mathbf{x}=\sum_{i=1}^{r}x_i\varepsilon_i$, 其中 x_i 均是实数. 那么利用第 1 节例 5, 可知

$$|\mathbf{v}-\mathbf{x}|^2=\sum_{i=1}^{r}(k_i-x_i)^2+\sum_{j=r+1}^{n}k_j^2. \tag{9}$$

显然当 $x_i=k_i\ (i=1,\cdots,r)$ 时 $|\mathbf{v}-\mathbf{x}|^2$ 取到最小值, 且最小值就是 $\sum_{j=r+1}^{n}k_j^2=|\mathbf{v}-\mathbf{w}|^2$. 由此还可见 $\mathbf{w}$ 的唯一性.

反之, 设 $\mathbf{w}=\sum_{i=1}^{r}\overline{k}_i\varepsilon_i$ 使得 $\mathbf{v}-\mathbf{w}\perp W$, 即 $(\sum_{i=1}^{r}(k_i-\overline{k}_i)\varepsilon_i+\sum_{j=r+1}^{n}k_j\varepsilon_j,\varepsilon_m)=0$ 对 $m=1,\cdots,r$ 都成立. 这就意味着 $k_i=\overline{k}_i$, $i=1\cdots,r$. 从而由前面证明, 此 $\mathbf{w}$ 确使得 $|\mathbf{v}-\mathbf{x}|^2$ 取到最小值. 证毕.

下面考虑线性方程组 (7). 由定理 1, 我们已经知道, $\mathbf{b}_0$ 需要满足的条件就是

$$\mathbf{b}-\mathbf{b}_0\perp R(A). \tag{10}$$

由前面的假设, 由于 A 列满秩, $R(A)$ 的一组基就是 $\mathbf{a}_1,\cdots,\mathbf{a}_n$. 所以 (10) 就相当于如下 n 个方程:

$$(\mathbf{b}-\mathbf{b}_0,\mathbf{a}_j)=0 \quad 或 \quad (\mathbf{b}_0,\mathbf{a}_j)=(\mathbf{b},\mathbf{a}_j),\ j=1,\cdots,n. \tag{11}$$

利用矩阵乘法 (回忆 AB 的第 i 行第 j 列的元素就是 A 的第 i 行与 B 的第 j 列都作为 n 维向量作内积), 把它们合写成矩阵方程的形式，就是

$$\mathbf{b}_0^T A = \mathbf{b}^T A. \tag{12}$$

由于 $\mathbf{b}_0 \in R(A)$, 故可假设

$$\mathbf{b}_0 = \sum_{i=1}^{n} x_i \mathbf{a}_i = A\mathbf{x}_0, \tag{13}$$

其中 $x_i\,(i = 1, \cdots, n)$ 是待求的数，$\mathbf{x}_0 = (x_1, x_2, \cdots, x_n)^T$. 于是 (12) 就是如下方程的解：

$$\mathbf{x}_0^T A^T A = \mathbf{b}^T A,$$

两边同时取转置，就得到

$$A^T A\mathbf{x}_0 = A^T \mathbf{b}. \tag{14}$$

由于 $r(A) = n$, 我们断言 $A^T A$ 可逆 (注意 $A^T A$ 是 n 阶方阵), 即 $r(A^T A) = n$, 或者等价地,

$$A^T A\mathbf{y} = \mathbf{0} \tag{15}$$

只有零解 $\mathbf{y} = \mathbf{0} \in \mathbf{R^n}$. 事实上，设 $\mathbf{y}$ 是一个解，那么在 (15) 两边左乘 $\mathbf{y}^T$ 就得到 $\mathbf{y}^T A^T A\mathbf{y} = 0$. 利用 $\mathbf{y}^T A^T A\mathbf{y} = (A\mathbf{y}, A\mathbf{y})$, 我们就得到 $A\mathbf{y} = \mathbf{0}$. 但是 $r(A) = n$, 从而必有 $\mathbf{y} = 0$. 这就证明了 $A^T A$ 可逆.

