高等代数笔记

讲课：谢启鸿
记录：Cowboy

更新：2025 年 3 月 5 日 18:11

目　录

1 行列式

1.1 行列式的展开和转置

定理 1.1 若

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \tag{1.1}$$

第 i 行第 j 列元素 a_{ij} 的代数余子式记为 A_{ij}, 则对任意的 $r(r=1,2,\cdots,n)$ 列有展开式:

$$|A| = a_{1r}A_{1r} + a_{2r}A_{2r} + \cdots + a_{nr}A_{nr}. \tag{1.2}$$

又对任意的 $s \neq r$, 有

$$a_{1r}A_{1s} + a_{2r}A_{2s} + \cdots + a_{nr}A_{ns} = 0 \tag{1.3}$$

证明:

$$\begin{aligned} |A| = &\begin{vmatrix} a_{11} & \cdots & a_{1r} & \cdots & a_{1n} \\ a_{21} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2r} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} \\ & + \cdots + \begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nr} & \cdots & a_{nn} \end{vmatrix} \\ = & a_{1r}A_{1r} + a_{2r}A_{2r} + \cdots + a_{nr}A_{nr}. \end{aligned}$$

所以，(1.2)式成立。再在(1.1)式基础上，将其第 s 列换成第 r 列，构建一个新行列式，其值为 0, 即：

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} & \cdots & a_{1r} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2r} & \cdots & a_{2r} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nr} & \cdots & a_{nr} & \cdots & a_{nn} \end{vmatrix} = 0$$

将上面行列式按第 s 列展开：

$$a_{1r}A_{1s} + a_{2r}A_{2s} + \cdots + a_{nr}A_{ns} = 0$$

所以(1.3)式成立。

同理可证，按照任意 $r(r=1,2,\cdots,n)$ 行展开有展开式：

$$|A| = a_{r1}A_{1r} + a_{r2}A_{r2} + \cdots + a_{rn}A_{rn}. \tag{1.4}$$

又对任意的 $s \neq r$, 有

$$a_{r1}A_{s1} + a_{r2}A_{s2} + \cdots + a_{rn}A_{sn} = 0 \tag{1.5}$$

□

1.2 Gramer 法则

定理 1.2 (Gramer 法则) 设有方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad \cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n, \end{cases} \tag{1.6}$$

记这个方程组的系数行列式为 $|A|$, 若 $|A| \neq 0$, 则方程组有且仅有一组解:

$$x_1 = \frac{|A_1|}{|A|}, x_2 = \frac{|A_2|}{|A|}, \cdots, x_n = \frac{|A_n|}{|A|}, \tag{1.7}$$

其中 $|A_j|(j = 1, 2, \cdots, n)$ 是一个 n 阶行列式, 它由 $|A|$ 去掉第 j 列换上方程组的常数项 $b_1, b_2, \cdots, b_n$ 组成的列而成.

证明:

方法一: 只需验证(1.7)式确为方程组(1.6)的解.

将 $|A_j|$ 按第 j 列展开, 得

$$|A_j| = b_1A_{1j} + b_2A_{2j} + \cdots + b_nA_{nj} = \sum_{i=1}^{n} b_iA_{ij}, \tag{1.8}$$

根据式(1.7)和式(1.8),

$$x_j = \frac{|A_j|}{|A|} = \frac{1}{|A|}\sum_{i=1}^{n} b_iA_{ij}, j = 1, 2, \cdots, n. \tag{1.9}$$

方程组(1.6)第 k 个方程为

$$\sum_{j=1}^{n} a_{kj}x_j = b_k, 1 \leq k \leq n. \tag{1.10}$$

将(1.9)代入(1.10)左边得

$$\begin{aligned} \sum_{j=1}^{n} a_{kj}x_j &= \sum_{j=1}^{n}(a_{kj}\frac{1}{|A|}\sum_{i=1}^{n} b_iA_{ij}) \\ &= \frac{1}{|A|}\sum_{j=1}^{n}\sum_{i=1}^{n} a_{kj}b_iA_{ij} \\ &= \frac{1}{|A|}\sum_{i=1}^{n} b_i(\sum_{j=1}^{n} a_{kj}A_{ij}) \end{aligned}$$

上式最后一个等式里 $\sum\limits_{j=1}^{n} a_{kj}A_{ij}$ 部分的值如下：

当 $i \neq k$ 时, 为行列式 $|A|$ 第 k 行每个元素与第 i 行同列元素对应的代数余子式的乘积之和, 其值为 0;

当 $i = k$ 时, 为行列式 $|A|$ 按第 k 行元素展开, 即第 k 行每个元素与其对应代数余子式的乘积之和, 其值为 $|A|$. 所以,

$$\sum_{j=1}^{n} a_{kj}x_j = \frac{1}{|A|}b_k\sum_{j=1}^{n} a_{kj}A_{kj} = \frac{b_k}{|A|}|A| = b_k$$

由于 k 为任意 1 至 n 的自然数, 因此(1.7)式是(1.6)式的解.

方法二：利用系数矩阵求逆的方法。方程组(1.6)可以写成矩阵形式

$$Ax = \beta,$$

其中，$A = (a_{ij})$. 若 $|A| \neq 0$, 则 $A^{-}1$ 必存在，因此

$$A^{-1}(Ax) = A^{-1}\beta,$$

即

$$x = A^{-1}\beta$$

将上式中的矩阵写出来就是

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

于是

$$x_1 = \frac{1}{|A|}(b_1A_{11} + b_2A_{21} + \cdots + b_nA_{n1})$$

其中

$$|A_1| = \begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

其余 x_i 同理可得. □

1.3 行列式的计算

例 1.1 计算 n 阶 $VanderMonde$（范德蒙德）行列式：

$$V_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-2} & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \cdots & x_{n-1}^{n-2} & x_{n-1}^{n-1} \\ 1 & x_{n-2} & x_{n-2}^2 & \cdots & x_{n-2}^{n-2} & x_{n-2}^{n-1} \end{vmatrix}$$

解: 将第 $n-1$ 列乘以 $-x_n$ 后加到第 n 列上, 再将第 $n-2$ 列乘以 $-x_n$ 加到第 $n-1$ 列上. 这样一直做下去, 直到将第一列乘以 $-x_n$ 加到第二列上为止. 这样变形后行列式值不变. 得到

$$V_n = \begin{vmatrix} 1 & x_1 - x_n & x_1^2 - x_1x_n & \cdots & x_1^{n-2} - x_1^{n-3}x_n & x_1^{n-1} - x_1^{n-2}x_n \\ 1 & x_2 - x_n & x_2^2 - x_2x_n & \cdots & x_2^{n-2} - x_2^{n-3}x_n & x_2^{n-1} - x_2^{n-2}x_n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_{n-1} - x_n & x_{n-1}^2 - x_{n-1}x_n & \cdots & x_{n-1}^{n-2} - x_{n-1}^{n-3}x_n & x_{n-1}^{n-1} - x_{n-1}^{n-2}x_n \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{vmatrix}$$

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

提取每行公因子后得到 $n-1$ 阶行列式恰好是一个 $x_1, x_2, \cdots, x_{n-1}$ 的 $n-1$ 阶 $VanderMonde$ 行列式, 将其标记为 V_{n-1}. 得到递推公式

$$\begin{aligned} V_n &= (-1)^{n+1}(x_1 - x_n)(x_2 - x_n)\cdots(x_{n-1} - x_n) \cdot \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-2} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \cdots & x_{n-1}^{n-2} \\ 1 & x_{n-2} & x_{n-2}^2 & \cdots & x_{n-2}^{n-2} \end{vmatrix} \\ &= (x_n - x_1)(x_n - x_2)\cdots(x_n - x_{n-1})V_{n-1} \end{aligned}$$

即

$$V_n = \prod_{1 \le i < j \le n} (x_j - x_i)$$

例 1.2 求下列行列式的值:

$$F_n = \begin{vmatrix} \lambda & 0 & 0 & \cdots & 0 & a_n \\ -1 & \lambda & 0 & \cdots & 0 & a_{n-1} \\ 0 & -1 & \lambda & \cdots & 0 & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & a_2 \\ 0 & 0 & 0 & \cdots & -1 & \lambda + a_1 \end{vmatrix}$$

解: 按第一行展开,a_n 的余子式是一个上三角行列式, 其值等于 $(-1)^{n-1}$, λ 的余子式是与 F_n 相类似的 $n-1$ 阶行列式, 记之为 F_{n-1}, 于是得到递推关系式:

$$F_n = \lambda F_{n-1} + (-1)^{1+n}(-1)^{n-1}a_n = \lambda F_{n-1} + a_n.$$

将递推公式不断代入, 最后求得

$$F_n = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_n.$$

1.4 *Laplace* 定理

定理 1.3 (*Laplace* 定理) 设 $|A|$ 是 n 阶行列式, 在 $|A|$ 中任取 k 行(列), 那么包含于这 k 行(列)的全部 k 阶子式与它们对应的代数余子式的成绩之和等于 $|A|$. 即若取定 k 个行: $1 \le i_1 < i_2 < \cdots < i_k \le n$, 则

$$|A| = \sum_{1\le j_1<j_2<\cdots<j_k\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k & n \\ j_1 & j_2 & \cdots & j_k & n \end{pmatrix} \widehat{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k & n \\ j_1 & j_2 & \cdots & j_k & n \end{pmatrix} \tag{1.11}$$

同样若取定 k 个列: $1 \le j_1 < j_2 < \cdots < j_k \le n$, 则

$$|A| = \sum_{1\le i_1<i_2<\cdots<i_k\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k & n \\ j_1 & j_2 & \cdots & j_k & n \end{pmatrix} \widehat{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k & n \\ j_1 & j_2 & \cdots & j_k & n \end{pmatrix} \tag{1.12}$$

证明: 如果能证明 $|A|$ 的每个 k 阶子式与其代数余子式之积的展开式中的每一项都互不相同且都属于 $|A|$ 的展开式, 两者总项数相等, 那么就证明了 *Laplace* 定理.

首先证明 $|A|$ 的任意 k 阶子式与其代数余子式之积的展开式中的每一项都属于 $|A|$ 的展开式.

先证明特殊情形: $i_1 = 1, i_2 = 2, \cdots, , i_k = k; j_1 = 1, j_2 = 2, \cdots, j_k = k$. 这时 $|A|$ 可写为:

$$\begin{vmatrix} A_1 & * \\ * & A_2 \end{vmatrix}, \tag{1.13}$$

其中

$$|A_1| = A\left(\begin{smallmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{smallmatrix}\right) = \begin{vmatrix} a_{11} & \cdots & a1k \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}, \tag{1.14}$$

$$|A_2| = \widehat{A}\left(\begin{smallmatrix}1&2&\cdots&k\\1&2&\cdots&k\end{smallmatrix}\right) = \begin{vmatrix} a_{k+1,k+1} & \cdots & ak+1,n \\ \vdots & & \vdots \\ a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}. \tag{1.15}$$

(1.14)式中的任一项具有形式:

$$(-1)^{N(j_1,j_2,\cdots,j_k)}a_{j_1 1}a_{j_2 2}\cdots a_{j_k k},$$

其中 $N(j_1,j_2,\cdots,j_k)$ 是排列 $(j_1,j_2,\cdots,j_k)$ 的逆序数. (1.15)式中的任一项具有形式:

$$(-1)^{N(j_{k+1},j_{k+2},\cdots,j_n)}a_{j_{k+1},k+1}a_{j_{k+2},k+2}\cdots a_{j_n n},$$

所以

$$A\left(\begin{smallmatrix}1&2&\cdots&k\\1&2&\cdots&k\end{smallmatrix}\right)\widehat{A}\left(\begin{smallmatrix}1&2&\cdots&k\\1&2&\cdots&k\end{smallmatrix}\right)$$

中的任一项具有下列形式:

$$(-1)^{\sigma}a_{j_1 1}a_{j_2 2}\cdots a_{j_k k}a_{j_{k+1}k+1}\cdots a_{j_n n}, \tag{1.16}$$

其中 $\sigma = N(j_1,\cdots,j_k)+N(j_{k+1},\cdots,j_n)$. 注意 $(j_1,\cdots,j_k)$ 是 $(1,\cdots,k)$ 的一个排列, $(j_{k+1},\cdots,$ $j_n)$ 是 $(k+1,\cdots,n)$ 的一个排列, 因此 $(j_1,\cdots,j_k,j_{k+1},\cdots,j_n)$ 是 $(1,2,\cdots,n)$ 的一个排列且

$$N(j_1,\cdots,j_k,j_{k+1},\cdots,j_n) = N(j_1,\cdots,j_k)+N(j_{k+1},\cdots,j_n).$$

因此,(1.16)式是 $|A|$ 中的某一项.

再证明一般情况, 设

$$1\le i_1<i_2<\cdots<i_k\le n; 1\le j_1<j_2<\cdots<j_k\le n.$$

经过 i_1-1 次相邻两行的对换, 可把 i_1 行调到第一行, 经过 i_2-2 次相邻两行的对换, 可把 i_2 行调到第二行,$\cdots$, 经过 $(i_1+\cdots+i_k)-\frac{1}{2}k(k+1)$ 次对换, 可把第 $i_1,i_2,\cdots,i_k$ 行调至前 k 行. 同理, 经过 $(j_1+\cdots+j_k)-\frac{1}{2}k(k+1)$ 次对换, 可把第 $j_1,j_2,\cdots,j_k$ 行调至前 k 列. 因此,$|A|$ 经过 $(i_1+\cdots+i_k)+(j_1+\cdots+j_k)-k(k+1)$ 次行列的对换 (其中 $k(k+1)$ 必为偶数), 得到了一个新的行列式:

$$|C| = \begin{vmatrix} D & * \\ * & B \end{vmatrix}$$

其中

$$|D| = A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$$

由以上分析可知,$|C| = (-1)^{p+q}|A|, p = i_1+\cdots+i_k, q = j_1+\cdots+j_k$. $|B|$ 是子式 $|D|$ 在 $|C|$ 中的代数余子式, 由前面特殊情况的分析可知, $|D||B|$ 中的任一项都是 $|C|$ 中的项. 由于

$$\widehat{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} = (-1)^{p+q}|B|,$$

因此,

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \widehat{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \tag{1.17}$$

中的任一项都是 $(-1)^{p+q}|C| = |A|$ 中的项.

以上分析了(1.17)式中任一项均属于 $|A|$ 的展开式。

最后证明二者总项数相等：当 $i_1, i_2, \cdots, i_k$ 固定时, 对不同的 $1 \leq j_1 < j_2 < \cdots < j_k \leq n$ 共有 C_n^k 个不同子式, 每个子式完全展开后均含有 $k!$ 项，相应的余子式也有 C_n^k 个，每个余子式展开后均含有 $(n-k)!$ 项. 由(1.17)式展开得到的项是没有重复的, 且一共有 $C_n^k \cdot k!(n-1)! = n!$ 项.$|A|$ 的展开式也有 $n!$ 项, 因此(1.11)式成立.(1.12)式同理可证. $\square$

2 矩阵

2.1 矩阵的逆

定理 2.1 设 A 为 n 阶方阵，A^* 为 A 的伴随矩阵, 若 $|A| \neq 0$, 则 A 是一个非异阵，且

$$A^{-1} = \frac{1}{|A|}A^*. \tag{2.1}$$

证明:

$$AA^* = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

上式中两个矩阵乘积的第 (i, j) 元素为

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}.$$

按照定理(1.1)可知：当 $i = j$ 时上式为 $|A|$, 当 $i \neq j$ 时上式为零。因此

$$AA^* = \begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & |A| \end{pmatrix} = |A| \cdot I_n.$$

同理可证

$$A^*A = |A| \cdot I_n.$$

因为 $|A| \neq 0$, 由以上二式同除以 $|A|$ 可以得到

$$A(\frac{1}{|A|}A^*) = (\frac{1}{|A|}A^*)A = I_n,$$

因此

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

A 是非异阵. □

2.2 $Cauchy - Binet$ 公式

定理 2.2 **($Cauchy - Binet$ 公式)** 设 $A = (a_{ij})$ 是 $m \times n$ 矩阵,$B = (b_{ij})$ 是 $n \times m$ 矩阵. $A\begin{pmatrix} i_1 & \cdots & i_s \\ j_1 & \cdots & j_s \end{pmatrix}$ 表示 A 的一个 s 阶子式, 它由 A 的第 $i_1, \cdots, i_s$ 行与第 $j_1, \cdots, j_s$ 列交点上的元素按原次序排列组成的行列式, 同理可定义 B 的 s 阶子式.

(1) 若 $m > n$, 则必有 $|AB| = 0$;

(2) 若 $m \le n$, 则必有

$$|AB| = \sum_{1\le j_1<j_2<\cdots<j_m\le n} A\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix} B\begin{pmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{pmatrix}$$

证明: 令 $C = \begin{pmatrix} A & O \\ -I_n & B \end{pmatrix}$. 用不同方法表示行列式 $|C|$:

首先, 对 C 进行第三类分块初等变换, 得到矩阵 $M = \begin{pmatrix} O & AB \\ -I_n & B \end{pmatrix}$, M 可以改写为 $M = \begin{pmatrix} I_m & A \\ O & I_n \end{pmatrix} C$. 因此, $|M| = |C|$.

用 Laplace 定理来计算 $|M|$, 由于 AB 占据 M 的前 m 行和第 $n+1$ 至第 $n+m$ 列, M 除 AB 外, 前 m 行的其余列都是 0, 因此 $|M|$ 按前 m 行展开得

$$\begin{aligned} |M| &= (-1)^{(n+1+n+2+\cdots+n+m)+(1+2+\cdots+m)}|-I_n||AB| \\ &= (-1)^{nm+m(m+1)+n}|AB| \end{aligned}$$

上面第二个等式是因为 $|-I_n| = (-1)^n$. 由于 $m(m+1)$ 总是偶数, 可以去掉该式, 且 $nm+n = n(m+1)$, 所以

$$|M| = (-1)^{n(m+1)}|AB|. \tag{2.2}$$

再计算 $|C|$, 用 *Laplace* 定理按前 m 行展开, 分两种情况:

若 $m > n$, 则 A 为瘦长型的矩阵,B 为矮胖型矩阵, 如图2.1所示:
前 m 行中任意一个 m 阶子式至少有一列全部为 0, 即要从 O 中获取 $m-n$ 列用于构建 m 阶子

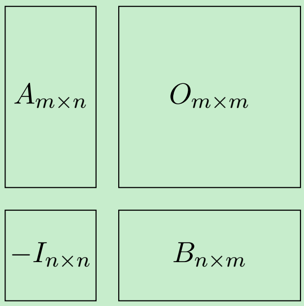

图 2.1: 瘦长型矩阵 A

式, 因此整个行列式的值为 0, 即 $|AB| = 0$.

若 $m \le n$, 则 A 为矮胖型矩阵,B 为瘦长型的矩阵, 如图2.2所示:

由 Laplace 定理得

$$|C| = \sum_{1<=j_1<j_2<\cdots<j_m<=n} A\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix} \hat{C}\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix}, \tag{2.3}$$

其中,$\hat{C}\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix}$ 是 $A\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix}$ 在矩阵 C 中的代数余子式. 显然,

$$\hat{C}\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix} = (-1)^{\frac{1}{2}m(m+1)+(j_1+j_2+\cdots+j_m)}|-e_{i_1}, -e_{i_2}, \cdots, -e_{i_{n-m}}, B|, \tag{2.4}$$

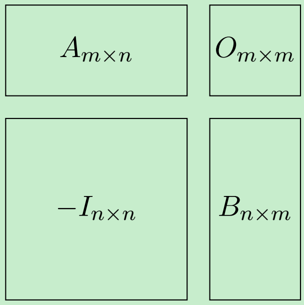

图 2.2: 矮胖型矩阵 A

其中 $i_1, i_2, \cdots, i_{n-m}$ 是 C 中前 n 列去掉 $j_1, j_2, \cdots, j_m$ 列后余下的列序数.e_{i_1}, e_{i_2}, $\cdots, e_{i_{n-m}}$ 是相应的 n 维标准单位列向量.$\hat{C}$ 除了包含全部 B, 还需要从 $-I_n$ 中获取 $n-m$ 个列. 记

$$|N| = |-e_{i_1}, -e_{i_2}, \cdots, -e_{i_{n-m}}, B|,$$

下面计算 $|N|$.$|N|$ 用 *Laplace* 定理按前 $n-m$ 列展开, 只有一个子式非零, 其值等于 $|-I_{n-m}| = (-1)^{n-m}$. 这个子式的余子式为 $B\left(\begin{smallmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{smallmatrix}\right)$. 因此,

$$|N| = (-1)^{(n-m)+(i_1+i_2+\cdots+i_{n-m})+(1+2+\cdots+n-m)} B\left(\begin{smallmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{smallmatrix}\right). \tag{2.5}$$

将(2.5)式代入(2.4)式, 再将(2.4)式代入(2.3)式得

$$|C| = \sum_{1\le j_1<j_2<\cdots<j_m\le n} (-1)^l A\left(\begin{smallmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{smallmatrix}\right) B\left(\begin{smallmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{smallmatrix}\right) \tag{2.6}$$

其中,$l = \frac{1}{2}m(m+1) + (j_1+j_2+\cdots+j_m) + (n-m) + (i_1+i_2+\cdots+i_{n-m}) + (1+2+\cdots+n-m)$. 从图(2.2)可知,$j_1+j_2+\cdots+j_m$ 是被选取到 $A\left(\begin{smallmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{smallmatrix}\right)$ 中的列的序数, 则 $i_1+i_2+\cdots+i_{n-m}$ 是前 n 列中剩余的列的序数, 所以 $j_1+j_2+\cdots+j_m+i_1+i_2+\cdots+i_{n-m} = 1+2+\cdots+n$. 因此, $l = (1+2+\cdots+m) + (1+2+\cdots+n) + (n-m) + (1+2+\cdots+n-m)$. 因为 $m \le n$, 所以 $l = 2(1+2+\cdots+m) + 2(n-m) + (m+1) + (m+2) + \cdots + n + (1+2+\cdots+n-m-1)$. 其中,$2(1+2+\cdots+m) + 2(n-m)$ 为偶数, 去掉该式后不影响 l 的奇偶性. 令 $l' = (m+1) + (m+2) + \cdots + n + (1+2+\cdots+n-m-1)$, 则 l 与 l' 有相同的奇偶性. 将 l' 首、尾对应项相加, 得 $l' = n(n-m)$. 又 $l' + n(m+1) = n(n-m) + n(m+1) = n(n+1)$, 因为 $n(n+1)$ 为偶数, 因此 l' 与 $n(m+1)$ 有相同的奇偶性, 因此,l 与 $n(m+1)$ 有相同的奇偶性.

最后, 由 $|M| = |C|$ 即(2.2)式 =(2.6)式, 得

$$(-1)^{n(m+1)}|AB| = \sum_{1\le j_1<j_2<\cdots<j_m\le n} (-1)^l A\left(\begin{smallmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{smallmatrix}\right) B\left(\begin{smallmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{smallmatrix}\right)$$

因为 l 取值和 $j_1, j_2, \cdots, j_m$ 无关, 所以

$$(-1)^{n(m+1)}|AB| = (-1)^l \sum_{1\le j_1<j_2<\cdots<j_m\le n} A\left(\begin{smallmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{smallmatrix}\right) B\left(\begin{smallmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{smallmatrix}\right)$$

上式中 l 与 $n(m+1)$ 有相同的奇偶性, 所以

$$|AB| = \sum_{1\leq j_1<j_2<\cdots<j_m\leq n} A\left(\begin{smallmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{smallmatrix}\right) B\left(\begin{smallmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{smallmatrix}\right)$$

当 $m=n$ 时, 上式右边只有一个组合. $\square$

3 线性空间

3.1 矩阵的秩

例 3.1 设 $C=\left(\begin{smallmatrix}A & O\\ O & B\end{smallmatrix}\right)$，求证：$\operatorname{rank}(C)=\operatorname{rank}(A)+\operatorname{rank}(B)$.

证明: 设初等矩阵 P_1,P_2,Q_1,Q_2 s. t. $P_1AQ_1=\left(\begin{smallmatrix}I_{r_1} & O\\ O & O\end{smallmatrix}\right)$, $P_2BQ_2=\left(\begin{smallmatrix}I_{r_2} & O\\ O & O\end{smallmatrix}\right)$. 于是,

$$\begin{pmatrix}P_1 & O\\ O & P_2\end{pmatrix}\begin{pmatrix}A & O\\ O & B\end{pmatrix}\begin{pmatrix}Q_1 & O\\ O & Q_2\end{pmatrix}=\begin{pmatrix}P_1AQ_1 & O\\ O & P_2BQ_2\end{pmatrix}=\begin{pmatrix}I_{r_1} & O & O & O\\ O & O & O & O\\ O & O & I_{r_2} & O\\ o & O & O & O\end{pmatrix}$$
$$\longrightarrow\begin{pmatrix}I_{r_1} & O & O & O\\ O & I_{r_2} & O & O\\ O & O & O & O\\ o & O & O & O\end{pmatrix}.$$

由矩阵乘以非异阵其秩不变以及分块矩阵在分块初等变换下秩不变，可得

$$\operatorname{rank}(C)=r_1+r_2=\operatorname{rank}(A)+\operatorname{rank}(B).$$

□

例 3.2 设 $C=\left(\begin{smallmatrix}A & D\\ O & B\end{smallmatrix}\right)$ 或 $\left(\begin{smallmatrix}A & O\\ D & B\end{smallmatrix}\right)$，求证：$\operatorname{rank}(C)\geq\operatorname{rank}(A)+\operatorname{rank}(B)$.

证明:

$$\begin{pmatrix}P_1 & O\\ O & P_2\end{pmatrix}\begin{pmatrix}A & D\\ O & B\end{pmatrix}\begin{pmatrix}Q_1 & O\\ O & Q_2\end{pmatrix}=\begin{pmatrix}P_1AQ_1 & P_1DQ_2\\ O & P_2BQ_2\end{pmatrix}$$
$$=\begin{pmatrix}I_{r_1} & O & D_{11} & D_{12}\\ O & O & D_{21} & D_{22}\\ O & O & I_{r_2} & O\\ O & O & O & O\end{pmatrix}\longrightarrow\begin{pmatrix}I_{r_1} & O & O & O\\ O & O & O & D_{22}\\ O & O & I_{r_2} & O\\ O & O & O & O\end{pmatrix}\longrightarrow\begin{pmatrix}I_{r_1} & O & O & O\\ O & D_{22} & O & O\\ O & O & I_{r_2} & O\\ O & O & O & O\end{pmatrix}$$

由上例可知，$\operatorname{rank}(C)=r_1+\operatorname{rank}(D_{22})+r_2\geq r_1+r_2=\operatorname{rank}(A)+\operatorname{rank}(B)$，当 $D_{22}=O$ 时等号成立. □

例 3.3 设 $A\in M_{m\times n}(\mathbb{K})$，$B\in M_{n\times p}(\mathbb{K})$，求证：$\operatorname{rank}(A)+\operatorname{rank}(B)-n\leq\operatorname{rank}(AB)\leq\min\{\operatorname{rank}(A),\operatorname{rank}(B)\}$.

证明: 先证右边. 令 $B=(\beta_1,\beta_2,\cdots,\beta_p)$, 设 $\{\beta_{i1},\beta_{i2},\cdots,\beta_{ir}\}$ 是其极大无关组且断言 $\forall\,1\leq j\leq p$，$A\beta_j$ 是 $A\beta_{i1},\beta_{i2},\cdots,\beta_{ir}$ 的线性组合. 因为 $\forall\,\beta_j$，$\beta_j=\lambda_1\beta_{i1}+\lambda_2\beta_{i2}+\cdots+\lambda_r\beta_{ir}\Longrightarrow$

$A\beta_j = \lambda_1 A\beta_{i1} + \lambda_2 A\beta_{i2} + \cdots + \lambda_r A\beta_{ir}$ 所以，$AB = (A\beta_1, A\beta_2, \cdots, A\beta_p)$ 的每一个列向量都可由 $\{A\beta_{i1}, A\beta_{i2}, \cdots, A\beta_{ir})\}$ 线性表示 $\Longrightarrow \operatorname{rank}(AB) \le r = \operatorname{rank}(B)$. 同理，$\operatorname{rank}(AB) = \operatorname{rank}(B'A') \le \operatorname{rank}(A') = \operatorname{rank}(A) \Longrightarrow \operatorname{rank}(AB) \le \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$

再证左边.

$\left(\begin{smallmatrix} A & O \\ -I_n & B \end{smallmatrix}\right) \longrightarrow \left(\begin{smallmatrix} O & AB \\ -I_n & B \end{smallmatrix}\right) \longrightarrow \left(\begin{smallmatrix} O & AB \\ -I_n & O \end{smallmatrix}\right) \longrightarrow \left(\begin{smallmatrix} AB & O \\ O & I_n \end{smallmatrix}\right)$. 由最右边矩阵可知，$\operatorname{rank}\left(\begin{smallmatrix} AB & O \\ O & I_n \end{smallmatrix}\right) = \operatorname{rank}(AB) + n = r\left(\begin{smallmatrix} A & O \\ -I_n & B \end{smallmatrix}\right) \ge \operatorname{rank}(A) + \operatorname{rank}(B) \Longrightarrow \operatorname{rank}(AB) \ge \operatorname{rank}(A) + \operatorname{rank}(B) - n$. □

例 3.4 （秩的降阶公式）$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$，则：

（1）若 A 非异，则 $\operatorname{rank}(M) = \operatorname{rank}(A) + \operatorname{rank}(D - CA^{-1}B)$;

（2）若 D 非异，则 $\operatorname{rank}(M) = \operatorname{rank}(D) + \operatorname{rank}(A - BD^{-1}C)$;

（3）若 A, D 非异，则 $\operatorname{rank}(A) + \operatorname{rank}(D - CA^{-1}B) = \operatorname{rank}(D) + \operatorname{rank}(A - BD^{-1}C)$.

证明: 只证（1）即可，其余类似可证。

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \longrightarrow \begin{pmatrix} A & B \\ O & D - CA^{-1}B \end{pmatrix} \longrightarrow \begin{pmatrix} A & O \\ O & D - CA^{-1}B \end{pmatrix}$$

由前面例题可知，$\operatorname{rank}(M) = \operatorname{rank}(A) + \operatorname{rank}(D - CA^{-1}B)$ □

3.2 线性方程组的解

例 3.5 设 $Ax = \beta(\beta \ne 0)$ 的特解为 γ, 相伴齐次线性方程组 $Ax = 0$ 的基础解系为 $\eta_1, \eta_2, \cdots, \eta_{n-r}$, 求证：

（1）$\gamma, \gamma + \eta_1, \gamma + \eta_2, \cdots, \gamma + \eta_{n-r}$ 线性无关；

（2）$Ax = \beta$ 的任一解必为如下形式：$c_0\gamma + c_1(\gamma + \eta_1) + c_2(\gamma + \eta_2) + \cdots + c_{n-r}(\gamma + \eta_{n-r})$, 其中 $c_0 + c_1 + \cdots + c_{n-r} = 1$.

证明:

（1）设 $\lambda_0\gamma + \lambda_1(\gamma + \eta_1) + \lambda_2(\gamma + \eta_2) + \lambda_2(\gamma + \eta_2) + \cdots + \lambda_{n-r}(\gamma + \eta_{n-r}) = 0$, 合并整理得下式：

$$(\sum_{i=0}^{n-r} \lambda_i)\gamma + \lambda_1\eta_1 + \lambda_2\eta_2 + \cdots + \lambda_{n-r}\eta_{n-r} = 0, \tag{3.1}$$

将 A 作用到上式两边，得

$$(\sum_{i=0}^{n-r} \lambda_i)A\gamma + \lambda_1 A\eta_1 + \lambda_2 A\eta_2 + \cdots + \lambda_{n-r} A\eta_{n-r} = 0,$$

$A\eta_i = 0 \Longrightarrow (\sum_{i=0}^{n-r} \lambda_i) A\gamma = 0 \xrightarrow{A\gamma=\beta} (\sum_{i=0}^{n-r} \lambda_i)\beta = 0 \xrightarrow{\beta\neq 0} \sum_{i=0}^{n-r} \lambda_i = 0,$

将 $\sum_{i=0}^{n-r} \lambda_i = 0$ 代入(3.1)式，可得 $\lambda_1\eta_1 + \lambda_2\eta_2 + \cdots + \lambda_{n-r}\eta_{n-r} = 0,$

$\because \eta_1, \eta_2, \cdots, \eta_{n-r}$ 是基础解系,$\therefore \lambda_1 = \lambda_2 = \cdots = \lambda_{n-r} = 0$,

由(3.1)式可知，$\gamma, \gamma + \eta_1, \gamma + \eta_2, \cdots, \gamma + \eta_{n-r}$ 线性无关.

(2) 设 α 为线性方程组 $Ax = \beta$ 的任一解，$k_i \in \mathbb{K}$, 那么

$$\begin{aligned}\alpha &= \gamma + k_1\eta_1 + k_2\eta_2 + \cdots + k_{n-r}\eta_{n-r} \\ &= (1 - k_1 - k_2 - \cdots - k_{n-r})\gamma + k_1(\gamma + \eta_1) + k_2(\gamma + \eta_2) + \cdots + k_{n-r}(\gamma + \eta_{n-r})\end{aligned}$$

令:

$$\begin{aligned}c_0 &= 1 - k_1 - k_2 - \cdots - k_{n-r} \\ c_1 &= k_1 \\ c_2 &= k_2 \\ &\cdots \\ c_{n-r} &= k_{n-r}\end{aligned}$$

显然，$\sum_{0}^{n-r} c_i = 1.$ □

例 3.6 $A^2 - A - 3I_n = 0$, 求证：$A - 2I_n$ 非异.

证明:

1、凑因子法：$(A - 2I_n)(A + I_n) = I_n$, 所以 $A - 2I_n$ 非异。

2、线性方程组解法：只要证明 $(A - 2I_n)x = 0$ 只有零解。

设 x_0 是线性方程组 $(A - 2I_n)x = 0$ 的解，那么

$Ax_0 = 2x_0, A^2x_0 = 2Ax_0 = 4x_0,$

$(A^2 - A - 3I_n)x_0 = 0 \Longrightarrow -x_0 = 0 \Longrightarrow x_0 = 0.$ □

例 3.7 假设 $A \in M_{m\times n}(\mathbb{R})$, 证明:$\mathrm{rank}(AA') = \mathrm{rank}(A'A) = \mathrm{rank}(A)$.

证明: 因为 $\mathrm{rank}(AA') = \mathrm{rank}(A'A)$, 故只需证明右边的等式。

设 $Ax = 0$ 的解空间为 V_A, $A'Ax = 0$ 的解空间为 $V_A'A$. $Ax = 0$ 时, $A'Ax$ 必为 0, 所以 $V_A \subseteq V_{A'A}$.

任取 $x_0 \in V_{A'A}$, 此时 $x_0 \in \mathbb{R}_n$ 且 $A'Ax_0 = 0$.

令 $Ax_0 = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_m \end{pmatrix} \in \mathbb{R}_m \Longrightarrow (Ax_0)'(Ax_0) = 0 \Longrightarrow (a_1, a_2, \cdots, a_m)\begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_m \end{pmatrix} = 0$

$$\Longrightarrow \sum_{1}^{m} a_i^2 = 0 \Longrightarrow \forall i, a_i = 0 \Longrightarrow Ax_0 = 0 \Longrightarrow V'_A A \subseteq V_A \xRightarrow{V_A \subseteq V_{A'A}} V_A = V_{A'A}$$

$$\Longrightarrow \operatorname{rank}(A) = \operatorname{rank}(A'A) \qquad \square$$

4 线性映射

4.1 线性映射与矩阵

定理 4.1 设 $V \in \mathbb{K}^n, U \in \mathbb{K}^m, W \in \mathbb{K}^p$, 它们的基分别为 $\{e_1, e_2, \cdots, e_n\}$, $\{f_1, f_2, \cdots, f_m\}$, $\{g_1, g_2, \cdots, g_p\}$, T 是从线性映射到其表示矩阵的线性同构. 设 $\varphi : V \longrightarrow U, \psi : U \longrightarrow W$ 为两个线性映射. 则 $\psi \circ \varphi : V \longrightarrow W$ 为线性映射, 且 $T(\psi \circ \varphi) = T(\psi) \cdot T(\varphi)$.

证明: 依题意设 $T(\varphi) = A = (a_{ij})_{m\times n}, T(\psi) = B = (b_{ij})_{p\times m}$. 只要证 $T(\varphi \circ \psi) = BA$ 即可.

$$
\begin{aligned}
&(\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n)) = (f_1, f_2, \cdots, f_m)A \\
&\Longleftrightarrow \varphi(e_j) = \sum_{i=1}^{m} a_{ij} f_i, \forall 1 \leq j \leq n \\
&\xRightarrow{\text{将}\psi\text{作用到上式两边}} \psi(\varphi(e_j)) = \psi(\sum_{i=1}^{m} a_{ij} f_i) = \sum_{i=1}^{m} a_{ij}\psi(f_i), \forall 1 \leq j \leq n \\
&\Longrightarrow (\psi\varphi(e_1), \psi\varphi(e_2), \cdots, \psi\varphi(e_n)) = (\psi(f_1), \psi(f_2), \cdots, \psi(f_m))A
\end{aligned}
$$

$\because (\psi(f_1), \psi(f_2), \cdots, \psi(f_n)) = (g_1, g_2, \cdots, g_p)B$,
$\therefore (\psi\varphi(e_1), \psi\varphi(e_2), \cdots, \psi\varphi(e_n)) = (g_1, g_2, \cdots, g_p)BA$.
$\therefore T(\psi \circ \varphi) = T(\psi) \cdot T(\varphi)$. □

推论 4.1 $T : \mathcal{L}(V) \longrightarrow M_n(\mathbb{K})$ 保持乘法, 从而 T 是一个 $\mathbb{K} -$ 代数 同构.

推论 4.2 上述 T 同构有下列性质:

(1)$T(I_V) = I_n$;

(2)$\varphi \in \mathcal{L}(V)$, 则 φ 是自同构的充要条件是 $T(\varphi)$ 可逆且 $T(\varphi^{-1}) = T(\varphi)^{-1}$.

