线 性 代 数

卢世荣
lshrlshr@163.com

作者所涂鸦的这本教材，希望能在便于阅读者学习《线性代数》这门课程上有所建树。作者认为，本教材重在课程思路的创建，希望线性代数某些内容的讲解方式使得阅读者能更好地掌握所要研究的问题，从而在学习这部分内容上具有清晰的思路；在合适的地方引入研究型学习方式，便于阅读者学习。作者认为本教材比较有特色的地方有如下几点：

1. 首先详细介绍了线性方程组及其高斯消元法，从而为通过讨论$2,3$元线性方程组的求解引入相应阶的行列式打下了基础，并可很自然地引入对行列式性质的讨论，另外也为矩阵的初等行变换、行阶梯形矩阵、行最简形矩阵打下了基础。
2. 通过对“线性表示的最简性“这一问题的讨论，我们可以很自然地引入向量组线性相关性的所有内容，且思路清晰，便于阅读者掌握线性代数中的这一难点；欧氏空间的相关内容则可通过”线性表示的有效性“来展开讨论，阅读者能够把以前的知识和这部分内容连接起来。
3. 改进了齐次线性方程组解空间维数定理的证明方法，在将齐次线性方程组的系数矩阵化为最简行标准形后，利用该方法可以直接给出该方程组的基础解系、通解。
4. 改变了部分概念、方法的叙述方法，将这些概念、方法与高中的概念、方法进行比较，便于阅读者理解。

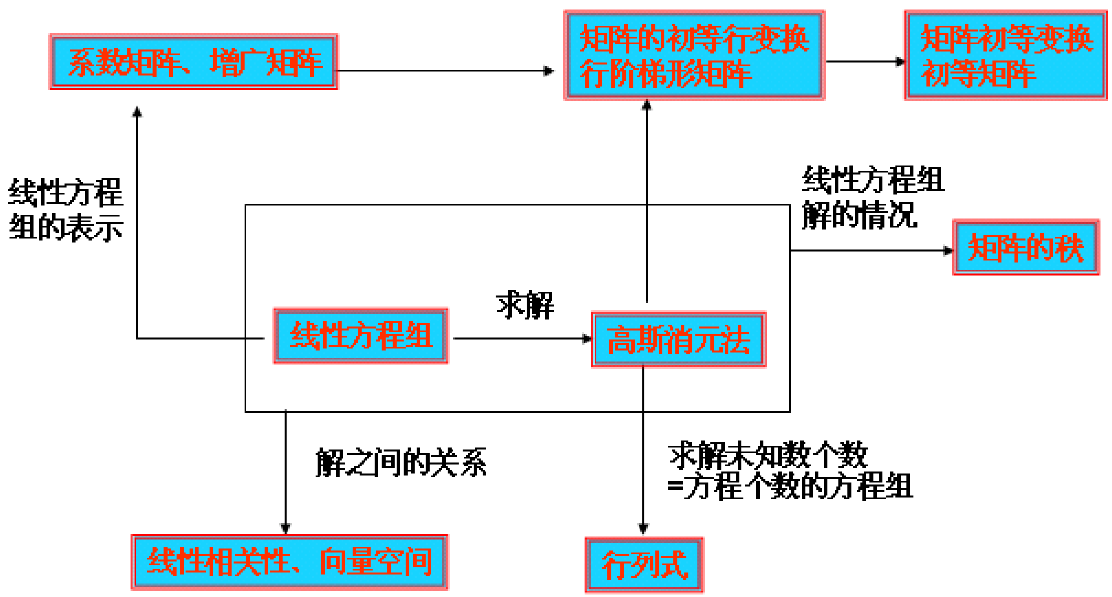
本教材的不足之处是明显的，首先是实际背景、与实际工程联系的例子很少，这就在提高大学新生的学习积极性上有不足；其次，在引入有些概念时铺垫不够，因此这些概念的引入不够自然；最后，学无止境，没有最好，只有更好，本教材在各个方面均具有很大的改进余地，对此作者深有自知之明。基于以上认识，故请各位阅读者不吝指教，多提宝贵意见建议，作者在此预先致谢！

目 录

作者：卢世荣　lshrlshr@163.com　　lshrlshr@gfkd.mtn　　欢迎任何意见与建议！

引论

有关线性方程组的内容，我们可通过下图来组织起来：



上图的说明：

1. 线性方程、线性方程组及其解；如何求解线性方程组——高斯消元法；我们能够通过计算判断线性方程组是否有解、解的情况（有唯一解、有无穷多解）
2. 在线性方程组中，我们通常用某个符号来代表未知数，而究竟用那个符号表示该未知数，却是可自由选择的，因此每个线性方程中，给出了关键信息的是各个未知数前面的系数、等号右边的常数，基于这种观察，就可给出线性方程、线性方程组的矩阵表示，这就是线性方程组的系数矩阵、增广矩阵；联系到高斯消元法，我们可得到行阶梯形矩阵、简化行阶梯形矩阵；
3. 既然可用增广矩阵来描述线性方程组，那么如果用增广矩阵来描述高斯消元法的过程，我们就可得到矩阵的初等行变换；结合前一点，可知我们可用初等行变换求解线性方程组；
4. 高斯消元法可求解一般的线性方程组，那么我们是否能够据此给出线性方程组的公式解？对于未知数个数与线性方程的个数相等的线性方程组，在一定情况下我们可以做到这一点，而这就导致了行列式的引入；
5. 从 1 可知，我们知道如何判断线性方程组的解的情况。但这是通过将线性方程组的增广矩阵化为行阶梯形后得到的，能否直接通过线性方程组的增广矩阵的某个度量来作出判断？这可通过引入矩阵秩的概念来作到；
6. 线性方程组可能有无穷多解，解的表达式不唯一，那么解的多种表达方式之间有何关系？按道理来讲，无论哪种表达式应该都能表示该线性方程组的所有解，因此我们需要讨论解的多种表达式之间的关系，这可通过线性相关性来讨论。

现在对以上的说明可能不太了解，但没有关系，在学习过程中会逐渐掌握。

第一章　　线性方程组

线性方程组在科学、工程中应用非常广泛，线性代数中许多概念都可以在对线性方程组的开展研究的过程中导出，因此我们先讨论线性方程组的基本内容。在学习完本章后，读者应该掌握如下内容：

（1）　线性方程组、解的概念；
（2）　如何表示一般的线性方程组？
（3）　高斯消元法的原理；
（4）　应用高斯消元法与高斯－约当法求解线性方程组；
（5）　如何判断线性方程组是否有解？

§1.1　线性方程组的表示、线性方程组的解

定义 1.1　　线性方程

含有 n 个未知数 $x_1, x_2, \cdots, x_n$ 的线性方程是具有如下形式的等式

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{1.1}$$

其中 $a_1, a_2, \cdots, a_n, b$ 为已知的常数，$a_i\left(i = 1, \cdots, n\right)$ 是未知数前 x_i 的系数，而 b 是等式右边的常数。

例 1.1　线性方程

$$2x = 2 \tag{1.2}$$

$$3x_1 + 2x_2 + 5x_3 = 8 \tag{1.3}$$

都是线性方程的例子。

定义 1.2　　线性方程的解

未知数 $x_1, x_2, \cdots, x_n$ 的一组值

$$\begin{cases} x_1 = s_1 \\ x_2 = s_2 \\ \quad\vdots \\ x_n = s_n \end{cases} \tag{1.4}$$

（$s_1, s_2, \cdots, s_n$ 为一组数）称为线性方程的(1.1)（一组）解，如果

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = b$$

成立。也可把(1.4)写为 $x_1 = s_1, x_2 = s_2, \cdots, x_n = s_n$ 的形式。

例如， $x=1$ 为(1.2)的解； $\begin{cases} x_1=1 \\ x_2=1 \\ x_3=1 \end{cases}$ 为(1.3)的一组解。

定义 1.3 线性方程组

多个含有相同未知数的线性方程构成一个线性方程组。

例 1.2 线性方程组的例子

$$\begin{cases} x_1+2x_2=3 \\ 2x_1+x_2=3 \end{cases} \tag{1.5}$$

$$\begin{cases} 2x_1-3x_2+2x_3=-5 \\ \qquad 2x_2-x_3=-2 \end{cases} \tag{1.6}$$

$$\begin{cases} 2x_1-3x_2+x_3=-5 \\ x_1-2x_2-x_3=-2 \\ x_1-x_2+4x_3=-11 \end{cases} \tag{1.7}$$

$$\begin{cases} x_1+3x_3+x_4=2 \\ x_1-3x_2+x_4=-1 \\ 2x_1+x_2+7x_3+2x_4=5 \\ 4x_1+2x_2+14x_3=6 \end{cases} \tag{1.8}$$

说明：加上大括号，表示这些线性方程是一个整体，即这些方程构成一个方程组。

线性方程组的一般形式。

由 m 个含有 n 个未知数 $x_1,x_2,\cdots,x_n$ 的线性方程所构成的线性方程组可表示为

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \qquad \vdots \\ a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i \longleftarrow \boxed{\text{第 } i \text{ 个方程}} \\ \qquad \vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases} \tag{1.9}$$

（1.9）

其中各个符号的意义为：

a_{ij}：第 i 个线性方程中第 j 个未知数（x_j）前的系数；

$i=1,2,\cdots,m$ 为方程序号；

$j=1,2,\cdots,n$ 为未知数的序号；

b_i：第 i 个线性方程中等号右边的常数；

称 $a_{ij}\left(i=1,2,\cdots,m;j=1,2,\cdots,n\right)$ 为方程组的系数，而称 $b_i\left(i=1,2,\cdots,m\right)$ 为常数项。

例如，对于线性方程组(1.6)，若表示为(1.9)的形式

$$\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3=b_1 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3=b_2 \end{cases}$$

则有

方程个数 $m=2$

未知数个数 $n=3$

$$a_{11}=2,\quad a_{12}=-3,\quad a_{13}=0,\quad b_1=-5$$
$$a_{21}=0,\quad a_{22}=2,\quad a_{23}=-1,\quad b_2=-2$$

定义 1.4 线性方程组的解

未知数 $x_1,x_2,\cdots,x_n$ 的一组值

$$\begin{cases} x_1=s_1 \\ x_2=s_2 \\ \quad\vdots \\ x_n=s_n \end{cases} \tag{1.10}$$

称为线性方程组（1.9）的解，如果（1.10）为线性方程组（1.9）中每个线性方程的解。

例如，$\begin{cases} x_1=1 \\ x_2=1 \end{cases}$ 为(1.5)的解，$\begin{cases} x_1=-6 \\ x_2=3 \\ x_3=8 \end{cases}$，$\begin{cases} x_1=-6 \\ x_2=3 \\ x_3=8 \end{cases}$，$\begin{cases} x_1=\frac{1}{2}(-9-c) \\ x_2=c \\ x_3=2+2c \end{cases}$，$\begin{cases} x_1=-4-\frac{c}{4} \\ x_2=-1+\frac{c}{2} \\ x_3=c \end{cases}$

（c 为任意常数）都是(1.6)的解，$\begin{cases} x_1=16 \\ x_2=11 \\ x_3=-4 \end{cases}$ 为(1.7)的解。

在线性方程组(1.9)中，若 $b_1=b_2=\cdots=b_m=0$，则称之为**齐次线性方程组**，否则称之为**非齐次线性方程组**。显然，若(1.9)为齐次线性方程组，则 $x_1=0,x_2=0,\cdots,x_n=0$ 必为它

的解，我们称之为齐次线性方程组(1.9)的**零解**（**平凡解**），齐次线性方程组的其它的解称为**非零解**（非平凡解）。也就是说，齐次线性方程组总是有解的，零解就是它的一个解。

若(1.9)为非齐次线性方程组，则令 $b_1 = b_2 = \cdots = b_m = 0$ 所得到的齐次线性方程组称为与之对应的齐次线性方程组，简称为对应的齐次线性方程组，或称为该非齐次线性方程组的**导出**（齐次线性方程）**组**。例如，与

$$\begin{cases} 2x_1 - 3x_2 + \quad x_3 = -5 \\ x_1 - 2x_2 - \quad x_3 = -2 \\ x_1 - \quad x_2 + 4x_3 = -11 \end{cases}$$

对应的齐次线性方程组为

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

对线性方程组而言，我们现在要研究的基本问题是：该线性方程组是否有解？如果该方程组有解，又该如何求解？

若线性方程组有解，该方程组也称为是**相容的**，否则称之为**不相容的**。

§1.2 线性方程组的求解——消元法

我们可以从简单的情形开始，以自然的方式引入求解线性方程组的解的表示方法与求解线性方程组的消元法，这种方式有助于读者理解高斯消元法、掌握解的表示方法。

2.1 一个方程的情形

2.1.1 一个方程、一个未知数

现在考虑线性方程 $ax = b$ 的求解。显然有

（1）当 $a \neq 0$ 时，线性方程 $ax = b$ 有唯一解 $x = \dfrac{b}{a}$；

（2）当 $a = 0$ 时

若 $b \neq 0$，线性方程即为 $0x = b \neq 0$，无论 x 取何值，方程都不可能成立，故此时方程无解；

若 $b = 0$，线性方程即为 $0x = 0$，这显然是恒成立的，因此任何数都是该方程的解；

2.1.2 一个方程、两个未知数

先以一个例子说明这种情况的处理方法。例如方程

$$x_1 + 2x_2 = 3$$

如果我们建立平面直角坐标系 $\boldsymbol{Ox_1x_2}$，则由解析几何的基本知识可知，该方程表示平面

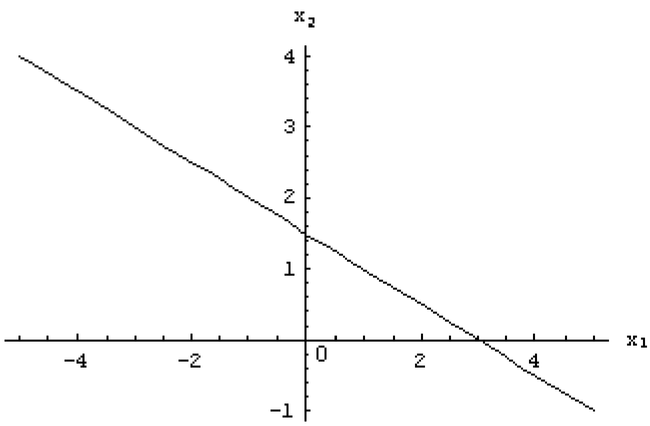


图 1-1

上的一条直线，该直线上的所有点的坐标$\left(x_1, x_2\right)$都满足该方程，因而可知该方程有解，且解不唯一，直线上的所有点的坐标都是该方程的解。

在这种情况下，虽然不能唯一确定该方程的解，但是如果我们能确定两个未知数x_1, x_2中某一个的值，譬如$x_1 = c$（c为任意常数），则有$x_2 = \dfrac{3-c}{2}$，此时可把该方程的解一般表示为

$$\begin{cases} x_1 = c \\ x_2 = \dfrac{3-c}{2} \end{cases} \qquad c \text{为任意常数}$$

由于c为任意常数，我们通常也直接将此线性方程的解表示为

$$\begin{cases} x_1 = x_1 \\ x_2 = \dfrac{3-x_1}{2} \end{cases} \qquad x_1 \text{为任意常数}$$

（注：在线性方程组解的这种表示方式下，出现在等号左边的符号为未知数，而凡是在等号右边出现的未知数都认为是已知的，它可任意取值，每取一个具体的值，就给出了方程的一个解。因此，上式中同一符号x_1有不同的意义。例如，令$x_1 = 1$，则得到一个解

$$\begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases}$$

）

当然，我们也能先确定x_2的值，则有$x_1 = 3 - 2x_2$，此时该线性方程的解又可表示为

含有两个未知数的线性方程$a_1x_1 + a_2x_2 = b$解的一般情况：

（1）$\left|a_1\right| + \left|a_2\right| > 0$（即$a_1, a_2$至少有一个不为 0），不妨设$a_1 \neq 0$，那么我们可先将$x_2$取任

意值，得到

$$x_1=\frac{b-a_2x_2}{a_1}$$

因此该线性方程的解可表示为

$$\begin{cases} x_1=\dfrac{b-a_2x_2}{a_1} \\ x_2=x_2 \end{cases}$$

（2）$|a_1|+|a_2|=0$，即 $a_1=a_2=0$，那么

（2.1） 若 $b=0$，方程此时为 $0x_1+0x_2=0$，显然 x_1,x_2 取任意值都是该方程的解；

（2.2）若 $b\neq 0$，方程此时为 $0x_1+0x_2=b\neq 0$，显然线性方程此时无解。

2.1.3 一个方程、多个未知数的情况

考虑线性方程(1.1)

$$a_1x_1+a_2x_2+\cdots+a_nx_n=b$$

求解该方程可分为如下三种情况：

（1）$|a_1|+|a_2|+\cdots+|a_n|>0$（即 $a_1,a_2,\cdots,a_n$ 至少有一个不为 0），不妨设 $a_i\neq 0$，那么我们可先选定未知数 $x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n$ 的值，可计算得 x_i 的值

$$x_i=\frac{1}{a_i}\left(b-\left(a_1x_1+\cdots+a_{i-1}x_{i-1}+a_{i+1}x_{i+1}+\cdots+a_nx_n\right)\right)$$

因此，该线性方程的解可表示为

$$\begin{cases} x_1=x_1 \\ \quad\vdots \\ x_{i-1}=x_{i-1} \\ x_i=\dfrac{1}{a_i}\left(b-\left(a_1x_1+\cdots+a_{i-1}x_{i-1}+a_{i+1}x_{i+1}+\cdots+a_nx_n\right)\right) \\ x_{i+1}=x_{i+1} \\ \quad\vdots \\ x_n=x_n \end{cases} \tag{1.11}$$

（2）$a_1=a_2=\cdots=a_n=0$，$b=0$，则 $x_1,x_2,\cdots,x_n$ 任意值都是该方程的解；

（3）$a_1=a_2=\cdots=a_n=0$，$b\neq 0$，此时方程无解。

为了给出一般的含有多个方程、多个未知数的线性方程组的求解方法，我们先考虑特殊

形式的线性方程组的求解，然后再过渡到一般情形线性方程组的求解。

2.2 含多个方程、多个未知数的特殊形式线性方程组的求解

我们先考虑如下几种特殊情况：

$$\begin{cases} 2x_1 \qquad\qquad = 32 \\ \quad x_2 \qquad = 11 \\ \qquad\quad 3x_3 = -12 \end{cases} \tag{1.12}$$

$$\begin{cases} x_1 - 2x_2 - x_3 = -2 \\ \qquad x_2 + 3x_3 = -1 \\ \qquad\qquad 2x_3 = -8 \end{cases} \tag{1.13}$$

$$\begin{cases} x_1 \qquad + x_3 = -4 \\ \quad x_2 - x_3 = -1 \end{cases} \tag{1.14}$$

$$\begin{cases} x_1 \qquad - x_3 - x_4 - 5x_5 = -16 \\ \quad x_2 + 2x_3 + 2x_4 + 6x_5 = 23 \end{cases} \tag{1.15}$$

$$\begin{cases} x_1 + 3x_2 + 6x_3 - x_4 - 5x_5 = 11 \\ \quad 2x_2 + x_3 + 3x_4 + 7x_5 = 23 \\ \qquad 3x_3 + 2x_4 + x_5 = 9 \end{cases} \tag{1.16}$$

$$\begin{cases} x_1 + 3x_2 + 6x_3 - x_4 - 5x_5 = 11 \\ \qquad x_3 + 3x_4 + 7x_5 = 23 \\ \qquad\quad 2x_4 + x_5 = 9 \end{cases} \tag{1.17}$$

通过简单观察，这几种类型的线性方程组容易求解。

(1.12)最容易求解，由于每个方程只含有一个未知数，且每个未知数的系数都不为 0，因此只需在每个方程两边都除以未知数前的系数，就可得到方程组的解

$$\begin{cases} x_1 = 16 \\ x_2 = 11 \\ x_3 = -4 \end{cases}$$

对于线性方程组(1.13)，我们观察到这三个方程中，注意到第三个方程中只含有一个未知数 x_3，因此可通过第三个方程求解出 x_3 的值；由于此时 x_3 已经由第三个方程解出，因此这时第二个方程实际上只含有一个未知数 x_2，因此可通过第二个方程求解出 x_2；类似地，由于 x_2, x_3 已知，因此第一个方程实际上只含有一个未知数 x_1，因此就能由第一个方程求解出 x_1。我们用式子描述这个过程：

首先，由第三个方程 $2x_3 = -8$ ，得到 $x_3 = -4$ ；

由第二个方程 $x_2+3x_3=-1 \Rightarrow x_2=-1-3x_3=-1-3\times(-4)=11$；

由第一个方程 $x_1-2x_2-\ \ x_3=-2$，因此

$$x_1=-2+2x_2+x_3=-2+2\times 11+(-4)=16$$

因此我们就得到线性方程组(1.13)的解

$$\begin{cases} x_1=16 \\ x_2=11 \\ x_3=-4 \end{cases}$$

对于线性方程组(1.14)，显然，如果我们将 x_3 视为已知量（即 x_3 可以取任意值），则(1.14)显然就化为(1.13)的情形，再联系到线性方程(1.1)有多个未知数的情况，我们可以如下处理该问题：把 x_3 视作已知量，将这两个方程中的未知数 x_3 移到等号右边，得到

$$\begin{cases} x_1 \qquad\ \ =-4-x_3 \\ \qquad x_2=-1+x_3 \end{cases}$$

因此我们能把该线性方程组的解表示为

$$\begin{cases} x_1=-4-x_3 \\ x_2=-1+x_3 \\ x_3=x_3 \end{cases} \qquad \boldsymbol{x}_3 \text{ 可取任意常数}$$

对于线性方程组（1.15），实际上类似于（1.14），我们可以把未知数 x_3,x_4,x_5 “视作”已知量并移到等式右边，得到

$$\begin{cases} x_1 \qquad =-16+x_3+\ x_4+5x_5 \\ \qquad x_2=23-2x_3-2x_4-6x_5 \end{cases}$$

因此，对于 x_3,x_4,x_5 的任意值，都给出了该线性方程组的一组解

$$\begin{cases} x_1=-16+x_3+\ x_4+5x_5 \\ x_2=23-2x_3-2x_4-6x_5 \\ x_3=x_3 \\ x_4=x_4 \\ x_5=x_5 \end{cases}$$

对于线性方程组(1.16)，联系到前面几种情况的处理，我们可以首先把各个方程中的未知数 x_4,x_5 移到等号右边，得到

$$\begin{cases} x_1+3x_2+6x_3=11+\ \ x_4+5x_5 \\ \quad 2x_2+x_3=23-3x_4-7x_5 \\ \qquad\quad 3x_3=9-2x_4-x_5 \end{cases}$$

这时我们把 x_4, x_5 当作已知量，上述方程组具有与（1.13）相同的形式，因此就可以用相同的方法求解，先从第三个方程求出 x_3，

$$x_3=3-\frac{2}{3}x_4-\frac{1}{3}x_5$$

然后从第二个方程求解出 x_2，

$$\begin{aligned} x_2&=\frac{1}{2}(23-3x_4-7x_5-x_3) \\ &=\frac{1}{2}\left(23-3x_4-7x_5-\left(3-\frac{2}{3}x_4-\frac{1}{3}x_5\right)\right) \\ &=\frac{1}{2}\left(20-\frac{7}{3}x_4-\frac{20}{3}x_5\right) \\ &=10-\frac{7}{6}x_4-\frac{10}{3}x_5 \end{aligned}$$

最后从第一个方程求解出 x_1

$$\begin{aligned} x_1&=11+\ \ x_4+5x_5-3x_2-6x_3 \\ &=11+\ \ x_4+5x_5-3\left(10-\frac{7}{6}x_4-\frac{10}{3}x_5\right)-6\left(3-\frac{2}{3}x_4-\frac{1}{3}x_5\right) \\ &=-27+\frac{22}{3}x_4+\frac{41}{3}x_5 \end{aligned}$$

因此，该线性方程组的解可表示为

$$\begin{cases} x_1=-27+\dfrac{22}{3}x_4+\dfrac{41}{3}x_5 \\ x_2=10-\dfrac{7}{6}x_4-\dfrac{10}{3}x_5 \\ x_3=3-\dfrac{2}{3}x_4-\dfrac{1}{3}x_5 \\ x_4=x_4 \\ x_5=x_5 \end{cases}$$

对于线性方程组(1.17)，其求解方式与(1.15)类似，我们只需要将未知数 x_2, x_5 移到等号右边，得到

$$\begin{cases} x_1+6x_3- x_4=11-3x_2+5x_5 \\ x_3+3x_4=23-7x_5 \\ 2x_4=9- x_5 \end{cases}$$

这也是方程组(1.13)的形式，因此其解法就与前面的例子相同。

根据上面的分析、介绍，我们得知上述这些线性方程组的解法是相同的：

（1） 如果必要，把某些未知数当作已知，并把所有方程中的这些未知数移到等号右边；这样，使得最后一个方程只含有一个未知数、倒数第二个方程只含有两个未知数，…；

（2） 从最后一个方程开始求解，逐步把所求解到的未知数的值代入到前一个方程中，使得该方程实际上只含有一个未知数，从而进行求解；

但是，我们应该意识到：并不是所有的线性方程组都可直接用上述方法求解。那么能用上述方法求解的线性方程组必须具有什么特点呢？

线性方程组(1.13)~(1.17)有一个共同点：**线性方程组中各方程所含未知数的个数越来越少！**我们称这种类型的线性方程组为**阶梯形（线性）方程组**。我们把以上求解阶梯形线性方程组的上述方法称作**“回代法”**，因为它是将已经求得的未知数的值代入前面的方程来求解新的未知数的过程。

（说明：可能有读者会注意到：线性方程组(1.14)中，每个方程都只有两个未知数，似乎并不符合阶梯形方程组的定义；实际上，如果在第一个方程中“引入”未知数 x_2，得到形如 $\begin{cases} x_1+x_2+x_3=-4 \\ x_2-x_3=-1 \end{cases}$ 的方程组，显然它能用回代法求解。基于如上事实，为了严格阶梯形线性方程组的定义，我们约定：凡是在后面方程组中出现的未知数，我们认为它在前面的方程中出现。例如在方程组(1.14)中，第二个方程中出现了未知数 x_2，即使原来第一个方程中没有出现未知数 x_2，我们也认为第一个方程中“包含”了 x_2）

问题：在以上几种情形中，我们需要把某些未知数“当作”已知，然后把这些未知量移到方程的等号右边，如何确定这些未知数？我们把这些“当作”已知、并移到等号右边的未知数叫做**自由未知量**，而其余的未知数叫做**约束未知量**。因此，这就是如何确定一个未知量是约束未知量还是自由未知量的问题。

注意：在上述阶梯形线性方程组中，在每个方程中出现的第一个未知数可选作约束未知量，而其它的未知量则作为自由未知量。在大多数教材中，约束未知量、自由未知量都是通过这种方式选择的。

（ 注：

当然，自由未知量、约束未知量的选取不是唯一的。例如，对于方程组(1.16)

$$\begin{cases} x_1+3x_2+6x_3- x_4-5x_5=11 \\ 2x_2+x_3+3x_4+7x_5=23 \\ 3x_3+2x_4+x_5=9 \end{cases}$$

可以把 x_1,x_2,x_4 作为约束未知量，而 x_3,x_5 为自由未知量，即把方程组视作为

$$\begin{cases} x_1+3x_2-\ x_4+6x_3-5x_5=11 \\ \quad 2x_2+3x_4+x_3+7x_5=23 \\ \qquad 2x_4+3x_3+x_5=9 \end{cases}$$

然后把每个方程中的 x_3, x_5 移到等号右边，得到

$$\begin{cases} x_1+3x_2-\ x_4=11-6x_3+5x_5 \\ \quad 2x_2+3x_4=23-x_3-7x_5 \\ \qquad 2x_4=9-3x_3-x_5 \end{cases}$$

然后用回代法求解。
）

2.3 多个方程的一般情形

对于线性方程组的一般情形(1.9)，其求解方法并不是那么一目了然，联系到上一部分所讲述的阶梯形线性方程组的求解，自然的想法：是否能够借助于特殊情形线性方程组的求解方法来求解一般形式的线性方程组？当然，要做到这一点，首先要做到的是：建立一般形式线性方程组与阶梯形线性方程组之间的关系，将求解一般形式的线性方程组以某种方式转化为求解阶梯形线性方程组？这就是高斯消元法。高斯消元法的目的是将一般形式的线性方程组转化为阶梯形线性方程组、然后进行求解，我们先举例说明这个方法，然后介绍该方法的理论基础。

例 1.3 高斯消元法

$$\begin{cases} 2x_1-\ x_2+7x_3=-7 & ① \\ x_1-2x_2-\ x_3=-2 & ② \\ x_1-\ x_2+4x_3=-11 & ③ \end{cases} \tag{1.18}$$

首先交换方程组中①、 ②两个方程的位置，表示为

①↔②

得到新的方程组（为何要交换着两个方程的位置？），得到

$$\begin{cases} x_1-2x_2-\ x_3=-2 & ① \\ 2x_1-\ x_2+7x_3=-7 & ② \\ x_1-\ x_2+4x_3=-11 & ③ \end{cases} \tag{1.19}$$

为了将该方程组化为阶梯形，即方程组中各方程所含未知数越来越少，我们需要减少方程②、③所含的未知数。我们首先“消掉”②、③中的未知数 x_1。为消掉②所含的未知数 x_1，我们将①乘以2，得到方程

$$2x_1-4x_2-\ 2x_3=-4$$

然后用（1.19）中的②减去该方程，得

$$2x_1-x_2+7x_3-(2x_1-4x_2-2x_3)=-7-(-4)$$

我们将该过程记作②－①×2。将上式化简，得到一个新方程

$$3x_2+9x_3=-3$$

我们用这个新的方程取代方程组(1.19)中的第二个方程，而(1.19)中的其余方程保持不变，得到一个新的方程组

$$\begin{cases} x_1-2x_2-\ \ x_3=-2 & ① \\ \qquad 3x_2+9x_3=-3 & ② \\ x_1-\ \ x_2+4x_3=-11 & ③ \end{cases} \tag{1.20}$$

类似地，我们可“去掉”③所含的未知数 x_1，这可通过③－①得到

$$(x_1-\ \ x_2+4x_3)-(x_1-2x_2-\ \ x_3)=-11-(-2)$$

化简得

$$x_2+5x_3=-9$$

这样，用这个新方程取代方程组(1.20)的第三个方程，而其余两个方程保持不变，这样得到一个新的方程组

$$\begin{cases} x_1-2x_2-\ \ x_3=-2 & ① \\ \qquad 3x_2+9x_3=-3 & ② \\ \qquad\ \ x_2+5x_3=-9 & ③ \end{cases} \tag{1.21}$$

注意到上述方程组中第二个方程的等号两边有一个公因子3，我们可以去掉公因子3，即该方程除以3，即②÷3，得到

$$x_2+3x_3=-1$$

用该方程取代(1.21)中的方程②，而保持其它两个方程保持不变，这样得到一个新的方程组

$$\begin{cases} x_1-2x_2-\ \ x_3=-2 & ① \\ \qquad x_2+3x_3=-1 & ② \\ \qquad x_2+5x_3=-9 & ③ \end{cases} \tag{1.22}$$

为了进一步减少方程组(1.22)中第三个方程所含的未知数，可以“去掉”方程组(1.22)中方程③所含的未知数 x_2，这可通过③－②得到

$$(x_2+5x_3)-(x_2+3x_3)=-9-2\times(-1)$$

即 $2x_3=-8$

这样，用这个新得到的方程取代(1.22)中的第三个方程，而其它方程保持不变，这样得到一个新的线性方程组

$$\begin{cases} x_1-2x_2-\ \ x_3=-2 & ① \\ \qquad x_2+3x_3=-1 & ② \\ \qquad\qquad 2x_3=-8 & ③ \end{cases} \tag{1.23}$$

这是一个阶梯形线性方程组，就可用回代法求解。我们把这个将一般线性方程组化为阶梯形

线性方程组的过程称为“消元”过程——逐步减少未知数的过程。

通过回代过程，我们可求得方程组(1.23)的解为

$$\begin{cases} x_1 = 16 \\ x_2 = 11 \\ x_3 = -4 \end{cases} \tag{1.24}$$

注意：(1.24)是线性方程组(1.23)的解，但它是否为线性方程组(1.18)的解？该问题转化为线性方程组(1.18)与(1.23)是否有相同的解。

我们注意到：将一般形式的线性方程组化为阶梯形线性方程组的过程中，我们得到了一序列的方程组(1.18)～(1.23)，而对于每相邻的两个线性方程组，后一个方程组（第二个方程组）是通过对前一个方程组（第一个方程组）做了下列三种运算中的某一个得到的：

1. **交换第一个方程组中第 i, j 个方程的位置，而方程组中的其余方程保持不变，这样得到第二个方程组；**
2. **在第一个方程组中，将第 j 个方程减去第 i 个方程的某个常数倍，得到的新方程作为第二个方程组中的第 j 个方程，而第二个方程组中是的其余方程与第一个方程组中对应的方程是相同的；**
3. **将第一个方程组中的第 i 方程除以某个非零常数，得到的新方程作为第二个方程组中的第 i 方程，而第二个方程组中的其余方程与第一个方程组中对应的方程相同；**

称以上三种运算为**线性方程组的初等变换**，即线性方程组的初等变换包含三种运算。

相邻的两个线性方程组是否具有相同的解？我们注意到如下两个显然的事实。

引理 1.1 线性方程组 I 经初等变换的一种运算化为线性方程组 II，则 I 的解必为 II 的解；

证明：（1）如果线性方程组 I 做第一种初等变换得到线性方程组 II，由于这两个方程组含有的方程是相同的，根据线性方程组解的定义，则 I 的解必为 II 的解；

（2）如果线性方程组 I 做一次第三种初等变换得到线性方程组 II，由于这两个方程组只有一个方程不同，我们只需要考虑方程组 I 的解是否满足方程组 II 中的这个新方程。设方程组 I 的第 i 个方程 $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$ 除以非零数 k，得到方程组 II 中的方程 $\frac{a_{i1}}{k}x_1 + \frac{a_{i2}}{k}x_2 + \cdots + \frac{a_{in}}{k}x_n = \frac{b_i}{k}$，这两个方程组中的其余方程相同，而 $x_1 = s_1, x_2 = s_2, \cdots, x_n = s_n$ 为方程组 I 的解，因此有 $a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n = b_i$ 成立，该等式除以非零数 k，得到 $\frac{a_{i1}}{k}s_1 + \frac{a_{i2}}{k}s_2 + \cdots + \frac{a_{in}}{k}s_n = \frac{b_i}{k}$，因此 $x_1 = s_1, x_2 = s_2, \cdots, x_n = s_n$ 满足 II 中的方程 $\frac{a_{i1}}{k}x_1 + \frac{a_{i2}}{k}x_2 + \cdots + \frac{a_{in}}{k}x_n = \frac{b_i}{k}$，而 II 中的其它方程与 I 的相同，因此 $x_1 = s_1, x_2 = s_2, \cdots, x_n = s_n$ 也是 II 的解；

（3）如果线性方程组 I 做一次第二种初等变换得到线性方程组 II，这两个方程组也只有一个方程不同。假设 I 中第 j 个方程 $a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j$ 减去第 i 个方程

$a_{i1}x_1+ a_{i2}x_2+\cdots+a_{in}x_n=b_i$ 的 k 倍，得到方程组Ⅱ中的第 j 个方程

$$\left(a_{j1}-ka_{i1}\right)x_1+\left(a_{j2}-ka_{i2}\right)x_2+\cdots+\left(a_{jn}-ka_{in}\right)x_n=b_j-kb_i$$

而Ⅱ中的其它方程与Ⅰ中的相应方程相同。设 $x_1=s_1, x_2=s_2,\cdots,x_n=s_n$ 为方程组Ⅰ的解，因此它满足方程组Ⅰ中的每一个方程，因此有

$$a_{j1}s_1+ a_{j2}s_2+\cdots+a_{jn}s_n=b_j$$

$$a_{i1}s_1+ a_{i2}s_2+\cdots+a_{in}s_n=b_i$$

均成立，这是两个恒等式。将第一个式子减去第二个式子的 k 倍，得到

$$\left(a_{j1}-ka_{i1}\right)s_1+\left(a_{j2}-ka_{i2}\right)s_2+\cdots+\left(a_{jn}-ka_{in}\right)s_n=b_j-kb_i$$

因此 $x_1=s_1, x_2=s_2,\cdots,x_n=s_n$ 满足Ⅱ中的第 j 个方程

$$\left(a_{j1}-ka_{i1}\right)x_1+\left(a_{j2}-ka_{i2}\right)x_2+\cdots+\left(a_{jn}-ka_{in}\right)x_n=b_j-kb_i$$

而显然 $x_1=s_1, x_2=s_2,\cdots,x_n=s_n$ 也满足Ⅱ中的其它所有方程，因此 $x_1=s_1, x_2=s_2,\cdots,x_n=s_n$ 是方程组Ⅱ的解。

综合上述三种情况，我们指导引理成立。 ■

引理 1.2　线性方程组Ⅰ经初等变换的一次运算化为线性方程组Ⅱ，则线性方程组Ⅱ可通过初等变换的一次运算化为线性方程组Ⅰ。

证明：（1）如果互换方程组Ⅰ的第 i、j 个方程得到方程组Ⅱ，则互换方程组Ⅱ的第 i、j 个方程得到方程组；

（2）如果方程组Ⅰ的第 j 个方程减去第 i 个方程的 k 倍得到方程组Ⅱ，则方程组Ⅱ的第 j 个方程加上第 i 个方程的 k 倍得到方程组Ⅰ；

（3）如果方程组Ⅰ的第 i 个方程除以非零数 k 得到方程组Ⅱ，则方程组Ⅱ的第 i 个方程除以非零数 $\dfrac{1}{k}$ 得到方程组Ⅰ。

因此有如下结论：

若线性方程组Ⅰ可经初等变换化为线性方程组Ⅱ，则线性方程组Ⅱ可经若干初等变换化为线性方程组Ⅰ，我们称这两个线性方程组**等价**。

定理 1.1 等价的线性方程组具有相同的解。

证明：由引理 1.1、引理 1.2，结论是显然的。 ■

根据以上的论述可知：在求解线性方程组时，我们可首先用消元法将线性方程组化为阶梯形线性方程组，由于原始的线性方程组与所得到的阶梯形方程组等价，因此它们有相同的解；然后用回代法求解该阶梯形线性方程组，就可得到原方程组的解。这就是求解线性方程

组的**高斯消元法**。

例1.4 求解线性方程组

（1）$\begin{cases} x_1+3x_3+x_4=2 \\ x_1-3x_2+x_4=-1 \\ 2x_1+x_2+7x_3+2x_4=5 \\ 4x_1+2x_2+14x_3=6 \end{cases}$ （2）$\begin{cases} x_1-x_2+3x_3-x_4=1 \\ 2x_1-x_2-x_3+4x_4=2 \\ 3x_1-2x_2+2x_3+3x_4=3 \\ x_1-4x_3+5x_4=-1 \end{cases}$

解：（1）$\begin{cases} x_1+3x_3+x_4=2 & ① \\ x_1-3x_2+x_4=-1 & ② \\ 2x_1+x_2+7x_3+2x_4=5 & ③ \\ 4x_1+2x_2+14x_3=6 & ④ \end{cases}$

先进行消元过程：利用①消除中②中的未知数 $\boldsymbol{x}_1$，相应的运算为②－①，得到方程组

$$\begin{cases} x_1+3x_3+x_4=2 \\ (x_1-3x_2+x_4)-(x_1+3x_3+x_4)=-1-2 \\ 2x_1+x_2+7x_3+2x_4=5 \\ 4x_1+2x_2+14x_3=6 \end{cases}$$

利用①消除中③中的未知数 x_1，相应的运算③－2×①，得到线性方程组，

$$\begin{cases} x_1+3x_3+x_4=2 \\ (x_1-3x_2+x_4)-(x_1+3x_3+x_4)=-1-2 \\ (2x_1+x_2+7x_3+2x_4)-2\times(x_1+3x_3+x_4)=5-2\times2 \\ 4x_1+2x_2+14x_3=6 \end{cases}$$

利用①消除中④中的未知数 x_1，相应的运算为④－4×①，得到线性方程组，

$$\begin{cases} x_1+3x_3+x_4=2 \\ (x_1-3x_2+x_4)-(x_1+3x_3+x_4)=-1-2 \\ (2x_1+x_2+7x_3+2x_4)-2\times(x_1+3x_3+x_4)=5-2\times2 \\ (4x_1+2x_2+14x_3)-4\times(x_1+3x_3+x_4)=6-4\times2 \end{cases}$$

这样写非常麻烦，因此我们把以上三步合并为一步，即：在题目所给方程组中，利用第一个方程①消去②，③，④中的未知数 x_1，也就是对所给方程组做运算②－①，③－2×①，④－4×①，得到

$$\begin{cases} x_1+3x_3+x_4=2 \\ (x_1-3x_2+x_4)-(x_1+3x_3+x_4)=-1-2 \\ (2x_1+x_2+7x_3+2x_4)-2\times(x_1+3x_3+x_4)=5-2\times 2 \\ (4x_1+2x_2+14x_3)-4\times(x_1+3x_3+x_4)=6-4\times 2 \end{cases}$$

化简为
$$\begin{cases} x_1 \quad +3x_3+x_4=2 & ① \\ -3x_2-3x_3=-3 & ② \\ x_2+x_3=1 & ③ \\ 2x_2+2x_3-4x_4=-2 & ④ \end{cases}$$

注意到此时的第二个方程有个共因子－3、第四个方程有共因子 2，去掉这两个因子，得到

$$\begin{cases} x_1 \quad +3x_3+x_4=2 & ① \\ x_2+x_3=1 & ② \\ x_2+x_3=1 & ③ \\ x_2+x_3-2x_4=-1 & ④ \end{cases}$$

然后利用②消除③、④中的未知数 $\boldsymbol{x}_3$，这可通过③－②，④－②得到

$$\begin{cases} x_1 \quad +3x_3+x_4=2 & ① \\ x_2+x_3=1 & ② \\ 0=0 & ③ \\ -2x_4=-2 & ④ \end{cases}$$

去掉④中的因子－2，得到

$$\begin{cases} x_1 \quad +3x_3+x_4=2 & ① \\ x_2+x_3=1 & ② \\ 0=0 & ③ \\ x_4=1 & ④ \end{cases}$$

可以认为 $0=0$ 这个方程不包含未知数，其所含未知数个数最少，因此可把它交换到最后：

$$\begin{cases} x_1 \quad +3x_3+x_4=2 \\ x_2+x_3=1 \\ x_4=1 \\ 0=0 \end{cases}$$

这已经是阶梯形线性方程组。上述方程组中的第四个方程 $0=0$ 总是成立的，因此在用回代法求解时可以忽略掉这个方程，因此得到

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

其中 x_1,x_2,x_4 为约束未知量， x_3 为自由未知量，将 x_3 移到方程右边，得到

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

利用回代法求得方程组的解为

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

这就是原方程组的解。

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

先进行消元过程：首先利用①消除中②、③、④中的未知数 x_1，相应的运算为 ②－2×①，③－3×①，④－①，得到

$$\begin{cases} x_1 - x_2 + 3x_3 - x_4 = 1 \\ (2x_1 - x_2 - x_3 + 4x_4) - 2\times(x_1 - x_2 + 3x_3 - x_4) = 2 - 2\times1 \\ (3x_1 - 2x_2 + 2x_3 + 3x_4) - 3\times(x_1 - x_2 + 3x_3 - x_4) = 3 - 3\times1 \\ (x_1 - 4x_3 + 5x_4) - (x_1 - x_2 + 3x_3 - x_4) = -1 - 1 \end{cases}$$

化简得

$$\begin{cases} x_1 - x_2 + 3x_3 - x_4 = 1 & ① \\ \quad x_2 - 7x_3 + 6x_4 = 0 & ② \\ \quad x_2 - 7x_3 + 6x_4 = 0 & ③ \\ \quad x_2 - 7x_3 + 6x_4 = -2 & ④ \end{cases}$$

利用第二个方程消掉中的未知数 x_2，相应的运算为③－②，④－②，得到

$$\begin{cases} x_1 - x_2 + 3x_3 - x_4 = 1 \\ \quad x_2 - 7x_3 + 6x_4 = 0 \\ \quad\quad 0 = 0 \\ \quad\quad 0 = -2 \end{cases}$$

最后一个方程显然不可能成立，因此该方程组无解。

问题：在消元过程中，为达到消去某个未知数，一个方程需要减去另一个方程的某个倍数，如何确定该倍数？

高斯消元法的变形：

经典的高斯消元法由“消元”、“回代”两个阶段构成，但是我们也可不需要“回代”过程，而是通过在将线性方程组化为阶梯形方程组之后，再从最后一个方程开始，继续用消元过程消掉前面各个方程中所含有的约束未知量，使得最后得到的方程组中，每个方程只含有一个约束未知量；进一步地，我们还把所有约束未知量前的系数化为 1。这样，我们只需要将每个方程中的自由未知量移到等号右边，就可得到线性方程组的解。

现举例说明这一方法。在例 1.3 中，将方程组化为(1.23)

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

后，根据约束未知量、自由未知量的选取方法可知，该方程组的三个未知数 x_1, x_2, x_3 均为约束未知量。首先去掉第三个方程的共因子 2，得到

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

利用第三个方程消掉①、②中的未知数 x_3，运算为②－3×③，①－（－1）③，得到线性方程组

$$\begin{cases} x_1 - 2x_2 = -6 \\ x_2 = 11 \\ x_3 = -4 \end{cases}$$

然后利用②消掉①中的未知数 x_2，相应运算为①－（－2）①，得到方程组

$$\begin{cases} x_1 = 16 \\ x_2 = 11 \\ x_3 = -4 \end{cases}$$

此时每个方程中只含有一个约束未知量，且每个约束未知量的系数为 1，实际上这就是线性方程组的解。

又例如假设一个线性方程组经消元过程化为如下阶梯形方程组：

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

当然可用回代过程继续求解。注意到约束未知量为 x_1, x_3, x_4，而自由未知量为 x_2，因此用第三个方程③消掉方程①、②中的未知数 x_4，运算为②－2×③，①－③，得到

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

然后用②消掉①中的未知数 x_3，相应运算为①－3×②，得到

$$\begin{cases} x_1 + x_2 \quad = 4 \\ \qquad x_3 \quad = -1 \\ \qquad\qquad x_4 = 1 \end{cases}$$

然后将自由未知量移到方程的右边，得到

$$\begin{cases} x_1 \quad = 4 - x_2 \\ \qquad x_3 \quad = -1 \\ \qquad\qquad x_4 = 1 \end{cases}$$

因此，线性方程组的解为

$$\begin{cases} x_1 = 4 - x_2 \\ x_2 = x_2 \\ x_3 = -1 \\ x_4 = 1 \end{cases}$$

在手工求解线性方程组组时常用这种方式求解。我们称这种方法为**高斯－约当消元法**。

注释：在消元过程中，消掉的未知数是可选的，一般情况下，我们首先使用第一个方程消掉后面各个方程中的第一个未知数（x_1），然后用第二个方程消掉后面各个方程中的第二个未知数，……，我们称该消元次序为**标准消元**。但并不是必须如此，消掉哪些未知数实际上是可以选择的。以线性方程组（1.18）为例，

$$\begin{cases} 2x_1 - 3x_2 + \quad x_3 = -5 & ① \\ \ \ x_1 - 2x_2 - \quad x_3 = -2 & ② \\ \ \ x_1 - \quad x_2 + 4x_3 = -11 & ③ \end{cases}$$

我们可先消掉第二、三个方程中的未知数 $\boldsymbol{x}_3$，

②＋①，③－4×①

得到

$$\begin{cases} 2x_1 \quad -3x_2 + \quad x_3 = -5 & ① \\ 3x_1 \quad -5x_2 \qquad = -7 & ② \\ -7x_1 + 11x_2 \qquad = 9 & ③ \end{cases}$$

然后用②消掉③中的一个未知数，譬如 $\boldsymbol{x}_1$，可用的运算为

③×3＋②×7

得到

①

②

③

$$\begin{cases} 2x_1 - 3x_2 + \ \ x_3 = -5 \\ 3x_1 - 5x_2 \qquad\ \ = -7 \\ \qquad -2x_2 \qquad = -22 \end{cases}$$

此时方程组中各方程所含未知数越来越少，现在可用回代过程求解；当然也可用高斯消元法的变形方法求解。

从现在开始，若没有特别说明，本教材中的消元过程都采用标准的消元过程。

到目前为止，我们已经学习了求解线性方程组的高斯消元法、高斯－约当消元法。对于一个给定的线性方程组，我们可用任何一种方法求解。为了解一个方程组解的情况，我们可通过用消元法将线性方程组化为阶梯形方程组，再来判断方程组解的情况。因此我们只需要对阶梯形线性方程组考虑如下问题

（1） 是否有解？

（2） 如果有解，是否是唯一解？

我们不考虑阶梯形方程组中“$0=0$”型的方程，由于这些方程总是成立的，因此去掉这些“$0=0$”型的方程并不会影响方程组的解的情况。为便于讨论方便起见，我们假设该阶梯形方程组中没有“$0=0$”型的方程。根据自由未知量、约束未知量的定义及选取方法可知，此时约束未知量的个数就为方程组中方程的个数。

（1） 如果最后一个方程没有未知数：既然最后一个方程不是“$0=0$”型的方程，该方程又不含有未知数，该方程显然是“$0=b\neq 0$”的形式，该方程显然不可能成立，此时方程组无解；例 1.4（2）就是这种情况；

（2） 如果最后一个方程含有未知数：方程组这时是有解的。

a) 如果方程个数小于未知数个数：由于约束未知量个数等于方程组中的方程个数，因此此时方程组中有自由未知量，这些约束未知量每取一组值，就可以计算得到约束未知量的值；由于约束未知量可以自由取值，因此此时方程组有无穷多解；

b) 如果方程个数等于未知数个数：每个未知数都是约束未知量，方程组中没有自由未知量。在该阶梯形方程组中，最后一个方程含有一个未知数，倒数第二个方程含有两个未知数，……，第一个方程含有所有的未知数。从最后一个方程求解得到最后一个未知数的值，从倒数第二个方程求解得到倒数第二个未知数的值，……，从第一个方程求解得到第一个未知数的值，因此该方程组只有唯一解。

（注：在上述讨论中，如果将方程组用高斯－约当消元法得到与之等价的方程组，则更便于讨论，读者可自行思考。）

根据如上结论，如果线性方程组的未知数个数大于方程个数，那么该方程组是不可能有唯一解的：或者无解，或者有无穷多解。只有当线性方程组的未知数个数不超过方程个数，该方程组才可能有唯一解。在下一章中，我们将讨论一种特殊情况：线性方程组的未知数个数等于方程个数、且该方程组有唯一解的情形。

这一章的主要内容为：

（1） 引入了线性方程、线性方程组及其解的概念；线性方程组的矩阵表示；

（2） 高斯消元法、高斯－约当法求解线性方程组；知道如何判断线性方程组解的情况。

习题 1

1、用高斯消元法、高斯－约当法求解下列线性方程组

（1）$\begin{cases} x_1+2x_2+x_3+4x_4=3 \\ 3x_1+7x_2+3x_3+15x_4=14 \\ 2x_1+2x_2+2x_3+3x_4=-2 \end{cases}$；

（2）$\begin{cases} 5x_1+5x_2+2x_3+3x_4=18 \\ 4x_1+5x_2-2x_3-2x_4=9 \\ -2x_1+2x_3-x_4=3 \\ 2x_2+4x_3-x_4=15 \end{cases}$；

（3）$\begin{cases} x_1+4x_2+3x_3+3x_4=19 \\ 5x_1+2x_3-2x_4=3 \\ -x_1+5x_2+x_3=14 \\ 2x_1+x_2-2x_3-x_4=9 \\ 5x_1-x_2+x_3=9 \end{cases}$；

（4）$\begin{cases} x_1+x_2+x_3+4x_4-3x_5=0 \\ 2x_1+x_2+3x_3+5x_4-5x_5=0 \\ x_1-x_2+3x_3-2x_4-x_5=0 \\ 3x_1+x_2+5x_3+6x_4-7x_5=0 \end{cases}$

2、求方程组$\begin{cases} x_1+x_2=0 \\ x_2-x_4=0 \end{cases}$与$\begin{cases} x_1-x_2+x_4=0 \\ x_2-x_3+x_4=0 \end{cases}$的公共解.

3、试问 t 取什么值时，线性方程组

$$\begin{cases} x_1-x_2+2x_3=-4 \\ x_1+x_2+tx_3=4 \\ x_1-tx_2-x_3=-t^2 \end{cases}$$

无解，有唯一解，有无穷多解.

第二章　行列式

在这一章中，我们将首先介绍方程个数等于未知数个数、且有唯一解的线性方程组的解的表达式，由此引入低阶行列式的概念与性质；然后在此基础上介绍一般的 n 阶行列式，并讨论其性质；最后应用行列式的性质给出克莱姆法则。在学习完本章后，应该做到

（1）　掌握行列式的定义与性质，能够给出行列式性质的证明；

（2）　掌握基本的行列式计算方法；

（3）　掌握克莱姆法则的基本应用方法；

行列式是非常重要的概念，希望同学们能够切实加以掌握，必须做到能够熟练使用教材上的符号与表达方式。

§2.1　特殊线性方程组的公式解

前面所介绍的高斯消元法可求解任意的线性方程组。但是正如一元二次方程

$$ax^2+bx+c=0 \qquad a\neq 0$$

我们知道可用配方法来求解，

$$a\left(x^2+\frac{b}{a}x\right)=-c$$

$$a\left[\left(x+\frac{b}{2a}\right)^2-\frac{b^2}{4a^2}\right]=-c$$

$$\left(x+\frac{b}{2a}\right)^2=\frac{b^2}{4a^2}-\frac{c}{a}=\frac{b^2-4ac}{4a^2}$$

所以，当 $b^2-4ac\geq 0$ 时，两边开方

$$x+\frac{b}{2a}=\pm\sqrt{\frac{b^2-4ac}{4a^2}}=\pm\frac{\sqrt{b^2-4ac}}{2a}$$

所以　$x=\dfrac{-b\pm\sqrt{b^2-4ac}}{2a}$。

显然我们没必要用配方法求解一元二次方程，直接用公式即可；并且，从推导过程我们还可得到有意义的结论。同样的道理，如果我们能够得到线性方程组的公式解，也可能是的求解线性方程组更为简单，并且还可能得到有意义的其它结论。但是，并不是所有的线性方程组都有公式解，而只有线性方程组的未知数个数与方程个数相等的时候才可能，在本节中我们只考虑这种情形，且设未知数个数为 n，因此线性方程组为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \\ \qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \tag{2.1}$$

现在我们利用高斯消元法来求公式解。

2.1.1　当 $n=1$ 时，线性方程组(2.1)为

$$a_{11}x_1 = b_1$$

根据上一章得到的结果可知：它有唯一解的充要条件是 $a_{11} \neq 0$，此时线性方程唯一解为 $x_1 = \dfrac{b_1}{a_{11}}$。

2.1.2　当 $n=2$ 时，线性方程组(2.1)为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{2.2}$$

我们总可设 $a_{11} \neq 0$（这是因为：我们要求该线性方程组有两个未知数，因此该方程组中未知数 x_1 的两个系数 a_{11}, a_{21} 不能都为零；如果 $a_{11} = 0$，则说明 $a_{21} \neq 0$，此时我们可互换这两个方程的位置，得到一个新方程组，在该新方程组中，第一个方程中未知数 x_1 前的系数就不为零；而这两个方程组有相同的解，因此我们就可把该新方程组作为我们的讨论对象），我们用高斯消元法将其化为阶梯形方程组

$$\begin{cases} a_{11}x_1 \qquad\qquad + a_{12}x_2 = b_1 \\ \qquad (a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1 \end{cases} \tag{2.3}$$

由于 $a_{11} \neq 0$，易知方程组(2.3)有唯一解的充要条件是：第二个方程

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1 \tag{2.4}$$

有唯一解。根据第一章中的讨论知，方程(2.4)有唯一解的充要条件是

$$a_{11}a_{22} - a_{12}a_{21} \neq 0$$

此时有

$$x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}$$

我们可以把 x_2 的值代入第一个方程，然后求解出 x_1。当然，我们也可用第一个方程消掉第二个方程中的未知数 x_2，

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ (a_{12}a_{21} - a_{11}a_{22})x_1 = a_{12}b_2 - a_{22}b_1 \end{cases}$$

故知当 $a_{11}a_{22} - a_{12}a_{21} \neq 0$ 时，x_1 有唯一解

$$x_1 = \frac{a_{12}b_2 - a_{22}b_1}{a_{12}a_{21} - a_{11}a_{22}} = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}$$

总结上述结论，我们有：线性方程组(2.2)有唯一解的充要条件是 $a_{11}a_{22} - a_{12}a_{21} \neq 0$，此时该方程组的解可表示为

$$\begin{cases} x_1 = \dfrac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}} \\ x_2 = \dfrac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}} \end{cases} \tag{2.5}$$

其分母中所出现的是线性方程组(2.6)中未知数前的四个系数，而分子与分母的区别与联系也是比较显然的：在 x_1 的表达式中，将分母中的 a_{11} 换成 b_1、a_{21} 换成 b_2，就得到分子；在 x_2 的表达式中，将分母中的 a_{22} 换成 b_2、a_{12} 换成 b_1，就得到分子。因此，如果我们定义四元函数

$$f(t_{11}, t_{12}, t_{21}, t_{22}) = t_{11}t_{22} - t_{12}t_{21}$$

那么解可表示成

$$\begin{cases} x_1 = \dfrac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}} = \dfrac{f(b_1, a_{12}, b_2, a_{22})}{f(a_{11}, a_{12}, a_{21}, a_{22})} \\ x_2 = \dfrac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}} = \dfrac{f(a_{11}, b_1, a_{21}, b_2)}{f(a_{11}, a_{12}, a_{21}, a_{22})} \end{cases}$$

由于该函数的分母只与线性方程组(2.2)的未知数前的系数有关，为了强调这一关系，我们将该函数用另外一种直观的方式来表示

$$f(t_{11}, t_{12}, t_{21}, t_{22}) = \begin{vmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{vmatrix} = t_{11}t_{22} - t_{12}t_{21} \tag{2.7}$$

因此，解就表示为

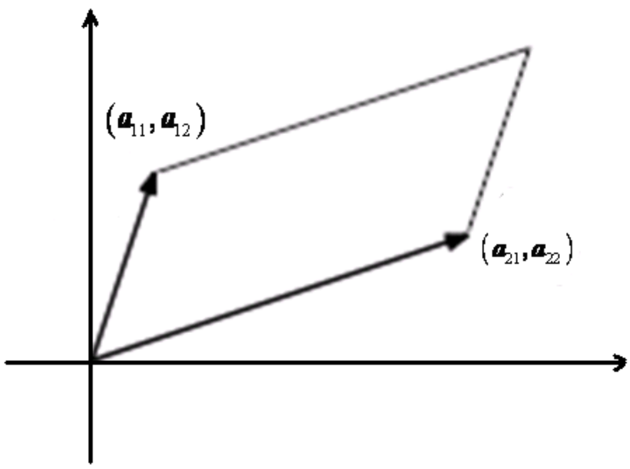
$$
\begin{cases}
x_1 = \dfrac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}} = \dfrac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \\[2ex]
x_2 = \dfrac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}} = \dfrac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}
\end{cases} \tag{2.8}
$$

我们把四元函数(2.7)称为**二阶行列式**。

　　请注意解的表达式(2.8)与方程组（2.3）中各个系数的关系。

二阶行列式的几何意义：

　　设(a_{11},a_{12})、(a_{21},a_{22})为平面上的两个点，则



则该平行四边形的面积就为

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

的绝对值。该命题的证明略。

假设线性方程组

$$
\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{2.9}
$$

有唯一解，对该方程组做一次线性方程组的初等变换，就得到如下三个线性方程组

$$
\begin{cases} a_{21}x_1 + a_{22}x_2 = b_2 \\ a_{11}x_1 + a_{12}x_2 = b_1 \end{cases} \tag{2.10}
$$

$$
\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ ka_{21}x_1 + ka_{22}x_2 = kb_2 \end{cases} \quad , \quad k \neq 0 \tag{2.11}
$$

作者：卢世荣　lshrlshr@163.com　　欢迎任何意见与建议！

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ (a_{21}+ka_{11})x_1 + (a_{22}+ka_{12})x_2 = b_2 + kb_1 \end{cases} \tag{2.12}$$

根据第一章的结论可知，这四个方程组(2.9)～(2.12)有相同的解，因此这四个方程组都有唯一解，根据以上部分的讨论可知，这四个方程组的解中，未知数 x_1 的值依次为

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, x_1 = \frac{\begin{vmatrix} b_2 & a_{22} \\ b_1 & a_{12} \end{vmatrix}}{\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}}, x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ kb_2 & ka_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ ka_{21} & ka_{22} \end{vmatrix}}, x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2+kb_1 & a_{22}+ka_{12} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21}+ka_{11} & a_{22}+ka_{12} \end{vmatrix}}$$

因此这四个解必须相等，就是有

$$\frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{\begin{vmatrix} b_2 & a_{22} \\ b_1 & a_{12} \end{vmatrix}}{\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}} = \frac{\begin{vmatrix} b_1 & a_{12} \\ kb_2 & ka_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ ka_{21} & ka_{22} \end{vmatrix}} = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2+kb_1 & a_{22}+ka_{12} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21}+ka_{11} & a_{22}+ka_{12} \end{vmatrix}}$$

首先考虑 $\frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{\begin{vmatrix} b_2 & a_{22} \\ b_1 & a_{12} \end{vmatrix}}{\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}}$，显然等号右边分母行列式是等号左边分母行列式互换两行得到的，分子等号右边的分子行列式也是等号左边分子行列式互换两行得到的，根据二阶行列式的定义，容易计算的

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

因此这两个分子行列式也有 $\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = -\begin{vmatrix} b_2 & a_{22} \\ b_1 & a_{12} \end{vmatrix}$，这就保证了 $\frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{\begin{vmatrix} b_2 & a_{22} \\ b_1 & a_{12} \end{vmatrix}}{\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}}$ 成立。

另外两种情况我们容易验证

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

因此对于二阶行列式，我们可得到其简单的性质：

（1）互换二阶行列式的两行，行列式改变符号；

（2）二阶行列式的一行乘以一个数 k，则行列式结果为原来的 k 倍；

（3）将二阶行列式的一行的某个倍数加到另一行上，不改变行列式的值。

2.1.3　当 $n=3$ 时，线性方程组(2.1)为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \tag{2.13}$$

与 $n = 2$ 时类似的道理，我们可假设 $a_{11} \neq 0$，利用高斯消元法，我们消掉后面两个方程中的未知数 x_1，得到

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ (a_{22}a_{11} - a_{12}a_{21})x_2 + (a_{23}a_{11} - a_{13}a_{21})x_3 = b_2a_{11} - b_1a_{21} \\ (a_{32}a_{11} - a_{12}a_{31})x_2 + (a_{33}a_{11} - a_{13}a_{31})x_3 = b_3a_{11} - b_1a_{31} \end{cases} \tag{2.14}$$

显然，线性方程组(2.14)有唯一解的充分必要条件是：由后面两个方程所组成的方程组

$$\begin{cases} (a_{22}a_{11} - a_{12}a_{21})x_2 + (a_{23}a_{11} - a_{13}a_{21})x_3 = b_2a_{11} - b_1a_{21} \\ (a_{32}a_{11} - a_{12}a_{31})x_2 + (a_{33}a_{11} - a_{13}a_{31})x_3 = b_3a_{11} - b_1a_{31} \end{cases} \tag{2.15}$$

有唯一解！由于该方程组只有两个未知数，因此，利用 $n = 2$ 时的结果，当且仅当

$$\begin{aligned} &\begin{vmatrix} (a_{22}a_{11} - a_{12}a_{21}) & (a_{23}a_{11} - a_{13}a_{21}) \\ (a_{32}a_{11} - a_{12}a_{31}) & (a_{33}a_{11} - a_{13}a_{31}) \end{vmatrix} \\ &= a_{22}a_{11}a_{33}a_{11} - a_{22}a_{11}a_{13}a_{31} - a_{12}a_{21}a_{33}a_{11} + a_{12}a_{21}a_{13}a_{31} \\ &-(a_{32}a_{11}a_{23}a_{11} - a_{32}a_{11}a_{13}a_{21} - a_{12}a_{31}a_{23}a_{11} + a_{12}a_{31}a_{13}a_{21}) \\ &= a_{22}a_{11}a_{33}a_{11} - a_{22}a_{11}a_{13}a_{31} - a_{12}a_{21}a_{33}a_{11} - a_{32}a_{11}a_{23}a_{11} \\ &\quad + a_{32}a_{11}a_{13}a_{21} + a_{12}a_{31}a_{23}a_{11} \\ &= a_{11}(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}) \\ &\neq 0 \end{aligned}$$

时，该方程组(2.15)有唯一解，因此原方程组(2.13)也就有唯一解。且 x_3 解为

$$x_3 = \frac{\begin{vmatrix} a_{23}a_{11} - a_{13}a_{21} & b_2a_{11} - b_1a_{21} \\ a_{33}a_{11} - a_{13}a_{31} & b_3a_{11} - b_1a_{31} \end{vmatrix}}{\begin{vmatrix} a_{22}a_{11} - a_{12}a_{21} & a_{23}a_{11} - a_{13}a_{21} \\ a_{32}a_{11} - a_{12}a_{31} & a_{33}a_{11} - a_{13}a_{31} \end{vmatrix}}$$

现在，我们比较一下上式中分子、分母的区别与联系，可注意到：如果将分母中的 a_{13}, a_{23}, a_{33} 分别换为 b_1, b_2, b_3，我们就可得到分子。根据二阶行列式的定义可得

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

所以

$$x_3=\frac{\begin{vmatrix} a_{23}a_{11}-a_{13}a_{21} & b_2a_{11}-b_1a_{21} \\ a_{33}a_{11}-a_{13}a_{31} & b_3a_{11}-b_1a_{31} \end{vmatrix}}{\begin{vmatrix} a_{22}a_{11}-a_{12}a_{21} & a_{23}a_{11}-a_{13}a_{21} \\ a_{32}a_{11}-a_{12}a_{31} & a_{33}a_{11}-a_{13}a_{31} \end{vmatrix}}$$
$$=\frac{a_{11}\left(a_{11}a_{22}b_3+a_{12}b_2a_{31}+b_1a_{21}a_{32}-a_{11}b_2a_{32}-a_{12}a_{21}b_3-b_1a_{22}a_{31}\right)}{a_{11}\left(a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}\right)}$$
$$=\frac{a_{11}a_{22}b_3+a_{12}b_2a_{31}+b_1a_{21}a_{32}-a_{11}b_2a_{32}-a_{12}a_{21}b_3-b_1a_{22}a_{31}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}$$

因此，如果定义 9 元函数

$$f\left(t_{11},t_{12},t_{13},t_{21},t_{22},t_{23},t_{31},t_{32},t_{33}\right)$$
$$=t_{11}t_{22}t_{33}+t_{12}t_{23}t_{31}+t_{13}t_{21}t_{32}-t_{11}t_{23}t_{32}-t_{12}t_{21}t_{33}-t_{13}t_{22}t_{31}$$

则 x_3 可表示为

$$x_3=\frac{a_{11}a_{22}b_3+a_{12}b_2a_{31}+b_1a_{21}a_{32}-a_{11}b_2a_{32}-a_{12}a_{21}b_3-b_1a_{22}a_{31}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}$$
$$=\frac{f\left(a_{11},a_{12},b_1,a_{21},a_{22},b_2,a_{31},a_{32},b_3\right)}{f\left(a_{11},a_{12},a_{13},a_{21},a_{22},a_{23},a_{31},a_{32},a_{33}\right)}$$

上式中的分母部分只与线性方程组(2.13)的未知数前的系数有关，因此也用一种直观的方式来表示该函数，定义 9 元函数

$$f\left(t_{11},t_{12},t_{13},t_{21},t_{22},t_{23},t_{31},t_{32},t_{33}\right)$$
$$=\begin{vmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{vmatrix} \tag{2.16}$$
$$=t_{11}t_{22}t_{33}+t_{12}t_{23}t_{31}+t_{13}t_{21}t_{32}-t_{11}t_{23}t_{32}-t_{12}t_{21}t_{33}-t_{13}t_{22}t_{31}$$

则 x_3 可表示为

$$x_3=\frac{a_{11}a_{22}b_3+a_{12}b_2a_{31}+b_1a_{21}a_{32}-a_{11}b_2a_{32}-a_{12}a_{21}b_3-b_1a_{22}a_{31}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}$$
$$=\frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \tag{2.17}$$

把上述对 x_3 的讨论应用于 x_1,x_2，就得到线性方程组(2.13)的解

$$x_1=\frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}},\quad x_2=\frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}},\quad x_3=\frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \tag{2.18}$$

显然，将解与线性方程组

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

相比较，我们可得到线性方程组的解的规律：每个未知数的解都是分数形式，分母为三阶行列式$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$，而 x_i 表达式中的分子部分则是将分母行列式的第 i 列（也就是三个方程中未知数 x_i 前的系数）换为方程组等号右边的三个常数 b_1, b_2, b_3 得到的结果。

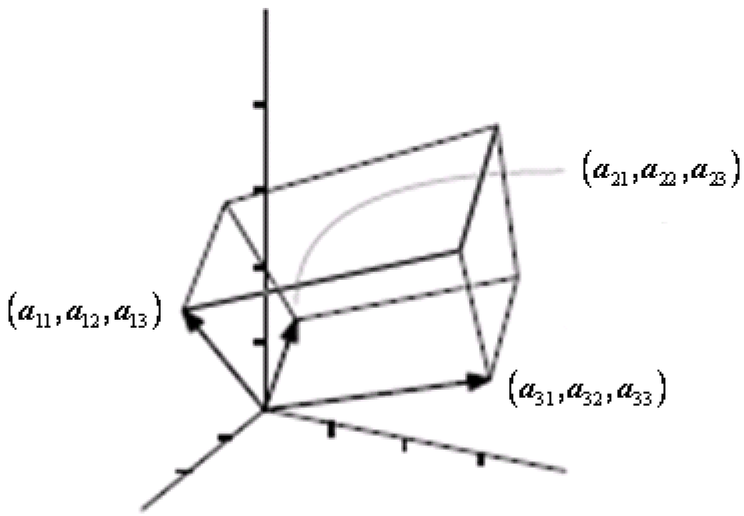
我们把 9 元函数(2.16)称为**三阶行列式**。

三阶行列式的几何意义：

设(a_{11}, a_{12}, a_{13})，(a_{21}, a_{22}, a_{23})，(a_{31}, a_{32}, a_{33})为空间直角坐标系中的三个点，则由这三条边所确定的平行六面体的体积为行列式

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

的绝对值。该结论的证明略。



我们可以讨论三阶行列式的简单性质。根据以上分析，对于线性方程组

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

若其系数行列式

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

则方程组有唯一解，且

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

线性方程组的初等变换不改变方程组的解，因此

$$\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3=b_1\\ a_{21}x_1+a_{22}x_2+a_{23}x_3=b_2\\ a_{31}x_1+a_{32}x_2+a_{33}x_3=b_3 \end{cases} \tag{2.19}$$

$$\begin{cases} a_{21}x_1+a_{22}x_2+a_{23}x_3=b_2\\ a_{11}x_1+a_{12}x_2+a_{13}x_3=b_1\\ a_{31}x_1+a_{32}x_2+a_{33}x_3=b_3 \end{cases} \tag{2.20}$$

$$\begin{cases} ka_{11}x_1 + ka_{12}x_2 + ka_{13}x_3 = kb_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \quad k \neq 0 \tag{2.21}$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ (a_{21}+ka_{11})x_1 + (a_{22}+ka_{12})x_2 + (a_{23}+ka_{13})x_3 = b_2 + kb_1 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \tag{2.22}$$

根据上述结论，这四个方程组(2.19)～(2.22)的解中，x_1的值分别为

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \tag{2.23}$$

$$x_1 = \frac{\begin{vmatrix} b_2 & a_{22} & a_{23} \\ b_1 & a_{12} & a_{13} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \tag{2.24}$$

$$x_1 = \frac{\begin{vmatrix} kb_1 & ka_{12} & ka_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \tag{2.25}$$

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 + kb_1 & a_{22} + kb_2 & a_{23} + kb_3 \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} + ka_{11} & a_{22} + ka_{12} & a_{23} + ka_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \tag{2.26}$$

由于这四个线性方程组等价，而等价的线性方程组有相同的解，因此就有

$$x_1=\frac{\begin{vmatrix}b_1 & a_{12} & a_{13}\\ b_2 & a_{22} & a_{23}\\ b_3 & a_{32} & a_{33}\end{vmatrix}}{\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}}=\frac{\begin{vmatrix}b_2 & a_{22} & a_{23}\\ b_1 & a_{12} & a_{13}\\ b_3 & a_{32} & a_{33}\end{vmatrix}}{\begin{vmatrix}a_{21} & a_{22} & a_{23}\\ a_{11} & a_{12} & a_{13}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}}=\frac{\begin{vmatrix}kb_1 & ka_{12} & ka_{13}\\ b_2 & a_{22} & a_{23}\\ b_3 & a_{32} & a_{33}\end{vmatrix}}{\begin{vmatrix}ka_{11} & ka_{12} & ka_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}}$$

$$=\frac{\begin{vmatrix}b_1 & a_{12} & a_{13}\\ b_2+kb_1 & a_{22}+kb_2 & a_{23}+kb_3\\ b_3 & a_{32} & a_{33}\end{vmatrix}}{\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21}+ka_{11} & a_{22}+ka_{12} & a_{23}+ka_{13}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}}$$

从上式可知道三阶行列式的一些性质。(2.24)的分母行列式是互换(2.23)的分母行列式的第 1、2 行的结果，通过定义(2.16)容易验证，有

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

所以对应的分子行列式有

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

所以相应的比值保持不变。实际上，通过定义(2.16)容易验证：

互换三阶行列式的任意两行，行列式改变符号；

(2.25)的分母行列式是(2.23)的分母行列式第一行的每个数都乘以常数k的结果，通过定义(2.16)可以验证

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

故比值保持不变。实际上，以常数k乘以三阶行列式某一行的每个数，所得的行列式是原来行列式的k倍。这也可通过定义(2.16)直接验证。

类似地，(2.26)的分母行列式是将(2.23)的分母行列式的第一行的每个数都乘以常数k、然后加到第二行对应数上得到的，用定义(2.16)可以验证，有

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

对应的分子行列式也有相同的结果，因此比值仍然保持不变。实际上，如果将三阶行列式某

一行的每个数乘以k倍后加到另外一行上去，所得的行列式与原来的行列式有相同的值。

根据上述讨论，我们可得三阶行列式的有如下性质：
（1）互换三阶行列式的两行，行列式改变符号；
（2）三阶行列式的一行乘以一个数k，则行列式结果为原来的k倍；
（3）将三阶行列式的一行的某个倍数加到另一行上，不改变行列式的值。

2.1.4　$n=4$的情况（*）

此时线性方程组(2.1)就为

$$\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3+a_{14}x_4=b_1 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3+a_{24}x_4=b_2 \\ a_{31}x_1+a_{32}x_2+a_{33}x_3+a_{34}x_4=b_3 \\ a_{41}x_1+a_{42}x_2+a_{43}x_3+a_{44}x_4=b_4 \end{cases} \tag{2.27}$$

同理可假设$a_{11}\neq 0$，利用高斯消元法，我们消掉后面三个方程中的未知数x_1，得到

$$\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3+a_{14}x_4=b_1 \\ \quad (a_{22}a_{11}-a_{12}a_{21})x_2+(a_{23}a_{11}-a_{13}a_{21})x_3+(a_{24}a_{11}-a_{14}a_{21})x_4=b_2a_{11}-b_1a_{21} \\ \quad (a_{32}a_{11}-a_{12}a_{31})x_2+(a_{33}a_{11}-a_{13}a_{31})x_3+(a_{34}a_{11}-a_{14}a_{31})x_4=b_3a_{11}-b_1a_{31} \\ \quad (a_{42}a_{11}-a_{12}a_{41})x_2+(a_{43}a_{11}-a_{13}a_{41})x_3+(a_{44}a_{11}-a_{14}a_{41})x_4=b_4a_{11}-b_1a_{41} \end{cases} \tag{2.28}$$

因此，整个线性方程组有唯一解的充要条件是：含有三个未知数、三个方程的线性方程组

$$\begin{cases} (a_{22}a_{11}-a_{12}a_{21})x_2+(a_{23}a_{11}-a_{13}a_{21})x_3+(a_{24}a_{11}-a_{14}a_{21})x_4=b_2a_{11}-b_1a_{21} \\ (a_{32}a_{11}-a_{12}a_{31})x_2+(a_{33}a_{11}-a_{13}a_{31})x_3+(a_{34}a_{11}-a_{14}a_{31})x_4=b_3a_{11}-b_1a_{31} \\ (a_{42}a_{11}-a_{12}a_{41})x_2+(a_{43}a_{11}-a_{13}a_{41})x_3+(a_{44}a_{11}-a_{14}a_{41})x_4=b_4a_{11}-b_1a_{41} \end{cases} \tag{2.29}$$

有唯一解。利用$n=3$时的结果，因此当

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

$$\begin{aligned} &=(a_{22}a_{11}-a_{12}a_{21})(a_{33}a_{11}-a_{13}a_{31})(a_{44}a_{11}-a_{14}a_{41}) \\ &+(a_{23}a_{11}-a_{13}a_{21})(a_{34}a_{11}-a_{14}a_{31})(a_{42}a_{11}-a_{12}a_{41}) \\ &+(a_{24}a_{11}-a_{14}a_{21})(a_{32}a_{11}-a_{12}a_{31})(a_{43}a_{11}-a_{13}a_{41}) \\ &-(a_{22}a_{11}-a_{12}a_{21})(a_{34}a_{11}-a_{14}a_{31})(a_{43}a_{11}-a_{13}a_{41}) \\ &-(a_{23}a_{11}-a_{13}a_{21})(a_{32}a_{11}-a_{12}a_{31})(a_{44}a_{11}-a_{14}a_{41}) \\ &-(a_{24}a_{11}-a_{14}a_{21})(a_{33}a_{11}-a_{13}a_{31})(a_{42}a_{11}-a_{12}a_{41}) \\ &\neq 0 \end{aligned}$$

时，该线性方程组有唯一解。由于该展开式中共有 48 项，为了便于合并同类项，我们将展开式每项中的元素首先按行标、然后按列标排序，例如$a_{22}a_{11}a_{13}a_{31}a_{14}a_{41}$排序成

$a_{11}a_{13}a_{14}a_{22}a_{31}a_{41}$，则

$$
\begin{aligned}
&(a_{22}a_{11}-a_{12}a_{21})(a_{33}a_{11}-a_{13}a_{31})(a_{44}a_{11}-a_{14}a_{41})\\
&+(a_{23}a_{11}-a_{13}a_{21})(a_{34}a_{11}-a_{14}a_{31})(a_{42}a_{11}-a_{12}a_{41})\\
&+(a_{24}a_{11}-a_{14}a_{21})(a_{32}a_{11}-a_{12}a_{31})(a_{43}a_{11}-a_{13}a_{41})\\
&-(a_{22}a_{11}-a_{12}a_{21})(a_{34}a_{11}-a_{14}a_{31})(a_{43}a_{11}-a_{13}a_{41})\\
&-(a_{23}a_{11}-a_{13}a_{21})(a_{32}a_{11}-a_{12}a_{31})(a_{44}a_{11}-a_{14}a_{41})\\
&-(a_{24}a_{11}-a_{14}a_{21})(a_{33}a_{11}-a_{13}a_{31})(a_{42}a_{11}-a_{12}a_{41})\\
&=(a_{11}{}^3a_{22}a_{33}a_{44}-a_{11}{}^2a_{14}a_{22}a_{33}a_{41}-a_{11}{}^2a_{13}a_{22}a_{31}a_{44}+a_{11}a_{13}a_{14}a_{22}a_{31}a_{41}\\
&-a_{11}{}^2a_{12}a_{21}a_{33}a_{44}+a_{11}a_{12}a_{14}a_{21}a_{33}a_{41}+a_{11}a_{12}a_{13}a_{21}a_{31}a_{44}-a_{12}a_{13}a_{14}a_{21}a_{31}a_{41})\\
&+(a_{11}{}^3a_{23}a_{34}a_{42}-a_{11}{}^2a_{12}a_{23}a_{34}a_{41}-a_{11}{}^2a_{14}a_{23}a_{31}a_{42}+a_{11}a_{12}a_{14}a_{23}a_{31}a_{41}\\
&-a_{11}{}^2a_{13}a_{21}a_{34}a_{42}+a_{11}a_{12}a_{13}a_{21}a_{34}a_{41}+a_{11}a_{13}a_{14}a_{21}a_{31}a_{42}-a_{12}a_{13}a_{14}a_{21}a_{31}a_{41})\\
&+(a_{11}{}^3a_{24}a_{32}a_{43}-a_{11}{}^2a_{13}a_{24}a_{32}a_{41}-a_{11}{}^2a_{12}a_{24}a_{31}a_{43}+a_{11}a_{12}a_{13}a_{24}a_{31}a_{41}\\
&-a_{11}{}^2a_{14}a_{21}a_{32}a_{43}+a_{11}a_{13}a_{14}a_{21}a_{32}a_{41}+a_{11}a_{12}a_{14}a_{21}a_{31}a_{43}-a_{12}a_{13}a_{14}a_{21}a_{31}a_{41})\\
&-(a_{11}{}^3a_{22}a_{34}a_{43}-a_{11}{}^2a_{13}a_{22}a_{34}a_{41}-a_{11}{}^2a_{14}a_{22}a_{31}a_{43}+a_{11}a_{13}a_{14}a_{22}a_{31}a_{41}\\
&-a_{11}{}^2a_{12}a_{21}a_{34}a_{43}+a_{11}a_{12}a_{13}a_{21}a_{34}a_{41}+a_{11}a_{12}a_{14}a_{21}a_{31}a_{43}-a_{12}a_{13}a_{14}a_{21}a_{31}a_{41})\\
&-(a_{11}{}^3a_{23}a_{32}a_{44}-a_{11}{}^2a_{14}a_{23}a_{32}a_{41}-a_{11}{}^2a_{12}a_{23}a_{31}a_{44}+a_{11}a_{12}a_{14}a_{23}a_{31}a_{41}\\
&-a_{11}{}^2a_{13}a_{21}a_{32}a_{44}+a_{11}a_{13}a_{14}a_{21}a_{32}a_{41}+a_{11}a_{12}a_{13}a_{21}a_{31}a_{44}-a_{12}a_{13}a_{14}a_{21}a_{31}a_{41})\\
&-(a_{11}{}^3a_{24}a_{33}a_{42}-a_{11}{}^2a_{12}a_{24}a_{33}a_{41}-a_{11}{}^2a_{13}a_{24}a_{31}a_{42}+a_{11}a_{12}a_{13}a_{24}a_{31}a_{41}\\
&-a_{11}{}^2a_{14}a_{21}a_{33}a_{42}+a_{11}a_{12}a_{14}a_{21}a_{33}a_{41}+a_{11}a_{13}a_{14}a_{21}a_{31}a_{42}-a_{12}a_{13}a_{14}a_{21}a_{31}a_{41})
\end{aligned}
$$

$$
\begin{aligned}
&=a_{11}{}^2(a_{11}a_{22}a_{33}a_{44}-a_{14}a_{22}a_{33}a_{41}-a_{13}a_{22}a_{31}a_{44}-a_{12}a_{21}a_{33}a_{44}\\
&+a_{11}a_{23}a_{34}a_{42}-a_{12}a_{23}a_{34}a_{41}-a_{14}a_{23}a_{31}a_{42}-a_{13}a_{21}a_{34}a_{42}\\
&+a_{11}a_{24}a_{32}a_{43}-a_{13}a_{24}a_{32}a_{41}-a_{12}a_{24}a_{31}a_{43}-a_{14}a_{21}a_{32}a_{43}\\
&-a_{11}a_{22}a_{34}a_{43}+a_{13}a_{22}a_{34}a_{41}+a_{14}a_{22}a_{31}a_{43}+a_{12}a_{21}a_{34}a_{43}\\
&-a_{11}a_{23}a_{32}a_{44}+a_{14}a_{23}a_{32}a_{41}+a_{12}a_{23}a_{31}a_{44}+a_{13}a_{21}a_{32}a_{44}\\
&-a_{11}a_{24}a_{33}a_{42}+a_{12}a_{24}a_{33}a_{41}+a_{13}a_{24}a_{31}a_{42}+a_{14}a_{21}a_{33}a_{42})
\end{aligned}
$$

$$
\begin{aligned}
&=a_{11}{}^2(a_{11}a_{22}a_{33}a_{44}-a_{14}a_{22}a_{33}a_{41}-a_{13}a_{22}a_{31}a_{44}-a_{12}a_{21}a_{33}a_{44}\\
&+a_{11}a_{23}a_{34}a_{42}-a_{12}a_{23}a_{34}a_{41}-a_{14}a_{23}a_{31}a_{42}-a_{13}a_{21}a_{34}a_{42}\\
&+a_{11}a_{24}a_{32}a_{43}-a_{13}a_{24}a_{32}a_{41}-a_{12}a_{24}a_{31}a_{43}-a_{14}a_{21}a_{32}a_{43}\\
&-a_{11}a_{22}a_{34}a_{43}+a_{13}a_{22}a_{34}a_{41}+a_{14}a_{22}a_{31}a_{43}+a_{12}a_{21}a_{34}a_{43}\\
&-a_{11}a_{23}a_{32}a_{44}+a_{14}a_{23}a_{32}a_{41}+a_{12}a_{23}a_{31}a_{44}+a_{13}a_{21}a_{32}a_{44}\\
&-a_{11}a_{24}a_{33}a_{42}+a_{12}a_{24}a_{33}a_{41}+a_{13}a_{24}a_{31}a_{42}+a_{14}a_{21}a_{33}a_{42})
\end{aligned}
$$

由于$a_{11}\neq 0$，因此，当

$$
\begin{aligned}
&a_{11}a_{22}a_{33}a_{44}-a_{14}a_{22}a_{33}a_{41}-a_{13}a_{22}a_{31}a_{44}-a_{12}a_{21}a_{33}a_{44}\\
&+a_{11}a_{23}a_{34}a_{42}-a_{12}a_{23}a_{34}a_{41}-a_{14}a_{23}a_{31}a_{42}-a_{13}a_{21}a_{34}a_{42}\\
&+a_{11}a_{24}a_{32}a_{43}-a_{13}a_{24}a_{32}a_{41}-a_{12}a_{24}a_{31}a_{43}-a_{14}a_{21}a_{32}a_{43}\\
&-a_{11}a_{22}a_{34}a_{43}+a_{13}a_{22}a_{34}a_{41}+a_{14}a_{22}a_{31}a_{43}+a_{12}a_{21}a_{34}a_{43}\\
&-a_{11}a_{23}a_{32}a_{44}+a_{14}a_{23}a_{32}a_{41}+a_{12}a_{23}a_{31}a_{44}+a_{13}a_{21}a_{32}a_{44}\\
&-a_{11}a_{24}a_{33}a_{42}+a_{12}a_{24}a_{33}a_{41}+a_{13}a_{24}a_{31}a_{42}+a_{14}a_{21}a_{33}a_{42}\neq 0
\end{aligned}
$$

时，线性方程组有唯一解，因此原线性方程组有唯一解。且 x_2 的值为

$$
x_2=\frac{\begin{vmatrix} b_2a_{11}-b_1a_{21} & a_{23}a_{11}-a_{13}a_{21} & a_{24}a_{11}-a_{14}a_{21}\\ b_3a_{11}-b_1a_{31} & a_{33}a_{11}-a_{13}a_{31} & a_{34}a_{11}-a_{14}a_{31}\\ b_4a_{11}-b_1a_{41} & a_{43}a_{11}-a_{13}a_{41} & a_{44}a_{11}-a_{14}a_{41}\end{vmatrix}}{\begin{vmatrix} a_{22}a_{11}-a_{12}a_{21} & a_{23}a_{11}-a_{13}a_{21} & a_{24}a_{11}-a_{14}a_{21}\\ a_{32}a_{11}-a_{12}a_{31} & a_{33}a_{11}-a_{13}a_{31} & a_{34}a_{11}-a_{14}a_{31}\\ a_{42}a_{11}-a_{12}a_{41} & a_{43}a_{11}-a_{13}a_{41} & a_{44}a_{11}-a_{14}a_{41}\end{vmatrix}}
$$

比较上式中的分子、分母，可知将分母中的 $a_{12},a_{22},a_{32},a_{42}$ 分别换为 b_1,b_2,b_3,b_4，我们就得到了分子。由于分母是一个三阶行列式，其展开后含有四个方程中各个未知数前的系数，因此，如果我们定义四阶行列式

$$
\begin{aligned}
&\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\\ a_{41} & a_{42} & a_{43} & a_{44}\end{vmatrix}\\
&=a_{11}a_{22}a_{33}a_{44}-a_{14}a_{22}a_{33}a_{41}-a_{13}a_{22}a_{31}a_{44}-a_{12}a_{21}a_{33}a_{44}\\
&+a_{11}a_{23}a_{34}a_{42}-a_{12}a_{23}a_{34}a_{41}-a_{14}a_{23}a_{31}a_{42}-a_{13}a_{21}a_{34}a_{42}\\
&+a_{11}a_{24}a_{32}a_{43}-a_{13}a_{24}a_{32}a_{41}-a_{12}a_{24}a_{31}a_{43}-a_{14}a_{21}a_{32}a_{43}\\
&-a_{11}a_{22}a_{34}a_{43}+a_{13}a_{22}a_{34}a_{41}+a_{14}a_{22}a_{31}a_{43}+a_{12}a_{21}a_{34}a_{43}\\
&-a_{11}a_{23}a_{32}a_{44}+a_{14}a_{23}a_{32}a_{41}+a_{12}a_{23}a_{31}a_{44}+a_{13}a_{21}a_{32}a_{44}\\
&-a_{11}a_{24}a_{33}a_{42}+a_{12}a_{24}a_{33}a_{41}+a_{13}a_{24}a_{31}a_{42}+a_{14}a_{21}a_{33}a_{42}
\end{aligned}
\tag{2.30}
$$

那么

$$x_2=\frac{\begin{vmatrix} b_2a_{11}-b_1a_{21} & a_{23}a_{11}-a_{13}a_{21} & a_{24}a_{11}-a_{14}a_{21} \\ b_3a_{11}-b_1a_{31} & a_{33}a_{11}-a_{13}a_{31} & a_{34}a_{11}-a_{14}a_{31} \\ b_4a_{11}-b_1a_{41} & a_{43}a_{11}-a_{13}a_{41} & a_{44}a_{11}-a_{14}a_{41}\end{vmatrix}}{\begin{vmatrix} a_{22}a_{11}-a_{12}a_{21} & a_{23}a_{11}-a_{13}a_{21} & a_{24}a_{11}-a_{14}a_{21} \\ a_{32}a_{11}-a_{12}a_{31} & a_{33}a_{11}-a_{13}a_{31} & a_{34}a_{11}-a_{14}a_{31} \\ a_{42}a_{11}-a_{12}a_{41} & a_{43}a_{11}-a_{13}a_{41} & a_{44}a_{11}-a_{14}a_{41}\end{vmatrix}}$$

$$=\frac{{a_{11}}^2\begin{vmatrix} a_{11} & b_1 & a_{13} & a_{14} \\ a_{21} & b_2 & a_{23} & a_{24} \\ a_{31} & b_3 & a_{33} & a_{34} \\ a_{41} & b_4 & a_{43} & a_{44}\end{vmatrix}}{{a_{11}}^2\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44}\end{vmatrix}}=\frac{\begin{vmatrix} a_{11} & b_1 & a_{13} & a_{14} \\ a_{21} & b_2 & a_{23} & a_{24} \\ a_{31} & b_3 & a_{33} & a_{34} \\ a_{41} & b_4 & a_{43} & a_{44}\end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44}\end{vmatrix}} \tag{2.31}$$

对于其它未知数有类似的结论。

从上述讨论可知，线性方程组(2.1)有唯一解是，在 $\boldsymbol{n}=2,3,4$ 时的解有类似的规律，实际上，在 $\boldsymbol{n}=1$ 时也有类似的规律，方程 $a_{11}x_1=b_1$ 在 $a_{11}\neq 0$ 时有唯一解 $x_1=\dfrac{b_1}{a_{11}}$，此时可定义一阶行列式 $|t|=t$，则该方程组有唯一解时，其解为 $x_1=\dfrac{b_1}{a_{11}}=\dfrac{|b_1|}{|a_{11}|}$。

从上述 $\boldsymbol{n}=1,2,3,4$ 的情况可知，我们可定义相应的行列式，使得相应线性方程组的**系数行列式**不为零时，线性方程组有唯一解，且在解得表达式中，未知数 $\boldsymbol{x_i}$ 可表示为分数形式，其分母为线性方程组的系数行列式，而分子为将系数行列式的第 $\boldsymbol{i}$ 列换为方程组右边常数的结果。

自然的问题：对于方程组（2.1）的一般情形，是否有类似规律？即：

问题 2.1　　对于线性方程组(2.1)，是否可以定义 n^2 个自变量

$$t_{ij}\left(i=1,2,\cdots,n;\ j=1,2,\cdots,n\right)$$

的函数（n 阶行列式）

$$\begin{vmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{vmatrix} \tag{2.32}$$

使得当线性方程组的系数行列式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0$$

时，线性方程组有唯一解，且解可表示为

$$x_k = \frac{\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,k-1} & b_1 & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{2,k-1} & b_i & a_{2,k+1} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{j,k-1} & b_j & a_{j,k+1} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,k-1} & b_n & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,k-1} & a_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{i,k-1} & a_{ik} & a_{2,k+1} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{j,k-1} & a_{jk} & a_{j,k+1} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,k-1} & a_{nk} & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}} \tag{2.33}$$

自然我们可以将 $n = 1,2,3,(4)$ 的讨论方法应用于推导 $n = 5,6,\cdots$ 的情形，但是显然运算越来越复杂，而且我们需要得到一般的结论，而不仅仅是 n 取某个具体值的情形。因此，我们需要进行抽象的逻辑推理，试图得到一般的结果。为达到此目的，我们希望能从已有的结论中找出普遍的规律，并应用此规律将已有结论推广到一般的情况。这是一个归纳－演绎的过程。

讨论：如果问题 2.1 的答案是肯定的，且

$$f = \begin{vmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{vmatrix} \tag{2.34}$$

为满足要求的函数，由于线性方程组的初等变换不改变它的解，假设互换线性方程组(2.1)的第 i, j 个方程，得到一个新的方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \qquad\vdots \\ a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j \\ \qquad\vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \\ \qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \tag{2.35}$$

这两个方程组有相同的解，因此根据问题 2.1 的结论，当

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0$$

时，的解为(2.33)，而方程组(2.35)的解为

$$x_k = \frac{\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,k-1} & b_1 & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{j,k-1} & b_j & a_{j,k+1} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{i,k-1} & b_i & a_{i,k+1} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,k-1} & b_n & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,k-1} & a_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{j,k-1} & a_{jk} & a_{j,k+1} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{i,k-1} & a_{ik} & a_{i,k+1} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,k-1} & a_{nk} & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}} \tag{2.36}$$

由于方程组(2.1)、(2.35)有相同的解，因此这两个解(2.33)、(2.36)应该相等，对比上述结果：(2.36)分子行列式是(2.33)分子行列式互换第 i, j 行的结果，而(2.36)分母行列式是(2.33)分母行列式互换第 i, j 行的结果，而要求(2.33)与(2.36)相等，那么互换

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

的两行时，其结果应该不会没有任何规律！

类似地，如果对线性方程组(2.1)做另外两种初等变换，得到新的方程组，对比两个方程组的解，同样会有类似的问题，总结为：

问题 2.2　如果存在满足问题 2.1 的函数 f (2.34)，那么

（1）互换 f 的两行，其结果如果变化？

（2） f 的一行乘以一个非零数 k，其结果如何变化？

（3） f 的一行的某个倍数加到另一行上，其结果如何变化？

对于前面所定义的 $n=2,3,4$ 阶行列式，通过计算容易得到如下结论：

（1）　互换行列式的两行，行列式的值改变符号；
（2）　行列式的一行乘以一个非零数 k，所得行列式的值为原来行列式的 k 倍；
（3）　行列式的一行的某个倍数加到另一行上，行列式的值保持不变。
我们自然会猜想：上述性质对(2.32)所给出的 n 阶行列式是否成立？对问题 2.2 的探讨就导出了对行列式性质的讨论。

有多种方式来展开行列式的讲述。一种方式是：分析低阶行列式展开式中每项的构成，然后将其规律推广到一般的行列式；而第二种方式是，注意到 4 阶行列式可通过 3 三阶行列式定义、3 阶行列式可通过 2 三阶行列式定义、2 阶行列式可通过 1 阶行列式定义，因此我们试图找出相邻阶次行列式之间的关系；第三种方式是通过公理来定义行列式。在此，我们主要讲述第一种方式，这些内容包含在§2.2、§2.3 中；对第二种方式，我们作为附录简单论述；至于第三种方式，有兴趣的读者可参考文献《代数与几何》（萧树铁 等， 高等教育出版社， ISBN 7－04－008524－0， 2000 年）。

为达到发现规律、然后加以推广的目的，首先观察低阶行列式的特点。行列式是一些项的代数和，其中每一项是一些数的乘积，

行列式阶数	项数	每项的构成	每项的符号
1	1	不同行、不同列元素的乘积	?
2	2	不同行、不同列元素的乘积	?
3	6	不同行、不同列元素的乘积	?
4	24	不同行、不同列元素的乘积	?

低阶行列式（阶数≤4）的项数与其阶数的关系是比较显然的，即 $\boldsymbol{n}$ 阶行列式展开后的项数为 $\boldsymbol{n!}$；其中每一项是不同行、不同列元素的乘积；关键是每项前的符号有何规律，其规律却不是那么明显。显然，如果我们规定每项中元素的次数如下：所有的元素按行标从小到大的次序排列，那么这些元素的列标就完全决定了这一项，因此也就决定了这一项的符号，因此我们可以从列标的排列方式来查找其规律；当然，我们亦可将一项中的元素首先按列标从小到大排列，此时这些数的行标就决定了这一项，因此这时行标的排列也就决定了这一项的符号，因此也可以从行标的排列方式来查找其规律。总而言之，我们需要考虑一个排列与一项的符号之间的关系，因此就引出了如下定义。

习题 2.1

1　请验证对于三阶行列式，如下性质成立

（1） 互换行列式的两行，行列式的值改变符号；

（2） 行列式的一行乘以一个非零数 k，所得行列式的值为原来行列式的 k 倍；

（3）行列式的一行的某个倍数加到另一行上，行列式的值保持不变。

§2.2　排列及逆序数

定义 2.1　n 元排列

n 个自然数 $1,2,\cdots,n$ 排成一个有序数组，称为一个 n 元排列。

例 2.1　排列的例子

13254	5 元排列
65(12)798(11)312(10)4	12 元排列
$(2n)(2n-2)\cdots42(2n-1)(2n-3)\cdots31$	$2n$ 元排列

显然，n 元排列共有 $n!$ 个。

为了对 n 元排列展开讨论，需要引入对 n 元排列的表示方法。n 元排列一般表示为：$i_1i_2\cdots i_n$，其中 $i_1,i_2,\cdots,i_n$ 分别为 $1,2,\cdots,n$ 中互不相同的数，i_k 为该排列 $i_1i_2\cdots i_n$ 中的第 k 个数，下标 k 表示 i_k 位于该排列中的第 k 位。（注意：符号 i 可用任何其它符号代替）

例如：若 13254 表示为 $i_1i_2i_3i_4i_5$，则 $i_1=1,\ i_2=3,\ i_3=2,\ i_4=5,\ i_5=4$。

在所有的 n 元排列中，只有 $12\cdots(n-1)n$ 是严格按照从小到大的次序排列的，我们称之为**自然排列**。在其它的 n 元排列中，必然会出现较大的数排在较小的数前面的情况。

定义 2.2　**逆序与逆序数**

在 n 元排列

$$i_1i_2\cdots i_j\cdots i_k\cdots i_n$$

中，如果 $i_j>i_k$（即前面的数大于后面的数），则称这两个数构成一个**逆序**；例如，在 5 元排列 13254 中，3 与 2 构成逆序，5 与 4 构成逆序，而 3 与 5 不构成逆序。一个排列 $i_1i_2\cdots i_n$ 的逆序总数称为该排列的**逆序数**，记作 $\tau\left(i_1i_2\cdots i_n\right)$。例如，5 元排列 13254 中，构成逆序的有：3 与 2、5 与 4，因此 13254 的逆序数为 $\tau\left(13254\right)=2$。

定义 2.3　**奇排列与偶排列**

逆序数为奇数的排列称为奇排列，逆序数为偶数的排列称为偶排列。

逆序数的计算

首先令 $t=0$，然后将排列中的每两个数拿出来进行比较，如果这两个数构成逆序，则 $\boldsymbol{t}$

增加 1，直到该排列中的任意两个数都进行了比较，则 t 的最后结果就是该排列的逆序数。

或

$$\tau\left(i_1 i_2 \cdots i_n\right) = （i_1 \text{后比} i_1 \text{小的数的个数}）$$

$$+（i_2 \text{后比} i_2 \text{小的数的个数}）$$

$$+\cdots$$

$$+（i_{n-1} \text{后比} i_{n-1} \text{小的数的个数}）$$

例 2.2　用上述方法计算排列的逆序数

$$\tau(14532) = 0+2+2+1 = 5,$$

$$\tau(n(n-1)\cdots 21) = n-1+n-2+\cdots+1 = \frac{n(n-1)}{2},$$

这说明，排列 $n(n-1)\cdots 21$ 的奇偶性为：

当 $n = 4k, 4k+1$ 时，为偶排列；
当 $n = 4k+2, 4k+3$，为奇排列。

$$\tau((2n)(2n-2)\cdots 42(2n-1)(2n-3)\cdots 31)$$

$$= (2n-1)+(2n-3)+\cdots+1+(n-1)+(n-2)+\cdots+1$$
$$= n^2 + \frac{n(n-1)}{2} = \frac{n(3n-1)}{2}$$

定义 2.4　　**对换与邻换**

将一个排列中的两个数互换位置，而其它的数保持位置不变，这样得到另外一个排列，这样的变换称为**对换**，而对换相邻的两个数称为**邻换**。

例 2.3　　5 元排列 13254，互换 3、5 的位置得到 15234。

定理 2.1　　任一排列经过一次对换必改变排列的奇偶性。
（例如：　13254 的逆序数为 2，为偶排列，
　　　　15234 的逆序数为 3，为奇排列；）
证明：首先考虑互换位置的数相邻的情况。设 n 元排列

$$i_1 i_2 \cdots i_{k-1} i_k i_{k+1} \cdots i_n \tag{2.37}$$

互换 i_k, i_{k+1} 的位置得到排列

$$i_1 i_2 \cdots i_{k-1} i_{k+1} i_k \cdots i_n \tag{2.38}$$

考虑这两个排列的逆序数的差别。显然，在这两个排列中，只有 i_k, i_{k+1} 这两个数的相对次序

发生了变化，因此根据逆序数的定义，这两个排列的逆序数相差 1：若i_k, i_{k+1}在排列(2.37)中构成逆序，则在排列(2.38)中不构成逆序，反过来，如果i_k, i_{k+1}在排列(2.37)中不构成逆序，则在排列(2.38)中构成逆序。故它们的奇偶性相反。

如果要互换位置的数不相邻，例如排列

$$i_1 i_2 \cdots i_{j-1} i_j i_{j+1} \cdots i_{k-1} i_k i_{k+1} \cdots i_n \tag{2.39}$$

互换i_j, i_k的位置得到排列

$$i_1 i_2 \cdots i_{j-1} i_k i_{j+1} \cdots i_{k-1} i_j i_{k+1} \cdots i_n \tag{2.40}$$

这个对换可通过如下方式实现：首先i_j依次与右边的$k-j-1$个数互换得到

$i_1 i_2 \cdots i_{j-1} i_{j+1} \cdots i_{k-1} i_j i_k i_{k+1} \cdots i_n$　　　共$k-j-1$次邻换

然后i_j, i_k互换位置，得到

$i_1 i_2 \cdots i_{j-1} i_{j+1} \cdots i_{k-1} i_k i_j i_{k+1} \cdots i_n$　　　1 次相邻数对换

然后i_k依次与左边的$k-j-1$个数互换得到

$i_1 i_2 \cdots i_{j-1} i_k i_{j+1} \cdots i_{k-1} i_j i_{k+1} \cdots i_n$　　　共$k-j-1$次邻换

总共$2(k-j-1)+1$次邻换，根据已有结论即得命题。

下面用例子说明这一过程：

132546　　交换 3、4 的位置得到　　142536

可通过如下方式达到此目的：

132546→ 123546 → 125346 → 125436 → 124536 → 142536

一共经过了 5 次邻换。

实际上，排列、逆序数的概念可以做简单的推广。设$a_1, a_2, \cdots, a_j, \cdots, a_k, \cdots, a_n$为任意$n$个自然数，把它们排成一行，得到$a_1 a_2 \cdots a_j \cdots a_k \cdots a_n$，称之为$n$元排列，若$a_j > a_k$，则称$a_j, a_k$构成一个逆序，而$a_1 a_2 \cdots a_j \cdots a_k \cdots a_n$中的所有逆序之和称为该排列的逆序数，即为$\tau\left(a_1 a_2 \cdots a_j \cdots a_k \cdots a_n\right)$；排列数为奇数的排列称为奇排列，排列数为偶数的排列称为偶排列。

例如，1532 为 4 元排列，其逆序数为 3，是奇排列。对于排列 164235，其逆序数为 6，

为偶排列。

问题：已知 164235 为偶排列，现在从中去除 3 这个数，得到排列 16425，请问它的逆序数为多少？其奇偶性与 164235 的奇偶性有何关系？如果在排列 16425 中所有比 3 大的数都减 1，得到排列 15324，这是一个严格意义上的 5 元排列，那么 15324、16425 的逆序数有有何关系？

一般情况下，从 n 元排列 $a_1a_2\cdots a_j\cdots a_{k-1}a_ka_{k+1}\cdots a_n$ 中去除 a_k，得到 $n-1$ 个数的排列 $a_1a_2\cdots a_j\cdots a_{k-1}a_{k+1}\cdots a_n$，已知 a_k 是 $a_1,a_2,\cdots,a_n$ 中第 l 个数（即若将 $a_1,a_2,\cdots,a_n$ 按从小到大排列，此时 a_k 为第 l 个数），请问这两个排列的奇偶性有何关系？如果把 $a_1a_2\cdots a_j\cdots a_{k-1}a_{k+1}\cdots a_n$ 中所有比 a_k 的数都减 1，而比 a_k 小的数都保持不变，得到 $b_1b_2\cdots b_j\cdots b_{k-1}b_{k+1}\cdots b_n$，那么这两个排列的逆序数又有何关系？能否得到一般的结论？

习题 2.2

1　请回答如下问题

（1） 全体 n 级排列中，奇排列和偶排列各占多少?

（2） n 元排列 $i_1i_2\cdots i_{n-1}i_n$ 与 $i_ni_{n-1}\cdots i_2i_1$ 的逆序数有何关系?

（3） 全体 n 级排列的逆序数总和等于多少?

