代数学基础

作者: 邹文杰

组织: 无

时间:2024/10/25

版本:ElegantBook-4.5

自定义: 信息

宠辱不惊, 闲看庭前花开花落;
去留无意, 漫随天外云卷云舒.

目录

第一章 行列式

1.1 行列式基本性质

命题 1.1 (行列式计算常识)

(1)
$$\begin{vmatrix} & & & a_n \\ & & \iddots & \\ & a_2 & & \\ a_1 & & & \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} a_1a_2\cdots a_n\,;\quad \begin{vmatrix} a_1 & & & & \\ & \ddots & & & \\ b_1 & \cdots & a_i & \cdots & b_n \\ & & & \ddots & \\ & & & & a_n \end{vmatrix} = a_1a_2\cdots a_n.$$

(2) 设 n 阶行列式 $D=\det(a_{ij})$, 把 D 上下翻转 (**行倒排**)、或左右翻转 (**列倒排**) 分别得到 D_1、D_2; 把 D **逆时针旋转** $90°$、或**顺时针旋转** $90°$ 分别得到 D_3、D_4; 把 D **依副对角线翻转**、或**依主对角线翻转**分别得到 D_5、D_6. 易知

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

$$D_4=\begin{vmatrix} a_{n1} & \cdots & a_{11} \\ \vdots & & \vdots \\ a_{nn} & \cdots & a_{1n} \end{vmatrix}, D_5=\begin{vmatrix} a_{nn} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{11} \end{vmatrix}, D_6=\begin{vmatrix} a_{nn} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{11} \end{vmatrix}.$$

则一定有

$$D_1=D_2=D_3=D_4=(-1)^{\frac{n(n-1)}{2}}D,$$
$$D_5=D_6=D.$$

(3) 设 $A=(a_{i,j})$ 为 n 阶复矩阵, 则一定有 $|\overline{A}|=\overline{|A|}$.

(4) 若 $|A|$ 是 n 阶行列式,$|B|$ 是 m 阶行列式, 它们的值都不为零, 则

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{vmatrix} = (-1)^{mn}\begin{vmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{O} \end{vmatrix}.$$

注 实际上, 令 $e_i,i=1,2,\cdots,n$ 分别为第 i 行元素为 1, 其余元素为零的列向量. 则由基本矩阵乘法, 不难发现

$$D_1=(e_n,e_{n-1},\cdots,e_1)\,D,\quad D_2=D\,(e_n,e_{n-1},\cdots,e_1)\,.$$

证明 (1) 运用行列式的定义即可得到结论.

(2)
$$D_1=\begin{vmatrix} a_{n1} & \cdots & a_{nn} \\ \vdots & & \vdots \\ a_{11} & \cdots & a_{1n} \end{vmatrix} \xlongequal[i=1,2,\cdots,n-1]{r_i\longleftrightarrow r_{i+1}} (-1)^{n-1}\begin{vmatrix} a_{n-1,1} & \cdots & a_{n-1,n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \xlongequal[i=1,2,\cdots,n-2]{r_i\longleftrightarrow r_{i+1}} (-1)^{n-1+n-2}\begin{vmatrix} a_{n-2,1} & \cdots & a_{n-2,n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

$$=\cdots=(-1)^{n-1+n-2+\cdots+1}\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}=(-1)^{\frac{n(n-1)}{2}}\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}=(-1)^{\frac{n(n-1)}{2}}D.$$

$$D_2 = \begin{vmatrix} a_{1n} & \cdots & a_{11} \\ \vdots & & \vdots \\ a_{nn} & \cdots & a_{n1} \end{vmatrix} \xlongequal[i=1,2,\cdots,n-1]{j_i \longleftrightarrow j_{i+1}} (-1)^{n-1} \begin{vmatrix} a_{1,n-1} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n,n-1} & \cdots & a_{nn} \end{vmatrix} \xlongequal[i=1,2,\cdots,n-2]{j_i \longleftrightarrow j_{i+1}} (-1)^{n-1+n-2} \begin{vmatrix} a_{1,n-2} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n,n-2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \cdots = (-1)^{n-1+n-2+\cdots+1} \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} D.$$

$$D_3 = \begin{vmatrix} a_{1n} & \cdots & a_{nn} \\ \vdots & & \vdots \\ a_{11} & \cdots & a_{n1} \end{vmatrix} \xlongequal{\text{行倒排}} (-1)^{\frac{n(n-1)}{2}} \begin{vmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} D^T = (-1)^{\frac{n(n-1)}{2}} D.$$

$$D_4 = \begin{vmatrix} a_{n1} & \cdots & a_{11} \\ \vdots & & \vdots \\ a_{nn} & \cdots & a_{1n} \end{vmatrix} \xlongequal{\text{列倒排}} (-1)^{\frac{n(n-1)}{2}} \begin{vmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} D^T = (-1)^{\frac{n(n-1)}{2}} D.$$

$$D_5 = \begin{vmatrix} a_{nn} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{11} \end{vmatrix} \xlongequal{\text{逆时针旋转}90^\circ} (-1)^{\frac{n(n-1)}{2}} \begin{vmatrix} a_{1n} & \cdots & a_{11} \\ \vdots & & \vdots \\ a_{nn} & \cdots & a_{n1} \end{vmatrix} \xlongequal{\text{列倒排}} (-1)^{\frac{n(n-1)}{2}} \cdot (-1)^{\frac{n(n-1)}{2}} \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = D.$$

$$D_6 = \begin{vmatrix} a_{nn} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{11} \end{vmatrix} \xlongequal{\text{顺时针旋转}90^\circ} (-1)^{\frac{n(n-1)}{2}} \begin{vmatrix} a_{1n} & \cdots & a_{nn} \\ \vdots & & \vdots \\ a_{11} & \cdots & a_{n1} \end{vmatrix} \xlongequal{\text{行倒排}} (-1)^{\frac{n(n-1)}{2}} \cdot (-1)^{\frac{n(n-1)}{2}} \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = D.$$

(3) 复数的共轭保持加法和乘法:$\overline{z_1+z_2} = \overline{z_1} + \overline{z_2}, \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$, 故由行列式的组合定义可得

$$\begin{aligned} |A| &= \sum_{1 \le k_1, k_2, \cdots, k_n \le n} (-1)^{\tau(k_1 k_2 \cdots k_n)} a_{k_1 1} a_{k_2 2} \cdots a_{k_n n} \\ &= \sum_{1 \le k_1, k_2, \cdots, k_n \le n} (-1)^{\tau(k_1 k_2 \cdots k_n)} \overline{a_{k_1 1}} \cdot \overline{a_{k_2 2}} \cdots \overline{a_{k_n n}} = |\overline{A}|. \end{aligned}$$

(4) 将 $|A|$ 的第一列依次和 $|B|$ 的第 m 列, 第 $m-1$ 列,⋯, 第一列对换, 共换了 m 次; 再将 $|A|$ 的第二列依次和 $|B|$ 的第 m 列, 第 $m-1$ 列,⋯, 第一列对换, 又换了 m 次;⋯. 依次类推, 经过 mn 次对换可将第二个行列式变为第一个行列式. 因此 $|D| = (-1)^{mn}|C|$, 于是由行列式的基本性质可得

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{vmatrix} = (-1)^{mn} \begin{vmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{O} \end{vmatrix}.$$

□

命题 1.2 (行列式的刻画)

设 f 为从 n 阶方阵全体构成的集合到数集上的映射, 使得对任意的 n 阶方阵 $\boldsymbol{A}$, 任意的指标 $1 \le i \le n$, 以及任意的常数 c, 满足下列条件:

(1) 设 $\boldsymbol{A}$ 的第 i 列是方阵 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 的第 i 列之和, 且 $\boldsymbol{A}$ 的其余列与 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 的对应列完全相同, 则 $f(\boldsymbol{A}) = f(\boldsymbol{B}) + f(\boldsymbol{C})$;

(2) 将 $\boldsymbol{A}$ 的第 i 列乘以常数 c 得到方阵 $\boldsymbol{B}$, 则 $f(\boldsymbol{B}) = cf(\boldsymbol{A})$;

(3) 对换 $\boldsymbol{A}$ 的任意两列得到方阵 $\boldsymbol{B}$, 则 $f(\boldsymbol{B}) = -f(\boldsymbol{A})$;

(4) $f(\boldsymbol{I}_n) = 1$, 其中 $\boldsymbol{I}_n$ 是 n 阶单位阵.

求证:$f(\boldsymbol{A}) = |\boldsymbol{A}|$.

笔记 这个命题给出了**行列式的刻画**: 在方阵 n 个列向量上的多重线性和反对称性, 以及正规性 (即单位矩阵处的取值为 1), 唯一确定了行列式这个函数.

证明 设 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$, 其中 $\boldsymbol{\alpha}_k$ 为 $\boldsymbol{A}$ 的第 k 列,$\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 为标准单位列向量, 则

$$\boldsymbol{\alpha}_j=a_{1j}\boldsymbol{e}_1+a_{2j}\boldsymbol{e}_2+\cdots+a_{nj}\boldsymbol{e}_n=\sum_{k=1}^{n}a_{kj}\boldsymbol{e}_k,\ j=1,2,\cdots,n.$$

从而由条件 (1) 和 (2) 可得

$$\begin{aligned}
f(\boldsymbol{A})&=f(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=f\left(\sum_{k_1=1}^{n}a_{k_1 1}\boldsymbol{e}_k,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\right)\\
&=a_{11}f(\boldsymbol{e}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)+a_{21}f(\boldsymbol{e}_2,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)+\cdots+a_{n1}f(\boldsymbol{e}_n,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\\
&=\sum_{k_1=1}^{n}a_{k_1 1}f\left(\boldsymbol{e}_{k_1},\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\right)=\sum_{k_1=1}^{n}a_{k_1 1}f\left(\boldsymbol{e}_{k_1},\sum_{k_2=1}^{n}a_{k_2 2}\boldsymbol{e}_{k_2},\cdots,\boldsymbol{\alpha}_n\right)\\
&=\sum_{k_1=1}^{n}a_{k_1 1}\left[a_{12}f\left(\boldsymbol{e}_{k_1},\boldsymbol{e}_1,\cdots,\boldsymbol{\alpha}_n\right)+a_{22}f\left(\boldsymbol{e}_{k_1},\boldsymbol{e}_2,\cdots,\boldsymbol{\alpha}_n\right)+\cdots+a_{n2}f\left(\boldsymbol{e}_{k_1},\boldsymbol{e}_n,\cdots,\boldsymbol{\alpha}_n\right)\right]\\
&=\sum_{k_1=1}^{n}a_{k_1 1}\sum_{k_2=1}^{n}a_{k_2 2}f\left(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{\alpha}_n\right)=\cdots=\sum_{k_1=1}^{n}a_{k1}\sum_{k_2=1}^{n}a_{k_2 2}\cdots\sum_{k_n=1}^{n}a_{k_n n}f\left(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n}\right)\\
&=\sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n}a_{k1}a_{k_2 2}\cdots a_{k_n n}f\left(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n}\right)=\sum_{(k_1,k_2,\cdots,k_n)}a_{k_1 1}a_{k_2 2}\cdots a_{k_n n}f\left(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n}\right).
\end{aligned}$$

若 $k_i=k_j$, 则 $(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n})$ 的第 i 列和第 j 列对换后仍然是 $(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n})$. 由条件 (3) 可知,$f(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n})=-f(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n})$, 于是 $f(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n})=0$. 因此在 $f(\boldsymbol{A})$ 的表示式中, 只剩下 $k_i(i=1,2,\cdots,n)$ 互不相同的项. 通过 $\tau(k_1k_2\cdots k_n)$ 次相邻对换可将 $(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n})$ 变成 $(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)=\boldsymbol{I}_n$, 故由条件 (3) 和 (4) 可得

$$f(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n})=(-1)^{\tau(k_1k_2\cdots k_n)}f(\boldsymbol{I}_n)=(-1)^{\tau(k_1k_2\cdots k_n)}.$$

于是由行列式的组合定义可知

$$f(\boldsymbol{A})=\sum_{(k_1,k_2,\cdots,k_n)}a_{k_1 1}a_{k_2 2}\cdots a_{k_n n}f(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\cdots,\boldsymbol{e}_{k_n})=\sum_{(k_1,k_2,\cdots,k_n)}(-1)^{\tau(k_1k_2\cdots k_n)}a_{k_1 1}a_{k_2 2}\cdots a_{k_n n}=|\boldsymbol{A}|.$$

□

1.2 降阶法

定义 1.1 (组合数定义的扩充)

$$\mathrm{C}_n^k\triangleq\begin{cases}0&,k<0,\\ \dfrac{n!}{k!\,(n-k)!}&,0\leqslant k\leqslant n,\\ 0&,k>n,\end{cases}.$$

♣

定理 1.1 (Vandermode 恒等式)

$$\mathrm{C}_{p+q}^l=\sum_{k=0}^{p+q}\mathrm{C}_p^k\mathrm{C}_q^{l-k},\quad l=1,2,\cdots,p+q.$$

♡

证明 注意到

$$(x+1)^p\,(x+1)^q=(x+1)^{p+q}\,.$$

由二项式定理可得

$$\sum_{r=0}^{p}\mathrm{C}_p^r x^r\cdot\sum_{r=0}^{p}\mathrm{C}_q^r x^r=\sum_{r=0}^{p+q}\mathrm{C}_{p+q}^r x^r.$$

对 $\forall l=\{1,2,\cdots,p+q\}$, 考虑上式 x^l 的系数, 就有

$$\mathrm{C}_{p+q}^{l}=\mathrm{C}_p^0\mathrm{C}_q^l+\mathrm{C}_p^1\mathrm{C}_q^{l-1}+\cdots+\mathrm{C}_p^l\mathrm{C}_q^0.$$

□

例题 1.1 计算 n 阶行列式:

$$|A|=\begin{vmatrix}1&1&\cdots&1\\1&\mathrm{C}_2^1&\cdots&\mathrm{C}_n^1\\1&\mathrm{C}_3^2&\cdots&\mathrm{C}_{n+1}^2\\\vdots&\vdots&&\vdots\\1&\mathrm{C}_n^{n-1}&\cdots&\mathrm{C}_{2n-2}^{n-1}\end{vmatrix}.$$

笔记 解法一的关键就是组合数公式: $\mathrm{C}_m^{k-1}+\mathrm{C}_m^k=\mathrm{C}_{m+1}^k$.

于是有

$$\mathrm{C}_m^k=\mathrm{C}_{m+1}^k-\mathrm{C}_m^{k-1}$$
$$\mathrm{C}_m^{k-1}=\mathrm{C}_{m+1}^k-\mathrm{C}_m^k$$

解法二的核心想法就是: 将Vandermode 恒等式与矩阵乘法的定义联系起来.

解 解法一:

$$|A|=\begin{vmatrix}1&1&\cdots&1\\1&\mathrm{C}_2^1&\cdots&\mathrm{C}_n^1\\1&\mathrm{C}_3^2&\cdots&\mathrm{C}_{n+1}^2\\\vdots&\vdots&&\vdots\\1&\mathrm{C}_n^{n-1}&\cdots&\mathrm{C}_{2n-2}^{n-1}\end{vmatrix}\xlongequal[i=n,\cdots,2]{(-1)\cdot r_{i-1}+r_i}\begin{vmatrix}\mathrm{C}_0^0&\mathrm{C}_1^0&\cdots&\mathrm{C}_{\mathrm{n}-1}^0\\0&\mathrm{C}_2^1-\mathrm{C}_1^0&\cdots&\mathrm{C}_n^1-\mathrm{C}_{\mathrm{n}-1}^0\\0&\mathrm{C}_3^2-\mathrm{C}_2^1&\cdots&\mathrm{C}_{n+1}^2-\mathrm{C}_n^1\\\vdots&\vdots&&\vdots\\0&\mathrm{C}_n^{n-1}-\mathrm{C}_{\mathrm{n}-1}^{\mathrm{n}-2}&\cdots&\mathrm{C}_{2n-2}^{n-1}-\mathrm{C}_{2\mathrm{n}-3}^{\mathrm{n}-2}\end{vmatrix}$$

$$=\begin{vmatrix}\mathrm{C}_0^0&\mathrm{C}_1^0&\cdots&\mathrm{C}_{\mathrm{n}-1}^0\\0&\mathrm{C}_1^1&\cdots&\mathrm{C}_{\mathrm{n}-1}^1\\0&\mathrm{C}_2^2&\cdots&\mathrm{C}_n^2\\\vdots&\vdots&&\vdots\\0&\mathrm{C}_{\mathrm{n}-1}^{\mathrm{n}-1}&\cdots&\mathrm{C}_{2\mathrm{n}-3}^{\mathrm{n}-1}\end{vmatrix}\xlongequal{\text{按第一列展开}}\begin{vmatrix}\mathrm{C}_1^1&\mathrm{C}_2^1&\cdots&\mathrm{C}_{\mathrm{n}-1}^1\\\mathrm{C}_2^2&\mathrm{C}_3^2&\cdots&\mathrm{C}_n^2\\\vdots&\vdots&&\vdots\\\mathrm{C}_{\mathrm{n}-1}^{\mathrm{n}-1}&\mathrm{C}_{\mathrm{n}}^{\mathrm{n}-1}&\cdots&\mathrm{C}_{2\mathrm{n}-3}^{\mathrm{n}-1}\end{vmatrix}$$

$$\xlongequal[i=n,\cdots,2]{(-1)\cdot j_{i-1}+j_i}\begin{vmatrix}\mathrm{C}_1^1&\mathrm{C}_2^1-\mathrm{C}_1^1&\cdots&\mathrm{C}_{\mathrm{n}-1}^1-\mathrm{C}_{\mathrm{n}-2}^1\\\mathrm{C}_2^2&\mathrm{C}_3^2-\mathrm{C}_2^2&\cdots&\mathrm{C}_n^2-\mathrm{C}_{\mathrm{n}-1}^2\\\vdots&\vdots&&\vdots\\\mathrm{C}_{\mathrm{n}-1}^{\mathrm{n}-1}&\mathrm{C}_{\mathrm{n}}^{\mathrm{n}-1}-\mathrm{C}_{\mathrm{n}-1}^{\mathrm{n}-1}&\cdots&\mathrm{C}_{2\mathrm{n}-3}^{\mathrm{n}-1}-\mathrm{C}_{2\mathrm{n}-4}^{\mathrm{n}-1}\end{vmatrix}=\begin{vmatrix}\mathrm{C}_1^1&\mathrm{C}_1^0&\cdots&\mathrm{C}_{\mathrm{n}-2}^0\\\mathrm{C}_2^2&\mathrm{C}_2^1&\cdots&\mathrm{C}_{\mathrm{n}-1}^1\\\vdots&\vdots&&\vdots\\\mathrm{C}_{\mathrm{n}-1}^{\mathrm{n}-1}&\mathrm{C}_{\mathrm{n}-1}^{\mathrm{n}-2}&\cdots&\mathrm{C}_{2\mathrm{n}-4}^{\mathrm{n}-2}\end{vmatrix}$$

$$=\begin{vmatrix}1&1&\cdots&1\\1&\mathrm{C}_2^1&\cdots&\mathrm{C}_{\mathrm{n}-1}^1\\\vdots&\vdots&&\vdots\\1&\mathrm{C}_{\mathrm{n}-1}^{\mathrm{n}-2}&\cdots&\mathrm{C}_{2\mathrm{n}-4}^{\mathrm{n}-2}\end{vmatrix}.$$

此时得到的行列式恰好是原行列式的左上角部分, 并具有相同的规律. 不断这样做下去, 最后可得 $|A|=1$

解法二:设 $A=(a_{ij})_{n\times n}$, 则 $a_{ij}=\mathrm{C}_{i+j-2}^{i-1}, i,j=1,2,\cdots,n$. 从而由Vandermode 恒等式及组合数定义的扩充可得

$$a_{ij}=\mathrm{C}_{i+j-2}^{i-1}=\sum_{k=0}^{i+j-2}\mathrm{C}_{i-1}^{i-1-k}\mathrm{C}_{j-1}^{k}=\sum_{k=1}^{i+j-1}\mathrm{C}_{i-1}^{i-k}\mathrm{C}_{j-1}^{k-1}$$

$$=\sum_{k=1}^{n}\mathrm{C}_{i-1}^{i-k}\mathrm{C}_{j-1}^{k-1}=\sum_{k=1}^{n}\mathrm{C}_{i-1}^{k-1}\mathrm{C}_{j-1}^{k-1}=\sum_{k=1}^{n}l_{ik}l_{jk},$$

其中 $l_{ij}=\begin{cases}\mathrm{C}_{i-1}^{j-1}, & 1\leqslant j\leqslant i\leqslant n,\\ 0, & 1\leqslant i<j\leqslant n,\end{cases}$. 记 $L=\left(l_{ij}\right)_{n\times n}$, 则根据矩阵乘法的定义可知

$$A=LL^T\Rightarrow |A|=|L|^2.$$

因为当 $1\leqslant i<j\leqslant n$ 时,$l_{ij}=0$, 所以 L 是上三角矩阵. 于是

$$|L|=\prod_{i=1}^{n}l_{ii}=\prod_{i=1}^{n}\mathrm{C}_{i-1}^{i-1}=1.$$

故 $|A|=|L|^2=1$. □

例题 1.2 计算 n 阶行列式:

$$|\boldsymbol{A}|=\begin{vmatrix}1 & 2 & 3 & \cdots & n\\ -1 & 0 & 3 & \cdots & n\\ -1 & -2 & 0 & \cdots & n\\ \vdots & \vdots & \vdots & & \vdots\\ -1 & -2 & -3 & \cdots & 0\end{vmatrix}$$

解

$$|\boldsymbol{A}|=\begin{vmatrix}1 & 2 & 3 & \cdots & n\\ -1 & 0 & 3 & \cdots & n\\ -1 & -2 & 0 & \cdots & n\\ \vdots & \vdots & \vdots & & \vdots\\ -1 & -2 & -3 & \cdots & 0\end{vmatrix}\xlongequal[i=2,\cdots,n]{r_1+r_i}\begin{vmatrix}1 & 2 & 3 & \cdots & n\\ 0 & 2 & * & \cdots & *\\ 0 & 0 & 3 & \cdots & *\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & n\end{vmatrix}=n!$$

□

例题 1.3 计算 n 阶行列式:

$$|\boldsymbol{A}|=\begin{vmatrix}a_1b_1 & a_1b_2 & a_1b_3 & \cdots & a_1b_n\\ a_1b_2 & a_2b_2 & a_2b_3 & \cdots & a_2b_n\\ a_1b_3 & a_2b_3 & a_3b_3 & \cdots & a_3b_n\\ \vdots & \vdots & \vdots & & \vdots\\ a_1b_n & a_2b_n & a_3b_n & \cdots & a_nb_n\end{vmatrix}.$$

解

$$|\boldsymbol{A}|=\begin{vmatrix}a_1b_1 & a_1b_2 & a_1b_3 & \cdots & a_1b_n\\ a_1b_2 & a_2b_2 & a_2b_3 & \cdots & a_2b_n\\ a_1b_3 & a_2b_3 & a_3b_3 & \cdots & a_3b_n\\ \vdots & \vdots & \vdots & & \vdots\\ a_1b_n & a_2b_n & a_3b_n & \cdots & a_nb_n\end{vmatrix}=a_1\begin{vmatrix}b_1 & b_2 & b_3 & \cdots & b_n\\ a_1b_2 & a_2b_2 & a_2b_3 & \cdots & a_2b_n\\ a_1b_3 & a_2b_3 & a_3b_3 & \cdots & a_3b_n\\ \vdots & \vdots & \vdots & & \vdots\\ a_1b_n & a_2b_n & a_3b_n & \cdots & a_nb_n\end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n]{(-a_i)r_1+r_i}a_1\begin{vmatrix}b_1 & b_2 & b_3 & \cdots & b_n\\ a_1b_2-a_2b_1 & 0 & 0 & \cdots & 0\\ a_1b_3-a_3b_1 & a_2b_3-a_3b_2 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ a_1b_n-a_nb_1 & a_2b_n-a_nb_2 & a_3b_n-a_nb_3 & \cdots & 0\end{vmatrix}$$

$$\xlongequal{\text{按第}n\text{列展开}}(-1)^{n+1}a_1b_n\begin{vmatrix}a_1b_2-a_2b_1 & 0 & \cdots & 0\\ a_1b_3-a_3b_1 & a_2b_3-a_3b_2 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ a_1b_n-a_nb_1 & a_2b_n-a_nb_2 & \cdots & a_{n-1}b_n-a_nb_{n-1}\end{vmatrix}$$

$$
\begin{aligned}
&= (-1)^{n-1}\, a_1 b_n \prod_{i=1}^{n-1} (a_i b_{i+1} - a_{i+1} b_i) \\
&= a_1 b_n \prod_{i=1}^{n-1} (a_{i+1} b_i - a_i b_{i+1}).
\end{aligned}
$$

□

命题 1.3 (" 爪" 型行列式)

证明 n 阶行列式:

$$
|\boldsymbol{A}| = \begin{vmatrix} a_1 & b_2 & \cdots & b_n \\ c_2 & a_2 & & \\ \vdots & & \ddots & \\ c_n & & & a_n \end{vmatrix} = a_1 a_2 \cdots a_n - \sum_{i=2}^{n} a_2 \cdots \widehat{a_i} \cdots a_n b_i c_i.
$$

笔记 记忆" 爪" 型行列式的计算方法和结论.

证明 当 $a_i \neq 0\,(\forall i \in [2,n] \cap \mathbb{N})$ 时, 我们有

$$
|\boldsymbol{A}| = \begin{vmatrix} a_1 & b_2 & \cdots & b_n \\ c_2 & a_2 & & \\ \vdots & & \ddots & \\ c_n & & & a_n \end{vmatrix} \xlongequal[i=2,\cdots,n]{\left(-\frac{c_i}{a_i}\right) j_i + j_1} \begin{vmatrix} a_1 - \sum\limits_{i=2}^{n} \dfrac{b_i c_i}{a_i} & b_2 & \cdots & b_n \\ 0 & a_2 & & \\ \vdots & & \ddots & \\ 0 & & & a_n \end{vmatrix}
$$

$$
= \left(a_1 - \sum_{i=2}^{n} \frac{b_i c_i}{a_i} \right) \prod_{i=2}^{n} a_i = a_1 a_2 \cdots a_n - \sum_{i=2}^{n} a_2 \cdots \widehat{a_i} \cdots a_n b_i c_i.
$$

当 $\exists i \in [2,n] \cap \mathbb{N}\ s.t.\ a_i = 0$ 时, 则 $a_1 a_2 \cdots a_n - \sum\limits_{i=2}^{n} a_2 \cdots \widehat{a_i} \cdots a_n b_i c_i = -a_2 \cdots \widehat{a_i} \cdots a_n b_i c_i$. 此时, 我们有

$$
|\boldsymbol{A}| = \begin{vmatrix} a_1 & b_2 & \cdots & b_{i-1} & b_i & b_{i+1} & \cdots & b_n \\ c_2 & a_2 & & & & & & \\ \vdots & & \ddots & & & & & \\ c_{i-1} & & & a_{i-1} & & & & \\ c_i & & & & 0 & & & \\ c_{i+1} & & & & & a_{i+1} & & \\ \vdots & & & & & & \ddots & \\ c_n & & & & & & & a_n \end{vmatrix} \xlongequal[\text{(按}c_i\text{所在行展开)}]{\text{按第}i\text{行展开}} (-1)^{i+1} c_i \begin{vmatrix} b_2 & \cdots & b_{i-1} & b_i & b_{i+1} & \cdots & b_n \\ a_2 & & & & & & \\ & \ddots & & & & & \\ & & a_{i-1} & 0 & 0 & & \\ & & 0 & 0 & a_{i+1} & & \\ & & & & & \ddots & \\ & & & & & & a_n \end{vmatrix}
$$

$$
\xlongequal[\text{(按}b_i\text{所在列展开)}]{\text{按第}i-1\text{列展开}} (-1)^{i+1}(-1)^{i} b_i c_i \begin{vmatrix} a_2 & & & & \\ & \ddots & & & \\ & & a_{i-1} & & \\ & & & a_{i+1} & \\ & & & & \ddots & \\ & & & & & a_n \end{vmatrix} = -a_2 \cdots \widehat{a_i} \cdots a_n b_i c_i.
$$

综上所述, 原命题得证. □

命题 1.4 (分块" 爪" 型行列式)

计算 n 阶行列式 $(a_{ii}\neq 0, i=k+1,k+2,\cdots,n)$:

$$|\boldsymbol{A}|=\begin{vmatrix} a_{11} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} & a_{k,k+1} & \cdots & a_{kn} \\ a_{k+1,1} & \cdots & a_{k+1,k} & a_{k+1,k+1} & & \\ \vdots & & \vdots & & \ddots & \\ a_{n1} & \cdots & a_{nk} & & & a_{nn} \end{vmatrix}.$$

笔记 记忆分块" 爪" 型行列式的计算方法即可, 计算方法和" 爪" 型行列式的计算方法类似.

解

$$|\boldsymbol{A}|=\begin{vmatrix} a_{11} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} & a_{k,k+1} & \cdots & a_{kn} \\ a_{k+1,1} & \cdots & a_{k+1,k} & a_{k+1,k+1} & & \\ \vdots & & \vdots & & \ddots & \\ a_{n1} & \cdots & a_{nk} & & & a_{nn} \end{vmatrix}$$

$$\xlongequal[i=k+1,k+2,\cdots,n]{-\frac{a_{i1}}{a_{ii}}j_i+j_1,-\frac{a_{i2}}{a_{ii}}j_i+j_2,\cdots,-\frac{a_{in}}{a_{ii}}j_i+j_k}\begin{vmatrix} c_{11} & \cdots & c_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{k1} & \cdots & c_{kk} & a_{k,k+1} & \cdots & a_{kn} \\ 0 & \cdots & 0 & a_{k+1,k+1} & & \\ \vdots & & \vdots & & \ddots & \\ 0 & \cdots & 0 & & & a_{nn} \end{vmatrix}$$

$$=\begin{vmatrix} C & B \\ O & \Lambda \end{vmatrix}=|C|\cdot|\Lambda|=|C|\prod_{i=k+1}^{n}a_{ii}.$$

其中 $C=\begin{pmatrix} c_{11} & \cdots & c_{1k} \\ \vdots & & \vdots \\ c_{k1} & \cdots & c_{kk} \end{pmatrix}, B=\begin{pmatrix} a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{k,k+1} & \cdots & a_{kn} \end{pmatrix}, \Lambda=\begin{pmatrix} a_{k+1} & & \\ & \ddots & \\ & & a_n \end{pmatrix}$. 并且 $c_{pq}=a_{pq}-\sum\limits_{i=k+1}^{n}\frac{a_{iq}a_{pi}}{a_{ii}}, p,q=1,2,\cdots,n.$ □

推论 1.1 (" 爪" 型行列式的推广)

计算 n 阶行列式:

$$|\boldsymbol{A}|=\begin{vmatrix} x_1-a_1 & x_2 & x_3 & \cdots & x_n \\ x_1 & x_2-a_2 & x_3 & \cdots & x_n \\ x_1 & x_2 & x_3-a_3 & \cdots & x_n \\ \vdots & \vdots & \vdots & & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n-a_n \end{vmatrix}.$$

笔记 这是一个有用的模板 (即**行列式除了主对角元素外, 每行都一样**).

记忆该命题的计算方法即可. 即先化为" 爪" 型行列式, 再利用" 爪" 型行列式的计算结果.

解 当 $a_i\neq 0\,(\forall i\in[2,n]\cap\mathbb{N})$ 时, 我们有

$$|\boldsymbol{A}|=\begin{vmatrix} x_1-a_1 & x_2 & x_3 & \cdots & x_n \\ x_1 & x_2-a_2 & x_3 & \cdots & x_n \\ x_1 & x_2 & x_3-a_3 & \cdots & x_n \\ \vdots & \vdots & \vdots & & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n-a_n \end{vmatrix}\xlongequal[i=2,\cdots,n]{(-1)r_1+r_i}\begin{vmatrix} x_1-a_1 & x_2 & x_3 & \cdots & x_n \\ a_1 & -a_2 & 0 & \cdots & 0 \\ a_1 & 0 & -a_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_1 & 0 & 0 & \cdots & -a_n \end{vmatrix}$$

$$\xlongequal{\text{命题1.3}}\left[(x_1-a_1)+\sum_{i=2}^{n}\frac{a_1x_i}{a_i}\right]\prod_{i=2}^{n}(-a_i)=(-1)^{n-1}\left[(x_1-a_1)+\sum_{i=2}^{n}\frac{a_1x_i}{a_i}\right]\prod_{i=2}^{n}a_i$$

$$=(-1)^{n-1}\left[(x_1-a_1)\prod_{i=2}^{n}a_i+\sum_{i=2}^{n}a_1a_2\cdots\widehat{a_i}\cdots a_nx_i\right].$$

当 $\exists i\in[2,n]\cap\mathbb{N}\ s.t.\ a_i=0$ 时, 我们有

$$|\boldsymbol{A}|=\begin{vmatrix} x_1-a_1 & x_2 & x_3 & \cdots & x_n \\ x_1 & x_2-a_2 & x_3 & \cdots & x_n \\ x_1 & x_2 & x_3-a_3 & \cdots & x_n \\ \vdots & \vdots & \vdots & & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n-a_n \end{vmatrix}\xlongequal[i=2,\cdots,n]{(-1)r_1+r_i}\begin{vmatrix} x_1-a_1 & x_2 & x_3 & \cdots & x_n \\ a_1 & -a_2 & 0 & \cdots & 0 \\ a_1 & 0 & -a_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_1 & 0 & 0 & \cdots & -a_n \end{vmatrix}$$

$$\xlongequal{\text{命题1.3}}(x_1-a_1)(-a_2)(-a_3)\cdots(-a_n)-\sum_{i=2}^{n}(-a_2)\cdots\widehat{(-a_i)}\cdots(-a_n)a_1x_i$$

$$=(-1)^{n-1}(x_1-a_1)\prod_{i=2}^{n}a_i+(-1)^{n-1}\sum_{i=2}^{n}a_1a_2\cdots\widehat{a_i}\cdots a_nx_i$$

$$=(-1)^{n-1}\left[(x_1-a_1)\prod_{i=2}^{n}a_i+\sum_{i=2}^{n}a_1a_2\cdots\widehat{a_i}\cdots a_nx_i\right].$$

综上所述,$|\boldsymbol{A}|=(-1)^{n-1}\left[(x_1-a_1)\prod_{i=2}^{n}a_i+\sum_{i=2}^{n}a_1a_2\cdots\widehat{a_i}\cdots a_nx_i\right]$. □

例题 1.4 计算 n 阶行列式:

$$|\boldsymbol{A}|=\begin{vmatrix} a & 0 & \cdots & 0 & 1 \\ 0 & a & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a & 0 \\ 1 & 0 & \cdots & 0 & a \end{vmatrix}.$$

解

$$|\boldsymbol{A}|=\begin{vmatrix} a & 0 & \cdots & 0 & 1 \\ 0 & a & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a & 0 \\ 1 & 0 & \cdots & 0 & a \end{vmatrix}\xlongequal{\text{按第一列展开}}a^n+(-1)^{n+1}\begin{vmatrix} 0 & 0 & \cdots & 0 & 1 \\ a & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a & 0 \end{vmatrix}=a^n+(-1)^{n+1+n}a^{n-2}=a^n-a^{n-2}.$$

□

注 本题也可由命题1.3直接得到,$|\boldsymbol{A}|=a^n-a^{n-2}$.

命题 1.5

设 $|A|=|a_i|$ 是一个 n 阶行列式,A_{ij} 是它的第 (i,j) 元素的代数余子式,$X=(x_1,x_2,\cdots,x_n)^T,Y=(y_1,y_2,\cdots,y_n)^T$,$z$ 是任意常数, 求证:

$$\begin{vmatrix} A & X \\ Y^T & z \end{vmatrix}=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & x_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n \\ y_1 & y_2 & \cdots & y_n & z \end{vmatrix}=z|A|-\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}x_iy_j.$$

进而得到

$$\begin{vmatrix} A & X \\ Y^T & 0 \end{vmatrix}=-Y^TA^*X.$$

笔记 根据这个命题可以得到一个**关于行列式 $|A|$ 的所有代数余子式求和的构造**:

$$-\sum_{i,j=1}^{n}A_{ij}=\begin{vmatrix} \boldsymbol{A} & \boldsymbol{1} \\ \boldsymbol{1}' & 0 \end{vmatrix}=\begin{vmatrix} \boldsymbol{\alpha_1} & \boldsymbol{\alpha_2} & \cdots & \boldsymbol{\alpha_n} & \boldsymbol{1} \\ 1 & 1 & \cdots & 1 & 0 \end{vmatrix}=\begin{vmatrix} \boldsymbol{\beta_1} & 1 \\ \boldsymbol{\beta_2} & 1 \\ \vdots & \vdots \\ \boldsymbol{\beta_n} & 1 \\ \boldsymbol{1}' & 0 \end{vmatrix}.$$

其中 $|\boldsymbol{A}|$ 的列向量依次为 $\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\cdots,\boldsymbol{\alpha_n}$,$|\boldsymbol{A}|$ 的行向量依次为 $\boldsymbol{\beta_1},\boldsymbol{\beta_2},\cdots,\boldsymbol{\beta_n}$. 并且 $\boldsymbol{1}$ 表示元素均为 1 的列向量,$\boldsymbol{1}'$ 表示 $\boldsymbol{1}$ 的转置. (令上述命题中的 $z=0,x_i=y_i=1,i=1,2,\cdots,n$ 即可得到.)

注 如果需要证明的是矩阵的代数余子式的相关命题, 我们可以考虑一下这种构造, 即令上述命题中的 $z=0$ 并且待定/任取 x_i,y_i.

证明 证法一: 将上述行列式先按最后一列展开, 展开式的第一项为

$$(-1)^{n+2}x_1\begin{vmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \\ y_1 & y_2 & \cdots & y_n \end{vmatrix}.$$

再将上式按最后一行展开得到

$$\begin{aligned} &(-1)^{n+2}x_1\left[(-1)^{n+1}(-1)^{1+1}y_1A_{11}+(-1)^{n+2}(-1)^{1+2}y_2A_{12}+\cdots+(-1)^{n+n}(-1)^{1+n}y_nA_{1n}\right] \\ &=(-1)^{n+2}x_1(-1)^{n+1}\left[(-1)^2y_1A_{11}+(-1)^4y_2A_{12}+\cdots+(-1)^{2n}y_nA_{1n}\right] \\ &=-x_1(y_1A_{11}+y_2A_{12}+\cdots+y_nA_{1n}) \\ &=-x_1\sum_{j=1}^{n}y_jA_{1j}. \end{aligned}$$

同理可得原行列式展开式的第 $i(i=1,2,\cdots,n-1)$ 项为

$$(-1)^{n+1+i}x_i\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \\ y_1 & y_2 & \cdots & y_n \end{vmatrix}.$$

将上式按最后一行展开得到 $z\,|\boldsymbol{A}|$.

$$\begin{aligned}
&(-1)^{n+1+i}x_i\left[(-1)^{n+1}(-1)^{i+1}y_1A_{i1}+(-1)^{n+2}(-1)^{i+2}y_2A_{i2}+\cdots+(-1)^{n+n}(-1)^{i+n}y_nA_{in}\right]\\
&=(-1)^{n+1+i}x_i(-1)^{n+1}\left[(-1)^{i+1}y_1A_{i1}+(-1)^{i+2+1}y_2A_{i2}+\cdots+(-1)^{i+n+n-1}y_nA_{in}\right]\\
&=(-1)^{2i+1}y_1A_{i1}+(-1)^{2i+3}y_2A_{i2}+\cdots+(-1)^{2i+2n-1}y_nA_{in}\\
&=-x_i\left(y_1A_{i1}+y_2A_{i2}+\cdots+y_nA_{in}\right)\\
&=-x_i\sum_{j=1}^{n}y_jA_{ij}.
\end{aligned}$$

而展开式的最后一项为 $z\,|\boldsymbol{A}|$.

因此, 原行列式的值为

$$z|\boldsymbol{A}|-\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}x_iy_j.$$

证法二:设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)',\boldsymbol{y}=(y_1,y_2,\cdots,y_n)'$. 若 A 是非异阵, 则由降阶公式可得

$$\begin{vmatrix} A & \boldsymbol{x} \\ \boldsymbol{y}' & z \end{vmatrix}=|A|(z-\boldsymbol{y}'A^{-1}\boldsymbol{x})=z|A|-\boldsymbol{y}'A^*\boldsymbol{x}.$$

对于一般的方阵 A, 可取到一列有理数 $t_k\to 0$, 使得 t_kI_n+A 为非异阵. 由非异阵情形的证明可得

$$\begin{vmatrix} t_kI_n+A & \boldsymbol{x} \\ \boldsymbol{y}' & z \end{vmatrix}=z|t_kI_n+A|-\boldsymbol{y}'(t_kI_n+A)^*\boldsymbol{x}.$$

注意到上式两边都是关于 t_k 的多项式, 从而关于 t_k 连续. 上式两边同时取极限, 令 $t_k\to 0$, 即有

$$\begin{vmatrix} A & \boldsymbol{x} \\ \boldsymbol{y}' & z \end{vmatrix}=z|A|-\boldsymbol{y}'A^*\boldsymbol{x}=z|A|-\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}x_iy_j.$$

□

例题 1.5 设 n 阶行列式 $|\boldsymbol{A}|=|a_{ij}|,A_{ij}$ 是元素 a_{ij} 的代数余子式, 求证:

$$|B|=\begin{vmatrix} a_{11}-a_{12} & a_{12}-a_{13} & \cdots & a_{1,n-1}-a_{1n} & 1 \\ a_{21}-a_{22} & a_{22}-a_{23} & \cdots & a_{2,n-1}-a_{2n} & 1 \\ a_{31}-a_{32} & a_{32}-a_{33} & \cdots & a_{3,n-1}-a_{3n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}-a_{n2} & a_{n2}-a_{n3} & \cdots & a_{n,n-1}-a_{nn} & 1 \end{vmatrix}=\sum_{i,j=1}^{n}A_{ij}.$$

证明 证法一:设 $|\boldsymbol{A}|$ 的列向量依次为 $\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\cdots,\boldsymbol{\alpha_n}$, 并且 $\boldsymbol{1}$ 表示元素均为 1 的列向量. 则

$$|\boldsymbol{B}|=|\boldsymbol{\alpha_1}-\boldsymbol{\alpha_2},\boldsymbol{\alpha_2}-\boldsymbol{\alpha_3},\cdots,\boldsymbol{\alpha_{n-1}}-\boldsymbol{\alpha_n},1|\xlongequal[i=n-1,n-2,\cdots,2]{j_i+j_{i-1}}|\boldsymbol{\alpha_1}-\boldsymbol{\alpha_n},\boldsymbol{\alpha_2}-\boldsymbol{\alpha_n},\cdots,\boldsymbol{\alpha_{n-1}}-\boldsymbol{\alpha_n},1|.$$

将最后一列写成 $(\boldsymbol{\alpha_n}+\boldsymbol{1})-\boldsymbol{\alpha_n}$, 进行拆分可得

$$\begin{aligned}
|\boldsymbol{B}|&=|\boldsymbol{\alpha_1}-\boldsymbol{\alpha_n},\boldsymbol{\alpha_2}-\boldsymbol{\alpha_n},\cdots,\boldsymbol{\alpha_{n-1}}-\alpha_n,(\boldsymbol{\alpha_n}+\boldsymbol{1})-\boldsymbol{\alpha_n}|\\
&=|\boldsymbol{\alpha_1}-\boldsymbol{\alpha_n},\boldsymbol{\alpha_2}-\boldsymbol{\alpha_n},\cdots,\boldsymbol{\alpha_{n-1}}-\boldsymbol{\alpha_n},\boldsymbol{\alpha_n}+\boldsymbol{1}|-|\boldsymbol{\alpha_1}-\boldsymbol{\alpha_n},\boldsymbol{\alpha_2}-\boldsymbol{\alpha_n},\cdots,\boldsymbol{\alpha_{n-1}}-\boldsymbol{\alpha_n},\boldsymbol{\alpha_n}|
\end{aligned}$$

$$= |\boldsymbol{\alpha_1}+\mathbf{1},\boldsymbol{\alpha_2}+\mathbf{1},\cdots,\boldsymbol{\alpha_{n-1}}+\mathbf{1},\boldsymbol{\alpha_n}+\mathbf{1}| - |\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\cdots,\boldsymbol{\alpha_{n-1}},\boldsymbol{\alpha_n}|.$$

根据行列式的性质将 $|\boldsymbol{\alpha_1}+\mathbf{1},\boldsymbol{\alpha_2}+\mathbf{1},\cdots,\boldsymbol{\alpha_{n-1}}+\mathbf{1},\boldsymbol{\alpha_n}+\mathbf{1}|$ 每一列都拆分成两列, 然后按 1 所在的列展开得到

$$\begin{aligned}|\boldsymbol{B}| &= |\boldsymbol{\alpha_1}+\mathbf{1},\boldsymbol{\alpha_2}+\mathbf{1},\cdots,\boldsymbol{\alpha_{n-1}}+\mathbf{1},\boldsymbol{\alpha_n}+\mathbf{1}| - |\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\cdots,\boldsymbol{\alpha_{n-1}},\boldsymbol{\alpha_n}|\\ &= |\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\cdots,\boldsymbol{\alpha_{n-1}},\boldsymbol{\alpha_n}| + \sum_{i,j=1}^{n} A_{ij} - |\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\cdots,\boldsymbol{\alpha_{n-1}},\boldsymbol{\alpha_n}| = \sum_{i,j=1}^{n} A_{ij}.\end{aligned}$$

证法二:设 $|\boldsymbol{A}|$ 的列向量依次为 $\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\cdots,\boldsymbol{\alpha_n}$, 并且 $\mathbf{1}$ 表示元素均为 1 的列向量.注意到

$$-\sum_{i,j=1}^{n} A_{ij} = \begin{vmatrix}\boldsymbol{\alpha_1} & \boldsymbol{\alpha_2} & \cdots & \boldsymbol{\alpha_n} & \mathbf{1}\\ 1 & 1 & \cdots & 1 & 0\end{vmatrix}.$$

依次将第 i 列乘以 -1 加到第 $i-1$ 列上去 $(i=2,3,\cdots,n)$, 再按第 $n+1$ 行展开可得

$$\begin{aligned}-\sum_{i,j=1}^{n} A_{ij} &= \begin{vmatrix}\boldsymbol{\alpha_1}-\boldsymbol{\alpha_2} & \boldsymbol{\alpha_2}-\boldsymbol{\alpha_3} & \cdots & \boldsymbol{\alpha_{n-1}}-\boldsymbol{\alpha_n} & \boldsymbol{\alpha_n} & 1\\ 0 & 0 & \cdots & 0 & 1 & 0\end{vmatrix}\\ &= -|\boldsymbol{\alpha_1}-\boldsymbol{\alpha_2},\boldsymbol{\alpha_2}-\boldsymbol{\alpha_3},\cdots,\boldsymbol{\alpha_{n-1}}-\boldsymbol{\alpha_n},1| = -|\boldsymbol{B}|.\end{aligned}$$

结论得证. □

例题 1.6 设 n 阶矩阵 A 的每一行、每一列的元素之和都为零, 证明:A 的每个元素的代数余子式都相等.

证明 证法一:设 $A=(a_{ij})$,$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)'$,$\boldsymbol{y}=(y_1,y_2,\cdots,y_n)'$, 不妨设 x_iy_j 均不相同,$i,j=1,2,\cdots,n$. 考虑如下 $n+1$ 阶矩阵的行列式求值:

$$B=\begin{pmatrix}A & \boldsymbol{x}\\ \boldsymbol{y}' & 0\end{pmatrix}$$

一方面, 由命题 1.5可得 $|B| = -\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}x_iy_j$. 另一方面, 先把行列式 $|B|$ 的第二行,···, 第 n 行全部加到第一行上; 再将第二列,···, 第 n 列全部加到第一列上, 可得

$$\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n} & x_1\\ a_{21} & a_{22} & \cdots & a_{2n} & x_2\\ \vdots & \vdots & & \vdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n\\ y_1 & y_2 & \cdots & y_n & 0\end{vmatrix} = \begin{vmatrix}0 & 0 & \cdots & 0 & \sum_{i=1}^{n} x_i\\ a_{21} & a_{22} & \cdots & a_{2n} & x_2\\ \vdots & \vdots & & \vdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n\\ y_1 & y_2 & \cdots & y_n & 0\end{vmatrix} = \begin{vmatrix}0 & 0 & \cdots & 0 & \sum_{i=1}^{n} x_i\\ 0 & a_{22} & \cdots & a_{2n} & x_2\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & a_{n2} & \cdots & a_{nn} & x_n\\ \sum_{j=1}^{n} y_j & y_2 & \cdots & y_n & 0\end{vmatrix}$$

依次按照第一行和第一列进行展开, 可得 $|B| = -A_{11}\sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j$. 比较上述两个结果, 又由于 x_iy_j 均不相同, 因此可得 A 的所有代数余子式都相等.

证法二:由假设可知 $|\boldsymbol{A}|=0$(每行元素全部加到第一行即得), 从而 $\boldsymbol{A}$ 是奇异矩阵. 若 $\boldsymbol{A}$ 的秩小于 $n-1$, 则 $\boldsymbol{A}$ 的任意一个代数余子式 A_{ij} 都等于零, 结论显然成立. 若 $\boldsymbol{A}$ 的秩等于 $n-1$, 则线性方程组 $\boldsymbol{Ax}=\mathbf{0}$ 的基础解系只含一个向量. 又因为 $\boldsymbol{A}$ 的每一行元素之和都等于零, 所以由命题 2.15可知, 我们可以选取 $\boldsymbol{\alpha}=(1,1,\cdots,1)'$ 作为 $\boldsymbol{Ax}=\mathbf{0}$ 的基础解系. 由命题 3.39 的证明可知 $\boldsymbol{A}^*$ 的每一列都是 $\boldsymbol{Ax}=\mathbf{0}$ 的解, 从而 $\boldsymbol{A}^*$ 的每一列与 $\boldsymbol{\alpha}$ 成比例, 特别地,$\boldsymbol{A}^*$ 的每一行都相等. 对 $\boldsymbol{A}'$ 重复上面的讨论, 可得 $(\boldsymbol{A}')^*$ 的每一行都相等. 注意到 $(\boldsymbol{A}')^*=(\boldsymbol{A}^*)'$, 从而 $\boldsymbol{A}^*$ 的每一列都相等, 于是 $\boldsymbol{A}$ 的所有代数余子式 A_{ij} 都相等. □

1.3 求和法

例题 1.7 设 x_1,x_2,x_3 是方程 $x^3+px+q=0$ 的 3 个根, 求下列行列式的值:

$$|\boldsymbol{A}|=\begin{vmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_1 \\ x_3 & x_1 & x_2 \end{vmatrix}.$$

解 由 *Vieta* 定理可知,$x_1+x_2+x_3=0$. 因此, 我们有

$$|\boldsymbol{A}|=\begin{vmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_1 \\ x_3 & x_1 & x_2 \end{vmatrix} \xlongequal[i=2,3]{r_i+r_1} \begin{vmatrix} 0 & 0 & 0 \\ x_2 & x_3 & x_1 \\ x_3 & x_1 & x_2 \end{vmatrix}=0.$$

□

例题 1.8 设 $b_{ij}=(a_{i1}+a_{i2}+\cdots+a_{in})-a_{ij}$, 求证:

$$\begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{vmatrix}=(-1)^{n-1}(n-1)\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

解

$$\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}=\begin{vmatrix} (a_{11}+a_{12}+\cdots+a_{1n})-a_{11} & (a_{11}+a_{12}+\cdots+a_{1n})-a_{12} & \cdots & (a_{11}+a_{12}+\cdots+a_{1n})-a_{1n} \\ (a_{21}+a_{22}+\cdots+a_{2n})-a_{21} & (a_{21}+a_{22}+\cdots+a_{2n})-a_{22} & \cdots & (a_{21}+a_{22}+\cdots+a_{2n})-a_{2n} \\ \vdots & \vdots & & \vdots \\ (a_{n1}+a_{n2}+\cdots+a_{nn})-a_{n1} & (a_{n1}+a_{n2}+\cdots+a_{nn})-a_{n2} & \cdots & (a_{n1}+a_{n2}+\cdots+a_{nn})-a_{nn} \end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n]{j_i+j_1}\begin{vmatrix} (n-1)(a_{11}+a_{12}+\cdots+a_{1n}) & (a_{11}+a_{12}+\cdots+a_{1n})-a_{12} & \cdots & (a_{11}+a_{12}+\cdots+a_{1n})-a_{1n} \\ (n-1)(a_{21}+a_{22}+\cdots+a_{2n}) & (a_{21}+a_{22}+\cdots+a_{2n})-a_{22} & \cdots & (a_{21}+a_{22}+\cdots+a_{2n})-a_{2n} \\ \vdots & \vdots & & \vdots \\ (n-1)(a_{n1}+a_{n2}+\cdots+a_{nn}) & (a_{n1}+a_{n2}+\cdots+a_{nn})-a_{n2} & \cdots & (a_{n1}+a_{n2}+\cdots+a_{nn})-a_{nn} \end{vmatrix}$$

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

$$\xlongequal[i=2,\cdots,n]{(-1)j_1+j_i}(n-1)\begin{vmatrix} (a_{11}+a_{12}+\cdots+a_{1n}) & -a_{12} & \cdots & -a_{1n} \\ (a_{21}+a_{22}+\cdots+a_{2n}) & -a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ (a_{n1}+a_{n2}+\cdots+a_{nn}) & -a_{n2} & \cdots & -a_{nn} \end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n]{j_i+j_1}(n-1)\begin{vmatrix} a_{11} & -a_{12} & \cdots & -a_{1n} \\ a_{21} & -a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & -a_{n2} & \cdots & -a_{nn} \end{vmatrix}$$

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

□

结论 第二个等号是行列式计算中的一个常用方法**求和法**:

将除第一列外的其余列全部加到第一列上 (或将除第一行外的其余行全部加到第一行上), 使第一列 (或列) 一样或者具有相同形式. 然后根据具体情况将第一列 (或行) 的倍数加到其余列 (或行) 上, 从而将行列式化为我们熟悉的形式.

应用该方法的一般情形:

1. 行列式每行 (或列) 和相等时;
2. 行列式每行 (或列) 和有一定规律时.

例题 1.9 计算 n 阶行列式:

$$|\boldsymbol{A}|=\begin{vmatrix}0&1&\cdots&1&1\\1&0&\cdots&1&1\\\vdots&\vdots&&\vdots&\vdots\\1&1&\cdots&0&1\\1&1&\cdots&1&0\end{vmatrix}.$$

解

$$|\boldsymbol{A}|=\begin{vmatrix}0&1&\cdots&1&1\\1&0&\cdots&1&1\\\vdots&\vdots&&\vdots&\vdots\\1&1&\cdots&0&1\\1&1&\cdots&1&0\end{vmatrix}\xlongequal[i=2,\cdots,n]{j_i+j_1}\begin{vmatrix}n-1&1&\cdots&1&1\\n-1&0&\cdots&1&1\\\vdots&\vdots&&\vdots&\vdots\\n-1&1&\cdots&0&1\\n-1&1&\cdots&1&0\end{vmatrix}=(n-1)\begin{vmatrix}1&1&\cdots&1&1\\1&0&\cdots&1&1\\\vdots&\vdots&&\vdots&\vdots\\1&1&\cdots&0&1\\1&1&\cdots&1&0\end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n]{(-1)r_1+r_i}(n-1)\begin{vmatrix}1&1&\cdots&1&1\\0&-1&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&-1&0\\0&0&\cdots&0&-1\end{vmatrix}=(-1)^{n-1}(n-1).$$

□

注 因为 $|\boldsymbol{A}|$ 除对角元素外, 每行都一样, 所以本题也可以看成命题1.1的应用, 利用命题1.1的计算方法直接得到结果.

$$|\boldsymbol{A}|=\begin{vmatrix}0&1&\cdots&1&1\\1&0&\cdots&1&1\\\vdots&\vdots&&\vdots&\vdots\\1&1&\cdots&0&1\\1&1&\cdots&1&0\end{vmatrix}\xlongequal[i=2,\cdots,n]{(-1)r_1+r_i}\begin{vmatrix}0&1&\cdots&1&1\\1&-1&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots\\1&0&\cdots&-1&0\\1&0&\cdots&0&-1\end{vmatrix}\xlongequal{\text{命题1.2}}-\sum_{i=2}^{n}(-1)^{n-2}=(-1)^{n-1}(n-1).$$

例题 1.10 计算 n 阶行列式:

$$|\boldsymbol{A}|=\begin{vmatrix}a_1+b&a_2&a_3&\cdots&a_n\\a_1&a_2+b&a_3&\cdots&a_n\\a_1&a_2&a_3+b&\cdots&a_n\\\vdots&\vdots&\vdots&&\vdots\\a_1&a_2&a_3&\cdots&a_n+b\end{vmatrix}.$$

笔记 既可以将 $|\boldsymbol{A}|$ 看作命题1.1的应用, 利用命题1.1的计算方法直接得到结果. 即下述解法一.