证明:

(1) 任取 V 的基 $\{e_1, e_2, \cdots, e_n\}, I_V(e_i) = e_i \Longrightarrow T(I_V) = I_n$.

(2) 证明必要性: 设 φ 为自同构, 则 $I_V = \varphi \cdot \varphi^{-1} = \varphi^{-1} \cdot \varphi$.
$I_n = T(I_V) = T(\varphi \cdot \varphi^{-1}) = T(\varphi) \cdot T(\varphi^{-1}) \Longrightarrow T(\varphi)$为可逆阵且$T(\varphi^{-1}) = T(\varphi)^{-1}$.

证明充分性: 设 $T(\varphi)$ 为可逆阵, 即 $T(\varphi)^{-1}$ 存在. $T : \mathcal{L}(V) \longrightarrow M_n(\mathbb{K})$ 为双射, 根据其满性: $\exists\psi \longmapsto T(\varphi)^{-1}$, 即 $T(\psi) = T(\varphi)^{-1}$, 即 $I_n = T(\varphi) \cdot T(\psi) = T(\psi) \cdot T(\varphi) \Longrightarrow T(I_V) = T(\varphi \circ \psi) = T(\psi \circ \varphi) \xRightarrow{\text{单性}} I_V = \varphi\psi = \psi\varphi \Longrightarrow \psi = \varphi^{-1}, \varphi$ 为自同构. □

4.2 线性映射的像与核

定理 4.2 设 $\varphi: V^n \longrightarrow U^m$ 为线性映射，设 V 的一组基为 $\{e_1, e_2, \cdots, e_n\}$, U 的一组基为 $\{f_1, f_2, \cdots, f_m\}$, φ 在给定基下的表示矩阵为 $A \in M_{m\times n}(\mathbb{K})$. 则 $\operatorname{rank}(\varphi) = \operatorname{rank}(A)$, $\dim\ker\varphi = n - \operatorname{rank}(A)$.

证明: 参看下面映射交换图: 上图中, $\eta_U \circ \varphi = \varphi_A \circ \eta_V$. 只需证 $\ker\varphi \cong \ker\varphi_A$, $\operatorname{Im}\varphi \cong \operatorname{Im}\varphi_A$.

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "\eta_V"', "\cong"] & U \arrow[d, "\eta_U", "\cong"'] \\ \mathbb{K}^n \arrow[r, "\varphi_A"] & \mathbb{K}^m \end{tikzcd}$$

图 4.1: 映射交换图

Step1: 先证 $\eta_V(\ker\varphi) \subseteq \ker\varphi_A$, $\eta_U(\operatorname{Im}\varphi) \subseteq \operatorname{Im}\varphi_A$.

（1）任取 $v \in \ker\varphi$, 只要证 $\eta_V(v) \in \ker\varphi_A$.
$\because \varphi_A(\eta_V(v)) = \eta_U(\varphi(v)) = \eta_U(0) = 0$, 得证.

(2) 任取 $\varphi(v) \in \operatorname{Im}\varphi$, $v \in V$, 只要证 $\eta_U(\varphi(v)) \in \operatorname{Im}\varphi_A$.
$\because \eta_U(\varphi(v)) = \varphi_A(\eta_V(v)) \in \operatorname{Im}\varphi_A$, 得证.

Step2: 作限制: $\eta_V: \ker\varphi \longrightarrow \ker\varphi_A$, $\eta_U: \operatorname{Im}\varphi \longrightarrow \operatorname{Im}\varphi_A$.
二者都是单线性映射, 只需证明以上两个映射是满射.

(1) 先证 $\eta_V: \ker\varphi \longrightarrow \ker\varphi_A$ 是满射:

$\because$ 任取 $x \in \ker\varphi_A$, 即 $\varphi_A(x) = 0$, $x \in \ker\varphi_A \subseteq \mathbb{K}^n$ 且 $\eta_V: V \longrightarrow \mathbb{K}^n$ 是满射, 从而 $\exists v \in V$ s.t. $\eta_V(v) = x$. $\eta_U(\varphi(v)) = \varphi_A(\eta_V(v)) = \varphi_A(x) = 0$, 由 η_U 为同构可知, $\varphi(v) = 0$, 即 $v \in \ker\varphi$.

(2) 再证 $\eta_U: \operatorname{Im}\varphi \longrightarrow \operatorname{Im}\varphi_A$ 是满射.

$\because$ 任取 $\varphi_A(x) \in \operatorname{Im}\varphi_A$, 其中 $x \in \mathbb{K}^n$. 取 $v \in V$ s.t. $\eta_V(v) = x$. $\eta_U(\varphi(v)) = \varphi_A(\eta_V(v)) = \varphi_A(x)$, $\varphi(v) \in \operatorname{Im}\varphi$. 从而 $\eta_U: \operatorname{Im}\varphi \longrightarrow \operatorname{Im}\varphi_A$ 是满射.

Step3: $\ker\varphi_A = \{x \in \mathbb{K}_n | Ax = 0\}$ 即 $\ker\varphi_A$ 为 $Ax = 0$ 的解空间 V_A.
$\ker\varphi \cong \ker\varphi_A \Longrightarrow \dim\ker\varphi = \dim\ker\varphi_A = \dim V_A = n - \operatorname{rank}(A)$.

作列分块 $A = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, 其中 $\alpha_i \in \mathbb{K}_m$,
另 $\operatorname{Im}\varphi_A = \{Ax | x \in \mathbb{K}_n\}$, 其中 $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}_n$,
从而推出 $\operatorname{Im}\varphi_A = \{x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n | x_i \in \mathbb{K}\} = L(\alpha_1, \alpha_2, \cdots, \alpha_n)$.
$\operatorname{Im}\varphi \cong \operatorname{Im}\varphi_A \Longrightarrow \operatorname{rank}(\varphi) = \dim\operatorname{Im}\varphi = \dim\operatorname{Im}\varphi_A = \dim\ L(\alpha_1, \alpha_2, \cdots, \alpha_n)$
$= \operatorname{rank}(\{\alpha_1, \alpha_2, \cdots, \alpha_n\}) = \operatorname{rank}(A)$. □

推论 4.3 (线性映射的维数公式) 设 $\varphi: V^n \longrightarrow U^m$ 为线性映射, 则 $\dim\ker\varphi + \dim\operatorname{Im}\varphi = \dim V$.

证明: $\because$ 根据定理4.2, $\dim\ker\varphi = n - \operatorname{rank}(A)$, $\dim\operatorname{Im}\varphi = \operatorname{rank}(A)$
$\therefore \dim\ker\varphi + \dim\operatorname{Im}\varphi = n = \dim V$. □

推论 4.4 设 $\varphi: V^n \longrightarrow U^m$ 为线性映射, φ 在某组基下的表示矩阵为 $A^{m\times n}$, 则

(1) φ 为满射 $\Longleftrightarrow$ A 为行满秩矩阵, 即 $\operatorname{rank}(A) = m$;

(2) φ 为单射 $\Longleftrightarrow$ A 为列满秩矩阵, 即 $\operatorname{rank}(A) = n$.

证明: (1) 根据定理4.2:

$$\begin{aligned}&\varphi\text{是满射} \Longleftrightarrow \dim\operatorname{Im}\varphi = \dim U = m\\ &\xLeftrightarrow{\operatorname{rank}(\varphi)=\operatorname{rank}(A)} \dim\operatorname{Im}\varphi = \operatorname{rank}(A) = m\end{aligned}$$

(2) 根据定理4.2:

$$\begin{aligned}&\varphi\text{是单射} \Longleftrightarrow \dim\ker\varphi = 0\\ &\xLeftrightarrow{\dim\ker\varphi=n-\operatorname{rank}(A)} \operatorname{rank}(A) = n\end{aligned}$$

□

推论 4.5 设 $\varphi: V^n \longrightarrow U^n$ 为线性映射且 $\dim V = \dim U = n$, 则以下结论等价:

(1)φ 为单射;

(2)φ 为满射;

(3)φ 为同构.

证明:

(1) 设 φ 为单射, 只要证 φ 为满射.

$$\begin{aligned}&\because 0 = \dim\ker\varphi = \dim V - \dim\operatorname{Im}\varphi\\ &\Longrightarrow \dim\operatorname{Im}\varphi = \dim V = \dim U\\ &\Longrightarrow \operatorname{Im}\varphi = U\end{aligned}$$

(2) 设 φ 为满射, 只要证 φ 为单射.

$$\begin{aligned}&\operatorname{Im}\varphi = U \Longrightarrow \dim\operatorname{Im}\varphi = \dim U = \dim V\\ &\Longrightarrow \dim\ker\varphi = 0 \Longrightarrow \ker\varphi = 0\end{aligned}$$

.

(3) 由以上 (1) 和 (2) 可得.

特别地，当 $\varphi \in \mathcal{L}(V^n)$ 时, φ 为自同构 $\Longleftrightarrow$ φ 为单射 $\Longleftrightarrow$ φ 为满射. □

推论 4.6 设 $\varphi \in \mathcal{L}(V^n)$, 则:
φ 为满射或单射 $\Longleftrightarrow$ φ 在任一组基下的表示矩阵可逆.

证明: 从几何角度证明：由推论4.5可知，φ 为满射或单射 $\Longleftrightarrow$ φ 为自同构 $\Longleftrightarrow$ φ 在任一组基下表示矩阵可逆.

从代数角度证明：φ 为满射或单射 $\Longleftrightarrow$ $A^{n\times n}$ 为行或列满秩矩阵 $\Longleftrightarrow$ A 为非异矩阵. □

例 4.1 设 $\varphi : V^n \Longrightarrow U^m$ 为线性映射，V 的一组基为 $\{e_1, e_2, \cdots, e_n\}$, U 的一组基为 $\{f_1, f_2, \cdots, f_m\}$. 计算:

(1)$\ker\varphi$ 用 e_i 的线性组合表示;

(2)$\mathrm{Im}\,\varphi$ 用 f_i 的线性组合表示.

解: 参考图4.1:

$\eta_V : \ker\varphi \longrightarrow \ker\varphi_A$ 为线性同构. 该映射将 V 中任意向量 v 映射成其坐标向量，即 $v \mapsto (x_1, x_2, \cdots, x_n)'$; 其逆映射为 $\eta_V^{-1} : \ker\varphi_A \longrightarrow \ker\varphi$, 该映射将核空间中向量的坐标映射回核空间中的向量，即 $(x_1, x_2, \cdots, x_n)' \mapsto x_1e_1 + x_2e_2 + \cdots + x_ne_n$.

同理，$\eta_U : \mathrm{Im}\,\varphi \longrightarrow \mathrm{Im}\,\varphi_A$ 为线性同构, 其逆映射为 $\eta_U^{-1} : \mathrm{Im}\,\varphi_A \longrightarrow \mathrm{Im}\,\varphi$, 将像的坐标向量映射回像向量，即 $(y_1, y_2, \cdots, y_m)' \mapsto y_1f_1 + y_2f_2 + \cdots + y_mf_m$.

根据定理4.2中的证明中 Step3 中所述，$\ker\varphi = \{x \in \mathbb{K}_n | Ax = 0\}$ 即 $\ker\varphi_A$ 为 $Ax = 0$ 的解空间, 又 $\mathrm{Im}\,\varphi_A = \{Ax | x \in \mathbb{K}_n\} = L(\alpha_1, \alpha_2, \cdots, \alpha_n)$，即 $\mathrm{Im}\,\varphi_A$ 为 A 的列向量张成的子空间.

计算方法:

(1) 对表示矩阵 A 实施初等行变换（必要时进行列对换):

得到 A 的列向量极大无关组 $\{\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}\}$, $r = \mathrm{rank}(A)$;

得到 $Ax = 0$ 的基础解系 $\{\beta_1, \beta_2, \cdots, \beta_{n-r}\}$.

(2) 算出 $\ker\varphi$ 和 $\mathrm{Im}\,\varphi$:

$$\begin{aligned}\ker\varphi &= k_1\eta_V^{-1}(\beta_1) + k_2\eta_V^{-1}(\beta_2) + \cdots + k_{n-r}\eta_V^{-1}(\beta_{n-r}), k \in \mathbb{K}\\ \mathrm{Im}\,\varphi &= l_1\eta_U^{-1}(\alpha_{i_1}) + l_2\eta_U^{-1}(\alpha_{i_2}) \cdots + l_r\eta_U^{-1}(\alpha_{i_r}), l \in \mathbb{K}\end{aligned}$$

例 4.2 V^5, U^4 为 $\mathbb{K}$ 上两个线性空间且 $\varphi : V \longrightarrow U$ 为线性映射，$\{e_1, e_2, e_3, e_4,$

$e_5\}$ 为 V 的基，$\{f_1, f_2, f_3, f_4\}$ 为 U 的基. φ 在上述基下的表示矩阵为

$$A=\begin{pmatrix}1&2&1&-3&2\\2&1&1&1&-3\\1&1&2&2&-2\\2&3&-5&-17&10\end{pmatrix},$$

求 $\operatorname{Im}\varphi$ 和 $\ker\varphi$.

解: 将表示矩阵 A 做行变换，

$$A=\begin{pmatrix}1&2&1&-3&2\\2&1&1&1&-3\\1&1&2&2&-2\\2&3&-5&-17&10\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0&1&-\frac{9}{4}\\0&1&0&-3&\frac{11}{4}\\0&0&1&2&-\frac{5}{4}\\0&0&0&0&0\end{pmatrix}$$

其中，A 的最大无关列向量组为

$$\{\alpha_1,\alpha_2,\alpha_3\}=\left\{\begin{pmatrix}1\\2\\1\\2\end{pmatrix},\begin{pmatrix}2\\1\\1\\3\end{pmatrix},\begin{pmatrix}1\\1\\2\\-5\end{pmatrix}\right\},$$

$\operatorname{Im}\varphi=k_1(f_1+2f_2+f_3+2f_4)+k_2(2f_1+f_2+f_3+3f_4)+k_3(f_1+f_2+2f_3-5f_4), k\in\mathbb{K};$
$Ax=0$ 的基础解系为

$$\beta_1=\begin{pmatrix}-1\\3\\-2\\1\\0\end{pmatrix},\beta_2=\begin{pmatrix}9\\-11\\5\\0\\4\end{pmatrix},$$

$$\ker\varphi=l_1(-e_1+3e_2-2e_3+e_4)+l_2(9e_1-11e_2+5e_3+4e_5), l\in\mathbb{K}.$$

4.3 不变子空间

例 4.3 设 $\varphi\in\mathcal{L}(V)$ 是纯量变换，即 $\varphi(\alpha)=k\alpha$, 其中 $k\in\mathbb{K}$ 固定, 则 V 的任一子空间都是 $\varphi-$ 不变子空间.

引理 4.1 设 $\varphi\in\mathcal{L}(V)$, $U=L(\alpha_1,\alpha_2,\cdots,\alpha_m)$, 其中，$\alpha_i\in V$, 则 U 为 $\varphi-$ 不变子空间 $\Longleftrightarrow\varphi(\alpha_i)\in U, \forall 1\le i\le m$.

证明:

先证必要性: U 是 $\varphi-$ 不变子空间, 由不变子空间的定义可知，$\varphi(\alpha_i)\in U$.

再证必要性: 设 $\varphi(\alpha_i) \in U, 1 \le i \le m$.
任取 $\alpha \in U$, 可设 $\alpha = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_m\alpha_m, \lambda_i \in \mathbb{K}$,
则 $\varphi(\alpha) = \lambda_1\varphi(\alpha_1) + \lambda_2\varphi(\alpha_2) + \cdots + \lambda_m\varphi(\alpha_m) \in U$. □

定理 4.3 设 $\varphi \in \mathcal{L}(V^n)$, U 是 $\varphi-$ 不变子空间, 取 U 的一组基 $\{e_1, e_2, \cdots, e_r\}$, 再扩张为 V 的基 $\{e_1, e_2, \cdots, e_r, e_{r+1}, \cdots, e_n\}$, 则 φ 在上述基下的表示矩阵为 $\left(\begin{smallmatrix} A & B \\ O & D \end{smallmatrix}\right)$.

证明: 由不变子空间的定义可知, $\varphi(e_i) \in U(1 \le i \le r)$,

$$\varphi(e_1) = a_{11}e_1 + a_{21}e_2 + \cdots + ar1e_r,$$
$$\varphi(e_2) = a_{12}e_1 + a_{22}e_2 + \cdots + ar2e_r,$$
$$\cdots$$
$$\varphi(e_r) = a_{1r}e_1 + a_{2r}e_2 + \cdots + arre_r,$$

$$\therefore (\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n)) = (e_1, e_2, \cdots, e_n)\begin{pmatrix} A & B \\ O & D \end{pmatrix}$$

□

注 定理4.3的逆命题也成立: 设 $\varphi \in \mathcal{L}(V^n)$, φ 在一组基 $\{e_1, e_2, \cdots, e_n\}$ 下的表示矩阵为 $\left(\begin{smallmatrix} A & B \\ O & D \end{smallmatrix}\right)$. 令 $U = L(e_1, e_2, \cdots, e_r)$, 则 U 是 $\varphi-$ 不变子空间.

证明: 由引理4.1 $\Longrightarrow$ 只要 $\varphi(e_i) \in U, \forall 1 \le i \le r$.
由表示矩阵定义可知:
$$(\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_r), \cdots, e_n) = (e_1, e_2, \cdots, e_r, \cdots, e_n)\begin{pmatrix} A^{r\times r} & B^{r\times(n-r)} \\ O & D^{(n-r)\times(n-r)} \end{pmatrix}$$
$\Longrightarrow \varphi(e_i) \in U, \forall 1 \le i \le r$. □

引理 4.2 设 $\varphi \in \mathcal{L}(V^n)$, $V = V_1 \oplus V_2$, 其中, V_1, V_2 都是 $\varphi-$ 不变子空间, 则可取 V_1, V_2 的一组基拼成 V 的一组基, 使得 φ 在这组基下的表示矩阵为 $\left(\begin{smallmatrix} A & O \\ O & D \end{smallmatrix}\right)$.

证明: 取 $V1$ 的基 $\{e_1, e_2, \cdots, e_r\}$ 及 $V2$ 的基 $\{e_{r+1}, e_{r+2}, \cdots, e_n\}$
$\Longrightarrow \varphi(e_i) \in V_1(1 \le i \le r), \varphi(e_j) \in V_2(r+1 \le j \le n)$
$$\Longrightarrow \varphi(e_1), \varphi(e_2), \cdots, \varphi(e_r), \cdots, e_n) = (e_1, e_2, \cdots, e_r, \cdots, e_n)\begin{pmatrix} A^{r\times r} & O \\ O & D^{(n-r)\times(n-r)} \end{pmatrix}$$ □

推广 设 $\varphi \in \mathcal{L}(V^n)$, $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$, 其中 V_i 均为 $\varphi-$ 不变子空间, 取定 V_i 的基, 设 $\varphi|_{V_i}$ 在给定基下的表示矩阵为 $A_i(1 \le i \le m)$. 将 V_i 的基拼成 V 的基, 则 φ 的表示矩阵为 $\mathrm{diag}\{A_1, A_2, \cdots, A_m\}$.

例 4.4 V^3 的基为 $\{e_1, e_2, e_3\}$, $\varphi \in \mathcal{L}(V)$ 的表示矩阵为 $A = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$. 求证:
$U = L(e_3, e_1 + e_2 + 2e_3)$ 是 $\varphi-$ 不变子空间.

证明: 由引理4.1 $\Longrightarrow$ 只要证 $\varphi(e_3) \in U, \varphi(e_1+e_2+2e_3) \in U$.

方法一：直接验证

$\because \eta : V \longrightarrow \mathbb{K}^3$ 为线性同构，$\therefore$ 只要对坐标向量进行验证即可.

$$\because e_3 \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, e_1+e_2+2e_3 \mapsto \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix},$$

$$\therefore \varphi(e_3) = A \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}, \varphi(e_1+e_2+2e_3) = A \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}.$$

因此 $\varphi(e_3) = -(e_1+e_2+2e_3)+2e_3, \varphi(e_1+e_2+2e_3) = 2(e_1+e_2+2e_3)$

所以, $\varphi(e_3) \in U, \varphi(e_1+e_2+2e_3) \in U$.

方法二：证明 $\varphi(e_3)$ 和 $\varphi(e_1+e_2+2e_3)$ 是 e_3 和 $e_1+e_2+2e_3$ 的线性组合

$$\because \operatorname{rank} U = \operatorname{rank}(e_3, e_1+e_2+2e_3) = \operatorname{rank} \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 2 \end{pmatrix} = 2,$$

$\therefore$ 只要证 $\operatorname{rank}(\varphi(e_3), e_3, e_1+e_2+2e_3) = \operatorname{rank}(\varphi(e_1+e_2+2e_3), e_3, e_1+e_2+2e_3) = 2$ 即可.

显然，$\operatorname{rank}(\varphi(e_3), e_3, e_1+e_2+2e_3) = \operatorname{rank} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} = 2,$

$$\operatorname{rank}(\varphi(e_1+e_2+2e_3), e_3, e_1+e_2+2e_3) = \operatorname{rank} \begin{pmatrix} 2 & 0 & 1 \\ 2 & 0 & 1 \\ 4 & 1 & 2 \end{pmatrix} = 2.$$ □

从图4.1可以看出，上面是线性空间表示的几何语言，下面是向量空间表示的代数语言，很多问题可以在几何和代数之间互相转化。

1. 几何转化成代数

例 4.5 设 $\varphi \in \mathcal{L}(V,U)$, $\operatorname{rank}(\varphi) = r \geq 1$, 求证: 存在 $\varphi_i \in \mathcal{L}(V,U)$, $\operatorname{rank}(\varphi_i) = 1$ 且 $\varphi = \varphi_1+\varphi_2+\cdots+\varphi_r$.

证明: 设 $A \in M_{m\times n}(\mathbb{K}), \operatorname{rank}(A) = r \geq 1$.

只需证明 $\exists A_i \in M_{m\times n}(\mathbb{K}), \operatorname{rank}(A_i) = 1$ 且 $A = A_1+A_2+\cdots+A_r$.

$\because \exists$ 非异阵 P,Q s. t. $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q,$

设 $E_{ij}(i=1,2,\cdots,m; j=1,2,\cdots,n)$ 是第 (i,j) 元素为 1、其余所有元素为 0 的 $m\times n$ 矩阵,

令 $A_i = PE_{ii}Q, 1 \leq i \leq r$, 则 $\operatorname{rank}(A_i) = 1$ 且 $A = A_1+A_2+\cdots+A_r$. □

例 4.6 设 $\varphi \in \mathcal{L}(V^m), \varphi^m = 0, n = mq+1$, 求证: $\operatorname{rank}(\varphi) \leq n-q-1$.

证明: 设 $A \in M_n(\mathbb{K})$, $A^m = 0$, $n = mq + 1$, 只要证 $\operatorname{rank}(A) \le n - q - 1$.
用反证法: 设 $\operatorname{rank}(A) \ge n - q$.
$\because 0 = \operatorname{rank}(A^m) = \operatorname{rank}(A^{m-1} \cdot A) \overset{\text{Sylvester 不等式}}{\ge} \operatorname{rank}(A^{m-1}) + \operatorname{rank}(A) - n \ge \operatorname{rank}(A^{m-1}) - q$
$\Longrightarrow \operatorname{rank}(A^{m-1}) \le q$
$\therefore q \ge \operatorname{rank}(A^{m-1}) = \operatorname{rank}(A^{m-2} \cdot A) \ge \operatorname{rank}(A^{m-2}) + \operatorname{rank}(A) - n \ge \operatorname{rank}(A^{m-2}) - q \Longrightarrow$
$\operatorname{rank}(A^{m-2}) \le 2q \xRightarrow{\text{反复应用以上过程}} \cdots \Longrightarrow \operatorname{rank}(A) \le (m-1)q$
这与 $\operatorname{rank}(A) \ge n - q \xlongequal{n=mq+1} (m-1)q + 1$ 矛盾. $\square$

2. 代数转化成几何

例 4.7 设 A 是 n 阶方阵, 求证: $\operatorname{rank}(A^n) = \operatorname{rank}(A^{n+1}) = \operatorname{rank}(A^{n+2}) = \cdots$.

证明: 转化成几何问题：设 $\varphi \in \mathcal{L}(V^n)$, 则存在 $m \in [0, n]$ 使得:

$$\operatorname{Im}\varphi^m = \operatorname{Im}\varphi^{m+1} = \cdots, \tag{4.1}$$

$$\ker\varphi^m = \ker\varphi^{m+1} = \cdots, \tag{4.2}$$

$$V = \ker\varphi^m \oplus \operatorname{Im}\varphi^m. \tag{4.3}$$

$\because \varphi^2(v) = \varphi(\varphi(v)) \in \operatorname{Im}\varphi$, $\therefore V \supseteq \operatorname{Im}\varphi \supseteq \operatorname{Im}\varphi^2 \supseteq \operatorname{Im}\varphi^3 \supseteq \cdots$,
同理, $\because v \in \ker\varphi$ 时, $\varphi^2(v) = \varphi(\varphi(v)) = \varphi(0) = 0$,
$\therefore \ker\varphi \subseteq \ker\varphi^2 \subseteq \ker\varphi^3 \subseteq \cdots \subseteq V$.
$\because V \supseteq \operatorname{Im}\varphi \supseteq \operatorname{Im}\varphi^2 \supseteq \cdots \supseteq \operatorname{Im}\varphi^n \supseteq \operatorname{Im}\varphi^{n+1}$,
$\therefore \dim V = n \ge \dim\operatorname{Im}\varphi \ge \dim\operatorname{Im}\varphi^2 \ge \dim\operatorname{Im}\varphi^n \ge \dim\operatorname{Im}\varphi^{n+1} \ge 0$,
由抽屉原理，以上 $n+2$ 个整数维数必有两个相等,
不妨设 $\exists m \in [0, n]$, s. t. $\dim\operatorname{Im}\varphi^m = \dim\operatorname{Im}\varphi^{m+1}$,
上式等价于 $\operatorname{Im}\varphi^m = \operatorname{Im}\varphi^{m+1}$, 只要证 $\forall k \ge m$, $\operatorname{Im}\varphi^k = \operatorname{Im}\varphi^{k+1}$.
$\because \operatorname{Im}\varphi^{k+1} \subseteq \operatorname{Im}\varphi^k$, 任取 $\varphi^k(v) \in \operatorname{Im}\varphi^k$:
$\varphi^m(v) \in \operatorname{Im}\varphi^m = \operatorname{Im}\varphi^{m+1}$, $\exists u \in V$ s. t. $\varphi^m(v) = \varphi^{m+1}(u)$.
$\varphi^k(v) = \varphi^{k-m}(\varphi^m(v)) = \varphi^{k-m}(\varphi^{m+1}(u)) = \varphi^{k+1}(u) \in \operatorname{Im}\varphi^{k+1}$.
$\forall k \ge m$, $\dim\ker\varphi^k + \dim\operatorname{Im}\varphi^k = \dim V = n$.
上式中，$\forall k \ge m$, $\operatorname{Im}\varphi^k$ 的值不变，所以 $\ker\varphi^k$ 的值亦不变. 因此, 4.1、4.2两式得证.
下面证4.3式:
①先证 $\ker\varphi^m \cap \operatorname{Im}\varphi^m = 0$.
$\because$ 任取 $\alpha \in \ker\varphi^m \cap \operatorname{Im}\varphi^m$, 即 $\varphi^m(\alpha) = 0$, $\alpha = \varphi^m(\beta) \Longrightarrow 0 = \varphi^m(\alpha) = \varphi^{2m}(\beta)$
即 $\beta \in \ker\varphi^{2m} = \ker\varphi^m \Longrightarrow \varphi^m(\beta) = 0$.
②再证 $V = \ker\varphi^m + \operatorname{Im}\varphi^m$.
$\because$ 任取 $\alpha \in V$, $\varphi^m(\alpha) \in \operatorname{Im}\varphi^m = \operatorname{Im}\varphi^{2m}$, $\exists\beta \in V$ s. t. $\varphi^m(\alpha) = \varphi^{2m}(\beta)$
$\varphi^m(\alpha - \varphi^m(\beta)) = 0$, 设 $\alpha - \varphi^m(\beta) = \gamma$, 即 $\gamma \in \ker\varphi^m \Longrightarrow \alpha = \gamma + \varphi^m(\beta) \in \ker\varphi^m + \operatorname{Im}\varphi^m$.
以上证毕。此外，由①亦可结合维数公式 $\Longrightarrow V = \ker\varphi^m \oplus \operatorname{Im}\varphi^m$. $\square$

例 4.8 证明线性映射的维数公式：$\varphi \in \mathcal{L}(V^n, U^m)$, 则 $\dim\ker\varphi + \dim\operatorname{Im}\varphi = \dim V = n$.

证明: 任取 $\ker\varphi$ 的一组基 $\{e_1, e_2, \cdots, e_k\}$, 扩展为 V 的一组基 $\{e_1, e_2, \cdots, e_k, e_{k+1}, \cdots, e_n\}$, $\forall \alpha \in V$, 设 $\alpha = \sum_{i=1}^{n} \lambda_i e_i$, 则:
$\varphi(\alpha) = \sum_{i=1}^{n} \lambda_i \varphi(e_i) = \lambda_{k+1}\varphi(e_{k+1}) + \lambda_{k+2}\varphi(e_{k+2}) + \cdots + \lambda_n \varphi(e_n)$.
$\Longrightarrow \operatorname{Im}\varphi = L(\varphi(e_{k+1}), \varphi(e_{k+2}), \cdots, \varphi(e_n))$, 只要证: $\varphi(e_{k+1}), \varphi(e_{k+2}), \cdots, \varphi(e_n)$ 线性无关.
设 $c_{k+1}\varphi(e_{k+1}) + c_{k+2}\varphi(e_{k+2}) + \cdots + c_n\varphi(e_n) = 0$, 则 $\varphi(c_{k+1}e_{k+1} + c_{k+2}e_{k+2} + \cdots + c_n e_n) = 0$,
因此, $c_{k+1}e_{k+1} + c_{k+2}e_{k+2} + \cdots + c_n e_n \in \ker\varphi$
$\Longrightarrow c_{k+1}e_{k+1} + c_{k+2}e_{k+2} + \cdots + c_n e_n = -c_1e_1 - c_2e_2 - \cdots - c_ke_k$
$\Longrightarrow c_1e_1 + c_2e_2 + \cdots + c_ke_k + c_{k+1}e_{k+1} + \cdots + c_ne_n = 0$.
$\because \{e_1, e_2, \cdots, e_k, e_{k+1}, \cdots, e_n\}$ 是 V 的一组基, $\therefore c_1 = c_2 = \cdots = c_k = c_{k+1} = \cdots = c_n = 0$ □

例 4.9 设 A,B 都是 $m \times n$ 矩阵, 求证: 方程组 $Ax = 0$, $Bx = 0$ 同解的充分必要条件是存在 m 阶非异阵 P, 使 $B = PA$.

证明: 转换成几何表述:
设 $\varphi, \psi \in \mathcal{L}(V^n, U^m)$, $\ker\varphi = \ker\psi$ 的充分必要条件是 $\exists$ 自同构 $\xi \in \mathcal{L}(U)$ s. t. $\psi = \xi\varphi$. □

5 多项式

5.1 复系数多项式

定理 5.1 (代数基本定理) 每个次数大于零的复数域上的多项式都至少有一个复数根.

推论 5.1 设 $n \geq 1$，以下记法等价:

(1) 任一 n 次复系数多项式至少有一个复根;

(2) 复系数不可约多项式都是一次多项式;

(3) 任一复系数多项式都是一次多项式的乘积;

(4) 任一 n 次复系数多项式恰有 n 个复根（计重根）.

证明: (1)$\Longrightarrow$(2): 设 $P(x)$ 在 $\mathbb{C}$ 上不可约.
由 (1) 可知，b 为 $P(x)$ 的根，则 $(x-b) \mid P(x)$, 即 $P(x)=(x-b)q(x)$,
由 $P(x)$ 不可约可知, $q(x)=c \neq 0$, 从而 $P(x)=c(x-b)$ 是一次多项式.

(2)$\Longrightarrow$(3): 因式分解定理即得.

(3)$\Longrightarrow$(4):$f(x)=c(x-b_1)(x-b_2)\cdots(x-b_n)$, 则 $f(x)$ 的根为 $b_1, b_2, \cdots, b_n$.

(4)$\Longrightarrow$(1): 显然. □

定理 5.2 (Osada) $f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 为整系数多项式，且 $|a_0|>1+\sum\limits_{i=1}^{n-1}|a_i|$, 其中 $|a_0|$ 为素数, 则 $f(x)$ 在有理数域上不可约.

5.2 结式和判别式

定理 5.3 $f(x)$ 与 $g(x)$ 在 $\mathbb{C}$ 上有公根 $\Longleftrightarrow R(f,g)=0$, $\quad (f(x),g(x))=1 \Longleftrightarrow R(f,g)\neq 0$.

定理 5.4 设 $f(x)$ 的 n 个根为 $x_1, x_2, \cdots, x_n$, $g(x)$ 的 m 个根为 $y_1, y_2, \cdots, y_n$, 则

$$R(f,g)=a_0^m b_0^n \prod_{i=1}^{n}\prod_{j=1}^{m}(x_i-y_j) \tag{†}$$

证明: Step 1 首一化: 令 $f(x)=a_0f_1(x), g(x)=b_0g_1(x)$, f_1, g_1 均为首一多项式.

$$R(f,g)=\begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_n & 0 & 0 & \cdots & 0 \\ 0 & a_0 & a_1 & \cdots & a_{n-1} & a_n & 0 & \cdots & 0 \\ 0 & 0 & a_0 & \cdots & a_{n-2} & a_{n-1} & a_n & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & a_0 & a_1 & \cdots & a_n \\ b_0 & b_1 & b_2 & \cdots & b_m & 0 & 0 & \cdots & 0 \\ 0 & b_0 & b_1 & \cdots & b_{m-1} & b_m & 0 & \cdots & 0 \\ 0 & 0 & b_0 & \cdots & b_{m-2} & b_{m-1} & b_m & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & b_0 & b_1 & \cdots & b_m \end{vmatrix}$$

从 $R(f,g)$ 前 m 行各提取一个 a_0, 从 $R(f,g)$ 后 n 行各提取一个 b_0, 得 $R(f,g)=a_0^m b_0^n R(f_1,g_1)$.
所以只要证明 $R(f_1,g_1)=\prod\limits_{i=1}^{n}\prod\limits_{j=1}^{m}(x_i-y_j)$.