2　计算以下排列的逆序数，并判断其奇偶性。

（1）47312658　　　（2）$135\cdots(2n-1)(2n)(2n-2)\cdots 42$

3　请试着回答本节最后一段所提出的问题。

§2.3　行列式的定义及性质

有了上一节所引入的奇排列、偶排列的概念，我们可以考虑低阶行列式中各项的符号，可以验证，其符号由如下方式决定：首先将各项的数按行标从小到大排列，然后由这些数的列标依次序排成一个排列，若该排列为偶排列，则该项取正号，若该排列为奇排列，则该项取负号。因此，我们可按得到的规律推广行列式的概念。

定义 2.5　　**n阶行列式**

设有 n^2 个数 $a_{ij},i=1,2,\cdots,n;\ j=1,2,\cdots,n$ 将它们排成 n 行、n 列

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

定义函数如下

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{\substack{j_1 j_2 \cdots j_n \\ \text{为}n\text{阶排列}}} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \tag{2.41}$$

称该函数为 **n 阶行列式**，上述行列式有时也记作 $\left|\boldsymbol{a_{ij}}\right|_{\boldsymbol{n}}$，或记作 $\left|\boldsymbol{a_{ij}}\right|$。

注释：$\sum$ 为求和号，对于每一个 n 元排列 $j_1 j_2 \cdots j_n$，都有一项

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

与该排列对应，由于 n 元排列共有 $n!$ 个，因此共有 $n!$ 个这样的项，然后把这 $n!$ 项相加，所得结果就是 n 阶行列式。（关于求和号 $\sum$ 的进一步说明参见附录）

根据 n 阶行列式的定义，n 阶行列式是所有不同行、不同列元素乘积的代数和。（由于其中每一项 $(-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$ 中的 n 个数 $a_{1j_1}, a_{2j_2}, \cdots, a_{nj_n}$ 位于不同行、不同列）

以 $n=3$ 为例，

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum_{\substack{j_1 j_2 j_3 \\ \text{为3阶排列}}} (-1)^{\tau(j_1 j_2 j_3)} a_{1j_1} a_{2j_2} a_{3j_3}$$

根据定义，三阶行列式是 3!=6 项求和的结果。

$j_1 j_2 j_3 = 123$ 时，对应项为

$$(-1)^{\tau(123)} a_{11}a_{22}a_{33} = (-1)^0 a_{11}a_{22}a_{33} = a_{11}a_{22}a_{33},$$

$j_1 j_2 j_3 = 132$ 时，对应项为

$$(-1)^{\tau(132)} a_{11}a_{23}a_{32} = (-1)^1 a_{11}a_{23}a_{32} = -a_{11}a_{23}a_{32},$$

$j_1 j_2 j_3 = 213$ 时，对应项为

$$(-1)^{\tau(213)} a_{12}a_{21}a_{33} = (-1)^1 a_{12}a_{21}a_{33} = -a_{12}a_{21}a_{33},$$

$j_1 j_2 j_3 = 231$ 时，对应项为

$$(-1)^{\tau(231)} a_{12}a_{23}a_{31} = (-1)^2 a_{12}a_{23}a_{31} = a_{12}a_{23}a_{31},$$

$j_1 j_2 j_3 = 312$ 时，对应项为

$$(-1)^{\tau(312)} a_{13}a_{21}a_{32} = (-1)^2 a_{13}a_{21}a_{32} = a_{13}a_{21}a_{32},$$

$j_1j_2j_3=321$时，对应项为

$$(-1)^{\tau(321)}a_{13}a_{22}a_{31}=(-1)^3a_{13}a_{22}a_{31}=-a_{13}a_{22}a_{31},$$

把这六项求和，就得到三阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=\sum_{\substack{j_1j_2j_3 \\ \text{为3阶排列}}}(-1)^{\tau(j_1j_2j_3)}a_{1j_1}a_{2j_2}a_{3j_3}$$

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

注意：在确定每一项的符号时，首先需要将这些项$a_{1j_1}a_{2j_2}\cdots a_{nj_n}$按行号从小到大排列，然后再看列号所构成排列的奇偶性。

某些特殊的n阶行列式可以用定义来计算。

例 2.4　　计算上三角行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & & \cdots & \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix}$$

解：根据行列式的定义可知，行列式是所有不同行、不同列元素乘积的代数和，

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=\sum_{\substack{j_1j_2\cdots j_n \\ \text{为}n\text{阶排列}}}(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$$

也就是$n!$个形如$(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$的项的和。如果其中某个项$(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$的$n$个数$a_{1j_1},a_{2j_2},\cdots,a_{nj_n}$中有一个为 0，则这一项乘积为 0，在求和时就不用考虑这一项，故在计算行列式的时候，只要考虑这n个数都不为 0 的项。在这个行列式中，在n行只有一个非零元素a_{nn}，由于第n列已经选了a_{nn}，在$n-1$行中只能选非零元素$a_{n-1,n-1}$，依此类推，在第i行只能选a_{ii}，因此这个行列式只有一个非零项

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

而其它的项都为零，因此该行列式的值为

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \cdots & & \cdots & & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn} \tag{2.42}$$

例 2.5　计算行列式

$$D_n = \begin{vmatrix} 0 & \cdots & 0 & 0 & a_{1n} \\ 0 & \cdots & 0 & a_{2,n-1} & a_{2n} \\ 0 & \cdots & a_{3,n-2} & a_{3,n-1} & a_{3,n} \\ \vdots & \iddots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-2} & a_{n,2-1} & a_{nn} \end{vmatrix}$$

解：与上例类似，这个行列式的展开式中，也只有一项不为零，有

$$D_n = (-1)^{\tau(n(n-1)\cdots 21)} a_{1n}a_{2,n-1}\cdots a_{n-1,2}a_{n1} = (-1)^{\frac{n(n-1)}{2}} a_{1n}a_{2,n-1}\cdots a_{n-1,2}a_{n1}$$

一般的n阶行列式，只有在阶数很低时才可能用定义式(2.41)计算！阶数稍高的行列式直接用定义式计算是不可行的，甚至是不可能的。各位可以估计一下用定义式计算n阶行列式需要多长时间。我们知道，n阶行列式是$n!$项的求和，其中每一项是n个数相乘，然后取相应的正负号，即使是不考虑计算逆序数的时间，那么用行列式的定义式计算n阶行列式需要做$n!$次乘法、$n!$次加法，我们也忽略掉计算$n!$次加法的时间，只考虑计算$n!$次乘法的时间。即使是$n=20$，在主频为 3GHz的个人计算机上（假设至多每秒计算3×2^{30}次乘法），那么用定义式计算 20 阶行列式至少需要$\dfrac{20!}{3\times 2^{30}\times 3600\times 24\times 365}\approx 2.4$年！基于如上分析，既然如此简单的行列式都无法计算出来，那么是不是行列式不具有实际的应用价值呢？实际上并非如此。为了使得行列式的计算在实践中可行，我们需要讨论行列式是否具有特别的性质，我们能否利用行列式的这些性质来计算行列式。而讨论行列式的性质在部分上也就是回答问题 2.2。

下面我们讨论行列式的性质。由于行列式是在求线性方程组的公式解时提出来的，而我们知道：对线性方程组作初等变换，不改变它的解。因此，对应的运算作用于行列式，又会有什么结论呢？这实际上就是回答问题 2.2。

设

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

将该行列式的行、列互换得到的行列式

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

称 D^{T} 为行列式 D 的**转置行列式**。

性质 1　行列式与它的转置行列式相等，即

$$D=D^{\mathrm{T}}$$

（说明：我们在考虑行列式定义式中某一项的符号时，我们是首先将一项中的数按行标从小到大排列，然后是根据这些数的列标所构成排列的奇偶性来确定其符号。实际上，我们也可以尝试首先将一项中的数按列标从小到大排列，然后考虑这一项的符号与这些数的行标所构成排列的奇偶性的关系。从低阶行列式的情况可知，一项的符号也是与此时的行标所构成排列的奇偶性所决定，若为偶排列，则该项的符号为正，否则为负。这就是本性质的背景。）

证明：令

$$D^{\mathrm{T}}=\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix} \overset{\Delta}{=} \begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}$$

显然有

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

对于行列式 D 中的每一项 $a_{1j_1}a_{2j_2}\cdots a_{nj_n}$，其符号为 $(-1)^{\tau(j_1j_2\cdots j_n)}$，那么这一项也在行列式 D^{T} 中出现，根据 b_{ij},a_{ji} 的对应关系，我们知道 $a_{1j_1}a_{2j_2}\cdots a_{nj_n}=b_{j_11}b_{j_22}\cdots b_{j_nn}$，因此这一项在 D^{T} 中的对应项为 $b_{j_11}b_{j_22}\cdots b_{j_nn}$，但是这一项 $b_{j_11}b_{j_22}\cdots b_{j_nn}$ 在 D^{T} 中的符号是否与它在 D 中的符号相同？根据行列式的定义，为确定这一项的符号，我们需要先将这 n 个数首先按行标从小到大排列，然后按考虑其列标所构成排列的奇偶性。虽然此时这 n 个数的行标并不是从小到大排列的，但是我们可以交换这 n 个数的次序，使得这 n 个数是按行标从小到大排序的，然后再考虑这 n 数的列标所构成的排列的奇偶性。当然这个重新排序的过程也可以逐步进行，首先将行标为 1 的数交换到第一个位置，然后将行标为 2 的数交换到第二个未知，如此下去，直到所有的数按行标从小到大排列。现在我们考虑这 n 个数行标、列标排列的逆序数之和。这一项为

$$\begin{aligned}&(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}=(-1)^{\tau(j_1j_2\cdots j_n)}b_{j_11}b_{j_22}\cdots b_{j_nn}\\&=(-1)^{\tau(j_1j_2\cdots j_n)+\tau(12\cdots n)}b_{j_11}b_{j_22}\cdots b_{j_nn}\end{aligned}$$

将 $b_{j_11},b_{j_22},\cdots,b_{j_nn}$ 重新排序后，行标排序成 $12\cdots n$，假设这时列标排成 $k_1k_2\cdots k_n$，则 $b_{j_11}b_{j_22}\cdots b_{j_nn}$ 就变为 $b_{1k_1}b_{2k_2}\cdots b_{nk_n}$，由于交换 $b_{1k_1}b_{2k_2}\cdots b_{nk_n}$ 中两个数的位置时，其行标、列

标所构成的排列同时做了一次对换，其奇偶性同变化，因此交换两个数的位置前后，其行标、列标所构成排列的逆序数之和的奇偶性是不变的，既然是交换两个数的位置时有这个结论，那么多次交换两个数的次序时，该结论仍然成立，因此

$$\begin{aligned}&(-1)^{\tau(j_1j_2\cdots j_n)+\tau(12\cdots n)}b_{j_11}b_{j_22}\cdots b_{j_nn}\\&=(-1)^{\tau(12\cdots n)+\tau(j_1j_2\cdots j_n)}b_{1k_1}b_{2k_2}\cdots b_{nk_n}\\&=(-1)^{\tau(k_1k_2\cdots k_n)}b_{1k_1}b_{2k_2}\cdots b_{nk_n}\end{aligned}$$

而这显然是 D^{T} 中的一项。既然 D 中的每一项都是 D^{T} 中的一项，且符号相同，因此我们就证明了这个结论。

■

因此，以后对行列式的行成立的性质，对行列式的列也成立。

性质 2　如果行列式中两行互换，则行列式改变符号。

$$\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\\vdots&\vdots&&\vdots\\a_{i1}&a_{i2}&\cdots&a_{in}\\\vdots&\vdots&&\vdots\\a_{k1}&a_{k2}&\cdots&a_{kn}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}=-\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\\vdots&\vdots&&\vdots\\a_{k1}&a_{k2}&\cdots&a_{kn}\\\vdots&\vdots&&\vdots\\a_{i1}&a_{i2}&\cdots&a_{in}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}$$

证明：考虑左边行列式中的一项，即在第 i 行中选取第 j_i 列（$i=1,2,\cdots,n$）数所确定的项

$$a_{1j_1}a_{2j_2}\cdots a_{ij_i}\cdots a_{kj_k}\cdots a_{nj_n}$$

其符号为

$$(-1)^{\tau(j_1j_2\cdots j_i\cdots j_k\cdots j_n)}$$

右边的行列式中也会出现 $a_{1j_1}a_{2j_2}\cdots a_{ij_i}\cdots a_{kj_k}\cdots a_{nj_n}$ 这一项——除了在第 k 行选第 j_i 列（a_{ij_i}）、在第 i 行选第 j_k 列（a_{kj_k}）外，其它各行（第 l 行）选第 j_l 列）。而在右边的行列式中，a_{ij_i} 位于第 k 行、第 j_i 列，a_{kj_k} 位于第 i 行、第 j_k 列，而这两个数的列标保持不变，这一项中其它的数的行标、列标都保持不变，因此该项在右边的行列式中的符号为

$$(-1)^{\tau(j_1j_2\cdots j_k\cdots j_i\cdots j_n)}$$

显然，$j_1j_2\cdots j_i\cdots j_k\cdots j_n$ 与 $j_1j_2\cdots j_k\cdots j_i\cdots j_n$ 相差一个对换——互换第 i,k 个数。据对换的性质，我们知道，对于左边行列式中的每一项

$$a_{1j_1}a_{2j_2}\cdots a_{ij_i}\cdots a_{kj_k}\cdots a_{nj_n}$$

在两边行列式中都有相同的一项，而只是符号相反，因此命题成立。

■

推论 1　如果行列式有两行(列)完全相同，则此行列式为零。

性质 3　以数 k 乘以行列式一行的元素，等于用 k 乘此行列式 。

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

证明：

$$\begin{aligned} \text{左边} &= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1}a_{2j_2}\cdots(ka_{ij_i})\cdots a_{nj_n} \\ &= k\sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1}a_{2j_2}\cdots a_{ij_i}\cdots a_{nj_n} \\ &= \text{右边} \end{aligned}$$

推论 2　行列式中如果有两行(列)成比例，则行列式为零。

性质 4　若行列式的某一行(列)的元素都是两数之和，则该行列式等于两个行列式的和。

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

证明：

$$\begin{aligned} \text{左边} &= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1}a_{2j_2}\cdots(b_{ij_i}+b_{ij_i})\cdots a_{nj_n} \\ &= \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1}a_{2j_2}\cdots b_{ij_i}\cdots a_{nj_n} + \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1}a_{2j_2}\cdots c_{ij_i}\cdots a_{nj_n} \\ &= \text{右边} \end{aligned}$$

性质 5　将行列式某一行的各个元素都乘以同一数 k 后加到另外一行对应的元素上去(简称：将一行的 k 倍加到另外一行)，行列式不变。

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

（将第 $\boldsymbol{i}$ 行的 $\boldsymbol{k}$ 倍加到第 j 行）

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

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

■

性质 6　　　行列式按行展开

根据定义，n 阶行列式是 $n!$ 项的代数和，其中每一项是 n 个不同行、不同列元素的乘积。在这 $n!$ 项中，含有 a_{ij} 的共有 $(n-1)!$ 项，这 $(n-1)!$ 项中的每一项提取出 a_{ij} 后，得到新的 $(n-1)!$ 项，显然，这新的 $(n-1)!$ 项中的每一项是原来行列式中 $n-1$ 个不同行、不同列元素的乘积，且这 $n-1$ 个元素位于原来行列式中的第 $1,\cdots,i-1,i+1,\cdots,n$ 行、第 $1,\cdots,j-1,j+1,\cdots,n$ 列。由于这 $n-1$ 行、$n-1$ 列元素可定义一个 $n-1$ 阶行列式，该 $n-1$ 阶行列式共有 $(n-1)!$ 项。自然的问题：这新的 $(n-1)!$ 项的和与原来行列式中去掉第 i 行、第 j 列后的 $n-1$ 阶行列式是否有密切的关系？我们也可以用低阶行列式来看一下是否有普遍的规律，各位读者可自行试验一下。

定义 2.6　　**余子式、代数余子式**

在 n 阶行列式 $\left|a_{ij}\right|$ 中，划去 a_{ij} 所在行(第 i 行)、所在列(第 j 列)的元素，其它元素的相对位置保持不变，这些元素定义一个 $n-1$ 阶行列式，该行列式称为元素 a_{ij} 的余子式，记做 M_{ij}，而 $A_{ij}=(-1)^{i+j}M_{ij}$ 称为 a_{ij} 的代数余子式。

5 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix}$$

的元素 $\boldsymbol{a}_{23}$ 的余子式、代数余子式分别为

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{14} & a_{15} \\ a_{31} & a_{32} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{54} & a_{55} \end{vmatrix}$$

元素 a_{23} 的余子式

$$A_{23}=(-1)^{2+3}\begin{vmatrix} a_{11} & a_{12} & a_{14} & a_{15} \\ a_{31} & a_{32} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{54} & a_{55} \end{vmatrix} = -\begin{vmatrix} a_{11} & a_{12} & a_{14} & a_{15} \\ a_{31} & a_{32} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{54} & a_{55} \end{vmatrix}$$

元素 a_{23} 的代数余子式

定理 2.2　　行列式 $D=\left|a_{ij}\right|_n$ 等于它的任意一行的元素与其代数余子式的乘积之和

$$D = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik} \tag{2.43}$$

（称之为行列式按第 i 行展开）

证明：首先考虑一种最简单情形——行列式的最后一行中只有最后一个数不为零：

$$D_1 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix} = \sum_{\substack{j_1 j_2 \cdots j_n \\ \text{为}n\text{阶排列}}} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

由于第 n 行中只有 $a_{nn} \neq 0$，因此在行列式定义的展开式中，只有这些项

$$(-1)^{\tau(j_1 j_2 \cdots j_{n-1} j_n)} a_{1j_1} a_{2j_2} \cdots a_{n-1,j_{n-1}} a_{nn} \qquad j_1 j_2 \cdots j_{n-1} n \text{ 为 } n \text{ 元排列}$$

才不为零，因此

$$\begin{aligned} D &= \sum_{\substack{j_1 j_2 \cdots j_{n-1} n \\ \text{为}n\text{阶排列}}} (-1)^{\tau(j_1 j_2 \cdots j_{n-1} n)} a_{1j_1} a_{2j_2} \cdots a_{n-1,j_{n-1}} a_{nn} \\ &= a_{nn} \sum_{\substack{j_1 j_2 \cdots j_{n-1} n \\ \text{为}n\text{阶排列}}} (-1)^{\tau(j_1 j_2 \cdots j_{n-1} n)} a_{1j_1} a_{2j_2} \cdots a_{n-1,j_{n-1}} a_{nn} \\ &= a_{nn} \sum_{\substack{j_1 j_2 \cdots j_{n-1} \\ \text{为}n-1\text{阶排列}}} (-1)^{\tau(j_1 j_2 \cdots j_{n-1})} a_{1j_1} a_{2j_2} \cdots a_{n-1,j_{n-1}} \\ &= a_{nn} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} \\ \vdots & \vdots & & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} \end{vmatrix} = a_{nn}M_{nn} = a_{nn}A_{nn} \end{aligned}$$

然后考虑第二种情况——某一行中只有一个数不为零：

$$D_2 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & 0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

这种行列式的某一行中只有一个数不为零（这里是在第 i 行中，只有第 j 个数不为零）。我们首先利用行列式的性质将其转化为第一种情形——将只有一个非零元素的行“换到”最后一行、这个非零元素“换到”第 n 列，同时保持其它各行、各列的相对位置不变；然后利用第一种情况的结论。我们可通过如下步骤达到该目的：首先将第 i 行与第 $i+1$ 行互换，然后

是$i+1$行与第$i+2$行互换，……，最后将$n-1$行与第n行互换，这样就将D_2中的第i行换到了第n行，而$i+1$行到第n行则整体上移到第i行到第$n-1$行，即

$$D_3=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \\ 0 & 0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \end{vmatrix}$$

显然这个过程共进行了$n-i$次行交换，因此根据行列式的性质有

$$D_3=(-1)^{n-i}D_2$$

类似地对列作交换，将a_{ij}所在的列换到最后一列，而其它各列的相对位置不变，即得到

$$D_4=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} & a_{1j} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} & a_{i-1,j} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} & a_{i+1,j} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} & a_{nj} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & a_{ij} \end{vmatrix}$$

显然，D_3经过$n-j$次交换相邻列的运算化为D_4，因此

$$D_4=(-1)^{n-j}D_3$$

D_4已经是第一种情况的类型，因此

$$D_4=a_{ij}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

$$=a_{ij}M_{ij}$$

故有

$$a_{ij}M_{ij} = D_4 = (-1)^{n-j} D_3 = (-1)^{n-j}(-1)^{n-i} D_2$$
$$= (-1)^{2n-i-j} D_2 = (-1)^{-i-j} D_2$$

而 $$D_2 = (-1)^{i+j} a_{ij}M_{ij} = a_{ij}(-1)^{i+j} M_{ij} = a_{ij}A_{ij}$$

而一般情形可用行列式的性质 4 来考虑：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & & \cdots \\ a_{i1}+0+\cdots 0 & 0+a_{i2}+0+\cdots+0 & \cdots & 0+\cdots+0+a_{in} \\ \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & & \cdots \\ a_{i1} & 0 & \cdots & 0 \\ \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & & \cdots \\ 0 & a_{i2} & \cdots & 0 \\ \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & & \cdots \\ 0 & 0 & \cdots & a_{in} \\ \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$
$$= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$$

这样我们就证明了该定理。 ■

上述定理我们也可以用行列式的定义直接加以证明，这需要对逆序数、行列式的定义有较深刻的理解。

$$D = \sum_{\substack{j_1 j_2 \cdots j_n \\ \text{为}n\text{阶排列}}} (-1)^{\tau(j_1 j_2 \cdots j_{i-1} j_i j_{i+1} \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{ij_i} \cdots a_{nj_n}$$

（在行列式的展开式中，把所有含有第 i 行的第 k 个元素 a_{ik} 的项合并到一起）

$$= \sum_{k=1}^{n} \sum_{\substack{j_1 j_2 \cdots j_{i-1} k j_{i+1} \cdots j_n \\ \text{为}n\text{阶排列}}} (-1)^{\tau(j_1 j_2 \cdots j_{i-1} k j_{i+1} \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{i-1,j_{i-1}} a_{ik} a_{i+1,j_{i+1}} \cdots a_{nj_n}$$

（提取公因式 a_{ik}，那么其余 $n-1$ 个数 $a_{1j_1}, a_{2j_2}, \cdots a_{i-1,j_{i-1}}, a_{i+1,j_{i+1}}, \cdots, a_{nj_n}$ 位于行列式 D 的除第 i 行的 $n-1$ 行、除第 k 列的 $n-1$ 列中）

$$= \sum_{k=1}^{n} a_{ik} \sum_{\substack{j_1 j_2 \cdots j_{i-1} k j_{i+1} \cdots j_n \\ \text{为}n\text{阶排列}}} (-1)^{\tau(j_1 j_2 \cdots j_{i-1} k j_{i+1} \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{i-1,j_{i-1}} a_{i+1,j_{i+1}} \cdots a_{nj_n}$$

（为了便于计算排列 $j_1 j_2 \cdots j_{i-1} k j_{i+1} \cdots j_n$ 的奇偶性，应用邻换将 k 从排列中的第 i 位换到第一位，根据关于置换对逆序数奇偶性的影响，）

$$=\sum_{k=1}^{n}a_{ik}\sum_{\substack{j_1j_2\cdots j_{i-1}kj_{i+1}\cdots j_n\\ \text{为}n\text{阶排列}}}(-1)^{\tau(kj_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n)}(-1)^{i-1}a_{1j_1}a_{2j_2}\cdots a_{i-1,j_{i-1}}a_{i+1,j_{i+1}}\cdots a_{nj_n}$$

（由于 $kj_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n$ 为 n 元排列，$kj_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n$ 中比 k 小的数有 $k-1$ 个，因此 $\tau(kj_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n)=\tau(j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n)+k-1$）

$$=\sum_{k=1}^{n}a_{ik}\sum_{\substack{j_1j_2\cdots j_{i-1}kj_{i+1}\cdots j_n\\ \text{为}n\text{阶排列}}}(-1)^{\tau(j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n)}(-1)^{i-1}(-1)^{k-1}a_{1j_1}a_{2j_2}\cdots a_{i-1,j_{i-1}}a_{i+1,j_{i+1}}\cdots a_{nj_n}$$

（把里层求和号中的公因子 $(-1)^{i-1}(-1)^{k-1}$ 提取出来）

$$=\sum_{k=1}^{n}(-1)^{i+k}a_{ik}\sum_{\substack{j_1j_2\cdots j_{i-1}kj_{i+1}\cdots j_n\\ \text{为}n\text{阶排列}}}(-1)^{\tau(j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{i-1,j_{i-1}}a_{i+1,j_{i+1}}\cdots a_{nj_n}$$

（排列 $j_1j_2\cdots j_{i-1}kj_{i+1}\cdots j_n$ 与排列 $j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n$ 一一对应）

$$=\sum_{k=1}^{n}(-1)^{i+k}a_{ik}\sum_{\substack{j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n\\ \text{为}1,2,\cdots,k-1,k+1,\cdots,n\text{的排列}}}(-1)^{\tau(j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{i-1,j_{i-1}}a_{i+1,j_{i+1}}\cdots a_{nj_n}$$

（把 $i+1,\cdots,n$ 行的行号减 1，在列号 $j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n$，把所有比 k 大的数减 1，根据行列式的定义可知，$\sum\limits_{j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{i-1,j_{i-1}}a_{i+1,j_{i+1}}\cdots a_{nj_n}$ 为 a_{ik} 的余子式 M_{ik}）

$$=\sum_{k=1}^{n}(-1)^{i+k}a_{ik}M_{ik}=\sum_{k=1}^{n}a_{ik}A_{ik}$$

该定理有一个重要的推论。

推论 3　一行元素与另外一行元素的代数余子式乘积之和为 0

$$a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=\sum_{k=1}^{n}a_{ik}A_{jk}=0\left(i\neq j\right),\tag{2.44}$$

证明：考虑如下两个行列式

$$D_1=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \qquad D_2=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

显然，D_2 是在 D_1 中，将其第 j 行元素换为第 i 行元素的结果，由于 D_2 有两行相同，因此该行列式为零。我们注意到：行列式 D_1、D_2 中，<u>其第 $\boldsymbol{j}$ 行对应元素的代数余子式是相同的</u>。即：如果令 D_1 中第 k 行、第 l 列元素的代数余子式为 A_{kl}，D_2 中第 k 行、第 l 列元素的代数余子式为 B_{kl}，则有

$$A_{jl}=B_{jl}, \qquad l=1,2,\cdots,n$$

因此，我们将 D_2 按第 j 行展开，就有

$$\begin{aligned} 0=D_2&=a_{i1}B_{j1}+a_{i2}B_{j2}+\cdots+a_{in}B_{jn} \\ &=a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=\sum_{k=1}^{n}a_{ik}A_{jk} \end{aligned}$$

命题得证。

■

根据性质行列式的性质 1，也有

$$a_{1i}A_{1j}+a_{2i}A_{2j}+\cdots+a_{ni}A_{nj}=\sum_{k=1}^{n}a_{ki}A_{kj}=0\left(\, i\neq j\right)$$

即：$i\neq j$ 时，所有第 i 列元素与第 j 列对应元素的代数余子式乘积之和等于 0。

综合以上的定义与推论，我们有

$$a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=\sum_{k=1}^{n}a_{ik}A_{jk}=\begin{cases}\left|a_{ij}\right| & i=j \\ 0 & i\neq j\end{cases} \tag{2.45}$$

$$a_{1i}A_{1j}+a_{2i}A_{2j}+\cdots+a_{ni}A_{nj}=\sum_{k=1}^{n}a_{ki}A_{kj}=\begin{cases}\left|a_{ij}\right| & i=j \\ 0 & i\neq j\end{cases} \tag{2.46}$$

现定义克雷勒克尔（Kronecker）符号

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

则有(2.45)、(2.46)分别可记作

$$\sum_{k=1}^{n} a_{ik} A_{jk} = |a_{ij}| \delta_{ij} \tag{2.47}$$

$$\sum_{k=1}^{n} a_{ki} A_{kj} = |a_{ij}| \delta_{ij} \tag{2.48}$$

Laplace 定理

定义 2.7 行列式的子式

设 D 是一个 n 阶行列式，在 D 中取 $k(1 \le k \le n)$ 行 $(i_1 < i_2 < \cdots < i_k)$、$k$ 列 $(j_1 < j_2 < \cdots < j_k)$，由这些行与列相交处的元素按原来相对位置所构成的 k 阶行列式，我们称之为 D 的一个 k 阶子式，记作 $N_D\begin{pmatrix} i_1, i_2, \cdots, i_k \\ j_1, j_2, \cdots, j_k \end{pmatrix}$，简记作 $N\begin{pmatrix} i_1, i_2, \cdots, i_k \\ j_1, j_2, \cdots, j_k \end{pmatrix}$。

定义 2.8 余子式、代数余子式、主子式

设 $N\begin{pmatrix} i_1, i_2, \cdots, i_k \\ j_1, j_2, \cdots, j_k \end{pmatrix}$ 是 n 阶行列式 D 的一个 k 阶子式，在 D 中划去 N 所在的行、列后，剩下的元素按原来的相对位置构成的 $n-k$ 阶子式称为子式 $N\begin{pmatrix} i_1, i_2, \cdots, i_k \\ j_1, j_2, \cdots, j_k \end{pmatrix}$ 的**余子式**，记作 $M\begin{pmatrix} i_1, i_2, \cdots, i_k \\ j_1, j_2, \cdots, j_k \end{pmatrix}$；称

$$A\begin{pmatrix} i_1, i_2, \cdots, i_k \\ j_1, j_2, \cdots, j_k \end{pmatrix} = (-1)^{i_1+i_2+\cdots+i_k+j_1+j_2+\cdots+j_k} M\begin{pmatrix} i_1, i_2, \cdots, i_k \\ j_1, j_2, \cdots, j_k \end{pmatrix}$$

为子式 $N\begin{pmatrix} i_1, i_2, \cdots, i_k \\ j_1, j_2, \cdots, j_k \end{pmatrix}$ 的**代数余子式**。若子式 $N\begin{pmatrix} i_1, i_2, \cdots, i_k \\ j_1, j_2, \cdots, j_k \end{pmatrix}$ 所在行的序数与所在列的序数相同，则称子式 N 为 D 的一个**主子式**。

在四级行列式

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

中选定第一、三行，第二、四列得到一个二阶子式 $N\begin{pmatrix}1,3\\2,4\end{pmatrix}=\begin{vmatrix}2&4\\0&1\end{vmatrix}$，$M$的余子式为

$M\begin{pmatrix}1,3\\2,4\end{pmatrix}=\begin{vmatrix}0&2\\0&1\end{vmatrix}$，$M$的代数余子式为 $M\begin{pmatrix}1,3\\2,4\end{pmatrix}=(-1)^{(1+3)+(2+4)}\begin{vmatrix}0&2\\0&1\end{vmatrix}$。

在五阶行列式

$$D=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{15}\\a_{21}&a_{22}&\cdots&a_{25}\\\vdots&\vdots&&\vdots\\a_{51}&a_{52}&\cdots&a_{55}\end{vmatrix}$$

中 $N\begin{pmatrix}1,2,5\\2,3,5\end{pmatrix}=\begin{vmatrix}a_{12}&a_{13}&a_{15}\\a_{22}&a_{23}&a_{25}\\a_{52}&a_{52}&a_{55}\end{vmatrix}$ 与 $M\begin{pmatrix}1,2,5\\2,3,5\end{pmatrix}=\begin{vmatrix}a_{31}&a_{34}\\a_{41}&a_{44}\end{vmatrix}$ 互为余子式。

引理2.1 行列式 D 的任一个子式 N 与它的代数余子式 A 的乘积中的每一项都是行列式 D 的展开式中的一项，而且符号也一致。

证明（*）：先证明 N 位于行列式D的左上方的情形，即 $N\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}$ 的情形：

$$D=\begin{vmatrix}a_{11}&\cdots&a_{1k}&a_{1,k+1}&\cdots&a_{1n}\\\vdots&M&\vdots&\vdots&&\vdots\\a_{k1}&\cdots&a_{kk}&a_{k,k+1}&\cdots&a_{kn}\\a_{k+1,1}&\cdots&a_{k+1,k}&a_{k+1,k+1}&\cdots&a_{k+1,n}\\\vdots&&\vdots&\vdots&M'&\vdots\\a_{n1}&\cdots&a_{nk}&a_{n,k+1}&\cdots&a_{nn}\end{vmatrix}$$

此时，$N\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}$ 的代数余子式A为

$$A\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}=(-1)^{(1+2+\cdots+n)+(1+2+\cdots+n)}M\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}=M\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}$$

$N\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}$的每一项可写作 $(-1)^{\tau(s_1s_2\cdots s_k)}a_{1s_1}a_{2s_2}\cdots a_{ks_k}$，其中 $s_1,s_2,\cdots,s_k$ 为 $1,2,\cdots,k$ 的一个排列；

$M\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}$ 的每一项可写作 $(-1)^{\tau((t_{k+1}-k)(t_{k+2}-k)\cdots(t_n-k))}a_{k+1,t_{k+1}}a_{k+2,t_{k+2}}\cdots a_{nt_n}$，其中

$t_{k+1},t_{k+2},\cdots,t_n$ 为 $k+1,k+2,\cdots,n$ 的一个排列。

因此 $N\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}A\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}$ 乘积中的一项为

$$
\begin{aligned}
&(-1)^{\tau(s_1s_2\cdots s_k)}a_{1s_1}a_{2s_2}\cdots a_{ks_k}(-1)^{\tau((t_{k+1}-k)(t_{k+2}-k)\cdots(t_n-k))}a_{k+1,t_{k+1}}a_{k+2,t_{k+2}}\cdots a_{nt_n}\\
&=(-1)^{\tau(s_1s_2\cdots s_k)+\tau((t_{k+1}-k)(t_{k+2}-k)\cdots(t_n-k))}a_{1s_1}a_{2s_2}\cdots a_{ks_k}a_{k+1,t_{k+1}}a_{k+2,t_{k+2}}\cdots a_{nt_n}\\
&=(-1)^{\tau(s_1s_2\cdots s_kt_{k+1}t_{k+2}\cdots t_n)}a_{1s_1}a_{2s_2}\cdots a_{ks_k}a_{k+1,t_{k+1}}a_{k+2,t_{k+2}}\cdots a_{nt_n}
\end{aligned}
$$

于是，这个乘积项是行列式D的展开式中的一项，而且符号也一致。
下证一般情形：

现考虑子式 $N\begin{pmatrix}i_1,i_2,\cdots,i_k\\j_1,j_2,\cdots,j_k\end{pmatrix}$ 与它的代数余子式的乘积。变动D中行列的次序使

$N\begin{pmatrix}i_1,i_2,\cdots,i_k\\j_1,j_2,\cdots,j_k\end{pmatrix}$ 位于D的左上角。为此，先把第 i_1 行依次与 $i_1-1,i_1-2,\cdots,1$ 行对换，这样

经过 i_1-1 次行变换将第 i_1 行换到第1行；如此继续下去，一共经过

$$(i_1-1)+(i_2-2)+\cdots+(i_k-k)=(i_1+i_2+\cdots+i_k)-(1+2+\cdots+k)$$

次行变换把第 $i_1,i_2,\cdots,i_k$ 行依次换到第 $1,2,\cdots,k$ 行。

利用类似的列变换，一共经过

$$(j_1-1)+(j_2-2)+\cdots+(j_k-k)=(j_1+j_2+\cdots+j_k)-(1+2+\cdots+k)$$

次列变换把第 $j_1,j_2,\cdots,j_k$ 列依次换到第 $1,2,\cdots,k$ 列。

用 D_1 表示对 D 做上述行列变换后所得的新行列式，则

$$D_1=(-1)^{(i_1+i_2+\cdots+i_k)-(1+2+\cdots+k)+(j_1+j_2+\cdots+j_k)-(1+2+\cdots+k)}D=(-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)}D$$

于是，D_1 和D的展开式中出现的项是一样的，只是每一项都差符号

$$(-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)}$$

现在 D 中子式 $N\begin{pmatrix}i_1,i_2,\cdots,i_k\\j_1,j_2,\cdots,j_k\end{pmatrix}$ 的元素位于位于 D_1 的左上角，因此

$N_{D_1}\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}A_{D_1}\begin{pmatrix}1,2,\cdots,k\\1,2,\cdots,k\end{pmatrix}$ 中每一项都是行列式 D_1 的展开式中的一项，而且符号也一致；又因

$$N\begin{pmatrix} i_1,i_2,\cdots,i_k \\ j_1,j_2,\cdots,j_k \end{pmatrix} A\begin{pmatrix} i_1,i_2,\cdots,i_k \\ j_1,j_2,\cdots,j_k \end{pmatrix} = (-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)} N\begin{pmatrix} i_1,i_2,\cdots,i_k \\ j_1,j_2,\cdots,j_k \end{pmatrix} M\begin{pmatrix} i_1,i_2,\cdots,i_k \\ j_1,j_2,\cdots,j_k \end{pmatrix}$$
$$= (-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)} N_{D_1}\begin{pmatrix} 1,2,\cdots,k \\ 1,2,\cdots,k \end{pmatrix} A_{D_1}\begin{pmatrix} 1,2,\cdots,k \\ 1,2,\cdots,k \end{pmatrix}$$

所以 $N\begin{pmatrix} i_1,i_2,\cdots,i_k \\ j_1,j_2,\cdots,j_k \end{pmatrix} A\begin{pmatrix} i_1,i_2,\cdots,i_k \\ j_1,j_2,\cdots,j_k \end{pmatrix}$ 中每一项都是行列式D的展开式中的一项，而且符号也一致。

定理 2.3　（Laplace定理）设在行列式D中任意取定了 $k\left(1\le k\le n-1\right)$ 行，由这 k 行元素所组成的一切 k 级子式与它们的代数余子式的乘积之和等于行列式D。

证明（*）：设D中取定 k 行后得到的子式为 $M_1,M_2,\cdots,M_t$ 它们的代数余子式分别为

$$A_1,A_2,\cdots,A_t$$

由引理得 M_iA_i 中每一项都是行列式D的展开式中的一项，且符号相同，而 $\sum_{i=1}^{t}M_iA_i$ 中有

$C_n^k k!(n-k)! = \dfrac{n!}{k!(n-k)!}k!(n-k)! = n!$ 项，又因 M_iA_i 和 $M_jA_j\left(i\ne j\right)$ 无公共项，所以

$D=\sum_{i=1}^{t}M_iA_i$ 。

例2.6 用Laplace定理计算行列式

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

解：在 D 中取定第一、二行，得到六个子式与相应的代数余子式：

$$N\begin{pmatrix} 1,2 \\ 1,2 \end{pmatrix} = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1, \quad A\begin{pmatrix} 1,2 \\ 1,2 \end{pmatrix} = (-1)^{(1+2)+(1+2)}\begin{vmatrix} 2 & -3 \\ 0 & 3 \end{vmatrix} = 6$$

$$N\begin{pmatrix} 1,2 \\ 1,3 \end{pmatrix} = \begin{vmatrix} 2 & 2 \\ 1 & -2 \end{vmatrix} = -6, \quad A\begin{pmatrix} 1,2 \\ 1,3 \end{pmatrix} = (-1)^{(1+2)+(1+3)}\begin{vmatrix} 0 & -3 \\ -1 & 3 \end{vmatrix} = 3$$

$$N\begin{pmatrix} 1,2 \\ 1,4 \end{pmatrix} = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1, \quad A\begin{pmatrix} 1,2 \\ 1,4 \end{pmatrix} = (-1)^{(1+2)+(1+4)}\begin{vmatrix} 0 & 2 \\ -1 & 0 \end{vmatrix} = 2$$

$$N\begin{pmatrix}1,2\\2,3\end{pmatrix}=\begin{vmatrix}1&2\\1&-2\end{vmatrix}=-4,\quad A\begin{pmatrix}1,2\\2,3\end{pmatrix}=(-1)^{(1+2)+(2+3)}\begin{vmatrix}0&-3\\-1&3\end{vmatrix}=-3$$

$$N\begin{pmatrix}1,2\\2,4\end{pmatrix}=\begin{vmatrix}1&1\\1&1\end{vmatrix}=0,\quad A\begin{pmatrix}1,2\\2,4\end{pmatrix}=(-1)^{(1+2)+(2+4)}\begin{vmatrix}0&2\\-1&0\end{vmatrix}=2$$

$$N\begin{pmatrix}1,2\\3,4\end{pmatrix}=\begin{vmatrix}2&1\\-2&1\end{vmatrix}=4,\quad A\begin{pmatrix}1,2\\3,4\end{pmatrix}=(-1)^{(1+2)+(3+4)}\begin{vmatrix}0&0\\-1&-1\end{vmatrix}=0$$

根据Laplace定理可知，四阶行列式 D 为上述这六个子式与其代数余子式乘积之和，即

$$D=1\times6+(-6)\times3+1\times2+(-4)\times(-3)+0\times2+1\times0=2$$

下面我们利用以上关于行列式的性质来计算一些行列式。首先，我们已经知道了有关行列式的一些基本结果，例如：

上、下三角行列式，

$$\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1,n-1}&a_{1n}\\0&a_{22}&\cdots&a_{2,n-1}&a_{2n}\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&a_{n-1,n-1}&a_{n-1,n}\\0&0&\cdots&0&a_{nn}\end{vmatrix}=\begin{vmatrix}a_{11}&0&\cdots&0&0\\a_{21}&a_{22}&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots\\a_{n-1,1}&a_{n-1,2}&\cdots&a_{n-1,n-1}&0\\a_{n1}&a_{n2}&\cdots&a_{n,n-1}&a_{nn}\end{vmatrix}=a_{11}a_{22}\cdots a_{nn}$$

类似地有，

$$\begin{vmatrix}0&0&\cdots&0&a_{1n}\\0&0&\cdots&a_{2,n-1}&a_{2n}\\\vdots&\vdots&&\vdots&\vdots\\0&a_{n-1,2}&\cdots&a_{n-1,n-1}&a_{n-1,n}\\a_{n1}&a_{n2}&\cdots&a_{n,n-1}&a_{nn}\end{vmatrix}=(-1)^{\frac{n(n-1)}{2}}a_{1n}a_{2,n-1}\cdots a_{n-1,2}a_{n1}$$

另外我们也已经讨论了行列式的性质，我们可把一般的行列式利用行列式的性质将它化作与之相等的上三角行列式。我们举例说明这一方法。

例 2.7　　计算四阶行列式 $D=\begin{vmatrix}1&1&-1&2\\-1&-1&-4&1\\2&4&-6&1\\1&2&4&2\end{vmatrix}$

解：$D=\begin{vmatrix}1&1&-1&2\\-1&-1&-4&1\\2&4&-6&1\\1&2&4&2\end{vmatrix}\overset{\substack{r_2+r_1\\r_3-2r_1\\r_4-r_1}}{=}\begin{vmatrix}1&1&-1&2\\0&0&-5&3\\0&2&-4&-3\\0&1&5&0\end{vmatrix}\overset{r_2\leftrightarrow r_4}{=}-\begin{vmatrix}1&1&-1&2\\0&1&5&0\\0&2&-4&-3\\0&0&-5&3\end{vmatrix}$

$$\overset{r_3-2r_2}{=}-\begin{vmatrix}1 & 1 & -1 & 2\\ 0 & 1 & 5 & 0\\ 0 & 0 & -14 & -3\\ 0 & 0 & -5 & 3\end{vmatrix}\overset{r_4-\frac{5}{14}r_3}{=}-\begin{vmatrix}1 & 1 & -1 & 2\\ 0 & 1 & 5 & 0\\ 0 & 0 & -14 & -3\\ 0 & 0 & 0 & 3+\dfrac{15}{14}\end{vmatrix}$$

$$=-1\times1\times(-14)\times\left(3+\frac{15}{14}\right)=57$$

用归纳法容易说明：

一切行列式都可化为上（下）三角行列式。

例 2.8　　计算行列式 $\begin{vmatrix}a^2 & ab & b^2\\ 2a & a+b & 2b\\ 1 & 1 & 1\end{vmatrix}$

解：$\begin{vmatrix}a^2 & ab & b^2\\ 2a & a+b & 2b\\ 1 & 1 & 1\end{vmatrix}\overset{r_2-2br_3}{\underset{r_1-b^2r_3}{=\!=}}\begin{vmatrix}a^2-b^2 & ab-b^2 & 0\\ 2(a-b) & a-b & 0\\ 1 & 1 & 1\end{vmatrix}$

$$\overset{c_1-c_2}{=\!=}(a-b)^2\begin{vmatrix}a & b & 0\\ 1 & 1 & 0\\ 0 & 1 & 1\end{vmatrix}\overset{c_2-c_3}{=\!=}(a-b)^2\begin{vmatrix}a & b & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{vmatrix}=(a-b)^3$$

例 2.9　　计算行列式 $D=\begin{vmatrix}a & b & c & d\\ a & a+b & a+b+c & a+b+c+d\\ a & 2a+b & 3a+2b+c & 4a+3b+2c+d\\ a & 3a+b & 6a+3b+c & 10a+6b+3c+d\end{vmatrix}$

解：$D\overset{\substack{r_4-r_3\\ r_3-r_2}}{\underset{r_2-r_1}{=}}\begin{vmatrix}a & b & c & d\\ 0 & a & a+b & a+b+c\\ 0 & a & 2a+b & 3a+2b+c\\ 0 & a & 3a+b & 6a+3b+c\end{vmatrix}$

$$\overset{r_4-r_3}{\underset{r_3-r_2}{=}}\begin{vmatrix}a & b & c & d\\ 0 & a & a+b & a+b+c\\ 0 & 0 & a & 2a+b\\ 0 & 0 & a & 3a+b\end{vmatrix}\overset{r_4-r_3}{=}\begin{vmatrix}a & b & c & d\\ 0 & a & a+b & a+b+c\\ 0 & 0 & a & 2a+b\\ 0 & 0 & 0 & a\end{vmatrix}=a^4$$

例 2.10　　计算行列式 $D_n = \begin{vmatrix} a_1 & x & x & \cdots & x \\ x & a_2 & x & \cdots & x \\ x & x & a_3 & \cdots & x \\ \vdots & \vdots & \vdots & & \vdots \\ x & x & x & \cdots & a_n \end{vmatrix}$

解：从第二行开始，每一行减去第一行

$$D_n = \begin{vmatrix} a_1 & x & x & \cdots & x \\ x-a_1 & a_2-x & 0 & \cdots & 0 \\ x-a_1 & 0 & a_3-x & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ x-a_1 & 0 & 0 & \cdots & a_n-x \end{vmatrix}$$

当$a_k \neq x(k=2,\cdots,n)$时，则可将第j列的$\dfrac{x-a_1}{x-a_j}$倍加到第一列

$$D_n = \begin{vmatrix} a_1+\dfrac{x-a_1}{x-a_2}x+\cdots+\dfrac{x-a_1}{x-a_n}x & x & x & \cdots & x \\ 0 & a_2-x & 0 & \cdots & 0 \\ 0 & 0 & a_3-x & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_n-x \end{vmatrix}$$

$$=\left(\frac{a_1}{x-a_1}+\frac{x}{x-a_2}+\cdots+\frac{x}{x-a_n}\right)(x-a_1)(a_2-x)\cdots(a_n-x)$$

若$a_1,\cdots,a_n$中只有一个等于x，不妨设$a_k = x$

$$D_n = \begin{vmatrix} a_1 & x & \cdots & x & \cdots & x \\ x & a_2 & \cdots & x & \cdots & x \\ \vdots & \vdots & & \vdots & & \vdots \\ x & x & \cdots & x & \cdots & x \\ \vdots & \vdots & & \vdots & & \vdots \\ x & x & \cdots & x & \cdots & a_n \end{vmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \text{第}k\text{行} \\ \\ \\ \end{array}$$

（每一行减去第 k 行）

作者：卢世荣　lshrlshr@163.com　　欢迎任何意见与建议！

$$=\begin{vmatrix} a_1-x & 0 & \cdots & 0 & \cdots & 0 \\ 0 & a_2-x & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ x & x & \cdots & x & \cdots & x \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & a_n-x \end{vmatrix}$$

$$=(a_1-x)\cdots(a_{k-1}-x)x(a_{k+1}-x)\cdots(a_n-x)$$

若$a_1,\cdots,a_n$中有多于一个等于x，则显然$D_n=0$

例 2.11　计算行列式 $D_n=\begin{vmatrix} 0 & 1 & 2 & \cdots & n-2 & n-1 \\ 1 & 0 & 1 & \cdots & n-3 & n-2 \\ 2 & 1 & 0 & \cdots & n-4 & n-3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ n-2 & n-3 & n-4 & \cdots & 0 & 1 \\ n-1 & n-2 & n-3 & \cdots & 1 & 0 \end{vmatrix}$。

解：（从最后一行开始，每一行减去前一行，得到）

$$=\begin{vmatrix} 0 & 1 & 2 & \cdots & n-2 & n-1 \\ 1 & -1 & -1 & \cdots & -1 & -1 \\ 1 & 1 & -1 & \cdots & -1 & -1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & -1 & -1 \\ 1 & 1 & 1 & \cdots & 1 & -1 \end{vmatrix} \overset{c_1+c_n}{=\!=} \begin{vmatrix} n-1 & 1 & 2 & \cdots & n-2 & n-1 \\ 0 & -1 & -1 & \cdots & -1 & -1 \\ 0 & 1 & -1 & \cdots & -1 & -1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & 1 & \cdots & -1 & -1 \\ 0 & 1 & 1 & \cdots & 1 & -1 \end{vmatrix}$$

（按第一列展开，得到）

$$=(n-1)\begin{vmatrix} -1 & -1 & \cdots & -1 & -1 \\ 1 & -1 & \cdots & -1 & -1 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & 1 & \cdots & -1 & -1 \\ 1 & 1 & \cdots & 1 & -1 \end{vmatrix}$$

（每一行都加上第一行）

$$=(n-1)\begin{vmatrix} -1 & -1 & \cdots & -1 & -1 \\ 0 & -2 & \cdots & -2 & -2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -2 & -2 \\ 0 & 0 & \cdots & 0 & -2 \end{vmatrix}$$

$$=(n-1)(-2)^{n-2}(-1)=(-1)^{n-1}2^{n-2}(n-1)$$

例 2.12　证明范德蒙行列式

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

$$\begin{aligned}&=(a_2-a_1)(a_3-a_1)(a_4-a_1)\cdots(a_n-a_1)\cdot\\ &\qquad(a_3-a_2)(a_4-a_2)\cdots(a_n-a_2)\cdot\\ &\qquad\cdots\\ &\qquad\qquad\qquad\cdot(a_n-a_{n-1})\\ &=\prod_{1\le i<j\le n}(a_j-a_i)\end{aligned}$$

证：用数学归纳法。$n=2$ 时命题显然成立。假设命题对 $n-1$ 阶范德蒙行列式成立，则

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

（从最后一行开始，每一行减去前一行的 a_1 倍）

$$=\begin{vmatrix}1 & 1 & \cdots & 1\\ 0 & a_2-a_1 & \cdots & a_n-a_1\\ 0 & a_2^2-a_1a_2 & \cdots & a_n^2-a_1a_n\\ \vdots & \vdots & & \vdots\\ 0 & a_2^{n-1}-a_1a_2^{n-2} & \cdots & a_n^{n-1}-a_1a_n^{n-1}\end{vmatrix}$$

（按第一列展开，得到）

$$=\begin{vmatrix}a_2-a_1 & \cdots & a_n-a_1\\ a_2^2-a_1a_2 & \cdots & a_n^2-a_1a_n\\ \vdots & & \vdots\\ a_2^{n-1}-a_1a_2^{n-2} & \cdots & a_n^{n-1}-a_1a_n^{n-1}\end{vmatrix}$$

（提取第 $i\left(1\le i\le n-1\right)$ 列的公因子 $a_{i+1}-a_1$）

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

而$\begin{vmatrix} 1 & \cdots & 1 \\ a_2 & \cdots & a_n \\ \vdots & & \vdots \\ a_2^{n-2} & \cdots & a_n^{n-1} \end{vmatrix}$是一个$n-1$阶范德蒙行列式，根据归纳假设，有

$$\begin{vmatrix} 1 & \cdots & 1 \\ a_2 & \cdots & a_n \\ \vdots & & \vdots \\ a_2^{n-2} & \cdots & a_n^{n-1} \end{vmatrix} = (a_3-a_2)(a_4-a_2)(a_5-a_2)\cdots(a_n-a_2)\cdot$$
$$(a_4-a_3)(a_5-a_3)\cdots(a_n-a_3)\cdot$$
$$\cdots$$
$$\cdot\ (a_n-a_{n-1})$$
$$= \prod_{2\le i<j\le n}(a_j-a_i)$$

所以

$$D(a_1,a_2,\cdots,a_n)$$
$$=(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)\cdot$$
$$(a_3-a_2)(a_4-a_2)(a_5-a_2)\cdots(a_n-a_2)\cdot$$
$$(a_4-a_3)(a_5-a_3)\cdots(a_n-a_3)\cdot$$
$$\cdots$$
$$\cdot\ (a_n-a_{n-1})$$
$$= \prod_{1\le i<j\le n}(a_j-a_i)$$

即命题对n阶范德蒙行列式成立。命题证毕。 ■

例 2.13　利用范德蒙行列式计算$D=\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{vmatrix}$的值。

解：这显然不是范德蒙行列式，但这个行列式与范德蒙行列式有关系，它是范德蒙行列式

$$D_1=\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ x & a & b & c & d \\ x^2 & a^2 & b^2 & c^2 & d^2 \\ x^3 & a^3 & b^3 & c^3 & d^3 \\ x^4 & a^4 & b^4 & c^4 & d^4 \end{vmatrix}$$

的第四行、第一个元素x^3的余子式，即M_{41}，根据范德蒙行列式的结论，有

$$D_1=\begin{vmatrix}1&1&1&1&1\\x&a&b&c&d\\x^2&a^2&b^2&c^2&d^2\\x^3&a^3&b^3&c^3&d^3\\x^4&a^4&b^4&c^4&d^4\end{vmatrix}$$

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

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

把 D_1 按第一列展开，

$$D_1=1\times A_{11}+x\times A_{21}+x^2\times A_{31}+x^3\times A_{41}+x^4\times A_{51}$$

即 D_1 是关于 x 的一个四次多项式，A_{41} 就是 D_1 中 x^3 的系数，比较系数得

$$A_{41}=(-1)^{4+1}M_{41}=(-1)^{4+1}D$$

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

$$\cdot(a+b+c+d)$$

所以，

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

例 2. 14　$D_4=\begin{vmatrix}1+x_1^2&x_1x_2&x_1x_3&x_1x_4\\x_2x_1&1+x_2^2&x_2x_3&x_2x_4\\x_3x_1&x_3x_2&1+x_3^2&x_3x_4\\x_4x_1&x_4x_2&x_4x_3&1+x_4^2\end{vmatrix}$

解：加边法

$$D_4=\begin{vmatrix}1&0&0&0&0\\x_1&1+x_1^2&x_1x_2&x_1x_3&x_1x_4\\x_2&x_2x_1&1+x_2^2&x_2x_3&x_2x_4\\x_3&x_3x_1&x_3x_2&1+x_3^2&x_3x_4\\x_4&x_4x_1&x_4x_2&x_4x_3&1+x_4^2\end{vmatrix}$$

（将第一列的 x_{j-1} 倍加到第 $j\left(2\le j\le 5\right)$ 列）

$$=\begin{vmatrix}1&-x_1&-x_2&-x_3&-x_4\\x_1&1&0&0&0\\x_2&0&1&0&0\\x_3&0&0&1&0\\x_4&0&0&0&1\end{vmatrix}$$

（将第 $i\left(2\le i\le 5\right)$ 行的 $-x_{i-1}$ 倍加到第一行）

$$=\begin{vmatrix} 1+x_1^2+x_2^2+x_3^2+x_4^2 & 0 & 0 & 0 & 0 \\ x_1 & 1 & 0 & 0 & 0 \\ x_2 & 0 & 1 & 0 & 0 \\ x_3 & 0 & 0 & 1 & 0 \\ x_4 & 0 & 0 & 0 & 1 \end{vmatrix}$$

$$=1+x_1^2+x_2^2+x_3^2+x_4^2$$

例 2.15　计算行列式

$$D=\begin{vmatrix} x-a & a & a & \cdots & a \\ a & x-a & a & \cdots & a \\ a & a & x-a & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & x-a \end{vmatrix}$$

解：

$$D\overset{c_1+c_2+\cdots+c_n}{=}\begin{vmatrix} x+(n-2)a & a & a & \cdots & a \\ x+(n-2)a & x-a & a & \cdots & a \\ x+(n-2)a & a & x-a & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ x+(n-2)a & a & a & \cdots & x-a \end{vmatrix}$$

$$\overset{c_1\div[x+(n-2)a]}{=}\left[x+(n-2)a\right]\begin{vmatrix} 1 & a & a & \cdots & a \\ 1 & x-a & a & \cdots & a \\ 1 & a & x-a & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a & a & \cdots & x-a \end{vmatrix}$$

$$\overset{\substack{r_2-r_1\\ r_3-r_1\\ \vdots\\ r_n-r_1}}{=}\left[x+(n-2)a\right]\begin{vmatrix} 1 & a & a & \cdots & a \\ 0 & x-2a & 0 & \cdots & 0 \\ 0 & 0 & x-2a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & x-2a \end{vmatrix}$$

$$=[x-(n-2)a](x-2a)^{n-1}$$

例 2.16　设 $D=\begin{vmatrix} a_{11} & \cdots & a_{1k} & & & \\ \vdots & & \vdots & & 0 & \\ a_{k1} & \cdots & a_{kk} & & & \\ c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{n1} & \cdots & c_{nk} & b_{n1} & \cdots & b_{nn} \end{vmatrix}$

$$D_1=\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}, \qquad D_2=\begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{vmatrix}$$

证明：$D=D_1D_2$

证：显然，本题用 Laplace 定理直接可得。另外也可用将值化为下三角行列式的方法来证明；也可以用行列式的定义来证明。现在讨论第二种方法。

对 D_1 作运算 r_i+kr_j，将 D_1 化为下三角行列式，设为

$$D_1=\begin{vmatrix} p_{11} & & 0 \\ \vdots & \ddots & \\ a_{k1} & \cdots & p_{kk} \end{vmatrix}=p_{11}\cdots p_{kk}$$

对 D_2 作运算 c_i+kc_j，把 D_2 化为下三角行列式，设为

$$D_2=\begin{vmatrix} q_{11} & & 0 \\ \vdots & \ddots & \\ q_{n1} & \cdots & q_{nn} \end{vmatrix}=q_{11}\cdots q_{nn}$$

对 D 的前 k 行作运算 r_i+kr_j，再对后 n 列作运算 c_i+kc_j，把 D 化为下三角行列式

$$D=\begin{vmatrix} p_{11} & & & & & \\ \vdots & \ddots & & & 0 & \\ p_{k1} & \cdots & p_{kk} & & & \\ c_{11} & \cdots & c_{1k} & q_{11} & & \\ \vdots & & \vdots & \vdots & \ddots & \\ c_{n1} & \cdots & c_{nk} & q_{n1} & \cdots & q_{nn} \end{vmatrix}$$

所以

$$D=p_{11}\cdots p_{kk}q_{11}\cdots q_{nn}=D_1D_2$$

实际上，本命题也可以用行列式的定义直接证明，有兴趣的读者可以尝试一下。

例 2.17　计算行列式 $D_n=\begin{vmatrix} a+b & a & 0 & \cdots & 0 & 0 \\ b & a+b & a & \cdots & 0 & 0 \\ 0 & b & a+b & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a+b & a \\ 0 & 0 & 0 & \cdots & b & a+b \end{vmatrix}$

解　将该行列式按最后一行展开 $D_n=(a+b)D_{n-1}-abD_{n-2}$，而显然 $D_1=a+b$，$D_2=a^2+ab+b^2$。由递推公式可得

$$D_n=\begin{cases} \dfrac{a^{n-1}-b^{n-1}}{a-b}D_2+\dfrac{ab\left(b^{n-2}-a^{n-2}\right)}{a-b}D_1, & a\neq b \\ (n-1)a^{n-2}D_2-(n-2)a^{n-1}D_1, & a=b \end{cases}$$

$$=\begin{cases} \dfrac{a^{n+1}-b^{n+1}}{a-b} & a\neq b \\ (n+1)a^n & a=b \end{cases}$$

习题 2.3

1. 请问在 7 阶行列式 $|a_{ij}|$ 中，$a_{33}a_{16}a_{72}a_{27}a_{55}a_{61}a_{44}$ 取正号还是取负号？在学习完本章之后，有几种方法来确定该项的符号？

2. 求 $\begin{vmatrix} 2x & -x & 1 & 3 \\ 2 & 3x & -1 & 2 \\ 1 & 2 & -x & 1 \\ 1 & 2 & 3 & x \end{vmatrix}$ 中 x^3 的系数。

3. 在 n 阶行列式 D 中，零元素的个数多于 n^2-n 个，则 $D=0$。

4. 已知 221，323，459 都能被 17 整除，不计算行列式的值，试证明：行列式 $D=\begin{vmatrix} 1 & 2 & 2 \\ 3 & 2 & 3 \\ 9 & 5 & 4 \end{vmatrix}$ 能被 17 整除。

5. 形如 $\begin{vmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ -a_{12} & 0 & a_{23} & \cdots & a_{2n} \\ -a_{13} & -a_{23} & 0 & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ -a_{1n} & -a_{2n} & -a_{3n} & \cdots & 0 \end{vmatrix}$ 的行列式称为**反对称行列式**，其特征是其元素满足 $a_{ij}=-a_{ji}$，特别是当 $i=j$ 时 $a_{ii}=0$。证明：奇数阶反对称行列式为 0。

6. 设 n 阶行列式 $D=\left|a_{ij}\right|$，把 D 上下翻转、或逆时针旋转 90，或依副对角线翻转，依次得

$$D_1=\begin{vmatrix} a_{n1} & \cdots & a_{nn} \\ \vdots & & \vdots \\ a_{11} & \cdots & a_{1n} \end{vmatrix}, D_2=\begin{vmatrix} a_{1n} & \cdots & a_{nn} \\ \vdots & & \vdots \\ a_{11} & \cdots & a_{n1} \end{vmatrix}, D_3=\begin{vmatrix} a_{nn} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{11} \end{vmatrix}$$

问：这三个行列式与 D 有何关系？

7. 求 $f(x)=\begin{vmatrix} x-2 & x-1 & x-2 & x-3 \\ 2x-2 & 2x-1 & 2x-2 & 2x-3 \\ 3x-3 & 3x-2 & 4x-2 & 3x-4 \\ 4x-1 & 4x-5 & 5x-3 & 4x-3 \end{vmatrix}=0$ 的根。

8. 计算下列行列式

（1）$\begin{vmatrix} 2 & 1 & 5 & 2 \\ 1 & 0 & 2 & 3 \\ 1 & 1 & 4 & 5 \\ -2 & 1 & 3 & 2 \end{vmatrix}$；　　（2）$\begin{vmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ b & b & \cdots & a \end{vmatrix}$；

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

（4）$\begin{vmatrix} 1 & -1 & 1 & x-1 \\ 1 & -1 & x+1 & -1 \\ 1 & x-1 & 1 & -1 \\ x+1 & -1 & 1 & -1 \end{vmatrix}$；

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

（6）$D_{n+1}=\begin{vmatrix} z_0 & x_1 & x_2 & \cdots & x_n \\ y_1 & z_1 & 0 & \cdots & 0 \\ y_2 & 0 & z_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ y_n & 0 & 0 & \cdots & z_n \end{vmatrix}$，　$z_1,z_2,\cdots,z_n\neq 0$；

（7）$D=\begin{vmatrix} a+x_1 & a & a & \cdots & a \\ a & a+x_2 & a & \cdots & a \\ a & a & a+x_3 & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & a+x_n \end{vmatrix}$；

（8）$D_{2n}=\begin{vmatrix} a & & & & & b \\ & \ddots & & & \iddots & \\ & & a & b & & \\ & & c & d & & \\ & \iddots & & & \ddots & \\ c & & & & & d \end{vmatrix}$；

（9）$\begin{vmatrix} a_0 & -1 & 0 & \cdots & 0 & 0 \\ a_1 & x & -1 & \cdots & 0 & 0 \\ a_2 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n-1} & 0 & 0 & \cdots & x & -1 \\ a_n & 0 & 0 & \cdots & 0 & x \end{vmatrix}=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$；

（10）$\begin{vmatrix} 1+a_1+b_1 & a_1+b_2 & \cdots & a_1+b_n \\ a_2+b_1 & 1+a_2+b_2 & \cdots & a_2+b_n \\ \vdots & \vdots & & \vdots \\ a_n+b_1 & a_n+b_2 & \cdots & 1+a_n+b_n \end{vmatrix}$；

（11）$\begin{vmatrix} \alpha & 0 & 0 & \cdots & 0 & \beta \\ \beta & \alpha & 0 & & 0 & 0 \\ 0 & \beta & \alpha & & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & \alpha & 0 \\ 0 & 0 & 0 & \cdots & \beta & \alpha \end{vmatrix}$；

（12）$D_{n+2}=\begin{vmatrix} \alpha & 0 & 0 & \cdots & 0 & \beta \\ x_1 & a & b & \cdots & b & y_1 \\ x_2 & b & a & \cdots & b & y_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ x_n & b & b & \cdots & a & y_n \\ \beta & 0 & 0 & \cdots & 0 & \alpha \end{vmatrix}$。

9. 证明：n 阶行列式 $\begin{vmatrix} 6 & 9 & & & & \\ 1 & 6 & 9 & & & \\ & 1 & 6 & 9 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 6 & 9 \\ & & & & 1 & 6 \end{vmatrix} = (n+1)3^n$。

10. 证明：n 阶行列式 $\begin{vmatrix} 2\cos\theta & 1 & 0 & \cdots & 0 & 0 \\ 1 & 2\cos\theta & 1 & \cdots & 0 & 0 \\ 0 & 1 & 2\cos\theta & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2\cos\theta & 1 \\ 0 & 0 & 0 & \cdots & 1 & 2\cos\theta \end{vmatrix} = \dfrac{\sin(n+1)\theta}{\sin\theta}$。

11. 计算 n 阶行列式 $\begin{vmatrix} 0 & x & x & \cdots & x \\ y & 0 & x & \cdots & x \\ y & y & 0 & \cdots & x \\ \vdots & \vdots & \vdots & & \vdots \\ y & y & y & \cdots & 0 \end{vmatrix}$。

12. 计算 n 阶行列式 $D_n = \begin{vmatrix} a & b & & & & \\ c & a & b & & & \\ & c & a & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & a & b \\ & & & & c & a \end{vmatrix}$。

13. 已知 n 阶行列式 $D = \begin{vmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n-1 \\ n & 0 & 0 & \cdots & 0 \end{vmatrix}$，试计算 D 的第 k 行元素的代数余子式之和 $A_{k1} + A_{k2} + \cdots + A_{kn}$。

14. 设 n 阶行列式 $D = \begin{vmatrix} 1 & 3 & 5 & \cdots & 2n-1 \\ 1 & 2 & 0 & \cdots & 0 \\ 1 & 0 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 2 & 3 & \cdots & n \end{vmatrix}$，求 $A_{11} + A_{12} + \cdots + A_{1n}$。

15. 设 n 阶行列式 $D=\begin{vmatrix}1 & 1 & \cdots & 1\\ 1 & 2 & \cdots & 1\\ \vdots & \vdots & & \vdots\\ 1 & 1 & \cdots & n\end{vmatrix}$，求行列式 D 的所有元素的代数余子式之和 $\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}$。

§2.4　克莱姆法则

在本章第一节引入了就未知数个数与方程个数相同的线性方程组有唯一解的情况进行了讨论，就 $n=1,2,3,4$ 给出了这种类型的线性方程组有唯一解的条件，并用低阶行列式给出了线性方程组的唯一解。在上述这几种情况下所得的结论，我们就一般情况提出了问题2.1，本节的内容就是给出这个问题的回答，这就是克莱姆定理。

定理 2.4　　克莱姆（Cramer）法则

若方程组

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

的系数行列式

$$D=\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}\neq 0$$

则该线性方程组有唯一解，且

$$x_j=\frac{\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn}\end{vmatrix}}{\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}}=\frac{D_j}{D} \tag{2.49}$$

说明：如果我们能够求解出这个线性方程组，自然就能够证明该结论。在用高斯消元法求解线性方程组时，其消元过程是逐步进行的，每次消去一个未知数。现在利用行列式的性质，

我们能一次性消去 $n-1$ 个未知数，使得仅剩下一个未知数，这样就能够得到方程组解的情况。

证明：对方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1j}x_j+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2j}x_j+\cdots+a_{2n}x_n=b_2 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\vdots \\ a_{i1}x_1+a_{i2}x_2+\cdots+a_{ij}x_j+\cdots+a_{in}x_n=b_i \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\vdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nj}x_j+\cdots+a_{nn}x_n=b_n \end{cases}$$

对于第 $i\left(i=1,2,\cdots n\right)$ 个方程，乘以方程组系数行列式中元素 a_{ij} 的代数余子式 A_{ij}，得

$$\begin{cases} a_{11}A_{1j}x_1+a_{12}A_{1j}x_2+\cdots+a_{1j}A_{1j}x_j+\cdots+a_{1n}A_{1j}x_n=b_1A_{1j} \\ a_{21}A_{2j}x_1+a_{22}A_{2j}x_2+\cdots+a_{2j}A_{2j}x_j+\cdots+a_{2n}A_{2j}x_n=b_2A_{2j} \\ \qquad\quad\vdots \\ a_{i1}A_{ij}x_1+a_{i2}A_{ij}x_2+\cdots+a_{ij}A_{ij}x_j+\cdots+a_{in}A_{ij}x_n=b_iA_{ij} \\ \qquad\quad\vdots \\ a_{n1}A_{nj}x_1+a_{n2}A_{nj}x_2+\cdots+a_{nj}A_{nj}x_j+\cdots+a_{nn}A_{nj}x_n=b_nA_{nj} \end{cases}$$

然后将这 n 个方程相加，我们得到

$$\begin{aligned} &\left(a_{11}A_{1j}+a_{21}A_{2j}+\cdots+a_{i1}A_{ij}+\cdots+a_{n1}A_{nj}\right)x_1 \\ &+\left(a_{12}A_{1j}+a_{22}A_{2j}+\cdots+a_{i2}A_{ij}+\cdots+a_{n2}A_{nj}\right)x_2 \\ &+\cdots \\ &+\left(a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{ij}A_{ij}+\cdots+a_{nj}A_{nj}\right)x_j \\ &+\cdots \\ &+\left(a_{1n}A_{1j}+a_{2n}A_{2j}+\cdots+a_{in}A_{ij}+\cdots+a_{nn}A_{nj}\right)x_n \\ &=b_1A_{1j}+b_2A_{2j}+\cdots+b_iA_{ij}+\cdots+b_nA_{nj} \end{aligned}$$

将等式左边用求和号表示

$$\sum_{k=1}^{n}\left(\sum_{i=1}^{n}a_{ik}A_{ij}\right)x_k=b_1A_{1j}+b_2A_{2j}+\cdots+b_iA_{ij}+\cdots+b_nA_{nj}$$

注意到：在上述等式的左边，未知数 x_k 前的系数是系数矩阵行列式的第 k 列元素乘以该行列式的第 j 列元素的代数余子式后的和，根据性质有

$$\sum_{i=1}^{n}a_{ik}A_{ij}=\begin{cases} D & k=j \\ 0 & k\neq j \end{cases}$$

由于 j 是选定的，因此 左边只有一项不为零，而等式右边显然为行列式 D_j 按第 j 列展开的

结果，因此就有

$$Dx_j = D_j$$

故当 $D \neq 0$ 时，x_j 有唯一解

$$x_j = \frac{D_j}{D}$$

上述步骤证明了若原始方程组有解，则其解必由(2.49)给出；下证明(2.49)确实给出了原始方程组的解，这只需要将(2.49)代入到每个方程中即可。考虑第 i 个方程

$$\sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} a_{ij} \frac{D_j}{D} = \frac{1}{D}\sum_{j=1}^{n} a_{ij} D_j \qquad \text{（将 } D_j \text{ 按第 } j \text{ 列展开）}$$

$$= \frac{1}{D}\sum_{j=1}^{n} a_{ij} \sum_{k=1}^{n} b_k A_{kj} \qquad \text{（交换求和次序）}$$

$$= \frac{1}{D}\sum_{k=1}^{n} b_k \sum_{j=1}^{n} a_{ij} A_{kj} \qquad \text{（应用结论）}$$

$$= \frac{1}{D}\sum_{k=1}^{n} b_k D \delta_{ik}$$

$$= \frac{1}{D} b_i D = b_i$$

故此时方程成立。

■

推论（1）如果齐次线性方程组

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

的系数行列式不为零，则该线性方程组只有零解。

（2）如果齐次线性方程组

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

有非零解，则其系数行列式等于 0。

问题：推论（2）的逆命题是否成立？即：如果上述齐次线性方程组的系数行列式等于零，其是否有非零解？目前我们还没有无法得到这个结论。

例 2. 18　　用克莱姆法则求解线性方程组

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

解：已知该线性方程组的系数行列式为

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

故该方程组有唯一解。根据克莱姆法则，有

$$x_1=\frac{\begin{vmatrix} 1 & 3 & 2 \\ 3 & -1 & 1 \\ 2 & 5 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 3 & 2 \\ 2 & -1 & 1 \\ 0 & 5 & 2 \end{vmatrix}}=\frac{15}{1}=15，\quad x_2=\frac{\begin{vmatrix} 1 & 1 & 2 \\ 2 & 3 & 1 \\ 0 & 2 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 3 & 2 \\ 2 & -1 & 1 \\ 0 & 5 & 2 \end{vmatrix}}=\frac{8}{1}=8，\quad x_3=\frac{\begin{vmatrix} 1 & 3 & 1 \\ 2 & -1 & 3 \\ 0 & 5 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 3 & 2 \\ 2 & -1 & 1 \\ 0 & 5 & 2 \end{vmatrix}}=\frac{-19}{1}=-19$$

例 2. 18　设平面上不在同一直线上的三点为$(x_i,y_i),i=1,2,3$，求通过这三点的圆方程的行列式形式。

解：设圆的一般方程为 $x^2+y^2+ax+by+c=0$

由题设，有等式

$$\begin{cases} (x^2+y^2)+ax+by+c=0 \\ (x_1^2+y_1^2)+ax_1+by_1+c=0 \\ (x_2^2+y_2^2)+ax_2+by_2+c=0 \\ (x_3^2+y_3^2)+ax_3+by_3+c=0 \end{cases}$$

将上式看成齐次线性方程组

$$\begin{cases} (x^2+y^2)z_1+xz_2+yz_3+z_4=0 \\ (x_1^2+y_1^2)z_1+x_1z_2+y_1z_3+z_4=0 \\ (x_2^2+y_2^2)z_1+x_2z_2+y_2z_3+z_4=0 \\ (x_3^2+y_3^2)z_1+x_3z_2+y_3z_3+z_4=0 \end{cases}$$

则$(1,a,b,c)$是该齐次线性方程组的一组非零解，则其系数行列式为零，即

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

这就是通过三点的圆的方程。

习题 2.4

1. 用 Cramer 法则求解下列方程组

（1）$\begin{cases} 3x_1 \quad\quad +2x_3=1 \\ x_1+7x_2+4x_3=4 \\ 7x_1+3x_2+2x_3=3 \end{cases}$　　（2）$\begin{cases} 2x_1+x_2+3x_3+x_4=6 \\ 3x_1+x_2-2x_3+x_4=3 \\ x_1+3x_2+4x_3-3x_4=5 \\ 5x_1+x_2+2x_3+5x_4=0 \end{cases}$

2. 试问当 k 取何值时，如下齐次线性方程组有非零解？

$$\begin{cases} kx_1+2x_1+x_3=0 \\ 6x_1+kx_2+x_3=0 \\ 2x_1+3x_2+4x_3=0 \end{cases}$$

3. 设 $x_1,x_2,\cdots,x_n$ 互不相同，$(x_i,y_i),i=1,2,\cdots,n$ 是平面上的 n 个点，证明：存在唯一的次数不超过 $n-1$ 的多项式 $p(x)$，使得 $p(x_i)=y_i,i=1,2,\cdots,n$。

4. 求空间 4 个平面 $a_ix+b_iy+c_iz+d_i=0\,(i=1,2,3,4)$ 相交于一点的条件。

5. 求通过不在同一个平面的 4 点 $(x_i,y_i,z_1),i=1,2,3,4$ 的球面的方程。

第三章　矩阵

§3.1　矩阵的概念

定义 3.1　　矩阵

数域$\mathbb{F}$（可以将$\mathbb{F}$认为是实数域$\mathbb{R}$或复数域$\mathbb{C}$）上的$m\times n$个数

$$a_{ij},\qquad i=1,2,\cdots,m;\ j=1,2,\cdots,n;$$

排成m行、n列的数表

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

为了表示这是一个整体，我们用方括号将之括起，记为

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{3.1}$$

称$\boldsymbol{A}$为$m\times n$矩阵，记为$\boldsymbol{A}=\left[a_{ij}\right]_{m\times n}$或$\boldsymbol{A}=\left[a_{ij}\right]$，$a_{ij}$表示矩阵中位于第$i$行、第$j$列的元素，$a_{ij}$第一个下标$i$称为**行标**，第二个下标$j$称为**列标**；矩阵$\boldsymbol{A}$的位于第$i$行、第$j$列的元素也记为$\left[\boldsymbol{A}\right]_{ij}$。元素属于实数域的矩阵称为**实矩阵**，元素属于复数域的矩阵称为**复矩阵**。在本书中，矩阵用大写斜体加粗字母来表示，如$\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\cdots$，而矩阵元素用小写斜体字母来表示，如$a_{ij},b_{ij},c_{ij},\cdots$。数域$\mathbb{F}$上的所有$m\times n$矩阵构成的集合记作$\mathbb{F}^{m\times n}$。

定义 3.2　　同型矩阵

若两个矩阵的行数、列数分别相等，则称它们为同型矩阵。

定义 3.3　　矩阵相等

如果矩阵$\boldsymbol{A},\boldsymbol{B}$为同型矩阵，且它们位于相同位置的元素相等，即若

$$\boldsymbol{A}=\left[a_{ij}\right]_{m\times n},\quad \boldsymbol{B}=\left[b_{ij}\right]_{m\times n}$$

且有

$$a_{ij}=b_{ij},\qquad i=1,2,\cdots,m;\ j=1,2,\cdots,n;$$

则 $\boldsymbol{A},\boldsymbol{B}$ 相等，记作 $\boldsymbol{A}=\boldsymbol{B}$。

有了矩阵相等的概念，我们就可定义一些特殊的矩阵：

(1)　**方阵**：行数与列数相等的矩阵称为方阵，行数为 n 的方阵称为 n 阶方阵，记为 $\boldsymbol{A}_{n\times n}$ 或 $\boldsymbol{A}_n$，在不强调方阵阶数的情况下也简记做 $\boldsymbol{A}$。

对于方阵 $\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}$，元素 $a_{11}, a_{22},\cdots,a_{nn}$ 决定的直线称为方阵的**主对角线**，而把 $a_{11}, a_{22},\cdots,a_{nn}$ 称为**主对角线元素**；而 $a_{1n}, a_{2,n-1},\cdots,a_{n-1,2},a_{n1}$ 决定的直线称为方阵的**副对角线**，并称 $a_{1n}, a_{2,n-1},\cdots,a_{n-1,2},a_{n1}$ 为**副对角线元素**。

(2)　**零矩阵**：所有元素都为 0 的矩阵，记为 $\boldsymbol{O}$。
注意：零矩阵可以不为方阵。
根据定义，不同型的零矩阵是不相等的。

(3)　**行矩阵**（行向量）：行数为 1 的矩阵。
行矩阵的列数称为行矩阵的**维数**。
例如，n 维行矩阵

$$\boldsymbol{A}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\end{bmatrix}$$

由于行矩阵只有一行，因此元素的行标可省略，即行矩阵可记为

$$\boldsymbol{A}=\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}$$

称 a_i 为该行矩阵的第 i 个**分量**。行矩阵也写作 $\boldsymbol{A}=\left[a_1,a_2,\cdots,a_n\right]$。

(4)　**列矩阵**（列向量）：列数为 1 的矩阵。
列矩阵的行数称为列矩阵的维数。例如，n 维列向量

$$\boldsymbol{A}=\begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{bmatrix}$$

称 a_i 为该列矩阵的第 i 个**分量**。

(5)　**对角矩阵**：主对角线之外的元素都为 0 的 n 阶方阵称为对角矩阵，记为

$$\boldsymbol{A}=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}=diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$$

约定：矩阵中凡是没有写出的元素都为 0。

(6)　**单位矩阵**：主对角线上的元素都为 1 的 n 阶对角矩阵称为 n 阶单位矩阵，记为 $\boldsymbol{E}_n$，即

$$\boldsymbol{E}=\begin{bmatrix}1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{bmatrix}=diag\left(1,1,\cdots,1\right)$$

如果不强调单位矩阵 $\boldsymbol{E}_n$ 的阶数，则简记为 $\boldsymbol{E}$。

(7)　**上（下）三角矩阵**

主对角线下方的元素都为零的方阵称为**上三角矩阵**，如

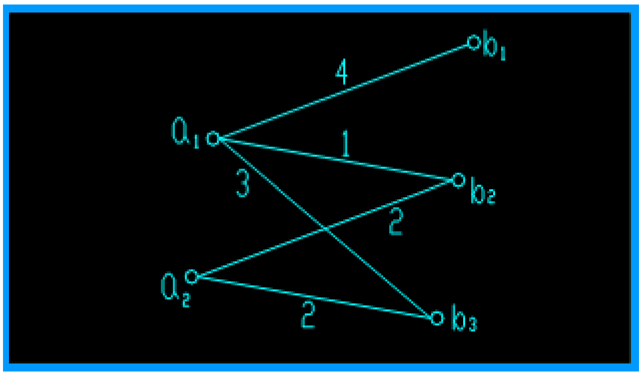
$$\boldsymbol{A}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn}\end{bmatrix}$$

主对角线上方的元素都为零的方阵称为**下三角矩阵**，如

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

矩阵应用的例子

例 3.1　设 a 省有两个城市 a_1,a_2 与 b 省的三个城市 b_1,b_2,b_3 的交通连接情况如图，



图中每条线上的数字表示联结该两城市的不同通路总数，由该图提供的通路信息，可用矩阵形式表示，称之为通路矩阵。

$$\begin{array}{cc} \begin{array}{cc} a_1 & a_2 \end{array} & \\ \begin{bmatrix}4 & 0 \\ 1 & 2 \\ 3 & 2\end{bmatrix} & \begin{array}{c} b_1 \\ b_2 \\ b_3 \end{array} \end{array}$$

该矩阵中第 i 行、第 j 列的元素表示 a_j,b_i 这两个城市之间的道路数目。

例 3.2　价格矩阵

四种食品(Food)在三家商店(Shop)中,单位量的售价(以某种货币单位计)可用以下矩阵给出

$$\begin{array}{cc} & \begin{array}{cccc} F_1 & F_2 & F_3 & F_4 \end{array} \\ \boldsymbol{A}= & \begin{bmatrix} 17 & 7 & 11 & 21 \\ 15 & 9 & 13 & 19 \\ 18 & 8 & 15 & 19 \end{bmatrix} \end{array} \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array}$$

其中行表示商店，而列表示食品种类。该矩阵中第 i 行、第 j 列的元素表示商店 S_i 中食品 F_j 的价格。

李四希望在某个商店一次性购齐四种食品，四种食品的数量分别为 2，3，2，1，那么他到哪个商店去采购，所花的钱最少？

显然，我们可以计算得到在每个商店中购齐这四种食品所需的花费：

S_1:$17\times2+7\times3+11\times2+21\times1=98$

S_2:$15\times2+9\times3+13\times2+19\times1=102$

S_3:$18\times2+8\times3+15\times2+19\times1=109$

因此在 S_1 中购买比较经济。如果我们也用矩阵来描述四种食品的购买量，

$$\boldsymbol{B}=\begin{bmatrix} 2 \\ 3 \\ 2 \\ 1 \end{bmatrix} \begin{array}{c} F_1 \\ F_2 \\ F_3 \\ F_4 \end{array}$$

那么在 S_1 中购买这四种食品所需的花费为：矩阵 $\boldsymbol{A}$ 第一行的四个数与矩阵 $\boldsymbol{B}$ 的这四个数对应相乘、然后相加的结果。

例 3.3　　　赢得矩阵

我国古代有“齐王赛马”的事例，说的是战国时代齐王与其大将田忌赛马，双方约定各出上、中、下 3 个等级的马各一匹进行比赛，这样共赛马 3 次，每次比赛的败者付给胜者一百金，已知在同一等级马的比赛中，齐王之马可稳操胜券，但田忌的上、中等级的马分别可胜齐王中、下等级的马。首先将马的出场次序进行编号，我们把三匹马的出场次序称为一个策略，因此可能采用的比赛策略为：

(上、中、下)	(中、上、下)	(下、中、上)
1	2	3
(上、下、中)	(中、下、上)	(下、上、中)
4	5	6

现在给出齐王的赢得矩阵（单位：百金），即齐王、田忌分别采取相应策略时齐王所赢得的

赌资：

$$\begin{array}{cc} & \rightarrow \text{田忌策略} \\ \downarrow \text{齐王策略} & \begin{bmatrix} 3 & 1 & 1 & 1 & 1 & -1 \\ 1 & 3 & 1 & -1 & 1 & 1 \\ 1 & -1 & 3 & 1 & 1 & 1 \\ 1 & 1 & -1 & 3 & 1 & 1 \\ -1 & 1 & 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 1 & -1 & 3 \end{bmatrix} \end{array}$$

在该矩阵中，行表示齐王所采取的策略，而列表示田忌所采取的策略。该矩阵中第 i 行、第 j 列的元素表示齐王采取第 i 种策略、田忌采取第 j 种策略时齐王所赢取的钱数。例如，该矩阵中的第二行、第三列的元素为 1，即说明当齐王采取第二个策略、田忌采用第三个策略时，齐王能赢 100 金。

例 3.4　**线性变换**的**系数矩阵**

n 个变量 $x_1, x_2, \cdots x_n$ 与 m 个变量 $y_1, y_2, \cdots y_m$ 之间的关系

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \quad \vdots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{cases} \tag{3.2}$$

（其中 a_{ij} 为常数）

我们称之为从 $y_1, y_2, \cdots y_m$ 到 $x_1, x_2, \cdots x_n$ 的**线性变换**，把

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

称为该线性变换的（**系数**）**矩阵**。

例3.5　线性方程组的**系数矩阵**、**增广矩阵**

对于线性方程组 $\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$，分别称矩阵

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} & b_i \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

为该方程组的**系数矩阵**与**增广矩阵**。

矩阵的另外一种视角：

我们知道，函数的**映射法则**（自变量、函数值之间的对应关系）一般有四种表示方式：解析法、图示法、列表法、说明法。在本质上，我们可把矩阵看作是函数的列表法表示。并且，凡是可用矩阵表示的函数须满足如下要求（我们只讨论矩阵的行数、列数均为有限数的情况）：

（1）函数有两个变量，且每个变量只有有限个取值；

（2）这两个变量的每一对值，函数都有定义，且函数的值域为数域$\mathbb{F}$；

例如，设函数 $f(x,y)$ 满足上述要求，且 x,y 的取值分别为

$$x: x_1, x_2, \cdots, x_m$$
$$y: y_1, y_2, \cdots, y_n$$

那么该函数自变量与函数值之间的对应关系可表示为

	y_1	y_2	$\cdots$	y_n
x_1	$f(x_1,y_1)$	$f(x_1,y_2)$	$\cdots$	$f(x_1,y_n)$
x_2	$f(x_2,y_1)$	$f(x_2,y_2)$	$\cdots$	$f(x_2,y_n)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x_m	$f(x_m,y_1)$	$f(x_m,y_2)$	$\cdots$	$f(x_m,y_n)$

因此，这显然是一个矩阵的形式

$$\begin{bmatrix} f(x_1,y_1) & f(x_1,y_2) & \cdots & f(x_1,y_n) \\ f(x_2,y_1) & f(x_2,y_2) & \cdots & f(x_2,y_n) \\ \vdots & \vdots & & \vdots \\ f(x_m,y_1) & f(x_m,y_2) & \cdots & f(x_m,y_n) \end{bmatrix}$$

到目前为止，我们学习了行列式与矩阵。从形式上看，行列式与矩阵都是很多数排成“方方正正”的形式，因此是“形似”，但根据行列式与矩阵的定义，其区别是显然的：

1、“联系”

从形式上看，这两者都是一些数排成若干行、若干列的形式；

2、区别

- 形式上：矩阵的行数、列数可以不相等，而行列式的行数、列数必须相等；
- 本质上而言，矩阵是一个数表，它只是一些数排成的若干行、若干列的形式，它可以看成一类二元函数的表示方式；而 n 阶行列式为本质上是一个具有 n^2 个自变量的函数，所得结果为一个数，其自变量就是行列式中的各个数，之所以把行列式这个函数写成 n 行、n 列的形式是为了便于地给出行列式的定义。
- 由于行列式的结果是一个数，所以不同阶数的行列式是可以相等的；而根据矩阵相等的定义，不同型的矩阵没有相等、不等的概念。
- 行列式运算性质是根据行列式的定义推导出来的，以后会看到，关于矩阵的运算是分别定义的，而不是由矩阵定义推导出来的；

对于一个数学对象，其所拥有的运算是由它所描述的客观对象所决定的，抽象地说一个符号具有何种运算并无意义。因此，可考虑如下问题：

问题 3.1：举一些能够用矩阵表示的例子，并说明对于这些矩阵，可以有哪些运算？

我们下面要讲述的矩阵运算是比较普遍的、应用最多的情形，并不是说我们穷尽了矩阵的所有可能运算，对于实际中的矩阵，根据其所描述对象的特性，可能具有其它的运算。

习题 3.1

1. 请举例能够用矩阵描述的例子，对于这些描述实际情况的矩阵例子，可以有哪些运算？

§3.2　矩阵的运算

3.2.1　矩阵加法与数乘运算

定义 3.4　矩阵加法

两个同型矩阵 $\boldsymbol{A}=\left[a_{ij}\right]_{m\times n}$ 与 $\boldsymbol{B}=\left[b_{ij}\right]_{m\times n}$ 的加法定义为

$$\boldsymbol{C}=\boldsymbol{A}+\boldsymbol{B}=\left[c_{ij}\right]_{m\times n}=\left[a_{ij}+b_{ij}\right]_{m\times n}$$

其中　$c_{ij}=a_{ij}+b_{ij}$。

也就是说：矩阵相加即为这两个矩阵的相应（相同位置的）元素相加。

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

根据矩阵加法的定义，我们很容易得到矩阵加法的运算规律。

矩阵加法的运算规律：

1. 交换律：$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$；
2. 结合律：$(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$；

3.　零矩阵：$\boldsymbol{A}+\boldsymbol{O}=\boldsymbol{A}$。

设 $\boldsymbol{A}=\left[\boldsymbol{a}_{ij}\right]$，定义

$$-\boldsymbol{A}=\left[-a_{ij}\right]$$

称 $-\boldsymbol{A}$ 为 $\boldsymbol{A}$ 的**负矩阵**。显然，$\boldsymbol{A}+(-\boldsymbol{A})=\boldsymbol{O}$。规定矩阵减法为

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

所以

$$\left[a_{ij}\right]_{m\times n}-\left[b_{ij}\right]_{m\times n}=\left[a_{ij}-b_{ij}\right]_{m\times n}$$

即矩阵减法为两个矩阵对应的元素相减。

定义 3.6　　矩阵的**数乘**

　　设 $k\in\mathbb{F}$，$\boldsymbol{A}=\left[a_{ij}\right]_{m\times n}\in\mathbb{F}^{m\times n}$，则数 k 与矩阵 $\boldsymbol{A}$ 的数乘定义为：

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

简记为　$k\boldsymbol{A}=k\left[a_{ij}\right]_{m\times n}=\left[ka_{ij}\right]_{m\times n}$。

因此，数与矩阵数乘相当于这个数乘以该矩阵的每个元素。

矩阵的加法、数乘统称为矩阵的**线性运算**。

根据定义，容易验证矩阵的线性运算符合如下运算规则，其中 $\boldsymbol{k},l\in\mathbb{F},\boldsymbol{A},\boldsymbol{B}\in\mathbb{F}^{m\times n}$：

1.　结合律：$(kl)\boldsymbol{A}=k(l\boldsymbol{A})=l(k\boldsymbol{A})$

2.　分配律：　$k(\boldsymbol{A}+\boldsymbol{B})=k\boldsymbol{A}+k\boldsymbol{B}$

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

3.　$1\boldsymbol{A}=\boldsymbol{A},\ \ 0\boldsymbol{A}=\boldsymbol{O}$

例 3.6　　已知 $\boldsymbol{A}=\begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 2 \end{bmatrix}$　$\boldsymbol{B}=\begin{bmatrix} -2 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$，求 $\boldsymbol{C}=2\boldsymbol{A}-3\boldsymbol{B}$。

解：$\boldsymbol{C}=2\boldsymbol{A}-3\boldsymbol{B}$

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

$$=\begin{bmatrix}2&0&2\\-2&0&4\end{bmatrix}-\begin{bmatrix}-6&3&0\\3&3&6\end{bmatrix}$$

$$=\begin{bmatrix}2-(-6)&0-3&2-0\\-2-3&0-3&4-6\end{bmatrix}$$

$$=\begin{bmatrix}8&-3&2\\-5&-3&-2\end{bmatrix}$$

3.2.2　矩阵乘法

矩阵的乘法是一个全新的概念，这与我们目前所学过的知识迥然不同。引入矩阵乘法的一个简单背景是线性变换。考虑如下两个线性变换及相应的矩阵

$$\begin{cases}y_1=a_{11}x_1+a_{12}x_2+a_{13}x_3\\y_2=a_{21}x_1+a_{22}x_2+a_{23}x_3\\y_3=a_{31}x_1+a_{32}x_2+a_{33}x_3\end{cases}\qquad\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}$$

$$\begin{cases}z_1=b_{11}y_1+b_{12}y_2+b_{13}y_3\\z_2=b_{21}y_1+b_{22}y_2+b_{23}y_3\end{cases}\qquad\begin{bmatrix}b_{11}&b_{12}&b_{13}\\b_{21}&b_{22}&b_{23}\end{bmatrix}$$

显然，以上两个变换决定了一个从 z_1,z_2 到 x_1,x_2,x_3 的线性变换，我们可以求得该线性变换的矩阵。

$$\begin{aligned}z_1&=b_{11}\left(a_{11}x_1+a_{12}x_2+a_{13}x_3\right)+b_{12}\left(a_{21}x_1+a_{22}x_2+a_{23}x_3\right)+b_{13}\left(a_{31}x_1+a_{32}x_2+a_{33}x_3\right)\\&=\left(b_{11}a_{11}+b_{12}a_{21}+b_{13}a_{31}\right)x_1+\left(b_{11}a_{12}+b_{12}a_{22}+b_{13}a_{32}\right)x_2+\left(b_{11}a_{13}+b_{12}a_{23}+b_{13}a_{33}\right)x_3\end{aligned}$$

$$\begin{aligned}z_2&=b_{21}\left(a_{11}x_1+a_{12}x_2+a_{13}x_3\right)+b_{22}\left(a_{21}x_1+a_{22}x_2+a_{23}x_3\right)+b_{23}\left(a_{31}x_1+a_{32}x_2+a_{33}x_3\right)\\&=\left(b_{21}a_{11}+b_{22}a_{21}+b_{23}a_{31}\right)x_1+\left(b_{21}a_{12}+b_{22}a_{22}+b_{23}a_{32}\right)x_2+\left(b_{21}a_{13}+b_{22}a_{23}+b_{23}a_{33}\right)x_3\end{aligned}$$

所以从 z_1,z_2 到 x_1,x_2,x_3 的线性变换的矩阵为

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

我们可观察这个矩阵与前两个线性变换的矩阵的关系，这就得到矩阵乘法的概念。

定义 3.6　矩阵乘法

设 $m\times k$ 矩阵 $\boldsymbol{A}$ 、 $k\times n$ 矩阵 $\boldsymbol{B}$ 分别为

$$\boldsymbol{A}=\left[a_{ij}\right]_{m\times k}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\boldsymbol{B}=\left[b_{ij}\right]_{k\times n}=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{bmatrix}$$

定义它们的乘法

$$\boldsymbol{C}=\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{AB}$$

其中$\boldsymbol{C}$为一个$m\times n$矩阵，若记$\boldsymbol{C}=\left[c_{ij}\right]_{m\times n}$，则其元素

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{ik}b_{kj}$$

因此在计算矩阵$\boldsymbol{C}$的第i行、j列元素c_{ij}时所用到的矩阵$\boldsymbol{A}$中的元素为

$$a_{i1}\quad a_{i2}\quad \cdots \quad a_{ik}$$

这是矩阵$\boldsymbol{A}$的第i行元素；而所用到的矩阵$\boldsymbol{B}$中的元素为

$$\begin{matrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{kj} \end{matrix}$$

这是矩阵$\boldsymbol{B}$的第j列元素。

因而矩阵乘法的规则用图示可表示为

$$\begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ik} \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix}\begin{bmatrix} \cdots & b_{1j} & \cdots \\ \cdots & b_{2j} & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & b_{kj} & \cdots \end{bmatrix}=\begin{bmatrix} & \vdots & \\ \cdots & c_{ij} & \cdots \\ & \vdots & \end{bmatrix}$$

根据矩阵乘法的定义，矩阵乘法$\boldsymbol{C}=\boldsymbol{AB}$有如下特点：

(1)　矩阵$\boldsymbol{A}$的列数必须等于矩阵$\boldsymbol{B}$的行数；

(2)　矩阵$\boldsymbol{C}$的行数等于矩阵$\boldsymbol{A}$的行数，$\boldsymbol{C}$的列数等于矩阵$\boldsymbol{B}$的列数；

(3)　矩阵$\boldsymbol{C}$的第i行、第j列元素等于矩阵$\boldsymbol{A}$的第i行元素与矩阵$\boldsymbol{B}$第j列对应元素乘积之和。

在$\boldsymbol{C}=\boldsymbol{AB}$中，称$\boldsymbol{A}$（左）乘$\boldsymbol{B}$，或$\boldsymbol{B}$右乘$\boldsymbol{A}$。

例 3.7　$\boldsymbol{A}=\begin{bmatrix}1&0&-1\\2&0&1\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}1&2\\-1&2\\0&0\end{bmatrix}$，求 $\boldsymbol{C}=\boldsymbol{AB}$ 、$\boldsymbol{D}=\boldsymbol{BA}$ 。

解：由于 $\boldsymbol{A}$ 的列数、$\boldsymbol{B}$ 的行数都为 3，因此 $\boldsymbol{AB}$ 有意义，且 $\boldsymbol{AB}$ 为一个 2×2 矩阵，

$$c_{11}=a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31}=1\times1+0\times(-1)+(-1)\times0=1$$

$$c_{12}=a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32}=1\times2+0\times2+(-1)\times0=2$$

$$c_{21}=a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31}=2\times1+0\times(-1)+1\times0=2$$

$$c_{22}=a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32}=2\times2+0\times2+1\times0=4$$

所以 $\boldsymbol{C}=\begin{bmatrix}1&2\\2&4\end{bmatrix}$。同理可得 $\boldsymbol{D}=\boldsymbol{BA}=\begin{bmatrix}5&0&1\\3&0&3\\0&0&0\end{bmatrix}$

矩阵乘法满足如下运算律：

1.　结合律：$(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC})$；

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

　　右分配律：$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}=\boldsymbol{AC}+\boldsymbol{BC}$

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

4.　设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，则

$$\boldsymbol{E_m A}=\boldsymbol{A}\qquad \boldsymbol{AE_n}=\boldsymbol{A}\text{。}$$

证明　　这里仅证明矩阵乘法的结合律，其余证明留给读者。设

$$\boldsymbol{A}=\left[a_{ij}\right]_{m\times k},\boldsymbol{B}=\left[b_{ij}\right]_{k\times l},\boldsymbol{C}=\left[c_{ij}\right]_{l\times n}$$

容易验证 $(\boldsymbol{AB})\boldsymbol{C}$ 与 $\boldsymbol{A}(\boldsymbol{BC})$ 为同型矩阵。

这里再一次声明：我们以 $[\boldsymbol{M}]_{ij}$ 表示位于矩阵 $\boldsymbol{M}$ 的第 i 行、第 j 列的元素。易知 $\boldsymbol{AB}$ 为 $m\times l$ 矩阵，$\boldsymbol{BC}$ 为 $k\times n$ 矩阵，因此

$$[(\boldsymbol{AB})\boldsymbol{C}]_{ij}=\sum_{t=1}^{l}[\boldsymbol{AB}]_{it}[\boldsymbol{C}]_{tj}=\sum_{t=1}^{l}\left(\sum_{s=1}^{k}a_{is}b_{st}\right)c_{tj}$$

$$=\sum_{t=1}^{l}\sum_{s=1}^{k}a_{is}b_{st}c_{tj}=\sum_{s=1}^{k}a_{is}\sum_{t=1}^{l}b_{st}c_{tj}=\sum_{s=1}^{k}a_{is}\sum_{t=1}^{l}b_{st}c_{tj}$$

$$= \sum_{s=1}^{k} a_{is} \left(\sum_{t=1}^{l} b_{st} c_{tj} \right) = \sum_{s=1}^{k} a_{is} \left[\boldsymbol{BC} \right]_{sj}$$

$$= \left[\boldsymbol{A}(\boldsymbol{BC}) \right]_{ij}$$

这里需要强调的是矩阵乘法与（实）数之间乘法的区别。

1、矩阵乘法一般不具有交换律。这分为三种情况：

a）$\boldsymbol{AB}$ 有定义，$\boldsymbol{BA}$ 没有定义；

例如，$\boldsymbol{A}$ 为 2×3 矩阵，$\boldsymbol{B}$ 为 3×5 矩阵；

b）$\boldsymbol{AB}$、$\boldsymbol{BA}$ 都有定义，但 $\boldsymbol{AB}$、$\boldsymbol{BA}$ 不是同型矩阵，因此根本不存在是否相等的问题。

例如，$\boldsymbol{A}$ 为 2×3 矩阵，$\boldsymbol{B}$ 为 3×2 矩阵；因此 $\boldsymbol{AB}$ 为 2 阶方阵，$\boldsymbol{BA}$ 为 3 阶方阵；

c）$\boldsymbol{AB}$、$\boldsymbol{BA}$ 都有意义，且两者是同型矩阵，但不相等。例如

$$\boldsymbol{A}=\begin{bmatrix}1 & 1\\ 2 & 1\end{bmatrix}, \boldsymbol{B}=\begin{bmatrix}3 & 2\\ 1 & 2\end{bmatrix}, \boldsymbol{AB}=\begin{bmatrix}4 & 4\\ 7 & 6\end{bmatrix}, \boldsymbol{BA}=\begin{bmatrix}7 & 5\\ 5 & 3\end{bmatrix};$$

d）$\boldsymbol{AB}$、$\boldsymbol{BA}$ 都有意义，且两者是同型矩阵，并且相等，此时称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$（乘法）可交换。例如

$$\boldsymbol{A}=\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}, \boldsymbol{B}=\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}, \boldsymbol{AB}=\begin{bmatrix}3 & 5\\ 5 & 8\end{bmatrix}=\boldsymbol{BA}$$

2、矩阵乘法一般不具有消去律，即由 $\boldsymbol{AB}=\boldsymbol{AC}$ 并不能无条件得到 $\boldsymbol{B}=\boldsymbol{C}$。例如

$$\boldsymbol{A}=\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}, \boldsymbol{B}=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}, \boldsymbol{C}=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}，显然 \boldsymbol{B}\neq\boldsymbol{C}，但 \boldsymbol{AB}=\boldsymbol{AC}=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}。$$

同样，仅从 $\boldsymbol{BA}=\boldsymbol{CA}$ 并不能得到 $\boldsymbol{B}=\boldsymbol{C}$；由 $\boldsymbol{AB}=\boldsymbol{O}$ 并不能得到 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{B}=\boldsymbol{O}$。例如

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

但是对矩阵某些矩阵 $\boldsymbol{A}$ 而言，则可从 $\boldsymbol{AB}=\boldsymbol{AC}$ 得到 $\boldsymbol{B}=\boldsymbol{C}$，我们称矩阵 $\boldsymbol{A}$ 满足（乘法）左消去律，例如，$\boldsymbol{A}=\begin{bmatrix}1 & \\ & 2\end{bmatrix}$；对某些矩阵 $\boldsymbol{A}$ 而言，从 $\boldsymbol{BA}=\boldsymbol{CA}$ 可得到 $\boldsymbol{B}=\boldsymbol{C}$，我们称矩阵 $\boldsymbol{A}$ 满足（乘法）右消去律。如果矩阵 $\boldsymbol{A}$ 同时满足左消去律、右消去律，则称 $\boldsymbol{A}$ 满足消去律。至于矩阵 $\boldsymbol{A}$ 满足乘法左、右消去律的充要条件，我们在以后将逐步展开讲述。

3、1 与任何数 a 相乘等于该数 a；单位矩阵与任何矩阵 $\boldsymbol{A}$ 相乘（如果乘法可以进行）等于该矩阵 $\boldsymbol{A}$。

我们这里考虑矩阵乘法一种有趣的情形。假设有三个矩阵

$$\boldsymbol{A}_{n\times1}=\left[a_i\right], \boldsymbol{B}_{1\times1}=\left[b\right], \boldsymbol{C}_{1\times n}=\left[c_j\right]$$

考虑 $\boldsymbol{AB}, \boldsymbol{BC}$ 的结果。根据矩阵乘法、矩阵数乘的定义，有

$$\boldsymbol{AB}=\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}[b]=\begin{bmatrix}a_1b\\a_2b\\\vdots\\a_nb\end{bmatrix}=b\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}=b\boldsymbol{A}$$

$$\begin{aligned}\boldsymbol{BC}&=[b][c_1\quad c_2\quad\cdots\quad c_n]=[bc_1\quad bc_2\quad\cdots\quad bc_n]\\&=b[c_1\quad c_2\quad\cdots\quad c_n]=b\boldsymbol{C}\end{aligned}$$

因此，如果某个矩阵乘以一个1×1矩阵（当然要求这个乘法是可行的），完全可以把这个1×1矩阵当作一个数，而不会影响最终的计算结果。后文中的例 3.7 就是这样的一个例子。

线性方程组的矩阵乘法表示。

对于线性方程组

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

矩阵 $\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}$ 为该方程组的系数矩阵，若记 $\boldsymbol{X}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}$，则该线性方程组可用矩阵乘法表示成

$$\boldsymbol{AX}=\boldsymbol{b}$$

因此，线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 与含有一个未知数的方程 $ax=b$ 在形式上是相同的，当然这两个式子中各个符号的意义是不同的，出现在这两个式子中的乘法也是不相同的，$\boldsymbol{AX}$ 是矩阵乘法，而 ax 是数的乘法。

方阵的幂、方阵的多项式

设 $\boldsymbol{A}$ 为 n 阶方阵，定义 $\boldsymbol{A}$ 的**正整数幂**为

$$\overbrace{\boldsymbol{AA}\cdots\boldsymbol{A}}^{k}=\boldsymbol{A}^k\qquad\qquad k\text{ 为正整数}$$

且规定

$$\boldsymbol{A}^0=\boldsymbol{E}$$

根据该定义，就有

$$\boldsymbol{A}^k\cdot\boldsymbol{A}^l=\boldsymbol{A}^{k+l},\quad\left(\boldsymbol{A}^k\right)^l=\boldsymbol{A}^{kl}\qquad\qquad k,l\text{ 为非负整数}$$

设 x 的 m 次多项式 $f(x)$ 为

$$f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$$

则方阵 $\boldsymbol{A}$ 的多项式 $f(\boldsymbol{A})$ 定义为

$$f(A)=a_m A^m + a_{m-1}A^{m-1}+\cdots+a_1 A + a_0 E$$

称为 n 阶方阵 A 的 m 次多项式。

例 3.8　已知 $A=\begin{bmatrix}1 & 2 & 3\end{bmatrix}$，$B=\begin{bmatrix}3\\2\\1\end{bmatrix}$，求 AB，$(BA)^2$，$(BA)^n$。

