第 1 章　行 列 式

行列式是线性代数中一个最基本的概念，它是研究线性代数的一个重要工具，在线性方程组、矩阵、二次型、线性变换等的讨论中都要用到行列式，在数学的其他分支以及一些实际问题中也常常用到行列式.

这一章的主要内容就是介绍行列式的定义、性质以及计算方法.

1.1　2 阶和 3 阶行列式

行列式是一种特定的算式，是根据线性方程组求解的需要而引进的. 首先介绍 2 阶和 3 阶行列式.

由两个方程式组成的二元线性方程组经过变形以后，可以化成一般形式：

$$\begin{cases} a_1x + b_1y = c_1, \\ a_2x + b_2y = c_2. \end{cases} \tag{1}$$

用消元法来解这个方程组，以 b_2 乘第 1 个方程，以 b_1 乘第 2 个方程，然后两式相减，便消去了 y，得到：

$$(a_1b_2 - a_2b_1)x = c_1b_2 - c_2b_1.$$

用同样的方法，可消去 y，得：

$$(a_1b_2 - a_2b_1)y = a_1c_2 - a_2c_1.$$

如果 $a_1b_2-a_2b_1\neq 0$，那么可得到：

$$\begin{cases} x = \dfrac{c_1b_2 - c_2b_1}{a_1b_2 - a_2b_1}, \\ y = \dfrac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}. \end{cases} \tag{2}$$

把这一组 x,y 的值代入方程组(1),可以验证它确是原方程组的解,而且方程组(1)只有这一组解.

这样,就可以用公式(2)来解二元线性方程组(1). 为了便于记忆这个公式,我们引进 2 阶行列式的概念,令:

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1.$$

这样规定的 $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ 称为 2 阶行列式. 根据这个规定,公式(2)中 x,y 的分子也可用 2 阶行列式表示:

$$D_1 = \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} = c_1b_2 - c_2b_1;$$

$$D_2 = \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} = a_1c_2 - a_2c_1.$$

我们把方程组(1)的系数组成的行列式称为(1)的系数行列式,于是上面的结论就可叙述为:**二元线性方程组(1)当它的系数行列式 $D\neq0$ 时有唯一解. 这个解可以用公式表示:**

$$x = \frac{D_1}{D}, \qquad y = \frac{D_2}{D}.$$

例 1 解方程组

$$\begin{cases} 2x + 3y = 7, \\ 3x + 7y = 6. \end{cases}$$

解: 这个方程组的系数行列式为:

$$D = \begin{vmatrix} 2 & 3 \\ 3 & 7 \end{vmatrix} = 14 - 9 = 5 \neq 0.$$

所以这个方程组有唯一解.又由

$$D_1 = \begin{vmatrix} 7 & 3 \\ 6 & 7 \end{vmatrix} = 49 - 18 = 31,$$

$$D_2 = \begin{vmatrix} 2 & 7 \\ 3 & 6 \end{vmatrix} = 12 - 21 = -9,$$

得到解:

$$x = \frac{31}{5} = 6\frac{1}{5}, \qquad y = \frac{-9}{5} = -1\frac{4}{5}.$$

下面讨论三元线性方程组:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \end{cases} \tag{3}$$

还用消元法来解这个方程组.分别用$(a_{22}a_{33}-a_{23}a_{32})$乘第 1 个方程,用$(a_{13}a_{32}-a_{12}a_{33})$乘第 2 个方程,用$(a_{12}a_{23}-a_{13}a_{22})$乘第 3 个方程,再把所得的 3 个式子相加,就消去了 x_2,x_3,得:

$$\begin{aligned} &(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\qquad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31})x_1 \\ &= b_1a_{22}a_{33} + a_{12}a_{23}b_3 + a_{13}b_2a_{32} \\ &\quad - b_1a_{23}a_{32} - a_{12}b_2a_{33} - a_{13}a_{22}b_3. \end{aligned} \tag{4}$$

用同样的方法消去 x_1,x_3,得:

$$\begin{aligned} &(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\qquad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31})x_2 \\ &= a_{11}b_2a_{33} + b_1a_{23}a_{31} + a_{13}a_{21}b_3 \\ &\quad - a_{11}a_{23}b_3 - b_1a_{21}a_{33} - a_{13}b_2a_{31}. \end{aligned} \tag{5}$$

消去 x_1,x_2 得:

$$\begin{aligned} &(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\qquad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31})x_3 \end{aligned}$$

$$
\begin{aligned}
&= a_{11}a_{22}b_3 + a_{12}b_2a_{33} + b_1a_{21}a_{32} \\
&\quad - a_{11}b_2a_{32} - a_{12}a_{21}b_3 - b_1a_{22}a_{31}.
\end{aligned} \tag{6}
$$

因此，当

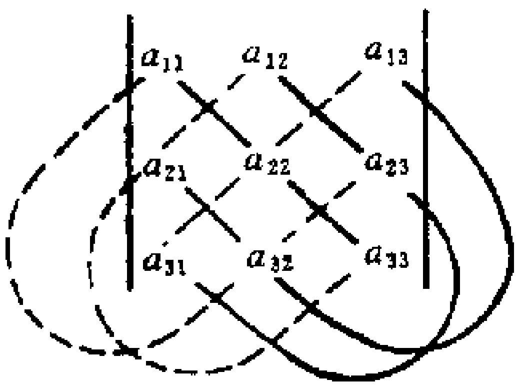
$$
\begin{aligned}
&a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\
&- a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \neq 0
\end{aligned}
$$

时，即可解出 x_1, x_2, x_3. 为了简单地表出这个方程组的解，我们引进 3 阶行列式的概念.

规定 3 阶行列式为：

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

可见，一个 3 阶行列式是由不同行不同列的 3 个数相乘而得到的 6 个项的代数和. 这些项前面所带的正负号可以从下图看出.



凡是实线上 3 个元素相乘所得到的项的前面带正号；虚线上 3 个元素相乘所得到的项的前面带负号.

例 2

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

$$
= 1 \times (-1) \times 2 + 2 \times 3 \times 3 + 5 \times 2 \times 0
$$

$$-1\times3\times0-2\times2\times2-5\times(-1)\times3$$
$$=-2+18+0-0-8-(-15)=23.$$

例 3

$$\begin{vmatrix}2 & 1 & -1\\ 3 & a & 0\\ b & -1 & c\end{vmatrix}$$
$$=2\times a\times c+1\times0\times b+(-1)\times3\times(-1)$$
$$-(-1)\times a\times b-0\times(-1)\times2-c\times1\times3$$
$$=2ac+ab-3c+3.$$

根据 3 阶行列式的定义,可以把消元后所得到的 3 个式子的右边用 3 阶行列式来表示:

$$\begin{vmatrix}b_1 & a_{12} & a_{13}\\ b_2 & a_{22} & a_{23}\\ b_3 & a_{32} & a_{33}\end{vmatrix}=b_1a_{22}a_{33}+a_{12}a_{23}b_3+a_{13}b_2a_{32}$$
$$-b_1a_{23}a_{32}-a_{12}b_2a_{33}-a_{13}b_2a_{32}.$$

同样地,可看出,(5)式及(6)式的右边分别等于 3 阶行列式:

$$\begin{vmatrix}a_{11} & b_1 & a_{13}\\ a_{21} & b_2 & a_{23}\\ a_{31} & b_3 & a_{33}\end{vmatrix}\quad 和 \quad\begin{vmatrix}a_{11} & a_{12} & b_1\\ a_{21} & a_{22} & b_2\\ a_{31} & a_{32} & b_3\end{vmatrix}.$$

于是得到三元一次方程组解的公式:

如果方程组(1)的系数行列式

$$D=\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}\neq0,$$

那么方程组(1)有唯一解:

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

其中 $D_j(j=1,2,3)$是把行列式 D 的第 j 列换成常数项 b_1,b_2,b_3 所得的行列式.

例 4 解三元线性方程组

$$\begin{cases} x_1 - 2x_2 + x_3 = -4, \\ 2x_1 + x_2 - 3x_3 = 7, \\ -x_1 + x_2 - x_3 = 2. \end{cases}$$

解：这个方程组的系数行列式

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

$$= -1 - 6 + 2 + 1 + 3 - 4 = -5 \neq 0.$$

因此有唯一解. 又因

$$D_1 = \begin{vmatrix} -4 & -2 & 1 \\ 7 & 1 & -3 \\ 2 & 1 & -1 \end{vmatrix} = -5;$$

$$D_2 = \begin{vmatrix} 1 & -4 & 1 \\ 2 & 7 & -3 \\ -1 & 2 & -1 \end{vmatrix} = -10;$$

$$D_3 = \begin{vmatrix} 1 & -2 & -4 \\ 2 & 1 & 7 \\ -1 & 1 & 2 \end{vmatrix} = 5.$$

所以解为 $x_1 = \dfrac{D_1}{D} = 1$，$x_2 = \dfrac{D_2}{D} = 2$，$x_3 = \dfrac{D_3}{D} = -1$.

从上面的例子看到：应用 2 阶、3 阶行列式来解系数行列式不等于零的二元、三元线性方程组是很方便的. 为了把这个结论推广到未知量更多的线性方程组，即一般线性方程组，我们需要把阶 2、3 阶行列式的概念推广，引进 n 阶行列式的概念. 为此，在下一节中先介绍 n 阶排列的概念.

习 题 1.1

1. 计算下列行列式：

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

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

(5) $\begin{vmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{vmatrix}$；　　(6) $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$.

2. 解下列线性方程组：

(1) $\begin{cases} 2x - 5y = 6, \\ x - y = 5; \end{cases}$　　(2) $\begin{cases} x_1 + 2x_2 + x_3 = -2, \\ 2x_1 + x_2 + 3x_3 = 1, \\ x_1 + x_2 + x_3 = 0; \end{cases}$

(3) $\begin{cases} 2x_1 - 4x_2 + x_3 = 1, \\ x_1 - 5x_2 + 3x_3 = 2, \\ x_1 - x_2 + x_3 = -1; \end{cases}$　　(4) $\begin{cases} 2x_1 - 3x_2 - x_3 = 10, \\ x_1 + 4x_2 - 2x_3 = -8, \\ 3x_1 + 2x_2 - x_3 = 1. \end{cases}$

1.2　n 阶排列

本节介绍 n 阶排列的概念和一些基本性质. 这一方面是为以后定义行列式作准备；另一方面，也因为排列本身就是一个重要的概念，在以后学习数学的其他分支，如概率论等都要用到.

定义 1　由 $1,2,\cdots,n$ 组成的一个有序数组称为一个 **n 阶排列**.

这里的排列就是以前所说的 n 个不同元素的全排列. 所以 n 阶排列一共有 $n!$ 个.

例 1　写出所有的 3 阶排列.

解：自然数 1,2,3 组成的有序数组共有下列 6 个：

123，　132，　213，　231，　312，　321.

这就是全部 3 阶排列.

例 2　写出全部 4 阶排列.

解：4 阶排列一共有 4！=24 个. 它们是：

1234，1243，1324，1342，
1423，1432，2134，2143，
2314，2341，2413，2431，
3124，3142，3214，3241，
3412，3421，4123，4132，
4213，4231，4312，4321.

4 阶排列 1234，它的各个数是按照由小到大的自然顺序排列的，称为 4 阶**自然序排列**. 一般地，1234…n 称为 n 阶自然序排列. 在其他的排列中，都可找到一个大数排在一个小数的前面. 例如在 4 阶排列 2143 中，2 排在 1 之前，4 排在 3 之前. 这样的排列顺序是与自然顺序相反的. 我们称它为逆序. 这就是下述定义：

定义 2 在一个 n 阶排列中，如果一个大数排在一个小数之前，就称这两个数组成一个**逆序**. 一个 n 阶排列中逆序的总数称为这个排列的**逆序数**. 反之，在一个排列中，如果一个小数排在一个大数之前，就称这两个数组成一个**顺序**.

例 3 求 4213 的逆序数.

解：在排列 4213 中共有 42，41，43，21 等 4 个逆序. 所以 4213 的逆序数等于 4.

例 4 排列 21534，54321，23154，21354 和 45132 的逆序和逆序数如下：

排　　列	逆　　序	逆　序　数
21534	21，53，54	3
54321	54，53，52，51，43，42，41，32，31，21	10
23154	21，31，54	3
21354	21，54	2
45132	41，43，42，51，53，52，32	7

为了方便起见，我们引进一个符号：如果 $a_1a_2\cdots a_n$ 是一个 n 阶排列，我们用 $\tau(a_1a_2\cdots a_n)$ 来表示 $a_1a_2\cdots a_n$ 的逆序数. 例如 $\tau(21\ 5\ 34)=3$, $\tau(45\ 1\ 32)=7$.

例 5 求 $\tau(n,n-1,\cdots,2,1)$.

解：在 $n,n-1,\cdots,2,1$ 中，n 与后面 $n-1$ 个数都组成逆序；$n-1$ 与它后面的 $n-2$ 个数组成逆序……一般地，$k(k>1)$ 与它后面 $k-1$ 个数组成逆序. 所以

$$\begin{aligned}\tau(n\ n-1\ \cdots\ 2\ 1)&=(n-1)+(n-2)+\cdots+2+1\\&=\frac{n(n-1)}{2}.\end{aligned}$$

例 5 中的方法也就是一般用来求排列的逆序数的方法.

习　题　1.2(1)

1. 写出第 1,2 个位置是 2,4 的全部 5 阶排列，并求它们的逆序数.
2. 求下列排列的逆序数：317428695；528497631；654321.

下面介绍排列的奇偶性.

定义 3 设 $a_1a_2\cdots a_n$ 是一个 n 阶排列. 如果 $\tau(a_1a_2\cdots a_n)$ 是一个偶数. 则称 $a_1a_2\cdots a_n$ 是一个**偶排列**；如果 $\tau(a_1a_2\cdots a_n)$ 是一个奇数，则称 $a_1a_2\cdots a_n$ 是一个**奇排列**.

也就是说，逆序数是偶数的排列称为偶排列，逆序数是奇数的排列称为奇排列.

例如：在例 4 中，54321 和 21354 是偶排列，其余的 3 个排列都是奇排列.

把一个排列中的某两个数互换位置，而其他的数保持不动，就得到另一个排列，这样的一种变换称为一个**对换**. 例如，在排列 21354 中将 1,3 两个数对换，得到排列 23154；在排列 32451 中将 3,5 对换，得到排列 52431.

下面讨论对换对于排列奇偶性的影响. 先看一下上面的例子,我们已知排列 21354 有 2 个逆序:21,54;而排列 23154 有 3 个逆序:21,31,54. 这说明偶排列 21354 经过一次对换后,成为奇排列 23154,而奇排列 32451 经过一次对换,成为偶排列 52431. 我们来分析一下对换改变这两个排列奇偶性的原因. 先来比较 21354 和 23154. 因为这两个排列只是 1 与 3 的位置对换了一下,而且 1 与 3 处于相邻的位置,因此除了 1 与 3 在 21354 中是顺序,而在 23154 变为逆序以外,其他数字之间的次序关系都没有改变,因此 23154 比 21354 多了一个逆序,它们的奇偶相反. 至于排列 32451 与 52431,它们对换的数字 3 与 5 不在相邻的位置,但是可以通过若干次相邻位置的对换而将 32451 变为 52431

$$32451 \xrightarrow{(2,3)} 23451 \xrightarrow{(3,4)} 24351 \xrightarrow{(3,5)} 24531$$
$$\xrightarrow{(4,5)} 25431 \xrightarrow{(2,5)} 52431.$$

从上面的式子看到,32451 经过 5 次相邻数字的变换变为 52431. 每经过一次相邻位置的对换排列改变一次奇偶,所以排列 32451 与 52431 奇偶相反.

我们按照这个想法来证明下面的定理.

定理 1 对换改变排列的奇偶性.

这就是说,经过一次对换,奇排列变成偶排列,偶排列变成奇排列.

证明:首先讨论对换的两个数 i 与 j 在排列中处于相邻位置的情形. 即排列

$$\cdots i \quad j \cdots \tag{1}$$

经过 i,j 对换,变成排列

$$\cdots j \quad i \cdots \tag{2}$$

显然,i,j 以外的数彼此间的逆序状况在排列(1)和(2)中是一样的;i,j 以外的数与 i(或 j)的逆序状况在排列(1)和(2)中也是一

样的.如果 i 与 j 在排列(1)中构成逆序,则它们的排列(2)中是顺序,这时(2)的逆序数比(1)的逆序数少 1;如果 i 与 j 在排列(1)中是顺序,则它们在(2)中构成逆序,这时(2)的逆序数比(1)的逆序数多 1.总之,排列(1)与(2)的逆序数相差 1 个,所以排列(1)与(2)的奇偶性相反.这就证明了:相邻两数的对换改变排列的奇偶性.

下面讨论一般的情况,设对换的两个数 i 与 j 之间还有 s 个数 $k_1,k_2,\cdots,k_s$.即排列:

$$\cdots\ i\ k_1\ k_2\ \cdots\ k_s\ j\ \cdots \tag{3}$$

经过 i 与 j 对换,变为排列

$$\cdots\ j\ k_1\ k_2\ \cdots\ k_s\ i\ \cdots \tag{4}$$

排列(3)变成排列(4)可以通过一系列相邻两数的对换来实现.先把排列(3)经过 $s+1$ 次相邻两数的对换变成排列(5):

$$\cdots\ k_1\ k_2\ \cdots\ k_s\ j\ i\ \cdots \tag{5}$$

再把排列(5)经过 s 次相邻两数的对换变成排列(4).于是总共经过 $2s+1$ 次相邻两数的对换,把排列(3)变成了排列(4).由于一次相邻两数的对换会改变排列的奇偶性,因此 $2s+1$ 次相邻两数的对换会改变排列的奇偶性.所以排列(3)与(4)奇偶相反.这就证明了对换改变排列的奇偶性.|

应用定理 1 可以证明以下重要的事实:

定理 2 在全部 $n!$ 个 n 阶排列中,奇、偶排列的个数相等,各有$\frac{n!}{2}$个.

证明:假设在 $n!$ 个 n 阶排列中有 s 个奇排列,t 个偶排列,下面来证明 $s=t$.

将这 s 个奇排列的头两个数字都对换一下,即将 $a_1a_2\cdots a_n$ 变为 $a_2a_1\cdots a_n$.就得到 s 个偶排列.而且这 s 个排列各不相同.但是偶排列一共有 t 个,所以 $s\leqslant t$.再将 t 个偶排列的头两个数字对换,得

到 t 个不同的奇排列. 因此 $t \leqslant s$. 由此得 $s=t$. 即奇排列的总数与偶排列的总数一样. 因为这两种排列一共有 $n!$ 个,所以它们各有$\frac{n!}{2}$个. |

例 6 在例 2 的 24 个 4 级排列中,第 1 排和第 4 排的 12 个排列是偶排列,而其余 12 个排列是奇排列.

最后,我们来证明一个以后常常用到的结论.

定理 3 任意一个 n 阶排列都可以经过一些对换变成自然序排列,并且所作对换的个数与这个排列有相同的奇偶性.

证明:只要依次将 $1,2,\cdots,n-1$ 经对换换到第 $1,2,\cdots,n-1$ 个位置即可将任一排列变为自然序排列. 又因为自然顺序 $1\ 2\cdots n$ 是一个偶排列,而且对换改变排列的奇偶. 所以将一个奇(偶)排列变到自然排列序,需要经过奇(偶)数次对换. |

例如,把排列 341562 经对换变为自然序排列:

$$341562 \to 143562 \to 123564 \\ \to 123465 \to 123456$$

所作的次数 4 与 $\tau(341562)=6$ 都是偶数.

推论:任意两个 n 级排列都可经过一些对换互变,而且如果这两个排列奇偶相同,则所作的对换次数是偶数;如果这两个排列奇偶相反,则所作的对换次数是奇数.

习 题 1.2(2)

1. 求下列排列的逆序数,并决定其奇偶性:

   4 2 6 7 3 5 1; 1 3 5 7 2 4 6; 5 4 7 8 2 1 3 6; 6 1 4 7 2 8 5 3.

2. 决定 i,j 使

   (1) 2 1 5 i 7 j 9 4 6 为奇排列;

   (2) 3 9 7 2 i 1 5 j 4 为偶排列.

3. 用对换将排列 315694278 变为自然序排列. 写出所作的对换,并由此决定这个排列的奇偶.

1.3 n阶行列式的定义

这一节介绍n阶行列式的定义.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

表示一个n阶行列式.行列式中横排称为**行**.竖排称为**列**.其中元素a_{ij}的第一个下标i表示这个元素位于第i行,称为**行标**;第二个下标j表示这个元素位于第j列,称为**列标**.例如a_{23}表示行列式中第2行第3列处的元素;a_{ij}表示第i行第j列处的元素.

从2阶和3阶行列式的定义可以看出:为了定义一个行列式,需要决定它有哪些项;以及每个项前面所带的正负号.

在给出n阶行列式的定义之前,先来回顾一下2阶和3阶行列式的定义:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \tag{1}$$

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

从2阶和3阶行列式的定义中可以看出,它们都是一些乘积的代数和,而第一项都是由行列式中位于不同行和不同列的元素构成的乘积,并且展开式恰恰就是由所有这种可能的乘积组成.在$n=2$时,由不同行不同列的元素构成的乘积只有$a_{11}a_{22}$与$a_{12}a_{21}$这2项,在$n=3$时也不难看出,只有(2)中的6项,这是2阶和3阶行列式的特征的一个方面;另一方面,每一项乘积都带有符号.这符号是按什么原则决定的呢?在3阶行列式的展开式(2)中,项的

一般形式可以写成

$$a_{1j_1} \quad a_{2j_2} \quad a_{2j_3} \tag{3}$$

其中 $j_1j_2j_3$ 是 1,2,3 的一个排列. 可以看出,当 $j_1j_2j_3$ 是偶排列时,对应的项在(2)中带有正号;当 $j_1j_2j_3$ 是奇排列时带有负号. 2 阶行列式显然也符合这个原则.

上面关于 2 阶和 3 阶行列式的分析对于我们理解一般的定义是有帮助的. 下面给出 n 阶行列式的定义.

定义 4 n 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{4}$$

等于所有取自不同行不同列的 n 个元素的乘积

$$a_{1j_1}a_{2j_2}\cdots a_{nj_n} \tag{5}$$

的代数和,其中 $j_1j_2\cdots j_n$ 是一个 n 阶排列. 每个项(5)的前面带有正负号:当 $j_1j_2\cdots j_n$ 是偶排列时,带正号;当 $j_1j_2\cdots j_n$ 是奇排列时,带负号. 因此,行列式(4)可表示成:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{(j_1j_2\cdots j_n)} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n}, \tag{6}$$

式中 $\sum\limits_{(j_1j_2\cdots j_n)}$ 表示对所有 n 阶排列求和.

(6) 式称为 n 阶行列式的展开式.

容易检验出来,当 $n=2,3$ 时,这个展开公式与前面 2 阶和 3 阶行列式的定义是一致的.

从定义可看出:n 阶行列式是 $n!$ 项的代数和,每个项是不同行不同列的 n 个元素的乘积. 为了不遗漏且不重复地找出这 $n!$ 个项,

可以利用 n 阶排列来写出各个项. 下面用 $n=4$ 的情形来说明.

例 1 写出 4 阶行列式:

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

的展开式.

解: 4 阶排列一共有 $4!=24$ 个,所以 4 阶行列式的展开式中共有 24 项. 根据这 24 个排列的奇偶性(参考上节例 2 和例 6),可以写出 4 阶行列式的展开式为:

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

从这个例子看出:直接应用行列式的定义来计算行列式是一件很麻烦的事. 以下几节将介绍行列式的一些重要性质,以及如何利用这些性质来简化行列式的计算. 为了熟悉和记住行列式的定义,下面先来举一些比较简单的、可以直接应用定义来计算的行列式的例子.

例 2 计算行列式：

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

解：这是一个 4 阶行列式，其展开式中应该有 24 个项，但是由于行列式中有很多零，所以有很多项等于零，只要找出那些不等于零的项就可以了. 因为行列式中一共只有 4 个元素不等于零，而且这 4 个元素刚好位于不同行不同列，所以这个行列式的展开式中只有一个项 $abcd$. 这个项前面所带的符号需要由这 4 个元素的位置决定. 将 a,b,c,d 按行的顺序排好，它们所在的列依次是 4，3，2，1. 所以

$$\begin{vmatrix} 0 & 0 & 0 & a \\ 0 & 0 & b & 0 \\ 0 & c & 0 & 0 \\ d & 0 & 0 & 0 \end{vmatrix} = (-1)^{\tau(4321)}abcd = abcd.$$

在 2 阶和 3 阶行列式中，反对角线（从右上角到左下角这条对角线）上的元素连乘积所成的项前面是带负号的. 这往往使我们得到一个印象，以为行列式中由反对角线上的元素所成的项总是带负号的. 但是上面的例子告诉我们，在 4 阶行列式中反对角线上的元素所成的项前面却是带正号. 读者不妨自己总结一下，n 阶行列式中由反对角线上的元素所成的项前面所带正负号的规律.

例 3 计算行列式：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

解：根据定义，n 阶行列式的项的一般形式是：

$$a_{1j_1}a_{2j_2}\cdots a_{nj_n}.$$

由于在这个行列式的第 n 行中，除 a_{nn}外，其他的元素都等于 0，所以 $j_n\neq n$ 的项都等于零，因而只要考虑 $j_n=n$ 的项即可；再看第 $n-1$ 行：这一行中除去 $a_{n-1,n-1}$及 $a_{n-1,n}$外，其他的元素都等于零，因此，j_{n-1}只有 $n-1$，n 这两个可能. 但因 $j_n=n$，而且 $j_{n-1}\neq j_n$，所以 $j_{n-1}=n-1$. 这样逐步推上去，可知：在展开式中，除去

$$a_{11}a_{22}\cdots a_{nn}$$

这一项外，其他的项都等于 0. 而这一项所在的列成自然顺序，所以这一项带正号. 于是

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}$$

这样的行列式叫做**上三角形行列式**. 这个例子说明：上三角形行列式等于**主对角线**（从左上角到右下角这条对角线）上的元素的乘积. 作为这种行列式的特殊情形，有

$$\begin{vmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_n \end{vmatrix} = a_1a_2\cdots a_n.$$

其中主对角线以外的元素都是零，称为**对角行列式**，它也等于主对角线上的元素的乘积.

同样可以证明下三角形行列式也等于主对角线上元素的乘积，即：

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

例 4 证明

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & 0 & 0 & 0 \\ a_{41} & a_{42} & 0 & 0 & 0 \\ a_{51} & a_{52} & 0 & 0 & 0 \end{vmatrix} = 0.$$

证明：这个行列式的元素满足

$$a_{33} = a_{34} = a_{35} = 0，即 a_{3j_3} = 0，当 j_3 \geqslant 3；$$

$$a_{43} = a_{44} = a_{45} = 0，即 a_{4j_4} = 0，当 j_4 \geqslant 3；$$

$$a_{53} = a_{54} = a_{55} = 0，即 a_{5j_5} = 0，当 j_5 \geqslant 3.$$

考察这个行列式的展开式中的一般项：

$$a_{1j_1}a_{2j_2}a_{3j_3}a_{4j_4}a_{5j_5}$$

如果 j_3,j_4,j_5 中有一个大于 2，那么这个项就等于 0. 但是 j_3,j_4,j_5 各不相等，所以至少有一个大于 2. 这说明这个行列式的每个项都等于零，所以这个行列式等于 0.

以上这些例子都可以作为公式应用. 下面介绍行列式的另一种展开式. 在行列式的定义中，为了决定每一项的正负，把 n 个元素按所在行的先后顺序排列起来，然后根据列标组成排列的奇偶性来决定这一项的正负. 如果给了一般的不同行不同列的 n 个元素的乘积

$$a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n},$$

其中 $i_1i_2\cdots i_n$；$j_1j_2\cdots j_n$ 都是 n 元排列. 在这种情况下，能不能根据

这两个排列的奇偶性直接决定这一项前面所带的正负号呢？事实上，为了根据定义来决定这一项前面的符号，先要把这一项的 n 个元素重新排列，使得它们的行标成自然顺序，即排成：

$$a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}=a_{1j_1'}a_{2j_2'}\cdots a_{nj_n'}.$$

于是这一项前面的符号就是

$$(-1)^{\tau(j_1'j_2'\cdots j_n')}.$$

例 5　决定 4 阶行列式中项 $a_{23}a_{42}a_{14}a_{31}$ 前面所带的符号.

解：

$$a_{23}a_{42}a_{14}a_{31}=a_{14}a_{23}a_{31}a_{42},$$

所以这一项前面所带的符号是：

$$(-1)^{\tau(4312)}=(-1)^5=-1$$

即这一项前面带负号.

为了讨论一般情况，需要分析 $j_1'j_2'\cdots j_n'$ 的奇偶性与 $i_1i_2\cdots i_n$ 及 $j_1j_2\cdots j_n$ 的关系. 把

$$a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}$$

写成

$$a_{1j_1'}a_{2j_2'}\cdots a_{nj_n'},$$

可以经过一系列元素的对换来实现. 在两个元素对换时，元素的行标与列标的排列也都同时作一次对换. 下面用一个具体例子来说明这个事实.

例如例 5 中 4 阶行列式中的一项 $a_{23}a_{42}a_{14}a_{31}$ 按行的自然顺序排列，是 $a_{14}a_{23}a_{31}a_{42}$. 这可以经过一系列对换，把 $a_{23}a_{42}a_{14}a_{31}$ 变成 $a_{14}a_{23}a_{31}a_{42}$. 这时，行标与列标所成的排列也经过了一系列对应的对换. 列表如下：

项　$a_{23}a_{42}a_{14}a_{31}\to a_{14}a_{42}a_{23}a_{31}\to a_{14}a_{23}a_{42}a_{31}\to a_{14}a_{23}a_{31}a_{42}$；

行标 2 4 1 3(奇)→1 4 2 3(偶)→1 2 4 3(奇)→1 2 3 4(偶)；

列标 3 2 4 1(偶)→4 2 3 1(奇)→4 3 2 1(偶)→4 3 1 2(奇).

从上面的过程可以看出：每作一次对换，行标与列标所成的排列同时改变奇偶性. 因此，行标与列标排列的逆序数的和的奇偶性是不变的，这些和总是奇数，所以 $a_{23}a_{42}a_{14}a_{31}$ 前面的正负号可以用 $(-1)^{\tau(\overbrace{2413}^{行标})+\tau(\overbrace{3241}^{列标})}$ 来表示.

一般地，当 $a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}$ 经过一系列对换变成 $a_{1j_1'}a_{2j_2'}\cdots a_{nj_n'}$ 时，每作一次对换，行标与列标所成的排列同时改变奇偶性，所以

$$(-1)^{\tau(i_1i_2\cdots i_n)+\tau(j_1j_2\cdots j_n)} = (-1)^{\tau(12\cdots n)+\tau(j_1'j_2'\cdots j_n')}$$
$$= (-1)^{\tau(j_1'j_2'\cdots j_n')}.$$

这说明
$$a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}$$
前面的正负号是
$$(-1)^{\tau(i_1i_2\cdots i_n)+\tau(j_1j_2\cdots j_n)}.$$

由此可知，如果 n 阶行列式的某一项按列的自然顺序将元素顺次排列：
$$a_{i_11}a_{i_22}\cdots a_{i_nn},$$
那么，这项前面的符号是
$$(-1)^{\tau(i_1i_2\cdots i_n)}.$$
因此，行列式的定义又可以写成
$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{(i_1i_2\cdots i_n)} (-1)^{\tau(i_1i_2\cdots i_n)} a_{i_11}a_{i_22}\cdots a_{i_nn}.$$

上面的结论说明行列式中行、列地位的对称性，这一事实在行列式的讨论中有重要的应用.

习　题　1.3

1. 决定下列各项前面所带的正负号：

　(1) $a_{12}a_{21}a_{34}a_{45}a_{53}$；

　(2) $a_{25}a_{34}a_{51}a_{72}a_{66}a_{17}a_{43}$；

　(3) $a_{13}a_{26}a_{32}a_{54}a_{41}a_{65}$.

2. 写出下面 5 阶行列式中包含 $a_{13}a_{25}$ 并带正号的所有项：

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

3. 求 i,k 使

　(1) $a_{12}a_{2i}a_{3k}a_{51}a_{44}$是 5 阶行列式中带正号的项；

　(2) $a_{21}a_{i4}a_{15}a_{k2}a_{33}$是 5 阶行列式中带负号的项.

4. 计算行列式：

(1) $\begin{vmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 7 & 0 \\ 0 & 3 & 6 & 9 & 0 \\ 4 & 10 & 11 & -5 & 0 \\ 8 & 1 & 3 & 7 & 5 \end{vmatrix}$；　　(2) $\begin{vmatrix} 0 & a_1 & 0 & 0 & 0 \\ 0 & 0 & a_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_3 \\ a_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_5 & 0 \end{vmatrix}$；

(3) $\begin{vmatrix} a_1 & 0 & b_1 & 0 \\ 0 & c_1 & 0 & d_1 \\ a_2 & 0 & b_2 & 0 \\ 0 & c_2 & 0 & d_2 \end{vmatrix}$.

5. 计算 n 阶行列式：

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

1.4 行列式的性质

从行列式的定义计算一个 n 阶行列式，需要计算 $n!$ 个项，而每个项又是 n 个元素的乘积，需要作 $n-1$ 次乘法，所以一共需作 $n!(n-1)$次乘法，当 n 比较大时，$n!(n-1)$就是一个惊人的数目，即使用电子计算机来进行计算也是难以实现的. 因此，必须对行列式作进一步的研究，找出其他切实可行的计算方法.

下面介绍的一些行列式的性质，它们不仅可以用来简化行列式的计算，并且对行列式的一些理论研究，也是极为重要的.

在上一节中，我们讨论了在行列式中，行与列的对称性，由此可得行列式的一个基本性质，设 n 阶行列式：

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

将 D 的行、列互换，得到一个行列式：

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

D^{T} 称为行列式 D 的转置行列式.

性质 1 行列互换，行列式的值不变. 即

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

证明：元素 a_{ij}位于上式右端的第 j 行第 i 列. 将上式右端按

列的自然顺序展开，得

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

$$= \text{左端按 } n \text{ 阶行列式的原定义展开}$$

$$= \text{左端}. \ |$$

这个性质证明了行列式与它的转置行列式相等，即 $D=D^T$，指出了行列式中行、列地位的对称性. 由此可知：行列式中有关行的性质对列也同样成立. 例如关于下三角形行列式的结果，可以应用性质 1 及关于上三角形行列式的结果（上节例 3）来得到而不必直接证明：

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

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

$$= a_{11} a_{22} \cdots a_{nn}$$

性质 2

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ ka_{p1} & ka_{p2} & \cdots & ka_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

这就是说，行列式中某一行的公因子可以提出来. 或者说，用一个

数来乘行列式的某一行(即用此数乘这一行的每个元素)就等于用这个数乘此行列式.

证明：上式左端$=\sum_{(j_1j_2\cdots j_n)}(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}\cdots(ka_{pj_p})\cdots a_{nj_n}$

$$=k\sum_{(j_1j_2\cdots j_n)}(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}\cdots a_{pj_p}\cdots a_{nj_n}$$

=右端. |

关于列也有类似的性质,即

$$\begin{vmatrix} a_{11} & \cdots & ka_{1q} & \cdots & a_{1n} \\ a_{21} & \cdots & ka_{2q} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & ka_{nq} & \cdots & a_{nn} \end{vmatrix} = k\begin{vmatrix} a_{11} & \cdots & a_{1q} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2q} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nq} & \cdots & a_{nn} \end{vmatrix}.$$

这可以由性质 1 和性质 2 加以证明：

$$\begin{vmatrix} a_{11} & \cdots & ka_{1q} & \cdots & a_{1n} \\ a_{21} & \cdots & ka_{2q} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & ka_{nq} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \cdots & \cdots & \cdots & \cdots \\ ka_{1q} & ka_{2q} & \cdots & ka_{nq} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix}$$

$$= k\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1q} & a_{2q} & \cdots & a_{nq} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix} = k\begin{vmatrix} a_{11} & \cdots & a_{1q} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2q} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nq} & \cdots & a_{nn} \end{vmatrix}. \ |$$

上述这个证明中反复地利用了行列式中行列的对称性. 以后讨论行列式的性质时,都是对行来说的. 对于列,都可以用与前述证明类似的方法来证明也有相同的性质,就不再重复了.

从性质 2 可以推出:如果行列式中有一行为零,那么行列式为零.

性质 3

$$
\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\cdots & \cdots & \cdots & \cdots \\
a_{p1}+a'_{p1} & a_{p2}+a'_{p2} & \cdots & a_{pn}+a'_{pn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
= \begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\cdots & \cdots & \cdots & \cdots \\
a_{p1} & a_{p2} & \cdots & a_{pn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
+ \begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\cdots & \cdots & \cdots & \cdots \\
a'_{p1} & a'_{p2} & \cdots & a'_{pn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}.
$$

这就是说，如果行列式中某一行(如第 p 行)是两组数的和，那么这个行列式就等于两个行列式的和.这两个行列式分别以这两组数为这一行(第 p 行)的元素，而除去这一行以外，这两个行列式的其他各行与原来行列式的对应各行都是相同的.

证明

$$
\begin{aligned}
\text{左端} &= \sum_{(j_1j_2\cdots j_n)} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}\cdots(a_{pj_p}+a'_{pj_p})\cdots a_{nj_n} \\
&= \sum_{(j_1j_2\cdots j_n)} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}\cdots a_{pj_p}\cdots a_{nj_n} \\
&\quad + \sum_{(j_1j_2\cdots j_n)} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}\cdots a'_{pj_p}\cdots a_{nj_n} \\
&= \text{右端}.
\end{aligned}
$$

这一性质可以推广到某一行为多组数的和的情形.

例 1 计算

$$
\begin{vmatrix}
2 & 1 & -1 \\
4 & -1 & 1 \\
201 & 102 & -99
\end{vmatrix}.
$$

解：

$$\begin{vmatrix} 2 & 1 & -1 \\ 4 & -1 & 1 \\ 201 & 102 & -99 \end{vmatrix} = \begin{vmatrix} 2 & 1 & -1 \\ 4 & -1 & 1 \\ 200+1 & 100+2 & -100+1 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 1 & -1 \\ 4 & -1 & 1 \\ 200 & 100 & -100 \end{vmatrix} + \begin{vmatrix} 2 & 1 & -1 \\ 4 & -1 & 1 \\ 1 & 2 & 1 \end{vmatrix}$$

$$= 100\begin{vmatrix} 2 & 1 & -1 \\ 4 & -1 & 1 \\ 2 & 1 & -1 \end{vmatrix} + \begin{vmatrix} 2 & 1 & -1 \\ 4 & -1 & 1 \\ 1 & 2 & 1 \end{vmatrix}$$

$$= 0 + (-18) = -18.$$

性质 4 对换行列式中两行的位置，行列式反号. 即

$$\begin{array}{c} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{q1} & a_{q2} & \cdots & a_{qn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \begin{array}{l} \\ \\ (\text{第 } p \text{ 行}) \\ \\ (\text{第 } q \text{ 行}) \\ \\ \\ \end{array} \end{array}$$

$$= -\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{q1} & a_{q2} & \cdots & a_{qn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \begin{array}{l} \\ \\ (\text{第 } p \text{ 行}) \\ \\ (\text{第 } q \text{ 行}) \\ \\ \\ \end{array}$$

证明： 已知

$$\text{左端} = \sum_{(j_1 j_2 \cdots j_n)} (-1)^{\tau(j_1 \cdots j_p \cdots j_q \cdots j_n)} a_{1j_1} \cdots a_{pj_p} \cdots a_{qj_q} \cdots a_{nj_n}.$$

现在 $a_{1j_1}, \cdots, a_{pj_p}, \cdots, a_{qj_q}, \cdots, a_{nj_n}$ 在右端的行列式中仍然是不同行不同列的. 所以它们的乘积 $a_{1j_1} \cdots a_{pj_p} \cdots a_{qj_q} \cdots a_{nj_n}$ 也是右端行列式

的一个项. 但是 a_{pj_p} 在右端位于第 q 行第 j_p 列；a_{qj_q} 在右端位于第 p 行第 j_q 列. 所以这个项的因子在这种顺序下它们的行标与列标所成的排列分别是：

$$\begin{array}{ccccccc} 1 & \cdots & q & \cdots & p & \cdots & n \\ & & (\text{第 } p \text{ 个}) & & (\text{第 } q \text{ 个}) & & \end{array}$$

和

$$\begin{array}{ccccccc} j_1 & \cdots & j_p & \cdots & j_q & \cdots & j_n. \\ & & (\text{第 } p \text{ 个}) & & (\text{第 } q \text{ 个}) & & \end{array}$$

排列 $1\cdots q\cdots p\cdots n$ 是从自然顺序中将 p,q 对换而得的，所以这是一个奇排列. 因此这一项作为右端行列式的展开式中的一项，前面的符号应是：

$$(-1)^{\tau(1\cdots q\cdots p\cdots n)}(-1)^{\tau(j_1\cdots j_p\cdots j_q\cdots j_n)}=-(-1)^{\tau(j_1\cdots j_p\cdots j_q\cdots j_n)},$$

从而

$$\text{右端}=\text{左端}. \mid$$

性质 5 如果行列式中有两行成比例，那么行列式等于零. 即

$$\begin{array}{r} \\ \\ (\text{第 } p \text{ 行}) \\ \\ (\text{第 } q \text{ 行}) \\ \\ \\ \end{array}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ ka_{p1} & ka_{p2} & \cdots & ka_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=0$$

证明：首先来证明一个特殊情形，即 $k=1$ 的情形. 此时，根据性质 4，把第 p,q 两行对换，得：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}\begin{array}{l} \\ \\ (\text{第 } p \text{ 行}) \\ \\ (\text{第 } q \text{ 行}) \\ \\ \\ \end{array}$$

$$
= -\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.
$$

所以
$$
\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0,
$$

即原行列式$=0$.

证明了这个特殊情形，一般情形就容易证了，根据性质 2：

$$
\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ ka_{p1} & ka_{p2} & \cdots & ka_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = k\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = k \cdot 0 = 0.
$$

性质 6 把某一行的倍数加到另一行，行列式不变，即

$$
\begin{array}{l} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{q1} & a_{q2} & \cdots & a_{qn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \begin{array}{l} \\ \\ (\text{第 } p \text{ 行}) \\ \\ (\text{第 } q \text{ 行}) \\ \\ \\ \end{array} \end{array}
$$

$$
=\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\cdots & \cdots & \cdots & \cdots \\
a_{p1}+ka_{q1} & a_{p2}+ka_{q2} & \cdots & a_{pn}+ka_{qn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{q1} & a_{q2} & \cdots & a_{qn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
\begin{array}{l}
\\ \\ (\text{第 } p \text{ 行}) \\ \\ (\text{第 } q \text{ 行}) \\ \\ \\
\end{array}
\quad (p \neq q)
$$

证明:

$$
\text{右端} = \begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\cdots & \cdots & \cdots & \cdots \\
a_{p1} & a_{p2} & \cdots & a_{pn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{q1} & a_{q2} & \cdots & a_{qn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix} + \begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\cdots & \cdots & \cdots & \cdots \\
ka_{q1} & ka_{q2} & \cdots & ka_{qn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{q1} & a_{q2} & \cdots & a_{qn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}
$$

$$
= \begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{n1} \\
\cdots & \cdots & \cdots & \cdots \\
a_{p1} & a_{p2} & \cdots & a_{pn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{q1} & a_{q2} & \cdots & a_{qn} \\
\cdots & \cdots & \cdots & \cdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{vmatrix} + 0 = \text{左端}. \mid
$$

行列式的性质就介绍到这里，在此再强调一遍：上面所讲的性质，是对行而言的；关于列也同样成立. 下面通过一些例题来说明如何应用这些性质计算行列式. 在例题解答中，一般不写出应用了哪一条性质，希望读者把每一步的理由想清楚. 这样做，可以比较快地把这些性质记熟.

例 2 证明：

$$\begin{vmatrix} a+b & b+c & c+a \\ a_1+b_1 & b_1+c_1 & c_1+a_1 \\ a_2+b_2 & b_2+c_2 & c_2+a_2 \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}.$$

证明:

$$\text{左端} = \begin{vmatrix} a & b+c & c+a \\ a_1 & b_1+c_1 & c_1+a_1 \\ a_2 & b_2+c_2 & c_2+a_2 \end{vmatrix} + \begin{vmatrix} b & b+c & c+a \\ b_1 & b_1+c_1 & c_1+a_1 \\ b_2 & b_2+c_2 & c_2+a_2 \end{vmatrix}$$

$$= \begin{vmatrix} a & b+c & c \\ a_1 & b_1+c_1 & c_1 \\ a_2 & b_2+c_2 & c_2 \end{vmatrix} + \begin{vmatrix} b & c & c+a \\ b_1 & c_1 & c_1+a_1 \\ b_2 & c_2 & c_2+a_2 \end{vmatrix}$$

$$= \begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} + \begin{vmatrix} b & c & a \\ b_1 & c_1 & a_1 \\ b_2 & c_2 & a_2 \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}.$$

例 3　计算

$$\begin{vmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{vmatrix}.$$

解: 在这个行列式中,各行元素的和是相同的,都是 $a+3$. 因此,如果逐次把第 2 列,第 3 列,第 4 列都加到第 1 列上,则第 1 列的元素就全等于 $a+3$. 应用性质 2,把第 1 列的公因子 $a+3$ 提出来,就可以把这个行列式化为便于计算的形式. 这是一个常用的方法. 把第 2,3,4 列加到第一列的几个步骤可以一次写出来.

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

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

$$=(a+3)(a-1)^3.$$

在上面的计算中，第 3 个等号也是把 3 个步骤一次写出的：把第一行乘 -1 后加到第 2，3，4 行上. 以后做习题时，也可这样写.

这个例子可以推广到 n 阶行列式的情形.

例 4 计算 n 阶行列式：

$$\begin{vmatrix}a&1&1&\cdots&1\\1&a&1&\cdots&1\\\cdots&\cdots&\cdots&\cdots&\cdots\\1&1&1&\cdots&a\end{vmatrix}.$$

解：

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

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

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

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

这个方法适用于各行或各列之和都相同的行列式.

例 5 计算

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

解：

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

$$= \begin{vmatrix} a^2 & a^2+2a+1 & a^2+4a+4 & a^2+6a+9 \\ b^2 & b^2+2b+1 & b^2+4b+4 & b^2+6b+9 \\ c^2 & c^2+2c+1 & c^2+4c+4 & c^2+6c+9 \\ d^2 & d^2+2d+1 & d^2+4d+4 & d^2+6d+9 \end{vmatrix}$$

$$= \begin{vmatrix} a^2 & 2a+1 & 4a+4 & 6a+9 \\ b^2 & 2b+1 & 4b+4 & 6b+9 \\ c^2 & 2c+1 & 4c+4 & 6c+9 \\ d^2 & 2d+1 & 4d+4 & 6d+9 \end{vmatrix}$$

$$= \begin{vmatrix} a^2 & 2a+1 & 2 & 6 \\ b^2 & 2b+1 & 2 & 6 \\ c^2 & 2c+1 & 2 & 6 \\ d^2 & 2d+1 & 2 & 6 \end{vmatrix} = 0.$$

例 6 计算 n 阶行列式

$$\begin{vmatrix} a_1-b & a_2 & \cdots & a_n \\ a_1 & a_2-b & \cdots & a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_1 & a_2 & \cdots & a_n-b \end{vmatrix}.$$

解：

$$
\begin{vmatrix}
a_1-b & a_2 & \cdots & a_n \\
a_1 & a_2-b & \cdots & a_n \\
\cdots & \cdots & \cdots & \cdots \\
a_1 & a_2 & \cdots & a_n-b
\end{vmatrix}
$$

$$
=\begin{vmatrix}
\sum\limits_{i=1}^{n}a_i-b & a_2 & \cdots & a_n \\
\sum\limits_{i=1}^{n}a_i-b & a_2-b & \cdots & a_n \\
\cdots & \cdots & \cdots & \cdots \\
\sum\limits_{i=1}^{n}a_i-b & a_2 & \cdots & a_n-b
\end{vmatrix}
$$

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

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

$$
=(-1)^{n-1}b^{n-1}\left(\sum_{i=1}^{n}a_i-b\right).
$$

习　题　1.4(1)

1. 计算下列行列式：

(1) $\begin{vmatrix} 12345 & 12245 \\ 67813 & 67913 \end{vmatrix}$;　　(2) $\begin{vmatrix} 246 & 427 & 327 \\ 1014 & 543 & 443 \\ -342 & 721 & 621 \end{vmatrix}$;

(3) $\begin{vmatrix} x & y & x-y \\ y & x-y & x \\ x+y & x & y \end{vmatrix}$;　　(4) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$;

(5) $\begin{vmatrix} 1+a & 1 & 1 & 1 \\ 1 & 1-a & 1 & 1 \\ 1 & 1 & 1+b & 1 \\ 1 & 1 & 1 & 1-b \end{vmatrix}$;

(6) $\begin{vmatrix} a & b & c & d \\ p & q & r & s \\ t & u & v & w \\ la+mp & lb+mq & lc+mr & ld+ms \end{vmatrix}$.

2. 试证

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ a_1 & a_1+a_2 & a_1+a_2+a_3 & a_1+a_2+a_3+a_4 \\ a_1 & 2a_1+a_2 & 3a_1+2a_2+a_3 & 4a_1+3a_2+2a_3+a_4 \\ a_1 & 3a_1+a_2 & 6a_1+3a_2+a_3 & 10a_1+6a_2+3a_3+a_4 \end{vmatrix} = a_1^4.$$

3. 计算 n 阶行列式

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

4. 计算

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

5. 计算 n+1 阶行列式：

$$\begin{vmatrix} a_1 & -a_1 & 0 & \cdots & 0 & 0 \\ 0 & a_2 & -a_2 & \cdots & 0 & 0 \\ 0 & 0 & a_3 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_n & -a_n \\ b & b & b & \cdots & b & b \end{vmatrix}.$$

下面介绍如何应用行列式的性质来解决数字行列式的计算问题.

在 1.3 节例 3 中曾经证明过，上三角形行列式：

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

就等于主对角线上的元素的乘积 $a_{11}a_{22}\cdots a_{nn}$. 因此，可以把一些行列式化为上三角形行列式来计算. 下面先举一个例子：

例 7 计算 $\begin{vmatrix} 1 & 1 & -1 & 3 \\ -1 & -1 & 2 & 1 \\ 2 & 5 & 2 & 4 \\ 1 & 2 & 3 & 2 \end{vmatrix}.$

解：我们把应用行列式性质的说明写在式子的推导和演算后面，请读者对照阅读.

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

$$\underset{②}{=}-\begin{vmatrix}1&1&-1&3\\0&1&4&-1\\0&3&4&-2\\0&0&1&4\end{vmatrix}\underset{③}{=}-\begin{vmatrix}1&1&-1&3\\0&1&4&-1\\0&0&-8&1\\0&0&1&4\end{vmatrix}$$

$$\underset{④}{=}\begin{vmatrix}1&1&-1&3\\0&1&4&-1\\0&0&1&4\\0&0&-8&1\end{vmatrix}\underset{⑤}{=}\begin{vmatrix}1&1&-1&3\\0&1&4&-1\\0&0&1&4\\0&0&0&33\end{vmatrix}=33.$$

① 根据性质 6：把第 1 行加到第 2 行上；第 1 行乘 -2 加到第 3 行上；第 1 行乘 -1 加到第 4 行上.

② 因为变化后的行列式的第 2 行第 2 列的元素等于零，所以必须先把这个位置的元素变成非零元素. 常用的方法，是将两行对换，或将某一行加到第 2 行上. 这里采用两行对换的方法，将第 2 行与第 4 行对换，第 2 行第 2 列位置的元素变为 1，不但以后计算简便，且可以避免出现分数.

③ 第 2 行乘 -3 加到第 3 行上.

④ 这里第 3 行第 3 列的元素原为 -8，本来可以将第 3 行的 $\frac{1}{8}$ 倍加到第 4 行上，使行列式化成上三角形. 但这样会出现分数. 所以先将第 3，4 行对换一下，使得第 3 行第 3 列位置上的元素变成 1，下一步再将第 4 行第 3 列位置上的 -8 化为零，就不会出现分数了. 这个方法在一般情形下是很有用的.

⑤ 将第 3 行的 8 倍加到第 4 行上.

例 8 计算

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

解：在这个行列式中有几个分数，而且它们的分母不同，所以计算时需要通分，这样做不但麻烦，而且容易出错，因此可以先将各行乘上适当的倍数，把分数化为整数，然后再进行计算. 在下面的演算过程中，不再将每一步所作的变换写出，读者可根据等式前后的改变，观察与分析所作的变换. 注意，有时为了方便起见，用的是列变换.

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

$$= \frac{1}{72}\begin{vmatrix} 2 & 1 & 0 & 2 \\ -1 & 0 & 6 & 3 \\ 2 & 0 & 2 & 3 \\ -2 & -2 & 0 & 1 \end{vmatrix} = \frac{-1}{72}\begin{vmatrix} 1 & 2 & 0 & 2 \\ 0 & -1 & 6 & 3 \\ 0 & 2 & 2 & 3 \\ -2 & -2 & 0 & 1 \end{vmatrix}$$

$$= \frac{-1}{72}\begin{vmatrix} 1 & 2 & 0 & 2 \\ 0 & -1 & 6 & 3 \\ 0 & 2 & 2 & 3 \\ 0 & 2 & 0 & 5 \end{vmatrix} = \frac{-1}{72}\begin{vmatrix} 1 & 2 & 0 & 2 \\ 0 & -1 & 6 & 3 \\ 0 & 0 & 14 & 9 \\ 0 & 0 & 12 & 11 \end{vmatrix}$$

$$=\frac{-1}{72}\begin{vmatrix}1 & 2 & 0 & 2\\0 & -1 & 6 & 3\\0 & 0 & 2 & -2\\0 & 0 & 12 & 11\end{vmatrix}=\frac{-1}{36}\begin{vmatrix}1 & 2 & 0 & 2\\0 & -1 & 6 & 3\\0 & 0 & 1 & -1\\0 & 0 & 12 & 11\end{vmatrix}$$

$$=\frac{-1}{36}\begin{vmatrix}1 & 2 & 0 & 2\\0 & -1 & 6 & 3\\0 & 0 & 1 & -1\\0 & 0 & 0 & 23\end{vmatrix}=\frac{23}{36}.$$

从以上两个例子可以看出一个一般的方法：应用性质 2，性质 4 以及性质 6，可以把一个行列式化为上三角形行列式. 从而可以很快地把这个行列式算出来.

对于一个 n 阶行列式，当 n 较大时，一般可以应用电子计算机来计算. 应用电子计算机计算时，主要考虑的是所作的乘法或除法的次数. 所以常常采用下面的步骤来计算：

$$\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\a_{21} & a_{22} & \cdots & a_{2n}\\\cdots & \cdots & \cdots & \cdots\\a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}=a_{11}\begin{vmatrix}1 & \frac{a_{12}}{a_{11}} & \frac{a_{13}}{a_{11}} & \cdots & \frac{a_{1n}}{a_{11}}\\a_{21} & a_{22} & a_{23} & \cdots & a_{2n}\\\cdots & \cdots & \cdots & \cdots & \cdots\\a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn}\end{vmatrix}.$$

$$=a_{11}\begin{vmatrix}1 & \frac{a_{12}}{a_{11}} & \frac{a_{13}}{a_{11}} & \cdots & \frac{a_{1n}}{a_{11}}\\0 & a_{22}-a_{21}\cdot\frac{a_{12}}{a_{11}} & a_{23}-a_{21}\cdot\frac{a_{13}}{a_{11}} & \cdots & a_{2n}-a_{21}\cdot\frac{a_{1n}}{a_{11}}\\0 & a_{32}-a_{31}\cdot\frac{a_{12}}{a_{11}} & a_{33}-a_{31}\cdot\frac{a_{13}}{a_{11}} & \cdots & a_{3n}-a_{31}\cdot\frac{a_{1n}}{a_{11}}\\\cdots & \cdots & \cdots & \cdots & \cdots\\0 & a_{n2}-a_{n1}\cdot\frac{a_{12}}{a_{11}} & a_{n3}-a_{n1}\cdot\frac{a_{13}}{a_{11}} & \cdots & a_{nn}-a_{n1}\cdot\frac{a_{1n}}{a_{11}}\end{vmatrix}$$

$$=a_{11}\begin{vmatrix}1 & 0 & 0 & \cdots & 0\\ 0 & a'_{22} & a'_{23} & \cdots & a'_{2n}\\ 0 & a'_{32} & a'_{33} & \cdots & a'_{3n}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & a'_{n2} & a'_{n3} & \cdots & a'_{nn}\end{vmatrix}.$$

这里假设了 $a_{11}\neq 0$. 如果 $a_{11}=0$，可以先将某两行(或某两列)对换，使 $a_{11}\neq 0$. 从上面的计算可以看出：经过 $n(n-1)$ 次乘法或除法，并与第 $2,3,\cdots,n$ 行作加法或减法运算，可以把 n 阶行列式中主对角线以下的第 1 列的元素全部化成零. 这样逐步进行下去，经过

$$n(n-1)+(n-1)(n-2)+\cdots+3\cdot 2+2\cdot 1$$
$$=(n^3-n)/3$$

次乘法或除法并分别与第 $2,3,\cdots,n$ 行作加法或减法运算后，就可以将 n 阶行列式化为上三角形行列式，而且主对角线上的元素除了最后一个外全等于 1. 不过，在行列式外面还有 $n-1$ 个因子，因此还要作 $n-1$ 次乘法. 所以总共需要作的乘法与除法的次数是 $(n^3+2n-3)/3$. 当 n 比较大时，这个数字比起 $n!$ 来，是小得无法比较的. 例如，当 $n=20$ 时，在本节开头曾提到，如果按照行列式的定义来逐项计算 20 阶行列式的话，即使应用快速的电子计算机，也是难以做到的. 但是，如果用把行列式化为上三角形行列式来计算，那么 $(20^3+2\times 20-3)/3=2679$，即只需做 2679 次乘法，应用电子计算机，几乎一秒种就可以算出来了. 在 n 比较小时，用这种方法，即使手算一个 n 阶行列式，也是不难的.

习　题　1.4(2)

计算下列行列式：

1. $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix}$.　　2. $\begin{vmatrix} 1 & 4 & 9 & 16 \\ 4 & 9 & 16 & 25 \\ 9 & 16 & 25 & 36 \\ 16 & 25 & 36 & 49 \end{vmatrix}$.

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

5. $\begin{vmatrix} 1 & \frac{1}{2} & 0 & 1 & -1 \\ 2 & 0 & -1 & 1 & 1 \\ 3 & 2 & 1 & \frac{1}{2} & -\frac{1}{2} \\ 1 & 0 & 1 & -1 & 1 \\ 1 & -1 & 0 & 1 & 2 \end{vmatrix}$.

1.5　行列式按一行(列)展开公式

从上节介绍的 n 阶行列式的性质及计算知道:行列式的阶数较低时,计算就比较容易.因此我们自然会想到:能否把一个阶数较高的行列式化成几个阶数较低的行列式来计算?我们知道:3 阶行列式是可以用 2 阶行列式表示的:

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

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

对于一般的 n 阶行列式也可以做到这一点,也就是说,可以将一个 n 阶行列式用一些 $n-1$ 阶行列式来表示.本节就论述这个问题.

首先引入下列定义：

定义 5 在 n 阶行列式

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

中，划去元素 a_{ij}所在的第 i 行第 j 列，剩下的元素按原来的排法，构成一个 $n-1$ 阶行列式

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

称为元素 a_{ij}的余子式，记为 M_{ij}.

例如，对于 3 阶行列式

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

各个元素的余子式分别为：

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

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

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

3 阶行列式 D 可以通过各行的余子式来表示：

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

也可以用各列的余子式来表示：

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

从以上等式看出：M_{ij}前面的符号，有时正，有时负．为了弄清这个问题，引入下述定义：

定义 6 令

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

A_{ij}称为元素 a_{ij}的**代数余子式**.

应用代数余子式的概念，3 阶行列式可以表示成：

$$D = a_{i1}A_{i1} + a_{i2}A_{i2} + a_{i3}A_{i3} \quad (i = 1,2,3)$$

或 $$D = a_{1j}A_{1j} + a_{2j}A_{2j} + a_{3j}A_{3j} \quad (j = 1,2,3).$$

在证明 n 阶行列式也可以用代数余子式来表示之前，为了熟悉与记住余子式和代数余子式的概念，先来举一些例子.

例 1 求

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

的余子式 M_{11}, M_{12}, M_{13}和代数余子式 A_{11}, A_{12}, A_{13}，并求 D.

解：$M_{11} = \begin{vmatrix} 2 & 0 \\ 3 & 2 \end{vmatrix} = 4$，　$M_{12} = \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} = 2$，

$$M_{13} = \begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix} = 5;$$

$A_{11} = (-1)^{1+1}M_{11} = 4$，$A_{12} = (-1)^{1+2}M_{12} = -2$，

$A_{13} = (-1)^{1+3}M_{13} = 5$；

$$D = 1 \cdot A_{11} + 0 \cdot A_{12} + (-1) \cdot A_{13} = -1.$$

例 2 求

$$\begin{vmatrix} a & b & c \\ 1 & 2 & 0 \\ -1 & 3 & 2 \end{vmatrix}.$$

解：这个行列式与例 1 的行列式 D 除去第一行外，其他位置上的元素都是相同的.因此，它的第一行的余子式与 D 的第一行的余子式也是相同的，因而它的第一行的代数余子式也与 D 的第一行的代数余子式相同.所以

$$\begin{vmatrix} a & b & c \\ 1 & 2 & 0 \\ -1 & 3 & 2 \end{vmatrix} = aA_{11} + bA_{12} + cA_{13} = 4a - 2b + 5c.$$

下面将证明如何用代数余子式表示 n 阶行列式的展开式.

定理 4 n 阶行列式

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

等于它任意一行的所有元素与它们的对应代数余子式的乘积的和，即

$$D = a_{k1}A_{k1} + a_{k2}A_{k2} + \cdots + a_{kn}A_{kn} \quad (k = 1,2,\cdots,n).$$

证明：(1) 首先讨论第 n 行除 a_{nn} 外，其余元素全等于零，即 $a_{n1}=a_{n2}=\cdots=a_{n,n-1}=0$ 的情形，证明

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

根据 n 阶行列式的定义，由于 $a_{n1}=a_{n2}=\cdots=a_{n,n-1}=0$，有

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

$$=\sum_{(j_1j_2\cdots j_{n-1}n)}(-1)^{\tau(j_1j_2\cdots j_{n-1}n)}a_{1j_1}a_{2j_2}\cdots a_{n-1\,j_{n-1}}a_{nn}$$

对于 n 阶排列 $j_1j_2\cdots j_{n-1}n$，$j_1j_2\cdots j_{n-1}$ 是一个 $n-1$ 阶排列，而且

$$\tau(j_1j_2\cdots j_{n-1}n)=\tau(j_1j_2\cdots j_{n-1}).$$

于是

$$D=a_{nn}\sum_{(j_1j_2\cdots j_{n-1})}(-1)^{\tau(j_1j_2\cdots j_{n-1})}a_{1j_1}a_{2j_2}\cdots a_{n-1\,j_{n-1}}$$

式中 $\sum\limits_{(j_1j_2\cdots j_{n-1})}$ 表示对所有 $n-1$ 阶排列求和，因此

$$D=a_{nn}M_{nn}=a_{nn}A_{nn}.$$

(2) 讨论第 k 行除 $a_{kj}(k,j=1,2,\cdots,n)$ 外，其余元素全等于零的情形，证明

$$\begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k-1,1} & \cdots & a_{k-1,j-1} & a_{k-1,j} & a_{k-1,j+1} & \cdots & a_{k-1,n} \\ 0 & \cdots & 0 & a_{kj} & 0 & \cdots & 0 \\ a_{k+1,1} & \cdots & a_{k+1,j-1} & a_{k+1,j} & a_{k+1,j+1} & \cdots & a_{k+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}=a_{kj}A_{kj}$$

我们设法改变行列式的行列次序，使 a_{kj} 位于第 n 行第 n 列的位置，并且保持 a_{kj} 的余子式不变，从而把情况(2)化为情况(1)。为此，把 D 的第 k 行依次与第 $k+1$ 行，第 $k+2$ 行，$\cdots$，第 $n-1$ 行对换。这样，一共进行了 $n-k$ 次两行互换的步骤，就把第 k 行换到第 n 行的位置，再将第 j 列依次与第 $j+1$ 列，第 $j+2$ 列，$\cdots$，第 $n-1$ 列对换，一共进行了 $n-j$ 次两列互换的步骤，就把 a_{kj} 换到第 n 行，第 n 列的位置上，因此

$$D=(-1)^{(n-k)+(n-j)}\begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} & a_{1j} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k-1,1} & \cdots & a_{k-1,j-1} & a_{k-1,j+1} & \cdots & a_{k-1,n} & a_{k-1,j} \\ a_{k+1,1} & \cdots & a_{k+1,j-1} & a_{k+1,j+1} & \cdots & a_{k+1,n} & a_{k+1,j} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} & a_{nj} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & a_{kj} \end{vmatrix}$$

$$=(-1)^{(2n-k-j)}a_{kj}\begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k-1,1} & \cdots & a_{k-1,j-1} & a_{k-1,j+1} & \cdots & a_{k-1,n} \\ a_{k+1,1} & \cdots & a_{k+1,j-1} & a_{k+1,j+1} & \cdots & a_{k+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

$$=(-1)^{k+j}a_{kj}M_{kj}=a_{kj}A_{kj}.$$

（3）最后证明一般情形，把 D 表成

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1}+0+\cdots+0 & 0+a_{k2}+0+\cdots+0 & \cdots & 0+\cdots+0+a_{kn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

应用行列式的性质 3，即得

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a_{k2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}+\cdots+\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$=a_{k1}A_{k1}+a_{k2}A_{k2}+\cdots+a_{kn}A_{kn}.\quad(k=1,2,\cdots,n)$$ |

这个定理今后时常会用到．通常称其为行列式按一行（第 k

行)展开的公式.

由于行列式中行与列的对称性,所以,同样也可以将行列式按一列展开,即

定理 4′ n 阶行列式

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

等于它任意一列的所有元素与它们的对应代数余子式的乘积的和,即

$$D=a_{1l}A_{1l}+a_{2l}A_{2l}+\cdots+a_{nl}A_{nl}\quad(l=1,2,\cdots,n).$$

从代数余子式的定义以及一些例题和习题可以看到:a_{ij}的代数余子式只与 a_{ij}所在的位置有关,而与 a_{ij}本身的数值无关. 利用这一点,可以证明关于代数余子式的另一个重要的性质,即

定理 5 n 阶行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

中某一行(列)的每个元素与另一行(列)相应元素的代数余子式的乘积的和等于零. 即

当 $k\neq i$ 时,

$$a_{k1}A_{i1}+a_{k2}A_{i2}+\cdots+a_{kn}A_{in}=0.$$

证明:在行列式

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \begin{array}{l} \\ \\ \\ (\text{第 } i \text{ 行}) \\ \\ (\text{第 } k \text{ 行}) \\ \\ \\ \end{array}$$

中将第 i 行的元素都换成第 $k(k\neq i)$ 行的元素，得到另一个行列式

$$D_0 = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \begin{array}{l} \\ \\ \\ (\text{第 } i \text{ 行}) \\ \\ (\text{第 } k \text{ 行}) \\ \\ \\ \end{array}$$

显然，D_0 的第 i 行的代数余子式与 D 的第 i 行的代数余子式是完全一样的．将 D_0 按第 i 行展开，得

$$D_0 = a_{k1}A_{i1} + a_{k2}A_{i2} + \cdots + a_{kn}A_{in}.$$

但是 D_0 中有两行元素相同，所以 $D_0=0$. 因此

$$a_{k1}A_{i1} + a_{k2}A_{i2} + \cdots + a_{kn}A_{in} = 0 \quad (i \neq k).$$

关于列也有相应的结果，即当 $l\neq j$ 时，

$$a_{1l}A_{1j} + a_{2l}A_{2j} + \cdots + a_{nl}A_{nj} = 0.$$

将定理 4 与定理 5 归纳起来，应用连加号，可以简写成：

$$\sum_{s=1}^{n} a_{ks}A_{is} = \begin{cases} D, & \text{当 } k = i \text{ 时；} \\ 0, & \text{当 } k \neq i \text{ 时.} \end{cases}$$

$$\sum_{s=1}^{n} a_{sl}A_{sj} = \begin{cases} D, & \text{当 } l = j \text{ 时;} \\ 0, & \text{当 } l \neq j \text{ 时.} \end{cases}$$

这两组公式是很重要的，下一节用行列式解线性方程组时就要用到.

行列式按一行(列)展开的公式可以用来计算 n 阶行列式，但是直接应用这两组公式只是把一个 n 阶行列式的计算换成计算 n 个 $n-1$ 阶行列式，计算量不一定减少许多. 而只有当行列式中某一行或某一列含有较多的零时，应用这两组公式才真正有意义.

以后我们将会看到关于行列式按一行(列)展开的公式不仅可用来简化行列式的计算，在理论上也有很重要的应用.

习 题 1.5

1. 求下述行列式的全部余子式及代数余子式：

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

2. 求下述行列式的第 1 列元素的代数余子式：

$$\begin{vmatrix} a & 1 & 2 & 3 \\ b & -1 & 0 & 1 \\ c & 0 & 2 & 3 \\ d & 1 & -1 & -2 \end{vmatrix}.$$

3. 求下述行列式的第 1 列元素的代数余子式：

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

4. 计算下列行列式中所有元素的代数余子式之和.

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

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

1.6 行列式的计算

在这一节，我们通过例题来说明利用行列式的定义、性质和一行(列)展开公式计算行列式的常用方法及主要技巧. 行列式的计算是一个专门的课题，有很多理论和计算方法，这里只是结合本课程的要求，介绍一些基本的方法.

例 1 计算行列式：

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

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

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

$$=-(-1)^{1+1}\begin{vmatrix} 1 & 6 \\ 5 & -1 \end{vmatrix} = 31.$$

例 2 计算 n 阶行列式

$$\begin{vmatrix} a & b & 0 & \cdots & 0 & 0 \\ 0 & a & b & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a & b \\ b & 0 & 0 & \cdots & 0 & a \end{vmatrix}.$$

解：将此行列式按第 1 列展开，得

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

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

例 3 试证

$$D = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j).$$

这个行列式叫做范德蒙(Vandermonde)行列式. 这个例题说明 n 阶范德蒙行列式等于 $a_1, a_2, \cdots, a_n$ 这 n 个数的所有可能的差 $a_i - a_j (1 \leqslant j < i \leqslant n)$ 的乘积.

证明：对 n 用归纳法来证明这个公式.

当 $n=2$ 时，

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

结论是对的. 假设对于 $n-1$ 阶的范德蒙行列式结论成立，现在来看 n 阶的情形：

在 D 中，从第 n 行减去第 $n-1$ 行的 a_1 倍，再从第 $n-1$ 行减去第 $n-2$ 行的 a_1 倍……依次由下而上地从每一行减去它上一行的 a_1 倍，得

$$D = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ 0 & a_2^2 - a_1a_2 & a_3^2 - a_1a_3 & \cdots & a_n^2 - a_1a_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_2^{n-1} - a_1a_2^{n-2} & a_3^{n-1} - a_1a_3^{n-2} & \cdots & a_n^{n-1} - a_1a_n^{n-2} \end{vmatrix}$$

$$= \begin{vmatrix} a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ a_2^2 - a_1a_2 & a_3^2 - a_1a_3 & \cdots & a_n^2 - a_1a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_2^{n-1} - a_1a_2^{n-2} & a_3^{n-1} - a_1a_3^{n-2} & \cdots & a_n^{n-1} - a_1a_n^{n-2} \end{vmatrix},$$

从第 1 列提出公因子 a_2-a_1，从第 2 列提出公因子 a_3-a_1，…，从最后一列提出公因子 a_n-a_1，得

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

后面这个行列式是一个 $n-1$ 阶范德蒙行列式，根据归纳法假设，它等于所有可能的差 $a_i-a_j(2\leqslant j<i\leqslant n)$ 的乘积；而包含 a_1 的差全在前面出现了. 从而

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

即结论对 n 阶范德蒙行列式也成立. 根据数学归纳法原理，等式普遍成立. ∎

例 4 计算 n 阶行列式

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

解：把行列式的第 2 列到第 n 列都加到第 1 列上，并提出公因子，得

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

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

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

$$=\frac{n(n+1)}{2}\begin{vmatrix} 1 & 1 & \cdots & 1-n \\ 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1-n & 1 & \cdots & 1 \end{vmatrix} (n-1\text{阶})$$

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

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

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

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

例 5 证明

$$\begin{vmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-2} & x+a_{n-1} \end{vmatrix}$$

$$=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0.$$

证明：对行列式的阶数作数学归纳法. 当 $n=2$ 时

$$\begin{vmatrix} x & -1 \\ a_0 & x+a_1 \end{vmatrix}=x^2+a_1x+a_0,$$

等式成立.

假设对 $n-1$ 阶行列式等式成立，则对 n 阶行列式

$$\begin{vmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-2} & x+a_{n-1} \end{vmatrix}$$

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

$$+a_0(-1)^{n+1}\begin{vmatrix} -1 & 0 & \cdots & 0 & 0 \\ x & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & x & -1 \end{vmatrix}$$

$$=x(x^{n-1}+a_{n-1}x^{n-2}+\cdots+a_2x+a_1)$$

$$+a_0(-1)^{n+1}(-1)^{n-1}$$

$$=x^n+a_{n-1}x^{n-1}+\cdots+a_2x^2+a_1x+a_0,$$

等式也成立.

根据归纳法原理，等式普遍成立.

例 6 试证

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{l1} & \cdots & c_{lk} & b_{l1} & \cdots & b_{ll} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{l1} & \cdots & b_{ll} \end{vmatrix}.$$

证明：对 k 用数学归纳法证明. 当 $k=1$ 时，上式左端成为：

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ c_{11} & b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots & \cdots \\ c_{l1} & b_{l1} & \cdots & b_{ll} \end{vmatrix},$$

按第 1 行展开，就得到所需结论.

假设对于 $k=s-1$，即左端的左上角是一个 $s-1$ 阶行列式时等式成立. 现在来看 $k=s$ 的情形：按第 1 行展开，得

$$\begin{vmatrix} a_{11} & \cdots & a_{1s} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{s1} & \cdots & a_{ss} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1s} & b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{l1} & \cdots & c_{ls} & b_{l1} & \cdots & b_{ll} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2s} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{s2} & \cdots & a_{ss} & 0 & \cdots & 0 \\ c_{12} & \cdots & c_{1s} & b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{l2} & \cdots & c_{ls} & b_{l1} & \cdots & b_{ll} \end{vmatrix} + \cdots$$

$$+ (-1)^{1+i} a_{1i} \begin{vmatrix} a_{21} & \cdots & a_{2,i-1} & a_{2,i+1} & \cdots & a_{2s} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{s1} & \cdots & a_{s,i-1} & a_{s,i+1} & \cdots & a_{ss} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1,i-1} & c_{1,i+1} & \cdots & c_{1s} & b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{l1} & \cdots & c_{l,i-1} & c_{l,i+1} & \cdots & c_{ls} & b_{l1} & \cdots & b_{ll} \end{vmatrix} + \cdots$$

$$+ (-1)^{1+s} a_{1s} \begin{vmatrix} a_{21} & \cdots & a_{2,s-1} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{s1} & \cdots & a_{s,s-1} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1,s-1} & b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{l1} & \cdots & c_{l,s-1} & b_{l1} & \cdots & b_{ll} \end{vmatrix}$$

$$= a_{11}\begin{vmatrix} a_{22} & \cdots & a_{2s} \\ \cdots & \cdots & \cdots \\ a_{s2} & \cdots & a_{ss} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{l1} & \cdots & b_{ll} \end{vmatrix} + \cdots$$

$$+ (-1)^{1+i}a_{1i}\begin{vmatrix} a_{21} & \cdots & a_{2,i-1} & a_{2,i+1} & \cdots & a_{2s} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{s1} & \cdots & a_{s,i-1} & a_{s,i+1} & \cdots & a_{ss} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{l1} & \cdots & b_{ll} \end{vmatrix} + \cdots$$

$$+ (-1)^{1+s}a_{1s}\begin{vmatrix} a_{21} & \cdots & a_{2,s-1} \\ \cdots & \cdots & \cdots \\ a_{s1} & \cdots & a_{s,s-1} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{l1} & \cdots & b_{ll} \end{vmatrix}$$

$$= \left(a_{11}\begin{vmatrix} a_{22} & \cdots & a_{2s} \\ \cdots & \cdots & \cdots \\ a_{s2} & \cdots & a_{ss} \end{vmatrix} + \cdots \right.$$

$$+ (-1)^{1+i}a_{1i}\begin{vmatrix} a_{21} & \cdots & a_{2,i-1} & a_{2,i+1} & \cdots & a_{2s} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{s1} & \cdots & a_{s,i-1} & a_{s,i+1} & \cdots & a_{ss} \end{vmatrix} + \cdots$$

$$\left. + (-1)^{1+s}a_{1s}\begin{vmatrix} a_{21} & \cdots & a_{2,s-1} \\ \cdots & \cdots & \cdots \\ a_{s1} & \cdots & a_{s,s-1} \end{vmatrix} \right) \cdot \begin{vmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{l1} & \cdots & b_{ll} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & \cdots & a_{1s} \\ \cdots & \cdots & \cdots \\ a_{s1} & \cdots & a_{ss} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{l1} & \cdots & b_{ll} \end{vmatrix}.$$

这里第 2 个等号是应用归纳法假设，最后一步是根据行列式按一行展开的公式. 这说明当行列式的左上角是一个 s 阶行列式时，结论也成立.

由归纳法原理可知，命题普遍成立.

例 7 证明

$$\begin{vmatrix} 0 & a_{12} & a_{13} & a_{14} & a_{15} \\ -a_{12} & 0 & a_{23} & a_{24} & a_{25} \\ -a_{13} & -a_{23} & 0 & a_{34} & a_{35} \\ -a_{14} & -a_{24} & -a_{34} & 0 & a_{45} \\ -a_{15} & -a_{25} & -a_{35} & -a_{45} & 0 \end{vmatrix} = 0$$

证明：记

$$D = \begin{vmatrix} 0 & a_{12} & a_{13} & a_{14} & a_{15} \\ -a_{12} & 0 & a_{23} & a_{24} & a_{25} \\ -a_{13} & -a_{23} & 0 & a_{34} & a_{35} \\ -a_{14} & -a_{24} & -a_{34} & 0 & a_{45} \\ -a_{15} & -a_{25} & -a_{35} & -a_{45} & 0 \end{vmatrix},$$

将 D 的每一行分别乘以 -1 得：

$$(-1)^5 D = \begin{vmatrix} 0 & -a_{12} & -a_{13} & -a_{14} & -a_{15} \\ a_{12} & 0 & -a_{23} & -a_{24} & -a_{25} \\ a_{13} & a_{23} & 0 & -a_{34} & -a_{35} \\ a_{14} & a_{24} & a_{34} & 0 & -a_{45} \\ a_{15} & a_{25} & a_{35} & a_{45} & 0 \end{vmatrix}$$

$$= D' = D$$

所以　$-D=D$，$D=0$.

例 7 中 D 的元素 a_{ij}满足 $a_{ij}=-a_{ji}$，这样的行列式称为反对称行列式. 可以将例 7 中的结果推广而得到：奇数阶反对称行列式等于零. 可用例 7 中的方法同样证明. 关于偶数阶反对称行列式有这样的结论：偶数阶反对称行列式等于一个平方，例如：

$$\begin{vmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{vmatrix} = a_{12}^2;$$

$$\begin{vmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{vmatrix} = (a_{12}a_{34} + a_{14}a_{23} - a_{13}a_{24})^2.$$

这些结论仅供参考，不属本课程的要求.

上面的一些例子，不仅介绍了一些常用的计算行列式的方法，有些例子还可作为公式应用.

例 8 计算行列式

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

解：这是一个范德蒙行列式，所以

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

$$=12.$$

例 9 设 a_1, a_2, a_3, a_4 是各不相同的数，求证

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ a_1 & a_2 & a_3 & a_4 \\ a_1^2 & a_2^2 & a_3^2 & a_4^2 \\ a_1^3 & a_2^3 & a_3^3 & a_4^3 \end{vmatrix} \neq 0.$$

证明：

$$D = \begin{vmatrix} 1 & 1 & 1 & 1 \\ a_1 & a_2 & a_3 & a_4 \\ a_1^2 & a_2^2 & a_3^2 & a_4^2 \\ a_1^3 & a_2^3 & a_3^3 & a_4^3 \end{vmatrix}$$

$$= (a_2 - a_1)(a_3 - a_1)(a_4 - a_1)(a_3 - a_2)(a_4 - a_2)(a_4 - a_3).$$

因为 a_1, a_2, a_3, a_4 各不相同，所以 $a_i - a_j \neq 0 \quad (i > j)$，因此 $D \neq 0$.

例 10 计算行列式

$$\begin{vmatrix} 1 & 2 & 0 & 0 \\ 4 & 5 & 0 & 0 \\ 21 & 31 & 8 & -7 \\ 15 & 16 & -5 & 4 \end{vmatrix}.$$

解：

$$\begin{vmatrix} 1 & 2 & 0 & 0 \\ 4 & 5 & 0 & 0 \\ 21 & 31 & 8 & -7 \\ 15 & 16 & -5 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \begin{vmatrix} 8 & -7 \\ -5 & 4 \end{vmatrix}$$

$$= (-3) \times (-3) = 9.$$

习 题 1.6

1. 计算行列式

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

(3) $\begin{vmatrix} \frac{1}{2} & 0 & -\frac{1}{3} & 1 \\ 2 & -1 & \frac{1}{2} & -1 \\ 3 & 2 & 1 & 0 \\ -2 & 3 & 2 & 1 \end{vmatrix}$；　(4) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & -2 \\ 1 & 1 & 4 & 4 \\ 1 & -1 & 8 & -8 \end{vmatrix}$

2. 计算

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

3. 计算

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

4. 计算 n 阶行列式

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

5. 计算 n 阶行列式

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

6. 计算 $2n$ 阶行列式

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

内 容 提 要

1. n 阶排列

(1) 基本概念：排列、逆序、逆序数、排列的奇偶性.

(2) 主要结论：

1) n 阶排列一共有 $n!$ 个，其中奇、偶排列各有$\frac{n!}{2}$个；

2) 对换改变排列的奇偶；

3) 任意一个 n 阶排列都可以经过一些对换变成自然顺序，并且所作对换的个数与这个排列有相同的奇偶性.

2. n 阶行列式

(1) n 阶行列式的定义：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$
$$= \sum_{(j_1j_2\cdots j_n)} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n}$$
$$= \sum_{(i_1i_2\cdots i_n)} (-1)^{\tau(i_1i_2\cdots i_n)} a_{i_11}a_{i_22}\cdots a_{i_nn}.$$

(2) 行列式的性质：

性质 1　行列互换，行列式不变；

性质 2　用一个数乘行列式的某一行(列)就等于用这个数乘此行列式；

性质 3　如果行列式中第 i 行(列)的元素是两组数的和，那么这个行列式就等于两个行列式的和，这两组数分别是这两个行列式第 i 行(列)的元素，除去第 i 行(列)外，这两个行列式其他各行(列)的元素与原行列式的元素都是相同的；

性质 4　对换行列式中两行(列)的位置，行列式反号；

性质 5　如果行列式中有两行(列)成比例，那么行列式等于零；

性质 6　把行列式的某一行(列)的倍数加到另一行(列)上，行列式不变.

(3) 行列式按某一行(列)展开：

1) 两个重要概念：

余子式 M_{ij}：在 n 阶行列式 D 中，划去元素 a_{ij}所在的行列，剩下的元素按原来的排列构成的 $n-1$ 阶行列式称为元素 a_{ij}的余子式，记为 M_{ij}.

代数余子式 $A_{ij}=(-1)^{i+j}M_{ij}$.

2) 行列式按某一行(列)展开的公式：

$$\sum_{s=1}^{n} a_{ks}A_{is} = \begin{cases} D, & \text{当 } k=i \\ 0, & \text{当 } k \neq i; \end{cases}$$

$$\sum_{s=1}^{n} a_{sl}A_{sj} = \begin{cases} D, & \text{当 } l=j \\ 0, & \text{当 } l \neq j. \end{cases}$$

(4) 行列式的计算

1) 化成上三角形行列式

2) 应用按一行(列)展开公式

3) 应用公式

范德蒙行列式(见 1.6 节例 3)；

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{l1} & \cdots & c_{lk} & b_{l1} & \cdots & b_{ll} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{l1} & \cdots & b_{ll} \end{vmatrix}$$

复习题 1

1. 解下列一次方程组：

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

(2) $\begin{cases} x_1+x_2+x_3=1, \\ ax_1+bx_2+cx_3=d, \\ a^2x_1+b^2x_2+c^2x_3=d^2. \end{cases}$ （a,b,c 各不相同）

2. 计算行列式：

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

(3) $\begin{vmatrix} 3 & 6 & 5 & 6 & 4 \\ 5 & 9 & 7 & 8 & 6 \\ 6 & 12 & 13 & 9 & 7 \\ 4 & 6 & 6 & 5 & 4 \\ 2 & 5 & 4 & 5 & 3 \end{vmatrix}$.

3. 计算行列式：

(1) $\begin{vmatrix} a & 1 & 0 & 0 \\ -1 & b & 1 & 0 \\ 0 & -1 & c & 1 \\ 0 & 0 & -1 & d \end{vmatrix}$；　　(2) $\begin{vmatrix} a & b & c & 1 \\ b & c & a & 1 \\ c & a & b & 1 \\ \frac{b+c}{2} & \frac{a+c}{2} & \frac{a+b}{2} & 1 \end{vmatrix}$.

4. 计算行列式

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

5. 计算行列式

$$\begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 2 & -2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & n-1 & -(n-1) \end{vmatrix}.$$

6. 计算行列式

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

7. 计算行列式

$$\begin{vmatrix} a_0 & 1 & 1 & \cdots & 1 \\ 1 & a_1 & 0 & \cdots & 0 \\ 1 & 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & 0 & \cdots & a_n \end{vmatrix}$$

$$(a_1 \cdots a_n \neq 0).$$

8. 计算 $n(n>1)$阶行列式:

$$\begin{vmatrix} 1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 1 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1 \end{vmatrix} (n\text{ 阶}).$$

9. 试证

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

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

10. 试证 n 阶行列式

$$\begin{vmatrix} a+b & ab & 0 & \cdots & 0 & 0 \\ 1 & a+b & ab & \cdots & 0 & 0 \\ 0 & 1 & a+b & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & a+b \end{vmatrix} = \frac{a^{n+1}-b^{n+1}}{a-b},$$

其中 $a\neq b$.

第 2 章　线性方程组

线性方程组是线性代数的基本内容，在数学的其他分支、自然科学、工程技术以及生产实际中都经常用到，是数学中一个非常重要的基础理论.

一般线性方程组是指形式为：

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

的方程组，其中 $x_1, x_2, \cdots, x_n$ 代表 n 个未知量，s 是方程的个数，a_{ij} $(i=1,2,\cdots,s; j=1,2,\cdots,n)$ 称为方程组的**系数**，$b_j(j=1,2,\cdots,s)$ 称为**常数项**. 系数 a_{ij} 的第一个下标 i 表示它在第 i 个方程，第二个下标 j 表示它是 x_j 的系数. 一般情况下，未知量的个数与方程的个数不一定相等. 因为(1)式包含 n 个未知量，所以称为 n 元**线性方程组**.

线性方程组的一个**解**是指由 n 个数 $c_1, c_2, \cdots, c_n$ 组成的有序数组 $(c_1, c_2, \cdots, c_n)$，当 $x_1, x_2, \cdots, x_n$ 分别用 $c_1, c_2, \cdots, c_n$ 代入后，式(1)中每个等式都成为恒等式. 方程组(1)的解的全体称为它的**解集合**. 如果两个线性方程组有相同的解集合，就称它们是**同解的**.

为了求解一个线性方程组，必须讨论以下一些问题：

(1) 这个方程组有没有解？

(2) 如果这个方程组有解，有多少个解？

(3) 在方程组有解时，解之间的关系，并求出全部解.

本章主要讨论线性方程组的求解问题. 首先应用行列式讨论

未知量个数与方程个数相同的一类特殊的线性方程组，其次介绍消元法. 应用消元法不仅简化了线性方程组求解的计算问题，还给出了线性方程组解的情况. 最后，应用 n 维向量的理论全面地解决了线性方程组的求解问题.

2.1 克莱姆法则

这一节讨论方程的个数与未知量的个数相等(即 $s=n$)的情形，更一般的情形放在下一节讨论. 以后我们会看到：这里所讨论的情形虽然是特殊的，但却是很重要的，在讨论一般情形时，也需要用到这里的结果.

这一节主要证明下述定理，它与 2 元、3 元线性方程组的结论是相仿的.

定理 1 (克莱姆(Cramer)法则)如果线性方程组

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

的系数行列式

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

那么这个方程组有解，并且解是唯一的，这个解可表示成：

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \cdots, x_n = \frac{D_n}{D}. \tag{3}$$

其中 D_i 是把 D 中第 i 列换成常数项 $b_1, b_2, \cdots, b_n$ 所得的行列式. 即

$$D_i = \begin{vmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{vmatrix} \quad (i=1,2,\cdots,n).$$

定理 1 一共有 3 个结论：①方程组有解；②解是唯一的；③解由公式(3)给出. 证明的步骤是：

第一，把 $x_i=\frac{D_i}{D}(i=1,2,\cdots,n)$代入方程组，验证它确实是解. 这样就证明了方程组有解，并且(3)是一个解，即证明了结论①与③.

第二，证明：如果 $x_1=c_1,x_2=c_2,\cdots,x_n=c_n$ 是方程组(2)的一个解，那么一定有 $c_1=\frac{D_1}{D},c_2=\frac{D_2}{D},\cdots,c_n=\frac{D_n}{D}$. 这就证明了解的唯一性，即证明了结论②.

在证明过程中，将反复应用行列式按一行(列)展开的公式.

证明：首先，证明(3)确是(2)的解：把 $x_i=\frac{D_i}{D}(i=1,2,\cdots,n)$ 代入第 k 个方程的左端，得：

$$\begin{aligned} & a_{k1}\frac{D_1}{D}+a_{k2}\frac{D_2}{D}+\cdots+a_{kn}\frac{D_n}{D} \\ = & \frac{1}{D}(a_{k1}D_1+a_{k2}D_2+\cdots+a_{kn}D_n). \end{aligned} \tag{4}$$

因为

$$D_i=b_1A_{1i}+b_2A_{2i}+\cdots+b_kA_{ki}+\cdots+b_nA_{ni} \quad (i=1,2,\cdots,n),$$

所以

$$\begin{aligned}(4)\text{ 式}=\frac{1}{D}[&a_{k1}(b_1A_{11}+b_2A_{21}+\cdots+b_kA_{k1}+\cdots+b_nA_{n1}) \\ &+a_{k2}(b_1A_{12}+b_2A_{22}+\cdots+b_kA_{k2}+\cdots+b_nA_{n2})+\cdots \\ &+a_{kn}(b_1A_{1n}+b_2A_{2n}+\cdots+b_kA_{kn}+\cdots+b_nA_{nn})]\end{aligned}$$

$$
\begin{aligned}
&=\frac{1}{D}[b_1(a_{k1}A_{11}+a_{k2}A_{12}+\cdots+a_{kn}A_{1n})\\
&\quad+b_2(a_{k1}A_{21}+a_{k2}A_{22}+\cdots+a_{kn}A_{2n})+\cdots\\
&\quad+b_k(a_{k1}A_{k1}+a_{k2}A_{k2}+\cdots+a_{kn}A_{kn})+\cdots\\
&\quad+b_n(a_{k1}A_{n1}+a_{k2}A_{n2}+\cdots+a_{kn}A_{nn})].
\end{aligned}
$$

根据行列式按一行展开的公式,可以看出:上式的方括号中只有 b_k 的系数是 D,而其他 $b_s(s\neq k)$的系数全为 0. 因此得:

$$
a_{k1}\frac{D_1}{D}+a_{k2}\frac{D_2}{D}+\cdots+a_{kn}\frac{D_n}{D}=\frac{1}{D}b_kD=b_k.
$$

这说明将(3)代入第 $k(k=1,2,\cdots,n)$个方程后,得到了一个恒等式,所以(3)是(2)的一个解.

其次,设 $x_1=c_1,x_2=c_2,\cdots,x_n=c_n$ 是方程组(2)的一个解. 那么,将 $x_i=c_i$ 代入(2)后,得到 n 个恒等式:

$$
\begin{cases}
a_{11}c_1+a_{12}c_2+\cdots+a_{1n}c_n=b_1,\\
a_{21}c_1+a_{22}c_2+\cdots+a_{2n}c_n=b_2,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
a_{n1}c_1+a_{n2}c_2+\cdots+a_{nn}c_n=b_n.
\end{cases}
\tag{5}
$$

用系数行列式的第 i 列的代数余子式 $A_{1i},A_{2i},\cdots,A_{ni}$依次去乘(5)中 n 个恒等式,得

$$
\begin{cases}
a_{11}A_{1i}c_1+a_{12}A_{1i}c_2+\cdots+a_{1n}A_{1i}c_n=b_1A_{1i},\\
a_{21}A_{2i}c_1+a_{22}A_{2i}c_2+\cdots+a_{2n}A_{2i}c_n=b_2A_{2i},\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
a_{n1}A_{ni}c_1+a_{n2}A_{ni}c_2+\cdots+a_{nn}A_{ni}c_n=b_nA_{ni}.
\end{cases}
$$

将此 n 个等式相加,得:

$$
\begin{aligned}
&(a_{11}A_{1i}+a_{21}A_{2i}+\cdots+a_{n1}A_{ni})c_1\\
&+(a_{12}A_{1i}+a_{22}A_{2i}+\cdots+a_{n2}A_{ni})c_2+\cdots\\
&+(a_{1i}A_{1i}+a_{2i}A_{2i}+\cdots+a_{ni}A_{ni})c_i+\cdots\\
&+(a_{1n}A_{1i}+a_{2n}A_{2i}+\cdots+a_{nn}A_{ni})c_n
\end{aligned}
$$

$$=b_1A_{1i}+b_2A_{2i}+\cdots+b_iA_{ii}+\cdots+b_nA_{ni}.$$

利用按一列展开的性质，根据

$$\sum_{s=1}^{n}a_{sl}A_{si}=\begin{cases}D, & \text{当 } l=i \text{ 时；}\\ 0, & \text{当 } l\neq i.\end{cases}$$

于是得：
$$c_iD=D_i,$$

因此
$$c_i=\frac{D_i}{D}.$$

这就是说，如果$(c_1,c_2,\cdots,c_n)$是方程组(2)的一个解，那么一定有$c_i=\dfrac{D_i}{D}$ $(i=1,2,\cdots,n)$，所以方程组只有一个解. |

例 1 解线性方程组

$$\begin{cases}3x_1+x_2-x_3+x_4=-3,\\ x_1-x_2+x_3+2x_4=4,\\ 2x_1+x_2+2x_3-x_4=7,\\ x_1\qquad\quad+2x_3+x_4=6.\end{cases}$$

解 方程组的系数行列式

$$D=\begin{vmatrix}3&1&-1&1\\1&-1&1&2\\2&1&2&-1\\1&0&2&1\end{vmatrix}=-13\neq 0.$$

所以根据克莱姆法则，这个线性方程组有唯一解. 又因

$$D_1=\begin{vmatrix}-3&1&-1&1\\4&-1&1&2\\7&1&2&-1\\6&0&2&1\end{vmatrix}=-13,$$

$$D_2=\begin{vmatrix}3&-3&-1&1\\1&4&1&2\\2&7&2&-1\\1&6&2&1\end{vmatrix}=26,$$

$$D_3=\begin{vmatrix}3&1&-3&1\\1&-1&4&2\\2&1&7&-1\\1&0&6&1\end{vmatrix}=-39,$$

$$D_4=\begin{vmatrix}3&1&-1&-3\\1&-1&1&4\\2&1&2&7\\1&0&2&6\end{vmatrix}=13.$$

所以这个线性方程组的唯一解为

$$x_1=\frac{D_1}{D}=1,\quad x_2=\frac{D_2}{D}=-2,$$

$$x_3=\frac{D_3}{D}=3,\quad x_4=\frac{D_4}{D}=-1.$$

需要注意的是，定理 1 所讨论的只是系数行列式不等于零的方程组；只有当系数行列式不等于零时，才能用克莱姆法则求解.

在线性方程组中，有一种特殊的线性方程组，即常数项全为零的方程组，称为**齐次线性方程组**. 显然，齐次线性方程组总是有解的，因为 $x_i=0(i=1,2,\cdots,n)$ 就是它的解，这个解称为**零解**；其他的，即 x_i 不全为零的解，称为**非零解**. 所以，对于齐次线性方程组，需要讨论的问题，不是有没有解，而是有没有非零解. 这个问题与齐次方程组解的个数是有密切关系的. 如果一个齐次线性方程组只有零解，那么这个方程组就有唯一解；反之，如果某个齐次方程组有唯一解，那么由于零解是一个解，所以这个方程组不可能有非零解.

对于方程个数与未知量个数相同的齐次线性方程组，应用克莱姆法则，有

定理 2 如果齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases} \tag{6}$$

的系数行列式 D 不等于零，那么(6)只有零解. 也就是说，如果方程组(6)有非零解，则必须 $D=0$.

证明：由于 $D\neq 0$，故可应用克莱姆法则. 因为行列式 $D_i(i=1,2,\cdots,n)$ 的第 i 列的元素全等于零，所以 $D_i=0(i=1,2,\cdots,n)$. 于是，方程组的唯一解是：

$$x_i = \frac{D_i}{D} = 0 \quad (i = 1,2,\cdots,n). \mid$$

关于齐次线性方程组的非零解问题将在 2.7 节中详细论述.

习　题　2.1

1. 用克莱姆法则解下列线性方程组：

(1) $\begin{cases} 3x_1+2x_2+4x_3-x_4=13, \\ 5x_1+x_2-x_3+2x_4=9, \\ 2x_1+3x_2-7x_3+3x_4=14, \\ 4x_1-4x_2+3x_3-5x_4=4; \end{cases}$

(2) $\begin{cases} 2x_1-x_2+3x_3-2x_4=-6, \\ x_1+7x_2+x_3-x_4=5, \\ 3x_1+5x_2-5x_3+3x_4=19, \\ x_1-x_2-2x_3+x_4=4; \end{cases}$

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

(4) $\begin{cases} 5x_1+6x_2=1 \\ x_1+5x_2+6x_3=-2, \\ x_2+5x_3+6x_4=2, \\ x_3+5x_4+6x_5=-2, \\ x_4+5x_5=-4. \end{cases}$

2. 证明下列齐次线性方程组只有零解：

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

2.2 消 元 法

对于解决未知量个数与方程个数相等而且系数行列式不等于零的线性方程组,克莱姆法则在理论上是一个非常完善的结果. 它不仅肯定了这种方程组有解,而且说明只有一个解,还用公式把解通过系数及常数项表示了出来. 这个结果不但完全解决了这一类特殊的线性方程组,而且还将在以后解决一般线性方程组的问题时起着重要的作用. 但是在具体应用克莱姆法则求解时,需要计算 $n+1$ 个 n 阶行列式,计算量是比较大的. 所以在具体解线性方程组时,一般不用克莱姆法则,而用消元法来计算.

在中学代数里,我们学过用消元法解 2 元、3 元线性方程组. 消元法的基本思想是:把方程组中一部分方程变成未知量较少的方程. 请看下面一个具体的例子.

例 1 解线性方程组

$$\begin{cases} 2x_1 + 2x_2 - 3x_3 = 9, \\ x_1 + 2x_2 + x_3 = 4, \\ 3x_1 + 9x_2 + 2x_3 = 19. \end{cases}$$

解: 将第 1、2 两个方程互换,方程组变为:

$$\begin{cases} x_1 + 2x_2 + x_3 = 4, \\ 2x_1 + 2x_2 - 3x_3 = 9, \\ 3x_1 + 9x_2 + 2x_3 = 19. \end{cases}$$

由第 2 个方程减去第 1 个方程的 2 倍;第 3 个方程减去第 1 个方程的 3 倍,得:

$$\begin{cases} x_1 + 2x_2 + x_3 = 4, \\ -2x_2 - 5x_3 = 1, \\ 3x_2 - x_3 = 7. \end{cases}$$

在新方程组中,把第 2 个方程的 3 倍与第三个方程的 2 倍相加,

得：

$$\begin{cases} x_1 + 2x_2 + x_3 = 4, \\ \quad - 2x_2 - 5x_3 = 1, \\ \qquad\quad - 17x_3 = 17. \end{cases}$$

至此，很容易求出方程组的解是(1，2，－1).

分析消元法的内容，不难看出：它是反复地将方程组进行变换，而所作的变换则是由下述 3 种基本变换构成的：

1. 用一个非零的数乘一个方程；

2. 用一个数乘一个方程后加到另一个方程上；

3. 互换两个方程的位置.

于是，引进下述定义.

定义 1 变换 1，2，3 称为线性方程组的初等变换.

对方程组进行变换是为了把方程组化简，但变换前后方程组的解集合必须要保持不变. 两个方程组如果有相同的解集合，就称这两个方程组是**同解的**. 下面证明：初等变换把方程组变成与它同解的方程组. 上述初等变换 1 和 3 不会改变方程组的解集合是明显的. 下面仅对初等变换 2 来证明：

对方程组

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

进行第二种初等变换：为简便起见，不妨设把第一个方程的 k 倍加到第二个方程上，得到新方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ (a_{21} + ka_{11})x_1 + (a_{22} + ka_{12})x_2 + \cdots + (a_{2n} + ka_{1n})x_n = b_2 + kb_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = b_s. \end{cases} \tag{2}$$

设$(c_1,c_2,\cdots,c_n)$是(1)的任一解. 把它代入(1)后,得到 s 个恒等式:

$$\begin{cases} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1, \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}c_1 + a_{s2}c_2 + \cdots + a_{sn}c_n = b_s. \end{cases}$$

把第一个等式的 k 倍加到第二个等式上,得到

$$\begin{cases} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1, \\ (a_{21} + ka_{11})c_1 + (a_{22} + ka_{12})c_2 + \cdots + (a_{2n} + ka_{1n})c_n = b_2 + kb_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}c_1 + a_{s2}c_2 + \cdots - a_{sn}c_n = b_s. \end{cases}$$

这说明$(c_1,c_2,\cdots,c_n)$也是方程组(2)的一个解. 用同样的方法可以证明,凡是方程组(2)的解也都是方程组(1)的解. 因此,(1)与(2)是同解的.

下面说明,如何利用初等变换来解线性方程组. 首先还是讨论 $s=n$ 并且系数行列式不等于零的方程组. 为此我们来证明下述定理:

定理 3 如果方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n, \end{cases} \tag{3}$$

的系数行列式

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

那么(3)可以用初等变换化成下列形式的同解方程组：

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + c_{13}x_3 + \cdots + c_{1n}x_n = d_1, \\ \qquad c_{22}x_2 + c_{23}x_3 + \cdots + c_{2n}x_n = d_2, \\ \qquad\qquad c_{33}x_3 + \cdots + c_{3n}x_n = d_3, \\ \qquad\qquad\cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad c_{nn}x_n = d_n, \end{cases} \tag{4}$$

其中系数 $c_{11},c_{22},\cdots,c_{nn}$全不为零.

证明：对 n 用数学归纳法：当 $n=1$ 时，结论显然成立. 假设定理对 $n-1$ 个未知量的方程组也是成立的. 下面考察 n 个未知量的方程组(3).

因为 $D\neq 0$，所以 $a_{11},a_{21},\cdots,a_{n1}$不会全等于零，因而利用上述初等变换 3，可以将方程组(3)变成同解方程组(目的是使新方程组中的 $a'_{11}\neq 0$)

$$\begin{cases} a'_{11}x_1 + a'_{12}x_2 + \cdots + a'_{1n}x_n = b'_1, \\ a'_{21}x_1 + a'_{22}x_2 + \cdots + a'_{2n}x_n = b'_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a'_{n1}x_1 + a'_{n2}x_2 + \cdots + a'_{nn}x_n = b'_n, \end{cases} \tag{5}$$

其中 $a'_{11}\neq 0$.

将(5)中第 1 个方程的$-\frac{a'_{21}}{a'_{11}}$倍加到第 2 个方程；第 1 个方程的$-\frac{a'_{31}}{a'_{11}}$倍加到第 3 个方程……第 1 个方程的$-\frac{a'_{n1}}{a'_{11}}$倍加到第 n 个方程. 于是方程组(5)就变成了下列方程组(6)

$$\begin{cases} a'_{11}x_1 + a'_{12}x_2 + \cdots + a'_{1n}x_n = b'_1, \\ \qquad a''_{22}x_2 + \cdots + a''_{2n}x_n = b''_2, \\ \qquad\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad a''_{n2}x_2 + \cdots + a''_{nn}x_n = b''_n, \end{cases} \tag{6}$$

其中 $a'_{11}\neq 0$.

方程组(6)的系数行列式是由(3)的系数行列式经过两行对换以及用一行的倍数加到另一行后而得的，因此

$$\begin{vmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ 0 & a''_{22} & \cdots & a''_{2n} \\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\ 0 & a''_{n2} & \cdots & a''_{nn} \end{vmatrix} = D' \neq 0.$$

因为

$$\begin{vmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ 0 & a''_{22} & \cdots & a''_{2n} \\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\ 0 & a''_{n2} & \cdots & a''_{nn} \end{vmatrix} = a'_{11} \begin{vmatrix} a''_{22} & \cdots & a''_{2n} \\ \multicolumn{3}{c}{\cdots\cdots\cdots\cdots} \\ a''_{n2} & \cdots & a''_{nn} \end{vmatrix},$$

所以

$$\begin{vmatrix} a''_{22} & \cdots & a''_{2n} \\ \multicolumn{3}{c}{\cdots\cdots\cdots\cdots} \\ a''_{n2} & \cdots & a''_{nn} \end{vmatrix} \neq 0.$$

考察 $n-1$ 个未知量的线性方程组

$$\begin{cases} a''_{22}x_2 + \cdots + a''_{2n}x_n = b''_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a''_{n2}x_2 + \cdots + a''_{nn}x_n = b''_n. \end{cases}$$

由于这个方程组的系数行列式不等于零，根据归纳法假设，可以用初等变换将它变为下列形式的线性方程组：

$$\begin{cases} c_{22}x_2 + c_{23}x_3 + \cdots + c_{2n}x_n = d_2, \\ \qquad\qquad c_{33}x_3 + \cdots + c_{3n}x_n = d_3, \\ \qquad\qquad\qquad \cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\qquad\qquad c_{nn}x_n = d_n. \end{cases}$$

其中 $c_{22}, c_{33}, \cdots, c_{nn}$ 全不为零. 将这些初等变换用于方程组(6)，就得到形式如(4)的同解方程. |

显然，方程组(4)是很容易求解的：从最后一个方程，可以直接得到 $x_n = \dfrac{d_n}{c_{nn}}$，代入倒数第 2 个方程，就可算出 x_{n-1}，逐次代上去，

就可依次求出 $x_{n-2},\cdots,x_3,x_1$.

例 2 解线性方程组

$$\begin{cases} x_1+2x_2-3x_3+4x_4-x_5=-1, \\ 2x_1-x_2+3x_3-4x_4+2x_5=8, \\ 3x_1+x_2-x_3+2x_4-x_5=3, \\ 4x_1+3x_2+4x_3+2x_4+2x_5=-2, \\ x_1-3x_2-x_3+2x_4-3x_5=-3; \end{cases}$$

解：用消元法求解：

$$\begin{cases} x_1+2x_2-3x_3+4x_4-x_5=-1, \\ 2x_1-x_2+3x_3-4x_4+2x_5=8, \\ 3x_1+x_2-x_3+2x_4-x_5=3, \\ 4x_1+3x_2+4x_3+2x_4+2x_5=-2, \\ x_1-3x_2-x_3+2x_4-3x_5=-3; \end{cases}$$

$$\to\begin{cases} x_1+2x_2-3x_3+4x_4-x_5=-1; \\ -5x_2+9x_3-12x_4+4x_5=10, \\ -5x_2+8x_3-10x_4+2x_5=6, \\ -5x_2+16x_3-14x_4+6x_5=2, \\ -5x_2+2x_3-2x_4-2x_5=-2; \end{cases}$$

$$\to\begin{cases} x_1+2x_2-3x_3+4x_4-x_5=-1, \\ -5x_2+9x_3-12x_4+4x_5=10, \\ -x_3+2x_4-2x_5=-4, \\ 7x_3-2x_4+2x_5=-8, \\ -7x_3+10x_4-6x_5=-12; \end{cases}$$

$$\to\begin{cases} x_1+2x_2-3x_3+4x_4-x_5=-1, \\ -5x_2+9x_3-12x_4+4x_5=10, \\ -x_3+2x_4-2x_5=-4, \\ 12x_4-12x_5=-36, \\ -4x_4+8x_5=16; \end{cases}$$

$$
\rightarrow\begin{cases}
x_1+2x_2-3x_3+4x_4-x_5=-1,\\
\qquad 5x_2+9x_3-12x_4+4x_5=10,\\
\qquad\qquad -x_3+2x_4-2x_5=-4,\\
\qquad\qquad\qquad x_4-x_5=-3,\\
\qquad\qquad\qquad\qquad x_5=1;
\end{cases}
$$

$$
\rightarrow\begin{cases}
x_1=2,\\
x_2=0,\\
x_3=-2,\\
x_4=-2,\\
x_5=1.
\end{cases}
$$

所以这个方程组的解是(2,0,−2,−2,1).

其实,在作初等变换化简方程的时候,只是对这些方程的系数进行变换,所以为了简便起见,可以将未知量略去不写,而将系数列成一个表来计算.这样做不但简单,而且不容易出错.下面举例说明.

例 3 解线性方程组

$$
\begin{cases}
2x_1+x_2+x_3+x_4+x_5=4,\\
x_1+2x_2+x_3+x_4+x_5=5,\\
x_1+x_2+3x_3+x_4+x_5=9,\\
x_1+x_2+x_3+4x_4+x_5=0,\\
x_1+x_2+x_3+x_4+5x_5=-5.
\end{cases}
$$

解:用消元法进行计算

$$
\left(\begin{array}{ccccc|c}
2 & 1 & 1 & 1 & 1 & 4\\
1 & 2 & 1 & 1 & 1 & 5\\
1 & 1 & 3 & 1 & 1 & 9\\
1 & 1 & 1 & 4 & 1 & 0\\
1 & 1 & 1 & 1 & 5 & -5
\end{array}\right)
\rightarrow
\left(\begin{array}{ccccc|c}
1 & 2 & 1 & 1 & 1 & 5\\
2 & 1 & 1 & 1 & 1 & 4\\
1 & 1 & 3 & 1 & 1 & 9\\
1 & 1 & 1 & 4 & 1 & 0\\
1 & 1 & 1 & 1 & 5 & -5
\end{array}\right)
$$

$$
\rightarrow\begin{pmatrix}1&2&1&1&1&5\\0&-3&-1&-1&-1&-6\\0&-1&2&0&0&4\\0&-1&0&3&0&-5\\0&-1&0&0&4&-10\end{pmatrix}
$$

$$
\rightarrow\begin{pmatrix}1&2&1&1&1&5\\0&1&-2&0&0&-4\\0&-3&-1&-1&-1&-6\\0&-1&0&3&0&-5\\0&-1&0&0&4&-10\end{pmatrix}
$$

$$
\rightarrow\begin{pmatrix}1&2&1&1&1&5\\0&1&-2&0&0&-4\\0&0&-7&-1&-1&-18\\0&0&-2&3&0&-9\\0&0&-2&0&4&-14\end{pmatrix}\rightarrow\begin{pmatrix}1&2&1&1&1&5\\0&1&-2&0&0&-4\\0&0&1&0&-2&7\\0&0&7&1&1&18\\0&0&-2&3&0&-9\end{pmatrix}
$$

$$
\rightarrow\begin{pmatrix}1&2&1&1&1&5\\0&1&-2&0&0&-4\\0&0&1&0&-2&7\\0&0&0&1&15&-31\\0&0&0&3&-4&5\end{pmatrix}\rightarrow\begin{pmatrix}1&2&1&1&1&5\\0&1&-2&0&0&-4\\0&0&1&0&-2&7\\0&0&0&1&15&-31\\0&0&0&0&-49&98\end{pmatrix}
$$

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

所以原方程组的解为$(1,2,3,-1,-2)$.

不知道读者有没有注意到这样一个问题:当我们用消元法解例 2 及例 3 的时候,并没有先计算这两个方程组的系数行列式.这

样做是不是会出现什么问题呢?不会的. 这是因为我们已经证明了经过初等变换后,方程组的解集合是保持不变的,所以最后一组方程与原来的方程组是同解的,因此只要解最后的方程组就可以了. 此外,还可以从另一个角度来看这个问题,原来的方程组经过一系列的初等变换变成新方程组时,这两个方程组的系数行列式之间有什么关系?从初等变换的定义可以看出:这两个系数行列式只相差一个非零常数倍,或者都等于零,或都不等于零. 如果原来的系数行列式等于零,那么在消元过程中所得到的每个方程组的系数行列式也一定都等于零,这一点一定会在消元过程中发现. 至于如果发现了某个线性方程组的系数行列式等于零,如何进行计算,这将在下面讨论.

习　题　2.2(1)

1. 解下列线性方程组:

(1) $\begin{cases} 2x_1-x_2+3x_3+2x_4=6, \\ 3x_1-3x_2+3x_3+2x_4=5, \\ 3x_1-x_2-x_3+2x_4=3, \\ 3x_1-x_2+3x_3-x_4=4; \end{cases}$　(2) $\begin{cases} x_1+3x_2+5x_3+7x_4=12, \\ 3x_1+5x_2+7x_3+x_4=0, \\ 5x_1+7x_2+x_3+3x_4=4, \\ 7x_1+x_2+3x_3+5x_4=16; \end{cases}$

(3) $\begin{cases} x_1+2x_2+3x_3+4x_4+5x_5=13, \\ 2x_1+x_2+2x_3+3x_4+4x_5=10, \\ 2x_1+2x_2+x_3+2x_4+3x_5=11, \\ 2x_1+2x_2+2x_3+x_4+2x_5=6, \\ 2x_1+2x_2+2x_3+2x_4+x_5=3; \end{cases}$

(4) $\begin{cases} x_1+4x_2+6x_3+4x_4+x_5=4, \\ x_1+x_2+4x_3+6x_4+4x_5=3, \\ 4x_1-x_2+x_3+4x_4+6x_5=3, \\ 6x_1+4x_2+x_3+x_4+4x_5=3, \\ 4x_1+6x_2+4x_3+x_4+x_5=3. \end{cases}$

以下通过消元法来讨论一般线性方程组的求解问题. 首先我

们通过一些例子来说明线性方程组的解可能出现的几种情况.

例 4 解线性方程组

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

解: 首先用初等变换化简方程组:

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

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

于是得到与原方程组同解的方程组:

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

显然,这个方程组有唯一解:

$$x_1 = -1,\ x_2 = -2,\ x_3 = 2.$$

这也是原方程组的解. 即原方程组有唯一解:(−1,−2,2).

例 5 解线性方程组

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

解: 用初等变换将方程组化简:

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

$$\to\begin{pmatrix}1&-2&3&-4&4\\0&1&-1&1&-3\\0&0&2&-4&12\\0&0&-4&8&-24\end{pmatrix}\to\begin{pmatrix}1&-2&3&-4&4\\0&1&-1&1&-3\\0&0&1&-2&6\\0&0&0&0&0\end{pmatrix}.$$

于是得到与原方程组同解的方程组：

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

把这个方程改写为：

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

可以把 x_1,x_2,x_3 用 x_4 唯一地表示出来：

$$\begin{cases}x_1=-8,\\x_2=3+x_4,\\x_3=6+2x_4.\end{cases}$$

由此可以看到：x_4 可以取任意值. 而当 x_4 取定了以后，x_1,x_2,x_3 就可以从上面这一组等式求出. 所以，原方程组的解可以表示成：

$$\begin{cases}x_1=-8,\\x_2=3+k,\\x_3=6+2k,\\x_4=k.\end{cases}$$

其中 k 可以是任意一个数，这个方程组有无穷多个解.

例 6 解线性方程组

$$\begin{cases} x_1 - 2x_2 + 3x_3 - x_4 - x_5 = 2, \\ x_1 + x_2 - x_3 + x_4 - 2x_5 = 1, \\ 2x_1 - x_2 + x_3 \qquad - 2x_5 = 2, \\ 2x_1 + 2x_2 - 5x_3 + 2x_4 - x_5 = 5. \end{cases}$$

解：首先进行初等变换

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

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

于是得到同解方程组

$$\begin{cases} x_1 - 2x_2 + 3x_3 - x_4 - x_5 = 2 \\ \qquad 3x_2 - 4x_3 + 2x_4 - x_5 = -1 \\ \qquad\qquad x_3 \qquad - x_5 = 1 \\ \qquad\qquad\qquad\qquad 0 = 6. \end{cases}$$

从最后一个方程可以看出：这个方程组是无解的．所以原方程组也无解．

从以上 3 个例子看出：一般线性方程组的解可能有 3 种情况：有唯一解，有无穷多解或无解．所以讨论线性方程组，主要是解决下面这 3 个问题：

1．给了一个线性方程组，如何判断它有没有解？也就是说，线性方程组有解的充分必要条件是什么？

2．如果已知一个线性方程组有解，那么它究竟有多少解？怎样求解？

3. 如果只有一个解，那么把这个解求出来就行了；如果不止一个解，那么这些解之间有什么关系？

这一章的主要内容就是解决这些问题.

给了一个线性方程组，一般地很难直接看出它有没有解以及有多少个解. 但从上述例题看出：用初等变换可以把线性方程组化简，直到可以看出这个方程组有没有解，并且可以通过同解的较为简单的方程组把原方程组的解求出来. 那么，任意给了一个线性方程组，通过初等变换能够化简到什么样子？我们首先来解决这个问题.

为了便于叙述及以后应用，下面引进矩阵以及矩阵的初等行变换等概念.

定义 2 由 sn 个数排成的 s 行 n 列的表

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} \end{bmatrix}, \tag{7}$$

称为一个 $s\times n$ **矩阵**.

例如

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

是一个 2×4 矩阵，

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

是一个 3×3 矩阵.

(7)中的数 $a_{ij}(i=1,2,\cdots,s;j=1,2,\cdots,n)$ 称为矩阵(7)的**元素**. i 称为 a_{ij} 的**行标**，说明 a_{ij} 位于矩阵(7)的第 i 行；j 称为 a_{ij} 的**列标**，说明 a_{ij} 位于矩阵(7)的第 j 列.

通常用大写拉丁字母 $A,B,\cdots$ 或者 $[a_{ij}],[b_{ij}],\cdots$ 来表示矩阵.

有时候,为了表明 $[a_{ij}]$ 是一个 $s\times n$ 矩阵,把它表成 $[a_{ij}]_{sn}$. 如果矩阵 A 的行数与列数相等,都等于 n,就称 A 是一个 n 阶矩阵或 n 阶**方阵**.

两个矩阵只有在它们的行、列数分别相等,并且对应的元素都相等时,才叫做**相等**. 设 $A=[a_{ij}]_{sn}$ 是一个 $s\times n$ 矩阵,$B=[b_{ij}]_{tm}$ 是一个 $t\times m$ 矩阵,那么 $A=B$ 就是说 $s=t,n=m$,并且 $a_{ij}=b_{ij}$ 对 $i=1,2,\cdots,s;j=1,2,\cdots,n$ 都成立. 即相等的矩阵是完全一样的.

给了一个线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{s1}x_1+a_{s2}x_2+\cdots+a_{sn}x_n=b_s.\end{cases}\tag{8}$$

把系数按原来的位置写成一个 $s\times n$ 矩阵

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

称为(8)的**系数矩阵**. 若把常数项也添成一列,则得到一个 $s\times(n+1)$ 矩阵

$$\overline{A}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}&b_1\\ a_{21}&a_{22}&\cdots&a_{2n}&b_2\\ \cdots&\cdots&\cdots&\cdots&\cdots\\ a_{s1}&a_{s2}&\cdots&a_{sn}&b_s\end{bmatrix}$$

称为(8)的**增广矩阵**.

显然,如果知道了一个线性方程组的全部系数和常数项,那么这个线性方程组就完全确定了. 也就是说,线性方程组(8)可以用它的增广矩阵 $\overline{A}$ 来表示. 至于用什么文字来表示未知量,当然不

是实质性的.

前面曾介绍了线性方程组的 3 种初等变换. 容易看出：对线性方程组(8)施行初等变换，就相当于对增广矩阵的行施行相应的变换. 因此，下面定义矩阵的初等行变换.

定义 3 矩阵的初等行变换是指下列 3 种变换：

1. 用一个非零的数乘矩阵的一行；
2. 把矩阵的某一行乘 k 后加到另一行上(k 是任意数)；
3. 互换矩阵中两行的位置.

当矩阵 A 经过初等变换，变成矩阵 B 时，我们写成

$$A \to B.$$

设 A 是一个矩阵，对 A 的任一非零行，其中第一个非零元称为此行的**首元**. 若 A 的前 r 行为非零，其余行全为零，并且第 1 至第 r 行的首元所在的列为 $j_1, j_2, \cdots, j_r$，满足

$$j_1 < j_2 < \cdots < j_r$$

则称 A 是一个**阶梯形矩阵**，简称**梯阵**. 例如

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

都是阶梯形矩阵. 如果在阶梯形矩阵 A 中，每个首元都等于 1，并且每个首元所在的列的其它元素都等于零，则称 A 是一个**约化阶梯形矩阵**，简称**约化梯阵**. 例如

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

都是约化梯阵.

可以证明，任意一个矩阵总可以经过一系列的初等行变换变成阶梯形矩阵.

事实上，设

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\ a_{s1} & a_{s2} & \cdots & a_{sn} \end{bmatrix}.$$

先看第 1 列元素 $a_{11},a_{21},\cdots,a_{s1}$. 如果这 s 个元素不全为零,就可以经过第 3 种初等行变换,而使第 1 列的第 1 个元素不等于零. 然后依次用一个适当的数乘第 1 行后加到第 2,…,s 行上,使得第 1 列除去第 1 个元素以外都等于零,也就是说,经过一系列初等行变换后,

$$A\to B=\begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\ 0 & a'_{s2} & \cdots & a'_{sn} \end{bmatrix}.$$

再对 B 的右下角的一块

$$\begin{bmatrix} a'_{22} & \cdots & a'_{2n} \\ \multicolumn{3}{c}{\cdots\cdots\cdots\cdots} \\ a'_{s2} & \cdots & a'_{sn} \end{bmatrix}$$

重复以上的变换(注意,在对这一块施行初等行变换时,B 的第 1 行和第 1 列不起变化). 这样下去,直到变成阶梯形为止. 如果矩阵 A 的第 1 列元素全为零,那么就考虑它的第 2 列元素,等等.

例如,设

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

对 A 施行初等行变换:

$$
A \to \begin{bmatrix} 0 & 1 & 3 & -2 & 2 & -1 & 1 \\ 0 & 0 & 0 & 2 & 2 & -1 & 1 \\ 0 & 2 & 6 & -4 & 5 & 7 & 1 \\ 0 & -1 & -3 & 4 & 0 & 5 & 1 \end{bmatrix}
$$

$$
\to \begin{bmatrix} 0 & 1 & 3 & -2 & 2 & -1 & 1 \\ 0 & 0 & 0 & 2 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 9 & -1 \\ 0 & 0 & 0 & 2 & 2 & 4 & 2 \end{bmatrix}
$$

$$
\to \begin{bmatrix} 0 & 1 & 3 & -2 & 2 & -1 & 1 \\ 0 & 0 & 0 & 2 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 9 & -1 \\ 0 & 0 & 0 & 0 & 0 & 5 & 1 \end{bmatrix}.
$$

这样就把 A 化成了一个阶梯形矩阵.

还可证明任一矩阵都可经初等行变换化为约化梯阵，而且矩阵的约化梯阵是唯一的.我们不予证明，只以上面例中的 A 来说明如何从 A 的梯阵再进一步化为约化梯阵：

$$
A \to \begin{bmatrix} 0 & 1 & 3 & -2 & 2 & -1 & 1 \\ 0 & 0 & 0 & 2 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 9 & -1 \\ 0 & 0 & 0 & 0 & 0 & 5 & 1 \end{bmatrix}
$$

$$
A \to \begin{bmatrix} 0 & 1 & 3 & -2 & 2 & -1 & 1 \\ 0 & 0 & 0 & 1 & 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 9 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{5} \end{bmatrix}
$$

$$\rightarrow\begin{bmatrix}0 & 1 & 3 & 0 & 0 & 0 & \frac{28}{5} \\ 0 & 0 & 0 & 1 & 0 & 0 & \frac{17}{5} \\ 0 & 0 & 0 & 0 & 1 & 0 & -\frac{14}{5} \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{5}\end{bmatrix}.$$

这就是 A 的约化梯阵.

现在回过来讨论线性方程组. 我们知道:线性方程组可以经过初等变换化为同解的方程组,而对线性方程组作初等变换就相当于对它的增广矩阵作相应的初等行变换. 由于每个矩阵都可以通过初等行变换化为阶梯形矩阵,所以每个线性方程组都可以用初等变换化成同解的**阶梯形方程组**.

设线性方程组(8)经过一系列初等变换化成了阶梯形方程组. 为了方便起见,不妨设所得的方程组为:

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1r}x_r + \cdots + c_{1n}x_n = d_1, \\ \qquad c_{22}x_2 + \cdots + c_{2r}x_r + \cdots + c_{2n}x_n = d_2, \\ \qquad\qquad \cdots\cdots\cdots\cdots \\ \qquad\qquad c_{rr}x_r + \cdots + c_{rn}x_n = d_r, \\ \qquad\qquad\qquad 0 = d_{r+1}, \\ \qquad\qquad\qquad 0 = 0, \\ \qquad\qquad\qquad \cdots\cdots \\ \qquad\qquad\qquad 0 = 0, \end{cases} \tag{9}$$

其中 $c_{ii}\neq 0(i=1,2,\cdots,r)$. 方程(9)中最后一些方程"$0=0$"是一些恒等式,可以去掉,并不影响方程组的解.

我们知道:(8)和(9)是同解的. 下面就来讨论方程组(8)的解的存在与个数的问题.

方程组(9)是否有解,取决于最后一个方程

$$0 = d_{r+1}$$

是否有解. 由于(8)与(9)同解，所以方程(9)有解的充分必要条件 $d_{r+1}=0$，也就是方程组(8)有解的充分必要条件.

在有解的情况下，我们来求解，分两种情形讨论：

1. 当 $r=n$ 时：阶梯形方程组为：

$$\begin{cases} c_{11}x_1+c_{12}x_2+\cdots+c_{1r}x_r=d_1, \\ \qquad\quad c_{22}x_2+\cdots+c_{2r}x_r=d_2, \\ \qquad\qquad\cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\qquad c_{nn}x_n=d_n, \end{cases}$$

其中 $c_{ii}\neq 0(i=1,2,\cdots,n)$. 这个方程组的系数行列式为

$$\begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1r} \\ 0 & c_{22} & \cdots & c_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & c_{nn} \end{vmatrix}=c_{11}c_{22}\cdots c_{nn}\neq 0.$$

所以这个方程组有唯一解. 解的方法是：从最后一个方程解出 $x_n=\dfrac{d_n}{c_{nn}}$，代入倒数第 2 个方程，解出 x_{n-1}，这样逐次代入，就得到了方程组的解.

前面所举的例 4 就是这种情形.

2. 当 $r<n$ 时：阶梯形方程组为：

$$\begin{cases} c_{11}x_1+c_{12}x_2+\cdots+c_{1r}x_r+\cdots+c_{1n}x_n=d_1, \\ \qquad\quad c_{22}x_2+\cdots+c_{2r}x_r+\cdots+c_{2n}x_n=d_2, \\ \qquad\qquad\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\qquad c_{rr}x_r+\cdots+c_{rn}x_n=d_r, \end{cases}$$

其中 $c_{ii}\neq 0\ (i=1,2,\cdots,r)$. 把它改写成

$$\begin{cases} c_{11}x_1+c_{12}x_2+\cdots+c_{1r}x_r=d_1-c_{1,r+1}x_{r+1}-\cdots-c_{1n}x_n, \\ \qquad\quad c_{22}x_2+\cdots+c_{2r}x_r=d_2-c_{2,r+1}x_{r-1}-\cdots-c_{2n}x_n, \\ \qquad\qquad\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\qquad c_{rr}x_r=d_r-c_{r,r+1}x_{r+1}-\cdots-c_{rn}x_n. \end{cases} \tag{10}$$

这时，任给 $x_{r+1},\cdots,x_n$ 的一组值，就能唯一地定出 $x_1,x_2,\cdots,x_r$ 的值，从而定出方程组(10)的解. 一般地，由(10)可以通过 $x_{r+1},\cdots,x_n$ 将 $x_1,x_2,\cdots,x_r$ 表示出来，这样一组表达式称为方程组(8)的**一般解**，而 $x_{r+1},\cdots,x_n$ 称为一组**自由未知量**. 这时，方程组(8)有无穷多解. 上面的例 5 就是这种情形，x_4 是自由未知量.

一般线性方程组化成阶梯形时，不一定就是(10)的样子，但是只要把方程中的某些项调动一下，也总可以化成(10)的形式. 看下面的例子.

例 7 解线性方程组

$$\begin{cases} x_1+2x_2-x_3+2x_4=1, \\ 2x_1+4x_2+x_3+x_4=5, \\ -x_1-2x_2-2x_3+x_4=-4. \end{cases}$$

解：将增广矩阵作初等行变换化为阶梯形：

$$\overline{A}=\begin{bmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & 1 & 5 \\ -1 & -2 & -2 & 1 & -4 \end{bmatrix}$$

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

于是得到与原方程组同解的阶梯形方程组

$$\begin{cases} x_1+2x_2-x_3+2x_4=1, \\ \qquad\qquad x_3-x_4=1. \end{cases}$$

我们把它改写成：

$$\begin{cases} x_1-x_3=1-2x_2-2x_4, \\ \qquad x_3=1+x_4. \end{cases}$$

可以得出一般解：

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

其中 x_2, x_4 是一组自由未知量.

自由未知量的取法不是唯一的，在上例中，也可以把阶梯形方程改写成

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

这样，把 x_1, x_4 取作自由未知量，得到一般解为

$$\begin{cases} x_2 \qquad = 1 - \dfrac{1}{2}x_1 - \dfrac{1}{2}x_4, \\ \ \ x_3 \quad = 1 \qquad\qquad + x_4. \end{cases}$$

由同解的阶梯形方程组求一般解的步骤可以通过将增广矩阵化为约化梯阵来进行. 例如在上例中，将增广矩阵进一步化为约化梯阵：

$$A \to \begin{bmatrix} 1 & 2 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

即得同解方程组：

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

改写成

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

这就是一般解，其中 x_2, x_4 是自由未知量.

可以看出，如果线性方程组(3)的系数行列式不等于 0，那么，这个线性方程组的增广矩阵可用初等行变换化为约化梯阵：

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & c_n \end{bmatrix},$$

其中$(c_1,c_2,\cdots,c_n)$就是这个线性方程组的唯一解.

最后提一下：$r>n$ 的情形是不可能出现的.

把以上所得的结果归纳一下：

用初等变换把线性方程组化为阶梯形方程组(这一步可以通过用初等行变换将增广矩阵化为阶梯形矩阵来实现)，把最后一些恒等式“0=0”(假若出现的话)去掉，如果剩下的方程中最后的一个是零等于非零的数，则方程组无解；如果没有这种情况，则有解. 在有解的情况下，如果阶梯形方程组中方程的个数 r 等于未知量的个数 n，那么方程组有唯一解；如果方程的个数 r 小于未知量的个数 n，那么方程组有无穷解. 把这个结果应用到齐次线性方程组，则有

定理 4 如果齐次线性方程组

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

中方程的个数少于未知量的个数，即 $s<n$，那么它有非零解.

证明：把这个方程组化为阶梯形方程组后，方程的个数 r 不会超过原方程组中方程的个数 s，即

$$r\leqslant s<n.$$

因此，它的解不是唯一的，所以它有非零解. |

在上一节中，已经应用克莱姆法则讨论了方程的个数与未知量的个数相等的齐次线性方程组

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

如果(11)的系数行列式

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

那么方程组(11)只有零解. 现在进一步来证明下述定理.

定理 5 齐次线性方程组(11)有非零解的充分必要条件是它的系数行列式 $D=0$.

证明：定理 2 已证明了条件的必要性，现在来证明充分性：将齐次线性方程组(11)用初等变换化为同解的阶梯形方程组. 如果这个阶梯形方程组包含 n 个方程：

$$\begin{cases} a'_{11}x_1 + a'_{12}x_2 + \cdots + a'_{1n}x_n = 0, \\ \qquad\qquad a'_{22}x_2 + \cdots + a'_{2n}x_n = 0, \\ \qquad\qquad\qquad \cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\qquad a'_{nn}x_n = 0, \end{cases} \tag{12}$$

其中 $a'_{ii}\neq 0(i=1,2,\cdots,n)$. 那么(12)的系数行列式

$$D' = \begin{vmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a'_{nn} \end{vmatrix} = a'_{11}a'_{22}\cdots a'_{nn} \neq 0.$$

但是 D' 是由 D 经过一系列初等行变换得到的. 所以 D' 与 D 只能相差一个非零的常数倍. 现在 $D=0$ 而 $D'\neq 0$，这是不可能的. 这说明方程组(11)经初等变换化成阶梯形方程组后，方程的个数(去掉恒等式 $0=0$ 后)必定小于未知量的个数，即 $r<n$. 根据定理 4 知，此方程组必有非零解，因此原方程组也一定有非零解. |

例 8 求 λ，使齐次线性方程组

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

有非零解，并求出它的一般解.

解：首先计算系数行列式：

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

所以这个方程组有非零解的充分必要条件是 $\lambda=0$ 或 1.

当 $\lambda=0$ 时，方程组为：

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

化为同解的阶梯形方程组

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

即

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

所以一般解为

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

其中 x_3 是自由未知量.

当 $\lambda=1$ 时，方程组为：

$$\begin{cases} 4x_1 + x_2 + 2x_3 = 0, \\ x_1 \qquad + x_3 = 0, \\ 6x_1 + x_2 + 4x_3 = 0. \end{cases}$$

化为阶梯形方程组

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

改写成

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

这就是一般解，其中 x_3 是自由未知量.

例 9 试讨论，a 取什么值时，方程组：

$$\begin{cases} ax_1 + x_2 + x_3 = 1, \\ x_1 + ax_2 - x_3 = a, \\ x_1 + x_2 + ax_3 = a^2. \end{cases}$$

有解，并求解.

解：对增广矩阵施行初等行变换：

$$\overline{A} = \begin{bmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & a \\ 1 & 1 & a & a^2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & a & a^2 \\ 1 & a & 1 & a \\ a & 1 & 1 & 1 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 1 & a & a^2 \\ 0 & a-1 & 1-a & a-a^2 \\ 0 & 1-a & 1-a^2 & 1-a^3 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 1 & a & a^2 \\ 0 & a-1 & 1-a & a-a^2 \\ 0 & 0 & 2-a-a^2 & 1+a-a^2-a^3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & a & a^2 \\ 0 & a-1 & 1-a & a(1-a) \\ 0 & 0 & (1-a)(2+a) & (1-a)(1+a)^2 \end{bmatrix} = B.$$

如果 $a \neq 1, -2$，那么 B 可变换成

$$\begin{bmatrix} 1 & 1 & a & a^2 \\ 0 & 1 & -1 & -a \\ 0 & 0 & a+2 & (1+a)^2 \end{bmatrix}.$$

其中 $a+2 \neq 0$. 从而得同解方程组

$$\begin{cases} x_1 + x_2 + ax_3 = a^2, \\ \qquad x_2 - \ x_3 = -a, \\ \quad (a+2)x_3 = (a+1)^2. \end{cases}$$

方程组有唯一解：

$$\begin{cases} x_1 = -\dfrac{a+1}{a+2}, \\ x_2 = \dfrac{1}{a+2}, \\ x_3 = \dfrac{(a+1)^2}{a+2}. \end{cases}$$

如果 $a=1$，那么

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

于是原方程组与

$$x_1 + x_2 + x_3 = 1$$

同解. 它的一般解为

$$x_1 = 1 - x_2 - x_3.$$

其中 x_2, x_3 是自由未知量.

如果 $a=-2$，那么

$$B = \begin{bmatrix} 1 & 1 & -2 & 4 \\ 0 & -3 & 3 & -6 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

所以方程组无解.

归纳上述得：

当 $a \neq 1, -2$ 时，方程组有唯一解：

$$\left(\frac{-a-1}{a+2}, \frac{1}{a+2}, \frac{(a+1)^2}{a+2} \right);$$

当 $a=1$ 时，方程组有无穷多解，其一般解为：

$$x_1 = 1 - x_2 - x_3,$$

其中 x_2, x_3 是自由未知量；

当 $a=-2$ 时，方程组无解.

习　题　2.2(2)

1. 用消元法解下列线性方程组：

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

(3) $\begin{cases} 2x_1+\ x_2+\ x_3=2, \\ x_1+3x_2+\ x_3=5, \\ x_1+\ x_2+5x_3=-7, \\ 2x_1+3x_2-3x_3=14; \end{cases}$

(4) $\begin{cases} 2x_1-\ 2x_2+\ x_3-\ x_4+\ x_5=2, \\ x_1-\ 4x_2+2x_3-2x_4+3x_5=3, \\ 4x_1-10x_2+3x_3-5x_4+7x_5=8, \\ x_1+\ 2x_2-\ x_3+\ x_4-2x_5=-1; \end{cases}$

(5) $\begin{cases} x_1+3x_2-5x_3-5x_4 \qquad =2, \\ x_1+2x_2+2x_3-2x_4+x_5=-2, \\ 2x_1\ +x_2+3x_3-3x_4 \qquad =2, \\ x_1-4x_2+\ x_3+\ x_4-x_5=3, \\ x_1 \qquad +3x_3-\ x_4+x_5=1; \end{cases}$　　(6) $\begin{cases} 2x_1+3x_2-\ x_3+5x_4=0, \\ 3x_1-\ x_2+2x_3-7x_4=0, \\ 4x_1+\ x_2-3x_3+6x_4=0, \\ x_1-2x_2+4x_3-7x_4=0. \end{cases}$

2. 求 λ，使方程组

$$\begin{cases} 2x_1-\ x_2+\ x_3+\ \ x_4=1, \\ x_1+2x_2-\ x_3+\ 4x_4=2, \\ x_1+7x_2-4x_3+11x_4=\lambda. \end{cases}$$

有解，并求解.

3. 求 a 和 b，使齐次方程组

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

有非零解，并求解.

4. 判断 a,b 取什么值时，线性方程组

$$\begin{cases} x_1 - x_2 + x_3 + x_4 + x_5 = 1, \\ 3x_1 - 2x_2 + x_3 + x_4 - 3x_5 = a, \\ \qquad x_2 + 2x_3 + 2x_4 + 6x_5 = 3, \\ 5x_1 - 4x_2 + 3x_3 + 3x_4 - x_5 = b, \end{cases}$$

有解？在有解的情形，求一般解.

5. 证明方程组

$$\begin{cases} x_1 - x_2 = a_1, \\ x_2 - x_3 = a_2, \\ x_3 - x_4 = a_3, \\ x_4 - x_5 = a_4, \\ x_5 - x_1 = a_5. \end{cases}$$

有解的充分必要条件为

$$\sum_{i=1}^{5} a_i = 0.$$

在有解的情形，求它的一般解.

2.3 数　　域

在上节中看到，如果一个线性方程组有无穷多个解，那么可以通过自由未知量来表出其一般解，而自由未知量可以取任意数. 但是，一般线性方程组有其实际背景，未知量的取值范围有一定的限制. 有时候，一个线性方程组有没有解，也和所取的数的范围有关系. 例如，1 元 1 次方程

$$2x = 1 \tag{1}$$

在有理数范围内是有解的：$x = 1/2$. 但是在整数范围内，方程(1)是无解的，这是因为解方程(1)时，必须在方程两边同除以未知量的系数 2，而 1 除以 2 就不是整数了，可见，关键在于两个整数相

除所得的商不一定是整数，这种情形称为整数集合对于除法是不封闭的. 而有理数集合对于除法是封闭的，即任何两个有理数相除（除数不为零）所得的商仍然是有理数. 同样地，实数集合对于除法是封闭的，复数集合也是. 由于解线性方程组时，需要对方程组的系数和常数项进行加、减、乘、除 4 种运算，所以所取的数的范围应当对加、减、乘、除都封闭.

除此以外，还有许多问题与考虑的数的范围有关，同一个问题在不同的数的范围内可能有不同的结论. 为了便于统一地讨论涉及到数的范围的一些问题，我们引进数域这个概念.

定义 4 设 **P** 是由一些数组成的一个集合，其中包含 0 与 1. 如果 **P** 中任意两个数（这两个数也可以相同）的和、差、积、商（除数不能为零）仍在 **P** 中，那么 **P** 就称为一个**数域**.

加、减、乘、除 4 种运算统称**有理运算**. 由数域的定义可知：一个数域中的数经过有理运算（0 不得作为除数）以后，其结果仍在这个数域之中. 也就是说，数域对于有理运算是封闭的.

从定义可以看出：全体有理数组成的集合 **Q**，全体实数组成的集合 **R**，全体复数组成的集合 **C**，都是数域. 但全体整数组成的集合不是数域，因为任意两个整数的商（除数不为 0）不一定还是整数.

下面再看一些例子：

例 1 所有形如

$$a+b\sqrt{2} \qquad (a,b\text{ 是有理数})$$

的数的全体记作 $\mathbf{Q}(\sqrt{2})$. $\mathbf{Q}(\sqrt{2})$ 是一个数域.

证明：首先，因为

$$0=0+0\sqrt{2}$$

$$1=1+0\sqrt{2}$$

所以 0 和 1 在 $\mathbf{Q}(\sqrt{2})$ 中.

任取 $\mathbf{Q}(\sqrt{2})$ 中两个数 $a+b\sqrt{2}$, $c+d\sqrt{2}$ (a,b,c,d 都是有理数)则由于

$$(a+b\sqrt{2})\pm(c+d\sqrt{2})=(a\pm c)+(b\pm d)\sqrt{2};$$

$$(a+b\sqrt{2})\cdot(c+d\sqrt{2})=(ac+2bd)+(ad+bc)\sqrt{2},$$

而且 $a\pm c, b\pm d, ac+2bd, ad+bc$ 都仍然是有理数，可知这两个数的和、差、积仍在 $\mathbf{Q}(\sqrt{2})$ 中.

最后证明：当 $c+d\sqrt{2}\neq 0$ 时，这两个数的商 $\dfrac{a+b\sqrt{2}}{c+d\sqrt{2}}$ 也在 $\mathbf{Q}(\sqrt{2})$ 中. 因为 $c+d\sqrt{2}\neq 0$，所以 c,d 不全为 0，因此 $c-d\sqrt{2}\neq 0$，$c^2-2d^2\neq 0$，于是

$$\frac{a+b\sqrt{2}}{c+d\sqrt{2}}=\frac{(a+b\sqrt{2})(c-d\sqrt{2})}{(c+d\sqrt{2})(c-d\sqrt{2})}$$
$$=\frac{ac-2bd}{c^2-2d^2}+\frac{bc-ad}{c^2-2d^2}\sqrt{2}.$$

而且其中 $\dfrac{ac-2bd}{c^2-2d^2}$, $\dfrac{bc-ad}{c^2-2d^2}$ 也都是有理数，所以可以令

$$\frac{ac-2bd}{c^2-2d^2}=A,\ \frac{bc-ad}{c^2-2d^2}=B.$$

因而 $\dfrac{a+b\sqrt{2}}{c+d\sqrt{2}}=A+B\sqrt{2}$ 也在 $\mathbf{Q}(\sqrt{2})$ 中.

这就证明了 $\mathbf{Q}(\sqrt{2})$ 是一个数域.

例 2 所有形如

$$a+bi \quad (a,b \text{ 是有理数})$$

的数的全体组成一个数域. 证明留给读者.

例 3 全体形如

$$a+b\sqrt{2} \quad (a,b \text{ 是整数})$$

的数组成的集合不是数域. 因为其中任意两个数的商(除数不为

0)不一定在这个集合中.例如 $1+\sqrt{2}$ 与 2 都在这个集合中，而 $\frac{1+\sqrt{2}}{2}=\frac{1}{2}+\frac{1}{2}\sqrt{2}$ 就不属于这个集合了.

数域有一个极为重要的性质：

定理 6 任何一个数域都包含有理数域.

证明：设 **P** 是一个数域.由定义，**P** 包含 1.因为 **P** 包含其中任意两个数的和，因此 **P** 一定包含 $1+1=2,2+1=3,\cdots,(n-1)+1=n,\cdots$.这说明 **P** 一定包含全体自然数.又由定义，0 在 **P** 中，再根据 **P** 中任意两数之差仍在 **P** 中，因此 **P** 一定包含 $0-n=-n(n=1,2,\cdots,)$，即 **P** 包含全体负数，因而 **P** 包含全体整数.而任何一个有理数可以表成两个整数的商，因此也在 **P** 中，所以 **P** 包含有理数域. |

这个定理说明了有理数域是最小的数域.

以前讨论行列式及线性方程组时，只用到系数的有理运算，所以可以把系数限制在某一个数域之中来讨论.系数在数域 **P** 中的矩阵称为数域 **P** 上的矩阵.对数域 **P** 上的矩阵进行初等变换时，所用的数必须是 **P** 中的数.数域 **P** 上矩阵经初等行变换后所得的梯阵及约化梯阵仍都是数域 **P** 上的矩阵.本章以后几节的讨论也假设在一个取定的数域中进行.

习　题　2.3

1. 检验下列集合是不是数域：

(1) 全体偶数；

(2) 全体正实数；

(3) $\mathbf{P}=\{a+b\sqrt{5}\mid a,b$ 是有理数$\}$(这个记号表示 P 是由全体形如 $a+b\sqrt{5}$ 的数组成的，其中 a,b 可以取任意有理数)；

(4) $\mathbf{P}=\{a+b\sqrt[3]{2}+c\sqrt[3]{4}\mid a,b,c$ 是有理数$\}$.

2.4 n 维向量空间

前面介绍了用消元法解一般线性方程组的方法.对于具体地解线性方程组,消元法是一个最有效和最基本的方法.但是,有时候需要直接从原方程组来看它是否有解,消元法就不够了;当方程组的系数中有很多未知参数时,消元法也很难进行;还有,用消元法化方程组成阶梯形时,所作的初等变换可以不同,那么,剩下来的方程个数是不是唯一确定呢?哪些未知量可以取作自由未知量?所有这些问题都还没有解决,因此,还需要对线性方程组作进一步的研究.

一个线性方程组的解的情况是由方程组中方程之间的关系所决定的.例如,在 2.2 节例 7 中,方程组

$$\begin{cases} x_1 + 2x_2 - x_3 + 2x_4 = 1, \\ 2x_1 + 4x_2 + x_3 + x_4 = 5, \\ -x_1 - 2x_2 - 2x_3 + x_4 = -4, \end{cases}$$

的第 1 个方程减去第 2 个方程就等于第 3 个方程.这就是说,第 3 个方程可以去掉而并不影响方程组的解.在那里用初等变换得到的阶梯形方程中只含有两个方程,正是反映了这种情况.因而可以认为,初等变换是揭露方程之间的关系的一种方法.但是初等变换的结果把原方程组改变了,不能从最后的阶梯形方程组看出原来方程组中方程之间的关系.因此,为了直接从原来的线性方程组来讨论它的解的情形,还必须直接研究方程之间的关系.

一个 n 元线性方程

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

可以用一个 $n+1$ 元有序数组

$$(a_1, a_2, \cdots, a_n, b)$$

来表示.方程之间的关系,实际上就是代表它们的 $n+1$ 元有序数

组之间的关系. 因而,在这里先来讨论多元有序数组.

应该指出,多元有序数组不仅可以代表线性方程,而且还与其他方面有极其广泛的联系. 我们在解析几何中已经看到,有些概念用一个数来描述是不够的. 例如,刻画一个点在平面上的位置(即坐标)需要两个数;一个点在空间中的位置需要 3 个数;力学中的力、速度、加速度等等,由于它们既有大小,又有方向,在取定坐标系之后,也需要用 3 个数来刻画,几何中向量的概念正是它们的抽象. 但是,有不少东西,用 3 个数来刻画还是不够的. 例如,刻画一个球的位置和大小(即球心坐标和半径)需要 4 个数;一个 n 元线性方程组的解是由 n 个数组成的,而这 n 个数作为方程组的解是一个整体,分开来谈就没有意义了. 在国民经济问题中,例如一个工厂生产多种产品,为了说明这个工厂的产量,就需要同时指出每种产品的产量;又如一个工厂的原料来自许多地方,于是原料的采购计划就需要同时指出从每个原料产地的采购量,等等. 由此可见,很多实际生活中的数量问题,物理的或数学的对象都可以用多元的有序数组来刻画. 不仅如此,对于这些物理的或数学的对象进行的运算,也常由刻画它们的有序数组的运算来实现. 例如,如果两个力向量的坐标分别为(x_1,y_1,z_1)和(x_2,y_2,z_2),那么这两个力的合力向量的坐标就是$(x_1+x_2,y_1+y_2,z_1+z_2)$,即这两个力的合力的坐标正是这两个力的对应的坐标的和;如果不改变力(x_1,y_1,z_1)的方向,而把它的大小改变成原来力的 k 倍,则所得的力的坐标就是(kx_1,ky_1,kz_1),即把原来的坐标都乘上数 k.

在解线性方程组时,经常用到两种初等变换:

1. 用一个非零常数 k 乘一个方程;

2. 把两个方程加起来.

在第一种变换下,方程

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

变为新方程 $$ka_1x_1+ka_2x_2+\cdots+ka_nx_n=kb.$$

刻画新方程的多元有序数组$(ka_1,ka_2,\cdots,ka_n,kb)$显然是由刻画原方程的有序数组$(a_1,a_2,\cdots,a_n,b)$把每个数乘上 k 而得到的；在第二种变换下，由方程

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

和

$$a'_1x_1+a'_2x_2+\cdots+a'_nx_n=b'$$

得到新方程

$$(a_1+a'_1)x_1+(a_2+a'_2)x_2+\cdots+(a_n+a'_n)x_n=(b+b').$$

这表明，刻画新方程的有序数组

$$(a_1+a'_1,a_2+a'_2,\cdots,a_n+a'_n,b+b')$$

是由刻画原方程的两个有序数组

$$(a_1,a_2,\cdots,a_n,b)$$

和

$$(a'_1,a'_2,\cdots,a'_n,b')$$

把对应位置的数相加而得到.

n 维向量的概念及其运算就是从大量的这一类现象中抽象出来的. 忽略掉各种不同对象的具体区别，保留了它们在数量及运算上的共同点，就得到了 n 维向量的概念以及它们的运算法则.

定义 5 数域 **P** 中 n 个数组成的有序数组

$$(a_1,a_2,\cdots,a_n) \tag{1}$$

称为数域 **P** 上的一个 **n 维向量**. 其中第 i 个数 a_i 称为向量(1)的第 i **个分量**.

几何中的向量可以认为是 n 维向量的特殊情形，即 $n=2,3$ 且 **P** 是实数域的情形. 当 $n>3$ 时，n 维向量就没有直接的几何意义了. 但仍旧沿用几何术语，把它称为向量. 这是因为，一方面，它包括通常的向量作为特殊情形；另一方面，它与通常的向量确实有许多相同的性质.