Step 2: 设 $f_1(x)=(x-x_1)(x-x_2)\cdots(x-x_n)$, $g_1(x)=(x-y_1)(x-y_2)\cdots(x-y_m)$.
若 $\exists 1\le i\le n, 1\le j\le m\,\mathrm{s.t.}\ x_i=y_j$, 由定理5.3$\Longrightarrow$(†)式成立.
下面设 $x_i\ne y_j, \forall 1\le i\le n, 1\le j\ne m$ 且 $x_1,x_2,\cdots,x_n$ 互不相同, $y_1,y_2,\cdots,y_m$ 互不相同,
设 $R(f_1,g_1)$ 在是 f_1 和 g_1 结式的方阵, 则

$$R(f_1,g_1)\begin{pmatrix} x_1^{m+n-1} & x_2^{m+n-1} & \cdots & x_n^{m+n-1} & y_1^{m+n-1} & y_2^{m+n-1} & \cdots & y_m^{m+n-1} \\ x_1^{m+n-2} & x_2^{m+n-2} & \cdots & x_n^{m+n-2} & y_1^{m+n-2} & y_2^{m+n-2} & \cdots & y_m^{m+n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1 & x_2 & \cdots & x_n & y_1 & y_2 & \cdots & y_m \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}=$$

$$\begin{pmatrix} x_1^{m-1}f_1(x_1) & x_2^{m-1}f_1(x_2) & \cdots & x_n^{m-1}f_1(x_n) & y_1^{m-1}f_1(y_1) & y_2^{m-1}f_1(y_2) & \cdots & y_m^{m-1}f_1(y_m) \\ x_1^{m-2}f_1(x_1) & x_2^{m-2}f_1(x_2) & \cdots & x_n^{m-2}f_1(x_n) & y_1^{m-2}f_1(y_1) & y_2^{m-2}f_1(y_2) & \cdots & y_m^{m-2}f_1(y_m) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ f_1(x_1) & f_1(x_2) & \cdots & f_1(x_n) & f_1(y_1) & f_1(y_2) & \cdots & f_1(y_m) \\ x_1^{n-1}g_1(x_1) & x_2^{n-1}g_1(x_2) & \cdots & x_n^{n-1}g_1(x_n) & y_1^{n-1}g_1(y_1) & y_2^{n-1}g_1(y_2) & \cdots & y_m^{n-1}g_1(y_m) \\ x_1^{n-2}g_1(x_1) & x_2^{n-2}g_1(x_2) & \cdots & x_n^{m-2}g_1(x_n) & y_1^{n-2}g_1(y_1) & y_2^{n-2}g_1(y_2) & \cdots & y_m^{n-2}g_1(y_m) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_1(x_1) & g_1(x_2) & \cdots & g_1(x_n) & g_1(y_1) & g_1(y_2) & \cdots & g_1(y_m) \end{pmatrix}$$

$\because x_i, y_j$ 分别是 $f_1(x), g_1(x)$ 的解, $\therefore f_1(x_i)=0, g_1(y_j)=0$, 但是 $f_1(y_j)\neq 0, g_1(x_i)\neq 0$. 于是,

$$
R(f_1,g_1)\begin{pmatrix}
x_1^{m+n-1} & x_2^{m+n-1} & \cdots & x_n^{m+n-1} & y_1^{m+n-1} & y_2^{m+n-1} & \cdots & y_m^{m+n-1}\\
x_1^{m+n-2} & x_2^{m+n-2} & \cdots & x_n^{m+n-2} & y_1^{m+n-2} & y_2^{m+n-2} & \cdots & y_m^{m+n-2}\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
x_1 & x_2 & \cdots & x_n & y_1 & y_2 & \cdots & y_m\\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}
$$

$$
=\begin{pmatrix}
0 & \cdots & 0 & y_1^{m-1}f_1(y_1) & \cdots & y_m^{m-1}f_1(y_m)\\
\vdots & & \vdots & \vdots & & \vdots\\
0 & \cdots & 0 & f_1(y_1) & \cdots & f_1(y_m)\\
x_1^{n-1}g_1(x_1) & \cdots & x_n^{n-1}g_1(x_n) & 0 & \cdots & 0\\
\vdots & & \vdots & \vdots & & \vdots\\
g_1(x_1) & \cdots & g_1(x_n) & 0 & \cdots & 0
\end{pmatrix} \tag{5.1}
$$

(5.1)式左边是结式方阵和一个降幂的 Vander Monde 矩阵的乘积，其乘积的行列式值为:

$$
R(f_1,g_1)\prod_{1\le i<k\le n}(x_i-x_k)\prod_{1\le j<l\le m}(y_j-y_l)\prod_{i=1}^{n}\prod_{j=1}^{m}(x_i-y_j) \tag{5.2}
$$

(5.1)式右边是一个分块对角方阵，前 n 列分别提出 $g_1(x_i)$, 后 m 列分别提出 $f_1(y_j)$ 之后，剩下非零部分是两个降幂 Vander Monde 矩阵, 按前 m 行、n 列进行 Laplace 展开，其行列式值为:

$$
(-1)^{mn}\prod_{i=1}^{n}g_1(x_i)\prod_{j=1}^{m}f_1(y_j)\prod_{1\le i<k\le n}(x_i-x_k)\prod_{1\le j<l\le m}(y_j-y_l) \tag{5.3}
$$

(5.2) = (5.3), 等式两边约去非零的相同部分后，得:

$$
R(f_1,g_1)\prod_{i=1}^{n}\prod_{j=1}^{m}(x_i-y_j)=(-1)^{mn}\prod_{i=1}^{n}g_1(x_i)\prod_{j=1}^{m}f_1(y_j) \tag{5.4}
$$

(5.4)式右边里，

$$
\begin{aligned}
(-1)^{mn}\prod_{j=1}^{m}f_1(y_j) &= \prod_{j=1}^{m}((-1)^n f_1(y_j))\\
&= \prod_{j=1}^{m}((-1)^n(y_j-x_1)\cdots(y_j-x_n))\\
&= \prod_{j=1}^{m}((x_1-y_j)\cdots(x_n-y_j))\\
&= \prod_{i=1}^{n}\prod_{j=1}^{m}(x_i-y_j)
\end{aligned} \tag{5.5}
$$

将(5.5)式代入(5.4)式右边并约去非零部分，得

$$
\begin{aligned}
R(f_1, g_1) &= \prod_{i=1}^{n} g_1(x_i) \\
&= \prod_{i=1}^{n} ((x_i - y_1) \cdots (x_i - y_m)) \\
&= \prod_{i=1}^{n} \prod_{j=1}^{m} (x_i - y_j)
\end{aligned}
$$

Step 3: 摄动. $\exists c_1, \cdots, c_n, d_1, \cdots, d_n \in \mathbb{C}$
s. t. $\forall 0 < t \ll 1, x_1 + c_1 t, \cdots, x_n + c_n t, y_1 + d_1 t, \cdots, y_m + d_m t$ 互不相同.
设 $f_t(x) = (x - x_1 - c_1 t) \cdots (x - x_n - c_n t), g_t(x) = (x - y_1 - d_1 t) \cdots (x - y_m - d_m t)$,
$\forall 0 < t \ll 1, f_t(x), g_t(x)$ 系数是 t 的多项式，且 $f_0(x) = f_1(x), g_0(x) = g_1(x)$, 由 Step 2 知,

$$
R(f_t, g_t) = \prod_{i=1}^{n} \prod_{j=1}^{m} (x_i + c_i t - y_j - d_j t)
$$

令 $t \longrightarrow 0^+$, 得

$$
R(f_1, g_1) = \prod_{i=1}^{n} \prod_{j=1}^{m} (x_i - y_j)
$$

$\square$

6 特征值

6.1 对角化

定理 6.1 设 φ 是 n 维线性空间 V 上的线性变换, 则存在 V 的一组基, 使得 φ 在这组基下的表示矩阵为对角阵的充分必要条件是 φ 有 n 个线性无关的特征向量 (这样的线性变换成为可对角化线性变换).

证明: 必要性. 设 φ 可对角化，根据定义即存在 V 的一组基 $\{e_1, e_2, \cdots, e_n\}$ 使得 φ 在此基下的表示矩阵为对角阵 $\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 则

$$(\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n)) = (e_1, e_2, \cdots, e_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix}$$

$$\Longrightarrow \varphi(e_1) = \lambda_1 e_1, \varphi(e_2) = \lambda_2 e_2, \cdots, \varphi = \lambda_n e_n$$

$$\xRightarrow{e_i\text{为基向量}} \varphi\text{有 } n \text{ 个线性无关的特征向量}.$$

充分性. 设 φ 有 n 个线性无关的特征向量 $e_1, e_2, \cdot, e_n$, 即 $\varphi(e_i) = \lambda_i e_i, (1 \le i \le n)$, 并且 $\{e_1, e_2, \cdot, e_n$ 是 V 的一组基, 所以

$$(\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n)) = (e_1, e_2, \cdots, e_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix}$$

$$\Longrightarrow \varphi\text{的表示矩阵为对角阵}$$

故 φ 可对角化. □

定理 6.2 若 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 为数域 $\mathbb{K}$ 上 n 维线性空间 V 上线性变换 φ 的不同特征值, 则 $V_1 + V_2 + \cdots + V_k = V_1 \oplus V_2 \oplus \cdots \oplus V_k$.

推论 6.1 线性变换 φ 属于不同特征值的特征向量必线性无关.

证明: 设 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 为 φ 的不同特征值, $v_1, v_2, \cdots, v_k$ 为对应的特征向量.
令 $c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0$, 那么, $c_1v_1 \in V_1, c_2v_2 \in V_2, \cdots, c_kv_k \in V_k$,
由定理6.2可知, $c_1v_1 = c_2v_2 = \cdots = c_kv_k = 0$, 由于 v_i 都是特征向量, 故 $v_i \ne 0$,
所以,$c_1 = c_2 = \cdots = c_k = 0$, 因此,$v_1, v_2, \cdots, v_k$ 线性无关. □

6.2 极小多项式与 Cayley-Hamilton 定理

命题 6.1 任一 n 阶矩阵的极小多项式必唯一.

命题 6.2 相似的矩阵具有相同的极小多项式.

命题 6.3 设 A 是一个分块对角阵

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \cdots & \\ & & & A_k \end{pmatrix},$$

其中 A_i 都是方阵, 则 A 的极小多项式等于诸 A_i 的极小多项式之最小公倍式.

例 6.1 设 A 的全体不同特征值为 $\lambda_1, \cdots, \lambda_k$, 求证: 若 A 可对角化, 则极小多项式为 $m(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_k)$.

证明: 存在非异矩阵 P, 使得

$$P^{-1}AP = \begin{pmatrix} \lambda_1 I & & & \\ & \lambda_2 I & & \\ & & \ddots & \\ & & & \lambda_k I \end{pmatrix} = B$$

由命题6.2可知,A 与 B 有相同极小多项式，即

$$m(x) = m_B(x).$$

因为 B 是分块对角阵, 由命题6.3可推出

$$m_B(x) = [x-\lambda_1, x-\lambda_2, \cdots, x-\lambda_k] = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_k).$$

□

注 A 可对角化 $\Longleftrightarrow$ A 的极小多项式无重根.

命题 6.4 设 $A = \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ & \lambda_2 & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix}$, 则

$$(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n) = 0, \tag{6.1}$$

即 A 适合其特征多项式

$$f(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n). \tag{6.2}$$

证明: 设 $e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ 是第 i 行为 1 的单位列向量, 可得如下 n 个等式:

$$
\begin{aligned}
&Ae_1 = \lambda_1 e_1, && (6.3)\\
&Ae_1 = a_{12}e_1 + \lambda_2 e_2, && (6.4)\\
&\cdots\cdots \\
&Ae_i = a_{1i}e_1 + \cdots + a_{i-1,i}e_{i-1} + \lambda_i e_i, && (6.5)\\
&\cdots\cdots \\
&Ae_n = a_{1n}e_1 + \cdots + a_{n-1,n}e_{n-1} + \lambda_n e_n. && (6.6)
\end{aligned}
$$

$\because f(A)g(A) = g(A)f(A)$,
$\therefore (A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n) = (A-\lambda_{i+1}I_n)(A-\lambda_{i+2}I_n)\cdots(A-\lambda_n I_n)(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_i I_n)$,
只要证

$$(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n)e_i = 0, \forall 1 \le i \le n \tag{6.7}$$

只要证

$$(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_i I_n)e_i = 0, \forall 1 \le i \le n \tag{6.8}$$

对 i 进行归纳:
$i=1$ 时, 由(6.3)式可知,$(A-\lambda_1 I_n)e_i = 0$ 成立.
现假设 $i-1$ 时结论成立, 则 i 时, 由(6.5)可知,

$$Ae_i - \lambda_i e_i = a_{1i}e_1 + \cdots + a_{i-1,i}e_{i-1} \tag{6.9}$$

即

$$(A-\lambda_i I_n)e_i = a_{1i}e_1 + \cdots + a_{i-1,i}e_{i-1} \tag{6.10}$$

将(6.10)式代入(6.8)式, 得

$$
\begin{aligned}
&(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_i I_n)e_i \\
=&(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_{i-1} I_n)(a_{1i}e_1 + \cdots + a_{i-1,i}e_{i-1}) \qquad (6.11)
\end{aligned}
$$

将(6.11)式前 $i-1$ 个因式和第 i 个因式中的每一项相乘, 可得

$$
\begin{aligned}
&(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_i I_n)e_i \\
=&\sum_{j=1}^{i-1} a_{ji}(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_{i-1} I_n)e_j \qquad (6.12)
\end{aligned}
$$

由归纳假设可知, (6.12)式的 $i-1$ 项全部为 0, 从而(6.8)式为 0. $\square$

定理 6.3 (Cayley-Hamilton(凯莱-哈密顿) 定理) 设 A 是数域 $\mathbb{K}$ 上的 n 阶矩阵, $f(x)$ 是 A 的特征多项式, 则 $f(A)=O$.

证明: $\exists$ 非异阵 $P\,\text{s.t.}\ P^{-1}AP=B=\begin{pmatrix}\lambda_1 & a_{12} & \cdots & a_{1n}\\ & \lambda_2 & \cdots & a_{2n}\\ & & \ddots & \vdots\\ & & & \lambda_n\end{pmatrix}$,
其特征多项式为 $f(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$.
由命题(6.4)可知 $f(B)=O$, 则
$f(A)=f(PBP^{-1})=Pf(B)P^{-1}=O$. □

推论 6.2 条件和假设同定理(6.3), 则 $f(\lambda)\mid m(\lambda)^n$.

证明: $f(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_k)^{m_k}$, 其中 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 是 A 的全体不同特征值,
$m(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_k)^{r_k}$, 其中 $r_1,r_2,\cdots,r_k\in\mathbb{Z}^+$.
$m_1+m_2+\cdots+m_k=n\Longrightarrow m_i\le n\le n\cdot r_i, 1\le i\le k\Longrightarrow f(\lambda)\mid m(\lambda)$. □

推论 6.3 (Cayley-Hamilton 定理 (几何领域)) 设 φ 是 n 维线性空间 V 上的线性变换, $f(x)$ 是 φ 的特征多项式, 则 $f(\varphi)=0$.

证明: 设 φ 的表示矩阵为 $A\in M_n(\mathbb{K})$.

先证 A 可对角化的情形:
设 $\Lambda=P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n\end{pmatrix}$, 其特征多项式为 $f(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$.
$f(\Lambda)=$
$\begin{pmatrix}0 & & & \\ & \lambda_2-\lambda_1 & & \\ & & \cdots & \\ & & & \lambda_n-\lambda_1\end{pmatrix}\begin{pmatrix}\lambda_1-\lambda_2 & & & \\ & 0 & & \\ & & \cdots & \\ & & & \lambda_n-\lambda_2\end{pmatrix}\cdots\begin{pmatrix}\lambda_1-\lambda_n & & & \\ & \lambda_2-\lambda_n & & \\ & & \cdots & \\ & & & 0\end{pmatrix}$
$=0$,
$f(A)=f(P\Lambda P^{-1})=Pf(\Lambda)P^{-1}=O$.

再证一般情形 (摄动法):
$\exists$ 非异阵 $P\,\text{s.t.}\ P^{-1}AP=\begin{pmatrix}\lambda_1 & * & * & *\\ & \lambda_2 & * & *\\ & & \cdots & *\\ & & & \lambda_n\end{pmatrix}$,
$\exists c_1,c_2,\cdots,c_n\in\mathbb{C}\,\text{s.t.}\ \forall 0<t\ll 1, \lambda_1+c_1t,\lambda_2+c_2t,\cdots,\lambda_n+c_nt$ 互不相同.

构造矩阵 $A_t = P\begin{pmatrix} \lambda_1+c_1t & * & * & * \\ & \lambda_2+c_2t & * & * \\ & & \cdots & * \\ & & & \lambda_n+c_nt \end{pmatrix}P^{-1}$,
当 $0<t\ll 1$ 时, A_t 可对角化, 当 $t=0$ 时, $A_0=A$.
$\forall 0<t\ll 1, (A_t-(\lambda_1+c_1t)I_n)(A_t-(\lambda_2+c_2t)I_n)\cdots(A_t-(\lambda_n+c_nt)I_n)=0$,
令 $t\longrightarrow 0$, 得 $(A-\lambda_1I_n)(A-\lambda_2I_n)\cdots(A-\lambda_nI_n)=0$ □

6.3 特征值的估计

例 6.2 已知 $A^2-A-3I_n=0$, 求证:$A-2I_n$ 可逆.

证明:

方法一、凑因子法:
$\because (A-2I_n)(A+I_n)=I_n, \therefore A-2I_n$ 可逆.

方法二、线性方程组求解法:
只要证 $(A-2I_n)x=0$ 只有零解.

方法三、互素多项式法:
$(x^2-x-3,x-2)=1$, A 代入 x^2-x-3 等于零, 则 A 代入 $x-2$ 必不为零.

方法四、特征值法:
反正法：设 $A-2I_n$ 不可逆 $\Longrightarrow$ 2 是 A 的特征值,
将 $\lambda=2$ 代入 A^2-A-3I_n 得 $2^2-2-3\neq 0$ 产生矛盾, 从而 $A-2I_n$ 可逆. □

例 6.3 A 是 4 阶方阵且满足 $\operatorname{tr}(A^i)=i, i=1,2,3,4$, 求 $|A|$.

解: 方法一：利用牛顿公式
设 A 的特征值为 $\lambda_1,\lambda_2,\lambda_3,\lambda_4$, 则 A^i 的特征值为 $\lambda_1^i,\lambda_2^i,\lambda_3^i,\lambda_4^i \Longrightarrow \sum\limits_{k=1}^{4}\lambda_k^i=i$.
根据上一章 Newton 公式:

$$s_k-s_{k-1}\sigma_1+s_{k-2}\sigma_2-\cdots+(-1)^kk\sigma_k=0, k\leq n=4 \tag{6.13}$$

上式中, $s_1=1, s_2=2, s_3=3, s_4=4$, 只要求 $|A|=\lambda_1\lambda_2\lambda_3\lambda_4=\sigma_4$.

$$\begin{aligned}
&\sigma_1=s_1=1\\
&s_2-s_1\sigma_1+2\sigma_2=0\Longrightarrow\sigma_2=-\frac{1}{2}\\
&s_3-s_2\sigma_2+s_1\sigma_1-3\sigma_3=0\Longrightarrow\sigma_3=\frac{1}{6}\\
&s_4-s_3\sigma_3+s_2\sigma_2-s_1\sigma_1+4\sigma_4=0\Longrightarrow\sigma_4=\frac{1}{24}
\end{aligned}$$

方法二：利用教材 P250 习题 6(1) 结论

$$\because \sigma_k = \frac{1}{k!}\begin{vmatrix} s_1 & 1 & 0 & \cdots & 0 \\ s_2 & s_1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ s_{k-1} & s_{k-2} & s_{k-3} & \vdots & k-1 \\ s_k & s_{k-1} & s_{k-2} & \cdots & s_1 \end{vmatrix}, \qquad \therefore \sigma_4 = \frac{1}{4!}\begin{vmatrix} 1 & 1 & 0 & 0 \\ 2 & 1 & 2 & 0 \\ 3 & 2 & 1 & 3 \\ 4 & 3 & 2 & 1 \end{vmatrix} = \frac{1}{24}.$$

例 6.4 $A^{m\times m}, B^{n\times n}$ 无公共特征值, 求证: 矩阵方程 $AX = XB$ 只有零解.

证明:

证法一: 设 $f(x) = |\lambda I - A|$ 为 A 的特征多项式, 由 Cayley-Hamilton 定理 $\Longrightarrow f(A) = 0$.
$A^2X = A(AX) = A(XB) = (AX)B = XB^2$, 因此 $A^nX = XB^n, n \in N$,
因为 $f(A)$ 是关于 A 的多项式,$f(B)$ 是关于 B 的多项式, 所以

$$0 = f(A)X = Xf(B) \tag{6.14}$$

设 B 的特征值为 $\mu_1, \mu_2, \cdots, \mu_n$, 则 $f(B)$ 的特征值为 $f(\mu_1), f(\mu_2), \cdots, f(\mu_n)$.
$\because A, B$ 无公共特征值,
$\therefore f(\mu_1), f(\mu_2), \cdots, f(\mu_n)$ 皆不为零 $\Longrightarrow f(B)$ 可逆 $\Longrightarrow$(6.14)式只有零解.

证法二: 设 B 的特征多项式为 $g(\lambda)$,
A, B 无公共特征值 $\Longrightarrow f(\lambda), g(\lambda)$ 无公共根 $\Longrightarrow (f(\lambda), g(\lambda)) = 1$
因而 $\exists$ 非零多项式 $u(\lambda), v(\lambda)$ s. t.

$$f(\lambda)u(\lambda) + g(\lambda)v(\lambda) = 1.$$

将 $\lambda = B$ 代入上式, 即

$$f(B)u(B) + g(B)v(B) = I \tag{6.15}$$

$\because g(\lambda)$ 为 B 的特征多项式,
$\therefore$(6.15)式中 $g(B) = O \Longrightarrow f(B)u(B) = I \Longrightarrow f(B)$ 可逆. □

应用 6.1 A, B 为 n 阶方阵且特征值全大于零, $A^2 = B^2$, 求证: $A = B$.

证明: $A(A-B) = A^2 - AB = B^2 - AB = (A-B)(-B)$
上式中,$A, -B$ 必无公共特征值, 由例6.4 $\Longrightarrow A - B = 0$. □

应用 6.2 A 为 n 阶方阵且特征值全为偶数, 求证：矩阵方程 $X + AX = XA^2$ 只有零解.

证明: $X + AX = XA^2 \Longrightarrow (A+I)X = XA^2$,
$\because A$ 的特征值都是偶数,$\therefore A + I$ 的特征值为奇数, A^2 的特征值为偶数,
因此 $A + I$ 与 A^2 无公共特征值, 由例6.4可知,
$(A+I)X = XA^2$ 只有零解, 从而 $X + AX = XA^2$ 只有零解. □

应用 6.3 A 是 n 阶方阵且适合多项式 $a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$, 其中 $|a_m| > \sum\limits_{i=0}^{m-1} |a_i|$, 求证: $2X + AX = XA$ 只有零解.

证明: 任取 A 的特征值 λ_0, 断言: $|\lambda_0| < 1$, 用反正法: 设 $|\lambda_0| \geq 1$.
由题意知: $a_m\lambda_0^m + a_{m-1}\lambda_0^{m-1} + \cdots + a_1\lambda_0 + a_0 = 0$, 则

$$a_m = -\frac{a_{m-1}}{\lambda_0} - \frac{a_{m-2}}{\lambda_0^2} - \cdots - \frac{a_1}{\lambda_0^{m-1}} - \frac{a_0}{\lambda_0^m} \Longrightarrow |a_m| \leq \sum_{i=0}^{m-1} |a_i|, \text{与题意矛盾}.$$

$\therefore A$ 的所有特征值必小于 1.
如图6.1所示, A 的所有特征值都位于左边的单位圆内, 且不包括圆周; $A + 2I_n$ 的所有特征值都位于右边的单位圆内, 且不包括圆周, 即左边的单位圆向右平移两个单位. 由于两个圆想切且不包括圆周, 因此 A 和 $A + 2I_n$ 没有公共特征值. 根据例6.4可知, $(A + 2I_n)X = XA$ 只有零解. □

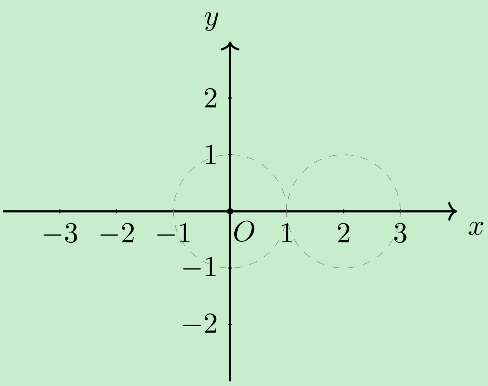

图 6.1

应用 6.4 $A^{m\times m}, B^{n\times n}$ 是数域 $\mathbb{K}$ 上的方阵且无公共特征值, 求证: 对 $\forall C \in M_{m\times n}(\mathbb{K})$, 矩阵方程 $AX - XB = C$ 有唯一解.

证明: 构建一个线性变换:

$$\begin{aligned} \varphi : M_{m\times n}(\mathbb{K}) &\mapsto M_{m\times n}(\mathbb{K}) \\ X &\mapsto AX - XB \end{aligned}$$

由题意 A, B 无公共特征值, 根据例6.4可知:
$AX = XB$ 只有零解, 即 $AX - XB = 0$ 只有零解, 从而只有 $X = 0$ 与 $AX - XB = 0$ 相对应,
即 $\ker\varphi = 0 \Longrightarrow \varphi$ 为单射, 由于 φ 为线性变换, 单射线性变换必为满射, 从而 φ 为线性同构,
因此, 对 $\forall C \in M_{m\times n}(\mathbb{K})$, 当 $AX - XB = C$ 时, 必有唯一的 X 与其对应,
即 $AX - XB = C$ 有唯一解. □

例 6.5 $A^{m\times m}, B^{n\times n}$ 是数域 $\mathbb{K}$ 上的方阵且无公共特征值, 若 A, B 可对角化, 则 $M = \left(\begin{smallmatrix} A & C \\ O & B \end{smallmatrix}\right)$ 可对角化.

证明: 由应用6.4 $\Longrightarrow \exists X_0\,\mathrm{s.t.}\ AX_0 - X_0B = C$, 则

$$\begin{pmatrix} I_m & X_0 \\ O & I_n \end{pmatrix}\begin{pmatrix} A & C \\ O & B \end{pmatrix}\begin{pmatrix} I_m & -X_0 \\ O & I_n \end{pmatrix} = \begin{pmatrix} A & C + X_0B - AX_0 \\ O & B \end{pmatrix}$$

上式等号右边, $C + X_0B - AX_0 = 0$, 所以, $M \sim \left(\begin{smallmatrix} A & O \\ O & B \end{smallmatrix}\right)$, A, B 可对角化, 故 M 也可对角化. □

例 6.6 A 为非异的循环矩阵, 求证: A^{-1} 也是循环矩阵.

证明:

证法一:

单位循环矩阵为:

$$J = \begin{pmatrix} & I_{n-1} \\ 1 & \end{pmatrix} \tag{6.16}$$

依题意可知:

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_n & a_1 & \cdots & a_{n-1} \\ \vdots & \vdots & & \vdots \\ a_2 & \cdots & a_n & a_1 \end{pmatrix} = a_1I_n + a_2J + \cdots + a_nJ^{n-1} \tag{6.17}$$

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

J 的特征多项式为 $f(\lambda) = \lambda^n - 1$, J 的 h 特征值为 $w_k = e^{\frac{2k\pi i}{n}}$, $0 \le k \le n-1$,

$A = g(J)$, A 的特征值为 $g(w_k), k = 0, 1, \cdots, n-1$,

又因为 A 非异, 所以 $g(w_i) \neq 0$, 但 $f(w_i) = 0$,

因此,

$$\begin{aligned} & f(x), g(x)\text{没有公共根} \Longrightarrow (f(x), g(x)) = 1 \\ \Longrightarrow & \exists\text{非零多项式}u, v\,\mathrm{s.t.}\ f(x)u(x) + g(x)v(x) = 1 \end{aligned} \tag{6.18}$$

将 $x = J$ 代入(6.18)式, 得

$$f(J)u(J) + g(J)v(J) = I_n \tag{6.19}$$

上式中, $f(J) = 0, g(J) = A$, 因而 $Av(J) = I_n \Longrightarrow A^{-1} = v(J)$ 为循环矩阵.

证法二:

A 非异 $\xRightarrow{Cayley-Hamilton\text{定理}} A^{-1} = h(A)$, $h(A)$ 为 A 的 $n-1$ 次多项式, 见 P282 习题 1.

$h(A) = h(g(J))$ 仍为循环矩阵. □

7 相似标准型

7.1 多项式矩阵

定义 7.1 $M(\lambda)=M_m\lambda^m+M_{m-1}\lambda^{m-1}+\cdots+M_1\lambda+M_0$, 其中 $M_i\in M_{r\times s}(\mathbb{K})$, 称 $M(\lambda)$ 为**矩阵多项式**. 若 $M_m\neq 0$, 定义 $\deg M(\lambda)=m$, 约定 $\deg 0=-\infty$.

$$N(\lambda)=N_n\lambda^n+N_{n-1}\lambda^{n-1}+\cdots+N_1\lambda+N_0, N_n\neq 0,\text{则}$$
$$M(\lambda)=N(\lambda)\overset{\text{def}}{\Longleftrightarrow} m=n, M_i=N_i, \forall 0\leq i\leq m$$

设 $m\geq n$,
定义加法:$M(\lambda)+N(\lambda)=M_m\lambda^m+\cdots+M_{n+1}\lambda^{n+1}+(M_n+N_n)\lambda^n+\cdots+(M_0+N_n)$
定义数乘:$k\cdot M(\lambda)=k\cdot M_m\lambda^m+\cdots+k\cdot M_0$
定义乘法:$M(\lambda)\cdot N(\lambda)=P(\lambda)=\sum P_k\lambda^k, P_k=\sum\limits_{i+j=k}M_iN_j$

引理 7.1 $\deg(M(\lambda)\cdot N(\lambda))\leq \deg M(\lambda)+\deg N(\lambda)$, 若 M_m 或 N_n 是可逆阵，则等号成立.

证明: $M(\lambda)\cdot N(\lambda)$ 的首项是 $M_m\cdot N_n\cdot\lambda^{m+n}$, 若 M_m 或 N_n 可逆，则 $M(\lambda)\cdot N(\lambda)\neq 0$. □

定理 7.1 设 A,B 是数域 $\mathbb{K}$ 上的矩阵, 则 A,B 相似的充分必要条件是 $\lambda-$ 矩阵 $\lambda I-A$ 与 $\lambda I-B$ 相抵.

7.2 矩阵的法式

定理 7.2 设 $A(\lambda)$ 是一个 n 阶 $\lambda-$ 矩阵, 则 $A(\lambda)$ 相抵于对角阵

$$\mathrm{diag}\{d_1(\lambda),d_1(\lambda),\cdots,d_r(\lambda);0,\cdots,0\},$$

其中 $d_i(\lambda)$ 是非零首一多项式且 $d_i(\lambda)\mid d_{i+1}(\lambda)(i=1,2,\cdots,r-1)$.

引理 7.2 设 $A(\lambda),B(\lambda)$ 为 n 阶 $\lambda-$ 矩阵, 则:

(1) $|A(\lambda)\cdot B(\lambda)|=|A(\lambda)|\cdot|B(\lambda)|$;

(2) $A(\lambda)A(\lambda)^*=A(\lambda)^*A(\lambda)=|A(\lambda)|\cdot I_n$.

证明: (1) 令 $f(\lambda)=|A(\lambda)\cdot B(\lambda)|-|A(\lambda)|\cdot|B(\lambda)|$ 是 λ 的多项式,
对任意的 $a\in\mathbb{K}$, $f(a)=|A(a)\cdot B(a)|-|A(a)|\cdot|B(a)|=0$, 从而 $f(\lambda)\equiv 0$.

(2) 令 $(f_{ij}(\lambda))_{n\times n}=A(\lambda)\cdot A(\lambda)^*-|A(\lambda)|\cdot I_n, f_{ij}(\lambda)\in\mathbb{K}[\lambda]$,
$\forall a\in\mathbb{K}, (f_{ij}(a))=A(a)\cdot A(a)^*-|A(a)|\cdot I_n=0\Longrightarrow f_{ij}(\lambda)\equiv 0,\forall i,j$ □

定理 7.3 设 $A(\lambda)$ 是 n 阶 $\lambda-$ 矩阵, 则下列结论等价:

(1) $A(\lambda)$ 是可逆 $\lambda-$ 阵;

(2) $|A(\lambda)|$ 是非零常数;

(3) $A(\lambda)$ 的相抵标准型为 I_n;

(4) $A(\lambda)$ 只通过初等行 (列) 变换可变为 I_n;

(5) $A(\lambda)$ 是初等 $\lambda-$ 矩阵的乘积.

证明: (1) $\Longrightarrow$ (2): $\exists B(\lambda)\,\text{s.t.}\ A(\lambda)B(\lambda)=B(\lambda)A(\lambda)=I_n$,
等式两边同取行列式, 得: $|A(\lambda)\cdot B(\lambda)|=|A(\lambda)|\cdot|B(\lambda)|=|I_n|=1$,
因为上式左边是两个关于 λ 的多项式相乘, 右边是 1,
因此 $|A(\lambda)|$ 和 $|B(\lambda)|$ 只能是非零常数.

(2) $\Longrightarrow$ (3): 由定理7.2 $\Longrightarrow \exists\lambda-$ 矩阵 $P(\lambda),Q(\lambda)\,\text{s.t.}\ P(\lambda)A(\lambda)Q(\lambda)=\text{diag}\{d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda);0,\cdots,0\}$, 其中, $d_i(\lambda)$ 皆为关于 λ 的首一多项式, 且 $d_1(\lambda)\mid d_2(\lambda)\mid\cdots\mid d_r(\lambda)$, $P(\lambda)$, $Q(\lambda)$ 为初等 $\lambda-$ 矩阵的乘积. 由引理7.2可知,

$$|P(\lambda)|\cdot|A(\lambda)|\cdot|Q(\lambda)|=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda)0\cdots0$$

上式左边, $P(\lambda)$,$Q(\lambda)$ 为初等 $\lambda-$ 矩阵的乘积, 所以皆不为零, 又 $|A(\lambda)|$ 为非零常数, 故等式左边不等于零, 从而等式右边的 $n-r$ 个 0 均不会出现, 因此 $r=n$ 且 $d_i(\lambda)$ 都是非零常数. 又因为 $d_i(\lambda)$ 皆是首一多项式, 故对 $\forall i, d_i(\lambda)=1$, 由此可推出 $A(\lambda)=I_n$.

(3) $\Longrightarrow$ (4): 设 $\exists P(\lambda),Q(\lambda)$ 是初等 $\lambda-$ 矩阵的乘积 s.t.

$$\begin{aligned}
&P(\lambda)A(\lambda)Q(\lambda)=I_n\\
\Longrightarrow &P(\lambda)A(\lambda)=Q(\lambda)^{-1}\\
\Longrightarrow &Q(\lambda)P(\lambda)A(\lambda)=I_n
\end{aligned}$$

即 $A(\lambda)$ 只需通过初等行变换即可变为 I_n. 同理可证 $A(\lambda)$ 只需通过初等列变换即可变为 I_n.

(4) $\Longrightarrow$ (5): 设 $P(\lambda)$ 是初等 $\lambda-$ 矩阵 s.t.

$$\begin{aligned}
&P_r(\lambda)P_{r-1}(\lambda)\cdots P_1(\lambda)A(\lambda)=I_n\\
\Longrightarrow &P_{r-1}(\lambda)P_{r-2}(\lambda)\cdots P_1(\lambda)A(\lambda)=P_r(\lambda)^{-1}\\
&\cdots\cdots\\
\Longrightarrow &A(\lambda)=P_1(\lambda)^{-1}P_2(\lambda)^{-1}\cdots P_r(\lambda)^{-1}\\
\Longrightarrow &A(\lambda)\text{是初等}\lambda-\text{矩阵的乘积.}
\end{aligned}$$

(5) $\Longrightarrow$ (1): 初等 $\lambda-$ 矩阵都是可逆阵, 且有限个可逆阵的乘积仍然是可逆阵, 因此推出 (1), 此时,

$$A(\lambda)^{-1}=\frac{1}{|A(\lambda)|}\cdot A(\lambda)^*.$$

证明: 由引理7.2可知,

$$
\begin{aligned}
&A(\lambda)\cdot A(\lambda)^* = A(\lambda)^*A(\lambda) = |A(\lambda)|\cdot I_n \\
\xRightarrow{A(\lambda)\text{可逆故}|A(\lambda)\neq 0|} &A(\lambda)(\frac{1}{|A(\lambda)|}A(\lambda)^*) = (\frac{1}{|A(\lambda)|}A(\lambda)^*)A(\lambda) = I_n \\
\Longrightarrow &A(\lambda)^{-1} = \frac{1}{|A(\lambda)|}A(\lambda)^*
\end{aligned}
$$

□

□

定理 7.4 设 $A\in M_n(\mathbb{K})$, 则 $\lambda - A$ 相抵于

$$\mathrm{diag}\{1,\cdots,1,d_1(\lambda),\cdots,d_m(\lambda)\},$$

其中,$d_i(\lambda)$ 为非常数首一多项式且 $d_1(\lambda)\mid d_2(\lambda)\mid\cdots\mid d_m(\lambda)$.

证明: 由定理7.2可知,$\exists P(\lambda),Q(\lambda)$ 为可逆 $\lambda-$ 矩阵 s. t.

$$P(\lambda)(\lambda I_n - A)Q(\lambda) = \mathrm{diag}\{d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda),0,\cdots,0\}$$

两边同时取行列式, 得

$$c\cdot|\lambda I_n - A| = d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda)0\cdots 0, c\neq 0$$

上式左边非零，所以右边也不等于零, 故右边的 $n-r$ 个零均不会出现, 从而 $r=n$ 且

$$
\begin{aligned}
&c|\lambda I_n - A| = d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda) \\
\Longrightarrow &c = 1 \\
\Longrightarrow &|\lambda I_n - A| = d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda) \qquad (7.1)
\end{aligned}
$$

把 $d_i(\lambda)$ 中的 1 都列出，即可得到结论. □

例 7.1 若 $\deg d_i(\lambda)\geq 1$, 则 $A = cI_n$.

证明: 由(7.1)两边取多项式次数, 得

$$n = \sum_{i=1}^{n}\deg d_i(\lambda)\geq n$$

从而 $d_1(\lambda),d_2(\lambda),\cdots,d_n(\lambda)$ 是一次首一多项式. 又由整除条件可知,

$$d_1(\lambda) = d_2(\lambda) = \cdots = d_n(\lambda) = \lambda - c$$

即 $\lambda I_n - A$ 相抵于矩阵

$$\mathrm{diag}\{\lambda - c,\lambda - c,\cdots,\lambda - c\} = \lambda I_n - cI_n,$$

由定理7.1可知,A 与 cI_n 相似, 从而 $cI_n = P^{-1}AP \Longrightarrow A = cI_n$. □

例 7.2 已知 $A=\begin{pmatrix}0&1&-1\\3&-2&0\\-1&1&-1\end{pmatrix}$, 求 λI_n-A 的相似标准型.

解:

$$\lambda I_3-A=\begin{pmatrix}\lambda&-1&1\\-3&\lambda+2&0\\1&-1&\lambda+1\end{pmatrix}\longrightarrow\begin{pmatrix}1&-1&\lambda+1\\-3&\lambda+2&0\\\lambda&-1&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&-1&\lambda+1\\0&\lambda-1&3(\lambda+1)\\0&\lambda-1&-\lambda^2-\lambda+1\end{pmatrix}$$
$$\longrightarrow\begin{pmatrix}1&0&0\\0&\lambda-1&3(\lambda+1)\\0&\lambda-1&-\lambda^2-\lambda+1\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0\\0&\lambda-1&6\\0&\lambda-1&-\lambda^2-4\lambda+4\end{pmatrix}$$
$$\longrightarrow\begin{pmatrix}1&0&0\\0&6&\lambda-1\\0&-\lambda^2-4\lambda+4&\lambda-1\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0\\0&6&6(\lambda-1)\\0&-\lambda^2-4\lambda+4&6(\lambda-1)\end{pmatrix}$$
$$\longrightarrow\begin{pmatrix}1&0&0\\0&6&0\\0&-\lambda^2-4\lambda+4&(\lambda-1)^2(\lambda^2+4\lambda+2)\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0\\0&1&0\\0&0&(\lambda-1)^2(\lambda^2+4\lambda+2)\end{pmatrix}$$

7.3 不变因子

引理 7.3 设 $A(\lambda)$ 的非零行列式因子为 $D_1(\lambda),D_2(\lambda),\cdots,D_r(\lambda)$, 则

$$D_i(\lambda)\mid D_{i+1}(\lambda),\forall 1\le i\le r-1.$$

证明: 任取 $A(\lambda)$ 的 $i+1$ 阶子式 M_{i+1}, 将 M_{i+1} 按某行 (列) 展开, 这里假设按第一行展开:

$$M_{i+1}=a_{11}(\lambda)M_{11}-a_{12}(\lambda)M_{12}+\cdots+(-1)^{i+2}a_{1,i+1}M_{1,i+1}\tag{7.2}$$

$$\begin{aligned}\text{由行列式因子定义}&\Longrightarrow D_i(\lambda)\mid(\lambda)M_{1j},1\le j\le i+1\\&\Longrightarrow D_i(\lambda)\mid M_{i+1}\\&\xRightarrow{D_i(\lambda)\text{是}M_{i+1}\text{公因式}}D_i(\lambda)\mid D_{i+1}(\lambda),\forall 1\le i\le r-1\end{aligned}$$

□

定理 7.5 相抵的 $\lambda-$ 矩阵有相同的行列式因子, 从而有相同的不变因子.