也可以利用求和法将 $|\boldsymbol{A}|$ 化为上三角形行列式. 即下述解法二.

解 解法一:

$$|A|=\begin{vmatrix}a_1+b & a_2 & a_3 & \cdots & a_n\\ a_1 & a_2+b & a_3 & \cdots & a_n\\ a_1 & a_2 & a_3+b & \cdots & a_n\\ \vdots & \vdots & \vdots & & \vdots\\ a_1 & a_2 & a_3 & \cdots & a_n+b\end{vmatrix}\xlongequal[i=2,\cdots,n]{-r_1+r_i}\begin{vmatrix}a_1+b & a_2 & a_3 & \cdots & a_n\\ -b & b & 0 & \cdots & 0\\ -b & 0 & b & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ -b & 0 & 0 & \cdots & b\end{vmatrix}$$

$$\xlongequal{\text{命题1.3}}(a_1+b)\,b^{n-1}-\sum_{i=2}^{n}b^{n-2}a_i\,(-b)=b^{n-1}\left[(a_1+b)+\sum_{i=2}^{n}a_i\right]=\left(b+\sum_{i=1}^{n}a_i\right)b^{n-1}.$$

解法二:

$$|A|=\begin{vmatrix}a_1+b & a_2 & a_3 & \cdots & a_n\\ a_1 & a_2+b & a_3 & \cdots & a_n\\ a_1 & a_2 & a_3+b & \cdots & a_n\\ \vdots & \vdots & \vdots & & \vdots\\ a_1 & a_2 & a_3 & \cdots & a_n+b\end{vmatrix}\xlongequal[i=2,\cdots,n]{j_i+j_1}(b+\sum_{i=1}^{n}a_i)\begin{vmatrix}1 & a_2 & a_3 & \cdots & a_n\\ 1 & a_2+b & a_3 & \cdots & a_n\\ 1 & a_2 & a_3+b & \cdots & a_n\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & a_2 & a_3 & \cdots & a_n+b\end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n]{-a_i\cdot j_1+j_i}(b+\sum_{i=1}^{n}a_i)\begin{vmatrix}1 & 0 & 0 & \cdots & 0\\ 1 & b & 0 & \cdots & 0\\ 1 & 0 & b & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & 0 & 0 & \cdots & b\end{vmatrix}=(b+\sum_{i=1}^{n}a_i)b^{n-1}.$$

□

例题 1.11 计算 n 阶行列式:

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

笔记 求和法的经典应用.

解

$$|\boldsymbol{A}| = \begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ n & 1 & 2 & \cdots & n-2 & n-1 \\ n-1 & n & 1 & \cdots & n-3 & n-2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 3 & 4 & 5 & \cdots & 1 & 2 \\ 2 & 3 & 4 & \cdots & n & 1 \end{vmatrix} \xlongequal[i=2,\cdots,n]{j_i+j_1} \frac{n(n+1)}{2} \begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 1 & 1 & 2 & \cdots & n-2 & n-1 \\ 1 & n & 1 & \cdots & n-3 & n-2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & 4 & 5 & \cdots & 1 & 2 \\ 1 & 3 & 4 & \cdots & n & 1 \end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n]{-r_1+r_i} \frac{n(n+1)}{2} \begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 0 & -1 & -1 & \cdots & -1 & -1 \\ 0 & n-2 & -2 & \cdots & -2 & -2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 2 & 2 & \cdots & 2-n & 2-n \\ 0 & 1 & 1 & \cdots & 1 & 1-n \end{vmatrix} \xlongequal{\text{按第一列展开}} \frac{n(n+1)}{2} \begin{vmatrix} -1 & -1 & \cdots & -1 & -1 \\ n-2 & -2 & \cdots & -2 & -2 \\ \vdots & \vdots & & \vdots & \vdots \\ 2 & 2 & \cdots & 2-n & 2-n \\ 1 & 1 & \cdots & 1 & 1-n \end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n]{-j_1+j_i} \frac{n(n+1)}{2} \begin{vmatrix} -1 & 0 & \cdots & 0 & 0 \\ n-2 & -n & \cdots & -n & -n \\ \vdots & \vdots & & \vdots & \vdots \\ 2 & 0 & \cdots & -n & -n \\ 1 & 0 & \cdots & 0 & -n \end{vmatrix} \xlongequal{\text{按第一行展开}} -\frac{n(n+1)}{2} \begin{vmatrix} -n & \cdots & -n & -n \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & -n & -n \\ 0 & \cdots & 0 & -n \end{vmatrix}$$

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

□

1.4 递推法与数学归纳法

命题 1.6 (三对角行列式)

求下列行列式的递推关系式 (空白处均为 0):

$$D_n = \begin{vmatrix} a_1 & b_1 & & & & \\ c_1 & a_2 & b_2 & & & \\ & c_2 & a_3 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & a_{n-1} & b_{n-1} \\ & & & & c_{n-1} & a_n \end{vmatrix}.$$

笔记 记忆三对角行列式的计算方法和结果: $\boldsymbol{D_n = a_n D_{n-1} - b_{n-1}c_{n-1}D_{n-2}(n \geq 2)}$, 即按最后一列 (或行) 展开得到递推公式.

解 显然 $D_0 = 1, D_1 = a_1$. 当 $n \geq 2$ 时, 我们有

$$D_n = \begin{vmatrix} a_1 & b_1 & & & & \\ c_1 & a_2 & b_2 & & & \\ & c_2 & a_3 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & a_{n-1} & b_{n-1} \\ & & & & c_{n-1} & a_n \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & & & & & \\ c_1 & a_2 & b_2 & & & & \\ & c_2 & a_3 & \ddots & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & a_{n-2} & b_{n-2} & \\ & & & & c_{n-2} & a_{n-1} & b_{n-1} \\ & & & & & c_{n-1} & a_n \end{vmatrix}$$

$$\xlongequal{\text{按最后一列展开}} a_n \begin{vmatrix} a_1 & b_1 & & & & \\ c_1 & a_2 & b_2 & & & \\ & c_2 & a_3 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & a_{n-2} & b_{n-2} \\ & & & & c_{n-2} & a_{n-1} \end{vmatrix} - b_{n-1} \begin{vmatrix} a_1 & b_1 & & & & & \\ c_1 & a_2 & b_2 & & & & \\ & c_2 & a_3 & \ddots & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & a_{n-3} & b_{n-3} & \\ & & & & c_{n-3} & a_{n-2} & b_{n-2} \\ & & & & & 0 & c_{n-1} \end{vmatrix}$$

$$\xlongequal{\text{第二项按最后展开}} a_n \begin{vmatrix} a_1 & b_1 & & & & \\ c_1 & a_2 & b_2 & & & \\ & c_2 & a_3 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & a_{n-2} & b_{n-2} \\ & & & & c_{n-2} & a_{n-1} \end{vmatrix} - b_{n-1}c_{n-1} \begin{vmatrix} a_1 & b_1 & & & & \\ c_1 & a_2 & b_2 & & & \\ & c_2 & a_3 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & a_{n-3} & b_{n-3} \\ & & & & c_{n-3} & a_{n-2} \end{vmatrix}$$

$$= a_n D_{n-1} - b_{n-1}c_{n-1}D_{n-2}.$$

□

推论 1.2

计算 n 阶行列式 $(bc \neq 0)$:

$$D_n = \begin{vmatrix} a & b & & & & \\ c & a & b & & & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & & & c & a & b \\ & & & & c & a \end{vmatrix}.$$

♡

笔记 解递推式:$D_n = aD_{n-1} - bcD_{n-2}(n \geq 2)$ 对应的特征方程:$x^2 - ax + bc = 0$ 得到两根 $\alpha = \dfrac{a+\sqrt{a^2-4bc}}{2}, \beta = \dfrac{a-\sqrt{a^2-4bc}}{2}$, 由 Vieta 定理可知 $a = \alpha + \beta, bc = \alpha\beta$.

若 a, b, c 均为复数, 则上述特征方程

解 由命题 1.6可知, 递推式为 $D_n = aD_{n-1} - bcD_{n-2}(n \geq 2)$. 又易知 $D_0 = 1, D_1 = a$. 令 $\alpha = \dfrac{a+\sqrt{a^2-4bc}}{2}, \beta = \dfrac{a-\sqrt{a^2-4bc}}{2}$, 则 $a = \alpha + \beta, bc = \alpha\beta$, 于是 $D_n = (\alpha+\beta)D_{n-1} - \alpha\beta D_{n-2}(n \geq 2)$. 从而

$$D_n - \alpha D_{n-1} = \beta(D_{n-1} - \alpha D_{n-2}), D_n - \beta D_{n-1} = \alpha(D_{n-1} - \beta D_{n-2}).$$

于是

$$D_n - \alpha D_{n-1} = \beta^{n-1}(D_1 - \alpha D_0) = \beta^{n-1}(a-\alpha) = \beta^n,$$
$$D_n - \beta D_{n-1} = \alpha^{n-1}(D_1 - \beta D_0) = \alpha^{n-1}(a-\beta) = \alpha^n.$$

因此, 若 $a^2 \neq 4bc$(即$\alpha \neq \beta$), 则联立上面两式, 解得

$$D_n = \frac{\alpha^{n+1}-\beta^{n+1}}{\alpha-\beta};$$

若 $a^2 = 4bc$(即$\alpha = \beta$), 则由 $a = \alpha+\beta$ 可知,$\alpha = \beta = \frac{a}{2}$. 又由 $D_n - \alpha D_{n-1} = \beta^n$ 可得

$$D_n = \left(\frac{a}{2}\right)^n + \frac{a}{2}D_{n-1} = \left(\frac{a}{2}\right)^n + \frac{a}{2}\left(\left(\frac{a}{2}\right)^{n-1} + \frac{a}{2}D_{n-2}\right) = 2\left(\frac{a}{2}\right)^n + \left(\frac{a}{2}\right)^2 D_{n-2} = \cdots = n\left(\frac{a}{2}\right)^n + \left(\frac{a}{2}\right)^n D_0 = (n+1)\left(\frac{a}{2}\right)^n.$$

综上, 我们有

$$D_n = \begin{cases} \frac{\alpha^{n+1}-\beta^{n+1}}{\alpha-\beta}, a^2 \neq 4bc, \\ (n+1)\left(\frac{\alpha}{2}\right)^n, a^2 = 4bc. \end{cases}$$

□

练习 1.1 求证:n 阶行列式

$$|\boldsymbol{A}| = \begin{vmatrix} \cos x & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 2\cos x & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 2\cos x & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2\cos x & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 2\cos x \end{vmatrix} = \cos nx.$$

解 解法一:

设 $|\boldsymbol{A}| = D_n$, 其中 n 表示 $|\boldsymbol{A}|$ 的阶数 $(n \geq 0)$. 易知 $D_0 = 1, D_1 = \cos x$.

从而 $|\boldsymbol{A}| = D_n \xlongequal[\text{命题1.6}]{\text{按最后一列展开}} 2\cos x D_{n-1} - D_{n-2}\,(n \geq 2)$.

其对应的特征方程为 $\lambda^2 = 2\cos x\lambda - 1$, 解得 $\lambda_1 = \cos x + i\sin x, \lambda_2 = \cos x - i\sin x$.

于是当 $n \geq 2$ 时, 我们有 $D_n = (\lambda_1+\lambda_2)D_{n-1} + \lambda_1\lambda_2 D_{n-2}$.

进而

$$\begin{aligned} D_n - \lambda_1 D_{n-1} &= \lambda_2(D_n - \lambda_1 D_{n-1}), \\ D_n - \lambda_2 D_{n-1} &= \lambda_1(D_n - \lambda_2 D_{n-1}). \end{aligned} \tag{1.1}$$

由此可得

$$D_n - \lambda_1 D_{n-1} = {\lambda_2}^{n-1}(D_1 - \lambda_1 D_0) = -i\sin x \cdot {\lambda_2}^{n-1},$$
$$D_n - \lambda_2 D_{n-1} = {\lambda_1}^{n-1}(D_1 - \lambda_2 D_0) = i\sin x \cdot {\lambda_1}^{n-1}.$$

若 $x \neq k\pi(k \in \mathbb{Z})$, 则联立上面两式, 解得

$$\begin{aligned} D_n &= \frac{i\sin x \cdot {\lambda_1}^n + i\sin x \cdot {\lambda_2}^n}{\lambda_1 - \lambda_2} = \frac{i\sin x \cdot (\cos x + i\sin x)^n + i\sin x \cdot (\cos x - i\sin x)^n}{2i\sin x} \\ &\xlongequal[e^{ix}=\cos x+i\sin x, e^{-ix}=\cos x-i\sin x.]{Euler\text{公式}} \frac{i\sin x \cdot e^{nxi} + i\sin x \cdot e^{-nxi}}{2i\sin x} = \frac{i\sin x \cdot (\cos nx + i\sin nx) + i\sin x \cdot (\cos nx - i\sin nx)}{2i\sin x} \\ &= \frac{2i\sin x \cdot \cos nx}{2i\sin x} = \cos nx. \end{aligned}$$

若 $x = k\pi(k \in \mathbb{Z})$, 则 $\lambda_1 = \lambda_2 = \cos k\pi$. 从而由(1.1)式可得,$D_n - \cos k\pi D_{n-1} = -i\sin x \cdot (\cos k\pi) = 0$.

于是

$$D_n = \cos k\pi D_{n-1} = (\cos k\pi)^2 D_{n-2} = \cdots = (\cos k\pi)^n D_0 = (\cos k\pi)^n = (-1)^{kn} = \cos(nk\pi) = \cos nx.$$

解法二:仿照练习1.14中的数学归纳法证明. □

练习 1.2 求下列 n 阶行列式的值:

$$D_n=\begin{vmatrix}1-a_1 & a_2 & 0 & 0 & \cdots & 0 & 0\\ -1 & 1-a_2 & a_3 & 0 & \cdots & 0 & 0\\ 0 & -1 & 1-a_3 & a_4 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & -1 & 1-a_n\end{vmatrix}.$$

笔记 观察原行列式我们可以得到,D_n 的每列和有一定的规律,即除了第一列和最后一列,中间每列和均为 0. 并且 D_n 是三对角行列式. 因此,我们既可以直接应用三对角行列式的结论 (即命题1.6),又可以使用求和法进行求解. 如果我们直接应用三对角行列式的结论 (即命题1.6),按照对一般的三对角行列式展开的方法能得到相应递推式,但是这样得到的递推式并不是相邻两项之间的递推,后续求解通项并不简便. 又因为使用求和法计算行列式后续计算一般比较简便所以我们先采用求和法进行尝试.

解 解法一:当 $n\geq 1$ 时,我们有

$$D_n=\begin{vmatrix}1-a_1 & a_2 & 0 & 0 & \cdots & 0 & 0\\ -1 & 1-a_2 & a_3 & 0 & \cdots & 0 & 0\\ 0 & -1 & 1-a_3 & a_4 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & -1 & 1-a_n\end{vmatrix}\xlongequal[i=2,\cdots,n]{r_i+r_1}\begin{vmatrix}-a_1 & 0 & 0 & 0 & \cdots & 0 & 1\\ -1 & 1-a_2 & a_3 & 0 & \cdots & 0 & 0\\ 0 & -1 & 1-a_3 & a_4 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & -1 & 1-a_n\end{vmatrix}$$

$$\xlongequal{\text{按第一行展开}}-a_1D_{n-1}+(-1)^{n+1}\begin{vmatrix}-1 & 1-a_2 & a_3 & 0 & \cdots & 0\\ 0 & -1 & 1-a_3 & a_4 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & 0 & \cdots & -1\end{vmatrix}$$

$$=-a_1D_{n-1}+(-1)^{n+1}(-1)^{n-1}$$

$$=1-a_1D_{n-1}.$$

其中 D_{n-i} 表示 D_{n-i+1} 去掉第一行和第一列得到的 $n-i$ 阶行列式,$i=1,2,\cdots,n-1$. (或者称 D_{n-i} 表示以 $a_{i+1},\cdots,a_n$ 为未定元的 $n-i$ 阶行列式,$i=1,2,\cdots,n-1$)

由递推不难得到

$$D_n=1-a_1(1-a_2D_{n-2})=1-a_1+a_1a_2D_{n-2}=\cdots=1-a_1+a_1a_2-a_1a_2a_3+\cdots+(-1)^n a_1a_2\cdots a_n.$$

解法二:仿照练习1.14中的数学归纳法证明. □

命题 1.7

计算 n 阶行列式:

$$D_n=\begin{vmatrix}x_1 & y & y & \cdots & y & y\\ z & x_2 & y & \cdots & y & y\\ z & z & x_3 & \cdots & y & y\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ z & z & z & \cdots & x_{n-1} & y\\ z & z & z & \cdots & z & x_n\end{vmatrix}.$$

笔记 解法二: $f(x) \triangleq \begin{vmatrix} x_1+x & y+x & \cdots & y+x \\ z+x & x_2+x & \cdots & y+x \\ \vdots & \vdots & & \vdots \\ z+x & z+x & \cdots & x_n+x \end{vmatrix} = \begin{vmatrix} x_1+x & y+x & \cdots & y+x \\ z-x_1 & x_2-y & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ z-x_1 & z-y & \cdots & x_n-y \end{vmatrix}$, 再按第一行展开可得 $f(x)$ 一定为关于 x 的线性函数.

解 解法一(小拆分法): 对第 n 列进行拆分即可得到递推式: (对第 1 或 n 行 (或列) 拆分都可以得到相同结果)

$$D_n = \begin{vmatrix} x_1 & y & y & \cdots & y & y+0 \\ z & x_2 & y & \cdots & y & y+0 \\ z & z & x_3 & \cdots & y & y+0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ z & z & z & \cdots & x_{n-1} & y+0 \\ z & z & z & \cdots & z & y+x_n-y \end{vmatrix} = \begin{vmatrix} x_1 & y & y & \cdots & y & y \\ z & x_2 & y & \cdots & y & y \\ z & z & x_3 & \cdots & y & y \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ z & z & z & \cdots & x_{n-1} & y \\ z & z & z & \cdots & z & y \end{vmatrix} + \begin{vmatrix} x_1 & y & y & \cdots & y & 0 \\ z & x_2 & y & \cdots & y & 0 \\ z & z & x_3 & \cdots & y & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ z & z & z & \cdots & x_{n-1} & 0 \\ z & z & z & \cdots & z & x_n-y \end{vmatrix}$$

$$= \begin{vmatrix} x_1-z & 0 & 0 & \cdots & 0 & 0 \\ 0 & x_2-z & 0 & \cdots & 0 & 0 \\ 0 & 0 & x_3-z & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x_{n-1}-z & 0 \\ z & z & z & \cdots & z & y \end{vmatrix} + (x_n-y)\,D_{n-1} = y\prod_{i=1}^{n-1}(x_i-z) + (x_n-y)\,D_{n-1}. \tag{1.2}$$

将原行列式转置后, 同理可得

$$D_n = D_n^T = \begin{vmatrix} x_1 & z & z & \cdots & z & z+0 \\ y & x_2 & z & \cdots & z & z+0 \\ y & y & x_3 & \cdots & z & z+0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ y & y & y & \cdots & x_{n-1} & z+0 \\ y & y & y & \cdots & y & z+x_n-z \end{vmatrix} = \begin{vmatrix} x_1 & z & z & \cdots & z & z \\ y & x_2 & z & \cdots & z & z \\ y & y & x_3 & \cdots & z & z \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ y & y & y & \cdots & x_{n-1} & z \\ y & y & y & \cdots & y & z \end{vmatrix} + \begin{vmatrix} x_1 & z & z & \cdots & z & 0 \\ y & x_2 & z & \cdots & z & 0 \\ y & y & x_3 & \cdots & z & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ y & y & y & \cdots & x_{n-1} & 0 \\ y & y & y & \cdots & y & x_n-z \end{vmatrix}$$

$$= \begin{vmatrix} x_1-y & 0 & 0 & \cdots & 0 & 0 \\ 0 & x_2-y & 0 & \cdots & 0 & 0 \\ 0 & 0 & x_3-y & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x_{n-1}-y & 0 \\ y & y & y & \cdots & y & z \end{vmatrix} + (x_n-z)\,D_{n-1}^T = z\prod_{i=1}^{n-1}(x_i-y) + (x_n-z)\,D_{n-1}. \tag{1.3}$$

若 $z \neq y$, 则联立(1.2)(1.3)式, 解得

$$D_n = \frac{1}{z-y}\left[z\prod_{i=1}^{n}(x_i-y) - y\prod_{i=1}^{n}(x_i-z)\right];$$

若 $z=y$, 则由(1.2)式递推可得

$$
\begin{aligned}
D_n &= y\prod_{i=1}^{n-1}(x_i-y)+(x_n-y)\,D_{n-1}\\
&= y\prod_{i=1}^{n-1}(x_i-y)+(x_n-y)\left(y\prod_{i=1}^{n-2}(x_i-y)+(x_{n-1}-y)\,D_{n-2}\right)\\
&= y\prod_{j\neq n}(x_i-y)+y\prod_{j\neq n-1}(x_i-y)+(x_n-y)\,(x_{n-1}-y)\,D_{n-2}\\
&= \cdots = y\sum_{i=1}^{n}\prod_{j\neq i}(x_j-y)+\prod_{i=1}^{n}(x_i-y)D_0\\
&= y\sum_{i=1}^{n}\prod_{j\neq i}(x_j-y)+\prod_{i=1}^{n}(x_i-y).
\end{aligned}
$$

解法二(大拆分法):令 $f(x)\triangleq\begin{vmatrix} x_1+x & y+x & \cdots & y+x\\ z+x & x_2+x & \cdots & y+x\\ \vdots & \vdots & & \vdots\\ z+x & z+x & \cdots & x_n+x\end{vmatrix}$, 则 $f(x)$ 一定是线性函数, 从而设 $f(x)=ax+b$. 注意到

$$
f(-z)=\begin{vmatrix} x_1-z & y-z & \cdots & y-z\\ 0 & x_2-z & \cdots & y-z\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & x_n-z\end{vmatrix}=\prod_{i=1}^{n}(x_i-z),\quad f(-y)=\begin{vmatrix} x_1-y & 0 & \cdots & 0\\ z-y & x_2-y & \cdots & 0\\ \vdots & \vdots & & \vdots\\ z-y & z-y & \cdots & x_n-y\end{vmatrix}=\prod_{i=1}^{n}(x_i-y).
$$

当 $y\neq z$ 时, 将上式代入 $f(x)=ax+b$(即线性函数 $f(x)$ 过两点 $(-y,f(-y))$,$(-z,f(-z))$, 再利用两点式) 解得

$$
f(x)=\frac{f(-z)-f(-y)}{-z-(-y)}(x+y)+f(-y)=\frac{\prod\limits_{i=1}^{n}(x_i-z)-\prod\limits_{i=1}^{n}(x_i-y)}{y-z}(x+y)+\prod_{i=1}^{n}(x_i-y).
$$

从而此时就有

$$
D_n=f(0)=\frac{y\prod\limits_{i=1}^{n}(x_i-z)-z\prod\limits_{i=1}^{n}(x_i-y)}{y-z}. \tag{1.4}
$$

当 $y=z$ 时, 将 D_n 看作关于 y 的连续函数, 记为 $g(y)=D_n$, 则此时由 g 的连续性及(1.4)式和 L'Hospital 法则可得

$$
\begin{aligned}
D_n = g\,(z) &= \lim_{y\to z}g\,(y)=\lim_{y\to z}\frac{y\prod\limits_{i=1}^{n}(x_i-z)-z\prod\limits_{i=1}^{n}(x_i-y)}{y-z}\\
&= \lim_{y\to z}\frac{\prod\limits_{i=1}^{n}(x_i-z)+y\sum\limits_{i=1}^{n}\prod\limits_{j\neq i}(x_j-y)}{1}=\prod_{i=1}^{n}(x_i-z)+z\sum_{i=1}^{n}\prod_{j\neq i}(x_j-z).
\end{aligned}
$$

□

例题 1.12

(1) 计算

$$
|B|=\begin{vmatrix} 2a_1 & a_1+a_2 & \cdots & a_1+a_n\\ a_2+a_1 & 2a_2 & \cdots & a_2+a_n\\ \vdots & \vdots & \ddots & \vdots\\ a_n+a_1 & a_n+a_2 & \cdots & 2a_n\end{vmatrix}.
$$

(2) 求下列 n 阶行列式的值:

$$|\boldsymbol{A}|=\begin{vmatrix}0 & a_1+a_2 & \cdots & a_1+a_{n-1} & a_1+a_n\\ a_2+a_1 & 0 & \cdots & a_2+a_{n-1} & a_2+a_n\\ \vdots & \vdots & & \vdots & \vdots\\ a_{n-1}+a_1 & a_{n-1}+a_2 & \cdots & 0 & a_{n-1}+a_n\\ a_n+a_1 & a_n+a_2 & \cdots & a_n+a_{n-1} & 0\end{vmatrix}.$$

笔记 第 (2) 问解法一中不仅使用了升阶法还使用了分块”爪”型行列式的计算方法. 观察到各行各列有不同的公共项, 因此可以利用升阶法将各行各列的公共项消去.

注 因为第 (2) 问中, 当 $a_i\neq 0(i=1,2,\cdots,n)$ 时, 最后的结果不含 a_i 的分式结构, 所以当存在 $a_i=0$, 其中 $i\in 1,2,\cdot,ns$ 时, 根据行列式 (可以看作多元多项式函数) 的连续性可知, 此时最后的结果就是将 a_i 中相应为零的值代入当 $a_i\neq 0(i=1,2,\cdots,n)$ 时的结果中. 因此我吗们可以直接不妨设 $a_i\neq 0(i=1,2,\cdots,n)$, 只需考虑这一种情况即可.

解

(1) 注意到 $B=\begin{pmatrix}2a_1 & a_1+a_2 & \cdots & a_1+a_n\\ a_2+a_1 & 2a_2 & \cdots & a_2+a_n\\ \vdots & \vdots & \ddots & \vdots\\ a_n+a_1 & a_n+a_2 & \cdots & 2a_n\end{pmatrix}=\begin{pmatrix}1 & a_1\\ 1 & a_2\\ \vdots & \vdots\\ 1 & a_n\end{pmatrix}\begin{pmatrix}1 & 1 & \cdots & 1\\ a_1 & a_2 & \cdots & a_n\end{pmatrix}$. 由 Cauchy-Binet 公式可知,$|B|=\begin{cases}0, & n\geq 3,\\ -(a_1-a_2)^2, & n=2,\\ 2a_1, & n=1.\end{cases}$.

(2) (i) 当 $a_i\neq 0(1\leq i\leq n)$ 时, 解法一 (升阶法):

$$|\boldsymbol{A}|\xlongequal{\text{升阶}}\begin{vmatrix}1 & -a_1 & -a_2 & \cdots & -a_{n-1} & -a_n\\ 0 & 0 & a_1+a_2 & \cdots & a_1+a_{n-1} & a_1+a_n\\ 0 & a_2+a_1 & 0 & \cdots & a_2+a_{n-1} & a_2+a_n\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & a_{n-1}+a_1 & a_{n-1}+a_2 & \cdots & 0 & a_{n-1}+a_n\\ 0 & a_n+a_1 & a_n+a_2 & \cdots & a_n+a_{n-1} & 0\end{vmatrix}$$

$$\xlongequal[i=1,2,\cdots,n+1]{r_1+r_i}\begin{vmatrix}1 & -a_1 & -a_2 & \cdots & -a_{n-1} & -a_n\\ 1 & -a_1 & a_1 & \cdots & a_1 & a_1\\ 1 & a_2 & -a_2 & \cdots & a_2 & a_2\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 1 & a_{n-1} & a_{n-1} & \cdots & -a_{n-1} & a_{n-1}\\ 1 & a_n & a_n & \cdots & a_n & -a_n\end{vmatrix}\xlongequal{\text{升阶}}\begin{vmatrix}1 & 0 & 0 & 0 & \cdots & 0 & 0\\ 0 & 1 & -a_1 & -a_2 & \cdots & -a_{n-1} & -a_n\\ -a_1 & 1 & -a_1 & a_1 & \cdots & a_1 & a_1\\ -a_2 & 1 & a_2 & -a_2 & \cdots & a_2 & a_2\\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\ -a_{n-1} & 1 & a_{n-1} & a_{n-1} & \cdots & -a_{n-1} & a_{n-1}\\ -a_n & 1 & a_n & a_n & \cdots & a_n & -a_n\end{vmatrix}$$

$$\xlongequal[i=1,3,4\cdots,n+2]{j_1+j_i}\begin{vmatrix}1 & 0 & 1 & 1 & \cdots & 1 & 1\\ 0 & 1 & -a_1 & -a_2 & \cdots & -a_{n-1} & -a_n\\ -a_1 & 1 & -2a_1 & 0 & \cdots & 0 & 0\\ -a_2 & 1 & 0 & -2a_2 & \cdots & 0 & 0_2\\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\ -a_{n-1} & 1 & 0 & 0 & \cdots & -2a_{n-1} & 0\\ -a_n & 1 & 0 & 0 & \cdots & 0 & -2a_n\end{vmatrix}$$

$$\xlongequal[i=3,4\cdots,n+2]{\substack{-\frac{1}{2}j_i+j_1 \\ \frac{1}{2a_{i-2}}j_i+j_2}} \begin{vmatrix} 1-\frac{n}{2} & \frac{S}{2} & 1 & 1 & \cdots & 1 & 1 \\ \frac{T}{2} & 1-\frac{n}{2} & -a_1 & -a_2 & \cdots & -a_{n-1} & -a_n \\ 0 & 0 & -2a_1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & -2a_2 & \cdots & 0 & 0_2 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -2a_{n-1} & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -2a_n \end{vmatrix}.$$

其中 $S=a_1+a_2+\cdots+a_n,T=\frac{1}{a_1}+\frac{1}{a_2}+\cdots+\frac{1}{a_n}$. 注意到上述行列式是分块上三角行列式, 从而可得

$$\begin{aligned} |\boldsymbol{A}| &= (-2)^n\prod_{i=1}^n a_i\cdot\frac{(n-2)^2-ST}{4} = (-2)^{n-2}\prod_{i=1}^n a_i[(n-2)^2-(\sum_{i=1}^n a_i)(\sum_{i=1}^n\frac{1}{a_i})] \\ &= (-2)^{n-2}\prod_{i=1}^n a_i\,(n-2)^2-(-2)^{n-2}\sum_{i=1}^n a_i\sum_{j=1}^n\prod_{k\neq j}a_k. \end{aligned}$$

解法二 (直接计算两个矩阵和的行列式)(不推荐使用!):

设 $\boldsymbol{B}=\begin{pmatrix} 2a_1 & a_1+a_2 & \cdots & a_1+a_n \\ a_2+a_1 & 2a_2 & \cdots & a_2+a_n \\ \vdots & \vdots & & \vdots \\ a_n+a_1 & a_n+a_2 & \cdots & 2a_n \end{pmatrix}$,$\boldsymbol{C}=\begin{pmatrix} -2a_1 & & & \\ & -2a_2 & & \\ & & \ddots & \\ & & & -2a_n \end{pmatrix}$, 则 $|\boldsymbol{A}|=|\boldsymbol{B}+\boldsymbol{C}|$.

从而利用直接计算两个矩阵和的行列式的结论得到

$$|\boldsymbol{A}|=|\boldsymbol{B}|+|\boldsymbol{C}|+\sum_{1\leqslant k\leqslant n-1}\left(\sum_{\substack{1\leqslant i_1<i_2<\cdots<i_k\leqslant n \\ 1\leqslant j_1<j_2<\cdots<j_k\leqslant n}}\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}\widehat{\boldsymbol{C}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}\right) \tag{1.5}$$

其中 $\widehat{\boldsymbol{C}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 是 $\boldsymbol{C}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 的代数余子式.

我们先来计算 $\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix},k=1,2,\cdots,n$. 拆分 $\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 的第一列得到

$$\begin{aligned} \boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} &= \begin{vmatrix} a_{i_1}+a_{j_1} & a_{i_1}+a_{j_2} & \cdots & a_{i_1}+a_{j_k} \\ a_{i_2}+a_{j_1} & a_{i_2}+a_{j_2} & \cdots & a_{i_2}+a_{j_k} \\ \vdots & \vdots & & \vdots \\ a_{i_k}+a_{j_1} & a_{i_k}+a_{j_2} & \cdots & a_{i_k}+a_{j_k} \end{vmatrix} \\ &= \begin{vmatrix} a_{i_1} & a_{i_1}+a_{j_2} & \cdots & a_{i_1}+a_{j_k} \\ a_{i_2} & a_{i_2}+a_{j_2} & \cdots & a_{i_2}+a_{j_k} \\ \vdots & \vdots & & \vdots \\ a_{i_k} & a_{i_k}+a_{j_2} & \cdots & a_{i_k}+a_{j_k} \end{vmatrix} + \begin{vmatrix} a_{j_1} & a_{i_1}+a_{j_2} & \cdots & a_{i_1}+a_{j_k} \\ a_{j_1} & a_{i_2}+a_{j_2} & \cdots & a_{i_2}+a_{j_k} \\ \vdots & \vdots & & \vdots \\ a_{j_1} & a_{i_k}+a_{j_2} & \cdots & a_{i_k}+a_{j_k} \end{vmatrix} \\ &= \begin{vmatrix} a_{i_1} & a_{j_2} & \cdots & a_{j_k} \\ a_{i_2} & a_{j_2} & \cdots & a_{j_k} \\ \vdots & \vdots & & \vdots \\ a_{i_k} & a_{j_2} & \cdots & a_{j_k} \end{vmatrix} + \begin{vmatrix} a_{j_1} & a_{i_1} & \cdots & a_{i_1} \\ a_{j_1} & a_{i_2} & \cdots & a_{i_2} \\ \vdots & \vdots & & \vdots \\ a_{j_1} & a_{i_k} & \cdots & a_{i_k} \end{vmatrix} \end{aligned}$$

因此当 $k\geqslant 3$ 时,$\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}=0$; 当 $k=2$ 时,$\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}=\boldsymbol{B}\begin{pmatrix} i_1 & i_2\\ j_1 & j_2\end{pmatrix}=\begin{vmatrix} a_{i_1} & a_{j_2}\\ a_{i_2} & a_{j_2}\end{vmatrix}+\begin{vmatrix} a_{j_1} & a_{i_1}\\ a_{j_1} & a_{i_2}\end{vmatrix}=(a_{i_1}a_{j_2}-a_{i_2}a_{j_2})(a_{i_2}a_{j_1}-a_{i_1}a_{j_1})$; 当 $k=1$ 时,$\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}=\boldsymbol{B}\begin{pmatrix} i_1\\ j_1\end{pmatrix}=a_{i_1}+a_{j_1}$.

又注意到 $|\boldsymbol{C}|$ 只有主子式非零,而其主子式 $\boldsymbol{C}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}=(-2)^k a_{i_1}a_{i_2}\cdots a_{i_k}$. 于是当 $\exists m\in\{1,2,\cdots,k\}$,

使得 $i_m\neq j_m$ 时,$\widehat{\boldsymbol{C}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}=0$; 当 $i_m\neq j_m, m=1,2,\cdots,k$ 时,$\widehat{\boldsymbol{C}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}=\widehat{\boldsymbol{C}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}=$
$(-2)^{n-k}a_1\cdots\hat{a}_{i_1}\cdots\hat{a}_{i_2}\cdots\hat{a}_{i_k}\cdots a_n$.

故当 $n\geqslant 3$ 时,(1.5)式可化为

$$|\boldsymbol{A}|=|\boldsymbol{B}|+|\boldsymbol{C}|+\sum_{1\leqslant k\leqslant n-1}\left(\sum_{\substack{1\leqslant i_1<i_2<\cdots<i_k\leqslant n\\ 1\leqslant j_1<j_2<\cdots<j_k\leqslant n}}\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\widehat{\boldsymbol{C}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\right)$$

$$=|\boldsymbol{C}|+\sum_{\substack{1\leqslant i_1\leqslant n\\ 1\leqslant j_1\leqslant n}}\boldsymbol{B}\binom{i_1}{j_1}\widehat{\boldsymbol{C}}\begin{pmatrix} i_1\\ j_1\end{pmatrix}+\sum_{\substack{1\leqslant i_1<i_2\leqslant n\\ 1\leqslant j_1<j_2\leqslant n}}\boldsymbol{B}\begin{pmatrix} i_1 & i_2\\ j_1 & j_2\end{pmatrix}\widehat{\boldsymbol{C}}\begin{pmatrix} i_1 & i_2\\ j_1 & j_2\end{pmatrix}$$

$$=|\boldsymbol{C}|+\sum_{1\leqslant i_1\leqslant n}\boldsymbol{B}\begin{pmatrix} i_1\\ i_1\end{pmatrix}\widehat{\boldsymbol{C}}\begin{pmatrix} i_1\\ i_1\end{pmatrix}+\sum_{1\leqslant i_1<i_2\leqslant n}\boldsymbol{B}\begin{pmatrix} i_1 & i_2\\ i_1 & i_2\end{pmatrix}\widehat{\boldsymbol{C}}\begin{pmatrix} i_1 & i_2\\ i_1 & i_2\end{pmatrix}=|\boldsymbol{C}|+\sum_{1\leqslant i\leqslant n}\boldsymbol{B}\begin{pmatrix} i\\ i\end{pmatrix}\widehat{\boldsymbol{C}}\begin{pmatrix} i\\ i\end{pmatrix}+\sum_{1\leqslant i<j\leqslant n}\boldsymbol{B}\begin{pmatrix} i & j\\ i & j\end{pmatrix}\widehat{\boldsymbol{C}}$$

$$=(-2)^n a_1a_2\cdots a_n+\sum_{1\leqslant i\leqslant n}2a_i(-2)^{n-1}a_1\cdots\hat{a}_i\cdots a_n+\sum_{1\leqslant i<j\leqslant n}[(a_ia_j-a_j^2)(a_ia_j-a_i^2)(-2)^{n-2}a_1\cdots\hat{a}_i\cdots\hat{a}_j\cdots a_n]$$

$$=(-2)^n a_1a_2\cdots a_n-(-2)^n\sum_{1\leqslant i\leqslant n}a_1a_2\cdots\cdots a_n+(-2)^{n-2}\sum_{1\leqslant i<j\leqslant n}[-(a_i-a_j)^2a_1\cdots\hat{a}_i\cdots\hat{a}_j\cdots a_n]$$

$$=(-2)^n a_1a_2\cdots a_n-(-2)^n na_1a_2\cdots\cdots a_n-(-2)^{n-2}\sum_{1\leqslant i<j\leqslant n}[(a_i-a_j)^2a_1\cdots\hat{a}_i\cdots\hat{a}_j\cdots a_n]$$

$$=(-2)^n\prod_{i=1}^n a_i(1-n)-(-2)^{n-2}\prod_{i=1}^n a_i\sum_{1\leqslant i<j\leqslant n}\frac{(a_i-a_j)^2}{a_ia_j}$$

$$=(-2)^{n-2}\prod_{i=1}^n a_i[(n-2)^2-(\sum_{i=1}^n a_i)(\sum_{i=1}^n\frac{1}{a_i})]$$

$$=(-2)^n\prod_{i=1}^n a_i(1-n)-(-2)^{n-2}\prod_{i=1}^n a_i\sum_{1\leqslant i<j\leqslant n}\frac{(a_i-a_j)^2}{a_ia_j}$$

$$=(-2)^{n-2}\prod_{i=1}^n a_i\left[4-4n-\sum_{1\leqslant i<j\leqslant n}\frac{(a_i-a_j)^2}{a_ia_j}\right]$$

$$=(-2)^{n-2}\prod_{i=1}^n a_i\left[4-4n-\sum_{1\leqslant i<j\leqslant n}\left(\frac{a_j}{a_i}+\frac{a_i}{a_j}-2\right)\right]$$

$$=(-2)^{n-2}\prod_{i=1}^n a_i\left[4-4n-\sum_{\substack{1\leqslant i,j\leqslant n\\ i\neq j}}\frac{a_i}{a_j}+\sum_{1\leqslant i<j\leqslant n}2\right]$$

$$=(-2)^{n-2}\prod_{i=1}^n a_i\left[4-4n-\left(\sum_{1\leqslant i,j\leqslant n}\frac{a_i}{a_j}-\sum_{i=1}^n\frac{a_i}{a_i}\right)+\sum_{i=1}^{n-1}\sum_{j=i+1}^n 2\right]$$

$$=(-2)^{n-2}\prod_{i=1}^n a_i\left[4-4n-\left(\sum_{1\leqslant i,j\leqslant n}\frac{a_i}{a_j}-n\right)+2\sum_{i=1}^{n-1}(n-i)\right]$$

$$
\begin{aligned}
&= (-2)^{n-2}\prod_{i=1}^{n} a_i \left[4-4n+n+n(n-1)-\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_i}{a_j}\right]\\
&= (-2)^{n-2}\prod_{i=1}^{n} a_i \left[n^2-4n+4-\sum_{i=1}^{n}a_i\sum_{j=1}^{n}\frac{1}{a_j}\right]\\
&= (-2)^{n-2}\prod_{i=1}^{n} a_i[(n-2)^2-(\sum_{i=1}^{n}a_i)(\sum_{i=1}^{n}\frac{1}{a_i})]\\
&= (-2)^{n-2}\prod_{i=1}^{n} a_i (n-2)^2-(-2)^{n-2}\sum_{i=1}^{n}a_i\sum_{j=1}^{n}\prod_{k\neq j}a_k.
\end{aligned}
$$

解法三 (降价公式)(推荐使用!): 令 $\Lambda=\begin{pmatrix} a_1 & 1\\ a_2 & 1\\ \vdots & \vdots\\ a_n & 1\end{pmatrix}$, $B=\begin{pmatrix} -2a_1 & & & \\ & -2a_2 & & \\ & & \ddots & \\ & & & -2a_n\end{pmatrix}$, 则

$$
A=\begin{pmatrix} -2a_1 & & & \\ & -2a_2 & & \\ & & \ddots & \\ & & & -2a_n\end{pmatrix}+\begin{pmatrix} a_1 & 1\\ a_2 & 1\\ \vdots & \vdots\\ a_n & 1\end{pmatrix} I_2^{-1}\begin{pmatrix} 1 & 1 & \cdots & 1\\ a_1 & a_2 & \cdots & a_n\end{pmatrix}=B+\Lambda I_2^{-1}\Lambda'.
$$

于是由降价公式 (打洞原理) 我们有

$$
|A|=|I|\left|B+\Lambda I_2^{-1}\Lambda'\right|=\begin{vmatrix} I_2 & \Lambda'\\ \Lambda & B\end{vmatrix}=|B|\left|I_2-\Lambda' B^{-1}\Lambda\right|
$$

$$
\begin{aligned}
&=\begin{vmatrix} -2a_1 & & & \\ & -2a_2 & & \\ & & \ddots & \\ & & & -2a_n\end{vmatrix}\cdot\left|I_2-\begin{pmatrix} 1 & 1 & \cdots & 1\\ a_1 & a_2 & \cdots & a_n\end{pmatrix}\begin{pmatrix} -\frac{1}{2a_1} & & & \\ & -\frac{1}{2a_2} & & \\ & & \ddots & \\ & & & -\frac{1}{2a_n}\end{pmatrix}\begin{pmatrix} a_1 & 1\\ a_2 & 1\\ \vdots & \vdots\\ a_n & 1\end{pmatrix}\right|\\
&=(-2)^n\prod_{i=1}^{n}a_i\left|I_2-\begin{pmatrix} -\frac{1}{2a_1} & -\frac{1}{2a_2} & \cdots & -\frac{1}{2a_n}\\ -\frac{1}{2} & -\frac{1}{2} & \cdots & -\frac{1}{2}\end{pmatrix}\begin{pmatrix} a_1 & 1\\ a_2 & 1\\ \vdots & \vdots\\ a_n & 1\end{pmatrix}\right|\\
&=(-2)^n\prod_{i=1}^{n}a_i\left|I_2-\begin{pmatrix} -\frac{n}{2} & -\frac{1}{2}\sum_{i=1}^{n}\frac{1}{a_i}\\ -\frac{1}{2}\sum_{i=1}^{n}a_i & -\frac{n}{2}\end{pmatrix}\right|=(-2)^n\prod_{i=1}^{n}a_i\begin{vmatrix} \frac{n+2}{2} & \frac{1}{2}\sum_{i=1}^{n}\frac{1}{a_i}\\ \frac{1}{2}\sum_{i=1}^{n}a_i & \frac{n+2}{2}\end{vmatrix}\\
&=(-2)^{n-2}\prod_{i=1}^{n}a_i\left[(n+2)^2-\left(\sum_{i=1}^{n}a_i\right)\left(\sum_{i=1}^{n}\frac{1}{a_i}\right)\right]\\
&=(-2)^{n-2}\prod_{i=1}^{n}a_i(n-2)^2-(-2)^{n-2}\sum_{i=1}^{n}a_i\sum_{j=1}^{n}\prod_{k\neq j}a_k.
\end{aligned}
$$

(ii) 当存在 $a_i=0$, 其中 $i\in 1,2,\cdot,ns$ 时, 不妨设只有 $a_{i_1},a_{i_2},\cdots,a_{i_m}=0,i_1,i_2,\cdots,i_m\in 1,2,\cdots,n$, 则可将 $|A|$ 看作关于 $a_{i_1},a_{i_2},\cdots,a_{i_m}$ 连续的多元多项式函数 $g(a_{i_1},a_{i_2},\cdots,a_{i_m})$, 于是由 g 的连续性可得

$$
g(0,0,\cdots,0)=\lim_{(a_{i_1},a_{i_2},\cdots,a_{i_m})\to(0,0,\cdots,0)}g(a_{i_1},a_{i_2},\cdots,a_{i_m})
$$

$$= \lim_{(a_{i_1},a_{i_2},\cdots,a_{i_m})\to(0,0,\cdots,0)} \left[(-2)^{n-2}\prod_{i=1}^{n} a_i\,(n-2)^2 - (-2)^{n-2}\sum_{i=1}^{n} a_i \sum_{j=1}^{n}\prod_{k\neq j} a_k\right] = 0.$$

即由行列式的连续性可知

$$|A| = (-2)^{n-2}\prod_{i=1}^{n} a_i\,(n-2)^2 - (-2)^{n-2}\sum_{i=1}^{n} a_i \sum_{j=1}^{n}\prod_{k\neq j} a_k.$$

对某些 a_i 为 0 时也成立.

□

结论 对角矩阵行列式的子式和余子式:

设 $|\boldsymbol{A}| = \begin{vmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{vmatrix}$, 则其 k 阶子式 $\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 除 k 阶主子式 $\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}$ 外都为零, 其中 $k = 1, 2, \cdots, n$.

记 $\widehat{\boldsymbol{A}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 为 $\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 的代数余子式 $(n-k$ 阶). 于是 $\widehat{\boldsymbol{A}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 除 $\widehat{\boldsymbol{A}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}$ 外也都为零, 其中 $k = 1, 2, \cdots, n$.

并且

$$\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} = a_{i_1}a_{i_2}\cdots a_{i_k},$$

$$\widehat{\boldsymbol{A}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} = a_1\cdots\hat{a}_{i_1}\cdots\hat{a}_{i_2}\cdots\hat{a}_{i_k}\cdots a_n$$

其中 $k = 1, 2, \cdots, n$.

命题 1.8 (Cauchy 行列式)

证明:

$$|\boldsymbol{A}| = \begin{vmatrix} (a_1+b_1)^{-1} & (a_1+b_2)^{-1} & \cdots & (a_1+b_n)^{-1} \\ (a_2+b_1)^{-1} & (a_2+b_2)^{-1} & \cdots & (a_2+b_n)^{-1} \\ \vdots & \vdots & & \vdots \\ (a_n+b_1)^{-1} & (a_n+b_2)^{-1} & \cdots & (a_n+b_n)^{-1} \end{vmatrix} = \frac{\prod\limits_{1\leq i<j\leq n}(a_j-a_i)(b_j-b_i)}{\prod\limits_{1\leqslant i<j\leqslant m}(a_i+b_j)}.$$

笔记 需要记忆 Cauchy 行列式的计算方法.

1. 分式分母有公共部分可以作差, 得到的分子会变得相对简便.

2. 行列式内行列做加减一般都是加减同一行 (或列). 但是在循环行列式中, 我们一般采取相邻两行 (或列) 相加减的方法.

证明

$$|\boldsymbol{A}| = \begin{vmatrix} \dfrac{1}{a_1+b_1} & \dfrac{1}{a_1+b_2} & \cdots & \dfrac{1}{a_1+b_n} \\ \dfrac{1}{a_2+b_1} & \dfrac{1}{a_2+b_2} & \cdots & \dfrac{1}{a_2+b_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{1}{a_n+b_1} & \dfrac{1}{a_n+b_2} & \cdots & \dfrac{1}{a_n+b_n} \end{vmatrix}$$

$$\xlongequal[i=n-1,\cdots,1]{-j_n+j_i}\begin{vmatrix} \frac{b_n-b_1}{(a_1+b_1)(a_1+b_n)} & \frac{b_n-b_2}{(a_1+b_2)(a_1+b_n)} & \cdots & \frac{b_n-b_{n-1}}{(a_1+b_{n-1})(a_1+b_n)} & \frac{1}{a_1+b_n} \\ \frac{b_n-b_1}{(a_2+b_1)(a_2+b_n)} & \frac{b_n-b_2}{(a_2+b_2)(a_2+b_n)} & \cdots & \frac{b_n-b_{n-1}}{(a_1+b_{n-1})(a_2+b_n)} & \frac{1}{a_2+b_n} \\ \vdots & \vdots & & \vdots & \vdots \\ \frac{b_n-b_1}{(a_n+b_1)(a_n+b_n)} & \frac{b_n-b_2}{(a_n+b_2)(a_n+b_n)} & \cdots & \frac{b_n-b_{n-1}}{(a_1+b_{n-1})(a_n+b_n)} & \frac{1}{a_n+b_n} \end{vmatrix}$$

$$=\frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)}\begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_{n-1}} & 1 \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_{n-1}} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \frac{1}{a_n+b_1} & \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_{n-1}} & 1 \end{vmatrix}$$

$$\xlongequal[i=n-1,\cdots,1]{-r_n+r_i}\frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)}\begin{vmatrix} \frac{a_n-a_1}{(a_1+b_1)(a_n+b_1)} & \frac{a_n-a_1}{(a_1+b_2)(a_n+b_2)} & \cdots & \frac{a_n-a_1}{(a_1+b_{n-1})(a_n+b_{n-1})} & 0 \\ \frac{a_n-a_2}{(a_2+b_1)(a_n+b_1)} & \frac{a_n-a_2}{(a_1+b_2)(a_n+b_2)} & \cdots & \frac{a_n-a_2}{(a_1+b_{n-1})(a_n+b_{n-1})} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \frac{a_n-a_{n-1}}{(a_{n-1}+b_1)(a_n+b_1)} & \frac{a_n-a_{n-1}}{(a_{n-1}+b_2)(a_n+b_2)} & \cdots & \frac{a_n-a_{n-1}}{(a_{n-1}+b_{n-1})(a_n+b_{n-1})} & 0 \\ \frac{1}{a_n+b_1} & \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_{n-1}} & 1 \end{vmatrix}$$

$$=\frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)}\cdot\frac{\prod\limits_{i=1}^{n-1}(a_n-a_i)}{\prod\limits_{k=1}^{n-1}(a_n+b_k)}\begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_{n-1}} & 0 \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_{n-1}} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \frac{1}{a_{n-1}+b_1} & \frac{1}{a_{n-1}+b_2} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{vmatrix}$$

$$\xlongequal{\text{按最后一列展开}}\frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)(a_n-a_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)\prod\limits_{k=1}^{n-1}(a_n+b_k)}\begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_{n-1}} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_{n-1}} \\ \vdots & \vdots & & \vdots \\ \frac{1}{a_{n-1}+b_1} & \frac{1}{a_{n-1}+b_2} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} \end{vmatrix}$$

$$=\frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)(a_n-a_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)\prod\limits_{k=1}^{n-1}(a_n+b_k)}\cdot D_{n-1}.$$

不断递推下去即得

$$D_n=\frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)(a_n-a_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)\prod\limits_{k=1}^{n-1}(a_n+b_k)}\cdot D_{n-1}=\frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)(a_n-a_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)\prod\limits_{k=1}^{n-1}(a_n+b_k)}\cdot\frac{\prod\limits_{i=1}^{n-2}(b_{n-1}-b_i)(a_{n-1}-a_i)}{\prod\limits_{j=1}^{n-1}\left(a_j+b_{n-1}\right)\prod\limits_{k=1}^{n-2}(a_{n-1}+b_k)}\cdot D_{n-2}$$

$$=\cdots=\frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)(a_n-a_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)\prod\limits_{k=1}^{n-1}(a_n+b_k)}\cdot\frac{\prod\limits_{i=1}^{n-2}(b_{n-1}-b_i)(a_{n-1}-a_i)}{\prod\limits_{j=1}^{n-1}\left(a_j+b_{n-1}\right)\prod\limits_{k=1}^{n-2}(a_{n-1}+b_k)}\cdots\cdots\frac{\prod\limits_{i=1}^{2}(b_3-b_i)(a_3-a_i)}{\prod\limits_{j=1}^{3}\left(a_j+b_3\right)\prod\limits_{k=1}^{2}(a_3+b_k)}\cdot D_2$$

$$
\begin{aligned}
&= \frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)(a_n-a_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)\prod\limits_{k=1}^{n-1}(a_n+b_k)}\cdot\frac{\prod\limits_{i=1}^{n-2}(b_{n-1}-b_i)(a_{n-1}-a_i)}{\prod\limits_{j=1}^{n-1}\left(a_j+b_{n-1}\right)\prod\limits_{k=1}^{n-2}(a_{n-1}+b_k)}\cdots\frac{\prod\limits_{i=1}^{2}(b_3-b_i)(a_3-a_i)}{\prod\limits_{j=1}^{3}\left(a_j+b_3\right)\prod\limits_{k=1}^{2}(a_3+b_k)}\cdot\frac{(b_2-b_1)(a_2-a_1)}{\prod\limits_{j=1}^{2}\left(a_j+b_2\right)(a_2+b_1)}\cdot D_1\\
&= \frac{\prod\limits_{i=1}^{n-1}(b_n-b_i)(a_n-a_i)}{\prod\limits_{j=1}^{n}\left(a_j+b_n\right)\prod\limits_{k=1}^{n-1}(a_n+b_k)}\cdot\frac{\prod\limits_{i=1}^{n-2}(b_{n-1}-b_i)(a_{n-1}-a_i)}{\prod\limits_{j=1}^{n-1}\left(a_j+b_{n-1}\right)\prod\limits_{k=1}^{n-2}(a_{n-1}+b_k)}\cdots\frac{\prod\limits_{i=1}^{2}(b_3-b_i)(a_3-a_i)}{\prod\limits_{j=1}^{3}\left(a_j+b_3\right)\prod\limits_{k=1}^{2}(a_3+b_k)}\cdot\frac{(b_2-b_1)(a_2-a_1)}{\prod\limits_{j=1}^{2}\left(a_j+b_2\right)(a_2+b_1)}\cdot\frac{1}{a_1+b_1}\\
&= \frac{\prod\limits_{1\le i<j\le n}(a_j-a_i)(b_j-b_i)}{\prod\limits_{1\le i\le j\le n}(a_i+b_j)\prod\limits_{1\le j<i\le n}(a_i+b_j)} = \frac{\prod\limits_{1\le i<j\le n}(a_j-a_i)(b_j-b_i)}{\prod\limits_{1\leqslant i<j\leqslant m}\left(a_i+b_j\right)}.
\end{aligned}
$$

□

例题 1.13 证明:

$$A=\left(\frac{1}{i+j}\right)_{1\leqslant i,j\leqslant n}\in\mathbb{R}^{n\times n}$$

是正定矩阵.