以后我们用小写希腊字母 α,β,γ 等表示向量.

定义 6 如果 n 维向量

$$\alpha=(a_1,a_2,\cdots,a_n),\quad \beta=(b_1,b_2,\cdots,b_n)$$

的对应分量都相等，即

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

就称这两个向量是**相等的**，记作 $\alpha=\beta$.

下面来定义 n 维向量的运算.

定义 7 设 $\alpha=(a_1,a_2,\cdots,a_n)$，$\beta=(b_1,b_2,\cdots,b_n)$ 是数域 **P** 上的两个 n 维向量. 向量 $(a_1+b_1,a_2+b_2,\cdots,a_n+b_n)$ 称为 α,β 的**和**，记作 $\alpha+\beta$；向量 $(a_1-b_1,a_2-b_2,\cdots,a_n-b_n)$ 称为 α,β 的**差**，记作 $\alpha-\beta$. 即

$$\alpha+\beta=(a_1+b_1,a_2+b_2,\cdots,a_n+b_n);$$

$$\alpha-\beta=(a_1-b_1,a_2-b_2,\cdots,a_n-b_n).$$

如果 k 是数域 **P** 中的数，那么向量 $(ka_1,ka_2,\cdots,ka_n)$ 称为 k 与 α 的**数量乘积**，记作 $k\alpha$，即

$$k\alpha=(ka_1,ka_2,\cdots,ka_n).$$

向量的加法、减法与数乘统称为向量的**线性运算**.

由定义可以推出：向量的加法满足：

交换律：

$$\alpha+\beta=\beta+\alpha. \tag{2}$$

结合律：

$$\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma. \tag{3}$$

向量的数量乘法满足：

$$k(l\alpha)=(kl)\alpha. \tag{4}$$

$$1\cdot\alpha=\alpha. \tag{5}$$

向量的数量乘法与加法满足：

$$k(\alpha+\beta)=k\alpha+k\beta. \tag{6}$$

$$(k+l)\alpha=k\alpha+l\alpha. \tag{7}$$

分量全为零的向量

$$(0,0,\cdots,0)$$

称为**零向量**，记作 **0**；向量

$$(-a_1, -a_2, \cdots, -a_n)$$

称为向量 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)$的**负向量**,记作$-\boldsymbol{\alpha}$.

要注意,维数不同的零向量是不同的向量,不要认为零向量都是一样的. 例如,2 维零向量为(0,0);3 维零向量为(0,0,0),等等. 为了简单起见,我们都用"**0**"来表示.

在向量的运算中,零向量具有与零在数的运算中相类似的性质. 对于任意 $\boldsymbol{\alpha}$,都有

$$\boldsymbol{\alpha}+\mathbf{0}=\boldsymbol{\alpha}. \tag{8}$$

$$\boldsymbol{\alpha}+(-\boldsymbol{\alpha})=\mathbf{0}. \tag{9}$$

(2)~(9)是向量运算的 8 条基本规则,这些规则都可以从定义直接验证,留给读者作为习题.

从定义还可推出:

$$0\boldsymbol{\alpha}=\mathbf{0}. \tag{10}$$

$$(-1)\boldsymbol{\alpha}=-\boldsymbol{\alpha}. \tag{11}$$

$$k\mathbf{0}=\mathbf{0}. \tag{12}$$

如果 $k\neq 0,\boldsymbol{\alpha}\neq\mathbf{0}$,那么

$$k\boldsymbol{\alpha}\neq\mathbf{0}. \tag{13}$$

定义 8 数域 **P** 上的 n 维向量的全体,同时考虑到定义在它们上面的线性运算,称为**数域 P 上的 n 维向量空间**.

在 $n=3$ 时,3 维实向量空间可以认为就是几何空间中全体向量所成的空间.

向量通常写成一行

$$\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n);$$

有时候也写成一列:

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

为了区别起见,前者称为**行向量**,后者称为**列向量**.

习　题　2.4

1. 设 $\alpha=(1,0,-1,2)$, $\beta=(3,2,4,-1)$,计算 $-\alpha,2\alpha,\alpha-\beta,5\alpha+4\beta$.

2. 设 $\alpha=(5,-1,3,2,4)$, $\beta=(3,1,-2,2,1)$,求向量 γ,使
$$3\alpha+\gamma=4\beta.$$

3. 已知
$$\alpha+\beta=(2,1,5,2,0),$$
$$\alpha-\beta=(3,0,1,-1,4).$$
求 α,β.

4. 设
$$3\alpha+4\beta=(2,1,1,2),$$
$$2\alpha+3\beta=(-1,2,3,1).$$
求 α,β.

5. 设 n 维向量
$$\varepsilon_1=(1,0,\cdots,0),$$
$$\varepsilon_2=(0,1,\cdots,0),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\varepsilon_n=(0,0,\cdots,1).$$
求
$$a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n.$$

6. 证明向量的线性运算满足运算规律(2)～(9)式.

2.5　线性相关性

以后讨论都是在一个固定的数域 **P** 上的 n 维向量空间中进行的,不再每次说明了.

在这一节进一步研究向量之间的关系.两个向量之间最简单

的关系是成比例.所谓向量 α 与 β **成比例**,就是说:有一个数 k,使

$$\alpha = k\beta.$$

例如向量(1,2,1,3)与(-2,-4,-2,-6)是成比例的,这是因为

$$(1,2,1,3) = -\frac{1}{2}(-2,-4,-2,-6).$$

在多个向量之间,成比例的关系表现为线性组合.

定义 9 $\alpha,\beta_1,\beta_2,\cdots,\beta_s$ 都是数域 **P** 上 n 维向量空间中的向量.如果有数域 **P** 中的数 $k_1,k_2,\cdots,k_s$,使得

$$\alpha = k_1\beta_1 - k_2\beta_2 + \cdots + k_s\beta_s.$$

就说 α 是 $\beta_1,\beta_2,\cdots,\beta_s$ 的一个**线性组合**.

例 1 设 $\alpha=(-1,-2,-2,1,-4)$,$\beta_1=(1,2,-1,2,1)$,$\beta_2=(2,4,1,1,5)$,因为

$$\alpha = \beta_1 - \beta_2,$$

所以 α 是 β_1,β_2 的线性组合.

例 2 零向量是任意向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 的线性组合.这是因为

$$\mathbf{0} = 0\beta_1 + 0\beta_2 + \cdots + 0\beta_s.$$

例 3 设

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

由于任一个 n 维向量 $\alpha=(a_1,a_2,\cdots,a_n)$都可以表示成

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

所以,任一个 α 都是 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的一个线性组合.

向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 称为 **n 维基本向量***.

如果向量 α 是向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 的一个线性组合,我们也说 α 可以由向量组 $\beta_1,\beta_2,\cdots,\beta_s$ **线性表出**. 例 2 说明零向量可以由任意向量组线性表出,而例 3 说明任一 n 维向量 α 都可以由 n 个 n 维基本向量线性表出,而且表式中的系数刚好是 α 的 n 个分量.

给了一个向量 α 及一组向量 $\beta_1,\beta_2,\cdots,\beta_s$,如何判断 α 能否由 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表出? 这个问题,根据定义,就是要解决能不能找到一组数 $k_1,k_2,\cdots,k_s$,使得 $\alpha=k_1\beta_1+k_2\beta_2+\cdots+k_s\beta_s$ 成立. 那么,这样的 $k_1,k_2,\cdots,k_s$ 又怎样去找呢? 下面通过具体例子来说明.

例 4 设

$$\begin{aligned}\alpha&=(2,3,-4,1),\\ \beta_1&=(1,2,3,-4),\\ \beta_2&=(2,-1,2,5),\\ \beta_3&=(2,-1,5,-4),\end{aligned}$$

问 α 能不能由 β_1,β_2,β_3 线性表出?

解: 如果 α 能由 β_1,β_2,β_3 线性表出,那么,存在 k_1,k_2,k_3,使得

$$\alpha=k_1\beta_1+k_2\beta_2+k_3\beta_3$$

成立,即

$$\begin{aligned}(2,3,-4,1)=&k_1(1,2,3,-4)+k_2(2,-1,2,5)\\ &+k_3(2,-1,5,-4).\end{aligned}$$

比较等号两端的分量,得

$$\begin{cases}2=k_1+2k_2+2k_3,\\ 3=2k_1-k_2-k_3,\\ -4=3k_1+2k_2+5k_3,\\ 1=-4k_1+5k_2-4k_3.\end{cases}$$

* 有的书上称 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为 **n 维单位向量**. 因为通常称长度等于 1 的向量为单位向量,为了区别起见,所以我们把 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 称为基本向量.

但是这个方程组没有解，这说明满足这个条件的 k_1,k_2,k_3 不存在，所以 α 不能被 β_1,β_2,β_3 线性表出.

例 5 设

$$\begin{aligned}\alpha &=(0,4,2,5),\\ \beta_1 &=(1,2,3,1),\\ \beta_2 &=(2,3,1,2),\\ \beta_3 &=(3,1,2,-2),\end{aligned}$$

问 α 能否表成 β_1,β_2,β_3 的线性组合？

解：设 $\alpha=k_1\beta_1+k_2\beta_2+k_3\beta_3$，则

$$\begin{cases} k_1+2k_2+3k_3=0,\\ 2k_1+3k_2+k_3=4,\\ 3k_1+k_2+2k_3=2,\\ k_1+2k_2-2k_3=5.\end{cases}$$

解方程组，得 $k_1=1,\ k_2=1,\ k_3=-1.$

所以 α 能够表成 β_1,β_2,β_3 的线性组合：

$$\alpha=\beta_1+\beta_2-\beta_3.$$

例 6 试将 $\alpha=(2,-1,3,4,-1)$ 表成

$\beta_1=(1,2,-3,1,2)$，$\beta_2=(5,-5,12,11,-5)$，

$\beta_3=(1,-3,6,3,-3)$ 的线性组合.

解：设 $\alpha=k_1\beta_1+k_2\beta_2+k_3\beta_3$，则

$$\begin{cases} k_1+5k_2+k_3=2,\\ 2k_1-5k_2-3k_3=-1,\\ -3k_1+12k_2+6k_3=3,\\ k_1+11k_2+3k_3=4,\\ 2k_1-5k_2-3k_3=-1.\end{cases}$$

这个方程组有无穷多个解. 任取一个解：

$$k_1=1,\ k_2=0,\ k_3=1.$$

则得 $$\alpha = \beta_1 + \beta_3.$$

例 6 说明，α 由 β_1,β_2,β_3 表出的形式不止一种. 如果取另一个解：

$$k_1 = -1,\ k_2 = 1,\ k_3 = -2.$$

那么 α 又可表成

$$\alpha = -\beta_1 + \beta_2 - 2\beta_3.$$

因为上述方程组有无穷多个解，所以 α 由 β_1,β_2,β_3 线性表出的形式也有无穷多种.

由上面一些例子可以看出，设

$$\begin{aligned} \alpha_1 &= (a_{11},a_{12},\cdots,a_{1n}), \\ \alpha_2 &= (a_{21},a_{22},\cdots,a_{2n}), \\ &\cdots\cdots\cdots\cdots \\ \alpha_s &= (a_{s1},a_{s2},\cdots,a_{sn}), \\ \beta &= (b_1,b_2,\cdots,b_n), \end{aligned}$$

则 β 可由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出的充分必要条件是：线性方程组

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \cdots + a_{s1}x_s = b_1, \\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{s2}x_s = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{sn}x_s = b_n \end{cases}$$

有解，而且这个线性方程组的每个解都可取作 β 表成 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的线性组合的系数. 因此，如果这个线性方程组有唯一解. 那么，β 由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出的方法是唯一的；如果这个线性方程组有无穷多解，那么 β 由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出的方法也有无穷多种.

定义 10 如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 中每一个向量 $\alpha_i(i=1,2,\cdots,t)$ 都可以由向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表出，那么就说向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 可以由 $\beta_1,\beta_2,\cdots,\beta_s$ **线性表出**. 如果两个向量组可以互相线性表出，就称它们是**等价的**.

由定义不难看出,每一个向量组都可以由它自身线性表出.如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 可以由向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表出,而向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 又可以由 $\gamma_1,\gamma_2,\cdots,\gamma_p$ 线性表出,那么向量组 $\alpha_1,\cdots,\alpha_t$ 可以由 $\gamma_1,\gamma_2,\cdots,\gamma_p$ 线性表出.这一点可以证明如下:

如果
$$\alpha_i=\sum_{j=1}^{s}k_{ij}\beta_j \quad (i=1,2,\cdots,t),$$
$$\beta_j=\sum_{m=1}^{p}l_{jm}\gamma_m \quad (j=1,2,\cdots,s),$$
那么
$$\alpha_i=\sum_{j=1}^{s}k_{ij}\sum_{m=1}^{p}l_{jm}\gamma_m$$
$$=\sum_{m=1}^{p}\left(\sum_{j=1}^{s}k_{ij}l_{jm}\right)\gamma_m \quad (i=1,2,\cdots,t).$$
这说明向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 中每一个向量都可以由向量组 $\gamma_1,\gamma_2,\cdots,\gamma_p$ 线性表出.因而向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 可以由向量组 $\gamma_1,\gamma_2,\cdots,\gamma_p$ 线性表出.

从上面的讨论可以知道:向量组之间的等价关系有下面 3 个性质:

1. 反身性:每一个向量组都与它自身等价.

2. 对称性:如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 与向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 等价,那么,向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 也与向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 等价.

3. 传递性:如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 与向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 等价;向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 与向量组 $\gamma_1,\gamma_2,\cdots,\gamma_p$ 等价,那么向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 与向量组 $\gamma_1,\gamma_2,\cdots,\gamma_p$ 等价.

根据定义,可知任一 n 维向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 都可由基本向量组 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性表出.

例 7 设

$$\begin{cases} \alpha_1 = (a_{11}, a_{12}, \cdots, a_{1n}), \\ \alpha_2 = (a_{21}, a_{22}, \cdots, a_{2n}), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_n = (a_{n1}, a_{n2}, \cdots, a_{nn}). \end{cases} \tag{1}$$

并设

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

求证：1. 任一 n 维向量都可由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表出.

2. 向量组(1)与基本向量组

$$\begin{cases} \varepsilon_1 = (1, 0, \cdots, 0), \\ \varepsilon_2 = (0, 1, \cdots, 0), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \varepsilon_n = (0, 0, \cdots, 1) \end{cases} \tag{2}$$

等价.

证明：1. 设

$$\beta = (b_1, b_2, \cdots, b_n)$$

是一个 n 维向量.现在证明 β 可以由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表出.

我们知道：β 可由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表出的充分必要条件是能够求得一组数 $k_1, k_2, \cdots, k_n$，使

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n \tag{3}$$

成立，也即满足

$$\begin{cases} a_{11}k_1 + a_{21}k_2 + \cdots + a_{n1}k_n = b_1, \\ a_{12}k_1 + a_{22}k_2 + \cdots + a_{n2}k_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1n}k_1 + a_{2n}k_2 + \cdots + a_{nn}k_n = b_n. \end{cases} \tag{4}$$

考虑线性方程组

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n = b_1, \\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \tag{5}$$

由题设知:它的系数行列式

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

由克莱姆法则知:方程组(5)有唯一解.设$(k_1,k_2,\cdots,k_n)$是它的解么,那么,$k_1,k_2,\cdots,k_n$满足恒等式(4),因而也满足(3),所以$\boldsymbol{\beta}$可以由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性表出.

2. 已知任一个n维向量都可以由向量组(2)线性表出,所以(1)可以由(2)线性表出;另一方面,已证任一n维向量都可由(1)线性表出,故(2)也可由(1)线性表出.所以(1)与(2)是等价的.

习 题 2.5(1)

1. 把向量$\boldsymbol{\beta}$表成向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$的线性组合.

(1) $\boldsymbol{\beta}=(1,2,1,1)$,

$\boldsymbol{\alpha}_1=(1,1,1,1)$, $\boldsymbol{\alpha}_2=(1,1,-1,-1)$,

$\boldsymbol{\alpha}_3=(1,-1,1,-1)$, $\boldsymbol{\alpha}_4=(1,-1,-1,1)$;

(2) $\boldsymbol{\beta}=(0,2,0,-1)$,

$\boldsymbol{\alpha}_1=(1,1,1,1)$, $\boldsymbol{\alpha}_2=(1,1,1,0)$,

$\boldsymbol{\alpha}_3=(1,1,0,0)$, $\boldsymbol{\alpha}_4=(1,0,0,0)$;

(3) $\boldsymbol{\beta}=(0,1,0,1,0)$,

$\boldsymbol{\alpha}_1=(1,1,1,1,1)$, $\boldsymbol{\alpha}_2=(1,2,1,3,1)$,

$\boldsymbol{\alpha}_3=(1,1,0,1,0)$, $\boldsymbol{\alpha}_4=(2,2,0,0,0)$;

2. 等价的向量组所包含的向量个数是否必须相等？试举例说明.

3. 试证明：如果 $\boldsymbol{\beta}$ 可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出，那么对于任意向量 $\boldsymbol{\alpha}_{s+1}$，$\cdots,\boldsymbol{\alpha}_t(t>s)$，$\boldsymbol{\beta}$ 总可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\alpha}_{s+1},\cdots,\boldsymbol{\alpha}_t$ 线性表出.

4. 设 $a_1,a_2,\cdots,a_n$ 是互不相同的数，令

$$\boldsymbol{\alpha}_1=(1,a_1,a_1^2,\cdots,a_1^{n-1})$$
$$\boldsymbol{\alpha}_2=(1,a_2,a_2^2,\cdots,a_2^{n-1})$$
$$\cdots\cdots\cdots\cdots$$
$$\boldsymbol{\alpha}_n=(1,a_n,a_n^2,\cdots,a_n^{n-1}).$$

求证：任一 n 维向量都可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表出.

5. 设

$$\boldsymbol{\beta}_1=a_{11}\boldsymbol{\alpha}_1+a_{12}\boldsymbol{\alpha}_2+\cdots+a_{1m}\boldsymbol{\alpha}_m$$
$$\boldsymbol{\beta}_2=a_{21}\boldsymbol{\alpha}_1+a_{22}\boldsymbol{\alpha}_2+\cdots+a_{2m}\boldsymbol{\alpha}_m$$
$$\cdots\cdots\cdots\cdots$$
$$\boldsymbol{\beta}_m=a_{m1}\boldsymbol{\alpha}_1+a_{m2}\boldsymbol{\alpha}_2+\cdots+a_{mm}\boldsymbol{\alpha}_m,$$

并设行列式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{vmatrix} \neq 0.$$

求证：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 等价.

下面介绍向量组线性相关及线性无关的概念.

定义 11 对于数域 P 上向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s(s\geqslant 1)$，如果有数域 P 中不全为零的数 $k_1,k_2,\cdots,k_s$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s=\mathbf{0},$$

则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ **线性相关**.

例 8 向量组

$$\boldsymbol{\alpha}_1=(2,-1,3,1),$$
$$\boldsymbol{\alpha}_2=(4,-2,5,4),$$
$$\boldsymbol{\alpha}_3=(2,-1,4,-1)$$

是线性相关的，因为有不全为零的数 3，-1，-1，使得

$$3\alpha_1+(-1)\alpha_2+(-1)\alpha_3=\mathbf{0}.$$

例 9 包含零向量的向量组一定是线性相关的. 特别地，一个向量线性相关的充分必要条件是这个向量为零向量.

证明：如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中包含零向量，设 $\alpha_l=\mathbf{0}(1\leqslant l\leqslant s)$，则有不全为零的数 $0,\cdots,0,1$（第 l 个），$0,\cdots,0$，使

$$0\alpha_1+\cdots+0\alpha_{l-1}+1\alpha_l+0\alpha_{l+1}+\cdots+0\alpha_s=\mathbf{0}.$$

所以这个向量组是线性相关的.

对于 $s=1$ 的情形：如果 $\alpha_1=\mathbf{0}$，那么有数 $1\neq0$，使 $1\alpha_1=\mathbf{0}$. 所以 α_1 是线性相关的. 反之，如果 α_1 线性相关，那么有不等于零的数 k，使 $k\alpha_1=\mathbf{0}$，因此必须 $\alpha_1=\mathbf{0}$. 所以 α_1 线性相关的充分必要条件是 $\alpha_1=\mathbf{0}$.

定义 12 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 如果不是线性相关的，就称为**是线性无关的**. 换句话说，如果等式

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=\mathbf{0},$$

只有当
$$k_1=k_2=\cdots=k_s=0$$
时才成立. 那么就称 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是线性无关的.

因为一个向量组或者是线性相关的，或者是线性无关的，二者必居其一，且仅居其一. 因此从关于线性相关的事实，可以推得相应的关于线性无关的事实，反之亦然. 如例 9 中关于线性相关的一些结论，用线性无关的语言来说，就成为：一个线性无关的向量组一定不能包含零向量. 特别地，一个向量是线性无关的充分必要条件是这个向量不等于零. 在以后的讨论中，有时候我们只用一种方式来叙述，希望读者能联想到另一种叙述，并且在应用或证明某个结论时，能够灵活地掌握线性相关或线性无关的概念.

例 10 证明：由 n 维基本向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 组成的向量组是线性无关的.

证明：从 $k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n=\mathbf{0}$，即

$$k_1(1,0,\cdots,0)+k_2(0,1,\cdots,0)+\cdots+k_n(0,0,\cdots,1)$$
$$=(k_1,k_2,\cdots,k_n)=(0,0,\cdots,0),$$

可以推出 $$k_1=k_2=\cdots=k_n=0.$$

因此 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性无关.

一般地,给了一个向量组

$$\begin{cases}\alpha_1=(a_{11},a_{12},\cdots,a_{1n}),\\ \alpha_2=(a_{21},a_{22},\cdots,a_{2n}),\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \alpha_s=(a_{s1},a_{s2},\cdots,a_{sn}),\end{cases} \tag{6}$$

如何来判断这个向量组是否线性相关?根据定义 11,就是要看方程

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_s\boldsymbol{\alpha}_s=\mathbf{0}$$

有没有非零解.按分量写出来,就是

$$\begin{cases}a_{11}x_1+a_{21}x_2+\cdots+a_{s1}x_s=0,\\ a_{12}x_1+a_{22}x_2+\cdots+a_{s2}x_s=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{1n}x_1+a_{2n}x_2+\cdots+a_{sn}x_s=0.\end{cases} \tag{7}$$

因此,向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关的充分必要条件是齐次线性方程组(7)有非零解.换句话说,向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关的充分必要条件是齐次线性方程组(7)只有零解.

从这个结论很容易看出:如果向量组(6)线性无关,那么在每个向量上任意添加一个分量,所得到的 $n+1$ 维的向量组

$$\boldsymbol{\beta}_i=(a_{i1},a_{i2},\cdots,a_{in},a_{i,n+1})\quad(i=1,2,\cdots,s) \tag{8}$$

也线性无关.

事实上,与向量组(8)相对应的齐次线性方程组为:

$$
\begin{cases}
a_{11}x_1 + a_{21}x_2 + \cdots + a_{s1}x_s = 0, \\
a_{12}x_1 + a_{22}x_2 + \cdots + a_{s2}x_s = 0, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{sn}x_s = 0, \\
a_{1,n+1}x_1 - a_{2,n+1}x_2 + \cdots + a_{s,n+1}x_s = 0.
\end{cases} \tag{9}
$$

显然，方程组(9)的解全是方程组(7)的解，如果向量组(6)线性无关，那么方程组(7)只有零解，因而方程组(9)也只有零解，所以向量组(8)也线性无关.

与这个结论相对应的关于线性相关的结论是：如果向量组(6)是线性相关的，那么在每个向量中减去一个分量后，所得到的$n-1$维向量组

$$\gamma_i = (a_{i1}, a_{i2}, \cdots, a_{i,n-1}) \quad (i = 1, 2, \cdots, s)$$

也是线性相关的.

这两个结果可以推广到增添或减少几个分量的情形.

例 11 设

$$\alpha_i = (a_{i1}, a_{i2}, \cdots, a_{in}) \quad (i = 1, \cdots, n),$$

证明向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性相关的充分必要条件是行列式

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

证明：$\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性相关的充分必要条件是齐次线性方程组

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

有非零解.根据本章定理 5 知:这个方程组有非零解的充分必要条件是系数行列式

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

因为 $D=D^T$,所以 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性相关的充分必要条件是 $D=0$.

利用本章定理 4,可以证明下述关于向量组的一个基本性质.

定理 7 设 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 与 $\beta_1,\beta_2,\cdots,\beta_s$ 是两个向量组,如果

1. 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 可以由 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表出,

2. $r>s$,

那么向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 一定线性相关.

证明:由 1,有

$$\alpha_i = \sum_{j=1}^{s} l_{ji}\beta_j \quad (i=1,2,\cdots,r).$$

为了证明 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性相关,只要证明可以找到不全为零的数 $k_1,k_2,\cdots,k_r$,使

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r = \mathbf{0}.$$

为此,作线性组合

$$\begin{aligned} & x_1\alpha_1 + x_2\alpha_2 + \cdots + x_r\alpha_r \\ = & \sum_{i=1}^{r} x_i \sum_{j=1}^{s} l_{ji}\beta_j = \sum_{i=1}^{r}\sum_{j=1}^{s} l_{ji}x_i\beta_j = \sum_{j=1}^{s}\Big(\sum_{i=1}^{r} l_{ji}x_i\Big)\beta_j. \end{aligned}$$

考虑齐次线性方程组

$$\begin{cases} l_{11}x_1 + l_{12}x_2 + \cdots + l_{1r}x_r = 0, \\ l_{21}x_1 + l_{22}x_2 + \cdots + l_{2r}x_r = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ l_{s1}x_1 + l_{s2}x_2 + \cdots - l_{sr}x_r = 0. \end{cases}$$

由题设:这个方程组中未知数的个数 r 大于方程的个数 s,根据定理 4 知这个方程组有非零解. 因而可以取一个非零解 $k_1,k_2,\cdots,k_r$,使

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r = 0.$$

所以 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是线性相关的. |

定理 7 的几何意义是很清楚的. 在 3 维向量的情形,如果 $s=2$,则由 β_1,β_2 线性表出的向量当然都在 β_1,β_2 所张成的平面上,因而这些向量都是共面的,也就是说,当 $r>2$ 时,这些向量是线性相关的.

定理 7 的另一个说法是

推论 1 如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 可由向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表出,而且 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关,那么 $r\leqslant s$.

从定理 7 可以推出.

推论 2 任意 $n+1$ 个 n 维向量必线性相关.

证明: 因为每个 n 维向量都可以被 n 个 n 维基本向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性表出,故由定理 7 可知,任意 $n+1$ 个 n 维向量一定是线性相关的. |

这个结论说明线性无关的 n 维向量组最多包含 n 个向量. 另一方面,例 10 和例 11 说明了的确可以找到 n 个线性无关的 n 维向量. 当 $n=2$ 或 3,且数域 **P** 是实数域时,这个结论的几何意义是:平面上最多有 2 个线性无关的向量,而且的确有 2 个线性无关的向量;空间里最多有 3 个线性无关的向量,而且的确可以找到 3 个线性无关的向量.

推论 3 两个等价的线性无关的向量组,一定包含相同个数的向量.

证明: 设

$$\alpha_1,\alpha_2,\cdots,\alpha_r \tag{10}$$

与

$$\beta_1,\beta_2,\cdots,\beta_s \tag{11}$$

是两个等价的线性无关的向量组. 因为(10)线性无关且可由(11)线性表出,故由推论 1 知 $r\leqslant s$. 同理可知 $s\leqslant r$,所以 $r=s$,即这两个向量组所包含的向量个数是相同的. |

线性相关有一个必要充分条件,这个条件说明了线性相关与线性表出的关系,即下述定理.

定理 8 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 2)$线性相关的充分必要条件是:$\alpha_1,\alpha_2,\cdots,\alpha_s$ 中有一个向量可以被其余的向量线性表出.

证明: 如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关,那么根据定义 11,有不全为零的数 $k_1,k_2,\cdots,k_s$,使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s=\mathbf{0},$$

其中 $k_i\neq 0$. 于是得

$$\boldsymbol{\alpha}_i=-\frac{k_1}{k_i}\boldsymbol{\alpha}_1-\cdots-\frac{k_{i-1}}{k_i}\boldsymbol{\alpha}_{i-1}-\frac{k_{i+1}}{k_i}\boldsymbol{\alpha}_{i+1}-\cdots-\frac{k_s}{k_i}\boldsymbol{\alpha}_s.$$

即 α_i 可被其余的向量线性表出. 这就证明了条件的必要性.

下面证条件的充分性. 设 $\alpha_i(1\leqslant i\leqslant s)$可以被其余的向量线性表出,即

$$\boldsymbol{\alpha}_i=l_1\boldsymbol{\alpha}_1+\cdots+l_{i-1}\boldsymbol{\alpha}_{i-1}+l_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+l_s\boldsymbol{\alpha}_s.$$

那么

$$l_1\boldsymbol{\alpha}_1+\cdots+l_{i-1}\boldsymbol{\alpha}_{i-1}+(-1)\boldsymbol{\alpha}_i+l_{i+1}\boldsymbol{\alpha}_{i+1}+\cdots+l_s\boldsymbol{\alpha}_s=\mathbf{0}.$$

其中系数 $l_1,\cdots,l_{i-1},-1,l_{i+1},\cdots,l_s$ 不全为零. 根据定义 11 可知,$\alpha_1,\alpha_2,\cdots,\alpha_s$ 是线性相关的. |

从定理 8 可知,如果 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是线性无关的,那么 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 中每一个向量都不可能被其余的向量线性表出. 显然,这个条件也是 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关的充分条件.

习 题 2.5(2)

1. 判断下列向量组是否线性相关:

(1) $\alpha_1=(2,2,7,-1)$, $\alpha_2=(3,-1,2,4)$, $\alpha_3=(1,1,3,1)$;

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

$\alpha_3=(3,5,-13,11)$;

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

$\alpha_3=(1,-3,0,1,-2)$, $\alpha_4=(1,5,2,-2,6)$;

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

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

$\alpha_5=(1,-1,3,-1,7)$.

2. 设 $a_1,a_2,\cdots,a_r(r\leqslant n)$是互不相同的数，

$$\alpha_i=(1,a_i,a_i^2,\cdots,a_i^{n-1})\quad(i=1,2,\cdots,r).$$

证明：$\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关.

3. 证明：如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关，而 $\alpha_1,\alpha_2,\cdots,\alpha_s,\beta$ 线性相关，则 β 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出，并且表法唯一.

4. 设 $\alpha_1,\alpha_2,\alpha_3$ 线性无关，证明：$\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_1$ 也线性无关.

5. $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是一组向量. 假设

(1) $\alpha_1\neq0$;

(2) 每个 $\alpha_i(i=2,3,\cdots,s)$都不能被 $\alpha_1,\alpha_2,\cdots,\alpha_{i-1}$线性表出，

求证：$\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关.

6. 用线性相关的概念写出与第 5 题等价的结论.

设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是一个向量组，由其中一部分向量组成的向量组称为这个向量组的一个**部分组**. 从线性相关和线性无关的定义可以看出：如果一个向量组是线性无关的，那么它的部分组也一定是线性无关的；反之，如果一个向量组有一个部分组是线性相关的，那么原来这个向量组也一定是线性相关的. 但是线性相关的向量组的部分组却不一定总是线性相关的. 例如向量组

$$\alpha_1=(2,-1,3,1),$$
$$\alpha_2=(4,-2,5,4),$$
$$\alpha_3=(2,-1,4,-1),$$

由于 $3\alpha_1-\alpha_2-\alpha_3=0$，所以是线性相关的. 但是其部分组 α_1 是线性无关的，部分组 α_1,α_2 也是线性无关的. 在线性无关的部分组中，最重要最有用的是所谓极大线性无关组，它的定义如下：

定义 13 向量组的一个部分组称为它的一个**极大线性无关组**，如果这个部分组本身是线性无关的，但是再从原向量组的其余向量(如果还有的话)中任取一个添进去以后，所得到的部分组都线性相关.

例 12 设 $\alpha_1=(2,-1,3,1)$，$\alpha_2=(4,-2,5,4)$，$\alpha_3=(2,-1,4,-1)$，那么，向量组 $\alpha_1,\alpha_2,\alpha_3$ 的部分组 α_1,α_2 就是一个极大线性无关组. 因为 α_1,α_2 本身线性无关，而将 α_3 添进去以后，$\alpha_1,\alpha_2,\alpha_3$ 就线性相关了. 此外，α_2,α_3 也是一个极大线性无关组.

例 12 说明了一个向量组的极大线性无关组不一定是唯一的.

从定义立刻可以看出：一个线性无关的向量组的极大线性无关组就是这个向量组本身. 这一点也是向量组线性无关的一个充分条件，因为如果一个向量组是线性相关的话，那么这个向量组的极大线性无关组所包含的向量一定少于原来向量组中向量的个数.

完全由零向量组成的向量组没有极大线性无关组，因为它的任何一个部分组都是线性相关的.

显然，任何一个向量组，只要含有非零向量，就一定有极大线性无关组. 下面给出一个求极大线性无关组的方法.

给了一个含有非零向量的向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$，设 α_{i_1} 是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中第一个非零向量. 考虑部分向量组 $\alpha_{i_1},\alpha_{i_1+1},\cdots,\alpha_s$：保留 α_{i_1}，然后从 α_{i_1+1} 起逐个检查. 如果有某个向量可以被前面留下来的那些向量线性表出，就把这个向量去掉；否则，就把这个向量留下来. 最后，如果留下来的向量是 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$. 那么这些向量就组成一个极大线性无关组. 因为 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 中第一个向量 $\alpha_{i_1}\neq\mathbf{0}$，而 $\alpha_{i_2},\cdots,\alpha_{i_r}$ 中每个都不能被前面的向量线性表出. 那么由习

题 2.5(2)第 5 题知，$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$是线性无关的. 而 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中其余的向量都可以被这个部分向量组线性表出，所以任取一个向量添入这个部分组以后，就一定线性相关了.

例 13 设

$\alpha_1=(0,0,0,0,0)$, $\alpha_2=(1,1,-1,0,1)$,

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

$\alpha_5=(0,0,-1,-1,1)$, $\alpha_6=(2,2,-1,1,1)$,

$\alpha_7=(1,0,-1,-1,1)$, $\alpha_8=(2,1,-2,-1,2)$,

求向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6,\alpha_7,\alpha_8$ 的一个极大线性无关组.

解：向量组中第一个向量 $\alpha_1=\mathbf{0}$，将它去掉，考虑 $\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6,\alpha_7,\alpha_8$.

$\alpha_2\neq\mathbf{0}$ 留下；

$\alpha_3=2\alpha_2$ 去掉；

α_4 不能被 α_2 线性表出，留下；

$\alpha_5=\alpha_2-\alpha_4$，去掉；

$\alpha_6=\alpha_2+\alpha_4$，去掉；

α_7 不能被 α_2,α_4 线性表出，留下；

$\alpha_8=\alpha_2+\alpha_7$，去掉.

这样，得到一个部分向量组 $\alpha_2,\alpha_4,\alpha_7$，这就是原向量组的一个极大线性无关组.

如果改变 $\alpha_1,\alpha_2,\cdots,\alpha_8$ 的次序，还能得到另外的极大线性无关组.

设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是任一个向量组. $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$是它的一个线性无关的部分向量组. 改变 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的次序，把 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 排在前面，即

$$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r},\cdots.$$

再用上述方法找出一个极大线性无关组. 从找的方法可以看出：这

个极大线性无关组一定包含向量 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_t}$. 这说明一个向量组的任何一个线性无关的部分组,都可以取作某个极大线性无关组的一部分. 也就是说:一个向量组的任何一个线性无关的部分向量组都可以**扩充**成一个极大线性无关组.

极大线性无关组有下述一些基本性质:

定理 9 (1) 向量组的任意一个极大线性无关组都与向量组本身等价.

(2) 向量组的任意两个极大线性无关组都是等价的,因此都包含相同个数的向量.

证明:(1) 设向量组为 $\alpha_1,\alpha_2,\cdots,\alpha_r,\cdots,\alpha_s$,而且 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是它的一个极大线性无关组. 现在证明这两个向量组是可以互相线性表出的. 因为

$$\alpha_i = 0\alpha_1 + \cdots + 0\alpha_{i-1} + 1\alpha_i + 0\alpha_{i+1} + \cdots + 0\alpha_s \quad (i = 1,2,\cdots,r),$$

所以 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 可以被 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出.

再证 $\alpha_1,\cdots,\alpha_r,\cdots,\alpha_s$ 可以被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表出. 其中,$\alpha_1,\alpha_2,\cdots,\alpha_r$ 中的每一个当然都可以被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表出. 接着再看 $\alpha_{r+1},\cdots,\alpha_s$:由题设,$\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关,根据极大线性无关组的定义,添上向量 $\alpha_l(r<l\leqslant s)$后就线性相关了. 故由习题 2.5(2)第 3 题知:α_l 可以被 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表出. 因此 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 与 $\alpha_1,\cdots,\alpha_r,\cdots,\alpha_s$ 等价.

(2) 因为每个极大线性无关组都与原向量组等价,由等价关系的传递性可知:同一向量组的任意两个极大线性无关组都是等价的,再由定理 7 的推论 3 知:它们包含相同个数的向量. |

定理 9 表明:一个向量组虽然可能有几个极大线性无关组,但各个极大线性无关组所含的向量个数却都是一样的,是与极大线性无关组的选择无关的,它直接反映了向量组本身的性质. 从而可引出向量组的秩的概念.

定义 14　向量组的极大线性无关组所含向量的个数称为这个向量组的**秩**.

只含零向量的向量组的秩规定为零.

我们用 $r\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$ 表示向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩. 例如

$$r\{(2,-1,3,1),(4,-2,5,4),(2,-1,4,-1)\}=2.$$

因为线性无关的向量组就是它自身的极大线性无关组，所以一个向量组线性无关的充分必要条件是它的秩等于它所含的向量个数.

从定理 9 我们知道，每一向量组都与它的极大线性无关组等价，且由等价的传递性可知，任意两个等价向量组的极大线性无关组也等价. 所以，等价的向量组必有相同的秩.

最后介绍关于 n 维向量的一个很有用的结论：设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 $s(0<s<n)$ 个线性无关的 n 维向量. 那么可以找到 $n-s$ 个 n 维向量 $\alpha_{s+1},\cdots,\alpha_n$，使得 $\alpha_1,\alpha_2,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$ 是线性无关的.

证明如下：任取 n 个线性无关的 n 维向量 $\beta_1,\beta_2,\cdots,\beta_n$. 构成向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s,\beta_1,\beta_2,\cdots,\beta_n$. 这个向量组的秩等于 n. 按照这个次序逐个去掉可以由前面的向量线性表出的向量，得到一个极大线性无关组 $\alpha_1,\alpha_2,\cdots,\alpha_s,\beta_{i_1},\beta_{i_2},\cdots,\beta_{i_{n-s}}$. 令

$$\alpha_{s+k}=\beta_{i_k}\quad(k=1,2,\cdots,n-s),$$

则 $\alpha_{s+1},\cdots,\alpha_n$ 即为满足要求的向量.

习　题　2.5(3)

1. 求下列向量组的一个极大线性无关组与秩：

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

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

(2) $\alpha_1=(1,1,1,1)$，　$\alpha_2=(1,1,-1,-1)$，

$\alpha_3=(1,-1,-1,1)$，　$\alpha_4=(-1,-1,-1,1)$；

(3) $\alpha_1=(0,4,10,1)$，　$\alpha_2=(4,8,18,7)$，

$\alpha_3=(10,18,40,17)$, $\alpha_4=(1,7,17,3)$.

2. 设 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个线性无关部分向量组，试证明：α_{i_1}，$\alpha_{i_2},\cdots,\alpha_{i_r}$ 是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个极大线性无关组的充分必要条件是：α_1，$\alpha_2,\cdots,\alpha_s$ 中每个向量都可以被它们线性表出.
3. 已知向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩为 r. 证明：$\alpha_1,\alpha_2,\cdots,\alpha_s$ 中任意 r 个线性无关的向量都构成它的一个极大线性无关组.
4. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩为 r，且 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中的 r 个向量，使得 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中每个向量都可被它们线性表出，证明：$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个极大线性无关组.
5. 设向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 可以由向量组 $\beta_1,\beta_2,\cdots,\beta_t$ 线性表出，求证：$r\{\alpha_1,\alpha_2,\cdots,\alpha_s\}\leqslant r\{\beta_1,\beta_2,\cdots,\beta_t\}$.
6. 已知 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 与 $\alpha_1,\alpha_2,\cdots,\alpha_t,\alpha_{t+1},\cdots,\alpha_s$ 有相同的秩，试证明：$\alpha_1,\alpha_2,\cdots,\alpha_t$ 与 $\alpha_1,\alpha_2,\cdots,\alpha_t,\alpha_{t+1},\cdots,\alpha_s$ 等价.

2.6 矩阵的秩

为了计算向量组的秩以及讨论线性方程组，我们引进矩阵的秩的概念.

设

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

是一个 $s\times n$ 矩阵.

矩阵 A 的每一行都是一个 n 元有序数组，因而是一个 n 维向量，称为矩阵 A 的**行向量**. A 有 s 行，因此 A 有 s 个行向量：

$$\begin{aligned} \alpha_1 &= (a_{11},a_{12},\cdots,a_{1n}), \\ &\cdots\cdots \\ \alpha_s &= (a_{s1},a_{s2},\cdots,a_{sn}). \end{aligned} \tag{1}$$

向量组 $\alpha_1,\cdots,\alpha_s$ 称为矩阵 A 的**行向量组**.

同样,矩阵 A 的每一列都是一个 s 元有序数组,因而是一个 s 维向量,称为矩阵 A 的**列向量**. A 有 n 列,因此,A 有 n 个列向量:

$$\beta_1=\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{s1}\end{pmatrix},\cdots,\quad \beta_n=\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{sn}\end{pmatrix}. \tag{2}$$

向量组 $\beta_1,\cdots,\beta_n$ 称为矩阵 A 的**列向量组**.

定义 15 矩阵 A 的行向量组的秩称为 A 的**行秩**,A 的列向量组的秩称为 A 的**列秩**.

例 1 设

$$A=\begin{bmatrix}1&-1&2&3\\0&3&5&7\\0&0&1&-3\end{bmatrix},$$

则 A 的行向量为:

$$\alpha_1=(1,-1,2,3),$$
$$\alpha_2=(0,3,5,7),$$
$$\alpha_3=(0,0,1,-3).$$

这 3 个向量线性无关,所以

$$A\text{ 的行秩}=\mathrm{r}\{\alpha_1,\alpha_2,\alpha_3\}=3.$$

A 的列向量为

$$\beta_1=(1,0,0),$$
$$\beta_2=(-1,3,0),$$
$$\beta_3=(2,5,1),$$
$$\beta_4=(3,7,-3).$$

容易看出 β_1,β_2,β_3 是列向量组 $\beta_1,\beta_2,\beta_3,\beta_4$ 的一个极大线性无关组,所以

$$A\text{ 的列秩} = r\{\beta_1,\beta_2,\beta_3,\beta_4\} = 3.$$

下面来证明矩阵的行秩与列秩总相等. 为此,用行列式把矩阵的行和列联系起来.

定义 16 在一个 $s\times n$ 矩阵 A 中,任意取定 k 行和 k 列,由位于这些行与列的交点上的 k^2 个元素按原来的次序组成的 k 阶行列式,称为 A 的一个 k 阶子式.

定义中的 k 当然必须满足 $k\leqslant \min(s,n)$.

例 2 在矩阵

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

中取定第 1 和 3 行,第 2 和 4 列,位于这些行与列交点上的 4 个元素组成的 2 阶行列式

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

就是 A 的一个 2 阶子式. 另外,如果选定第 1,2,4 行,第 2,4,5 列,就得到 A 的一个 3 阶子式

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

因为行与列的选法很多,所以 k 阶子式也是很多的. 特别是,A 的每个元素都是 A 的一个 1 阶子式.

定义 17 矩阵 A 的不等于零的子式的最高阶数称为 A 的**行列式秩**.

元素全为 0 的矩阵称为零矩阵. 零矩阵的行列式秩规定为 0.

例如在例 1 中,A 有一个 3 阶子式

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

而 A 的子式最多只能是 3 阶，所以

$$A \text{ 的行列式秩} = 3.$$

例 2 中的 A，其行秩、列秩和行列式秩都等于 3. 读者可以自己计算一下.

矩阵的秩与行列式的关系，可以从下面的定理看出.

定理 10 矩阵的行秩、列秩和行列式秩都相等.

证明：如果矩阵 $A=0$，显然 A 的行秩、列秩和行列式秩都等于 0. 下面设 $A\neq 0$. 先来证明矩阵的行秩等于行列式秩. 设矩阵

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

的行秩等于 r，那么 A 的任意 $r+1$ 个行向量都是线性相关的，所以 A 的任一个 $r+1$ 阶子式都等于零. 因此，A 的阶数大于 r 的子式都等于零.

下面证明 A 有一个 r 阶子式不等于零：因为 A 的秩等于 r，所以 A 有 r 个线性无关的行向量，不妨设 A 的前 r 个行向量

$$\begin{aligned} \alpha_1 &= (a_{11}, a_{12}, \cdots, a_{1n}), \\ \alpha_2 &= (a_{21}, a_{22}, \cdots, a_{2n}), \\ &\cdots\cdots\cdots\cdots \\ \alpha_r &= (a_{r1}, a_{r2}, \cdots, a_{rn}) \end{aligned}$$

线性无关. 于是从

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r = 0$$

可推出

$$k_1 = k_2 = \cdots = k_r = 0.$$

即齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \cdots + a_{r1}x_r = 0 \\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{r2}x_r = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{rn}x_r = 0 \end{cases}$$

只有零解. 因此由定理 4，这个方程组的系数矩阵

$$A_1 = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{r1} \\ a_{12} & a_{22} & \cdots & a_{r2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{rn} \end{bmatrix}$$

的行秩等于未知量的个数 r，所以 A_1 中有 r 个行线性无关，即 A_1 有一个 r 阶子式不等于零. 这个子式当然也是 A 的一个子式. 这就证明了 A 有一个 r 阶子式不等于零.

下面证明 A 的列秩也等于行列式秩.

把 A 的行写成列，得到另一个矩阵

$$A^{\mathrm{T}} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{s1} \\ a_{12} & a_{22} & \cdots & a_{s2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{sn} \end{bmatrix}.$$

根据 A^{T} 的做法，把 A 的一个 k 阶子式行列互换，就得到 A^{T} 的一个 k 阶子式；把 A' 的一个 k 阶子式行列互换，就得到 A 的一个 k 阶子式. 因此，A 和 A' 有相同的子式. 所以 A 的行列式秩等于 A^{T} 的行列式秩.

因为 A^{T} 的行向量组就是 A 的列向量组，所以 A 的列秩就等于 A 的行列式秩. |

根据定理 10. 把矩阵 A 的行秩、列秩和行列式秩统称为矩阵 A 的**秩**，记作 $r(A)$.

容易看出，矩阵的初等行变换把矩阵 A 的行向量组化为等价

的向量组，因此矩阵的初等行变换不改变矩阵的**秩**，可以用来计算矩阵的秩和向量组的秩.

例 3 设

$$\alpha_1=(1,-2,3,-1,-1),$$
$$\alpha_2=(2,-1,1,0,-2),$$
$$\alpha_3=(-2,-5,8,-4,3),$$
$$\alpha_4=(1,1,-1,1,-2),$$

求向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的秩.

解：以 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 为行构成一个矩阵：

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

则 $r\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}=A$ 的行秩. 对 A 作初等行变换：

$$A=\begin{bmatrix}1&-2&3&-1&-1\\2&-1&1&0&-2\\-2&-5&8&-4&3\\1&1&-1&1&-2\end{bmatrix}\to\begin{bmatrix}1&-2&3&-1&-1\\0&3&-5&2&0\\0&-6&9&-4&1\\0&3&-4&2&-1\end{bmatrix}$$

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

所以 A 的行秩$=3$. 即 $r\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}=3$. 并可看出 $\alpha_1,\alpha_2,\alpha_3$ 是一个极大线性无关组.

既然矩阵 A 的秩也等于 A 的列向量组的秩，所以也可以通过化简矩阵的列向量组来计算矩阵的秩. 类似于矩阵的初等行变换，可以定义数域 $\mathbf{P}$ 上矩阵的**初等列变换**如下：

1. 以 $\mathbf{P}$ 中一个非零的数乘矩阵的一列；

2. 把矩阵的某一列乘 k 后加到另一列上(这里 k 是数域 $\mathbf{P}$ 中任意一个数);

3. 互换矩阵中两列的位置.

矩阵经过初等列变换后,所得的矩阵的列向量组与原矩阵的列向量组是等价的,所以初等列变换不改变矩阵的秩,故也可以应用初等列变换来化简矩阵,求矩阵的秩.有时候,同时施行初等行变换和初等列变换,计算矩阵的秩更为方便.

矩阵的初等行变换和初等列变换统称为矩阵的**初等变换**.

定理 10 给出了一个不必经过初等变换就可直接从矩阵的元素来求它的秩的方法.这一结论在线性方程组的讨论中要用到.

例 4 设 a,b,c,d 是 4 个不同的数,试证明

$$\begin{aligned}\alpha_1&=(1,1,1,1),\\ \alpha_2&=(a,b,c,d),\\ \alpha_3&=(a^2,b^2,c^2,d^2),\\ \alpha_4&=(a^3,b^3,c^3,d^3)\end{aligned}$$

线性无关.

证明:作矩阵

$$A=\begin{bmatrix}1&1&1&1\\ a&b&c&d\\ a^2&b^2&c^2&d^2\\ a^3&b^3&c^3&d^3\end{bmatrix}.$$

因为 a,b,c,d 各不相同,所以 A 的行列式

$$\begin{vmatrix}1&1&1&1\\ a&b&c&d\\ a^2&b^2&c^2&d^2\\ a^3&b^3&c^3&d^3\end{vmatrix}\neq 0.$$

A 的秩 $=A$ 的行列式秩 $=4$,A 的行秩 $=4$. A 的行线性无关,即 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 线性无关.

习 题 2.6

1. 设

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

(1) 计算 A 的全部 4 阶子式；

(2) 求 $r(A)$.

2. 计算下列矩阵的秩：

(1) $\begin{bmatrix}3&-7&6&1&5\\1&-2&4&-1&3\\-1&1&-10&5&-7\\4&-11&-2&8&0\end{bmatrix}$；

(2) $\begin{bmatrix}3&2&-1&2&0&1\\4&1&0&-3&0&2\\2&-1&-2&1&1&-3\\3&1&3&-9&-1&6\\3&-1&5&7&2&-7\end{bmatrix}$.

3. 计算下列向量组的秩，并且判断该向量组是否线性相关.

(1) $\alpha_1=(1,-1,2,3,4)$， $\alpha_2=(3,-7,8,9,13)$，
$\alpha_3=(-1,-3,0,-3,-3)$， $\alpha_4=(1,-9,6,3,6)$；

(2) $\alpha_1=(1,-3,2,-1)$， $\alpha_2=(-2,1,5,3)$，
$\alpha_3=(4,-3,7,1)$， $\alpha_4=(-1,-11,8,-3)$，
$\alpha_5=(2,-12,30,6)$.

4. 设

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

求 α_3,α_4 使得 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 线性无关.

2.7 线性方程组有解判别定理与解的结构

前面几节论述了关于 n 维向量以及矩阵的秩的一些基本性质. 这一节将应用这些结论来讨论线性方程组有解的条件以及求解的方法.

定理 11 (线性方程组有解判别定理)线性方程组

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

有解的充分必要条件是:它的系数矩阵

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

与增广矩阵

$$\bar{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} & b_s \end{bmatrix}$$

有相同的秩.

证明: 令

$$\alpha_1 = (a_{11}, a_{21}, \cdots, a_{s1}),$$
$$\alpha_2 = (a_{12}, a_{22}, \cdots, a_{s2}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_n = (a_{1n}, a_{2n}, \cdots, a_{sn}),$$
$$\beta = (b_1, b_2, \cdots, b_s),$$

则线性方程组(1)有解的充分必要条件是 β 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出.

如果 β 可以由 $\alpha_1,\cdots,\alpha_n$ 线性表出,则向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n,\beta$ 等价,因此 A 与 $\overline{A}$ 的秩相等.

如果 A 的秩等于 $\overline{A}$ 的秩,则 $\mathrm{r}\{\alpha_1,\alpha_2,\cdots,\alpha_n\}=\mathrm{r}\{\alpha_1,\cdots,\alpha_n,\beta\}=r$,取 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的一个极大线性无关组 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$. 这也是 $\alpha_1,\cdots,\alpha_n,\beta$ 的一个线性无关的部分组. 由 $\mathrm{r}\{\alpha_1,\cdots,\alpha_n,\beta\}=r$,知 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是 $\alpha_1,\cdots,\alpha_n,\beta$ 的一个极大线性无关组. 因此,β 可由 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性表出,也可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出. |

例 1 设 a,b,c,d 是各不相同的数,试证:线性方程组

$$\begin{cases} x_1+x_2+x_3=1, \\ ax_1+bx_2+cx_3=d, \\ a^2x_1+b^2x_2+c^2x_3=d^2, \\ a^3x_1+b^3x_2+c^3x_3=d^3, \end{cases}$$

无解.

证明: 这个线性方程组的系数矩阵 A 及增广矩阵 $\overline{A}$ 分别为:

$$A=\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{bmatrix}, \qquad \overline{A}=\begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{bmatrix}.$$

因为 a,b,c,d 各不相同,所以范德蒙行列式

$$|\overline{A}|=\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{vmatrix}\neq 0.$$

$\mathrm{r}(\overline{A})=4$.

而系数矩阵 A 只有 3 列,所以 $\mathrm{r}(A)\leqslant 3$,因此 A 与 $\overline{A}$ 的秩不相同,

从而此线性方程组无解.

例 2 讨论 a 取什么值时,线性方程组:

$$\begin{cases} ax_1 + x_2 + x_3 = 1, \\ x_1 + ax_2 + x_3 = a, \\ x_1 + x_2 + ax_3 = a^2, \end{cases}$$

有解? 并在有解时,求出一般解.

解: 先计算系数行列式

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

(1) 当 $a \neq 1$,并且 $a \neq -2$ 时,$D \neq 0$,据克莱姆法则知此方程组有唯一解.

又因

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

$$D_2 = \begin{vmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & a^2 & a \end{vmatrix} = (a-1)^2,$$

$$D_3 = \begin{vmatrix} a & 1 & 1 \\ 1 & a & a \\ 1 & 1 & a^2 \end{vmatrix} = (a+1)^2(a-1)^2,$$

所以解为

$$x_1 = \frac{D_1}{D} = -\frac{a+1}{a+2},$$

$$x_2 = \frac{D_2}{D} = \frac{1}{a+2},$$

$$x_3=\frac{D_3}{D}=\frac{(a+1)^2}{a+2}.$$

(2) 当 $a=1$ 时，系数矩阵 A 为

$$A=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}.$$

显然 $r(A)=1$. 增广矩阵 $\overline{A}$ 这时成为

$$\overline{A}=\begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix}.$$

同理，$r(\overline{A})=1$. 所以此方程组有解. 此时，方程组与

$$x_1+x_2+x_3=1$$

同解. 因此，一般解为

$$x_1=1-x_2-x_3\quad(x_2,x_3\text{ 为自由未知量})$$

(3) 当 $a=-2$ 时，增广矩阵 $\overline{A}$ 成为

$$\overline{A}=\begin{bmatrix}-2&1&1&1\\1&-2&1&-2\\1&1&-2&4\end{bmatrix}.$$

$\overline{A}$ 的下述 3 阶子式

$$\begin{vmatrix}1&1&1\\-2&1&-2\\1&-2&4\end{vmatrix}=9\neq 0.$$

所以 $r(\overline{A})=3$. 而 A 的唯一的 3 阶子式 $D=0$，因此 $r(A)<3$，从而 A 与 $\overline{A}$ 的秩不相等，所以此时方程组无解.

当线性方程组有解时，可以用系数矩阵的秩判断它何时有唯一解，何时有无穷多个解.

应用消元法求解时曾看到，线性方程组(1)有解时，如果它化成的阶梯形方程组的方程个数 r 等于未知量个数 n，则方程组(1)

有唯一解；如果 $r<n$，则方程组(1)有无穷多个解. 因为阶梯形方程组的方程个数等于相应的阶梯形矩阵的非零行行数，即等于原方程组的增广矩阵的秩，而在方程组(1)有解时，它也就等于系数矩阵的秩. 因此得到：

命题 线性方程组(1)有解时，如果它的系数矩阵的秩 r 等于未知量个数 n，则方程组(1)有唯一解；如果 $r<n$，则方程组(1)有无穷多个解.

把这个命题用到齐次线性方程组上就得到下述定理.

定理 12 齐次线性方程组

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

有非零解的充分必要条件是：它的系数矩阵的秩 r 小于未知量个数 n.

推论 齐次线性方程组

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

有非零解的充分必要条件是：它的系数行列式等于 0.

例 3 求 λ 使齐次线性方程组

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

有非零解，并求一般解.

解：计算系数行列式：

$$D=\begin{vmatrix}\lambda+3 & 1 & 2\\ \lambda & \lambda-1 & 1\\ 3(\lambda+1) & \lambda & \lambda+3\end{vmatrix}=\begin{vmatrix}\lambda & 1 & 2\\ 0 & \lambda-1 & 1\\ \lambda & \lambda & \lambda+3\end{vmatrix}$$

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

所以，当 $D=0$，即 $\lambda=0$ 或 1 时，此方程组有非零解.

当 $\lambda=0$ 时，方程组成为：

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

其一般解为：

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

其中 x_3 为自由未知量.

当 $\lambda=1$ 时，方程组成为：

$$\begin{cases}4x_1+x_2+2x_3=0,\\ x_1\qquad\ +x_3=0,\\ 6x_1+x_2+4x_3=0.\end{cases}$$

其一般解为：

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

其中 x_3 为自由未知量.

应用矩阵的行列式秩的概念，可以直接从矩阵的元素来计算矩阵的秩，因而也就给出了一个直接从线性方程组的系数和常数项来判断这个线性方程组有没有解的方法. 不但如此，在线性方程组有解的情形下，还可以直接从原方程组求出一般解. 下面就来介绍这个方法.

设线性方程组

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

有解,其系数矩阵 A 和增广矩阵 $\overline{A}$ 的秩都等于 r. 根据定理 10,矩阵 A 有一个不等于零的 r 阶子式 D,当然,D 也是 $\overline{A}$ 的一个不等于零的子式. 为了讨论方便,不妨假设 D 位于 $\overline{A}$ 的左上角,即

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

于是,$\overline{A}$ 的前 r 行就是 $\overline{A}$ 的行向量组的一个极大线性无关组,第 $r+1,\cdots,s$ 行都可以由它们线性表出,因此,方程组(1)与方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1+a_{r2}x_2+\cdots+a_{rn}x_n=b_r, \end{cases} \tag{2}$$

同解.

如果 $r=n$,则由克莱姆法则知:方程组(2)有唯一解,也就是方程组(1)有唯一解. 这个解可以应用克莱姆法则从方程组(2)解出来.

如果 $r<n$,把方程组(2)改写成

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1r}x_r=b_1-a_{1,r+1}x_{r+1}-\cdots-a_{1n}x_n, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2r}x_r=b_2-a_{2,r+1}x_{r+1}-\cdots-a_{2n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1+a_{r2}x_2+\cdots+a_{rr}x_r=b_r-a_{r,r+1}x_{r+1}-\cdots-a_{rn}x_n. \end{cases} \tag{3}$$

把方程组(3)看作 $x_1,x_2,\cdots,x_r$ 的一个方程组,它的系数行列式 $D \neq 0$. 任意给定了 $x_{r+1},x_{r+2},\cdots,x_n$ 的一组值后,由克莱姆法则,方程组(3)也即方程组(1)都有唯一解. $x_{r+1},\cdots,x_n$ 就是方程组(1)的一组自由未知量. 从而得线性方程组(1)的一般解为:

$$\begin{cases} x_1 = d_1 + c_{1,r+1}x_{r+1} + \cdots + c_{1n}x_n, \\ x_2 = d_2 + c_{2,r+1}x_{r+1} + \cdots + c_{2n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_r = d_r + c_{r,r+1}x_{r+1} + \cdots + c_{rn}x_n. \end{cases} \tag{4}$$

其中 $x_{r+1},\cdots,x_n$ 是自由未知量.

习　题　2.7(1)

1. 判断下列线性方程组有没有解.

$$\begin{cases} x_1 + x_2 + x_3 = 1, \\ 3x_1 + 5x_2 + 2x_3 = 4, \\ 9x_1 + 25x_2 + 4x_3 = 16, \\ 27x_1 + 125x_2 + 8x_3 = 64. \end{cases}$$

2. 讨论 a 取什么值时,线性方程组:

$$\begin{cases} 3ax_1 + (2a+1)x_2 + (a+1)x_3 = a, \\ (2a-1)x_1 + (2a-1)x_2 + (a-2)x_3 = a+1, \\ (4a-1)x_1 + 3ax_2 + 2ax_3 = 1, \end{cases}$$

有解,并在有解时求出一般解.

3. λ 取何值时,下述齐次线性方程组:

$$\begin{cases} (\lambda-2)x_1 - 3x_2 - 2x_3 = 0, \\ -x_1 + (\lambda-8)x_2 - 2x_3 = 0, \\ 2x_1 + 14x_2 + (\lambda+3)x_3 = 0, \end{cases}$$

有非零解,并且求出它的一般解.

现在讨论线性方程组的解的结构问题. 在方程组只有唯一解的情形,当然没有什么结构问题,只要把这个解求出来就行了. 在

方程组有无穷多个解的情况下，所谓解的结构问题，就是解与解之间的关系问题. 下面将证明，虽然这时有无穷多个解，但是全部的解都可以用有限多个解表示出来. 这就是下面要讨论的中心问题以及所得到的结果.

以下的讨论当然都是对于有解的情况讲的，这一点就不再每次都说明了.

n 元线性方程组的解可以看作 n 维向量. 在解不是唯一的情况下，同一个线性方程组的解向量之间有什么关系呢?先来讨论齐次线性方程组的情形. 设

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = 0, \end{cases} \tag{5}$$

是一个齐次线性方程组，它的解集合具有下面两个重要性质：

1. 两个解的和还是方程组的解.

证明：设 $(k_1, k_2, \cdots, k_n)$ 与 $(l_1, l_2, \cdots, l_n)$ 是方程组(5)的两个解，这就是说，把它们代入方程组(5)后，每个方程都成为恒等式，即

$$\sum_{j=1}^{n} a_{ij}k_j = 0 \quad (i = 1, 2, \cdots, s);$$

$$\sum_{j=1}^{n} a_{ij}l_j = 0 \quad (i = 1, 2, \cdots, s).$$

把这两个解的和

$$(k_1 + l_1, k_2 + l_2, \cdots, k_n + l_n) \tag{6}$$

代入方程组，得

$$\sum_{j=1}^{n} a_{ij}(k_j + l_j) = \sum_{j=1}^{n} a_{ij}k_j + \sum_{j=1}^{n} a_{ij}l_j = 0 + 0 = 0$$
$$(i = 1, 2, \cdots, s).$$

这说明(6)确实是方程组(5)的解.|

2. 一个解的倍数还是方程组的解.

证明:设$(k_1,k_2,\cdots,k_n)$是方程组(5)的一个解,那么

$$\sum_{j=1}^{n}a_{ij}(ck_j)=c\sum_{j=1}^{n}a_{ij}k_j=c\cdot 0=0\quad(i=1,2,\cdots,s).$$

这说明这个解的倍数$(ck_1,ck_2,\cdots,ck_n)$还是方程组的解.|

综合以上两点,即知:对于齐次线性方程组,解的线性组合还是方程组的解.这个性质说明:如果找出了方程组的几个解,那么这些解的所有可能的线性组合就给出了很多的解.因此,我们要问:齐次线性方程组的全部解是否能够通过它的有限的几个解的线性组合得出来?回答是肯定的.为此,下面先给出一个定义.

定义 18 设$\eta_1,\eta_2,\cdots,\eta_t$是齐次线性方程组(5)的一组解,如果

1. $\eta_1,\eta_2,\cdots,\eta_t$线性无关;

2. 方程组(5)的任一个解都能表成$\eta_1,\eta_2,\cdots,\eta_t$的线性组合.

则$\eta_1,\eta_2,\cdots,\eta_t$称为方程组(5)的一个**基础解系**.

定义中的条件2保证了方程组(5)的全部解都可以由$\eta_1,\eta_2,\cdots,\eta_t$线性表出.而条件1是为了保证基础解系中没有多余的解.否则,若$\eta_1,\eta_2,\cdots,\eta_t$是线性相关的话,那么其中有一个可以表成其他解的线性组合,譬如说,η_t可以表成$\eta_1,\eta_2,\cdots,\eta_{t-1}$的线性组合.于是$\eta_1,\eta_2,\cdots,\eta_{t-1}$也具有性质2.

下面来证明只要齐次线性方程组有非零解,那么它一定有基础解系.

定理 13 在齐次线性方程组有非零解的情况下,它一定有基础解系,并且基础解系所含解的个数等于$n-r$,这里r表示系数矩阵的秩.

证明:设方程组(5)的系数矩阵的秩为r,不妨假设左上角的r阶子式不等于零.于是按上节最后一段的分析,方程组(5)可以

改写成

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1r}x_r=-a_{1,r+1}x_{r+1}-\cdots-a_{1n}x_n,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2r}x_r=-a_{2,r+1}x_{r+1}-\cdots-a_{2n}x_n,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\a_{r1}x_1+a_{r2}x_2+\cdots+a_{rr}x_r=-a_{r,r+1}x_{r+1}-\cdots-a_{rn}x_n.\end{cases}\tag{7}$$

如果 $r=n$，那么方程组有唯一解，即只有零解，当然没有基础解系；因此，如果方程组(5)有非零解，那么一定 $r<n$. 此时便有 $n-r$ 个自由未知量. 我们知道，把自由未知量的任意一组值(c_{r+1}，$\cdots,c_n$)代入(7)，应用克莱姆法则解出唯一的 $x_1,x_2,\cdots,x_r$，就得到方程组(5)的一个解. 这就是说，只要自由未知量的值一样，方程组(5)的两个解就完全一样. 特别地，如果在一个解中，自由未知量的值全为零，那么这个解就一定是零解.

在(7)中分别用 $n-r$ 组数

$$(1,0,\cdots,0),(0,1,\cdots,0),\cdots,(0,0,\cdots,1)\tag{8}$$

代替自由未知量($x_{r+1},\cdots,x_n$)，得出方程组(7)，也就是方程组(5)的 $n-r$ 个解：

$$\begin{cases}\eta_1=(c_{11},\cdots,c_{1r},1,0,\cdots,0),\\\eta_2=(c_{21},\cdots,c_{2r},0,1,\cdots,0),\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\\eta_{n-r}=(c_{n-r,1},\cdots,c_{n-r,r},0,0,\cdots,1).\end{cases}\tag{9}$$

现在来证明：(9)就是方程组(5)的一个基础解系. 首先证明 $\eta_1,\eta_2,\cdots,\eta_{n-r}$ 是线性无关的. 因为向量组(9)可以看成是由向量组(8)在每个向量中添加 r 个分量而得到的. 而向量组(8)是 $n-r$ 个 $n-r$ 维基本向量，所以是线性无关的，因此向量组(9)也是线性无关的.

下面来证明方程组(5)的任一个解都可以由 $\eta_1,\eta_2,\cdots,\eta_{n-r}$ 线性表出. 设

$$\eta=(c_1,\cdots,c_r,c_{r+1},\cdots,c_n) \tag{10}$$

是(5)的一个解. 由于 $\eta_1,\eta_2,\cdots,\eta_{n-r}$都是方程组(5)的解,所以它们的线性组合

$$c_{r+1}\eta_1+c_{r+2}\eta_2+\cdots+c_n\eta_{n-r} \tag{11}$$

也是方程组(5)的一个解. 比较(10)和(11)的最后 $n-r$ 个分量可知:这两个解的自由未知量有相同的值,因而这两个解完全一样,即

$$\eta=c_{r+1}\eta_1+c_{r+2}\eta_2+\cdots+c_n\eta_{n-r}.$$

这就是说,方程组(5)的任意一个解都可以表成 $\eta_1,\eta_2,\cdots,\eta_{n-r}$的线性组合. 综合以上两点,就证明了 $\eta_1,\eta_2,\cdots,\eta_{n-r}$确是方程组(5)的一个基础解系,也就是证明了齐次线性方程组的确有基础解系.

证明中给出的这个基础解系是由 $n-r$ 个解组成的. 至于方程组(5)其他的基础解系,由定义可知,都是与这个基础解系等价的. 同时它们也都是线性无关的,因此包含的向量个数也是相同的. 这就证明了方程组(5)的基础解系所包含的解的个数都等于 $n-r$. |

定理的证明给出了一个具体找基础解系的方法. 从证明中可以看出,$n-r$ 也就是自由未知量的个数,并且从基础解系的定义及齐次线性方程组解的性质可知齐次线性方程组(5)的全部解就是

$$k_1\eta_1+k_2\eta_2+\cdots+k_{n-r}\eta_{n-r}.$$

其中 $k_1,k_2,\cdots,k_{n-r}$是数域 **P** 中任意数.

例 4 求下列线性方程组的一个基础解系,并用基础解系表示出全部解.

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

解: 首先将系数矩阵化为阶梯形:

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

$$\to\begin{bmatrix}1&1&1&1&1\\0&1&2&3&6\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix}.$$

得同解方程组

$$\begin{cases}x_1+x_2+x_3+x_4+x_5=0,\\ \qquad x_2+2x_3+3x_4+6x_5=0,\\ \qquad\qquad x_5=0.\end{cases}$$

取 x_3、x_4 为自由未知量，将方程组改写成

$$\begin{cases}x_1+x_2+x_5=-x_3-x_4,\\ \qquad x_2+6x_5=-2x_3-3x_4,\\ \qquad\qquad x_5=0.\end{cases}$$

将 $x_3=1,x_4=0$ 代入，得

$$\begin{cases}x_1+x_2+x_5=-1,\\ \qquad x_2+6x_5=-2,\\ \qquad\qquad x_5=0.\end{cases}$$

于是得一个解：$(1,-2,1,0,0)$.

将 $x_3=0,x_4=1$ 代入，得

$$\begin{cases}x_1+x_2+x_5=-1,\\ \qquad x_2+6x_5=-3,\\ \qquad\qquad x_5=0.\end{cases}$$

于是得另一解：$(2,-3,0,1,0)$.

从而得到线性方程组的一个基础解系：$(1,-2,1,0,0)$，$(2,$

$-3,0,1,0)$. 而这个线性方程组的全部解为：

$$k_1\eta_1 + k_2\eta_2$$

其中 k_1,k_2 为数域 P 中任意数.

由于齐次线性方程组的常数项都是零，所以在上例中只是对系数矩阵进行了初等变换，而没有用增广矩阵. 最后列出同解的阶梯形方程组时，必须注意到这一点.

当齐次线性方程组有非零解时，基础解系有很多种取法. 这一点可以从下述例题看出：

例 5 证明与基础解系等价的线性无关向量组也是基础解系.

证明：设 $\eta_1,\eta_2,\cdots,\eta_t$是一个基础解系；$\alpha_1,\alpha_2,\cdots,\alpha_t$是一个与$\eta_1,\eta_2,\cdots,\eta_t$等价的线性无关向量组(因为等价的线性无关向量组所含向量的个数是相同的，所以这两组向量包含的向量个数相同).

由于解的线性组合还是解，而每个 α_i 都可以表成 $\eta_1,\cdots,\eta_t$ 的线性组合，所以每个 $\alpha_i(i=1,2,\cdots,t)$都是解.

如果 η 是任一个解，那么 η 可以由基础解系 $\eta_1,\eta_2,\cdots,\eta_t$ 线性表出. 根据等价性：$\eta_1,\eta_2,\cdots,\eta_t$ 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 线性表出，所以 η 也可以由 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 线性表出.

又由假设：$\alpha_1,\alpha_2,\cdots,\alpha_t$ 是线性无关的，所以 $\alpha_1,\alpha_1,\cdots,\alpha_t$ 也是一个基础解系.

习　题　2.7(2)

1. 求下列齐次线性方程组的一个基础解系，并用基础解系表出方程组的全部解：

$$(1)\begin{cases} x_1+2x_2+3x_3+3x_4+7x_5=0, \\ 3x_1+2x_2+x_3+x_4-3x_5=0, \\ x_2+2x_3+2x_4+6x_5=0, \\ 5x_1+4x_2+3x_3+3x_4-x_5=0; \end{cases}$$

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

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

(4) $x_1+x_2+x_3+x_4+x_5=0.$

2. 设下列齐次线性方程组的秩为 r，

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

求证：方程组的任意 $n-r$ 个线性无关的解都是它的一个基础解系.

下面来讨论一般线性方程组的解的结构. 如果把一般线性方程组

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

的常数项换成 0，就得到齐次线性方程组(5). 方程组(5)称为方程组(1)的**导出组**. 一般方程组的解与它的导出组的解之间有密切的关系：

1. 线性方程组(1)的两个解的差是它的导出组(5)的解.

证明：设$(k_1,k_2,\cdots,k_n)$，$(l_1,l_2,\cdots,l_n)$是方程组(1)的两个解，即

$$\sum_{j=1}^{n} a_{ij}k_j=b_i \quad (i=1,2,\cdots,s);$$

$$\sum_{j=1}^{n} a_{ij} l_j = b_i \quad (i = 1, 2, \cdots, s).$$

它们的差是 $$(k_1 - l_1, k_2 - l_2, \cdots, k_n - l_n).$$

将它代入方程组(5)的第 i 个方程的左边，得

$$\sum_{j=1}^{n} a_{ij}(k_j - l_j) = \sum_{j=1}^{n} a_{ij} k_j - \sum_{j=1}^{n} a_{ij} l_j = b_i - b_j = 0$$
$$(i = 1, 2, \cdots, s).$$

这就证明(1)的两个解的差$(k_1-l_1, k_2-l_2, \cdots, k_n-l_n)$是导出组(5)的解. |

2. 线性方程组(1)的一个解与它的导出组(5)的一个解的和还是方程组(1)的一个解.

证明：设$(k_1, k_2, \cdots, k_n)$是方程组(1)的一个解，即

$$\sum_{j=1}^{n} a_{ij} k_j = b_i \quad (i = 1, 2, \cdots, s);$$

再设$(l_1, l_2, \cdots, l_n)$是导出组(5)的一个解，即

$$\sum_{j=1}^{n} a_{ij} l_j = 0 \quad (i = 1, 2, \cdots, s).$$

于是 $$\sum_{j=1}^{n} a_{ij}(k_j + l_j) = \sum_{j=1}^{n} a_{ij} k_j + \sum_{j=1}^{n} a_{ij} l_j = b_i + 0 = b_i$$
$$(i = 1, 2, \cdots, s).$$

这就证明了$(k_1+l_1, k_2+l_2, \cdots, k_n+l_n)$即$(k_1, k_2, \cdots, k_n)+(l_1, l_2, \cdots, l_n)$是方程组(1)的一个解. |

根据上述两点很容易证明：

定理 14 如果 γ_0 是方程组(1)的一个解，那么方程组(1)的任一个解都可以表示成：

$$\gamma = \gamma_0 + \eta \tag{12}$$

的形式，其中 η 是导出组(5)的一个解.

证明：γ 与 γ_0 的差是导出组(5)的一个解.

令 $$\eta = \gamma - \gamma_0,$$

即得 $$\gamma = \gamma_0 + \eta. \ |$$

既然方程组(1)的任一个解都能表示成(12)的形式，而且形如(12)的向量当然都是方程组(1)的解. 那么当 η 取遍导出组(5)的全部解的时候，

$$\gamma = \gamma_0 + \eta$$

就取遍方程组(1)的全部解. 因此，要找出一个线性方程组的全部解，只要找出它的一个解及它的导出组的全部解就行了. 由于导出组是一个齐次线性方程组，所以它的解的全体可以用基础解系来表出. 于是，可以用导出组的基础解系来表出一般线性方程组的全部解：取定方程组(1)的一个解 γ_0，再找出导出组的一个基础解系 $\eta_1, \eta_2, \cdots, \eta_{n-r}$，那么方程组(1)的任一个解 γ 都可以表示成：

$$\gamma = \gamma_0 + k_1\eta_1 + k_2\eta_2 + \cdots + k_{n-r}\eta_{n-r}.$$

其中 γ_0 称为方程组(1)的一个**特解**(方程组(1)的任一个解都可取作特解). 为了方便，我们把线性方程组的导出组的基础解系就称为这个线性方程组的基础解系.

推论 在方程组(1)有解的前提下，解是唯一的充分必要条件是：它的导出组(5)只有零解.

证明：充分性：如果方程组(1)有两个不同的解，那么它们的差就是导出组的一个非零解. 因此，如果导出组只有零解，那么方程组(1)只有唯一解.

必要性：如果导出组有非零解，那么这个解与方程组(1)的一个解的和就是方程组(1)的另一个解，这说明方程组(1)不只一个解. 因此，如果方程组(1)有唯一解，那么它的导出组只有零解. ▏

例 6 求方程组

$$\begin{cases} x_1 - 2x_2 + x_3 - x_4 + x_5 = 1, \\ 2x_1 + x_2 - x_3 + 2x_4 - 3x_5 = 2, \\ 3x_1 - 2x_2 - x_3 + x_4 - 2x_5 = 2, \\ 2x_1 - 5x_2 + x_3 - 2x_4 + 2x_5 = 1 \end{cases}$$

的全部解.

解：用初等变换把增广矩阵化为阶梯形：

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

$$\to\begin{bmatrix}1&-2&1&-1&1&1\\0&1&1&0&0&1\\0&0&-8&4&-5&-5\\0&0&-8&4&-5&-5\end{bmatrix}\to\begin{bmatrix}1&-2&1&-1&1&1\\0&1&1&0&0&1\\0&0&8&-4&5&5\\0&0&0&0&0&0\end{bmatrix}.$$

得同解方程组

$$\begin{cases}x_1-2x_2+x_3-x_4+x_5=1,\\ \qquad x_2+x_3=1,\\ \qquad\qquad 8x_3-4x_4+5x_5=5.\end{cases}$$

取 x_4,x_5 作为自由未知量，将方程改写成：

$$\begin{cases}x_1-2x_2+x_3=1+x_4-x_5,\\ \qquad x_2+x_3=1,\\ \qquad\qquad 8x_3=5+4x_4-5x_5.\end{cases}$$

求得方程组的一般解为：

$$\begin{cases}x_1=\dfrac{9}{8}-\dfrac{1}{2}x_4+\dfrac{7}{8}x_5,\\[2mm] x_2=\dfrac{3}{8}-\dfrac{1}{2}x_4+\dfrac{5}{8}x_5,\\[2mm] x_3=\dfrac{5}{8}+\dfrac{1}{2}x_4+\dfrac{5}{8}x_5.\end{cases}$$

其中 x_1,x_2,x_3 为自由未知量

将 $x_4=x_5=0$ 代入，得得到一个特解：

$$\left(\frac{9}{8},\frac{3}{8},\frac{5}{8},0,0\right).$$

下面求其导出组的基础解系. 因为导出组是由原方程组把常

数项改成 0 而得到的，而它的系数矩阵和原方程组仍是一样的，因此，用和上面同样的初等变换可以得到与导出组同解的方程组

$$\begin{cases} x_1 - 2x_2 + x_3 - x_4 + x_5 = 0, \\ x_2 + x_3 = 0, \\ 8x_3 - 4x_4 + 5x_5 = 0. \end{cases}$$

仍取 x_4, x_5 为自由未知量，得一般解为：

$$\begin{cases} x_1 = -\dfrac{1}{2}x_4 + \dfrac{7}{8}x_5, \\ x_2 = -\dfrac{1}{2}x_4 + \dfrac{5}{8}x_5, \\ x_3 = \dfrac{1}{2}x_4 - \dfrac{5}{8}x_5. \end{cases}$$

其中 x_4, x_5 为自由变量.

分别用(1,0)，(0,1)代替自由未知量(x_4, x_5)，得到一个基础解系：

$$\left(-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, 1, 0\right), \left(\frac{7}{8}, \frac{5}{8}, -\frac{5}{8}, 0, 1\right).$$

所以，原方程组的全部解为：

$$\left(\frac{9}{8}, \frac{3}{8}, \frac{5}{8}, 0, 0\right) + k_1\left(-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, 1, 0\right) + k_2\left(\frac{7}{8}, \frac{5}{8}, -\frac{5}{8}, 0, 1\right).$$

其中 k_1, k_2 是数域 P 中任意数.

习　题　2.7(3)

1. 用基础解系表出下列方程组的全部解：

(1) $\begin{cases} 2x_1 + x_2 - x_3 + x_4 = 1, \\ x_1 + 2x_2 + x_3 - x_4 = 2, \\ x_1 + x_2 + 2x_3 + x_4 = 3; \end{cases}$　(2) $\begin{cases} x_1 + x_2 - 3x_4 - x_5 = 2, \\ x_1 - x_2 + 2x_3 - x_4 = 1, \\ 4x_1 - 2x_2 + 6x_3 + 3x_4 - 4x_5 = 8, \\ 2x_1 + 4x_2 - 2x_3 + 4x_4 - 7x_5 = 9; \end{cases}$

(3) $\begin{cases} x_1+2x_2+x_3-x_4 = 6, \\ 2x_1-x_2+x_3+3x_4+4x_5=-7, \\ 2x_1-x_2+2x_3+x_4-2x_5=-4, \\ 2x_1-3x_2+x_3+2x_4-2x_5=-9, \\ x_1+x_3-2x_4-6x_5=4; \end{cases}$

(4) $\begin{cases} 3x_1+2x_2-x_3+2x_4 = 1, \\ 4x_1+x_2-3x_4 = 2, \\ 2x_1-x_2-2x_3+x_4+x_5=-3, \\ 3x_1+x_2+3x_3-9x_4-x_5=6, \\ 3x_1-x_2-5x_3+7x_4+2x_5=-7. \end{cases}$

2. 线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n-1,1}x_1+a_{n-1,2}x_2+\cdots+a_{n-1,n}x_n=0 \end{cases}$$

的系数矩阵为

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \end{bmatrix}.$$

设 $M_j(j=1,2,\cdots,n)$是在矩阵 A 中划去第 j 列所得到的 $n-1$ 阶子式. 试证:

(1) $(M_1,-M_2,\cdots(-1)^{n-1}M_n)$是方程组的一个解;

(2) 如果 A 的秩为 $n-1$,那么方程组的解全是$(M_1,-M_2,\cdots,(-1)^{n-1}M_n)$的倍数.

3. 假设 $\eta_1,\eta_2,\cdots,\eta_t$ 是某个线性方程组的解,且常数 $u_1,u_2,\cdots,u_t$ 的和等于 1,求证: $u_1\eta_1+u_2\eta_2+\cdots+u_t\eta_t$ 也是这个方程组的一个解.

内 容 提 要

1. n 维向量空间

(1) n 维向量空间

1) n 维向量及其运算；

2) 运算的基本规律.

(2) 线性相关性

基本概念：线性组合、线性表出、线性相关、线性无关、等价、极大线性无关组、向量组的秩.

主要结论：

1) 设

$$\begin{aligned}\boldsymbol{\alpha}_1 &= (a_{11}, a_{12}, \cdots, a_{1n}),\\ \boldsymbol{\alpha}_2 &= (a_{21}, a_{22}, \cdots, a_{2n}),\\ &\cdots\cdots\cdots\cdots\\ \boldsymbol{\alpha}_s &= (a_{s1}, a_{s2}, \cdots, a_{sn}),\\ \boldsymbol{\beta} &= (b_1, b_2, \cdots, b_n).\end{aligned}$$

$\boldsymbol{\beta}$ 可以被 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 线性表出的充分必要条件是下述线性方程组有解：

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \cdots + a_{s1}x_s = b_1,\\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{s2}x_s = b_2,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{sn}x_s = b_n. \end{cases}$$

2) 设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 如上. 向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 线性相关的充分必要条件是下述齐次线性方程组有非零解：

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \cdots + a_{s1}x_s = 0,\\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{s2}x_s = 0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{sn}x_s = 0. \end{cases}$$

特别地，当 $s=n$ 时，$\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性相关的充分必要条件是：

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

3) 如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关，而 $\alpha_1,\alpha_2,\cdots,\alpha_s,\beta$ 线性相关，那么 β 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出，并且表法唯一.

4) 一个向量组线性相关的充分必要条件是：其中有一个向量可以由其他的向量线性表出.

5) 如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 可由 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表出，而且 $r>s$，那么 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 一定线性相关.

6) 多于 n 个的 n 维向量一定线性相关.

7) 等价的线性无关的向量组所含的向量个数是相同的.

8) 向量组的极大线性无关组与原向量组是等价的.

9) 向量组的任一线性无关的部分组都可以扩充成一个极大线性无关组.

10) 等价的向量组有相同的秩.

2. 矩阵的秩

(1) 矩阵的概念

(1) 矩阵的秩的概念

矩阵的秩＝行秩＝列秩＝行列式秩

(3) 矩阵的秩的求法

1) 用初等行变换将矩阵化为阶梯形；

2) 找出阶数最大的非零子式.

3. 克莱姆法则

(1) 如果线性方程组

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

的系数行列式 $D\neq0$，那么这个方程组有解，并且解是唯一的，表示成

$$x_i = \frac{D_i}{D} \quad (i = 1,2,\cdots,n).$$

其中 D 是系数行列式；D_i 是把 D 中第 i 列（即 x_i 的系数）的元素换成常数 $b_i(i=1,2,\cdots,n)$ 所构成的行列式.

(2) 将克莱姆法则应用于齐次线性方程组

如果齐次线性方程组

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

的系数行列式不等于零，那么这个方程组只有零解.

4. 线性方程组有解判别定理

线性方程组

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

有解的充分必要条件是：它的系数矩阵

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

与增广矩阵

$$\overline{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} & b_s \end{bmatrix}$$

有相同的秩,即 $\mathrm{r}(A)=\mathrm{r}(\overline{A})$.

5. 线性方程组解的个数

(1) 齐次线性方程组

1) 如果齐次线性方程组

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

的系数矩阵的 $\mathrm{r}(A)=n$,那么方程组(2)只有零解.

2) 如果 $\mathrm{r}(A)=r<n$,那么方程组(2)有非零解.此时方程组有 $n-r$ 个自由未知量,有无穷多解.

3) 特别地,当 $s=n$ 时,齐次方程组有非零解的充分必要条件是:系数行列式等于零.

(2) 一般线性方程组

1) 如果 $\mathrm{r}(A)=\mathrm{r}(\overline{A})=n$,则线性方程组(1)有唯一解.

2) 如果 $\mathrm{r}(A)=\mathrm{r}(\overline{A})=r<n$,则线性方程组(1)有 $n-r$ 个自由未知量,有无穷多解.

6. 线性方程组解的结构

(1) 齐次线性方程组

解的性质:

1) 如果 α,β 是齐次线性方程组的解,那么 $\alpha+\beta$ 也是这个方程组的解.

2) 如果 α 是齐次方程组的解,那么 $k\alpha$ 也是这个方程组的解.

归纳 1)和 2)知:齐次线性方程组的解的线性组合还是这个方程组的解.

解的结构:

1) 基础解系的定义及存在性.

2) 如果 $r(A)=r$,那么齐次线性方程组(2)的基础解系包含 $n-r$ 个解.

3) 如果 $\eta_1,\eta_2,\cdots,\eta_{n-r}$是齐次线性方程组(2)的一个基础解系,那么方程组(2)的全部解就是

$$k_1\eta_1+k_2\eta_2+\cdots+k_{n-r}\eta_{n-r},$$

其中 $k_1,k_2,\cdots,k_{n-r}$可以是指定数域中的任意数.

(2) 一般线性方程组

齐次线性方程组(2)称为线性方程组(1)的导出组.

解的性质:

1) 如果 α,β 都是方程组(1)的解,那么 $\alpha-\beta$ 是导出组(2)的解.

2) 如果 α 是方程组(1)的解,γ 是导出组(2)的解,那么 $\alpha+\gamma$ 是方程组(1)的解.

解的结构:

如果 γ_0 是方程组(1)的一个解(称为特解),$\eta_1,\eta_2,\cdots,\eta_{n-r}$是一个基础解系(即导出组(2)的基础解系).那么方程组(1)的全部解就是

$$\gamma_0+k_1\eta_1+k_2\eta_2+\cdots+k_{n-r}\eta_{n-r},$$

其中 $k_1,k_2,\cdots,k_{n-r}$是指定数域中的任意数.

7. 线性方程组解的求法

(1) 将线性方程组用初等变换化为阶梯形方程组.

(2) 找出 A 的阶数最大的子式,应用克莱姆法则求解.

复习题2

1. 用基础解系表示出下列线性方程组的全部解：

(1) $\begin{cases} x_1+2x_2+x_3-3x_4+2x_5=1, \\ 2x_1+x_2+x_3+x_4-3x_5=6, \\ x_1+x_2+2x_3+2x_4-2x_5=2, \\ 2x_1+3x_2-5x_3-17x_4+10x_5=5; \end{cases}$

(2) $\begin{cases} x_1+2x_2-x_3+3x_4-x_5+2x_6=-5, \\ 2x_1-x_2+3x_3-4x_4+x_5-x_6=14, \\ 3x_1+x_2-x_3+2x_4+x_5+3x_6=-11, \\ 4x_1-7x_2+8x_3-15x_4+6x_5-5x_6=32, \\ 5x_1+5x_2-6x_3+11x_4+9x_6=-41; \end{cases}$

(3) $\begin{cases} x_1-5x_2+2x_3-3x_4=11, \\ -3x_1+x_2-4x_3+2x_4=-5, \\ -x_1-9x_2-4x_4=17, \\ 5x_1+3x_2+6x_3-x_4=-1; \end{cases}$

(4) $x_1+2x_2+3x_3+4x_4+5x_5=1$.

2. 讨论 a,b 取什么值时下列线性方程组有解，并在有解时求出全部解：

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

(2) $\begin{cases} x_1+2x_2+x_3+x_4+x_5=1, \\ 3x_1+4x_2+x_3+x_4+3x_5=a, \\ x_1-x_3-x_4+x_5=-2, \\ 5x_1+4x_2-x_3-x_4+5x_5=b. \end{cases}$

3. 假设向量 β 可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出. 证明：表法是唯一的充分必要条件为 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关.

4. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是一组 n 维向量，已知 n 维基本向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 都可由它们线性表出. 求证：$\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关.

5. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是一组 n 维向量. 试证：$\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关的充分必要条件是：任一 n 维向量都可由它们线性表出.

6. 试证:线性方程组

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

对任何 $b_1, b_2, \cdots, b_n$ 都有解的充分必要条件是:系数行列式

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

7. 设

$$\begin{aligned} \boldsymbol{\beta}_1 &= \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3 + \cdots + \boldsymbol{\alpha}_t \\ \boldsymbol{\beta}_2 &= \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_3 + \cdots + \boldsymbol{\alpha}_t \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \boldsymbol{\beta}_t &= \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \cdots + \boldsymbol{\alpha}_{t-1} \qquad (t > 1). \end{aligned}$$

试证:$\mathrm{r}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_t\} = \mathrm{r}\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_t\}$.

8. 证明:若向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_t$ 与向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_t, \boldsymbol{\beta}$ 有相同的秩,则向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_t$ 的任一个极大线性无关组也是向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_t, \boldsymbol{\beta}$ 的一个极大线性无关组.

9. 证明:若向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_t$ 与向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_t, \boldsymbol{\beta}$ 有相同的秩,则 $\boldsymbol{\beta}$ 可由向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_t$ 线性表出.

10. 已知两向量组有相同的秩,并且其中之一可以由另一组线性表出,试证明:这两个向量组等价.

第 3 章　矩　　阵

在第 2 章讨论线性方程组的时候，曾经引入了矩阵的概念. 从那里已经知道，不仅线性方程组可以用矩阵来表示，而且线性方程组的一些重要性质也可以通过它的系数矩阵和增广矩阵的性质来反映，甚至解方程组的过程也是通过变换矩阵来进行的. 以后还可以看到：除了线性方程组以外，还有许多问题不但可以用矩阵来表现，而且还可以利用矩阵来研究和解决这些问题；有些性质完全不同的、表面上毫无联系的问题，归结成矩阵问题以后，却可能是相同的了. 这就使矩阵成为数学中一个极其重要而且应用广泛的工具，因而矩阵就成为代数特别是线性代数的一个主要研究对象.

这一章介绍矩阵的运算，并讨论矩阵运算的一些基本性质. 这些性质在以后各章中都要用到.

本章所讨论的矩阵都是某个取定的数域 P 上的矩阵，所提到的数也都是同一个数域 P 中的数.

3.1　矩阵的运算

首先来定义矩阵的运算，即矩阵的加法、乘法、矩阵与数的乘法以及矩阵的转置.

1. 矩阵的加法

定义 1　设

$$A=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots}\\ a_{s1} & a_{s2} & \cdots & a_{sn}\end{bmatrix},$$

$$B=\begin{bmatrix}b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots}\\ b_{s1} & b_{s2} & \cdots & b_{sn}\end{bmatrix}$$

是两个 $s\times n$ 矩阵，则 $s\times n$ 矩阵

$$C=\begin{bmatrix}a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n}\\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n}\\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots}\\ a_{s1}+b_{s1} & a_{s2}+b_{s2} & \cdots & a_{sn}+b_{sn}\end{bmatrix}$$

称为 A 与 B 的**和**，记作

$$C=A+B.$$

从定义可以看出：两个矩阵必须在行数与列数分别相同的情况下才能相加.

例 1 $\begin{bmatrix}1 & 2 & 3 & 1\\ 1 & -1 & 2 & 0\\ 3 & 1 & 2 & -2\end{bmatrix}+\begin{bmatrix}0 & 1 & 2 & 3\\ -1 & 2 & 4 & 1\\ -3 & 2 & 0 & 1\end{bmatrix}$

$$=\begin{bmatrix}1+0 & 2+1 & 3+2 & 1+3\\ 1+(-1) & (-1)+2 & 2+4 & 0+1\\ 3+(-3) & 1+2 & 2+0 & (-2)+1\end{bmatrix}$$

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

例 2 $$\begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{bmatrix}+\begin{bmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{bmatrix}=\begin{bmatrix}a_1+b_1\\ a_2+b_2\\ \vdots\\ a_n+b_n\end{bmatrix}.$$

由于矩阵的加法就是把矩阵对应的元素相加，因此，矩阵的加法满足

交换律：$A+B=B+A$.

结合律：$A+(B+C)=(A+B)+C$.

元素都是零的矩阵称为**零矩阵**，记为 $O_{s\times n}$. 在不致于混淆的情况下，可以简单地记为 O. 显然，对任意矩阵 A，都有

$$A+O=A.$$

当然，这里的 O 是表示与 A 的行数与列数都相同的那个零矩阵.

矩阵

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

称为矩阵 A 的**负矩阵**，记作 $-A$. 显然有

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

两个行数与列数相同的矩阵可以相减. 设 A,B 如上述，那么

$$A-B=\begin{bmatrix} a_{11}-b_{11} & a_{12}-b_{12} & \cdots & a_{1n}-b_{1n} \\ a_{21}-b_{21} & a_{22}-b_{22} & \cdots & a_{2n}-b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1}-b_{s1} & a_{s2}-b_{s2} & \cdots & a_{sn}-b_{sn} \end{bmatrix}.$$

矩阵的减法可以用负矩阵表示为

$$A-B=A+(-B).$$

于是,矩阵方程 $X+A=B$ 总有唯一解 $X=B-A$.

例 3 求矩阵 X,使

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

解:

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

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

习　题　3.1(1)

1. 计算:

(1) $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{bmatrix}+\begin{bmatrix} 3 & 1 & -2 \\ 3 & -2 & 1 \\ -3 & 1 & -1 \end{bmatrix}$;

(2) $\begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 3 & 1 & 4 \\ 2 & 0 & 3 & 1 \end{bmatrix}+\begin{bmatrix} 3 & 2 & 1 & 0 \\ 2 & -1 & -1 & 1 \\ 0 & -1 & 3 & 2 \end{bmatrix}-\begin{bmatrix} -1 & 2 & 3 & 4 \\ 0 & 2 & 0 & -1 \\ -1 & 1 & 3 & 1 \end{bmatrix}$.

2. 求 X，使

$$\begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix} + X - \begin{bmatrix} 2 & 3 & 0 \\ -1 & 0 & -1 \\ 2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -3 & -1 & 2 \end{bmatrix}.$$

3. 求 X，使

$$\begin{bmatrix} a & b & c \\ 2b & 3c-1 & a-b \\ 1 & 2+a & a-c \end{bmatrix} + X = \begin{bmatrix} 1-a & b+c & c-3 \\ 2a & 3b+c & 3 \\ 2 & -1 & c \end{bmatrix}.$$

2. 矩阵的乘法

在叙述矩阵乘法的定义以前，先看一个引出矩阵乘法的例子.

设 $x_1, x_2, x_3; y_1, y_2, y_3; z_1, z_2, z_3$ 是三组变量，且 x_1, x_2, x_3 与 y_1, y_2, y_3 之间的关系为

$$\begin{cases} x_1 = a_{11}y_1 + a_{12}y_2 + a_{13}y_3, \\ x_2 = a_{21}y_1 + a_{22}y_2 + a_{23}y_3, \\ x_3 = a_{31}y_1 + a_{32}y_2 + a_{33}y_3, \end{cases} \tag{1}$$

显然，这个关系是由系数 a_{ij} 完全确定的，把系数写成一个矩阵

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix};$$

再设 $y_1, y_2, y_3; z_1, z_2, z_3$ 之间的关系为：

$$\begin{cases} y_1 = b_{11}z_1 + b_{12}z_2 + b_{13}z_3, \\ y_2 = b_{21}z_1 + b_{22}z_2 + b_{23}z_3, \\ y_3 = b_{31}z_1 + b_{32}z_2 + b_{33}z_3. \end{cases} \tag{2}$$

把系数写成一个矩阵

$$B=\begin{bmatrix}b_{11} & b_{12} & b_{13}\\ b_{21} & b_{22} & b_{23}\\ b_{31} & b_{32} & b_{33}\end{bmatrix}.$$

最后，设 x_1,x_2,x_3 与 z_1,z_2,z_3 之间的关系为

$$\begin{cases}x_1=c_{11}z_1+c_{12}z_2+c_{13}z_3,\\ x_2=c_{21}z_1+c_{22}z_2+c_{23}z_3,\\ x_3=c_{31}z_1+c_{32}z_2+c_{33}z_3.\end{cases}\tag{3}$$

其系数所成的矩阵为

$$C=\begin{bmatrix}c_{11} & c_{12} & c_{13}\\ c_{21} & c_{22} & c_{23}\\ c_{31} & c_{32} & c_{33}\end{bmatrix}.$$

但是 x_1,x_2,x_3 与 z_1,z_2,z_3 之间的关系由(1)和(2)所决定，将(2)代入(1)，得

$$\begin{aligned}x_i&=\sum_{k=1}^{3}a_{ik}y_k=\sum_{k=1}^{3}a_{ik}\Big(\sum_{y=1}^{3}b_{kj}z_j\Big)=\sum_{k=1}^{3}\sum_{j=1}^{3}a_{ik}b_{kj}z_j\\ &=\sum_{j=1}^{3}\sum_{k=1}^{3}a_{ik}b_{kj}z_j=\sum_{j=1}^{3}\Big(\sum_{k=1}^{3}a_{ik}b_{kj}\Big)z_j\quad(i=1,2,3).\end{aligned}$$

与(3)式比较，即得

$$c_{ij}=\sum_{k=1}^{3}a_{ik}b_{kj}\quad(i,j=1,2,3).$$

满足这个等式的矩阵 C 就称做矩阵 A 与 B 的乘积.

一般地，定义矩阵的乘法如下：

定义 2 设 A 是一个 $s\times n$ 矩阵：

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

B 是一个 $n\times m$ 矩阵：

$$B=\begin{bmatrix}b_{11} & b_{12} & \cdots & b_{1m}\\ b_{21} & b_{22} & \cdots & b_{2m}\\ \cdots & \cdots & \cdots & \cdots\\ b_{n1} & b_{n2} & \cdots & b_{nm}\end{bmatrix}.$$

作 $s\times m$ 矩阵

$$C=\begin{bmatrix}c_{11} & c_{12} & \cdots & c_{1m}\\ c_{21} & c_{22} & \cdots & c_{2m}\\ \cdots & \cdots & \cdots & \cdots\\ c_{s1} & c_{s2} & \cdots & c_{sm}\end{bmatrix},$$

其中

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}=\sum_{k=1}^{n}a_{ik}b_{kj}$$
$$(i=1,2,\cdots,s;\ j=1,2,\cdots,m). \tag{4}$$

矩阵 C 称为矩阵 A 与 B 的**乘积**，记为

$$C=AB.$$

在矩阵乘积的定义中，要求第一个矩阵的列数必须等于第二个矩阵的行数．公式(4)说明乘积 C 的第 i 行第 j 列的元素等于 A 的第 i 行与 B 的第 j 列的对应元素的乘积的和．

例 4 设

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

那么

$$AB=\begin{bmatrix}1\times1+0\times2+2\times0+(-1)\times1 & 1\times2+0\times1+2\times3+(-1)\times4\\0\times1+1\times2+(-1)\times0+3\times1 & 0\times2+1\times1+(-1)\times3+3\times\\(-1)\times1+2\times2+0\times0+1\times1 & (-1)\times2+2\times1+0\times3+1\times\end{bmatrix}$$

$$=\begin{bmatrix}0&4\\5&10\\4&4\end{bmatrix}.$$

矩阵的乘法与数的乘法有一个极不相同的地方，就是矩阵的乘法不满足交换律，也就是说：矩阵的乘积 AB 与 BA 不一定相等. 看下面的例子.

例 5 设

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

那么

$$AB=\begin{bmatrix}5&5&-5\\5&0&-5\\4&4&-6\end{bmatrix};$$

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

例 6 设

$$A=\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix},\quad B=[b_1\quad b_2\quad\cdots\quad b_n],$$

那么
$$AB=\begin{bmatrix}a_1b_1 & a_1b_2 & \cdots & a_1b_n\\a_2b_1 & a_2b_2 & \cdots & a_2b_n\\\cdots & \cdots & \cdots & \cdots\\a_nb_1 & a_nb_2 & \cdots & a_nb_n\end{bmatrix};$$

$$BA=\left[\sum_{i=1}^{n}a_ib_i\right].$$

例 7 给了线性方程组

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

它的系数矩阵为

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

再令
$$X=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix};\quad B=\begin{bmatrix}b_1\\b_2\\\vdots\\b_s\end{bmatrix}.$$

那么线性方程组就可以写成一个矩阵方程：

$$AX=B.$$

由于矩阵的乘法不满足交换律，所以作矩阵乘法时必须注意．首先，当 AB 有意义时，BA 却不一定有意义．例 4 就是这种情形；其次，即使 AB 与 BA 都有意义，它们的级数也不一定相等．例 6 就是这种情形；最后，即使 A,B 都是 $n\times n$ 矩阵，AB 与 BA 都有意义，而且都是 $n\times n$ 矩阵，但是它们也不一定相等．如例 5 中 A 与 B 都是 3×3 矩阵，但是 $AB\neq BA$．矩阵的乘法还有一个特点：两个不等于零的矩阵之积可以是零矩阵．

例 8 $\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

因此，在讨论矩阵的时候必须注意：从 $AB=O$，不能推出 $A=O$ 或 $B=O$.

例 9 如果 $AB=BA$，就说矩阵 A 与 B **可交换**．设

$$A=\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

求所有与 A 可交换的矩阵．

解：设

$$X=\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}.$$

那么

$$AX=\begin{bmatrix} x_1+x_2 & y_1+y_2 & z_1+z_2 \\ x_2+x_3 & y_2+y_3 & z_2+z_3 \\ x_3 & y_3 & z_3 \end{bmatrix};$$

$$XA=\begin{bmatrix} x_1 & x_1+y_1 & y_1+z_1 \\ x_2 & x_2+y_2 & y_2+z_2 \\ x_3 & x_3+y_3 & y_3+z_3 \end{bmatrix}.$$

因此,X 与 A 可交换的充分必要条件是

$$\begin{cases} x_1+x_2=x_1, \\ y_1+y_2=x_1+y_1, \\ z_1+z_2=y_1+z_1, \\ x_2+x_3=x_2, \\ y_2+y_3=x_2+y_2, \\ z_2+z_3=y_2+z_2, \\ y_3=x_3+y_3, \\ z_3=y_3+z_3. \end{cases}$$

而这一组条件又等价于

$$\begin{cases} x_2=x_3=y_3=0, \\ x_1=y_2=z_3, \\ y_1=z_2. \end{cases}$$

所以与 A 可交换的全部矩阵都可表成:

$$X=\begin{bmatrix} x_1 & y_1 & z_1 \\ 0 & x_1 & y_1 \\ 0 & 0 & x_1 \end{bmatrix},$$

其中 x_1,y_1,z_1 可以取数域 $\mathbf{P}$ 中任意的数.

上面通过例子指出矩阵的乘法有一些与数的乘法不同的地方.这是需要经常注意的.但是矩阵的乘法也有许多与数的乘法相类似的地方,即矩阵的乘法满足以下一些规律.这些规律可以简化矩阵的运算.

(1) 矩阵乘法满足结合律，即

$$A(BC) = (AB)C.$$

当然，这里的 AB 与 BC 都被认为在乘法运算中是有意义的. 这一点以后就不再说明了.

证明：设

$$A = [a_{ij}]_{sn}; \quad B = [b_{jk}]_{nm}; \quad C = [c_{kl}]_{mt}.$$

令
$$U = BC = [u_{jl}]_{nt};$$
$$V = AB = [v_{ik}]_{sm}.$$

根据乘法公式，知

$$u_{jl} = \sum_{k=1}^{m} b_{jk}c_{kl} \quad (j = 1,2,\cdots,n;\ l = 1,2,\cdots,t);$$

$$v_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} \quad (i = 1,2,\cdots,s;\ k = 1,2,\cdots,m).$$

于是 $A(BC)=AU$ 的第 i 行第 l 列处的元素为：

$$\sum_{j=1}^{n} a_{ij}u_{jl} = \sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{m} b_{jk}c_{kl}\right) = \sum_{j=1}^{n}\sum_{k=1}^{m} a_{ij}b_{jk}c_{kl};$$

而 $(AB)C=VC$ 的第 i 行第 l 列处的元素为：

$$\sum_{k=1}^{m} v_{ik}c_{kl} = \sum_{k=1}^{m}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)c_{kl} = \sum_{k=1}^{m}\sum_{j=1}^{n} a_{ij}b_{jk}c_{kl}.$$

由于双重连加号可以交换次序，所以

$$\sum_{j=1}^{n}\sum_{k=1}^{m} a_{ij}b_{jk}c_{kl} = \sum_{k=1}^{m}\sum_{j=1}^{n} a_{ij}b_{jk}c_{kl}.$$

这就证明了 $A(BC)=(AB)C$.

(2) 矩阵中有一类特殊的矩阵，起着与数的乘法中 1 相同的作用，即所谓单位矩阵. 主对角线上的元素全是 1，其余的元素全是 0 的 n 阶方阵

$$\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots}\\ 0 & 0 & \cdots & 1\end{bmatrix}$$

称为 **n 阶单位矩阵**，记作 E_n. 在不致混淆的情况下，可以简单地写成 E. 很容易检验

$$A_{sn}E_n = A_{sn};$$

$$E_sA_{sn} = A_{sn}.$$

(3) 矩阵的乘法和加法满足分配律，即

$$A(B+C) = AB + AC;$$

$$(B+C)A = BA + CA.$$

由于矩阵的乘法不满足交换律，所以这是两条不同的规律. 证明留给读者.

因为矩阵的乘法满足结合律，所以可以定义矩阵的方幂. 设 A 是一个 n 阶矩阵. 用 $A^k(k>0)$表示 k 个 A 的连乘积，称为 **A 的 k 次方幂**. 规定 $A^0=E$，容易看出

$$A^k \cdot A^l = A^{k+l},$$

$$(A^k)^l = A^{kl},$$

$$(k, l \geqslant 0).$$

需要注意的是：由于矩阵的乘法不满足交换律，所以等式

$$(AB)^k = A^kB^k$$

一般不成立.

习　题　3.1(2)

1. 求下列矩阵的乘积：

(1) $\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}1 & -1\\1 & 2\end{bmatrix}$；

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

(3) $\begin{bmatrix}a & b & c\\c & a & b\\1 & 1 & 1\end{bmatrix}\begin{bmatrix}a & c & 1\\b & a & 1\\c & b & 1\end{bmatrix}$；

(4) $\begin{bmatrix}\mathrm{i} & 3 & \frac{1}{2}\\0 & 1+\mathrm{i} & 1\\-1 & 0 & 1\end{bmatrix}\begin{bmatrix}\mathrm{i} & 1-\mathrm{i} & -1\\1 & 4 & \mathrm{i}\\0 & 1+\mathrm{i} & 2\end{bmatrix}$.

2. 计算

$$\begin{bmatrix}1 & 2 & 3\\-1 & 2 & 1\\1 & -3 & 2\end{bmatrix}\begin{bmatrix}1 & 2 & 4\\2 & -4 & 1\\-1 & 1 & 0\end{bmatrix}+\begin{bmatrix}2 & 4 & 5\\5 & 1 & -1\\3 & -2 & 7\end{bmatrix}.$$

3. 设

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

求 $AB,(AB)C,BC,A(BC)$.

4. (1) 举出两个 2 阶矩阵 A 与 B，使 $(AB)^2=A^2B^2$；

(2) 举出两个 2 阶矩阵 A 与 B，使 $(AB)^2\neq A^2B^2$.

5. 计算

(1) $\begin{bmatrix}2 & 1 & 1\\3 & 1 & 0\\0 & 1 & 2\end{bmatrix}^2$；　　(2) $\begin{bmatrix}1 & 2\\-2 & 1\end{bmatrix}^5$；

(2) $\begin{bmatrix}1 & 1 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}^n$ $(n>0)$;　　(4) $\begin{bmatrix}1 & -1 & -1 & -1\\-1 & 1 & -1 & -1\\-1 & -1 & 1 & -1\\-1 & -1 & -1 & 1\end{bmatrix}^n$ $(n>0)$;

6. 求与下列矩阵 A 可交换的矩阵：

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

(2) $A=\begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\\1 & 0 & 0\end{bmatrix}$;　　(3) $A=\begin{bmatrix}1 & 0 & 0\\0 & 2 & 0\\0 & 0 & 3\end{bmatrix}$.

7. (1) 试证：如果 B_1,B_2 都与 A 可交换，那么 B_1+B_2,B_1B_2 也与 A 可交换；

(2) 试证：如果 B 与 A 可交换，那么 B 的 $k(k\geqslant 0)$次幂也与 A 可交换.

3. 矩阵与数的乘法

定义 3 设 A 是一个 $s\times n$ 矩阵：$A=[a_{ij}]_{sn}$；k 是一个数，矩阵

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

称为 A 与 k 的**数量乘积**，记作 kA.

也就是说，用数 k 乘矩阵 A，就是把 A 的每个元素都乘上 k. 根据定义可以直接验证数量乘积满足下列规律：

$$(k+l)A=kA+lA;$$

$$k(A+B)=kA+kB;$$

$$k(lA)=(kl)A;$$

$$1 \cdot A = A;$$
$$k(AB) = (kA)B = A(kB).$$

矩阵
$$kE = \begin{bmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & k \end{bmatrix}$$
称为**数量矩阵**. 对于 $s \times n$ 矩阵 A，有
$$kA = (kE_s)A = A(kE_n).$$
特别地，如果 A 是一个 $n \times n$ 矩阵；kE 是一个 n 阶数量矩阵，那么
$$(kE)A = A(kE).$$
这个式子说明：n 阶数量矩阵与所有的 $n \times n$ 矩阵作乘法是可交换的.

另外还有：
$$kE + lE = (k + l)E;$$
$$(kE) \cdot (lE) = (kl)E.$$
这就是说，数量矩阵的加法与乘法完全可以归结为数的加法与乘法.

4. 矩阵的转置

把一个矩阵 A 的行列互换，所得到的矩阵称为这个矩阵的转置，即

定义 4 设 A 是一个 $s \times n$ 矩阵
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} \end{bmatrix},$$

那么，$n\times s$ 矩阵

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

称为 A 的**转置矩阵**，简称 A 的转置，记作 A^{T} 或 A'.

矩阵的转置满足以下规律：

$$(A^{\mathrm{T}})^{\mathrm{T}} = A;$$

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

$$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}};$$

$$(kA)^{\mathrm{T}} = kA^{\mathrm{T}}.$$

下面只证明第三个等式，其余的留给读者验证.

设

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix},$$

首先 $(AB)^{\mathrm{T}}$ 与 $B^{\mathrm{T}}A^{\mathrm{T}}$ 都是 $m\times s$ 矩阵，而且 AB 中第 i 行第 j 列的元素为：

$$\sum_{k=1}^{n} a_{ik}b_{kj}.$$

所以 $(AB)^{\mathrm{T}}$ 中第 i 行第 j 列的元素是：

$$\sum_{k=1}^{n} a_{jk}b_{ki}.$$

B^{T} 的第 i 行第 k 列的元素是 b_{ki}；A^{T} 的第 k 行第 j 列的元素是 a_{jk}，因此，$B^{\mathrm{T}}A^{\mathrm{T}}$ 中第 i 行第 j 列的元素是：

$$\sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n} a_{jk}b_{ki}.$$

这就证明了$(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}$.

例 10 设

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

那么

$$AB=\begin{bmatrix}5 & 6\\ 10 & 7\\ 0 & -5\end{bmatrix},$$

$$A^{\mathrm{T}}=\begin{bmatrix}1 & 1 & -1\\ -1 & 0 & 2\\ 2 & 3 & -1\end{bmatrix},\quad B^{\mathrm{T}}=\begin{bmatrix}1 & 2 & 3\\ 1 & -1 & 2\end{bmatrix},$$

$$B^{\mathrm{T}}A^{\mathrm{T}}=\begin{bmatrix}5 & 10 & 0\\ 6 & 7 & -5\end{bmatrix}=(AB)^{\mathrm{T}}.$$

习　题　3.1(3)

1. 设

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

计算 $A^{\mathrm{T}}, B^{\mathrm{T}}, A+B, A^{\mathrm{T}}+B^{\mathrm{T}}, AB, BA, A^{\mathrm{T}}B^{\mathrm{T}}, B^{\mathrm{T}}A^{\mathrm{T}}, A^2, (A^{\mathrm{T}})^2$.

2. 用两种方法求$(ABC)^{\mathrm{T}}$，其中

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

3. 设 $f(x)$ 是一个系数在数域 P 中的多项式：

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0;$$

A 是数域 P 上的一个 n 阶矩阵，

$$f(A) = a_m A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 E,$$

(其中 E 是 n 阶单位矩阵)称为矩阵 A 的多项式. 求 $f(A)$：

(1) $f(x) = x^2 + x - 1$， $A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 3 \\ 2 & -1 & -4 \end{bmatrix}$；

(2) $f(x) = x^3 - 3x^2 + 3x - 1$， $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

4. 试证：如果 A 与 B 可交换，那么 A 的任一个多项式也与 B 可交换.

5. 设 $AB=BA$. 求证：

(1) $(A+B)^2=A^2+2AB+B^2$；

(2) $A^2-B^2=(A+B)(A-B)$.

3.2 矩阵的分块

这一节介绍在处理阶数较高的矩阵时常用的一个方法，即矩阵的分块. 把一个大矩阵看成是由一些小矩阵组成的，就如矩阵是由数组成的一样，这就是所谓**矩阵的分块**.

先举一个例子说明. 假设给了两个矩阵：

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

将矩阵 A 分成一些小块：

$$A=\left[\begin{array}{cc:cc}1&0&0&0\\0&1&0&0\\ \hdashline -1&2&1&0\\1&1&0&1\end{array}\right]=\begin{bmatrix}E_2&O\\A_1&E_2\end{bmatrix},$$

其中 E_2 是 2 阶单位矩阵,O 是 2 阶零矩阵,而

$$A_1=\begin{bmatrix}-1&2\\1&1\end{bmatrix}.$$

再将 B 分成一些小块:

$$B=\left[\begin{array}{cc:cc}1&0&3&2\\-1&2&0&1\\ \hdashline 1&0&4&1\\-1&-1&2&0\end{array}\right]=\begin{bmatrix}B_{11}&B_{12}\\B_{21}&B_{22}\end{bmatrix},$$

其中 $$B_{11}=\begin{bmatrix}1&0\\-1&2\end{bmatrix},\quad B_{12}=\begin{bmatrix}3&2\\0&1\end{bmatrix},$$

$$B_{21}=\begin{bmatrix}1&0\\-1&-1\end{bmatrix},\quad B_{22}=\begin{bmatrix}4&1\\2&0\end{bmatrix}.$$

在计算 $A+B$ 及 AB 时,把 A,B 看成是由这些小矩阵组成的,于是可以按 2 阶矩阵来运算:

$$A+B=\begin{bmatrix}E_2&O\\A_1&E_2\end{bmatrix}+\begin{bmatrix}B_{11}&B_{12}\\B_{21}&B_{22}\end{bmatrix}=\begin{bmatrix}E_2+B_{11}&B_{12}\\A_1+B_{21}&E_2+B_{22}\end{bmatrix}$$

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

$$AB=\begin{bmatrix}E_2&O\\A_1&E_2\end{bmatrix}\begin{bmatrix}B_{11}&B_{12}\\B_{21}&B_{22}\end{bmatrix}$$

$$= \begin{bmatrix} B_{11} & B_{12} \\ A_1B_{11}+B_{21} & A_1B_{12}+B_{22} \end{bmatrix},$$

其中 $A_1B_{11}+B_{21}$，$A_1B_{12}+B_{22}$ 可以按 2 阶矩阵计算：

$$A_1B_{11}+B_{21} = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} -3 & 4 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ -1 & 1 \end{bmatrix};$$

$$A_1B_{12}+B_{22} = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 1 \\ 2 & 0 \end{bmatrix}$$

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

于是得
$$AB = \begin{bmatrix} 1 & 0 & 3 & 2 \\ -1 & 2 & 0 & 1 \\ -2 & 4 & 1 & 1 \\ -1 & 1 & 5 & 3 \end{bmatrix}.$$

读者可以验证一下，直接按矩阵乘法的定义来计算，结果是一样的.

需要注意的是，应用矩阵分块来进行运算，必须将矩阵分成可以作运算的一些小块. 用分块矩阵来做矩阵加法时，必须将矩阵分成大小相同的小块. 一般地，设 A，B 都是 $s\times n$ 矩阵. 把 A，B 分成同样的小块矩阵：

$$A = \begin{array}{c} \\ s_1 \\ s_2 \\ \vdots \\ s_t \end{array}\!\!\begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \cdots & n_l \end{array} \\ \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1l} \\ A_{21} & A_{22} & \cdots & A_{2l} \\ \cdots & \cdots & \cdots & \cdots \\ A_{t1} & A_{t2} & \cdots & A_{tl} \end{bmatrix} \end{array}$$

$$B=\begin{array}{c} \\ s_1 \\ s_2 \\ \vdots \\ s_t \end{array}\begin{array}{c} \begin{matrix} n_1 & n_2 & \cdots & n_l \end{matrix} \\ \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1l} \\ B_{21} & B_{22} & \cdots & B_{2l} \\ \cdots & \cdots & \cdots & \cdots \\ B_{t1} & B_{t2} & \cdots & B_{tl} \end{bmatrix} \end{array} \quad \begin{array}{l} (s_1+s_2+\cdots+s_t=s; \\ n_1+n_2+\cdots+n_l=n), \end{array}$$

其中 A_{ij} 与 B_{ij} 都是 $s_i \times n_j$ 矩阵($i=1,2,\cdots,t;j=1,2,\cdots,l$),那么

$$A+B=\begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} & \cdots & A_{1l}+B_{1l} \\ A_{21}+B_{21} & A_{22}+B_{22} & \cdots & A_{2l}+B_{2l} \\ \cdots & \cdots & \cdots & \cdots \\ A_{t1}+B_{t1} & A_{t2}+B_{t2} & \cdots & A_{tl}+B_{tl} \end{bmatrix}.$$

如果要用分块矩阵作乘法,由于两个矩阵相乘时,第 1 个矩阵的列数必须等于第 2 个矩阵的行数,所以用分块矩阵计算 AB 时,对矩阵 A 的列的分法必须与矩阵 B 的行的分法相一致.设 A 是一个 $s\times n$ 矩阵,B 是一个 $n\times m$ 矩阵,把 A,B 分成一些小矩阵:

$$A=\begin{array}{c} \\ s_1 \\ s_2 \\ \vdots \\ s_t \end{array}\begin{array}{c} \begin{matrix} n_1 & n_2 & \cdots & n_l \end{matrix} \\ \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1l} \\ A_{21} & A_{22} & \cdots & A_{2l} \\ \cdots & \cdots & \cdots & \cdots \\ A_{t1} & A_{t2} & \cdots & A_{tl} \end{bmatrix} \end{array} \quad \begin{array}{l} (s_1+s_2+\cdots+s_t=s; \\ n_1+n_2+\cdots+n_l=n), \end{array}$$

$$B=\begin{array}{c} \\ n_1 \\ n_2 \\ \vdots \\ n_l \end{array}\begin{array}{c} \begin{matrix} m_1 & m_2 & \cdots & m_r \end{matrix} \\ \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1r} \\ B_{21} & B_{22} & \cdots & B_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ B_{l1} & B_{l2} & \cdots & B_{lr} \end{bmatrix} \end{array} \quad \begin{array}{l} (n_1+n_2+\cdots+n_l=n; \\ m_1+m_2+\cdots+m_n=m). \end{array}$$

其中 A_{ij}是 $s_i\times n_j$ 矩阵($i=1,2,\cdots,t,j=1,2,\cdots,l$);$B_{jk}$是 $n_j\times m_k$ 矩阵($j=1,2,\cdots,l,k=1,2,\cdots,r$),于是

$$AB=\begin{array}{c} \\ s_1 \\ s_2 \\ \vdots \\ s_t \end{array}\begin{array}{c} \begin{array}{cccc} m_1 & m_2 & \cdots & m_r \end{array} \\ \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1r} \\ C_{21} & C_{22} & \cdots & C_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ C_{t1} & C_{t2} & \cdots & C_{tr} \end{bmatrix} \end{array},$$

其中

$$C_{ij}=A_{i1}B_{1j}+A_{i2}B_{2j}+\cdots+A_{il}B_{lj}=\sum_{k=1}^{l}A_{ik}B_{kj}$$
$$(i=1,2,\cdots,t;\ j=1,2,\cdots,r).$$

以上用分块矩阵作加法和乘法运算的两个结果都可以用矩阵的加法与乘法的定义来验证,这里就不详细说明了.

以后会看到,把矩阵分块运算有许多方便之处.因为在分块之后,矩阵间的相互关系可以看得更清楚.作为例子,下面应用分块矩阵来证明关于矩阵的秩的两个定理.

定理 1 两个矩阵的和的秩不超过这两个矩阵的秩的和,即

$$\mathrm{r}(A+B)\leqslant \mathrm{r}(A)+\mathrm{r}(B)$$

证明: 设 A,B 是两个 $s\times n$ 矩阵.用 $A_1,A_2,\cdots,A_s$ 及 $B_1,B_2,\cdots,B_s$ 表示 A 及 B 的行向量.于是 A 与 B 都可以表成分块矩阵:

$$A=\begin{bmatrix} A_1 \\ A_2 \\ \cdots \\ A_s \end{bmatrix},\qquad B=\begin{bmatrix} B_1 \\ B_2 \\ \cdots \\ B_s \end{bmatrix}.$$

从而
$$A+B=\begin{bmatrix}A_1+B_1\\A_2+B_2\\\cdots\cdots\\A_s+B_s\end{bmatrix}.$$
这说明 $A+B$ 的行向量组可以由向量组 $A_1,A_2,\cdots,A_s$ 及 $B_1,B_2,\cdots,B_s$ 线性表出，因此
$$\begin{aligned}\mathrm{r}(A+B)&\leqslant \mathrm{r}\{A_1,A_2,\cdots,A_s,B_1,B_2,\cdots,B_s\}\\&\leqslant \mathrm{r}\{A_1,A_2,\cdots,A_s\}+\mathrm{r}\{B_1,B_2,\cdots,B_s\}\\&=\mathrm{r}(A)+\mathrm{r}(B).\quad |\end{aligned}$$

推论
$$\begin{aligned}&\mathrm{r}(A_1+A_2+\cdots+A_t)\\\leqslant&\mathrm{r}(A_1)+\mathrm{r}(A_2)+\cdots+\mathrm{r}(A_t)\end{aligned}$$

定理 2 矩阵乘积的秩不超过各因子的秩，即
$$\mathrm{r}(AB)\leqslant\min\{\mathrm{r}(A),\mathrm{r}(B)\}.$$

证明：设
$$A=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{s1}&a_{s2}&\cdots&a_{sn}\end{bmatrix},\qquad B=\begin{bmatrix}b_{11}&b_{12}&\cdots&b_{1m}\\b_{21}&b_{22}&\cdots&b_{2m}\\\cdots&\cdots&\cdots&\cdots\\b_{n1}&b_{n2}&\cdots&b_{nm}\end{bmatrix},$$
用 $B_1,B_2,\cdots,B_n$ 表示 B 的行向量，那么 B 可以表成分块矩阵
$$B=\begin{bmatrix}B_1\\B_2\\\cdots\\B_n\end{bmatrix}.$$
于是

$$AB=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} \end{bmatrix}\begin{bmatrix} B_1 \\ B_2 \\ \cdots \\ B_n \end{bmatrix}$$

$$=\begin{bmatrix} a_{11}B_1+a_{12}B_2+\cdots+a_{1n}B_n \\ a_{21}B_1+a_{22}B_2+\cdots+a_{2n}B_n \\ \cdots\cdots \\ a_{s1}B_1+a_{s2}B_2+\cdots+a_{sn}B_n \end{bmatrix}.$$

这说明 AB 的行向量组可以由 B 的行向量组线性表出，所以

$$\mathrm{r}(AB)\leqslant \mathrm{r}(B).$$

再用 $A_1,A_2,\cdots,A_n$ 表示 A 的列向量.那么 A 可以表成分块矩阵

$$A=[A_1 \quad A_2 \quad \cdots \quad A_n].$$

于是

$$AB=[A_1 \quad A_2 \quad \cdots \quad A_n]\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix}$$

$$=\left(\sum_{k=1}^{n} b_{k1}A_k,\ \sum_{k=1}^{n} b_{k2}A_k,\cdots,\sum_{k=1}^{n} b_{km}A_k\right).$$

这说明 AB 的列向量组可以由 A 的列向量组线性表出，所以

$$\mathrm{r}(AB)\leqslant \mathrm{r}(A).$$

综合以上两点，即得

$$\mathrm{r}(AB)\leqslant \min\{\mathrm{r}(A),\mathrm{r}(B)\}. \quad |$$

应用数学归纳法，可以将定理 2 推广到多个因子的情况.

推论 $\mathrm{r}(A_1A_2\cdots A_t)\leqslant \min\{\mathrm{r}(A_1),\ \mathrm{r}(A_2),\cdots,\mathrm{r}(A_t)\}.$

从定理 1,2 的证明可以看出矩阵的和、积的行与列与原矩阵的行列有以下关系.

命题 设矩阵 A 的行向量组是 $A_1,A_2,\cdots,A_s$;列向量组是 $C_1,C_2,\cdots,C_n$;B 的行向量组是 $B_1,B_2,\cdots,B_s$,列向量组是 $D_1,D_2,\cdots,D_n$,则 $A+B$ 的行向量组是 $A_1+B_1,A_2+B_2,\cdots,A_s+B_s$,列向量组是 $C_1+D_1,C_2+D_2,\cdots,C_n+D_n$.

命题 设

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

$$B=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix},$$

则 AB 的行都是 B 的行向量组的线性组合,而且 AB 的第 i 行表成 B 的行向量的线性组合时系数为 A 的第 i 行,即 $a_{i1},a_{i2},\cdots,a_{in}$ $(i=1,2,\cdots,s)$. AB 的列都是 A 的列向量组的线性组合,而且 AB 的第 j 列表成 A 的列向量的线性组合时,系数为 B 的第 j 列,即 $b_{1j},b_{2j},\cdots,b_{nj}(j=1,2,\cdots,m)$.

从定理 2 可以推出:当 A,B 都是 n 阶矩阵时,如果矩阵 A,B 的行列式中有一个等于零,那么这个矩阵的秩就小于 n. 因此,AB 的秩也就小于 n,AB 的行列式也等于零. 比这个结果更确切的,关于矩阵乘积的行列式有下述定理:

定理 3 矩阵乘积的行列式等于矩阵因子的行列式的乘积,即

$$|AB| = |A| \cdot |B|.$$

证明：设

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix},$$

它们的乘积 $D = AB = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix},$

其中 $d_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \quad (i,j = 1,2,\cdots,n).$

另作一个 $2n$ 级矩阵

$$C = \begin{bmatrix} A & O \\ -E & B \end{bmatrix}.$$

由第 1 章 1.6 节的例 6 可推知：

$$|C| = |A| \cdot |B|.$$

下面来证 $|C| = |D|$. 为此，对 C 进行初等列变换：将第 1 列的 b_{11} 倍，第 2 列的 b_{21} 倍，…，第 n 列的 b_{n1} 倍加到第 $n+1$ 列；再将第 1 列的 b_{12} 倍，第 2 列的 b_{22} 倍，…，第 n 列的 b_{n2} 倍加到第 $n+2$ 列；一般地，把第 1 列的 b_{1k} 倍，第 2 列的 b_{2k} 倍，…，第 n 列的 b_{nk} 倍加到第 $n+k$ 列；最后，把第 1 列的 b_{1n} 倍，第 2 列的 b_{2n} 倍，…，第 n 列的 b_{nn} 倍加到第 $2n$ 列. 这样就把矩阵 C 变成矩阵

$$C_1 = \begin{bmatrix} A & D \\ -E & O \end{bmatrix}.$$

因为初等变换不改变矩阵的行列式，所以

$$|C| = |C_1|.$$

现在

$$\begin{aligned} |C_1| &= (-1)^n |D| \cdot |-E| \\ &= (-1)^n \cdot |D| \cdot (-1)^n = |D|. \end{aligned}$$

因此 $$|C| = |D|.$$

这就证明了 $$|AB| = |A| \cdot |B|. \quad |$$

这个定理也可以推广到多个因子的情形：

推论 $|A_1 A_2 \cdots A_t| = |A_1| \cdot |A_2| \cdots |A_t|$.

当然，这里的 $A_i (i = 1, 2, \cdots, t)$ 都是 n 阶矩阵.

定义 5 如果 n 级矩阵 A 的行列式 $|A|$ 不等于零，则称 A 为**非退化的**. 否则，称为**退化的**.

根据定义，可知：

命题 一个 n 阶矩阵是非退化的充分必要条件是：它的秩等于 n.

因此，非退化矩阵也称**满秩矩阵**.

从定理 3，可以推得：

命题 设 A, B 都是 n 阶矩阵，矩阵 AB 是非退化的充分必要条件是：A, B 都是非退化的.

习 题 3.2

1. 用矩阵的分块方法计算 AB，其中

$$A = \begin{bmatrix} 1 & -2 & 7 & 0 & 0 \\ -1 & 3 & 6 & 0 & 0 \\ -3 & 2 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 0 & 0 & 1 & 2 \\ 0 & 3 & 0 & 3 & 4 \\ 0 & 0 & 3 & 5 & 6 \\ 0 & 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 5 & -1 \end{bmatrix}.$$

2. 设 $A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{bmatrix}$，其中 A_i 是 n_i 阶方阵，求证 $|A| = |A_1||A_2|\cdots|A_r|$.

3. (1) 试找两个矩阵 A、B，使

$$r(A+B) = r(A) + r(B);$$

(2) 试找两个矩阵 A、B，使

$$r(A+B) < r(A) + r(B).$$

4. (1) 试找两个矩阵 A、B，使

$$r(AB) = \min[r(A), r(B)];$$

(2) 试找两个矩阵 A、B，使

$$r(AB) < \min[r(A), r(B)].$$

5. 设 A 是一个 n 阶矩阵. 试证：存在一个 n 阶非零矩阵 B，使得 $AB=0$ 的充分必要条件是：$|A|=0$.

6. 设 A，B 都是 n 阶矩阵. 试证：如果 $AB=0$，那么

$$r(A) + r(B) \leqslant n.$$

3.3 矩阵的逆

以上两节介绍了矩阵的运算. 当两个矩阵的行数、列数适当的时候，可以相加、相减及相乘. 那么给了两个矩阵，能不能做除法呢？说得具体一些，就是：给了两个矩阵 A，B，能不能找到满足

$$AX = B, \quad YA = B$$

的矩阵 X 与 Y 呢？这一节就是讨论这个问题.

根据矩阵乘积的秩的定理，我们知道不是对于任意的 A，B 都可找到满足上述方程的 X 与 Y 的. 那么需要 A，B 满足什么条件

时，X 与 Y 才存在呢？为此，首先讨论方程

$$AX = E, \quad YA = E$$

的解.

这一节讨论的矩阵，如不特别说明，都是指 n 阶方阵.

定义 6 对于矩阵 A，如果有矩阵 B，使得

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

则 A 称为**可逆的**；B 称为 A 的**逆矩阵**，记作 A^{-1}.

关于这个定义，应注意以下两点：首先，满足(1)式的矩阵 B 是唯一的(如果存在的话). 这一点可以这样来证明：如果有 B_1，B_2 两个满足条件(1)的矩阵，那么

$$\begin{aligned} B_1AB_2 &= B_1(AB_2) = B_1E = B_1 \\ &= (B_1A)B_2 = EB_2 = B_2, \end{aligned}$$

即

$$B_1 = B_2.$$

这就证明了 A 的逆矩阵是唯一的.

其次，如果矩阵 B 满足

$$BA = E,$$

那么，B 一定也满足

$$AB = E.$$

由于矩阵的乘法一般是不可交换的，所以在定义 6 中，特地强调指出逆矩阵满足 $AB=BA=E$.

例 1 单位矩阵 E 是可逆矩阵，而且

$$E^{-1} = E.$$

这是因为 $EE=E$ 的缘故.

例 2 因为对任意矩阵 A，$OA=AO=O$，所以零矩阵不是可逆矩阵.

例 3 设 $a_1, a_2, \cdots, a_n$ 都不等于零，求证：对角矩阵

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

是可逆矩阵，并且

$$A^{-1}=\begin{bmatrix}a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a_n^{-1}\end{bmatrix}.$$

证明：因为

$$\begin{bmatrix}a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n\end{bmatrix}\begin{bmatrix}a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a_n^{-1}\end{bmatrix}=\boldsymbol{E},$$

$$\begin{bmatrix}a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a_n^{-1}\end{bmatrix}\begin{bmatrix}a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n\end{bmatrix}=\boldsymbol{E},$$

所以根据定义 6，结论成立.

下面来讨论矩阵 A 可逆的条件是什么？如果 A 可逆，怎样来求 A^{-1}？

从可逆矩阵的定义可以看出，A 可逆的条件必须是非退化的.反过来，如果 A 是非退化的，那么 A 是否一定是可逆的呢？换句话说，当 A 为非退化时，是否一定能够找到满足条件(1)的矩阵 B 呢？为此，先来观察一下 A 的逆矩阵的元素与 A 的元素之间有什

么关系.

设 $$A=[a_{ij}],\ B=[b_{ij}],$$
如果 $AB=E$,那么 B 的第 j 列元素 $b_{1j},b_{2j},\cdots,b_{nj}$适合下列等式:
$$\sum_{i=1}^{n}a_{ik}b_{kj}=\begin{cases}1, & i=j;\\ 0, & i\neq j.\end{cases}$$
因此根据行列式按一行展开的公式,可以把 b_{ij}求出来.继续讨论,需要用到下述定义.

定义 7 设 A_{ij}是矩阵
$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}$$
中元素 a_{ij}的代数余子式.矩阵
$$A^*=\begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \cdots & \cdots & \cdots & \cdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{bmatrix}$$
称为 A 的**伴随矩阵**.

根据按一行(列)展开的公式,可得到
$$AA^*=A^*A=\begin{bmatrix} |A| & 0 & \cdots & 0\\ 0 & |A| & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & |A|\end{bmatrix}=|A|E.$$

例 4 设

$$A=\begin{bmatrix}1 & -4 & -3\\ 2 & -5 & -3\\ -1 & 2 & 1\end{bmatrix},$$

求 A^*.

解：因为

$$A_{11}=1,\quad A_{12}=1,\quad A_{13}=-1,$$
$$A_{21}=-2,\quad A_{22}=-2,\quad A_{23}=2,$$
$$A_{31}=-3,\quad A_{32}=-3,\quad A_{33}=3,$$

所以

$$A^*=\begin{bmatrix}1 & -2 & -3\\ 1 & -2 & -3\\ -1 & 2 & 3\end{bmatrix}.$$

读者可以验证一下

$$AA^*=A^*A=|A|E=O.$$

如果 $|A|\neq 0$，那么就有

$$A\left(\frac{1}{|A|}A^*\right)=\left(\frac{1}{|A|}A^*\right)A=E, \tag{2}$$

即得 $A^{-1}=\frac{1}{|A|}A^*$.

因此，有以下定理：

定理 4 矩阵 A 是可逆的充分必要条件是：A 是非退化的，而且当 A 可逆时，

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

证明：如果 A 可逆，那么有 A^{-1}，使

$$AA^{-1}=E.$$

取等式两端的行列式，得

$$|A|\cdot|A^{-1}|=|E|=1,$$

所以$|A|\neq 0$,即 A 是非退化的.

如果 A 是非退化的,那么$|A|\neq 0$. 由(2)式即知 A^{-1}存在,且

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

定理中的公式(3)就是求逆阵的公式.

例 5 判断矩阵

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

是否可逆. 如果可逆,求 A^{-1}.

解: 因为

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

所以 A 是可逆的.

又因 $A_{11}=2$, $A_{12}=2$, $A_{13}=-4$,

$A_{21}=-1$, $A_{22}=-1$, $A_{23}=3$,

$A_{31}=1$, $A_{32}=-1$, $A_{33}=-1$,

所以 $$A^{-1}=\frac{1}{2}\begin{bmatrix}2 & -1 & 1\\ 2 & -1 & -1\\ -4 & 3 & -1\end{bmatrix}.$$

可逆矩阵有下述一些性质:

性质 1 根据逆矩阵的定义

$$AB=BA=E$$

可以看到:A,B 互为逆矩阵,即$(A^{-1})^{-1}=A$.

性质 2 $|A^{-1}| = \frac{1}{|A|}$.

性质 3 如果矩阵 A 可逆，那么转置矩阵 A^T 也可逆，而且

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

这是因为 $A^T \cdot (A^{-1})^T = (A^{-1}A)^T = E^T = E.$

性质 4 如果 A,B 都可逆，那么 AB 也可逆. 而且

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

证明：因为

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1}\\ &= AEA^{-1} = AA^{-1} = E.\end{aligned}$$ |

性质 4 可以推广到几个因子的情形：

如果 $A_1,A_2,\cdots,A_s$ 是同阶可逆矩阵，则 $A_1A_2\cdots A_s$ 也可逆，并且

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

性质 5 如果 $A_1,A_2,\cdots,A_s$ 都是可逆矩阵，则

$$A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{bmatrix}$$

也可逆，并且

$$A^{-1} = \begin{bmatrix} A_1^{-1} & & & \\ & A_2^{-1} & & \\ & & \ddots & \\ & & & A_s^{-1} \end{bmatrix}.$$

定理 4 不但给出了一个矩阵可逆的条件，同时也给出了求逆

矩阵的公式(3). 按照这个公式来求逆矩阵，计算量一般是非常大的，在以后我们将给出另一种求逆矩阵的方法.

关于矩阵及可逆矩阵乘积的秩之间有下述关系：

定理 5 设 A 是一个 $s\times n$ 矩阵，P 是 s 阶可逆矩阵，Q 是 n 阶可逆矩阵，那么

$$\mathrm{r}(A)=\mathrm{r}(PA)=\mathrm{r}(AQ)=\mathrm{r}(PAQ).$$

证明：令

$$B=PA,$$

由定理 2 知：
$$\mathrm{r}(B)\leqslant \mathrm{r}(A).$$

但是
$$A=P^{-1}B,$$

故再根据定理 2，又可得

$$\mathrm{r}(A)\leqslant \mathrm{r}(B),$$

所以
$$\mathrm{r}(A)=\mathrm{r}(B)=\mathrm{r}(PA).$$

同样可以证明等式

$$\mathrm{r}(A)=\mathrm{r}(AQ).$$

综合这两个等式，即得

$$\mathrm{r}(A)=\mathrm{r}(PAQ). \quad |$$

最后，讨论矩阵方程

$$AX=B.$$

如果 A 是一个可逆方阵，那么这个方程有唯一解

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

这里的 B 可以推广到 $n\times m$ 矩阵的情形，即：如果 A 是一个 n 阶可逆矩阵，B 是一个 $n\times m$ 矩阵，那么方程

$$AX=B$$

有唯一解
$$X=A^{-1}B.$$

且解 X 也是一个 $n\times m$ 矩阵.

特别地，当 B 是一个 $n\times 1$ 矩阵，即列矩阵时，这个公式给出克莱姆法则的另一个证明：

线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n, \end{cases} \tag{4}$$

可以写成矩阵方程

$$AX = B. \tag{5}$$

其中

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

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

如果 $|A| \neq 0$，那么 A 可逆，方程(5)有唯一解

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

这也是方程组(4)的解. 将 $A^{-1} = \dfrac{1}{|A|}A^*$ 代入，乘出来，就是克莱姆法则给出的公式.

如果 A 不是可逆方阵，我们来讨论更一般的情形. 设 A 是一个 $s\times n$ 矩阵，B 是一个 $s\times m$ 矩阵，考虑方程

$$AX = B. \tag{6}$$

这里的未知矩阵 X 是一个 $n\times m$ 矩阵. 现在用分块矩阵来讨论这

个问题，将 B,X 用列向量表成

$$B=(B_1 \quad B_2 \quad \cdots \quad B_m);$$

$$X=(X_1 \quad X_2 \quad \cdots \quad X_m),$$

其中 B_i 都是 $s\times 1$ 列矩阵；X_i 都是 $n\times 1$ 列矩阵. 于是方程(6)可写成

$$A(X_1 \quad X_2 \quad \cdots \quad X_m)=(B_1 \quad B_2 \quad \cdots \quad B_m).$$

它等价于方程组

$$\begin{cases} AX_1=B_1, \\ AX_2=B_2, \\ \cdots\cdots\cdots\cdots \\ AX_m=B_m. \end{cases} \tag{7}$$

再将 A 表成分块矩阵：

$$A=(A_1 \quad A_2 \quad \cdots \quad A_n),$$

其中 A_j 是 A 的第 j 个列向量，是 $s\times 1$ 矩阵. 于是方程组(7)可写成

$$(A_1 \quad A_2 \quad \cdots \quad A_n)X_k=B_k \quad (k=1,2,\cdots,m). \tag{8}$$

因为方程组(8)有解的条件是 $B_k(k=1,2,\cdots,m)$ 可以被 $A_1,A_2,\cdots,A_n$ 线性表出，所以方程(6)有解的充分必要条件是 B 的列向量组可以由 A 的列向量组线性表出. 这个结论当然也包含 A 是可逆矩阵的情形.

方程 $YA=B$ 也可以类似地讨论，这里就不详细讲了.

例 6 求 X，使

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

解：

$$X=\begin{bmatrix}3&0&8\\3&-1&6\\-2&0&-5\end{bmatrix}^{-1}\begin{bmatrix}1&-1&2\\-1&3&4\\-2&0&5\end{bmatrix}$$

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

$$=\begin{bmatrix}11&5&-50\\10&0&-40\\-4&-2&19\end{bmatrix}.$$

习 题 3.3

1. 求下列矩阵 A 的伴随矩阵 A^*，并验证 $A^*A=AA^*=|A|\cdot E$.

(1) $A=\begin{bmatrix}3&1\\0&2\end{bmatrix}$；　　(2) $A=\begin{bmatrix}3&7&-3\\-2&-5&2\\-4&-10&3\end{bmatrix}$；

(3) $A=\begin{bmatrix}1&2&-1\\0&5&-3\\-1&2&4\end{bmatrix}$.

2. 求下列矩阵的逆矩阵：

(1) $\begin{bmatrix}a&b\\c&d\end{bmatrix}$　$(ad-bc\neq 0)$；

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

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

3. 试证：如果 $A^k=0$，那么 $E-A$ 可逆，并且

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

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

5. 试证：如果 A 与 B 可交换，并且 A 是可逆的，那么 A^{-1} 与 B 也可交换.

6. 求满足下列条件的 X：

(1) $\begin{bmatrix}1 & -5\\ -1 & 4\end{bmatrix}X=\begin{bmatrix}3 & 2\\ 1 & 4\end{bmatrix}$；

(2) $X\begin{bmatrix}1 & -1 & 1\\ 1 & 1 & 0\\ 2 & 1 & 1\end{bmatrix}=\begin{bmatrix}1 & 2 & -3\\ 2 & 0 & 4\\ 0 & -1 & 5\end{bmatrix}$；

(3) $\begin{bmatrix}1 & -1 & 1\\ 1 & 1 & 0\\ 3 & 2 & 1\end{bmatrix}X\begin{bmatrix}1 & -1 & 1\\ 1 & 1 & 0\\ 3 & 2 & 1\end{bmatrix}=\begin{bmatrix}4 & 2 & 3\\ 0 & -1 & 5\\ 2 & 1 & 1\end{bmatrix}$；

(4) $\begin{bmatrix}1 & -1 & 0\\ 2 & 0 & 1\end{bmatrix}X=\begin{bmatrix}2 & 5\\ 1 & 4\end{bmatrix}$.

7. (1) 假设 A，B 都可逆. 求证：$\begin{bmatrix}O & A\\ B & O\end{bmatrix}$ 可逆，并且

$$\begin{bmatrix}O & A\\ B & O\end{bmatrix}^{-1}=\begin{bmatrix}O & B^{-1}\\ A^{-1} & O\end{bmatrix};$$

(2) 计算

$$\begin{bmatrix}3 & -2 & 0 & 0\\ 5 & -3 & 0 & 0\\ 0 & 0 & 3 & 4\\ 0 & 0 & 1 & 1\end{bmatrix}^{-1} \text{ 及 } \begin{bmatrix}0 & 0 & 0 & 1 & 2\\ 0 & 0 & 0 & 2 & 3\\ 1 & 1 & 0 & 0 & 0\\ 0 & 1 & 1 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\end{bmatrix}^{-1}.$$

8. 假设 A，B 都可逆，求证

$$\begin{bmatrix}A & O\\ C & B\end{bmatrix} \text{ 及 } \begin{bmatrix}A & D\\ O & B\end{bmatrix}$$

也都可逆，并且

$$\begin{bmatrix} A & O \\ C & B \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & O \\ -B^{-1}CA^{-1} & B^{-1} \end{bmatrix},$$

$$\begin{bmatrix} A & D \\ O & B \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & A^{-1}DB^{-1} \\ O & B^{-1} \end{bmatrix}.$$

3.4 等价矩阵

在上一章里，曾经介绍过应用初等变换求矩阵的秩的方法. 在那里，我们看到，用初等变换可以将矩阵化简，而且变换前后的矩阵具有许多共同的性质. 可以用初等变换互相变换的矩阵称为**等价矩阵**. 这一节讨论等价矩阵的标准形，并在此基础上给出用初等变换求逆矩阵的方法.

定义 8 如果矩阵 B 可以从矩阵 A 经过一系列初等变换而得到，则称矩阵 A 与 B **是等价的**.

等价是矩阵间的一种关系. 不难证明，这种关系具有下述 3 个性质：

1. 反身性：矩阵 A 与本身总是等价的.

2. 对称性：如果矩阵 A 与 B 等价，那么矩阵 B 也与 A 等价.

3. 传递性：如果 A 与 B 等价，B 与 C 等价，那么 A 与 C 也等价.

为了将矩阵间的等价关系通过矩阵运算来表示，下面先来建立矩阵的初等变换与矩阵乘法的联系.

定义 9 由单位矩阵 E 经过一次初等变换而得到的矩阵称为**初等矩阵**.

因为初等变换有 3 种，所以初等矩阵也有 3 类. 每个初等行变换都有一个初等矩阵与之对应.

1. 把矩阵 E 的 i 行与 j 行互换后，得到初等矩阵

$$
P(i,j)=\begin{array}{c}\left[\begin{array}{ccccccccc}
1 & & & & & & & & \\
 & \ddots & & & & & & & \\
 & & 1 & & & & & & \\
 & & & 0 & \cdots & \cdots & 1 & & \\
 & & & \vdots & 1 & & \vdots & & \\
 & & & \vdots & & \ddots & \vdots & & \\
 & & & \vdots & & 1 & \vdots & & \\
 & & & 1 & \cdots & \cdots & 0 & & \\
 & & & & & & & 1 & \\
 & & & & & & & & \ddots \\
 & & & & & & & & & 1
\end{array}\right]\end{array}
\begin{array}{l} \\ \\ \\ \text{第 } i \text{ 行} \\ \\ \\ \\ \text{第 } j \text{ 行} \\ \\ \\ \\ \end{array}.
$$

2. 用数域 $\mathbf{P}$ 中的非零数 c 乘 E 的第 i 行，得到初等矩阵

$$
P(i(c))=\left[\begin{array}{ccccccc}
1 & & & & & & \\
 & \ddots & & & & & \\
 & & 1 & & & & \\
 & & & c & & & \\
 & & & & 1 & & \\
 & & & & & \ddots & \\
 & & & & & & 1
\end{array}\right]
\begin{array}{l} \\ \\ \\ \text{第 } i \text{ 行}. \\ \\ \\ \\ \end{array}
$$

3. 把矩阵 E 的第 j 行的 k 倍加到第 i 行上，得到初等矩阵

$$
\begin{array}{cc}
 & \begin{array}{ccc} \text{第 } i \text{ 列} & & \text{第 } j \text{ 列} \end{array} \\
P(i,j(k))= & \left[\begin{array}{ccccccc}
1 & & & & & & \\
 & \ddots & & & & & \\
 & & 1 & \cdots & k & & \\
 & & & \ddots & \vdots & & \\
 & & & & 1 & & \\
 & & & & & \ddots & \\
 & & & & & & 1
\end{array}\right]
\begin{array}{l} \\ \\ \text{第 } i \text{ 行} \\ \\ \text{第 } j \text{ 行} \\ \\ \\ \end{array}.
\end{array}
$$

同样可以得到与列变换相应的初等矩阵，请读者自己试作一

下. 可以看到,对单位矩阵作 1 次初等列变换所得到的矩阵也包括在上面这 3 类矩阵之中. 譬如说,把 E 的第 i 列的 k 倍加到第 j 列上,仍然得到 $P(i,j(k))$. 因此,这 3 类矩阵也就是全部初等矩阵.

不难看出,初等矩阵都是可逆的,它们的逆矩阵还是初等矩阵:

$$P(i,j)^{-1} = P(i,j);$$
$$P(i(c))^{-1} = P(i(c^{-1}));$$
$$P(i,j(k))^{-1} = P(i,j(-k)).$$

我们对矩阵进行初等变换,可以通过用乘以初等矩阵的形式来表示. 下述定理说明了这一点.

定理 6 设 A 是一个 $s\times n$ 矩阵. 对 A 施行 1 次初等行变换,就相当于在 A 的左边乘上一个相应的 s 阶初等矩阵;对 A 施行 1 次初等列变换,就相当于在 A 的右边乘上一个相应的 n 阶初等矩阵.

证明: 在此只证明行变换的情形,列变换的情形可以同样证明.