证明: 设

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$$

其中,$P(\lambda),Q(\lambda)$ 是可逆 $\lambda-$ 矩阵, 设 $A(\lambda),B(\lambda)$ 的行列式因子分别为

$$D_k(\lambda),E_k(\lambda),1\le k\le n$$

要证
$$D_k(\lambda)=E_k(\lambda),\forall 1\le k\le n.$$

$$B(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \xlongequal{\text{Cauchy-Binet 公式}} P(\lambda)A(\lambda)Q(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \tag{7.3}$$

$$= \sum_{\substack{1\le r_1<\cdots<r_k\le n\\ 1\le s_1<\cdots<s_k\le n}} P(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ r_1 & r_2 & \cdots & r_k \end{pmatrix} A(\lambda)\begin{pmatrix} r_1 & r_2 & \cdots & r_k \\ s_1 & s_2 & \cdots & s_k \end{pmatrix} Q(\lambda)\begin{pmatrix} s_1 & s_2 & \cdots & s_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \tag{7.4}$$

1°. 若 $D_k(\lambda)=0$: 由(7.4)式可得:
$$B(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}=0,$$
从而 $E_k(\lambda)=0$;

2°. 若 $D_k(\lambda)\ne 0$: 由行列式因子定义可知:
$$D_k(\lambda)\mid A(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$$
从而由(7.4)式可知:
$$D_k(\lambda)\mid B(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$$
从而 $D_k(\lambda)\mid E_k(\lambda)$.

由于 $A(\lambda)=P(\lambda)^{-1}B(\lambda)Q(\lambda)^{-1}$, 因此同理可证,

1°. $E_k(\lambda)=0\Longrightarrow D_k(\lambda)=0$;

2°. 若 $E_k(\lambda)\ne 0$, 则 $D_k(\lambda)\ne 0$ 且 $E_k(\lambda)\mid D_k(\lambda)$.

由以上可知,
$$E_k(\lambda)=0\Longleftrightarrow D_k(\lambda)=0$$
或者
$$D_k(\lambda)\mid E_k(\lambda), E_k(\lambda)\mid D_k(\lambda)$$

结合以上可推出,
$$\exists 0\ne c\, \text{s.t. } D_r(\lambda)=c\cdot E_k(\lambda),$$
又因为 $D_k(\lambda),E_k(\lambda)$ 都是首一多项式, 因此 $c=1\Longrightarrow D_k(\lambda)=E_k(\lambda)$. □

推论 7.1 设 n 阶 $\lambda-$ 矩阵 $A(\lambda)$ 的法式为
$$\Lambda=\operatorname{diag}\{d_1(\lambda),d_1(\lambda),\cdots,d_r(\lambda),0,\cdots,0\},$$
其中 d_i 是非零首一多项式且 $d_i(\lambda)\mid d_{i+1}(\lambda)(i=1,2,\cdots,r-1)$, 则 $A(\lambda)$ 的不变因子为 $d_1(\lambda)$, $d_2(\lambda),\cdots,d_r(\lambda)$. 特别, 法式和不变因子之间相互唯一确定.

证明: 因为 $A(\lambda)$ 相抵于 Λ, 由定理7.5可知 $A(\lambda)$ 的不变因子即为 Λ 的不变因子, 而 Λ 的不变因子即 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$. □

推论 7.2 设 $A(\lambda), B(\lambda)$ 为 n 阶 $\lambda-$ 矩阵, 则 $A(\lambda)$ 与 $B(\lambda)$ 相抵当且仅当它们有相同的法式.

证明: 若 $A(\lambda)$ 和 $B(\lambda)$ 有相同的法式, 设为 Λ, 则 $A(\lambda) \sim \Lambda \sim B(\lambda)$, 由相抵的传递性可知, $A(\lambda) \sim B(\lambda)$.

若 $A(\lambda)$ 和 $B(\lambda)$ 相抵, 设 $A(\lambda)$ 的法式为 Λ_1, 设 $B(\lambda)$ 的法式为 Λ_2, 则

$$A(\lambda) \sim \Lambda_1, B(\lambda) \sim \Lambda_2, \text{且} A(\lambda) \sim B(\lambda)$$

由相抵的传递性可知, $\Lambda_1 \sim \Lambda_2$, 由定理7.5可知, Λ_1 与 Λ_2 有相同的不变因子. 由于 Λ_1 与 Λ_2 都是法式, 所以它们主对角线上的非零首一多项式元素都相同, 从而 $\Lambda_1 = \Lambda_2$. □

推论 7.3 $\lambda-$ 矩阵在相抵关系下的全系不变量是它的行列式因子组或不变因子组.

证明: 由定理7.5可知, 相抵的 $\lambda-$ 矩阵有相同的行列式因子 (组), 从而有相同的不变因子 (组).

反过来, 如果两个 $\lambda-$ 矩阵有相同的行列式因子或不变因子, 由推论7.1可知, 它们必具有相同的法式, 从而由推论7.2可知, $A(\lambda) \sim B(\lambda)$. □

推论 7.4 n 阶 $\lambda-$ 矩阵 $A(\lambda)$ 的法式与初等变换的选取无关.

证明: 设 $A(\lambda)$ 可通过某初等变换过程变换为法式 Λ_1, 设 $A(\lambda)$ 可通过另一初等变换过程变换为法式 Λ_2, 只要证 $\Lambda_1 = \Lambda_2$.

$$A(\lambda) \sim \Lambda_1, A(\lambda) \sim \Lambda_2 \Longrightarrow \Lambda_1 \sim \Lambda_2$$

由推论7.2可知,$\Lambda_1 = \Lambda_2$. □

定理 7.6 数域 $\mathbb{K}$ 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 $\lambda I - A$ 和 $\lambda I - B$ 具有相同的行列式因子或不变因子.

证明: 由定理7.1可知,

$$A\text{与}B\text{相似} \Longleftrightarrow \lambda I - A\text{与}\lambda I - B\text{相抵}.$$

由推论7.3可知, $\lambda I - A$ 与 $\lambda I - B$ 有相同的行列式因子或不变因子. □

注 特征矩阵 $\lambda I - A$ 的行列式因子及不变因子均简称为 A 的行列式因子及不变因子.

推论 7.5 设 $\mathbb{F} \subseteq \mathbb{K}$ 是两个数域, A, B 是 $\mathbb{F}$ 上的两个矩阵, 则 A 与 B 在 $\mathbb{F}$ 上相似的充分必要条件是他们在 $\mathbb{K}$ 上相似.

证明: A 与 B 在 $\mathbb{F}$ 上相似 $\overset{\text{def}}{\Longleftrightarrow}$ $\exists$ 非异矩阵 $P \in M_n(\mathbb{F})$ s. t. $B = P^{-1}AP$, A 与 B 在 $\mathbb{K}$ 上相似 $\overset{\text{def}}{\Longleftrightarrow}$ $\exists$ 非异矩阵 $Q \in M_n(\mathbb{K})$ s. t. $B = Q^{-1}AQ$.

必要性: 设 $\exists$ 非异矩阵 $P \in M_n(\mathbb{F})$ s. t. $B = P^{-1}AP$, 可知 $P^{-1} \in \mathbb{F}$. 把 P, P^{-1} 看成是 $\mathbb{K}$ 上的矩阵, 从而 P 在 $\mathbb{K}$ 上可逆, 可以推出 A 与 B 在 $\mathbb{K}$ 上也相似.

充分性: 若 A 与 B 在 $\mathbb{K}$ 上相似, 由推论7.6可知, $\lambda I - A$ 与 $\lambda I - B$ 作为 $\mathbb{K}$ 上的 $\lambda-$ 矩阵有相同的不变因子组 (法式). 实际上 $\lambda I - A$ 与 $\lambda I - B$ 都是 $\mathbb{F}$ 上的 $\lambda-$ 矩阵, 由推论7.4可知, 在求法式的过程中, 只要取 $\mathbb{F}[\lambda]$ 上的初等变换, 就可以得到 $\lambda I - A$ 与 $\lambda I - B$ 的法式, 且法式中的不变因子多项式 $d_i(\lambda)$ 仍是 $\mathbb{F}$ 上的多项式. 因此, 存在 $\mathbb{F}$ 上的可逆 $\lambda-$ 矩阵 $P(\lambda), Q(\lambda), M(\lambda), N(\lambda)$ s. t.

$$P(\lambda)(\lambda I - A)Q(\lambda) = M(\lambda)(\lambda I - B)N(\lambda) = \operatorname{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)\},$$

从而

$$(\lambda I - B) = M(\lambda)^{-1}P(\lambda)(\lambda I - A)Q(\lambda)N(\lambda)^{-1},$$

即 $\lambda I - A$ 与 $\lambda I - B$ 在 $\mathbb{F}$ 相抵，由定理7.1可知, A 与 B 在 $\mathbb{F}$ 上相似. □

推论 7.6 矩阵的不变因子组在基域扩张下不变.

7.4 有理标准型

引理 7.4 设 r 阶矩阵

$$F = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_r & -a_{r-1} & -a_{r-2} & \cdots & -a_1 \end{pmatrix}, \tag{7.5}$$

则

(1) F 的行列式因子为

$$1, \cdots, 1, f(\lambda), \tag{7.6}$$

其中共有 $r-1$ 个 1, $f(\lambda) = \lambda^r + a_1\lambda^{r-1} + \cdots + a_r$, F 的不变因子组也由(7.6)式给出;

(2) F 的极小多项式和特征多项式都等于 $f(\lambda)$.

证明: (1)

$$\lambda I_r - F = \begin{pmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \lambda & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda & -1 \\ a_r & a_{r-1} & a_{r-2} & \cdots & \lambda + a_1 \end{pmatrix}$$

$$D_r(\lambda)=|\lambda I_r-F|=\lambda^r+a_1\lambda^{r-1}+\cdots+a_r=f(\lambda)$$

对 $\forall 1\le k<r, \lambda I-F$ 总有一个 k 阶子式 $=(-1)^k$, 从而

$$D_k(\lambda)=1,\forall 1\le k<r.$$

(2) 设 F 极小多项式为 $m(\lambda)$, 只要证 $m(\lambda)=f(\lambda)$.

由定理6.3(Cayley-Hamilton 定理)$\Longrightarrow m(\lambda)\mid f(\lambda)\Longrightarrow \deg m(\lambda)\le r$.

1°. 若 $\deg m(\lambda)=r$, 则 $m(\lambda)=f(\lambda)$;

2°. 若 $\deg m(\lambda)<r$, 用反正法: 设

$$m(\lambda)=c_{r-1}\lambda^{r-1}+c_{r-2}\lambda^{r-2}+\cdots+c_1\lambda+c_0, c_i\text{不全为}0,$$

设 e_i 为第 i 个元素为 1, 其余元素为 0 的单位行向量, 则

$$e_1F=e_2,e_2F=e_3,\cdots,e_{r-1}F=e_r,$$

即

$$e_1F=e_2,e_1F^2=e_3,\cdots,e_1F^{r-1}=e_r. \tag{7.7}$$

由于 $m(\Lambda)$ 是 F 的极小多项式, 所以

$$0=m(F)=c_{r-1}F^{r-1}+c_{r-2}F^{r-2}+\cdots+c_1F+c_0I_r,$$

上式左右两边同时左乘 e_1, 可得

$$0=c_{r-1}e_1F^{r-1}+c_{r-2}e_1F^{r-2}+\cdots+c_1e_1F+c_0e_1,$$

上式各项分别应用(7.7)式的结论, 可得

$$0=c_{r-1}e_r+c_{r-2}e_{r-1}+\cdots+c_1e_2+c_0e_1=(c_0,c_1,\cdots,c_{r-1})\Longrightarrow c_0=c_1=\cdots=c_{r-1}=0$$

与假设矛盾. □

引理 7.5 设 $\lambda-$ 矩阵 $A(\lambda)$ 相抵于对角 $\lambda-$ 矩阵

$$\operatorname{diag}\{d_1(\lambda),d_2(\lambda),\cdots,d_n(\lambda)\}, \tag{7.8}$$

$\lambda-$ 矩阵 $B(\lambda)$ 相抵于对角 $\lambda-$ 矩阵

$$\operatorname{diag}\{d_{i_1}(\lambda),d_{i_2}(\lambda),\cdots,d_{i_n}(\lambda)\}, \tag{7.9}$$

且 $i_1,i_2,\cdots,i_n$ 是 $1,2,\cdots,n$ 的全排列, 则 $A(\lambda)\sim B(\lambda)$.

证明: 利用行对换及列对换即可将(7.8)式变换成(7.9)式, 因此(7.8)式所表示的矩阵和(7.9)式所表示的矩阵相抵, 所以 $A(\lambda)\sim B(\lambda)$. □

定理 7.7 设 A 是数域 $\mathbb{K}$ 上的 n 阶方阵, A 的不变因子组为

$$1, \cdots, 1, d_1(\lambda), \cdots, d_k(\lambda),$$

其中 $\deg d_i(\lambda) = m_i$, 则 A 相似于下列分块对角阵:

$$F = \begin{pmatrix} F_1 & & & \\ & F_2 & & \\ & & \ddots & \\ & & & F_k \end{pmatrix}$$

其中 F_i 的阶等于 m_i, 且 F_i 是形如(7.5)式的矩阵, F_i 的最后一行由 $d_i(\lambda)$ 的系数 (除最高次项) 的负值组成.

证明:

$$\begin{aligned} &\lambda I_n - A \sim \operatorname{diag}\{1, \cdots, 1, d_1(\lambda), \cdots, d_k(\lambda)\} \qquad (7.10) \\ \Longrightarrow &|\lambda I_n - A| = d_1(\lambda) \cdots d_k(\lambda) \end{aligned}$$

则 $m_1 + m_2 + \cdots + m_k = n$,

$$\lambda I - F_i = \lambda I - F(d_i(\lambda)) \sim \operatorname{diag}\{1, \cdots, 1, d_i(\lambda)\},$$

上式中 $d_i(\lambda)$ 前有 $m_i - 1$ 个 1, 则

$$\lambda I - F = \begin{pmatrix} \lambda I - F(d_1(\lambda)) & & & \\ & \lambda I - F(d_2(\lambda)) & & \\ & & \ddots & \\ & & & \lambda I - F(d_k(\lambda)) \end{pmatrix}$$

即 $\lambda I - A$ 的法式为

$$\operatorname{diag}\{1, \cdots, 1, d_1(\lambda); 1, \cdots, 1, d_2(\lambda); \cdots; 1, \cdots, 1, d_k(\lambda)\}, \qquad (7.11)$$

上式中, 每个 $d_i(\lambda)$ 前都配有 $m_i - 1$ 个 1, 即共有 $m_1 - 1 + m_2 - 1 + \cdots + m_k - 1 = n - k$ 个 1. 因此, (7.11)式是(7.10)式的一个置换. 由引理7.5可知, $\lambda I_n - A \sim \lambda I_n - F$, 由定理7.1可知, $A \sim F$. □

定理 7.8 设数域 $\mathbb{K}$ 上的 n 阶矩阵 A 的不变因子为

$$1, \cdots, 1, d_1(\lambda), \cdots, d_k(\lambda),$$

其中 $d_i(\lambda) \mid d_{i+1}(\lambda) (i = 1, \cdots, k-1)$, 则 A 的极小多项式 $m(\lambda) = d_k(\lambda)$.

证明: 由定理7.6可知,

$$A \sim \operatorname{diag}\{F(d_1(\lambda)), F(d_2(\lambda)), \cdots, F(d_k(\lambda))\},$$

1°. $m_A(\lambda) = m_F(\lambda)$;

2°. 令 $F_i = F(d_i(\lambda))$, 则

$$m_F(\lambda) = [m_{F_1}(\lambda), m_{F_2}(\lambda), \cdots, m_{F_k}(\lambda)].$$

$$\begin{aligned}\text{由引理7.4} \Longrightarrow & m_{F_i}(\lambda) = d_i(\lambda), 1 \leq i \leq k \\ \Longrightarrow & m_A(\lambda) = [d_1(\lambda), d_2(\lambda), \cdots, d_k(\lambda)] \xlongequal{d_i(\lambda)|d_{i+1}(\lambda)} d_k(\lambda).\end{aligned}$$

□

推论 7.7 设 $A \in M_n(\mathbb{F})$, 其极小多项式为 $m_{\mathbb{F}}(\lambda)$. $\mathbb{F} \subseteq \mathbb{K}$, 将 A 看成是 $\mathbb{K}$ 上的矩阵, 极小多项式为 $m_{\mathbb{K}}(\lambda)$, 则 $m_{\mathbb{K}}(\lambda) = m_{\mathbb{F}}(\lambda)$.

7.5 初等因子

定义 7.2 设 $f(\lambda) \in \mathbb{K}[\lambda], p(\lambda)$ 为不可约多项式, 若存在 $e \in \mathbb{Z}^+$, 使得 $p(\lambda)^e \mid f(\lambda)$ 但 $p(\lambda)^{e+1} \nmid f(\lambda)$, 则称 $p(\lambda)^e$ 为 $f(\lambda)$ 的准素因子.

定义 7.3 设 $d_1(\lambda), d_2(\lambda), \cdots, d_k(\lambda)$ 是数域 $\mathbb{K}$ 上矩阵 A 的非常数不变因子, 在 $\mathbb{K}$ 上把 $d_i(\lambda)$ 分解成不可约因式之积:

$$\begin{aligned} d_1(\lambda) &= p_1(\lambda)^{e_{11}} p_2(\lambda)^{e_{12}} \cdots p_t(\lambda)^{e_{1t}}, \\ d_2(\lambda) &= p_1(\lambda)^{e_{21}} p_2(\lambda)^{e_{22}} \cdots p_t(\lambda)^{e_{2t}}, \\ &\cdots\cdots \\ d_k(\lambda) &= p_1(\lambda)^{e_{k1}} p_2(\lambda)^{e_{k2}} \cdots p_t(\lambda)^{e_{kt}}, \end{aligned} \tag{7.12}$$

其中 e_{ij} 是非负整数 (注意 e_{ij} 可以为零!). 由于 $d_i(\lambda) \mid d_{i+1}$, 因此

$$e_{1j} \leq e_{2j} \leq \cdots \leq e_{kj} (j = 1, 2, \cdots, t).$$

若(7.12)式中的 $e_{ij} > 0$, 则称 $p_j(\lambda)^{e_{ij}}$ 为 A 的一个**初等因子**, A 的全体初等因子成为 A 的**初等因子组**.

由因式分解的唯一性可知 A 的初等因子被 A 的不变因子唯一确定. 反过来, 若给定一组初等因子 $p_j(\lambda)^{e_{ij}}$, 适当添加一些 1(表示为 $p_j(\lambda)^{e_{ij}} = 0$), 则可将这组初等因子按降幂排列如下:

$$\begin{aligned} & p_1(\lambda)^{e_{k1}}, p_1(\lambda)^{e_{k-1,1}}, \cdots, p_1(\lambda)^{e_{11}}, \\ & p_2(\lambda)^{e_{k2}}, p_1(\lambda)^{e_{k-1,2}}, \cdots, p_1(\lambda)^{e_{12}}, \\ & \cdots\cdots \\ & p_t(\lambda)^{e_{kt}}, p_1(\lambda)^{e_{k-1,t}}, \cdots, p_1(\lambda)^{e_{1t}}, \end{aligned} \tag{7.13}$$

令

$$\begin{aligned}
d_k(\lambda) &= p_1(\lambda)^{e_{k1}} p_2(\lambda)^{e_{k2}} \cdots p_t(\lambda)^{e_{kt}}, \\
d_{k-1}(\lambda) &= p_1(\lambda)^{e_{k-1,1}} p_2(\lambda)^{e_{k-1,2}} \cdots p_t(\lambda)^{e_{k-1,t}}, \\
&\cdots\cdots \\
d_1(\lambda) &= p_1(\lambda)^{e_{11}} p_2(\lambda)^{e_{12}} \cdots p_t(\lambda)^{e_{1t}},
\end{aligned}$$

则 $d_i(\lambda) \mid d_{i+1}(\lambda)(i = 1, 2, \cdots, k-1)$, 且 $d_1(\lambda), \cdots, d_k(\lambda)$ 的初等因子组就如 (7.13)式所示, 因此, 给定 A 的初等因子组, 我们可唯一地确定 A 的不变因子组. 这一事实表明, A 的不变因子组与初等因子组在讨论矩阵相似关系中的作用是相同的. 因此, 有下述定理.

定理 7.9 数域 $\mathbb{K}$ 上的两个矩阵 A 与 B 相似的充分必要条件是他们具有相同的初等因子组, 即矩阵的初等因子组是矩阵相似关系的全系不变量.

7.6 Jordan 标准型

引理 7.6 r 阶矩阵

$$J = \begin{pmatrix}
\lambda_0 & 1 & & & \\
& \lambda_0 & 1 & & \\
& & \ddots & \ddots & \\
& & & \ddots & 1 \\
& & & & \lambda_0
\end{pmatrix}$$

的初等因子组为 $(\lambda - \lambda_0)^r$.

证明:

$$\lambda I - J = \left(\begin{array}{c|cccc}
\lambda - \lambda_0 & -1 & & & \\
& \lambda - \lambda_0 & -1 & & \\
& & \ddots & \ddots & \\
& & & \ddots & -1 \\
\hline
& & & & \lambda - \lambda_0
\end{array}\right), \tag{7.14}$$

$$D_r(\lambda) = |\lambda I_r - J| = (\lambda - \lambda_0)^r,$$

如(7.14)式矩阵中虚线右上角部分所示,

$$\forall 1 \le k < r, \exists k\text{阶子式} = (-1)^k$$
$$\Longrightarrow D_k(\lambda) = 1, \forall 1 \le k < r.$$

(7.14)式的行列式因子为

$$1, \cdots, 1, (\lambda - \lambda_0)^r,$$

(7.14)式的不变因子也为上式, 因此矩阵 J 的初等因子组为 $(\lambda - \lambda_0)^r$. □

引理 7.7 设特征矩阵 $\lambda I - A$ 经过初等变换化为下列对角阵:

$$\begin{pmatrix} f_1(\lambda) & & & \\ & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \tag{7.15}$$

其中, $f(\lambda)(i=1,\cdots,n)$ 为非零首一多项式. 将 $f_i(\lambda)$ 作不可约分解, 若 $(\lambda-\lambda_0)^k$ 能整除 $f_i(\lambda)$, 但 $(\lambda-\lambda_0)^{k+1}$ 不能整除 $f_i(\lambda)$, 就称 $(\lambda-\lambda_0)^k$ 是 $f_i(\lambda)$ 的一个准素因子, 则矩阵 A 的初等因子组等于所有 $f_i(\lambda)$ 的准素因子的集合.

证明: Step 1: 对 $\forall 1 \le i < j \le n$, 设

$$(f_i(\lambda), f_j(\lambda)) = g(\lambda), [f_i(\lambda), f_j(\lambda)] = h(\lambda),$$

先证

$$\begin{pmatrix} f_1(\lambda) & & & & & & \\ & \cdots & & & & & \\ & & f_i(\lambda) & & & & \\ & & & \cdots & & & \\ & & & & f_j(\lambda) & & \\ & & & & & \cdots & \\ & & & & & & f_n(\lambda) \end{pmatrix} \sim \begin{pmatrix} f_1(\lambda) & & & & & & \\ & \cdots & & & & & \\ & & g(\lambda) & & & & \\ & & & \cdots & & & \\ & & & & h(\lambda) & & \\ & & & & & \cdots & \\ & & & & & & f_n(\lambda) \end{pmatrix} \tag{7.16}$$

并且两个对角阵主对焦元素的准素因子全体相同.

不妨设 $i=1, j=2$, 设

$$f_1(\lambda) = g(\lambda)q(\lambda), \tag{7.17}$$

$$h(\lambda) = f_1(\lambda)f_2(\lambda)/g(\lambda) = f_2(\lambda)q(\lambda), \tag{7.18}$$

$$\exists u(\lambda), v(\lambda)\, \mathrm{s.t.}\ g(\lambda) = f_1(\lambda)u(\lambda) + f_2(\lambda)v(\lambda). \tag{7.19}$$

由于对角阵第 $1,2$ 两行 (列) 的初等行 (列) 变换和其他行 (列) 无关，因此，下面只对前 2 行 (列) 进行操作:

$$\begin{pmatrix} f_1(\lambda) & 0 \\ 0 & f_2(\lambda) \end{pmatrix} \xrightarrow{(7.19)\text{式}} \begin{pmatrix} f_1(\lambda) & g(\lambda) \\ 0 & f_2(\lambda) \end{pmatrix} \xrightarrow{(7.17)\text{式}} \begin{pmatrix} 0 & g(\lambda) \\ -f_2(\lambda)q(\lambda) & f_2(\lambda) \end{pmatrix}$$

$$\xrightarrow{g(\lambda)\text{是}f_2(\lambda)\text{的因式}} \begin{pmatrix} 0 & g(\lambda) \\ -h(\lambda) & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} g(\lambda) & 0 \\ 0 & h(\lambda) \end{pmatrix}.$$

在数域 $\mathbb{K}$ 上做公共因式分解:

$$\begin{aligned} f_1(\lambda) &= p_1(\lambda)^{r_1} p_2(\lambda)^{r_2} \cdots p_t(\lambda)^{r_t}, \\ f_2(\lambda) &= p_1(\lambda)^{s_1} p_2(\lambda)^{s_2} \cdots p_t(\lambda)^{s_t}, \end{aligned} \tag{7.20}$$

其中 $p_i(\lambda)$ 是首一互异的不可约多项式且 $r_i, s_i \geq 0$, 令 $e_i = \min\{r_i, s_i\}, k_i = \max\{r_i, s_i\}$, 则

$$\begin{aligned} g(\lambda) &= p_1(\lambda)^{e_1} p_2(\lambda)^{e_2} \cdots p_t(\lambda)^{e_t}, \\ h(\lambda) &= p_1(\lambda)^{k_1} p_2(\lambda)^{k_2} \cdots p_t(\lambda)^{k_t}, \end{aligned} \tag{7.21}$$

(7.21)式和(7.20)式中, e_i, k_i 和对应的 r_i, s_i 相同或只是互换个位置, 由构造过程可知:

$$\mathrm{diag}\{f_1(\lambda), f_2(\lambda)\}\text{与}\,\mathrm{diag}\{g(\lambda), h(\lambda)\}\text{有相同的准素因子组.}$$

Step 2: 反复利用 Step 1, 将 $\mathrm{diag}\{f_1(\lambda), \cdots, f_n(\lambda)$ 变为法式. 依次将第 $(1,1)$ 元素与 $(2,2)$ 元素, 第 $(3,3)$ 元素,..., 第 (n,n) 元素进行 Step 1 的操作，从而使第 $(1,1)$ 元素的多项式幂次均是最小的. 一般地, 设第 $(1,1)$ 元素,..., 第 $(i-1,i-1)$ 元素已完成如上操作, 第 (i,i) 元素, 第 $(i+1,i+1)$ 元素,..., 第 (n,n) 元素重复进行 Step 1 的操作, 可使原矩阵相抵于 $\mathrm{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)\}$ 且满足 $d_1(\lambda) \mid d_2(\lambda) \mid \cdots \mid d_n(\lambda)$, 即原矩阵和其法式相抵. □

例 7.3 $\lambda I - A$ 相抵于 $\mathrm{diag}\{1, (\lambda-1)^2(\lambda+2), \lambda+2, 1, \lambda-1\}$, 求它的所有初等因子.

解: 由引理7.7可知, 初等因子为 $\lambda-1, (\lambda-1)^2, \lambda+2, \lambda+2$.

引理 7.8 设 $A = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_k}(\lambda_k)\}$, 则 A 的初等因子组为 $(\lambda-\lambda_1)^{r_1}, (\lambda-\lambda_2)^{r_2}, \cdots, J_{r_k}(\lambda)^{r_k}$.

证明:

$$\lambda I - A = \begin{pmatrix} \lambda I - J_{r_1}(\lambda_1) & & & \\ & \lambda I - J_{r_2}(\lambda_2) & & \\ & & \ddots & \\ & & & \lambda I - J_{r_k}(\lambda_k) \end{pmatrix} \tag{7.22}$$

由引理7.6可推出 $\lambda I - J_{r_i}(\lambda_i) \sim \mathrm{diag}\{1, \cdots, 1, (\lambda-\lambda_i)^{r_i}\}$, 因此

$$\lambda I - A \sim \mathrm{diag}\{1, \cdots, 1, (\lambda-\lambda_1)^{r_1}; 1, \cdots, 1, (\lambda-\lambda_2)^{r_2}; \cdots; 1, \cdots, 1, (\lambda-\lambda_k)^{r_k}\},$$

由引理7.7可推出, A 的初等因子组为 $(\lambda-\lambda_1)^{r_1}, (\lambda-\lambda_2)^{r_2}, \cdots, (\lambda-\lambda_k)^{r_k}$. □

定理 7.10 设 $A \in M_n(\mathbb{C})$ 的初等因子组为 $(\lambda-\lambda_1)^{r_1}, (\lambda-\lambda_2)^{r_2}, \cdots, (\lambda-\lambda_k)^{r_k}$, 则 A 相似于 $J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_k}(\lambda_k)\}$.

证明: 由引理7.8可知 J 的初等因子组也为 $(\lambda-\lambda_1)^{r_1}, (\lambda-\lambda_2)^{r_2}, \cdots, (\lambda-\lambda_k)^{r_k}$, 从而

$$A \simeq J.$$

□

定义 7.4 定理7.10中的 $J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_k}(\lambda_k)$ 称为 A 的 Jordan 标准型, $J_{r_i}(\lambda_i)$ 称为 Jordan 块.

注 Jordan 标准型中, Jordan 块次序可以互换. 每个 Jordan 块由一个初等因子唯一确定.

$$J_r(\lambda_0) \longleftrightarrow (\lambda-\lambda_0)^r,$$

从而在不考虑 Jordan 块的次序时, Jordan 标准型由 A 唯一确定.

定理 7.11 设 $\varphi \in \mathcal{L}(V_{\mathbb{C}}^n)$, 则存在 V 的一组基 $\{e_1, e_2, \cdots, e_n\}$, 使得 φ 在这组基下的表示矩阵为 Jordan 标准型 $J = \mathrm{diag}\{J_{r_1}(\lambda_1), \cdots, J_{r_k}(\lambda_k)\}$.

推论 7.8 设 $A \in M_n(\mathbb{C})$, 则下列结论等价:

(1) A 可对角化;

(2) A 的极小多项式无重根;

(3) A 的初等因子都是一次多项式.

证明: (1)$\Longrightarrow$(2): 由例题6.1可得.

(2)$\Longrightarrow$(3): 设 A 的不变因子组为 $1, \cdots, 1, d_1(\lambda), \cdots, d_k(\lambda)$, 由有理标准型理论可知, A 的极小多项式为 $m(\lambda) = d_k(\lambda)$, 从而 $d_k(\lambda)$ 无重根, 因此 $d_1(\lambda), \cdots, d_k(\lambda)$ 均无重根, 由此可推出, $d_i(\lambda)$ 的准素因子都是一次的, 即 A 的初等因子都是一次多项式.

(3)$\Longrightarrow$(1): 设 A 的初等因子组为 $\lambda-\lambda_1, \lambda-\lambda_2, \cdots, \lambda-\lambda_n$, 由定理7.11可知,

$$A \simeq J = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

所以, A 可对角化, 即 A 的 Jordan 块都是一阶的. □

命题 7.1 设 $\varphi \in \mathcal{L}(V_{\mathbb{C}}^n)$, V_o 是 φ 的不变子空间. 若 φ 可对角化, 则 $\varphi|_{V_0}$ 也可对角化.

证明: 设 φ 的极小多项式为 $m(\lambda)$, $\varphi|_{V_0}$ 的极小多项式为 $g(\lambda)$, 则

$$m(\varphi|_{V_0}) = m(\varphi)|_{V_0} = 0.$$

$$\begin{aligned} &\text{由极小多项式性质} \Longrightarrow g(\lambda)|m(\lambda) \\ &\varphi\text{可对角化、推论7.6} \Longrightarrow m(\lambda)\text{无重根} \Longrightarrow g(\lambda)\text{也无重根} \\ &\xRightarrow{\text{推论7.6}} \varphi|_{V_0}\text{也可对角化.} \end{aligned}$$

□

推论 7.9 若

$$M = \begin{pmatrix} A^{m\times m} & C^{m\times n} \\ O & B^{n\times n} \end{pmatrix}$$

可对角化, 则 A, B 均可对角化.

证明: 设 M 的极小多项式为 $m(\lambda)$, 则

$$0 = m(M) = \begin{pmatrix} m(A) & * \\ O & m(B) \end{pmatrix},$$

从而

$$m(A) = m(B) = 0.$$

由极小多项式性质可以推出

$$m_A(\lambda) \mid m(\lambda), m_B(\lambda) \mid m(\lambda).$$

$$\text{由推论7.6} \Longrightarrow m(\lambda)\text{无重根} \Longrightarrow m_A(\lambda), m_B(\lambda)\text{无重根} \Longrightarrow A, B\text{可对角化}.$$

□

命题 7.2 $\varphi \in \mathcal{L}(V), V = V_1 \oplus V_2 \oplus \cdots \oplus V_k, V_i$ 为 φ 的不变子空间. 则

$$\varphi\text{可对角化} \Longleftrightarrow \varphi|_{V_i}\text{均可对角化}.$$

证明: **必要性**由命题7.1可得结论.

充分性$\exists V_i$ 的一组基 $\{e_{i1}, e_{i2}, \cdots, e_{in_i}\}$ s. t. $\varphi|_{V_i}$ 的表示矩阵为对角阵.

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k \Longrightarrow V_i\text{的基可拼成}V\text{的一组基},$$

$\because \varphi$ 在每个 $\varphi|_{V_i}$ 的表示矩阵在对应的基下都是对角阵,
$\therefore \varphi$ 在这组基下的表示矩阵也为对角阵. □

推论 7.10

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \cdots & \\ & & & A_k \end{pmatrix},$$

则 A 可对角化 $\Longleftrightarrow A_1, A_2, \cdots, A_k$ 均可对角化.

证明:

$$m_A(\lambda) = [m_{A_1}(\lambda), m_{A_2}(\lambda), \cdots, m_{A_k}(\lambda)].$$

□

推论 7.11 设 $A \in M_n(\mathbb{K})$ 的特征值全部在 $\mathbb{K}$ 中, 则 A 在 $\mathbb{K}$ 上相似于其 Jordan 标准型.

证明: 有定理7.10可推出, 在复数域上

$$A \simeq J = \mathrm{diag}\{J_{r_1}(\lambda_1), \cdots, J_{r_K}(\lambda_k)\},$$

由相似关系的基域扩张不变形可知, 在数域 $\mathbb{K}$ 上

$$A \simeq J.$$

□

例 7.4 A 的初等因子组为 $(\lambda-1),(\lambda-1)^3,(\lambda+2)^2,(\lambda+2)^3$, 求 A 的 Jordan 标准型.

解:

$$J=\operatorname{diag}\left\{1,\begin{pmatrix}1&1&\\&1&1\\&&1\end{pmatrix},\begin{pmatrix}-2&1\\&-2\end{pmatrix},\begin{pmatrix}-2&1&\\&-2&1\\&&-2\end{pmatrix}\right\}$$

例 7.5 设 $\varphi\in\mathcal{L}(V_{\mathbb{C}}^4)$ 在基 $\{e_1,e_2,e_3,e_4\}$ 下的表示矩阵为

$$A=\begin{pmatrix}3&1&0&0\\-4&-1&0&0\\6&1&2&1\\-14&-5&-1&0\end{pmatrix},$$

φ 在新基 $\{f_1,f_2,f_3,f_4\}$ 下的表示矩阵为 Jordan 标准型 J, 求过渡矩阵 P.

解:

$$P^{-1}AP=J,$$

$$\lambda I-A\longrightarrow\cdots\longrightarrow\operatorname{diag}\{1,1,(\lambda-1)^2,(\lambda-1)^2\},$$

所以 A 的初等因子组为 $(\lambda-1)^2,(\lambda-1)^2$, 由此可算出

$$J=\begin{pmatrix}1&1&&\\0&1&&\\&&1&1\\&&0&1\end{pmatrix}$$

设 $P=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$,

$$P^{-1}AP=J\Longrightarrow AP=PJ\Longrightarrow(A\alpha_1,A\alpha_2,A\alpha_3,A\alpha_4)=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)\begin{pmatrix}1&1&&\\0&1&&\\&&1&1\\&&0&1\end{pmatrix}$$

$$\Longrightarrow\begin{cases}A\alpha_1=\alpha_1\\A\alpha_2=\alpha_1+\alpha_2\Longrightarrow(A-I_4)\alpha_2=\alpha_1\\A\alpha_3=\alpha_3\\A\alpha_4=\alpha_3+\alpha_4\Longrightarrow(A-I_4)\alpha_4=\alpha_3\end{cases}\Longrightarrow\alpha_1=\begin{pmatrix}-1\\2\\4\\0\end{pmatrix},\alpha_3=\begin{pmatrix}-1\\2\\0\\4\end{pmatrix},$$

$$\Longrightarrow\alpha_2=\begin{pmatrix}\frac{5}{4}\\-\frac{7}{2}\\0\\0\end{pmatrix},\alpha_4=\begin{pmatrix}\frac{1}{4}\\-\frac{3}{2}\\0\\0\end{pmatrix}\Longrightarrow P=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=\begin{pmatrix}-1&\frac{5}{4}&-1&\frac{1}{4}\\2&-\frac{7}{2}&2&-\frac{3}{2}\\4&0&0&0\\0&0&4&0\end{pmatrix}$$

7.7 Jordan 标准型的进一步讨论和应用

设 V 是 n 维复线性空间, φ 是 V 上的线性变换. 设 φ 的初等因子组为

$$(\lambda-\lambda_1)^{r_1},(\lambda-\lambda_2)^{r_2},\cdots,(\lambda-\lambda_k)^{r_k}, \tag{7.23}$$

由定理7.11可知, 存在 V 的一组基 $\{e_{11},e_{12},\cdots,e1r_1;e_{21},e_{22},\cdots,e_{2r_2};\cdots;e_{k1},e_{k2},\cdots,e_{kr_k}\}$, 使得 φ 在这组基下的表示矩阵为

$$J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \cdots & \\ & & & J_k \end{pmatrix}.$$

上式中每个 J_i 是相应于初等因子 $(\lambda-\lambda_i)^{r_i}$ 的 Jordan 块, 其阶正好为 r_i. 令 V_i 是由基元 $e_{i1},e_{i2},\cdots,e_{ir_i}$ 生成的子空间, 则

$$\begin{aligned} \varphi(e_{i1})&=\lambda_i e_{i1},\\ \varphi(e_{i2})&=e_{i1}+\lambda_i e_{i2},\\ &\cdots\cdots\\ \varphi(e_{ir_i})&=e_{ir_i-1}\lambda_i e_{ir_i}. \end{aligned} \tag{7.24}$$

这表明 $\varphi(V_i)\subseteq V_i$, 即 $V_i(i=1,2,\cdots,k)$ 是 φ 的不变子空间, 且有

$$V=V_1\oplus V_2\oplus\cdots\oplus V_k.$$

定理 7.12 线性变换 φ 的特征值 λ_1 的代数重数等于属于 λ_1 的 Jordan 块的阶数之和; λ_1 的几何重数等于属于 λ_1 的 Jordan 块的个数.