证明 由Cauchy 行列式可知, 对 A 的所有 m 阶顺序主子式, 我们都有

$$
\begin{vmatrix}
(1+1)^{-1} & (1+2)^{-1} & \cdots & (1+m)^{-1}\\
(2+1)^{-1} & (2+2)^{-1} & \cdots & (2+m)^{-1}\\
\vdots & \vdots & & \vdots\\
(m+1)^{-1} & (m+2)^{-1} & \cdots & (m+m)^{-1}
\end{vmatrix}
=\frac{\prod\limits_{1\leqslant i<j\leqslant m}(j-i)^2}{\prod\limits_{1\leqslant i<j\leqslant m}(i+j)}>0.
$$

故 A 是正定矩阵. □

例题 1.14 设 n 阶行列式

$$
A_n=\begin{vmatrix}
a_0+a_1 & a_1 & 0 & 0 & \cdots & 0 & 0\\
a_1 & a_1+a_2 & a_2 & 0 & \cdots & 0 & 0\\
0 & a_2 & a_2+a_3 & a_3 & \cdots & 0 & 0\\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & 0 & 0 & \cdots & a_{n-1} & a_{n-1}+a_n
\end{vmatrix},
$$

求证:

$$A_n=a_0a_1\cdots a_n\left(\frac{1}{a_0}+\frac{1}{a_1}+\cdots+\frac{1}{a_n}\right).$$

笔记 用**数学归纳法**证明与行列式有关的结论.

练习1.1和练习1.2都可同理使用用数学归纳法证明 (对阶数 n 进行归纳即可).

证明 (数学归纳法) 对阶数 n 进行归纳. 当 $n=1,2$ 时, 结论显然成立. 假设阶数小于 n 结论成立.

现证明 n 阶的情形. 注意到

$$
A_n=\begin{vmatrix}
a_0+a_1 & a_1 & 0 & 0 & \cdots & 0 & 0\\
a_1 & a_1+a_2 & a_2 & 0 & \cdots & 0 & 0\\
0 & a_2 & a_2+a_3 & a_3 & \cdots & 0 & 0\\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & 0 & 0 & \cdots & a_{n-1} & a_{n-1}+a_n
\end{vmatrix}=(a_{n-1}+a_n)A_{n-1}-a_{n-1}^2A_{n-2}.
$$

将归纳假设代入上面的式子中得

$$
\begin{aligned}
A_n&=(a_{n-1}+a_n)A_{n-1}-a_{n-1}^2A_{n-2}\\
&=(a_{n-1}+a_n)a_0a_1\cdots a_{n-1}\left(\frac{1}{a_0}+\frac{1}{a_1}+\cdots+\frac{1}{a_{n-1}}\right)-a_{n-1}^2a_0a_1\cdots a_{n-2}\left(\frac{1}{a_0}+\frac{1}{a_1}+\cdots+\frac{1}{a_{n-2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= a_0a_1\cdots a_n\left(\frac{1}{a_0}+\frac{1}{a_1}+\cdots+\frac{1}{a_{n-1}}\right)+a_0a_1\cdots a_{n-2}a_{n-1}^2\frac{1}{a_{n-1}}\\
&= a_0a_1\cdots a_{n-1}\left[a_n\left(\frac{1}{a_0}+\frac{1}{a_1}+\cdots+\frac{1}{a_{n-1}}\right)+1\right]\\
&= a_0a_1\cdots a_{n-1}a_n\left(\frac{1}{a_0}+\frac{1}{a_1}+\cdots+\frac{1}{a_{n-1}}+\frac{1}{a_n}\right).
\end{aligned}
$$

故由数学归纳法可知, 结论对任意正整数 n 都成立. □

例题 1.15 设 $n(n>2)$ 阶行列式 $|A|$ 的所有元素或为 1 或为 -1, 求证:$|A|$ 的绝对值小于等于 $\frac{2}{3}n!$.

解 对阶数 n 进行归纳. 当 $n=3$ 时, 将 $|A|$ 的第一列元素为-1 的行都乘以-1, 再将 $|A|$ 的第一行元素为 1 的列都乘以-1,$|A|$ 的绝对值不改变.

因此不妨设 $|A|=\begin{vmatrix}1&-1&-1\\1&a_0&b_0\\1&c_0&d_0\end{vmatrix}$, 其中$a_0,b_0,c_0,d_0=1$或$-1$.

从而

$$
|A|=\begin{vmatrix}1&-1&-1\\1&a_0&b_0\\1&c_0&d_0\end{vmatrix}\xlongequal[i=2,3]{j_1+j_i}\begin{vmatrix}1&0&0\\1&a&b\\1&c&d\end{vmatrix}, \text{其中} a,b,c,d=0\text{或}2.
$$

于是

$$
abs\,(|A|)=abs\left(\begin{vmatrix}1&0&0\\1&a&b\\1&c&d\end{vmatrix}\right)=abs\,(ad-bc)\leqslant 4=\frac{2}{3}\cdot 3!
$$

假设 n-1 阶时结论成立, 现证 n 阶的情形. 将 $|A|$ 按第一行展开得

$$
|A|=a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n}, \text{其中} a_{1i}=1\text{或}-1\,(i=1,2\cdots,n).
$$

从而由归纳假设可得

$$
\begin{aligned}
abs\,(|A|)=abs\,(a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n})&\leqslant abs\,(A_{11})+abs\,(A_{12})+\cdots+abs\,(A_{1n})\\
&\leqslant \frac{2}{3}(n-1)!+\frac{2}{3}(n-1)!+\cdots+\frac{2}{3}(n-1)!\\
&= n\cdot\frac{2}{3}(n-1)!=\frac{2}{3}n!.
\end{aligned}
$$

故由数学归纳法可知结论对任意正整数都成立. □

命题 1.9 (行列式的求导运算)

设 $f_{ij}(t)$ 是可微函数,

$$
F(t)=\begin{vmatrix}f_{11}(t)&f_{12}(t)&\cdots&f_{1n}(t)\\f_{21}(t)&f_{22}(t)&\cdots&f_{2n}(t)\\\vdots&\vdots&&\vdots\\f_{n1}(t)&f_{n2}(t)&\cdots&f_{nn}(t)\end{vmatrix}
$$

求证:$\frac{d}{dt}F(t)=\sum_{j=1}^{n}F_j(t)$, 其中

$$
F_j(t)=\begin{vmatrix}f_{11}(t)&f_{12}(t)&\cdots&\frac{d}{dt}f_{1j}(t)&\cdots&f_{1n}(t)\\f_{21}(t)&f_{22}(t)&\cdots&\frac{d}{dt}f_{2j}(t)&\cdots&f_{2n}(t)\\\vdots&\vdots&&\vdots&&\vdots\\f_{n1}(t)&f_{n2}(t)&\cdots&\frac{d}{dt}f_{nj}(t)&\cdots&f_{nn}(t)\end{vmatrix}
$$

证明 证法一 (数学归纳法):对阶数 n 进行归纳. 当 $n=1$ 时结论显然成立. 假设 $n-1$ 阶时结论成立, 现证 n 阶的情形.

将 $F(t)$ 按第一列展开得

$$F(t)=f_{11}(t)A_{11}(t)+f_{21}(t)A_{21}(t)+\cdots+f_{n1}(t)A_{n1}(t).$$

其中 $A_{i1}(t)$ 是元素 $f_{i1}(t)$ 的代数余子式.$(i=1,2,\cdots,n)$

从而由归纳假设可得

$$A'_{i1}(t)=\frac{d}{dt}A_{i1}(t)=\sum_{k=2}^{n}A_{i1}^{k}(t), i=1,2,\cdots,n.$$

$$\text{其中}A_{i1}^{k}(t)=\begin{vmatrix} f_{12}(t) & \cdots & \frac{d}{dt}f_{1k}(t) & \cdots & f_{1n}(t)\\ \vdots & & \vdots & & \vdots\\ f_{i-1,2}(t) & \cdots & \frac{d}{dt}f_{i-1,k}(t) & \cdots & f_{i-1,n}(t)\\ f_{i+1,2}(t) & \cdots & \frac{d}{dt}f_{i+1,k}(t) & \cdots & f_{i+1,n}(t)\\ \vdots & & \vdots & & \vdots\\ f_{n2}(t) & \cdots & \frac{d}{dt}f_{nk}(t) & \cdots & f_{nn}(t)\end{vmatrix}, k=2,3,\cdots,n.$$

于是, 我们就有

$$\begin{aligned}\frac{d}{dt}F(t)&=\frac{d}{dt}\left[f_{11}(t)A_{11}(t)+f_{21}(t)A_{21}(t)+\cdots+f_{n1}(t)A_{n1}(t)\right]\\&=f'_{11}(t)A_{11}(t)+f'_{21}(t)A_{21}(t)+\cdots+f'_{n1}(t)A_{n1}(t)+f_{11}(t)A'_{11}(t)+f_{21}(t)A'_{21}(t)+\cdots+f_{n1}(t)A'_{n1}(t)\\&=\sum_{i=1}^{n}f'_{i1}(t)A_{i1}(t)+f_{11}(t)\sum_{k=2}^{n}A_{11}^{k}(t)+f_{21}(t)\sum_{k=2}^{n}A_{21}^{k}(t)+\cdots+f_{n1}(t)\sum_{k=2}^{n}A_{n1}^{k}(t)\\&=\sum_{i=1}^{n}f'_{i1}(t)A_{i1}(t)+\sum_{i=1}^{n}\left(f_{i1}(t)\sum_{k=2}^{n}A_{i1}^{k}(t)\right)\\&=\sum_{i=1}^{n}f'_{i1}(t)A_{i1}(t)+\sum_{i=1}^{n}f_{i1}(t)\left(A_{i1}^{2}+A_{i1}^{3}+\cdots+A_{i1}^{n}\right)\\&=\sum_{i=1}^{n}f'_{i1}(t)A_{i1}(t)+\sum_{i=1}^{n}f_{i1}(t)A_{i1}^{2}+\sum_{i=1}^{n}f_{i1}(t)A_{i1}^{3}+\cdots+\sum_{i=1}^{n}f_{i1}(t)A_{i1}^{n}\\&=F_1(t)+F_2(t)+F_3(t)+\cdots+F_n(t)\\&=\sum_{j=1}^{n}F_j(t).\end{aligned}$$

故由数学归纳法可知结论对任意正整数都成立.

证法二 (行列式的组合定义):由行列式的组合定义可得

$$F(t)=\sum_{1\leq k_1,k_2,\cdots,k_n\leq n}(-1)^{\tau(k_1k_2\cdots k_n)}f_{k_1 1}(t)f_{k_2 2}(t)\cdots f_{k_n n}(t).$$

因此

$$\begin{aligned}\frac{d}{dt}F(t)=&\sum_{1\leq k_1,k_2,\cdots,k_n\leq n}(-1)^{\tau(k_1k_2\cdots k_n)}f_{k_{11}}(t)f_{k_{22}}(t)\cdots f_{k_{nn}}(t)\\&+\sum_{1\leq k_1,k_2,\cdots,k_n\leq n}(-1)^{\tau(k_1k_2\cdots k_n)}f_{k_{11}}(t)f'_{k_{22}}(t)\cdots f_{k_{nn}}(t)\\&+\cdots+\sum_{1\leq k_1,k_2,\cdots,k_n\leq n}(-1)^{\tau(k_1k_2\cdots k_n)}f_{k_{11}}(t)f_{k_{22}}(t)\cdots f'_{k_{nn}}(t)\\=&F_1(t)+F_2(t)+\cdots+F_n(t).\end{aligned}$$

□

1.5 拆分法

命题 1.10 (大拆分法)

设 t 是一个参数,

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

求证:

$$|A(t)|=|A(0)|+t\sum_{i,j=1}^{n}A_{ij},$$

其中 A_{ij} 是 a_{ij} 在 $|A(0)|$ 中的代数余子式.

笔记 大拆分法的想法: **将行列式的每一行/列拆分成两行/列**, 得到

$$|\boldsymbol{A}(t)|=|\boldsymbol{A}(0)|+t\sum_{j=1}^{n}|A_j|.\text{其中}A_j=\begin{array}{c} \begin{matrix} 1 & \cdots & i & \cdots & n \end{matrix} \\ \begin{pmatrix} a_{11} & \cdots & t & \cdots & a_{1n} \\ a_{21} & \cdots & t & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & t & \cdots & a_{nn} \end{pmatrix} \end{array},j=1,2,\cdots,n.$$

大拆分法的关键是**拆分**, 根据行列式的性质将原行列式拆分成 2^n 个行列式.(不一定需要公共的 t). 不仅要熟悉大拆分法的想法还要记住大拆分法的这个命题.

注 大拆分法后续计算不一定要按行/列展开, 拆分的方式一般比较多, 只要拆分的方式方便后续计算即可.

证明 将行列式第一列拆成两列再展开得到

$$|\boldsymbol{A}(t)|=\begin{vmatrix} a_{11} & a_{12}+t & \cdots & a_{1n}+t \\ a_{21} & a_{22}+t & \cdots & a_{2n}+t \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2}+t & \cdots & a_{nn}+t \end{vmatrix}+\begin{vmatrix} t & a_{12}+t & \cdots & a_{1n}+t \\ t & a_{22}+t & \cdots & a_{2n}+t \\ \vdots & \vdots & & \vdots \\ t & a_{n2}+t & \cdots & a_{nn}+t \end{vmatrix}.$$

将上式右边第二个行列式的第一列乘-1 加到后面每一列上, 得到

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

再对上式右边第一个行列式的第二列拆成两列展开, 不断这样做下去就可得到

$$|\boldsymbol{A}(t)|=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}+\begin{vmatrix} t & a_{12} & \cdots & a_{1n} \\ t & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ t & a_{n2} & \cdots & a_{nn} \end{vmatrix}+\cdots+\begin{vmatrix} a_{11} & a_{1n} & \cdots & t \\ a_{21} & a_{2n} & \cdots & t \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{nn} & \cdots & t \end{vmatrix}=|\boldsymbol{A}(0)|+\sum_{j=1}^{n}|A_j|.$$

其中 $A_j = \begin{array}{c} \begin{matrix} 1 & \cdots & i & \cdots & n \end{matrix} \\ \begin{pmatrix} a_{11} & \cdots & t & \cdots & a_{1n} \\ a_{21} & \cdots & t & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & t & \cdots & a_{nn} \end{pmatrix} \end{array}$, $j = 1, 2, \cdots, n$. 将 A_j 按第 j 列展开可得

$$A_j = \begin{vmatrix} a_{11} & \cdots & t & \cdots & a_{1n} \\ a_{21} & \cdots & t & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & t & \cdots & a_{nn} \end{vmatrix} = t\left(A_{1j} + A_{2j} + \cdots + A_{nj}\right) = t\sum_{i=1}^{n} A_{ij}.$$

从而

$$|\boldsymbol{A}(t)| = |\boldsymbol{A}(0)| + \sum_{i=1}^{n} A_i = |\boldsymbol{A}(0)| + t\sum_{i=1}^{n}\sum_{i=1}^{n} A_{ij} = |\boldsymbol{A}(0)| + t\sum_{i,j=1}^{n} A_{ij}.$$

□

推论 1.3 (推广的大拆分法)

设

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

则

$$|A(t_1, t_2, \cdots, t_n)| = \begin{vmatrix} a_{11}+t_1 & a_{12}+t_2 & \cdots & a_{1n}+t_n \\ a_{21}+t_1 & a_{22}+t_2 & \cdots & a_{2n}+t_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+t_1 & a_{n2}+t_2 & \cdots & a_{nn}+t_n \end{vmatrix} = |A| + \sum_{j=1}^{n}\left(t_j\sum_{i=1}^{n} A_{ij}\right).$$

♡

笔记 记忆这种推广的大拆分法的想法 (即**将行列式的每一行/列拆分成两行/列**).

这里推广的大拆分法的关键也是**要找到合适的** $t_1, t_2, \cdots, t_n$ 进行拆分将原行列式拆分成更好处理的形式.

注 大拆分法后续计算不一定要按行/列展开, 拆分的方式一般比较多, 只要拆分的方式方便后续计算即可.

证明 运用大拆分法的证明方法不难得到. □

命题 1.11 (小拆分法)

设

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

并且 a_{in} 可以拆分成 $b_{in} + c_{in}$, $i = 1, 2, \cdots, n$.

则

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

笔记 记忆小拆分法的想法 (即**拆边列/行, 再展开得到递推式**).

注 若已知的拆分不是最后一列而是其他的某一行或某一列, 则可以通过倒排、旋转、翻转、两行或两列对换的方法将这一行或一列变成最后一列, 再按照上述方法进行拆分即可.

小拆分法后续计算也不一定要按行/列展开, 拆分的方式一般比较多, 只要拆分的方式方便后续计算即可.

证明 由行列式的性质可直接得到结论. □

例题 1.16 计算 n 阶行列式:

$$|\boldsymbol{A}|=\begin{vmatrix}a&b&\cdots&b\\b&a&\cdots&b\\\vdots&\vdots&&\vdots\\b&b&\cdots&a\end{vmatrix}.$$

解 解法一 (大拆分法): 注意到

$$|\boldsymbol{A}|=\begin{vmatrix}a&b&\cdots&b\\b&a&\cdots&b\\\vdots&\vdots&&\vdots\\b&b&\cdots&a\end{vmatrix}=\begin{vmatrix}b+(a-b)&b+0&\cdots&b+0\\b+0&b+(a-b)&\cdots&b+0\\\vdots&\vdots&&\vdots\\b+0&b+0&\cdots&b+(a-b)\end{vmatrix}$$

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

其中 A_i 是第 i 行元素全为 b, 主对角元素除了 (i,i) 元外都为 $a-b$, 其他元素都为 0 的 n 阶行列式.

又因为

$$A_i=\begin{array}{c}1\\\vdots\\i\\\vdots\\n\end{array}\begin{vmatrix}a-b&&&&\\&\ddots&&&\\b&\cdots&b&\cdots&b\\&&&\ddots&\\&&&&a-b\end{vmatrix}=b\,(a-b)^{n-1},i=1,2,\cdots,n.$$

所以

$$|\boldsymbol{A}|=(a-b)^n+\sum_{i=1}^{n}A_i=(a-b)^n+nb\,(a-b)^{n-1}=[a+(n-1)\,b]\,(a-b)^{n-1}.$$

解法二 (小拆分法): 记原行列式为 D_n, 其中 n 为原行列式的阶数. 则将原行列式按第一列拆开为两个行列式得

$$D_n=\begin{vmatrix}a&b&\cdots&b\\b&a&\cdots&b\\\vdots&\vdots&&\vdots\\b&b&\cdots&a\end{vmatrix}=\begin{vmatrix}b+(a-b)&b&\cdots&b\\b+0&a&\cdots&b\\\vdots&\vdots&&\vdots\\b+0&b&\cdots&a\end{vmatrix}=\begin{vmatrix}b&b&\cdots&b\\b&a&\cdots&b\\\vdots&\vdots&&\vdots\\b&b&\cdots&a\end{vmatrix}+\begin{vmatrix}a-b&b&\cdots&b\\0&a&\cdots&b\\\vdots&\vdots&&\vdots\\0&b&\cdots&a\end{vmatrix}$$

$$= \begin{vmatrix} b & b & \cdots & b \\ 0 & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a-b \end{vmatrix} + (a-b) D_{n-1} = b(a-b)^{n-1} + (a-b) D_{n-1}.(n \geq 2)$$

从而由上式递推可得

$$\begin{aligned} D_n &= b(a-b)^{n-1} + (a-b) D_{n-1} \\ &= b(a-b)^{n-1} + (a-b)\left[b(a-b)^{n-2} + (a-b) D_{n-2}\right] = 2b(a-b)^{n-1} + (a-b)^2 D_{n-2} \\ &= \cdots = (n-1) b (a-b)^{n-1} + (a-b)^{n-1} D_1 \\ &= (n-1) b (a-b)^{n-1} + (a-b)^{n-1} a \\ &= [a + (n-1) b] (a-b)^{n-1}. \end{aligned}$$

解法三 (求和法):

$$|A| = \begin{vmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ b & b & \cdots & a \end{vmatrix} \xlongequal[i=2,3,\cdots,n]{j_i + j_1} \begin{vmatrix} a+(n-1)b & b & \cdots & b \\ a+(n-1)b & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ a+(n-1)b & b & \cdots & a \end{vmatrix} = [a+(n-1)b] \begin{vmatrix} 1 & b & \cdots & b \\ 1 & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ 1 & b & \cdots & a \end{vmatrix}$$

$$\xlongequal[i=2,3,\cdots,n]{-r_1 + r_i} [a+(n-1)b] \begin{vmatrix} 1 & b & \cdots & b \\ 0 & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a-b \end{vmatrix} = [a+(n-1)b](a-b)^{n-1}.$$

解法四 (" 爪" 型行列式的推广):

$$|A| = \begin{vmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ b & b & \cdots & a \end{vmatrix} \xlongequal[i=2,3,\cdots,n]{-r_1 + r_i} \begin{vmatrix} a & b & \cdots & b \\ b-a & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ b-a & 0 & \cdots & a-b \end{vmatrix}$$

$$\xlongequal[i=2,3,\cdots,n]{-j_i + j_1} \begin{vmatrix} a-(n-1)b & b & \cdots & b \\ 0 & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a-b \end{vmatrix} = [a-(n-1)b](a-b)^{n-1}.$$

□

例题 1.17 计算 n 阶行列式:

$$|A| = \begin{vmatrix} a & b & \cdots & b \\ c & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ c & c & \cdots & a \end{vmatrix}.$$

解 解法一 (大拆分法): 令

$$|A(t)| = \begin{vmatrix} a+t & b+t & \cdots & b+t \\ c+t & a+t & \cdots & b+t \\ \vdots & \vdots & & \vdots \\ c+t & c+t & \cdots & a+t \end{vmatrix} = |A| + tu, u = \sum_{i,j=1}^{n} A_{ij}.$$

当 $t=-b$ 时, 可得

$$|\boldsymbol{A}(-b)|=\begin{vmatrix} a-b & 0 & \cdots & 0 \\ c-b & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ c-b & c-b & \cdots & a-b \end{vmatrix}=|\boldsymbol{A}|-bu=(a-b)^n.$$

当 $t=-c$ 时, 可得

$$|\boldsymbol{A}(-c)|=\begin{vmatrix} a-c & b-c & \cdots & b-c \\ 0 & a-c & \cdots & b-c \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a-c \end{vmatrix}=|\boldsymbol{A}|-cu=(a-c)^n.$$

若 $b\neq c$, 则联立上面两式可得

$$|\boldsymbol{A}|=\frac{b\,(a-c)^n-c\,(a-b)^n}{b-c}.$$

若 $b=c$, 则由练习1.16可知

$$|\boldsymbol{A}|=[a+(n-1)\,b]\,(a-b)^{n-1}.$$

解法二 (小拆分法): 记原行列式为 D_n, 其中 n 为原行列式的阶数. 则将原行列式分别按第一行、第一列拆开为两个行列式得

$$D_n=\begin{vmatrix} a & b & \cdots & b \\ c & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ c & c & \cdots & a \end{vmatrix}=\begin{vmatrix} b+(a-b) & b+0 & \cdots & b+0 \\ c & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ c & c & \cdots & a \end{vmatrix}=\begin{vmatrix} b & b & \cdots & b \\ c & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ c & c & \cdots & a \end{vmatrix}+\begin{vmatrix} a-b & 0 & \cdots & 0 \\ c & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ c & c & \cdots & a \end{vmatrix}$$

$$=b\begin{vmatrix} 1 & 1 & \cdots & 1 \\ c & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ c & c & \cdots & a \end{vmatrix}+(a-b)\,D_{n-1}=b\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & a-c & \cdots & b-c \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a-c \end{vmatrix}+(a-b)\,D_{n-1}$$

$$=b\,(a-c)^{n-1}++(a-b)\,D_{n-1}.\,(n\geq 2)$$

$$D_n=\begin{vmatrix} a & b & \cdots & b \\ c & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ c & c & \cdots & a \end{vmatrix}=\begin{vmatrix} c+(a-c) & b & \cdots & b \\ c+0 & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ c+0 & c & \cdots & a \end{vmatrix}=\begin{vmatrix} c & b & \cdots & b \\ c & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ c & c & \cdots & a \end{vmatrix}+\begin{vmatrix} a-c & b & \cdots & b \\ 0 & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ 0 & c & \cdots & a \end{vmatrix}$$

$$=c\begin{vmatrix} 1 & b & \cdots & b \\ 1 & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ 1 & c & \cdots & a \end{vmatrix}+(a-c)\,D_{n-1}=c\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 1 & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & c-b & \cdots & a-b \end{vmatrix}+(a-c)\,D_{n-1}$$

$$=c\,(a-b)^{n-1}++(a-c)\,D_{n-1}.\,(n\geq 2)$$

若 $b\neq c$, 则联立上面两式可得

$$|\boldsymbol{A}|=D_n=\frac{b\,(a-c)^n-c\,(a-b)^n}{b-c}.$$

若 $b=c$, 则由上面式子递推可得

$$\begin{aligned}|\boldsymbol{A}|=D_n&=b\,(a-b)^{n-1}+(a-b)\,D_{n-1}\\&=b\,(a-b)^{n-1}+(a-b)\left[b\,(a-b)^{n-2}+(a-b)\,D_{n-2}\right]=2b\,(a-b)^{n-1}+(a-b)^2\,D_{n-2}\end{aligned}$$

$$
\begin{aligned}
&= \cdots = (n-1)\,b\,(a-b)^{n-1} + (a-b)^{n-1}\,D_1 \\
&= (n-1)\,b\,(a-b)^{n-1} + (a-b)^{n-1}\,a \\
&= [a+(n-1)\,b]\,(a-b)^{n-1}\,.
\end{aligned}
$$

当 $b=c$ 时, 也可以由练习1.16可知

$$|\boldsymbol{A}| = [a+(n-1)\,b]\,(a-b)^{n-1}\,.$$

□

例题 1.18 设 $f_1(x), f_2(x), \cdots, f_n(x)$ 是次数不超过 $n-2$ 的多项式, 求证: 对任意 n 个数 $a_1, a_2, \cdots, a_n$, 均有

$$\begin{vmatrix} f_1(a_1) & f_2(a_1) & \cdots & f_n(a_1) \\ f_1(a_2) & f_2(a_2) & \cdots & f_n(a_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(a_n) & f_2(a_n) & \cdots & f_n(a_n) \end{vmatrix} = 0.$$

证明 证法一 (大拆分法): 因为 $f_k(x)(1\leq k\leq n)$ 的次数不超过 $n-2$, 所以它们都是单项式 $1, x, \cdots, x^{n-2}$ 的线性组合. 将原行列式中每一列的多项式都按这 $n-1$ 个单项式进行拆分, 最后得到至多 $(n-1)!$ 个简单行列式之和, 这些行列式中每一列的多项式只是单项式. 由于每个简单行列式都有 n 列, 根据抽屉原理, 每个简单行列式中至少有两列是共用同一个单项式 (可能相差一个常系数), 于是这两列成比例, 从而所有这样的简单行列式都等于零, 因此原行列式也等于零.

证法二 (多项式根的有限性): 令 $f(x) = \begin{vmatrix} f_1(x) & f_2(a_1) & \cdots & f_n(a_1) \\ f_1(x) & f_2(a_2) & \cdots & f_n(a_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x) & f_2(a_n) & \cdots & f_n(a_n) \end{vmatrix}$, 则将 $f(x)$ 按第一列展开得到

$$f(x) = k_1 f_1(x) + k_2 f_2(x) + \cdots + k_n f_n(x)\,.$$

其中 k_i 为行列式 $f(x)$ 的第 $(i,1)$ 元素的代数余子式,$i=1,2,\cdots,n$.

注意 k_i 与 x 无关, 均为常数. 若 $f(x)$ 不恒为 0, 则又因为 $f_k(x)(1\leq k\leq n)$ 的次数不超过 $n-2$, 所以 $deg f(x)\leq n-2$. 但是, 注意到 $f(a_2)=f(a_3)=\cdots=f(a_n)=0$, 即 $f(x)$ 有 $n-1$ 个根. 于是由余数定理可知,$(x-a_2)\cdots(x-a_n)\,|\,f(x)$. 从而 $n-1 = deg\,(x-a_2)\cdots(x-a_n) \geq deg f(x)$. 这与 $deg f(x)\leq n-2$ 矛盾. 故 $f(x)\equiv 0$, 当然也有 $f(a_1)=0$.

证法三:

设多项式

$$f_k(x) = c_{k,n-2}x^{n-2} + \cdots + c_{k1}x + c_{k0}, 1\leq k\leq n.$$

则有如下的矩阵分解:

$$\begin{pmatrix} f_1(a_1) & f_2(a_1) & \cdots & f_n(a_1) \\ f_1(a_2) & f_2(a_2) & \cdots & f_n(a_2) \\ \vdots & \vdots & & \vdots \\ f_1(a_n) & f_2(a_n) & \cdots & f_n(a_n) \end{pmatrix} = \begin{pmatrix} 1 & a_1 & \cdots & a_1^{n-2} \\ 1 & a_2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \cdots & a_n^{n-2} \end{pmatrix} \begin{pmatrix} c_{10} & c_{20} & \cdots & c_{n0} \\ c_{11} & c_{21} & \cdots & c_{n1} \\ \vdots & \vdots & & \vdots \\ c_{1,n-2} & c_{2,n-2} & \cdots & c_{n,n-2} \end{pmatrix}.$$

注意到上式右边的两个矩阵分别是 $n\times(n-1)$ 和 $(n-1)\times n$ 矩阵, 故由Cauchy - Binet 公式马上得到左边矩阵的行列式值等于零. □

例题 1.19 求下列 n 阶行列式的值:

$$D_n = \begin{vmatrix} 1+a_1^2 & a_1a_2 & \cdots & a_1a_n \\ a_2a_1 & 1+a_2^2 & \cdots & a_2a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_na_1 & a_na_2 & \cdots & 1+a_n^2 \end{vmatrix}$$

笔记 本题行列式每行或每列求和后得到的结果不具备明显的规律性, 故不适合使用求和法.

本题行列式难以找到合适的 t 对其进行大拆分, 故也不适合使用大拆分法.(并且因为难以找到合适的 t_i, 所以推广的大拆分也不行)

解 (小拆分法) 将 D_n 最后一列拆成两列得

$$D_n=\begin{vmatrix}1+a_1^2 & a_1a_2 & \cdots & a_1a_n\\ a_2a_1 & 1+a_2^2 & \cdots & a_2a_n\\ \vdots & \vdots & \ddots & \vdots\\ a_na_1 & a_na_2 & \cdots & 1+a_n^2\end{vmatrix}=\begin{vmatrix}1+a_1^2 & a_1a_2 & \cdots & a_1a_n\\ a_2a_1 & 1+a_2^2 & \cdots & a_2a_n\\ \vdots & \vdots & \ddots & \vdots\\ a_na_1 & a_na_2 & \cdots & a_n^2\end{vmatrix}+\begin{vmatrix}1+a_1^2 & a_1a_2 & \cdots & 0\\ a_2a_1 & 1+a_2^2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ a_na_1 & a_na_2 & \cdots & 1\end{vmatrix}$$

$$=\begin{vmatrix}1+a_1^2 & a_1a_2 & \cdots & a_1a_n\\ a_2a_1 & 1+a_2^2 & \cdots & a_2a_n\\ \vdots & \vdots & \ddots & \vdots\\ a_na_1 & a_na_2 & \cdots & a_n^2\end{vmatrix}+D_{n-1}.$$

若 $a_n\neq 0$, 则由上式可得

$$D_n=a_n\begin{vmatrix}1+a_1^2 & a_1a_2 & \cdots & a_1\\ a_2a_1 & 1+a_2^2 & \cdots & a_2\\ \vdots & \vdots & \ddots & \vdots\\ a_na_1 & a_na_2 & \cdots & a_n\end{vmatrix}+D_{n-1}\xlongequal[i=1,2,\cdots,n]{\text{对第一个行列式:}-a_ij_n+j_i}a_n\begin{vmatrix}1 & 0 & \cdots & a_1\\ 0 & 1 & \cdots & a_2\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & a_n\end{vmatrix}+D_{n-1}=a_n^2+D_{n-1}.\,(n\geq 2)$$

若 $a_n=0$, 则上面第一个行列式等于 0, 进而 $D_n=D_{n-1}(n\geq 0)$. 仍然满足上述递推式.

从而由上式递推可得

$$D_n=a_n^2+D_{n-1}=a_n^2+\left(a_{n-1}^2+D_{n-2}\right)=\cdots=\sum_{i=2}^{n}a_i^2+D_1=1+\sum_{i=1}^{n}a_i^2.$$

□

1.6 Vandermode 行列式

本节我们用 $V_n(x_1,x_2,\cdots,x_n)$ 表示 n 阶 Vandermode 行列式.

定义 1.2

对 $1\leqslant i\leqslant n, V_n^{(i)}(x_1,x_2,\cdots,x_n)$ 表示删除 $V_n(x_1,x_2,\cdots,x_n)$ 的第 i 行 $(x_1^{i-1},x_2^{i-1},\cdots,x_n^{i-1})$ 之后新添第 n 行 $(x_1^n,x_2^n,\cdots,x_n^n)$ 所得 n 阶行列式. ♣

定义 1.3

$\Delta_n(x_1,x_2,\cdots,x_n)$ 表示将 $V_n(x_1,x_2,\cdots,x_n)$ 的第 n 行换成 $(x_1^{n+1},x_2^{n+1},\cdots,x_n^{n+1})$ 所得 n 阶行列式. ♣

例题 1.20 设初等对称多项式

$$\sigma_j=\sum_{1\leqslant k_1<k_2<\cdots<k_j\leqslant n}x_{k_1}x_{k_2}\cdots x_{k_j},\ j=1,2,\cdots,n, \tag{1.6}$$

我们有

$$V_n^{(i)}(x_1,x_2,\cdots,x_n)=\sigma_{n-i+1}V_n(x_1,x_2,\cdots,x_n),\ i=1,2,\cdots,n. \tag{1.7}$$

证明 (加边法) 不妨设 $x_i, 1 \leqslant i \leqslant n$ 互不相同. 设

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

由行列式性质我们知道 D_n 是 n 次多项式且有 n 个根 $-x_1, -x_2, \cdots, -x_n$. 于是我们有

$$D_n(x) = c(x+x_1)(x+x_2)\cdots(x+x_n). \tag{1.8}$$

把 $D_n(x)$ 按第一列展开得

$$D_n(x) = \sum_{i=1}^{n} V_n^{(i)}(x_1, x_2, \cdots, x_n)x^{i-1} + V_n(x_1, x_2, \cdots, x_n)x^n. \tag{1.9}$$

于是比较(1.8)式和(1.9)式最高次项系数, 我们有 $c = V_n(x_1, x_2, \cdots, x_n)$. 定义 $\sigma_0 = 1$, 利用根和系数的关系 (Vieta 定理), 结合(1.8)式和(1.9)式得

$$D_n(x) = \sum_{i=1}^{n+1} \sigma_{n-i+1} V_n(x_1, x_2, \cdots, x_n)x^{i-1} = \sum_{i=1}^{n} V_n^{(i)}(x_1, x_2, \cdots, x_n)x^{i-1} + V_n(x_1, x_2, \cdots, x_n)x^n,$$

比较上式等号两边 $x^i(1 \leq i \leq n)$ 的系数就能得到(1.7). □

例题 1.21 证明:

$$\Delta_n(x_1, x_2, \cdots, x_n) = \left(\sum_{k=1}^{n} x_k^2 + \sum_{1 \leq i < j \leq n} x_i x_j\right) V_n(x_1, x_2, \cdots, x_n) \tag{1.10}$$

证明 不妨设 $x_i, 1 \leq i \leq n$ 互不相同. 设 $n+1$ 次多项式

$$P_{n+1}(x) \triangleq \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ -x & x_1 & x_2 & \cdots & x_n \\ (-x)^2 & x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (-x)^{n-1} & x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \\ (-x)^{n+1} & x_1^{n+1} & x_2^{n+1} & \cdots & x_n^{n+1} \end{vmatrix}$$

注意到有 n 个根 $-x_1, -x_2, \cdots, -x_n$. 我们用 $-x_{n+1}$ 表示 P_{n+1} 第 $n+1$ 个根. 于是我们有

$$P_{n+1}(x) = c(x+x_1)(x+x_2)\cdots(x+x_n)(x+x_{n+1}). \tag{1.11}$$

将 $P_{n+1}(x)$ 按第一列展开得

$$P_{n+1}(x) = -V_n(x_1, x_2, \cdots, x_n)x^{n+1} + \Delta_n(x_1, x_2, \cdots, x_n)x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_0 \tag{1.12}$$

其中 $a_{n-2}, \cdots, a_0$ 是某些与 x_j 有关的 n 阶行列式. 比较(1.11)和(1.12)式的系数可知 $c = -V_n(x_1, x_2, \cdots, x_n)$. 于是结合(1.11)式, 并利用 Vieta 定理得

$$P_{n+1}(x) = -V_n(x_1, x_2, \cdots, x_n)(x^{n+1} + \delta_1 x^n + \delta_2 x^{n-1} + \cdots + \delta_{n-1}) \tag{1.13}$$

这里 δ_j 类似(1.6)式定义是 $x_1, x_2, \cdots, x_n, x_{n+1}$ 的初等对称多项式. 比较(1.12)(1.13)式的 x^{n-1} 系数可得 $\Delta_n(x_1, x_2, \cdots, x_n) = -\delta_2 V_n(x_1, x_2, \cdots, x_n)$. 因为 $P_{n+1}(x)$ 没有 x^n 的项, 所以

$$\delta_1 = x_1 + x_2 + \cdots + x_{n+1} = 0 \Rightarrow x_{n+1} = -(x_1 + x_2 + \cdots + x_n).$$

从而

$$\delta_2 = \sum_{1 \leq i < j \leq n+1} x_i x_j = \sum_{1 \leq i < j \leq n} x_i x_j + x_{n+1} \sum_{i=1}^{n} x_i$$

$$
\begin{aligned}
&= \sum_{1\le i<j\le n} x_i x_j - (x_1+x_2+\cdots+x_n)\sum_{i=1}^{n} x_i \\
&= \sum_{1\le i<j\le n} x_i x_j - \left(\sum_{i=1}^{n} x_i\right)^2 = -\sum_{i=1}^{n} x_i^2 - \sum_{1\le i<j\le n} x_i x_j .
\end{aligned}
$$

现在就有(1.10)成立. □

命题 1.12

设 $A=(a_{ij})_{n\times n}, f_i(x)=a_{i1}+a_{i2}x+\cdots+a_{in}x^{n-1}(i=1,2,\cdots,n)$, 证明: 对任何复数 $x_1,x_2,\cdots,x_n$, 都有

$$
\begin{vmatrix}
f_1(x_1) & f_1(x_2) & \cdots & f_1(x_n)\\
f_2(x_1) & f_2(x_2) & \cdots & f_2(x_n)\\
\vdots & \vdots & \ddots & \vdots\\
f_n(x_1) & f_n(x_2) & \cdots & f_n(x_n)
\end{vmatrix} = |A|\cdot V_n(x_1,x_2,\cdots,x_n)
$$

这里 $V_n(x_1,x_2,\cdots,x_n)$ 表示 $x_1,x_2,\cdots,x_n$ 的 Vandermonde 行列式.

笔记 关键是利用命题 2.35.

证明 直接由矩阵乘法观察知显然. □

推论 1.4

设 $f_k(x)=x^k+a_{k1}x^{k-1}+a_{k2}x^{k-2}+\cdots+a_{kk}$, 求下列行列式的值:

$$
\begin{vmatrix}
1 & f_1(x_1) & f_2(x_1) & \cdots & f_{n-1}(x_1)\\
1 & f_1(x_2) & f_2(x_2) & \cdots & f_{n-1}(x_2)\\
\vdots & \vdots & \vdots & & \vdots\\
1 & f_1(x_n) & f_2(x_n) & \cdots & f_{n-1}(x_n)
\end{vmatrix}.
$$

笔记 知道这类行列式化简的操作即可. 以后这种行列式化简操作不再作额外说明.

注 也可以由命题 1.12直接得到.

解 解法一: 利用行列式的性质可得

$$
\begin{vmatrix}
1 & f_1(x_1) & f_2(x_1) & \cdots & f_{n-1}(x_1)\\
1 & f_1(x_2) & f_2(x_2) & \cdots & f_{n-1}(x_2)\\
\vdots & \vdots & \vdots & & \vdots\\
1 & f_1(x_n) & f_2(x_n) & \cdots & f_{n-1}(x_n)
\end{vmatrix}
=
\begin{vmatrix}
1 & x_1+a_{11} & x_1^2+a_{21}x_1+a_{22} & \cdots & x_1^{n-1}+a_{n-1,1}x_1^{n-2}+\cdots+a_{n-1,n-2}x_1+a_{n-1,n-1}\\
1 & x_2+a_{11} & x_2^2+a_{21}x_2+a_{22} & \cdots & x_2^{n-1}+a_{n-1,1}x_2^{n-2}+\cdots+a_{n-1,n-2}x_2+a_{n-1,n-1}\\
\vdots & \vdots & \vdots & & \vdots\\
1 & x_n+a_{11} & x_n^2+a_{21}x_n+a_{22} & \cdots & x_n^{n-1}+a_{n-1,1}x_n^{n-2}+\cdots+a_{n-1,n-2}x_n+a_{n-1,n-1}
\end{vmatrix}
$$

$$
\xlongequal[\substack{\cdots\\ -a_{i,i-(n-3)}j_{n-2}+j_{i+1},\, i=n-2,n-1\\ -a_{n-1,1}j_{n-1}+j_n}]{\substack{-a_{ii}j_1+j_{i+1},\, i=1,2,\cdots n-1\\ -a_{i,i-1}j_2+j_{i+1},\, i=2,3,\cdots,n-1}}
\begin{vmatrix}
1 & x_1 & x_1^2 & \cdots & x_1^{n-1}\\
1 & x_2 & x_2^2 & \cdots & x_2^{n-1}\\
\vdots & \vdots & \vdots & & \vdots\\
1 & x_n & x_n^2 & \cdots & x_n^{n-1}
\end{vmatrix}
= \prod_{1\le i<j\le n}(x_j-x_i).
$$

解法二:由命题 1.12可得

$$
\begin{vmatrix}
1 & f_1(x_1) & f_2(x_1) & \cdots & f_{n-1}(x_1)\\
1 & f_1(x_2) & f_2(x_2) & \cdots & f_{n-1}(x_2)\\
\vdots & \vdots & \vdots & & \vdots\\
1 & f_1(x_n) & f_2(x_n) & \cdots & f_{n-1}(x_n)
\end{vmatrix}
=
\begin{vmatrix}
1 & 1 & \cdots & 1\\
f_1(x_1) & f_1(x_2) & \cdots & f_1(x_n)\\
f_2(x_1) & f_2(x_2) & \cdots & f_2(x_n)\\
\vdots & \vdots & & \vdots\\
f_{n-1}(x_1) & f_{n-1}(x_2) & \cdots & f_{n-1}(x_n)
\end{vmatrix}
=
\begin{vmatrix}
1 & 1 & \cdots \\
x_1+a_{11} & x_2+a_{11} & \cdots \\
x_1^2+a_{21}x_1+a_{22} & x_2^2+a_{21}x_2+a_{22} & \cdots \\
\vdots & \vdots & \\
x_1^{n-1}+\cdots+a_{n-1,n-1} & x_2^{n-1}+\cdots+a_{n-1,n-1} & \cdots \; x
\end{vmatrix}
$$

$$= V_n(x_1,x_2,\cdots,x_n)\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{vmatrix} = V_n(x_1,x_2,\cdots,x_n) = \prod_{1\le i<j\le n}(x_j - x_i).$$

□

例题 1.22 计算

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1+1 & x_2+1 & x_3+1 & \cdots & x_n+1 \\ x_1^2+x_1 & x_2^2+x_2 & x_3^2+x_3 & \cdots & x_n^2+x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1}+x_1^{n-2} & x_2^{n-1}+x_2^{n-2} & x_3^{n-1}+x_3^{n-2} & \cdots & x_n^{n-1}+x_n^{n-2} \end{vmatrix}$$

证明 由命题 1.12我们知道

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1+1 & x_2+1 & x_3+1 & \cdots & x_n+1 \\ x_1^2+x_1 & x_2^2+x_2 & x_3^2+x_3 & \cdots & x_n^2+x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1}+x_1^{n-2} & x_2^{n-1}+x_2^{n-2} & x_3^{n-1}+x_3^{n-2} & \cdots & x_n^{n-1}+x_n^{n-2} \end{vmatrix} = V_n(x_1,x_2,\cdots,x_n)\cdot\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1 \end{vmatrix}$$

$$= \prod_{1\le i<j\le n}(x_j - x_i)\cdot\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1 \end{vmatrix} = \prod_{1\le i<j\le n}(x_j - x_i).$$

□

命题 1.13

计算下列行列式的值:

$$|\boldsymbol{A}| = \begin{vmatrix} a_1^{n-1} & a_1^{n-2}b_1 & \cdots & a_1b_1^{n-2} & b_1^{n-1} \\ a_2^{n-1} & a_2^{n-2}b_2 & \cdots & a_2b_2^{n-2} & b_2^{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_n^{n-1} & a_n^{n-2}b_n & \cdots & a_nb_n^{n-2} & b_n^{n-1} \end{vmatrix}.$$

解 若所有的 $a_i(i=1,2,\cdots,n)$ 都不为 0, 则有

$$|\boldsymbol{A}| = \begin{vmatrix} a_1^{n-1} & a_1^{n-2}b_1 & \cdots & a_1b_1^{n-2} & b_1^{n-1} \\ a_2^{n-1} & a_2^{n-2}b_2 & \cdots & a_2b_2^{n-2} & b_2^{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_n^{n-1} & a_n^{n-2}b_n & \cdots & a_nb_n^{n-2} & b_n^{n-1} \end{vmatrix} = \prod_{i=1}^{n} a_i^{n-1}\begin{vmatrix} 1 & \frac{b_1}{a_1} & \cdots & \frac{b_1^{n-2}}{a_1^{n-2}} & \frac{b_1^{n-1}}{a_1^{n-1}} \\ 1 & \frac{b_2}{a_2} & \cdots & \frac{b_2^{n-2}}{a_2^{n-2}} & \frac{b_2^{n-1}}{a_2^{n-1}} \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \frac{b_n}{a_n} & \cdots & \frac{b_n^{n-2}}{a_n^{n-2}} & \frac{b_n^{n-2}}{a_n^{n-2}} \end{vmatrix}$$

$$=\prod_{i=1}^{n}a_i^{n-1}\prod_{1\le i<j\le n}\left(\frac{b_j}{a_j}-\frac{b_i}{a_i}\right)=\prod_{i=1}^{n}a_i^{n-1}\prod_{1\le i<j\le n}\frac{a_ib_j-a_jb_i}{a_ja_i}=\prod_{1\le i<j\le n}(a_ib_j-a_jb_i).$$

若只有一个 a_i 为 0, 则将原行列式按第 i 行展开得到具有相同类型的 $n-1$ 阶行列式

$$|\boldsymbol{A}|=\begin{vmatrix}a_1^{n-1}&a_1^{n-2}b_1&\cdots&a_1b_1^{n-2}&b_1^{n-1}\\a_2^{n-1}&a_2^{n-2}b_2&\cdots&a_2b_2^{n-2}&b_2^{n-1}\\\vdots&\vdots&&\vdots&\vdots\\a_i^{n-1}&a_i^{n-2}b_i&\cdots&a_ib_i^{n-2}&b_i^{n-1}\\\vdots&\vdots&&\vdots&\vdots\\a_n^{n-1}&a_n^{n-2}b_n&\cdots&a_nb_n^{n-2}&b_n^{n-1}\end{vmatrix}=\begin{vmatrix}a_1^{n-1}&a_1^{n-2}b_1&\cdots&a_1b_1^{n-2}&b_1^{n-1}\\a_2^{n-1}&a_2^{n-2}b_2&\cdots&a_2b_2^{n-2}&b_2^{n-1}\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&0&b_i^{n-1}\\\vdots&\vdots&&\vdots&\vdots\\a_n^{n-1}&a_n^{n-2}b_n&\cdots&a_nb_n^{n-2}&b_n^{n-1}\end{vmatrix}$$

$$\xlongequal{\text{按第}i\text{行展开}}(-1)^{n+i}b_i^{n-1}\begin{vmatrix}a_1^{n-1}&a_1^{n-2}b_1&\cdots&a_1b_1^{n-2}\\a_2^{n-1}&a_2^{n-2}b_2&\cdots&a_2b_2^{n-2}\\\vdots&\vdots&&\vdots\\a_{i-1}^{n-1}&a_{i-1}^{n-2}b_{i-1}&\cdots&a_{i-1}b_{i-1}^{n-2}\\a_{i+!}^{n-1}&a_{i+1}^{n-2}b_{i+1}&\cdots&a_{i+1}b_{i+1}^{n-2}\\\vdots&\vdots&&\vdots\\a_n^{n-1}&a_n^{n-2}b_n&\cdots&a_nb_n^{n-2}\end{vmatrix}.$$

此时同理可得

$$|\boldsymbol{A}|=(-1)^{n+i}b_i^{n-1}\begin{vmatrix}a_1^{n-1}&a_1^{n-2}b_1&\cdots&a_1b_1^{n-2}\\a_2^{n-1}&a_2^{n-2}b_2&\cdots&a_2b_2^{n-2}\\\vdots&\vdots&&\vdots\\a_{i-1}^{n-1}&a_{i-1}^{n-2}b_{i-1}&\cdots&a_{i-1}b_{i-1}^{n-2}\\a_{i+1}^{n-1}&a_{i+1}^{n-2}b_{i+1}&\cdots&a_{i+1}b_{i+1}^{n-2}\\\vdots&\vdots&&\vdots\\a_n^{n-1}&a_n^{n-2}b_n&\cdots&a_nb_n^{n-2}\end{vmatrix}=(-1)^{n+i}b_i^{n-1}\prod_{\substack{1\le k\le n\\k\ne i}}a_k^{n-1}\begin{vmatrix}1&\frac{b_1}{a_1}&\cdots&\frac{b_1^{n-2}}{a_1^{n-2}}\\1&\frac{b_2}{a_2}&\cdots&\frac{b_2^{n-2}}{a_2^{n-2}}\\\vdots&\vdots&&\vdots\\1&\frac{b_{i-1}}{a_{i-1}}&\cdots&\frac{b_{i-1}^{n-2}}{a_{i-1}^{n-2}}\\1&\frac{b_{i+1}}{a_{i+1}}&\cdots&\frac{b_{i+1}^{n-2}}{a_{i+1}^{n-2}}\\\vdots&\vdots&&\vdots\\1&\frac{b_n}{a_n}&\cdots&\frac{b_n^{n-2}}{a_n^{n-2}}\end{vmatrix}$$

$$=(-1)^{n+i}b_i^{n-1}\prod_{\substack{1\le k\le n\\k\ne i}}a_k^{n-1}\prod_{\substack{1\le k<l\le n\\k,l\ne i}}\left(\frac{b_l}{a_l}-\frac{b_k}{a_k}\right)=(-1)^{n+i}b_i^{n-1}\prod_{\substack{1\le k\le n\\k\ne i}}a_k^{n-1}\prod_{\substack{1\le k<l\le n\\k,l\ne i}}\frac{a_kb_l-a_lb_k}{a_ka_l}$$

$$=(-1)^{n+i}b_i^{n-1}\prod_{\substack{1\le k\le n\\k\ne i}}a_k\cdot\prod_{\substack{1\le k<l\le n\\k,l\ne i}}(a_kb_l-a_lb_k)=(-1)^{n-i}b_i^{n-1}\prod_{\substack{1\le k\le n\\k\ne i}}a_k\cdot\prod_{\substack{1\le k<l\le n\\k,l\ne i}}(a_kb_l-a_lb_k)$$

$$=\prod_{1\le k<i}a_kb_i\prod_{i<l\le n}(-a_lb_i)\cdot\prod_{\substack{1\le k<l\le n\\k,l\ne i}}(a_kb_l-a_lb_k)$$

$$=\prod_{1\le k<l\le n}(a_kb_l-a_lb_k).\ (a_i=0)\,.$$

若至少有两个 $a_i=a_j=0$, 则第 i 行与第 j 行成比例, 因此行列式的值等于 0. 经过计算发现, 后面两种情形的答案都可以统一到第一种情形的答案.

综上所述,$|\boldsymbol{A}|=\prod\limits_{1\le i<j\le n}(a_ib_j-a_jb_i).$

□

结论 连乘号计算小结论:

(1) $\prod\limits_{1\le i<j\le n} a_ia_j = \prod\limits_{i=1}^{n} a_i^{n-1}$.

证明: $$\prod_{1\le i<j\le n} a_ia_j = \underbrace{a_2a_1\cdot a_3a_2a_3a_1\cdot a_4a_3a_4a_2a_4a_1\cdots\cdots \overbrace{a_ka_{k-1}a_ka_{k-2}\cdots a_ka_1}^{k-1\text{对}}\cdots\cdots\overbrace{a_na_{n-1}a_na_{n-2}\cdots a_na_1}^{n-1\text{对}}}_{n-1\text{组}}$$

$$\xlongequal{\text{从左往右按组计数}} a_1^{n-1}a_2^{1+n-2}a_3^{2+n-3}a_4^{3+n-4}\cdots a_k^{k-1+n-k}\cdots a_n^{n-1} = \prod_{i=1}^{n} a_i^{n-1}.$$

(2) $\prod\limits_{\substack{1\le i<j\le n\\ i,j\ne k}} a_ia_j = \prod\limits_{\substack{1\le i\le n\\ i\ne k}} a_i^{n-2}$, 其中 $k\in[1,n]\cap\mathbb{N}_+$.

证明: $$\prod_{\substack{1\le i<j\le n\\ i,j\ne k}} a_ia_j = \underbrace{a_2a_1\cdot a_3a_2a_3a_1\cdots\cdots\overbrace{a_{k-1}a_{k-2}\cdots a_{k-1}a_1}^{k-2\text{对}}\cdot\overbrace{a_{k+1}a_{k-1}\cdots a_{k+1}a_1}^{k-1\text{对}}\cdots\cdots\overbrace{a_na_{n-1}\cdots a_na_{k+1}a_na_{k-1}\cdots a_na_1}^{n-2\text{对}}}_{n-2\text{组}}$$

$$\xlongequal{\text{从左往右按组计数}} a_1^{n-2}a_2^{1+n-3}a_3^{2+n-4}a_4^{3+n-4}\cdots a_{k-1}^{k-2+n-k}a_{k+1}^{k-1+n-k-1}\cdots a_n^{n-2} = \prod_{\substack{1\le i\le n\\ i\ne k}} a_i^{n-2}.$$

注意: 从第 $k-1$ 组开始, 后面每组都比原来少一对 (后面每组均缺少原本含 a_k 的那一对).

例题 1.23 计算 $D_{n+1} = \begin{vmatrix} (a_0+b_0)^n & (a_0+b_1)^n & \cdots & (a_0+b_n)^n \\ (a_1+b_0)^n & (a_1+b_1)^n & \cdots & (a_1+b_n)^n \\ \vdots & \vdots & \vdots & \\ (a_n+b_0)^n & (a_n+b_1)^n & \cdots & (a_n+b_n)^n \end{vmatrix}$.