设 A 的行向量为 $A_1, A_2, \cdots, A_s$. 将 A 表成分块矩阵

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_s \end{bmatrix}.$$

用 s 阶初等矩阵 $P(i,j)$ 左乘 A,得

$$P(i,j)\cdot A=\left[\begin{array}{c}A_1\\ \vdots\\ A_j\\ \vdots\\ A_i\\ \vdots\\ A_s\end{array}\right]\begin{array}{l} \\ \\ i\text{ 行}\\ \\ j\text{ 行}\\ \\ \\ \end{array}.$$

这相当于把第 i 行与第 j 行互换;用 $P(i(c))$左乘 A,得

$$P(i(c))\cdot A=\left[\begin{array}{c}A_1\\ \vdots\\ cA_i\\ \vdots\\ A_s\end{array}\right]\begin{array}{l} \\ \\ \text{第 } i\text{ 行}\\ \\ \\ \end{array}$$

这相当于用 c 乘 A 的第 i 行;用 $P(i,j(k))$左乘 A,得

$$P(i,j(k))\cdot A=\left[\begin{array}{c}A_1\\ \vdots\\ A_i+kA_j\\ \vdots\\ A_j\\ \vdots\\ A_s\end{array}\right]\begin{array}{l} \\ \\ i\text{ 行}\\ \\ j\text{ 行}\\ \\ \\ \end{array},$$

这相当于把 A 的第 j 行的 k 倍加到第 i 行上.

根据这个定理,可以把矩阵的等价关系用矩阵的乘法方式表

示出来.即

推论 矩阵 A,B 等价的充分必要条件是：有初等矩阵 $P_1,\cdots,P_l,Q_1,\cdots,Q_t$，使

$$A = P_1P_2\cdots P_lBQ_1Q_2\cdots Q_t.$$

我们已知，用初等行变换可以将矩阵化成梯阵和简化梯阵.如果同时施行初等行变换与初等列变换于矩阵上，那么，矩阵还可以进一步化简.看下述定理.

定理 7 任意一个矩阵 A 都与一个形如

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

的矩阵等价，它称为矩阵 A（在等价关系下）的**标准形**，主对角线上 1 的个数等于 A 的秩（1 的个数可以是零）.

证明：如果 $A=0$，那么它已经是标准形了.以下设 $A\neq 0$.经过互换两行及互换两列，A 一定可以变成一个左上角的元素不为零的矩阵，因此可设 $a_{11}\neq 0$.

从第 $i(i=2,\cdots,s)$行减去第 1 行的 $a_{11}^{-1}a_{i1}$倍；从第 $j(j=2,\cdots,n)$列减去第 1 列的 $a_{11}^{-1}a_{1j}$倍后，再用 a_{11}^{-1}乘第 1 行，A 就变成

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ & & A_1 & \\ 0 & & & \end{bmatrix}.$$

其中 A_1 是一个 $(s-1)\times(n-1)$ 矩阵. 对 A_1 再重复以上的步骤. 这样继续变换下去,就可得到所要的标准形.

标准形的秩就等于它的主对角线上 1 的个数. 我们知道,初等变换不改变矩阵的秩,所以标准形中 1 的个数也就是矩阵 A 的秩. |

例 1 用初等变换把矩阵 A 化为标准形:

$$A=\begin{bmatrix} 0 & 0 & 3 & 1 \\ 2 & 1 & -1 & 2 \\ 4 & 2 & 3 & 1 \\ -2 & -1 & 4 & -3 \end{bmatrix}.$$

解:

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

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

$$\to\begin{bmatrix}1&0&0&0\\0&0&3&1\\0&0&5&-3\\0&0&3&-1\end{bmatrix}\to\begin{bmatrix}1&0&0&0\\0&1&3&0\\0&-3&5&0\\0&-1&3&0\end{bmatrix}$$

$$\to\begin{bmatrix}1&0&0&0\\0&1&3&0\\0&0&14&0\\0&0&6&0\end{bmatrix}\to\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&14&0\\0&0&0&0\end{bmatrix}$$

$$\to\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix}.$$

从前面两节可看到，n 阶可逆矩阵的秩为 n，所以可逆矩阵的标准形为单位矩阵；显然，反过来也是对的. 这说明一个矩阵可逆的充分必要条件是：它与单位矩阵等价. 因此，由定理 6 的推论有：

定理 8 n 阶方阵 A 可逆的充分必要条件是：它能表成一些初等矩阵的乘积：

$$A = Q_1Q_2\cdots Q_m.$$

把这个式子改写一下，即得

$$Q_m^{-1}Q_{m-1}^{-1}\cdots Q_2^{-1}Q_1^{-1}A = E.$$

因为初等矩阵的逆矩阵还是初等矩阵，而在矩阵 A 的左边乘上一个初等矩阵就相当于对 A 作一个初等行变换. 因此，上式说明了可逆矩阵的一个重要性质：

推论 1 可逆矩阵可以经过一系列初等行变换化成单位矩阵.

从定理 6 的推论及定理 8，还可推得：

推论 2 两个 $s\times n$ 矩阵 A、B 等价的充分必要条件是：存在可逆的 s 阶矩阵 P 和可逆的 n 阶矩阵 Q，使得

$$B = PAQ.$$

上一节的定理 4 给出了当矩阵 A 可逆时，计算逆矩阵 A^{-1}的公式 $A^{-1}=\frac{1}{|A|}A^*$. 但按照这个公式来求逆矩阵时，计算量一般是非常大的. 下面介绍另一种求逆矩阵的方法，即用初等变换求逆矩阵的方法.

根据定理 8 的推论 1，如果 A 可逆，那么 A 可以经过一系列初等行变换化成单位矩阵. 而对一个矩阵作一次初等行变换就相当于将一个相应的初等矩阵左乘这个矩阵. 因此，存在一系列初等矩阵 $P_1,P_2,\cdots,P_m$，使

$$P_m\cdots P_2P_1A = E.$$

于是

$$P_m\cdots P_2P_1E = A^{-1}.$$

这两个式子说明：如果用一系列初等行变换把可逆矩阵 A 化为单位矩阵，那么用同样的初等行变换作用于单位矩阵 E，就可得到 A 的逆矩阵 A^{-1}. 从而得到一个用初等变换求逆矩阵的方法：

作一个 $n\times 2n$ 矩阵

$$[AE].$$

对这个矩阵施行初等行变换，将它的左半部的矩阵 A 化为单位矩阵. 那么，右半部的单位矩阵 E 就同时化成了 A^{-1}：

$$[AE]\longrightarrow[EA^{-1}].$$

例 2 设

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

求 A^{-1}.

解：对$[AE]$施行初等行变换：

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

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

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

$$\rightarrow\begin{bmatrix}-1 & 1 & 4 & 0 & 1 & 0\\ 0 & 1 & 5 & 0 & 2 & 1\\ 0 & 0 & -9 & 1 & -4 & -2\end{bmatrix}$$

$$\rightarrow\begin{bmatrix}1 & -1 & -4 & 0 & -1 & 0\\ 0 & 1 & 5 & 0 & 2 & 1\\ 0 & 0 & 1 & -\frac{1}{9} & \frac{4}{9} & \frac{2}{9}\end{bmatrix}$$

$$\rightarrow\begin{bmatrix}1 & 0 & 0 & \frac{1}{9} & \frac{5}{9} & \frac{7}{9}\\ 0 & 1 & 0 & \frac{5}{9} & -\frac{2}{9} & -\frac{1}{9}\\ 0 & 0 & 1 & -\frac{1}{9} & \frac{4}{9} & \frac{2}{9}\end{bmatrix},$$

所以 $$A^{-1}=\begin{bmatrix}\frac{1}{9} & \frac{5}{9} & \frac{7}{9}\\ \frac{5}{9} & -\frac{2}{9} & -\frac{1}{9}\\ -\frac{1}{9} & \frac{4}{9} & \frac{2}{9}\end{bmatrix}.$$

习 题 3.4

1. 写出 4 阶方阵的全部可能的等价标准形.

2. 试证：矩阵 A,B 等价的充分必要条件是：它们的标准形相同.

3. 试证：矩阵 A,B 等价的充分必要条件是：它们的秩相等.

4. 用初等变换将下列矩阵化为标准形：

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

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

5. 求下列矩阵的逆矩阵：

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

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

3.5 正交矩阵

为了以后的应用，这一节介绍一类特殊的矩阵：正交矩阵. 本节中所提到的矩阵、向量及数都是实的.

定义 10 如果实数域上的矩阵 A 满足 $AA^{\mathrm{T}}=A^{\mathrm{T}}A=E$，那么，$A$ 称为**正交矩阵**.

例如，实矩阵 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$，$\begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$和$\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$

都是正交矩阵.

正交矩阵有以下一些性质：

1. 正交矩阵的行列式等于 1 或－1.

证明：设 A 是正交矩阵，则

$$AA^{\mathrm{T}}=E$$

两边取行列式得：

$$|A||A^{\mathrm{T}}| = |E| = 1,$$

于是
$$|A|^2 = 1,$$
$$|A| = \pm 1.$$

2. 如果 A 是正交矩阵，则 $A^{-1} = A^{\mathrm{T}}$.

3. 如果 A 是正交矩阵，则 A^{-1} 也是正交矩阵.

4. 如果 A,B 是同阶正交矩阵，则它们的乘积 AB 也是正交矩阵.

这些性质的证明留给读者作为习题.

下面来讨论正交矩阵的元素之间的关系.

设

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

是一个正交矩阵，根据正交矩阵的定义：

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

$$= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

推出
$$\begin{cases} a_{i1}^2 + a_{i2}^2 + \cdots + a_{in}^2 = 1, \\ a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{in}a_{jn} = 0 \quad (\text{当 } i \neq j \text{ 时}), \end{cases}$$

即
$$\sum_{k=1}^{n} a_{ik}a_{jk} = \delta_{ij} \qquad (i,j = 1,2,\cdots,n).$$
式中的 δ_{ij}，当 $i=j$ 时，为 1；当 $i\neq j$ 时，为 0. 这组等式表明正交矩阵的行向量之间的重要关系，通常称为**正交条件**.

同样地，正交矩阵的列向量也满足正交条件：
$$\sum_{k=1}^{n} a_{ki}a_{kj} = \delta_{ij}.$$
这组等式可以从 $A^{T}A=E$ 得到.

为了进一步讨论有关正交矩阵的问题，引入下列定义：

定义 11 两个 n 维实向量
$$\boldsymbol{\alpha} = (a_1,a_2,\cdots,a_n), \quad \boldsymbol{\beta} = (b_1,b_2,\cdots,b_n)$$
的**内积** (α,β) 定义为
$$(\boldsymbol{\alpha},\boldsymbol{\beta}) = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$
当 $(\alpha,\beta)=0$ 时，则称 α 与 β **正交**.

由定义可知，零向量与任何向量都是正交的.

定义 12 如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中任意两个向量都正交，而且每个 $\alpha_i(i=1,2,\cdots,s)$ 都不是零向量，那么，这个向量组就称为**正交向量组**.

例如 n 个基本向量组成一个正交向量组.

如果 A 是一个正交矩阵，那么正交条件
$$\sum_{k=1}^{n} a_{ik}a_{jk} = 0 \quad (i\neq j, i,j = 1,2,\cdots,n)$$
说明了 A 的 n 个行向量是两两正交的. 又由于 A 可逆，所以它的行向量都不是零向量，这说明正交矩阵 A 的行向量组成一个正交向量组. 同样地，正交条件

$$\sum_{k=1}^{n} a_{ki}a_{kj} = 0 \quad (i \neq j, i, j = 1, 2, \cdots, n)$$

说明了 A 的列向量也组成一个正交向量组.

为了刻划条件

$$\sum_{k=1}^{n} a_{ik}^2 = 1; \qquad \sum_{k=1}^{n} a_{ki}^2 = 1,$$

引入以下定义:

定义 13 设 $\boldsymbol{\alpha} = (a_1, a_2, \cdots, a_n)$ 是一个 n 维实向量,令 $|\boldsymbol{\alpha}| = \sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}$. $|\boldsymbol{\alpha}|$ 称为 $\boldsymbol{\alpha}$ 的**长度**. 如果 $|\boldsymbol{\alpha}| = 1$,则 $\boldsymbol{\alpha}$ 称为**单位向量**.

因为 $\boldsymbol{\alpha} = (a_1, a_2, \cdots, a_n)$ 是实向量,所以 $(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = a_1^2 + a_2^2 + \cdots + a_n^2 \geqslant 0$, $|\boldsymbol{\alpha}| = \sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}$ 总是有意义的.

从定义可知: $|\boldsymbol{\alpha}| = 0$ 的充分必要条件是 $\boldsymbol{\alpha} = \mathbf{0}$; $\boldsymbol{\alpha}$ 是单位向量的充分必要条件是 $(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = 1$.

例如,基本向量以及正交矩阵的行向量与列向量都是单位向量.

根据定义,有:

定理 9 n 阶矩阵 A 是正交矩阵的充分必要条件是: 它的行(列)向量组是正交的单位向量组.

因此,只要找出 n 个正交的 n 维单位向量,以它们为行(或列)作成的矩阵一定是正交矩阵. 下面将介绍一种找正交向量组的方法——正交化方法. 在介绍这个方法以前,先对向量的内积进行一些讨论.

向量的内积具有下列性质:

1. $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\boldsymbol{\beta}, \boldsymbol{\alpha})$;

2. $(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2, \boldsymbol{\beta}) = (\boldsymbol{\alpha}_1, \boldsymbol{\beta}) + (\boldsymbol{\alpha}_2, \boldsymbol{\beta})$;

3. $(k\boldsymbol{\alpha},\boldsymbol{\beta})=k(\boldsymbol{\alpha},\boldsymbol{\beta})$ （k 是任意实数）.

这些性质都可以直接从内积的定义推得.请读者自己验证一下.

从这 3 个性质又可以推出：

$$(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2,\boldsymbol{\beta})=k_1(\boldsymbol{\alpha}_1,\boldsymbol{\beta})+k_2(\boldsymbol{\alpha}_2,\boldsymbol{\beta});$$

$$(\boldsymbol{\alpha},l_1\boldsymbol{\beta}_1+l_2\boldsymbol{\beta}_2)=l_1(\boldsymbol{\alpha},\boldsymbol{\beta}_1)+l_2(\boldsymbol{\alpha},\boldsymbol{\beta}_2).$$

这两个性质均可推广到一些向量的线性组合的情形.

下面证明关于正交向量组的一个重要性质.

定理 10 正交向量组一定是线性无关的.

证明：设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是一个正交向量组，如果

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s=0,$$

那么 $(\boldsymbol{\alpha}_i,k_1\boldsymbol{\alpha}_1+\cdots+k_s\boldsymbol{\alpha}_s)=0\quad(i=1,2,\cdots,s),$

展开得 $k_1(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_1)+k_2(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_2)+\cdots+k_s(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_s)=0.$

因为 $\boldsymbol{\alpha}_i$ 与 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_s$ 都正交，所以

$$k_i(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)=0\quad(i=1,2,\cdots,s).$$

又因 $\boldsymbol{\alpha}_i\neq0$，所以 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)\neq0$. 由此得

$$k_i=0\quad(i=1,2,\cdots,s).$$

这就证明了 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是线性无关的. |

定理 11 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是一组线性无关的向量，那么，可以找到一组正交的向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$，使得 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_i$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_i$ $(i=1,2,\cdots,s)$ 等价.

证明：只要令

$$
\begin{cases}
\beta_1 = \alpha_1 \\
\beta_2 = \alpha_2 - \dfrac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1 \\
\beta_3 = \alpha_3 - \dfrac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1 - \dfrac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\beta_s = \alpha_s - \dfrac{(\alpha_s,\beta_1)}{(\beta_1,\beta_1)}\beta_1 - \cdots - \dfrac{(\alpha_s,\beta_{s-1})}{(\beta_{s-1},\beta_{s-1})}\beta_{s-1}
\end{cases}
$$

即可. ┆

这个证明给出了求与已知线性无关向量组等价的正交向量组的方法. 通常称为**施密特正交化方法**. 如果再将所得的正交向量组单位化,即令

$$\gamma_i = \frac{\beta_i}{|\beta_i|} \qquad (i = 1,2,\cdots,s).$$

就得到一组与 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 等价的正交单位向量组 $\gamma_1,\gamma_2,\cdots,\gamma_s$.

例 1 求与向量组

$$\alpha_1(1,1,1,1),$$

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

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

等价的正交单位向量组.

解: 令

$$\beta_1 = \alpha_1 = (1,1,1,1),$$

$$
\begin{aligned}
\beta_2 &= \alpha_2 - \frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1 \\
&= (1,-2,-3,-4) + 2(1,1,1,1) \\
&= (3,0,-1,-2)
\end{aligned}
$$

$$\beta_3 = \alpha_3 - \frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1 - \frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2$$

$$= (1,2,2,3) - 2(1,1,1,1) + \frac{5}{14}(3,0,-1,-2)$$

$$= \left(\frac{1}{14},0,\frac{-5}{14},\frac{4}{14}\right).$$

β_1,β_2,β_3 就是与 $\alpha_1,\alpha_2,\alpha_3$ 等价的正交向量组. 再令

$$\gamma_1 = \frac{\beta_1}{|\beta_1|} = \left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),$$

$$\gamma_2 = \frac{\beta_2}{|\beta_2|} = \left(\frac{3}{\sqrt{14}},0,\frac{-1}{\sqrt{14}},\frac{-2}{\sqrt{14}}\right)$$

$$= \left(\frac{3}{14}\sqrt{14},0,\frac{-1}{14}\sqrt{14},\frac{-1}{7}\sqrt{14}\right),$$

$$\gamma_3 = \frac{\beta_3}{|\beta_3|} = \left(\frac{1}{42}\sqrt{42},0,-\frac{5}{42}\sqrt{42},-\frac{2}{21}\sqrt{42}\right).$$

那么,$\gamma_1,\gamma_2,\gamma_3$ 就是与 $\alpha_1,\alpha_2,\alpha_3$ 等价的正交单位向量组.

我们在上一章中曾经证明了：任意给了一组线性无关的 n 维向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$. 总可以找到 $n-s$ 个向量 $\alpha_{s+1},\alpha_{s+2},\cdots,\alpha_n$,使得 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关. 于是以 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为行(或列)作一个矩阵,就得到一个可逆矩阵. 也就是说：任给 s 个线性无关的向量,总可以作为某个可逆矩阵的一部分行(列)向量. 对于正交单位向量组有更进一步的结论. 下面举例说明这一事实.

例 2 已知正交单位向量

$$\alpha_1 = \left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\alpha_2 = \left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right).$$

(1) 求 α_3,α_4 使 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 是正交单位向量组;

(2) 求一个以 α_3,α_4 为第 1,2 列的正交矩阵.

解:(1) 由于 α_1,α_2 是线性无关的,所以可取两个向量

$$\beta_3=(1,0,0,0),\qquad \beta_4=(0,0,1,0),$$

使 $\alpha_1,\alpha_2,\beta_3,\beta_4$ 线性无关.

将 $\alpha_1,\alpha_2,\beta_3,\beta_4$ 正交化,得一个正交向量组:

$$\alpha_1=\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\qquad \alpha_2=\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right),$$

$$\begin{aligned}\gamma_3&=\beta_3-\frac{(\beta_3,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1-\frac{(\beta_3,\alpha_2)}{(\alpha_2,\alpha_2)}\alpha_2\\&=\beta_3-\frac{1}{2}\alpha_1-\frac{1}{2}\alpha_2\\&=\left(\frac{1}{2},-\frac{1}{2},0,0\right),\end{aligned}$$

$$\begin{aligned}\gamma_4&=\beta_4-\frac{(\beta_4,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1-\frac{(\beta_4,\alpha_2)}{(\alpha_2,\alpha_2)}\alpha_2-\frac{(\beta_4,\gamma_3)}{(\gamma_3,\gamma_3)}\gamma_3\\&=\beta_4-\frac{1}{2}\alpha_1+\frac{1}{2}\alpha_2\\&=\left(0,0,\frac{1}{2},-\frac{1}{2}\right).\end{aligned}$$

再将这组向量单位化,即得到一个正交单位向量组:

$$\alpha_1=\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\qquad \alpha_2=\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right),$$

$$\alpha_3=\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},0,0\right),\qquad \alpha_4=\left(0,0,\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right),$$

其中向量 α_3,α_4 即为所求.

(2) 以 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 为列作一个矩阵 T:

$$T=\begin{bmatrix}\frac{1}{2} & \frac{1}{2} & \frac{\sqrt{2}}{2} & 0\\ \frac{1}{2} & \frac{1}{2} & -\frac{\sqrt{2}}{2} & 0\\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{\sqrt{2}}{2}\\ \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{\sqrt{2}}{2}\end{bmatrix}$$

则因 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 是正交单位向量组,所以 T 是一个正交矩阵,而且以 α_1,α_2 为第 1,2 列.

从这个例题可以看出:正交单位向量 α_1 与 α_2 在正交化及单位化的过程中都不会改变.这说明:任意 $s(s\leqslant n)$ 个 n 维正交的单位向量都可以作为某个 n 阶正交矩阵的 s 个行(或列).例题还介绍了计算这个正交矩阵的方法.这是一个很重要的结论.

习　题　3.5

1. 判断下列矩阵是否正交矩阵:

(1) $\begin{bmatrix}\frac{\sqrt{3}}{2} & -\frac{1}{2}\\ \frac{1}{2} & \frac{\sqrt{3}}{2}\end{bmatrix}$;

(2) $\begin{bmatrix}\frac{1}{9} & -\frac{8}{9} & -\frac{4}{9}\\ -\frac{8}{9} & \frac{1}{9} & -\frac{4}{9}\\ -\frac{4}{9} & -\frac{4}{9} & \frac{7}{9}\end{bmatrix}$;

$$(3)\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & \frac{\sqrt{2}}{3} \\ 0 & -\frac{2\sqrt{2}}{3} & \frac{1}{3} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & \frac{2}{3} \end{bmatrix};$$

$$(4)\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & -\frac{1}{2} \\ 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{6} & -\frac{1}{2} \\ 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

2. 试证：若 A 是正交矩阵，则 A^{T} 也是正交矩阵.

3. 试证：若 A、B 是同阶正交矩阵，则 AB 也是正交矩阵.

4. 计算$(\boldsymbol{\alpha},\boldsymbol{\beta})$：

(1) $\boldsymbol{\alpha}=(-1,0,3,-5)$，$\boldsymbol{\beta}=(4,-2,0,1)$；

(2) $\boldsymbol{\alpha}=(\sqrt{3}/2,-1/3,\sqrt{3}/4,-1)$，$\boldsymbol{\beta}=(-\sqrt{3}/2,-2,\sqrt{3},2/3)$.

5. 把下列向量单位化$\left(即求\frac{1}{|\boldsymbol{\alpha}|}\boldsymbol{\alpha}\right)$：

(1) $\boldsymbol{\alpha}=(3,0,-1,4)$；　　(2) $\boldsymbol{\alpha}=(5,1,-2,0)$.

6. 试证：如果 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 都正交，那么 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 的任一个线性组合也正交.

7. 试证：若 n 维实向量 α 与任意 n 维实向量都正交，则 α 必定是零向量.

8. 试证：若 A 是实对称矩阵，T 是正交矩阵，则 $T^{-1}AT$ 也是实对称矩阵.

9. 求与向量组(1,0,1,1),(1,1,1,−1),(1,2,3,1)等价的正交单位向量组.

10. 求两个正交矩阵,以$\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)$及$\left(\frac{1}{6},\frac{1}{6},\frac{1}{2},-\frac{5}{6}\right)$为前两行.

内 容 提 要

1. 矩阵的运算

(1) 矩阵运算的定义

1) 矩阵的加法:

$$[a_{ij}]_{sn}+[b_{ij}]_{sn}=[a_{ij}+b_{ij}]_{sn}.$$

2) 矩阵的乘法:

$$[a_{ij}]_{sn}\cdot[b_{ij}]_{nm}=[c_{ij}]_{sm},$$

其中 $c_{ij}=\sum\limits_{k=1}^{n}a_{ik}b_{kj}\quad(i=1,2,\cdots,s;\ j=1,2,\cdots,m).$

3) 矩阵与数的乘法:

$$k[a_{ij}]=[ka_{ij}].$$

4) 矩阵的转置:

$$[a_{ij}]_{sn}^{\mathrm{T}}=[a'_{ij}]_{ns},$$

其中 $\qquad a'_{ij}=a_{ji}\quad(i=1,2,\cdots,s;\ j=1,2,\cdots,n).$

(2) 矩阵运算满足的一些规律

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

$$A+B=B+A,$$

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

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

$$(B+C)A=BA+CA,$$

$$(k+l)A=kA+lA,$$

$$k(A+B)=kA+kB,$$

$$k(lA)=(kl)A,$$

$$1\cdot A=A,$$

$$k(AB)=(kA)B=A(kB),$$

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

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

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

$$(kA)^{\mathrm{T}}=kA^{\mathrm{T}}.$$

(3) 零矩阵、负矩阵及单位矩阵

$$A+O=O+A=A,$$

$$A+(-A)=O,$$

$$A-B=A+(-B),$$

对 $s\times n$ 矩阵 A，有

$$E_sA=AE_n=A.$$

2. 矩阵的分块运算.

3. 关于矩阵的秩和行列式的定理

$$\mathrm{r}(A+B)\leqslant\mathrm{r}(A)+\mathrm{r}(B),$$

$$\mathrm{r}(AB)\leqslant\min\{\mathrm{r}(A),\mathrm{r}(B)\},$$

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

当 A 可逆时，

$$\mathrm{r}(AB)=\mathrm{r}(B),$$

$$r(CA) = r(C).$$

4. 矩阵的等价

(1) 初等变换与初等矩阵

由单位矩阵经过一次初等变换所得到的矩阵称为初等矩阵. 对一个矩阵作一次初等行(列)变换,就相当于用相应的初等矩阵去左(右)乘这个矩阵.

(2) 矩阵的等价

1) 矩阵等价的定义:如果矩阵 B 可以从矩阵 A 经过一系列初等变换而得到,则称矩阵 A 与 B 是等价的.

2) 等价关系满足以下 3 个规则:

① 反身性:矩阵 A 与自身等价;

② 对称性:如果矩阵 A 与 B 等价,那么矩阵 B 也与 A 等价;

③ 传递性:如果 A 与 B 等价,B 与 C 等价,那么 A 与 C 也等价.

3) 矩阵(在等价关系下)的标准形:

每一个矩阵都可以经初等变换化为标准形:

$$\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix},$$

其中 r 是原矩阵的秩.

4) 等价的充分必要条件:

$$
\text{矩阵 } A \text{ 与 } B \text{ 等价}
\begin{cases}
\Longleftrightarrow \text{存在初等矩阵 } P_1, P_2, \cdots, P_s; \\
\quad Q_1, Q_2, \cdots, Q_t, \text{使得} \\
\quad B = P_s \cdots P_2 P_1 A Q_1 Q_2 \cdots Q_t. \\
\Longleftrightarrow \text{存在可逆矩阵 } P, Q, \text{使得} \\
\qquad B = PAQ. \\
\Longleftrightarrow A, B \text{ 的标准形相同}. \\
\Longleftrightarrow A, B \text{ 的秩相等}.
\end{cases}
$$

5. 矩阵的逆

(1) **定义**：如果有矩阵 B，使

$$AB = BA = E,$$

那么 A 称为**可逆矩阵**，B 称为 A 的**逆矩阵**，记作 A^{-1}.

(2) 逆矩阵的性质：

$$
\begin{aligned}
(A^{-1})^{-1} &= A \\
(AB)^{-1} &= B^{-1}A^{-1} \\
(A')^{-1} &= (A^{-1})'
\end{aligned}
$$

(3) 矩阵可逆的条件：

$$
\text{矩阵 } A \text{ 可逆}
\begin{cases}
\Longleftrightarrow A \text{ 非退化，即 } |A| \neq 0. \\
\Longleftrightarrow A \text{ 的 } n \text{ 个行(列) 向量线性无关}. \\
\Longleftrightarrow A \text{ 可以表成一些初等矩阵的乘积}. \\
\Longleftrightarrow A \text{ 可以经初等行变换化为单位矩阵}.
\end{cases}
$$

(4) 逆矩阵的求法：

1) 用伴随矩阵求逆：

$$A^{-1}=\frac{1}{|A|}A^{*}=\frac{1}{|A|}\begin{bmatrix}A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \cdots & \cdots & \cdots & \cdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{bmatrix}.$$

2) 用初等行变换求逆：

$$(AE)\xrightarrow{\text{初等行变换}}(EA^{-1}).$$

6. 正交矩阵

(1) 正交矩阵的定义：

实数矩阵 A，如果满足

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

即
$$A^{-1}=A^{\mathrm{T}}.$$

则称 A 为正交矩阵.

(2) n 阶实矩阵 $A=[a_{ij}]$是正交矩阵的条件：

$$A\text{ 是正交矩阵}\begin{cases}\Longleftrightarrow \displaystyle\sum_{k=1}^{n}a_{ik}a_{jk}=\delta_{ij} \qquad (i,j=1,2,\cdots,n).\\ \Longleftrightarrow \displaystyle\sum_{k=1}^{n}a_{ki}a_{kj}=\delta_{ij} \qquad (i,j=1,2,\cdots,n).\\ \Longleftrightarrow A\text{ 的 }n\text{ 个行(列)向量组成正交单位向量组.}\end{cases}$$

(3) 正交化方法

设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是一组线性无关的实向量，令

$$\beta_1=\alpha_1;$$

$$\beta_i=\alpha_i-\sum_{j=1}^{i-1}\frac{(\alpha_i,\beta_j)}{(\beta_j,\beta_j)}\beta_j \qquad (i=2,3,\cdots,s),$$

则 $\beta_1,\beta_2,\cdots,\beta_s$ 是一个正交向量组，使得 $\alpha_1,\alpha_2,\cdots,\alpha_i$ 与 $\beta_1,\beta_2,\cdots,\beta_i(i=1,2,\cdots,s)$等价.

如果再令

$$\gamma_j = \frac{\beta_j}{|\beta_j|} \quad (j = 1,2,\cdots,n).$$

则 $\gamma_1,\gamma_2,\cdots,\gamma_n$ 是使得 $\alpha_1,\cdots,\alpha_i$ 与 $\gamma_1,\cdots,\gamma_i (i=1,2,\cdots,s)$ 等价的正交单位向量组.

复习题 3

1. 求满足下列条件的矩阵 A：

(1) $\begin{bmatrix} 1 & 0 & -2 \\ -3 & 4 & -1 \\ 2 & 1 & 3 \end{bmatrix} A = \begin{bmatrix} 5 & -1 \\ -2 & 3 \\ 1 & 4 \end{bmatrix}$；

(2) $A \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & -1 \\ 2 & 1 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 6 \\ 2 & 1 & 0 \end{bmatrix}$；

(3) $\begin{bmatrix} 1 & -2 & 0 \\ 4 & -2 & -1 \\ -3 & 1 & 2 \end{bmatrix} A \begin{bmatrix} 3 & -1 & 2 \\ 1 & 0 & -1 \\ -2 & 1 & 4 \end{bmatrix}$

$= \begin{bmatrix} 5 & 0 & -1 \\ 1 & -3 & 0 \\ -2 & 1 & 3 \end{bmatrix}$；

(4) $\begin{bmatrix} 1 & 2 & 3 \\ -2 & 1 & -1 \\ 3 & 2 & 1 \end{bmatrix} A + \begin{bmatrix} 2 & -1 & 0 \\ -1 & 3 & 1 \\ 2 & 1 & 1 \end{bmatrix}$

$= 3A + \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 3 \\ 3 & 2 & 1 \end{bmatrix}$.

2. 设

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

其中 $a_1,a_2,\cdots,a_n$ 两两不同.求证与 A 可交换的矩阵一定是对角矩阵.

3. 证明与一切 n 阶矩阵都可交换的矩阵一定是 n 阶数量矩阵.

4. 矩阵 A 如果满足 $A^T=A(A^T=-A)$,则称 A 为对称(反对称)矩阵.证明:

(1) 两个对称(反对称)矩阵的和也是对称(反对称)矩阵;

(2) 一个数与一个对称(反对称)矩阵的乘积也是对称(反对称)矩阵.

5. 设 A,B 是同阶对称矩阵,则 AB 是对称矩阵的充分必要条件是:A 与 B 可交换.

6. 设 A,B 是两个反对称矩阵.试证:

(1) A^2 是对称矩阵;

(2) $AB-BA$ 是反对称矩阵;

(3) AB 是对称矩阵的充要条件是:$AB=BA$.

7. (1) 试证:任一 n 阶矩阵都可表为一个对称矩阵与一个反对称矩阵之和;

(2) 把矩阵

$$A=\begin{bmatrix}1 & 2 & 3\\ 0 & 2 & 4\\ 0 & 0 & 5\end{bmatrix}$$

表成一个对称矩阵与一个反对称矩阵之和.

8. (1) 如果 A 是实对称矩阵,且 $A^2=O$,则 $A=O$;

(2) 举例说明在一般情况下,由 $A^2=O$ 推不出 $A=O$.

9. 设 A 是一个 n 阶方阵,如果对任一 n 维向量 X 都有 $AX=0$,那么 $A=O$.

10. 设 A 是一个 n 阶矩阵($n\geqslant 2$). 试证:

$$\mathrm{r}(A^*)=\begin{cases} n, & \text{当 } \mathrm{r}(A)=n; \\ 1, & \text{当 } \mathrm{r}(A)=n-1; \\ 0, & \text{当 } \mathrm{r}(A)<n-1. \end{cases}$$

11. 设 A 是一个 n 阶矩阵($n\geqslant 2$). 求证:

$$|A^*|=|A|^{n-1}.$$

12. 设 A 是一个 n 阶矩阵($n>2$). 求证:

$$(A^*)^*=|A|^{n-2}A.$$

13. 试证: 如果 $A^2=E$, 那么

$$\mathrm{r}(A+E)+\mathrm{r}(A-E)=n.$$

14. 试证: 如果 $A^2=A$, 那么

$$\mathrm{r}(A)+\mathrm{r}(A-E)=n.$$

15. 设 A 是一个 n 阶矩阵, $\mathrm{r}(A)=1$. 试证:

(1) A 可表成:

$$A=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}(b_1\ b_2\ \cdots\ b_n);$$

(2) $A^2=kA$.

16. 设 α 是一个 n 维实向量. 已知 α 是一个单位向量, 令矩阵 $T=E-2\alpha^{\mathrm{T}}\alpha$. 试证: T 是一个对称的正交矩阵.

17. 设

$$A=\begin{bmatrix} 0 & 10 & 6 \\ 1 & -3 & -3 \\ -2 & 10 & 8 \end{bmatrix}, \quad P=\begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}.$$

(1) 求 $P^{-1}AP$;

(2) 求 A^k(k 为正整数).

18. (1) 证明：任一 n 阶实可逆矩阵 A 都可表成一个正交矩阵 T 与一个上三角矩阵 P 的乘积：

$$A = TP;$$

(2) 设

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

求正交矩阵 T 及上三角矩阵 P 使 $A=TP$.

第 4 章　矩阵的对角化问题

在上一章中我们定义了矩阵的等价. 从那里看到，有些问题，特别是矩阵的秩的问题可以应用矩阵在等价关系下的标准形来解决.

这一章讨论矩阵的相似问题. 矩阵 $X^{-1}AX$ 与 A 称为相似. 矩阵的相似关系可以用来简化运算. 例如，如果 $B=X^{-1}AX$，那么 $B^k=X^{-1}A^kX$，$A^k=XB^kX^{-1}$. 因此，当 B 比较简单时，可以利用 B^k 来计算 A^k. 相似矩阵还可以用来简化线性方程组及微分方程组. 当然，相似矩阵的应用还远不止于此.

那么，给了一个矩阵 A，在和 A 相似的矩阵之中，最简单的矩阵是什么样的呢？怎样去找可逆矩阵 X，使 $X^{-1}AX$ 最简单呢？这就是求矩阵的标准形问题，也就是这一章要讨论的主要内容.

矩阵在相似关系下的标准形问题不仅可以简化运算，在理论上也是非常重要的. 在其他学科中也有极广泛的应用. 另外，本章还将介绍的特征值和特征向量等概念，也是非常重要的，应用也很广泛.

本章将在某个确定的数域 **P** 上进行讨论，所讨论的矩阵都是数域 **P** 上的矩阵.

4.1　相似矩阵

定义 1　设 A,B 是两个同阶方阵，如果存在可逆矩阵 X，使得 $B=X^{-1}AX$，就说 **A 相似于 B**，记作 $A\sim B$.

例如，因为

$$\begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 10 & 6 \\ 1 & -3 & -3 \\ -2 & 10 & 8 \end{bmatrix} \begin{bmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$$

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

所以
$$\begin{bmatrix} 0 & 10 & 6 \\ 1 & -3 & -3 \\ -2 & 10 & 8 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

“相似”是矩阵之间的一种关系，这种关系具有下面 3 个性质：

1. 反身性：$A \sim A$.

这是因为 $A = E^{-1}AE$.

2. 对称性：如果 $A \sim B$，那么 $B \sim A$.

这是因为，如果 $A \sim B$，那么有 X，使 $B = X^{-1}AX$. 令 $Y = X^{-1}$，就有 $A = XBX^{-1} = Y^{-1}BY$，所以 $B \sim A$.

3. 传递性：如果 $A \sim B$，$B \sim C$，那么 $A \sim C$.

这是因为，已知 $A \sim B$，$B \sim C$，所以有 X 和 Y，使 $B = X^{-1}AX$，$C = Y^{-1}BY$，于是

$$C = Y^{-1}(X^{-1}AX)Y = (XY)^{-1}A(XY),$$

因此 $A \sim C$.

根据相似关系的对称性，当 $A \sim B$ 时，既可以说 A 与 B 相似，也可以说 B 与 A 相似.

相似矩阵还有下面的一些性质：

1. 相似矩阵有相同的行列式.

证明：设 $A \sim B$，那么有 X，使 $B = X^{-1}AX$. 于是 $|B| = |X^{-1}AX| = |X^{-1}||A||X| = |X|^{-1}|A||X| = |A|$.

因为矩阵可逆的充分必要条件是它的行列式不等于零. 由此可知：

2. 相似矩阵或者同时可逆，或者都不可逆. 而且，如果 $B=X^{-1}AX$，那么，当它们可逆时，它们的逆矩阵也相似，即 $B^{-1}=X^{-1}A^{-1}X$.

3. 如果 $B_1=X^{-1}A_1X$，$B_2=X^{-1}A_2X$，那么

$$B_1+B_2=X^{-1}(A_1+A_2)X;$$

$$B_1B_2=X^{-1}(A_1A_2)X;$$

$$kB_1=X^{-1}(kA_1)X.$$

因此，如果 $B=X^{-1}AX$，$f(x)$是系数在数域 **P** 中的一个多项式，那么

$$f(B)=X^{-1}f(A)X.$$

这 4 个等式都可以根据矩阵运算的规律直接验证. 请读者自己证明一下. 根据相似矩阵的这些性质，就可以简化矩阵的运算. 问题是：给了一个矩阵 A，如何去找矩阵 X，使 $X^{-1}AX$ 最简单？最简单的矩阵当然是数量矩阵 kE. 但是与数量矩阵 kE 相似的矩阵只有它自己：

$$X^{-1}(kE)X=kE.$$

由此可知，不是每个矩阵都可以与某个数量矩阵相似的. 于是，退而求其次. 比数量矩阵稍为复杂些的是对角矩阵，那么，是不是任一个矩阵都可与某个对角矩阵相似呢？现在首先来分析矩阵 A 与对角矩阵相似的条件.

为了简便起见，以后用$[a_1,a_2,\cdots,a_n]$表示对角矩阵：

$$\begin{bmatrix} a_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_n \end{bmatrix};$$

用$[A_1,A_2,\cdots,A_l]$表示准对角矩阵：

$$\begin{bmatrix} A_1 & O & O & \cdots & O \\ O & A_2 & O & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & O & O & \cdots & A_s \end{bmatrix}.$$

设 A 是一个 n 阶矩阵：

$$A = (a_{ij}),$$

X 是一个 n 阶可逆矩阵：

$$X = (x_{ij}).$$

并设

$$X^{-1}AX = [\lambda_1, \lambda_2, \cdots, \lambda_n]$$

是对角矩阵.

用 $X_1, X_2, \cdots, X_n$ 表示 X 的 n 个列向量，并将 X 表成分块矩阵：

$$X = [X_1 \quad X_2 \quad \cdots \quad X_n].$$

于是从 $X^{-1}AX=[\lambda_1, \lambda_2, \cdots, \lambda_n]$得到

$$AX = X[\lambda_1, \lambda_2, \cdots, \lambda_n],$$

即 $$A[X_1 \quad X_2 \quad \cdots \quad X_n] = [X_1 \quad X_2 \quad \cdots \quad X_n]\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & 0 & \\ & 0 & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

$$= [\lambda_1 X_1 \quad \lambda_2 X_2 \quad \cdots \quad \lambda_n X_n].$$

等式左右两边的列向量应该依次相等，所以

$$AX_i = \lambda_i X_i \quad (i = 1, 2, \cdots, n).$$

这说明：如果可逆矩阵 X 使 $X^{-1}AX$ 为对角矩阵，那么 X 的列向量必须满足上述条件. 满足这个条件的向量就是特征向量. 下一节将讨论特征向量的性质及求法，并在这个基础上进一步解决矩阵 A 能否相似于对角矩阵的问题.

习 题 4.1

1. 试证：如果 $A\sim B$，那么 $A^T\sim B^T$.

2. 试证：如果 A 与 B 可交换，那么 $X^{-1}AX$ 与 $X^{-1}BX$ 也可交换.

3. 设 $A\sim B$，$C\sim D$. 试证：

$$\begin{bmatrix} A & O \\ O & C \end{bmatrix} \sim \begin{bmatrix} B & O \\ O & D \end{bmatrix}.$$

4. 试证：如果 A 可逆，那么 $AB\sim BA$.

5. 矩阵 A 称作幂零的，如果有正整数 k 使 $A^k=O$；A 称作幂等的，如果 $A^2=A$，从而对正整数 k 都有 $A^k=A$；A 称为幺幂的，如果有正整数 k 使 $A^k\geqslant E$. 试证：

(1) 与幂零矩阵相似的矩阵都是幂零矩阵；

(2) 与幂等矩阵相似的矩阵都是幂等矩阵；

(3) 与幺幂矩阵相似的矩阵都是幺幂矩阵.

4.2 特征值与特征向量

定义 2 设 A 是数域 $\mathbf{P}$ 上一个 n 阶矩阵，λ_0 是 $\mathbf{P}$ 中一个数，如果有 $\mathbf{P}$ 上非零列向量（即 $n\times 1$ 矩阵）$\boldsymbol{\alpha}$，使得

$$A\boldsymbol{\alpha} = \lambda_0\boldsymbol{\alpha}, \tag{1}$$

就称 λ_0 是 A 的一个**特征值**，$\boldsymbol{\alpha}$ 是 A 的属于特征值 λ_0 的**特征向量**，简称特征向量.

例如，因为

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

所以，2 是矩阵

$$\begin{bmatrix} 1 & -2 & 2 \\ -2 & -2 & 4 \\ -2 & 4 & -2 \end{bmatrix}$$

的一个特征值，$\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$是属于特征值 2 的一个特征向量.

定理 1 n 阶矩阵 A 与一个对角矩阵相似的充分必要条件是：A 有 n 个线性无关的特征向量.

证明：必要性：如果 A 与对角矩阵$[\lambda_1,\lambda_2,\cdots,\lambda_n]$相似，那么有可逆矩阵 X，使 $X^{-1}AX=[\lambda_1,\lambda_2,\cdots,\lambda_n]$. X 的列向量 $X_1,X_2,\cdots,X_n$ 满足

$$AX_i=\lambda_i X_i \quad (i=1,2,\cdots,n).$$

X_i 当然都不是零向量，所以 $X_i(i=1,2,\cdots,n)$都是 A 的特征向量. 而且因为 X 可逆，所以这 n 个特征向量是线性无关的.

充分性：如果 A 有 n 个线性无关的特征向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 它们所对应的特征值依次为 $\lambda_1,\lambda_2,\cdots,\lambda_n$：

$$A\alpha_i=\lambda_i\alpha_i \quad (i=1,2,\cdots,n).$$

以 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为列向量作一个矩阵 X：

$$X=[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n],$$

因为 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是线性无关的，所以 X 是可逆矩阵. 而且

$$AX=X[\lambda_1,\lambda_2,\cdots,\lambda_n],$$

即

$$X^{-1}AX=[\lambda_1,\lambda_2,\cdots,\lambda_n].$$

所以 A 与一个对角矩阵相似. |

因此，是否每一个矩阵都与某一个对角矩阵相似的问题就归结为：是否每一个 n 阶矩阵都有 n 个线性无关的特征向量的问题了. 为了解决这个问题，下面先讨论特征向量的求法.

设
$$\alpha=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}\neq 0$$
是矩阵
$$A=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{bmatrix}$$
的属于特征值 λ_0 的特征向量，那么
$$A\alpha=\lambda_0\alpha.$$
具体写出来，就是
$$\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{bmatrix}\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}=\lambda_0\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}.$$
将等式两端乘开，比较分量，得：
$$\begin{cases}a_{11}c_1+a_{12}c_2+\cdots+a_{1n}c_n=\lambda_0c_1,\\a_{21}c_1+a_{22}c_2+\cdots+a_{2n}c_n=\lambda_0c_2,\\\cdots\cdots\cdots\cdots\cdots\cdots\\a_{n1}c_1+a_{n2}c_2+\cdots+a_{nn}c_n=\lambda_0c_n.\end{cases}$$
移项，得：
$$\begin{cases}(\lambda_0-a_{11})c_1-a_{12}c_2-\cdots-a_{1n}c_n=0,\\-a_{21}c_1+(\lambda_0-a_{22})c_2-\cdots-a_{2n}c_n=0,\\\cdots\cdots\cdots\cdots\cdots\cdots\\-a_{n1}c_1-a_{n2}c_2-\cdots+(\lambda_0-a_{nn})c_n=0.\end{cases}$$
这说明$(c_1,c_2,\cdots,c_n)$是齐次线性方程组

$$
\begin{cases}
(\lambda_0 - a_{11})x_1 - a_{12}x_2 - \cdots - a_{1n}x_n = 0, \\
-a_{21}x_1 + (\lambda_0 - a_{22})x_2 - \cdots - a_{2n}x_n = 0, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
-a_{n1}x_1 - a_{n2}x_2 + \cdots + (\lambda_0 - a_{nn})x_n = 0,
\end{cases} \tag{2}
$$

的一个解.这个齐次方程组既然有非零解，所以它的系数行列式等于零：

$$
\begin{vmatrix}
\lambda_0 - a_{11} & -a_{12} & \cdots & -a_{1n} \\
-a_{21} & \lambda_0 - a_{22} & \cdots & -a_{2n} \\
\cdots & \cdots & \cdots & \cdots \\
-a_{n1} & -a_{n2} & \cdots & \lambda_0 - a_{nn}
\end{vmatrix} = 0,
$$

即

$$|\lambda_0 E - A| = 0.$$

定义 3 A 是一个 n 阶矩阵，λ 是一个未知量，矩阵 $\lambda E - A$ 称为 A 的**特征矩阵**，它的行列式

$$
|\lambda E - A| = \begin{vmatrix}
\lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\
-a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\
\cdots & \cdots & \cdots & \cdots \\
-a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn}
\end{vmatrix}
$$

是 λ 的一个多项式，称为 **A 的特征多项式**.

上面的分析说明：如果 λ_0 是矩阵 A 的一个特征值，那么 λ_0 一定是 A 的特征多项式的一个根；反过来，如果 λ_0 是 A 的特征多项式的一个根，即 $|\lambda_0 E - A| = 0$，那么齐次线性方程组(2)就有非零解，因此，λ_0 是 A 的特征值，而方程组(2)的每一个非零解都是 A 的属于 λ_0 的特征向量.这就是说：矩阵 A 的特征值就是 A 的特征多项式的根，所以特征值也叫做**特征根**.而且也可以用此作为特征值的定义.

归纳以上讨论，可总结出矩阵 A 的特征值和特征向量的求法：

(1) 计算 A 的特征多项式 $f(\lambda) = |\lambda E - A|$；

(2) 求出 $f(\lambda)$在数域 P 中的全部根，就是 A 的全部特征值；

(3) 对于每个特征值 λ_0，求出齐次线性方程组(2)的非零解，就是属于 λ_0 的特征向量.

例 1 求矩阵 A 的特征值与特征向量：

$$A=\begin{bmatrix}1 & -2 & 2\\ -2 & -2 & 4\\ 2 & 4 & -2\end{bmatrix}.$$

解：先求 A 的特征多项式

$$|\lambda E-A|=\begin{vmatrix}\lambda-1 & 2 & -2\\ 2 & \lambda+2 & -4\\ -2 & -4 & \lambda+2\end{vmatrix}=(\lambda-2)^2(\lambda+7).$$

所以 A 的特征值为 2(2 重)、-7.

把 $\lambda=2$ 代入齐次线性方程组(2)，得

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

化简，即得：

$$x_1+2x_2-2x_3=0.$$

它的一个基础解系是

$$\begin{bmatrix}2\\0\\1\end{bmatrix},\quad\begin{bmatrix}0\\1\\1\end{bmatrix}.$$

把 $\lambda=-7$ 代入(2)，得

$$\begin{cases}-8x_1-2x_2-2x_3=0,\\ 2x_1-5x_2-4x_3=0,\\ -2x_1-4x_2-5x_3=0.\end{cases}$$

化简，得：

$$\begin{cases} 2x_1 - 5x_2 - 4x_3 = 0 \\ \qquad\quad\; x_2 + x_3 = 0. \end{cases}$$

它的一个基础解系是

$$\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}.$$

因此,A 的特征值为 2,-7. 属于 2 的特征向量是

$$k_1\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + k_2\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad (k_1, k_2 \text{ 不全为 } 0);$$

属于 -7 的特征向量是

$$k\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \quad (k \neq 0).$$

例 2 求矩阵 B 的特征值与特征向量:

$$B = \begin{bmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{bmatrix}.$$

解:B 的特征多项式为

$$|\lambda E - B| = \begin{vmatrix} \lambda - 3 & -1 & 0 \\ 4 & \lambda + 1 & 0 \\ -4 & 8 & \lambda + 2 \end{vmatrix} = (\lambda - 1)^2(\lambda + 2).$$

所以 B 的特征值为 1(2 重),-2.

把 $\lambda=1$ 代入齐次线性方程组(2),得

$$\begin{cases} -2x_1 - x_2 = 0, \\ 4x_1 + 2x_2 = 0, \\ -4x_1 + 8x_2 + 3x_3 = 0. \end{cases}$$

其基础解系为 $\begin{bmatrix} 3 \\ -6 \\ 20 \end{bmatrix}$;

把 $\lambda=-2$ 代入(2),得

$$\begin{cases} -5x_1 - x_2 = 0 \\ 4x_1 - x_2 = 0 \\ -4x_1 + 8x_2 = 0. \end{cases}$$

其基础解系为 $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

所以,B 的特征值为 1,-2. 属于 1 的特征向量为:

$$k\begin{bmatrix} 3 \\ -6 \\ 20 \end{bmatrix} \quad (k \neq 0);$$

属于 -2 的特征向量为:

$$l\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (l \neq 0).$$

因为凡是 A 的属于 λ_0 的特征向量都是齐次线性方程组(2)的解;反过来,凡是方程组(2)的非零解一定都是 A 的属于 λ_0 的特征向量,所以,为了求 A 的属于 λ_0 的全部特征向量,只需找出方程组(2)的一个基础解系,设为 $\alpha_1,\alpha_2,\cdots,\alpha_s$,那么 A 的属于 λ_0 的全部特征向量就是

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s.$$

其中 $k_1,k_2,\cdots,k_s$ 可以取数域 P 中任意的数. 需要注意的是:因为特征向量是非零向量,所以 $k_1,k_2,\cdots,k_s$ 必须不全为零.

下面对矩阵的特征多项式作几点说明. 首先来证明下述定理.

定理 2 相似矩阵有相同的特征多项式.

证明：设 $A \sim B$，那么有可逆矩阵 X，使

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

于是 A,B 的特征矩阵有下述关系：

$$\lambda E - B = \lambda E - X^{-1}AX = X^{-1}(\lambda E - A)X.$$

求等式两端的行列式，得

$$\begin{aligned}|\lambda E - B| &= |X^{-1}(\lambda E - A)X| = |X^{-1}||\lambda E - A||X| \\ &= |\lambda E - A|.\end{aligned}$$

这就证明了 A,B 有相同的特征多项式. |

但是，特征多项式相等的矩阵却并不一定相似，例如矩阵

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

的特征多项式都等于 $(\lambda-1)^2$，但它们显然不相似，因为与 A 相似的矩阵只有它自己.

由于矩阵的特征值就是它的特征多项式的根，因此，从定理 2 可得出下述推论.

推论 相似矩阵有相同的特征值.

如果 A 是一个 n 阶矩阵，那么它的特征多项式一定是一个 n 次多项式. 事实上，如果

$$A = [a_{ij}].$$

那么

$$f(\lambda) = |\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}.$$

其中包含 λ 的元素只有 $\lambda-a_{11},\lambda-a_{22},\cdots,\lambda-a_{nn}$. 由于这些元素位于不同行不同列，所以

$$f(\lambda) = (\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn}) + \cdots.$$

其他的项至少包含一个因子 $(-a_{ij})$. 如果包含 $-a_{ij}$，那么就不能包

含 $\lambda-a_{ii}$(因为这个元素与$-a_{ij}$在同一行),也不能包含 $\lambda-a_{jj}$(因为这个元素与$-a_{ij}$在同一列),所以其他的项最多是 $n-2$ 次. 于是

$$f(\lambda)=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+c_{n-2}\lambda^{n-2}$$
$$+\cdots+c_1\lambda+c_0.$$

这说明:$f(\lambda)$是 λ 的一个 n 次多项式,首项系数是 1,λ^{n-1}的系数是

$$-(a_{11}+a_{22}+\cdots+a_{nn}).$$

其中 $a_{11}+a_{22}+\cdots+a_{nn}$是 A 的主对角线上的元素之和. 称为 A 的迹.

定义 4 设 $A=(a_{ij})$是一个 n 阶矩阵,A 的对角线元素之和称为 A 的**迹**,记 $\mathrm{tr}A$:

$$\mathrm{tr}A=a_{11}+a_{22}+\cdots+a_{nn}.$$

由于 n 阶矩阵的特征多项式中 λ^{n-1}的系数等于这个矩阵的迹的负值. 所以当 A 与 B 是两个相似矩阵时,它们的特征多项式相等,因而它们的迹也相等. 这就是说:**相似矩阵有相同的迹**.

A 的特征多项式 $f(\lambda)$的常数项可以很容易地计算出来:在等式

$$\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}=\lambda^n-(a_{11}+\cdots+a_{nn})\lambda^{n-1}$$
$$+\cdots+c_1\lambda+c_0$$

中,两边用 $\lambda=0$ 代入,即得

$$\begin{vmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & -a_{nn} \end{vmatrix}=c_0.$$

所以 $$c_0=(-1)^n|A|.$$

因为矩阵 A 的特征值是 A 的特征多项式的根，而多项式的求根问题是与所考虑的数域有关的，因此，矩阵 A 的特征值与特征向量的存在也与所考虑的数域有关，在不同的数域中，可以得到不同的结果. 下面举例说明.

例 3 求矩阵 A 的特征值与特征向量：

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

解：A 的特征多项式为

$$f(\lambda)=|\lambda E-A|=\begin{vmatrix}\lambda-5 & -6 & 3\\ 1 & \lambda & -1\\ -1 & -2 & \lambda+1\end{vmatrix}$$

$$=(\lambda-2)(\lambda^2-2\lambda-2).$$

如果在有理数域中考虑，那么 A 只有一个特征值 $\lambda=2$. 把 $\lambda=2$ 代入方程组(2)中，得

$$\begin{cases}-3x_1-6x_2+3x_3=0,\\ x_1+2x_2-x_3=0,\\ -x_1-2x_2+3x_3=0.\end{cases}$$

求得一般解为：

$$x_1=-2x_2$$

$$x_3=0$$

其中 x_2 为自由未知量.

其基础解系为：$\begin{bmatrix}2\\ -1\\ 0\end{bmatrix}$.

所以 A 的全部特征向量为：

$$k\begin{bmatrix}2\\-1\\0\end{bmatrix}\quad (k\text{ 为不等于零的有理数}).$$

如果在实数域中考虑，那么除去 $\lambda=2$ 外，A 还有两个特征值 $\lambda=1\pm\sqrt{3}$，把 $\lambda=1\pm\sqrt{3}$ 分别代入方程组(2)中，得

$$\begin{cases}(-4+\sqrt{3})x_1-6x_2+3x_3=0,\\ x_1+(1+\sqrt{3})x_2-x_3=0,\\ -x_1-2x_2+(2+\sqrt{3})x_3=0,\end{cases}$$

和

$$\begin{cases}(-4-\sqrt{3})x_1-6x_2+3x_3=0,\\ x_1+(1-\sqrt{3})x_2-x_3=0,\\ -x_1-2x_2+(2-\sqrt{3})x_3=0.\end{cases}$$

它们的基础解系分别为：

$$\begin{bmatrix}6+3\sqrt{3}\\-2-\sqrt{3}\\1\end{bmatrix}\text{和}\begin{bmatrix}6-3\sqrt{3}\\-2+\sqrt{3}\\1\end{bmatrix}.$$

所以 A 的全部特征向量为

$$k\begin{bmatrix}2\\-1\\0\end{bmatrix},\quad l\begin{bmatrix}6+3\sqrt{3}\\-2-\sqrt{3}\\1\end{bmatrix},\quad m\begin{bmatrix}6-3\sqrt{3}\\-2+\sqrt{3}\\1\end{bmatrix}.$$

其中 k,l,m 都是非零实数.

因为复系数多项式在复数域中一定有根，所以在复数域中考虑时，任一个矩阵一定都有特征向量.

最后，我们指出如下一个重要结论而不加证明.

定理 3(汉密尔顿-凯莱(Hamilton-Caylay)定理) 设 A 是一个 n 阶矩阵，$f(\lambda)=|\lambda E-A|$ 是 A 的特征多项式，那么 $f(A)=O$.

习 题 4.2

1. 求下列复系数矩阵的特征值和特征向量：

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

(3) $\begin{bmatrix}3 & 7 & -3\\ -2 & -5 & 2\\ -4 & -10 & 3\end{bmatrix}$；　　(4) $\begin{bmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}$；

(5) $\begin{bmatrix}0 & 2 & 1\\ -2 & 0 & 3\\ -1 & -3 & 0\end{bmatrix}$.

2. 试证：

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

(2) $\operatorname{tr}(kA)=k\operatorname{tr}A$；

(3) $\operatorname{tr}(AB)=\operatorname{tr}(BA)$.

3. 设 α 是 A 的属于特征值 λ_0 的特征向量. 求证：

(1) 对任意数非零 k，α 是 kA 的属于特征值 $k\lambda_0$ 的特征向量；

(2) 对正整数 m，α 是 A^m 的属于特征值 λ_0^m 的特征向量；

(3) 如果 A 可逆，则 $\lambda_0\neq 0$，且 α 是 A^{-1}的属于 λ_0^{-1} 的特征向量.

4. 试证：

(1) 幂零矩阵的特征值都等于 0；

(2) 幂等矩阵的特征值等于 0 或 1；

(3) 幺幂矩阵的特征值都是单位根.（如果复数 λ_0 满足 $\lambda_0^k=1$，(k 是正整数)，则称 λ_0 是一个 m 次单位根）

5. 证明：如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是矩阵 A 的属于特征值 λ_0 的一组线性无关的特征向量；$k_1,k_2,\cdots,k_s$ 是一组不全为零的数，则 $k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s$ 也是 A 的属于特征值 λ_0 的特征向量.

4.3 矩阵可对角化的条件

这一节讨论矩阵与某一个对角矩阵相似的条件. 如果 A 与一个对角矩阵相似,就说 A 对角化. 定理 1 说明了:n 阶矩阵 A 可对角化的充分必要条件是:A 有 n 个线性无关的特征向量. 然而,上节的例题和习题告诉我们:并不是所有的 n 阶矩阵都有 n 个线性无关的特征向量. 因而并不是每一个矩阵都可以对角化,现在引用上节中的例 1 和例 2 来核对这个论断. 例 1 中的 3 阶矩阵 A 有 3 个线性无关的特征向量:

$$\begin{bmatrix}2\\0\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix} \quad 与 \quad \begin{bmatrix}1\\2\\-2\end{bmatrix},$$

以它们为列作一个矩阵

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

那么,$X^{-1}AX$ 就一定是对角矩阵:

$$X^{-1}AX=\begin{bmatrix}2 & 0 & 0\\0 & 2 & 0\\0 & 0 & -7\end{bmatrix}.$$

读者可以验证一下这个等式.

但是,例 2 中的 3 阶矩阵 B 只有 2 个线性无关的特征向量,所以 B 一定不能与一个对角矩阵相似. 这个例子明确地回答了在第一节中提出的问题. 就是说,并不是任一个矩阵 A 都可找到可逆矩阵 X,使 $X^{-1}AX$ 成为对角矩阵. 或者简单地说,并不是任意矩阵都可以化为对角形.

如果 A 可以化为对角形 B:$B=X^{-1}AX$. 那么,B 与 A 有相同

的特征值. 而 B 的特征值就是它的对角线上的元素，所以 B 除了对角线上的元素的排列次序可以不同外，是唯一确定的. 但是，因为特征向量不是唯一的，所以 X 的取法有无穷多种.

又因为特征值、特征向量的存在与所考虑的数域有关，所以一个矩阵能不能对角化，也与所考虑的数域有关. 例如上节例 3 中的矩阵 A，如果把它考虑为有理系数的矩阵，那么它只有一个线性无关的特征向量，所以不能对角化. 即，不能找到可逆的系数为有理数的矩阵 X，使 $X^{-1}AX$ 为对角矩阵. 但是，如果把 A 看成一个实数域上的矩阵，那么就可以找到可逆矩阵

$$X=\begin{bmatrix}2 & 6+3\sqrt{3} & 6-3\sqrt{3}\\ -1 & -2-\sqrt{3} & -2+\sqrt{3}\\ 0 & 1 & 1\end{bmatrix},$$

使 $X^{-1}AX$ 为对角矩阵：

$$X^{-1}AX=\begin{bmatrix}2 & 0 & 0\\ 0 & 1+\sqrt{3} & 0\\ 0 & 0 & 1-\sqrt{3}\end{bmatrix}.$$

在应用时常常要求特征向量是线性无关的. 对于同一个特征值，在求特征向量的时候，只要取一个基础解系，那么对应的特征向量一定包含了尽可能多的线性无关的特征向量. 对于不同的特征值，当分别取了线性无关的特征向量(例如，可以取基础解系)以后，把它们放在一起合成一个向量组，是不是还线性无关呢? 答案是肯定的. 下面的定理就证明了这一点.

定理 4 如果 $\lambda_1,\lambda_2,\cdots,\lambda_t$ 是矩阵 A 的不同的特征值；$\alpha_{i1},\alpha_{i2},\cdots,\alpha_{is_i}$，是 A 的属于 λ_i 的线性无关的特征向量($i=1,2,\cdots,t$). 那么向量组 $\alpha_{11},\alpha_{12},\cdots,\alpha_{1s_1},\cdots,\alpha_{t1},\alpha_{t2},\cdots,\alpha_{ts_t}$ 也是线性无关的.

证明：对特征值的个数 t 用归纳法.

当 $t=1$ 时，由假设知 $\alpha_{11},\alpha_{12},\cdots,\alpha_{1s_1}$ 是线性无关的. 设 $t=k$ 时

结论成立，即 $\alpha_{11},\alpha_{12},\cdots,\alpha_{1s_1},\cdots,\alpha_{k1},\alpha_{k2},\cdots,\alpha_{ks_k}$是线性无关的，下面证 $t=k+1$ 时结论也成立.

假设有关系式

$$\begin{aligned} a_{11}\alpha_{11}+\cdots+a_{1s_1}\alpha_{1s_1}+\cdots+a_{k1}\alpha_{k1}+\cdots+a_{ks_k}\alpha_{ks_k} \\ +a_{k+1,1}\alpha_{k+1,1}+\cdots+a_{k+1,s_{k+1}}\alpha_{k+1,s_{k+1}}=0 \end{aligned} \tag{1}$$

成立. 在等式两边乘以 λ_{k+1}，得：

$$\begin{aligned} & a_{11}\lambda_{k+1}\alpha_{11}+\cdots+a_{1s_1}\lambda_{k+1}\alpha_{1s_1}+\cdots+a_{k1}\lambda_{k-1}\alpha_{k1}+\cdots \\ & +a_{ks_k}\lambda_{k+1}\alpha_{ks_k}+a_{k+1,1}\lambda_{k+1}\alpha_{k+1,1}+\cdots \\ & +a_{k+1,s_{k+1}}\lambda_{k+1}\alpha_{k+1,s_{k+1}}=0. \end{aligned} \tag{2}$$

在(1)式两端同用矩阵 A 左乘，得

$$\begin{aligned} a_{11}\lambda_1\alpha_{11}+\cdots+a_{1s_1}\lambda_1\alpha_{1s_1}+\cdots+a_{k1}\lambda_k\alpha_{k1}+\cdots+a_{ks_k}\lambda_k\alpha_{ks_k} \\ +a_{k+1,1}\lambda_{k+1}\alpha_{k+1,1}+\cdots+a_{k+1,s_{k+1}}\lambda_{k+1}\alpha_{k+1,s_{k+1}}=0. \end{aligned} \tag{3}$$

从(2)式减去(3)式，得

$$\begin{aligned} & a_{11}(\lambda_{k+1}-\lambda_1)\alpha_{11}+\cdots+a_{1s_1}(\lambda_{k+1}-\lambda_1)\alpha_{1s_1}+\cdots \\ & +a_{k1}(\lambda_{k+1}-\lambda_k)\alpha_{k1}+\cdots+a_{ks_k}(\lambda_{k+1}-\lambda_k)\alpha_{ks_k}=0. \end{aligned}$$

根据归纳法假设：$\alpha_{11},\cdots,\alpha_{1s_1},\cdots,\alpha_{k1},\cdots,\alpha_{ks_k}$是线性无关的，于是

$$a_{ij}(\lambda_{k+1}-\lambda_i)=0 \quad (i=1,2,\cdots,k;\ j=1,2,\cdots,s_i).$$

但是 $\lambda_{k+1}-\lambda_i\neq 0(i=1,2,\cdots,k)$，所以

$$a_{ij}=0 \quad (i=1,2,\cdots,k;\ j=1,2,\cdots,s_i).$$

于是(1)式变成

$$a_{k+1,1}\alpha_{k+1,1}+\cdots+a_{k+1,s_{k-1}}\alpha_{k+1,s_{k+1}}=0.$$

由题设：$\alpha_{k+1,1},\cdots,\alpha_{k+1,s_{k+1}}$是线性无关的，所以

$$a_{k+1,j}=0 \quad (j=1,2,\cdots,s_{k+1}).$$

这就证明了 $\alpha_{11},\cdots,\alpha_{1s_1},\cdots,\alpha_{k+1,1},\cdots,\alpha_{k+1,s_{k+1}}$是线性无关的.

根据归纳法原理，定理得证. |

在定理 4 中如果每个 λ_i 取一个特征向量 $\alpha_i(i=1,\cdots,t)$，那

么，$\alpha_1,\alpha_2,\cdots\alpha_t$ 线性无关．这说明属于不同特征值的特征向量是线性无关的．由此得到矩阵 A 与对角矩阵相似的一个充分条件：

定理 5 如果 n 阶矩阵有 n 个不同的特征值，那么 A 可以对角化．

在复数域中，因为每个 n 次多项式都有 n 个根，所以从定理 5 可以得到：

推论 如果复数域上的矩阵 A 的特征多项式没有重根，那么 A 可以化为对角形．

需要注意的是，这两个条件都不是必要的．例如，n 阶单位矩阵 E 本身就是对角形，可是 E 只有唯一的特征值 1，它是 E 的特征多项式的 n 重根．

习　题　4.3

1. 习题 4.2 第 1 题中的矩阵哪些可以对角化？哪些不能？对于能对角化的矩阵 A，试找出可逆矩阵 X，使 $X^{-1}AX$ 为对角矩阵．
2. 设 α,β 分别是 A 的属于 λ_1,λ_2 的特征向量，而且 $\lambda_1\neq\lambda_2$．试证：

   (1) α,β 线性无关；

   (2) $\alpha+\beta$ 不可能是 A 的特征向量．
3. 设上三角矩阵

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

   的对角线元素 $a_{11},a_{22},\cdots,a_{nn}$ 各不相同，试证：A 可以对角化．
4. A 是一个 3 阶方阵．已知它的特征值为 $\lambda_1=1,\lambda_2=-1,\lambda_3=0$．$A$ 的属于特征值 $\lambda_1,\lambda_2,\lambda_3$ 的特征向量依次为：

$$\alpha_1=\begin{bmatrix}1\\2\\1\end{bmatrix},\quad \alpha_2=\begin{bmatrix}0\\-2\\1\end{bmatrix},\quad \alpha_3=\begin{bmatrix}1\\1\\2\end{bmatrix}$$

求 A.

5. 计算

$$\begin{bmatrix}1 & 2 & 2\\2 & 1 & 2\\2 & 2 & 1\end{bmatrix}^k \quad (k > 0,\text{整数}).$$

4.4 实对称矩阵的对角化

这一节讨论实对称矩阵的标准形问题. 主要证明：任一个实对称矩阵都可以化为对角形. 为了证明这个结论，必须先对实对称矩阵的特征值加以讨论，首先来证明下述定理.

定理 6 实对称矩阵的特征多项式的根都是实数.

证明： 设 $A=[a_{ij}]$是一个 n 阶实对称矩阵；λ_0 是 A 的特征多项式 $f(\lambda)=|\lambda E-A|$ 的一个根. $\alpha=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}$是齐次线性方程组 $(\lambda_0 E-A)X=0$ 的一个非零解，即 α 满足

$$(\lambda_0 E-A)\alpha=0.$$

其中 $\lambda_0,c_1,c_2,\cdots,c_n$ 都是复数. 于是

$$A\alpha=\lambda_0\alpha.$$

乘开后比较各分量，得

$$\begin{cases}a_{11}c_1+a_{12}c_2+\cdots+a_{1n}c_n=\lambda_0 c_1,\\a_{21}c_1+a_{22}c_2+\cdots+a_{2n}c_n=\lambda_0 c_2,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\a_{n1}c_1+a_{n2}c_2+\cdots+a_{nn}c_n=\lambda_0 c_n.\end{cases}$$

将第 i 个式子乘上 c_i 的共轭复数 $\bar{c}_i(i=1,2,\cdots,n)$，并将所得到的 n 个式子相加，得

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\bar{c}_i c_j = \sum_{i=1}^{n}\lambda_0\bar{c}_i c_i = \lambda_0\sum_{i=1}^{n}\bar{c}_i c_i.$$

因为 α 是非零向量，所以 $c_1,c_2,\cdots,c_n$ 不全为零，$\sum_{i=1}^{n}\bar{c}_i c_i$ 是一个非零实数，可解出 λ_0：

$$\lambda_0 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\bar{c}_i c_j}{\sum_{i=1}^{n}\bar{c}_i c_i}.$$

要证明 λ_0 是一个实数，只要证明：

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\bar{c}_i c_j$$

是一个实数就行了. 为此，下面证明这个数的共轭复数就等于它自己：

$$\begin{aligned}\overline{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\bar{c}_i c_j} &= \sum_{i=1}^{n}\sum_{j=1}^{n}\overline{a_{ij}\bar{c}_i c_j} = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}c_i\bar{c}_j \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ji}c_i\bar{c}_j = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\bar{c}_i c_j.\end{aligned}$$

其中倒数第二个等式的成立是由于 A 是一个实对称矩阵，所以 $a_{ij}=a_{ji}$. 而最后一个等式成立是由于双重连加号的可交换性. 于是定理得到证明. |

推论 n 阶实对称矩阵有 n 个实特征值（重根按重数计算）.

应用定理 6，可以证明下述定理.

定理 7 设 A 是一个 n 阶实对称矩阵，那么可以找到 n 阶正交矩阵 T，使得 $T^{-1}AT$ 为对角矩阵.

证明：对实对称矩阵 A 的阶数 n 作归纳法：当 $n=1$ 时，结论是显然成立的. 假设对于 $n-1$ 阶实对称矩阵结论成立. 下面证明对于 n 阶实对称矩阵结论也成立.

设 λ_1 是 A 的一个特征值，并设 α 是 A 的属于 λ_1 的一个特征

向量. 于是

$$A\alpha = \lambda_1\alpha.$$

因为特征向量的非零倍数还是特征向量. 所以可设 α 是一个单位向量.

以 α 为第 1 列作一个正交矩阵 T_1，那么

$$T_1^{-1}AT_1 = \begin{bmatrix} \lambda_1 & b_2 & \cdots & b_n \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}.$$

但是，因为 T_1 是正交矩阵，所以

$$(T_1^{-1}AT_1)^{\mathrm{T}} = T_1^{\mathrm{T}}A^{\mathrm{T}}(T_1^{-1})^{\mathrm{T}} = T_1^{-1}AT_1.$$

即 $T_1^{-1}AT_1$ 也是对称矩阵，所以 $b_2=\cdots=b_n=0$，而且 A_1 是 $n-1$ 阶实对称矩阵：

$$T_1^{-1}AT_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix}.$$

根据归纳法假设，有 $n-1$ 阶正交矩阵 T_2，使得 $T_2^{-1}A_1T_2$ 为对角矩阵.

$$T_2^{-1}A_1T_2 = [\lambda_2,\cdots,\lambda_n].$$

令
$$T_3 = \begin{bmatrix} 1 & 0 \\ 0 & T_2 \end{bmatrix},$$

则
$$T_3^{-1}\begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix}T_3 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & T_2^{-1}A_1T_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}.$$

再令
$$T = T_1T_3,$$

即得
$$T^{-1}AT = [\lambda_1,\lambda_2,\cdots,\lambda_n].$$

其中 T 是两个正交矩阵的乘积，仍是一个正交矩阵. |

定理 7 说明了使 A 化为对角形的正交矩阵 T 的存在性.那么具体的 T 怎样去找呢?因为 $T^{-1}AT$ 是对角形,所以 T 的每个列都是 A 的特征向量.而且,因为 T 是一个正交矩阵,所以 T 的 n 个列组成一个正交的单位向量组.因此在求出 A 的 n 个线性无关的特征向量以后,还需要再把这 n 个向量正交化与单位化.因为属于不同特征值的特征向量的和不再是特征向量(见习题 4.3 第 2 题),所以为了保证正交化后所得到的向量仍是特征向量,需要用到下述定理:

定理 8 设 A 是一个实对称矩阵.那么属于 A 的不同特征值的特征向量一定是正交的.

证明:设 α_1,α_2 分别是 A 的属于不同特征值 λ_1,λ_2 的实特征向量:

$$A\alpha_1=\lambda_1\alpha_1,\ A\alpha_2=\lambda_2\alpha_2,\ \lambda_1\neq\lambda_2.$$

于是
$$(A\alpha_1,\alpha_2)=(\lambda_1\alpha_1,\alpha_2)=\lambda_1(\alpha_1,\alpha_2).$$

而
$$\begin{aligned}(A\alpha_1,\alpha_2)&=(A\alpha_1)'\alpha_2=\alpha_1'A'\alpha_2=\alpha_1'A\alpha_2=(\alpha_1,A\alpha_2)\\&=(\alpha_1,\lambda_2\alpha_2)=\lambda_2(\alpha_1,\alpha_2),\end{aligned}$$

所以
$$\lambda_1(\alpha_1,\alpha_2)=\lambda_2(\alpha_1,\alpha_2).$$

但是 $\lambda_1\neq\lambda_2$,所以 $(\alpha_1,\alpha_2)=0$,即 α_1 与 α_2 是正交的. |

根据上面的分析,对于 n 阶实对称矩阵 A,正交矩阵 T 的求法可按照以下步骤进行:

1. 求出特征多项式 $f(\lambda)=|\lambda E-A|$ 的全部根,即 A 的特征值,设 A 的全部不同的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_t$.

2. 对每个 $\lambda_i(i=1,2,\cdots,t)$ 解齐次线性方程组

$$(\lambda_iE-A)X=0,$$

找出一个基础解系 $\alpha_{i1},\alpha_{i2},\cdots,\alpha_{is_i}$.

3. 将 $\alpha_{i1},\alpha_{i2},\cdots,\alpha_{is_i}$ 正交化,单位化,得到一组正交的单位向量 $\eta_{i1},\eta_{i2},\cdots,\eta_{is_i}$. 它们是 A 的属于 λ_i 的线性无关的特征向量.

4. 因为 $\lambda_1,\lambda_2,\cdots,\lambda_t$ 各不相同,向量组

$$\eta_{11},\cdots,\eta_{1s_1},\eta_{21},\cdots,\eta_{2s_2},\cdots,\eta_{t1},\cdots,\eta_{ts_t}.$$

仍是正交的单位向量组. 它们的总数共 n 个. 以这一组向量为列，作一个矩阵 T，那么 T 即为所求的正交矩阵.

例 1 设

$$A=\begin{bmatrix}2&-1&-1&1\\-1&2&1&-1\\-1&1&2&-1\\1&-1&-1&2\end{bmatrix},$$

求正交矩阵 T，使 $T^{-1}AT$ 为对角形.

解：首先求 A 的特征值. 因为

$$|\lambda E-A|=\begin{vmatrix}\lambda-2&1&1&-1\\1&\lambda-2&-1&1\\1&-1&\lambda-2&1\\-1&1&1&\lambda-2\end{vmatrix}=(\lambda-1)^3(\lambda-5),$$

所以 A 的特征值为 1(3 重)，5.

其次，求属于 1 的特征向量：把 $\lambda_1=1$ 代入齐次线性方程组 $(\lambda E-A)X=0$，得

$$\begin{cases}-x_1+x_2+x_3-x_4=0,\\ \quad x_1-x_2-x_3+x_4=0,\\ \quad x_1-x_2-x_3+x_4=0,\\ -x_1+x_2+x_3-x_4=0.\end{cases}$$

求得一个基础解系：

$$\begin{cases}\alpha_1=(1,1,0,0),\\ \alpha_2=(1,0,1,0),\\ \alpha_3=(1,0,0,-1).\end{cases}$$

把它正交化，得

$$\begin{cases}\beta_1=\alpha_1,\\ \beta_2=\alpha_2-\dfrac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1=\left(\dfrac{1}{2},-\dfrac{1}{2},1,0\right),\\ \beta_3=\alpha_3-\dfrac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\dfrac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2=\left(\dfrac{1}{3},-\dfrac{1}{3},-\dfrac{1}{3},-1\right)\end{cases}$$

再单位化，得

$$\eta_1=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},0,0\right),$$

$$\eta_2=\left(\frac{\sqrt{6}}{6},-\frac{\sqrt{6}}{6},\frac{\sqrt{6}}{3},0\right),$$

$$\eta_3=\left(\frac{\sqrt{3}}{6},-\frac{\sqrt{3}}{6},-\frac{\sqrt{3}}{6},-\frac{\sqrt{3}}{2}\right).$$

再求属于 5 的特征向量：把 $\lambda_2=5$ 代入 $(\lambda E-A)X=0$，得

$$\begin{cases}3x_1+x_2+x_3-x_4=0,\\ x_1+3x_2-x_3+x_4=0,\\ x_1-x_2+3x_3+x_4=0,\\ -x_1+x_2+x_3+3x_4=0.\end{cases}$$

求得基础解系为

$$\alpha_4=(1,-1,-1,1).$$

α_4 一定与 η_2,η_2,η_3 正交. 再将 α_4 单位化，得

$$\eta_4=\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right).$$

$\eta_1,\eta_2,\eta_3,\eta_4$ 是 A 的一组单位正交的特征向量. 以它们为列，作一个矩阵

$$T=\begin{bmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{2}\\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & -\frac{1}{2}\\ 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{6} & -\frac{1}{2}\\ 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2}\end{bmatrix}$$

T 是一个正交矩阵,而且

$$T^{-1}AT=[1,1,1,5].$$

习 题 4.4

1. 求正交矩阵 T,使 $T^{-1}AT$ 为对角矩阵:

(1) $A=\begin{bmatrix}2 & -2 & 0\\ -2 & 1 & -2\\ 0 & -2 & 0\end{bmatrix}$; (2) $A=\begin{bmatrix}1 & -2 & 2\\ -2 & -2 & 4\\ 2 & 4 & -2\end{bmatrix}$;

(3) $A=\begin{bmatrix}1 & 2 & 4\\ 2 & -2 & 2\\ 4 & 2 & 1\end{bmatrix}$;

(4) $A=\begin{bmatrix}3 & 1 & 0 & -1\\ 1 & 3 & -1 & 0\\ 0 & -1 & 3 & 1\\ -1 & 0 & 1 & 3\end{bmatrix}$;

(5) $A=\begin{bmatrix}1 & -1 & 0 & 0\\ -1 & 1 & 0 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & -1 & 1\end{bmatrix}$.

2. 试证:若 A 是 n 阶实对称矩阵,并且 $A^2=A$,则存在正交矩阵 T,使得

$$T^{-1}AT=\begin{bmatrix}E_r & \\ & O\end{bmatrix}.$$

3. 试证：若 A 是 n 阶实对称矩阵，并且 $A^2=E$，则存在正交矩阵 T，使得

$$T^{-1}AT=\begin{bmatrix} E_r & \\ & -E_{n-r} \end{bmatrix}.$$

4. 试证：若 A 是 n 阶实对称矩阵，并且 A 是幂零矩阵，则 $A=O$.

5. 试证：实反对称矩阵的特征多项式的根是零或是纯虚数.

4.5 约当标准形简单介绍

我们知道，并不是每个矩阵都与一个对角矩阵相似. 这一节介绍，已知一个矩阵 A，与 A 相似的较简单的矩阵是什么形状. 本节讨论在复数域中进行.

定义 5 形式为

$$J=\begin{bmatrix} \lambda_0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_0 \end{bmatrix}$$

的矩阵称为**约当(Jordan)块**. 其中 λ_0 是复数. 由若干个约当块组成的准对角矩阵称为**约当形矩阵**.

1 阶约当块就是 1 阶矩阵，所以对角矩阵是约当形矩阵的特例.

本节要介绍的主要结果是：

定理 9 每个 n 阶复系数矩阵 A 都与一个约当形矩阵相似. 这个约当形矩阵除去其中约当块的排列次序外，是被矩阵 A 唯一决定的，它称为 A 的**约当标准形**.

因为约当形矩阵是上三角矩阵，所以不难算出，在矩阵 A 的约当标准形中，对角线上的元素正好就是 A 的特征多项式的全部根(重根按重数计算).

这个定理的证明比较复杂，超出了本册所要求的范围，这里就不讲了．下面证明一个较为容易的定理．这个定理的证明与定理 7 的证明是完全相仿的．

定理 10 每个复系数矩阵都与一个上三角矩阵相似．

证明：对于矩阵的阶数 n 作归纳法：当 $n=1$ 时，结论是显然成立的．假设结论对 $n-1$ 阶矩阵成立，下面证明结论对 n 阶矩阵也成立．

设 A 是一个 n 阶复矩阵．并设 α_1 是 A 的属于 λ_1 的特征向量：

$$A\alpha_1 = \lambda_1\alpha_1.$$

再选 $n-1$ 个向量 $\alpha_2,\cdots,\alpha_n$，使得 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关．以 α_1，$\alpha_2,\cdots,\alpha_n$ 为例，作一个矩阵 X_1．X_1 是一个可逆矩阵，而且

$$X_1^{-1}AX_1 = \begin{bmatrix} \lambda_1 & b \\ O & A_1 \end{bmatrix}.$$

其中 O 是 $(n-1)\times 1$ 零矩阵；b 是 $1\times(n-1)$ 矩阵；A_1 是 $n-1$ 阶方阵．

根据归纳假设，结论对 $n-1$ 阶矩阵成立，所以有 $n-1$ 阶可逆矩阵 X_2，使 $X_2^{-1}A_1X_2$ 为上三角矩阵：

$$X_2^{-1}A_1X_2 = \begin{bmatrix} \lambda_2 & & * \\ & \ddots & \\ O & & \lambda_n \end{bmatrix}.$$

（$*$ 代表矩阵中主对角线上方的元素）．

令
$$X_3 = [1, \quad X_2],$$

那么

$$X_3^{-1}\begin{bmatrix} \lambda_1 & b \\ O & A_1 \end{bmatrix}X_3 = \begin{bmatrix} 1 & O \\ O & X_2^{-1} \end{bmatrix}\begin{bmatrix} \lambda_1 & b \\ O & A_1 \end{bmatrix}\begin{bmatrix} 1 & O \\ O & X_2 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & bX_2 \\ O & X_2^{-1}A_1X_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & * & \\ & O & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

再令 $$X = X_1X_3,$$
那么 $$X^{-1}AX = (X_1X_3)^{-1}A(X_1X_3) = X_3^{-1}(X_1^{-1}AX_1)X_3$$
$$= X_3^{-1}\begin{bmatrix}\lambda_1 & b \\ O & A_1\end{bmatrix}X_3 = \begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & * & \\ & O & \ddots & \\ & & & \lambda_n\end{bmatrix}.$$
所以 A 相似于一个上三角矩阵.

同样,如果 A 与上三角矩阵 B 相似,那么 B 的对角线元素正好是 A 的特征多项式的全部根(重根按重数计算).

下面我们来举例说明这个定理的应用.

例 1 设 n 阶矩阵 A 的特征多项式的 n 个根是 $\lambda_1,\lambda_2,\cdots,\lambda_n$;$f(x)$是一个多项式. 求证:$f(A)$的特征多项式的 n 个根是 $f(\lambda_1)$,$f(\lambda_2),\cdots,f(\lambda_n)$.

证明:由题设,根据定理 10,A 与一个上三角形 B 相似;
$$B = \begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & * \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}.$$
于是 $f(A)$与 $f(B)$相似:
$$f(B) = \begin{bmatrix}f(\lambda_1) & & & \\ & f(\lambda_2) & & * \\ & & \ddots & \\ & & & f(\lambda_n)\end{bmatrix}.$$
$f(A)$与 $f(B)$有相同的特征多项式,所以 $f(A)$的特征多项式的 n 个根是 $f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$.

习　题　4.5

1. 计算

$$\begin{bmatrix} \lambda_0 & 1 \\ & \lambda_0 \end{bmatrix}^k, \quad \begin{bmatrix} \lambda_0 & 1 & \\ & \lambda_0 & 1 \\ & & \lambda_0 \end{bmatrix}^k \quad (k > 0, \text{整数}).$$

2. 试证 $\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$的约当标准形是$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

3. 求可逆矩阵 T,使 $T^{-1}\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

4. 求$\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}^k$ $(k > 0$,整数).

内 容 提 要

1. 相似矩阵

(1) 矩阵相似的定义

(1) 相似矩阵的公共性质

2. 特征值与特征向量

(1) 基本概念：特征值、特征向量、特征矩阵、特征多项式.

(2) 矩阵 A 的特征向量的求法：

1) 计算 A 的特征多项式 $f(\lambda)=|\lambda E-A|$.

2) 求出 $f(\lambda)$的全部根，它们也就是 A 的全部特征值.

3) 把 A 的特征值逐个代入齐次线性方程组

$$(\lambda E - A)X = 0,$$

求出这个方程组的一个基础解系. 那么，这个基础解系的非零线性组合就是 A 的属于这个特征值的全部特征向量.

(3) 特征向量的重要性质：

1) 属于不同特征值的特征向量是线性无关的.

2) 属于同一个特征值的特征向量是非零线性组合还是属于这个特征值的特征向量；属于不同特征值的特征向量的和不再是特征向量.

3．相似于对角矩阵的条件

（1）相似于对角矩阵的条件：

1）n 阶矩阵 A 相似于一个对角矩阵的必要充分条件是：A 有 n 个线性无关的特征向量；

2）如果 n 阶矩阵 A 有 n 个不同的特征值，那么，A 与一个对角矩阵相似.

3）如果复数域上的矩阵 A 的特征多项式没有重根，那么 A 可对角化.

（2）求矩阵 X，使 $X^{-1}AX$ 为对角矩阵的方法：

如果 n 阶矩阵相似于一个对角矩阵，那么，A 有 n 个线性无关的特征向量. 以 n 个线性无关的特征向量为列，作矩阵 X，那么，矩阵 X 可逆，并且

$$X^{-1}AX = [\lambda_1, \lambda_2, \cdots, \lambda_n],$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 X 的 n 个列所对应的特征值.

4．实对称矩阵的对角化问题

（1）主要结果：

设 A 是一个 n 阶实对称矩阵，那么，可以找到一个 n 阶正交矩阵 T，使得

$$T^{-1}AT = [\lambda_1, \lambda_2, \cdots, \lambda_n],$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的特征值.

（2）实对称矩阵的特征值和特征向量的重要性质：

1）实对称矩阵的特征多项式的根都是实数；

2）n 阶实对称矩阵有 n 个实特征值；

3）实对称矩阵的属于不同特征值的特征向量是正交的.

（3）T 的求法：

1）找出实对称矩阵 A 的 n 个线性无关的特征向量；

2）把属于同一特征值的特征向量正交化，单位化，得到 n 个单位正交的特征向量 $\eta_1, \eta_2, \cdots, \eta_n$；

3) 以 $\eta_1,\eta_2,\cdots,\eta_n$ 为列作一个矩阵 T,则 T 就是正交矩阵,并且

$$T^{-1}AT = [\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n],$$

其中 λ_i 是 η_i 所对应的特征值.

5. 约当标准形

(1) 每个复矩阵 A 都与一个约当形矩阵相似,而且与 A 相似的约当形矩阵除去其中约当块的排列次序外是唯一的.此约当形矩阵称为 A 的约当标准形.

(2) 每个复矩阵都与一个上三角矩阵相似.

(3) 如果 A 与上三角矩阵 B 相似,那么 B 的对角线元素就是 A 的特征多项式的全部根(重根按重数计算).

复 习 题 4

1. 求下列矩阵的全部特征值与特征向量:

(1) $A = \begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$;　　(2) $A = \begin{bmatrix} 6 & 2 & 4 \\ 2 & 3 & 2 \\ 4 & 2 & 6 \end{bmatrix}$;

(3) $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 3 & 6 \end{bmatrix}$;　　(4) $A = \begin{bmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$.

2. 找出 1. 题中可对角化的矩阵 A,并求可逆矩阵 X 使 $X^{-1}AX$ 为对角矩阵.

3. 求正交矩阵 T 使 $T^{-1}AT$ 为对角矩阵:

(1) $A = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$;　　(2) $A = \begin{bmatrix} 0 & -2 & 2 \\ -2 & -3 & 4 \\ 2 & 4 & -3 \end{bmatrix}$;

(3) $A = \begin{bmatrix} 1 & 3 & -3 & 3 \\ 3 & 1 & 3 & -3 \\ -3 & 3 & 1 & 3 \\ 3 & -3 & 3 & 1 \end{bmatrix}$;

(4) $A=\begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 \\ 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & 1 \end{bmatrix}$.

4. 试证：矩阵 A 可逆的充分必要条件是：它的特征值都不等于零.

5. 设 n 阶可逆矩阵 A 的特征值是 $\lambda_1,\lambda_2,\cdots,\lambda_n$，证明：$A^{-1}$的特征值是 λ_1^{-1}，$\lambda_2^{-1},\cdots,\lambda_n^{-1}$.

6. 设 λ_0 是 n 阶矩阵 A 的特征多项式的 k 重根，求证：A 的属于 λ_0 的特征向量最多有 k 个是线性无关的.

7. 设 A 是一个 n 阶复矩阵，$\lambda_1,\lambda_2,\cdots,\lambda_s$ 是它的全部不同特征值，重数分别为 $n_1,n_2,\cdots,n_s$. 求证：A 可以对角化的充分必要条件是：对每个 λ_i，A 恰有 n_i 个线性无关的特征向量，$i=1,2,\cdots,s$.

8. 如果任一个 n 维非零向量都是 n 阶矩阵 A 的特征向量，则 A 是一个数量矩阵.

9. A 是一个 n 阶实对称矩阵，试证：如果 λ_0 是 A 的 k 重特征值，则矩阵 λ_0E-A 的秩等于 $n-k$.

10. 设 A,B 都是实对称矩阵，试证：存在正交矩阵 T，使 $T^{-1}AT=B$ 的充分必要条件是：A 与 B 的特征多项式相等.

第5章　二次型

二次型的问题起源于化二次曲线和二次曲面为标准形的问题.在解析几何中,曾介绍过,当坐标原点与中心重合时,有心二次曲线的一般方程是:

$$ax^2 + 2bxy + cy^2 = f.$$

为了便于研究这个二次曲线,可以将坐标轴适当旋转而把上述方程化成标准方程.在二次曲面的研究中也有类似的情况.

上述方程式的左端是一个二次齐次多项式.二次齐次多项式不但在几何中有,而且在数学的其他分支以及物理、力学中也常常会碰到.这一章的内容就是介绍二次齐次多项式的一些重要性质.

5.1　二次型及其矩阵表示

设 **P** 是一个数域,以 **P** 中的数作系数的 $x_1, x_2, \cdots, x_n$ 的二次齐次多项式

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = & a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n \\ & + a_{22}x_2^2 + \cdots + 2a_{2n}x_2x_n + \cdots + a_{nn}x_n^2 \end{aligned} \tag{1}$$

称为数域 **P** 上的一个 n 元二次型,在不致引起混淆的情况下,也简称二次型.例如

$$x_1^2 - 2x_1x_2 + 8x_1x_3 - x_2^2 + 5x_2x_3 + x_3^2$$

就是实数域上的一个3元二次型.为了以后讨论方便,把(1)中 $x_ix_j(i<j)$ 的系数写成 $2a_{ij}$,而不简单地写成 a_{ij}.

与在几何中一样,讨论二次型问题的主要内容是:用变量的线性变换来化简二次型.为此,首先引入下述定义:

定义 1 设 $x_1,x_2,\cdots,x_n;y_1,y_2,\cdots,y_n$ 是两组变量，系数在数域 **P** 中的一组关系式

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n, \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n, \end{cases} \tag{2}$$

称为由 $x_1,x_2,\cdots,x_n$ 到 $y_1,y_2,\cdots,y_n$ 的一个**线性替换**，简称**线性替换**，如果系数矩阵

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

是非退化的，就称线性替换(2)是**非退化的**或**可逆的**. 当 **P** 是实数域时，如果系数矩阵 C 是正交的，就称(2)为**正交的**. 正交的线性替换简称正交替换.

不难看出：如果把(2)代入(1)，那么得到的 $y_1,y_2,\cdots,y_n$ 的多项式仍然是二次齐次的. 这就是说，二次型经过线性替换后还是二次型. 这一章的主要内容就是研究二次型在非退化线性替换下的变化情况.

线性替换可以用它的系数矩阵来表示. 令

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

那么(2)就可以写成

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

即

$$X = CY. \tag{3}$$

如果(2)是一个可逆的线性替换，那么它的系数矩阵 C 是可逆的，用 C^{-1}左乘(3)的两边，得

$$Y = C^{-1}X.$$

设

$$C^{-1} = \begin{bmatrix} c'_{11} & c'_{12} & \cdots & c'_{1n} \\ c'_{21} & c'_{22} & \cdots & c'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c'_{n1} & c'_{n2} & \cdots & c'_{nn} \end{bmatrix},$$

那么 $y_1, y_2, \cdots, y_n$ 也可以由 $x_1, x_2, \cdots, x_n$ 表出：

$$\begin{cases} y_1 = c'_{11}x_1 + c'_{12}x_2 + \cdots + c'_{1n}x_n, \\ y_2 = c'_{21}x_1 + c'_{22}x_2 + \cdots + c'_{2n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = c'_{n1}x_1 + c'_{n2}x_2 + \cdots + c'_{nn}x_n. \end{cases} \tag{4}$$

(4)与(2)表示同一个线性替换. 以后有时候也用(4)给出线性替换.

如果再设

$$\begin{cases} y_1 = d_{11}z_1 + d_{12}z_2 + \cdots + d_{1n}z_n, \\ y_2 = d_{21}z_1 + d_{22}z_2 + \cdots + d_{2n}z_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = d_{n1}z_1 + d_{n2}z_2 + \cdots + d_{nn}z_n, \end{cases} \tag{5}$$

是 $y_1, y_2, \cdots, y_n$ 到 $z_1, z_2, \cdots, z_n$ 的一个线性替换，它的系数矩阵是：

$$D = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix}.$$

令
$$Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}.$$
那么(5)可表成
$$Y = DZ. \tag{6}$$
将(5)代入(2),得到 $x_1,x_2,\cdots,x_n$ 到 $z_1,z_2,\cdots,z_n$ 的一个线性替换,这个线性替换的系数可以通过系数矩阵求出来.因为将(5)代入(2)相当于将(6)代入(3),因此得到 $x_1,\cdots,x_n$ 与 $z_1,z_2,\cdots,z_n$ 间的关系式:
$$X = CY = C(DZ) = (CD)Z.$$
这说明,两个线性替换连续施行(或称相乘)的结果,还是一个线性替换,它的系数矩阵等于原来两个线性替换的矩阵的乘积.

因为可逆矩阵的乘积还是可逆矩阵,所以,可逆的线性替换连续施行的结果还是可逆的.

因为正交矩阵都是可逆的,所以正交替换也一定是可逆的,而且由于正交矩阵的乘积是正交的,所以两个正交替换连续施行的结果,还是一个正交替换.

在讨论二次型时,矩阵是一个有力的工具,因此我们先把二次型用矩阵来表示.

令
$$a_{ji} = a_{ij} \qquad (i < j).$$
因为
$$x_i x_j = x_j x_i,$$
所以二次型(1)可以写成
$$\begin{aligned} f(x_1,x_2,\cdots,x_n) = & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n + \cdots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \\ = & \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j. \end{aligned} \tag{7}$$

把(7)的系数排成一个矩阵

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

它称为**二次型(7)的矩阵**. 因为 $a_{ij}=a_{ji}$, 所以 A 是一个对称矩阵. 也就是说, 二次型的矩阵都是对称矩阵.

令
$$X=\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix},$$

于是二次型就可以用矩阵的乘积表示出来:

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)&=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\\ &=(x_1,x_2,\cdots,x_n)\begin{bmatrix}\sum_{j=1}^{n}a_{1j}x_j\\ \sum_{j=1}^{n}a_{2j}x_j\\ \vdots\\ \sum_{j=1}^{n}a_{nj}x_j\end{bmatrix}\\ &=(x_1,x_2,\cdots,x_n)\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}\\ &=X^{\mathrm{T}}AX.\end{aligned}$$

即
$$f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX.$$

容易看到，二次型(7)的矩阵 A 的对角线元素 $a_{11},a_{22},\cdots,a_{nn}$ 刚好就是(7)中 $x_1^2,x_2^2,\cdots,x_n^2$ 的系数；而 $a_{ij}=a_{ji}(i\neq j)$ 刚好就是 x_ix_j 的系数的一半. 因此二次型与它的矩阵是相互唯一决定的.

二次型经过非退化线性替换后仍是二次型，现在看替换前后的二次型之间有什么关系. 为此，我们来找出替换前后二次型的矩阵之间的关系.

设

$$f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX \qquad (A^{\mathrm{T}}=A) \tag{8}$$

是一个二次型. 作非退化的线性替换

$$X=CY. \tag{9}$$

得到 $y_1,y_2,\cdots,y_n$ 的一个二次型

$$Y^{\mathrm{T}}BY.$$

现在来看矩阵 A 与 B 的关系：

把(9)代入(8)，有

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)&=X^{\mathrm{T}}AX=(CY)^{\mathrm{T}}A(CY)=Y^{\mathrm{T}}C^{\mathrm{T}}ACY\\&=Y^{\mathrm{T}}(C^{\mathrm{T}}AC)Y=Y^{\mathrm{T}}BY.\end{aligned}$$

上式中的矩阵 $C^{\mathrm{T}}AC$ 也是对称的. 这是因为

$$(C^{\mathrm{T}}AC)^{\mathrm{T}}=C^{\mathrm{T}}A^{\mathrm{T}}(C^{\mathrm{T}})^{\mathrm{T}}=C^{\mathrm{T}}AC.$$

所以

$$B=C^{\mathrm{T}}AC.$$

这就是替换前后两个二次型的矩阵之间的关系. 与此相应，有下述定义：

定义 2 对于数域 **P** 上 n 阶矩阵 A,B，如果有 **P** 上的 n 阶可逆矩阵 C，使得

$$B=C^{\mathrm{T}}AC,$$

那么，就说 A 与 B **是合同的**，记作 $A\simeq B$.

因此，经过非退化的线性替换后，新二次型的矩阵与原二次型的矩阵是合同的，所以可以把二次型的替换通过矩阵表示出来，为以后的讨论提供了有力的工具. 二次型的矩阵的秩称为这个二次

型的秩.

合同是矩阵之间的一种关系,这种关系具有:

1. 反身性: $A\simeq A$.

这是因为 $$A=E^{\mathrm{T}}AE.$$

2. 对称性: 如果 $A\simeq B$,那么 $B\simeq A$.

这是因为: 由 $B=C^{\mathrm{T}}AC$ 可得出 $A=(C^{-1})^{\mathrm{T}}BC^{-1}$.

3. 传递性: 如果 $A_1\simeq A_2, A_2\simeq A_3$,那么 $A_1\simeq A_3$.

这是因为,由 $A_2=C_1^{\mathrm{T}}A_1C_1, A_3=C_2^{\mathrm{T}}A_2C_2$,可得:

$$A_3 = (C_1C_2)^{\mathrm{T}}A_1(C_1C_2).$$

合同矩阵还有下列重要性质: 如果 A 与 B 合同: $B=C^TAC$,则 $|B|=|C|^2|A|$,因此合同矩阵都可逆或都不可逆. 不仅如此,合同矩阵有相同的秩.

最后指出,在对二次型进行替换时,总是要求所作的线性替换是非退化的. 于是,如果对二次型

$$X^{\mathrm{T}}AX$$

作非退化线性替换

$$X = CY$$

后,得到一个新的二次型

$$Y^{\mathrm{T}}BY,$$

其中 $$B = C^{\mathrm{T}}AC.$$

把 $X=CY$ 写成:

$$Y = C^{-1}X,$$

这也是一个线性替换,并且它把所得的二次型还原. 这样就使我们可从所得到的二次型的性质推出原来二次型的一些性质.

习　题　5.1

1. 证明:

$$\begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix} \simeq \begin{bmatrix} a_2 & 0 & 0 \\ 0 & a_3 & 0 \\ 0 & 0 & a_1 \end{bmatrix}.$$

2. 写出下列二次型的矩阵表示：

(1) $f(x_1,x_2,x_3)=-4x_1x_2+2x_1x_3+2x_2x_3$；

(2) $f(x_1,x_2,x_3)=x_1^2+2x_1x_2-x_1x_3+2x_3^2$；

(3) $f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2+x_1x_2+x_1x_3+x_2x_3$；

(4) $f(x_1,x_2,x_3,x_4)=x_1x_2-x_3x_4$.

3. 设 A 是一个 n 阶对称矩阵，如果对任一个 n 维列向量 X 都有 $X^{\mathrm{T}}AX=0$，则 $A=0$.

4. 试证：

(1) 当 **P** 是实数域时，对称矩阵 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 与 $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ 不是合同的；

(2) 当 **P** 是复数域时，对称矩阵 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 与 $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ 是合同的.

5.2　用正交替换化实二次型为标准形

这一节讨论用正交的线性替换化简二次型的问题.

因为二次型

$$X^{\mathrm{T}}AX$$

经线性替换

$$X = CY$$

后，成为

$$Y^{\mathrm{T}}BY = Y^{\mathrm{T}}(C^{\mathrm{T}}AC)Y.$$

所以用线性替换化简二次型的问题，用矩阵来说，就是给了对称矩阵 A，如何选择可逆矩阵 C，使

$$B = C^{\mathrm{T}}AC$$

最简单的问题.

在第 4 章 4.4 节讨论实对称矩阵时，曾经证明了任给一个实对称矩阵 A，总可找到一个正交矩阵 T，使得 $T^{-1}AT$ 为对角矩阵

$$T^{-1}AT = [\lambda_1, \lambda_2, \cdots \lambda_n].$$

因为 T 是正交矩阵，所以

$$T^{-1} = T^{\mathrm{T}}.$$

于是 $$T^{\mathrm{T}}AT = T^{-1}AT = [\lambda_1, \lambda_2, \cdots \lambda_n].$$

而一个二次型的系数矩阵如果是对角矩阵，那么，这个二次型就是平方和的形式. 因此，有下述定理：

定理 1 任意一个实二次型

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = X^{\mathrm{T}}AX \quad (a_{ij} = a_{ji}, A^{\mathrm{T}} = A) \tag{1}$$

都可以经正交替换化成平方和

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2, \tag{2}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的全部特征根.

(2)式称为二次型(1)的**标准形**

例 1 用正交替换化实二次型

$$f(x_1, x_2, x_3) = x_1^2 + 4x_2^2 + x_3^2 - 4x_1x_2 - 8x_1x_3 - 4x_2x_3$$

为标准形.

解：$f(x_1, x_2, x_3)$的矩阵为

$$A = \begin{bmatrix} 1 & -2 & -4 \\ -2 & 4 & -2 \\ -4 & -2 & 1 \end{bmatrix}.$$

首先找一个正交矩阵 T，使 $T^{-1}AT$ 为对角形.

先求 A 的特征多项式：

$$|\lambda E - A| = \begin{vmatrix} \lambda - 1 & 2 & 4 \\ 2 & \lambda - 4 & 2 \\ 4 & 2 & \lambda - 1 \end{vmatrix} = (\lambda - 5)^2(\lambda + 4),$$

得 A 的特征值是 5(2 重)，-4.

由 $\lambda=5$ 得到的齐次方程组是：

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

得基础解系：$\boldsymbol{\alpha}_1=(1,0,-1)$，$\boldsymbol{\alpha}_2=(1,-2,0)$.

正交化后，得 $\boldsymbol{\beta}_1=(1,0,-1)$，$\boldsymbol{\beta}_2=\left(\frac{1}{2},-2,\frac{1}{2}\right)$.

单位化后，得 $\boldsymbol{\gamma}_1=\left(\frac{\sqrt{2}}{2},0,-\frac{\sqrt{2}}{2}\right)$，

$$\boldsymbol{\gamma}_2=\left(\frac{\sqrt{2}}{6},-\frac{2\sqrt{2}}{3},\frac{\sqrt{2}}{6}\right).$$

由 $\lambda=-4$ 得到的齐次方程组是

$$\begin{cases}-5x_1+2x_2+4x_3=0,\\2x_1-8x_2+2x_3=0,\\4x_1+2x_2-5x_3=0.\end{cases}$$

得基础解系：

$$\boldsymbol{\alpha}_3=(2,1,2).$$

单位化后，得

$$\boldsymbol{\gamma}_3=\left(\frac{2}{3},\frac{1}{3},\frac{2}{3}\right).$$

由 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3$ 组成正交矩阵

$$T=\begin{bmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & \frac{2}{3}\\ 0 & -\frac{2\sqrt{2}}{3} & \frac{1}{3}\\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & \frac{2}{3}\end{bmatrix},$$

则 $f(x_1,x_2,x_3)$经正交替换

$$X=TY,$$

即

$$\begin{cases} x_1 = \dfrac{\sqrt{2}}{2}y_1 + \dfrac{\sqrt{2}}{6}y_2 + \dfrac{2}{3}y_3, \\ x_2 = \qquad\quad -\dfrac{2\sqrt{2}}{3}y_2 + \dfrac{1}{3}y_3, \\ x_3 = -\dfrac{\sqrt{2}}{2}y_1 + \dfrac{\sqrt{2}}{6}y_2 + \dfrac{2}{3}y_3, \end{cases}$$

化为标准形 $\qquad 5y_1^2 + 5y_2^2 - 4y_3^2.$

需要注意的是，矩阵的合同关系与矩阵的相似关系是两种不同的关系，合同关系只是对称矩阵间的关系. 即使对于对称矩阵来说，也可以找到合同但不相似的矩阵以及相似而不合同的矩阵.

例 2 设 $A=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$, $B=\begin{bmatrix}1 & 0\\0 & 4\end{bmatrix}$.

对此，有可逆矩阵$\begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}$，使得

$$B = \begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}^{\mathrm{T}} A \begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}.$$

所以 B 与 A 是合同的. 但是它们的特征值并不相等，所以它们是不相似的.

然而，如果 A，B 是两个相似的实对称矩阵，那么它们一定是合同的. 这个事实可以这样来证明：因为 A，B 相似，所以 A 与 B 的特征多项式的根全部都相等，因而可找到正交矩阵 T，使得

$$B = T^{-1}AT = T^{\mathrm{T}}AT$$

(见复习题 4 第 8 题)这就证明了 A 与 B 是合同的. 这里证明的关键在于 T 是正交矩阵，所以 $T^{\mathrm{T}}=T^{-1}$. 这一点也是用正交替换化二次型为平方和的主要根据.

习 题 5.2

用正交替换把下列二次型化为平方和，并写出所作的正交替换：

1. $2x_1^2+5x_1^2+5x_3^2+4x_1x_2-4x_1x_3-8x_2x_3$.
2. $3x_1^2+3x_2^2+6x_3^2+8x_1x_2-4x_1x_3+4x_2x_3$.
3. $2x_1x_2-2x_3x_4$.
4. $x_1^2+x_2^2+x_3^2+x_4^2+4x_1x_2+4x_1x_3+4x_1x_4-4x_2x_3-4x_2x_4-4x_3x_4$.
5. $x_1x_2+x_2x_3+x_3x_4+x_4x_1$.

5.3 用非退化线性替换化二次型为标准形

这一节讨论一般数域上的二次型的化简问题．二次型中最简单的一种是只包含平方项的形式，即平方和的形式

$$d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2. \tag{1}$$

本节的主要结果是：

定理 2 数域 P 上任意一个二次型都可以经过非退化的线性替换化成平方和(1)的形式.

证明：这个证明实际上是一个具体地将二次型化成平方和的方法——配方法.

对变量的个数 n 用归纳法：

当 $n=1$ 时，二次型就是：

$$f(x_1)=a_{11}x_1^2,$$

已经是平方和了. 假定对 $n-1$ 元的二次型定理成立. 下面讨论 n 个变量的二次型：

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_ix_j \quad (a_{ij}=a_{ji}). \tag{2}$$

分 3 种情形讨论：

1. $a_{11}\neq 0$，这时

$$f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2+\sum_{j=2}^n a_{1j}x_1x_j+\sum_{i=2}^n a_{i1}x_ix_1$$
$$+\sum_{i=2}^n\sum_{j=2}^n a_{ij}x_ix_j$$

$$
\begin{aligned}
&= a_{11}x_1^2 + 2\sum_{j=2}^{n} a_{1j}x_1x_j + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j \\
&= a_{11}\Big(x_1 + \sum_{j=2}^{n} a_{11}^{-1}a_{1j}x_j\Big)^2 - a_{11}^{-1}\Big(\sum_{j=2}^{n} a_{1j}x_j\Big)^2 \\
&\quad + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j \\
&= a_{11}\Big(x_1 + \sum_{j=2}^{n} a_{11}^{-1}a_{1j}x_j\Big)^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} b_{ij}x_ix_j.
\end{aligned}
$$

其中 $\sum_{i=2}^{n}\sum_{j=2}^{n} b_{ij}x_ix_j = -a_{11}^{-1}\Big(\sum_{j=2}^{n} a_{1j}x_j\Big)^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j$ 是 $x_2,x_3,\cdots,x_n$ 的一个二次型. 令

$$
\begin{cases}
y_1 = x_1 + \sum_{j=2}^{n} a_{11}^{-1}a_{1j}x_j, \\
y_1 = x_2, \\
\cdots\cdots\cdots \\
y_n = x_n,
\end{cases}
\quad \text{即} \quad
\begin{cases}
x_1 = y_1 - \sum_{j=2}^{n} a_{11}^{-1}a_{1j}y_j, \\
x_2 = y_2, \\
\cdots\cdots\cdots \\
x_n = y_n.
\end{cases}
\tag{3}
$$

这是一个非退化的线性替换. $f(x_1,x_2,\cdots,x_n)$经过这个替换后,化成

$$
f(x_1,x_2,\cdots,x_n) = a_{11}y_1^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} b_{ij}y_iy_j. \tag{4}
$$

根据归纳法假设,有非退化的线性替换:

$$
\begin{cases}
y_2 = c_{22}z_2 + c_{23}z_3 + \cdots + c_{2n}z_n, \\
y_3 = c_{32}z_2 + c_{33}z_3 + \cdots + c_{3n}z_n, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
y_n = c_{n2}z_2 + c_{n3}z_3 + \cdots + c_{nn}z_n.
\end{cases}
$$

把二次型 $\sum_{i=2}^{n}\sum_{j=2}^{n} b_{ij}y_iy_j$ 化成平方和

$$
d_2z_2^2 + d_3z_3^2 + \cdots + d_nz_n^2.
$$

于是非退化线性替换

$$\begin{cases} y_1 = z_1, \\ y_2 = c_{22}z_2 + \cdots + c_{2n}z_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = c_{n2}z_2 + \cdots + c_{nn}z_n, \end{cases} \tag{5}$$

就把(4)式变成平方和

$$f(x_1,x_2,\cdots,x_n) = a_{11}z_1^2 + d_2z_2^2 + \cdots + d_nz_n^2. \tag{6}$$

因为线性替换(3)和(5)都是非退化的，所以由 $x_1,x_2,\cdots,x_n$ 到 $z_1,z_2,\cdots,z_n$ 的线性替换

$$\begin{cases} x_1 = z_1 - \sum\limits_{j=2}^{n} a_{11}^{-1}a_{1j}z_j \\ z_2 = c_{22}z_2 + c_{23}z_3 + \cdots + c_{2n}z_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = c_{n2}z_2 + c_{n3}z_3 + \cdots + c_{nn}z_n \end{cases}$$

也是非退化的，而且这个线性替换直接把(1)式化为(6)式右端的平方和．

2. $a_{11}=0$. 但是至少有一个 $a_{1j}\neq 0(j=2,3,\cdots,n)$，不失普遍性，可设 $a_{12}\neq 0$. 令

$$\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = y_3, \\ \cdots\cdots\cdots \\ x_n = y_n, \end{cases}$$

这是一个非退化的线性替换，它使

$$\begin{aligned} f(x_1,x_2,\cdots,x_n) &= 2a_{12}x_1x_2 + \cdots \\ &= 2a_{12}(y_1 + y_2)(y_1 - y_2) + \cdots \\ &= 2a_{12}y_1^2 - 2a_{12}y_2^2 + \cdots. \end{aligned}$$

这时上式右端是 $y_1,y_2,\cdots,y_n$ 的二次型，且 y_1^2 的系数不为零，属于第一种情况，可化为平方和.

3. $a_{11}=a_{12}=\cdots=a_{1n}=0$. 由于对称性，有

$$a_{21}=a_{31}=\cdots=a_{n1}=0.$$

这时
$$f(x_1,x_2,\cdots,x_n)=\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}x_ix_j$$
是一个 $n-1$ 元的二次型，根据归纳法假定，它能用非退化的线性替换化成平方和.

这样就完成了定理的证明. ▌

二次型 $f(x_1,x_2,\cdots,x_n)$ 经过非退化线性替换所变成的平方和称为 $f(x_1,x_2,\cdots,x_n)$ 的**标准形**.

例 1 化二次型

$$f(x_1,x_2,x_3)=x_1x_2+x_1x_3-3x_2x_3$$

为标准形.

解：作非退化线性替换：

$$\begin{cases}x_1=y_1+y_2,\\x_2=y_1-y_2,\\x_3=y_3,\end{cases}$$

则

$$\begin{aligned}f(x_1,x_2,x_3)&=(y_1-y_2)(y_1+y_2)+(y_1+y_2)y_3\\&\quad-3(y_1-y_2)y_3\\&=y_1^2-y_2^2-2y_1y_3+4y_2y_3\\&=(y_1-y_3)^2-y_2^2-y_3^2+4y_2y_3.\end{aligned}$$

再令 $\begin{cases}y_1-y_3=z_1,\\y_2=z_2,\\y_3=z_3,\end{cases}$ 即 $\begin{cases}y_1=z_1+z_3,\\y_2=z_2,\\y_3=z_3,\end{cases}$

得 $f(x_1,x_2,x_3)=z_1^2-z_2^2-z_3^2+4z_2z_3=z_1^2-(z_2-2z_3)^2+3z_3^2.$

最后，令

$$\begin{cases} z_1 = w_1, \\ z_2 - 2z_3 = w_2, \\ z_3 = w_3, \end{cases} \quad 即 \quad \begin{cases} z_1 = w_1, \\ z_2 = w_2 + 2w_3, \\ z_3 = w_3, \end{cases}$$

就得到 $f(x_1, x_2, x_3) = w_1^2 - w_2^2 + 3w_3^2$.

也就是将 $f(x_1, x_2, x_3)$ 化成了平方和.

上面所作的几次线性替换相当于作一个总的线性替换：

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

$$= \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix},$$

即

$$\begin{cases} x_1 = w_1 - w_2 + 3w_3, \\ x_2 = w_1 + w_2 - w_3, \\ x_3 = w_3. \end{cases}$$

平方和(1)的系数矩阵是对角矩阵：

$$d_1x_1^2 + d_2x_2^2 + \cdots + d_nx_n^2$$

$$= (x_1, x_2, \cdots, x_n)\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix};$$

反过来，系数矩阵为对角形的二次型只含有平方项. 因为二次型经过非退化替换后，它的矩阵变到一个合同矩阵. 因此，定理 2 关于矩阵的相应的结果为：

定理 3 任一个对称矩阵都合同于一个对角矩阵.

也就是说，对于任一个对称矩阵，总可找到一个可逆矩阵 C，使 $C^{\mathrm{T}}AC$ 为对角矩阵. 矩阵 C 的找法与定理 2 的证明过程类似，只要找出与线性替换相应的矩阵即可. 举例说明如下：

例 2 设
$$A = \begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & 5 \\ -2 & 5 & 0 \end{bmatrix}.$$
求可逆矩阵 C，使 $C^{\mathrm{T}}AC$ 为对角矩阵.

解：因为 A 的第 1 行第 1 列处的元素等于零，所以令
$$C_1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
则
$$C_1^{\mathrm{T}}AC_1 = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & 5 \\ -2 & 5 & 0 \end{bmatrix}\begin{bmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 0 & 3 \\ 0 & -2 & -7 \\ 3 & -7 & 0 \end{bmatrix} = A_1.$$

再令
$$C_2=\begin{bmatrix}1 & 0 & -\frac{3}{2}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix},$$
则
$$C_2^{\mathrm{T}}A_1C_2=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -\frac{3}{2} & 0 & 1\end{bmatrix}\begin{bmatrix}2 & 0 & 3\\ 0 & -2 & -7\\ 3 & -7 & 0\end{bmatrix}\begin{bmatrix}1 & 0 & -\frac{3}{2}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$
$$=\begin{bmatrix}2 & 0 & 0\\ 0 & -2 & 7\\ 0 & -7 & -\frac{9}{2}\end{bmatrix}=A_2$$
最后令
$$C_3=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & -\frac{7}{2}\\ 0 & 0 & 1\end{bmatrix},$$
即得
$$C_3^{\mathrm{T}}A_2C_3=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -\frac{7}{2} & 1\end{bmatrix}\begin{bmatrix}2 & 0 & 0\\ 0 & -2 & -7\\ 0 & -7 & -\frac{9}{2}\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & -\frac{7}{2}\\ 0 & 0 & 1\end{bmatrix}$$
$$=\begin{bmatrix}2 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & 20\end{bmatrix}=A_3.$$
A_3 已经是对角矩阵了.将 C_1,C_2,C_3 相乘,即得 C:
$$C=C_1C_2C_3=\begin{bmatrix}1 & 1 & 0\\ 1 & -1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & -\frac{3}{2}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & -\frac{7}{2}\\ 0 & 0 & 1\end{bmatrix}$$

$$=\begin{bmatrix}1 & 1 & -5\\1 & -1 & 2\\0 & 0 & 1\end{bmatrix}.$$

于是 C 是可逆矩阵. $C^{T}AC=\begin{bmatrix}2 & 0 & 0\\0 & -2 & 0\\0 & 0 & 20\end{bmatrix}$.

也可以用化二次型为标准形的方法来解例 2,以 A 为矩阵作一个二次型:

$$f(x_1,x_2,x_3)=X^{T}AX=2x_1x_2-4x_1x_3+10x_2x_3$$

令

$$\begin{cases}x_1=y_1+y_2,\\x_2=y_1-y_2,\\x_3=y_3,\end{cases}$$

即

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1 & 1 & 0\\1 & -1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix},$$

则

$$\begin{aligned}f(x_1,x_2,x_3)&=2y_1^2-2y_2^2+6y_1y_3-14y_2y_3.\\&=2\left(y_1+\frac{3}{2}y_3\right)^2-2y_2^2-\frac{9}{2}y_3^2-14y_2y_3\\&=2\left(y_1+\frac{3}{2}y_3\right)^2-2\left(y_2+\frac{7}{2}y_3\right)^2+20y_3^2\end{aligned}$$

再令

$$\begin{cases}y_1+\dfrac{3}{2}y_3=z_1,\\y_2+\dfrac{7}{2}y_3=z_2,\\y_3=z_3,\end{cases}$$

即
$$\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}1&0&-\frac{3}{2}\\0&1&-\frac{7}{2}\\0&0&1\end{bmatrix}\begin{bmatrix}z_1\\z_2\\z_3\end{bmatrix}$$
则 $f(x_1,x_2,x_3)$化为平方和：
$$2z_1^2-2z_2^2+20z_3^2.$$
所作的非退化线性替换为：
$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1&1&0\\1&-1&0\\0&0&1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$$
$$=\begin{bmatrix}1&1&0\\1&-1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&-\frac{3}{2}\\0&1&-\frac{7}{2}\\0&0&1\end{bmatrix}\begin{bmatrix}z_1\\z_2\\z_3\end{bmatrix}$$
$$=\begin{bmatrix}1&1&-5\\1&-1&2\\0&0&1\end{bmatrix}\begin{bmatrix}z_1\\z_2\\z_3\end{bmatrix}.$$
考虑替换前后的矩阵，令
$$C=\begin{bmatrix}1&1&-5\\1&-1&2\\0&0&1\end{bmatrix}$$
则
$$C^{\mathrm{T}}AC=\begin{bmatrix}2&0&0\\0&-2&0\\0&0&20\end{bmatrix}.$$

由于把二次型用非退化的线性替换化为标准形的问题，就是把这个二次型的矩阵（对称矩阵）在合同关系下化为对角形的问题，所以也可以用矩阵的方法把二次型化为标准形．但在例 2 中所用的方法基本上是线性替换的方法，不过只是写成矩阵的形式而

已，并没有显示出矩阵方法的优越性来．下面要介绍如何利用初等变换及初等矩阵的概念来简化运算．这里主要解决两个问题：第一是要弄清楚每次的 C_i 怎样求？第二，由于每次需要记下 C_i，最后还要算出这些 C_i 的乘积，很麻烦．故需想法简化这个运算过程．

在第 3 章 3.4 节中曾经介绍过用初等变换将矩阵化成对角形的方法．在那里，变换前后的矩阵是等价的，故对矩阵可以随意作初等变换．在这里，要求变换前后的矩阵是合同的，所以就不能随意地作初等变换了．必须对行、列同时作初等变换，并且要自始至终地使矩阵保持合同关系．那么怎样才能作到这一点呢？我们知道，对 A 施行一次初等列替换，其结果等于在 A 的右边乘上一个初等矩阵 P_1．为了保证所得到的矩阵与 A 合同，就必须在左边乘上 P_1^T，而得 $P_1^TAP_1$．这就需要弄清楚用 P_1 右乘 A 所对应的初等列替换与用 P_1^T 左乘 A 所对应的初等行替换之间的关系．由于初等矩阵有 3 类，所以下面分 3 种情形来说明：

$$
1.\ P_1=P(i,j)=\begin{array}{c}\left[\begin{array}{ccccccccc}1&&&&&&&&\\&\ddots&&&&&&&\\&&1&&&&&&\\&&&0&\cdots&\cdots&1&&\\&&&\vdots&1&&\vdots&&\\&&&\vdots&&\ddots&\vdots&&\\&&&\vdots&&1&\vdots&&\\&&&1&\cdots&\cdots&0&&\\&&&&&&&1&\\&&&&&&&&\ddots\\&&&&&&&&&1\end{array}\right]\end{array}\begin{array}{l}\\\\\\\text{第 } i \text{ 行}\\\\\\\\\text{第 } j \text{ 行}\\\\\\\\\end{array}.
$$

此时，$P_1^T=P_1$．用 P_1 右乘 A，相当于把 A 的第 i,j 列互换；而用 P_1^T 左乘 A，相当于把 A 的第 i,j 行互换．

$$
2.\ P_1=P(i(c))=\left[\begin{array}{ccccccc}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&c&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{array}\right]\text{第 } i \text{ 行}\ (c\neq 0).
$$

此时，$P_1^{\mathrm{T}}=P_1$. 用 P_1 右乘 A，相当于用非零常数 c 乘 A 的第 i 列；用 P_1^{T} 左乘 A，相当于用 c 乘 A 的第 i 行.

$$3.\ P_1=P(i,j(k))=\begin{array}{cl}\begin{bmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&\cdots&k&&\\&&&\ddots&\vdots&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{bmatrix}&\begin{matrix}\\ \\ \text{第 } i \text{ 行}\\ \\ \text{第 } j \text{ 行}\\ \\ \\ \end{matrix}\end{array}$$

此时，
$$P_1^{\mathrm{T}}=\begin{array}{cl}\begin{bmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&\vdots&\ddots&&&\\&&k&\cdots&1&&\\&&&&&\ddots&\\&&&&&&1\end{bmatrix}&\begin{matrix}\\ \\ \text{第 } i \text{ 行}\\ \\ \text{第 } j \text{ 行}\\ \\ \\ \end{matrix}\end{array}$$

用 P_1 右乘 A，相当于把 A 的第 i 列的 k 倍加到第 j 列上；而用 P_1^{T} 左乘 A，相当于把 A 的第 i 行的 k 倍加到第 j 行上.

弄清楚了这 3 对初等矩阵所对应的初等列、行替换之间的关系后，只要每次成对地施行初等列、行替换，就可以保证所得到的矩阵与原来的矩阵合同.

例 3 用初等替换将对称矩阵

$$A=\begin{bmatrix}1&2&-2\\2&4&-1\\-2&-1&0\end{bmatrix}$$

化为与 A 合同的对角矩阵.

解：首先要把第 1 行的后两个元素化成零. 为此先从 A 的第 2 列减去第 1 列的 2 倍：

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

为了保证与 A 合同，再从这个矩阵的第 2 行减去第 1 行的 2 倍：

$$\begin{bmatrix}1&0&-2\\2&0&-1\\-2&3&0\end{bmatrix}\to\begin{bmatrix}1&0&-2\\0&0&3\\-2&3&0\end{bmatrix}=A_1.$$

A_1 是与 A 合同的，而且

$$A_1=P_1^{\mathrm{T}}AP_1,$$

其中 P_1 是一个初等矩阵：

$$P_1=\begin{bmatrix}1&-2&0\\0&1&0\\0&0&1\end{bmatrix}.$$

将 A_1 的第 1 列的 2 倍加到第 3 列上；再将所得到的矩阵的第 1 行的 2 倍加到第 3 行上：

$$A_1\to\begin{bmatrix}1&0&0\\0&0&3\\-2&3&-4\end{bmatrix}\to\begin{bmatrix}1&0&0\\0&0&3\\0&3&-4\end{bmatrix}=A_2.$$

由初等变换与初等矩阵的关系知：

$$A_2=P_2^{\mathrm{T}}A_1P_2,$$

其中
$$P_2=\begin{bmatrix}1&0&2\\0&1&0\\0&0&1\end{bmatrix}.$$

至此，A_2 的第 1 行及第 1 列的元素除去对角线上的元素以外，都等于零．再对下面的小矩阵$\begin{bmatrix}0&3\\3&-4\end{bmatrix}$进行变换，由于在 A_2 中，这个小矩阵的上方及左边都是零，所以对这个小矩阵进行初等变换时，这些零不会改变．

因为$\begin{bmatrix}0&3\\3&-4\end{bmatrix}$中左上方那个元素等于零，必须先将这个地方换成非零元素．这有两个办法：一是利用行、列互换，把 -4 换到左上角；另一个办法是将 3 加到左上角去．下面分别用这两个方法

对 A_2 进行初等变换：把 A_2 的第 2,3 列互换，再把所得到的矩阵的第 2,3 行互换：

$$A_2 \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & -4 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 3 \\ 0 & 3 & 0 \end{bmatrix} = A_3.$$

$$A_3 = P_3^{\mathrm{T}} A_2 P_3.$$

这里
$$P_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

或者，将 A_2 的第 3 列加到第 2 列上，再将所得到的矩阵的第 3 行加到第 2 行上：

$$A_2 \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 3 \\ 0 & -1 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & -4 \end{bmatrix} = B_3$$

$$B_3 = Q^{\mathrm{T}} A_2 Q.$$

这里
$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

一般说来，互换行列比较方便一些. 但是如果对角线上的元素都等于零时，就必须采用第二种方法了.

最后，将 A_3 的第 2 列的 $\frac{3}{4}$ 倍加到第 3 列上，再将第 2 行的 $\frac{3}{4}$ 倍加到第 3 行上：

$$A_3 \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 3 & \frac{9}{4} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & \frac{9}{4} \end{bmatrix} = A_4.$$

$$A_4 = P_4^{\mathrm{T}} A_3 P_4.$$

这里
$$P_4=\begin{bmatrix}1&0&0\\0&1&\frac{3}{4}\\0&0&1\end{bmatrix}.$$

于是
$$P_4^{\mathrm{T}}P_3^{\mathrm{T}}P_2^{\mathrm{T}}P_1^{\mathrm{T}}AP_1P_2P_3P_4=\begin{bmatrix}1&0&0\\0&-4&0\\0&0&\frac{9}{4}\end{bmatrix}.$$

令
$$P=P_1P_2P_3P_4,$$

即有
$$P^{\mathrm{T}}AP=\begin{bmatrix}1&0&0\\0&-4&0\\0&0&\frac{9}{4}\end{bmatrix}.$$

这里
$$P=P_1P_2P_3P_4=\begin{bmatrix}1&2&-\frac{1}{2}\\0&0&1\\0&1&\frac{3}{4}\end{bmatrix}.$$

上述的过程说明了：为了找 P，需要每次记下 P_i，还需将几个 P_i 相乘．虽然每个 P_i 都比较简单，但是乘到最后还是比较麻烦的．下面介绍一个不找出 P_i 而直接求出 P 的方法．基本的思路是与利用初等变换求逆矩阵的方法一样的．就用上面这个例题中的 P 来说明．将 P 写成：

$$P=P_1P_2P_3P_4=EP_1P_2P_3P_4.$$

由于用初等矩阵右乘，就相当于作一次初等列变换，所以矩阵 P 是由单位矩阵经过 4 次初等列变换而得到的．这 4 次列变换正好就是对 A 所作的 4 次列变换．所以当对 A 施行成对的行、列变换化成对角形时，同样的列变换施行于 E 就得到 P．当然，不必将每次所作的变换记下来，只要对 A 和 E 同时施行初等列变换．因为

所作的是列变换，故把 E 写在 A 下面. 下面仍用上述的例 3 来说明这个方法：

$$
\begin{bmatrix} A \\ E \end{bmatrix} = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 4 & -1 \\ -2 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[\text{第 1 列的 2 倍}]{\text{第 2 列减去}} \begin{bmatrix} 1 & 0 & -2 \\ 2 & 0 & -1 \\ -2 & 3 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

$$
\xrightarrow{\substack{\text{第 2 行减去} \\ \text{第 1 行的 2 倍}}} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 0 & 3 \\ -2 & 3 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[\text{加到第 3 列上}]{\text{第 1 列的 2 倍}}
$$

$$
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 3 \\ -2 & 3 & -4 \\ 1 & -2 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{\substack{\text{第 1 行的 2 倍} \\ \text{加到第 3 行上}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 3 \\ 0 & 3 & -4 \\ 1 & -2 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

$$
\xrightarrow[\text{列互换}]{\text{第 2、3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & -4 & 3 \\ 1 & 2 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{\substack{\text{第 2、3} \\ \text{行互换}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 3 \\ 0 & 3 & 0 \\ 1 & 2 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}
$$

$$\xrightarrow[\text{加到第 3 列上}]{\text{第 2 列的}\frac{3}{4}\text{倍}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 3 & \frac{9}{4} \\ 1 & 2 & -\frac{1}{2} \\ 0 & 0 & 1 \\ 0 & 1 & \frac{3}{4} \end{bmatrix} \xrightarrow[\text{加到第 3 行上}]{\text{第 2 行的}\frac{3}{4}\text{倍}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & \frac{9}{4} \\ 1 & 2 & -\frac{1}{2} \\ 0 & 0 & 1 \\ 0 & 1 & \frac{3}{4} \end{bmatrix},$$

所以
$$P = \begin{bmatrix} 1 & 2 & -\frac{1}{2} \\ 0 & 0 & 1 \\ 0 & 1 & \frac{3}{4} \end{bmatrix},$$

使得
$$P^{\mathrm{T}}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & \frac{9}{4} \end{bmatrix}.$$

这个例子说明了 P 的求法及其理论根据. 这种方法在求出 P 的同时,把 A 化成了对角形,而且不必写出每个 P_i,省去了很多计算.

因为化简二次型的问题可以通过化简它的矩阵来实现,所以上面介绍的方法也可以用来化简二次型. 举例如下:

例 4 设

$f(x_1,x_2,x_3;x_4)=x_1^2+x_2^2-2x_1x_2+4x_1x_4+2x_2x_3-2x_3x_4+2x_3x_4$,
用非退化线性替换将 $f(x_1,x_2,x_3,x_4)$ 化为标准形,并写出所作的线性替换.

解: $f(x_1,x_2,x_3,x_4)$的矩阵为

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

将 A 用初等变换化为合同的对角矩阵：

$$\begin{bmatrix}A\\ E\end{bmatrix}=\begin{bmatrix}1 & -1 & 0 & 2\\ -1 & 1 & 1 & -1\\ 0 & 1 & 0 & 1\\ 2 & -1 & 1 & 0\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\to\begin{bmatrix}1 & 0 & 0 & 0\\ -1 & 0 & 1 & 1\\ 0 & 1 & 0 & 1\\ 2 & 1 & 1 & -4\\ 1 & 1 & 0 & -2\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

$$\to\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 1 & 0 & 1\\ 0 & 1 & 1 & -4\\ 1 & 1 & 0 & -2\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\to\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 1 & 1\\ 0 & 1 & 0 & 1\\ 0 & 2 & 1 & -4\\ 1 & 1 & 0 & -2\\ 0 & 1 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

$$
\rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 1 & -4 \\ 1 & 1 & 0 & -2 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 2 & 0 & -6 \\ 1 & 1 & -\frac{1}{2} & -3 \\ 0 & 1 & -\frac{1}{2} & -1 \\ 0 & 1 & \frac{1}{2} & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

$$
\rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -6 \\ 1 & 1 & -\frac{1}{2} & -3 \\ 0 & 1 & -\frac{1}{2} & -1 \\ 0 & 1 & \frac{1}{2} & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
$$

令
$$P=\begin{bmatrix}1 & 1 & -\frac{1}{2} & -3\\ 0 & 1 & -\frac{1}{2} & -1\\ 0 & 1 & \frac{1}{2} & -1\\ 0 & 0 & 0 & 1\end{bmatrix},$$
则 P 可逆，并且
$$P^TAP=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & -\frac{1}{2} & 0\\ 0 & 0 & 0 & -6\end{bmatrix}.$$

所以原二次型通过线性替换
$$X=PY,$$
即
$$\begin{cases}x_1=y_1+y_2-\frac{1}{2}y_3-3y_4,\\ x_2=\qquad y_2-\frac{1}{2}y_3-y_4,\\ x_3=\qquad y_2+\frac{1}{2}y_3-y_4,\\ x_4=\qquad\qquad\qquad y_4,\end{cases}$$
化为标准形：
$$y_1^2+2y_2^2-\frac{1}{2}y_3^2-6y_4^2.$$

希望读者在作习题的时候，把配方法和用初等变换化对称矩阵为对角形这两种方法都练习一下.

习　题　5.3

用非退化线性替换化下列二次型为标准形，并写出所作的线性替换(用配方法和用初等变换法分别练习)：

1. $f(x_1,x_2,x_3)=x_1^2-2x_2^2+2x_1x_2-2x_1x_3$.

2. $f(x_1,x_2,x_3)=x_1^2-x_3^2+2x_1x_2+2x_2x_3$.

3. $f(x_1,x_2,x_3)=x_1x_2+x_1x_3+x_2x_3$.

4. $f(x_1,x_2,x_3,x_4)=x_1^2+2x_2^2+x_4^2+4x_1x_2+4x_1x_3+2x_1x_4+2x_2x_3+2x_2x_4$
$+2x_3x_4$.

5. $f(x_1,x_2,x_3,x_4)=x_1x_2+x_1x_3+x_1x_4+x_2x_3+x_2x_4+x_3x_4$.

5.4 规　范　形

一个二次型经过非退化线性替换化成标准形时，如果所作的替换不同，那么所得的标准形也可能不同. 例如，二次型

$$f(x_1,x_2,x_3) = x_1x_2 + x_2x_3 + x_1x_3,$$

经非退化线性替换

$$\begin{cases} x_1 = y_1 - y_2 - y_3, \\ x_2 = y_1 + y_2 - y_3, \\ x_3 = y_3, \end{cases}$$

化为
$$y_1^2 - y_2^2 - y_3^2;$$

而经非退化线性替换

$$\begin{cases} x_1 = z_2, \\ x_2 = 3z_1 - z_2 - 2z_3, \\ x_3 = 3z_1 - z_2 + 2z_3, \end{cases}$$

化为
$$9z_1^2 - z_2^2 - 4z_3^2.$$

这就是说，二次型的标准形不是唯一的.

但是，经过非退化线性替换后，二次型的矩阵变成了一个与它合同的矩阵. 我们知道，合同的矩阵有相同的秩，所以二次型变换前后的矩阵的秩是相同的. 二次型的矩阵的秩就称为这个**二次型的秩**.

由于标准形的矩阵是对角矩阵，对角矩阵的秩等于对角线上

非零元素的个数，因此，可一个二次型的标准形虽然不是唯一的，然而标准形中系数不等于零的平方项的项数却是相同的，就等于二次型的秩.

下面进一步在复数域和实数域中来讨论标准形的唯一性问题. 先看复数域的情形：

设 $f(x_1,x_2,\cdots,x_n)$是一个复系数的二次型，经过适当的非退化线性替换后，化成标准形

$$d_1y_1^2+d_2y_2^2+\cdots+d_ry_r^2\quad(d_i\neq 0,\ i=1,2,\cdots,r),\tag{1}$$

其中 r 是 $f(x_1,x_2,\cdots,x_n)$的秩. 因为复数可以开平方，故可再作一次非退化线性替换：

$$\begin{cases}y_1=\dfrac{1}{\sqrt{d_1}}z_1,\\ \cdots\cdots\cdots\\ y_r=\dfrac{1}{\sqrt{d_r}}z_r,\\ y_{r+1}=z_{r+1},\\ \cdots\cdots\cdots\\ y_n=z_n,\end{cases}$$

将(1)变成

$$z_1^2+z_2^2+\cdots+z_r^2.\tag{2}$$

(2) 称为复二次型 $f(x_1,x_2,\cdots,x_n)$的**规范形**. 显然，规范形完全由原二次型的秩所决定，因此有：

定理 4 复系数二次型可以经过适当的非退化线性替换变成规范形；规范形是唯一的.

定理 4 按矩阵来说，就是：任一复系数对称矩阵都合同于一个形式为：

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

的对角矩阵，其中 1 的个数等于这个矩阵的秩，因此，两个同阶复对称矩阵合同的充分必要条件是它们的秩相等.

再来看实数域的情形：

设 $f(x_1,x_2,\cdots,x_n)$是一个实系数二次型.经过一个非退化线性替换，再适当排列变量的次序(这也可看成作一次可逆线性替换)，可使 $f(x_1,x_2,\cdots,x_n)$变成标准形

$$d_1y_1^2+\cdots+d_py_p^2-d_{p+1}y_{p+1}^2-\cdots-d_ry_r^2, \tag{3}$$

其中 $d_i>0(i=1,2,\cdots,r)$，r 是 $f(x_1,x_2,\cdots,x_n)$的秩.因为在实数域中，正数可以开平方，而且其平方根不等于 0，所以可再作一次非退化线性替换

$$\begin{cases} y_1=\dfrac{1}{\sqrt{d_1}}z_1, \\ \cdots\cdots\cdots\cdots \\ y_r=\dfrac{1}{\sqrt{d_r}}z_r, \\ y_{r+1}=z_{r+1}, \\ \cdots\cdots\cdots\cdots \\ y_n=z_n, \end{cases}$$

将(3)就变成

$$z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2. \tag{4}$$

(4) 称为实系数二次型 $f(x_1,\cdots,x_n)$的**规范形**.显然，规范形由 r，

p 这两个数所决定.

定理 5(惯性定理) 任意一个实数域上的二次型,总可以经过非退化线性替换化成规范形;规范形是唯一的.

证明:上面的讨论已经证明了:任意一个实二次型都可以化成规范形.下面来证明规范形的唯一性:

设实二次型 $f(x_1,x_2,\cdots,x_n)$ 经过非退化线性替换

$$X = BY$$

化为规范形

$$f(x_1,x_2,\cdots,x_n) = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2;$$

而经过非退化线性替换

$$X = CZ,$$

化为规范形

$$f(x_1,x_2,\cdots,x_n) = z_1^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2.$$

要证明规范形是唯一的,就需要证明 $p=q$.

用反证法:设 $p>q$,根据上面的假设,有

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2=z_1^2+\cdots+z_q^2-z_{q+1}^2-\cdots-z_r^2, \tag{5}$$

其中

$$Z = C^{-1}BY. \tag{6}$$

设

$$C^{-1}B = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix},$$

于是(6)式成为

$$\begin{cases} z_1 = g_{11}y_1 + g_{12}y_2 + \cdots + g_{1n}y_n, \\ z_2 = g_{21}y_1 + g_{22}y_2 + \cdots + g_{2n}y_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ z_n = g_{n1}y_1 + g_{n2}y_2 + \cdots + g_{nn}y_n. \end{cases} \tag{7}$$

考察齐次线性方程组

$$
\begin{cases}
g_{11}y_1 + g_{12}y_2 + \cdots + g_{1n}y_n = 0, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
g_{q1}y_1 + g_{q2}y_2 + \cdots + g_{qn}y_n = 0, \\
y_{p+1} = 0, \\
\cdots\cdots\cdots \\
y_n = 0.
\end{cases} \tag{8}
$$

这个方程组含有 n 个未知量，而含有

$$q + (n - p) = n - (p - q) < n$$

个方程，所以它有非零解. 设

$$y_1 = k_1, y_2 = k_2, \cdots, y_p = k_p, y_{p+1} = k_{p+1}, \cdots, y_n = k_n,$$

是它的一个非零解. 显然

$$k_{p+1} = \cdots = k_n = 0,$$

所以 $k_1, k_2, \cdots, k_p$ 不全为零，因此，把 $y_i = k_i (i = 1, 2, \cdots, n)$ 代入(5)式左端，得到的值为

$$k_1^2 + k_2^2 + \cdots + k_p^2 > 0.$$

另一方面，因为 $y_i = k_i (i = 1, 2, \cdots, n)$ 是方程组(8)的解，故把它们代入关系式(7)，得到一组 $z_i (i = 1, 2, \cdots, n)$，其中

$$z_1 = z_2 = \cdots = z_q = 0.$$

将这一组 z_i 代入(5)式右端，得到的值是

$$- z_{q+1}^2 - \cdots - z_r^2 \leqslant 0.$$

这是一个矛盾，这说明 $p > q$ 是不可能的.

同理，可证明 $q > p$ 也是不可能的，因而 $p = q$. 这就证明了规范形的唯一性. |

这个定理通常称为惯性定理.

定义 3 在实系数二次型 $f(x_1, x_2, \cdots, x_n)$ 的规范形中，正平方项的个数 p 称为 $f(x_1, x_2, \cdots, x_n)$ 的**正惯性指数**；负平方项的个数 $r - p$ 称为 $f(x_1, x_2, \cdots, x_n)$ 的**负惯性指数**；它们的差 $p - (r - p) = 2p - r$ 称为 $f(x_1, x_2, \cdots, x_n)$ 的**符号差**.

虽然实二次型的标准形不是唯一的，但是根据规范形的唯一性以及上面化标准形为规范形的过程可以看出：标准形中系数为正的平方项个数是唯一确定的，它等于正惯性指数；标准形中系数为负的平方项的个数也是唯一确定的，它等于负惯性指数. 因此，两个实系数二次型可以经非退化线性替换互变的充分必要条件是：它们有相同的秩和正惯性指数.

定理 5 用矩阵的语言描述是：任一个实对称矩阵 A 合同于一个形式为

$$\begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{bmatrix}.$$

的对角矩阵. 其中 ± 1 的总数等于 A 的秩，1 的个数由 A 唯一确定，称为 A 的正惯性指数，-1 的个数称为 A 的负惯性指数. 两个同阶的实对称矩阵合同的充分必要条件是：它们的秩和正惯性指数分别相等.

习　题　5.4

1. 把习题 5.3 中所有二次型化为规范形(分实系数及复系数两种情形)，并写出所作的线性替换.
2. 如果把 n 阶实对称矩阵按合同关系分类(即两个 n 阶实对称矩阵属于同一类当且仅当它们是合同的)，问共有几类?
3. 试证：一个实二次型可以分解成两个实系数一次多项式的乘积的充分必

要条件是：它的秩等于2，而且符号差等于0，或者秩等于1.

4. 设实二次型 $f(x_1,x_2,\cdots,x_n)$ 的正、负惯性指数分别是 k,l；$a_1,a_2,\cdots,a_k$ 是任意 k 个正数；$b_1,b_2,\cdots,b_l$ 是任意 l 个负数. 试证：$f(x_1,x_2,\cdots,x_n)$ 可经非退化线性替换化成：

$$a_1y_1^2+\cdots+a_ky_k^2+b_1y_{k+1}^2+\cdots+b_ly_{k+l}^2.$$

5.5 正定二次型

在实二次型中，正定二次型有着特殊的地位. 这一节介绍正定二次型的定义及判别条件.

定义 4 实二次型 $f(x_1,x_2,\cdots,x_n)$ 如果对于任意一组不全为零的实数 $c_1,c_2,\cdots,c_n$，都有 $f(c_1,c_2,\cdots,c_n)>0$，就称为**正定的**.

下面讨论实二次型是否正定的判别方法. 首先，如果 $f(x_1,x_2,\cdots,x_n)$ 是平方和

$$f(x_1,x_2,\cdots,x_n)=a_1x_1^2+a_2x_2^2+\cdots+a_nx_n^2.$$

那么，很容易看出：$f(x_1,x_2,\cdots,x_n)$ 是正定的必要充分条件是 $a_i>0(i=1,2,\cdots,n)$. 我们知道，一般地，任一个二次型 $f(x_1,x_2,\cdots,x_n)$ 都可以经过非退化线性替换化为平方和的形式

$$f(x_1,x_2,\cdots,x_n)=b_1y_1^2+b_2y_2^2+\cdots+b_ny_n^2.$$

能不能利用 $b_i(i=1,2,\cdots,n)$ 的正负来判断 $f(x_1,x_2,\cdots,x_n)$ 是否正定呢？答案是肯定的. 对此，有下述定理：

定理 6 n 元实二次型 $f(x_1,x_2,\cdots,x_n)$ 是正定的充分必要条件是：它的正惯性指数等于 n.

证明：设 $f(x_1,x_2,\cdots,x_n)$ 经非退化线性替换

$$X=CY \tag{1}$$

化为

$$f(x_1,x_2,\cdots,x_n)=b_1y_1^2+b_2y_2^2+\cdots+b_ny_n^2. \tag{2}$$

如果 $f(x_1,x_2,\cdots,x_n)$ 的正惯性指数等于 n. 那么，$b_i>0(i=1,

$2,\cdots,n$). 设 $c_1,c_2,\cdots,c_n$ 是任意 n 个不全为零的实数. 由于线性替换(1)是非退化的，所以可找到 $y_i(i=1,2,\cdots,n)$ 的一组值

$$y_i = k_i \quad (i = 1,2,\cdots,n)$$

使

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = C \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}.$$

显然 $k_i(i=1,2,\cdots,n)$ 不全为零.

将 $x_i=c_i(i=1,2,\cdots,n)$ 代入(2)式左边；将 $y_i=k_i(i=1,2,\cdots,n)$ 代入(2)式右边，得到的值应该相等，因此

$$f(c_1,c_2,\cdots,c_n) = b_1k_1^2 + b_2k_2^2 + \cdots + b_nk_n^2 > 0,$$

所以 $f(x_1,x_2,\cdots,x_n)$ 是正定的.

如果 $f(x_1,x_2,\cdots,x_n)$ 的正惯性指数小于 n，那么，$b_1,b_2,\cdots,b_n$ 不能全大于零，不妨设 $b_n\leqslant 0$. 令

$$y_1 = y_2 = \cdots = y_{n-1} = 0, y_n = 1,$$

代入(1)，得 x_i 的一组值

$$x_i = c_i \quad (i = 1,2,\cdots,n).$$

因为线性替换(1)是可逆的，所以 $c_i(i=1,2,\cdots,n)$ 不全为零，将 x_i 及 y_i 相应的值代入(2)式两边，得：

$$f(c_1,c_2,\cdots,c_n) \leqslant 0.$$

所以 $f(x_1,x_2,\cdots,x_n)$ 不是正定的. |

定理 6 说明：

推论 正定二次型的规范形是

$$y_1^2 + y_2^2 + \cdots + y_n^2.$$

定义 5 A 是一个实对称矩阵，如果实二次型

$$X^{\mathrm{T}}AX$$

是正定的，则 A 称为**正定矩阵**.

因为正定二次型的规范形的矩阵是单位矩阵 E，所以一个实对称矩阵是正定的充分必要条件是它与单位矩阵是合同的. 即有：

定理 7 实对称矩阵 A 是正定矩阵的充分必要条件是：有可逆矩阵 C，使

$$A = C^{\mathrm{T}}C$$

证明作为习题留给读者.

我们知道，任给一个实二次型

$$f(x_1, x_2, \cdots, x_n) = X^{\mathrm{T}}AX,$$

可以找到一个正交变换

$$X = TY,$$

把 $f(x_1, x_2, \cdots, x_n)$ 化为平方和

$$f(x_1, x_2, \cdots, x_n) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,$$

式中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是矩阵 A 的特征值. 由此可得：

定理 8 实二次型

$$X^{\mathrm{T}}AX$$

正定的充分必要条件是：A 的特征值全大于零.

当然，相应地有：实对称矩阵 A 正定的充分必要条件是：它的特征值全大于零.

可以利用行列式来说明二次型或实对称矩阵正定的条件，先证明下述引理.

引理 正定矩阵的行列式大于零.

证明：设 A 是一个正定矩阵. 那么 A 与单位矩阵 E 合同，所以有可逆矩阵 C，使

$$A = C^{\mathrm{T}}EC = C^{\mathrm{T}}C.$$

两边取行列式，就得：

$$|A| = |C^{\mathrm{T}}C| = |C^{\mathrm{T}}||C| = |C|^2 > 0. \quad |$$

定义 6 设 $A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$是一个 n 阶矩阵. 行标和列标相同的子式

$$\begin{vmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \cdots & a_{i_1 i_k} \\ a_{i_2 i_1} & a_{i_2 i_2} & \cdots & a_{i_2 i_k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_k i_1} & a_{i_k i_2} & \cdots & a_{i_k i_k} \end{vmatrix} \quad (1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n)$$

称为 A 的**主子式**;其中,主子式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{21} & a_{22} & \cdots & a_{2i} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{vmatrix} \quad (i=1,2,\cdots,n)$$

称为 A 的**顺序主子式**.