令 $\psi=\varphi-\lambda_i I_V$, 则(7.24)式可化为下式:

$$\begin{aligned} \psi(e_{i1})&=0,\\ \psi(e_{i2})&=e_{i1},\\ &\cdots\cdots\\ \psi(e_{ir_i})&=e_{ir_i-1}. \end{aligned} \tag{7.25}$$

即有如下循环链:

$$e_{ir_i}\xrightarrow{\psi}e_{ir_i-1}\xrightarrow{\psi}e_{ir_i-2}\xrightarrow{\psi}\cdots\xrightarrow{\psi}e_{i2}\xrightarrow{\psi}e_{i1}\xrightarrow{\psi}0.$$

从而

$$\{e_{ir_i},\psi(e_{ir_i}),\cdots,\psi^{r_i-1}(e_{ir_i})\}$$

构成了 V_i 的一组基. 从以下定义可知, V_i 是关于 $\psi=\varphi-\lambda_i I_V$ 的循环子空间.

定义 7.5 设 $\varphi\in\mathcal{L}(V)$, V_0 是 r 维 $\varphi-$ 不变子空间. 若存在 $0\neq\alpha\in V_0$ 使得 $\{\alpha,\varphi(\alpha),\cdots,\varphi^{r-1}(\alpha)\}$ 是 V_0 的一组基, 则称 V_0 为关于关于 φ 的循环子空间, α 称为循环向量.

引理 7.9 设 $R(\lambda_1)=V_1\oplus V_2\oplus\cdots\oplus V_s$, 则

$$R(\lambda_1)=\ker(\varphi-\lambda_1 I_V)^n=\{v\in V\mid(\varphi-\lambda_1 I_v)^n(v)=0\}.$$

证明: 先证 $R(\lambda_1)\subseteq\ker(\varphi-\lambda_i I_V)^n$:

任取 $v\in R(\lambda_1), v=v_1+v_2+\cdots+v_s, v_i\in V_i$, 则 V_i 的基有如下循环链:

$$e_{ir_i}\xrightarrow{\varphi-\lambda_1 I_V}e_{ir_i-1}\xrightarrow{\varphi-\lambda_1 I_V}e_{ir_i-2}\xrightarrow{\varphi-\lambda_1 I_V}\cdots\xrightarrow{\varphi-\lambda_1 I_V}e_{i2}\xrightarrow{\varphi-\lambda_1 I_V}e_{i1}\xrightarrow{\varphi-\lambda_1 I_V}0.$$

可得以下结论:

$$\begin{aligned}&(\varphi-\lambda_1 I_V)^{r_i}(e_{ij})=0,\forall j\Longrightarrow(\varphi-\lambda_1 I_V)^{r_i}(v_i)=0\\ \Longrightarrow&(\varphi-\lambda_1 I_V)^n(v)=(\varphi-\lambda_1 I_V)^n(v_1)+\cdots+(\varphi-\lambda_1 I_V)^n(v_s)=0\end{aligned}$$

上式中的 n 可缩小到 $\max\{r_1,r_2,\cdots,r_s\}$.

再证 $\ker(\varphi-\lambda_1 I)^n\subseteq R(\lambda_1)$:

设 $v\in V$ 在上述基下的坐标向量 $x=(x_{11},x_{12},\cdots,x_{1r_1},\cdots)'$, 求 $(J-\lambda_1 I_n)^n x=0$ 的解.

$$(J_{r_i}(\lambda_i)-\lambda_1 I_{r_i})^n=\begin{pmatrix}\lambda_i-\lambda_1 & 1 & & & \\ & \lambda_i-\lambda_1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ & & & & \lambda_i-\lambda_1\end{pmatrix}^n=\begin{cases}0,1\le i\le s\\ \text{非异},s<i\le k\end{cases}$$

由上式可推出

$$\begin{cases}x_{i1}-x_{i2}=\cdots=x_{ir_i}=0,s<i\le k\\ (x_{11},x_{12}\cdots,x_{1r_1};\cdots;x_{s1},x_{s2},\cdots,x_{sr_s})\text{为任意解}\end{cases}$$

由上式可得

$$\ker(\varphi-\lambda_1 I)^n=V_1\oplus V_2\oplus\cdots\oplus V_s=R(\lambda_1).$$

□

定义 7.6 $R(\lambda_1)=\ker(\varphi-\lambda_1 I)^n=V_1\oplus V_2\oplus\cdots\oplus V_s$ 称为特征值 λ_1 的根子空间.

注 (1) $R(\lambda_1)=\ker(\varphi-\lambda_1 I)^m, m=\max\{r_1,r_2,\cdots,r_s\}$ 或 m 为 λ_1 的代数重数.

(2) 特征子空间 $V_{\lambda_1}=\ker(\varphi-\lambda_1 I_V)\subseteq R(\lambda_1)$.

推论 7.12 $\varphi\in\mathcal{L}(V_{\mathbb{C}})$ 可对角化 $\Longleftrightarrow$ 对任意特征值 $\lambda_0, R(\lambda_0)=V_{\lambda_0}$.

证明:

$$\begin{aligned}&\forall\lambda_0,R(\lambda_0)=V_{\lambda_0}\Longleftrightarrow\dim R(\lambda_0)=\dim V_{\lambda_0}\\ \Longleftrightarrow&\lambda_0\text{的代数重数}=\text{几何重数}\\ \Longleftrightarrow&\varphi\text{有完全的特征向量系}\Longleftrightarrow\varphi\text{可对角化}\end{aligned}$$

$\square$

定理 7.13 $\varphi \in \mathcal{L}(V_{\mathbb{C}}^n)$.

(1) 设 φ 的初等因子组为 $(\lambda-\lambda_1)^{r_1}, (\lambda-\lambda_1)^{r_2}, \cdots, (\lambda-\lambda_1)^{r_k}$, 则 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_K$, V_i 是关于 $\varphi - \lambda_i I_V$ 的循环子空间;

(2) 设 φ 的不同特征值为 $\lambda_1, \cdots, \lambda_s$, 则 $V = R(\lambda_1) \oplus R(\lambda_2) \oplus \cdots \oplus R(\lambda_s)$.

评论 矩阵问题, 如果条件和结论在相似关系不改变, 则可将此问题化简到 Jordan 标准型, 进一步化简到 Jordan 块进行证明.

三段论: Jordan 块成立 $\Longrightarrow$ Jordan 标准型成立 $\Longrightarrow$ 一般矩阵成立.

例 7.6 设 $A \in M_n(\mathbb{C})$, 则 $A = BC$, 其中 B, C 为对称阵.

证明: $\forall$ 非异阵 P s. t. $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1), \cdots, J_{r_k}(\lambda_k)$,

$$J_{r_i}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix} = \begin{pmatrix} & & & 1 & \lambda_i \\ & & 1 & \lambda_i & \\ & \iddots & \iddots & & \\ 1 & \iddots & & & \\ \lambda_i & & & & \end{pmatrix} \begin{pmatrix} & & & 1 \\ & & \iddots & \\ & 1 & & \\ 1 & & & \end{pmatrix} = S_i \cdot T_i.$$

令

$$\begin{aligned} S = \mathrm{diag}\{S_1, S_2, \cdots, S_k\}, T = \mathrm{diag}\{T_1, T_2, \cdots, T_k\} \\ \Longrightarrow J = S \cdot T \end{aligned}$$

上式中, S, T 为对称阵.

$$\therefore A = PJP^{-1} = PSTP^{-1} = (PSP')((P^{-1})'TP^{-1}),$$

上式中, PSP' 和 $(P^{-1})'TP^{-1}$ 皆为对称阵. $\square$

例 7.7 $A \in M_n(\mathbb{C})$, 求 $A^k (k \geq 1)$.

解: $\exists$ 非异阵 P s. t. $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_k}(\lambda_k)\}$,

$$A^k = (PJP^{-1})^k = PJ^kP-1 = P\,\mathrm{diag}\{J_{r_1}(\lambda_1)^k, \cdots, J_{r_k}(\lambda_k)^k\}P^{-1}, \tag{7.26}$$

$$J_n(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 & & & \\ & \lambda_0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_0 \end{pmatrix} = \lambda_0 I_n + N \tag{7.27}$$

$$N = J_n(0) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix} \Longrightarrow N^n = 0 \tag{7.28}$$

N 为幂零矩阵, 且有如下性质:

$$N^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ & 0 & 0 & 1 & 0 \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 0 \\ & & & & 0 \end{pmatrix}, \cdots, N^{n-1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 & 0 \\ & & \ddots & \ddots & 0 \\ & & & \ddots & 0 \\ & & & & 0 \end{pmatrix} \tag{7.29}$$

因为 I 与 N 可互相交换位置相乘结果不变, 所以

$$\begin{aligned} J_n(\lambda_0)^k &= (\lambda_0 I_n + N)^k, k \geq n-1 \\ &= \lambda_0^k I_n + C_k^1 \lambda_0^{k-1} N + C_k^2 \lambda_0^{k-2} N^2 + \cdots + C_k^{n-1} \lambda_0^{k-n+1} N^{n-1} \\ &= \begin{pmatrix} \lambda_0^k & C_k^1 \lambda_0^{k-1} & C_k^2 \lambda_0^{k-2} & \cdots & C_k^{n-1} \lambda_0^{k-n+1} \\ & \lambda_0^k & C_k^1 \lambda_0^{k-1} & & \vdots \\ & & \ddots & \ddots & C_k^2 \lambda_0^{k-2} \\ & & & \ddots & C_k^1 \lambda_0^{k-1} \\ & & & & \lambda_0^k \end{pmatrix} \end{aligned} \tag{7.30}$$

上式中如果 $k < n-1$, 则右上角某些斜排会全为 0, 且 k 越小, 右上角的 0 斜行会越多. 将上式代入(7.26)式右边, 替换各 $J_{r_i}(\lambda_i)$, 即可求出 A^k.

引理 7.10 设 $A, B \in M_n(\mathbb{C})$, A, B 均可对角化, 且 $AB = BA$, 则 A, B 可同时对角化, 即 $\exists$ 非异阵 P s. t. $P^{-1}AP, P^{-1}BP$ 均为对角阵.

证明: 将引理换成几何语言: $\varphi, \psi \in \mathcal{L}(V_{\mathbb{C}}^n)$, φ, ψ 可对角化且 $\varphi \cdot \psi = \psi \cdot \varphi$, 则 φ, ψ 可同时对角化, 即 $\exists$ 公共基 s. t. φ, ψ 表示矩阵都为对角阵.

对维数 n 进行归纳. $n = 1$ 时, 显然成立. 设维数 $< n$ 时成立, 现证维数为 n 时的情形.

设 φ 的全体特征值为 $\lambda_1, \cdots, \lambda_s$:

Case 1: 若 $s = 1 \Longrightarrow \varphi = \lambda_1 I_V$ 为纯量阵.

$$\psi\text{可对角化} \Longrightarrow \exists\text{一组基}\{e_1, \cdots, d_n\}\text{ s. t. }\psi\text{ 的表示矩阵为对角阵.}$$

由于纯量变换在任何基下表示矩阵都为对角阵, $\therefore$ 在基$\{e_1, \cdots, e_n\}$下φ的表示矩阵也为对角阵.

Case 2: $s>1$, 设 φ 的各特征子空间为 $V_1,\cdots,V_s$.

$$\varphi\text{可对角化} \Longrightarrow V=V_1\oplus V_2\oplus\cdots\oplus V_s \Longrightarrow \dim V_i<n$$

下面证明 $\varphi\psi=\psi\varphi \Longrightarrow V_i$ 都是 $\psi-$ 不变子空间:

$\because$ 任取 $v_i\in V_i$, 即 $\varphi(v_i)=\lambda_i v_i$,

$$\begin{aligned}&\varphi(\psi(v_i))=\psi(\varphi(v_i))=\psi(\lambda_i v_i)=\lambda_i\psi(v_i)\\ \Longrightarrow&\psi(v_i)\in V_i\end{aligned}$$

做限制: $\varphi|_{V_i},\psi|_{V_i}$ 都可对角化.

$$\because \varphi\psi=\psi\varphi \Longrightarrow \varphi|_{V_i}\psi|_{V_i}=\psi|_{V_i}\varphi|_{V_i},$$

由归纳假设可知, $\exists V_i$ 的一组基 s. t. $\varphi|_{V_i},\psi|_{V_i}$ 在该基下的表示矩阵为对角阵.

将 V_i 的基拼成 V 的基, φ,ψ 的表示矩阵均为对角阵. □

定理 7.14 (Jordan-Chevalley 分解定理) 设 $A\in M_n(\mathbb{C})$, 则 $A=B+C$, 其中:

(1) B 可对角化;

(2) C 为幂零阵;

(3) $BC=CB$;

(4) B,C 均可表示为 A 的多项式.

并且满足条件 (1) (3) 的分解必定唯一.

证明: 先证存在性: 设 P 为非异阵 s. t. $P^{-1}AP=J=\mathrm{diag}\{J_1,J_2,\cdots,J_s\}$, 其中 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 A 的全体不同特征值,$J_1,J_2,\cdots,J_s$ 是 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 对应的根子空间的块.

$$J_i=\begin{pmatrix}J_{r_1}(\lambda_i)&&&\\&J_{r_2}(\lambda_i)&&\\&&\ddots&\\&&&J_{r_k}(\lambda_i)\end{pmatrix}=\lambda_i I+\begin{pmatrix}0&1&&&&&&&\\&0&1&&&&&&\\&&0&0&&&&&\\&&&0&1&&&&\\&&&&0&1&&&\\&&&&&0&0&&\\&&&&&&\ddots&\ddots&\\&&&&&&&0&1\\&&&&&&&&0\end{pmatrix}$$

上式右边 $\lambda_i I$ 为纯量阵可对角化, 令其为 M_i; 上式右边最后的矩阵主对角线上全部为 0, 上次对角线上在 Jordan 块交界处的元素为 0, 其余元素全部为 1, 显然为幂零阵, 令其为 N_i. 又因为纯量

阵与其他方阵相乘时皆可交换位置, 所以 $M_iN_i = N_iM_i$. 令

$$M = \mathrm{diag}\{M_1, M_2, \cdots, M_s\}, N = \mathrm{diag}\{N_1, N_2, \cdots, N_s\},$$

则 M 为对角阵, N 为幂零阵, 且

$$MN = NM, J = M + N.$$

$$J_i\text{的特征多项式} = (\lambda-\lambda_i)^{m_i} \xRightarrow{\text{Cayley-Hamilton 定理}} (J_i-\lambda_iI)^{m_i} = 0, \tag{7.31}$$

又因为 $(\lambda-\lambda_1)^{m_1}, \cdots, (\lambda-\lambda_s)^{m_s}$ 两两互素, 根据中国剩余定理可知:

$$\exists g(\lambda)\,\mathrm{s.\,t.}\ g(\lambda) = (\lambda-\lambda_i)^{m_i}q_i(\lambda) + \lambda_i, \forall 1 \le i \le s.$$

将 J_i 代入上式可得:

$$g(J_i) = (J_i-\lambda_iI)^{m_i}q_i(J_i) + \lambda_iI$$

由(7.31)式可知 $(J_i-\lambda_iI)^{m_i} = 0$, 代入上式可得:

$$g(J_i) = \lambda_iI = M_i$$

所以

$$g(J) = \mathrm{diag}\{g(J_1), \cdots, g(J_s)\} = \mathrm{diag}\{M_1, \cdots, M_s\} = M,$$

$$N = J - M = J - g(J),$$

从而 M, N 均可表示为 J 的多项式. 所以

$$A = PJP^{-1} = P(M+N)P^{-1} = PMP^{-1} + PNP^{-1},$$

令 $B = PMP^{-1}, C = PNP^{-1}$, 则 A 仍为可对角化矩阵,B 仍为幂零矩阵.

$$BC = (PMP^{-1})(PNP^{-1}) = PMNP^{-1} = PNMP^{-1} = (PNP^{-1})(PMP^{-1}) = CB.$$

$$g(A) = g(PJP^{-1})Pg(J)P^{-1} = PMP^{-1} = B, C = A - B = A - g(A).$$

从而 B, C 均可表示为 A 的多项式.

再证唯一性: 设 $A = B_1 + C_1$ 满足条件 (1) (3).

$$B_1C_1 = C_1B_1 \Longrightarrow AB_1 = (B_1+C_1)B_1 = B_1^2 + C_1B_1 = B_1^2 + B_1C_1 = B_1(B_1+C_1) = B_1A.$$

$$B = g(A) \Longrightarrow BB_1 = B_1B,$$

同理可推出 $CC_1 = C_1C$.

$$A = B + C = B_1 + C_1 \Longrightarrow B - B_1 = C_1 - C.$$

令 $C_1^r = 0, C^s = 0$, 则

$$(C_1-C)^{r+s} = \sum_{i+j=r+s} C_{r+s}^i C_1^i C^j$$

上式中, 若 $i \geq r$, 则 $C_1^i = 0$, 若 $j \geq s$, 则 $C^j = 0$; 且若 $i \leq r$, 则 $j \geq s$, 若 $j \leq s$, 则 $i \geq r$.

$$\therefore (C_1 - C)^{r+s} \equiv 0,$$

同理可得 $(B - B_1)^{r+s} \equiv 0$. 由引理7.10可知,∃ 非异阵 P s. t.

$$P^{-1}BP = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}, P^{-1}B_1P = \begin{pmatrix} b_1 & & & \\ & b_2 & & \\ & & \ddots & \\ & & & b_n \end{pmatrix}$$

$$\begin{aligned} 0 &= P^{-1}(B - B_1)^{r+s}P = (P^{-1}BP - P^{-1}B_1P)^{r+s} \\ &= \begin{pmatrix} (a_1 - b_1)^{r+s} & & & \\ & (a_2 - b_2)^{r+s} & & \\ & & \ddots & \\ & & & (a_n - b_n)^{r+s} \end{pmatrix} \\ &\implies a_i = b_i, \forall i \implies B = B_1. \end{aligned}$$

同理可证 $C_1 = C$. □

7.8 Jordan 标准型练习题

7.8.1 TOPIC 1

如何合理选取特征向量求出广义特征向量.

例 7.8 $A = \begin{pmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{pmatrix}$, 求 P s. t. $P^{-1}AP$ 为 Jordan 型.

解: 经计算可得 A 的初等因子组为 $(\lambda + 1), (\lambda + 1)^2$, 由此可推出 A 的 Jordan 标准型为

$$J = \begin{pmatrix} -1 & & \\ & -1 & 1 \\ & & -1 \end{pmatrix}.$$

设 $P = (\alpha_1, \alpha_2, \alpha_3)$, 由 $AP = PJ$ 可得

$$A\alpha_1 = -\alpha_1, A\alpha_2 = -\alpha_2, A\alpha_3 = \alpha_2 - \alpha_3.$$

由上式可知, α_1, α_2 是 A 的特征向量且

$$(A + I_3)\alpha_3 = \alpha_2 \tag{7.32}$$

其中 α_3 即广义特征向量.

$$A+I_3=\begin{pmatrix}3&6&-15\\1&2&-5\\1&2&-5\end{pmatrix}\longrightarrow\begin{pmatrix}1&2&-5\\0&0&0\\0&0&0\end{pmatrix}$$

$$\Longrightarrow A\text{的两个无关特征向量为}\beta_1=\begin{pmatrix}-2\\1\\0\end{pmatrix},\beta_2=\begin{pmatrix}5\\0\\1\end{pmatrix}.$$

将 β_1 或 β_2 代入(7.32)式代替 α_2, 例如

$$\left(A+I_3 \vdots \beta_2\right)=\left(\begin{array}{ccc:c}3&6&-15&5\\1&2&-5&0\\1&2&-5&0\end{array}\right)$$

上式代表的线性方程组无解, 将 β_1 代入也同样无解.

令 $\alpha_2=k_1\beta_1+k_2\beta_2=(-2k_1+5k_2,k_1,k_2)$, 要使(7.32)式有解, 则

$$\operatorname{rank}\left(\begin{array}{ccc:c}3&6&-15&-2k_1+5k_2\\1&2&-5&k_1\\1&2&-5&k_2\end{array}\right)=1\Longrightarrow k_1=k_2.$$

因此, 可令 $\alpha_2=\beta_1+\beta_2=(3,1,1)$, 可解出 $\alpha_3=(1,0,0)'$, 令 $\alpha_1=\beta_1$, 可求出

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

7.8.2 TOPIC 2

带参数矩阵 Jordan 标准型的确定.

(1) 选取特殊子式求出行列式因子;

(2) 计算几何重数以确定 Jordan 块的个数;

(3) 计算极小多项式以确定最大 Jordan 块的阶数.

例 7.9 (教材 P303 第 3 题) 求下列 n 阶上三角阵的行列式因子和不变因子:

$$A=\begin{pmatrix}a&1&1&\cdots&1\\0&a&1&\cdots&1\\0&0&a&\cdots&1\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&a\end{pmatrix}.$$

解: 方法一:

$$\lambda I - A = \begin{pmatrix} \lambda - a & -1 & -1 & \cdots & -1 \\ & \lambda - a & -1 & \cdots & -1 \\ & & \lambda - a & \cdots & -1 \\ & & & \ddots & \vdots \\ & & & & \lambda - a \end{pmatrix},$$

$$(\lambda I - A)\begin{pmatrix} 1 & \cdots & n-1 \\ 1 & \cdots & n-1 \end{pmatrix} = (\lambda - a)^{n-1}, \tag{7.33}$$

$$(\lambda I - A)\begin{pmatrix} 1 & \cdots & n-1 \\ 2 & \cdots & n \end{pmatrix} = g(\lambda). \tag{7.34}$$

因为 $g(a) = (-1)^{n-1} \neq 0$, 所以(7.33)式和(7.34)式无公根, 从而(7.33)式和(7.34)式互素. 由此可知,

$$D_{n-1}(\lambda) = 1, D_n(\lambda) = (\lambda - a)^n \Longrightarrow A\text{的行列式因子为}1, \cdots, 1, (\lambda - a)^n, \text{Jordan 标准型为}J_n(a).$$

方法二:

A 的特征值为 a, 代数重数为 n, 其几何重数 $= n - \operatorname{rank}(A - aI_n) = n - (n-1) = 1$. 由此可知 A 的 Jordan 块的个数为 1, 其 Jordan 标准型为 $J_n(a)$.

方法三:

A 的特征多项式为 $f(\lambda) = (\lambda - a)^n$, 设其极小多项式为 $m(\lambda) = (\lambda - a)^k, k \leq n$.

$$\begin{aligned} & A = aI_n + N + N^2 + \cdots + N^{n-1}, \text{其中}N = J_n(0) \\ \Longrightarrow & (A - aI_n)^{n-1} = (N + N^2 + \cdots + N^{n-1})^{n-1} = N^{n-1} \neq 0 \\ \Longrightarrow & m(\lambda) = (\lambda - a)^n \\ \Longrightarrow & A\text{的不变因子为}1, \cdots, 1, (\lambda - a)^n \\ \Longrightarrow & A\text{的 Jordan 标准型为}J_n(a). \end{aligned}$$

例 7.10

$$A = \begin{pmatrix} 1 & & & \\ a+2 & 1 & & \\ 5 & 3 & 1 & \\ 7 & 6 & b+4 & 1 \end{pmatrix},$$

求 A 的 Jordan 标准型.

解: A 的特征值全为 1,

$$\text{特征值 1 的几何重数} = 4 - \operatorname{rank}(A - I_4) = \begin{cases} 1, \text{若 } a+2 \neq 0 \text{ 且 } b+4 \neq 0 \\ 2, \text{若 } a+2 = 0 \text{ 或 } b+4 = 0 \end{cases}$$

当几何重数为 1 时, 只有一个 Jordan 块, $J = J_4(1)$.

当几何重数为 2 时, 有两个 Jordan 块, 设

$$J = \mathrm{diag}\{J_k(1), J_l(1)\}, \text{其中} 1 \le k \le l, k + l = 4 \Longrightarrow 2 \le l \le 3,$$

此时, 如果 $l = 2$, 则将 A 代入极小多项式 $(\lambda - 1)^2$ 将等于 0, 否则 $l = 3$.

$$(A - I_4)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 3(a+2) & 0 & 0 & 0 \\ 6(a+2)+5(b+4) & 3(b+4) & 0 & 0 \end{pmatrix}$$

$$= \begin{cases} 0, & \text{若 } a+2=0 \text{ 且 } b+4=0 \Longrightarrow k=l=2 \Longrightarrow J = \mathrm{diag}\{J_2(1), J_2(1)\} \\ \neq 0, & \text{若 } a+2=0 \text{ 和 } b+4=0 \text{ 之一成立} \Longrightarrow k=1, l=3 \Longrightarrow J = \mathrm{diag}\{1, J_3(1)\} \end{cases}$$

7.8.3 TOPIC 3

循环子空间的应用

定义 7.7 设 $\varphi \in \mathcal{L}(V_{\mathbb{K}}^n)$, $0 \neq \alpha \in V$, 由 $\{\alpha, \varphi(\alpha), \varphi^2(\alpha), \cdots\}$ 张成的子空间记为 $C(\varphi, \alpha)$, 称为 φ 关于循环向量 α 的循环子空间.

引理 7.11 设 $\dim C(\varphi, \alpha) = m$, 则 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{m-1}(\varphi)\}$ 是 $C(\varphi, \alpha)$ 的一组基.

证明: 设 $k = \max\{i \in \mathbb{Z}^+ \mid \alpha, \varphi(\alpha), \cdots, \varphi^{i-1}(\alpha)\text{线性无关}\}$

$$\Longrightarrow \begin{cases} \alpha, \varphi(\alpha), \cdots, \varphi^{k-1}(\alpha)\text{线性无关} \\ \alpha, \varphi(\alpha), \cdots, \varphi^{k-1}(\alpha), \varphi^k(\alpha)\text{线性相关} \end{cases}$$
$$\Longrightarrow \varphi^k(\alpha)\text{是}\alpha, \varphi(\alpha), \cdots, \varphi^{k-1}(\alpha)\text{的线性组合}$$

用数学归纳法可证明: $\forall i \ge k, \varphi^i(\alpha)$是$\alpha, \varphi(\alpha), \cdots, \varphi^{k-1}(\alpha)$的线性组合,

$$\Longrightarrow \{\alpha, \varphi(\alpha), \cdots, \varphi^{k-1}(\alpha)\}\text{是 } C(\varphi, \alpha) \text{ 的一组基}$$
$$\Longrightarrow \dim C(\varphi, \alpha) = k = m.$$

□

注 设

$$\varphi^m(\alpha) = -a_0\alpha - a_1\varphi(\alpha) - \cdots - a_{m-1}\varphi^{m-1}(\alpha),$$

令

$$g(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 \in k[x].$$

$C(\varphi,\alpha)$ 实际上是包含 α 的最小的 $\varphi-$ 不变子空间, $\varphi|_{C(\varphi,\alpha)}$ 在基 $\{\alpha,\varphi(\alpha),\cdots,\varphi^{m-1}(\alpha)\}$ 下的表示矩阵为

$$F=F(g(\lambda))=\begin{pmatrix} & & & -a_0 \\ 1 & & & -a_1 \\ & \ddots & & \vdots \\ & & 1 & -a_{m-1} \end{pmatrix},$$

即 $g(x)$ 的友阵.

例 7.11 (1) $\varphi(\alpha)=\lambda_0\alpha,(\lambda_0\in\mathbb{K},\alpha\neq 0)$, 则 $C(\varphi,\alpha)=L(\alpha)$.

(2) $\varphi(\alpha_i)=\lambda_i\alpha_i,\lambda_1,\cdots,\lambda_n\in\mathbb{K}$ 且互异, $\alpha\neq 0$, 令 $\alpha=\alpha_1+\alpha_2+\cdots+\alpha_n$, 则 $C(\varphi,\alpha)=V$.

(3) $\varphi\in\mathcal{L}(V_k^n)$, 不变因子为 $1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)$, 则

$$\exists\text{一组基}\{e_1,\cdots,e_n\}\,\text{s. t. }\varphi\text{的表示矩阵为}F=\operatorname{diag}\{F(d_1(\lambda)),\cdots,F(d_k(\lambda))\}. \tag{7.35}$$

如果一个线性映射的表示矩阵为一个友阵, 则其特征子空间必是循环子空间.

$$(\varphi(e_1),\cdots,\varphi(e_n))=(e_1,\cdots,e_n)\begin{pmatrix} & & & -a_0 \\ 1 & & & -a_1 \\ & \ddots & & \vdots \\ & & 1 & -a_{m-1} \end{pmatrix}$$

$$\varphi(e_1)=e_2,\varphi(e_2)=e_3,\cdots,\varphi(e_{n-1})=e_n,$$

$$\varphi(e_n)=-a_0e_1-a_1e_2-\cdots-a_{n-1}e_n$$

$$e_1\xrightarrow{\varphi}e_2\xrightarrow{\varphi}e_3\xrightarrow{\varphi}\cdots\xrightarrow{\varphi}e_n\xrightarrow{\varphi}-\sum_{i=1}^n a_{i-1}e_i$$

$$\text{由(7.35)}\Longrightarrow V=C(\varphi,e_{i1})\oplus C(\varphi,e_{i2})\oplus\cdots\oplus C(\varphi,e_{ik})$$

(4) $\varphi\in\mathcal{L}(V_{\mathbb{C}}^n)$, 初等因子为 $(\lambda-\lambda_1)^{r_1},\cdots,(\lambda-\lambda_k)^{r_k}$, $\exists$一组基$\{e_1,\cdots,e_n\}$ s. t. φ 的表示矩阵为 $J=\operatorname{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_k}(\lambda_k)\}\Longrightarrow V=C(\varphi-\lambda_1Ie_{i_1})\oplus C(\varphi-\lambda_2Ie_{i_2})\oplus\cdots\oplus C(\varphi-\lambda_kIe_{i_k})$.

命题 7.3 设 $V=C(\varphi,\alpha)$ 为循环空间,$\psi\in\mathcal{L}(V)$ 且 $\varphi\psi=\psi\varphi$, 则

(1) ψ 由 $\psi(\alpha)$ 的值唯一决定;

(2) $\exists g(\lambda)\in\mathbb{K}[\lambda]$ s. t. $\psi=g(\varphi)$.

证明: (1) $V=C(\varphi,\alpha)$ 的基为 $\{\alpha,\varphi(\alpha),\cdots,\varphi^{n-1}(\alpha)\}$, ψ 由 ψ 在基上的作用唯一决定.

$$\psi(\varphi^i(\alpha))=\varphi^i(\psi(\alpha)),0\le i\le n-1$$

$$\Longrightarrow\psi\text{由}\psi(\alpha)\text{唯一决定}.$$

(2)

$$\psi(\alpha)=a_0\alpha+a_1\varphi(\alpha)+\cdots+a_{n-1}\varphi^{n-1}(\alpha),$$

令 $g(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$,

$$\left.\begin{array}{l}\psi, g(\varphi)\text{可交换}\\ \psi(\alpha)=g(\varphi)(\alpha)\end{array}\right\}\Longrightarrow \psi=g(\varphi).$$

$\square$

推论 7.13 $A=F(g(\lambda)), AB=BA$, 则 $B=g(A)$.

推论 7.14 $A=J_n(\lambda_0), AB=BA$, 则 $B=g(A)$.

7.9 矩阵函数

定理 7.15 幂级数 $f(x)=\sum\limits_{n=0}^{\infty}$ 在 $|x|<r$ 时逐项可导, 即

$$f'(x)=\frac{d}{dx}(\sum_{n=0}^{\infty}a_nx^n)=\sum_{n=0}^{\infty}\frac{d}{dx}(a_nx^n)=\sum_{n=1}^{\infty}na_nx^{n-1},$$

且 $f'(x)$ 的收敛半径也是 r.

定义 7.8 $f(z)=a_0+a_1z+\cdots+a_pz^p\in\mathbb{C}[z], A\in M_n(\mathbb{C})$, 则 $f(A)\xlongequal{\text{def}}a_0I_n+a_1A+\cdots+a_pA^p$.

设 P 为非异阵 s.t. $P^{-1}AP=J=\operatorname{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_k}(\lambda_k)\}$, 则

$$f(A)=f(PJP^{-1})=Pf(J)P^{-1}=P\operatorname{diag}\{f(J_{r_1}(\lambda_1)),\cdots,f(J_{r_k}(\lambda_k))\}P^{-1},$$

其中

$$J_{r_i}(\lambda_i)=\lambda_iI+N, N=J_{r_i}(0), N^{r_i}=0.$$

在 $z=\lambda_i$ 处对 $f(z)$ 进行泰勒展开:

$$f(z)=f(\lambda_i)+\frac{f'(\lambda_i)}{1!}(z-\lambda_i)+\cdots+\frac{f^{(p)}(\lambda_i)}{p!}(z-\lambda_i)^p,$$

将 $z=J_{r_i}(\lambda_i), J_{r_i}(\lambda_i)=\lambda_iI+N$ 代入上式可得

$$f(J_{r_i}(\lambda_i))=f(\lambda_i)I+\frac{f'(\lambda_i)}{1!}N+\cdots+\frac{f^{(r_i-1)}(\lambda_i)}{(r_i-1)!}N^{r_i-1}=\begin{pmatrix}f(\lambda_i)&\frac{f'(\lambda_i)}{1!}&\cdots&\frac{f^{(r_i-1)}(\lambda_i)}{(r_i-1)!}\\ &f(\lambda_i)&\cdots&\frac{f^{(r_i-2)}(\lambda_i)}{(r_i-2)!}\\ &&\ddots&\vdots\\ &&&f(\lambda_i)\end{pmatrix}\tag{7.36}$$

例 7.12 设 $A\in M_n(\mathbb{C})$, 求 e^{tA}, 其中 t 是数值变量.

解: 设 $f(z)=e^{tz}$, 则 $f^{(i)}(z)=t^ie^{tz}$, 再设

$$P^{-1}AP=J=\operatorname{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_k}(\lambda_k)\},$$

则

$$e^{tA}=f(z)=Pf(J)P^{-1}=P\{f(J_{r_1}(\lambda_1)),\cdots,f(J_{r_k}(\lambda_k))\}P^{-1}, \tag{7.37}$$

上式中 $f(J_{r_i}(\lambda_i))=e^{tJ_{r_i}(\lambda_i)}$, 根据(7.36)式可得

$$\begin{aligned}
f(J_{r_i}(\lambda_i))=e^{tJ_{r_i}(\lambda_i)}=&\begin{pmatrix} e^{t\lambda_i} & \frac{t}{1!}e^{t\lambda_i} & \frac{t^2}{2!}e^{t\lambda_i} & \cdots & \frac{t^{r_i-1}}{(r_i-1)!}e^{t\lambda_i} \\ & e^{t\lambda_i} & \frac{t}{1!}e^{t\lambda_i} & \cdots & \frac{t^{r_i-2}}{(r_i-2)!}e^{t\lambda_i} \\ & & e^{t\lambda_i} & \cdots & \frac{t^{r_i-2}}{(r_i-2)!}e^{t\lambda_i} \\ & & & \ddots & \vdots \\ & & & & e^{t\lambda_i} \end{pmatrix} \\
=&e^{t\lambda_i}\cdot\begin{pmatrix} 1 & \frac{t}{1!} & \frac{t^2}{2!} & \cdots & \frac{t^{r_i-1}}{(r_i-1)!} \\ & 1 & \frac{t}{1!} & \cdots & \frac{t^{r_i-2}}{(r_i-2)!} \\ & & 1 & \cdots & \frac{t^{r_i-3}}{(r_i-3)!} \\ & & & \ddots & \vdots \\ & & & & 1 \end{pmatrix}.
\end{aligned} \tag{7.38}$$

将每个(7.38)式代入(7.37)式, 即可求出 e^{tA} 的表达式.

8 二次型

8.1 二次型的化简与矩阵的合同

引理 8.1 下列变换都是合同变换, 称为对称的初等变换:

(1) 对换 A 的第 i 行与第 j 行, 再对换第 i 列与第 j 列, 即 $P_{ij}AP_{ij}, P'_{ij}=P_{ij}$;

(2) A 的第 i 行乘以常数 C, 再将第 i 列乘以常数 C, 即 $P_i(C)AP_i(C), P'_i(C)=P_i(C)$;

(3) A 的第 i 行乘以常数 C 加到第 j 行上, 再将第 i 列乘以常数 C 加到第 j 列上, 即

$$T_{ij}(C)AT_{ji}(C), T'_{ij}(C)=T_{ji}(C).$$

引理 8.2 设 A 是数域 $\mathbb{K}$ 上的非零对称阵, 则必存在非异阵 C, 使 $C'AC$ 的第 $(1,1)$ 元素不等于零.