解　$AB=\begin{bmatrix}1 & 2 & 3\end{bmatrix}\begin{bmatrix}3\\2\\1\end{bmatrix}=\begin{bmatrix}1\times3+2\times2+3\times1\end{bmatrix}=\begin{bmatrix}10\end{bmatrix}$ 是 1×1 矩阵。

$$BA=\begin{bmatrix}3\\2\\1\end{bmatrix}\begin{bmatrix}1 & 2 & 3\end{bmatrix}=\begin{bmatrix}3 & 6 & 9\\2 & 4 & 6\\1 & 2 & 3\end{bmatrix}$$ 是 3 阶方阵；

$$(BA)^2=(BA)(BA)=\begin{bmatrix}3 & 6 & 9\\2 & 4 & 6\\1 & 2 & 3\end{bmatrix}\begin{bmatrix}3 & 6 & 9\\2 & 4 & 6\\1 & 2 & 3\end{bmatrix}$$

我们当然可以直接计算这两个矩阵的乘积，但这是比较麻烦的。我们可以利用矩阵乘法的结合律，

$$(BA)^2=(BA)(BA)=BABA=B(AB)A$$

（注意到 AB 是一个一阶方阵，根据于前文得到的结果，可将该一阶方阵当作一个数，而不会影响最终结果）

$$=(AB)BA$$

$$=10\begin{bmatrix}3 & 6 & 9\\2 & 4 & 6\\1 & 2 & 3\end{bmatrix}=\begin{bmatrix}30 & 60 & 90\\20 & 40 & 60\\10 & 20 & 30\end{bmatrix}$$

类似地，$(BA)^n=10^{n-1}(BA)=10^{n-1}\begin{bmatrix}3 & 6 & 9\\2 & 4 & 6\\1 & 2 & 3\end{bmatrix}$

$$\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}\begin{bmatrix}\mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_n\end{bmatrix}=\begin{bmatrix}\lambda_1\mu_1 & & & \\ & \lambda_2\mu_2 & & \\ & & \ddots & \\ & & & \lambda_n\mu_n\end{bmatrix}$$

定义 3.7　**方阵的行列式**

设 $\boldsymbol{A}$ 为 n 阶方阵

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

由于矩阵是一个数表，对这 n^2 个数，我们称 n 阶行列式

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

为方阵 $\boldsymbol{A}$ 的行列式，记为 $|\boldsymbol{A}|$，或 $\det(\boldsymbol{A})$，$\det\boldsymbol{A}$。方阵的行列式具有如下性质：

1. $|k\boldsymbol{A}|=k^n|\boldsymbol{A}|$ ，n 为方阵的阶数，k 为任意一个数；

2. 设 $\boldsymbol{A},\boldsymbol{B}$ 为两个 n 阶方阵，则有 $|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|$，因此就有 $|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|=|\boldsymbol{B}||\boldsymbol{A}|=|\boldsymbol{BA}|$。

证明：（1）根据矩阵运算的定义与行列式的性质，这是显然的。

（2）构造 $2n$ 阶行列式

$$D=\begin{vmatrix} a_{11} & \cdots & a_{1n} & & & \\ \vdots & & \vdots & & & \\ a_{n1} & \cdots & a_{nn} & & & \\ -1 & & & b_{11} & \cdots & b_{1n} \\ & \ddots & & \vdots & & \vdots \\ & & -1 & b_{n1} & \cdots & b_{nn} \end{vmatrix}$$

根据行列式的相关结论，我们有

$$D=|\boldsymbol{A}||\boldsymbol{B}|$$

在 D 中，第 1 列的 b_{1j} 倍、第 2 列的 b_{2j} 倍，$\cdots$，第 n 列的 b_{nj} 倍加到第 $n+j$ 列（$j=1,2,\cdots,n$）上，将 D 中 $\boldsymbol{B}$ 位置的元素化为零，得到

$$D=\begin{vmatrix} \boldsymbol{A} & \boldsymbol{C} \\ -\boldsymbol{E} & \boldsymbol{O} \end{vmatrix}$$

其中 $\boldsymbol{C}=\left[c_{ij}\right]_{n\times n},c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}$，故

$$\boldsymbol{C}=\boldsymbol{AB}$$

将 $\begin{vmatrix} \boldsymbol{A} & \boldsymbol{C} \\ -\boldsymbol{E} & \boldsymbol{O} \end{vmatrix}$ 的第 i 行与第 $n+i$ 行（$i=1,2,\cdots,n$）互换，得到

$$D=(-1)^n\begin{vmatrix}-E & O\\ A & C\end{vmatrix}=(-1)^n|-E||C|=(-1)^n(-1)^n|C|=|C|$$

故　　$|AB|=|A||B|$

（注：以后利用分块矩阵的初等变换，此定理的证明更为简单易懂）

以上结论说明：对于 n 阶方阵 A,B，即使 $AB=BA$ 不一定成立，但其行列式却相等。

定义 3.8　矩阵的**转置**

　　设 A 为一个 $m\times n$ 矩阵，

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

将 A 的 $m\times n$ 个元素重新排列，将 A 的第一行元素排成第一列，第二行元素作为第二列，…，第 m 行元素作为第 m 列，这样得到一个 $n\times m$ 矩阵

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

我们称该矩阵为 A 的**转置矩阵**，简称为 A 的**转置**，记为 A^{T}，即

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

因此，矩阵 A 与其转置矩阵 A^{T} 的元素之间有如下的对应关系

$$[A]_{ij}=\left[A^{\mathrm{T}}\right]_{ji} \tag{3.1}$$

例 3.9　　设矩阵 $A=\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix}$，则 $A^{\mathrm{T}}=\begin{bmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{bmatrix}$。

例 3.10　　设 n 维行向量 $A=[a_1\ a_2\ \cdots\ a_n]$，则

$$\boldsymbol{A}^{\mathrm{T}}=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

因此，n 维列向量 $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ 亦写作 $\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}^{\mathrm{T}}$。

转置运算的运算律：

1、$\left(\boldsymbol{A}^{\mathrm{T}}\right)^{\mathrm{T}}=\boldsymbol{A}$

2、$\left(\boldsymbol{A}+\boldsymbol{B}\right)^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}$

3、$\left(k\boldsymbol{A}\right)^{\mathrm{T}}=k\boldsymbol{A}^{\mathrm{T}}$，$k$ 为任意一个数；

4、$\left(\boldsymbol{AB}\right)^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$

因此，$\left(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k\right)^{\mathrm{T}}=\boldsymbol{A}_k^{\mathrm{T}}\cdots\boldsymbol{A}_2^{\mathrm{T}}\boldsymbol{A}_1^{\mathrm{T}}$

5、对 n 阶方阵 $\boldsymbol{A}$，有 $\left|\boldsymbol{A}^{\mathrm{T}}\right|=\left|\boldsymbol{A}\right|$。

证明　　我们这里仅证明性质 1、4，其余证明留给读者。

（1）　容易验证 $\left(\boldsymbol{A}^{\mathrm{T}}\right)^{\mathrm{T}},\boldsymbol{A}$ 为同型矩阵，然后验证其对应的元素相等。利用结论(3.1)，有

$$\left[\left(\boldsymbol{A}^{\mathrm{T}}\right)^{\mathrm{T}}\right]_{ij}=\left[\boldsymbol{A}^{\mathrm{T}}\right]_{ji}=\left[\boldsymbol{A}\right]_{ij}$$

因此这两个矩阵 $\left(\boldsymbol{A}^{\mathrm{T}}\right)^{\mathrm{T}},\boldsymbol{A}$ 对应的元素相等；

（2）　设 $\boldsymbol{A}=\left[a_{ij}\right]_{m\times n},\boldsymbol{B}=\left[b_{ij}\right]_{m\times n}$，容易验证 $\left(\boldsymbol{AB}\right)^{\mathrm{T}},\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 是同型矩阵，则

$$\left[\left(\boldsymbol{AB}\right)^{\mathrm{T}}\right]_{ij}=\left[\boldsymbol{AB}\right]_{ji}=\sum_{t=1}^{k}\boldsymbol{a}_{jt}\boldsymbol{b}_{ti}=\sum_{t=1}^{k}\left[\boldsymbol{A}^{\mathrm{T}}\right]_{tj}\left[\boldsymbol{B}^{\mathrm{T}}\right]_{it}=\sum_{t=1}^{k}\left[\boldsymbol{B}^{\mathrm{T}}\right]_{it}\left[\boldsymbol{A}^{\mathrm{T}}\right]_{tj}=\left[\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\right]_{ij}$$

定义 3.9　　**对称矩阵、反对称矩阵**

若方阵 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$$

则称 $\boldsymbol{A}$ 为**对称矩阵**；

若方阵 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$$

则称 $\boldsymbol{A}$ 为**反对称矩阵**。

因此，方阵 $\boldsymbol{A}=\left[a_{ij}\right]_n$ 为对称矩阵的充要条件为：

$$\begin{aligned}&\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}\Leftrightarrow\left[\boldsymbol{A}^{\mathrm{T}}\right]_{ij}=\left[\boldsymbol{A}\right]_{ij},i,j=1,2,\cdots,n\\&\Leftrightarrow\left[\boldsymbol{A}^{\mathrm{T}}\right]_{ij}=\left[\boldsymbol{A}\right]_{ji}=\left[\boldsymbol{A}\right]_{ij},i,j=1,2,\cdots,n\\&\Leftrightarrow a_{ij}=a_{ji},\ \ i,j=1,2,\cdots,n\end{aligned}$$

在方阵 $\boldsymbol{A}$ 中，元素 a_{ij} 与 a_{ji} 是关于主对角线对称的，因此方阵为对称矩阵的充要条件是：关于主对角线对称的元素相等。

类似地，方阵 $\boldsymbol{A}=\left[a_{ij}\right]_n$ 为反对称矩阵的充要条件为

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

即方阵为反对称矩阵的充要条件是：关于主对角线对称的元素互为相反数。若令 $i=j$，由此可得 $a_{ii}=0,\ \ i=1,2,\cdots,n$，即反对称矩阵的主对角线元素都为 0。

例 3.11　设 n 阶实方阵 $\boldsymbol{A},\boldsymbol{B}$ 满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，$\boldsymbol{B}^{\mathrm{T}}=-\boldsymbol{B}$，且 $\boldsymbol{A}^2=\boldsymbol{B}^2$，则 $\boldsymbol{A}=\boldsymbol{B}=\boldsymbol{O}$。

证明：设 $\boldsymbol{A}=\left[a_{ij}\right]_n$，$\boldsymbol{B}=\left[b_{ij}\right]_n$。由于 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，$\boldsymbol{B}^{\mathrm{T}}=-\boldsymbol{B}$，所以

$$a_{ij}=a_{ji}$$

$$b_{ij}=-b_{ji},\ \ b_{ii}=0$$

所以

$$\begin{aligned}&[\boldsymbol{A}^2]_{11}=[\boldsymbol{A}\boldsymbol{A}]_{11}=\sum_{i=1}^n[\boldsymbol{A}]_{1i}\ [\boldsymbol{A}]_{i1}\\&=\sum_{i=1}^n a_{1i}a_{1i}=\sum_{i=1}^n a_{1i}^2\geq 0\end{aligned}$$

$$\begin{aligned}&[\boldsymbol{B}^2]_{11}=[\boldsymbol{B}\boldsymbol{B}]_{11}=\sum_{i=1}^n[\boldsymbol{B}]_{1i}\ [\boldsymbol{B}]_{i1}\\&=\sum_{i=1}^n b_{1i}\left(-b_{1i}\right)=-\sum_{i=1}^n b_{1i}^2\leq 0\end{aligned}$$

由 $[\boldsymbol{A}^2]_{11}=[\boldsymbol{B}^2]_{11}$ 以及 $\boldsymbol{A},\boldsymbol{B}$ 为实矩阵可得 $\sum_{i=1}^n a_{1i}^2=\sum_{i=1}^n b_{1i}^2=0$，故矩阵的第一行元素都为 0。

同理可证 $\boldsymbol{A},\boldsymbol{B}$ 的其它各行元素都为 0。故结论成立。

例 3.12　任意方阵都可表示为一个对称矩阵与一个反对称矩阵之和。
（请注意高等数学中的类似结论：任意一个定义域关于坐标原点对称的实函数都可表示为一个奇函数与一个偶函数之和）
分析：设 $\boldsymbol{A}$ 为方阵，且它可表示为对称矩阵 $\boldsymbol{B}$ 与反对称矩阵 $\boldsymbol{C}$ 之和，

$$\boldsymbol{A}=\boldsymbol{B}+\boldsymbol{C}$$

其中 $\boldsymbol{B}=\boldsymbol{B}^{\mathrm{T}},\boldsymbol{C}=-\boldsymbol{C}^{\mathrm{T}}$。则

$$\boldsymbol{A}^{\mathrm{T}}=\left(\boldsymbol{B}+\boldsymbol{C}\right)^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}+\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{B}-\boldsymbol{C}$$

因此可得

$$\boldsymbol{B}=\frac{\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}}{2},\quad \boldsymbol{C}=\frac{\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}}}{2}$$

只需要验证 $\boldsymbol{B}=\boldsymbol{B}^{\mathrm{T}},\boldsymbol{C}=-\boldsymbol{C}^{\mathrm{T}}$ 是否成立，而显然这是成立的。

证明　　略。

定义 3.10　　**共轭矩阵**

设矩阵 $\boldsymbol{A}=\left[a_{ij}\right]_{m\times n}$ 是复矩阵，以 $\overline{a}_{ij}$ 表示 a_{ij} 的共轭复数，则定义

$$\overline{\boldsymbol{A}}=\left[\overline{a}_{ij}\right]$$

为矩阵 $\boldsymbol{A}$ 的**共轭矩阵**。

因此矩阵 $\boldsymbol{A}$ 与其共轭矩阵 $\overline{\boldsymbol{A}}$ 的元素之间有关系

$$\left[\overline{\boldsymbol{A}}\right]_{ij}=\overline{\left[\boldsymbol{A}\right]_{ij}} \tag{3.2}$$

共轭矩阵具有如下的运算性质（$\boldsymbol{A},\boldsymbol{B}$ 为复矩阵，$\lambda\in\mathbb{C}$，且运算都是可行的）

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

（2）$\overline{\lambda\boldsymbol{A}}=\overline{\lambda}\ \overline{\boldsymbol{A}}$

（3）$\overline{\boldsymbol{AB}}=\overline{\boldsymbol{A}}\ \overline{\boldsymbol{B}}$

证明　　这里仅证明性质（3）。

设 $\boldsymbol{A}=\left[a_{ij}\right]_{m\times n},\boldsymbol{B}=\left[b_{ij}\right]_{m\times n}$，显然 $\overline{\boldsymbol{AB}},\ \overline{\boldsymbol{A}}\ \overline{\boldsymbol{B}}$ 是同型矩阵。则有

$$\left[\overline{\boldsymbol{AB}}\right]_{ij}=\overline{\sum_{t=1}^{k}a_{it}b_{tj}}=\sum_{t=1}^{k}\overline{a_{it}}\,\overline{b_{tj}}=\sum_{t=1}^{k}\left[\overline{\boldsymbol{A}}\right]_{it}\left[\overline{\boldsymbol{B}}\right]_{tj}=\left[\overline{\boldsymbol{A}}\,\overline{\boldsymbol{B}}\right]_{ij}$$

因此 $\overline{\boldsymbol{AB}}=\overline{\boldsymbol{A}}\ \overline{\boldsymbol{B}}$。

习题 3.2

1、 计算如下矩阵运算

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

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

（3）已知 $\boldsymbol{A} = \begin{bmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{bmatrix}$，求 $\boldsymbol{A}^n$。

（4）对于一般的 n 阶方阵 $\boldsymbol{A} = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$，求 $\boldsymbol{A}^k$。

（5）设 $\boldsymbol{A} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 2 & 5 & 3 \end{bmatrix}$，计算 $\boldsymbol{AB}, \boldsymbol{BA}, (\boldsymbol{AB})^k, (\boldsymbol{BA})^k$。

2、 设 $\boldsymbol{A}, \boldsymbol{B}$ 是同阶方阵，证明：$(\boldsymbol{A}+\boldsymbol{B})^2 = \boldsymbol{A}^2 + 2\boldsymbol{AB} + \boldsymbol{B}^2 \Leftrightarrow \boldsymbol{AB} = \boldsymbol{BA}$。

3、 证明：对任意矩阵 $\boldsymbol{A}$， $\boldsymbol{AA}^{\mathrm{T}}$ 与 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 都是对称矩阵。

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

5、 设 $\boldsymbol{A} = \left[a_{ij}\right]_{n\times n}$，定义 $\boldsymbol{A}$ 的迹 $tr(\boldsymbol{A})$ 为：$tr(\boldsymbol{A}) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}$。证明：

任意 $m\times n$ 矩阵 $\boldsymbol{A}$ 和任意 $n\times m$ 矩阵 $\boldsymbol{B}$，均有 $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$。

6、 设 $\boldsymbol{A}, \boldsymbol{B}$ 为任意两个 n 阶方阵，证明：$\boldsymbol{AB} - \boldsymbol{BA} \neq \boldsymbol{E}$。

§3.3　矩阵的逆

在数域 $\mathbb{F}$ 上，我们有除法的概念。这是因为 $\forall a, b \in \mathbb{F}, a \neq 0$，则存在唯一的 $x \in \mathbb{F}$，使得

$$a \times x = b$$

这样我们就可以定义除法

$$b \div a = x$$

对于上式我们有另外一种视角。根据上一段的论述，$\forall a \in \mathbb{F}, a \neq 0$，由于$1 \in \mathbb{F}$，故唯一存在$d \in \mathbb{F}$，使得

$$a \times d = 1$$

我们将该d记做$d := a^{-1} = \dfrac{1}{a}$，并称之为$a$的逆（倒数）。

因而在求解方程

$$a \times x = b$$

时，我们可在上式两边同时乘以d，得

$$d \times (a \times x) = d \times b$$

再由数域上乘法的结合律，有

$$(d \times a) \times x = d \times b$$

$$1 \times x = d \times b$$

由于 1 乘以任何数等于该数本身，因此有

$$x = d \times b = a^{-1} \times b = \frac{1}{a} \times b \triangleq \frac{b}{a}$$

换句话说，数域上的除法可以通过乘法来定义，但需要首先引入数的倒数的概念。

到目前为止，我们已经定义了矩阵乘法。那么我们是否能类似地定义其逆运算——除法呢？

对于线性方程组（1.9），可表示为

$$\boldsymbol{AX} = \boldsymbol{b}$$

因此，如果矩阵也能象实数那样做“除法”，则有解

$$“\ \boldsymbol{X} = \frac{\boldsymbol{b}}{\boldsymbol{A}}\ ”$$

我们在第一章中已经知道，不是所有线性方程组都是有解的；即使线性方程组有解，其解也未必是唯一的。因此，如果可以定义矩阵“除法”，那么矩阵“除法”的定义也不可能是无条件的。

我们在上一节已经指出，矩阵乘法中消去律不是无条件成立的，$\boldsymbol{AB} = \boldsymbol{AC} \Rightarrow \boldsymbol{B} = \boldsymbol{C}$ 不是无条件成立的。即满足方程

$$\boldsymbol{AX} = \boldsymbol{D}$$

的$\boldsymbol{X}$不一定唯一（如果该矩阵方程有解的话），因此我们就无法据此定义矩阵除法。但矩阵“除法”与消去律是等价的。也就是说，矩阵乘法若有左消去律

$$\boldsymbol{AB} = \boldsymbol{AC} \Rightarrow \boldsymbol{B} = \boldsymbol{C}$$

这相当于等式两边同时“除以”$\boldsymbol{A}$后得到的结果，也就是说，矩阵$\boldsymbol{A}$可以“除掉”；反过来，如果在$\boldsymbol{AB} = \boldsymbol{AC}$中矩阵$\boldsymbol{A}$可以“除掉”，则得到$\boldsymbol{B} = \boldsymbol{C}$，这说明此时$\boldsymbol{AB} = \boldsymbol{AC}$满足左消去律。由于上述原因，既然直接定义矩阵“除法”是不可能的。我们可通过讨论矩阵乘法的消去律来研究定义矩阵“除法”的可行性。

问题：（1）当 $\boldsymbol{A}$ 满足什么条件时，有

$$\boldsymbol{AB}=\boldsymbol{AC}\Rightarrow \boldsymbol{B}=\boldsymbol{C}$$

如果 $\boldsymbol{A}$ 满足上述要求，则称 $\boldsymbol{A}$ 满足左消去律。

（2）当 $\boldsymbol{A}$ 满足什么条件时，有

$$\boldsymbol{BA}=\boldsymbol{CA}\Rightarrow \boldsymbol{B}=\boldsymbol{C}$$

如果 $\boldsymbol{A}$ 满足上述要求，则称 $\boldsymbol{A}$ 满足右消去律。

在数的乘法中，1 乘以任何数都等于该数；而在矩阵乘法中，单位矩阵与任何矩阵相乘都等于该矩阵。因此，对于

$$\boldsymbol{AB}=\boldsymbol{AC}$$

若有矩阵 $\boldsymbol{D}$，使得

$$\boldsymbol{DA}=\boldsymbol{E} \tag{3.3}$$

则在 $\boldsymbol{AB}=\boldsymbol{AC}$ 两边同时左乘 $\boldsymbol{D}$ 得：

$$\boldsymbol{DAB}=\boldsymbol{DAC}$$

运用矩阵乘法的结合律，可得 $\left(\boldsymbol{DA}\right)\boldsymbol{B}=\left(\boldsymbol{DA}\right)\boldsymbol{C}$

$$\boldsymbol{EB}=\boldsymbol{EC} \qquad \Rightarrow \boldsymbol{B}=\boldsymbol{C}$$

因此，当(3.3)成立时，$\boldsymbol{A}$ 满足左消去律。

类似地，对于

$$\boldsymbol{BA}=\boldsymbol{CA}$$

若有矩阵 $\boldsymbol{F}$，使得

$$\boldsymbol{AF}=\boldsymbol{E} \tag{3.4}$$

则

$$\boldsymbol{BAF}=\boldsymbol{CAF}$$

$$\boldsymbol{B}\left(\boldsymbol{AF}\right)=\boldsymbol{C}\left(\boldsymbol{AF}\right)$$

$$\boldsymbol{BE}=\boldsymbol{CE}\Rightarrow \boldsymbol{B}=\boldsymbol{C}$$

此时 $\boldsymbol{A}$ 满足右消去律。

注意：我们所给出的上述条件是充分条件。

进一步的问题是：当 $\boldsymbol{A}_{m\times n}$ 满足什么条件时，它同时满足左消去律、右消去律——即满足消去律？

充分条件：存在 $\boldsymbol{D}_{n\times m},\boldsymbol{F}_{m\times n}$，使得

$$\boldsymbol{D}_{n\times m}\boldsymbol{A}_{m\times n}=\boldsymbol{E}_n, \qquad \boldsymbol{A}_{m\times n}\boldsymbol{F}_{n\times m}=\boldsymbol{E}_m$$

问题是 $m=n$？。

$m\neq n$ 的情况更为复杂，目前我们无法开展讨论，我们将把该问题留到以后解决。目前先讨论 $m=n$、即 $\boldsymbol{A}$ 为方阵的情形。

设方阵 $\boldsymbol{A}$ 满足消去律，根据以上结论，其充分条件是存在方阵 $\boldsymbol{B},\boldsymbol{C}$ 使得

$$\boldsymbol{BA}=\boldsymbol{E},\qquad \boldsymbol{AC}=\boldsymbol{E}$$

我们可以断言，此时必有

$$\boldsymbol{B}=\boldsymbol{C}$$

这是因为

$$\boldsymbol{B}=\boldsymbol{BE}=\boldsymbol{B}(\boldsymbol{AC})=(\boldsymbol{BA})\boldsymbol{C}=\boldsymbol{EC}=\boldsymbol{C}$$

这实际上是方阵可逆的概念。

定义 3.11　　**可逆矩阵、逆矩阵**

　　对于 n 阶方阵 $\boldsymbol{A}$，若存在 n 阶方阵 $\boldsymbol{B}$，使得

$$\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{E}_n$$

则称 $\boldsymbol{A}$ 为**可逆矩阵**（**非奇异矩阵**），也称 $\boldsymbol{A}$ 是可逆的，$\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的逆矩阵，简称为 $\boldsymbol{A}$ 的逆，否则称 $\boldsymbol{A}$ 是不可逆的。

注意：可逆、不可逆概念都是仅对方阵定义的。

命题　　可逆矩阵的逆矩阵的是唯一的。

证明：设 $\boldsymbol{A}$ 可逆，若 $\boldsymbol{A}$ 的逆矩阵不唯一，不妨设 $\boldsymbol{B},\boldsymbol{C}$ 均为 $\boldsymbol{A}$ 的任意两个逆矩阵，即

$$\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{E},\qquad \boldsymbol{AC}=\boldsymbol{CA}=\boldsymbol{E}$$

那么，

$$\boldsymbol{B}=\boldsymbol{EB}=(\boldsymbol{CA})\boldsymbol{B}=\boldsymbol{C}(\boldsymbol{AB})=\boldsymbol{CE}=\boldsymbol{C}$$

故

$$\boldsymbol{B}=\boldsymbol{C}$$

因此，可逆矩阵的任意两个逆矩阵相等，因而可逆矩阵的逆矩阵唯一。

基于逆矩阵的唯一性，我们把可逆矩阵 $\boldsymbol{A}$ 的逆矩阵记做 $\boldsymbol{A}^{-1}$，

　　若 $\boldsymbol{A}$ 可逆，$\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的逆，则有 $\boldsymbol{AB}=\boldsymbol{E}_n$，取行列式

$$|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|=|\boldsymbol{E}_n|=1$$

$$\Rightarrow|\boldsymbol{A}|\neq 0$$

　　因此，方阵 $\boldsymbol{A}$ 可逆 $\Rightarrow|\boldsymbol{A}|\neq 0$。

　　实际上逆命题也成立：方阵 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|\neq 0\Rightarrow\boldsymbol{A}$ 可逆。为证明该命题，我们先引入如下伴随矩阵概念。

定义 3.12　**伴随矩阵**

设 $\boldsymbol{A}$ 为 n 阶方阵，则以 $|\boldsymbol{A}|$ 的代数余子式 $\boldsymbol{A}_{ij}$ 为元素的矩阵

$$\boldsymbol{A}^* = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{n1} \\ \boldsymbol{A}_{12} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{n2} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{1n} & \boldsymbol{A}_{2n} & \cdots & \boldsymbol{A}_{nn} \end{bmatrix} \tag{3.5}$$

称为矩阵 $\boldsymbol{A}$ 的伴随矩阵。

注意有：

$$\left[\boldsymbol{A}^*\right]_{ij} = \boldsymbol{A}_{ji} \tag{3.6}$$

如下定理 3.1 给出了伴随矩阵的一个重要性质。

定理 3.1　设 $\boldsymbol{A}$ 为 n 阶方阵，$\boldsymbol{A}^*$ 为其伴随矩阵，则

$$\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A} = |\boldsymbol{A}|\boldsymbol{E}_n$$

证明：先证明 $\boldsymbol{A}\boldsymbol{A}^* = |\boldsymbol{A}|\boldsymbol{E}_n$。

根据行列式展开定理

$$\sum_{k=1}^{n} a_{ik}\boldsymbol{A}_{jk} = \begin{cases} |\boldsymbol{A}| & i = j \\ 0 & i \neq j \end{cases}$$

所以

$$\left[\boldsymbol{A}\boldsymbol{A}^*\right]_{ij} = \sum_{k=1}^{n}[\boldsymbol{A}]_{ik}\left[\boldsymbol{A}^*\right]_{kj} = \sum_{k=1}^{n} a_{ik}\boldsymbol{A}_{jk} = \begin{cases} |\boldsymbol{A}| & i = j \\ 0 & i \neq j \end{cases}$$

这说明矩阵

$$\boldsymbol{A}\boldsymbol{A}^* = \begin{bmatrix} |\boldsymbol{A}| & & & \\ & |\boldsymbol{A}| & & \\ & & \ddots & \\ & & & |\boldsymbol{A}| \end{bmatrix} = |\boldsymbol{A}|\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} = |\boldsymbol{A}|\boldsymbol{E}$$

$\boldsymbol{A}^*\boldsymbol{A} = |\boldsymbol{A}|\boldsymbol{E}_n$ 可以类似地加以证明。 ■

根据定理 3.1，容易得到如下定理。

定理 3.2　n 阶方阵 $\boldsymbol{A}$ 可逆的充要条件是 $|\boldsymbol{A}| \neq 0$。

证明：必要性已经证明。现证明充分性。当 $|\boldsymbol{A}| \neq 0$ 时，由于 $\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A} = |\boldsymbol{A}|\boldsymbol{E}_n$，同时除

以$|\boldsymbol{A}|$，得到

$$\boldsymbol{A}\left(\frac{\boldsymbol{A}^*}{|\boldsymbol{A}|}\right)=\left(\frac{\boldsymbol{A}^*}{|\boldsymbol{A}|}\right)\boldsymbol{A}=\boldsymbol{E}_n$$

根据定义，$\boldsymbol{A}$可逆，且$\dfrac{\boldsymbol{A}^*}{|\boldsymbol{A}|}$为$\boldsymbol{A}$的逆矩阵，即$\boldsymbol{A}^{-1}=\dfrac{\boldsymbol{A}^*}{|\boldsymbol{A}|}$。

考虑矩阵方程$\boldsymbol{AX}=\boldsymbol{C}$，其中$\boldsymbol{A},\boldsymbol{C}$为已知矩阵，且$\boldsymbol{A}$可逆，$\boldsymbol{X}$未知，则在

$$\boldsymbol{AX}=\boldsymbol{C}$$

两边同时左乘$\boldsymbol{A}^{-1}$，可得

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

$$\Rightarrow(\boldsymbol{A}^{-1}\boldsymbol{A})\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{C}\Rightarrow\boldsymbol{EX}=\boldsymbol{A}^{-1}\boldsymbol{C}$$

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

这可看作是在$\boldsymbol{AB}=\boldsymbol{C}$“除掉“$\boldsymbol{A}$所得的结果。

注意：　$\boldsymbol{A}^{-1}\boldsymbol{C}$不能写为$\dfrac{\boldsymbol{C}}{\boldsymbol{A}}$。这是由于仅从$\dfrac{\boldsymbol{C}}{\boldsymbol{A}}$中我们无法判断$\dfrac{\boldsymbol{C}}{\boldsymbol{A}}$究竟是$\boldsymbol{A}^{-1}\boldsymbol{C}$还是$\boldsymbol{CA}^{-1}$，而$\boldsymbol{A}^{-1}\boldsymbol{C}=\boldsymbol{CA}^{-1}$一般是不成立的（假设这两个乘法有意义），所以不能这样写。

定理 3.3　若方阵$\boldsymbol{A},\boldsymbol{B}$满足$\boldsymbol{AB}=\boldsymbol{E}$，则$\boldsymbol{A},\boldsymbol{B}$都可逆，且

$$\boldsymbol{A}^{-1}=\boldsymbol{B},\qquad \boldsymbol{B}^{-1}=\boldsymbol{A}$$

证明：由于$\boldsymbol{AB}=\boldsymbol{E}$，等式两边取行列式，得到

$$|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|=|\boldsymbol{E}|=1$$

因而$|\boldsymbol{A}|\neq 0$，根据定理 3.2 可知$\boldsymbol{A}$可逆，在$\boldsymbol{AB}=\boldsymbol{E}$两边同时乘以$\boldsymbol{A}^{-1}$得

$$\boldsymbol{A}^{-1}\boldsymbol{AB}=\boldsymbol{A}^{-1}\quad\Rightarrow\boldsymbol{EB}=\boldsymbol{A}^{-1}\Rightarrow\boldsymbol{B}=\boldsymbol{A}^{-1}$$

有关$\boldsymbol{B}$的结论类似可证。■

同理可得如下结论：

（1）$\boldsymbol{A},\boldsymbol{B}$为$n$阶方阵，则$\boldsymbol{AB}$可逆$\Leftrightarrow\boldsymbol{A},\boldsymbol{B}$都可逆。

（2）$\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$都为$n$阶方阵，则$\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_n$可逆$\Leftrightarrow\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$都可逆。

方阵逆的运算规律（其中 $\boldsymbol{A},\boldsymbol{B}$ 为可逆矩阵）

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

(2) $\left(\boldsymbol{A}^{\mathrm{T}}\right)^{-1}=\left(\boldsymbol{A}^{-1}\right)^{\mathrm{T}}$

(3) $\left(k\boldsymbol{A}\right)^{-1}=k^{-1}\boldsymbol{A}^{-1}$，其中 k 为非零常数。(4) $\left|\boldsymbol{A}^{-1}\right|=\left|\boldsymbol{A}\right|^{-1}$

(5) $\left(\boldsymbol{AB}\right)^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$，进一步有：若 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_{m-1},\boldsymbol{A}_m$ 都可逆，则 $\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{m-1}\boldsymbol{A}_m$ 可逆，且有

$$\left(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_{m-1}\boldsymbol{A}_m\right)^{-1}=\boldsymbol{A}_m^{-1}\boldsymbol{A}_{m-1}^{-1}\cdots\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}$$

证明　　应用定理 3.4 可证明性质（1)、（2)、（3)、（5)。例如性质（5)，由于 $\boldsymbol{A},\boldsymbol{B}$ 可逆，因此 $\boldsymbol{A}^{-1},\boldsymbol{B}^{-1}$ 都存在，而

$$\left(\boldsymbol{AB}\right)\left(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\right)=\boldsymbol{A}\left(\boldsymbol{BB}^{-1}\right)\boldsymbol{A}^{-1}=\boldsymbol{AEA}^{-1}=\boldsymbol{AA}^{-1}=\boldsymbol{E}$$

根据定理 3.4 可知，$\boldsymbol{AB}$ 可逆，且 $\left(\boldsymbol{AB}\right)^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$。

对于性质（4)，由 $\boldsymbol{AA}^{-1}=\boldsymbol{E}$，两边取行列式得 $\left|\boldsymbol{AA}^{-1}\right|=\left|\boldsymbol{E}\right|\Rightarrow\left|\boldsymbol{A}\right|\left|\boldsymbol{A}^{-1}\right|=1$，因此结论成立。

例 3.13　证明如下结论

(1) 设 $\boldsymbol{B}=diag\left(\lambda_1,\lambda_2,\cdots\lambda_n\right)$ 为对角矩阵，证明 $\boldsymbol{B}^k=diag\left(\lambda_1^k,\lambda_2^k,\cdots\lambda_n^k\right)\left(k\in\mathbb{N}\right)$；

(2) 若方阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{P}$ 满足 $\boldsymbol{A}=\boldsymbol{PBP}^{-1}$，证明：$\boldsymbol{A}^k=\boldsymbol{PB}^k\boldsymbol{P}^{-1}\left(k\in\mathbb{N}\right)$；

(3) 设 $f\left(x\right)$ 是关于 x 的多项式，方阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{P}$ 满足 $\boldsymbol{A}=\boldsymbol{PBP}^{-1}$，证明：$f\left(\boldsymbol{A}\right)=\boldsymbol{P}f\left(\boldsymbol{B}\right)\boldsymbol{P}^{-1}$。

证明　　（1）根据矩阵乘法的定义，应用数学归纳法可很容易地证明本结论。证明过程略。
（2）利用数学归纳法证明即可。

（3）设 $f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$，利用结论（2)，有

$$\begin{aligned}&f\left(\boldsymbol{A}\right)=a_m\boldsymbol{A}^m+a_{m-1}\boldsymbol{A}^{m-1}+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{E}\\&=a_m\boldsymbol{PB}^m\boldsymbol{P}^{-1}+a_{m-1}\boldsymbol{PB}^{m-1}\boldsymbol{P}^{-1}+\cdots+a_1\boldsymbol{PBP}^{-1}+a_0\boldsymbol{PEP}^{-1}\\&=\boldsymbol{P}\left(a_m\boldsymbol{B}^m+a_{m-1}\boldsymbol{B}^{m-1}+\cdots+a_1\boldsymbol{B}+a_0\boldsymbol{E}\right)\boldsymbol{P}^{-1}\\&=\boldsymbol{P}f\left(\boldsymbol{B}\right)\boldsymbol{P}^{-1}\end{aligned}$$

关于逆矩阵的例题

1、利用定理 3.4 证明矩阵可逆并求逆矩阵。典型题型为：

给出一个方阵 $f(\boldsymbol{A})=\boldsymbol{O}$ 的多项式，需要证明矩阵 $\boldsymbol{A}$ 的某个多项式 $\mathrm{g}(\boldsymbol{A})$ 可逆。

方法：

利用 $f(\boldsymbol{A})=\boldsymbol{O}$，通过因式分解，得到 $g(\boldsymbol{A})h(\boldsymbol{A})=\boldsymbol{E}$

根据定理 3.4，因而 $\mathrm{g}(\boldsymbol{A})$ 可逆，且 $\left[g(\boldsymbol{A})\right]^{-1}=h(\boldsymbol{A})$。

例 3.14 已知 $\boldsymbol{A}^m+a_1\boldsymbol{A}^{m-1}+\cdots+a_{m-1}\boldsymbol{A}+a_m\boldsymbol{E}=\boldsymbol{O}$，其中 $a_m\neq 0$，证明 $\boldsymbol{A}$ 可逆，并求 $\boldsymbol{A}$ 的逆。

证明：由于 $\boldsymbol{A}^m+a_1\boldsymbol{A}^{m-1}+\cdots+a_{m-1}\boldsymbol{A}+a_m\boldsymbol{E}=\boldsymbol{O}$，所以

$$\boldsymbol{A}^m+a_1\boldsymbol{A}^{m-1}+\cdots+a_{m-1}\boldsymbol{A}=-a_m\boldsymbol{E}$$

$$\Rightarrow \boldsymbol{A}\left(\boldsymbol{A}^{m-1}+a_1\boldsymbol{A}^{m-2}+\cdots+a_{m-1}\boldsymbol{E}\right)=-a_m\boldsymbol{E}$$

由于 $a_m\neq 0$，故有

$$\Rightarrow \boldsymbol{A}\left[-\frac{1}{a_m}\left(\boldsymbol{A}^{m-1}+a_1\boldsymbol{A}^{m-2}+\cdots+a_{m-1}\boldsymbol{E}\right)\right]=\boldsymbol{E}$$

根据定理 3.4 可得：$\boldsymbol{A}$ 可逆，且

$$\boldsymbol{A}^{-1}=-\frac{1}{a_m}\left(\boldsymbol{A}^{m-1}+a_1\boldsymbol{A}^{m-2}+\cdots+a_{m-1}\boldsymbol{E}\right)$$

例 3.15 已知 $\boldsymbol{A}^2-2\boldsymbol{A}-3\boldsymbol{E}=\boldsymbol{O}$，求证矩阵 $\boldsymbol{A}+3\boldsymbol{E}$ 可逆，并求其逆。

证明：$(\boldsymbol{A}+3\boldsymbol{E})(\boldsymbol{A}-5\boldsymbol{E})=\boldsymbol{A}^2-2\boldsymbol{A}-15\boldsymbol{E}$

$$=\boldsymbol{A}^2-2\boldsymbol{A}-3\boldsymbol{E}-12\boldsymbol{E}=\boldsymbol{O}-12\boldsymbol{E}=-12\boldsymbol{E}$$

所以

$$(\boldsymbol{A}+3\boldsymbol{E})\left[-\frac{1}{12}(\boldsymbol{A}-5\boldsymbol{E})\right]=\boldsymbol{E}$$

由定理 3.4 可知 $\boldsymbol{A}+3\boldsymbol{E}$ 可逆，且

$$(\boldsymbol{A}+3\boldsymbol{E})^{-1}=-\frac{1}{12}(\boldsymbol{A}-5\boldsymbol{E})$$

例 3.16 设矩阵 $\boldsymbol{A},\boldsymbol{B}$ 与 $\boldsymbol{A}+\boldsymbol{B}$ 都可逆，证明：$\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$ 也可逆。

证明　$A^{-1}+B^{-1}=A^{-1}\left(E+AB^{-1}\right)=A^{-1}\left(BB^{-1}+AB^{-1}\right)=A^{-1}\left(B+A\right)B^{-1}$

由于 A^{-1},B^{-1} 与 $A+B$ 都可逆，因此它们的乘积 $A^{-1}+B^{-1}$ 也可逆，且有

$$\left(A^{-1}+B^{-1}\right)^{-1}=\left(A^{-1}\left(B+A\right)B^{-1}\right)^{-1}=B\left(B+A\right)^{-1}A$$

本题还可以通过比较矩阵乘法、求逆与数之间的乘法、求倒数运算规律的异同来思考。矩阵乘法、数乘法运算律的比较：
（1）相同点：结合律、分配律；
（2）不同点：交换律；
（3）单位矩阵的性质与数 1 的性质类似，方阵可逆与数不为零（存在倒数）相对应。
基于以上考虑，我们将原问题中的矩阵分别用数来代替，得到如下问题：

设 x,y 及 $x+y$ 都不为零，证明 $x^{-1}+y^{-1}$ 也不为零，并求其倒数。

$$x^{-1}+y^{-1}=\frac{1}{x}+\frac{1}{y}=\frac{x+y}{xy}$$

所以

$$\left(x^{-1}+y^{-1}\right)^{-1}=\left(\frac{x+y}{xy}\right)^{-1}=\left(x+y\right)^{-1}xy$$

将以上结果“返回”到矩阵，即

$$\left(A^{-1}+B^{-1}\right)^{-1}=\left(A+B\right)^{-1}AB \qquad \text{是否成立？}$$

注意到矩阵乘法、求逆与数的乘法、求倒数运算律的异同，我们可得到结论：$\left(A^{-1}+B^{-1}\right)^{-1}$ 必然是 $\left(B+A\right)^{-1},A,B$ 这三者的乘积，但是这三者的相对次序则是需要进一步确定的。这种类比思维方式可以给我们提供一点启示。

2、求解矩阵方程

典型题型为：

对于矩阵方程 $AX=C$，其中 A,C 已知，且 A 可逆，求 X。

由于 A 可逆，所以　$X=A^{-1}C$

对于　$XA=C$，若 A 可逆，则有　$X=CA^{-1}$。

例3.17　A,B 满足关系式 $AB=2B+A$，且

$$A=\begin{bmatrix}3&0&1\\1&1&0\\0&1&4\end{bmatrix}$$

求 $\boldsymbol{B}$ 。

解：由 $\boldsymbol{AB}=2\boldsymbol{B}+\boldsymbol{A}$ 得：　$\boldsymbol{AB}-2\boldsymbol{B}=\boldsymbol{A}$

$$\Rightarrow\left(\boldsymbol{A}-2\boldsymbol{E}\right)\boldsymbol{B}=\boldsymbol{A}$$

$$\boldsymbol{A}-2\boldsymbol{E}=\begin{bmatrix}1&0&1\\1&-1&0\\0&1&2\end{bmatrix}$$

$$\left|\boldsymbol{A}-2\boldsymbol{E}\right|=\begin{vmatrix}1&0&1\\1&-1&0\\0&1&2\end{vmatrix}=\begin{vmatrix}1&0&1\\0&-1&-1\\0&1&2\end{vmatrix}=\begin{vmatrix}1&0&1\\0&-1&-1\\0&0&1\end{vmatrix}=-1\neq 0,$$

故 $\boldsymbol{A}-2\boldsymbol{E}$ 可逆，因此 $\boldsymbol{B}=(\boldsymbol{A}-2\boldsymbol{E})^{-1}\boldsymbol{A}$ 。为求 $(\boldsymbol{A}-2\boldsymbol{E})^{-1}$，先求 $\boldsymbol{A}-2\boldsymbol{E}$ 的各个代数余子式：

$$\boldsymbol{A}-2\boldsymbol{E}=\begin{bmatrix}1&0&1\\1&-1&0\\0&1&2\end{bmatrix}$$

$$\boldsymbol{A}_{11}=(-1)^{1+1}\begin{vmatrix}-1&0\\1&2\end{vmatrix}=-2,\quad \boldsymbol{A}_{12}=(-1)^{1+2}\begin{vmatrix}1&0\\0&2\end{vmatrix}=-2,$$

$$\boldsymbol{A}_{13}=(-1)^{1+3}\begin{vmatrix}1&-1\\0&1\end{vmatrix}=1,$$

$$\boldsymbol{A}_{21}=(-1)^{2+1}\begin{vmatrix}0&1\\1&2\end{vmatrix}=1,\quad \boldsymbol{A}_{22}=(-1)^{2+2}\begin{vmatrix}1&1\\0&2\end{vmatrix}=2,$$

$$\boldsymbol{A}_{23}=(-1)^{2+3}\begin{vmatrix}1&0\\0&1\end{vmatrix}=-1,$$

$$\boldsymbol{A}_{31}=(-1)^{3+1}\begin{vmatrix}0&1\\-1&0\end{vmatrix}=1,\quad \boldsymbol{A}_{32}=(-1)^{3+2}\begin{vmatrix}1&1\\1&0\end{vmatrix}=1,$$

$$\boldsymbol{A}_{33}=(-1)^{3+3}\begin{vmatrix}1&0\\1&-1\end{vmatrix}=-1,$$

$$(\boldsymbol{A}-2\boldsymbol{E})^{-1}=\frac{1}{-1}(\boldsymbol{A}-2\boldsymbol{E})^{*}=\frac{1}{-1}\begin{bmatrix}-2&1&1\\-2&2&1\\1&-1&-1\end{bmatrix}$$

$$=\begin{bmatrix}2&-1&-1\\2&-2&-1\\-1&1&1\end{bmatrix}$$

所以

$$B=(A-2E)^{-1}A=\begin{bmatrix}2&-1&-1\\2&-2&-1\\-1&1&1\end{bmatrix}\begin{bmatrix}3&0&1\\1&1&0\\0&1&4\end{bmatrix}$$

$$=\begin{bmatrix}5&-2&-2\\4&-3&-2\\-2&2&3\end{bmatrix}$$

定义 3.13　可逆方阵的幂

设方阵 A 可逆，$k\in\mathbb{N}$，定义

$$A^{-k}=\left(A^{-1}\right)^{k}$$

因此，若 A 可逆，则 $\forall k,l\in\mathbb{Z}$，都有

$$A^{k}A^{l}=A^{k+l}$$

$$\left(A^{k}\right)^{l}=A^{kl}$$

成立。

习题 3.3

1. 用余子式法求矩阵的 $A=\begin{bmatrix}0&2&-1\\1&1&2\\-1&-1&-1\end{bmatrix}$ 的逆矩阵。

2. 若矩阵 $\begin{bmatrix}a&b\\c&d\end{bmatrix}$ 可逆，求其逆矩阵。

3. 若方阵 A 的每行元素之和都为 a，证明：

（1） A^{k} 的每行元素之和为 a^{k}；

（2）若 A 可逆，则 $a\neq 0$，且 A^{-1} 的每行元素之和等于 a^{-1}。

4. 若 $A^{2}=B^{2}=E$，且 $|A|+|B|=0$，证明：$A+B$ 不可逆。

5. 已知 $A=\begin{bmatrix}1&0&0\\0&1&0\\-3&0&3\end{bmatrix}$，求 $\left(2E+A\right)^{\mathrm{T}}\left(2E-A\right)^{-1}\left(4E-A^{2}\right)$ 的行列式。

6. 求解矩阵方程 $AX+E=A^{2}+X$，其中 $A=\begin{bmatrix}1&0&0\\0&2&0\\1&6&1\end{bmatrix}$。

7. 设 $\boldsymbol{A}=\begin{bmatrix}3&0&1\\1&1&0\\0&1&4\end{bmatrix}$，$\boldsymbol{AX}=\boldsymbol{A}+2\boldsymbol{X}$，求 $\boldsymbol{X}$。

8. 已知矩阵 $\boldsymbol{A}=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}$, $\boldsymbol{B}=\begin{bmatrix}0&1&1\\1&0&1\\1&1&0\end{bmatrix}$，且矩阵 $\boldsymbol{X}$ 满足

$$\boldsymbol{AXA}+\boldsymbol{BXB}=\boldsymbol{AXB}+\boldsymbol{BXA}+\boldsymbol{E}$$

求 $\boldsymbol{X}$。

9. 设矩阵 $\boldsymbol{A}$ 的伴随矩阵 $\boldsymbol{A}^*=\begin{bmatrix}1&0&0&0\\0&1&0&0\\1&0&1&0\\0&-3&0&8\end{bmatrix}$，并且 $\boldsymbol{ABA}^{-1}=\boldsymbol{BA}^{-1}+3\boldsymbol{E}$，求 $\boldsymbol{B}$。

10. 已知方阵 $\boldsymbol{A},\boldsymbol{B}$ 满足 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{AB}$，证明：$\boldsymbol{AB}=\boldsymbol{BA}$。

11. 设 $\boldsymbol{A}=\left[a_{ij}\right]$ 为 n 阶方阵，A_{ij} 为行列式 $|\boldsymbol{A}|$ 中元素 a_{ij} 的代数余子式，且有 $A_{ij}=a_{ij}$，证明：$|\boldsymbol{A}|=\pm1$。

12. 设 $\boldsymbol{A}$ 是 $n(n>2)$ 阶方阵，证明：（1）$|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$；（2）$(\boldsymbol{A}^*)^*=|\boldsymbol{A}|^{n-2}\boldsymbol{A}$。

13. 设方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2-2\boldsymbol{A}+3\boldsymbol{E}=\boldsymbol{O}$，证明：$\boldsymbol{A}+\boldsymbol{E}$ 可逆，并求其逆。

14. 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{A}+\boldsymbol{B}$ 都可逆，证明 $\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$ 也可逆，并求其逆矩阵。

15. 已知 $\boldsymbol{A}^3=2\boldsymbol{E}$，$\boldsymbol{B}=\boldsymbol{A}^2-2\boldsymbol{A}+2\boldsymbol{E}$，证明 $\boldsymbol{B}$ 可逆，并求出其逆矩阵。

16. 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{B}$ 是 $n\times m$ 矩阵，$\boldsymbol{P}$ 为 n 阶可逆矩阵，若 $\boldsymbol{E}+\boldsymbol{APB}$ 可逆，则 $\boldsymbol{P}^{-1}+\boldsymbol{BA}$ 也可逆，并求其逆矩阵。$(\boldsymbol{P}^{-1}+\boldsymbol{BA})^{-1}=\boldsymbol{P}-\boldsymbol{PB}(\boldsymbol{E}+\boldsymbol{APB})^{-1}\boldsymbol{AP}$

17. 设 $\boldsymbol{A}$ 为可逆矩阵，$\boldsymbol{X},\boldsymbol{Y}$ 均为 $n\times1$ 矩阵，且 $\boldsymbol{Y}^T\boldsymbol{A}^{-1}\boldsymbol{X}\neq-1$，证明 $\boldsymbol{A}+\boldsymbol{XY}^{\mathrm{T}}$ 可逆，且

$$(\boldsymbol{A}+\boldsymbol{XY}^{\mathrm{T}})^{-1}=\boldsymbol{A}^{-1}-\frac{\boldsymbol{A}^{-1}\boldsymbol{XY}^{\mathrm{T}}\boldsymbol{A}^{-1}}{1+\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{X}}$$

18. 证明：如果可逆矩阵 $\boldsymbol{A}$ 的元素均为整数，则 $\boldsymbol{A}^{-1}$ 的元素均为整数的充要条件是：$|\boldsymbol{A}|=\pm1$。

19. 设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶对称矩阵，且 $\boldsymbol{A}$ 及 $\boldsymbol{AB}+\boldsymbol{E}$ 都可逆，证明：$(\boldsymbol{AB}+\boldsymbol{E})^{-1}\boldsymbol{A}$ 为对称矩阵。

20. 设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶下三角矩阵，证明：$\boldsymbol{C}=\boldsymbol{AB}$ 也是下三角矩阵；若下三角矩阵 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{-1}$ 也为下三角矩阵。

§3.4　分块矩阵

根据矩阵乘法的定义，在矩阵乘法 $\boldsymbol{C}=\boldsymbol{AB}$ 中，$\boldsymbol{C}$ 的第 i 行元素只与 $\boldsymbol{A}$ 的第 i 行元素相关，而与 $\boldsymbol{A}$ 的其它行的元素无关；$\boldsymbol{C}$ 的第 j 列元素只与 $\boldsymbol{B}$ 的第 j 列元素相关，而与 $\boldsymbol{B}$ 的其它元素无关。将上述事实推广，我们有分块矩阵的概念。

分块矩阵的思想是将一个“大”矩阵化为“小”矩阵。例如

$$\boldsymbol{A}=\begin{bmatrix} 5 & -2 & -2 & 1 & 3 \\ 4 & -3 & -2 & 0 & 2 \\ -2 & 2 & 3 & 1 & 6 \\ 1 & 4 & 2 & 3 & -2 \end{bmatrix}$$

我们将该矩阵划分为多个部分，下面给出了一种方式，

$$\boldsymbol{A}=\left[\begin{array}{cc|ccc} 5 & -2 & -2 & 1 & 3 \\ 4 & -3 & -2 & 0 & 2 \\ \hline -2 & 2 & 3 & 1 & 6 \\ 1 & 4 & 2 & 3 & -2 \end{array}\right]=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{21} \end{bmatrix}$$

其中

$$\boldsymbol{A}_{11}=\begin{bmatrix} 5 & -2 \\ 4 & -3 \end{bmatrix},\quad \boldsymbol{A}_{12}=\begin{bmatrix} -2 & 1 & 3 \\ -2 & 0 & 2 \end{bmatrix},$$

$$\boldsymbol{A}_{21}=\begin{bmatrix} -2 & 2 \\ 1 & 4 \end{bmatrix},\quad \boldsymbol{A}_{22}=\begin{bmatrix} 3 & 1 & 6 \\ 2 & 3 & -2 \end{bmatrix}$$

注意：这里 $\boldsymbol{A}_{ij}$ 不是代数余子式，而是矩阵 $\boldsymbol{A}$ 的一部分元素构成的矩阵，称为 $\boldsymbol{A}$ 的**子矩阵**（**子块**）。

定义 3.13　　分块矩阵

对于一个 $m\times n$ 矩阵 $\boldsymbol{A}$，在 $\boldsymbol{A}$ 的行之间加入 $r-1$ 条横线 $(1\le r\le m)$，在 $\boldsymbol{A}$ 的列之间加入 $s-1$ 条竖线 $(1\le s\le n)$，则 $\boldsymbol{A}$ 被分成 $r\times s$ 个小矩阵，分别记为 $\boldsymbol{A}_{ij}\,(i=1,2,\cdots,r;\ j=1,2,\cdots,s)$，则 $\boldsymbol{A}$ 可写为

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1s}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rs}\end{bmatrix}=\left[\boldsymbol{A}_{ij}\right]_{r\times s}$$

把 $\boldsymbol{A}$ 视作以 $\boldsymbol{A}_{ij}\left(i=1,2,\cdots,r;\ j=1,2,\cdots,s\right)$ 为元素的矩阵，称为**分块矩阵**，$\boldsymbol{A}_{ij}\left(i=1,2,\cdots,r;\ j=1,2,\cdots,s\right)$ 称为 $\boldsymbol{A}$ 的**子块**。

为了叙述的方便，我们称被分块的矩阵为**母矩阵**，其元素为一个数；而以子矩阵为元素的矩阵称为**分块矩阵**。例如，在上例中，$\boldsymbol{A}=\begin{bmatrix}5 & -2 & -2 & 1 & 3\\ 4 & -3 & -2 & 0 & 2\\ -2 & 2 & 3 & 1 & 6\\ 1 & 4 & 2 & 3 & -2\end{bmatrix}$ 为母矩阵，而 $\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}$ 为分块矩阵，它是 $\boldsymbol{A}$ 的一种分块形式。

显然，对于给定的矩阵，如果没有附加一些约束条件，其分块的方法是多样的。

如果分块矩阵没有相应的运算，则“分块矩阵”这个概念是没有意义的。与矩阵的运算相对应，我们考虑分块矩阵的运算：加法、数乘、乘法、转置。当然基本要求是：“分块前的运算结果与分块后的计算结果应该是一样的“，以矩阵加法为例，设有矩阵 $\boldsymbol{A},\boldsymbol{B}$，其分块后分别记为 $\boldsymbol{A}_1,\boldsymbol{B}_1$，则 $\boldsymbol{A}_1+\boldsymbol{B}_1$ 应该为 $\boldsymbol{A}+\boldsymbol{B}$ 某个分块下的结果。另外我们希望：分块矩阵的运算法则最好与通常矩阵的对应运算有类似法则。例如，矩阵加法 $\boldsymbol{A}+\boldsymbol{B}$ 定义为这两个矩阵的对应元素相加，因而对于分块矩阵相加，我们希望其运算法则为对应的子矩阵相加；矩阵乘法 $\boldsymbol{AB}$ 中，$\boldsymbol{AB}$ 的一个元素为 $\boldsymbol{A}$ 的一行元素与 $\boldsymbol{B}$ 的一列元素对应相乘、相加，因而分块矩阵乘法 $\boldsymbol{AB}$，我们希望其运算法则为 $\boldsymbol{A}$ 的一行子矩阵与 $\boldsymbol{B}$ 的一列子矩阵相乘、相加；其它运算可类似地考虑。但目前需要指出的是：这只是我们的“希望”，这种想法是否可行，需要考虑以下两个问题：

(1) 如何对矩阵分块，才能使得这样的运算规律**形式上**能够进行？

(2) 当第一步可行时，这样的运算规律是否成立？即分块矩阵的运算结果是否等于未分块矩阵的运算结果？

在本节的以下论述中，我们主要集中(1)，而验证（2）的过程则省略。并且要求（1）导致了在不同矩阵运算时，对矩阵分块所需遵循的准则。

下面讨论矩阵各种运算下相应的分块准则：

1、矩阵加法　　$\boldsymbol{A}+\boldsymbol{B}$

$$A=\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{bmatrix}=\left[A_{ij}\right]_{r\times s}$$

如何对 $\boldsymbol{B}$ 进行分块，才能使得分块矩阵 $\boldsymbol{A},\boldsymbol{B}$ 能相加？显然，$\boldsymbol{A},\boldsymbol{B}$ 的分块方式必须相同，即有

$$B=\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{r1} & B_{r2} & \cdots & B_{rs} \end{bmatrix}=\left[B_{ij}\right]_{r\times s}$$

且 $A_{ij},B_{ij}\left(i=1,2,\cdots,r;\ j=1,2,\cdots,s\right)$ 为同型矩阵。此时分块矩阵相加即为对应子矩阵相加

$$A+B=\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{bmatrix}+\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{r1} & B_{r2} & \cdots & B_{rs} \end{bmatrix}$$

$$=\begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} & \cdots & A_{1s}+B_{1s} \\ A_{21}+B_{21} & A_{22}+B_{22} & \cdots & A_{2s}+B_{2s} \\ \vdots & \vdots & & \vdots \\ A_{r1}+B_{r1} & A_{r2}+B_{r2} & \cdots & A_{rs}+B_{rs} \end{bmatrix}=\left[A_{ij}+B_{ij}\right]_{r\times s}$$

2、数乘 kA

由于此时只涉及一个矩阵，因此 $\boldsymbol{A}$ 可任意分块

$$A=\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{bmatrix}=\left[A_{ij}\right]_{r\times s}$$

运算时，k 乘以每一个子矩阵。

$$kA=\begin{bmatrix} kA_{11} & kA_{12} & \cdots & kA_{1s} \\ kA_{21} & kA_{22} & \cdots & kA_{2s} \\ \vdots & \vdots & & \vdots \\ kA_{r1} & kA_{r2} & \cdots & kA_{rs} \end{bmatrix}=\left[kA_{ij}\right]_{r\times s}$$

3、转置

这时也只涉及一个矩阵，因此可对该矩阵任意分块，

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1s}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rs}\end{bmatrix}=\left[\boldsymbol{A}_{ij}\right]_{r\times s}$$

运算时先对分块矩阵做转置，然后对每一个子矩阵做转置

$$\boldsymbol{A}^{\mathrm{T}}=\begin{bmatrix}\boldsymbol{A}_{11}^{\mathrm{T}} & \boldsymbol{A}_{21}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{r1}^{\mathrm{T}}\\ \boldsymbol{A}_{12}^{\mathrm{T}} & \boldsymbol{A}_{22}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{r2}^{\mathrm{T}}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{1s}^{\mathrm{T}} & \boldsymbol{A}_{2s}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{rs}^{\mathrm{T}}\end{bmatrix}$$

所以有 $\left[\boldsymbol{A}^{\mathrm{T}}\right]_{ij}=\left(\left[\boldsymbol{A}\right]_{ji}\right)^{\mathrm{T}}=\boldsymbol{A}_{ji}^{\mathrm{T}}$

4、矩阵乘法　　$\boldsymbol{AB}$

根据分块矩阵运算的原则，母矩阵 $\boldsymbol{A},\boldsymbol{B}$ 分块后， 分块矩阵 $\boldsymbol{A}$ 的列数必须等于分块矩阵 $\boldsymbol{B}$ 的行数，即若

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1s}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2s}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rs}\end{bmatrix}=\left[\boldsymbol{A}_{ij}\right]_{r\times s}$$

则分块矩阵 $\boldsymbol{B}$ 为

$$\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1k}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2k}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{B}_{s1} & \boldsymbol{B}_{s2} & \cdots & \boldsymbol{B}_{sk}\end{bmatrix}=\left[\boldsymbol{B}_{ij}\right]_{s\times k}$$

要使得

$$\boldsymbol{C}_{ij}=\sum_{t=1}^{s}\boldsymbol{A}_{it}\boldsymbol{B}_{tj}$$

有意义，则 $\boldsymbol{A}_{it}$ 的列数必须等于 $\boldsymbol{B}_{tj}$ 行数，因此，对 $\boldsymbol{A}$ 的列分块必须与 $\boldsymbol{B}$ 的行分块相同，而对 $\boldsymbol{A}$ 的行分块方式、$\boldsymbol{B}$ 的列分块方式没有要求。

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

假设对 $\boldsymbol{A}$ 作如下分块

$$A=\left[\begin{array}{cc:ccc} 5 & -2 & -2 & 1 & 3 \\ 4 & -3 & -2 & 0 & 2 \\ \hdashline -2 & 2 & 3 & 1 & 6 \\ 1 & 4 & 2 & 3 & -2 \end{array}\right]=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

那么 B 的分块方式则不唯一，而只有 B 的列分块方式与 A 的行分块方式相同即可，例如分块方式

$$B=\left[\begin{array}{c:cc} 5 & -2 & -2 \\ 4 & -3 & -2 \\ \hdashline -2 & 2 & 3 \\ 1 & 4 & 2 \\ 2 & 1 & 4 \end{array}\right], \qquad B=\left[\begin{array}{ccc} 5 & -2 & -2 \\ 4 & -3 & -2 \\ \hdashline -2 & 2 & 3 \\ 1 & 4 & 2 \\ 2 & 1 & 4 \end{array}\right]$$

都是可接受的。

对上述四种运算，分块矩阵的运算规则可总结如下：

对于涉及到两个矩阵的运算（矩阵加法、乘法），首先，将分块矩阵 A,B 各个子矩阵当成数，运用矩阵的运算法则进行运算，子矩阵之间的运算再次运用矩阵的运算法则。因此，这相当于嵌套运用矩阵的运算法则。对于只涉及一个矩阵的运算有类似的规律。

5、求逆矩阵

分块矩阵求逆没有普遍的规律，但是对于以下情形则比较简单。

定义 3.14　分块对角矩阵（准对角矩阵）

称方阵

$$A=\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{bmatrix}$$

其中 A_i 都为方阵

为分块对角矩阵（准对角矩阵）。

由于此时有 $\det A=|A_1||A_2|\cdots|A_s|$，因此有

$$A \text{ 可逆} \Leftrightarrow \text{所有 } A_i\,(i=1,2,\cdots,s) \text{ 都可逆。}$$

设准对角矩阵 A 的逆矩阵为 B

$$B=\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{ss} \end{bmatrix}$$

其中 $B_{11},B_{22},\cdots,B_{ss}$ 分别为与 $A_1,A_2,\cdots,A_s$ 同阶的方阵，则

$$AB=\begin{bmatrix} A_1B_{11} & A_1B_{12} & \cdots & A_1B_{1s} \\ A_2B_{21} & A_2B_{22} & \cdots & A_2B_{2s} \\ \vdots & \vdots & & \vdots \\ A_sB_{s1} & A_sB_{s2} & \cdots & A_sB_{ss} \end{bmatrix}=\begin{bmatrix} E_1 & & & \\ & E_2 & & \\ & & \ddots & \\ & & & E_n \end{bmatrix}$$

由于 $A_1, A_2, \cdots, A_s$ 都可逆，所以有

$$A^{-1}=\begin{bmatrix} A_1^{-1} & & & \\ & A_2^{-1} & & \\ & & \ddots & \\ & & & A_s^{-1} \end{bmatrix}$$

因此准对角矩阵的求逆只需要对每个主对角线上的子方阵求逆矩阵即可。

分块矩阵的应用

分块矩阵具有非常重要的应用，其应用主要体现在两个方面：

（1）　简化计算；

（2）　提供观察问题的新视角，抓住问题的本质；

但是对于一个具体的问题，矩阵是否应该分块？以何种思想指导分块？应该指出的是：决定对矩阵如何分块的出发点是矩阵乘法，只有根据矩阵乘法的情况进行适当分块，才可能到达矩阵分块的目的。分块矩阵除乘法之外的运算是辅助性的，是为了配套分块矩阵乘法这一运算而导出的。

我们首先给出一个将分块矩阵用于简化运算的例子。

例 3.18　　设有矩阵

$$A=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix},\qquad B=\begin{bmatrix} 1 & 0 & 1 & 0 \\ -1 & 2 & 0 & 1 \\ 1 & 0 & 4 & 1 \\ -1 & -1 & 2 & 0 \end{bmatrix}$$

计算 AB。

解：直接用矩阵乘法的定义计算，需要 64 次乘法、48 次加法。如果我们将这两个矩阵作如下分块：

$$A=\left[\begin{array}{cc:cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hdashline -1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right]=\begin{bmatrix} E & O \\ A_1 & E \end{bmatrix}\qquad A_1=\begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}$$

$$B=\left[\begin{array}{cc:cc}1 & 0 & 1 & 0\\ -1 & 2 & 0 & 1\\ \hdashline 1 & 0 & 4 & 1\\ -1 & -1 & 2 & 0\end{array}\right]=\begin{bmatrix}B_{11} & E\\ B_{21} & B_{22}\end{bmatrix}$$

$$B_{11}=\begin{bmatrix}1 & 0\\ -1 & 2\end{bmatrix},\quad B_{21}=\begin{bmatrix}1 & 0\\ -1 & -1\end{bmatrix},\quad B_{22}=\begin{bmatrix}4 & 1\\ 2 & 0\end{bmatrix}$$

根据分块矩阵乘法的运算规则，有

$$AB=\begin{bmatrix}E & O\\ A_1 & E\end{bmatrix}\begin{bmatrix}B_{11} & E\\ B_{21} & B_{22}\end{bmatrix}=\begin{bmatrix}EB_{11}+OB_{21} & E\times E+OB_{22}\\ A_1B_{11}+B_{21} & A_1E+IB_{21}\end{bmatrix}$$
$$=\begin{bmatrix}B_{11} & E\\ A_1B_{11}+B_{21} & A_1+B_{21}\end{bmatrix}$$

从上式可知，只需要做两个二阶方阵的乘法。因此根据上式计算 AB，只需要做四次乘法、6 次加法，比直接计算 AB 所需要的运算量大大减少。

注意：用分块矩阵简化矩阵运算，只有在矩阵乘法中才有可能达到，并且需要矩阵经过适当分块后有较多的 0 块、单位矩阵子块。

对于某些问题而言，分块矩阵可为解决该问题提供新的视角。我们主要考虑下面三个例子。

1、线性方程组的分块

设有线性方程组

$$A_{m\times n}X_{n\times 1}=b_{m\times 1}$$

其中 A 为线性方程组的系数矩阵，X 为未知数所构成的列矩阵，b 为方程组右边的常数所构成的列矩阵。我们可采取如下分块方法，

$$A=\left[\begin{array}{c:c:c:c}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{array}\right]=\begin{bmatrix}A_1 & A_2 & \cdots & A_n\end{bmatrix}$$

则 $AX=b$ 可写作

$$x_1A_1+x_2A_2+\cdots+x_nA_n=b$$

也就是

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

该式称为线性方程组的**向量表示**。

2、矩阵方程

$\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times k}=\boldsymbol{B}_{m\times k}$ ，$\boldsymbol{A},\boldsymbol{B}$ 已知求 $\boldsymbol{X}$ 。问：该矩阵方程是否有解，有解时如何求解？

注意到若 $m=n$，即 $\boldsymbol{A}$ 为方阵，且 $\boldsymbol{A}$ 可逆时，有 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$ 。但现在并不满足该条件，我们可用矩阵分块方法将该问题进行转化为已知类型的问题。

$\boldsymbol{A}$ 不用分块，而把 $\boldsymbol{X},\boldsymbol{B}$ 的每一列作为一个子矩阵，即

$$\boldsymbol{X}=\begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_k\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & \boldsymbol{B}_k\end{bmatrix}$$

则

$$\begin{aligned}\boldsymbol{AX}&=\boldsymbol{A}\begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_k\end{bmatrix}\\&=\begin{bmatrix}\boldsymbol{AX}_1 & \boldsymbol{AX}_2 & \cdots & \boldsymbol{AX}_k\end{bmatrix}\\&=\begin{bmatrix}\boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & \boldsymbol{B}_k\end{bmatrix}\end{aligned}$$

根据矩阵相等的定义，有

$$\boldsymbol{AX}_i=\boldsymbol{B}_i,\qquad i=1,2,\cdots,k$$

其中 $\boldsymbol{A}$ 为矩阵，$\boldsymbol{X}_i,\boldsymbol{B}_i$ 为列向量，且 $\boldsymbol{A}$ 、$\boldsymbol{B}_i$ 已知，$\boldsymbol{X}_i$ 未知，因此 $\boldsymbol{AX}_i=\boldsymbol{B}_i$ 可视作线性方程组。因此，即使 $\boldsymbol{A}$ 不可逆，我们也可以求 $\boldsymbol{X}$——它的第 i 列 $\boldsymbol{X}_i$ 是线性方程组

$$\boldsymbol{AX}_i=\boldsymbol{B}_i$$

的解。因此，我们把求解矩阵方程转化为了求解一系列的线性方程组的问题。

3、对于等式 $\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times k}=\boldsymbol{B}_{m\times k}$ ，如果 $\boldsymbol{A},\boldsymbol{B}$ 分别分块为

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n\end{bmatrix}$$

$$\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & \boldsymbol{B}_k\end{bmatrix}$$

$$\begin{aligned}\boldsymbol{AX}&=\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n\end{bmatrix}\begin{bmatrix}x_{11} & x_{12} & \cdots & x_{1k}\\ x_{21} & x_{22} & \cdots & x_{2k}\\ \vdots & \vdots & & \vdots\\ x_{n1} & x_{n2} & \cdots & x_{nk}\end{bmatrix}\\&=\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & \boldsymbol{B}_k\end{bmatrix}\end{aligned}$$

则得到

$$\boldsymbol{B}_i=\boldsymbol{A}_1x_{1i}+\boldsymbol{A}_2x_{2i}+\cdots+\boldsymbol{A}_nx_{ni}\ ,\qquad i=1,\cdots,k$$

4、对于 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times k}$ 而如果 $\boldsymbol{A},\boldsymbol{B}$ 分别分块为

$$A=\begin{bmatrix}A_1\\A_2\\\vdots\\A_m\end{bmatrix},\quad B=\begin{bmatrix}B_1 & B_2 & \cdots & B_k\end{bmatrix}$$

则

$$AB=\begin{bmatrix}A_1\\A_2\\\vdots\\A_m\end{bmatrix}\begin{bmatrix}B_1 & B_2 & \cdots & B_k\end{bmatrix}=\begin{bmatrix}A_1B_1 & A_1B_2 & \cdots & A_1B_k\\A_2B_1 & A_2B_2 & \cdots & A_2B_k\\\vdots & \vdots & & \vdots\\A_mB_1 & A_mB_2 & \cdots & A_mB_k\end{bmatrix}$$

一些特殊的结果

设 e_i 为第 i 个分量为 1、其余分量为 0 的 n 维列向量。$A_{m\times n}$ 为 $m\times n$ 矩阵，将 A 按列分块 $A=\begin{bmatrix}A_1 & A_2 & \cdots & A_n\end{bmatrix}$，则

$$Ae_i=A_i$$

若 B 为 $n\times m$ 矩阵，将 B 按行分块 $B=\begin{bmatrix}B_1\\B_2\\\vdots\\B_n\end{bmatrix}$，则

$$e_i^{\mathrm{T}}B=B_i$$

问题：为何分块时只能在矩阵的行、列之间画直线，而不画折线？例如，下述分块可不可以？

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

为什么？

习题 3.4

1.　请确实理解各种分块矩阵运算的分块原则与运算规则。

§3.5　矩阵的初等变换与初等矩阵

在线性方程组的高斯消元法中，第一步是将线性方程组化为阶梯形线性方程组。由于每个线性方程组都可用它的增广矩阵来描述，因此将线性方程组化为阶梯形线性方程组的过

程就可以用矩阵来描述。我们回到第一章的例子。

例 3.19　求解线性方程组 $\begin{cases} 2x_1 - 3x_2 + \ \ x_3 = -5 \\ x_1 - 2x_2 - \ \ x_3 = -2 \\ x_1 - \ \ x_2 + 4x_3 = -11 \end{cases}$。

下面左边给出了高斯消元法各个步骤中得到的线性方程组，而每两个线性方程组之间的箭头旁的注释说明了将前一个线性方程组化为后一个线性方程组所用的线性方程组的初等变换；而右边为相应方程组的增广矩阵。根据左边相邻线性方程组之间的关系，我们可得到其相邻的两个增广矩阵之间的关系。

$$\begin{cases} 2x_1 - \ \ x_2 + 7x_3 = -7 & ① \\ x_1 - 2x_2 - \ \ x_3 = -2 & ② \\ x_1 - \ \ x_2 + 4x_3 = -11 & ③ \end{cases} \qquad \begin{bmatrix} 2 & -1 & 7 & -7 \\ 1 & -2 & -1 & -2 \\ 1 & -1 & 4 & -11 \end{bmatrix}$$

$$\Big\downarrow ①\leftrightarrow② \qquad\qquad \Big\downarrow r_1 \leftrightarrow r_2$$

$$\begin{cases} x_1 - 2x_2 - \ \ x_3 = -2 & ① \\ 2x_1 - \ \ x_2 + 7x_3 = -7 & ② \\ x_1 - \ \ x_2 + 4x_3 = -11 & ③ \end{cases} \qquad \begin{bmatrix} 1 & -2 & -1 & -2 \\ 2 & -1 & 7 & -7 \\ 1 & -1 & 4 & -11 \end{bmatrix}$$

$$\Big\downarrow \begin{matrix} ②-①\times 2 \\ ③-① \end{matrix} \qquad\qquad \Big\downarrow \begin{matrix} r_2 - 2r_1 \\ r_3 - r_1 \end{matrix}$$

$$\begin{cases} x_1 - 2x_2 - \ \ x_3 = -2 & ① \\ \qquad 3x_2 + 9x_3 = -3 & ② \\ \qquad\ \ x_2 + 5x_3 = -9 & ③ \end{cases} \qquad \begin{bmatrix} 1 & -2 & -1 & -2 \\ 0 & 3 & 9 & -3 \\ 0 & 1 & 5 & -9 \end{bmatrix}$$

$$\Big\downarrow ②\div 3 \qquad\qquad \Big\downarrow \frac{1}{3} r_3$$

$$\begin{cases} x_1 - 2x_2 - \ \ x_3 = -2 & ① \\ \qquad\ \ x_2 + 3x_3 = -1 & ② \\ \qquad\ \ x_2 + 5x_3 = -9 & ③ \end{cases} \qquad \begin{bmatrix} 1 & -2 & -1 & -2 \\ 0 & 1 & 3 & -1 \\ 0 & 1 & 5 & -9 \end{bmatrix}$$

$$\Big\downarrow ③-② \qquad\qquad \Big\downarrow r_3 - r_2$$

$$\begin{cases} x_1 - 2x_2 - \quad x_3 = -2 \\ \qquad\ \ x_2 + 3x_3 = -1 \\ \qquad\qquad\ \ 2x_3 = -8 \end{cases} \qquad \begin{bmatrix} 1 & -2 & -1 & -2 \\ 0 & 1 & 3 & -1 \\ 0 & 0 & 2 & -8 \end{bmatrix}$$

符号说明：

1. $r_i \leftrightarrow r_j$：互换矩阵的第 i 行、j 行；

2. kr_i：将第 i 行的所有数同时乘以一个非零常数 k；

   简称：一行乘以一个非零常数 k

3. $r_j + kr_i$：将第 i 行的所有数同时乘以 k、再加到 j 行相应的数上；

   简称：将 i 行的 k 倍加到 j 行

定义 3.15　**矩阵的初等行变换**。

我们称上述三种运算为矩阵的初等行变换。

因此，高斯消元法是将线性方程组的增广矩阵用初等行变换化为阶梯形线性方程组的增广矩阵。而阶梯形线性方程组的增广矩阵有何特点？

定义 3.16　**行阶梯形矩阵**

如果矩阵 $\boldsymbol{A}$ 满足

(1)若有零行，则零行都在矩阵的最下方；

(2)从第一行起，每一行的第一个非零元前面零的个数逐行增加；

则称 $\boldsymbol{A}$ 为行阶梯形矩阵。

下面为两个行阶梯形矩阵的例子，

$$\begin{bmatrix} 1 & 3 & 2 & 8 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 1 & 0 & 5 & 2 \\ 0 & 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

图 3.1　行阶梯形矩阵的例子

显然，阶梯形线性方程组的增广矩阵为行阶梯形矩阵。高斯消元法求解线性方程组就是将线性方程组的增广矩阵用初等行变换化为行阶梯形矩阵。

定义 3.17　**行标准形矩阵（行最简形矩阵）**

如果矩阵 $\boldsymbol{A}$ 满足

(1)　$\boldsymbol{A}$ 为一个行阶梯形矩阵；

(2)每行的第一个非零元等于 1；

(3)每行的第一个非零元所在列的其它元素为 0；

则称 $\boldsymbol{A}$ 为**行最简形矩阵(行标准形矩阵)**。例如，下面为两个行标准形矩阵，

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & -1 & -6 \\ 0 & 1 & 0 & 2 & 8 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

图 3.2　行最简形矩阵的例子

用高斯－约当法求解线性方程组时，最后所得到线性方程组的增广矩阵为**行最简形矩阵**。

我们需要掌握如何用初等行变换将矩阵化为行阶梯形、行标准形。

用初等行变换将矩阵化为行阶梯形矩阵的方法：
(1)若矩阵的第一列元素都为 0，则第一列不需要再加以处理；否则，在第一列中选一个非零元素，然后通过交换该元素所在行与第一行，将该元素所在的行交换到第一行，得到一个新矩阵，位于该矩阵的第一行、第一列的元素不为 0，此时通过将第一行的某个倍数加到后面各行，就可将第一列的其它元素都化为 0；如下图所示：（其中 $\boldsymbol{B}$ 也是一个矩阵）

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \to \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{m2} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \alpha \\ 0 & \boldsymbol{B} \end{bmatrix} \text{（假设 } a_{11}\neq 0 \text{）}$$

(2)对矩阵 $\boldsymbol{B}$ 运用(1)；
(3)重复(1)、(2)即可将矩阵化为行阶梯形。

用初等行变换将矩阵化为行最简形矩阵
（1）　首先用初等行变换将矩阵化为行阶梯形矩阵；
（2）　然后从最后一个非零行开始，一直到第二个非零行为止，用初等行变换将与该行第一个非零元所在列的其它数化为零；
（3）　然后每一个非零行除以该行的第一个非零元；
（注意：其中第二步、第三步可调换次序）

下面举一个将矩阵化为行阶梯形、行最简形的例子。

例 3.20　用初等行变换将矩阵 $\boldsymbol{A}=\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 1 & -3 \\ 0 & 1 & 3 & 2 & 5 \\ 5 & 4 & 3 & 3 & -1 \end{bmatrix}$ 化为行阶梯形、行最简形。

解
$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 1 & -3 \\ 0 & 1 & 3 & 2 & 5 \\ 5 & 4 & 3 & 3 & -1 \end{bmatrix} \xrightarrow[r_4-5r_1]{r_2-3r_1} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -2 & -6 \\ 0 & 1 & 3 & 2 & 5 \\ 0 & -1 & -2 & -2 & -6 \end{bmatrix}$$

$$\xrightarrow[r_4-r_2]{r_3+r_2}\begin{bmatrix}1&1&1&1&1\\0&-1&-2&-2&-6\\0&0&1&0&-1\\0&0&0&0&0\end{bmatrix}$$

此时已经把该矩阵化为行阶梯形矩阵。下一步再用初等行变换将其化为行最简形矩阵：

$$\begin{bmatrix}1&1&1&1&1\\0&-1&-2&-2&-6\\0&0&1&0&-1\\0&0&0&0&0\end{bmatrix}\xrightarrow[r_1-r_3]{r_2+2r_3}\begin{bmatrix}1&1&0&1&2\\0&-1&0&-2&-8\\0&0&1&0&-1\\0&0&0&0&0\end{bmatrix}$$

$$\xrightarrow{r_1+r_2}\begin{bmatrix}1&0&0&-1&-6\\0&-1&0&-2&-8\\0&0&1&0&-1\\0&0&0&0&0\end{bmatrix}\xrightarrow{-r_2}\begin{bmatrix}1&0&0&-1&-6\\0&1&0&2&8\\0&0&1&0&-1\\0&0&0&0&0\end{bmatrix}$$

类似于矩阵的初等行变换，我们引入矩阵的初等列变换。

定义 3.18　　矩阵的**初等列变换**

我们把矩阵的下述三种运算统称为矩阵的初等列变换：

1. 对换变换：互换矩阵的第 i 列、 j 列的位置，记为 $c_i \leftrightarrow c_j$；

2. 数乘变换：将一列所有数同时乘以一个非零常数 k，简称：一列乘以一个非零常数 k；记为 kc_i；

3. 倍加变换：将第 i 列的所有数同时乘以 k、再加到 j 列相应的数上，简称为：将一列的 k 倍加到另一列；记为 $c_j + kc_i$；

矩阵的初等行变换、初等列变换统称为矩阵的**初等变换**。若矩阵 $\boldsymbol{A}$ 可经过有限次初等行（列）变换化为矩阵 $\boldsymbol{B}$，则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 行（列）**等价**；若矩阵 $\boldsymbol{A}$ 可经过有限次初等变换化为矩阵 $\boldsymbol{B}$，则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价，记作 $\boldsymbol{A} \cong \boldsymbol{B}$。矩阵等价也满足三条性质：

（1）　　反身性： $\boldsymbol{A} \cong \boldsymbol{A}$；

（2）　　对称性：若 $\boldsymbol{A} \cong \boldsymbol{B}$，则 $\boldsymbol{B} \cong \boldsymbol{A}$；

（3）　　传递性：若 $\boldsymbol{A} \cong \boldsymbol{B}$， $\boldsymbol{B} \cong \boldsymbol{C}$，则 $\boldsymbol{A} \cong \boldsymbol{C}$。

在将矩阵化为行阶梯形、行标准形的过程中，只要应用矩阵的初等行变换即可，在这个变换过程中所得的任何两个矩阵都是不相等的。例如， $\boldsymbol{A}$ 经过一个初等变换化为矩阵 $\boldsymbol{B}$，这两个矩阵之间不是相等关系，但如果能够建立关于这两者的等式，则可能比较有意义，这一目的可通过使用**初等矩阵**达到。

定义 3.19　　**初等矩阵**

单位矩阵 $\boldsymbol{E}$ 经过一次初等变换得到的矩阵称为**初等矩阵**。

定义 3.20　　**行(列）初等矩阵**

单位矩阵 $\boldsymbol{E}$ 经过一次初等行（列）变换得到的矩阵称为**行(列）初等矩阵**。初等行矩阵、初等列矩阵都是初等矩阵。

与矩阵的三种初等行变换相对应，有三种行初等矩阵：

(1)互换单位矩阵 $\boldsymbol{E}$ 的两行，互换第 i 行、第 j 行得到

$$R_{ij}=\begin{bmatrix}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 0 & & & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & & & 0 & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{bmatrix}\begin{matrix} \\ \\ \\ \longleftarrow \textbf{第 } i \textbf{ 行} \\ \\ \\ \\ \longleftarrow \textbf{第 } j \textbf{ 行} \\ \\ \\ \\ \end{matrix}$$

$$=\begin{bmatrix}
\boldsymbol{e}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{e}_{i-1}^{\mathrm{T}} \\ \boldsymbol{e}_j^{\mathrm{T}} \\ \boldsymbol{e}_{i+1}^{\mathrm{T}} \\ \vdots \\ \boldsymbol{e}_{j-1}^{\mathrm{T}} \\ \boldsymbol{e}_i^{\mathrm{T}} \\ \boldsymbol{e}_{j+1}^{\mathrm{T}} \\ \vdots \\ \boldsymbol{e}_n^{\mathrm{T}}
\end{bmatrix}$$

(2)单位矩阵 $\boldsymbol{E}$ 的某一行乘以一个非零常数 k，第 i 行乘以非零常数 k 得到

$$R_{i(k)}=\begin{array}{c}\begin{bmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&k&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{bmatrix}\end{array}\begin{array}{l}\\\\\\\leftarrow \text{第 } i \text{ 行}\\\\\\\\\end{array}$$

$$=\begin{bmatrix}\boldsymbol{e}_1^{\mathrm{T}}\\\vdots\\\boldsymbol{e}_{i-1}^{\mathrm{T}}\\k\boldsymbol{e}_i^{T}\\\boldsymbol{e}_{i+1}^{\mathrm{T}}\\\vdots\\\boldsymbol{e}_n^{\mathrm{T}}\end{bmatrix}$$

(3) 将 $\boldsymbol{E}$ 一行的某个倍数加到另外一行。将第 j 行的 k 倍加到第 i 行，得到

$$R_{(i,j(k))}=\begin{bmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&k&&\\&&&\ddots&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{bmatrix}\begin{array}{l}\\\\\leftarrow \text{第 } i \text{ 行}\\\\\leftarrow \text{第 } j \text{ 行}\\\\\\\end{array}$$

$$=\begin{bmatrix}\boldsymbol{e}_1^{\mathrm{T}}\\\vdots\\\boldsymbol{e}_{i-1}^{\mathrm{T}}\\\boldsymbol{e}_i^{\mathrm{T}}+k\boldsymbol{e}_j^{\mathrm{T}}\\\boldsymbol{e}_{i+1}^{\mathrm{T}}\\\vdots\\\boldsymbol{e}_{j-1}^{\mathrm{T}}\\\boldsymbol{e}_j^{\mathrm{T}}\\\boldsymbol{e}_{j+1}^{\mathrm{T}}\\\vdots\\\boldsymbol{e}_n^{\mathrm{T}}\end{bmatrix}$$

类似地，三种列初等矩阵为

(1) 互换单位矩阵 $\boldsymbol{E}$ 的两列，互换第 i 列、第 j 列得到

$$
\begin{array}{cc}
C_{ij}= & \begin{bmatrix}
1 & & & & & & & & & \\
& \ddots & & & & & & & & \\
& & 1 & & & & & & & \\
& & & 0 & & & & 1 & & \\
& & & & 1 & & & & & \\
& & & & & \ddots & & & & \\
& & & & & & 1 & & & \\
& & & 1 & & & & 0 & & \\
& & & & & & & & 1 & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{bmatrix} \\
& \begin{array}{ccc} \uparrow & \qquad\qquad & \uparrow \\ \textbf{第 } i \textbf{ 列} & & \textbf{第 } j \textbf{ 列} \end{array}
\end{array}
$$

$$=\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{i-1} & \boldsymbol{e}_j & \boldsymbol{e}_{i+1} & \cdots & \boldsymbol{e}_{j-1} & \boldsymbol{e}_i & \boldsymbol{e}_{j+1} & \cdots & \boldsymbol{e}_n\end{bmatrix}$$

(2) 单位矩阵 $\boldsymbol{E}$ 的某一列乘以一个非零常数 k。第 i 列乘以非零常数 k，得到

$$
\begin{array}{cc}
C_{i(k)}= & \begin{bmatrix}
1 & & & & & \\
& \ddots & & & & \\
& & 1 & & & \\
& & & k & & & \\
& & & & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{bmatrix} \\
& \begin{array}{c} \uparrow \\ \textbf{第 } i \textbf{ 列} \end{array}
\end{array}
$$

$$=\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{i-1} & k\boldsymbol{e}_i & \boldsymbol{e}_{i+1} & \cdots & \boldsymbol{e}_n\end{bmatrix}$$

(3) 将 $\boldsymbol{E}$ 一列的某个倍数加到另外一列，例如，第 j 列的 k 倍加到第 i 列，得到

$$C_{(i,j(k))}=\begin{array}{c}\left[\begin{array}{ccccccc}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&\ddots&&&\\&&k&&1&&\\&&&&&\ddots&\\&&&&&&1\end{array}\right]\\ \begin{array}{ccccccc}&&\uparrow&&\uparrow&&\\ &&\text{第 }i\text{ 列}&&\text{第 }j\text{ 列}&&\end{array}\end{array}$$

$$=\begin{bmatrix}\boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{i-1} & \boldsymbol{e}_i+k\boldsymbol{e}_j & \boldsymbol{e}_{i+1} & \cdots & \boldsymbol{e}_{j-1} & \boldsymbol{e}_j & \boldsymbol{e}_{j+1} & \cdots & \boldsymbol{e}_n\end{bmatrix}$$

显然，对应的行初等矩阵、列初等矩阵之间有如下关系：

$$\boldsymbol{R}_{ij}=\boldsymbol{C}_{ij}\qquad \boldsymbol{R}_{i(k)}=\boldsymbol{C}_{i(k)}\qquad \boldsymbol{R}^{\mathrm{T}}_{(i,j(k))}=\boldsymbol{C}_{(i,j(k))}=\boldsymbol{R}_{(j,i(k))}$$

因而，一个初等矩阵即可看成是行初等矩阵，也可看成列初等矩阵，但是需要注意从单位矩阵变到该行（列）初等矩阵所做的初等变换。

初等矩阵的行列式不等于零，因而它们都是可逆的。其逆在稍后给出。

初等矩阵的作用

若矩阵 $\boldsymbol{A}$ 经过一次初等变换化为矩阵 $\boldsymbol{B}$，则可用初等矩阵建立关于 $\boldsymbol{A},\boldsymbol{B}$ 的方程。这反映为如下定理：

定理 3.5　　对矩阵 $\boldsymbol{A}$ 施行某种初等行(列)变换，相当于用相应的行(列)初等矩阵左(右)乘矩阵 $\boldsymbol{A}$。

定理内容说明：

(1)互换 $\boldsymbol{A}$ 的第 i 行、j 行得到 $\boldsymbol{B}$，则 $\boldsymbol{R}_{ij}\boldsymbol{A}=\boldsymbol{B}$；

(2) $\boldsymbol{A}$ 的第 i 行乘以非零常数 k 得到 $\boldsymbol{B}$，则 $\boldsymbol{R}_{i(k)}\boldsymbol{A}=\boldsymbol{B}$；

(3) 将 $\boldsymbol{A}$ 的第 j 行的 k 倍加到第 i 行得到 $\boldsymbol{B}$，则 $\boldsymbol{R}_{(i,j(k))}\boldsymbol{A}=\boldsymbol{B}$；

(4)互换 $\boldsymbol{A}$ 的第 i 列、j 两列得到 $\boldsymbol{B}$，则 $\boldsymbol{A}\boldsymbol{C}_{ij}=\boldsymbol{B}$；

(5) $\boldsymbol{A}$ 的第 i 列乘以非零常数 k 得到 $\boldsymbol{B}$，则 $\boldsymbol{A}\boldsymbol{R}_{i(k)}=\boldsymbol{B}$；

(6) 将 $\boldsymbol{A}$ 的第 j 列的 k 倍加到第 i 列得到 $\boldsymbol{B}$，则 $\boldsymbol{A}\boldsymbol{C}_{(i,j(k))}=\boldsymbol{B}$；

证明：我们只证明第一种情况，其它情况可类似地加以证明。设交换矩阵 $\boldsymbol{A}$ 的第 $\boldsymbol{i}$ 行、$\boldsymbol{j}$ 行得到 $\boldsymbol{B}$ 。我们要证明 $\boldsymbol{R}_{ij}\boldsymbol{A}=\boldsymbol{B}$ 。

证明：由于做初等行变换都是以一行为单位，因而我们可以将 $\boldsymbol{A}$ 、$\boldsymbol{B}$ 按行分块，即

$$A=\begin{bmatrix} A_1 \\ \vdots \\ A_{i-1} \\ A_i \\ A_{i+1} \\ \vdots \\ A_{j-1} \\ A_j \\ A_{j+1} \\ \vdots \\ A_m \end{bmatrix}\begin{matrix} \\ \\ \\ \longleftarrow \text{第 } i \text{ 行} \\ \\ \\ \\ \longleftarrow \text{第 } j \text{ 行} \\ \\ \\ \\ \end{matrix}\qquad B=\begin{bmatrix} A_1 \\ \vdots \\ A_{i-1} \\ A_j \\ A_{i+1} \\ \vdots \\ A_{j-1} \\ A_i \\ A_{j+1} \\ \vdots \\ A_m \end{bmatrix}\begin{matrix} \\ \\ \\ \longleftarrow \text{第 } i \text{ 行} \\ \\ \\ \\ \longleftarrow \text{第 } j \text{ 行} \\ \\ \\ \\ \end{matrix}$$

现考虑 $\boldsymbol{R}_{ij}$ 的第 s 行与 $\boldsymbol{A}$ 相乘的结果：

$$\left[\boldsymbol{R}_{ij}\boldsymbol{A}\right]_{s1}=\sum_{t=1}^{m}\left[\boldsymbol{R}_{ij}\right]_{st}\left[\boldsymbol{A}\right]_{t1}=\sum_{t=1}^{m}\left[\boldsymbol{R}_{ij}\right]_{st}\boldsymbol{A}_t$$

(1) 当 $s\neq i,j$ 时，在 $\boldsymbol{R}_{ij}$ 的第 s 行元素中，只有第 s 列元素为 1，其它元素为 0，因此上式等于 $\boldsymbol{A}_s$；

(2) 当 $s=i$ 时，在 $\boldsymbol{R}_{ij}$ 的第 i 行元素中，只有第 j 列元素为 1，其它元素为 0，因此上式等于 $\boldsymbol{A}_j$；

(3) 当 $s=j$ 时，在 $\boldsymbol{R}_{ij}$ 的第 j 行元素中，只有第 i 列元素为 1，其它元素为 0，因此上式等于 $\boldsymbol{A}_i$；

综合上面三种情况，我们得到

$$\boldsymbol{R}_{ij}\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1\\ \vdots\\ \boldsymbol{A}_{i-1}\\ \boldsymbol{A}_j\\ \boldsymbol{A}_{i+1}\\ \vdots\\ \boldsymbol{A}_{j-1}\\ \boldsymbol{A}_i\\ \boldsymbol{A}_{j+1}\\ \vdots\\ \boldsymbol{A}_m\end{bmatrix}=\boldsymbol{B}$$

对于其它的情形可以类似地进行证明。

或利用分块矩阵进行证明：

由于 $\boldsymbol{R}_{ij}=\begin{bmatrix}\boldsymbol{e}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{e}_{i-1}^{\mathrm{T}}\\ \boldsymbol{e}_j^{\mathrm{T}}\\ \boldsymbol{e}_{i+1}^{\mathrm{T}}\\ \vdots\\ \boldsymbol{e}_{j-1}^{\mathrm{T}}\\ \boldsymbol{e}_i^{\mathrm{T}}\\ \boldsymbol{e}_{j+1}^{\mathrm{T}}\\ \vdots\\ \boldsymbol{e}_n^{\mathrm{T}}\end{bmatrix}$，而根据已有结论，将 $\boldsymbol{A}$ 按行分块 $\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1\\ \boldsymbol{A}_2\\ \vdots\\ \boldsymbol{A}_n\end{bmatrix}$ 时，$\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}_i$。所以

$$\boldsymbol{R}_{ij}\boldsymbol{A}=\begin{bmatrix}\boldsymbol{e}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{e}_{i-1}^{\mathrm{T}}\\ \boldsymbol{e}_j^{\mathrm{T}}\\ \boldsymbol{e}_{i+1}^{\mathrm{T}}\\ \vdots\\ \boldsymbol{e}_{j-1}^{\mathrm{T}}\\ \boldsymbol{e}_i^{\mathrm{T}}\\ \boldsymbol{e}_{j+1}^{\mathrm{T}}\\ \vdots\\ \boldsymbol{e}_n^{\mathrm{T}}\end{bmatrix}\boldsymbol{A}=\begin{bmatrix}\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{A}\\ \vdots\\ \boldsymbol{e}_{i-1}^{\mathrm{T}}\boldsymbol{A}\\ \boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{A}\\ \boldsymbol{e}_{i+1}^{\mathrm{T}}\boldsymbol{A}\\ \vdots\\ \boldsymbol{e}_{j-1}^{\mathrm{T}}\boldsymbol{A}\\ \boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{A}\\ \boldsymbol{e}_{j+1}^{\mathrm{T}}\boldsymbol{A}\\ \vdots\\ \boldsymbol{e}_n^{\mathrm{T}}\boldsymbol{A}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_1\\ \vdots\\ \boldsymbol{A}_{i-1}\\ \boldsymbol{A}_j\\ \boldsymbol{A}_{i+1}\\ \vdots\\ \boldsymbol{A}_{j-1}\\ \boldsymbol{A}_i\\ \boldsymbol{A}_{j+1}\\ \vdots\\ \boldsymbol{A}_n\end{bmatrix}=\boldsymbol{B}$$

命题证毕。 ■

定理 3.5 的应用：

(1)　　由定理 3.4 容易求得初等矩阵的逆矩阵如下：

$$\boldsymbol{R}_{ij}^{-1}=\boldsymbol{R}_{ij},\quad \boldsymbol{R}_{i(k)}^{-1}=\boldsymbol{R}_{i\left(k^{-1}\right)},\quad \boldsymbol{R}_{(i,j(k))}^{-1}=\boldsymbol{R}_{(i,j(-k))}$$

$$\boldsymbol{C}_{ij}^{-1}=\boldsymbol{C}_{ij},\quad \boldsymbol{C}_{i(k)}^{-1}=\boldsymbol{C}_{i\left(k^{-1}\right)},\quad \boldsymbol{C}_{(i,j(k))}^{-1}=\boldsymbol{C}_{(i,j(-k))}$$

我们以 $\boldsymbol{R}_{ij}^{-1}=\boldsymbol{R}_{ij}$ 为例说明。由于 $\boldsymbol{R}_{ij}\xrightarrow{r_i\to r_j}\boldsymbol{E}$，根据定理 3.5，有

$$\boldsymbol{R}_{ij}\boldsymbol{R}_{ij}=\boldsymbol{E}$$

故 $\boldsymbol{R}_{ij}^{-1}=\boldsymbol{R}_{ij}$。

因此，初等矩阵的逆矩阵还是初等矩阵。

(2)　　矩阵标准形

如果只使用初等行变换，那么矩阵只能化为行阶梯形、行标准形，但是如果对行标准形矩阵再使用初等列变换，则还能做到使得每一行也只具有一个非零元 1，然后交换各列的位置，可使得第 i 行的唯一非零元 1 位于第 i 列，即此时所得的矩阵的左上角为一个单位矩阵，其它元素为 0，

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}=\begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}$$

我们称之为矩阵的标准型。如果矩阵 $\boldsymbol{A}$ 经过初等变换可化为上式，则称 $\boldsymbol{A}$ 等价于标准形。

注意：矩阵的标准形具有不同的表现形式，例如零矩阵的标准形为

$$\begin{bmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}=\boldsymbol{O}$$

此时左上角没有出现 r 阶单位矩阵；另外，对于单位矩阵 $\boldsymbol{E}$，其标准形就为单位矩阵，而没有零行。

如何化矩阵为标准形：

(1) 如果矩阵为零矩阵，则不需要任何操作；

(2) 先将矩阵化为行阶梯形；

(3) 如果行阶梯形有 r 个非零行，则矩阵标准形的左上角为一个 r 阶单位矩阵，其它元素为

0；

由于矩阵 $\boldsymbol{A}$ 经过初等变换可化为标准型，设其标准型为

$$\begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}$$

根据定理 3.5，与将 $\boldsymbol{A}$ 化为标准形所用的初等变换相对应，有一序列的行初等矩阵 $\boldsymbol{P}_t, \boldsymbol{P}_{t-1}, \cdots, \boldsymbol{P}_1$，列初等矩阵 $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_s$，使得

$$\boldsymbol{P}_t \boldsymbol{P}_{t-1} \cdots \boldsymbol{P}_1 \boldsymbol{A} \boldsymbol{Q}_1 \boldsymbol{Q}_2 \cdots \boldsymbol{Q}_s = \begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}$$

其中行初等矩阵 $\boldsymbol{P}_i\left(i=1,\cdots,t\right)$ 的阶数等于 $\boldsymbol{A}$ 的行数，列初等矩阵 $\boldsymbol{Q}_j\left(j=1,\cdots,s\right)$ 的阶数等于 $\boldsymbol{A}$ 的列数。若令 $\boldsymbol{P}=\boldsymbol{P}_t \boldsymbol{P}_{t-1} \cdots \boldsymbol{P}_1$，$\boldsymbol{Q}=\boldsymbol{Q}_1 \boldsymbol{Q}_2 \cdots \boldsymbol{Q}_s$，则有

$$\boldsymbol{PAQ} = \begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}$$

显然，$\boldsymbol{P}, \boldsymbol{Q}$ 都可逆。实际上其逆命题也成立，即：

若有可逆矩阵 $\boldsymbol{P}, \boldsymbol{Q}$ 使得 $\boldsymbol{PAQ} = \begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}$ 成立，则 $\boldsymbol{A}$ 等价于标准型 $\begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}$。

而该命题的证明需要如下定理。

定理 3.6　　设 $\boldsymbol{A}$ 为 n 阶可逆矩阵，则 $\boldsymbol{A}$ 必等价于 n 阶单位矩阵。

证明：反证法。若 $\boldsymbol{A}$ 的标准型不为 n 阶单位矩阵，则 $\boldsymbol{A}$ 的标准型为

$$\begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}$$

其中 $r<n$。

根据以上讨论可知，存在可逆矩阵 $\boldsymbol{P}, \boldsymbol{Q}$ 使得

$$\boldsymbol{PAQ} = \begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}$$

这里的 $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{A}$ 都为 n 阶方阵，两边取行列式得

$$\left|\boldsymbol{PAQ}\right| = \left|\boldsymbol{P}\right|\left|\boldsymbol{A}\right|\left|\boldsymbol{Q}\right| = \begin{vmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{vmatrix} = 0$$

由于 $\boldsymbol{P}, \boldsymbol{Q}$ 可逆，故 $\left|\boldsymbol{P}\right| \neq 0,\ \left|\boldsymbol{Q}\right| \neq 0$，因此得到

$$|A| = 0$$

这与 $\boldsymbol{A}$ 可逆矛盾。故命题成立。

■

由于可逆矩阵与单位矩阵等价，因此可逆矩阵可经过一序列的初等行变换、初等列变换化为单位矩阵，因此对于 n 阶可逆矩阵 $\boldsymbol{A}$，存在行初等矩阵 $\boldsymbol{P}_t,\boldsymbol{P}_{t-1},\cdots,\boldsymbol{P}_1$，列初等矩阵 $\boldsymbol{Q}_1,\boldsymbol{Q}_2,\cdots,\boldsymbol{Q}_s$，使得

$$\boldsymbol{P}_t\boldsymbol{P}_{t-1}\cdots\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_s = \boldsymbol{E}_n$$

因此

$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{P}_1^{-1}\cdots\boldsymbol{P}_{t-1}^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{E}_n\boldsymbol{Q}_s^{-1}\boldsymbol{Q}_{s-1}^{-1}\cdots\boldsymbol{Q}_1^{-1}\\ &= \boldsymbol{P}_1^{-1}\cdots\boldsymbol{P}_{t-1}^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{Q}_s^{-1}\boldsymbol{Q}_{s-1}^{-1}\cdots\boldsymbol{Q}_1^{-1}\\ &= \boldsymbol{P}_1^{-1}\cdots\boldsymbol{P}_{t-1}^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{Q}_s^{-1}\boldsymbol{Q}_{s-1}^{-1}\cdots\boldsymbol{Q}_1^{-1}\boldsymbol{E}\\ &= \boldsymbol{E}\boldsymbol{P}_1^{-1}\cdots\boldsymbol{P}_{t-1}^{-1}\boldsymbol{P}_t^{-1}\boldsymbol{Q}_s^{-1}\boldsymbol{Q}_{s-1}^{-1}\cdots\boldsymbol{Q}_1^{-1}\end{aligned}$$

由于初等矩阵可逆且其逆还是初等矩阵，因而有即 $\boldsymbol{A}$ 可表示成有限个初等矩阵的乘积。

已知可逆矩阵与单位矩阵等价，可综合使用初等行变换、初等列变换将可逆矩阵化为单位矩阵。由于同一个初等矩阵即可看作行初等矩阵，也可看作列初等矩阵，因此从上式中的最后两个式子可知：对于可逆矩阵，仅用初等行变换就可将之化为单位矩阵（可逆矩阵的标准形），也可仅用初等列变换将之化为单位矩阵（可逆矩阵的标准形）。

定理 3.7　　方阵 $\boldsymbol{A}$ 可逆 $\Leftrightarrow$ $\boldsymbol{A}$ 可表示为一序列初等矩阵的乘积。

结合定理 3.5，就有以下重要结论：

一个可逆矩阵左乘矩阵 $\boldsymbol{A}$ 相当于对 $\boldsymbol{A}$ 做一序列的初等行变换，一个可逆矩阵右乘 $\boldsymbol{A}$ 相当于对 $\boldsymbol{A}$ 做一序列的初等列变换。

由此可得到求逆矩阵的初等变换法。

设 $\boldsymbol{A}$ 可逆，为求得 $\boldsymbol{A}$ 的逆矩阵，构造分块矩阵 $\boldsymbol{B}$。

$$\boldsymbol{B} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{E}\end{bmatrix}$$

则

$$\begin{aligned}\boldsymbol{A}^{-1}\boldsymbol{B} &= \boldsymbol{A}^{-1}\begin{bmatrix}\boldsymbol{A} & \boldsymbol{E}\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}^{-1}\boldsymbol{A} & \boldsymbol{A}^{-1}\boldsymbol{E}\end{bmatrix}\\ &= \begin{bmatrix}\boldsymbol{E} & \boldsymbol{A}^{-1}\end{bmatrix}\end{aligned}$$

由于左乘一个可逆矩阵相当于对该矩阵做一序列的初等行变换，因此从上式可得：这些初等行变换将矩阵 $\boldsymbol{A}$ 变为单位矩阵 $\boldsymbol{E}$，则同样的初等行变换将单位矩阵 $\boldsymbol{E}$ 变为 $\boldsymbol{A}^{-1}$，而初等行变换是对整个矩阵 $\begin{bmatrix}\boldsymbol{A} & \boldsymbol{E}\end{bmatrix}$ 进行的。

如果构造分块矩阵

$$B=\begin{bmatrix}A\\E\end{bmatrix}$$

则可考虑

$$BA^{-1}=\begin{bmatrix}A\\E\end{bmatrix}A^{-1}=\begin{bmatrix}AA^{-1}\\EA^{-1}\end{bmatrix}=\begin{bmatrix}E\\A^{-1}\end{bmatrix}$$