例如,设 $A=\begin{bmatrix} 1 & -1 & 2 & 3 \\ -1 & 0 & -1 & 1 \\ 2 & -1 & 2 & 0 \\ 3 & 1 & 0 & -1 \end{bmatrix}$,

那么 $1,2,\begin{vmatrix} 1 & -1 \\ -1 & 0 \end{vmatrix},\begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix},\begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix},$

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

都是 A 的主子式;而 A 的顺序主子式共有 4 个,即

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

定理 9 实二次型

$$f(x_1,x_2,\cdots,x_n) = X^TAX = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$$

是正定的充分必要条件是：矩阵 A 的顺序主子式全大于零.

证明：先证必要性：设二次型

$$f(x_1,x_2,\cdots,x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$$

是正定的. 令

$$f_k(x_1,x_2,\cdots,x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \qquad (k=1,2,\cdots,n).$$

下面证明 $f_k(x_1,x_2,\cdots,x_k)$是 k 元的正定二次型. 对于任意 k 个不全为零的实数 $c_1,c_2,\cdots,c_k$，有

$$\begin{aligned} f_k(c_1,c_2,\cdots,c_k) &= \sum_{i=1}^{k}\sum_{j=1}^{k}a_{ij}c_ic_j \\ &= f(c_1,\cdots,c_k,0,\cdots,0) > 0. \end{aligned}$$

所以 $f_k(x_1,x_2,\cdots,x_k)$是正定的，从而 $f_k(x_1,x_2,\cdots,x_k)$的矩阵

$$A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$$

是正定的，根据引理知：$|A_k|>0$. 容易看出：$|A_k|$就是 A 的第 k 个顺序主子式. 这就证明了必要性.

再来证明充分性：对 n 作数学归纳法：

当 $n=1$ 时，

$$f(x_1) = a_{11}x_1^2.$$

由条件 $a_{11}>0$，显然 $f(x_1)$是正定的.

假设充分性的论断对 $n-1$ 元的二次型是成立的，下面来证明 n 元的情形. 令

$$f_{n-1}(x_1, x_2, \cdots, x_{n-1}) = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} a_{ij}x_ix_j.$$

它的矩阵是

$$A_{n-1} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} \end{bmatrix}$$

A_{n-1}的顺序主子式就是 A 的前 $n-1$ 个顺序主子式. 现在 A 的顺序主子式全大于零，所以 A_{n-1}的顺序主子式也全大于零. 由归纳法假设知：$f_{n-1}(x_1, x_2, \cdots, x_{n-1})$是正定的，因此有可逆的线性替换

$$\begin{cases} x_1 = g_{11}y_1 + g_{12}y_2 + \cdots + g_{1,n-1}y_{n-1}, \\ x_2 = g_{21}y_1 + g_{22}y_2 + \cdots + g_{2,n-1}y_{n-1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_{n-1} = g_{n-1,1}y_1 + g_{n-1,2}y_2 + \cdots + g_{n-1,n-1}y_{n-1}, \end{cases}$$

使 $$f_{n-1}(x_1, x_2, \cdots, x_{n-1}) = y_1^2 + y_2^2 + \cdots + y_{n-1}^2.$$

作线性替换

$$\begin{cases} x_1 = g_{11}y_1 + g_{12}y_2 + \cdots + g_{1,n-1}y_{n-1}, \\ x_2 = g_{21}y_1 + g_{22}y_2 + \cdots + g_{2,n-1}y_{n-1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_{n-1} = g_{n-1,1}y_1 + g_{n-1,2}y_2 + \cdots + g_{n-1,n-1}y_{n-1}, \\ x_n = y_n. \end{cases}$$

这是一个 n 元非退化的线性替换，而且 $f(x_1, x_2, \cdots, x_n)$经过这个线性替换化成：

$$
\begin{aligned}
f(x_1,x_2,\cdots,x_n) &= \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}a_{ij}x_ix_j + 2\sum_{i=1}^{n-1}a_{in}x_ix_n + a_{nn}x_n^2 \\
&= y_1^2 + y_2^2 + \cdots + y_{n-1}^2 + 2\sum_{i=1}^{n-1}b_{in}y_iy_n + a_{nn}y_n^2 \\
&= (y_1 + b_{1n}y_n)^2 + (y_2 + b_{2n}y_n)^2 + \cdots \\
&\quad + (y_{n-1} + b_{n-1,n}y_n)^2 + b_{nn}y_n^2.
\end{aligned}
$$

再令
$$
\begin{cases}
y_1 + b_{1n}y_n = z_1, \\
y_2 + b_{2n}y_n = z_2, \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
y_{n-1} + b_{n-1,n}y_n = z_{n-1}, \\
y_n = z_n,
\end{cases}
$$

这也是一个可逆的线性替换，$f(x_1,x_2,\cdots,x_n)$再经过这个线性替换后，化成

$$f(x_1,x_2,\cdots,x_n) = z_1^2 + z_2^2 + \cdots + z_{n-1}^2 + b_{nn}z_n^2.$$

$f(x_1,x_2,\cdots,x_n)$经过非退化线性替换化成 $z_1^2+\cdots+z_{n-1}^2+b_{nn}z_n^2$，因此矩阵 A 与矩阵

$$
\begin{bmatrix}
1 & & & \\
& \ddots & & \\
& & 1 & \\
& & & b_{nn}
\end{bmatrix}
$$

是合同的，故有可逆矩阵 C，使

$$
C^TAC = \begin{bmatrix}
1 & & & \\
& \ddots & & \\
& & 1 & \\
& & & b_{nn}
\end{bmatrix}.
$$

两边取行列式，即得

$$b_{nn} = |C^TAC| = |C|^2 \cdot |A| > 0,$$

所以 $b_{nn}>0$. $f(x_1,x_2,\cdots,x_n)$的正惯性指数等于 n,从而 $f(x_1,x_2,\cdots,x_n)$是正定的.

根据归纳法原理,充分性得证. |

从定理 9 知道:实对称矩阵 A 是正定的当且仅当它的顺序主子式全大于零.

例 1 判别实二次型

$$f(x_1,x_2,x_3)=x_1^2+2x_1x_2+2x_2^2+4x_2x_3+x_3^2$$

是否正定.

解:
$$\begin{aligned}f(x_1,x_2,x_3)&=x_1^2+2x_1x_2+2x_2^2+4x_2x_3+x_3^2\\&=(x_1+x_2)^2+(x_2+2x_3)^2-3x_3^2.\end{aligned}$$

令
$$\begin{cases}x_1+x_2=y_1,\\x_2+2x_3=y_2,\\x_3=y_3,\end{cases}$$

这是一个非退化线性替换.经过这个替换,$f(x_1,x_2,x_3)$化为

$$f(x_1,x_2,x_3)=y_1^2+y_2^2-3y_3^2,$$

所以 $f(x_1,x_2,x_3)$的正惯性指数等于 2,由定理 6 知:$f(x_1,x_2,x_3)$不是正定的.

例 2 判别二次型

$$f(x_1,x_2,x_3)=3x_1^2+4x_2^2+5x_3^2+4x_1x_2-4x_2x_3$$

是否正定.

解: $f(x_1,x_2,x_3)$的矩阵为

$$A=\begin{bmatrix}3&2&0\\2&4&-2\\0&-2&5\end{bmatrix}.$$

A 的顺序主子式

$$3>0,\begin{vmatrix}3&2\\2&4\end{vmatrix}=8>0,\begin{vmatrix}3&2&0\\2&4&-2\\0&-2&5\end{vmatrix}=28>0.$$

根据定理 9 知，$f(x_1,x_2,x_3)$是正定的.

与实二次型的正定性相仿的，有下述一些概念：

定义 7 设 $f(x_1,x_2,\cdots,x_n)$是一个实二次型，如果对于任意一组不全为零的实数 $c_1,c_2,\cdots,c_n$，都有 $f(c_1,c_2,\cdots,c_n)<0$，就称 $f(x_1,x_2,\cdots,x_n)$是**负定**的；如果对任意一组实数 $c_1,c_2,\cdots,c_n$，都有 $f(c_1,c_2,\cdots,c_n)\geqslant 0$，就称 $f(x_1,x_2,\cdots,x_n)$是**半正定**的；如果对任意一组实数 $c_1,c_2,\cdots,c_n$，都有 $f(c_1,c_2,\cdots,c_n)\leqslant 0$，就称 $f(x_1,x_2,\cdots,x_n)$是**半负定**的；如果 $f(x_1,x_2,\cdots,x_n)$既不是半正定的，又不是半负定的，就称它是**不定的**.

设 A 是实对称矩阵，如果二次型 $X^{\mathrm{T}}AX$ 是负定的，就称 A 是负定的；如果 $X^{\mathrm{T}}AX$ 是半正定的或半负定的，就称 A 是半正定的或半负定的.

显然，如果 $f(x_1,x_2,\cdots,x_n)$是负定的，那么$-f(x_1,x_2,\cdots,x_n)$就是正定的，因此可推得负定的二次型的判别条件如下：

实二次型 $f(x_1,x_2,\cdots,x_n)$是负定的充分必要条件是：它的负惯性指数等于 n. 因此，负定实二次型的规范形是

$$-y_1^2-y_2^2-\cdots-y_n^2.$$

从定理 8 可推知：实二次型是负定的，当且仅当它的系数矩阵的特征值全小于零.

从定理 9 可以推出用行列式判别一个二次型是不是负定的方法. 设 $D_i(i=1,2,\cdots,n)$表示实二次型 $f(x_1,x_2,\cdots,x_n)$的顺序主子式. 那么 $f(x_1,x_2,\cdots,x_n)$是负定的充分必要条件是：它的顺序主子式满足

$$(-1)^iD_i>0\quad(i=1,2,\cdots,n).$$

根据半正定二次型的定义，可以得到与正定条件类似的判别法. 下面只写出结果，证明作为习题留给读者.

实二次型 $f(x_1,x_2,\cdots,x_n)$是半正定的充分必要条件是 $f(x_1,x_2,\cdots,x_n)$的正惯性指数等于它的秩；或 $f(x_1,x_2,\cdots,x_n)$的系数矩

阵的特征值全大于或等于零；或 $f(x_1,x_2,\cdots,x_n)$ 的系数矩阵的主子式全大于或等于零.

需要注意的是，一个实二次型必须当它的系数矩阵的主子式全大于或等于零时才是半正定的. 只有顺序主子式大于或等于零是不够的. 下面举例说明这一点：

例 3 二次型

$$f(x_1,x_2,x_3) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 4x_1x_3 + 4x_2x_3$$

的矩阵是
$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}.$$

它的顺序主子式全大于或等于零：

$$1 > 0, \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0, \begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 2 & 1 \end{vmatrix} = 0.$$

但是
$$f(x_1,x_2,x_3) = (x_1 + x_2 + 2x_3)^2 - 3x_3^2$$

不是半正定的.

实二次型是不是半负定的也有类似的条件，这里就不仔细讲了.

关于实对称矩阵是不是负定的或半正(负)定的也都有与实二次型相类似的判别法则，读者自己把这些条件讨论一下.

习 题 5.5

1. 判别下列二次型是否正定：

(1) $5x_1^2+6x_2^2+4x_3^2-4x_1x_2-4x_2x_3$；

(2) $99x_1^2-12x_1x_2+48x_1x_3+130x_2^2-60x_2x_3+71x_3^2$；

(3) $10x_1^2+8x_1x_2+24x_1x_3+2x_2^2-28x_2x_3+x_3^2$；

(4) $x_1^2+x_2^2+4x_3^2+7x_4^2+6x_1x_3+4x_1x_4-4x_2x_3+2x_2x_4+4x_3x_4$.

2. t 满足什么条件时，下列二次型是正定的：

(1) $f(x_1,x_2,x_3)=x_1^2+x_2^2+5x_3^2+2tx_1x_2-2x_1x_3+4x_2x_3$;

(2) $f(x_1,x_2,x_3)=x_1^2+4x_2^2+x_3^2+2tx_1x_2+10x_1x_3+6x_2x_3$.

3. 试证：如果 A 是正定矩阵，那么 A 的主子式全大于零.

4. 试证：如果 A 是正定矩阵，那么 A^{-1}和 A^* 也都是正定矩阵.

5. 试证：如果 A,B 都是 n 阶正定矩阵，那么 $A+B$ 也是正定矩阵.

6. 试证：实二次型 $f(x_1,x_2,\cdots,x_n)$是半正定的充分必要条件是 $f(x_1,x_2,\cdots,x_n)$的正惯性指数等于它的秩.

7. 试证：实二次型 $f(x_1,x_2,\cdots,x_n)=X^TAX$ 是半正定的充分必要条件是：A 的特征值全大于或等于零.

内 容 提 要

1. 二次型及其矩阵

(1) 二次型：

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j \qquad (a_{ij}=a_{ji})$$
$$=X^TAX;$$

$A=A^T$称为二次型 $f(x_1,x_2,\cdots,x_n)$的矩阵.

(2) 线性替换、非退化(可逆)线性替换、正交替换.

(3) 二次型 $f(x_1,x_2,\cdots,x_n)=X^TAX$ 经线性替换

$$X=CY,$$

化成
$$f(x_1,x_2,\cdots,x_n)=Y^TBY,$$

则
$$B=C^TAC.$$

(4) 矩阵的合同关系.

2. 二次型的标准形

(1) 主要结果：

数域 P 上任一个二次型可以经过系数在数域 P 中的非退化线性替换化成标准形

$$d_1y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2.$$

其中,不等于零的系数的个数等于这个二次型的矩阵的秩(也称这个二次型的秩).

(2) 计算方法:

配方法、初等变换的方法.

3. 规范形

(1) 复系数二次型可以经过非退化线性替换化成规范形

$$y_1^2 + y_2^2 + \cdots + y_r^2.$$

规范形是唯一的.

(2) 实系数二次型可以经过实系数的非退化线性替换化成规范形

$$y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2,$$

规范形是唯一的(惯性定理).

规范形中正系数的个数 p 称为 $f(x_1,x_2,\cdots,x_n)$的正惯性指数;负系数的个数 $r-p$ 称为负惯性指数,它们的差 $p-(r-p)=2p-r$ 称为符号差.

4. 用正交替换化实二次型为平方和

(1) 主要结果:

任一个实二次型 $f(x_1,x_2,\cdots,x_n)=X^TAX$ 可经过正交替换化成平方和

$$f(x_1,x_2,\cdots,x_n) = \lambda_1y_1^2 + \lambda_2y_2^2 + \cdots + \lambda_ny_n^2,$$

其中,$\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 A 的特征多项式的全部根.

(2) 计算方法:

求正交矩阵 T,使 $T^{-1}AT$ 为对角形:

$$T^{-1}AT = T^TAT = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

则正交替换 $X=TY$

将 $f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX$ 化成平方和：

$$f(x_1,x_2,\cdots,x_n)=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2.$$

5. 正定二次型

(1) 正定二次型的定义：

实二次型 $f(x_1,x_2,\cdots,x_n)$如果对于任意一组不全为零的实数 $c_1,c_2,\cdots,c_n$，都有 $f(c_1,c_2,\cdots,c_n)>0$，就称 $f(x_1,x_2,\cdots,x_n)$为正定的.

(2) 正定二次型的判别条件：

下列条件都是实二次型 $f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX$ 为正定二次型的充分必要条件：

1) $f(x_1,x_2,\cdots,x_n)$的正惯性指数等于 n；

2) A 的顺序主子式全大于零；

3) A 的特征值全大于零.

(3) 正定矩阵.

(4) 其他一些概念：

负定、半正定、半负定、不定二次型.

6. 关于对称矩阵的相应定义及结论

复习题5

1. 用非退化线性替换把下列二次型化为标准形和规范形(分实数域、复数域两种情况)，并写出所作的线性替换

(1) $f(x_1,x_2,x_3)=x_1^2+2x_2^2+5x_3^2+2x_1x_2+2x_1x_3+8x_2x_3$；

(2) $f(x_1,x_2,x_3)=2x_1x_2+4x_1x_3$；

(3) $f(x_1,x_2,x_3,x_4)=x_1x_2+x_1x_3-x_1x_4-x_2x_3+x_2x_4+x_3x_4$；

(4) $f(x_1,x_2,x_3,x_4)=x_1^2+5x_2^2+4x_3^2-x_4^2+6x_1x_2-4x_1x_3-4x_2x_4$
$+6x_3x_4$.

2. 写出第 1 题的矩阵结果.

3. 用正交线性替换把下列实二次型化为标准形，写出所作的正交替换和该二次型的正负惯性指数及符号差.

(1) $f(x_1,x_2,x_3)=x_1^2-2x_2^2-2x_3^2-4x_1x_2-4x_1x_3+8x_2x_3$;

(2) $f(x_1,x_2,x_3)=2x_1x_2+2x_3x_4$;

(3) $f(x_1,x_2,x_3,x_4)=x_1^2+x_2^2+x_3^2+x_4^2-2x_1x_2+6x_1x_3-4x_1x_4-4x_2x_3$
$+6x_2x_4-2x_3x_4$.

4. 判断下列实二次型是否正定：

(1) $f(x_1,x_2,x_3)=5x_1^2+x_2^2+5x_3^2+4x_1x_2-8x_1x_3-4x_2x_3$;

(2) $f(x_1,x_2,x_3)=x_1^2+2x_2^2-3x_3^2+4x_1x_2-2x_2x_3$.

5. 求 t，使下列实二次型是正定二次型：

(1) $f(x_1,x_2,x_3)=x_1^2+4x_2^2+2x_3^2+2tx_1x_2+2x_1x_3$;

(2) $f(x_1,x_2,x_3)=x_1^2+2x_2^2+3x_3^2+2tx_1x_2+2x_2x_3$.

6. 设实二次型

$$f(x_1,x_2,\cdots,x_n)=l_1^2+\cdots+l_s^2-l_{s+1}^2-\cdots-l_{s+t}^2,$$

其中，$l_i(i=1,2,\cdots,s+t)$是 $x_1,x_2,\cdots,x_n$ 的实系数一次齐次多项式. 试证：$f(x_1,x_2,\cdots,x_n)$的正惯性指数$\leqslant s$；负惯性指数$\leqslant t$.

7. A 是一个实对称矩阵，试证：A 是正定矩阵的充分必要条件是有实可逆矩阵 C 使 $A=C^TC$.

8. A 是一个实对称矩阵，证明：A 是半正定矩阵的充分必要条件是有实矩阵 C 使 $A=C^TC$.

9. 设 A 是实对称矩阵，试证：t 充分大之后，$tE+A$ 是正定矩阵.

10. 设 A 是一个实对称矩阵，且$|A|<0$，试证：必有实 n 维向量 X，使 $X^TAX<0$.

11. 设 $f(x_1,x_2,\cdots,x_n)=X^TAX$ 是一个实二次型，有实 n 维向量 X_1,X_2，使

$$X_1^TAX_1>0,X_2^TAX_2<0.$$

试证：必存在实 n 维向量 $X_0\neq 0$，使 $X_0^TAX_0=0$.

12. 设 A,B 是两个 n 阶实对称矩阵，并且 A 是正定的，试证：存在一个 n 阶实可逆矩阵 T，使 T^TAT 及 T^TBT 都是对角矩阵.

习题答案与提示

第1章　行列式

习　题　1.1

1. (1) 22；　(2) $a^2+2ab-b^2$；　(3) -25；

(4) 56；　(5) 0；　(6) $3abc-a^3-b^3-c^3$.

2. (1) $x_1=\dfrac{19}{3}$，$x_2=\dfrac{4}{3}$；　(2) $x_1=3$，$x_2=-2$，$x_3=-1$；

(3) $x_1=-\dfrac{11}{8}$，$x_2=-\dfrac{9}{8}$，$x_3=-\dfrac{6}{8}$；　(4) $x_1=2$，$x_2=-1$，$x_3=3$.

习　题　1.2(1)

1.

排　列	逆序数
2　4　1　3　5	3
2　4　1　5　3	4
2　4　3　1　5	3
2　4　3　5　1	4
2　4　5　1　3	5
2　4　5　3　1	6

2. 11;22;15.

习　题　1.2(2)

1. 12，偶；6，偶；16，偶；13，奇.

2. (1) $i=3$，$j=8$；　(2) $i=6$，$j=8$.

3. 3 1 5 6 9 4 2 7 8 $\xrightarrow{(3,1)}$ 1 3 5 6 9 4 2 7 8

$\xrightarrow{(3,2)}$ 1 2 5 6 9 4 3 7 8 $\xrightarrow{(5,3)}$ 1 2 3 6 9 4 5 7 8

$\xrightarrow{(6,4)}$ 1 2 3 4 9 6 5 7 8 $\xrightarrow{(9,5)}$ 1 2 3 4 5 6 9 7 8

$\xrightarrow{(9,7)}$ 1 2 3 4 5 6 7 9 8 $\xrightarrow{(9,8)}$ 1 2 3 4 5 6 7 8 9.

是奇排列.

习 题 1.3

1. (1) $-$; (2) $+$; (3) $-$.
2. $a_{13}a_{25}a_{31}a_{44}a_{52}$, $a_{13}a_{25}a_{32}a_{41}a_{54}$, $a_{13}a_{25}a_{34}a_{42}a_{51}$.
3. (1) $i=3$, $k=5$; (2) $i=1$, $k=5$.
4. (1) 120; (2) $a_1a_2a_3a_4a_5$; (3) $(a_1b_2-a_2b_1)(c_1d_2-c_2d_1)$.
5. (1) $(-1)^{n-1}n!$; (2) $(-1)^{(n-1)(n-2)/2}n!$.

习 题 1.4(1)

1. (1) 8015800; (2) -29400000; (3) $-2(x^3+y^3)$;
   (4) 160; (5) a^2b^2; (6) 0.
3. $[a+(n-1)b](a-b)^{n-1}$.
4. $n!$.
5. $(n+1)a_1a_2\cdots a_nb$.

习 题 1.4(2)

1. 1. 2. 0. 3. -42. 4. 780. 5. $-9\frac{3}{4}$.

习 题 1.5

1. $M_{11}=3, M_{12}=12, M_{13}=12, M_{14}=21, M_{21}=6, M_{22}=-4, M_{23}=10, M_{24}=14, M_{31}=3, M_{32}=-2, M_{33}=-16, M_{34}=7, M_{41}=12, M_{42}=6, M_{43}=6, M_{44}=0$;

   $A_{11}=3, A_{12}=-12, A_{13}=12, A_{14}=-21, A_{21}=-6, A_{22}=-4, A_{23}=-10, A_{24}=14, A_{31}=3, A_{32}=2, A_{33}=-16, A_{34}=-7, A_{41}=-12, A_{42}=6, A_{43}=$

$-6, A_{44}=0.$

2. $A_{11}=-1, A_{21}=1, A_{31}=2, A_{41}=2.$

3. $A_{11}=-1, A_{21}=1, A_{31}=2, A_{41}=2.$

4. (1) $a_1a_2\cdots a_n\sum_{i=1}^{n}\frac{1}{a_i}$;

(2) $(-1)^{\frac{n(n-1)}{2}}a_1a_2\cdots a_n\sum_{i=1}^{n}\frac{1}{a_i}$.

习 题 1.6

1. (1) 900; (2) 483; (3) -16; (4) 72.

2. -72.

3. $-2\cdot(n-2)!$.

4. $1+(-1)^{n+1}$.

5. $(n+1)a^n$.

6. $(a^2-b^2)^n$.

复习题 1

1. (1) $x_1=\frac{7}{13}, x_2=-\frac{16}{13}, x_3=-\frac{4}{13}$;

(2) $x_1=\frac{(b-d)(c-d)}{(b-a)(c-a)}, x_2=\frac{(d-a)(c-d)}{(b-a)(c-b)}, x_3=\frac{(d-a)(d-b)}{(c-a)(c-b)}$.

2. (1) 24; (2) $\frac{22}{3}$; (3) 5.

3. (1) $(ab+1)(cd+1)+ad$; (2) 0.

4. $\frac{1}{4}(n+2)!$.

5. $(-1)^{n-1}\cdot\frac{1}{2}(n+1)!$.

6. $(-1)^{n-1}(n+1)2^{n-2}$.

7. $a_1a_2\cdots a_n\left(a_0-\sum_{i=1}^{n}\frac{1}{a_i}\right)$.

8. $D=\begin{cases}(-1)^k, & \text{当 } n=3k \text{ 或 } 3k+1;\\ 0, & \text{当 } n=3k+2.\end{cases}$

第 2 章　线性方程组

习　题　2.1

1. (1) $(2,4,-1,-3)$；　(2) $(1,1,-1,2)$；　(3) $(1,2,-1,-2)$；
   (4) $(-1,1,-1,1,-1)$.

习　题　2.2(1)

1. (1) $(1,1,1,1)$；　(2) $(1,-1,0,2)$；　(3) $(0,2,-2,0,3)$；
   (4) $(1,-2,3,-2,1)$.

习　题　2.2(2)

1.　(1) $\begin{cases} x_1=2x_2-x_3 \\ x_4=1, \end{cases}$

其中 x_2,x_3 是自由未知量；

(2) 无解；

(3) $x_1=1,x_2=2,x_3=-2$；

(4) $\begin{cases} x_1=\dfrac{1}{3}(1-x_5) \\ x_2=\dfrac{1}{6}(-4-3x_4+5x_5) \\ x_3=0, \end{cases}$

其中 x_4,x_5 是自由未知量；

(5) $x_1=-8\dfrac{2}{3},x_2=-3\dfrac{2}{3},\ x_3=1\dfrac{2}{3},\ x_4=-6,\ x_5=-1\dfrac{1}{3}$；

(6) $x_1=x_2=x_3=x_4=0$.

2. $\lambda=5$ 时有解，一般解为：

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

其中 x_3,x_4 为自由未知量.

3. 当 $a=1$ 或 $b=0$ 时，方程组有非零解.

当 $a=1$ 时，一般解为：

$$\begin{cases} x_1 = -x_3 \\ x_2 = 0, \end{cases}$$

其中 x_3 为自由未知量；

当 $b=0$ 时，一般解为：

$$\begin{cases} x_1 = -x_3 \\ x_2 = (a-1)x_3, \end{cases}$$

其中 x_3 为自由未知量.

4. $a=0,b=2$ 时有解，一般解为：

$$\begin{cases} x_1 = x_3 + x_4 + 5x_5 - 2 \\ x_2 = -2x_3 - 2x_4 - 6x_5 + 3, \end{cases}$$

其中 x_3,x_4,x_5 为自由未知量.

5. 一般解为：

$$\begin{cases} x_1 = a_1 + a_2 + a_3 + a_4 + x_5 \\ x_2 = a_2 + a_3 + a_4 + x_5 \\ x_3 = a_3 + a_4 + x_5 \\ x_4 = a_4 + x_5 \end{cases}$$

其中 x_5 是自由未知量.

习　题　2.3

1. (1) 不是； (2) 不是； (3) 是； (4) 是.

习　题　2.4

1. $-\alpha=(-1,0,1,-2)$；$2\alpha=(2,0,-2,4)$；$\alpha-\beta=(-2,-2,-5,3)$；

$5\alpha+4\beta=(17,8,11,6)$.

2. $\gamma=(-3,7,-17,2,-8)$.

3. $\alpha=\left(\frac{5}{2},\frac{1}{2},3,\frac{1}{2},2\right)$，$\beta=\left(-\frac{1}{2},\frac{1}{2},2,\frac{3}{2},-2\right)$

4. $\alpha=(10,-5,-9,2)$，$\beta=(-7,4,7,-1)$.

5. $(a_1,a_2,\cdots,a_n)$

习　题　2.5(1)

1. (1) $\beta=\frac{5}{4}\alpha_1+\frac{1}{4}\alpha_2-\frac{1}{4}\alpha_3-\frac{1}{4}\alpha_4$;

   (2) $\beta=-\alpha_1+\alpha_2+2\alpha_3-2\alpha_4$;

   (3) $\beta=-\alpha_1+\alpha_2-\alpha_3+\frac{1}{2}\alpha_4$.

2. 不一定.

5. 提示:用 $A_{ij}(i,j=1,2,\cdots,n)$表示行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{vmatrix}$$

   的元素 a_{ij}的代数余子式,那么

$$\alpha_i=\frac{A_{1i}}{D}\beta_1+\frac{A_{2i}}{D}\beta_2+\cdots+\frac{A_{mi}}{D}\beta_m \quad (i=1,2,\cdots,m).$$

习　题　2.5(2)

1. (1) 线性无关; (2) 线性相关; (3) 线性相关; (4) 线性无关.

2. 提示:证明满足 $k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r=0$ 的数 $k_1,k_2,\cdots,k_r$ 必须全等于 0.

5. 用反证法证明:如果 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性相关,那么有不全为零的数 $k_1,k_2,\cdots,k_s$,使 $k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0$.

设 $k_s,k_{s-1},\cdots,k_1$ 中第一个不等于零的数是 $k_l(1\leqslant l\leqslant s)$,即 $k_l\neq0$,而 $k_{l+1}=\cdots=k_s=0$. 上式可写成

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_l\alpha_l=0.$$

而且,其中 $k_l\neq0$. 因为 $\alpha_1\neq0$,所以 $l\neq1$. 于是

$$\alpha_l=-\frac{k_1}{k_l}\alpha_1-\frac{k_2}{k_l}\alpha_2-\cdots-\frac{k_{l-1}}{k_l}\alpha_{l-1}.$$

这说明 α_l 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_{l-1}$线性表出,这与题中假设(2)相矛盾. 所以 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是线性无关的.

习　题　2.5(3)

1. (1) 秩＝2；α_1,α_2 是一个极大线性无关组(极大线性无关组不是唯一的，这里只列出一个.)

 (2) 秩＝4；$\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 是极大线性无关组.

 (3) 秩＝2；α_1,α_2 是一个极大线性无关组.

2. 提示：证明 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 满足极大线性无关组的条件.
3. 提示：应用习题 2.5(2)第 3 题.
4. 提示：应用上题.
5. 提示：应用定理 7 推论 1.
6. 提示：应用第 3 题证明 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 的极大线性无关组也是 $\alpha_1,\alpha_2,\cdots,\alpha_t,\alpha_{t+1},\cdots,\alpha_s$ 的极大线性无关组.

习　题　2.6

1. (1) 矩阵 A 共有 5 个 4 阶子式，全等于 0；　(2) $r(A)=3$.
2. (1) 2；　(2) 4.
3. (1) 秩＝2，线性相关；　(2) 秩＝3，线性相关.
4. α_3,α_4 的取法很多，例如可取 $\alpha_3=(1,0,0,0)$，$\alpha_4=(0,1,0,0)$

习　题　2.7(1)

1. 无解.
2. (1) 当 $a\neq\pm1$ 时，方程组有唯一解：

$$x_1=\frac{4a+1}{(a-1)(a+1)},x_2=\frac{a(2a-7)}{(a-1)(a+1)},x_3=\frac{-3a}{a+1};$$

 (2) 当 $a=1$ 时，方程组有无穷多解，一般解为：

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

其中 x_2 为自由未知量；

 (3) 当 $a=-1$ 时，方程组无解.

3. $\lambda=1,3$ 时有非零解.

 当 $\lambda=1$ 时，一般解为：

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

x_3 为自由未知量.

当 $\lambda=3$ 时,一般解为:

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

x_3 为自由未知量.

习 题 2.7(2)

1. (1) 基础解系:(1,−2,1,0,0),(1,−2,0,1,0),(5,−6,0,0,1),

全部解:$k_1(1,-2,1,0,0)+k_2(1,-2,0,1,0)+k_3(5,-6,0,0,1)$;$k_1$,$k_2$,$k_3$ 是数域 **P** 中任意数;

(2) 基础解系:(0,1,1,0,0),(0,1,0,1,0),$\left(\frac{1}{3},-\frac{5}{3},0,0,1\right)$

全部解:$k_1(0,1,1,0,0)+k_2(0,1,0,1,0)+k_3\left(\frac{1}{3},-\frac{5}{3},0,0,1\right)$,$k_1$,$k_2$,$k_3$ 是数域 **P** 中任意数;

(3) 基础解系:(0,1,2,1),

全部解:$k(0,1,2,1)$,k 是数域 **P** 中任意数;

(4) 基础解系:(1,−1,0,0,0),(1,0,−1,0,0),(1,0,0,−1,0),(1,0,0,0,−1),

全部解:$k_1(1,-1,0,0,0)+k_2(1,0,-1,0,0)+k_3(1,0,0,-1,0)+k_4(1,0,0,0,-1)$,$k_1$,$k_2$,$k_3$,$k_4$ 是数域 **P** 中任意数.

(注意,答案不是唯一的,但同一个齐次方程组的基础解系必等价.)

习 题 2.7(3)

1. (1) $(1,0,1,0)+k(3,-3,1,-2)$,k 为数域 **P** 中任意数;

(2) $\left(\frac{13}{6},\frac{5}{6},0,\frac{1}{3},0\right)+k_1(-1,1,1,0,0)+k_2(7,5,0,2,6)$,$k_1$,$k_2$ 是数域 **P** 中任意数;

(3) $(1,2,-1,-2,0)+k(0,-20,14,-26,11)$,$k$ 是数域 **P** 中任意数;

(4) $\left(\frac{5}{16},\frac{3}{4},\frac{23}{16},0,0\right)+k_1\left(\frac{9}{8},\frac{-3}{2},\frac{19}{8},1,0\right)+k_2\left(\frac{-1}{16},\frac{1}{4},\frac{5}{16},0,1\right)$，$k_1,k_2$ 是数域 $\mathbf{P}$ 中任意数.

复习题 2

1. (1) $\left(\frac{15}{4},-\frac{5}{4},-\frac{1}{4},0,0\right)+k_1(-1,3,-2,1,0)+k_2(9,-11,5,0,4)$，$k_1,k_2$ 是数域 $\mathbf{P}$ 中任意数；

(2) $(1,2,3,-1,-2,-3)+k_1(0,-1,1,1,0,0)+k_2(-8,14,5,0,15,0)+k_3(-2,-1,2,0,0,3)$，$k_1,k_2,k_3$ 是数域 $\mathbf{P}$ 中任意数；

(3) $(1,-2,0,0)+k_1(-9,1,7,0)+k_2(1,-1,0,2)$，$k_1,k_2$ 是数域 $\mathbf{P}$ 中任意数；

(4) $(1,0,0,0,0)+k_1(-2,1,0,0,0)+k_2(-3,0,1,0,0)+k_3(-4,0,0,1,0)+k_4(-5,0,0,0,1)$，$k_1,k_2,k_3,k_4$ 是数域 $\mathbf{P}$ 中任意数.

2. (1) 当 $a\neq1,b\neq0$ 时有唯一解：

$$x_1=\frac{2b-1}{b(a-1)},\ x_2=\frac{1}{b},\ x_3=\frac{2ab-4b+1}{b(a-1)};$$

当 $a=1,b=\frac{1}{2}$ 时，有无穷多解：

$$(2,2,0)+k(1,0,-1),$$

k 是数域 $\mathbf{P}$ 中任意数；

(2) 当 $a=0,b=-4$ 时，有无穷多解：

$\left(-2,\frac{3}{2},0,0,0\right)+k_1(1,-1,1,0,0)+k_2(1,-1,0,1,0)+k_3(-1,0,0,0,1)$

k_1,k_2,k_3 为数域 $\mathbf{P}$ 中任意数.

4. 提示：证明 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 等价.

5. 提示：应用上题.

6. 提示：把线性方程组看成向量方程组，再应用上题.

7. 提示：将 $\alpha_i(i=1,2,\cdots,t)$ 表成 $\beta_1,\beta_2,\cdots,\beta_t$ 的线性组合.

8. 提示：应用上题.

10. 提示：应用第 8 题.

第3章 矩 阵

习 题 3.1(1)

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

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

3. $\begin{bmatrix} 1-2a & c & -3 \\ 2a-2b & 3b-2c+a & 3-a+b \\ 1 & -2-2a & -3+2c \end{bmatrix}$.

习 题 3.1(2)

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

(3) $\begin{bmatrix} a^2+b^2+c^2 & ac+ab+bc & a+b+c \\ ac+ab+bc & a^2+b^2+c^2 & a+b+c \\ a+b+c & a+b+c & 3 \end{bmatrix}$；

(4) $\begin{bmatrix} 2 & 13\frac{1}{2}+1\frac{1}{2}\mathrm{i} & 1+2\mathrm{i} \\ 1+\mathrm{i} & 5+5\mathrm{i} & 1+\mathrm{i} \\ -\mathrm{i} & 2\mathrm{i} & 3 \end{bmatrix}$.

2. $\begin{bmatrix} 4 & 1 & 11 \\ 7 & -8 & -3 \\ -4 & 14 & 8 \end{bmatrix}$.

3. $AB=\begin{bmatrix} 12 & 10 & 2 \\ 0 & 2 & 1 \\ 8 & 10 & 3 \end{bmatrix}$； $(AB)C=\begin{bmatrix} 62 & 2 & 18 \\ 7 & -2 & -3 \\ 53 & -2 & 7 \end{bmatrix}$；

$$BC=\begin{bmatrix}15 & -1 & 1\\ 8 & 1 & 4\end{bmatrix},\quad A(BC)=\begin{bmatrix}62 & 2 & 18\\ 7 & -2 & -3\\ 53 & -2 & 7\end{bmatrix}.$$

4. (1) 例如,令 $A=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$, $B=\begin{bmatrix}1 & -1\\ 1 & -1\end{bmatrix}$,则 $(AB)^2=A^2B^2=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$

(2) 例如,令 $A=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$, $B=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$,则 $(AB)^2=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$,

$A^2B^2=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}$,所以 $(AB)^2\neq A^2B^2$.

5. (1) $\begin{bmatrix}7 & 4 & 4\\ 9 & 4 & 3\\ 3 & 3 & 4\end{bmatrix}$; (2) $\begin{bmatrix}41 & -38\\ 38 & 41\end{bmatrix}$; (3) $\begin{bmatrix}1 & n & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$;

(4) 当 $n=2k+1$: $\begin{bmatrix}4^k & -4^k & -4^k & -4^k\\ -4^k & 4^k & -4^k & -4^k\\ -4^k & -4^k & 4^k & -4^k\\ -4^k & -4^k & -4^k & 4^k\end{bmatrix}$;

当 $n=2k$: $\begin{bmatrix}4^k & 0 & 0 & 0\\ 0 & 4^k & 0 & 0\\ 0 & 0 & 4^k & 0\\ 0 & 0 & 0 & 4^k\end{bmatrix}$;

6. (1) $\begin{bmatrix}x & y-x\\ 0 & y\end{bmatrix}$, 其中 x,y 可以是数域 **P** 中的任意数.

(2) $\begin{bmatrix}a & b & c\\ c & a & b\\ b & c & a\end{bmatrix}$ (a,b,c 任意); (3) $\begin{bmatrix}a & 0 & 0\\ 0 & b & 0\\ 0 & 0 & c\end{bmatrix}$ (a,b,c 任意).

习　题　3.1(3)

1. $A^{\mathrm{T}}=\begin{bmatrix}1 & 2 & -1\\ 2 & 3 & 0\\ -1 & 2 & 2\end{bmatrix}$, $B^{\mathrm{T}}=\begin{bmatrix}0 & 2 & -1\\ 1 & -1 & -1\\ 2 & 0 & 3\end{bmatrix}$,

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

$$AB=\begin{bmatrix}5&0&-1\\4&-3&10\\-2&-3&4\end{bmatrix},\ BA=\begin{bmatrix}0&3&6\\0&1&-4\\-6&-5&5\end{bmatrix},$$

$$A^TB^T=\begin{bmatrix}0&0&-6\\3&1&-5\\6&-4&5\end{bmatrix},\ B^TA^T=\begin{bmatrix}5&4&-2\\0&-3&-3\\-1&10&4\end{bmatrix},$$

$$A^2=\begin{bmatrix}6&8&1\\6&13&8\\-3&-2&5\end{bmatrix},\ (A^T)^2=\begin{bmatrix}6&6&-3\\8&13&-2\\1&8&5\end{bmatrix}.$$

2. $(ABC)^T=C^TB^TA^T=\begin{bmatrix}82&140\\31&42\end{bmatrix}$.

3. (1) $\begin{bmatrix}4&4&4\\9&3&-10\\-3&5&6\end{bmatrix}$； (2) O.

习 题 3.2

1. $\begin{bmatrix}3&6&21&30&36\\-3&9&18&38&46\\-9&6&-15&-22&-28\\0&0&0&13&2\\0&0&0&25&-5\end{bmatrix}$

3. (1) 例如，取 $A=\begin{bmatrix}1&0\\0&0\end{bmatrix}$，$B=\begin{bmatrix}0&0\\0&1\end{bmatrix}$；

(2) 例如，取 $A=\begin{bmatrix}1&0\\0&1\end{bmatrix}$，$B=\begin{bmatrix}0&0\\0&1\end{bmatrix}$.

4. (1) 例如，取 $A=\begin{bmatrix}1&0\\0&1\end{bmatrix}$，$B=\begin{bmatrix}1&0\\0&0\end{bmatrix}$；

(2) 例如，取 $A=\begin{bmatrix}0&0\\0&1\end{bmatrix}$，$B=\begin{bmatrix}1&0\\0&0\end{bmatrix}$.

5. 提示：$AB=0$ 的充分必要条件是矩阵 B 的每个列都是齐次线性方程组 $AX=0$ 的解.

6. 提示：同上题.

习 题 3.3

1. (1) $A^*=\begin{bmatrix}2 & -1\\ 0 & 3\end{bmatrix}$; (2) $A^*=\begin{bmatrix}5 & 9 & -1\\ -2 & -3 & 0\\ 0 & 2 & -1\end{bmatrix}$;

(3) $A^*=\begin{bmatrix}26 & -10 & -1\\ 3 & 3 & 3\\ 5 & -4 & 5\end{bmatrix}$.

2. (1) $\begin{bmatrix}\dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc}\\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc}\end{bmatrix}$;

(2) $\begin{bmatrix}\dfrac{4}{5} & \dfrac{1}{10} & -\dfrac{11}{10}\\ \dfrac{7}{5} & -\dfrac{1}{5} & -\dfrac{9}{5}\\ -\dfrac{6}{5} & \dfrac{1}{10} & \dfrac{19}{10}\end{bmatrix}$; (3) $\begin{bmatrix}-\dfrac{17}{27} & -\dfrac{1}{27} & \dfrac{44}{27}\\ \dfrac{10}{27} & -\dfrac{1}{27} & -\dfrac{10}{27}\\ -\dfrac{1}{9} & \dfrac{1}{9} & \dfrac{1}{9}\end{bmatrix}$;

(4) $\dfrac{1}{4}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\end{bmatrix}$; (5) $\begin{bmatrix}1 & -2 & 1 & 0\\ 0 & 1 & -2 & 1\\ 0 & 0 & 1 & -2\\ 0 & 0 & 0 & 1\end{bmatrix}$.

3. 提示：验证$(E-A)(E+A+A^2+\cdots+A^{k-1})=E$.

4. $(A^*)^{-1}=\dfrac{1}{|A|}A$.

5. 提示：由 $AB=BA$ 可得 $BA^{-1}=A^{-1}B$.

6. (1) $\begin{bmatrix}-17 & -28\\ -4 & -6\end{bmatrix}$; (2) $\begin{bmatrix}2 & 9 & -5\\ -2 & -8 & 6\\ -4 & -14 & 9\end{bmatrix}$;

(3) $\begin{bmatrix}-13 & -75 & 30\\ 9 & 52 & -21\\ 21 & 120 & -47\end{bmatrix}$; (4) $\begin{bmatrix}a & b\\ a-2 & b-5\\ -2a+1 & -2b-4\end{bmatrix}$ (a,b 任意).

7. (2) $\begin{bmatrix}-3&2&0&0\\-5&3&0&0\\0&0&-1&4\\0&0&1&-3\end{bmatrix}$; $\begin{bmatrix}0&0&1&-1&1\\0&0&0&1&-1\\0&0&0&0&1\\-3&2&0&0&0\\2&-1&0&0&0\end{bmatrix}$.

习 题 3.4

1. $\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$, $\begin{bmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$, $\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$,

$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix}$, $\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$.

2. 提示：应用矩阵等价关系的对称性，传递性.

3. 提示：应用上题.

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

5. (1) $\dfrac{1}{21}\begin{bmatrix}-3&23&22\\6&-32&-37\\3&-2&-1\end{bmatrix}$; (2) $\begin{bmatrix}-2&-7&-2&9\\-2&-6&-1&7\\\frac{4}{5}&3&\frac{4}{5}&-\frac{18}{5}\\1&3&1&-4\end{bmatrix}$;

(3) $\begin{bmatrix} -\frac{95}{99} & \frac{4}{11} & -\frac{28}{99} \\ -\frac{20}{33} & \frac{6}{11} & \frac{8}{33} \\ \frac{20}{3} & 0 & \frac{4}{3} \end{bmatrix}$； (4) $\begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{6} & -\frac{1}{6} \\ \frac{2}{3} & -\frac{5}{3} & \frac{7}{6} & -\frac{1}{6} \\ \frac{4}{3} & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{6} & \frac{1}{6} \end{bmatrix}$.

习 题 3.5

1. (1) 是； (2) 是； (3) 不是； (4) 是.

4. (1) -9； (2) 0.

5. (1) $|\alpha|=\sqrt{26}$； (2) $|\alpha|=\sqrt{30}$.

7. 提示：α 与 α 正交.

9. $\left[\frac{\sqrt{3}}{3},0,\frac{\sqrt{3}}{3},\frac{\sqrt{3}}{3}\right]$，$\left[\frac{2\sqrt{33}}{33},\frac{\sqrt{33}}{11},\frac{2\sqrt{33}}{33},\frac{-4\sqrt{33}}{33}\right]$，

$\left[\frac{-7\sqrt{110}}{110},\frac{3\sqrt{110}}{55},\frac{2\sqrt{110}}{55},\frac{3\sqrt{110}}{110}\right]$（注：答案不唯一.）

10. 这样的正交矩阵很多，下面举出两个：

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{2} & -\frac{5}{6} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 & 0 \\ \frac{\sqrt{2}}{3} & \frac{\sqrt{2}}{3} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{6} \end{bmatrix},$$

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{2} & -\frac{5}{6} \\ 0 & -\frac{2\sqrt{26}}{13} & \frac{3\sqrt{26}}{26} & \frac{\sqrt{26}}{26} \\ -\frac{\sqrt{26}}{6} & \frac{5\sqrt{26}}{78} & \frac{\sqrt{26}}{13} & \frac{\sqrt{26}}{39} \end{bmatrix}.$$

复习题3

1. (1) $\frac{1}{35}\begin{bmatrix} 77 & 13 \\ 28 & 42 \\ -49 & 24 \end{bmatrix}$;　(2) $\frac{1}{35}\begin{bmatrix} -79 & -4 & 16 \\ 33 & 3 & 23 \end{bmatrix}$;

(3) $\frac{1}{7}\begin{bmatrix} 2 & -37 & -8 \\ -1 & -34 & -6 \\ 3 & -38 & -6 \end{bmatrix}$;　(4) $\frac{1}{20}\begin{bmatrix} -28 & 18 & -26 \\ 16 & 9 & 17 \\ -36 & 26 & -22 \end{bmatrix}$.

7. (1) $A=\frac{1}{2}(A+A^{\mathrm{T}})+\frac{1}{2}(A-A^{\mathrm{T}})$;

(2) $\begin{bmatrix} 1 & 1 & \frac{3}{2} \\ 1 & 2 & 2 \\ \frac{3}{2} & 2 & 5 \end{bmatrix}+\begin{bmatrix} 0 & 1 & \frac{3}{2} \\ -1 & 0 & 2 \\ -\frac{3}{2} & -2 & 0 \end{bmatrix}$.

9. 提示：作一个可逆矩阵 B，使 $AB=0$.

10. 提示：$AA^*=|A|\cdot E$.

11. 提示：应用 $AA^*=|A|\cdot E$，分 A 为可逆或不可逆两种情形进行讨论.

12. 提示：同上题.

13. 提示：如果 $A^2=E$，那么 $A^2-E=(A-E)(A+E)=0$，应用习题 3.2 第 6 题及定理 1.

14. 提示：如果 $A^2=A$，则 $A(A-E)=0$，余同上题.

17. (1) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$；(2) $\begin{bmatrix} 2(1-2^{k-1}) & 10(2^k-1) & 6(2^k-1) \\ 2^k-1 & 5-2^{k+2} & 3(1-2^k) \\ 2(1-2^k) & 10(2^k-1) & 7\cdot 2^k-6 \end{bmatrix}$.

18. (1) 提示：将 A 的列正交化，单位化，得正交矩阵 T. 考虑 A 与 T 的关系；

(2) $$T=\begin{bmatrix} \frac{\sqrt{3}}{3} & 0 & \frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \end{bmatrix},$$

$$P=\begin{bmatrix} \sqrt{3} & \sqrt{3} & 0 \\ 0 & \sqrt{2} & \frac{\sqrt{2}}{2} \\ 0 & 0 & \frac{\sqrt{6}}{2} \end{bmatrix}$$

第 4 章　矩阵的对角化问题

习　题　4.1

3. 证明：因为 $A\sim B, C\sim D$，所以有可逆矩阵 X, Y，使 $X^{-1}AX=B, Y^{-1}CY=D$. 矩阵 $\begin{bmatrix} X & O \\ O & Y \end{bmatrix}$ 也可逆，而且 $\begin{bmatrix} X & O \\ O & Y \end{bmatrix}^{-1}=\begin{bmatrix} X^{-1} & O \\ O & Y^{-1} \end{bmatrix}$，于是

$$\begin{bmatrix} X & O \\ O & Y \end{bmatrix}^{-1}\begin{bmatrix} A & O \\ O & C \end{bmatrix}\begin{bmatrix} X & O \\ O & Y \end{bmatrix}=\begin{bmatrix} X^{-1} & O \\ O & Y^{-1} \end{bmatrix}\begin{bmatrix} A & O \\ O & C \end{bmatrix}\begin{bmatrix} X & O \\ O & Y \end{bmatrix}$$

$$=\begin{bmatrix} X^{-1}AX & O \\ O & Y^{-1}CY \end{bmatrix}=\begin{bmatrix} B & O \\ O & D \end{bmatrix}.$$

所以 $\begin{bmatrix} A & O \\ O & C \end{bmatrix}\sim\begin{bmatrix} B & O \\ O & D \end{bmatrix}$.

习　题　4.2

1. (1) 特征值：7，−2；

属于特征值 7 的特征向量：

$$k\begin{bmatrix}1\\1\end{bmatrix},$$

其中 k 为任意非零复数

属于特征值—2 的特征向量：

$$l\begin{bmatrix}4\\-5\end{bmatrix},$$

其中 l 为任意非零复数.

(2) 特征值：—1(三重)；

特征向量：

$$k\begin{bmatrix}1\\1\\-1\end{bmatrix},$$

其中 k 为任意非零复数.

(3) 特征值：1,i,—i；

属于 1 的特征向量：

$$k_1\begin{bmatrix}2\\-1\\-1\end{bmatrix},$$

其中 k_1 是任意非零复数；

属于 i 的特征向量：

$$k_2\begin{bmatrix}-1+2i\\1-i\\2\end{bmatrix},$$

其中 k_2 是任意非零复数；

属于—i 的特征向量：

$$k_3\begin{bmatrix}-1-2i,\\1+i\\2\end{bmatrix},$$

其中 k_3 是任意非零复数.

(4) 特征值；1,1,—1；

属于 1 的特征向量：

$$k_1\begin{bmatrix}0\\1\\0\end{bmatrix}+k_2\begin{bmatrix}1\\0\\1\end{bmatrix},$$

其中 k_1,k_2 是不全为零的任意复数；

属于 -1 的特征向量：

$$l\begin{bmatrix}1\\0\\-1\end{bmatrix},$$

其中 l 是任意非零复数.

(5) 特征值：$0,\sqrt{14}\mathrm{i},-\sqrt{14}\mathrm{i}$；

属于 0 的特征向量：

$$k_1\begin{bmatrix}3\\-1\\2\end{bmatrix},$$

其中 k_1 是任意非零复数；

属于 $\sqrt{14}\mathrm{i}$ 的特征向量：

$$k_2\begin{bmatrix}-6-\sqrt{14}\mathrm{i}\\2-3\sqrt{14}\mathrm{i}\\10\end{bmatrix},$$

其中 k_2 是任意非零复数；

属于 $-\sqrt{14}\mathrm{i}$ 的特征向量：

$$k_3\begin{bmatrix}-6-\sqrt{14}\mathrm{i}\\2+3\sqrt{14}\mathrm{i}\\10\end{bmatrix},$$

其中 k_3 为任意非零复数.

习　题　4.3

1. (1) 能，令 $X=\begin{bmatrix}1&4\\1&-5\end{bmatrix}$，则 $X^{-1}AX=[7,-2]$（X 的取法不是唯一的）；

(2) 不能；

(3) 能，令

$$X=\begin{bmatrix} 2 & -1+2\mathrm{i} & -1-2\mathrm{i} \\ -1 & 1-\mathrm{i} & 1+\mathrm{i} \\ -1 & 2 & 2 \end{bmatrix}，则 X^{-1}AX=[1,\mathrm{i},-\mathrm{i}]；$$

(4) 能，令 $X=\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$，则 $X^{-1}AX=[1,1,-1]$；

(5) 能，令 $X=\begin{bmatrix} 3 & -6-\sqrt{14}\mathrm{i}, & -6+\sqrt{14}\mathrm{i} \\ -1 & 2-3\sqrt{14}\mathrm{i} & 2+3\sqrt{14}\mathrm{i} \\ 2 & 10 & 10 \end{bmatrix}$，则

$X^{-1}AX=[0,\sqrt{14}\mathrm{i},-\sqrt{14}\mathrm{i}]$.

2. (2) 提示：用反证法.

4. $\begin{bmatrix} 5 & -1 & -2 \\ 16 & 4 & -6 \\ 2 & 0 & -1 \end{bmatrix}$.

5. $\dfrac{1}{3}\begin{bmatrix} (-1)^k\cdot 2+5^k & (-1)^{k+1}+5^k & (-1)^{k+1}+5^k \\ (-1)^{k+1}+5^k & (-1)^k\cdot 2+5^k & (-1)^{k+1}+5^k \\ (-1)^{k+1}+5^k & (-1)^{k+1}+5^k & (-1)^k\cdot 2+5^k \end{bmatrix}$.

习　题　4.4

1. (1) 令 $T=\begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$，则 $T^{-1}AT=[4,1,-2]$；

(2) 令 $T=\begin{bmatrix} \frac{2}{5}\sqrt{5} & \frac{2}{15}\sqrt{5} & \frac{1}{3} \\ -\frac{\sqrt{5}}{5} & \frac{4}{15}\sqrt{5} & \frac{2}{3} \\ 0 & \frac{\sqrt{5}}{3} & -\frac{2}{3} \end{bmatrix}$，则 $T^{-1}AT=[2,2,-7]$；

(3) 令 $T=\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & \frac{2}{3} \\ 0 & -\frac{2\sqrt{2}}{3} & \frac{1}{3} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6} & \frac{2}{3} \end{bmatrix}$，则 $T^{-1}AT=[-3,-3,6]$；

(4) 令 $T=\begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{2}}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$，则 $T^{-1}AT=[3,3,5,1]$；

(5) 令 $T=\begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{bmatrix}$，

则 $T^{-1}AT=[2,2,0,0]$

（T 的取法不是唯一的）.

5．提示：参照定理 6 的证明.

习 题 4.5

1. $\begin{bmatrix}\lambda_0^k & k\lambda_0^{k-1}\\ & \lambda_0^k\end{bmatrix}$; $\begin{bmatrix}\lambda_0^k & k\lambda_0^{k-1} & \frac{1}{2}k(k-1)\lambda_0^{k-2}\\ & \lambda_0^k & k\lambda_0^{k-1}\\ & & \lambda_0^k\end{bmatrix}$.

2. 提示：$\begin{bmatrix}2 & 1\\ -1 & 0\end{bmatrix}$的特征值 $\lambda=1,1$,但不能与对角矩阵相似.

3. $\begin{bmatrix}1 & 0\\ -1 & 1\end{bmatrix}$(答案不唯一).

4. $\begin{bmatrix}k+1 & k\\ -k & 1-k\end{bmatrix}$.

复习题 4

1. (1) 特征值 1,1,2;属于特征值 1 的特征向量:$k\begin{pmatrix}1\\0\\0\end{pmatrix}$,

k 为数域 P 中任意非零数.

属于特征值 2 的全部特征向量:

$$l\begin{pmatrix}0\\0\\1\end{pmatrix},$$

l 为不等于零的任意数.

(2) 特征值：2,2,11;

属于特征值 2 的特征向量：

$$k_1\begin{pmatrix}1\\-2\\0\end{pmatrix}+k_2\begin{pmatrix}1\\0\\-1\end{pmatrix},$$

其中 k_1,k_2 为不全为零的任意数.

属于特征值 11 的特征向量：

$$l\begin{pmatrix}2\\1\\2\end{pmatrix},$$

l 为任意非零数.

(3) A 的特征值为：$-1,9,0$；

属于 -1 的特征向量：

$$k_1\begin{pmatrix}1\\-1\\0\end{pmatrix},$$

k 为任意非零数.

属于 9 的特征向量：

$$k_2\begin{pmatrix}1\\1\\2\end{pmatrix}$$

k_2 是任意非零数.

属于特征值 0 的特征向量：

$$k_3\begin{pmatrix}1\\1\\-1\end{pmatrix},$$

k_3 是任意非零数.

(4) 特征值：1,1,2；

属于特征值 1 的特征向量：

$$k\begin{pmatrix}1\\2\\-1\end{pmatrix}$$

k 是任意非零数.

属于特征值 2 的特征向量：

$$l\begin{pmatrix}0\\0\\1\end{pmatrix},$$

l 是任意非零数.

2. (2) 令 $X=\begin{bmatrix}1 & 1 & 2\\ -2 & 0 & 1\\ 0 & -1 & 2\end{bmatrix}$，则 $X^{-1}AX=\begin{bmatrix}2 & & \\ & 2 & \\ & & 11\end{bmatrix}$；

(3) 令 $X=\begin{bmatrix}1 & 1 & 1\\ -1 & 1 & 1\\ 0 & 2 & -1\end{bmatrix}$，则 $X^{-1}AX=\begin{bmatrix}-1 & & \\ & 9 & \\ & & 0\end{bmatrix}$.

3. (1) $T=\frac{1}{3}\begin{bmatrix}2 & 2 & 1\\ -2 & 1 & 2\\ -1 & 2 & -2\end{bmatrix}$，$T^{-1}AT=\begin{bmatrix}1 & & \\ & 1 & \\ & & 10\end{bmatrix}$；

(2) $T=\begin{bmatrix}-\frac{2\sqrt{5}}{5} & \frac{2\sqrt{5}}{15} & \frac{1}{3}\\ \frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} & \frac{2}{3}\\ 0 & \frac{\sqrt{5}}{3} & -\frac{2}{3}\end{bmatrix}$，$T^{-1}AT=\begin{bmatrix}1 & & \\ & 1 & \\ & & -8\end{bmatrix}$；

(3) $T=\begin{bmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{2}\\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & -\frac{1}{2}\\ 0 & -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{6} & \frac{1}{2}\\ 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2}\end{bmatrix}$，

$T^{-1}AT=\begin{bmatrix}4 & & & \\ & 4 & & \\ & & 4 & \\ & & & -8\end{bmatrix}$；

(4) $T=\begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{\sqrt{2}}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$, $T^{-1}AT=\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & 3 \end{bmatrix}$.

4. 提示：A 的全部特征值的乘积等于 $|A|$.

5. 提示：把 A 化为上三角形.

7. 提示：利用上题.

8. 提示：证明 A 可对角化，并且 A 的特征值全相等.

9. 提示：应用第 7 题.

10. 提示：应用定理 7.

第 5 章　二 次 型

习　题　5.1

2. (1) $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 0 & -2 & 1 \\ -2 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$;

(2) $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 1 & 1 & -\frac{1}{2} \\ 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$;

(3) $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$;

(4) $\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$.

3. 提示:分别令

$$X=\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\text{第 } i \text{ 个}(1\leqslant i\leqslant n) \text{ 及 } X=\begin{array}{l}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\end{array}\begin{array}{l} \\ \\ \\ \text{第 } i \text{ 个} \\ \\ \\ \\ \text{第 } j \text{ 个} \\ \\ \\ \\ \end{array}(1\leqslant i<j\leqslant n).$$

代入 $X^{T}AX=0$.

习 题 5.2

1. 正交替换 $\begin{cases} x_1=\dfrac{2\sqrt{5}}{5}y_1+\dfrac{2\sqrt{5}}{15}y_2+\dfrac{1}{3}y_3, \\ x_2=-\dfrac{\sqrt{5}}{5}y_1+\dfrac{4\sqrt{5}}{15}y_2+\dfrac{2}{3}y_3, \\ x_3=\dfrac{\sqrt{5}}{3}y_2-\dfrac{2}{3}y_3, \end{cases}$

将原二次型化为 $y_1^2+y_2^2+10y_3^2$.(所作的正交替换不是唯一的.)

2. 正交替换 $\begin{cases} x_1=\dfrac{\sqrt{2}}{2}y_1+\dfrac{\sqrt{2}}{6}y_2+\dfrac{2}{3}y_3, \\ x_2=\dfrac{\sqrt{2}}{2}y_1-\dfrac{\sqrt{2}}{6}y_2-\dfrac{2}{3}y_3, \\ x_3=-\dfrac{2\sqrt{2}}{3}y_2+\dfrac{1}{3}y_3 \end{cases}$

将原二次型化为 $7y_1^2+7y_2^2-2y_3^2$.

3. 正交替换 $\begin{cases} x_1=\dfrac{\sqrt{2}}{2}y_1+\dfrac{\sqrt{2}}{2}y_3, \\ x_2=\dfrac{\sqrt{2}}{2}y_1-\dfrac{\sqrt{2}}{2}y_3, \\ x_3=\dfrac{\sqrt{2}}{2}y_2+\dfrac{\sqrt{2}}{2}y_4, \\ x_4=-\dfrac{\sqrt{2}}{2}y_2+\dfrac{\sqrt{2}}{2}y_4, \end{cases}$

将原二次型化为 $y_1^2+y_2^2-y_3^2-y_4^2$.

4. 正交替换 $\begin{cases} x_1=\dfrac{\sqrt{2}}{2}y_1+\dfrac{\sqrt{6}}{6}y_2+\dfrac{\sqrt{3}}{6}y_3+\dfrac{1}{2}y_4, \\ x_2=\dfrac{\sqrt{2}}{2}y_1-\dfrac{\sqrt{6}}{6}y_2-\dfrac{\sqrt{3}}{6}y_3-\dfrac{1}{2}y_4, \\ x_3=\dfrac{\sqrt{6}}{3}y_2-\dfrac{\sqrt{3}}{6}y_3-\dfrac{1}{2}y_4, \\ x_4=\dfrac{\sqrt{3}}{2}y_3-\dfrac{1}{2}y_4, \end{cases}$

将原二次型化为 $3y_1^2+3y_2^2+3y_3^2-5y_4^2$.

5. 正交替换 $\begin{cases} x_1=\dfrac{1}{2}y_1+\dfrac{1}{2}y_2+\dfrac{\sqrt{2}}{2}y_3, \\ x_2=\dfrac{1}{2}y_1-\dfrac{1}{2}y_2+\dfrac{\sqrt{2}}{2}y_4, \\ x_3=\dfrac{1}{2}y_1+\dfrac{1}{2}y_2-\dfrac{\sqrt{2}}{2}y_3, \\ x_4=\dfrac{1}{2}y_1-\dfrac{1}{2}y_2-\dfrac{\sqrt{2}}{2}y_4, \end{cases}$

将原二次型化为 $y_1^2-y_2^2$.

5.3 化二次型为标准形

习 题 5.3

1. 非退化线性替换 $\begin{cases} x_1=y_1-y_2+2y_3, \\ x_2=y_2-y_3, \\ x_3=y_3, \end{cases}$

将原二次型化为 $y_1^2+y_2^2-2y_3^2$.(标准形及所作非退化线性替换都不是唯一的.)

2. 非退化线性替换 $\begin{cases} x_1=y_1-y_2-y_3, \\ x_2=y_2+y_3, \\ x_3=y_3, \end{cases}$

将原二次型化为 $y_1^2-y_2^2$.

3. 非退化线性替换 $\begin{cases} x_1=y_1-y_2-y_3, \\ x_2=y_1+y_2-y_3, \\ x_3=y_3, \end{cases}$

将原二次型化为 $y_1^2-y_2^2-y_3^2$.

4. 原二次型经线性替换 $\begin{cases} x_1=y_1-2y_2+y_3-y_4, \\ x_2=y_2-\dfrac{3}{2}y_3+y_4, \\ x_3=y_3-y_4, \\ x_4=y_4, \end{cases}$

化为标准形 $y_1^2-2y_2^2+\frac{1}{2}y_3^2$.

5. 原二次型经非退化线性替换 $\begin{cases} x_1=y_1-\frac{1}{2}y_2-y_3-\frac{1}{2}y_4, \\ x_2=y_1+\frac{1}{2}y_2-y_3-\frac{1}{2}y_4, \\ x_3=y_3-\frac{1}{2}y_4, \\ x_4=y_4, \end{cases}$

化为 $y_1^2-\frac{1}{4}y_2^2-y_3^2-\frac{3}{4}y_4^2$.

习 题 5.4

1. (1) 实二次型 $f(x_1,x_2,x_3)$ 经线性替换

$$\begin{cases} x_1=y_1-y_2+\sqrt{2}y_3, \\ x_2=y_2-\frac{\sqrt{2}}{2}y_3, \\ x_3=\frac{\sqrt{2}}{2}y_3, \end{cases}$$

化为规范形 $y_1^2+y_2^2-y_3^2$;

复二次型 $f(x_1,x_2,x_3)$ 经线性替换

$$\begin{cases} x_1=z_1-z_2-\sqrt{2}\mathrm{i}z_3, \\ x_2=z_2+\frac{\sqrt{2}\mathrm{i}}{2}z_3, \\ x_3=-\frac{\sqrt{2}\mathrm{i}}{2}z_3, \end{cases}$$

化为规范形 $z_1^2+z_2^2+z_3^2$.(所作替换不是唯一的.)

(2) 实二次型 $f(x_1,x_2,x_3)$ 经线性替换

$$\begin{cases} x_1=y_1-y_2-y_3, \\ x_2=y_2+y_3, \\ x_3=y_3, \end{cases}$$

化为规范形 $y_1^2-y_2^2$;

复二次型 $f(x_1,x_2,x_3)$ 经线性替换

$$\begin{cases}x_1 = z_1 + iz_2 - z_3,\\ x_2 = -iz_2 + z_3,\\ x_3 = z_3,\end{cases}$$

化为规范形 $z_1^2+z_2^2$.

(3) 实二次型 $f(x_1,x_2,x_3)$ 经线性替换

$$\begin{cases}x_1 = y_1 - y_2 - y_3,\\ x_2 = y_1 + y_2 - y_3,\\ x_3 = y_3,\end{cases}$$

化为规范形 $y_1^2-y_2^2-y_3^2$;

复二次型 $f(x_1,x_2,x_3)$ 经线性替换

$$\begin{cases}x_1 = z_1 + iz_2 + iz_3,\\ x_2 = z_1 - iz_2 + iz_3,\\ x_3 = -iz_3,\end{cases}$$

化为规范形 $z_1^2+z_2^2+z_3^2$.

(4) $f(x_1,x_2,x_3,x_4)$ 看作实二次型,可经线性替换

$$\begin{cases}x_1 = y_1 + \sqrt{2}\,y_2 - \sqrt{2}\,y_3 - y_4,\\ x_2 = -\dfrac{3\sqrt{2}}{2}y_2 + \dfrac{\sqrt{2}}{2}y_3 + y_4\\ x_3 = \sqrt{2}\,y_2 - y_4,\\ x_4 = y_4,\end{cases}$$

化成规范形 $y_1^2+y_2^2-y_3^2$;

$f(x_1,x_2,x_3,x_4)$ 看作复二次型可经线性替换

$$\begin{cases}x_1 = z_1 + \sqrt{2}\,z_2 + \sqrt{2}\,iz_3 - z_4,\\ x_2 = -\dfrac{3\sqrt{2}}{2}z_2 - \dfrac{\sqrt{2}}{2}iz_3 + z_4,\\ x_3 = \sqrt{2}\,z_2 - z_4,\\ x_4 = z_4,\end{cases}$$

化成规范形 $z_1^2+z_2^2+z_3^2$.

(5) $f(x_1,x_2,x_3,x_4)$看作实二次型可经线性替换

$$\begin{cases} x_1 = y_1 - y_2 - y_3 - \dfrac{\sqrt{3}}{3}y_4, \\ x_2 = y_1 + y_2 - y_3 - \dfrac{\sqrt{3}}{3}y_4, \\ x_3 = y_3 - \dfrac{\sqrt{3}}{3}y_4, \\ x_4 = \dfrac{2\sqrt{3}}{3}y_4, \end{cases}$$

化为规范形 $y_1^2-y_2^2-y_3^2-y_4^2$;

$f(x_1,x_2,x_3,x_4)$看作复二次型可经线性替换

$$\begin{cases} x_1 = z_1 + \mathrm{i}z_2 + \mathrm{i}z_3 + \dfrac{\sqrt{3}\,\mathrm{i}}{3}z_4, \\ x_2 = z_1 - \mathrm{i}z_2 + \mathrm{i}z_3 + \dfrac{\sqrt{3}\,\mathrm{i}}{3}z_4, \\ x_3 = -\mathrm{i}z_3 + \dfrac{\sqrt{3}\,\mathrm{i}}{3}z_4, \\ x_4 = -\dfrac{2\sqrt{3}\,\mathrm{i}}{3}z_4, \end{cases}$$

化为规范形 $z_1^2+z_2^2+z_3^2+z_4^2$.

2. $\dfrac{(n+1)(n+2)}{2}$.

4. 提示：应用 $f(x_1,x_2,\cdots,x_n)$的规范形.

习　题　5.5

1. (1) 是； (2) 是； (3) 不是； (4) 不是.

2. (1) $-\dfrac{4}{5}<t<0$;

   (2) 不论 t 取什么值，这个二次型都不是正定的.

3. 提示：与正定矩阵合同的矩阵也是正定矩阵.

4. 提示：A 正定 $\rightleftharpoons A$ 的特征值全大于 0.

5. 提示：根据定义.

复 习 题 5

1. (1) 标准形：$y_1^2+y_2^2-5y_3^2$；

所作非退化线性替换：

$$\begin{cases} x_1 = y_1 - y_2 + 2y_3, \\ x_2 = \qquad\quad y_2 - 3y_3, \\ x_3 = \qquad\qquad\quad\ \ y_3; \end{cases}$$

实规范形：$z_1^2+z_2^2-z_3^2$；

所作非退化线性替换：

$$\begin{cases} x_1 = z_1 - z_2 + \dfrac{2\sqrt{5}}{5}z_3, \\ x_2 = \qquad\ \ z_2 - \dfrac{3\sqrt{5}}{5}z_3, \\ x_3 = \qquad\qquad\quad\ \dfrac{\sqrt{5}}{5}z_3; \end{cases}$$

复规范形：$w_1^2+w_2^2+w_3^2$；

所作非退化线性替换：

$$\begin{cases} x_1 = w_1 - w_2 - \dfrac{2\sqrt{5}\,\mathrm{i}}{5}w_3, \\ x_2 = \qquad\quad z_2 + \dfrac{3\sqrt{5}\,\mathrm{i}}{5}w_3, \\ x_3 = \qquad\qquad\quad -\dfrac{\sqrt{5}\,\mathrm{i}}{5}w_3. \end{cases}$$

(2) 标准形：$2y_1^2-2y_2^2$；

所作非退化线性替换：

$$\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2 - 2y_3, \\ x_3 = z_3; \end{cases}$$

实规范形：$z_1^2-z_2^2$

所作非退化线性替换：

$$\begin{cases} x_1 = \dfrac{\sqrt{2}}{2}z_1 + \dfrac{\sqrt{2}}{2}z_2 \\ x_2 = \dfrac{\sqrt{2}}{2}z_1 - \dfrac{\sqrt{2}}{2}z_2 - 2z_3 \\ x_3 = z_3 \end{cases}$$

复规范形：$w_1^2+w_2^2$

所作非退化线性替换：

$$\begin{cases} x_1 = \dfrac{\sqrt{2}}{2}w_1 - \dfrac{\sqrt{2}\,\mathrm{i}}{2}w_2 \\ x_2 = \dfrac{\sqrt{2}}{2}w_1 + \dfrac{\sqrt{2}\,\mathrm{i}}{2}w_2 - 2w_3 \\ x_3 = w_3 \end{cases}$$

(3) 标准形：$y_1^2 - y_2^2 + y_3^2 + \dfrac{3}{4}y_4^2$

所作非退化线性替换：

$$\begin{cases} x_1 = y_1 + y_2 + y_3 - \dfrac{1}{2}y_4 \\ x_2 = y_1 - y_2 - y_3 + \dfrac{1}{2}y_4 \\ x_3 = y_3 + \dfrac{1}{2}y_4 \\ x_4 = y_4 \end{cases}$$

实规范形：$z_1^2 + z_2^2 + z_3^2 - z_4^2$

所作非退化线性替换

$$\begin{cases} x_1 = z_1 - \dfrac{\sqrt{3}}{3}z_2 + z_3 + z_4 \\ x_2 = z_1 + \dfrac{\sqrt{3}}{3}z_2 - z_3 - z_4 \\ x_3 = \dfrac{\sqrt{3}}{3}z_2 + z_3 \\ x_4 = \dfrac{2\sqrt{3}}{3}z_2 \end{cases}$$

复规范形：$w_1^2+w_2^2+w_3^2+w_4^2$；

所作的非退化线性替换：

$$\begin{cases} x_1 = w_1 - \dfrac{\sqrt{3}}{3}w_2 + w_3 - \mathrm{i}w_4, \\ x_2 = w_1 + \dfrac{\sqrt{3}}{3}w_2 - w_3 + \mathrm{i}w_4, \\ x_3 = \dfrac{\sqrt{3}}{3}w_2 + w_3, \\ x_4 = \dfrac{2\sqrt{3}}{3}w_2; \end{cases}$$

（4）标准形：$y_1^2-y_2^2+9y_3^2$；

所作非退化线性替换：

$$\begin{cases} x_1 = y_1^2 - \dfrac{3}{2}y_2 - \dfrac{5}{2}y_3 + \dfrac{3}{2}y_4, \\ x_2 = \dfrac{1}{2}y_2 + \dfrac{3}{2}y_3 - \dfrac{1}{2}y_4, \\ x_3 = y_3, \\ x_4 = y_4; \end{cases}$$

实规范形：$z_1^2+z_2^2-z_3^2$；

所作非退化线性替换：

$$\begin{cases} x_1 = z_1 - \dfrac{5}{6}z_2 - \dfrac{3}{2}z_3 + \dfrac{3}{2}z_4, \\ x_2 = \dfrac{1}{2}z_2 + \dfrac{1}{2}z_3 - \dfrac{1}{2}z_4, \\ x_3 = \dfrac{1}{3}z_3, \\ x_4 = z_4; \end{cases}$$

复规范形：$w_1^2+w_2^2+w_3^2$；

所作非退化线性替换：

$$\begin{cases} x_1 = w_1 - \dfrac{5}{6}w_2 + \dfrac{3}{2}\mathrm{i}w_3 - \dfrac{3}{2}w_4, \\ x_2 = \dfrac{1}{2}w_2 - \dfrac{1}{2}\mathrm{i}w_3 - \dfrac{1}{2}w_4, \\ x_3 = \dfrac{1}{3}w_2, \\ x_4 = w_4. \end{cases}$$

2. (1) $A=\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 5 \end{bmatrix}$;

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & -5 \end{bmatrix};$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \dfrac{2\sqrt{5}}{5} & -\dfrac{3\sqrt{5}}{5} & \dfrac{\sqrt{5}}{5} \end{bmatrix} A \begin{bmatrix} 1 & -1 & \dfrac{2\sqrt{5}}{5} \\ 0 & 1 & -\dfrac{3\sqrt{5}}{5} \\ 0 & 0 & \dfrac{\sqrt{5}}{5} \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \end{bmatrix};$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -\dfrac{2\sqrt{5}}{5}\mathrm{i} & \dfrac{3\sqrt{5}}{5}\mathrm{i} & -\dfrac{\sqrt{5}}{5}\mathrm{i} \end{bmatrix} A \begin{bmatrix} 1 & -1 & -\dfrac{2\sqrt{5}}{5}\mathrm{i} \\ 0 & 1 & \dfrac{3\sqrt{5}}{5}\mathrm{i} \\ 0 & 0 & \dfrac{\sqrt{5}}{5}\mathrm{i} \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}.$$

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

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

$$
\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & -2 & 1 \end{bmatrix} A \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & -1 & \\ & & 0 \end{bmatrix};
$$

$$
\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2}\mathrm{i} & \frac{\sqrt{2}}{2}\mathrm{i} & 0 \\ 0 & -2 & 1 \end{bmatrix} A \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\mathrm{i} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\mathrm{i} & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix}.
$$

(3) $A = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$;

$$
\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} A \begin{bmatrix} 1 & 1 & 1 & -\frac{1}{2} \\ 1 & -1 & -1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \frac{3}{4} \end{bmatrix};
$$

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

$$\begin{bmatrix}1&&&\\&1&&\\&&1&\\&&&-1\end{bmatrix};$$

$$\begin{bmatrix}1&1&0&0\\-\frac{\sqrt{3}}{3}&\frac{\sqrt{3}}{3}&\frac{\sqrt{3}}{3}&\frac{2\sqrt{3}}{3}\\1&-1&1&0\\-\mathrm{i}&\mathrm{i}&0&0\end{bmatrix}A\begin{bmatrix}1&-\frac{\sqrt{3}}{3}&1&-\mathrm{i}\\1&\frac{\sqrt{3}}{3}&-1&\mathrm{i}\\0&\frac{\sqrt{3}}{3}&1&0\\0&\frac{2\sqrt{3}}{3}&0&0\end{bmatrix}=$$

$$\begin{bmatrix}1&&&\\&1&&\\&&1&\\&&&1\end{bmatrix}.$$

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

$$\begin{bmatrix}1&0&0&0\\-\frac{3}{2}&\frac{1}{2}&0&0\\-\frac{5}{2}&\frac{3}{2}&1&0\\\frac{3}{2}&-\frac{1}{2}&0&1\end{bmatrix}A\begin{bmatrix}1&-\frac{3}{2}&-\frac{5}{2}&\frac{3}{2}\\0&\frac{1}{2}&\frac{3}{2}&-\frac{1}{2}\\0&0&1&0\\0&0&0&1\end{bmatrix}=\begin{bmatrix}1&&&\\&-1&&\\&&9&\\&&&0\end{bmatrix};$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{5}{6} & \frac{1}{2} & 0 & 0 \\ -\frac{3}{2} & \frac{1}{2} & \frac{1}{3} & 0 \\ \frac{3}{2} & -\frac{1}{2} & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -\frac{5}{6} & -\frac{3}{2} & \frac{3}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & 0 \end{bmatrix};$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{5}{6} & \frac{1}{2} & \frac{1}{3} & 0 \\ -\frac{3i}{2} & -\frac{i}{2} & 0 & 0 \\ -\frac{3}{2} & -\frac{1}{2} & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -\frac{5}{6} & -\frac{3i}{2} & -\frac{3}{2} \\ 0 & \frac{1}{2} & -\frac{i}{2} & -\frac{1}{2} \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{bmatrix}$$

3. (1) 标准形：$2y_1^2+2y_2^2-7y_3^2$；

所作正交线性替换：

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{2}{5}\sqrt{5} & \frac{2}{15}\sqrt{5} & \frac{1}{3} \\ \frac{1}{5}\sqrt{5} & \frac{4}{15}\sqrt{5} & \frac{2}{3} \\ 0 & \frac{1}{3}\sqrt{5} & -\frac{2}{3} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix};$$

正惯性指数$=2$,负惯性指数$=1$,符号差$=1$.

(2) 标准形：$y_1^2+y_2^2-y_3^2-y_4^2$；

所作正交替换：

$$\begin{cases} x_1 = \frac{\sqrt{2}}{2}y_1 - \frac{\sqrt{2}}{2}y_3, \\ x_2 = \frac{\sqrt{2}}{2}y_1 + \frac{\sqrt{2}}{2}y_3, \\ x_3 = \frac{\sqrt{2}}{2}y_2 - \frac{\sqrt{2}}{2}y_4; \\ x_4 = \frac{\sqrt{2}}{2}y_2 + \frac{\sqrt{2}}{2}y_4; \end{cases}$$

正惯性指数=负惯性指数=2,符号差=0.

(3) 标准形:$y_1^2+7y_2^2-y_3^2-3y_4^2$;

所作正交替换:

$$
\begin{cases}
x_1=\frac{1}{2}y_1-\frac{1}{2}y_2-\frac{1}{2}y_3+\frac{1}{2}y_4,\\
x_2=\frac{1}{2}y_1+\frac{1}{2}y_2-\frac{1}{2}y_3-\frac{1}{2}y_4,\\
x_3=\frac{1}{2}y_1-\frac{1}{2}y_2+\frac{1}{2}y_3-\frac{1}{2}y_4,\\
x_4=\frac{1}{2}y_1+\frac{1}{2}y_2+\frac{1}{2}y_3+\frac{1}{2}y_4;
\end{cases}
$$

正惯性指数=负惯性指数=2,符号差=0.

4. (1) 是; (2) 否.

5. (1) $-\sqrt{2}<t<\sqrt{2}$; (2) $-\frac{\sqrt{15}}{3}<t<\frac{\sqrt{15}}{3}$.

6. 提示:参照定理 5 的证明.

7. 提示:A 是正定矩阵的充分必要条件是 A 与单位矩阵合同.

8. 提示:A 是半正定矩阵的充分必要条件是 A 合同于矩阵[1,1,…,1,0,…,0].

9. 提示:证明可找到 t 使 $tE+A$ 的特征值全大于 0.

12. 证明:因为 A 是正定矩阵,所以存在可逆矩阵 T_1,使 $T_1^{\mathrm{T}}AT_1=E$. 设 $T_1^{\mathrm{T}}BT_1=B_1$,找正交矩阵 T_2 使 $T_2^{\mathrm{T}}B_1T_2=B_2$ 为对角形. 令 $T=T_1T_2$,则 T 是可逆的,并且 $T^{\mathrm{T}}AT=T_2^{\mathrm{T}}T_1^{\mathrm{T}}AT_1T_2=T_2^{\mathrm{T}}ET_2=E$,$T^{\mathrm{T}}BT=T_2^{\mathrm{T}}T_1^{\mathrm{T}}BT_1T_2=T_2^{\mathrm{T}}B_1T_2=B_2$ 都是对角矩阵.

第6章 多 项 式

多项式是一类最常见、最简单的函数，它的应用非常广泛. 关于多项式的一些重要定理和方法，不仅在解决实际问题时常常用到，而且在进一步学习代数以及其他学科时也经常会碰到. 在中学代数里，我们曾经学习过多项式的运算. 这一章将在复习这些内容的基础上，进一步讨论关于一元多项式的一些理论，例如最大公因式的性质、求法，以及多项式的因式分解等问题.

6.1 多项式及其运算

这一节讨论一元多项式的基本概念以及一元多项式的运算.

定义 1 设 x 是一个变量（或称文字），n 是一个非负整数，表示式

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \tag{1}$$

其中 $a_n, a_{n-1}, \cdots, a_1, a_0$ 全属于数域 P，称为**系数在数域 P 中的一元多项式，简称数域 P 上的一元多项式**. 例如：

$$\frac{1}{2}x^3 + 5x - 5, \quad \sqrt{2}x^4 + 2ix^2 + 1, 4$$

都是 x 的多项式，而

$$2x^{1/2} - 1, \quad x^2 - 3x + x^{-1} + 2$$

都不是 x 的多项式.

这里定义的多项式只含有一个变量，所以也称一元多项式. 我们讨论的都是一元多项式，以后不再声明，就简单地称作多项式. 我们以后常常用 $f(x), g(x)$ 等表示多项式. 有时候为了方便起

见，在不致引起误解的情况下，简单地用 f,g 等表示多项式.

设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

是数域 **P** 上一个多项式. $a_n,a_{n-1},\cdots,a_1,a_0$ 称为 $f(x)$ 的**系数**；a_kx^k $(k=n,n-1,\cdots,1,0)$ 称为 $f(x)$ 的 **k 次项**；a_k 称为 $f(x)$ 的 **k 次项系数**. $f(x)$ 的零次项也称作 $f(x)$ 的**常数项**.

如果 c 是数域 **P** 中一个数，用 $f(c)$ 表示当 $x=c$ 时 $f(x)$ 的值，即

$$f(c)=a_nc^n+a_{n-1}c^{n-1}+\cdots+a_1c+a_0$$

容易看出，$f(c)$ 也是数域 **P** 中的数.

定义 2 如果两个多项式 $f(x)$ 和 $g(x)$ 的同次项的系数全相等，就称 $f(x)$ 与 $g(x)$ 相等，记作 $f(x)=g(x)$.

系数全等于零的多项式称为**零多项式**，记作 0.

定义 3 如果在(1)中，$a_n\neq0$，那么 a_nx^n 就称为多项式(1)的**首项**，a_n 称为多项式(1)的**首项系数**，n 称为多项式(1)的**次数**.

为了方便起见，我们用"$\deg f(x)$"(或者简单地，用"$\deg f$")来表示 $f(x)$ 的次数. 非零常数是零次多项式. 零多项式是唯一的无法确定次数的多项式，因此必须注意：只有在 $f(x)\neq0$ 时，符号"$\deg f(x)$"才有意义.

两个多项式相加(或相减)，就是把它们的同次项的系数相加(或相减)；两个多项式相乘，就是把第 1 个多项式的各个项与第 2 个多项式的各个项分别相乘，然后合并同类项. 多项式 $f(x)$ 与 $g(x)$ 相加、相减与相乘的结果分别称为这两个多项式的**和**、**差**与**积**，记作 $f(x)+g(x)$，$f(x)-g(x)$ 与 $f(x)\cdot g(x)$(或 $f(x)g(x)$). 下面举例说明多项式的运算.

例 1 设 $f(x)=2x^2+3x-1$，$g(x)=x^3+2x^2-3x+2$.

求 $f(x)+g(x)$， $f(x)-g(x)$，$f(x)g(x)$.

解：$f(x)+g(x)=(2x^2+3x-1)+(x^3+2x^2-3x+2)$

$$=x^3+(2+2)x^2+(3-3)x+(-1+2)$$

$$=x^3+4x^2+1.$$

$$f(x)-g(x)=(2x^2+3x-1)-(x^3+2x^2-3x+2)$$

$$=(0-1)x^3+(2-2)x^2+[3-(-3)]x$$

$$+(-1-2)=-x^3+6x-3.$$

$$f(x)g(x)=(2x^2+3x-1)\cdot(x^3+2x^2-3x+2)$$

$$=2x^5+3x^4-x^3+4x^4+6x^3-2x^2$$

$$-6x^3-9x^2+3x+4x^2+6x-2$$

$$=2x^5+(3+4)x^4+(-1+6-6)x^3$$

$$+(-2-9+4)x^2+(3+6)x-2$$

$$=2x^5+7x^4-x^3-7x^2+9x-2.$$

多项式的乘法比较复杂，应用竖式就方便多了. 例如上面的例子可以这样来计算.

$$\begin{array}{rrrrrrl}
 & & & f(x)= & 2x^2 & +3x & -1 \\
\times) & & & g(x)= x^3 & +2x^2 & -3x & +2 \\
\hline
2x^5 & +3x^4 & -x^3 & & & & \\
 & 4x^4 & +6x^3 & -2x^2 & & & \\
 & & -6x^3 & -9x^2 & +3x & & \\
 & & & 4x^2 & +6x & -2 & \\
\hline
\end{array}$$

$$f(x)\cdot g(x)=2x^5+7x^4-x^3-7x^2+9x-2.$$

这样计算不仅简单明了，而且不容易出错. 有时候，为了更加方便，我们还可以应用分离系数法来作乘法，就是把 x 的方幂略去不写，只把系数按次序写出来进行运算. 仍用上面的例子来说明，上面的竖式可以应用分离系数法简化为：

$$
\begin{array}{rrrrrrr}
 & 2 & 3 & -1 & & & \\
\times) & 1 & 2 & -3 & 2 & & \\
\hline
 & 2 & 3 & -1 & & & \\
 & & 4 & 6 & -2 & & \\
 & & & -6 & -9 & 3 & \\
 & & & & 4 & 6 & -2 \\
\hline
 & 2 & 7 & -1 & -7 & 9 & -2
\end{array}
$$

最后所得的一列数就是 $f(x)\cdot g(x)$ 按降幂排列的系数.

应用这种方法进行运算的时候，一定要注意：必须把系数是"0"的项补上. 我们再举一个例子.

例 2 $f(x)=x^3-1, g(x)=5x^2+x-1$.

求 $f(x)\cdot g(x)$.

解：应用分离系数法进行计算

$$
\begin{array}{rrrrrrr}
 & 1 & 0 & 0 & -1 & & \\
\times) & 5 & 1 & -1 & & & \\
\hline
 & 5 & 0 & 0 & -5 & & \\
 & & 1 & 0 & 0 & -1 & \\
 & & & -1 & 0 & 0 & 1 \\
\hline
 & 5 & 1 & -1 & -5 & -1 & 1
\end{array}
$$

所以 $f(x)g(x)=5x^5+x^4-x^3-5x^2-x+1$.

在有些时候，$f(x)\cdot g(x)$ 可能是按升幂排列的，那么我们当然也可以按升幂排法写出各项系数来计算，得出 $f(x)g(x)$ 按升幂排法的系数. 还有，从上面的例子可以看出，如果 $g(x)$ 中有一些系数为 0 的项，在计算时可以不必写出一整行 0，只要在那一行写上第一个 0，然后接着算下一行就可以了.

例 3 假设 $f(x)=2-x+3x^3, g(x)=3+x^2-2x^3$. 求 $f(x)g(x)$.

解：

$$
\begin{array}{rrrrrrrr}
 & 2 & -1 & 0 & 3 & & & \\
\times) & 3 & 0 & 1 & -2 & & & \\
\hline
 & 6 & -3 & 0 & 9 & & & \\
 & & 0 & 2 & -1 & 0 & 3 & \\
 & & & & -4 & 2 & 0 & -6 \\
\hline
 & 6 & -3 & 2 & 4 & 2 & 3 & -6
\end{array}
$$

所以　$f(x)g(x)=6-3x+2x^2+4x^3+2x^4+3x^5-6x^6$.

上面算式中第四行的第一个 0 就是由 $g(x)$ 中的那个 0 系数得出的. 我们把这一行其余的 0 都略去不写而接着计算下一行. 因此必须注意，在计算再下面一行的时候，一定要把第一个数字写在这一行第 3 个数字下面. 对此读者可以用数的乘法中乘数含有 0 的情形进行比较.

根据以上的方法，任意给了两个具体的多项式，都可以求出它们的和、差与积. 但是为了统一地讨论多项式的问题，还需要将多项式的运算用一般公式表示出来. 下面来给出和、差与积的公式.

设　$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$,

$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0$.

首先写出 $f(x)+g(x)$ 与 $f(x)-g(x)$ 的表达式. 因为在一个多项式中可以任意加入一些系数为 0 的项，如果 $m\leqslant n$，当 $m<n$ 时可设 $b_{m+1}=b_{m+2}=\cdots=b_n=0$. 于是

$$
\begin{aligned}
f(x)+g(x) &= (a_n+b_n)x^n+(a_{n-1}+b_{n-1})x^{n-1} \\
&\quad +\cdots+(a_1+b_1)x+(a_0+b_0) \\
f(x)-g(x) &= (a_n-b_n)x^n+(a_{n-1}+b_{n-1})x^{n-1} \\
&\quad +\cdots+(a_1-b_1)x+a_0
\end{aligned}
$$

如果 $m>n$，可以类似地处理. 读者可以作为练习推导一下.

根据上面的 $f(x)+g(x)$ 与 $f(x)-g(x)$ 的表示式，可以看

出：$f(x)+g(x)$的 k 次项系数是 a_k+b_k；而 $f(x)-g(x)$的 k 次项系数是 a_k-b_k.

下面来推导 $f(x)g(x)$的公式.

$$\begin{aligned}f(x)g(x) &= (a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0)\\ &\quad\cdot(b_mx^m + b_{m-1}x^{m-1} + \cdots + b_1x + b_0)\\ &= (a_nb_mx^{n+m} + a_{n-1}b_mx^{n+m-1} + \cdots +\\ &\quad + a_1b_mx^{m+1} + a_0b_mx^m) + (a_nb_{m-1}x^{n+m-1}\\ &\quad + a_{n-1}b_{m-1}x^{n+m-2} + \cdots + a_1b_{m-1}x^m\\ &\quad + a_0b_{m-1}x^{m-1}) + \cdots + (a_nb_0x^n + a_{n-1}b_0x^{n-1}\\ &\quad + \cdots + a_1b_0x + a_0b_0).\end{aligned}$$

在合并同次项以前，我们先来分析 $f(x)g(x)$中各个项的系数与 x 的幂次的关系. 由于 $f(x)g(x)$中各项是由 $f(x)$的一个项与 $g(x)$的一个项相乘而得. 所以可以表示成：

$$(a_ix^i)(b_jx^j) = a_ib_jx^{i+j}.$$

因此 $f(x)g(x)$中 $x^k(0\leqslant k\leqslant n+m)$的系数是

$$a_0b_k + a_1b_{k-1} + \cdots + a_{k-1}b_1 + a_kb_0.$$

其中，如果 $i>n$，则 $a_i=0$；如果 $j>m$，则 $b_j=0$.（以后遇到类似情况，都这么理解而不再重复说明.）因此

$$\begin{aligned}f(x)g(x) &= a_nb_mx^{n+m} + (a_{n-1}b_m + a_nb_{m-1})x^{n+m-1}\\ &\quad + \cdots + (a_0b_1 + a_1b_0)x + a_0b_0.\end{aligned}$$

为了简便起见，我们常常把多项式的运算公式用和号来表示. 如果

$$f(x) = \sum_{i=0}^{n}a_ix^i,\quad g(x) = \sum_{i=0}^{m}b_jx^j,$$

则
$$f(x) \pm g(x) = \sum_{i=0}^{n}(a_i \pm b_i)x^i \quad (n \geqslant m);$$

$$f(x)g(x) = \sum_{k=0}^{n+m}\Big(\sum_{i+j=k}a_ib_j\Big)x^k.$$

从多项式的运算公式，很容易得出下述次数公式：

命题 1 1) $\deg(f(x)\pm g(x))\leqslant\max(\deg f(x),\deg g(x))$；

2) $\deg(f(x)g(x))=\deg f(x)+\deg g(x)$.

多项式的乘积的次数不仅满足上述公式，从乘积公式还可看出：

命题 2 如果 $f(x)\neq 0$，$g(x)\neq 0$，那么 $f(x)g(x)\neq 0$. 而且 $f(x)g(x)$ 的首项就等于 $f(x)$ 的首项与 $g(x)$ 的首项之积；$f(x)g(x)$ 的首项系数等于 $f(x)$ 的首项系数与 $g(x)$ 的首项系数之积.

以上我们复习了多项式的运算，在以后的讨论以及一些实际问题中，我们不只简单地用到两个多项式的加法、减法或乘法，有时会遇到多个多项式的复杂运算. 因此为了简化运算，必须掌握一些运算的规律. 其实我们在算术与初等代数中，经常应用一些规律来比较巧妙地进行计算. 例如

$$\begin{aligned}&(a+1)(a+b-1)+(a+1)(a-b-1)\\&=(a+1)[(a+b-1)+(a-b-1)]\\&=(a+1)(2a-2)=2(a+1)(a-1)\\&=2(a^2-1).\end{aligned}$$

在上面的计算过程中，第 1 个等号我们应用了加法对乘法的分配律；第 2 个等号用到加法的交换律与结合律；第 3 个等号应用了乘法的交换律与结合律；最后一个等号用到了平方差公式（其实这个公式的证明也需应用好几个规律）. 与数的运算类似，多项式的运算也满足下列一些规律.

命题 3 多项式的运算满足：

加法交换律：

$$f(x)+g(x)=g(x)+f(x);$$

加法结合律：

$$[f(x)+g(x)]+h(x)=f(x)+[g(x)+h(x)];$$

乘法交换律：

$$f(x)g(x)=g(x)f(x);$$

乘法结合律：

$$[f(x)g(x)]h(x)=f(x)[g(x)h(x)];$$

加乘分配律：

$$f(x)[g(x)+h(x)]=f(x)g(x)+f(x)h(x).$$

这些规律是需要逐条验证的.但是由于我们在初等代数中已经熟悉并经常运用这些规律，而且这些规律的证明又比较呆板和烦琐，因此我们不一一验证，而只是对其中的几条加以证明.目的是想介绍一下如何验证这类规律，并通过这些规律的检验复习多项式的运算公式.

1）加法交换律的证明　设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$
$$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0,$$

不妨设 $m\leqslant n$，于是

$$\begin{aligned}f(x)+g(x)&=(a_n+b_n)x^n+(a_{n-1}+b_{n-1})x^{n-1}\\&\quad+\cdots+(a_1+b_1)x+a_0+b_0\\&=(b_n+a_n)x^n+(b_{n-1}+a_{n-1})x^{n-1}\\&\quad+\cdots+(b_1+a_1)x+b_0+a_0\\&=g(x)+f(x).\end{aligned}$$

因为多项式的加法比较简单，所以可以直接验证其交换律.加法结合律及乘法交换律都可以类似地证明.读者不妨自己证明一下.至于其余几条规律，其中的系数比较复杂，要直接从等式左边推导到右边，写起来比较麻烦，再用上面的方法来证明就不大方便了，必须采用别的方法.由于两个多项式相等的意思是指它们的同次项系数相等，因此为了证明某个等式，只要证明等式两边同次项的系数相等就行了.下面我们以乘法结合律为例来说明这一方法.

2）乘法结合律的证明　设

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$
$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0,$$
$$h(x) = c_l x^l + c_{l-1}x^{l-1} + \cdots + c_1 x + c_0.$$

为了方便起见，我们再设

$$f(x)g(x) = d_{n+m}x^{n+m} + \cdots + d_1 x + d_0;$$
$$g(x)h(x) = e_{m+l}x^{m+l} + \cdots + e_1 x + e_0,$$

其中

$$d_i = \sum_{s+t=i} a_s b_t, \quad i = 0,1,\cdots,n+m,$$
$$e_j = \sum_{t+u=j} b_t c_u, \quad j = 0,1,\cdots,m+l.$$

于是，$[f(x)g(x)]h(x)$的 $k(1\leqslant k\leqslant n+m+l)$次项系数为

$$\begin{aligned}\sum_{i+u=k} d_i c_u &= \sum_{i+u=k}\Big(\sum_{s+t=i} a_s b_t\Big) c_u \\ &= \sum_{s+t+u=k}(a_s b_t)c_u = \sum_{s+t+u=k} a_s b_t c_u.\end{aligned}$$

$f(x)[g(x)h(x)]$的 $k(1\leqslant k\leqslant n+m+l)$次项系数为

$$\begin{aligned}\sum_{s+j=k} a_s e_j &= \sum_{s+j=k} a_s\Big(\sum_{t+u=j} b_t c_u\Big) \\ &= \sum_{s+t+u=k} a_s(b_t c_u) = \sum_{s+t+u=k} a_s b_t c_u.\end{aligned}$$

所以$[f(x)g(x)]h(x)$与 $f(x)[g(x)h(x)]$的同次项系数都相等.根据多项式相等的定义，即得

$$[f(x)g(x)]h(x) = f(x)[g(x)h(x)].$$

乘法结合律得证.

多项式的乘法还满足消去律：

命题 4 如果 $f(x)g(x)=f(x)h(x)$，并且 $f(x)\neq 0$. 那么可以消去 $f(x)$而得 $g(x)=h(x)$.

证明：由 $f(x)g(x)=f(x)h(x)$可得

$$f(x)[g(x) - h(x)] = 0.$$

如果 $g(x)-h(x)\neq0$. 那么由于 $f(x)\neq0$,故由命题 2 有

$$f(x)[g(x)-h(x)]\neq0,$$

与假设矛盾,所以 $g(x)-h(x)=0$,即

$$g(x)=h(x). \quad |$$

以上介绍了有关多项式的基本概念及运算. 在多项式的定义中,$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 的系数 $a_n,a_{n-1},\cdots,a_1,a_0$ 是一些常数. 如果这些系数都是复数,就称 $f(x)$是一个复系数多项式;如果这些系数都是实数,就称 $f(x)$是一个实系数多项式;如果这些系数都是有理数,就称 $f(x)$是一个有理系数多项式. 在我们讨论多项式的某些理论和求解高次方程时,有时需要在复数范围内进行,有时需要在实数范围内进行,有时还需要在有理数或其他指定的范围内进行. 但是有些问题,不论在复数或实数或有理数范围,都可得到同样的结论,因此为了统一地进行研究,我们可以取定一个数域 **P** 来讨论系数在 **P** 中的多项式.

设 **P** 是一个数域,用 **P**$[x]$表示系数在 **P** 中的所有一元多项式组成的集合。**P**$[x]$称为**数域 P 上的一元多项式环**,**P** 称为 **P**$[x]$的**系数域**.

例如:**Q**$[x]$表示全体有理系数多项式所成的集合;**R**$[x]$表示全体实系数多项式所成的集合;**C**$[x]$表示全体复系数多项式所成的集合.

系数在数域 **P** 中的多项式也称为数域 **P** 上的多项式. 在这定义中应用了一个"上"字是没有什么意义的. 由于这种说法在代数学中已经是根深蒂固了,我们也就沿用了这个习惯的名称.

因为多项式的加法、减法与乘法可以通过系数的加法、减法与乘法来进行,所以如果 $f(x)$,$g(x)$是 **P**$[x]$中两个多项式,那么它们的和、差与积也都在 **P**$[x]$中,因此我们可以像以前讨论线性方程组以及矩阵的时候一样,取定一个数域 **P** 来讨论 **P**$[x]$中的多项式.

习　题　6.1

1. 计算 $f(x)+g(x)$，$f(x)-g(x)$以及 $f(x)g(x)$.

 (1) $f(x)=x^4+2x^3-x^2-4x-2$，

 $g(x)=2x^3-x^2-5x+4$；

 (2) $f(x)=x^4-x^3-4x^2+2x+1$，

 $g(x)=x^2-x+1$；

 (3) $f(x)=x^4-2x^3-6x^2+5x+9$，

 $g(x)=-x^4+2x^3-x^2-2x+2$；

 (4) $f(x)=2x^4-x^3+2x^2-x+1$

 $g(x)=x^2-3x+1$；

 (5) $f(x)=x^3+(1+\mathrm{i})x^2-(1-\mathrm{i})x-1$，

 $g(x)=x^3-2\mathrm{i}x+1-2\mathrm{i}$.

2. 计算

 (1) $(x^3+x^2-x-1)(x^3-2x-1)-8x(x^2-5)$；

 (2) $(x^3+ax-b)(x^2-1)+(x^3-ax+b)(x^2-1)$.

3. 求 k,l,m，使$(2x^2+lx-1)(x^2-kx+1)=2x^4+5x^3+mx^2-x-1$.

4. 设 $f(x)$，$g(x)$是两个非零多项式且 $f(x)+g(x)\neq 0$. 问 $f(x)$，$g(x)$的系数满足什么条件时，次数公式

$$\deg(f(x)+g(x))\leqslant\max(\deg f(x),\deg g(x))$$

 中等号成立？满足什么条件时，小于号成立？

5. 假设 $\deg f_1(x)\leqslant\deg g_1(x)$，$\deg f_2(x)\leqslant\deg g_2(x)$，

 (1) 求证：$\deg f_1(x)f_2(x)\leqslant\deg g_1(x)g_2(x)$.

 (2) 问：是否一定有 $\deg(f_1(x)+f_2(x))\leqslant\deg(g_1(x)+g_2(x))$？

 举例说明您的结论.

6.2　整除性理论

1. 带余除法

上一节我们介绍了多项式的加法、减法和乘法. 现在来讨论多

项式的除法. 我们知道，任意给了 $\mathbf{P}[x]$中两个多项式 $f(x)$，$g(x)$. 总可进行加法、减法与乘法而得到它们的和、差与积. 而且 $f(x)+g(x)$，$f(x)-g(x)$及 $f(x)g(x)$仍是 $\mathbf{P}[x]$中的多项式. 但是用 $g(x)$($g(x)\neq 0$)去除 $f(x)$并不是总是能除得尽的. 不过我们在初等代数中学过长除法，会用 $g(x)$去除 $f(x)$而求出商式与余式. 我们通过例子来复习一下长除法.

例 1

设
$$f(x) = 2x^4 + 4x^2 - 5x + 6,$$
$$g(x) = x^2 - 3x + 1.$$

用 $g(x)$除 $f(x)$：

$$
\begin{array}{r|l}
 & 2x^2+6x+20 \\
\hline
x^2-3x+1 & 2x^4 \qquad +4x^2-5x+6 \\
 & \underline{2x^4-6x^3+2x^2} \\
 & 6x^3+2x^2-5x+6 \\
 & \underline{6x^3-18x^2+6x} \\
 & 20x^2-11x+6 \\
 & \underline{20x^2-60x+20} \\
 & 49x-14
\end{array}
$$

从上式看出：$g(x)$不能除尽 $f(x)$. 用 $g(x)$去除 $f(x)$得到一个商式 $2x^2+6x+20$ 及一个余式 $49x-14$. 可以写成

$$2x^4 + 4x^2 - 5x + 6 = (2x^2 + 6x + 20)(x^2 - 3x + 1) + (49x - 14).$$

即
$$f(x) = (2x^2 + 6x + 20)g(x) + (49x - 14).$$

这个例子说明了关于多项式的一个基本性质：

定理 1(带余除法)　对于任意两个多项式 $f(x)$，$g(x)$，其中 $g(x)\neq 0$. 总可以找到多项多 $q(x)$及 $r(x)$，使得

$$f(x) = q(x)g(x) + r(x),$$

其中 $r(x)=0$，或 $\deg r(x)<\deg g(x)$；而且满足这两个条件的多

项式 $q(x)$ 及 $r(x)$ 是唯一的.

证明：$q(x)$ 与 $r(x)$ 的存在性可以由上面例子中的算法直接得到.

下面来证明唯一性：假设有两对多项式 $q(x),r(x)$ 及 $q_1(x)$，$r_1(x)$，都满足定理中的两个条件，即

$$\begin{cases} f(x)=q(x)g(x)+r(x) \\ r(x)=0 \text{ 或 } \deg r(x)<\deg g(x); \end{cases}$$

$$\begin{cases} f(x)=q_1(x)g(x)+r_1(x) \\ r_1(x)=0 \text{ 或 } \deg r_1(x)<\deg g(x). \end{cases}$$

我们要证明 $q(x)=q_1(x)$ 及 $r(x)=r_1(x)$. 从 $f(x)$ 的两个等式得到

$$q(x)g(x)+r(x)=q_1(x)g(x)+r_1(x).$$

因此 $$[q(x)-q_1(x)]g(x)=r_1(x)-r(x).$$

如果 $q_1(x)\neq q(x)$，那么 $q_1(x)-q(x)\neq 0$. 又因 $g(x)\neq 0$，所以它们的乘积也不等于 0，而且等式两边的次数应该相等，但是

$$\begin{aligned} &\deg[q(x)-q_1(x)]g(x) \\ =&\deg[q(x)-q_1(x)]+\deg g(x) \\ \geqslant&\deg g(x). \end{aligned}$$

而

$$\begin{aligned} &\deg[r_1(x)-r(x)] \\ \leqslant&\max[\deg r_1(x),\deg r(x)] \\ <&\deg g(x). \end{aligned}$$

这是不可能的，所以 $q_1(x)=q(x)$，因而 $r_1(x)=r(x)$. |

定理 1 中的 $q(x)$ 称为 $g(x)$ 除 $f(x)$ 所得的**商式**；$r(x)$ 称为**余式**.

在做带余除法时，要注意：余式 $r(x)$（如果 $r(x)\neq 0$）的次数必须小于 $g(x)$ 的次数. 否则，除法还没有做完，还须继续进行.

习　题　6.2(1)

1. 计算 $f(x)$被 $g(x)$除所得的商式及余式：

(1) $f(x)=2x^4-3x^3+4x^2-5x+6$,

$g(x)=x^2-3x+1$;

(2) $f(x)=x^3-3x^2-x-1$,

$g(x)=3x^2-2x+1$;

(3) $f(x)=x^5-x^4+4x^3+2x^2+10x-3$,

$g(x)=x^3-2x^2+3x-1$;

(4) $f(x)=x^3-(1-\mathrm{i})x^2+(5+3\mathrm{i})x-\mathrm{i}$,

$g(x)=x^2+\mathrm{i}x-1$;

(5) $f(x)=x^4-x^3-8x^2+11x-3$,

$g(x)=x^2-3x+1$.

2. 设 $f(x)=3x^4-10x^2-5x-4$, $g(x)=x-2$.

(1) 求 $g(x)$除 $f(x)$所得的商式及余式；

(2) 求 $f(2)$.

3. 设 $f(x)=a_3x^3+a_2x^2+a_1x+a_0$.

(1) 求 $f(x)$被 $x-c$ 除所得的商式及式余；

(2) 求 $f(x)$被 $2(x-c)$除所得的商式及余式.

比较这两个小题的结果，您能得出什么结论?

2. 整除的概念

上面介绍了带余除法. 用非零多项式 $g(x)$去除多项式 $f(x)$，可以得到一个商式及一个余式，余式一般不等于零. 当余式等于 0 的时候，我们就得到**整除**的概念：

定义 4　$f(x)$与 $g(x)$是两个多项式. 如果有一个多项式 $q(x)$ 使得

$$f(x)=q(x)g(x),$$

就称 $g(x)$**整除**(或**除尽**)$f(x)$或 $f(x)$被 $g(x)$所**整除**(或**除尽**). 这时，$f(x)$称为 $g(x)$的一个**倍式**；$g(x)$称为 $f(x)$的一个**因式**.

我们用“$g(x)\mid f(x)$”来表示 $g(x)$整除 $f(x)$，用“$g(x)\nmid f(x)$”

表示 $g(x)$ 不能整除 $f(x)$.

当 $g(x)\mid f(x)$ 并且 $g(x)\neq 0$ 时，可以用 $\dfrac{f(x)}{g(x)}$ 表示 $g(x)$ 除 $f(x)$ 所得的商.

例 2 因为 $x^3-1=(x-1)(x^2+x+1)$，所以 $x-1$ 能整除 x^3-1. $x-1$ 是 x^3-1 的因式，x^3-1 是 $x-1$ 的倍式. 一般地，由于

$$x^n-1=(x-1)(x^{n-1}+\cdots+x+1).$$

所以 $x-1$ 是 x^n-1 的因式.

与带余除法的情形不同，在定义 4 中，我们并没有假设 $g(x)\neq 0$. 从定义看出，如果 $g(x)=0$，那么 $g(x)$ 的倍式一定也等于 0. 我们把几种虽然简单，然而常常用到的特殊情形总结如下：

命题 1 (1) 任一多项式都能整除它自己；

(2) 零多项式是任意多项式的倍式，而零多项式的倍式只有零多项式；

(3) 零次多项式(非零常数)是任意多项式的因式，而零次多项式的因式只有零次多项式.

证明：(1) 这是因为对任一多项式 $f(x)$，都有

$$f(x)=f(x)\cdot 1.$$

(2) 对于任一多项式 $f(x)$，有

$$0=0\cdot f(x).$$

所以 $f(x)$ 整除 0，即 $f(x)$ 是 0 的因式. 另外，如果 $f(x)$ 可被 0 整除，那么由定义，存在 $q(x)$ 使

$$f(x)=0\cdot q(x).$$

因此 $f(x)=0$，所以只有零多项式才能被零多项式整除.

(3) 设 c 是一个非零常数，由于对任一多项式 $f(x)$ 都有

$$f(x)=c\cdot\left(\frac{1}{c}f(x)\right),$$

所以 $c\mid f(x)$. 反之，如果 $f(x)$ 能整除 c，那么存在 $q(x)$ 使

$$c=f(x)q(x).$$

比较两边的次数知 $f(x)$ 必为零次多项式.

命题 1 给出了当 $g(x)$ 为零多项式或非零常数时，判断 $g(x)$ 能否整除 $f(x)$ 的方法. 一般情形，如果 $g(x)\neq 0$，可以用带余除法来判断 $g(x)$ 能否整除 $f(x)$.

定理 2 如果 $g(x)$ 是一个非零多项多，那么 $g(x)$ 整除 $f(x)$ 的充分必要条件是 $g(x)$ 除 $f(x)$ 所得的余式为 0.

证明：如果 $g(x)$ 除 $f(x)$ 的余式为 0，那么 $f(x)$ 可表成

$$f(x) = g(x)q(x).$$

所以根据定义，$g(x)\mid f(x)$.

如果 $g(x)$ 整除 $f(x)$，那么根据定义 4，存在 $q(x)$ 使

$$f(x) = g(x)q(x) = g(x)q(x) + 0.$$

由定理 1 知，$g(x)$ 除 $f(x)$ 的余式为 0. $\blacksquare$

定理 2 给出了当 $g(x)\neq 0$ 时，判断 $g(x)$ 能否整除 $f(x)$ 的一个具体的方法. 当 $g(x)\neq 0$ 且 $g(x)\mid f(x)$ 时，用 $g(x)$ 除 $f(x)$ 所得的商可以用

$$\frac{f(x)}{g(x)}$$

来表示

例 3 设 $f(x)=2x^4-3x^3+5x^2-6$，$g(x)=x^2-x-1$，判断 $g(x)$ 能否整除 $f(x)$.

解：用分离系数法进行除法：

$$
\begin{array}{r|rrrrr}
 & & & 2 & -1 & 6 \\ \cline{2-6}
1\ -1\ -1 & 2 & -3 & 5 & 0 & -6 \\
 & 2 & -2 & -2 & & \\ \cline{2-6}
 & & -1 & 7 & 0 & \\
 & & -1 & 1 & 1 & \\ \cline{3-6}
 & & & 6 & -1 & -6 \\
 & & & 6 & -6 & -6 \\ \cline{4-6}
 & & & & 5 & 0
\end{array}
$$

因为 $g(x)$除 $f(x)$所得的余式为 $5x\neq 0$,所以 $g(x)\nmid f(x)$.

有时我们还可以应用已知公式来求出 $r(x)$,从而判断 $g(x)$ 是否是 $f(x)$的因式.

例 4 问 x^2+1 能否整除 x^4-2?

解:因为

$$x^4-2=(x^2+1)(x^2-1)-1,$$

所以
$$x^2+1\nmid x^4-2.$$

例 5 求 k,l,使 x^2+x+l 能整除 x^3+kx+1.

解:用 x^2+x+l 除 x^3+kx+1,得商式 $x-1$,余式为$(k-l+1)x+1+l$. 所以 x^2+x+l 整除 x^3+kx+1 的充要条件为

$$\begin{cases} k-l+1=0, \\ 1+l=0, \end{cases}$$

即
$$\begin{cases} k=-2, \\ l=-1. \end{cases}$$

有时我们还会遇到由一些已知条件来证明 $g(x)$整除 $f(x)$或 $g(x)$不能整除 $f(x)$的情形. 这时候,就要直接根据定义来证明.

例 6 证明:如果

$$g(x)\mid f_1(x)+f_2(x), g(x)\mid f_1(x)-f_2(x),$$

则
$$g(x)\mid f_1(x), \quad g(x)\mid f_2(x).$$

证:由假设,$g(x)\mid f_1(x)+f_2(x)$,$g(x)\mid f_1(x)-f_2(x)$,故有 $q_1(x)$,$q_2(x)$使

$$f_1(x)+f_2(x)=g(x)q_1(x),$$
$$f_1(x)-f_2(x)=g(x)q_2(x).$$

因此

$$f_1(x)=g(x)\left[\frac{1}{2}q_1(x)+\frac{1}{2}q_2(x)\right],$$

$$f_2(x)=g(x)\left[\frac{1}{2}q_1(x)-\frac{1}{2}q_2(x)\right].$$

根据定义

$$g(x) \mid f_1(x), \; g(x) \mid f_2(x).$$

在这个证明中反复应用了整除的定义，通过这个例题的证明，我们知道，假设 $g(x) \mid f(x)$，就是给了等式

$$f(x) = g(x)q(x).$$

而为了证明 $g(x) \mid f(x)$，只要把 $f(x)$ 表示成 $g(x)q(x)$ 的形式就可以了.

在讨论整除性问题时，除了具体计算或直接从定义验证外，还可应用命题 1 来讨论. 此外，还有一些以后经常用到的性质，我们把这些性质总结成一个命题，以便于以后应用.

命题 2 (1) 如果 $g(x) \mid f(x)$，$f(x) \mid g(x)$，那么 $f(x) = cg(x)$，式中 c 是一个非零常数.

(2) 如果 $g(x) \mid f(x)$，$h(x) \mid g(x)$，那么 $h(x) \mid f(x)$.

(3) 如果 $g(x) \mid f_1(x)$，$g(x) \mid f_2(x)$，那么对任意多项式 $u(x)$，$v(x)$，都有

$$g(x) \mid u(x)f_1(x) + v(x)f_2(x).$$

证明：(1) 设 $g(x) \mid f(x)$，$f(x) \mid g(x)$. 如果 $g(x)$，$f(x)$ 中有一个是零多项式，那么另一个一定也等于 0，因此任取一个非零常数 c，即有 $f(x) = cg(x)$.

如果 $f(x)$，$g(x)$ 都不等于零. 由假设有 $q_1(x)$，$q_2(x)$ 使

$$f(x) = g(x)q_1(x), \quad g(x) = f(x)q_2(x).$$

因此

$$f(x) = f(x)q_1(x)q_2(x).$$

由于 $f(x) \neq 0$，所以 $q_1(x)q_2(x)$ 也不为 0. 比较两边次数，得

$$\deg(q_1(x)q_2(x)) = 0.$$

因此 $\deg q_1(x) = 0$，$q_1(x)$ 是一个非零常数 c，$f(x) = cg(x)$.

(2) 如果 $g(x) \mid f(x)$，$h(x) \mid g(x)$，则有 $q_1(x)$，$q_2(x)$ 使

$$f(x) = g(x)q_1(x), \quad g(x) = h(x)q_2(x).$$

因此

$$f(x) = h(x)[q_2(x)q_1(x)],$$

所以

$$h(x)\mid f(x).$$

(3) 如果 $g(x)\mid f_1(x)$，$g(x)\mid f_2(x)$，则有 $q_1(x)$，$q_2(x)$使

$$f_1(x)=g(x)q_1(x),\quad f_2(x)=g(x)q_2(x).$$

因此

$$u(x)f_1(x)+v(x)f_2(x)=g(x)(u(x)q_1(x)+v(x)q_2(x)).$$

所以

$$g(x)\mid(u(x)f_1(x)+v(x)f_2(x)).\ |$$

命题 2 可以推广到几个多项式的情形.

命题 2′ (1) 如果 $f_1(x)\mid f_2(x),\cdots,f_{s-1}(x)\mid f_s(x),f_s(x)\mid f_1(x)(s\geqslant 2)$，则 $f_1(x)=cf_s(x)$，c 是一个非零常数.

(2) 如果 $f_1(x)\mid f_2(x),f_2(x)\mid f_3(x),\cdots,f_{s-1}(x)\mid f_s(x)(s\geqslant 2)$，则 $f_1(x)\mid f_s(x)$.

(3) 如果 $g(x)\mid f_1(x),g(x)\mid f_2(x),\cdots,g(x)\mid f_s(x)(s\geqslant 2)$，则对任意多项式 $u_1(x),u_2(x),\cdots,u_s(x)$都有：

$$g(x)\mid u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_s(x)f_s(x).$$

通常称 $u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_s(x)f_s(x)$为$f_1(x)$，$f_2(x),\cdots,f_s(x)$的一个组合. 因此命题 2′的(3)就可说成：如果 $g(x)$ 整除 $f_i(x)$ $(i=1,2,\cdots,s)$，那么 $g(x)$能整除 $f_1(x)$，$f_2(x),\cdots,f_s(x)$的任一个组合.

命题 2′可以和命题 2 同样地证明. 这个命题在以后经常用到，读者可以自己证明一下以加深印象.

最后，我们再来举几个例子.

例 7 问 l,m,n 满足什么条件，x^2+lx+1 能整除 x^3+mx+n?

解：用 x^2+lx+1 除 x^3+mx+n，得商式为 $x-l$，余式为$(l^2+m-1)x+(l+n)$，即

$$\begin{aligned}x^3+mx+n=&(x-l)(x^2+lx+1)\\&+[(l^2+m-1)x+(l+n)].\end{aligned}$$

所以 x^2+lx+1 整除 x^3+mx+n 的充分必要条件为

$$\begin{cases} l^2+m-1=0, \\ l+n=0, \end{cases}$$

即
$$\begin{cases} n=-l, \\ m=1-l^2. \end{cases}$$

例 8 证明:如果 $g(x)\mid f(x)$, $g(x)\nmid h(x)$,则 $g(x)\nmid f(x)+h(x)$.

证明: 由假设知

$$f(x)=g(x)q_1(x), \quad h(x)=g(x)q_2(x)+r_2(x),$$

其中 $r_2(x)\neq 0$,且 $\deg r_2(x)<\deg g(x)$. 于是

$$f(x)-h(x)=g(x)[q_1(x)+q_2(x)]+r_2(x),$$

$$\deg r_2(x)<\deg g(x).$$

所以 $g(x)\nmid f(x)+h(x)$.

需要注意的是,如果 $g(x)\nmid f(x)$, $g(x)\nmid h(x)$,那么 $g(x)$ 能否整除 $f(x)+h(x)$ 就不是一定的了. 请读者自己举例说明这一事实.

习　题　6.2(2)

1. 检验 $g(x)$ 是不是 $f(x)$ 的因式:

 (1) $f(x)=x^4-4x^3-1$, $g(x)=x^2-3x-1$;

 (2) $f(x)=x^5+x^4+x^3+x^2-3x+1$, $g(x)=x^3+2x+1$.

2. 求 $g(x)=x^3-x+1$ 除 $f(x)=x^5+3x^4-3x^3+kx^2+5x+l$ 所得的商式及余式. 并且确定 k,l 的值,使 $g(x)$ 整除 $f(x)$.

3. 求 l,m,使 $f(x)=x^3+lx^2+5x+2$ 能被 $g(x)=x^2+mx+1$ 整除.

4. 已知 x^2+x-2 能整除 x^4+x^3+lx+m. 求 l,m.

5. 证明:如果 $g(x)\mid f_1(x)+2f_2(x)$, $g(x)\mid 3f_1(x)+4f_2(x)$,那么 $g(x)\mid f_1(x)$, $g(x)\mid f_2(x)$.

6. (1)证明:如果 $h(x)\mid f(x)$, $h(x)\nmid g(x)$,那么 $h(x)\nmid f(x)+g(x)$;

 (2) 如果 $h(x)\nmid f(x)$, $h(x)\nmid g(x)$,那么 $h(x)$ 能整除 $f(x)+g(x)$ 吗? 试举

例说明.

3. 余数定理及其应用

当 $g(x)=ax+b(a\neq0)$ 是一个一次多项式时，用 $g(x)$ 除 $f(x)$ 所得的余式是一个常数，因此可以称其为**余数**. 在这种情形，商式和余都有简单的算法，就是有些读者已经熟悉的综合除法. 下面来介绍余数定理及综合除法，然后介绍将一个多项式表示成一次多项式的方幂和的方法.

设 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$，$c$ 是一个常数. 用 $f(c)$ 表示当 $x=c$ 时多项式 $f(x)$ 的值：

$$f(c)=a_nc^n+a_{n-1}c^{n-1}+\cdots+a_1c+a_0.$$

如果 $f(c)=0$，则称 c 是 $f(x)$ 的一个**根**. 因此 $f(x)$ 的根也就是方程式 $f(x)=0$ 的根(或解).

例 9 $f(x)=2x^4-6x^3+3x^2-2x+5$，

(1) 求 $f(x)$ 被 $x-2$ 除所得的商式及余数；

(2) 求 $f(2)$；

(3) 问：2 是不是 $f(x)$ 的根？

解：(1) 作长除法：

$$\begin{array}{r|rrrrr}
 & & 2 & -2 & -1 & -4 \\ \hline
1\ \ -2 & 2 & -6 & 3 & -2 & 5 \\
 & 2 & -4 & & & \\ \cline{2-6}
 & & -2 & 3 & & \\
 & & -2 & 4 & & \\ \cline{3-6}
 & & & -1 & -2 & \\
 & & & -1 & 2 & \\ \cline{4-6}
 & & & & -4 & 5 \\
 & & & & -4 & 8 \\ \cline{5-6}
 & & & & & -3
\end{array}$$

所以商式 $=2x^3-2x^2-x-4$，余数 $=-3$.

(2) $f(2)=2\cdot2^4-6\cdot2^3+3\cdot2^2-2\cdot2+5=-3$.

(3) 因为 $f(2)\neq0$,所以 2 不是 $f(x)$的根.

在这个例子中,$f(x)$被 $x-2$ 除所得的余数恰好等于 $f(2)$. 习题 6.2(1)第 7 题也给出了一个 $f(x)$被 $x-c$ 除的余数等于 $f(c)$ 的例子. 事实上,这是一个一般的规律. 我们来证明这个结论.

定理 3(余数定理) 多项式 $f(x)$被 $x-c$ 除,所得的余数就等于 $f(c)$.

证明:因为 $x-c$ 是一个一次多项式,所以,根据带余除法,存在一个多项式 $q(x)$及一个常数 r,使得

$$f(x)=q(x)(x-c)+r.$$

r 就是 $f(x)$被 $x-c$ 除所得的余数. 等式两边用 $x=c$ 代入,得

$$f(c)=q(c)(c-c)+r=r.$$

即

$$r=f(c). \mid$$

这个定理给出了一次多项式 $x-c$ 整除 $f(x)$的条件.

推论 (1) $f(x)$能被 $x-c$ 整除的充要条件是 $f(c)=0$;

(2) $f(x)$以 c 为根的充要条件是 $x-c$ 整除 $f(x)$.

以上结论对于多项式的讨论及求根问题都是很有用处的.

例 10 求 $f(x)=x^n+a^n$ 被 $x+a$ 除所得的余式.

解:因为 $f(-a)=(-a)^n+a^n=\begin{cases}2a^n, & \text{如果 } n \text{ 是偶数;}\\ 0, & \text{如果 } n \text{ 是奇数.}\end{cases}$

所以,当 n 是偶数时,余式是 $2a^n$;当 n 是奇数时,余式是 0.

综合除法是用一次多项式来除任一多项式的一种简便的方法. 因此在进一步讨论以一个一般的一次多项式除 $f(x)$的商式及余式问题以前,我们先来介绍一下综合除法. 为了说明综合除法的理论根据,首先来观察 $f(x)$被 $x-c$ 除所得的商式及余式与 $f(x)$的系数之间的关系. 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0 \quad (a_n\neq0).$$

由次数公式知,$f(x)$被 $x-c$ 除的商式是一个 $n-1$ 次多项式,

设为
$$q(x)=b_{n-1}x^{n-1}+b_{n-2}x^{n-2}+\cdots+b_1x+b_0,$$
记余数为 r. 于是
$$f(x)=(x-c)q(x)+r,$$
即
$$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$
$$=(x-c)(b_{n-1}x^{n-1}+\cdots+b_1x+b_0)+r.$$
比较同次项系数,得
$$\begin{cases} b_{n-1}=a_n, \\ b_{n-2}-cb_{n-1}=a_{n-1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ b_0-cb_1=a_1, \\ r-cb_0=a_0. \end{cases}$$
由此,得
$$\begin{cases} b_{n-1}=a_n, \\ b_{n-2}=a_{n-1}+cb_{n-1} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ b_0=a_1+cb_1 \\ r=a_0+cb_0. \end{cases}$$
为了求出 $q(x)$ 及 r,只要求出 $b_{n-1},b_{n-2},\cdots,b_2,b_1$ 及 r 就可以了. 综合除法就是一种计算 $b_{n-1},b_{n-2},\cdots,b_2,b_1,r$ 的简单方法. 方法是这样的:把 $a_n,a_{n-1},\cdots,a_1,a_0$ 写在第 1 行,把 c 写在右边,然后按照上面的递推关系:$b_{n-1}=a_n,b_k=a_{k+1}+cb_{k+1},k=n-2,n-1,\cdots,2,1,0;r=a_0+cb_0$,逐次算出 b_i 及 r:
$$\begin{array}{ccccc|c} a_n & a_{n-1} & \cdots & a_1 & a_0 & c \\ & cb_{n-1} & \cdots & cb_1 & cb_0 & \\ \hline a_n & a_{n-1}+cb_{n-1} & \cdots & a_1+cb_1 & a_0+cb_0 & \\ \| & \| & & \| & \| & \\ b_{n-1} & b_{n-2} & \cdots & b_0 & r & \end{array}$$

其中第 1 行是 $f(x)$按降幂排列的系数;第 2 行的数字是由第 3 行前面一个数乘上 c 而得;和 3 行是第 1,2 两行相加而得. 最后一行即为所求的 $q(x)$按降幂排列的系数及余数 r.

我们把例 9 用综合除法来重新计算一下.

$$\begin{array}{rrrrr|l} 2 & -6 & 3 & -2 & 5 & 2 \\ & 4 & -4 & -2 & -8 & \\ \hline 2 & -2 & -1 & -4 & -3 & \end{array}$$

所以商式$=2x^3-2x^2-x-4$,余数$=-3$.

可以看出,综合除法不仅书写起来比长除法简单得多,而且,由于综合除法把长除法中的减法变成了加法,计算起来也方便多了. 我们知道,$f(x)$被 $x-c$ 除的余数 $r=f(c)$. 所以综合除法可以用来计算 $f(c)$. 一般说来,这比把 c 直接代入要简便. 由此可知,综合除法也可用来判断 c 是否是多项式 $f(x)$的根.

例 11 求 k,使 $f(x)=x^4-5x^3+5x^2+kx+3$ 以 3 为根.

解: $f(x)$以 3 为根的条件是 $f(3)=0$. 用综合除法求 $f(3)$

$$\begin{array}{rrrrr|l} 1 & -5 & 5 & k & 3 & 3 \\ & 3 & -6 & -3 & 3k-9 & \\ \hline 1 & -2 & -1 & k-3 & 3k-6 & \end{array}$$

所以 $f(3)=3k-6$. 当 $k=2$ 时,$f(x)$以 3 为根.

当除式为 $x-c$ 形式时,综合除法是一个计算商式与余式的简便方法. 如果除式为一般的一次式 $ax-b(a\neq 0)$,能不能应用综合除法来计算呢?为了解决这个问题,首先需要弄清除式为 $ax-b$ 与 $x-c$ 时商式及余数之间的关系. 如果 $f(x)$被 $ax-b$ 除所得的商式为 $q(x)$,余数为 r,那么

$$f(x)=(ax-b)q(x)+r.$$

我们把这个等式改写成

$$f(x)=\left(x-\frac{b}{a}\right)[aq(x)]+r.$$

因此 $aq(x)$ 就是 $f(x)$ 被 $x-\dfrac{b}{a}$ 除所得的商式，而 r 是 $f(x)$ 被 $x-\dfrac{b}{a}$ 除所得的余数. 所以也可应用综合除法来进行计算.

命题 3 如果 $f(x)$ 被 $x-c$ 除所得的商式为 $q(x)$，余数为 r. 那么 $f(x)$ 被 $a(x-c)(a\neq 0)$ 除所得的商式为 $\dfrac{1}{a}q(x)$，余数仍为 r.

例 12 求 $2x^4-5x^3+4x^2-3x+1$ 被 $2x-3$ 除所得的商式及余数.

解：作综合除法：

$$\begin{array}{r|rrrrr|l}
 & 2 & -5 & 4 & -3 & 1 & \frac{3}{2} \\
 & & 3 & -3 & \frac{3}{2} & -\frac{9}{4} & \\
\hline
\frac{1}{2} & 2 & -2 & 1 & -\frac{3}{2} & -\frac{5}{4} & \\
\hline
 & 1 & -1 & \frac{1}{2} & -\frac{3}{4} & &
\end{array}$$

(最后一行的计算是把 $q(x)$ 乘上 $1/2$). 所以

$$商式=x^3-x^2+\frac{1}{2}x-\frac{3}{4}，余数=-\frac{5}{4}.$$

最后，作为综合除法的一个应用，介绍将一个多项式表示成一次多项式 $x-c$ 的方幂和的方法. 所谓把 n 次多项式 $f(x)$ 表示成 $x-c$ 的方幂和，就是要把 $f(x)$ 表示成：

$$f(x)=b_n(x-c)^n+b_{n-1}(x-c)^{n-1}+\cdots+b_1(x-c)+b_0$$

的形式. 问题是如何求系数 $b_n,b_{n-1},\cdots,b_1,b_0$. 我们把上式改写成

$$f(x)=[b_n(x-c)^{n-1}+b_{n-1}(x-c)^{n-2}+\cdots+b_1]\times(x-c)+b_0,$$

就可看出，b_0 就是 $f(x)$ 被 $x-c$ 除所得的余数；而 $b_n(x-c)^{n-1}+b_{n-1}(x-c)^{n-2}+\cdots+b_1$ 就是 $f(x)$ 被 $x-c$ 除所得的商式，我们把

它记作 $q_1(x)$. 又因

$$q_1(x) = b_n(x-c)^{n-1} + b_{n-1}(x-c)^{n-2} + \cdots + b_1$$
$$= [b_n(x-c)^{n-2} + b_{n-1}(x-c)^{n-3} + \cdots + b_2] \times (x-c) + b_1,$$

所以又可看到 b_1 是商式 $q_1(x)$ 被 $x-c$ 除所得的余数，而 $q_2(x) = b_n(x-c)^{n-2}+b_{n-1}(x-c)^{n-3}+\cdots+b_2$ 就是 $q_1(x)$ 被 $x-c$ 除所得的商. 这样逐次用 $x-c$ 除所得的商，所得到的余数就是 $b_0, b_1, \cdots, b_{n-1}, b_n$. 我们来举例说明应用综合除法将一个多项式表成 $x-c$ 的方幂和的方法.

例 13 将

$$f(x) = x^4 - 3x^2 + 3$$

表成 $x-1$ 的方幂和.

解:

1	0	-3	0	3	1
	1	1	-2	-2	
1	1	-2	-2	$1(=b_0)$	
	1	2	0		
1	2	0	$-2(=b_1)$		
	1	3			
1	3	$3(=b_2)$			
	1				
$1(=b_4)$	$4(=b_3)$				

所以，$f(x) = (x^3+x^2-2x-2)(x-1)+1$

$$= [(x^2+2x)(x-1)-2](x-1)+1$$

$$
\begin{aligned}
&=(x^3+2x)(x-1)^2-2(x-1)+1\\
&=[(x+3)(x-1)+3](x-1)^2-2(x-1)+1\\
&=(x+3)(x-1)^3+3(x-1)^2-2(x-1)+1\\
&=[(x-1)+4](x-1)^3\\
&\quad+3(x-1)^2-2(x-1)+1\\
&=(x-1)^4+4(x-1)^3\\
&\quad+3(x-1)^2-2(x-1)+1
\end{aligned}
$$

根据上面的分析和例子，可以看出：用 $x-c$ 去除 k 次多项式 $f(x)$，得商 q_1 及余 r_1；用 $x-c$ 去除 q_1，得商 q_2 及余 r_2；…；用 $x-c$ 去除 q_{k-1}，得商 q_k 及余 r_k. $q_k,r_k,r_{k-1},\cdots,r_2,r_1$ 就是在 $f(x)$ 表成 $x-c$ 的方幂和中从高次幂到低次幂按降幂排列的各项系数. 掌握了这个规则，就可以根据综合除法直接写出答案，而不必像在例 13 中那样逐步推导了. 下面我们再举一个例子.

例 14 将 $f(x)=2x^4+6x^3+3x^2+4x+1$ 表示成 $x+2$ 的方幂和.

解：

2	6	3	4	1	-2
	-4	-4	2	-12	
2	2	-1	6	$-11(=b_0)$	
	-4	4	-6		
2	-2	3	$0(=b_1)$		
	-4	12			
2	-6	$15(=b_2)$			
	-4				
$2(=b_4)$	$-10(=b_3)$				

所以 $2x^4+6x^3+3x^2+4x+1$

$$=2(x+2)^4-10(x+2)^3+15(x+2)^2-11.$$

例 15　将 $f(x)=(x-2)^4+(x-2)^3-3(x-2)^2+5$ 展成 x 的多项式.

解：当然可以应用二项展开公式来计算这个问题，但是这样做太麻烦.这个例题与例 13 有类似的地方，能不能应用例 13 中的方法来解决这个问题呢？为此，令 $y=x-2$，于是 $x=y+2$. 而

$$f(x)=f_1(y)=y^4+y^3-3y^2+5.$$

问题就归结为将 $f_1(y)$ 表成 $y+2$（即 x）的方幂和，因此可以用综合除法来解决这个问题：

$$\begin{array}{ccccc|c}
1 & 1 & -3 & 0 & 5 & -2 \\
 & -2 & 2 & 2 & -4 & \\
\hline
1 & -1 & -1 & 2 & 1 & \\
 & -2 & 6 & -10 & & \\
\hline
1 & -3 & 5 & -8 & & \\
 & -2 & 10 & & & \\
\hline
1 & -5 & 15 & & & \\
 & -2 & & & & \\
\hline
1 & -7 & & & &
\end{array}$$

所以　　$f(x)=x^4-7x^3+15x^2-8x+1.$

需要注意的是：如果要将 $f(x)$ 表成 $x-c$ 的方幂和，那么，就要用 $x-c$ 去除，所以作综合除法时，要把 c 写在右边；而如果要将 $x-c$ 的方幂和展成 x 的多项式，那么相当于将它表成 $(x-c)+c$ 的方幂和，所以要把 $-c$ 写在右边进行计算.

习　题　6.2(3)

1. 计算 $f(c)$：

(1) $f(x)=x^4-3x^3+5x^2-6x-1$，$c=2$；

(2) $f(x)=x^5-2x^4+3x^2-6$，$c=-3$；

(3) $f(x)=x^4+(1+i)x^3+(2-3i)x^2+x-1+i$，$c=1+i$.

2. 求 k,使 $f(x)=x^5-3k^2x^3+5kx^2-x+2$ 能被 $x-1$ 整除.
3. 将 $f(x)$表成 $x-c$ 的方幂和：
   (1) $f(x)=x^4+2x^2-5$, $c=3$;
   (2) $f(x)=x^3-6x^2+7x+8$, $c=-1$;
   (3) $f(x)=x^5$, $c=1$.
4. 将 $f(x)$展成 x 的多项式：
   (1) $f(x)=(x+2)^4+3(x+2)^3-5(x+2)+7$;
   (2) $f(x)=(x-\mathrm{i})^5+(1+3\mathrm{i})(x-\mathrm{i})^3-(x-\mathrm{i})^2+(x-\mathrm{i})+4\mathrm{i}$.

6.3 最大公因式

1. 最大公因式

如果多项式 $h(x)$既是 $f(x)$的因式,又是 $g(x)$的因式,那么 $h(x)$就称为 $f(x)$与 $g(x)$的一个**公因式**.任意两个多项式,它们总是有公因式的,例如非零常数就是它们的公因式.在公因式中,最重要的就是最大公因式.

定义 5 设 $f(x)$与 $g(x)$是两个多项式,如果多项式 $d(x)$满足下面两个条件：

(1) $d(x)$是 $f(x)$与 $g(x)$的公因式;

(2) $f(x)$与 $g(x)$的任何公因式都是 $d(x)$的因式.

就称 $d(x)$是 $f(x)$与 $g(x)$的一个**最大公因式**.

这一节介绍最大公因式的求法和一些重要性质.

首先来看几个特殊情形：

(1) 如果 $f(x)$与 $g(x)$都等于 0.那么 $f(x)$与 $g(x)$的最大公因式就等于 0.以后仅就 $f(x)$,$g(x)$不全为 0 的情形进行讨论.

(2) 如果 $g(x)\mid f(x)$,那么 $g(x)$就是 $f(x)$与 $g(x)$的一个公因式.而且,凡是 $f(x)$与 $g(x)$的公因式一定是 $g(x)$的因式,所以 $g(x)$就是 $f(x)$与 $g(x)$的一个最大公因式.

因为 0 是任意多项式的倍式,所以

(3)对于任何多项式 $f(x)$，$f(x)$是其本身与 0 的一个最大公因式. 因为非零常数 c 是任意多项式 $f(x)$的一个因式，所以，

(4) 设 c 是非零常数，则 c 是 c 与 $f(x)$的一个最大公因式.

下面讨论在一般的情况下，怎样求 $f(x)$与 $g(x)$的最大公因式. 根据上述(3)可以设 $f(x)$，$g(x)$不全为 0.

在讨论前，为叙述方便起见，我们先引进一个符号. 设 $f(x)$，$g(x)$是两个不全为 0 的多项式，如果 $d_1(x)$是它们的一个最大公因式. 那么根据最大公因式的定义可知，对于任何非零常数 c，$cd_1(x)$也是 $f(x)$与 $g(x)$的一个最大公因式. 另一方面，如果 $d_2(x)$也是 $f(x)$与 $g(x)$的一个最大公因式，那么，根据最大公因式的定义，有 $d_1(x)\mid d_2(x)$及 $d_2(x)\mid d_1(x)$. 因此 $d_1(x)$与 $d_2(x)$只能相差一个非零常数因子，所以，$f(x)$与 $g(x)$的全部最大公因式就是 $cd_1(x)$，这里 c 可以是任意非零常数. 在这些最大公因式之中，有一个且仅有一个，其首项系数是 1. 我们用$(f(x),g(x))$表示 $f(x)$与 $g(x)$的那个首项系数是 1 的最大公因式. 要注意：在应用符号$(f(x),g(x))$的时候，必须假设 $f(x)$，$g(x)$不全为零，否则，$f(x)$与 $g(x)$的最大公因式就等于零.

给了两个不全为零的多项式 $f(x)$与 $g(x)$. 假设 $g(x)\neq 0$，先用 $g(x)$去除 $f(x)$，如果 $g(x)$能整除 $f(x)$，那么 $g(x)$就是 $f(x)$与 $g(x)$的一个最大公因式，问题就解决了；如果 $g(x)$不能整除 $f(x)$，假设

$$f(x)=q(x)g(x)+r(x)\quad (r(x)\neq 0).$$

于是
$$r(x)=f(x)-q(x)g(x).$$

由于 $r(x)$是 $f(x)$与 $g(x)$的一个组合，因此它们的因式之间必然有密切的关系，下面证明一个引理.

引理 设 $f(x)$与 $g(x)$是不全为零的多项式，如果

$$f(x)=q(x)g(x)+r(x).$$

那么
$$(f(x),g(x))=(g(x),r(x)).$$

证明：因为 $f(x)$ 是 $g(x)$ 与 $r(x)$ 的一个组合，所以凡是 $g(x)$ 与 $r(x)$ 的公因式，一定也是 $f(x)$ 的因式，因而是 $f(x)$ 与 $g(x)$ 的公因式. 特别地，$(g(x),r(x))$ 也是 $f(x)$ 与 $g(x)$ 的公因式. 根据最大公因式的定义，有 $(g(x),r(x))\mid(f(x),g(x))$.

而且，由于

$$r(x)=f(x)-q(x)g(x),$$

$r(x)$ 也是 $f(x)$ 与 $g(x)$ 的一个组合. 因此，由同样的论证，可得 $(f(x),g(x))\mid(g(x),r(x))$. 于是 $(f(x),g(x))$ 与 $(g(x),r(x))$ 只能相差一个常数因子，因为它们的首项系数都等于 1，所以 $(f(x),g(x))=(g(x),r(x))$. ▌

这样，就可以把求 $(f(x),g(x))$ 的问题，化为求 $(g(x),r(x))$ 的问题.

例 1 求 $f(x)=x^5+3x^4+x^3+x^2+3x+1$

与 $\qquad g(x)=x^4+2x^3+x+2$

的最大公因式.

解：先用 $g(x)$ 除 $f(x)$：

$$\begin{array}{r|l}
 & x+1 \\
\hline
x^4+2x^3+x+2 & x^5+3x^4+x^3+x^2+3x+1 \\
 & x^5+2x^4\qquad +x^2+2x \\
\cline{2-2}
 & x^4+x^3\qquad +x+1 \\
 & x^4+2x^3\qquad +x+2 \\
\cline{2-2}
 & -x^3\qquad -1
\end{array}$$

得商 $q(x)=x+1$，余式 $r(x)=-x^3-1$. 即

$$f(x)=(x+1)g(x)+(-x^3-1).$$

由引理：$\qquad (f(x),g(x))=(g(x),r(x)).$

为了求 $(g(x),r(x))$，可以再用 $r(x)$ 去除 $g(x)$：

$$
\begin{array}{r|l}
 & -x-2 \\
\cline{2-2}
-x^3-1 & x^4+2x^3\quad +x+2 \\
 & x^4\qquad\quad\ \ +x \\
\cline{2-2}
 & \qquad 2x^3\quad\ \ +2 \\
 & \qquad 2x^3\quad\ \ +2 \\
\cline{2-2}
 & \qquad\qquad\qquad 0
\end{array}
$$

这说明 $r(x)\mid g(x)$，$r(x)$是 $g(x)$与 $r(x)$的一个最大公因式. 因此

$$(f(x),g(x)) = (g(x),r(x)) = x^3 + 1.$$

要注意，在引理中，并不要求 $r(x)$的次数比 $g(x)$的次数低，也就是说，$r(x)$不一定是 $f(x)$被 $g(x)$除所得的余式. 这个事实，在以后的应用中可以看到，是比较有用的. 但是，在具体计算 $f(x)$与 $g(x)$的最大公因式时，一般总是取 $r(x)$为 $g(x)$除 $f(x)$所得的余式，这样，可以把多项式的次数降低，从而算出$(f(x),g(x))$. 这一点在下面介绍的辗转相除法中是很关键的.

在上面的例题中，我们进行了两次除法，就求出了 $f(x)$与 $g(x)$的最大公因式. 如果 $r(x)\nmid g(x)$，那么，这时又有一个商式 $q_1(x)$及余式 $r_1(x)$：

$$g(x) = q_1(x)r(x) + r_1(x),$$

而
$$(g(x),r(x)) = (r(x),r_1(x)).$$

这个步骤可以一直进行下去，直到求出最大公因式为止. 下面写出一般的作法：

为了求出 $f(x)$与 $g(x)$的最大公因式，先用 $g(x)(\neq 0)$除 $f(x)$，得到商 $q_1(x)$及余式 $r_1(x)$；如果 $r_1(x)\neq 0$，再用 $r_1(x)$除 $g(x)$，得商 $q_2(x)$及余式 $r_2(x)$；如果 $r_2(x)\neq 0$，再用 $r_2(x)$除 $r_1(x)$，得商 $q_3(x)$及余式 $r_3(x)$；这样继续下去. 因为

$$\deg(g(x)) > \deg(r_1(x)) > \deg(r_2(x)) > \cdots \geqslant 0.$$

因此在有限步之后，一定有一个余数等于 0. 于是我们得到一串等

式：

$$
\begin{aligned}
f(x) &= q_1(x)g(x) + r_1(x)\\
g(x) &= q_2(x)r_1(x) + r_2(x)\\
r_1(x) &= q_3(x)r_2(x) + r_3(x)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
r_{k-2}(x) &= q_k(x)r_{k-1}(x) + r_k(x)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
r_{s-3}(x) &= q_{s-1}(x)r_{s-2}(x) + r_{s-1}(x)\\
r_{s-2}(x) &= q_s(x)r_{s-1}(x) + r_s(x)\\
r_{s-1}(x) &= q_{s+1}(x)r_s(x) + 0,
\end{aligned}
\tag{1}
$$

其中 $r_1(x),r_2(x),\cdots,r_s(x)$都不等于 0. 因此

$$
\begin{aligned}
(f(x),g(x)) &= (g(x),r_1(x)) = (r_1(x),r_2(x)) = \cdots\\
&= (r_{k-1}(x),r_k(x)) = \cdots\\
&= (r_{s-3}(x),r_{s-2}(x))\\
&= (r_{s-2}(x),r_{s-1}(x))\\
&= (r_{s-1}(x),r_s(x))\\
&= cr_s(x).
\end{aligned}
$$

这里 c 是一个非零常数，使得 $cr_s(x)$是一个首项系数为 1 的多项式.

这种方法叫做**辗转相除法**. 因为要用到余式反复作除法，所以在作除法时，最好把商写在两边，以便逐次进行运算. 下面举例说明这个方法.

例 2 设$f(x)=x^7+x^5+x^4-x^3+2x^2-x+1$,

$g(x)=x^6+x^5+3x^3+2x+1$.

求$(f(x),g(x))$.

解：作辗转相除法：

$q_2(x)=\frac{x}{2}+1$	$g(x)=x^6+x^5\quad+3x^3\quad+2x+1$	$f(x)=x^7\quad+x^5+x^4-x^3+2x^2-x+1$	$q_1(x)=x-1$
	$x^6-x^5+x^4\quad+x$	$x^7+x^6\quad+3x^4\quad+2x^2+x$	
	$2x^5-x^4+3x^3\quad+x+1$	$-x^6+x^5-2x^4-x^3\quad-2x+1$	
	$2x^5-2x^4+2x^3\quad+2$	$-x^6-x^5\quad-3x^3\quad-2x-1$	
$q_4(x)=\frac{x}{6}+\frac{2}{9}$	$r_2(x)=x^4+x^3\quad+x-1$	$r_1(x)=2x^5-2x^4+2x^3\quad+2$	$q_3(x)=2x-4$
	$x^4-\frac{1}{3}x^3+\quad x^2-\frac{1}{3}x$	$2x^5+2x^4\quad+2x^2-2x$	
	$\frac{4}{3}x^3-\quad x^2+\frac{4}{3}x-1$	$-4x^4+2x^3-2x^2+2x+2$	
	$\frac{4}{3}x^3-\frac{4}{9}x^2+\frac{4}{3}x-\frac{4}{9}$	$-4x^4-4x^3\quad-4x+4$	
	$r_4(x)=-\frac{5}{9}x^2\quad-\frac{5}{9}$	$r_3(x)=6x^3-2x^2+6x-2$	$q_5(x)=\frac{-54}{5}x+\frac{18}{5}$
		$6x^3\quad+6x$	
		$-2x^2\quad-2$	
		$-2x^2\quad-2$	
		$r_5(x)=0$	

第 1 步：用 $g(x)$除 $f(x)$，得商式 $q_1(x)$，余式 $r_1(x)\neq 0$；
第 2 步：用 $r_1(x)$除 $g(x)$，得商式 $q_2(x)$，余式 $r_2(x)\neq 0$；
第 3 步：用 $r_2(x)$除 $r_1(x)$，得商式 $q_3(x)$，余式 $r_3(x)\neq 0$；
第 4 步：用 $r_3(x)$除 $r_2(x)$，得商式 $q_4(x)$，余式 $r_4(x)\neq 0$；
第 5 步：用 $r_4(x)$除 $r_3(x)$，得商式 $q_5(x)$，余式 $r_5(x)=0$.
因此 $r_4(x)\mid r_3(x)$，计算到此结束. 我们得到一串等式：

$$\begin{aligned}
f(x) &= q_1(x)g(x) + r_1(x) \\
g(x) &= q_2(x)r_1(x) + r_2(x) \\
r_1(x) &= q_3(x)r_2(x) + r_3(x) \\
r_2(x) &= q_4(x)r_3(x) + r_4(x) \\
r_3(x) &= q_5(x)r_4(x).
\end{aligned}$$

所以
$$\begin{aligned}
(f(x),g(x)) &= (g(x),r_1(x)) = (r_1(x),r_2(x)) \\
&= (r_2(x),r_3(x)) = (r_3(x),r_4(x)) \\
&= -\frac{9}{5}r_4(x) = x^2+1.
\end{aligned}$$

辗转相除法给出了求最大公因式的一个具体方法，因而也就证明了最大公因式的存在性.