证明: 设 $A=(a_{ij})_{n\times n}$, 若 $a_{11}\neq 0$, 则结论成立;

下设 $a_{11}=0$, 若 $\exists 2\leq i\leq n\,\text{s.t.}\ a_{ii}\neq 0$, 则

$$A=\begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & a_{ii} & \\ & & & \ddots \end{pmatrix} \xrightarrow[P_{ij}AP_{ij}]{\text{第 (1) 类合同变换}} \begin{pmatrix} a_{ii} & & & \\ & \ddots & & \\ & & 0 & \\ & & & \ddots \end{pmatrix}; \tag{8.1}$$

下设 $\forall i, a_{ii}=0$, 由 $A\neq 0$ 可知 $\exists a_{ij}\neq 0, 1\leq i<j\leq n$, 则

$$A=\begin{pmatrix} & \vdots & & \vdots & \\ \cdots & 0 & \cdots & a_{ij} & \cdots \\ & \vdots & & \vdots & \\ \cdots & a_{ji} & \cdots & 0 & \cdots \\ & \vdots & & \vdots & \end{pmatrix} \xrightarrow[T_{ji}(1)AT_{ij}(1)]{\text{第 (3) 类合同变换}} \begin{pmatrix} & \vdots & & \vdots & \\ \cdots & 2a_{ij} & \cdots & a_{ij} & \cdots \\ & \vdots & & \vdots & \\ \cdots & a_{ji} & \cdots & 0 & \cdots \\ & \vdots & & \vdots & \end{pmatrix},$$

再重复(8.1)式的操作即可得结论. □

定理 8.1 设 A 是数域 $\mathbb{K}$ 上的 n 阶对称阵, 则必存在 $\mathbb{K}$ 上的 n 阶非异阵 C, 使 $C'AC$ 为对角阵.

推论 8.1 合同变换是若干次对称初等变换的复合.

证明: 设矩阵 A 可通过合同变换 $C'AC$ 化为对角阵, 其中 C 为非异阵.

$$\begin{aligned} &\because C\text{为非异阵, 故可令}C=P_1P_2\cdots P_r, P_i\text{为初等矩阵},\\ &\therefore C'AC=(P_1P_2\cdots P_r)'A(P_1P_2\cdots P_r)=(P'_r\cdots(P'_2(P'_1AP_1)P_2)\cdots P_r) \end{aligned} \tag{8.2}$$

□

定义 8.1 下列变换为合同变换, 称为分块对称初等变换:

(1) 对换 A 的第 i 分块行与第 j 分块行, 再对换第 i 分块列与第 j 分块列, 即 $P_{ij}AP'ij$;

(2) 第 i 分块行左乘非异阵 M, 第 i 分块列右乘 M', 即 $P_i(M)AP_i(M'), P_i'(M) = P_i(M')$;

(3) A 的第 i 分块行左乘矩阵 M 加到第 j 分块行上, 第 i 分块列右乘 M' 加到第 j 分块列上, 即 $T_{ij}(M)AT_{ji}(M'), T_{ij}(M)' = T_{ji}(M')$.

8.2 二次型的化简

8.2.1 配方法

基本公式

$$(\sum_{i=1}^{n} x_i)^2 = \sum_{i=1}^{n} x_i^2 + 2 \sum_{1 \le i < j \le n} x_i x_j$$

步骤 (1) 将含 x_1 的所有项凑成完全平方式 (消去 x_1);

(2) 剩余项中将含 x_2 的所有项凑成完全平方式 (消去 x_2);

……

(n-1) 剩余项中将含含 x_{n-1} 的所有项凑成完全平方式.

8.2.2 对称初等变换法

A 为对称阵, 则 $\exists$ 非异阵 C s. t. $C'AC = \Lambda$ 为对角阵, C 的形式如(8.2)所示.

方法 (1) 构造一个 $n \times 2n$ 的矩阵 $(\ A \,\vdots\, I\ )$, 对该矩阵整体实施行变换;

(2) 对左边的分块矩阵 A 实施对称的列变换, 直到左边的分块矩阵变为对角阵 Λ, 此时右边的分块阵是过度阵 C 的转置, 即 C'.

$$\begin{pmatrix} A \,\vdots\, I_n \end{pmatrix} \xrightarrow[\text{整体实施行变换}]{\text{整体左乘}P_1'} \begin{pmatrix} P_1'A \,\vdots\, P_1' \end{pmatrix} \xrightarrow[\text{对左边分块矩阵实施对称的列变换}]{\text{仅对左边分块矩阵右乘}P_1} \begin{pmatrix} P_1'AP_1 \,\vdots\, P_1' \end{pmatrix}$$
$$\xrightarrow{\text{多次实施对称初等变换}} \cdots \longrightarrow \begin{pmatrix} P_r' \cdots P_2'P_1'AP_1P_2 \cdots P_r \,\vdots\, P_r' \cdots P_2'P_1' \end{pmatrix}$$

此时, 左边的分块矩阵即为对角阵 Λ, 右边的分块矩阵即为 C'.

8.3 惯性定理

8.3.1 实二次型

设是对称阵 $A = \operatorname{diag}\{d_1, d_2, \cdots, d_r, 0, \cdots, 0\}$, 把正项和负项分别放在一起, 可设 $d_1 > 0, \cdots, d_p > 0; d_{p+1} < 0, \cdots, d_r < 0$. A 所代表的二次型为

$$f(x_1, x_2, \cdots, x_n) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_r x_r^2. \tag{8.3}$$

令

$$y_1 = \sqrt{d_1} x_1 1, \cdots, y_p = \sqrt{d_p} x_p;$$
$$y_{p+1} = \sqrt{-d_{p+1}} x_{p+1}, \cdots, y_r = \sqrt{-d_r} x_r;$$
$$y_j = x_j (j = r+1, \cdots, n),$$

则(8.3)式变为

$$f = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2. \tag{8.4}$$

该事实等价于 A 合同于下列对角阵:

$$\operatorname{diag}\{1, \cdots, 1; -1, \cdots, -1; 0, \cdots, 0\}, \tag{8.5}$$

上式中有 p 个 1,q 个 -1,$n-r$ 个零. (8.4)式中的二次型称为 f 的**规范标准型**.

定理 8.2 设 $f(x_1, x_2, \cdots, x_n)$ 是一个 n 元实二次型, 且 f 可化为两个标准型:

$$c_1 y_1^2 + \cdots + c_p y_p^2 - c_{p+1} y_{p+1}^2 - \cdots - c_r y_r^2,$$
$$d_1 z_1^2 + \cdots + d_k z_k^2 - d_{k+1} z_{k+1}^2 - \cdots - d_r z_r^2,$$

其中 $c_i > 0, d_i > 0$, 则必有 $p = k$.

定义 8.2 设 $f(x_1, x_2 \cdots, x_n)$ 是一个实二次型, 若它能化为形如(8.4)式的形状, 则称 r 是该二次型的秩, p 是它的正惯性指数, $q = r - p$ 是它的负惯性指数, $s = p - q$ 称为 f 的符号差.

定理 8.3 秩与符号差 (或正负惯性指数) 是是对称阵在合同关系下的全系不变量.

8.3.2 复二次型

因为复二次型

$$f(x_1, x_2, \cdots, x_n) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_r x_r^2$$

必可化为

$$z_1^2 + z_2^2 + \cdots + z_r^2,$$

其中 $z_i = \sqrt{d_i} x_i (i = 1, 2, \cdots, r), z_j = x_j (j = r+1, \cdots, n)$. 所以复对称阵的合同关系只有一个全系不变量, 那就是秩 r.

8.4 正定型与正定矩阵

例 8.1

$(1) f = x_1^2 + x_2^2 + \cdots + x_n^2$	正定型
$(2) f = x_1^2 + x_2^2 + \cdots + x_r^2 (r \leq n)$	半正定型
$(3) f = -x_1^2 - x_2^2 - \cdots - x_n^2$	负定型
$(4) f = -x_1^2 - x_2^2 - \cdots - x_r^2 (r \leq n)$	半负定型
$(5) f = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 (1 \leq p < r \leq n)$	不定型

定理 8.4 设实二次型 $f(x_1, \cdots, x_n)$ 的秩为 r, 正负惯性指数分别为 p, q, 则

(1) f 正定 $\Longleftrightarrow p = n$;

(2) f 半正定 $\Longleftrightarrow p = r$;

(3) f 负定 $\Longleftrightarrow q = n$;

(4) f 半负定 $\Longleftrightarrow q = r$;

(5) f 不定 $\Longleftrightarrow p > 0$ 且 $q > 0$.

证明: (1) 充分性: 设 $p = n$, 则存在非异线性变换

$$x = Cy \,\text{s.t.}\; f = y_1^2 + y_2^2 = \cdots + y_n^2.$$

任取 $0 \neq \alpha = (a_1, a_2, \cdots, a_n) \in \mathbb{R}$, 从而

$$\beta = C^{-1}\alpha\text{是非零实列向量}, \beta = (b_1, b_2, \cdots, b_n)',$$

则

$$f(a_1, a_2, \cdots, a_n) = b_1^2 + b_2^2 + \cdots + b_n^2 > 0 \Longrightarrow f\text{正定}.$$

必要性: 用反证法, 设 $p < n$, 则 $\exists$ 非异线性变换 $x = Cy$ 使得

$$f = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 (p \leq r).$$

令

$$b_1 = \cdots = b_p = 0, b_{p+1} = \cdots = b_n = 1,$$

从而 $\beta = (b_1, b_2, \cdots, b_n)$ 是非零实列向量, 故

$$\alpha = (a_1, a_2, \cdots, a_n)' = C\beta \neq 0\text{是非零列向量},$$

于是

$$f(a_1, a_2, \cdots, a_n) = b_1^2 + \cdots + b_p^2 - b_{p+1}^2 - \cdots - b_r^2 \leq 0,$$

这与 f 是正定型矛盾.

(2)(3)(4)(5) 可类似证明. □

定理 8.5 设 A 为实对称阵, $\text{rank}(A)=r$, 则

(1) A 正定阵 $\Leftrightarrow$ A 合同于 I_n;

(2) A 半正定阵 $\Longleftrightarrow$ A 合同于 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$;

(3) A 负定阵 $\Longleftrightarrow$ A 合同于 $-I_n$;

(4) A 半负定阵 $\Longleftrightarrow$ A 合同于 $\begin{pmatrix} -I_r & O \\ O & O \end{pmatrix}$;

(5) A 不定阵 $\Longleftrightarrow$ A 合同于 $\begin{pmatrix} I_p & & \\ & -I_q & \\ & & O \end{pmatrix}$.

评论

$$\begin{aligned} & f(A)\text{负定} && \Longleftrightarrow && -f(-A)\text{正定;} \\ & f(A)\text{半负定} && \Longleftrightarrow && -f(-A)\text{半正定.} \end{aligned}$$

定理 8.6 设 A 为 n 阶是对称阵, 则

$$A\text{正定} \Longleftrightarrow A\text{的 } n \text{ 个顺序主子式全大于零.}$$

证明: 必要性: 设 $A=(a_{ij})$ 为正定阵, 则

$$f(x_1,\cdots,x_n)=x'Ax=\sum_{j=1}^{n}\sum_{i=1}^{n}a_{ij}x_ix_j$$

为正定型, 令

$$f_k(x_1,\cdots,x_k)\xlongequal{\text{def}}f(x_1,\cdots,x_k,0,\cdots,0)=\sum_{j=1}^{k}\sum_{i=1}^{k}a_{ij}x_ix_j$$

f_k 的相伴对称矩阵为 A_k, 对 $\forall 0\neq(C_1,C_2,\cdots,C_k)'in\mathbb{R}^k$, 有

$$f_k(C_1,C_2,\cdots,C_k)=f(C_1,\cdots,C_k,0,\cdots,0)>0,$$

注意到 $(C_1,C_2,\cdots,C_k,0,\cdots,0)'$ 为非零 n 维列向量, 由此可知

$$f_k\text{为正定型} \Longleftrightarrow A_k\text{为正定阵}(1\leq k\leq n).$$

由定理8.5(1) 可知, 存在非异阵 $B_k\in M_k(\mathbb{R})$ s. t.

$$\begin{aligned} & B_k'A_kB_k=I_k \Longrightarrow 1=|I_k|=|B_k'A_kB_k|=|B_k|^2\cdot|A_k| \\ \Longrightarrow & |A_k|>0(1\leq k\leq n) \end{aligned}$$

充分性: 设 A 的 n 个顺序主子式全大于零, 对 n 进行归纳. 当 $n=1$ 时,$A=(a_{11})\Longleftrightarrow f=a_{11}x_1^2$, 由 $a_{11}>0$ 可知 $f=a_{11}x_1^2$ 为正定型, 即 A 为正定型. 设阶数 $<n$ 时结论成立, 现证 n 阶的情形. 此时 A 可写成如下形式:

$$A=\begin{pmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{pmatrix} \xrightarrow{\text{第一分块行乘以 } -\alpha'A_{n-1}^{-1} \text{ 加到第二分块行}} \begin{pmatrix} A_{n-1} & \alpha \\ O & a_{nn}-\alpha'A_{n-1}^{-1}\alpha \end{pmatrix}$$

$$\xrightarrow{\text{第一分块列乘以 } -A_{n-1}^{-1}\alpha \text{ 加到第二分块列}} \begin{pmatrix} A_{n-1} & O \\ O & a_{nn}-\alpha'A_{n-1}^{-1}\alpha \end{pmatrix}=B$$

A_{n-1} 的 $n-1$ 个顺序主子式是 A 的前 $n-1$ 个顺序主子式, 从而它们全大于零, 由归纳假设可知 A_{n-1} 为正定阵. 上式是是合同变换, 属于第三类分块对称初等变换, 所以不改变行列式的值. 因此

$$|A| = |A_{n-1}| \cdot (a_{nn} - \alpha' A_{n-1}^{-1} \alpha) = |B|.$$

上式中,

$$\because |A| > 0, |A_{n-1}| > 0, \therefore \alpha' A_{n-1}^{-1} \alpha > 0.$$

因为 A_{n-1} 正定, 故存在非异阵 $C_{n-1} \in M_{n-1}(\mathbb{R})$ s. t. $C'_{n-1} A_{n-1} C_{n-1} = I_{n-1}$, 因此

$$\begin{pmatrix} C'_{n-1} & O \\ O & 1 \end{pmatrix} \cdot B \cdot \begin{pmatrix} C_{n-1} & O \\ O & 1 \end{pmatrix} = \begin{pmatrix} I_{n-1} & O \\ O & a_{nn} - \alpha' A_{n-1}^{-1} \alpha \end{pmatrix}$$
$$\Longrightarrow A\text{的正惯性指数} = n$$

即 A 为正定阵. □

例 8.2 试求 t 的取值范围, 使下列二次型为正定型:

$$f(x_1, x_2, x_3, x_4) = x_1^2 + 4x_2^2 + 4x_3^2 + 3x_4^2 + 2t x_1 x_2 - 2x_1 x_3 + 4x_2 x_3.$$

解: 该二次型的相伴矩阵为

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

A 的顺序主子式为

$$|A_1| = 1 > 0, |A_2| = \begin{vmatrix} 1 & t \\ t & 4 \end{vmatrix} = 4 - t^2,$$

$$|A_3| = \begin{vmatrix} 1 & t & -1 \\ t & 4 & 2 \\ -1 & 2 & 4 \end{vmatrix} = -4(t-1)(t+2), |A_4| = -12(t-1)(t+2).$$

要使 A 正定, 必须

$$\begin{cases} 4 - t^2 > 0 \\ -4(t-1)(t+2) > 0 \end{cases}$$

解得 $-2 < t < 1$.

推论 8.2 设 A 为 n 阶正定实对称阵, 则

(1) A 的所有主子阵都是正定阵;

(2) A 的所有主子式全大于零, 特别地, 主对角元素全大于零;

(3) A 中之元素绝对值最大值只出现在主对角线上.

证明: (1) 取 A 中第 $i_1, i_2, \cdots, i_k$ 行、列构成主子阵 B, 设 A 对应的相伴二次型为

$$f(x_1, \cdots, x_n) = x'Ax,$$

令

$$f_B(x_{i_1}, x_{i_2}, \cdots, x_{i_k}) \xlongequal{\text{def}} f(0, \cdots, 0, x_{i_1}, 0, \cdots, 0, x_{i_2}, 0, \cdots, 0, x_{i_k}).$$

由 f 正定 $\Longrightarrow f_B$ 正定 $\Longrightarrow$ 相伴矩阵 B 正定.

(2) 是 (1) 的推论.

(3) 用反证法: 设绝对值最大的元素为 $|a_{ij}|$, 其中 $i < j$. 考虑 2 阶主子式

$$A\begin{pmatrix} i & j \\ i & j \end{pmatrix} = \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix} = a_{ii}a_{jj} - a_{ij}^2, \tag{8.6}$$

$$\because |a_{ij}| \geq a_{ii} > 0, |a_{ij}| \geq a_{jj} > 0, \therefore (8.6)\text{式小于零}.$$

这和 (2) 矛盾. □

推论 8.3 设 A 为 n 阶正定实对称阵, 则 A 的特征值全大于零.

证明: **事实**: 设

$$\xi = (a_1, a_2, \cdots, a_n)' \neq 0 \in \mathbb{C}^n,$$

则

$$\overline{\xi'} \cdot \xi = (\overline{a_1}, \overline{a_2}, \cdots, \overline{a_n}) \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \sum_{i=1}^{n} |a_i|^2 > 0.$$

任取 A 的特征值 λ, 对应特征向量为 ξ.

$$A\text{正定} \Longrightarrow \exists\text{非异阵}C \in M_n(\mathbb{R})\,\text{s.t.}\ A = C'I_nC = C'C.$$

$$A\xi = \lambda_0\xi \Longrightarrow \lambda_0\overline{\xi'}\xi = \overline{\xi'}A\xi = \overline{\xi'}C'C\xi = \overline{(C\xi)'}(C\xi).$$

$$\xi \neq 0 \Longrightarrow \overline{\xi'}\xi > 0, C\xi \neq 0 \Longrightarrow \overline{(C\xi)'}(C\xi) > 0 \Longrightarrow \lambda_0 = \overline{(C\xi)'}(C\xi)/\overline{\xi'}\xi > 0.$$

□

定理 8.7 设 A 为 n 阶实对称阵, 则下列结论等价:

(1) A 正定;

(2) A 合同于 I_n;

(3) $\exists$ 非异阵 $C \in M_n(\mathbb{R})\,\text{s.t.}\ A = C'C$;

(4) A 的所有主子式全大于零;

(5) A 的 n 个顺序主子式全大于零;

(6) A 的特征值全大于零.

证明: (1) $\xLeftrightarrow{\text{定理8.5}}$ (2) $\xLeftrightarrow{\text{定义}}$ 3

(1) $\xRightarrow{\text{推论8.2}}$ (4) $\Longrightarrow$ (5) $\xRightarrow{\text{定理8.6}}$ (1)

(1) $\xRightarrow{\text{推论8.3}}$ (6) $\xRightarrow{\text{第九章讲}}$ (1) □

8.5 半正定型和半正定矩阵

引理 8.3 设 A 是 n 阶实对称阵, 则

$$A\text{半正定} \Longleftrightarrow \exists C \in M_n(\mathbb{R})\ \text{s. t.}\ A = C'C.$$

证明: 先证充分性: 由定理8.5(2) 可知,

$$A\text{合同于}\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}, r = \operatorname{rank}(A),$$

即 $\exists$ 非异阵 $B \in M_n(\mathbb{R})$ s. t.

$$A = B'\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}B.$$

令

$$C = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}B,$$

则 $A = C'C$.

再证必要性: 设 $A = C'C$, 对 $0 \neq \alpha \in \mathbb{R}^n$ 有

$$\alpha' A\alpha = \alpha' C' C\alpha = (C\alpha)'(C\alpha).$$

设 $C\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}$, 则上式可写为:

$$\alpha' A\alpha = a_1^2 + a_2^2 + \cdots + a_n^2 \geq 0.$$

□

引理 8.4 设 A 为 n 阶实对称阵, 则

$$A\text{半正定} \Longleftrightarrow \forall t \in \mathbb{R}^+, A + tI_n\text{为正定阵}.$$

证明: 先证充分性: $\forall 0 \neq \alpha \in \mathbb{R}^n$, 有

$$\alpha'(A + tI_n)\alpha = \alpha' A\alpha + t\alpha'\alpha.$$

$$\because \alpha' A\alpha \geq 0, t\alpha'\alpha > 0, \therefore \alpha'(A+tI_n)\alpha > 0.$$

再证必要性: $\forall 0 \neq \alpha \in \mathbb{R}^n$, 有

$$\alpha'(A+tI_n)\alpha = \alpha' A\alpha + t\alpha'\alpha > 0.$$

令 $t \longrightarrow 0^+$, 则 $\alpha' A\alpha \geq 0$. □

例 8.3 $A = \mathrm{diag}\{1, 0, -1\}, |A_1| = 1, |A_2| = |A_3| = 0$, 不能推出 A 为半正定阵.

引理 8.5 设 A 是 n 阶实对称阵, 则

$$A\text{半正定} \Longleftrightarrow A\text{的所有主子式全大于或等于零}.$$

证明: 先证充分性: 令

$$f(x_1, x_2, \cdots, x_n) = x' Ax,$$

则 f 为半正定型. 取 A 第 $i_1, \cdots, i_r$ 行, 列交叉处的元素构成的主子阵 B, 构造二次型:

$$f_B(x_{i_1}, \cdots, x_{i_r}) = f(0, \cdots, 0, x_{i_1}, 0, \cdots, 0, x_{i_2}, 0, \cdots, 0, x_{i_r}).$$

$$f\text{半正定} \Longrightarrow f_B\text{半正定} \Longleftrightarrow B\text{半正定} \Longrightarrow A\begin{pmatrix} i_1 & \cdots & i_r \\ i_1 & \cdots & i_r \end{pmatrix} = |B| \geq 0.$$

再证必要性:

$$|A+tI_n| = t^n + C_1 t^{n-1} + \cdots + C_{n-1}t + C_n,$$

其中 C_i 为 A 的所有 i 阶主子式之和, 因此

$$C_i \geq 0, \forall 1 \leq i \leq n \Longrightarrow \forall t \in \mathbb{R}^+, |A+tI_n| > 0.$$

设 A_k 为 A 的顺序主子式 $(1 \leq k \leq n)$, 可得 A_k 的所有主子式都大于或等于零, 重复上述计算, 可得

$$|A_k + tI_k| > 0, \forall 1 \geq k \geq n, \forall t \in \mathbb{R}^+ \Longrightarrow A+tI_n\text{是正定阵}, \forall t \in \mathbb{R}^+ \xRightarrow{\text{引理8.4}} A\text{是半正定阵}.$$

□

引理 8.6 设 A 为 n 阶实对称阵, 则

$$A\text{半正定} \Longleftrightarrow A\text{的所有特征值全大于或等于零}.$$

证明: 设 A 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$.

$$\begin{aligned} A\text{半正定} &\xLeftrightarrow{\text{引理8.4}} A+tI_n\text{正定}, \forall t \in \mathbb{R}+ \\ &\Longleftrightarrow \lambda_i + t > 0, \forall 1 \leq i \leq n, \forall t > 0 \\ &\Longrightarrow \lambda_i \geq 0, \forall 1 \geq i \geq n. \end{aligned}$$

□

定理 8.8 设 A 为 n 阶实对称阵, 则下列结论等价:

(1) A 半正定;

(2) A 合同于 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$;

(3) $\exists C \in M_n(\mathbb{R})$ s. t. $A = C'C$;

(4) A 的所有主子式全部大于或等于零;

(5) A 的所有特征值全大于或等于零.

性质 8.5.1 设 $A = (a_{ij})$ 为半正定阵, 若 $a_{ii} = 0$, 则 A 的第 i 行和第 i 列元素全为零.

证明: 任取 $a_{ij}, a_{ji} (i \neq j)$ 所在的二阶主子式, 即

$$\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix} = -a_{ij}^2 \geq 0,$$

上式中 $A_{ii} = 0$, 从而 $\begin{cases} a_{ij} = 0 \\ a_{ji} = 0 \end{cases}$ □

性质 8.5.2 A 为半正定阵, $\alpha \in \mathbb{R}^n$ s. t. $\alpha' A \alpha = 0$, 则 $A\alpha = 0$.

证明: $\exists C \in M_n(\mathbb{R})$ s. t. $A = C'C$, 所以

$$0 = \alpha' A \alpha = \alpha' C' C \alpha = (C\alpha)'(C\alpha) \Longrightarrow C\alpha = 0 \Longrightarrow A\alpha = C'C\alpha = C'(C\alpha) = 0.$$

□

8.6 Hermite 型

二次型

$$f(x_1, x_2, \cdots, x_n) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_r x_r^2,$$

其中 $d_1, d_2, \cdots, d_r \in \mathbb{C}$, $x_1, x_2, \cdots, x_n$ 为复变元. 则 $\forall (a_1, a_2, \cdots, a_n)' \in \mathbb{C}^2$,

$$f(a_1, a_2, \cdots, a_n) = d_1 a_1^2 + d_2 a_2^2 + \cdots + d_r a_r^2 \in \mathbb{C}.$$

做如下调整:

(1) 加限制性条件, 使得 $d_1, d_2, \cdots, d_r$ 为实数;

(2) 将 x_i^2 变为 $|x_i|^2 = \overline{x_i} x_i$.

此时 $f \in \mathbb{R}$, 可写成如下形式:

$$f(x_1, x_2, \cdots, x_n) = d_1 \overline{x_1} x_1 + d_2 \overline{x_2} x_2 + \cdots + d_r \overline{x_r} x_r.$$

调整后, f 已经不是多项式了, 不是二次型, 而是 Hermite 型.

定义 8.3 下列 n 个复变元的二次齐次函数称为 Hermite 型:

$$f(x_1,x_2,\cdots,x_n)=\sum_{j=1}^{n}\sum_{i=1}^{n}a_{ij}\overline{x_i}x_j,$$

其中 $\overline{a_{ij}}=a_{ji}(\forall 1\le i,j\le n)$. 此时, f 为复变元实值函数.

引理 8.7 定义从 $\{n$阶 Hermite 阵$\}\longrightarrow\{n$元 Hermite 型$\}$ 的映射 φ:

$$A\mapsto f=\overline{x'}Ax,$$

求证:φ 为一一映射.

证明: 每个 Hermite 矩阵的各个元素都可写成一个 Hermite 型的对应项系数, 因此可知 φ 显然为满射. 下证 φ 为单射:

设 A,B 为 Hermite 阵且 $\overline{x'}Ax=\overline{x'}Bx$, 则

$$\overline{x'}(A-B)x=0. \tag{8.7}$$

令 $C=(c_{ij})=A-B$, x 可在复数域内取任意值(8.7)式都成立, 因此可令 x 分别取下列值:

$$x=\begin{array}{cc} & 1\\ \begin{array}{c}1\\ \\ i\\ \\ n\end{array} & \begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix}\end{array}\Longrightarrow C_{ii}=0(\forall i),\ x=\begin{array}{cc} & 1\\ \begin{array}{c}1\\ \\ i\\ \\ j\\ \\ n\end{array} & \begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix}\end{array}\Longrightarrow C_{ij}+C_{ji}=0(\forall i,j),$$

$$x=\begin{array}{cc} & 1\\ \begin{array}{c}1\\ \\ i\\ \\ j\\ \\ n\end{array} & \begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ \sqrt{-1}\\ \vdots\\ 0\end{pmatrix}\end{array}\Longrightarrow C_{ij}-C_{ji}=0(\forall i,j).$$

由以上可知, $A-B=C=0$, 因此, $A=B$. $\square$

定理 8.9 设 A 为 Hermite 阵, 则 $\exists$ 非异阵 $C \in \mathbb{C}^n$ s. t. $\overline{C'}AC$ 为是对角阵.

证明: 共轭对称初等变换为负相合变换: (1) 对换第 i 行与第 j 行, 对换第 i 列与第 j 列:

$$P_{ij}AP_{ij}, \overline{P'_{ij}} = P_{ij};$$

(2) 第 i 行乘以非零复数 C, 第 i 列乘以 $\overline{C}$:

$$P_i(C)AP_i(\overline{C}), \overline{P_i(C)'} = P_i(\overline{C});$$

(3) 第 i 行乘以复数 C 加到第 j 行上, 再将第 i 列乘以 $\overline{C}$ 加到第 j 列上:

$$T_{ij}(C)AT_{ji}(\overline{C}), \overline{T_{ij}(C)'} = T_{ji}(\overline{C}).$$

证明概要: 令 $A = (a_{ij})$.

(I) 若 $a_{11} \neq 0$, 则符合要求.

(II) 若 $a_{11} = 0, a_{ii} \neq 0 (2 \leq i \leq n)$, 则经过 $P_{1i}AP_{1i}$ 之后,$a_{11} \neq 0$.

(III) 若 $a_{ii} = 0 (\forall 1 \leq i \leq n)$, 任取 $a_{ij} \neq 0 (i < j)$.

$$A = \begin{array}{c} \\ i \\ j \end{array} \!\!\begin{array}{c} \begin{array}{cc} i & j \end{array} \\ \begin{pmatrix} 0 & a_{ij} \\ \overline{a_{ij}} & 0 \end{pmatrix} \end{array} \xrightarrow[\text{第 } j \text{ 行乘以 } a_{ij} \text{ 再加到第 } i \text{ 行上}]{\text{第 } j \text{ 列乘以 } \overline{a_{ij}} \text{ 再加到第 } i \text{ 列上}} \begin{array}{c} \\ i \\ j \end{array} \!\!\begin{array}{c} \begin{array}{cc} i & j \end{array} \\ \begin{pmatrix} 2|a_{ij}|^2 & a_{ij} \\ \overline{a_{ij}} & 0 \end{pmatrix} \end{array}$$

此时 a_{ii} 已不为零, 再通过上述第 (III) 步调整, 可使 $a_{11} \neq 0$. A 可变为下列形式:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & & \\ \vdots & & * & \\ a_{n1} & & & \end{pmatrix} \xrightarrow[\text{第 1 列分别乘以适当常数后加到第 } 2-n \text{ 列上}]{\text{第 1 行分别乘以适当常数后加到第 } 2-n \text{ 行上}} \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_{n-1} & \\ 0 & & & \end{pmatrix}$$

上式中, A_{n-1} 为 $n-1$ 阶 Hermite 阵, 再用归纳法即可得证. □

定理 8.10 (惯性定理) 设 $f(x_1, x_2, \cdots, x_n)$ 是一个 Hermite 型, 则它总可化为如下标准型:

$$\overline{y_1}y_1 + \cdots + \overline{y_p}y_p - \overline{y_{p+1}}y_{p+1} - \cdots - \overline{y_r}y_r,$$

且若 f 又可化为另一个标准型:

$$\overline{z_1}z_1 + \cdots + \overline{z_k}z_k - \overline{z_{k+1}}z_{k+1} - \cdots - \overline{z_r}z_r,$$

则 $p = k$.

定理 8.11 秩 r 与符号差 s(或正负惯性指数 p,q) 是 Hermite 型 (阵) 在复相合关系下的全系不变量.

定义 8.4 $f(x_1,\cdots,x_n)=\overline{x}'Ax$ 是 hermite 型, 其相伴 Hermite 矩阵为 A, $\forall 0\neq\alpha,\beta\in\mathbb{C}^n$, 则有如下定义和结论:

$\overline{\alpha}'A\alpha$	f	A
>0	正定	正定
≥ 0	半正定	半正定
<0	负定	负定
≤ 0	半负定	半负定
$\overline{\alpha}'A\alpha>0,\overline{\beta}'A\beta<0$	不定型	不定型

定理 8.12

$$(1) f\text{正定} \iff p=n \quad A\text{正定} \xLeftrightarrow{\text{复相合于}} I_n;$$

$$(2) f\text{半正定} \iff p=r \quad A\text{半正定} \xLeftrightarrow{\text{复相合于}} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix};$$

$$(3) f\text{负定} \iff q=n \quad A\text{负定} \xLeftrightarrow{\text{复相合于}} -I_n;$$

$$(4) f\text{半负定} \iff q=r \quad A\text{半负定} \xLeftrightarrow{\text{复相合于}} \begin{pmatrix} -I_r & O \\ O & O \end{pmatrix};$$

$$(5) f\text{不定} \iff p>0\text{且}q>0 \quad A\text{不定} \xLeftrightarrow{\text{复相合于}} \begin{pmatrix} I_p & & \\ & -I_q & \\ & & O \end{pmatrix}.$$

评论

$$f(A)\text{负定 (半负定)} \iff -f(-A)\text{正定 (半正定)}.$$

定理 8.13 设 A 为 Hermite 阵, 则下列结论等价:

(1) A 正定;

(2) A 复相合于 I_n;

(3) $\exists$ 非异阵 $C\in M_n(\mathbb{C})$ s. t. $A=\overline{C}'C$;

(4) A 的所有主子式全大于零;

(5) A 的 n 个顺序主子式全大于零;

(6) A 的特征值全大于零.

定理 8.14 设 A 为 Hermite 阵, 则下列结论等价:

(1) A 半正定;

(2) A 复相合于 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$;

(3) $\exists C \in M_n(\mathbb{C})\,\mathrm{s.\,t.}\ A = \overline{C'}C$;

(4) A 的所有主子式全大于或等于零;

(5) A 的特征值全大于或等于零.

9 内积空间

9.1 内积空间的概念

$$\mathbb{R}^3: \quad \vec{u}=(x_1,x_2,x_3), \quad \vec{v}=(y_1,y_2,y_3),$$

$$\text{点积 (内积):} \quad \vec{u}\cdot\vec{v}=x_1y_1+xy_2+xy_3,$$

$$\text{长度:} \quad \|\vec{u}\|=(\vec{u}\cdot\vec{u})=\sqrt{x_1^2+x_2^2+x_3^2},$$

$$\vec{u},\vec{v}\text{ 的距离:} \quad \|\vec{u}\cdot\vec{v}\|=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+(x_3-y_3)^2}.$$

$\mathbb{R}^3$ 中向量内积的性质

(1) $\vec{u}\cdot\vec{v}=\vec{v}\cdot\vec{u}$;

(2) $(\vec{u}+\vec{w})\cdot\vec{v}=\vec{u}\cdot\vec{v}+\vec{w}\cdot\vec{v}$;

(3) $(c\vec{u})\cdot\vec{v}=c(\vec{u}\cdot\vec{v})$;

(4) $\vec{u}\cdot\vec{u}\geq 0$, 且 $\vec{u}\cdot\vec{u}=0$ 当且仅当 $\vec{u}=0$.

其中 $\vec{u},\vec{v},\vec{w}$ 是 $\mathbb{R}^3$ 中的任意向量, c 是任一实数.

定义 9.1 设 V 为实线性空间, 若存在二元运算

$$\begin{aligned}(-,-):&V\times V\longrightarrow\mathbb{R}\\&\boldsymbol{\alpha}\times\boldsymbol{\beta}\mapsto(\boldsymbol{\alpha},\boldsymbol{\beta})\text{且满足如下性质:}\end{aligned}$$

$$\begin{aligned}&(1)(\boldsymbol{\beta},\boldsymbol{\alpha})=(\boldsymbol{\alpha},\boldsymbol{\beta});\text{对称性}\\&\left.\begin{aligned}&(2)(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha}+\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma});\\&(3)(c\boldsymbol{\alpha},\boldsymbol{\beta})=c(\boldsymbol{\alpha},\boldsymbol{\beta}),c\text{为任一实数;}\end{aligned}\right\}\text{第一变元的线性}\\&(4)(\boldsymbol{\alpha},\boldsymbol{\alpha})\geq 0\text{且等号成立当且仅当}\boldsymbol{\alpha}=0.\text{正定性}\end{aligned}$$

则称 $(-,-)$ 为 V 上的一个内积,$(\boldsymbol{\alpha},\boldsymbol{\beta})$ 称为向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的内积. 给定一个内积结构的实线性空间称为实内积空间. 有限维的内积空间称为 Euclid 空间 (欧氏空间).

定义 9.2 设 V 为复线性空间, 若存在二元运算

$$\begin{aligned}(-,-):&V\times V\longrightarrow\mathbb{C}\\&\boldsymbol{\alpha}\times\boldsymbol{\beta}\mapsto(\boldsymbol{\alpha},\boldsymbol{\beta})\text{且满足如下性质:}\end{aligned}$$

$$(1)(\boldsymbol{\beta},\boldsymbol{\alpha})=\overline{(\boldsymbol{\alpha},\boldsymbol{\beta})};\ \text{共轭对称性}$$

$$\left.\begin{array}{l}(2)(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha}+\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma});\\(3)(c\boldsymbol{\alpha},\boldsymbol{\beta})=c(\boldsymbol{\alpha},\boldsymbol{\beta}),c\text{为任一复数};\end{array}\right\}\text{第一变元的线性}$$

$$(4)(\boldsymbol{\alpha},\boldsymbol{\alpha})\geq 0\text{且等号成立当且仅当}\boldsymbol{\alpha}=0.\text{正定性}$$

则称 $(-,-)$ 为 V 上的一个内积,$(\boldsymbol{\alpha},\boldsymbol{\beta})$ 称为向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的内积. 给定一个内积结构的复线性空间称为复内积空间. 有限维的复内积空间称为酉空间.

注 (1) 由共轭对称性 $\Longrightarrow (\boldsymbol{\alpha},\boldsymbol{\alpha})\in\mathbb{R}\Longrightarrow$ (4) 的定义有意义;

(2) 对于实线性空间, 第二变元仍线性, 对于复线性空间, 第二变元为共轭线性, 即

$$(\boldsymbol{\gamma},c\boldsymbol{\alpha}+\boldsymbol{d\beta})=\overline{c}(\boldsymbol{\gamma},\boldsymbol{\alpha})+\overline{d}(\boldsymbol{\gamma},\boldsymbol{\beta}).$$

$$\because(\boldsymbol{\gamma},c\boldsymbol{\alpha}+\boldsymbol{d\beta})=\overline{(c\boldsymbol{\alpha}+\boldsymbol{d\beta},\boldsymbol{\gamma})}=\overline{c(\boldsymbol{\alpha},\boldsymbol{\gamma})+d(\boldsymbol{\beta},\boldsymbol{\gamma})}=\overline{c}\overline{(\boldsymbol{\alpha},\boldsymbol{\gamma})}+\overline{d}\overline{(\boldsymbol{\beta},\boldsymbol{\gamma})}=\overline{c}(\boldsymbol{\gamma},\boldsymbol{\alpha})+\overline{d}(\boldsymbol{\gamma},\boldsymbol{\beta}).$$

(3) 实内积空间的定义相容于复内积空间的定义, 因为实数的共轭等于它本身, 实内积空间是复内积空间的特殊情况.