因此，对 $B=\begin{bmatrix}A\\E\end{bmatrix}$ 作初等列变换，若这些初等列变换将 A 化为单位矩阵，则这些初等列变换将单位矩阵化为 A^{-1}。

我们可以进一步地推广上述结论。设方阵 A 可逆，则为了求 $A^{-1}B$，我们先构造分块矩阵 $\begin{bmatrix}A & B\end{bmatrix}$，则

$$A^{-1}\begin{bmatrix}A & B\end{bmatrix}=\begin{bmatrix}A^{-1}A & A^{-1}B\end{bmatrix}=\begin{bmatrix}E & A^{-1}B\end{bmatrix}$$

上式说明：对 $\begin{bmatrix}A & B\end{bmatrix}$ 做初等行变换，若该初等行变换将 A 化为单位矩阵，则该初等行变换将 B 化为 $A^{-1}B$。

类似的，为了求 BA^{-1}，构造分块矩阵 $\begin{bmatrix}A\\B\end{bmatrix}$，对它做初等列变换，若该初等列变换将 A 化为单位矩阵 E，则该初等列变换将 B 化为 BA^{-1}。

我们也可应用上述结果来说明高斯消元法。线性方程组可表示为 $AX=b$，高斯消元法是用初等行变换将该方程组的增广矩阵 $\begin{bmatrix}A & b\end{bmatrix}$ 化为行阶梯形 $\begin{bmatrix}\bar{A} & \bar{b}\end{bmatrix}$，因此存在可逆矩阵 P，使得 $\begin{bmatrix}\bar{A} & \bar{b}\end{bmatrix}=P\begin{bmatrix}A & b\end{bmatrix}$，故有 $\bar{A}=PA,\quad \bar{b}=PB$，因此以 $\begin{bmatrix}\bar{A} & \bar{b}\end{bmatrix}$ 为增广矩阵的线性方程组 $\bar{A}X=\bar{b}$ 可表示为 $PAX=Pb$，由于 P 可逆，故

$$\bar{A}X=\bar{b}\Leftrightarrow PAX=Pb\Leftrightarrow AX=b$$

故方程组 $AX=b$ 与 $\bar{A}X=\bar{b}$ 同解。

例 3.22　求 $\boldsymbol{A}=\begin{bmatrix}1 & -1 & 1\\ 2 & -4 & 1\\ 1 & -5 & 3\end{bmatrix}$ 的逆矩阵。

解
$$\begin{bmatrix}1 & -1 & 1 & 1 & 0 & 0\\ 2 & -4 & 1 & 0 & 1 & 0\\ 1 & -5 & 3 & 0 & 0 & 1\end{bmatrix}\xrightarrow[r_3-r_1]{r_2-2r_1}\begin{bmatrix}1 & -1 & 1 & 1 & 0 & 0\\ 0 & -2 & -1 & -2 & 1 & 0\\ 0 & -4 & 2 & -1 & 0 & 1\end{bmatrix}$$

$$\xrightarrow{r_3-2r_1}\begin{bmatrix}1 & -1 & 1 & 1 & 0 & 0\\ 0 & -2 & -1 & -2 & 1 & 0\\ 0 & 0 & 4 & 3 & -2 & 1\end{bmatrix}\xrightarrow{\frac{1}{4}r_3}\begin{bmatrix}1 & -1 & 1 & 1 & 0 & 0\\ 0 & -2 & -1 & -2 & 1 & 0\\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{2} & \frac{1}{4}\end{bmatrix}$$

$$\xrightarrow[r_2+r_3]{r_1-r_3}\begin{bmatrix}1 & -1 & 0 & \frac{1}{4} & \frac{1}{2} & -\frac{1}{4}\\ 0 & -2 & 0 & -\frac{5}{4} & \frac{1}{2} & \frac{1}{4}\\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{2} & \frac{1}{4}\end{bmatrix}$$

$$\xrightarrow{\frac{-1}{2}r_2}\begin{bmatrix}1 & -1 & 0 & \frac{1}{4} & \frac{1}{2} & -\frac{1}{4}\\ 0 & 1 & 0 & \frac{5}{8} & -\frac{1}{4} & -\frac{1}{8}\\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{2} & \frac{1}{4}\end{bmatrix}$$

$$\xrightarrow{r_1+r_2}\begin{bmatrix}1 & 0 & 0 & \frac{7}{8} & \frac{1}{4} & -\frac{3}{8}\\ 0 & 1 & 0 & \frac{5}{8} & -\frac{1}{4} & -\frac{1}{8}\\ 0 & 0 & 1 & \frac{3}{4} & -\frac{1}{2} & \frac{1}{4}\end{bmatrix}$$

因此 $\boldsymbol{A}$ 的逆为

$$\boldsymbol{A}^{-1}=\begin{bmatrix}\frac{7}{8} & \frac{1}{4} & -\frac{3}{8}\\ \frac{5}{8} & -\frac{1}{4} & -\frac{1}{8}\\ \frac{3}{4} & -\frac{1}{2} & \frac{1}{4}\end{bmatrix}=\frac{1}{8}\begin{bmatrix}7 & 2 & -3\\ 5 & -2 & -1\\ 6 & -4 & 2\end{bmatrix}$$

例 3.23　设方阵 $\boldsymbol{A}$ 可逆，将 $\boldsymbol{A}$ 的第二列加到第一列得到 $\boldsymbol{B}$，则 $\boldsymbol{A}^*$ 与 $\boldsymbol{B}^*$ 有何关系？

解：根据已有定理可知 $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{C}_{(1,2(1))}$ ，由于 $\boldsymbol{A}$ 可逆，故 $\boldsymbol{B}$ 也可逆。且

$$
\begin{aligned}
\boldsymbol{B}^* &= |\boldsymbol{B}|\boldsymbol{B}^{-1} = \left|\boldsymbol{A}\boldsymbol{C}_{(1,2(1))}\right|\left(\boldsymbol{A}\boldsymbol{C}_{(1,2(1))}\right)^{-1} \\
&= |\boldsymbol{A}|\left|\boldsymbol{C}_{(1,2(1))}\right|\left[\boldsymbol{C}_{(1,2(1))}\right]^{-1}\boldsymbol{A}^{-1} = |\boldsymbol{A}|\boldsymbol{C}_{(1,2(-1))}\boldsymbol{A}^{-1} \\
&= \boldsymbol{C}_{(1,2(-1))}|\boldsymbol{A}|\boldsymbol{A}^{-1} = \boldsymbol{C}_{(1,2(-1))}\boldsymbol{A}^* \\
&= \boldsymbol{R}_{(2,1(-1))}\boldsymbol{A}^*
\end{aligned}
$$

所以 $\boldsymbol{B}^*$ 可通过将 $\boldsymbol{A}^*$ 的第二行减去第一行得到。

问题：如果 $\boldsymbol{A}$ 不可逆，则上述结论是否还成立？

习题 3.5

1. 设 n 阶方阵 $\boldsymbol{A}$ 可逆，将 $\boldsymbol{A}$ 的第二列加到第一列得到 $\boldsymbol{B}$，问它们的伴随矩阵 $\boldsymbol{A}^*$、$\boldsymbol{B}^*$ 之间有何关系?若 $\boldsymbol{A}$ 不可逆， $\boldsymbol{A}^*$ 与 $\boldsymbol{B}^*$ 之间又有何关系？

2. 将矩阵 $\boldsymbol{A}=\begin{bmatrix}1&0&0\\2&0&-1\\0&-1&0\end{bmatrix}$ 表示成初等矩阵的乘积。

3. 设 $\boldsymbol{A}$ 是 n 阶可逆矩阵，将 $\boldsymbol{A}$ 的第 i 列与第 j 列互换后得到矩阵 $\boldsymbol{B}$，证明：（1）证明 $|\boldsymbol{B}|\neq 0$；（2）计算 $\boldsymbol{B}^{-1}\boldsymbol{A}$。

§3.6　分块矩阵的初等变换

我们已经学习了矩阵的初等变换，实际上，初等变换的概念也可以应用于分块矩阵，这就是分块矩阵的初等变换。

定义 3.21　分块矩阵的初等变换

设 $\boldsymbol{A}$ 是一个分块矩阵，如下的三种运算称为分块矩阵的初等行变换：

（1）　互换两行；

（2）　用一个可逆矩阵左乘 $\boldsymbol{A}$ 的某一行的所有子块；

（3）　用一个矩阵左乘 $\boldsymbol{A}$ 的某一行中的每个子块，并加到另外一行的对应子块上；

假设分块矩阵 $\boldsymbol{A}=\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{i1} & \boldsymbol{A}_{i1} & \cdots & \boldsymbol{A}_{it} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{j1} & \boldsymbol{A}_{j2} & \cdots & \boldsymbol{A}_{jt} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \cdots & \boldsymbol{A}_{st} \end{bmatrix}$，则它互换 i,j 两行的初等变换所得结果为

$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{j1} & \boldsymbol{A}_{j1} & \cdots & \boldsymbol{A}_{jt} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{i1} & \boldsymbol{A}_{i2} & \cdots & \boldsymbol{A}_{it} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \cdots & \boldsymbol{A}_{st} \end{bmatrix}$；第 i 行乘以矩阵可逆矩阵 $\boldsymbol{B}$ 所得的结果为

$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{B}\boldsymbol{A}_{i1} & \boldsymbol{B}\boldsymbol{A}_{i1} & \cdots & \boldsymbol{B}\boldsymbol{A}_{it} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{j1} & \boldsymbol{A}_{j2} & \cdots & \boldsymbol{A}_{jt} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \cdots & \boldsymbol{A}_{st} \end{bmatrix}$，而将第 i 行的每个子块都用矩阵 $\boldsymbol{B}$ 去乘（设 $\boldsymbol{B}$ 的列数等于子块 $\boldsymbol{A}_{i1}$ 的行数）、并加到第 j 行的结果为

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{i1} & \boldsymbol{A}_{i1} & \cdots & \boldsymbol{A}_{it} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{j1}+\boldsymbol{B}\boldsymbol{A}_{i1} & \boldsymbol{A}_{j2}+\boldsymbol{B}\boldsymbol{A}_{i2} & \cdots & \boldsymbol{A}_{jt}+\boldsymbol{B}\boldsymbol{A}_{it} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \cdots & \boldsymbol{A}_{st} \end{bmatrix}$$

类似地可以定义分块矩阵的初等列变换：

（1）　互换两列；

（2）　用一个可逆矩阵右乘 $\boldsymbol{A}$ 的某一列的所有子块；

（3）　用一个矩阵右乘 $\boldsymbol{A}$ 的某一列中的每个子块，并加到另外一列的对应子块上；

分块矩阵的初等行变换、初等列变换统称为分块矩阵的初等变换。

定义 3.22 分块初等矩阵

将单位矩阵分块，使得所得的分块矩阵为一准对角矩阵，然后对该准对角矩阵应用一次分块矩阵的初等变换所得的矩阵称为分块初等矩阵；经过分块矩阵的初等行变换所得到的矩阵称为分块行初等矩阵，经过分块矩阵的初等列变换所得到的矩阵称为初等列分块矩阵。根据定义可知，分块初等矩阵是可逆的。

为描述简单起见，我们不妨假设将单位矩阵分为 2×2 的准对角矩阵 $\boldsymbol{E}=\begin{bmatrix} \boldsymbol{E}_m & \\ & \boldsymbol{E}_n \end{bmatrix}$，

那么该矩阵做初等变换所得的初等分块矩阵分别为 $\begin{bmatrix} & \boldsymbol{E}_n \\ \boldsymbol{E}_m & \end{bmatrix}$，$\begin{bmatrix} \boldsymbol{B} & \\ & \boldsymbol{E}_n \end{bmatrix}$（$\boldsymbol{B}$ 可逆），

$\begin{bmatrix} \boldsymbol{E}_m & \\ \boldsymbol{B} & \boldsymbol{E}_n \end{bmatrix}$，且有 $\begin{vmatrix} & \boldsymbol{E}_n \\ \boldsymbol{E}_m & \end{vmatrix}=(-1)^{mn}\neq 0, \begin{vmatrix} \boldsymbol{B} & \\ & \boldsymbol{E}_n \end{vmatrix}=|\boldsymbol{B}|\neq 0, \begin{vmatrix} \boldsymbol{E}_m & \\ \boldsymbol{B} & \boldsymbol{E}_n \end{vmatrix}=1\neq 0$，即这三个分块初等矩阵都可逆。

关于初等分块矩阵与分块的初等变换，我们有类似于初等矩阵与矩阵的初等变换的定理：

定理 3. 8 对分块矩阵 $\boldsymbol{A}$ 施行某种初等行(列)变换，相当于用相应的行(列) 初等分块矩阵左(右)乘矩阵 $\boldsymbol{A}$ 。

例如 $\begin{bmatrix} & \boldsymbol{E}_n \\ \boldsymbol{E}_m & \end{bmatrix}\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}=\begin{bmatrix} \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \\ \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \end{bmatrix}$，$\begin{bmatrix} \boldsymbol{B} & \\ & \boldsymbol{E}_n \end{bmatrix}\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}=\begin{bmatrix} \boldsymbol{B}\boldsymbol{A}_{11} & \boldsymbol{B}\boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}$，

$\begin{bmatrix} \boldsymbol{E}_m & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{E}_n \end{bmatrix}\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}=\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{B}\boldsymbol{A}_{11}+\boldsymbol{A}_{21} & \boldsymbol{B}\boldsymbol{A}_{12}+\boldsymbol{A}_{22} \end{bmatrix}$（均假设分块矩阵乘法可以进行）。

现在，我们利用分块矩阵的初等变换证明定理：

若 $\boldsymbol{A},\boldsymbol{B}$ 为同阶方阵，则 $|\boldsymbol{A}\boldsymbol{B}|=|\boldsymbol{A}||\boldsymbol{B}|$ 。

我们讨论矩阵 $\begin{bmatrix} \boldsymbol{A} & \boldsymbol{O} \\ -\boldsymbol{E} & \boldsymbol{B} \end{bmatrix}$ 的行列式，根据分块矩阵的初等变换，

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{O} \\ -\boldsymbol{E} & \boldsymbol{B} \end{bmatrix}\xrightarrow{\text{行}}\begin{bmatrix} \boldsymbol{O} & \boldsymbol{A}\boldsymbol{B} \\ -\boldsymbol{E} & \boldsymbol{B} \end{bmatrix}$$

根据如上定理，即有

$$\begin{bmatrix} \boldsymbol{E} & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E} \end{bmatrix}\begin{bmatrix} \boldsymbol{A} & \\ -\boldsymbol{E} & \boldsymbol{B} \end{bmatrix}=\begin{bmatrix} \boldsymbol{O} & \boldsymbol{A}\boldsymbol{B} \\ -\boldsymbol{E} & \boldsymbol{B} \end{bmatrix}$$

取行列式，有

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

而

$$\begin{vmatrix} E & A \\ O & E \end{vmatrix}\begin{vmatrix} A & O \\ -E & B \end{vmatrix} = 1\begin{vmatrix} A & O \\ -E & B \end{vmatrix} = |A||B|$$

故$|AB| = |A||B|$成立。

例 3.24　设 A, B 均为 n 阶方阵，则有$|E - AB| = |E - BA|$。

证明　　考虑分块矩阵$\begin{bmatrix} E & A \\ B & E \end{bmatrix}$，分别用初等变换将它化为“**准上三角形**”：

$$\begin{bmatrix} E & O \\ -B & E \end{bmatrix}\begin{bmatrix} E & A \\ B & E \end{bmatrix} = \begin{bmatrix} E & A \\ O & E - BA \end{bmatrix}$$

$$\begin{bmatrix} E & A \\ B & E \end{bmatrix}\begin{bmatrix} E & O \\ -B & E \end{bmatrix} = \begin{bmatrix} E - AB & A \\ O & E \end{bmatrix}$$

对上两个式子取行列式，得到

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

故命题成立。

例 3.25　已知 m 阶方阵 A 、n 阶方阵 B 均可逆，证明$\begin{bmatrix} O & A \\ B & O \end{bmatrix}$可逆，并求其逆。

分析：假设其逆矩阵为$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$，则有

$$\begin{bmatrix} E_m & O \\ O & E_n \end{bmatrix} = \begin{bmatrix} O & A \\ B & O \end{bmatrix}\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} AX_{21} & AX_{22} \\ BX_{11} & BX_{12} \end{bmatrix}$$

即$\begin{cases} AX_{21} = E \\ AX_{22} = 0 \\ BX_{11} = 0 \\ BX_{12} = E \end{cases}$，由于 A, B 可逆，故有$\begin{cases} X_{21} = A^{-1} \\ X_{22} = O \\ X_{11} = O \\ X_{12} = B^{-1} \end{cases}$，即$\begin{bmatrix} O & A \\ B & O \end{bmatrix}^{-1} = \begin{bmatrix} O & B^{-1} \\ A^{-1} & O \end{bmatrix}$。

证明：直接验证$\begin{bmatrix} O & B^{-1} \\ A^{-1} & O \end{bmatrix}\begin{bmatrix} O & A \\ B & O \end{bmatrix} = E$ 即可。

习题 3.6

1. 已知方阵 $\boldsymbol{A},\boldsymbol{B}$ 都可逆，证明 $\begin{bmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B} \end{bmatrix}$ 可逆，并求其逆矩阵。

2. 设 $\boldsymbol{A},\boldsymbol{B}$ 是同阶方阵，证明：$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{vmatrix} = |\boldsymbol{A}+\boldsymbol{B}||\boldsymbol{A}-\boldsymbol{B}|$。

3. $m\times n$ 矩阵 $\boldsymbol{A}$ 和任意 $n\times m$ 矩阵 $\boldsymbol{B}$，证明：

（1）$\begin{vmatrix} \boldsymbol{E}_m & \boldsymbol{B} \\ \boldsymbol{A} & \boldsymbol{E}_n \end{vmatrix} = |\boldsymbol{E}_n-\boldsymbol{BA}| = |\boldsymbol{E}_m-\boldsymbol{AB}|$

（2）对任意数 λ，有 $\lambda^m|\boldsymbol{E}_n-\boldsymbol{BA}| = \lambda^n|\boldsymbol{E}_m-\boldsymbol{AB}|$。

4. 将 n 阶矩阵 $\boldsymbol{A}$ 分块为 $\boldsymbol{A}=\begin{bmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{\beta} \\ \boldsymbol{\alpha} & a_{nn} \end{bmatrix}$，其中 $\boldsymbol{A}_{n-1}$ 是 $n-1$ 阶可逆矩阵。若 $\boldsymbol{A}$ 可逆且已知 $\boldsymbol{A}_{n-1}^{-1}$，求 $\boldsymbol{A}^{-1}$。

§3.7　线性方程组解的情况与矩阵的秩

从我们以前所接触到的线性方程组的例子可知，有的线性方程组无解，有的线性方程组有解；在有解时，有的线性方程组有唯一解，而有的线性方程组有无穷多解。我们希望能够判断一个线性方程组到底属于哪种情况：无解、有唯一解、有无穷多解。是否除了这三种情况外，还有其它可能的情形——线性方程组有有限个解。当然，如果我们求解出该线性方程组，自然就知道其解的情况。我们知道，在用高斯消元法求解线性方程组时，先将线性方程组化为阶梯形方程组，然后再用回代法求解；但如果用高斯-约当法求解线性方程组，则相当于用矩阵的初等行变换将线性方程组中的增广矩阵化为行标准形矩阵。由于原始的线性方程组与以该行标准形矩阵为增广矩阵的线性方程组有相同的解，因此我们可首先讨论增广矩阵为行标准形矩阵的线性方程组解的情况。

我们假设线性方程组为（1.9），其增广矩阵经过初等行变换化为了行标准形矩阵。为便于讨论，我们假设该线性方程组的增广矩阵具有如下行标准形的形式[注 1]

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

该行标准型矩阵的特点是：其左上角 r 行、r 列是一个单位矩阵。以该行标准形矩阵为增广矩阵的线性方程组为

$$\begin{cases} x_1 \qquad\qquad + c_{1,r+1}x_{r+1} + \cdots + c_{1n}x_n = d_1 \\ \qquad x_2 \qquad + c_{2,r+1}x_{r+1} + \cdots + c_{2n}x_n = d_2 \\ \qquad\qquad\qquad\qquad \vdots \\ \qquad\qquad x_r + c_{r,r+1}x_{r+1} + \cdots + c_{rn}x_n = d_r \\ \qquad\qquad\qquad\qquad 0 = d_{r+1} \\ \qquad\qquad\qquad\qquad 0 = 0 \\ \qquad\qquad\qquad\qquad \vdots \\ \qquad\qquad\qquad\qquad 0 = 0 \end{cases}$$

此时很容易讨论该线性方程组解的情况：

1　若 $d_{r+1} \neq 0$，则方程

$$0 = d_{r+1}$$

显然不能成立，因此线性方程组无解；

2　若 $d_{r+1} = 0$，则还可分为两种情况

2.1　若 $r = n$，此时线性方程组的增广矩阵为

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & d_1 \\ 0 & 1 & \cdots & 0 & d_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & d_n \end{bmatrix}$$

对应的线性方程组为

$$\begin{cases} x_1 \qquad\qquad = d_1 \\ \qquad x_2 \qquad = d_2 \\ \qquad\qquad \vdots \\ \qquad\qquad x_n = d_n \end{cases}$$

所以线性方程组有惟一解。

2.2）若 $r < n$，则增广矩阵为

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

我们不需考虑$0=0$类型的方程，而只需考虑线性方程组

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

我们将自由未知量$x_{r+1},\cdots,x_n$移到等号右边，得到

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

因此，对于给定的$x_{r+1},\cdots,x_n$的值，我们都可以得到$x_1,x_2,\cdots,x_r$的值，因而此时线性方程组有无穷多解，

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

根据以上分析可知，不存在线性方程组有多于一个的有限个解的情况。

因而对于增广矩阵为行标准形矩阵的线性方程组，我们能够很容易地判断其解的情况。但是如果线性方程组的增广矩阵不为行标准形矩阵，那么该如何判断其解的情况呢？当然我们可以首先将其增广矩阵化为行标准形矩阵，然后利用以上结论来判断。但是否能够根据增广矩阵的“特点”来判断该方程组解的情况呢？更直观地说，是否能够给出增广矩阵的某个指标，而使得我们能够根据该指标来判断方程组解的情况。显然，为了达到我们的目标，满足该要求的矩阵指标与其行标准形矩阵的指标应该有某种联系，这样才能利用上述分析的结果。而满足以上要求的一个候选者就是矩阵的秩。

矩阵的秩是矩阵的一个重要特征，在很多情况下起着非常重要的作用。

定义 3.23　　矩阵的k阶子式

设有$m\times n$矩阵

$$A=\left[a_{ij}\right]_{m\times n}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

从矩阵 $\boldsymbol{A}$ 中选取 k 行（$i_1<i_2<\cdots<i_k$）、k 列（$j_1<j_2<\cdots<j_k$），位于这些行和列相交处的元素所定义的 k 阶行列式

$$\begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_k} \\ \vdots & \vdots & & \vdots \\ a_{i_k j_1} & a_{i_k j_2} & \cdots & a_{i_k j_k} \end{vmatrix}$$

称为矩阵 $\boldsymbol{A}$ 的一个 k 阶子式，其中 $k\le\min\{m,n\}$。比如，对于矩阵 5×4 矩阵

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \\ 17 & 18 & 19 & 20 \end{bmatrix}$$

我们可以给出其若干子式

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \\ 17 & 18 & 19 & 20 \end{bmatrix} \qquad \begin{vmatrix} 2 & 4 \\ 10 & 12 \end{vmatrix}=-16$$ 是一个不为零的二阶子式

为非零子式

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \\ 17 & 18 & 19 & 20 \end{bmatrix} \qquad \begin{vmatrix} 5 & 7 & 8 \\ 9 & 11 & 12 \\ 13 & 15 & 16 \end{vmatrix}=0$$ 是一个为零的三阶子式

显然，对于 $m\times n$ 矩阵 $\boldsymbol{A}$，共有 $C_m^k C_n^k$ 个 k 阶子式。

定义 3.24　　矩阵的秩

在矩阵 $\boldsymbol{A}$ 的所有不为 0 的子式中，阶数最高的子式的阶数称为矩阵的秩，记为 $r(\boldsymbol{A})$，

或秩（A）， rank$A$ 。

由于零矩阵没有不为零的子式，我们规定：零矩阵的秩为 0。

根据矩阵秩的定义，我们能得到关于矩阵秩的一些基本结论：

1、若 A 为 $m\times n$ 矩阵，则 $r(A)\le \min\{m,n\}$，即矩阵的秩不会超过其行数、列数的最小值；

2、若矩阵 A 的所有 $r+1$ 阶子式为 0，则由行列式展开定理可知 A 的所有 $r+2$ 阶子式为 0，依此类推可知 A 的所有阶数大于 r 的子式全为零，因此有 $r(A)\le r$；若 A 有一个 r 阶子式不为零，则 $r(A)\ge r$；

3、阶梯形矩阵的秩等于矩阵中非零行的个数；

4、$k\ne 0$ 时，$r(A)=r(kA)$；

5、$r(A)=r(A^{\mathrm{T}})$；

6、$r(A)\le r([A\quad B])$，$r(A)\le r\left(\begin{bmatrix} A \\ B \end{bmatrix}\right)$。

例 3.26　求下述矩阵的秩.

$$A=\begin{bmatrix} 1 & 3 & -9 & 3 \\ 0 & 1 & -3 & 4 \\ -2 & -3 & 9 & 6 \end{bmatrix}$$

有四个三阶子式。注意到该矩阵的第 2 列、第 3 列成比例，因而同时含有这两列元素的三阶子式为零，因而只需要考虑两个三阶子式

$$\begin{vmatrix} 1 & 3 & 3 \\ 0 & 1 & 4 \\ -2 & -3 & 6 \end{vmatrix}=\begin{vmatrix} 1 & 3 & 3 \\ 0 & 1 & 4 \\ 0 & 3 & 12 \end{vmatrix}=0$$

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

因而，$r(A)<3$。又由于二阶子式

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

所以，$r(A)=2$。

问题：如何求矩阵的秩？

根据定义直接求(如上述例子)矩阵的秩，其缺点是运算量大，而且只对简单矩阵适用。对于一般的矩阵，我们根据以下定理与性质 3 来求矩阵的秩。

定理 3.9　初等变换不改变矩阵的秩。

证明：考虑初等行变换。设矩阵 $\boldsymbol{A}$ 经过一次初等行变换化为 $\boldsymbol{B}$，我们证明 $r(\boldsymbol{A})\le r(\boldsymbol{B})$。

设 $r(\boldsymbol{A})=r$，我们证明 $r(\boldsymbol{B})\ge r$，只需要证明 $\boldsymbol{B}$ 至少有一个 r 阶不为零的子式。由于 $r(\boldsymbol{A})=r$，故 $\boldsymbol{A}$ 至少有一个 r 阶子式 $D_r\ne 0$。

(1)如果 $\boldsymbol{A}$ 互换两行得到 $\boldsymbol{B}$，则总能在 $\boldsymbol{B}$ 中找到与 D_r 对应的子式 $\overline{D_r}$，由于 $\overline{D_r}=D_r$ 或 $\overline{D_r}=-D_r$，因此 $\overline{D_r}\ne 0$，所以

$$r(\boldsymbol{B})\ge r$$

（2）如果 $\boldsymbol{A}\xrightarrow{r_i\times k}\boldsymbol{B}$，则在 $\boldsymbol{B}$ 中找一个与 D_r 位置相同的子式 $\overline{D_r}$，显然 $\overline{D_r}=kD_r$ 或 $\overline{D_r}=D_r$，所以 $\overline{D_r}\ne 0$；

（3）如果 $\boldsymbol{A}\xrightarrow{r_i+kr_j}\boldsymbol{B}$，分为三种情况

（3.1） D_r 中不含有第 i 行，

（3.2） D_r 中同时含有第 i,j 行；

（3.3） D_r 中含有第 i 行但不含有第 j 行；

对于（3.1）、（3.2）两种情况，显然 $\boldsymbol{B}$ 中与 D_r 对应的子式 $\overline{D_r}=D_r\ne 0$，所以 $r(\boldsymbol{B})\ge r$。

对于情况（3.3），由

$$\begin{aligned}\overline{D_r}&=\begin{vmatrix}\cdots & \cdots & \cdots & \cdots\\ a_{is_1}+ka_{js_1} & a_{is_2}+ka_{js_2} & \cdots & a_{is_r}+ka_{js_r}\\ \cdots & \cdots & \cdots & \cdots\end{vmatrix}\\ &=\begin{vmatrix}\cdots & \cdots & \cdots & \cdots\\ a_{is_1} & a_{is_2} & \cdots & a_{is_r}\\ \cdots & \cdots & \cdots & \cdots\end{vmatrix}+k\begin{vmatrix}\cdots & \cdots & \cdots & \cdots\\ a_{js_1} & a_{js_2} & \cdots & a_{js_r}\\ \cdots & \cdots & \cdots & \cdots\end{vmatrix}\\ &=D_r+k\hat{D}_r\end{aligned}$$

故若 $\hat{D}_r=0$，则 $\overline{D}_r=D_r\ne 0$，$\boldsymbol{B}$ 有一个不为零的子式 $\overline{D}_r$，所以 $r(\boldsymbol{B})\ge r$；

若 $\hat{D}_r \neq 0$，则因 $\hat{D}_r$ 中不含有第 i 行知 $\boldsymbol{A}$ 中有不含有第 i 行的 r 阶非零子式，从而由（3.1）知 $r(\boldsymbol{B}) \geq r$。

以上证明了若 $\boldsymbol{A}$ 经过一次初等行变换变为 $\boldsymbol{B}$，则 $r(\boldsymbol{A}) \leq r(\boldsymbol{B})$。由于 $\boldsymbol{B}$ 也可经过一次初等行变换变为 $\boldsymbol{A}$，故也有 $r(\boldsymbol{B}) \leq r(\boldsymbol{A})$。故 $r(\boldsymbol{A}) = r(\boldsymbol{B})$。

■

根据该定理，为求 $r(\boldsymbol{A})$，我们可将矩阵 $\boldsymbol{A}$ 通过初等变换化为阶梯形 $\boldsymbol{B}$，由定理 3.9 有 $r(\boldsymbol{A}) = r(\boldsymbol{B})$，由于 $\boldsymbol{B}$ 为阶梯形矩阵，故 $r(\boldsymbol{B})$ 等于 $\boldsymbol{B}$ 的非零行的个数。

例 3.27　求矩阵 $\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 0 & 5 & 0 \\ 3 & -2 & 3 & 6 & -1 \\ 2 & 0 & 1 & 5 & -3 \\ 1 & 6 & -4 & -1 & 4 \end{bmatrix}$ 的秩，并求 $\boldsymbol{A}$ 的一个最高阶非零子式。

解：方法 1　首先用初等行变换将矩阵 $\boldsymbol{A}$ 化为阶梯形矩阵

$$\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 0 & 5 & 0 \\ 3 & -2 & 3 & 6 & -1 \\ 2 & 0 & 1 & 5 & -3 \\ 1 & 6 & -4 & -1 & 4 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_4} \begin{bmatrix} 1 & 6 & -4 & -1 & 4 \\ 3 & -2 & 3 & 6 & -1 \\ 2 & 0 & 1 & 5 & -3 \\ 3 & 2 & 0 & 5 & 0 \end{bmatrix}$$

$$\xrightarrow[r_4-3r_1]{\substack{r_2-3r_1 \\ r_3-2r_1}} \begin{bmatrix} 1 & 6 & -4 & -1 & 4 \\ 0 & -20 & 15 & 9 & -13 \\ 0 & -12 & 9 & 7 & -11 \\ 0 & -16 & 12 & 8 & -12 \end{bmatrix} \xrightarrow{r_4 \div (-4)} \begin{bmatrix} 1 & 6 & -4 & -1 & 4 \\ 0 & -20 & 15 & 9 & -13 \\ 0 & -12 & 9 & 7 & -11 \\ 0 & 4 & -3 & -2 & 3 \end{bmatrix}$$

$$\xrightarrow{r_2 \leftrightarrow r_4} \begin{bmatrix} 1 & 6 & -4 & -1 & 4 \\ 0 & 4 & -3 & -2 & 3 \\ 0 & -12 & 9 & 7 & -11 \\ 0 & -20 & 15 & 9 & -13 \end{bmatrix} \xrightarrow{\substack{r_3+3r_2 \\ r_4+5r_2}} \begin{bmatrix} 1 & 6 & -4 & -1 & 4 \\ 0 & 4 & -3 & -2 & 3 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

$$\xrightarrow{r_4+r_3} \begin{bmatrix} 1 & 6 & -4 & -1 & 4 \\ 0 & 4 & -3 & -2 & 3 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \boldsymbol{B}$$

该阶梯形矩阵 $\boldsymbol{B}$ 有三个非零行，因此 $r(\boldsymbol{A}) = r(\boldsymbol{B}) = 3$；在 $\boldsymbol{B}$ 中，第 1，2，3 行、1，2，4

列所确定的子式$\begin{vmatrix} 1 & 6 & -1 \\ 0 & 4 & -2 \\ 0 & 0 & 1 \end{vmatrix} = 4$不为零；但是题目要求是在 $\boldsymbol{A}$ 中找一个三阶的不为零子式，为此我们在 $\boldsymbol{A}$ 选 1，2，4 列，然后看 $\boldsymbol{B}$ 的第 1，2，3 行是与 $\boldsymbol{A}$ 的哪些行对应的。在上述用初等行变换将 $\boldsymbol{A}$ 化为 $\boldsymbol{B}$ 的过程中，我们做了两次行交换 $r_1 \leftrightarrow r_4, r_2 \leftrightarrow r_4$，因此 $\boldsymbol{B}$ 的第 1 行对应 $\boldsymbol{A}$ 的第 4 行、$\boldsymbol{B}$ 的第 2 行对应 $\boldsymbol{A}$ 的第 1 行、$\boldsymbol{B}$ 的第 4 行对应 $\boldsymbol{A}$ 的第 2 行，$\boldsymbol{B}$ 的第 3 行对应 $\boldsymbol{A}$ 的第 3 行，因此在 $\boldsymbol{A}$ 取第 4、1、3 行。在 $\boldsymbol{A}$ 中，由第 4、1、3 行、1，2，4 列确定的子式一定不为零，该子式为

$$\begin{vmatrix} 3 & 2 & 5 \\ 2 & 0 & 5 \\ 1 & 6 & -1 \end{vmatrix} = -16 \neq 0$$

方法 2　用初等行变换中的第 2、3 种运算(不使用互换两行的运算)，将 $\boldsymbol{A}$ 化为一种特殊类型的矩阵 $\boldsymbol{B}$，对 $\boldsymbol{B}$ 只需经过互换行的运算就可以化为阶梯形矩阵，

$$\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 0 & 5 & 0 \\ 3 & -2 & 3 & 6 & -1 \\ 2 & 0 & 1 & 5 & -3 \\ 1 & 6 & -4 & -1 & 4 \end{bmatrix} \xrightarrow[r_3-2r_4]{\substack{r_1-3r_4 \\ r_2-3r_4}} \begin{bmatrix} 0 & -16 & 12 & 8 & -12 \\ 0 & -20 & 15 & 9 & -13 \\ 0 & -12 & 9 & 7 & -11 \\ 1 & 6 & -4 & -1 & 4 \end{bmatrix}$$

$$\xrightarrow{r_1 \div 4} \begin{bmatrix} 0 & -4 & 3 & 2 & -3 \\ 0 & -20 & 15 & 9 & -13 \\ 0 & -12 & 9 & 7 & -11 \\ 1 & 6 & -4 & -1 & 4 \end{bmatrix} \xrightarrow{\substack{r_2-5r_1 \\ r_3-3r}} \begin{bmatrix} 0 & -4 & 3 & 2 & -3 \\ 0 & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 & -2 \\ 1 & 6 & -4 & -1 & 4 \end{bmatrix}$$

$$\xrightarrow{r_2+r_3} \begin{bmatrix} 0 & -4 & 3 & 2 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 \\ 1 & 6 & -4 & -1 & 4 \end{bmatrix} = \boldsymbol{B}$$

所以 $r(\boldsymbol{A}) = r(\boldsymbol{B})$。这个矩阵经过互换两行的初等行变换就可把它化为行阶梯形，它有 3 个非零行，故 $r(\boldsymbol{B}) = 3$。显然，我们在 $\boldsymbol{B}$ 中可找一个不为零的三阶子式:第 1、3、4 行，第 1、2、4 列，即

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

则 $\boldsymbol{A}$ 对应位置的子式也不为零，即

作者：卢世荣　lshrlshr@163.com　欢迎任何意见与建议！

$$\begin{vmatrix} 3 & 2 & 5 \\ 2 & 0 & 5 \\ 1 & 6 & -1 \end{vmatrix} = -16 \neq 0$$

问题：为什么有这个结论？这样做的依据是什么？

定理 3.10　　若 $\boldsymbol{P},\boldsymbol{Q}$ 可逆，则 $r\left(\boldsymbol{PAQ}\right)=r\left(\boldsymbol{A}\right)$。

证明：可逆矩阵 $\boldsymbol{P}$ 左乘 $\boldsymbol{A}$ 相当于对 $\boldsymbol{A}$ 做一序列的初等行变换，可逆矩阵 $\boldsymbol{Q}$ 右乘 $\boldsymbol{A}$ 相当于对 $\boldsymbol{A}$ 做一序列的初等列变换，而初等变换不改变矩阵的秩，因此本定理成立。

根据增广矩阵为行标准形矩阵的线性方程组的解的情况的讨论以及矩阵秩的定义与性质，我们有如下定理：

定理 3.11　　设 $m\times n$ 型线性方程组的系数矩阵、增广矩阵分别为 $\boldsymbol{A}$，$\tilde{\boldsymbol{A}}$，则

2. 若 $r\left(\boldsymbol{A}\right)<r\left(\tilde{\boldsymbol{A}}\right)$，则线性方程组无解；

3. 若 $r\left(\boldsymbol{A}\right)=r\left(\tilde{\boldsymbol{A}}\right)$，则线性方程组有解；

   a） 若 $r\left(\boldsymbol{A}\right)<r\left(\tilde{\boldsymbol{A}}\right)=n$，则线性方程组有唯一解；

   b） 若 $r\left(\boldsymbol{A}\right)<r\left(\tilde{\boldsymbol{A}}\right)<n$，则线性方程组有无穷多解；

推论　　设 $m\times n$ 型齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的系数矩阵的秩 $r\left(\boldsymbol{A}\right)=r$，则

（1）　$\boldsymbol{AX}=\boldsymbol{0}$ 有无穷多解 $\Leftrightarrow r\left(\boldsymbol{A}\right)=r<n$；

（2）　$\boldsymbol{AX}=\boldsymbol{0}$ 仅有零解 $\Leftrightarrow r\left(\boldsymbol{A}\right)=r=n$。

例 3.28　判断如下方程组解的情况，如果方程组有解，则求其解。

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

解：该线性方程组的增广矩阵为

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

利用矩阵的初等行变换将其划为阶梯形

$$\begin{bmatrix} 1 & -5 & 2 & -3 & 11 \\ 0 & 14 & -2 & 7 & -28 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

所以$r(\boldsymbol{A}) < r(\tilde{\boldsymbol{A}}) = 2$，所以该线性方程组有解；将增广矩阵化为行标准形

$$\begin{bmatrix} 1 & 0 & \frac{9}{7} & -\frac{1}{2} & 1 \\ 0 & 1 & -\frac{1}{7} & \frac{1}{2} & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

取x_3, x_4为自由未知量，其解为

$$\begin{cases} x_1 = 1 - \frac{9}{7}x_3 + \frac{1}{2}x_4 \\ x_2 = -2 + \frac{1}{7}x_3 - \frac{1}{2}x_4 \\ x_3 = x_3 \\ x_4 = x_4 \end{cases}$$

例 3.29 当k取何值时，方程组$\begin{cases} x_1 + x_2 + kx_3 = 4 \\ -x_1 + kx_2 + x_3 = k^2 \\ x_1 - x_2 + 2x_3 = -4 \end{cases}$有唯一解、无解、有无穷多组解？并在有无穷多组解时，求通解表达式

解 方法（1）回答本问题，就是分析k取不同值时线性方程组系数矩阵的秩与增广矩阵的秩的关系。增广矩阵为

$$\begin{bmatrix} 1 & 1 & k & 4 \\ -1 & k & 1 & k^2 \\ 1 & -1 & 2 & -4 \end{bmatrix}$$

用初等行变换将其化为行阶梯形

$$\begin{bmatrix} 1 & 1 & k & 4 \\ -1 & k & 1 & k^2 \\ 1 & -1 & 2 & -4 \end{bmatrix} \overset{\substack{r_1 \leftrightarrow r_3 \\ r_2 \leftrightarrow r_3}}{\sim} \begin{bmatrix} 1 & -1 & 2 & -4 \\ 1 & 1 & k & 4 \\ -1 & k & 1 & k^2 \end{bmatrix} \overset{\substack{r_2 - r_1 \\ r_3 + r_1}}{\sim} \begin{bmatrix} 1 & -1 & 2 & -4 \\ 0 & 2 & k-2 & 8 \\ 0 & k-1 & 3 & k^2-4 \end{bmatrix}$$

$$\overset{r_3-\frac{k-1}{2}r_2}{\sim}\begin{bmatrix}1 & -1 & 2 & -4\\ 0 & 2 & k-2 & 8\\ 0 & 0 & 3-\dfrac{(k-1)(k-2)}{2} & k^2-4-4(k-1)\end{bmatrix}$$

$$=\begin{bmatrix}1 & -1 & 2 & -4\\ 0 & 2 & k-2 & 8\\ 0 & 0 & -\dfrac{k^2-3k-4}{2} & k^2-4k\end{bmatrix}$$

因此，当$k^2-3k-4\neq 0$，即$k\neq 4,-1$时，该线性方程组有唯一解；

当$k=4$时，增广矩阵化为$\begin{bmatrix}1 & -1 & 2 & -4\\ 0 & 2 & 2 & 8\\ 0 & 0 & 0 & 0\end{bmatrix}\sim\begin{bmatrix}1 & -1 & 2 & -4\\ 0 & 1 & 1 & 4\\ 0 & 0 & 0 & 0\end{bmatrix}\sim\begin{bmatrix}1 & 0 & 3 & 0\\ 0 & 1 & 1 & 4\\ 0 & 0 & 0 & 0\end{bmatrix}$，故解为

$$\begin{cases}x_1=-3x_3\\ x_2=4-x_3\\ x_3=x_3\end{cases}$$

而当$k=-1$时，增广矩阵化为$\begin{bmatrix}1 & -1 & 2 & -4\\ 0 & 2 & -3 & 8\\ 0 & 0 & 0 & 5\end{bmatrix}$，显然无解。

方法（2）由于该线性方程组的系数矩阵为一个方阵，而线性方程组有唯一解的充要条件是

系数矩阵的秩＝增广矩阵的秩＝未知数个数

因此在线性方程组的增广矩阵为方阵的情况下，线性方程组有唯一解的充要条件是系数矩阵可逆，即系数行列式不为零！

$$\begin{vmatrix}1 & 1 & k\\ -1 & k & 1\\ 1 & -1 & 2\end{vmatrix}\overset{\substack{r_2+r_1\\ r_3-r_1}}{=}\begin{vmatrix}1 & 1 & k\\ 0 & k+1 & k+1\\ 0 & -2 & 2-k\end{vmatrix}=(k+1)\begin{vmatrix}1 & 1 & k\\ 0 & 1 & 1\\ 0 & -2 & 2-k\end{vmatrix}$$

$$\overset{r_3+2r_2}{=}(k+1)\begin{vmatrix}1 & 1 & k\\ 0 & 1 & 1\\ 0 & 0 & 4-k\end{vmatrix}=(k+1)(4-k)$$

故$k\neq 4,-1$时，线性方程组有唯一解；

当$k=4$时，增广矩阵为$\begin{bmatrix}1&1&4&4\\-1&4&1&16\\1&-1&2&-4\end{bmatrix}$，化为行最简形得$\begin{bmatrix}1&0&3&0\\0&1&1&4\\0&0&0&0\end{bmatrix}$，因此方程组此时有无穷多解，易求得其解。

当$k=-1$时，增广矩阵为$\begin{bmatrix}1&1&-1&4\\-1&-1&1&1\\1&-1&2&-4\end{bmatrix}$，化为行阶梯形$\begin{bmatrix}1&-1&2&-4\\0&2&-3&8\\0&0&0&5\end{bmatrix}$，故线性方程组无解。

注：此种方法适合于线性方程组的系数矩阵为方阵、且用初等变换不容易把它化为行阶梯形的情况。

在介绍逆矩阵时，我们提出了矩阵乘法满足消去律的问题。现在，我们能够完整地回答该问题。

例 3.30 (1) $\boldsymbol{A}_{m\times k}\boldsymbol{B}_{k\times n}=\boldsymbol{A}_{m\times k}\boldsymbol{C}_{k\times n}\Rightarrow\boldsymbol{B}=\boldsymbol{C}$的充要条件是$r(\boldsymbol{A})=k$（$\boldsymbol{A}$是列满秩的）;

（2）$\boldsymbol{B}_{m\times k}\boldsymbol{A}_{k\times n}=\boldsymbol{C}_{m\times k}\boldsymbol{A}_{k\times n}\Rightarrow\boldsymbol{B}=\boldsymbol{C}$的充要条件是$r(\boldsymbol{A})=k$（$\boldsymbol{A}$是行满秩的）。

证明　我们先证明（1）。根据以前的分析，我们已知$\boldsymbol{A}_{m\times k}\boldsymbol{B}_{k\times n}=\boldsymbol{A}_{m\times k}\boldsymbol{C}_{k\times n}\Rightarrow\boldsymbol{B}=\boldsymbol{C}$的一个充要条件是

$$\boldsymbol{A}_{m\times k}\boldsymbol{X}_{k\times n}=\boldsymbol{O}\Rightarrow\boldsymbol{X}=\boldsymbol{O}$$

即$\boldsymbol{A}_{m\times k}\boldsymbol{X}_{k\times n}=\boldsymbol{O}$只有解$\boldsymbol{X}=\boldsymbol{O}$。

我们将$\boldsymbol{X},\boldsymbol{O}$分别按列分块，记$\boldsymbol{X}=\begin{bmatrix}\boldsymbol{X}_1&\boldsymbol{X}_2&\cdots&\boldsymbol{X}_k\end{bmatrix}$，就得到

$$\boldsymbol{A}\boldsymbol{X}_i=\boldsymbol{0},\qquad i=1,2,\cdots,n$$

若$\boldsymbol{A}_{m\times k}\boldsymbol{X}_{k\times n}=\boldsymbol{O}$只有解$\boldsymbol{X}=\boldsymbol{O}$，则说明$\boldsymbol{X}_i=\boldsymbol{0},i=1,2,\cdots,n$，故齐次线性方程组$\boldsymbol{A}\boldsymbol{X}_i=\boldsymbol{0},i=1,2,\cdots,n$只有零解，因此$r(\boldsymbol{A})=k$。

反之，若$r(\boldsymbol{A})=k$，则齐次线性方程组

$$\boldsymbol{A}\boldsymbol{X}_i=\boldsymbol{0},\qquad i=1,2,\cdots,n$$

仅有零解，即$\boldsymbol{X}_i=\boldsymbol{0},i=1,2,\cdots,n$，因此$\boldsymbol{A}_{m\times k}\boldsymbol{X}_{k\times n}=\boldsymbol{O}$仅有解$\boldsymbol{X}=\boldsymbol{O}$。

（2）由于$\boldsymbol{B}_{m\times k}\boldsymbol{A}_{k\times n}=\boldsymbol{C}_{m\times k}\boldsymbol{A}_{k\times n}\Rightarrow\boldsymbol{B}=\boldsymbol{C}$可等价地写成$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\Rightarrow\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{C}^{\mathrm{T}}$，根据（1）的结论，并注意到$\boldsymbol{A}^{\mathrm{T}}$是一个$n\times k$矩阵，因此其充要条件是$r(\boldsymbol{A}^{\mathrm{T}})=k$，即$r(\boldsymbol{A})=k$。

下面我们讨论关于矩阵秩的不等式的一些结论。

定理 3.12　　两个矩阵乘积的秩不超过每个因子的秩，即

$$r(\boldsymbol{AB}) \le \min\{r(\boldsymbol{A}), r(\boldsymbol{B})\}$$

证明：

设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，$\boldsymbol{B}$ 为 $n\times k$ 矩阵，$r(\boldsymbol{A})=r$。故存在 m 阶可逆矩阵 $\boldsymbol{P}$、k 阶可逆矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{PAQ}=\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}$$

因此

$$\boldsymbol{A}=\boldsymbol{P}^{-1}\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{Q}^{-1}$$

$$\boldsymbol{AB}=\boldsymbol{P}^{-1}\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{Q}^{-1}\boldsymbol{B}$$

将 $\boldsymbol{Q}^{-1}\boldsymbol{B}$ 分块为

$$\boldsymbol{Q}^{-1}\boldsymbol{B}=\begin{bmatrix}\boldsymbol{C}_1\\ \boldsymbol{C}_2\end{bmatrix}$$

其中 $\boldsymbol{C}_1$ 为 $r\times k$ 矩阵，$\boldsymbol{C}_2$ 为 $(n-r)\times k$ 矩阵，于是

$$\boldsymbol{AB}=\boldsymbol{P}^{-1}\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\begin{bmatrix}\boldsymbol{C}_1\\ \boldsymbol{C}_2\end{bmatrix}=\boldsymbol{P}^{-1}\begin{bmatrix}\boldsymbol{C}_1\\ \boldsymbol{O}\end{bmatrix}$$

故

$$r(\boldsymbol{AB})=r\left(\boldsymbol{P}^{-1}\begin{bmatrix}\boldsymbol{C}_1\\ \boldsymbol{O}\end{bmatrix}\right)=r\left(\begin{bmatrix}\boldsymbol{C}_1\\ \boldsymbol{O}\end{bmatrix}\right)=r(\boldsymbol{C}_1)\le r=r(\boldsymbol{A})。$$

类似可证 $r(\boldsymbol{AB})\le r(\boldsymbol{B})$。 ■

定理 3.13（Sylvester 公式）设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，$\boldsymbol{B}$ 为 $n\times k$ 矩阵，则

$$r(\boldsymbol{AB})\ge r(\boldsymbol{A})+r(\boldsymbol{B})-n$$

特别地，若 $\boldsymbol{AB}=\boldsymbol{O}$，则 $r(\boldsymbol{A})+r(\boldsymbol{B})\le n$。

证明：由于

$$\begin{bmatrix} E_n & O \\ -A & E_m \end{bmatrix}\begin{bmatrix} E_n & B \\ A & O \end{bmatrix}\begin{bmatrix} E_n & -B \\ O & E_k \end{bmatrix}=\begin{bmatrix} E_n & O \\ O & -AB \end{bmatrix}$$

且$\begin{bmatrix} E_n & O \\ -A & E_k \end{bmatrix},\begin{bmatrix} E_n & -B \\ O & E_k \end{bmatrix}$均可逆，所以

$$r\left(\begin{bmatrix} E_n & B \\ A & O \end{bmatrix}\right)=r\left(\begin{bmatrix} E_n & O \\ O & -AB \end{bmatrix}\right)=r(E_n)+r(-AB)=n+r(AB)$$

但

$$r\left(\begin{bmatrix} E_n & B \\ A & O \end{bmatrix}\right)=r\left(\begin{bmatrix} B & E_n \\ O & A \end{bmatrix}\right)$$

显然，若B有一个t_1阶子式不为零，A有一个t_2阶子式不为零，则$\begin{bmatrix} E_n & B \\ A & O \end{bmatrix}$一定有一个相应的$t_1+t_2$阶子式不为零，因此

$$r\left(\begin{bmatrix} E_n & B \\ A & O \end{bmatrix}\right)\ge r(A)+r(B)$$

故$r(AB)\ge r(A)+r(B)-n$。 ■

定理 3.14　设A,B,C分别为$m\times n,n\times s,s\times t$矩阵，则

$$r(AB)+r(BC)\le r(B)+r(ABC)$$

证　构造矩阵$\begin{bmatrix} AB & O \\ B & BC \end{bmatrix}$，因为

$$\begin{bmatrix} O & E_n \\ E_m & O \end{bmatrix}\begin{bmatrix} E_m & -A \\ O & E_n \end{bmatrix}\begin{bmatrix} AB & O \\ B & BC \end{bmatrix}\begin{bmatrix} E_s & O \\ O & -E_t \end{bmatrix}=\begin{bmatrix} B & 0 \\ 0 & ABC \end{bmatrix}$$

且$\begin{bmatrix} O & E_n \\ E_m & O \end{bmatrix},\begin{bmatrix} E_m & -A \\ O & E_n \end{bmatrix},\begin{bmatrix} E_s & O \\ O & -E_t \end{bmatrix}$都可逆，因此

$$r\left(\begin{bmatrix} AB & O \\ B & BC \end{bmatrix}\right)=r\left(\begin{bmatrix} B & O \\ O & ABC \end{bmatrix}\right)$$

故

$$r(AB)+r(BC)\le r\left(\begin{bmatrix} AB & O \\ B & BC \end{bmatrix}\right)=r\left(\begin{bmatrix} B & O \\ O & ABC \end{bmatrix}\right)$$
$$=r(B)+r(ABC)$$

在上式中令 $B=E_n$，即得 $r(A)+r(C)\le n+r(AC)$。■

定理 3.15　设 A,B 均为 $m\times n$ 矩阵，则

$$r(A+B)\le r(A)+r(B)$$

证明：注意到

$$\begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} E_n \\ E_n \end{bmatrix}=A+B$$

因此有

$$r(A+B)\le r(\begin{bmatrix} A & B \end{bmatrix})$$

另外

$$\begin{bmatrix} E_m & E_m \\ O & E_m \end{bmatrix}\begin{bmatrix} A & O \\ O & B \end{bmatrix}=\begin{bmatrix} A & B \\ O & B \end{bmatrix}$$

由于 $\begin{bmatrix} E_m & E_m \\ O & E_m \end{bmatrix}$ 可逆，所以

$$r\left(\begin{bmatrix} A & B \\ O & B \end{bmatrix}\right)=r\left(\begin{bmatrix} A & O \\ O & B \end{bmatrix}\right)=r(A)+r(B)$$

而

$$r(\begin{bmatrix} A & B \end{bmatrix})\le r\left(\begin{bmatrix} A & B \\ 0 & B \end{bmatrix}\right)$$

综上可得

$$r(A+B)\le r(\begin{bmatrix} A & B \end{bmatrix})\le r\left(\begin{bmatrix} A & B \\ 0 & B \end{bmatrix}\right)=r\left(\begin{bmatrix} A & O \\ O & B \end{bmatrix}\right)=r(A)+r(B)$$

命题得证。■

例 3.31　设 A 为 n 阶幂等矩阵，即 $A^2=A$，证明：

$$r(A)+r(E-A)=n$$

证明：由 $A^2=A$，得到

$$A(E-A)=O$$

所以

$$r(A)+r(E-A)\le n$$

另外

$$r(A)+r(E-A)\ge r(A+E-A)=r(E)=n$$

综上可得

$$r(A)+r(E-A)=n$$

■

例 3.26　$A_{m\times n}B_{n\times k}=C_{m\times k}$，且$r(A)=n$，必有$r(B)=r(C)$。

证明：根据定理 3.11，有

$$r(C)\le r(B)$$

根据定理 3.12，有

$$r(C)\ge r(A)+r(B)-n=n+r(B)-n=r(B)$$

综合上述两个不等式即得

$$r(B)=r(C)$$

类似地，若$r(B)=n$，必有$r(A)=r(C)$。

定理 3.16　矩阵方程$AX=B$有解的充要条件是$r(A)=r([A\quad B])$。

证明：设A为$m\times n$矩阵，B为$n\times k$矩阵，对矩阵X,B如下分块：

$$X=[X_1\quad X_2\quad \cdots\quad X_k]$$

$$B=[B_1\quad B_2\quad \cdots\quad B_k]$$

则有

$$\begin{aligned}
&A_{m\times n}X_{n\times k}=B_{m\times k}\\
&\Leftrightarrow A[X_1\quad X_2\quad \cdots\quad X_k]=[B_1\quad B_2\quad \cdots\quad B_k]\\
&\Leftrightarrow [AX_1\quad AX_2\quad \cdots\quad AX_k]=[B_1\quad B_2\quad \cdots\quad B_k]\\
&\Leftrightarrow AX_i=B_i,\quad i=1,\cdots,k
\end{aligned}$$

充分性：若$r(A)=r([A\quad B])$，则由

$$r(A)\le r([A\quad B_i])\le r([A\quad B])=r(A)$$

可得

$$r(\boldsymbol{A})=r([\boldsymbol{A}\quad \boldsymbol{B}_i])$$

$\Rightarrow$ $\boldsymbol{AX}_i=\boldsymbol{B}_i\ (i=1,\cdots,k)$ 有解，从而故 $\boldsymbol{AX}=\boldsymbol{B}$ 有解。

必要性：若 $\boldsymbol{AX}=\boldsymbol{B}$ 有解，设

$$\boldsymbol{X}=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

则

$$\boldsymbol{AX}=[\boldsymbol{A}_1\quad \boldsymbol{A}_2\quad \cdots\quad \boldsymbol{A}_n]\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$$=\boldsymbol{B}=[\boldsymbol{B}_1\quad \boldsymbol{B}_2\quad \cdots\quad \boldsymbol{B}_k]$$

因此有

$$\boldsymbol{B}_i=x_{1i}\boldsymbol{A}_1+x_{2i}\boldsymbol{A}_2+\cdots+x_{ni}\boldsymbol{A}_n,\quad i=1,\cdots,k$$

对矩阵 $[\boldsymbol{A}\quad \boldsymbol{B}]$ 作初等列变换 $c_{n+i}-x_{1i}c_1-x_{2i}c_2-\cdots-x_{ni}c_n,\quad i=1,\cdots,k$ ，则将 $[\boldsymbol{A}\quad \boldsymbol{B}]$ 的第 $n+1,\cdots,n+k$ 列都化为零向量，即 $[\boldsymbol{A}\quad \boldsymbol{B}]\sim[\boldsymbol{A}\quad \boldsymbol{O}]$，所以 $r(\boldsymbol{A})=r([\boldsymbol{A}\quad \boldsymbol{B}])$。

■

习题 3.7

1、 计算以下矩阵的秩，并求一个阶数最大的非零子式。

$$\begin{bmatrix} 2 & 1 & 8 & 3 & 7 \\ 2 & -3 & 0 & 7 & -5 \\ 3 & -2 & 5 & 8 & 0 \\ 1 & 0 & 3 & 2 & 0 \end{bmatrix}$$

2、 设 $\boldsymbol{A}$ 为 n 阶方阵，证明

$$r(\boldsymbol{A}^*)=\begin{cases} n & r(\boldsymbol{A})=n \\ 1 & r(\boldsymbol{A})=n-1 \\ 0 & r(\boldsymbol{A})<n-1 \end{cases}$$

3、 设 $\boldsymbol{A}$ 为 n 阶方阵，证明

（1）如果 $\boldsymbol{A}^2=\boldsymbol{E}$，则 $r(\boldsymbol{A}+\boldsymbol{E})+r(\boldsymbol{A}-\boldsymbol{E})=n$；

（2）如果 $\boldsymbol{A}^2=\boldsymbol{A}$，则 $r(\boldsymbol{A})+r(\boldsymbol{A}-\boldsymbol{E})=n$。

4、 证明：矩阵 $\boldsymbol{A}_{m\times n}$ 可表示为 $\boldsymbol{\alpha}_{m\times 1}$ 与 $\boldsymbol{\beta}_{1\times n}$ 的乘积，即 $\boldsymbol{A}_{m\times n}=\boldsymbol{\alpha}_{m\times 1}\boldsymbol{\beta}_{1\times n}$ 的充要条件是 $r(\boldsymbol{A})\le 1$。

5、 设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，证明：

（1）方程 $\boldsymbol{AX}=\boldsymbol{E}_m$ 有解的充要条件是 $r(\boldsymbol{A})=m$。

（2）方程 $\boldsymbol{YA}=\boldsymbol{E}_n$ 有解的充要条件是 $r(\boldsymbol{A})=n$。

6、 设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，证明：若 $\boldsymbol{AX}=\boldsymbol{AY}$，且 $r(\boldsymbol{A})=n$，则 $\boldsymbol{X}=\boldsymbol{Y}$。

7、 对于方程组 $\begin{cases}(2-\lambda)x_1+2x_2-2x_3=1\\ 2x_1+(5-\lambda)x_2-4x_3=2\\ -2x_1-4x_2+(5-\lambda)x_3=-\lambda-1\end{cases}$，问 λ 取何值时，此方程组有唯一解、无解、有无穷多解？并在有无穷多解时给出其解。

8、 设矩阵 $\begin{bmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{bmatrix}$ 是满秩的，则直线

$$L_1:\frac{x-a_3}{a_1-a_2}=\frac{y-b_3}{b_1-b_2}=\frac{z-c_3}{c_1-c_2}$$

与直线

$$L_2:\frac{x-a_1}{a_2-a_3}=\frac{y-b_1}{b_2-b_3}=\frac{z-c_1}{c_2-c_3}$$

的关系为

（A）相交于一点　　（B）重合　　（C）平行但不重合　　（D）异面

9、 设 $\boldsymbol{A}=\left[a_{ij}\right]_{m\times n}$，$\boldsymbol{Y}=\left[y_1,y_2,\cdots,y_n\right]^{\mathrm{T}}$，$\boldsymbol{b}=\left[b_1,b_2,\cdots,b_m\right]^{\mathrm{T}}$，$\boldsymbol{X}=\left[x_1,x_2,\cdots,x_m\right]^{\mathrm{T}}$，证明：方程组 $\boldsymbol{AY}=\boldsymbol{b}$ 有解的充要条件是方程组 $\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{b}^{\mathrm{T}}\end{bmatrix}\boldsymbol{X}=\begin{bmatrix}\boldsymbol{0}\\ 1\end{bmatrix}$ 无解。

注 1：并不是所有的线性方程组的增广矩阵都可化为的形式，例如

$$\begin{cases} x_1 \quad +x_2 & =2 \\ \qquad\qquad x_3 & =4 \\ \qquad\qquad\qquad x_4 & =2 \\ \qquad\qquad\qquad 0 & =0 \end{cases}$$

其增广矩阵为

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

它的左上角显然不是一个单位矩阵。但是如果我们交换未知数次序，

$$\begin{cases} x_1 \qquad\qquad\quad x_2 & =2 \\ \quad x_3 & =4 \\ \qquad\quad x_4 & =2 \\ \qquad\qquad\quad 0 & =0 \end{cases}$$

其增广矩阵为

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

这就是我们所讨论的标准形式。

第四章　n维向量空间

§4.1　向量的引入与推广

向量概念具有悠久的历史。在牛顿研究力的作用原理时，由于力具有方向、大小两个特性，因此他用有向线段来描述一个力：该线段的长度描述力的大小，而该线段的方向即为力的作用方向。以M_1为起点、M_2为终点的向量记作$\overrightarrow{M_1M_2}$；此时所定义的向量，只有大小与方向，向量的“值”与描述它的有向线段的起点没有关系，即与描述该向量的有向线段的起点无关，这种向量称为**自由向量**。描述向量的有向线段是可以移动的，只要保持该有向的大小、方向不变。两个向量$\overrightarrow{M_1M_2}$与$\overrightarrow{M_3M_4}$是方向相同的，如果

（1）　线段$\overrightarrow{M_1M_2}$与$\overrightarrow{M_3M_4}$平行；

（2）　将$\overrightarrow{M_3M_4}$平行移动，使得M_3与M_1重合后，M_1不在线段$\overrightarrow{M_2M_4}$的内部；

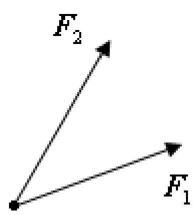
向量$\overrightarrow{M_1M_2}$与$\overrightarrow{M_3M_4}$是方向相同的，如果

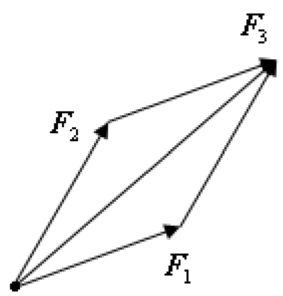
（1）　线段$\overrightarrow{M_1M_2}$与$\overrightarrow{M_3M_4}$平行；

（2）　将$\overrightarrow{M_3M_4}$平行移动，使得M_3与M_1重合后，M_1在线段$\overrightarrow{M_2M_4}$的内部；

大小为零的向量称为零向量，记作$\mathbf{0}$，零向量没有方向，或者说零向量可以取任意方向。如果两个向量的大小、方向均相同，则这两个向量相等。由于力可用向量表示，我们从此对力与向量不加区分。向量$\boldsymbol{\alpha}$的大小（表示向量的有向线段的长度）记为$\|\boldsymbol{\alpha}\|$。用$-\boldsymbol{\alpha}$表示大小与$\boldsymbol{\alpha}$相同、而方向与$\boldsymbol{\alpha}$的向量。

通过实验，牛顿得出了合力定律：作用到一个质点上的两个力$\boldsymbol{F}_1,\boldsymbol{F}_2$的效果等价于一个力$\boldsymbol{F}_3$作用于该质点，$\boldsymbol{F}_3$由$\boldsymbol{F}_1,\boldsymbol{F}_2$根据平行四边形法则确定，我们把$\boldsymbol{F}_3$称为$\boldsymbol{F}_1,\boldsymbol{F}_2$的**合力**，记作$\boldsymbol{F}_3=\boldsymbol{F}_1+\boldsymbol{F}_2$。基于该物理规律，平行四边形法则就成为了向量运算的基本法则。





另外，定义一个数k乘以一个向量$\boldsymbol{F}$，得到向量$k\boldsymbol{F}$：即为该向量的大小（长度）为原来的k倍，如果$k>0$，则$k\boldsymbol{F}$与$\boldsymbol{F}$方向相同，否则方向相反。向量的加法、数乘统称为向量的**线性运算**。

在上述定义下，通过几何关系可以验证向量运算具有如下性质：

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

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

（3） $\boldsymbol{\alpha}+\mathbf{0}=\boldsymbol{\alpha}$；

（4） $\boldsymbol{\alpha}+(-\boldsymbol{\alpha})=\mathbf{0}$，称$-\boldsymbol{\alpha}$为$\boldsymbol{\alpha}$的负元素；

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

（6） $k(l\boldsymbol{\alpha})=(kl)\boldsymbol{\alpha}$；

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

（8） $k(\boldsymbol{\alpha}+\boldsymbol{\beta})=k\boldsymbol{\alpha}+k\boldsymbol{\beta}$。

验证上述性质的过程从略，有兴趣的读者可参考相关教材，例如科学出版社的《线性代数与解析几何》，冯良贵等编著，ISBN978-7-03-020081-5.

在经典力学中，有“功”的概念。一个常力$\boldsymbol{F}$沿直线把一物体从M_1移动到M_2，记$\boldsymbol{S}=\overrightarrow{M_1M_2}$，$\boldsymbol{F}$与$\boldsymbol{S}$之间的夹角为$\theta(0\le\theta\le\pi)$，则力$\boldsymbol{F}$对位移$\boldsymbol{S}$所做的功为

$$W=\|\boldsymbol{F}\|\ \|\boldsymbol{S}\|\cos\theta$$

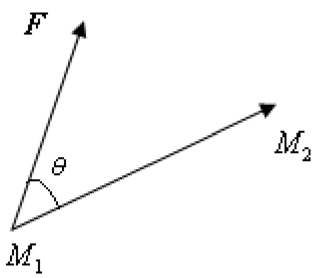


图 4.1　功的定义

可从该物理规律定义向量的**内积**。

平移两个向量$\boldsymbol{\alpha},\boldsymbol{\beta}$，使它们的起点重合，此时这两条有向线段所构成的不大于π的夹角，称为向量$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$之间的夹角，记作$\angle(\boldsymbol{\alpha},\boldsymbol{\beta})$。如果$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$中有一个是零向量，则规定$\angle(\boldsymbol{\alpha},\boldsymbol{\beta})$可为$[0,\pi)$中的任何值。

两个向量$\boldsymbol{\alpha},\boldsymbol{\beta}$的点积$\boldsymbol{\alpha}\bullet\boldsymbol{\beta}$（也称为内积）定义为

$$\boldsymbol{\alpha}\bullet\boldsymbol{\beta}=\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|\cos\angle(\boldsymbol{\alpha},\boldsymbol{\beta}) \tag{4.1}$$

向量的点积满足如下运算规律（$\boldsymbol{\alpha},\boldsymbol{\beta}$ 为向量，k 为任意实数）：

（1）　交换律：$\boldsymbol{\alpha}\bullet\boldsymbol{\beta}=\boldsymbol{\beta}\bullet\boldsymbol{\alpha}$；

（2）　分配律：$\boldsymbol{\alpha}\bullet(\boldsymbol{\beta}+\boldsymbol{\gamma})=\boldsymbol{\alpha}\bullet\boldsymbol{\beta}+\boldsymbol{\alpha}\bullet\boldsymbol{\gamma}$；

（3）　齐次性：$k(\boldsymbol{\alpha}\bullet\boldsymbol{\beta})=(k\boldsymbol{\alpha})\bullet\boldsymbol{\beta}=\boldsymbol{\alpha}\bullet(k\boldsymbol{\beta})$；

（4）　非负性：当 $\boldsymbol{\alpha}\neq\boldsymbol{0}$ 时，$\boldsymbol{\alpha}\bullet\boldsymbol{\alpha}\geq 0$，$\boldsymbol{\alpha}\bullet\boldsymbol{\alpha}=0\Leftrightarrow\boldsymbol{\alpha}=\boldsymbol{0}$；

（5）　Cauchy-Schwarz 不等式：$|\boldsymbol{\alpha}\bullet\boldsymbol{\beta}|\leq\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|$，其中等号当且仅当 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 共线时成立。

向量点积的上述性质的验证过程从略。

从向量点积的定义可知如下简单性质：

（1）如果两个向量垂直（这两个向量之间的夹角为 $\frac{\pi}{2}$），则这两个向量的点积为零；

（2）如果向量 $\boldsymbol{\alpha}$ 为非零向量，则 $\boldsymbol{\alpha}\bullet\boldsymbol{\alpha}=\|\boldsymbol{\alpha}\|^2>0$。

在物理中还有力矩的概念。以 O 为杠杆的固定点，力 $\boldsymbol{F}$ 作用在杠杆上的 P 点，$\boldsymbol{F}$ 与 $\overrightarrow{OP}$ 的夹角为 θ，力 $\boldsymbol{F}$ 对 O 点的力矩是一个向量 $\boldsymbol{M}$，大小为 $\|\boldsymbol{M}\|=\|\boldsymbol{F}\|\bullet\|\overrightarrow{OP}\|\sin\theta$，$\boldsymbol{M}$ 的方向垂直于 $\boldsymbol{F}$ 与 $\overrightarrow{OP}$ 决定的平面（即 $\boldsymbol{M}$ 同时垂直于 $\boldsymbol{F}$ 与 $\boldsymbol{r}$）、且 $\overrightarrow{OP},\boldsymbol{F},\boldsymbol{M}$ 成右手系。我们把力矩 $\boldsymbol{M}$ 定义 $\overrightarrow{OP}$ 与 $\boldsymbol{F}$ 的叉积，记作 $\boldsymbol{M}=\overrightarrow{OP}\times\boldsymbol{F}$。

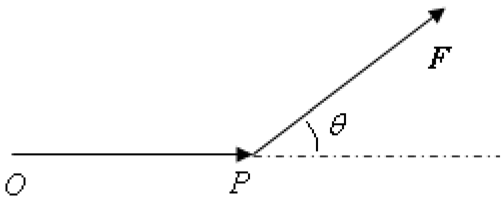


图 4.2　力矩的定义

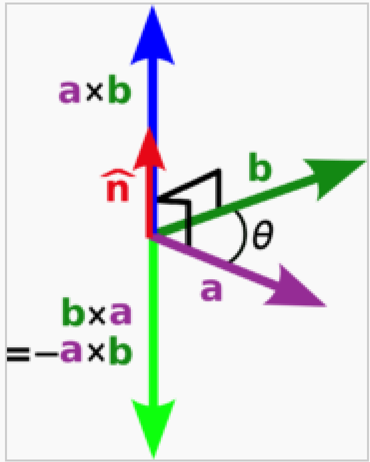


图 4.3　叉积的定义

到目前为止，向量的定义、运算是通过几何方式来给出的，其运算规律都是通过几何方式来证明的。显然，这种方式并不利于我们的应用。

笛卡尔引入坐标系的概念是数学史上非常重要的事件，极大地促进了数学的发展。

通过在空间的一点 O 引三条不共面的数轴，这三条数轴称为**坐标轴**，分别为 $\boldsymbol{Ox}$ 轴（ $\boldsymbol{x}$ 轴，横轴）、$\boldsymbol{Oy}$ 轴（ $\boldsymbol{y}$ 轴，纵轴）、$\boldsymbol{Oz}$ 轴（ $\boldsymbol{z}$ 轴，竖轴），这三条坐标轴的原点都是 O 点；我们把满足上述条件的三条数轴、点 O 视作一个整体，称为**坐标系**，记作 $\boldsymbol{Oxyz}$ 。通常，我们要求该坐标系为**右手系**——以右手握住 z 轴，大拇指指向 z 轴正向，其余四指开始指向 $\boldsymbol{x}$ 轴正向，然后将四指向 $\boldsymbol{y}$ 轴正向，在该过程中四指所转角度不超过 π ）。类似地可以定义左手系——以右手握住 z 轴，大拇指指向 z 轴正向，其余四指开始指向 $\boldsymbol{y}$ 轴正向，然后将四指向 $\boldsymbol{x}$ 轴正向，在该过程中四指所转角度不超过 π ）。本文中的坐标系均采用右手系。

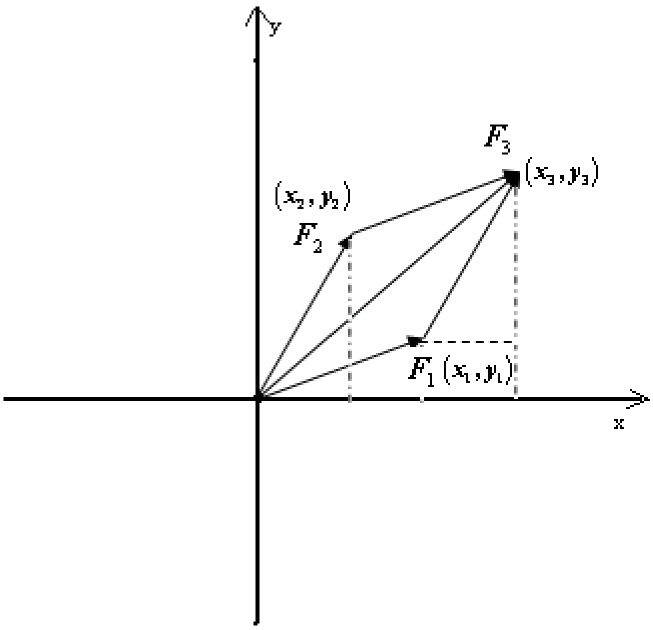
每两条坐标轴确定的平面称为**坐标平面**， $\boldsymbol{x}$ 轴与 $\boldsymbol{y}$ 轴确定的平面称为 $\boldsymbol{Oxy}$ 平面， $\boldsymbol{x}$ 轴与 $\boldsymbol{z}$ 轴确定的平面称为 $\boldsymbol{Oxz}$ 平面， $\boldsymbol{y}$ 轴与 $\boldsymbol{z}$ 轴确定的平面称为 $\boldsymbol{Oyz}$ 平面。

在取定坐标系 $\boldsymbol{Oxyz}$ 后，对于空间中的任意一点 M ，可作过 M 、且与 $\boldsymbol{Oxy}$ 平行的平面，该平面必与 $\boldsymbol{z}$ 轴相交，交点在 $\boldsymbol{z}$ 轴上的坐标 z 称为 M 的 $\boldsymbol{z}$ 坐标（竖坐标）；类似地，作过 M 、且与 $\boldsymbol{Oxz}$ 平行的平面，该平面必与 $\boldsymbol{y}$ 轴相交，交点在 $\boldsymbol{y}$ 轴上的坐标 y 称为 M 的 $\boldsymbol{y}$ 坐标（纵坐标）；作过 M 、且与 $\boldsymbol{Oyz}$ 平行的平面，该平面必与 $\boldsymbol{x}$ 轴相交，交点在 $\boldsymbol{x}$ 轴上的坐标 x 称为 M 的 $\boldsymbol{x}$ 坐标（横坐标）。而把有序数组 (x,y,z) 称为 M （在坐标系坐标系 $\boldsymbol{Oxyz}$ 下）的**坐标**。反之，对于给定的有序数组 (x,y,z)，过 $\boldsymbol{x}$ 轴上坐标为 x 的点作与 $\boldsymbol{Oyz}$ 平行的平面、 $\boldsymbol{y}$ 轴上坐标为 y 的点作与 $\boldsymbol{Oxz}$ 平行的平面、 $\boldsymbol{z}$ 轴上坐标为 z 的点作与 $\boldsymbol{Oxy}$ 平行的平面，这三个平面有唯一的交点，该交点的坐标就是 (x,y,z)。因此，在建立了坐标系之后，空间中

的点与有序数组构成一一对应关系。

有了坐标系后，可以将向量与坐标系联系起来。为了描述向量（自由向量），我们将自由向量的起点（表示向量的有向线段的起点）置于坐标系的原点（以坐标系原点为起点的向量称为**径矢**），那么该向量就由径矢的终点唯一地确定，这样点与向量建立了一一对应的关系，而点由其坐标唯一确定，所以这样向量就与点的坐标——三元有序数组建立了一一对应关系。对于一个向量$\boldsymbol{\alpha}$，将其起点平移到坐标原点，若其终点的坐标为(x,y,z)，我们就把该向量记为$\boldsymbol{\alpha}(x,y,z)$，或$\boldsymbol{\alpha}=(x,y,z)$，也记作$(x,y,z)$。下图就是这种对应关系的脉络。

	平移向量，将起点置于坐标原点		径矢与其终点一一对应			
（自由）向量	$\longleftrightarrow$	径矢	$\longleftrightarrow$	终点	$\longleftrightarrow$ 坐标系	终点坐标(x,y,z)



通过简单的几何推理可知，对于两个向量$\boldsymbol{F}_1(x_1,y_1,z_1),\boldsymbol{F}_2(x_2,y_2,z_2)$，根据平行四边形法则所确定的$\boldsymbol{F}_3=\boldsymbol{F}_1+\boldsymbol{F}_2$，其终点的坐标为$(x_1+x_2,y_1+y_2,z_1+z_2)$，也就是说，向量加法就是这两个向量的终点的每个坐标分别相加。易知若$\boldsymbol{F}=(x,y,z)$，则$k\boldsymbol{F}=(kx,ky,kz)$。因此向量的运算就简化为其终点坐标的运算。

通过引入坐标系，可以比较完美地解决向量线性运算的描述问题。为了便于描述向量点积、叉积，我们引入一种特殊的坐标系——直角坐标系。

前文中引入坐标系时，仅要求三条坐标轴不共面。现在进一步要求坐标系的这三条坐标轴相互垂直，所得到的坐标系就为直角坐标系。在空间直角坐标系的$\boldsymbol{x}$轴、$\boldsymbol{y}$轴、$\boldsymbol{z}$轴上分别取单位向量，依次记为$\vec{\boldsymbol{i}},\vec{\boldsymbol{j}},\vec{\boldsymbol{k}}$（称为基本向量），则终点坐标为$\boldsymbol{F}(x,y,z)$的向量可表示为$\boldsymbol{F}=x\vec{\boldsymbol{i}}+y\vec{\boldsymbol{j}}+z\vec{\boldsymbol{k}}$；根据勾股定理易知，向量$\boldsymbol{\alpha}(x,y,z)$的长度为$\sqrt{x^2+y^2+z^2}$，即

$\|\boldsymbol{\alpha}\| = \sqrt{x^2+y^2+z^2}$ 。

空间直角坐标系的好处是能够很好地解决向量点积、叉积的计算，这是直角坐标系的优势。根据直角坐标系的定义、点积与叉积的定义，我们有

$$\vec{\boldsymbol{i}}\bullet\vec{\boldsymbol{j}} = \vec{\boldsymbol{i}}\bullet\vec{\boldsymbol{k}} = \vec{\boldsymbol{j}}\bullet\vec{\boldsymbol{k}} = 0$$
$$\vec{\boldsymbol{i}}\bullet\vec{\boldsymbol{i}} = \vec{\boldsymbol{j}}\bullet\vec{\boldsymbol{j}} = \vec{\boldsymbol{k}}\bullet\vec{\boldsymbol{k}} = 1$$

$$\begin{aligned}
&\vec{\boldsymbol{i}}\times\vec{\boldsymbol{j}} = \vec{\boldsymbol{k}}\\
&\vec{\boldsymbol{j}}\times\vec{\boldsymbol{k}} = \vec{\boldsymbol{i}}\\
&\vec{\boldsymbol{k}}\times\vec{\boldsymbol{i}} = \vec{\boldsymbol{j}}\\
&\vec{\boldsymbol{i}}\times\vec{\boldsymbol{i}} = \vec{\boldsymbol{j}}\times\vec{\boldsymbol{j}} = \vec{\boldsymbol{k}}\times\vec{\boldsymbol{k}} = \mathbf{0}
\end{aligned}$$

因此对于两个向量$\boldsymbol{\alpha}(x_1,y_1,z_1),\boldsymbol{\beta}(x_2,y_2,z_2)$，其点积为

$$\begin{aligned}
\boldsymbol{\alpha}\bullet\boldsymbol{\beta} &= \left(x_1\vec{\boldsymbol{i}}+y_1\vec{\boldsymbol{j}}+z_1\vec{\boldsymbol{k}}\right)\bullet\left(x_2\vec{\boldsymbol{i}}+y_2\vec{\boldsymbol{j}}+z_2\vec{\boldsymbol{k}}\right)\\
&= x_1x_2\vec{\boldsymbol{i}}\bullet\vec{\boldsymbol{i}}+x_1y_2\vec{\boldsymbol{i}}\bullet\vec{\boldsymbol{j}}+x_1z_2\vec{\boldsymbol{i}}\bullet\vec{\boldsymbol{k}}\\
&\quad+y_1x_2\vec{\boldsymbol{j}}\bullet\vec{\boldsymbol{i}}+y_1y_2\vec{\boldsymbol{j}}\bullet\vec{\boldsymbol{j}}+y_1z_2\vec{\boldsymbol{j}}\bullet\vec{\boldsymbol{k}}\\
&\quad+z_1x_2\vec{\boldsymbol{k}}\bullet\vec{\boldsymbol{i}}+z_1y_2\vec{\boldsymbol{k}}\bullet\vec{\boldsymbol{j}}+z_1z_2\vec{\boldsymbol{k}}\bullet\vec{\boldsymbol{k}}\\
&= x_1x_2+y_1y_2+z_1z_2
\end{aligned}$$

因此两个向量之间的夹角很容易计算。由于$\boldsymbol{\alpha}\bullet\boldsymbol{\beta} = \|\boldsymbol{\alpha}\|\ \|\boldsymbol{\beta}\|\cos\angle(\boldsymbol{\alpha},\boldsymbol{\beta})$，故当$\boldsymbol{\alpha},\boldsymbol{\beta}$均不为零向量时，有

$$\cos\angle(\boldsymbol{\alpha},\boldsymbol{\beta}) = \frac{\boldsymbol{\alpha}\bullet\boldsymbol{\beta}}{\|\boldsymbol{\alpha}\|\ \|\boldsymbol{\beta}\|} = \frac{x_1x_2+y_1y_2+z_1z_2}{\sqrt{x_1^2+y_1^2+z_1^2}\sqrt{x_2^2+y_2^2+z_2^2}}$$

对于向量$\boldsymbol{\alpha}(x_1,y_1,z_1),\boldsymbol{\beta}(x_2,y_2,z_2)$的叉积，有

$$\begin{aligned}
\boldsymbol{\alpha}\times\boldsymbol{\beta} &= \left(x_1\vec{\boldsymbol{i}}+y_1\vec{\boldsymbol{j}}+z_1\vec{\boldsymbol{k}}\right)\times\left(x_2\vec{\boldsymbol{i}}+y_2\vec{\boldsymbol{j}}+z_2\vec{\boldsymbol{k}}\right)\\
&= x_1x_2\vec{\boldsymbol{i}}\times\vec{\boldsymbol{i}}+x_1y_2\vec{\boldsymbol{i}}\times\vec{\boldsymbol{j}}+x_1z_2\vec{\boldsymbol{i}}\times\vec{\boldsymbol{k}}\\
&\quad+y_1x_2\vec{\boldsymbol{j}}\times\vec{\boldsymbol{i}}+y_1y_2\vec{\boldsymbol{j}}\times\vec{\boldsymbol{j}}+y_1z_2\vec{\boldsymbol{j}}\times\vec{\boldsymbol{k}}\\
&\quad+z_1x_2\vec{\boldsymbol{k}}\times\vec{\boldsymbol{i}}+z_1y_2\vec{\boldsymbol{k}}\times\vec{\boldsymbol{j}}+z_1z_2\vec{\boldsymbol{k}}\times\vec{\boldsymbol{k}}\\
&= x_1y_2\vec{\boldsymbol{k}}-x_1z_2\vec{\boldsymbol{j}}-y_1x_2\vec{\boldsymbol{k}}+y_1z_2\vec{\boldsymbol{i}}+z_1x_2\vec{\boldsymbol{j}}-z_1y_2\vec{\boldsymbol{i}}\\
&= (y_1z_2-y_2z_1)\vec{\boldsymbol{i}}+(x_2z_1-x_1z_2)\vec{\boldsymbol{j}}+(x_1y_2-x_2y_1)\vec{\boldsymbol{k}}
\end{aligned}$$

由上面的关于向量的历史进程可知，向量可用有序数组（径矢终点的坐标）来表示，其运算转化为有序数组（终点坐标）的相应运算。因此向量与三元有序数组等价，为此我们将向量推广到一般的情形，这就是我们将要介绍的n维向量。

§4.2　n维向量的定义与运算

定义 4.1　设$a_1,a_2,\cdots,a_n$为数域$\mathbb{F}$中的n个数，它们组成的有序数组

$$\boldsymbol{\alpha}=[a_1,a_2,\cdots,a_n] \quad \text{或} \quad \boldsymbol{\alpha}=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

称为n维向量，a_i称为$\boldsymbol{\alpha}$的第i个分量或坐标。$\mathbb{F}=\mathbb{R}$时，$\boldsymbol{\alpha}$称为n维实向量；$\mathbb{F}=\mathbb{C}$时，$\boldsymbol{\alpha}$称为n维复向量。

$[a_1,a_2,\cdots,a_n]$称为n维**行向量**，$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$称为$n$维**列向量**，$n$称为该向量的**维数**。

实际上，n维行向量是$1\times n$矩阵，而n维列向量是$n\times 1$矩阵，并且$\boldsymbol{\alpha}=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$可用矩阵的转置运算表示为$\boldsymbol{\alpha}=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}=[a_1,a_2,\cdots,a_n]^{\mathrm{T}}$。在本书以后的论述中，若没有特别说明，$n$维向量均指$n$维列向量，通常用$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\cdots$来表示，而行向量则表示为列向量的转置，即表示为$\boldsymbol{\alpha}^{\mathrm{T}},\boldsymbol{\beta}^{\mathrm{T}},\boldsymbol{\gamma}^{\mathrm{T}},\cdots$。

令$\mathbb{F}^n=\left\{[a_1,a_2,\cdots,a_n]^{\mathrm{T}} \,\middle|\, a_1,a_2,\cdots,a_n\in\mathbb{F}\right\}$表示$\mathbb{F}$上所有$n$维向量所构成的集合；特别地，$\mathbb{R}^n=\left\{[a_1,a_2,\cdots,a_n]^{\mathrm{T}} \,\middle|\, a_1,a_2,\cdots,a_n\in\mathbb{R}\right\}$表示所有$n$维实向量所构成的集合。

定义 4.2　向量的加法与数乘

设$\boldsymbol{\alpha}=[a_1,a_2,\cdots,a_n]^{\mathrm{T}}$，$\boldsymbol{\beta}=[b_1,b_2,\cdots,b_n]^{\mathrm{T}}$，$k\in\mathbb{R}$，则向量的加法$\boldsymbol{\alpha}+\boldsymbol{\beta}$，数乘向量$k\boldsymbol{\alpha}$分别定义为

$$\boldsymbol{\alpha}+\boldsymbol{\beta}=[a_1+b_1,a_2+b_2,\cdots,a_n+b_n]^{\mathrm{T}}$$

$$k\boldsymbol{\alpha}=\left[ka_1,ka_2,\cdots,ka_n\right]^{\mathrm{T}}$$

向量的加法、数乘统称为向量的**线性运算**。

根据该定义可知，n 维向量的线性运算实际上是向量作为 $n\times1$ 矩阵的线性运算。

对于 $\boldsymbol{\alpha}=\left[a_1,a_2,\cdots,a_n\right]^{\mathrm{T}}$，根据定义有

$$(-1)\boldsymbol{\alpha}=\left[-a_1,-a_2,\cdots,-a_n\right]^{\mathrm{T}}$$

我们记

$$-\boldsymbol{\alpha}=(-1)\boldsymbol{\alpha}$$

所以 $-\boldsymbol{\alpha}+\boldsymbol{\alpha}=\left[0,0,\cdots,0\right]^{\mathrm{T}}\overset{\Delta}{=}\mathbf{0}$，其中 $\mathbf{0}=\left[0,0,\cdots,0\right]^{\mathrm{T}}$ 为**零向量**。

为了便于以后的应用，我们把向量线性运算的性质列举如下：

$\forall\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathbb{F}^n;k,l\in\mathbb{F}$，有

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

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

（3） $\boldsymbol{\alpha}+\mathbf{0}=\boldsymbol{\alpha}$；

（4） $\boldsymbol{\alpha}+(-\boldsymbol{\alpha})=\mathbf{0}$，称 $-\boldsymbol{\alpha}$ 为 $\boldsymbol{\alpha}$ 的负元素；

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

（6） $k(l\boldsymbol{\alpha})=(kl)\boldsymbol{\alpha}$；

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

（8） $k(\boldsymbol{\alpha}+\boldsymbol{\beta})=k\boldsymbol{\alpha}+k\boldsymbol{\beta}$；

定义 4.3　　向量组

若干个同维向量组成的集合称为向量组。

例如，$\begin{bmatrix}2\\-5\\3\\0\end{bmatrix},\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$是一个由五个四维列向量所构成的向量组。

我们称$\boldsymbol{e}_1=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\boldsymbol{e}_2=\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix},\cdots,\boldsymbol{e}_n=\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$为$\mathbb{R}^n$的**基本向量组**。

向量组与矩阵有密切的关系。考虑矩阵

$$\boldsymbol{A}_{m\times n}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}$$

它的每一列元素构成一个m维列向量，因此该矩阵有n个列向量

$$\boldsymbol{\alpha}_1=\begin{bmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{bmatrix},\boldsymbol{\alpha}_2=\begin{bmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{bmatrix},\cdots,\boldsymbol{\alpha}_n=\begin{bmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{bmatrix}$$

这n个列向量构成向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$，在这种意义上，该矩阵可看作是由一个向量组所构成的，即有

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}$$

这实际上是对矩阵的按列分块。我们把上述向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$称作矩阵$\boldsymbol{A}$的列向量组。

类似地，矩阵的每一行元素构成一个行向量，因此该矩阵有m个行向量

$$\begin{aligned}\boldsymbol{\beta}_1^{\mathrm{T}}&=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\end{bmatrix}=\begin{bmatrix}a_{11},a_{12},\cdots,a_{1n}\end{bmatrix}\\ \boldsymbol{\beta}_2^{\mathrm{T}}&=\begin{bmatrix}a_{21} & a_{22} & \cdots & a_{2n}\end{bmatrix}=\begin{bmatrix}a_{21},a_{22},\cdots,a_{2n}\end{bmatrix}\\ &\vdots\\ \boldsymbol{\beta}_m^{\mathrm{T}}&=\begin{bmatrix}a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}=\begin{bmatrix}a_{m1},a_{m2},\cdots,a_{mn}\end{bmatrix}\end{aligned}$$

这m个行向量构成一个向量组$\boldsymbol{\beta}_1^{\mathrm{T}},\boldsymbol{\beta}_2^{\mathrm{T}},\cdots,\boldsymbol{\beta}_m^{\mathrm{T}}$，此时矩阵可写作

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\beta}_1^{\mathrm{T}}\\\boldsymbol{\beta}_2^{\mathrm{T}}\\\vdots\\\boldsymbol{\beta}_m^{\mathrm{T}}\end{bmatrix}$$

这实际上是将矩阵按行分块。我们把$\boldsymbol{\beta}_1^{\mathrm{T}},\boldsymbol{\beta}_2^{\mathrm{T}},\cdots,\boldsymbol{\beta}_m^{\mathrm{T}}$称作矩阵$\boldsymbol{A}$的行向量组。

反过来，对于向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$，我们可用它构成矩阵

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}$$

使得$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$为$\boldsymbol{A}$的列向量组；对于行向量组$\boldsymbol{\beta}_1^{\mathrm{T}},\boldsymbol{\beta}_2^{\mathrm{T}},\cdots,\boldsymbol{\beta}_m^{\mathrm{T}}$，我们也可构成矩阵

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\beta}_1^{\mathrm{T}}\\ \boldsymbol{\beta}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\beta}_m^{\mathrm{T}}\end{bmatrix}$$

此时$\boldsymbol{\beta}_1^{\mathrm{T}},\boldsymbol{\beta}_2^{\mathrm{T}},\cdots,\boldsymbol{\beta}_m^{\mathrm{T}}$为$\boldsymbol{A}$的行向量组。

定义 4.4　　**线性组合、线性表示**

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\in\mathbb{F}^n$为一个向量组，$\boldsymbol{\beta}\in\mathbb{F}^n$，若存在$k_1,k_2,\cdots,k_m\in\mathbb{F}$，使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m \tag{4.1}$$

则称$\boldsymbol{\beta}$为向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的**线性组合**，或称$\boldsymbol{\beta}$可由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$**线性表示**，$k_1,k_2,\cdots,k_m$为$\boldsymbol{\beta}$在该向量组下的**组合系数**。

借用分块矩阵乘法，$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m$上可表示成

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_m\end{bmatrix}\begin{bmatrix}k_1\\ k_2\\ \vdots\\ k_m\end{bmatrix} \tag{4.2}$$

例 4.1　给定向量组$\boldsymbol{e}_1=\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\boldsymbol{e}_2=\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\boldsymbol{e}_3=\begin{bmatrix}0\\0\\1\\0\end{bmatrix},\boldsymbol{e}_4=\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$与向量$\boldsymbol{\beta}=\begin{bmatrix}2\\-5\\3\\0\end{bmatrix}$可知

$$\boldsymbol{\beta}=2\boldsymbol{e}_1-5\boldsymbol{e}_2+3\boldsymbol{e}_3+0\boldsymbol{e}_4$$

所以$\boldsymbol{\beta}$是向量组$\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4$的线性组合，而$2,-5,3,0$是$\boldsymbol{\beta}$在向量组$\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4$下的组合系数。

问题 4.1　如何判断$\boldsymbol{\beta}$是否是$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的线性组合？我们先考虑一个具体例子。

例4.2　判断 $\boldsymbol{\beta}=[2,-4,-3,-7]^{\mathrm{T}}$ 是否可由向量组

$$\boldsymbol{\alpha}_1=[1,1,1,2]^{\mathrm{T}},\boldsymbol{\alpha}_2=[0,2,1,3]^{\mathrm{T}},\boldsymbol{\alpha}_3=[3,1,0,1]^{\mathrm{T}}$$

线性表示。如果能，则求一组组合系数。

解：若有 k_1,k_2,k_3，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3=\boldsymbol{\beta}, \qquad 即 \quad [\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\boldsymbol{\alpha}_3]\begin{bmatrix}k_1\\k_2\\k_3\end{bmatrix}=\boldsymbol{\beta}$$

这说明 $[k_1,k_2,k_3]^{\mathrm{T}}$ 是线性方程组

$$[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\boldsymbol{\alpha}_3]\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\boldsymbol{\beta}$$

的解。因此考虑该线性方程组的增广矩阵

$$[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\boldsymbol{\alpha}_3\quad\boldsymbol{\beta}]=\begin{bmatrix}1&0&3&2\\1&2&1&-4\\1&1&0&-3\\2&3&1&-7\end{bmatrix}$$

将其化为行阶梯形得

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

由此可知该线性方程组有解，故 $\boldsymbol{\beta}$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，为求组合系数，我们将其化为行标准形，得

$$\begin{bmatrix}1&0&0&-1\\0&1&0&-2\\0&0&1&1\\0&0&0&0\end{bmatrix}$$

所以 $k_1=-1,k_2=-2,k_3=1$，即 $\boldsymbol{\beta}=-\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$。

实际上，借助于线性方程组解的理论以及线性方程组的分块，很显然有如下结论：

定理 4.1　　$\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的线性组合 $\Leftrightarrow$ 存在 $k_1,k_2,\cdots,k_m\in\mathbb{F}$，使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m$$

$\Leftrightarrow$线性方程组$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_m\end{bmatrix}\begin{bmatrix}k_1\\k_2\\\vdots\\k_m\end{bmatrix}=\boldsymbol{\beta}$有解；

$\Leftrightarrow r\left(\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_m\end{bmatrix}\right)=r\left(\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_m & \boldsymbol{\beta}\end{bmatrix}\right)$。

因此，该问题就转化为求矩阵秩的问题。

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\in\mathbb{F}^n$，则集合

$$\begin{aligned}&L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}=span\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}\\&=\{\boldsymbol{\alpha}=c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_m\boldsymbol{\alpha}_m|c_1,c_2,\cdots,c_m\in\mathbb{F}\}\end{aligned}$$

为所有可由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性表示的向量所构成的集合。

例如，所有n维实向量的集合$\mathbb{R}^n=L\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$。

习题 4.2

1. 有甲、乙、丙三种规格的钢条，已知甲种 2 根、乙种 1 根、丙种 3 根共长 23 米；甲种 1 根、乙种 4 根、丙种 5 根共长 36 米。问：甲种 1 根、乙种 2 根、丙种 3 根共长多少？
2. 已知 $\boldsymbol{\alpha}_1=[1,0,2,3]^{\mathrm{T}},\boldsymbol{\alpha}_2=[1,1,3,5]^{\mathrm{T}},\boldsymbol{\alpha}_3=[1,-1,a+2,1]^{\mathrm{T}},\boldsymbol{\alpha}_4=[1,2,4,a+8]^{\mathrm{T}}$，

   $\boldsymbol{\beta}=[1,1,b+3,5]^{\mathrm{T}}$，

   （1）a,b为何值时，$\boldsymbol{\beta}$不能表示为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$的线性组合；

   （2）a,b为何值时，$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$唯一地线性表示？并写出该表达式。

§4.3　线性表示的最简性

这一节的内容，我们将围绕一个问题逐步展开讨论，在解决这个问题的过程中导出相应的概念与定义。首先我们介绍一个例子。

设

$$\boldsymbol{\alpha}_1=[1,\ 1,\ 2,\ 2,\ 1]^{\mathrm{T}},\qquad \boldsymbol{\alpha}_2=[0,\ 2,\ 1,\ 5,\ -1]^{\mathrm{T}}$$

$$\boldsymbol{\alpha}_3=[2,\ 0,\ 3,\ -1,\ 3]^{\mathrm{T}},\qquad \boldsymbol{\beta}=[1,\ -3,\ 0,\ -8,\ 3]^{\mathrm{T}}$$

判断$\boldsymbol{\beta}$是否可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表示?