为了计算简便起见，在作辗转相除法时，可以应用分离系数除法，即只写出各项的系数，不必写出 x 的方幂. 在应用分离系数法时，一定要注意在缺项的地方补上零系数.

还有，从例 2 可以看到：在作除法过程中，有时会出现分数，这样，不但计算复杂，而且容易出错，能不能避免这个困难呢？我们知道：$f(x)$与 $cf(x)$（c 是非零常数）有相同的因式，因此$(f(x),g(x))=(cf(x),g(x))$. 这说明，在计算过程中，必要时，可以用一个非零常数去乘某一个多项式，并不影响所得的结果. 下面举例说明.

例 3 求 $f(x)=x^5+2x^4-7x^3-8x-2$

与
$$g(x)=2x^4-2x^3+5x^2-2x+3$$

的最大公因式.

解:

		1	2	−7	0	−8	−2		2	−2	5	−2	3	2 −4
1 3	(×2)	2	4	−14	0	−16	−4		2	2	2	2		
		2	−2	5	−2	3				−4	3	−4	3	
			6	−19	2	−19	−4			−4	−4	−4	−4	
			6	−6	15	−6	9				7	0	7	
	$\left(\times\frac{-1}{13}\right)$			−13	−13	−13	−13	$\left(\times\frac{1}{7}\right)$			1	0	1	
				1	1	1	1							
				1	0	1								
					1	0	1							
					1	0	1							
							0							

所以 $$(f(x),g(x))=x^2+1.$$

习　题　6.3(1)

1. 求$(f(x),g(x))$:

(1) $f(x)=x^4+2x^3-2x^2-4x+1$, $g(x)=x^3+2x^2+1$;

(2) $f(x)=x^4-4x^3+x+2$, $g(x)=x^3-2x+1$;

(3) $f(x)=2x^4-5x^2+2$, $g(x)=2x^4-6\sqrt{2}x^3+5x^2+4\sqrt{2}x-2$;

(4) $f(x)=x^4+x^3+2x^2+x+1$, $g(x)=x^3+2x^2+2x+1$.

2. 设$f(x)$,$g(x)$是不全为0的多项式.证明:

(1) $(f(x),g(x))=(f(x)+g(x),g(x))$;

(2) $(f(x),g(x))=(f(x)+g(x),f(x)-g(x))$.

下面介绍最大公因式的一个重要性质:

定理 4　设$d(x)$是多项式$f(x)$,$g(x)$的一个最大公因式,那么$d(x)$一定可以表成$f(x)$与$g(x)$的组合,即可找到多项式$u(x)$与$v(x)$,使得

$$d(x)=u(x)f(x)+v(x)g(x).$$

证明 如果 $f(x)$与 $g(x)$中有一个能整除另一个，譬如说：如果 $g(x)\mid f(x)$. 那么 $g(x)$就是 $f(x),g(x)$的一个最大公因式. $d(x)$与 $g(x)$只能相差一个非零的常数因子，因此

$$d(x)=cg(x)=0\cdot f(x)+cg(x).$$

$f(x)\mid g(x)$的情形与此相仿.

现在设 $f(x),g(x)$互相都不能整除. 我们来考察辗转相除法的等式组(1). 首先，从(1)中倒数第 2 个等式可得到

$$r_s(x)=r_{s-2}(x)-q_s(x)r_{s-1}(x).$$

再由倒数第 3 个等式，得到

$$r_{s-1}(x)=r_{s-3}(x)-q_{s-1}(x)r_{s-2}(x),$$

代入上式，消去 $r_{s-1}(x)$，得

$$r_s(x)=[1+q_{s-1}(x)q_s(x)]r_{s-2}(x)-q_s(x)r_{s-3}(x).$$

然后应用同样的方法，逐次消去 $r_{s-2}(x),r_{s-3}(x),\cdots,r_1(x)$. 得

$$r_s(x)=u_1(x)f(x)+v_1(x)g(x).$$

设
$$d(x)=cr_s(x),$$

其中 c 是一个非零常数. 于是令

$$u(x)=cu_1(x),v(x)=cv_1(x).$$

从而得
$$d(x)=u(x)f(x)+v(x)g(x).\ |$$

从定理 4 可知：当 $f(x),g(x)$不全为 0 时，$(f(x),g(x))$可以表成 $f(x),g(x)$的组合，而且定理 4 的证明给出了具体计算 $u(x),v(x)$的方法. 下面举例说明.

例 4 将例 2 中的$(f(x),g(x))$表成 $f(x),g(x)$的组合.

解：由例 2 中的一些等式，得

$$\begin{aligned}r_4(x)&=r_2(x)-q_4(x)r_3(x)\\&=r_2(x)-q_4(x)[r_1(x)-q_3(x)r_2(x)]\\&=[1+q_3(x)q_4(x)]r_2(x)-q_4(x)r_1(x)\\&=[1+q_3(x)q_4(x)][g(x)-q_2(x)r_1(x)]-q_4(x)r_1(x)\\&=[1+q_3(x)q_4(x)]g(x)\end{aligned}$$

$$
\begin{aligned}
&\quad -[q_2(x)+q_4(x)+q_2(x)q_3(x)q_4(x)]r_1(x)\\
&=[1+q_3(x)q_4(x)]g(x)\\
&\quad -[q_2(x)+q_4(x)+q_2(x)q_3(x)q_4(x)]\\
&\quad \times[f(x)-q_1(x)g(x)]\\
&=[-q_2(x)-q_4(x)-q_2(x)q_3(x)q_4(x)]f(x)\\
&\quad +[1+q_1(x)q_2(x)+q_1(x)q_4(x)+q_3(x)q_4(x)\\
&\quad +q_1(x)q_2(x)q_3(x)q_4(x)]g(x).
\end{aligned}
$$

将 $q_1(x),q_2(x),q_3(x),q_4(x)$代入，经过计算整理，得

$$
\begin{aligned}
r_4(x)&=-\frac{5}{9}x^2-\frac{5}{9}\\
&=\left(-\frac{1}{6}x^3-\frac{2}{9}x^2-\frac{1}{3}\right)f(x)\\
&\quad+\left(\frac{1}{6}x^4+\frac{1}{18}x^3+\frac{1}{9}x^2+\frac{1}{9}x-\frac{2}{9}\right)g(x).
\end{aligned}
$$

因此

$$
\begin{aligned}
(f(x),g(x))&=x^2+1\\
&=\left(\frac{3}{10}x^3+\frac{2}{5}x^2+\frac{3}{5}\right)f(x)\\
&\quad+\left(-\frac{3}{10}x^4-\frac{1}{10}x^3-\frac{1}{5}x^2-\frac{1}{5}x+\frac{2}{5}\right)g(x).
\end{aligned}
$$

注意，在计算过程中，每次只消去一个 $r_i(x)$. 例如在第 1 步中只将 $r_3(x)$的表示式代入，而保留 $r_2(x)$不动. 因为 $r_3(x)$的表示式中包含 $r_2(x)$，可以在下一步中一次将 $r_2(x)$消去. 由于 $r_2(x)$中还包含 $g(x)$，若在第 1 步中就将 $r_2(x)$的表示式也代入，则会同时出现 $r_2(x),r_1(x),g(x)$等 3 个多项式，计算起来反而麻烦了.

多项式的最大公因式在实际问题中常常会用到，而且有时还需要求出表示式$(f(x),g(x))=u(x)f(x)+v(x)g(x)$中的 $u(x)$与 $v(x)$. 在 $f(x),g(x)$的次数较高时，$(f(x),g(x)),u(x),v(x)$的计算是比较复杂的，常常需要用电子计算机来计算. 在用电子计

算机计算$(f(x),g(x))$及$u(x),v(x)$时，若按证明定理的方法来进行，那就需要逐次把$f(x),g(x),q_1(x),r_1(x),\cdots,q_s(x),r_s(x)$都存储起来，然后再求$u(x)$及$v(x)$. 这样，存储量比较大，所以，通常采用下述顺次计算$r_1(x),r_2(x),\cdots$的方法.

设 $\qquad f(x)=q_1(x)g(x)+r_1(x),$

记 $\qquad r_1(x)=u_1(x)f(x)+v_1(x)g(x);$

$$g(x)=q_2(x)r_1(x)+r_2(x),$$

$$r_2(x)=u_2(x)f(x)+v_2(x)g(x);$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$r_{k-3}(x)=q_{k-1}(x)r_{k-2}(x)+r_{k-1}(x),$$

$$r_{k-1}(x)=u_{k-1}(x)f(x)+v_{k-1}(x)g(x);$$

$$r_{k-2}(x)=q_k(x)r_{k-1}(x)+r_k(x),$$

$$r_k(x)=u_k(x)f(x)+v_k(x)g(x);$$

$$r_{k-1}(x)=q_{k+1}(x)r_k(x)+r_{k+1}(x),$$

$$r_{k+1}(x)=u_{k+1}(x)f(x)+v_{k+1}(x)g(x);$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

下面推导$u_k(x),v_k(x)$的递推公式：

因为
$$\begin{aligned}r_{k+1}(x)&=r_{k-1}(x)-q_{k+1}(x)r_k(x)\\&=[u_{k-1}(x)f(x)+v_{k-1}(x)g(x)]\\&\quad-q_{k+1}(x)[u_k(x)f(x)+v_k(x)g(x)]\\&=[u_{k-1}(x)-q_{k+1}(x)u_k(x)]f(x)\\&\quad+[v_{k-1}(x)-q_{k+1}(x)v_k(x)]g(x).\end{aligned}$$

所以
$$\begin{cases}u_{k+1}(x)=u_{k-1}(x)-q_{k+1}(x)u_k(x),\\v_{k+1}(x)=v_{k-1}(x)-q_{k+1}(x)v_k(x).\end{cases}$$

因此，$u_{k+1}(x),v_{k+1}(x)$的计算只依赖于$u_{k-1}(x),v_{k-1}(x),u_k(x),v_k(x)$及$q_{k+1}(x)$. 而$q_{k+1}(x)$及$r_{k-1}(x)$可由$r_{k-1}(x)$及$r_k(x)$算出. 由此看出：在计算过程中，只要存储$r_{k-1}(x),r_k(x);u_{k-1}(x),$

$u_k(x)$;$v_{k-1}(x)$,$v_k(x)$等 6 个多项式即可. 而且每一步的计算步骤都是一样的:用 $r_k(x)$去除 $r_{k+1}(x)$,求出 $q_{k+1}(x)$及 $r_{k+1}(x)$,然后再代入上述公式,求出 $u_{k+1}(x)$及 $v_{k+1}(x)$.

还需要交代一下初始的情况. 因为

$$f(x) = 1 \cdot f(x) + 0 \cdot g(x),$$
$$g(x) = 0 \cdot f(x) + 1 \cdot g(x).$$

而且 $f(x)$与 $g(x)$的地位相当于 $r_{k-1}(x)$与 $r_k(x)$的地位. 所以如果令

$$r_{-1}(x) = f(x), u_{-1}(x) = 1, v_{-1}(x) = 0;$$
$$r_0(x) = g(x), u_0(x) = 0, v_0(x) = 1.$$

那么上面总结的计算方法对 $k=0,1,2,\cdots$都能应用. 故可以先将这 6 个多项式输入,然后应用循环程序,逐次求出 $r_1(x)$,$u_1(x)$,$v_1(x)$;$r_2(x)$,$u_2(x)$,$v_2(x)$;$\cdots$. 直到求出最大公因式 $r_s(x)$及 $u_s(x)$,$v_s(x)$为止. 然后再将这 3 个多项式分别乘上一个适当的常数(为了使 $r_s(x)$的首项系数变为 1)即得到$(f(x),g(x))$,$u(x)$及 $v(x)$. 在每步过程中,都只需存储 6 个多项式. 且用循环计算,程序也很简单. 但是由于上述的方法每次都要代入两个多项式,比较繁琐,故在定理的证明中,采用直接推导 $r_s(x)$的表示式的方法.

下面证明定理 4 也是最大公因式的一个特征性质. 即

定理 5 如果 $d(x)$是 $f(x)$,$g(x)$的一个公因式,而且 $d(x)$可以表成 $f(x)$与 $g(x)$的组合,那么 $d(x)$一定是 $f(x)$与 $g(x)$的一个最大公因式.

证明: 只要证明 $d(x)$满足最大公因式定义中的第 2 个条件,即 $f(x)$与 $g(x)$的任一个公因式一定是 $d(x)$的因式. 而这一点,可以从 $d(x)$是 $f(x)$与 $g(x)$的组合这一性质直接得到. ▎

最后,举几个例子说明如何简化最大公因式的计算以及证明有关最大公因式的一些命题.

例 5 设 $f(x)=x^5+x^4-x^3+2x+3$,

$$g(x)=x^4-x^3+2x+3.$$

求$(f(x),g(x))$.

解：因为 $f(x)-g(x)=x^5$，

所以$(f(x),g(x))=(x^5,g(x))$.

而 x^5 的因式只有 $cx^k(c\neq 0,k=0,1,2,3,4,5)$；$g(x)$的常数项不等于零，所以$(x^5,g(x))=1$. 因此

$$(f(x),g(x))=1.$$

例 6 假设$(f(x),g(x))=u(x)f(x)+v(x)g(x)$.

求证 $$(u(x),v(x))=1.$$

证明：设

$$f(x)=(f(x),g(x))\,q_1(x),$$

$$g(x)=(f(x),g(x))q_2(x).$$

用$(f(x),g(x))$除$(f(x),g(x))=u(x)f(x)+v(x)g(x)$两端，得

$$1=u(x)q_1(x)+v(x)q_2(x).$$

因此，凡是 $u(x),v(x)$的公因式都是 1 的因式，只能是非零常数.

所以 $$(u(x),v(x))=1.$$

前面例 3 中介绍了如何避免在辗转相除过程中出现分数的方法. 是用一个非零常数去乘某一个多项式. 这个方法在只求$(f(x),g(x))$而不要求计算 $u(x)$和 $v(x)$时，是完全行得通的. 但当需要计算 $u(x)$和 $v(x)$时，由于用一个非零常数乘某一个多项式，会引起商式和余式相应的改变. 所以在计算时，必须将所作的变换记下来，以便能准确地计算 $u(x),v(x)$. 下面我们通过例 7 来说明这一点.

例 7 设$f(x)=x^6-7x^4+8x^3-7x+7$，

$$g(x)=3x^5-7x^3+3x^2-7.$$

求$(f(x),g(x))$，并求 $u(x)$和 $v(x)$，使

$$(f(x),g(x))=u(x)f(x)+v(x)g(x).$$

解：作辗转相除法：

$q_2=6$	9	$g=3$		0	-7	3	0	-7
		$4g=12$		0	-28	12	0	-28
			12	-18	0	12	-18	
				18	-28	0	18	-28
				18	-27	0	18	-27
		$r_2=$			-1	0	0	-1
		$-r_2=$			1	0	0	1

$f=1$	0	-7	8	0	-7	7	
$3f=3$	0	-21	24	0	-21	21	$q_1=1 \quad 0$
3	0	-7	3	0	-7		
$r_1=$		-14	21	0	-14	21	
$-\frac{1}{7}$ $r_1=$		2	-3	0	2	-3	$q_3=2 \quad -3$
		2	0	0	2		
			-3	0	0	-3	
			-3	0	0	-3	
						$r_3=0$	

所以 $$(f(x),g(x))=-r_2(x)=x^3+1.$$

为了计算 $u(x)$ 和 $v(x)$，先写下每一步的带余除法. 必须注意：要把所乘的非零常数写入等式.

$$3f(x)=q_1(x)g(x)+r_1(x),$$

$$4g(x)=q_2(x)\left[-\frac{1}{7}r_1(x)\right]+r_2(x),$$

$$-\frac{1}{7}r_1(x)=q_3(x)[-r_2(x)]+0.$$

因此 $$\begin{aligned} r_2(x)&=4g(x)-q_2(x)\left[-\frac{1}{7}r_1(x)\right]\\ &=4g(x)+\frac{1}{7}q_2(x)[3f(x)-q_1(x)g(x)]\\ &=\frac{3}{7}q_2(x)f(x)+\left[4-\frac{1}{7}q_1(x)q_2(x)\right]g(x).\end{aligned}$$

因为 $$(f(x),g(x))=-r_2(x)=x^3+1,$$

所以 $$u(x)=-\frac{3}{7}q_2(x)=-\frac{18}{7}x-\frac{27}{7},$$

$$\begin{aligned} v(x)&=-4+\frac{1}{7}q_1(x)q_2(x)=-4+\frac{1}{7}x(6x+9)\\ &=\frac{6}{7}x^2+\frac{9}{7}x-4.\end{aligned}$$

习 题 6.3(2)

1. 求 $(f(x),g(x))$，并求 $u(x)$ 和 $v(x)$，使 $(f(x),g(x))=u(x)f(x)+v(x)g(x)$.

   (1) $f(x)=x^4+2x^3-4x-4$，$g(x)=x^4+2x^3-x^2-4x-2$；

   (2) $f(x)=2x^4-x^3-8x^2+3x+4$，$g(x)=2x^3-x^2-4x-3$；

   (3) $f(x)=x^4-2x^3+4x^2+4x+1$，$g(x)=x^2+x-1$.

2. 证明 $\left(1+x+\frac{1}{2!}x^2+\cdots+\frac{x^n}{n!},1+x+\frac{1}{2!}x^2+\cdots+\frac{1}{(n-1)!}x^{n-1}\right)=1$

3. 设 $f(x)$，$g(x)$ 是不全为零的多项式，求证：

$$\left(\frac{f(x)}{(f(x),g(x))},\frac{g(x)}{(f(x),g(x))}\right)=1.$$

4. $h(x)$是一个首项系数为 1 的多项式. 求证

$$(f(x)h(x),g(x)h(x))=(f(x),g(x))h(x).$$

2. 互素

定义 6 如果多项式 $f(x)$与 $g(x)$的最大公因式是非零常数，即$(f(x),g(x))=1$. 那么就称 $f(x)$与 $g(x)$**互素**.

从定义可以看出：两个互素的多项式除去零次多项式外，没有其他的公因式；反之也对.

定理 6 两个多项式 $f(x)$与 $g(x)$互素的充分必要条件是存在多项式 $u(x)$与 $v(x)$，使得

$$u(x)f(x)+v(x)g(x)=1.$$

证明：如果 $f(x)$与 $g(x)$互素，那么$(f(x),g(x))=1$. 因此根据定理 4，存在 $u(x)$与 $v(x)$，使得

$$u(x)f(x)+v(x)g(x)=1.$$

反之，如果$u(x)f(x)+v(x)g(x)=1$. 由于 1 是 $f(x)$与 $g(x)$的一个公因式，所以从定理 5 可知：$(f(x),g(x))=1$，即 $f(x)$与 $g(x)$是互素的. ▌

下面介绍关于互素多项式的一些重要性质，这些性质在以后经常用到.

定理 7 如果$(f(x),g(x))=1$，且 $f(x)\mid g(x)h(x)$. 那么 $f(x)\mid h(x)$.

证明：因为$(f(x),g(x))=1$，所以存在 $u(x)$与 $v(x)$，使

$$u(x)f(x)+v(x)g(x)=1.$$

等式两端同乘 $h(x)$，得

$$u(x)f(x)h(x)+v(x)g(x)h(x)=h(x).$$

因为 $f(x)$能整除 $f(x)$及 $g(x)h(x)$，所以 $f(x)$能整除它们的组合 $h(x)$. ▌

定理 8 如果 $g_1(x)\mid f(x)$，$g_2(x)\mid f(x)$，并且，$(g_1(x),g_2(x))=1$，那么 $g_1(x)g_2(x)\mid f(x)$.

证明：由于 $g_1(x)\mid f(x)$，故有 $q_1(x)$ 使 $f(x)=q_1(x)g_1(x)$. 现在 $g_2(x)\mid f(x)$，即 $g_2(x)\mid q_1(x)g_1(x)$，并且 $(g_1(x),g_2(x))=1$. 故由定理 7，有 $g_2(x)\mid q_1(x)$. 设 $q_1(x)=q_2(x)g_2(x)$，于是得 $f(x)=g_1(x)g_2(x)q_2(x)$，即

$$g_1(x)g_2(x)\mid f(x). \mid$$

应用这两个定理以及类似的方法，还可以证明关于互素多项式的其他一些性质. 下面把这些结论作为习题，留给读者.

习 题 6.3(3)

1. 证明如果 $f(x)$，$g(x)$ 不全为 0，且 $u(x)f(x)+v(x)g(x)=(f(x),g(x))$，那么 $(u(x),v(x))=1$.
2. 证明如果 $(f(x),g(x))=1$，$(f(x),h(x))=1$，那么 $(f(x),g(x)h(x))=1$.
3. 证明如果 $(f(x),g(x))=1$，那么 $(f(x)+g(x),f(x)g(x))=1$.

3. 几个多项式的最大公因式

上面介绍的最大公因式及互素的概念，都是对于两个多项式定义的. 事实上，对于多个多项式 $f_1(x),f_2(x),\cdots,f_s(x)$ $(s\geqslant 2)$ 也同样可以定义最大公因式及互素.

定义 5′ $d(x)$ 称为 $f_1(x),f_2(x),\cdots,f_s(x)$ $(s\geqslant 2)$ 的一个**最大公因式**，如果

1) $d(x)$ 是 $f_1(x),f_2(x),\cdots,f_s(x)$ 的公因式；

2) $f_1(x),f_2(x),\cdots,f_s(x)$ 的任何公因式都是 $d(x)$ 的因式.

如果 $f_1(x),f_2(x),\cdots,f_s(x)$ 全等于 0，那么它们的最大公因式等于 0；如果 $f_1(x),f_2(x),\cdots,f_s(x)$ 不全为 0，那么它们的最大公因式不等于 0，而且从定义 5′ 的第(2)个条件知道：与 $s=2$ 的情形一样，$f_1(x),f_2(x),\cdots,f_s(x)$ 的任意两个最大公因式都只能相

差一个非零的常数因子. 我们仍用$(f_1(x),f_2(x),\cdots,f_s(x))$表示$f_1(x),f_2(x),\cdots,f_s(x)$的首项系数为 1 的最大公因式. 于是$f_1(x),f_2(x),\cdots,f_s(x)$的任一最大公因式都可表成$c(f_1(x),f_2(x),\cdots,f_s(x))$,其中$c$是一个非零常数. 当然,对于任意的非零常数$c$,$c(f_1(x),f_2(x),\cdots,f_s(x))$都是$f_1(x),f_2(x),\cdots,f_s(x)$的最大公因式.

关于多个多项式的最大公因式的求法,有下述定理. 以后我们总假设$f_1(x),f_2(x),\cdots,f_s(x)$不全为 0. 一般情况下,不再重复申明.

定理 9 $(f_1(x),f_2(x),\cdots,f_s(x))=((f_1(x),\cdots,f_{s-1}(x)),f_s(x))$.

证明: 记$(f_1(x),f_2(x),\cdots,f_s(x))=d(x)$;

$(f_1(x),\cdots,f_{s-1}(x))=d_1(x)$;$(d_1(x),f_s(x))=d_2(x)$.

因为$d(x)\mid f_i(x)(i=1,2,\cdots,s)$. 所以$d(x)\mid(f_1(x),\cdots,f_{s-1}(x))$,而且$d(x)\mid f_s(x)$. 因此$d(x)\mid(d_1(x),f_s(x))$,即$d(x)\mid d_2(x)$. 另一方面,因为$d_2(x)\mid d_1(x)$,$d_2(x)\mid f_s(x)$,所以$d_2(x)\mid f_i(x)(i=1,2,\cdots,s)$,因此$d_2(x)\mid d(x)$. 于是$d(x)$与$d_2(x)$只能相差一个非零的常数因子. 然而$d(x)$与$d_1(x)$的首项系数都等于 1,所以$d(x)=d_1(x)$,即

$$\begin{aligned}&(f_1(x),f_2(x),\cdots,f_s(x))\\=&((f_1(x),\cdots,f_{s-1}(x)),f_s(x)).\ \blacksquare\end{aligned}$$

这个定理给出了求几个多项式的最大公因式的一个方法:

$$\begin{aligned}&(f_1(x),f_2(x),\cdots,f_s(x))\\=&((f_1(x),f_2(x),\cdots,f_{s-1}(x)),f_s(x))=\cdots\\=&((\cdots(f_1(x),f_2(x)),f_3(x)),\cdots,f_{s-1}(x)),f_s(x)).\end{aligned}$$

例 8 设$f_1(x)=x^6-7x^4+8x^3-7x+7$,

$f_2(x)=3x^5-7x^3+3x^2-7$,

$$f_3(x)=x^4+x^3-3x^2-4x-1,$$
$$f_4(x)=x^3+x^2-x-1.$$
求$(f_1(x),f_2(x),f_3(x),f_4(x))$.

解：根据$(f_1(x),f_2(x),f_3(x),f_4(x))=(((f_1(x),f_2(x)),f_3(x)),f_4(x))$逐步计算：

由于 $$(f_1(x),f_2(x))=x^3+1,$$
$$(x^3+1,f_3(x))=x+1,$$
$$(x+1,f_4(x))=x+1,$$
所以 $$(f_1(x),f_2(x),f_3(x),f_4(x))=x+1.$$

与两个多项式的情形一样，$f_1(x),\cdots,f_s(x)$的最大公因式也可以表示成$f_1(x),f_2(x),\cdots,f_s(x)$的组合. 而且，定理5也可以推广到多个多项式的情形. 这个事实可以由$(f_1(x),f_2(x),\cdots,f_s(x))=((f_1(x),\cdots,f_{s-1}(x)),f_s(x))$用归纳法证明，留给读者作为习题（习题6.3(4)第3题）.

互素的概念也可以推广：

定义 6′ 如果$(f_1(x),f_2(x),\cdots,f_s(x))=1$，就称$f_1(x),f_2(x),\cdots,f_s(x)$互素.

定理6也可以推广到多个多项式的情形. 也留作习题（习题6.3(4)第4题）.

最后提一下两两互素的概念：

多项式$f_1(x),f_2(x),\cdots,f_s(x)(s\geqslant2)$称为**两两互素的**，如果它们中的任意两个都是互素的.

当然，两两互素的多项式一定是互素的（为什么?）；但是互素的多项式不一定是两两互素的. 这个事实可以从下面的例子看出来.

例 9

(1) $f_1(x)=x+1,f_2(x)=x-1,f_3(x)=x^2+1$，这3个多项式是两两互素的.

(2) $g_1(x)=x+1, g_2(x)=x-1, g_3(x)=x^2-1$,这 3 个多项式是互素的,但不是两两互素的.

习 题 6.3(4)

1. 计算$(f_1(x),f_2(x),f_3(x))$,并求$u_1(x),u_2(x),u_3(x)$,使得$(f_1(x),f_2(x),f_3(x))=u_1(x)f_1(x)+u_2(x)f_2(x)+u_3(x)f_3(x)$.

   (1) $f_1(x)=x^5+3x^4+x^3+x^2+3x+1$,

   $f_2(x)=x^4+2x^3+x+2, f_3(x)=x^3+3x^2+x-1$;

   (2) $f_1(x)=x^5+3x^4+x^3-5x^2+2x-2$,

   $f_2(x)=x^5+2x^4+3x^2-4x-2, f_3(x)=x^4-x^3-x^2+2x-1$.

2. 证明:$(f_1(x),f_2(x),f_3(x),f_4(x))=((f_1(x),f_2(x)),((f_3(x),f_4(x)))$.

3. (1) 证明$f_1(x),f_2(x),\cdots,f_s(x)$的任一个最大公因式都可以表成$f_1(x),f_2(x),\cdots,f_s(x)$的组合;

   (2) 如果$d(x)$是$f_1(x),f_2(x),\cdots,f_s(x)$的一个公因式,而且$d(x)$可以表示成$f_1(x),f_2(x),\cdots,f_s(x)$的组合,那么$d(x)$一定是$f_1(x),f_2(x),\cdots,f_s(x)$的一个最大公因式.

4. 多项式$f_1(x),f_2(x),\cdots,f_s(x)$互素的充分必要条件是:存在多项式$u_1(x),u_2(x),\cdots,u_s(x)$,使得

$$u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_s(x)f_s(x)=1.$$

5. 已知$g_1(x)\mid f(x), g_2(x)\mid f(x), g_3(x)\mid f(x)$.

   (1) 证明:如果$g_1(x),g_2(x),g_3(x)$两两互素,那么一定有

$$g_1(x)g_2(x)g_3(x)\mid f(x);$$

   (2) 举例说明:如果$g_1(x),g_2(x),g_3(x)$互素,但不是两两互素,那么不一定有

$$g_1(x)g_2(x)g_3(x)\mid f(x).$$

6.4 因式分解定理

在这一节中,取定一个数域 **P** 作为系数域,来讨论数域 **P** 上多项式环 **P**$[x]$中的多项式的因式分解问题.

1. 不可约多项式

任意一个多项式 $f(x)$ 总有以下两类明显的因式：非零常数 c 及 $f(x)$ 的非零常数倍 $cf(x)$. 如果除此以外不再有其他因式，那么，这样的多项式就称为不可约的；否则，就称为可约的. 我们把这个概念写成一个定义.

定义 7 设 $p(x)$ 是数域 **P** 上一个次数 $\geqslant 1$ 的多项式. 如果 $p(x)$ 不能表示成数域 **P** 上两个次数比 $p(x)$ 低的多项式的乘积，那么 $p(x)$ 就称为数域 **P** 上的**不可约多项式**.

从定义可知：一次多项式总是不可约的. 由定义还可推知：如果 $p(x)$ 是一个不可约多项式，那么 $p(x)$ 的因式只有 c 及 $cp(x)$（这里 c 可以是任意非零常数），而且 $cp(x)$ 也是不可约的.

一个多项式是不是不可约多项式，依赖于系数域. 例如：x^2+2 是实数域上的不可约多项式. 而在复数域上则可以分解成两个一次因式的乘积：

$$x^2+2=(x+\sqrt{2}i)(x-\sqrt{2}i),$$

就成为可约的了. 又如：x^2-2 作为有理系数多项式是不可约的，而作为实系数多项式则是可约的：

$$x^2-2=(x-\sqrt{2})(x+\sqrt{2}).$$

不可约多项式有下述重要性质：

定理 10 设 $p(x)$ 是一个不可约多项式，那么

(1) 对于任一多项式 $f(x)$，或者 $p(x)\mid f(x)$；或者 $f(x)$ 与 $p(x)$ 互素.

(2) 对于任意两个多项式 $f(x)$ 与 $g(x)$，从 $p(x)\mid f(x)g(x)$ 可推出 $p(x)\mid f(x)$ 或 $p(x)\mid g(x)$.

证明：(1) 设 $(f(x),p(x))=d(x)$，则因 $d(x)\mid p(x)$，所以 $d(x)=1$ 或 $cp(x)$（$c\neq 0$，$cp(x)$ 的首项系数是 1）. 如果 $d(x)=1$，那么 $p(x)$ 与 $f(x)$ 互素；如果 $d(x)=cp(x)$，那么 $p(x)\mid f(x)$.

(2) 如果 $p(x)\mid f(x)$，那么结论已经成立；如果 $p(x)\nmid f(x)$，

那么由(1)知：$(p(x),f(x))=1$. 由定理 7，即得 $p(x)\mid g(x)$. |

定义 10 中的第(2)部分可以推广到多个多项式的情形：

如果不可约多项式 $p(x)$ 可以整除一些多项式 $f_1(x),f_2(x),\cdots,f_s(x)$ 的乘积：$p(x)\mid f_1(x)f_2(x)\cdots f_s(x)$，那么 $p(x)$ 一定能整除这些多项式中的一个.

我们用数学归纳法来证明这个结论.

当 $s=2$ 时，就是定理 10 中的情形，已经证明.

假设 $s=k$ 时结论成立. 我们来证 $s=k+1$ 时结论也成立. 设 $p(x)\mid f_1(x)f_2(x)\cdots f_{k+1}(x)$，即

$$p(x)\mid (f_1(x)\cdots f_k(x))f_{k+1}(x).$$

把 $(f_1(x)\cdots f_k(x))f_{k+1}(x)$ 看成 $f_1(x)\cdots f_k(x)$ 与 f_{k+1} 的乘积，那么由定理 10，得到 $p(x)\mid f_1(x)\cdots f_k(x)$ 或 $p(x)\mid f_{k+1}(x)$. 如果 $p(x)\mid f_{k+1}(x)$，那么结论已经成立；如果 $p(x)\mid f_1(x),\cdots,f_k(x)$，由归纳法假设，$p(x)$ 必能整除 $f_1(x),\cdots,f_k(x)$ 中的一个. 因此，$p(x)$ 总能整除 $f_1(x),\cdots,f_k(x),f_{k+1}(x)$ 中的一个.

由数学归纳法原则知结论普遍成立. |

最后，有必要指出：定理 10 中关于不可约多项式的两条性质，也是多项式不可约的充分条件. 把这两个结论留作习题（习题 6.4(1)第 2,3 题）.

习　题　6.4(1)

1. 设 $p(x),q(x)$ 都是不可约多项式. 证明：如果 $p(x)\mid q(x)$，那么 $p(x)=cq(x)$，式中 c 是一个非零常数.
2. 设 $p(x)$ 是 $\mathbf{P}[x]$ 中一个次数 $\geqslant 1$ 的多项式. 如果对于 $\mathbf{P}[x]$ 中任意多项式 $f(x)$，都有 $p(x)\mid f(x)$ 或 $(p(x),f(x))=1$，那么，$p(x)$ 是数域 $\mathbf{P}$ 上的不可约多项式.
3. 设 $p(x)$ 是 $\mathbf{P}[x]$ 中一个次数 $\geqslant 1$ 的多项式. 如果对于 $\mathbf{P}[x]$ 中任意两个多项式 $f(x)$ 与 $g(x)$，从 $p(x)\mid f(x)g(x)$ 可以推出 $p(x)\mid f(x)$ 或 $p(x)\mid$

$g(x)$，那么，$p(x)$是 $\mathbf{P}[x]$中的一个不可约多项式.

2. 因式分解定理

下面讨论多项式的因式分解定理.

设 $f(x)$是 $\mathbf{P}[x]$中的一个多项式. 如果 $f(x)$是一个不可约多项式，那就不可能再进行分解；如果不是，那么 $f(x)$可以分解为两个次数低于 $f(x)$的多项式的乘积：$f(x)=f_1(x)f_2(x)$. 如果 $f_1(x)$，$f_2(x)$都是不可约多项式，那么 $f(x)$就表示成了两个不可约多项式的乘积，这就分解完毕；如果不然，可再对 $f_1(x)$，$f_2(x)$进行分解. 将此步骤继续下去，最后总可以把 $f(x)$表示成一些不可约多项式的乘积. 下面来证明这个结论.

定理 11（因式分解定理） 数域 $\mathbf{P}$ 上每个次数$\geqslant 1$ 的多项式 $f(x)$，都可以分解成数域 $\mathbf{P}$ 上一些不可约多项式的乘积，并且分解是唯一的，即：如果有两个分解式

$$f(x) = p_1(x)p_2(x)\cdots p_s(x) = q_1(x)q_2(x)\cdots q_t(x),$$

其中 $p_i(x)(i=1,2,\cdots,s)$及 $q_j(x)(j=1,2,\cdots,t)$都是数域 $\mathbf{P}$ 上的不可约多项式. 那么必有 $s=t$，而且可以适当排列因式的次序，使得

$$p_i(x) = c_i q_i(x) \quad (i = 1,2,\cdots,s).$$

其中 $c_i(i=1,2,\cdots,s)$是非零常数.

证明：先证明分解式的存在. 对 $f(x)$的次数 n 作归纳法：

因为 1 次多项式都是不可约的，所以 $n=1$ 时，结论是成立的.

假设结论对于次数低于 n 的多项式成立.

设 $f(x)$是一个 n 次多项式. 如果 $f(x)$是不可约的，结论显然是成立的；如果 $f(x)$是可约的，那么 $f(x)$可以表示成

$$f(x) = f_1(x)f_2(x).$$

其中 $f_1(x)$，$f_2(x)$的次数都低于 n. 由归纳法假设知：$f_1(x)$，$f_2(x)$都可以分解成数域 $\mathbf{P}$ 上一些不可约多项式的乘积. 把 $f_1(x)$，

$f_2(x)$的分解式相乘，就得到 $f(x)$的一个分解式.

由归纳法原理，结论普遍成立.

下面来证唯一性：设 $f(x)$可用两种方法表示成不可约多项式的乘积：

$$f(x) = p_1(x)p_2(x)\cdots p_s(x) = q_1(x)q_2(x)\cdots q_t(x). \quad (1)$$

其中 $p_i(x)(i=1,2,\cdots,s)$及 $q_j(x)(j=1,2,\cdots,t)$都是不可约多项式. 我们仍对 n 用归纳法来证明分解的唯一性：

当 $n=1$ 时，$f(x)$是一个不可约多项式. 根据不可约多项式的定义，可得

$$s = t = 1 \text{ 以及 } f(x) = p_1(x) = q_1(x).$$

所以分解法是唯一的. 现在假设结论对于次数小于 n 的多项式成立.

如果 $f(x)$是一个不可约多项式，那么由不可约多项式的定义，即可得出分解的唯一性，如果 $f(x)$是可约的，那么(1)式中的 $s \geqslant 2$，并且从(1)式得到：$p_1(x)|q_1(x)q_2(x)\cdots q_t(x)$. 因为 $p_1(x)$是不可约的，所以它一定能整除 $q_1(x),q_2(x),\cdots,q_t(x)$中的一个. 不妨设(必要时可以重新排列 $q_j(x)(j=1,\cdots,t)$的次序)

$$p_1(x)|q_1(x).$$

又因 $q_1(x)$也是不可约的，所以有

$$p_1(x) = c_1q_1(x) \quad (c_1 \text{ 是一个非零常数}).$$

在(1)式中消去 $p_1(x)$，得

$$p_2(x)p_3(x)\cdots p_s(x) = c_1^{-1}q_2(x)q_3(x)\cdots q_t(x).$$

这个多项式的次数显然小于 n. 由归纳法假设知道：分解法是唯一的，从而有

$$s - 1 = t - 1, \text{即 } s = t.$$

适当排列因式的次数之后，可得

$$p_2(x) = c_2^1c_1^{-1}q_2(x) = c_2q_2(x),$$

$$p_i(x) = c_iq_i(x)(i = 3,\cdots,s).$$

这就证明了分解法是唯一的.定理证毕. ▎

在多项式 $f(x)$ 的分解式中,可以取每个不可约多项式的首项系数为 1,再把相同的不可约因式合并,于是 $f(x)$ 的分解式就成为

$$f(x)=ap_1^{n_1}(x)p_2^{n_2}(x)\cdots p_s^{n_s}(x). \tag{2}$$

其中 a 是 $f(x)$ 的首项系数;$p_1(x),p_2(x),\cdots,p_s(x)$ 是各不相同的、首项系数为 1 的不可约多项式;$n_1,n_2,\cdots,n_s$ 是正整数.这种分解式称为**标准分解式**.

多项式的标准分解式可用来计算最大公因式及讨论整除关系.

假设 $f(x)$ 的标准分解式如(2).那么根据因式分解的唯一性可知:$f(x)$ 的因式一定具有形式

$$bp_1^{m_1}(x)p_2^{m_2}(x)\cdots p_s^{m_s}(x).$$

其中 b 是任意的非零常数;$m_1,m_2,\cdots,m_s$ 是非负整数,满足

$$0\leqslant m_i\leqslant n_i \quad (i=1,\cdots,s).$$

如果知道了 $f(x)$ 与 $g(x)$ 的标准分解式,根据以上的分析可知:$f(x)$ 与 $g(x)$ 的最大公因式 $d(x)$ 就等于那些同时在 $f(x)$ 与 $g(x)$ 的标准分解式中出现的不可约多项式的方幂的乘积,所带的方幂等于它在 $f(x)$ 与 $g(x)$ 中所带的方幂的较小的一个.而 $g(x)|f(x)$ 的充分必要条件是 $g(x)$ 的不可约因式 $p(x)$ 都是 $f(x)$ 的因式,而且 $p(x)$ 在 $g(x)$ 中的幂指数小于或等于在 $f(x)$ 中的幂指数.

例 1 设 $f(x)=x^7-x^6-6x^5+2x^4+13x^3+3x^2-8x-4$,

$$g(x)=-2x^7-6x^6+20x^4+30x^3+18x^2+4x.$$

是两个实系数多项式,求 $f(x),g(x)$ 及 $(f(x),g(x))$ 的标准分解式.

解:先求 $f(x)$ 的分解式.由观察知:$f(1)=0$,所以 $f(x)$ 以 $x-1$ 为因式.用综合除法求商 $q(x)$:

$$
\begin{array}{rrrrrrrr|r}
1 & -1 & -6 & 2 & 13 & 3 & -8 & -4 & 1 \\
 & 1 & 0 & -6 & -4 & 9 & 12 & 4 & \\
\hline
1 & 0 & -6 & -4 & 9 & 12 & 4 & 0 &
\end{array}
$$

得　　$q(x) = x^6 - 6x^4 - 4x^3 + 9x^2 + 12x + 4$,

所以　　$f(x) = (x-1)q(x)$.

再将 $q(x)$ 进行分解:因为 $q(-1)=0$,所以 $x+1 \mid q(x)$. 仍用综合除法计算:

$$
\begin{array}{rrrrrrr|r}
1 & 0 & -6 & -4 & 9 & 12 & 4 & -1 \\
 & -1 & 1 & 5 & -1 & -8 & -4 & \\
\hline
1 & -1 & -5 & 1 & 8 & 4 & 0 & -1 \\
 & -1 & 2 & 3 & -4 & -4 & & \\
\hline
1 & -2 & -3 & 4 & 4 & 0 & & -1 \\
 & -1 & 3 & 0 & -4 & & & \\
\hline
1 & -3 & 0 & 4 & 0 & & & -1 \\
 & -1 & 4 & -4 & & & & \\
\hline
1 & -4 & 4 & 0 & & & &
\end{array}
$$

得　　$f(x) = (x-1)(x+1)^4(x^2 - 4x + 4)$

$$= (x-1)(x+1)^4(x-2)^2.$$

用同样的方法,可以求出 $g(x)$ 的标准分解式为

$$g(x) = -2x(x+1)^5(x-2).$$

由此可得　　$(f(x), g(x)) = (x+1)^4(x-2)$.

还可看出:因为 x 是 $g(x)$ 的因式,不是 $f(x)$ 的因式,所以 $g(x) \nmid f(x)$.

最后提一下,由于一个多项式是否可约同系数域有关,所以,一个多项式的标准分解式也是与系数域有关的.

习　题　6.4(2)

1. 求下列实系数多项式的标准分解式：

　(1) $f(x)=-x^7+3x^6-5x^5+7x^4-7x^3+5x^2-3x+1$；

　(2) $g(x)=x^6-15x^4+8x^3+51x^2-72x+27$；

　(3) $h(x)=2x^5-10x^4+16x^3-16x^2+14x-6$.

2. $f(x),g(x),h(x)$同上题. 求$(f(x),g(x))$，$(f(x),h(x))$，$(g(x),h(x))$及$(f(x),g(x),h(x))$.

3. 插值法

因式分解定理指出：一个 n 次多项式最多有 n 个不可约因式，因此，最多只能有 n 个不同的一次因式，即最多有 n 个不同的根. 根据这一点，可以证明多项式的下述重要性质.

定理 12　如果多项式 $f(x),g(x)$的次数都不超过 n，而它们对 $n+1$ 个不同的数 $c_1,c_2,\cdots,c_{n+1}$有相同的值，即

$$f(c_i)=g(c_i)(i=1,2,\cdots,n+1).$$

那么 $f(x)=g(x)$.

这个定理说明，一个次数不超过 n 的多项式，被它在 $n+1$ 个数上的值所唯一确定. 这就是多项式插值的理论根据.

在实际问题中，我们常常用 $y=F(x)$来表示某种规律的数量关系. 但是在许多问题中，函数 $F(x)$往往是通过实验或观测得到的，一般地只是给出了 $F(x)$在一些点 x_i 上的函数值 $y_i=F(x_i)$，$i=1,2,\cdots,n+1$. 有时函数 $F(x)$虽然有解析表达式，但是如果表式比较复杂，计算起来就比较麻烦，使用起来也不方便. 因此，我们希望根据给定点 x_i 上的函数值 $F(x_i)$，$i=1,2,\cdots,n+1$，求出一个既能反映 $F(x)$的特性，又便于计算的简单函数 $f(x)$，用 $f(x)$近似地代替 $F(x)$. $f(x)$称为 $F(x)$的**插值函数**；$x_1,\cdots,x_{n+1}$称为**插值节点**；求插值函数的方法称为**插值法**.

由于多项式是一类简单的初等函数，而且任意给了两组数 c_1，

$c_2,\cdots,c_{n+1};b_1,b_2,\cdots,b_{n+1}$，其中 $c_1,\cdots,c_{n+1}$ 各不相同，总有唯一的一个次数不超过 n 的多项式 $f(x)$，使 $f(c_i)=b_i,i=1,2,\cdots,n+1$. 因此在实际应用中常常取多项式作为插值函数，称为**插值多项式**. 下面通过具体例子介绍几种求插值多项式的方法.

例 2 求一个三次多项式 $f(x)$，使其满足 $f(0)=-1,f(1)=2,f(-1)=-2,f(2)=19$.

解：设 $f(x)=a_3x^3+a_2x^2+a_1x+a_0$. 根据假设得

$$\begin{aligned} a_0 &= -1, \\ a_3+a_2+a_1+a_0 &= 2, \\ -a_3+a_2-a_1+a_0 &= -2, \\ 8a_3+4a_2+2a_1+a_0 &= 19. \end{aligned}$$

解方程组，得

$$a_3=2,a_2=1,a_1=0,a_0=-1.$$

所以

$$f(x)=2x^3+x^2-1.$$

这个方法叫做**待定系数法**. 原理比较简单. 对一般的情形，如果已知次数$\leqslant n$ 的多项式 $f(x)$在 $n+1$ 个不同数 $c_1,c_2,\cdots,c_{n+1}$处的值为 $b_1,b_2,\cdots,b_{n+1}$. 为了求出 $f(x)$，可设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0.$$

由 $\qquad f(c_i)=b_i,i=1,2,\cdots,n+1$，得

$$\begin{gathered} c_1^na_n+c_1^{n-1}a_{n-1}+\cdots+c_1a_1+a_0=b_1, \\ c_2^na_n+c_2^{n-1}a_{n-1}+\cdots+c_2a_1+a_0=b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{n+1}^na_n+c_{n+1}^{n-1}a_{n-1}+\cdots+c_{n+1}a_1+a_0=b_{n+1}. \end{gathered}$$

这是关于 $a_n,a_{n+1},\cdots,a_1,a_0$ 的一个线性方程组，包含 $n+1$ 个线性方程. 它的系数行列式为

$$d=\begin{vmatrix} c_1^n & c_1^{n-1} & \cdots & c_1 & 1 \\ c_2^n & c_2^{n-1} & \cdots & c_2 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n+1}^n & c_{n+1}^{n-1} & \cdots & c_{n+1} & 1 \end{vmatrix}.$$

这是一个 $n+1$ 阶范德蒙行列式. 因为 $c_1, c_2, \cdots, c_{n+1}$ 各不相同, 所以 $d \neq 0$. 因此这个线性方程组有唯一解, 这个解就是 $f(x)$ 的系数.

当 $f(x)$ 的次数比较高, 即 n 比较大的时候, 应用这个方法解一个未知量较多的线性方程组, 计算量是比较大的. 但是这个方程组的系数比较有规律, 如果应用电子计算机来求解, 还是很方便的.

例 3 求一个次数不超过 3 的多项式 $f(x)$, 使 $f(0)=1$, $f(-1)=5$, $f(1)=-1$, $f(2)=5$.

解: 设

$f(x)=a_0+a_1x+a_2x(x+1)+a_3x(x+1)(x-1)$.

由 $f(0)=1$, 得 $a_0=1$.

再由 $f(-1)=5$, 得 $1-a_1=5$, 所以 $a_1=-4$.

再由 $f(1)=-1$, 得 $1-4+2a_2=-1$, 所以 $a_2=1$.

再由 $f(2)=5$, 得 $1-8+6+6a_3=5$, 所以 $a_3=1$.

因此

$$\begin{aligned} f(x) &= 1-4x+x(x+1)+x(x+1)(x-1) \\ &= x^3+x^2-4x+1. \end{aligned}$$

应用这种方法求 $f(x)$, 可以逐次求出 $a_0, a_1, \cdots, a_n$ 而不必解联立线性方程组. 一般方法是: 已知 n 次多项式 $f(x)$ 满足 $f(c_i)=b_i(i=1,2,\cdots,n+1)$, 则设

$$\begin{aligned} f(x) = a_0 &+ a_1(x-c_1)+a_2(x-c_1)(x-c_2)+\cdots \\ &+ a_n(x-c_1)\cdots(x-c_n). \end{aligned}$$

然后将 $x=c_i(i=1,2,\cdots,n+1)$ 逐次代入 $f(x)$ 求出 $a_i(i=0,1,\cdots,$

n). 最后，如果需要的话，将 $f(x)$ 展开并重新整理.

我们把这种解法称为逐次逼近法. 这种解法具有下述优点. 当我们求出插值多项式后，如果精度不满足要求，需要增加插值节点时，若用待定系数法来计算，则原来算出的数据就不能利用，必须重新计算；而用逐次逼近法来计算就可避免这个矛盾. 举例如下.

例 4. 求四次多项式 $g(x)$. $g(x)$ 除满足例 3 中的条件，即 $g(0)=1, g(-1)=5, g(1)=-1, g(2)=5$ 以外，还满足 $g(1/2)=1/2$.,

解: 设

$$
\begin{aligned}
g(x) = & a_0 + a_1 x + a_2 x(x+1) + a_3 x(x+1)(x-1) \\
& + a_4 x(x+1)(x-1)(x-2).
\end{aligned}
$$

那么，a_0, a_1, a_2, a_3 满足的条件仍与例 3 中一样，因此可以不必重新计算而知 $a_0=1, a_1=-4, a_2=1, a_3=1$. 我们只要算出 a_4 就可以了. 由 $g(1/2)=1/2$，得，

$$1 - 2 + \frac{3}{4} - \frac{3}{8} + \frac{9}{16}a_4 = \frac{1}{2}.$$

所以

$$a_4 = 2,$$

于是

$$
\begin{aligned}
g(x) & = 1 - 4x + x(x+1) + x(x+1)(x-1) \\
& \quad + 2x(x+1)(x-1)(x-2) \\
& = 2x^4 - 3x^3 - x^2 + 1.
\end{aligned}
$$

逐次逼近法还有一个优点是可以通过已经算出的数据进一步推出如何选择新的插值节点. 具体作法就不在这里介绍了.

最后，我们举例说明如何应用拉格朗日插值法来计算插值多项式. 还是先从具体例子做起.

例 5 求一个次数不超过 3 的多项式 $f(x)$，使 $f(1)=1, f(2)=0, f(3)=5, f(4)=22$.

解：设

$$\begin{aligned} f(x) = & k_1(x-2)(x-3)(x-4) \\ & + k_2(x-1)(x-3)(x-4) \\ & + k_3(x-1)(x-2)(x-4) \\ & + k_4(x-1)(x-2)(x-3). \end{aligned}$$

则由 $f(1)=1$，得 $-6k_1=1$，所以 $k_1=-1/6$.

由 $f(2)=0$，得 $2k_2=0$，所以 $k_2=0$.

由 $f(3)=5$，得 $-2k_3=5$，所以 $k_3=-5/2$.

由 $f(4)=22$，得 $6k_4=22$，所以 $k_4=11/3$.

从而得

$$\begin{aligned} f(x) &= -\frac{1}{6}(x-2)(x-3)(x-4) \\ &\quad -\frac{5}{2}(x-1)(x-2)(x-4) \\ &\quad +\frac{11}{3}(x-1)(x-2)(x-3) \\ &= x^3-3x^2+x+2. \end{aligned}$$

这种解法的一般方法是：如果已知次数 $<n$ 的多项式 $f(x)$ 在 $x=c_i(i=1,2,\cdots,n)$ 的值 $f(c_i)=b_i$，则可设

$$f(x)=\sum_{i=1}^{n}k_i(x-c_1)\cdots(x-c_{i-1})(x-c_{i+1})\cdots(x-c_n).$$

依次令 $x=c_i(i=1,2,\cdots,n)$ 代入 $f(x)$，得

$$\begin{aligned} f(c_i) &= k_i(c_i-c_1)\cdots(c_i-c_{i-1})(c_i-c_{i+1})\cdots(c_i-c_n) \\ &= b_i. \end{aligned}$$

所以

$$k_i=\frac{b_i}{(c_i-c_1)\cdots(c_i-c_{i-1})(c_i-c_{i+1})\cdots(c_i-c_n)},$$

$$f(x)=\sum_{i=1}^{n}\frac{b_i(x-c_1)\cdots(x-c_{i-1})(x-c_{i+1})\cdots(x-c_n)}{(c_i-c_1)\cdots(c_i-c_{i-1})(c_i-c_{i+1})\cdots(c_i-c_n)}.$$

为了把 $f(x)$ 的表达式写得简单一些，令

$$F(x) = (x - c_1)(x - c_2)\cdots(x - c_n).$$

于是

$$(x - c_1)\cdots(x - c_{i-1})(x - c_{i+1})\cdots(x - c_n) = \frac{F(x)}{x - c_i},$$

$$(c_i - c_1)\cdots(c_i - c_{i-1})(c_i - c_{i+1})\cdots(c_i - c_n) = F'(c_i),$$

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

$F'(c_i)$就是$F(x)$在$x=c_i$处的导数(微商),因此$f(x)$可表成

$$f(x) = \sum_{i=1}^{n} \frac{b_i F(x)}{(x - c_i)F'(c_i)}.$$

这个公式称为**拉格朗日插值公式**.利用它可以直接算出满足$f(c_i)=b_i(i=1,2,\cdots,n)$的次数小于$n$的多项式$f(x)$.

习 题 6.4(3)

1. 求一个次数<4的多项式$f(x)$,它适合条件:$f(-1)=-5,f(2)=-5,f(3)=-1,f(4)=15$.

2. 求一个二次多项式,使它在$x=0,\frac{\pi}{2},\pi$处与函数$\sin x$有相同的值.

3. 求一个次数尽可能低的多项式$f(x)$,使$f(0)=1,f(1)=3,f(2)=11,f(3)=61,f(-2)=-9$.

6.5 重 因 式

这一节继续讨论数域$\mathbf{P}$上多项式的因式分解问题.首先介绍k重因式的概念.

定义 8 不可约多项式$p(x)$称为多项式$f(x)$的k**重因式**,如果$p^k(x)\mid f(x)$,而$p^{k+1}(x)\nmid f(x)$.

需要强调的是:定义中的$p(x)$必须是不可约的.

如果$k=0$,那么$p(x)$根本不是$f(x)$的因式;如果$k=1$,那么$p(x)$称为$f(x)$的**单因式**;如果$k>1$,那么$p(x)$称为$f(x)$的**重因**

式. 如果一次因式 $x-c$ 是 $f(x)$ 的 $k(k>0)$ 重因式，就称 c 是 $f(x)$ 的 **k 重根**. 特别地，当 $k=1$ 时，称为**单根**；$k>1$ 时，称为**重根**.

例如，设 $f(x)=-5(x-1)^3(x+2)^2(x+3)(x-5)$，则 $x-1$ 是 $f(x)$ 的 3 重因式；$x+2$ 是 $f(x)$ 的 2 重因式；$x+3$ 与 $x-5$ 都是 $f(x)$ 的 1 重因式. 所以 $x-1$ 及 $x+2$ 是 $f(x)$ 的重因式；$x+3$ 及 $x-5$ 是 $f(x)$ 的单因式。相应地，1 是 $f(x)$ 的 3 重根；-2 是 $f(x)$ 的 2 重根；-3 及 5 都是 $f(x)$ 的 1 重根. 所以，1 及 -2 是 $f(x)$ 的重根；-3 及 5 是 $f(x)$ 的单根.

在许多问题中，要求所讨论的多项式是没有重因式的. 多项式的一些近似求根法也要求所考虑的多项式是没有重根的. 这一节介绍如何判断多项式有没有重因式，求多项式的重因式以及去掉重因式的方法.

当然，如果知道了 $f(x)$ 的标准分解式

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x).$$

那么也就知道了 $p_1(x),p_2(x),\cdots,p_s(x)$ 分别是 $f(x)$ 的 $r_1,r_2,\cdots,r_s$ 重因式. 那些指数 $r_i=1$ 的不可约因式 $p_i(x)$ 是 $f(x)$ 的单因式；指数 $r_i>1$ 的不可约因式 $p_i(x)$ 是 $f(x)$ 的重因式. 但是，由于没有一个一般的方法能将多项式分解成标准分解式，所以需要用另外的方法来解决多项式的重因式问题. 下面我们利用微积分中的导数（微商）的概念来解决这个问题. 为此，先把多项式的导数复习一下.

设多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0.$$

我们知道 $f(x)$ 的导数 $f'(x)$ 是比 $f(x)$ 低一次的多项式*

---

* 如果有的读者没有学过微积分. 那么可以定义 $f(x)$ 的导数为

$$f'(x)=na_nx^n+(n-1)a_{n-1}x^{n-1}+\cdots+a_1.$$

后面的 4 个公式可以直接从这个定义验证.

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.$$

多项式的导数满足下面几个公式：

$$[f(x) + g(x)]' = f'(x) + g'(x);$$

$$[cf(x)]' = cf'(x);$$

$$[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x);$$

$$[f^m(x)]' = mf^{m-1}(x)f'(x).$$

还要用到高阶导数的概念：$f'(x)$称为 $f(x)$的一阶导数；$f'(x)$的导数 $f''(x)$称为 $f(x)$的二阶导数；等等. 一般地，$f(x)$的 **k 阶导数**是 $f(x)$的 $k-1$ 阶导数的导数，记作 $f^{(k)}(x)$.

例 1 设 $f(x)=3x^5+4x^4-2x^2+3x-1$，求 $f(x)$的各阶导数.

解： $f'(x)=15x^4+16x^3-4x+3$，

$f''(x)=60x^3+48x^2-4$，

$f'''(x)=180x^2+96x$，

$f^{(4)}(x)=360x+96$，

$f^{(5)}(x)=360$，

$f^{(6)}(x)=0$.

一个 n 次多项式的导数是一个 $n-1$ 次多项式；它的 n 阶导数是一个常数；它的 $n+1$ 阶导数等于零. 特别地，$f'(x)=0$ 的充分必要条件是 $f(x)$是一个常数.

下面介绍利用导数来求多项式的重因式的方法.

定理 13 如果不可约多项式 $p(x)$是 $f(x)$的 k 重因式($k\geqslant 1$)，那么 $p(x)$是 $f'(x)$的 $k-1$ 重因式.

证明： 因为 $p(x)$是 $f(x)$的 k 重因式，所以可设

$$f(x) = p^k(x)g(x),$$

其中 $g(x)$不能被 $p(x)$整除，求 $f(x)$的导数

$$f'(x) = [p^k(x)]'g(x) + p^k(x)g'(x)$$
$$= p^{k-1}(x)[kp'(x)g(x) + p(x)g'(x)].$$

这说明 $p^{k-1}(x)\mid f'(x)$. 现在来观察 $kp'(x)g(x)+p(x)g'(x)$，显然 $p(x)$能整除第二项. 但是 $p(x)\nmid p'(x)$，$p(x)\nmid g(x)$因此 $p(x)\nmid p'(x)g(x)$，所以 $p(x)\nmid kp'(x)g(x)+p(x)g(x)$. 这就证明了 $p(x)$是 $f'(x)$的 $k-1$ 重因式. |

从这个定理可以得到下面一些推论.

推论 1 如果不可约因式 $p(x)$是 $f(x)$的 $k(k\geqslant1)$重因式，那么 $p(x)$分别是 $f'(x),f''(x),\cdots,f^{(k-1)}(x)$的 $k-1,k-2,\cdots,1$ 重因式，而不是 $f^{(k)}(x)$的因式. 特别地，$f(x)$的单因式不再是 $f'(x)$的因式.

推论 2 不可约多项式 $p(x)$是 $f(x)$的重因式的充分必要条件为：$p(x)$是 $f(x)$与 $f'(x)$的公因式.

证明：如果不可约多项式 $p(x)$是 $f(x)$的重因式，那么 $p(x)$一定是 $f'(x)$的因式，因而是 $f(x)$与 $f'(x)$的公因式；反之，如果 $p(x)$是 $f(x)$的单因式，那么 $p(x)$不再是 $f'(x)$的因式，因而不是 $f(x)$与 $f'(x)$的公因式. |

推论 3 多项式 $f(x)$没有重因式的充分必要条件是 $f(x)$与 $f'(x)$互素.

这个推论给出了判断一个多项式有无重因式的具体方法. 由于两个多项式是否互素，不因系数域的扩大而改变，所以一个多项式有无重因式，也不会因系数域的扩大而改变.

因为 $f(x)$与 $f'(x)$的最大公因式可以用辗转相除法求出，所以可以用辗转相除法来判断 $f(x)$有无重因式，并在有重因式的时候求出重因式来. 这个方法既然可以求出重因式，故在某些情况下，也可用来分解因式.

例 2 求 $f(x)=x^4+x^3-3x^2-5x-2$ 的重因式.

解：
$$f'(x)=4x^3+3x^2-6x-5,$$
用辗转相除法，求得
$$(f(x),f'(x))=x^2+2x+1=(x+1)^2.$$

所以 $x+1$ 是 $f(x)$的 3 重因式，由此得到 $f(x)$的标准分解式为

$$f(x)=(x+1)^3(x-2).$$

例 3 试证多项式 $f(x)=1+x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n$ 没有重根.

证明: $f'(x)=1+x+\frac{1}{2!}x^2+\cdots+\frac{1}{(n-1)!}x^{n-1}$.

由
$$f(x)-f'(x)=\frac{1}{n!}n^n$$
知
$$(f(x),f'(x))=(f(x),f(x)-f'(x))=1.$$
所以 $f(x)$没有重根.

例 4 求 a,b，使$(x+1)^2|ax^4+bx^2+1$.

解: 记 $f(x)=ax^4+bx^2+1$，则 $f'(x)=4ax^3+2bx$.

由定理 13 知，$(x+1)^2|f(x)$的充要条件为：$x+1|f(x)$，$x+1|f'(x)$，即

$$f(-1)=a+b+1=0, f'(-1)=-4a-2b.$$

解方程，得 $a=1,b=-2$.

最后，我们来讨论如何去掉重因式的问题. 由于在有些问题中，要求所讨论的多项式没有重因式，例如，在多项式的根的一些近似计算方法中，要求该多项式没有重根. 如果我们要计算 $f(x)$的根，但是 $f(x)$却有重根，那么怎么办呢？我们当然会想到找一个多项式 $g(x)$来代替它. $g(x)$必须满足两个条件：首先，$g(x)$的根必须与 $f(x)$的根相同；其次，$g(x)$必须没有重根. 与这两个条件等价的条件是：第一，$g(x)$与 $f(x)$有相同的不可约因式；第二，$g(x)$没有重因式. 问题是这样的 $g(x)$是否存在，又如何去找呢？假如知道了 $f(x)$的标准分解式：

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x).$$

那么满足这两个条件的多项式 $g(x)$必须等于

$$kp_1(x)p_2(x)\cdots p_s(x),$$

其中 k 是一个非零常数. 根据定理 13，

$$f'(x) = c' p_1^{n_1-1}(x) p_2^{n_2-1}(x) \cdots p_s^{n_s-1}(x) q(x),$$

其中 $q(x)$ 不能被 $p_i(x)(i=1,2,\cdots,s)$ 整除，因此

$$(f(x), f'(x)) = p_1^{n_1-1}(x) p_2^{n_2-1}(x) \cdots p_s^{n_s-1}(x).$$

(上式中的 $n_1-1, n_2-1, \cdots, n_s-1$ 可能有一些等于 0，因此上式不一定是 $(f(x), f'(x))$ 的标准分解式.) 由此得

$$\frac{f(x)}{(f(x), f'(x))} = c p_1(x) p_2(x) \cdots p_s(x).$$

这个多项式就可以取作 $g(x)$.

这样，我们就得出了去掉重因式的方法：给了一个非零多项式 $f(x)$，我们先求 $f'(x)$，再求 $f(x)$ 与 $f'(x)$ 的最大公因式 $(f(x), f'(x))$. 用 $(f(x), f'(x))$ 去除 $f(x)$，所得的商 $\frac{f(x)}{(f(x), g(x))}$ 是一个没有重因式的多项式，它的不可约因式与 $f(x)$ 的不可约因式完全相同.

例 5 设 $f(x)=x^5-6x^4+16x^3-24x^2+20x-8$，求一个多项式与 $f(x)$ 有相同的根，但无重根.

解： $f'(x)=5x^4-24x^3+48x^2-48x+20$，

用辗转相除法，求得

$$(f(x), f'(x)) = x^2 - 2x + 2.$$

多项式

$$\frac{f(x)}{(f(x), f'(x))} = x^3 - 4x^2 + 6x - 4$$

即为所求.

习 题 6.5

1. 判断下述多项式有无重因式：

(1) $x^4+x^3+2x^2+x+1$；

(2) $x^5-5x^4+8x^3-5x^2+4x-4$；

(3) x^4+4x^2-4x-3.

2. 求 t，使 x^3-3x^2+tx-1 有重根，并求出相应的重根及重数.

3. 求多项式 x^3+px+q 有重根的条件.

4. 试证 c 是 $f(x)$的 k 重根的充分必要条件是：

$$f(c)=f'(c)=f''(c)=\cdots=f^{(k-1)}(c)=0,\text{而 } f^{(k)}(c)\neq 0.$$

5. 试证 $x^n+ax^{n-m}+b(n>2,m<n)$不能有非零的重数大于 2 的根.

6.6 复系数与实系数多项式的因式分解

前面讨论了任意数域 P 上多项式的因式分解问题. 这一节讨论在复数域与实数域上多项式的因式分解. 复数域与实数域既然都是数域，所以前面所得的结论对它们也都成立. 由因式分解定理可得：任一复(实)系数多项式都可以唯一地分解成不可约复(实)系数多项式的乘积. 这一节要进一步讨论怎样的复系数或实系数多项式是不可约的.

首先讨论复数域的情形. 关于复系数多项式，有下述重要定理.

代数基本定理 每个次数≥1 的复系数多项式，在复数域中有一个根.

这个定理是高斯(Gauss)1799 年第一个证明的，尔后他又给出了 4 个证明. 由于解方程的问题是当时代数学的中心问题，所以这个定理获得了代数基本定理这个名称. 至今，这个定理的证明已经非常之多. 遗憾的是，这个名为代数基本定理的定理的证明都不可避免地要用到复数域的解析性质. 这个定理的证明在本书中不讲，将来读者学了复变函数就可以很容易地给予证明了.

根据根与一次因式的关系(定理 3 的推论)，从代数基本定理可得出.

每个次数≥1 的复系数多项式，在复数域上一定有一个一次因式.

由此可知：在复数域上，所有次数>1 的多项式都是可约的. 换句话说，复系数不可约多项式只有一次多项式，从而因式分解定理在复数域上可以叙述成：

复系数多项式因式分解定理 每个次数≥1 的复系数多项式都可以唯一地分解成一次因式的乘积.

因此，复系数多项式的标准分解式为

$$f(x) = a(x - a_1)^{l_1}(x - a_2)^{l_2}\cdots(x - a_s)^{l_s}.$$

其中 $a_1, a_2, \cdots, a_s$ 是不同的复数；$l_1, l_2, \cdots, l_s$ 是正整数；$l_1 + l_2 + \cdots + l_s$ 等于 $f(x)$ 的次数. 标准分解式说明了每个 n 次复系数多项式恰有 n 个复根(重根按重数计算).

例 1 求复系数多项式

$$\begin{aligned}f(x) = &2\mathrm{i}x^8 + (2 + 3\mathrm{i})x^7 + (3 + 4\mathrm{i})x^6 + (4 + 5\mathrm{i})x^5 \\ &+ (5 + 2\mathrm{i})x^4 + (2 + \mathrm{i})x^3 + x^2 - \mathrm{i}x - 1\end{aligned}$$

的标准分解式.

解：

$$\begin{aligned}f(x) &= 2\mathrm{i}\Big[x^8 + \left(\frac{3}{2} - \mathrm{i}\right)x^7 + \left(2 - \frac{3}{2}\mathrm{i}\right)x^6 + \left(\frac{5}{2} - 2\mathrm{i}\right)x^5 \\ &\quad + \left(1 - \frac{5}{2}\mathrm{i}\right)x^4 + \left(\frac{1}{2} - \mathrm{i}\right)x^3 - \frac{\mathrm{i}}{2}x^2 - \frac{1}{2}x + \frac{1}{2}\mathrm{i}\Big] \\ &= 2\mathrm{i}(x - \mathrm{i})\left(x - \frac{1}{2}\right)(x + 1)^2(x^2 + 1)^2 \\ &= 2\mathrm{i}\left(x - \frac{1}{2}\right)(x + 1)^2(x + \mathrm{i})^2(x - \mathrm{i})^3.\end{aligned}$$

例 2 求复系数多项式

$$f(x) = x^8 - 1$$

的标准分解式.

解：

$$f(x) = (x^4 - 1)(x^4 + 1) = (x^2 - 1)(x^2 + 1)(x^2 - \mathrm{i})(x^2 + \mathrm{i})$$

$$=(x+1)(x-1)(x+\mathrm{i})(x-\mathrm{i})$$
$$\times\left(x+\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}\mathrm{i}\right)\left(x+\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}\mathrm{i}\right)$$
$$\times\left(x-\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}\mathrm{i}\right)\left(x-\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}\mathrm{i}\right).$$

例 3 求复系数多项式

$$f(x)=x^n-1$$

的标准分解式.

注意,如果将 $f(x)=x^n-1$ 直接进行分解,第一步容易看出:$x-1$ 是 $f(x)$的因式,于是得

$$f(x)=(x-1)(x^{n-1}+x^{n-2}+\cdots+x+1).$$

再进一步将 $x^{n-1}+x^{n-2}+\cdots+x+1$ 进行分解就比较困难了.根据根与一次因式的关系可知:找出 $f(x)$的所有一次因式,就相当于找出 $f(x)$的所有根.因而,为了将 x^n-1 表成一次因式的乘积,只要求出 $x^n-1=0$ 的所有根就行了.

解:首先来求 $f(x)=x^n-1$ 的所有根.为此,将 1 表示成三角表示式

$$1=\cos 2k\pi+\mathrm{i}\sin 2k\pi.$$

因为 $x^n-1=0$ 的解满足 $x^n=1$.因而可知 x^n-1 的全部根为

$$\cos\frac{2k\pi}{n}+\mathrm{i}\sin\frac{2k\pi}{n}\ (k=0,1,\cdots,n-1).$$

于是得到

$$f(x)=\prod_{k=0}^{n-1}\left[x-\left(\cos\frac{2k\pi}{n}+\mathrm{i}\sin\frac{2k\pi}{n}\right)\right].$$

习　题　6.6(1)

求下列多项式在复数域上的分解式.

1. $x^7-3x^6+5x^5-7x^4+7x^3-5x^2+3x-1$.
2. $2x^5-20x^3-40x^2-30x-8$.

3. $2ix^6+x^5-6ix^4+(-3+2i)x^3+(1+4i)x^2+(2-2i)x-1$.

4. $x^{12}-1$.

下面讨论实系数多项式的因式分解问题. 首先证明实系数多项式的根的一个重要性质:

定理 14 设 $f(x)$是一个实系数多项式. 如果复数 α 是 $f(x)$的一个根,那么 α 的共轭数 $\bar{\alpha}$ 也是 $f(x)$的一个根.

证明: 设

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0.$$

式中系数 $a_n,a_{n-1},\cdots,a_1,a_0$ 都是实数. 因为 α 是 $f(x)$的一个根,所以

$$f(\alpha) = a_n\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0 = 0.$$

由于 $a_i(i=n,n-1,\cdots,1,0)$都是实数,所以

$$\begin{aligned} f(\bar{\alpha}) &= a_n\bar{\alpha}^n + a_{n-1}\bar{\alpha}^{n-1} + \cdots + a_1\bar{\alpha} + a_0 \\ &= \overline{a_n\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0} \\ &= \bar{0} = 0. \end{aligned}$$

即 $\bar{\alpha}$ 也是 $f(x)$的根. |

如果 α 不是实数,那么 $\bar{\alpha}\neq\alpha$,而

$$(x-\alpha)(x-\bar{\alpha}) = x^2-(\alpha+\bar{\alpha})x+\alpha\bar{\alpha}$$

是一个实系数二次多项式,而且在实数域上是不可约的.

由此可以证明下述关于实系数多项式的因式分解定理.

实系数多项式因式分解定理 每个次数$\geqslant 1$的实系数多项式在实数域上都可以唯一地分解成一次与二次不可约因式的乘积.

证明: 根据因式分解定理,只要证明实系数不可约多项式都是一次或二次的.

设 $f(x)$是一个实系数不可约多项式. 把 $f(x)$看成一个复系数多项式,由代数基本定理知:$f(x)$有一个复根 α. 如果这个 α 是一个实数,那么 $f(x)$可以被实系数多项式 $x-\alpha$ 整除,由于 $f(x)$

是不可约的，故得

$$f(x) = c(x - \alpha),$$

即 $f(x)$是一次的.

如果 α 是一个虚数，那么 $\bar{\alpha}$ 也是 $f(x)$的一个根，而且 $\alpha \neq \bar{\alpha}$. 于是$(x-\alpha)(x-\bar{\alpha})$是一个实系数不可约二次多项式，而且

$$(x - \alpha)(x - \bar{\alpha}) \mid f(x),$$

再根据 $f(x)$是不可约多项式，得

$$f(x) = c(x - \alpha)(x - \bar{\alpha}) = c[x^2 - (\alpha + \bar{\alpha})x + \alpha\bar{\alpha}].$$

所以 $f(x)$是二次的. ▌

从这个定理可知：实系数多项式的标准分解式为

$$f(x) = a(x - c_1)^{l_1} \cdots (x - c_s)^{l_s}(x^2 + p_1 x + q_1)^{k_1} \cdots (x^2 + p_t x + q_t)^{k_t}.$$

其中 $c_1, \cdots, c_s$ 是相异实数；$p_1, q_1, \cdots, p_t, q_t$ 是相异实数对，而且因为 $x^2 + p_i x + q_i$ 都是不可约的，所以 p_i, q_i 满足 $p_i^2 < 4q_i (i = 1, 2, \cdots, t)$. $l_1, \cdots, l_s$ 及 $k_1, \cdots, k_t$ 都是正整数，$l_1 + \cdots + l_s + 2k_1 + \cdots + 2k_t$ 等于 $f(x)$的次数 n.

从实系数多项式的标准分解式还可以看出：如果虚数 α 是实系数多项式 $f(x)$的一个根，那么不但 $\bar{\alpha}$ 也是 $f(x)$的根，而且它们的重数也是相等的.

例 4 将 $f(x) = x^n - 1$ 分解成实系数不可约多项式的乘积.

解：已知 $x^n - 1$ 的 n 个复根为

$$x_k = \cos\frac{2k}{n}\pi + \mathrm{i}\sin\frac{2k}{n}\pi \quad (k = 0, 1, \cdots, n - 1).$$

为了将 $f(x)$分解为实系数不可约多项式的乘积，必须将这 n 个根区别为实根还是虚根，而且对虚根要找出其共轭根.

容易看出，x_k 是实数的充要条件是 $\sin\frac{2k}{n}\pi = 0$. 因此，当 n 是奇数时，只有一个实根 $x_0 = 1$；当 n 是偶数时，有两个实根 $x_0 = 1$ 及

$x_{\frac{n}{2}}=-1$.

因为 $x_k=\cos\frac{2k}{n}\pi+\mathrm{i}\sin\frac{2k}{n}\pi\left(k=0,1,\cdots,\left[\frac{n-1}{2}\right]^*\right)$的共轭数为

$$x_{n-k}=\cos\frac{2(n-k)}{n}\pi+\mathrm{i}\sin\frac{2(n-k)}{n}\pi$$
$$=\cos\frac{2k}{n}\pi-\mathrm{i}\sin\frac{2k}{n}\pi,$$

而且

$$\left[x-\left(\cos\frac{2k}{n}\pi+\mathrm{i}\sin\frac{2k}{n}\pi\right)\right]$$
$$\cdot\left[x-\left(\cos\frac{2k}{n}\pi-\mathrm{i}\sin\frac{2k}{n}\pi\right)\right]$$
$$=x^2-\left(2\cos\frac{2k}{n}\pi\right)x+1.$$

所以 x^n-1 在实数域上的分解式为：

当 n 是奇数时，

$$x^n-1=(x-1)\prod_{k=1}^{\frac{n-1}{2}}\left[x^2-\left(2\cos\frac{2k}{n}\pi\right)x+1\right];$$

当 n 是偶数时，

$$x^n-1=(x-1)(x+1)\prod_{k=1}^{\frac{n}{2}-1}\left[x^2-\left(2\cos\frac{2k}{n}\pi\right)x+1\right].$$

习　题　6.6(2)

1. 求习题 6.6(1)中第 1,2,4 题的多项式在实数域上的标准分解式.
2. 证明奇数次实系数多项式一定有实根.
3. 求一个次数尽可能低的实系数多项式，以 1,0,i,i,1−i 为根.

---

* 设 a 是一个实数. 用$[a]$表示不大于 a 的最大的那个整数，称为 a 的整数部分. 例如$[3]=3$；$\left[\frac{5}{2}\right]=2$；$[-3.15]=-4$.

6.7 有理系数多项式

上一节讨论了实系数与复系数多项式的因式分解.这一节讨论有理系数多项式的因式分解.首先,根据因式分解定理,每个次数$\geqslant 1$的有理系数多项式都能唯一地分解成不可约的有理系数多项式的乘积.但是,与实系数多项式及复系数多项式的情况不同,任意给了一个有理系数多项式,很难判断它是不是可约的.而且,对于任何正整数n,都有n次不可约有理系数多项式,这一节首先指出有理系数多项式的因式分解,可以归结为整系数的因式分解,进而解决求有理系数多项式的有理根的问题.最后,举例说明n次不可约有理系数多项式的存在性.

设 $$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

是一个有理系数多项式.将$f(x)$乘上一个适当的整数c,使得$cf(x)$是一个整系数多项式.如果$cf(x)$的各项系数的最大公因数是d,那么把d提出来,得到

$$cf(x)=dg(x),$$

即

$$f(x)=\frac{d}{c}\cdot g(x).$$

这里$g(x)$是一个整系数多项式,而且它的各项系数除了±1以外没有其他公因数.

例如,如果

$$f(x)=\frac{5}{3}x^4-10x^2-\frac{5}{4}x+5,$$

那么将$f(x)$乘上12,得

$$12f(x)=20x^4-120x^2-15x+60.$$

系数有公因数5,提出5,得

$$12f(x)=5(4x^4-24x^2-3x+12),$$

于是 $$f(x)=\frac{5}{12}(4x^4-24x^2-3x+12).$$

其中 $4x^4-24x^2-3x+12$ 是一个整系数多项式，它的各项系数是互素的.

定义 9 设 $g(x)=b_nx^n+b_{n-1}x^{n-1}+\cdots+b_1x+b_0$ 是一个非零的整系数多项式. 如果它的各项系数 $b_n,b_{n-1},\cdots,b_1,b_0$ 除去 ±1 外没有其他的公因数，那么 $g(x)$ 就称为一个**本原多项式**.

例如，上面的 $4x^4-24x^2-3x+12$ 就是一个本原多项式.

前面的分析说明：每个有理系数多项式都可以表成一个有理数与一个本原多项式的乘积，而且根据本原多项式的定义可知：这样的表示法除了差一个正负号外是唯一的，即如果

$$f(x)=r_1g_1(x)=r_2g_2(x),$$

其中 r_1,r_2 都是有理数，$g_1(x)$ 与 $g_2(x)$ 都是本原多项式，那么一定有

$$r_1=\pm r_2, g_1(x)=\pm g_2(x).$$

因为 $f(x)$ 与本原多项式 $g(x)$ 只差一个常数因子，所以 $f(x)$ 的因式分解可以归结为本原多项式 $g(x)$ 的因式分解. 重要的是，下面要证明：如果一个本原多项式可以分解为两个次数较低的有理系数多项式的乘积，那么它一定可以分解成两个次数较低的整系数多项式的乘积. 于是有理系数多项式的因式分解就可归结为整系数多项式的因式分解.

下面先证明关于本原多项式的一个重要性质.

定理 15 两个本原多项式的乘积还是本原多项式.

证明：设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$
$$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0$$

是两个本原多项式. 现在证明它们的乘积也是一个本原多项式. 设

$$f(x)g(x)=h(x)=c_{m+n}x^{m+n}+c_{m+n-1}x^{m+n-1}+\cdots+c_1x+c_0.$$

$h(x)$当然是一个整系数多项式. 用反证法来证明结论. 如果$h(x)$不是一个本原多项式，那么它的各项系数 $c_{m+n}, c_{m+n-1}, \cdots, c_1, c_0$ 就有一个异于± 1的公因数，因此，$c_{m+n}, c_{m+n-1}, \cdots, c_1, c_0$ 就能被一个素数(或称质数)p除尽.

因为$f(x)$是本原的，所以p不能除尽$f(x)$的全部系数，设a_l $(0 \leqslant l \leqslant n)$是第一个不能被$p$除尽的系数，即

$$p \mid a_n, p \mid a_{n-1}, \cdots, p \mid a_{l+1}, p \nmid a_l.$$

同样地，因为$g(x)$是本原的，所以可设$b_k(0 \leqslant k \leqslant m)$是$g(x)$的第一个不能被$p$除尽的系数，即

$$p \mid b_m, p \mid b_{m-1}, \cdots, p \mid b_{k+1}, p \nmid b_k.$$

现在来看$h(x)$中x^{l+k}的系数c_{l+k}:

$$\begin{aligned} c_{l+k} = & a_{l+k}b_0 + a_{l+k-1}b_1 + \cdots + a_{l-1}b_{k-1} + a_l b_k \\ & + a_{l-1}b_{k+1} + \cdots + a_1 b_{l+k-1} + a_0 b_{l+k}. \end{aligned}$$

上式右边的各项除了$a_l b_k$以外，其他各项都可以被p除尽，但因$p \nmid a_l, p \nmid b_k$，而且p是素数，所以$p \nmid a_l b_k$，因此$p \nmid c_{l+k}$. 这与p的取法相矛盾，这说明$h(x)$只能是本原多项式. ∎

我们利用这个定理来证明下述重要结果:

定理 16 如果非零整系数多项式能够分解成两个次数较低的有理系数多项式的乘积，那么它一定能分解成两个次数较低的整系数多项式的乘积.

证明: 设整系数多项式$f(x)$可以分解成

$$f(x) = g(x)h(x).$$

其中$g(x), h(x)$都是有理系数多项式，而且

$$\deg g(x) < \deg f(x), \deg h(x) < \deg f(x).$$

再设 $f(x) = af_1(x), g(x) = bg_1(x), h(x) = ch_1(x).$

其中$f_1(x), g_1(x), h_1(x)$都是本原多项式; a是整数; b, c是有理数. 于是

$$af_1(x) = bcg_1(x)h_1(x),$$

$$\frac{a}{bc}f_1(x) = g_1(x)h_1(x).$$

由定理 15，$g_1(x)h_1(x)$是一个本原多项式. 所以

$$\frac{a}{bc} = \pm 1,$$

即
$$bc = \pm a.$$

bc 是一个整数. 将 $f(x)$表成

$$f(x) = [bcg_1(x)] \cdot h_1(x).$$

那么 $bcg_1(x)$及 $h_1(x)$都是整系数多项式，而且它们的次数都低于 $f(x)$. 定理证毕. |

从定理的证明可以得到下述推论：

推论 如果整系数多项式 $f(x)$分解成

$$f(x) = g(x)h(x).$$

其中 $g(x)$是本原多项式，$h(x)$是有理系数多项式，那么 $h(x)$一定是整系数多项式.

证明：由于 $g(x)$是本原的，所以在定理 16 的证明中，$b=1$，从而 $c=\pm a$ 是一个整数，因此 $h(x)=ch_1(x)$是一个整系数多项式. |

应用这个推论，可以得出一个求有理系数多项式的有理根的方法. 这个方法的根据是下述定理.

定理 17 设

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

是一个整系数多项式，如果它有一个有理根$\frac{r}{s}$，其中 r,s 是互素的整数，那么必有 $s|a_n,r|a_0$. 特别地，如果 $f(x)$的首项系数 $a_n=1$，那么 $f(x)$的有理根都是整根，而且是 a_0 的因数.

证明：把 $f(x)$看成有理系数多项式. 因为$\frac{r}{s}$是 $f(x)$的有理根，所以根据根与一次因式的关系，在有理数域上，

$$\left(x-\frac{r}{s}\right)\bigg| f(x),$$

即
$$(sx-r)\,|\,f(x).$$

设 $f(x)=(sx-r)(b_{n-1}x^{n-1}+b_{n-2}x^{n-2}+\cdots+b_1x+b_0)$.

其中 $b_{n-1},\cdots,b_1,b_0$ 都是有理数. 因为 r,s 互素,所以 $rx+s$ 是一个本原多项式,由上述推论可知:$b_{n-1},\cdots,b_1,b_0$ 都是整数. 比较 x^n 的系数及常数项,得

$$a_n=sb_{n-1},a_0=-rb_0.$$

所以
$$s\,|\,a_n,\quad r\,|\,a_0.$$

如果 $a_n=1$,那么显然有 $s=\pm1$,$\frac{r}{s}=\pm r$ 是一个整数,而且是 a_0 的因数. |

例 1 解方程

$$2x^4-2x^3-9x^2-8x-3=0.$$

解: 先求此方程的有理根. 根据定理 17,这个方程的有理根只可能是 $\pm1,\pm3,\pm\frac{1}{2},\pm\frac{3}{2}$. 用综合除法检验,知 -1 和 3 是这个方程的根,而且

$$2x^4-2x^3-9x^2-8x-3=(x+1)(x-3)(2x^2+2x+1).$$

所以原方程的根为

$$-1,3,\frac{-1+\mathrm{i}}{2},\quad\frac{-1-\mathrm{i}}{2}.$$

例 2 求 $x^4+\frac{3}{2}x^3-3x^2+4x-\frac{3}{2}$ 的有理根.

解: 因为

$$x^4+\frac{3}{2}x^3-3x^2+4x-\frac{3}{2}$$

$$=\frac{1}{2}(2x^4+3x^3-6x^2+8x-3).$$

所以 $x^4+\frac{3}{2}x^3-3x^2+4x-\frac{3}{2}$ 与 $f(x)=2x^4+3x^3-6x^2+8x-3$ 有相同的根. 根据定理 17, $f(x)$ 的有理根只可能是 ±1, ±3, $\pm1/2$, $\pm3/2$. 用综合除法检验知只有 -3 和 $1/2$ 是它的根. 这也就是

$$x^4+\frac{3}{2}x^3-3x^2+4x-\frac{3}{2}$$

的全部有理根.

例 3 证明

$$f(x)=3x^3+2x-2$$

在有理数域上不可约.

证明: 如果 $f(x)$ 可约, 那么 $f(x)$ 可以分解成两个次数较低的有理系数多项式的乘积. 因为 $f(x)$ 的次数等于 3, 所以 $f(x)$ 必有一个一次有理因式, 因此有一个有理根. $f(x)$ 的有理根只可能是 ±1, ±2, $\pm1/3$, $\pm2/3$. 经检验, 这些数都不是 $f(x)$ 的根, 所以 $f(x)$ 没有一次因式, 因而是不可约的.

最后, 我们给出一个整系数多项式为不可约多项式的一个充分条件.

定理 18 (艾森斯坦(Eisenstein)判别法) 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

是一个整系数多项式. 如果有一个素数 p, 使得

(1) $p\nmid a_n$;

(2) $p\mid a_{n-1}, a_{n-2}, \cdots, a_1, a_0$;

(3) $p^2\nmid a_0$.

那么 $f(x)$ 在有理数域上是不可约的.

证明: 如果 $f(x)$ 在有理数域上是可约的. 那么根据定理 16, $f(x)$ 可表成两个次数较低的整系数多项式的乘积:

$$f(x)=(b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0)(c_lx^l+c_{l-1}x^{l-1}+\cdots+c_1x+c_0)\quad(l,m<n;\ l+m=n).$$

于是

$$a_n = b_m c_l, \quad a_0 = b_0 c_0.$$

因为 $p|a_0$，所以 p 能整除 b_0 或 c_0. 又因 $p^2 \nmid a_0$，所以 p 不能同时整除 b_0 及 c_0. 因此 p 恰能整除 b_0, c_0 中的一个. 不妨设 $p|b_0$，$p\nmid c_0$. 又因为 $p\nmid a_n$，所以 $p\nmid b_m$. 假设 $b_0, b_1, \cdots, b_m$ 中第一个不能被 p 整除的是 $b_s(0<s\leqslant m)$，即

$$p|b_0, p|b_1, \cdots, p|b_{s-1}, p\nmid b_s.$$

比较 $f(x)$ 中 x^s 的系数，得

$$a_s = b_s c_0 + b_{s-1} c_1 + \cdots + b_1 c_{s-1} + b_0 c_s$$

(如果 $i>l$，则令 $c_i=0$). 上式中 $a_s, b_{s-1}, \cdots, b_1, b_0$ 都可被 p 整除，因此 $b_s c_0$ 也可被 p 整除. 但是 b_s 与 c_0 却都不能被 p 整除. 这是一个矛盾. 所以 $f(x)$ 在有理数域上是不可约的.

例 4 对于任意正整数 n，多项式 x^n+2 在有理数域上是不可约的.

这个例子说明，在有理数域上存在任意次数的不可约多项式，而且应用艾森斯坦判别法可以作出许多不可约的有理系数多项式.

习 题 6.7

1. 把下列多项式表示成一个有理数与一个本原多项式的乘积.

(1) $x^3+\frac{1}{2}x^2-\frac{5}{3}x+3$；

(2) $3x^3+6x-12$；

(3) $5x^3+\frac{5}{2}x^2-\frac{5}{3}x+\frac{10}{3}$.

2. 设 $f(x)$ 是一个 n 次本原多项式，试证：对于任意整数 a，多项式 $ax^{n+1}+f(x)$ 及 $xf(x)+a$ 都是本原多项式.

3. 求下列多项式的有理根：

(1) $x^5-x^3+x^2-2$；

(2) $6x^5+3x^4-4x^3-2x^2+2x+1$；

(3) $6x^4-x^3+5x^2-x-1$；

(4) $x^4-\frac{40}{3}x^3+\frac{130}{3}x^2-40x+9$.

4. 判断下列多项式在有理数域上是否可约：

(1) $x^3-6x^2+16x-14$;

(2) $x^4-8x^3+11x^2+4$;

(3) $x^4+4x^3+6x^2+4x+8$.

5. 求下列多项式的标准分解式(分实数域与复数域两种情形)：

(1) $2x^5+3x^4-2x^2-2x-1$;

(2) $2x^5+5x^4-2x^3+2x^2-4x-3$.

6. 设 $f(x)$是一个首项系数为 1 的整系数多项式，证明：如果 $f(x)$在有理数域上可约，则 $f(x)$一定能表示成两个次数较低的，首项系数为 1 的整系数多项式的乘积.

内容提要

1. 一元多项式的运算

设
$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$
$$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0\quad(m\leqslant n).$$

则
$$f(x)+g(x)=(a_n+b_n)x^n+(a_{n-1}+b_{n-1})x^{n-1}+\cdots+(a_1+b_1)x+(a_0+b_0),$$
$$f(x)-g(x)=(a_n-b_n)x^n+(a_{n-1}-b_{n-1})x^{n-1}+\cdots+(a_1-b_1)x+(a_0-b_0)$$

(如果 $m<n$，则其中 $b_{m+1}=\cdots=b_n=0$).

$$f(x)g(x)=c_{n+m}x^{n+m}+c_{n+m-1}x^{n+m-1}+\cdots+c_1x+c_0,$$

其中
$$c_k=\sum_{i=0}^{k}a_{k-i}b_i.$$

多项式的运算满足

(1) 加法交换律：
$$f(x)+g(x)=g(x)+f(x).$$

(2) 加法结合律：

$$[f(x)+g(x)]+h(x)=f(x)+[g(x)+h(x)].$$

(3) 乘法交换律：

$$f(x)g(x)=g(x)f(x).$$

(4) 乘法结合律：

$$[f(x)g(x)]h(x)=f(x)[g(x)h(x)].$$

(5) 加乘分配律：

$$f(x)[g(x)+h(x)]=f(x)g(x)+f(x)h(x).$$

(6) 乘法消去律：

如果 $f(x)g(x)=f(x)h(x)$，并且 $f(x)\neq 0$. 那么必有 $g(x)=h(x)$.

2. 整除性理论

(1) 带余除法

带余除法：任给 $f(x)$，$g(x)$，且 $g(x)\neq 0$. 必有 $q(x)$，$r(x)$存在，满足

$$f(x)=q(x)g(x)+r(x),$$

其中，或者 $r(x)=0$，或者 $\deg r(x)<\deg g(x)$，而且满足这个条件的 $q(x)$及 $r(x)$只有一对.

如果 $f(x)=q(x)g(x)$，则称 $g(x)$整除 $f(x)$，记作 $g(x)\mid f(x)$.

余数定理：用 $x-c$ 除 $f(x)$所得的余数等于 $f(c)$.

推论：$x-c\mid f(x)$的充要条件是 $f(c)=0$.

(2) 关于整除的一些结论

1) $f(x)\mid g(x)$，$g(x)\mid f(x)\Rightarrow f(x)=cg(x)$（$c$ 是一个非零常数）.

2) $f_1(x)\mid f_2(x)$，$f_2(x)\mid f_3(x)$，$\cdots$，$f_{k-1}(x)\mid f_k(x)\Rightarrow f_1(x)\mid f_k(x)$ $(k\geqslant 2)$.

3) $g(x)\mid f_1(x)$，$g(x)\mid f_2(x)$，$\cdots$，$g(x)\mid f_s(x)$

$$\Rightarrow g(x) \mid u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots$$
$$+u_k(x)f_k(x)\ (k\geqslant 2)$$

(式中 $u_1(x),u_2(x),\cdots,u_k(x)$ 可以是任意的多项式).

3. 最大公因式

(1) 最大公因式的定义.

(2) $f(x),g(x)$ 的最大公因式可以表示成 $f(x)$ 与 $g(x)$ 的组合:

$$(f(x),g(x)) = u(x)f(x) + v(x)g(x).$$

$(f(x),g(x))$ 与 $u(x),v(x)$ 都可以应用辗转相除法求得.

(3) 互素的概念.

(4) 互素的多项式有下述一些重要性质:

1) $f(x)$ 与 $g(x)$ 互素的充分必要条件是存在 $u(x),v(x)$,使

$$u(x)f(x) + v(x)g(x) = 1;$$

2) $(f(x),g(x))=1, f(x) \mid g(x)h(x) \Rightarrow f(x) \mid h(x)$;

3) $g_1(x) \mid f(x), g_2(x) \mid f(x), (g_1(x),g_2(x))=1$
$\Rightarrow g_1(x)g_2(x) \mid f(x)$;

4) $(f(x),g(x))=1, (f(x),h(x))=1$
$\Rightarrow (f(x),g(x)h(x))=1$.

4. 因式分解定理

(1) 不可约多项式的概念.

(2) 不可约多项式的两个重要性质:

1) $p(x)$ 不可约,$f(x)$ 是任一多项式 $\Rightarrow (p(x),f(x))=1$ 或 $p(x) \mid f(x)$;

2) $p(x)$ 不可约,$p(x) \mid f_1(x)f_2(x)\cdots f_s(x) \Rightarrow p(x)$ 能整除 $f_1(x),f_2(x),\cdots,f_s(x)$ 中的一个.

(3) 因式分解定理:

$\mathbf{P}[x]$ 中每一个次数 $\geqslant 1$ 的多项式都可以唯一地表示成 $\mathbf{P}$ 上不可约多项式的乘积.

(4) 插值法.

5. 重因式

(1) 重因式的定义.

(2) 重因式的判断：

$$f(x)\text{没有重因式} \Longleftrightarrow (f(x), f'(x)) = 1.$$

(3) 重因式的去除：

$\dfrac{f(x)}{(f(x), f'(x))}$是一个没有重因式的多项式，它与 $f(x)$有相同的不可约因式.

6. 复系数、实系数多项式的因式分解

(1) 复系数多项式可以唯一地表示成一次多项式的乘积.

(2) 实系数多项式可以唯一地表示成一次多项式及二次不可约多项式的乘积.

7. 有理系数多项式

(1) 本原多项式.

(2) 有理系数多项式的因式分解与整系数多项式的因式分解间的关系.

(3) 有理多项式的有理根.

(4) 艾森斯坦判断法.

(5) 对于任意的正整数 n，都有 n 次不可约的有理系数多项式.

复习题 6

1. 求证 $x^{3n}+x^{3m+1}+x^{3l+2}$能被 x^2+x+1 整除.

2. 求 x^2-x+1 能整除 $x^{3n}+x^{3m+1}+x^{3l+2}$的条件.

3. 证明：如果 $x^2+x+1 \mid f_1(x^3)+xf_2(x^3)$，那么

$$x-1 \mid f_1(x), x-1 \mid f_2(x).$$

4. 设 $f_1(x)=af(x)+bg(x)$，$g_1(x)=cf(x)+dg(x)$，且 $ad-bc\neq 0$.

证明：$(f(x),g(x))=(f_1(x),g_1(x))$.

5. 求$(f(x),g(x))$和 $u(x),v(x)$，使 $u(x)f(x)+v(x)g(x)=(f(x),g(x))$：

(1) $f(x)=x^4+2x^3-x^2-4x-2$,

$g(x)=x^4+x^3-x^2-2x-2$;

(2) $f(x)=x^5+3x^4+x^3+x^2+3x+1$,

$g(x)=x^4+2x^3+x+2$;

(3) $f(x)=4x^4-2x^3-16x^2+5x+9$,

$g(x)=2x^3-x^2-5x+4$.

6. 已知有理系数多项式 $f(x)=x^3+(1+t)x^2-3x+2u$ 与 $g(x)=x^3+tx+u$ 的最大公因式是一个二次多项式. 求 t,u 的值和$(f(x),g(x))$.

7. 设多项式 $f_1(x),\cdots,f_m(x);g_1(x),\cdots,g_n(x)$满足

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

求证$(f_1(x)f_2(x)\cdots f_m(x),g_1(x)g_2(x)\cdots g_n(x))=1$.

8. 判断下列多项式有没有重因式，如有，求其重因式和重数：

(1) $x^4-x^3-5x^2+12$;

(2) x^6+x^4+2;

(3) $x^5-5x^4+7x^3-2x^2+4x-8$.

9. 证明实系数多项式

$$x^3-3x^2+3x+a^2$$

没有重因式.

10. 求 a 使下列多项式有重根，并求出重根及其重数：

$$x^3-3x+a.$$

11. 判断下列多项式在有理数域上是否不可约：

(1) x^3+3x^2-x+2;

(2) $2x^3+x^2-3x+1$;

(3) $x^5+4x^4-12x^3+6x^2+8x-2$.

12. a 是一个整数，已知多项式

$$x^3-\frac{1}{2}x^2+ax-2$$

有有理根，求 a 和相应的有理根.

13. 求下列多项式的标准分解式(分实数域与复数域两种情况)：

(1) $2x^5-3x^3-3x^2+x+1$；

(2) $3x^5-2x^4-x^3-6x^2-14x-4$.

14. 试证：次数≥1 且首项系数为 1 的多项式 $f(x)$ 是一个不可约多项式的方幂的充分必要条件是：对于任意的多项式 $g(x)$，都有 $(f(x),g(x))=1$ 或者有某一个正整数 m，使 $f(x)\mid g^m(x)$.

15. 试证：次数≥1 且首项系数为 1 的多项式 $f(x)$ 是某一个不可约多项式的方幂的充分必要条件是：对于任意的多项式 $g(x)$，$h(x)$，由 $f(x)\mid g(x)h(x)$ 可以推出 $f(x)\mid g(x)$ 或者有一个正整数 m，使 $f(x)\mid h^m(x)$.

16. 如果 a 是 $f'''(x)$ 的一个 k 重根，证明：a 是

$$g(x)=\frac{x-a}{2}[f'(x)+f'(a)]-f(x)+f(a)$$

的一个 $k+3$ 重根.

17. 设 $f(x)$ 是一个 n 次多项式，如果 $f'(x)\mid f(x)$，那么 $f(x)$ 有 n 重根.

18. 设 $f(x)$ 是一个整系数多项式，证明：如果 $f(0)$ 与 $f(1)$ 都是奇数，那么 $f(x)$ 不可能有整数根.

19. 多项式 $h(x)$ 称为 $f(x)$ 与 $g(x)$ 的一个公倍式，如果 $h(x)$ 既是 $f(x)$ 的倍式又是 $g(x)$ 的倍式，多项式 $m(x)$ 称为 $f(x)$ 与 $g(x)$ 的一个最小公倍式，如果(1) $m(x)$ 是 $f(x)$ 与 $g(x)$ 的公倍式；(2) $f(x)$ 与 $g(x)$ 的任一个公倍式都是 $m(x)$ 的倍式.

   (1) 证明：如果 $m_1(x)$，$m_2(x)$ 都是 $f(x)$，$g(x)$ 的最小公倍式，那么 $m_1(x)=cm_2(x)$，c 是非零常数；

   (2) 当 $f(x)$，$g(x)$ 不全为 0 时，用 $[f(x),g(x)]$ 表示 $f(x)$，$g(x)$ 的那个首项系数为 1 的最小公倍式，证明：

$$[f(x),g(x)]=\frac{f(x)g(x)}{(f(x),g(x))}.$$

20. 对第 5 题中的 $f(x)$，$g(x)$，求 $[f(x),g(x)]$.

*第 7 章　λ-矩　阵

在第 4 章中，我们曾经讨论过矩阵对角化的问题. 证明了矩阵可对角化的一些条件，并且简单地介绍了约当标准形. 这一章将讨论 λ-矩阵的性质，并应用关于 λ-矩阵的一些结论进一步讨论矩阵相似的条件，从而得到矩阵可对角化的另一些条件；证明关于约当形的主要定理，并应用初等因子给出计算约当标准形的方法.

7.1　λ-矩　阵

设 **P** 是一个数域，λ 是一个文字，用 **P**[λ]表示数域 **P** 上 λ 的多项式的全体组成的集合，称为数域 **P** 上 λ 多项式环. 和 **P**[x]一样，**P**[λ]中的元素可以作加、减、乘 3 种运算，并且它们与数的运算有相同的运算规律. **P**[λ]中元素还可作带余除法. 总之，上一章中的一切结论，对 **P**[λ]都是成立的.

一个矩阵，如果它的元素都是 λ 的多项式，即都是 **P**[λ]中的元素，就称为 λ-矩阵. 因为数域 **P** 中的数也是 **P**[λ]的元素，所以在 λ-矩阵中也包括以数为元素的矩阵. 为了与 λ-矩阵相区别，有时我们把以数域 **P** 中的数为元素的矩阵称为**数字矩阵**. 仍用 A，B 等表示数字矩阵，而用 $A(\lambda)$，$B(\lambda)$等表示 λ-矩阵.

矩阵加法与乘法的定义只是用到其中元素的加法与乘法，因此，可以同样定义 λ-矩阵的加法与乘法，它们与数字矩阵的运算有相同的运算规律. 这些就不重复叙述与证明了.

行列式的定义也只用到其中元素的加法与乘法，因此，同样可以定义一个 $n\times n$ 即 n 阶 λ-矩阵的行列式. 一般地，λ-矩阵的行列

式是 λ 的一个多项式，它与数字矩阵的行列式有相同的性质，例如，对于 λ-矩阵的行列式，矩阵乘积的行列式等于行列式的乘积这一结论，显然是对的.

既然有行列式，也就有 λ-矩阵的子式的概念. 利用这个概念，可以定义 λ-矩阵的秩：

定义 1 如果 λ-矩阵 $A(\lambda)$ 中有一个 $r(r\geqslant 1)$ 阶子式不为零，而所有 $r+1$ 阶子式(如果有的话)全为零，则称 $A(\lambda)$ 的**秩**为 r. 零矩阵的秩规定为零.

例 1

$$A(\lambda)=\begin{bmatrix} 1 & 2 & 2 & 3 \\ \lambda+1 & \lambda & 2\lambda+2 & 2\lambda+1 \\ \lambda^2 & \lambda^2 & 2\lambda^2 & 2\lambda^2 \end{bmatrix}.$$

因为 $A(\lambda)$ 的 4 个 3 阶子式全等于 0，而 $A(\lambda)$ 有一个二阶子式：

$$\begin{bmatrix} 1 & 2 \\ \lambda+1 & \lambda \end{bmatrix}=-\lambda-2\neq 0,$$

所以 $A(\lambda)$ 的秩等于 2.

对于数字矩阵，这与以前的定义是一致的.

与以前一样，我们还有

定义 2 一个 n 阶的 λ-矩阵 $A(\lambda)$ 称为**可逆的**，如果有一个 n 阶的 λ-矩阵 $B(\lambda)$ 使

$$A(\lambda)B(\lambda)=B(\lambda)A(\lambda)=E, \tag{1}$$

这里 E 是 n 级单位矩阵，适合(1)的矩阵 $B(\lambda)$(它是唯一的)称为 $A(\lambda)$ 的**逆矩阵**，记为 $A^{-1}(\lambda)$.

关于 λ-矩阵可逆的条件有

定理 1 一个 n 阶的 λ-矩阵 $A(\lambda)$ 是可逆的充分必要条件为行列式 $|A(\lambda)|$ 是一个非零常数

证明：先证充分性. 设

$$d=|A(\lambda)|$$

是一个非零常数. $A^*(\lambda)$是 $A(\lambda)$的伴随矩阵，它也是一个 λ-矩阵，而

$$A(\lambda)\frac{1}{d}A^*(\lambda)=\frac{1}{d}A^*(\lambda)A(\lambda)=E,$$

因此，$A(\lambda)$可逆.

反过来，如果 $A(\lambda)$可逆，在(1)的两边取行列式，

$$|A(\lambda)||B(\lambda)|=|E|=1.$$

因为$|A(\lambda)|$与$|B(\lambda)|$都是 λ 的多项式，所以由它们的乘积是 1 可以推知，它们都是零次多项式，也就是非零常数. ┃

从证明可以看出，关于可逆矩阵有：

推论 如果 $A(\lambda)$可逆，则

$$A^{-1}(\lambda)=\frac{1}{d}A^*(\lambda).$$

其中 $d=|A(\lambda)|$是数域 $\mathbf{P}$ 中一个非零常数.

例 2 设

$$A(\lambda)=\begin{bmatrix}1 & 2 & 3\\ \lambda+1 & \lambda & 2\lambda+1\\ \lambda^2 & \lambda^2 & 2\lambda^2\end{bmatrix}.$$

因为$|A(\lambda)|=0$，所以 $A(\lambda)$不可逆.

例 3 设

$$B(\lambda)=\begin{bmatrix}1 & 2 & 3\\ \lambda+1 & \lambda & 2\lambda\\ \lambda^2 & \lambda^2 & \lambda^2\end{bmatrix},$$

因为 $$|B(\lambda)|=\lambda^2(\lambda+1)$$

不是一个非零常数，所以 $B(\lambda)$不可逆.

例 4 设

$$C(\lambda)=\begin{bmatrix}1 & 2 & 3\\ \lambda^2+1 & \lambda^2 & \lambda^2\\ \lambda^2 & \lambda^2+1 & \lambda^2+3\end{bmatrix}$$

因为
$$|C(\lambda)| = -3$$
是一个非零常数，所以 $C(\lambda)$ 可逆，并且

$$C^{-1}(\lambda) = \frac{1}{-3}C^{*}(\lambda)$$
$$= \begin{bmatrix} -\frac{2}{3}\lambda^2 & -\frac{1}{3}\lambda^2+1 & \frac{1}{3}\lambda^2 \\ \frac{4}{3}\lambda^2+1 & \frac{2}{3}\lambda^2-1 & -\frac{2}{3}\lambda^2-1 \\ -\frac{2}{3}\lambda^2-\frac{1}{3} & -\frac{1}{3}\lambda^2+\frac{1}{3} & \frac{1}{3}\lambda^2+\frac{2}{3} \end{bmatrix}.$$

如果一个矩阵的行列式不等于 0，则称这个矩阵为**非退化矩阵**，对于数字矩阵 A 来说，A 可逆的充分必要条件是 A 是非退化的. 然而，从定理 1 及例题可以看出，一个 λ-矩阵 A 如果可逆，则 $A(\lambda)$ 一定是非退化的. 但是，当 $A(\lambda)$ 是非退化的，$A(\lambda)$ 不一定是可逆的.

最后，我们引进 λ-矩阵的另一种表示法——多项式表法. 这种表法在下一节中将用到.

设 λ-矩阵 $A(\lambda)$ 的元素的最高次数为 m，则 $A(\lambda)$ 可以表示成：
$$A(\lambda) = \lambda^m A_m + \lambda^{m-1}A_{m-1} + \cdots + \lambda A_1 + A_0.$$
其中 $A_m \neq 0, A_{m-1}, \cdots, A_1, A_0$ 都是数字矩阵，m 称为这个 λ-矩阵的次数，这种表式可看成以数字矩阵为系数的多项式，叫做矩阵多项式.

例 5

$$A(\lambda) = \begin{bmatrix} \lambda^2 & 1 & 2 \\ 1 & \lambda^3+2 & \lambda^2+1 \\ \lambda^3+\lambda^2 & \lambda^2+1 & 2\lambda^3+3\lambda^2+\lambda-1 \end{bmatrix}$$
$$= \lambda^3\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} + \lambda^2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

$$+\lambda\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}+\begin{bmatrix}0&1&2\\1&2&1\\0&1&-1\end{bmatrix}.$$

$A(\lambda)$的次数为 3.

如果 $A(\lambda)$，$B(\lambda)$是两个 λ-矩阵，它们的次数分别为 m，n. 表示成多项式矩阵：

$$A(\lambda)=\lambda^m A_m+\lambda^{m-1}A_{m-1}+\cdots+\lambda A_1+A_0,$$

$$B(\lambda)=\lambda^n B_n+\lambda^{n-1}B_{n-1}+\cdots+\lambda B_1+B_0.$$

则 $A(\lambda)=B(\lambda)$，当且仅当 $m=n$ 而且 $A_i=B_i, i=m,\cdots,1,0$.

两个矩阵多项式的相加、相减、相乘与一般多项式的运算一样. 但是，由于矩阵的乘法不满足消去律，因此，两个矩阵多项式相乘不满足多项式乘法的次数规律.

例 6 设

$$A(\lambda)=\lambda^3\begin{bmatrix}1&0\\0&0\end{bmatrix}+\lambda\begin{bmatrix}1&2\\-1&0\end{bmatrix}+\begin{bmatrix}1&1\\0&0\end{bmatrix}$$

$$B(\lambda)=\lambda^2\begin{bmatrix}0&0\\2&1\end{bmatrix}+\begin{bmatrix}1&2\\-1&-2\end{bmatrix}$$

则 $A(\lambda)$，$B(\lambda)$的次数分别为 3，2. 它们的乘积为：

$$A(\lambda)B(\lambda)=\lambda^3\begin{bmatrix}5&4\\0&0\end{bmatrix}+\lambda^2\begin{bmatrix}2&1\\0&0\end{bmatrix}+\lambda\begin{bmatrix}-1&-2\\-1&-2\end{bmatrix}.$$

$A(\lambda)B(\lambda)$的次数为 3.

和数字矩阵一样，还可以出现 $A(\lambda)\neq 0$，$B(\lambda)\neq 0$，而 $A(\lambda)B(\lambda)=0$的情形.

例 7 设

$$A(\lambda)=\begin{bmatrix}\lambda-1&\lambda+1\\-\lambda+1&-\lambda-1\end{bmatrix}$$

$$=\begin{bmatrix}1&1\\-1&-1\end{bmatrix}\lambda+\begin{bmatrix}-1&1\\1&-1\end{bmatrix}$$

$$B(\lambda)=\begin{bmatrix}\lambda+1 & \lambda^2+\lambda\\ -\lambda+1 & -\lambda^2+\lambda\end{bmatrix}$$
$$=\lambda^2\begin{bmatrix}0 & 1\\ 0 & -1\end{bmatrix}+\lambda\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}+\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}$$

就有 $A(\lambda)\neq 0,B(\lambda)\neq 0$,而 $A(\lambda)B(\lambda)=0$

习 题 7.1

1. 求下列 λ-矩阵的秩:

$$(1)\begin{bmatrix}1 & -1 & 2\\ \lambda & -\lambda & 2\lambda\\ \lambda+1 & \lambda-1 & 2\lambda+2\end{bmatrix};\quad (2)\begin{bmatrix}1 & \lambda & \lambda^2\\ \lambda+1 & 2\lambda & 3\lambda^2\\ \lambda & \lambda+1 & \lambda^2\end{bmatrix}.$$

2. 求第 1 题中 λ-矩阵的伴随矩阵.

3. 判断下列 λ-矩阵是否可逆. 如可逆,求其逆矩阵:

$$(1)\begin{bmatrix}1 & -1 & \lambda\\ 2 & \lambda+1 & 2\lambda\\ \lambda & -1 & \lambda+1\end{bmatrix};\quad (2)\begin{bmatrix}\lambda+1 & \lambda-1 & -2\\ -1 & \lambda^2+\lambda & \lambda^2+\lambda+1\\ 1 & \lambda & \lambda-1\end{bmatrix};$$

$$(3)\begin{bmatrix}1 & \lambda & \lambda+1\\ \lambda & \lambda^2+2 & 2\lambda^2+\lambda\\ 3 & 3\lambda+2 & \lambda^2+3\lambda+4\end{bmatrix}.$$

4. A 是一个数字矩阵. 试证:A 的特征矩阵作为 λ-矩阵是非退化的,但不是可逆的.

7.2 最小多项式

首先证明关于数字矩阵的一个重要性质:

定理 2(哈密尔顿-凯莱 Hamilton-Caylay 定理). 设 A 是数域 $\mathbf{P}$ 上一个 n 级方阵. A 的特征多项式 $|\lambda E-A|=f(\lambda)$. 则 $f(A)=O$.

证明: 设 $f(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$. 用 $B(\lambda)$表示 λE

$-A$ 的伴随矩阵. 根据伴随矩阵的行列式的性质，有

$$\begin{aligned}B(\lambda)(\lambda E-A)&=|\lambda E-A|E=f(\lambda)E\\&=\lambda^nE+a_{n-1}\lambda^{n-1}E+\cdots+a_1\lambda E+a_0E.\end{aligned}\tag{1}$$

因为矩阵 $B(\lambda)$的元素是 $\lambda E-A$ 的各个代数余子式，都是 λ 的多项式，其次数不超过 $n-1$，所以 $B(\lambda)$的矩阵多项式表示为：

$$B(\lambda)=\lambda^{n-1}B_{n-1}+\lambda^{n-2}B_{n-2}+\cdots+\lambda B_1+B_0.$$

于是

$$\begin{aligned}B(\lambda)(\lambda E-A)&=(\lambda^{n-1}B_{n-1}+\cdots+\lambda B_1+B_0)(\lambda E-A)\\&=\lambda^nB_{n-1}+\lambda^{n-1}(B_{n-2}-B_{n-1}A)\\&\quad+\cdots+\lambda(B_0-B_1A)-B_0A\end{aligned}\tag{2}$$

比较(1)，(2)两式，得

$$\begin{cases}B_{n-1}=E,\\B_{n-2}-B_{n-1}A=a_{n-1}E,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\B_0-B_1A=a_1E,\\-B_0A=a_0E.\end{cases}\tag{3}$$

以 $A^n,A^{n-1},\cdots,A,E$ 依次从右边乘(3)的第 1 式，第 2 式，…，第 n 式，第 $n+1$ 式，得：

$$\begin{cases}B_{n-1}A^n=A^n,\\B_{n-2}A^{n-1}-B_{n-1}A^n=a_{n-1}A^{n-1},\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\B_0A-B_1A^2=a_1A,\\-B_0A=a_0E.\end{cases}\tag{4}$$

把(4)中 $n+1$ 个式子相加，左边为 O，右边为

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

因此 $f(A)=O$，定理证毕. ▎

例 1 设

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

A 的特征多项式为：

$$f(\lambda)=|\lambda E-A|=\lambda^3-4\lambda^2+4\lambda+5$$

$$f(A)=A^3-4A^2+4A+5E$$

$$=\begin{bmatrix}-1&4&-8\\8&11&-8\\32&32&-9\end{bmatrix}-4\begin{bmatrix}2&3&-3\\4&5&-2\\11&9&1\end{bmatrix}+4\begin{bmatrix}1&2&-1\\2&1&0\\3&1&2\end{bmatrix}$$

$$+5\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}=0$$

根据定理 2，任给数域 **P** 上一个 n 阶矩阵 A，总可找到数域 **P** 上一个多项式 $f(x)$ 使 $f(A)=O$. 如果多项式 $f(x)$ 使 $f(A)=O$. 我们就称 $f(x)$ 以 A 为根.

定义 3 A 是一个方阵，以 A 为根的次数最低，首项系数为 1 的多项式称为 A 的**最小多项式**.

这一节讨论如何应用最小多项式来判断一个矩阵能否对角化的问题.

首先介绍最小多项式的一些基本性质.

命题 1 矩阵 A 的最小多项式是唯一的.

证明：设 $g_1(x)$ 和 $g_2(x)$ 都是 A 的最小多项式. 则它们的次数相同. 根据带余除法，$g_1(x)$ 可表示成：

$$g_1(x)=q(x)g_2(x)+r(x).$$

其中 $r(x)=0$ 或 $\deg r(x)<\deg q_2(x)$. 于是

$$g_1(A)=q(A)g_2(A)+r(A).$$

因为 $g_1(A)=g_2(A)=O$，所以 $r(A)=O$. 根据最小多项式的定义，必有 $r(x)=0$. 因此 $g_2(x)\mid g_1(x)$. 又因 $\deg g_1(x)=\deg g_2(x)$，所

以$g_1(x)=cg_2(x)$. 根据 $g_1(x)$,$g_2(x)$的首项系数都等于 1,所以 $c=1$,即 $g_1(x)=g_2(x)$.

命题 2 设 $g(x)$是矩阵 A 的最小多项式,那么 $f(x)$以 A 为根的充分必要条件是 $f(x)$可被 $g(x)$整除.

由命题 2 可得,证法和命题 1 相同,留给读者作为习题.

命题 3 矩阵 A 的最小多项式是 A 的特征多项式的一个因式.

例 2 数量矩阵 kE 的最小多项式为 $x-k$. 特别地,单位矩阵的最小多项式为 $x-1$,零矩阵的最小多项式为 x.

例 3 设

$$A=\begin{bmatrix}1 & 1 & \\ & 1 & \\ & & 1\end{bmatrix}$$

求 A 的最小多项式.

解:因为

$$|\lambda E-A|=(\lambda-1)^3$$

所以 A 的最小多项式为$(\lambda-1)^3$ 的因式,又因 $A-E\neq O$,而 $(A-E)^2=O$,所以 A 的最小多项式为$(\lambda-1)^2$.

如果矩阵 A 与 B 相似,那么有可逆矩阵 X 使 $B=X^{-1}AX$. 对任一多项式 $f(x)$有 $f(B)=X^{-1}f(A)X$,因此 $f(B)=O$ 当且仅当 $f(A)=O$. 这说明相似的矩阵有相同的最小多项式. 但是,这个条件不是充分的,即最小多项式相同的矩阵不一定是相似的. 下面的例子说明了这一事实.

例 4 设

$$A=\begin{bmatrix}1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 2\end{bmatrix},\quad B=\begin{bmatrix}1 & 1 & & \\ & 1 & & \\ & & 2 & \\ & & & 2\end{bmatrix},$$

则 A 与 B 有相同的最小多项式$(\lambda-1)^2(\lambda-2)$. 但是它们的特征多项式不同,所以不是相似的.

例 5 设

$$A=\begin{bmatrix}1&1&&\\&1&&\\&&1&1\\&&&1\end{bmatrix},\quad B=\begin{bmatrix}1&1&&\\&1&&\\&&1&\\&&&1\end{bmatrix},$$

则 A 与 B 的最小多项式及特征多项式都相同. 但是 A 有 2 个线性无关的特征向量,B 有 3 个线性无关的特征向量,所以 A 与 B 不是相似的.

为了讨论矩阵对角化问题,还需要用到下述命题.

命题 4 A 是一个准对角矩阵

$$A=\begin{bmatrix}A_1&\\&A_2\end{bmatrix}.$$

设 A_1 的最小多项式为 $g_1(x)$,A_2 的最小多项式为 $g_2(x)$,则 A 的最小多项式为 $g_1(x)$,$g_2(x)$的最小公倍式$[g_1(x),g_2(x)]$.

证明:记 $g(x)=[g_1(x),g_2(x)]$首先

$$g(A)=\begin{bmatrix}g(A_1)&\\&g(A_2)\end{bmatrix}=0.$$

因此 $g(x)$以 A 为根,其次,如果 $h(A)=O$,那么

$$h(A)=\begin{bmatrix}h(A_1)&\\&h(A_2)\end{bmatrix}=0$$

因此 $h(A_1)=O$,$h(A_2)=O$,从而 $g_1(x)\mid h(x)$,$g_2(x)\mid h(x)$. 由此得 $g(x)\mid h(x)$. 又因 $g(x)$的首项系数为 1,所以 $g(x)$是 A 的最小多项式.

这个结论可以推广到 A 为 n 个矩阵组成的准对角矩阵的情形. 即,如果

$$A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{bmatrix},$$

A_i 的最小多项式为 $g_i(x), i=1,2,\cdots,s$，则 A 的最小多项式为 $[g_1(x), g_2(x), \cdots, g_s(x)]$.

命题 5 k 级若当块

$$J = \begin{bmatrix} a & 1 & & & \\ & a & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & a \end{bmatrix}$$

的最小多项式为 $(x-a)^k$.

证明：J 的特征多项式为 $(x-a)^k$，而

$$J - aE = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} (k \text{ 阶})$$

所以

$$(J-aE)^k = 0, (J-aE)^{k-1} \neq 0.$$

因此 J 的最小多项式为 $(x-a)^k$.

定理 3 数域 $\mathbf{P}$ 上 n 阶矩阵 A 可对角化的充分必要条件是：A 的最小多项式是 $\mathbf{P}$ 上互素的 1 次因式的乘积.

证明：根据命题 4 推广的情形，条件的必要性是显然的.

反之，如果 A 不能对角化，则 A 看成复数域上矩阵，其若当形中有 $k>1$ 的约当块，因此由命题 4 和 5 知，A 的最小多项式中有 $(x-a)^k, k>1$ 的因式，不可能表示成互素的一次因式的乘积.

推论 复系数矩阵 A 可对角化的充分必要条件是 A 的最小

多项式没有重根.

例 6 求矩阵 A 的约当形，

$$A=\begin{bmatrix}3 & 1 & 0\\ -4 & -1 & 0\\ 4 & -8 & -2\end{bmatrix}.$$

解： A 的特征多项式为

$$|\lambda E-A|=\begin{vmatrix}\lambda-3 & -1 & 0\\ 4 & \lambda+1 & 0\\ -4 & 8 & \lambda+2\end{vmatrix}$$

$$=(\lambda+2)(\lambda^2-2\lambda+1)$$

$$=(\lambda+2)(\lambda-1)^2.$$

因为

$$(A+2E)(A-E)=\begin{bmatrix}5 & 1 & 0\\ -4 & 1 & 0\\ 4 & -8 & 0\end{bmatrix}\begin{bmatrix}2 & 1 & 0\\ -4 & -2 & 0\\ 4 & -8 & -3\end{bmatrix}$$

$$\neq 0$$

所以 A 不能对角化. 又因 A 的特征值为 1，1，-2，所以 A 的约当形为

$$\begin{bmatrix}1 & 1 & \\ & 1 & \\ & & -2\end{bmatrix}.$$

最后我们来证明最小多项式的一个重要性质：

定理 4 矩阵 A 的特征值都是 A 的最小多项式的根.

证明： 设 A 的特征值为 $\overbrace{\lambda_1,\cdots,\lambda_1}^{k_1\text{个}},\overbrace{\lambda_2,\cdots,\lambda_2}^{k_2\text{个}},\cdots,\overbrace{\lambda_s,\cdots,\lambda_s}^{k_s},k_1+k_2+\cdots+k_s=n$（$A$ 的阶数），则 A 的约当形为

$$\begin{bmatrix} \overbrace{\begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{matrix}}^{k_1\text{个}} & & * \\ & \ddots & \\ & & \overbrace{\begin{matrix}\lambda_s & & \\ & \ddots & \\ & & \lambda_s\end{matrix}}^{k_s\text{个}} \end{bmatrix}$$

因此,如果 $h(A)=O$,则 $h(\lambda)$必须有一次因式 $\lambda-\lambda_i, i=1,2,\cdots,s$,所以 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 都是 $h(\lambda)$的根. 由此可知,A 的特征值都是 A 的最小多项式的根. |

推论 矩阵 A 的特征多项式与最小多项式有相同的不可约因式.

习 题 7.2

1. 求下列矩阵的最小多项式:

$$(1)\ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix};\quad (2)\ \begin{bmatrix} 3 & -1 & -3 & 1 \\ -1 & 3 & 1 & -3 \\ 3 & -1 & -3 & 1 \\ -1 & 3 & 1 & -3 \end{bmatrix}.$$

2. 讨论第 1 题中两个矩阵能否对角化.
3. 证明:多项式 $f(x)$以 A 为根的充分必要条件是 $f(x)$可以被 A 的最小多项式整除.
4. 证明:矩阵 A 可逆的充分必要条件是 A 的最小多项式的常数项不等于 0.

7.3 λ-矩阵的等价标准形

这一节讨论 λ-矩阵在等价关系下的标准形. 和数字矩阵的情形一样,λ-矩阵也有初等变换:

定义 4 下面的 3 种变换叫做 λ-矩阵的**初等变换**:

(1) 矩阵的两行(列)互换位置;

(2) 矩阵的某一行(列)乘以非零的常数 c;

(3) 矩阵的某一行(列)加另一行(列)的 $\varphi(\lambda)$ 倍,$\varphi(\lambda)$ 是一个多项式.

定义 5 如果 λ-矩阵 $A(\lambda)$ 可以经过一系列初等变换化为 $B(\lambda)$,则称 $A(\lambda)$ 与 $B(\lambda)$ **等价**.

等价是 λ-矩阵之间的一种关系,这个关系,显然具有下列 3 个性质:

(1) 反身性:每一个 λ-矩阵与自己等价.

(2) 对称性:若 $A(\lambda)$ 与 $B(\lambda)$ 等价,则 $B(\lambda)$ 与 $A(\lambda)$ 等价.这是由于初等变换具有可逆性的缘故.

(3) 传递性:若 $A(\lambda)$ 与 $B(\lambda)$ 等价,$B(\lambda)$ 与 $C(\lambda)$ 等价,则 $A(\lambda)$ 与 $C(\lambda)$ 等价.

为了将 λ-矩阵的等价关系通过矩阵运算来表示,必须建立初等变换与矩阵的关系.为此,和数字矩阵一样,首先介绍初等矩阵的概念.

定义 6 由单位矩阵 E 经过一次 λ-矩阵的初等变换而得到的 λ-矩阵称为**初等 λ-矩阵**,简称**初等矩阵**.

例如,将单位矩阵的第 j 行的 $\varphi(\lambda)$ 倍加到第 i 行上得:

$$P(i,j(\varphi))=\begin{array}{l}\begin{array}{ccccccc} & & i\text{列} & & j\text{列} & & \end{array}\\ \left[\begin{array}{ccccccc}1&&&&&&\\&\ddots&&&&&\\&&1&\cdots&\varphi(\lambda)&&\\&&&\ddots&\vdots&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{array}\right]\begin{array}{l}\\ \\ i\text{行}\\ \\ j\text{行}\\ \\ \\ \end{array}\end{array}.$$

仍用 $P(i,j)$ 表示由单位矩阵经过第 i 行第 j 行互换位置所得的初

等矩阵,用 $P(i(c))$ 表示用非零常数 c 乘单位矩阵第 i 行所得的初等矩阵. 同样地,对一个 $s\times n$ 的 λ-矩阵 $A(\lambda)$ 作一次初等行变换就相当于在 $A(\lambda)$ 的左边乘上相应的 $s\times s$ 初等矩阵;对 $A(\lambda)$ 作一次初等列变换就相当于在 $A(\lambda)$ 的右边乘上相应的 $n\times n$ 的初等矩阵.

初等矩阵都是可逆的,并且它的逆矩阵也是初等矩阵:

$$P(i,j)^{-1}=P(i,j),P(i(c))^{-1}=P(i(c^{-1})),$$
$$P(i,j(\varphi))^{-1}=P(i,j(-\varphi)).$$

由此得出初等变换具有可逆性:设 λ-矩阵 $A(\lambda)$ 用初等变换变成 $B(\lambda)$,这相当于对 $A(\lambda)$ 左乘或右乘一个初等矩阵. 再用此初等矩阵的逆矩阵来乘 $B(\lambda)$ 就变回 $A(\lambda)$,而这逆矩阵仍是初等矩阵,因而由 $B(\lambda)$ 可用初等变换变回 $A(\lambda)$. 我们还可看出,在第(2)种初等变换中,规定只能乘以一个非零常数,这也是为了使 $P(i(c))$ 可逆的缘故.

应用初等变换与初等矩阵的关系即得下述定理:

定理 5 矩阵 $A(\lambda)$ 与 $B(\lambda)$ 等价的充分必要条件为有一系列初等矩阵 $P_1(\lambda),P_2(\lambda),\cdots,P_l(\lambda);Q_1(\lambda),Q_2(\lambda),\cdots,Q_t(\lambda)$ 使

$$B(\lambda)=P_l(\lambda)\cdots P_2(\lambda)P_1(\lambda)A(\lambda)Q_1(\lambda)Q_2(\lambda)\cdots Q_t(\lambda).$$

推论 如果 $A(\lambda)$ 与 $B(\lambda)$ 等价,那么 $|A(\lambda)|=c|B(\lambda)|$,$c$ 是一个非零常数.

这一节主要是证明任意一个 λ-矩阵可以经过初等变换化为某种对角形.

为了写起来方便起见,我们采用以下的记号:

$[i,j]$ 代表 i,j 行(列)互换位置;

$[i(c)]$ 代表用非零的数 c 去乘 i 行(列);

$[i+j(\varphi)]$ 代表把 j 行(列)的 $\varphi(\lambda)$ 倍加到 i 行(列).

首先证明下述命题:

命题 1 设 λ-矩阵 $A(\lambda)$ 的左上角元素 $a_{11}(\lambda)\neq 0$,并且 $A(\lambda)$

中至少有一个元素不能被它除尽，那么一定可以找到一个与 $A(\lambda)$ 等价的矩阵 $B(\lambda)$，它的左上角元素也不为零，但是次数比 $a_{11}(\lambda)$ 的次数低.

证明：根据 $A(\lambda)$ 中不能被 $a_{11}(\lambda)$ 除尽的元素所在的位置，分 3 种情形来讨论：

(1) 若在 $A(\lambda)$ 的第 1 列中有一个元素 $a_{i1}(\lambda)$ 不能被 $a_{11}(\lambda)$ 除尽，则有

$$a_{i1}(\lambda) = a_{11}(\lambda)q(\lambda) + r(\lambda),$$

其中余式 $r(\lambda)\neq 0$，且次数比 $a_{11}(\lambda)$ 的次数低.

对 $A(\lambda)$ 作下列初等行变换：

$$A(\lambda) = \begin{bmatrix} a_{11}(\lambda) & \cdots \\ \cdots\cdots & \cdots \\ a_{i1}(\lambda) & \cdots \\ \cdots\cdots & \cdots \end{bmatrix} \xrightarrow{[i-1(q)]^*} \begin{bmatrix} a_{11}(\lambda) & \cdots \\ \cdots\cdots & \cdots \\ r(\lambda) & \cdots \\ \cdots\cdots & \cdots \end{bmatrix}$$

$$\xrightarrow{[1,i]} \begin{bmatrix} r(\lambda) & \cdots \\ \cdots\cdots & \cdots \\ a_{11}(\lambda) & \cdots \\ \cdots\cdots & \cdots \end{bmatrix} = B(\lambda).$$

$B(\lambda)$ 的左上角元素 $r(\lambda)$ 符合命题 1 的要求，故 $B(\lambda)$ 即为所求的矩阵.

(2) 在 $A(\lambda)$ 的第 1 行中有一个元素 $a_{1i}(\lambda)$ 不能被 $a_{11}(\lambda)$ 除尽，这种情况的证明与(1)类似.

(3) $A(\lambda)$ 的第 1 行与第 1 列中的元素都可以被 $a_{11}(\lambda)$ 除尽，但 $A(\lambda)$ 中有另一个元素 $a_{ij}(\lambda)(i>1, j>1)$ 不能被 $a_{11}(\lambda)$ 除尽. 我们设 $a_{i1}(\lambda)=a_{11}(\lambda)\varphi(\lambda)$. 对 $A(\lambda)$ 作下述初等行变换：

---

*变换记号写在"→"的上边表示行变换，写在下边表示列变换.

$$A(\lambda)=\begin{bmatrix} a_{11}(\lambda) & \cdots & a_{1j}(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1}(\lambda) & \cdots & a_{ij}(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$\xrightarrow{[i-1(\varphi)]}\begin{bmatrix} a_{11}(\lambda) & \cdots & a_{1j}(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & a_{ij}(\lambda)-a_{1j}(\lambda)\varphi(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$\xrightarrow{[1+i]}\begin{bmatrix} a_{11}(\lambda) & \cdots & a_{1j}(\lambda)+(1-\varphi)a_{1j}(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & a_{ij}(\lambda)-a_{1j}(\lambda)\varphi(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$=A_1(\lambda).$$

矩阵 $A_1(\lambda)$的第 1 行中，有一个元素 $a_{ij}(\lambda)+(1-\varphi)a_{1j}(\lambda)$不能被左上角元素$a_{11}(\lambda)$除尽，这就化为已经证明了的情况(2).

定理 6 任意一个非零的 $s\times n$ 的 λ-矩阵 $A(\lambda)$都等价于一个下列形式的矩阵

$$\begin{bmatrix} d_1(\lambda) & & & & \\ & d_2(\lambda) & & O & \\ & & \ddots & & \\ & & & d_r(\lambda) & \\ & O & & & O \end{bmatrix},$$

其中 $r\geqslant 1$，$d_i(\lambda)(i=1,2,\cdots,r)$是首项系数为 1 的多项式，且

$$d_i(\lambda)\mid d_{i+1}(\lambda) \quad (i=1,2,\cdots,r-1).$$

证明：经过行列调动之后，可以使得 $A(\lambda)$的左上角元素 $a_{11}(\lambda)\neq 0$. 如果 $a_{11}(\lambda)$不能除尽 $A(\lambda)$的全部元素，由命题 1 可以找到与 $A(\lambda)$等价的 $B_1(\lambda)$，它的左上角元素 $b_1(\lambda)\neq 0$，并且次数比

$a_{11}(\lambda)$低. 如果$b_1(\lambda)$还不能除尽$B_1(\lambda)$的全部元素，由命题1又可以找到与$B_1(\lambda)$等价的$B_2(\lambda)$，它的左上角元素$b_2(\lambda)\neq 0$，并且次数比$b_1(\lambda)$低. 如此下去，将得到一系列彼此等价的λ-矩阵$A(\lambda)$，$B_1(\lambda)$，$B_2(\lambda)$，…. 它们的左上角元素皆不为零，而且次数越来越低. 但次数是非负整数，不可能无止境地降低，因此，在有限步以后，将终止于一个λ-矩阵$B_s(\lambda)$，它的左上角元素$b_s(\lambda)\neq 0$，而且可以除尽$B_s(\lambda)$的全部元素$b_{ij}(\lambda)$，即

$$b_{ij}(\lambda) = b_s(\lambda)q_{ij}(\lambda),$$

对$B_s(\lambda)$作初等变换：

$$B_s(\lambda) = \begin{bmatrix} b_s(\lambda) & \cdots & b_{1j}(\lambda) & \cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ b_{i1}(\lambda) & \cdots & \cdots & \cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \end{bmatrix}$$

$$\xrightarrow[{[2-1(q_{12})],[3-1(q_{13})],\cdots\cdots}]{{[2-1(q_{21})],[3-1(q_{31})],\cdots\cdots}} \begin{bmatrix} b_s(\lambda) & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_1(\lambda) & \\ 0 & & & \end{bmatrix}.$$

在右下角的λ-矩阵$A_1(\lambda)$中，全部元素都是可以被$b_s(\lambda)$除尽的，因为它们都是$B_s(\lambda)$中元素的组合.

如果$A_1(\lambda)\neq 0$，则对于$A_1(\lambda)$可以重复上述过程，进而把矩阵化成

$$\begin{bmatrix} d_1(\lambda) & 0 & \cdots & 0 \\ 0 & d_2(\lambda) & \cdots & 0 \\ 0 & 0 & & \\ \vdots & \vdots & A_2(\lambda) & \\ 0 & 0 & & \end{bmatrix},$$

其中$d_1(\lambda)$与$d_2(\lambda)$都是首项系数为1的多项式($d_1(\lambda)$与$b_s(\lambda)$只差一个常数倍数)，而且$d_1(\lambda)\mid d_2(\lambda)$，$d_2(\lambda)$能除尽$A_2(\lambda)$的全部

元素.

如此下去,$A(\lambda)$最后就化成了所要求的形式. ∎

最后化成的这个矩阵称为 $A(\lambda)$的**标准形**.

例 1 用初等变换化 λ-矩阵

$$A(\lambda)=\begin{bmatrix}1-\lambda & 2\lambda-1 & \lambda \\ \lambda & \lambda^2 & -\lambda \\ 1+\lambda^2 & \lambda^3+\lambda-1 & -\lambda^2\end{bmatrix}$$

为标准形.

解:

$$A(\lambda)\xrightarrow[{[3+1]}]{}\begin{bmatrix}1-\lambda & 2\lambda-1 & 1 \\ \lambda & \lambda^2 & 0 \\ 1+\lambda^2 & \lambda^3+\lambda-1 & 1\end{bmatrix}$$

$$\xrightarrow[{[1,3]}]{}\begin{bmatrix}1 & 2\lambda-1 & 1-\lambda \\ 0 & \lambda^2 & \lambda \\ 1 & \lambda^3+\lambda-1 & 1+\lambda^2\end{bmatrix}$$

$$\xrightarrow{[3-1]}\begin{bmatrix}1 & 2\lambda-1 & 1-\lambda \\ 0 & \lambda^2 & \lambda \\ 0 & \lambda^3-\lambda & \lambda^2+\lambda\end{bmatrix}$$

$$\xrightarrow[{[2-1(2\lambda-1)],[3+1(\lambda-1)]}]{}\begin{bmatrix}1 & 0 & 0 \\ 0 & \lambda^2 & \lambda \\ 0 & \lambda^3-\lambda & \lambda^2+\lambda\end{bmatrix}$$

$$\xrightarrow[{[2,3]}]{}\begin{bmatrix}1 & 0 & 0 \\ 0 & \lambda & \lambda^2 \\ 0 & \lambda^2+\lambda & \lambda^3-\lambda\end{bmatrix}$$

$$\xrightarrow[{[3-2(\lambda)]}]{}\begin{bmatrix}1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & \lambda^2+\lambda & -\lambda^2-\lambda\end{bmatrix}$$

$$\xrightarrow[{[3(-1)]}]{[3-2(\lambda+1)]}\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda^2+\lambda\end{bmatrix}=B(\lambda).$$

$B(\lambda)$就是$A(\lambda)$的标准形.

因为一个λ-矩阵可逆的充分必要条件是它的行列式是一个非零常数,因此由定理 6 可以得到下列一些结论.

定理 7 λ-矩阵$A(\lambda)$可逆的充分必要条件是:$A(\lambda)$的标准形为单位矩阵E.

定理 8 λ-矩阵$A(\lambda)$可逆的充分必要条件是:它能表成一些初等矩阵的乘积.

定理 9 两个$s\times n$ λ-矩阵$A(\lambda)$,$B(\lambda)$等价的充分必要条件是:存在可逆的s阶λ-矩阵$P(\lambda)$及可逆的n阶λ-矩阵$Q(\lambda)$,使得

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda).$$

例 2 对例 1 中的$A(\lambda)$及其标准形$B(\lambda)$,求可逆矩阵$P(\lambda)$及$Q(\lambda)$,使

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda).$$

解:根据例 1 的计算有:

$$\begin{aligned}B(\lambda)=&P[3,2(-\lambda-1)]P[3,1(-1)]A(\lambda)(P[3,1(1)])^{T}\\ &(P[1,3])^{T}P([2,1(-2\lambda+1)])^{T}\\ &\cdot P([3,1(\lambda-1)])^{T}(P[2,3])^{T}(P[3,2(-\lambda)])^{T}\\ &\cdot(P[3(-1)])^{T}\end{aligned}$$

所以

$$\begin{aligned}P(\lambda)&=P[3,2(-\lambda-1)]P[3,1(-1)]\\ &=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -\lambda-1 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -1 & 0 & 1\end{bmatrix}\end{aligned}$$

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

$$Q(\lambda)=(P[3,1(1)])^{T}(P[1,3])^{T}(P[2,1(-2\lambda+1)])^{T}$$
$$\cdot(P[3,1(\lambda-1)])^{T}(P[2,3])^{T}(P[3,2(-\lambda)])^{T}$$
$$\cdot(P[3(-1)])^{T}$$

$$=\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}\begin{bmatrix}1 & -2\lambda+1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$
$$\cdot\begin{bmatrix}1 & 0 & \lambda-1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & -\lambda\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{bmatrix}$$

$$=\begin{bmatrix}1 & \lambda & \lambda^2+2\lambda-1\\ 0 & 0 & -1\\ 1 & \lambda-1 & \lambda^2+\lambda-1\end{bmatrix}$$

习 题 7.3

用初等变换把 $A(\lambda)$ 化为标准形 $B(\lambda)$，并求可逆矩阵 $P(\lambda)$ 和 $Q(\lambda)$，使 $B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$：

1. $A(\lambda)=\begin{bmatrix}\lambda^2-\lambda & 2\lambda^2\\ \lambda^2 & 3\lambda\end{bmatrix}$

2. $A(\lambda)=\begin{bmatrix}1-\lambda & \lambda^2 & \lambda\\ \lambda & \lambda & -\lambda\\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{bmatrix}$

3. $A(\lambda)=\begin{bmatrix}\lambda^2+\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & (\lambda+1)^2\end{bmatrix}$

4. $A(\lambda)=\begin{bmatrix}\lambda+1 & 1 & 0 & 0\\ \lambda^2-1 & \lambda-1 & 0 & 0\\ 0 & 0 & \lambda+1 & 1\\ 0 & 0 & \lambda^2-1 & \lambda-1\end{bmatrix}$

7.4 不变因子

这一节讨论 λ-矩阵等价标准形的唯一性. 为此，首先介绍下述定义:

定义 7 设 λ-矩阵 $A(\lambda)$的秩为 r. 对于正整数 k，$1\leqslant k\leqslant r$，$A(\lambda)$中必有非零的 k 阶子式. $A(\lambda)$的全部 k 阶子式的首项系数为 1 的最大公因式 $D_k(\lambda)$称为 $A(\lambda)$的 **k 阶行列式因子**.

例 1 设

$$A(\lambda)=\begin{bmatrix}1-\lambda & \lambda^2 & \lambda\\ \lambda & \lambda & -\lambda\\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{bmatrix}$$

的各阶行列式因子.

解:

$$|A(\lambda)|=-\lambda^2(\lambda+1),$$

所以 $P_3=\lambda^2(\lambda+1)$.

$A(\lambda)$有一个 2 阶子式

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

而 $A(\lambda)$中除第 1 列外，其它元素都是 λ 的信式，故 A 的 2 阶子式必以 λ 为因式，因此

$$D_2(\lambda)=\lambda.$$

由于 $A(\lambda)$中有 2 个 1 阶子式互素，所以 $D_1(\lambda)=1$.

由定义可知，对于秩为 r 的 λ-矩阵，行列式因子一共有 r 个.

行列式因子的意义就在于，它在初等变换下是不变的.

定理 10 等价的矩阵具有相同的秩与相同的各阶行列式因子.

证明：我们只需要证明，λ-矩阵经过一次初等变换，秩与行列式因子是不变的.

设 λ-矩阵 $A(\lambda)$ 经过一次初等行变换变成 $B(\lambda)$. $f(\lambda)$ 与 $g(\lambda)$ 分别是 $A(\lambda)$ 与 $B(\lambda)$ 的 k 阶行列式因子. 为了证明 $f(\lambda)=g(\lambda)$，分下面 3 种情形讨论：

(1) $A(\lambda)\xrightarrow{[i,j]}B(\lambda)$

这时，$B(\lambda)$ 的每个 k 阶子式或者等于 $A(\lambda)$ 的某个 k 阶子式，或者与 $A(\lambda)$ 的某一个 k 阶子式反号，因此 $f(\lambda)$ 是 $B(\lambda)$ 的 k 阶子式的公因式，从而 $f(\lambda)\mid g(\lambda)$.

(2) $A(\lambda)\xrightarrow{[i(c)]}B(\lambda)$

这时，$B(\lambda)$ 的每个 k 阶子式或者等于 $A(\lambda)$ 的某一个 k 阶子式. 或者等于 $A(\lambda)$ 的某一个 k 阶子式的 c 倍. 因此 $f(\lambda)$ 是 $B(\lambda)$ 的 k 阶子式的公因式，从而 $f(\lambda)\mid g(\lambda)$.

(3) $A(\lambda)\xrightarrow{[i+j(\varphi)]}B(\lambda)$

这时 $B(\lambda)$ 中那些包含 i 行与 j 行的 k 阶子式和那些不包含 i 行的 k 阶子式都等于 $A(\lambda)$ 中对应的 k 阶子式；$B(\lambda)$ 中那些包含 i 行但不包含 j 行的 k 阶子式，按 i 行分成两部分，而等于 $A(\lambda)$ 的一个 k 阶子式与另一个 k 阶子式的 $\pm\varphi(\lambda)$ 倍的和，也就是 $A(\lambda)$ 的两个 k 阶子式的组合. 因此 $f(\lambda)$ 是 $B(\lambda)$ 的 k 阶子式的公因式，从而 $f(\lambda)\mid g(\lambda)$.

对于列变换，可以完全一样地讨论. 总之，如果 $A(\lambda)$ 经过一次初等变换变成 $B(\lambda)$，那么 $f(\lambda)\mid g(\lambda)$. 但由初等变换的可逆性，$B(\lambda)$ 也可以经过一次初等变换变成 $A(\lambda)$. 由上面的讨论，同样应有 $g(\lambda)\mid f(\lambda)$，于是 $f(\lambda)=g(\lambda)$.

当 $A(\lambda)$ 的全部 k 阶子式为零时，$B(\lambda)$ 的全部 k 阶子式也就等于零；反之亦然. 因此，$A(\lambda)$ 与 $B(\lambda)$ 有相同的各阶行列式因子和秩. ▏

现在来计算标准形矩阵的行列式因子. 设标准形为

$$\begin{bmatrix} d_1(\lambda) & & & & \\ & d_2(\lambda) & & O & \\ & & \ddots & & \\ & & & d_r(\lambda) & \\ & O & & & O \end{bmatrix}, \tag{1}$$

其中 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 是首项系数为 1 的多项式，且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ $(i=1,2,\cdots,r-1)$. 不难证明，在这种形式的矩阵中，如果一个 k 阶子式包含的行与列的标号不完全相同，那么这个 k 阶子式一定为零. 因此，为了计算 k 阶行列式因子，只要看由 $i_1, i_2, \cdots, i_k$ 行与 $i_1, i_2, \cdots, i_k$ 列 $(1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant r)$ 组成的 k 阶子式就行了，而这个 k 阶子式等于

$$d_{i_1}(\lambda) d_{i_2}(\lambda) \cdots d_{i_k}(\lambda).$$

显然，这些 k 阶子式的最大公因式就是

$$D_k(\lambda) = d_1(\lambda) d_2(\lambda) \cdots d_k(\lambda).$$

定理 11 λ-矩阵的标准形是唯一的.

证明：设(1)是 $A(\lambda)$ 的标准形. 由于 $A(\lambda)$ 与(1)等价，它们有相同的秩与相同的行列式因子，因此 $A(\lambda)$ 的秩就是标准形的主对角线上非零元素的个数 r；$A(\lambda)$ 的 k 阶行列式因子就是

$$D_k(\lambda) = d_1(\lambda) d_2(\lambda) \cdots d_k(\lambda) \, (k = 1, 2, \cdots, r). \tag{2}$$

于是

$$d_1(\lambda) = D_1(\lambda), d_2(\lambda) = \frac{D_2(\lambda)}{D_1(\lambda)}, \cdots, d_r(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}. \tag{3}$$

这说明 $A(\lambda)$ 的标准形(1)的主对角线上的非零元素是被 $A(\lambda)$ 的行列式因子所唯一决定的，所以 $A(\lambda)$ 的标准形是唯一的. ▏

定义 8　$A(\lambda)$的标准形的主对角线上非零元素 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$称为 λ-矩阵 $A(\lambda)$的**不变因子**.

作为例子，我们指出：n 阶单位矩阵的行列式因子及不变因子全都等于 1. 而例 1 中的 $A(\lambda)$的不变因子为：$1, \lambda, \lambda(\lambda+1)$.

定理 12　两个 $m\times n$ λ-矩阵等价的充分必要条件是它们有相同的行列式因子，或者，它们有相同的不变因子.

证明：等式(2)与(3)给出了 λ-矩阵的行列式因子与不变因子之间的关系. 这些关系式说明，行列式因子与不变因子是相互确定的. 因此，说两个矩阵有相同的各阶行列式因子，就等于说它们有相同的各阶不变因子.

必要性已由定理 10 证明.

充分性是很明显的. 事实上，若 λ-矩阵 $A(\lambda)$与$B(\lambda)$有相同的不变因子，则 $A(\lambda)$与 $B(\lambda)$和同一个标准形等价，因而 $A(\lambda)$与 $B(\lambda)$等价. ▌

由(3)式可以看出，在 λ-矩阵的行列式因子之间，有关系

$$D_k(\lambda) \mid D_{k+1}(\lambda) \quad (k = 1, 2, \cdots, r-1). \tag{4}$$

在计算 λ-矩阵的行列式因子时，常常是先计算最高阶的行列式因子. 这样，由(4)式就大致有了低阶行列式因子的范围了.

习　题　7.4

求下列 λ-矩阵 $A(\lambda)$的行列式因子及不变因子：

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

2. $A(\lambda)=\begin{bmatrix} \lambda & -1 & 0 & 0 \\ 0 & \lambda & -1 & 0 \\ 0 & 0 & \lambda & -1 \\ 5 & 4 & 3 & \lambda+2 \end{bmatrix}$.

3. $A(\lambda)=\begin{bmatrix} \lambda+a & b & 1 & 0 \\ -b & \lambda+a & 0 & 1 \\ 0 & 0 & \lambda+a & b \\ 0 & 0 & -b & \lambda+a \end{bmatrix}$.

7.5 初等因子

为了讨论数字矩阵在相似关系下的标准形，还需要用到初等因子的概念.