解 由二项式定理可知

$$\left(a_i+b_j\right)^n = {a_i}^n + \mathrm{C}_n^1{a_i}^{n-1}b_j + \cdots + \mathrm{C}_n^{n-1}a_i{b_j}^{n-1} + {b_j}^n, \text{其中} i,j=0,1,\cdots,n.$$

从而

$$D_{n+1} = \begin{vmatrix} {a_0}^n+\mathrm{C}_n^1{a_0}^{n-1}b_0+\cdots+\mathrm{C}_n^{n-1}a_0{b_0}^{n-1}+{b_0}^n & \cdots & {a_0}^n+\mathrm{C}_n^1{a_0}^{n-1}b_n+\cdots+\mathrm{C}_n^{n-1}a_0{b_n}^{n-1}+{b_n}^n \\ {a_1}^n+\mathrm{C}_n^1{a_1}^{n-1}b_0+\cdots+\mathrm{C}_n^{n-1}a_1{b_0}^{n-1}+{b_0}^n & \cdots & {a_1}^n+\mathrm{C}_n^1{a_1}^{n-1}b_n+\cdots+\mathrm{C}_n^{n-1}a_1{b_n}^{n-1}+{b_n}^n \\ \vdots & & \vdots \\ {a_{n-1}}^n+\mathrm{C}_n^1{a_{n1}}^{n-1}b_0+\cdots+\mathrm{C}_n^{n-1}a_{n-1}{b_0}^{n-1}+{b_0}^n & \cdots & {a_{n-1}}^n+\mathrm{C}_n^1{a_{n-1}}^{n-1}b_n+\cdots+\mathrm{C}_n^{n-1}a_{n-1}{b_n}^{n-1}+{b_n}^n \\ {a_n}^n+\mathrm{C}_n^1{a_n}^{n-1}b_0+\cdots+\mathrm{C}_n^{n-1}a_n{b_0}^{n-1}+{b_0}^n & \cdots & {a_n}^n+\mathrm{C}_n^1{a_n}^{n-1}b_n+\cdots+\mathrm{C}_n^{n-1}a_n{b_n}^{n-1}+{b_n}^n \end{vmatrix}$$

$$= \begin{vmatrix} {a_0}^n & {a_0}^{n-1} & \cdots & a_0 & 1 \\ {a_1}^n & {a_1}^{n-1} & \cdots & a_1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ {a_{n-1}}^n & {a_{n-1}}^{n-1} & \cdots & a_{n-1} & 1 \\ {a_n}^n & {a_n}^{n-1} & \cdots & a_n & 1 \end{vmatrix} \cdot \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ \mathrm{C}_n^1b_0 & \mathrm{C}_n^1b_1 & \cdots & \mathrm{C}_n^1b_{n-1} & \mathrm{C}_n^1b_n \\ \vdots & \vdots & & \vdots & \vdots \\ \mathrm{C}_n^{n-1}{b_0}^{n-1} & \mathrm{C}_n^{n-1}{b_1}^{n-1} & \cdots & \mathrm{C}_n^{n-1}{b_{n-1}}^{n-1} & \mathrm{C}_n^{n-1}{b_n}^{n-1} \\ {b_0}^n & {b_1}^n & \cdots & {b_{n-1}}^n & {b_n}^n \end{vmatrix}$$

$$\xlongequal{\text{列倒排}} (-1)^{\frac{n(n+1)}{2}} \begin{vmatrix} 1 & a_0 & \cdots & {a_0}^{n-1} & {a_0}^n \\ 1 & a_1 & \cdots & {a_1}^{n-1} & {a_1}^n \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & a_n & \cdots & {a_n}^{n-1} & {a_n}^n \end{vmatrix} \cdot \prod_{i=1}^{n-1}\mathrm{C}_n^i \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ b_0 & b_1 & \cdots & b_{n-1} & b_n \\ \vdots & \vdots & & \vdots & \vdots \\ {b_0}^{n-1} & {b_1}^{n-1} & \cdots & {b_{n-1}}^{n-1} & {b_n}^{n-1} \\ {b_0}^n & {b_1}^n & \cdots & {b_{n-1}}^n & {b_n}^n \end{vmatrix}$$

$$= (-1)^{\frac{n(n+1)}{2}} \prod_{0\le j<i\le n}\left(a_i-a_j\right)\prod_{i=1}^{n-1}\mathrm{C}_n^i\prod_{0\le j<i\le n}\left(b_i-b_j\right) == \prod_{i=1}^{n-1}\mathrm{C}_n^i\prod_{0\le j<i\le n}\left(a_j-a_i\right)\left(b_i-b_j\right).$$

□

例题 1.24 求下列行列式的值:

$$|\boldsymbol{A}|=\begin{vmatrix}1 & \cos\theta_1 & \cos 2\theta_1 & \cdots & \cos(n-1)\theta_1\\ 1 & \cos\theta_2 & \cos 2\theta_2 & \cdots & \cos(n-1)\theta_2\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & \cos\theta_n & \cos 2\theta_n & \cdots & \cos(n-1)\theta_n\end{vmatrix}.$$

解 由 De Moivre 公式及二项式定理, 可得

$$\begin{aligned}&\cos k\theta+\mathrm{i}\sin k\theta=(\cos\theta+\mathrm{i}\sin\theta)^k\\&=\cos^k\theta+\mathrm{i}\mathrm{C}_k^1\cos^{k-1}\theta\sin\theta-\mathrm{C}_k^2\cos^{k-2}\theta\sin^2\theta+\mathrm{i}\mathrm{C}_k^3\cos^{k-3}\theta\sin^3\theta-\cdots\\&=\cos^k\theta+\mathrm{i}\mathrm{C}_k^1\cos^{k-1}\theta\sin\theta-\mathrm{C}_k^2\cos^{k-2}\theta\left(1-\cos^2\theta\right)+\mathrm{i}\mathrm{C}_k^3\cos^{k-3}\theta\sin^3\theta-\cdots\end{aligned}$$

比较实部可得

$$\begin{aligned}\cos k\theta&=\cos^k\theta\left(1+\mathrm{C}_k^2+\mathrm{C}_k^4+\cdots\right)-\mathrm{C}_k^2\cos^{k-2}+\mathrm{C}_k^4\cos^{k-4}-\cdots\\&=2^{k-1}\cos^k\theta-\mathrm{C}_k^2\cos^{k-2}+\mathrm{C}_k^4\cos^{k-4}-\cdots\end{aligned}$$

利用这个事实, 依次将原行列式各列表示成 $\cos\theta_j(j=2,3,\cdots,n)$ 的多项式.

再利用行列式的性质, 可依次将第 $3,4,\cdots,n$ 列消去除最高次项外的其他项, 从而得到

$$\begin{aligned}|\boldsymbol{A}|&=\begin{vmatrix}1 & \cos\theta_1 & 2\cos^2\theta_1 & \cdots & 2^{n-2}\cos^{n-1}\theta_1\\ 1 & \cos\theta_2 & 2\cos^2\theta_2 & \cdots & 2^{n-2}\cos^{n-1}\theta_2\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & \cos\theta_n & 2\cos^2\theta_n & \cdots & 2^{n-2}\cos^{n-1}\theta_n\end{vmatrix}=2^{\frac{1}{2}(n-1)(n-2)}\begin{vmatrix}1 & \cos\theta_1 & \cos^2\theta_1 & \cdots & \cos^{n-1}\theta_1\\ 1 & \cos\theta_2 & \cos^2\theta_2 & \cdots & \cos^{n-1}\theta_2\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & \cos\theta_n & \cos^2\theta_n & \cdots & \cos^{n-1}\theta_n\end{vmatrix}\\&=2^{\frac{1}{2}(n-1)(n-2)}\prod_{1\le i<j\le n}\left(\cos\theta_j-\cos\theta_i\right).\end{aligned}$$

□

结论 组合式计算常用公式:

(1)$\mathrm{C}_n^m=\mathrm{C}_{n-1}^m+\mathrm{C}_{n-1}^{m-1}$

(2)$\mathrm{C}_n^0+\mathrm{C}_n^2+\cdots=\mathrm{C}_n^1+\mathrm{C}_n^3+\cdots=2^{n-1}$

证明:(1)

$$\begin{aligned}\mathrm{C}_n^m&=\frac{n!}{m!(n-m)!}=\frac{(n-1)!(n-m+m)}{m!(n-m)!}=\frac{(n-1)!(n-m)}{m!(n-m)!}+\frac{(n-1)!m}{m!(n-m)!}\\&=\frac{(n-1)!}{m!(n-m-1)!}+\frac{(n-1)!}{(m-1)!(n-m)!}=\mathrm{C}_{n-1}^m+\mathrm{C}_{n-1}^{m-1}\end{aligned}$$

(2)(i) 当 n 为奇数时, 由 $C_n^m=C_{n-1}^{m-1}+C_{n-1}^m$, 可得

$$\begin{aligned}C_n^0+C_n^2+C_n^4\cdots+C_n^{n-1}&=C_{n-1}^0+C_{n-1}^1+C_{n-1}^2+C_{n-1}^3+C_{n-1}^4\cdots+C_{n-1}^{n-2}+C_{n-1}^{n-1}\\C_n^1+C_n^3+C_n^5\cdots+C_n^n&=C_{n-1}^0+C_{n-1}^1+C_{n-1}^2+C_{n-1}^3+C_{n-1}^4+C_{n-1}^5+\cdots+C_{n-1}^{n-1}+C_{n-1}^n\end{aligned}$$

由于 $C_{n-1}^n=0$, 再对比上面两式每一项可知, 上面两式相等.

而上面两式相加, 得 $C_n^0+C_n^1+C_n^2\cdots+C_n^{n-1}+C_n^n=(1+1)^n=2^n$.

故 $C_n^0+C_n^2+C_n^4\cdots+C_n^{n-1}=C_n^1+C_n^3+C_n^5\cdots+C_n^n=2^{n-1}$.

(ii) 当 n 为偶数时, 由 $C_n^m=C_{n-1}^{m-1}+C_{n-1}^m$, 可得

$$\begin{aligned}C_n^0+C_n^2+C_n^4\cdots+C_n^n&=C_{n-1}^0+C_{n-1}^1+C_{n-1}^2+C_{n-1}^3+C_{n-1}^4\cdots+C_{n-1}^{n-1}+C_{n-1}^n\\C_n^1+C_n^3+C_n^5\cdots+C_n^{n-1}&=C_{n-1}^0+C_{n-1}^1+C_{n-1}^2+C_{n-1}^3+C_{n-1}^4+C_{n-1}^5+\cdots+C_{n-1}^{n-2}+C_{n-1}^{n-1}\end{aligned}$$

由于 $C_{n-1}^n=0$, 再对比上面两式每一项可知, 上面两式相等.

而上面两式相加, 得 $C_n^0+C_n^1+C_n^2\cdots+C_n^{n-1}+C_n^n=(1+1)^n=2^n$.

故 $C_n^0+C_n^2+C_n^4\cdots+C_n^{n-1}=C_n^1+C_n^3+C_n^5\cdots+C_n^n=2^{n-1}$.

综上所述,$C_n^0+C_n^2+\cdots=C_n^1+C_n^3+\cdots=2^{n-1}$.

例题 1.25 求下列行列式的值:

$$|\boldsymbol{A}|=\begin{vmatrix}\sin\theta_1 & \sin 2\theta_1 & \cdots & \sin n\theta_1\\ \sin\theta_2 & \sin 2\theta_2 & \cdots & \sin n\theta_2\\ \vdots & \vdots & & \vdots\\ \sin\theta_n & \sin 2\theta_n & \cdots & \sin n\theta_n\end{vmatrix}.$$

笔记 可以利用上一题类似的方法求解. 但我们给出另外一种解法, 目的是直接利用上一题的结论.

解 根据和差化积公式, 可得

$$\sin k\theta-\sin(k-2)\theta=2\sin\theta\cos(k-1)\theta,\ k=2,3,\cdots,n.$$

再结合上一题结论, 可得

$$\begin{aligned}|\boldsymbol{A}|&=\begin{vmatrix}\sin\theta_1 & \sin 2\theta_1 & \cdots & \sin n\theta_1\\ \sin\theta_2 & \sin 2\theta_2 & \cdots & \sin n\theta_2\\ \vdots & \vdots & & \vdots\\ \sin\theta_n & \sin 2\theta_n & \cdots & \sin n\theta_n\end{vmatrix}=\begin{vmatrix}\sin\theta_1 & 2\sin\theta_1\cos\theta_1 & \cdots & 2\sin\theta_1\cos(n-1)\theta_1\\ \sin\theta_2 & 2\sin\theta_2\cos\theta_2 & \cdots & 2\sin\theta_2\cos(n-1)\theta_2\\ \vdots & \vdots & & \vdots\\ \sin\theta_n & 2\sin\theta_n\cos\theta_n & \cdots & 2\sin\theta_n\cos(n-1)\theta_n\end{vmatrix}\\
&=2^{n-1}\prod_{i=1}^{n}\sin\theta_i\begin{vmatrix}\cos\theta_1 & \cos 2\theta_1 & \cdots & \cos(n-1)\theta_1\\ \cos\theta_2 & \cos 2\theta_2 & \cdots & \cos(n-1)\theta_2\\ \vdots & \vdots & & \vdots\\ \cos\theta_n & \cos 2\theta_n & \cdots & \cos(n-1)\theta_n\end{vmatrix}=2^{\frac{1}{2}(n-2)(n-1)+n-1}\prod_{i=1}^{n}\sin\theta_i\prod_{1\leq i<j\leq n}\left(\cos\theta_j-\cos\theta_i\right)\\
&=2^{\frac{1}{2}n(n-1)}\prod_{i=1}^{n}\sin\theta_i\prod_{1\leq i<j\leq n}\left(\cos\theta_j-\cos\theta_i\right).\end{aligned}$$

□

命题 1.14 (多项式根的有限性)

设多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

若 $f(x)$ 有 $n+1$ 个不同的根 $b_1,b_2,\cdots,b_{n+1}$, 即 $f(b_1)=f(b_2)=\cdots=f(b_{n+1})=0$, 求证:$f(x)$ 是零多项式, 即 $a_n=a_{n-1}=\cdots=a_1=a_0=0$.

证明 由 $f(b_1)=f(b_2)=\cdots=f(b_{n+1})=0$, 可知 $x_0=a_0,x_1=a_1,\cdots,x_{n-1}=a_{n-1},x_n=a_n$ 是下列线性方程组的解:

$$\begin{cases}x_0+b_1x_1+\cdots+b_1^{n-1}x_{n-1}+b_1^nx_n=0,\\ x_0+b_2x_1+\cdots+b_2^{n-1}x_{n-1}+b_2^nx_n=0,\\ \quad\cdots\cdots\cdots\cdots\\ x_0+b_{n+1}x_1+\cdots+b_{n+1}^{n-1}x_{n-1}+b_{n+1}^nx_n=0.\end{cases}$$

上述线性方程组的系数行列式是一个 Vandermode 行列式, 由于 $b_1,b_2,\cdots,b_{n+1}$ 互不相同, 所以系数行列式不等于零. 由 Crammer 法则可知上述方程组只有零解. 即有 $a_n=a_{n-1}=\cdots=a_1=a_0=0$. □

1.7 升阶法

例题 1.26 计算 n 阶行列式:

$$|\boldsymbol{A}| = \begin{vmatrix} 1+x_1 & 1+x_2 & \cdots & 1+x_1^n \\ 1+x_1^2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & & \vdots \\ 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{vmatrix}.$$

笔记 本题也可以使用大拆分法进行求解. 但我们以本题为例介绍利用**升阶法**计算行列式.

解 解法一升阶法:

$$\begin{aligned}
|\boldsymbol{A}| &= \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1+x_1 & 1+x_1^2 & \cdots & 1+x_1^n \\ 1 & 1+x_2 & 1+x_2^2 & \cdots & 1+x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1+x_n & 1+x_n^2 & \cdots & 1+x_n^n \end{vmatrix} = \begin{vmatrix} 1 & -1 & -1 & \cdots & -1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} \\
&\xlongequal{\text{小拆分法}} \begin{vmatrix} 2 & 0 & 0 & \cdots & 0 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} + \begin{vmatrix} -1 & -1 & -1 & \cdots & -1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} \\
&= 2\begin{vmatrix} x_1 & x_1^2 & \cdots & x_1^n \\ x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} - \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} \\
&= 2x_1x_2\cdots x_n \begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{vmatrix} - (x_1-1)(x_2-1)\cdots(x_n-1)\prod_{1\le i<j\le n}(x_j-x_i) \\
&= 2x_1x_2\cdots x_n \prod_{1\le i<j\le n}(x_j-x_i) - (x_1-1)(x_2-1)\cdots(x_n-1)\prod_{1\le i<j\le n}(x_j-x_i) \\
&= [2x_1x_2\cdots x_n - (x_1-1)(x_2-1)\cdots(x_n-1)]\prod_{1\le i<j\le n}(x_j-x_i).
\end{aligned}$$

解法二 (大拆分法): 设 $|\boldsymbol{B}(t)| = \begin{vmatrix} x_1+t & x_1^2+t & \cdots & x_1^n+t \\ x_2+t & x_2^2+t & \cdots & x_2^n+t \\ \vdots & \vdots & & \vdots \\ x_n+t & x_n^2+t & \cdots & x_n^n+t \end{vmatrix}$, 且 B_{ij} 是 $|\boldsymbol{B}(0)|$ 的第 (i,j) 元素的代数余子式.

根据行列式的性质将 $|\boldsymbol{A}|$ 每一列都拆分成两列, 然后按 t 所在的列展开得到

$$|\boldsymbol{A}| = |\boldsymbol{B}(1)| = |\boldsymbol{B}(0)| + \sum_{i,j=1}^{n} B_{ij},$$

$$|\boldsymbol{B}(-1)| = |\boldsymbol{B}(0)| - \sum_{i,j=1}^{n} B_{ij}.$$

于是 $|\boldsymbol{A}|=2|\boldsymbol{B}(0)|-|\boldsymbol{B}(-1)|$. 注意到

$$|\boldsymbol{B}(0)|=\begin{vmatrix} x_1 & x_1^2 & \cdots & x_1^n \\ x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ x_n & x_n^2 & \cdots & x_n^n \end{vmatrix}=x_1x_2\cdots x_n\begin{vmatrix} 1 & x_1 & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^n \end{vmatrix}=x_1x_2\cdots x_n\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

又由推论 1.4或命题 1.12可得

$$|\boldsymbol{B}(-1)|=\begin{vmatrix} x_1-1 & x_1^2-1 & \cdots & x_1^n-1 \\ x_2-1 & x_2^2-1 & \cdots & x_2^n-1 \\ \vdots & \vdots & & \vdots \\ x_n-1 & x_n^2-1 & \cdots & x_n^n-1 \end{vmatrix}=(x_1-1)(x_2-1)\cdots(x_n-1)\begin{vmatrix} 1 & x_1+1 & \cdots & x_1^{n-1}+x_1^{n-2}\cdots+x_1+1 \\ 1 & x_2+1 & \cdots & x_2^{n-1}+x_2^{n-2}\cdots+x_2+1 \\ \vdots & \vdots & & \vdots \\ 1 & x_n+1 & \cdots & x_n^{n-1}+x_n^{n-2}\cdots+x_n+1 \end{vmatrix}$$

$$=(x_1-1)(x_2-1)\cdots(x_n-1)\begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{vmatrix}=(x_1-1)(x_2-1)\cdots(x_n-1)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

故可得

$$\begin{aligned}|\boldsymbol{A}|&=2|\boldsymbol{B}(0)|-|\boldsymbol{B}(-1)|=2x_1x_2\cdots x_n\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)-(x_1-1)(x_2-1)\cdots(x_n-1)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)\\&=[2x_1x_2\cdots x_n-(x_1-1)(x_2-1)\cdots(x_n-1)]\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).\end{aligned}$$

□

结论 **升阶法:** 将原行列式加上一行和一列使得到到新行列式的阶数比原行列式要高一阶.

升阶法的应用:

(1) 当原行列式每一行具有相同的结构时, 我们可以在原行列式的基础上加上一行和一列, 新加上的一列和一行需要满足: 新的一列除了与新的一行交叉位置的元素为 1 外其余全为 0(这样才能保证新的行列式按新的一行或一列展开后与原行列式相同), 并且新加上的一行除 1 以外其他位置的元素就取原行列式中每一行所具有的相同结构 (这样可以利用行列式的性质将每一行中的相同的结构减去, 进而达到简化原行列式的目的). 具体例子见练习1.26.

(2) 当原行列式是我们由熟悉的行列式去掉某一行、或某一列、或某一行和一列得到的, 我们可以在原行列式的基础上补充上缺少的那一行和一列, 再进行计算得到新行列式的式子. 再将新行列式按照新添加的一行或一列展开得到的对应元素乘与其对应的代数余子式, 而新添加的一行和一列交叉位置的元素对应的余子式就是原行列式, 最后两边式子比较系数一般就能得到原行列式的值. 具体例子见练习1.27.

例题 1.27 求下列 n 阶行列式的值 $(1\leq i\leq n-1)$:

$$|\boldsymbol{A}|=\begin{vmatrix} 1 & x_1 & \cdots & x_1^{i-1} & x_1^{i+1} & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^{i-1} & x_2^{i+1} & \cdots & x_2^n \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{i-1} & x_n^{i+1} & \cdots & x_n^n \end{vmatrix}.$$

解 令

$$|\boldsymbol{B}|=\begin{vmatrix} 1 & x_1 & \cdots & x_1^{i-1} & x_1^{i} & x_1^{i+1} & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^{i-1} & x_2^{i} & x_2^{i+1} & \cdots & x_2^n \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{i-1} & x_n^{i} & x_n^{i+1} & \cdots & x_n^n \\ 1 & y & \cdots & y^{i-1} & y^{i} & y^{i+1} & \cdots & y^n \end{vmatrix}=(y-x_1)(y-x_2)\cdots(y-x_n)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i). \tag{1.14}$$

而上式右边是关于 y 的 n 次多项式, 并且其 y^i 前的系数是

$$\sum_{1\leqslant k_1<k_2<\cdots<k_{n-i}\leqslant n}(-1)^{n-i}x_{k_1}x_{k_2}\cdots x_{k_{n-i}}\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

将 $|\boldsymbol{B}|$ 按最后一行展开, 得

$$|\boldsymbol{B}|=A_{n1}+A_{n2}y+\cdots+A_{ni}y^i+\cdots+A_{nn}y^n,$$

其中 A_{nk} 为 $|\boldsymbol{B}|$ 的 (n,k) 位置元素的代数余子式,$k=1,2,\cdots,n$.

注意到 A_{nk} 均与 y 无关. 因此 $|\boldsymbol{B}|$ 作为关于 y 的 n 次多项式, 其 y^i 前的系数是

$$A_{ni}=(-1)^{n+1+i+1}|\boldsymbol{A}|=(-1)^{n+i}|\boldsymbol{A}|.$$

再结合(1.14)式, 可知

$$(-1)^{n+i}|\boldsymbol{A}|=\sum_{1\leqslant k_1<k_2<\cdots<k_{n-i}\leqslant n}(-1)^{n-i}x_{k_1}x_{k_2}\cdots x_{k_{n-i}}\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

故 $|\boldsymbol{A}|=x_{k_1}x_{k_2}\cdots x_{k_{n-i}}\prod\limits_{1\leqslant i<j\leqslant n}(x_j-x_i).$ □

1.8 求根法

1.9 行列式的组合定义

例题 1.28 若 n 阶行列式 $|A|$ 中零元素的个数超过 n^2-n 个, 证明:$|A|=0$.

证明 由行列式的组合定义可得

$$|A|=\sum_{1\leq k_1k_2\cdots k_n\leq n}(-1)^{\tau(k_1,k_2,\cdots,k_n)}a_{k_11}a_{k_22}\cdots a_{k_nn}$$

由于 $|A|$ 中零元素的个数超过 n^2-n 个, 故 $a_{k_11},a_{k_22},\cdots,a_{k_nn}$ 中至少有一个为零, 从而 $a_{k_11}a_{k_22}\cdots a_{k_nn}=0$, 因此 $|A|=0$. 如直接利用行列式的性质, 也可以这样来证明: 因为 $|A|$ 中零元素的个数超过 n^2-n 个, 由抽屉原理可知,$|A|$ 至少有一列其零元素的个数大于等于 $\left\lfloor\dfrac{n^2-n}{n}\right\rfloor+1=n$, 即 $|A|$ 至少有一列其元素全为零, 因此 $|A|=0$. □

例题 1.29 设 $A=(a_{ij})$ 是 $n(n\geq 2)$ 阶非异整数方阵, 满足对任意的 i,j,$|A|$ 均可整除 a_{ij}, 证明:$|A|=\pm 1$.

解 $|A|$ 可整除每个元素 $a_{i,j}$, 故由行列式的组合定义

$$\sum_{1\leq k_1,k_2,\cdots,k_n\leq n}(-1)^{\tau(k_1k_2\cdots k_n)}a_{k_11}a_{k_22}\cdots a_{k_nn}$$

可知 $|A|^n$ 可整除 $|A|$ 中每个单项 $a_{k_11}a_{k_22}\cdots a_{k_nn}$, 从而 $|A|^n$ 可整除 $|A|$, 即有 $|A|^{n-1}$ 可整除 1, 于是 $|A|^{n-1}=\pm 1$. 又由行列式的组合定义可知 $|A|$ 是整数, 从而只能是 $|A|=\pm 1$. □

命题 1.15 (奇数阶反对称行列式的值等于零)

如果 n 阶行列式 $|A|$ 的元素满足 $a_{ij}=-a_{ji}(1\leq i,j\leq n)$, 则称为反对称行列式. 求证: 奇数阶反对称行列式的值等于零.

笔记 证法二的想法是将行列式按组合的定义写成 (n-1)! 个单项的和. 然后将其两两分组再求和 (因为一共有 (n-1)! 个单项, 即和式中共有偶数个单项, 所以只要使用合适的分组方式就一定能够将其两两分组再求和), 最后发现每组的和均为 0.

构造的这个映射 φ 的目的是为了更加准确、严谨地说明分组的方式. 证明这个映射 φ 是一个双射是为了保证原来的和式中的每一个单项都能与和式中另一个单项一一对应. 然后利用反证法证明了这两个一一对应的单项一定互不相同 **(注: 我认为这步有些多余. 这里应该只需要说明这两个一一对应的单项是原和式中不同的单项即可, 即这两个单项的角标不完全相同就行, 其实, 这个在我们定义映射 φ 的时候就已经满足了. 满足这个条件就足**

以说明原和式可以按照这种方式进行分组. 并且利用反对称行列式的性质也能够证明这两个单项不仅互不相同, 还能进一步得到这两个单项互为相反数). 于是我们就可以将原和式中的每一个单项与其在双射 φ 作用下的像看成一组, 按照这种方式就可以将原和式进行分组再求和.

证明 证法一 (行列式的性质): 由反对称行列式的定义可知,$|A|$ 的转置 $|A'|$ 与 $|A|$ 的每个元素都相差一个符号, 将 $|A'|$ 的每一行都提出公因子 -1 可得 $|A| = |A'| = (-1)^n|A| = -|A|$, 从而 $|A| = 0$.

证法二 (行列式的组合定义): 由于 $|A|$ 的主对角元全为 0, 故由组合定义, 只需考虑下列单项:

$$T = \{a_{k_1 1}a_{k_2 2}\cdots a_{k_n n} \mid k_i \neq i(1 \leq i \leq n)\}$$

定义映射 $\varphi : T \to T, a_{k_1 1}a_{k_2 2}\cdots a_{k_n n} \mapsto a_{1k_1}a_{2k_2}\cdots a_{nk_n}$. 显然 $\varphi^2 = \mathrm{Id}_T$, 于是 φ 是一个双射. 我们断言:$a_{k_1 1}a_{k_2 2}\cdots a_{k_n n}$ 和 $a_{1k_1}a_{2k_2}\cdots a_{nk_n}$ 作为 $|A|$ 的单项不相同, 否则 $\{1,2,\cdots,n\}$ 必可分成若干对 $(i_1, j_1),\cdots,(i_t, j_t)$, 使得 $a_{k_1 1}a_{k_2 2}\cdots a_{k_n n} = a_{i_1 j_1}a_{j_1 i_1}\cdots a_{i_t j_t}a_{j_t i_t}$, 这与 n 为奇数矛盾. 将上述两个单项看成一组, 则它们在 $|A|$ 中符号均为 $(-1)^{\tau(k_1k_2\cdots k_n)}$. 由于 $|A|$ 反对称, 故

$$a_{1k_1}a_{2k_2}\cdots a_{nk_n} = (-1)^n a_{k_1 1}a_{k_2 2}\cdots a_{k_n n} = -a_{k_1 1}a_{k_2 2}\cdots a_{k_n n}$$

从而每组和为 0, 于是 $|A| = 0$. □

命题 1.16 (直接计算两个矩阵和的行列式)

设 A, B 都是 n 阶矩阵, 求证:

$$|\boldsymbol{A}+\boldsymbol{B}| = |\boldsymbol{A}| + |\boldsymbol{B}| + \sum_{1\leq k\leq n-1}\left(\sum_{\substack{1\leq i_1<i_2<\cdots<i_k\leq n\\1\leq j_1<j_2<\cdots<j_k\leq n}} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\widehat{\boldsymbol{B}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\right).$$

其中 $\widehat{\boldsymbol{B}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}$ 是 $|\boldsymbol{B}|$ 的 k 阶子式 $\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}$ 的代数余子式.

笔记 当 $\boldsymbol{A},\boldsymbol{B}$ 之一是比较简单的矩阵 (例如对角矩阵或秩较小的矩阵) 时, 可利用这个命题计算 $|\boldsymbol{A}+\boldsymbol{B}|$.

证明 设 $|\boldsymbol{A}| = |\alpha_1, \alpha_2, \cdots, \alpha_n|, |\boldsymbol{B}| = |\beta_1, \beta_2, \cdots, \beta_n|$, 其中 $\alpha_j, \beta_j (j = 1, 2, \cdots, n)$ 分别是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的列向量. 注意到

$$|\boldsymbol{A}+\boldsymbol{B}| = |\alpha_1+\beta_1, \alpha_2+\beta_2, \cdots, \alpha_n+\beta_n|.$$

对 $|\boldsymbol{A}+\boldsymbol{B}|$, 按列用行列式的性质展开, 使每个行列式的每一列或者只含有 α_j, 或者只含有 β_j(即利用大拆分法按列向量将行列式完全拆分开), 则 $|\boldsymbol{A}+\boldsymbol{B}|$ 可以表示为 2^n 个这样的行列式之和. 即 (并且单独把 $k = 0, n$ 的项分离出来, 即将 $|\boldsymbol{A}|$、$|\boldsymbol{B}|$ 分离出来)

$$\begin{aligned}|\boldsymbol{A}+\boldsymbol{B}| &= |\alpha_1+\beta_1, \alpha_2+\beta_2, \cdots, \alpha_n+\beta_n|\\ &= |\boldsymbol{A}| + |\boldsymbol{B}| + \sum_{1\leqslant k\leqslant n-1}\sum_{1\leq j_1\leq j_2\leq\cdots\leq j_k\leq n} \begin{array}{c} 1 \ \cdots\ j_1\ \cdots\ j_2\ \cdots\ j_k\ \cdots\ n\\ |\beta_1,\cdots,\alpha_{j_1},\cdots,\alpha_{j_2},\cdots,\alpha_{j_k},\cdots,\beta_n|\end{array}.\end{aligned}$$

再对上式右边除 $|\boldsymbol{A}|$、$|\boldsymbol{B}|$ 外的每个行列式用 *Laplace* 定理按含有 $\boldsymbol{A}$ 的列向量的那些列展开得到

$$\begin{aligned}|\boldsymbol{A}+\boldsymbol{B}| &= |\boldsymbol{A}| + |\boldsymbol{B}| + \sum_{1\leqslant k\leqslant n-1}\sum_{1\leq j_1\leq j_2\leq\cdots\leq j_k\leq n} \begin{array}{c} 1 \ \cdots\ j_1\ \cdots\ j_2\ \cdots\ j_k\ \cdots\ n\\ |\beta_1,\cdots,\alpha_{j_1},\cdots,\alpha_{j_2},\cdots,\alpha_{j_k},\cdots,\beta_n|\end{array}\\ &= |\boldsymbol{A}| + |\boldsymbol{B}| + \sum_{1\leqslant k\leqslant n-1}\sum_{1\leqslant j_1,j_2,\cdots,j_k\leqslant n}\sum_{1\leqslant i_1,i_2,\cdots,i_k\leqslant n} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\widehat{\boldsymbol{B}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\\ &= |\boldsymbol{A}| + |\boldsymbol{B}| + \sum_{1\leq k\leq n-1}\left(\sum_{\substack{1\leq i_1<i_2<\cdots<i_k\leq n\\1\leq j_1<j_2<\cdots<j_k\leq n}} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\widehat{\boldsymbol{B}}\begin{pmatrix} i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\right).\end{aligned}$$

□

例题 1.30 设

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

其中 x 是未定元,a_{ij} 是常数. 证明:$f(x)$ 是一个最高次项系数为 1 的 n 次多项式, 且其 $n-1$ 次项的系数等于 $-(a_{11}+a_{22}+\cdots+a_{nn})$.

笔记 注意 $f(x)$ 的每行每列除主对角元素外, 其他元素均不相同. 因此 $f(x)$ 并不是推广的" 爪" 型行列式.

解 由行列式的组合定义可知,$f(x)$ 的最高次项出现在组合定义展开式中的单项 $(x-a_{11})(x-a_{22})\cdots(x-a_{nn})$ 中, 且展开式中的其他单项作为 x 的多项式其次数小于等于 $n-2$. 因此 $f(x)$ 是一个最高次项系数为 1 的 n 次多项式, 且其 $n-1$ 次项的系数等于 $-(a_{11}+a_{22}+\cdots+a_{nn})$. □

注 将这个例题进行推广再结合直接计算两个矩阵和的行列式的结论可以得到下述推论.

推论 1.5

设 $A=(a_{ij})$ 为 n 阶方阵,x 为未定元,

$$f(x)=|xI_n-A|=\begin{vmatrix} x-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & x-a_{nn} \end{vmatrix}$$

证明:$f(x)=x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$, 其中

$$a_k=(-1)^k\sum_{1\le i_1<i_2<\cdots<i_k\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}, 1\le k\le n.$$

笔记 需要注意上述推论中 $a_1=-(a_{11}+a_{22}+\cdots+a_{nn}), a_n=(-1)^n|\boldsymbol{A}|$.

证明 注意到 xI_n 非零的 $n-k$ 阶子式只有 $n-k$ 阶主子式, 并且其值为 x^{n-k}, 其余 $n-k$ 阶子式均为零. 记 $\widehat{xI_n}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 是 $xI_n\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 的代数余子式, 则 $\widehat{xI_n}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$ 是 xI_n 非零的 $n-k$ 阶子式. 于是我们有

$$\widehat{xI_n}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}=x^{n-k}.$$

再利用直接计算两个矩阵和的行列式的结论就可以得到

$$f(x)=|x\boldsymbol{I}_n-\boldsymbol{A}|=\begin{vmatrix} x-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & x-a_{nn} \end{vmatrix}=\left|\begin{pmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & -a_{nn} \end{pmatrix}+\begin{pmatrix} x & 0 & \cdots & 0 \\ 0 & x & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & x \end{pmatrix}\right|$$

$$=\begin{vmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & -a_{nn} \end{vmatrix}+\begin{vmatrix} x & 0 & \cdots & 0 \\ 0 & x & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & x \end{vmatrix}+\sum_{1\le k\le n-1}\sum_{\substack{1\le i_1,i_2,\cdots,i_k\le n \\ 1\le j_1,j_2,\cdots,j_k\le n}}(-\boldsymbol{A})\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}\widehat{x\boldsymbol{I}_n}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$$

$$=(-1)^n|\boldsymbol{A}|+x^n+\sum_{1\le k\le n-1}\sum_{1\le i_1,i_2,\cdots,i_k\le n}(-1)^k\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}\widehat{x\boldsymbol{I}_n}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}$$

$$= x^n + \sum_{1\le k\le n-1} (-1)^k \sum_{1\le i_1,i_2,\cdots,i_k\le n} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} \cdot x^{n-k} + (-1)^n |\boldsymbol{A}|$$

$$= x^n + \sum_{1\le k\le n-1} x^{n-k} (-1)^k \sum_{1\le i_1,i_2,\cdots,i_k\le n} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} + (-1)^n |\boldsymbol{A}|.$$

因此 $f(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$, 其中

$$a_k = (-1)^k \sum_{1\le i_1<i_2<\cdots<i_k\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}, 1 \le k \le n.$$

□

1.10 Laplace 定理

例题 1.31 利用行列式的 Laplace 定理证明恒等式:

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

解 显然下列行列式的值为零:

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

利用 *Laplace* 定理按第一、二列展开得

$$\begin{vmatrix} a & a' & a & a' \\ b & b' & b & b' \\ c & c' & c & c' \\ d & d' & d & d' \end{vmatrix} = (-1)^{1+2+1+2}\begin{vmatrix} a & a' \\ b & b' \end{vmatrix}\begin{vmatrix} c & c' \\ d & d' \end{vmatrix} + (-1)^{1+2+1+3}\begin{vmatrix} a & a' \\ c & c' \end{vmatrix}\begin{vmatrix} b & b' \\ d & d' \end{vmatrix} + (-1)^{1+2+1+4}\begin{vmatrix} a & a' \\ d & d' \end{vmatrix}\begin{vmatrix} b & b' \\ c & c' \end{vmatrix}$$

$$+ (-1)^{1+2+2+3}\begin{vmatrix} b & b' \\ c & c' \end{vmatrix}\begin{vmatrix} a & a' \\ d & d' \end{vmatrix} + (-1)^{1+2+2+4}\begin{vmatrix} b & b' \\ d & d' \end{vmatrix}\begin{vmatrix} a & a' \\ c & c' \end{vmatrix}$$

$$+ (-1)^{1+2+3+4}\begin{vmatrix} c & c' \\ d & d' \end{vmatrix}\begin{vmatrix} a & a' \\ b & b' \end{vmatrix}$$

$$= 2\begin{vmatrix} a & a' \\ b & b' \end{vmatrix}\begin{vmatrix} c & c' \\ d & d' \end{vmatrix} - 2\begin{vmatrix} a & a' \\ c & c' \end{vmatrix}\begin{vmatrix} b & b' \\ d & d' \end{vmatrix} + 2\begin{vmatrix} a & a' \\ d & d' \end{vmatrix}\begin{vmatrix} b & b' \\ c & c' \end{vmatrix} = 0.$$

上式等价于

$$\begin{vmatrix} a & a' \\ b & b' \end{vmatrix}\begin{vmatrix} c & c' \\ d & d' \end{vmatrix} - \begin{vmatrix} a & a' \\ c & c' \end{vmatrix}\begin{vmatrix} b & b' \\ d & d' \end{vmatrix} + \begin{vmatrix} a & a' \\ d & d' \end{vmatrix}\begin{vmatrix} b & b' \\ c & c' \end{vmatrix} = 0.$$

整理可得

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

□

例题 1.32 求 $2n$ 阶行列式的值 (空缺处都是零):

$$\begin{vmatrix} a & & & & & b \\ & \ddots & & & \iddots & \\ & & a & b & & \\ & & b & a & & \\ & \iddots & & & \ddots & \\ b & & & & & a \end{vmatrix}.$$

解 设原行列式为 D_{2n}, 其中 $2n$ 为行列式的阶数. 不断用 $Laplace$ 定理按第一行及最后一行展开, 可得

$$D_{2n}=\begin{vmatrix} a & & & & & b \\ & \ddots & & & \iddots & \\ & & a & b & & \\ & & b & a & & \\ & \iddots & & & \ddots & \\ b & & & & & a \end{vmatrix} \xlongequal{\text{按第一行及最后一行展开}} \begin{vmatrix} a & b \\ b & a \end{vmatrix} D_{2n-2}=\left(a^2-b^2\right)D_{2(n-1)}.$$

进而, 由上述递推式可得

$$\begin{aligned} D_{2n} &= \left(a^2-b^2\right)D_{2(n-1)}=\left(a^2-b^2\right)^2 D_{2(n-2)}=\cdots=\left(a^2-b^2\right)^{n-1}D_2 \\ &= \left(a^2-b^2\right)^{n-1}\begin{vmatrix} a & b \\ b & a \end{vmatrix}=\left(a^2-b^2\right)^n. \end{aligned}$$

□

1.11 循环行列式

循环行列式关于单位根的计算公式见命题 2.3.

例题 1.33 设 a,n 是给定互素正整数, 按 Build 除法, 存在唯一确定的整数对 (s,t) 使得 $a=sn+t, 0\le t\le n-1$. 令

$$u_i=\begin{cases} s+1, & 0\le i<t \\ s, & t\le i\le n-1 \end{cases}$$

若 t 与 n 互素, 计算

$$D_n=\begin{vmatrix} u_0 & u_1 & \cdots & u_{n-1} \\ u_{n-1} & u_0 & \cdots & u_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ u_1 & u_2 & \cdots & u_0 \end{vmatrix}$$

证明 记 $f(x)\triangleq\sum\limits_{i=1}^{n}u_ix^i, w_j\triangleq e^{\frac{2\pi ji}{n}}, j=0,1,2,\cdots,n-1$, 则由命题 2.3可知

$$D_n=\prod_{k=0}^{n}f\left(w_k\right)=f\left(1\right)\prod_{k=0}^{n-1}f\left(w_k\right).$$

由条件可知

$$f(1)=\sum_{i=0}^{t-1}(s+1)+\sum_{i=t}^{n-1}s=(s+1)t+(n-t)s=ns+t=a.$$

从而

$$D_n = a\prod_{i=0}^{n-1} f(w_k) = a\prod_{i=0}^{n-1}\left[\sum_{i=0}^{t-1}(s+1)w_k^i + \sum_{i=t}^{n-1} sw_k^i\right]$$

$$= a\prod_{i=0}^{n-1}\left[\sum_{i=0}^{t-1} w_k^i + s\sum_{i=0}^{n-1} w_k^i\right] = a\prod_{i=0}^{n-1}\left(\frac{1-w_k^t}{1-w_k} + s\frac{1-w_k^n}{1-w_k}\right).$$

由 $w_k^n = 1, w_k = w_1^k$ 可知

$$D_n = a\prod_{i=0}^{n-1}\frac{1-w_1^{kt}}{1-w_1^k}.$$

由群论可知 $\{1, w_1, \cdots, w_1^{n-1}\}$ 是一个循环群且 w_1 的阶为 n, 再根据群论的 Lagrange 定理及 $(t,n)=1$ 可知, w_1^t 的阶为 $\frac{n}{(n,t)} = n$. 因此 $w_1^k = w_1$, 故 $\{w_1, w_1^2, \cdots, w_1^{n-1}\} = \{w_1^t, w_2^{2t}, \cdots, w_1^{(n-1)t}\}$. 于是 $w_1^k = w_1^{tk}$, 故

$$D_n = a\prod_{i=0}^{n-1}\frac{1-w_1^{kt}}{1-w_1^k} = a.$$

□

例题 1.34 计算:

$$D_n = \begin{vmatrix} a & a+d & \cdots & a+(n-1)d \\ a+(n-1)d & a & \cdots & a+(n-2)d \\ \vdots & \vdots & & \vdots \\ a+d & a+2d & \cdots & a \end{vmatrix}.$$

证明 记 $f(x) = \sum_{j=0}^{n-1}(a+jd)w_k^j$, 其中 $w_k = e^{\frac{2\pi k i}{n}}, k = 0,1,2,\cdots,n-1$. 由命题 2.3可知

$$D_n = f(w_0)f(w_1)\cdots f(w_{n-1}) = \prod_{k=0}^{n-1}\sum_{j=0}^{n-1}(a+jd)w_k^j = \frac{2an+n(n-1)d}{2}\prod_{k=1}^{n-1}\sum_{j=0}^{n-1}(a+jd)w_k^j$$

$$\xlongequal{\text{错位相减}} \frac{2an+n(n-1)d}{2}\prod_{k=1}^{n-1}\frac{aw_k^{n+1}-aw_k^n-aw_k+a-dw_k^{n+1}-dnw_k^n+dw_k}{(w_k-1)^2}$$

$$\xlongequal{w_k^n=1} \frac{2an+n(n-1)d}{2}\prod_{k=1}^{n-1}\frac{dn}{w_k-1} = \frac{2an+n(n-1)d}{2}\cdot(dn)^{n-1}\prod_{k=1}^{n-1}\frac{1}{w_k-1}.$$

注意到 $w_k - 1, k = 1,2,\cdots,n-1$ 是 $(x+1)^n - 1 = 0$ 的 $n-1$ 个复根, 这些根和 $w_0 - 1 = 0$ 一起就是 $(x+1)^n - 1 = 0$ 的全部复根. 从而由 Vieta 定理可得, $(x+1)^n - 1$ 的一次项系数乘 $(-1)^{n-1}$ 为

$$(-1)^{n-1}n = \sum_{0\leqslant i_1<i_2<\cdots<i_{n-1}\leqslant n-1}(w_{i_1}-1)(w_{i_2}-1)\cdots(w_{i_{n-1}}-1) = \prod_{k=1}^{n-1}(w_k-1).$$

故

$$D_n = \frac{2an+n(n-1)d}{2}\cdot(dn)^{n-1}\prod_{k=1}^{n-1}\frac{1}{w_k-1} = (-1)^{n-1}\cdot\frac{2a+(n-1)d}{2}\cdot(nd)^{n-1}.$$

□

1.12 行列式综合问题

例题 1.35 求下列 n 阶行列式的值:

$$|A|=\begin{vmatrix}(x-a_1)^2 & a_2^2 & \cdots & a_n^2\\ a_1^2 & (x-a_2)^2 & \cdots & a_n^2\\ \vdots & \vdots & \ddots & \vdots\\ a_1^2 & a_2^2 & \cdots & (x-a_n)^2\end{vmatrix}.$$

笔记 注意到这个行列式每行元素除了主对角元素外, 其余位置元素都相同. 因此这个行列式是推广的"爪"型行列式.

解

$$|A|=\begin{vmatrix}(x-a_1)^2 & a_2^2 & \cdots & a_n^2\\ a_1^2 & (x-a_2)^2 & \cdots & a_n^2\\ \vdots & \vdots & \ddots & \vdots\\ a_1^2 & a_2^2 & \cdots & (x-a_n)^2\end{vmatrix}=\begin{vmatrix}(x-a_1)^2 & a_2^2 & \cdots & a_n^2\\ 2a_1x-x^2 & x^2-2a_2x & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 2a_1x-x^2 & 0 & \cdots & x^2-2a_nx\end{vmatrix}$$

$$\xlongequal{\text{"爪"型行列式}}(x-a_1)^2\prod_{i=2}^{n}\left(x^2-2a_ix\right)-\sum_{i=2}^{n}a_i^2\left(2a_1x-x^2\right)\left(x^2-2a_2x\right)\cdots\widehat{\left(x^2-2a_ix\right)}\cdots\left(x^2-2a_nx\right)$$

$$=(x-a_1)^2\prod_{i=2}^{n}\left(x^2-2a_ix\right)+\sum_{i=2}^{n}a_i^2\left(x^2-2a_1x\right)\left(x^2-2a_2x\right)\cdots\widehat{\left(x^2-2a_ix\right)}\cdots\left(x^2-2a_nx\right)$$

$$=(x-a_1)^2\prod_{i=2}^{n}\left(x^2-2a_ix\right)+\sum_{i=2}^{n}\left(x^2-2a_1x\right)\cdots\left(x^2-2a_{i-1}x\right)a_i^2\left(x^2-2a_{i+1}x\right)\cdots\left(x^2-2a_nx\right)$$

$$=\left[\left(x^2-2a_1x\right)+a_1^2\right]\prod_{i=2}^{n}\left(x^2-2a_ix\right)+\sum_{i=2}^{n}\left(x^2-2a_1x\right)\cdots\left(x^2-2a_{i-1}x\right)a_i^2\left(x^2-2a_{i+1}x\right)\cdots\left(x^2-2a_nx\right)$$

$$=\prod_{i=1}^{n}\left(x^2-2a_ix\right)+\sum_{i=1}^{n}\left(x^2-2a_1x\right)\cdots\left(x^2-2a_{i-1}x\right)a_i^2\left(x^2-2a_{i+1}x\right)\cdots\left(x^2-2a_nx\right).$$

□

例题 1.36 求下列行列式式的值:

$$|\boldsymbol{A}|=\begin{vmatrix}(a+b)^2 & c^2 & c^2\\ a^2 & (b+c)^2 & a^2\\ b^2 & b^2 & (c+a)^2\end{vmatrix}.$$

解 解法一:

$$|\boldsymbol{A}|=\begin{vmatrix}(a+b)^2 & c^2 & c^2\\ a^2 & (b+c)^2 & a^2\\ b^2 & b^2 & (c+a)^2\end{vmatrix}\xlongequal[i=1,2]{-j_1+j_i}\begin{vmatrix}(a+b)^2-c^2 & c^2 & 0\\ a^2-(b+c)^2 & (b+c)^2 & a^2-(b+c)^2\\ 0 & b^2 & (c+a)^2-b^2\end{vmatrix}$$

$$=(a+b+c)^2\begin{vmatrix}a+b-c & c^2 & 0\\ a-b-c & (b+c)^2 & a-b-c\\ 0 & b^2 & a+c-b\end{vmatrix}\xlongequal[i=1,2]{-r_i+r_2}(a+b+c)^2\begin{vmatrix}a+b-c & c^2 & 0\\ -2b & 2bc & -2c\\ 0 & b^2 & a+c-b\end{vmatrix}$$

$$\xlongequal[\frac{b}{2}j_3+j_2]{\frac{c}{2}j_1+j_2}(a+b+c)^2\begin{vmatrix}a+b-c & \frac{c}{2}(a+b+c) & 0\\ -2b & 0 & -2c\\ 0 & \frac{b}{2}(a+b+c) & a+c-b\end{vmatrix}=(a+b+c)^3\begin{vmatrix}a+b-c & \frac{c}{2} & 0\\ -2b & 0 & -2c\\ 0 & \frac{b}{2} & a+c-b\end{vmatrix}$$

$$=2abc(a+b+c)^3.$$

解法二 (求根法):

□

例题 1.37 证明: 若一个 $n(n>1)$ 阶行列式中元素或为 1 或为 -1, 则其值必为偶数.

证明 将该行列式的任意一行加到另一行上去得到的行列式有一行元素全是偶数 (注意: 零也是偶数), 由行列式的基本性质知道, 可将因子 2 提出, 剩下的行列式的元素都是整数, 其值也是整数, 乘以 2 后必是偶数. □

例题 1.38 n 阶行列式 $|\boldsymbol{A}|$ 的值为 c, 若从第二列开始每一列加上它前面的一列, 同时对第一列加上 $|\boldsymbol{A}|$ 的第 n 列, 求得到的新行列式 $|\boldsymbol{B}|$ 的值.

解

$$
\begin{aligned}
|\boldsymbol{B}| &= |\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_2+\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n+\boldsymbol{\alpha}_{n-1}| \\
&= |\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n| + |\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{n-1}| + \sum_{1\leqslant k\leqslant n-2}\ \sum_{2\le j_1\le j_2\le\cdots\le j_k\le n} \begin{array}{c} \begin{array}{ccccccccc} 1 & \cdots & j_1 & \cdots & j_2 & \cdots & j_k & \cdots & n \end{array} \\ |\boldsymbol{\alpha}_n, \cdots, \boldsymbol{\alpha}_{j_1+1}, \cdots, \boldsymbol{\alpha}_{j_2+1}, \cdots, \boldsymbol{\alpha}_{j_k+1}, \cdots, \boldsymbol{\alpha}_{n-1}| \end{array} \\
&\quad + \sum_{1\leqslant k\leqslant n-2}\ \sum_{2\le j_1\le j_2\le\cdots\le j_k\le n} \begin{array}{c} \begin{array}{ccccccccc} 1 & \cdots & j_1 & \cdots & j_2 & \cdots & j_k & \cdots & n \end{array} \\ |\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{j_1+1}, \cdots, \boldsymbol{\alpha}_{j_2+1}, \cdots, \boldsymbol{\alpha}_{j_k+1}, \cdots, \boldsymbol{\alpha}_{n-1}| \end{array} . \\
&= |\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n| + |\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_{n-1}| \\
&= c + (-1)^{n-1} |\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n| \\
&= c + (-1)^{n-1} c \\
&= \begin{cases} 0\ , n\text{为偶数} \\ 2c, n\text{为奇数} \end{cases}
\end{aligned}
$$

□

例题 1.39 令

$$
(a_1a_2\cdots a_n) = \begin{vmatrix} a_1 & 1 & & & \\ -1 & a_2 & 1 & & \\ & -1 & a_3 & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & -1 & a_n \end{vmatrix},
$$

证明关于连分数的如下等式成立:

$$
a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}} = \frac{(a_1a_2\cdots a_n)}{(a_2a_3\cdots a_n)}.
$$

解 假设等式对 $\forall n \le k-1, k\in\mathbb{N}_+$ 都成立. 则当 $n=k$ 时, 将行列式 $(a_1a_2,\cdots,a_k)$ 按第一列展开得

$$
\begin{aligned}
(a_1a_2\cdots a_k) &= \begin{vmatrix} a_1 & 1 & & & \\ -1 & a_2 & 1 & & \\ & -1 & a_3 & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & -1 & a_k \end{vmatrix} = a_1 \begin{vmatrix} a_2 & 1 & & \\ -1 & a_3 & \ddots & \\ & \ddots & \ddots & 1 \\ & & -1 & a_k \end{vmatrix} + \begin{vmatrix} a_3 & 1 & & \\ -1 & a_4 & \ddots & \\ & \ddots & \ddots & 1 \\ & & -1 & a_k \end{vmatrix} \\
&= a_1(a_2a_3\cdots a_k) + (a_3a_4\cdots a_k).
\end{aligned}
$$

从而

$$
\frac{(a_1a_2\cdots a_k)}{(a_2a_3\cdots a_k)} = a_1 + \frac{(a_3a_4\cdots a_k)}{(a_2a_3\cdots a_k)} = a_1 + \frac{1}{\frac{(a_2a_3\cdots a_k)}{(a_3a_4\cdots a_k)}}.
$$

于是由归纳假设可知

$$\frac{(a_1a_2\cdots a_k)}{(a_2a_3\cdots a_k)}=a_1+\frac{1}{\frac{(a_2a_3\cdots a_k)}{(a_3a_4\cdots a_k)}}=a_1+\frac{1}{a_2+\frac{1}{a_3+\cdots+\frac{1}{a_{n-1}+\frac{1}{a_n}}}}.$$

故由数学归纳法可知结论成立. □

例题 1.40 设 $|A|$ 是 n 阶行列式,$|A|$ 的第 (i,j) 元素 $a_{ij}=\max\{i,j\}$, 试求 $|A|$ 的值.

解

$$|\boldsymbol{A}|=\begin{vmatrix}1&2&3&\cdots&n\\2&2&3&\cdots&n\\3&3&3&\cdots&n\\\vdots&\vdots&\vdots&&\vdots\\n&n&n&\cdots&n\end{vmatrix}\xlongequal[i=n,n-1,\cdots,2]{-r_i+r_{i-1}}\begin{vmatrix}-1&0&0&\cdots&0\\2&-1&0&\cdots&0\\3&3&-1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\n&n&n&\cdots&n\end{vmatrix}=(-1)^{n-1}n.$$

□

例题 1.41 设 $|A|$ 是 n 阶行列式,$|A|$ 的第 (i,j) 元素 $a_{ij}=|i-j|$, 试求 $|A|$ 的值.

笔记 注意: 这只是一个**对称行列式**, 不是循环行列式. 类似这种每行、每列元素有一定的等差递进关系的行列式, 都可以先尝试用每一列减去前面一列.

解

$$|\boldsymbol{A}|=\begin{vmatrix}0&1&2&\cdots&n-2&n-1\\1&0&1&\cdots&n-3&n-2\\2&1&0&\cdots&n-4&n-3\\\vdots&\vdots&\vdots&&\vdots&\vdots\\n-1&n-2&n-3&\cdots&1&0\end{vmatrix}\xlongequal[i=n,n-1,\cdots,2]{-j_{i-1}+j_i}\begin{vmatrix}0&1&1&\cdots&1&1\\1&-1&1&\cdots&1&1\\2&-1&-1&\cdots&1&1\\\vdots&\vdots&\vdots&&\vdots&\vdots\\n-1&-1&-1&\cdots&-1&-1\end{vmatrix}$$

$$\xlongequal[i=n-1,n-2,\cdots,1]{r_n+r_i}\begin{vmatrix}n-1&0&0&\cdots&0&0\\n&-2&0&\cdots&0&0\\n+1&-2&-2&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\n-1&-1&-1&\cdots&-1&-1\end{vmatrix}=(-2)^{n-2}(n-1).$$

□

例题 1.42 求下列 n 阶行列式的值:

$$|\boldsymbol{A}|=\begin{vmatrix}1&x_1(x_1-a)&x_1^2(x_1-a)&\cdots&x_1^{n-1}(x_1-a)\\1&x_2(x_2-a)&x_2^2(x_2-a)&\cdots&x_2^{n-1}(x_2-a)\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&x_n(x_n-a)&x_n^2(x_n-a)&\cdots&x_n^{n-1}(x_n-a)\end{vmatrix}.$$

笔记 当行列式的行或列有一定的规律性时, 但是由于缺少一行或一列导致这个行列式行或列的规律性并不完整. 此时我们可以尝试**升阶法**补全这个行列式行或列的规律, 再对行列式进行化简.

本题若直接使用**大拆分法**会得到比较多的行列式, 而且每个行列式并不是完整的 *Vandermode* 行列式. 后续求解很繁琐, 因此不采取**大拆分法**.