容易验证

$$\boldsymbol{\beta}=-\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$$

$$\boldsymbol{\beta}=3\boldsymbol{\alpha}_1-3\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$$

因此，$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表示。实际上还有

$$\boldsymbol{\beta}=\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2$$

$$\boldsymbol{\beta}=-3\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_3$$

$$\boldsymbol{\beta}=-\frac{3}{2}\boldsymbol{\alpha}_2+\frac{1}{2}\boldsymbol{\alpha}_3$$

即$\boldsymbol{\beta}$可由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$线性表示，也可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3$线性表示，还可由$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表示。因此，如果只是需要线性表示$\boldsymbol{\beta}$，则不需要同时使用$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$这三个向量，而只需要用其中的两个就可以达到线性表示$\boldsymbol{\beta}$的目的；向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$有三个向量，而$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$、$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3$、$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$、这三个向量组都只含有两个向量，从用来线性表示$\boldsymbol{\beta}$的向量组所含向量的个数来讲，后面三种表示方式更为简单——因为向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$、$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3$、$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$所含向量个数$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$所含向量的个数更少。

为讨论上述问题的一般情形，先引入部分组的概念。

定义 4.5　部分组

由向量组中的一部分向量所构成的向量组称为该向量组的**部分组**。

从上述这个例子，我们可讨论如下一般的问题：

问题 4.0　若向量$\boldsymbol{\beta}$可由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性表示，是否存在$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的部分组$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\,(r<m)$，使得$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$线性表示？如果存在这样的部分组，如何找出它？进一步地，若$\boldsymbol{\beta}$可由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性表示，能否从该向量组中找出**所含向量个数尽可能少**的部分组来线性表示$\boldsymbol{\beta}$？

为说明上述问题，再看一个例子。

例 4.3　考虑由 $\boldsymbol{\alpha}_1=\begin{bmatrix}1\\0\end{bmatrix}$，$\boldsymbol{\alpha}_2=\begin{bmatrix}0\\1\end{bmatrix}$ 构成的向量组，对于两个向量 $\begin{bmatrix}2\\0\end{bmatrix}$，$\begin{bmatrix}1\\1\end{bmatrix}$，易知有

$$\begin{bmatrix}2\\0\end{bmatrix}=2\boldsymbol{\alpha}_1$$

$$\begin{bmatrix}1\\1\end{bmatrix}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$$

因此 $\begin{bmatrix}1\\1\end{bmatrix}$ 可由向量组 $\boldsymbol{\alpha}_1$，$\boldsymbol{\alpha}_2$ 线性表示，且容易知道 $\begin{bmatrix}1\\1\end{bmatrix}$ 不可能仅由 $\boldsymbol{\alpha}_1$ 来线性表示，也不能仅由 $\boldsymbol{\alpha}_2$ 线性表示，即：$\begin{bmatrix}1\\1\end{bmatrix}$ 必须用 $\boldsymbol{\alpha}_1$，$\boldsymbol{\alpha}_2$ 才能线性表示，$\boldsymbol{\alpha}_1$，$\boldsymbol{\alpha}_2$ 的任意一个部分组都不能线性表示 $\begin{bmatrix}1\\1\end{bmatrix}$。而向量 $\begin{bmatrix}2\\0\end{bmatrix}$ 就只需要用 $\boldsymbol{\alpha}_1$，$\boldsymbol{\alpha}_2$ 的部分组 $\boldsymbol{\alpha}_1$ 线性表示。

上述这个例子说明：对于同一个向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$，如果所取的 $\boldsymbol{\beta}$ 不同，问题 4.0 就可能有不同的答案。为了得到普遍意义上的结果，我们希望去除这一因素的影响，而仅仅考虑向量组本身的性质，因此我们转而考虑如下问题：

问题 4.2　对于向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$，能否在其中找到所含向量个数最少的部分组 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\ (r<m)$，使得所有能够由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表示的向量也能够由 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性表示？

这一节的内容将围绕这个问题展开讨论，并在讨论这个问题的过程中自然而然地导出相应的概念与方法。

所有能被 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表示的向量构成集合 $L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$，所有能够由 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性表示的向量构成集合 $L\{\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\}$，如果所有能够由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表示的向量也能够由它的部分组 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性表示，即有

$$L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}\subseteq L\{\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\}$$

注意到 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的一个部分组，因此任意能够由 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性表示的向量，一定能够由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表示，即必然有

$$L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}\supseteq L\left\{\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\right\}$$

这就得到此时有

$$L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}=L\left\{\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\right\}$$

在这种意义上，我们称$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$与$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$具有相同的线性表示能力。

我们先讨论一种最简单的情况，那就是部分组与所属向量组只相差一个向量、且它们具有相同线性表示能力的情况，即：

问题 4.3　$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$满足什么条件，才能使得

$$L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}=L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}?$$

对于该问题，很容易给出一个必要条件。若

$$L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}=L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$$

则由于

$$\boldsymbol{\alpha}_i=0\boldsymbol{\alpha}_1+\cdots0\boldsymbol{\alpha}_{i-1}+1\boldsymbol{\alpha}_i+0\boldsymbol{\alpha}_{i+1}+\cdots+0\boldsymbol{\alpha}_m\in L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$$

所以$\boldsymbol{\alpha}_i\in L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}$，因此$\boldsymbol{\alpha}_i$可由$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$线性表示。

实际上该条件也是充分的。若$\boldsymbol{\alpha}_i$可由$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$线性表示，我们证明$L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}=L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$。由于

$$L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}\subseteq L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$$

总是成立的，故只需要证明$L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}\supseteq L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$，也就是证明$\forall\boldsymbol{\beta}\in L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$，有$\boldsymbol{\beta}\in L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}$。现在任取$\boldsymbol{\beta}\in L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$，存在$k_1,k_2,\cdots,k_m\in\mathbb{F}$，使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+\cdots+k_{i-1}\boldsymbol{\alpha}_{i-1}+k_i\boldsymbol{\alpha}_i+k_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+k_m\boldsymbol{\alpha}_m$$

由于$\boldsymbol{\alpha}_i$可由$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$线性表示，故存在$c_1,\cdots,c_{i-1},c_{i+1},\cdots,c_m\in\mathbb{F}$，使得

$$\boldsymbol{\alpha}_i=c_1\boldsymbol{\alpha}_1+\cdots+c_{i-1}\boldsymbol{\alpha}_{i-1}+c_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+c_m\boldsymbol{\alpha}_m$$

所以

$$\begin{aligned}\boldsymbol{\beta}&=k_1\boldsymbol{\alpha}_1+\cdots+k_{i-1}\boldsymbol{\alpha}_{i-1}+k_i\boldsymbol{\alpha}_i+k_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+k_m\boldsymbol{\alpha}_m\\&=k_1\boldsymbol{\alpha}_1+\cdots+k_{i-1}\boldsymbol{\alpha}_{i-1}\\&\quad+k_i\left(c_1\boldsymbol{\alpha}_1+\cdots+c_{i-1}\boldsymbol{\alpha}_{i-1}+c_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+c_m\boldsymbol{\alpha}_m\right)\\&\quad+k_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+k_m\boldsymbol{\alpha}_m\end{aligned}$$

$$=\left(k_1+k_ic_1\right)\boldsymbol{\alpha}_1+\cdots+\left(k_{i-1}+k_ic_{i-1}\right)\boldsymbol{\alpha}_{i-1}+\left(k_{i+1}+k_ic_{i+1}\right)\boldsymbol{\alpha}_{i+1}+\cdots\quad+\left(k_m+k_ic_m\right)\boldsymbol{\alpha}_m$$

故$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$线性表示，所以$\boldsymbol{\beta}\in L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}$。由$\boldsymbol{\beta}$的任意性，我们有$L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}\subseteq L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}$，前文已经说明$L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}\subseteq L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$总是成立的，所以

$$L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}=L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$$

因此命题成立。

这就得到如下结论：

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$与$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$有相同的线性表示能力$\Leftrightarrow$ $\boldsymbol{\alpha}_i$可由$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$线性表示。

该命题纯粹用数学符号表示为：

$$L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}=L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$$

$$\Leftrightarrow\boldsymbol{\alpha}_i\in L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}$$

由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$中除$\boldsymbol{\alpha}_i$之外的向量$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$显然可由向量组$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$线性表示，因此我们得到一个等价的结论：

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$与$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$有相同的线性表示能力$\Leftrightarrow$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$中的每个向量都可由$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m$线性表示。

或用数学符号表示为

$$L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}=L\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}$$

$$\Leftrightarrow\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right\}\subset L\left\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\right\}$$

我们将该结论推广到一般情况，就得到：

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$与$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\left(r\le m\right)$有相同的线性表示能力$\Leftrightarrow$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$中的每

个向量都可由$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$线性表示。

该定理用数学符号表示为

$$L\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_m\}=L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$$

$$\Leftrightarrow\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}\subset L\{\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\}$$

从上述讨论我们可以得到如下结论：

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$与$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\ (r\le m)$有相同的线性表示能力，而$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$的任意部分组都不具有与之相同的线性表示能力$\Leftrightarrow$ $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$中的每个向量都可由$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$线性表示，且$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$中的任意一个向量都不能由$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$中的其余向量线性表示（否则，若$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$中的某个向量可由该向量组中除该向量外的其它向量线性表示，则根据所得结论有：除这个向量外的其它向量所构成的向量组就是满足条件的$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的部分组，而此时该向量组所含向量的个数更少，因此与假设矛盾）。

但是如何判断一个向量组中的任意一个向量都不能由该向量组中的其它向量线性表示？直接考虑这个问题不太容易，我们转而考虑其否命题：向量组中有一个向量可由该向量组中的其余向量线性表示，这个向量组需要满足什么条件？

定理 4.2　　向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$中有一个向量可由该向量组中的其余向量线性表示⇔存在不全为零的数$k_1,k_2,\cdots,k_m$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0}$$

证明：必要性：若有一个向量能由该向量组中的其余向量所组成的向量组线性表示，不妨设$\boldsymbol{\alpha}_m$能由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性表示，因此，存在$c_1,c_2,\cdots,c_{m-1}$，使得

$$\boldsymbol{\alpha}_m=c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_{m-1}\boldsymbol{\alpha}_{m-1}$$

因此有

$$c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_{m-1}\boldsymbol{\alpha}_{m-1}+(-1)\boldsymbol{\alpha}_m=\mathbf{0}$$

而这组数$k_1=c_1,\cdots,k_{m-1}=c_{m-1},k_m=-1$不全为零，故必要性成立。

充分性：设存在不全为零的数$k_1,k_2,\cdots,k_m$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0}$$

不妨设 $k_m \neq 0$，则有

$$-\left(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_{m-1}\boldsymbol{\alpha}_{m-1}\right) = k_m\boldsymbol{\alpha}_m$$

$$\boldsymbol{\alpha}_m = -\frac{k_1}{k_m}\boldsymbol{\alpha}_1 - \frac{k_2}{k_m}\boldsymbol{\alpha}_2 - \cdots \frac{k_{m-1}}{k_m}\boldsymbol{\alpha}_{m-1}$$

所以 $\boldsymbol{\alpha}_m$ 能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1}$ 线性表示。

因此，向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 中的任一向量不能由该向量组的其余向量所构成的部分组线性表示⇔不存在不全为零的数 $k_1,k_2,\cdots,k_m$，使得

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}$$

换句话说，如果向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 中的任一向量不能由该向量组的其余向量所构成的部分组线性表示，且有 $k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}$，则必然有 $k_1 = k_2 = \cdots = k_m = 0$。

由于“一个向量组的任一向量不能由该向量组的其余向量所构成的部分组线性表示”这一条件不容易验证，且为了便于给出这个问题的回答，考虑到如上命题，引入如下定义。

定义 4.6　　**线性相关、线性无关**

对于给定的向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$，如果存在不全为零的数 $k_1,k_2,\cdots,k_m \in \mathbb{F}$，使得

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}$$

则称向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ **线性相关**，否则称该向量组**线性无关**。

换句话说，若向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关，且有 $k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}$，则必有 $k_1 = k_2 = \cdots = k_m = 0$。

向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关性的判断。

定理 4.3 （1）向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关⇔存在不全为零的数 $k_1,k_2,\cdots,k_m$，使得

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}$$

即以 $\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_m\end{bmatrix}$ 为系数矩阵的齐次线性方程组

$$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_m\end{bmatrix}\begin{bmatrix}k_1\\k_2\\\vdots\\k_m\end{bmatrix}=\mathbf{0}$$

有非零解$\Leftrightarrow r\left(\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_m\end{bmatrix}\right)<m$。

（2）向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性无关$\Leftrightarrow r\left(\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_m\end{bmatrix}\right)=m$。

利用线性相关、线性无关的概念，我们有：

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$与$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\ (r\le m)$有相同的线性表示能力，而$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$的任意一部分都不具有与之相同的线性表示能力$\Leftrightarrow\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$中的每个向量都可由$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$线性表示，且$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$线性无关。

为刻划向量组的具有上述性质的部分组，引入如下定义：

定义 4.7　　**极大（线性）无关组**

若向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的部分组$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$满足

（1）$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$中的每个向量都可由$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$线性表示；

（2）$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$线性无关，

则称$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的**极（最）大线性无关组**，简称为**极（最）大无关组**。

因此：

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$与$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\ (r\le m)$有相同的线性表示能力，而$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$的任意一部分都不具有与之相同的线性表示能力$\Leftrightarrow\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的极大无关组。

到目前为止，我们得到结论：所有能由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性表示的向量，都可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的极大无关组$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$线性表示，且其极大无关组的任意部分组（该部分组所含向量个数严格小于该极大无关组所含向量个数）都不再具有该性质，在这种意义上，极大无关组是所含向量个数最少的满足问题 4.1 要求的$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的部分组。但是，一

个向量组的极大无关组是否唯一？如果向量组的极大无关组是唯一的，那么解决问题 4.1 就只需要找该向量组的极大无关组就可以了。若向量组的极大无关组不唯一，自然有如下随之而来的问题：

同一向量组的不同极大无关组所含向量的个数是否相同？

如果这个问题的答案是肯定的，那么解决问题 4.1 就只需要找向量组的任意一个极大无关组就可以了；如果该问题的答案是否定的，则我们还需要在该向量组的所有极大无关组中，找出所含向量个数最少的极大无关组。故我们需要首先回答这个问题：

问题 4.4　　向量组的极大无关组是否是唯一的？

首先，我们用例子说明一个向量组的极大无关组一般是不唯一的。

例 4.4　设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 线性无关，则对于向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$，容易验证 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$、$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$、$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$ 都是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$ 的极大无关组。

因此向量组的极大无关组一般是不唯一的，故需要考虑如下问题：

问题 4.5　　同一向量组的不同极大无关组所含向量的个数是否都是相同的？

首先注意到以下事实：

设 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}\ (r\le m)$、$\boldsymbol{\alpha}_{j_1},\boldsymbol{\alpha}_{j_2},\cdots,\boldsymbol{\alpha}_{j_s}\ (s\le m)$ 都是向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的极大无关组，则根据极大无关组的定义，$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 可线性表示 $\boldsymbol{\alpha}_{j_1},\boldsymbol{\alpha}_{j_2},\cdots,\boldsymbol{\alpha}_{j_s}$ 中的每一个向量，反之也成立。

为了解决这个问题，我们先引入如下定义与定理。

定义 4.8　　向量组的线性表示、向量组等价

若向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 中的每个向量都可由向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示，则称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ **线性表示**。如果两个向量组能够相互线性表示，则称这两个向量组**等价**。

因此有命题：

同一向量组的不同极大无关组是等价的。

如何判断向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是否可由向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示？

定理 4.4　　$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示 $\Leftrightarrow$ 存在矩阵 $\boldsymbol{C}_{s\times r}$，使得

$$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_s\end{bmatrix}\boldsymbol{C}_{s\times r}$$

证明：必要性：若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$线性表示，由定义可知，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$中的每个向量都可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$线性表示，因此对于$\boldsymbol{\alpha}_i$，存在系数$c_{1i},c_{2i},\cdots,c_{si}$，使得

$$\boldsymbol{\alpha}_i=c_{1i}\boldsymbol{\beta}_1+\cdots+c_{si}\boldsymbol{\beta}_s=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_s\end{bmatrix}\begin{bmatrix}c_{1i}\\ c_{2i}\\ \vdots\\ c_{si}\end{bmatrix}$$

所以

$$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_s\end{bmatrix}\begin{bmatrix}c_{11} & c_{12} & \cdots & c_{1r}\\ c_{21} & c_{22} & \cdots & c_{2r}\\ \vdots & \vdots & & \vdots\\ c_{s1} & c_{s2} & \cdots & c_{sr}\end{bmatrix}$$

令

$$\boldsymbol{C}=\begin{bmatrix}c_{11} & c_{12} & \cdots & c_{1r}\\ c_{21} & c_{22} & \cdots & c_{2r}\\ \vdots & \vdots & & \vdots\\ c_{s1} & c_{s2} & \cdots & c_{sr}\end{bmatrix}$$

即可。

充分性：显然，若有$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_s\end{bmatrix}\boldsymbol{C}_{s\times r}$，则说明

$$\boldsymbol{\alpha}_i=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_s\end{bmatrix}\begin{bmatrix}c_{1i}\\ c_{2i}\\ \vdots\\ c_{si}\end{bmatrix}=c_{1i}\boldsymbol{\beta}_1+\cdots+c_{si}\boldsymbol{\beta}_s$$

所以$\boldsymbol{\alpha}_i$可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$线性表示。由$\boldsymbol{\alpha}_i$的任意性可知充分性得证。 ■

有了上述结论，容易证明向量组线性表示具有如下意义的传递性：

向量组线性表示的传递性。

设向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$可由向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$线性表示，向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$可由向量组$\gamma_1,\gamma_2,\cdots,\gamma_t$线性表示，则向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$可由向量组$\gamma_1,\gamma_2,\cdots,\gamma_t$线性表示。

证明：根据定理 4.4 及题设条件，可知存在矩阵$\boldsymbol{C},\boldsymbol{D}$，分别使得

$$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_s\end{bmatrix}\boldsymbol{C}$$

$$[\boldsymbol{\beta}_1 \quad \boldsymbol{\beta}_2 \quad \cdots \quad \boldsymbol{\beta}_s] = [\boldsymbol{\gamma}_1 \quad \boldsymbol{\gamma}_2 \quad \cdots \quad \boldsymbol{\gamma}_t]\boldsymbol{D}$$

所以

$$\begin{aligned}[\boldsymbol{\alpha}_1 \quad \boldsymbol{\alpha}_2 \quad \cdots \quad \boldsymbol{\alpha}_r] &= [\boldsymbol{\beta}_1 \quad \boldsymbol{\beta}_2 \quad \cdots \quad \boldsymbol{\beta}_s]\boldsymbol{C} \\ &= [\boldsymbol{\gamma}_1 \quad \boldsymbol{\gamma}_2 \quad \cdots \quad \boldsymbol{\gamma}_t]\boldsymbol{DC}\end{aligned}$$

根据定理 4.4，这说明$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$可由$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_t$线性表示。

由以上结论就得到向量组等价的传递性：

设向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$与向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$等价，向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$与向量组$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_t$等价，则向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$与向量组$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_t$等价。

容易验证向量组的等价满足反身性、对称性、传递性，因此它是一种等价关系。

我们给出向量组线性表示的另外一个充要条件。

定理 4.5　向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 可由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表示的充要条件是矩阵$\boldsymbol{A} = [\boldsymbol{\alpha}_1 \ \boldsymbol{\alpha}_2 \ \cdots \ \boldsymbol{\alpha}_m]$的秩等于矩阵$[\boldsymbol{A} \quad \boldsymbol{B}] = [\boldsymbol{\alpha}_1 \quad \boldsymbol{\alpha}_2 \quad \cdots \quad \boldsymbol{\alpha}_m \quad \boldsymbol{\beta}_1 \quad \boldsymbol{\beta}_2 \quad \cdots \quad \boldsymbol{\beta}_s]$的秩，即$r(\boldsymbol{A}) = r([\boldsymbol{A} \quad \boldsymbol{B}])$。

证明：根据定理 4.4，向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性表示⇔存在矩阵$\boldsymbol{K}$，使得

$$\boldsymbol{B} = \boldsymbol{AK}$$

即矩阵方程

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

有解$\boldsymbol{K}$。（第三章定理 3.16 ）⇔$r(\boldsymbol{A}) = r([\boldsymbol{A} \quad \boldsymbol{B}])$。 ■

推论 4.1　向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$等价的一个充要条件是

$$r[\boldsymbol{\alpha}_1 \ \boldsymbol{\alpha}_2 \ \cdots \ \boldsymbol{\alpha}_m] = r[\boldsymbol{\beta}_1 \ \boldsymbol{\beta}_2 \ \cdots \ \boldsymbol{\beta}_s] = r[\boldsymbol{\alpha}_1 \quad \boldsymbol{\alpha}_2 \quad \cdots \quad \boldsymbol{\alpha}_m \quad \boldsymbol{\beta}_1 \quad \boldsymbol{\beta}_2 \quad \cdots \quad \boldsymbol{\beta}_s]$$

（本段放在这里合适？）设矩阵$\boldsymbol{A}$经过初等行变换变成矩阵$\boldsymbol{B}$，则$\boldsymbol{B}$的每个行向量都可由$\boldsymbol{A}$的行向量组线性表示，即$\boldsymbol{B}$的行向量组可由$\boldsymbol{A}$的行向量组线性表示；由初等行变换的可逆性可知，$\boldsymbol{B}$经过初等行变换可变成矩阵$\boldsymbol{A}$，因此$\boldsymbol{A}$的行向量组可由$\boldsymbol{B}$的行向量组线性表示，这说明这两个矩阵的行向量组等价。同样，如果矩阵$\boldsymbol{A}$经过初等列变换变成矩阵$\boldsymbol{B}$，则这两个矩阵的列向量组等价。

下面这个定理在解决我们目前所面临的问题上起着基础性作用，这表现为如下定理：

定理 4.6　设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 能由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示，且 $r>s$，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性相关。

证明：由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示，根据定理 4.4 可知，存在矩阵 $\boldsymbol{C}_{s\times r}$，使得

$$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_s\end{bmatrix}\boldsymbol{C}_{s\times r}$$

由于 $r>s$，故 $r\left(\boldsymbol{C}_{s\times r}\right)\le\min\{s,r\}=s<r$，因此齐次线性方程组

$$\boldsymbol{C}_{s\times r}\begin{bmatrix}x_1\\x_2\\\vdots\\x_r\end{bmatrix}=\mathbf{0}$$

有非零解，为简单起见，我们将该非零解还记作 $x_1,x_2,\cdots,x_r$，这说明有不全为零的数 $x_1,x_2,\cdots,x_r$，使得

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_r\boldsymbol{\alpha}_r=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_r\end{bmatrix}$$

$$=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_s\end{bmatrix}\boldsymbol{C}_{s\times r}\begin{bmatrix}x_1\\x_2\\\vdots\\x_r\end{bmatrix}==\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_s\end{bmatrix}\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}=\mathbf{0}$$

这说明有不全为零的数 $x_1,x_2,\cdots,x_r$，使得

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_r\boldsymbol{\alpha}_r=\mathbf{0}$$

成立，因此 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性相关。　■

推论 4.2　若向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 可由向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示，且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，则 $r\le s$。

定理 4.7　如果两个向量组等价，且这两个向量组均线性无关，那么这两个向量组所含向

量的个数相等。

证明：设$\alpha_1,\alpha_2,\cdots,\alpha_r$与$\beta_1,\beta_2,\cdots,\beta_s$为满足命题条件的两个向量组，由于它们等价，这说明$\alpha_1,\alpha_2,\cdots,\alpha_r$可由$\beta_1,\beta_2,\cdots,\beta_s$线性表示，由于$\alpha_1,\alpha_2,\cdots,\alpha_r$线性无关，由推论 4.1 可知$r\le s$；同样地，$\beta_1,\beta_2,\cdots,\beta_s$可由$\alpha_1,\alpha_2,\cdots,\alpha_r$线性表示，而$\beta_1,\beta_2,\cdots,\beta_s$线性无关，由推论 4.1 可知$s\le r$。综合可得$s=r$，命题得证。

■

推论 4.3　　同一向量组的不同极大无关组所含向量的个数是相同的。
证明　　由于同一个向量组的不同极大无关组是等价的，且极大无关组是线性无关的，由定理 4.7 即得本结论。

这个结论回答了我们最初提出的问题：

$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$是向量组$\alpha_1,\alpha_2,\cdots,\alpha_m$的满足$L\{\alpha_1,\alpha_2,\cdots,\alpha_m\}=L\{\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}\}$的所含向量个数最少的部分组$\Leftrightarrow\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$为$\alpha_1,\alpha_2,\cdots,\alpha_m$的极大无关组。

由于向量组的所有极大无关组所含向量的个数是相同的，故如下定义是有意义的：

定义 4.9　向量组的**秩**

向量组的极大无关组所含向量的个数称为向量组的**秩**。向量组$\alpha_1,\alpha_2,\cdots,\alpha_m$的秩记作$r\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$。

根据以上论述，容易得到如下定理。

定理 4.8（1）$\alpha_1,\alpha_2,\cdots,\alpha_m$线性相关$\Leftrightarrow r\{\alpha_1,\alpha_2,\cdots,\alpha_m\}<m$。

(2) 向量组$\alpha_1,\alpha_2,\cdots,\alpha_m$能由向量组$\beta_1,\beta_2,\cdots,\beta_n$线性表示，则$r\{\alpha_1,\alpha_2,\cdots,\alpha_m\}\le r\{\beta_1,\beta_2,\cdots,\beta_n\}$。

(3)等价的向量组有相同的秩。

证明　　（1）根据线性相关以及向量组秩的定义，这是显然的。

（2）设$r\{\alpha_1,\alpha_2,\cdots,\alpha_m\}=s$，$r\{\beta_1,\beta_2,\cdots,\beta_n\}=t$，$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_s}$为$\alpha_1,\alpha_2,\cdots,\alpha_m$的一个极大无关组，$\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_t}$为$\beta_1,\beta_2,\cdots,\beta_n$的一个极大无关组。因此，$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_s}$可由$\alpha_1,\alpha_2,\cdots,\alpha_m$线性表示，而$\alpha_1,\alpha_2,\cdots,\alpha_m$能由向量组$\beta_1,\beta_2,\cdots,\beta_n$线性表示，

$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 可由 $\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_t}$ 线性表示，由向量组线性表示的传递性可知，$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_s}$ 可由 $\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_t}$ 线性表示，由于 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_s}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的极大无关组，故 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_s}$ 线性无关，根据推论 4.1 可知 $s\le t$，因此命题成立。

（3）只需要应用向量组等价的定义以及（2）的结论即可。

■

如何求向量组的秩？

到现在为止，我们引入了矩阵的秩、向量组的秩。而向量组可构成矩阵，那么这两者的秩有何关系？

定理 4.9　矩阵的秩等于它的列向量组的秩，也等于它的行向量组的秩。

证明　设 $\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_k\end{bmatrix}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1k}\\ a_{21} & a_{22} & \cdots & a_{2k}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nk}\end{bmatrix}$，且设 $r(\boldsymbol{A})=r$，因此在 $\boldsymbol{A}$ 中存在一个不为零的 r 阶子式，不妨设该 r 阶子式为

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

所以

$$\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1r}\\ a_{21} & a_{22} & \cdots & a_{2r}\\ \vdots & \vdots & & \vdots\\ a_{r1} & a_{r2} & \cdots & a_{rr}\end{bmatrix}$$

的秩为 r，根据矩阵秩的定义与性质可知

$$\boldsymbol{B}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1r}\\ a_{21} & a_{22} & \cdots & a_{2r}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nr}\end{bmatrix}$$

的秩也为 r。所以齐次线性方程组

$$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_r\end{bmatrix}=\boldsymbol{BX}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1r}\\ a_{21} & a_{22} & \cdots & a_{2r}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nr}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_r\end{bmatrix}=\begin{bmatrix}0\\ 0\\ \vdots\\ 0\end{bmatrix}$$

只有零解，故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，故 $r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k\}\ge r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}=r=r(\boldsymbol{A})$。

反过来，设 $r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k\}=r$，且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 为它的一个极大无关组，这说明

$$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_r\end{bmatrix}=\boldsymbol{BX}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1r}\\a_{21} & a_{22} & \cdots & a_{2r}\\\vdots & \vdots & & \vdots\\a_{n1} & a_{n2} & \cdots & a_{nr}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_r\end{bmatrix}=\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}$$

只有零解，所以 $r(\boldsymbol{B})=r$，由于 $\boldsymbol{B}$ 的子式都是 $\boldsymbol{A}$ 的子式，所以 $r(\boldsymbol{B})\le r(\boldsymbol{A})$，故 $r(\boldsymbol{A})\ge r(\boldsymbol{B})=r=r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k\}$。综合上述结果可知

$$r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k\}=r(\boldsymbol{A})$$

由于 $r(\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}})$，而 $\boldsymbol{A}$ 的行向量组就是 $\boldsymbol{A}^{\mathrm{T}}$ 的列向量组，所以可得到行向量组的情况。

■

由于向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的秩 $r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k\}$ 等于矩阵 $\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}$ 的秩 $r\left(\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}\right)$，因此在以后对这两者就不加以区别。

根据上述定理，为求 $\mathbb{F}^n$ 中的向量组的秩，我们用这些向量构成矩阵，然后求该矩阵的秩即得到该向量组的秩。另外，利用该方法，我们还可以求得向量组的一个极大无关组。

定理 4.10　矩阵 $\boldsymbol{A}$ 经过初等行变换化为矩阵 $\boldsymbol{B}$，则 $\boldsymbol{A}$ 的列向量组与 $\boldsymbol{B}$ 对应的列向量组有相同的线性组合关系。

证明：对矩阵 $\boldsymbol{A}$ 做初等行变换得到 $\boldsymbol{B}$，相当于用一个可逆矩阵 $\boldsymbol{P}$ 左乘 $\boldsymbol{A}$，即 $\boldsymbol{PA}=\boldsymbol{B}$。对 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 按列分块，即

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n\end{bmatrix},\qquad \boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & \boldsymbol{B}_n\end{bmatrix}$$

则有

$$\begin{aligned}\boldsymbol{PA}&=\boldsymbol{P}\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{PA}_1 & \boldsymbol{PA}_2 & \cdots & \boldsymbol{PA}_n\end{bmatrix}\\&=\begin{bmatrix}\boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & \boldsymbol{B}_n\end{bmatrix}=\boldsymbol{B}\end{aligned}$$

这说明 $\boldsymbol{PA}_i=\boldsymbol{B}_i, i=1,2,\cdots,n$，。设 $\boldsymbol{A}$ 的某些列 $\boldsymbol{A}_{i1},\boldsymbol{A}_{i2},\cdots,\boldsymbol{A}_{ik}$ 的线性组合为零向量

$$x_1\boldsymbol{A}_{i1}+x_2\boldsymbol{A}_{i2}+\cdots+x_k\boldsymbol{A}_{ik}=\boldsymbol{0}$$

就有

$$\begin{aligned}&x_1\boldsymbol{B}_{i1}+x_2\boldsymbol{B}_{i2}+\cdots+x_k\boldsymbol{B}_{ik}\\&=x_1\boldsymbol{PA}_{i1}+x_2\boldsymbol{PA}_{i2}+\cdots+x_k\boldsymbol{PA}_{ik}\\&=\boldsymbol{P}\left(x_1\boldsymbol{A}_{i1}+x_2\boldsymbol{A}_{i2}+\cdots+x_k\boldsymbol{A}_{ik}\right)\\&=\boldsymbol{P0}=\boldsymbol{0}\end{aligned}$$

反过来，若 $x_1\boldsymbol{B}_{i1}+x_2\boldsymbol{B}_{i2}+\cdots+x_k\boldsymbol{B}_{ik}=\boldsymbol{P}\left(x_1\boldsymbol{A}_{i1}+x_2\boldsymbol{A}_{i2}+\cdots+x_k\boldsymbol{A}_{ik}\right)=\boldsymbol{0}$，由于矩阵 $\boldsymbol{P}$ 可逆，因此 $x_1\boldsymbol{A}_{i1}+x_2\boldsymbol{A}_{i2}+\cdots+x_k\boldsymbol{A}_{ik}=\boldsymbol{0}$。故矩阵 $\boldsymbol{A}$、$\boldsymbol{B}$ 对应的列向量有相同的线性组合关系。

根据定理 4.9 就得到求向量组的一个极大无关组的方法：以向量组中的向量作为列向量构成矩阵 $\boldsymbol{A}$，将矩阵 $\boldsymbol{A}$ 化为行标准形矩阵 $\boldsymbol{B}$，根据行标准形矩阵的列向量的线性组合关系就可得到原来向量组中的向量极大无关组。

例 4.5　求向量组

$$\boldsymbol{\alpha}_1=[1,4,1,0,2]^{\mathrm{T}},\boldsymbol{\alpha}_2=[2,5,-1,-3,2]^{\mathrm{T}},\boldsymbol{\alpha}_3=[-1,2,5,6,2]^{\mathrm{T}},\boldsymbol{\alpha}_4=[0,2,2,-1,0]^{\mathrm{T}}$$

的秩，找出它的一个极大无关组，并将其它向量用该极大无关组线性表示。

解：用这些向量作为列向量构成矩阵

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3 & \boldsymbol{\alpha}_4\end{bmatrix}=\begin{bmatrix}1&2&-1&0\\4&5&2&2\\1&-1&5&2\\0&-3&6&-1\\2&2&2&0\end{bmatrix}$$

利用初等行变换将其化为行最简形矩阵

$$\begin{bmatrix}1&0&3&0\\0&1&-2&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{bmatrix}=\boldsymbol{B}=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \boldsymbol{\beta}_3 & \boldsymbol{\beta}_4\end{bmatrix}$$

所以，$r(\boldsymbol{B})=3$，故向量组的秩为 3，且 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_4$ 线性无关，因此 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4$ 也线性无关。注意到 $\boldsymbol{\beta}_3=3\boldsymbol{\beta}_1-2\boldsymbol{\beta}_2$，根据定理 4.9 可得 $\boldsymbol{\alpha}_3=3\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2$。

到目前为止，我们已经完美地回答了问题 4.2，下面我们补充一些结论。

定理 4.11　若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}$ 线性相关，则 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示，且表示方式唯一。

证明：由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}$ 线性相关，故存在不全为零的数 $k_1,k_2,\cdots,k_n,k\in\mathbb{F}$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n+k\boldsymbol{\gamma}=\mathbf{0}$$

我们可断言$k\neq 0$，实际上，若$k=0$，则可得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\mathbf{0}$$

由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性无关，得到

$$k_1=k_2=\cdots=k_n=0$$

因而

$$k_1=k_2=\cdots=k_n=k=0$$

这与$k_1,k_2,\cdots,k_n,k$的取法相矛盾，故确有$k\neq 0$，所以

$$\boldsymbol{\beta}=-\frac{k_1}{k}\boldsymbol{\alpha}_1-\frac{k_2}{k}\boldsymbol{\alpha}_2\cdots-\frac{k_n}{k}\boldsymbol{\alpha}_n$$

故$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性表示。

下证线性表示的唯一性。若有两种表示方式

$$\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n$$

$$\boldsymbol{\beta}=y_1\boldsymbol{\alpha}_1+y_2\boldsymbol{\alpha}_2+\cdots+y_n\boldsymbol{\alpha}_n$$

两式相减得

$$\mathbf{0}=(x_1-y_1)\boldsymbol{\alpha}_1+(x_2-y_2)\boldsymbol{\alpha}_2+\cdots+(x_n-y_n)\boldsymbol{\alpha}_n$$

由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$此线性无关，故由上式可知

$$x_1-y_1=x_2-y_2=\cdots=x_n-y_n=0$$

$$\Rightarrow x_1=y_1,x_2=y_2,\cdots,x_n=y_n$$

所以表示方式唯一。

推论 4.4　若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性无关，且$\boldsymbol{\beta}$不能由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性表示，则$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}$线性无关。

定理 4.12　若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性无关，向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性表示，且

$$\begin{aligned}\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_r\end{bmatrix} &= \begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\boldsymbol{X} \\ &= \begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_r\end{bmatrix}\end{aligned}$$

则 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性相关 $\Leftrightarrow \boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_r$ 线性相关。

证明：若有 $k_1,k_2,\cdots,k_r\in\mathbb{F}$，使得

$$\mathbf{0}=\sum_{i=1}^{r}k_i\boldsymbol{\beta}_i=\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_r\end{bmatrix}\begin{bmatrix}k_1\\k_2\\\vdots\\k_r\end{bmatrix}$$

$$=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_r\end{bmatrix}\begin{bmatrix}k_1\\k_2\\\vdots\\k_r\end{bmatrix}$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，故上式成立的充要条件是

$$\begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_r\end{bmatrix}\begin{bmatrix}k_1\\k_2\\\vdots\\k_r\end{bmatrix}=\mathbf{0}$$

故：$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性相关 $\Leftrightarrow \boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_r$ 线性相关。

推论 4.5（1）题设条件同定理 4.12，则有 $r\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_r\end{bmatrix}=r(\boldsymbol{X})$；

（2）若 $\boldsymbol{C}=\boldsymbol{AB}$，且矩阵 $\boldsymbol{A}$ 列满秩，则 $r(\boldsymbol{C})=r(\boldsymbol{B})$；

（3）若 $\boldsymbol{C}=\boldsymbol{AB}$，且矩阵 $\boldsymbol{B}$ 行满秩，则 $r(\boldsymbol{C})=r(\boldsymbol{A})$。

定理 4.13（1）若向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关，则它的任一部分组也线性无关；若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关，则对于任意向量 $\boldsymbol{\beta}$，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}$ 线性相关。

（2）若 $\begin{bmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{bmatrix},\begin{bmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{bmatrix},\cdots,\begin{bmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{bmatrix}$ 线性无关，则 $\begin{bmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\\a_{m+1,1}\end{bmatrix},\begin{bmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\\a_{m+1,2}\end{bmatrix},\cdots,\begin{bmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\\a_{m+1,n}\end{bmatrix}$ 也线性无关；

反之若$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{m+1,1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \\ a_{m+1,2} \end{bmatrix}, \cdots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \\ a_{m+1,n} \end{bmatrix}$线性相关，则$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \cdots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$线性相关。

（3）任意$n+1$个n维向量所构成的向量组都是线性相关的。

证明（1）只证明第一个结论，第二个结论是第一个结论的逆否命题。用反证法。设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$线性无关，若它的一个部分组线性相关，不妨设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$为$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$的线性相关的部分组，根据定义，存在不全为零的数$k_1, k_2, \cdots, k_s$，使得$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_s\boldsymbol{\alpha}_s = \boldsymbol{0}$。把上式写成

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_s\boldsymbol{\alpha}_s + 0\boldsymbol{\alpha}_{s+1} + \cdots + 0\boldsymbol{\alpha}_m = \boldsymbol{0}$$

这m个数$k_1, k_2, \cdots, k_s, 0, \cdots, 0$不全为零，故$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$线性相关，这与题设矛盾。

（2）考虑矩阵$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$，$\boldsymbol{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \\ a_{m+1,1} & a_{m+1,2} & \cdots & a_{m+1,n} \end{bmatrix}$。由于$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \cdots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$线性无关，故矩阵$\boldsymbol{A}$的秩$r(\boldsymbol{A}) = n$；由于$\boldsymbol{A}$为$\boldsymbol{B}$的子矩阵，故$r(\boldsymbol{A}) \le r(\boldsymbol{B})$，所以$n = r(\boldsymbol{A}) \le r(\boldsymbol{B}) \le n$，所以$r(\boldsymbol{B}) = n$，故$\boldsymbol{B}$的列向量组线性无关，命题得证。

（3）设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_{n+1}$为$n$维向量所构成的向量组，由于任意一个$n$维向量都可以由$n$维基本向量组$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$线性表示，故$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_{n+1}$可由$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$线性表示，而$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_{n+1}$所含向量的个数大于$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$所含向量的个数，故$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_{n+1}$线性相关。

习题 4.3

1. 已知$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$线性无关，试判断$3\boldsymbol{\alpha}_1 + 2\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_3, 4\boldsymbol{\alpha}_3 - 5\boldsymbol{\alpha}_1$ 是否线性无关.

2. 设向量组$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$能由向量组$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_t$线性表示为

$$\left[\boldsymbol{\beta}_1 \quad \boldsymbol{\beta}_2 \quad \cdots \quad \boldsymbol{\beta}_s\right]=\left[\boldsymbol{\alpha}_1 \quad \boldsymbol{\alpha}_2 \quad \cdots \quad \boldsymbol{\alpha}_t\right]\boldsymbol{K}$$

其中$\boldsymbol{K}$为$t\times s$矩阵，且$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_t$组线性无关。证明：$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$组线性无关的充要条件是：矩阵$\boldsymbol{K}$的秩$r\left(\boldsymbol{K}\right)=s$。

3. 已知向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s(s\ge 2)$ 线性无关，设$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$，$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\cdots,$ $\boldsymbol{\beta}_{s-1}=\boldsymbol{\alpha}_{s-1}+\boldsymbol{\alpha}_s,\boldsymbol{\beta}_s=\boldsymbol{\alpha}_s+\boldsymbol{\alpha}_1$，讨论$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$的线性相关性.

4. 已知向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性相关，$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$线性无关，问 （1）$\boldsymbol{\alpha}_1$能否由$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表示？证明你的结论； （2）$\boldsymbol{\alpha}_4$能否由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示？证明你的结论.

5. 设$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 线性无关，又设$\boldsymbol{\delta}\neq\mathbf{0}$，$\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\delta}$线性相关，$\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta}$也线性相关，证明$\boldsymbol{\delta}=k\boldsymbol{\gamma}$ ，其中$k\neq 0$是常数。

6. 已知$\mathbb{R}^4$中，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性无关，$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$线性无关，（1）若$\boldsymbol{\gamma}\in\mathbb{R}^4$不能由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表出，则$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\gamma}$线性无关；（2）证明$\boldsymbol{\delta}\in\mathbb{R}^4$使得$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\delta}$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\delta}$均线性无关.

7. 设在向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$中，$\boldsymbol{\alpha}_1\neq\mathbf{0}$，并且每一个$\boldsymbol{\alpha}_i$都不能由前面的$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{i-1}$线性表示，证明$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性无关。

8. 设$\boldsymbol{A}$是n阶方阵，若存在正整数k，使得线性方程组$\boldsymbol{A}^k\boldsymbol{X}=\mathbf{0}$有解向量$\boldsymbol{\alpha}$，且$\boldsymbol{A}^{k-1}\boldsymbol{\alpha}\neq\mathbf{0}$，证明向量组$\boldsymbol{\alpha},\boldsymbol{A}\boldsymbol{\alpha},\cdots,\boldsymbol{A}^{k-1}\boldsymbol{\alpha}$是线性无关的。

9. 已知向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$构成的矩阵分别为$\boldsymbol{A}=\left[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\right]=\begin{bmatrix}2&3\\0&-2\\-1&1\\3&-1\end{bmatrix}$，

$\boldsymbol{B}=\left[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\right]=\begin{bmatrix}-5&4\\6&-4\\-5&3\\9&-5\end{bmatrix}$，证明向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$等价。

10. 设有向量组$(\mathrm{I}):\boldsymbol{\alpha}_1=[1,0,2]^{\mathrm{T}},\boldsymbol{\alpha}_2=[1,1,3]^{\mathrm{T}},\boldsymbol{\alpha}_3=[1,-1,a+2]^{\mathrm{T}}$和向量组$(\mathrm{II}):\boldsymbol{\beta}_1=[1,2,a+3]^{\mathrm{T}},\boldsymbol{\beta}_2=[2,1,a+6]^{\mathrm{T}},\boldsymbol{\beta}_3=[2,1,a+4]^{\mathrm{T}}$.

(1) 当 a 为何值时，向量组 (I) 与 (II) 等价？说明理由.

(2) 当 a 为何值时，向量组 (I) 与 (II) 不等价？说明理由。

11. 求向量组

$\alpha_1=\left[1,-1,0,0\right]^{\mathrm{T}},\alpha_2=\left[-1,2,1,-1\right]^{\mathrm{T}},\alpha_3=\left[0,1,1,-1\right]^{\mathrm{T}},\alpha_4=\left[-1,3,2,1\right]^{\mathrm{T}},\alpha_5=\left[-2,6,4,1\right]^{\mathrm{T}}$

的秩，并给出该向量组的一个极大无关组，并把该向量组中的其它向量用该极大无关组线性表示。

12. 设向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示，证明 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关。

13. 向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 可被 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性表出，且秩相等，证明 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 也可被 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出。

14. 设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的秩为 r_1， $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 的秩为 r_2，向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_t$ 的秩为 r_3，证明：$\max\left\{r_1,r_2\right\}\le r_3\le r_1+r_2$。

15. 已知 3 阶方阵 $\boldsymbol{A}$ 与 3 维列向量 $\boldsymbol{X}$ 满足 $\boldsymbol{A}^3\boldsymbol{X}=3\boldsymbol{AX}-\boldsymbol{A}^2\boldsymbol{X}$，且向量组 $\boldsymbol{X},\boldsymbol{AX},\boldsymbol{A}^2\boldsymbol{X}$ 线性无关。（1）记 $\boldsymbol{P}=\left[\boldsymbol{X},\boldsymbol{AX},\boldsymbol{A}^2\boldsymbol{X}\right]$，求三阶方阵 $\boldsymbol{B}$，使得 $\boldsymbol{AP}=\boldsymbol{PB}$；（2）求 $\left|\boldsymbol{A}\right|$。

§4.4　　向量空间

初等数学的一个特点是对具体的对象加以研究，而更高级的数学则是对抽象的对象加以研究，因此需要明确这些对象所具有的性质与运算，如何明确这些对象的性质与运算？这就是空间的概念。所谓空间，就是具有特定性质的对象的集合。

定义 4.10　　向量空间

设 V 为数域 $\mathbb{F}$ 上的 n 维向量构成的非空集合，且满足

1. $\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in V$， $\boldsymbol{\alpha}+\boldsymbol{\beta}\in V$；

2. $\forall\boldsymbol{\alpha}\in V,k\in\mathbb{F},$ 有 $k\boldsymbol{\alpha}\in V$；

则称集合 V 为数域 $\mathbb{F}$ 上的向量空间。若 $\mathbb{F}$ 为实(复)数域，则称 V 为实(复)向量空间。

称性质(1)、(2)称为向量空间 V 关于向量加法、向量数乘封闭。因此向量空间也就是对向量加法、向量数乘封闭的向量集合，因此向量空间中任意向量组的任意线性组合还在该向量空间中。

命题　　向量空间 V 必然含有零向量。

证明　　由于V为向量空间，故$V \neq \phi$，任取$\boldsymbol{\alpha} \in V$，由于向量空间满足数乘封闭性，故$0\boldsymbol{\alpha} \in V$，即$\mathbf{0} \in V$。

例 4.6　$\mathbb{R}^n$是一个向量空间。

例 4.7　判断

（1）$V_1 = \left\{ x = \left[0, x_2, \cdots, x_n\right]^{\mathrm{T}} \middle| x_2, \cdots, x_n \in \mathbb{F} \right\}$；

（2）$V_2 = \left\{ x = \left[1, x_2, \cdots, x_n\right]^{\mathrm{T}} \middle| x_2, \cdots, x_n \in \mathbb{F} \right\}$

是否为向量空间？

解：（1）只需要判断该集合是否满足加法、数乘封闭性。

$$\forall \boldsymbol{\alpha} = \left[0, a_2, \cdots, a_n\right]^{\mathrm{T}}, \boldsymbol{\beta} = \left[0, b_2, \cdots, b_n\right]^{\mathrm{T}} \in V_1，其中 a_i, b_j \in \mathbb{F}，$$

有　$\boldsymbol{\alpha} + \boldsymbol{\beta} = \left[0, a_2 + b_2, \cdots, a_n + b_n\right]^{\mathrm{T}}$，且$a_i + b_i \in \mathbb{F}$，故

$$\boldsymbol{\alpha} + \boldsymbol{\beta} = \left[0, a_2 + b_2, \cdots, a_n + b_n\right]^{\mathrm{T}} \in V_1$$

设$k \in \mathbb{F}$，则$k\boldsymbol{\alpha} = \left[0, ka_2, \cdots, ka_n\right]^{\mathrm{T}} \in V_1$。所以$V_1$是向量空间。

（2）由于$\mathbf{0} \notin V_2$，所以V_2不是向量空间。

例 4.8　设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$为向量空间$V$中的向量组，证明：

$$L\left\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m\right\} = span\left\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m\right\} = \left\{c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_m\boldsymbol{\alpha}_m \middle| c_1, c_2, \cdots, c_m \in \mathbb{F}\right\}$$

是一个向量空间，称为由向量组$\alpha_1, \alpha_2, \cdots, \alpha_m$生成的向量空间。

证明　　证明略。

例 4.9　矩阵的**零空间**（null space）、**列空间**（ column space ）

　　齐次线性方程组$\boldsymbol{AX} = \mathbf{0}$的解集

$$\boldsymbol{N}(\boldsymbol{A}) = \left\{\boldsymbol{X} \middle| \boldsymbol{AX} = \mathbf{0}\right\}$$

是一个向量空间，称为齐次线性方程组$\boldsymbol{AX} = \mathbf{0}$的解空间，也称为矩阵$\boldsymbol{A}$的零空间，记做$N(\boldsymbol{A})$，而由矩阵$\boldsymbol{A}$的列向量组所生成的向量空间称为$\boldsymbol{A}$的列空间，记做$R(\boldsymbol{A})$。

例 4.10　　设向量组$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$与$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$等价，记

$$V_1 = L\left\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m\right\}，V_2 = L\left\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s\right\}$$

则$V_1=V_2$。

证明：由于向量空间V_1,V_2是向量的集合，为证明$V_1=V_2$，只需要证明这两个集合相互包含即可。我们先证明$V_1\subseteq V_2$。任取$\gamma\in V_1$，故γ可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性表示，由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$与$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$等价，故$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$线性表示，根据线性表示的传递性可知$\boldsymbol{\gamma}$可由$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$线性表示，因而$\boldsymbol{\gamma}\in V_2$。由$\boldsymbol{\gamma}$的任意性可得$V_1\subseteq V_2$。类似地可以证明$V_2\subseteq V_1$，故$V_1=V_2$。

设$V\neq\{\mathbf{0}\}$是$n(n<+\infty)$维向量所构成的向量空间，则在V中存在所含向量个数不超过n的线性无关组，使得V中的任意向量都可由该线性无关组线性表示。其理由可以叙述如下：由于$V\neq\{\mathbf{0}\}$，则可取非零向量$\boldsymbol{\alpha}_1\in V$；如果$\forall\boldsymbol{\beta}\in V$，$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1$线性表示，则$\boldsymbol{\alpha}_1$就是满足要求的向量组；否则，取$\boldsymbol{\alpha}_2\in V$，使得$\boldsymbol{\alpha}_2$不能由$\boldsymbol{\alpha}_1$线性表示，我们得到向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$，显然此时$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$线性无关；如果$\forall\boldsymbol{\beta}\in V$，$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$线性表示，则$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$就是满足要求的向量组，否则取$\boldsymbol{\alpha}_3\in V$，使得$\boldsymbol{\alpha}_3$不能$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$由线性表示，我们得到向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$，此时$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性无关；我们重复上述过程，由于任意$n+1$个$n$维向量所构成的向量组都是线性相关的，因此上述过程必然在$d\leq n$步后停止，我们得到线性无关的向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$，且$\forall\boldsymbol{\beta}\in V$，$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$线性表示，因此$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$就是满足要求的线性无关组，即有$V=L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$。

根据上述结论，我们有如下定义：

定义 4.11 向量空间的**基**、**维数**

设V是由n维向量所构成的向量空间，若V中的向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$满足

（1）$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$ 线性无关；

（2）V中的任一向量都可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$线性表示；

则称$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$为向量空间$V$的一组**基**，$d$为向量空间$\alpha_1,\alpha_2,\cdots,\alpha_d$的**维数**，记为$\dim V=d$，并称$V$为$d$维向量空间。

注意：由于只由零向量构成的向量空间$\{\mathbf{0}\}$没有基存在，我们将它的维数规定为 0。

从上述定义可知，向量空间的基与向量组的极大无关组有类似的性质，而向量空间的维数与向量组的秩有可比性。实际上，我们有如下结论：

命题　设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$为一向量组，则$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的极大无关组为向量空间$V=L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$的一组基，且$r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}=\dim V$。

证明　这是显然的。

例 4.11　$\dim\mathbb{R}^n=n$，标准向量组$\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$为一组基；而例 4.6 中的$V_1$，则有$\dim V_1=n-1$，$\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$为其一组基。

需要注意两种维数定义的区别：向量空间的维数与向量的维数。例如例 4.6 中向量空间V_1中的向量是n维的，但向量空间V_1的维数$\dim V_1=n-1$。

定理 4.14　d维向量空间V中的任意d个线性无关的向量都构成V的一组基。

证明：设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\in V$线性无关，我们要证明$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$为$V$的一组基，只要证明$V$中的任意向量都可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$线性表示。由于$V$中的任意一个向量都可由该向量空间的基线性表示，由于V为d维向量空间，它的基含有d个向量，因此V中任意$d+1$向量构成的向量组都线性相关。任取$\boldsymbol{\beta}\in V$，故向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d,\boldsymbol{\beta}$线性相关，根据前面所得的结论可知：$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$唯一地线性表示，因此$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$为向量空间$V$的一组基。证毕。　■

根据上述定理可知，向量空间的基是不唯一的。设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$为$d$维向量空间$V$的一组基，而$\boldsymbol{C}$为$d$阶可逆矩阵，则由$\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_d\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_d\end{bmatrix}\boldsymbol{C}$所得到的向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\in V$且线性无关，因此$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d$也是$V$的基，由于可逆矩阵$\boldsymbol{C}$是无穷多的，所以一个向量空间的基也是无穷多的。

根据前面的分析，对于d维向量空间V中的一个线性无关向量组

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\,(r<d)$，存在$\boldsymbol{\alpha}_{r+1}\in V$，使得$\boldsymbol{\alpha}_{r+1}$不能由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性表示，所以$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1}$线性无关，依此类推，我们可得到线性无关的向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_d$，因此它就是$\boldsymbol{V}$的一组基。也就是说：

向量空间中的任意一个线性无关的向量组都可扩充为该向量空间的一组基。

定义 4.12　**坐标**

若$B=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$为$d$维向量空间$\boldsymbol{V}$的一组基，则$\forall\boldsymbol{\beta}\in V$，存在唯一的一组数$x_1,x_2,\cdots,x_d$，使得

$$\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_d\boldsymbol{\alpha}_d=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_d\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_d\end{bmatrix}$$

则称$[x_1,x_2,\cdots,x_d]^{\mathrm{T}}$为$\boldsymbol{\beta}$在基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$下的坐标。

基变换与坐标变换

向量空间中，基是不唯一的，在实际应用中，我们可根据需要用某个指标来衡量各组基的优劣，从而选择最适合需要的基。因此，这就可能涉及到基变换的问题。

设$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$、$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\}$是$d$维向量空间$V_d$的两组基，因此它们等价，所以存在方阵$\boldsymbol{C}_{d\times d}$，使得

$$\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_d\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_d\end{bmatrix}\boldsymbol{C}_{d\times d}$$

定义 4.13　**过渡矩阵**

设$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$、$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\}$是$d$维向量空间$\boldsymbol{V}_d$的两组基，必有方阵矩阵$\boldsymbol{C}_{d\times d}$，使得

$$\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_d\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_d\end{bmatrix}\boldsymbol{C}_{d\times d}$$

称$\boldsymbol{C}_{d\times d}$为从基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$到$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\}$的过渡矩阵（基变换矩阵）。

定理 4.15　过渡矩阵是可逆的。

证明：反证法。设$\boldsymbol{C}$为从基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$到$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\}$的过渡矩阵，故有

$$\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_d\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_d\end{bmatrix}\boldsymbol{C}_{d\times d}$$

如果 $\boldsymbol{C}$ 不可逆，则以 $\boldsymbol{C}$ 为系数矩阵的齐次线性方程组

$$\boldsymbol{CX}=\boldsymbol{0}$$

有非零解，设一个非零解为 $\boldsymbol{X}_0=\left[x_1,x_2,\cdots,x_d\right]^{\mathrm{T}}$，因此

$$\begin{aligned}
x_1\boldsymbol{\beta}_1+x_2\boldsymbol{\beta}_2+\cdots+x_d\boldsymbol{\beta}_d &= \left[\boldsymbol{\beta}_1\ \ \boldsymbol{\beta}_2\ \ \cdots\ \ \boldsymbol{\beta}_d\right]\begin{bmatrix}x_1\\x_2\\\vdots\\x_d\end{bmatrix}\\
&=\left[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_d\right]\boldsymbol{C}\begin{bmatrix}x_1\\x_2\\\vdots\\x_d\end{bmatrix}\\
&=\left[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_d\right]\boldsymbol{CX}_0\\
&=\left[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_d\right]\left(\boldsymbol{CX}_0\right)\\
&=\left[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_d\right]\boldsymbol{0}\\
&=\boldsymbol{0}
\end{aligned}$$

这说明 $\left\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\right\}$ 线性相关，这与 $\left\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\right\}$ 为一组基矛盾。故 $\boldsymbol{C}$ 可逆。

我们也用另外一种方法来证明这个命题。设 D 为从基 $\left\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\right\}$ 到 $\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\right\}$ 的过渡矩阵，故有

$$\left[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_n\right]=\left[\boldsymbol{\beta}_1\ \ \boldsymbol{\beta}_2\ \ \cdots\ \ \boldsymbol{\beta}_n\right]D$$

因此

$$\begin{aligned}
\left[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_n\right]&=\left[\boldsymbol{\beta}_1\ \ \boldsymbol{\beta}_2\ \ \cdots\ \ \boldsymbol{\beta}_n\right]\boldsymbol{D}\\
&=\left[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_n\right]\boldsymbol{CD}
\end{aligned}$$

由于 $\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\right\}$ 为一组基，故每个向量在这组基下的坐标是唯一的，这说明

$$\boldsymbol{CD}=\boldsymbol{E}$$

因此，$\boldsymbol{C}$ 可逆，且 $\boldsymbol{C}^{-1}=\boldsymbol{D}$。 ■

过渡矩阵的一个重要作用是建立一个向量在两组基下的坐标之间的关系。

定理 4.16　设向量空间 V_d 一组基 $\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\right\}$ 到另一组基 $\left\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\right\}$ 的过渡矩阵为 $\boldsymbol{C}$，而向量 $\boldsymbol{\gamma}\in V$ 在这两组基下的坐标分别为 $\boldsymbol{X}$、$\boldsymbol{Y}$，则有

$$\boldsymbol{X}=\boldsymbol{CY}$$

证明： $\boldsymbol{C}$ 为从基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$ 到 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\}$ 的过渡矩阵，因此

$$[\boldsymbol{\beta}_1\ \ \boldsymbol{\beta}_2\ \ \cdots\ \ \boldsymbol{\beta}_d]=[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_d]\boldsymbol{C}$$

由于 $\boldsymbol{\gamma}$ 在这两组基下的坐标分别为 $\boldsymbol{X}$ 、$\boldsymbol{Y}$ ，因而

$$\boldsymbol{\gamma}=[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_d]\boldsymbol{X}$$

$$\boldsymbol{\gamma}=[\boldsymbol{\beta}_1\ \ \boldsymbol{\beta}_2\ \ \cdots\ \ \boldsymbol{\beta}_d]\boldsymbol{Y}$$

故

$$\begin{aligned}&[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_d]\boldsymbol{X}=[\boldsymbol{\beta}_1\ \ \boldsymbol{\beta}_2\ \ \cdots\ \ \boldsymbol{\beta}_d]\boldsymbol{Y}\\&=[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_d]\boldsymbol{CY}\end{aligned}$$

由表示的唯一性得

$$\boldsymbol{X}=\boldsymbol{CY}$$

■

例 4.12　设 $\mathbb{R}^3$ 的两组基为 $\boldsymbol{\alpha}_1=[1,0,-1]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=[2,1,1]^{\mathrm{T}}$，$\boldsymbol{\alpha}_3=[1,1,1]^{\mathrm{T}}$ 和 $\boldsymbol{\beta}_1=[0,1,1]^{\mathrm{T}}$，$\boldsymbol{\beta}_2=[-1,1,0]^{\mathrm{T}}$，$\boldsymbol{\beta}_3=[1,2,1]^{\mathrm{T}}$。

(1) 求从基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$ 到基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}$ 的过渡矩阵；

(2) 求向量 $\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2-3\boldsymbol{\alpha}_3$ 在基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}$ 下的坐标；

解 (1) 设 $\boldsymbol{C}$ 为从基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$ 到基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}$ 的过渡矩阵，因此

$$[\boldsymbol{\beta}_1\ \ \boldsymbol{\beta}_2\ \ \boldsymbol{\beta}_3]=[\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \boldsymbol{\alpha}_3]\boldsymbol{C}$$

即：

$$\begin{bmatrix}0&-1&1\\1&1&2\\1&0&1\end{bmatrix}=\begin{bmatrix}1&2&1\\0&1&1\\-1&1&1\end{bmatrix}\boldsymbol{C}$$

因此

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

由于向量 $\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2-3\boldsymbol{\alpha}_3$ 在 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$ 下的坐标为 $[1,2,-3]^{\mathrm{T}}$，因此它在基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}$ 下的坐标为

$$\boldsymbol{Y}=\boldsymbol{C}^{-1}\boldsymbol{X}=\begin{bmatrix}0&1&1\\-1&-3&-2\\2&4&4\end{bmatrix}^{-1}\begin{bmatrix}1\\2\\-3\end{bmatrix}=\frac{1}{2}\begin{bmatrix}-7\\-1\\3\end{bmatrix}$$

习题 4.3

1. 下列向量集合是否构成向量空间，若是，求其一组基及维数。

a) $V=\left\{[a,a,a,b]^{\mathrm{T}}\,\middle|\,a,b\in\mathbb{R}\right\}$；

b) $V=\left\{[1,a,b,c]^{\mathrm{T}}\,\middle|\,a,b,c\in\mathbb{R}\right\}$；

c) $V=\left\{[a,2a,3a,4a]^{\mathrm{T}}\,\middle|\,a\in\mathbb{R}\right\}$；

d) $V=\left\{[0,a,b,c]^{\mathrm{T}}\,\middle|\,a,b,c\in\mathbb{R}\right\}$；

e) $V=\left\{[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\,\middle|\,\sum_{i=1}^{n}x_i=0,\ x_1,x_2,\cdots,x_n\in\mathbb{R}\right\}$；

f) $V=\left\{[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\,\middle|\,\sum_{i=1}^{n}x_i=1,\ x_1,x_2,\cdots,x_n\in\mathbb{R}\right\}$；

g) $V=\left\{[x_1,x_2,x_3]^{\mathrm{T}}\,\middle|\,x_1=\frac{1}{2}x_2=\frac{1}{3}x_3,\ x_1,x_2,x_3\in\mathbb{R}\right\}$

2. 设 $\boldsymbol{\alpha}_1=[1,1,0,0]^{\mathrm{T}},\boldsymbol{\alpha}_2=[1,0,1,1]^{\mathrm{T}},\boldsymbol{\beta}_1=[2,-1,3,3]^{\mathrm{T}},\boldsymbol{\beta}_2=[0,1,-1,-1]^{\mathrm{T}}$，记 $V_1=L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2),V_2=L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)$，证明 $V_1=V_2$。

3. 已知三维线性空间的一个基为 $\boldsymbol{\alpha}_1=[1,1,0]^{\mathrm{T}},\boldsymbol{\alpha}_2=[1,0,1]^{\mathrm{T}},\boldsymbol{\alpha}_3=[0,1,1]^{\mathrm{T}}$。求 $\boldsymbol{\alpha}=[2,0,0]^{\mathrm{T}}$ 在这个基下的坐标.

4. 已知 $\mathbb{R}^3$ 的两个基为 $\boldsymbol{\alpha}_1=[1,1,1]^{\mathrm{T}},\boldsymbol{\alpha}_2=[1,0,-1]^{\mathrm{T}},\boldsymbol{\alpha}_3=[1,0,1]^{\mathrm{T}}$ 和 $\boldsymbol{\beta}_1=[1,2,1]^{\mathrm{T}},\boldsymbol{\beta}_2=[2,3,4]^{\mathrm{T}},\boldsymbol{\beta}_3=[3,4,3]^{\mathrm{T}}$，求由基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 到基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 的过渡矩阵。

5. 已知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 是 4 维向量空间 V 的一个基，$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4$，$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4$，$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4$，$\boldsymbol{\beta}_4=\boldsymbol{\alpha}_4$

（1） 证明 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$ 是 V 的一个基；

（2） 求由基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$ 到基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的过渡矩阵；

（3） 求在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$和基$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$下坐标相同的向量.

6. 已知$\mathbb{R}^3$的向量$\boldsymbol{\gamma}$在基$\boldsymbol{\alpha}_1=[1,0,1]^{\mathrm{T}},\boldsymbol{\alpha}_2=[1,1,1]^{\mathrm{T}},\boldsymbol{\alpha}_3=[1,0,0]^{\mathrm{T}}$下的坐标是$[1,0,-1]^{\mathrm{T}}$，求$\boldsymbol{\gamma}$在基$\boldsymbol{\beta}_1=[1,2,0]^{\mathrm{T}},\boldsymbol{\beta}_2=[1,-1,2]^{\mathrm{T}},\boldsymbol{\beta}_3=[0,1,-1]^{\mathrm{T}}$的坐标.

7. 设向量空间$\mathbb{R}^3$中两组基为：$\boldsymbol{\alpha}_1=[1,0,-1]^{\mathrm{T}},\boldsymbol{\alpha}_2=[2,1,1]^{\mathrm{T}},\boldsymbol{\alpha}_3=[1,1,1]^{\mathrm{T}}$和$\boldsymbol{\beta}_1=[0,1,1]^{\mathrm{T}},\boldsymbol{\beta}_2=[-1,1,0]^{\mathrm{T}},\boldsymbol{\beta}_3=[1,2,1]^{\mathrm{T}}$，

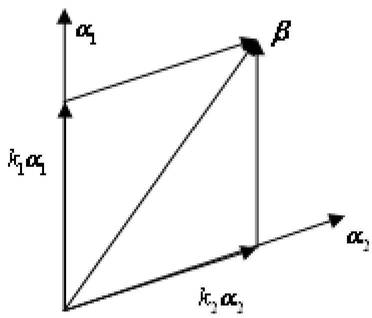
（1）求从基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$到基$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$的过渡矩阵；

（2）求向量$\boldsymbol{\alpha}=3\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$在基$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$下的坐标。

§4.4　内积空间与正交基

在本章第二节中解决了线性表示的最简性问题，这一节我们讨论线性表示的方便性。

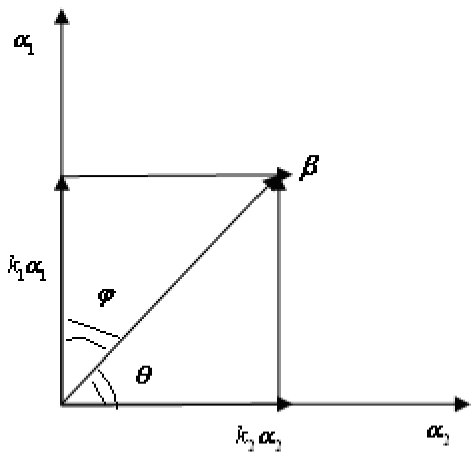
我们知道，$n(n\le 3)$向量加法遵循并行四边形法则。以 2 维向量为例，已知$\boldsymbol{\beta}$可由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$线性表示，为求$k_1,k_2$，使得$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2$，从所学知识可知，可通过求解以$[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\boldsymbol{\beta}]$为增广矩阵的线性方程组来得到$k_1,k_2$。但我们也可以从几何上考虑该问题。显然，$k_1\boldsymbol{\alpha}_1,k_2\boldsymbol{\alpha}_2,\boldsymbol{\beta}$满足平行四边形法则，$k_1\boldsymbol{\alpha}_1,k_2\boldsymbol{\alpha}_2$是分别与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$平行的向量，其长度分别为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$的$k_1,k_2$倍，因此为求$k_1$，只要求得$k_1\boldsymbol{\alpha}_1$，然后将$k_1\boldsymbol{\alpha}_1$的长度与$\boldsymbol{\alpha}_1$的长度比较可知$|k|$，如果$k_1\boldsymbol{\alpha}_1$与$\boldsymbol{\alpha}_1$同向，则$k_1>0$，否则$k_1\le 0$；类似可讨论$k_2$。



一般情况下，$k_1\boldsymbol{\alpha}_1,k_2\boldsymbol{\alpha}_2$不容易求，这种几何方法并不具有多大优越性，但是当$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$这

两个向量垂直时，$k_1\boldsymbol{\alpha}_1, k_2\boldsymbol{\alpha}_2$ 则容易求，显然 $\|k_1\boldsymbol{\alpha}_1\| = \|\boldsymbol{\beta}\|\cos\theta$，故 $|k_1| = \dfrac{\|k_1\boldsymbol{\alpha}_1\|}{\|\boldsymbol{\alpha}_1\|} = \dfrac{\|\boldsymbol{\beta}\|\cos\theta}{\|\boldsymbol{\alpha}_1\|}$，

再根据 $k_1\boldsymbol{\alpha}_1$ 与 $\boldsymbol{\alpha}_1$ 方向的关系确定 k_1 的符号。k_2 可类似地确定。



实际上，由于 $\boldsymbol{\beta} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2$，且 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 垂直，根据向量点积的定义可知这两个向量之间的内积 $\boldsymbol{\alpha}_1 \bullet \boldsymbol{\alpha}_2 = 0$，因此

$$\boldsymbol{\beta}\bullet\boldsymbol{\alpha}_1 = (k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2)\bullet\boldsymbol{\alpha}_1 = k_1\boldsymbol{\alpha}_1\bullet\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2\bullet\boldsymbol{\alpha}_1 = k_1\|\boldsymbol{\alpha}_1\|^2$$

故得到 $k_1 = \dfrac{\boldsymbol{\beta}\bullet\boldsymbol{\alpha}_1}{\|\boldsymbol{\alpha}_1\|^2}$；类似可得 $k_2 = \dfrac{\boldsymbol{\beta}\bullet\boldsymbol{\alpha}_2}{\|\boldsymbol{\alpha}_2\|^2}$。在这个结论中，向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 的点积为零起到了非常重要的作用，我们知道，在直角坐标系表示下，向量 $\boldsymbol{\alpha}(x_1, y_1), \boldsymbol{\beta}(x_2, y_2)$ 的内积为 $\boldsymbol{\alpha}\bullet\boldsymbol{\beta} = x_1x_2 + y_1y_2$。我们能否能够上述情形推广到一般的 n 维向量？

（过渡过程应该进一步完善，似乎还不够自然）

定义 4.14　　**内积，欧氏空间**

$\boldsymbol{\alpha} = [x_1, x_2, \cdots, x_n]^{\mathrm{T}}, \boldsymbol{\beta} = [y_1, y_2, \cdots, y_n]^{\mathrm{T}} \in \mathbb{R}^n$ 的内积（点积）$(\boldsymbol{\alpha}, \boldsymbol{\beta})$ 定义为

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta} \tag{4.3}$$

定义了内积的向量空间 $\mathbb{R}^n$ 称为**欧式空间**。

容易验证，上述定义的 $\mathbb{R}^n$ 上的内积具有如下性质：

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

(2) 线性性：$(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta})=(\boldsymbol{\alpha}_1,\boldsymbol{\beta})+(\boldsymbol{\alpha}_2,\boldsymbol{\beta})$，

$$(k\boldsymbol{\alpha},\boldsymbol{\beta})=k(\boldsymbol{\alpha},\boldsymbol{\beta}),\qquad k\in\mathbb{R}\text{；}$$

(3) 非负性：$(\boldsymbol{\alpha},\boldsymbol{\alpha})\geq 0$，且$(\boldsymbol{\alpha},\boldsymbol{\alpha})=0\Leftrightarrow\boldsymbol{\alpha}=\mathbf{0}$。

定义 4.15　　向量**长度**

$\boldsymbol{\alpha}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\in\mathbb{R}^n$的长度（范数）$\|\alpha\|$为

$$\|\boldsymbol{\alpha}\|=\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}\tag{4.4}$$

所以$\|\boldsymbol{\alpha}\|\geq 0$，$\|\boldsymbol{\alpha}\|=0\Leftrightarrow\boldsymbol{\alpha}=\mathbf{0}$，且$\|k\boldsymbol{\alpha}\|=|k|\|\boldsymbol{\alpha}\|$。

定义 4.16　　**单位向量**

长度为 1 的向量称为单位向量。

显然，当$\boldsymbol{\alpha}\neq\mathbf{0}$时，则$\left\|\dfrac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}\right\|=1$，我们将向量$\boldsymbol{\alpha}$化为单位向量$\dfrac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}$的过程称为向量的单位化。

关于向量内积、长度的一个重要定理为柯西（Cauchy）不等式。

定理 4.17　　柯西（Cauchy）不等式

设$\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{R}^n$，则有

$$|(\boldsymbol{\alpha},\boldsymbol{\beta})|^2\leq(\boldsymbol{\alpha},\boldsymbol{\alpha})(\boldsymbol{\beta},\boldsymbol{\beta})\tag{4.5}$$

其中等号当且仅当$\boldsymbol{\alpha}=s\boldsymbol{\beta}\,(s\in\mathbb{R})$时成立。

证明：如果$\boldsymbol{\alpha}=\mathbf{0}$或$\boldsymbol{\beta}=\mathbf{0}$，结论显然成立。设$\boldsymbol{\beta}\neq\mathbf{0}$，取向量$\boldsymbol{\alpha}-k\boldsymbol{\beta}$，其中$k$为任意实数，那么

$$\begin{aligned}
0&\leq(\boldsymbol{\alpha}-k\boldsymbol{\beta},\boldsymbol{\alpha}-k\boldsymbol{\beta})\\
&=(\boldsymbol{\alpha}-k\boldsymbol{\beta},\boldsymbol{\alpha})+(\boldsymbol{\alpha}-k\boldsymbol{\beta},-k\boldsymbol{\beta})\\
&=(\boldsymbol{\alpha},\boldsymbol{\alpha})+(-k\boldsymbol{\beta},\boldsymbol{\alpha})+(\boldsymbol{\alpha},-k\boldsymbol{\beta})+(-k\boldsymbol{\beta},-k\boldsymbol{\beta})\\
&=(\boldsymbol{\alpha},\boldsymbol{\alpha})-k(\boldsymbol{\beta},\boldsymbol{\alpha})-k(\boldsymbol{\alpha},\boldsymbol{\beta})+k^2(\boldsymbol{\beta},\boldsymbol{\beta})\\
&=(\boldsymbol{\alpha},\boldsymbol{\alpha})-2k(\boldsymbol{\beta},\boldsymbol{\alpha})+k^2(\boldsymbol{\beta},\boldsymbol{\beta})
\end{aligned}\tag{4.6}$$

取$k=\dfrac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\beta},\boldsymbol{\beta})}$，则有

$$0\le(\alpha,\alpha)-2\frac{(\alpha,\beta)}{(\beta,\beta)}(\beta,\alpha)+\left(\frac{(\alpha,\beta)}{(\beta,\beta)}\right)^2(\beta,\beta)$$
$$=(\alpha,\alpha)-2\frac{(\alpha,\beta)^2}{(\beta,\beta)}+\frac{(\alpha,\beta)^2}{(\beta,\beta)}$$
$$=(\alpha,\alpha)-\frac{(\alpha,\beta)^2}{(\beta,\beta)}$$

所以

$$|(\alpha,\beta)|^2\le(\alpha,\alpha)(\beta,\beta) \tag{4.7}$$

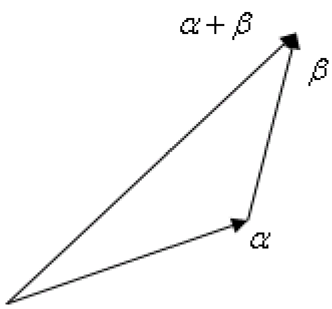
故(4.5)成立。容易验证，当$\boldsymbol{\alpha}=s\boldsymbol{\beta}$时，(4.5)中的等号成立；反之，若(4.7)中的等号成立，由证明过程可知(4.6)中有$0=(\boldsymbol{\alpha}-k\boldsymbol{\beta},\boldsymbol{\alpha}-k\boldsymbol{\beta})$，其中$k=\dfrac{(\alpha,\beta)}{(\beta,\beta)}$，根据内积的性质可知$\boldsymbol{\alpha}-k\boldsymbol{\beta}=\mathbf{0}$。命题得证。

■

关于向量的长度，有如下的三角不等式

$$\|\alpha+\beta\|\le\|\alpha\|+\|\beta\|$$

（该不等式的名字在$\mathbb{R}^n\ (n\le3)$中可用如下语句来表示：“两点之间直线最短”。



而对于一般的情况，我们可利用柯西不等式对该式加以证明。）

证明：

$$\|\alpha+\beta\|\le\|\alpha\|+\|\beta\|$$
$$\Leftrightarrow\|\alpha+\beta\|^2\le(\|\alpha\|+\|\beta\|)^2$$
$$\Leftrightarrow(\alpha+\beta,\alpha+\beta)\le\|\alpha\|^2+2\|\alpha\|\bullet\|\beta\|+\|\beta\|^2$$
$$\Leftrightarrow(\alpha,\alpha)+2(\alpha,\beta)+(\beta,\beta)\le(\alpha,\alpha)+2\|\alpha\|\bullet\|\beta\|+(\beta,\beta)$$
$$\Leftrightarrow(\alpha,\beta)\le\|\alpha\|\bullet\|\beta\|$$

而最后一式显然成立。

■

根据柯西不等式，我们有

$$\left|(\boldsymbol{\alpha},\boldsymbol{\beta})\right| \le \|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|$$

故当$\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{R}^n$为非零向量时，

$$\theta = \mathbf{arccos}\left(\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|}\right)$$

有意义，我们将之定义为$\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{R}^n$之间的**夹角**，记为$\angle(\boldsymbol{\alpha},\boldsymbol{\beta})$。

内积空间中的一个重要概念就是正交。

定义 4.17　　**向量正交**

若$\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{R}^n$且$(\boldsymbol{\alpha},\boldsymbol{\beta})=0$，则称向量$\boldsymbol{\alpha},\boldsymbol{\beta}$**正交**，记做$\boldsymbol{\alpha}\perp\boldsymbol{\beta}$。

由于若$(\boldsymbol{\alpha},\boldsymbol{\beta})=0$，则$\boldsymbol{\alpha},\boldsymbol{\beta}$之间的夹角为$\dfrac{\pi}{2}$，因此，正交就是$\mathbb{R}^2$、$\mathbb{R}^3$中向量垂直概念的推广。

我们将正交概念推广到多个向量的情形，就有正交向量组的概念。

定义 4.18　　**正交向量组、标准正交向量组**

向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in\mathbb{R}^n$，若它们两两正交，即

$$(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=0,\qquad i\ne j$$

则称向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$为**正交向量组**。如果正交向量组中的每一个向量都是单位向量，即

$$\|\boldsymbol{\alpha}_i\|=1,\qquad i=1,2,\cdots,r$$

则该向量组称为**标准正交向量组**或**规范正交向量组**。

正交向量组的应用。

若已知向量$\boldsymbol{\beta}$可由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性表示，求$\boldsymbol{\beta}$在该向量组下的组合系数，即求$k_1,k_2,\cdots,k_r$，使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r$$

这可通过求解增广矩阵为$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r & \boldsymbol{\beta}\end{bmatrix}$的线性方程组得到组合系数。而若

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 为正交向量组，则组合系数很容易求，例如为求 $\boldsymbol{k}_1$，可将 $\boldsymbol{\beta}$ 与 $\boldsymbol{\alpha}_1$ 作内积得

$$\begin{aligned}(\boldsymbol{\beta},\boldsymbol{\alpha}_1)&=(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_1)\\&=k_1(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)+k_2(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1)+\cdots+k_r(\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_1)\\&=k_1(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)\end{aligned}$$

若 $\boldsymbol{\alpha}_1\neq\boldsymbol{0}$，则 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)\neq 0$，所以

$$k_1=\frac{(\boldsymbol{\beta},\boldsymbol{\alpha}_1)}{(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)}$$

类似可得

$$k_i=\frac{(\boldsymbol{\beta},\boldsymbol{\alpha}_i)}{(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)},\qquad i=1,2,\cdots,r$$

故此时有

$$\boldsymbol{\beta}=\frac{(\boldsymbol{\beta},\boldsymbol{\alpha}_1)}{(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)}\boldsymbol{\alpha}_1+\frac{(\boldsymbol{\beta},\boldsymbol{\alpha}_2)}{(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2)}\boldsymbol{\alpha}_2+\cdots+\frac{(\boldsymbol{\beta},\boldsymbol{\alpha}_r)}{(\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_r)}\boldsymbol{\alpha}_r$$

而如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 为标准正交向量组，还有 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)=1$，因而有

$$k_i=(\boldsymbol{\beta},\boldsymbol{\alpha}_i),\qquad i=1,2,\cdots,r$$

因此

$$\boldsymbol{\beta}=(\boldsymbol{\beta},\boldsymbol{\alpha}_1)\boldsymbol{\alpha}_1+(\boldsymbol{\beta},\boldsymbol{\alpha}_2)\boldsymbol{\alpha}_2+\cdots+(\boldsymbol{\beta},\boldsymbol{\alpha}_r)\boldsymbol{\alpha}_r$$

因此，很容易求得组合系数。下面将这个结果可推广到向量空间的基、向量在一组基下的坐标上去。

问题 1：对于向量空间 $\mathbb{R}^n$，是否能够找到一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$，使得 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为一个正交向量组？或为标准正交向量组？进而如何找一个这样的(标准)正交向量组？

显然，若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为满足要求正交向量组，则将该向量组中的每个向量单位化就得满足的标准正交向量组。

正交向量组有一个重要特性，这就是如下定理。

定理 4.18　　若正交向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 中不含有零向量，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关。

证明：若有

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r=\boldsymbol{0}$$

因此，根据以上部分得到的结果，有

$$k_i=\frac{(\mathbf{0},\boldsymbol{\alpha}_i)}{(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)}=0, \qquad i=1,2,\cdots,r$$

所以$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性无关。

因此，若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\in\mathbb{R}^n$为不含有零向量的正交向量组，那么它就是$\mathbb{R}^n$的一组基，称为$\mathbb{R}^n$的**正交基**。而称$\mathbb{R}^n$中的标准正交向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$为$\mathbb{R}^n$的**标准正交基**，即：

$$(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases} \tag{4.8}$$

但是，在$\mathbb{R}^n$中是否存在正交基？显然，标准向量组$\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$就是一组正交基，$\mathbb{R}^n$中是否还有其它正交基？我们可如下讨论该问题。

例 4.13　设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$是$\mathbb{R}^n$中的正交向量组，若$r<n$，则存在非零向量$\boldsymbol{X}$，使得$\boldsymbol{X}$与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$都正交，即$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{X}$为正交向量组。

解：由于要求$\boldsymbol{X}$与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$都正交，由于

$$(\boldsymbol{\alpha}_i,\boldsymbol{X})=\mathbf{0}=\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{X}, \qquad i=1,2,\cdots,r,$$

因此可将上述条件写成矩阵形式

$$\begin{bmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\ \boldsymbol{\alpha}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\alpha}_r^{\mathrm{T}}\end{bmatrix}\boldsymbol{X}=\mathbf{0},$$

这是一个齐次线性方程组，系数矩阵

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\ \boldsymbol{\alpha}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\alpha}_r^{\mathrm{T}}\end{bmatrix}$$

是一个$r\times n$矩阵，由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$是不含零向量的正交向量组，故$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性无关，因此$r(\boldsymbol{A})=r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}=r$，故当$r<n$时，$r(\boldsymbol{A})=r<n$，该方程组有非零解，该方程组的任意非零解都满足题目要求。故命题成立。

■

从以上结论可得到寻找$\mathbb{R}^n$的一个正交基的方法：

首先找一个非零向量$\boldsymbol{\alpha}_1$，由例 4.12 的方法寻找非零向量$\boldsymbol{\alpha}_2$，使得$\boldsymbol{\alpha}_2$与$\boldsymbol{\alpha}_1$正交，得到正交向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$；再由例 4.12 的方法寻找与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$都正交的非零向量$\boldsymbol{\alpha}_3$，则得到正交向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$；⋯；最后得到正交向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_n$，该向量组就是$\mathbb{R}^n$的一组基，就是$\mathbb{R}^n$的一组正交基。这是一个“由无到有”的过程。

上述方法说明了如何构造$\mathbb{R}^n$的正交基，但是向量空间$V\neq\mathbb{R}^n$中是否有正交基？直接用上述方法并不能简单地回答这个问题，这是因为在上述方法中，需要用求解线性方程组的方法求得与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$正交的向量$\boldsymbol{X}$，但是如何保证求得的向量$\boldsymbol{X}\in V$？说明这一点相对比较繁琐，在此不予赘述。

另外，例 4.12 中的证明方法用到了内积的定义(4.3)，为了将该结论推广到一般的内积空间，我们用另外一种方法来证明该结论，该方法只用到内积的性质，而不用到$\mathbb{R}^n$内积的具体定义，并能够简要回答$V\neq\mathbb{R}^n$中正交基是否存在的问题，这就是施密特正交化方法，其构造方法是一个“从有到有”的过程。

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$是$\mathbb{R}^n$的一组基，则可从$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$出发得到$\mathbb{R}^n$的一组(标准)正交基，这就是 Schmidt 正交化方法。

定理 4.19 (Schmidt **正交化方法**)　设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$是欧氏空间$\mathbb{R}^n$中线性无关的向量组，则由如下方法：

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1;$$

$$\begin{aligned}\boldsymbol{\beta}_k&=\boldsymbol{\alpha}_k-\frac{(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2-\cdots-\frac{(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_{k-1})}{(\boldsymbol{\beta}_{k-1},\boldsymbol{\beta}_{k-1})}\boldsymbol{\beta}_{k-1}\\&=\boldsymbol{\alpha}_k-\sum_{i=1}^{k-1}\frac{(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_i)}{(\boldsymbol{\beta}_i,\boldsymbol{\beta}_i)}\boldsymbol{\beta}_i\end{aligned},\qquad k=2,\cdots,r;$$

所得的向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$为正交向量组。

证明：用归纳法。

$$(\boldsymbol{\beta}_2,\boldsymbol{\beta}_1)=\left(\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1,\boldsymbol{\beta}_1\right)$$

$$=(\alpha_2,\alpha_1)-\left(\frac{(\alpha_2,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1,\alpha_1\right)$$

$$=(\alpha_2,\alpha_1)-\frac{(\alpha_2,\alpha_1)}{(\alpha_1,\alpha_1)}(\alpha_1,\alpha_1)$$

$$=(\alpha_2,\alpha_1)-(\alpha_2,\alpha_1)=0$$

所以 β_1,β_2 正交。设向量组 $\beta_1,\beta_2,\cdots,\beta_{k-1}$ 相互正交，下面证明 β_k 与 $\beta_1,\beta_2,\cdots,\beta_{k-1}$ 中的每个向量正交。任取 $j<k$，

$$(\beta_k,\beta_j)=\left(\alpha_r-\sum_{i=1}^{k-1}\frac{(\alpha_k,\beta_i)}{(\beta_i,\beta_i)}\beta_i,\beta_j\right)$$

$$=(\alpha_k,\beta_j)-\sum_{i=1}^{k-1}\frac{(\alpha_k,\beta_i)}{(\beta_i,\beta_i)}(\beta_i,\beta_j)$$

$$(\beta_k,\beta_j)=(\alpha_k,\beta_j)-\sum_{i=1}^{k-1}\frac{(\alpha_k,\beta_i)}{(\beta_i,\beta_i)}(\beta_i,\beta_j)$$

$$=(\alpha_k,\beta_j)-\frac{(\alpha_k,\beta_j)}{(\beta_j,\beta_j)}(\beta_j,\beta_j)$$

$$=(\alpha_r,\beta_j)-(\alpha_r,\beta_j)$$
$$=0$$

所以， β_k 与 $\beta_1,\beta_2,\cdots,\beta_{k-1}$ 正交。由归纳法可知，向量组 $\beta_1,\beta_2,\cdots,\beta_k$ 是正交向量组。

■

从上述 Schmidt 正交化的公式可知， β_i 是 $\alpha_1,\alpha_2,\cdots,\alpha_i$ 的线性组合，且有

$$\beta_k=\alpha_k-\sum_{i=1}^{k-1}\frac{(\alpha_k,\beta_i)}{(\beta_i,\beta_i)}\beta_i=\alpha_k-\left(d_{k1}\alpha_1+d_{k2}\alpha_2+\cdots+d_{k,k-1}\alpha_{k-1}\right)$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关，所以

$$\beta_k\neq\mathbf{0},\qquad k=1,\cdots,r$$

且有

$$\begin{bmatrix}\beta_1 & \beta_2 & \cdots & \beta_r\end{bmatrix}=\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_r\end{bmatrix}\begin{bmatrix}1 & -d_{21} & \cdots & -d_{r-1,1} & -d_{r1}\\ & 1 & & -d_{r-1,2} & -d_{r2}\\ & & \ddots & & \vdots\\ & & & 1 & -d_{r,r-1}\\ & & & & 1\end{bmatrix}$$

由于r阶方阵

$$\begin{bmatrix}1 & -d_{21} & \cdots & -d_{r-1,1} & -d_{r1}\\ & 1 & & -d_{r-1,2} & -d_{r2}\\ & & \ddots & & \vdots\\ & & & 1 & -d_{r,r-1}\\ & & & & 1\end{bmatrix}$$

是一个上三角矩阵，且对角线上元素为 1，所以其可逆。因此，$\alpha_1,\alpha_2,\cdots,\alpha_r$ 与 $\beta_1,\beta_2,\cdots,\beta_r$ 等价。故若$\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关，$\beta_1,\beta_2,\cdots,\beta_r$ 也线性无关，故正交向量组$\beta_1,\beta_2,\cdots,\beta_r$ 中没有零向量。反若$\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性相关，因而$\beta_1,\beta_2,\cdots,\beta_r$ 也线性相关，正交向量组 $\beta_1,\beta_2,\cdots,\beta_r$ 中一定有零向量。

当$r=n$时，可从$\mathbb{R}^n$的一组基$\alpha_1,\alpha_2,\cdots,\alpha_n$出发，通过 Schmidt 过程得到一个正交向量组$\beta_1,\beta_2,\cdots,\beta_n$， $\beta_1,\beta_2,\cdots,\beta_n$线性无关，故$\beta_1,\beta_2,\cdots,\beta_n$为一个正交基。再将所得的正交向量组$\beta_1,\beta_2,\cdots,\beta_n$单位化，就可得到一组标准正交基。

当然，也可正交化、单位化同步进行：

$$\begin{cases}\beta_1=\alpha_1 & \varepsilon_1=\dfrac{\beta_1}{\sqrt{(\beta_1,\beta_1)}}\\ \beta_2=\alpha_2-(\alpha_2,\varepsilon_1)\varepsilon_1 & \varepsilon_2=\dfrac{\beta_2}{\sqrt{(\beta_2,\beta_2)}}\\ \quad\vdots & \\ \beta_n=\alpha_n-\sum\limits_{i=1}^{n-1}(\alpha_n,\varepsilon_i)\varepsilon_i & \varepsilon_n=\dfrac{\beta_n}{\sqrt{(\beta_n,\beta_n)}}\end{cases}$$

一般来讲，先正交化、再标准化计算较为简单。

我们需要注意到上述两种求正交向量组（正交基）方法的差异：第一种方法运用的$\mathbb{R}^n$中内积的定义，而第二种方法只用到了内积的性质，因此，第二种方法可应用到一般的内积空间。

定理 4.20　（1）（勾股定理）设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是正交向量组，则

$$\boldsymbol{\beta}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_r=\sum_{i=1}^{r}\boldsymbol{\alpha}_i$$

的长度为

$$\|\boldsymbol{\beta}\|=\sqrt{\|\boldsymbol{\alpha}_1\|^2+\|\boldsymbol{\alpha}_2\|^2+\cdots+\|\boldsymbol{\alpha}_r\|^2}\ ;$$

（2）若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是标准正交向量组，则向量

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r=\sum_{i=1}^{r}k_i\boldsymbol{\alpha}_i$$

的长度为 $\|\boldsymbol{\beta}\|=\sqrt{k_1^2+k_2^2+\cdots+k_r^2}$ 。

（3）如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为 $\mathbb{R}^n$ 的一组基，而 $\boldsymbol{\beta}$ 与该组基中的每个向量都正交，即

$$(\boldsymbol{\beta},\boldsymbol{\alpha}_i)=0,\qquad i=1,2,\cdots,n$$

则有 $\boldsymbol{\beta}=\mathbf{0}$ 。

证明　（1）

$$\|\boldsymbol{\beta}\|^2=(\boldsymbol{\beta},\boldsymbol{\beta})=\left(\sum_{i=1}^{r}\boldsymbol{\alpha}_i,\sum_{j=1}^{r}\boldsymbol{\alpha}_j\right)$$
$$=\sum_{i=1}^{r}\sum_{j=1}^{r}(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=\sum_{i=1}^{r}\left(\sum_{j=1}^{r}(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)\right)=\sum_{i=1}^{r}\|\boldsymbol{\alpha}_i\|^2$$

（2）利用（1）所得的结论即可。由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是标准正交向量组，故 $k_1\boldsymbol{\alpha}_1,k_2\boldsymbol{\alpha}_2,\cdots,k_r\boldsymbol{\alpha}_r$ 也是正交向量组，因此根据（1）的结论有

$$\|\boldsymbol{\beta}\|^2=\sum_{i=1}^{r}\|k_i\boldsymbol{\alpha}_i\|^2=\sum_{i=1}^{r}k_i^2\|\boldsymbol{\alpha}_i\|^2=\sum_{i=1}^{r}k_i^2=k_1^2+k_2^2+\cdots+k_r^2\ 。$$

命题得证。

（3）由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为 $\mathbb{R}^n$ 的一组基，$\boldsymbol{\beta}\in\mathbb{R}^n$ ，故 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示，因而存在 $k_1,k_2,\cdots,k_n\in\mathbb{R}^n$，使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i\ ,$$

从而$\left(\boldsymbol{\beta},\boldsymbol{\beta}\right)=\left(\boldsymbol{\beta},\sum_{i=1}^{n}k_i\boldsymbol{\alpha}_i\right)=\sum_{i=1}^{n}k_i\left(\boldsymbol{\beta},\boldsymbol{\alpha}_i\right)=\sum_{i=1}^{n}k_i 0=0$，根据内积的性质有$\boldsymbol{\beta}=\mathbf{0}$。

■

在欧氏空间$\mathbb{R}^n$中，我们可以推广正交基的概念，这就是对偶基（双正交基）的概念。

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$是向量空间$\mathbb{R}^n$的一组非正交基，$\forall\boldsymbol{\gamma}\in\mathbb{R}^n$，存在$k_1,k_2,\cdots,k_n$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\boldsymbol{\gamma},$$

由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$不是正交基，因此$k_1,k_2,\cdots,k_n$可通过求解线性方程组来得到。借助于正交基的思想，可考虑如下问题：能否找到一个向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$，满足

$$\begin{aligned}\left(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j\right)&=0, & i\neq j\\ \left(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j\right)&\neq 0, & i=j\end{aligned}\tag{4.9}$$

？若$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$满足该条件，则将

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\boldsymbol{\gamma}$$

两边同时与$\boldsymbol{\beta}_j$作内积，得到

$$k_1\left(\boldsymbol{\alpha}_1,\boldsymbol{\beta}_j\right)+\cdots+k_{j-1}\left(\boldsymbol{\alpha}_{j-1},\boldsymbol{\beta}_j\right)+k_j\left(\boldsymbol{\alpha}_j,\boldsymbol{\beta}_j\right)+k_{j+1}\left(\boldsymbol{\alpha}_{j+1},\boldsymbol{\beta}_j\right)+\cdots+k_n\left(\boldsymbol{\alpha}_n,\boldsymbol{\beta}_j\right)=\left(\boldsymbol{\gamma},\boldsymbol{\beta}_j\right)$$

因而有

$$k_j\left(\boldsymbol{\alpha}_j,\boldsymbol{\beta}_j\right)=\left(\boldsymbol{\gamma},\boldsymbol{\beta}_j\right),$$

$$k_j=\frac{\left(\boldsymbol{\gamma},\boldsymbol{\beta}_j\right)}{\left(\boldsymbol{\alpha}_j,\boldsymbol{\beta}_j\right)},$$

这样就可求得$k_1,k_2,\cdots,k_n$。问题是：是否存在满足要求(4.9)的向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$？

下面说明满足(4.9)的向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$是存在的。根据前面的论述，对于向量组$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_n$，存在非零向量$\boldsymbol{\beta}_i$，使得

$$\left(\boldsymbol{\alpha}_j,\boldsymbol{\beta}_i\right)=0,\qquad j\neq i。$$

我们只需要证明$\left(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i\right)\neq 0$即可。这是显然的，因为若$\left(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i\right)=0$，则说明$\boldsymbol{\beta}_i$与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$中的每个向量正交，$\Rightarrow\boldsymbol{\beta}_i=\mathbf{0}$，这与$\boldsymbol{\beta}_i$的取法相矛盾。故满足要求的向量组

$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 存在。再把上述 $\boldsymbol{\beta}_i$ 规范化，使得

$$\left(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i\right)=1$$

这样求得的 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 满足

$$\begin{aligned}\left(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j\right)&=0, \quad i\neq j\\ \left(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j\right)&\neq 0, \quad i=j\end{aligned}$$

我们把上述的 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 称为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的对偶基。显然，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 也是 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 的对偶基。

实际上，若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为 $\mathbb{R}^n$ 的一组基，故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，因此方阵

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}$$

可逆，其逆矩阵 $\boldsymbol{A}^{-1}$ 存在，且

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

对 $\boldsymbol{A}^{-1}$ 做如下分块

$$\boldsymbol{A}^{-1}=\begin{bmatrix}\boldsymbol{\beta}_1^{\mathrm{T}}\\ \boldsymbol{\beta}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\beta}_n^{\mathrm{T}}\end{bmatrix}$$

则

$$\begin{aligned}\boldsymbol{A}^{-1}\boldsymbol{A}&=\begin{bmatrix}\boldsymbol{\beta}_1^{\mathrm{T}}\\ \boldsymbol{\beta}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\beta}_n^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\\ &=\begin{bmatrix}\boldsymbol{\beta}_1^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\beta}_1^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\beta}_1^{\mathrm{T}}\boldsymbol{\alpha}_n\\ \boldsymbol{\beta}_2^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\beta}_2^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\beta}_2^{\mathrm{T}}\boldsymbol{\alpha}_n\\ \vdots & \vdots & & \vdots\\ \boldsymbol{\beta}_n^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\beta}_n^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\beta}_n^{\mathrm{T}}\boldsymbol{\alpha}_n\end{bmatrix}=\boldsymbol{E}\end{aligned}$$

故

$$\boldsymbol{\beta}_i^{\mathrm{T}}\boldsymbol{\alpha}_j=\left(\boldsymbol{\beta}_i,\boldsymbol{\alpha}_j\right)=\delta_{ij}$$

因此 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的对偶基。

应用上面的符号，我们考虑$\forall \gamma \in \mathbb{R}^n$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的坐标$k_1,k_2,\cdots,k_n$（组合系数）的实质。

$$\begin{aligned}\gamma &= k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_n\boldsymbol{\alpha}_n \\ &= \begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\begin{bmatrix}k_1\\k_2\\\vdots\\k_n\end{bmatrix} = \boldsymbol{AK}\end{aligned}$$

因此$k_1,k_2,\cdots,k_n$是非齐次线性方程组$\boldsymbol{AK}=\gamma$的解。设$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$的对偶基，则$\gamma$在$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的坐标为$\begin{bmatrix}k_1\\k_2\\\vdots\\k_n\end{bmatrix}=\begin{bmatrix}(\boldsymbol{\beta}_1,\gamma)\\(\boldsymbol{\beta}_2,\gamma)\\\vdots\\(\boldsymbol{\beta}_n,\gamma)\end{bmatrix}=\begin{bmatrix}\boldsymbol{\beta}_1^{\mathrm{T}}\gamma\\\boldsymbol{\beta}_2^{\mathrm{T}}\gamma\\\vdots\\\boldsymbol{\beta}_n^{\mathrm{T}}\gamma\end{bmatrix}=\begin{bmatrix}\boldsymbol{\beta}_1^{\mathrm{T}}\\\boldsymbol{\beta}_2^{\mathrm{T}}\\\vdots\\\boldsymbol{\beta}_n^{\mathrm{T}}\end{bmatrix}\gamma=\boldsymbol{A}^{-1}\gamma$。因此，这实质上求解方程组$\boldsymbol{AX}=\gamma$的过程，由于此时$\boldsymbol{A}$可逆，故$\boldsymbol{K}=\boldsymbol{A}^{-1}\gamma$。

如果$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$是标准正交向量组，由(4.8)易知

$$\begin{bmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\\boldsymbol{\alpha}_2^{\mathrm{T}}\\\vdots\\\boldsymbol{\alpha}_n^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_n\\\boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{\alpha}_n\\\vdots & \vdots & & \vdots\\\boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{\alpha}_n\end{bmatrix}=\boldsymbol{E}$$

即矩阵$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}$的逆矩阵为$\boldsymbol{A}^{\mathrm{T}}$。

定义 4.19　　**正交矩阵**

若方阵$\boldsymbol{A}$满足$\boldsymbol{AA}^{\mathrm{T}}=\boldsymbol{E}$，则称$\boldsymbol{A}$为正交矩阵。

定理 4.21　　方阵$\boldsymbol{A}$为正交矩阵的充要条件是$\boldsymbol{A}$的列（行）向量组是标准正交向量组。

对偶基的存在使得计算一个向量在非正交基下的坐标这一任务显得比较简单，其优越性在$\mathbb{R}^n$中可能并不那么明显，而在一般的内积空间中就有巨大的作用。

对于一般的线性无关的向量组，我们也可定义与其对偶的向量组，并且可以类似地加以应用。这里不再赘述。

例 4.14　已知在三维向量空间$\mathbb{R}^3$中两个向量$\boldsymbol{\alpha}_1=[1,1,1]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=[1,-2,1]^{\mathrm{T}}$正交，试求一个非零向量$\boldsymbol{\alpha}_3$，使$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$两两正交。

解：记

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\ \boldsymbol{\alpha}_2^{\mathrm{T}}\end{bmatrix}=\begin{bmatrix}1 & 1 & 1\\ 1 & -2 & 1\end{bmatrix}$$

则 $\boldsymbol{\alpha}_3$ 应满足齐次线性方程组 $\boldsymbol{AX}=\mathbf{0}$，即

$$\begin{bmatrix}1 & 1 & 1\\ 1 & -2 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=0$$

由

$$\boldsymbol{A}=\begin{bmatrix}1 & 1 & 1\\ 1 & -2 & 1\end{bmatrix}\overset{r}{\sim}\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}$$

得到 $\begin{cases}x_1=-x_3\\ x_2=0\end{cases}$，因而可得一个非零解 $[-1,0,1]^{\mathrm{T}}$，取 $\alpha_3=[-1,0,1]^{\mathrm{T}}$ 即符合要求。

例 4.15　已知 $\boldsymbol{\alpha}_1=[1,2,-1]^{\mathrm{T}},\boldsymbol{\alpha}_2=[-1,3,1]^{\mathrm{T}},\boldsymbol{\alpha}_3=[4,-1,0]^{\mathrm{T}}$，使用施密特正交化过程把这组向量规范正交化。

解：根据施密特正交化方法，有

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1$$

$$\tilde{\boldsymbol{\beta}}_2=\boldsymbol{\alpha}_2-\frac{[\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1]}{[\boldsymbol{\beta}_1,\boldsymbol{\beta}_1]}\boldsymbol{\beta}_1=\begin{bmatrix}-1\\ 3\\ 1\end{bmatrix}-\frac{4}{6}\begin{bmatrix}1\\ 2\\ -1\end{bmatrix}=\begin{bmatrix}-\frac{5}{3}\\ \frac{5}{3}\\ \frac{5}{3}\end{bmatrix}=\frac{5}{3}\begin{bmatrix}-1\\ 1\\ 1\end{bmatrix}$$

取　$$\boldsymbol{\beta}_2=\begin{bmatrix}-1\\ 1\\ 1\end{bmatrix}$$

$$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\frac{[\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1]}{[\boldsymbol{\beta}_1,\boldsymbol{\beta}_1]}\boldsymbol{\beta}_1-\frac{[\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2]}{[\boldsymbol{\beta}_2,\boldsymbol{\beta}_2]}\boldsymbol{\beta}_2=\begin{bmatrix}4\\ -1\\ 0\end{bmatrix}-\frac{2}{6}\begin{bmatrix}1\\ 2\\ -1\end{bmatrix}-\frac{-5}{3}\begin{bmatrix}-1\\ 1\\ 1\end{bmatrix}$$

$$=\begin{bmatrix}2\\ 0\\ 2\end{bmatrix}=2\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}$$

再把它们单位化，得到

$$\gamma_1=\frac{\beta_1}{\|\beta_1\|}=\frac{1}{\sqrt{6}}\begin{bmatrix}1\\2\\-1\end{bmatrix},\quad \gamma_2=\frac{\beta_2}{\|\beta_2\|}=\frac{1}{\sqrt{3}}\begin{bmatrix}-1\\1\\1\end{bmatrix},\quad \gamma_3=\frac{\beta_3}{\|\beta_3\|}=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix}$$

例 4.16　　已知$\boldsymbol{\alpha}_1=[1,1,1]^{\mathrm{T}}$，求非零向量$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$，使得$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$两两正交。

解：由于$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$与$\boldsymbol{\alpha}_1$正交，所以$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$满足方程

$$\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{X}=0$$

即

$$x_1+x_2+x_3=0$$

取它的两个线性无关的解

$$\xi_1=[1,0,-1]^{\mathrm{T}},\xi_2=[0,1,-1]^{\mathrm{T}}$$

然后将该向量组正交化，即

$$\boldsymbol{\alpha}_2=\xi_1=[1,0,-1]^{\mathrm{T}}$$

$$\boldsymbol{\alpha}_3=\xi_2-\frac{(\xi_2,\boldsymbol{\alpha}_2)}{(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2)}\boldsymbol{\alpha}_2=[0,1,-1]^{\mathrm{T}}-\frac{1}{2}[1,0,-1]^{\mathrm{T}}=\frac{1}{2}[-1,2,-1]^{\mathrm{T}}$$

就满足要求。

我们也可应用例 4.12 给出的方法来求解该问题，只不过更为麻烦。首先利用例 4.12 的方法求与$\boldsymbol{\alpha}_1$正交的$\boldsymbol{\alpha}_2$，然后再用同样的方法求与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$正交的$\boldsymbol{\alpha}_3$，这一过程需要求解两个线性方程组。

习题 4.4

1. 设向量$\boldsymbol{\alpha}=[1,2,2]^{\mathrm{T}},\boldsymbol{\beta}=[1,-2,2]^{\mathrm{T}}$，求一个单位向$\boldsymbol{\gamma}$量，使其与已知的两个向量都正交。
2. 设$\boldsymbol{\alpha}=[1,2,2]^{\mathrm{T}},\boldsymbol{\beta}=[1,-2,2]^{\mathrm{T}}$，求一个单位向量$\boldsymbol{\gamma}$，使其与已知的两个向量都夹$60^\circ$角。
3. 设n维向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$线性无关，$\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$线性无关，若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$都分别与$\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$正交。证明$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$线性无关。
4. 利用 Schmidt 正交化方法将向量组$\boldsymbol{\alpha}_1=[0,1,1]^{\mathrm{T}},\boldsymbol{\alpha}_2=[1,0,1]^{\mathrm{T}},\boldsymbol{\alpha}_3=[1,1,0]^{\mathrm{T}}$化为标准正交向量组。

5.　已知$\boldsymbol{\alpha}_1=[0,1,1]^{\mathrm{T}}$，求向量$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$，使得$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$为正交向量组。

§4.5　子空间

定义 4.20　**子空间**

设V为数域$\mathbb{F}$上的向量空间，$W\subset V$且$W\neq\phi$，若W的所有元素关于V中的加法与数乘运算也构成一个向量空间，则称W为V的子空间。

例 4.17　任何向量空间$\boldsymbol{V}$都有两个子空间：$\boldsymbol{V}$与$\{\boldsymbol{0}\}$，称这两个子空间为$\boldsymbol{V}$的**平凡子空间**。

例 4.18

$$\left\{\begin{bmatrix}\boldsymbol{x}_1\\\boldsymbol{x}_2\\\vdots\\\boldsymbol{x}_n\end{bmatrix}\middle|\ \boldsymbol{x}_1+\boldsymbol{x}_2+\cdots+\boldsymbol{x}_n=0\right\}$$

为$\mathbb{R}^{\boldsymbol{n}}$的子空间。

定理 4.22　W为向量空间V的非空子集，在V的加法、数乘运算下，如下命题是等价的：

（1）W为V的子空间；

（2）W对加法、数乘运算封闭：

a)　若$\boldsymbol{\alpha},\boldsymbol{\beta}\in W$，$\boldsymbol{\alpha}+\boldsymbol{\beta}\in W$；

b)　若$\boldsymbol{\alpha}\in W,k\in\mathbb{R}$，则$k\boldsymbol{\alpha}\in W$；

（3）W对其中任意两个向量的线性运算是封闭的：$\forall\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in W,k_1,k_2\in\mathbb{F}$，有$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\in W$；

（4）W对其中任意向量组的线性运算是封闭的：$\forall\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in W,k_1,k_2,\cdots,k_r\in\mathbb{F}$，有$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r\in W$；

证明：根据向量空间及子空间的定义，$(1)\Leftrightarrow(2)$是显然的。

$(2)\Rightarrow(3)$：由于$\forall\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in W,k_1,k_2\in\mathbb{F}$，由于$W$对数乘封闭，所以$k_1\boldsymbol{\alpha}_1,k_2\boldsymbol{\alpha}_2\in W$，再利

用W对向量加法封闭，所以$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\in W$；

(3)⇒(4)：利用归纳法。显然，当$r=1$时命题成立，假设当$r=m-1$时命题成立，下证$r=m$时命题也成立。由归纳假设，$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{m-1}\boldsymbol{\alpha}_{m-1}\in W$，由于$\alpha_m\in W$，因此

$$1\times\left(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{m-1}\boldsymbol{\alpha}_{m-1}\right)+k_m\boldsymbol{\alpha}_m\in W$$

所以命题在$r=m$时也成立。

(4)⇒(2)：这是显然的。

向量空间中任意集合所生成的向量空间是该空间的子空间。

定理 4.23　**交空间、和空间**

设W_1,W_2是向量空间V的子空间，则有

（1）$W_1\bigcap W_2=\left\{\boldsymbol{\alpha}\middle|\boldsymbol{\alpha}\in W_1\text{且}\boldsymbol{\alpha}\in W_2\right\}$是$V$的子空间，称为$W_1$与$W_2$的**交空间**；

（2）$W_1+W_2=\left\{\boldsymbol{\alpha}\middle|\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\ \boldsymbol{\alpha}_1\in W_1,\ \boldsymbol{\alpha}_2\in W_2\right\}$是$V$的子空间，称为$W_1$与$W_2$的**和空间**；

证明：利用以上定理证明即可。

（1）$\forall\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in W_1\bigcap W_2,k_1,k_2\in\mathbb{F}$，由于$W_1,W_2$是向量空间$V$的子空间，故$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\in W_1$，$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\in W_2$，因此$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\in W_1\bigcap W_2$，这说明$W_1\bigcap W_2$对向量的线性组合封闭，根据定理 4.17（3）可知$W_1\bigcap W_2$为子空间。

（2）设$\boldsymbol{\alpha},\boldsymbol{\beta}\in W_1+W_2$，根据定义，分别存在$\boldsymbol{\alpha}_1\in W_1,\boldsymbol{\alpha}_2\in W_2,\boldsymbol{\beta}_1\in W_1,\boldsymbol{\beta}_2\in W_2$，使得

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\quad \boldsymbol{\beta}=\boldsymbol{\beta}_1+\boldsymbol{\beta}_2$$

因此$\forall k_1,k_2\in\mathbb{F}$，

$$\begin{aligned}k_1\boldsymbol{\alpha}+k_2\boldsymbol{\beta}&=k_1\left(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2\right)+k_2\left(\boldsymbol{\beta}_1+\boldsymbol{\beta}_2\right)\\&=\left(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\beta}_1\right)+\left(k_1\boldsymbol{\alpha}_2+k_2\boldsymbol{\beta}_2\right)\end{aligned}$$

显然有$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\beta}_1\in W_1$，$k_1\boldsymbol{\alpha}_2+k_2\boldsymbol{\beta}_2\in W_2$，根据定义因此有$k_1\boldsymbol{\alpha}+k_2\boldsymbol{\beta}\in W_1+W_2$，因此$W_1+W_2$是$V$的子空间。

例 4.19　设

$$\boldsymbol{\alpha}_1=\left[1,2,1,0\right]^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=\left[-1,1,1,1\right]^{\mathrm{T}}$$
$$\boldsymbol{\beta}_1=\left[2,-1,0,1\right]^{\mathrm{T}},\boldsymbol{\beta}_2=\left[1,-1,3,7\right]^{\mathrm{T}}$$
$$W_1=span\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\right\},\qquad W_2=span\left\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\right\}$$

求$W_1\bigcap W_2,W_1+W_2$。

解：设$\boldsymbol{\gamma}\in W_1\bigcap W_2$，因此存在$k_1,k_2,-k_3,-k_4$，使得

$$\boldsymbol{\gamma}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2=-k_3\boldsymbol{\beta}_1-k_4\boldsymbol{\beta}_2$$

所以

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\beta}_1+k_4\boldsymbol{\beta}_2=0$$

所以

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

所以

$$\begin{cases}k_1=k_4\\k_2=-4k_4\\k_3=-3k_4\\k_4=k_4\end{cases}$$

所以

$$\begin{aligned}\boldsymbol{\gamma}&=k_4\boldsymbol{\alpha}_1-4k_4\boldsymbol{\alpha}_2=k_4\left(\boldsymbol{\alpha}_1-4\boldsymbol{\alpha}_2\right)\\&=k_4\left[5,-2,-3,-4\right]^{\mathrm{T}}\end{aligned}$$

即

$$W_1\bigcap W_2=L\left\{\left[5,-2,-3,-4\right]^{\mathrm{T}}\right\}$$

显然，$W_1+W_2=span\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\right\}$，所以$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$的一个极大无关组就是$W_1+W_2$的一组基，根据前面的计算结果知

$$W_1+W_2=span\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1\right\}$$

两个子空间的和空间、交空间可以推广到多个子空间。设$W_1,W_2,\cdots,W_m$是V的子空间，则

$$\bigcap_{i=1}^{m}W_i=\left\{\boldsymbol{\alpha}\middle|\boldsymbol{\alpha}\in W_i,i=1,2,\cdots,m\right\}$$

也是V的子空间，称为$W_1,W_2,\cdots,W_m$的交空间；

$$\mathop{+}_{i=1}^{m}W=\sum_{i=1}^{m}W_i=\left\{\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_m\,\middle|\,\boldsymbol{\alpha}_i\in W_i,i=1,2,\cdots,m\right\}$$

也是V的子空间，称为$W_1,W_2,\cdots,W_m$和空间。

定理 4.24（维数公式）　设W_1,W_2是向量空间V的子空间，则有

$$\dim W_1+\dim W_2=\dim\left(W_1+W_2\right)+\dim\left(W_1\bigcap W_2\right)$$

证明　设$dim\left(W_1\bigcap W_2\right)=r$，$dim\left(W_1\right)=m$，$dim\left(W_2\right)=n$，显然$r\le\min\{m,n\}$。设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$为$W_1\bigcap W_2$的一组基，因此$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性无关，由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$是$W_1$中的线性无关组，因此可将之扩充为$W_1$的一组基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_m$；同理$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$可扩充为$W_2$的一组基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_{r+1},\cdots,\boldsymbol{\beta}_n$。下面我们证明：向量组

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_{r+1},\cdots,\boldsymbol{\beta}_n$$

是W_1+W_2的一组基。首先，显然有

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_{r+1},\cdots,\boldsymbol{\beta}_n\in W_1+W_2$$

（即该向量组中的每个向量都在W_1+W_2中）

然后证明该向量组能够线性表示W_1+W_2中的每个向量。$\forall\boldsymbol{\gamma}\in W_1+W_2$，存在$\boldsymbol{\gamma}_1\in W_1,\boldsymbol{\gamma}_2\in W_2$，使得$\boldsymbol{\gamma}=\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2$。由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_m$为$W_1$的基，故存在数$k_1,k_2,\cdots,k_r,k_{r+1},\cdots,k_m$，使得

$$\boldsymbol{\gamma}_1=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r+k_{r+1}\boldsymbol{\alpha}_{r+1},+\cdots+k_m\boldsymbol{\alpha}_m$$

同理，存在数$l_1,l_2,\cdots,l_r,l_{r+1},\cdots,l_n$，使得

$$\boldsymbol{\gamma}_2=l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_r\boldsymbol{\alpha}_r+l_{r+1}\boldsymbol{\beta}_{r+1},+\cdots+l_n\boldsymbol{\beta}_n$$

因此

$$
\begin{aligned}
\gamma &= \gamma_1 + \gamma_2 \\
&= \left(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_r\boldsymbol{\alpha}_r + k_{r+1}\boldsymbol{\alpha}_{r+1}, + \cdots + k_m\boldsymbol{\alpha}_m\right) \\
&\quad + \left(l_1\boldsymbol{\alpha}_1 + l_2\boldsymbol{\alpha}_2 + \cdots + l_r\boldsymbol{\alpha}_r + l_{r+1}\boldsymbol{\beta}_{r+1}, + \cdots + l_n\boldsymbol{\beta}_n\right) \\
&= \left(k_1 + l_1\right)\boldsymbol{\alpha}_1 + \left(k_2 + l_2\right)\boldsymbol{\alpha}_2 + \cdots + \left(k_r + l_r\right)\boldsymbol{\alpha}_r + k_{r+1}\boldsymbol{\alpha}_{r+1}, + \cdots + k_m\boldsymbol{\alpha}_m + l_{r+1}\boldsymbol{\beta}_{r+1}, + \cdots + l_n\boldsymbol{\beta}_n
\end{aligned}
$$

因此$W_1 + W_2$中的任意向量都可由$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}_{r+1}, \cdots, \boldsymbol{\beta}_n$线性表示。下证明该向量组线性无关。假设$k_1, k_2, \cdots, k_r, k_{r+1}, \cdots, k_m, l_{r+1}, \cdots, l_n$，使得

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_r\boldsymbol{\alpha}_r + k_{r+1}\boldsymbol{\alpha}_{r+1} + \cdots + k_m\boldsymbol{\alpha}_m + l_{r+1}\boldsymbol{\beta}_{r+1} + \cdots + l_n\boldsymbol{\beta}_n = \mathbf{0}$$

因此有

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_r\boldsymbol{\alpha}_r + k_{r+1}\boldsymbol{\alpha}_{r+1}, + \cdots + k_m\boldsymbol{\alpha}_m = -l_{r+1}\boldsymbol{\beta}_{r+1} - \cdots - l_n\boldsymbol{\beta}_n$$

显然有

$$
\begin{aligned}
k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_r\boldsymbol{\alpha}_r + k_{r+1}\boldsymbol{\alpha}_{r+1}, + \cdots + k_m\boldsymbol{\alpha}_m &\in W_1 \\
-l_{r+1}\boldsymbol{\beta}_{r+1} - \cdots - l_n\boldsymbol{\beta}_n &\in W_2
\end{aligned}
$$

因此

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_r\boldsymbol{\alpha}_r + k_{r+1}\boldsymbol{\alpha}_{r+1}, + \cdots + k_m\boldsymbol{\alpha}_m = -l_{r+1}\boldsymbol{\beta}_{r+1} - \cdots - l_n\boldsymbol{\beta}_n \in W_1 \cap W_2$$

由于$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$为$W_1 \cap W_2$的基，故存在$c_1, c_2, \cdots, c_r$，使得

$$-l_{r+1}\boldsymbol{\beta}_{r+1} - \cdots - l_n\boldsymbol{\beta}_n = c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_r\boldsymbol{\alpha}_r$$

因此$c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_r\boldsymbol{\alpha}_r + l_{r+1}\boldsymbol{\beta}_{r+1} + \cdots + l_n\boldsymbol{\beta}_n = \mathbf{0}$，由于$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_{r+1}, \cdots, \boldsymbol{\beta}_n$是$W_2$的基，因此该向量组线性无关，因此得到$c_1 = c_2 = \cdots = c_r = l_{r+1} = \cdots = l_n = 0$，因此

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_r\boldsymbol{\alpha}_r + k_{r+1}\boldsymbol{\alpha}_{r+1}, + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}$$

由于$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \cdots, \boldsymbol{\alpha}_m$是$W_1$的基，因此该向量组线性无关，因此得到

$$k_1 = k_2 = \cdots = k_r = k_{r+1}, = \cdots = k_m = 0$$

因此$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}_{r+1}, \cdots, \boldsymbol{\beta}_n$是线性无关的。综合上述结果可知，向量组

$$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}_{r+1}, \cdots, \boldsymbol{\beta}_n$$

是$W_1 + W_2$的一组基，即得$\dim W_1 + W_2 = m + n - r$，故结论是成立的。

■

定义 4.21　**直和**

设W_1,W_2是向量空间V的子空间，$W=W_1+W_2$为其和空间，若$\forall\boldsymbol{\alpha}\in W$，存在唯一的$\boldsymbol{\alpha}_1\in W_1$，$\boldsymbol{\alpha}_2\in W_2$，使得$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$，则称$W_1+W_2$为$W_1$与$W_2$的直和，记作

$$W=W_1\oplus W_2$$

且称$\boldsymbol{\alpha}_1$为$\boldsymbol{\alpha}$在W_1上的投影。

定理 4.25　W_1,W_2是向量空间V的子空间，则

$$W=W_1\oplus W_2\Leftrightarrow W_1\cap W_2=\{\mathbf{0}\}$$

证明：必要性：若$W=W_1\oplus W_2$，要证明$W_1\cap W_2=\{\mathbf{0}\}$。这可用反证法，设有$\mathbf{0}\neq\boldsymbol{\gamma}\in W_1\cap W_2$，那么对于任意的$\boldsymbol{\alpha}_1\in W_1$，$\boldsymbol{\alpha}_2\in W_2$，则有

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2=(\boldsymbol{\alpha}_1-\boldsymbol{\gamma})+(\boldsymbol{\alpha}_2+\boldsymbol{\gamma})$$

这说明$\boldsymbol{\alpha}$的表示方式不唯一，这与$W=W_1\oplus W_2$矛盾。

充分性：若$W_1\cap W_2=\{\mathbf{0}\}$，我们要证明$W=W_1+W_2$是直和。若有$\boldsymbol{\alpha}\in W$有两种表示方法，即存在$\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1\in W_1$，$\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2\in W_2$，使得

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2=\boldsymbol{\beta}_1+\boldsymbol{\beta}_2$$

因此有

$$\boldsymbol{\alpha}_1-\boldsymbol{\beta}_1=\boldsymbol{\alpha}_2-\boldsymbol{\beta}_2$$

由于$\boldsymbol{\alpha}_1-\boldsymbol{\beta}_1\in W_1$，$\boldsymbol{\alpha}_2-\boldsymbol{\beta}_2\in W_2$，所以$\boldsymbol{\alpha}_1-\boldsymbol{\beta}_1=\boldsymbol{\alpha}_2-\boldsymbol{\beta}_2\in W_1\cap W_2$，所以$\boldsymbol{\alpha}_1-\boldsymbol{\beta}_1=\boldsymbol{\alpha}_2-\boldsymbol{\beta}_2=\mathbf{0}$，所以$\boldsymbol{\alpha}_1=\boldsymbol{\beta}_1$，$\boldsymbol{\alpha}_2=\boldsymbol{\beta}_2$，因此$\boldsymbol{\alpha}$的表示方式唯一。因此$W=W_1+W_2$是直和。

■

结合上述两个定理，我们有

定理 4.26　W_1,W_2是向量空间$\boldsymbol{V}$的子空间，W为W_1与W_2的和空间，则

$$W=W_1\oplus W_2\Leftrightarrow\dim(W)=\dim(W_1)+\dim(W_2)$$

证明：由于$\dim W_1+\dim W_2=\dim(W_1+W_2)+\dim(W_1\cap W_2)$总是成立的，故

$$W = W_1 \oplus W_2 \Leftrightarrow W_1 \cap W_2 = \{\mathbf{0}\} \Leftrightarrow \dim W_1 \cap W_2 = 0$$
$$\Leftrightarrow \dim(W_1 + W_2) = \dim(W_1) + \dim(W_2)$$

■

同样地，直和也可以推广到多个子空间的情形。设设$W_1, W_2, \cdots, W_m$是V的子空间，

$$\sum_{i=1}^{m} W_i = \left\{\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m \middle| \alpha_i \in W_i, i = 1, 2, \cdots, m\right\}$$

是$W_1, W_2, \cdots, W_m$和空间，若$\forall \alpha \in \sum_{i=1}^{m} W_i$，存在唯一的$\alpha_i \in W_i, i = 1, 2, \cdots, m$，使得

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m$$

则称$\sum_{i=1}^{m} W_i$为$W_1, W_2, \cdots, W_m$的直和，记作$\bigoplus_{i=1}^{m} W_i$。

类似地，关于直和有如下结论：

定理 4.26　以下结论是等价的

（1）$W = \bigoplus_{i=1}^{m} W_i$是$W_1, W_2, \cdots, W_m$的直和；

（2）$W_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{m} W_j = \{0\}, i = 1, 2, \cdots, m$；

（3）$\dim\left(\bigoplus_{i=1}^{m} W_i\right) = \sum_{i=1}^{m} \dim W_i$.