定义 9 把数域 **P** 上 λ-矩阵 $A(\lambda)$ 的次数大于 1 的不变因式分解成标准分解式，即分解成数域 **P** 上首项系数为 1 的互不相同的不可约多项式的方幂的乘积，所有这些不可约多项式的方幂（相同的必须按出现的次数计算）称为 $A(\lambda)$ 的**初等因子**.

例 1 设 $A(\lambda)$ 的等价标准形为：

$$\begin{bmatrix} 1 & & & & \\ & \lambda^2(\lambda+1) & & O & \\ & & \lambda^2(\lambda+1)^3(\lambda^2+4)^2 & & \\ & O & & & O \end{bmatrix}$$

如果在实数域中讨论，则 $A(\lambda)$ 的初等因子为：

$$\lambda^2, \lambda^2, \lambda+1, (\lambda+1)^3, (\lambda^2+4)^2.$$

如果在复数域中讨论，则 $A(\lambda)$ 的初等因子为：

$$\lambda^2, \lambda^2, \lambda+1, (\lambda+1)^3, (\lambda+2\mathrm{i})^2, (\lambda-2\mathrm{i})^2.$$

这个例子说明，与行列式因子，不变因子不同，初等因子与系数域 **P** 有关.

下面进一步说明不变因子和初等因子的关系. 首先，假设 $A(\lambda)$ 的秩及不变因子 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 为已知，为了讨论方便起见，把它们统一地表成不同的首项系数为 1 的不可约多项式的方幂的乘积：

$$d_1(\lambda) = (e_1(\lambda))^{k_{11}}(e_2(\lambda))^{k_{12}}\cdots(e_t(\lambda))^{k_{1t}}, k_{11}, k_{12}, \cdots, k_{1t} \geqslant 0;$$

$$d_2(\lambda) = (e_1(\lambda))^{k_{21}}(e_2(\lambda))^{k_{22}}\cdots(e_t(\lambda))^{k_{2t}}, k_{21}, k_{22}, \cdots, k_{2t} \geqslant 0;$$

$\cdots\cdots\cdots\cdots\cdots\cdots$

$$d_r(\lambda) = (e_1(\lambda))^{k_{r1}}(e_2(\lambda))^{k_{r2}}\cdots(e_t(\lambda))^{k_{rt}}, k_{r1}, k_{r2}, \cdots, k_{rt} \geqslant 0.$$

则其中对应于 $k_{ij}>1$ 的那些方幂

$$(e_i(\lambda))^{k_{ij}}, \quad k_{ij} > 0$$

就是 $A(\lambda)$的全部初等因子，根据不变因子有一个整除一个的性质：

$$d_i(\lambda) | d_{i+1}(\lambda), \quad i = 1, 2, \cdots, r-1.$$

因此

$$(e_{ij}(\lambda))^{k_{ij}} | (e_{ij}(\lambda))^{k_{i+1,j}}, \quad i = 1, 2, \cdots, r-1;\ j = 1, 2, \cdots, t.$$

因此，在 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$的分解式中，属于同一个不可约因式的幂指数有递升的性质，即

$$k_{1j} \leqslant k_{2j} \leqslant \cdots \leqslant k_{rj}, \quad j = 1, 2, \cdots, t.$$

这说明，同一个不可约因式的方幂作成的初等因子中，幂次最高的一定出现在 $d_r(\lambda)$的分解式中，幂次次高的必定出现在 $d_{r-1}(\lambda)$的分解式中，如此顺推下去，可知属于同一个不可约因式的方幂所成的初等因子在不变因子中出现的位置是唯一确定的. 因此，不变因子也由初等因子及秩唯一确定.

上面的分析还给出了如何从 λ-矩阵的初等因子及秩作出不变因子的方法. 设 n 阶矩阵 $A(\lambda)$的秩为 r，$A(\lambda)$的全部初等因子为已知，在全部初等因子中将同一个不可约因式的方幂的那些初等因子按降幂排列. 如果这些初等因子不足 r 个，就在后面补上一些 1，使凑成 r 个. 设所得的初等因子排列为：

$$(e_j(\lambda))^{k_{rj}}, (e_j(\lambda))^{k_{r-1,j}}, \cdots, (e_j(\lambda))^{k_{1j}}, \quad j = 1, 2, \cdots, t$$

令

$$d_i(\lambda) = (e_1(\lambda))^{k_{i1}}(e_2(\lambda))^{k_{i2}}\cdots(e_t(\lambda))^{k_{it}}, \quad i = 1, 2, \cdots, r,$$

则 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$就是 $A(\lambda)$的不变因子.

以上讨论说明：如果两个 λ-矩阵有相同的初等因子及秩，那么它们就有相同的不变因子，因此它们等价；反之，如果两个 λ-矩阵等价，则它们有相同的不变因子，因此有相同的初等因子. 即有：

定理 13 两个 $m\times n$ λ-矩阵 $A(\lambda)$，$B(\lambda)$等价的充分必要条件是它们有相同的秩及初等因子.

例 2 设 $A(\lambda)$是一个 4×5 实系数 λ-矩阵，已知 $A(\lambda)$的秩为3，全部初等因子为：

$$\lambda,\lambda+1,(\lambda+1)^2,(\lambda^2+1)^2,(\lambda^2+1)^3,$$

求 $A(\lambda)$的标准形.

解：$A(\lambda)$的不变因子为：

$$d_3(\lambda)=\lambda(\lambda+1)^2(\lambda^2+1)^3,$$
$$d_2(\lambda)=(\lambda+1)(\lambda^2+1)^2,$$
$$d_1(\lambda)=1,$$

所以 $A(\lambda)$的标准形为：

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & (\lambda+1)(\lambda^2+1)^2 & 0 & 0 & 0 \\ 0 & 0 & \lambda(\lambda+1)^2(\lambda+1)^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

初等因子、不变因子及行列式因子都是 λ-矩阵的等价不变量，但是初等因子有比不变因子及行列式因子有较为方便的计算方法.

定理 14 设 $A(\lambda)$是数域 $\mathbf{P}$ 上的一个秩为 r 的对角形 λ-矩阵（不必为标准形）：

$$\begin{bmatrix} g_1(\lambda) & & & & \\ & g_2(\lambda) & & O & \\ & & \ddots & & \\ & & & g_r(\lambda) & \\ O & & & & O \end{bmatrix}.$$

其中 $g_i(\lambda)$ 的首项系数为 $1, i=1,2,\cdots,r$. 将 $g_i(\lambda)$ 分解成 $\mathbf{P}$ 上首项系数为 1 的不可约多项式的方幂的乘积：

$$g_i(\lambda) = (e_1(\lambda))^{k_{i1}}(e_2(\lambda))^{k_{i2}}\cdots(e_t(\lambda))^{k_{it}}$$

$$k_{ij} \geqslant 0, \quad i = 1,2,\cdots,r, \quad j = 1,2,\cdots,t,$$

则所有

$$(e_j(\lambda))^{k_{ij}}, \quad i = 1,2,\cdots,r, \ j = 1,2,\cdots,t,$$

中那些 $k_{ij}>0$ 的方幂就是 $A(\lambda)$ 的全部初等因子.

证明：首先计算 $A(\lambda)$ 的 k 阶行列式因子，$A(\lambda)$ 的不等于零的 l 阶子式一定是主子式，而且由 $g_1(\lambda), g_2(\lambda), \cdots, g_r(\lambda)$ 中的 l 个相乘而得：

$$g_{i_1}(\lambda)g_{i_2}(\lambda)\cdots g_{i_l}(\lambda) = (e_1(\lambda))^{k_{i_11}+k_{i_21}+\cdots k_{i_l1}}\cdot(e_2(\lambda))^{k_{i_12}+k_{i_22}+\cdots+k_{i_l2}}$$
$$\cdots(e_t(\lambda))^{k_{i_1t}+k_{i_2t}+\cdots+k_{i_lt}}$$

将 $k_{1j}, k_{2j}, \cdots, k_{rj}(r=1,2,\cdots,t)$ 按由小到大的次序排好：

$$k_{p_{j1}j} \leqslant k_{p_{j2}j} \leqslant \cdots \leqslant k_{p_{jr}j},$$

其中 $p_{j1}, p_{j2}, \cdots, p_{jr}$ 是 $1,2,\cdots,r$ 的一个排列，则

$$\begin{aligned} D_l(\lambda) &= (e_1(\lambda))^{k_{p_{11}1}+k_{p_{12}1}+\cdots+k_{p_{1l}1}} \\ &\cdot (e_2(\lambda))^{k_{p_{21}2}+k_{p_{22}2}+\cdots+k_{p_{2l}2}} \\ &\cdots\cdots \\ &\cdot (e_t(\lambda))^{k_{p_{t1}t}+k_{p_{t2}t}+\cdots+k_{p_{tl}t}}, \ l = 1,2,\cdots,r, \end{aligned}$$

由

$$d_1(\lambda) = D_1(\lambda), d_i(\lambda) = \frac{D_{i1}(\lambda)}{D_{i-1}(\lambda)}, \ i = 2,\cdots,r,$$

即得

$$\begin{aligned} d_1(\lambda) &= (e_1(\lambda))^{k_{p_{11}1}}(e_2(\lambda))^{k_{p_{21}2}}\cdots(e_t(\lambda))^{k_{p_{t1}t}}, \\ d_2(\lambda) &= (e_1(\lambda))^{k_{p_{12}1}}(e_2(\lambda))^{k_{p_{22}2}}\cdots(e_t(\lambda))^{k_{p_{t2}}t}, \\ &\cdots\cdots\cdots \\ d_r(\lambda) &= (e_1(\lambda))^{k_{p_{1r}1}}(e(\lambda))^{k_{p_{2r}2}}\cdots(e_t(\lambda))^{k_{p_{tr}}t}. \end{aligned}$$

因此 $A(\lambda)$的全部初等因子就是

$$(e_j(\lambda))^{k_{ij}} \quad i=1,2,\cdots r, \quad j=1,2,\cdots,t,$$

中不等于 1 的那些方幂. ▎

根据定理 14，只要把一个 λ-矩阵 $A(\lambda)$用初等变换化为对角形(不必化为标准形)就可以求出 $A(\lambda)$的全部初等因子，从而求出 $A(\lambda)$的标准形.

推论 设 $A(\lambda)$是一个准对角矩阵：

$$A=\begin{bmatrix} A_1(\lambda) & & & \\ & A_2(\lambda) & & \\ & & \ddots & \\ & & & A_t(\lambda) \end{bmatrix},$$

则 $A_1(\lambda),A_2(\lambda),\cdots,A_t(\lambda)$的全部初等因子合起来就是 $A(\lambda)$的全部初等因子.

证明：应用定理 14，只要将 $A_1(\lambda),A_2(\lambda),\cdots,A_t(\lambda)$分别化为对角形即可.

例 3 求实系数矩阵

$$A(\lambda)=\begin{bmatrix} \lambda-1 & \lambda^2 & 0 & 0 \\ 2\lambda & \lambda-1 & 0 & 0 \\ 0 & 0 & \lambda^2+3 & 0 \\ 0 & 0 & 0 & \lambda^2(\lambda^2+1) \end{bmatrix}$$

的全部初等因子及标准形

解：因为

$$\begin{bmatrix} \lambda-1 & \lambda^2 \\ 2\lambda & \lambda-1 \end{bmatrix} \to \begin{bmatrix} 1 & \\ (2\lambda-1) & (\lambda^2+1) \end{bmatrix},$$

所以 $A(\lambda)$的全部初等因子为：

$$2\lambda-1, \quad \lambda^2+1, \quad \lambda^2+3, \quad \lambda^2, \quad \lambda^2+1.$$

$A(\lambda)$的标准形为：

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \lambda^2+1 & \\ & & & \lambda^2(2\lambda-1)(\lambda^2+1)(\lambda^2+3) \end{bmatrix}.$$

习 题 7.5

求下列 λ-矩阵的初等因子及标准形：

1. $\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}$.

2. $\begin{bmatrix} 0 & 0 & 0 & \lambda^2 \\ 0 & 0 & \lambda(\lambda-1) & 0 \\ 0 & (\lambda-1)^2 & 0 & 0 \\ \lambda(\lambda-1) & 0 & 0 & 0 \end{bmatrix}$.

3. $\begin{bmatrix} \lambda^2+2 & \lambda^2+1 & \lambda^2+1 \\ 3 & \lambda^2+1 & 3 \\ \lambda^2+1 & \lambda^2+1 & \lambda^2+1 \end{bmatrix}$（分有理数域实数域，复数域 3 种情形）

7.6 矩阵相似的条件

在计算数字矩阵 A 的特征值与特征向量时，曾介绍过 A 的特征矩阵$\lambda E-A$. 这是一类特殊的 λ-矩阵. 这一节主要证明两个 n 阶数字矩阵 A 和 B 相似的充分必要条件是它们的特征矩阵 $\lambda E-A$ 和 $\lambda E-B$ 等价.

命题 1 如果有 n 阶数字矩阵 P_0,Q_0 使

$$\lambda E - A = P_0(\lambda E - B)Q_0, \tag{1}$$

则 A 与 B 相似.

证明： 因 $P_0(\lambda E-B)Q_0=\lambda P_0Q_0-P_0BQ_0$，它又与 $\lambda E-A$ 相

等，进行比较后应有 $P_0Q_0=E, P_0BQ_0=A$. 由此 P_0 可逆，$Q_0=P_0^{-1}$，而 $A=P_0BP_0^{-1}$，故 A 与 B 相似.

命题 2 对于任何不为零的 n 阶数字矩阵 A 和 λ-矩阵 $U(\lambda)$ 与 $V(\lambda)$，一定存在 λ-矩阵 $Q(\lambda)$ 与 $R(\lambda)$ 以及数字矩阵 U_0 和 V_0，使

$$U(\lambda)=(\lambda E-A)Q(\lambda)+U_0, \tag{2}$$

$$V(\lambda)=R(\lambda)(\lambda E-A)+V_0. \tag{3}$$

证明：把 $U(\lambda)$ 改写成：

$$U(\lambda)=D_0\lambda^m+D_1\lambda^{m-1}+\cdots+D_{m-1}\lambda+D_m.$$

这里 $D_0,D_1,\cdots,D_m$ 都是 n 阶数字矩阵，而且 $D_0\neq 0$. 如 $m=0$，则令 $Q(\lambda)=O$ 及 $U_0=D_0$，它们显然满足命题 2 的要求.

设 $m>0$，令

$$Q(\lambda)=Q_0\lambda^{m-1}+Q_1\lambda^{m-2}+\cdots+Q_{m-2}\lambda+Q_{m-1}.$$

这里 Q_j 都是待定的数字矩阵. 于是

$$\begin{aligned}(\lambda E-A)Q(\lambda)=&Q_0\lambda^m+(Q_1-AQ_0)\lambda^{m-1}+\cdots\\&+(Q_k-AQ_{k-1})\lambda^{m-k}+\cdots\\&+(Q_{m-1}-AQ_{m-2})\lambda-AQ_{m-1}.\end{aligned}$$

要想使(2)式成立，只需取

$$\begin{aligned}&Q_0=D_0,\\&Q_1=D_1+AQ_0,\\&Q_2=D_2+AQ_1,\\&\cdots\cdots\cdots\cdots\cdots\\&Q_k=D_k+AQ_{k-1},\\&\cdots\cdots\cdots\cdots\cdots\\&Q_{m-1}=D_{m-1}+AQ_{m-2},\\&U_0=D_m+AQ_{m-1}\end{aligned}$$

就行了. 用完全相同的办法可以求得 $R(\lambda)$ 和 V_0. 命题证毕.

定理 15 设 A,B 是数域 $\mathbf{P}$ 上两上 n 阶矩阵，A 与 B 相似的

充分必要条件是它们的特征矩阵 $\lambda E-A$ 和 $\lambda E-B$ 等价.

证明：由定理 9 知道 $\lambda E-A$ 与 $\lambda E-B$ 等价就是有可逆的 λ-矩阵$U(\lambda)$和 $V(\lambda)$，使

$$\lambda E - A = U(\lambda)(\lambda E - B)V(\lambda). \tag{4}$$

先证必要性. 设 A 与 B 相似，即有可逆矩阵 T，使

$$A = T^{-1}BT.$$

于是

$$\lambda E - A = \lambda E - T^{-1}BT = T^{-1}(\lambda E - B)T,$$

从而 $\lambda E-A$ 与 $\lambda E-B$ 等价.

再证充分性. 设 $\lambda E-A$ 与 $\lambda E-B$ 等价，即有可逆的 λ-矩阵 $U(\lambda)$，$V(\lambda)$使(4)式成立. 用命题 2，存在 λ-矩阵 $Q(\lambda)$和 $R(\lambda)$以及数字矩阵 U_0 和 V_0 使

$$U(\lambda) = (\lambda E - A)Q(\lambda) + U_0, \tag{5}$$

$$V(\lambda) = R(\lambda)(\lambda E - A) + V_0 \tag{6}$$

成立. 把(4)式改写成

$$U(\lambda)^{-1}(\lambda E - A) = (\lambda E - B)V(\lambda),$$

式中的 $V(\lambda)$用(6)式代入，再移项，得

$$[U(\lambda)^{-1} - (\lambda E - B)R(\lambda)](\lambda E - A) = (\lambda E - B)V_0.$$

右端次数等于 1 或 $V_0=0$，因此 $U(\lambda)^{-1}-(\lambda E-B)R(\lambda)$是一个数字矩阵(后一情形下应是零矩阵)，记作 T，即

$$T = U(\lambda)^{-1} - (\lambda E - B)R(\lambda),$$

$$T(\lambda E - A) = (\lambda E - B)V_0. \tag{7}$$

现在我们来证明 T 是可逆的. 由(7)式的第 1 式有

$$\begin{aligned}
E &= U(\lambda)T + U(\lambda)(\lambda E - B)R(\lambda) \\
&= U(\lambda)T + (\lambda E - A)V(\lambda)^{-1}R(\lambda) \\
&= [(\lambda E - A)Q(\lambda) + U_0]T \\
&\quad + (\lambda E - A)V(\lambda)^{-1}R(\lambda) \\
&= U_0T + (\lambda E - A)[Q(\lambda)T + V(\lambda)^{-1}R(\lambda)].
\end{aligned}$$

等式右端的第 2 项必须为零，否则它的次数至少是 1，由于 E 和 U_0T 都是数字矩阵，等式不可能成立，因此

$$E = U_0T.$$

这就是说，T 是可逆的. 由(7)式的第 2 式得：

$$\lambda E - A = T^{-1}(\lambda E - B)V_0.$$

再用命题 1，A 与 B 相似. ▎

矩阵 A 的特征矩阵 $\lambda E - A$ 的不变因子、行列式因子及初等因子以后就简称为 A 的不变因子、行列式因子及初等因子. 因为两个 λ-矩阵等价的充分必要条件是它们有相同的不变因子，所以由定理即得：

推论 矩阵 A 与 B 相似的充分必要条件是：(1) 它们有相同的不变因子；或(2)它们有相同的行列式因子；或(3)它们有相同的初等因子.

应该指出 n 阶矩阵的特征矩阵的秩一定是 n. 因此，n 阶矩阵的不变因子总是有 n 个，并且，它们的乘积就等于这个矩阵的特征多项式.

例 1 讨论

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & -3 & 6 \\ 2 & -2 & 4 \end{bmatrix}$$

能否对角化

解：

$$\lambda E - A = \begin{vmatrix} \lambda - 1 & 1 & -2 \\ -3 & \lambda + 3 & -6 \\ -2 & 2 & \lambda - 4 \end{vmatrix} = \lambda^2(\lambda - 2).$$

如果 A 能对角化，则 A 与

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

相似.因为 $\lambda E-A$ 的不变因子与 $\lambda E-B$ 的不变因子都等于 $1,\lambda,\lambda(\lambda-2)$,所以 $\lambda E-A$ 与 $\lambda E-B$ 等价,因此 A 与 B 相似,即 A 可对角化.

习 题 7.5

1. 设 A 是一个 n 阶数字矩阵.证明:A 与 A^{T} 相似.

判断下列数字矩阵 A 能否对角化(2—4):

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

3. $A=\begin{bmatrix}13&16&16\\-5&-7&-6\\-6&-8&-7\end{bmatrix}$.

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

5. 证明:矩阵 A 的全部初等因子的乘积等于 A 的特征多项式.

7.7 约当标准形

这一节应用初等因子的理论来解决约当标准形的存在及计算问题.并且给出数字矩阵可对角化的另一些条件.

这一节的讨论都是在复数域中进行的.

因为复系数不可约多项式都是 1 次多项式,所以复系数矩阵的初等因子都是 1 次因式的方幂.

首先计算约当形的初等因子.先对约当块进行计算.设

$$J_0 = \begin{bmatrix} \lambda_0 & 1 & & & \\ & \lambda_0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_0 \end{bmatrix}$$

是一个 k 阶约当块，它的特征矩阵为：

$$\lambda E - J_0 = \begin{bmatrix} \lambda - \lambda_0 & -1 & & & \\ & \lambda - \lambda_0 & -1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & -1 \\ & & & & \lambda - \lambda_0 \end{bmatrix}$$

容易看出 $|\lambda E - J_0| = (\lambda - \lambda_0)^k$. 这就是 $\lambda E - J_0$ 的 k 阶行列式因子. 由于 $\lambda E - J_0$ 有一个 $k-1$ 阶子式.

$$\begin{vmatrix} -1 & & & & \\ \lambda - \lambda_0 & -1 & & & \\ & \lambda - \lambda_0 & -1 & & \\ & & \ddots & \ddots & \\ & & & \lambda - \lambda_0 & -1 \end{vmatrix} = (-1)^{k-1},$$

所以它的 $k-1$ 阶行列式等于 1，从而，它的以下各阶行列式因子全等于 1. 于是，它的不变因子是：

$$d_1(\lambda) = \cdots = d_{k-1}(\lambda) = 1,$$

$$d_k(\lambda) = (\lambda - \lambda_0)^k,$$

由此得出 $\lambda E - J_0$ 的初等因子是 $(\lambda - \lambda_0)^k$.

根据定理 14 的推论，可以很容易地算出约当形的初等因子. 设

$$J=\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}$$

是一个约当形矩阵，其中 J_i 是一个 k_i 阶约当块：

$$J_i=\begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix},\quad i=1,2,\cdots,s,$$

那么，J_i 的初等因子就是$(\lambda-\lambda_i)^{k_i}$. 于是，$J$ 的全部初等因子就是：

$$(\lambda-\lambda_1)^{k_1},\quad (\lambda-\lambda_2)^{k_2},\cdots,(\lambda-\lambda_s)^{k_s}.$$

由上面的讨论，可看出，约当块被它的初等因子唯一确定，而约当形矩阵除去其中约当块的次序外，被它的初等因子唯一决定.

定理 16 每个 n 阶复矩阵 A 都与一个 n 阶约当形矩阵相似. 这个约当形矩阵除去其中约当块的排列次序外是被矩阵 A 唯一决定的，称为 A 的**约当标准形**.

证明： 设 n 阶矩阵 A 的初等因子为：

$$(\lambda-\lambda_1)^{k_1},\ (\lambda-\lambda_2)^{k_2},\cdots,(\lambda-\lambda_s)^{k_s} \tag{1}$$

其中 $k_i>0, i=1,2,\cdots,s$. 对应于每个初等因子$(\lambda-\lambda_i)^{k_i}$，作一个 k_i 阶约当块：

$$J_i=\begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix},\quad i=1,2,\cdots,s.$$

将这些约当块构成一个约当形矩阵：

$$J=\begin{bmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{bmatrix}$$

则 J 的全部初等因子也是式(1). 因为 J 与 A 有相同的初等因子，所以它们相似.

如果有另一个约当形矩阵 $\tilde{J}$ 与 A 相似，那么，$\tilde{J}$ 与 A，因此 $\tilde{J}$ 与 J 有相同的初等因子，所以 $\tilde{J}$ 与 J 除了其中约当块的排列次序外是相同的. 这样就证明了唯一性. ▌

定理 16 的证明给出了求矩阵的约当标准形的方法.

例 1 求矩阵

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

的约当形.

解：首先求 A 的初等因子：

$$E-A=\begin{bmatrix}\lambda+1 & 2 & -6\\ 1 & \lambda & -3\\ 1 & 1 & \lambda-4\end{bmatrix}\to\begin{bmatrix}0 & -\lambda+1 & -\lambda^2+3\lambda-2\\ 0 & \lambda-1 & -\lambda+1\\ 1 & 1 & \lambda-4\end{bmatrix}$$

$$\to\begin{bmatrix}0 & 0 & -\lambda^2+2\lambda-1\\ 0 & \lambda-1 & -\lambda+1\\ 1 & 1 & \lambda-4\end{bmatrix}$$

$$\to\begin{bmatrix}0 & 0 & -\lambda^2+2\lambda-1\\ 0 & \lambda-1 & -\lambda+1\\ 1 & 0 & 0\end{bmatrix}$$

$$\to\begin{bmatrix}0 & 0 & -\lambda^2+2\lambda-1\\ 0 & \lambda-1 & 0\\ 1 & 0 & 0\end{bmatrix}\to\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda-1 & 0\\ 0 & 0 & (\lambda-1)^2\end{bmatrix}$$

所以，A 的初等因子是 $\lambda-1,(\lambda-1)^2$，A 的约当标准形是

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}.$$

例 2 求矩阵

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

的约当标准形.

解：A 是一个准对角形矩阵：

$$A=\begin{bmatrix}A_1 & \\ & A_2\end{bmatrix}$$

其中

$$A_1=\begin{bmatrix}0 & -1\\ 1 & 2\end{bmatrix},\quad A_2=\begin{bmatrix}-3 & 8\\ -2 & 5\end{bmatrix}.$$

分别求 A_1,A_2 的初等因子.

$$\lambda E-A_1=\begin{bmatrix}\lambda & 1\\ -1 & \lambda-2\end{bmatrix}\to\begin{bmatrix}0 & \lambda^2-2\lambda+1\\ -1 & \lambda-2\end{bmatrix}$$

$$\to\begin{bmatrix}1 & 0\\ 0 & (\lambda-1)^2\end{bmatrix}.$$

所以 $\lambda E-A_1$ 的初等因子是 $(\lambda-1)^2$.

$$\lambda E-A_2=\begin{bmatrix}\lambda+3 & -8\\ 2 & \lambda-5\end{bmatrix}$$

容易看出 $\lambda E-A_2$ 的 1 阶行列式因子 $D_1(\lambda)=1$，而 $D_2(\lambda)=|\lambda E-A_2|=\lambda^2-2\lambda+1=(\lambda-1)^2$，所以 $\lambda E-A_2$ 的不变因式为 $d_1(\lambda)=1$，$d_2(\lambda)=(\lambda-1)^2$，初等因子为 $(\lambda-1)^2$.

于是得 A 的初等因子为 $(\lambda-1)^2,(\lambda-1)^2$. A 的约当标准

形为：

$$\begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}.$$

约当标准形包括对角矩阵为特殊情形.就是由 1 阶约当块构成的约当形矩阵.因为 1 阶约当块对应于 1 次初等因子,因此可得到矩阵可对角化的另一些条件.

定理 17 复系数矩阵 A 与对角矩阵相似的充分必要条件是：A 的初等因子全是 1 次的.

因为初等因子是由不变因子分解成不同一次因式的方幂而得,因此有：

定理 18 复系数矩阵与对角矩阵相似的充分必要条件是：A 的不变因子都没有重根.

又因为不变因子有一个整除一个的性质：

$$d_i(\lambda)\,|\,d_{i+1}(\lambda), \quad i = 1,2,\cdots,n-1.$$

因此定理 18 中的条件只要对 $d_n(\lambda)$ 成立就可以了.而根据约当标准形的作法,可以看出矩阵 A 的最后一个不变因子 $d_n(\lambda)$ 就是 A 的最小多项式.

定理 19 复系数矩阵 A 可对角化的充分必要条件是 A 的最小多项式没有重根.

在这一节我们证明了每个矩阵 A 都与一个约当形矩阵相似,并且介绍了具体求 A 的约当标准形的方法.但是并没有讲到如何求可逆矩阵 T 使 $T^{-1}AT$ 成的约当标准形的问题.由于 T 的确定涉及到比较复杂的计算问题,这里就不讨论了.

如果规定下三角矩阵：

$$\begin{bmatrix} \lambda_0 & & & & \\ 1 & \lambda_0 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda_0 \end{bmatrix}$$

为约当块，所有的讨论也是类似的.读者在应用的时候，可以灵活地选择合适的约当形矩阵.

习 题 7.7

求下列矩阵 A 的约当标准形：

1. $A=\begin{bmatrix} -1 & 1 & 1 \\ -5 & 21 & 17 \\ 6 & -26 & -21 \end{bmatrix}$.

2. $A=\begin{bmatrix} 8 & -3 & 6 \\ 3 & -2 & 0 \\ -4 & 2 & -2 \end{bmatrix}$.

3. $A=\begin{bmatrix} -7 & -12 & -6 \\ 3 & 5 & 3 \\ 3 & 6 & 2 \end{bmatrix}$.

4. $A=\begin{bmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 2 \end{bmatrix}$.

内 容 提 要

1. λ-矩阵

(1) λ-矩阵的秩的定义

(2) λ-矩阵可逆及逆矩阵的定义

λ-矩阵 $A(\lambda)$可逆的充分必要条件是$|A(\lambda)|$是一个非零常数，

而且当 $A(\lambda)$ 可逆时 $A^{-1}(\lambda)=\frac{1}{|A(\lambda)|}A^*(\lambda)$.

2. 最小多项式

(1) 哈密尔顿-凯莱定理

(2) 最小多项式

1) 定义

2) 一些性质

① A 的最小多项式是 A 的特征多项式的一个因式.

② 相似矩阵有相同的最小多项式.

③ A 的特征值都是最小多项式的根.

④ A 的特征多项式与最小多项式有相同的不可约因式.

⑤ k 阶约当块

$$J_0=\begin{bmatrix}\lambda_0 & 1 & & & \\ & \lambda_0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_0\end{bmatrix}$$

的最小多项式为 $(\lambda-\lambda_0)^k$.

⑥ 准对角矩阵

$$A=\begin{bmatrix}A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s\end{bmatrix}$$

的最小多项式是 $A_i(i=1,\cdots,s)$ 的最小多项式 $f_i(\lambda)$ 的最小公倍式 $[f_1(\lambda),f_2(\lambda),\cdots,f_s(\lambda)]$.

(3) 数字矩阵 A 可对角化的一些条件

1) 数域 **P** 上矩阵 A 可对角化的充分必要条件是:A 的最小多项式是数域 **P** 上一些互素的一次因式的乘积.

2）复系数矩阵 A 可对角化的充分必要条件是：A 的最小多项式没有重根.

3. λ-矩阵的标准形

（1）基本概念

1）λ-矩阵的初等变换

2）λ-矩阵的等价

3）λ-矩阵的标准形

4）λ-矩阵的行列式因子、不变因子、初等因子以及它们之间的关系.

（2）λ-矩阵等价的一些条件

两个 $m\times n$ λ-矩阵 $A(\lambda)$与 $B(\lambda)$等价

$\rightleftharpoons$ 可找到可逆的 λ-矩阵 $P(\lambda)$及 $Q(\lambda)$使 $B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$

$\rightleftharpoons$ $A(\lambda)$与 $B(\lambda)$有相同的标准形

$\rightleftharpoons$ $A(\lambda)$与 $B(\lambda)$有相同的不变因子

$\rightleftharpoons$ $A(\lambda)$与 $B(\lambda)$有相同的行列式因子

$\rightleftharpoons$ $A(\lambda)$与 $B(\lambda)$有相同的初等因子及秩

4. 约当标准形

（1）n 阶数字矩阵 A 与 B 相似的充分必要条件是：它们的特征矩阵 $\lambda E-A$ 与 $\lambda E-B$ 等价.

A 的特征矩阵的不变因子、行列式因子及初等因子称为 A 的不变因子、行列式因子及初等因子.

（2）每个复系数矩阵 A 都与一个约当形矩阵相似，而且这个约当形矩阵除去其中约当块的排列次序外被矩阵 A 唯一确定. 此称为 A 的约当标准形.

（3）n 阶复矩阵 A 的约当标准形的求法

1）求出 A 的全部初等因子，设为

$$(\lambda-\lambda_1)^{k_1},(\lambda-\lambda_2)^{k_2},\cdots,(\lambda-\lambda_s)^{k_s}$$

$$k_i > 0, i = 1, \cdots, s;\ k_1 + k_2 + \cdots + k_s = n.$$

2）对每个初等因子$(\lambda-\lambda_i)^{k_i}(i=1,\cdots,s)$，作一个 k_i 阶约当块：

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

3）用约当块 $J_1, J_2, \cdots, J_s$ 排成一个准对角矩阵

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix},$$

那么 J 就是 A 的约当标准形.

(4) 复系数矩阵 A 可对角化的另一些充分必要条件

1）A 的初等因子都是 1 次的.

2）A 的不变因子都没有重根.

复习题 7

1. 求下列 λ-矩阵的标准形：

(1) $\begin{bmatrix} 3\lambda^2+2\lambda-3 & 2\lambda-1 & \lambda^2+2\lambda-3 \\ 4\lambda^2+3\lambda-5 & 3\lambda-2 & \lambda^2+3\lambda-4 \\ \lambda^2+\lambda-4 & \lambda-2 & \lambda-1 \end{bmatrix}$；

(2) $\begin{bmatrix} 2\lambda & 3 & 0 & 1 & \lambda \\ 4\lambda & 3\lambda & 0 & \lambda+2 & 2\lambda \\ 0 & 6\lambda & \lambda & 2\lambda & 0 \\ \lambda-1 & 0 & \lambda-1 & 0 & 0 \\ 3\lambda-3 & 1-\lambda & 2\lambda-2 & 0 & 0 \end{bmatrix}$.

2. 求下列 λ-矩阵的不变因子、初等因子(分复数域及实数域两种情形)及秩：

(1) $\begin{bmatrix} 0 & 0 & 1 & \lambda+2 \\ 0 & 1 & \lambda+2 & 0 \\ 1 & \lambda+2 & 0 & 0 \\ \lambda+2 & 0 & 0 & 0 \end{bmatrix}$； (2) $\begin{bmatrix} \lambda+1 & 0 & 0 & 0 \\ 0 & \lambda+2 & 0 & 0 \\ 0 & 0 & \lambda+1 & 0 \\ 0 & 0 & 0 & \lambda+2 \end{bmatrix}$；

(3) $\begin{bmatrix} \lambda & 2\lambda+1 & \lambda^2 \\ 1 & \lambda+1 & \lambda^2-1 \\ \lambda-1 & \lambda & 1 \end{bmatrix}$.

3. 求下列复系数矩阵的约当标准形：

(1) $\begin{bmatrix} 1 & 1 & -1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{bmatrix}$； (2) $\begin{bmatrix} -4 & 2 & 10 \\ -4 & 3 & 7 \\ -3 & 1 & 7 \end{bmatrix}$；

(3) $\begin{bmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{bmatrix}$； (4) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$；

(5) $\begin{bmatrix} 3 & 7 & -3 \\ -2 & -5 & 2 \\ -4 & -10 & 3 \end{bmatrix}$； (6) $\begin{bmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -7 & -6 & -1 & 0 \end{bmatrix}$.

4. 设 A 为 3 阶幂零矩阵，求 A 可能的约当标准形.

5. 设 A 为 3 阶幂等矩阵，求 A 可能的约当标准形.

6. 设 A 为 3 阶矩阵，满足 $A^2=E$，求 A 可能的约当标准形.

7. 设 A 是一个 n 阶复系数矩阵，证明：A 可以表示成一个幂零矩阵及一个可对角化矩阵之和.

8. 设 $f(\lambda)$是 n 阶矩阵 A 的特征多项式，$g(\lambda)$是 A 的最小多项式，证明：$f(\lambda)\mid(g(\lambda))^n$.

第 8 章　线性空间

从这一章开始讨论有关线性空间的一些基本问题. 线性空间是线性代数中最基本的一个概念. 我们在本书上册第 2 章中曾经介绍过 n 维向量空间的概念，n 维向量空间是几何空间的推广. 通过 n 维向量空间使我们有可能把一些反映不同研究对象的有序数组在线性运算下的性质统一在一起讨论. 但是在一些数学问题和实际问题中，还会遇到许多其它对象，对它们也可进行这两种线性运算. 为了对它们统一地加以研究，有必要使向量的概念更为一般化，推广向量的涵意，将 n 维向量空间抽象化，就得到线性空间的概念.

本章将介绍线性空间、它的维数和基、线性子空间等基本概念，并讨论它们的一些重要性质.

8.1　线性空间的定义与简单性质

以前介绍过数域 $\mathbf{P}$ 上的 n 维向量空间 $\mathbf{P}^n$ 是数域 $\mathbf{P}$ 上的所有 n 维向量组成的集合，$\mathbf{P}^n$ 中两个向量可以作加法；$\mathbf{P}$ 中一个数和 $\mathbf{P}^n$ 中一个向量可以作数乘运算，而且这两种运算满足一些基本的规律. 我们还曾经讨论过一些其他的集合，也可以作这两种线性运算. 例如，元素属于 $\mathbf{P}$ 的 $m\times n$ 矩阵，可以作加法以及 $\mathbf{P}$ 中数与矩阵的乘法；数域 $\mathbf{P}$ 上一元多项式环 $\mathbf{P}[x]$ 中两个元素可以作加法，$\mathbf{P}$ 中一个数与 $\mathbf{P}[x]$ 中一个元素可以作数量乘法，更一般地，对于函数也可以定义加法以及数与函数的乘法. 仔细分析，可以发现这些集合对于这两种运算都具有 n 维向量空间的一些性质（参看本

书上册第 2 章 2.2 节)．为了能把这些对象统一地加以研究，从而使所得的结果有更广泛的应用，我们引入线性空间的概念．

定义 1 设 V 是一个非空集合，$\mathbf{P}$ 是一个数域．在集合 V 的元素之间定义了一种运算，叫做**加法**．也就是说，给出了一个法则，按照这个法则，对于 V 中任意两个元素 α 与 β，在 V 中都有唯一的一个元素 γ 与它们对应，称为 α 与 β 的**和**，记作 $\gamma=\alpha+\beta$；在数域 $\mathbf{P}$ 的数与集合 V 的元素之间也定义了一种运算，叫做数量乘法．也就是说，对于数域 $\mathbf{P}$ 中任一数 k 与 V 中任一元素 α，在 V 中都有唯一的一个元素 δ 与它们对应，称为 k 与 α 的**数量乘积**，记作 $\delta=k\alpha$．如果加法与数量乘法满足下述一些规则，那么 V 就称为数域 $\mathbf{P}$ 上的**线性空间**．

加法满足下面 4 条规则：

(1) $\alpha+\beta=\beta+\alpha$；

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

(3) 在 V 中有一个元素 $\mathbf{0}$，对于 V 中任一元素 α 都有

$$\alpha+\mathbf{0}=\alpha,$$

具有这个性质的元素 $\mathbf{0}$ 称为 V 的**零元素**；

(4) 对于 V 中每一个元素 α，都有 V 中的元素 β，使得

$$\alpha+\beta=\mathbf{0}.$$

β 称为 α 的**负元素**，记作 $-\alpha$．

数量乘法满足下面两条规则：

(5) $1\cdot\alpha=\alpha$；

(6) $k(l\alpha)=(kl)\alpha$．

数量乘法与加法满足下面两条规则：

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

(8) $k(\alpha+\beta)=k\alpha+k\beta$．

在以上规则中，k,l 等表示数域 $\mathbf{P}$ 中的任意数，α,β,γ 等表示集合 V 中任意元素．

根据定义，几何空间中全部向量组成的集合，是一个实数域上的线性空间；分量属于数域 **P** 的全体 n 元数组构成数域 **P** 上的一个线性空间，这个线性空间仍用 $\mathbf{P}^n$ 表示；定义在区间 $[a,b]$ 上的连续函数构成实数域上的线性空间.

下面再举几个例子：

例 1 数域 **P** 上的一元多项式环 $\mathbf{P}[x]$，按通常的多项式加法和数与多项式的乘法，构成数域 **P** 上的一个线性空间.

例 2 元素属于数域 **P** 的 $m\times n$ 矩阵全体，按矩阵的加法及矩阵与数的数量乘法，构成数域 **P** 上的一个线性空间，用 $\mathbf{P}^{m\times n}$ 表示.

例 3 数域 **P** 按照本身的加法及乘法，构成一个 **P** 上的线性空间.

例 4 $V=\{a\}$ 只包含一个元素. 对任意数域 **P**，定义

$$a \oplus a = a$$
$$k \circ a = a$$

（为了与通常的加法与乘法相区别，分别用 $\oplus$ 及"$\circ$"表示所定义的加法和数量乘法，以后在必要时也将采用这种符号.）

容易验证 V 对这两种运算满足定义中的 8 条规则，因此是数域 **P** 上的一个线性空间，a 就是 V 的零元素，这种由一个零向量组成的线性空间称为**零空间**.

例 5 **P** 为实数域，V 为全体正实数组成的集合，定义 V 中两个元素的加法 $\oplus$ 为：

$$a \oplus b = ab, \qquad a,b \in V.$$

定义 **P** 中元素与 V 中元素的数乘运算 $\circ$ 为：

$$k \circ a = a^k, \qquad k \in P, a \in V.$$

下面来检验 V 对于这两种运算满足定义 1 中的 8 个条件：

(1) $a\oplus b=ab=ba=b\oplus a$

(2) $(a\oplus b)\oplus c=ab\oplus c=abc=a(bc)=a\oplus(b\oplus c)$

(3) 1 是零元素：

$$1 \oplus a = 1 \cdot a = a$$

(4) a 的负元素是 a^{-1}：

$$a \oplus a^{-1} = aa^{-1} = 1$$

(5) $1 \circ a = a^1 = a$

(6) $k \circ (l \circ a) = k \circ a^l = (a^l)^k = a^{lk} = lk \circ a$

(7) $(k+l) \circ a = a^{k+l} = a^k a^l = a^k \oplus a^l = k \circ a \oplus l \circ a$

(8) $k \circ (a \oplus b) = k \circ ab = (ab)^k = a^k b^k$

$$= a^k \oplus b^k = k \circ a \oplus k \circ b$$

所以 V 对这两种运算构成实数域上的一个线性空间.

线性空间的元素也称为**向量**. 当然,这里的所谓向量,比几何中的向量涵义要广泛得多. 线性空间有时也称为**向量空间**. 以后用小写的希腊字母 $\alpha,\beta,\gamma,\cdots$ 代表线性空间 V 中的元素;用小写的拉丁字母 $a,b,c,\cdots$ 代表数域 $\mathbf{P}$ 中的数. 实数域上的线性空间简称为**实线性空间**;复数域上的线性空间简称为**复线性空间**.

V 的零元素称为**零向量**;$-\alpha$ 称为 α 的**负向量**.

在深入讨论线性空间之前,我们先从定义出发来证明一些有关线性空间的基本性质. 这些性质在今后的讨论中常常要用到. 首先由线性空间定义中的(1),(2)两条规律可得:

命题 1 任意有限多个向量相加时,可以不计其次序及先后.

命题 2 零向量是唯一的.

这就是说,满足定义 1 中条件(3)的向量只有一个.

证明: 如果线性空间 V 中有两个向量 $\mathbf{0}_1,\mathbf{0}_2$ 都满足条件(3). 考虑它们的和 $\mathbf{0}_1+\mathbf{0}_2$: 一方面,由于 $\mathbf{0}_1$ 是零向量,所以

$$\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2;$$

另一方面,又因 $\mathbf{0}_2$ 也是零向量,所以

$$\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_1.$$

因此得 $\mathbf{0}_1=\mathbf{0}_2$. 即零向量是唯一的. |

命题 3 向量 α 的负向量是由 α 唯一确定的.

这就是说，满足条件(4)的 β 只有一个.

证明：如果除 β 外还有一个 α 的负向量 γ，那么

$$(\beta+\alpha)+\gamma=0+\gamma=\gamma,$$

$$\beta+(\alpha+\gamma)=\beta+0=\beta.$$

根据条件(2)，即得

$$\beta=\gamma.$$

利用负向量，可以定义减法为：

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

命题 4 从 $\alpha+\beta=\alpha+\gamma$，可推出 $\beta=\gamma$.

证明：在等式

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

两边都加上 $-\alpha$，即得

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

$$\mathbf{0}+\beta=\mathbf{0}+\gamma$$

即

$$\beta=\gamma.$$

命题 5 当且仅当 $k=0$ 或 $\alpha=\mathbf{0}$ 时，才有 $k\alpha=\mathbf{0}$.

这就是说：$0\alpha=\mathbf{0}$，$k\cdot\mathbf{0}=\mathbf{0}$；从 $k\alpha=\mathbf{0}$ 可推出 $k=0$ 或 $\alpha=\mathbf{0}$.

证明：因为

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

又

$$0\alpha+\mathbf{0}=0\alpha$$

故由命题 4，即得

$$0\alpha=\mathbf{0}.$$

又由

$$k\mathbf{0}=k(\mathbf{0}+\mathbf{0})=k\mathbf{0}+k\mathbf{0}.$$

同上可得

$$k\mathbf{0}=\mathbf{0}.$$

为了证明从 $k\alpha=\mathbf{0}$ 可推出 $k=0$ 或 $\alpha=\mathbf{0}$. 只要证明从 $k\neq 0$, $k\alpha=\mathbf{0}$可推出 $\alpha=\mathbf{0}$. 如果 $k\alpha=\mathbf{0}$ 而 $k\neq 0$. 那么用 k^{-1}乘等式

$$k\alpha=\mathbf{0}$$

两边,即得左端为

$$k^{-1}(k\alpha)=(k^{-1}k)\alpha=1\cdot\alpha=\alpha$$

而右端为

$$k^{-1}\mathbf{0}=\mathbf{0}$$

因此得

$$\alpha=\mathbf{0}$$

命题证毕. ▏

命题 6 $(-k)\alpha=-k\alpha$. 特别地,当 $k=1$ 时,有

$$(-1)\alpha=-\alpha.$$

证明:因为

$$(-k)\alpha+k\alpha=(-k+k)\alpha=0\alpha=\mathbf{0}$$

故由负向量的定义,即得

$$(-k)\alpha=-k\alpha. \quad ▏$$

这些命题及其证明方法在以后经常要用到.

习 题 8.1

1. 检验以下集合对于所指的运算,是否构成数域 **P** 上的线性空间:

(1) 集合:数域 **P** 上次数小于 n 的多项式全体再添上零多项式;运算:多项式的加法及数量乘法.

(2) 集合:数域 **P** 上 n 次多项式全体;运算:多项式的加法与数量乘法.

(3) 集合:数域 **P** 上全体 n 阶实对称矩阵;运算:矩阵的加法与数量乘法.

(4) 集合:数域 **P** 上的 n 阶矩阵 A 的多项式全体;运算:矩阵的加法和数量乘法.

(5) **P**:实数域;集合:实数的二元数列全体;运算:

$$(a_1,b_1)\oplus(a_2,b_2)=(a_1+a_2,b_1+b_2+a_1a_2),$$

$$k\circ(a_1,b_1)=\left(ka_1,kb_1+\frac{k(k-1)}{2}a_1^2\right).$$

2. 在例 5 的线性空间中,计算

$$3\circ 4\oplus(-1)\circ 2\oplus(-2)\circ 7.$$

3. 在第 1 题(5)的线性空间中,计算

$$3\circ(0,1)\oplus(-2)\circ(2,3)\oplus 2\circ(1,0).$$

4. 试证: $-(\alpha+\beta)=-\alpha-\beta$.

5. 试证: $k(\alpha-\beta)=k\alpha-k\beta$.

6. 试证:对于线性空间 V 中的任意两个向量 α,β,有 V 中唯一的一个向量 γ,使

$$\alpha+\gamma=\beta.$$

即方程

$$\alpha+x=\beta$$

在 V 中有唯一解.

8.2 向量组的线性关系

在讨论 n 维向量空间时,曾经介绍了线性组合、线性相关、线性无关等概念.这些概念是非常重要的.它们在线性方程组和矩阵等问题的研究中,起了很大的作用.为了对线性空间及其有关问题进行讨论,也需要对线性空间中的向量引入这些概念.

定义 2 设 V 是数域 $\mathbf{P}$ 上的一个线性空间,$\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 V 中一组向量;$k_1,k_2,\cdots,k_s$ 是数域 $\mathbf{P}$ 中的数.向量

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

称为 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个**线性组合**,$k_1,k_2,\cdots,k_s$ 称为这个线性组合的系数.

定义 3 如果向量 β 可以表成向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 的一个线性组合,即如果有数域 $\mathbf{P}$ 中的数 $k_1,k_2,\cdots,k_s$,使得

$$\beta=k_1\alpha_2+k_2\alpha_2+\cdots+k_s\alpha_s,$$

则称 β 可由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ **线性表出**.

根据定义,在任一线性空间 V 中,零向量是任意向量组 α_1,$\alpha_2,\cdots,\alpha_s$ 的线性组合:

$$\mathbf{0}=0\alpha_1+0\alpha_2+\cdots+0\alpha_s.$$

也就是说,零向量可由任意向量组线性表出.

下面再来举几个例子.

例 1 在 $\mathbf{P}^{2\times 2}$ 中,任一向量都可由下列向量组线性表出:

$$\begin{bmatrix}1&0\\0&0\end{bmatrix},\quad \begin{bmatrix}0&1\\0&0\end{bmatrix},\quad \begin{bmatrix}0&0\\1&0\end{bmatrix},\quad \begin{bmatrix}0&0\\0&1\end{bmatrix}.$$

这是因为:

$$\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}=a_{11}\begin{bmatrix}1&0\\0&0\end{bmatrix}+a_{12}\begin{bmatrix}0&1\\0&0\end{bmatrix}+a_{21}\begin{bmatrix}0&0\\1&0\end{bmatrix}+a_{22}\begin{bmatrix}0&0\\0&1\end{bmatrix}.$$

例 2 在多项式组成的线性空间 $\mathbf{P}[x]$中,容易看出,任一个二次多项式都可以由 $1,x,x^2$ 线性表出,而 x^2 不能由 $1,x$ 线性表出.

线性组合及线性表出这两个概念是我们所熟悉的,它们几乎是逐字逐句地重复了上册第 2 章中关于 n 维向量的相应的定义. 不但如此,还有在第 2 章中介绍过的其他一些关于向量组的线性关系的定义、结果及其证明也可以推广到这里的数域 $\mathbf{P}$ 上的线性空间中来. 为了便于在以后应用,我们把这些内容加以复习和补充. 有些在以前曾经证明过的结论,我们只是叙述一下而不再重复证明了.

首先,我们重述一个非常简单然而十分有用的结果.

命题 1 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中任一个向量 $\alpha_i(1\leqslant i\leqslant s)$都可以由这个向量组线性表出.

下面来复习关于向量组线性相关性的一些定义和结论.

定义 4 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是数域 $\mathbf{P}$ 上线性空间 V 中的一组向量，如果在数域 $\mathbf{P}$ 中有 s 个不全为 0 的数 $k_1,k_2,\cdots,k_s$，使得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=\mathbf{0},$$

就称向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是**线性相关**的.

我们知道，任何一组向量的系数全为零的线性组合一定是零向量. 定义 4 是说，一个线性相关的向量组除了系数全为零的线性组合是零向量外，它还可以有系数不全为零的线性组合也等于零向量. 因此，零向量由一个线性相关的向量组线性表出的方法不止一种.

命题 2 (1) 包含零向量的向量组一定是线性相关的.

(2) 单个向量 α 线性相关的充要条件是 $\alpha=\mathbf{0}$.

证明：(1) 设 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant1)$是一个含有零向量的向量组，其中 $\alpha_i=0$ $(1\leqslant i\leqslant s)$. 于是

$$0\alpha_1+\cdots+0\alpha_{i-1}+1\alpha_i+0\alpha_{i+1}+\cdots+0\alpha_s=\mathbf{0}.$$

因此由定义可知这个向量组是线性相关的.

(2) 充分性可由(1)得出. 如果 α 线性相关，则有不等于零的数 k，使得 $k\alpha=\mathbf{0}$，因此 $\alpha=\mathbf{0}$.

定义 5 一个向量组如果不是线性相关的，就称为是**线性无关**的. 也就是说，向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 称为线性无关的，如果只有系数 $k_1,k_2,\cdots,k_s$ 全等于 0，才能使得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=\mathbf{0}.$$

一个线性无关的向量组只有系数全为零的线性组合才是零向量. 因此，零向量由一个线性无关的向量组线性表出的方法只有一种. 根据这一点，线性无关的定义也可叙述成：

定义 5′ $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是数域 $\mathbf{P}$ 上线性空间 V 中的一组向量. 如果从

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=\mathbf{0}\qquad(k_1,k_2,\cdots,k_s\in\mathbf{P})$$

可推出

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

就称 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是线性无关的.

例 3 试证,在习题 8.1 第 5 题的线性空间中,下述向量组是线性无关的:

$$\alpha_1 = (1,0), \qquad \alpha_2 = (0,1).$$

证明 如果

$$k_1 \circ \alpha_1 \oplus k_2 \circ \alpha_2 = \mathbf{0},$$

即

$$k_1 \circ (1,0) \oplus k_2 \circ (0,1) = (0,0).$$

则由运算的定义可得:

$$\left(k_1, \frac{k_1(k_1 - 1)}{2} + k_2 \right) = (0,0).$$

因此

$$k_1 = 0, \quad \frac{k_1(k_1 - 1)}{2} + k_2 = 0.$$

所以

$$k_1 = k_2 = 0.$$

根据定义 5′,知 α_1, α_2 是线性无关的.

由一个向量组中的一部分向量组成的向量组称为原向量组的一个**部分组**.下述命题是今后常常用到的.

命题 3 (1) 如果向量组有一个部分组是线性相关的,那么这个向量组也是线性相关的;

(2) 线性无关向量组的任一个部分组都是线性无关的.

需要注意的是:命题 3 的逆命题不成立.即:一个线性相关的向量组不一定有一个线性相关的真部分组;即使一个向量组的任一个真部分组都线性无关,也不能推断这个向量组一定是线性无关的.我们通过例子来说明这一事实.

例 4 在 $\mathbf{P}^3$ 中向量组

$$(1,1,0),\qquad (1,0,0),\qquad (0,1,0)$$

是线性相关的，但是它的任一个真部分组都是线性无关的. 而且这个向量组也说明了即使一个向量组的任一个真部分组都是线性无关的，它本身也可能是线性相关的.

因为一个向量组或者是线性相关的，或者是线性无关的，二者必居其一，且仅居其一. 因此从关于线性相关的一个结论，可以推得关于线性无关的一个相应的结论. 反之亦然. 例如命题 2 中关于线性相关的结论，用线性无关的语言来叙述，就是：一个线性无关的向量组一定不能包含零向量；单个向量 α 线性无关的充要条件是 $\alpha\neq\mathbf{0}$. 又如命题 3 中的两个结论其实是同一个事实，不过根据线性相关及线性无关这两个概念有两种说法而已. 所以例 4 中的同一个例子就同时说明了这两个结论的逆命题不成立. 以后，有时只用一种方式来叙述. 希望读者能联想到另一种叙述，并且在应用或证明某个结论的时候，能灵活地使用线性相关或线性无关这两个概念.

为了简便起见，以下提到的向量如不特别声明，都理解为数域 $\mathbf{P}$ 上线性空间 V 中的向量，而不再另加说明了.

线性相关和线性表出有下述关系.

定理 1 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 2)$ 线性相关的充分必要条件是：$\alpha_1,\alpha_2,\cdots,\alpha_s$ 中有一个向量可以由其余的向量线性表出.

将定理 1 用于线性无关，就得：

向量 $\alpha_1,\alpha_2,\cdots,\alpha_s(s\geqslant 2)$ 线性无关的充分必要条件，是其中任一个向量 $\alpha_i(i=1,2,\cdots,s)$ 都不能由其余的向量线性表出.

定理 2 如果向量组

$$\alpha_1,\alpha_2,\cdots,\alpha_s$$

线性无关，而向量组

$$\alpha_1,\alpha_2,\cdots,\alpha_s,\beta$$

线性相关，那么 β 可由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出，而且表法是

唯一的.

证明：因为向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s,\beta$ 是线性相关的，所以有不全为零的数 $k_1,k_2,\cdots,k_s,l$，使得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s+l\beta=\mathbf{0}.$$

假如 $l=0$，则 $k_1,k_2,\cdots,k_s$ 不全为零. 并且由上式可得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=\mathbf{0}.$$

这与 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是线性无关的假设相矛盾. 因此 $l\neq 0$. 于是

$$\beta=-\frac{k_1}{l}\alpha_1-\frac{k_2}{l}\alpha_2-\cdots-\frac{k_s}{l}\alpha_s.$$

所以 β 可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出.

下面证明表法的唯一性. 如果 β 可表成

$$\beta=a_1\alpha_1+a_2\alpha_2+\cdots+a_s\alpha_s,$$

又可表成

$$\beta=b_1\alpha_1+b_2\alpha_2+\cdots+b_s\alpha_s.$$

两式相减，得

$$(a_1-b_1)\alpha_1+(a_2-b_2)\alpha_2+\cdots+(a_s-b_s)\alpha_s=\mathbf{0}.$$

因为向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是线性无关的，所以由上式可得

$$a_1-b_1=a_2-b_2=\cdots=a_s-b_s=0.$$

即

$$a_1=b_1,\quad a_2=b_2,\quad \cdots,\quad a_s=b_s.$$

所以 β 由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出的表法只有一种. |

现在再证明一个今后常常用到的定理.

定理 3 设 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是线性空间 V 中一组线性无关的向量，$\beta_1,\beta_2,\cdots,\beta_s$ 都是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的线性组合：

$$\beta_1=a_{11}\alpha_1+a_{12}\alpha_2+\cdots+a_{1s}\alpha_s,$$
$$\beta_2=a_{21}\alpha_1+a_{22}\alpha_2+\cdots+a_{2s}\alpha_s,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\beta_s=a_{s1}\alpha_1+a_{s2}\alpha_2+\cdots+a_{ss}\alpha_s.$$

那么，向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 线性无关的充分必要条件是：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{vmatrix} \neq 0.$$

这个定理的一个等价的说法是，在定理 3 的假设下，$\beta_1,\beta_2,\cdots,\beta_s$ 线性相关的充要条件是：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{vmatrix} = 0.$$

下面就通过证明这一点来证明定理 3.

证明：$\beta_1,\beta_2,\cdots,\beta_s$线性相关的充分必要条件是：可以找到不全为零的数 $k_1,k_2,\cdots,k_s$，使得

$$k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_s\boldsymbol{\beta}_s=\mathbf{0}.$$

把 $\beta_1,\beta_2,\cdots,\beta_s$ 的表式代入上式，得

$$\begin{aligned} & k_1(a_{11}\boldsymbol{\alpha}_1+a_{12}\boldsymbol{\alpha}_2+\cdots+a_{1s}\boldsymbol{\alpha}_s) \\ & \quad +k_2(a_{21}\boldsymbol{\alpha}_1+a_{22}\boldsymbol{\alpha}_2+\cdots+a_{2s}\boldsymbol{\alpha}_s) \\ & \quad +\cdots+k_s(a_{s1}\boldsymbol{\alpha}_1+a_{s2}\boldsymbol{\alpha}_2+\cdots+a_{ss}\boldsymbol{\alpha}_s) \\ = & (k_1a_{11}+k_2a_{21}+\cdots+k_sa_{s1})\boldsymbol{\alpha}_1 \\ & \quad +(k_1a_{12}+k_2a_{22}+\cdots+k_sa_{s2})\boldsymbol{\alpha}_2 \\ & \quad +\cdots+(k_1a_{1s}+k_2a_{2s}+\cdots+k_sa_{ss})\boldsymbol{\alpha}_s \\ = & \mathbf{0}. \end{aligned}$$

因为向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$线性无关，所以上式成立的充分必要条件是：

$$\begin{cases} k_1a_{11}+k_2a_{21}+\cdots+k_sa_{s1}=0, \\ k_1a_{12}+k_2a_{22}+\cdots+k_sa_{s2}=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ k_1a_{1s}+k_2a_{2s}+\cdots+k_sa_{ss}=0. \end{cases}$$

因此，$\beta_1,\beta_2,\cdots,\beta_s$ 线性相关的充分必要条件是：存在一组不全为零的数 $k_1,k_2,\cdots,k_s$，使上面 s 个等式都成立。即：齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{21}x_2+\cdots+a_{s1}x_s=0, \\ a_{12}x_1+a_{22}x_2+\cdots+a_{s2}x_s=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1s}x_1+a_{2s}x_2+\cdots-a_{ss}x_s=0. \end{cases}$$

有非零解。亦即其系数行列式

$$\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{s1} \\ a_{12} & a_{22} & \cdots & a_{s2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1s} & a_{2s} & \cdots & a_{ss} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{vmatrix} = 0.$$

所以 $\beta_1,\beta_2,\cdots,\beta_s$ 线性无关的充分必要条件是：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{vmatrix} \neq 0.$$

以上讨论了向量组的线性相关性的一些重要性质.这些内容都是我们熟悉的.它们几乎是重复了第 2 章关于 n 维向量空间的讨论.不过，这里的向量和线性运算比以前广泛多了.这一点从我们所举的例子可以看出来.由于这些内容都是线性空间理论中最基本的概念，也是今后讨论中常常用到的结论，因此我们不厌其烦地提出这些概念，加以讨论并作必要的补充.为了定义线性空间的维数和基等概念，我们还需要复习关于向量组的等价及向量组的

秩等概念和结论.

定义 6 设

$$\alpha_1, \alpha_2, \cdots, \alpha_s \tag{1}$$

及

$$\beta_1, \beta_2, \cdots, \beta_t \tag{2}$$

是线性空间 V 中两个向量组. 如果向量组(1)中每个向量都可由向量组(2)线性表出,就称向量组(1)可由向量组(2)线性表出. 如果向量组(1)可由(2)线性表出,向量组(2)也可由(1)线性表出,就称这两个向量组是**等价的**.

根据定义,可以看出向量组之间的等价关系是对称的. 容易证明,这个关系还具有反身性和传递性.

例 5 在 $\mathbf{P}^{2\times 2}$ 中向量组

$$\begin{pmatrix}1&0\\0&0\end{pmatrix},\quad \begin{pmatrix}0&1\\0&0\end{pmatrix},\quad \begin{pmatrix}0&0\\1&0\end{pmatrix},\quad \begin{pmatrix}0&0\\0&1\end{pmatrix} \tag{①}$$

与向量组

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

是等价的.

向量组

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

可以由向量组①线性表出,也可由向量组②线性表出. 但是向量组①及②都不能由向量组③线性表出. 因此,向量组①与③不是等价的;向量组②与③也不是等价的.

从定义 6 和命题 1,可知向量组的部分组一定可由这个向量组线性表出. 例如,例 5 中的向量组③是向量组①的一个部分组,它可由向量组①线性表出.

一般说来,一个向量组能否由另一个向量组线性表出与它们所含的向量个数没有关系. 然而,对于线性无关的向量组来说,向

量组中所含向量的个数对于线性表出的问题是有影响的.

定理 4 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 与 $\beta_1,\beta_2,\cdots,\beta_t$ 是线性空间 V 中两个向量组,如果

(1) 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 可由向量组 $\beta_1,\beta_2,\cdots,\beta_t$ 线性表出;

(2) $s>t$,

那么向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 一定线性相关.

定理 4 的另一种说法是

推论 1 如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 可由向量组 $\beta_1,\beta_2,\cdots,\beta_t$ 线性表出,而且向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是线性无关的,那么 $s\leqslant t$.

从定理 4 还可推出:

推论 2 两个线性无关的向量组如果等价,则它们包含的向量个数一定相等.

证明 可以应用推论 1 加以证明.

下面我们来讨论向量组的秩.

定义 7 向量组的一个部分组称为一个**极大线性无关组**,如果这个部分组本身是线性无关的,但是再从原向量组的其余向量(如果还有的话)中任取一个添进去以后,所得到的部分组都是线性相关的.

例 6 在多项式空间中,设

$$\alpha_1=1,\ \alpha_2=x^2,\ \alpha_3=x^2+1,\ \alpha_4=3x^2-2,$$

则向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的部分组 α_1,α_2 是它们的一个极大线性无关组,因为向量组 α_1,α_2 显然是线性无关的.由于

$$\alpha_3=\alpha_1+\alpha_2,\qquad \alpha_4=-2\alpha_1+3\alpha_2.$$

所以在部分组 α_1,α_2 中分别添入 α_3 或 α_4 后,所得的部分组 $\alpha_1,\alpha_2,\alpha_3$ 及 $\alpha_1,\alpha_2,\alpha_4$ 都是线性相关的了,因此,α_1,α_2 是向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的一个极大线性无关组.同样可证 α_1,α_3 也是 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的极大线性无关组.向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 除去这两个极大线性无关组外,还有其他极大线性无关组.读者不妨自己找一下并加以

证明.

从例 6 可以看出,应用下述命题来求向量组的极大线性无关组,有时要比直接用定义 7 方便一些.

命题 4 设 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个线性无关的部分组,如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中每个向量都可由 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性表出,那么 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个极大线性无关组.

很容易看到,一个线性无关的向量组就是自己的一个极大线性无关组,而且是它的唯一的极大线性无关组. 而一个线性相关向量组的极大线性无关组中所含向量个数一定少于原向量组中向量的个数.

完全由零向量组成的向量组(简称零向量组)没有极大线性无关组,因为它的任一个部分组都是线性相关的. 除此以外,任何向量组只要包含非零向量就一定有极大线性无关组. 下面介绍求向量组的极大线性无关组的一个方法. 这个方法需要用到下述命题.

命题 5 设向量组

$$\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_2,\ \cdots,\ \boldsymbol{\beta}_t$$

满足

(1) $\boldsymbol{\beta}_1\neq\mathbf{0}$;

(2) 每个 $\boldsymbol{\beta}_i(i=2,3,\cdots,t)$ 都不能被 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{i-1}$ 线性表出,

那么向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性无关.

证明: 用反证法. 如果 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性相关,那么有不全为零的数 $k_1,k_2,\cdots,k_t$,使得

$$k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_t\boldsymbol{\beta}_t=\mathbf{0}.$$

设 $k_t,k_{t-1},\cdots,k_2,k_1$ 中第一个不等于零的是 $k_l(1\leqslant l\leqslant t)$,即 $k_l\neq0$,而 $k_{l+1}=\cdots=k_t=0$. 于是上式可写成:

$$k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_l\boldsymbol{\beta}_l=\mathbf{0}.$$

如果 $l=1$,那么上式成为:

$$k_1\beta_1 = \mathbf{0} \qquad (k_1 \neq 0).$$

与 $\beta_1 \neq \mathbf{0}$ 的假设相矛盾，所以 $l>1$. 于是

$$\beta_l = -\frac{k_1}{k_l}\beta_1 - \frac{k_2}{k_l}\beta_2 - \cdots - \frac{k_{l-1}}{k_l}\beta_{l-1}.$$

这说明 β_l 可由 $\beta_1,\beta_2,\cdots,\beta_{l-1}$ 线性表出，与假设(2)相矛盾，所以 $\beta_1,\beta_2,\cdots,\beta_s$ 一定是线性无关的.

根据命题 5，可以得到求向量组的极大线性无关组的方法如下：

给了一个向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$，如果都是零向量，就没有极大线性无关组. 如果其中含有非零向量，设 α_{i_1} 是其中第一个非零向量，即 $\alpha_1,\cdots,\alpha_{i_1-1}$ 都是零向量，而 $\alpha_{i_1}\neq 0$，于是考虑部分向量组 $\alpha_{i_1},\alpha_{i_1+1},\cdots,\alpha_s$. 保留 α_{i_1}，然后从 α_{i_1+1} 起逐个检查. 设 $\alpha_{i_1+1},\cdots,\alpha_{i_2-1}$ 都可被 α_{i_1} 线性表出，而 α_{i_2} 不能被 α_{i_1} 线性表出，那么再去掉向量 $\alpha_{i_1+1},\cdots,\alpha_{i_2-1}$，进而考虑部分组 $\alpha_{i_1},\alpha_{i_2},\alpha_{i_2-1},\cdots,\alpha_s$. 一般地，找出 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_u}$ 以后，就从 α_{i_u+1} 起逐个检查，如果有某个向量可以由前面的向量线性表出，就把这个向量去掉，否则就把这个向量留下来取作 $\alpha_{i_{u+1}}$，再从 $\alpha_{i_{u+1}+1}$ 起逐个检查. 最后，如果留下来的部分组是 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$，那么这个部分组就是原向量组的一个极大线性无关部分组. 因为 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 中第一个向量 $\alpha_{i_1}\neq 0$，而 α_{i_2}，$\alpha_{i_3},\cdots,\alpha_{i_r}$ 中每一个都不能被前面的向量线性表出，故由命题 11 知，$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是线性无关的. 又因 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中不属于 α_{i_1}，$\alpha_{i_2},\cdots,\alpha_{i_r}$ 的那些向量都可由 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 中的一部分向量线性表出，当然也可被 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 线性表出. 因此，$\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_r}$ 是原向量组的一个极大无关部分组.

例 7 在 $\mathbf{P}^n$ 中设

$$\alpha_1 = (0,0,0,0,0), \qquad \alpha_2 = (1,0,-1,1,0),$$
$$\alpha_3 = (2,0,-2,2,0), \quad \alpha_4 = (1,0,0,1,0),$$

$$\alpha_5=(-2,0,3,-2,0),\quad \alpha_6=(1,1,1,1,1),$$
$$\alpha_7=(1,1,0,1,1),\quad \alpha_8=(2,5,8,2,5),$$
求向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6,\alpha_7,\alpha_8$ 的一个极大线性无关组.

解：向量组中第一个向量 $\alpha_1=\mathbf{0}$，将它去掉. 考虑 $\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6,\alpha_7,\alpha_8$.

$\alpha_2\neq\mathbf{0}$，留下；$\alpha_3=2\alpha_2$，去掉；α_4 不能由 α_2 线性表出，留下；$\alpha_5=-3\alpha_2+\alpha_4$，去掉；$\alpha_6$ 不能被 α_2,α_4 线性表出，留下；$\alpha_7=\alpha_2-\alpha_4+\alpha_6$，去掉；$\alpha_8=-3\alpha_2+5\alpha_6$，去掉. 这样，就得到原向量组的一个极大线性无关组：$\alpha_2,\alpha_4,\alpha_6$.

如果改变 $\alpha_1,\alpha_2,\alpha_3,\cdots,\alpha_8$ 的次序，用这个方法还可得到另外的极大线性无关组. 设 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_t}$ 是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个线性无关的部分组，改变 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的次序，把 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_t}$ 排在前面，即排成

$$\alpha_{j_1},\ \alpha_{j_2},\ \cdots,\ \alpha_{j_t},\ \cdots.$$

再用上面介绍的方法找出一个极大线性无关组. 从找的方法可以看出：这个极大线性无关组一定包含向量 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_t}$. 这说明一个向量组的任何一个线性无关的部分组，都可以取作某个极大线性无关组的一部分. 也就是说，一个向量组的任何一个线性无关的部分组都可以**扩充**成一个极大线性无关组.

向量组的极大线性无关组有以下一个重要的性质：

定理 5 (1) 向量组的任一个极大线性无关组都与原向量组等价.

(2) 向量组的任意两个极大线性无关组都是等价的，因此所包含的向量个数相同.

我们从例 6 可看到，一个向量组的极大线性无关组不一定是唯一的，然而定理 4 表明：同一个向量组的各个极大线性无关组所含的向量个数却是一样的，都与极大线性无关组的选择无关，它直接反映了向量组本身的性质. 因而可引出向量组的秩的概念.

定义 8 向量组的极大线性无关组中所包含向量的个数称为这个向量组的**秩**.

只含零向量的向量组的秩认为是 0.

我们用 $r\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$表示向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩. 例如，例 7 中向量组 $\alpha_1,\alpha_2,\cdots,\alpha_8$ 的秩等于 3，就可记作

$$r\{\alpha_1,\alpha_2,\cdots,\alpha_8\}=3.$$

关于向量组的秩有下列一些重要结论. 这些结论的证明方法与以前关于 n 维向量的结论的证明类似，现把结果写在下面.

命题 6 向量组线性无关的充分必要条件是：它的秩等于它所含向量的个数.

命题 7 (1) 如果向量组(Ⅰ) $\alpha_1,\alpha_2,\cdots,\alpha_r$ 可以由向量组(Ⅱ) $\beta_1,\beta_2,\cdots,\beta_t$ 线性表出，则向量组(Ⅰ)的秩≤向量组(Ⅱ)的秩.

(2) 等价的向量组有相同的秩.

(3) 如果两个向量组的秩相等，并且其中的一个向量组可以由另一个向量组线性表出，则这两个向量组等价.

以上关于向量的线性相关性的这些理论，将在下一节中用来定义线性空间的维数、基和坐标.

习　题　8.2

1. 在习题 8.1 第(5)题的 V 中，把

$$\beta_1=(2,0),\quad \beta_2=(1,1)$$

表成

$$\alpha_1=(1,0),\quad \alpha_2=(0,1)$$

的线性组合.

2. 如果 α 可由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个部分组线性表出，则 α 也能由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$线性表出.

3. $\alpha,\alpha_1,\alpha_2,\cdots,\alpha_s$都是数域 **P** 上线性空间 V 中的向量，试证：α 能由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$线性表出的充分必要条件是：存在数域 **P** 中的数 $k,k_1,k_2,$

$\cdots, k_s$，其中 $k \neq 0$，使得

$$k\alpha + k_1\alpha_1 + \cdots + k_s\alpha_s = \mathbf{0}.$$

4. 试证：在上节例 5 的 V 中，任何两个向量都是线性相关的.

5. 在线性空间 $\mathbf{P}^{2\times 2}$ 中

(1) 证明：向量组

$$\boldsymbol{E}_{11}, \boldsymbol{E}_{12}, \boldsymbol{E}_{21}, \boldsymbol{E}_{22}$$

与向量组

$$\boldsymbol{E}_{11}, \boldsymbol{E}_{22}, \boldsymbol{E}_{12} + \boldsymbol{E}_{21}, \boldsymbol{E}_{12} - \boldsymbol{E}_{21}$$

等价.

(2) 将 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 表成 $\boldsymbol{E}_{11}, \boldsymbol{E}_{22}, \boldsymbol{E}_{12}+\boldsymbol{E}_{21}, \boldsymbol{E}_{12}-\boldsymbol{E}_{21}$ 的线性组合.

6. 设

$$\begin{aligned}
\boldsymbol{\beta}_1 &= \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \cdots + \boldsymbol{\alpha}_{s-1} - \boldsymbol{\alpha}_s, \\
\boldsymbol{\beta}_2 &= \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \cdots - \boldsymbol{\alpha}_{s-1} + \boldsymbol{\alpha}_s, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\boldsymbol{\beta}_{s-1} &= \boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2 + \cdots + \boldsymbol{\alpha}_{s-1} + \boldsymbol{\alpha}_s, \\
\boldsymbol{\beta}_s &= -\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \cdots + \boldsymbol{\alpha}_{s-1} + \boldsymbol{\alpha}_s, \qquad (s \geqslant 3).
\end{aligned}$$

试证：秩$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s\}$＝秩$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s\}$.

7. 设

$$\begin{aligned}
\boldsymbol{\alpha}_1 &= (1,0,1,2), & \boldsymbol{\alpha}_2 &= (1,-2,-11,0), \\
\boldsymbol{\alpha}_3 &= (3,-1,-3,5), & \boldsymbol{\alpha}_4 &= (2,4,0,3), \\
\boldsymbol{\alpha}_5 &= (-1,1,5,-1),
\end{aligned}$$

求向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4, \boldsymbol{\alpha}_5$ 的一个极大线性无关组，其中包含向量 $\boldsymbol{\alpha}_3$.

8.3 维数、基与坐标

根据上节中关于向量组的线性相关性的一些结论，可以来定义线性空间的维数、基与坐标.

定义 9 如果数域 $\mathbf{P}$ 上线性空间 V 中有 n 个线性无关的向量，但是 V 中任意 $n+1$ 个向量都是线性相关的，那么 V 就称为 n

维线性空间，n 称为 V 的**维数**. 如果在 V 中存在任意多个线性无关的向量，那么 V 就称为**无限维线性空间**.

在几何空间中，可找到 3 个线性无关的向量，而其中任意 4 个向量都是线性相关的，所以根据定义 9，它的维数等于 3. 由数域 $\mathbf{P}$ 上 n 维向量组成的线性空间 $\mathbf{P}^n$ 中，n 个 n 维基本向量是线性无关的，而任意 $n+1$ 个向量都是线性相关的，所以按照定义 9，$\mathbf{P}^n$ 的维数为 n. 这与以前的理解是一致的.

零空间的维数规定为 0.

在 $\mathbf{P}[x]$中对任意正整数 n，都可找到 $n+1$ 个线性无关的向量：

$$1,\ x,\ x^2,\ \cdots,\ x^n.$$

因此 $\mathbf{P}[x]$是一个无限维线性空间.

n 维线性空间通称有限维线性空间. 本书主要讨论有限维线性空间. 无限维线性空间的研究，属于泛函分析学科，它与有限维线性空间有较大的差别，超出了本书的范围. 为了方便起见，我们用 $\dim V$ 表示线性空间的维数.

为了证明一个线性空间的维数为 n，从定义看，需要做到下述两点：

(1) 找出 V 中 n 个线性无关的向量；

(2) 证明 V 中任意 $n+1$ 个向量都是线性相关的.

这样做是比较麻烦的. 应用下面的定理可以把第二点简化.

定理 6 如果在线性空间 V 中可以找到 n 个线性无关的向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$，而且 V 中任一个向量都可以由它们线性表出，那么 V 是 n 维的.

证明：任取 V 中 $n+1$ 个向量 $\beta_1,\beta_2,\cdots,\beta_{n+1}$. 由假设，它们都可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出. 因此根据定理 4，$\beta_1,\beta_2,\cdots,\beta_{n+1}$线性相关，故由定义 9，$V$ 的维数为 n. |

例 1 求 $\mathbf{P}^{m\times n}$ 的维数.

解：用 E_{ij} 表示第 i 行第 j 列处的元素为 1，而其它元素全为零的 $m\times n$ 矩阵. 于是 E_{ij}，$i=1,2,\cdots,m$；$j=1,2,\cdots,n$ 是 $\mathbf{P}^{m\times n}$ 中 mn 个线性无关的向量. 而且，$\mathbf{P}^{m\times n}$ 中任一向量都可被它们线性表出：

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}E_{ij}.$$

因此维($\mathbf{P}^{m\times n}$)$=mn$.

例 2 数域 $\mathbf{P}$ 上次数小于 n 的多项式全体再添上零多项式组成的集合，对于多项式的加法及多项式与数的乘法构成数域 $\mathbf{P}$ 上一个线性空间，记作 $\mathbf{P}_n[x]$. 在 $\mathbf{P}_n[x]$中，有 n 个线性无关的向量

$$1,\ x,\ x^2,\ \cdots,\ x^{n-1}.$$

而且 $\mathbf{P}_n[x]$中任一向量都可表成它们的线性组合，所以 $\mathbf{P}_n[x]$是 n 维的.

在解析几何中我们看到，坐标是研究向量的一个有力工具. 在有限维线性空间中，坐标也有同样的作用.

定义 10 在 n 维线性空间 V 中，n 个线性无关的向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 称为 V 的一组**基**.

定义 11 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 n 维线性空间 V 的一组基，α 是 V 中一个向量，则 α 可由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性表出，而且表法是唯一的. 如果

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

则 $a_1,a_2,\cdots,a_n$ 称为 α 关于基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的**坐标**，记作$(a_1,a_2,\cdots,a_n)$.

例 3 我们已知 $\mathbf{P}_n[x]$是 n 维的，而且

$$1,\ x,\ x^2,\ \cdots,\ x^{n-1}.$$

是 $\mathbf{P}_n[x]$中 n 个线性无关的向量，因此这是 $\mathbf{P}_n[x]$的一组基. 对于

这组基，$\mathbf{P}_n[x]$中多项式

$$f(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}$$

的坐标就是 $f(x)$按升幂排列的系数$(a_0, a_1, \cdots, a_{n-1})$.

例 4 在 n 维线性空间 $\mathbf{P}^n$ 中，显然

$$\varepsilon_1 = (1, 0, \cdots, 0), \quad \varepsilon_2 = (0, 1, \cdots, 0), \cdots,$$
$$\varepsilon_n = (0, 0, \cdots, 1)$$

是一组基. 对于这组基，向量 $\alpha = (a_1, a_2, \cdots, a_n)$的坐标就是它的分量 $a_1, a_2, \cdots, a_n$. 令

$$\varepsilon_1' = (1, 0, \cdots, 0), \quad \varepsilon_2' = (1, 1, \cdots, 0), \cdots,$$
$$\varepsilon_n' = (1, 1, \cdots, 1),$$

则因

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

所以向量组 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 线性无关，因此也是 $\mathbf{P}^n$ 的一组基. 对于这组基，因为向量 $\alpha = (a_1, a_2, \cdots, a_n)$可表成

$$\alpha = (a_1 - a_2)\varepsilon_1' + (a_2 - a_3)\varepsilon_2' + \cdots$$
$$+ (a_{n-1} - a_n)\varepsilon_{n-1}' + a_n \varepsilon_n',$$

所以 α 关于这组基的坐标是

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

根据定义 10，n 维线性空间中任意 n 个线性无关的向量都可取作这个线性空间的一组基. 下列命题更进一步说明基的普遍性：

命题 8 设 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 是 n 维线性空间 V 中 m 个线性无关的向量，那么总可找到 V 中 $n-m$ 个向量 $\alpha_{m+1}, \alpha_{m+2}, \cdots, \alpha_n$，使得 $\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$ 成为 V 的一组基.

证明：只要取 $\alpha_{m+1}, \cdots, \alpha_n$ 使 $\alpha_1, \alpha_2, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$ 线性无关即可.

这个命题在以后经常要用到.一般叙述为：n 维线性空间中任意一组线性无关的向量都可以扩充成 V 的一组基.

习　题　8.3

1. 试证：如果在线性空间 V 中可找到 n 个线性无关的向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$，使得对 V 中任意向量 β，向量组 $\alpha_1,\cdots,\alpha_n,\beta$ 都是线性相关的，则 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一组基.
2. 求(1) 8.1 节中例 5 与(2) 习题 8.1 第 1 题(5)中的线性空间的维数和一组基.
3. 证明下列集合对所定义的运算构成线性空间，并求其维数和一组基.

   (1) 数域 $\mathbf{P}$ 上的 n 阶对称矩阵全体；运算：矩阵的加法及 $\mathbf{P}$ 中的数与矩阵的数量乘法.

   (2) 数域 $\mathbf{P}$ 上的 n 阶反对称矩阵全体；运算同(1)；

   (3) 数域 $\mathbf{P}$ 上的 n 阶上三角形矩阵全体；运算同(1).
4. 在线性空间 V 中求向量 ξ 关于基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的坐标.

   (1) $V=\mathbf{P}^4$；$\varepsilon_1=(1,1,1,1)$，　$\varepsilon_2=(1,1,-1,-1)$，

   $\varepsilon_3=(1,-1,1,-1)$，　　$\varepsilon_4=(1,-1,-1,1)$；

   $\xi=(1,2,-2,-1)$.

   (2) V 是 8.1 例 5 中的线性空间，
   $$\varepsilon_1=2;\ \xi=3,$$

   (3) V 为习题 8.1 第 1 题(5)中的线性空间，
   $$\varepsilon_1=(1,0),\ \varepsilon_2=(0,1),\ \xi=(a,b);$$

   (4) $V=\mathbf{P}_n[x]$；$\varepsilon_1=1$，$\varepsilon_2=1+x$，
   $$\varepsilon_3=1+x+x^2,\cdots,\ \varepsilon_n=1+x+\cdots+x^{n-1},$$
   $$\xi=a_{n-1}x^{n-1}+a_{n-2}x^{n-2}+\cdots+a_1x+a_0.$$
5. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与 $\beta_1,\beta_2,\cdots,\beta_n$ 是 n 维线性空间 V 中两组向量，已知 $\beta_1,\beta_2,\cdots,\beta_n$ 可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出：

$$\beta_1 = a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{n1}\alpha_n$$
$$\beta_2 = a_{12}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{n2}\alpha_n$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\beta_n = a_{1n}\alpha_1 + a_{2n}\alpha_2 + \cdots + a_{nn}\alpha_n.$$

令

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

试证：(1) 如果 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一组基，那么，当且仅当矩阵 A 可逆时，$\beta_1,\beta_2,\cdots,\beta_n$ 也是 V 的一组基；

(2) 如果矩阵 A 可逆，那么，或者 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与 $\beta_1,\beta_2,\cdots,\beta_n$ 都是 V 的基；或者都不是.

8.4 基变换与坐标变换

在 n 维空间中，任意 n 个线性无关的向量都可取作这个线性空间的一组基. 显然，关于不同的基，同一个向量的坐标一般是不同的. 因此，在处理一些问题的时候，选择适当的基，以使所要讨论向量的坐标比较简单，是一个关键问题. 为此，首先需要知道同一个向量在两组基下的坐标之间有什么关系.

设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 及 $\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n'$ 是 n 维线性空间 V 的两组基，它们之间有下述关系：

$$\begin{cases} \varepsilon_1' = a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + \cdots + a_{n1}\varepsilon_n, \\ \varepsilon_2' = a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{n2}\varepsilon_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \varepsilon_n' = a_{1n}\varepsilon_1 + a_{2n}\varepsilon_2 + \cdots + a_{nn}\varepsilon_n. \end{cases} \tag{1}$$

再设 V 中向量 ξ 对于这两组基的坐标分别是 $(x_1,x_2,\cdots,x_n)$ 及 $(x_1',x_2',\cdots,x_n')$，即

$$\xi = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n$$
$$= x_1'\varepsilon_1' + x_2'\varepsilon_2' + \cdots + x_n'\varepsilon_n'. \tag{2}$$

下面求 $x_1, x_2, \cdots, x_n$ 与 $x_1', x_2', \cdots, x_n'$ 之间的关系.

将(1)式代入(2)式，得

$$\begin{aligned}\xi &= x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n \\ &= x_1'(a_{11}\varepsilon_1 + \cdots + a_{n1}\varepsilon_n) \\ &\quad + x_2'(a_{12}\varepsilon_1 + \cdots + a_{n2}\varepsilon_n) \\ &\quad + \cdots + x_n'(a_{1n}\varepsilon_1 + \cdots + a_{nn}\varepsilon_n) \\ &= (a_{11}x_1' + \cdots + a_{1n}x_n')\varepsilon_1 \\ &\quad + (a_{21}x_1' + \cdots + a_{2n}x_n')\varepsilon_2 \\ &\quad + \cdots + (a_{n1}x_1' + \cdots + a_{nn}x_n')\varepsilon_n.\end{aligned}$$

因为 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性无关的，所以 ξ 由它们线性表出的表法唯一，对应的系数应相等. 即得

$$\begin{cases} x_1 = a_{11}x_1' + a_{12}x_2' + \cdots + a_{1n}x_n', \\ x_2 = a_{21}x_1' + a_{22}x_2' + \cdots + a_{2n}x_n', \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = a_{n1}x_1' + a_{n2}x_2' + \cdots + a_{nn}x_n'. \end{cases} \tag{3}$$

这就是 ξ 关于这两组基的坐标间的关系. 为了便于记忆和应用，把(3)式写成矩阵的形式. 令

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

则(3)式可表成

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix}.$$

由于 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 也是 V 的一组基，所以 A 是可逆的（习题 8.3 第 6 题），故从上式可得

$$\begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} = A^{-1} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

下面介绍一种形式的记法. 这种记法不但写起来很方便，应用起来也很方便. 将向量

$$\xi = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n$$

写成

$$\xi = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

也就是把基写成一个 $1\times n$ 矩阵，把 ξ 在此基下的坐标写成一个 $n\times 1$ 矩阵，而把向量 ξ 看成这两个矩阵的乘积. 所以说这个记法是"形式"的，是由于式中把向量作为一个矩阵的元素，一般说来是没有意义的. 不过在应用的特殊情形下，这种约定的记法是不会出问题的.

相仿地，把式(1)写成

$$[\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'] = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

在利用形式表法来作计算之前，先来指出这种写法所具有的一些运算规律.

设 $\alpha_1, \alpha_2, \cdots, \alpha_n$; $\beta_1, \beta_2, \cdots, \beta_n$ 是数域 **P** 上线性空间 V 中两个向量组,则对于数域 **P** 中的数 $a_1, a_2, \cdots, a_n$; $b_1, b_2, \cdots, b_n$,有

(1) $$[\alpha_1, \alpha_2, \cdots, \alpha_n]\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}+[\alpha_1, \alpha_2, \cdots, \alpha_n]\begin{bmatrix}b_1\\b_2\\\vdots\\b_n\end{bmatrix}$$
$$=[\alpha_1, \alpha_2, \cdots, \alpha_n]\begin{bmatrix}a_1+b_1\\a_2+b_2\\\vdots\\a_n+b_n\end{bmatrix}.$$

(2) 如果向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性无关的,则
$$[\alpha_1, \alpha_2, \cdots, \alpha_n]\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}=[\alpha_1, \alpha_2, \cdots, \alpha_n]\begin{bmatrix}b_1\\b_2\\\vdots\\b_n\end{bmatrix}$$
$$\rightleftharpoons\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\\vdots\\b_n\end{bmatrix}.$$

对于数域 **P** 上的 n 阶矩阵 A,B,有

(3) 如果向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性无关的,则
$$[\alpha_1, \alpha_2, \cdots, \alpha_n]A=[\alpha_1, \alpha_2, \cdots, \alpha_n]B$$
$$\rightleftharpoons A=B.$$

(4) $$[\alpha_1, \alpha_2, \cdots, \alpha_n]A+[\alpha_1, \alpha_2, \cdots, \alpha_n]B$$
$$=[\alpha_1, \alpha_2, \cdots, \alpha_n](A+B).$$

(5) $([\alpha_1, \alpha_2, \cdots, \alpha_n]A)B=[\alpha_1, \alpha_2, \cdots, \alpha_n](AB)$.(这个等式当 B 是一个 $n\times 1$ 矩阵时也成立.)

(6) $[\alpha_1, \alpha_2, \cdots, \alpha_n]A+[\beta_1, \beta_2, \cdots, \beta_n]A$

$$=[\alpha_1+\beta_1,\ \alpha_2+\beta_2,\ \cdots,\ \alpha_n+\beta_n]A.$$

这些等式根据线性组合的性质，很容易直接验证，请读者自己证明.

现在回到本节所要解决的坐标变换的问题上来.

定义 12 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 及 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 是线性空间 V 的两组基. 并设它们之间有关系式(1)，则矩阵

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

称为由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 的**过渡矩阵**.

过渡矩阵一定是可逆的.

定理 7 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 及 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 是线性空间 V 的两组基，由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 的过渡矩阵是 A. ξ 是 V 中一个向量. 再设 ξ 对于基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 及基 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 的坐标分别为 $(x_1, x_2, \cdots, x_n)$ 及 $(x_1', x_2', \cdots, x_n')$，则

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix}, \quad 即 \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} = A^{-1} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{4}$$

证明：这个结论在前面已经证明过，现在用形式表达式再证一下.

因为由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 的过渡矩阵是 A，故有

$$[\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'] = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]A. \tag{5}$$

再根据ξ对于这两组基的坐标,有表式

$$\xi=[\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n]\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$$

$$=[\varepsilon_1',\ \varepsilon_2',\ \cdots,\ \varepsilon_n']\begin{bmatrix}x_1'\\x_2'\\\vdots\\x_n'\end{bmatrix}. \qquad (6)$$

把(5)式代入(6)式,得

$$\xi=[\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n]\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$$

$$=([\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n]A)\begin{bmatrix}x_1'\\x_2'\\\vdots\\x_n'\end{bmatrix}$$

$$=[\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n]\left(A\begin{bmatrix}x_1'\\x_2'\\\vdots\\x_n'\end{bmatrix}\right).$$

因为$\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n$是线性无关的,故得

$$\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=A\begin{bmatrix}x_1'\\x_2'\\\vdots\\x_n'\end{bmatrix}.$$

(4) 式称为对应于基变换(5)式(即(1)式)的坐标变换公式.

例 1 上节例 4 中,给了 $\mathbf{P}^n$ 的两组基:

$$\begin{aligned}\varepsilon_1 &= (1,0,\cdots,0), \\ \varepsilon_2 &= (0,1,\cdots,0), \\ &\cdots\cdots\cdots\cdots \\ \varepsilon_n &= (0,0,\cdots,1);\end{aligned} \quad \text{及} \quad \begin{aligned}\varepsilon_1' &= (1,0,\cdots,0), \\ \varepsilon_2' &= (1,1,\cdots,0), \\ &\cdots\cdots\cdots\cdots \\ \varepsilon_n' &= (1,1,\cdots,1).\end{aligned}$$

它们之间有下述关系:

$$\varepsilon_1' = \varepsilon_1,\ \varepsilon_2' = \varepsilon_1 + \varepsilon_2,\ \cdots,\ \varepsilon_n' = \varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_n.$$

因此,由基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到基 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 的过渡矩阵是

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

易知,$\alpha=(a_1,a_2,\cdots,a_n)$对于基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$的坐标为$(a_1,a_2,\cdots,a_n)$,所以,设 α 对于基 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$的坐标为$(a_1', a_2', \cdots, a_n')$,则有

$$\begin{bmatrix} a_1' \\ a_2' \\ \vdots \\ a_n' \end{bmatrix} = A^{-1} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$=\begin{bmatrix} a_1 - a_2 \\ a_2 - a_3 \\ \cdots\cdots\cdots \\ a_{n-1} - a_n \\ a_n \end{bmatrix}.$$

在计算向量在不同基下的坐标时，下述命题很有用.

命题 1 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$; $\eta_1, \eta_2, \cdots, \eta_n$ 及 $\gamma_1, \gamma_2, \cdots, \gamma_n$ 都是线性空间 V 的基.

(1) 如果由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\eta_1, \eta_2, \cdots, \eta_n$ 的过渡矩阵是 A，那么由 $\eta_1, \eta_2, \cdots, \eta_n$ 到 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的过渡矩阵是 A^{-1}；

(2) 如果由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\eta_1, \eta_2, \cdots, \eta_n$ 的过渡矩阵是 A；由 $\eta_1, \eta_2, \cdots, \eta_n$ 到 $\gamma_1, \gamma_2, \cdots, \gamma_n$ 的过渡矩阵是 B，则由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\gamma_1, \gamma_2, \cdots, \gamma_n$ 的过渡矩阵是 AB.

证明：(1) 由假设

$$[\eta_1, \eta_2, \cdots, \eta_n] = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]A,$$

因此

$$[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n] = [\eta_1, \eta_2, \cdots, \eta_n]A^{-1}.$$

即由 $\eta_1, \eta_2, \cdots, \eta_n$ 到 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的过渡矩阵是 A^{-1}.

(2) 由假设

$$[\eta_1, \eta_2, \cdots, \eta_n] = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]A,$$
$$[\gamma_1, \gamma_2, \cdots, \gamma_n] = [\eta_1, \eta_2, \cdots, \eta_n]B,$$

因此

$$\begin{aligned}[\gamma_1, \gamma_2, \cdots, \gamma_n] &= ([\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]A)B \\ &= [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n](AB).\end{aligned}$$

所以由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\gamma_1, \gamma_2, \cdots, \gamma_n$ 的过渡矩阵是 AB.

例 2 设 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 是线性空间 V 的一组基.

$$\begin{cases} \alpha_1 = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4, \\ \alpha_2 = \varepsilon_1 + \varepsilon_2 - \varepsilon_3 - \varepsilon_4, \\ \alpha_3 = \varepsilon_1 - \varepsilon_2 + \varepsilon_3 - \varepsilon_4, \\ \alpha_4 = \varepsilon_1 - \varepsilon_2 - \varepsilon_3 + \varepsilon_4; \end{cases} \quad \begin{cases} \beta_1 = \varepsilon_1 - \varepsilon_2 + 2\varepsilon_3 + \varepsilon_4, \\ \beta_2 = \varepsilon_1 \qquad - 2\varepsilon_3 - \varepsilon_4, \\ \beta_3 = \varepsilon_1 + 2\varepsilon_2 - \varepsilon_3, \\ \beta_4 = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + 2\varepsilon_4. \end{cases}$$

(1) 证明 α_1, α_2, α_3, α_4 及 β_1, β_2, β_3, β_4 都是 V 的基;

(2) 求由基 α_1, α_2, α_3, α_4 到基 β_1, β_2, β_3, β_4 的过渡矩阵;

(3) 求由基 α_1, α_2, α_3, α_4 到基 β_1, β_2, β_3, β_4 的坐标变换公式.

解 (1) 因为

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{vmatrix} = -16 \neq 0,$$

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

所以 α_1, α_2, α_3, α_4; β_1, β_2, β_3, β_4 都是 V 的基.

(2) 直接从定义求 α_1, α_2, α_3, α_4 到 β_1, β_2, β_3, β_4 的过渡矩阵,需要把每个 $\beta_i(i=1,2,3,4)$表成 α_1, α_2, α_3, α_4 的线性组合,这是比较麻烦的.下面应用命题 1 来解决这个问题.

因为由 ε_1, ε_2, ε_3, ε_4 到 α_1, α_2, α_3, α_4 的过渡矩阵是

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

由 ε_1, ε_2, ε_3, ε_4 到 β_1, β_2, β_3, β_4 的过渡矩阵是

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

所以由 α_1，α_2，α_3，α_4 到 ε_1，ε_2，ε_3，ε_4 的过渡矩阵是 A^{-1}；由 α_1，α_2，α_3，α_4 到 β_1，β_2，β_3，β_4 的过渡矩阵是 $A^{-1}B$：

$$A^{-1}B=\frac{1}{4}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\end{bmatrix}\cdot\begin{bmatrix}1 & 1 & 1 & 1\\ -1 & 0 & 2 & 1\\ 2 & -2 & -1 & 1\\ 1 & -1 & 0 & 2\end{bmatrix}$$

$$=\begin{bmatrix}\frac{3}{4} & -\frac{1}{2} & \frac{1}{2} & 1\frac{1}{4}\\ -\frac{3}{4} & 1 & 1 & -\frac{1}{4}\\ \frac{3}{4} & 0 & -\frac{1}{2} & -\frac{1}{4}\\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4}\end{bmatrix}.$$

(3) 设 ξ 是 V 中任一个向量. 如果它对于这两组基的坐标分别为(x_1,x_2,x_3,x_4)及(x_1', x_2', x_3', x_4')，则

$$\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}=\begin{bmatrix}\frac{3}{4} & -\frac{1}{2} & \frac{1}{2} & 1\frac{1}{4}\\ -\frac{3}{4} & 1 & 1 & -\frac{1}{4}\\ \frac{3}{4} & 0 & -\frac{1}{2} & -\frac{1}{4}\\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4}\end{bmatrix}\begin{bmatrix}x_1'\\ x_2'\\ x_3'\\ x_4'\end{bmatrix}$$

或

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{bmatrix} = (A^{-1}B)^{-1}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{3}{2} & -\frac{1}{2} \\ -\frac{3}{10} & \frac{1}{10} & -\frac{1}{10} & \frac{3}{2} \\ \frac{7}{10} & \frac{11}{10} & \frac{9}{10} & -\frac{3}{2} \\ \frac{1}{10} & -\frac{7}{10} & -\frac{13}{10} & \frac{3}{2} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

最后指出：基变换确定以后，坐标变换也就随之而确定了；反之，如果已经知道了两组基所对应的坐标之间的关系，那么这两组基之间的过渡矩阵也可以从坐标变换公式得出.

习 题 8.4

1. 在 $\mathbf{P}^4$ 中，求由基 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 到 $\eta_1, \eta_2, \eta_3, \eta_4$ 的过渡矩阵，并求向量 ξ 对于所指定的基的坐标：

(1) $\begin{cases} \varepsilon_1 = (1,0,0,0), \\ \varepsilon_2 = (0,1,0,0), \\ \varepsilon_3 = (0,0,1,0), \\ \varepsilon_4 = (0,0,0,1), \end{cases}$ $\begin{cases} \eta_1 = (2,1,0,1), \\ \eta_2 = (0,1,2,2), \\ \eta_3 = (-2,1,2,1), \\ \eta_4 = (1,3,1,2); \end{cases}$

$\xi=(1,1,1,1)$ 对于基 $\eta_1, \eta_2, \eta_3, \eta_4$ 的坐标；

(2) $\begin{cases} \varepsilon_1 = (1,2,-1,0), \\ \varepsilon_2 = (1,-1,1,1), \\ \varepsilon_3 = (-1,2,1,1), \\ \varepsilon_4 = (-1,-1,0,1), \end{cases}$ $\begin{cases} \eta_1 = (2,0,1,1), \\ \eta_2 = (0,1,1,2), \\ \eta_3 = (1,-1,2,3), \\ \eta_4 = (1,-3,2,0); \end{cases}$

$\xi=(0,1,1,2)$ 对于基 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 的坐标；

(3) $\begin{cases} \varepsilon_1 = (1,1,1,1,), \\ \varepsilon_2 = (1,1,-1,-1), \\ \varepsilon_3 = (1,-1,1,-1), \\ \varepsilon_4 = (1,-1,-1,1), \end{cases}$ $\begin{cases} \eta_1 = (1,2,3,1), \\ \eta_2 = (2,1,0,1), \\ \eta_3 = (1,-1,0,-1), \\ \eta_4 = (2,1,1,-2); \end{cases}$

$\xi=(a_1,a_2,a_3,a_4)$对于基 ε_1, ε_2, ε_3, ε_4 的坐标.

2. 已知 $P_4[x]$的一组基:

$$1, x, x^2, x^3$$

(1) 证明

$$1, 1+x, (1+x)^2, (1+x)^3$$

也是 $P_4[x]$的一组基;

(2) 求由基 $1,x,x^2,x^3$ 到基 $1,1+x,(1+x)^2,(1+x)^3$ 的过渡矩阵;

(3) 求由基 $1,1+x,(1+x)^2,(1+x)^3$ 到基 $1,x,x^2,x^3$ 的过渡矩阵;

(4) 求

$$a_3x^2 + a_2x^2 + a_1x + a_0$$

对于基 $1,1+x,(1+x)^2,(1+x)^3$ 的坐标.

3. 设 ε_1, ε_2, ε_3 是数域 P 上 3 维线性空间 V 的一组基. 已知 V 中向量 α 对于这组基的坐标为(x_1,x_2,x_3),

(1) 求 α 对于 $\varepsilon_2,\varepsilon_3,\varepsilon_1$ 的坐标;

(2) 求 α 对于基 $\varepsilon_1,k\varepsilon_2,\varepsilon_3$ 的坐标(k 是 P 中非零常数);

(3) 求 α 对于基 $\varepsilon_1,\varepsilon_2+k\varepsilon_3,\varepsilon_3$ 的坐标.

4. (1) 证明: E_{11}, E_{22}, E_{33}, $E_{12}+E_{21}$, $E_{13}+E_{31}$, $E_{23}+E_{32}$, $E_{12}-E_{21}$, $E_{13}-E_{31}$, $E_{23}-E_{32}$

是 $P^{3\times 3}$ 的一组基;

(2) 求 $\xi=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

对于这组基的坐标.

8.5 线性子空间

有时,一个线性空间的一个子集合也是一个线性空间. 例如,

$\mathbf{P}^{n\times n}$ 是数域 $\mathbf{P}$ 上一个线性空间，其中全体对称矩阵也组成数域 $\mathbf{P}$ 上的一个线性空间. 又如 $\mathbf{P}_n[x]$是 $\mathbf{P}[x]$的一个子集合，它们都是数域 $\mathbf{P}$ 上的线性空间. 还有，在通常的 3 维几何空间中，一个通过原点的平面上的向量对于加法与数量乘法组成一个 2 维实线性空间. 这就是子空间的概念.

定义 13 设 V 是数域 $\mathbf{P}$ 上一个线性空间，W 是 V 的一个非空子集合，如果 W 对于 V 的两种运算(加法与数量乘法)也构成一个线性空间，那么就称 W 为 V 的一个**线性子空间**，简称**子空间**.

下面分析线性空间 V 的一个非空子集合成为子空间的条件.

设 W 是 V 的一个非空子集. 因为 V 是线性空间，而 W 是 V 的一部分，所以对原有的运算，W 中的向量也满足线性空间定义中的(1)，(2)，(5)，(6)，(7)，(8)这 6 条规则. 为了使 W 成为一个线性空间，条件是要求 W 对 V 的运算是封闭的，并且规则(3)与(4)成立. 把这些条件列在下面，即

(1) 如果 W 包含向量 α,β，那么一定也包含 α,β 之和 $\alpha+\beta$；

(2) 如果 W 包含向量 α，那么对数域 $\mathbf{P}$ 中任一个数 k，W 一定也包含数量乘积 $k\alpha$；

(3) V 的零向量在 W 中；

(4) 如果 W 包含向量 α，那么 $-\alpha$ 也在 W 中.

不难看出，条件(3)与(4)可以从条件(2)推出. 只要令 $k=0$，-1，即可知零向量及负向量 $-\alpha$ 在 W 中. 因此得：

定理 8 数域 $\mathbf{P}$ 上线性空间 V 的非空子集 W 是 V 的一个子空间的必要充分条件是：W 对于 V 的两种运算是封闭的. 即

(1) 如 α,β 属于 W，则 $\alpha+\beta$ 也属于 W；

(2) 如 α 属于 W，则对 $\mathbf{P}$ 中任意数 k，$k\alpha$ 也属于 W.

条件(1)与(2)可以合并成：如果 α 与 β 属于 W，那么 α,β 的任意一个线性组合也属于 W.

这个条件还可推广成：

命题 1 设 W 是数域 $\mathbf{P}$ 上线性空间 V 的一个子空间. 那么，对于 W 中任一组向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 及 $\mathbf{P}$ 中任一组数 $k_1, k_2, \cdots, k_s$，线性组合

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

在 W 中.

根据定义，线性空间 V 本身是 V 的一个子空间；由 V 的单个零向量组成的集合也是 V 的子空间，叫做**零子空间**. 这两个子空间叫做**平凡子空间**，其他的子空间叫做**非平凡子空间**.

下面再来举几个例子.

例 1 $\mathbf{P}_n[x]$是 $\mathbf{P}[x]$的子空间. 当 $n \leqslant m$ 时，$\mathbf{P}_n[x]$是 $\mathbf{P}_m[x]$的子空间.

例 2 考虑系数在 $\mathbf{P}$ 中的齐次线性方程组

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

这个方程组的全部解构成 $\mathbf{P}$ 上的一个线性空间，是 $\mathbf{P}^n$ 的一个子空间，称为方程组(1)的**解空间**. 方程组(1)的任一个基础解系(如果有非零解的话)都是解空间的一组基，因此解空间的维数等于 $n-r$，r 是方程组(1)的系数矩阵的秩.

例 3 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是线性空间 V 中的一组向量. 这些向量的全部线性组合所成的集合是非空的，而且对两种线性运算是封闭的，因此是 V 的一个子空间. 这个子空间叫做由 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 生成的子空间，记作

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s).$$

例 3 给出了作线性空间的子空间的一个方法，而且在有限维线性空间中，任何一个子空间都可以由这个方法得到. 事实上，设

W 是 n 维线性空间 V 的一个子空间，再设 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 是 W 的一组基，于是 W 中任一个向量都可表成 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 的线性组合.因而

$$W = L(\alpha_1, \alpha_2, \cdots, \alpha_r).$$

定理 9 (1) 两个向量组生成相同的子空间的充分必要条件是：这两个向量组等价；

(2) $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 的维数等于向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的秩.

证明：(1) 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 及 $\beta_1, \beta_2, \cdots, \beta_t$ 是线性空间 V 中的两个向量组.如果

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s) = L(\beta_1, \beta_2, \cdots, \beta_t),$$

那么每个 $\alpha_i (i=1,2,\cdots,s)$ 看作 $L(\beta_1, \beta_2, \cdots, \beta_t)$ 中的向量，都可以由向量组 $\beta_1, \beta_2, \cdots, \beta_t$ 线性表出；而每个 $\beta_j (j=1,2,\cdots,t)$ 看作 $L(\alpha_1, \cdots, \alpha_s)$ 中的向量，都可由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性表出，所以这两个向量组等价.

反之，如果这两个向量组等价，那么由于 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 中的向量都是 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的线性组合，因此也都可以由向量组 $\beta_1, \beta_2, \cdots, \beta_t$ 线性表出，因而属于 $L(\beta_1, \beta_2, \cdots, \beta_t)$.这说明

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s) \subseteq L(\beta_1, \beta_2, \cdots, \beta_t).$$

同样可证，$L(\beta_1, \beta_2, \cdots, \beta_t)$ 中的向量都是 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的线性组合，因此

$$L(\beta_1, \beta_2, \cdots, \beta_t) \subseteq L(\alpha_1, \alpha_2, \cdots, \alpha_s),$$

所以

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s) = L(\beta_1, \beta_2, \cdots, \beta_t).$$

(2) 设向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的秩为 r，并设 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 是它的一个极大线性无关组，那么 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 与 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 等价，$L(\alpha_1, \cdots, \alpha_s)$ 中每个向量都可由向量组 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 线性表出.根据定理 6，$L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 的维数等于 r.

从定理 9 的证明可得：

推论 如果向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 可以由向量组 $\beta_1, \beta_2, \cdots, \beta_t$ 线性表出，则

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s) \subseteq L(\beta_1, \beta_2, \cdots, \beta_t).$$

因此，$L(\alpha_1, \alpha_2, \cdots, \alpha_s)$的维数小于或等于 $L(\beta_1, \beta_2, \cdots, \beta_t)$的维数.

由于线性空间 V 的子空间 W 是 V 的一个子集合，因此 W 中不可能存在比 V 中更多数目的线性无关的向量，所以，W 的维数小于或等于 V 的维数. 而且，当且仅当 $W=V$ 时，W 的维数才能等于 V 的维数.

下面的定理是以后经常要用到的.

定理 10 设 W 是 n 维线性空间 V 的一个 m 维子空间. $\alpha_1, \alpha_2, \cdots, \alpha_m$ 是 W 的一组基. 那么

(1) 向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 可以扩充为整个空间 V 的一组基. 也就是说，可以找到 V 中 $n-m$ 个向量 $\alpha_{m+1}, \cdots, \alpha_n$，使得

$$\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$$

是 V 的一组基.

(2) 设 α 是 V 中一个向量，则 $\alpha \in W$ 的充分必要条件是：α 关于基 $\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$ 的坐标为

$$(a_1, \cdots, a_m, 0, \cdots, 0),$$

即其中最后 $n-m$ 个坐标都等于 0.

证明 (1) 因为 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 是 W 的一组基，所以是线性无关的，它们可扩充成 V 的一组基.

(2) 如果 $\alpha \in W$，则 α 可表成 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 的线性组合：

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_m\alpha_m.$$

把上式写成

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_m\alpha_m + 0\alpha_{m+1} + \cdots + 0\alpha_n,$$

即知 α 对于基 $\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$ 的坐标为$(a_1, \cdots, a_m, 0,$

$\cdots, 0)$.

如果 α 对于基 $\alpha_1, \cdots, \alpha_m, \cdots, \alpha_n$ 的坐标如上，则根据坐标的定义，即得

$$\begin{aligned}\alpha &= a_1\alpha_1 + \cdots + a_m\alpha_m + 0\alpha_{m+1} + \cdots + 0\alpha_n \\ &= a_1\alpha_1 + \cdots + a_m\alpha_m \in W.\end{aligned}$$

习 题 8.5

1. 检验线性空间 V 的子集 W 是否是 V 的线性子空间. 如果 W 是子空间，求其维数和一组基：

   (1) $V=\mathbf{P}^4$, $W=\{(a_1, a_2, a_3, a_4) \mid a_1+a_2+a_3=a_4\}$；

   (2) $V=\mathbf{P}^4$, $W=\{(a_1, a_2, a_3, a_4) \mid a_1^2=a_2\}$；

   (3) $V=\mathbf{P}^{n\times n}$, $W=\{V$ 中反对称矩阵全体$\}$；

   (4) $V=\mathbf{R}^{n\times n}$（实数域上 n 阶方阵全体），
   $W=\{V$ 中正交矩阵全体$\}$；

   (5) $V=\mathbf{P}[x]$, $W=\{$数域 $\mathbf{P}$ 上 n 次多项式全体$\}$.

2. 在 $\mathbf{P}^{(4)}$ 中求出由向量 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 生成的线性子空间的维数和一维基：

   (1) $\alpha_1=(2,0,1,2)$, $\alpha_2=(-1,1,0,3)$,
   $\alpha_3=(0,2,1,8)$, $\alpha_4=(5,-1,2,1)$；

   (2) $\alpha_1=(2,1,-1,-2)$, $\alpha_2=(1,0,-3,2)$,
   $\alpha_3=(2,2,1,-1)$, $\alpha_4=(3,3,3,-5)$；

   (3) $\alpha_1=(2,1,3,-1)$, $\alpha_2=(1,-1,3,-1)$,
   $\alpha_3=(4,-1,9,-3)$, $\alpha_4=(1,5,-3,-1)$；

   (4) $\alpha_1=(2,1,3,1)$, $\alpha_2=(-1,1,2,3)$,
   $\alpha_3=(0,1,2,1)$, $\alpha_4=(1,1,2,-2)$.

3. 求齐次线性方程组

$$\begin{cases}3x_1 + 2x_2 - 5x_3 + 4x_4 = 0 \\ 3x_1 - 4x_2 + 11x_3 - 10x_4 = 0 \\ 3x_1 + 5x_2 - 13x_3 + 11x_4 = 0\end{cases}$$

   的解空间的维数和一组基.

4. 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$; $\beta_1, \beta_2, \cdots, \beta_t$ 都是线性空间 V 中的向量. 令

$$W_1 = L(\alpha_1, \alpha_2, \cdots, \alpha_s);$$
$$W_2 = L(\beta_1, \beta_2, \cdots, \beta_t);$$
$$W_3 = L(\alpha_1, \alpha_2, \cdots, \alpha_s, \beta_1, \beta_2, \cdots, \beta_t)$$

试证：$\dim W_3 \leqslant \dim W_1 + \dim W_2$.

5. 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性空间 V 的一组基，$\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 中 n 个向量，已知

$$(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A,$$

求证：$L(\alpha_1, \alpha_2, \cdots, \alpha_n)$的维数等于 A 的秩.

8.6 子空间的交与和

这一节介绍子空间的交与和、直和等概念以及其重要性质.

定理 11 设 W_1 和 W_2 都是数域 $\mathbf{P}$ 上线性空间 V 的子空间，用 $W_1 \cap W_2$ 表示 W_1 与 W_2 中全部公共元素的集合，那么 $W_1 \cap W_2$ 也是 V 的子空间.

证明：首先，由 $\mathbf{0} \in W_1, \mathbf{0} \in W_2$，得 $\mathbf{0} \in W_1 \cap W_2$，所以 $W_1 \cap W_2$ 是非空的. 其次，如果 $\alpha, \beta \in W_1 \cap W_2$，那么 $\alpha, \beta \in W_1, \alpha, \beta \in W_2$，由于 W_1 及 W_2 是子空间，所以 $\alpha + \beta \in W_1, \alpha + \beta \in W_2$，故 $\alpha + \beta \in W_1 \cap W_2$. 最后，对任意 $k \in \mathbf{P}$，有 $k\alpha \in W_1, k\alpha \in W_2$，所以 $k\alpha \in W_1 \cap W_2$. 根据定理 8，$W_1 \cap W_2$ 是 V 的一个子空间. ∎

例 1 设 W_1 是 $\mathbf{P}^{n\times n}$ 的由上三角矩阵全体组成的子空间，W_2 是 $\mathbf{P}^{n\times n}$ 中全部对称矩阵组成的子空间，则 $W_1 \cap W_2$ 是由全部对角矩阵组成的子空间. 如果再设 W_3 是 $\mathbf{P}^{n\times n}$ 中全部反对称矩阵组成的子空间，则 $W_1 \cap W_3, W_2 \cap W_3$ 都是零子空间.

例 2 在线性空间 $\mathbf{P}^n$ 中，用 W_1 与 W_2 分别表示齐次线性方程组

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

$$(a_{ij} \in \mathbf{P};\ i = 1,2,\cdots,s;\ j = 1,2,\cdots,n)$$

及

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n = 0, \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_{t1}x_1 + b_{t2}x_2 + \cdots + b_{tn}x_n = 0 \end{cases}$$

$$(b_{ij} \in \mathbf{P};\ i = 1,2,\cdots,t;\ j = 1,2,\cdots,n)$$

的解空间. 那么 $W_1 \cap W_2$ 就是齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = 0, \\ b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_{t1}x_1 + b_{t2}x_2 + \cdots + b_{tn}x_n = 0 \end{cases}$$

的解空间.

定理 12 设 W_1, W_2 是数域 $\mathbf{P}$ 上线性空间 V 的两个子空间. 用 W_1+W_2 表示全体形如 $\alpha_1+\alpha_2$, 其中 $\alpha_1 \in W_1$, $\alpha_2 \in W_2$, 的向量组成的集合, 那么 W_1+W_2 也是 V 的一个子空间.

证明 首先, W_1+W_2 显然是非空的. 因为 $\mathbf{0} \in W_1, \mathbf{0} \in W_2$, 所以 $\mathbf{0}=\mathbf{0}+\mathbf{0} \in W_1+W_2$. 其次, 如果 $\alpha, \beta \in W_1+W_2$, 那么 α, β 可表成

$$\alpha = \alpha_1 + \alpha_2, \qquad \alpha_1 \in W_1,\ \alpha_2 \in W_2;$$
$$\beta = \beta_1 + \beta_2, \qquad \beta_1 \in W_1,\ \beta_2 \in W_2.$$

于是

$$\alpha+\beta=(\alpha_1+\alpha_2)+(\beta_1+\beta_2)=(\alpha_1+\beta_1)+(\alpha_2+\beta_2).$$

并且对数域 $\mathbf{P}$ 中的数 k 有

$$k\alpha=k\alpha_1+k\alpha_2.$$

由于 W_1,W_2 都是子空间,所以

$$\alpha_1+\beta_1\in W_1,\qquad k\alpha_1\in W_1,$$
$$\alpha_2+\beta_2\in W_2,\qquad k\alpha_2\in W_2.$$

于是由 W_1+W_2 的定义,有

$$\alpha+\beta=(\alpha_1+\beta_1)+(\alpha_2+\beta_2)\in W_1+W_2;$$
$$k\alpha=k\alpha_1+k\alpha_2\in W_1+W_2.$$

所以 W_1+W_2 是 V 的一个子空间. |

定义 14 设 W_1,W_2 是线性空间 V 的子空间,$W_1\cap W_2$ 称为 W_1 与 W_2 的**交**;W_1+W_2 称为 W_1 与 W_2 的**和**.

子空间的交与和可以看成由已知子空间构作新的子空间的一种方法.

例 3 如果 $\alpha_1,\alpha_2,\cdots,\alpha_s;\beta_1,\beta_2,\cdots,\beta_t$ 都是线性空间 V 中的向量,那么有

$$L(\alpha_1,\alpha_2,\cdots,\alpha_s)+L(\beta_1,\beta_2,\cdots,\beta_t)$$
$$=L(\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t).$$

例 4 仍用 W_1,W_2,W_3 表示例 1 中 $\mathbf{P}^{n\times n}$ 的子空间,那么,由于每个 n 阶矩阵都可表示成一个对称矩阵及一个反对称矩阵的和,所以

$$W_2+W_3=\mathbf{P}^{n\times n}.$$

同样可证

$$W_1+W_2=\mathbf{P}^{n\times n};\ W_1+W_3=\mathbf{P}^{n\times n}.$$

例 5 求 $L(\alpha_1,\alpha_2)$ 及 $L(\beta_1,\beta_2)$ 的交与和;并分别求它们的维数与一组基,其中

$$\alpha_1=(2,5,-1,-5,), \qquad \beta_1=(2,0,-1,2),$$
$$\alpha_2=(-1,2,-2,-3), \qquad \beta_2=(1,3,2,-4).$$

解 $L(\alpha_1,\alpha_2)+L(\beta_1,\beta_2)=L(\alpha_1,\alpha_2,\beta_1,\beta_2)$. 因为向量组 $\alpha_1,\alpha_2,\beta_1,\beta_2$ 的秩等于 3；$\alpha_1,\alpha_2,\beta_1$ 是它的一个极大线性无关组，所以 $L(\alpha_1,\alpha_2)+L(\beta_1,\beta_2)$的维数为 3；$\alpha_1,\alpha_2,\beta$ 是它的一组基.

现在求 $L(\alpha_1,\alpha_2)\cap L(\beta_1,\beta_2)$以及它的维数与基. 首先来看一个向量 α 属于 $L(\alpha_1,\alpha_2)\cap L(\beta_1,\beta_2)$的条件. 如果 $\alpha\in L(\alpha_1,\alpha_2)\cap L(\beta_1,\beta_2)$，那么 α 既属于 $L(\alpha_1,\alpha_2)$又属于 $L(\beta_1,\beta_2)$. 故有常数 $k_1,k_2;l_1,l_2$ 使得

$$\alpha=k_1\alpha_1+k_2\alpha_2=l_1\beta_1+l_2\beta_2.$$

于是

$$k_1\alpha_1+k_2\alpha_2-l_1\beta_1-l_2\beta_2=0.$$

比较分量，得

$$\begin{cases} 2k_1-k_2-2l_1-l_2=0, \\ 5k_1+2k_2-3l_2=0, \\ -k_1-2k_2+l_1-2l_2=0, \\ -5k_1-3k_2-2l_1+4l_2=0. \end{cases}$$

解方程组，得基础解系$(1,-1,1,1)$. 即

$$k_1=1,\ k_2=-1,\ l_1=1,\ l_2=1$$

得一个向量

$$\alpha_0=\alpha_1-\alpha_2=\beta_1+\beta_2=(3,3,1,-2).$$

这说明，如果 $\alpha=L(\alpha_1,\alpha_2)\cap L(\beta_1,\beta_2)$，那么 α 必须是 α_0 的倍数. 当然，这样的向量一定在 $L(\alpha_1,\alpha_2)\cap L(\beta_1,\beta_2)$中. 于是

$$L(\alpha_1,\alpha_2)\cap L(\beta_1,\beta_2)=L(\alpha_0)$$

它的维数为 1，$\alpha_0=(3,3,1,-2)$是它的一组基.

关于子空间的交与和的维数，有下述重要定理：

定理 13（维数公式） 如果 W_1,W_2 是线性空间 V 的两个子空

间，那么

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$$

证明：设 $W_1, W_2, W_1 \cap W_2$ 的维数分别为 n_1, n_2, m，则因为 $W_1 \cap W_2 \subseteq W_1$，$W_1 \cap W_2 \subseteq W_2$，所以

$$m \leqslant n_1, \qquad m \leqslant n_2$$

如果 $m \neq 0$，取 $W_1 \cap W_2$ 的一组基

$$\alpha_1, \alpha_2, \cdots, \alpha_m.$$

把它扩充成 W_1 的一组基：

$$\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1-m}.$$

把它扩充成 W_2 的一组基：

$$\alpha_1, \cdots, \alpha_m, \gamma_1, \cdots, \gamma_{n_2-m}.$$

下面来证明

$$\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1-m}, \gamma_1, \cdots, \gamma_{n_2-m}$$

是 $W_1 + W_2$ 的一组基.

因为

$$W_1 = L(\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1-m}),$$

$$W_2 = L(\alpha_1, \cdots, \alpha_m, \gamma_1, \cdots, \gamma_{n_2-m}),$$

所以

$$W_1 + W_2 = L(\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1-m}, \gamma_1, \cdots, \gamma_{n_2-m}).$$

因此，只要证明 $\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1-m}, \gamma_1, \cdots, \gamma_{n_2-m}$ 是线性无关的. 设

$$k_1\alpha_1 + \cdots + k_m\alpha_m + l_1\beta_1 + \cdots + l_{n_1-m}\beta_{n_1-m} + p_1\gamma_1 + \cdots + p_{n_2-m}\gamma_{n_2-m} = \mathbf{0}.$$

令

$$\alpha = k_1\alpha_1 + \cdots + k_m\alpha_m + l_1\beta_1 + \cdots + l_{n_1-m}\beta_{n_1-m} = -p_1\gamma_1 - \cdots - p_{n_2-m}\gamma_{n_2-m}. \tag{1}$$

于是由第一个等号知 $\alpha_1\in W_1$；由第二个等号知 $\alpha\in W_2$. 因此，$\alpha\in W_1\cap W_2$，α 可以由 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性表出. 设

$$\alpha = q_1\alpha_1 + \cdots + q_m\alpha_m,$$

则

$$q_1\alpha_1 + \cdots + q_m\alpha_m = -p_1\gamma_1 - \cdots - p_{n_2-m}\gamma_{n_2-m},$$

即

$$q_1\alpha_1 + \cdots + q_m\alpha_m + p_1\gamma_1 + \cdots + p_{n_2-m}\gamma_{n_2-m} = \mathbf{0}.$$

由于 $\alpha_1,\cdots,\alpha_m,\gamma_1,\cdots,\gamma_{n_2-m}$ 是线性无关的，所以必有

$$q_1 = \cdots = q_m = p_1 = \cdots = p_{n_2-m} = 0.$$

再由(1)式得

$$k_1\alpha_1 + \cdots + k_m\alpha_m + l_1\beta_1 + \cdots + l_{n_1-m}\beta_{n_1-m} = \mathbf{0}.$$

又因 $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_{n_1-m}$ 是线性无关的，所以

$$k_1 = \cdots = k_m = l_1 = \cdots = l_{n_1-m} = 0.$$

这就证明了 $\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1-m}, \gamma_1,\cdots,\gamma_{n_2-m}$ 是线性无关的，因此它们是 W_1+W_2 的一组基. 由此得

$$\begin{aligned}\dim(W_1 + W_2) &= n_1 + n_2 - m \\ &= \dim W_1 + \dim W_2 - \dim(W_1\cap W_2),\end{aligned}$$

维数公式成立.

如果 $m=0$，即 $W_1\cap W_2=\{\mathbf{0}\}$. 取 W_1 的一组基 $\beta_1, \beta_2, \cdots, \beta_{n_1}$ 和 W_2 的一组基 $\gamma_1, \gamma_2, \cdots, \gamma_{n_2}$，同样可证 $\beta_1,\cdots,\beta_{n_1}, \gamma_1,\cdots,\gamma_{n_2}$ 是 W_1+W_2 的一组基，仍可得到维数公式. |

例 6 例 1 中的 $W_i(i=1,2,3)$ 及其交与和的维数如下：

$$\dim W_1 = \frac{n(n+1)}{2},\ \dim W_2 = \frac{n(n+1)}{2},\ \dim W_3 = \frac{n(n-1)}{2};$$

$$\dim(W_1 + W_2) = n^2, \qquad \dim(W_1\cap W_2) = n;$$

$$\dim(W_1 + W_3) = n^2, \qquad \dim(W_1\cap W_3) = 0;$$

$$\dim(W_2+W_3)=n^2,\quad \dim(W_2\cap W_3)=0.$$

这些数据给出维数公式的一些具体例子.

例 5 也给出了维数公式的一个例子.

维数公式有下述重要推论：

命题 1 如果 n 维线性空间 V 的两个子空间 W_1,W_2 的维数之和大于 n，则 W_1 与 W_2 必有非零的公共向量.

证明：由维数公式

$$\dim(W_1\cap W_2)=\dim W_1+\dim W_2-\dim(W_1+W_2)$$

又将假设

$$\dim W_1+\dim W_2>n$$

代入上式，得

$$\dim(W_1\cap W_2)>n-\dim(W_1+W_2).$$

因为 W_1+W_2 是 V 的子空间，它的维数小于或等于 n，所以

$$\dim(W_1\cap W_2)>0.$$

于是 $W_1\cap W_2$ 不是零子空间，$W_1\cap W_2$ 中必有非零向量，即 W_1 与 W_2 有非零公共向量.

根据定义，显然有

$$(W_1\cap W_2)\subseteq W_1,\qquad (W_1\cap W_2)\subseteq W_2;$$

$$(W_1+W_2)\supseteq W_1,\qquad (W_1+W_2)\supseteq W_2.$$

不仅如此，我们还有：

命题 2 设 W_1,W_2 都是线性空间 V 的子空间，则

(1) $W_1\cap W_2$ 是 V 的既包含在 W_1 内又包含在 W_2 内的最大的子空间. 即

$$(W_1\cap W_2)\subseteq W_1,\qquad (W_1\cap W_2)\subseteq W_2.$$

并且，如果 W 是 V 的子空间，则由

$$W\subseteq W_1,\qquad W\subseteq W_2$$

可推出

$$W \subseteq (W_1 \cap W_2).$$

(2) W_1+W_2 是 V 的既包含 W_1 又包含 W_2 的最小的子空间，即

$$(W_1 + W_2) \supseteq W_1, \qquad (W_1 + W_2) \supseteq W_2.$$

并且，如果 W 是 V 的子空间，则由

$$W \supseteq W_1, \qquad W \supseteq W_2,$$

可推出

$$W \supseteq (W_1 + W_2).$$

命题 2 可以从线性子空间的定义及子空间的交与和的定义直接得出.

习　题　8.6 (1)

1. 求由向量 α_1，α_2 生成的子空间 W_1 及 β_1，β_2(或 β_1，β_2，β_3)生成的子空间 W_2 的交与和，并分别求它们的维数和一组基.

(1) $\alpha_1 = (1,1,0,0)$, $\qquad \beta_1 = (0,0,1,1,)$,

$\alpha_2 = (0,1,1,1)$, $\qquad \beta_2 = (1,1,1,0)$;

(2) $\alpha_1 = (2,0,1,3,-1)$, $\qquad \beta_1 = (1,1,0,-1,1)$,

$\alpha_2 = (0,-2,1,5,-3)$, $\qquad \beta_2 = (1,-3,2,0,5)$;

(3) $\alpha_1 = (2,5,-1,-5)$, $\qquad \beta_1 = (1,2,-1,-2)$,

$\alpha_2 = (-1,2,-2,3)$, $\qquad \beta_2 = (3,1,-1,1)$,

$\beta_3 = (-1,0,1,-1)$.

2. 设线性空间 V 中向量 α_1，α_2，α_3 满足

$$k_1\alpha_1 + k_2\alpha_2 + k_3\alpha_3 = 0$$

其中 $k_1k_2 \neq 0$. 证明

$$L(\alpha_1,\alpha_3) = L(\alpha_2,\alpha_3)$$

3. 设 W_1,W_2,W_3 都是线性空间 V 的子空间. 证明下列条件都是等价的：

(1) $W_1 \subseteq W_2$;

(2) $W_1 \cap W_2 = W_1$;

(3) $W_1 + W_2 = W_2$.

下面介绍子空间的直和.子空间的直和是子空间的和的一个特殊情形,是一个很重要的概念,将在下一章讨论线性变换时用到.

定义 15 设 W_1,W_2 是线性空间 V 的两个子空间.如果和 W_1+W_2 中每个向量 α 表成

$$\alpha=\alpha_1+\alpha_2 \quad (\alpha_i\in W_i,\ i=1,2)$$

的方法是唯一的,那么和 W_1+W_2 称为**直和**,记作 $W_1\dot{+}W_2$.

如在例 4 中,对于 $\mathbf{P}^{n\times n}$ 的子空间 W_1,W_2,W_3 有

$$W_1+W_2=W_1+W_3=W_2+W_3=V.$$

由于任一个矩阵表成一个对称矩阵与一个反对称矩阵之和的方法是唯一的,所以 W_2+W_3 为直和.但是一个矩阵可以用两种不同的方法表示成一个上三角矩阵与一个对称矩阵之和,例如

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

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

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

所以 W_1+W_2 不是直和.请读者检验一下,此例中的 W_1+W_3 是否是直和.

下面的一些定理给出了两个子空间的和是直和的一些充分必要条件，因此也就给出了一些简单的判别子空间的和是否是直和的方法.

定理 14 线性空间 V 的两个子空间 W_1, W_2 之和 W_1+W_2 是直和的充分必要条件是：V 中零向量表示成

$$\mathbf{0} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2, \qquad \boldsymbol{\alpha}_1 \in W_1, \boldsymbol{\alpha}_2 \in W_2$$

的表法是唯一的. 即由

$$\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 = \mathbf{0}, \qquad \boldsymbol{\alpha}_1 \in W_1, \boldsymbol{\alpha}_2 \in W_2$$

可推出

$$\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \mathbf{0}.$$

证明：必要性是显然的，可以从直和的定义直接得出. 下面用反证法证明充分性. 如果 W_1+W_2 不是直和，则 W_1+W_2 中有某个向量 $\boldsymbol{\alpha}$ 有两种不同的表法：

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 \qquad (\boldsymbol{\alpha}_1 \in W_1, \boldsymbol{\alpha}_2 \in W_2),$$
$$\boldsymbol{\alpha} = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 \qquad (\boldsymbol{\beta}_1 \in W_1, \boldsymbol{\beta}_2 \in W_2),$$

其中，或者 $\boldsymbol{\alpha}_1 \neq \boldsymbol{\beta}_1$，或者 $\boldsymbol{\alpha}_2 \neq \boldsymbol{\beta}_2$. 把两式相减，得

$$\mathbf{0} = (\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1) + (\boldsymbol{\alpha}_2 - \boldsymbol{\beta}_2),$$

其中

$$\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1 \in W_1, \boldsymbol{\alpha}_2 - \boldsymbol{\beta}_2 \in W_2.$$

而且 $\boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1$，$\boldsymbol{\alpha}_2 - \boldsymbol{\beta}_2$ 不全为零，因此零向量的表示法也不唯一. |

定理 15 W_1+W_2 是直和的充分必要条件是

$$W_1 \cap W_2 = \{\mathbf{0}\}.$$

证明：必要性. 设 W_1+W_2 是直和，任取 $\boldsymbol{\alpha} \in W_1 \cap W_2$，则零向量可表成

$$\mathbf{0} = \boldsymbol{\alpha} + (-\boldsymbol{\alpha}) \quad (\boldsymbol{\alpha} \in W_1, -\boldsymbol{\alpha} \in W_2).$$

故由定理 14，得 $\boldsymbol{\alpha} = \mathbf{0}$. 因此，$W_1 \cap W_2 = \{\mathbf{0}\}$.

充分性. 假设 $W_1 \cap W_2 = \{\mathbf{0}\}$，如果零向量可表成

$$\mathbf{0} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2, \ \boldsymbol{\alpha}_1 \in W_1, \boldsymbol{\alpha}_2 \in W_2,$$

则

$$\alpha_1 = -\alpha_2 \in W_2.$$

所以

$$\alpha_1 \in W_1 \cap W_2.$$

由假设 $W_1 \cap W_2 = \{\mathbf{0}\}$. 得 $\alpha_1 = \mathbf{0}$. 于是 $\alpha_2 = \mathbf{0} - \alpha_1 = \mathbf{0}$. 这说明零向量的表法是唯一的. 于是应用定理 14 知 $W_1 + W_2$ 为直和. |

由定理 15 及维数公式可得:

定理 16 $W_1 + W_2$ 是直和的充分必要条件是

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2.$$

因此,如果 $W_1 + W_2$ 是直和,在 W_1 中取一组基

$$\alpha_1, \alpha_2, \cdots, \alpha_{n_1};$$

在 W_2 中取一组基

$$\beta_1, \beta_2, \cdots, \beta_{n_2}.$$

那么

$$\alpha_1, \cdots, \alpha_{n_1}, \beta_1, \cdots, \beta_{n_2}$$

就是 $W_1 + W_2$ 的一组基.

特别地,如果线性空间 V 可表成子空间 W_1 及 W_2 的直和

$$V = W_1 \dotplus W_2,$$

那么,如果取 W_1 的一组基 $\alpha_1, \cdots, \alpha_{n_1}$,取 W_2 的一组基 $\beta_1, \cdots, \beta_{n_2}$,则 $\alpha_1, \cdots, \alpha_{n_1}, \beta_1, \cdots, \beta_{n_2}$就是 V 的一组基. 对于这组基,W_1 中向量的坐标为$(a_1, \cdots, a_{n_1}, 0, \cdots, 0)$;$W_2$ 中向量的坐标为$(0, \cdots, 0, b_1, \cdots, b_{n_2})$.

定理 17 设 W_1 是 n 维线性空间 V 的一个子空间,则一定存在 V 的一个子空间 W_2,使

$$V = W_1 \dotplus W_2.$$

证明: 如果 $W_1 = V$,则取 $W_2 = \{\mathbf{0}\}$;如果 $W_1 = \{\mathbf{0}\}$,则取 $W_2 = V$. 如果$\{\mathbf{0}\} \subset W_1 \subset V$,取 W_1 的一组基 $\alpha_1, \alpha_2, \cdots, \alpha_m$, $(m<n)$把

它扩充成 V 的一组基 $\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$. 令

$$W_2 = L(\alpha_{m+1}, \cdots, \alpha_n)$$

则

$$V = W_1 + W_2. \quad \mid$$

子空间的交与和的概念可以推广到几个子空间的情况.

设 $W_1, W_2, \cdots, W_s$ 是数域 $\mathbf{P}$ 上线性空间 V 的 $s(s\geqslant 2)$ 个线性子空间. 用 $W_1 \cap W_2 \cap \cdots \cap W_s$ 表示 $W_1, W_2, \cdots, W_s$ 的全部公共元素组成的集合. 显然,由 $\mathbf{0} \in W_i (i=1,2,\cdots,s)$ 可得

$$\mathbf{0} \in (W_1 \cap W_2 \cap \cdots \cap W_s).$$

因此 $W_1 \cap W_2 \cap \cdots \cap W_s$ 不是空集. 如果 V 中两个向量属于 $W_1 \cap W_2 \cap \cdots \cap W_s$,则由 $W_1 \cap W_2 \cap \cdots \cap W_s$ 的定义,

$$\alpha, \beta \in W_i, \qquad i = 1,2,\cdots,s.$$

由于 $W_i (1\leqslant i\leqslant s)$ 都是 V 的线性子空间,因此

$$\alpha + \beta \in W_i, \qquad i = 1,2,\cdots,s.$$

并且对 $\mathbf{P}$ 中任意的数 k,有

$$k\alpha \in W_i, \qquad i = 1,2,\cdots,s.$$

从而

$$\alpha + \beta \in (W_1 \cap W_2 \cap \cdots \cap W_s);$$
$$k\alpha \in (W_1 \cap W_2 \cap \cdots \cap W_s).$$

因此 $W_1 \cap W_2 \cap \cdots \cap W_s$ 是 V 的一个子空间.

再用 $W_1 + W_2 + \cdots + W_s$ 表示 V 中形如

$$\alpha_1 + \alpha_2 + \cdots + \alpha_s$$

的元素的全体所成的集合,其中

$$\alpha_i \in W_i, \qquad i = 1,2,\cdots,s.$$

则同样可证 $W_1 + W_2 + \cdots + W_s$ 也是 V 的子空间.

定义 14′ 设 $W_1, W_2, \cdots, W_s (s\geqslant 2)$ 都是 V 的子空间. 则子空间 $W_1 \cap W_2 \cap \cdots \cap W_s$ 称为子空间 $W_1, W_2, \cdots, W_s$ 的**交**;子空间 W_1

$+W_2+\cdots+W_s$ 称为子空间 $W_1,W_2,\cdots,W_s$ 的**和**.

两个子空间 W_1,W_2 的交与和是上述定义当 $s=2$ 时的特殊情形. 与 $s=2$ 时一样,关于 s 个子空间的交与和也有类似于命题 2 的结论:

命题 2′ 设 $W_1,W_2,\cdots,W_s(s\geqslant 2)$ 都是线性空间 V 的线性子空间,则

(1) $W_1\cap W_2\cap\cdots\cap W_s$ 是 V 的包含于每个 $W_i(i=1,2,\cdots,s)$ 内的最大的子空间,即

$$(W_1\cap W_2\cap\cdots\cap W_s)\subseteq W_i,\qquad i=1,2,\cdots,s.$$

并且由

$$W\subseteq W_i,\qquad i=1,2,\cdots,s$$

可推出

$$W\subseteq(W_1\cap W_2\cap\cdots\cap W_s).$$

(2) $W_1+W_2+\cdots+W_s$ 是 V 的同时包含每个 $W_i(i=1,2,\cdots,s)$ 的最小的子空间,即

$$(W_1+W_2+\cdots+W_s)\supseteq W_i\quad(i=1,2,\cdots,s).$$

并且由

$$W\supseteq W_i,\qquad i=1,2,\cdots,s$$

可推出

$$W\supseteq(W_1+W_2+\cdots+W_s).$$

例 7 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是 4 维线性空间 V 的一组基. 令

$$W_1=L(\varepsilon_2,\varepsilon_3,\varepsilon_4),$$
$$W_2=L(\varepsilon_1,\varepsilon_3,\varepsilon_4),$$
$$W_3=L(\varepsilon_1,\varepsilon_2,\varepsilon_4),$$
$$W_4=L(\varepsilon_1,\varepsilon_2,\varepsilon_3),$$

则

$$W_1\cap W_2=L(\varepsilon_3,\varepsilon_4),$$

$$W_1 \cap W_2 \cap W_3 = L(\varepsilon_4),$$
$$W_2 \cap W_3 \cap W_4 = L(\varepsilon_1),$$
$$W_1 \cap W_2 \cap W_3 \cap W_4 = \{\mathbf{0}\}.$$

而

$$W_1 + W_2 = W_1 + W_3 = W_1 + W_2 + W_3 + W_4 = V.$$

例 8 仍设 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 是例 7 中线性空间 V 的一组基. 再令

$$V_i = L(\varepsilon_i), \qquad i = 1,2,3,4,$$

则

$$V_i \cap V_j = \{\mathbf{0}\} \qquad (i \neq j,\ i,j = 1,2,3,4),$$
$$V_1 + V_2 + V_3 = W_4,$$
$$V_1 + V_3 + V_4 = W_2,$$
$$V_1 + V_2 + V_3 + V_4 = V.$$

例 9 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 n 维线性空间 V 中 n 个线性无关的向量. 令

$$V_i = L(\alpha_i) \quad (i = 1,2,\cdots,n),$$

则

$$V = V_1 + V_2 + \cdots + V_n.$$

如果令

$$W_1 = L(\alpha_1, \alpha_2, \cdots, \alpha_m),$$
$$W_2 = L(\alpha_{m+1}, \cdots, \alpha_n) \qquad (0 < m < n),$$

那么

$$V = W_1 + W_2.$$

其中

$$\dim W_1 = m, \qquad \dim W_2 = n - m.$$

例 9 说明一个一般事实：任一个 n 维线性空间可表示成 n 个一维子空间的和；一个 n 维线性空间可表示成一个 m 维子空间及一个 $n-m$ 维子空间的和.

子空间的直和也可推广到多个子空间的情况.

定义 15′ 设 $W_1, W_2, \cdots, W_s$ 是线性空间 V 的 $s(s\geqslant 2)$ 个子空间. 如果它们的和 $W_1+W_2+\cdots+W_s$ 中每个向量 α 表示成

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_s \quad (\alpha_i \in W_i,\ i = 1,2,\cdots,s)$$

的方法是唯一的,那么 $W_1+\cdots+W_s$ 称为**直和**,记作 $W_1\dot{+}\cdots\dot{+}W_s$.

同两个子空间的情况一样,多个子空间的直和也有下列一些判别方法:

命题 3 下列一些条件都是 $W_1+\cdots+W_s(s\geqslant 2)$ 是直和的充分必要条件:

(1) 零向量表示成

$$\mathbf{0} = \alpha_1 + \alpha_2 + \cdots + \alpha_s \quad (\alpha_i \in W_i,\ i = 1,2,\cdots,s)$$

的方法是唯一的,即由上式可推出

$$\alpha_1 = \alpha_2 = \cdots = \alpha_s = \mathbf{0}.$$

(2) $W_i \cap (\sum\limits_{j\neq i} W_j) = \{\mathbf{0}\}$.

(3) $\dim(W_1 + W_2 + \cdots + W_s) = \sum\limits_{i=1}^{s}\dim W_i$.

根据命题 3 可知,对于 s 个线性子空间,与 $s=2$ 的情况一样,如果 $W_1\dot{+}\cdots\dot{+}W_s$ 是直和,在 $W_i(i=1,2,\cdots,s)$ 中各取一组基

$$\alpha_{i1},\ \alpha_{i2},\ \cdots,\ \alpha_{in_i},$$

那么

$$\alpha_{11},\cdots,\alpha_{1n_1},\alpha_{21},\cdots,\alpha_{2n_2},\cdots,\alpha_{s1},\cdots,\alpha_{sn_s}$$

就是 $W_1\dot{+}W_2\dot{+}\cdots\dot{+}W_s$ 的一组基.

已知:任一个 n 维线性空间可以表示成 n 个一维子空间的和. 根据命题 3,这个结果可以改进成:任一个 n 维线性空间可以表示成 n 个一维子空间的直和.

习　题　8.6(2)

1. 设 $\alpha_1,\alpha_2,\alpha_3$ 都是 $\mathbf{P}^4$ 中的向量：

$$\alpha_1=(1,-1,0,2),$$
$$\alpha_2=(2,3,-1,1),$$
$$\alpha_3=(1,9,-2,-4).$$

求 $\mathbf{P}^4$ 的一个子空间 W，使

$$\mathbf{P}^4=W\dotplus L(\alpha_1,\alpha_2,\alpha_3)$$

2. 设 W_1,W_2 分别是齐次方程组

$$x_1+x_2+\cdots+x_n=0$$

及

$$\begin{cases} x_1-x_2 & =0 \\ \quad x_2-x_3 & =0 \\ \quad\cdots\cdots & \\ \quad\quad x_{n-1}-x_n & =0 \end{cases}$$

的解空间，求 W_1,W_2 并证 $\mathbf{P}^n=W_1\dotplus W_2$.

3. 设 V 是数域 $\mathbf{P}$ 上全体 n 阶准对角矩阵

$$\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{bmatrix}$$

所成的线性空间，其中 A_i 为 n_i 阶矩阵，$i=1,2,\cdots,s$；$n_1,n_2,\cdots,n_s$ 是确定的正整数，$n_1+n_2+\cdots+n_s=n$. 用 W_i 表示 V 中全体形如

$$\begin{bmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & A_i & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \text{第 } i \text{ 个}$$

的准对角矩阵所成的子空间，其中 A_i 为 n_i 阶矩阵，$i=1,2,\cdots,s$，求证：

$$V=W_1\dotplus W_2\dotplus\cdots\dotplus W_s.$$

4. 求证：如果 $W=W_1\dotplus W_2$，$W_2=V_1\dotplus V_2$，则

$$W = W_1 + V_1 + V_2.$$

8.7 线性空间的同构

设 V 是数域 $\mathbf{P}$ 上的一个 n 维线性空间，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组基. 对于这组基，V 中每个向量都有确定的坐标. 向量的坐标是一个行向量，因此也可以看成 $\mathbf{P}^n$ 中的元素. 如果 V 中的向量 α,β 的坐标分别是$(a_1,a_2,\cdots,a_n)$及$(b_1,b_2,\cdots,b_n)$，那么

$$\begin{aligned}\alpha+\beta &= (a_1\varepsilon_1+\cdots+a_n\varepsilon_n)+(b_1\varepsilon_1+\cdots+b_n\varepsilon_n)\\ &= (a_1+b_1)\varepsilon_1+\cdots+(a_n+b_n)\varepsilon_n;\end{aligned}$$

$$k\alpha = k(a_1\varepsilon_1+\cdots+a_n\varepsilon_n) = ka_1\varepsilon_1+\cdots+ka_n\varepsilon_n.$$

所以 $\alpha+\beta$ 和 $k\alpha$ 的坐标分别为

$$(a_1+b_1,\cdots,a_n+b_n) = (a_1,\cdots,a_n)+(b_1,\cdots,b_n)$$

及

$$(ka_1,\cdots,ka_n) = k(a_1,\cdots,a_n).$$

这说明：在 V 中取定一组基后，V 的向量可以用坐标表示，而且向量的运算也可以归结为它们的坐标的运算，因而线性空间的讨论可以归结为向量空间 $\mathbf{P}^n$ 的讨论. 为了确切地说明这个事实，需要用到同构的概念. 为此，首先介绍映射的概念及有关性质，这些内容在以后也要用到.

定义 16 设 M 与 M'是两个集合. 所谓集合 M 到 M'的一个**映射**，是指一个规则，根据这个规则，对于 M 中的每个元素，都有 M' 中的一个确定的元素与它对应. 如果映射 σ 使 M 中元素 a 与 M' 中元素 a' 相对应，则称 a' 是 a 在映射 σ 下的**象**，而 a 称为 a' 的一个**原象**. 记作

$$\sigma(a) = a'.$$

与通常使用的表示法一样，用

$$a \in M$$

表示 a 是集合 M 中的一个元素，读作"a 属于 M". 用

$$a \overline{\in} M$$

表示 a 不是集合 M 中的元素，读作"a 不属于 M".

例 1 M 是全体整数的集合；M' 是全体偶数的集合. 定义

$$\sigma(n) = 2n \quad (n \in M).$$

这是 M 到 M' 的一个映射.

例 2 M 是全体整数的集合. 令

$$\sigma(n) = n + 1 \quad (n \in M).$$

σ 是 M 到自身的一个映射.

例 3 $M = \mathbf{P}^{n \times n}(n > 1)$. 定义

$$\sigma(A) = |A| \quad (A \in M).$$

这是 M 到 $\mathbf{P}$ 的一个映射.

例 4 设 V 是一个线性空间. $\mathbf{0}$ 是它的零向量. 令

$$\sigma(\alpha) = \mathbf{0} \quad (\alpha \in V),$$

则 σ 是 V 到自身的一个映射.

例 5 对于 $f(x) \in \mathbf{P}_n[x]$，定义

$$\sigma(f(x)) = f'(x).$$

这是 $\mathbf{P}_n[x]$ 到自身的一个映射.

例 6 设 M 是一个非空集合，定义

$$\sigma(a) = a \quad (a \in M).$$

即 σ 把 M 中每个元素都映到自身，σ 称为 M 的**恒等映射**或**单位映射**，记作 1_M. 在不致混淆的情况下，也可以简单地记作 1.

集合 M 到自身的映射也称作 M 的**变换**，1_M 也称 M 的**恒等变换**.

根据定义，如果 σ 是集合 M 到 M' 的一个映射，那么对于 M 中的任一个元素 a，$\sigma(a)$ 是 M' 中的一个**唯一确定**的元素，但是从上面的例子可以看出，给了 M' 中的一个元素 a'，M 中却不一定只有一个元素与它对应. 例如，在例 3 中，取定 $\mathbf{P}$ 中一个数，譬如说，

取定数 0，那么由于 M 中有许多矩阵的行列式都等于 0，因此，M 中有许多元素 A，都满足 $\sigma(A)=0$. 这个事实说明：原象不一定是唯一的. 从例 4 可以看到另一种情况：如果 V 不是零空间，那么对于 V 中非零向量 α_0，在 V 中找不到一个向量 α，使 $\sigma(\alpha)=\alpha_0$. 即 α_0 不是 V 中任何元素的象. 为了讨论这几种特殊情形，先介绍几个名词.

设 σ 是 M 到 M' 的一个映射. 用

$$\sigma(M)$$

表示 M 中元素在映射 σ 下的象的全体，称为 σ 的**象集合**. 显然，$\sigma(M)$ 是 M' 的一部分：$\sigma(M)\subseteq M'$. 如果 $\sigma(M)=M'$，就称映射 σ 是一个**满射**. 也就是说，如果 M' 中每个元素都是 M 中某个元素在映射 σ 下的象，那么 σ 就是一个满射. 例如，例 1,2,3,6 中的 σ 都是满射；例 5 中的 σ 不是满射；例 4 中的 σ 当 V 是零空间时是满射，否则不是满射.

如果在映射 σ 下，M 中不同元素的象也不同，即由 $a_1\neq a_2$ (a_1，$a_2\in M$) 可得 $\sigma(a_1)\neq\sigma(a_2)$，那么 σ 称为一个**单射**. 例如，例 1,2,6 中的映射是单射；例 3 与 5 中的映射不是单射. 同样地，当且仅当 V 是零空间时，例 4 中的 σ 才是单射.