解 (升阶法) 考虑 $n+1$ 阶行列式 $|\boldsymbol{B}|=\begin{vmatrix}1 & x_1-a & x_1(x_1-a) & x_1^2(x_1-a) & \cdots & x_1^{n-1}(x_1-a)\\ 1 & x_2-a & x_2(x_2-a) & x_2^2(x_2-a) & \cdots & x_2^{n-1}(x_2-a)\\ \vdots & \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n-a & x_n(x_n-a) & x_n^2(x_n-a) & \cdots & x_n^{n-1}(x_n-a)\\ 1 & y-a & y(y-a) & y^2(y-a) & \cdots & y^{n-1}(y-a)\end{vmatrix}$, 则

$$|\boldsymbol{B}|=\begin{vmatrix}1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^n\\ 1 & x_2 & x_2^2 & x_2^3 & \cdots & x_2^n\\ \vdots & \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^n\\ 1 & y & y^2 & y^3 & \cdots & y^n\end{vmatrix}=\prod_{k=1}^{n}(y-x_k)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

由上式可知,$|\boldsymbol{B}|$ 可以看作一个关于 y 的 n 次多项式. 将 $|\boldsymbol{B}|$ 按最后一行展开得到

$$|\boldsymbol{B}|=\sum_{i=1}^{n+1}(-1)^{n+i}B_{n+1,i}y^{i-1}, \text{其中} B_{ni} \text{是} |\boldsymbol{B}| \text{的第} (n+1,i) \text{元的余子式}, i=1,2,\cdots,n+1.$$

从而

$$|\boldsymbol{B}|=(-1)^{n+2}B_{n+1,1}+\sum_{i=2}^{n+1}(-1)^{n+i+1}B_{n+1,i}y^{i-2}(y-a)=\prod_{k=1}^{n}(y-x_k)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i). \tag{1.15}$$

又易知 $B_{n+1,2}=|\boldsymbol{A}|$, 而当 $a=0$ 时, 由等式(1.15)可知,$|\boldsymbol{B}|$ 中 y 前面的系数只有 $B_{n+1,2}$. 比较等式(1.15)两边 y 的系数可得

$$(-1)^{n+3}|\boldsymbol{A}|=(-1)^{n+3}B_{n+1,2}=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)\left(\sum_{i=1}^{n}(-x_1)\cdots(-x_{i-1})(-x_{i+1})\cdots(-x_n)\right).$$

于是 $|\boldsymbol{A}|=(-1)^{n+3}(-1)^{n-1}\prod\limits_{1\leqslant i<j\leqslant n}(x_j-x_i)\left(\sum\limits_{i=1}^{n}x_1\cdots x_{i-1}x_{i+1}\cdots x_n\right)=\prod\limits_{1\leqslant i<j\leqslant n}(x_j-x_i)\left(\sum\limits_{i=1}^{n}x_1\cdots x_{i-1}x_{i+1}\cdots x_n\right).$

当 $a\neq 0$ 时, 由等式(1.15)可知,$|\boldsymbol{B}|$ 中 y 前面的系数不只有 $B_{n+1,2}$, 但是, 我们比较等式(1.15)两边的常数项可得

$$(-1)^{n+2}B_{n+1,1}-a(-1)^{n+3}B_{n+1,2}=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)\prod_{k=1}^{n}(-x_k). \tag{1.16}$$

又因为

$$\begin{aligned}B_{n+1,1}&=\begin{vmatrix}x_1-a & x_1(x_1-a) & x_1^2(x_1-a) & \cdots & x_1^{n-1}(x_1-a)\\ x_2-a & x_2(x_2-a) & x_2^2(x_2-a) & \cdots & x_2^{n-1}(x_2-a)\\ \vdots & \vdots & \vdots & & \vdots\\ x_n-a & x_n(x_n-a) & x_n^2(x_n-a) & \cdots & x_n^{n-1}(x_n-a)\end{vmatrix}\\ &=\prod_{i=1}^{n}(x_i-a)\begin{vmatrix}1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^{n-1}\\ 1 & x_2 & x_2^2 & x_2^3 & \cdots & x_2^{n-1}\\ \vdots & \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^{n-1}\end{vmatrix}=\prod_{i=1}^{n}(x_i-a)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).\end{aligned}$$

所以再结合等式(1.16)可得

$$\begin{aligned}-a(-1)^{n+3}|\boldsymbol{A}|=-a(-1)^{n+3}B_{n+1,2}&=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)\prod_{k=1}^{n}(-x_k)-(-1)^{n+2}B_{n+1,1}\\ &=(-1)^n\prod_{k=1}^{n}x_k\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)+(-1)^{n+1}\prod_{i=1}^{n}(x_i-a)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)\end{aligned}$$

$$= (-1)^n \prod_{1\leqslant i<j\leqslant n}(x_j - x_i)\left[\prod_{k=1}^{n} x_k - \prod_{i=1}^{n}(x_i - a)\right].$$

故此时 $|\boldsymbol{A}| = \prod\limits_{1\leqslant i<j\leqslant n}(x_j - x_i)\left(\prod\limits_{k=1}^{n} x_k - \prod\limits_{i=1}^{n}(x_i - a)\right)$. □

例题 1.43 求下列行列式式的值 (n 为偶数)

$$I = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2^2 & \cdots & 2^n & 2^{n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n & n^2 & \cdots & n^n & n^{n+1} \\ \frac{n}{2} & \frac{n^2}{3} & \cdots & \frac{n^n}{n+1} & \frac{n^{n+1}}{n+2} \end{vmatrix}.$$

笔记 应用行列式函数求导求行列式的值.

解 令 $G(x) = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2^2 & \cdots & 2^n & 2^{n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n & n^2 & \cdots & n^n & n^{n+1} \\ \frac{x^2}{2} & \frac{x^3}{3} & \cdots & \frac{x^{n+1}}{n+1} & \frac{x^{n+2}}{n+2} \end{vmatrix}$, 则 $I = \dfrac{G(n)}{n}$ 且 $G(0) = 0$. 利用行列式求导公式, 可得

$$G'(x) = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2^2 & \cdots & 2^n & 2^{n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n & n^2 & \cdots & n^n & n^{n+1} \\ x & x^2 & \cdots & x^n & x^{n+1} \end{vmatrix} = n!x\begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 2^{n-1} & 2^n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & n & \cdots & n^{n-1} & n^n \\ 1 & x & \cdots & x^{n-1} & x^n \end{vmatrix} = n!\prod_{1\leqslant i<j\leqslant n}(j-i)\prod_{k=0}^{n}(x-k).$$

因此

$$\begin{aligned} I &= \frac{G(n)}{n} = \frac{\int_0^n G'(x)dx}{n} = (n-1)!\prod_{1\leqslant i<j\leqslant n}(j-i)\int_0^n \prod_{k=0}^{n}(x-k)dx \\ &\overset{\text{区间再现}}{=} (n-1)!\prod_{1\leqslant i<j\leqslant n}(j-i)\int_0^n \prod_{k=0}^{n}(n-k-x)dx \\ &= (-1)^{n+1}(n-1)!\prod_{1\leqslant i<j\leqslant n}(j-i)\int_0^n \prod_{k=0}^{n}(x-k)dx \\ &= (-1)^{n+1}I. \end{aligned}$$

由于 n 为偶数, 所以 $(-1)^{n+1} = -1$. 于是 $I = -I$. 故 $I = 0$. □

第二章 矩阵

2.1 矩阵的运算

命题 2.1 (标准单位向量和基础矩阵)

1. 标准单位向量

n 维标准单位列向量是指下列 n 个 n 维列向量:

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

向量组 $\boldsymbol{e}_1',\boldsymbol{e}_2',\cdots,\boldsymbol{e}_n'$ 则被称为 n 维标准单位行向量, 容易验证标准单位向量有下列基本性质:

1. 若 $i\neq j$, 则 $\boldsymbol{e}_i'\boldsymbol{e}_j=0$, 而 $\boldsymbol{e}_i'\boldsymbol{e}_i=1$;
2. 若 $A=(a_{ij})$ 是 $m\times n$ 矩阵, 则 $A\boldsymbol{e}_i$ 是 A 的第 i 个列向量;$\boldsymbol{e}_i'A$ 是 A 的第 i 个行向量;
3. 若 $A=(a_{ij})$ 是 $m\times n$ 矩阵, 则 $\boldsymbol{e}_i'A\boldsymbol{e}_j=a_{ij}$;
4. **判定准则:** 设 A,B 都是 $m\times n$ 矩阵, 则 $A=B$ 当且仅当 $A\boldsymbol{e}_i=B\boldsymbol{e}_i(1\leq i\leq n)$ 成立, 也当且仅当 $\boldsymbol{e}_i'A=\boldsymbol{e}_i'B(1\leq i\leq m)$ 成立.

2. 基础矩阵

n 阶基础矩阵 (又称初级矩阵) 是指 n^2 个 n 阶矩阵 $\{E_{ij},1\leq i,j\leq n\}$. 这里 E_{ij} 是一个 n 阶矩阵, 它的第 (i,j) 元素等于 1, 其他元素全为 0. 基础矩阵也可以看成是标准单位向量的积:$E_{ij}=\boldsymbol{e}_i\boldsymbol{e}_j^T$. 由此不难证明基础矩阵的下列性质:

1. 若 $j\neq k$, 则 $E_{ij}E_{kl}=0$;
2. 若 $j=k$, 则 $E_{ij}E_{kl}=E_{il}$;
3. 若 A 是 n 阶矩阵且 $A=(a_{ij})$, 则 $A=\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}a_{ij}E_{ij}$;
4. 若 A 是 n 阶矩阵且 $A=(a_{ij})$, 则 $E_{ij}A$ 的第 i 行是 A 的第 j 行,$E_{ij}A$ 的其他行全为零;
5. 若 A 是 n 阶矩阵且 $A=(a_{ij})$, 则 AE_{ij} 的第 j 列是 A 的第 i 列,AE_{ij} 的其他列全为零;
6. 若 A 是 n 阶矩阵且 $A=(a_{ij})$, 则 $E_{ij}AE_{kl}=a_{jk}E_{il}$.

笔记 标准单位向量和基础矩阵虽然很简单, 但如能灵活应用就可以得到意外的结果. 我们在今后将经常应用它们, 因此请读者熟记这些结论.

一些常见的想法:

1. 可以将一般的矩阵写成标准单位列向量或基础矩阵的形式 (这个形式可以是和式的形式, 也可以是分块的形式).

2. 如果要证明两个矩阵相等, 那么我们就可以考虑判定法则.

3. 如果某种等价关系蕴含了一种递减的规律 (项数减少, 阶数降低等), 那么我们就可以考虑数学归纳法, 去尝试根据这个规律得到一些结论.

定义 2.1 (循环矩阵)

1. 下列形状的 n 阶矩阵称为 n 阶基础循环矩阵:

$$\boldsymbol{J}=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\\1&0&0&\cdots&0\end{pmatrix}=\begin{pmatrix}O&I_{n-1}\\1&O\end{pmatrix}.$$

2. 下列形状的矩阵称为循环矩阵:

$$\begin{pmatrix}a_1&a_2&a_3&\cdots&a_n\\a_n&a_1&a_2&\cdots&a_{n-1}\\a_{n-1}&a_n&a_1&\cdots&a_{n-2}\\\vdots&\vdots&\vdots&&\vdots\\a_2&a_3&a_4&\cdots&a_1\end{pmatrix}.$$

♣

笔记 记 $C_n(\mathbb{K})$ 为 $\mathbb{K}$ 上所有 n 阶循环矩阵构成的集合.

命题 2.2 (循环矩阵的性质)

1. 若 $\boldsymbol{J}$ 为 n 阶基础循环矩阵, 则

$$\boldsymbol{J}^k=\begin{pmatrix}O&I_{n-k}\\I_k&O\end{pmatrix},1\le k\le n.$$

2. 若 A 是循环矩阵, 即

$$A=\begin{pmatrix}a_1&a_2&a_3&\cdots&a_n\\a_n&a_1&a_2&\cdots&a_{n-1}\\a_{n-1}&a_n&a_1&\cdots&a_{n-2}\\\vdots&\vdots&\vdots&&\vdots\\a_2&a_3&a_4&\cdots&a_1\end{pmatrix},$$

   则循环矩阵 A 可以表示为基础循环矩阵 J 的多项式:

$$A=a_1I_n+a_2J+a_3J^2+\cdots+a_nJ^{n-1}.$$

   反之, 若一个矩阵能表示为基础循环矩阵 J 的多项式, 则它必是循环矩阵.
3. 同阶循环矩阵之积仍是循环矩阵.
4. 基础循环矩阵 $\boldsymbol{J}=\begin{pmatrix}O&I_{n-k}\\I_k&O\end{pmatrix}$ $(1\le k\le n)$ 的逆仍是循环矩阵, 并且

$$\boldsymbol{J}^{-1}=\begin{pmatrix}O&I_k\\I_{n-k}&O\end{pmatrix},1\le k\le n.$$

♠

笔记 循环矩阵的性质及应用详见谢启鸿博客.

证明

1. 将 $\boldsymbol{J}$ 写作 $(e_n,e_1,\cdots,e_{n-1})$, 其中 e_i 是标准单位列向量 $(i=1,2,\cdots,n)$. 由分块矩阵乘法并注意到 $\boldsymbol{J}e_i$ 就是 $\boldsymbol{J}$ 的第 i 列, 可得

$$\boldsymbol{J}^2=\boldsymbol{J}\left(e_n,e_1,\cdots,e_{n-1}\right)=\left(\boldsymbol{J}e_n,\boldsymbol{J}e_1,\cdots,\boldsymbol{J}e_{n-1}\right)=\left(e_{n-1},e_n,\cdots,e_{n-2}\right).$$

   不断这样做下去就可以得到结论.
2. 由循环矩阵和基础循环矩阵的定义和循环矩阵的性质 1容易得到证明.

3. 由循环矩阵的性质 2可知两个循环矩阵之积可写为基础循环矩阵 J 的两个多项式之积. 又由循环矩阵的性质 1, 可知 $J^n = I_n$. 因此两个循环矩阵之积可以表示为基础循环矩阵 J 的多项式, 故由循环矩阵的性质 1即得结论.
4. 利用矩阵初等行变换可得

$$\begin{pmatrix} O & \boldsymbol{I}_{n-k} & \boldsymbol{I}_{n-k} & O \\ \boldsymbol{I}_k & O & O & \boldsymbol{I}_k \end{pmatrix} \to \begin{pmatrix} \boldsymbol{I}_k & O & O & \boldsymbol{I}_k \\ O & \boldsymbol{I}_{n-k} & \boldsymbol{I}_{n-k} & O \end{pmatrix}, 1 \leq k \leq n.$$

从而 $\boldsymbol{J}^{-1} = \begin{pmatrix} O & \boldsymbol{I}_k \\ \boldsymbol{I}_{n-k} & O \end{pmatrix}, 1 \leq k \leq n.$

□

命题 2.3 (循环行列式关于 n 次方根的计算公式)

已知下列循环矩阵 A:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}.$$

则 A 的行列式的值为:

$$|A| = f(\varepsilon_1)f(\varepsilon_2)\cdots f(\varepsilon_n).$$

其中 $f(x) = a_1 + a_2x + a_3x^2 + \cdots + a_nx^{n-1}$,$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 1 的所有 n 次方根.

笔记 关键是要注意到

$$AV = \begin{pmatrix} f(\varepsilon_1) & f(\varepsilon_2) & f(\varepsilon_3) & \cdots & f(\varepsilon_n) \\ \varepsilon_1 f(\varepsilon_1) & \varepsilon_2 f(\varepsilon_2) & \varepsilon_3 f(\varepsilon_3) & \cdots & \varepsilon_n f(\varepsilon_n) \\ \varepsilon_1^2 f(\varepsilon_1) & \varepsilon_2^2 f(\varepsilon_2) & \varepsilon_3^2 f(\varepsilon_3) & \cdots & \varepsilon_n^2 f(\varepsilon_n) \\ \vdots & \vdots & \vdots & & \vdots \\ \varepsilon_1^{n-1} f(\varepsilon_1) & \varepsilon_2^{n-1} f(\varepsilon_2) & \varepsilon_3^{n-1} f(\varepsilon_3) & \cdots & \varepsilon_n^{n-1} f(\varepsilon_n) \end{pmatrix}.$$

然后再利用命题 1.12就能得到分解 $AV = V\Lambda$.

证明 作多项式 $f(x) = a_1 + a_2x + a_3x^2 + \cdots + a_nx^{n-1}$, 令 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 1 的所有 n 次方根. 又令

$$V = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ \varepsilon_1 & \varepsilon_2 & \varepsilon_3 & \cdots & \varepsilon_n \\ \varepsilon_1^2 & \varepsilon_2^2 & \varepsilon_3^2 & \cdots & \varepsilon_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ \varepsilon_1^{n-1} & \varepsilon_2^{n-1} & \varepsilon_3^{n-1} & \cdots & \varepsilon_n^{n-1} \end{pmatrix}, \Lambda = \begin{pmatrix} f(\varepsilon_1) & 0 & 0 & \cdots & 0 \\ 0 & f(\varepsilon_2) & 0 & \cdots & 0 \\ 0 & 0 & f(\varepsilon_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & f(\varepsilon_n) \end{pmatrix}$$

则

$$AV = \begin{pmatrix} f(\varepsilon_1) & f(\varepsilon_2) & f(\varepsilon_3) & \cdots & f(\varepsilon_n) \\ \varepsilon_1 f(\varepsilon_1) & \varepsilon_2 f(\varepsilon_2) & \varepsilon_3 f(\varepsilon_3) & \cdots & \varepsilon_n f(\varepsilon_n) \\ \varepsilon_1^2 f(\varepsilon_1) & \varepsilon_2^2 f(\varepsilon_2) & \varepsilon_3^2 f(\varepsilon_3) & \cdots & \varepsilon_n^2 f(\varepsilon_n) \\ \vdots & \vdots & \vdots & & \vdots \\ \varepsilon_1^{n-1} f(\varepsilon_1) & \varepsilon_2^{n-1} f(\varepsilon_2) & \varepsilon_3^{n-1} f(\varepsilon_3) & \cdots & \varepsilon_n^{n-1} f(\varepsilon_n) \end{pmatrix} = V\Lambda$$

从而 $V^{-1}AV = \Lambda$, 又因为 ε_i 互不相同, 所以 $|V| \neq 0$, 故

$$|A| = |V^{-1}AV| = |\Lambda| = f(\varepsilon_1)f(\varepsilon_2)\cdots f(\varepsilon_n).$$

□

命题 2.4 (b-循环矩阵)

设 b 为非零常数, 下列形状的矩阵称为 b -循环矩阵:

$$A=\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ ba_n & a_1 & a_2 & \cdots & a_{n-1} \\ ba_{n-1} & ba_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ ba_2 & ba_3 & ba_4 & \cdots & a_1 \end{pmatrix}$$

(1) 证明: 同阶 b -循环矩阵的乘积仍然是 b-循环矩阵;

(2) 求上述 b-循环矩阵 A 的行列式的值.

证明

(1) (证明类似于循环矩阵的性质3.) 设 $J_b=\begin{pmatrix} O & I_{n-1} \\ b & O \end{pmatrix}$, 则 $J_b^k=\begin{pmatrix} O & I_{n-k} \\ bI_k & O \end{pmatrix}$, $0\le k\le n-1$. 从而 $J_b^n=bI_n$ 且 $A=a_1I_n+a_2J_b+a_3J_b^2+\cdots+a_nJ_b^{n-1}$. 因此同阶 $b-$ 循环阵的乘积仍然可以写成 J_b 的 $n-1$ 次多项式, 故同阶 $b-$ 循环阵的乘积仍然是 $b-$ 循环矩阵.

(2) (证明完全类似循环行列式计算公式的证明) 作多项式 $f(x)=a_1+a_2x+a_3x^2+\cdots+a_nx^{n-1}$, 令 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 b 的所有 n 次方根. 又令

$$V=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ \varepsilon_1 & \varepsilon_2 & \varepsilon_3 & \cdots & \varepsilon_n \\ \varepsilon_1^2 & \varepsilon_2^2 & \varepsilon_3^2 & \cdots & \varepsilon_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ \varepsilon_1^{n-1} & \varepsilon_2^{n-1} & \varepsilon_3^{n-1} & \cdots & \varepsilon_n^{n-1} \end{pmatrix},\Lambda=\begin{pmatrix} f(\varepsilon_1) & 0 & 0 & \cdots & 0 \\ 0 & f(\varepsilon_2) & 0 & \cdots & 0 \\ 0 & 0 & f(\varepsilon_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & f(\varepsilon_n) \end{pmatrix}$$

则

$$AV=\begin{pmatrix} f(\varepsilon_1) & f(\varepsilon_2) & f(\varepsilon_3) & \cdots & f(\varepsilon_n) \\ \varepsilon_1f(\varepsilon_1) & \varepsilon_2f(\varepsilon_2) & \varepsilon_3f(\varepsilon_3) & \cdots & \varepsilon_nf(\varepsilon_n) \\ \varepsilon_1^2f(\varepsilon_1) & \varepsilon_2^2f(\varepsilon_2) & \varepsilon_3^2f(\varepsilon_3) & \cdots & \varepsilon_n^2f(\varepsilon_n) \\ \vdots & \vdots & \vdots & & \vdots \\ \varepsilon_1^{n-1}f(\varepsilon_1) & \varepsilon_2^{n-1}f(\varepsilon_2) & \varepsilon_3^{n-1}f(\varepsilon_3) & \cdots & \varepsilon_n^{n-1}f(\varepsilon_n) \end{pmatrix}=V\Lambda$$

从而 $V^{-1}AV=\Lambda$, 又因为 ε_i 互不相同, 所以 $|V|\neq 0$, 故

$$|A|=|V^{-1}AV|=|\Lambda|=f(\varepsilon_1)f(\varepsilon_2)\cdots f(\varepsilon_n).$$

□

命题 2.5 (基础幂零 Jordan 块)

设 n 阶基础幂零 Jordan 块

$$\boldsymbol{A}=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

则

$$A^k = \begin{pmatrix} O & I_{n-k} \\ O & O \end{pmatrix}, 1 \le k \le n.$$

证明 将 A 写为 $A = (0, \mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_{n-1})$, 其中 $\mathbf{e}_i$ 是标准单位列向量. 由分块矩阵乘法并注意 $A\mathbf{e}_i$ 就是 A 的第 i 列, 因此

$$A^2 = (0, A\mathbf{e}_1, A\mathbf{e}_2, \cdots, A\mathbf{e}_{n-1}) = (0, 0, \mathbf{e}_1, \cdots, \mathbf{e}_{n-2})$$

不断这样做下去就可得到结论. □

例题 2.1 设 A 是 n 阶矩阵,A 适合 $A^n = O$ 时,$I_n - A$ 必是可逆矩阵.

证明 注意到

$$I_n = I_n - A^n = (I_n - A)\left(I_n + A + A^2 + \cdots + A^{n-1}\right).$$

故此时 $I_n - A$ 必是可逆矩阵. □

例题 2.2 设 A 是 n 阶矩阵,A 适合 $AB = B(I_n - A)$ 对任意 n 阶矩阵 B 成立, 那么 $B = O$.

笔记 若已知矩阵乘法的相关等式, 可以尝试得到一些递推等式.

证明 假设 $A^k = O$, 其中 k 为某个正整数. 由条件可得 $AB = B(I_n - A)$, 于是 $O = A^k B = B(I_n - A)^k$. 由上一题知 $I_n - A$ 是可逆矩阵, 从而 $B = O$. □

命题 2.6 (多项式的友矩阵和 Frobenius 块)

设首一多项式 $f(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$,$f(x)$ 的友阵

$$C(f(x)) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & -a_n \\ 1 & 0 & \cdots & 0 & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & 0 & -a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & -a_2 \\ 0 & 0 & \cdots & 0 & 1 & -a_1 \end{pmatrix},$$

则 $|xI_n - C(f(x))| = f(x)$.

$C(f(x))$ 的转置 $F(f(x))$ 称为 $f(x)$ 的 Frobenius 块. 即

$$C^T(f(x)) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_2 & -a_1 \end{pmatrix}.$$

并且容易验证 $C(f(x))$ 具有以下性质, 其中 e_i 是标准单位列向量 $(i = 1, 2, \cdots, n)$:

$$C(f(x))e_i = e_{i+1}\ (1 \le i \le n-1),\ C(f(x))e_n = -\sum_{i=1}^{n} a_{n-i+1}e_i.$$

证明 $|x\boldsymbol{I}_n - \boldsymbol{C}(f(x))| = f(x)$ 的证明见友矩阵的特征多项式. □

例题 2.3 求下列矩阵的逆矩阵 ($a_n \neq 0$):

$$F = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}.$$

解 用初等变换法不难求得

$$F^{-1} = \begin{pmatrix} -\dfrac{a_{n-1}}{a_n} & 1 & 0 & \cdots & 0 \\ -\dfrac{a_{n-2}}{a_n} & 0 & 1 & \cdots & 0 \\ -\dfrac{a_{n-3}}{a_n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\dfrac{1}{a_n} & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

□

命题 2.7

和所有 n 阶对角矩阵乘法可交换的矩阵必是对角矩阵.

证明 由矩阵乘法易得. □

命题 2.8 (纯量矩阵的刻画)

(1) 和所有 n 阶奇异阵乘法可交换的矩阵必是纯量阵 kI_n.
(2) 和所有 n 阶非奇异阵乘法可交换的矩阵必是纯量阵 kI_n.
(3) 和所有 n 阶正交阵乘法可交换的矩阵必是纯量阵 kI_n.
(4) 和所有 n 阶矩阵乘法可交换的矩阵必是纯量阵 kI_n.

证明 首先设 $A = (a_{ij})_{n\times n}$.

1. 设 $E_{ij}(1 \le i \neq j \le n)$ 为基础矩阵, 因为基础矩阵都是奇异阵, 所以由条件可知 $E_{ij}A = AE_{ij}$. 注意到 $E_{ij}A$ 是将 A 的第 j 行变为第 i 行而其他行都是零的 n 阶矩阵,AE_{ij} 是将 A 的第 i 列变为第 j 列而其他列都是零的 n 阶矩阵, 于是我们有

$$\begin{array}{c} \\ i \end{array}\!\!\begin{array}{c} \begin{array}{ccccc} \phantom{a_{j1}} & \phantom{a_{j2}} & & j & \phantom{\cdots\ a_{jn}} \end{array} \\ \begin{pmatrix} & & & & & \\ a_{j1} & a_{j2} & \cdots & a_{jj} & \cdots & a_{jn} \\ & & & & & \end{pmatrix} \end{array} = \begin{array}{c} \\ \\ \\ i \\ \\ \\ \end{array}\!\!\begin{array}{c} j \\ \begin{pmatrix} & a_{1i} & \\ & a_{2i} & \\ & \vdots & \\ & a_{ii} & \\ & \vdots & \\ & a_{ni} & \end{pmatrix} \end{array}.$$

从而比较上述等式两边矩阵的每个元素可得 $a_{ij} = 0(i \neq j)$,$a_{ii} = a_{jj}(1 \le i \neq j \le n)$, 因此 A 是纯量阵.

2. 设 $D = \mathrm{diag}\{1, 2, \cdots, n\}$ 为对角阵, 因为 D 为非奇异阵, 所以由条件可知 $AD = DA$. 进而

$$AD = DA$$
$$\Leftrightarrow A(\boldsymbol{e}_1, 2\boldsymbol{e}_2, \cdots, n\boldsymbol{e}_n) = (\boldsymbol{e}_1, 2\boldsymbol{e}_2, \cdots, n\boldsymbol{e}_n)A$$
$$\Leftrightarrow (A\boldsymbol{e}_1, 2A\boldsymbol{e}_2, \cdots, nA\boldsymbol{e}_n) = (\boldsymbol{e}_1 A, 2\boldsymbol{e}_2 A, \cdots, n\boldsymbol{e}_n A)$$

$$\Leftrightarrow\begin{pmatrix} a_{11} & 2a_{12} & \cdots & na_{1n} \\ a_{21} & 2a_{22} & \cdots & na_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & 2a_{n2} & \cdots & na_{nn} \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 2a_{21} & 2a_{22} & \cdots & 2a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ na_{n1} & na_{n2} & \cdots & na_{nn} \end{pmatrix}.$$

比较上述等式两边矩阵的每个元素可得 $ja_{ij}=ia_{ij}\,(i\neq j)$, 从而 $(i-j)\,a_{ij}=0\,(i\neq j)$, 于是 $a_{ij}=0\,(i\neq j)$. 故 $A=\mathrm{diag}\{a_{11},a_{22},\cdots,a_{nn}\}$ 也为对角阵.

设 $P_{ij}(1\le i\neq j\le n)$ 为第一类初等阵, 因为第一类初等阵均为非奇异阵, 所以由条件可知 $AP_{ij}=P_{ij}A$. 进而可得

$$\begin{array}{cc} & \begin{matrix} \phantom{a_{11}} & & i & & j & & \phantom{a_{nn}} \end{matrix} \\ \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} & \begin{pmatrix} a_{11} & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & a_{jj} & & \\ & & \vdots & \ddots & \vdots & & \\ & & a_{ii} & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & a_{nn} \end{pmatrix} \end{array}=\begin{array}{cc} & \begin{matrix} \phantom{a_{11}} & & i & & j & & \phantom{a_{nn}} \end{matrix} \\ \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} & \begin{pmatrix} a_{11} & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & a_{ii} & & \\ & & \vdots & \ddots & \vdots & & \\ & & a_{jj} & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & a_{nn} \end{pmatrix} \end{array}.$$

从而比较上述等式两边矩阵的每个元素可得 $a_{ii}=a_{jj}(1\le i\neq j\le n)$, 于是 A 为纯量阵.

3. 设第二类初等阵 $P_i(-1)(1\le i\le n)$, 因为 $P_i(-1)(1\le i\le n)$ 都是正交阵, 所以由条件可知 $P_i(-1)A=AP_i(-1)$. 进而可得

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ -a_{i1} & -a_{i2} & \cdots & -a_{ii} & \cdots & -a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix}=\begin{pmatrix} a_{11} & \cdots & -a_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & -a_{2i} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & -a_{ii} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & -a_{ni} & \cdots & a_{nn} \end{pmatrix}.$$

比较上述等式两边矩阵的每个元素可得 $a_{ij}=-a_{ij}\,(i\neq j)$, 从而 $a_{ij}=0\,(i\neq j)$. 于是 $A=\mathrm{diag}\{a_{11},a_{22},\cdots,a_{nn}\}$ 为对角阵.

设 $P_{ij}(1\le i\neq j\le n)$ 为第一类初等阵, 因为第一类初等阵均为正交阵, 所以由条件可知 $AP_{ij}=P_{ij}A$. 进而可得

$$\begin{array}{cc} & \begin{matrix} \phantom{a_{11}} & & i & & j & & \phantom{a_{nn}} \end{matrix} \\ \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} & \begin{pmatrix} a_{11} & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & a_{jj} & & \\ & & \vdots & \ddots & \vdots & & \\ & & a_{ii} & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & a_{nn} \end{pmatrix} \end{array}=\begin{array}{cc} & \begin{matrix} \phantom{a_{11}} & & i & & j & & \phantom{a_{nn}} \end{matrix} \\ \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} & \begin{pmatrix} a_{11} & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & a_{ii} & & \\ & & \vdots & \ddots & \vdots & & \\ & & a_{jj} & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & a_{nn} \end{pmatrix} \end{array}.$$

从而比较上述等式两边矩阵的每个元素可得 $a_{ii}=a_{jj}(1\le i\neq j\le n)$, 于是 A 为纯量阵.

4. 可以由上面 (1)(2)(3) 中任意一个证明得到. 注意如果此时用 (3) 的证明方法, 那么我们可以先考虑 A 与第一

类初等矩阵 $P_i(c)(c \neq 1, 1 \le i \le n)$ 的乘法交换性. 而不是像 (3) 中只能考虑 $P_i(-1)(1 \le i \le n)$.

□

命题 2.9 (零矩阵的充要条件)

1. $m \times n$ 实矩阵 $A = O$ 的充要条件是适合条件 $AA' = O$ 或 $\mathrm{tr}(AA') \ge 0$, 等号成立;
2. $m \times n$ 复矩阵 $A = O$ 的充要条件是适合条件 $A\overline{A}' = O$ 或 $\mathrm{tr}(A\overline{A}') \ge 0$, 等号成立.

证明

1. (1) 设 $A = (a_{ij})_{m\times n}$, 则 AA' 的第 (i,i) 元素等于零, 即

$$a_{i1}^2 + a_{i2}^2 + \cdots + a_{in}^2 = 0, i = 1, 2, \cdots, m.$$

又因为 a_{ij} 都是实数, 所以必有 $a_{ij} = 0, i = 1, 2, \cdots, m, j = 1, 2, \cdots, n$. 故 $A = O$.

(2) 设 $A = (a_{ij})$ 为 $m \times n$ 实矩阵, 则通过计算可得

$$\mathrm{tr}(AA') = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2 \ge 0,$$

等号成立当且仅当 $a_{ij} = 0(1 \le i \le m, 1 \le j \le n)$, 即 $A = O$.

2. (1) 设 $A = (a_{ij})_{m\times n}$, 则 $A\overline{A}'$ 的第 (i,i) 元素等于零, 即

$$|a_{i1}|^2 + |a_{i2}|^2 + \cdots + |a_{in}|^2 = 0, i = 1, 2, \cdots, m.$$

又因为 a_{ij} 都是复数, 所以可设 $a_{ij} = b_{ij} + \mathrm{i}c_{ij}$, 其中 $b_{ij}, c_{ij} \in \mathbb{R}, i = 1, 2, \cdots, m, j = 1, 2, \cdots, n$. 于是

$$b_{i1}^2 + c_{i1}^2 + b_{i2}^2 + c_{i2}^2 + \cdots + b_{in}^2 + c_{in}^2 = 0, i = 1, 2, \cdots, m.$$

再结合 $b_{ij}, c_{ij} \in \mathbb{R}$, 可知 $b_{ij} = c_{ij} = 0$. 即 $a_{ij} = 0, i = 1, 2, \cdots, m, j = 1, 2, \cdots, n$. 故 $A = O$.

(2) 设 $A = (a_{ij})$ 为 $m \times n$ 复矩阵, 则通过计算可得

$$\mathrm{tr}(A\overline{A}') = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2 \ge 0,$$

等号成立当且仅当 $a_{ij} = 0(1 \le i \le m, 1 \le j \le n)$, 即 $A = O$.

□

命题 2.10 (对称阵是零矩阵的充要条件)

设 $\boldsymbol{A}$ 为 n 阶对称阵, 则 $\boldsymbol{A}$ 是零矩阵的充要条件是对任意的 n 维列向量 $\boldsymbol{\alpha}$, 有

$$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} = 0.$$

证明 只要证明充分性. 设 $\boldsymbol{A} = (a_{ij})$, 令 $\alpha = \boldsymbol{e}_i$, 是第 i 个标准单位列向量. 因为 $\boldsymbol{e}_i'\boldsymbol{A}\boldsymbol{e}_i$ 是 $\boldsymbol{A}$ 的第 (i,i) 元素, 故 $a_{ii} = 0$. 又令 $\boldsymbol{\alpha} = \boldsymbol{e}_i + \boldsymbol{e}_j (i \neq j)$, 则

$$0 = (\boldsymbol{e}_i + \boldsymbol{e}_j)'\boldsymbol{A}(\boldsymbol{e}_i + \boldsymbol{e}_j) = a_{ii} + a_{jj} + a_{ij} + a_{ji}.$$

由于 $\boldsymbol{A}$ 是对称阵, 故 $a_{ij} = a_{ji}$, 又上面已经证明 $a_{ii} = a_{jj} = 0$, 从而 $a_{ij} = 0$, 这就证明了 $\boldsymbol{A} = \boldsymbol{O}$. □

命题 2.11 (反对称阵的刻画)

设 $\boldsymbol{A}$ 为 n 阶方阵, 则 $\boldsymbol{A}$ 是反称阵的充要条件是对任意的 n 维列向量 $\boldsymbol{\alpha}$, 有

$$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} = 0.$$

证明 必要性 ($\Rightarrow$): 若 $\boldsymbol{A}$ 是反称阵, 则对任意的 n 维列向量 $\boldsymbol{\alpha}$, 有 $(\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha})' = -\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}$. 而 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}$ 是数, 因此 $(\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha})' = \boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}$. 比较上面两个式子便有 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} = 0$.

充分性 ($\Leftarrow$): 若上式对任意的 n 维列向量 $\boldsymbol{\alpha}$ 成立, 则由 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}$ 是数, 可知 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} = \left(\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}\right)' = \boldsymbol{\alpha}'\boldsymbol{A}'\boldsymbol{\alpha} = 0$, 故 $\boldsymbol{\alpha}'(\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{\alpha} = 0$. 因为矩阵 $\boldsymbol{A} + \boldsymbol{A}'$ 是对称阵, 故由对称阵是零矩阵的充要条件可得 $\boldsymbol{A} + \boldsymbol{A}' = \boldsymbol{O}$, 即 $\boldsymbol{A}' = -\boldsymbol{A}$,$\boldsymbol{A}$ 是反称阵. □

命题 2.12

任一 n 阶方阵均可表示为一个对称阵与一个反对称阵之和.

笔记 构造思路: 设 $A = B + C$, 且 B 为对称矩阵,C 为反称矩阵. 则两边取转置可得

$$\begin{cases} A = B + C \\ A' = (B + C)' = B - C \end{cases}$$

解得:$B = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}'), C = \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}')$.

证明 设 $\boldsymbol{A}$ 是 n 阶方阵, 则 $\boldsymbol{A} + \boldsymbol{A}'$ 是对称阵,$\boldsymbol{A} - \boldsymbol{A}'$ 是反对称阵, 并且

$$\boldsymbol{A} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}') + \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}').$$

□

注 上例中的 $\frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}')$ 称为 $\boldsymbol{A}$ 的对称化,$\frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}')$ 称为 $\boldsymbol{A}$ 的反对称化.

命题 2.13 (上三角阵性质)

(1) 设 $\boldsymbol{A}$ 是 n 阶上三角阵且主对角线上元素全为零, 则 $\boldsymbol{A}^n = O$.

(2) 设 $\boldsymbol{A}$ 是 $n(n \geq 2)$ 阶上三角阵, 若 $i < j$, 则 $A_{ij} = M_{ij} = 0$.

(3) 上 (下) 三角阵的加减、数乘、乘积 (幂)、多项式、伴随和求逆仍然是上 (下) 三角阵, 并且所得上 (下) 三角阵的主对角元是原上 (下) 三角阵对应主对角元的加减、数乘、乘积 (幂)、多项式、伴随和求逆.

证明 (1) 证法一 (抽屉原理): 设 $A = (a_{ij})$, 当 $i \geq j$ 时,$a_{ij} = 0$. 将 A 表示为基础矩阵 E_{ij} 之和:

$$A = \sum_{i>j} a_{ij} E_{ij}$$

因为当 $j \neq k$ 时,$E_{ij}E_{kl} = \boldsymbol{O}$, 故在 A^n 的乘法展开式中, 可能非零的项只能是具有形式 $E_{i_1j_1}E_{i_2j_2}\cdots E_{i_{n-1}j_{n-1}}$, 但足标必须满足条件 $1 \leq i_1 < j_1 < i_2 < j_2 < \cdots < j_{n-1} \leq n$. 根据可知, 这样的项也不存在, 因此 $A^n = \boldsymbol{O}$.

证法二 (数学归纳法): 由假设 $Ae_i = a_{i1}e_1 + \cdots + a_{i,i-1}e_{i-1}(1 \leq i \leq n)$, 我们只要用归纳法证明:$A^k e_k = 0$ 对任意的 $1 \leq k \leq n$ 都成立, 则 $A^n e_i = A^{n-i} \cdot A^i e_i = A^{n-i} \cdot 0 = 0$ 对任意的 $1 \leq i \leq n$ 都成立, 从而由判定法则可知 $A^n = O$ 成立. 显然,$Ae_1 = 0$ 成立. 假设 $A^k e_k = 0$ 对任意的 $1 < k < n$ 都成立, 则

$$\begin{aligned} A^k e_k &= A^{k-1}(Ae_k) = A^{k-1}(a_{k1}e_1 + \cdots + a_{k,k-1}e_{k-1}) \\ &= a_{k1}A^{k-1}e_1 + \cdots + a_{k,k-1}A^{k-1}e_{k-1} = 0. \end{aligned}$$

(2) 根据条件可设 $|\boldsymbol{A}| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}$, 则当 $i < j$ 时, 有

$$M_{ij} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i} & a_{1,i+1} & \cdots & a_{1j} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2i} & a_{2,i+1} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{ii} & a_{i,i+1} & \cdots & a_{ij} & \cdots & a_{in} \\ 0 & 0 & \cdots & 0 & a_{i+1,i+1} & \cdots & a_{i+1,j} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & a_{nn} \end{vmatrix}$$

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

故 $A_{ij} = M_{ij} = 0$.

(3) 只证上三角阵的情形, 下三角阵的情形完全类似. 上三角阵的加减、数乘、乘积 (幂) 以及多项式结论的证明是显然的. 下面我们来证明伴随和求逆的结论. 设 $\boldsymbol{A} = (a_{ij})$ 为 n 阶上三角阵, 即满足 $a_{ij} = 0, (\forall i > j)$. 由(2)可知 $\boldsymbol{A}$ 的代数余子式 $A_{ij} = 0, \forall i < j$. 于是

$$
\boldsymbol{A}^* = \begin{pmatrix}
A_{11} & A_{21} & \cdots & A_{n1} \\
0 & A_{22} & \cdots & A_{n2} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & A_{nn}
\end{pmatrix}.
$$

故 $\boldsymbol{A}^*$ 也是上三角阵. 而对 $\forall i \in [1,n] \cap N$, 有

我们又将 $A_{ii} = a_{11} \cdots \widehat{a_{ii}} \cdots a_{nn}$ 这个数称为 a_{ii} 的伴随. 这就完成了 $\boldsymbol{A}^*$ 结论的证明.

$$
A_{ii} = (-1)^{2i} M_{ii} = M_{ii} = \begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1,i-1} & a_{1i} & a_{1,i+1} & \cdots & a_{1n} \\
0 & a_{22} & \cdots & a_{2,i-1} & a_{2i} & a_{2,i+1} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & a_{i-1,i-1} & a_{i-1,i} & a_{i-1,i+1} & \cdots & a_{i-1,n} \\
0 & 0 & \cdots & 0 & a_{ii} & a_{i,i+1} & \cdots & a_{in} \\
0 & 0 & \cdots & 0 & 0 & a_{i+1,i+1} & \cdots & a_{i+1,n} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & a_{nn}
\end{vmatrix}
$$

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

由于当 $|\boldsymbol{A}| \neq 0$ 时, 我们有 $\boldsymbol{A}^{-1} = \dfrac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*$, 故由上三角阵的数乘结论可知,$\boldsymbol{A}^{-1}$ 也是上三角阵, 其主对角元为 $\dfrac{1}{|\boldsymbol{A}|}A_{ii} = a_{ii}^{-1}$. 结论得证. □

命题 2.14

若 $\boldsymbol{A},\boldsymbol{B}$ 都是由非负实数组成的矩阵且 AB 有一行等于零, 则或者 $\boldsymbol{A}$ 有一行为零, 或者 $\boldsymbol{B}$ 有一行为零. ♠

证明 设 $\boldsymbol{A}=(a_{ij})_{n\times m}$,$\boldsymbol{B}=(b_{ij})_{m\times s}$. 假设 $\boldsymbol{C}=\boldsymbol{AB}$,$\boldsymbol{C}=(c_{ij})_{n\times s}$ 的第 i 行全为零. 则对 $\forall j\in[1,s]\cap N$, 都有
$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{im}b_{nj}=0.$$
已知对 $\forall i\in[1,n]\cap N,j\in[1,m]\cap N$, 有 $a_{ij}\geq 0$; 对 $\forall i\in[1,m]\cap N,j\in[1,s]\cap N$, 有 $b_{ij}\geq 0$. 从而
$$a_{i1}b_{1j}=a_{i2}b_{2j}=\cdots=a_{im}b_{nj}=0,\forall j\in[1,s]\cap N.$$
若 $\boldsymbol{A}$ 的第 i 行不全为零, 不妨设 $a_{ik}\neq 0,k\in[1,m]\cap N$, 则由 $a_{ik}b_{kj}=0,\forall j\in[1,s]\cap N$ 可得 $b_{kj}=0$, 对 $\forall j\in[1,s]\cap N$ 都成立, 即 $\boldsymbol{B}$ 的第 k 行全为零. □

命题 2.15 (矩阵行和和列和的一种刻画)

(1) n 阶矩阵 A 第 i 行元素之和为 $a_i(i=1,2,\cdots,n)$ 当且仅当
$$A\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}.$$
特别地,n 阶矩阵 A 的每一行元素之和等于 c 当且仅当 $A\alpha=c\cdot\alpha$, 其中 $\alpha=(1,1,\cdots,1)'$.

(2) n 阶矩阵 A 第 i 列元素之和为 $a_i(i=1,2,\cdots,n)$ 当且仅当
$$\begin{pmatrix}1&1&\cdots&1\end{pmatrix}A=\begin{pmatrix}a_1&a_2&\cdots&a_n\end{pmatrix}.$$
特别地,n 阶矩阵 A 的每一列元素之和等于 c 当且仅当 $\alpha A=c\cdot\alpha$, 其中 $\alpha=(1,1,\cdots,1)$.

证明 由矩阵乘法容易得到证明. □

例题 2.4 设 n 阶方阵 $\boldsymbol{A}$ 的每一行元素之和等于常数 c, 求证:

(1) 对任意的正整数 k,$\boldsymbol{A}^k$ 的每一行元素之和等于 c^k;

(2) 若 $\boldsymbol{A}$ 为可逆阵, 则 $c\neq 0$ 并且 $\boldsymbol{A}^{-1}$ 的每一行元素之和等于 c^{-1}.

笔记 核心想法是利用命题 2.15.

证明 设 $\boldsymbol{\alpha}=(1,1,\cdots,1)'$, 则由矩阵乘法可知,$\boldsymbol{A}$ 的每一行元素之和等于 c 当且仅当 $\boldsymbol{A\alpha}=c\cdot\boldsymbol{\alpha}$ 成立.

(1) 由 $\boldsymbol{A\alpha}=c\cdot\boldsymbol{\alpha}$ 不断递推可得 $\boldsymbol{A}^k\boldsymbol{\alpha}=c^k\cdot\boldsymbol{\alpha}$, 故结论成立.

(2) 若 $c=0$, 则由 $\boldsymbol{A}$ 可逆以及 $\boldsymbol{A\alpha}=\boldsymbol{0}$ 可得 $\boldsymbol{\alpha}=\boldsymbol{0}$, 矛盾. 在 $\boldsymbol{A\alpha}=c\cdot\boldsymbol{\alpha}$ 的两边同时左乘 $c^{-1}\boldsymbol{A}^{-1}$, 可得 $\boldsymbol{A}^{-1}\boldsymbol{\alpha}=c^{-1}\cdot\boldsymbol{\alpha}$, 由此即得结论. □

命题 2.16 (矩阵可逆的等价命题)

(1)n 阶方阵 $\boldsymbol{A}$ 可逆.

(2) 存在矩阵 $\boldsymbol{B}$, 使得 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}_n$(这个等式同时也说明 $\boldsymbol{B}$ 可逆).

(3)$\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|\neq 0$.

(4)$\boldsymbol{A}$ 等价 (相抵) 于 n 阶单位矩阵.

(5)$\boldsymbol{A}$ 可以表示为有限个初等矩阵的积.

(6)$\boldsymbol{A}$ 的 n 个行向量 (列向量) 线性无关.

命题 2.17

(1) 若已知 $\lambda_1\boldsymbol{A}^2+\lambda_2\boldsymbol{A}+\lambda_3\boldsymbol{I}_n=\boldsymbol{O}$, 其中 $\lambda_1,\lambda_2,\lambda_3\in\mathbb{R}$,$\lambda_1\neq 0$, 并且 $\lambda_1x^2+\lambda_2x+\lambda_3=0$ 无实根 (即原等式左边不可因式分解成 $(a_1\boldsymbol{I}_n+a_2\boldsymbol{A})(b_1\boldsymbol{I}_n+b_2\boldsymbol{A})$), 则对任何 $c,d\in\mathbb{R}$, 都有 $c\boldsymbol{A}+d\boldsymbol{I}_n$ 可逆.

(2) 若已知 $\lambda_1\boldsymbol{A}^2+\lambda_2\boldsymbol{A}+\lambda_3\boldsymbol{I}_n=(a_1\boldsymbol{A}+b_1\boldsymbol{I}_n)(a_2\boldsymbol{A}+b_2\boldsymbol{I}_n)=\boldsymbol{O}$, 其中 $a_1,a_2,b_1,b_2\in\mathbb{R}$, 且 $\lambda_1=a_1a_2,\lambda_2=a_1b_2+a_2b_1,\lambda_3=b_1b_2,a_1\neq 0,a_2\neq 0$. 则对任何实数对 $(c,d)\neq(a_1,b_1),(a_2,b_2)$, 都有 $c\boldsymbol{A}+d\boldsymbol{I}_n$ 可逆.

笔记 **构造逆矩阵的方法:** 不妨设 $k(c\boldsymbol{A}+d\boldsymbol{I}_n)^{-1}=(p\boldsymbol{A}+q\boldsymbol{I}_n)$, 其中 k,p,q 为待定系数. 则
$$(c\boldsymbol{A}+d\boldsymbol{I}_n)\cdot k(c\boldsymbol{A}+d\boldsymbol{I}_n)^{-1}=(c\boldsymbol{A}+d\boldsymbol{I}_n)(p\boldsymbol{A}+q\boldsymbol{I}_n)=pc\boldsymbol{A}^2+(cq+dp)\boldsymbol{A}+dq\boldsymbol{I}_n=k\boldsymbol{I}_n.$$

令 $pc=\lambda_1, cq+dp=\lambda_2$, 则 $p=\frac{\lambda_1}{c}, q=\frac{\lambda_2}{c}-\frac{\lambda_1 d}{c^2}$. 于是由已知条件可得

$$(c\boldsymbol{A}+d\boldsymbol{I}_n)(p\boldsymbol{A}+q\boldsymbol{I}_n)=(c\boldsymbol{A}+d\boldsymbol{I}_n)\left(\frac{\lambda_1}{c}\boldsymbol{A}+\left(\frac{\lambda_2}{c}-\frac{\lambda_1 d}{c^2}\right)\boldsymbol{I}_n\right)=\lambda_1\boldsymbol{A}^2+\lambda_2\boldsymbol{A}+d\left(\frac{\lambda_2}{c}-\frac{\lambda_1 d}{c^2}\right)\boldsymbol{I}_n=\left(\frac{\lambda_2 d}{c}-\frac{\lambda_1 d^2}{c^2}-\lambda_3\right)\boldsymbol{I}_n.$$

从而 $k=\frac{\lambda_2 d}{c}-\frac{\lambda_1 d^2}{c^2}-\lambda_3$. 因此 $(c\boldsymbol{A}+d\boldsymbol{I}_n)^{-1}=\frac{1}{k}(p\boldsymbol{A}+q\boldsymbol{I}_n)=\frac{1}{\frac{\lambda_2 d}{c}-\frac{\lambda_1 d^2}{c^2}-\lambda_3}\left(\frac{\lambda_1}{c}\boldsymbol{A}+\left(\frac{\lambda_2}{c}-\frac{\lambda_1 d}{c^2}\right)\boldsymbol{I}_n\right)$.

实际做题中只需要先设 $k(c\boldsymbol{A}+d\boldsymbol{I}_n)^{-1}=(p\boldsymbol{A}+q\boldsymbol{I}_n)$, 其中 k,p,q 为待定系数. 则有 $(c\boldsymbol{A}+d\boldsymbol{I}_n)(p\boldsymbol{A}+q\boldsymbol{I}_n)=k\boldsymbol{I}_n$. 然后通过比较二次项和一次项的系数得到方程组 $\begin{cases} pc=\lambda_1 \\ cq+dp=\lambda_2 \end{cases}$ (即要凑出合适的 p,q, 使得 $(c\boldsymbol{A}+d\boldsymbol{I}_n)(p\boldsymbol{A}+q\boldsymbol{I}_n)$ 与 $\lambda_1\boldsymbol{A}^2+\lambda_2\boldsymbol{A}+\lambda_3\boldsymbol{I}_n$ 的二次项和一次项的系数相等), 解出 p,q 的值. 最后将已知条件 $\lambda_1\boldsymbol{A}^2+\lambda_2\boldsymbol{A}+\lambda_3\boldsymbol{I}_n=\boldsymbol{O}$ 代入 $(c\boldsymbol{A}+d\boldsymbol{I}_n)(p\boldsymbol{A}+q\boldsymbol{I}_n)=k\boldsymbol{I}_n$ 即可得到 k 的值.

熟悉这种方式之后就能快速构造出我们需要的逆矩阵.

证明 (1) 和 (2) 的证明相同. 如下 (这里我们是利用了上述构造逆矩阵的方法直接构造出逆矩阵, 再根据逆矩阵的定义直接得到证明):

当 $c=0$ 时,$c\boldsymbol{A}+d\boldsymbol{I}_n=d\boldsymbol{I}_n$ 显然可逆.

当 $c\neq 0$ 时, 注意到 $(c\boldsymbol{A}+d\boldsymbol{I}_n)\left(\frac{\lambda_1}{c}\boldsymbol{A}+\left(\frac{\lambda_2}{c}-\frac{\lambda_1 d}{c^2}\right)\boldsymbol{I}_n\right)=\left(\frac{\lambda_2 d}{c}-\frac{\lambda_1 d^2}{c^2}-\lambda_3\right)\boldsymbol{I}_n$, 故 $c\boldsymbol{A}+d\boldsymbol{I}_n$ 可逆. □

例题 2.5 设 n 阶方阵 $\boldsymbol{A}$ 适合等式 $\boldsymbol{A}^2-3\boldsymbol{A}+2\boldsymbol{I}_n=\boldsymbol{O}$, 求证:$\boldsymbol{A}$ 和 $\boldsymbol{A}+\boldsymbol{I}_n$ 都是可逆阵, 而若 $\boldsymbol{A}\neq\boldsymbol{I}_n$, 则 $\boldsymbol{A}-2\boldsymbol{I}_n$ 必不是可逆阵.

笔记 这里构造逆矩阵利用了命题 2.17.

证明 由已知得 $\boldsymbol{A}(\boldsymbol{A}-3\boldsymbol{I}_n)=-2\boldsymbol{I}_n$, 因此 $\boldsymbol{A}$ 是可逆阵. 又 $\boldsymbol{A}^2-3\boldsymbol{A}-4\boldsymbol{I}_n=-6\boldsymbol{I}_n$, 于是 $(\boldsymbol{A}+\boldsymbol{I}_n)(\boldsymbol{A}-4\boldsymbol{I}_n)=-6\boldsymbol{I}_n$, 故 $\boldsymbol{A}+\boldsymbol{I}_n$ 也是可逆阵.

另一方面, 由已知等式可得 $(\boldsymbol{A}-\boldsymbol{I}_n)(\boldsymbol{A}-2\boldsymbol{I}_n)=\boldsymbol{O}$, 如果 $\boldsymbol{A}-2\boldsymbol{I}_n$ 可逆, 则 $\boldsymbol{A}-\boldsymbol{I}_n=\boldsymbol{O},\boldsymbol{A}=\boldsymbol{I}_n$ 和假设不合, 因此 $\boldsymbol{A}-2\boldsymbol{I}_n$ 不是可逆阵. □

命题 2.18

(1) 若已知 $\lambda_1\boldsymbol{AB}+\lambda_2\boldsymbol{A}+\lambda_3\boldsymbol{B}+\lambda_4\boldsymbol{I}_n=\boldsymbol{O}$, 其中 $\lambda_1,\lambda_2,\lambda_3,\lambda_4\in\mathbb{R},\lambda_1\neq 0$, 并且 $\lambda_1x^2+(\lambda_2+\lambda_3)x+\lambda_4=0$ 无实根 (即原等式左边不可因式分解成 $(a_1\boldsymbol{I}_n+a_2\boldsymbol{A})(b_1\boldsymbol{I}_n+b_2\boldsymbol{B})$), 则对任何 $c,d\in\mathbb{R}$, 都有 $a\boldsymbol{I}_n+b\boldsymbol{A},c\boldsymbol{I}_n+d\boldsymbol{B}$ 可逆.

(2) 若已知 $\lambda_1\boldsymbol{AB}+\lambda_2\boldsymbol{A}+\lambda_3\boldsymbol{B}+\lambda_4\boldsymbol{I}_n=(a_1\boldsymbol{I}_n+b_1\boldsymbol{A})(a_2\boldsymbol{I}_n+b_2\boldsymbol{B})=\boldsymbol{O}$, 其中 $a_1,a_2,b_1,b_2\in\mathbb{R}$, 且 $\lambda_1=b_1b_2,\lambda_2=a_2b_1,\lambda_3=a_1b_2,\lambda_4=a_1a_2,b_1\neq 0,b_2\neq 0$. 则对任何实数对 $(a,b),(c,d)\neq(a_1,b_1),(a_2,b_2)$, 都有 $a\boldsymbol{I}_n+b\boldsymbol{A},c\boldsymbol{I}_n+d\boldsymbol{B}$ 可逆.