定义 4.22　向量与空间正交，空间正交

如果向量$\boldsymbol{\alpha}$与欧式空间V中的每个向量正交，即$\forall \boldsymbol{\beta} \in V$，有$\boldsymbol{\alpha} \perp \boldsymbol{\beta}$，则称$\boldsymbol{\alpha}$与$V$正交，记做$\boldsymbol{\alpha} \perp V$。

设V_1, V_2都是欧式空间$\mathbb{R}^n$的子空间，若$\forall \boldsymbol{\alpha} \in V_1, \boldsymbol{\beta} \in V_2$，有$\boldsymbol{\alpha} \perp \boldsymbol{\beta}$，则称$V_1, V_2$正交，记做$V_1 \perp V_2$。

显然若$V_1 \perp V_2$，则$V_1 + V_2$为直和。

定理 4.27　设V是欧式空间$\mathbb{R}^n$的子空间，则

$$V^{\perp} = \left\{\alpha \in \mathbb{R}^n \mid \alpha \perp V\right\}$$

也是$\mathbb{R}^n$的子空间，称为V的**正交补空间**，简称为**正交补**。

证明　显然 $\mathbf{0}\in V^{\perp}$，故 $V^{\perp}\neq\phi$。下证 $V^{\perp}$ 对向量线性运算封闭。设 $\boldsymbol{\alpha},\boldsymbol{\beta}\in\boldsymbol{V}^{\perp}$，故 $\forall\boldsymbol{\gamma}\in\boldsymbol{V}$，有 $\boldsymbol{\alpha}\perp\gamma,\boldsymbol{\beta}\perp\gamma$，即 $(\boldsymbol{\alpha},\gamma)=(\boldsymbol{\beta},\gamma)=0$，所以 $\forall k,l\in\mathbb{R}$，有

$$(k\boldsymbol{\alpha}+l\boldsymbol{\beta},\gamma)=k(\boldsymbol{\alpha},\gamma)+l(\boldsymbol{\beta},\gamma)=k0+l0=0$$

所以 $k\boldsymbol{\alpha}+l\boldsymbol{\beta}\in V^{\perp}$，因此 $V^{\perp}$ 为一个向量空间，从而是 $\mathbb{R}^n$ 的子空间。 ■

定理 4.28　若 V 是欧式空间 $\mathbb{R}^n$ 的子空间，则

$$\mathbb{R}^n=V\oplus V^{\perp}$$

证明：设 $\dim V=r<n$，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 为 V 的一组正交基，故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in\mathbb{R}^n$，根据前面所得到的结论，我们可找到 $\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n$，使得 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n$ 为 $\mathbb{R}^n$ 的正交基，所以

$$\begin{aligned}\mathbb{R}^n&=L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n\}\\&=L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}+L\{\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n\}\\&=V+L\{\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n\}\end{aligned}$$

我们证明 $V^{\perp}=L\{\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n\}$。首先由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n$ 为 $\mathbb{R}^n$ 的正交基，所以 $L\{\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n\}\subseteq V^{\perp}$；其次，若 $\boldsymbol{\beta}\in V^{\perp}\subseteq\mathbb{R}^n$，则 $\boldsymbol{\beta}$ 可由 $\mathbb{R}^n$ 的基线性表示，故存在 $k_1,k_2,\cdots,k_n\in\mathbb{R}$，使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r+k_{r+1}\boldsymbol{\alpha}_{r+1}+k_{r+2}\boldsymbol{\alpha}_{r+2}+\cdots+k_n\boldsymbol{\alpha}_n$$

由于 $\boldsymbol{\beta}\perp V$，故 $(\boldsymbol{\alpha}_i,\boldsymbol{\beta})=0,\ i=1,2,\cdots,r$，所以 $k_i=0,\ i=1,2,\cdots,r$，即

$$\boldsymbol{\beta}=k_{r+1}\boldsymbol{\alpha}_{r+1}+k_{r+2}\boldsymbol{\alpha}_{r+2}+\cdots+k_n\boldsymbol{\alpha}_n$$

故 $\boldsymbol{\beta}\in L\{\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n\}$，因此 $V^{\perp}\subseteq L\{\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n\}$。就有 $V^{\perp}=L\{\boldsymbol{\alpha}_{r+1},\boldsymbol{\alpha}_{r+2},\cdots,\boldsymbol{\alpha}_n\}$。故 $\mathbb{R}^n=V\oplus V^{\perp}$。 ■

由于 $\mathbb{R}^n=V\oplus V^{\perp}$，故 $\forall\boldsymbol{\alpha}\in\mathbb{R}^n$，存在唯一的 $\boldsymbol{\alpha}_1\in V$，$\boldsymbol{\alpha}_2\in V^{\perp}$，使得

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$$

我们称之为 $\boldsymbol{\alpha}$ 的**正交分解**，$\boldsymbol{\alpha}_1$ 为 $\boldsymbol{\alpha}$ 在子空间 V 上的**正交投影**。

显然有如下结论

定理 4.29　$\boldsymbol{\alpha}_1$ 为 $\boldsymbol{\alpha}$ 在子空间 V 上的正交投影 $\Leftrightarrow \boldsymbol{\alpha}-\boldsymbol{\alpha}_1 \in V^{\perp}$

正交投影有一个重要的性质：

定理 4.30　设 V 是欧氏空间 $\mathbb{R}^n$ 的子空间，$\boldsymbol{\alpha}\in\mathbb{R}^n$，$\boldsymbol{\alpha}_1$ 为 $\boldsymbol{\alpha}$ 在子空间 V 上的正交投影 $\Leftrightarrow \|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1\| = \min\limits_{\boldsymbol{\beta}\in V}\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|$。（因此：$\boldsymbol{\alpha}$ 在子空间 V 上的正交投影 $\boldsymbol{\alpha}_1$ 是 V 中距离 $\boldsymbol{\alpha}$ “最近”的向量。）

证明　设 $\dim V = r$，$\gamma_1,\gamma_2,\cdots,\gamma_r$ 是 V 的一组标准正交基，将 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_r$ 扩充为 $\mathbb{R}^n$ 的一组标准正交基 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_r,\boldsymbol{\gamma}_{r+1},\cdots,\boldsymbol{\gamma}_n$，因此存在唯一的一组数 $k_1,k_2,\cdots,k_r,k_{r+1},\cdots,k_n$，使得

$$\boldsymbol{\alpha} = k_1\boldsymbol{\gamma}_1 + k_2\boldsymbol{\gamma}_2 + \cdots + k_r\boldsymbol{\gamma}_r + k_{r+1}\boldsymbol{\gamma}_{r+1} + \cdots + k_n\boldsymbol{\gamma}_n$$

显然 $k_1\boldsymbol{\gamma}_1 + k_2\boldsymbol{\gamma}_2 + \cdots + k_r\boldsymbol{\gamma}_r \in V$，$k_{r+1}\boldsymbol{\gamma}_{r+1} + \cdots + k_n\boldsymbol{\gamma}_n \in V^{\perp}$，故 $\boldsymbol{\alpha}$ 在子空间 V 上的正交投影 $\boldsymbol{\alpha}_1$ 为

$$\boldsymbol{\alpha}_1 = k_1\boldsymbol{\gamma}_1 + k_2\boldsymbol{\gamma}_2 + \cdots + k_r\boldsymbol{\gamma}_r$$

对于 V 中的任何其它向量 $\boldsymbol{\beta}$，存在 $l_1,l_2,\cdots,l_r$，使得

$$\boldsymbol{\beta} = l_1\boldsymbol{\alpha}_1 + l_2\boldsymbol{\alpha}_2 + \cdots + l_r\boldsymbol{\alpha}_r$$

因此

$$\begin{aligned}
\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|^2 &= \left\|\left(k_1\boldsymbol{\gamma}_1 + k_2\boldsymbol{\gamma}_2 + \cdots + k_r\boldsymbol{\gamma}_r + k_{r+1}\boldsymbol{\gamma}_{r+1} + \cdots + k_n\boldsymbol{\gamma}_n\right) - \left(l_1\boldsymbol{\alpha}_1 + l_2\boldsymbol{\alpha}_2 + \cdots + l_r\boldsymbol{\alpha}_r\right)\right\|^2 \\
&= \left\|\left(k_1-l_1\right)\boldsymbol{\gamma}_1 + \left(k_2-l_2\right)\boldsymbol{\gamma}_2 + \cdots + \left(k_r-l_r\right)\boldsymbol{\gamma}_r + k_{r+1}\boldsymbol{\gamma}_{r+1} + \cdots + k_n\boldsymbol{\gamma}_n\right\|^2 \\
&= \left(k_1-l_1\right)^2 + \left(k_2-l_2\right)^2 + \cdots + \left(k_r-l_r\right)^2 + k_{r+1}^2 + \cdots + k_n^2 \\
&\geq k_{r+1}^2 + \cdots + k_n^2 = \|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1\|^2
\end{aligned}$$

由于上述不等式中的等号当且仅当 $k_1-l_1 = k_2-l_2 = \cdots = k_r-l_r = 0$ 时成立，即 $\boldsymbol{\beta}=\boldsymbol{\alpha}_1$ 时成立，即 $\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|$ 在 $\boldsymbol{\beta}=\boldsymbol{\alpha}_1$ 时取最小值，因此必要性成立。

充分性：若 $\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1\| = \min\limits_{\boldsymbol{\beta}\in V}\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|$，由于

$$\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|^2=(k_1-l_1)^2+(k_2-l_2)^2+\cdots+(k_r-l_r)^2+k_{r+1}^2+\cdots+k_n^2$$
$$\geq k_{r+1}^2+\cdots+k_n^2$$

其最小值在$k_1-l_1=k_2-l_2=\cdots=k_r-l_r=0$时成立，即$\boldsymbol{\alpha}_1=\boldsymbol{\beta}=k_1\boldsymbol{\gamma}_1+k_2\boldsymbol{\gamma}_2+\cdots+k_r\boldsymbol{\gamma}_r$，而该向量显然是$\boldsymbol{\alpha}$在$V$上的正交投影。

由该定理的证明过程可知：对任给的$\boldsymbol{\alpha}$与子空间$V\neq\{\mathbf{0}\}$，$\boldsymbol{\alpha}$在子空间V上的正交投影总是存在的。

定理的意义：设V是欧式空间$\mathbb{R}^n$的子空间，$\boldsymbol{\alpha}\in\mathbb{R}^n$但$\boldsymbol{\alpha}\notin V$，我们考虑用$V$中的向量$\boldsymbol{\beta}$来逼近$\boldsymbol{\alpha}$。既然是用$\boldsymbol{\beta}$来逼近$\boldsymbol{\alpha}$，自然有逼近误差存在，如果定义用$\boldsymbol{\beta}$来逼近$\boldsymbol{\alpha}$时的逼近误差为

$$\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|$$

而V中使得上述误差最小的向量$\boldsymbol{\gamma}$就称为$\boldsymbol{\alpha}$在V中的**最佳逼近元**：

$$\|\boldsymbol{\alpha}-\boldsymbol{\gamma}\|=\min_{\boldsymbol{\beta}\in V}\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|$$

显然，$\boldsymbol{\alpha}$在V中的**最佳逼近元**就是$\boldsymbol{\alpha}_1$，它为$\boldsymbol{\alpha}$在V上的正交投影。

这个问题是有意义的。考虑线性方程组$\boldsymbol{AX}=\boldsymbol{b}$，如果该方程组有解，说明向量$\boldsymbol{b}$可由矩阵$\boldsymbol{A}$的列向量组线性表示，也就是$\boldsymbol{b}\in R(\boldsymbol{A})$（$\boldsymbol{A}$的列空间）；若$\boldsymbol{AX}=\boldsymbol{b}$无解，则说明向量$\boldsymbol{b}$不能由矩阵$A$的列向量组线性表示，也就是$\boldsymbol{b}\notin R(\boldsymbol{A})$（$\boldsymbol{A}$的列空间）。如果$\boldsymbol{b}$不能由矩阵$\boldsymbol{A}$的列向量组精确地线性表示，我们考虑用$\boldsymbol{A}$的列向量组的线性组合（也就是$R(\boldsymbol{A})$中的向量）来逼近$\boldsymbol{b}$。如上给出的误差下，$\boldsymbol{b}$在$R(\boldsymbol{A})$中的最佳逼近元就是$\boldsymbol{b}$在$R(\boldsymbol{A})$上的正交投影。

正交投影的确定

对于线性方程组$\boldsymbol{AX}=\boldsymbol{b}$，考虑$\boldsymbol{b}$在$R(\boldsymbol{A})$上的正交投影。设$\boldsymbol{b}$在$R(\boldsymbol{A})$上的正交投影为$\boldsymbol{A}\hat{\boldsymbol{X}}$，则由定理可知，其充要条件为

$$\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{X}}\in R(\boldsymbol{A})^{\perp}$$

由于$R(\boldsymbol{A})$是$\boldsymbol{A}$的列向量组所生成的向量空间，故$\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{X}}\in R(\boldsymbol{A})^{\perp}$的充要条件是：

$\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{X}}$ 与 $\boldsymbol{A}$ 的每个列向量正交。

设 $\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n\end{bmatrix}$，故有

$$\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{X}} \perp \boldsymbol{A}_i, i=1,2,\cdots,n$$

即

$$\left(\boldsymbol{A}_i,\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{X}}\right)=\boldsymbol{A}_i^{\mathrm{T}}\left(\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{X}}\right)=0,\qquad i=1,2,\cdots,n$$

由此有

$$\begin{bmatrix}\boldsymbol{A}_1^{\mathrm{T}}\\ \boldsymbol{A}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{A}_n^{\mathrm{T}}\end{bmatrix}\left(\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{X}}\right)=0$$

即 $\boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{X}}\right)=0$，化简得

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\hat{\boldsymbol{X}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$$

根据前文的论述可知，正交投影总是存在的，所以上述方程组是有解的。以后将用另外的方法证明该线性方程组是有解的。

如果 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 可逆，就有 $\hat{\boldsymbol{X}}=\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$，而 $\boldsymbol{b}$ 在 $R\left(\boldsymbol{A}\right)$ 上的正交投影为 $\boldsymbol{A}\hat{\boldsymbol{X}}=\boldsymbol{A}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$。若 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 不可逆，则可用消元法求解。

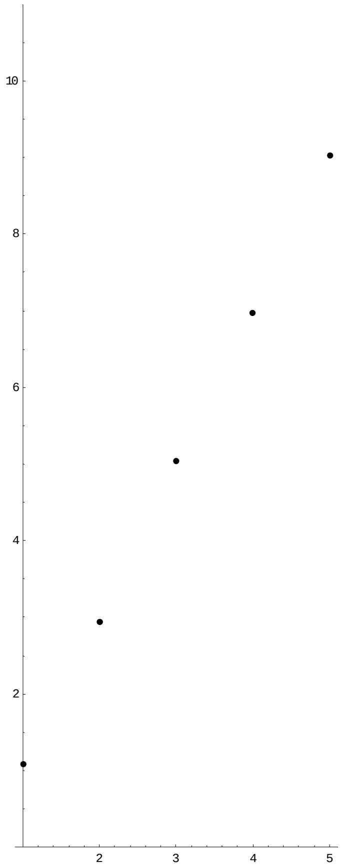
例 4.20　正交投影的应用的例子

在实际应用中，我们需要对数据进行测量，然后用相应的模型来拟合。考虑一个简单问题。有一个单输入、单输出的黑箱系统，其输入/输出满足如下关系

输入	1	2	3	4	5
输出	1.09	2.95	5.04	6.97	9.03

我们能否总结出其输入、输出关系？

首先，我们把数据绘出图形，试图对这些数据的规律有直观的把握。



容易看出，这些数据点大致在一条直线上，设直线方程为

$$cx+d=y$$

把这五个数据代入，得到

$$\begin{cases} c+d=1.09 \\ 2c+d=2.95 \\ 3c+d=5.04 \\ 4c+d=6.97 \\ 5c+d=9.03 \end{cases} \quad \text{即} \quad \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix}=\begin{bmatrix} 1.09 \\ 2.95 \\ 5.04 \\ 6.97 \\ 9.03 \end{bmatrix}$$

此线性方程组无解，我们应用上述正交投影的理论来解决该问题。此时

$$\boldsymbol{A}=\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix}, \qquad \boldsymbol{b}=\begin{bmatrix} 1.09 \\ 2.95 \\ 5.04 \\ 6.97 \\ 9.03 \end{bmatrix}$$

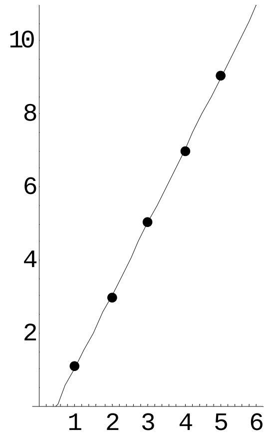
易知 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{bmatrix} 65 & 15 \\ 15 & 5 \end{bmatrix}$，故 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 可逆，

$$\hat{\boldsymbol{X}}=\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}=\begin{bmatrix} 1.99 \\ -0.954 \end{bmatrix}$$

所得到的直线为

$$y = 1.99x - 0.954$$

所得到的直线与原始数据的拟合程度，可用图形显示为



逼近误差为$\left[0.054, -0.076, 0.024, -0.036, 0.034\right]^{\mathrm{T}}$。

习题 4.5

1. 已知设

$$\boldsymbol{\alpha}_1 = \left[-6, 3, -3, -2\right]^{\mathrm{T}}, \quad \boldsymbol{\alpha}_2 = \left[7, 4, 5, -3\right]^{\mathrm{T}}$$

$$\boldsymbol{\beta}_1 = \left[4, 0, 2, -1\right]^{\mathrm{T}}, \boldsymbol{\beta}_2 = \left[-3, 1, 0, 3\right]^{\mathrm{T}}$$

$$W_1 = span\left\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\right\}, \qquad W_2 = span\left\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2\right\}$$

求$W_1 \bigcap W_2, W_1 + W_2$。且问：$W_1 + W_2$是否是直和？

2. 已知$\boldsymbol{\alpha}_1 = \left[1, 3, 2, -1\right]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2 = \left[2, -1, 0, 1\right]^{\mathrm{T}}$，$W_1 = span\left\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\right\}$，求$W_1$的正交补空间。

3. 设$\boldsymbol{\alpha}_1 = \left[-3, 3, 5, -4\right]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2 = \left[0, 3, -1, -1\right]^{\mathrm{T}}$，$\boldsymbol{\alpha}_3 = \left[3, 4, -3, 3\right]^{\mathrm{T}}$，$\boldsymbol{\beta} = \left[2, 4, -2, 1\right]^{\mathrm{T}}$，求$\boldsymbol{\beta}$在$span\left\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\right\}$上的正交投影。

第五章　线性方程组解的结构

§5.1　线性方程组解的表示

线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases} \tag{5.1}$$

可写成矩阵乘法的形式

$$\boldsymbol{AX}=\boldsymbol{b} \tag{5.2}$$

其中

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},\quad \boldsymbol{X}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

若记 $\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n\end{bmatrix}$，则线性方程组(5.1)可写作

$$x_1\boldsymbol{A}_1+x_2\boldsymbol{A}_2+\cdots+x_n\boldsymbol{A}_n=\boldsymbol{b} \tag{5.3}$$

这称为线性方程组的向量表示形式。
线性方程组的解

$$x_1=s_1, x_2=s_2, \cdots, x_n=s_n$$

可写成向量的形式

$$\boldsymbol{X}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

称之为线性方程组的**解向量**。$\boldsymbol{AX}=\boldsymbol{b}$ 的任何一个解称为它的一个**特解**。

上一章我们定义了向量的线性运算，运用向量的线性运算，我们可把线性方程组的解用向量的线性运算来表示。例如，对于线性方程组（1.15）的解

$$\begin{cases} x_1 = -16 + x_3 + x_4 + 5x_5 \\ x_2 = 23 - 2x_3 - 2x_4 - 6x_5 \\ x_3 = x_3 \\ x_4 = x_4 \\ x_5 = x_5 \end{cases}$$

我们将其表示为向量的形式

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -16 + x_3 + \ x_4 + 5x_5 \\ 23 - 2x_3 - 2x_4 - 6x_5 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

运用向量的线性运算，上式可写作

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -16 \\ 23 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} x_3 \\ 2x_3 \\ x_3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} x_4 \\ 2x_4 \\ 0 \\ x_4 \\ 0 \end{bmatrix} + \begin{bmatrix} 5x_5 \\ -6x_5 \\ 0 \\ 0 \\ x_5 \end{bmatrix}$$

$$= \begin{bmatrix} -16 \\ 23 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 5 \\ -6 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

由于自由未知量取任意的值时，我们都可以得到线性方程组的一个解，因此，若令 $x_3 = x_4 = x_5 = 0$，可知 $\begin{bmatrix} -16 \\ 23 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ 为线性方程组（1.15）的一个解，因此它为（1.15）的一个特解。

以后我们都把线性方程组的解写成如上形式。例如线性方程组（1.16）的解可写成

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -27 \\ 10 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} \frac{22}{3} \\ -\frac{7}{6} \\ -\frac{2}{3} \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} \frac{41}{3} \\ -\frac{10}{3} \\ -\frac{1}{3} \\ 0 \\ 1 \end{bmatrix}$$

其中$\begin{bmatrix} -27 \\ 10 \\ 0 \\ 0 \\ 0 \end{bmatrix}$为一个特解。

由于线性方程组(5.1)也可写作(5.3)的形式，这是向量组的线性组合的形式，这使得我们可用向量组的相关理论来讨论线性方程组的相关问题。

定理 5.1　对于非齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times 1} = \boldsymbol{b}_{m\times 1}$，

（1） $\boldsymbol{AX} = \boldsymbol{b}$ 有解 $\Leftrightarrow$ 增广矩阵的秩与系数矩阵的秩相等，即 $r\left(\begin{bmatrix}\boldsymbol{A} & \boldsymbol{b}\end{bmatrix}\right) = r\left(\boldsymbol{A}\right)$；

（2） $\boldsymbol{AX} = \boldsymbol{b}$ 有唯一解 $\Leftrightarrow r\left(\begin{bmatrix}\boldsymbol{A} & \boldsymbol{b}\end{bmatrix}\right) = r\left(\boldsymbol{A}\right) = n$。

证明：该定理已经在第三章中得到证明，这里用第四章的知识进行证明。

（1）　必要性：若方程组有解，将线性方程组

$$\boldsymbol{AX} = \boldsymbol{b}$$

写作

$$\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n\end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boldsymbol{b}$$

故 $\boldsymbol{b}$ 是矩阵 $\boldsymbol{A}$ 的列向量组 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n$ 的线性组合，因而 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n$ 与 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n, \boldsymbol{b}$ 等价，这两个向量组有相同的秩，

$$r\left(\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n\right) = r\left(\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n, \boldsymbol{b}\right)$$

由于矩阵的秩等于它的列向量组的秩，即有 $r\left(\begin{bmatrix}\boldsymbol{A} & \boldsymbol{b}\end{bmatrix}\right) = r\left(\boldsymbol{A}\right)$。

充分性：若

$$r\left(\begin{bmatrix}\boldsymbol{A} & \boldsymbol{b}\end{bmatrix}\right) = r\left(\boldsymbol{A}\right)$$

由于矩阵的秩等于它的列向量组的秩，故

$$r\left(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n\right)=r\left(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n,\boldsymbol{b}\right)$$

故 $\boldsymbol{b}$ 必能由 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$ 的极大无关组线性表示，故 $\boldsymbol{b}$ 能由 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$ 线性表示，故有 $x_1,x_2,\cdots,x_n$，使得

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

即线性方程组

$$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$$

有解；

（2）必要性：若 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$ 有唯一解，故有

$$r\left(\left[\boldsymbol{A}\quad\boldsymbol{b}\right]\right)=r\left(\boldsymbol{A}\right)$$

设其唯一解向量为 $\left[x_1,x_2,\cdots,x_n\right]^{\mathrm{T}}$，则

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

若 $r\left(\boldsymbol{A}\right)<n$，则其列向量组 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$ 线性相关，因此存在不全为零的数 $k_1,k_2,\cdots,k_n$，使得

$$k_1\boldsymbol{A}_1+k_2\boldsymbol{A}_2+\cdots+k_n\boldsymbol{A}_n=\boldsymbol{0}$$

因此有

$$\begin{aligned}\boldsymbol{b}&=x_1\boldsymbol{A}_1+x_2\boldsymbol{A}_2+\cdots+x_n\boldsymbol{A}_n\\&=\left(x_1\boldsymbol{A}_1+x_2\boldsymbol{A}_2+\cdots+x_n\boldsymbol{A}_n\right)\\&\quad+\left(k_1\boldsymbol{A}_1+k_2\boldsymbol{A}_2+\cdots+k_n\boldsymbol{A}_n\right)\\&=\left(x_1+k_1\right)\boldsymbol{A}_1+\left(x_2+k_2\right)\boldsymbol{A}_2+\cdots+\left(x_n+k_n\right)\boldsymbol{A}_n\end{aligned}$$

因此 $\left[x_1+k_1,x_2+k_2,\cdots,x_n+k_n\right]^{\mathrm{T}}$ 也是 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$ 的解，而显然

$$\left[x_1+k_1,x_2+k_2,\cdots,x_n+k_n\right]^{\mathrm{T}}\neq\left[x_1,x_2,\cdots,x_n\right]^{\mathrm{T}}$$

这与 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$ 有唯一解矛盾。

充分性：若 $r\left(\left[\boldsymbol{A}\quad\boldsymbol{b}\right]\right)=r\left(\boldsymbol{A}\right)=n$，即

$$r\left(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n\right)=r\left(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n,\boldsymbol{b}\right)=n$$

这说明 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$ 线性无关，而 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n,\boldsymbol{b}$ 线性相关，故 $\boldsymbol{b}$ 可由 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$ 线性表示，且表示方式唯一，即 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$ 有唯一解。

推论 5.1　齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times 1}=\boldsymbol{0}$，

（1） $\boldsymbol{AX}=\boldsymbol{0}$ 有无穷多解 $\Leftrightarrow r(\boldsymbol{A})<n$；

（2） $\boldsymbol{AX}=\boldsymbol{b}$ 有唯一解 $\Leftrightarrow r(\boldsymbol{A})=n$。

例 5.1　设有线性方程组

$$\begin{cases}(1+\lambda)x_1+x_2+x_3=0\\ x_1+(1+\lambda)x_2+x_3=3\\ x_1+x_2+(1+\lambda)x_3=\lambda\end{cases}$$

问 λ 取何值时，此方程组（1）有唯一解；（2）无解；（3）有无穷多解？并在有无穷多解时求其通解。

解一：根据定理 5.1，我们只要讨论此线性方程组系数矩阵与增广矩阵的秩的关系，而求矩阵秩的通常方法是将它化为行阶梯形矩阵。

对增广矩阵 $\boldsymbol{B}=[\boldsymbol{A}\quad \boldsymbol{b}]$ 做初等行变换，将它化为行阶梯形矩阵

$$\boldsymbol{B}=\begin{bmatrix}1+\lambda & 1 & 1 & 0\\ 1 & 1+\lambda & 1 & 3\\ 1 & 1 & 1+\lambda & \lambda\end{bmatrix}\overset{r_1\leftrightarrow r_3}{\sim}\begin{bmatrix}1 & 1 & 1+\lambda & \lambda\\ 1 & 1+\lambda & 1 & 3\\ 1+\lambda & 1 & 1 & 0\end{bmatrix}$$

$$\overset{\substack{r_2-r_1\\ r_3-(1+\lambda)r_1}}{\sim}\begin{bmatrix}1 & 1 & 1+\lambda & \lambda\\ 0 & \lambda & -\lambda & 3-\lambda\\ 0 & -\lambda & -\lambda(2+\lambda) & -\lambda(1+\lambda)\end{bmatrix}$$

$$\overset{r_3+r_2}{\sim}\begin{bmatrix}1 & 1 & 1+\lambda & \lambda\\ 0 & \lambda & -\lambda & 3-\lambda\\ 0 & 0 & -\lambda(3+\lambda) & (1-\lambda)(3+\lambda)\end{bmatrix}$$

因此，

（1）当 $\lambda\neq 0$ 且 $\lambda\neq -3$ 时，$r(\boldsymbol{A})=r(\boldsymbol{B})=3$，方程组有唯一解；

（2）当 $\lambda=0$ 时，增广矩阵化为 $\begin{bmatrix}1 & 1 & 1 & 0\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0\end{bmatrix}$，所以 $r(\boldsymbol{A})=1, r(\boldsymbol{B})=2$，方程组无解；

（3） $\lambda=-3$ 时，增广矩阵化为 $\begin{bmatrix}1 & 1 & -2 & -3\\ 0 & -3 & 3 & 6\\ 0 & 0 & 0 & 0\end{bmatrix}$， $r(\boldsymbol{A})=r(\boldsymbol{B})=2$，方程组有无穷多解；这时

$$\boldsymbol{B}\sim\begin{bmatrix}1&1&-2&-3\\0&-3&3&6\\0&0&0&0\end{bmatrix}\sim\begin{bmatrix}1&1&-2&-3\\0&-1&1&2\\0&0&0&0\end{bmatrix}\sim\begin{bmatrix}1&0&-1&-1\\0&-1&1&2\\0&0&0&0\end{bmatrix}$$
$$\sim\begin{bmatrix}1&0&-1&-1\\0&1&-1&-2\\0&0&0&0\end{bmatrix}$$

由此得到

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

亦即

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=x_3\begin{bmatrix}1\\1\\1\end{bmatrix}+\begin{bmatrix}-1\\-2\\0\end{bmatrix}$$

解二：由于此时线性方程组的系数矩阵为方阵，故方程组有唯一解的充要条件是系数行列式 $|\boldsymbol{A}|\neq 0$，而

$$|\boldsymbol{A}|=\begin{vmatrix}1+\lambda&1&1\\1&1+\lambda&1\\1&1&1+\lambda\end{vmatrix}=\begin{vmatrix}3+\lambda&3+\lambda&3+\lambda\\1&1+\lambda&1\\1&1&1+\lambda\end{vmatrix}$$
$$=(3+\lambda)\begin{vmatrix}1&1&1\\1&1+\lambda&1\\1&1&1+\lambda\end{vmatrix}=(3+\lambda)\begin{vmatrix}1&1&1\\0&\lambda&0\\0&0&\lambda\end{vmatrix}=(3+\lambda)\lambda^2$$

故当 $\lambda\neq 0$ 且 $\lambda\neq -3$ 时，线性方程组有唯一解。

对于 $\lambda=0$、$\lambda=-3$ 时的情形，我们可以计算方程组的系数矩阵与增广矩阵的秩，从而判断解的情况。具体计算过程在此不再赘述。

习题 5.1

1.　求解齐次线性方程组

（1）$\begin{cases}x_1+x_2+x_3+4x_4-3x_5=0\\2x_1+x_2+3x_3+5x_4-5x_5=0\\x_1-x_2+3x_3-2x_4-x_5=0\\3x_1+x_2+5x_3+6x_4-7x_5=0\end{cases}$　　（2）$\begin{cases}x_1-8x_2+10x_3+2x_4=0\\2x_1+4x_2+5x_3-x_4=0\\3x_1+8x_2+6x_3-2x_4=0\end{cases}$

2.　求解非齐次线性方程组

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

（3）$\begin{bmatrix} 1 & 1 & -1 & 2 \\ 2 & 1 & 0 & -3 \\ -2 & 0 & -2 & 10 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}$

3. 设 $\boldsymbol{A}$ 为 $m \times n$ 矩阵，$\boldsymbol{AX} = \boldsymbol{0}$ 是 $\boldsymbol{AX} = \boldsymbol{b}$ 对应的齐次线性方程组，则下列结论成立的是（　）

（A）若 $\boldsymbol{AX} = \boldsymbol{0}$ 仅有零解，则 $\boldsymbol{AX} = \boldsymbol{b}$ 有唯一解；
（B）若 $\boldsymbol{AX} = \boldsymbol{0}$ 有非零解，则 $\boldsymbol{AX} = \boldsymbol{b}$ 有无穷多解；
（C）若 $\boldsymbol{AX} = \boldsymbol{0}$ 有无穷多解，则 $\boldsymbol{AX} = \boldsymbol{b}$ 仅有零解；
（D）若 $\boldsymbol{AX} = \boldsymbol{b}$ 有无穷多解，则 $\boldsymbol{AX} = \boldsymbol{0}$ 有非零解。

4. 设有齐次线性方程组 $\boldsymbol{AX} = \boldsymbol{0}$ 和 $\boldsymbol{BX} = \boldsymbol{0}$,其中 $\boldsymbol{A}, \boldsymbol{B}$ 为 $m \times n$ 矩阵，现有 4 个命题：

(1)　若 $\boldsymbol{AX} = \boldsymbol{0}$ 的解都是 $\boldsymbol{BX} = \boldsymbol{0}$ 的解，则 $r(\boldsymbol{A}) \geq r(\boldsymbol{B})$；

(2)　若 $r(\boldsymbol{A}) \geq r(\boldsymbol{B})$，则 $\boldsymbol{AX} = \boldsymbol{0}$ 的解都是 $\boldsymbol{BX} = \boldsymbol{0}$ 的解;

(3)　若 $\boldsymbol{AX} = \boldsymbol{0}$ 和 $\boldsymbol{BX} = \boldsymbol{0}$ 同解，则 $r(\boldsymbol{A}) = r(\boldsymbol{B})$；

(4)　若 $r(\boldsymbol{A}) = r(\boldsymbol{B})$，则 $\boldsymbol{AX} = \boldsymbol{0}$ 和 $\boldsymbol{BX} = \boldsymbol{0}$ 同解.

以上命题正确的是(　)

(A)　(1)(2)　　(B)　(1)(3)　(C)　(2)(4)　　(D)　(3)(4).

5. 设 $\boldsymbol{A}, \boldsymbol{B}$ 为满足 $\boldsymbol{AB} = \boldsymbol{O}$ 的任意两个非零矩阵,则必有 ( )

(A)　$\boldsymbol{A}$ 的列向量组线性相关，$\boldsymbol{B}$ 的行向量组线性相关.
(B)　$\boldsymbol{A}$ 的列向量组线性相关，$\boldsymbol{B}$ 的列向量组线性相关.
(C)　$\boldsymbol{A}$ 的行向量组线性相关，$\boldsymbol{B}$ 的行向量组线性相关.
(D)　$\boldsymbol{A}$ 的行向量组线性相关，$\boldsymbol{B}$ 的列向量组线性相关.

6. 设矩阵 $\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$ 是满秩的，则直线 $\dfrac{x-a_3}{a_1-a_2} = \dfrac{y-b_3}{b_1-b_2} = \dfrac{z-c_3}{c_1-c_2}$ 与直线

$\dfrac{x-a_1}{a_2-a_3} = \dfrac{y-b_1}{b_2-b_3} = \dfrac{z-c_1}{c_2-c_3}$　　　　( )

(A)　　相交于一点　　(B) 重合　　(C)　平行但不重合　　(D)　异面

7. 参数a取何值时，线性方程组$\begin{cases} ax_1+(a+3)x_2+x_3=-2 \\ x_1+ax_2+x_3=a \\ x_1+x_2+ax_3=a^2 \end{cases}$有无穷多解？求出通解。

8. 参数a取何值时，线性方程组$\begin{cases} ax_1+(a+3)x_2+x_3=-2 \\ x_1+ax_2+x_3=a \\ x_1+x_2+ax_3=a^2 \end{cases}$有无穷多解？求出通解。

9. 已知某非齐次线性方程组的增广矩阵为$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 & 1 \\ 2 & 3 & a+2 & 4 & b+3 \\ 3 & 5 & 1 & a+8 & 5 \end{bmatrix}$，请判断其解的情况。

10. 相容的线性方程组$\boldsymbol{AX}=\boldsymbol{b}$在怎样的条件下，其解中第$k$个未知量$x_k$都是同一个值？你给的条件是否是充分必要的？

11. 已知平面上的三条不同的直线方程分别为$\begin{matrix} l1: & ax+2by+3c=0 \\ l2: & bx+2cy+3a=0 \\ l3: & cx+2ay+3b=0 \end{matrix}$，证明：如果$a+b+c=0$，则三条直线交于一点。

12. 已知三阶非零矩阵$\boldsymbol{B}$的每个列向量都是方程组

$$\begin{cases} x_1+x_2+(\lambda+1)x_3=0 \\ x_1-x_3=0 \\ 3x_1-2x_2-x_3=0 \\ 5x_1-x_2+(\lambda-1)x_3=0 \end{cases}$$

的解，（1）求λ的值；（2）证明$|\boldsymbol{B}|=0$。

13. 设$\boldsymbol{A}$是n阶方阵，$\boldsymbol{\beta}$是n维列向量，$\boldsymbol{B}=\begin{bmatrix} \boldsymbol{A} & \boldsymbol{\beta} \\ \boldsymbol{\beta}^{\mathrm{T}} & 0 \end{bmatrix}$，若$r(\boldsymbol{A})=r(\boldsymbol{B})$，则$\boldsymbol{AX}=\boldsymbol{\beta}$有解。

14. 设$\boldsymbol{A}$为$m\times n$矩阵，记$\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n]$，如果$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-1}$线性相关，而$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_n$线性无关，又$\boldsymbol{\beta}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_n$，证明

（1）线性方程组$\boldsymbol{AX}=\boldsymbol{\beta}$必有无穷多解；（2）记$\boldsymbol{X}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$为$\boldsymbol{AX}=\boldsymbol{\beta}$的任意解，则必有$x_n=1$。

§5.2　齐次线性方程组解的结构

线性方程组的解有三种情况：无解、唯一解、无穷多解。若线性方程组有为穷多解，则它必有自由未知量，消元过程不同，最后选取的自由未知量也就不一样，因此对于具有无穷多解的线性方程组，其解的表达方式是不唯一的。我们这里的问题是：线性方程组的这些解之间有何关系、解的这些不同的表达方式之间有何关系？

定理 5.2　设 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解，则 $\boldsymbol{X}_1+\boldsymbol{X}_2$，$k\boldsymbol{X}_1$ 都是 $\boldsymbol{AX}=\boldsymbol{0}$ 的解。

证明：由于 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解，故

$$\boldsymbol{AX}_1=\boldsymbol{0},\qquad \boldsymbol{AX}_2=\boldsymbol{0}$$

所以

$$\begin{aligned}&\boldsymbol{A}\left(\boldsymbol{X}_1+\boldsymbol{X}_2\right)=\boldsymbol{AX}_1+\boldsymbol{AX}_2=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}\\&\boldsymbol{A}\left(k\boldsymbol{X}_1\right)=k\left(\boldsymbol{AX}_1\right)=k\boldsymbol{0}=\boldsymbol{0}\end{aligned}$$

因此命题成立。■

由于 $\boldsymbol{AX}=\boldsymbol{0}$ 至少具有零解，再由定理 5.2 可知齐次线性方程组所有的解构成一个向量空间。

定义 5.1　齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解集

$$N\left(\boldsymbol{A}\right)=\left\{\boldsymbol{X}\,\middle|\,\boldsymbol{AX}=\boldsymbol{0}\right\}\tag{5.4}$$

为一个向量空间，称之为 $\boldsymbol{AX}=\boldsymbol{0}$ 的**解空间**。

定义 5.2　齐次线性方程组的**基础解系**、**通解**

齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 解空间 $N\left(\boldsymbol{A}\right)$ 的基称为该方程组的**基础解系**。设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系，则它的任一解 $\boldsymbol{X}$ 都可表示为

$$\boldsymbol{X}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s\tag{5.5}$$

称之为 $\boldsymbol{AX}=\boldsymbol{0}$ 的**通解**。

既然齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 解集 $N\left(\boldsymbol{A}\right)$ 为一个向量空间，我们自然关心该向量空间的维数、以及如何给出它的一组基，如下定理则给出了该问题的答案。

定理 5.3　线性方程组

$$\boldsymbol{A}_{m\times n}\boldsymbol{X}_{n\times 1}=\boldsymbol{0}_{m\times 1}$$

的系数矩阵的秩 $r(\boldsymbol{A})=r$，则其解空间 $N(\boldsymbol{A})$ 的维数

$$\dim N(\boldsymbol{A})=n-r\text{。}$$

证明：设 $r(\boldsymbol{A})=r$，则 $\boldsymbol{A}$ 的列向量组 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$ 的秩为 r，假设 $\boldsymbol{A}$ 的列向量组的一个极大无关组为 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_r$，其余 $n-r$ 个列向量都可由该极大无关组线性表示

$$\boldsymbol{A}_{r+i}=k_{1i}\boldsymbol{A}_1+k_{2i}\boldsymbol{A}_2+\cdots+k_{ri}\boldsymbol{A}_r,\qquad i=1,2,\cdots,n-r$$

即

$$-k_{1i}\boldsymbol{A}_1-k_{2i}\boldsymbol{A}_2-\cdots-k_{ri}\boldsymbol{A}_r+\boldsymbol{A}_{r+i}=\boldsymbol{0},\qquad i=1,2,\cdots,n-r$$

上式可写作

$$-k_{1i}\boldsymbol{A}_1-k_{2i}\boldsymbol{A}_2-\cdots-k_{ri}\boldsymbol{A}_r+0\boldsymbol{A}_{r+1}+\cdots+0\boldsymbol{A}_{r+i}+1\boldsymbol{A}_{r+i}+0\boldsymbol{A}_{r+i+1}+\cdots+0\boldsymbol{A}_n=\boldsymbol{0}$$

因此

$$\boldsymbol{\alpha}_1=\begin{bmatrix}-k_{11}\\-k_{21}\\\vdots\\-k_{r1}\\1\\0\\\vdots\\0\end{bmatrix},\boldsymbol{\alpha}_2=\begin{bmatrix}-k_{12}\\-k_{22}\\\vdots\\-k_{r2}\\0\\1\\\vdots\\0\end{bmatrix},\cdots,\boldsymbol{\alpha}_{n-r}=\begin{bmatrix}-k_{1,n-r}\\-k_{2,n-r}\\\vdots\\-k_{r,n-r}\\0\\0\\\vdots\\1\end{bmatrix}$$

都是方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解，显然该向量组线性无关。将矩阵 $\boldsymbol{A}$ 用初等行变换化为行标准形矩阵 $\boldsymbol{B}$，由于初等行变换不改变列向量的线性组合关系，由 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_r$ 线性无关及

$$\boldsymbol{A}_{r+i}=k_{1i}\boldsymbol{A}_1+k_{2i}\boldsymbol{A}_2+\cdots+k_{ri}\boldsymbol{A}_r\qquad i=1,2,\cdots,n-r$$

可知有

$$\boldsymbol{B}=\begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_r & \boldsymbol{K}_1 & \boldsymbol{K}_2 & \cdots & \boldsymbol{K}_{n-r}\end{bmatrix}$$

$$\boldsymbol{K}_i=k_{1i}\boldsymbol{e}_1+k_{2i}\boldsymbol{e}_2+\cdots+k_{ri}\boldsymbol{e}_r=\begin{bmatrix}k_{1i} & k_{2i} & \cdots & k_{ri} & 0 & \cdots & 0\end{bmatrix}^{\mathrm{T}},\qquad i=1,2,\cdots,n-r$$

此时 $\boldsymbol{BX}=\boldsymbol{0}$ 与 $\boldsymbol{AX}=\boldsymbol{0}$ 等价。

$$\boldsymbol{B}=\begin{bmatrix}1 & & & & k_{11} & k_{12} & \cdots & k_{1,n-r}\\ & 1 & & & k_{21} & k_{22} & \cdots & k_{2,n-r}\\ & & \ddots & & \vdots & \vdots & & \vdots\\ & & & 1 & k_{r1} & k_{r2} & \cdots & k_{r,n-r}\\ & & & & & & & 0\\ & & & & & & & \vdots\\ & & & & & & & 0\end{bmatrix}$$

方程组 $\boldsymbol{BX}=\boldsymbol{0}$ 为

$$\begin{cases}x_1 & & & & +k_{11}x_{r+1} & +k_{12}x_{r+2} & \cdots & +k_{1,n-r}x_n & =0\\ & x_2 & & & +k_{21}x_{r+1} & +k_{22}x_{r+2} & \cdots & +k_{2,n-r}x_n & =0\\ & & \ddots & & \vdots & \vdots & & \vdots & \vdots\\ & & & x_r & +k_{r1}x_{r+1} & +k_{r2}x_{r+2} & \cdots & +k_{r,n-r}x_n & =0\\ & & & & & & & 0 & =0\\ & & & & & & & \vdots & \vdots\\ & & & & & & & 0 & =0\end{cases}$$

在上述方程组中，当 $n-r(A)$ 个自由未知量取值 $x_{r+1},x_{r+2},\cdots,x_n$ 时，约束未知量 $x_1,x_2,\cdots,x_r$ 为

$$\begin{cases}x_1=-k_{11}x_{r+1}-k_{12}x_{r+2}-\cdots-k_{1,n-r}x_n\\ x_2=-k_{21}x_{r+1}-k_{22}x_{r+2}-\cdots-k_{2,n-r}x_n\\ \vdots\\ x_r=-k_{r1}x_{r+1}-k_{r2}x_{r+2}-\cdots-k_{r,n-r}x_n\end{cases}$$

所以解为

$$\begin{bmatrix}x_1\\x_2\\\vdots\\x_r\\x_{r+1}\\x_{r+2}\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}-k_{11}x_{r+1}-k_{12}x_{r+2}-\cdots-k_{1,n-r}x_n\\-k_{21}x_{r+1}-k_{22}x_{r+2}-\cdots-k_{2,n-r}x_n\\\vdots\\-k_{r1}x_{r+1}-k_{r2}x_{r+2}-\cdots-k_{r,n-r}x_n\\x_{r+1}\\x_{r+2}\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}-k_{11}x_{r+1}\\-k_{21}x_{r+1}\\\vdots\\-k_{r1}x_{r+1}\\x_{r+1}\\0\\\vdots\\0\end{bmatrix}+\begin{bmatrix}-k_{12}x_{r+2}\\-k_{22}x_{r+2}\\\vdots\\-k_{r2}x_{r+2}\\0\\x_{r+2}\\\vdots\\0\end{bmatrix}+\cdots+\begin{bmatrix}-k_{1,n-r}x_n\\-k_{2,n-r}x_n\\\vdots\\-k_{r,n-r}x_n\\0\\0\\\vdots\\x_n\end{bmatrix}$$

$$=x_{r+1}\begin{bmatrix}-k_{11}\\-k_{21}\\\vdots\\-k_{r1}\\1\\0\\\vdots\\0\end{bmatrix}+x_{r+2}\begin{bmatrix}-k_{12}\\-k_{22}\\\vdots\\-k_{r2}\\0\\1\\\vdots\\0\end{bmatrix}+\cdots+x_n\begin{bmatrix}-k_{1,n-r}\\-k_{2,n-r}\\\vdots\\-k_{r,n-r}\\0\\0\\\vdots\\1\end{bmatrix}=x_{r+1}\boldsymbol{\alpha}_1+x_{r+2}\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_{n-r}$$

因此，该线性方程组的任意一个解都可表示为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 的线性组合，而矩阵 $\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_{n-r}\end{bmatrix}$ 的后 $n-r$ 行构成一个 $n-r$ 阶单位矩阵，故 $r\left(\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_{n-r}\end{bmatrix}\right)=n-r$，故向量组 $r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}\}=n-r$，因此 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 线性无关，即 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 为 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系，因此 $\dim N(\boldsymbol{A})=n-r(\boldsymbol{A})$。

证明方法 2　　设 $r(\boldsymbol{A})=r$，用初等行变换将 $\boldsymbol{A}$ 化为行标准形矩阵 $\boldsymbol{B}$，则 $\boldsymbol{B}$ 有 r 个非零行，$\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1\\\boldsymbol{O}_{(m-r)\times n}\end{bmatrix}$，其中 $\boldsymbol{B}_1$ 为 $r\times n$ 行标准形矩阵，$\boldsymbol{O}_{(m-r)\times n}$ 为 $(m-r)\times n$ 零矩阵，因此 $\boldsymbol{B}_1$ 的列向量组中有 r 维基本向量 $\boldsymbol{e}_1^{(r)},\boldsymbol{e}_2^{(r)},\ldots,\boldsymbol{e}_r^{(r)}$，不妨设 $\boldsymbol{B}_1$ 前 r 列就为 $\boldsymbol{e}_1^{(r)},\boldsymbol{e}_2^{(r)},\ldots,\boldsymbol{e}_r^{(r)}$。构造矩阵 $\boldsymbol{C}=\begin{bmatrix}\boldsymbol{C}_1\\\boldsymbol{O}_{(n-r)\times n}\end{bmatrix}_{n\times n}$，则矩阵 $\boldsymbol{C}$ 具有如下形式

$$\boldsymbol{C}=\begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_r & \boldsymbol{K}_1 & \boldsymbol{K}_2 & \cdots & \boldsymbol{K}_{n-r}\end{bmatrix}$$

其中 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$ 为 n 维基本向量中的前 r 个，$\boldsymbol{K}_1,\boldsymbol{K}_2,\ldots,\boldsymbol{K}_{n-r}$ 是后 $n-r$ 个分量为 0 的 n 维列向量，$\boldsymbol{C}$ 是行标准形矩阵。因此有

$$\boldsymbol{AX}=\boldsymbol{0}\Leftrightarrow\boldsymbol{BX}=\boldsymbol{0}\Leftrightarrow\boldsymbol{B}_1\boldsymbol{X}=\boldsymbol{0}\Leftrightarrow\boldsymbol{CX}=\boldsymbol{0}$$

即 $\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{CX}=\boldsymbol{0}$ 同解。齐次线性方程组 $\boldsymbol{CX}=\boldsymbol{0}$ 的向量表示为

$$x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots x_r\boldsymbol{e}_r+x_{r+1}\boldsymbol{K}_1+x_{r+2}\boldsymbol{K}_2+\cdots+x_n\boldsymbol{K}_{n-r}=0$$

故有

$$x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots x_r\boldsymbol{e}_r=-x_{r+1}\boldsymbol{K}_1-x_{r+2}\boldsymbol{K}_2-\cdots-x_n\boldsymbol{K}_{n-r}$$

行标准形矩阵 $\boldsymbol{C}$ 所对应的方程组中，当 $n-r(\boldsymbol{A})$ 个自由未知量取值 $x_{r+1},x_{r+2},\cdots,x_n$ 时，解

的一般表达式为

$$
\begin{aligned}
\boldsymbol{X} &= \begin{bmatrix} x_1 & x_2 & \cdots & x_r & x_{r+1} & x_{r+2} & \cdots & x_n \end{bmatrix}^{\mathrm{T}} \\
&= x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_r\boldsymbol{e}_r + x_{r+1}\boldsymbol{e}_{r+1} + x_{r+2}\boldsymbol{e}_{r+2} + \cdots + x_n\boldsymbol{e}_n \\
&= -x_{r+1}\boldsymbol{K}_1 - x_{r+2}\boldsymbol{K}_2 - \cdots - x_n\boldsymbol{K}_{n-r} + x_{r+1}\boldsymbol{e}_{r+1} + x_{r+2}\boldsymbol{e}_{r+2} + \cdots + x_n\boldsymbol{e}_n \\
&= x_{r+1}\left(\boldsymbol{e}_{r+1} - \boldsymbol{K}_1\right) + x_{r+2}\left(\boldsymbol{e}_{r+2} - \boldsymbol{K}_2\right) + \cdots + x_n\left(\boldsymbol{e}_n - \boldsymbol{K}_{n-r}\right) \\
&= x_{r+1}\boldsymbol{\alpha}_1 + x_{r+2}\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_{n-r}
\end{aligned}
$$

故该线性方程组的任意一个解都可表示为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 的线性组合，而矩阵 $\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_{n-r}\end{bmatrix}$ 的后 $n-r$ 行构成一个 $n-r$ 阶单位矩阵，故 $r\left(\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_{n-r}\end{bmatrix}\right)=n-r$，故向量组 $r\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}\right\}=n-r$，即 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 线性无关，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 为 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系，因此 $\dim N\left(\boldsymbol{A}\right)=n-r\left(\boldsymbol{A}\right)$。

方法 3　　本方法是方法 2 的一般情形

设 $r\left(\boldsymbol{A}\right)=r$，用初等行变换将 $\boldsymbol{A}$ 化为行标准形矩阵 $\boldsymbol{B}$，则 $\boldsymbol{B}$ 有 r 个非零行，$\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1 \\ \boldsymbol{O}_{(m-r)\times n}\end{bmatrix}$，其中 $\boldsymbol{B}_1$ 为 $r\times n$ 行标准形矩阵，$\boldsymbol{O}_{(m-r)\times n}$ 为 $\left(m-r\right)\times n$ 零矩阵，因此 $\boldsymbol{B}_1$ 的列向量组中有 r 维基本向量 $\boldsymbol{e}_1^{(r)},\boldsymbol{e}_2^{(r)},\ldots,\boldsymbol{e}_r^{(r)}$（这些列向量对应约束未知量），不妨设这 r 个维基本向量 $\boldsymbol{e}_1^{(r)},\boldsymbol{e}_2^{(r)},\ldots,\boldsymbol{e}_r^{(r)}$ 依次位于第 $c_1,c_2,\ldots,c_r$ 列，构造矩阵行阶梯形矩阵 $\boldsymbol{C}=\begin{bmatrix}\boldsymbol{B}_1 \\ \boldsymbol{O}_{(n-r)\times n}\end{bmatrix}_{n\times n}$，矩阵 $\boldsymbol{C}$ 具有如下形式

$$
\begin{array}{cccccccc}
 & & c_1\text{列} & & c_2\text{列} & & c_r\text{列} & \\
\boldsymbol{C}=[\cdots & & \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_r & \cdots \quad]
\end{array}
$$

其中 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$ 为 n 维基本向量中的前 r 个，其它各列依次记为 $\boldsymbol{K}_{d_1},\boldsymbol{K}_{d_2},\ldots,\boldsymbol{K}_{d_{n-r}}\left(d_1<d_2<\cdots<d_{n-r}\right)$，其下标 $d_1,d_2,\ldots,d_{n-r}$ 分别为 $\boldsymbol{K}_{d_1},\boldsymbol{K}_{d_1},\ldots,\boldsymbol{K}_{d_{n-r}}$ 所在列的列号。由于有

$$\boldsymbol{AX}=\boldsymbol{0}\Leftrightarrow\boldsymbol{BX}=\boldsymbol{0}\Leftrightarrow\boldsymbol{B}_1\boldsymbol{X}=\boldsymbol{0}\Leftrightarrow\boldsymbol{CX}=\boldsymbol{0}$$

即 $\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{CX}=\boldsymbol{0}$ 同解。齐次线性方程组 $\boldsymbol{CX}=\boldsymbol{0}$ 的向量表示为

$$\sum_{i=1}^{r}x_{c_i}\boldsymbol{e}_i+\sum_{i=1}^{n-r}x_{d_i}\boldsymbol{K}_{d_i}=\boldsymbol{0}$$

故有

$$\sum_{i=1}^{r} x_{c_i}\boldsymbol{e}_i = -\sum_{i=1}^{n-r} x_{d_i}\boldsymbol{K}_{d_i}$$

令 $\boldsymbol{f}_{d_i}$ 为第 d_i 个 n 维基本向量，则

$$\begin{aligned}\boldsymbol{X} &= \begin{bmatrix} x_1 & x_2 & \cdots & x_r & x_{r+1} & x_{r+2} & \cdots & x_n \end{bmatrix}^{\mathrm{T}} \\ &= \sum_{i=1}^{r} x_{c_i}\boldsymbol{e}_i + \sum_{i=1}^{n-r} x_{d_i}\boldsymbol{f}_{d_i} \\ &= -\sum_{j=1}^{n-r} x_{d_i}\boldsymbol{K}_{d_i} + \sum_{i=1}^{n-r} x_{d_i}\boldsymbol{f}_{d_i} \\ &= \sum_{i=1}^{n-r} x_{d_i}\left(-\boldsymbol{K}_{d_i} + \boldsymbol{f}_{d_i}\right)\end{aligned}$$

在齐次线性方程组 $\boldsymbol{CX}=\boldsymbol{0}$ 中，当 $n-r(\boldsymbol{A})$ 个自由未知量取值 $x_{d_1}, x_{d_1}, \ldots, x_{d_{n-r}}$ 时，解的一般表达式为

$$\boldsymbol{X} = \sum_{i=1}^{n-r} x_{d_i}\left(-\boldsymbol{K}_{d_i} + \boldsymbol{f}_{d_i}\right)$$

由行阶梯形矩阵的性质知，$-\boldsymbol{K}_{d_i} + \boldsymbol{f}_{d_i}, i=1,2,\ldots,n-r$ 线性无关，因而由上式知 $-\boldsymbol{K}_{d_i} + \boldsymbol{f}_{d_i}, i=1,2,\ldots,n-r$ 为 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系，故得 $\dim N(\boldsymbol{A}) = n-r(\boldsymbol{A})$。证毕。

■

（上述“$-\boldsymbol{K}_{d_i} + \boldsymbol{f}_{d_i}, i=1,2,\ldots,n-r$ 线性无关”说明得似乎还不够充分）

注：分析第三种方法得到的结论可知，为求得齐次线性方程组的基础解系，可首先将系数矩阵化为行标准形矩阵，不需要考虑每个非零行的第一个非零元所在的列；对于其它的列，只需要把该列的每个元素改变符号，若该列为第 d_i 列，则把该列的第 d_i 个元素变为 1，且该向量前乘以 x_{d_i}；每一列如此操作，然后相加，则得到通解。（注意：必须考虑解向量的维数！）

推论 5.1　　$m\times n$ 型齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的任意 $n-r(\boldsymbol{A})$ 个线性无关的解都是 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系。

推论 5.2　　$m\times n$ 型齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 有非零解的充要条件是 $r(\boldsymbol{A})<n$。

推论 5.3　　$n\times n$ 型齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 有非零解的充要条件是 $|\boldsymbol{A}|=0$。

例 5.2　　求齐次线性方程组的通解与基础解系

$$\begin{cases} x_1 & +x_2 & +x_3 & +x_1 & =0 \\ 2x_1 & +2x_2 & +x_3 & 3x_4 & =0 \\ x_1 & +x_2 & +2x_3 & & =0 \end{cases}$$

齐次线性方程组的系数矩阵为

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 3 \\ 1 & 1 & 2 & 0 \end{bmatrix}$$

用初等行变换将之化为行标准形矩阵

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{5.6}$$

法 1：与第一种证明方法相对应。

以(5.6)为增广矩阵的线性方程组为

$$\begin{cases} x_1+x_2 \quad +2x_4=0 \\ \qquad\qquad x_3-x_4=0 \\ \qquad\qquad\qquad 0=0 \end{cases}$$

取 x_2,x_4 为自由未知量，其解为

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

写为向量形式

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}=\begin{bmatrix} -x_2-2x_4 \\ x_2 \\ x_4 \\ x_4 \end{bmatrix}=x_2\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}+x_4\begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

这就是通解，而方程组的一个基础解系为

$$\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix},\begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

法 2：与第三种证明方法相对应。

齐次线性方程组的系数矩阵化为行标准形矩阵

作者：卢世荣　lshrlshr@163.com　　欢迎任何意见与建议！

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

故第一、三列对应约束未知量，第一、三个列向量为基本向量，我们只需要考虑第二、四列，由于解向量为 4 维向量，把第二列的每个元素改变符号，得到向量$\begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$，由于该向量位于矩阵的第二列，然后把该向量的第二个元素变为 1，得到向量$\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$，并乘以 x_2；对第四列也如法炮制，首先第四列的每个元素改变符号，得到向量$\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$，然后把第四个元素改为 1，得到向量$\begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$，然后乘以 x_4；因此通解为

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

而基础解系即为相应的一组向量。
显然这种方法更为简单、方便。

例 5.3　设 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times k} = \boldsymbol{O}$，证明 $r(\boldsymbol{A}) + r(\boldsymbol{B}) \le n$。

证明　　令 $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_1 & \boldsymbol{B}_2 & \cdots & \boldsymbol{B}_k \end{bmatrix}$，则由 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times k} = \boldsymbol{O}$ 可知 $\boldsymbol{A}\boldsymbol{B}_i = \boldsymbol{0}$，　$i = 1, 2, \cdots, k$，这说明 $\boldsymbol{B}$ 的列向量都是齐次线性方程组 $\boldsymbol{AX} = \boldsymbol{0}$ 的解。设 $r(\boldsymbol{A}) = r$，则 $\dim N(\boldsymbol{A}) = n - r$，不妨设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{n-r}$ 为 $\boldsymbol{AX} = \boldsymbol{0}$ 的基础解系，则 $\boldsymbol{B}_1, \boldsymbol{B}_2, \cdots, \boldsymbol{B}_k$ 可由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{n-r}$ 线性表示，故 $r\{\boldsymbol{B}_1, \boldsymbol{B}_2, \cdots, \boldsymbol{B}_k\} \le r\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{n-r}\} = n - r$，即 $r(\boldsymbol{B}) \le n - r(\boldsymbol{A})$，因此 $r(\boldsymbol{A}) + r(\boldsymbol{B}) \le n$。

例 5.4　证明：矩阵 $\boldsymbol{A}_{m\times n}$ 与 $\boldsymbol{B}_{l\times n}$ 的行向量组等价的充分必要条件是，齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{BX}=\boldsymbol{0}$ 同解。

证明（1）必要性：考虑方程组

$$\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{bmatrix}\boldsymbol{X}=\boldsymbol{0}$$

矩阵 $\boldsymbol{A}_{m\times n}$ 与 $\boldsymbol{B}_{l\times n}$ 的行向量组等价，故 $\boldsymbol{B}$ 的行向量组可由 $\boldsymbol{A}$ 的行向量组线性表示，故通过初等行变换，可将其系数矩阵化为

$$\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{bmatrix}\sim\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{bmatrix}$$

因此

$$\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{bmatrix}\boldsymbol{X}=\boldsymbol{0},\quad \begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{bmatrix}\boldsymbol{X}=\boldsymbol{0}$$

同解，而显然

$$\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{bmatrix}\boldsymbol{X}=\boldsymbol{0},\quad \boldsymbol{AX}=\boldsymbol{0}$$

同解，故 $\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{bmatrix}\boldsymbol{X}=\boldsymbol{0},\quad \boldsymbol{AX}=\boldsymbol{0}$。同理可证

$$\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{bmatrix}\boldsymbol{X}=\boldsymbol{0},\quad \boldsymbol{BX}=\boldsymbol{0}$$

同解，命题得证。

（2）充分性：由 $\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{BX}=\boldsymbol{0}$ 同解，则

$$\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{bmatrix}\boldsymbol{X}=\boldsymbol{0},\quad \boldsymbol{AX}=\boldsymbol{0},\quad \boldsymbol{BX}=\boldsymbol{0}$$

同解，所以

$$r\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{bmatrix}=r(\boldsymbol{A})=r(\boldsymbol{B})\Rightarrow r\left(\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}} & \boldsymbol{B}^{\mathrm{T}}\end{bmatrix}\right)=r\left(\boldsymbol{A}^{\mathrm{T}}\right)=r\left(\boldsymbol{B}^{\mathrm{T}}\right)$$

故根据定理 3.10 可知，$\boldsymbol{A}^{\mathrm{T}}$ 与 $\boldsymbol{B}^{\mathrm{T}}$ 的列向量组等价，即 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的行向量组等价。

例 5.5　设 n 阶方阵 $\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}$ 满足 $\left|a_{ii}\right|>\sum\limits_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right|,\quad i=1,2,\cdots,n$，证明 $\boldsymbol{A}$ 是可逆矩阵。

证明　用反证法。假设 $\boldsymbol{A}$ 不可逆，则齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 有非零解

$\boldsymbol{X}=\left[x_1,x_2,\cdots,x_n\right]^{\mathrm{T}}$。设 $\left|x_k\right|=\max\left\{\left|x_i\right|,i=1,2,\cdots,n\right\}$，由于 $\boldsymbol{X}\neq\boldsymbol{0}$，故 $\left|x_k\right|>0$；则由 $\boldsymbol{AX}=\boldsymbol{0}$ 的第 k 个方程 $\sum_{j=1}^{n}a_{kj}x_j=0$ 得到，$a_{kk}x_k=-\sum_{\substack{j=1\\j\neq k}}^{n}a_{kj}x_j$，因此

$$\left|a_{kk}x_k\right|=\left|-\sum_{\substack{j=1\\j\neq k}}^{n}a_{kj}x_j\right|=\left|\sum_{\substack{j=1\\j\neq k}}^{n}a_{kj}x_j\right|\le\sum_{\substack{j=1\\j\neq k}}^{n}\left|a_{kj}x_j\right|=\sum_{\substack{j=1\\j\neq k}}^{n}\left|a_{kj}\right|\left|x_j\right|$$
$$\le\sum_{\substack{j=1\\j\neq k}}^{n}\left|a_{kj}\right|\left|x_k\right|$$

因此 $\left|a_{kk}\right|\le\sum_{\substack{j=1\\j\neq k}}^{n}\left|a_{kj}\right|$，这与题设矛盾。因此假设不成立。命题得证。

例 5.6　设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，证明 $r\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)=r\left(\boldsymbol{A}\right)$。

证：$\boldsymbol{A}$ 为 $m\times n$ 矩阵，则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 为 n 阶方阵，取齐次线性方程组

$$m\times n\text{ 型}\qquad \boldsymbol{AX}=\boldsymbol{0}$$
$$n\times n\text{ 型}\qquad \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{X}=\boldsymbol{0}$$

先证 $\boldsymbol{AX}=\boldsymbol{0}$ 与 $\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{X}=\boldsymbol{0}$ 为等价的线性方程组。

若 $\boldsymbol{AX}=\boldsymbol{0}$，则 $\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{AX}\right)=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{0}=\boldsymbol{0}$，所以 $\boldsymbol{AX}=\boldsymbol{0}$ 的解为 $\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{X}=\boldsymbol{0}$ 的解，即 $N\left(\boldsymbol{A}\right)\subseteq N\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$。

反过来，若 $\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{X}=\boldsymbol{0}$，则

$$\boldsymbol{X}^{\mathrm{T}}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{X}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{0}=0$$

即　$0=\boldsymbol{X}^{\mathrm{T}}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{X}=\left(\boldsymbol{AX}\right)^{\mathrm{T}}\left(\boldsymbol{AX}\right)=\left\|\boldsymbol{AX}\right\|^2$，所以 $\boldsymbol{AX}=\boldsymbol{0}$。故 $\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{X}=\boldsymbol{0}$ 的解为 $\boldsymbol{AX}=\boldsymbol{0}$ 的解。所以这两个方程组的解空间相同，即

$$N\left(\boldsymbol{A}\right)=N\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$$

故　$$\dim N\left(\boldsymbol{A}\right)=\dim N\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$$

因此　$$n-r\left(\boldsymbol{A}\right)=\dim N\left(\boldsymbol{A}\right)=\dim N\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)=n-r\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$$

即　$$r\left(\boldsymbol{A}\right)=r\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$$

由于 $r(\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}})=r\left((\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\right)=r(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})$，所以

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

例 5.7　线性方程组 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 有解。

证明　只需要证明 $r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=r\left(\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} & \boldsymbol{A}\boldsymbol{b}\end{bmatrix}\right)$。

显然有 $r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\le r\left(\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\end{bmatrix}\right)$。另外

$$r\left(\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\end{bmatrix}\right)=r\left(\boldsymbol{A}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{A} & \boldsymbol{b}\end{bmatrix}\right)\le r(\boldsymbol{A}^{\mathrm{T}})=r(\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})。$$

因此 $r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=r\left(\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} & \boldsymbol{A}\boldsymbol{b}\end{bmatrix}\right)$ 成立，故本定理成立。■

因此该定理回答了正交投影存在性的问题。

习题 5.2

1. 设 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的两个不同的解，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是对应的齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系，k_1,k_2 为任意常数，则 $\boldsymbol{AX}=\boldsymbol{b}$ 的通解为（　）B

（A）$k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)+\frac{1}{2}(\boldsymbol{\beta}_1-\boldsymbol{\beta}_2)$　（B）$k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)+\frac{1}{2}(\boldsymbol{\beta}_1+\boldsymbol{\beta}_2)$

（C）$k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\beta}_1+\boldsymbol{\beta}_2)+\frac{1}{2}(\boldsymbol{\beta}_1-\boldsymbol{\beta}_2)$　（D）$k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\beta}_1-\boldsymbol{\beta}_2)+\frac{1}{2}(\boldsymbol{\beta}_1+\boldsymbol{\beta}_2)$

2. 求齐次线性方程组

$$\begin{cases}x_1+x_2+x_3+4x_4-3x_5=0\\2x_1+x_2+3x_3+5x_4-5x_5=0\\x_1-x_2+3x_3-2x_4-x_5=0\\3x_1+x_2+5x_3+6x_4-7x_5=0\end{cases}$$

的一个基础解系.

3. 已知四元线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+a_{13}x_3+a_{14}x_4=b_1\\a_{21}x_1+a_{22}x_2+a_{23}x_3+a_{24}x_4=b_2\\a_{31}x_1+a_{32}x_2+a_{33}x_3+a_{34}x_4=b_3\end{cases}$$

的 5 个解 $X_1=[1,1,1,0]^{\mathrm{T}},X_2=[2,2,1,0]^{\mathrm{T}},X_3=[0,0,1,0]^{\mathrm{T}},X_4=[3,1,2,0]^{\mathrm{T}}$，$X_5=[0,1,1,1]^{\mathrm{T}}$，试求此方程组的一般解，并写出该方程组。

4. 设齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的系数矩阵 $\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}$ 且 $|\boldsymbol{A}|=0$，但 $\boldsymbol{A}$ 中某一元素 a_{ij} 的

代数余子式 $A_{ij}\neq 0$。证明：方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解恒可表示为

$$x_1=kA_{1i},x_2=kA_{2i},\cdots,x_n=kA_{ni}$$

5. 设方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 都是 $\boldsymbol{BX}=\boldsymbol{0}$ 的解，且 $r(\boldsymbol{A})=r(\boldsymbol{B})$，试证：$\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{BX}=\boldsymbol{0}$ 同解。

6. 设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，$\boldsymbol{B}$ 为 $n\times s$ 矩阵，证明：$(\boldsymbol{AB})\boldsymbol{X}=\boldsymbol{0}$ 与 $\boldsymbol{BX}=\boldsymbol{0}$ 同解的充要条件是 $r(\boldsymbol{AB})=r(\boldsymbol{B})$。

7. 设 $\boldsymbol{A},\boldsymbol{B}$ 分别为 $m\times n$、$n\times k$ 矩阵，若 $\boldsymbol{AB}=\boldsymbol{O}$，用齐次线性方程组解的理论证明矩阵秩的不等式 $r(\boldsymbol{A})+r(\boldsymbol{B})\leq n$。

8. 设 $\boldsymbol{A}$ 是 n 阶方阵，$r(\boldsymbol{A})=r<n$，证明：存在秩为 $n-r$ 的 n 阶方阵 $\boldsymbol{B},\boldsymbol{C}$，使得 $\boldsymbol{AB}=\boldsymbol{CA}=\boldsymbol{O}$。

9. 设四元方程组（I）$\begin{cases}2x_1+3x_2-x_3=0\\x_1+2x_2+x_3-x_4=0\end{cases}$，另一个方程组（II）的基础解系为 $\boldsymbol{\alpha}_1=[2,-1,a+2,1]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=[-1,2,4,a+8]^{\mathrm{T}}$。求（I）的一个基础解系；当 a 为何值时，（I）与（II）有非零公共解，并求解。

10. 已知齐次线性方程组

$$\begin{cases}(a_1+b)x_1+a_2x_2+a_3x_3+\cdots+a_nx_n=0,\\a_1x_1+(a_2+b)x_2+a_3x_3+\cdots+a_nx_n=0,\\a_1x_1+a_2x_2+(a_3+b)x_3+\cdots+a_nx_n=0,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\a_1x_1+a_2x_2+a_3x_3+\cdots+(a_n+b)x_n=0.\end{cases}$$

其中 $\sum\limits_{i=1}^{n}a_i\neq 0.$ 试讨论 $a_1,a_2,\cdots,a_n$ 和 b 满足何种关系时，

(1)方程组仅有零解；

(2)方程组有非零解，在有非零解时，求此方程组的一个基础解系.

11. 设有齐次线性方程组

$$\begin{cases}(1+a)x_1+x_2+\cdots+x_n=0\\2x_1+(2+a)x_2+\cdots+2x_n=0\\\qquad\vdots\\nx_1+nx_2+\cdots+(n+a)x_n=0\end{cases}\quad(n\geq 2)$$

试问 a 取何值时，该方程组有非零解，并求出其通解。

12. 设 $\boldsymbol{A}$ 是 n 阶方阵，证明：存在 n 阶非零矩阵 $\boldsymbol{B}$，使 $\boldsymbol{AB}=\boldsymbol{O}$ 的充要条件是：$|\boldsymbol{A}|=0$。

13. 设 ξ_1,ξ_2,ξ_3 是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的一个基础解系，试证明 $\xi_1-\xi_2,\xi_1+\xi_2+\xi_3,\xi_1+2\xi_3$ 也是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的一个基础解系.

14. 问：a,b,c,d,e 满足什么条件，齐次线性方程组

$$\begin{cases} \quad\ \ x_2 + ax_3 + bx_4 = 0 \\ -x_1 + cx_3 + dx_4 = 0 \\ ax_1 + cx_2 - ex_4 = 0 \\ bx_1 + dx_2 + ex_3 = 0 \end{cases}$$

的一般解中，可取 x_3 和 x_4 作为自由未知量？

15. 设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，$r(\boldsymbol{A})=m$，$\boldsymbol{B}$ 是 m 阶可逆矩阵，已知 $\boldsymbol{A}$ 的行空间是方程组 $\boldsymbol{CX}=\boldsymbol{0}$ 的解空间。证明：$\boldsymbol{BA}$ 的行向量组也是 $\boldsymbol{CX}=\boldsymbol{0}$ 的一个基础解系。

16. 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$m<n$，$r(\boldsymbol{A})=m$，齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的一个基础解系为

$$\boldsymbol{B}_i=(b_{i1},b_{i2},\cdots,b_{in})^{\mathrm{T}},i=1,2,\cdots,n-m$$

试求齐次线性方程组

$$\begin{cases} b_{11}y_1+b_{12}y_2+\cdots+b_{1n}y_n=0 \\ b_{21}y_1+b_{22}y_2+\cdots+b_{2n}y_n=0 \\ \qquad\qquad\vdots \\ b_{n-m,1}y_1+b_{n-m,2}y_2+\cdots+b_{n-m,n}y_n=0 \end{cases}$$

的解空间的维数，以及一个基础解系。

17. 设方程组 $\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1,2n}x_{2n}=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2,2n}x_{2n}=0 \\ \qquad\qquad\vdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{n,2n}x_{2n}=0 \end{cases}$ 的一个基础解系为

$$\boldsymbol{\beta}_1=\left[b_{11},b_{12},\ldots,b_{1,2n}\right]^{\mathrm{T}},\boldsymbol{\beta}_2=\left[b_{21},b_{22},\ldots,b_{2,2n}\right]^{\mathrm{T}},\cdots,\boldsymbol{\beta}_n=\left[b_{n1},b_{n2},\ldots,b_{n,2n}\right]^{\mathrm{T}}$$

写出方程组

$$\begin{cases} b_{11}x_1+b_{12}x_2+\cdots+b_{12n}x_{2n}=0 \\ b_{21}x_1+b_{22}x_2+\cdots+b_{22n}x_{2n}=0 \\ \qquad\qquad\vdots \\ b_{n1}x_1+b_{n2}x_2+\cdots+b_{n2n}x_{2n}=0 \end{cases}$$

的通解.

§5.3　非齐次线性方程组解的结构

非齐次线性方程组与其对应的齐次线性方程组的解有相应的关系。本节讨论这一关系。

定理 5.4　设 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 是非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的解，则 $\boldsymbol{X}_1-\boldsymbol{X}_2$ 是对应的齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解。

证明：$\boldsymbol{X}_1,\boldsymbol{X}_2$ 是 $\boldsymbol{AX}=\boldsymbol{b}$ 的解，故

$$\boldsymbol{AX}_1=\boldsymbol{b}$$
$$\boldsymbol{AX}_2=\boldsymbol{b}$$

两式相减得

$$\boldsymbol{A}\left(\boldsymbol{X}_1-\boldsymbol{X}_2\right)=\boldsymbol{0}$$

故 $\boldsymbol{X}_1-\boldsymbol{X}_2$ 为 $\boldsymbol{AX}=\boldsymbol{0}$ 的解。　■

定理 5.5　非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的解可写成 $\boldsymbol{\beta}+\boldsymbol{\gamma}$ 的形式，其中 $\boldsymbol{\beta}$ 为该线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的一个特解，而 $\boldsymbol{\gamma}$ 为其对应的齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解。

证明：设 $\boldsymbol{X}$ 是非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的任意一个解，根据上述定理可知，$\boldsymbol{X}-\boldsymbol{\beta}$ 是它所对应的齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解，而

$$\boldsymbol{X}=\left(\boldsymbol{X}-\boldsymbol{\beta}\right)+\boldsymbol{\beta}$$

令 $\boldsymbol{\gamma}=\boldsymbol{X}-\boldsymbol{\beta}$ 即可。　■

因此，若 $\dim N\left(\boldsymbol{A}\right)=s$，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系，$\boldsymbol{X}_0$ 为 $\boldsymbol{AX}=\boldsymbol{b}$ 的一个特解，则非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 其任一解 $\boldsymbol{X}$ 都可表示为

$$\boldsymbol{X}=\boldsymbol{X}_0+k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s \tag{5.7}$$

称为 $\boldsymbol{AX}=\boldsymbol{b}$ 的**通解**。

例 5.8　求非齐次线性方程组的通解。

$$\begin{cases} x_1-x_2-x_3+x_4=0 \\ x_1-x_2+x_3-3x_4=1 \\ x_1-x_2-2x_3+3x_4=-\dfrac{1}{2} \end{cases}$$

解：该方程组的增广矩阵为

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

用初等行变换将其化为行最简形矩阵，得

$$\begin{bmatrix}1 & -1 & 0 & -1 & 1/2\\ 0 & 0 & 1 & -2 & 1/2\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

可见 $r(\boldsymbol{A})=r(\boldsymbol{B})=2$，故方程组有解，且

$$\begin{cases}x_1=x_2+x_4+\dfrac{1}{2}\\ x_3=2x_4+\dfrac{1}{2}\end{cases}$$

写成解向量的形式：

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}x_2+x_4+1/2\\x_2\\2x_4+1/2\\x_4\end{bmatrix}=\begin{bmatrix}1/2\\0\\1/2\\0\end{bmatrix}+\begin{bmatrix}x_2\\x_2\\0\\0\end{bmatrix}+\begin{bmatrix}x_4\\0\\2x_4\\x_4\end{bmatrix}=\begin{bmatrix}1/2\\0\\1/2\\0\end{bmatrix}+x_2\begin{bmatrix}1\\1\\0\\0\end{bmatrix}+x_4\begin{bmatrix}1\\0\\2\\1\end{bmatrix}$$

为线性方程组的通解。

注：上述求非齐次线性方程组的通解的过程是标准的，容易理解且不易出错。但注意到非齐次线性方程组的通解可表示为“特解＋对应的齐次线性方程组的通解”的形式，且在§5.2已经给出了求齐次线性方程组通解的简便方法，因此只需要给出求非齐次线性方程组特解的方法即可，而这是很容易的。假设该非齐次线性方程组的增广矩阵已经化为行最简形，那么它的一个特解可通过如下简单规律给出：每个约束未知量取等号右边的常数值，而自由未知量则取 0。例如上例中，增广矩阵的行最简形为 $\begin{bmatrix}1 & -1 & 0 & -1 & 1/2\\ 0 & 0 & 1 & -2 & 1/2\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$，约束未知量为 x_1,x_3，这两个约束未知量所在方程等号右边的常数为 $\dfrac{1}{2},\dfrac{1}{2}$，故令 $x_1=\dfrac{1}{2},x_3=\dfrac{1}{2}$，$x_2,x_4$ 为自由未知量，故令 $x_2=x_4=0$，因此该非齐次线性方程组的一个特解为 $\left[\dfrac{1}{2},0,\dfrac{1}{2},0\right]^{\mathrm{T}}$。

例 5.9　问 k 取何值时，线性方程组

$$\begin{cases}kx_1+x_2+x_3=5\\ 3x_1+2x_2+kx_3=18-5k\\ \qquad x_2+2x_3=2\end{cases}$$

有唯一解,无穷多解或无解？有无穷多解时求出通解。
解：
法 1：直接求系数矩阵、增广矩阵的秩，该线性方程组的增广矩阵为

$$[\boldsymbol{A}\quad \boldsymbol{b}]=\begin{bmatrix} k & 1 & 1 & 5 \\ 3 & 2 & k & 18-5k \\ 0 & 1 & 2 & 2 \end{bmatrix}$$

作初等行变换得

$$\sim\begin{bmatrix} 3 & 2 & k & 18-5k \\ k & 1 & 1 & 5 \\ 0 & 1 & 2 & 2 \end{bmatrix}\sim\begin{bmatrix} 3 & 2 & k & 18-5k \\ 0 & 0 & 1-\frac{k^2}{3}-2+\frac{4k}{3} & 5-\frac{k}{3}(18-5k)-2\left(1-\frac{2k}{3}\right) \\ 0 & 1 & 2 & 2 \end{bmatrix}$$

$$\sim\begin{bmatrix} 3 & 2 & k & 18-5k \\ 0 & 1 & 2 & 2 \\ 0 & 0 & -\frac{k^2}{3}+\frac{4k}{3}-1 & \frac{5k^2}{3}-\frac{14k}{3}+3 \end{bmatrix}\sim\begin{bmatrix} 3 & 2 & k & 18-5k \\ 0 & 1 & 2 & 2 \\ 0 & 0 & -k^2+4k-3 & 5k^2-14k+9 \end{bmatrix}$$

所以，当$-k^2+4k-3\neq 0$即$k\neq 1,3$时，

$r(\boldsymbol{A})=r([\boldsymbol{A}\quad \boldsymbol{b}])=3$，线性方程组有惟一解。

当$k=1$时

$$-k^2+4k-3=0 \qquad 5k^2-14k+9=0$$

$r(\boldsymbol{A})=r([\boldsymbol{A}\quad \boldsymbol{b}])=2$，线性方程组有无穷多组解；

当$k=3$时

$-k^2+4k-3=0$，$5k^2-14k+9\neq 0$，$2=r(\boldsymbol{A})\neq r([\boldsymbol{A}\quad \boldsymbol{b}])=3$，线性方程组无解。

法 2：注意到该线性方程组的系数矩阵为一个方阵，故可利用 Cramer 法则

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

当$D\neq 0$，即$k\neq 1,3$时线性方程组有唯一解。

$k=1$时，线性方程组的增广矩阵为

$$\begin{bmatrix}A & b\end{bmatrix}=\left[\begin{array}{ccc|c}1&1&1&5\\3&2&1&13\\0&1&2&2\end{array}\right]\sim\left[\begin{array}{ccc|c}1&0&-1&3\\0&1&2&2\\0&0&0&0\end{array}\right]$$

故有无穷多解。由

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

其通解为

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}3\\2\\0\end{bmatrix}+c\begin{bmatrix}1\\-2\\1\end{bmatrix}$$

$k=3$时，$\begin{bmatrix}\boldsymbol{A} & \boldsymbol{b}\end{bmatrix}=\left[\begin{array}{ccc|c}3&1&1&5\\3&2&3&3\\0&1&2&2\end{array}\right]\sim\left[\begin{array}{ccc|c}3&1&1&5\\0&1&2&-2\\0&0&0&4\end{array}\right]$，因而$r(\boldsymbol{A})<r\left(\begin{bmatrix}\boldsymbol{A} & \boldsymbol{b}\end{bmatrix}\right)$，线性方程组无解。

例 5.10　给出一个通解为

$$\boldsymbol{X}=\begin{bmatrix}-16\\23\\0\\0\\0\end{bmatrix}+k_1\begin{bmatrix}1\\-2\\1\\0\\0\end{bmatrix}+k_2\begin{bmatrix}1\\-2\\0\\1\\0\end{bmatrix}+k_3\begin{bmatrix}5\\-6\\0\\0\\1\end{bmatrix}$$

的线性方程组。

解：设所求线性方程组为$\boldsymbol{AX}=\boldsymbol{b}$，线性方程组有 5 个未知数，而其导出方程组的基础解系含有 3 个向量

$$\boldsymbol{\alpha}_1=\begin{bmatrix}1\\-2\\1\\0\\0\end{bmatrix},\quad \boldsymbol{\alpha}_2=\begin{bmatrix}1\\-2\\0\\1\\0\end{bmatrix},\quad \boldsymbol{\alpha}_3=\begin{bmatrix}5\\-6\\0\\0\\1\end{bmatrix},$$

$$\Rightarrow \boldsymbol{r}(\boldsymbol{A})=5-3=2$$

可设$\boldsymbol{A}$有两行，且这两个行向量线性无关。由于

$$\boldsymbol{A\alpha}_1=\boldsymbol{0},\quad \boldsymbol{A\alpha}_2=\boldsymbol{0},\quad \boldsymbol{A\alpha}_3=\boldsymbol{0}$$

$$\Rightarrow \boldsymbol{A}\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3\end{bmatrix}=\begin{bmatrix}\boldsymbol{A\alpha}_1 & \boldsymbol{A\alpha}_2 & \boldsymbol{A\alpha}_3\end{bmatrix}=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}=\boldsymbol{O}$$

令$\boldsymbol{B}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3\end{bmatrix}$，则有$\boldsymbol{AB}=\boldsymbol{O}$，取转置得 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{O}$，故$\boldsymbol{A}^{\mathrm{T}}$的列向量为齐次线性

方程组 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{0}$ 的解。

$$\boldsymbol{B}^{\mathrm{T}}=\begin{bmatrix}1&-2&1&0&0\\1&-2&0&1&0\\5&-6&0&0&1\end{bmatrix}\overset{r}{\sim}\begin{bmatrix}1&0&0&-\frac{3}{2}&\frac{1}{2}\\0&1&0&-\frac{5}{4}&\frac{1}{4}\\0&0&1&-1&0\end{bmatrix}$$

因此 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{0}$ 的通解为

$$\boldsymbol{Y}=k_1\begin{bmatrix}\frac{3}{2}\\\frac{5}{4}\\1\\1\\0\end{bmatrix}+k_2\begin{bmatrix}-\frac{1}{2}\\-\frac{1}{4}\\0\\0\\1\end{bmatrix}$$

取它的两个线性无关的解

$$\begin{bmatrix}6\\5\\4\\4\\0\end{bmatrix},\qquad\begin{bmatrix}-2\\-1\\0\\0\\4\end{bmatrix}$$

则得到　$\boldsymbol{A}^{\mathrm{T}}=\begin{bmatrix}6&-2\\5&-1\\4&0\\4&0\\0&4\end{bmatrix}$

所以　　　$\boldsymbol{A}=\begin{bmatrix}6&5&4&4&0\\-2&-1&0&0&4\end{bmatrix}$

因为 $\begin{bmatrix}-16\\23\\0\\0\\0\end{bmatrix}$ 是 $\boldsymbol{AX}=\boldsymbol{b}$ 的一个特解，所以将这个解代入方程组，

得到 $\begin{bmatrix}6&5&4&4&0\\-2&-1&0&0&4\end{bmatrix}\boldsymbol{X}=\begin{bmatrix}19\\9\end{bmatrix}$ 为所求的方程组。

例 5.10　设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，$r(\boldsymbol{A})=r$， $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 为齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系， $\boldsymbol{\xi}$ 为非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的一个特解，证明：

(1)　$\boldsymbol{\xi},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 线性无关；

(2)　$\boldsymbol{\xi},\boldsymbol{\alpha}_1+\boldsymbol{\xi},\boldsymbol{\alpha}_2+\boldsymbol{\xi},\cdots,\boldsymbol{\alpha}_{n-r}+\boldsymbol{\xi}$ 为 $\boldsymbol{AX}=\boldsymbol{b}$ 的线性无关的解；

(3)　$\boldsymbol{AX}=\boldsymbol{b}$ 的任意 $n-r+2$ 个解线性相关。

(4)　$\boldsymbol{AX}=\boldsymbol{b}$ 的任意一个解 $\boldsymbol{X}$ 可写成

$$\boldsymbol{X}=k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2+\cdots+k_{n-r+1}\boldsymbol{X}_{n-r+1}$$

其中 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_{n-r+1}$ 为 $\boldsymbol{AX}=\boldsymbol{b}$ 的 $n-r+1$ 个线性无关的解，且

$$k_1+k_2+\cdots+k_{n-r+1}=1$$

证明：(1) 由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 为 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系，故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 线性无关。若 $\boldsymbol{\xi},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 线性相关，则 $\boldsymbol{\xi}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 线性表示，故 $\boldsymbol{\xi}$ 为 $\boldsymbol{AX}=\boldsymbol{0}$ 的解，$\Rightarrow \boldsymbol{b}=\boldsymbol{0}$。矛盾。

(2) 根据非齐次线性方程组解的结构可知，向量组 $\boldsymbol{\xi},\boldsymbol{\alpha}_1+\boldsymbol{\xi},\boldsymbol{\alpha}_2+\boldsymbol{\xi},\cdots,\boldsymbol{\alpha}_{n-r}+\boldsymbol{\xi}$ 中的每个向量都为 $\boldsymbol{AX}=\boldsymbol{b}$ 的解；下证明证明线性无关性。

设有 $k,k_1,k_2,\cdots,k_{n-r}$ 使得

$$k\boldsymbol{\xi}+k_1(\boldsymbol{\alpha}_1+\boldsymbol{\xi})+k_2(\boldsymbol{\alpha}_2+\boldsymbol{\xi})+\cdots+k_{n-r}(\boldsymbol{\alpha}_{n-r}+\boldsymbol{\xi})=\boldsymbol{0}$$

$$(k+k_1+k_2+\cdots+k_{n-r})\boldsymbol{\xi}+k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r}=\boldsymbol{0}$$

由于 $\boldsymbol{\xi},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 线性无关，得到

$$\begin{cases} k+k_1+k_2+\cdots+k_{n-r}=0 \\ k_1=0 \\ k_2=0 \\ \vdots \\ k_{n-r}=0 \end{cases}$$

$$\Rightarrow k=k_1=k_2=\cdots=k_{n-r}=0$$

故 $\boldsymbol{\xi},\boldsymbol{\alpha}_1+\boldsymbol{\xi},\boldsymbol{\alpha}_2+\boldsymbol{\xi},\cdots,\boldsymbol{\alpha}_{n-r}+\boldsymbol{\xi}$ 线性无关。

也可如此证明线性无关性：

由于$\boldsymbol{\xi},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$线性无关，因此$r\left(\begin{bmatrix}\boldsymbol{\xi} & \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_{n-r}\end{bmatrix}\right)=n-r+1$，然后对矩阵$\begin{bmatrix}\boldsymbol{\xi} & \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_{n-r}\end{bmatrix}$做初等列变换：从第二列到最后一列的每一列都加上第一列，由于初等变换不改变矩阵的秩，因此有

$$r\left(\begin{bmatrix}\boldsymbol{\xi} & \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_{n-r}\end{bmatrix}\right)=r\left(\begin{bmatrix}\boldsymbol{\xi} & \boldsymbol{\xi}+\boldsymbol{\alpha}_1 & \boldsymbol{\xi}+\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\xi}+\boldsymbol{\alpha}_{n-r}\end{bmatrix}\right)$$

$$\Rightarrow r\left(\begin{bmatrix}\boldsymbol{\xi} & \boldsymbol{\xi}+\boldsymbol{\alpha}_1 & \boldsymbol{\xi}+\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\xi}+\boldsymbol{\alpha}_{n-r}\end{bmatrix}\right)=n-r+1$$

故$\boldsymbol{\xi},\boldsymbol{\alpha}_1+\boldsymbol{\xi},\boldsymbol{\alpha}_2+\boldsymbol{\xi},\cdots,\boldsymbol{\alpha}_{n-r}+\boldsymbol{\xi}$线性无关。

（3）由于非齐次线性方程组$\boldsymbol{AX}=\boldsymbol{b}$的通解为

$$\boldsymbol{X}=\boldsymbol{\xi}+k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r}$$

因此$\boldsymbol{AX}=\boldsymbol{b}$的任一解都可由向量组$\boldsymbol{\xi},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$线性表示，由定理 3.6 可知，它的任意$n-r+2$个解都线性相关。

(4) 由(2)知，$\boldsymbol{AX}=\boldsymbol{b}$有$n-r+1$个线性无关的解，设为$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_{n-r+1}$；由(3)，$\boldsymbol{AX}=\boldsymbol{b}$任意$n-r+2$个解线性相关，因此任取$\boldsymbol{AX}=\boldsymbol{b}$的解$\boldsymbol{X}$，则$\boldsymbol{X},\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_{n-r+1}$线性相关，故$\boldsymbol{X}$可由$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_{n-r+1}$线性表示，因此有$k_1,k_2,\cdots,k_{n-r+1}$，使得

$$\boldsymbol{X}=k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2+\cdots+k_{n-r+1}\boldsymbol{X}_{n-r+1}$$

两边同时乘以$\boldsymbol{A}$，得到

$$\begin{aligned}\boldsymbol{b}=\boldsymbol{AX}&=\boldsymbol{A}\left(k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2+\cdots+k_{n-r+1}\boldsymbol{X}_{n-r+1}\right)\\&=k_1\boldsymbol{AX}_1+k_2\boldsymbol{AX}_2+\cdots+k_{n-r+1}\boldsymbol{AX}_{n-r+1}\\&=k_1\boldsymbol{b}+k_2\boldsymbol{b}+\cdots+k_{n-r+1}\boldsymbol{b}=\left(k_1+k_2+\cdots+k_{n-r+1}\right)\boldsymbol{b}\end{aligned}$$

由于$\boldsymbol{b}\neq\boldsymbol{0}$，所以得到

$$k_1+k_2+\cdots+k_{n-r+1}=1。$$

（
线性图形（直线、平面）的几何问题两条直线的关系：
相交、共面、异面
两个平面之间的关系：平行、相交、重合三个平面的位置关系
）

习题 5.3

1. 设$\boldsymbol{\alpha}_1=[1,2,0]^{\mathrm{T}},\boldsymbol{\alpha}_2=[1,a+2,-3a]^{\mathrm{T}},\boldsymbol{\alpha}_3=[-1,-b-2,a+2b]^{\mathrm{T}},\boldsymbol{\beta}=[1,3,-3]^{\mathrm{T}}$，请讨论当$a,b$为何值时，

（1）$\boldsymbol{\beta}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示；

（2）$\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 唯一地线性表示，并求出表示式；

（3）$\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，但表示方式不唯一，并求出表示式。

2. 求实系数线性方程组 $\begin{cases} x_1-x_2=a_1 \\ x_2-x_3=a_2 \\ \vdots \\ x_{n-1}-x_n=a_{n-1} \\ x_n-x_1=a_n \end{cases}$ 有解的充要条件，并在它有解时求出通解。

3. 设 $\boldsymbol{A}=\begin{bmatrix} -6 & 2 & 10 \\ -4 & 1 & 7 \\ -3 & 1 & a-2 \end{bmatrix}$ 是齐次线性方程组（Ⅰ）的系数矩阵，$\boldsymbol{\beta}=[b,c,1]^{\mathrm{T}}$ 是齐次线性方程组（Ⅱ）的基础解系，已知线性方程组（Ⅰ）与（Ⅱ）同解。

（1）求 a,b,c；（2）求非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{\beta}$ 的通解。

4. 已知 4 元线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的三个解 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$，且 $\boldsymbol{\xi}_1+\boldsymbol{\xi}_2=[1,0,1,2]^{\mathrm{T}},\boldsymbol{\xi}_3=[3,2,-1,4]^{\mathrm{T}}$，而 $r(\boldsymbol{A})=3$，求该线性方程组的通解。

5. 已知线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的系数矩阵 $\boldsymbol{A}=[\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\ \boldsymbol{\alpha}_3\ \boldsymbol{\alpha}_4]$，其中 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关，$\boldsymbol{\alpha}_1=2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3,\boldsymbol{b}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4$，求 $\boldsymbol{AX}=\boldsymbol{b}$ 的通解。

6. 已知 4 阶方阵 $\boldsymbol{A}=[\boldsymbol{\alpha}_1\quad \boldsymbol{\alpha}_2\quad \boldsymbol{\alpha}_3\quad \boldsymbol{\alpha}_4]$，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 均为 4 维列向量，其中 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关，$\boldsymbol{\alpha}_1=2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$，如果 $\boldsymbol{\beta}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4$，求线性方程组 $\boldsymbol{AX}=\boldsymbol{\beta}$ 的通解.

7. 设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 对任意 $\boldsymbol{b}\in\mathbb{R}^m$ 都有解的充要条件是 $r(\boldsymbol{A})=m$。

8. 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{\beta}=[b_1,b_2,\cdots,b_n]^{\mathrm{T}}$ 是 m 维列向量，证明下属命题相互等价：

（1）线性方程组 $\boldsymbol{AX}=\boldsymbol{\beta}$ 有解；

（2）齐次方程组 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{0}$ 的任一解 $\boldsymbol{X}=[x_1,x_2,\cdots,x_m]^{\mathrm{T}}$ 必满足

$$x_1b_1+x_2b_2+\cdots+x_mb_m=0$$

（3）线性方程组$\begin{bmatrix} \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{\beta}^{\mathrm{T}} \end{bmatrix} \boldsymbol{X} = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix}$无解，其中 0 是$n$维零向量。

9. 设$\boldsymbol{A}$是$m \times n$实矩阵，b是m维实列向量，证明

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

（2）线性方程组$(\boldsymbol{A}^{T}\boldsymbol{A})X = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$一定有解。

10. 设$\boldsymbol{A}$是$m \times n$实矩阵，$m > n$，非齐次线性方程组$\boldsymbol{AX} = \boldsymbol{b}$有唯一解。证明：$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$为可逆矩阵，且该方程组的解为$\boldsymbol{X} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$。

11. 设$\boldsymbol{A}$是$m \times n$矩阵，$r(\boldsymbol{A}) = r$，$b_1, b_2, \cdots, b_m$是一组不全为零的实数，如果矩阵$\boldsymbol{B}$满足

$$\boldsymbol{AB} = \begin{bmatrix} b_1 & b_1 & \cdots & b_1 \\ b_2 & b_2 & \cdots & b_2 \\ \vdots & \vdots & & \vdots \\ b_m & b_m & \cdots & b_m \end{bmatrix}$$

证明：$r(\boldsymbol{B}) \le n - r + 1$。

12. 设n阶方阵$\boldsymbol{A}$的秩为r，并且非齐次线性方程组$\boldsymbol{AX} = \boldsymbol{b}$有无穷多组解，其所有解的集合记为$S$.