于是，我们可从 (14) 写出

$$\mathbf{x}_0 = (A^T A)^{-1} A^T \mathbf{b}. \tag{16}$$

这就是最小二乘法意义下无解方程 (6) 的广义解，简称 **最小二乘解**.

例 1. *求方程组* $\begin{cases} x + y = 3, \\ 2x + 3y = 4, \\ 3x + 4y = 5 \end{cases}$ *的最小二乘解.*

解. 可写出 $A=\begin{pmatrix}1&1\\2&3\\3&4\end{pmatrix}$, $\mathbf{b}=\begin{pmatrix}3\\4\\5\end{pmatrix}$, 所以 $A^TA=\begin{pmatrix}14&19\\19&26\end{pmatrix}$, 从而 $\mathbf{x}_0=\frac{1}{3}\begin{pmatrix}26&-19\\-19&14\end{pmatrix}\begin{pmatrix}1&2&3\\1&3&4\end{pmatrix}\begin{pmatrix}3\\4\\5\end{pmatrix}=\begin{pmatrix}\frac{11}{3}\\-\frac{4}{3}\end{pmatrix}$, 即得所求.

习题 1. 求如下方程组的最小二乘解：

$$\begin{cases}0.39x-1.89y=1,\\0.61x-1.80y=1,\\0.93x-1.68y=1,\\1.35x-1.50y=1.\end{cases}$$

复习题

1. 在 $\mathbf{R}^4$ 中，求 $\mathbf{a},\mathbf{b}$ 之间的夹角 $<\mathbf{a},\mathbf{b}>$. 设

(1) $\mathbf{a}=(2,1,3,2)^T,\mathbf{b}=(1,2,-2,1)^T$;

(2) $\mathbf{a}=(1,2,2,3)^T,\mathbf{b}=(3,1,5,1)^T$;

(3) $\mathbf{a}=(1,1,1,2)^T,\mathbf{b}=(3,1,-1,0)^T$.

2. 设 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 是 $\mathbf{R}^3$ 的一组标准正交基. 证明： $\mathbf{a}_1=\frac{1}{3}(2\mathbf{e}_1+2\mathbf{e}_2-\mathbf{e}_3)$, $\mathbf{a}_2=\frac{1}{3}(2\mathbf{e}_1-\mathbf{e}_2+2\mathbf{e}_3)$, $\mathbf{a}_3=\frac{1}{3}(\mathbf{e}_1-2\mathbf{e}_2-2\mathbf{e}_3)$ 也是 $\mathbf{R}^3$ 的一组标准正交基.

3. 求齐次线性方程组 $\begin{cases}2x_1+x_2-x_3+x_4-3x_5=0,\\x_1+x_2-x_3+x_5=0,\end{cases}$ 的解空间 (作为 $\mathbf{R}^5$ 的子空间) 的一组标准正交基.

4. 证明欧式空间中的平行四边形等式：$|\mathbf{x}+\mathbf{y}|^2+|\mathbf{x}-\mathbf{y}|^2=2|\mathbf{x}|^2+2|\mathbf{y}|^2$.

5. 向量 $(1,1,-1,-1)^T,(1,1,0,0)^T,(0,0,-1,-1)^T,(2,2,-1,-1)^T,(1,0,-1,-1)^T$ 在 $\mathbf{R}^4$ 中的正交补空间的维数是多少？

6. 求如下问题的最小二乘解：$\begin{cases}a+1.2b=1.6,\\a+1.4b=1.7,\\a+1.8b=2.0.\end{cases}$

7. 设 $S\subset \mathbf{R}^n$ 是齐次线性方程组 $A\mathbf{x}=\mathbf{0}$ 的解空间，其中 A 是 $m\times n$ 矩阵. 则 $S^{\perp}$ 是什么方程组的解空间？

8. 设 $\alpha_1,\cdots,\alpha_m$ 是 n 维欧式空间的 m 个向量. 定义

$$A=\begin{pmatrix} (\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_m) \\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_m) \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_m,\alpha_1) & (\alpha_m,\alpha_2) & \cdots & (\alpha_m,\alpha_m) \end{pmatrix}.$$

证明： $\alpha_1,\cdots,\alpha_m$ 线性无关当且仅当 $|A|\neq 0$.