证明 证明方法与命题 2.17类似, 构造逆矩阵的方法也与其类似. 这里不再赘述. □

例题 2.6

1. 求证: 不存在 n 阶奇异矩阵 A, 适合条件 $A^2+A+I_n=O$.
2. 设 A 是 n 阶矩阵, 且 $A^2=A$, 求证:I_n-2A 是可逆矩阵.
3. 若 A 是 n 阶矩阵, 且 $2A(A-I_n)=A^3$, 求证:I_n-A 可逆.

笔记 这类问题构造逆矩阵的方法 (以3.为例): 已知条件等价于 $A^3-2A^2+2A=O$, 设 $(I_n-A)^{-1}=aA^2+bA+cI_n$, 其中 a,b,c 为待定系数, 使得

$$(I_n-A)\left(aA^2+bA+cI_n\right)=A^3-2A^2+2A+kI_n=kI_n, k\text{为待定常数}.$$

比较等式两边系数可得

$$\begin{cases}-a=1\\a-b=-2\\b-c=2\\c=k\end{cases}\Rightarrow\begin{cases}a=-1\\b=1\\k=c=-1\end{cases}$$

于是 $(I_n-A)\left(-A^2+A-I_n\right)=-I_n$. 从而 $(I_n-A)^{-1}=A^2-A+I_n$.

证明

1. 由已知 $A^2+A+I_n=O$, 则 $(A-I_n)(A^2+A+I_n)=A^3-I_n=O$, 即 $A^3=I_n$, 于是 A 是可逆矩阵.
2. 因为 $(I_n-2A)^2=I_n-4A+4A^2=I_n$, 故 I_n-2A 是可逆矩阵.
3. 由已知 $A^3-2A^2+2A-I_n=-I_n$, 即 $(A-I_n)(A^2-A+I_n)=-I_n$, 于是 $(I_n-A)^{-1}=A^2-A+I_n$.

□

例题 2.7 设 n 阶方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 满足 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{AB}$, 求证:$\boldsymbol{I}_n-\boldsymbol{A}$ 是可逆阵且 $\boldsymbol{AB}=\boldsymbol{BA}$.

证明 因为

$$(\boldsymbol{I}_n-\boldsymbol{A})(\boldsymbol{I}_n-\boldsymbol{B})=\boldsymbol{I}_n-\boldsymbol{A}-\boldsymbol{B}+\boldsymbol{AB}=\boldsymbol{I}_n,$$

所以 $\boldsymbol{I}_n-\boldsymbol{A}$ 是可逆阵. 另一方面, 由上式可得 $(\boldsymbol{I}_n-\boldsymbol{A})^{-1}=(\boldsymbol{I}_n-\boldsymbol{B})$, 故

$$\boldsymbol{I}_n=(\boldsymbol{I}_n-\boldsymbol{B})(\boldsymbol{I}_n-\boldsymbol{A})=\boldsymbol{I}_n-\boldsymbol{B}-\boldsymbol{A}+\boldsymbol{BA},$$

从而 $\boldsymbol{BA}=\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{AB}$. □

命题 2.19 (矩阵转置的性质)

设矩阵 A,B, 则有

1. $(A')'=A$;
2. $(A+B)'=A'+B'$;
3. $(kA)'=kA'$;
4. $(AB)'=B'A'$.

证明 由矩阵的性质易证. □

命题 2.20 (矩阵的逆运算)

设矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 可逆, 则有

常规逆运算:

1. $(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.
2. $(\boldsymbol{AC}+\boldsymbol{BC})^{-1}=\boldsymbol{C}^{-1}(\boldsymbol{A}+\boldsymbol{B})^{-1}$.
3. $(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{C}=\left(\boldsymbol{C}^{-1}\boldsymbol{A}+\boldsymbol{C}^{-1}\boldsymbol{B}\right)^{-1}$.
4. $\boldsymbol{C}(\boldsymbol{A}+\boldsymbol{B})^{-1}=\left(\boldsymbol{AC}^{-1}+\boldsymbol{BC}^{-1}\right)^{-1}$.

凑因子:

1. $\boldsymbol{A}=\left(\boldsymbol{AB}^{-1}\right)\boldsymbol{B}=(\boldsymbol{AB})\boldsymbol{B}^{-1}=\boldsymbol{B}\left(\boldsymbol{B}^{-1}\boldsymbol{A}\right)=\boldsymbol{B}^{-1}(\boldsymbol{BA})$.
2. $\boldsymbol{A}+\boldsymbol{B}=\left(\boldsymbol{AC}^{-1}+\boldsymbol{BC}^{-1}\right)\boldsymbol{C}=(\boldsymbol{AC}+\boldsymbol{BC})\boldsymbol{C}^{-1}=\boldsymbol{C}\left(\boldsymbol{C}^{-1}\boldsymbol{A}+\boldsymbol{C}^{-1}\boldsymbol{B}\right)=\boldsymbol{C}^{-1}(\boldsymbol{CA}+\boldsymbol{CB})$.

笔记 无需额外记忆这些公式, 只需要知道**凑因子**的想法, 即**在矩阵可逆的条件下, 我们可以利用矩阵 $\boldsymbol{I}_n=\boldsymbol{AA}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}$ 的性质, 将原本矩阵没有的因子凑出来, 然后提取我们需要的矩阵因子到矩阵逆的外面或将其乘入矩阵逆的内部, 从而达到化简原矩阵的目的.**

证明 由矩阵的运算性质不难证明. □

注 凑因子想法的应用:例题 2.10,例题 2.11,例题 2.12.

例题 2.8 设 $A, B, A-B$ 都是 n 阶可逆阵, 证明:

$$B^{-1}-A^{-1}=(B+B(A-B)^{-1}B)^{-1}.$$

笔记 直接运用逆矩阵的定义验证即可.

证明

$$\begin{aligned}
&\left(B^{-1}-A^{-1}\right)\left(B+B(A-B)^{-1}B\right)\\
&=I_n+(A-B)^{-1}B-A^{-1}B-A^{-1}B(A-B)-1B\\
&=I_n+(A-B)^{-1}B-A^{-1}B(I_n+(A-B)-1B)\\
&=\left(I_n-A^{-1}B\right)\left(I_n+(A-B)^{-1}B\right)\\
&=A^{-1}(A-B)\left[(A-B)^{-1}(A-+B)\right]\\
&=A^{-1}(A-B)(A-B)^{-1}A=I_n.
\end{aligned}$$

□

例题 2.9 Sherman-Morrison 公式 设 A 是 n 阶可逆阵,α, β 是 n 维列向量, 且 $1+\beta' A^{-1}\alpha \neq 0$. 求证:

$$(A+\alpha\beta')^{-1}=A^{-1}-\frac{1}{1+\beta' A^{-1}\alpha}A^{-1}\alpha\beta' A^{-1}.$$

笔记 直接运用逆矩阵的定义验证即可, 注意 $\beta' A^{-1}\alpha$ 是一个数可以提出来.

证明

$$\begin{aligned}
&(A+\alpha\beta')\left(A^{-1}-\frac{1}{1+\beta' A^{-1}\alpha}A^{-1}\alpha\beta' A^{-1}\right)\\
&=I_n-\frac{1}{1+\beta' A^{-1}\alpha}\alpha\beta' A^{-1}+\alpha\beta' A^{-1}-\frac{1}{1+\beta' A^{-1}\alpha}\alpha\left(\beta' A^{-1}\alpha\right)\beta' A^{-1}\\
&=I_n+\alpha\beta' A^{-1}-\frac{1}{1+\beta' A^{-1}\alpha}\alpha\beta' A^{-1}-\frac{\beta' A^{-1}\alpha}{1+\beta' A^{-1}\alpha}\alpha\beta' A^{-1}\\
&=I_n+\alpha\beta' A^{-1}-\frac{1+\beta' A^{-1}\alpha}{1+\beta' A^{-1}\alpha}\alpha\beta' A^{-1}=I_n.
\end{aligned}$$

□

命题 2.21 (一些矩阵等式)

1. 设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵,$\boldsymbol{B}$ 为 $n\times m$ 矩阵. 则有 $\boldsymbol{A}\left(\boldsymbol{I_n}+\boldsymbol{BA}\right)=\left(\boldsymbol{I_m}+\boldsymbol{AB}\right)\boldsymbol{A}$.
2. 设 $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶可逆矩阵, 则有 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{A}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})\boldsymbol{B}$.
3. 若 n 阶矩阵 A, B 满足 $A^2=B^2$, 则 $A(A+B)=A^2+AB=B^2+AB=(A+B)B$.

笔记 这是一些常见的矩阵等式. 可以通过反复凑因子得到.

证明 由矩阵的运算性质不难证明. □

例题 2.10 设 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{AB}-\boldsymbol{I}_n$ 都是 n 阶可逆阵, 证明:$\boldsymbol{A}-\boldsymbol{B}^{-1}$ 与 $(\boldsymbol{A}-\boldsymbol{B}^{-1})^{-1}-\boldsymbol{A}^{-1}$ 均可逆, 并求它们的逆矩阵.

笔记 核心想法是利用命题 2.20和命题 2.21.

证明 注意到 $\boldsymbol{A}-\boldsymbol{B}^{-1}=(\boldsymbol{AB}-\boldsymbol{I}_n)\boldsymbol{B}^{-1}$, 故 $\boldsymbol{A}-\boldsymbol{B}^{-1}$ 是可逆矩阵, 并且 $(\boldsymbol{A}-\boldsymbol{B}^{-1})^{-1}=\boldsymbol{B}(\boldsymbol{AB}-\boldsymbol{I}_n)^{-1}$. 注意到如下变形:

$$\begin{aligned}
&(\boldsymbol{A}-\boldsymbol{B}^{-1})^{-1}-\boldsymbol{A}^{-1}\\
=&\boldsymbol{B}(\boldsymbol{AB}-\boldsymbol{I}_n)^{-1}-\boldsymbol{A}^{-1}=\boldsymbol{A}^{-1}(\boldsymbol{AB}(\boldsymbol{AB}-\boldsymbol{I}_n)^{-1}-\boldsymbol{I}_n)\\
=&\boldsymbol{A}^{-1}(\boldsymbol{AB}-(\boldsymbol{AB}-\boldsymbol{I}_n))(\boldsymbol{AB}-\boldsymbol{I}_n)^{-1}=\boldsymbol{A}^{-1}(\boldsymbol{AB}-\boldsymbol{I}_n)^{-1}.
\end{aligned}$$

故 $(\boldsymbol{A}-\boldsymbol{B}^{-1})^{-1}-\boldsymbol{A}^{-1}$ 可逆, 并且 $((\boldsymbol{A}-\boldsymbol{B}^{-1})^{-1}-\boldsymbol{A}^{-1})^{-1}=(\boldsymbol{AB}-\boldsymbol{I}_n)\boldsymbol{A}$. □

命题 2.22

设 A 为 $m\times n$ 矩阵,B 为 $n\times m$ 矩阵, 使得 I_m+AB 可逆, 则 I_n+BA 也可逆, 并且 $(I_n+BA)^{-1}=I_n-B(I_m+AB)^{-1}A$.

笔记 命题 2.22的应用: 一般对于求只含有两项的矩阵和式的逆矩阵, 我们可以利用矩阵的逆运算 (凑因子)的方法将原矩阵和式转化为 $C(I_n+AB)$ 或 $(I_n+AB)C$ 的形式, 再利用这个命题求得原矩阵的逆.

注 证法一只能得到 I_n+BA 可逆, 并不能得到具体的逆矩阵. 而证法二可以求出 $(I_n+BA)^{-1}=I_n-B(I_m+AB)^{-1}A$.

证明 证法一 (打洞原理):根据分块矩阵的初等变换可得

$$\begin{vmatrix} I_m & -A \\ B & I_n \end{vmatrix} = \begin{vmatrix} I_m & O \\ -B & I_n \end{vmatrix}\begin{vmatrix} I_m & -A \\ B & I_n \end{vmatrix} = \left|\begin{pmatrix} I_m & O \\ -B & I_n \end{pmatrix}\begin{pmatrix} I_m & -A \\ B & I_n \end{pmatrix}\right| = \begin{vmatrix} I_m & -A \\ O & I_n+BA \end{vmatrix} = |I_n+BA|.$$

$$\begin{vmatrix} I_m & -A \\ B & I_n \end{vmatrix} = \begin{vmatrix} I_m & A \\ O & I_n \end{vmatrix}\begin{vmatrix} I_m & -A \\ B & I_n \end{vmatrix} = \left|\begin{pmatrix} I_m & A \\ O & I_n \end{pmatrix}\begin{pmatrix} I_m & -A \\ B & I_n \end{pmatrix}\right| = \begin{vmatrix} I_m+AB & O \\ B & I_n \end{vmatrix} = |I_m+AB|.$$

故 $\begin{vmatrix} I_m & -A \\ B & I_n \end{vmatrix} = |I_m+AB| = |I_n+BA|$. 又因为 I_m+AB 可逆, 所以 $|I_n+BA|=|I_m+AB|\neq 0$. 因此 I_n+BA 也可逆.

证法二 (矩阵的逆运算): 注意到 $A(I_n+BA)=(I_m+AB)A$, 故 $(I_m+AB)^{-1}A(I_n+BA)=A$, 于是 $B(I_m+AB)^{-1}A(I_n+BA)=BA$, 从而

$$\begin{aligned} I_n &= I_n+BA-BA=(I_n+BA)-B(I_m+AB)^{-1}A(I_n+BA) \\ &= (I_n-B(I_m+AB)^{-1}A)(I_n+BA). \end{aligned}$$

于是 $(I_n+BA)^{-1}=I_n-B(I_m+AB)^{-1}A$. □

例题 2.11 设 A,B 均为 n 阶可逆阵, 使得 $A^{-1}+B^{-1}$ 可逆, 证明:$A+B$ 也可逆, 并且

$$(A+B)^{-1}=A^{-1}-A^{-1}(A^{-1}+B^{-1})^{-1}A^{-1}.$$

证明 注意到 $A+B=A(A^{-1}+B^{-1})B$, 故 $A+B$ 可逆. 由命题 2.22可得

$$(I_n+A^{-1}B)^{-1}=I_n-A^{-1}(I_n+BA^{-1})^{-1}B=I_n-A^{-1}(A^{-1}+B^{-1})^{-1},$$

于是

$$\begin{aligned} (A+B)^{-1} &= (A(I_n+A^{-1}B))^{-1}=(I_n+A^{-1}B)^{-1}A^{-1} \\ &= A^{-1}-A^{-1}(A^{-1}+B^{-1})^{-1}A^{-1}.\square \end{aligned}$$

□

例题 2.12 Sherman-Morrison-Woodbury 公式

设 A 为 n 阶可逆阵,C 为 m 阶可逆阵,B 为 $n\times m$ 矩阵,D 为 $m\times n$ 矩阵, 使得 $C^{-1}+DA^{-1}B$ 可逆. 求证:$A+BCD$ 也可逆, 并且

$$(A+BCD)^{-1}=A^{-1}-A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}.$$

注 若已知矩阵逆的表达式, 也可以采取利用矩阵逆的定义直接验证的方法进行证明.

证明 注意到 $A+BCD=A(I_n+A^{-1}BCD)$, 将 $A^{-1}B$ 和 CD 分别看成整体, 此时 $I_m+(CD)(A^{-1}B)=C(C^{-1}+DA^{-1}B)$ 可逆, 故由命题 2.22的结论可知 $I_n+(A^{-1}B)(CD)$ 也可逆, 并且

$$\begin{aligned} (I_n+A^{-1}BCD)^{-1} &= I_n-A^{-1}B(I_m+CDA^{-1}B)^{-1}CD \\ &= I_n-A^{-1}B(C^{-1}+DA^{-1}B)^{-1}D. \end{aligned}$$

于是 $A+BCD=A(I_n+A^{-1}BCD)$ 也可逆, 并且

$$(A+BCD)^{-1}=A^{-1}-A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}.$$

□

2.1.1 练习

练习 2.1 计算下列矩阵的 k 次幂, 其中 k 为正整数:

$$(1)\ A=\begin{pmatrix} a & 1 & 0\\ 0 & a & 1\\ 0 & 0 & a\end{pmatrix};\quad (2)\ A=\begin{pmatrix} 1 & 2 & 4\\ 2 & 4 & 8\\ 3 & 6 & 12\end{pmatrix}.$$

笔记 第 (2) 问核心想法是利用命题 2.34.

解 (1) 设 $J=\begin{pmatrix} 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}$, 则 $A=aI_3+J$. 注意到 aI_3 和 J 乘法可交换,J 是幂零阵并且 $J^3=O$, 因此我们可用二项式定理来求 A 的 k 次幂:

$$\begin{aligned}A^k&=(aI_3+J)^k=(aI_3)^k+C_k^1(aI_3)^{k-1}J+C_k^2(aI_3)^{k-2}J^2\\&=a^kI_3+C_k^1a^{k-1}J+C_k^2a^{k-2}J^2=\begin{pmatrix} a^k & C_k^1a^{k-1} & C_k^2a^{k-2}\\ 0 & a^k & C_k^1a^{k-1}\\ 0 & 0 & a^k\end{pmatrix}\end{aligned}$$

(2) 注意到 A 的列向量成比例, 故可设 $\alpha=(1,2,3),\beta=(1,2,4)$, 则 $A=\alpha\beta'$. 由矩阵乘法的结合律并注意到 $\beta\alpha'=17$, 可得

$$\begin{aligned}A^k&=(\alpha\beta')(\alpha\beta')\cdots(\alpha\beta')=\alpha(\beta'\alpha)(\beta'\alpha)\cdots(\beta'\alpha)\beta'\\&=(\beta'\alpha)^{k-1}\alpha\beta'=17^{k-1}A=\begin{pmatrix} 17^{k-1} & 2\cdot 17^{k-1} & 4\cdot 17^{k-1}\\ 2\cdot 17^{k-1} & 4\cdot 17^{k-1} & 8\cdot 17^{k-1}\\ 3\cdot 17^{k-1} & 6\cdot 17^{k-1} & 12\cdot 17^{k-1}\end{pmatrix}\end{aligned}$$

□

练习 2.2 设 k 是正整数, 计算 $\begin{pmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}^k$.

解 已知 $k=1$ 时, 有 $\begin{pmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$. 假设 $k=n$ 时, 有 $\begin{pmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}^n=\begin{pmatrix} \cos n\theta & \sin n\theta\\ -\sin n\theta & \cos n\theta\end{pmatrix}$. 则当 $k=n+1$ 时, 有

$$\begin{aligned}&\begin{pmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}^{n+1}=\begin{pmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}^n\begin{pmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}=\begin{pmatrix} \cos n\theta & \sin n\theta\\ -\sin n\theta & \cos n\theta\end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\\&=\begin{pmatrix} \cos n\theta\cos\theta-\sin n\theta\sin\theta & \cos n\theta\sin\theta+\sin n\theta\cos\theta\\ -(\cos n\theta\sin\theta+\sin n\theta\cos\theta) & \cos n\theta\cos\theta-\sin n\theta\sin\theta\end{pmatrix}=\begin{pmatrix} \cos(n+1)\theta & \sin(n+1)\theta\\ -\sin(n+1)\theta & \cos(n+1)\theta\end{pmatrix}.\end{aligned}$$

从而由数学归纳法可知, 对任意正整数 k, 有 $\begin{pmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}^k=\begin{pmatrix} \cos k\theta & \sin k\theta\\ -\sin k\theta & \cos k\theta\end{pmatrix}$. □

练习 2.3 求矩阵 A 的逆阵:

$$A=\begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n\\ n & 1 & 2 & \cdots & n-2 & n-1\\ n-1 & n & 1 & \cdots & n-3 & n-2\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 2 & 3 & 4 & \cdots & n & 1\end{pmatrix}.$$

解 对 $\begin{pmatrix} A & I_n \end{pmatrix}$ 用初等变换法, 将所有行加到第一行上, 再将第一行乘以 s^{-1}, 其中 $s=\dfrac{1}{2}n(n+1)$, 得到

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

$$\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 & 1 & \frac{1}{s} & \frac{1}{s} & \frac{1}{s} & \cdots & \frac{1}{s} & \frac{1}{s}\\
n & 1 & 2 & \cdots & n-2 & n-1 & 0 & 1 & 0 & \cdots & 0 & 0\\
n-1 & n & 1 & \cdots & n-3 & n-2 & 0 & 0 & 1 & \cdots & 0 & 0\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
2 & 3 & 4 & \cdots & n & 1 & 0 & 0 & 0 & \cdots & 0 & 1
\end{pmatrix}.$$

从第二行起依次减去下一行, 得到

$$\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 & 1 & \frac{1}{s} & \frac{1}{s} & \frac{1}{s} & \cdots & \frac{1}{s} & \frac{1}{s}\\
1 & 1-n & 1 & \cdots & 1 & 1 & 0 & 1 & -1 & \cdots & 0 & 0\\
1 & 1 & 1-n & \cdots & 1 & 1 & 0 & 0 & 1 & \cdots & 0 & 0\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
2 & 3 & 4 & \cdots & n & 1 & 0 & 0 & 0 & \cdots & 0 & 1
\end{pmatrix}.$$

消去第一列除第一行外的所有元素后, 得到

$$\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 & 1 & \frac{1}{s} & \frac{1}{s} & \frac{1}{s} & \cdots & \frac{1}{s} & \frac{1}{s}\\
0 & -n & 0 & \cdots & 0 & 0 & -\frac{1}{s} & \frac{s-1}{s} & -\frac{s+1}{s} & \cdots & -\frac{1}{s} & -\frac{1}{s}\\
0 & 0 & -n & \cdots & 0 & 0 & -\frac{1}{s} & -\frac{1}{s} & \frac{s-1}{s} & \cdots & -\frac{1}{s} & -\frac{1}{s}\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 1 & 2 & \cdots & n-2 & -1 & -\frac{2}{s} & -\frac{2}{s} & -\frac{2}{s} & \cdots & -\frac{2}{s} & \frac{s-2}{s}
\end{pmatrix}.$$

从第二行到第 $n-1$ 行分别乘以 $-\dfrac{1}{n}$, 得到

$$\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 & 1 & \frac{1}{s} & \frac{1}{s} & \frac{1}{s} & \cdots & \frac{1}{s} & \frac{1}{s}\\
0 & 1 & 0 & \cdots & 0 & 0 & \frac{1}{ns} & \frac{1-s}{ns} & \frac{s+1}{ns} & \cdots & \frac{1}{ns} & \frac{1}{ns}\\
0 & 0 & 1 & \cdots & 0 & 0 & \frac{1}{ns} & \frac{1}{ns} & \frac{1-s}{ns} & \cdots & \frac{1}{ns} & \frac{1}{ns}\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 1 & 2 & \cdots & n-2 & -1 & -\frac{2}{s} & -\frac{2}{s} & -\frac{2}{s} & \cdots & -\frac{2}{s} & \frac{s-2}{s}
\end{pmatrix}.$$

将第一行依次减去第二行, 第三行,$\cdots$, 第 $n-1$ 行, 得到

$$\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 & 1 & \frac{2}{ns} & \frac{s+2}{ns} & \frac{2}{ns} & \cdots & \frac{2}{ns} & \frac{2-s}{ns}\\
0 & 1 & 0 & \cdots & 0 & 0 & \frac{1}{ns} & \frac{1-s}{ns} & \frac{s+1}{ns} & \cdots & \frac{1}{ns} & \frac{1}{ns}\\
0 & 0 & 1 & \cdots & 0 & 0 & \frac{1}{ns} & \frac{1}{ns} & \frac{1-s}{ns} & \cdots & \frac{1}{ns} & \frac{1}{ns}\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 1 & 2 & \cdots & n-2 & -1 & -\frac{2}{s} & -\frac{2}{s} & -\frac{2}{s} & \cdots & -\frac{2}{s} & \frac{s-2}{s}
\end{pmatrix}.$$

将最后一行加到第一行, 再将最后一行乘以 -1, 得到

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & \frac{1-s}{ns} & \frac{1+s}{ns} & \frac{1}{ns} & \cdots & \frac{1}{ns} & \frac{1}{ns} \\ 0 & 1 & 0 & \cdots & 0 & 0 & \frac{1}{ns} & \frac{1-s}{ns} & \frac{s+1}{ns} & \cdots & \frac{1}{ns} & \frac{1}{ns} \\ 0 & 0 & 1 & \cdots & 0 & 0 & \frac{1}{ns} & \frac{1}{ns} & \frac{1-s}{ns} & \cdots & \frac{1}{ns} & \frac{1}{ns} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & \frac{s+1}{ns} & \frac{1}{ns} & \frac{1}{ns} & \cdots & \frac{1}{ns} & \frac{1-s}{ns} \end{pmatrix}.$$

因此

$$A^{-1} = \frac{1}{ns}\begin{pmatrix} 1-s & 1+s & 1 & \cdots & 1 & 1 \\ 1 & 1-s & 1+s & \cdots & 1 & 1 \\ 1 & 1 & 1-s & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1+s & 1 & 1 & \cdots & 1 & 1-s \end{pmatrix}.$$

□

练习 2.4 求下列 n 阶矩阵的逆阵, 其中 $a_i \neq 0(1 \leq i \leq n)$:

$$A = \begin{pmatrix} 1+a_1 & 1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & 1 & \cdots & 1 \\ 1 & 1 & 1+a_3 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1+a_n \end{pmatrix}.$$

解 对 $\begin{pmatrix} A & I_n \end{pmatrix}$ 用初等变换法, 将第 i 行乘以 $a_i^{-1}(1 \leq i \leq n)$, 有

$$\begin{pmatrix} 1+a_1 & 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 1+a_2 & 1 & \cdots & 1 & 0 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1+a_3 & \cdots & 1 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1+a_n & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \to$$

$$\begin{pmatrix} 1+\frac{1}{a_1} & \frac{1}{a_1} & \frac{1}{a_1} & \cdots & \frac{1}{a_1} & \frac{1}{a_1} & 0 & 0 & \cdots & 0 \\ \frac{1}{a_2} & 1+\frac{1}{a_2} & \frac{1}{a_2} & \cdots & \frac{1}{a_2} & 0 & \frac{1}{a_2} & 0 & \cdots & 0 \\ \frac{1}{a_3} & \frac{1}{a_3} & 1+\frac{1}{a_3} & \cdots & \frac{1}{a_3} & 0 & 0 & \frac{1}{a_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{a_n} & \frac{1}{a_n} & \frac{1}{a_n} & \cdots & 1+\frac{1}{a_n} & 0 & 0 & 0 & \cdots & \frac{1}{a_n} \end{pmatrix}.$$

将下面的行都加到第一行上, 并令 $s = 1 + \frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}$, 则上面的矩阵变为

$$\begin{pmatrix} s & s & s & \cdots & s & \frac{1}{a_1} & \frac{1}{a_2} & \frac{1}{a_3} & \cdots & \frac{1}{a_n} \\ \frac{1}{a_2} & 1+\frac{1}{a_2} & \frac{1}{a_2} & \cdots & \frac{1}{a_2} & 0 & \frac{1}{a_2} & 0 & \cdots & 0 \\ \frac{1}{a_3} & \frac{1}{a_3} & 1+\frac{1}{a_3} & \cdots & \frac{1}{a_3} & 0 & 0 & \frac{1}{a_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{a_n} & \frac{1}{a_n} & \frac{1}{a_n} & \cdots & 1+\frac{1}{a_n} & 0 & 0 & 0 & \cdots & \frac{1}{a_n} \end{pmatrix} \to$$

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & \frac{1}{sa_1} & \frac{1}{sa_2} & \frac{1}{sa_3} & \cdots & \frac{1}{sa_n} \\ \frac{1}{a_2} & 1+\frac{1}{a_2} & \frac{1}{a_2} & \cdots & \frac{1}{a_2} & 0 & \frac{1}{a_2} & 0 & \cdots & 0 \\ \frac{1}{a_3} & \frac{1}{a_3} & 1+\frac{1}{a_3} & \cdots & \frac{1}{a_3} & 0 & 0 & \frac{1}{a_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{a_n} & \frac{1}{a_n} & \frac{1}{a_n} & \cdots & 1+\frac{1}{a_n} & 0 & 0 & 0 & \cdots & \frac{1}{a_n} \end{pmatrix} \to$$

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & \frac{1}{sa_1} & \frac{1}{sa_2} & \frac{1}{sa_3} & \cdots & \frac{1}{sa_n} \\ 0 & 1 & 0 & \cdots & 0 & -\frac{1}{sa_1a_2} & \frac{sa_2-1}{sa_2^2} & -\frac{1}{sa_3a_2} & \cdots & -\frac{1}{sa_na_2} \\ 0 & 0 & 1 & \cdots & 0 & -\frac{1}{sa_1a_3} & -\frac{1}{sa_2a_3} & \frac{sa_3-1}{sa_3^2} & \cdots & -\frac{1}{sa_na_3} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\frac{1}{sa_1a_n} & -\frac{1}{sa_2a_n} & -\frac{1}{sa_3a_n} & \cdots & \frac{sa_n-1}{sa_n^2} \end{pmatrix}.$$

再消去第一行的后 $n-1$ 个 1 就得到

$$A^{-1} = -\frac{1}{s}\begin{pmatrix} \frac{1}{a_1} & \frac{1}{a_2} & \frac{1}{a_3} & \cdots & \frac{1}{a_n} \\ -\frac{1}{a_1a_2} & \frac{sa_2-1}{a_2^2} & -\frac{1}{a_3a_2} & \cdots & -\frac{1}{a_na_2} \\ -\frac{1}{a_1a_3} & -\frac{1}{a_2a_3} & \frac{sa_3-1}{a_3^2} & \cdots & -\frac{1}{a_na_3} \\ \vdots & \vdots & \vdots & & \vdots \\ -\frac{1}{a_1a_n} & -\frac{1}{a_2a_n} & -\frac{1}{a_3a_n} & \cdots & \frac{sa_n-1}{a_n^2} \end{pmatrix}.$$

□

练习 2.5 求下列 n 阶矩阵的逆矩阵:

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

笔记 解法一和解法二的核心想法是: 先假设 (猜测) 矩阵 A 的逆矩阵与其具有相似的结构, 再结合逆矩阵的定义, 使用待定系数法求出矩阵 A 的逆矩阵.

解 解法一: 设 $\alpha = (1, 1, \cdots, 1)'$, 则 $A = -I_n + \alpha\alpha'$. 设 $B = cI_n + d\alpha\alpha'$, 其中 c, d 为待定系数. 则 $AB = -cI_n + (c + (n-1)d)\alpha\alpha'$. 令 $c = -1, c + (n-1)d = 0$, 则 $d = \frac{1}{n-1}$. 于是 $AB = I_n$, 从而 $A^{-1} = B = -I_n + \frac{1}{n-1}\alpha\alpha'$.

解法二(Sherman-Morrison 公式):设 $\alpha = (1, 1, \cdots, 1)'$, 则 $A = -I_n + \alpha\alpha'$. 由Sherman-Morrison 公式可得

$$\boldsymbol{A}^{-1} = \left(-\boldsymbol{I}_n + \boldsymbol{\alpha}\boldsymbol{\alpha}'\right)^{-1} = (-\boldsymbol{I}_n)^{-1} - \frac{1}{1+\boldsymbol{\alpha}'(-\boldsymbol{I}_n)^{-1}\boldsymbol{\alpha}}(-\boldsymbol{I}_n)^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}'(-\boldsymbol{I}_n)^{-1} = -\boldsymbol{I}_n + \frac{1}{n-1}\boldsymbol{\alpha}\boldsymbol{\alpha}'.$$

解法三 (循环矩阵): 设 J 为基础循环矩阵, 则 $A = J + J^2 + \cdots + J^{n-1}$. 设 $B = cI_n + J + J^2 + \cdots + J^{n-1}$(因为循环矩阵的逆仍是循环矩阵), 其中 c 为待定系数. 则

$$AB = (n-1)I_n + (c + n - 2)(J + J^2 + \cdots + J^{n-1})$$

只要令 $c = 2 - n$, 则 $AB = (n-1)I_n$. 于是 $A^{-1} = \frac{1}{n-1}B = \frac{2-n}{n-1}I_n + J + J^2 + \cdots + J^{n-1}$.

解法四 (初等变换):本题是练习 2.4的特例, 都利用相同的初等变换方法求逆矩阵. □

练习 2.6 设 $\boldsymbol{A}$ 是非零实矩阵且 $\boldsymbol{A}^* = \boldsymbol{A}'$. 求证:$\boldsymbol{A}$ 是可逆阵.

证明 设 $\boldsymbol{A}=(a_{ij})$,a_{ij} 的代数余子式记为 A_{ij}, 由已知,$a_{ij}=A_{ij}$. 由于 $\boldsymbol{A}$ 是非零实矩阵, 故必有某个 $a_{rs}\neq 0$, 将 $|\boldsymbol{A}|$ 按第 r 行展开, 可得

$$|\boldsymbol{A}|=a_{r1}A_{r1}+\cdots+a_{rs}A_{rs}+\cdots+a_{rn}A_{rn}=a_{r1}^2+\cdots+a_{rs}^2+\cdots+a_{rn}^2>0.$$

特别地,$|\boldsymbol{A}|\neq 0$, 即 $\boldsymbol{A}$ 是可逆阵. □

练习 2.7 设 $\boldsymbol{A}$ 是奇数阶矩阵, 满足 $\boldsymbol{A}\boldsymbol{A}'=\boldsymbol{I}_n$ 且 $|\boldsymbol{A}|>0$, 证明:$\boldsymbol{I}_n-\boldsymbol{A}$ 是奇异阵.

证明 由 $1=|\boldsymbol{I}_n|=|\boldsymbol{A}\boldsymbol{A}'|=|\boldsymbol{A}||\boldsymbol{A}'|=|\boldsymbol{A}|^2$ 以及 $|\boldsymbol{A}|>0$ 可得 $|\boldsymbol{A}|=1$. 因为

$$|\boldsymbol{I}_n-\boldsymbol{A}|=|\boldsymbol{A}\boldsymbol{A}'-\boldsymbol{A}|=|\boldsymbol{A}||\boldsymbol{A}'-\boldsymbol{I}_n|=|(\boldsymbol{A}-\boldsymbol{I}_n)'|=|\boldsymbol{A}-\boldsymbol{I}_n|=(-1)^n|\boldsymbol{I}_n-\boldsymbol{A}|.$$

又 n 是奇数, 故 $|\boldsymbol{I}_n-\boldsymbol{A}|=-|\boldsymbol{I}_n-\boldsymbol{A}|$, 从而 $|\boldsymbol{I}_n-\boldsymbol{A}|=0$, 即 $\boldsymbol{I}_n-\boldsymbol{A}$ 是奇异阵. □

练习 2.8 设 $\boldsymbol{A},\boldsymbol{B}$ 为 n 阶可逆阵, 满足 $\boldsymbol{A}^2=\boldsymbol{B}^2$ 且 $|\boldsymbol{A}|+|\boldsymbol{B}|=0$, 求证:$\boldsymbol{A}+\boldsymbol{B}$ 是奇异阵.

证明 由已知 $\boldsymbol{A},\boldsymbol{B}$ 都是可逆阵且 $|\boldsymbol{B}|=-|\boldsymbol{A}|$, 因此

$$|\boldsymbol{A}||\boldsymbol{A}+\boldsymbol{B}|=|\boldsymbol{A}^2+\boldsymbol{A}\boldsymbol{B}|=|\boldsymbol{B}^2+\boldsymbol{A}\boldsymbol{B}|=|\boldsymbol{B}+\boldsymbol{A}||\boldsymbol{B}|=-|\boldsymbol{A}||\boldsymbol{A}+\boldsymbol{B}|.$$

于是 $|\boldsymbol{A}||\boldsymbol{A}+\boldsymbol{B}|=0$. 因为 $|\boldsymbol{A}|\neq 0$, 故 $|\boldsymbol{A}+\boldsymbol{B}|=0$, 即 $\boldsymbol{A}+\boldsymbol{B}$ 是奇异阵. □

2.2 矩阵的初等变换

2.2.1 相抵标准型

定义 2.2 (矩阵相抵的定义)

设矩阵 A,B, 若 A 经有限次初等变换后变成 B, 则称 A 与 B **相抵**, 记作 $A\sim B$.

笔记 容易验证相抵是 $M_{s\times n}(K)$ 上的一个等价关系. 在相抵关系下, 矩阵 A 的等价类称为 A 的相抵类.

命题 2.23 (矩阵相抵的等价命题)

数域 K 上 $s\times n$ 矩阵 A 和 B 相抵等价于:

1. A可以经过初等行变换和初等列变换变成B.
2. 存在K上s级初等矩阵$P_1,P_2,\cdots,P_t$与n级初等矩阵$Q_1,Q_2,\cdots,Q_m$,使得 $P_t\cdots P_2P_1AQ_1Q_2\cdots Q_m=B$.
3. 存在K上s级可逆矩阵P与n级可逆矩阵Q,使得: $PAQ=B$.

定理 2.1 (相抵标准型)

设数域 K 上 $s\times n$ 矩阵 A 的秩为 r. 如果 $r>0$, 那么 A 相抵于下述形式的矩阵:

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \tag{2.1}$$

称矩阵(2.1)为 A 的相抵标准形; 如果 $r=0$, 那么 A 相抵于零矩阵, 此时称 A 的相抵标准形是零矩阵.

推论 2.1

1. 数域 K 上 $s\times n$ 矩阵 A 和 B 相抵当且仅当它们的秩相等.
2. 设数域 K 上 $s\times n$ 矩阵 A 的秩为 $r(r>0)$, 则存在 K 上的 s 级、n 级可逆矩阵 P、Q, 使得 $A=P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q$.

命题 2.24 (奇异阵的充要条件)

数域 K 上的 n 阶矩阵 A 是奇异阵的充要条件有:

1. 存在数域 K 上不为零的同阶方阵 B, 使得 $AB = O$.
2. 存在数域 K 上的 n 维非零列向量 x, 使得 $Ax = 0$.

证明

1. 充分性 ($\Leftarrow$): 显然若 A 可逆, 则从 $AB = O$ 可得到 $B = O$, 因此充分性成立.

   必要性 ($\Rightarrow$): 反之, 若 A 是奇异阵, 则存在数域 K 上的可逆阵 P, Q, 使得 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 其中 $r < n$. 令 $C = \begin{pmatrix} O & O \\ O & I_{n-r} \end{pmatrix}$, 则 $PAQC = O$. 又因为 P 可逆, 故 $AQC = O$. 只要令 $B = QC \in K$ 就得到了结论.
2. 充分性 ($\Leftarrow$): 显然若 A 可逆, 则从 $Ax = 0$ 可得到 $x = 0$, 因此充分性成立.

   必要性 ($\Rightarrow$): 反之, 若 A 是奇异阵, 则存在数域 K 上的可逆阵 P, Q, 使 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 其中 $r < n$. 令 $y = (0, \cdots, 0, 1)'$ 为 n 维列向量, 则 $PAQy = 0$. 又因为 P 可逆, 故 $AQy = 0$. 只要令 $x = Qy \in K$ 就得到了结论.

□

2.2.2 练习

练习 2.9 设 A 为 n 阶实反对称阵, 证明:$I_n - A$ 是非异阵.

证明 (反证法) 假设是 $I_n - A$ 是奇异阵, 则由命题 2.24 的 2, 可知存在 n 维非零实列向量 x, 使得 $(I_n - A)x = 0$, 即 $Ax = x$. 设 $x = (a_1, a_2, \cdots, a_n)'$, 其中 a_i 都是实数, 则由 A 的反对称性以及命题 2.11, 可知

$$0 = x'Ax = x'x = a_1^2 + a_2^2 + \cdots + a_n^2.$$

从而 $a_1 = a_2 = \cdots = a_n = 0$, 即 $x = 0$, 这与已知矛盾. □

练习 2.10 设 A 为 n 阶可逆阵, 求证: 只用第三类初等变换就可以将 A 化为如下形状:

$$\mathrm{diag}\{1, \cdots, 1, |A|\}.$$

证明 假设 A 的第 $(1,1)$ 元素等于零, 因为 A 可逆, 故第一行必有元素不为零. 用第三初等变换将非零元素所在的列加到第一列, 则到的矩阵中第 $(1,1)$ 元素不为零. 因此不设 A 的第 $(1,1)$ 元素非零, 于是可用三类初等变换将 A 的第一行及第一列其余素都消为零. 这就是说,A 经过第三类初变换可化为如下形状:

$$\begin{pmatrix} a & O \\ O & A_1 \end{pmatrix}.$$

再对 A_1 同样处理, 不断做下去, 可将 A 化为对角阵, 并且对角元素均非零. 因此我们只要对对角阵证明结论即可. 为简化讨论, 我们先考虑二阶对角阵:

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

将其第一行乘以 $(1-a)a^{-1}$ 加到第行上, 再将第二行加到第一行上得到:

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

将其第一列乘以 $-b$ 加到第二列上, 再将第行乘以 $a-1$ 加到第二行上得到:

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

从而原结论对二阶对角阵成立. 对于 n 阶对角阵 $B = diag\{a_1, a_2, \cdots, a_n\}$ 而言, 按照上述方法对 $B\begin{pmatrix}1 & 2\\1 & 2\end{pmatrix}$ 所对应的子矩阵进行第三类初等变换得到

$$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & & & \\ & a_1a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}.$$

按照上述方法对再对 $B\begin{pmatrix}2 & 3\\2 & 3\end{pmatrix}$ 所对应的子矩阵进行第三类初等变换得到

$$\begin{pmatrix} 1 & & & & \\ & a_1a_2 & & & \\ & & a_3 & & \\ & & & \ddots & \\ & & & & a_n \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & a_1a_2a_3 & & \\ & & & \ddots & \\ & & & & a_n \end{pmatrix}.$$

同理依次对 $B\begin{pmatrix}k & k+1\\k & k+1\end{pmatrix}, k = 1, 2\cdots, n-1$ 所对应的子矩阵按照上述方法进行第三类初等变换, 最后得到

$$B = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & a_1a_2\cdots a_n \end{pmatrix}.$$

于是原结论对对角阵也成立. 而我们所用的初等变换始终是第三类初等变换. 这就得到了结论. □

练习 2.11 求证: 任一 n 阶矩阵均可表示为形如 $I_n + a_{ij}E_{ij}$ 这样的矩阵之积, 其中 E_{ij} 是 n 阶基础矩阵.

证明 由命题 2.1可知任意一个 n 阶矩阵都可表示为有限个初等阵和具有下列形状的对角阵 D 之积:

$$D = \mathrm{diag}\{1, \cdots, 1, 0, \cdots, 0\},$$

故只要对初等阵和 D 证明结论即可. 对 D, 假设 D 有 r 个 1, 则

$$D = (I_n - E_{r+1,r+1})\cdots(I_n - E_{nn}).$$

第三类初等阵已经是这种形状了, 即 $P_{ij}(c) = I_n + cE_{ij}$. 对第二类初等阵 $P_i(c)$, 显然我们有 $P_i(c) = I_n + (c-1)E_{ii}$. 对第一类初等阵 P_{ij}, 由练习 2.10可知, 只用第三类初等变换就可以将 P_{ij} 化为 $P_n(-1) = \mathrm{diag}\{1, \cdots, 1, -1\}$, 因此对第一类初等阵结论也成立. 具体地, 我们有

$$P_{ij} \cdot P_{ij}(-1)\,P_j(-1)\,P_{ji}(-1)\,P_{ij}(1) = I_n.$$

由此可得

$$\begin{aligned} P_{ij} &= \left[P_{ij}(-1)\,P_j(-1)\,P_{ji}(-1)\,P_{ij}(1)\right]^{-1} = P_{ij}^{-1}(1)\,P_{ji}^{-1}(-1)\,P_j^{-1}(-1)\,P_{ij}^{-1}(-1) \\ &= P_{ij}(-1)\,P_{ji}(1)\,P_j(-1)\,P_{ij}(1) = \left(I_n - E_{ij}\right)\left(I_n + E_{ji}\right)\left(I_n - 2E_{jj}\right)\left(I_n + E_{ij}\right). \end{aligned}$$

□

2.3 伴随矩阵

定义 2.3 (伴随矩阵定义)

设 $A=(a_{ij})_{n\times n}$, 若

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

其中 A_{ij} 是 a_{ij} 的代数余子式. 则称 A^* 为 A 的**伴随矩阵**.

定理 2.2

设 A 为 n 阶矩阵,$n\geq 2$, 则

(i) $AA^*=A^*A=|A|\,I_n$.

(ii) 当 A 可逆时, 有 $A^{-1}=\dfrac{1}{|A|}A^*$.

证明 由伴随矩阵的定义不难证明. □

命题 2.25

设 A 为 n 阶矩阵, 满足 $A^m=I_n$, 则 $(A^*)^m=I_n$.

证明 由 $A^m=I_n$ 得 $|A|^m=1\neq 0$, 于是矩阵 A 可逆. 又 $A^*=|A|A^{-1}$, 故 $(A^*)^m=|A|^m(A^{-1})^m=(A^m)^{-1}=I_n$. □

定理 2.3 (矩阵乘积的伴随)

设 A,B 为 n 阶矩阵,$n\geq 2$, 则 $(AB)^*=B^*A^*$.

证明 证法一 (Cauchy-Binet 公式推论):设 $C=AB$. 记 M_{ij},N_{ij},P_{ij} 分别是 A,B,C 中第 (i,j) 元素的余子式,A_{ij},B_{ij},C_{ij} 分别是 A,B,C 中第 (i,j) 元素的代数余子式. 注意到

$$A^*=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix},\quad B^*=\begin{pmatrix} B_{11} & B_{21} & \cdots & B_{n1} \\ B_{12} & B_{22} & \cdots & B_{n2} \\ \vdots & \vdots & & \vdots \\ B_{1n} & B_{2n} & \cdots & B_{nn} \end{pmatrix},$$

B^*A^* 的第 (i,j) 元素为 $\sum\limits_{k=1}^n B_{ki}A_{jk}$. 而 C^* 的第 (i,j) 元素就是 $C_{ji}=(-1)^{j+i}P_{ji}$.

由Cauchy-Binet 公式推论可得

$$\begin{aligned} C_{ji}&=(-1)^{j+i}P_{ji}=(-1)^{j+i}\sum_{k=1}^n M_{jk}N_{ki} \\ &=\sum_{k=1}^n(-1)^{j+k}M_{jk}(-1)^{i+k}N_{ki}=\sum_{k=1}^n A_{jk}B_{ki} \end{aligned}$$

故结论成立.

证法二 (摄动法):若 A,B 均为非异阵, 则 $A^*=|A|A^{-1},B^*=|B|B^{-1}$, 从而

$$(AB)^*=|AB|(AB)^{-1}=|A||B|(B^{-1}A^{-1})=(|B|B^{-1})(|A|A^{-1})=B^*A^*.$$

由命题 2.42, 可知对于一般的方阵 A,B, 可取到一列有理数 $t_k\to 0$, 使得 t_kI_n+A 与 t_kI_n+B 均为非异阵. 由

非异阵情形的证明可得

$$((t_kI_n+A)(t_kI_n+B))^*=(t_kI_n+B)^*(t_kI_n+A)^*.$$

注意到上式两边均为 n 阶方阵, 其元素都是 t_k 的多项式, 从而关于 t_k 连续. 上式两边同时取极限, 令 $t_k\to 0$, 即有 $(AB)^*=B^*A^*$ 成立. □

定理 2.4 (伴随矩阵的秩)

设 A 为 n 阶矩阵,$n\geq 2$, 则

$$\operatorname{rank}A^*=\begin{cases}n, & \operatorname{rank}A=n,\\ 1, & \operatorname{rank}A=n-1,\\ 0, & \operatorname{rank}A<n-1.\end{cases}$$

证明 当 $\operatorname{rank}A=n$ 时, 则 $|A|\neq 0$,A 可逆, 又 $AA^*=A^*A=|A|I_n$, 两边同时取行列式, 可得 $|A^*|\cdot|A|=|A^*A|=||A|I_n|=|A|^n$, 于是 $|A^*|=|A|^{n-1}\neq 0$. 所以 $\operatorname{rank}A^*=n$.

当 $\operatorname{rank}A=n-1$ 时,A 至少存在一个 $n-1$ 阶子式不等于 0, 故 $A^*\neq 0$, 即 $\operatorname{rank}A^*\geq 1$; 由 $\operatorname{rank}A<n$ 知 $|A|=0$, 从而 $AA^*=|A|E=0$, 故由定理 3.3可知 $\operatorname{rank}A^*\leq n-\operatorname{rank}A=1$, 于是 $\operatorname{rank}A^*=1$. (另证: 若 $\boldsymbol{A}$ 的秩等于 $n-1$, 则由命题 3.39可知 $\boldsymbol{A}^*$ 的 n 个列向量都成比例且至少有一列不为零, 故 $\boldsymbol{A}^*$ 的秩等于 1.)

当 $\operatorname{rank}A<n-1$ 时,A 的所有 $n-1$ 阶子式均等于 0, 即 $A^*=0$, 故 $\operatorname{rank}A^*=0$. □

命题 2.26 (伴随矩阵的性质)

设 A 为 n 阶矩阵,$n\geq 2$, 则

1. $(A^{\mathrm{T}})^*=(A^*)^{\mathrm{T}}$.
2. $(kA)^*=k^{n-1}A^*$,k 为常数.
3. 若 A 为可逆阵, 则 A^* 也可逆, 并且 $(A^{-1})^*=(A^*)^{-1}$.
4. $(A^m)^*=(A^*)^m$,m 为正整数.
5. $|A^*|=|A|^{n-1}$.
6. $(A^*)^*=|A|^{n-2}A$.

证明

1. 由伴随矩阵的定义及行列式的性质即得.
2. 由伴随矩阵的定义及行列式的性质即得.
3. 由定理 2.3可知 $A^*\left(A^{-1}\right)^*=\left(A^{-1}A\right)^*=I_n^*=I_n$. 从而 $(A^{-1})^*=(A^*)^{-1}$.
4. 多次利用定理 2.3即得.
5. 证法一:当 A 可逆时, 有 $A^*=|A|A^{-1}$, 从而 $|A^*|=|A|^{n-1}$; 当 A 不可逆时, 有 $\operatorname{rank}A<n$, 由定理 2.4知 $\operatorname{rank}A^*<n$. 于是 $|A^*|=|A|=0$, 故 $|A^*|=|A|^{n-1}$.

   若 A 是非异阵, 有 $A^*=|A|A^{-1}$, 从而 $|A^*|=|A|^{n-1}$. 对于一般的方阵 A, 由命题 2.42可知, 可取到一列有理数 $t_k\to 0$, 使得 t_kI_n+A 为非异阵. 由非异阵情形的证明可得

   $$|(t_kI_n+A)^*|=|t_kI_n+A|^{n-1}.$$

   注意到上式两边均为行列式的幂次, 其值都是 t_k 的多项式, 从而关于 t_k 连续. 上式两边同时取极限 (上式两边都是关于 t_k 的多项式函数), 令 $t_k\to 0$, 即有 $|A^*|=|A|^{n-1}$ 成立.

   证法二:见白皮书.
6. 证法一:当 A 可逆时,A^* 也可逆, 且 $(A^*)^{-1}=\dfrac{1}{|A|}A$, 从而由 伴随矩阵的性质5.得

   $$(A^*)^*=|A^*|(A^*)^{-1}=|A|^{n-1}\frac{1}{|A|}A=|A|^{n-2}A.$$

   当 A 不可逆时, 则 $|A|=0$, 且由定理 2.4及 $n\geq 2$ 知 $\operatorname{rank}A^*\leq 1<n-1$, 从而 $\operatorname{rank}(A^*)^*=0$, 即 $(A^*)^*=0$, 因

此 $(A^*)^* = |A|^{n-2}A$.

若 A 是非异阵,A^* 也可逆, 且 $(A^*)^{-1} = \dfrac{1}{|A|}A$, 从而由 伴随矩阵的性质5.得

$$(A^*)^* = |A^*|(A^*)^{-1} = |A|^{n-1}\frac{1}{|A|}A = |A|^{n-2}A.$$

对于一般的方阵 A, 由命题 2.42可知, 可取到一列有理数 $t_k \to 0$, 使得 $t_kI_n + A$ 为非异阵. 由非异阵情形的证明可得

$$((t_kI_n + A)^*)^* = |t_kI_n + A|^{n-2}(t_kI_n + A).$$

注意到上式两边均为 n 阶方阵, 其元素都是 t_k 的多项式 (上式两边的矩阵每个元素都是关于 t_k 的多项式函数), 从而关于 t_k 连续. 上式两边同时取极限, 令 $t_k \to 0$, 即有 $(A^*)^* = |A|^{n-2}A$ 成立.

证法二:见白皮书.

□

命题 2.27 (伴随矩阵的继承性)

1. 对角矩阵的伴随矩阵是对角矩阵;
2. 对称矩阵的伴随矩阵是对称矩阵;
3. 上 (下) 三角矩阵的伴随矩阵是上 (下) 三角矩阵;
4. 可逆矩阵的伴随矩阵是可逆;
5. 正交矩阵的伴随矩阵是正交矩阵;
6. 半正定 (正定) 矩阵的伴随矩阵是半正定 (正定) 矩阵;
7. 可对角化矩阵的伴随矩阵是可对角化矩阵.

♠

证明

1. 设 n 阶矩阵 $A = (a_{ij})$, $n \geq 2$.
2. 若 A 为对角矩阵, 则 $a_{ij} = 0(i \neq j)$, 从而 $i \neq j$ 时,M_{ij} 是对角行列式, 且主对角元必有零, 即 $M_{ij} = 0$, 故 $A_{ij} = 0$, 于是 A^* 为对角矩阵.
3. 若 A 为对称矩阵, 则 $a_{ij} = a_{ji}(i, j = 1, 2, \cdots, n)$, 因此 $i, j = 1, 2, \cdots, n$ 时,M_{ij} 是对称行列式, 从而 $A_{ij} = A_{ji}$, 即 A^* 为对称矩阵.
4. 若 A 为上三角矩阵, 则 $1 \leq j < i \leq n$ 时,$a_{ij} = 0$, 所以 $1 \leq i < j \leq n$ 时,M_{ij} 是上三角行列式, 且主对角元必有零, 即 $M_{ij} = 0$, 从而 $A_{ij} = 0$, 所以 A^* 为上三角矩阵. 同理可证: 下三角矩阵的伴随矩阵是下三角矩阵.
5. 由 $|A| \neq 0$ 和 $A^* = |A|A^{-1}$ 即知.
6. 因为 A 为正交矩阵等价于 $A^{-1} = A^{\mathrm{T}}$, 所以 $|A|^{-1} = |A|$. 从而由定理 2.2(ii), 有
$$(A^*)^{-1} = |A|^{-1}A = (|A|A^{\mathrm{T}})^{\mathrm{T}} = (A^*)^{\mathrm{T}},$$
故 A^* 为正交矩阵.
7. 由于 A 为半正定矩阵等价于存在实矩阵 C, 使得 $A = C^{\mathrm{T}}C$, 因此由定理 2.3和伴随矩阵的性质1., 有
$$A^* = (C^{\mathrm{T}}C)^* = C^*(C^{\mathrm{T}})^* = C^*(C^*)^{\mathrm{T}},$$
于是 A^* 为半正定矩阵. 当 A 为正定矩阵时, 同理可证 A^* 为正定矩阵.
8. 若 A 可对角化, 则存在可逆矩阵 P, 使得 $A = P\Lambda P^{\mathrm{T}}$, 其中 Λ 为对角矩阵, 从而由定理 2.3和伴随矩阵的性质1., 有
$$A^* = (P^{\mathrm{T}})^*\Lambda^*P^* = (P^*)^{\mathrm{T}}\Lambda^*P^*,$$
再根据伴随矩阵的继承性1.和伴随矩阵的性质3., 知 Λ^* 为对角矩阵,P^* 为可逆矩阵, 故 A^* 可对角化.