（1）集合S是否为向量空间？说明理由.

（2）证明集合S的极大线性无关组的秩为$n - r + 1$.

第六章　　相似矩阵

特征值的介绍、背景。

§6.1　特征值与特征向量

定义 6.1　设 $\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}\in\mathbb{C}^{n\times n}$ 为 n 阶方阵，若有数 $\lambda\in\mathbb{C}$ 与非零向量 $\boldsymbol{X}\in\mathbb{C}^{n}$，使得

$$\boldsymbol{AX}=\lambda\boldsymbol{X} \tag{6.1}$$

则称 λ 为 $\boldsymbol{A}$ 的特征值，$\boldsymbol{X}$ 为方阵 $\boldsymbol{A}$ 的对应于特征值 λ 的特征向量，简称为对应于特征值 λ 的特征向量。

根据上述定义，若方阵 $\boldsymbol{A}$ 不可逆，则齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 有非零解 $\boldsymbol{X}_0$，所以

$$\boldsymbol{AX}_0=\boldsymbol{0}=0\boldsymbol{X}_0$$

这说明 0 是 $\boldsymbol{A}$ 的特征值，而该方程组的非零解 $\boldsymbol{X}_0$ 就是 $\boldsymbol{A}$ 的对应于特征值 0 的特征向量。

定理 6.1　　同一方阵的不同特征值没有相同的特征向量。

证明：用反证法。设 λ_1,λ_2 方阵 $\boldsymbol{A}$ 的两个不同的特征值，若它们有相同的特征向量 $\boldsymbol{X}$，则

$$\boldsymbol{AX}=\lambda_1\boldsymbol{X}$$
$$\boldsymbol{AX}=\lambda_2\boldsymbol{X}$$

两式相减得

$$\boldsymbol{0}=\left(\lambda_1-\lambda_2\right)\boldsymbol{X}$$

由于 $\lambda_1\neq\lambda_2$，故 $\lambda_1-\lambda_2\neq 0$，$\Rightarrow\boldsymbol{X}=0$，这与 $\boldsymbol{X}$ 为特征向量矛盾。

根据上述命题，对于同一个方阵而言，其特征值与特征向量是相对应的，所有我们有定义 6.1 中的概念。（若 $\boldsymbol{AX}=\lambda\boldsymbol{X}$，则称 $\boldsymbol{X}$ 为 $\boldsymbol{A}$ 的对应于特征值 λ 的特征向量，而 λ 是 $\boldsymbol{A}$ 的对应于特征向量 $\boldsymbol{X}$ 的特征值。）

下面求给定方阵 $\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}$ 的特征值与特征向量。设 λ 为 $\boldsymbol{A}$ 的特征值，$\boldsymbol{X}$ 为方阵 $\boldsymbol{A}$ 的对应于特征值 λ 的特征向量，则有

$$\boldsymbol{AX}=\lambda\boldsymbol{X}$$

变形得

$$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$$

由于 $\boldsymbol{X}\neq\boldsymbol{0}$，这说明齐次线性方程组$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$有非零解，从而方阵$\lambda\boldsymbol{E}-\boldsymbol{A}$不可逆，所以其行列式

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=0$$

记

$$f(\lambda)=|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix} \tag{6.2}$$

根据行列式的定义可知，$|\lambda \boldsymbol{E}-\boldsymbol{A}|$是关于$\lambda$的$n$次多项式，且$\lambda^n$的系数为 1，即有

$$\begin{aligned} f(\lambda)=|\lambda \boldsymbol{E}-\boldsymbol{A}|&=\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix} \\ &=\lambda^n+b_1\lambda^{n-1}+\cdots+b_{n-1}\lambda+b_n \end{aligned}$$

我们称之为$\boldsymbol{A}$的**特征多项式**，而关于λ的方程$f(\lambda)=|\lambda \boldsymbol{E}-\boldsymbol{A}|=0$称为$\boldsymbol{A}$的**特征方程**。因此，$\boldsymbol{A}$的特征值必为其特征方程的根；反过来，若$\lambda$为$\boldsymbol{A}$的特征方程的根，则说明

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=0$$

因此方阵$\lambda\boldsymbol{E}-\boldsymbol{A}$不可逆，即$r(\lambda \boldsymbol{E}-\boldsymbol{A})<n$，因此$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$有非零解$\boldsymbol{X}_0$，即存在非零向量$\boldsymbol{X}_0$，使得

$$\boldsymbol{A}\boldsymbol{X}_0=\lambda\boldsymbol{X}_0$$

所以λ为$\boldsymbol{A}$特征值。

根据上述分析可知，求$\boldsymbol{A}$的特征方程的根就得到它的特征值。由于n阶方阵$\boldsymbol{A}$的特征多项式为$|\lambda \boldsymbol{E}-\boldsymbol{A}|$为$n$次多项式，因此其特征方程$|\lambda \boldsymbol{E}-\boldsymbol{A}|=0$在复数范围内一定有$n$个根（重根按重数计算），因此$n$阶方阵$\boldsymbol{A}$一定有$n$个特征值。

对于$\boldsymbol{A}$的特征值λ，求齐次线性方程组

$$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0} \tag{6.3}$$

的非零解就得到$\boldsymbol{A}$的对应于特征值λ的特征向量。由第五章知识可知，齐次线性方程组$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$的解集构成向量空间，称之为特征值$\lambda$的特征子空间，记为$V_\lambda$，所以有

$$V_\lambda = \{\boldsymbol{X} | (\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}\} = \{\boldsymbol{X} | \boldsymbol{AX}=\lambda\boldsymbol{X}\}$$
$$= \{\boldsymbol{X} | \boldsymbol{AX}=\lambda\boldsymbol{X}, \boldsymbol{X}\neq 0\} \cup \{\boldsymbol{0}\}$$

因此V_λ的元素包括$\boldsymbol{A}$的对应于λ的特征向量与零向量。$\dim V_\lambda$称为λ的**几何重数**，根据向量空间的维数的定义可知，λ的**几何重数**就是对应于特征值λ的线性无关的特征向量的个数。

例 6.1　求矩阵

（1）$\boldsymbol{A}=\begin{bmatrix}-1&1&0\\-4&3&0\\1&0&2\end{bmatrix}$；　　（2）$\boldsymbol{B}=\begin{bmatrix}5&-4&1\\2&-4&2\\-1&4&3\end{bmatrix}$；

（3）$\boldsymbol{C}=\begin{bmatrix}0&1&0\\-1&0&0\\0&0&1\end{bmatrix}$。

的特征值与特征向量。

解：（1）$\boldsymbol{A}$的特征多项式为

$$|\lambda\boldsymbol{E}-\boldsymbol{A}| = \begin{vmatrix}\lambda+1&-1&0\\4&\lambda-3&0\\-1&0&\lambda-2\end{vmatrix} = (\lambda-2)\begin{vmatrix}\lambda+1&-1\\4&\lambda-3\end{vmatrix}$$
$$=(\lambda-2)\left[(\lambda+1)(\lambda-3)+4\right]=(\lambda-2)(\lambda-1)^2$$

故$\boldsymbol{A}$的特征值为$\lambda_1=2,\lambda_2=\lambda_3=1$，因此特征值$\lambda_1=2$的代数重数为 1，而$\lambda_2=1$的代数重数为 2。

当$\lambda_1=2$时，求解线性方程组$(\lambda_1\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=0$就可得到对应于特征值$\lambda_1=2$的特征向量

$$\lambda_1\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}3&-1&0\\4&-1&0\\-1&0&0\end{bmatrix}\overset{r}{\sim}\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$$

其基础解系为$\boldsymbol{\alpha}_1=[0,0,1]^{\mathrm{T}}$，故$k\boldsymbol{\alpha}_1\,(k\neq 0)$为对应于$\lambda_1=2$的所有特征向量。

$\lambda_2=\lambda_3=1$，求解线性方程组$(\lambda_2\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=0$就可得到对应于特征值$\lambda_2=\lambda_3=1$的特征向量

$$\lambda_2\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}2&-1&0\\4&-2&0\\-1&0&-1\end{bmatrix}\overset{r}{\sim}\begin{bmatrix}1&0&1\\0&1&2\\0&0&0\end{bmatrix}$$

故基础解系为$\boldsymbol{\alpha}_2=[-1,-2,1]^{\mathrm{T}}$，故$k\boldsymbol{\alpha}_2\,(k\neq 0)$为对应于$\lambda_2=\lambda_3=1$的所有特征向量。故特

征值 1 的代数重数是 2，而其几何重数为 1；

（2）

$$|\lambda E-B|=\begin{vmatrix}\lambda-5 & 4 & -1\\ -2 & \lambda+4 & -2\\ 1 & -4 & \lambda-3\end{vmatrix}\overset{c_1-c_3}{=\!=}\begin{vmatrix}\lambda-4 & 4 & -1\\ 0 & \lambda+4 & -2\\ 4-\lambda & -4 & \lambda-3\end{vmatrix}$$

$$=(\lambda-4)\begin{vmatrix}1 & 4 & -1\\ 0 & \lambda+4 & -2\\ -1 & -4 & \lambda-3\end{vmatrix}\overset{r_3+r_1}{=\!=}(\lambda-4)\begin{vmatrix}1 & 4 & -1\\ 0 & \lambda+4 & -2\\ 0 & 0 & \lambda-4\end{vmatrix}$$

$$=(\lambda+4)(\lambda-4)^2$$

故 $\boldsymbol{B}$ 的特征值为 $\lambda_1=-4,\lambda_2=\lambda_3=4$。

当 $\lambda_1=-4$ 时，$\lambda_1\boldsymbol{E}-\boldsymbol{B}=-4\boldsymbol{E}-\boldsymbol{B}=\begin{bmatrix}-9 & 4 & -1\\ -2 & 0 & -2\\ 1 & -4 & -7\end{bmatrix}\xrightarrow{\text{初等行变换}}\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 2\\ 0 & 0 & 0\end{bmatrix}$

因此 $\lambda_1=-4$ 对应的特征向量为 $k[-1,-2,1]^{\mathrm{T}},k\neq0$；

当 $\lambda_2=\lambda_3=4$ 时，$\lambda_2\boldsymbol{E}-\boldsymbol{B}=4\boldsymbol{E}-\boldsymbol{B}=\begin{bmatrix}-1 & 4 & -1\\ -2 & 8 & -2\\ 1 & -4 & 1\end{bmatrix}\xrightarrow{\text{行}}\begin{bmatrix}1 & -4 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$；

故 $\lambda_2=\lambda_3=4$ 对应的特征向量为 $k_1[4,1,0]^{\mathrm{T}}+k_2[-1,0,1]^{\mathrm{T}},|k_1|+|k_2|\neq0$。因此特征值 4 的代数重数是 2，其几何重数也是 2，也就是它的几何重数与代数重数相等。

（3） $|\lambda\boldsymbol{E}-\boldsymbol{C}|=\begin{vmatrix}\lambda & -1 & 0\\ 1 & \lambda & 0\\ 0 & 0 & \lambda-1\end{vmatrix}=(\lambda^2+1)(\lambda-1)=0$，故 $\boldsymbol{C}$ 的特征值为 $\lambda_1=i,\lambda_2=-i,\lambda_3=1$。

当 $\lambda_1=i$ 时，$\lambda_1\boldsymbol{E}-\boldsymbol{C}=i\boldsymbol{E}-\boldsymbol{B}=\begin{bmatrix}i & -1 & 0\\ 1 & i & 0\\ 0 & 0 & i-1\end{bmatrix}\xrightarrow{\text{初等行变换}}\begin{bmatrix}1 & i & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}$，故 $\lambda_1=i$ 对应的特征向量为 $k[-i,1,0]^{\mathrm{T}},k\neq0$；

当 $\lambda_2=-i$ 时，$\lambda_2\boldsymbol{E}-\boldsymbol{C}=-i\boldsymbol{E}-\boldsymbol{B}=\begin{bmatrix}-i & -1 & 0\\ 1 & -i & 0\\ 0 & 0 & -i-1\end{bmatrix}\xrightarrow{\text{初等行变换}}\begin{bmatrix}1 & -i & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}$，故 $\lambda_2=-i$ 对应的特征向量为 $k[i,1,0]^{\mathrm{T}},k\neq0$；

当$\lambda_3=1$时，$\lambda_3\boldsymbol{E}-\boldsymbol{C}=\boldsymbol{E}-\boldsymbol{B}=\begin{bmatrix}1&-1&0\\1&1&0\\0&0&0\end{bmatrix}\xrightarrow{\text{初等行变换}}\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$，故$\lambda_3=1$对应的特征向量为$k[0,0,1]^{\mathrm{T}},k\neq 0$。

定义 6.3　　对于方阵$\boldsymbol{A}$，$tr(\boldsymbol{A})=\sum_{i=1}^{n}a_{ii}$称为方阵$\boldsymbol{A}$的**迹**。

由于n阶方阵的特征多项式是一个n次多项式，方阵的特征值就是特征方程的根，根据多项式方程根与系数的关系，有如下定理：

定理 6.2　　设n阶方阵$\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}$的$n$个特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$，则

（1）　$\lambda_1\lambda_2\cdots\lambda_n=|\boldsymbol{A}|$；

（2）　$\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}=tr(\boldsymbol{A})$。

证明：由于n阶方阵$\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}$的$n$个特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$，且特征多项式$|\lambda\boldsymbol{E}-\boldsymbol{A}|$的首系数为 1，所以

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-a_{11}&a_{12}&\cdots&a_{1n}\\-a_{21}&\lambda-a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\-a_{n1}&-a_{n2}&\cdots&\lambda-a_{nn}\end{vmatrix}=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)\qquad(6.4)$$

（1）在上式中令$\lambda=0$，得

$$|-\boldsymbol{A}|=(-1)^n\lambda_1\lambda_2\cdots\lambda_n$$

故$|\boldsymbol{A}|=\lambda_1\lambda_2\cdots\lambda_n$；

（2）为证明这个等式，我们考虑λ^{n-1}的系数，由于行列式展开后，每一项取自不同行、不同列的n个元素的乘积，因此只有对角线上的n个元素$\lambda-a_{11},\lambda-a_{22},\cdots,\lambda-a_{nn}$相乘才会出现$\lambda^{n-1}$这一项，展开式中的其它的项最多只包含$n-2$个对角线上的元素，因此不会有$\lambda^{n-1}$这一项，因此特征多项式中包含$\lambda^{n-1}$的项只在$(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})$中出现，而$(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})$中$\lambda^{n-1}$中的系数为$-(a_{11}+a_{22}+\cdots+a_{nn})$，$(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$展开式中$\lambda^{n-1}$的系数为$-(\lambda_1+\lambda_2+\cdots+\lambda_n)$，比较两边$\lambda^{n-1}$前

的系数，所以

$$\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}。$$

■

上述定理称为韦达定理。它反映的是多项式的根与系数的关系。

根据如上定理，有

推论 6.1　n 阶方阵 $\boldsymbol{A}$ 可逆的充要条件为：$\boldsymbol{A}$ 的 n 个特征值非零。

定义 6.3　特征值的**代数重数**

设 n 阶方阵 $\boldsymbol{A}$ 有 s 个互不相同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_s$，则 $\boldsymbol{A}$ 的特征多项式可表示为

$$f(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda-\lambda_1)^{t_1}(\lambda-\lambda_2)^{t_2}\cdots(\lambda-\lambda_s)^{t_s}，\qquad t_1,t_2,\cdots,t_s \text{ 是正整数}$$

我们称 t_i 为特征值 λ_i 的**代数重数**。因此，若特征值 λ_i 的**代数重数**为 t_i，则 λ_i 为特征方程 $f(\lambda)=0$ 的 t_i 重根。根据代数重数的定义，显然有

$$t_1+t_2+\cdots+t_s=n \tag{6.5}$$

例 6.2　设 λ 是方阵 $\boldsymbol{A}$ 对应于特征向量 $\boldsymbol{X}$ 的特征值，证明：

(1) $k\lambda$ 是 $k\boldsymbol{A}$ 对应于特征向量 $\boldsymbol{X}$ 的特征值；

(2) 对正整数 $m(m\geq 2)$，λ^m 是 $\boldsymbol{A}^m$ 对应于特征向量 $\boldsymbol{X}$ 的特征值；

(3) $f(x)=a_kx^k+\cdots+a_1x+a_0$ 为 x 的多项式，λ 为方阵 $\boldsymbol{A}$ 的特征值，$\boldsymbol{X}$ 为 $\boldsymbol{A}$ 的对应于 λ 的特征向量，则 $f(\lambda)$ 为方阵 $f(\boldsymbol{A})$ 的特征值，则 $\boldsymbol{X}$ 为 $f(\boldsymbol{A})$ 的对应于特征值 $f(\lambda)$ 的特征向量。

(4) 若 $\boldsymbol{A}$ 是可逆的，则 λ^{-1} 是 $\boldsymbol{A}^{-1}$ 对应于特征向量 $\boldsymbol{X}$ 的特征值。

证明：(1) 由于 $\boldsymbol{AX}=\lambda\boldsymbol{X}$，因而

$$(k\boldsymbol{A})\boldsymbol{X}=k(\boldsymbol{AX})=k\lambda\boldsymbol{X}=(k\lambda)\boldsymbol{X}$$

所以结论(1)成立。

(2)

$$\begin{aligned}\boldsymbol{A}^m\boldsymbol{X}&=(\boldsymbol{A}^{m-1}\boldsymbol{A})\boldsymbol{X}=\boldsymbol{A}^{m-1}(\boldsymbol{AX})\\&=\boldsymbol{A}^{m-1}(\lambda\boldsymbol{X})=\lambda\boldsymbol{A}^{m-1}\boldsymbol{X}\\&=\cdots=\lambda^m\boldsymbol{X}\end{aligned}$$

结论(2)成立。

（3）

$$f\left(\boldsymbol{A}\right)X=\left(a_k\boldsymbol{A}^k+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{E}\right)\boldsymbol{X}$$
$$=a_k\boldsymbol{A}^k\boldsymbol{X}+\cdots+a_1\boldsymbol{A}\boldsymbol{X}+a_0\boldsymbol{E}\boldsymbol{X}$$
$$=a_k\lambda^k\boldsymbol{X}+\cdots+a_1\lambda\boldsymbol{X}+a_0\boldsymbol{X}$$
$$=\left(a_k\lambda^k+\cdots+a_1\lambda+a_0\right)\boldsymbol{X}$$
$$=f\left(\lambda\right)\boldsymbol{X}$$

（4）由于 $\boldsymbol{A}$ 可逆，所以的所有特征值都不等于 0，故 $\lambda\neq 0$，因而

$$\boldsymbol{AX}=\lambda\boldsymbol{X}$$

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

结论(3)成立。

例 6.3　设 n 阶方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2=\boldsymbol{A}$，证明：$\boldsymbol{A}$ 的特征值为 0 或 1。

证明：设 λ 为 $\boldsymbol{A}$ 的特征值，而 $\boldsymbol{X}$ 为与 λ 对应的特征向量，故

$$\boldsymbol{AX}=\lambda\boldsymbol{X}$$

由于　$$\boldsymbol{A}^2=\boldsymbol{A}$$

所以　$$\boldsymbol{A}^2\boldsymbol{X}=\boldsymbol{AX}$$

$$\lambda^2\boldsymbol{X}=\lambda\boldsymbol{X}$$

由于 $\boldsymbol{X}\neq\boldsymbol{0}$，所以

$$\lambda^2=\lambda$$

故命题成立。

定理 6.3*　（Hamilton-Caylay 定理）设 A 是一个 n 阶方阵，$f(\lambda)=\left|\lambda\boldsymbol{E}-\boldsymbol{A}\right|$ 为 $\boldsymbol{A}$ 的特征多项式，则 $f\left(\boldsymbol{A}\right)=\boldsymbol{O}$（零矩阵）。

证明　设 $\boldsymbol{B}\left(\lambda\right)$ 是 $\lambda\boldsymbol{E}-\boldsymbol{A}$ 的伴随矩阵，由行列式的性质，有

$$\boldsymbol{B}\left(\lambda\right)\left(\lambda\boldsymbol{E}-\boldsymbol{A}\right)=\left|\lambda\boldsymbol{E}-\boldsymbol{A}\right|\boldsymbol{E}=f\left(\lambda\right)\boldsymbol{E}$$

因为矩阵 $\boldsymbol{B}\left(\lambda\right)$ 的元素是行列式 $\left|\lambda\boldsymbol{E}-\boldsymbol{A}\right|$ 的各个余子式，都是 λ 的多项式，其次数不超过 $n-1$，因此由矩阵的运算性质，$\boldsymbol{B}\left(\lambda\right)$ 可写成

$$\boldsymbol{B}\left(\lambda\right)=\lambda^{n-1}\boldsymbol{B}_0+\lambda^{n-2}\boldsymbol{B}_1+\cdots+\lambda\boldsymbol{B}_{n-2}+\boldsymbol{B}_{n-1}$$

其中$\boldsymbol{B}_0,\boldsymbol{B}_1,\cdots,\boldsymbol{B}_{n-2},\boldsymbol{B}_{n-1}$都是$n-1$阶数字方阵。再设

$$f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$$

因而

$$\begin{aligned}\boldsymbol{B}(\lambda)(\lambda\boldsymbol{E}-\boldsymbol{A})&=\left(\lambda^{n-1}\boldsymbol{B}_0+\lambda^{n-2}\boldsymbol{B}_1+\cdots+\lambda\boldsymbol{B}_{n-2}+\boldsymbol{B}_{n-1}\right)(\lambda\boldsymbol{E}-\boldsymbol{A})\\&=\lambda^n\boldsymbol{B}_0+\lambda^{n-1}(\boldsymbol{B}_1-\boldsymbol{B}_0\boldsymbol{A})+\lambda^{n-2}(\boldsymbol{B}_2-\boldsymbol{B}_1\boldsymbol{A})+\cdots+\lambda(\boldsymbol{B}_{n-1}-\boldsymbol{B}_{n-2}\boldsymbol{A})-\boldsymbol{B}_{n-1}\boldsymbol{A}\end{aligned}$$

比较上述两式，得到

$$\begin{cases}\boldsymbol{B}_0=\boldsymbol{E}\\\boldsymbol{B}_1-\boldsymbol{B}_0\boldsymbol{A}=a_1\boldsymbol{E}\\\boldsymbol{B}_2-\boldsymbol{B}_1\boldsymbol{A}=a_2\boldsymbol{E}\\\qquad\vdots\\\boldsymbol{B}_{n-1}-\boldsymbol{B}_{n-2}\boldsymbol{A}=a_{n-1}\boldsymbol{E}\\-\boldsymbol{B}_{n-1}\boldsymbol{A}=a_n\boldsymbol{E}\end{cases}$$

以$\boldsymbol{A}^n,\boldsymbol{A}^{n-1},\cdots,\boldsymbol{A},\boldsymbol{E}$依次从右边乘上式中的第一式、第二式，…，第$n$式，第$n+1$式，得

$$\begin{cases}\boldsymbol{B}_0\boldsymbol{A}^n=\boldsymbol{E}\boldsymbol{A}^n=\boldsymbol{A}^n\\(\boldsymbol{B}_1-\boldsymbol{B}_0\boldsymbol{A})\boldsymbol{A}^{n-1}=\boldsymbol{B}_1\boldsymbol{A}^{n-1}-\boldsymbol{B}_0\boldsymbol{A}^n=a_1\boldsymbol{E}\boldsymbol{A}^{n-1}=a_1\boldsymbol{A}^{n-1}\\(\boldsymbol{B}_2-\boldsymbol{B}_1\boldsymbol{A})\boldsymbol{A}^{n-2}=\boldsymbol{B}_2\boldsymbol{A}^{n-1}-\boldsymbol{B}_1\boldsymbol{A}^n=a_2\boldsymbol{E}\boldsymbol{A}^{n-2}=a_2\boldsymbol{A}^{n-2}\\\qquad\vdots\\(\boldsymbol{B}_{n-1}-\boldsymbol{B}_{n-2}\boldsymbol{A})\boldsymbol{A}=\boldsymbol{B}_{n-1}\boldsymbol{A}-\boldsymbol{B}_{n-2}\boldsymbol{A}^2=a_{n-1}\boldsymbol{E}\boldsymbol{A}=a_{n-1}\boldsymbol{A}\\-\boldsymbol{B}_{n-1}\boldsymbol{A}=a_n\boldsymbol{E}\end{cases}$$

把上式中的$n+1$个式子加起来，左边变为零矩阵，右边即为$f(\boldsymbol{A})$，故$f(\boldsymbol{A})=\boldsymbol{0}$。

■

例 6.4　设三阶方阵$\boldsymbol{A}$的三个特征值分别为 2，3，7，求行列式$|5\boldsymbol{A}+\boldsymbol{E}|$。

解：由于$\boldsymbol{A}$的三个特征值分别为 2，3，7，所以$5\boldsymbol{A}+\boldsymbol{E}$的三个特征值分别为 5×2+1， 5×3+1 ，5×7+1，所以$|5\boldsymbol{A}+\boldsymbol{E}|=11\times16\times36=6336$。

下面讨论特征向量的性质，我们集中于特征向量的线性相关性。

定理 6.4　(1)设$\boldsymbol{X}_1,\boldsymbol{X}_2$为对应于方阵$\boldsymbol{A}$的特征值$\lambda$的特征向量，则$\boldsymbol{X}_1,\boldsymbol{X}_2$的非零线性组合也是对应于特征值$\lambda$的特征向量；

(2)设$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_r$为是方阵$\boldsymbol{A}$的不同特征值所对应的特征向量，则$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_r$线性

无关；

(3)方阵 $\boldsymbol{A}$ 的 s 个不同特征值所对应的 s 组各自线性无关的特征向量并在一起仍是线性无关的。

(4)设 λ_0 是 n 阶方阵 $\boldsymbol{A}$ 的一个 t 重特征值，则与 λ_0 对应的线性无关特征向量的最大个数 $\le t$，即 $\dim V_{\lambda_0} \le t$。

证明：

(1)结论是显然的；

(2) 证法 1　用归纳法。

设 $r=2$，$\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 为方阵 $\boldsymbol{A}$ 对应于特征值 λ_1 与 λ_2 的两个特征向量，即

$$\boldsymbol{A}\boldsymbol{X}_1=\lambda_1\boldsymbol{X}_1, \qquad \boldsymbol{A}\boldsymbol{X}_2=\lambda_2\boldsymbol{X}_2$$

设有两个数 k_1,k_2 使得

$$k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2=\boldsymbol{0} \tag{6.6}$$

为证明 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 线性无关，只需要证明上式中 $k_1=k_2=0$。在**错误！未找到引用源。**式两边左乘方阵 $\boldsymbol{A}$，得到

$$\boldsymbol{A}\left(k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2\right)=\boldsymbol{A0}=\boldsymbol{0}$$

即

$$\boldsymbol{0}=k_1\boldsymbol{A}\boldsymbol{X}_1+k_2\boldsymbol{A}\boldsymbol{X}_2=k_1\lambda_1\boldsymbol{X}_1+k_2\lambda_2\boldsymbol{X}_2 \tag{6.7}$$

错误！未找到引用源。式乘以数 λ_1 得

$$\boldsymbol{0}=k_1\lambda_1\boldsymbol{X}_1+k_2\lambda_1\boldsymbol{X}_2 \tag{6.8}$$

(6.7)－(6.8)得：

$$k_2\boldsymbol{X}_2\left(\lambda_2-\lambda_1\right)=\boldsymbol{0}$$

因为 $\lambda_2-\lambda_1\ne 0$，而 $\boldsymbol{X}_2$ 为一个特征向量，因而不为零向量，因此 $k_2=0$，进而可得 $k_1=0$，因此 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2$ 线性无关，故命题在 $r=2$ 时成立。

假设命题在 r 时成立。即：　r 个不同特征值所对应的特征向量线性无关，则当 $r+1$ 时，考虑

$$k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2+\cdots+k_r\boldsymbol{X}_r+k_{r+1}\boldsymbol{X}_{r+1}=\boldsymbol{0} \tag{6.9}$$

在上式两边同时左乘方阵 $\boldsymbol{A}$ 得

$$\boldsymbol{A}\left(k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2+\cdots+k_r\boldsymbol{X}_r+k_{r+1}\boldsymbol{X}_{r+1}\right)=\boldsymbol{0}$$

$$k_1\boldsymbol{AX}_1+k_2\boldsymbol{AX}_2+\cdots+k_r\boldsymbol{AX}_r+k_{r+1}\boldsymbol{AX}_{r+1}=\mathbf{0}$$

$$k_1\lambda_1\boldsymbol{X}_1+k_2\lambda_2\boldsymbol{X}_2+\cdots+k_r\lambda_r\boldsymbol{X}_r+k_{r+1}\lambda_{r+1}\boldsymbol{X}_{r+1}=\mathbf{0} \tag{6.10}$$

在(6.9)式两边同时乘以 λ_{r+1} 得

$$k_1\lambda_{r+1}\boldsymbol{X}_1+k_2\lambda_{r+1}\boldsymbol{X}_2+\cdots+k_r\lambda_{r+1}\boldsymbol{X}_r+k_{r+1}\lambda_{r+1}\boldsymbol{X}_{r+1}=\mathbf{0} \tag{6.11}$$

(6.11)－ (6.10)得

$$k_1(\lambda_{r+1}-\lambda_1)\boldsymbol{X}_1+k_2(\lambda_{r+1}-\lambda_2)\boldsymbol{X}_2+\cdots+k_r(\lambda_{r+1}-\lambda_r)\boldsymbol{X}_r=\mathbf{0}$$

根据归纳假设可知， $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_r$ 线性无关，故

$$k_1(\lambda_{r+1}-\lambda_1)=k_2(\lambda_{r+1}-\lambda_2)=\cdots=k_r(\lambda_{r+1}-\lambda_r)=0$$

由于各特征值互不相同，因此 $\lambda_{r+1}-\lambda_i\neq 0$ ，故

$$k_1=k_2=\cdots=k_r=0$$

再由(6.9)得到 $k_{r+1}=0$，故特征向量 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_r,\boldsymbol{X}_{r+1}$ 线性无关。命题得证。

证法 2　设 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_r$ 是 $\boldsymbol{A}$ 的不同特征值对应的特征向量

设有数 $s_1,s_2,\cdots,s_r$，使得 $s_1\boldsymbol{X}_1+s_2\boldsymbol{X}_2+\cdots+s_r\boldsymbol{X}_r=\mathbf{0}$，则

$$A^k(s_1\boldsymbol{X}_1+s_2\boldsymbol{X}_2+\cdots+s_r\boldsymbol{X}_r)=\mathbf{0}(k=1,2,\cdots,r-1)$$

即

$$\lambda_1^k s_1\boldsymbol{X}_1+\lambda_2^k s_2\boldsymbol{X}_2+\cdots+\lambda_r^k s_r\boldsymbol{X}_r=\mathbf{0}(k=1,2,\cdots,r-1)$$

将上述各式写成矩阵形式，有

$$\begin{bmatrix} s_1\boldsymbol{X}_1 & s_2\boldsymbol{X}_2 & \cdots & s_r\boldsymbol{X}_r \end{bmatrix}\begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{r-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_r & \cdots & \lambda_r^{r-1} \end{bmatrix}=\begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$$

由于 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 互不相同，故方阵 $\begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{r-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_r & \cdots & \lambda_r^{r-1} \end{bmatrix}$ 可逆，所以

$$\begin{bmatrix} s_1\boldsymbol{X}_1 & s_2\boldsymbol{X}_2 & \cdots & s_r\boldsymbol{X}_r \end{bmatrix}=\boldsymbol{O}$$

因此 $s_i\boldsymbol{X}_i=\mathbf{0}(i=1,2,\cdots,r)$，注意到 $\boldsymbol{X}_i\neq\mathbf{0}(i=1,2,\cdots,r)$，故 $s_i=\mathbf{0}(i=1,2,\cdots,r)$，因此

$\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_r$ 线性无关。

（3）设 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 为是方阵 $\boldsymbol{A}$ 的 s 个不同特征值，而

$$\boldsymbol{X}_{i1}, \boldsymbol{X}_{i2}, \cdots, \boldsymbol{X}_{im_i}$$

为方阵 $\boldsymbol{A}$ 的对应于特征值 λ_i 的一组线性无关的特征向量，我们要证明

$$\boldsymbol{X}_{11}, \boldsymbol{X}_{12}, \cdots, \boldsymbol{X}_{1m_1}, \boldsymbol{X}_{21}, \boldsymbol{X}_{22}, \cdots, \boldsymbol{X}_{2m_2}, \cdots, \boldsymbol{X}_{s1}, \boldsymbol{X}_{s2}, \cdots, \boldsymbol{X}_{sm_s}$$

线性无关，设有常数

$$c_{i1}, c_{i2}, \cdots, c_{im_i}, i = 1, 2, \cdots, s$$

使得

$$\begin{aligned}
& c_{11}\boldsymbol{X}_{11} + c_{12}\boldsymbol{X}_{12} + \cdots + c_{1m_1}\boldsymbol{X}_{1m_1} \\
& + c_{21}\boldsymbol{X}_{21} + c_{22}\boldsymbol{X}_{22} + \cdots + c_{2m_1}\boldsymbol{X}_{2m_2} \\
& \cdots \\
& + c_{s1}\boldsymbol{X}_{s1} + c_{s2}\boldsymbol{X}_{s2} + \cdots + c_{sm_s}\boldsymbol{X}_{sm_s} = \boldsymbol{0}
\end{aligned}$$

即

$$\sum_{j=1}^{m_1} c_{1j}\boldsymbol{X}_{1j} + \sum_{j=1}^{m_2} c_{2j}\boldsymbol{X}_{2j} + \cdots + \sum_{j=1}^{m_s} c_{sj}\boldsymbol{X}_{sj} = \boldsymbol{0}$$

对于任意的 $1 \le i \le s$，如果 $\sum_{j=1}^{m_1} c_{ij}\boldsymbol{X}_{ij} \neq \boldsymbol{0}$，则 $\sum_{j=1}^{m_1} c_{ij}\boldsymbol{X}_{ij}$ 为矩阵 $\boldsymbol{A}$ 的对应于特征值 λ_i 的特征向量。根据本定理结论 2，这些“特征向量”线性无关，因此，如果

$$\sum_{j=1}^{m_i} c_{ij}X_{ij}, \qquad i = 1, 2, \cdots, s$$

中有一个不为零向量，这与结论矛盾，故

$$\sum_{j=1}^{m_i} c_{ij}X_{ij} = \boldsymbol{0}, \qquad i = 1, 2, \cdots, s$$

由于 $\boldsymbol{X}_{ij}, j = 1, 2, \cdots, m_i$ 为对应于特征值 λ_i 的线性无关的特征向量，因此得到

$$c_{ij} = 0 \quad j = 1, 2, \cdots, m_i; i = 1, 2, \cdots, s,$$

故命题得证。

(4) 设 n 阶方阵 $\boldsymbol{A}$ 的特征值 λ_0 有 r 个线性无关的特征向量 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_r$，再找 $n-r$ 个向量 $\boldsymbol{X}_{r+1}, \boldsymbol{X}_{r+2}, \cdots, \boldsymbol{X}_n$，使得向量组 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ 线性无关，因此 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ 为向量空

间 $\mathbb{R}^n$ 的一组基，且

$$\boldsymbol{A}\boldsymbol{X}_i = \lambda_0 \boldsymbol{X}_i \qquad i = 1, 2, \cdots, r$$

$$\boldsymbol{A}\boldsymbol{X}_i = c_{i1}\boldsymbol{X}_1 + c_{i2}\boldsymbol{X}_2 + \cdots + c_{in}\boldsymbol{X}_n \qquad i = r+1, \cdots, n$$

因此

$$\boldsymbol{A}\begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_r & \boldsymbol{X}_{r+1} & \cdots & \boldsymbol{X}_n\end{bmatrix}$$

$$=\begin{bmatrix}\boldsymbol{A}\boldsymbol{X}_1 & \boldsymbol{A}\boldsymbol{X}_2 & \cdots & \boldsymbol{A}\boldsymbol{X}_r & \boldsymbol{A}\boldsymbol{X}_{r+1} & \cdots & \boldsymbol{A}\boldsymbol{X}_n\end{bmatrix}$$

$$=\begin{bmatrix}\lambda_0\boldsymbol{X}_1 & \lambda_0\boldsymbol{X}_2 & \cdots & \lambda_0\boldsymbol{X}_r & \sum_{j=1}^{n} c_{r+1,j}\boldsymbol{X}_j & \cdots & \sum_{j=1}^{n} c_{nj}\boldsymbol{X}_j\end{bmatrix}$$

$$=\begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_r & \boldsymbol{X}_{r+1} & \cdots & \boldsymbol{X}_n\end{bmatrix}\begin{bmatrix}\lambda_0 & & & c_{r+1,1} & \cdots & c_{n,1}\\ & \ddots & & \vdots & \vdots & \vdots\\ & & \lambda_0 & c_{r+1,r} & \cdots & c_{n,r}\\ & & & c_{r+1,r+1} & \cdots & c_{n,r+1}\\ & & & \vdots & \vdots & \vdots\\ & & & c_{r+1,n} & \cdots & c_{n,n}\end{bmatrix}$$

令 $\boldsymbol{P} = \begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_r & \boldsymbol{X}_{r+1} & \cdots & \boldsymbol{X}_n\end{bmatrix}$

由于 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ 线性无关，所以 $\boldsymbol{P}$ 可逆，上式可写成

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\begin{bmatrix}\lambda_0 & & & c_{r+1,1} & \cdots & c_{n,1}\\ & \ddots & & \vdots & \vdots & \vdots\\ & & \lambda_0 & c_{r+1,r} & \cdots & c_{n,r}\\ & & & c_{r+1,r+1} & \cdots & c_{n,r+1}\\ & & & \vdots & \vdots & \vdots\\ & & & c_{r+1,n} & \cdots & c_{n,n}\end{bmatrix}$$

$$\boldsymbol{A} = \boldsymbol{P}\begin{bmatrix}\lambda_0 & & & c_{r+1,1} & \cdots & c_{n,1}\\ & \ddots & & \vdots & \vdots & \vdots\\ & & \lambda_0 & c_{r+1,r} & \cdots & c_{n,r}\\ & & & c_{r+1,r+1} & \cdots & c_{n,r+1}\\ & & & \vdots & \vdots & \vdots\\ & & & c_{r+1,n} & \cdots & c_{n,n}\end{bmatrix}\boldsymbol{P}^{-1} = \boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}$$

由于

$$\left|\lambda\boldsymbol{E}-\boldsymbol{A}\right|=\left|\lambda\boldsymbol{E}-\boldsymbol{PBP}^{-1}\right|=\left|\lambda\boldsymbol{PP}^{-1}-\boldsymbol{PBP}^{-1}\right|=\left|\boldsymbol{P}\right|\left|\lambda\boldsymbol{E}-\boldsymbol{B}\right|\left|\boldsymbol{P}^{-1}\right|$$
$$=\left|\boldsymbol{PP}^{-1}\right|\left|\lambda\boldsymbol{E}-\boldsymbol{B}\right|=\left|\lambda\boldsymbol{E}-\boldsymbol{B}\right|$$

所以方阵 $\boldsymbol{A},\boldsymbol{B}$ 的特征多项式相同，因此有相同的特征值，而显然 λ_0 至少为方阵 $\boldsymbol{B}$ 的 r 重特征值，所以命题成立。

由结论（3）可知：设 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 是方阵 $\boldsymbol{A}$ 的不同特征值， $V_{\lambda_1},V_{\lambda_2},\cdots,V_{\lambda_m}$ 分别为 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 特征子空间，则这 m 个子空间的和是直和，即有

$$\dim\left(V_{\lambda_1}+V_{\lambda_2}+\cdots+V_{\lambda_m}\right)=\sum_{i=1}^{m}\dim V_{\lambda_i}$$

习题 6.1

1. 设 $\boldsymbol{A}$ 是 n 阶实对称矩阵， $\boldsymbol{P}$ 是 n 阶可逆矩阵，已知 n 维列向量 $\boldsymbol{\alpha}$ 是 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量，则矩阵 $(\boldsymbol{P}^{-1}\boldsymbol{AP})^{\mathrm{T}}$ 属于特征值 λ 的特征向量是（　　）.

　（A） $\boldsymbol{P}^{-1}\boldsymbol{\alpha}$　　（B） $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\alpha}$　　（C） $\boldsymbol{P\alpha}$　　（D） $(\boldsymbol{P}^{-1})^{\mathrm{T}}\boldsymbol{\alpha}$

2. 设 λ_1,λ_2 是三阶矩阵 $\boldsymbol{A}$ 的两个不同的特征值， $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是 $\boldsymbol{A}$ 的属于 λ_1 的线性无关的特征向量， $\boldsymbol{\alpha}_3$ 是 $\boldsymbol{A}$ 的属于 λ_2 的特征向量，则 $\boldsymbol{\alpha}_1+\boldsymbol{A\alpha}_3,\boldsymbol{A}\left(\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3\right),\boldsymbol{A\alpha}_1+\boldsymbol{\alpha}_3$ 线性相关的充要条件是

　（A） $\lambda_1=0$ 或 $\lambda_1\lambda_2=1$　　（B） $\lambda_2=0$ 或 $\lambda_1\lambda_2=1$

　（C） $\lambda_1\neq0$ 或 $\lambda_1\lambda_2\neq1$　　（D） $\lambda_2\neq0$ 或 $\lambda_1\lambda_2\neq1$

3. n 阶方阵 $\boldsymbol{A}$ 的 n 个特征值中恰有 1 个为 0，则 $r\left(\boldsymbol{A}\right)$ 为（　　）

　(A)　n　　(B) $n-1$　　(C)　1　　(D)　0.

4. 计算下列矩阵的特征值与特征向量。

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

5. 已知 $\boldsymbol{A}$ 为三阶方阵，且 $\dfrac{1}{2}\boldsymbol{E}-\boldsymbol{A},,\dfrac{1}{2}\boldsymbol{E}+\boldsymbol{A},\boldsymbol{E}+\boldsymbol{A}$ 都不可逆，求 $|\boldsymbol{A}|$。

6. 设四阶方阵 $\boldsymbol{A}$ 满足 $\left|\boldsymbol{E}+3\boldsymbol{A}\right|=0,\boldsymbol{AA}^{\mathrm{T}}=2\boldsymbol{E},\left|\boldsymbol{A}\right|<0$，求 $\boldsymbol{A}^*$ 的一个特征值。

7. 设矩阵 $\boldsymbol{A}=\begin{bmatrix}2&1&1\\1&2&1\\1&1&a\end{bmatrix}$ 可逆，向量 $\boldsymbol{\alpha}=\begin{bmatrix}1\\b\\1\end{bmatrix}$ 是矩阵 $\boldsymbol{A}^*$ 的一个特征向量，λ 是 $\boldsymbol{\alpha}$ 对应的特征值，其中 $\boldsymbol{A}^*$ 是 $\boldsymbol{A}$ 的伴随矩阵，试求 a,b 和 λ 的值.

8. 设矩阵 $\boldsymbol{A}=\begin{bmatrix}a&-1&c\\5&b&3\\1-c&0&-a\end{bmatrix}$，$|\boldsymbol{A}|=-1$，又 $\boldsymbol{A}$ 的伴随矩阵 $\boldsymbol{A}^*$ 有一个特征值 λ_0，属于 λ_0 的特征向量为 $\alpha=[-1,-1,1]^{\mathrm{T}}$，求 a,b,c,λ_0 的值。

9. $\boldsymbol{A},\boldsymbol{B}$ 为四阶方阵，$\boldsymbol{AB}+2\boldsymbol{B}=\boldsymbol{O}$，$r(\boldsymbol{B})=2$，$|\boldsymbol{E}+\boldsymbol{A}|=|2\boldsymbol{E}-\boldsymbol{A}|=0$。(1) 求矩阵 $\boldsymbol{A}$ 的特征值；(2) 求 $|\boldsymbol{A}+3\boldsymbol{E}|$。

10. 已知 3 阶矩阵 $\boldsymbol{A}$ 满足 $|\boldsymbol{A}-\boldsymbol{E}|=|\boldsymbol{A}-2\boldsymbol{E}|=|\boldsymbol{A}+\boldsymbol{E}|=\lambda$，

(1) 当 $\lambda=0$ 时，求行列式 $|\boldsymbol{A}+3\boldsymbol{E}|$；

(2) 当 $\lambda=2$ 时，求行列式 $|\boldsymbol{A}+3\boldsymbol{E}|$。

11. 已知 $\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_3$ 是三阶非零矩阵，并且 $\boldsymbol{A}_i^2=\boldsymbol{A}_i\,(i=1,2,3),\boldsymbol{A}_i\boldsymbol{A}_j=\boldsymbol{O}\,(i\neq j)$，则 $\boldsymbol{A}_i\,(i=1,2,3)$ 的特征值为 0 与 1.

12. 设 $\boldsymbol{A},\boldsymbol{B}$ 为 n 阶方阵，满足 $\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}$，证明：

(1) 证明 $\boldsymbol{A},\boldsymbol{B}$ 的特征值不为 1；

(2) 设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值，求 $\boldsymbol{B}$ 的特征值。

13. 设向量 $\boldsymbol{\alpha}=[a_1,a_2,\cdots,a_n]^{\mathrm{T}},\boldsymbol{\beta}=[b_1,b_2,\cdots,b_n]^{\mathrm{T}}$ 都是非零向量，且满足条件 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}=0$。记 n 阶矩阵 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}$，求：

(1) $\boldsymbol{A}^2$；　　　　(2) 矩阵 $\boldsymbol{A}$ 的特征值和特征向量.

14. 设 n 阶方阵 $\boldsymbol{A}=[a_{ij}]$，若 $\sum_{j=1}^{n}|a_{ij}|<1,i=1,2,\cdots,n$，则 $\boldsymbol{A}$ 的所有特征值 $\lambda_i\,(i=1,2,\cdots,n)$ 的模小于 1，即 $|\lambda_i|<1$。

15. 设 $\boldsymbol{A},\boldsymbol{B}$ 为 n 阶方阵，满足 $\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}$，证明：

（1）证明 $\boldsymbol{A},\boldsymbol{B}$ 的特征值不为 1；

（2）设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值，求 $\boldsymbol{B}$ 的特征值。

16. 已知 $\boldsymbol{A}$ 是反对称矩阵，证明：$\boldsymbol{E}-\boldsymbol{A}^2$ 可逆。

§6.2　矩阵相似对角化

定义 6.4　相似矩阵

对于 n 阶方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，若有可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$$

则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **相似**，或 $\boldsymbol{A}$ **相似于** $\boldsymbol{B}$，记做 $\boldsymbol{A}\sim\boldsymbol{B}$，可逆矩阵 $\boldsymbol{P}$ 称为**相似变换矩阵**。

矩阵相似显然满足下述三条性质：

（1）　反身性：$\boldsymbol{A}\sim\boldsymbol{A}$；

（2）　对称性：若 $\boldsymbol{A}\sim\boldsymbol{B}$，则 $\boldsymbol{B}\sim\boldsymbol{A}$；

（3）　传递性：若 $\boldsymbol{A}\sim\boldsymbol{B}$，$\boldsymbol{B}\sim\boldsymbol{C}$，则 $\boldsymbol{A}\sim\boldsymbol{C}$。

因此矩阵相似也是一种等价关系。

定理 6.5　设 n 阶方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，则有以下结论：

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

(2) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式相同，进而特征值相同；

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=|\lambda\boldsymbol{E}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}|=|\lambda\boldsymbol{P}\boldsymbol{P}^{-1}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}|=|\boldsymbol{P}||\lambda\boldsymbol{E}-\boldsymbol{B}||\boldsymbol{P}^{-1}|$$
$$=|\boldsymbol{P}\boldsymbol{P}^{-1}||\lambda\boldsymbol{E}-\boldsymbol{B}|=|\lambda\boldsymbol{E}-\boldsymbol{B}|$$

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

(4) $tr(\boldsymbol{A})=tr(\boldsymbol{B})$；

(5) 若 $f(x)=a_kx^k+\cdots+a_1x+a_0$ 为 x 的多项式，则 $f(\boldsymbol{A})$ 与 $f(\boldsymbol{B})$ 相似。

相似矩阵的一种特殊情形就是相似于对角矩阵，即：对于方阵 $\boldsymbol{A}$，能否找到可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}=\boldsymbol{\Lambda}$$

这也是我们在这一章将要重点研究的问题。

定义 6.5　　**相似对角化**

如果n阶方$\boldsymbol{A}$与对角矩阵相似，则称$\boldsymbol{A}$相似于对角矩阵，或称$\boldsymbol{A}$**可（相似）对角化**。

该问题的讨论可如下进行：先假定方阵$\boldsymbol{A}$可相似对角化，我们分析矩阵$\boldsymbol{P}$、对角矩阵$\boldsymbol{\Lambda}$需要满足哪些性质，再据此展开进一步的讨论。

若$\boldsymbol{A}$可相似对角化，则存在可逆矩阵$\boldsymbol{P}$与对角矩阵$\boldsymbol{\Lambda}$，使得$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}$。将$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}$改写成

$$\boldsymbol{AP}=\boldsymbol{P\Lambda} \tag{6.12}$$

将可逆矩阵$\boldsymbol{P}$按列分块得

$$\boldsymbol{P}=\begin{bmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_n\end{bmatrix}$$

则(6.12)式变为

$$\boldsymbol{A}\begin{bmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}$$

$$\begin{bmatrix}\boldsymbol{AP}_1 & \boldsymbol{AP}_2 & \cdots & \boldsymbol{AP}_n\end{bmatrix}=\begin{bmatrix}\lambda_1\boldsymbol{P}_1 & \lambda_2\boldsymbol{P}_2 & \cdots & \lambda_n\boldsymbol{P}_n\end{bmatrix}$$

这说明：矩阵$\boldsymbol{A},\boldsymbol{P},\boldsymbol{\Lambda}$需满足

$$\boldsymbol{AP}_i=\lambda_i\boldsymbol{P}_i\ ,\qquad i=1,2,\cdots,n \tag{6.13}$$

由于$\boldsymbol{P}$可逆，其列向量组线性无关，所以$\boldsymbol{P}_i\neq\boldsymbol{0}$，故$\lambda_i$是$\boldsymbol{A}$的特征值，$\boldsymbol{P}_i$为$\boldsymbol{A}$的对应于$\lambda_i$的特征值。因此，若$\boldsymbol{A}$可相似对角化，则$\boldsymbol{A}$有$n$个线性无关的特征向量$\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_n$。反之，若$\boldsymbol{A}$有$n$个线性无关的特征向量$\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_n$，因此有

$$\boldsymbol{AP}_i=\lambda_i\boldsymbol{P}_i\ ,\qquad i=1,2,\cdots,n$$

得到

$$\begin{bmatrix}\boldsymbol{AP}_1 & \boldsymbol{AP}_2 & \cdots & \boldsymbol{AP}_n\end{bmatrix}=\begin{bmatrix}\lambda_1\boldsymbol{P}_1 & \lambda_2\boldsymbol{P}_2 & \cdots & \lambda_n\boldsymbol{P}_n\end{bmatrix}$$

即得

$$\boldsymbol{A}\begin{bmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}$$

令$\boldsymbol{P}=\begin{bmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_n\end{bmatrix}$，$\boldsymbol{\Lambda}=diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$，因此有$\boldsymbol{AP}=\boldsymbol{P\Lambda}$。由于$\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_n$线

性无关，因此方阵 $\boldsymbol{P}$ 可逆，因此有 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}$，故 $\boldsymbol{A}$ 相似于对角矩阵 $\boldsymbol{\Lambda}$。

■

根据上述分析，若 $\boldsymbol{A}$ 可对角化，则存在可逆矩阵 $\boldsymbol{P}$、对角矩阵 $\boldsymbol{\Lambda}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}$，其中对角矩阵 $\boldsymbol{\Lambda}=diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$ 主对角线上的元素 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值，$\boldsymbol{P}=\begin{bmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_n\end{bmatrix}$，$\boldsymbol{P}_i$ 为 $\boldsymbol{A}$ 的对应于 λ_i 的特征值。即：如果方阵 $\boldsymbol{A}$ 相似于对角矩阵 $\boldsymbol{\Lambda}$，则对角矩阵 $\boldsymbol{\Lambda}$ 的对角线上的元素为方阵的 n 个特征值，$\boldsymbol{P}$ 的第 $\boldsymbol{i}$ 个列向量为方阵 $\boldsymbol{A}$ 的对应于特征值 λ_i 的特征向量。

从如上分析得到如下定理：

定理 6.5　n 阶方阵 $\boldsymbol{A}$ 相似于对角矩阵的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量。

推论 6.2　若 n 阶方阵 $\boldsymbol{A}$ 有 n 个互不相同的特征值，则 $\boldsymbol{A}$ 必相似于对角矩阵。

推论 6.3　n 阶方阵 $\boldsymbol{A}$ 相似于对角矩阵的充要条件是，所有特征值的代数重数等于几何重数。

推论 6.4　n 阶方阵 $\boldsymbol{A}$ 相似于对角矩阵 $\Leftrightarrow$ 对于 $\boldsymbol{A}$ 的代数重数为 t_i 的特征值 λ_i，有 $r\left(\lambda_i\boldsymbol{E}-\boldsymbol{A}\right)=n-t_i$。

由推论 6.3 可知，例 6.5 中的矩阵 $\boldsymbol{A}$ 不能相似于对角矩阵，而矩阵 $\boldsymbol{B}$ 则相似于对角矩阵，且可令 $\boldsymbol{\Lambda}=diag\left(-4,4,4\right),\boldsymbol{P}=\begin{bmatrix}-1 & 4 & -1\\ -2 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}$，则 $\boldsymbol{B}=\boldsymbol{P\Lambda P}^{-1}$。

例 6.5　设 $\boldsymbol{A}=\begin{bmatrix}2 & -1\\ -1 & 2\end{bmatrix}$，求 $\boldsymbol{A}^n$。

解：如果 $\boldsymbol{A}$ 可对角化，即存在可逆矩阵 $\boldsymbol{P}$、对角矩阵 $\boldsymbol{\Lambda}$，使得 $\boldsymbol{A}=\boldsymbol{P\Lambda P}^{-1}$，根据第二章的结论可知 $\boldsymbol{A}^n=\boldsymbol{P\Lambda}^n\boldsymbol{P}^{-1}$。

由于对角矩阵 $\boldsymbol{\Lambda}$ 的 n 次幂容易计算，因此 $\boldsymbol{A}^n$ 容易计算。

首先看 $\boldsymbol{A}$ 是否可对角化。

由$\left|\lambda\boldsymbol{E}-\boldsymbol{A}\right|=\begin{vmatrix}\lambda-2 & 1\\ 1 & \lambda-2\end{vmatrix}=(\lambda-1)(\lambda-3)$，故$\boldsymbol{A}$的特征值为$\lambda_1=1,\lambda_2=3$，二阶方阵$\boldsymbol{A}$有两个不同的特征值，故$\boldsymbol{A}$可对角化。$\lambda_1=1$时，$\lambda_1\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}-1 & 1\\ 1 & -1\end{bmatrix}\overset{r}{\sim}\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}$，故一个特征向量为$[1,1]^{\mathrm{T}}$；$\lambda_2=3$时，$\lambda_2\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\overset{r}{\sim}\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$，故一个特征向量为$[1,-1]^{\mathrm{T}}$。因此可令

$$\boldsymbol{P}=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix},\qquad \boldsymbol{\Lambda}=\begin{bmatrix}1 & \\ & 3\end{bmatrix}$$

有$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$。故

$$\boldsymbol{A}^n=\boldsymbol{P}\boldsymbol{\Lambda}^n\boldsymbol{P}^{-1}=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}1 & \\ & 3\end{bmatrix}^n\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}^{-1}=\frac{1}{2}\begin{bmatrix}1+3^n & 1-3^n\\ 1-3^n & 1+3^n\end{bmatrix}$$

例 6.6　设三阶方阵$\boldsymbol{A}$的特征值为$\lambda_1=1,\lambda_2=2,\lambda_3=3$，对应的特征向量分别为$\boldsymbol{\alpha}_1=\begin{bmatrix}1\\1\\1\end{bmatrix},\boldsymbol{\alpha}_2=\begin{bmatrix}1\\2\\4\end{bmatrix},\boldsymbol{\alpha}_3=\begin{bmatrix}1\\3\\9\end{bmatrix}$又设$\boldsymbol{\beta}=\begin{bmatrix}1\\1\\3\end{bmatrix}$。

(1)将$\boldsymbol{\beta}$用$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表出；

(2)求$\boldsymbol{A}^n\boldsymbol{\beta}$（$n$为正整数）。

解（1）首先将$\boldsymbol{\beta}$用$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表出，即求解以

$$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3 & \boldsymbol{\beta}\end{bmatrix}=\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 2 & 3 & 1\\ 1 & 4 & 9 & 3\end{bmatrix}$$

为增广矩阵的线性方程组，用初等行变换将之化为行最简形矩阵

$$\to\begin{bmatrix}1 & 0 & 0 & 2\\ 0 & 1 & 0 & -2\\ 0 & 0 & 1 & 1\end{bmatrix}$$

故它的解为$[2,-2,1]^{\mathrm{T}}$，因此有$\boldsymbol{\beta}=2\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$；

（2）为求$\boldsymbol{A}^n\boldsymbol{\beta}$，我们可先求$\boldsymbol{A}^n$。由于三阶方阵$\boldsymbol{A}$有 3 个互不相同的特征值

$\lambda_1=1,\lambda_2=2,\lambda_3=3$，因此 $\boldsymbol{A}$ 可相似对角化，且根据上述理论有

$$\boldsymbol{A}=\begin{bmatrix}1&1&1\\1&2&3\\1&4&9\end{bmatrix}\begin{bmatrix}1&&\\&2&\\&&3\end{bmatrix}\begin{bmatrix}1&1&1\\1&2&3\\1&4&9\end{bmatrix}^{-1}=\begin{bmatrix}0&1&0\\0&0&1\\6&-11&6\end{bmatrix}$$

因此

$$\boldsymbol{A}^n=\begin{bmatrix}1&1&1\\1&2&3\\1&4&9\end{bmatrix}\begin{bmatrix}1&&\\&2&\\&&3\end{bmatrix}^n\begin{bmatrix}1&1&1\\1&2&3\\1&4&9\end{bmatrix}^{-1}=\begin{bmatrix}1&1&1\\1&2&3\\1&4&9\end{bmatrix}\begin{bmatrix}1&&\\&2^n&\\&&3^n\end{bmatrix}\begin{bmatrix}1&1&1\\1&2&3\\1&4&9\end{bmatrix}^{-1}$$

$$=\begin{bmatrix}3-3\times2^n+3^n & -\dfrac{5}{2}+2^{n+2}-\dfrac{3^{n+1}}{2} & \dfrac{1}{2}-2^n+\dfrac{3^n}{2}\\ 3-3\times2^{n+1}+3^{n+1} & -\dfrac{5}{2}+2^{n+3}-\dfrac{3^{n+2}}{2} & \dfrac{1}{2}-2^{n+1}+\dfrac{3^{n+1}}{2}\\ 3-3\times2^{n+2}+3^{n+2} & -\dfrac{5}{2}+2^{n+4}-\dfrac{3^{n+3}}{2} & \dfrac{1}{2}-2^{n+2}+\dfrac{3^{n+2}}{2}\end{bmatrix}$$

容易求得

$$\boldsymbol{A}^n\boldsymbol{\beta}=\boldsymbol{A}^n\begin{bmatrix}1\\1\\3\end{bmatrix}=\begin{bmatrix}2-2^{n+1}+3^n\\2-2^{n+2}+3^{n+1}\\2-2^{n+3}+3^{n+2}\end{bmatrix}$$

我们也可以如此求 $\boldsymbol{A}^n\boldsymbol{\beta}$：

$$\boldsymbol{A}^n\boldsymbol{\beta}=\boldsymbol{A}^n\left(2\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3\right)$$

$$=2\boldsymbol{A}^n\boldsymbol{\alpha}_1-2\boldsymbol{A}^n\boldsymbol{\alpha}_2+\boldsymbol{A}^n\boldsymbol{\alpha}_3$$
$$=2\times\lambda_1^n\boldsymbol{\alpha}_1-2\lambda_2^n\boldsymbol{\alpha}_2+\lambda_3^n\boldsymbol{\alpha}_3$$
$$=2\boldsymbol{\alpha}_1-2\times2^n\boldsymbol{\alpha}_2+3^n\boldsymbol{\alpha}_3$$
$$=\begin{bmatrix}2-2^{n+1}+3^n\\2-2^{n+2}+3^{n+1}\\2-2^{n+3}+3^{n+2}\end{bmatrix}$$

例 6.7　$\boldsymbol{A}$ 有三个线性无关的特征向量，求 x,y 满足的条件。

$$\boldsymbol{A}=\begin{bmatrix}0&0&1\\x&1&y\\1&0&0\end{bmatrix}$$

解：先求 $\boldsymbol{A}$ 的特征值

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda & 0 & -1\\ -x & \lambda-1 & -y\\ -1 & 0 & \lambda\end{vmatrix}=(\lambda-1)\begin{vmatrix}\lambda & -1\\ -1 & \lambda\end{vmatrix}$$

$$=(\lambda-1)(\lambda^2-1)=(\lambda-1)^2(\lambda+1)$$

$\boldsymbol{A}$ 的特征值为 $\lambda_1=\lambda_2=1,\lambda_3=-1$，$\boldsymbol{A}$ 有三个线性无关的特征向量⇔二重特征值 1 有两个线性无关的特征向量⇔

$$r(1\,\boldsymbol{E}-\boldsymbol{A})=3-2=1$$

$$\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}1 & 0 & -1\\ -x & 0 & -y\\ -1 & 0 & 1\end{bmatrix}$$

$$r(\boldsymbol{E}-\boldsymbol{A})=1\Leftrightarrow x+y=0$$

例 6.8　解微分方程组

$$\begin{cases}\dfrac{dx_1(t)}{dt}=x_1+2x_2+2x_3\\ \dfrac{dx_2(t)}{dt}=2x_1+x_2+2x_3\\ \dfrac{dx_3(t)}{dt}=2x_1+2x_2+x_3\end{cases}$$

解　记 $\boldsymbol{X}(t)=[x_1(t),x_2(t),x_3(t)]^{\mathrm{T}}=\boldsymbol{X},\ \boldsymbol{A}=\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{bmatrix}$，规定

$$\frac{d\boldsymbol{X}(t)}{dt}=\left[\frac{dx_1(t)}{dt},\frac{dx_2(t)}{dt},\frac{dx_3(t)}{dt}\right]^{\mathrm{T}}$$

则题目中所给微分方程组可表示为

$$\frac{d\boldsymbol{X}(t)}{dt}=\boldsymbol{A}\boldsymbol{X}$$

通过计算可知，方阵 $\boldsymbol{A}$ 可相似于对角矩阵，存在对角矩阵 $\boldsymbol{\Lambda}=\begin{bmatrix}5 & & \\ & -1 & \\ & & -1\end{bmatrix}$ 与可逆矩阵 $\boldsymbol{P}=\begin{bmatrix}1 & -1 & -1\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}$，使得 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$。

因此$\frac{d\boldsymbol{X}(t)}{dt}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\boldsymbol{X}$，$\Rightarrow \boldsymbol{P}^{-1}\frac{d\boldsymbol{X}(t)}{dt}=\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\boldsymbol{X}$，由于$\boldsymbol{P}^{-1}$的每个元素均为常数，与$t$无关，故$\boldsymbol{P}^{-1}\frac{d\boldsymbol{X}(t)}{dt}=\frac{d\left(\boldsymbol{P}^{-1}\boldsymbol{X}(t)\right)}{dt}$，令$\boldsymbol{Y}(t)=\boldsymbol{P}^{-1}\boldsymbol{X}(t)$，则得到

$$\frac{d\boldsymbol{Y}(t)}{dt}=\boldsymbol{\Lambda}\boldsymbol{Y}=\begin{bmatrix}5 & & \\ & -1 & \\ & & -1\end{bmatrix}\begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}$$

也就是

$$\begin{cases}\dfrac{dy_1(t)}{dt}=5y_1\\ \dfrac{dy_2(t)}{dt}=-y_2\\ \dfrac{dy_3(t)}{dt}=-y_3\end{cases}$$

这是三个相互独立的微分方程，由微分方程基本理论可求得

$$\begin{cases}y_1=C_1e^{5t}\\ y_2=C_2e^{-t}\\ y_3=C_3e^{-t}\end{cases}$$

由$\boldsymbol{Y}(t)=\boldsymbol{P}^{-1}\boldsymbol{X}(t)$得到

$$\boldsymbol{X}(t)=\boldsymbol{P}\boldsymbol{Y}(t)=\begin{bmatrix}1 & -1 & -1\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}\begin{bmatrix}C_1e^{5t}\\ C_2e^{-t}\\ C_3e^{-t}\end{bmatrix}=\begin{bmatrix}C_1e^{5t}-C_2e^{-t}-C_3e^{-t}\\ C_1e^{5t}+C_2e^{-t}\\ C_1e^{5t}+C_3e^{-t}\end{bmatrix}。$$

习题 6.2

1. 设$\boldsymbol{A}=\begin{bmatrix}2 & a & 2\\ 5 & b & 3\\ -1 & 1 & -1\end{bmatrix}$有特征值$\lambda=\pm1$，问$\boldsymbol{A}$是否能相似对角化，试说明之。

2. 设矩阵$\boldsymbol{A}=\begin{bmatrix}a & 1 & 1\\ 1 & a & -1\\ 1 & -1 & a\end{bmatrix}$，求可逆矩阵$\boldsymbol{P}$，使$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$为对角矩阵，并计算行列式$|\boldsymbol{A}-\boldsymbol{E}|$的值。

3. 设 3 阶方阵$\boldsymbol{A}$的特征值为$\lambda_1=2,\lambda_2=-2,\lambda_3=1$，对应的特征向量依次为$\boldsymbol{P}_1=[0,1,1]^{\mathrm{T}},\boldsymbol{P}_2=[1,1,1]^{\mathrm{T}},\boldsymbol{P}_3=[1,1,0]^{\mathrm{T}}$，求$\boldsymbol{A}$。

4. 某试验生产线每年一月份进行熟练工与匪熟练工的人数统计，然后将 1/4 的熟练工支援其他生产部门，其缺额由招收新的非熟练工补齐。新、老非熟练工经过培训及实践至年终考核有 5/6 成为熟练工。设第 n 年一月份统计的熟练工和非熟练工所占百分比分别为 x_n 和 y_n，记成向量 $\begin{bmatrix} x_n \\ y_n \end{bmatrix}$。（1）求 $\begin{bmatrix} x_n \\ y_n \end{bmatrix}$ 与 $\begin{bmatrix} x_{n-1} \\ y_{n-1} \end{bmatrix}$ 的关系式，并写成矩阵形式：$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} x_{n-1} \\ y_{n-1} \end{bmatrix}$；（2）求 $\begin{bmatrix} x_n \\ y_n \end{bmatrix}$。

5. 求下列矩阵的 k 次幂。（1）$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \\ -1 & 1 & -1 \end{bmatrix}$（2）$\boldsymbol{B} = \begin{bmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 1 & -1 \end{bmatrix}$。

6. 试确定 x, y，使得 $\begin{bmatrix} 1 & -2 & -4 \\ -2 & x & -2 \\ -4 & -2 & 1 \end{bmatrix}$ 与 $\begin{bmatrix} 5 & & \\ & y & \\ & & -4 \end{bmatrix}$ 相似。

7. 设矩阵 $\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 1 \\ x & 4 & y \\ -3 & -3 & 5 \end{bmatrix}$，已知 $\boldsymbol{A}$ 有三个线性无关的特征向量，$\lambda = 2$ 是 $\boldsymbol{A}$ 的二重特征值，求可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为对角阵。

8. 设矩阵 $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -3 \\ -1 & 4 & -3 \\ 1 & a & 5 \end{bmatrix}$ 的特征方称有一个二重根，求 a 的值，并讨论 $\boldsymbol{A}$ 是否可相似于对角阵。

9. 设 n 阶方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2 - 5\boldsymbol{A} + 6\boldsymbol{E} = \boldsymbol{O}$，试证明矩阵 $\boldsymbol{A}$ 和对角矩阵相似。

10. 设 $\boldsymbol{A}, \boldsymbol{B}$ 都是 n 阶方阵，$\boldsymbol{A}$ 有 n 个不同的特征值。证明：$\boldsymbol{AB} = \boldsymbol{BA} \Leftrightarrow \boldsymbol{A}$ 的特征向量也是 $\boldsymbol{B}$ 的特征向量。

11. 设 $a_0, a_1, \cdots, a_{n-1} \in \mathbb{R}$，$\boldsymbol{B} = \left[b_{ij}\right]_{n\times n}$，其中

$$b_{ij} = \begin{cases} -a_{j-1}, & i = n \\ 1, & i = j-1, j = 2, 3, \cdots, n, \\ 0, & \text{其余} \end{cases}$$

（1）若 λ 为 $\boldsymbol{B}$ 的特征值，证明 $\left[1, \lambda, \lambda^2, \cdots, \lambda^{n-1}\right]^{\mathrm{T}}$ 为 $\boldsymbol{B}$ 的特征向量。

（2）若 $\boldsymbol{B}$ 有 n 个不同的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$，求一可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{AP} = diag\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right)$。

12. 设 $\boldsymbol{A}$ 为 n 阶上三角矩阵，（1）若 $a_{ii} \neq a_{jj}\left(i \neq j, i, j=1,2,\cdots,n\right)$，则 $\boldsymbol{A}$ 可相似对角化；

（2）若 $a_{11}=a_{22}=\cdots=a_{nn}$ 且至少有一个 $a_{i_0 j_0} \neq 0\left(i_0 \neq j_0\right)$，则 $\boldsymbol{A}$ 不可相似对角化。

13. 设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 为 $\mathbb{R}^n$ 中两个非零的正交向量，证明：矩阵 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}$ 的特征值全为零，且 $\boldsymbol{A}$ 不可对角化。

14. 设 $\boldsymbol{\alpha}$ 为 $\mathbb{R}^n$ 中的非零向量，证明：矩阵 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}$ 可对角化，并求可逆矩阵 $\boldsymbol{P}$ 与对角矩阵 $\boldsymbol{\Lambda}$，使得 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$。

15. 设 $\boldsymbol{A}$ 是 n 阶矩阵，$\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n$ 是 n 维列向量。若 $\boldsymbol{\xi}_n \neq \mathbf{0}$，且 $\boldsymbol{A}\boldsymbol{\xi}_1=\boldsymbol{\xi}_2, \boldsymbol{A}\boldsymbol{\xi}_2=\boldsymbol{\xi}_3,\cdots,\boldsymbol{A}\boldsymbol{\xi}_{n-1}=\boldsymbol{\xi}_n, \boldsymbol{A}\boldsymbol{\xi}_n=\mathbf{0}$，证明：

（1）$\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n$ 线性无关；

（2）$\boldsymbol{A}$ 不能相似于对角矩阵。

第七章　二次型

§7.1　二次型及其矩阵表示

定义 7.1　n 元二次型

称 n 个实变量 $x_1, x_2, \cdots, x_n$ 的二次齐次多项式

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) &= a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ &+ a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ &+ \cdots + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \end{aligned} \tag{7.1}$$

为 n 元实二次型，简称 n 元二次型。(注意：(7.1)中，并无 $a_{ij} = a_{ji}$ 的要求。)

上述二次型还可以写成

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) &= a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ &+ a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ &+ \cdots + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \end{aligned}$$

$$\begin{aligned} = a_{11}x_1^2 &+ (a_{12}+a_{21})x_1x_2 + (a_{13}+a_{31})x_1x_3 + \cdots + (a_{1n}+a_{n1})x_1x_n \\ &+ a_{22}x_2^2 + (a_{23}+a_{32})x_2x_3 + (a_{24}+a_{42})x_2x_4 + \cdots + (a_{2n}+a_{n2})x_2x_n \\ &+ \cdots \\ &+ a_{n-1,n-1}x_{n-1}^2 + (a_{n-1,n}+a_{n,n-1})x_{n-1}x_n \\ &+ a_{nn}x_n^2 \end{aligned}$$

例 7.1　以下都是二次型

$$f_1(x_1, x_2, x_3) = 3x_1^2 + x_2^2 - x_3^2 + 2x_1x_2 - 4x_2x_3,$$

$$f_2(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 - 2x_2x_3,$$

$$f_3(x_1, x_2, x_3) = x_1^2 + 3x_2^2 - 5x_3^2。$$

二次型的矩阵表示：

二次型可用矩阵乘法来表示，从而可将二次型与矩阵联系起来。令

$$\boldsymbol{X}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}},\qquad \boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

则

$$\begin{aligned} f(x_1,x_2,\cdots,x_n) &= a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n \\ &\quad +a_{21}x_2x_1+a_{22}x_2^2+\cdots+a_{2n}x_2x_n \\ &\quad +\cdots+a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2 \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \end{aligned}$$

$$\begin{aligned} &= x_1\left(a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\right) \\ &+x_2\left(a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n\right) \\ &+\cdots \\ &+x_n\left(a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n\right) \\ &= [x_1\ \ x_2\ \ \cdots\ \ x_n]\begin{bmatrix} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n \\ \vdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n \end{bmatrix} \\ &= [x_1\ \ x_2\ \ \cdots\ \ x_n]\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X} \end{aligned}$$

由于 $a_{ij}x_ix_j+a_{ji}x_jx_i$ 可合并为 $\left(a_{ij}+a_{ji}\right)x_ix_j$，如果令

$$b_{ij}=\begin{cases} a_{ii}, & i=j \\ \dfrac{a_{ij}+a_{ji}}{2}, & i\neq j \end{cases} \tag{7.2}$$

则矩阵 $\boldsymbol{B}=\left[b_{ij}\right]_{n\times n}$ 满足 $b_{ij}=b_{ji}$，因此 $\boldsymbol{B}=\boldsymbol{B}^{\mathrm{T}}$ 为对称矩阵。

$$f\left(x_1,x_2,\cdots,x_n\right)=[x_1,x_2,\cdots,x_n]\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$$

$$=\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \boldsymbol{X}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{X}$$

称对称矩阵 $\boldsymbol{B}$ 为二次型(7.1)所对应的矩阵,或二次型(7.1)的矩阵。

从而，二次型与对称矩阵建立了联系，因此能用矩阵理论研究二次型问题。

从上述讨论可知如何求二次型所对应的实对称矩阵？实际上根据(7.2)即可，即

$$b_{ij}=\begin{cases} x_i^2\text{系数}, & i=j \\ x_ix_j\text{系数的一半}, & i\neq j \end{cases}$$

显然，二次型的矩阵是唯一的。以后在将二次型表示为 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 的形式时，其中 $\boldsymbol{A}$ 为实对称矩阵。

例 7.1 中的三个二次型所对应的矩阵分别为

$f_1(x_1,x_2,x_3)=3x_1^2+x_2^2-x_3^2+2x_1x_2-4x_2x_3$ 的矩阵为 $\boldsymbol{A}_1=\begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & -2 \\ 0 & -2 & -1 \end{bmatrix}$；

$f_2(x_1,x_2,x_3)=2x_1x_2+2x_1x_3-2x_2x_3$ 的矩阵为 $\boldsymbol{A}_2=\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$；

$f_3(x_1,x_2,x_3)=x_1^2+3x_2^2-5x_3^2$ 的矩阵为 $\boldsymbol{A}_3=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{bmatrix}$。

在这三个二次型中，f_3 最简单，因为它只含有单个变量的平方项，而不含有不同变量的乘积(交叉乘积)。

定义 7.2　　二次型的秩

二次型矩阵的秩称为二次型的秩。

习题 7.1

1.　给出下述各二次型的矩阵，并求其秩。

a)　$f\left(x_1,x_2,x_3\right)=5x_1^2+5x_2^2+3x_3^2-2x_1x_2+6x_1x_3-6x_2x_3$；

b)　$f\left(x_1,x_2,x_3\right)=\left(x_1+x_2\right)^2+\left(x_2-x_3\right)^2+\left(x_3+x_1\right)^2$

c)　$f=3\sum_{i=1}^{n}x_i^2+\sum_{1\le i<j\le n}2x_ix_j$；

d)　设 $\boldsymbol{A}$ 为 n 阶实对称矩阵，$r(\boldsymbol{A})=n$，$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{A_{ij}}{|A|}x_ix_j$。

§7.2　二次型的标准形

定义 7.2　　二次型的**标准形，规范形**

如果二次型

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$$

经过非退化线性变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$，变成平方和

$$\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{Y}=d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2$$

称之为 $f(x_1,x_2,\cdots,x_n)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$ 的**标准形或法式**；而经过非退化线性变换

将 $f(x_1,x_2,\cdots,x_n)$ 化为

$$z_1^2+z_2^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_{p+q}^2,\quad p+q\le n$$

称之为二次型 $f(x_1,x_2,\cdots,x_n)$ 的**规范形**。

身为标准形、规范形的二次型的矩阵分别为

$$\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}$$

$$\begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{bmatrix}$$

关于二次型，本章只讲两个问题。第一个问题：能否将一个二次型“化为”标准形？如果能，又如何做到这一点？准确地说，对于二次型**错误！未找到引用源。**，是否存在可逆的线性变换

$$\begin{cases} x_1 = p_{11}y_1 + p_{12}y_2 + \cdots + p_{1n}y_n \\ x_2 = p_{21}y_1 + p_{22}y_2 + \cdots + p_{2n}y_n \\ \quad \vdots \\ x_n = p_{n1}y_1 + p_{n2}y_2 + \cdots + p_{nn}y_n \end{cases}$$

即 $\boldsymbol{X} = \boldsymbol{PY}$，其中

$$\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \qquad \boldsymbol{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

其中 $\boldsymbol{P}$ 为可逆矩阵，使得

$$f\left(x_1, x_2, \cdots, x_n\right) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$$

$$\underline{\underline{\quad \boldsymbol{X} = \boldsymbol{PY} \quad}} \left(\boldsymbol{PY}\right)^{\mathrm{T}} \boldsymbol{A}\left(\boldsymbol{PY}\right) = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{APY} = \boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}\right)\boldsymbol{Y}$$

为标准形？

由于标准形二次型的矩阵为对角矩阵，并且二次型 $f\left(x_1, x_2, \cdots, x_n\right) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 的矩阵 $\boldsymbol{A}$ 为实对称矩阵，因此这个问题可用矩阵描述为：对于一个实对称矩阵 $\boldsymbol{A}_{n\times n}$，是否存在一个可逆矩阵 $\boldsymbol{P} = \left[p_{ij}\right]_{n\times n}$，使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} \tag{7.3}$$

为对角矩阵。

定义 7.3　　矩阵的**合同**

设 $\boldsymbol{A}, \boldsymbol{B}$ 为 n 阶方阵，若存在 n 阶可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \boldsymbol{B} \tag{7.4}$$

则称 $\boldsymbol{A}$ 合同于 $\boldsymbol{B}$，记做 $\boldsymbol{A} \simeq \boldsymbol{B}$。

矩阵的合同也有下面三个性质(可与矩阵相似、向量组的等价相比较)

(1) 自反性：$\boldsymbol{A} \simeq \boldsymbol{A}$；

(2) 对称性：若 $\boldsymbol{A} \simeq \boldsymbol{B}$，则 $\boldsymbol{B} \simeq \boldsymbol{A}$；

(3) 传递性：若 $\boldsymbol{A} \simeq \boldsymbol{B}$，$\boldsymbol{B} \simeq \boldsymbol{C}$，则 $\boldsymbol{A} \simeq \boldsymbol{C}$。

有了矩阵合同的概念，二次型的标准化问题可叙述为：对一个实对称矩阵 $\boldsymbol{A}$ ，求一个可逆矩阵 $\boldsymbol{P}$ ，使得 $\boldsymbol{A}$ 合同于一个对角矩阵，即

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$$

定理 7.1（标准形的存在性定理）　任何一个实对称矩阵 $\boldsymbol{A}$ 都合同于对角矩阵，即存在可逆矩阵 $\boldsymbol{P}$ ，使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$$

因此，任何一个二次型都可用可逆线性变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$ 化为标准型：

$$f(x_1,x_2,\cdots,x_n)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$$

$$\xlongequal{\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}}\boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\right)\boldsymbol{Y}=d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2$$

证明

有三种方法将二次型化为标准形：

(1)拉格朗日(Lagrange)配方法

(2)初等变换合同法

(3)正交变换法

例 7.2　用拉格朗日(Lagrange)配方法将二次型化为标准形

$$f(x_1,x_2,x_3)=x_1^2+2x_2^2+5x_3^2+2x_1x_2+2x_1x_3+6x_2x_3$$

$$=\left(x_1^2+2x_1x_2+2x_1x_3\right)+2x_2^2+5x_3^2+6x_2x_3$$

$$=\left(\left(x_1^2+x_2+x_3\right)^2-2x_2x_3-x_2^2-x_3^2\right)+2x_2^2+5x_3^2+6x_2x_3$$

$$=\left(x_1^2+x_2+x_3\right)^2+x_2^2+4x_3^2+4x_2x_3$$

$$=\left(x_1^2+x_2+x_3\right)^2+(x_2^2+4x_2x_3+4x_3^2)$$

$$=\left(x_1+x_2+x_3\right)^2+(x_2+2x_3)^2$$

若令

$$\begin{cases} y_1=x_1+x_2+x_3 \\ y_2=x_2+2x_3 \end{cases}$$

则

$$f(x_1,x_2,x_3)=y_1^2+y_2^2$$

由于我们要找一个可逆线性变换，该线性变换的矩阵必须是可逆矩阵，因此该矩阵首先

必须是方阵，因此还需要加上一个做变换的式子：

$$\begin{cases} y_1 = x_1 + x_2 + x_3 \\ y_2 = x_2 + 2x_3 \\ y_3 = x_3 \end{cases}$$

显然，该线性变换的矩阵为

$$\boldsymbol{P} = \begin{bmatrix} 1 & 1 & 1 \\ & 1 & 2 \\ & & 1 \end{bmatrix}$$

为可逆矩阵。所用的线性变换为

$$\boldsymbol{Y} = \boldsymbol{PX}，\quad 亦即\ \boldsymbol{X} = \boldsymbol{P}^{-1}\boldsymbol{Y}$$

注意：构成可逆变换的第三个式子 $y_3 = x_3$ （即前一矩阵的第三行）的取法是不唯一的，可随意取值，只要所得的矩阵 $\boldsymbol{P}$ 可逆即可。

如果二次型不含平方项，则首先“构造”出平方项。

例 7.3　用配方法将二次型化为标准形

$$f(x_1, x_2, x_3) = x_1x_2 + 4x_1x_3 + x_2x_3$$

解：首先令

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \end{cases}$$

则有

$$\begin{aligned} f &= (y_1 + y_2)(y_1 - y_2) + 4(y_1 + y_2)x_3 + (y_1 - y_2)x_3 \\ &= y_1^2 - y_2^2 + 4y_1x_3 + 4y_2x_3 + y_1x_3 - y_2x_3 \\ &= y_1^2 - y_2^2 + 5y_1x_3 + 3y_2x_3 \end{aligned}$$

这里含有两种变量 x, y，为方便表示起见，令 $x_3 = y_3$，因而相当于做可逆线性变换

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \end{cases} \quad 即 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

二次型化为

$$f = y_1^2 - y_2^2 + 5y_1y_3 + 3y_2y_3$$

此时已经出现平方项，可以运用例 7.2 中的方法配方，

$$f = (y_1^2 - y_2^2 + 5y_1y_3) + 3y_2y_3$$

$$=\left(y_1+\frac{5}{2}y_3\right)^2-\frac{25}{4}y_3^2-y_2^2+3y_2y_3$$

$$=\left(y_1+\frac{5}{2}y_3\right)^2-\left(y_2^2-3y_2y_3+\frac{9}{4}y_3^2\right)-\frac{16}{4}y_3^2$$

$$=\left(y_1+\frac{5}{2}y_3\right)^2-\left(y_2-\frac{3}{2}y_3\right)^2-4y_3^2$$

因此所用的线性变换为

$$\begin{cases} z_1=y_1+\frac{5}{2}y_3 \\ z_2=y_2-\frac{3}{2}y_3 \\ z_3=y_3 \end{cases}$$

其逆变换为

$$\begin{cases} y_1=z_1-\frac{5}{2}z_3 \\ y_2=z_2+\frac{3}{2}z_3 \\ y_3=z_3 \end{cases}$$

二次型的一个标准形为

$$f=z_1^2-z_2^2-4z_3^2$$

由 x_1,x_2,x_3 到 z_1,z_2,z_3 的变换矩阵为：

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}=\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -\frac{5}{2} \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}=\begin{bmatrix} 1 & 1 & -\frac{5}{2} \\ 1 & -1 & \frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

初等变换合同法

将二次型 $f=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 化为标准型，相当于找一个可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$$

为对角矩阵。因为 $\boldsymbol{P}$ 可逆，因此 $\boldsymbol{P}$ 可表示为一序列初等矩阵得乘积：$\boldsymbol{P}=\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m$，其

中 $\boldsymbol{P}_i\left(i=1,2,\cdots,m\right)$ 为初等矩阵，因此有

$$\boldsymbol{P}_m^{\mathrm{T}}\cdots\boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m=\boldsymbol{\Lambda}$$

注意到行初等矩阵、列初等矩阵有如下关系：

$$\boldsymbol{R}_{ij}=\boldsymbol{C}_{ij}=\boldsymbol{C}_{ij}^{\mathrm{T}}$$

$$\boldsymbol{R}_{i(k)}=\boldsymbol{C}_{i(k)}=\boldsymbol{C}_{i(k)}^{\mathrm{T}}$$

$$\boldsymbol{R}_{(i,j(k))}=\boldsymbol{C}_{(j,i(k))}=\boldsymbol{C}_{(i,j(k))}^{\mathrm{T}}$$

先考虑 $\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1$，根据 $\boldsymbol{P}_1$ 为三种不同的列初等矩阵，考虑其意义：

1. 若 $\boldsymbol{P}_1=\boldsymbol{C}_{ij}$，则 $\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1=\boldsymbol{C}_{ij}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}_{ij}=\boldsymbol{R}_{ij}\left(\boldsymbol{A}\boldsymbol{C}_{ij}\right)$，$\boldsymbol{A}\boldsymbol{C}_{ij}$ 相当于互换 $\boldsymbol{A}$ 的第 $\boldsymbol{i},\boldsymbol{j}$ 列，而 $\boldsymbol{R}_{ij}\left(\boldsymbol{A}\boldsymbol{C}_{ij}\right)$ 相当于互换 $\boldsymbol{A}\boldsymbol{C}_{ij}$ 的第 $\boldsymbol{i},\boldsymbol{j}$ 行；
2. 若 $\boldsymbol{P}_1=\boldsymbol{C}_{i(k)}$，则 $\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1=\boldsymbol{C}_{i(k)}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}_{i(k)}=\boldsymbol{R}_{i(k)}\left(\boldsymbol{A}\boldsymbol{C}_{i(k)}\right)$，$\boldsymbol{A}\boldsymbol{C}_{i(k)}$ 相当于将 $\boldsymbol{A}$ 的第 $\boldsymbol{i}$ 列乘以常数 k，而 $\boldsymbol{R}_{i(k)}\left(\boldsymbol{A}\boldsymbol{C}_{i(k)}\right)$ 相当于 $\boldsymbol{A}\boldsymbol{C}_{i(k)}$ 的第 $\boldsymbol{i}$ 行乘以常数 k；
3. 若 $\boldsymbol{P}_1=\boldsymbol{C}_{(i,j(k))}$，$\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1=\boldsymbol{C}_{(i,j(k))}^{\mathrm{T}}\left(\boldsymbol{A}\boldsymbol{C}_{(i,j(k))}\right)=\boldsymbol{R}_{(i,j(k))}\left(\boldsymbol{A}\boldsymbol{C}_{(i,j(k))}\right)$，$\boldsymbol{A}\boldsymbol{C}_{(i,j(k))}$ 相当于将 $\boldsymbol{A}$ 第 j 列的 k 倍加到第 i 列，$\boldsymbol{R}_{(i,j(k))}\left(\boldsymbol{A}\boldsymbol{C}_{(i,j(k))}\right)$ 相当于将 $\boldsymbol{A}\boldsymbol{C}_{(i,j(k))}$ 的第 j 行的 k 倍加到第 i 行；

因此，无论 $\boldsymbol{P}_1$ 是哪种类型的列初等矩阵，$\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1$ 都相当于对 $\boldsymbol{A}$ 做了一个初等列变换，然后做“相应”的初等行变换，而 $\boldsymbol{P}_m^{\mathrm{T}}\cdots\boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m=\boldsymbol{\Lambda}$ 相当于对 $\boldsymbol{A}$ 做一序列的初等列变换和相应的初等行变换(显然做这两种变换的次序可交换。Why？)，这些初等变换把 $\boldsymbol{A}$ 化为对角形，由 $\boldsymbol{P}=\boldsymbol{E}\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m$ 可知，这些初等列变换把单位矩阵化为矩阵 $\boldsymbol{P}$。由以上分析就得到初等变换合同法：

$$\begin{bmatrix}\boldsymbol{A}\\ \cdots\cdots\\ \boldsymbol{E}\end{bmatrix}\xrightarrow[\boldsymbol{E}\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m]{\boldsymbol{P}_m^{\mathrm{T}}\cdots\boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_m}\begin{bmatrix}\boldsymbol{\Lambda}\\ \cdots\cdots\\ \boldsymbol{P}\end{bmatrix}$$

即：对矩阵 $\begin{bmatrix}\boldsymbol{A}\\ \cdots\cdots\\ \boldsymbol{E}\end{bmatrix}$ 做初等列变换，同时对 $\boldsymbol{A}$ 做相应的行变换，目标是用这些初等变换将 $\boldsymbol{A}$ 化为对角矩阵，则在该过程中我们在单位矩阵 $\boldsymbol{E}$ 的位置得到可逆矩阵 $\boldsymbol{P}$。下面我们用上述方法将二次型标准化。

例 7.4　用初等变换合同法将下述二次型化为标准型

（1） $f(x_1,x_2,x_3)=x_1^2+2x_2^2+5x_3^2+2x_1x_2+2x_1x_3+6x_2x_3$；

（2） $f(x_1,x_2,x_3)=x_1x_2+4x_1x_3+x_2x_3$。

解：（1）该二次型的矩阵为

$$\begin{bmatrix}1&1&1\\1&2&3\\1&3&5\end{bmatrix}$$

用初等变换

$$\to\begin{bmatrix}1&0&0\\1&1&2\\1&2&4\\ \cdots&\cdots&\cdots\\1&-1&-1\\0&1&0\\0&0&1\end{bmatrix}\to\begin{bmatrix}1&0&0\\1&1&2\\1&2&4\\ \cdots&\cdots&\cdots\\1&-1&-1\\0&1&0\\0&0&1\end{bmatrix}\to\begin{bmatrix}1&0&0\\0&1&2\\0&2&4\\ \cdots&\cdots&\cdots\\1&-1&-1\\0&1&0\\0&0&1\end{bmatrix}\to\begin{bmatrix}1&0&0\\0&1&0\\0&2&0\\ \cdots&\cdots&\cdots\\1&-1&1\\0&1&-2\\0&0&1\end{bmatrix}\to\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\\ \cdots&\cdots&\cdots\\1&-1&1\\0&1&-2\\0&0&1\end{bmatrix}$$

因此，二次型的标准形为

$$f=y_1^2+y_2^2$$

相应线性变换为

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1&-1&1\\0&1&-2\\0&0&1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$$

（2）该二次型的矩阵为

$$\begin{bmatrix}0&\frac{1}{2}&2\\ \frac{1}{2}&0&\frac{1}{2}\\ 2&\frac{1}{2}&0\end{bmatrix}$$

用初等变换法标准化二次型

$$
\left[\begin{array}{ccc} 0 & \frac{1}{2} & 2 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 2 & \frac{1}{2} & 0 \\ \hdashline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc} \frac{1}{2} & \frac{1}{2} & 2 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{5}{2} & \frac{1}{2} & 0 \\ \hdashline 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc} 1 & \frac{1}{2} & \frac{5}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{5}{2} & \frac{1}{2} & 0 \\ \hdashline 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc} 1 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{4} & -\frac{3}{4} \\ \frac{5}{2} & -\frac{3}{4} & -\frac{25}{4} \\ \hdashline 1 & -\frac{1}{2} & -\frac{5}{2} \\ 1 & \frac{1}{2} & -\frac{5}{2} \\ 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & -\frac{3}{4} \\ 0 & -\frac{3}{4} & -\frac{25}{4} \\ \hdashline 1 & -\frac{1}{2} & -\frac{5}{2} \\ 1 & \frac{1}{2} & -\frac{5}{2} \\ 0 & 0 & 1 \end{array}\right]
$$

$$
\to \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & -\frac{3}{4} & -4 \\ \hdashline 1 & -\frac{1}{2} & -1 \\ 1 & \frac{1}{2} & -4 \\ 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -4 \\ \hdashline 1 & -\frac{1}{2} & -1 \\ 1 & \frac{1}{2} & -4 \\ 0 & 0 & 1 \end{array}\right]
$$

注意：二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 的标准型中，系数非零的个数与所作的可逆线性变换无关，等于矩阵 A 的秩。但是标准型平方项前面的系数却与所作的可逆线性变换有关。

习题 7.2

1. 设矩阵 $\boldsymbol{A}=\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$，则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$（　　）

　（A）合同且相似　　　　　（B）合同，但不相似

　（C）不合同，但相似　　　（D）既不合同，也不相似

2. 用配方法化二次型为标准形，并写出所作的可逆线性替换。

　a） $f\left(x_1,x_2,x_3\right)=x_1^2+2x_2^2+5x_3^2+2x_1x_2+2x_1x_3+8x_2x_3$

　b） $f\left(x_1,x_2,x_3\right)=x_1^2+x_2^2+2x_3^2+2x_1x_2+4x_1x_3-x_2x_3$

　c） $f\left(x_1,x_2,x_3\right)=x_1x_2+4x_1x_3-2x_2x_3$

d）　$f\left(x_1,x_2,x_3\right)=2x_1^2+5x_2^2+5x_3^2+4x_1x_2-4x_1x_3-8x_2x_3$

2. 用合同变换法将题 1 中的二次型化为标准形。

§7.3　用正交变换化二次型为标准形

定义 7.3　设 $\boldsymbol{C}$ 为 n 阶实矩阵，如果 $\boldsymbol{C}$ 满足

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}=\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{E} \tag{7.5}$$

则称 $\boldsymbol{C}$ 为正交矩阵。

定理 7.2　正交矩阵的性质

1. 正交矩阵的行列式为 1 或−1；
2. 若 $\boldsymbol{C}$ 为正交矩阵，则 $\boldsymbol{C}$ 可逆，且 $\boldsymbol{C}^{-1}=\boldsymbol{C}^{\mathrm{T}}$ 也为正交矩阵；
3. 若 $\boldsymbol{A},\boldsymbol{B}$ 均为正交矩阵，则 $\boldsymbol{AB}$ 也为正交矩阵；
4. $\boldsymbol{C}$ 为正交矩阵的充要条件是 $\boldsymbol{C}$ 的列(行)向量组是标准正交向量组。

证明：

1. $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}=\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{E}$

$$\left|\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\right|=\left|\boldsymbol{E}\right|$$

$$1=\left|\boldsymbol{E}\right|=\left|\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}\right|=\left|\boldsymbol{C}^{\mathrm{T}}\right|\left|\boldsymbol{C}\right|=\left|\boldsymbol{C}\right|^2$$

所以　$\left|\boldsymbol{C}\right|=\pm1$；

2. 由 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}=\boldsymbol{E}$，根据定理 2.4 可知 $\boldsymbol{C}$ 可逆，且 $\boldsymbol{C}^{-1}=\boldsymbol{C}^{\mathrm{T}}$。而

$$\boldsymbol{C}^{\mathrm{T}}\left(\boldsymbol{C}^{\mathrm{T}}\right)^{\mathrm{T}}=\left(\boldsymbol{C}^{\mathrm{T}}\right)^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{E}$$

所以 $\boldsymbol{C}^{\mathrm{T}}$ 也为正交矩阵。

3. 若 $\boldsymbol{A},\ \boldsymbol{B}$ 均为正交矩阵，则

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{E},\qquad \boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}=\boldsymbol{B}\,\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{E}$$

因此

$$\left(\boldsymbol{AB}\right)^{\mathrm{T}}\left(\boldsymbol{AB}\right)=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AB}=\boldsymbol{B}^{\mathrm{T}}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{B}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{B}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}=\boldsymbol{E}$$

类似可证 $\left(\boldsymbol{AB}\right)\left(\boldsymbol{AB}\right)^{\mathrm{T}}=\boldsymbol{E}$

故 $\boldsymbol{AB}$ 为正交矩阵。

4. 设$\boldsymbol{C}$的列分块为$\boldsymbol{C}=[\boldsymbol{C}_1\ \boldsymbol{C}_2\ \cdots\ \boldsymbol{C}_n]$，则

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}=[\boldsymbol{C}_1\ \boldsymbol{C}_2\ \cdots\boldsymbol{C}_n]^{\mathrm{T}}[\boldsymbol{C}_1\ \boldsymbol{C}_2\ \cdots\boldsymbol{C}_n]$$

$$=\begin{bmatrix}\boldsymbol{C}_1^{\mathrm{T}}\\ \boldsymbol{C}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{C}_n^{\mathrm{T}}\end{bmatrix}[\boldsymbol{C}_1\ \boldsymbol{C}_2\ \cdots\boldsymbol{C}_n]$$

$$=\begin{bmatrix}\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{C}_1 & \boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{C}_2 & \cdots & \boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{C}_n\\ \boldsymbol{C}_2^{\mathrm{T}}\boldsymbol{C}_1 & \boldsymbol{C}_2^{\mathrm{T}}\boldsymbol{C}_2 & \cdots & \boldsymbol{C}_2^{\mathrm{T}}\boldsymbol{C}_n\\ \vdots & \vdots & & \vdots\\ \boldsymbol{C}_n^{\mathrm{T}}\boldsymbol{C}_1 & \boldsymbol{C}_n^{\mathrm{T}}\boldsymbol{C}_2 & \cdots & \boldsymbol{C}_n^{\mathrm{T}}\boldsymbol{C}_n\end{bmatrix}=\boldsymbol{E}=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}$$

故
$$\boldsymbol{C}_i^{\mathrm{T}}\boldsymbol{C}_j=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}$$

因此$\boldsymbol{C}$的列向量组为标准正交向量组。

当$\boldsymbol{C}$为正交矩阵时，$\boldsymbol{C}^{\mathrm{T}}$也为正交矩阵，故$\boldsymbol{C}^{\mathrm{T}}$的列向量组——也就是$\boldsymbol{C}$的行向量组为正交向量组。

反之，如果矩阵$\boldsymbol{C}$的行(列)向量组为正交向量组，则$\boldsymbol{C}$显然为正交矩阵。

■

定义 7.4　　正交变换

设$\boldsymbol{C}$为n阶正交矩阵，$\boldsymbol{X},\boldsymbol{Y}$是欧氏空间$\mathbb{R}^n$中的$n$维向量，则称线性变换$\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$是$\mathbb{R}^n$上的正交变换。

显然正交变换是可逆线性变换。

定理 7.3　　正交变换的性质

设$\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$是欧氏空间$\mathbb{R}^n$上的线性变换，则下列命题等价：

1. 线性变换$\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$是正交变换；
2. 在线性变换$\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$下，向量的内积不变，即$\boldsymbol{X}_1=\boldsymbol{C}\boldsymbol{Y}_1$，$\boldsymbol{X}_2=\boldsymbol{C}\boldsymbol{Y}_2$时，

   $(\boldsymbol{X}_1,\boldsymbol{X}_2)=(\boldsymbol{Y}_1,\boldsymbol{Y}_2)$；
3. 线性变换$\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$把$\mathbb{R}^n$中的标准正交基变成标准正交基；

证明：1 $\Rightarrow$ 2

$$\begin{aligned}(\boldsymbol{X}_1,\boldsymbol{X}_2)&=\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2=(\boldsymbol{CY}_1)^{\mathrm{T}}(\boldsymbol{CY}_2)\\&=\boldsymbol{Y}_1^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{CY}_2=\boldsymbol{Y}_1^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C})\boldsymbol{Y}_2=\boldsymbol{Y}_1^{\mathrm{T}}\boldsymbol{EY}_2=\boldsymbol{Y}_1^{\mathrm{T}}\boldsymbol{Y}_2\\&=(\boldsymbol{Y}_1,\boldsymbol{Y}_2)\end{aligned}$$

2 $\Rightarrow$ 3

任取 $\mathbb{R}^n$ 中的标准正交基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$，经过正交变换 $\boldsymbol{X}=\boldsymbol{CY}$ 得到 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$，即

$$\boldsymbol{\beta}_i=\boldsymbol{C\alpha}_i,\qquad i=1,2,\cdots,n$$

由于正交变换保持内积不变，因此

$$(\boldsymbol{\beta}_i,\boldsymbol{\beta}_j)=(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=\begin{cases}1, & i=j\\0, & i\neq j\end{cases}$$

所以 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 也是标准正交基。

3 $\Rightarrow$1

假设线性变换 $\boldsymbol{X}=\boldsymbol{CY}$ 把 $\mathbb{R}^n$ 中的标准正交基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 变为标准正交基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$，其中

$$\boldsymbol{\beta}_i=\boldsymbol{C\alpha}_i,\qquad i=1,2,\cdots,n$$

即有

$$[\boldsymbol{\beta}_1\quad\boldsymbol{\beta}_2\quad\cdots\quad\boldsymbol{\beta}_n]=\boldsymbol{C}[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\cdots\quad\boldsymbol{\alpha}_n]$$

令

$$\boldsymbol{B}=[\boldsymbol{\beta}_1\quad\boldsymbol{\beta}_2\quad\cdots\quad\boldsymbol{\beta}_n]\qquad\boldsymbol{A}=[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\cdots\quad\boldsymbol{\alpha}_n]$$

则方阵 $\boldsymbol{A},\boldsymbol{B}$ 均为正交矩阵，因此 $\boldsymbol{C}=\boldsymbol{BA}^{-1}$ 也为正交矩阵，故 $\boldsymbol{X}=\boldsymbol{CY}$ 为正交变换。

正交变换的例子（要求具有比较明显的几何意义）

由正交变换的性质 2 可知，

$$(\boldsymbol{CX},\boldsymbol{CX})=(\boldsymbol{X},\boldsymbol{X})$$

即正交变换不改变向量的长度，进一步地有

$$\frac{(\boldsymbol{CX}_1,\boldsymbol{CX}_2)}{\|\boldsymbol{CX}_1\|\bullet\|\boldsymbol{CX}_2\|}=\frac{(\boldsymbol{X}_1,\boldsymbol{X}_2)}{\|\boldsymbol{X}_1\|\bullet\|\boldsymbol{X}_2\|}$$

因此正交变换也不改变任意向量之间的夹角。我们看一下 $\mathbb{R}^2$ 上的正交变换有何特点？设 $\boldsymbol{X}=\boldsymbol{CY}$ 是 $\mathbb{R}^2$ 上的正交变换，故 $\boldsymbol{C}=\begin{bmatrix}c_{11}&c_{12}\\c_{21}&c_{22}\end{bmatrix}$ 为二阶正交矩阵，其列向量组为标准正交

向量组，即有

$$c_{11}^2+c_{21}^2=1$$
$$c_{12}^2+c_{22}^2=1$$
$$c_{11}c_{12}+c_{21}c_{22}=0$$

因此可设

$$c_{11}=\cos\theta,\ c_{21}=\sin\theta,\ c_{12}=\cos\varphi,\ c_{22}=\sin\varphi$$

由

$$0=c_{11}c_{12}+c_{21}c_{22}=\cos\theta\cos\varphi+\sin\theta\sin\varphi=\cos(\theta-\varphi)$$

因此　$\theta-\varphi=\pm\dfrac{\pi}{2}$，即 $\varphi=\theta\pm\dfrac{\pi}{2}$。

当 $\varphi=\theta+\dfrac{\pi}{2}$ 时，$\boldsymbol{C}=\begin{bmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{bmatrix}=\begin{bmatrix}\cos\theta & \cos\left(\theta+\dfrac{\pi}{2}\right)\\ \sin\theta & \sin\left(\theta+\dfrac{\pi}{2}\right)\end{bmatrix}=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$；

当 $\varphi=\theta-\dfrac{\pi}{2}$ 时，$\boldsymbol{C}=\begin{bmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{bmatrix}=\begin{bmatrix}\cos\theta & \cos\left(\theta-\dfrac{\pi}{2}\right)\\ \sin\theta & \sin\left(\theta-\dfrac{\pi}{2}\right)\end{bmatrix}=\begin{bmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{bmatrix}$。

对于第一种情况。考虑向量 $\boldsymbol{\alpha}=[r\sin\phi,r\cos\phi]^{\mathrm{T}}$ 在正交变换下的结果 $\boldsymbol{C\alpha}$ 与 $\boldsymbol{\alpha}$ 的关系。

$$\boldsymbol{C\alpha}=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}r\cos\phi\\ r\sin\phi\end{bmatrix}=r\begin{bmatrix}\cos\theta\cos\phi-\sin\theta\sin\phi\\ \sin\theta\cos\phi+\cos\theta\sin\phi\end{bmatrix}$$
$$=r\begin{bmatrix}\cos(\phi+\theta)\\ \sin(\phi+\theta)\end{bmatrix}=\begin{bmatrix}r\cos(\phi+\theta)\\ r\sin(\phi+\theta)\end{bmatrix}$$

因此 $\boldsymbol{C\alpha}$ 是向量 $\boldsymbol{\alpha}$ 逆时针旋转角度 θ 得到的向量。

对于第二种情况。考虑向量 $\boldsymbol{\alpha}=[r\sin\phi,r\cos\phi]^{\mathrm{T}}$ 在正交变换下的结果 $\boldsymbol{C\alpha}$ 与 $\boldsymbol{\alpha}$ 的关系。

$$\boldsymbol{C\alpha}=\begin{bmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{bmatrix}\begin{bmatrix}r\cos\phi\\ r\sin\phi\end{bmatrix}=r\begin{bmatrix}\cos\theta\cos\phi+\sin\theta\sin\phi\\ \sin\theta\cos\phi-\cos\theta\sin\phi\end{bmatrix}$$
$$=r\begin{bmatrix}\cos(\theta-\phi)\\ \sin(\theta-\phi)\end{bmatrix}=\begin{bmatrix}r\cos(\theta-\phi)\\ r\sin(\theta-\phi)\end{bmatrix}=\begin{bmatrix}r\cos(\phi-\theta)\\ -r\sin(\phi-\theta)\end{bmatrix}$$

因此 $\boldsymbol{C\alpha}$ 可这样得到：首先将 $\boldsymbol{\alpha}$ 顺时针旋转 θ 得到向量 $\boldsymbol{\beta}$，$\boldsymbol{C\alpha}$ 是与 $\boldsymbol{\beta}$ 关于 x 轴对称的向量。

正交变换化二次型为标准形

假设正交变换 $\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$ 可将二次型 $f=\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X}$ 为标准形，则相当于$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵，故 $\boldsymbol{A}$ 可相似对角化，矩阵$\boldsymbol{C}$ 的列向量为 $\boldsymbol{A}$ 的特征向量，由于$\boldsymbol{C}$ 为正交矩阵，它的列向量组为正交向量组，因此此时 $\boldsymbol{A}$ 必有 n 个相互正交的特征向量。假设 $\boldsymbol{A}$ 的特征值 λ 有 k 个线性无关的特征向量，我们可应用 Schmidt 正交化过程得到一个标准正交向量组，其中每个向量都为对应于 λ 的特征向量。关键是：不同特征值对应的特征向量是否正交？

定理 6.4　设 $\boldsymbol{A}_{n\times n}$ 为实对称矩阵，则有

(1) $\boldsymbol{A}$ 的特征值都是实数；
(2) $\boldsymbol{A}$ 的属于不同特征值的特征向量必正交。

证明(1)设 λ 为 $\boldsymbol{A}$ 的特征值，而 $\boldsymbol{X}$ 为对应于 λ 的特征向量，即

$$\boldsymbol{A}\boldsymbol{X}=\lambda\boldsymbol{X}$$

对上式取共轭：

$$\overline{\boldsymbol{A}\boldsymbol{X}}=\overline{\lambda\boldsymbol{X}}$$

$$\overline{\boldsymbol{A}}\ \overline{\boldsymbol{X}}=\overline{\lambda}\ \overline{\boldsymbol{X}}$$

$$\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{A}}\ \overline{\boldsymbol{X}}=\boldsymbol{X}^{\mathrm{T}}\overline{\lambda}\ \overline{\boldsymbol{X}}\ =\overline{\lambda}\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}}$$

由于 $\boldsymbol{A}$ 为实对称矩阵，故

$$\overline{\boldsymbol{A}}=\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$$

因此

$$\begin{aligned}\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{A}}\ \overline{\boldsymbol{X}}&=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\ \overline{\boldsymbol{X}}\ =(\boldsymbol{A}\boldsymbol{X})^{\mathrm{T}}\overline{\boldsymbol{X}}=(\lambda\boldsymbol{X})^{\mathrm{T}}\overline{\boldsymbol{X}}\\&=\lambda\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}}\end{aligned}$$

$$\overline{\lambda}\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}}=\lambda\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}}$$

由于 $\boldsymbol{X}\neq\boldsymbol{0}$，从而 $\boldsymbol{X}^{\mathrm{T}}\overline{\boldsymbol{X}}\neq 0$，所以 $\overline{\lambda}=\lambda$，即 λ 为实数。

(2)设 λ_1,λ_2 为 $\boldsymbol{A}$ 的两个不同特征值，而 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 为分别与之对应的特征向量，则

$$\boldsymbol{A}\boldsymbol{X}_1=\lambda_1\boldsymbol{X}_1$$

$$\boldsymbol{A}\boldsymbol{X}_2=\lambda_2\boldsymbol{X}_2$$

因此有

$$\begin{aligned}&\lambda_1\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2=(\lambda_1\boldsymbol{X}_1)^{\mathrm{T}}\boldsymbol{X}_2=(\boldsymbol{A}\boldsymbol{X}_1)^{\mathrm{T}}\boldsymbol{X}_2=\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}_2\\&=\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}_2=\boldsymbol{X}_1^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{X}_2)=\boldsymbol{X}_1^{\mathrm{T}}(\lambda_2\boldsymbol{X}_2)\\&=\lambda_2\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2\end{aligned}$$

即　$\lambda_1\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2=\lambda_2\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2$

由于 $\lambda_1 \neq \lambda_2$ ，故 $\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2 = \left(\boldsymbol{X}_1, \boldsymbol{X}_2\right) = 0$，命题得证。

还有一个问题：n 阶实对称矩阵是否有 n 个线性无关的特征向量？也就是说，对于代数重数为 t 的特征值 λ，其几何重数是否也为 t，即 λ 是否有 t 个线性无关的特征向量，或者说 $r\left(\lambda\boldsymbol{E}-\boldsymbol{A}\right)=n-t$ 成立？

定理 7.5　对 n 阶实对称矩阵 $\boldsymbol{A}$，必存在正交矩阵 $\boldsymbol{C}$，使

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC} = \boldsymbol{C}^{-1}\boldsymbol{AC} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值，$\boldsymbol{C}$ 的 n 个列向量是 $\boldsymbol{A}$ 的对应于特征值的标准正交特征向量。

证明：对 n 做归纳法。$n=1$ 时结论显然成立。假设定理对 $n-1$ 时成立，下证明定理在 n 时也成立。

设 λ_1 是 $\boldsymbol{A}$ 的一个特征值，$\boldsymbol{AX}_1=\lambda_1\boldsymbol{X}_1$，其中 $\boldsymbol{X}_1$ 为单位向量，现将 $\boldsymbol{X}_1$ 扩充为 $\mathbb{R}^n$ 的一组标准正交基 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$，于是 $\boldsymbol{AX}_j$ 可由这组标准正交基 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$ 线性表示，从而有

$$\boldsymbol{A}\left[\boldsymbol{X}_1 \quad \boldsymbol{X}_2 \quad \cdots \quad \boldsymbol{X}_n\right] = \left[\boldsymbol{X}_1 \quad \boldsymbol{X}_2 \quad \cdots \quad \boldsymbol{X}_n\right]\begin{bmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{bmatrix}$$

设 $\boldsymbol{P}=\left[\boldsymbol{X}_1 \quad \boldsymbol{X}_2 \quad \cdots \quad \boldsymbol{X}_n\right]$，则 $\boldsymbol{P}$ 为正交矩阵。将等式右边的矩阵用分块矩阵表示，则有

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{bmatrix} \lambda_1 & \boldsymbol{b} \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix}$$

由于 $\boldsymbol{P}^{-1}=\boldsymbol{P}^{\mathrm{T}}$，故 $\left(\boldsymbol{P}^{-1}\boldsymbol{AP}\right)^{\mathrm{T}} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{P}^{-1}\right)^{\mathrm{T}} = \boldsymbol{P}^{-1}\boldsymbol{AP}$，因而 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为实对称矩阵，因此上式中 $\boldsymbol{b}=\boldsymbol{0}$，$\boldsymbol{B}$ 为 $n-1$ 阶实对称矩阵。根据归纳假设，存在 $n-1$ 阶正交矩阵 $\boldsymbol{Q}_1$，使得

$$\boldsymbol{Q}_1^{-1}\boldsymbol{BQ}_1 = \boldsymbol{diag}\left(\lambda_2,\cdots,\lambda_n\right)$$

取 $\boldsymbol{S}=\begin{bmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_1 \end{bmatrix}$，则 $\boldsymbol{S}$ 是正交矩阵。则有

$$S^{-1}\left(P^{-1}AP\right)S=\begin{bmatrix}1 & \mathbf{0}\\ \mathbf{0} & Q_1^{-1}\end{bmatrix}\begin{bmatrix}\lambda_1 & 0\\ 0 & B\end{bmatrix}\begin{bmatrix}1 & \mathbf{0}\\ \mathbf{0} & Q_1\end{bmatrix}$$

$$=\begin{bmatrix}\lambda_1 & \mathbf{0}\\ \mathbf{0} & Q_1^{-1}BQ_1\end{bmatrix}=diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$$

为对角矩阵。令 $Q=PS$，则 Q 也是正交矩阵。故有 $Q^{-1}AQ=diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$。命题得证。

定理 7.6　（主轴定理）实二次型 $X^{\mathrm{T}}AX$ 必可由正交变换 $X=CY$ 化为标准形，即

$$f\left(x_1,x_2,\cdots,x_n\right)=X^{\mathrm{T}}AX\ \underline{\underline{\quad X=PY\quad}}\ \lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2$$

结合上一章矩阵相似对角化的知识与本节的内容可知，正交变换化二次型为标准型的计算过程为：

1、由 $|\lambda E-A|=0$，计算 A 的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$；

2、对 λ_i，求解齐次线性方程组 $\left(\lambda_iE-A\right)X=\mathbf{0}$，得到该方程组的一个基础解系，也就得到 A 的关于特征值 λ_i 的线性无关的特征向量；

3、对于单特征值，将其特征向量单位化；对 $t\left(t>1\right)$ 重特征值 λ_i，用 Schmidt 正交化方法将其 t 个线性无关的特征向量正交化，然后单位化；

例 7.5　用正交变换将二次型

$$f(x_1,x_2,x_3,x_4)=2x_1x_2+2x_1x_3-2x_1x_4-2x_2x_3+2x_2x_4+2x_3x_4$$

化为标准形。

解：该二次型的矩阵为

$$\begin{bmatrix}0 & 1 & 1 & -1\\ 1 & 0 & -1 & 1\\ 1 & -1 & 0 & 1\\ -1 & 1 & 1 & 0\end{bmatrix}$$

（1）先求该矩阵的特征值

$$f(\lambda)=|\lambda E-A|=\begin{vmatrix}\lambda & -1 & -1 & 1\\ -1 & \lambda & 1 & -1\\ -1 & 1 & \lambda & -1\\ 1 & -1 & -1 & \lambda\end{vmatrix}$$

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

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

$$=(\lambda-1)^2\left((\lambda+1)^2-4\right)=(\lambda-1)^3(\lambda+3)$$

令　　$f(\lambda)=(\lambda-1)^3(\lambda+3)=0$

得到特征值

$$\lambda_1=\lambda_2=\lambda_3=1 \qquad \lambda_4=-3$$

（2）求特征向量：

对应于特征值 1 的特征向量，即求解齐次线性方程组

$$\left(\boldsymbol{E}-\boldsymbol{A}\right)\boldsymbol{X}=\begin{bmatrix}1&-1&-1&1\\-1&1&1&-1\\-1&1&1&-1\\1&-1&-1&1\end{bmatrix}\boldsymbol{X}=\boldsymbol{0}$$

得到它的通解为

$$\boldsymbol{X}=k_1\begin{bmatrix}1\\1\\0\\0\end{bmatrix}+k_2\begin{bmatrix}1\\0\\1\\0\end{bmatrix}+k_3\begin{bmatrix}-1\\0\\0\\1\end{bmatrix}$$

它的一个基础解系为

$$\boldsymbol{\alpha}_1=\begin{bmatrix}1\\1\\0\\0\end{bmatrix},\boldsymbol{\alpha}_2=\begin{bmatrix}1\\0\\1\\0\end{bmatrix},\boldsymbol{\alpha}_3=\begin{bmatrix}-1\\0\\0\\1\end{bmatrix}$$

用 Schmidt 正交化过程正交化、单位化：

$$\boldsymbol{X}_1^0=\frac{\boldsymbol{X}_1}{|\boldsymbol{X}_1|}=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\\0\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\\0\end{bmatrix}$$

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\left(\boldsymbol{\alpha}_2,\boldsymbol{X}_1^0\right)\boldsymbol{X}_1^0=\begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\\0\end{bmatrix},\qquad \boldsymbol{X}_2^0=\frac{\boldsymbol{\beta}_2}{|\boldsymbol{\beta}_2|}=\sqrt{\frac{2}{3}}\begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\\0\end{bmatrix}=\begin{bmatrix}\sqrt{\frac{1}{6}}\\-\sqrt{\frac{1}{6}}\\\sqrt{\frac{2}{3}}\\0\end{bmatrix}$$

$$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\left(\boldsymbol{\alpha}_3,\boldsymbol{X}_1^0\right)\boldsymbol{X}_1^0-\left(\boldsymbol{\alpha}_3,\boldsymbol{X}_2^0\right)\boldsymbol{X}_2^0=\frac{1}{3}\begin{bmatrix}-1\\1\\1\\3\end{bmatrix},\qquad \boldsymbol{X}_3^0=\frac{\boldsymbol{\beta}_3}{|\boldsymbol{\beta}_3|}=\begin{bmatrix}\frac{-1}{\sqrt{12}}\\\frac{1}{\sqrt{12}}\\\frac{1}{\sqrt{12}}\\\frac{3}{\sqrt{12}}\end{bmatrix}$$

求关于特征值–3 的特征值，即求解齐次线性方程组

$$\left(-3\boldsymbol{E}-\boldsymbol{A}\right)\boldsymbol{X}=\begin{bmatrix}-3&-1&-1&1\\-1&-3&1&-1\\-1&1&-3&-1\\1&-1&-1&-3\end{bmatrix}\boldsymbol{X}=\boldsymbol{0}$$

求得一个解为 $\boldsymbol{X}=[1,-1,-1,1]^{\mathrm{T}}$，它就是对应于–3 的特征向量，单位化得

$$\boldsymbol{X}_4^0=\left[\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right]^{\mathrm{T}}$$

因此，有

$$\boldsymbol{C}=\begin{bmatrix}\boldsymbol{X}_1^0&\boldsymbol{X}_2^0&\boldsymbol{X}_3^0&\boldsymbol{X}_4^0\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}&\frac{-1}{\sqrt{12}}&\frac{1}{2}\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{12}}&-\frac{1}{2}\\0&\sqrt{\frac{2}{3}}&\frac{1}{\sqrt{12}}&-\frac{1}{2}\\0&0&\frac{3}{\sqrt{12}}&\frac{1}{2}\end{bmatrix}$$

所用的正交变换为 $\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$，所得的标准型为

$$f(x_1,x_2,x_3,x_4)\underline{\underline{\ \boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}\ }}y_1^2+y_2^2+y_3^2-3y_4^2$$

习题 7. 3

1. 用正交变换将习题 7.2 中题 1 中的二次型化为标准形。

2. 设 3 阶实对称矩阵 $\boldsymbol{A}$ 的各行元素之和均为 3, 向量 $\boldsymbol{\alpha}_1=[-1,2,-1]^{\mathrm{T}}$,$\boldsymbol{\alpha}_2=[0,-1,1]^{\mathrm{T}}$ 是线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的两个解, (1)求 $\boldsymbol{A}$ 的特征值与特征向量 ；（Ⅱ)求正交矩阵 $\boldsymbol{Q}$ 和对角矩阵 $\boldsymbol{\Lambda}$,使得 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{AQ}=\boldsymbol{\Lambda}$.（3）求 $\boldsymbol{A}$ 及 $(\boldsymbol{A}-\boldsymbol{E})^6$ 。

3. 设 $\boldsymbol{X}=[1,1,2]^{\mathrm{T}}$ 是 $\boldsymbol{A}=\begin{bmatrix}0&1&2\\1&0&a\\2&a&b\end{bmatrix}$ 的特征向量，求 a,b 的值，并求正交矩阵 $\boldsymbol{P}$ ，使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为对角阵。

4. 已知三阶实对称矩阵 $\boldsymbol{A}$ 的特征值为 $1,1,-1$，特征值 1 对应的线性无关的特征向量为 $[1,-1,0]^{\mathrm{T}},[1,0,-1]^{\mathrm{T}}$.