例 9.1 设 $\mathbb{R}_n$ 是 n 维实列向量空间, $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)'$, $\boldsymbol{\beta}=(y_1,y_2,\cdots,y_n)'$, 定义

$$(\boldsymbol{\alpha},\boldsymbol{\beta})\triangleq\boldsymbol{\alpha}'\cdot\boldsymbol{\beta}=x_1y_1+x_2y_2+\cdots+x_ny_n,$$

则我们定义了一个内积, 这个内积称为 $\mathbb{R}_n$ 中的标准内积.

设 $\mathbb{R}^n$ 是 n 维实行向量空间, $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)$, $\boldsymbol{\beta}=(y_1,y_2,\cdots,y_n)$, 定义

$$(\boldsymbol{\alpha},\boldsymbol{\beta})\triangleq\boldsymbol{\alpha}\cdot\boldsymbol{\beta}'=x_1y_1+x_2y_2+\cdots+x_ny_n,$$

则我们定义了一个内积, 这个内积称为 $\mathbb{R}^n$ 中的标准内积.

例 9.2 设 $\mathbb{C}_n$ 是 n 维复列向量空间, $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)'$, $\boldsymbol{\beta}=(y_1,y_2,\cdots,y_n)'$, 定义

$$(\boldsymbol{\alpha},\boldsymbol{\beta})\triangleq\boldsymbol{\alpha}'\cdot\overline{\boldsymbol{\beta}}=x_1\overline{y_1}+x_2\overline{y_2}+\cdots+x_n\overline{y_n},$$

$$\overline{(\boldsymbol{\alpha},\boldsymbol{\beta})}=\overline{(\boldsymbol{\alpha}'\cdot\overline{\boldsymbol{\beta}})}=\overline{\boldsymbol{\alpha}}'\cdot\boldsymbol{\beta}=(\overline{\boldsymbol{\alpha}}'\cdot\boldsymbol{\beta})'=\boldsymbol{\beta}'\cdot\overline{\boldsymbol{\alpha}}=(\boldsymbol{\beta},\boldsymbol{\alpha}),$$

$$(\boldsymbol{\alpha},\boldsymbol{\alpha})=x_1\overline{x_1}+x_2\overline{x_2}+\cdots+x_n\overline{x_n}=|x_1|^2+|x_2|^2+\cdots+|x_n|^2\geq 0.$$

则在此定义下 $\mathbb{C}_n$ 成为一个酉空间, 上述内积称为 $\mathbb{C}_n$ 的标准内积.

设 $\mathbb{C}^n$ 是 n 维复行向量空间, $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)$, $\boldsymbol{\beta}=(y_1,y_2,\cdots,y_n)$, 定义

$$(\boldsymbol{\alpha},\boldsymbol{\beta})\triangleq\boldsymbol{\alpha}\cdot\overline{\boldsymbol{\beta}}'=x_1\overline{y_1}+x_2\overline{y_2}+\cdots+x_n\overline{y_n},$$

则在此定义下 $\mathbb{C}^n$ 成为一个酉空间, 上述内积称为 $\mathbb{C}^n$ 的标准内积.

例 9.3 设 $V=\mathbb{R}^2$ 为二维实行向量空间, 若 $\boldsymbol{\alpha}=(x_1,x_2), \boldsymbol{\beta}=(1,2)$, 定义:

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=x_1y_1-x_2y_1-x_1y_2+4x_2y_2,$$

容易验证定义9.1中的 (1),(2),(3) 都成立. 当 $\boldsymbol{\beta}=\boldsymbol{\alpha}$ 时, 上式为 (4) 也成立. 因此, $\mathbb{R}^2$ 在此内积下成为二维欧氏空间, 此内积称为**非标准内积**.

例 9.4 设 V 是 n 维实列向量空间, G 是 n 阶正定实对称阵, 对 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$, 定义 V 上内积:

$$(\boldsymbol{\alpha},\boldsymbol{\beta})\triangleq\boldsymbol{\alpha}'G\boldsymbol{\beta}.$$

证明:

$$(1)(\boldsymbol{\beta},\boldsymbol{\alpha})=\boldsymbol{\beta}'G\boldsymbol{\alpha}=(\boldsymbol{\beta}'G\boldsymbol{\alpha})'=\boldsymbol{\alpha}'G'\boldsymbol{\beta}=\boldsymbol{\alpha}'G\boldsymbol{\beta}=(\boldsymbol{\alpha},\boldsymbol{\beta})\text{对称性成立;}$$
$$(2)(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}'G\boldsymbol{\beta}\text{根据矩阵性质, 第一变元的线性显然成立;}$$
$$(3)(\boldsymbol{\alpha},\boldsymbol{\alpha})=\boldsymbol{\alpha}'G\boldsymbol{\alpha}\text{根据正定阵性质, 正定性显然成立.}$$

□

注 (1) 当 $G=I$ 时, V 上内积就是例9.1中的标准内积, 是例9.4的特例;

(2) 当 $G=\begin{pmatrix}1&-1\\-1&4\end{pmatrix}$ 时, V 上内积就是例9.3中的非标准内积;

(3) 设 V 是 n 维实行向量空间, G 是 n 阶正定实对称阵, 定义 V 上内积:

$$(\boldsymbol{\alpha},\boldsymbol{\beta})\triangleq\boldsymbol{\alpha}G\boldsymbol{\beta}';$$

例 9.5 对 n 维复列向量空间 U, 若有正定 Hermite 矩阵 H, 则可定义 U 上的内积:

$$(\boldsymbol{\alpha},\boldsymbol{\beta})\triangleq\boldsymbol{\alpha}'H\overline{\boldsymbol{\beta}}.$$

证明:

$$(1)\overline{(\boldsymbol{\alpha},\boldsymbol{\beta})}=\overline{\boldsymbol{\alpha}}'\overline{H}\boldsymbol{\beta}=(\overline{\boldsymbol{\alpha}}'\overline{H}\boldsymbol{\beta})'=\boldsymbol{\beta}'\overline{H}'\overline{\boldsymbol{\alpha}}=\boldsymbol{\beta}'H\overline{\boldsymbol{\alpha}}=(\boldsymbol{\beta},\boldsymbol{\alpha})\text{对称性成立;}$$
$$(2)\boldsymbol{\alpha}'H\overline{\boldsymbol{\beta}}\text{由矩阵性质可知, 第一变元线性成立;}$$
$$(3)(\boldsymbol{\alpha},\boldsymbol{\alpha})=\boldsymbol{\alpha}'H\overline{\boldsymbol{\alpha}}\text{由正定 Hermite 阵性质可知, 正定性成立.}$$

□

例 9.6 $V=C[a,b]$, 是 $[a,b]$ 上连续函数全体构成的实线性空间, $f(x),g(x)\in V$, 定义

$$(f(t),g(t))\triangleq\int_a^b f(t)g(t)\mathrm{d}t.$$

证明: 对称性由积分性质可知显然成立; 第一变元的线性由积分的线性可得; 下面证正定性:

$$(f(t),f(t))=\int_a^b f(t)^2\mathrm{d}t\geq 0\text{等号成立当且仅当}f(t)\equiv 0.$$

□

例 9.7 $V=\mathbb{R}[x]$, 假设 $n \geq m$,

$$f(x)=a_0+a_1x+\cdots+a_nx^n,$$
$$g(x)=b_0+b_1x+\cdots+b_mx^m.$$

$$(f(x),g(x))=a_0b_0+a_1b_1+\cdots+a_mb_m.$$

例 9.8 $V=M_n(\mathbb{R}), A,B,C \in V, c,d$ 为任一实数:

$$(A,B) \triangleq \operatorname{Tr}(AB').$$

证明:

(1) $(B,A)=\operatorname{Tr}(BA')=\operatorname{Tr}((BA')')=\operatorname{Tr}(AB')=(A,B)$;

(2) $(cA+dB,C)=\operatorname{Tr}((cA+dB)C')=c\operatorname{Tr}(AC')+d\operatorname{Tr}(BC')$;

(3) $(A,A)=\operatorname{Tr}(AA')=\sum\limits_{i,j=1}^{n}a_{ij}^2 \geq 0$等号成立当且仅当 $A=0$, 此处假设 $A=(a_{ij})$;

此内积称为 Frobenius 内积. □

定义 9.3 设 V 是一个内积空间,$\boldsymbol{\alpha} \in V$, $\boldsymbol{\alpha}$ 的长度 (范数) 为

$$\|\boldsymbol{\alpha}\|=(\boldsymbol{\alpha},\boldsymbol{\alpha})^{\frac{1}{2}},$$

$$\mathrm{d}(\boldsymbol{\alpha},\boldsymbol{\beta})=\|\boldsymbol{\alpha}-\boldsymbol{\alpha}\|.$$

例 9.9 $V=\mathbb{R}_n$, 标准内积, $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)'$,

$$\|\boldsymbol{\alpha}\|=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}.$$

$V=\mathbb{C}_n$, 标准内积, $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)'$,

$$\|\boldsymbol{\alpha}\|=\sqrt{|x_1|^2+|x_2|^2+\cdots+|x_n|^2}.$$

范数的性质:

(1) 非负性: $\|\boldsymbol{\alpha}\| \geq 0$且等号成立当且仅当 $\boldsymbol{\alpha}=0$;

(2) 对称性: $\mathrm{d}(\boldsymbol{\alpha},\boldsymbol{\beta})=\mathrm{d}(\boldsymbol{\beta},\boldsymbol{\alpha})$.

定理 9.1 设 V 是实或复的内积空间, $\boldsymbol{\alpha},\boldsymbol{\beta} \in V$, c 是任一常数（实数或复数）, 则

(1) $\|c\boldsymbol{\alpha}\|=|c|\|\boldsymbol{\alpha}\|$;范数的齐次性

(2) $|(\boldsymbol{\alpha},\boldsymbol{\beta})| \leq \|\boldsymbol{\alpha}\| \cdot \|\boldsymbol{\beta}\|$; Cauchy-Schwarz 不等式

(3) $\|\boldsymbol{\alpha}+\boldsymbol{\beta}\| \leq \|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\|$.三角不等式

证明: (1) $\|c\boldsymbol{\alpha}\|^2 = (c\boldsymbol{\alpha}, c\boldsymbol{\alpha}) = c \cdot \bar{c}(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = |c|^2 \cdot \|\boldsymbol{\alpha}\|^2$, 取平方根可得结论.

(2)(i) $\boldsymbol{\alpha} = \mathbf{0} : (\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\mathbf{0}, \boldsymbol{\beta}) = 0, \|\boldsymbol{\alpha}\| \cdot \|\boldsymbol{\beta}\| = 0\|\boldsymbol{\beta}\| = 0;$

(ii) $\boldsymbol{\alpha} \neq \mathbf{0}$ **:** 令

$$\boldsymbol{\gamma} = \boldsymbol{\beta} - \frac{(\boldsymbol{\beta}, \boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2} \cdot \boldsymbol{\alpha},$$

$$(\boldsymbol{\gamma}, \boldsymbol{\alpha}) = \left(\boldsymbol{\beta} - \frac{(\boldsymbol{\beta}, \boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2} \cdot \boldsymbol{\alpha}, \boldsymbol{\alpha}\right) = (\boldsymbol{\beta}, \boldsymbol{\alpha}) - \frac{(\boldsymbol{\beta}, \boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2}(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = 0.$$

计算 $\|\boldsymbol{\gamma}\|$:

$$\begin{aligned}
\|\boldsymbol{\gamma}\| &= (\boldsymbol{\gamma}, \boldsymbol{\gamma}) = (\boldsymbol{\gamma}, \boldsymbol{\beta} - \frac{(\boldsymbol{\beta}, \boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2} \cdot \boldsymbol{\alpha}) \\
&= (\boldsymbol{\gamma}, \boldsymbol{\beta}) - \frac{\overline{(\boldsymbol{\beta}, \boldsymbol{\alpha})}}{\|\boldsymbol{\alpha}\|^2}(\boldsymbol{\gamma}, \boldsymbol{\alpha}) \\
&= (\boldsymbol{\gamma}, \boldsymbol{\beta}) = (\boldsymbol{\beta} - \frac{(\boldsymbol{\beta}, \boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2} \cdot \boldsymbol{\alpha}, \boldsymbol{\beta}) = (\boldsymbol{\beta}, \boldsymbol{\beta}) - \frac{(\boldsymbol{\beta}, \boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2}(\boldsymbol{\alpha}, \frac{(\boldsymbol{\beta}, \boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \\
&= \|\boldsymbol{\beta}\|^2 - \frac{|(\boldsymbol{\alpha}, \boldsymbol{\beta})|^2}{\|\boldsymbol{\alpha}\|^2} \\
&\geq 0 \\
&\Longrightarrow |(\boldsymbol{\alpha}, \boldsymbol{\beta})| \leq \|\boldsymbol{\alpha}\| \cdot \|\boldsymbol{\beta}\|
\end{aligned}$$

上式中等号成立当且仅当 $\boldsymbol{\gamma} = 0 \Longleftrightarrow \boldsymbol{\alpha}, \boldsymbol{\beta}$ 线性相关.

(3)

$$\begin{aligned}
\|\boldsymbol{\alpha} + \boldsymbol{\beta}\|^2 &= (\boldsymbol{\alpha} + \boldsymbol{\beta}, \boldsymbol{\alpha} + \boldsymbol{\beta}) \\
&= (\boldsymbol{\alpha}, \boldsymbol{\alpha}) + (\boldsymbol{\beta}, \boldsymbol{\beta}) + (\boldsymbol{\alpha}, \boldsymbol{\beta}) + (\boldsymbol{\beta}, \boldsymbol{\alpha}) \\
&= \|\boldsymbol{\alpha}\|^2 + \|\boldsymbol{\beta}\|^2 + 2\mathrm{Re}(\boldsymbol{\alpha}, \boldsymbol{\beta}).
\end{aligned}$$

$$\begin{aligned}
&\because |\mathrm{Re}(\boldsymbol{\alpha}, \boldsymbol{\beta})| \leq |(\boldsymbol{\alpha}, \boldsymbol{\beta})| \leq \|\boldsymbol{\alpha}\| \cdot \|\boldsymbol{\beta}\|, \\
&\therefore \|\boldsymbol{\alpha} + \boldsymbol{\beta}\|^2 \leq \|\boldsymbol{\alpha}\|^2 + \|\boldsymbol{\beta}\|^2 + 2\|\boldsymbol{\alpha}\| \cdot \|\boldsymbol{\beta}\| = (\|\boldsymbol{\alpha}\| + \|\boldsymbol{\beta}\|)^2 \\
&\therefore \|\boldsymbol{\alpha} + \boldsymbol{\beta}\| \leq \|\boldsymbol{\alpha}\| + \|\boldsymbol{\beta}\|.
\end{aligned}$$

□

例 9.10 $V = \mathbb{R}_n$, 标准内积,$x_i, y_i \in \mathbb{R}$, Cauchy 不等式:

$$(x_1y_1 + x_2y_2 + \cdots + x_ny_n)^2 \leq (x_1^2 + x_2^2 + \cdots + x_n^2)(y_1^2 + y_2^2 + \cdots + y_n^2).$$

$V = C[a, b]$, 积分内积, Schwarz 不等式:

$$\left(\int_a^b f(t)g(t)\,\mathrm{d}t\right)^2 \leq \left(\int_a^b f(t)^2\,\mathrm{d}t\right)\left(\int_a^b g(t)^2\,\mathrm{d}t\right).$$

定义 9.4 设 $\mathbf{0}\neq\boldsymbol{\alpha},\mathbf{0}\neq\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\beta}\in V$, $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的夹角为:

$$\begin{aligned}
&\text{实内积} \qquad \cos\theta=\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|} \qquad \theta\in[0,\pi];\\
&\text{复内积} \qquad \cos\theta=\frac{|(\boldsymbol{\alpha},\boldsymbol{\beta})|}{\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|} \qquad \theta\in[0,\frac{\pi}{2}].
\end{aligned}$$

上式中, $|(\boldsymbol{\alpha},\boldsymbol{\beta})|$ 不能换成 $\mathrm{Re}(\boldsymbol{\alpha},\boldsymbol{\beta})$, 这样将失去几何意义.

定义 9.5 若 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$, 则称 $\boldsymbol{\alpha}$和$\boldsymbol{\beta}$ 正交 (垂直), 记为 $\boldsymbol{\alpha}\perp\boldsymbol{\beta}$,

$$(\boldsymbol{\alpha},\mathbf{0})=(\mathbf{0},\boldsymbol{\beta})=0\Longrightarrow\mathbf{0}\perp\boldsymbol{\alpha},\mathbf{0}\perp\boldsymbol{\beta}.$$

$$\text{若}\boldsymbol{\alpha}\neq\mathbf{0},\boldsymbol{\beta}\neq\mathbf{0}\text{且}\boldsymbol{\alpha}\perp\boldsymbol{\beta}\Longrightarrow\text{夹角}\theta=\frac{\pi}{2}.$$

定理 9.2 (勾股定理) 若 $\boldsymbol{\alpha}\perp\boldsymbol{\beta}$, 则

$$\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2.$$

证明:

$$\begin{aligned}
\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2&=(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta})\\
&=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2+(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\beta},\boldsymbol{\alpha})\\
&\xlongequal{(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})=0}\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2.
\end{aligned}$$

□

9.2 内积的表示和正交基

设 V 是内积空间,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n\}$ 是 V 的一组基. 令 $g_{ij}=(\boldsymbol{e}_i,\boldsymbol{e}_j),G=(g_{ij})_{n\times n}$. 取

$$\boldsymbol{\alpha}=\sum_{i=1}^n a_i\boldsymbol{e}_i\longleftrightarrow\boldsymbol{x}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},$$

$$\boldsymbol{\beta}=\sum_{j=1}^n b_j\boldsymbol{e}_j\longleftrightarrow\boldsymbol{y}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}.$$

Case 1: V 是欧氏空间:

$$\begin{aligned}
(\boldsymbol{\alpha},\boldsymbol{\beta}) &= (\sum_{i=1}^{n} a_i\boldsymbol{e}_i, \sum_{j=1}^{n} b_j\boldsymbol{e}_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_j(\boldsymbol{e}_i,\boldsymbol{e}_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_jg_{ij} \\
&= \sum_{j=1}^{n}(\sum_{i=1}^{n} a_ig_{ij})b_j = (a_1,a_2,\cdots,a_n)G\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \\
&= \boldsymbol{x}'G\boldsymbol{y}.
\end{aligned}$$

上式中, G 称 V 关于基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n\}$ 的内积矩阵或度量矩阵或 Gram 矩阵.

$$g_{ij} = (\boldsymbol{e}_i,\boldsymbol{e}_j) = (\boldsymbol{e}_j,\boldsymbol{e}_i) = g_{ji} \Longrightarrow G\text{是实对称矩阵}.$$

$$(\boldsymbol{\alpha},\boldsymbol{\alpha}) = \boldsymbol{x}'G\boldsymbol{x} > 0, \mathbb{R}^n \ni \forall\boldsymbol{x} \neq \mathbf{0}(\Longleftrightarrow V \ni \forall\boldsymbol{\alpha} \neq \mathbf{0}) \Longrightarrow G\text{是正定阵}.$$

反之, 设 V 是实线性空间,$\{e_1,e_2,\cdots,e_n\}$ 是给定的一组基, G 为 n 阶正定实对称阵, $\forall\boldsymbol{\alpha} \longleftrightarrow \boldsymbol{x}, \boldsymbol{\beta} \longleftrightarrow \boldsymbol{y}$, 则 V 上的内积

$$(\boldsymbol{\alpha},\boldsymbol{\beta})_G \triangleq \boldsymbol{\alpha}'G\boldsymbol{\beta} \Longrightarrow (-,-)_G\text{是 } V \text{ 上一个内积}.$$

设 V 是实线性空间,$\{e_1,e_2,\cdots,e_n\}$ 是给定的一组基,

$$\begin{gathered}
\{V\text{上所有内积结构}\} \xrightarrow{\text{一一对应}} \{n\text{阶正定实对称矩阵}\} \\
\varphi : (-,-) \longmapsto G = (\boldsymbol{e}_i,\boldsymbol{e}_j)_{n\times n} \\
\psi : (\boldsymbol{\alpha},\boldsymbol{\beta}) \triangleq \boldsymbol{\alpha}'G\boldsymbol{\beta} \longleftarrow G
\end{gathered}$$

Case 2: V 是酉空间:

$$\begin{aligned}
(\boldsymbol{\alpha},\boldsymbol{\beta}) &= (\sum_{i=1}^{n} a_i\boldsymbol{e}_i, \sum_{j=1}^{n} b_j\boldsymbol{e}_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\overline{b_j}(\boldsymbol{e}_i,\boldsymbol{e}_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\overline{b_j}g_{ij} \\
&= \sum_{j=1}^{n}(\sum_{i=1}^{n} a_ig_{ij})\overline{b_j} = (a_1,a_2,\cdots,a_n)G\begin{pmatrix} \overline{b_1} \\ \overline{b_2} \\ \vdots \\ \overline{b_n} \end{pmatrix} \\
&= \boldsymbol{x}'G\overline{\boldsymbol{y}}.
\end{aligned}$$

$$\overline{g_{ij}} = \overline{(\boldsymbol{e}_i,\boldsymbol{e}_j)} = (\boldsymbol{e}_j,\boldsymbol{e}_i) = g_{ji} \Longrightarrow G' = \overline{G}, \overline{G}' = G \Longrightarrow G\text{为 Hermite 阵}.$$

$$
\begin{aligned}
(\boldsymbol{\alpha},\boldsymbol{\alpha}) &= \boldsymbol{x}'G\overline{\boldsymbol{x}} > 0, 0 \neq \boldsymbol{x} \in \mathbb{C}^n \\
&= \boldsymbol{y}'G\overline{\boldsymbol{y}} > 0, 0 \neq \boldsymbol{y} \in \mathbb{C}^n \\
\Longrightarrow G\text{为正定 Hermite 阵}.
\end{aligned}
$$

反之, G 为 n 阶正定 Hermite 阵, 则 V 上的内积

$$(\boldsymbol{\alpha},\boldsymbol{\beta})_G \triangleq \boldsymbol{\alpha}'G\overline{\boldsymbol{\beta}}.$$

设 V 是复线性空间,$\{e_1,e_2,\cdots,e_n\}$ 是给定的一组基,

$$
\begin{aligned}
\{V\text{上所有内积结构}\} &\xrightarrow{\text{一一对应}} \{n\text{阶正定 Hermite 矩阵}\} \\
\varphi : (-,-) &\longmapsto G = (\boldsymbol{e}_i,\boldsymbol{e}_j)_{n\times n} \\
\psi : (\boldsymbol{\alpha},\boldsymbol{\beta}) \triangleq \boldsymbol{\alpha}'G\overline{\boldsymbol{\beta}} &\longleftarrow G
\end{aligned}
$$

命题 9.1

$$\exists\text{一组基}\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}\,\text{s. t. Gram 矩阵为}I_n \Longleftrightarrow (\boldsymbol{e}_i,\boldsymbol{e}_j) = \delta_{ij}, \forall i,j = 1,2,\cdots,n.$$

定义 9.6 设 V 是 n 维内积空间,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是 V 的一组基, 若 $(\boldsymbol{e}_i,\boldsymbol{e}_j)=0, \forall i\neq j$, 则称 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是 V 的一组正交基. 进一步, 若 $\|\boldsymbol{e}_i\|=1, \forall i=1,2,\cdots,n$, 则称 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是 V 的一组标准正交基.

例 9.11 $V=\mathbb{R}_n$(或$\mathbb{C}_n$), 标准内积, $\boldsymbol{e}_i=(0,0,\cdots,1,\cdots,0)'$, 其中 1 在第 i 个位置, 其他位置全为 0, 则 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是 V 的一组标准正交基.

例 9.12 $V=M_n(\mathbb{R})$,Frobenius 内积, E_{ij} 为 n 阶标准基矩阵, 其 (i,j) 元为 1, 其他元素为 0, 则 $\{E_{ij}\}$ 是 V 的一组标准正交基.

证明:

$$
\begin{aligned}
(E_{ij},E_{kl}) &= \mathrm{Tr}(E_{ij}E_{kl}') = \mathrm{Tr}(E_{ij}E_{lk}) \\
&= \delta_{jl}\,\mathrm{Tr}(E_{ik}) = \delta_{jl}\delta_{ik} \\
&= \begin{cases} 1, & (i,j)=(k,l), \\ 0, & (i,j)\neq(k,l). \end{cases}
\end{aligned}
$$

□

引理 9.1 两两正交的非零向量必线性无关.

证明: 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是两两正交的非零向量且

$$c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_m\boldsymbol{\alpha}_m=0.$$

$$
\begin{aligned}
0 &= (\sum_1^m c_i\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j) = \sum_1^m c_i(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j) = c_j(\boldsymbol{\alpha}_j,\boldsymbol{\alpha}_j) \\
&\Longrightarrow c_j = 0 \forall 1\leq j\leq m
\end{aligned}
$$

□

推论 9.1 n 维内积空间中两两正交的非零向量至多只有 n 个.

引理 9.2 设 $\boldsymbol{\beta}$ 与 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 正交, 则 $\boldsymbol{\beta}$ 与 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的线性组合也正交, 即 $\forall \boldsymbol{\gamma} \in L(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m)$ 都有 $(\boldsymbol{\beta}, \boldsymbol{\gamma}) = 0$.

证明: 设 $\boldsymbol{\gamma} = c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_m\boldsymbol{\alpha}_m$,

$$\begin{aligned}(\boldsymbol{\beta}, \boldsymbol{\gamma}) &= (\boldsymbol{\beta}, c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_m\boldsymbol{\alpha}_m) \\ &= c_1(\boldsymbol{\beta}, \boldsymbol{\alpha}_1) + c_2(\boldsymbol{\beta}, \boldsymbol{\alpha}_2) + \cdots + c_m(\boldsymbol{\beta}, \boldsymbol{\alpha}_m) \\ &= 0.\end{aligned}$$

□

定理 9.3 (Gram-Schmidt 正交化方法) 设 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 是 V 中线性无关的向量, 则存在一组两两正交的非零向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$, 使得 $L(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m) = L(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m)$.

证明: 对 m 进行归纳:

(1) 当 $m = 1$ 时, $\boldsymbol{v}_1 = \boldsymbol{u}_1$;

(2) 假设对 $m = k$ 成立, 即存在一组两两正交的非零向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$, 使得 $L(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k) = L(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k)$;

(3) 当 $m = k+1$ 时, 先从三维情况获得思路和启发. 假设 V 为三维实内积空间, 如下图所示: 设 $\boldsymbol{u}_3 = a\boldsymbol{v}_1 + b\boldsymbol{v}_2 + \boldsymbol{v}_3$, 则

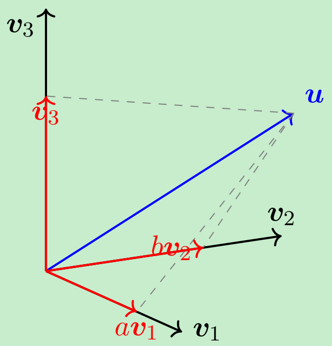

图 9.1: 向量 $\mathbf{u}$ 在正交坐标系 v_1, v_2, v_3 上的投影

$$\begin{aligned}&(\boldsymbol{u}_3, \boldsymbol{v}_1) = (a\boldsymbol{v}_1 + b\boldsymbol{v}_2 + \boldsymbol{v}_3, \boldsymbol{v}_1) = a(\boldsymbol{v}_1, \boldsymbol{v}_1) = a\|\boldsymbol{v}_1\|^2 \\ &\Longrightarrow a = \frac{(\boldsymbol{u}_3, \boldsymbol{v}_1)}{\|\boldsymbol{v}_1\|^2}, b = \frac{(\boldsymbol{u}_3, \boldsymbol{v}_2)}{\|\boldsymbol{v}_2\|^2} \\ &\Longrightarrow \boldsymbol{v}_3 = \boldsymbol{u}_3 - \frac{(\boldsymbol{u}_3, \boldsymbol{v}_1)}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 - \frac{(\boldsymbol{u}_3, \boldsymbol{v}_2)}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2.\end{aligned}$$

因此, 令

$$\boxed{\boldsymbol{v}_{k+1} = \boldsymbol{u}_{k+1} - \sum_{i=1}^{k} \frac{(\boldsymbol{u}_{k+1}, \boldsymbol{v}_i)}{(\boldsymbol{v}_i, \boldsymbol{v}_i)}\boldsymbol{v}_i.} \tag{9.1}$$

先证 $\boldsymbol{v}_{k+1} \neq 0$. 用反正法, 若 $v_{k+1}=0$, 则

$$\boldsymbol{u}_{k+1} \in L(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k) = L(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k)$$

这与 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{k+1}$ 线性无关矛盾. 再证 $\boldsymbol{v}_{k+1}$ 与 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$ 正交:

$$\begin{aligned}
\forall 1 \leq j \leq k, (\boldsymbol{v}_{k+1}, \boldsymbol{v}_j) &= (\boldsymbol{u}_{k+1} - \sum_{i=1}^{k} \frac{(\boldsymbol{u}_{k+1}, \boldsymbol{v}_i)}{(\boldsymbol{v}_i, \boldsymbol{v}_i)} \boldsymbol{v}_i, \boldsymbol{v}_j) \\
&= (\boldsymbol{u}_{k+1}, \boldsymbol{v}_j) - \sum_{i=1}^{k} \frac{(\boldsymbol{u}_{k+1}, \boldsymbol{v}_i)}{(\boldsymbol{v}_i, \boldsymbol{v}_i)} (\boldsymbol{v}_i, \boldsymbol{v}_j) \\
&= (\boldsymbol{u}_{k+1}, \boldsymbol{v}_j) - \frac{(\boldsymbol{u}_{k+1}, \boldsymbol{v}_j)}{(\boldsymbol{v}_j, \boldsymbol{v}_j)} (\boldsymbol{v}_j, \boldsymbol{v}_j) = 0.
\end{aligned}$$

从(9.1)式可知, $\boldsymbol{v}_{k+1} \in L(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{k+1})$,

$$\begin{aligned}
L(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{k+1}) &= L(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k) + L(\boldsymbol{u}_{k+1}) \\
&= L(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k) + L(\boldsymbol{u}_{k+1}) \subseteq L(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{k+1}).
\end{aligned}$$

$$\begin{aligned}
L(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{k+1}) &= L(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k) + L(\boldsymbol{v}_{k+1}) \\
&\subseteq L(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k) + L(\boldsymbol{u}_{k+1}) \\
&= L(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k) + L(\boldsymbol{u}_{k+1}) \\
&= L(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{k+1}).
\end{aligned}$$

□

注 从内积空间的一组基出发, 经过 Gram-Schmidt 正交化方法, 可得到一组正交基, 即

$$\boldsymbol{u} = L(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n) = L(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n).$$

设过度矩阵:

$$(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n) = (\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n) A.$$

$$\begin{aligned}
\boldsymbol{u}_1 &= \boldsymbol{v}_1, \\
\boldsymbol{u}_2 &= \boldsymbol{v}_2 - \frac{(\boldsymbol{v}_2, \boldsymbol{v}_1)}{(\boldsymbol{v}_1, \boldsymbol{v}_1)} \boldsymbol{v}_1, \\
\boldsymbol{u}_3 &= \boldsymbol{v}_3 - \frac{(\boldsymbol{v}_3, \boldsymbol{v}_1)}{(\boldsymbol{v}_1, \boldsymbol{v}_1)} \boldsymbol{v}_1 - \frac{(\boldsymbol{v}_3, \boldsymbol{v}_2)}{(\boldsymbol{v}_2, \boldsymbol{v}_2)} \boldsymbol{v}_2, \\
&\cdots \\
\boldsymbol{u}_{k+1} &= \boldsymbol{v}_{k+1} - \sum_{i=1}^{k} \frac{(\boldsymbol{u}_{k+1}, \boldsymbol{v}_i)}{(\boldsymbol{v}_i, \boldsymbol{v}_i)} \boldsymbol{v}_i
\end{aligned}$$

$$\Longrightarrow A = \begin{pmatrix} 1 & * & * & \cdots & * \\ 0 & 1 & * & \cdots & * \\ 0 & 0 & 1 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

单位化 (标准化):

$$\boldsymbol{v} \neq 0 \longrightarrow \boldsymbol{e} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \longleftarrow \left\| \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \right\| = 1.$$

令

$$\boldsymbol{w}_i = \frac{\boldsymbol{v}_i}{\|\boldsymbol{v}_i\|}, i = 1, 2, \cdots, n,$$

则 $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n\}$ 是 V 的一组标准正交基.

$$\begin{aligned}
(\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_n}) &= (\boldsymbol{w_1}, \boldsymbol{w_2}, \cdots, \boldsymbol{w_n})B \\
&= (\boldsymbol{w_1}, \boldsymbol{w_2}, \cdots, \boldsymbol{w_n}) \begin{pmatrix} \|\boldsymbol{v}_1\| & 0 & 0 & \cdots & 0 \\ 0 & \|\boldsymbol{v}_2\| & 0 & \cdots & 0 \\ 0 & 0 & \|\boldsymbol{v}_3\| & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \|\boldsymbol{v}_n\| \end{pmatrix} \\
\Longrightarrow (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n) &= (\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n) \cdot C.
\end{aligned}$$

其中,

$$C = B \cdot A = \begin{pmatrix} \|\boldsymbol{v}_1\| & * & * & \cdots & * \\ 0 & \|\boldsymbol{v}_2\| & * & \cdots & * \\ 0 & 0 & \|\boldsymbol{v}_3\| & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \|\boldsymbol{v}_n\| \end{pmatrix}.$$

定理 9.4 任一有限维内积空间必存在标准正交基.

证明:

$$\begin{aligned}
&\text{任取内积空间 } V \text{ 的一组基} \{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\} \\
&\xrightarrow{\text{Gram-Schmidt 正交化方法}} V \text{ 的一组正交基} \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\} \\
&\xrightarrow{\text{标准化 (单位化)}} V \text{ 的一组标准正交基} \{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n\}.
\end{aligned}$$

□

例 9.13 $V = \mathbb{R}^3$, 标准内积, $\boldsymbol{u}_1 = (3, 0, 4), \boldsymbol{u}_2 = (-1, 0, 7), \boldsymbol{u}_3 = (2, 9, 11)$, 用 Gram-Schmidt 正交化方法求 V 的一组标准正交基.

解:

$$
\begin{aligned}
&\boldsymbol{v}_1=\boldsymbol{u}_1=(3,0,4),\\
&\boldsymbol{v}_2=\boldsymbol{u}_2-\frac{(\boldsymbol{u}_2,\boldsymbol{v}_1)}{(\boldsymbol{v}_1,\boldsymbol{v}_1)}\boldsymbol{v}_1=(-1,0,7)-\frac{(-1,0,7)\cdot(3,0,4)}{(3,0,4)\cdot(3,0,4)}\cdot(3,0,4)\\
&=(-1,0,7)-\frac{25}{25}\cdot(3,0,4)=(-1,0,7)-(3,0,4)=(-4,0,3),\\
&\boldsymbol{v}_3=\boldsymbol{u}_3-\frac{(\boldsymbol{u}_3,\boldsymbol{v}_1)}{(\boldsymbol{v}_1,\boldsymbol{v}_1)}\boldsymbol{v}_1-\frac{(\boldsymbol{u}_3,\boldsymbol{v}_2)}{(\boldsymbol{v}_2,\boldsymbol{v}_2)}\boldsymbol{v}_2\\
&=(2,9,11)-\frac{(2,9,11)\cdot(3,0,4)}{(3,0,4)\cdot(3,0,4)}\cdot(3,0,4)-\frac{(2,9,11)\cdot(-4,0,3)}{(-4,0,3)\cdot(-4,0,3)}\cdot(-4,0,3)\\
&=(2,9,11)-\frac{50}{25}\cdot(3,0,4)-\frac{25}{25}\cdot(-4,0,3)=(2,9,11)-2\cdot(3,0,4)-(-4,0,3)=(0,9,0),\\
&\boldsymbol{w}_1=\frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}=\frac{1}{5}(3,0,4),\\
&\boldsymbol{w}_2=\frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|}=\frac{1}{5}(-4,0,3),\\
&\boldsymbol{w}_3=\frac{\boldsymbol{v}_3}{\|\boldsymbol{v}_3\|}=(0,1,0).
\end{aligned}
$$

引理 9.3 设 V 是 n 维内积空间,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$, $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 是 V 的两组基, 过度矩阵为 C, 即

$$
(\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n)=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)C,
$$

则

$$
\begin{aligned}
&V\text{ 为实内积空间时:}\quad G_{\boldsymbol{f}}=C'G_{\boldsymbol{e}}C,\\
&V\text{ 为复内积空间时:}\quad G_{\boldsymbol{f}}=C'G_{\boldsymbol{e}}\overline{C}.
\end{aligned}
$$

证明: 设

$$
C=(c_{ij})_{n\times n},G=(g_{ij})_{n\times n},
$$

$$
\begin{aligned}
&\because(\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n)=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)C,\\
&\therefore\boldsymbol{f}_k=\sum_{i=1}^{n}c_{ik}\boldsymbol{e}_i,\boldsymbol{f}_l=\sum_{j=1}^{n}c_{jl}\boldsymbol{e}_j,k,l=1,2,\cdots,n.
\end{aligned}
$$

$$
\begin{aligned}
(\boldsymbol{f}_k,\boldsymbol{f}_l)&=(\sum_{i=1}^{n}c_{ik}\boldsymbol{e}_i,\sum_{j=1}^{n}c_{jl}\boldsymbol{e}_j)\\
&=\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ik}c_{jl}(\boldsymbol{e}_i,\boldsymbol{e}_j)=\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ik}c_{jl}g_{ij}\\
&=\sum_{j=1}^{n}(\sum_{i=1}^{n}g_{ij}c_{ik})c_{jl}.
\end{aligned}
$$

□

注 在一组基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下，对于任一正定实对称矩阵 G_e, $\exists$非异阵C s.t. $C'G_eC=I_n$. 令

$$(\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n)=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)C,$$

由引理9.3可知,

$$G_{\boldsymbol{f}}=C'G_eC=I_n.$$

即 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 是 V 的一组标准正交基.

定义 9.7 设 U 是内积空间 V 的子空间,

$$U^{\perp}=\{\boldsymbol{v}\in V|(\boldsymbol{v},\boldsymbol{u})=0,\forall\boldsymbol{u}\in U\}$$

易证 $U^{\perp}$ 是 V 的子空间, 称为 U 的正交补空间.