□

命题 2.28 (分块矩阵的伴随矩阵)

设 A 为 m 阶矩阵,B 为 n 阶矩阵, 分块对角阵 C 为

$$C = \begin{pmatrix} A & O \\ O & B \end{pmatrix}.$$

则分块对角阵 C 的伴随矩阵为:

$$C^* = \begin{pmatrix} |B|A^* & O \\ O & |A|B^* \end{pmatrix}.$$

证明 证法一: 设 $A = (a_{ij})_{m\times m}$, 元素 a_{ij} 的余子式和代数余子式分别记为 M_{ij} 和 A_{ij};$B = (b_{ij})_{n\times n}$, 元素 b_{ij} 的余子式和代数余子式分别记为 N_{ij} 和 B_{ij}. 利用 Laplace 定理可以容易地计算出: 当 $1 \le i, j \le m$ 时,C 的第 (i, j) 元素的代数余子式为 $(-1)^{i+j}M_{ij}|B| = |B|A_{ij}$; 当 $m+1 \le i, j \le m+n$ 时, 由 Laplace 定理, 可知 C 的第 (i, j) 元素的代数余子式为 $(-1)^{i+j}N_{i-m,j-m}|A| = |A|B_{i-m,j-m}$; 当 i, j 属于其他范围时, 由 Laplace 定理, 当 $1 \le i \le m, m \le j \le m+n$ 时, 将其按前 m 列展开, 当 $m \le i \le m+n, 1 \le j \le m$ 时, 将其按前 m 行展开, 可得 C 的第 (i, j) 元素的代数余子式等于零. 因此我们有

$$C^* = \begin{pmatrix} |B|A^* & O \\ O & |A|B^* \end{pmatrix}.$$

证法二:若 A, B 均为非异阵, 则

$$C\begin{pmatrix} |B|A^* & O \\ O & |A|B^* \end{pmatrix} = \begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} |B|A^* & O \\ O & |A|B^* \end{pmatrix} = \begin{pmatrix} |B|AA^* & O \\ O & |A|BB^* \end{pmatrix} = \begin{pmatrix} |A||B|I_m & O \\ O & |A||B|I_n \end{pmatrix} = |C|I_{m+n} = CC^*,$$

注意到 C 非异, 故由上式可得

$$C^* = \begin{pmatrix} A & O \\ O & B \end{pmatrix}^* = \begin{pmatrix} |B|A^* & O \\ O & |A|B^* \end{pmatrix}.$$

对于一般的方阵 A, B, 由命题 2.42可知, 可取到一列有理数 $t_k \to 0$, 使得 $t_kI_m + A$ 与 $t_kI_n + B$ 均为非异阵. 由非异阵情形的证明可得

$$\begin{pmatrix} t_kI_m + A & O \\ O & t_kI_n + B \end{pmatrix}^* = \begin{pmatrix} |t_kI_n + B|(t_kI_m + A)^* & O \\ O & |t_kI_m + A|(t_kI_n + B)^* \end{pmatrix}.$$

注意到上式两边均为 $m+n$ 阶方阵, 其元素都是 t_k 的多项式, 从而关于 t_k 连续. 上式两边同时取极限, 令 $t_k \to 0$, 即有 $\begin{pmatrix} A & O \\ O & B \end{pmatrix}^* = \begin{pmatrix} |B|A^* & O \\ O & |A|B^* \end{pmatrix}$ 成立. □

2.3.1 练习

练习 2.12 设 A, B 为 n 阶方阵, 满足 $AB = BA$, 证明:$AB^* = B^*A$.

证明 若 B 为非异阵, 则由 $AB = BA$ 可得 $AB^{-1} = B^{-1}A$. 又 $B^* = |B|B^{-1}$, 于是 $AB^* = B^*A$ 成立. 对于一般的方阵 B, 可取到一列有理数 $t_k \to 0$, 使得 $t_kI_n + B$ 为非异阵, 此时 $A(t_kI_n + B) = (t_kI_n + B)A$ 仍然成立. 由非异阵情形的证明可得

$$A(t_kI_n + B)^* = (t_kI_n + B)^*A.$$

注意到上式两边均为 n 阶方阵, 其元素都是 t_k 的多项式, 从而关于 t_k 连续. 上式两边同时取极限, 令 $t_k \to 0$, 即有

$AB^* = B^*A$ 成立. □

练习 2.13 设 n 阶矩阵

$$A = \begin{pmatrix} 2 & 2 & 2 & \cdots & 2 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

求 $\sum\limits_{i,j=1}^{n} A_{ij}$.

解 解法一:显然 $|A| = 2$, 用初等变换不难求出

$$A^{-1} = \begin{pmatrix} \frac{1}{2} & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

故

$$A^* = 2A^{-1} = \begin{pmatrix} 1 & -2 & 0 & \cdots & 0 & 0 \\ 0 & 2 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -2 \\ 0 & 0 & 0 & \cdots & 0 & 2 \end{pmatrix}.$$

将 A^* 的所有元素加起来, 可得 $\sum\limits_{i,j=1}^{n} A_{ij} = 1$.

解法二:由命题 1.5可得

$$-\sum_{i,j=1}^{n} A_{ij} = \begin{vmatrix} 2 & 2 & 2 & \cdots & 2 & 1 \\ 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 2 & 2 & \cdots & 2 & 1 \\ 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -\frac{1}{2} \end{vmatrix} = -1.$$

于是 $\sum\limits_{i,j=1}^{n} A_{ij} = 1$.

解法三:由大拆分法可得 $|A(-1)| = |A| - \sum\limits_{i,j=1}^{n} A_{ij}$, 且

$$|A(-1)| = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ -1 & 0 & \cdots & 0 & 0 \\ -1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ -1 & -1 & \cdots & -1 & 0 \end{vmatrix} = (-1)^{n+1} \begin{vmatrix} -1 & 0 & \cdots & 0 \\ -1 & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ -1 & -1 & \cdots & -1 \end{vmatrix} = 1.$$

故 $\sum\limits_{i,j=1}^{n} A_{ij} = |A(-1)| - |A|$.

解法四:由例题 1.5可得

$$\sum_{i,j=1}^{n} A_{ij} = \begin{vmatrix} 0 & 0 & \cdots & 0 & 1 \\ -1 & 0 & \cdots & 0 & 1 \\ 0 & -1 & \cdots & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{vmatrix} = (-1)^{n+1} \begin{vmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 \end{vmatrix} = 1.$$

□

2.4 矩阵的迹

命题 2.29 (矩阵迹的性质)

设 A, B 是 n 阶矩阵, 则有:

1. (线性)$\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$,$\mathrm{tr}(kA) = k\mathrm{tr}(A)$;
2. (对称性)$\mathrm{tr}(A') = \mathrm{tr}(A)$;
3. (交换性)$\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

证明 根据矩阵迹的定义及矩阵乘法的定义容易验证. □

命题 2.30 (矩阵迹的刻画)

设 f 是数域 $\mathbb{F}$ 上 n 阶矩阵集合到 $\mathbb{F}$ 的一个映射, 它满足下列条件:

(1) 对任意的 n 阶矩阵 $A, B, f(A+B) = f(A) + f(B)$;

(2) 对任意的 n 阶矩阵 A 和 $\mathbb{F}$ 中的数 $k, f(kA) = kf(A)$;

(3) 对任意的 n 阶矩阵 $A, B, f(AB) = f(BA)$;

(4) $f(I_n) = n$.

求证:f 就是迹, 即 $f(A) = \mathrm{tr}(A)$ 对一切 $\mathbb{F}$ 上 n 阶矩阵 A 成立.

笔记 这个命题给出了迹的刻画, 它告诉我们迹函数由线性、交换性和正规性 (即单位矩阵处的取值为其阶数) 唯一决定.

证明 设 E_{ij} 是 n 阶基础矩阵. 由 (1) 和 (4), 有

$$n = f(I_n) = f(E_{11} + E_{22} + \cdots + E_{nn}) = f(E_{11}) + f(E_{22}) + \cdots + f(E_{nn}).$$

又由 (3), 有

$$f(E_{ii}) = f(E_{ij}E_{ji}) = f(E_{ji}E_{ij}) = f(E_{jj}),$$

所以 $f(E_{ii}) = 1(1 \le i \le n)$. 另一方面, 若 $i \ne j$, 则

$$f(E_{ij}) = f(E_{i1}E_{1j}) = f(E_{1j}E_{i1}) = f(O) = f(0 \cdot I_n) = 0 \cdot f(I_n) = 0.$$

设 n 阶矩阵 $A = (a_{ij})$, 则

$$f(A) = f\left(\sum_{i,j=1}^{n} a_{ij}E_{ij}\right) = \sum_{i,j=1}^{n} a_{ij}f(E_{ij}) = \sum_{i=1}^{n} a_{ii} = \mathrm{tr}(A).$$

□

例题 2.13 求证: 不存在 n 阶矩阵 A, B, 使得 $AB - BA = kI_n(k \ne 0)$.

证明 用反证法证明. 若存在 n 阶矩阵 A, B 满足条件 $AB - BA = kI_n(k \ne 0)$, 则

$$kn = \mathrm{tr}(kI_n) = \mathrm{tr}(AB - BA) = \mathrm{tr}(AB) - \mathrm{tr}(BA) = 0$$

矛盾. □

例题 2.14 设 A 是 n 阶矩阵,P 是同阶可逆阵, 求证:$\mathrm{tr}(P^{-1}AP)=\mathrm{tr}(A)$, 即相似矩阵具有相同的迹.

证明 因为 $\mathrm{tr}(AB)=\mathrm{tr}(BA)$, 故 $\mathrm{tr}(P^{-1}AP)=\mathrm{tr}(APP^{-1})=\mathrm{tr}(A)$. □

例题 2.15 设 $A_1,A_2,\cdots,A_k$ 是实对称阵且 $A_1^2+A_2^2+\cdots+A_k^2=O$, 证明: 每个 $A_i=O$.

证明 对题设中的等式两边同时取迹, 可得

$$0=\mathrm{tr}(O)=\mathrm{tr}(A_1^2+A_2^2+\cdots+A_k^2)=\mathrm{tr}(A_1A_1')+\mathrm{tr}(A_2A_2')+\cdots+\mathrm{tr}(A_kA_k').$$

又由于 $\mathrm{tr}(A_iA_i')\geq 0$, 从而只可能是 $\mathrm{tr}(A_iA_i')=0(1\leq i\leq k)$, 再次由零矩阵的充要条件可得 $A_i=O(1\leq i\leq k)$. □

命题 2.31

1. 设 n 阶实矩阵 A 适合 $A'=-A$, 如果存在同阶实矩阵 B, 使得 $AB=B$, 则 $B=O$;
2. 设 n 阶复矩阵 A 适合 $\overline{A}'=-A$, 如果存在同阶复矩阵 B, 使得 $AB=B$, 则 $B=O$.

证明

1. 在等式 $AB=B$ 两边同时左乘 B' 可得

$$B'AB=B'B.$$

上式两边同时转置并注意到 $A'=-A$, 可得

$$B'B=(B'B)'=(B'AB)'=B'A'B=-B'AB=-B'B,$$

从而有 $B'B=O$. 两边同时取迹, 由零矩阵的充要条件可得 $B=O$.

2. 证明与 1 类似.

□

命题 2.32

设 A 为 n 阶实矩阵, 求证:$\mathrm{tr}(A^2)\leq\mathrm{tr}(AA')$, 等号成立当且仅当 A 是对称阵.

证明 若已知 $\mathrm{tr}(AA')\geq\mathrm{tr}(A^2)$, 则由迹的线性、对称性、交换性和正定性可得

$$\begin{aligned}&\mathrm{tr}((A-A')(A-A')')\\=&\mathrm{tr}((A-A')(A'-A))=\mathrm{tr}(AA'-A^2-(A')^2+A'A)\\=&2\mathrm{tr}(AA')-2\mathrm{tr}(A^2)\geq 0,\end{aligned}$$

故要证的不等式成立. 若上述不等式的等号成立, 则由迹的正定性可知 $A-A'=O$, 即 A 为对称阵. 若已知 A 为对称阵, 则 $\mathrm{tr}(AA')=\mathrm{tr}(A^2)$ 显然成立. □

命题 2.33

设 A,B 是两个 n 阶矩阵, 使得 $\mathrm{tr}(ABC)=\mathrm{tr}(CBA)$ 对任意 n 阶矩阵 C 成立, 求证:$AB=BA$.

证明 设 $AB=(d_{ij}),BA=(e_{ij})$, 令 $C=E_{kl}(1\leq k,l\leq n)$, 则

$$\mathrm{tr}(ABC)=d_{lk},\mathrm{tr}(CBA)=e_{lk},$$

因此 $d_{lk}=e_{lk}(1\leq k,l\leq n)$, 即有 $AB=BA$. □

注 注若 A,B 是实 (复) 矩阵, 我们还可以通过迹的正定性来证明结论. 事实上, 由迹的交换性和线性可得 $\mathrm{tr}((AB-BA)C)=0$, 令 C 为 $AB-BA$ 的转置 (共轭转置), 再由零矩阵的充要条件即得结论.

例题 2.16 若 n 阶实方阵 A 满足 $AA'=I_n$, 则称为正交矩阵. 证明: 不存在 n 阶正交矩阵 A,B 满足 $A^2=cAB+B^2$, 其中 c 是非零常数.

证明 用反证法, 设存在 n 阶正交阵 A,B, 使得 $A^2=cAB+B^2(c\neq 0)$. 在等式两边同时左乘 A', 右乘 B', 可得 $AB'=cI_n+A'B$, 从而 $cI_n=A'B-AB'$. 两边同时取迹, 可得 $0\neq nc=\mathrm{tr}(cI_n)=\mathrm{tr}(A'B)-\mathrm{tr}(AB')=\mathrm{tr}((A'B)')-$

$\mathrm{tr}(AB')=\mathrm{tr}(B'A)-\mathrm{tr}(AB')=0$, 矛盾. □

例题 2.17 设 A,B 为 n 阶实对称阵, 证明:$\mathrm{tr}((AB)^2)\le\mathrm{tr}(A^2B^2)$, 并求等号成立的充要条件.

证明 由命题 2.32, 再结合 A,B 的对称性可得

$$\mathrm{tr}\left((AB)^2\right)\leqslant\mathrm{tr}\left((AB)(AB)'\right)=\mathrm{tr}(ABBA)=\mathrm{tr}\left(A^2B^2\right).$$

等号成立当且仅当 AB 也为实对称矩阵, 即 $AB=B'A'=BA$. □

2.5 矩阵乘法与行列式计算

命题 2.34 (可以写成两个矩阵 (向量) 乘积的矩阵)

若已知矩阵

$$A=\begin{pmatrix} a_{11}b_{11}+a_{12}b_{12}+\cdots+a_{1n}b_{1n} & a_{21}b_{11}+a_{22}b_{12}+\cdots+a_{2n}b_{1n} & \cdots & a_{n1}b_{11}+a_{n2}b_{12}+\cdots+a_{nn}b_{1n}\\ a_{11}b_{21}+a_{12}b_{22}+\cdots+a_{1n}b_{2n} & a_{21}b_{21}+a_{22}b_{22}+\cdots+a_{2n}b_{2n} & \cdots & a_{n1}b_{21}+a_{n2}b_{22}+\cdots+a_{nn}b_{2n}\\ \vdots & \vdots & & \vdots\\ a_{11}b_{n1}+a_{12}b_{n2}+\cdots+a_{1n}b_{nn} & a_{21}b_{n1}+a_{22}b_{n2}+\cdots+a_{2n}b_{nn} & \cdots & a_{n1}b_{n1}+a_{n2}b_{n2}+\cdots+a_{nn}b_{nn}\end{pmatrix}.$$

则矩阵 A 可以写成 BC, 其中

$$B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots\\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{pmatrix},C=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1}\\ a_{12} & a_{22} & \cdots & a_{n2}\\ \vdots & \vdots & & \vdots\\ a_{1n} & a_{2n} & \cdots & a_{nn}\end{pmatrix}.$$

即

$$A=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots\\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{pmatrix}\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1}\\ a_{12} & a_{22} & \cdots & a_{n2}\\ \vdots & \vdots & & \vdots\\ a_{1n} & a_{2n} & \cdots & a_{nn}\end{pmatrix}=BC.$$

特别地, 若矩阵 A 的行/列向量成比例, 不妨设

$$A=\begin{pmatrix} a_1b_1 & a_2b_1 & \cdots & a_nb_1\\ a_1b_2 & a_2b_2 & \cdots & a_nb_2\\ \vdots & \vdots & & \vdots\\ a_1b_n & a_2b_n & \cdots & a_nb_n\end{pmatrix}.$$

则令 $\alpha=(a_1,a_2,\cdots,a_n),\beta=(b_1,b_2,\cdots,b_n)$, 就有 $A=\beta'\alpha=\begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}\begin{pmatrix} a_1 & a_2 & \cdots & a_n\end{pmatrix}.$

注 若矩阵的列向量成比例, 则行向量也一定成比例. 反之也成立.

笔记 观察原矩阵 A 不难发现: **矩阵 A 的每一行沿行方向只有 a_{ij}(的角标) 改变, 而 b_{kl}(的角标) 并不改变; 而矩阵 A 的每一列沿列方向只有 b_{kl}(的角标) 改变,a_{ij}(的角标) 并不改变.** 因此具有这种性质的矩阵, 都可以按照这个命题将其写成两个矩阵的乘积. 特别地, **若矩阵的行/列向量成比例, 则一定可以将其写成两个向量的乘积.**

记忆小技巧: 只需要记住矩阵 B 的形式 (沿行方向不变的项写在前面作为矩阵 B 的元素), 然后结合原矩阵, 利用矩阵乘法就能写出矩阵 C. 即**按行变化的项写左边 (作为矩阵 B 的元素), 按列变化的项写右边 (作为矩阵 C 的元素).**

注 拆分后的矩阵 B 的行数与原矩阵 A 相同, 矩阵 C 的列数与原矩阵 A 相同. 但是矩阵 B 的列数与矩阵 C 的行数可以任意选取, 只要满足 $BC = A$ 即可.

证明 利用矩阵乘法容易得到证明. □

命题 2.35 (一些能写成两个向量乘积的矩阵)

1. $\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} = \alpha\alpha'$, 其中 $\alpha = (1, 1, \cdots, 1)'$.

2. 若矩阵 A 的行/列向量成比例, 不妨设

$$A = \begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{pmatrix}.$$

则有 $A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix}$.

笔记 这里的 a_i 可以是行向量,b_i 可以是列向量. 此时矩阵 A 的元素就是 a_ib_i 仍然是一个数. 并且此时矩阵 A 能够分解的条件应该改为**矩阵 A 的每一行都有公共的行向量 a_i, 每一列都有公共的列向量 b_i.**

注 若 a_i, b_i 是上述向量, 则根据矩阵乘法, 可知 a_i 的列数可以任意选取,b_i 的行数可以任意选取. 此时只要确定每个向量 a_i, b_i 就可以确定矩阵 A 的分解式.

例题 2.18 设 $s_k = x_1^k + x_2^k + \cdots + x_n^k (k \geq 1), s_0 = n$,

$$S = \begin{vmatrix} s_0 & s_1 & s_2 & \cdots & s_{n-1} \\ s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-2} \end{vmatrix}.$$

求 $|S|$ 的值并证明若 x_i 是实数, 则 $|S| \geq 0$.

解 设

$$V = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}$$

则 $S = VV'$, 因此

$$|S| = |V|^2 = \prod_{1 \leq i < j \leq n} (x_j - x_i)^2 \geq 0.$$

□

例题 2.19 设 $s_k = x_1^k + x_2^k + \cdots + x_n^k (k \geq 1), s_0 = n$, 计算矩阵 A 的行列式的值:

$$A = \begin{pmatrix} s_0 & s_1 & \cdots & s_{n-1} & 1 \\ s_1 & s_2 & \cdots & s_n & x \\ \vdots & \vdots & & \vdots & \vdots \\ s_n & s_{n+1} & \cdots & s_{2n-1} & x^n \end{pmatrix}.$$

解 将矩阵 A 分解为两个矩阵的乘积:

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ x_1 & x_2 & \cdots & x_n & x \\ \vdots & \vdots & & \vdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} & x^{n-1} \\ x_1^n & x_2^n & \cdots & x_n^n & x^n \end{pmatrix} \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} & 0 \\ 1 & x_2 & \cdots & x_2^{n-1} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

因此 $|A| = (x-x_1)(x-x_2)\cdots(x-x_n) \prod\limits_{1 \leq i < j \leq n} (x_j - x_i)^2$. □

例题 2.20 计算下列矩阵 A 的行列式的值:

$$A = \begin{pmatrix} x & -y & -z & -w \\ y & x & -w & z \\ z & w & x & -y \\ w & -z & y & x \end{pmatrix}.$$

解 解法一: 注意到

$$AA' = \begin{pmatrix} x & -y & -z & -w \\ y & x & -w & z \\ z & w & x & -y \\ w & -z & y & x \end{pmatrix} \begin{pmatrix} x & y & z & w \\ -y & x & w & -z \\ -z & -w & x & y \\ -w & z & -y & x \end{pmatrix} = \begin{pmatrix} u & 0 & 0 & 0 \\ 0 & u & 0 & 0 \\ 0 & 0 & u & 0 \\ 0 & 0 & 0 & u \end{pmatrix}$$

其中 $u = x^2 + y^2 + z^2 + w^2$, 因此

$$|A|^2 = (x^2 + y^2 + z^2 + w^2)^4.$$

故

$$|A| = (x^2 + y^2 + z^2 + w^2)^2.$$

在矩阵 A 中令 $x = 1, y = z = w = 0$, 显然 $|A| = 1$.

解法二: 令

$$B = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}, C = \begin{pmatrix} z & w \\ w & -z \end{pmatrix},$$

则 $|A| = \begin{vmatrix} B & -C \\ C & B \end{vmatrix}$. 由命题 2.41可得

$$|A| = |B + \mathrm{i}C||B - \mathrm{i}C| = \begin{vmatrix} x+\mathrm{i}z & -y+\mathrm{i}w \\ y+\mathrm{i}w & x-\mathrm{i}z \end{vmatrix} \begin{vmatrix} x-\mathrm{i}z & -y-\mathrm{i}w \\ y-\mathrm{i}w & x+\mathrm{i}z \end{vmatrix} = (x^2 + y^2 + z^2 + w^2)^2.$$

□

例题 2.21 计算下列矩阵 A 的行列式的值:

$$A=\begin{pmatrix}\cos\theta & \cos 2\theta & \cos 3\theta & \cdots & \cos n\theta\\ \cos n\theta & \cos\theta & \cos 2\theta & \cdots & \cos(n-1)\theta\\ \cos(n-1)\theta & \cos n\theta & \cos\theta & \cdots & \cos(n-2)\theta\\ \vdots & \vdots & \vdots & & \vdots\\ \cos 2\theta & \cos 3\theta & \cos 4\theta & \cdots & \cos\theta\end{pmatrix}.$$

解 解由上面的结论可知

$$|A|=f(\varepsilon_1)f(\varepsilon_2)\cdots f(\varepsilon_n),$$

其中 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 1 的所有 n 次方根,$f(x)=\cos\theta+x\cos 2\theta+\cdots+x^{n-1}\cos n\theta$. 令

$$g(x)=\sin\theta+x\sin 2\theta+\cdots+x^{n-1}\sin n\theta,$$

则由 De Moivre 公式可得

$$\begin{aligned}
&f(x)+\mathrm{i}g(x)=(\cos\theta+\mathrm{i}\sin\theta)+x(\cos\theta+\mathrm{i}\sin\theta)^2+\cdots+x^{n-1}(\cos\theta+\mathrm{i}\sin\theta)^n\\
&=\frac{(\cos\theta+\mathrm{i}\sin\theta)\left[1-x^n(\cos\theta+\mathrm{i}\sin\theta)^n\right]}{1-x(\cos\theta+\mathrm{i}\sin\theta)}=\frac{1-x^n(\cos\theta+\mathrm{i}\sin\theta)^n}{\cos\theta-x-\mathrm{i}\sin\theta}=\frac{(1-x^n\cos n\theta+\mathrm{i}x^n\sin n\theta)(\cos\theta-x+\mathrm{i}\sin\theta)}{[(\cos\theta-\mathrm{i}\sin\theta)-x][(\cos\theta+\mathrm{i}\sin\theta)-x]}\\
&=\frac{(1-x^n\cos n\theta+\mathrm{i}x^n\sin n\theta)(\cos\theta-x+\mathrm{i}\sin\theta)}{(\cos\theta-x)^2+\sin^2\theta}=\frac{(1-x^n\cos n\theta+\mathrm{i}x^n\sin n\theta)(\cos\theta-x+\mathrm{i}\sin\theta)}{x^2-2x\cos\theta+1}\\
&=\frac{\left[x^{n+1}\cos n\theta-x^n\cos(n+1)\theta-x+\cos\theta\right]+\mathrm{i}\left[x^{n+1}\sin n\theta+x^n\sin(n-1)\theta+\sin\theta\right]}{x^2-2x\cos\theta+1}.
\end{aligned}$$

再比较实部, 可得

$$f(x)=\frac{\cos n\theta\cdot x^{n+1}-\cos(n+1)\theta\cdot x^n-x+\cos\theta}{x^2-2\cos\theta\cdot x+1}.$$

对任意的 ε_i, 经计算并化简, 可得

$$f(\varepsilon_i)=\frac{(\cos\theta-\cos(n+1)\theta)-\varepsilon_i(1-\cos n\theta)}{[(\cos\theta+\mathrm{i}\sin\theta)-\varepsilon_i][(\cos\theta-\mathrm{i}\sin\theta)-\varepsilon_i]}.$$

注意到对任意的 a,b, 有 $a^n-b^n=(a-\varepsilon_1 b)(a-\varepsilon_2 b)\cdots(a-\varepsilon_n b)$, 因此

$$\begin{aligned}
|A|=\prod_{i=1}^n f(\varepsilon_i)&=\frac{(\cos\theta-\cos(n+1)\theta)^n-(1-\cos n\theta)^n}{(\cos n\theta+\mathrm{i}\sin n\theta-1)(\cos n\theta-\mathrm{i}\sin n\theta-1)}\\
&=\frac{(\cos\theta-\cos(n+1)\theta)^n-(1-\cos n\theta)^n}{2(1-\cos n\theta)}\\
&=\frac{2^n\sin^n\frac{n\theta}{2}\sin^n\frac{(n-2)\theta}{2}-2^n\sin^{2n}\frac{n\theta}{2}}{4\sin^2\frac{n\theta}{2}}\\
&=2^{n-2}\sin^{n-2}\frac{n\theta}{2}\left(\sin^n\frac{(n+2)\theta}{2}-\sin^n\frac{n\theta}{2}\right).
\end{aligned}$$

□

2.6 Cauchy-Binet 公式

定理 2.5 (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, 它是由 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. 同理定义 B 的 s 阶子式.

(1) 若 $m>n$, 则有 $|AB|=0$;

(2) 若 $m \le n$, 则有

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

证明

(1) 若 $m > n$, 则 $r(AB) \le \min\{r(A), r(B)\} \le \min\{m, n\} = n < m$, 故 $|-AB| = 0$.

(2)

□

推论 2.2 (Cauchy-Binet 公式推论)

设 $A = (a_{ij})$ 是 $m \times n$ 矩阵,$B = (b_{ij})$ 是 $n \times m$ 矩阵,r 是一个正整数且 $r \le m$.

(1) 若 $r > n$, 则 AB 的任意 r 阶子式都等于零;

(2) 若 $r \le n$, 则 AB 的 r 阶子式

$$AB\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = \sum_{1\le k_1<k_2<\cdots<k_r\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \cdots & k_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}.$$

例题 2.22 设 $n \ge 3$, 证明下列矩阵是奇异阵:

$$A = \begin{pmatrix} \cos(\alpha_1-\beta_1) & \cos(\alpha_1-\beta_2) & \cdots & \cos(\alpha_1-\beta_n) \\ \cos(\alpha_2-\beta_1) & \cos(\alpha_2-\beta_2) & \cdots & \cos(\alpha_2-\beta_n) \\ \vdots & \vdots & & \vdots \\ \cos(\alpha_n-\beta_1) & \cos(\alpha_n-\beta_2) & \cdots & \cos(\alpha_n-\beta_n) \end{pmatrix}.$$

解 注意到

$$A = \begin{pmatrix} \cos(\alpha_1-\beta_1) & \cos(\alpha_1-\beta_2) & \cdots & \cos(\alpha_1-\beta_n) \\ \cos(\alpha_2-\beta_1) & \cos(\alpha_2-\beta_2) & \cdots & \cos(\alpha_2-\beta_n) \\ \vdots & \vdots & & \vdots \\ \cos(\alpha_n-\beta_1) & \cos(\alpha_n-\beta_2) & \cdots & \cos(\alpha_n-\beta_n) \end{pmatrix} = \begin{pmatrix} \cos\alpha_1 & \sin\alpha_1 \\ \cos\alpha_2 & \sin\alpha_2 \\ \vdots & \vdots \\ \cos\alpha_n & \sin\alpha_n \end{pmatrix} \begin{pmatrix} \cos\beta_1 & \cos\beta_2 & \cdots & \cos\beta_n \\ \sin\beta_1 & \sin\beta_2 & \cdots & \sin\beta_n \end{pmatrix}.$$

由Cauchy-Binet 公式, 可知 $|A| = 0$. □

例题 2.23 设 A 是 $m \times n$ 实矩阵, 求证: 矩阵 AA' 的任一主子式都非负.

证明 若 $r \le n$, 则由Cauchy-Binet 公式推论可得

$$AA'\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} = \sum_{1\le j_1<j_2<\cdots<j_r\le n} \left(A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}\right)^2 \ge 0;$$

若 $r > n$, 则 AA' 的任一 r 阶主子式都等于零, 结论也成立. □

例题 2.24 设 A 是 n 阶实方阵且 $AA' = I_n$. 求证: 若 $1 \le i_1 < i_2 < \cdots < i_r \le n$, 则

$$\sum_{1\le j_1<j_2<\cdots<j_r\le n} \left(A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}\right)^2 = 1.$$

证明 对等式 $AA' = I_n$ 两边同时求 r 阶子式, 因为 $r \le n$, 所以由Cauchy-Binet 公式即得结论成立. □

例题 2.25 设 A, B 分别是 $m \times n$,$n \times m$ 矩阵, 求证:AB 和 BA 的 r 阶主子式之和相等, 其中 $1 \le r \le \min\{m, n\}$.

证明 由Cauchy-Binet 公式可得

$$\begin{aligned} &\sum_{1\le i_1<i_2<\cdots<i_r\le m} AB\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} \\ =& \sum_{1\le i_1<i_2<\cdots<i_r\le m}\ \sum_{1\le j_1<j_2<\cdots<j_r\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} B\begin{pmatrix} j_1 & j_2 & \cdots & j_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} \end{aligned}$$

$$= \sum_{1\le j_1<j_2<\cdots<j_r\le n}\sum_{1\le i_1<i_2<\cdots<i_r\le m} B\begin{pmatrix} j_1 & j_2 & \cdots & j_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}$$
$$= \sum_{1\le j_1<j_2<\cdots<j_r\le n} BA\begin{pmatrix} j_1 & j_2 & \cdots & j_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}.$$

□

引理 2.1 (Lagrange 恒等式)

证明 Lagrange 恒等式 $(n \ge 2)$:

$$\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) - \left(\sum_{i=1}^n a_ib_i\right)^2 = \sum_{1\le i<j\le n}(a_ib_j - a_jb_i)^2.$$

♡

证明 左边的式子等于

$$\begin{vmatrix} \sum_{i=1}^n a_i^2 & \sum_{i=1}^n a_ib_i \\ \sum_{i=1}^n a_ib_i & \sum_{i=1}^n b_i^2 \end{vmatrix},$$

这个行列式对应的矩阵可化为

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix}.$$

由 Cauchy - Binet 公式可得

$$\begin{vmatrix} \sum_{i=1}^n a_i^2 & \sum_{i=1}^n a_ib_i \\ \sum_{i=1}^n a_ib_i & \sum_{i=1}^n b_i^2 \end{vmatrix} = \sum_{1\le i<j\le n}\begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix}\begin{vmatrix} a_i & b_i \\ a_j & b_j \end{vmatrix} = \sum_{1\le i<j\le n}(a_ib_j - a_jb_i)^2.$$

□

定理 2.6 (Cauchy - Schwarz 不等式)

设 a_i, b_i 都是实数, 证明 Cauchy - Schwarz 不等式:

$$\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) \ge \left(\sum_{i=1}^n a_ib_i\right)^2.$$

♡

证明 由Lagrange 恒等式, 恒等式右边总非负, 即得结论. □

例题 2.26 设 A, B 都是 $m \times n$ 实矩阵, 求证:

$$|AA'||BB'| \ge |AB'|^2.$$

证明 若 $m > n$, 则 $|AA'| = |BB'| = |AB'| = 0$, 结论显然成立. 若 $m \le n$, 则由 Cauchy - Binet 公式可得

$$|AA'| = \sum_{1\le j_1<j_2<\cdots<j_m\le n}\left(A\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix}\right)^2;$$
$$|BB'| = \sum_{1\le j_1<j_2<\cdots<j_m\le n}\left(B\begin{pmatrix} 1 & 2 & \cdots & m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix}\right)^2;$$

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

再由Cauchy - Schwarz 不等式即得结论. □

2.7 分块矩阵的初等变换与降价公式 (打洞原理)

命题 2.36 (打洞原理)

(1) 设

$$\boldsymbol{M} = \begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}_{(n+m)\times(n+m)}$$

是一个方阵, 并且 $\boldsymbol{A}$ 为 n 阶可逆子方阵, 那么

$$|\boldsymbol{M}| = |\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}|.$$

(2) 设

$$\boldsymbol{M} = \begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}_{(n+m)\times(n+m)}$$

是一个方阵, 并且 $\boldsymbol{D}$ 为 n 阶可逆子方阵, 那么

$$|\boldsymbol{M}| = |\boldsymbol{D}|\cdot|\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}|.$$

笔记 打洞原理是一个重要结论, 必须要熟练掌握. 但是在实际解题中我们一般不会直接套用打洞原理的结论, 而是利用分块矩阵的初等变换书写过程.

记忆打洞原理公式的小技巧: 先记住一个模板 $|\square| = |\square|\left|\square-\square\square^{-1}\square\right|$, 然后从左往右填入子矩阵 (每个子矩阵只能填一次), 第一个 □ 填相应的可逆子矩阵, 再从主对角线上另外一个子矩阵开始, 按顺时针顺序将子矩阵填入 □ 即可.

证明 (核心想法: 利用分块矩阵的初等变换消去 $\boldsymbol{B}$ 或 $\boldsymbol{C}$)

(1) 根据分块矩阵的初等变换, 对 $\boldsymbol{M}$ 的第一行左乘 $(-\boldsymbol{C}\boldsymbol{A}^{-1})$ 再加到第二行得到

$$\begin{pmatrix}\boldsymbol{I}_n & \boldsymbol{O}\\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I}_m\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix} = \begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\end{pmatrix}.$$

然后两边同时取行列式就得到

$$|\boldsymbol{M}| = \begin{vmatrix}\boldsymbol{I}_n & \boldsymbol{O}\\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I}_m\end{vmatrix}\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix} = \left|\begin{pmatrix}\boldsymbol{I}_n & \boldsymbol{O}\\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I}_m\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}\right| = \begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\end{vmatrix} = |\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}|.$$

(2) 根据分块矩阵的初等变换, 对 $\boldsymbol{M}$ 的第二行左乘 $(-\boldsymbol{B}\boldsymbol{D}^{-1})$ 再加到第一行得到

$$\begin{pmatrix}\boldsymbol{I}_n & -\boldsymbol{B}\boldsymbol{D}^{-1}\\ \boldsymbol{O} & \boldsymbol{I}_m\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix} = \begin{pmatrix}\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C} & \boldsymbol{O}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}.$$

然后两边同时取行列式就得到

$$|\boldsymbol{M}| = \begin{vmatrix}\boldsymbol{I}_n & -\boldsymbol{B}\boldsymbol{D}^{-1}\\ \boldsymbol{O} & \boldsymbol{I}_m\end{vmatrix}\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix} = \left|\begin{pmatrix}\boldsymbol{I}_n & -\boldsymbol{B}\boldsymbol{D}^{-1}\\ \boldsymbol{O} & \boldsymbol{I}_m\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}\right| = \begin{vmatrix}\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C} & \boldsymbol{O}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix} = |\boldsymbol{D}|\cdot|\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}|.$$

□

推论 2.3 (打洞原理推论)

设 $\boldsymbol{A}$ 是 $n\times m$ 矩阵,$\boldsymbol{B}$ 是 $m\times n$ 矩阵, 则

$$\lambda^m|\lambda\boldsymbol{I}_n-\boldsymbol{A}\boldsymbol{B}| = \lambda^n|\lambda\boldsymbol{I}_m-\boldsymbol{B}\boldsymbol{A}|.$$

笔记 这个推论能将原本复杂的矩阵 AB 通过交换顺序变成相对简单的矩阵 BA. 例如:例题 2.35.

注 这是由打洞原理得到的一个重要结论, 也需要熟练掌握. 同样地, 在实际解题中如果不能直接套用打洞原理推论的结论, 就需要利用分块矩阵的初等变换书写过程.

证明 当 $\lambda=0$ 时, 结论显然成立.

当 $\lambda\neq 0$ 时, 根据分块矩阵的初等变换可知

$$\begin{pmatrix} \boldsymbol{I}_n & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}=\begin{pmatrix} \lambda\boldsymbol{I}_n-\boldsymbol{AB} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix},$$

$$\begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\frac{1}{\lambda}\boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}=\begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m-\frac{1}{\lambda}\boldsymbol{BA} \end{pmatrix}.$$

再对上式两边分别取行列式得到

$$\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix}=\begin{vmatrix} \boldsymbol{I}_n & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m \end{vmatrix}\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix}=\left|\begin{pmatrix} \boldsymbol{I}_n & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}\right|=\begin{vmatrix} \lambda\boldsymbol{I}_n-\boldsymbol{AB} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix}=|\lambda\boldsymbol{I}_n-\boldsymbol{AB}|.$$

$$\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix}=\begin{vmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\frac{1}{\lambda}\boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix}\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix}=\left|\begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\frac{1}{\lambda}\boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}\right|=\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m-\frac{1}{\lambda}\boldsymbol{BA} \end{vmatrix}=\lambda^n\left|\boldsymbol{I}_m-\frac{1}{\lambda}\boldsymbol{BA}\right|=\lambda^{n-m}|\lambda\boldsymbol{I}_m-\boldsymbol{BA}|.$$

于是 $\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix}=|\lambda\boldsymbol{I}_n-\boldsymbol{AB}|=\lambda^{n-m}|\lambda\boldsymbol{I}_m-\boldsymbol{BA}|$.. 故 $\lambda^m|\lambda\boldsymbol{I}_n-\boldsymbol{AB}|=\lambda^n|\lambda\boldsymbol{I}_m-\boldsymbol{BA}|$. □

例题 2.27 求下列矩阵的行列式的值:

$$A=\begin{pmatrix} a_1^2 & a_1a_2+1 & \cdots & a_1a_n+1 \\ a_2a_1+1 & a_2^2 & \cdots & a_2a_n+1 \\ \vdots & \vdots & & \vdots \\ a_na_1+1 & a_na_2+1 & \cdots & a_n^2 \end{pmatrix}.$$

解 令 $\boldsymbol{\Lambda}=\begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}$, 则由降价公式 (打洞原理) 我们有

$$\boldsymbol{A}=-\boldsymbol{I}_n+\begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}\boldsymbol{I}_2^{-1}\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}=(-1)^n|\boldsymbol{I}_2|\left|\boldsymbol{I}_n-\begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}\boldsymbol{I}_2^{-1}\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}\right|$$

$$=(-1)^n\begin{vmatrix} \boldsymbol{I}_2 & \boldsymbol{\Lambda}' \\ \boldsymbol{\Lambda} & \boldsymbol{I}_n \end{vmatrix}=(-1)^n|\boldsymbol{I}_n|\left|\boldsymbol{I}_2-\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}\boldsymbol{I}_n^{-1}\begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}\right|$$

$$=(-1)^n\left|\boldsymbol{I}_2-\begin{pmatrix} \sum\limits_{i=1}^n a_i^2 & \sum\limits_{i=1}^n a_i \\ \sum\limits_{i=1}^n a_i & n \end{pmatrix}\right|=(-1)^n\left[(1-n)\left(1-\sum_{i=1}^n a_i^2\right)-\left(\sum_{i=1}^n a_i\right)^2\right].$$

□

例题 2.28 计算矩阵 A 的行列式的值:

$$A=\begin{pmatrix}1+a_1^2 & a_1a_2 & \cdots & a_1a_n\\ a_2a_1 & 1+a_2^2 & \cdots & a_2a_n\\ \vdots & \vdots & & \vdots\\ a_na_1 & a_na_2 & \cdots & 1+a_n^2\end{pmatrix}.$$

解 注意到

$$A-I_n=\begin{pmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{pmatrix}\begin{pmatrix}a_1 & a_2 & \cdots & a_n\end{pmatrix}.$$

从而由降价公式可得

$$|A|=\left|I_n+\begin{pmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{pmatrix}\begin{pmatrix}a_1 & a_2 & \cdots & a_n\end{pmatrix}\right|=|I_n|\left|1+\begin{pmatrix}a_1 & a_2 & \cdots & a_n\end{pmatrix}I_n^{-1}\begin{pmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{pmatrix}\right|=1+\sum_{i=1}^n a_i^2.$$

□

例题 2.29 计算矩阵 A 的行列式的值:

$$A=\begin{pmatrix}a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n\\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n\\ \vdots & \vdots & & \vdots\\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n\end{pmatrix}.$$

解 注意到

$$A=\begin{pmatrix}a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n\\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n\\ \vdots & \vdots & & \vdots\\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n\end{pmatrix}=\begin{pmatrix}a_1 & -1\\ a_2 & -1\\ \vdots & \vdots\\ a_n & -1\end{pmatrix}\begin{pmatrix}1 & 1 & \cdots & 1\\ b_1 & b_2 & \cdots & b_n\end{pmatrix}.$$

当 $n>2$ 时, 由 Cauchy-Binet 公式可知 $|A|=0$. 当 $n=2$ 时,$|A|=a_1b_1+a_2b_2-a_1b_2-b_1a_2$. 当 $n=1$ 时,$|A|=a_1-b_1$.

□

例题 2.30 求下列矩阵的行列式的值:

$$A=\begin{vmatrix}0 & 2 & 3 & \cdots & n\\ 1 & 0 & 3 & \cdots & n\\ 1 & 2 & 0 & \cdots & n\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & 2 & 3 & \cdots & 0\end{vmatrix}.$$

解 将 A 化为

$$A=\begin{pmatrix}-1 & 0 & \cdots & 0\\ 0 & -2 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & -n\end{pmatrix}+\begin{pmatrix}1\\ 1\\ \vdots\\ 1\end{pmatrix}(1,2,\cdots,n),$$

利用降阶公式容易求得 $|A|=(-1)^n n!(1-n)$. □

命题 2.37

设 A, B 是 n 阶矩阵, 求证:

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix} = |A+B||A-B|.$$

证明 将分块矩阵的第二行加到第一行上, 再将第二列减去第一列, 可得

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix} \to \begin{pmatrix} A+B & A+B \\ B & A \end{pmatrix} \to \begin{pmatrix} A+B & O \\ B & A-B \end{pmatrix}.$$

第三类分块初等变换不改变行列式的值, 因此可得

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix} = \begin{vmatrix} A+B & O \\ B & A-B \end{vmatrix} = |A+B||A-B|.$$

□

例题 2.31 计算:

$$|A| = \begin{vmatrix} x & y & z & w \\ y & x & w & z \\ z & w & x & y \\ w & z & y & x \end{vmatrix}.$$

解 解法一: 令

$$B = \begin{pmatrix} x & y \\ y & x \end{pmatrix}, C = \begin{pmatrix} z & w \\ w & z \end{pmatrix},$$

则 $|A| = \begin{vmatrix} B & C \\ C & B \end{vmatrix}$. 由命题 2.40可得

$$\begin{aligned} |A| = |B+C||B-C| &= \begin{vmatrix} x+z & y+w \\ y+w & x+z \end{vmatrix} \begin{vmatrix} x-z & y-w \\ y-w & x-z \end{vmatrix} \\ &= (x+y+z+w)(x+z-y-w)(x+y-z-w)(x+w-y-z). \end{aligned}$$

解法二 (求根法):

□

命题 2.38

设 A, B 是 n 阶复矩阵, 求证:

$$\begin{vmatrix} A & -B \\ B & A \end{vmatrix} = |A+\mathrm{i}B||A-\mathrm{i}B|.$$

证明 将分块矩阵的第二行乘以 i 加到第一行上, 再将第一列乘以 $-\mathrm{i}$ 加到第二列上, 可得

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix} \to \begin{pmatrix} A+\mathrm{i}B & \mathrm{i}A-B \\ B & A \end{pmatrix} \to \begin{pmatrix} A+\mathrm{i}B & O \\ B & A-\mathrm{i}B \end{pmatrix}.$$

第三类分块初等变换不改变行列式的值, 因此可得

$$\begin{vmatrix} A & -B \\ B & A \end{vmatrix} = \begin{vmatrix} A+\mathrm{i}B & O \\ B & A-\mathrm{i}B \end{vmatrix} = |A+\mathrm{i}B||A-\mathrm{i}B|.$$

□

例题 2.32 设 A,B,C,D 都是 n 阶矩阵, 求证:

$$|M|=\begin{vmatrix}A&B&C&D\\B&A&D&C\\C&D&A&B\\D&C&B&A\end{vmatrix}=|A+B+C+D||A+B-C-D||A-B+C-D||A-B-C+D|.$$

解 反复利用命题 2.40的结论可得

$$|M|=\left|\begin{pmatrix}A&B\\B&A\end{pmatrix}+\begin{pmatrix}C&D\\D&C\end{pmatrix}\right|\cdot\left|\begin{pmatrix}A&B\\B&A\end{pmatrix}-\begin{pmatrix}C&D\\D&C\end{pmatrix}\right|=\begin{vmatrix}A+C&B+D\\B+D&A+C\end{vmatrix}\cdot\begin{vmatrix}A-C&B-D\\B-D&A-C\end{vmatrix}$$
$$=|A+B+C+D||A-B+C-D||A+B-C-D||A-B-C+D|.$$

□

例题 2.33 设 A,B 是 n 阶矩阵且 $AB=BA$, 求证:

$$\begin{vmatrix}A&-B\\B&A\end{vmatrix}=|A^2+B^2|.$$

证明 由命题 2.41的结论可得

$$\begin{vmatrix}A&-B\\B&A\end{vmatrix}=|A+\mathrm{i}B|\cdot|A-\mathrm{i}B|=|(A+\mathrm{i}B)(A-\mathrm{i}B)|=|A^2+B^2-\mathrm{i}(AB-BA)|=|A^2+B^2|.$$

□

例题 2.34 设 A,B 是 n 阶实矩阵, 求证:

$$\begin{vmatrix}A&-B\\B&A\end{vmatrix}\geq 0.$$

证明 注意到 A,B 都是实矩阵, 故 $\overline{|A+\mathrm{i}B|}=|\overline{A+\mathrm{i}B}|=|A-\mathrm{i}B|$, 再由命题 2.41的结论可得

$$\begin{vmatrix}A&-B\\B&A\end{vmatrix}=|A+\mathrm{i}B|\cdot|A-\mathrm{i}B|=|A+\mathrm{i}B|\cdot\overline{|A+\mathrm{i}B|}\geq 0.$$

□

2.8 分块矩阵的初等变换与降价公式 (打洞原理)

命题 2.39 (打洞原理)

(1) 设

$$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{B}\\\boldsymbol{C}&\boldsymbol{D}\end{pmatrix}_{(n+m)\times(n+m)}$$

是一个方阵, 并且 $\boldsymbol{A}$ 为 n 阶可逆子方阵, 那么

$$|\boldsymbol{M}|=|\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}|.$$

(2) 设

$$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{B}\\\boldsymbol{C}&\boldsymbol{D}\end{pmatrix}_{(n+m)\times(n+m)}$$

是一个方阵, 并且 $\boldsymbol{D}$ 为 n 阶可逆子方阵, 那么

$$|\boldsymbol{M}|=|\boldsymbol{D}|\cdot|\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}|.$$

♠

笔记 打洞原理是一个重要结论, 必须要熟练掌握. 但是在实际解题中我们一般不会直接套用打洞原理的结论, 而

是利用分块矩阵的初等变换书写过程.

记忆打洞原理公式的小技巧: 先记住一个模板 $|\square| = |\square|\left|\square - \square\square^{-1}\square\right|$, 然后从左往右填入子矩阵 (每个子矩阵只能填一次), 第一个 □ 填相应的可逆子矩阵, 再从主对角线上另外一个子矩阵开始, 按顺时针顺序将子矩阵填入 □ 即可.

证明 (核心想法: 利用分块矩阵的初等变换消去 $\boldsymbol{B}$ 或 $\boldsymbol{C}$)

(1) 根据分块矩阵的初等变换, 对 $\boldsymbol{M}$ 的第一行左乘 $(-\boldsymbol{C}\boldsymbol{A}^{-1})$ 再加到第二行得到

$$\begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} \end{pmatrix}.$$

然后两边同时取行列式就得到

$$|\boldsymbol{M}| = \begin{vmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I}_m \end{vmatrix}\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = \left|\begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}\right| = \begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} \end{vmatrix} = |\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}|.$$

(2) 根据分块矩阵的初等变换, 对 $\boldsymbol{M}$ 的第二行左乘 $(-\boldsymbol{B}\boldsymbol{D}^{-1})$ 再加到第一行得到

$$\begin{pmatrix} \boldsymbol{I}_n & -\boldsymbol{B}\boldsymbol{D}^{-1} \\ \boldsymbol{O} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}.$$

然后两边同时取行列式就得到

$$|\boldsymbol{M}| = \begin{vmatrix} \boldsymbol{I}_n & -\boldsymbol{B}\boldsymbol{D}^{-1} \\ \boldsymbol{O} & \boldsymbol{I}_m \end{vmatrix}\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = \left|\begin{pmatrix} \boldsymbol{I}_n & -\boldsymbol{B}\boldsymbol{D}^{-1} \\ \boldsymbol{O} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}\right| = \begin{vmatrix} \boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = |\boldsymbol{D}|\cdot|\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}|.$$

□

推论 2.4 (打洞原理推论)

设 $\boldsymbol{A}$ 是 $n\times m$ 矩阵,$\boldsymbol{B}$ 是 $m\times n$ 矩阵, 则

$$\lambda^m|\lambda\boldsymbol{I}_n-\boldsymbol{A}\boldsymbol{B}| = \lambda^n|\lambda\boldsymbol{I}_m-\boldsymbol{B}\boldsymbol{A}|.$$

♡

笔记 这个推论能将原本复杂的矩阵 AB 通过交换顺序变成相对简单的矩阵 BA. 例如:例题 2.35.

注 这是由打洞原理得到的一个重要结论, 也需要熟练掌握. 同样地, 在实际解题中如果不能直接套用打洞原理推论的结论, 就需要利用分块矩阵的初等变换书写过程.

证明 当 $\lambda=0$ 时, 结论显然成立.

当 $\lambda\neq 0$ 时, 根据分块矩阵的初等变换可知

$$\begin{pmatrix} \boldsymbol{I}_n & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix} = \begin{pmatrix} \lambda\boldsymbol{I}_n-\boldsymbol{A}\boldsymbol{B} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix},$$

$$\begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\frac{1}{\lambda}\boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix} = \begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m-\frac{1}{\lambda}\boldsymbol{B}\boldsymbol{A} \end{pmatrix}.$$

再对上式两边分别取行列式得到

$$\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix} = \begin{vmatrix} \boldsymbol{I}_n & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m \end{vmatrix}\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix} = \left|\begin{pmatrix} \boldsymbol{I}_n & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}\right| = \begin{vmatrix} \lambda\boldsymbol{I}_n-\boldsymbol{A}\boldsymbol{B} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix} = |\lambda\boldsymbol{I}_n-\boldsymbol{A}\boldsymbol{B}|.$$

$$\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix} = \begin{vmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\frac{1}{\lambda}\boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix}\begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{vmatrix} = \left|\begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ -\frac{1}{\lambda}\boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_m \end{pmatrix}\right| = \begin{vmatrix} \lambda\boldsymbol{I}_n & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I}_m-\frac{1}{\lambda}\boldsymbol{B}\boldsymbol{A} \end{vmatrix} = \lambda^n\left|\boldsymbol{I}_m-\frac{1}{\lambda}\boldsymbol{B}\boldsymbol{A}\right| = \lambda^{n-m}|\lambda\boldsymbol{I}_m-\boldsymbol{B}\boldsymbol{A}|.$$

于是 $\begin{vmatrix} \lambda I_n & A \\ B & I_m \end{vmatrix} = |\lambda I_n - AB| = \lambda^{n-m} |\lambda I_m - BA|$.. 故 $\lambda^m |\lambda I_n - AB| = \lambda^n |\lambda I_m - BA|$. □

例题 2.35 求下列矩阵的行列式的值:

$$A = \begin{pmatrix} a_1^2 & a_1a_2+1 & \cdots & a_1a_n+1 \\ a_2a_1+1 & a_2^2 & \cdots & a_2a_n+1 \\ \vdots & \vdots & & \vdots \\ a_na_1+1 & a_na_2+1 & \cdots & a_n^2 \end{pmatrix}.$$

解 令 $\Lambda = \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}$, 则由降价公式 (打洞原理) 我们有

$$A = -I_n + \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix} I_2^{-1} \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} = (-1)^n |I_2| \left| I_n - \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix} I_2^{-1} \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \right|$$

$$= (-1)^n \begin{vmatrix} I_2 & \Lambda' \\ \Lambda & I_n \end{vmatrix} = (-1)^n |I_n| \left| I_2 - \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} I_n^{-1} \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix} \right|$$

$$= (-1)^n \left| I_2 - \begin{pmatrix} \sum\limits_{i=1}^n a_i^2 & \sum\limits_{i=1}^n a_i \\ \sum\limits_{i=1}^n a_i & n \end{pmatrix} \right| = (-1)^n \left[(1-n)\left(1 - \sum_{i=1}^n a_i^2\right) - \left(\sum_{i=1}^n a_i\right)^2 \right].$$

□

例题 2.36 计算矩阵 A 的行列式的值:

$$A = \begin{pmatrix} 1+a_1^2 & a_1a_2 & \cdots & a_1a_n \\ a_2a_1 & 1+a_2^2 & \cdots & a_2a_n \\ \vdots & \vdots & & \vdots \\ a_na_1 & a_na_2 & \cdots & 1+a_n^2 \end{pmatrix}.$$

解 注意到

$$A - I_n = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}.$$

从而由降价公式可得

$$|A| = \left| I_n + \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \right| = |I_n| \left| 1 + \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} I_n^{-1} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \right| = 1 + \sum_{i=1}^n a_i^2.$$

□

例题 2.37 计算矩阵 A 的行列式的值:

$$A=\begin{pmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \vdots & \vdots & & \vdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{pmatrix}.$$

解 注意到

$$A=\begin{pmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \vdots & \vdots & & \vdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{pmatrix}=\begin{pmatrix} a_1 & -1 \\ a_2 & -1 \\ \vdots & \vdots \\ a_n & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}.$$

当 $n>2$ 时, 由 Cauchy-Binet 公式可知 $|A|=0$. 当 $n=2$ 时,$|A|=a_1b_1+a_2b_2-a_1b_2-b_1a_2$. 当 $n=1$ 时,$|A|=a_1-b_1$.