如果映射 σ 既是满射又是单射，就称为一个**一一对应**或**一一映射**. 如例 1,2,6 中的映射都是一一对应. 如果集合 M 与 M' 中的元素个数都是有限的，那么存在 M 到 M' 的一一对应的充分必要条件是：它们包含的元素个数相等.

定义 17 设 V 与 V' 都是数域 $\mathbf{P}$ 上的线性空间，σ 是由 V 到 V' 的一个一一对应. 如果 σ 满足下列条件：

(1) $\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta)$ (α,β 是 V 中任意向量)；

(2) $\sigma(k\alpha)=k\sigma(\alpha)$ (α 是 V 中任意向量，k 是 $\mathbf{P}$ 中任意数)，

那么，σ 称为一个**同构映射**. 如果存在 V 到 V' 的一个同构映射，则称 V 与 V' 是**同构的**，记作 $V\cong V'$.

例 7 设 V 是数域 $\mathbf{P}$ 上的一个 n 维线性空间，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组基. 定义 V 到 $\mathbf{P}^n$ 的映射如下：如果 V 中向量 α 对于这组基的坐标是 $(a_1,a_2,\cdots,a_n)$，则令

$$\sigma(\alpha)=(a_1,a_2,\cdots,a_n).$$

即 σ 把 V 中向量

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

映到 $\mathbf{P}^n$ 中向量 $(a_1,a_2,\cdots,a_n)$.

根据坐标的性质，易知这是一个一一对应，而且

$$\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta),$$

$$\sigma(k\alpha)=k\sigma(\alpha).$$

所以 σ 是一个同构映射，V 与 $\mathbf{P}^n$ 是同构的.

例 8 关于 $\mathbf{P}$ 上线性空间 $\mathbf{P}^{n\times n}$，定义

$$\sigma(A)=A^{\mathrm{T}}$$

（A 是 $\mathbf{P}^{n\times n}$ 中任一个矩阵，A^{T} 是 A 的转置矩阵），则 σ 是 $\mathbf{P}^{n\times n}$ 到自身的一个同构映射.

定义 18 线性空间 V 到自身的同构映射称为 V 的**自同构映射**.

同构关系是线性空间之间的一个等价关系，即

命题 1 设 V,V',V'' 都是数域 $\mathbf{P}$ 上的线性空间，则

(1) $V\cong V$；

(2) 如果 $V\cong V'$，则 $V'\cong V$；

(3) 如果 $V\cong V'$，$V'\cong V''$，则 $V\cong V''$.

证明：(1) 易见 V 到自身的恒等映射是一个自同构映射. 故 $V\cong V$.

(2) 如果 $V\cong V'$，则有 V 到 V' 的同构映射 σ. 定义 V' 到 V 的映射 σ' 如下：设 α' 是 V' 中任一个向量，因为 σ 是一个一一对应，所以 V 中有唯一的一个向量 α，使得 $\sigma(\alpha)=\alpha'$. 令

$$\sigma'(\alpha')=\alpha.$$

σ'是一个 V'到 V 的映射.(这个映射称为 σ 的逆映射,可记作 σ^{-1}.)

首先证明 σ'是一个一一对应.对于 V 中任一向量 α,都有

$$\sigma'(\sigma(\alpha)) = \alpha, \quad \sigma(\alpha) \in V'.$$

所以 σ'是满射.对于 V'中两个不同的向量 α_1',α_2',用 α_1,α_2 分别表示它们在 σ 下的原象.由于 σ 是一一对应,所以

$$\alpha_1 \neq \alpha_2.$$

即

$$\sigma'(\alpha_1') \neq \sigma'(\alpha_2').$$

因此 σ'是单射,从而 σ'是一个一一对应.

下面来证 σ'是 V'到 V 的一个同构映射.设 α',β'是 V'中任意两个向量,k 是数域 $\mathbf{P}$ 中任意一个数.再设 V 中向量 α,β 分别是 α',β'对于 σ 的原像,即 $\sigma(\alpha)=\alpha',\sigma(\beta)=\beta'$.由于 σ 是一个同构映射,因此有

$$\sigma(\alpha+\beta) = \sigma(\alpha) + \sigma(\beta) = \alpha' + \beta';$$

$$\sigma(k\alpha) = k\sigma(\alpha) = k\alpha'.$$

于是

$$\sigma'(\alpha' + \beta') = \alpha + \beta = \sigma'(\alpha') + \sigma'(\beta');$$

$$\sigma'(k\alpha') = k\alpha = k\sigma'(\alpha'),$$

所以 σ'是一个同构映射,$V'\cong V$.

(3) 由 $V\cong V',V'\cong V''$ 可知,有由 V 到 V'的同构映射 σ 及 V'到 V'' 的同构映射 τ.对于 V 中向量 α,如果

$$\sigma(\alpha) = \alpha', \quad \tau(\alpha') = \alpha''.$$

则令

$$\rho(d) = d''.$$

这样定义的 ρ 是 V 到 V'' 的一个映射.(这个映射 ρ 称为映射 σ 与 τ 的乘积,通常记作 $\tau\sigma$.)显然 ρ 是 V 到 V'' 的一个一一对应.并且,对于 V 中向量 α,β,如果

$$\sigma(\alpha)=\alpha',\quad \tau(\alpha')=\alpha'',$$
$$\sigma(\beta)=\beta',\quad \tau(\beta')=\beta'',$$

则

$$\sigma(\alpha+\beta)=\alpha'+\beta',\quad \tau(\alpha'+\beta')=\alpha''+\beta''.$$

故
$$\rho(\alpha+\beta)=\alpha''+\beta''=\rho(\alpha)+\rho(\beta).$$

而且对于数域 $\mathbf{P}$ 中任意数 k，有

$$\sigma(k\alpha)=k\alpha',\quad \tau(k\alpha')=k\tau(\alpha')=k\alpha''.$$

因此

$$\rho(k\alpha)=k\alpha''.$$

所以 ρ 是一个同构映射，$V\cong V''$.

从命题 1 的证明可知：同构映射的逆映射是同构映射；两个同构映射的乘积还是一个同构映射.

需要注意的是：必须是同一个数域 $\mathbf{P}$ 上的线性空间，才能考虑同构问题.

由同构映射的定义立即可以证明，同构映射具有以下一些重要性质.

命题 2 设 σ 是数域 $\mathbf{P}$ 上线性空间 V 到 V' 的一个同构映射，则

(1) $\sigma(\mathbf{0})=\mathbf{0}$（式中左边的 $\mathbf{0}$ 是 V 中零向量，右边的 $\mathbf{0}$ 是 V' 中的零向量）；

(2) $\sigma(-\alpha)=-\sigma(\alpha)\quad(\alpha\in V)$；

(3) $\sigma(k_1\alpha_1+\cdots+k_s\alpha_s)=k_1\sigma(\alpha_1)+\cdots+k_s\sigma(\alpha_s)$
$$(\alpha_1,\cdots,\alpha_s\in V;\ k_1,\cdots,k_s\in\mathbf{P}).$$

证明留给读者作为练习.

命题 3 设 σ 是数域 $\mathbf{P}$ 上线性空间 V 到 V' 的同构映射，α_1，α_2，…，α_s 是 V 中向量，则 α_1，α_2，…，α_s 线性相关的充分必要条件是：它们的象 $\sigma(\alpha_1)$，$\sigma(\alpha_2)$，…，$\sigma(\alpha_s)$ 线性相关.

证明：如果 α_1，α_2，…，α_s 线性相关，则有 $\mathbf{P}$ 中不全为零的数

$k_1, k_2, \cdots, k_s$，使得

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

于是，

$$k_1\sigma(\alpha_1) + k_2\sigma(\alpha_2) + \cdots + k_s\sigma(\alpha_s) = 0.$$

所以 $\sigma(\alpha_1)$，$\sigma(\alpha_2)$，$\cdots$，$\sigma(\alpha_s)$也线性相关.

同构映射 σ 的逆映射 σ^{-1}是 V'到 V 的同构映射，它把$\sigma(\alpha_1)$，$\sigma(\alpha_2)$，$\cdots$，$\sigma(\alpha_s)$映到 α_1，α_2，$\cdots$，α_s. 故由 $\sigma(\alpha_1)$，$\sigma(\alpha_2)$，$\cdots$，$\sigma(\alpha_s)$线性相关可推出 α_1，α_2，$\cdots$，α_s 线性相关. |

命题 4 设 σ 是数域 $\mathbf{P}$ 上线性空间 V 到 V'的一个同构映射，W 是 V 的一个非空子集，则当且仅当 W 是 V 的子空间时，$\sigma(W)$ 是 V'的子空间.

证明：设 W 是 V 的一个子空间. 任取 $\alpha',\beta'\in\sigma(W)$，$k\in\mathbf{P}$，则有 W 中的向量 α,β，使

$$\sigma(\alpha) = \alpha', \quad \sigma(\beta) = \beta'.$$

于是

$$\sigma(\alpha+\beta) = \alpha' + \beta', \quad \sigma(k\alpha) = k\sigma(\alpha) = k\alpha'.$$

因为 W 是 V 的子空间，故 $\alpha+\beta$，$k\alpha$ 都在 W 中，因此

$$\alpha' + \beta' \in \sigma(W),\ k\alpha' \in \sigma(W),$$

所以 $\sigma(W)$是 V'的子空间.

同样可证，当 $\sigma(W)$是 V'的子空间时，W 是 V 的子空间.

由命题 3，V 中一个向量组 α_1，α_2，$\cdots$，α_s 线性无关的充分必要条件是：它们的象元素 $\sigma(\alpha_1)$，$\sigma(\alpha_2)$，$\cdots$，$\sigma(\alpha_s)$线性无关. 因为线性空间的维数是其中线性无关向量的最多个数，所以由同构映射的性质可推知：同构的线性空间有相同的维数. 另一方面，例 7 说明数域 $\mathbf{P}$ 上任一个 n 维线性空间都与 $\mathbf{P}^n$ 同构. 因此，由同构关系的对称性及传递性可知，数域 $\mathbf{P}$ 上任两个维数相同的线性空间都是同构的. 这样，我们就得到下述重要定理：

定理 18 数域 $\mathbf{P}$ 上两个有限维数线性空间同构的充分必要

条件是它们的维数相等.

事实上,两个 n 维线性空间 V 与 V' 之间的同构映射可以这样得到:取 V 的一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$;$V'$ 的一组基 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$,令

$$\sigma(k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n) = k_1\varepsilon_1' + k_2\varepsilon_2' + \cdots + k_n\varepsilon_n'.$$

则 σ 就是 V 到 V' 的一个同构映射.

下面我们再举几个同构映射的例子.

例 9 设 V_1 是实数域,看成实线性空间;V_2 是本章 8.1 节例 5 中的线性空间. 定义 V_1 到 V_2 的映射 σ 为:

$$a \to 2^a, \quad a \in V_1,$$

则 σ 是 V_1 到 V_2 的一个一一对应. 而且,由于

$$\begin{aligned}\sigma(a+b) &= 2^{a+b} = 2^a \cdot 2^b = 2^a \oplus 2^b \\ &= \sigma(a) \oplus \sigma(b) \quad (a,b \in V_1);\end{aligned}$$

$$\begin{aligned}\sigma(ka) &= 2^{ka} = (2^a)^k = k \circ 2^a \\ &= k \circ \sigma(a) \quad (a \in V_1, k \in R).\end{aligned}$$

所以 σ 是一个同构映射.

σ 的逆映射 σ^{-1} 为:

$$b \to \log_2 b \quad (b \in V_2, \text{即 } b \text{ 是正实数}).$$

根据命题 1 的证明,σ^{-1} 是 V_2 到 V_1 的一个同构映射. 请读者根据定义验证 σ^{-1} 确是 V_2 到 V_1 的同构映射.

例 10 V_1 是复数域,看成实数域 $\mathbf{R}$ 上的线性空间. 令

$$\sigma(a+b\mathrm{i}) = (a,b) \quad (a,b \in \mathbf{R}),$$

则 σ 是 V_1 到 $\mathbf{R}^2$ 的一个一一对应,而且

$$\begin{aligned}\sigma[(a+b\mathrm{i}) + (c+d\mathrm{i})] &= \sigma(a+c+(b+d)\mathrm{i}) \\ &= (a+c, b+d) = (a,b) + (c,d) \\ &= \sigma(a+b\mathrm{i}) + \sigma(c+d\mathrm{i}) \quad (a,b \in \mathbf{R});\end{aligned}$$

$$\begin{aligned}\sigma[k(a+b\mathrm{i})] &= \sigma(ka+kb\mathrm{i}) \\ &= (ka,kb) = k(a,b) \\ &= k\sigma(a+b\mathrm{i}) \quad (a,b \in \mathbf{R}, k \in \mathbf{R}),\end{aligned}$$

因此 σ 是 V_1 到 $\mathbf{R}^2$ 的一个同构映射.

用 W_1 表示 V_1 的由全部实数组成的子空间,W_2 表示 V_1 中全部纯虚数及 0 组成的子空间,则

$$V_1 = W_1 \dotplus W_2;$$

$$\sigma(W_1) = \{(a,0) \mid a \in \mathbf{R}\} = U_1,$$

$$\sigma(W_2) = \{(0,b) \mid b \in \mathbf{R}\} = U_2,$$

U_1, U_2 都是 $\mathbf{R}^2$ 的子空间,且

$$\mathbf{R}^2 = U_1 \dotplus U_2.$$

如果把复数域看成复数域上的线性空间,则是一维的. 因此在讨论线性空间时,一定要注意用来作数量乘法的基础数域 $\mathbf{P}$. 同一个集合作为不同数域上的线性空间是不同的.

在无限维线性空间的情况下,有些与有限维线性空间截然不同的结果. 下面举一个例子.

例 11 仍用 $\mathbf{P}[x]$表示数域 $\mathbf{P}$ 上全部多项式对多项式的加法及 $\mathbf{P}$ 中之数与多项式的乘法所成的线性空间. 我们知道,它是一个无限维线性空间. 用 V 表示 $\mathbf{P}[x]$中全部常数项为零的多项式所成的子空间. V 是 $\mathbf{P}[x]$的一个非平凡子空间,也是一个无限维线性空间. 定义 $\mathbf{P}[x]$到 V 的一个映射 σ:

$$f(x) \to xf(x), \quad f(x) \in \mathbf{P}[x],$$

则容易验证 σ 是一个同构映射.

这个例子说明无限维线性空间可以与它的一个非平凡子空间同构. 在有限维线性空间的情况,由于非平凡子空间的维数一定小于原线性空间的维数,这种情况是不可能发生的.

在线性空间的讨论中,有些问题只涉及到线性空间在线性运算下的代数性质,而不考虑线性空间的元素是什么,也不考虑其中运算是怎样定义的,在这种情况下,同构的空间具有相同的性质,因此,定理 18 说明维数是有限维线性空间的一个本质特征.

特别地,每个数域 $\mathbf{P}$ 上的 n 维线性空间都与 n 元有序数组所

构成的线性空间 $\mathbf{P}^n$ 同构，而同构的空间具有相同的性质，因此，我们以前讨论的关于 n 维向量的一些结论，在一般线性空间中也都是成立的. 而且，由于 $\mathbf{P}^n$ 的构造比较具体，关于一般线性空间的问题，如果不牵涉到线性空间中元素的特点，也可以在 $\mathbf{P}^n$ 中进行讨论.

习　题　8.7

1. 设 V_1 是数域 $\mathbf{P}$ 上 n 阶对称矩阵全体所成的线性空间；V_2 是 $\mathbf{P}$ 上三角形矩阵全体所成的线性空间，给出由 V_1 到 V_2 的两个同构映射.
2. 设 V_1 与 V_2 是数域 $\mathbf{P}$ 上两个线性空间；σ 是 V_1 到 V_2 的一个同构映射；α_1, α_2, …, α_n 是 V_1 中一组向量，证明：α_1, α_2, …, α_n 是 V_1 的一组基的充分必要条件是：$\sigma(\alpha_1)$, $\sigma(\alpha_2)$, …, $\sigma(\alpha_n)$是 V_2 的一组基.
3. 设 V_1 与 V_2 是数域 $\mathbf{P}$ 上的两个线性空间；α_1, α_2, …, α_s 是 V_1 中一组线性无关的向量；β_1, β_2, …, β_s 是 V_2 中一组线性无关的向量，证明：存在 V_1 到 V_2 的同构映射 σ，使
$$\sigma(\alpha_i) = \beta_i, \qquad i = 1,2,\cdots,s.$$
4. 找 $\mathbf{P}^4$ 的一个自同构映射 σ 满足：
$$\sigma((1,2,1,1)) = (1,2,0,1)$$
$$\sigma((0,1,0,-1)) = (-1,-2,1,1)$$

内 容 提 要

1. 线性空间

(1) 线性空间的定义

(2) 线性空间的简单性质

2. 线性空间的维数、基与坐标

(1) 关于向量组线性相关、线性无关、等价、向量组的极大线性无关组及秩等概念与重要性质的复习与补充

(2) 线性空间的维数、基与坐标

1) 线性空间维数的定义

2) 线性空间维数的判别法

如果在线性空间 V 中有 n 个线性无关的向量,而 V 中任一向量都可由这 n 个向量线性表出,那么 V 的维数为 n

3) 线性空间的基与坐标的定义

4) 线性空间的基与坐标的计算

5) 线性空间 V 中任一组线性无关的向量都可扩充为 V 的一组基

(3) 基变换与坐标变换

1) 由一组基到另一组基的过渡矩阵

过渡矩阵是可逆矩阵.

2) 同一个向量对于两组基的坐标之间的关系

设由基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到基 $\eta_1, \eta_2, \cdots, \eta_n$ 的过渡矩阵为 A,α 对于这两组基的坐标分别为$(a_1, a_2, \cdots, a_n)$及$(b_1, b_2, \cdots, b_n)$,则

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = A \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad \text{即} \quad \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = A^{-1} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

3) 几组基之间过渡矩阵的关系

设由基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到基 $\eta_1, \eta_2, \cdots, \eta_n$ 的过渡矩阵为 A;由基 $\eta_1, \eta_2, \cdots, \eta_n$ 到 $\xi_1, \xi_2, \cdots, \xi_n$ 的过渡矩阵为 B,则由 $\eta_1, \eta_2, \cdots, \eta_n$ 到 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的过渡矩阵为 A^{-1};由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\xi_1, \xi_2, \cdots, \xi_n$ 的过渡矩阵为 AB.

3. 线性子空间

(1) 线性子空间的定义

(2) 线性空间的子集构成子空间的条件

(3) 线性子空间的生成

1) 由向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 生成的子空间 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$

2）$L(\alpha_1, \alpha_2, \cdots, \alpha_s)$的一些性质

① 如果向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 可由向量组 $\beta_1, \beta_2, \cdots, \beta_t$ 线性表出，则

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s) \subseteq L(\beta_1, \beta_2, \cdots, \beta_t).$$

② 两个向量组生成相同子空间的充分必要条件是：这两个向量组等价.

③ $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$的维数等于向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的秩

3）齐次线性方程组的解空间

(4) 线性子空间的基可以扩充成原空间的基

4. 子空间的交与和

(1) 子空间交与和的定义

(2) 关于子空间交与和的维数公式

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$$

(3) 子空间的直和

1）子空间的直和的定义

2）子空间的直和的判别法

以下一些条件都是 $W_1 + W_2$ 是直和的充分必要条件：

① 零向量表成

$$\mathbf{0} = \alpha_1 + \alpha_2, \quad \alpha_1 \in W_1, \alpha_2 \in W_2$$

的表法是唯一的，即由上式能推出

$$\alpha_1 = \alpha_2 = \mathbf{0}.$$

② $W_1 \cap W_2 = \{\mathbf{0}\}$

③ $\dim(W_1 + W_2) = \dim W_1 + \dim W_2$.

3）W_1 是线性空间 V 的一个子空间，则存在 V 的一个子空间 W_2，使

$$V = W_1 \dotplus W_2.$$

(4) 多个子空间的交与和、以及直和

(5) n 维线性空间可以表示成 n 个 1 维子空间的直和

5. 线性空间的同构

(1) 同构、同构映射以及自同构的概念

(2) 同构映射的一些性质

(3) 数域 $\mathbf{P}$ 上两个有限维线性空间同构的充分必要条件是：它们的维数相同

(4) 数域 $\mathbf{P}$ 上任一 n 维线性空间都与 $\mathbf{P}^n$ 同构

复习题 8

1. 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是数域 $\mathbf{P}$ 上线性空间 V 中一组线性相关的向量，并设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 中任意 $s-1$ 个向量都线性无关，试证：

(1) 存在数域 $\mathbf{P}$ 中 s 个全不为零的数 $k_1, k_2, \cdots, k_s$ 使

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = \mathbf{0};$$

(2) 如果 $\mathbf{P}$ 中两组数 $k_1, k_2, \cdots, k_s$ 及 $l_1, l_2, \cdots, l_s$ 满足

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = \mathbf{0};$$

及

$$l_1\alpha_1 + l_2\alpha_2 + \cdots + l_s\alpha_s = \mathbf{0};$$

则 $(k_1, k_2, \cdots, k_s)$ 与 $(l_1, l_2, \cdots, l_s)$ 成比例.

2. $\alpha_1, \alpha_2, \cdots, \alpha_s$; $\beta_1, \beta_2, \cdots, \beta_t$ 是线性空间 V 中两组向量. 已知向量组 $\beta_1, \beta_2, \cdots, \beta_t$ 可由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性表出：

$$\begin{aligned}\beta_1 &= a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1s}\alpha_s,\\ \beta_2 &= a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2s}\alpha_s,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \beta_t &= a_{t1}\alpha_1 + a_{t2}\alpha_2 + \cdots + a_{ts}\alpha_s.\end{aligned}$$

令

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ a_{t1} & a_{t2} & \cdots & a_{ts} \end{bmatrix},$$

试证：

(1) $r\{\beta_1, \beta_2, \cdots, \beta_s\} \leqslant r(A)$;

(2) 如果 $\alpha_1, \alpha_2, \cdots, \alpha_t$ 线性无关,则 $r\{\beta_1, \beta_2, \cdots, \beta_s\} = r(A)$.

3. 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 n 维线性空间 V 中 n 个向量,如果 V 中任一向量都可由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表出,则 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的一组基.

4. 设

$$
\begin{aligned}
\alpha_1 &= (1,0,1,2), & \beta_1 &= (1,-1,1,-1),\\
\alpha_2 &= (-1,1,1,0), & \beta_2 &= (-2,3,-1,1),\\
\alpha_3 &= (2,2,1,1), & \beta_3 &= (1,0,2,1),\\
\alpha_4 &= (3,1,2,3); & \beta_4 &= (0,2,-1,2).
\end{aligned}
$$

(1) 证明 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 及 $\beta_1, \beta_2, \beta_3, \beta_4$ 都是 $\mathbf{P}^4$ 的基;

(2) 求由基 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 到 $\beta_1, \beta_2, \beta_3, \beta_4$ 的过渡矩阵;

(3) 求向量 $\xi=(1,3,-2,1)$ 对于基 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的坐标.

5. 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$ 是线性空间 V 的两组基,用 W 表示 V 中对于这两组基有相同坐标的向量全体组成的集合,

(1) 证明:W 是 V 的一个子空间;

(2) 对上题中的两组基求 W.

6. 证明:数域 $\mathbf{P}$ 上线性空间 V 的一个非空子集 W 是 V 的子空间的充分必要条件是:如果 $\alpha, \beta \in W$,则对 $\mathbf{P}$ 中任意数 k,都有

$$\alpha + k\beta \in W.$$

7. 求下列齐次方程组的解空间,并求其维数和一组基:

$$
(1)\begin{cases}
x_1 - 3x_2 + 2x_3 + x_4 + x_5 = 0,\\
x_1 - 3x_2 + 2x_3 - 3x_4 - 3x_5 = 0,\\
x_1 - 3x_2 + 2x_3 + 5x_4 + 5x_5 = 0;
\end{cases}
$$

$$
(2)\begin{cases}
x_1 + x_2 - 2x_3 - x_4 + x_5 = 0,\\
3x_1 - x_2 + x_3 + 4x_4 + 3x_5 = 0,\\
x_1 + 5x_2 - 9x_3 - 8x_4 + x_5 = 0.
\end{cases}
$$

8. 求上题中两个解空间的交与和,并分别求它们的维数和一组基.

9. 给了 $\mathbf{P}^5$ 中两组向量:

$$
\begin{aligned}
\alpha_1 &= (1,2,1,0,1), & \beta_1 &= (2,-1,0,1,1),\\
\alpha_2 &= (-1,1,1,1,0); & \beta_2 &= (1,-1,3,7,2);
\end{aligned}
$$

令 $W_1=L(\alpha_1,\alpha_2), W_2=L(\beta_1,\beta_2)$. 求 $W_1+W_2, W_1\cap W_2$,并分别求它们的

维数和一组基.

10. A 是 $\mathbf{P}^{n\times n}$ 中一个矩阵,令

$$F(A)=\{f(A)\mid f(x)\in\mathbf{P}[x]\}.$$

(1) 证明:$F(A)$是 $\mathbf{P}^{n\times n}$ 的一个子空间;

(2) 证明:可找到非负整数 m,使

$$E,\ A,\ A^2,\ \cdots,\ A^m$$

是 $F(A)$的一组基;

(3) 证明:$F(A)$的维数等于 A 的最小多项式的次数.

11. 设

$$A=\begin{bmatrix}0&0&0\\0&0&1\\0&0&0\end{bmatrix},$$

求 $\mathbf{P}^{3\times 3}$ 的子空间 $F(A)$,并求 $F(A)$的维数和一组基.

12. A 是 $\mathbf{P}^{n\times n}$ 中一个矩阵,令

$$C(A)=\{B\in\mathbf{P}^{n\times n}\mid AB=BA\}.$$

证明:(1) $C(A)$是 $\mathbf{P}^{n\times n}$ 的一个子空间;

(2) $C(A)\supseteq F(A)$.

13. A 同第 11 题.求 $\mathbf{P}^{3\times 3}$ 的子空间 $C(A)$,并求 $C(A)$的维数和一组基.

14. 设 A 是 $\mathbf{P}^{n\times n}$ 中一个矩阵,令

$$N(A)=\{B\in\mathbf{P}^{n\times n}\mid AB=O\}.$$

(1) 证明:$N(A)$是 $\mathbf{P}^{n\times n}$ 的一个子空间;

(2) 已知 A 的秩等于 r,求 $N(A)$的维数.

15. A 同第 11 题,求 $\mathbf{P}^{3\times 3}$ 的子空间 $N(A)$,并求 $S(A)$的维数和一组基.

16. A 同第 11 题,

(1) 求 $F(A)\cap C(A)$,$F(A)+C(A)$,并分别求它们的维数和一组基;

(2) 求 $F(A)\cap N(A)$,$F(A)+N(A)$,并分别求它们的维数和一组基;

(3) 求 $N(A)\cap C(A)$,$N(A)+C(A)$,并分别求它们的维数和一组基.

17. 考虑实数域 $\mathbf{R}$ 上齐次方程组

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

用 V_1 表示这个方程组的解空间，令

$$\alpha_1 = (a_{11}, a_{12}, \cdots, a_{1n})$$

$$\alpha_2 = (a_{21}, a_{22}, \cdots, a_{2n})$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\alpha_s = (a_{s1}, a_{s2}, \cdots, a_{sn})$$

再令 V_2 为 $\mathbf{R}^n$ 的由 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 生成的子空间：

$$V_2 = L(\alpha_1, \alpha_2, \cdots, \alpha_s).$$

试证：

$$\mathbf{R}^n = V_1 \dotplus V_2.$$

18. 设 V 是习题 8.1 第 1 题(5)中的线性空间，求 V 到 $\mathbf{R}^2$ 的一个同构映射.

第 9 章　线性变换

设 V 是数域 $\mathbf{P}$ 上的一个线性空间，V 到自身的映射称为 V 的变换. 线性变换是线性空间的一种最基本的变换，是线性代数研究的中心问题之一.

本章主要讨论线性变换的矩阵问题，包括线性变换对于一组基的矩阵表法，同一个线性变换对于不同基的矩阵之间的关系，如何选择线性空间的基使得给定的线性变换在这组基下的矩阵最简单等问题. 这些结论给出了矩阵相似关系的几何背景，提供了解决矩阵相似问题的一种方法. 这一章中介绍的线性变换的特征值及特征向量等概念不仅在讨论线性变换的矩阵时要用到，在处理其他数学问题及应用问题时，也是很有用的数学工具.

9.1　线性变换的定义与简单性质

下面如果不特别声明，总假设 V 是数域 $\mathbf{P}$ 上的一个线性空间.

定义 1　线性空间 V 的一个变换 $\boldsymbol{A}$ 称为 V 的一个**线性变换**，如果对于 V 中任意向量 α,β 及 $\mathbf{P}$ 中任意数 k，都有

(1)　$$\boldsymbol{A}(\alpha+\beta)=\boldsymbol{A}(\alpha)+\boldsymbol{A}(\beta);$$

(2)　$$\boldsymbol{A}(k\alpha)=k\boldsymbol{A}(\alpha).$$

以后用粗体大写拉丁字母 $\boldsymbol{A},\boldsymbol{B},\cdots$表示 V 的变换；用 $\boldsymbol{A}(\alpha)$或简单地用 $\boldsymbol{A}\alpha$ 表示向量 α 在变换 $\boldsymbol{A}$ 下的象.

定义 1 中条件(1)说明线性变换保持向量的加法；条件(2)说明线性变换保持数量乘法. 因此，线性空间的线性变换就是保持线性运算的变换.

例 1　线性空间 V 的**恒等变换**显然是线性变换. 恒等变换也称**单位变换**,记作 $\boldsymbol{E}$,即

$$\boldsymbol{E}(\alpha)=\alpha \quad (\alpha\in V).$$

线性空间的**零变换** 0,即

$$O(\alpha)=O \quad (\alpha\in V).$$

也是线性变换。

例 2　设 V 是数域 $\mathbf{P}$ 上的一个线性空间,k 是数域 $\mathbf{P}$ 中一个取定的数,定义 V 的变换如下:

$$\alpha\to k\alpha \quad (\alpha\in V).$$

不难验证,这是 V 的一个线性变换,称为由数 k 决定的**数乘变换**,记作 $k\boldsymbol{E}$. 当 $k=1$ 时,即得恒等变换;当 $k=0$ 时,即得零变换. 因此,恒等变换和零变换都是数乘变换.

例 3　平面上的定位向量构成实数域上的一个二维线性空间. 把平面围绕坐标原点按反时针方向旋转 θ 角,就是一个线性变换,记作 $\boldsymbol{T}_\theta$. 如果平面上的一个向量 α 在直角坐标系下的坐标是 (x,y),那么 α 在旋转后的象 $\boldsymbol{T}_\theta(\alpha)$ 的坐标 (x',y') 可以用公式

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

来计算.

同样地,几何空间(实数域上的三维线性空间)绕轴的旋转也是一个线性变换.

例 4　在线性空间 $\mathbf{P}[x]$ 或 $\mathbf{P}_n[x]$ 中,用 $\boldsymbol{D}$ 表示求导数的变换:

$$\boldsymbol{D}f(x)=f'(x).$$

由于

$$\boldsymbol{D}[f(x)+g(x)]=\boldsymbol{D}f(x)+\boldsymbol{D}g(x);$$
$$\boldsymbol{D}[kf(x)]=k\boldsymbol{D}f(x).$$

因此这是一个线性变换.

例 5　在 $\mathbf{P}^{n\times n}$ 中取定一个矩阵 A. 定义 $\mathbf{P}^{n\times n}$ 的变换 $\boldsymbol{A}$ 为

$$\mathbf{A}(X)=AX,\quad X\in \mathbf{P}^{n\times n}.$$

因为对 $\mathbf{P}^{n\times n}$ 中元素 X,Y，以及 $\mathbf{P}$ 中数 k，有

$$\begin{aligned}\mathbf{A}(X+Y)&=A(X+Y)=AX+AY\\&=\mathbf{A}(X)+\mathbf{A}(Y);\\ \mathbf{A}(kX)&=A(kX)=k(AX)=k\mathbf{A}(X).\end{aligned}$$

所以 $\mathbf{A}$ 是 $\mathbf{P}^{n\times n}$ 的一个线性变换.

例 6 在 $\mathbf{P}^{n\times n}$ 中取定一个矩阵 A. 定义 $\mathbf{P}^{n\times n}$ 的变换 $\mathbf{B}$ 为

$$\mathbf{B}(X)=X+A.$$

由于

$$\begin{gathered}\mathbf{B}(X+Y)=(X+Y)+A,\\ \mathbf{B}X+\mathbf{B}Y=(X+A)+(Y+A)=X+Y+2A;\\ \mathbf{B}(kX)=kX+A,\\ k(\mathbf{B}X)=k(X+A)=kX+kA.\end{gathered}$$

因此，当 $A\neq O$ 时，$\mathbf{B}$ 不是线性变换；当 $A=O$ 时，$\mathbf{B}$ 是一个线性变换.

从定义可以推出线性变换的一些简单而重要的性质：

命题 1 设 $\mathbf{A}$ 是线性空间 V 的一个线性变换，则

$$\mathbf{A}(\mathbf{0})=\mathbf{0};\mathbf{A}(-\alpha)=-\mathbf{A}(\alpha),\alpha\in V.$$

证明：

$$\begin{gathered}\mathbf{A}(\mathbf{0})=\mathbf{A}(0\cdot\alpha)=0\cdot\mathbf{A}(\alpha)=\mathbf{0}.\\ \mathbf{A}(-\alpha)=\mathbf{A}[(-1)\alpha]=(-1)\mathbf{A}(\alpha)=-\mathbf{A}(\alpha).\end{gathered}$$

命题 2 设 $\mathbf{A}$ 是线性空间 V 的一个线性变换. $\alpha_1,\cdots,\alpha_s$，及 β 都是 V 中向量. 如果 β 可由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出：

$$\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s,$$

则 $\mathbf{A}(\beta)$ 也可由 $\mathbf{A}(\alpha_1),\mathbf{A}(\alpha_2),\cdots,\mathbf{A}(\alpha_s)$ 线性表出：

$$\mathbf{A}(\beta)=k_1\mathbf{A}(\alpha_1)+k_2\mathbf{A}(\alpha_2)+\cdots+k_s\mathbf{A}(\alpha_s).$$

证明 由假设

$$\mathbf{A}(\beta)=\mathbf{A}(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)$$

$$= k_1\boldsymbol{A}(\alpha_1) + k_2\boldsymbol{A}(\alpha_2) + \cdots + k_s\boldsymbol{A}(\alpha_s).$$

命题 2 说明线性变换保持线性组合与线性关系不变. 由此并可推出:

命题 3 线性相关的向量组经过线性变换后,所得的象向量组也是线性相关的.

证明: 设 V 是数域 $\mathbf{P}$ 上的一个线性空间,$\boldsymbol{A}$ 是 V 的一个线性变换. $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 V 中一组向量. 如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是线性相关的,则有数域 $\mathbf{P}$ 中不全为零的数 $k_1,k_2,\cdots,k_s$,使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = \mathbf{0}.$$

于是,由命题 2,得

$$k_1\boldsymbol{A}(\alpha_1) + k_2\boldsymbol{A}(\alpha_2) + \cdots + k_s\boldsymbol{A}(\alpha_s) = \boldsymbol{A}(\mathbf{0}) = \mathbf{0}.$$

因此,$\boldsymbol{A}(\alpha_1),\boldsymbol{A}(\alpha_2),\cdots,\boldsymbol{A}(\alpha_s)$也是线性相关的.

命题 3 的逆命题是不成立的. 即:线性无关的向量组经过线性变换后,可能会变成线性相关的. 例如:任何向量经零变换后都变成零向量,因此零变换把任一个线性无关的向量组都变成线性相关的向量组. 又如下例说明,有的线性变换把一些线性无关的向量组变成线性相关的向量组,而把某些线性无关的向量组仍变为线性无关的向量组.

例 7 考虑 $\mathbf{P}^n$ 的变换 $\boldsymbol{A}$:

$$A(a_1,a_2,\cdots,a_n) = (a_2,a_3,\cdots,a_n,0).$$

因为

$$\begin{aligned}
&A[(a_1,a_2,\cdots,a_n) + (b_1,b_2,\cdots,b_n)] \\
&\quad = A(a_1, + b_1,a_2 + b_2,\cdots,a_n + b_n) \\
&\quad = (a_2 + b_2,a_3 + b_3,\cdots,a_n + b_n,0) \\
&\quad = (a_2,a_3,\cdots,a_n,0) + (b_2,b_3,\cdots,b_n,0) \\
&\quad = A(a_1,a_2,\cdots,a_n) + A(b_1,b_2,\cdots,b_n);
\end{aligned}$$

$$\begin{aligned}
A[k(a_1,a_2,\cdots,a_n)] &= A(ka_1,ka_2,\cdots,ka_n) \\
&= (ka_2,ka_3,\cdots,ka_n,0)
\end{aligned}$$

$$= k(a_2, a_3, \cdots, a_n, 0)$$
$$= kA(a_1, a_2, \cdots, a_n).$$

因此,A 是一个线性变换. 设 $n \geqslant 3$. 令

$$\alpha_1 = (1,1,\cdots,1), \qquad \alpha_2 = (2,1,\cdots,1);$$
$$\beta_1 = (1,1,\cdots,1), \qquad \beta_2 = (1,1,\cdots,2);$$

则

$$A\alpha_1 = (1,\cdots,1,0), \qquad A\alpha_2 = (1,\cdots,1,0);$$
$$A\beta_1 = (1,\cdots,1,0), \qquad A\beta_2 = (1,\cdots,1,2,0).$$

容易看出,向量组 α_1, α_2 线性无关;而向量组 $A\alpha_1, A\alpha_2$ 线性相关. 向量组 β_1, β_2 线性无关,而向量组 $A\beta_1, A\beta_2$ 仍线性无关.

命题 2 说明,如果 V 中一组向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 在线性变换 A 下的象已经知道,那么这组向量的任一个线性组合在 A 下的象也就知道了. 特别地,如果 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 V 的一组基,那么由于 V 中任一向量都可表成 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的线性组合,因此,如果知道了 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 在线性变换 A 下的象,那么 V 中任一个向量在 A 下的象也就知道了. 这就是说,一个线性变换的作用完全由它在一组基上的作用所确定. 我们还可证明,基向量的象可以是任意指定的. 即有下述定理.

定理 1 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性空间 V 的一组基,$\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 中任意取定的 n 个向量,那么有唯一的一个线性变换 A,使得

$$A\varepsilon_i = \alpha_i, \qquad i = 1, 2, \cdots, n.$$

证明: 首先证明存在性. 为此作一个满足要求的线性变换. 设 ξ 是 V 中任一向量,那么 ξ 可以唯一地表成 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的线性组合:

$$\xi = k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n.$$

定义 V 的变换 A 为

$$A\xi = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n.$$

现证 A 是 V 的一个线性变换.

对于 V 中任意两个向量 α,β,设

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n,$$
$$\beta = b_1\varepsilon_1 + b_2\varepsilon_2 + \cdots + b_n\varepsilon_n,$$

则

$$\alpha + \beta = (a_1 + b_1)\varepsilon_1 + (a_2 + b_2)\varepsilon_2 + \cdots + (a_n + b_n)\varepsilon_n.$$

于是

$$\begin{aligned} A(\alpha + \beta) &= (a_1 + b_1)\varepsilon_1 + (a_2 + b_2)\varepsilon_2 + \cdots + (a_n + b_n)\varepsilon_n \\ &= (a_1\varepsilon_1 + \cdots + a_n\varepsilon_n) + (b_1\varepsilon_1 + \cdots + b_n\varepsilon_n) \\ &= A\alpha + A\beta. \end{aligned}$$

如果 k 是 $\mathbf{P}$ 中一个数,那么

$$k\alpha = ka_1\varepsilon_1 + ka_2\varepsilon_2 + \cdots + ka_n\varepsilon_n$$

于是

$$\begin{aligned} A(k\alpha) &= ka_1\varepsilon_1 + ka_2\varepsilon_2 + \cdots + ka_n\varepsilon_n \\ &= k(a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n) \\ &= kA\alpha. \end{aligned}$$

所以 A 是 V 的一个线性变换.而且,因为

$$\varepsilon_i = 0\varepsilon_1 + \cdots + 0\varepsilon_{i-1} + \varepsilon_i + 0\varepsilon_{i+1} + \cdots + 0\varepsilon_n \quad (i = 1,2,\cdots,n),$$

所以

$$\begin{aligned} A\varepsilon_i &= 0\varepsilon_1 + \cdots + 0\varepsilon_{i-1} + \varepsilon_i + 0\varepsilon_{i+1} + \cdots + 0\varepsilon_n \\ &= \varepsilon_i \quad (i = 1,2,\cdots,n). \end{aligned}$$

因此 A 即为所求的线性变换.

下面证明唯一性.设有 V 的两个线性变换 A,B 都满足条件,即

$$A\varepsilon_i = \alpha_i, \qquad i = 1,2,\cdots,n;$$
$$B\varepsilon_i = \alpha_i, \qquad i = 1,2,\cdots,n;$$

那么对于 V 中任一向量 α:

$$\alpha = k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n,$$

有

$$\begin{aligned}\boldsymbol{A}\alpha &= k_1\boldsymbol{A}\varepsilon_1 + k_2\boldsymbol{A}\varepsilon_2 + \cdots + k_n\boldsymbol{A}\varepsilon_n \\ &= k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n; \\ \boldsymbol{B}\alpha &= k_1\boldsymbol{B}\varepsilon_1 + k_2\boldsymbol{B}\varepsilon_2 + \cdots + k_n\boldsymbol{B}\varepsilon_n \\ &= k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n.\end{aligned}$$

因此

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

由于 α 是 V 中任意向量，故必有 $\boldsymbol{A}=\boldsymbol{B}$. ▏

由于定理 1 中的唯一性的部分经常用到，因此我们再把它独立地写出来.

推论 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是线性空间 V 的一组基. 如果 V 的两个线性变换 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 在这组基上的作用相同，即

$$\boldsymbol{A}\varepsilon_i = \boldsymbol{B}\varepsilon_i, \qquad i = 1,2,\cdots,n,$$

那么必有

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

即线性变换由它在一组基上的作用唯一确定.

例 8 求 $\mathbf{P}^3$ 的一个线性变换 $\boldsymbol{A}$，满足

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

解：令

$$\varepsilon_1 = (1,1,1), \quad \varepsilon_2 = (1,1,0), \quad \varepsilon_3 = (1,0,0).$$

因为

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} \neq 0,$$

所以 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 线性无关，是 $\mathbf{P}^3$ 的一组基. $\mathbf{P}^3$ 中任一向量

$$\alpha = (a_1,a_2,a_3)$$

可表成 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 的线性组合：

$$\alpha = a_3\varepsilon_1 + (a_2 - a_3)\varepsilon_2 + (a_1 - a_2)\varepsilon_3.$$

定义 V 的变换 A 为

$$\begin{aligned} A(\alpha) &= a_3(1,2,3) + (a_2 - a_3)(-1,1,1) \\ &\quad + (a_1 - a_2)(1,0,-2) \\ &= (a_1 - 2a_2 + 2a_3, a_2 + a_3, -2a_1 + 3a_2 + 2a_3). \end{aligned}$$

则 A 是 V 的一个线性变换，并且满足所述条件.

下面的命题在以后经常要用到.

命题 4 设 $\xi_1,\xi_2,\cdots,\xi_s$ 是 n 维线性空间 V 中一组线性无关的向量；$\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 V 中任意 s 个取定的向量，则存在 V 的线性变换 A，使得

$$A\xi_i = \alpha_i, \qquad i = 1,2,\cdots,s.$$

证明：因为 $\xi_1,\xi_2,\cdots,\xi_s$ 线性无关，故可以扩充为 V 的一组基：

$$\xi_1,\xi_2,\cdots,\xi_s,\xi_{s+1},\cdots,\xi_n.$$

再在 V 中任取 $n-s$ 个向量 $\alpha_{s+1},\cdots,\alpha_n$. 则由定理 1，知有 V 的线性变换 A，使得

$$A\xi_i = \alpha_i, \qquad i = 1,2,\cdots,n.$$

这个线性变换 A 当然满足所需条件.

从命题 4 的证明可看出，当 $s<n$ 时，满足要求的线性变换不止一个.

习 题 9.1

1. 判断下面的一些变换，是否是线性变换：

(1) 在线性空间 V 中，定义 $A\alpha=\alpha_0 \quad (\alpha\in V)$，其中 α_0 是 V 中一个固定的向量；

(2) 在线性空间 V 中，定义 $A\alpha=\alpha+\alpha_0(\alpha\in V)$，其中 α_0 是 V 中一个固定的向量；

(3) 在 $\mathbf{P}^3$ 中定义 $\boldsymbol{A}(a_1,a_2,a_3)=(a_1^2,a_2+a_3,a_3)$

(4) 在 $\mathbf{P}^4$ 中定义

$$\boldsymbol{A}(a_1,a_2,a_3,a_4) = (a_1 + a_2, a_1 + a_3, a_1 + a_4, a_2 - a_4)$$

2. 证明：线性变换把等价的向量组变为等价的向量组.

3. 设 $\boldsymbol{A}$ 是线性空间 V 的一个线性变换，证明下列两个条件是等价的：

(1) $\boldsymbol{A}$ 把 V 中某一组线性无关的向量变成一组线性相关的向量；

(2) $\boldsymbol{A}$ 把 V 中某个非零向量变成零向量.

4. 求 $\mathbf{P}^3$ 的一个线性变换 $\boldsymbol{A}$ 满足

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

5. 找出 $\mathbf{P}^3$ 的把(1,0,0)变成(1,2,3)，把(0,1,0)变成(1,−1,0)的两个线性变换.

9.2 线性变换的运算

这一节介绍线性变换的运算及其简单性质

1. 线性变换的加法

首先来定义线性变换的加法

定义 2 设 $\boldsymbol{A},\boldsymbol{B}$ 是数域 $\mathbf{P}$ 上线性空间 V 的两个线性变换. 定义它们的和 $\boldsymbol{A}+\boldsymbol{B}$ 为

$$(\boldsymbol{A} + \boldsymbol{B})(\alpha) = \boldsymbol{A}(\alpha) + \boldsymbol{B}(\alpha), \qquad \alpha \in V.$$

线性变换 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和 $\boldsymbol{A}+\boldsymbol{B}$ 也是线性变换. 这是因为，对于 V 中向量 α,β 及 $\mathbf{P}$ 中数 k，有

$$\begin{aligned}(\boldsymbol{A} + \boldsymbol{B})(\alpha + \beta) &= \boldsymbol{A}(\alpha + \beta) + \boldsymbol{B}(\alpha + \beta)\\ &= [\boldsymbol{A}(\alpha) + \boldsymbol{A}(\beta)] + [\boldsymbol{B}(\alpha) + \boldsymbol{B}(\beta)]\\ &= [\boldsymbol{A}(\alpha) + \boldsymbol{B}(\alpha)] + [\boldsymbol{A}(\beta) + \boldsymbol{B}(\beta)]\\ &= (\boldsymbol{A} + \boldsymbol{B})(\alpha) + (\boldsymbol{A} + \boldsymbol{B})(\beta);\end{aligned}$$

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

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

对于加法，零变换有着特殊的地位，它与任一线性变换 $\boldsymbol{A}$ 的和仍等于 $\boldsymbol{A}$：

$$\boldsymbol{A} + \boldsymbol{O} = \boldsymbol{A}.$$

对于线性变换 $\boldsymbol{A}$，可以定义它的**负变换** $-\boldsymbol{A}$：

$$(-\boldsymbol{A})(\alpha) = -\boldsymbol{A}(\alpha), \quad \alpha \in V.$$

容易验证，线性变换的负变换也是线性变换. 并且

$$\boldsymbol{A} + (-\boldsymbol{A}) = \boldsymbol{O}.$$

因此可定义线性变换的减法为

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

例 1 在 $\mathbf{P}^{n\times n}$ 中取定矩阵 A,B. 定义 V 的变换 $\boldsymbol{A},\boldsymbol{B}$ 如下：

$$\boldsymbol{A}(X) = AX, \quad X \in V;$$
$$\boldsymbol{B}(X) = BX, \quad X \in V.$$

已知 $\boldsymbol{A},\boldsymbol{B}$ 都是线性变换. 它们的和 $\boldsymbol{A}+\boldsymbol{B}$ 为

$$(\boldsymbol{A} + \boldsymbol{B})X = \boldsymbol{A}X + \boldsymbol{B}X = AX + BX$$
$$= (A + B)X, \quad X \in V.$$

$\boldsymbol{A}$ 的负变换 $-\boldsymbol{A}$ 为

$$(-\boldsymbol{A})(X) = -(\boldsymbol{A}X) = -AX, \quad X \in V.$$

命题 1 线性变换的加法适合结合律及交换律. 即，对于线性空间 V 的任意线性变换 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$，有

$$\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C};$$
$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}.$$

2. 线性变换的数量乘法

定义 3 设 V 是数域 $\mathbf{P}$ 上的一个线性空间，$\boldsymbol{A}$ 是 V 的一个线性变换，k 是数域 $\mathbf{P}$ 中的一个数. 定义 k 与 $\boldsymbol{A}$ 的数量乘积 $k\boldsymbol{A}$ 为

$$(k\boldsymbol{A})\alpha = k\boldsymbol{A}(\alpha), \quad \alpha \in V.$$

当然，kA 也是线性变换. 应用数乘变换，可以把数量乘积 kA 表成

$$kA = (kE)A.$$

关于数量乘法，有以下结论：

命题 2 线性变换的数量乘法适合以下一些规律：

$$(kl)A = k(lA);$$

$$(k + l)A = kA + lA;$$

$$k(A + B) = kA + kB;$$

$$1A = A.$$

根据以上的讨论，可知：

命题 3 设 V 是数域 $\mathbf{P}$ 上的一个线性空间，则 V 的全体线性变换对于加法与数量乘法也构成数域 $\mathbf{P}$ 上的一个线性空间，记作 $L(V)$.

3. 线性变换的乘法

下面讨论线性变换的乘法.

定义 4 设 A, B 是数域 $\mathbf{P}$ 上线性空间 V 的两个线性变换. 定义它们的乘积 AB 为

$$AB(\alpha) = A(B(\alpha)), \quad \alpha \in V.$$

AB 当然是 V 的一个变换. 下面证明 AB 是 V 的线性变换.

对于 V 中任意两个向量 α, β，由于 A, B 都是线性变换，故有

$$\begin{aligned} AB(\alpha + \beta) &= A(B(\alpha + \beta)) = A(B(\alpha) + B(\beta)) \\ &= A(B(\alpha)) + A(B(\beta)) \\ &= AB(\alpha) + AB(\beta). \end{aligned}$$

对于 V 中向量 α 及 $\mathbf{P}$ 中数 k，有

$$\begin{aligned} AB(k\alpha) &= A(B(k\alpha)) = A(kB(\alpha)) \\ &= kA(B(\alpha)) = kAB(\alpha). \end{aligned}$$

因此，AB 是 V 的一个线性变换.

我们知道，单位变换 E 是 V 的线性变换. 对于任一变换 A，都

有 $\boldsymbol{EA}=\boldsymbol{AE}=\boldsymbol{A}$. 对于零变换 $\boldsymbol{O}$,则有

$$\boldsymbol{OA}=\boldsymbol{AO}=\boldsymbol{O}.$$

例 2 已知 $\boldsymbol{D}$ 是 $\mathbf{P}[x]$的一个线性变换. 再定义变换 $\boldsymbol{A}$ 为

$$\boldsymbol{A}f(x)=xf(x).$$

因为

$$\begin{aligned}\boldsymbol{A}(f(x)+g(x))&=x(f(x)+g(x))=xf(x)+xg(x)\\&=\boldsymbol{A}f(x)+\boldsymbol{A}g(x),\quad f(x),g(x)\in\mathbf{P}[x];\\\boldsymbol{A}(kf(x))&=x[kf(x)]=kxf(x)\\&=k\boldsymbol{A}f(x),\quad k\in\mathbf{P},\quad f(x)\in\mathbf{P}[x].\end{aligned}$$

所以 $\boldsymbol{A}$ 是 $\mathbf{P}[x]$的一个线性变换.

根据定义

$$\boldsymbol{DA}f(x)=\boldsymbol{D}(\boldsymbol{A}f(x))=\boldsymbol{D}(xf(x))=f(x)+xf'(x);$$

$$\boldsymbol{AD}f(x)=\boldsymbol{A}(\boldsymbol{D}f(x))=\boldsymbol{A}f'(x)=xf'(x).$$

所以

$$\boldsymbol{AD}\neq\boldsymbol{DA}.$$

例 3 在 $\mathbf{P}^{2\times2}$中,定义线性变换 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 为

$$\boldsymbol{A}(X)=\begin{bmatrix}1&0\\0&0\end{bmatrix}X,\quad \boldsymbol{B}(X)=\begin{bmatrix}1&1\\0&0\end{bmatrix}X,$$

$$\boldsymbol{C}(X)=\begin{bmatrix}0&1\\0&0\end{bmatrix}X,$$

则

$$\begin{aligned}\boldsymbol{AB}&=\boldsymbol{B},&\quad \boldsymbol{BA}&=\boldsymbol{A},\\\boldsymbol{AC}&=\boldsymbol{C},&\quad \boldsymbol{CA}&=\boldsymbol{O},\\\boldsymbol{BC}&=\boldsymbol{C},&\quad \boldsymbol{CB}&=\boldsymbol{O}.\end{aligned}$$

例 4 设 A,B 是 $\mathbf{P}^{n\times n}$中两个取定的矩阵. 定义

$$\boldsymbol{A}(X)=AX,\qquad X\in\mathbf{P}^{n\times n};$$

$$\boldsymbol{B}(X)=XB,\qquad X\in\mathbf{P}^{n\times n};$$

则 $\boldsymbol{A},\boldsymbol{B}$ 都是 $\mathbf{P}^{n\times n}$ 的线性变换. 根据定义

$$\boldsymbol{AB}(X) = \boldsymbol{A}(\boldsymbol{B}(X)) = \boldsymbol{A}(XB) = \boldsymbol{A}(XB) = AXB;$$
$$\boldsymbol{BA}(X) = \boldsymbol{B}(\boldsymbol{A}(X)) = \boldsymbol{B}(AX) = (AX)B = AXB.$$

所以

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

上面的一些例子说明线性变换的乘法不适合交换律及消去律;两个非零线性变换的乘积可能是零变换.这些结果与矩阵的乘法很类似.同样地,与矩阵的乘法一样,我们也有:

命题 4 线性变换的乘法满足结合律.即对于线性空间 V 的任意 3 个线性变换 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$,都有

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

证明:任取 V 中向量 α,都有

$$[\boldsymbol{A}(\boldsymbol{BC})](\alpha) = \boldsymbol{A}((\boldsymbol{BC})(\alpha)) = \boldsymbol{A}(\boldsymbol{B}(\boldsymbol{C}(\alpha)));$$
$$[(\boldsymbol{AB})\boldsymbol{C}](\alpha) = (\boldsymbol{AB})(\boldsymbol{C}(\alpha)) = \boldsymbol{A}(\boldsymbol{B}(\boldsymbol{C}(\alpha))).$$

所以

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

以后,就把$(\boldsymbol{AB})\boldsymbol{C}$记成$\boldsymbol{ABC}$,而不必写上括号.根据结合律,当若干个线性变换$\boldsymbol{A}$相乘时,其最终结果与乘积的结合方法无关,因此,当$n$个线性变换$\boldsymbol{A}$相乘时,可以用

$$\underbrace{\boldsymbol{AA}\cdots\boldsymbol{A}}_{m个}(m \geqslant 1).$$

来表示所得的乘积,而不必在中间加上任何括号.为了方便起见,我们把这个乘积记作$\boldsymbol{A}^m$,称为$\boldsymbol{A}$的m次方幂.

规定

$$\boldsymbol{A}^0 = \boldsymbol{E}.$$

根据定义,可知$\boldsymbol{A}$的方幂满足指数法则:

$$\boldsymbol{A}^m\boldsymbol{A}^n = \boldsymbol{A}^{m+n},(\boldsymbol{A}^m)^n = \boldsymbol{A}^{mn} \quad (m,n \geqslant 0).$$

需要注意的是,对于线性变换,乘积的指数法则是不成立的.即,一般来说,线性变换$\boldsymbol{A}$与$\boldsymbol{B}$,$(\boldsymbol{AB})^n$与$\boldsymbol{A}^n\boldsymbol{B}^n$不一定相等.

单位变换 $\boldsymbol{E}$ 在线性变换的乘法运算中起着单位元素的作用.我们还可引入逆变换的概念.

定义 5 设 $\boldsymbol{A}$ 是线性空间 V 的一个变换.如果有 V 的变换 $\boldsymbol{B}$,使得

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{E}.$$

就称变换 $\boldsymbol{A}$ 是**可逆的**.$\boldsymbol{B}$ 称为 $\boldsymbol{A}$ 的**逆变换**.

$\boldsymbol{A}$ 的逆变换记作 $\boldsymbol{A}^{-1}$.因此 $\boldsymbol{AA}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{E}$.

命题 5 如果数域 $\mathbf{P}$ 上线性空间 V 的线性变换 $\boldsymbol{A}$ 是可逆的,那么 $\boldsymbol{A}$ 的逆变换 $\boldsymbol{A}^{-1}$也是线性变换.

证明: 对于 V 中任意向量 α,β 及 $\mathbf{P}$ 中任意数 k,有

$$\begin{aligned}\boldsymbol{A}^{-1}(\alpha+\beta) &= \boldsymbol{A}^{-1}[\boldsymbol{AA}^{-1}(\alpha)+\boldsymbol{AA}^{-1}(\beta)]\\ &= \boldsymbol{A}^{-1}[\boldsymbol{A}(\boldsymbol{A}^{-1}(\alpha))+\boldsymbol{A}(\boldsymbol{A}^{-1}(\beta))]\\ &= \boldsymbol{A}^{-1}[\boldsymbol{A}(\boldsymbol{A}^{-1}(\alpha)+\boldsymbol{A}^{-1}(\beta))]\\ &= (\boldsymbol{A}^{-1}\boldsymbol{A})[\boldsymbol{A}^{-1}(\alpha)+\boldsymbol{A}^{-1}(\beta)]\\ &= \boldsymbol{A}^{-1}(\alpha)+\boldsymbol{A}^{-1}(\beta);\end{aligned}$$

$$\begin{aligned}\boldsymbol{A}^{-1}(k\alpha) &= \boldsymbol{A}^{-1}[k\boldsymbol{AA}^{-1}(\alpha)] = \boldsymbol{A}^{-1}[k\boldsymbol{A}(\boldsymbol{A}^{-1}(\alpha))]\\ &= \boldsymbol{A}^{-1}[\boldsymbol{A}(k\boldsymbol{A}^{-1}(\alpha))] = (\boldsymbol{A}^{-1}\boldsymbol{A})(k\boldsymbol{A}^{-1}(\alpha))\\ &= k\boldsymbol{A}^{-1}(\alpha),\end{aligned}$$

所以 $\boldsymbol{A}^{-1}$是一个线性变换.

例 5 在线性空间 $\mathbf{P}^{n\times n}$ 中取定一个可逆矩阵 A.定义 $\mathbf{P}^{n\times n}$的线性变换 $\boldsymbol{A},\boldsymbol{B}$ 如下:

$$\boldsymbol{A}X = OX, \qquad X \in \mathbf{P}^{n\times n};$$
$$\boldsymbol{B}X = A^{-1}X, \quad X \in \mathbf{P}^{n\times n};$$

于是

$$(\boldsymbol{AB})X = \boldsymbol{A}(\boldsymbol{B}X) = \boldsymbol{A}(A^{-1}X) = A(A^{-1}X) = X;$$
$$(\boldsymbol{BA})X = \boldsymbol{B}(\boldsymbol{A}X) = \boldsymbol{B}(AX) = A^{-1}(AX) = X;$$

所以

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{E}.$$

因此，$\boldsymbol{A}$ 是一个可逆变换，$\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆变换.

下述定理给出了判断线性变换 $\boldsymbol{A}$ 是否可逆的一个方法.

定理 2 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是线性空间 V 的一组基，$\boldsymbol{A}$ 是 V 的一个线性变换，则 $\boldsymbol{A}$ 是可逆变换的充分必要条件是：向量组 $\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 线性无关.

证明：必要性. 当 $\boldsymbol{A}$ 可逆时，$\boldsymbol{A}$ 的逆变换把 $\boldsymbol{A}\varepsilon_i$ 变为 ε_i：

$$\boldsymbol{A}^{-1}(\boldsymbol{A}\varepsilon_i) = (\boldsymbol{A}^{-1}\boldsymbol{A})(\varepsilon_i) = \boldsymbol{E}(\varepsilon_i) = \varepsilon_i, \qquad i = 1,2,\cdots,n.$$

因为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性无关，因此根据定理 2，$\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 必须线性无关.

充分性. 如果 $\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 线性无关，则由定理 1 存在 V 的一个线性变换 $\boldsymbol{B}$，使得

$$\boldsymbol{B}(\boldsymbol{A}\varepsilon_i) = \varepsilon_i, \quad i = 1,2,\cdots,n.$$

于是

$$\boldsymbol{B}\boldsymbol{A}\varepsilon_i = \varepsilon_i = \boldsymbol{E}\varepsilon_i, \quad i = 1,2,\cdots,n.$$

因此，由定理 1 的推论，得

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

再对等式 $\boldsymbol{B}\boldsymbol{A}\varepsilon_i=\varepsilon_i$ 两边用 $\boldsymbol{A}$ 作用，得

$$\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}\varepsilon_i = \boldsymbol{A}\varepsilon_i, \quad i = 1,2,\cdots,n.$$

即

$$(\boldsymbol{A}\boldsymbol{B})(\boldsymbol{A}\varepsilon_i) = \boldsymbol{E}(\boldsymbol{A}\varepsilon_i) \quad i = 1,2,\cdots,n.$$

所以，同样由定理 1 的推论可得

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

故由定义，知 $\boldsymbol{A}$ 可逆. |

定理 2 不但说明当 $\boldsymbol{A}$ 可逆时，$\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 线性无关，并且它的证明给出了求逆变换的一个方法.

例 6 定义 $\mathbf{P}^3$ 的变换 $\boldsymbol{A}$ 为

$$\boldsymbol{A}(a_1,a_2,a_3) = (a_1 + a_2,a_2 + a_3,a_3).$$

容易验证 $\boldsymbol{A}$ 是一个线性变换. 取 $\mathbf{P}^3$ 的一组基：

$$\varepsilon_1=(1,0,0),\quad \varepsilon_2(0,1,0),\quad \varepsilon_3=(0,0,1),$$

则

$$A\varepsilon_1=(1,0,0),\quad A\varepsilon_2=(1,1,0),\quad A\varepsilon_3=(0,1,1).$$

因为 $A\varepsilon_1,A\varepsilon_2,A\varepsilon_3$ 线性无关，所以 A 是一个可逆变换.

根据定理 2，A 的逆变换 A^{-1} 应该满足

$$A^{-1}(1,0,0)=(1,0,0),$$
$$A^{-1}(1,1,0)=(0,1,0),$$
$$A^{-1}(0,1,1)=(0,0,1).$$

因为(1,0,0),(1,1,0),(0,1,1)也是 $\mathbf{P}^3$ 的一组基，因此根据定理 1，A^{-1} 由这些条件唯一确定. 经计算，可以求得 A^{-1} 为

$$A^{-1}(a_1,a_2,a_3)=(a_1-a_2+a_3,a_2-a_3,a_3).$$

定理 2 有一些很有用的推论：

推论 1 数域 $\mathbf{P}$ 上有限维线性空间 V 的线性变换 A 是可逆变换的充分必要条件是：A 把非零向量变为非零向量.

这个结论当 V 是无限维空间时也是成立的. 下面我们对有限维空间 V 来加以证明.

证明： 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组基.

如果 A 是可逆的，那么由定理 2 知，$A\varepsilon_1,A\varepsilon_2,\cdots,A\varepsilon_n$ 线性无关. 设 α 是 V 中任一个非零向量，把 α 表成 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的线性组合：

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

其中系数 $a_1,a_2,\cdots,a_n$ 不全为零. 于是

$$A\alpha=a_1A\varepsilon_1+a_2A\varepsilon_2+\cdots+a_nA\varepsilon_n\neq 0,$$

故 A 把非零向量变为非零向量.

假设 A 把非零向量变成非零向量. 我们只要证 $A\varepsilon_1,A\varepsilon_2,\cdots,A\varepsilon_n$ 是线性无关的. 因此由定理 2 知，A 是可逆变换. 用反证法. 如果 $A\varepsilon_1,A\varepsilon_2,\cdots,A\varepsilon_n$ 线性相关，则有 $\mathbf{P}$ 中不全为零的数 $k_1,k_2,\cdots,k_n$，使得

$$k_1A\varepsilon_1 + k_2A\varepsilon_2 + \cdots + k_nA\varepsilon_n = 0.$$

于是，由于 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性无关的，故

$$k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n$$

是 V 中一个非零向量. 但是它在 A 下的象为：

$$\begin{aligned} &A(k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n) \\ &\quad = k_1A\varepsilon_1 + k_2A\varepsilon_2 + \cdots + k_nA\varepsilon_n = 0. \end{aligned}$$

与假设矛盾.

推论 2 可逆线性变换把线性无关的向量组变为线性无关的向量组.

证明：设 A 是线性空间 V 的一个可逆线性变换. $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是 V 中一组线性无关的向量. 用推论 1 中的证明方法可证，如果 $A\alpha_1, A\alpha_2, \cdots, A\alpha_s$ 是线性相关的，则 A 把 V 中某个非零向量变为零向量. 这是不可能的.

在下一节我们介绍了线性变换的矩阵之后，还将介绍判断一个线性变换是否可逆的，以及求逆变换的另一种方法.

下述命题说明矩阵的乘法与其他运算间的关系.

命题 6 (1) 线性变换的加法与乘法满足左、右分配律，即

$$A(B + C) = AB + AC;$$

$$(B + C)A = BA + CA.$$

(2) 线性变换的乘法与数量乘法满足

$$k(AB) = (kA)B = A(kB).$$

证明： 我们只证明左分配律，其它规律的证明请读者作为练习，自己证明.

设 A, B, C 都是线性空间 V 的线性变换. 对于 V 中任意向量 α，有

$$\begin{aligned} (A(B + C))\alpha &= A((B + C)\alpha) = A(B\alpha + C\alpha) \\ &= A(B\alpha) + A(C\alpha) = (AB)(\alpha) + (AC)(\alpha) \\ &= (AB + AC)\alpha, \end{aligned}$$

所以结合律成立.

最后,我们介绍线性变换的多项式的概念. 设 $\boldsymbol{A}$ 是数域 $\mathbf{P}$ 上线性空间 V 的一个线性变换,$f(x)$是 $\mathbf{P}[x]$中的一个多项式:

$$f(x)=a_m x^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0.$$

规定

$$f(\boldsymbol{A})=a_m\boldsymbol{A}^m+a_{m-1}\boldsymbol{A}^{m-1}+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{E},$$

则 $f(\boldsymbol{A})$也是 V 的一个线性变换,称为线性变换 $\boldsymbol{A}$ 的一个多项式.

容易验证,如果 $f(x),g(x)$都是 $\mathbf{P}[x]$中的多项式,设

$$h(x)=f(x)+g(x),\quad p(x)=f(x)g(x),$$

则

$$h(\boldsymbol{A})=f(\boldsymbol{A})+g(\boldsymbol{A}),\qquad p(\boldsymbol{A})=f(\boldsymbol{A})g(\boldsymbol{A}).$$

同一个线性变换的多项式对于乘法是可交换的,即

$$f(\boldsymbol{A})g(\boldsymbol{A})=g(\boldsymbol{A})f(\boldsymbol{A}).$$

例 7 设 $\boldsymbol{A}$ 是 $\mathbf{P}^4$ 的线性变换:

$$\boldsymbol{A}(x_1,x_2,x_3,x_4)=(0,x_1,x_2,x_3).$$

则

$$\boldsymbol{A}^2(x_1,x_2,x_3,x_4)=(0,0,x_1,x_2);$$
$$\boldsymbol{A}^3(x_1,x_2,x_3,x_4)=(0,0,0,x_1);$$
$$\boldsymbol{A}^4(x_1,x_2,x_3,x_4)=(0,0,0,0);$$

所以

$$\boldsymbol{A}^k=\boldsymbol{O},\qquad k=4,5,\cdots.$$

如再设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

是数域 $\mathbf{P}$ 上的一个多项式,则

$$f(\boldsymbol{A})(x_1,x_2,x_3,x_4)=(a_0x_1,a_0x_2+a_1x_1,$$
$$a_0x_3+a_1x_2+a_2x_1,a_0x_4+a_1x_3+a_2x_2+a_3x_1),$$

其中仍约定当 $k>n$ 时,$a_k=0$.

例 8 仍用 $\boldsymbol{A}$ 表示线性空间 $\mathbf{P}^{n\times n}$ 的由下式决定的线性变换：

$$\boldsymbol{A}(X) = AX, \qquad X \in \mathbf{P}^{n\times n},$$

则

$$(k\boldsymbol{A})(X) = (kA)X;$$

$$\boldsymbol{A}^k(X) = A^kX;$$

$$(f(\boldsymbol{A}))(X) = f(A)X$$

$$(X \in \mathbf{P}^{n\times n}, \quad f(X) \in \mathbf{P}[x]).$$

习　题　9.2

1. 设线性空间 $\mathbf{P}^3$ 的线性变换 $\boldsymbol{A},\boldsymbol{B}$ 如下：

$$\boldsymbol{A}(x_1,x_2,x_3) = (x_1,x_2,x_1+x_2)$$

$$\boldsymbol{B}(x_1,x_2,x_3) = (x_1+x_2-x_3,0,-x_1-x_2)$$

(1) 求 $\boldsymbol{A}+\boldsymbol{B},\boldsymbol{A}-\boldsymbol{B},2\boldsymbol{A}$；

(2) 求 $\boldsymbol{AB},\boldsymbol{BA},\boldsymbol{A}^2$.

2. 已知线性空间 $\mathbf{P}^3$ 的线性变换 $\boldsymbol{A}$ 为.

$$\boldsymbol{A}(x_1,x_2,x_3) = (x_1+x_2+x_3,x_2+x_3,x_3)$$

证明 $\boldsymbol{A}$ 为可逆变换，并求 $\boldsymbol{A}^{-1}$.

3. 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 都是线性空间 V 的线性变换，试证：

(1) 如果 $\boldsymbol{A},\boldsymbol{B}$ 都与 $\boldsymbol{C}$ 可交换，则 $\boldsymbol{AB},\boldsymbol{A}^2$ 也都与 $\boldsymbol{C}$ 可交换；

(2) 如果 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 可交换，且 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{-1}$与 $\boldsymbol{C}$ 也可交换；

(3) 如果 $\boldsymbol{A},\boldsymbol{B}$ 与 $\boldsymbol{C}$ 可交换，则 $\boldsymbol{A}+\boldsymbol{B},\boldsymbol{A}-\boldsymbol{B},k\boldsymbol{A}$ 也与 $\boldsymbol{C}$ 可交换；

(4) 如果 $\boldsymbol{A}+\boldsymbol{B},\boldsymbol{A}-\boldsymbol{B}$ 都与 $\boldsymbol{C}$ 可交换，则 $\boldsymbol{A},\boldsymbol{B}$ 也都与 $\boldsymbol{C}$ 可交换.

4. 设 $\boldsymbol{A},\boldsymbol{B}$ 都是数域 $\mathbf{P}$ 上线性空间 V 的线性变换，$f(x)$是 $\mathbf{P}[x]$中一个多项式，试证：如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可交换，则 $f(\boldsymbol{A})$与 $\boldsymbol{B}$ 也可交换.

5. 设 V 是数域 $\mathbf{P}$ 上一个线性空间，$\boldsymbol{A}$ 是 V 的一个线性变换，用 $L(V)$表示 V 的全部线性变换构成的数域 $\mathbf{P}$ 上的线性空间，试证：

(1) V 的与 $\boldsymbol{A}$ 可交换的线性变换全体组成 $L(V)$的一个线性子空间；

(2) $\boldsymbol{A}$ 的系数在 $\mathbf{P}$ 中的多项式全体组成 $L(V)$的一个线性子空间；

(3) V 满足 $\boldsymbol{AX}=\boldsymbol{O}$ 的线性变换 $\boldsymbol{X}$ 全体组成 $L(V)$的一个线性子空间.

9.3 线性变换的矩阵

这一节讨论线性变换的矩阵表示.

设 $\boldsymbol{A}$ 是数域 $\mathbf{P}$ 上线性空间 V 的一个线性变换. 上节中证明了 $\boldsymbol{A}$ 被它在 V 的一组基上的作用唯一确定. 根据这一点,我们引入线性变换的矩阵的概念.

定义 6 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是数域 $\mathbf{P}$ 上 n 维线性空间 V 的一组基;$\boldsymbol{A}$ 是 V 的一个线性变换,$\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 可以唯一地由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性表出,设为

$$\begin{cases}\boldsymbol{A}\varepsilon_1 = a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + \cdots + a_{n1}\varepsilon_n,\\ \boldsymbol{A}\varepsilon_2 = a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{n2}\varepsilon_n,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \boldsymbol{A}\varepsilon_n = a_{1n}\varepsilon_1 + a_{2n}\varepsilon_2 + \cdots + a_{nn}\varepsilon_n.\end{cases} \tag{1}$$

矩阵

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

称为 **$\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵**.

例 1 设 V 是 n 维线性空间,那么 V 的恒等变换 $\boldsymbol{E}$ 在任一组基下的矩阵都是 n 阶单位矩阵 E;零变换 $\boldsymbol{O}$ 在任一组基下的矩阵都是 n 阶零矩阵 0;数乘变换 $k\boldsymbol{E}$ 在任一组基下的矩阵都是数量矩阵 kE.

例 2 设 $\mathbf{P}^3$ 的线性变换 $\boldsymbol{A}$ 为

$$\boldsymbol{A}(x_1,x_2,x_3) = (x_1,x_2,x_1 + x_2);$$

取 $\mathbf{P}^3$ 的一组基

$$\varepsilon_1 = (1,0,0), \qquad \varepsilon_2 = (0,1,0), \qquad \varepsilon_3 = (0,0,1).$$

则

$$\begin{aligned} A\varepsilon_1 &= \varepsilon_1 \qquad + \varepsilon_3, \\ A\varepsilon_2 &= \qquad \varepsilon_2 + \varepsilon_3, \\ A\varepsilon_3 &= \mathbf{0}. \end{aligned}$$

所以 $\boldsymbol{A}$ 在 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

例 3 在 $\mathbf{P}_n[x]$中取定一组基

$$\varepsilon_1 = 1, \varepsilon_2 = x, \varepsilon_3 = x^2, \cdots, \varepsilon_n = x^{n-1}.$$

$\boldsymbol{D}$ 为 $\mathbf{P}_n[x]$的微分变换：

$$\boldsymbol{D}f(x) = f'(x).$$

因为

$$\begin{aligned} \boldsymbol{D}\varepsilon_1 &= \mathbf{0}, \\ \boldsymbol{D}\varepsilon_2 &= \quad \varepsilon_1, \\ \boldsymbol{D}\varepsilon_3 &= \qquad 2\varepsilon_2, \\ &\cdots\cdots\cdots\cdots\cdots \\ \boldsymbol{D}\varepsilon_n &= \qquad\qquad (n-1)\varepsilon_{n-1}. \end{aligned}$$

所以 $\boldsymbol{D}$ 在这一组基下的矩阵为

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & n-1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

如果另取一组基

$$\eta_1 = 1, \eta_2 = 1 + x, \eta_3 = 1 + x + \frac{x^2}{2!}, \cdots,$$

$$\eta_n = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{n-1}}{(n-1)!}.$$

则因

$$\boldsymbol{D}\eta_1 = \boldsymbol{0}, \qquad \boldsymbol{D}\eta_2 = \eta_1, \cdots, \boldsymbol{D}\eta_n = \eta_{n-1},$$

所以 $\boldsymbol{D}$ 在这一组基下的矩阵是

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

为了以后应用起来方便起见，我们仍用形式的记法来表示线性变换的矩阵，用 $\boldsymbol{A}(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ 表示 $(\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2, \cdots, \boldsymbol{A}\varepsilon_n)$，于是(1)式可表成

$$\begin{aligned} \boldsymbol{A}[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n] &= [\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2, \cdots, \boldsymbol{A}\varepsilon_n] \\ &= [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \\ &= [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n] A. \end{aligned}$$

这就是说，如果线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵为 A，那么

$$\boldsymbol{A}(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) A. \tag{2}$$

此外，如果已知 $\boldsymbol{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵为 A，则 $\boldsymbol{A}$ 的作用可由(2)式决定.

下面这个定理是应用矩阵来讨论线性变换的理论根据.

定理 3 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是数域 $\mathbf{P}$ 上 n 维线性空间 V 的一组基. 在这组基下，V 的每个线性变换都对应于数域 $\mathbf{P}$ 上的一个 n 阶矩阵. 这个对应具有以下的性质：

(1) 线性变换的和对应于矩阵的和；

(2) 线性变换的乘积对应于矩阵的乘积；

(3) 线性变换的数量乘积对应于矩阵的数量乘积.

证明： 设 $\boldsymbol{A},\boldsymbol{B}$ 是 V 的两个线性变换，它们在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵分别是 A,B. 即

$$\boldsymbol{A}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]A,$$

$$\boldsymbol{B}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]B.$$

(1) 由

$$\begin{aligned}(\boldsymbol{A}+\boldsymbol{B})&[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]\\&=\boldsymbol{A}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]+\boldsymbol{B}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]\\&=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]A+[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]B\\&=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n](A+B)\end{aligned}$$

即得线性变换 $\boldsymbol{A}+\boldsymbol{B}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是 $A+B$.

(2) 因为

$$\begin{aligned}(\boldsymbol{AB})[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]&=\boldsymbol{A}(\boldsymbol{B}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n])\\&=\boldsymbol{A}([\varepsilon_1,\cdots,\varepsilon_n]B)\\&=(\boldsymbol{A}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n])B\\&=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]AB,\end{aligned}$$

所以线性变换 $\boldsymbol{AB}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是 AB.

(3) 因为

$$\begin{aligned}(k\boldsymbol{A})[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]&=k(\boldsymbol{A}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n])\\&=k[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]A\\&=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n](kA),\end{aligned}$$

所以，线性变换 $k\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是 kA. |

在数域 **P** 上 n 维线性空间 V 中取定了一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 之后，V 的每个线性变换 $\boldsymbol{A}$ 都对应于一个数域 **P** 上的 n 阶矩阵，即 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵. 根据定理 1，不同的线性变换对应于不

同的矩阵. 而且, 任给一个数域 $\mathbf{P}$ 上的 n 阶矩阵 A, 根据(2)式, 一定有一个 V 的线性变换 $\boldsymbol{A}$, 使得 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵就等于 A. 这样就建立了由 V 的线性变换到数域 $\mathbf{P}$ 上 n 级矩阵间的一个一一对应. 不仅如此, 定理 3 说明这个映射还保持运算, 因此可以应用矩阵来讨论线性变换. 由于这个映射保持加法和数量乘法, 所以有下述定理:

定理 4 设 V 是数域 $\mathbf{P}$ 上的一个 n 维线性空间. V 的全体线性变换所构成的数域 $\mathbf{P}$ 上的线性空间 $L(V)$ 与线性空间 $\mathbf{P}^{n\times n}$ 是同构的. 由此可知, $L(V)$ 的维数等于 n^2.

可以应用线性变换的矩阵来判断这个线性变换是否可逆.

定理 5 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是线性空间 V 的一组基, $\boldsymbol{A}$ 是 V 的一个线性变换. 并设 $\boldsymbol{A}$ 在这组基下的矩阵是 A, 则 $\boldsymbol{A}$ 可逆的充分必要条件是 A 可逆. 并且, 当 $\boldsymbol{A}$ 可逆时, $\boldsymbol{A}^{-1}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是 A^{-1}.

证明: 已知 $\boldsymbol{A}$ 可逆的充分必要条件是 $\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 线性无关. 由假设

$$[\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n]=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]A,$$

得知, $\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 线性无关的充要条件是 A 可逆.

当 $\boldsymbol{A}$ 可逆时, 矩阵 A 可逆. 作 V 的一个线性变换 $\boldsymbol{B}$, 使 $\boldsymbol{B}$ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 A^{-1}. 于是, $\boldsymbol{AB}$ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 $AA^{-1}=E$; $\boldsymbol{BA}$ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 $A^{-1}A=E$, 所以

$$\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{E},\quad \boldsymbol{B}=\boldsymbol{A}^{-1}.$$

定理证毕. |

例 4 设 $\mathbf{P}_4$ 的线性变换 $\boldsymbol{A}$ 为

$$\boldsymbol{A}(x_1,x_2,x_3,x_4)=(x_1+x_2,x_1-x_2,x_3+x_4,x_3-x_4),$$

求 $\boldsymbol{A}$ 在基

$$\varepsilon_1=(1,0,0,0),\qquad \varepsilon_2=(0,1,0,0),$$
$$\varepsilon_3=(0,0,1,0),\qquad \varepsilon_4=(0,0,0,1)$$

下的矩阵.并判断 $\boldsymbol{A}$ 是否为可逆变换.

解：因为

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{\varepsilon}_1 &= (1,1,0,0) = \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2,\\
\boldsymbol{A}\boldsymbol{\varepsilon}_2 &= (1,-1,0,0) = \boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2,\\
\boldsymbol{A}\boldsymbol{\varepsilon}_3 &= (0,0,1,1) = \boldsymbol{\varepsilon}_3 + \boldsymbol{\varepsilon}_4,\\
\boldsymbol{A}\boldsymbol{\varepsilon}_4 &= (0,0,1,-1) = \boldsymbol{\varepsilon}_3 - \boldsymbol{\varepsilon}_4,
\end{aligned}
$$

所以 $\boldsymbol{A}$ 在 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4$ 下的矩阵为

$$
\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}.
$$

因为 A 可逆,故 $\boldsymbol{A}$ 可逆;且可知 $\boldsymbol{A}^{-1}$在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4$ 下的矩阵为

$$
A^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix},
$$

所以 $\boldsymbol{A}^{-1}$为

$$
\boldsymbol{A}^{-1}[\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4] = [\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4]A^{-1},
$$

即

$$
\begin{aligned}
\boldsymbol{A}^{-1}(x_1,x_2,x_3,x_4) = \Big(& \frac{1}{2}x_1 + \frac{1}{2}x_2, \frac{1}{2}x_1 - \frac{1}{2}x_2,\\
& \frac{1}{2}x_3 + \frac{1}{2}x_4, \frac{1}{2}x_3 - \frac{1}{2}x_4 \Big).
\end{aligned}
$$

应用线性变换的矩阵,可以直接计算向量在此变换下的象.

定理 6 设线性空间 V 的线性变换 $\boldsymbol{A}$ 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 下的矩

阵为 A. 如果 V 中向量 ξ 对于基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的坐标为 $(x_1,x_2,\cdots,x_n)$，则 $\mathcal{A}\xi$ 对于基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的坐标 $(y_1,y_2,\cdots,y_n)$ 可按公式

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

计算.

证明：由假设

$$\mathcal{A}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n] = [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]A,$$

$$\xi = [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

因此

$$\mathcal{A}\xi = \mathcal{A}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]A\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

另一方面，由假设又有

$$\mathcal{A}\xi = [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

因为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是线性无关的，所以

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

线性变换的矩阵是与所取的基有关的.同一个线性变换在不同基下的矩阵一般是不相等的.因此，如何选择一组基使得一个线性变换在这组基下的矩阵最简单，就成为一个重要的问题.为了解决这个问题，以便于利用矩阵来研究线性变换，必须弄清楚同一个线性变换在不同基下的矩阵间的关系.

定理 7　(1)同一个线性变换在两组基下的矩阵是相似的.确切地说，如果线性空间 V 的线性变换 $\boldsymbol{A}$ 在 V 的两组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n;$$

$$\eta_1,\eta_2,\cdots,\eta_n$$

下的矩阵分别是 A 和 B，由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到 $\eta_1,\eta_2,\cdots,\eta_n$ 的过渡矩阵为 X，则

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

(2) 相似矩阵可以看作同一个线性变换在不同基下的矩阵.

证明：(1)由假设

$$\boldsymbol{A}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n] = [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]A,$$

$$\boldsymbol{A}[\eta_1,\eta_2,\cdots,\eta_n] = [\eta_1,\eta_2,\cdots,\eta_n]B;$$

$$[\eta_1,\eta_2,\cdots,\eta_n] = [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]X,$$

于是，

$$\begin{aligned}\boldsymbol{A}[\eta_1,\eta_2,\cdots,\eta_n] &= \boldsymbol{A}([\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]X) \\ &= (\boldsymbol{A}[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n])X = [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]AX \\ &= [\eta_1,\eta_2,\cdots,\eta_n]X^{-1}AX.\end{aligned}$$

由此即得

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

(2) 设 n 阶矩阵 A,B 相似:

$$B = C^{-1}AC \qquad (|C| \neq 0).$$

任取 V 的一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. 设 V 的线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 A. 以 C 为过渡矩阵,由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 得一组基 $\eta_1,\eta_2,\cdots,\eta_n$:

$$[\eta_1,\eta_2,\cdots,\eta_n] = [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]C,$$

则 $\boldsymbol{A}$ 在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的矩阵为 $C^{-1}AC$. |

例 5 在例 3 中,由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到 $\eta_1,\eta_2,\cdots,\eta_n$ 的过渡矩阵为 C,

$$C = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & \frac{1}{2!} & \cdots & \frac{1}{2!} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \frac{1}{(n-1)!} \end{bmatrix}.$$

请读者自己验证:

$$C^{-1}\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & n-1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

定理 7 指出,如果线性变换 $\boldsymbol{A}$ 在一组基下的矩阵为 A,那么,线性变换 $\boldsymbol{A}$ 在其他基下的矩阵就给出了与 A 相似的全部矩阵. 因此,求 $\boldsymbol{A}$ 的最简单的矩阵就是求与 A 相似的最简单的矩阵. 问题归结为相似矩阵的标准形问题. 我们以前曾经讨论过这个问题. 在以后几节中,我们将在复习已有结果的基础上,进一步讨论这个问

题.

习 题 9.3

1. 设 $\mathbf{P}^3$ 的线性变换 $\boldsymbol{A}$ 定义如下:

$$\boldsymbol{A}(x_1,x_2,x_3)=(2x_1-x_2,x_2-x_3,x_2+x_3)$$

求 $\boldsymbol{A}$ 在基

$$\boldsymbol{\varepsilon}_1=(1,0,0),\ \boldsymbol{\varepsilon}_2=(0,1,0),\ \boldsymbol{\varepsilon}_3=(0,0,1)$$

及基

$$\boldsymbol{\eta}_1=(1,1,0),\ \boldsymbol{\eta}_2=(0,1,1),\ \boldsymbol{\eta}_3=(0,0,1)$$

下的矩阵.

2. 在 $\mathbf{P}^3$ 中,线性变换 $\boldsymbol{B}$ 定义如下:

$$\boldsymbol{B}\xi_1=(-5,0,3),\ \boldsymbol{B}\xi_2=(0,-1,6),\ \boldsymbol{B}\xi_3=(-5,-1,0),$$

其中

$$\xi_1=(-1,0,2),\ \xi_2=(0,1,1),\ \xi_3=(3,-1,0).$$

求 $\boldsymbol{B}$ 在基 ξ_1,ξ_2,ξ_3 下的矩阵及 $\boldsymbol{B}$ 在基 $\boldsymbol{\varepsilon}_1=(1,0,0),\boldsymbol{\varepsilon}_2=(0,1,0),\boldsymbol{\varepsilon}_3=(0,0,1)$ 下的矩阵.

3. $\boldsymbol{A},\boldsymbol{B}$ 如第 1,2 题,

(1) 证明:$\boldsymbol{A},\boldsymbol{B}$ 都是可逆线性变换;

(2) 求 $\boldsymbol{A}^{-1},\boldsymbol{AB},\boldsymbol{A}+\boldsymbol{B}$ 在基(1,0,0),(0,1,0),(0,0,1)下的矩阵.

4. 在 $\mathbf{P}^{2\times 2}$ 中定义线性变换 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 为:

$$\boldsymbol{A}(X)=\begin{bmatrix}a & b\\ c & d\end{bmatrix}X,$$

$$\boldsymbol{B}(X)=X\begin{bmatrix}a & b\\ c & d\end{bmatrix},$$

$$\boldsymbol{C}(X)=\begin{bmatrix}a & b\\ c & d\end{bmatrix}X\begin{bmatrix}a & b\\ c & d\end{bmatrix}\quad (X\in\mathbf{P}^{2\times 2}),$$

求 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 在基 $E_{11},E_{12},E_{21},E_{22}$ 下的矩阵.

5. (1) 证明:上题中线性变换 $\boldsymbol{A}$ 可逆的条件为

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\neq 0;$$

(2) 当 A 可逆时，求 A^{-1} 及 A^{-1} 在基 $E_{11}E_{12}E_{21}E_{22}$ 下的矩阵.

6. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 是线性空间 V 的一组基. 已知线性变换 A 在 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为

$$A=\begin{bmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix},$$

(1) 求 A 在基 $\varepsilon_1,\varepsilon_3,\varepsilon_2$ 下的矩阵;

(2) 求 A 在基 $\varepsilon_1,k\varepsilon_2,\varepsilon_3$ 下的矩阵($k\neq 0$);

(3) 求 A 在基 $\varepsilon_1,\varepsilon_1+\varepsilon_2,\varepsilon_3$ 下的矩阵.

9.4 线性变换的特征值与特征向量

在有限维线性空间 V 中取定了一组基之后，V 的线性变换可以用矩阵表示. 为了利用矩阵研究线性变换，对于每个给定的线性变换，希望能找到一组基，使得这个线性变换在这组基下的矩阵具有最简单的形式. 由于同一个线性变换在不同基下的矩阵是相似的，因此这就归结为相似矩阵的标准形问题. 这个问题我们以前曾经讨论过. 为了把关于相似矩阵的一些结果应用于线性变换，必须引入线性变换的特征值与特征向量等概念.

定义 7 设 V 是数域 $\mathbf{P}$ 上一个线性空间；A 是 V 的一个线性变换，α 是 V 中一个非零向量. 如果

$$A(\alpha)=\lambda_0\alpha$$

λ_0 是 $\mathbf{P}$ 中一个常数，则 λ_0 称为 A 的一个**特征值**(**特征根**)，α 称为 A 的一个**属于特征值 λ_0 的特征向量**，简称**特征向量**.

因此，λ_0 是 A 的特征值的意思就是可找到 V 中一个非零向量 α，使得

$$A\alpha=\lambda_0\alpha.$$

例 1 任何非零向量既是 E 的属于特征值 1 的特征向量，又是零变换的属于特征值 0 的特征向量.

例 2 考虑 $P[x]$的微分变换 $\boldsymbol{D}$,因为对于 $\mathbf{P}$ 中非零常数 c,有

$$\boldsymbol{D}(c)=0.$$

因此 c 是 $\boldsymbol{D}$ 的属于特征值 0 的特征向量.而且由于当次 $f(x)\geqslant 1$ 时:

$$\deg(\boldsymbol{D}f(x))=\deg f'(x)=\deg f(x)-1.$$

因此,此时 $f(x)$不可能是 $\boldsymbol{D}$ 的特征向量,所以 $\boldsymbol{D}$ 只有一个特征值: 0,而且,$\mathbf{P}$ 中非零常数是 $\boldsymbol{D}$ 的全部特征向量.

下面讨论线性变换的特征值、特征向量与它的矩阵的特征值、特征向量之间的关系.

设 $\boldsymbol{A}$ 是 n 维线性空间 V 的一个线性变换.取定 V 的一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$.设 $\boldsymbol{A}$ 在这组基下的矩阵为 A:

$$\boldsymbol{A}[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]=[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]A.$$

再设 α 是 $\boldsymbol{A}$ 的一个属于特征值 λ_0 的特征向量:

$$\boldsymbol{A}\alpha=\lambda_0\alpha.$$

如果 α 对于基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的坐标为$(x_1, x_2, \cdots, x_n)$,则 $\boldsymbol{A}\alpha$ 对于基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的坐标为

$$A\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}.$$

又因为 $\boldsymbol{A}\alpha=\lambda_0\alpha$,所以 $\boldsymbol{A}\alpha$ 对于 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的坐标应为

$$\lambda_0\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}.$$

因此

$$A\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\lambda_0\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}.$$

由于 $\alpha\neq 0$,故$(x_1,x_2,\cdots,x_n)$是非零向量.这说明 λ_0 是矩阵 A 的特征值.$(x_1,x_2,\cdots,x_n)$是 A 的属于特征值 λ_0 的特征向量.把上面推导的过程逆推过去,还可看出,如果 A 是线性变换 $\boldsymbol{A}$ 在基 $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$, $\cdots$, $\boldsymbol{\varepsilon}_n$ 下的矩阵,λ_0 是 A 的一个特征值,$(x_1,x_2,\cdots,x_n)$是 A 的属于特征值 λ_0 的特征向量,那么,λ_0 是 $\boldsymbol{A}$ 的一个特征值,而

$$\begin{aligned}\alpha&=[\boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2,\ \cdots,\ \boldsymbol{\varepsilon}_n]\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\\&=x_1\boldsymbol{\varepsilon}_1+x_2\boldsymbol{\varepsilon}_2+\cdots+x_n\boldsymbol{\varepsilon}_n\end{aligned}$$

是 $\boldsymbol{A}$ 的一个属于 λ_0 的特征向量.

这样,我们就可以应用求线性变换的矩阵的特征值和特征向量的方法来计算线性变换的特征值及特征向量.为此,首先引入下列概念.

定义 8 线性变换 $\boldsymbol{A}$ 在任一组基下的矩阵的特征多项式称为线性变换 $\boldsymbol{A}$ 的**特征多项式**.

由于 $\boldsymbol{A}$ 在不同基下的矩阵都是相似的,而相似矩阵的特征多项式是相同的,所以 $\boldsymbol{A}$ 的特征多项式不依赖于基的选择.$\boldsymbol{A}$ 的特征多项式的根就是 $\boldsymbol{A}$ 的特征值.

我们把求线性变换的特征值与特征向量的方法叙述于下:

设 $\boldsymbol{A}$ 是线性空间 V 的一个线性变换.

(1) 任取 V 的一组基 $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$, $\cdots$, $\boldsymbol{\varepsilon}_n$,求出 $\boldsymbol{A}$ 在这组基下的矩阵为 A.

(2) 计算 A 的特征多项式 $f(\lambda)=|\lambda E-A|$.

(3) 求出 $f(\lambda)$的全部根，就是 A 的全部特征值.

(4) 设 λ_0 是 A 的一个特征值，解齐次线性方程组

$$(\lambda_0 E - A)\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0.$$

求出一个基础解系：

$$(c_{11}, c_{12}, \cdots, c_{1n}), (c_{21}, c_{22}, \cdots, c_{2n}),$$
$$\cdots, (c_{s1}, c_{s2}, \cdots, c_{sn}).$$

令

$$\xi_1 = c_{11}\varepsilon_1 + c_{12}\varepsilon_2 + \cdots + c_{1n}\varepsilon_n,$$
$$\xi_2 = c_{21}\varepsilon_1 + c_{22}\varepsilon_2 + \cdots + c_{2n}\varepsilon_n,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\xi_s = c_{s1}\varepsilon_1 + c_{s2}\varepsilon_2 + \cdots + c_{sn}\varepsilon_n,$$

则

$$\xi_1, \xi_2, \cdots, \xi_s$$

就是 A 的属于 λ_0 的一组线性无关的特征向量. A 的任一个属于 λ_0 的特征向量都可表示成它们的线性组合，即可表成

$$k_1\xi_1 + k_2\xi_2 + \cdots + k_s\xi_s$$

其中，$k_1, k_2, \cdots, k_s$ 是数域 $\mathbf{P}$ 中 s 个不全为零的数. 这就是 A 的属于特征值 λ_0 的全部特征向量.

(5) 对 A 的所有不同的特征值求出属于每个特征值的全部特征向量，就得到 A 的全部特征向量.

例 3 设 A 是数域 $\mathbf{P}$ 上 3 维线性空间 V 的一个线性变换. ε_1, ε_2, ε_3 是 V 的一组基. 已知

$$A\varepsilon_1 = \varepsilon_1 + 2\varepsilon_2 - 2\varepsilon_3,$$
$$A\varepsilon_2 = 2\varepsilon_1 + \varepsilon_2 - 2\varepsilon_3,$$
$$A\varepsilon_3 = 2\varepsilon_1 - 2\varepsilon_2 + \varepsilon_3.$$

求 $\boldsymbol{A}$ 的全部特征值及特征向量.

解: 由假设,知 $\boldsymbol{A}$ 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 下的矩阵是

$$A=\begin{bmatrix}1&2&2\\2&1&-2\\-2&-2&1\end{bmatrix}.$$

A 的特征多项式为

$$|\lambda E-A|=\begin{vmatrix}\lambda-1&-2&-2\\-2&\lambda-1&2\\2&2&\lambda-1\end{vmatrix}$$
$$=(\lambda-1)(\lambda+1)(\lambda-3),$$

所以 $\boldsymbol{A}$ 的特征值为 1,−1,3.

对 $\lambda=1$,解齐次方程组

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

得基础解系(1,−1,1),所以 $\boldsymbol{A}$ 的属于特征值 1 的全部特征向量为

$$k_1(\boldsymbol{\varepsilon}_1-\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_3)\qquad(k_1\text{ 为 }\mathbf{P}\text{ 中任意非零常数}).$$

对 $\lambda=-1$,解齐次方程组

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

得基础解系(1,−1,0),所以 $\boldsymbol{A}$ 的属于特征值−1 的全部特征向量为

$$k_2(\boldsymbol{\varepsilon}_1-\boldsymbol{\varepsilon}_2)\quad(k_2\neq 0).$$

由 $\lambda=3$,解齐次方程组

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

得基础解系(0,1,−1),所以 $\boldsymbol{A}$ 的属于特征值 3 的全部特征向量为

$$k_3(\boldsymbol{\varepsilon}_2 - \boldsymbol{\varepsilon}_3) \quad (k_3 \neq 0).$$

线性变换 $\boldsymbol{A}$ 的特征向量关于基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 的坐标就是 $\boldsymbol{A}$ 在 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 下的矩阵的特征向量. 因此关于矩阵的特征向量的结论对于线性变换也成立. 我们把一些重要结论叙述于下:

(1) 线性变换 $\boldsymbol{A}$ 的属于同一个特征值的特征向量的非零线性组合仍是 $\boldsymbol{A}$ 的属于这个特征值的特征向量.

(2) $\boldsymbol{A}$ 的属于不同特征值的特征向量的和不再是 $\boldsymbol{A}$ 的特征向量.

(3) $\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的. 即: 设 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是 $\boldsymbol{A}$ 的不同的特征值, $\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{is_i}$ 是 $\boldsymbol{A}$ 的属于 $\lambda_i (i=1,2,\cdots,m)$ 的线性无关的特征向量, 则向量组

$$\alpha_{11}, \cdots, \alpha_{1s_1}, \alpha_{21}, \cdots, \alpha_{2s_2}, \cdots, \alpha_{m1}, \cdots, \alpha_{ms_m}$$

是线性无关的.

我们知道, 并不是任一个矩阵都可对角化的, 因此, 任给一个线性变换 $\boldsymbol{A}$, 并不是总可能找到一组基使得 $\boldsymbol{A}$ 在这组基下的矩阵是对角形矩阵. 因为同一个线性变换在不同基下的矩阵是相似的, 而且相似矩阵可看成同一个线性变换在不同基下的矩阵, 因此可以根据矩阵可对角化的条件得出线性变换在某组基下的矩阵是对角形矩阵的条件. 即

(4) $\boldsymbol{A}$ 是 n 维线性空间 V 的一个线性变换. 存在 V 的一组基使得 $\boldsymbol{A}$ 在这组基下的矩阵是对角矩阵的条件是:

1) 充分必要条件: $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

2) 充分条件: $\boldsymbol{A}$ 有 n 个不同的特征值.

3）充分条件：当 P 是复数域时，$\boldsymbol{A}$ 的特征多项式没有重根.

例 4 例 3 中的线性变换 $\boldsymbol{A}$ 有 3 个线性无关的特征向量. 取

$$\eta_1=\varepsilon_1-\varepsilon_2+\varepsilon_3,\qquad \eta_2=\varepsilon_1-\varepsilon_2,\qquad \eta_3=\varepsilon_2-\varepsilon_3$$

为基，则因

$$\boldsymbol{A}\eta_1=\eta_1,\qquad \boldsymbol{A}\eta_2=-\eta_2,\qquad \boldsymbol{A}\eta_3=3\eta_3,$$

所以 $\boldsymbol{A}$ 在这组基下的矩阵

$$\begin{bmatrix}1 & & \\ & -1 & \\ & & 3\end{bmatrix}$$

是一个对角矩阵.

应用矩阵在相似关系下的标准形的结果，我们还可得到：

（5）$\boldsymbol{A}$ 是 n 维复线性空间 V 的一个线性变换，则

1）$\boldsymbol{A}$ 一定有特征值及特征向量；

2）可找到 V 的一组基，使 $\boldsymbol{A}$ 在此基下的矩阵是上三角形矩阵；

3）可找到 V 的一组基，使 $\boldsymbol{A}$ 在此基下的矩阵是约当形矩阵.

以上介绍了用相似矩阵的结论讨论线性变换所得的一些结果. 从另一方面来看，我们也可以应用线性空间和线性变换的一些概念和方法来讨论矩阵. 有些矩阵问题用线性空间及线性变换的方法比较容易解决. 下面我们举几个简单的例子. 更进一步的讨论将在下一节中进行.

例 5 证明数域 P 上的矩阵

$$\begin{bmatrix}a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n\end{bmatrix} \text{ 与 } \begin{bmatrix}a_{i_1} & & & \\ & a_{i_2} & & \\ & & \ddots & \\ & & & a_{i_n}\end{bmatrix}$$

相似，其中 $a_1, a_2, \cdots, a_n$ 是 P 中 n 个数；$i_1, i_2, \cdots, i_n$ 是 $1,2,\cdots,$

n 的一个排列.

证明: 任取 $\mathbf{P}^n$ 的一组基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$. 作 $\mathbf{P}^n$ 的线性变换 $\boldsymbol{A}$:

$$\boldsymbol{A}\boldsymbol{\varepsilon}_k = a_k\boldsymbol{\varepsilon}_k, \qquad k = 1,2,\cdots,n,$$

则 $\boldsymbol{A}$ 在基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 下的矩阵是

$$A = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}.$$

令

$$\boldsymbol{\eta}_1 = \boldsymbol{\varepsilon}_{i_1}, \boldsymbol{\eta}_2 = \boldsymbol{\varepsilon}_{i_2}, \cdots, \boldsymbol{\eta}_n = \boldsymbol{\varepsilon}_{i_n},$$

则 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n$ 也是 $\mathbf{P}^n$ 的一组基. 由于

$$\boldsymbol{A}\boldsymbol{\eta}_1 = a_{i_1}\boldsymbol{\eta}_1,$$

$$\boldsymbol{A}\boldsymbol{\eta}_2 = a_{i_2}\boldsymbol{\eta}_2,$$

$$\cdots\cdots\cdots\cdots$$

$$\boldsymbol{A}\boldsymbol{\eta}_n = a_{i_n}\boldsymbol{\eta}_n.$$

因此 $\boldsymbol{A}$ 在基 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n$ 下的矩阵是

$$B = \begin{bmatrix} a_{i_1} & & & \\ & a_{i_2} & & \\ & & \ddots & \\ & & & a_{i_n} \end{bmatrix}.$$

所以 A 与 B 是相似的.

例 6 设 λ_0 是 A 的特征多项式的 k 重根,试证:A 的属于 λ_0 的线性无关的特征向量最多有 k 个.

证明: 设 V 是一个 n 维线性空间,$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 V 的一组基. 再设线性变换 $\boldsymbol{A}$ 在 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵为 A,则 A 的特征多项式即为 $\boldsymbol{A}$ 的特征多项式.

如果 A 有 l 个线性无关的属于 λ_0 的特征向量,那么 $\boldsymbol{A}$ 也有 l

个线性无关的属于 λ_0 的特征向量，设为 $\eta_1,\eta_2,\cdots,\eta_l$. 把 $\eta_1,\eta_2,\cdots,\eta_l$ 扩充成 V 的一组基：

$$\eta_1,\ \eta_2,\ \cdots,\ \eta_l,\eta_{l+1},\ \cdots,\ \eta_n.$$

则因

$$\begin{aligned} A\eta_1 &= \lambda_0\eta_1, \\ A\eta_2 &= \lambda_0\eta_2, \\ &\cdots\cdots \\ A\eta_l &= \lambda_0\eta_l, \end{aligned}$$

所以 A 在这组基下的矩阵 B 具有下述形式：

$$B=\begin{bmatrix} \lambda_0 & 0 & \cdots & 0 & b_{1,l+1} & \cdots & b_{1n} \\ 0 & \lambda_0 & \cdots & 0 & b_{2,l+1} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_0 & b_{l,l+1} & \cdots & b_{ln} \\ 0 & 0 & \cdots & 0 & b_{l+1,l+1} & \cdots & b_{l+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & b_{n,l+1} & \cdots & b_{nn} \end{bmatrix}.$$

于是

$$|\lambda E-B|=(\lambda-\lambda_0)^l g(\lambda).$$

因为 B 与 A 是相似的，所以 $|\lambda E-B|=|\lambda E-A|$，因此 $l\leqslant k$.

这个例子还说明了如何应用线性变换的特征向量化简这个线性变换的矩阵的一个一般方法.

习　题　9.4

1. 设 $\mathbf{A}$ 是 $\mathbf{P}^3$ 的一个线性变换，已知

$$\begin{aligned} \mathbf{A}(1,0,0) &= (5,6,-3), \\ \mathbf{A}(0,1,0) &= (-1,0,1), \\ \mathbf{A}(0,0,1) &= (1,2,1), \end{aligned}$$

求 $\mathbf{A}$ 的特征值和特征向量. 并回答能否找到 $\mathbf{P}^3$ 的一组基，使 $\mathbf{A}$ 在这组基

下的矩阵为对角矩阵？

2. V 是 4 维实线性空间，$\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是 V 的一组基. 已知 V 的线性变换 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的矩阵为

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

(1) 求 $\boldsymbol{A}$ 的全部特征值及特征向量；

(2) 求 V 的一组基，使 $\boldsymbol{A}$ 在这组基下的矩阵是对角矩阵.

3. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是 4 维线性空间 V 的一组基，V 的线性变换 $\boldsymbol{A}$ 在这组基下的矩阵为

$$A=\begin{bmatrix} 5 & -2 & -4 & 3 \\ 3 & -1 & -3 & 2 \\ -3 & 1/2 & 9/2 & -5/2 \\ -10 & 3 & 11 & -7 \end{bmatrix}.$$

(1) 求 $\boldsymbol{A}$ 在基

$$\begin{aligned} \eta_1 &= \varepsilon_1+2\varepsilon_2+\varepsilon_3+\varepsilon_4, \\ \eta_2 &= 2\varepsilon_1+3\varepsilon_2+\varepsilon_3, \\ \eta_3 &= \varepsilon_3, \\ \eta_4 &= \varepsilon_4 \end{aligned}$$

下的矩阵；

(2) 求 $\boldsymbol{A}$ 的全部特征值与特征向量；

(3) 求 V 的一组基使 $\boldsymbol{A}$ 在这组基下的矩阵是对角矩阵.

4. $\boldsymbol{A}$ 是 n 维空间 V 的一个线性变换. 试证：$\boldsymbol{A}$ 是可逆变换的充分必要条件为：$\boldsymbol{A}$ 的特征值全不等于 0.

5. 设 $\boldsymbol{A}$ 是线性空间 V 的一个线性变换，λ_0 是 $\boldsymbol{A}$ 的特征多项式的一个 k 重根. 再设

$$W=\{\alpha\in V\,|\,\boldsymbol{A}\alpha=\lambda_0\alpha\};$$

(1) 证明：W 是 V 的一个子空间；

(2) 证明：$\dim W\leqslant k$.

9.5 不变子空间

上节中最后一个例子说明线性变换的特征向量可以用来化简这个线性变换的矩阵. 这一节将要介绍的概念——线性变换的不变子空间是特征向量的推广,它更深入地说明线性变换的矩阵的化简与线性变换的内在联系.

定义 9 设 A 是数域 $\mathbf{P}$ 上线性空间 V 的一个线性变换,W 是 V 的一个子空间. 如果 W 中的向量在 A 下的象仍在 W 中,即对于 W 中任一个向量 ξ,都有 $A\xi \in W$,就称 W 是 A 的一个**不变子空间**.

线性变换 A 的不变子空间也称 **A-子空间**.

例 1 V 的平凡子空间: V 及零子空间都是 V 的任一个线性变换的不变子空间.

例 2 V 的任一个子空间都是数乘变换的不变子空间.

例 3 设 V 中向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 都是 V 的线性变换 A 的特征向量,那么 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$是一个 A-子空间.

证明: 设 α_i 对应的特征值为 λ_i,即

$$A\alpha_i = \lambda_i\alpha_i, \qquad i = 1, 2, \cdots, s.$$

那么,由于 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$中任一向量 ξ 都可表示成

$$\xi = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s,$$

于是

$$\begin{aligned} A\xi &= A(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s) \\ &= \lambda_1 k_1\alpha_1 + \lambda_2 k_2\alpha_2 + \cdots + \lambda_s k_s\alpha_s \\ &\in L(\alpha_1, \alpha_2, \cdots, \alpha_s), \end{aligned}$$

所以 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$是 A 的一个不变子空间.

特别地,由 A 的一个特征向量生成的子空间是一个一维 A-子空间.

例 4 设 $\boldsymbol{A}$ 是 V 的一个线性变换，λ_0 是 $\boldsymbol{A}$ 的一个特征值. $\boldsymbol{A}$ 的属于 λ_0 的全部特征向量添上零向量组成一个子空间，记作 V_{λ_0}：

$$V_{\lambda_0} = \{\alpha \in V \mid \boldsymbol{A}\alpha = \lambda_0\alpha\},$$

则 V_{λ_0}是一个 $\boldsymbol{A}$-子空间，称为 $\boldsymbol{A}$ 的一个**特征子空间**.

例 5 仍用 $\boldsymbol{D}$ 表示 $\mathbf{P}_n[x]$的微分变换. 证明：$\mathbf{P}_k[x]$($k\leqslant n$)是 $\boldsymbol{D}$ 的全部不变子空间.

证明：由于

$$\deg(\boldsymbol{D}f(x)) = \deg(f(x)) - 1, \qquad \text{当} \deg f(x) \geqslant 1;$$

$$\boldsymbol{D}f(x) = 0, \qquad \text{当} \deg f(x) = 0.$$

由此可见，$\mathbf{P}_k[x]$($k\leqslant n$)都是 $\boldsymbol{D}$-不变子空间.

反之，设 W 是 $\boldsymbol{D}$ 的一个不变子空间，且 $W=\{0\}$，则 $W=\mathbf{P}_0[x]$；否则，可取 W 中一个次数最高的多项式，设为

$$f(x) = a_lx^l + a_{l-1}x^{l-1} + \cdots + a_1x + a_0 \quad (l < n, a_l \neq 0).$$

由于 W 是 $\boldsymbol{D}$ 的不变子空间，所以 $f'(x)$，$f''(x)$，…都在 W 中，即

$$f'(x) = la_lx^{l-1} + (l-1)a_{l-1}x^{l-2} + \cdots + a_1 \in W,$$

$$f''(x) = l(l-1)a_lx^{l-2} + \cdots \in W,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$f^{(l-1)}(x) = l!a_lx + (l-1)!a_{l-1} \in W,$$

$$f^{(l)}(x) = l!a_l \in W.$$

由这些等式可推出 $1, x, x^2, \cdots, x^l$ 都在 W 中. 又因 W 中多项式的次数最高为 l，所以

$$W = \mathbf{P}_{l+1}[x] \qquad (l+1 \leqslant n).$$

可以用基来判断一个子空间是否是不变子空间.

命题 1 $\boldsymbol{A}$ 是 V 的一个线性变换，W 是 V 的一个非零子空间，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_s$ 是 W 的一组基，则 W 是 $\boldsymbol{A}$-子空间的充分必要条件是 $\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2, \cdots, \boldsymbol{A}\varepsilon_s$ 全在 W 中.

命题 2 $\boldsymbol{A}$ 是数域 $\mathbf{P}$ 上线性空间 V 的一个线性变换，则 $\boldsymbol{A}$-子

空间的交与和仍都是 A-子空间.

证明：设 W_1, W_2 是两个 A-子空间. 已知 $W_1 \cap W_2$ 及 $W_1 + W_2$ 都是 V 的子空间.

任取 $\alpha \in W_1 \cap W_2$，则由它的定义，知

$$\alpha \in W_1, \alpha \in W_2.$$

由于 W_1, W_2 都是 A-子空间，因此

$$A\alpha \in W_1,\ A\alpha \in W_2,$$

于是

$$A\alpha \in W_1 \cap W_2,$$

故 $W_1 \cap W_2$ 是 A 的不变子空间.

同样可证 $W_1 + W_2$ 是 A-子空间，将这个证明留给读者作为习题.

下面介绍如何应用 A-子空间来化简 A 的矩阵.

设 A 是数域 $\mathbf{P}$ 上 n 维线性空间 V 的一个线性变换，W 是 A 的一个非平凡不变子空间. 在 W 中任取一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m$ $(0 < m < n)$，把它扩充成 V 的一组基：

$$\varepsilon_1, \cdots, \varepsilon_m, \varepsilon_{m+1}, \cdots, \varepsilon_n.$$

由于 $A\varepsilon_i \in W\ (i = 1, 2, \cdots, m)$，故可设

$$\begin{aligned}
A\varepsilon_1 &= a_{11}\varepsilon_1 + \cdots + a_{m1}\varepsilon_m, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
A\varepsilon_m &= a_{1m}\varepsilon_1 + \cdots + a_{mm}\varepsilon_m, \\
A\varepsilon_{m+1} &= a_{1,m+1}\varepsilon_1 + \cdots + a_{m,m+1}\varepsilon_m + \cdots + a_{n,m+1}\varepsilon_n, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
A\varepsilon_n &= a_{1n}\varepsilon_1 + \cdots + a_{mn}\varepsilon_m + \cdots + a_{nn}\varepsilon_n.
\end{aligned}$$

因此，A 在这一组基下的矩阵是

$$A=\begin{bmatrix} a_{11} & \cdots & a_{1m} & a_{1,m+1} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mm} & a_{m,m+1} & \cdots & a_{mn} \\ 0 & \cdots & 0 & a_{m+1,m+1} & \cdots & a_{m+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & a_{n,m+1} & \cdots & a_{nn} \end{bmatrix}.$$

可以把 A 写成分块矩阵：

$$A=\begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}. \tag{1}$$

反之，如果线性变换 $\boldsymbol{A}$ 在基 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_m, \cdots, \boldsymbol{\varepsilon}_n$ 下的矩阵具有(1)式的形式，其中 A_1 是一个 $m(1\leqslant m\leqslant n)$级矩阵，那么由 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_m$ 生成的子空间是一个 $\boldsymbol{A}$-子空间.

更进一步，如果 V 可以分解成两个非平凡 $\boldsymbol{A}$-子空间的直和：

$$V=W_1 \dotplus W_2.$$

取 W_1 的一组基 $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_m(1\leqslant m<n)$，再取 W_2 的一组基 $\boldsymbol{\varepsilon}_{m+1}, \cdots, \boldsymbol{\varepsilon}_n$，则

$$\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_m, \boldsymbol{\varepsilon}_{m+1}, \cdots, \boldsymbol{\varepsilon}_n \tag{2}$$

是 V 的一组基. 由于 W_1, W_2 都是 $\boldsymbol{A}$-子空间，因此有

$$\begin{aligned} \boldsymbol{A}\boldsymbol{\varepsilon}_1 &= a_{11}\boldsymbol{\varepsilon}_1+\cdots+a_{m1}\boldsymbol{\varepsilon}_m, \\ &\cdots\cdots \\ \boldsymbol{A}\boldsymbol{\varepsilon}_m &= a_{1m}\boldsymbol{\varepsilon}_1+\cdots+a_{mm}\boldsymbol{\varepsilon}_m, \\ \boldsymbol{A}\boldsymbol{\varepsilon}_{m+1} &= a_{m+1,m+1}\boldsymbol{\varepsilon}_{m+1}+\cdots+a_{n,m+1}\boldsymbol{\varepsilon}_n, \\ &\cdots\cdots \\ \boldsymbol{A}\boldsymbol{\varepsilon}_n &= a_{m+1,n}\boldsymbol{\varepsilon}_{m+1}+\cdots+a_{nn}\boldsymbol{\varepsilon}_n. \end{aligned}$$

所以，$\boldsymbol{A}$ 在这组基下的矩阵为

$$A=\begin{bmatrix} a_{11} & \cdots & a_{1m} & 0 & \cdots & 0 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1} & \cdots & a_{mm} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & a_{m+1,m+1} & \cdots & a_{m+1,n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ 0 & \cdots & 0 & a_{n,m+1} & \cdots & a_{nn} \end{bmatrix}.$$

这是一个准对角矩阵，可以写成

$$A=\begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix}. \tag{3}$$

同样地，如果线性变换 $\boldsymbol{A}$ 在基(2)下的矩阵具有(3)式的形式，则

$$W_1=L(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_m),\quad W_2=L(\boldsymbol{\varepsilon}_{m+1},\cdots,\boldsymbol{\varepsilon}_n)$$

都是 $\boldsymbol{A}$ 的不变子空间，并且

$$V=W_1\dotplus W_2.$$

这个结论还可以推广到多个 $\boldsymbol{A}$-子空间的情形：如果线性空间 V 可以分解成若干个非平凡 $\boldsymbol{A}$-子空间的直和：

$$V=W_1\dotplus W_2\dotplus\cdots\dotplus W_s\quad(s\geqslant 2),$$

在每一个 $\boldsymbol{A}$-子空间 W_i 中取定一组基：

$$\boldsymbol{\varepsilon}_{i1},\boldsymbol{\varepsilon}_{i2},\cdots,\boldsymbol{\varepsilon}_{im_i}\quad(i=1,2,\cdots,s), \tag{4}$$

把它们合起来成为 V 的一组基：

$$\boldsymbol{\varepsilon}_{11},\cdots,\boldsymbol{\varepsilon}_{1m_1},\boldsymbol{\varepsilon}_{21},\cdots,\boldsymbol{\varepsilon}_{2m_2},\cdots,\boldsymbol{\varepsilon}_{s1},\cdots,\boldsymbol{\varepsilon}_{sm_s}. \tag{5}$$

那么由于 $W_i(i=1,2,\cdots,s)$ 是不变子空间，$\boldsymbol{A}\boldsymbol{\varepsilon}_{ij}(j=1,2,\cdots,m_i)$ 仍在 W_i 中，可以表成 $\boldsymbol{\varepsilon}_{i1},\boldsymbol{\varepsilon}_{i2},\cdots,\boldsymbol{\varepsilon}_{im_i}$ 的线性组合，所以 $\boldsymbol{A}$ 在这组基下的矩阵是准对角矩阵

$$\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{bmatrix} \tag{6}$$

其中 A_i 是 m_i 阶矩阵　$(i=1,2,\cdots,s)$.

反之，如果 $\boldsymbol{A}$ 在基(5)下的矩阵是准对角矩阵(6)，其中 $A_i(i=1,2,\cdots,s)$ 是 m_i 阶矩阵. 那么由(4)生成的子空间 $W_i(i=1,2,\cdots,s)$ 是 $\boldsymbol{A}$ 的不变子空间，且 V 是它们的直和.

由此可知，线性变换的矩阵的化简与不变子空间有着密切的联系.

为了更进一步应用不变子空间来讨论线性变换，下面这个概念是很有用的. 设 $\boldsymbol{A}$ 是线性空间 V 的一个线性变换，W 是 $\boldsymbol{A}$ 的一个不变子空间. 由于 W 中向量在 $\boldsymbol{A}$ 下的象仍在 W 中，因此，如果只在线性空间 W 中考虑 $\boldsymbol{A}$ 的作用，就可以把 $\boldsymbol{A}$ 看成 W 的一个线性变换，这个变换称为 $\boldsymbol{A}$ 在不变子空间 W 上引起的变换. 为了区别起见，我们用符号 $\boldsymbol{A}|W$ 来表示它；但是在不会引起误解的情况下，仍可以用 $\boldsymbol{A}$ 来表示.

必须在概念上弄清楚 $\boldsymbol{A}$ 与 $\boldsymbol{A}|W$ 的差别：$\boldsymbol{A}$ 是线性空间 V 的线性变换，V 中每个向量在 $\boldsymbol{A}$ 下有确定的象；$\boldsymbol{A}|W$ 是线性空间 W 的线性变换，对于 W 中每个向量 ξ，有

$$(\boldsymbol{A}|W)\xi = \boldsymbol{A}\xi.$$

但是对于 V 中不属于 W 的向量 η，$(\boldsymbol{A}|W)\eta$ 是没有意义的.

应用线性变换在不变子空间上所引起的线性变换这个概念，可以更深刻地理解不变子空间与线性变换的矩阵的化简之间的关系. 在前面的讨论中，矩阵(3)中左上角的 m 阶矩阵 A_1 就是线性变换 $\boldsymbol{A}|W$ 在 W 的基 $\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_m$ 下的矩阵. 而在线性空间 V 分解成线性变换 $\boldsymbol{A}$ 的若干个不变子空间的直和的情形，准对角矩阵(6)中的 m_i 阶矩阵 $A_i(i=1,2,\cdots,s)$ 就是线性变换 $\boldsymbol{A}|W_i$ 在 W_i 的

基(4)下的矩阵.因此把 $\boldsymbol{A}$ 的矩阵更进一步化简的问题可以归结为化简 W_i 的线性变换 $\boldsymbol{A}|W_i$ 的矩阵的问题,从而可以在子空间 W_i 中进行讨论.

例 6 证明:如果复线性空间 V 的线性变换 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可交换,那么它们有公共的特征向量.

证明:因为 V 是一个复数域上的线性空间,所以 $\boldsymbol{A}$ 一定有特征值.设 λ_0 是 $\boldsymbol{A}$ 的一个特征值.现在我们来证 $\boldsymbol{A}$ 的特征子空间 V_{λ_0} 是 $\boldsymbol{B}$ 的一个不变子空间:

任取 $\alpha\in V_{\lambda_0}$,则

$$\boldsymbol{A}(\boldsymbol{B}\alpha)=\boldsymbol{B}\boldsymbol{A}\alpha=\boldsymbol{B}(\lambda_0\alpha)=\lambda_0(\boldsymbol{B}\alpha),$$

所以 $\boldsymbol{B}\alpha\in V_{\lambda_0}$.

于是可以考虑 $\boldsymbol{B}$ 在 V_{λ_0}上引起的线性变换 $\boldsymbol{B}|V_{\lambda_0}$.因为 V_{λ_0}是一个复线性空间,所以 $\boldsymbol{B}$ 在 V_{λ_0}中有一个特征向量 α_0,而 α_0 也是 $\boldsymbol{A}$ 的特征向量,所以 $\boldsymbol{A},\boldsymbol{B}$ 有公共的特征向量.

把这个结论应用于矩阵,就得到一个非常有用的结果:如果复系数矩阵 A 与 B 可交换,那么它们有公共特征向量.

下面再介绍两个重要的不变子空间——线性变换的值域和核.

定义 10 设 $\boldsymbol{A}$ 是线性空间 V 的一个线性变换.由 $\boldsymbol{A}$ 的全体象组成的集合称为 $\boldsymbol{A}$ 的**值域**,记作 $\boldsymbol{A}V$;由所有被 $\boldsymbol{A}$ 变成零向量的向量组成的集合称为 $\boldsymbol{A}$ 的**核**,记作 $\ker\boldsymbol{A}$.

也就是说:

$$\boldsymbol{A}V=\{\boldsymbol{A}\xi|\xi\in V\};$$

$$\ker\boldsymbol{A}=\{\xi|\xi\in V,\boldsymbol{A}\xi=\mathbf{0}\}.$$

命题 3 线性变换 $\boldsymbol{A}$ 的值域及核都是 $\boldsymbol{A}$ 的不变子空间.

证明:先证 $\boldsymbol{A}V$ 是 $\boldsymbol{A}$ 的不变子空间:由于

$$\boldsymbol{A}\alpha+\boldsymbol{A}\beta=\boldsymbol{A}(\alpha+\beta);$$

$$k\boldsymbol{A}\alpha=\boldsymbol{A}(k\alpha),$$

所以 $\boldsymbol{A}V$ 对于 V 的线性运算是封闭的. 当然 $\boldsymbol{A}V$ 是非空的，因此 $\boldsymbol{A}V$ 是 V 的一个子空间. 又，从 $\boldsymbol{A}V\subset V$ 可推出 $\boldsymbol{A}(\boldsymbol{A}V)\subset \boldsymbol{A}V$，所以 $\boldsymbol{A}V$ 是 $\boldsymbol{A}$ 的不变子空间.

次证 $\ker\boldsymbol{A}$ 是 $\boldsymbol{A}$ 的不变子空间. 首先，零向量在 $\ker\boldsymbol{A}$ 中，所以 $\ker\boldsymbol{A}$ 是非空的. 如果 $\boldsymbol{\alpha},\boldsymbol{\beta}\in\ker\boldsymbol{A}$，那么

$$\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{0},\qquad \boldsymbol{A}\boldsymbol{\beta}=\boldsymbol{0}.$$

于是

$$\boldsymbol{A}(\boldsymbol{\alpha}+\boldsymbol{\beta})=\boldsymbol{A}(\boldsymbol{\alpha})+\boldsymbol{A}(\boldsymbol{\beta})=\boldsymbol{0};$$
$$\boldsymbol{A}(k\boldsymbol{\alpha})=k\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{0},$$

所以 $\ker\boldsymbol{A}$ 是 V 的一个子空间. 因为 $\ker\boldsymbol{A}$ 中向量在 $\boldsymbol{A}$ 下的象都是零向量，所以 $\ker\boldsymbol{A}$ 是 $\boldsymbol{A}$ 的不变子空间.

例 7 线性空间 V 的零变换的值域是零子空间，核是整个空间 V.

例 8 线性空间 V 的可逆变换的值域是 V，核是零子空间.

例 9 线性空间 $\mathbf{P}_n[x]$的微分变换 $\boldsymbol{D}$ 的值域是 $\mathbf{P}_{n-1}[x]$，核是一维子空间 $\mathbf{P}$.

定义 11 $\boldsymbol{A}V$ 的维数称为线性变换 $\boldsymbol{A}$ 的秩，记作 $\mathrm{r}(\boldsymbol{A})$；$\ker\boldsymbol{A}$ 的维数称为 $\boldsymbol{A}$ 的**零度**，记作 $\mathrm{nul}(\boldsymbol{A})$.

命题 4 设 $\boldsymbol{A}$ 是 n 维线性空间 V 的一个线性变换，$\boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2,\ \cdots,\ \boldsymbol{\varepsilon}_n$ 是 V 的一组基；$\boldsymbol{A}$ 在这组基下的矩阵是 A，则

(1) $\boldsymbol{A}$ 的值域是由这组基的象生成的子空间，即

$$\boldsymbol{A}V=L(\boldsymbol{A}\boldsymbol{\varepsilon}_1,\ \boldsymbol{A}\boldsymbol{\varepsilon}_2,\ \cdots,\ \boldsymbol{A}\boldsymbol{\varepsilon}_n);$$

(2) $\boldsymbol{A}$ 的秩等于 A 的秩.

证明：(1) 任取 $\boldsymbol{\xi}\in V$，则 $\boldsymbol{\xi}$ 可表成 $\boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2,\ \cdots,\ \boldsymbol{\varepsilon}_n$ 的线性组合：

$$\boldsymbol{\xi}=a_1\boldsymbol{\varepsilon}_1+a_2\boldsymbol{\varepsilon}_2+\cdots+a_n\boldsymbol{\varepsilon}_n.$$

于是

$$\begin{aligned}\boldsymbol{A}\boldsymbol{\xi}&=a_1\boldsymbol{A}\boldsymbol{\varepsilon}_1+a_2\boldsymbol{A}\boldsymbol{\varepsilon}_2+\cdots+a_n\boldsymbol{A}\boldsymbol{\varepsilon}_n\\&\in L(\boldsymbol{A}\boldsymbol{\varepsilon}_1,\ \boldsymbol{A}\boldsymbol{\varepsilon}_2,\ \cdots,\ \boldsymbol{A}\boldsymbol{\varepsilon}_n),\end{aligned}$$

故

$$\boldsymbol{A}V \subset L(\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2, \cdots, \boldsymbol{A}\varepsilon_n).$$

显然有

$$L(\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2, \cdots, \boldsymbol{A}\varepsilon_n) \subset \boldsymbol{A}V,$$

所以

$$\boldsymbol{A}V = L(\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2, \cdots, \boldsymbol{A}\varepsilon_n).$$

(2) 由(1)得

$$\mathrm{r}(\boldsymbol{A}) = \mathrm{r}\{\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2, \cdots, \boldsymbol{A}\varepsilon_n\} = \mathrm{r}(A).$$

命题 5 设 $\boldsymbol{A}$ 是 n 维线性空间 V 的一个线性变换. 下述的一些条件都是 $\boldsymbol{A}$ 为可逆变换的充分必要条件.

(1) $\boldsymbol{A}V = V$;

(2) $\ker\boldsymbol{A} = \{\boldsymbol{0}\}$;

(3) $\mathrm{r}(\boldsymbol{A}) = n$;

(4) $\mathrm{nul}(\boldsymbol{A}) = 0$.

定理 8 设 $\boldsymbol{A}$ 是 n 维线性空间 V 的一个线性变换,则

$$\mathrm{r}(A) + \mathrm{nul}\boldsymbol{A} = n.$$

证明: 在 V 中取定一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$. 设 $\boldsymbol{A}$ 在这组基下的矩阵为 A,则由命题 4 知,$\boldsymbol{A}$ 的秩等于 A 的秩.

如果向量 $\xi \in \ker\boldsymbol{A}$,那么

$$\boldsymbol{A}\xi = 0,$$

因此 ξ 对于基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的坐标$(x_1, x_2, \cdots, x_n)$满足

$$A\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{7}$$

而且任取方程(7)的一个解$(a_1, a_2, \cdots, a_n)$,则以此为坐标(对于基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$)的向量一定在 $\ker\boldsymbol{A}$ 中. 所以 $\ker\boldsymbol{A}$ 的维数等于齐次线性方程组(7)的解空间的维数,即 $n-$秩(A),故

$$r(A)+\mathrm{nul}A=n.$$

定理 8 的证明指出了求线性变换的核的一个方法.

例 10 设 $\boldsymbol{A}$ 是线性空间 V 的一个线性变换. 已知 $\boldsymbol{A}$ 在基 ε_1, ε_2, ε_3, ε_4 下的矩阵为

$$A=\begin{bmatrix}1 & -1 & 2 & 1\\ 2 & 0 & 1 & -1\\ 4 & -2 & 5 & 1\\ 5 & -1 & 4 & -1\end{bmatrix},$$

求 $\boldsymbol{A}$ 的值域和核,并求 $\boldsymbol{A}$ 的秩及零度.

解: $\boldsymbol{A}V=L(\boldsymbol{A}\varepsilon_1, \boldsymbol{A}\varepsilon_2, \boldsymbol{A}\varepsilon_3, \boldsymbol{A}\varepsilon_4)$. 因为 $r(A)=2$,而且 $\boldsymbol{A}\varepsilon_1$, $\boldsymbol{A}\varepsilon_2$ 线性无关,所以 $\boldsymbol{A}V$ 维数为 2; $\varepsilon_1+2\varepsilon_2+4\varepsilon_3+5\varepsilon_4$, $\varepsilon_1+2\varepsilon_3+\varepsilon_4$ 是它的一组基.

下面求 $\ker\boldsymbol{A}$: 如果 $\alpha=x_1\varepsilon_1+x_2\varepsilon_2+x_3\varepsilon_3+x_4\varepsilon_4\in\ker\boldsymbol{A}$,则 $\boldsymbol{A}\alpha=0$. 因此 $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的坐标为(0,0,0,0)即

$$\begin{bmatrix}1 & -1 & 2 & 1\\ 2 & 0 & 1 & -1\\ 4 & -2 & 5 & 1\\ 5 & -1 & 4 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix}.$$

解这个齐次线性方程组,得基础解系:

$$(1,-3,-2,0),\quad (1,3,0,2).$$

因此

$\ker\boldsymbol{A}=L(\varepsilon_1-3\varepsilon_2-2\varepsilon_3, \varepsilon_1+3\varepsilon_2+2\varepsilon_4)$,维数为 2, $\varepsilon_1-2\varepsilon_2-2\varepsilon_3$, $\varepsilon_1+3\varepsilon_2+2\varepsilon_4$ 是它的一组基.

最后,我们举一些例子说明线性变换及不变子空间在矩阵问题中的应用.

例 11 设 A 是一个 n 阶幂等矩阵:

$$A^2=A,$$

求证: A 相似于一个对角矩阵

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix},$$

其中 1 的个数等于 A 的秩.

证明：设 V 是一个 n 维线性空间，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 V 的一组基. 定义 V 的线性变换 $\boldsymbol{A}$ 为

$$\boldsymbol{A}(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A.$$

由 $A^2=A$，可知 $\boldsymbol{A}^2=\boldsymbol{A}$. 如果 $\alpha\in \boldsymbol{A}V$，则有 V 中向量 β，使得

$$\alpha = \boldsymbol{A}\beta.$$

于是

$$\boldsymbol{A}\alpha = \boldsymbol{A}(\boldsymbol{A}\beta) = \boldsymbol{A}^2\beta = \boldsymbol{A}\beta = \alpha, \tag{8}$$

因此

$$\boldsymbol{A}V \cap \ker\boldsymbol{A} = \{\mathbf{0}\}.$$

由定理 8，

$$V = \boldsymbol{A}V + \ker\boldsymbol{A}.$$

在 $\boldsymbol{A}V$ 中取一组基 $\eta_1, \eta_2, \cdots, \eta_r$，在 $\ker\boldsymbol{A}$ 中取一组基 $\eta_{r+1}, \cdots, \eta_n$，则

$$\eta_1, \cdots, \eta_r, \eta_{r+1}, \cdots, \eta_n \tag{9}$$

是 V 的一组基. 由(8)式及 $\ker\boldsymbol{A}$ 的定义知

$$\begin{aligned} \boldsymbol{A}\eta_1 &= \eta_1, \\ &\cdots\cdots \\ \boldsymbol{A}\eta_r &= \eta_r, \\ \boldsymbol{A}\eta_{r+1} &= \mathbf{0}, \\ &\cdots\cdots \\ \boldsymbol{A}\eta_n &= \mathbf{0}. \end{aligned}$$

因此 $\mathcal{A}$ 在基(9)下的矩阵为

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}, \tag{10}$$

其中 1 的个数等于 $\mathcal{A}V$ 的维数,即 A 的秩.

因为 $\mathcal{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵为 A,故 A 与矩阵(10)相似.

例 12 设 A 是一个 n 阶矩阵,满足 $A^2=E$,求证:A 相似于一个对角矩阵

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{bmatrix}.$$

证明:设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 n 维线性空间 V 的一组基,再设 $\mathcal{A}$ 是 V 的在此基下以 A 为矩阵的线性变换.令

$$W_1 = \{\alpha + \mathcal{A}\alpha \mid \alpha \in V\},$$
$$W_2 = \{\alpha - \mathcal{A}\alpha \mid \alpha \in V\},$$

则 W_1, W_2 都是 V 的子空间.而且因

$$\left.\begin{aligned} \mathcal{A}(\alpha + \mathcal{A}\alpha) &= \mathcal{A}\alpha + \mathcal{A}^2\alpha = \mathcal{A}\alpha + \alpha, \\ \mathcal{A}(\alpha - \mathcal{A}\alpha) &= \mathcal{A}\alpha - \mathcal{A}^2\alpha = -(\alpha - \mathcal{A}\alpha) \\ &= (-\alpha) - \mathcal{A}(-\alpha), \end{aligned}\right\} \tag{11}$$

所以 W_1, W_2 都是 V 的 $\mathcal{A}$-子空间.又因 V 中任一向量 β 都可表示成

$$\beta = \frac{1}{2}(\beta + A\beta) + \frac{1}{2}(\beta - A\beta),$$

所以

$$V = W_1 + W_2.$$

由(11)式知,W_1 中任一个向量 γ 都满足

$$A\gamma = \gamma,$$

而 W_2 中任一个向量 δ 都满足

$$A\delta = -\delta,$$

因此 $W_1 \cap W_2 = \{\mathbf{0}\}, V = W_1 \dotplus W_2$.

取 W_1 的一组基 $\varepsilon_1, \cdots, \varepsilon_m; W_2$ 的一组基 $\varepsilon_{m+1}, \cdots, \varepsilon_n$,合成 V 的一组基,则 A 在此基下的矩阵为

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{bmatrix}.$$

因为 A 与这个矩阵是线性变换 A 在不同基下的矩阵,所以是相似的.

可以把值域的概念加以推广:设 A 是线性空间 V 的一个线性变换,W 是 V 的一个子空间.用 AW 表示由 W 中向量的象组成的集合:

$$AW = \{A\alpha \mid \alpha \in W\}.$$

易证 AW 是 V 的一个子空间.用证明定理 8 的同样方法可证:

$$\dim(AW) + \dim(\ker A \cap W) = \dim W.$$

应用这个等式,可证下述例题.

例 13 设 A, B 是 n 维线性空间 V 的两个线性变换,则

$$\mathrm{r}(AB) \geqslant \mathrm{r}(A) + \mathrm{r}(B) - n.$$

证明：令 $W=\boldsymbol{B}V$，则

$$\begin{aligned}\mathrm{r}(\boldsymbol{AB}) &= \dim(\boldsymbol{AB}V) = \dim(\boldsymbol{A}W) \\ &= \dim W - \dim(\ker\boldsymbol{A} \cap W) \\ &\geqslant \dim W - \dim(\ker\boldsymbol{A}) \\ &= \mathrm{r}(\boldsymbol{B}) + \mathrm{r}(\boldsymbol{A}) - n.\end{aligned}$$

与此相应的矩阵结果就是：对于 n 阶矩阵 A,B，有

$$\mathrm{r}(AB) \geqslant \mathrm{r}(A) + \mathrm{r}(B) - n.$$

习 题 9.5

1. 设 $\boldsymbol{A}$ 是线性空间 V 的一个线性变换，α 是 V 中一个非零向量，试证：如果

$$\alpha, \boldsymbol{A}\alpha, \boldsymbol{A}^2\alpha, \cdots, \boldsymbol{A}^{k-1}\alpha \quad (k \geqslant 1)$$

线性无关，而

$$\alpha, \boldsymbol{A}\alpha, \boldsymbol{A}^2\alpha, \cdots, \boldsymbol{A}^k\alpha$$

线性相关，那么

(1) $L(\alpha, \boldsymbol{A}\alpha, \cdots, \boldsymbol{A}^{k-1}\alpha)$是 $\boldsymbol{A}$ 的一个不变子空间；

(2) $L(\alpha, \boldsymbol{A}\alpha, \cdots, \boldsymbol{A}^{k-1}\alpha)$是包含 α 的最小的 $\boldsymbol{A}$-子空间.

2. 设 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 是线性空间 V 的一组基，$\boldsymbol{A}$ 是 V 的一个线性变换. 已知 $\boldsymbol{A}$ 在基 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 下的矩阵是

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

求包含 ε_1 的最小的 $\boldsymbol{A}$-子空间.

3. $\boldsymbol{A}$ 是数域 $\mathbf{P}$ 上线性空间 V 的一个线性变换，W 是 $\boldsymbol{A}$ 的一个不变子空间，试证：

(1) W 是 $\boldsymbol{A}^2$ 的不变子空间；

(2) 对于 $\mathbf{P}[x]$中任一多项式 $f(x)$，W 是 $f(\boldsymbol{A})$的不变子空间；

(3) 如果 $\boldsymbol{A}$ 可逆，则 W 是 $\boldsymbol{A}^{-1}$的不变子空间.

4. 证明：如果 W_1、W_2 都是 $\boldsymbol{A}$-子空间，那么 W_1+W_2 也是 $\boldsymbol{A}$-子空间.

5. 设 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 是 4 维线性空间 V 的一组基. 已知 V 的线性变换 $\boldsymbol{A}$ 在这组

基下的矩阵为

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

(1) 求 $\boldsymbol{A}$ 的值域和核，并分别求它们的维数和一组基；

(2) 求 $\boldsymbol{A}V \cap \ker\boldsymbol{A}$ 及 $\boldsymbol{A}V + \ker\boldsymbol{A}$，并求它们的维数和一组基.

6. $\boldsymbol{A}$ 是线性空间 V 的一个线性变换，W 是 $\boldsymbol{A}$ 的一个不变子空间，试证：$\boldsymbol{A}|W$ 的核等于 $\ker\boldsymbol{A} \cap W$.

7. 设 V 的两个线性变换 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是可交换的，试证：$\boldsymbol{B}$ 的值域和核都是 $\boldsymbol{A}$ 的不变子空间.

8. $\boldsymbol{A}, \boldsymbol{B}$ 是 n 维线性空间 V 的两个线性变换，证明：如果 $\boldsymbol{A}$ 的秩与 $\boldsymbol{B}$ 的秩之和小于 n，则 $\boldsymbol{A}, \boldsymbol{B}$ 有公共特征向量.

9. 设 V 的线性变换 $\boldsymbol{A}$ 在基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 下的矩阵为

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

试证：(1) $\boldsymbol{A}$ 的包含 $\boldsymbol{\varepsilon}_n$ 的不变子空间只有 V；

(2) $\boldsymbol{A}$ 的任一个非零不变子空间一定包含 $\boldsymbol{\varepsilon}_1$；

(3) $\{\boldsymbol{0}\}, L(\boldsymbol{\varepsilon}_1), L(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2), \cdots, L(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_{n-1}), V$ 是 $\boldsymbol{A}$ 的全部不变子空间.

10. 设 $\boldsymbol{A}$ 是 n 维线性空间 V 的一个线性变换，证明：可找到 V 的一组基，使得 $\boldsymbol{A}$ 在这组基下的矩阵为对角矩阵的充分必要条件是：V 是 $\boldsymbol{A}$ 的特征子空间的直和.

内 容 提 要

1. 线性变换的定义

2. 线性变换的性质

(1) 向量组的线性关系及它们在线性变换下的象向量组的线性关系之间的关系.

(2) 线性变换被它在一组基上的作用唯一确定：定理1及其推论.

3. 线性变换的运算

(1) 线性变换的加法、数量乘法及乘法的定义以及它们满足的规律.

(2) 可逆变换及逆变换.

1) 线性变换可逆及逆变换的定义.

2) 线性变换为可逆变换的条件.

3) 逆变换的求法.

(3) 数域 **P** 上 n 维线性空间 V 的全部线性变换对加法和数量乘法构成数域 **P** 上一个线性空间 $L(V)$；$L(V)$ 与 $\mathbf{P}^{n\times n}$ 同构.

4. 线性变换的矩阵

(1) 线性变换的矩阵的定义.

(2) 线性变换的和、数量乘积及乘积的矩阵.

(3) 可逆线性变换及其逆变换的矩阵.

(4) 同一个线性变换在不同基下矩阵的关系.

(5) 线性空间中向量的坐标与其在线性变换下的象的坐标之间的关系.

5. 线性变换的特征值与特征向量

(1) 特征值与特征向量的定义及性质.

(2) 特征值与特征向量的求法.

(3) 线性变换的特征值对化简线性变换的应用.

(4) n 维线性空间 V 的一个线性变换在 V 的某组基下的矩阵是对角矩阵的条件.

(5) 如何找 n 维线性空间 V 的一组基，使 V 的一个给定的线

性变换在这组基下的矩阵是对角形.

6. 线性变换的不变子空间

(1) 不变子空间的定义.

(2) 线性变换的不变子空间与线性变换的矩阵的化简之间的关系.

(3) 特征子空间.

(4) 值域和核.

1) 定义和求法.

2) 秩及零度的定义.

3) 值域和核的维数关系.

复习题 9

1. 设线性空间 P^3 的线性变换 $\boldsymbol{A}$ 为

$$\boldsymbol{A}(1,0,0)=(1,0,-2),$$
$$\boldsymbol{A}(0,1,0)=(-1,1,2),$$
$$\boldsymbol{A}(0,0,1)=(1,-1,1),$$

(1) 求 $\boldsymbol{A}$ 在基$(1,0,0)$, $(0,1,0)$, $(0,0,1)$下的矩阵;

(2) 求 P^3 的另一组基,使 $\boldsymbol{A}$ 在这组基下的矩阵与 $\boldsymbol{A}$ 在基$(1,0,0)$,$(0,1,0)$,$(0,0,1)$下的矩阵相同.

2. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是数域 $\mathbf{P}$ 上线性空间 V 的一组基,$\boldsymbol{A}$ 是 V 的一个线性变换.已知

$$\boldsymbol{A}\varepsilon_1=\varepsilon_1+2\varepsilon_2,$$
$$\boldsymbol{A}\varepsilon_2=4\varepsilon_1-\varepsilon_2,$$
$$\boldsymbol{A}\varepsilon_3=7\varepsilon_3-2\varepsilon_4,$$
$$\boldsymbol{A}\varepsilon_4=-4\varepsilon_3+5\varepsilon_4,$$

(1) 求 $\boldsymbol{A}$ 的全部特征值与特征向量;

(2) 求 V 的一组基,使 $\boldsymbol{A}$ 在这组基下的矩阵是对角矩阵;

(3) 证明 $\boldsymbol{A}$ 是可逆线性变换,并求 $\boldsymbol{A}^{-1}$.

3. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是线性空间 V 的一组基,并设 V 的线性变换 $\boldsymbol{A}$ 如下:

$$\begin{aligned}
A\varepsilon_1 &= \quad 5\varepsilon_1 - \varepsilon_2 \qquad - \varepsilon_4,\\
A\varepsilon_2 &= -2\varepsilon_1 + \varepsilon_2 - \varepsilon_3 \qquad ,\\
A\varepsilon_3 &= \quad \varepsilon_1 + \varepsilon_2 - 2\varepsilon_3 - \varepsilon_4,\\
A\varepsilon_4 &= \quad \varepsilon_1 - 2\varepsilon_2 + 3\varepsilon_3 + \varepsilon_4,
\end{aligned}$$

(1) 求 A 的值域 AV 及核 $\ker A$，并分别求它们的维数和一组基；

(2) 求 A 的秩及零度；

(3) 求 $AV+\ker A$ 及 $AV\cap\ker A$，并分别求它们的维数和一组基.

4. 设 A 是线性空间 V 的一个线性变换，ξ 是 V 中一个向量. 已知 $A^{k-1}\xi\neq\mathbf{0}$，$A^k\xi=\mathbf{0}$. 试证：向量组 $\xi,A\xi,A^2\xi,\cdots,A^{k-1}\xi$ 线性无关.

5. 设 V 是一个 n 维线性空间，A 是 V 的一个线性变换. 已知 V 中向量 ξ 使 $A^{n-1}\xi\neq\mathbf{0}$，而 $A^n\xi=\mathbf{0}$，求证：

(1) A 在某组基下的矩阵是

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

(2) A 在 V 的任一组基下的矩阵都不是对角矩阵.

6. 设 A 是数域 $\mathbf{P}$ 上 n 维线性空间 V 的一个线性变换，证明：如果 A 与 V 的任一线性变换都可交换，则 A 是数乘变换.

7. 设 A 是数域 $\mathbf{P}$ 上 n 维线性空间 V 的一个线性变换，证明：如果 A 在 V 的任一组基下的矩阵都相等，则 A 是一个数乘变换.

8. A 是线性空间 V 的一个线性变换，证明：如果 V 中每个非零向量都是 A 的特征向量，那么 A 是一个数乘变换.

9. 设 V 是数域 $\mathbf{P}$ 上一个 n 维线性空间，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组基. 定义 V 的线性变换 $E_{ij}(i,j=1,2,\cdots,n)$ 为

$$E_{ij}(a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n)=a_j\varepsilon_i.$$

(1) 求 $E_{ij}(i,j=1,2,\cdots,n)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵；

(2) 证明：$E_{11},\cdots,E_{1n},E_{21},\cdots,E_{2n},\cdots,E_{n1},\cdots,E_{nn}$ 是 $L(V)$ 的一组基.

10. 设 A 是数域 $\mathbf{P}$ 上 n 维线性空间 V 的一个线性变换. 证明：

(1) 存在 $\mathbf{P}[x]$ 中一个次数 $\leqslant n^2$ 的多项式 $f(x)$，使 $f(A)=O$；

(2) 设 $f(x),g(x)$是 $\mathbf{P}[x]$中两个多项式,$d(x)$是 $f(x),g(x)$的一个最大公因式. 如果 $f(\boldsymbol{A})=\boldsymbol{O},g(\boldsymbol{A})=\boldsymbol{O}$,则 $d(\boldsymbol{A})=\boldsymbol{O}$.

11. 设 $\boldsymbol{A}$ 是数域 $\mathbf{P}$ 的 n 维线性空间 V 的一个线性变换,证明:$\boldsymbol{A}$ 可逆的充分必要条件是:存在一个 $\mathbf{P}[x]$中常数项不等于零的多项式 $f(x)$,使得 $f(\boldsymbol{A})=0$.

12. 设 $\boldsymbol{A}$ 是 n 维线性空间 V 的一个线性变换,$\eta_1,\eta_2,\cdots,\eta_r$ 是 $\boldsymbol{A}V$ 的一组基,且

$$\boldsymbol{A}\varepsilon_i = \eta_i, \qquad i = 1,2,\cdots,r,$$

证明:(1) $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$ 线性无关.

(2) $V=L(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r)\dotplus\ker\boldsymbol{A}$

13. $\boldsymbol{A}$ 是 n 维线性空间 V 的一个线性变换,证明:如果 $\boldsymbol{A}^2=\boldsymbol{O}$,并且 $\boldsymbol{A}$ 的秩为 r,则可找到 V 中一组基,使得 $\boldsymbol{A}$ 在这组基下的矩阵为

$$\begin{bmatrix} O & N \\ O & O \end{bmatrix},$$

其中 N 是一个$(n-r)\times r$ 矩阵,并且 N 的秩等于 r.

14. 设 $\boldsymbol{A},\boldsymbol{B}$ 是线性空间 V 的两个线性变换,试证:

(1) 如果 $\boldsymbol{AB}=\boldsymbol{B}$, $\boldsymbol{BA}=\boldsymbol{A}$,则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的值域;

(2) 如果 $\boldsymbol{AB}=\boldsymbol{A}$, $\boldsymbol{BA}=\boldsymbol{B}$,则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的核.

15. 设 $\boldsymbol{A},\boldsymbol{B}$ 都是线性空间 V 的线性变换,并且 $\boldsymbol{AB}=\boldsymbol{BA}$. 求证:

(1) $\boldsymbol{A}$ 的特征子空间一定是 $\boldsymbol{B}$ 的不变子空间.

(2) 如果 V 是复数域上的线性空间,则 $\boldsymbol{A},\boldsymbol{B}$ 一定有公共特征向量.

16. 设矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可交换,试证:如果 $\boldsymbol{A}$ 有特征向量,则 $\boldsymbol{A},\boldsymbol{B}$ 一定有公共特征向量.

第 10 章 欧氏空间

前两章介绍了线性空间和线性变换的概念.线性空间的许多例子说明,它应用的范围是很广泛的.但是在线性空间中,向量之间的运算只有加法与数量乘法这两种线性运算.与几何空间相比较,就会发现向量的度量性质,如向量的长度、向量间的夹角、距离等,在线性空间的理论中都没有得到反映.向量的度量性质在有些问题中是很重要的,因此有必要在线性空间中引入度量的概念.