（1）求 $\boldsymbol{A}$;

（2）求出正交变换 $\boldsymbol{X}=\boldsymbol{CY}$ ，使得二次型 $f(x_1,x_2,x_3)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 化为标准形.

5. 已知三阶实对称矩阵 $\boldsymbol{A}$ 的三个特征值为 $1,1,-1$，特征值－1 对应的的特征向量为 $[0,1,1]^{\mathrm{T}}$，求特征值 1 对应的特征向量和矩阵 $\boldsymbol{A}$.

6. 设 $\boldsymbol{A}$ 为三阶实对称矩阵，已知 $|\boldsymbol{A}|=-12$ ，$\boldsymbol{A}$ 的三个特征值之和为 1。又 $\boldsymbol{\alpha}=\begin{bmatrix}1\\0\\-2\end{bmatrix}$ 是齐次线性方程组 $(\boldsymbol{A}^*-4\boldsymbol{E})\boldsymbol{X}=\boldsymbol{0}$ 的一个解向量,

(1)求 $\boldsymbol{A}$ ；(2)求 $(\boldsymbol{A}^*+10\boldsymbol{E})\boldsymbol{X}=\boldsymbol{0}$ 的通解；(3)求正交变换矩阵 $\boldsymbol{Q}$ ，化二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 为标准形.

7. 设 n 阶实矩阵 $\boldsymbol{A}$ 有 n 个两两正交的特征向量 $\xi_1,\xi_2,\cdots,\xi_n$ ，证明：$\boldsymbol{A}$ 是对称矩阵。

8. 若 $\boldsymbol{A}$ 是 n 阶实反对称矩阵，证明：（1）$\boldsymbol{A}$ 的特征值只能是 0 或纯虚数；（2）$\boldsymbol{E}+\boldsymbol{A}$ 是可逆矩阵；（3）$\boldsymbol{B}=(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A})^{-1}$ 为正交矩阵。

9. 设 $f(x_1,x_2,\cdots,x_n)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 为一 n 元实二次型，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值，且 $\lambda_1\le\lambda_2\le\cdots\le\lambda_n$ ，证明：$\forall\boldsymbol{X}\in\mathbb{R}^n$ ，均有 $\lambda_1\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\le\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}\le\lambda_n\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 。

10. 设 $\boldsymbol{A}$ 是 n 阶实对称矩阵，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个互异的特征值，$\boldsymbol{X}_1$ 是对应于 λ_1 的

单位特征向量。证明：$\boldsymbol{A}-\lambda_1\boldsymbol{X}_1\boldsymbol{X}_1^{\mathrm{T}}$ 的特征值为 $0,\lambda_2,\cdots,\lambda_n$。

11. 设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶实对称矩阵，且 $\boldsymbol{AB}=\boldsymbol{BA}$。证明：存在一个 n 阶正交矩阵 $\boldsymbol{Q}$，使得 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{AQ}$ 与 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{BQ}$ 同时为对角矩阵。

§7.4　二次型的正定性

n 元二次型

$$
\begin{aligned}
f\left(x_1,x_2,\cdots,x_n\right) &= a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n \\
&\quad +a_{21}x_2x_1+a_{22}x_2^2+\cdots+a_{2n}x_2x_n \\
&\quad +\cdots+a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j
\end{aligned}
$$

$$
=\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
$$

$$
=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}
$$

是 n 个变量 $x_1,x_2,\cdots,x_n$ 的二次齐次多项式；若每个实变量取一个值，则可由二次型可算得一个实数，譬如三元二次型

$$
f\left(x_1,x_2,x_3\right)=3x_1^2+x_2^2-x_3^2+2x_1x_2-4x_2x_3
$$

若令

$$
x_1=x_2=x_3=1
$$

则
$$
f\left(1,1,1\right)=1
$$

显然这时把二次型看作是多元函数。

对于二次型

$$
f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}
$$

显然有 $f\left(0,0,\cdots,0\right)=0$

定义 7.5　正定二次型，负定二次型，半正定二次型，半负定二次型，不定二次型

如果二次型 $f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 满足

$$f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}>0\ ,\qquad \boldsymbol{X}\neq\mathbf{0}$$

则称该二次型为**正定二次型**，而称正定二次型的矩阵 $\boldsymbol{A}$ 为**正定矩阵**；

如果二次型 $f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 满足

$$f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}<0\ ,\qquad \boldsymbol{X}\neq\mathbf{0}$$

则称该二次型为**负定二次型**，而称负定二次型的矩阵 $\boldsymbol{A}$ 为**负定矩阵**；

如果二次型 $f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 满足

$$f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\geq 0$$

则称该二次型为**半正定二次型**，而称二次型的矩阵 $\boldsymbol{A}$ 为**半正定矩阵**；

如果二次型 $-f$ 为半正定二次型，则称 f 为**半负定二次型**，称其矩阵为**半负定矩阵**。

对于二次型 $f\left(\boldsymbol{X}\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ ，如果同时存在 $\boldsymbol{X}_0,\boldsymbol{Y}_0$，使得

$$f\left(\boldsymbol{X}_0\right)>0,f\left(\boldsymbol{Y}_0\right)<0$$

则称该二次型是**不定二次型**。

容易验证，下列各个二次型的特性。

$f_1(x_1,x_2,x_3,x_4)=2x_1^2+x_2^2+x_3^2+5x_4^2$　　正定

$f_2(x_1,x_2,x_3)=x_1^2+2x_3^2$　　半正定

$f_3(x_1,x_2,x_3)=x_1^2-4x_2^2+3x_3^2$　　不定

以下我们讨论如何判断二次型的正定性。

首先考虑最简单的情况，从二次型的标准形判断一个二次型是否正定。

定理 7.7　二次型

$$f\left(\boldsymbol{X}\right)=f\left(x_1,x_2,\cdots,x_n\right)=d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2$$

为正定二次型 $\Leftrightarrow d_i>0\quad i=1,2,\cdots,n$ 成立。

证明：必要性：若 $f\left(\boldsymbol{X}\right)=f\left(x_1,x_2,\cdots,x_n\right)=d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2$ 正定，则由 $\boldsymbol{e}_i\neq\mathbf{0}$，

可知　$f\left(\boldsymbol{e}_i\right)=d_i>0$；

充分性：由于 $d_i>0\quad i=1,2,\cdots,n$，故当 $\boldsymbol{X}\neq\mathbf{0}$ 时，至少有某个 $x_i\neq 0$，所以

$$f\left(\boldsymbol{X}\right)=f\left(x_1,x_2,\cdots,x_n\right)=d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2\geq d_ix_i^2>0$$

因此 f 是正定二次型。

二次型都可以化为标准形，那么我们能否通过二次型的标准形来判断其正定性呢？

上述定理说明：我们可将二次型化为标准形来判断它的正定性。但是二次型的标准形不唯一，但是同一个二次型的所有标准型所含正的平方项的个数、负平方项的个数不会改变。这就是惯性定理。

定理 6.8(惯性定理)　实二次型

$$f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$$

经可逆线性变换化为标准形时，其标准形中正、负平方项的个数是唯一确定的，它们的和等于矩阵 $\boldsymbol{A}$ 的秩。

证明：设 f 经过可逆线性变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$ 化为标准形

$$f=d_1y_1^2+\cdots+d_ky_k^2-d_{k+1}y_{k+1}^2-\cdots-d_{k+l}y_{k+l}^2 \tag{7.6}$$

$$d_i>0,\quad i=1,2,\cdots,k+l,k+l\leq n$$

设 f 经过可逆线性变换 $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Z}$ 化为标准形

$$f=\lambda_1z_1^2+\cdots+\lambda_sz_s^2-\lambda_{s+1}z_{s+1}^2-\cdots-\lambda_{s+t}z_{s+t}^2 \tag{7.7}$$

$$\lambda_i>0,i=1,2,\cdots,s+t,s+t\leq n$$

显然 $k+l=s+t=r\left(\boldsymbol{A}\right)$ 。

如果 $k\neq s$，不妨设 $k<s$，由于 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$， $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Z}$，故 $\boldsymbol{Y},\boldsymbol{Z}$ 之间有可逆线性变换

$$\boldsymbol{Y}=\boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{Z}$$

考虑以 $z_1,z_2,\cdots,z_n$ 为未知数的齐次线性方程组

$$\begin{cases}y_1=0\\ \vdots\\ y_k=0\\ z_{s+1}=0\\ \vdots\\ z_n=0\end{cases}$$

由于$k<s$，故该方程组中方程的个数$<n$，因此该方程组有非零解，设该非零解为

$$\boldsymbol{Z}=\left[z_1,z_2,\cdots,z_s,0,\cdots,0\right]^{\mathrm{T}}$$

由(4.2)得：$f>0$；由(4.1)得$f\le 0$，矛盾。显然$k>s$也不成立，因此必有$k=s$。

推论：任何一个实二次型$f(x_1,x_2,\cdots,x_n)$经可逆线性变换化为规范形

$$f\left(x_1,x_2,\cdots,x_n\right)\xlongequal{\boldsymbol{X}=\boldsymbol{PY}}y_1^2+y_2^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2$$

时，该规范型是唯一的。

定义 7.6　实二次型$f\left(x_1,x_2,\cdots,x_n\right)$的标准形或规范形中正平方项的个数称为二次型$f\left(x_1,x_2,\cdots,x_n\right)$的**正惯性指数**，负平方项的个数称为二次型$f\left(x_1,x_2,\cdots,x_n\right)$的**负惯性指数**。正惯性指数与负惯性指数之差称为二次型$f\left(x_1,x_2,\cdots,x_n\right)$的**符号差**。

定理 6.9　n元实二次型正定⇔它的正惯性指数等于n。

证明：

$$f\left(x_1,x_2,\cdots,x_n\right)\xlongequal{\boldsymbol{X}=\boldsymbol{PY}}d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2$$

必要性：由于$f\left(x_1,x_2,\cdots,x_n\right)$正定，如果它的正惯性指数$<n$，则$d_1,d_2,\cdots,d_n$至少有一个$\le 0$，不妨设$d_1\le 0$，则令$\boldsymbol{Y}=\boldsymbol{e}_1$，$\boldsymbol{X}=\boldsymbol{PY}\ne\boldsymbol{0}$，则$f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}\right)\boldsymbol{Y}=d_1\le 0$，这与$f\left(x_1,x_2,\cdots,x_n\right)$正定矛盾。

充分性：如果$f\left(x_1,x_2,\cdots,x_n\right)$的正惯性指数为$n$，则$\forall\boldsymbol{X}\ne\boldsymbol{0}$，取$\boldsymbol{Y}=\boldsymbol{P}^{-1}\boldsymbol{X}$，则$\boldsymbol{Y}\ne\boldsymbol{0}$，所以

$$f\left(x_1,x_2,\cdots,x_n\right)\xlongequal{\boldsymbol{X}=\boldsymbol{PY}}d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2>0$$

故$f\left(x_1,x_2,\cdots,x_n\right)$正定。

定理 7.10　n元实二次型

$$f\left(x_1,x_2,\cdots,x_n\right)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$$

正定⇔$\boldsymbol{A}$的全部特征值都是正数。

证明：由于存在正交变换$\boldsymbol{X}=\boldsymbol{CY}$，使得

$$f\left(x_1,x_2,\cdots,x_n\right)\xlongequal{\boldsymbol{X}=\boldsymbol{CY}}\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2$$

其中λ_i为$\boldsymbol{A}$的特征值；由定理 10，该二次型正定的充要条件为它的正惯性指数为n，即

$\lambda_i > 0,\quad i = 1, 2, \cdots, n$。

定理 6.11　实对称矩阵$\boldsymbol{A}$正定⇔存在可逆矩阵$\boldsymbol{P}$，使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{E}$$

推论：正定矩阵的行列式大于零。

定义 7.7　设$\boldsymbol{A} = \left[a_{ij}\right]_{n\times n}$为$n$阶实对称矩阵，则顺序取$\boldsymbol{A}$的前$k$行、$k$列构成的矩阵

$$\boldsymbol{A}_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}, \qquad k = 1, 2, \cdots, n$$

称为$\boldsymbol{A}$的k阶**顺序主子阵**，它的行列式$\left|\boldsymbol{A}_k\right|$称为$\boldsymbol{A}$的$k$阶**顺序主子式**。

定理 7.12　实二次型

$$f\left(x_1, x_2, \cdots, x_n\right) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$$

正定⇔$\boldsymbol{A}$的所有顺序主子式都大于 0。

证明　先证明必要性。若$f\left(x_1, x_2, \cdots, x_n\right) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$正定，那么对于每个$k\left(1 \le k \le n\right)$，二次型$f\left(x_1, x_2, \cdots, x_k, 0, \cdots, 0\right)$是关于$x_1, x_2, \cdots, x_k$的正定二次型，该二次型$f\left(x_1, x_2, \cdots, x_k, 0, \cdots, 0\right)$的矩阵即为$\boldsymbol{A}$的$k$阶顺序主子阵，由于正定矩阵的行列式大于零，故$\boldsymbol{A}$的$k$阶顺序主子阵的行列式，即$\boldsymbol{A}$的$k$阶顺序主子式大于零。

充分性：用归纳法。当$n = 1$时命题显然成立，假设命题在$n-1$时成立，下证明命题在n时成立。由于$\boldsymbol{A}$的所有顺序主子式都大于 0，因此$\boldsymbol{A}$的$k\left(1 \le k \le n-1\right)$阶顺序主子式都大于零，因此二次型$f\left(x_1, x_2, \cdots, x_{n-1}, 0\right)$是关于变量$x_1, x_2, \cdots, x_{n-1}$的正定二次型，而其矩阵就是$\boldsymbol{A}$的$n-1$阶顺序主子阵$\boldsymbol{A}_{n-1}$是正定矩阵，故$\left|\boldsymbol{A}_{n-1}\right| > 0$，其逆矩阵$\boldsymbol{A}_{n-1}^{-1}$也是对称矩阵，并且还是正定矩阵。根据分块矩阵的初等变换知

$$\begin{bmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_{nn} \end{bmatrix} \sim \begin{bmatrix} \boldsymbol{E} & \boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_{nn} \end{bmatrix} \sim \begin{bmatrix} \boldsymbol{E} & \boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \\ \boldsymbol{0} & a_{nn} - \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \end{bmatrix}$$

根据分块初等矩阵与分块初等变换的对应关系，有

$$\begin{bmatrix} \boldsymbol{E} & \boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \\ \boldsymbol{0} & a_{nn}-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} \boldsymbol{E} & \boldsymbol{O} \\ -\boldsymbol{\alpha}^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_{n-1}^{-1} & \boldsymbol{O} \\ \boldsymbol{0} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_{nn} \end{bmatrix}$$

两边取行列式，得到

$$\begin{vmatrix} \boldsymbol{E} & \boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \\ \boldsymbol{0} & a_{nn}-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \end{vmatrix} = \begin{vmatrix} \boldsymbol{E} & \boldsymbol{O} \\ -\boldsymbol{\alpha}^{\mathrm{T}} & 1 \end{vmatrix} \begin{vmatrix} \boldsymbol{A}_{n-1}^{-1} & \boldsymbol{O} \\ \boldsymbol{0} & 1 \end{vmatrix} \begin{vmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_{nn} \end{vmatrix}$$

即

$$a_{nn}-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} = \left|\boldsymbol{A}_{n-1}^{-1}\right|\left|\boldsymbol{A}\right|$$

由于$\left|\boldsymbol{A}_{n-1}\right|^{-1} > 0, \left|\boldsymbol{A}\right| > 0$，因此有$a_{nn}-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} > 0$。

由于$\boldsymbol{A}_{n-1}$是正定矩阵，故存在矩阵可逆矩阵$\boldsymbol{C}$，使得$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}_{n-1}\boldsymbol{C} = \boldsymbol{E}$，构造矩阵

$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{C} & -\boldsymbol{A}_{n-1}^{-1}\alpha \\ & 1 \end{bmatrix}$，则

$$\begin{aligned}
\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} &= \begin{bmatrix} \boldsymbol{C}^{\mathrm{T}} & \boldsymbol{0} \\ -\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_{nn} \end{bmatrix} \begin{bmatrix} \boldsymbol{C} & -\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \\ \boldsymbol{0} & 1 \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}_{n-1} & \boldsymbol{C}^{\mathrm{T}}\boldsymbol{\alpha} \\ -\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{A}_{n-1}+\boldsymbol{\alpha}^{\mathrm{T}} & -\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}+a_{nn} \end{bmatrix} \begin{bmatrix} \boldsymbol{C} & -\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \\ \boldsymbol{0} & 1 \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}_{n-1} & \boldsymbol{C}^{\mathrm{T}}\boldsymbol{\alpha} \\ 0 & -\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}+a_{nn} \end{bmatrix} \begin{bmatrix} \boldsymbol{C} & -\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \\ \boldsymbol{0} & 1 \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}_{n-1}\boldsymbol{C} & \boldsymbol{0} \\ \boldsymbol{0} & a_{nn}-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} \boldsymbol{E} & 0 \\ \boldsymbol{0} & a_{nn}-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} \end{bmatrix}
\end{aligned}$$

由于$a_{nn}-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha} > 0$，因此$\boldsymbol{A}$的正惯性指数为$n$，故$\boldsymbol{A}$正定。

定理 7.13　n元实二次型$f(\boldsymbol{X}) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$为负定的充要条件是下列条件之一：

1、$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$的负惯性指数为$n$；

2、存在可逆矩阵$\boldsymbol{P}$，使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = -\boldsymbol{E}$；

3、$\boldsymbol{A}$的奇数阶顺序主子式都小于零，A的偶数阶顺序主子式都大于零；

利用定理 7.12 确定二次型中参数的取值范围，使得该二次型正定(负定)。

例 7.6　　求t的取值范围，使得下面二次型为正定二次型

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + 5x_3^2 + 2tx_1x_2 - 2x_1x_3 + 4x_2x_3$$

解　该二次型的矩阵为

$$\boldsymbol{A}=\begin{bmatrix}1 & t & -1\\ t & 1 & 2\\ -1 & 2 & 5\end{bmatrix}$$

由定理 6.12，该二次型正定⇔各个顺序主子式都必须大于 0。

$$|\boldsymbol{A}_1|=|1|=1>0$$

$$|\boldsymbol{A}_2|=\begin{vmatrix}1 & t\\ t & 1\end{vmatrix}=1-t^2>0 \qquad \Rightarrow -1<t<1$$

$$|\boldsymbol{A}_3|=\begin{vmatrix}1 & t & -1\\ t & 1 & 2\\ -1 & 2 & 5\end{vmatrix}=\begin{vmatrix}1 & 0 & 0\\ t & 1-t^2 & 2+t\\ -1 & 2+t & 5-1\end{vmatrix}$$

$$=4(1-t^2)-(2+t)^2=-5t^2-4t>0$$

$$\Rightarrow -\frac{4}{5}<t<0$$

求解上述三个不等式得：$-\dfrac{4}{5}<t<0$

例 7.7　已知实对称矩阵 $\boldsymbol{A}_{n\times n}$ 满足 $|\boldsymbol{A}|<0$，证明存在非零向量 $X_0\in\mathbb{R}^n$，使二次型 $f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 满足 $f(\boldsymbol{X}_0)<0$。

证明：这种题型一般都通过化为标准形来讨论。

由于

$$|\boldsymbol{A}|=\prod_{i=1}^{n}\lambda_i<0 \qquad \lambda_i \text{ 为 } \boldsymbol{A} \text{ 的特征值}$$

因此 $\boldsymbol{A}$ 至少有一个负特征值，不妨设 $\lambda_1<0$，由于存在正交矩阵 $\boldsymbol{C}$，使得

$$f(x_1,x_2,\cdots,x_n)\xlongequal{\boldsymbol{X}=\boldsymbol{CY}}\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2$$

令　$\boldsymbol{Y}=\boldsymbol{e}_1=[1,0,\cdots,0]^{\mathrm{T}}$

$$\boldsymbol{X}_0=\boldsymbol{CY}$$

则

$$f(\boldsymbol{X}_0)\xlongequal{\boldsymbol{X}_0=\boldsymbol{CY}}\lambda_1<0$$

由于 $\boldsymbol{Y}\neq\boldsymbol{0},\boldsymbol{C}$ 可逆，故 $\boldsymbol{X}_0=\boldsymbol{CY}\neq\boldsymbol{0}$。命题证毕。

习题 7.4

1. 设 $\boldsymbol{A}$ 为 n 阶实对称矩阵，且满足 $\boldsymbol{A}^3+\boldsymbol{A}^2+\boldsymbol{A}=3\boldsymbol{E}$，证明 $\boldsymbol{A}$ 是正定矩阵。

2. 设 $f(x_1,x_2,x_3)=x_1^2+2x_2^2+(1-k)x_3^2+2kx_1x_2+2x_1x_3$，若二次型 f 正定，求 k 的取值范围.

3. 求矩阵 $\boldsymbol{A}=\begin{bmatrix}1&0&1\\0&2&0\\1&0&1\end{bmatrix}$，矩阵 $\boldsymbol{B}=(k\boldsymbol{E}+\boldsymbol{A})^2$，其中 k 为实数，求对角矩阵 $\boldsymbol{\varLambda}$，使得 $\boldsymbol{B}$ 与 $\boldsymbol{\varLambda}$ 相似，并求 k 为何止时，$\boldsymbol{B}$ 为正定矩阵。

4. 设 5 阶实对称矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2-5\boldsymbol{A}+6\boldsymbol{E}=\boldsymbol{O}$，$r(\boldsymbol{A}-2\boldsymbol{E})=2$。

（1）求 $|\boldsymbol{A}-\boldsymbol{E}|$；（2）判断 $\boldsymbol{A}$ 是否为正定矩阵？证明你的结论。

5. 设有 n 元实二次型

$$f(x_1,x_2,\cdots,x_n)=(x_1+a_1x_2)^2+(x_2+a_2x_3)^2+\cdots+(x_{n-1}+a_{n-1}x_n)^2+(x_n+a_nx_1)^2$$

其中 $a_i\,(i=1,2,\cdots,n)$ 为实数。试问当 $a_i\,(i=1,2,\cdots,n)$ 满足何种条件时，二次型 $f(x_1,x_2,\cdots,x_n)$ 为正定二次型。

6. 设 $\boldsymbol{M}=\begin{bmatrix}\boldsymbol{A}&\boldsymbol{B}\\\boldsymbol{B}^{\mathrm{T}}&\boldsymbol{D}\end{bmatrix}$ 是 n 阶正定矩阵，其中 $\boldsymbol{A}$ 是 $r(r<n)$ 阶方阵，证明：$\boldsymbol{A},\boldsymbol{D},\boldsymbol{D}-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B}$ 都是正定矩阵。

7. 实对称矩阵 $\boldsymbol{A}$ 正定的充要条件是：有实的上三角矩阵 $\boldsymbol{B}$，且 $\boldsymbol{B}$ 的主对角元全大于零，使得 $\boldsymbol{A}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$。

8. n 阶实对称矩阵 $\boldsymbol{A}$ 正定的充要条件是：有 $m\times n$ 的列满秩实矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{A}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$。

9. 分别证明下列各题

（1）设 $\boldsymbol{A}$ 为 n 阶实对称矩阵，证明：$\boldsymbol{A}$ 正定的充要条件是存在可逆矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{A}=\boldsymbol{B}^2$；

（2）设 $\boldsymbol{A}$ 为 n 阶实对称矩阵，证明：$\boldsymbol{A}$ 正定的充要条件是存在可逆矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{A}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$。

（3）n 阶实对称矩阵 $\boldsymbol{A}$ 正定的充要条件是：有 $m\times n$ 的列满秩实矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{A}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$。

10. 已知 n 阶方阵 $\boldsymbol{A}$ 正定，$\boldsymbol{X}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$，证明：二次型

$$f\left(x_1, x_2, \cdots, x_n\right) = \det\begin{bmatrix} \boldsymbol{A} & \boldsymbol{X} \\ \boldsymbol{X}^{\mathrm{T}} & 0 \end{bmatrix}$$负定。

11. 设$\boldsymbol{A}$为m阶实对称矩阵且正定，$\boldsymbol{B}$为$m\times n$实矩阵，试证$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{AB}$为正定矩阵的充要条件是$r\left(\boldsymbol{B}\right)=n$。

12. 对于任一实可逆矩阵$\boldsymbol{A}$，存在一个正交矩阵$\boldsymbol{T}$与两个正定矩阵$\boldsymbol{S}_1, \boldsymbol{S}_2$，使得$\boldsymbol{A}=\boldsymbol{TS}_1=\boldsymbol{S}_2\boldsymbol{T}$，并且这种分解是唯一的。

13. 设n阶方阵$\boldsymbol{A}$可逆，则$\boldsymbol{A}$可表示为$\boldsymbol{A}=\boldsymbol{PB}$，其中$\boldsymbol{P}$为正交矩阵，$\boldsymbol{B}$是可逆对称矩阵。

14. 对于任一n阶实可逆矩阵$\boldsymbol{A}$，都存在正交矩阵$\boldsymbol{T}_1, \boldsymbol{T}_2$，使得

$$\boldsymbol{A}=\boldsymbol{T}_1\boldsymbol{diag}\left\{\lambda_1, \lambda_2, \cdots, \lambda_{\boldsymbol{n}}\right\}\boldsymbol{T}_2$$

并且$\lambda_1^2, \lambda_2^2, \cdots, \lambda_n^2$是$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$的特征值。

15. 设$\boldsymbol{A}$是n阶实对称矩阵，$\boldsymbol{B}$是n阶正定矩阵。证明：存在可逆矩阵$\boldsymbol{C}$，使得$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$与$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{BC}$都为对角矩阵。

第八章　线性空间与线性变换

§8.1　线性空间的概念与例子

高级数学研究的一个主要内容是映射。就一般的概念而言，映射有三个要素：定义域、值域、映射法则。对于一般的映射，其定义域可以是任意定义的。但是，如果映射的定义域具有良好的性质，那么该映射也就更容易具有好的性质，即使映射法则相同，如果定义域不同，则映射也具有不同的性质，例如，连续函数在闭区间上有界，但连续函数在开区间上就没有这个结论。这一章将要介绍的线性空间是映射的最典型的定义域。

对于向量，我们可以定义其线性运算：向量加法、数乘。实际上，我们可以在许多不同的集合上定义“类似”的结构。在此首先我们将给出线性空间的定义，然后给出它的一些基本性质。

定义 8.1　**线性空间**

设V是一个非空集合，$\mathbb{F}$为一个数域（实数域或复数域，可认为$\mathbb{F}$就是实数域$\mathbb{R}$）．在V中定义了两种运算，

（1）加法：$\forall \boldsymbol{\alpha},\boldsymbol{\beta}\in V$，存在唯一的元素$\boldsymbol{\gamma}\in V$与它们之对应，$\boldsymbol{\gamma}$称为$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的和，记作$\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{\beta}$；

（2）数乘：$\forall \boldsymbol{\alpha}\in V, k\in\mathbb{F}$，存在唯一的元素$\boldsymbol{\delta}\in V$与之对应，称为$k$与$\boldsymbol{\alpha}$的数乘，记作$\boldsymbol{\delta}=k\boldsymbol{\alpha}$；

如果这两个运算满足如下性质：（其中$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in V; k,l\in\mathbb{F}$）

（1）加法交换律：$\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\beta}+\boldsymbol{\alpha}$；

（2）加法结合律：$(\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\gamma}=\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma})$；

（3）V中存在零元素：存在$\boldsymbol{\alpha}_0\in V$，使得$\forall\boldsymbol{\alpha}\in V$，$\boldsymbol{\alpha}+\boldsymbol{\alpha}_0=\boldsymbol{\alpha}$成立。我们称$\boldsymbol{\alpha}_0$为$\boldsymbol{V}$的**零元素**；简记作$\boldsymbol{\theta}$；

（4）负元素存在：$\forall\boldsymbol{\alpha}\in V$，存在$\boldsymbol{\beta}\in V$，使得$\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\theta}$，称$\boldsymbol{\beta}$为$\boldsymbol{\alpha}$的负元素，记做$\boldsymbol{\beta}=-\boldsymbol{\alpha}$；

（5）数域$\mathbb{F}$中存在单位元：存在数$k_1\in\mathbb{F}$，使得$\forall\boldsymbol{\alpha}\in V$，有$k_1\boldsymbol{\alpha}=\boldsymbol{\alpha}$；

（6）数乘结合律：$k(l\boldsymbol{\alpha})=(kl)\boldsymbol{\alpha}$；

（7）　分配律：$(k+l)\boldsymbol{\alpha}=k\boldsymbol{\alpha}+l\boldsymbol{\alpha}$；

（8）　分配律：$k(\boldsymbol{\alpha}+\boldsymbol{\beta})=k\boldsymbol{\alpha}+k\boldsymbol{\beta}$；

则称V为$\mathbb{F}$上的**线性空间**，简称V为**线性空间**。线性空间中的元素称为**向量**。线性空间中任意两个向量的和属于该线性空间，该性质称为线性空间对（向量）加法运算封闭；任意数与线性空间中向量的数乘也属于该线性空间，该性质称为线性空间对数乘运算封闭。

注意：线性空间中的两个运算我们一般叫做加法与数乘。但是对于不同的线性空间，其运算有各自的意义，一定要注意向量空间中的加法、数乘的具体定义。

例 8.1　线性空间$\mathbb{R}^n$

考虑由所有n维列向量所构成的集合

$$\mathbb{R}^n=\left\{[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\,\middle|\,x_1,x_2,\cdots,x_n\in\mathbb{R}\right\}$$

对于向量$[x_1,x_2,\cdots,x_n]^{\mathrm{T}},[y_1,y_2,\cdots,y_n]^{\mathrm{T}}\in\mathbb{R}^n$，定义其加法为

$$[x_1,x_2,\cdots,x_n]^{\mathrm{T}}+[y_1,y_2,\cdots,y_n]^{\mathrm{T}}=[x_1+y_1,x_2+y_2,\cdots,x_n+y_n]^{\mathrm{T}}$$

定义数$k\in\mathbb{R}$与$[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\in\mathbb{R}^n$的数乘为

$$k[x_1,x_2,\cdots,x_n]^{\mathrm{T}}=[kx_1,kx_2,\cdots,kx_n]^{\mathrm{T}}$$

容易验证此时$\mathbb{R}^n$在这两种运算下为一个线性空间。该线性空间的零元素为零向量$[0,0,\cdots,0]^{\mathrm{T}}$，而元素$[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$的负元素为$[-x_1,-x_2,\cdots,-x_n]^{\mathrm{T}}$。

例 8.2　$\mathbb{R}^n$中的集合

$$\left\{[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\in\mathbb{R}^n\,\middle|\,x_1+x_2+\cdots+x_n=0\right\}$$

在通常的向量加法、数乘运算下是一个线性空间。

例 8.3　多项式所构成的线性空间

由所有次数不超过n的多项式构成的集合

$$\mathbf{P}_n[x]=\left\{a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\,\middle|\,a_n,a_{n-1},\cdots,a_1,a_0\in\mathbb{F}\right\}$$

$\mathbf{P}_n[x]$中的定义加法为

$$\begin{aligned}&\left(a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\right)+\left(b_nx^n+b_{n-1}x^{n-1}+\cdots+b_1x+b_0\right)\\&=(a_n+b_n)x^n+(a_{n-1}+b_{n-1})x^{n-1}+\cdots+(a_1+b_1)x+(a_0+b_0)\end{aligned}$$

$\mathbf{P}_n[x]$中的数乘为

$$\begin{aligned}&k\left(a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\right)\\&=\left(ka_n\right)x^n+\left(ka_{n-1}\right)x^{n-1}+\cdots+\left(ka_1\right)x+\left(ka_0\right)\end{aligned}$$

显然，$\mathbf{P}_n[x]$对这两种运算封闭。容易验证$\mathbf{P}_n[x]$中的这两种运算满足线性空间定义中所要求的 8 条性质，因此$\mathbf{P}_n[x]$为一个线性空间。其零元素为零多项式$0x^n+0x^{n-1}+\cdots+0x+0$，而$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$的负元素为$-a_nx^n-a_{n-1}x^{n-1}-\cdots-a_1x-a_0$。

例 8.4　定义于$\mathbb{R}$上的实函数集

$$V=\left\{f\,\middle|\,f:\mathbb{R}\to\mathbb{R}\right\}$$

在运算

$$\left(f+g\right)\left(x\right)=f\left(x\right)+g\left(x\right),\qquad \left(kf\right)\left(x\right)=kf\left(x\right)$$

下是一个线性空间。该线性空间的零元素为零函数$\theta\left(x\right)=0$；而f的负元素为$-f$，$-f$定义为$\left(-f\right)\left(x\right)=-f\left(x\right)$。

例 8.5　在$\left[a,b\right]$上k次连续可微的函数所组成的集合

$$\left\{f\,\middle|\,f:\mathbb{R}\to\mathbb{R}\text{且}f\in C^k\left[a,b\right]\right\}$$

在

$$\left(f+g\right)\left(x\right)=f\left(x\right)+g\left(x\right),\qquad \left(kf\right)\left(x\right)=kf\left(x\right)$$

下是一个线性空间。

例 8.6　数域$\mathbb{F}$上的所有$m\times n$矩阵所构成的集合

$$\mathbb{F}^{m\times n}=\left\{\boldsymbol{A}=\left[a_{ij}\right]_{m\times n}\middle|a_{ij}\in\mathbb{F},i=1,2,\cdots,m;j=1,2,\cdots,n\right\}$$

在矩阵加法、数乘运算下是线性空间。

下面这个例子就与我们通常的认识有所不同。

例 8.7　正实数的全体，记作$\mathbb{R}^+$，在其中定义加法及乘数运算为

$$a\oplus b=ab,\quad k\circ a=a^k,\left(k\in\mathbb{R},a,b\in\mathbb{R}^+\right)$$

$\mathbb{R}^+$在上述两种运算下构成一个线性空间。

证明：　$\forall a,b\in\mathbb{R}^{+}\Rightarrow a\oplus b=ab\in\mathbb{R}^{+}$，

$\forall k\in\mathbb{R},a\in\mathbb{R}^{+}\Rightarrow k\circ a=a^{k}\in\mathbb{R}^{+}$，

所以$\mathbb{R}^{+}$对上述定义的加法与数乘运算封闭。下面逐一验证八条线性运算规律

（1）　$a\oplus b=ab=ba=b\oplus a$；

（2）　$(a\oplus b)\oplus c=(ab)\oplus c=(ab)c=a(bc)=a\oplus(b\oplus c)$；

（3）　数 1 是$\mathbb{R}^{+}$中的零元素：$\forall a\in\mathbb{R}^{+}$，有$a\oplus 1=a\cdot 1=a$；

（4）　$\forall\alpha\in\mathbb{R}^{+}$，其负元素为$a^{-1}\in\mathbb{R}^{+}$，s. t.　$a\oplus a^{-1}=a\cdot a^{-1}=1$；

（5）　$1\circ a=a^{1}=a$，因此 1 是$\mathbb{R}$上的单位元；

（6）　$k\circ(l\circ a)=k\circ a^{l}=\left(a^{l}\right)^{k}=a^{kl}=(kl)\circ a$；

（7）　$(k+l)\circ a=a^{k+l}=a^{k}a^{k}=a^{k}\oplus a^{l}=(k\circ a)\oplus(l\circ a)$；

（8）　$k\circ(a\oplus b)=k\circ(ab)=(ab)^{k}=a^{k}b^{k}=a^{k}\oplus b^{k}=(k\circ a)\oplus(k\circ b)$

我们将这个线性空间记作$\mathbb{R}^{+}$。

定义 8.2　　**向量组，线性组合**

设V为一线性空间，线性空间中若干个向量组成的集合称为**向量组**；设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in V$为一个向量组，$\boldsymbol{\beta}\in V$，若存在$k_1,k_2,\cdots,k_r\in\mathbb{F}$，使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r$$

则称$\boldsymbol{\beta}$为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$的（一个）**线性组合**，或称$\boldsymbol{\beta}$可由向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$**线性表示**，$k_1,k_2,\cdots,k_r$为$\boldsymbol{\beta}$在该向量组下的**组合系数**。

借用矩阵乘法的定义，$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r$“形式”上可表示成

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}\begin{bmatrix}k_1\\k_2\\\vdots\\k_r\end{bmatrix}$$

即把该线性组合写成矩阵乘法的形式。但应该注意的是：由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$不再是$\mathbb{R}^{n}$中的向量，因此$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r\end{bmatrix}$本质上不再是矩阵，而仅仅是一个记号。

例 8. 8　设 V 为一线性空间，而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r \in V$，则集合

$$L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}=span\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}=\{c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_r\boldsymbol{\alpha}_r | c_1,c_2,\cdots,c_r\in\mathbb{F}\}$$

为一个线性空间，称为由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 生成的线性空间。这个空间由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 的所有线性组合构成。

例8. 9　设 S 是线性空间 V 的子集，那么

$$[S]=\{k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r | k_1,k_2,\cdots,k_r\in\mathbb{F},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in S,r\in\mathbb{Z}^+\}$$

是一个线性空间，称为由 $\boldsymbol{S}$ 所生成的线性空间。$\boldsymbol{S}$ 由这样的向量 $\boldsymbol{\beta}$ 构成：$\boldsymbol{\beta}$ 可表示为 $\boldsymbol{S}$ 中某个向量组的线性组合。

线性空间的基本性质

（1）　零元素是唯一的；

（2）　负元素是唯一的；

（3）　满足消去律：如果 $\boldsymbol{\alpha}+\boldsymbol{\gamma}=\boldsymbol{\beta}+\boldsymbol{\gamma}$，则 $\boldsymbol{\alpha}=\boldsymbol{\beta}$；

（4）　$0\boldsymbol{\alpha}=\boldsymbol{\theta}$;　$(-1)\boldsymbol{\alpha}=-\boldsymbol{\alpha}$;　$k\boldsymbol{\theta}=\boldsymbol{\theta}$；

（5）　如果 $k\boldsymbol{\alpha}=\boldsymbol{\theta}$，则 $k=0$ 或 $\boldsymbol{\alpha}=\boldsymbol{\theta}$；

（6）　1 为数域 $\mathbb{F}$ 中的单位元，即 $1\boldsymbol{\alpha}=\boldsymbol{\alpha}$；

证明（1）设线性空间 $\boldsymbol{V}$ 中有两个零元素，分别为 $\boldsymbol{\theta}_1,\boldsymbol{\theta}_2$，那么根据零元素的定义，$\forall\alpha\in\boldsymbol{V}$，$\boldsymbol{\theta}+\boldsymbol{\alpha}=\boldsymbol{\alpha}$ 成立。因此在 $\boldsymbol{\theta}_1+\boldsymbol{\theta}_2$ 中，由于 $\boldsymbol{\theta}_1$ 为零元素，因此 $\boldsymbol{\theta}_1+\boldsymbol{\theta}_2=\boldsymbol{\theta}_2$；同样由于 $\boldsymbol{\theta}_2$ 为零元素，因此 $\boldsymbol{\theta}_1+\boldsymbol{\theta}_2=\boldsymbol{\theta}_1$，故 $\boldsymbol{\theta}_1=\boldsymbol{\theta}_1+\boldsymbol{\theta}_2=\boldsymbol{\theta}_2$，故 $\boldsymbol{\theta}_1=\boldsymbol{\theta}_2$。故零元素唯一。

（2）设 $\boldsymbol{\alpha}\in V$，$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\in V$ 为 $\boldsymbol{\alpha}$ 的两个负元素，那么根据定义

$$\boldsymbol{\alpha}+\boldsymbol{\beta}_1=\boldsymbol{\alpha}+\boldsymbol{\beta}_2=\boldsymbol{\theta}$$

因此

$$\boldsymbol{\beta}_1=\boldsymbol{\beta}_1+\boldsymbol{\theta}=\boldsymbol{\beta}_1+(\boldsymbol{\alpha}+\boldsymbol{\beta}_2)=\boldsymbol{\beta}_1+\boldsymbol{\alpha}+\boldsymbol{\beta}_2=(\boldsymbol{\beta}_1+\boldsymbol{\alpha})+\boldsymbol{\beta}_2=\boldsymbol{\theta}+\boldsymbol{\beta}_2=\boldsymbol{\beta}_2$$

故命题成立。

由于元素的负元素唯一，我们将 $\boldsymbol{\alpha}\in V$ 的负元素记为 $-\boldsymbol{\alpha}$；

（3）由于 $\boldsymbol{\alpha}+\boldsymbol{\gamma}=\boldsymbol{\beta}+\boldsymbol{\gamma}$，在上式两边同时加上 $\boldsymbol{\gamma}$ 的负元素 $-\boldsymbol{\gamma}$，得到

$$\boldsymbol{\alpha}+\boldsymbol{\gamma}+(-\boldsymbol{\gamma})=\boldsymbol{\beta}+\boldsymbol{\gamma}+(-\boldsymbol{\gamma})$$

因此　$\boldsymbol{\alpha}+(\boldsymbol{\gamma}+(-\boldsymbol{\gamma}))=\boldsymbol{\beta}+(\boldsymbol{\gamma}+(-\boldsymbol{\gamma}))$，即

$$\boldsymbol{\alpha}+\boldsymbol{\theta}=\boldsymbol{\beta}+\boldsymbol{\theta}$$

根据零元素的定义即得

$$\boldsymbol{\alpha}=\boldsymbol{\beta}$$

（4）$k\boldsymbol{\alpha}+\boldsymbol{\theta}=k\boldsymbol{\alpha}=(k+0)\boldsymbol{\alpha}=k\boldsymbol{\alpha}+0\boldsymbol{\alpha}$，故$0\boldsymbol{\alpha}=\boldsymbol{\theta}$。

$$\boldsymbol{\alpha}+(-\boldsymbol{\alpha})=\boldsymbol{\theta}=0\boldsymbol{\alpha}=(1+(-1))=1\boldsymbol{\alpha}+(-1)\boldsymbol{\alpha}=\boldsymbol{\alpha}+(-1)\boldsymbol{\alpha};$$

依据消去律得：$(-1)\boldsymbol{\alpha}=-\boldsymbol{\alpha}$。

由　$k\boldsymbol{\alpha}+\boldsymbol{\theta}=k\boldsymbol{\alpha}=k(\boldsymbol{\alpha}+\boldsymbol{\theta})=k\boldsymbol{\alpha}+k\boldsymbol{\theta}$

依据消去律得　$k\boldsymbol{\theta}=\boldsymbol{\theta}$。

（5）如果$k\neq 0$，那么在$k\boldsymbol{\alpha}=\boldsymbol{\theta}$两边同时乘以$k^{-1}$，得到

$$k^{-1}(k\boldsymbol{\alpha})=k^{-1}\boldsymbol{\theta}=\boldsymbol{\theta}$$

得到

$$(k^{-1}k)\boldsymbol{\alpha}=1\boldsymbol{\alpha}=\boldsymbol{\alpha}=\boldsymbol{\theta}$$

(6) 由于$\mathbb{F}$中存在单位元k_1，使得$\forall\boldsymbol{\alpha}\in V$，有$k_1\boldsymbol{\alpha}=\boldsymbol{\alpha}$；因此

$$\boldsymbol{\alpha}=k_1\boldsymbol{\alpha}=(1\times k_1)\boldsymbol{\alpha}=1(k_1\boldsymbol{\alpha})=1\boldsymbol{\alpha}$$

但请注意：单位元未必是唯一的。对于线性空间$V=\{\boldsymbol{\theta}\}$，数域$\mathbb{F}$中的任意一个数都是单位元；但若$V\neq\{\boldsymbol{\theta}\}$，则在$V$中至少存在非零元$\boldsymbol{\alpha}$。因此，若$k_1$为单位元，

$$k_1\boldsymbol{\alpha}=\boldsymbol{\alpha}=1\boldsymbol{\alpha}$$

$$(k_1-1)\boldsymbol{\alpha}=\boldsymbol{\theta}$$

由于$\boldsymbol{\alpha}\neq\boldsymbol{\theta}$，根据(5)的结论可知$k_1-1=0$，即$k_1=1$。因此，若线性空间$V\neq\{\boldsymbol{\theta}\}$，则数域$\mathbb{F}$中的单位元是唯一的，这个数一定是 1。

定义 8.3　**子空间**

设V为一线性空间，$W\subset V$且$W\neq\phi$，若W的所有元素关于V中的加法与数乘运算也构成一个线性空间，则称W为V的子空间。

例 8.10　任何线性空间 V 都有两个子空间：V 与 $\{\boldsymbol{\theta}\}$。

例 8.11　$\left\{[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\in\mathbb{R}^n \middle| x_1+x_2+\cdots+x_n=0\right\}$ 为 $\mathbb{R}^n$ 的子空间。

例 8.12　当 $m>n$ 时，$P_n[x]$ 为 $P_m[x]$ 的子空间。

例 8.13　线性空间 $\mathbb{R}^+$ 不是线性空间 $\mathbb{R}$ 的子空间。

虽然 $\mathbb{R}^+$ 是 $\mathbb{R}$ 的子集，但是这两个空间中加法、数乘的定义不相同。

定理 8.1　W 为线性空间 V 的非空子集，在 V 的加法、数乘运算下，如下命题是等价的

（1） W 为 V 的子空间；

（2） W 对加法、数乘运算封闭：

a)　若 $\boldsymbol{\alpha},\boldsymbol{\beta}\in W$，$\boldsymbol{\alpha}+\boldsymbol{\beta}\in W$；

b)　若 $\boldsymbol{\alpha}\in W, k\in\mathbb{F}$，则 $k\boldsymbol{\alpha}\in W$；

（3） W 对其中任意两个向量的线性运算是封闭的：$\forall\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in W, k_1,k_2\in\mathbb{F}$，有 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\in W$；

（4） W 对其中任意向量组的线性运算是封闭的：$\forall\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in W, k_1,k_2,\cdots,k_r\in\mathbb{F}$，有 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_r\boldsymbol{\alpha}_r\in W$；

证明：根据线性空间的定义，$(1)\Rightarrow(2)$ 是显然的。我们现证明 $(2)\Rightarrow(1)$，我们只需要证明 W 满足线性空间定义中的八条性质。这里以加法结合律为例。设 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in W$，由于 $W\subset V$，所以 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in V$，且由于 V 是一个线性空间，所以在 V 中有

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

由于 W 对 V 中的向量加法、数乘封闭，所以

$$(\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\gamma}=\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma})\in W$$

这说明 W 对 $\boldsymbol{V}$ 中的向量加法满足结合律。其它的性质可类似地证明：首先在 V 中该性质成立，然后由 W 对向量加法、数乘封闭，所以该性质在 W 中成立。在此省略。

$(2)\Rightarrow(3)$：由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in W, k_1,k_2\in\mathbb{F}$，由于 W 对数乘封闭，所以 $k_1\boldsymbol{\alpha}_1,k_2\boldsymbol{\alpha}_2\in W$，再利

用W对向量加法封闭，所以$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\in W$；

(3)⇒(4)：利用归纳法。显然，当$r=1$时命题成立，假设当$r=m-1$时命题成立，下证$r=m$时命题也成立。由归纳假设，$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{m-1}\boldsymbol{\alpha}_{m-1}\in W$，由于$\boldsymbol{\alpha}_m\in V$，因此

$$1\times\left(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{m-1}\boldsymbol{\alpha}_{m-1}\right)+k_m\boldsymbol{\alpha}_m\in V$$

所以命题在$r=m$时也成立。

(4)⇒(2)：这是显然的。

线性空间中任意集合所生成的线性空间都是该空间的子空间。

例 8.14　在线性空间$\mathbb{R}^{n\times n}$中，判断集合

（1）$W_1=\left\{\boldsymbol{A}\in\mathbb{R}^{n\times n}\middle|\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}\right\}$；

（2）$W_2=\left\{\boldsymbol{B}\in\mathbb{R}^{n\times n}\middle||\boldsymbol{B}|\neq 0\right\}$

是否为子空间。

解：根据以上定理容易验证：W_1为子空间，而W_2不为子空间。

定理 8.2　　设W_1,W_2是线性空间V的子空间，则有

（1）　$W_1\cap W_2=\left\{\boldsymbol{\alpha}\middle|\boldsymbol{\alpha}\in W_1\text{且}\boldsymbol{\alpha}\in W_2\right\}$是$V$的子空间，称为$W_1$与$W_2$的**交空间**；

（2）　$W_1+W_2=\left\{\boldsymbol{\alpha}\middle|\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\ \boldsymbol{\alpha}_1\in W_1,\ \boldsymbol{\alpha}_2\in W_2\right\}$是$V$的子空间，称为$\boldsymbol{W}_1$与$\boldsymbol{W}_2$的**和空间**；

证明：利用定理 8.1 证明即可。

（1）设$\boldsymbol{\alpha},\boldsymbol{\beta}\in W_1\cap W_2$，故$\boldsymbol{\alpha},\boldsymbol{\beta}\in W_1$，由于$W_1$是线性空间$V$的子空间，因此$\boldsymbol{\alpha}+\boldsymbol{\beta}\in W_1$；类似地，$\boldsymbol{\alpha}+\boldsymbol{\beta}\in W_2$。因此，$\boldsymbol{\alpha}+\boldsymbol{\beta}\in W_1\cap W_2$。

设$\boldsymbol{\alpha}\in W_1,k\in F$，则由$W_1$是线性空间$V$的子空间，因此$k\boldsymbol{\alpha}\in W_1$；类似地，$k\boldsymbol{\alpha}\in W_2$。因此，$k\boldsymbol{\alpha}\in W_1\cap W_2$。故$W_1\cap W_2$为$V$的子空间。

（2）证明方法与步骤类似于（1），故省略。

例 8.15　设

$$\boldsymbol{\alpha}_1=\left[1,2,1,0\right]^{\mathrm{T}},\boldsymbol{\alpha}_2=\left[-1,1,1,1\right]^{\mathrm{T}},$$
$$\boldsymbol{\beta}_1=\left[2,-1,0,1\right]^{\mathrm{T}},\boldsymbol{\beta}_2=\left[1,-1,3,7\right]^{\mathrm{T}},$$
$$W_1=span\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\right\},\qquad W_2=span\left\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\right\}$$

求$W_1\bigcap W_2,W_1+W_2$。

解：设$\boldsymbol{\gamma}\in W_1\bigcap W_2$，因此存在$k_1,k_2,-k_3,-k_4$，使得

$$\boldsymbol{\gamma}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2=k_3\boldsymbol{\beta}_1+k_4\boldsymbol{\beta}_2$$

所以

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\beta}_1+k_4\boldsymbol{\beta}_2=\mathbf{0}$$

所以

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

所以

$$\begin{cases}k_1=k_4\\k_2=-4k_4\\k_3=-3k_4\\k_4=k_4\end{cases}$$

所以

$$\boldsymbol{\gamma}=k_4\boldsymbol{\alpha}_1-4k_4\boldsymbol{\alpha}_2=k_4\left(\boldsymbol{\alpha}_1-4\boldsymbol{\alpha}_2\right)$$
$$=k_4\left[5,-2,-3,-4\right]^{\mathrm{T}}$$

显然，$W_1+W_2=span\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\right\}$，所以$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$的一个极大无关组就是$\boldsymbol{W}_1+\boldsymbol{W}_2$的一组基，根据前面的计算结果知

$$W_1+W_2=span\left\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1\right\}$$

习题 8.1

1. 判断下列命题是否成立。
   a） n阶对称矩阵关于矩阵的加法、数乘运算构成线性空间；
   b） $W=\left\{\boldsymbol{A}\middle|\boldsymbol{A}\in\mathbb{F}^{n\times n},\left|\boldsymbol{A}\right|=0\right\}$是$\mathbb{F}^{n\times n}$的子空间。
   c） $V=\left\{\left[a+bi,c+di\right]^{\mathrm{T}}\middle|a,b,c,d\in\mathbb{R}\right\}$为$\mathbb{C}^2$的子空间。

d） $\left\{ \boldsymbol{A} \in \mathbb{F}^{n\times n} \middle| tr\left(\boldsymbol{A}\right)=0 \right\}$ 是 $\mathbb{F}^{n\times n}$ 的子空间。

e） 设 $S=\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$，定义 S 上的加法 $\oplus$ 、数乘 $\circ$ 如下：

$$x \oplus y = arctg\left(tgx + tgy\right)$$
$$k \circ x = arctg\left(k\ tgx\right)$$

问：S 在上述两种运算下是否为线性空间。

2. V 是定义在实数域上的函数构成的线性空间，令

$$W_1=\left\{ f\left(x\right) \middle| f\left(x\right)\in V, f\left(x\right)=f\left(-x\right) \right\}$$
$$W_2=\left\{ f\left(x\right) \middle| f\left(x\right)\in V, f\left(x\right)=-f\left(-x\right) \right\}$$

证明：W_1,W_2 都是 V 的子空间

3. 设 W,W_1,W_2 都是向量空间 V 的子空间，其中 $W_1 \subseteq W_2$ 且 $W \cap W_1 = W \cap W_2$，$W+W_1=W+W_2$。证明：$W_1=W_2$。

§8.2　线性空间的基与维数

在线性空间中，可类似于向量空间 $\mathbb{R}^n$ 的情形，可定义线性相关、线性无关、极大线性无关组、向量组等价、向量组的秩等概念，并且有类似的结论。这里不再赘述，而只是引用这些概念与结果。

定义 8.4　基、维数

设 V 是线性空间，若向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d \in V$ 满足

（1） $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$ 线性无关；

（2） V 中的任一向量都可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$ 线性表示；

则称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d$ 为线性空间 V 的一组**基**，d 为线性空间 V 的维数，记为 $\dim V = d$，并称 V 为 d 维线性空间，V 称为有限维线性空间，我们在本课程中只考虑有限维线性空间。零空间 $\{\boldsymbol{\theta}\}$ 的维数规定为 0。

定理 8.3　　d 维线性空间 V 中任意 d 个线性无关的向量都是 d 的基。

证明：证明过程与$\mathbb{R}^n$中的情形类似，请参见第三章。

例 8.16　设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 为一向量组，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的极大无关组为线性空间 $span\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$ 的一组基，且 $\dim V = r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$。

证明：设 $r\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}=r$，且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的极大无关组，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关；且 $\forall\boldsymbol{\beta}\in V = L\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$，$\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表示，由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 等价，故由线性表示的传递性可知，β 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示。根据定义可知结论成立。

例 8.17　对于矩阵 $\boldsymbol{A}_{m\times n}$，求零空间 $N(\boldsymbol{A})$ 与列空间 $R(\boldsymbol{A})$ 的基与维数。

解：$\dim N(\boldsymbol{A}) = n - r(\boldsymbol{A})$，而齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的基础解系为 $N(\boldsymbol{A})$ 的一组基；$R(\boldsymbol{A})$ 的维数等于 $\boldsymbol{A}$ 的列向量组的秩，因此 $\dim R(\boldsymbol{A}) = r(\boldsymbol{A})$，且 $\boldsymbol{A}$ 的列向量组的极大无关组为 $R(\boldsymbol{A})$ 的一组基。

例 8.18　求矩阵空间 $\mathbb{R}^{m\times n}$ 的基与维数。

解：　考虑矩阵序列 $\boldsymbol{A}_{kl}, k=1,\cdots,m; l=1,\cdots,n$，其元素为

$$[\boldsymbol{A}_{kl}]_{ij} = \delta_{ki}\delta_{lj} = \begin{cases} 1 & i=k,\quad j=l \\ 0 & \text{其它} \end{cases}$$

因此，对于 $c_{kl}, k=1,\cdots,m; l=1,\cdots,n$，则若令

$$\boldsymbol{B} = \sum_{k=1}^{m}\sum_{l=1}^{n} c_{kl}\boldsymbol{A}_{kl} = [c_{kl}] = 0$$

则得到 $c_{kl}=0\,(k=1,\cdots,m; l=1,\cdots,n)$，因此 $\boldsymbol{A}_{kl}, k=1,\cdots,m; l=1,\cdots,n$ 线性无关。

而 $\forall\boldsymbol{B}=[b_{ij}]\in\mathbb{R}^{m\times n}$，则有 $\boldsymbol{B}=\sum_{k=1}^{m}\sum_{l=1}^{n} b_{kl}\boldsymbol{A}_{kl}$，因此 $\boldsymbol{B}$ 可由 $\boldsymbol{A}_{kl}, k=1,\cdots,m; l=1,\cdots,n$ 线性表示。根据定义，$\dim\mathbb{R}^{m\times n} = m\times n$，$\boldsymbol{A}_{kl}, k=1,\cdots,m; l=1,\cdots,n$ 为 $\mathbb{R}^{m\times n}$ 的一组基。

定义 8.5　**坐标**

若 $B=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$ 为 d 维线性空间 V 的一组基，则 $\forall\boldsymbol{\beta}\in V$，存在唯一的一组数 $x_1,x_2,\cdots,x_d$，使得

$$\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_d\boldsymbol{\alpha}_d=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_d\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_d\end{bmatrix}$$

则称 $[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$ 为 $\boldsymbol{\beta}$ 在基 $B=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$ 下的坐标。

基变换与坐标变换

线性空间中，基是不唯一的，在实际应用中，我们可根据需要用某个指标来衡量各组基的优劣，从而选择最适合需要的基。因此，这就可能涉及到基变换的问题。

设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$、$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\}$ 是 d 维线性空间 V_d 的两组基，因此它们等价，所以存在方阵 $\boldsymbol{C}$，使得

$$\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_d\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_d\end{bmatrix}\boldsymbol{C}_{d\times d}$$

定义 8.6　**过渡矩阵**

设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$、$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\}$ 是 d 维线性空间 V_d 的两组基，必有方阵矩阵 $\boldsymbol{C}_{\boldsymbol{d\times d}}$，使得

$$\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_d\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_d\end{bmatrix}\boldsymbol{C}_{d\times d}$$

称 $\boldsymbol{C}_{\boldsymbol{d\times d}}$ 为从基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$ 到 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\}$ 的过渡矩阵（基变换矩阵）。

定理 8.4　过渡矩阵是可逆的。
证明：证明方法参见第三章。

定理 8.5　设线性空间 $\boldsymbol{V_d}$ 一组基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_d\}$ 到另一组基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_d\}$ 的过渡矩阵为 $\boldsymbol{C}$，而向量 $\gamma\in\boldsymbol{V}$ 在这两组基下的坐标分别为 $\boldsymbol{X}$、$\boldsymbol{Y}$，则有

$$\boldsymbol{X}=\boldsymbol{CY}$$

证明：参见第三章。

实内积空间

定义 8.7　设 V 是实数域 $\mathbb{R}$ 上的线性空间，如果存在一个映射 f，

$$f: V\times V\to\mathbb{R}$$

我们将该映射记作

$$f(\boldsymbol{\alpha},\boldsymbol{\beta})\triangleq(\boldsymbol{\alpha},\boldsymbol{\beta}),\qquad \boldsymbol{\alpha},\boldsymbol{\beta}\in V$$

它满足：

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

(2) 线性性：$(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta})=(\boldsymbol{\alpha}_1,\boldsymbol{\beta})+(\boldsymbol{\alpha}_2,\boldsymbol{\beta})$，

$$(k\boldsymbol{\alpha},\boldsymbol{\beta})=k(\boldsymbol{\alpha},\boldsymbol{\beta}),\qquad k\in\mathbb{R};$$

(3) 非负性：$(\boldsymbol{\alpha},\boldsymbol{\alpha})\ge 0$，且$(\boldsymbol{\alpha},\boldsymbol{\alpha})=0\Leftrightarrow\boldsymbol{\alpha}=\boldsymbol{\theta}$，

则称$(\boldsymbol{\alpha},\boldsymbol{\beta})$为线性空间$V$中的内积，并称$V$为**实内积空间或欧氏空间**，记做$\left[V;(\boldsymbol{\alpha},\boldsymbol{\beta})\right]$。

例8.19　设$\boldsymbol{A}$为n阶正定矩阵，则

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta},\qquad \boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{R}^n$$

为$\mathbb{R}^n$上的内积。

例 8.20　在连续函数空间$C[a,b]$中，定义如下内积

$$(f(x),g(x))=\int_a^b f(x)g(x)dx,\qquad f(x),g(x)\in C[a,b]$$

则$C[a,b]$为欧式空间。

对于一般内积空间中的内积，它也有柯西不等式与三角不等式，其证明过程与$\mathbb{R}^n$中的证明完全相同。

有了内积的概念，欧氏空间$\mathbb{R}^n$中的相关概念都可以推广到一般的内积空间中，如向量的长度、单位向量、正交、正交向量组、标准正交向量组、正交基、标准正交基。而$\mathbb{R}^n$中的 Schmidt 方法也可原封不动地照搬到一般的欧式空间中。

习题 8.2

1.　在四维线性空间$\mathbb{R}^{2\times2}$中，证明：

$$\begin{bmatrix}1&1\\1&1\end{bmatrix},\begin{bmatrix}1&1\\-1&-1\end{bmatrix},\begin{bmatrix}1&-1\\1&-1\end{bmatrix},\begin{bmatrix}-1&1\\1&-1\end{bmatrix}$$

是$\mathbb{R}^{2\times2}$的一组基，并求矩阵$\begin{bmatrix}1&2\\3&4\end{bmatrix}$在这组基下的坐标。

2. 设$W=\left\{\boldsymbol{A}\middle|\boldsymbol{A}\in\mathbb{R}^{2\times2},\boldsymbol{A}^T=\boldsymbol{A}\right\}$，

（1）证明W是$\mathbb{R}^{2\times2}$的子空间；

（2）给出W的一组基，并求出W的维数。

3. 线性空间$\mathrm{R}^{2\times2}$中，

（1）求由基$\boldsymbol{E}_1=\begin{bmatrix}1&0\\0&0\end{bmatrix},\boldsymbol{E}_2=\begin{bmatrix}0&1\\0&0\end{bmatrix},\boldsymbol{E}_3=\begin{bmatrix}0&0\\1&0\end{bmatrix},\boldsymbol{E}_4=\begin{bmatrix}0&0\\0&1\end{bmatrix}$到基

$\boldsymbol{F}_1=\begin{bmatrix}1&1\\2&1\end{bmatrix},\boldsymbol{F}_2=\begin{bmatrix}0&2\\1&2\end{bmatrix},\boldsymbol{F}_3=\begin{bmatrix}0&0\\3&1\end{bmatrix},\boldsymbol{F}_4=\begin{bmatrix}0&0\\0&4\end{bmatrix}$过渡矩阵；

（2）求$\begin{bmatrix}1&1\\1&3\end{bmatrix}$在$\boldsymbol{F}_1,\boldsymbol{F}_2,\boldsymbol{F}_3,\boldsymbol{F}_4$下的坐标向量。

4. 在$\mathbb{F}^3$中，线性变换T关于基$\boldsymbol{\alpha}_1=[-1,1,1]^{\mathrm{T}},\boldsymbol{\alpha}_2=[1,0,-1]^{\mathrm{T}},\boldsymbol{\alpha}_3=[0,1,1]^{\mathrm{T}}$的矩阵为

$\boldsymbol{A}=\begin{bmatrix}1&0&1\\1&1&0\\-1&2&1\end{bmatrix}$，（1）求$T$关于标准基$\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$的矩阵；（2）设$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+6\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$，

求$T(\boldsymbol{\alpha})$关于基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$的坐标。

5. 在线性空间$\mathbb{F}^{2\times2}$中，$\boldsymbol{A}_1=\begin{bmatrix}1&2\\1&0\end{bmatrix},\boldsymbol{A}_2=\begin{bmatrix}-1&1\\1&1\end{bmatrix},\boldsymbol{B}_1=\begin{bmatrix}2&-1\\0&1\end{bmatrix},\boldsymbol{B}_2=\begin{bmatrix}1&-1\\3&7\end{bmatrix}$，

a) 求$L(\boldsymbol{A}_1,\boldsymbol{A}_2)\cap L(\boldsymbol{B}_3,\boldsymbol{B}_4)$的维数与一组基；

b) 求$L(\boldsymbol{A}_1,\boldsymbol{A}_2)+L(\boldsymbol{B}_3,\boldsymbol{B}_4)$的维数与一组基。

6. 设$W=\left\{f(x)\middle|f(x)\in\mathbf{P}_4[x],f(2)=0\right\}$，（1）证明$W$是$\mathbf{P}_4[x]$的子空间；（2）求$W$的维数与一组基。

7. 设$\boldsymbol{M}$是复数域上取定的方阵，令$S(\boldsymbol{M})=\left\{\boldsymbol{N}\middle|\boldsymbol{MN}+\boldsymbol{N}^{\mathrm{T}}\boldsymbol{M}=\boldsymbol{O}\right\}$，（1）证明：$S(\boldsymbol{M})$是复数域上的向量空间；（2）证明：如果$\boldsymbol{M}=\begin{bmatrix}\boldsymbol{O}&\boldsymbol{E}_n\\-\boldsymbol{E}_n&\boldsymbol{O}\end{bmatrix}$是$2n$阶方阵，则

$\dim S(\boldsymbol{M}) = 2n^2 + n$。

8. 设W是$\mathbb{F}^n$的一个非零子空间，若对于W的每一个向量$[a_1, a_2, \cdots, a_n]^{\mathrm{T}}$来说，或者$a_1 = a_2 = \cdots = a_n = 0$，或者每一个$a_i$都不为零，证明：$\dim W = 1$。

§8.3　线性变换

定义 8.8　线性空间上的**变换**与**线性变换**

设V是一线性空间，若有映射T，使得对V中的每一个向量$\boldsymbol{\alpha}$，都有确定的$T(\boldsymbol{\alpha}) \in V$与之对应，则称$T$为$V$上的一个**变换**，$T(\boldsymbol{\alpha})$称为$\boldsymbol{\alpha}$在变换$T$下的像，$\boldsymbol{\alpha}$称为$T(\boldsymbol{\alpha})$的原像。又若$T$满足：

(1)　$\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$，$T(\boldsymbol{\alpha} + \boldsymbol{\beta}) = T(\boldsymbol{\alpha}) + T(\boldsymbol{\beta})$；

(2)　$\forall \boldsymbol{\alpha} \in V, k \in \mathbb{F}$，$T(k\boldsymbol{\alpha}) = kT(\boldsymbol{\alpha})$，

则称T为V上的一个**线性变换**。

定理 8.6　设T是线性空间V上的一个变换，则T为线性变换的充要条件是

$$T(k_1\boldsymbol{\alpha} + k_2\boldsymbol{\beta}) = k_1T(\boldsymbol{\alpha}) + k_2T(\boldsymbol{\beta})$$

对任意的$k_1, k_2 \in \mathbb{F}, \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$成立。

证明　这是显然的，证明略。

例 8.21 线性空间V中，定义变换

$$\forall \boldsymbol{\alpha} \in V, \quad T(\boldsymbol{\alpha}) = \lambda \boldsymbol{\alpha}$$

其中λ为给定的数，证明T是线性变换。

证明：任取$k_1, k_2 \in \mathbb{F}, \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$，则

$$\begin{aligned}
&T(k_1\boldsymbol{\alpha} + k_2\boldsymbol{\beta}) = \lambda(k_1\boldsymbol{\alpha} + k_2\boldsymbol{\beta}) \\
&= \lambda k_1\boldsymbol{\alpha} + \lambda k_2\boldsymbol{\beta} = k_1\lambda\boldsymbol{\alpha} + k_2\lambda\boldsymbol{\beta} \\
&= k_1(\lambda\boldsymbol{\alpha}) + k_2(\lambda\boldsymbol{\beta}) \\
&= k_1T(\boldsymbol{\alpha}) + k_2T(\boldsymbol{\beta})
\end{aligned}$$

因此T是线性变换。

若取$T_1(\boldsymbol{\alpha}) = \boldsymbol{\lambda\alpha} + \boldsymbol{\gamma}$，其中$\boldsymbol{\gamma} \neq \boldsymbol{\theta}$为给定向量。则

$$T_1(k_1\boldsymbol{\alpha}+k_2\boldsymbol{\beta})=\boldsymbol{\lambda}(k_1\boldsymbol{\alpha}+k_2\boldsymbol{\beta})+\boldsymbol{\gamma}=k_1\boldsymbol{\lambda}\boldsymbol{\alpha}+k_2\boldsymbol{\lambda}\boldsymbol{\beta}+\boldsymbol{\gamma}$$

而

$$\begin{aligned}&k_1T(\boldsymbol{\alpha})+k_2T(\boldsymbol{\beta})=k_1(\boldsymbol{\lambda}\boldsymbol{\alpha}+\boldsymbol{\gamma})+k_2(\boldsymbol{\lambda}\boldsymbol{\beta}+\boldsymbol{\gamma})\\&=k_1\boldsymbol{\lambda}\boldsymbol{\alpha}+k_2\boldsymbol{\lambda}\boldsymbol{\beta}+(k_1+k_2)\boldsymbol{\gamma}\end{aligned}$$

由于$\boldsymbol{\gamma}\neq\boldsymbol{\theta}$，所以当$k_1+k_2\neq 1$时，

$$T_1(k_1\boldsymbol{\alpha}+k_2\boldsymbol{\beta})\neq k_1T_1(\boldsymbol{\alpha})+k_2T_1(\boldsymbol{\beta}),$$

因此T_1不是线性变换。

例 8.22　求导运算$\frac{d}{dx}$是线性空间$\mathbf{P}_n[x]$上的线性变换。

证明：显然$\frac{d}{dx}$是$\mathbf{P}_n[x]$上的变换。$\forall k_1,k_2\in\mathbb{F}, f(x),g(x)\in\mathbf{P}_n[x]$，根据求导法则可知

$$\frac{d}{dx}\left(k_1f(x)+k_2g(x)\right)=k_1\frac{df(x)}{dx}+k_2\frac{dg(x)}{dx}$$

成立。因此$\frac{d}{dx}$是线性变换。

例 8.23　线性空间$\mathbb{R}^n$中，定义变换$T_{\boldsymbol{A}}$如下

$$\forall \boldsymbol{X}\in\mathbb{R}^n,\quad T(\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X},$$

其中$\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}$为给定矩阵。证明$T_{\boldsymbol{A}}$是$\mathbb{R}^n$上的线性变换。

证明：因为$\forall\boldsymbol{X}\in\mathbb{R}^n$，　$T(\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X}\in\mathbb{R}^n$，因此$T_{\boldsymbol{A}}$是$\mathbb{R}^n$上的变换。

根据矩阵乘法的运算规律，容易验证：

$$\forall\ k_1,k_2\in\mathbb{R},\boldsymbol{X}_1,\boldsymbol{X}_2\in\mathbb{R}^n,\quad T_A(k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2)=k_1T_A(\boldsymbol{X}_1)+k_2T_A(\boldsymbol{X}_2)$$

因此它是线性变换。

例 8.24　$\mathbb{R}\to\mathbb{R}$的变换

$$T:x\to x^2$$

不是线性变换。

证明：取$k,x\in\mathbb{R}$，则$T(kx)=(kx)^2=k^2x^2$，而$kT(x)=kx^2$，因此当$k\neq 1,x\neq 0$时，这两者不相等，因此该映射不是线性变换。

例8.25　考虑 8.7 所定义的线性空间$\mathbb{R}^+$，定义变换：

$$T:\mathbb{R}^+\to\mathbb{R}^+：\ \forall x\in\mathbb{R}^+,\ T(x)=x^2$$

证明：T 是一个线性变换。

任取$a,b\in\mathbb{R}^+$，则

$$T(a\oplus b)=T(ab)=(ab)^2=a^2b^2=T(a)T(b)=T(a)\oplus T(b)$$

对于$a\in\mathbb{R}^+$，$k\in\mathbb{R}$，则

$$T(k\circ a)=T\left(a^k\right)=\left(a^k\right)^2=a^{2k}=\left(a^2\right)^k=k\circ a^2=k\circ T(a)$$

因此T 是一个线性变换。

实际上，根据$\mathbb{R}^+$中数乘的定义可知$T(x)=x^2=2\circ x$，实际上这是一个$\mathbb{R}^+$中的数乘运算，而根据例 8.17 可知，线性空间中的数乘运算总是线性变换。

定理 8.7　设T 是线性空间V 上的线性变换，则T 具有如下性质：

（1）$T(\boldsymbol{\theta})=\boldsymbol{\theta}$；

（2）$T(-\boldsymbol{\alpha})=-T(\boldsymbol{\alpha})$；

（3）$T\left(\sum_{i=1}^{r}k_i\boldsymbol{\alpha}_i\right)=\sum_{i=1}^{r}k_iT(\boldsymbol{\alpha}_i)$

（4）若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性相关，则$T(\boldsymbol{\alpha}_1),T(\boldsymbol{\alpha}_2),\cdots,T(\boldsymbol{\alpha}_r)$线性相关。

证明：略。

定理 8.7　设T 是线性空间V 上的线性变换，则

(1) $\mathrm{Im}(V)=T(V)=\{\boldsymbol{\beta}\,|\,\exists\boldsymbol{\alpha}\in V,\ s.t.\ \boldsymbol{\beta}=T(\boldsymbol{\alpha})\}$是$V$ 的子空间，称为T 的**像空间**；像空间的维数称为线性变换$\boldsymbol{T}$ 的**秩**。

(2) $Ker(T)=N(T)=\{\boldsymbol{\alpha}\in V\,|\,T(\boldsymbol{\alpha})=\boldsymbol{\theta}\}$是$\boldsymbol{V}$ 的子空间，称为T 的**零空间**（**核**）。零空间的维数称为线性变换T 的**零度**（nullity）。

证明：（1）显然$\boldsymbol{\theta}\in T(V)$，故$T(V)\neq\phi$，任取$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\in T(V)$，则存在$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in V$，使得

$$T(\boldsymbol{\alpha}_i)=\boldsymbol{\beta}_i,\qquad i=1,2$$

由于T 是线性空间V 上的线性变换，因此$\forall k_1,k_2\in\mathbb{F}$，有

$$\begin{aligned} k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2 &= k_1T(\boldsymbol{\alpha}_1)+k_2T(\boldsymbol{\alpha}_2) \\ &= T(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2) \end{aligned}$$

由于 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2 \in V$，故 $T(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2)\in T(V)$。因此 $T(V)$ 为 V 的子空间。

（2）显然 $\boldsymbol{\theta}\in N(T)$，故 $N(T)\neq\boldsymbol{\phi}$。任取 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in N(V)$，故

$$T(\boldsymbol{\alpha}_i)=\boldsymbol{\beta}_i,\qquad i=1,2$$

$\forall k_1,k_2\in\mathbb{F}$，有 $T(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2)=k_1T(\boldsymbol{\alpha}_1)+k_2T(\boldsymbol{\alpha}_2)=k_1\boldsymbol{\theta}+k_2\boldsymbol{\theta}=\boldsymbol{\theta}$，故得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\in N(T)。$$

因此 $N(T)$ 为 V 的子空间。

例 8.26　求例 8.21 中所给线性变换 T_A 的像空间 $T_A(\mathbb{R}^n)$ 与零空间 $N(T_A)$。

解：设矩阵 $\boldsymbol{A}$ 的列向量组为 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$，$\boldsymbol{X}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$，则

$$\boldsymbol{T}_A(\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X}=\sum_{i=1}^{n}x_i\boldsymbol{A}_i$$

故 $T_A(\mathbb{R}^n)=span\{\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n\}=R(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的列空间。

而 $N(T_A)=\left\{\boldsymbol{X}\in\mathbb{R}^n \,\middle|\, T_A(\boldsymbol{X})=\boldsymbol{\theta}\right\}=\left\{\boldsymbol{X}\in\mathbb{R}^n \,\middle|\, \boldsymbol{A}\boldsymbol{X}=\boldsymbol{\theta}\right\}=N(\boldsymbol{A})$。

习题 8.3

1. 在 $\mathbb{R}^3$ 空间中定义了下面四个变换，哪个不是线性变换(　　)

   ① 平移变换
   ② 将任意向量向 OXY 平面做垂直投影
   ③ 以原点为中心的旋转变换
   ④ 将任意向量做关于 X 轴的对称变换

2. 设 V 是 n 维线性空间，W 是 V 的子空间，证明：存在 V 上的线性变换 T,S，使得

   $T(V)=W, KerS=W$。

§8.4　线性变换的矩阵表示

T 是 n 维线性空间 V 上的线性变换，因而 T 是一种映射。而我们知道，要刻画一个映射，需要给出它的定义域、值域、对应法则。对于线性空间 V 上的线性变换 T，其定义域、

值域都是V，所以只需要给出它的对应法则即可。一般而言，要给出T所确定的对应法则，我们需要给出任意$\boldsymbol{\alpha}\in V$的像$T\left(\boldsymbol{\alpha}\right)$。但对于有限维线性空间上的线性变换，我们可以用更简单的方法给出线性变换所确定的对应法则。

设T是n维线性空间V上的线性变换，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$为$V$的一组基，$\forall\boldsymbol{\alpha}\in V$设$\boldsymbol{\alpha}$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的坐标为$\boldsymbol{X}=\left[x_1,x_2,\cdots,x_n\right]^{\mathrm{T}}$，即

$$\begin{aligned}\boldsymbol{\alpha}&=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n\\&=\begin{bmatrix}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\cdots&\boldsymbol{\alpha}_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\end{aligned}$$

则

$$\begin{aligned}T\left(\boldsymbol{\alpha}\right)&=T\left(x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n\right)\\&=x_1T\left(\boldsymbol{\alpha}_1\right)+x_2T\left(\boldsymbol{\alpha}_2\right)+\cdots+x_nT\left(\boldsymbol{\alpha}_n\right)\\&=\begin{bmatrix}T\left(\boldsymbol{\alpha}_1\right)&T\left(\boldsymbol{\alpha}_2\right)&\cdots&T\left(\boldsymbol{\alpha}_n\right)\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\end{aligned}$$

因此，只要知道了$T\left(\boldsymbol{\alpha}_i\right),i=1,2,\cdots,n$，就可知道$T\left(\boldsymbol{\alpha}\right)$。由于$T\left(\boldsymbol{\alpha}_i\right)\in V$，故它可表示为$V$的基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$的线性组合，因此向量组$T\left(\boldsymbol{\alpha}_1\right),T\left(\boldsymbol{\alpha}_2\right),\cdots,T\left(\boldsymbol{\alpha}_n\right)$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性表示，因而存在矩阵$\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}$，使得

$$\begin{bmatrix}T\left(\boldsymbol{\alpha}_1\right)&T\left(\boldsymbol{\alpha}_2\right)&\cdots&T\left(\boldsymbol{\alpha}_n\right)\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\cdots&\boldsymbol{\alpha}_n\end{bmatrix}\boldsymbol{A}$$

根据上述定义可知$\boldsymbol{A}_i=\left[a_{1i},a_{2i},\cdots,a_{ni}\right]^{\mathrm{T}}$为$T\left(\boldsymbol{\alpha}_i\right)$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的坐标。

记$\begin{bmatrix}T\left(\boldsymbol{\alpha}_1\right)&T\left(\boldsymbol{\alpha}_2\right)&\cdots&T\left(\boldsymbol{\alpha}_n\right)\end{bmatrix}=T\begin{bmatrix}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\cdots&\boldsymbol{\alpha}_n\end{bmatrix}$，就有

$$T\begin{bmatrix}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\cdots&\boldsymbol{\alpha}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\cdots&\boldsymbol{\alpha}_n\end{bmatrix}\boldsymbol{A}$$

因此

$$
\begin{aligned}
T(\boldsymbol{\alpha}) &= T(x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n) \\
&= x_1T(\boldsymbol{\alpha}_1) + x_2T(\boldsymbol{\alpha}_2) + \cdots + x_nT(\boldsymbol{\alpha}_n) \\
&= \begin{bmatrix} T(\boldsymbol{\alpha}_1) & T(\boldsymbol{\alpha}_2) & \cdots & T(\boldsymbol{\alpha}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n \end{bmatrix} \boldsymbol{A} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n \end{bmatrix} \boldsymbol{AX}
\end{aligned}
$$

即$T(\boldsymbol{\alpha})$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的坐标为$\boldsymbol{AX}$。

定义 8.9　设T是n维线性空间V上的线性变换，$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$为$V$的一组基，则有矩阵$\boldsymbol{A}$，使得

$$
T\begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n \end{bmatrix} \boldsymbol{A} \tag{8.1}
$$

称$\boldsymbol{A}$为线性变换在基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$下的矩阵。

对于给定的线性变换T，其在一组基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$下的矩阵是唯一的；反过来，对于给定的一组基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$及线性变换$T$在这组基下的矩阵$\boldsymbol{A}$，就唯一地决定了这个线性变换。即

$$
T \xleftrightarrow{\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}} \boldsymbol{A}_{n\times n}
$$

根据以上推导可知，若线性变换T在基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$下的矩阵为$\boldsymbol{A}$，$\boldsymbol{\alpha}$在基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$下的坐标为$\boldsymbol{X}$，则$T(\boldsymbol{\alpha})$在基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$下的坐标为$\boldsymbol{AX}$。因此有下图成立

$$
\boldsymbol{\alpha} \xrightarrow{T} T(\boldsymbol{\alpha})
$$

$$
\boldsymbol{X} \xrightarrow{T} \boldsymbol{AX}
$$

下面讨论如何求线性变换在一组基下的矩阵。再一次强调：设$\boldsymbol{A}$为线性变换T在基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$下的矩阵，则$\boldsymbol{A}$的第$i$个列向量$\boldsymbol{A}_i$为$T(\boldsymbol{\alpha}_i)$在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的坐标向量，这就给出了求线性变换在一组基下的矩阵的方法——求出$T(\boldsymbol{\alpha}_i)$，并将$T(\boldsymbol{\alpha}_i)$用

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示，求得 $T(\boldsymbol{\alpha}_i)$ 在 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 下的坐标作为 $\boldsymbol{A}$ 的第 i 列。

情形 1、如果已经给出线性变换，则只需要求基向量在该变换下的象，再求出该像在这组基下的坐标。

例8.26 $\mathbb{R}^3$ 的线性变换 T 由

$$T\left([x,y,z]^{\mathrm{T}}\right)=[x+y,y-z,z-x]^{\mathrm{T}}$$

给出。求 T 在标准基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$ 下的矩阵。

解： $T(\boldsymbol{e}_1)=T\left([1,0,0]^{\mathrm{T}}\right)=[1,0,-1]^{\mathrm{T}}=[\boldsymbol{e}_1\quad \boldsymbol{e}_2\quad \boldsymbol{e}_3]\begin{bmatrix}1\\0\\-1\end{bmatrix}$

$$T(\boldsymbol{e}_2)=T\left([0,1,0]^{\mathrm{T}}\right)=[1,1,0]^{\mathrm{T}}=[\boldsymbol{e}_1\quad \boldsymbol{e}_2\quad \boldsymbol{e}_3]\begin{bmatrix}1\\1\\0\end{bmatrix}$$

$$T(\boldsymbol{e}_3)=T\left([0,0,1]^{\mathrm{T}}\right)=[0,-1,1]^{\mathrm{T}}=[\boldsymbol{e}_1\quad \boldsymbol{e}_2\quad \boldsymbol{e}_3]\begin{bmatrix}0\\-1\\1\end{bmatrix}$$

所以，线性变换 T 在标准基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$ 下的矩阵为

$$\begin{bmatrix}1&1&0\\0&1&-1\\-1&0&1\end{bmatrix}$$

例 8.27　设四维线性空间 $\mathbf{P}_3[x]$ 上的线性变换 T 定义如下：

$$\forall\ f(x)\in\mathbf{P}_3[x]\ ,\qquad T(f(x))=\frac{df(x)}{dx}-f(x)$$

求 T 在基 $\{1,x,x^2,x^3\}$ 下的矩阵。

解：

$$T(1)=\frac{d(1)}{dx}-1=0-1=-1=[1\quad x\quad x^2\quad x^3]\begin{bmatrix}-1\\0\\0\\0\end{bmatrix}$$

$$T(x)=\frac{dx}{dx}-x=1-x=\begin{bmatrix}1 & x & x^2 & x^3\end{bmatrix}\begin{bmatrix}1\\-1\\0\\0\end{bmatrix}$$

$$T(x^2)=\frac{dx^2}{dx}-x^2=2x-x^2=\begin{bmatrix}1 & x & x^2 & x^3\end{bmatrix}\begin{bmatrix}0\\2\\-1\\0\end{bmatrix}$$

$$T(x^3)=\frac{dx^3}{dx}-x^3=3x^2-x^3=\begin{bmatrix}1 & x & x^2 & x^3\end{bmatrix}\begin{bmatrix}0\\0\\3\\-1\end{bmatrix}$$

故该线性变换T在所给基下的矩阵为

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

情形 2：给出了n个线性无关的向量在线性变换下的象，求该线性变换在一组基下的矩阵。这种情况可转化为$\mathbb{R}^n$中的同类问题。

例 8.28　设T是三维线性空间V_3的线性变换，$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$是$V_3$的基，线性变换由下列方程给出，其中括号内向量是关于基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$的坐标向量，

$$T\left([3,2,1]^{\mathrm{T}}\right)=[0,0,2]^{\mathrm{T}}$$
$$T\left([0,1,1]^{\mathrm{T}}\right)=[1,1,1]^{\mathrm{T}}$$
$$T\left([2,0,1]^{\mathrm{T}}\right)=[1,0,1]^{\mathrm{T}}$$

求T在这组基下的矩阵。

解：设T在这组基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$下的矩阵为$\boldsymbol{A}$，则根据坐标公式

$$\begin{bmatrix}0\\0\\2\end{bmatrix}=\boldsymbol{A}\begin{bmatrix}3\\2\\1\end{bmatrix}\qquad\begin{bmatrix}1\\1\\1\end{bmatrix}=\boldsymbol{A}\begin{bmatrix}0\\1\\1\end{bmatrix}\qquad\begin{bmatrix}1\\0\\1\end{bmatrix}=\boldsymbol{A}\begin{bmatrix}2\\0\\1\end{bmatrix}$$

根据分块矩阵乘法，有

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} 3 & 0 & 2 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

由于$\begin{vmatrix} 3 & 0 & 2 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 5 \neq 0$，故$\begin{bmatrix} 3 & 0 & 2 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$可逆，所以

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 2 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}^{-1} = \frac{1}{5} \begin{bmatrix} -1 & -2 & 7 \\ -2 & 1 & 4 \\ 1 & 2 & 3 \end{bmatrix}$$

习题 8. 4

1. 线性空间 $P_{n-1}[x]$ 有基 $1, x, x^2, \cdots, x^{n-1}$. 定义 $P_{n-1}[x]$ 上的线性变换 T : 对任意 $f \in P_{n-1}[x],\quad T(f) = \dfrac{\mathrm{d}f}{\mathrm{d}x} - f$. 求 T 在基 $1, x, x^2, \cdots, x^{n-1}$ 下的矩阵表示.
2. 函数集合 $V = \left\{ \left(a_1 x^2 + a_2 x + a_3 \right) e^x \middle| a_1, a_2, a_3 \in \mathbb{R} \right\}$ 对于函数的线性运算构成 3 维线性空间,，在 V 中取一个基 $x^2 e^x, xe^x, e^x$ 求微分运算 $\dfrac{d}{dx}$ 在这个基下的矩阵。
3. 设 $W = \left\{ \boldsymbol{A} \middle| \boldsymbol{A} \in \mathbb{R}^{2\times 2}, \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A} \right\}$，

    （1）证明 W 是 $\mathbb{R}^{2\times 2}$ 的子空间；

    （2）给出 W 的一组基，并求出 W 的维数；

    （3）定义 W 上的线性变换 T ： $T\left(\begin{bmatrix} a & b \\ b & c \end{bmatrix} \right) = \begin{bmatrix} a+b & b \\ b & a+c \end{bmatrix}$，求 T 在您（2）中给出的基下的矩阵。
4. 在次数不大于 3 的实系数多项式的全体添上零多项式对于多项式的加法与多项式的数乘构成的实数域上的线性空间 $\mathbb{R}_3[x]$ 中定义映射

    $T\left(f\left(x \right) \right) = f\left(x + a \right), \forall f\left(x \right) \in \mathbb{R}_3[x]$，（其中实数 $a \neq 0$）

    （1）证明：T 是 $\mathbb{R}_3[x]$ 上的线性变换；（2）求 T 在基 $1, x, x^2, x^3$ 下的矩阵。
5. 在 $\mathbb{R}^{2\times 2}$ 中定义线性变换 T 如下：取定 $\boldsymbol{P} = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}, \forall \boldsymbol{A} \in \mathbb{R}^{2\times 2}$，定义 $T\left(\boldsymbol{A} \right) = \boldsymbol{AP}$，

    （1）T 为 $\mathbb{R}^{2\times 2}$ 的线性变换；（2）求 T 在基

$$E_1=\begin{bmatrix}1&0\\0&0\end{bmatrix},E_2=\begin{bmatrix}1&1\\0&0\end{bmatrix},E_3=\begin{bmatrix}1&1\\1&0\end{bmatrix},E_4=\begin{bmatrix}1&1\\1&1\end{bmatrix}$$

下的矩阵。

6. 2 阶实对称矩阵的全体$W=\left\{\boldsymbol{A}\middle|\boldsymbol{A}\in\mathbb{R}^{2\times2},\boldsymbol{A}^T=\boldsymbol{A}\right\}$对于矩阵的线性运算构成 3 维线性空间。求$W$中的线性变换

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

在组基

$$\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}$$

下的矩阵。

§8.5　线性变换在不同基下矩阵的关系

线性变换可用其在一组基下的矩阵来刻画，如果选择了不同的基，那么该线性变换在不同基下的矩阵有何关系？

设V为n维线性空间，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$、$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$为$V$的两组基，从$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$到$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$的过渡矩阵为$\boldsymbol{C}$，即有

$$[\boldsymbol{\beta}_1\quad\boldsymbol{\beta}_2\quad\cdots\quad\boldsymbol{\beta}_n]=[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\cdots\quad\boldsymbol{\alpha}_n]\boldsymbol{C}$$

设线性变换T在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$、$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$下的矩阵分别为$\boldsymbol{A},\boldsymbol{B}$，现在我们讨论$\boldsymbol{A}$与$\boldsymbol{B}$的关系。

$$T[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\cdots\quad\boldsymbol{\alpha}_n]=[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\cdots\quad\boldsymbol{\alpha}_n]\boldsymbol{A}$$

$$T[\boldsymbol{\beta}_1\quad\boldsymbol{\beta}_2\quad\cdots\quad\boldsymbol{\beta}_n]=[\boldsymbol{\beta}_1\quad\boldsymbol{\beta}_2\quad\cdots\quad\boldsymbol{\beta}_n]\boldsymbol{B}$$

由于

$$[\boldsymbol{\beta}_1\quad\boldsymbol{\beta}_2\quad\cdots\quad\boldsymbol{\beta}_n]=[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\cdots\quad\boldsymbol{\alpha}_n]\boldsymbol{C}$$

所以

$$\begin{aligned}&T[\boldsymbol{\beta}_1\quad\boldsymbol{\beta}_2\quad\cdots\quad\boldsymbol{\beta}_n]=T\left([\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\cdots\quad\boldsymbol{\alpha}_n]\boldsymbol{C}\right)\\&=T\left([\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\cdots\quad\boldsymbol{\alpha}_n]\right)\boldsymbol{C}\\&=[\boldsymbol{\alpha}_1\quad\boldsymbol{\alpha}_2\quad\cdots\quad\boldsymbol{\alpha}_n]\boldsymbol{AC}\end{aligned}$$

而

$$T[\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n]=[\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n]\boldsymbol{B}=[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n]\boldsymbol{CB}$$

所以

$$[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n]\boldsymbol{CB}=[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n]\boldsymbol{AC}$$

由于$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性无关，所以

$$\boldsymbol{CB}=\boldsymbol{AC}$$

得到

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

或

$$\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{AC}$$

因此一个线性变换在不同基下的矩阵是相似的。

例 8.29　设V是数域$\mathbb{F}$上的一个 3 维线性空间，V上的线性变换T在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$下的矩阵为$\boldsymbol{A}=\begin{bmatrix}1&3&-2\\2&1&1\\-1&1&2\end{bmatrix}$，请计算$T$在基$[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}-3&2&-1\\-3&3&-2\\-2&-3&1\end{bmatrix}$下的矩阵$\boldsymbol{B}$。

解　根据题意可知，从基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$到基$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$的过渡矩阵为$\boldsymbol{C}=\begin{bmatrix}-3&2&-1\\-3&3&-2\\-2&-3&1\end{bmatrix}$，根据前文所得到的结论可知

$$\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{AC}=\begin{bmatrix}-3&2&-1\\-3&3&-2\\-2&-3&1\end{bmatrix}^{-1}\begin{bmatrix}1&3&-2\\2&1&1\\-1&1&2\end{bmatrix}\begin{bmatrix}-3&2&-1\\-3&3&-2\\-2&-3&1\end{bmatrix}$$
$$=\frac{1}{8}\begin{bmatrix}17&-42&23\\11&114&-51\\35&218&-99\end{bmatrix}$$

习题 8.5

1.　在四维空间$\mathrm{R}^{2\times2}$中，设线性变换 T 在基$\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\}$下 的矩阵

为$\begin{bmatrix} 1 & 3 & -3 & 2 \\ -1 & -5 & 2 & 1 \\ 0 & -2 & 5 & 0 \\ 0 & 1 & -1 & 1 \end{bmatrix}$，求T在基$\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$下的矩阵.

2. 已知线性变换T在基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$下的矩阵为$\boldsymbol{A}=\left[a_{ij}\right]_{n\times n}$，求$T$在基$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_3$下的矩阵。