定理 9.5 设 V 是 n 维内积空间,U 是 V 的子空间,

(1) $V=U\oplus U^{\perp}$;

(2) U 的标准正交基可扩充为 V 的标准正交基.

证明: (1) 设 $\dim U=m$, 将 V 上内积限制在 U 上, 则 U 是 m 维内积空间, 由定理9.3可知,U 有一组标准正交基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m\}$, 令

$$w=v-\sum_{i=1}^{m}(\boldsymbol{v},\boldsymbol{e}_i)\boldsymbol{e}_i,\forall\boldsymbol{v}\in V,$$

则

$$\begin{aligned}(\boldsymbol{w},\boldsymbol{e}_j)&=(\boldsymbol{v}-\sum_{i=1}^{m}(\boldsymbol{v},\boldsymbol{e}_i)\boldsymbol{e}_i,\boldsymbol{e}_j)\\&=(\boldsymbol{v},\boldsymbol{e}_j)-\sum_{i=1}^{m}(\boldsymbol{v},\boldsymbol{e}_i)(\boldsymbol{e}_i,\boldsymbol{e}_j)\\&=0,j=1,2,\cdots,m.\end{aligned}$$

从而 $(\boldsymbol{w},U)=0$, 即 $\boldsymbol{w}\in U^{\perp}$, 于是

$$U+U^{\perp}\ni\boldsymbol{v}=\sum_{i=1}^{m}(\boldsymbol{v},\boldsymbol{e}_i)\boldsymbol{e}_i+\boldsymbol{w}.$$

任取 $\boldsymbol{v}\in U\cap U^{\perp}$, 则 $(\boldsymbol{v},\boldsymbol{v})=0$, 即 $\boldsymbol{v}=0$, 从而 $U\cap U^{\perp}=\{\boldsymbol{0}\}$, 即 $V=U\oplus U^{\perp}$.

(2) 取 $U^{\perp}$ 的一组标准正交基 $\{\boldsymbol{e}_{m+1},\cdots,\boldsymbol{e}_n\}$, 则 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m,\boldsymbol{e}_{m+1},\cdots,\boldsymbol{e}_n\}$ 是 V 的一组标准正交基. □

定义 9.8 设 V 是 n 维内积空间,$V_1,\cdots,V_m$ 是 V 的子空间, 若对 $\forall\boldsymbol{\alpha}\in V_i,\forall\boldsymbol{\beta}\in V_j$, 有 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$, 则称 V_1 与 V_2 正交. 若 $V=V_1+V_2+\cdots+V_m$ 且 V_i 两两正交, 则称上述和为正交和, 记为 $V=V_1\perp V_2\perp\cdots\perp V_m$. 例如 $V=U+U^{\perp}$ 即为正交和.

引理 9.4 正交和都是直和, 称为正交直和.

证明: 先证 V_i 与 $\sum\limits_{j\neq i} V_j$ 正交. 任取 $\boldsymbol{v_i} \in V_i$ 有

$$(\boldsymbol{v_i}, \sum_{j\neq i} \boldsymbol{v_j}) = \sum_{j\neq i} (\boldsymbol{v_i}, \boldsymbol{v_j}) = 0.$$

任取 $\boldsymbol{\alpha} \in V_i \cap \sum\limits_{j\neq i} V_j$, 则 $(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = 0$, 由内积的正定性可知 $\boldsymbol{\alpha} = \mathbf{0}$, 从而

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_m.$$

$\square$

定义 9.9 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$, 对任一 $\boldsymbol{v} \in V$, $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2 + \cdots + \boldsymbol{v}_m, (\boldsymbol{v}_i \in V_i)$ 的分解唯一. 构造 $E_i \in \mathscr{L}(V), E_i(\boldsymbol{v}) = \boldsymbol{v}_i, 1 \leq i \leq m$, 称 E_i 为 V 的**投影变换**.

投影变换的性质

(1) $E_i^2 = E_i$;

(2) $E_i E_j = 0, i \neq j$;

(3) $E_1 + E_2 + \cdots + E_m = I_V$;

进一步, 设 $V = V_1 \perp V_2 \perp \cdots \perp V_m$ 为正交直和, 则投影变换 E_i 称为从 V 到 V_i 的**正交投影变换**.

引理 9.5 设 V 为 n 维内积空间, $V = U \perp U^{\perp}$, 记 E 为从 V 到 U 的正交投影变换, 则 $\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$,

$$(E(\boldsymbol{\alpha}), \boldsymbol{\beta}) = (\boldsymbol{\alpha}, E(\boldsymbol{\beta})).$$

证明: 设 $\boldsymbol{\alpha} = \boldsymbol{u}_1 + \boldsymbol{w}_1, \boldsymbol{\beta} = \boldsymbol{u}_2 + \boldsymbol{w}_2, \boldsymbol{u}_1, \boldsymbol{u}_2 \in U, \boldsymbol{w}_1, \boldsymbol{w}_2 \in U^{\perp}$, 则 $E(\boldsymbol{\alpha}) = \boldsymbol{u}_1, E(\boldsymbol{\beta}) = \boldsymbol{u}_2$,

$$\begin{aligned}(E(\boldsymbol{\alpha}), \boldsymbol{\beta}) &= (\boldsymbol{u}_1, \boldsymbol{u}_2 + \boldsymbol{w}_2) = (\boldsymbol{u}_1, \boldsymbol{u}_2)\\ (\boldsymbol{\alpha}, E(\boldsymbol{\beta})) &= (\boldsymbol{u}_1 + \boldsymbol{w}_1, \boldsymbol{u}_2) = (\boldsymbol{u}_1, \boldsymbol{u}_2).\end{aligned}$$

$\square$

定理 9.6 (Bessel 不等式) 设 V 是内积空间,$\boldsymbol{v_1}, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ 是 V 中两两正交的非零向量组, 则 $\forall \boldsymbol{y} \in V$, 有

$$\sum_{i=1}^{m} \frac{|(\boldsymbol{y}, \boldsymbol{v}_i)|^2}{\|\boldsymbol{v}_i\|^2} \leq \|\boldsymbol{y}\|^2.$$

等号成立当且仅当 $\boldsymbol{y}$ 可由 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m\}$ 线性表示.

证明: $\because \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m\}$ 两两正交, $\boldsymbol{y} \in V$, 设

$$\boldsymbol{z} = \boldsymbol{y} - \sum_{i=1}^{m} \frac{(\boldsymbol{y}, \boldsymbol{v}_i)}{\|\boldsymbol{v}_i\|^2} \boldsymbol{v}_i,$$

再令 $\boldsymbol{x} = \sum\limits_{i=1}^{m} \frac{(\boldsymbol{y},\boldsymbol{v}_i)}{\|\boldsymbol{v}_i\|^2} \boldsymbol{v}_i$, 则

$$\begin{aligned}
&(\boldsymbol{z}, \boldsymbol{v}_i) = (\boldsymbol{y}, \boldsymbol{v}_i) - \sum_{j=1}^{m} \frac{(\boldsymbol{y}, \boldsymbol{v}_j)}{\|\boldsymbol{v}_j\|^2} (\boldsymbol{v}_j, \boldsymbol{v}_i) = 0 \\
\Longrightarrow &(\boldsymbol{z}, \boldsymbol{x}) = \sum_{i=1}^{m} \frac{(\boldsymbol{y}, \boldsymbol{v}_i)}{\|\boldsymbol{v}_i\|^2} (\boldsymbol{z}, \boldsymbol{v}_i) = 0 \\
\Longrightarrow &\|\boldsymbol{y}\|^2 \xlongequal{\boldsymbol{y}=\boldsymbol{x}+\boldsymbol{z}} \|\boldsymbol{z}+\boldsymbol{x}\|^2 \xlongequal[\text{勾股定理}]{(\boldsymbol{z},\boldsymbol{x})=0} \|\boldsymbol{z}\|^2 + \|\boldsymbol{x}\|^2 \geq \|\boldsymbol{x}\|^2 \xlongequal[\text{勾股定理}]{} \sum_{i=1}^{m} \frac{|(\boldsymbol{y}, \boldsymbol{v}_i)|^2}{\|\boldsymbol{v}_i\|^2} \\
\Longrightarrow &\sum_{i=1}^{m} \frac{|(\boldsymbol{y}, \boldsymbol{v}_i)|^2}{\|\boldsymbol{v}_i\|^2} \leq \|\boldsymbol{y}\|^2.
\end{aligned}$$

上式中等号成立当且仅当 $\boldsymbol{z} = \boldsymbol{0}$, 即 $\boldsymbol{y} = \boldsymbol{x}$, 即 $\boldsymbol{y}$ 可由 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m\}$ 线性表示. □

注 上面定理对于无限维内积空间也成立: $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m, \cdots\}$ 为一列两两正交的非零向量组, 则

$$\sum_{i=1}^{\infty} \frac{|(\boldsymbol{y}, \boldsymbol{v}_i)|^2}{\|\boldsymbol{v}_i\|^2} \leq \|\boldsymbol{y}\|^2.$$

9.3 伴随

线性变换又称为线性算子, operator.

设 V 是 n 维内积空间,$\mathscr{L}(V)$ 是 V 上的线性变换全体, $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ 是 V 的一组标准正交基, $\varphi \in \mathscr{L}(V)$ 是线性算子. 任取 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, 分别是 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 在 $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ 下的坐标向量, 则

$$\boldsymbol{\alpha} = \sum_{i=1}^{n} a_i \boldsymbol{e}_i \longleftrightarrow \boldsymbol{x} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix},$$

$$\boldsymbol{\beta} = \sum_{i=1}^{n} b_i \boldsymbol{e}_i \longleftrightarrow \boldsymbol{y} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

$$\begin{aligned}
\varphi &\xleftrightarrow{\text{表示矩阵}} A = (a_{ij})_{n\times n} \in \mathbb{C}, \\
\varphi(\boldsymbol{\alpha}) &\longleftrightarrow A\boldsymbol{x}, \\
(\varphi(\boldsymbol{\alpha}), \boldsymbol{\beta}) &= (A\boldsymbol{x})'\overline{\boldsymbol{y}} = \boldsymbol{x}'A'\overline{\boldsymbol{y}} = \boldsymbol{x}'\overline{(\overline{A}'\boldsymbol{y})}.
\end{aligned}$$

$$
\begin{aligned}
\text{定义}\psi \in \mathscr{L}(V) &\xleftrightarrow{\text{表示矩阵}} \overline{A}' \\
\text{即}\psi(\boldsymbol{\beta}) &\longleftrightarrow \overline{A}'\boldsymbol{y}.
\end{aligned}
$$

$$
\begin{aligned}
&(\boldsymbol{\alpha}, \psi(\boldsymbol{\beta})) = \boldsymbol{x}'\overline{(\overline{A}'\boldsymbol{y})} \\
\Longrightarrow &(\varphi(\boldsymbol{\alpha}, \boldsymbol{\beta})) = (\boldsymbol{\alpha}, \psi(\boldsymbol{\beta})), \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V
\end{aligned}
$$

特别的, 当 V 为欧氏空间时: $\psi \in \mathscr{L}(V) \xleftrightarrow{\text{表示矩阵}} A'$.

定义 9.10 设 φ 是内积空间 V 上的线性算子, 若存在线性算子 φ^*, 使得 $\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$

$$(\varphi(\boldsymbol{\alpha}), \boldsymbol{\beta}) = (\boldsymbol{\alpha}, \varphi^*(\boldsymbol{\beta})),$$

则称 φ^* 是 φ 的**伴随算子**, 或简称**伴随**.

定理 9.7 (1) 伴随算子若存在, 必唯一;

(2) 有限维内积空间上任一线性算子, 必有伴随算子.

证明: (1) 设 φ^* 和 $\varphi^\#$ 都是 φ 的伴随算子, 即

$$
\begin{aligned}
&(\varphi(\boldsymbol{\alpha}), \boldsymbol{\beta}) = (\boldsymbol{\alpha}, \varphi^*(\boldsymbol{\beta})) = (\boldsymbol{\alpha}, \varphi^\#(\boldsymbol{\beta})) \\
\Longrightarrow &(\boldsymbol{\alpha}, \varphi^*(\boldsymbol{\beta}) - \varphi^\#(\boldsymbol{\beta})) = 0, \forall \boldsymbol{\alpha} \in V
\end{aligned}
$$

取 $\boldsymbol{\alpha} = \varphi^*(\boldsymbol{\beta}) - \varphi^\#(\boldsymbol{\beta})$, 则

$$(\varphi^*(\boldsymbol{\beta}) - \varphi^\#(\boldsymbol{\beta}), \varphi^*(\boldsymbol{\beta}) - \varphi^\#(\boldsymbol{\beta})) = 0,$$

由正定性可知 $\varphi^*(\boldsymbol{\beta}) = \varphi^\#(\boldsymbol{\beta})$, 从而

$$\varphi^* = \varphi^\#.$$

(2) 由伴随的唯一性及本节开始部分的构造可知,

$$\psi = \varphi^*,$$

所以, 有限维内积空间上任一线性算子必有伴随算子.

$$
\begin{aligned}
&\varphi \in \mathscr{L}(V) \xleftrightarrow{\text{表示矩阵}} A, \\
\Longrightarrow &\begin{cases} \text{酉空间} & \varphi^* \xleftrightarrow{\text{表示矩阵}} \overline{A}', \\ \text{欧氏空间} & \varphi^* \xleftrightarrow{\text{表示矩阵}} A'. \end{cases}
\end{aligned}
$$

□

定理 9.8 设 V 为 n 维内积空间,$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 是 V 的一组标准正交基,$\varphi \in \mathscr{L}(V)$ 的表示矩阵为 A,

定理 9.9 设 V 为内积空间,$\varphi, \psi \in \mathscr{L}$ 且 φ^*, ψ^* 存在, c 为常数, 则

(1) $(\varphi+\psi)^*=\varphi^*+\psi^*$;

(2) $(c\varphi)^*=\overline{c}\varphi^*$;

(3) $(\varphi\psi)^*=\psi^*\varphi^*$;

(4) $(\varphi^*)^*=\varphi$;

(5) 若 φ 可逆, 则 φ^* 也可逆, 且 $(\varphi^*)^{-1}=(\varphi^{-1})^*$.

证明: 若 V 是有限维内积空间, 设 φ 的表示矩阵为 A, ψ 的表示矩阵为 B. $\{\boldsymbol{e}_1,\boldsymbol{e}_2,cdots,\boldsymbol{e}_n\}$ 是 V 的一组标准正交基, 则

(1) $\overline{(A+B)}'=\overline{A}'+\overline{B}'$;

(2) $\overline{(cA)}'=\overline{c}\overline{A}'$;

(3) $\overline{(AB)}'=\overline{B}'\overline{A}'$;

(4) $\overline{\overline{A}'}'=A$;

(5) 若 A 可逆, 则 $\overline{A}'$ 也可逆, 且 $(\overline{A}')^{-1}=\overline{A^{-1}}$.

若 V 是无限维内积空间,(1)-(2) 很容易证明, 下面证 (3)-(5):

(3) $(\varphi\psi(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\psi(\boldsymbol{\alpha}),\varphi^*(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\psi^*\varphi^*(\boldsymbol{\beta}))\Longrightarrow(\varphi\psi)^*=\psi^*\varphi^*$.

(4) $(\varphi^*(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\varphi(\boldsymbol{\beta}))\Longrightarrow(\varphi^*)^*=\varphi$.

(5) $(\varphi\varphi^{-1}(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\varphi^{-1}(\boldsymbol{\alpha}),\varphi^*(\boldsymbol{\beta}))$.

□

命题 9.2 设 V 是 n 维内积空间, $\varphi\in\mathscr{L}(V)$,

(1) 若 U 是 $\varphi-$ 不变子空间, 则 $U^\perp$ 是 φ^*- 不变子空间;

(2) 若 φ 的全体特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则 φ^* 的全体特征值为 $\overline{\lambda_1},\overline{\lambda_2},\cdots,\overline{\lambda_n}$.

证明: (1) 任取 $\boldsymbol{u}\in U,\boldsymbol{w}\in U^\perp$, 则

$$\begin{aligned}&(\varphi(\boldsymbol{u}),\boldsymbol{w})=(\boldsymbol{u},\varphi^*(\boldsymbol{w}))=0,\\ \Longrightarrow&\varphi^*(\boldsymbol{w})\in U^\perp,\forall\boldsymbol{w}\in U^\perp.\end{aligned}$$

(2) 取标准正交基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$, φ 的表示矩阵为 A, φ^* 的表示矩阵为 $\overline{A}'$, 由题意可知,

$$|\lambda I_n-A|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n),$$

则

$$\begin{aligned}
|\lambda I_n - \overline{A}'| &= |\lambda I_n - \overline{A}| \\
\xlongequal{令\lambda=\overline{\mu}} \overline{|\mu I_n - A|} &= \overline{(\mu-\lambda_1)(\mu-\lambda_2)\cdots(\mu-\lambda_n)} \\
&=(\overline{\mu}-\overline{\lambda_1})(\overline{\mu}-\overline{\lambda_2})\cdots(\overline{\mu}-\overline{\lambda_n}) \\
&=(\lambda_1-\mu)(\lambda_2-\mu)\cdots(\lambda_n-\mu).
\end{aligned}$$

□

例 9.14 $V = U \perp U^{\perp}$, E 是从 U 到 V 的正交投影变换, 且

$$(E(\boldsymbol{\alpha}), \boldsymbol{\beta}) = (\boldsymbol{\alpha}, E(\boldsymbol{\beta})), \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V.$$

则

$$E^* = E.$$

E 称为**自伴算子**.

例 9.15 $V = M_n(\mathbb{R})$, Frobenius 内积, 且 $\varphi \in \mathscr{L}(V), \varphi(A) = PAQ, P, Q \in M_n(\mathbb{R})$. 则

$$\begin{aligned}
(\varphi(A), B) &= \mathrm{Tr}(\varphi(A)B') = \mathrm{Tr}((PAQ)B') \\
&= \mathrm{Tr}(AQB'P') = \mathrm{Tr}(A(P'BQ')') = (A, \varphi^*(B)).
\end{aligned}$$

9.4 内积空间的同构，正交变换和酉变换

设 V 为欧氏空间, $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 为标准正交基. 线性同构:

$$\varphi : V \longrightarrow \mathbb{R}^n(\text{标准内积})$$

$$\boldsymbol{\alpha} = \sum_{i=1}^{n} a_i \boldsymbol{e}_i \longmapsto \boldsymbol{x} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

$$\boldsymbol{\beta} = \sum_{i=1}^{n} b_i \boldsymbol{e}_i \longmapsto \boldsymbol{y} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \boldsymbol{x}'\boldsymbol{y},$$

$$(\varphi(\boldsymbol{\alpha}), \varphi(\boldsymbol{\beta})) = \boldsymbol{x}'\boldsymbol{y}.$$

$$\therefore (\varphi(\boldsymbol{\alpha}), \varphi(\boldsymbol{\beta}))_{\mathbb{R}^n} = (\boldsymbol{\alpha}, \boldsymbol{\beta})_V, \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V.$$

设 V 为酉空间, $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 为标准正交基. 线性同构:

$$\varphi : V \longrightarrow \mathbb{C}^n(\text{标准内积})$$

$$(\varphi(\boldsymbol{\alpha}), \varphi(\boldsymbol{\beta}))_{\mathbb{C}^n} = (\boldsymbol{\alpha}, \boldsymbol{\beta})_V, \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V.$$

定义 9.11 设 V, U 为实内积 (复内积) 空间, $\varphi: V \longrightarrow U$ 是线性映射, 若 $\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$ 有

$$(\varphi(\boldsymbol{\alpha}), \varphi(\boldsymbol{\beta}))_U = (\boldsymbol{\alpha}, \boldsymbol{\beta})_V,$$

则称 φ 是保持内积的线性映射. 若进一步, φ 为线性同构, 则称 φ 为保积同构.

注 (1) 线性空间的保积同构是等价关系;

(2) 保积同构的线性映射必为单射;

$$\begin{aligned}
&\because 任取\boldsymbol{\alpha} \in \ker \varphi, 即\varphi(\boldsymbol{\alpha}) = \mathbf{0}, 则(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = (\varphi(\boldsymbol{\alpha}), \varphi(\boldsymbol{\alpha})) = 0 \\
&\therefore 由正定性可知\boldsymbol{\alpha} = \mathbf{0}.
\end{aligned}$$

例 9.16 嵌入映射

$$\begin{aligned}
\varphi: \mathbb{R}^2 &\longrightarrow \mathbb{R}^3(\text{标准内积}) \\
(a, b)' &\longmapsto (a, b, 0)'
\end{aligned}$$

这个映射只能填满 $x-y$ 平面, 不能填满整个 $\mathbb{R}^3$.

命题 9.3 保积映射必是保范 (保距) 映射.

证明: 令 $\boldsymbol{\alpha} = \boldsymbol{\beta}$,

$$\begin{aligned}
&\|\varphi(\boldsymbol{\alpha})\|^2 = (\varphi(\boldsymbol{\alpha}), \varphi(\boldsymbol{\alpha})) = (\boldsymbol{\alpha}, \boldsymbol{\alpha}) = \|\boldsymbol{\alpha}\|^2 \\
\Longrightarrow &\|\varphi(\boldsymbol{\alpha})\| = \|\boldsymbol{\alpha}\|, \forall \boldsymbol{\alpha}.
\end{aligned}$$

□

命题 9.4 设 $\varphi: V \longrightarrow U$ 是实 (复) 内积空间上的线性映射, 若 φ 保持范数, 则 φ 保持内积.

证明:

$$\|\boldsymbol{\alpha} + \boldsymbol{\beta}\|^2 = \|\boldsymbol{\alpha}\|^2 + \|\boldsymbol{\beta}\|^2 + (\boldsymbol{\alpha}, \boldsymbol{\beta}) + (\boldsymbol{\beta}, \boldsymbol{\alpha}) \tag{9.2}$$

$$\|\boldsymbol{\alpha} - \boldsymbol{\beta}\|^2 = \|\boldsymbol{\alpha}\|^2 + \|\boldsymbol{\beta}\|^2 - (\boldsymbol{\alpha}, \boldsymbol{\beta}) - (\boldsymbol{\beta}, \boldsymbol{\alpha}) \tag{9.3}$$

若 V 为欧式空间, 则 (9.2) 式-(9.3) 式, 得

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{4}(\|\boldsymbol{\alpha} + \boldsymbol{\beta}\|^2 - \|\boldsymbol{\alpha} - \boldsymbol{\beta}\|^2).$$

则

$$\begin{aligned}
&(\varphi(\boldsymbol{\alpha}), \varphi(\boldsymbol{\beta})) = \frac{1}{4}(\|\varphi(\boldsymbol{\alpha}) + \varphi(\boldsymbol{\beta})\|^2 - \|\varphi(\boldsymbol{\alpha}) - \varphi(\boldsymbol{\beta})\|^2) \\
=&\frac{1}{4}(\|\varphi(\boldsymbol{\alpha} + \boldsymbol{\beta})\|^2 - \|\varphi(\boldsymbol{\alpha} - \boldsymbol{\beta})\|^2) \\
=&\frac{1}{4}(\|\boldsymbol{\alpha} + \boldsymbol{\beta}\|^2 - \|\boldsymbol{\alpha} - \boldsymbol{\beta}\|^2) \\
=&(\boldsymbol{\alpha}, \boldsymbol{\beta}).
\end{aligned}$$

若 V 为酉空间, 则 (9.2) 式 $-$(9.3) 式, 得

$$\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2-\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|^2=2((\boldsymbol{\alpha},\boldsymbol{\beta})+2(\boldsymbol{\beta},\boldsymbol{\alpha})) \tag{9.4}$$

$$\|\boldsymbol{\alpha}+\boldsymbol{i\beta}\|^2-\|\boldsymbol{\alpha}-\boldsymbol{i\beta}\|^2=-2((\boldsymbol{\alpha},\boldsymbol{\beta})+2i(\boldsymbol{\beta},\boldsymbol{\alpha})) \tag{9.5}$$

将 (9.4) 式 +(9.5) 式 $\times i$, 得

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\frac{1}{4}(\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2-\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|^2+i\|\boldsymbol{\alpha}+i\boldsymbol{\beta}\|^2-i\|\boldsymbol{\alpha}-i\boldsymbol{\beta}\|^2).$$

则

$$\begin{aligned}
&(\varphi(\boldsymbol{\alpha}),\varphi(\boldsymbol{\beta}))\\
=&\frac{1}{4}(\|\varphi(\boldsymbol{\alpha})+\varphi(\boldsymbol{\beta})\|^2-\|\varphi(\boldsymbol{\alpha})-\varphi(\boldsymbol{\beta})\|^2+i\|\varphi(\boldsymbol{\alpha})+i\varphi(\boldsymbol{\beta})\|^2-i\|\varphi(\boldsymbol{\alpha})-i\varphi(\boldsymbol{\beta})\|^2)\\
=&\frac{1}{4}(\|\varphi(\boldsymbol{\alpha}+\boldsymbol{\beta})\|^2-\|\varphi(\boldsymbol{\alpha}-\boldsymbol{\beta})\|^2+i\|\varphi(\boldsymbol{\alpha}+i\boldsymbol{\beta})\|^2-i\|\varphi(\boldsymbol{\alpha}-i\boldsymbol{\beta})\|^2)\\
=&\frac{1}{4}(\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2-\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|^2+i\|\boldsymbol{\alpha}+i\boldsymbol{\beta}\|^2-i\|\boldsymbol{\alpha}-i\boldsymbol{\beta}\|^2)\\
=&(\boldsymbol{\alpha},\boldsymbol{\beta}).
\end{aligned}$$

□

注 由以上两个命题可知, 保积映射和保范 (保距) 映射等价.

定理 9.10 $\varphi: V^n \longrightarrow U^n$ 是 n 维实 (复) 内积空间上的线性映射, 则下列命题等价:

(1) φ 保持内积;

(2) φ 是保积同构;

(3) φ 将 V 的任一组标准正交基映为 U 的一组标准正交基;

(4) φ 将 V 的某一组标准正交基映为 U 的一组标准正交基

证明: (1) $\Longrightarrow$ (2):

$$\varphi\text{保持内积} \Longrightarrow \ker\varphi=0.$$

由维数公式可知:

$$\dim\operatorname{Im}\varphi=\dim V=\dim U$$

从而 $\operatorname{Im}\varphi=U$, 即 φ 是满射, 于是 $\varphi: V\longrightarrow U$ 是线性同构.

(2) $\Longrightarrow$ (3): 任取 V 的标准正交基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$, 则

$$(\varphi(\boldsymbol{e}_i),\varphi(\boldsymbol{e}_j))=(\boldsymbol{e}_i,\boldsymbol{e}_j)=\delta_{ij},\forall i,j$$

即 $\varphi(\boldsymbol{e}_1),\varphi(\boldsymbol{e}_2),\cdots,\varphi(\boldsymbol{e}_n)$ 是 U 的一组标准正交基.

(3) $\Longrightarrow$ (4): 显然.

(4) $\Longrightarrow$ (1): 设 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是 V 的标准正交基且 $\{\varphi(\boldsymbol{e}_1),\varphi(\boldsymbol{e}_2),\cdots,\varphi(\boldsymbol{e}_n)\}$ 是 U 的标准正交基, 任取

$$\boldsymbol{\alpha}=\sum a_i\boldsymbol{e}_i\longleftrightarrow\boldsymbol{x}=(a_1,a_2,\cdots,a_n)'$$
$$\boldsymbol{\beta}=\sum b_i\boldsymbol{e}_i\longleftrightarrow\boldsymbol{y}=(b_1,b_2,\cdots,b_n)'$$

则

$$\varphi(\boldsymbol{\alpha})=\sum a_i\varphi(\boldsymbol{e}_i),\varphi(\boldsymbol{\beta})=\sum b_i\varphi(\boldsymbol{e}_i).$$

$$(\varphi(\boldsymbol{\alpha}),\varphi(\boldsymbol{\beta}))=(\sum a_i\varphi(\boldsymbol{e}_i),\sum b_i\varphi(\boldsymbol{e}_i))=\sum a_i\overline{b_i}\tag{9.6}$$
$$(\boldsymbol{\alpha},\boldsymbol{\beta})=(\sum a_i\boldsymbol{e}_i,\sum b_i\boldsymbol{e}_i)=\sum a_i\overline{b_i}\tag{9.7}$$

由 (9.6) 式 = (9.7) 式可知, φ 保持内积. □

推论 9.2 设 V,U 为实 (复) 内积空间, 则存在保积同构

$$\varphi:V\longrightarrow U\Longleftrightarrow\dim V=\dim U.$$

证明: 必要性显然. 下面证明充分性: 设 $\dim V=\dim U=n$, 取 V 的标准正交基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$, 取 U 的标准正交基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$, 定义一个线性映射:

$$\begin{aligned}\varphi:V&\longrightarrow U\\ \boldsymbol{e}_i&\longmapsto\boldsymbol{f}_i\end{aligned}$$

扩张为 $\varphi\in\mathscr{L}(V,U)$. 由定理9.10(4) 可知, φ 是保积同构. □

定义 9.12 设 φ 是内积空间 V 上的线性算子, 若 φ 保持内积, 则

$$\text{称 }\varphi\text{ 为}\begin{cases}\text{正交变换} & \text{欧氏空间 (实内积空间)},\\ \text{酉变换} & \text{酉空间 (复内积空间)}.\end{cases}$$

由定理9.9可知, φ 为可逆算子.

定理 9.11 设 φ 为有限维空间 V 上的线性算子, 则

$$\varphi\text{为正交算子 (酉算子)}\Longleftrightarrow\varphi\text{可逆且}\varphi^*=\varphi^{-1}.$$

证明: 必要性 ($\Rightarrow$)

$$(\varphi(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\varphi(\boldsymbol{\alpha}),\varphi(\varphi^{-1}(\boldsymbol{\beta})))=(\boldsymbol{\alpha},\varphi^{-1}(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\varphi^*(\boldsymbol{\beta})).$$

充分性 ($\Leftarrow$)

$$(\varphi(\boldsymbol{\alpha}),\varphi(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\varphi^*(\varphi\boldsymbol{\beta}))=(\boldsymbol{\alpha},\varphi^{-1}(\varphi(\boldsymbol{\beta})))=(\boldsymbol{\alpha},\boldsymbol{\beta}).$$

□

定义 9.13 设 $A \in M_n(\mathbb{R})$, 若 $A' = A^{-1}$, 则称 A 为正交矩阵; 设 $C \in M_n(\mathbb{C})$, 若 $\overline{C}' = C^{-1}$, 则称 C 为酉矩阵.

定理 9.12 设 φ 为 n 维内积空间 V 上的线性算子, 则

$$\varphi\text{为正交算子 (酉算子)} \Longleftrightarrow \varphi\text{在任 (某) 一组标准正交基下的表示矩阵为正交矩阵 (酉矩阵)}.$$

证明: 任取 V 的标准正交基 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$, 设 φ 的表示矩阵为 A, 则 φ^* 在同一组基下的表示矩阵为 $\begin{cases} A' & V\text{为欧氏空间}, \\ \overline{A}' & V\text{为酉空间}. \end{cases}$

φ 为正交算子 (酉算子), 由定理9.11可知, φ 可逆且 $\varphi^* = \varphi^{-1}$. 由线性变换及其表示矩阵之间的一一对应性可知,φ^* 与 φ^{-1} 在同一组标准正交基下有相同的表示矩阵, 即

$$\begin{cases} A' = A^{-1}(\text{即 } A \text{ 为正交矩阵}) & V\text{为欧氏空间时}, \\ \overline{A}' = A^{-1}(\text{即 } A \text{ 为酉矩阵}) & V\text{为酉空间时}. \end{cases}$$

□

定理 9.13 设 $A \in M_n(\mathbb{R})$, 则下列命题等价:

(1) A 为正交矩阵;

(2) A 的 n 个行分块向量是 $\mathbb{R}_n$(标准内积) 的一组标准正交基;

(3) A 的 n 个列分块向量是 $\mathbb{R}^n$(标准内积) 的一组标准正交基.

证明: 这里只简要证明 (2)-(3).

(2) 设 $A = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \cdots \\ \boldsymbol{\alpha}_n \end{pmatrix}$, 其中, $\boldsymbol{\alpha}_i$ 是 A 的行分块向量. 依题意,

$$I_n = AA' = A = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \cdots \\ \boldsymbol{\alpha}_n \end{pmatrix} (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n),$$

$$a_{ij} = \boldsymbol{\alpha}_i \cdot \boldsymbol{\alpha}_j' = (\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j)_{\mathbb{R}_n} = \delta_{ij}$$

上式中,a_{ij} 即 A 的第 (i, j) 元素.

(3) 设 $A = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)$, 其中, $\boldsymbol{\beta}_i$ 是 A 的列分块向量. 依题意,

$$I_n = A'A = \begin{pmatrix} \boldsymbol{\beta}_1' \\ \boldsymbol{\beta}_2' \\ \cdots \\ \boldsymbol{\beta}_n' \end{pmatrix},$$

$$a_{ij} = \boldsymbol{\beta}_i' \cdot \boldsymbol{\beta}_j = (\boldsymbol{\beta}_i, \boldsymbol{\beta}_j)_{\mathbb{R}^n} = \delta_{ij}.$$

□

定理 9.14 设 $A \in M_n(\mathbb{C})$, 则下列命题等价:

(1) A 为酉矩阵;

(2) A 的 n 个行分块向量是 $\mathbb{C}_n$(标准内积) 的一组标准正交基;

(3) A 的 n 个列分块向量是 $\mathbb{C}^n$(标准内积) 的一组标准正交基.

证明: 证明和上个定理完全一致, 下面是简要证明思路:

$$A\text{为酉矩阵} \iff A\overline{A}' = I_n \iff \overline{A}'A = I_n.$$

□

例 9.17 (1) 正交矩阵是特殊的酉矩阵;

(2) I_n 是正交矩阵或酉矩阵;

(3) $D = \{d_1, d_2, \cdots, d_n\} \in \mathbb{R}$ 是正交阵 $\iff d_i = \pm 1, 1 \leq i \leq n$;

(4) 二阶正交阵分类:

(I) 旋转 θ 角度

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, |A| = 1;$$

(II) 反射

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}, |A| = -1.$$

命题 9.5 (1) 酉阵的行列式的模长为 1, 正交阵的行列式值为 ± 1;

(2) 酉阵的特征值模长为 1.

证明: (1)

$$\begin{aligned} A\overline{A}' = I_n \Longrightarrow 1 = |I_n| = |A\overline{A}'| = |\det(A)|^2 \\ \Longrightarrow |\det(A)| = 1 \end{aligned}$$

(2) 设

$$A\boldsymbol{\alpha} = \lambda_0 \boldsymbol{\alpha}, 0 \neq \boldsymbol{\alpha} \in \mathbb{C}^n, \lambda_n \in \mathbb{C}, \tag{9.8}$$

则

$$\overline{\boldsymbol{\alpha}}'\overline{A}' = \overline{\lambda}_0 \overline{\boldsymbol{\alpha}}' \tag{9.9}$$

将(9.9)式左右两边分别左乘到(9.8)式左右两边, 得:

$$\overline{\boldsymbol{\alpha}}'\boldsymbol{\alpha} = |\lambda_0|^2\overline{\boldsymbol{\alpha}}'\boldsymbol{\alpha},$$

从而 $|\lambda_0| = 1$. □

定理 9.15 (QR 分解定理) 设 A 为 n 阶实 (复) 方阵, 则

$$A = QR$$

其中 Q 为正交阵 (酉阵),R 为上三角阵且主对角元素全大于等于 0. 进一步, 若 A 非异, 则上述分解必唯一.

证明: 设 $A = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n)$, 其中 $\boldsymbol{u}_i$ 是 A 的列分块向量. 作一个推广的 Gram-Schmidt 正交化过程 (因为不能确定 $\boldsymbol{u}_i$ 是否线性无关): 目标是最终得到一个列向量组 $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n\}$ 是正交向量组, 其中 $\boldsymbol{w}_i$ 是 $\mathbf{0}$ 或者是单位向量.

归纳法: 设 $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_{k-1}\}$ 已经构建好，现求 $\boldsymbol{w}_k$. 令

$$\boldsymbol{v}_k = \boldsymbol{u}_k - \sum_{j=1}^{k-1}(\boldsymbol{u}_k, \boldsymbol{w}_j)\boldsymbol{w}_j,$$

$$\left.\begin{cases} \text{若}\boldsymbol{v}_k = \mathbf{0}, \text{则令}\boldsymbol{w}_k = \mathbf{0}; \\ \text{若}\boldsymbol{v}_k \neq \mathbf{0}, \text{则令}\boldsymbol{w}_k = \dfrac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}; \end{cases}\right\} \text{得到正交向量组}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_k\}.$$

因此,

$$\boldsymbol{u}_k = \sum_{j=1}^{k}(\boldsymbol{u}_k, \boldsymbol{w}_j)\boldsymbol{w}_j + \|\boldsymbol{u}_k\|\boldsymbol{w}_k, (1 \leq k \leq n).$$

$$A = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n) = (\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n)\begin{pmatrix} \|\boldsymbol{v}_1\| & * & * & * \\ & \|\boldsymbol{v}_2\| & * & * \\ & & \ddots & * \\ & & & \|\boldsymbol{v}_n\| \end{pmatrix}.$$

令上式中的矩阵为 R, 则若某个 $\boldsymbol{w}_i = \mathbf{0}$, 则 R 的第 i 行全为 0. 设 $\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n$ 中有 r 个非零向量 $\boldsymbol{w}_{i_1}, \boldsymbol{w}_{i_2}, \cdots, \boldsymbol{w}_{i_r}$, 将其扩张为全空间 $\mathbb{R}^n$ 的一组标准正交基 $\{\widetilde{\boldsymbol{w}_1}, \widetilde{\boldsymbol{w}_2}, \cdots, \widetilde{\boldsymbol{w}_n}\}$, 其中 $\widetilde{\boldsymbol{w}_j} = \boldsymbol{w}_j, j = i_1, i_2, \cdots, i_r$, 令 $(\widetilde{\boldsymbol{w}_1}, \widetilde{\boldsymbol{w}_2}, \cdots, \widetilde{\boldsymbol{w}_n}) = Q$, 则 Q 为正交阵, 且

$$A = QR.$$

□

9.5 自伴随算子