在解析几何中我们看到:向量的长度、夹角等度量性质都可以通过向量的内积表示出来,而且向量的内积有比较明显的代数性质,容易计算.因此在抽象的讨论中,我们就取内积作为度量性质的基本概念.

在实线性空间中引入内积以后就成为欧氏空间.本章讨论欧氏空间的一些基本问题.因此,本章所讨论的问题都在实数域 **R** 上的线性空间中进行,所提到的数也都是实数.

10.1 欧氏空间的定义与简单性质

首先介绍欧氏空间的概念.

定义 1 设 V 是实数域 **R** 上的一个线性空间.在 V 上定义了一个二元实函数 (α,β),如果它具有以下性质:

(1) $(\alpha,\beta)=(\beta,\alpha)$;

(2) $(k\alpha,\beta)=k(\alpha,\beta)$;

(3) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$;

(4) $(\alpha,\alpha)\geqslant 0$;当且仅当 $\alpha=\mathbf{0}$ 时,$(\alpha,\alpha)=0$,

这里 α,β,γ 是 V 中任意向量，k 是任意实数. 就称(α,β)是线性空间 V 的一个**内积**；V 对内积(α,β)构成一个**欧几里得空间**，简称**欧氏空间**.

定义中所说的二元函数的意思是：对 V 中任意两个向量 α,β，都有唯一的一个实数与之对应，记作(α,β).

几何空间中向量的内积显然具有定义 1 中列举的 4 条性质，因此几何空间是一个欧氏空间.

下面再看几个例子.

例 1 在实线性空间 $\mathbf{R}^n$ 中，定义向量

$$\alpha = (a_1,a_2,\cdots,a_n), \quad \beta = (b_1,b_2,\cdots,b_n)$$

的内积为

$$(\alpha,\beta) := a_1b_1 + a_2b_2 + \cdots + a_nb_n. \tag{1}$$

这个内积就是在本书上册第 3 章中讨论正交化方法时定义的内积. 我们曾经证明过这样定义的内积满足定义 1 中的条件. 定义了内积以后，$\mathbf{R}^n$ 就成为一个欧氏空间. 我们仍用 $\mathbf{R}^n$ 来表示这个欧氏空间.

当 $n=3$ 时，(1)式就是几何空间向量的内积在直角坐标系中的坐标表达式.

例 2 设 V 是一个 n 维实线性空间. 在 V 中取定一组基. 设 A 是一个正定矩阵. 定义 V 的内积如下：设

$$\alpha = [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

及

$$\beta = [\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

是 V 中两个向量. 令

$$(\boldsymbol{\alpha},\boldsymbol{\beta}) = [x_1,x_2,\cdots,x_n]A\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}. \tag{2}$$

下面来证明这样定义的二元函数$(\boldsymbol{\alpha},\boldsymbol{\beta})$满足定义 1 中的条件.

(1) 因为 A 是正定矩阵,所以 A 是对称的. 于是

$$\begin{aligned}(\boldsymbol{\alpha},\boldsymbol{\beta}) &=[x_1,x_2,\cdots,x_n]A\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ &=[y_1,y_2,\cdots,y_n]A^{\mathrm{T}}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &=[y_1,y_2,\cdots,y_n]A\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &=(\boldsymbol{\beta},\boldsymbol{\alpha}).\end{aligned}$$

(2) $$(k\boldsymbol{\alpha},\boldsymbol{\beta}) =[kx_1,kx_2,\cdots,kx_n]A\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$=k\left([x_1,x_2,\cdots,x_n]A\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}\right)=k(\alpha,\beta).$$

(3) 再设 $\gamma=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]\begin{bmatrix}z_1\\z_2\\\vdots\\z_n\end{bmatrix}$,则因

$$\alpha+\beta=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]\begin{bmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{bmatrix},$$

故有

$$(\alpha+\beta,\gamma)=[x_1+y_1,x_2+y_2,\cdots,x_n+y_n]A\begin{bmatrix}z_1\\z_2\\\vdots\\z_n\end{bmatrix}$$

$$=[x_1,x_2,\cdots,x_n]A\begin{bmatrix}z_1\\z_2\\\vdots\\z_n\end{bmatrix}+[y_1,y_2,\cdots,y_n]A\begin{bmatrix}z_1\\z_2\\\vdots\\z_n\end{bmatrix}$$

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

(4) 因为 A 是正定矩阵,所以有

$$(\alpha,\alpha)=[x_1,x_2,\cdots,x_n]A\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\geqslant 0.$$

而且,当且仅当 $\alpha=\mathbf{0}$ 时 $(\alpha,\alpha)=0$.

因此，V 对内积(2) 构成一个欧氏空间.

例 2 说明每一个实线性空间都可以有许多方法定义内积. 需要注意的是，由不同内积构成的欧氏空间是不同的.

例 3 在实线性空间 $\mathbf{R}[x]$中，定义二元函数

$$(f(x),g(x)) = \int_{-1}^{+1} f(x)g(x)\mathrm{d}x.$$

易证，这个二元函数具有定义 1 中的 4 个性质. 于是 $\mathbf{R}[x]$对这个内积构成一个欧氏空间.

同样地，对 n 维实线性空间 $\mathbf{R}_n[x]$也可同样定义内积而成为一个欧氏空间.

定义 1 中性质(1)说明内积是对称的. 因此，与性质(2)及(3)相应地有

(2′) $(\alpha,k\beta)=\kappa(\alpha,\beta)$;

(3′) $(\alpha,\beta+\gamma)=(\alpha,\beta)+(\alpha+\gamma)$.

理由是：

$$(\alpha,k\beta) = (k\beta,\alpha) = k(\beta,\alpha) = k(\alpha,\beta);$$

$$\begin{aligned}(\alpha,\beta+\gamma) &= (\beta+\gamma,\alpha) = (\beta,\alpha)+(\gamma,\alpha)\\ &= (\alpha,\beta)+(\alpha,\gamma).\end{aligned}$$

由此可知，在欧氏空间 V 中，对任意向量 $\alpha_1,\alpha_2,\cdots,\alpha_s;\beta_1,\beta_2,\cdots,\beta_t$ 及任意实数 $k_1,k_2,\cdots,k_s;l_1,l_2,\cdots,l_t$，都有

$$\left(\sum_{i=1}^{s}k_i\alpha_i,\sum_{j=1}^{t}l_j\beta_j\right) = \sum_{i=1}^{s}\sum_{j=1}^{t}k_il_j(\alpha_i,\beta_j).$$

下面我们定义向量的长度及夹角.

首先，由性质(4)，对欧氏空间 V 中任意一向量 α，都有 $(\alpha,\alpha)\geqslant 0$，因此 $\sqrt{(\alpha,\alpha)}$ 是有意义的. 故可引进长度的概念.

定义 2 设 α 是欧氏空间中一个向量. 非负实数 $\sqrt{(\alpha,\alpha)}$ 称为向量 α 的**长度**，记作 $|\alpha|$.

显然，向量的长度一般都是正数，只有零向量的长度才等于零. 这样定义的长度与几何空间中向量的长度是一致的. 向量的长度也具有下述熟知的性质：

$$|k\alpha| = |k| \cdot |\alpha|, \quad k \in \mathbf{R}, \alpha \in V.$$

这个等式可以从定义直接得出：

$$|k\alpha| = \sqrt{(k\alpha, k\alpha)} = \sqrt{k^2(\alpha,\alpha)} = |k| \cdot |\alpha|.$$

与以前一样，我们把长度为 1 的向量称为**单位向量**. 如果 $\alpha \neq \mathbf{0}$，则向量 $\frac{1}{|\alpha|}\alpha$ 就是一个单位向量. 给了一个非零向量 α，用 α 的长度去除向量 α，得到一个与 α 同向的单位向量，称为把 α **单位化**.

在解析几何中，我们知道，向量 α,β 的夹角 $<\alpha,\beta>$ 的余弦可以通过内积来表示：

$$\cos < \alpha, \beta > = \frac{(\alpha,\beta)}{|\alpha||\beta|}. \tag{3}$$

由于

$$|\cos\theta| \leqslant 1,$$

因此为了在欧氏空间中应用(3)式引入夹角的概念，必须证明

$$\frac{|(\alpha,\beta)|}{|\alpha||\beta|} \leqslant 1.$$

为此，我们证明：

定理 1 (**柯西-布涅可夫斯基**(Cauchy-Буняковский)**不等式**)对于欧氏空间 V 中任意两个向量 α,β，都有

$$|\alpha,\beta| \leqslant |\alpha||\beta|.$$

当且仅当 α,β 线性相关时等号成立.

证明： 分 α,β 线性相关或线性无关两种情况加以证明.

如果 α,β 是线性相关的，则无妨假设

$$\beta = k\alpha, \quad k \in R.$$

于是

$$|(\alpha,\beta)| = |(\alpha,k\alpha)| = |k||(\alpha,\alpha)| = |\alpha||k\alpha| = |\alpha||\beta|.$$

如果 α,β 是线性无关的，那么对任意实数 k，

$$\alpha + k\beta \neq 0.$$

因此

$$(\alpha + k\beta,\alpha + k\beta) = (\beta,\beta)k^2 + 2(\alpha,\beta)k + (\alpha,\alpha) > 0.$$

这说明实系数方程

$$(\beta,\beta)x^2 + 2(\alpha,\beta)x + (\alpha,\alpha) = 0$$

无实数解. 因此

$$(\alpha,\beta)^2 - (\beta,\beta)(\alpha,\alpha) < 0,$$

即

$$(\alpha,\beta)^2 < (\beta,\beta)(\alpha,\alpha).$$

两边开方，即得

$$|(\alpha,\beta)| < |\alpha||\beta|. \mid$$

柯西-布涅可夫斯基不等式可用来证明一些不等式. 如应用这个不等式于例 1 中的内积，可得

$$|a_1b_1 + a_2b_2 + \cdots + a_nb_n|$$

$$\leqslant \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2},$$

式中的 $a_1,a_2,\cdots,a_n;b_1,b_2,\cdots,b_n$ 可以是任意实数.

对于例 2 中的内积，设

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

应用柯西-布涅可夫斯基不等式，即得

$$\left|\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_iy_j\right| \leqslant \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}y_iy_j},$$

式中的实数 a_{ij} 要求满足

$$a_{ij} = a_{ji}, \quad i,j = 1,2,\cdots,n;$$

并且 A 是正定矩阵.

例 3 中内积相应的不等式为

$$\left|\int_{-1}^{+1} f(x)g(x)\mathrm{d}x\right| \leqslant \left(\int_{-1}^{+1} f^2(x)\mathrm{d}x\right)^{1/2}\left(\int_{-1}^{+1} g^2(x)\mathrm{d}x\right)^{1/2}.$$

这些不等式都是很有用的不等式.

有了定理 1,就可以定义两个向量之间的夹角.

定义 3 欧氏空间 V 中两个非零向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 之间的**夹角** $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$ 规定为:

$$<\boldsymbol{\alpha},\boldsymbol{\beta}> = \cos^{-1}\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|}, \quad 0 \leqslant <\boldsymbol{\alpha},\boldsymbol{\beta}> \leqslant \pi.$$

定义 4 如果向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的内积为零,即

$$(\boldsymbol{\alpha},\boldsymbol{\beta}) = 0,$$

那么 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 称为**正交或垂直**,记作 $\boldsymbol{\alpha}\perp\boldsymbol{\beta}$.

根据定义有:

命题 1 (1) 零向量与任何向量都正交;

(2) 零向量是唯一的与它自己正交的向量;

(3) 如果 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 都是非零向量,那么,当且仅当 $<\boldsymbol{\alpha},\boldsymbol{\beta}> = \pi/2$ 时,$\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 正交.

命题 2 向量的长度有下述关系:

1) $|\boldsymbol{\alpha}+\boldsymbol{\beta}| \leqslant |\boldsymbol{\alpha}| + |\boldsymbol{\beta}|$.

2) 当且仅当 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 正交时,$|\boldsymbol{\alpha}+\boldsymbol{\beta}|^2 = |\boldsymbol{\alpha}|^2 + |\boldsymbol{\beta}|^2$.

证明: 因为

$$|\boldsymbol{\alpha}+\boldsymbol{\beta}|^2 = (\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta}) = (\boldsymbol{\alpha},\boldsymbol{\alpha}) + 2(\boldsymbol{\alpha},\boldsymbol{\beta}) + (\boldsymbol{\beta},\boldsymbol{\beta}),$$

所以

(1) 根据柯西-布涅可夫斯基不等式,有

$$\begin{aligned}|\boldsymbol{\alpha}+\boldsymbol{\beta}|^2 &= |\boldsymbol{\alpha}|^2 + |\boldsymbol{\beta}|^2 + 2(\boldsymbol{\alpha},\boldsymbol{\beta}) \\ &\leqslant |\boldsymbol{\alpha}|^2 + |\boldsymbol{\beta}|^2 + 2|\boldsymbol{\alpha}||\boldsymbol{\beta}|\end{aligned}$$

$$=(|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|)^2. \tag{4}$$

两边开方，即得

$$|\alpha+\beta| \leqslant |\alpha|+|\beta|.$$

(2) 由(4)式即可看到，等式

$$|\alpha+\beta|^2=|\alpha|^2+|\beta|^2$$

成立的充分必要条件是：$(\alpha,\beta)=0$，即 α 与 β 正交.

命题 2 中的第一个不等式就是通常的三角不等式；第二个等式说明在欧氏空间中，勾股定理同样成立. 这两个结论都可以推广到多个向量的情形，对于欧氏空间中任意 $s(s\geqslant 2)$ 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$，有

1) $|\alpha_1+\alpha_2+\cdots+\alpha_s|\leqslant|\alpha_1|+|\alpha_2|+\cdots+|\alpha_s|$；

2) 当 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 两两正交时，下述等式成立：

$$|\alpha_1+\alpha_2+\cdots+\alpha_s|^2=|\alpha_1|^2+|\alpha_2|^2+\cdots+|\alpha_s|^2.$$

例 4 在欧氏空间 $\mathbf{R}^4$ 中，给了向量

$$\alpha=(1,-2,0,2),\quad \beta=(2,0,0,2),$$

求 $<\alpha,\beta>$.

解： 因为在欧氏空间 $\mathbf{R}^4$ 中，内积定义为

$$((a_1,a_2,a_3,a_4),(b_1,b_2,b_3,b_4))=a_1b_1+a_2b_2+a_3b_3+a_4b_4,$$

所以

$$<\alpha,\beta>=\cos^{-1}\frac{(\alpha,\beta)}{|\alpha||\beta|}=\cos^{-1}\frac{6}{3\cdot 2\sqrt{2}}$$

$$=\cos^{-1}\frac{1}{\sqrt{2}}=\frac{\pi}{4}.$$

例 5 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 是欧氏空间 V 的一组基. 已知

$$(\varepsilon_1,\varepsilon_1)=(\varepsilon_2,\varepsilon_2)=(\varepsilon_3,\varepsilon_3)=1,$$

$$(\varepsilon_1,\varepsilon_2)=(\varepsilon_1,\varepsilon_3)=(\varepsilon_2,\varepsilon_3)=0.$$

再设

$$\alpha=3\varepsilon_1+2\varepsilon_2+4\varepsilon_3,\quad \beta=\varepsilon_1-2\varepsilon_2.$$

(1) 求与 α,β 都正交的全部向量；

(2) 求与 α,β 都正交的单位向量.

解：(1) 设

$$\gamma = x_1\varepsilon_1 + x_2\varepsilon_2 + x_3\varepsilon_3.$$

γ 与 α 及 β 都正交的充分必要条件为

$$\begin{aligned}(\gamma,\alpha) &= 3x_1 + 2x_2 + 4x_3 = 0,\\ (\gamma,\beta) &= x_1 - 2x_2 \qquad\quad = 0.\end{aligned}$$

解方程，得

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = k\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix},$$

所以与 α,β 都正交的全部向量为

$$k(2\varepsilon_1 + \varepsilon_2 - 2\varepsilon_3) \quad (k \text{ 为任意实数}).$$

(2) 把 $2\varepsilon_1+\varepsilon_2-2\varepsilon_3$ 单位化，得

$$\frac{2}{3}\varepsilon_1 + \frac{1}{3}\varepsilon_2 - \frac{2}{3}\varepsilon_3.$$

因此，与 α,β 都正交的向量共有两个，即

$$\pm\left(\frac{2}{3}\varepsilon_1 + \frac{1}{3}\varepsilon_2 - \frac{2}{3}\varepsilon_3\right).$$

最后，我们引进距离的概念.

定义 5 设 α,β 是欧氏空间 V 中的两个向量，α 与 β 的距离 $d(\alpha,\beta)$ 定义为

$$d(\alpha,\beta) = |\alpha - \beta|.$$

欧氏空间中向量间的距离也具有通常几何空间中向量间距离的性质.

命题 3 设 α,β,γ 是欧氏空间 V 中的 3 个向量，则

(1) $d(\alpha,\beta)=d(\beta,\alpha)$；

(2) $d(\alpha,\beta)\geqslant 0$，当且仅当 $\alpha=\beta$ 时等号成立；

(3) $d(\alpha,\gamma)\leqslant d(\alpha,\beta)+d(\beta,\gamma)$.

证明：(1)与(2)可以由定义直接得出.

(3) 根据三角不等式：

$$d(\alpha,\gamma)=|\alpha-\gamma|=|(\alpha-\beta)+(\beta-\gamma)|$$
$$\leqslant|\alpha-\beta|+|\beta-\gamma|=d(\alpha,\beta)+d(\beta,\gamma).$$

这样，我们通过内积定义了其他度量概念，使得欧氏空间中的向量有长度、夹角、距离等度量关系. 而且这些度量概念具有三维几何空间中向量的度量关系的一些重要性质.

习　题　10.1

1. 在 $\mathbf{R}^4$ 中，内积按通常定义：如果 $\alpha=(a_1,a_2,a_3,a_4)$，$\beta=(b_1,b_2,b_3,b_4)$，则

$$(\alpha,\beta)=a_1b_1+a_2b_2+a_3b_3+a_4b_4.$$

设

(1) $\alpha=(2,1,3,2)$，　　$\beta=(1,2,-2,1)$；

(2) $\alpha=(1,2,2,-3)$，　　$\beta=(1,3,-2,-1)$；

(3) $\alpha=(1,1,-1,-1)$，　$\beta=(1,1,0,-1)$.

求(α,β)和 α,β 之间的夹角$\langle\alpha,\beta\rangle$.

2. 设 α,β 是欧氏空间 V 中两个向量，k 是一个正实数，试证：

(1) $\langle\alpha,k\beta\rangle=\langle\alpha,\beta\rangle$.

(2) $\langle\alpha,\beta\rangle+\langle\alpha,-\beta\rangle=\pi$.

3. 试证：如果 α 与 $\beta_1,\beta_2,\cdots,\beta_t$ 都正交，则 α 与 $\beta_1,\beta_2,\cdots,\beta_t$ 的任一线性组合都正交.

4. 在欧氏空间 $\mathbf{R}^4$ 中，求一个单位向量与

$$(1,1,0,0),\quad(1,1,-1,-1),\quad(1,-1,1,-1)$$

都正交.

5. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是欧氏空间 V 的一组基，试证：

(1) 如果 $\gamma\in V$，使$(\gamma,\alpha_i)=0$，$i=1,2,\cdots,n$，那么 $\gamma=\mathbf{0}$.

(2) 如果 $\gamma_1,\gamma_2\in V$，使$(\gamma_1,\alpha_i)=(\gamma_2,\alpha_i)$，$i=1,2,\cdots,n$，那么 $\gamma_1=\gamma_2$.

10.2 度量矩阵

这一节讨论的欧氏空间都是有限维的.

设 V 是一个 n 维线性空间,在 V 中取定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. 对 V 中两个向量

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n,$$

$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n.$$

由内积的性质,有

$$(\alpha,\beta) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j(\varepsilon_i,\varepsilon_j). \tag{1}$$

这说明 α,β 的内积可以通过它们的坐标及基向量之间的内积表示出来. 由此还可知,只要基向量之间的内积知道了,那么 V 中任意两个向量的内积也就唯一确定了.

为了便于计算和应用,下面引进度量矩阵的概念.

设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是欧氏空间 V 的一组基. 令

$$a_{ij} = (\varepsilon_i,\varepsilon_j)$$

作一个 n 阶矩阵

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

如果 α,β 对基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的坐标分别是

$$X = (x_1,x_2,\cdots,x_n)$$

及

$$Y = (y_1,y_2,\cdots,y_n),$$

则(α,β)可表成

$$(\alpha,\beta)=(x_1,x_2,\cdots,x_n)A\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}$$

$$=XAY^{\mathrm{T}}.$$

定义 6 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是欧氏空间 V 的一组基. 矩阵

$$\begin{bmatrix}(\varepsilon_1,\varepsilon_1) & (\varepsilon_1,\varepsilon_2) & \cdots & (\varepsilon_1,\varepsilon_n)\\(\varepsilon_2,\varepsilon_1) & (\varepsilon_2,\varepsilon_2) & \cdots & (\varepsilon_2,\varepsilon_n)\\ \cdots & \cdots & \cdots & \cdots\\(\varepsilon_n,\varepsilon_1) & (\varepsilon_n,\varepsilon_2) & \cdots & (\varepsilon_n,\varepsilon_n)\end{bmatrix}$$

称为基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的**度量矩阵**.

上面的讨论表明:知道了一组基的度量矩阵以后,任意两个向量的内积就可以通过坐标,根据(1)式或(2)式来计算,因而度量矩阵完全确定了内积.

例 1 欧氏空间 $\mathbf{R}^n$ 中,n 个基本向量

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

组成一组基. 因为

$$(\varepsilon_i,\varepsilon_j)=\delta_{ij},\quad i,j=1,2,\cdots,n,$$

所以这组基的度量矩阵是单位矩阵.

例 2 V 是 10.1 节例 2 中的欧氏空间. 由于

$$
(\boldsymbol{\varepsilon}_i,\boldsymbol{\varepsilon}_j)=[0,\cdots,0,\underset{\substack{\text{第}\\ i\\ \text{个}}}{1},0,\cdots,0]A\begin{array}{cl}\begin{bmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{bmatrix} & \begin{array}{l}\\ \\ \\ \text{第 } j \text{ 个}\\ \\ \\ \\ \end{array}\end{array}
$$

$$
=a_{ij}\qquad (i,j=1,2,\cdots,n),
$$

所以基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 的度量矩阵就是 A.

例 3 在欧氏空间 $\mathbf{R}^3$ 中求下列基的度量矩阵：

$\boldsymbol{\alpha}_1=(1,1,-1)$， $\boldsymbol{\alpha}_2=(1,0,2)$， $\boldsymbol{\alpha}_3=(1,-1,-2)$.

解： 因为

$$
(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)=3,\ (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2)=5,\ (\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3)=6,
$$

$$
(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)=(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1)=-1,
$$

$$
(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3)=(\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_1)=2,
$$

$$
(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=-3,
$$

所以这组基的度量矩阵为：

$$
\begin{bmatrix}3 & -1 & 2\\ -1 & 5 & -3\\ 2 & -3 & 6\end{bmatrix}
$$

例 4 设欧氏空间 $\mathbf{R}_3[x]$的内积为

$$
(f(x),g(x))=\int_{-1}^{+1}f(x)g(x)\mathrm{d}(x),
$$

求基

$$
\boldsymbol{\varepsilon}_1=1,\ \boldsymbol{\varepsilon}_2=x,\ \boldsymbol{\varepsilon}_3=x^2
$$

的度量矩阵.

解： 因为

$$(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_1)=\int_{-1}^{+1}\mathrm{d}x=2;$$

$$(\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_2)=\int_{-1}^{+1}x^2\mathrm{d}x=\frac{2}{3};$$

$$(\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_3)=\int_{-1}^{+1}x^4\mathrm{d}x=\frac{2}{5};$$

$$(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2)=(\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_1)=\int_{-1}^{+1}x\mathrm{d}x=0;$$

$$(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_3)=(\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_1)=\int_{-1}^{+1}x^2\mathrm{d}x=\frac{2}{3};$$

$$(\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3)=(\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_2)=\int_{-1}^{+1}x^3\mathrm{d}x=0.$$

所以这组基的度量矩阵为

$$\begin{bmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{bmatrix}.$$

度量矩阵有下述重要性质：

定理 2 度量矩阵是正定矩阵.

证明： 设 V 是一个 n 维欧氏空间；$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 是 V 的一组基；A 是这组基的度量矩阵. 由内积的性质(1)知

$$(\boldsymbol{\varepsilon}_i,\boldsymbol{\varepsilon}_j)=(\boldsymbol{\varepsilon}_j,\boldsymbol{\varepsilon}_i),\quad i,j=1,2,\cdots,n,$$

所以 A 是一个实对称矩阵. 而由性质(4)知，对于非零向量 $\boldsymbol{\alpha}$，有

$$(\boldsymbol{\alpha},\boldsymbol{\alpha})>0.$$

即当

$$X=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\neq 0$$

时，

$$X^{\mathrm{T}}AX>0,$$

所以 A 是正定的. |

下面讨论，不同基的度量矩阵之间有什么关系. 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$；$\eta_1,\eta_2,\cdots,\eta_n$ 是 n 维欧氏空间 V 的两组基，它们的度量矩阵分别是

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

及

$$B=\begin{bmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\b_{21}&b_{22}&\cdots&b_{2n}\\\cdots&\cdots&\cdots&\cdots\\b_{n1}&b_{n2}&\cdots&b_{nn}\end{bmatrix}.$$

再设由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到 $\eta_1,\eta_2,\cdots,\eta_n$ 的过渡矩阵是

$$C=\begin{bmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\cdots&\cdots&\cdots&\cdots\\c_{n1}&c_{n2}&\cdots&c_{nn}\end{bmatrix},$$

那么

$$b_{ij}=(\eta_i,\eta_j)=\left(\sum_{k=1}^{n}c_{ki}\varepsilon_k,\sum_{l=1}^{n}c_{lj}\varepsilon_l\right)$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} c_{ki} c_{lj} (\varepsilon_k, \varepsilon_l)$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} c_{ki} a_{kl} c_{lj} \quad (i,j = 1,2,\cdots,n).$$

上式右端恰好是 $C^{\mathrm{T}}AC$ 的第 i 行第 j 列的元素，所以

$$B = C^{\mathrm{T}}AC.$$

因此我们有下列定理：

定理 3 不同基的度量矩阵是合同的. 确切地说，如果基 ε_1, $\varepsilon_2,\cdots,\varepsilon_n$ 与 $\eta_1,\eta_2,\cdots,\eta_n$ 的度量矩阵分别是 A 及 B，由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到 $\eta_1,\eta_2,\cdots,\eta_n$ 的过渡矩阵是 C，那么

$$B = C^{\mathrm{T}}AC.$$

例如，在例 3 中，欧氏空间 $\mathbf{R}^3$ 的基

$$\varepsilon_1 = (1,0,0), \quad \varepsilon_2 = (0,1,0), \quad \varepsilon_3 = (0,0,1)$$

的矩阵为单位矩阵. 而由 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 到 $\alpha_1,\alpha_2,\alpha_3$ 的过渡矩阵为

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

因此，根据定理 3，$\alpha_1,\alpha_2,\alpha_3$ 的度量矩阵应为

$$A^{\mathrm{T}}EA = A^{\mathrm{T}}A = \begin{bmatrix} 3 & -1 & 2 \\ -1 & 5 & -3 \\ 2 & -3 & 6 \end{bmatrix}.$$

这与例 3 中直接计算的结果是一样的.

由于 n 阶正定矩阵都是合同的，所以还有：

命题 1 V 是一个 n 维欧氏空间，则任一正定矩阵 D 都可看成 V 的某一组基的度量矩阵.

证明： 任取 V 的一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，设它的度量矩阵为 A. 由于 A 与 D 都是 n 阶正定矩阵，所以是合同的，即可找到一个可逆矩阵 C，使得

$$D=C^{\mathrm{T}}AC.$$

令

$$[\gamma_1,\gamma_2,\cdots,\gamma_n]=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]C,$$

则 $\gamma_1,\gamma_2,\cdots,\gamma_n$ 也是 V 的基，且这组基的矩阵就是 $C^{\mathrm{T}}AC$，即 D.

习 题 10.2

1. 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是欧氏空间 V 的一组基，试证：如果 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 两两正交，那么这组基的度量矩阵是对角矩阵.

2. 在欧氏空间 $\mathbf{R}^4$ 中，求下述基的度量矩阵：

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

   (2) $\beta_1=\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\beta_2=\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right),$

   $\beta_3=\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right),\beta_4=\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right).$

3. 设 V 是一个欧氏空间 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是 V 的一组基. 已知基

$$\alpha_1=\varepsilon_1-\varepsilon_2,\alpha_2=-\varepsilon_1+2\varepsilon_2,\alpha_3=\varepsilon_2-2\varepsilon_3+\varepsilon_4,\alpha_4=\varepsilon_1+\varepsilon_3+\varepsilon_4$$

的度量矩阵为

$$\begin{bmatrix} 2 & -3 & 0 & 1 \\ -3 & 6 & 0 & -1 \\ 0 & 0 & 13 & 9 \\ 1 & -1 & 9 & 7 \end{bmatrix}.$$

   (1) 求基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 的度量矩阵；

   (2) 求 a 使

$$\alpha=\varepsilon_1-\varepsilon_2+2\varepsilon_3$$

与

$$\beta=\varepsilon_1+a\varepsilon_2+2\varepsilon_3-\varepsilon_4$$

正交；

   (3) 求一个单位向量 ξ_4 与 $\xi_1=\varepsilon_1-\varepsilon_2-\varepsilon_3+\varepsilon_4,\xi_2=\varepsilon_1-\varepsilon_2-\varepsilon_3+\varepsilon_4,\xi_3=2\varepsilon_1+\varepsilon_2+\varepsilon_3+3\varepsilon_4$ 都正交；

   (4) 求 ξ_1,ξ_2,ξ_3,ξ_4 的度量矩阵.

4. 求欧氏空间 $\mathbf{R}^4$ 的一组基以

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 2 & 1 & -1 \\ 0 & 1 & 2 & -1 \\ 1 & -1 & -1 & 10 \end{bmatrix}$$

为度量矩阵.

10.3 标准正交基

欧氏空间与线性空间的主要差别是在欧氏空间中有度量性质，而度量性质的依据是在欧氏空间中定义了内积. 内积可以通过度量矩阵来表示. 因此，如何选择基，使得度量矩阵最简单，是一个重要的问题. 我们在上节中曾经证明了度量矩阵是正定矩阵，而且两组基的度量矩阵是合同的. 因为正定矩阵与单位矩阵是合同的，所以一定可以找到一组基，使这组基的度量矩阵是单位矩阵 E. 什么样的基才能以 E 为度量矩阵呢？如果基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的度量矩阵是 E，那么根据定义，立即可得

$$(\varepsilon_i,\varepsilon_j)=\begin{cases}1, & 当\ i=j;\\ 0, & 当\ i\neq j.\end{cases}$$

这说明这组基中的向量是两两正交的，并且都是单位向量. 据此，我们引入下列定义.

定义 7 欧氏空间 V 中由两两正交的非零向量组成的向量组称为一个**正交向量组**.

根据定义，由一个非零向量组成的向量组也是正交向量组.

可以与本书上册第 3 章定理 6 一样地证明：正交向量组一定是线性无关的. 这里就不重复了.

因此，在 n 维欧氏空间中，最多有 n 个两两正交的非零向量. 而且，n 维欧氏空间中任意 n 个正交的非零向量组成一组基.

定义 8 在 n 维欧氏空间中，由 n 个正交向量组成的基称为

正交基. 由单位向量组成的正交基称为**标准正交基**.

由前面的分析可知:

定理 4 n 维欧氏空间的一组基是标准正交基的充分必要条件是:它的度量矩阵是单位矩阵.

前面的分析也说明了标准正交基的存在及求法. 下面举例说明.

例 1 三维实线性空间 V 中,取定一组基 $\alpha_1,\alpha_2,\alpha_3$. 定义内积为

$$(\alpha,\beta)=X^{\mathrm{T}}\begin{bmatrix}1&-1&1\\-1&2&0\\1&0&4\end{bmatrix}Y,$$

其中 X,Y 分别是 α,β 对于基 $\alpha_1,\alpha_2,\alpha_3$ 的坐标. 从而 V 成为一个欧氏空间. 求欧氏空间 V 的一组标准正交基.

解: 令

$$A=\begin{bmatrix}1&-1&1\\-1&2&0\\1&0&4\end{bmatrix},$$

则 A 是 $\alpha_1,\alpha_2,\alpha_3$ 的度量矩阵,是一个正定矩阵,因此可以找到一个可逆矩阵 T,使得

$$T^{\mathrm{T}}AT=E.$$

令

$$[\eta_1,\eta_2,\eta_3]=[\alpha_1,\alpha_2,\alpha_3]T,$$

则 η_1,η_2,η_3 也是 V 的一组基. 这组基的度量矩阵是 $T^{\mathrm{T}}AT$,即单位矩阵,所以这是一组标准正交基. 下面应用初等变换法来求 T:

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

$$
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & -\sqrt{2} \\ 0 & 1 & -\frac{\sqrt{2}}{2} \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix},
$$

因此，令

$$
T = \begin{bmatrix} 1 & 1 & -\sqrt{2} \\ 0 & 1 & -\frac{\sqrt{2}}{2} \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix},
$$

即有

$$
T^{\mathrm{T}} A T = E.
$$

所以得到一组标准正交基

$$
\eta_1 = \alpha_1,
$$

$$
\eta_2 = \alpha_1 + \alpha_2,
$$

$$
\eta_3 = -\sqrt{2}\alpha_1 - \frac{\sqrt{2}}{2}\alpha_2 + \frac{\sqrt{2}}{2}\alpha_3.
$$

也可以用施密特正交化过程来求标准正交基. 我们通过下面

的例题来说明这一方法.

例 2 设 V 是一个四维欧氏空间，$\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 是 V 的一组基. 这组基的度量矩阵为

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

求 V 的一组标准正交基.

解： 先将 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 用施密特正交化过程化为等价的正交向量组：

$$\beta_1=\alpha_1,$$

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1=-\alpha_1+\alpha_2,$$

$$\beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2=-\alpha_1+\alpha_2+\alpha_3,$$

$$\begin{aligned}\beta_4&=\alpha_4-\frac{(\alpha_4,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_4,\beta_2)}{(\beta_2,\beta_2)}\beta_2-\frac{(\alpha_4,\beta_3)}{(\beta_3,\beta_3)}\beta_3\\&=\alpha_1-2\alpha_2-\alpha_3+\alpha_4.\end{aligned}$$

$\beta_1,\beta_2,\beta_3,\beta_4$ 是 V 的一组正交基，再将它们单位化，即得到 V 的一组标准正交基：

$$\varepsilon_1=\frac{\beta_1}{|\beta_1|}=\alpha_1,$$

$$\varepsilon_2=\frac{\beta_2}{|\beta_2|}=-\alpha_1+\alpha_2,$$

$$\varepsilon_3=\frac{\beta_3}{|\beta_3|}=-\frac{\sqrt{2}}{2}\alpha_1+\frac{\sqrt{2}}{2}\alpha_2+\frac{\sqrt{2}}{2}\alpha_3,$$

$$\varepsilon_4=\frac{\beta_4}{|\beta_4|}=\alpha_1-2\alpha_2-\alpha_3+\alpha_4.$$

应用施密特正交化过程还证明了下述定理.

定理 5 对于 n 维欧氏空间 V 中任意一组线性无关的向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$,都可找到一组正交的单位向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s$,使

$$L(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_i)=L(\alpha_1,\alpha_2,\cdots,\alpha_i),\ i=1,2,\cdots,s.$$

特别地,当 $s=n$ 时,$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基.

例如,例 2 中的标准正交基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 就满足

$$L(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_i)=L(\alpha_1,\alpha_2,\cdots,\alpha_i)\ i=1,2,3,4.$$

命题 1 n 维欧氏空间 V 中任一个正交向量组都可扩充成 V 的一组正交基;任一个正交单位向量组都可扩充成 V 的一组标准正交基.

例 3 在欧氏空间 $\mathbf{R}^4$ 中,给了两个正交的单位向量

$$\varepsilon_1=\left(\frac{1}{3},\frac{2}{3},\frac{2}{3},0\right)$$

$$\varepsilon_2=\left(0,\frac{1}{\sqrt{3}},\frac{-1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)$$

将 $\varepsilon_1,\varepsilon_2$ 扩充成 $\mathbf{R}^4$ 的一组标准正交基.

解: 首先找出两个向量

$$\alpha_1=(1,0,0,0),$$
$$\alpha_2=(0,1,0,0).$$

使 $\varepsilon_1,\varepsilon_2,\alpha_1,\alpha_2$ 线性无关,将 $\varepsilon_1,\varepsilon_2,\alpha_1,\alpha_2$ 正交化,得

$$\beta_1=\varepsilon_1$$

$$\beta_2=\varepsilon_2$$

$$\beta_3=\alpha_1-\frac{(\alpha_1,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_1,\beta_2)}{(\beta_2,\beta_2)}\beta_2$$

$$=\left(\frac{8}{9},-\frac{2}{9},-\frac{2}{9},0\right)$$

$$\beta_4=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_2,\beta_2)}{(\beta_2,\beta_2)}\beta_2-\frac{(\alpha_2,\beta_3)}{(\beta_3,\beta_3)}\beta_3$$

$$=\left(0,\frac{1}{6},-\frac{1}{6},-\frac{1}{3}\right)$$

再将 $\beta_1,\beta_2,\beta_3,\beta_4$ 单位化：

$$\varepsilon_1=\frac{1}{|\beta_1|}\beta_1$$

$$\varepsilon_2=\frac{1}{|\beta_2|}\beta_2$$

$$\varepsilon_3=\frac{1}{|\beta_3|}\beta_3=\left(\frac{2}{3}\sqrt{2},-\frac{1}{6}\sqrt{2},-\frac{1}{6}\sqrt{2},0\right)$$

$$\varepsilon_4=\frac{1}{|\beta_4|}\beta_4=\left(0,\frac{1}{6}\sqrt{6},-\frac{1}{6}\sqrt{6},-\frac{1}{3}\sqrt{6}\right)$$

则 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 就是 $\varepsilon_1,\varepsilon_2$ 扩充成的一组标准正交基.

包含 $\varepsilon_1,\varepsilon_2$ 的标准正交基是非常多的，还可以先求出与 $\varepsilon_1,\varepsilon_2$ 正交的两个线性无关的向量 α_3,α_4，然后再将 α_3,α_4 正交化，单位化，求出一组 $\varepsilon_1,\varepsilon_2$ 扩充成的标准正交基. 这种算法也很方便. 读者不妨自己计算一下.

在标准正交基下，向量的坐标可以通过内积简单地表示出来.

命题 2 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是欧氏空间 V 的一组标准正交基，则 V 中向量 α 对于这组基的坐标为

$$((\alpha,\varepsilon_1),(\alpha,\varepsilon_2),\cdots,(\alpha,\varepsilon_n)).$$

即，α 可表示成

$$\alpha=(\alpha,\varepsilon_1)\varepsilon_1+(\alpha,\varepsilon_2)\varepsilon_2+\cdots+(\alpha,\varepsilon_n)\varepsilon_n.$$

证明： 设

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n.$$

用 ε_i 与上式两边做内积，即得

$$(\alpha,\varepsilon_i)=x_i,\quad i=1,2,\cdots,n.$$

在标准正交基下，内积的表达式也特别简单.

命题 3 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是欧氏空间 V 的一组标准正交基. 又

设 α,β 是 V 中两个向量. 如果它们对于 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的坐标分别为

$$(x_1,x_2,\cdots,x_n)\ 及\ (y_1,y_2,\cdots,y_n),$$

即

$$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n,$$
$$\beta=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n,$$

那么

$$(\alpha,\beta)=x_1y_1+x_2y_2+\cdots+x_ny_n.$$

下面讨论从一组标准正交基到另一组标准正交基的过渡矩阵.

定理 6 (1) 由标准正交基到标准正交基的过渡矩阵是正交矩阵;

(2) 如果两组基间的过渡矩阵是正交矩阵,那么从其中一组基是标准正交基可推出另一组基也是标准正交基.

证明: 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 及 $\eta_1,\eta_2,\cdots,\eta_n$ 是欧氏空间 V 的两组标准正交基,并设由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到 $\eta_1,\eta_2,\cdots,\eta_n$ 的过渡矩阵是 A.

(1) 如果 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 及 $\eta_1,\eta_2,\cdots,\eta_n$ 都是标准正交基,那么它们的度量矩阵都是单位矩阵 E. 又因由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到 $\eta_1,\eta_2,\cdots,\eta_n$ 的过渡矩阵是 A,故由定理 3,得

$$E=A^{\mathrm{T}}EA=A^{\mathrm{T}}A,$$

所以 A 是一个正交矩阵.

(2) 设 A 是正交矩阵. 如果 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是一组标准正交基,那么它的度量矩阵是单位矩阵 E,于是 $\eta_1,\eta_2,\cdots,\eta_n$ 的度量矩阵为

$$A^{\mathrm{T}}EA=A^{\mathrm{T}}A=E,$$

所以 $\eta_1,\eta_2,\cdots,\eta_n$ 也是标准正交基.

由于由 $\eta_1,\eta_2,\cdots,\eta_n$ 到 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的过渡矩阵是 A^{-1},也是一个正交矩阵,所以当 $\eta_1,\eta_2,\cdots,\eta_n$ 是标准正交基时,$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 一定也是标准正交基.

标准正交基之间的这些关系,告诉我们从一组标准正交基得出另一组标准正交基的方法,这在处理一些问题的时候是很有用

的.

习 题 10.3

1. 证明：

$$\alpha_1=\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\quad \alpha_2=\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right),$$

$$\alpha_3=\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right),\quad \alpha_4=\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right)$$

是欧氏空间 $\mathbf{R}^4$ 的一组标准正交基.

2. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 是 3 维欧氏空间 V 的一组标准正交基，试证：

$$\alpha_1=\frac{2}{3}\varepsilon_1+\frac{2}{3}\varepsilon_2+\frac{1}{3}\varepsilon_3,$$

$$\alpha_2=\frac{2}{3}\varepsilon_1-\frac{1}{3}\varepsilon_2-\frac{2}{3}\varepsilon_3,$$

$$\alpha_3=\frac{1}{3}\varepsilon_1-\frac{2}{3}\varepsilon_2+\frac{2}{3}\varepsilon_3$$

也是 V 的一组标准正交基.

3. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 是欧氏空间 V 的一组基，这组基的度量矩阵是

$$\begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & 6 \end{bmatrix}.$$

再设

$$\alpha_1=\varepsilon_1+\varepsilon_2,$$

(1) 证明 α_1 是一个单位向量；

(2) 求 a 使 α_1 与

$$\beta_2=\varepsilon_1+\varepsilon_2+a\varepsilon_3$$

正交；

(3) 把所求出的 β_2 单位化，记作 α_2；

(4) 把 α_1,α_2 扩充成 V 的一组标准正交基.

4. 设

$$\alpha_1=(1,1,1,1),\quad \alpha_2=(3,3,1,1),$$

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

求欧氏空间 $\mathbf{R}^4$ 的一组标准正交基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$，使

$$L(\varepsilon_1,\cdots,\varepsilon_i)=L(\alpha_1,\cdots,\alpha_i),\ i=1,2,3,4.$$

10.4 子 空 间

欧氏空间的子空间对于原空间的内积显然也是一个欧氏空间. 作为线性空间，第 8 章中讨论的关于子空间的所有结论，对于欧氏空间的子空间也都是成立的. 现在讨论欧氏空间中子空间的正交关系.

定义 9 设 W 是欧氏空间 V 的一个子空间，α 是 V 中一个向量. 如果对于 W 中任一向量 β 都有

$$(\alpha,\beta)=0,$$

则称 α 与子空间 W **正交**（或称 α **垂直**于子空间 W），记作 $\alpha\perp W$.

定义 10 设 W_1,W_2 是欧氏空间 V 的两个子空间，如果对于 W_1 中任一向量 α 及 W_2 中任一向量 β，都有

$$(\alpha,\beta)=0,$$

则称 W_1 与 W_2 是**正交**的，记作 $W_1\perp W_2$.

根据定义，可知当且仅当 W_1 中每个向量都与 W_2 正交时，$W_1\perp W_2$. 并且，如果 $W_1\perp W_2$，那么 $W_2\perp W_1$.

因为只有零向量与它自己正交，因此有：

命题 1 (1) 由 $\alpha\in W,\alpha\perp W$ 可推出 $\alpha=\mathbf{0}$；

(2) 由 $W_1\perp W_2$ 可推出 $W_1\cap W_2=\{\mathbf{0}\}$.

例 1 如果 α 与 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 都正交，则可推出 α 与 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一切线性组合都正交，因此

$$\alpha\perp L(\alpha_1,\alpha_2,\cdots,\alpha_s).$$

如果向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 中每个向量都与向量组 $\beta_1,\beta_2,\cdots,\beta_t$ 中每个向量正交，则

$$L(\alpha_1,\alpha_2,\cdots,\alpha_s)\perp L(\beta_1,\beta_2,\cdots,\beta_t).$$

例 2 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s$ 是欧氏空间中一组正交的向量，那么有

$$\varepsilon_i\perp L(\varepsilon_1,\cdots,\varepsilon_{i-1},\varepsilon_{i+1},\cdots,\varepsilon_s),\quad i=1,2,\cdots,s;$$

$$L(\varepsilon_1,\cdots,\varepsilon_i)\perp L(\varepsilon_{i+1},\cdots,\varepsilon_s),\quad i=1,2,\cdots,s-1.$$

因为正交向量组是线性无关的，所以有：

定理 7 如果欧氏空间 V 的子空间 $W_1,W_2,\cdots,W_s(s\geqslant2)$ 两两相交，那么 $W_1+W_2+\cdots+W_s$ 是直和.

证明： 在 $W_i(i=1,2,\cdots,s)$ 中取一组正交基

$$\alpha_{i1},\alpha_{i2},\cdots,\alpha_{in_i}.$$

则由假设，

$$\alpha_{11},\cdots,\alpha_{1i_1},\alpha_{21},\cdots,\alpha_{2i_2},\cdots,\alpha_{s1},\cdots,\alpha_{si_s}$$

也是一个正交向量组，因此是线性无关的，组成 $W_1+W_2+\cdots+W_s$ 的一组基. 这说明

$$\dim(W_1+W_2+\cdots+W_s)=\sum_{i=1}^{s}\dim(W_i),$$

所以 $W_1+W_2+\cdots+W_s$ 是直和. |

定义 11 设 W_1 与 W_2 是欧氏空间 V 的两个子空间. 如果 $W_1\perp W_2$，并且 $W_1+W_2=V$，则称 W_2 为 W_1 的**正交补**.

显然，如果 W_2 是 W_1 的正交补，那么 W_1 是 W_2 的正交补. 根据定理 7，如果 W_1 是 W_2 的正交补，那么有

$$W_1\dot{+}W_2=V.$$

例 3 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是欧氏空间 V 的一组正交基，那么，$L(\varepsilon_1,\cdots,\varepsilon_i)$ 是 $L(\varepsilon_{i+1},\cdots,\varepsilon_n)$ 的正交补 $(i=1,2,\cdots,n-1)$.

例 3 说明了子空间的正交补的存在与求法. 一般地，有下述定理：

定理 8 n 维欧氏空间的每一个子空间 W，都有唯一的正交补.

证明： 先证明正交补的存在. 如果 $W_1=\{\mathbf{0}\}$，则 V 是 W 的正交补. 如果 $W=V$，则 $\{\mathbf{0}\}$ 是 W 的正交补. 因此设 $W\neq\{\mathbf{0}\}$，$W\neq V$. 取 W 的一组正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m(1\leqslant m<n)$. 把它扩充成 V 的一组正交基

$$\varepsilon_1,\cdots,\varepsilon_m,\varepsilon_{m+1},\cdots,\varepsilon_n,$$

那么子空间 $L(\varepsilon_{m+1},\cdots,\varepsilon_n)$ 就是 W 的一个正交补.

下面证明唯一性. 如果 W_1,W_2 都是 W 的正交补，那么

$$V=W\dotplus W_1,\quad V=W\dotplus W_2.$$

任取 W_1 中一个向量 α_1，由第二个等式，α_1 可表示成

$$\alpha_1=\alpha+\alpha_2,\quad \alpha\in W,\quad \alpha_2\in W_2.$$

因为 $W_1\perp W$，$W_2\perp W$，所以 $\alpha\perp\alpha_1$，$\alpha\perp\alpha_2$. 于是

$$(\alpha,\alpha)=(\alpha,\alpha_1-\alpha_2)=(\alpha,\alpha_1)-(\alpha,\alpha_2)=0.$$

由此得

$$\alpha=\mathbf{0}.$$

因此

$$\alpha_1=\alpha_2\in W_2.$$

这说明 $W_1\subseteq W_2$；同样可证明 $W_2\subseteq W_1$. 所以 $W_1=W_2$. 唯一性得证. ┃

我们以后用 $W^{\perp}$ 表示 W 的正交补. 易知下述命题成立.

命题 2 设 W 是有限维欧氏空间 V 的一个子空间，则

$$\dim W+\dim W^{\perp}=\dim V.$$

命题 3 设 W 是 n 维欧氏空间 V 的一个子空间，则 $W^{\perp}$ 恰好由 V 中与 W 正交的全部向量组成.

证明： 当 $W=\{\mathbf{0}\}$ 或 $W=V$ 时，结论显然成立. 设 $W\neq\{\mathbf{0}\}$，$W\neq V$. 取 W 的一组正交基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m\quad(1\leqslant m<n),$$

将其扩充成 V 的一组正交基

$$\varepsilon_1,\cdots,\varepsilon_m,\varepsilon_{m+1},\cdots,\varepsilon_n,$$

则

$$W^{\perp}=L(\varepsilon_{m+1},\cdots,\varepsilon_n).$$

现在我们来证明 V 中任一个与 W 正交的向量一定属于 $W^{\perp}$. 设 $\alpha\perp W$. 把 α 表示成 $\varepsilon_1,\cdots,\varepsilon_m,\varepsilon_{m+1},\cdots,\varepsilon_n$ 的线性组合:

$$\alpha=a_1\varepsilon_1+\cdots+a_m\varepsilon_m+a_{m+1}\varepsilon_{m+1}+\cdots+a_n\varepsilon_n.$$

依次用 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m$ 与上式两边作内积,即得

$$a_i(\varepsilon_i,\varepsilon_i)=0,\quad i=1,2,\cdots,m.$$

因为 $(\varepsilon_i,\varepsilon_i)\neq 0$,所以

$$a_1=a_2=\cdots=a_m=0,$$

$$\alpha=a_{m+1}\varepsilon_{m+1}+\cdots+a_n\varepsilon_n\in W^{\perp}.$$

而且,$W^{\perp}$ 中任一向量都与 W 正交,因此 $W^{\perp}$ 是由 V 中与 W 正交的全部向量组成的.

推论 设 W 是欧氏空间 V 的一个非零子空间,$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m$ 是 W 的一组基,则 $W^{\perp}$ 由 V 中与 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m$ 都正交的全部向量组成.

定义 12 设 W 是欧氏空间 V 的一个子空间. 根据

$$V=W\dotplus W^{\perp},$$

V 中每个向量 α 都可以唯一地分解成

$$\alpha=\beta+\gamma,$$

其中,$\beta\in W,\gamma\in W^{\perp}$. β 叫做向量 α 在子空间 W 上的**内射影**.

例 4 W 是欧氏空间 $\mathbf{R}^4$ 的一个子空间. 已知

$$W=L(\alpha_1,\alpha_2),$$

其中

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

求 $W^{\perp}$;并求向量

$$\alpha=(1,-3,2,2)$$

在 W 上的内射影.

解： 根据命题 3 的推论，$W^{\perp}$ 由 $\mathbf{R}^4$ 中与 α_1,α_2 正交的全部向量组成. 向量 (x_1,x_2,x_3,x_4) 与 α_1,α_2 正交的充分必要条件为：

$$\begin{cases} x_1 + x_2 + 2x_3 + x_4 = 0, \\ x_1 \qquad\qquad\qquad - 2x_4 = 0. \end{cases}$$

取这个齐次方程组的一个基础解系：

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

则 $\mathbf{R}^4$ 中任一个与 α_1,α_2 正交的向量都是 α_3,α_4，的线性组合. α_3,α_4 是 $W^{\perp}$的一组基，即

$$W^{\perp} = L(\alpha_3, \alpha_4).$$

为了求 α 在 W 上的内射影，把 α 表示成 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的线性组合：

$$\begin{aligned} \alpha &= \frac{1}{2}\alpha_1 - \frac{1}{2}\alpha_2 + \alpha_3 + \frac{1}{2}\alpha_4 \\ &= \left(\frac{1}{2}\alpha_1 - \frac{1}{2}\alpha_2\right) + \left(\alpha_3 + \frac{1}{2}\alpha_4\right), \end{aligned}$$

其中

$$\frac{1}{2}\alpha_1 - \frac{1}{2}\alpha_2 \in W, \quad \alpha_3 + \frac{1}{2}\alpha_4 \in W^{\perp}.$$

所以 α 在 W 上的内射影为

$$\frac{1}{2}\alpha_1 - \frac{1}{2}\alpha_2 = \left(0, \frac{1}{2}, 1, 1\frac{1}{2}\right).$$

习 题 10.4

1. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是欧氏空间 V 的一组标准正交基，$W = L(\alpha_1,\alpha_2)$，其中

$$\alpha_1 = \varepsilon_1 + \varepsilon_3, \quad \alpha_2 = 2\varepsilon_1 - \varepsilon_2 + \varepsilon_4.$$

(1) 求 W 的一组标准正交基；

(2) 求 $W^{\perp}$的一组标准正交基.

2. 求实系数齐次线性方程组

$$\begin{cases} 2x_1 + x_2 - x_3 + x_4 = 0, \\ x_1 + x_2 - x_3 \qquad\quad = 0, \end{cases}$$

的解空间 W(作为欧氏空间 $\mathbf{R}^4$ 的子空间)的一组标准正交基,并求 $W^{\perp}$.

3. 已知 $\mathbf{R}^4$ 的子空间 W 的一组基:

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

求向量

$$\alpha = (1, -3, 1, -3)$$

在 W 上的内射影.

4. 设 α 是 n 维欧氏空间 V 中一个非零向量.用 W 表示 V 中全部与 α 正交的向量的集合,试证 W 是 V 的一个 $n-1$ 维子空间.

10.5 欧氏空间的同构

在讨论线性空间的时候,曾经证明了数域 $\mathbf{P}$ 上维数相同的线性空间都是同构的.因此,有关线性空间的一般问题可以通过 P^n 来讨论.关于欧氏空间也有相当的结论.不过由于欧氏空间的定义中除了具有线性空间原有的线性运算以外,还有内积,所以在欧氏空间的同构的定义中必须要把这一点反映出来.

定义 13 设 V 与 V' 是两个欧氏空间.σ 是由 V 到 V' 的一个一一对应,如果 σ 满足

(1) $\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta), \alpha, \beta \in V$;

(2) $\sigma(k\alpha)=k\sigma(\alpha), \alpha \in V, k \in R$;

(3) $(\sigma(\alpha), \sigma(\beta))=(\alpha, \beta), \quad \alpha, \beta \in V$.

就称 σ 是 V 到 V' 的一个**同构映射**;如果存在 V 到 V' 的同构映射,则称欧氏空间 V 与 V' 是**同构的**.

从欧氏空间的同构的定义可以看出,如果 σ 是欧氏空间 V 到 V' 的一个同构映射,那么 σ 也是线性空间 V 到 V' 的一个同构映射.因此,以前关于线性空间同构的一些结论对欧氏空间也都成立.而且,同样可证,欧氏空间的同构关系也具有反身性、对称性和

传递性.

定理 9 两个有限维欧氏空间同构的充分必要条件是：它们的维数相等.

证明： 必要性. 如果欧氏空间 V 与 V' 是同构的，那么它们作为线性空间也是同构的，所以它们的维数相等.

充分性. 设 V 与 V' 是两个欧氏空间，它们的维数相等，都等于 n. 在 V 中取一组标准正交基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n;$$

在 V' 中取一组标准正交基

$$\varepsilon'_1, \varepsilon'_2, \cdots, \varepsilon'_n.$$

定义 V 到 V' 的映射 σ 为

$$\sigma(k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n) = k_1\varepsilon'_1 + k_2\varepsilon'_2 + \cdots + k_n\varepsilon'_n,$$

则 σ 是一个一一对应，而且

$$\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta);$$

$$\sigma(k\alpha) = k\sigma(\alpha).$$

如果

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n,$$

$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n,$$

则

$$\sigma(\alpha) = x_1\varepsilon'_1 + x_2\varepsilon'_2 + \cdots + x_n\varepsilon'_n,$$

$$\sigma(\beta) = y_1\varepsilon'_1 + y_2\varepsilon'_2 + \cdots + y_n\varepsilon'_n.$$

由于 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n; \varepsilon'_1, \varepsilon'_2, \cdots, \varepsilon'_n$ 都是标准正交基，所以

$$(\sigma(\alpha), \sigma(\beta)) = x_1y_1 + x_2y_2 + \cdots + x_ny_n = (\alpha, \beta).$$

因此 σ 是由 V 到 V' 的一个同构映射，故 V 与 V' 是同构的. ▮

这个定理说明，从抽象的观点看，欧氏空间的结构完全由它的维数决定. 如果所讨论的问题不涉及欧氏空间的元素及运算的特殊性质，那么维数相同的欧氏空间就具有相同的性质. 由于每个 n 维欧氏空间都与欧氏空间 $\mathbf{R}^4$ 同构，因此，关于 $\mathbf{R}^4$ 中的一些结论在

任一个 n 维欧氏空间中也都成立. 所以当我们讨论关于欧氏空间的一般问题时,可以在 $\mathbf{R}^4$ 中进行讨论而不失普遍性.

定理 9 的证明指出了建立欧氏空间的同构映射的方法. 下面举例说明.

例 1 已知实线性空间 $\mathbf{R}_4[x]$对于内积

$$(f(x),g(x)) = \int_{-1}^{+1} f(x)g(x)\mathrm{d}x$$

构成一个欧氏空间. 求欧氏空间 $\mathbf{R}_4$ 到 $\mathbf{R}_4[x]$的一个同构映射.

为此,首先要求 $\mathbf{R}_4[x]$的一组标准正交基. 任取 $\mathbf{R}_4[x]$的一组基:

$$1, x, x^2, x^3.$$

把它们正交化:

$$f_1 = 1,$$

$$f_2 = x - \frac{(x,f_1)}{(f_1,f_1)}f_1 = x,$$

$$f_3 = x^2 - \frac{(x^2,f_1)}{(f_1,f_1)}f_1 - \frac{(x^2,f_2)}{(f_2,f_2)}f_2 = x^2 - \frac{1}{3},$$

$$f_4 = x^3 - \frac{(x^3,f_1)}{(f_1,f_1)}f_1 - \frac{(x^3,f_2)}{(f_2,f_2)}f_2 - \frac{(x^3,f_3)}{(f_3,f_3)}f_3 = x^3 - \frac{3}{5}x.$$

再将 f_1, f_2, f_3, f_4 单位化,得

$$h_1 = \frac{f_1}{|f_1|} = \frac{\sqrt{2}}{2},$$

$$h_2 = \frac{f_2}{|f_2|} = \frac{\sqrt{6}}{2}x,$$

$$h_3 = \frac{f_3}{|f_3|} = \frac{3\sqrt{10}}{4}x^2 - \frac{\sqrt{10}}{4},$$

$$h_4 = \frac{f_4}{|f_4|} = \frac{5\sqrt{14}}{4}x^3 - \frac{3\sqrt{14}}{4}x.$$

h_1, h_2, h_3, h_4 是 $\mathbf{R}_4[x]$的一组标准正交基.

定义 $\mathbf{R}^4$ 到 $\mathbf{R}_4[x]$的一个映射 σ：

$$(a_1,a_2,a_3,a_4) \to a_1h_1 + a_2h_2 + a_3h_3 + a_4h_4,$$

则 σ 是一个同构映射.

习　题　10.5

1. 设 σ 是欧氏空间 V 到 V' 的一个同构映射，试证：如果 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基，则 $\sigma(\varepsilon_1),\sigma(\varepsilon_2),\cdots,\sigma(\varepsilon_n)$是 V' 的一组标准正交基.
2. 设 V 与 V' 是两个 n 维欧氏空间. $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 及 $\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n'$分别是 V 及 V' 的基，σ 是线性空间 V 到 V' 的同构映射，满足 $\sigma(\varepsilon_i)=\varepsilon_i'$，$i=1,2,\cdots,n$. 试证：$\sigma$ 是欧氏空间 V 到 V' 的同构映射的充分必要条件是：$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 与 $\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n'$的度量矩阵相同.
3. 设 σ 是欧氏空间 V 到 V' 的一个同构映射. W 是 V 的一个子空间，α 是 V 中一个向量. 证明：如果 $\alpha\perp W$，则 $\sigma(\alpha)\perp\sigma(W)$.
4. 设 σ 是欧氏空间 V 到 V' 的一个同构映射. W 是 V 的一个子空间，试证：
   (1) $\sigma(W^{\perp})=(\sigma(W))^{\perp}$；
   (2) 如果 β 是 α 在 W 中的内射影，那么 $\sigma(\beta)$是 $\sigma(\alpha)$在 $\sigma(W)$中的内射影.

10.6　正交变换与对称变换

我们在第 9 章讨论了线性空间的线性变换. 线性变换是线性空间的保持两种线性运算的变换. 至于欧氏空间，除了这两种线性运算以外，还有度量性质. 因此，对于欧氏空间，我们需要讨论保持度量关系的线性变换. 由于度量关系是通过内积来定义的，所以我们引入下列概念.

定义 14　如果欧氏空间 V 的线性变换 A 保持向量的内积不变，即对于 V 中任意两个向量 α 与 β，都有

$$(A\alpha,A\beta) = (\alpha,\beta),$$

则称 A 为一个**正交变换**.

由于欧氏空间中向量的长度、夹角、距离等都可以通过内积表达出来，因此，正交变换保持向量的长度、向量间的夹角和距离不变. 在解析几何中我们也有正交变换的概念，那里的正交变换是保持两个定位向量终点之间距离不变的变换. 这说明欧氏空间的正交变换是几何空间的正交变换的推广.

此外，因为内积也可以通过其他度量性质，如长度、夹角等来表达，因此，正交变换可以通过几个方面来加以描述.

定理 10 $\boldsymbol{A}$ 是 n 维欧氏空间 V 的一个线性变换. 下面的三个条件都是 $\boldsymbol{A}$ 为正交变换的充分必要条件：

(1) $\boldsymbol{A}$ 保持长度不变，即对于 V 中任一向量 α，都有

$$|\boldsymbol{A}\alpha| = |\alpha|;$$

(2) 如果 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组标准正交基，那么 $\boldsymbol{A}\varepsilon_1,\boldsymbol{A}\varepsilon_2,\cdots,\boldsymbol{A}\varepsilon_n$ 也是 V 的标准正交基.

(3) $\boldsymbol{A}$ 在任一组标准正交基下的矩阵是正交矩阵.

证明：(1) 必要性. 如果 $\boldsymbol{A}$ 是正交变换，那么 $\boldsymbol{A}$ 保持内积不变，所以

$$(\boldsymbol{A}\alpha,\boldsymbol{A}\alpha) = (\alpha,\alpha).$$

两边开方，即得

$$|\boldsymbol{A}\alpha| = |\alpha|,$$

所以 $\boldsymbol{A}$ 保持长度不变.

充分性. 如果线性变换 $\boldsymbol{A}$ 保持长度不变，那么对 V 中任意向量 α,β，都有

$$(\boldsymbol{A}\alpha,\boldsymbol{A}\alpha) = (\alpha,\alpha),$$
$$(\boldsymbol{A}\beta,\boldsymbol{A}\beta) = (\beta,\beta),$$
$$(\boldsymbol{A}(\alpha+\beta),\boldsymbol{A}(\alpha+\beta)) = (\alpha+\beta,\alpha+\beta).$$

把最后一式两边展开，得

$$\begin{aligned}(\boldsymbol{A}\alpha+\boldsymbol{A}\beta,\boldsymbol{A}\alpha+\boldsymbol{A}\beta) &= (\boldsymbol{A}\alpha,\boldsymbol{A}\alpha)+(\boldsymbol{A}\beta,\boldsymbol{A}\beta)+2(\boldsymbol{A}\alpha,\boldsymbol{A}\beta)\\ &= (\alpha,\alpha)+(\beta,\beta)+2(\alpha,\beta).\end{aligned}$$

再由前两式，即得

$$(A\alpha, A\beta) = (\alpha, \beta),$$

所以 A 是一个正交变换.

(2) 必要性. 如果 A 是正交变换，那么 A 保持内积与长度不变，因此 A 把标准正交基变为标准正交基.

充分性. 如果 A 把标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 变成标准正交基 $A\varepsilon_1, A\varepsilon_2, \cdots, A\varepsilon_n$，那么

$$(\varepsilon_i, \varepsilon_j) = \begin{cases} 1, & \text{当 } i = j, \\ 0, & \text{当 } i \neq j, \end{cases} \qquad (i, j = 1, 2, \cdots, n),$$

$$(A\varepsilon_i, A\varepsilon_j) = \begin{cases} 1, & \text{当 } i = j, \\ 0, & \text{当 } i \neq j, \end{cases} \qquad (i, j = 1, 2, \cdots, n).$$

任取 V 中两个向量 α, β. 如果

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n,$$

$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n,$$

则

$$A\alpha = x_1A\varepsilon_1 + x_2A\varepsilon_2 + \cdots + x_nA\varepsilon_n,$$

$$A\beta = y_1A\varepsilon_1 + y_2A\varepsilon_2 + \cdots + y_nA\varepsilon_n.$$

于是

$$(A\alpha, A\beta) = x_1y_1 + x_2y_2 + \cdots + x_ny_n = (\alpha, \beta).$$

所以 A 是一个正交变换.

(3) 必要性. 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 V 的一组标准正交基. 那么根据 (2)，$A\varepsilon_1, A\varepsilon_2, \cdots, A\varepsilon_n$ 也是标准正交基，所以 A 在 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵 A 是正交矩阵.

充分性. 如果线性变换 A 在标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵 A 是一个正交矩阵，那么 $A\varepsilon_1, A\varepsilon_2, \cdots, A\varepsilon_n$ 一定也是标准正交基. 故由(2)，A 是一个正交变换. |

因为正交矩阵是可逆的，并且其逆矩阵也是正交矩阵；两个正交矩阵的乘积仍是正交矩阵. 因此有：

命题 1 (1) 正交变换是可逆变换，其逆变换仍是正交变换；

(2) 两个正交变换的乘积也是正交变换.

我们知道正交矩阵的行列式为 1 或 -1. 所以正交变换的矩阵的行列式也等于 1 或 -1. 如果正交变换 $\boldsymbol{A}$ 的行列式等于 1，则称 $\boldsymbol{A}$ 是一个**旋转**，或称 $\boldsymbol{A}$ 是一个**第一类正交变换**；如果正交变换 $\boldsymbol{A}$ 的行列式等于 -1，则称 $\boldsymbol{A}$ 为一个**第二类正交变换**.

例 1 平面上的定位向量构成一个二维欧氏空间. 把平面围绕坐标原点按反时针方向旋转 θ 角，就是一个线性变换，记作 $\boldsymbol{T}_\theta$. 这个线性变换把基 $(1,0),(0,1)$ 变为

$$\boldsymbol{T}_\theta(1,0) = (\cos\theta, \sin\theta),$$
$$\boldsymbol{T}_\theta(0,1) = (-\sin\theta, \cos\theta).$$

所以 $\boldsymbol{T}_\theta$ 在基 $(1,0),(0,1)$ 下的矩阵是

$$T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

因为 $(1,0),(0,1)$ 是一组标准正交基，T 是一个正交矩阵，所以 $\boldsymbol{T}_\theta$ 是一个正交变换. 又因

$$|T| = 1,$$

所以 $\boldsymbol{T}_\theta$ 是一个第一类正交变换.

例 2 在平面上取定一个直角坐标系，把每个向量对于 x 轴作反射，即把坐标为 (x,y) 的向量对应于 $(x,-y)$. 这是一个线性变换. 它在标准正交基 $(1,0),(0,1)$ 下的矩阵是

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

所以这是一个第二类的正交变换.

一般地，如果在 n 维欧氏空间 V 中取定一组标准正交基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$. 定义 V 的线性变换 $\boldsymbol{A}$ 为

$$\begin{aligned} &\boldsymbol{A}(x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + \cdots + x_n\boldsymbol{\varepsilon}_n) \\ &= x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + \cdots + x_{n-1}\boldsymbol{\varepsilon}_{n-1} - x_n\boldsymbol{\varepsilon}_n, \end{aligned}$$

那么 $\boldsymbol{A}$ 在这组标准正交基下的矩阵是

$$\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -1 \end{bmatrix},$$

所以这是一个第二类的正交变换.

我们知道,在 n 维线性空间 V 中取定一组基以后,V 的线性变换与 n 阶矩阵之间就建立了一个一一对应.通过这个对应关系,就有可能利用矩阵来讨论线性变换.而且,这个对应关系也给了 n 阶矩阵以几何意义.我们在以前曾多次看到,有时候应用线性变换的概念来解决某些矩阵问题也是非常方便的.下面的例子再一次说明这个事实.

例 3 证明:正交矩阵的实特征值都等于 1 或 -1.

证明: 设 A 是一个 n 阶正交矩阵.在欧氏空间 $\mathbf{R}^n$ 中取定一组标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$.定义 $\mathbf{R}^n$ 的线性变换 $\boldsymbol{A}$ 为

$$\boldsymbol{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\mathrm{A},$$

即 $\boldsymbol{A}$ 是在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下以 A 为矩阵的线性变换.因为 A 是正交矩阵,所以 $\boldsymbol{A}$ 是 V 的一个正交变换.

设 λ_0 是 A 的一个实特征值.那么 λ_0 也是线性变换 $\boldsymbol{A}$ 的一个特征值,故有非零向量 $\alpha\in R^4$,使得

$$\boldsymbol{A}\alpha=\lambda_0\alpha.$$

于是

$$\begin{aligned}(\alpha,\alpha)&=(\boldsymbol{A}\alpha,\boldsymbol{A}\alpha)\\&=(\lambda_0\alpha,\lambda_0\alpha)\\&=\lambda_0^2(\alpha,\alpha).\end{aligned}$$

由于 $(\alpha,\alpha)\neq 0$,所以

$$\lambda_0^2=1,\quad \lambda_0=\pm 1.$$

这个例题也证明了正交变换的实特征值都等于 1 或 -1.

下面介绍对称变换及其与对称矩阵的关系.

定义 15 $\boldsymbol{A}$ 是欧氏空间 V 的一个线性变换. 如果对 V 中任意两个向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$,都有

$$(\boldsymbol{A\alpha},\boldsymbol{\beta}) = (\boldsymbol{\alpha},\boldsymbol{A\beta}),$$

则 $\boldsymbol{A}$ 称为一个**对称变换**。

如果 $\boldsymbol{A}$ 是 n 维欧氏空间 V 的一个对称变换,$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 是 V 的一组标准正交基,并设 $\boldsymbol{A}$ 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 下的矩阵是

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

于是

$$(\boldsymbol{A\varepsilon}_i,\boldsymbol{\varepsilon}_j) = (a_{1i}\boldsymbol{\varepsilon}_1 + a_{2i}\boldsymbol{\varepsilon}_2 + \cdots + a_{ni}\boldsymbol{\varepsilon}_n,\boldsymbol{\varepsilon}_j) = a_{ji},$$

$$(\boldsymbol{\varepsilon}_i,\boldsymbol{A\varepsilon}_j) = (\boldsymbol{\varepsilon}_i,a_{1j}\boldsymbol{\varepsilon}_1 + a_{2j}\boldsymbol{\varepsilon}_2 + \cdots + a_{nj}\boldsymbol{\varepsilon}_n) = a_{ij},$$

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

因为 $\boldsymbol{A}$ 是一个对称变换,所以

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

这说明 A 是一个对称矩阵,这个条件也是 $\boldsymbol{A}$ 为对称变换的充分条件,即

定理 11 n 维欧氏空间 V 的线性变换 $\boldsymbol{A}$ 是对称变换的充要条件为:$\boldsymbol{A}$ 在任一组标准正交基下的矩阵都是对称矩阵.

证明: 条件的必要性已在前面证明.

如果 $\boldsymbol{A}$ 在标准正交基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 下的矩阵 A 是对称的,那么对于 V 中任意两个向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$:

$$\boldsymbol{\alpha} = [\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n]X; \quad \boldsymbol{\beta} = [\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n]Y$$

(注意,X,Y 分别是 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 对于基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 的坐标所对应的列向量),都有

$$(A\alpha,\beta) = (AX)^{\mathrm{T}}Y = X^{\mathrm{T}}A^{\mathrm{T}}Y = X^{\mathrm{T}}AY = (\alpha,A\beta).$$

所以 $\boldsymbol{A}$ 是一个对称变换.

这样,在 n 维欧氏空间 V 中取定一组基后,V 的对称变换与 n 阶实对称矩阵之间就建立起一个一一对应.

我们以前曾经证明过:如果 A 是一个实对称矩阵,那么一定有一个正交矩阵 T,使得 $T^{-1}AT$ 为对角矩阵(第 4 章定理 7). 把这个结论用于对称变换,即得:

定理 12 如果 $\boldsymbol{A}$ 是 n 维欧氏空间 V 的一个对称变换,那么可找到 V 的一组标准正交基,使 $\boldsymbol{A}$ 在这组基下的矩阵是对角矩阵.

证明: 任取 V 的一组标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,则 $\boldsymbol{A}$ 在这组基下的矩阵 A 是一个实对称矩阵,因此有正交矩阵 T 使得 $T^{-1}AT$ 是对角矩阵. 令

$$[\eta_1,\eta_2,\cdots,\eta_n]=[\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n]T.$$

因 T 是正交矩阵,所以 $\eta_1,\eta_2,\cdots,\eta_n$ 也是 V 的标准正交基. 而且,$\boldsymbol{A}$ 在 $\eta_1,\eta_2,\cdots,\eta_n$ 下的矩阵就是对角矩阵 $T^{-1}AT$. |

当然,如果 V 的线性变换 $\boldsymbol{A}$ 在某一组标准正交基下的矩阵是对角形的,那么 $\boldsymbol{A}$ 一定是对称变换.

例 4 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是欧氏空间 V 的一组标准正交基. $\boldsymbol{A}$ 是 V 的一个线性变换. 已知

$$\boldsymbol{A}\varepsilon_1=\varepsilon_1+\varepsilon_2-\varepsilon_4,\quad \boldsymbol{A}\varepsilon_2=\varepsilon_1+\varepsilon_2-\varepsilon_3,$$
$$\boldsymbol{A}\varepsilon_3=-\varepsilon_2+\varepsilon_3+\varepsilon_4,\quad \boldsymbol{A}\varepsilon_4=-\varepsilon_1+\varepsilon_3+\varepsilon_4,$$

(1) 证明 $\boldsymbol{A}$ 是一个对称变换;

(2) 求 V 的一组标准正交基,使 $\boldsymbol{A}$ 在这组基下的矩阵是对角矩阵.

解: (1) $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的矩阵为

$$A=\begin{bmatrix}1&1&0&-1\\1&1&-1&0\\0&-1&1&1\\-1&0&1&1\end{bmatrix},$$

因为 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是标准正交基，A 是对称矩阵，所以 $\boldsymbol{A}$ 是一个对称变换.

(2)

$$|\lambda E - A| = \begin{vmatrix} \lambda-1 & -1 & 0 & 1 \\ -1 & \lambda-1 & 1 & 0 \\ 0 & 1 & \lambda-1 & -1 \\ 1 & 0 & -1 & \lambda-1 \end{vmatrix}$$
$$=(\lambda-1)^2(\lambda-3)(\lambda+1),$$

所以 $\boldsymbol{A}$ 的特征值为 $1,1,3,-1$.

解得 $\boldsymbol{A}$ 的属于 1 的特征向量为

$$k_1(\varepsilon_1+\varepsilon_3)+k_2(\varepsilon_2+\varepsilon_4),$$

k_1,k_2 是不全为零的任意实数.

$\boldsymbol{A}$ 的属于 3 的特征向量为

$$k(\varepsilon_1+\varepsilon_2-\varepsilon_3-\varepsilon_4),$$

k 是任意非零实数.

$\boldsymbol{A}$ 的属于 -1 的特征向量为

$$l(\varepsilon_1-\varepsilon_2-\varepsilon_3+\varepsilon_4),$$

l 是任意非零实数.

将 $\varepsilon_1+\varepsilon_3,\varepsilon_2+\varepsilon_4$（因为这两个向量已经正交，所以不必正交化），$\varepsilon_1+\varepsilon_2-\varepsilon_3-\varepsilon_4,\varepsilon_1-\varepsilon_2-\varepsilon_3+\varepsilon_4$ 单位化，得

$$\eta_1=\frac{\sqrt{2}}{2}\varepsilon_1+\frac{\sqrt{2}}{2}\varepsilon_3,\ \eta_2=\frac{\sqrt{2}}{2}\varepsilon_2+\frac{\sqrt{2}}{2}\varepsilon_4,$$

$$\eta_3=\frac{1}{2}\varepsilon_1+\frac{1}{2}\varepsilon_2-\frac{1}{2}\varepsilon_3-\frac{1}{2}\varepsilon_4,$$

$$\eta_4=\frac{1}{2}\varepsilon_1-\frac{1}{2}\varepsilon_2-\frac{1}{2}\varepsilon_3+\frac{1}{2}\varepsilon_4.$$

$\eta_1,\eta_2,\eta_3,\eta_4$ 是 V 的一组标准正交基. 线性变换 $\boldsymbol{A}$ 在这组基下的矩

阵为对角矩阵

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 3 & \\ & & & -1 \end{bmatrix}.$$

关于实对称矩阵的其他一些性质也可以得到对称变换的相应的性质,这里就不重复叙述了.

习 题 10.6

1. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 及 $\beta_1,\beta_2,\cdots,\beta_s$ 是欧氏空间 V 中两个由单位向量组成的正交向量组,试证:存在 V 的一个正交变换把 α_i 变为 $\beta_i(i=1,2,\cdots,s)$.
2. 设

$$\alpha_1=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},0,0\right),\qquad \beta_1=\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),$$

$$\alpha_2=\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},0,0\right),\quad \beta_2=\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right),$$

$$\alpha_3=\left(0,0,\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right),\qquad \beta_3=\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right).$$

(1) 求欧氏空间 $\mathbf{R}^4$ 的一个正交变换 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}\alpha_i=\beta_i,\quad i=1,2,3;$$

(2) 求 $\boldsymbol{A}$ 在基

$$\varepsilon_1=(1,0,0,0),\quad \varepsilon_3=(0,0,1,0),$$
$$\varepsilon_2=(0,1,0,0),\quad \varepsilon_4=(0,0,0,1)$$

下的矩阵;

(3) 问满足这个条件的正交变换共有几个?

3. 设 $\boldsymbol{A}$ 是 n 维欧氏空间 V 的一个正交变换,W 是 $\boldsymbol{A}$ 的一个不变子空间,证明:W 的正交补 $W^{\perp}$ 也是 $\boldsymbol{A}$ 的不变子空间.
4. 证明:奇数维欧氏空间的旋转一定以 1 作为它的一个特征值.
5. 证明:第二类正交变换一定以 -1 作为它的特征值.

6. 试证：如果 A 是 n 维欧氏空间 V 的一个对称变换，W 是 A 的一个不变子空间，则 W 的正交补 $W^{\perp}$ 也是 A 的不变子空间.

7. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是欧氏空间 V 的一组标准正交基，A 是 V 的一个线性变换. 已知

$$\begin{aligned}
A\varepsilon_1 &= \varepsilon_1 - \varepsilon_2 - \varepsilon_3 + \varepsilon_4 \\
A\varepsilon_2 &= -\varepsilon_1 + \varepsilon_2 + \varepsilon_3 - \varepsilon_4 \\
A\varepsilon_3 &= -\varepsilon_1 + \varepsilon_2 + \varepsilon_3 - \varepsilon_4 \\
A\varepsilon_4 &= \varepsilon_1 - \varepsilon_2 - \varepsilon_3 + \varepsilon_4
\end{aligned}$$

(1) 证明 A 是一个对称变换；

(2) 求 V 的一组标准正交基，使 A 在这组基下的矩阵为对角矩阵.

*10.7 最小二乘法

这一节介绍欧氏空间的一个应用——最小二乘法问题.

我们知道，实系数线性方程组

$$\begin{cases}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s = b_1, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s = b_2, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{ns}x_s = b_n
\end{cases} \tag{1}$$

可能无解，即任何一组实数 $x_1,x_2,\cdots,x_s$，都使

$$\sum_{i=1}^{n}(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{is}x_s - b_i)^2 \tag{2}$$

不等于零. 我们设法找 $x_1^0,x_2^0,\cdots,x_s^0$ 使(2)式最小，用它作为方程组(1)的近似解. 这样的 $x_1^0,x_2^0,\cdots,x_s^0$ 称为方程组(1)的**最小二乘解**. 这种问题就叫做最小二乘法问题.

在应用欧氏空间的概念来解决最小二乘法问题以前，我们先讨论一个向量与一个子空间的距离问题.

在中学几何中，我们学过一个点到一个平面(或一条直线)上

所有点的距离以垂线为最短. 下面证明欧氏空间中一个向量与某个子空间中各向量间的距离也以"垂线"为最短.

设 W 是欧氏空间 V 的一个子空间,它是由向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 生成的,即 $W=L(\alpha_1,\alpha_2,\cdots,\alpha_s)$. 我们以前曾定义过 V 中一个向量 β 垂直于子空间 W,就是指 β 垂直于 W 中任何一个向量,因此 β 垂直于子空间 W 的充分必要条件是:β 垂直于每个 $\alpha_i(i=1,2,\cdots,s)$.

命题 1 设 W 是欧氏空间 V 的一个子空间,α 是 V 中一个向量. 再设 β 是 W 中一个向量,使 $\alpha-\beta$ 垂直于 W,则对 W 中任一向量 γ,都有

$$|\alpha-\beta|\leqslant|\alpha-\gamma|$$

参见图 10-1.

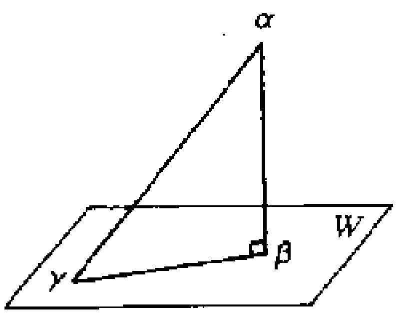


图 10-1

证明: $\alpha-\gamma=(\alpha-\beta)+(\beta-\gamma)$. 因为 W 是一个子空间,所以由 $\beta\in W,\gamma\in W$ 知 $\beta-\gamma\in W$,故 $\alpha-\beta$ 垂直于 $\beta-\gamma$. 由勾股定理,即得

$$|\alpha-\beta|^2+|\beta-\gamma|^2=|\alpha-\gamma|^2.$$

所以

$$|\alpha-\beta|\leqslant|\alpha-\gamma|.$$

命题 1 的意思就是向量到子空间中各向量的距离以垂线为最短.

由于 α 可表示成

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

其中 $\beta\in W,\alpha-\beta\in W^{\perp}$. 因此 β 就是 α 在子空间 W 上的内射影. 由此可知,向量到子空间各向量的距离以该向量到它在此空间上的内射影的距离为最短.

下面利用这个事实讨论最小二乘法,并给出最小二乘解所满足的代数条件.

令

$$A=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1s}\\ a_{21} & a_{22} & \cdots & a_{2s}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{ns}\end{bmatrix},\quad B=\begin{bmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{bmatrix},$$

$$X=\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_s\end{bmatrix},\quad Y=\begin{bmatrix}\sum\limits_{j=1}^{s}a_{1j}x_j\\ \sum\limits_{j=1}^{s}a_{2j}x_j\\ \vdots\\ \sum\limits_{j=1}^{s}a_{nj}x_j\end{bmatrix}=AX. \tag{3}$$

应用距离的概念,(2)式就是

$$|Y-B|^2.$$

最小二乘法就是找 $x_1^0,x_2^0,\cdots,x_s^0$,使 Y 与 B 的距离为最短. 但是由(3)式,知道向量 Y 可表示成 A 的列向量的线性组合:

$$Y=x_1\begin{bmatrix}a_{11}\\ a_{21}\\ \vdots\\ a_{n1}\end{bmatrix}+x_2\begin{bmatrix}a_{12}\\ a_{22}\\ \vdots\\ a_{n2}\end{bmatrix}+\cdots+x_s\begin{bmatrix}a_{1s}\\ a_{2s}\\ \vdots\\ a_{ns}\end{bmatrix}.$$

把 A 的各列向量依次记成 $\alpha_1,\alpha_2,\cdots,\alpha_s$,用 W 表示由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 生成的子空间:

$$W=L(\alpha_1,\alpha_2,\cdots,\alpha_s),$$

则 $Y\in W$.

因此,为了找 X 使(2)式为最小,就要在 $W=L(\alpha_1,\alpha_2,\cdots,\alpha_s)$ 中找一个向量 Y,使得 B 到 Y 的距离比 B 到 W 中其他向量的距离都短.

应用前面的讨论，如果

$$Y = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_s\alpha_s$$

就是所求的向量，那么

$$C = B - Y = B - AX$$

必垂直于子空间 W. 而 C 垂直于子空间 W 的充要条件为

$$(\alpha_1, C) = (\alpha_2, C) = \cdots = (\alpha_s, C) = 0.$$

根据矩阵乘法规则，这些等式可写成矩阵相乘的式子

$$\alpha_1^{\mathrm{T}}C = \alpha_2^{\mathrm{T}}C = \cdots = \alpha_s^{\mathrm{T}}C = 0.$$

而 $\alpha_1^{\mathrm{T}}, \alpha_2^{\mathrm{T}}, \cdots, \alpha_s^{\mathrm{T}}$ 恰好是 A^{T} 的各行，因此，由上面一些等式可得

$$A^{\mathrm{T}}C = 0,$$

即

$$A^{\mathrm{T}}(B - AX) = 0$$

或

$$A^{\mathrm{T}}AX = A^{\mathrm{T}}B.$$

这就是最小二乘解所满足的线性方程组. 它的系数矩阵是 $A^{\mathrm{T}}A$，常数项是 $A^{\mathrm{T}}B$.

例 1 求下列方程组的最小二乘解

$$\begin{cases} 0.13x_1 - 0.29x_2 = 0.22, \\ 0.55x_1 + 1.72x_2 = -0.78, \\ 1.23x_1 - 0.79x_2 = 0.74, \\ 0.37x_1 + 0.48x_2 = -0.23. \end{cases}$$

解： 令

$$A = \begin{bmatrix} 0.13 & -0.29 \\ 0.55 & 1.72 \\ 1.23 & -0.79 \\ 0.37 & 0.48 \end{bmatrix}, \quad B = \begin{bmatrix} 0.22 \\ -0.78 \\ 0.74 \\ -0.23 \end{bmatrix},$$

则原方程组可表示成

$$AX = B.$$

因为

$$A^{\mathrm{T}}A=\begin{bmatrix}1.97 & 0.11\\ 0.11 & 3.90\end{bmatrix},\quad A^{\mathrm{T}}B=\begin{bmatrix}0.42\\ -2.10\end{bmatrix},$$

解线性方程组

$$\begin{cases}1.97x_1+0.11x_2=0.42,\\ 0.11x_1+3.90x_2=-2.10,\end{cases}$$

即得原方程组的最小二乘解为：

$$x_1=0.24,\quad x_2=-0.55.$$

习　题　10.7

1. 求下列方程组的最小二乘解

$$\begin{cases}0.84x_1+0.56x_2=0.37,\\ 0.51x_1+0.36x_2=0.28,\\ 0.71x_1+0.53x_2=0.43,\\ 0.23x_1+0.75x_2=0.25.\end{cases}$$

内容提要

1. 欧氏空间

(1) 欧氏空间的定义

(2) 内积的性质

(3) 长度、夹角、距离的定义以及它们之间的关系

(4) 柯西-布涅可夫斯基不等式

2. 度量矩阵

(1) 度量矩阵的定义

(2) 度量矩阵是正定矩阵

(3) 两组基的度量矩阵之间的关系

如果基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 及 $\eta_1,\eta_2,\cdots,\eta_n$ 的度量矩阵分别是 A 及

B,由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 到 $\eta_1,\eta_2,\cdots,\eta_n$ 的过渡矩阵是 C,那么

$$B = C^{\mathrm{T}}AC.$$

(4) 任一正定矩阵都可看作某组基的度量矩阵

3. 标准正交基

(1) 正交向量组、单位正交向量组

1) 定义

2) 任给一个线性无关向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$,可找到一个单位正交向量组 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s$,使得

$$L(\alpha_1,\cdots,\alpha_i) = L(\varepsilon_1,\cdots,\varepsilon_i), \quad i = 1,2,\cdots,s.$$

(2) 正交基、标准正交基的定义

(3) 标准正交基的判别方法

1) 基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是标准正交基的充分必要条件是:它的度量矩阵是单位矩阵

2) 由标准正交基到标准正交基的过渡矩阵是正交矩阵

3) 如果两组基之间的过渡矩阵是正交矩阵,那么从其中一组基是标准正交基可推出另一组基也是标准正交基

(4) 标准正交基的求法

(5) n 维欧氏空间中任一组正交的单位向量组都可扩充成 V 的一组标准正交基

(6) 任给一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$,都可找到一组标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,使

$$L(\varepsilon_1,\cdots,\varepsilon_i) = L(\alpha_1,\cdots,\alpha_i),\ i = 1,2,\cdots,n$$

(7) 标准正交基对于坐标和内积的计算

设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是标准正交基,那么

1) 对任一向量 α 有:

$$\alpha = (\alpha,\varepsilon_1)\varepsilon_1 + (\alpha,\varepsilon_2)\varepsilon_2 + \cdots + (\alpha,\varepsilon_n)\varepsilon_n.$$

即 α 对基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的坐标是

$$((\alpha,\varepsilon_1),(\alpha,\varepsilon_2),\cdots,(\alpha,\varepsilon_n))$$

2）如果

$$\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n,$$
$$\beta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n,$$

那么

$$(\alpha,\beta) = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

4. 子空间

(1) 向量与子空间正交(垂直)的定义

(2) 两个子空间正交的定义

(3) 子空间的正交补

1）定义、存在性、唯一性及求法

2）向量 α 在子空间 W 中的内射影的定义和求法

5. 欧氏空间的同构

(1) 欧氏空间同构的定义

(2) 两个有限维欧氏空间同构的充分必要条件是：它们的维数相同

(3) 任一个 n 维欧氏空间都与欧氏空间 $\mathbf{R}^n$ 同构

6. 正交变换

(1) 正交变换的定义

(2) 欧氏空间 V 的线性变换 $\boldsymbol{A}$ 是正交变换的一些充分必要条件

(3) 第一类正交变换与第二类正交变换

7. 对称变换

(1) 对称变换的定义

(2) 欧氏空间 V 的线性变换 $\boldsymbol{A}$ 是对称变换的充分必要条件

(3) 对于欧氏空间 V 的对称变换 $\boldsymbol{A}$，可找 V 的一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，使 $\boldsymbol{A}$ 在此基下的矩阵是对角矩阵：存在性和求法

*8. 最小二乘法

实线性方程组最小二乘解的求法

复 习 题 10

1. 在 $\mathbf{R}^4$ 中设 $\alpha=(1,1,-1,1)$，$\beta=(1,-1,2,3)$，$\gamma=(1,2,3,a)$．

(1) 求 (α,β)；

(2) 求 a 使 $(\alpha,\gamma)=(\beta,\gamma)$；

(3) 求 a 使 $d(\alpha,\gamma)=d(\beta,\gamma)$．

2. 设 α,β 是欧氏空间中两个向量．试证：

$$|\alpha-\beta|^2=|\alpha|^2+|\beta|^2-2|\alpha|\ |\beta|\cos<\alpha,\beta>.$$

3. 证明：欧氏空间中两个向量 α、β 正交的充分必要条件是：对任意实数 t，都有

$$|\alpha+t\beta|\geqslant|\alpha|.$$

4. 设 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 是欧氏空间中一组向量，令

$$D=\begin{bmatrix}(\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_m)\\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_m)\\ \cdots & \cdots & \cdots & \cdots\\ (\alpha_m,\alpha_1) & (\alpha_m,\alpha_2) & \cdots & (\alpha_m,\alpha_m)\end{bmatrix},$$

证明：当且仅当 D 是可逆矩阵时，向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 是线性无关的．

5. 设 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 是欧氏空间 V 的一组基，已知 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的度量矩阵为

$$A=\begin{bmatrix}1 & 1 & -1 & 0\\ 1 & 2 & -1 & 1\\ -1 & -1 & 2 & -1\\ 0 & 1 & -1 & 4\end{bmatrix},$$

(1) 求 V 的一组标准正交基；

(2) 求基

$$\beta_1=\alpha_1+2\alpha_2+\alpha_3,\qquad \beta_2=\alpha_1-\alpha_2+\alpha_4,$$
$$\beta_3=\alpha_2+\alpha_3+\alpha_4,\qquad \beta_4=\alpha_1+2\alpha_2-\alpha_3.$$

的度量矩阵．

6. 在欧氏空间 $\mathbf{R}^4$ 中，设

$$\alpha_1=(1,-1,1,0),\qquad \alpha_2=(1,0,2,-1),$$
$$\alpha_3=(1,1,1,1).$$

(1) 求正交的单位向量组 $\varepsilon_1,\varepsilon_2,\varepsilon_3$,使

$$L(\varepsilon_1,\cdots,\varepsilon_i) = L(\alpha_1,\cdots,\alpha_i), i = 1,2,3;$$

(2) 将 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 扩充成 $\mathbf{R}^4$ 的一组标准正交基.

7. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是欧氏空间 V 的一组标准正交基,再设

$$\alpha_1 = \varepsilon_1 + \varepsilon_2 - \varepsilon_3 + 2\varepsilon_4,$$
$$\alpha_2 = \varepsilon_1 - \varepsilon_2 - \varepsilon_3 - 4\varepsilon_4,$$
$$\alpha_3 = \varepsilon_1 + 3\varepsilon_2 - \varepsilon_3 + 8\varepsilon_4,$$
$$W = L(\alpha_1,\alpha_2,\alpha_3).$$

(1) 求 W 的一组标准正交基;

(2) 求 $W^{\perp}$的一组标准正交基;

(3) 求 $\alpha = \varepsilon_1 + 4\varepsilon_2 - 4\varepsilon_3 - \varepsilon_4$ 在 W 上的内射影.

8. 在欧氏空间 $\mathbf{R}^n$ 中,设

$$\alpha_1 = (a_{11},a_{12},\cdots,a_{1n}),$$
$$\alpha_2 = (a_{21},a_{22},\cdots,a_{2n}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_s = (a_{s1},a_{s2},\cdots,a_{sn}),$$
$$W_1 = L(\alpha_1,\alpha_2,\cdots,\alpha_s).$$

再设 W_2 是实系数线性方程组

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

的解空间,证明 W_2 是 W_1 的正交补.

9. 证明:$\mathbf{R}^n$ 的任一个子空间都是某个实系数齐次线性方程组的解空间.

10. 设

$$\alpha_i = (a_{1i},a_{2i},\cdots,a_{ni}), i = 1,2,\cdots,m,$$

及 $\qquad \beta = (b_1,b_2,\cdots,b_n)$

都是 n 维欧氏空间 $\mathbf{R}^n$ 中的向量. 令

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}, \qquad B=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

$$W=L(\alpha_1,\alpha_2,\cdots,\alpha_m).$$

试证：$\beta\in W^{\perp}$ 的充分必要条件是：$A^{T}B=0$.

11. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s;\beta_1,\beta_2,\cdots,\beta_t$ 都是 n 维实向量. 试证：如果

$$(\alpha_i,\beta_j)=0,\quad (i=1,2,\cdots,s;\quad j=1,2,\cdots,t),$$

则

$$\mathrm{r}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}+\mathrm{r}\{\beta_1,\beta_2,\cdots,\beta_t\}\leqslant n.$$

12. 设 W_1,W_2 是欧氏空间 V 的两个子空间.

试证：

(1) $(W_1+W_2)^{\perp}=W_1^{\perp}\cap W_2^{\perp}$；

(2) $(W_1\cap W_2)^{\perp}=W_1^{\perp}+W_2^{\perp}$.

13. 设 $\boldsymbol{A}$ 是欧氏空间 V 的一个变换，试证：如果 $\boldsymbol{A}$ 保持内积不变，即对于 V 中任意两个向量 α,β，都有

$$(\boldsymbol{A}\alpha,\boldsymbol{A}\beta)=(\alpha,\beta),$$

那么，它一定是线性的，因而是一个正交变换.

14. 设 η 是 n 维欧氏空间 V 中一个单位向量. 定义 V 的变换 $\boldsymbol{A}$ 为：

$$\boldsymbol{A}\alpha=\alpha-2(\eta,\alpha)\eta,\quad (\alpha\in V),$$

证明：

(1) $\boldsymbol{A}$ 是 V 的一个正交变换(这样的正交变换称为镜面反射)；

(2) $\boldsymbol{A}$ 是第二类正交变换；

(3) $\boldsymbol{A}$ 的对于特征值 1 的特征子空间是 $n-1$ 维的；

(4) $\boldsymbol{A}$ 在 V 的某一组标准正交基下的矩阵是

$$\begin{bmatrix} -1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 1 \end{bmatrix}.$$

15. 证明：n 维欧氏空间 V 的正交变换 $\boldsymbol{A}$ 是镜面反射的充分必要条件为：$\boldsymbol{A}$ 以 1 为特征值，并且恰有 $n-1$ 个属于 1 的线性无关的特征向量.

16. 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 及 $\beta_1,\beta_2,\cdots,\beta_s$ 是欧氏空间 V 中的两组向量，证明：存在 V 的一个正交变换 $\boldsymbol{A}$，使

$$\boldsymbol{A}\alpha_i = \beta_i, \qquad i = 1,2,\cdots,s,$$

的必要充分条件是

$$(\alpha_i,\alpha_j) = (\beta_i,\beta_j), \qquad i,j = 1,2,\cdots,s.$$

17. 证明：如果 n 维欧氏空间 V 的线性变换 $\boldsymbol{A}$ 既是正交变换，又是对称变换，则可找到 V 的一组标准正交组，使 $\boldsymbol{A}$ 在这组基下的矩阵为

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{bmatrix}.$$

18. 设 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 是欧氏空间 V 的一组标准正交基，$\boldsymbol{A}$ 是 V 的一个线性变换. 已知

$$\begin{aligned} \boldsymbol{A}\varepsilon_1 &= 2\varepsilon_1 + 3\varepsilon_2 + 3\varepsilon_3 + 3\varepsilon_4, \\ \boldsymbol{A}\varepsilon_2 &= 3\varepsilon_1 + 2\varepsilon_2 - 3\varepsilon_3 - 3\varepsilon_4, \\ \boldsymbol{A}\varepsilon_3 &= 3\varepsilon_1 - 3\varepsilon_2 + 2\varepsilon_3 - 3\varepsilon_4, \\ \boldsymbol{A}\varepsilon_4 &= 3\varepsilon_1 - 3\varepsilon_2 - 3\varepsilon_3 + 2\varepsilon_4. \end{aligned}$$

(1) 证明 $\boldsymbol{A}$ 是一个正交变换；

(2) 求 V 的一组标准正交基，使 $\boldsymbol{A}$ 在这组基下的矩阵为对角矩阵.

习题答案与提示

第6章　多项式

习　题　6.1

1. (1) $f(x)+g(x)=x^4+4x^3-2x^2-9x+2$, $f(x)-g(x)=x^4+x-6$,
$f(x)g(x)=2x^7+3x^6-9x^5-13x^4+13x^3+18x^2-6x-8$;
(2) $f(x)+g(x)=x^4-x^3-3x^2+x+2$, $f(x)-g(x)=x^4-x^3-5x^2+3x$,
$f(x)g(x)=x^6-2x^5-2x^4+5x^3-5x^2+x+1$;
(3) $f(x)+g(x)=-7x^2+3x+11$, $f(x)-g(x)=2x^4-4x^3-5x^2+7x+7$,
$f(x)g(x)=-x^8+4x^7+x^6-17x^5+13x^4+21x^3-31x^2-8x+18$;
(4) $f(x)+g(x)=2x^4-x^3+3x^2-4x+2$, $f(x)-g(x)=2x^4-x^3+x^2+2x$,
$f(x)g(x)=2x^8-7x^5+7x^4-8x^3+6x^2-4x+1$;
(5) $f(x)+g(x)=2x^3+(1+i)x^2-(1+i)x-2i$, $f(x)-g(x)=(1+i)x^2-(1-3i)x-(2-2i)$,
$f(x)g(x)=x^6+(1+i)x^5-(1+i)x^4+(2-4i)x^3+(5+i)x^2+(1+5i)x-(1-2i)$;
2. (1) $x^6+x^5-3x^4-12x^3+x^2+43x+1$;　(2) $2x^5-2x^3$.
3. $k=-2, l=1, m=3$.
4. 当 $f(x)$, $g(x)$的次数相等而且它们的首项系数之和等于零时,次数公式中小于号成立,否则,等号成立.
5. (2) 不一定.

习　题　6.2(1)

1. (1) $q(x)=2x^2+3x+11$, $r(x)=25x-5$;

(2) $q(x)=\frac{1}{3}x-\frac{7}{9}$，$r(x)=-2\frac{8}{9}x-\frac{2}{9}$；

(3) $q(x)=x^2-x+3$，$r(x)=0$；

(4) $q(x)=x+1-2i$，$r(x)=(4+2i)x+(1-3i)$；

(5) $q(x)=x^2+2x-3$，$r(x)=0$.

2. (1) $q(x)=3x^3+6x^2+2x-1$，$r(x)=-6$；(2) $f(2)=-6$.

3. (1) $q_1(x)=a_3x^2+(a_3c+a_2)x+a_3c^2+a_2c+a_1$，

$r_1(x)=a_3c^3+a_2c^2+a_1c+a_0$.

(2) $q_2(x)=\frac{1}{2}a_3x^2+\frac{1}{2}(a_3c+a_2)x+\frac{1}{2}(a_3c^2+a_2c+a_1)$，$r_2(x)=a_3c^3+a_2c^2+a_1c+a_0$.

$q_2(x)=\frac{1}{2}q_1(x)$，$r_2(x)=r_1(x)$.

习　题　6.2(2)

1. (1) 不是；(2) 不是.

2. $q(x)=x^2+3x-2$，$r(x)=(k+2)x^2+(l+2)$；当 $k=l=-2$ 时，$g(x)\mid f(x)$.

3. $l=4$，$m=2$.

4. $l=2$，$m=-4$.

5. 提示：应用命题 2.

6. (1) 提示：用反证法；(2) 不一定.

习　题 6.2(3)

1. (1) -1；(2) -384；(3) $-2+6i$.

2. $k=2$ 或 $-\frac{1}{3}$.

3. (1) $(x-3)^4+12(x-3)^3+56(x-3)^2+120(x-3)+94$；

(2) $(x+1)^3-9(x+1)^2+22(x+1)-6$；

(3) $(x-1)^5+5(x-1)^4+10(x-1)^3+10(x-1)^2+5(x-1)+1$.

4. (1) $x^4+11x^3+42x^2+63x+37$；

(2) $x^5-5ix^4+(-9+3i)x^3+(8+7i)x^2+(3-7i)x-2+3i$.

习 题 6.3(1)

1. (1) $(f(x),g(x))=1$; (2) $(f(x),g(x))=x-1$;

(3) $(f(x),g(x))=x^2-\frac{\sqrt{2}}{2}x-1$;

(4) $(f(x),g(x))=x^2+x+1$.

2. 提示：应用关于最大公因式的引理.

习 题 6.3(2)

1. (1) $(f(x),g(x))=x^2-2$; $u(x)=1$, $v(x)=-1$.

(2) $(f(x),g(x))=1$; $u(x)=\frac{1}{3}x^2+\frac{1}{6}x$, $v(x)=-\frac{1}{3}x^3-\frac{1}{6}x^2+\frac{2}{3}x-\frac{1}{3}$. (3) $(f(x),g(x))=1$; $u(x)=\frac{7}{95}x+\frac{16}{95}$, $v(x)=-\frac{7}{95}x^3+\frac{1}{19}x^2-\frac{8}{95}x-\frac{79}{95}$. (满足 $u(x)f(x)+v(x)g(x)=(f(x),g(x))$的多项式 $u(x)$与 $v(x)$不是唯一的. 这里只举出一组 $u(x)$, $v(x)$. 如果读者的答案与这里的答案不一样，可以直接验证，看看是否适合$u(x)f(x)+v(x)g(x)=(f(x),g(x))$.)

2. 提示：应用习题 6.3(1)第 2 题，有$(f(x),g(x))=(f(x),f(x)-g(x))$.

3. 提示：应用定理 4 和定理 5.

4. 提示：应用定理 5.

习 题 6.3(3)

1. 提示：应用定理 6.

2. 提示：应用定理 6.

3. 提示：应用习题 6.3(1)第 2 题及上题的结果.

习 题 6.3(4)

1. (1) $(f_1(x),f_2(x),f_3(x))=x+1$; $u_1(x)=-\frac{3}{7}x-\frac{8}{7}$,

$u_2(x)=\frac{3}{7}x^2+\frac{11}{7}x+\frac{8}{7}$, $u_3(x)=-\frac{3}{7}x+\frac{1}{7}$;

(2) $(f_1(x),f_2(x),f_3(x))=x-1$; $u_1(x)=-\frac{3}{35}x^2-\frac{6}{35}x+\frac{1}{7}$,

$u_2(x)=0,u_3(x)=\frac{3}{35}x^3+\frac{18}{35}x^2+\frac{37}{35}x+\frac{5}{7}$.

2. 提示：证明$(f_1(x),f_2(x),f_3(x),f_4(x))$是$(f_1(x),f_2(x))$与$(f_3(x),f_4(x))$的公因式，而且$(f_1(x),f_2(x))$与$(f_3(x),f_4(x))$的任一个公因式都能够整除$(f_1(x),f_2(x),f_3(x),f_4(x))$.
3. 提示：对 s 作归纳法.
4. 提示：应用上题.
5. 提示：应用定理 8.

习　题　6.4(1)

1. 提示：应用定理 10.
2. 提示：用反证法.
3. 提示：用反证法.

习　题　6.4(2)

1. (1) $f(x)=-(x-1)^3(x^2+1)^2$；　(2) $g(x)=(x-1)^3(x+3)^2(x-3)$；

   (3) $h(x)=2(x-1)^2(x-3)(x^2+1)$.
2. $(f(x),g(x))=(x-1)^3$；$(f(x),h(x))=(x-1)^2(x^2+1)$；

   $(g(x),h(x))=(x-1)^2(x-3)$；$(f(x),g(x),h(x))=(x-1)^2$.

习　题　6.4(3)

1. $f(x)=x^3-3x^2-1$.
2. $f(x)=-\frac{4}{\pi^2}x^2+\frac{4}{\pi}x$.
3. $f(x)=x^4-4x^2+5x+1$.

习　题　6.5

1. (1)无；　(2) $f(x)$有 2 重因式 $x-2$；　(3) 无.
2. $t=3$ 时有 3 重根 1；$t=-\frac{15}{4}$时，有 2 重根$-\frac{1}{2}$.

3. $4p^3+27q^2=0$.

4. 提示：应用定理 13.

5. 提示：证明 $f(x)=x^n+ax^{n-m}+b$ 与 $f'(x),f''(x)$ 没有不等于零的公共根.

6.6 复系数与实系数多项式的因式分解

习　题　6.6(1)

1. $(x-1)^3(x+\mathrm{i})^2(x-\mathrm{i})^2$.

2. $2(x+1)^4(x-4)$.

3. $2\mathrm{i}(x-1)^2(x+1)\left(x+\dfrac{1+\sqrt{5}}{2}\right)\left(x+\dfrac{1-\sqrt{5}}{2}\right)\left(x-\dfrac{1}{2}\mathrm{i}\right)$.

4. $$(x-1)(x+1)(x-\mathrm{i})(x+\mathrm{i})\left(x-\frac{\sqrt{3}}{2}-\frac{1}{2}\mathrm{i}\right)\left(x-\frac{\sqrt{3}}{2}+\frac{1}{2}\mathrm{i}\right)$$
$$\cdot\left(x-\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i}\right)\left(x-\frac{1}{2}+\frac{\sqrt{3}}{2}\mathrm{i}\right)\left(x+\frac{1}{2}-\frac{\sqrt{3}}{2}\mathrm{i}\right)$$
$$\cdot\left(x+\frac{1}{2}+\frac{\sqrt{3}}{2}\mathrm{i}\right)\left(x+\frac{\sqrt{3}}{2}-\frac{1}{2}\mathrm{i}\right)\left(x+\frac{\sqrt{3}}{2}+\frac{1}{2}\mathrm{i}\right).$$

习　题　6.6(2)

1. (1) $f(x-1)^3(x^2+1)^2$；　(2) $2(x+1)^4(x-4)$；

(4) $(x-1)(x+1)(x^2+1)(x^2-\sqrt{3}x+1)(x^2-x+1)(x^2+x+1)$
$\cdot(x^2+\sqrt{3}x+1)$.

2. 提示：因为实系数多项式的虚根的共轭数也是这个多项式的根，而且重数相同.

3. $x(x-1)(x^2+1)^2(x^2-2x+2)$.

习　题　6.7

1. (1) $\dfrac{1}{6}(6x^3+3x^2-10x+8)$；　(2) $3(x^3+2x-4)$；

(3) $\dfrac{5}{6}(6x^3+3x^2-2x+4)$.

3. (1) 无有理根； (2) $-\frac{1}{2}$； (3) $\frac{1}{2}$及$-\frac{1}{3}$； (4) $\frac{1}{3}$,1,3,9.

4. (1) 不可约； (2) 可约； (3) 不可约.

5. (1) 实数域：$2(x+1)^2(x-1)\left(x^2+\frac{1}{2}x+\frac{1}{2}\right)$

复数域：$2(x+1)^2(x-1)\left(x+\frac{1+\sqrt{7}\,\mathrm{i}}{4}\right)\left(x+\frac{1-\sqrt{7}\,\mathrm{i}}{4}\right)$

(2) 实数域：$2(x-1)(x+3)\left(x+\frac{1}{2}\right)(x^2+1)$

复数域：$2(x-1)(x+3)\left(x+\frac{1}{2}\right)(x-\mathrm{i})(x+\mathrm{i})$

复习题 6

1. 提示：证明 x^2+x+1 的根都是 $x^{3n}+x^{3m+1}+x^{3l+2}$的根.

2. n 与 l 奇偶相同，而与 m 奇偶相反.

3. 证明：设 ω 是 x^2+x+1 的一个根，那么 ω^2 是 x^2+x+1 的另一个根，而且 $\omega^3=1,\omega\neq\omega^2$.

由假设： $x^2+x+1\mid f_1(x^3)+xf_2(x^3)$,

所以 x^2+x+1 的根都是 $f_1(x^3)+xf_2(x^3)$的根：

$$f_1(\omega^3)+\omega f_2(\omega^3)=0,\quad f_1(\omega^6)+\omega^2 f_2(\omega^6)=0.$$

即 $\qquad f_1(1)+\omega f_2(1)=0,\quad f_1(1)+\omega^2 f_2(1)=0,$

因此 $f_1(1)=f_2(1)=0$， 即 $x-1\mid f_1(x)$， $x-1\mid f_2(x)$.

4. 提示：因为 $ad-bc\neq0$，所以 $f(x),g(x)$也都是 $f_1(x),g_1(x)$的组合.

5. (1) $(f(x),g(x))=x^2-2$; $u(x)=-x-1$, $v(x)=x+2$.

(2) $(f(x),g(x))=x^3+1$; $u(x)=-1$, $v(x)=x+1$.

(3) $(f(x),g(x))=x-1$;

$$u(x)=-\frac{1}{3}x+\frac{1}{3},\ v(x)=\frac{2}{3}x^2-\frac{2}{3}x-1.$$

6. $t=1,u=10$，此时$(f(x),g(x))=x^2-2x+5$；

或 $t=-2,u=1$，此时$(f(x),g(x))=x^2+x-1$.

7. 提示：应用习题 6.3(3)第 2 题.

8. (1) 有：2 重因式 $x-2$； (2) 没有； (3) 有：3 重因式 $x-2$.

10. 当 $a=2$ 时有 2 重根 1；当 $a=-2$ 时有 2 重根-1.

11. (1) 不可约； (2) 可约； (3) 不可约.

12. $a=-2$,有理根为 2;或 $a=-6$,有理根为 -2;或 $a=4$,有理根为 $\frac{1}{2}$.

13. (1) 实数域：$2(x-1)^2(x+1)\left(x^2+x+\frac{1}{2}\right)$,

复数域：$2(x-1)^2(x+1)\left(x+\frac{1+i}{2}\right)\left(x+\frac{1-i}{2}\right)$;

(2) 实数域：$3(x+1)(x-2)\left(x+\frac{1}{3}\right)(x^2+2)$,

复数域：$3(x+1)(x-2)\left(x+\frac{1}{3}\right)(x+\sqrt{2}\,i)(x-\sqrt{2}\,i)$.

14. 提示：用反证法证明条件的充分性.

15. 提示：同上题.

16. 提示：计算 $g(x)$的导数 $g'(x),g''(x)$.

17. 提示：$f(x)$与$\frac{f(x)}{(f(x),f'(x))}$有相同的根.

18. 提示：用反证法.应用定理 3 推论 2.

20. (1) $x^6+3x^5+2x^4-3x^3-7x^2-6x-2$;

(2) $x^6+5x^5+7x^4+3x^3+5x^2+7x+2$;

(3) $x^6-\frac{25}{4}x^4+\frac{1}{4}x^3+\frac{87}{8}x^2-\frac{11}{8}x-\frac{9}{2}$.

第 7 章 λ-矩阵

习 题 7.1

1. (1) 2； (2) 3.

2. (1) $\begin{bmatrix}-4\lambda^2 & 4\lambda & 0\\ 0 & 0 & 0\\ 2\lambda^2 & -2\lambda & 0\end{bmatrix}$;

(2) $\begin{bmatrix}-\lambda^3-3\lambda^2 & \lambda^2 & \lambda^2\\ 2\lambda^3-\lambda^2 & -\lambda^3+\lambda^2 & \lambda^3-2\lambda^2\\ -\lambda^2+2\lambda+1 & \lambda^2-\lambda-1 & -\lambda^2+\lambda\end{bmatrix}$.

3. (1) 不可逆； (2) 不可逆；

(3) 可逆;逆矩阵为：

$$\frac{1}{2}\begin{bmatrix}\lambda^4-3\lambda^3-\lambda^2+4\lambda+8 & -\lambda^3+\lambda+2 & \lambda^3-2\lambda-2\\ -\lambda^3+3\lambda^2-\lambda & \lambda^2+1 & -\lambda^2\\ 2\lambda-6 & -2 & 2\end{bmatrix}.$$

4. 提示：如果 A 是 n 阶数字矩阵，则 $|\lambda E-A|$ 是一个 n 次多项式.

习　题　7.2

1. (1) $(\lambda-1)(\lambda+1)$；　(2) λ^2.

2. (1) 能对角化；　(2) 不能对角化.

3. 提示：参考命题 1 的证明.

4. 提示：应用定理 4.

习　题　7.3

1. $B(\lambda)=\begin{bmatrix}\lambda & \\ & \lambda^3-\frac{3}{2}\lambda^2+\frac{3}{2}\lambda\end{bmatrix}$，

$P(\lambda)=\begin{bmatrix}-1 & 1\\ \frac{1}{2}\lambda & -\frac{1}{2}\lambda+\frac{1}{2}\end{bmatrix}$，　$Q(\lambda)=\begin{bmatrix}1 & 2\lambda-3\\ 0 & 1\end{bmatrix}$；

($P(\lambda)$，$Q(\lambda)$ 不是唯一的)

2. $B(\lambda)=\begin{bmatrix}1 & & \\ & \lambda & \\ & & \lambda^2+\lambda\end{bmatrix}$，

$P(\lambda)=\begin{bmatrix}0 & -\lambda & 1\\ 0 & 1 & 0\\ 0 & \lambda+1 & -1\end{bmatrix}$，　$Q(\lambda)=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 1 & -1 & 1\end{bmatrix}$；

3. $B(\lambda)=\begin{bmatrix}1 & & \\ & \lambda(\lambda+1) & \\ & & \lambda(\lambda+1)^2\end{bmatrix}$，

$P(\lambda)=\begin{bmatrix}0 & 1 & 1\\ 1 & 0 & 0\\ 0 & (\lambda+1)^2 & \lambda^2+2\lambda\end{bmatrix}$，

$$Q(\lambda)=\begin{bmatrix}0 & 1 & 0\\ -\lambda-2 & 0 & (\lambda+1)^2\\ 1 & 0 & -\lambda\end{bmatrix};$$

4. $$B(\lambda)=\begin{bmatrix}1 & & & \\ & 1 & & \\ & & 0 & \\ & & & 0\end{bmatrix},$$

$$P(\lambda)=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ -\lambda+1 & 1 & 0 & 0\\ 0 & 0 & -\lambda+1 & 1\end{bmatrix},$$

$$Q(\lambda)=\begin{bmatrix}0 & 0 & 1 & 0\\ 1 & 0 & -\lambda-1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 1 & 0 & -\lambda-1\end{bmatrix}.$$

习　题　7.4

1. $D_1(\lambda)=D_2(\lambda)=1$，　$D_3(\lambda)=(\lambda-2)^3$；

$d_1(\lambda)=d_2(\lambda)=1$，　$d_3(\lambda)=(\lambda-2)^3$.

2. $D_1(\lambda)=D_2(\lambda)=D_3(\lambda)=1$，　$D_4(\lambda)=\lambda^4+2\lambda^3+3\lambda^2+4\lambda+5$；

$d_1(\lambda)=d_2(\lambda)=d_3(\lambda)=1$，　$d_4(\lambda)=\lambda^4+2\lambda^3+3\lambda^2+4\lambda+5$.

3. 当 $b\neq 0$ 时，

$D_1(\lambda)=D_2(\lambda)=D_3(\lambda)=1$，　$D_4(\lambda)=[(\lambda+a)^2+b^2]^2$；

$d_1(\lambda)=d_2(\lambda)=d_3(\lambda)=1$，　$d_4(\lambda)=[(\lambda+a)^2+b^2]^2$.

当 $b=0$ 时，

$D_1(\lambda)=D_2(\lambda)=1$，　$D_3(\lambda)=(\lambda+a)^2$，　$D_4(\lambda)=(\lambda+a)^4$；

$d_1(\lambda)=d_2(\lambda)=1$，　$d_3(\lambda)=(\lambda+a)^2$，　$d_4(\lambda)=(\lambda+a)^2$.

习　题　7.5

1. 初等因子：λ^4，

标准形：$\begin{bmatrix}1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \lambda^4\end{bmatrix}$.

2. 初等因子：$\lambda,\lambda,\lambda^2,\lambda-1,\lambda-1,(\lambda-1)^2$

标准形：$\begin{bmatrix}1 & & & \\ & \lambda(\lambda-1) & & \\ & & \lambda(\lambda-1) & \\ & & & \lambda^2(\lambda-1)^2\end{bmatrix}$.

3. 初等因子：有理数域 λ^2-2,λ^2+1；

实数域 $\lambda-\sqrt{2}$，$\lambda+\sqrt{2}$，λ^2+1；

复数域 $\lambda-\sqrt{2},\lambda+\sqrt{2},\lambda-i,\lambda+i$.

标准形：$\begin{bmatrix}1 & & \\ & 1 & \\ & & (\lambda^2+1)(\lambda^2-2)\end{bmatrix}$.

习 题 7.6

2. 能. 3. 不能. 4. 不能.

习 题 7.7

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

4. $\begin{bmatrix}2 & 1 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 4 & 1\\ 0 & 0 & 0 & 4\end{bmatrix}$.

复习题7

1. (1) $\begin{bmatrix}1 & & \\ & \lambda-1 & \\ & & (\lambda-1)^2(\lambda+1)\end{bmatrix}$;

(2) $\begin{bmatrix}1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & \lambda^2(\lambda-1)(\lambda-12)\end{bmatrix}$.

2. (1) 不变因子：$1,1,1,(\lambda+2)^4$，

初等因子：$(\lambda+2)^4$，

秩：4；

(2) 不变因子：$1,1,(\lambda+1)(\lambda+2),(\lambda+1)(\lambda+2)$，

初等因子：$\lambda+1,\lambda+1,\lambda+2,\lambda+2$，

秩：4；

(3) 不变因子：$1,\ \lambda^2-\lambda-1$，

初等因子：实数域：$\lambda^2-\lambda-1$，

复数域：$\left(\lambda-\dfrac{1+\sqrt{3}\,\mathrm{i}}{2}\right),\left(\lambda-\dfrac{1-\sqrt{3}\,\mathrm{i}}{2}\right)$，

秩：2.

3. (1) $\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$;　(2) $\begin{bmatrix}2 & 1 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{bmatrix}$;　(3) $\begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{bmatrix}$;

(4) $\begin{bmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 2\end{bmatrix}$;　(5) $\begin{bmatrix}1 & 0 & 0\\ 0 & \mathrm{i} & 0\\ 0 & 0 & -\mathrm{i}\end{bmatrix}$;　(6) $\begin{bmatrix}1 & 1 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 1\end{bmatrix}$.

4. 提示：幂零矩阵的特征值为 0.

5. 提示：幂等矩阵能对角化，并且特征值为 1 或 0.

6. 提示：A 可对角化且其特征值为 1 或 -1.

7. 提示：应用 A 的约当标准形.

8. 提示：$f(x)$ 与 $g(x)$ 有相同的不可约因式.

第8章 线性空间

习 题 8.1

1. (1) 是(这个线性空间用 $\mathbf{P}_n[x]$表示).

   (2) $n>0$ 时,不是;$n=0$ 时,是.

   (3) 是. (4) 是. (5) 是.

2. $\frac{32}{49}$. 3. $(-2,2)$. 6. 提示:$\beta-\gamma$ 是解.

习 题 8.2

1. $\beta_1=2\circ\alpha_1\oplus(-1)\circ\alpha_2$;

   $\beta_2=\alpha_1\oplus\alpha_2$.

5. (2) $\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}=a_{11}E_{11}+a_{22}E_{22}+\frac{a_{12}+a_{21}}{2}(E_{12}+E_{21})+\frac{a_{12}-a_{21}}{2}(E_{12}-E_{21})$.

6. 提示:证明 $\beta_1,\beta_2,\cdots,\beta_s$ 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 线性表出.

7. $\alpha_1,\alpha_3,\alpha_4$.

习 题 8.3

2. (1) 维数:1,基:2;

   (2) 维数:2,基:(1,0),(0,1).

3. (1) 维数:$\frac{n(n+1)}{2}$,

   基:$\{E_{ii},E_{ij}+E_{ji}\,|\,i,j=1,2,\cdots,n;\ j>i\}$;

   (2) 维数:$\frac{n(n-1)}{2}$.

   基:$\{E_{ij}-E_{ji}\,|\,i,j=1,2,\cdots,n;\ j>i\}$;

   (3) 维数:$\frac{n(n+1)}{2}$,

   基:$\{E_{ij}\,|\,i,j=1,2,\cdots,n;\ j\geqslant i\}$.

4. (1) $\left(0,\frac{3}{2},-\frac{1}{2},0\right)$;

   (2) $\lg_2 3$;

(3) $\left(a, b-\frac{a(a-1)}{2}\right)$;

(4) $(a_0-a_1, a_1-a_2, \cdots, a_{n-2}-a_{n-1}, a_{n-1})$.

习 题 8.4

1. (1) $\begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 2 & 1 \\ 1 & 2 & 1 & 2 \end{bmatrix}$, 坐标：$(3,-1,2,-1)$;

(2) $\frac{1}{13}\begin{bmatrix} 5 & 6 & 4 & -15 \\ 17 & 10 & 24 & 14 \\ 1 & 9 & 6 & -3 \\ -5 & 7 & 9 & -11 \end{bmatrix}$, 坐标：$\left(\frac{6}{13},\frac{10}{13},\frac{9}{13},\frac{7}{13}\right)$;

(3) $\frac{1}{4}\begin{bmatrix} 7 & 4 & -1 & 2 \\ -1 & 2 & 1 & 4 \\ 1 & 0 & 3 & 4 \\ -3 & 2 & 1 & -2 \end{bmatrix}$, 坐标：$\left(-\frac{1}{8}a_2+\frac{3}{8}a_3+\frac{1}{8}a_4, \frac{1}{3}a_1\right.$

$\left.+\frac{1}{12}a_2-\frac{1}{4}a_3+\frac{1}{4}a_4, \frac{1}{3}a_1-\frac{19}{24}a_2+\frac{3}{8}a_3+\frac{1}{8}a_4, \frac{3}{8}a_2-\frac{1}{8}a_3-\frac{3}{8}a_4\right)$.

2. (2) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$; (3) $\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$;

(4) $(a_0-a_1+a_2-a_3, a_1-2a_2+3a_3, a_2-3a_3, a_3)$.

3. (1) (x_2, x_3, x_1); (2) $\left(x_1, \frac{1}{k}x_2, x_3\right)$; (3) (x_1, x_2, x_3-kx_2).

4. (2) $\left(a_{11}, a_{22}, a_{33}, \frac{a_{12}+a_{21}}{2}, \frac{a_{13}+a_{31}}{2}, \frac{a_{23}+a_{32}}{2}, \frac{a_{12}-a_{21}}{2}, \frac{a_{13}-a_{31}}{2},\right.$

$\left.\frac{a_{23}-a_{32}}{2}\right)$

习 题 8.5

1. (1) 是;维数：3；基：(1,0,0,1),(0,1,0,1),(0,0,1,1).

(2) 不是.

(3) 是;维数: $\frac{n(n-1)}{2}$;基: $\{E_{ij}-E_{ji} \mid i,j=1,2,\cdots,n; i<j\}$.

(4) 不是.

(5) $n=0$ 时,是;维数: 1;基:1.

$n>0$ 时,不是.

2. (1) 维数: 2; 基: α_1,α_2.

(2) 维数: 3; 基: $\alpha_1,\alpha_2,\alpha_3$.

(3) 维数: 3; 基: $\alpha_1,\alpha_2,\alpha_4$.

(4) 维数: 4; 基: $\alpha_1,\alpha_2,\alpha_3,\alpha_4$.

3. 维数: 2;基: $(1,-24,-9,0),(2,-21,0,9)$.

习　题　8.6(1)

1. (1) $W_1+W_2=L(\alpha_1,\alpha_2,\beta_1,\beta_2)$;维数:4;　基: $\alpha_1,\alpha_2,\beta_1,\beta_2$.

$W_1\cap W_2=\{0\}$;维数:0;　没有基.

(2) $W_1+W_2=L(\alpha_1,\alpha_2,\beta_2)$;维数:3;　基: $\alpha_1,\alpha_2,\beta_2$.

$W_1\cap W_2=L(\beta_1)$;维数:1;　基: β_1.

(3) $W_1+W_2=L(\alpha_1,\alpha_2,\beta_1,\beta_2)$;维数:4;　基: $\alpha_1,\alpha_2,\beta_1,\beta_2$.

$W_1\cap W_2=L(\alpha)$, $\alpha=(53,119,-19,-134)$;

维数:1;　基: α.

2. 提示:证明向量 α_1,α_3 与 α_2,α_3 等价.

习　题　8.6(2)

1. 可令 $\beta_1=(0,0,1,0)$, $\beta_2=(0,0,0,1)$

$W=L(\beta_1,\beta_2)$,　(答案不唯一)

2. $W_1=L(\alpha_1,\alpha_2,\cdots,\alpha_{n-1})$,其中 $\alpha_1=(1,-1,0,\cdots,0,0)$, $\alpha_2=(0,1,-1,\cdots,0,0)$, $\alpha_{n-1}=(0,0,0,\cdots,1,-1)$;

$W_2=L(\beta)$,其中 $\beta-(1,1,\cdots,1)$.

3. 提示:证明 $V=W_1+W_2+\cdots+W_s$ 且 $\dim V=\dim W_1+\dim W_2+\cdots+\dim W_s$

习 题 8.7

1. σ_1: $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \to \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$;

σ_2: $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \to \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & 2a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ & & & na_{nn} \end{bmatrix}$.

3. 提示：把 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 及 $\beta_1,\beta_2,\cdots,\beta_t$ 分别扩充成 V_1 及 V_2 的基.

4. 提示：应用第 3 题的方法.例如

$(a_1,a_2,a_3,a_4) \to (2a_1 - a_2 + a_3, 3a_1 - a_2 + a_4, -2a_1 + a_2, -a_1 + a_2)$

就是一个满足条件的自同构.

复习题 8

4. (1) 提示：只要证明 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 及 $\beta_1,\beta_2,\beta_3,\beta_4$ 都线性无关；

(2) $\frac{1}{4}\begin{bmatrix} -24 & 36 & -12 & 32 \\ 3 & -2 & 4 & -5 \\ -13 & 22 & -8 & 19 \\ 19 & -30 & 12 & -25 \end{bmatrix}$;

(3) $(8,-2,6,-7)$.

5. (2) $W=L(\alpha)$, $\alpha=(2,-3,-2,0)$；维数：1；基：α.

7. (1) 解空间 $W_1=L(\alpha_1,\alpha_2,\alpha_3)$，其中 $\alpha_1=(3,1,0,0,0)$，$\alpha_2=(2,0,-1,0,0)$，$\alpha_3=(0,0,0,1,-1)$；维数：3；基：$\alpha_1,\alpha_2,\alpha_3$.

(2) 解空间 $W_2=L(\beta_1,\beta_2,\beta_3)$，其中 $\beta_1=(1,7,4,0,0)$，$\beta_2=(3,-7,0,-4,0)$，$\beta_3=(1,0,0,0,-1)$；维数：3；基：β_1,β_2,β_3.

8. $W_1+W_2=L(\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2)$；维数：5；基：$\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2$.

$W_1\cap W_2=L(\alpha)$，$\alpha=(1,7,10,-6,6)$. 维数：1；基：α.

9. $W_1+W_2=L(\alpha_1,\alpha_2,\beta_1)$；维数：3；基：$\alpha_1,\alpha_2,\beta_1$.

$W_1\cap W_2=L(\alpha)$，$\alpha=(5,-2,-3,-4,1)$；维数：1；基：α.

11. $F(A)=L(E,A)$；维数：2；基：E,A.

13. $C(A)=\left\{\begin{bmatrix} x_1 & 0 & z_1 \\ x_2 & y_2 & z_2 \\ 0 & 0 & y_2 \end{bmatrix} \middle| x_1,x_2,y_2,z_1,z_2 \text{ 为数域 } \mathbf{P} \text{ 中任意数}\right\}$;

维数:5;基: $E_{11},E_{13},E_{21},E_{23},E_{22}+E_{33}$.

15. $\ker(A)=\left\{\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ 0 & 0 & 0 \end{bmatrix} \middle| x_1,x_2,x_3,y_1,y_2,y_3 \text{ 为 } \mathbf{P} \text{ 中任意数}\right\}$;维数:6;基: $E_{11},E_{12},E_{13},E_{21},E_{22},E_{33}$.

16. (1) $F(A)\cap C(A)=F(A)$, $F(A)+C(A)=C(A)$.

(2) $F(A)\cap\ker(A)=L(A)$;维数:1;基:A.

$F(A)+\ker(A)$

$=\left\{\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ 0 & 0 & z_3 \end{bmatrix} \middle| x_1,y_1,z_1,x_2,y_2,z_2,z_3 \text{ 是 } \mathbf{P} \text{ 中任意数}\right\}$;

维数: 7;基: $E_{11},E_{12},E_{13},E_{21},E_{22},E_{23},E_{33}$.

(3) $\ker(A)\cap C(A)$

$=\left\{\begin{bmatrix} x_1 & 0 & z_1 \\ x_2 & 0 & z_2 \\ 0 & 0 & 0 \end{bmatrix} \middle| x_1,x_2,z_1,z_2 \text{ 是 } \mathbf{P} \text{ 中任意数}\right\}$; 维 数: 4; 基: $E_{11},E_{13},E_{21},E_{23}$.

$\ker(A)+C(A)$

$=\left\{\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ 0 & 0 & z_3 \end{bmatrix} \middle| x_1,y_1,z_1,x_2,y_2,z_2,z_3 \text{ 是 } \mathbf{P} \text{ 中任意数}\right\}$;

维数:7;基: $E_{11},E_{12},E_{13},E_{21},E_{22},E_{23},E_{33}$.

18. $(a,b)\to\left(a,b-\dfrac{a(a-1)}{2}\right)$.

第9章 线性变换

习 题 9.1

1. (1) 当 $\alpha_0=0$ 时,是;当 $\alpha_0\neq 0$ 时,不是.

(2) 同(1). (3) 不是. (4) 是.

2. 提示：应用命题 2.

4. $\boldsymbol{A}(a_1,a_2,a_3)=(a_1,-a_1+2a_2-a_3,2a_1-6a_2+3a_3)$.

5. $(a_1,a_2,a_3)\to(a_1+a_2+a_3,2a_1-a_2,3a_1)$及$(a_1,a_2,a_3)\to(a_1+a_2,2a_1-a_2,3a_1)$.

习 题 9.2

1. (1) $\boldsymbol{A}+\boldsymbol{B}(x_1,x_2,x_3)=(2x_1+x_2-x_3,x_2,0)$;

$\boldsymbol{A}-\boldsymbol{B}(x_1,x_2,x_3)=(-x_2+x_3,x_2,2x_1+2x_2)$;

$2\boldsymbol{A}(x_1,x_2,x_3)=(2x_1,2x_2,2x_1+2x_2)$.

(2) $\boldsymbol{AB}(x_1,x_2,x_3)=(x_1+x_2-x_3,0,x_1+x_2-x_3)$;

$\boldsymbol{BA}(x_1,x_2,x_3)=(0,0,-x_1-x_2)$;

$\boldsymbol{A}^2(x_1,x_2,x_3)=(x_1,x_2,x_1+x_2)$.

2. $\boldsymbol{A}^{-1}(x_1,x_2,x_3)=(x_1-x_2,x_2-x_3,x_3)$.

习 题 9.3

1. $\boldsymbol{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为

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

$\boldsymbol{A}$ 在基 η_1,η_2,η_3 下的矩阵为

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

2. $\boldsymbol{B}$ 在基 ξ_1,ξ_2,ξ_3 下的矩阵为

$$\begin{bmatrix}2&3&1\frac{1}{7}\\-1&0&-2\frac{2}{7}\\-1&1&-1\frac{2}{7}\end{bmatrix};$$

$\boldsymbol{B}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为

$$\frac{1}{7}\begin{bmatrix} -5 & 20 & -20 \\ -4 & -5 & -2 \\ 9 & 27 & 15 \end{bmatrix}.$$

3. (2) $\boldsymbol{A}^{-1}$,在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 下的矩阵为

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix};$$

$\boldsymbol{AB}$ 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 下的矩阵为

$$\frac{1}{7}\begin{bmatrix} -6 & 45 & -38 \\ -13 & -32 & -17 \\ 5 & 22 & 13 \end{bmatrix};$$

$\boldsymbol{A}+\boldsymbol{B}$ 在基 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 下的矩阵为

$$\frac{1}{7}\begin{bmatrix} 9 & 13 & -20 \\ -4 & 2 & -9 \\ 9 & 34 & 22 \end{bmatrix}.$$

4. 在基 $E_{11},E_{12},E_{21},E_{22}$下,$\boldsymbol{A}$ 的矩阵是

$$\begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix};$$

$\boldsymbol{B}$ 的矩阵是

$$\begin{bmatrix} a & c & 0 & 0 \\ b & d & 0 & 0 \\ 0 & 0 & a & c \\ 0 & 0 & b & d \end{bmatrix};$$

$\boldsymbol{AB}$ 的矩阵是

$$\begin{bmatrix} a^2 & ac & ab & bc \\ ab & ad & b^2 & bd \\ ac & c^2 & ad & cd \\ bc & cd & bd & d^2 \end{bmatrix}.$$

5. (2) $\boldsymbol{A}^{-1}(X)=\dfrac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}X,\quad X\in P^{2\times 2}.$

$\boldsymbol{A}^{-1}$在基 $E_{11},E_{12},E_{21},E_{22}$下的矩阵为

$$\frac{1}{ad-bc}\begin{bmatrix} d & 0 & -b & 0 \\ 0 & d & 0 & -b \\ -c & 0 & a & 0 \\ 0 & -c & 0 & d \end{bmatrix}.$$

6. (1) $\begin{bmatrix} a_{11} & a_{13} & a_{12} \\ a_{31} & a_{33} & a_{32} \\ a_{21} & a_{23} & a_{22} \end{bmatrix}$;

(2) $\begin{bmatrix} a_{11} & ka_{12} & a_{13} \\ k^{-1}a_{21} & a_{22} & k^{-1}a_{23} \\ a_{31} & ka_{32} & a_{33} \end{bmatrix}$;

(3) $\begin{bmatrix} a_{11}-a_{21} & a_{11}+a_{12}-a_{21}-a_{22} & a_{13}-a_{23} \\ a_{21} & a_{21}+a_{22} & a_{23} \\ a_{31} & a_{31}+a_{32} & a_{33} \end{bmatrix}$.

习　题　9.4

1. $\boldsymbol{A}$ 的特征值:2(3 重)

特征向量:$k(1,3,0)+l(0,1,1)$,k 和 l 是不全为零的任意数.

2. (1) $\boldsymbol{A}$ 的特征值:2(3 重),-2;

属于 2 的特征向量: $k_1(\varepsilon_1+\varepsilon_2)+k_2(\varepsilon_1+\varepsilon_3)+k_3(\varepsilon_1+\varepsilon_4)$, k_1,k_2, k_3 是不全为零的任意数;

属于-2 的特征向量:$k(\varepsilon_1-\varepsilon_2-\varepsilon_3-\varepsilon_4)$,$k$ 是任意非零数.

(2) $\boldsymbol{A}$ 在基 $\varepsilon_1+\varepsilon_2,\varepsilon_1+\varepsilon_3,\varepsilon_1+\varepsilon_4,\varepsilon_1-\varepsilon_2-\varepsilon_3-\varepsilon_4$ 下的矩阵为对角矩阵

$$\begin{bmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & -2 \end{bmatrix}.$$

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

(2) 特征值：$0,0,\frac{1}{2},1$；

属于 0 的特征向量：$k_1(\varepsilon_1+2\varepsilon_2+\varepsilon_3+\varepsilon_4)+k_2(2\varepsilon_1+3\varepsilon_2+\varepsilon_3)$，$k_1,k_2$ 是不全为零的任意数；

属于 1/2 的特征向量：$k_3(4\varepsilon_1+2\varepsilon_2-\varepsilon_3-6\varepsilon_4)$，$k_3\neq 0$，任意；

属于 1 的特征向量：$k_4(3\varepsilon_1+\varepsilon_2+\varepsilon_3-2\varepsilon_4)$，$k_4\neq 0$，任意.

(3) $\boldsymbol{A}$ 在基 $\varepsilon_1+2\varepsilon_2+\varepsilon_3-\varepsilon_4, 2\varepsilon_1+3\varepsilon_2+\varepsilon_3, 4\varepsilon_1+2\varepsilon_2-\varepsilon_3-6\varepsilon_4, 3\varepsilon_1+\varepsilon_2+\varepsilon_3-2\varepsilon_4$ 下的矩阵为对角矩阵

$$\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \frac{1}{2} & \\ & & & 1 \end{bmatrix}.$$

习 题 9.5

2. $L(\varepsilon_1,\varepsilon_3,\varepsilon_4)$.

5. (1) $\boldsymbol{A}V=L(\varepsilon_1-\varepsilon_2+\varepsilon_3+2\varepsilon_4,\varepsilon_2+\varepsilon_3-\varepsilon_4)$；维数：2；基：$\varepsilon_1-\varepsilon_2+\varepsilon_3+2\varepsilon_4$，$\varepsilon_2+\varepsilon_3-\varepsilon_4$.

$\ker\boldsymbol{A}=L(4\varepsilon_1+3\varepsilon_2-2\varepsilon_3,\varepsilon_1+2\varepsilon_2-\varepsilon_4)$；维数：2；基：$4\varepsilon_1+3\varepsilon_2-2\varepsilon_3$，$\varepsilon_1+2\varepsilon_2-\varepsilon_4$.

(2) $\boldsymbol{A}V+\ker\boldsymbol{A}=V$；维数：$n$；基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$.

$\boldsymbol{A}V\cap\ker\boldsymbol{A}=\{\boldsymbol{0}\}$；维数：0；没有基.

8. 提示：应用维数公式，证明 $\ker\boldsymbol{A}$，$\ker\boldsymbol{B}$ 中有公共非零向量.

复习题 9

1. (1) $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ -2 & 2 & 1 \end{bmatrix}$;

(2) $(1,0,-2),(-1,1,2),(1,-1,1)$ （解答不唯一）.

2. (1) 矩阵：$\begin{bmatrix} 1 & 4 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & 7 & -4 \\ 0 & 0 & -2 & 5 \end{bmatrix}$;

特征值：$3,3,-3,9$；

属于 3 的特征向量：$2k_1\boldsymbol{\varepsilon}_1+k_1\boldsymbol{\varepsilon}_2+k_2\boldsymbol{\varepsilon}_3+k_2\boldsymbol{\varepsilon}_4$ （k_1,k_2 不全为 0，任意数）；

属于 -3 的特征向量：$k\boldsymbol{\varepsilon}_1-k\boldsymbol{\varepsilon}_2$ （k 为任意非零常数）；

属于 9 的特征向量：$2l\boldsymbol{\varepsilon}_3-l\boldsymbol{\varepsilon}_4$ （l 为任意非零常数）.

(2) $\boldsymbol{A}$ 在基 $2\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3+\boldsymbol{\varepsilon}_4,\boldsymbol{\varepsilon}_1-\boldsymbol{\varepsilon}_2,2\boldsymbol{\varepsilon}_3-\boldsymbol{\varepsilon}_4$ 下的矩阵为对角矩阵

$$\begin{bmatrix} 3 & & & \\ & 3 & & \\ & & -3 & \\ & & & 9 \end{bmatrix}.$$

3. (1) $\boldsymbol{A}V=L(5\boldsymbol{\varepsilon}_1-\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_4,-2\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_3)$；维数：2；一组基：$5\boldsymbol{\varepsilon}_1-\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_4$，$2\boldsymbol{\varepsilon}_1-\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_3$.

$\ker\boldsymbol{A}=L(\boldsymbol{\varepsilon}_1+2\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_1+3\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_4)$；维数：2；一组基：$\boldsymbol{\varepsilon}_1+2\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_3$，$\boldsymbol{\varepsilon}_1+3\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_4$.

(2) $r(\boldsymbol{A})=2,\mathrm{nul}(\boldsymbol{A})=2$.

(3) $\boldsymbol{A}V+\ker\boldsymbol{A}=L(5\boldsymbol{\varepsilon}_1-2\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_4,2\boldsymbol{\varepsilon}_1-\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_1+2\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_3)$；维数：3；一组基：$5\boldsymbol{\varepsilon}_1-2\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_4,2\boldsymbol{\varepsilon}_1-\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_1+2\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_3$.

$\boldsymbol{A}V\cap\ker\boldsymbol{A}=L(\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2-2\boldsymbol{\varepsilon}_3-\boldsymbol{\varepsilon}_4)$；维数：1；基：$\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2-2\boldsymbol{\varepsilon}_3-\boldsymbol{\varepsilon}_4$.

5. 提示：取基 $\boldsymbol{A}^{n-1}\xi,\boldsymbol{A}^{n-2}\xi,\cdots,\boldsymbol{A}\xi,\xi$.

8. 提示：证明所有 V 中非零向量作为 $\boldsymbol{A}$ 的特征向量都有相同的特征值.

9. (1) $\boldsymbol{E}_{ij}$.

10. 提示：$\boldsymbol{E},\boldsymbol{A},\boldsymbol{A}^2,\cdots,\boldsymbol{A}^{n^2}$是线性相关的.

第10章　欧氏空间

习　题　10.1

1. (1) $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$, $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\frac{\pi}{2}$；

   (2) $(\boldsymbol{\alpha},\boldsymbol{\beta})=6$, $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\cos^{-1}\frac{\sqrt{30}}{15}$；

   (3) $(\boldsymbol{\alpha},\boldsymbol{\beta})=3$, $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\frac{\pi}{6}$.

4. $\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)$或$\left(-\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right)$.

5. (1) 提示：证明$(\boldsymbol{\gamma},\boldsymbol{\gamma})=0$.

   (2) 提示：应用(1).

习　题　10.2

2. (1) $\begin{bmatrix}1&1&1&1\\1&2&2&2\\1&2&3&3\\1&2&3&4\end{bmatrix}$；　(2) E.

3. (1) $\begin{bmatrix}2&1&0&-1\\1&2&-1&0\\0&-1&2&1\\-1&0&1&3\end{bmatrix}$；　(2) $3\frac{1}{3}$.

   (3) $\pm\left(\frac{11\sqrt{43}}{86},-\frac{3\sqrt{43}}{43},-\frac{\sqrt{43}}{86},0\right)$；　(4) $\begin{bmatrix}9&1&9&0\\1&1&3&0\\9&3&35&0\\0&0&0&1\end{bmatrix}$.

4. (1,0,0,0),(1,1,0,0),(0,1,1,0),(1,−2,1,2).(答案不唯一)

习　题　10.3

3. (2) $a=-1$；

(3) $\alpha_2=\frac{\sqrt{5}}{5}\varepsilon_1+\frac{\sqrt{5}}{5}\varepsilon_2-\frac{\sqrt{5}}{5}\varepsilon_3$;

(4) 令$\alpha_3=\frac{7\sqrt{5}}{5}\varepsilon_1+\frac{2\sqrt{5}}{5}\varepsilon_2-\frac{2\sqrt{5}}{5}\varepsilon_3$

$$\left(\text{或}-\frac{7\sqrt{5}}{5}\varepsilon_1-\frac{2\sqrt{5}}{5}\varepsilon_2+\frac{2\sqrt{5}}{5}\varepsilon_3\right),$$

则 $\alpha_1,\alpha_2,\alpha_3$ 是 V 的一组标准正交基.

4. $\varepsilon_1=\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)$, $\varepsilon_2=\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right)$,

$\varepsilon_3=\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right)$, $\varepsilon_4=\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)$.

习 题 10.4

1. (1) $\frac{\sqrt{2}}{2}\varepsilon_1+\frac{\sqrt{2}}{2}\varepsilon_2$, $\frac{1}{2}\varepsilon_1-\frac{1}{2}\varepsilon_2-\frac{1}{2}\varepsilon_3+\frac{1}{2}\varepsilon_4$;

(2) $\frac{\sqrt{6}}{6}\varepsilon_1+\frac{\sqrt{6}}{3}\varepsilon_2-\frac{\sqrt{6}}{6}\varepsilon_3$, $-\frac{\sqrt{3}}{6}\varepsilon_1+\frac{\sqrt{3}}{6}\varepsilon_2+\frac{\sqrt{3}}{6}\varepsilon_3+\frac{\sqrt{3}}{2}\varepsilon_4$.

2. $\left(0,\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},0\right)$, $\left(\frac{\sqrt{10}}{5},-\frac{\sqrt{10}}{10},\frac{\sqrt{10}}{10},-\frac{\sqrt{10}}{5}\right)$;

$W^{\perp}=L((2,1,-1,1),(1,1,-1,0))$.

3. $(2,-3,1,-2)$.

习 题 10.5

2. 提示:证明充分性时应用 10.2 节第(2)式.

4. 提示:应用上题.

习 题 10.6

2. (1) $\boldsymbol{A}(a_1,a_2,a_3,a_4)=\frac{\sqrt{2}}{2}(a_1+a_3,a_2+a_4,a_2-a_4,a_1-a_3)$;

(2) $\begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & 0 \end{bmatrix}$;

(3) 共有 2 个，另一个为 $\boldsymbol{B}$：

$$\boldsymbol{B}(a_1,a_2,a_3,a_4)=\frac{\sqrt{2}}{2}(a_1+a_4,a_2+a_3,a_2-a_3,a_1-a_4).$$

3. 提示：任取 $\alpha\in W^{\perp}$，证明对 $\beta\in W$，都有 $(\boldsymbol{A}\alpha,\beta)=0$.

6. 提示：同第 3 题.

7. (2) $\frac{\sqrt{2}}{2}\varepsilon_1+\frac{\sqrt{2}}{2}\varepsilon_2,\frac{\sqrt{6}}{6}\varepsilon_1-\frac{\sqrt{6}}{6}\varepsilon_2+\frac{\sqrt{6}}{3}\varepsilon_3,$

$\frac{\sqrt{3}}{6}\varepsilon_1-\frac{\sqrt{3}}{6}\varepsilon_2-\frac{\sqrt{3}}{6}\varepsilon_3-\frac{\sqrt{3}}{2}\varepsilon_4,\frac{1}{2}\varepsilon_1-\frac{1}{2}\varepsilon_2-\frac{1}{2}\varepsilon_3+\frac{1}{2}\varepsilon_4.$

习 题 10.7

1. $x_1=0.35,x_2=0.24$.

复习题 10

1. (1) 1;(2) $-\frac{5}{2}$;(3) $\frac{1}{4}$.

5. (1) $\alpha_1,-\alpha_1+\alpha_2,\alpha_1+\alpha_3,\sqrt{2}\alpha_1-\frac{\sqrt{2}}{2}\alpha_2+\frac{\sqrt{2}}{2}\alpha_3+\frac{\sqrt{2}}{2}\alpha_4$;

(2) $\begin{bmatrix} 9 & -1 & 4 & 11 \\ -1 & 3 & 2 & 1 \\ 4 & 2 & 6 & 4 \\ 11 & 1 & 4 & 21 \end{bmatrix}$.

6. (1) $\frac{\sqrt{3}}{3}(1,-1,1,0),\frac{\sqrt{3}}{3}(0,1,1,-1),\frac{\sqrt{30}}{30}(2,3,1,4)$;

(2) $\frac{\sqrt{30}}{30}(4,1,-3,-2)$.

7. (1) $\frac{\sqrt{7}}{7}(\varepsilon_1+\varepsilon_2-2\varepsilon_3+\varepsilon_4)$, [illegible] $(\varepsilon_1-\varepsilon_3-\varepsilon_4)$;

(2) $\frac{\sqrt{2}}{2}(\varepsilon_1+\varepsilon_2)$, $\frac{\sqrt{42}}{42}(\varepsilon_1-6\varepsilon_2-\varepsilon_3+2\varepsilon_4)$;

(3) $3\varepsilon_1+\varepsilon_2-3\varepsilon_3$.

13. 提示:证明$(A(\alpha+\beta)-A\alpha-A\beta, A(\alpha+\beta)-A\alpha-A\beta)=0$,

$$(A(k\alpha)-kA\alpha, A(k\alpha)-kA\alpha)=0.$$

18. (2) $\frac{\sqrt{2}}{2}\varepsilon_1+\frac{\sqrt{2}}{2}\varepsilon_2$, $\frac{\sqrt{6}}{6}\varepsilon_1-\frac{\sqrt{6}}{6}\varepsilon_2+\frac{\sqrt{6}}{3}\varepsilon_3$,

$\frac{\sqrt{3}}{6}\varepsilon_1-\frac{\sqrt{3}}{6}\varepsilon_2-\frac{\sqrt{3}}{6}\varepsilon_3+\frac{\sqrt{3}}{2}\varepsilon_4$,

$\frac{1}{2}\varepsilon_1-\frac{1}{2}\varepsilon_2-\frac{1}{2}\varepsilon_3-\frac{1}{2}\varepsilon_4$.