□

例题 2.38 求下列矩阵的行列式的值:

$$A=\begin{vmatrix} 0 & 2 & 3 & \cdots & n \\ 1 & 0 & 3 & \cdots & n \\ 1 & 2 & 0 & \cdots & n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 2 & 3 & \cdots & 0 \end{vmatrix}.$$

解 将 A 化为

$$A=\begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & -2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -n \end{pmatrix}+\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}(1,2,\cdots,n),$$

利用降阶公式容易求得 $|A|=(-1)^n n!(1-n)$. □

命题 2.40

设 A,B 是 n 阶矩阵, 求证:

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix}=|A+B||A-B|.$$

证明 将分块矩阵的第二行加到第一行上, 再将第二列减去第一列, 可得

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix}\to\begin{pmatrix} A+B & A+B \\ B & A \end{pmatrix}\to\begin{pmatrix} A+B & O \\ B & A-B \end{pmatrix}.$$

第三类分块初等变换不改变行列式的值, 因此可得

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix}=\begin{vmatrix} A+B & O \\ B & A-B \end{vmatrix}=|A+B||A-B|.$$

□

例题 2.39 计算:

$$|A|=\begin{vmatrix} x & y & z & w \\ y & x & w & z \\ z & w & x & y \\ w & z & y & x \end{vmatrix}.$$

解 解法一: 令

$$B=\begin{pmatrix}x & y\\ y & x\end{pmatrix}, C=\begin{pmatrix}z & w\\ w & z\end{pmatrix},$$

则 $|A|=\begin{vmatrix}B & C\\ C & B\end{vmatrix}$. 由命题 2.40可得

$$\begin{aligned}|A|=|B+C||B-C|&=\begin{vmatrix}x+z & y+w\\ y+w & x+z\end{vmatrix}\begin{vmatrix}x-z & y-w\\ y-w & x-z\end{vmatrix}\\&=(x+y+z+w)(x+z-y-w)(x+y-z-w)(x+w-y-z).\end{aligned}$$

解法二 (求根法):

□

命题 2.41

设 A,B 是 n 阶复矩阵, 求证:

$$\begin{vmatrix}A & -B\\ B & A\end{vmatrix}=|A+\mathrm{i}B||A-\mathrm{i}B|.$$

证明 将分块矩阵的第二行乘以 i 加到第一行上, 再将第一列乘以 $-\mathrm{i}$ 加到第二列上, 可得

$$\begin{pmatrix}A & -B\\ B & A\end{pmatrix}\to\begin{pmatrix}A+\mathrm{i}B & \mathrm{i}A-B\\ B & A\end{pmatrix}\to\begin{pmatrix}A+\mathrm{i}B & O\\ B & A-\mathrm{i}B\end{pmatrix}.$$

第三类分块初等变换不改变行列式的值, 因此可得

$$\begin{vmatrix}A & -B\\ B & A\end{vmatrix}=\begin{vmatrix}A+\mathrm{i}B & O\\ B & A-\mathrm{i}B\end{vmatrix}=|A+\mathrm{i}B||A-\mathrm{i}B|.$$

□

例题 2.40 设 A,B,C,D 都是 n 阶矩阵, 求证:

$$|M|=\begin{vmatrix}A & B & C & D\\ B & A & D & C\\ C & D & A & B\\ D & C & B & A\end{vmatrix}=|A+B+C+D||A+B-C-D||A-B+C-D||A-B-C+D|.$$

解 反复利用命题 2.40的结论可得

$$\begin{aligned}|M|&=\left|\begin{pmatrix}A & B\\ B & A\end{pmatrix}+\begin{pmatrix}C & D\\ D & C\end{pmatrix}\right|\cdot\left|\begin{pmatrix}A & B\\ B & A\end{pmatrix}-\begin{pmatrix}C & D\\ D & C\end{pmatrix}\right|=\begin{vmatrix}A+C & B+D\\ B+D & A+C\end{vmatrix}\cdot\begin{vmatrix}A-C & B-D\\ B-D & A-C\end{vmatrix}\\&=|A+B+C+D||A-B+C-D||A+B-C-D||A-B-C+D|.\end{aligned}$$

□

例题 2.41 设 A,B 是 n 阶矩阵且 $AB=BA$, 求证:

$$\begin{vmatrix}A & -B\\ B & A\end{vmatrix}=|A^2+B^2|.$$

证明 由命题 2.41的结论可得

$$\begin{vmatrix}A & -B\\ B & A\end{vmatrix}=|A+\mathrm{i}B|\cdot|A-\mathrm{i}B|=|(A+\mathrm{i}B)(A-\mathrm{i}B)|=|A^2+B^2-\mathrm{i}(AB-BA)|=|A^2+B^2|.$$

□

例题 2.42 设 A,B 是 n 阶实矩阵, 求证:

$$\begin{vmatrix}A & -B\\ B & A\end{vmatrix}\geq 0.$$

证明 注意到 A,B 都是实矩阵, 故 $\overline{|A+\mathrm{i}B|}=|\overline{A+\mathrm{i}B}|=|A-\mathrm{i}B|$, 再由命题 2.41的结论可得

$$\begin{vmatrix} A & -B \\ B & A \end{vmatrix} = |A+\mathrm{i}B|\cdot|A-\mathrm{i}B| = |A+\mathrm{i}B|\cdot\overline{|A+\mathrm{i}B|} \geq 0.$$

□

2.9 摄动法

摄动法的原理

(1) 证明矩阵问题对非异阵成立.

(2) 对任意的 n 阶矩阵 A, 由上例可知, 存在一列有理数 $t_k\to 0$, 使得 t_kI_n+A 都是非异阵. 验证 t_kI_n+A 仍满足矩阵问题的条件, 从而该问题对 t_kI_n+A 成立.

(3) 若矩阵问题关于 t_k 连续, 则可取极限令 $t_k\to 0$, 从而得到该问题对一般的矩阵 A 也成立.

注

1. 矩阵问题对非异阵成立以及矩阵问题关于 t_k 连续, 这两个要求缺一不可, 否则将不能使用摄动法进行证明.
2. 验证摄动矩阵仍然满足矩阵问题的条件是必要的. 例如, 若矩阵问题中有 $AB=-BA$ 这一条, 但 $(t_kI_n+A)B\neq -B(t_kI_n+A)$, 因此便不能使用摄动法.
3. 根据实际问题的需要, 也可以使用其他非异阵来替代 I_n 对 A 进行摄动.

笔记 关于伴随矩阵的问题中经常会使用摄动法.

命题 2.42

设 A 是一个 n 阶方阵, 求证: 存在一个正数 a, 使得对任意的 $0<t<a$, 矩阵 tI_n+A 都是非异阵.

笔记 这个命题告诉我们对任意的 n 阶矩阵 A, 经过微小的一维摄动之后,tI_n+A 总能成为一个非异阵.

证明 通过简单的计算可得

$$|tI_n+A|=t^n+a_1t^{n-1}+\cdots+a_{n-1}t+a_n,$$

这是一个关于未定元 t 的 n 次多项式. 由多项式根的有限性可知上述多项式至多只有 n 个不同的根. 若上述多项式的根都是零, 则不妨取 $a=1$; 若上述多项式有非零根, 则令 a 为 $|tI_n+A|$ 所有非零根的模长的最小值. 因此对任意的 $0<t_0<a$, t_0 都不是 $|tI_n+A|$ 的根, 即 $|t_0I_n+A|\neq 0$, 从而 t_0I_n+A 是非异阵. □

例题 2.43 设 A,B,C,D 是 n 阶矩阵且 $AC=CA$, 求证:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD-CB|.$$

笔记 本题也给出了例题 2.41的摄动法证明.

证明 若 A 为非异阵, 则由降阶公式, 再结合条件 $AC=CA$ 可得

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A|\left|D-CA^{-1}B\right| = \left|AD-ACA^{-1}B\right| = |AD-CB|.$$

对于一般的方阵 A, 由命题 2.11 可知, 存在一列有理数 $t_k\to 0$, 使得 t_kI_n+A 是非异阵, 并且条件 $(t_kI_n+A)C=C(t_kI_n+A)$ 仍然成立. 于是

$$\begin{vmatrix} t_kI_n+A & B \\ C & D \end{vmatrix} = |(t_kI_n+A)D-CB|\,.$$

上式两边都是行列式, 其值都是 t_k 的多项式, 从而都关于 t_k 连续. 上式两边同时令 $t_k\to 0$, 即有 $\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|AD-CB|$ 成立. □

2.10 练习

练习 2.14 设 $A=(a_{ij})$ 为 n 阶方阵, 定义函数 $f(A)=\sum\limits_{i,j=1}^{n}a_{ij}^2$. 设 P 为 n 阶可逆矩阵, 使得对任意的 n 阶方阵 A 成立:$f(PAP^{-1})=f(A)$. 证明: 存在非零常数 c, 使得 $P'P=cI_n$.

证明 由假设知 $f(A)=\mathrm{tr}(AA')$, 因此

$$f(PAP^{-1})=\mathrm{tr}(PAP^{-1}(P')^{-1}A'P')=\mathrm{tr}((P'P)A(P'P)^{-1}A')=\mathrm{tr}(AA').$$

以下设 $P'P=(c_{ij}),(P'P)^{-1}=(d_{ij})$. 注意 $P'P$ 是对称矩阵, 后面要用到. 令 $A=E_{ij}$, 其中 $1\le i,j\le n$. 并将其代入 $(P'P)A(P'P)^{-1}A'$ 可得

$$
\begin{aligned}
(P'P)A(P'P)^{-1}A' &= (P'P)E_{ij}(P'P)^{-1}E_{ji}\\
&= \begin{array}{c} j \\ \begin{pmatrix} c_{1i}\\ c_{2i}\\ \vdots\\ c_{ii}\\ \vdots\\ c_{ni}\end{pmatrix}\end{array}
\begin{array}{c} i \\ \begin{pmatrix} d_{1j}\\ d_{2j}\\ \vdots\\ d_{jj}\\ \vdots\\ d_{nj}\end{pmatrix}\end{array}
= \begin{array}{c} i \\ \begin{pmatrix} c_{1i}d_{jj}\\ c_{2i}d_{jj}\\ \vdots\\ c_{ii}d_{jj}\\ \vdots\\ c_{ni}d_{jj}\end{pmatrix}\end{array}
\end{aligned}
$$

于是 $\mathrm{tr}\left(\left(P'P\right)A\left(P'P\right)^{-1}A'\right)=c_{ii}d_{jj}$. 而 $\mathrm{tr}\left(AA\prime\right)=\mathrm{tr}\left(E_{ij}E_{ji}\right)=\mathrm{tr}\left(E_{ii}\right)=1$. 则由 $\mathrm{tr}((P'P)A(P'P)^{-1}A')=\mathrm{tr}(AA')$ 可知

$$c_{ii}d_{jj}=1. \tag{2.2}$$

再令 $A=E_{ij}+E_{kl}$, 其中 $1\le i,j,k,l\le n$. 不妨设 $k\ge i, l\ge j$, 将其代入 $(P'P)A(P'P)^{-1}A'$ 可得

$$(P'P)A(P'P)^{-1}A'=(P'P)(E_{ij}+E_{kl})(P'P)^{-1}(E_{ji}+E_{lk})$$

$$
=\left[\begin{array}{c} j \\ \begin{pmatrix} c_{1i}\\ c_{2i}\\ \vdots\\ c_{ii}\\ \vdots\\ c_{ki}\\ \vdots\\ c_{ni}\end{pmatrix}\end{array}
+
\begin{array}{c} l \\ \begin{pmatrix} c_{1k}\\ c_{2k}\\ \vdots\\ c_{ik}\\ \vdots\\ c_{kk}\\ \vdots\\ c_{nk}\end{pmatrix}\end{array}\right]
\left[\begin{array}{c} i \\ \begin{pmatrix} d_{1j}\\ d_{2j}\\ \vdots\\ d_{jj}\\ \vdots\\ d_{lj}\\ \vdots\\ d_{nj}\end{pmatrix}\end{array}
+
\begin{array}{c} k \\ \begin{pmatrix} d_{1l}\\ d_{2l}\\ \vdots\\ d_{jl}\\ \vdots\\ d_{ll}\\ \vdots\\ d_{nl}\end{pmatrix}\end{array}\right]
$$

$$
= \begin{array}{c} \begin{array}{ccc} \quad j \quad & & \quad l \quad \end{array} \\ \begin{pmatrix} c_{1i} & \cdots & c_{1k} \\ c_{2i} & \cdots & c_{2k} \\ \vdots & & \vdots \\ c_{ii} & \cdots & c_{ik} \\ \vdots & & \vdots \\ c_{ki} & \cdots & c_{kk} \\ \vdots & & \vdots \\ c_{ni} & \cdots & c_{nk} \end{pmatrix} \end{array} \begin{array}{c} \begin{array}{ccc} \quad i \quad & & \quad k \quad \end{array} \\ \begin{pmatrix} d_{1j} & \cdots & d_{1l} \\ d_{2j} & \cdots & d_{2l} \\ \vdots & & \vdots \\ d_{jj} & \cdots & d_{jl} \\ \vdots & & \vdots \\ d_{lj} & \cdots & d_{ll} \\ \vdots & & \vdots \\ d_{nj} & \cdots & d_{nl} \end{pmatrix} \end{array} = \begin{array}{c} \begin{array}{ccc} \qquad\qquad i \qquad\qquad & & \qquad\qquad k \qquad\qquad \end{array} \\ \begin{pmatrix} c_{1i}d_{jj} + c_{1k}d_{lj} & \cdots & c_{1i}d_{jl} + c_{1k}d_{ll} \\ c_{2i}d_{jj} + c_{2k}d_{lj} & \cdots & c_{2i}d_{jl} + c_{2k}d_{ll} \\ \vdots & & \vdots \\ c_{ii}d_{jj} + c_{ik}d_{lj} & \cdots & c_{ii}d_{jl} + c_{ik}d_{ll} \\ \vdots & & \vdots \\ c_{ki}d_{jj} + c_{kk}d_{lj} & \cdots & c_{ki}d_{jl} + c_{kk}d_{ll} \\ \vdots & & \vdots \\ c_{ni}d_{jj} + c_{nk}d_{lj} & \cdots & c_{ni}d_{jl} + c_{nk}d_{ll} \end{pmatrix} \end{array}
$$

从而 $\mathrm{tr}\left(\left(P'P\right)A\left(P'P\right)^{-1}A'\right) = c_{ii}d_{jj} + c_{kk}d_{ll} + c_{ki}d_{jl} + c_{ik}d_{lj}$. 而

$$
\mathrm{tr}\left(AA'\right) = \mathrm{tr}\left(\left(E_{ij} + E_{kl}\right)\left(E_{ji} + E_{lk}\right)\right) = \mathrm{tr}\left(E_{ij}E_{ji} + E_{ij}E_{lk} + E_{kl}E_{ji} + E_{kl}E_{lk}\right) = 2 + 2\delta_{ik}\delta_{jl}.
$$

于是由 $\mathrm{tr}((P'P)A(P'P)^{-1}A') = \mathrm{tr}(AA')$ 可知

$$
c_{ii}d_{jj} + c_{kk}d_{ll} + c_{ki}d_{jl} + c_{ik}d_{lj} = 2 + 2\delta_{ik}\delta_{jl}, \tag{2.3}
$$

其中 δ_{ik} 是 Kronecker 符号. 由上述(2.2)(2.3)两式可得

$$
c_{ki}d_{jl} + c_{ik}d_{lj} = 2\delta_{ik}\delta_{jl}.
$$

在上式中令 $j = l, i \neq k$, 注意到 $d_{jj} \neq 0$, 故有 $c_{ik} + c_{ki} = 0$, 又因为 $P'P$ 是对称矩阵, 所以 $c_{ik} = c_{ki}$. 故 $c_{ik} = 0, \forall i \neq k$. 于是 $P'P$ 是一个对角矩阵, 从而由(2.2)式可得 $d_{jj} = c_{jj}^{-1}$, 由此可得 $c_{ii} = c_{jj}, \forall i, j$. 因此 $P'P = cI_n$, 其中 $c = c_{11} \neq 0$.

□

第三章 线性空间与线性方程组

3.1 向量的线性关系

定理 3.1

设 $\alpha_1,\alpha_2,\cdots,\alpha_m,\beta$ 是线性空间 V 中的向量.

(1) 若 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关, 则任意一组包含这组向量的向量组必线性相关; 若 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性无关, 则从这组向量中任意取出一组向量必线性无关.

(2) 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关的充要条件是其中至少有一个向量可以表示为其余向量的线性组合.

(3) 若 β 可表示为 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的线性组合, 即

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

则表示唯一的充要条件是向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性无关.

证明

□

定理 3.2

(1) 设 A,B 是两组向量,A 含有 r 个向量,B 含有 s 个向量, 且 A 中每个向量均可用 B 中向量线性表示. 若 A 中向量线性无关, 则 $r \le s$.

(2) 设 A,B 是两组向量,A 含有 r 个向量,B 含有 s 个向量, 且 A 中每个向量均可用 B 中向量线性表示. 若 $r > s$, 则 A 中向量线性相关.

证明 (2) 是 (1) 的逆否命题, 因此我们只证明 (1). □

命题 3.1

设 A 是 $m\times n$ 阶梯形矩阵, 证明:A 的秩等于其非零行的个数, 且阶梯点所在的列向量是 A 的列向量的极大无关组.

证明 设

$$A = \begin{pmatrix} 0 & \cdots & a_{1k_1} & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \cdots & a_{2k_2} & \cdots & \cdots & \cdots \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & a_{rk_r} & \cdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots \\ \vdots & & \vdots & & \vdots & & \vdots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots \end{pmatrix}.$$

其中 $a_{1k_1},a_{2k_2},\cdots,a_{rk_r}$ 是 A 的阶梯点. 设 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是 A 的前 r 行, 我们先证明它们线性无关. 设

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r = \mathbf{0}$$

其中 $c_1,c_2,\cdots,c_r$ 是常数. 上式是关于 n 维行向量的等式, 先考察行向量的第 k_1 分量, 可得 $c_1a_{1k_1} = 0$. 因为 $a_{1k_1} \ne 0$, 故 $c_1 = 0$; 再依次考察行向量的第 $k_2,\cdots,k_r$ 分量, 最后可得 $c_1 = c_2 = \cdots = c_r = 0$. 因此 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关, 从而 A 的秩等于 r, 即其非零行的个数.

再将 r 个阶梯点所在的列向量取出, 拼成一个新的矩阵:

$$B=\begin{pmatrix} a_{1k_1} & \cdots & \cdots & \cdots \\ 0 & a_{2k_2} & \cdots & \cdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{rk_r} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

采用相同的方法可证明矩阵 B 的前 r 行线性无关, 因此 $\mathrm{r}(B)=r$, 从而阶梯点所在的列向量组的秩也等于 r. 又因为 $\mathrm{r}(A)=r$, 故它们是 A 的列向量的极大无关组. □

命题 3.2

设 $A=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 是一个 $m\times n$ 矩阵,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是列向量. P 是一个 m 阶可逆矩阵,$B=PA=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$, 其中 $\boldsymbol{\beta}_j=P\boldsymbol{\alpha}_j(1\le j\le n)$. 证明: 若 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 是 A 的列向量的极大无关组, 则 $\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_r}$ 是 B 的列向量的极大无关组.

证明 先证明向量组 $\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_r}$ 线性无关. 设

$$c_1\boldsymbol{\beta}_{i_1}+c_2\boldsymbol{\beta}_{i_2}+\cdots+c_r\boldsymbol{\beta}_{i_r}=\mathbf{0},$$

即

$$c_1P\boldsymbol{\alpha}_{i_1}+c_2P\boldsymbol{\alpha}_{i_2}+\cdots+c_rP\boldsymbol{\alpha}_{i_r}=\mathbf{0}.$$

已知 P 是可逆矩阵, 因此

$$c_1\boldsymbol{\alpha}_{i_1}+c_2\boldsymbol{\alpha}_{i_2}+\cdots+c_r\boldsymbol{\alpha}_{i_r}=\mathbf{0}.$$

而向量组 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性无关, 故 $c_1=c_2=\cdots=c_r=0$, 这证明了向量组 $\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_r}$ 线性无关. 要证这是 B 的列向量的极大无关组, 只需证明 B 的任意一个列向量都是这些向量的线性组合即可. 设 $\boldsymbol{\beta}_j$ 是 B 的任意一个列向量, 则 $\boldsymbol{\beta}_j=P\boldsymbol{\alpha}_j$. 因为 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 是 A 的列向量的极大无关组, 故 $\boldsymbol{\alpha}_j$ 可用 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性表示. 不妨设

$$\boldsymbol{\alpha}_j=b_1\boldsymbol{\alpha}_{i_1}+b_2\boldsymbol{\alpha}_{i_2}+\cdots+b_r\boldsymbol{\alpha}_{i_r},$$

则

$$P\boldsymbol{\alpha}_j=b_1P\boldsymbol{\alpha}_{i_1}+b_2P\boldsymbol{\alpha}_{i_2}+\cdots+b_rP\boldsymbol{\alpha}_{i_r},$$

即

$$\boldsymbol{\beta}_j=b_1\boldsymbol{\beta}_{i_1}+b_2\boldsymbol{\beta}_{i_2}+\cdots+b_r\boldsymbol{\beta}_{i_r}.$$

□

推论 3.1

设 n 维向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关,A 为 n 阶可逆矩阵, 求证:$A\boldsymbol{\alpha}_1,A\boldsymbol{\alpha}_2,\cdots,A\boldsymbol{\alpha}_m$ 线性无关.

证明 由命题 3.2即得. □

命题 3.3

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是线性空间 V 中一组线性无关的向量,$\boldsymbol{\beta}$ 是 V 中的向量. 求证: 或者 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}$ 线性无关, 或者 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的线性组合.

证明 若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}$ 线性无关,则结论得证. 若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}$ 线性相关,则存在不全为零的数 $c_1,c_2,\cdots,c_m,d$,

使得

$$c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_m\boldsymbol{\alpha}_m+d\boldsymbol{\beta}=\mathbf{0}.$$

若 $d=0$, 则 $c_1,c_2,\cdots,c_m$ 不全为零且 $c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_m\boldsymbol{\alpha}_m=\mathbf{0}$, 这与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关矛盾. 因此 $d\neq 0$, 从而

$$\boldsymbol{\beta}=-\frac{c_1}{d}\boldsymbol{\alpha}_1-\frac{c_2}{d}\boldsymbol{\alpha}_2-\cdots-\frac{c_m}{d}\boldsymbol{\alpha}_m,$$

即 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的线性组合. □

推论 3.2

若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关且 $\boldsymbol{\beta}\notin L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)$, 则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}$ 线性无关.

笔记 这个推论与上一个命题 3.3等价. 虽然这个等价命题很简单, 但后面经常会用到.

命题 3.4

设向量 $\boldsymbol{\beta}$ 可由向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表示, 但不能由其中任何个数少于 m 的部分向量线性表示, 则这 m 个向量线性无关.

证明 用反证法, 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关, 则至少有一个向量是其余向量的线性组合. 不妨设 $\boldsymbol{\alpha}_m$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1}$ 的线性组合, 则由线性组合的传递性可知,$\boldsymbol{\beta}$ 也是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1}$ 的线性组合, 这与假设矛盾. □

命题 3.5

设线性空间 V 中向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关, 已知有序向量组 $\{\boldsymbol{\beta},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}$ 线性相关, 求证: 最多只有一个 $\boldsymbol{\alpha}_i$ 可以表示为前面向量的线性组合.

证明 用反证法, 设存在 $1\leq i<j\leq r$, 使得

$$\boldsymbol{\alpha}_i=b\boldsymbol{\beta}+a_1\boldsymbol{\alpha}_1+a_2\boldsymbol{\alpha}_2+\cdots+a_{i-1}\boldsymbol{\alpha}_{i-1},$$

$$\boldsymbol{\alpha}_j=d\boldsymbol{\beta}+c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_{j-1}\boldsymbol{\alpha}_{j-1}.$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关, 故 $b\neq 0$. 若 $b=0$, 则 $\boldsymbol{\alpha}_i$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{i-1}$ 的线性组合, 这与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关矛盾. 将第一个等式乘以 $-\dfrac{d}{b}$ 加到第二个等式上, 可得 $\boldsymbol{\alpha}_j$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{j-1}$ 的线性组合, 这与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关矛盾. □

命题 3.6

设 A 是 $n\times m$ 矩阵,B 是 $m\times n$ 矩阵, 满足 $AB=I_n$, 求证:B 的 n 个列向量线性无关.

笔记 实际上, 由 $AB=I_n$ 可知 A,B 互为逆矩阵, 从而 A,B 满秩, 结论得证.

证明 设 $B=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$ 为列分块, 则 $AB=(A\boldsymbol{\beta}_1,A\boldsymbol{\beta}_2,\cdots,A\boldsymbol{\beta}_n)$. 由 $AB=I_n$ 可得 $A\boldsymbol{\beta}_i=\boldsymbol{e}_i(1\leq i\leq n)$, 其中 $\boldsymbol{e}_i$ 是 n 维标准单位列向量. 设

$$c_1\boldsymbol{\beta}_1+c_2\boldsymbol{\beta}_2+\cdots+c_n\boldsymbol{\beta}_n=\mathbf{0},$$

上式两边同时左乘 A, 可得

$$\mathbf{0}=c_1A\boldsymbol{\beta}_1+c_2A\boldsymbol{\beta}_2+\cdots+c_nA\boldsymbol{\beta}_n=c_1\boldsymbol{e}_1+c_2\boldsymbol{e}_2+\cdots+c_n\boldsymbol{e}_n=(c_1,c_2,\cdots,c_n)',$$

因此 $c_1=c_2=\cdots=c_n=0$, 即 B 的 n 个列向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 线性无关. □

命题 3.7 (缩短向量与延伸向量)

1. 设 $\{\boldsymbol{\alpha}_i=(a_{i1},a_{i2},\cdots,a_{in}),1\leq i\leq m\}$ 是一组 n 维行向量,$1\leq j_1<j_2<\cdots<j_t\leq n$ 是给定的 $t(t<n)$ 个指标. 定义 $\widetilde{\boldsymbol{\alpha}}_i=(a_{ij_1},a_{ij_2},\cdots,a_{ij_t})$, 称 $\widetilde{\boldsymbol{\alpha}}_i$ 为 $\boldsymbol{\alpha}_i$ 的 t 维**缩短向量**. 则

(1) 若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关, 则 $\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_m$ 也线性相关;

(2) 设 n 维行向量 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的线性组合, 则 $\widetilde{\alpha}$ 也是 $\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_m$ 的线性组合.

2. 设 $\{\boldsymbol{\alpha}_i=(a_{i1},a_{i2},\cdots,a_{in}),1\le i\le m\}$ 是一组 n 维行向量,$j_1,j_2,\cdots,j_t\ge 1$ 是给定的 $t(t>n)$ 个指标. 定义 $\overline{\alpha}_i=(a_{ij_1},a_{ij_2},\cdots,a_{ij_t})$, 称 $\overline{\alpha}_i$ 为 $\boldsymbol{\alpha}_i$ 的 t 维**延伸向量**. 则

(1) 若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关, 则 $\overline{\alpha}_1,\overline{\alpha}_2,\cdots,\overline{\alpha}_m$ 也线性无关.

笔记 这个命题告诉我们: **线性相关向量组的任意缩短组也是线性相关的, 线性无关向量组的任意延伸组也是线性无关的.**

证明 1. (1) 由假设存在不全为零的数 $c_1,c_2,\cdots,c_m$, 使得

$$\mathbf{0}=c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_m\boldsymbol{\alpha}_m=\left(\sum_{i=1}^m c_ia_{i1},\sum_{i=1}^m c_ia_{i2},\cdots,\sum_{i=1}^m c_ia_{in}\right).$$

在等式两边同时取 t 维缩短向量, 可得

$$\mathbf{0}=\left(\sum_{i=1}^m c_ia_{ij_1},\sum_{i=1}^m c_ia_{ij_2},\cdots,\sum_{i=1}^m c_ia_{ij_t}\right)=c_1\widetilde{\alpha}_1+c_2\widetilde{\alpha}_2+\cdots+c_m\widetilde{\alpha}_m,$$

从而结论成立.

(2) 设 $\boldsymbol{\alpha}=c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_m\boldsymbol{\alpha}_m$, 则

$$\boldsymbol{\alpha}=c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_m\boldsymbol{\alpha}_m=\left(\sum_{i=1}^m c_ia_{i1},\sum_{i=1}^m c_ia_{i2},\cdots,\sum_{i=1}^m c_ia_{in}\right).$$

在等式两边同时取 t 维缩短向量, 可得

$$\widetilde{\alpha}=\left(\sum_{i=1}^m c_ia_{ij_1},\sum_{i=1}^m c_ia_{ij_2},\cdots,\sum_{i=1}^m c_ia_{ij_t}\right)=c_1\widetilde{\alpha}_1+c_2\widetilde{\alpha}_2+\cdots+c_m\widetilde{\alpha}_m,$$

从而结论成立.

2. 这个命题就是 1(1) 的逆否命题, 从而结论成立. □

例题 3.1 设 V 是实数域上连续函数全体构成的实线性空间, 求证下列函数线性无关:

(1) $\sin x,\sin 2x,\cdots,\sin nx$;

(2) $1,\cos x,\cos 2x,\cdots,\cos nx$;

(3) $1,\sin x,\cos x,\sin 2x,\cos 2x,\cdots,\sin nx,\cos nx$.

证明 证法一:根据向量线性无关的基本性质, 我们只要证明 (3) 即可. 对 n 进行归纳, 当 $n=0$ 时, 显然 1 作为一个函数线性无关. 假设命题对小于 n 的自然数成立, 现证明等于 n 的情形. 设

$$a+b_1\sin x+c_1\cos x+b_2\sin 2x+c_2\cos 2x+\cdots+b_n\sin nx+c_n\cos nx=0,$$

其中 a,b_i,c_i 都是实数. 对上式两次求导, 可得

$$-b_1\sin x-c_1\cos x-4b_2\sin 2x-4c_2\cos 2x-\cdots-n^2b_n\sin nx-n^2c_n\cos nx=0,$$

再将第一个式子乘以 n^2 加到第二个式子上, 可得

$$an^2+\sum_{i=1}^{n-1}b_i(n^2-i^2)\sin ix+\sum_{i=1}^{n-1}c_i(n^2-i^2)\cos ix=0,$$

由归纳假设即得 $a=b_1=c_1=\cdots=b_{n-1}=c_{n-1}=0$. 将此结论代入第一个式子可得 $b_n\sin nx+c_n\cos nx=0$. 若 $b_n\neq 0\ (c_n\neq 0)$, 则 $\tan nx=-c_n/b_n\ (\cot nx=-b_n/c_n)$ 为常数 ($\tan nx,\cot nx$ 都不是常函数), 矛盾. 因此,$b_n=c_n=0$.

证法二: 设

$$f(x)=a+b_1\sin x+c_1\cos x+b_2\sin 2x+c_2\cos 2x+\cdots+b_n\sin nx+c_n\cos nx=0,$$

其中 a,b_i,c_i 都是实数. 依次设 $g(x)=1,\sin x,\cos x,\sin 2x,\cos 2x,\cdots,\sin nx,\cos nx$, 并分别计算定积分 $\int_0^{2\pi}f(x)g(x)\mathrm{d}x$, 可得 $a=b_1=c_1=\cdots=b_n=c_n=0$. □

命题 3.8

设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关, 又

$$\begin{cases}\boldsymbol{\beta}_1=a_{11}\boldsymbol{\alpha}_1+a_{12}\boldsymbol{\alpha}_2+\cdots+a_{1r}\boldsymbol{\alpha}_r,\\ \boldsymbol{\beta}_2=a_{21}\boldsymbol{\alpha}_1+a_{22}\boldsymbol{\alpha}_2+\cdots+a_{2r}\boldsymbol{\alpha}_r,\\ \cdots\cdots\cdots\cdots\\ \boldsymbol{\beta}_r=a_{r1}\boldsymbol{\alpha}_1+a_{r2}\boldsymbol{\alpha}_2+\cdots+a_{rr}\boldsymbol{\alpha}_r.\end{cases}$$

则:$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性相关的充要条件是系数矩阵 $A=(a_{ij})_{r\times r}$ 的行列式为零.

证明 记 A 的行向量为 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_r$. 若 $|A|=0$, 则 A 的行向量线性相关, 即存在不全为零的 r 个数 $c_1,c_2,\cdots,c_r$, 使得

$$c_1\boldsymbol{\gamma}_1+c_2\boldsymbol{\gamma}_2+\cdots+c_r\boldsymbol{\gamma}_r=\mathbf{0}.$$

设 $\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)'$, 则经简单计算可得

$$c_1\boldsymbol{\beta}_1+c_2\boldsymbol{\beta}_2+\cdots+c_r\boldsymbol{\beta}_r=c_1\boldsymbol{\gamma}_1\boldsymbol{\alpha}+c_1\boldsymbol{\gamma}_2\boldsymbol{\alpha}+\cdots+c_1\boldsymbol{\gamma}_r\boldsymbol{\alpha}=\left(c_1\boldsymbol{\gamma}_1+c_2\boldsymbol{\gamma}_2+\cdots+c_r\boldsymbol{\gamma}_r\right)\boldsymbol{\alpha}=\mathbf{0}\boldsymbol{\alpha}=\mathbf{0},$$

从而 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性相关.

反之, 若 A 可逆, 如有 $k_1,k_2,\cdots,k_r$, 使得

$$k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_r\boldsymbol{\beta}_r=\mathbf{0},$$

将 $\boldsymbol{\beta}_i$ 代入, 并利用 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 的线性无关性, 可得以 k_i 为未知数的线性方程组:

$$\begin{cases}a_{11}k_1+a_{21}k_2+\cdots+a_{r1}k_r=0,\\ a_{12}k_1+a_{22}k_2+\cdots+a_{r2}k_r=0,\\ \cdots\cdots\cdots\cdots\\ a_{1r}k_1+a_{2r}k_2+\cdots+a_{rr}k_r=0.\end{cases}$$

因为 A 可逆, 所以该方程组只有零解, 从而 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性无关. □

命题 3.9

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是一组线性无关的向量, 向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_k$ 可用 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表示如下:

$$\begin{cases}\boldsymbol{\beta}_1=a_{11}\boldsymbol{\alpha}_1+a_{12}\boldsymbol{\alpha}_2+\cdots+a_{1m}\boldsymbol{\alpha}_m,\\ \boldsymbol{\beta}_2=a_{21}\boldsymbol{\alpha}_1+a_{22}\boldsymbol{\alpha}_2+\cdots+a_{2m}\boldsymbol{\alpha}_m,\\ \cdots\cdots\cdots\cdots\\ \boldsymbol{\beta}_k=a_{k1}\boldsymbol{\alpha}_1+a_{k2}\boldsymbol{\alpha}_2+\cdots+a_{km}\boldsymbol{\alpha}_m.\end{cases}$$

记表示矩阵 $A=(a_{ij})_{k\times m}$, 求证: 向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_k$ 的秩等于 $\mathrm{r}(A)$.

证明 证法一: 设 $\mathrm{r}(A)=r$, 记 A 的 k 个行向量为 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_k$. 不失一般性, 可假设 A 的前 r 个行向量线性无关, 其余向量均可用前 r 个行向量线性表示. 若

$$\boldsymbol{\gamma}_i=c_1\boldsymbol{\gamma}_1+c_2\boldsymbol{\gamma}_2+\cdots+c_r\boldsymbol{\gamma}_r,$$

设 $\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)'$, 则经过简单计算可得

$$\begin{aligned}\boldsymbol{\beta}_i=\boldsymbol{\gamma}_i\boldsymbol{\alpha}&=\left(c_1\boldsymbol{\gamma}_1+c_2\boldsymbol{\gamma}_2+\cdots+c_r\boldsymbol{\gamma}_r\right)\boldsymbol{\alpha}\\ &=c_1\boldsymbol{\gamma}_1\boldsymbol{\alpha}+c_1\boldsymbol{\gamma}_2\boldsymbol{\alpha}+\cdots+c_1\boldsymbol{\gamma}_r\boldsymbol{\alpha}\\ &=c_1\boldsymbol{\beta}_1+c_2\boldsymbol{\beta}_2+\cdots+c_r\boldsymbol{\beta}_r.\end{aligned}$$

另一方面, 若

$$c_1\boldsymbol{\beta}_1 + c_2\boldsymbol{\beta}_2 + \cdots + c_r\boldsymbol{\beta}_r = \mathbf{0},$$

则

$$c_1(a_{11}\boldsymbol{\alpha}_1 + \cdots + a_{1m}\boldsymbol{\alpha}_m) + \cdots + c_r(a_{r1}\boldsymbol{\alpha}_1 + \cdots + a_{rm}\boldsymbol{\alpha}_m) = \mathbf{0},$$

即

$$(c_1a_{11} + \cdots + c_ra_{r1})\boldsymbol{\alpha}_1 + \cdots + (c_1a_{1m} + \cdots + c_ra_{rm})\boldsymbol{\alpha}_m = \mathbf{0}.$$

因为 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 线性无关, 故可得

$$\begin{cases} a_{11}c_1 + a_{21}c_2 + \cdots + a_{r1}c_r = 0, \\ a_{12}c_1 + a_{22}c_2 + \cdots + a_{r2}c_r = 0, \\ \cdots\cdots\cdots\cdots \\ a_{1m}c_1 + a_{2m}c_2 + \cdots + a_{rm}c_r = 0. \end{cases}$$

将上述方程组看成是未知数 c_i 的齐次线性方程组, 其系数矩阵的秩为 r, 未知数个数也是 r, 因此只有唯一一组解, 即零解. 这表明 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$ 是向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_k$ 的极大无关组, 因此向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_k$ 的秩等于 r.

证法二:令 V 是由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 生成的向量空间. 因为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性无关, 故它们组成 V 的一组基,V 的维数等于 m. 注意到 $\boldsymbol{\beta}_i$ 在这组基下的坐标向量为 $(a_{i1}, a_{i2}, \cdots, a_{im})'$, 故由这些列向量组成的矩阵就是 $\boldsymbol{A}'$, 从而

$$\mathrm{r}\left(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_k\right) = \mathrm{r}\left(\boldsymbol{A}'\right) = \mathrm{r}\left(\boldsymbol{A}\right).$$

□

命题 3.10

设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是向量空间 V 中一组向量, 向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_k$ 可用 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表出, 求证: 向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_k$ 的秩小于等于向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的秩.

注 如果将向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 称为原向量组, 将向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_k$ 称为表出向量组, 则这个命题可简述为: "**表出向量组的秩不超过原向量组的秩.**" 从几何上看, 这是一个自然的结果. 因为每个 $\boldsymbol{\beta}_i$ 都属于由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 生成的子空间, 故它们的秩不会超过该子空间的维数.

证明 不失一般性, 可设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 是向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的极大无关组,$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 是向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_k$ 的极大无关组. 因为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 可用 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性表出, 所以 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 也可用 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性表出, 从而由定理 3.2(1)可知 $s \le r$, 结论成立. □

命题 3.11

设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是向量空间 V 中一组向量且其秩等于 r,$\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 是其中 r 个向量. 若下列条件之一成立:

(1) $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 线性无关;

(2) 任一 $\boldsymbol{\alpha}_i$ 均可由 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 线性表示.

则 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 是向量组的极大无关组.

证明 (1) 设 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 线性无关, 又设 $\boldsymbol{\alpha}_{j_1}, \boldsymbol{\alpha}_{j_2}, \cdots, \boldsymbol{\alpha}_{j_r}$ 是向量组的极大无关组. 对任意的 $1 \le i \le m$,$\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}, \boldsymbol{\alpha}_i$ 均可由 $\boldsymbol{\alpha}_{j_1}, \boldsymbol{\alpha}_{j_2}, \cdots, \boldsymbol{\alpha}_{j_r}$ 线性表示, 由定理 3.2(2)可知 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}, \boldsymbol{\alpha}_i$ 必线性相关. 再由命题 3.3可知 $\boldsymbol{\alpha}_i$ 可由 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 线性表示, 从而 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 也是向量组的极大无关组.

(2) 设任一 $\boldsymbol{\alpha}_i$ 均可由 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 线性表示. 不失一般性, 可设 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_s}$ 是向量组 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 的极大无关组. 因此,$\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_s}$ 线性无关. 再由线性组合的传递性可知, 任一 $\boldsymbol{\alpha}_i$ 均可由 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_s}$ 线性表示, 故 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_s}$ 是原向量组的极大无关组, 从而 $s = r$, 即 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 是原向量组的极大无关组. □

命题 3.12

若 $\boldsymbol{A}$ 是对称矩阵或反称矩阵, 并且 $\boldsymbol{A}$ 的第 $i_1,\cdots,i_r$ 行是 $\boldsymbol{A}$ 的行向量的极大无关组, 则它的第 $i_1,\cdots,i_r$ 列也是 $\boldsymbol{A}$ 的列向量的极大无关组.

证明 设 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$ 为矩阵 $\boldsymbol{A}$ 的行分块和列分块, 则由条件可知,$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性无关. 从而存在一组不全为零的数 $k_1,k_2,\cdots,k_r$, 使得

$$k_1\boldsymbol{\alpha}_{i_1}+k_2\boldsymbol{\alpha}_{i_2}+\cdots+k_r\boldsymbol{\alpha}_{i_r}=0.$$

又因为 $\boldsymbol{A}$ 为对称或反称矩阵, 所以 $\boldsymbol{\alpha}_i=\pm\boldsymbol{\beta}_i',i=1,2,\cdots,n$. 代入上式可得

$$k_1\boldsymbol{\beta}_{i_1}'+k_2\boldsymbol{\beta}_{i_2}'+\cdots+k_r\boldsymbol{\beta}_{i_r}'=0.$$

再对上式两边同时取转置可得

$$k_1\boldsymbol{\beta}_{i_1}+k_2\boldsymbol{\beta}_{i_2}+\cdots+k_r\boldsymbol{\beta}_{i_r}=0.$$

故 $\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_r}$ 线性无关. □

命题 3.13 (向量组等价的充要条件)

设有两个向量组 $A=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$ 和 $B=\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$. 求证: 它们等价的充要条件是它们的秩相等且其中一组向量可以用另外一组向量线性表示.

笔记 遇到向量组相关的问题, 一般都会先设出各个向量组的极大无关组.

证明 必要性由向量组等价的定义和命题 3.10即得, 下证充分性. 假设向量组 A 可用向量组 B 线性表示, 且它们的秩都等于 r. 不失一般性, 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是向量组 A 的极大无关组,$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 是向量组 B 的极大无关组. 考虑向量组 $C=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r\}$. 因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 可用 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性表示, 故 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 是向量组 C 的极大无关组, 从而向量组 C 的秩等于 r. 又因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关, 故由命题 3.10(1)可知,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 也是向量组 C 的极大无关组, 从而 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 可用 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示, 于是向量组 B 也可用向量组 A 线性表示. 因此, 向量组 A 与向量组 B 等价. □

命题 3.14

设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_p\}\subseteq\mathbb{K}^m$ 是 p 个线性无关的 m 维列向量, $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_q\}\subseteq\mathbb{K}^n$ 是 q 个线性无关的 n 维列向量. 求证: $\{\boldsymbol{\alpha}_i\cdot\boldsymbol{\beta}_j'\mid 1\le i\le p,1\le j\le q\}$ 是 pq 个线性无关的 $m\times n$ 矩阵.

证明 设 $c_{ij}\in\mathbb{K}$, 使得 $\sum\limits_{i=1}^{p}\sum\limits_{j=1}^{q}c_{ij}\boldsymbol{\alpha}_i\cdot\boldsymbol{\beta}_j'=\boldsymbol{O}$, 则有

$$\sum_{i=1}^{p}\boldsymbol{\alpha}_i\cdot\left(\sum_{j=1}^{q}c_{ij}\boldsymbol{\beta}_j'\right)=\boldsymbol{O}. \tag{3.1}$$

设 $\sum\limits_{j=1}^{q}c_{ij}\boldsymbol{\beta}_j'=(a_{i1},a_{i2},\cdots,a_{in})(1\le i\le p)$, 则(3.1)式可化为

$$\sum_{i=1}^{p}\boldsymbol{\alpha}_i(a_{i1},a_{i2},\cdots,a_{in})=\sum_{i=1}^{p}(a_{i1}\boldsymbol{\alpha}_i,a_{i2}\boldsymbol{\alpha}_i,\cdots,a_{in}\boldsymbol{\alpha}_i)=\left(\sum_{i=1}^{p}a_{i1}\boldsymbol{\alpha}_i,\sum_{i=1}^{p}a_{i2}\boldsymbol{\alpha}_i,\cdots,\sum_{i=1}^{p}a_{in}\boldsymbol{\alpha}_i\right)=\boldsymbol{O}. \tag{3.2}$$

从而比较(3.2)式两边矩阵的第 k 列有 $\sum\limits_{i=1}^{p}a_{ik}\boldsymbol{\alpha}_i=\boldsymbol{0}$. 由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_p$ 线性无关可得 $a_{ik}=0(1\le i\le p,1\le k\le n)$,

于是 $\sum_{j=1}^{q} c_{ij}\boldsymbol{\beta}'_j = \boldsymbol{0}(1 \le i \le p)$. 再由 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_q$ 线性无关可得 $c_{ij} = 0(1 \le i \le p, 1 \le j \le q)$, 因此 $\{\boldsymbol{\alpha}_i \cdot \boldsymbol{\beta}'_j \mid 1 \le i \le p, 1 \le j \le q\}$ 线性无关. □

3.2 线性同构和几何问题代数化

我们有一类特别重要的线性同构. 设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 是 V 的一组基并固定次序. 对任一向量 $\boldsymbol{\alpha} \in V$, 设 $\boldsymbol{\alpha} = \lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 + \cdots + \lambda_n\boldsymbol{e}_n$, 则映射 $\eta : V \to \mathbb{K}^n$ 定义为:$\eta(\boldsymbol{\alpha}) = (\lambda_1, \lambda_2, \cdots, \lambda_n)'$, 即 η 将 V 中的向量映射到它在给定基下的坐标向量. 容易验证 $\eta : V \to \mathbb{K}^n$ 是一个线性同构. 因此, 通过这个线性同构, 我们可将抽象的线性空间 V 和具体的列向量空间 $\mathbb{K}^n$ 等同起来.

定理 3.3

(1) 同构关系是一种等价关系;

(2) 线性同构不仅将线性相关的向量组映射为线性相关的向量组, 而且将线性无关的向量组映射为线性无关的向量组;

(3) 同一个数域 $\mathbb{F}$ 上的线性空间同构的充要条件是它们具有相同的维数.

♡

证明

□

定理 3.4

定理假设和记号同上, 设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}$ 是 V 中向量, 它们在给定基下的坐标向量记为 $\widetilde{\boldsymbol{\alpha}}_1, \widetilde{\boldsymbol{\alpha}}_2, \cdots, \widetilde{\boldsymbol{\alpha}}_m, \widetilde{\boldsymbol{\beta}}$, 则

(1) $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性无关的充要条件是 $\widetilde{\boldsymbol{\alpha}}_1, \widetilde{\boldsymbol{\alpha}}_2, \cdots, \widetilde{\boldsymbol{\alpha}}_m$ 线性无关.

(2) $\boldsymbol{\beta}$ 可以用 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表示的充要条件是 $\widetilde{\boldsymbol{\beta}}$ 可以用 $\widetilde{\boldsymbol{\alpha}}_1, \widetilde{\boldsymbol{\alpha}}_2, \cdots, \widetilde{\boldsymbol{\alpha}}_m$ 线性表示, 并且线性表示的系数不变. 即

$$\boldsymbol{\beta} = c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_m\boldsymbol{\alpha}_m \Leftrightarrow \widetilde{\boldsymbol{\beta}} = c_1\widetilde{\boldsymbol{\alpha}}_1 + c_2\widetilde{\boldsymbol{\alpha}}_2 + \cdots + c_m\widetilde{\boldsymbol{\alpha}}_m.$$

(3) $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}$ 是向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的极大无关组的充要条件是 $\widetilde{\boldsymbol{\alpha}}_{i_1}, \widetilde{\boldsymbol{\alpha}}_{i_2}, \cdots, \widetilde{\boldsymbol{\alpha}}_{i_r}$ 是向量组 $\widetilde{\boldsymbol{\alpha}}_1, \widetilde{\boldsymbol{\alpha}}_2, \cdots, \widetilde{\boldsymbol{\alpha}}_m$ 的极大无关组. 特别地, 我们有

$$\mathrm{r}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m) = \mathrm{r}(\widetilde{\boldsymbol{\alpha}}_1, \widetilde{\boldsymbol{\alpha}}_2, \cdots, \widetilde{\boldsymbol{\alpha}}_m).$$

♡

笔记 由上述定理, 我们可以将抽象线性空间 V 中向量组线性关系的判定和秩的计算, 转化为具体列向量空间 $\mathbb{K}^n$ 中由它们的坐标向量构成的列向量组线性关系的判定和秩的计算. 由于后者通常可以通过矩阵的方法来处理, 故上述过程被称为"几何问题代数化".

证明 将定理 3.5运用到线性同构 η上就能得到证明. □

命题 3.15

设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k; \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 是向量空间 V 中的向量, 且满足:

$$\begin{cases} \boldsymbol{\beta}_1 = c_{11}\boldsymbol{\alpha}_1 + c_{12}\boldsymbol{\alpha}_2 + \cdots + c_{1k}\boldsymbol{\alpha}_k, \\ \boldsymbol{\beta}_2 = c_{21}\boldsymbol{\alpha}_1 + c_{22}\boldsymbol{\alpha}_2 + \cdots + c_{2k}\boldsymbol{\alpha}_k, \\ \cdots\cdots\cdots\cdots \\ \boldsymbol{\beta}_m = c_{m1}\boldsymbol{\alpha}_1 + c_{m2}\boldsymbol{\alpha}_2 + \cdots + c_{mk}\boldsymbol{\alpha}_k. \end{cases}$$

记上述表示式中的系数矩阵为 $\boldsymbol{C} = (c_{ij})_{m\times k}$, 则

(1) 若 $\mathrm{r}(\boldsymbol{C}) = k$, 则这两组向量等价.

(2) 若 $\mathrm{r}(\boldsymbol{C})=r$, 则向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 的秩不超过 r.

证明

(1) 在 V 中取定一组基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$, 假设在这组基下 $\boldsymbol{\alpha}_i$ 的坐标向量是 $\widetilde{\alpha}_i(1\le i\le k)$,$\boldsymbol{\beta}_j$ 的坐标向量是 $\widetilde{\boldsymbol{\beta}}_j(1\le j\le m)$, 则

$$\begin{cases}\widetilde{\beta}_1=c_{11}\widetilde{\alpha}_1+c_{12}\widetilde{\alpha}_2+\cdots+c_{1k}\widetilde{\alpha}_k,\\ \widetilde{\beta}_2=c_{21}\widetilde{\alpha}_1+c_{22}\widetilde{\alpha}_2+\cdots+c_{2k}\widetilde{\alpha}_k,\\ \cdots\cdots\cdots\cdots\\ \widetilde{\beta}_m=c_{m1}\widetilde{\alpha}_1+c_{m2}\widetilde{\alpha}_2+\cdots+c_{mk}\widetilde{\alpha}_k,\end{cases}$$

写成矩阵形式为

$$(\widetilde{\beta}_1,\widetilde{\beta}_2,\cdots,\widetilde{\beta}_m)=(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_k)\boldsymbol{C}'.$$

因为 $\boldsymbol{C}'$ 是一个行满秩 $k\times m$ 矩阵, 故由行满秩矩阵性质可知, 存在 $m\times k$ 矩阵 $\boldsymbol{T}$, 使得 $\boldsymbol{C}'\boldsymbol{T}=\boldsymbol{I}_k$, 于是

$$(\widetilde{\beta}_1,\widetilde{\beta}_2,\cdots,\widetilde{\beta}_m)\boldsymbol{T}=(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_k).$$

这表明 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 可用 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 来线性表示, 于是这两组向量等价.

(2) 在 V 中取定一组基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$, 假设在这组基下 $\boldsymbol{\alpha}_i$ 的坐标向量是 $\widetilde{\alpha}_i(1\le i\le k)$,$\boldsymbol{\beta}_j$ 的坐标向量是 $\widetilde{\boldsymbol{\beta}}_j(1\le j\le m)$, 则

$$\begin{cases}\widetilde{\beta}_1=c_{11}\widetilde{\alpha}_1+c_{12}\widetilde{\alpha}_2+\cdots+c_{1k}\widetilde{\alpha}_k,\\ \widetilde{\beta}_2=c_{21}\widetilde{\alpha}_1+c_{22}\widetilde{\alpha}_2+\cdots+c_{2k}\widetilde{\alpha}_k,\\ \cdots\cdots\cdots\cdots\\ \widetilde{\beta}_m=c_{m1}\widetilde{\alpha}_1+c_{m2}\widetilde{\alpha}_2+\cdots+c_{mk}\widetilde{\alpha}_k,\end{cases}$$

写成矩阵形式为

$$(\widetilde{\beta}_1,\widetilde{\beta}_2,\cdots,\widetilde{\beta}_m)=(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_k)\boldsymbol{C}'.$$

由于两个矩阵乘积的秩不超过每个矩阵的秩, 因此

$$\mathrm{r}\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m\right)=\mathrm{r}\left(\widetilde{\beta}_1,\widetilde{\beta}_2,\cdots,\widetilde{\beta}_m\right)=\mathrm{r}\left(\left(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_k\right)\boldsymbol{C}'\right)\leqslant\mathrm{r}\left(\boldsymbol{C}'\right)=\mathrm{r}\left(\boldsymbol{C}\right)=r.$$

□

命题 3.16

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 中的 m 个向量, 且已知它们的秩等于 r. 求证: 全体满足 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_m\boldsymbol{\alpha}_m=\mathbf{0}$ 的列向量 $(x_1,x_2,\cdots,x_m)'(x_i\in\mathbb{F})$ 构成数域 $\mathbb{F}$ 上 m 维列向量空间 $\mathbb{F}^m$ 的 $m-r$ 维子空间.

证明 在 V 中引进基以后, 记 $\widetilde{\alpha}_i$ 是 $\boldsymbol{\alpha}_i$ 的坐标向量, 则 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_m\boldsymbol{\alpha}_m=\mathbf{0}$ 等价于 $x_1\widetilde{\alpha}_1+x_2\widetilde{\alpha}_2+\cdots+x_m\widetilde{\alpha}_m=\mathbf{0}$. 而后者是一个齐次线性方程组, 其系数矩阵的秩等于 r(将 x_i 视为未知数), 故其解构成 $\mathbb{F}^m$ 的 $m-r$ 维子空间. □

例题 3.2 设 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是线性空间 V 的一组基, 问:$\{\boldsymbol{e}_1,\boldsymbol{e}_1+\boldsymbol{e}_2,\cdots,\boldsymbol{e}_1+\boldsymbol{e}_2+\cdots+\boldsymbol{e}_n\}$ 是否也是 V 的基?

笔记 利用定理 3.5即可.

解 将 $\{\boldsymbol{e}_1,\boldsymbol{e}_1+\boldsymbol{e}_2,\cdots,\boldsymbol{e}_1+\boldsymbol{e}_2+\cdots+\boldsymbol{e}_n\}$ 对应的坐标向量拼成如下矩阵:

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

显然 $|\boldsymbol{A}|=1$, 从而 $\boldsymbol{A}$ 是满秩阵, 于是 $\{\boldsymbol{e}_1,\boldsymbol{e}_1+\boldsymbol{e}_2,\cdots,\boldsymbol{e}_1+\boldsymbol{e}_2+\cdots+\boldsymbol{e}_n\}$ 也是 V 的一组基. □

例题 3.3 已知向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s\}(s>1)$ 是线性空间 V 的一组基, 设 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\beta}_s=\boldsymbol{\alpha}_s+\boldsymbol{\alpha}_1$. 讨论向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 的线性相关性.

笔记 利用定理 3.5即可.

解 将 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 对应的坐标向量拼成如下矩阵:

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

经计算可得 $|\boldsymbol{A}|=1+(-1)^{s+1}$. 因此当 s 为偶数时,$|\boldsymbol{A}|=0$, 从而向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性相关; 当 s 为奇数时,$|\boldsymbol{A}|=2$, 从而向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性无关. □

例题 3.4 设 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4\}$ 是线性空间 V 的一组基, 已知

$$\begin{cases}\boldsymbol{\alpha}_1=\boldsymbol{e}_1+\boldsymbol{e}_2+\boldsymbol{e}_3+3\boldsymbol{e}_4,\\\boldsymbol{\alpha}_2=-\boldsymbol{e}_1-3\boldsymbol{e}_2+5\boldsymbol{e}_3+\boldsymbol{e}_4,\\\boldsymbol{\alpha}_3=3\boldsymbol{e}_1+2\boldsymbol{e}_2-\boldsymbol{e}_3+4\boldsymbol{e}_4,\\\boldsymbol{\alpha}_4=-2\boldsymbol{e}_1-6\boldsymbol{e}_2+10\boldsymbol{e}_3+2\boldsymbol{e}_4,\end{cases}$$

求 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的一个极大无关组.

笔记 利用定理 3.5即可.

解 将 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 对应的坐标向量拼成如下矩阵, 并用初等行变换将其化为阶梯形矩阵:

$$\boldsymbol{A}=\begin{pmatrix}1&-1&3&-2\\1&-3&2&-6\\1&5&-1&10\\3&1&4&2\end{pmatrix}\to\begin{pmatrix}1&-1&3&-2\\0&-2&-1&-4\\0&0&-7&0\\0&0&0&0\end{pmatrix}.$$

因此, 矩阵 $\boldsymbol{A}$ 的第一列、第二列和第三列是坐标向量组的极大无关组, 从而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的一个极大无关组. □

例题 3.5 设 $a_1,a_2,\cdots,a_n$ 是 n 个不同的数,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是线性空间 V 的一组基, 已知

$$\begin{cases}\boldsymbol{\alpha}_1=\boldsymbol{e}_1+a_1\boldsymbol{e}_2+\cdots+a_1^{n-1}\boldsymbol{e}_n,\\\boldsymbol{\alpha}_2=\boldsymbol{e}_1+a_2\boldsymbol{e}_2+\cdots+a_2^{n-1}\boldsymbol{e}_n,\\\cdots\cdots\cdots\cdots\\\boldsymbol{\alpha}_n=\boldsymbol{e}_1+a_n\boldsymbol{e}_2+\cdots+a_n^{n-1}\boldsymbol{e}_n,\end{cases}$$

求证:$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 也是 V 的一组基.

笔记 利用定理 3.5即可.

证明 将 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 对应的坐标向量拼成如下矩阵:

$$\boldsymbol{A}=\begin{pmatrix}1&1&\cdots&1\\a_1&a_2&\cdots&a_n\\\vdots&\vdots&&\vdots\\a_1^{n-1}&a_2^{n-1}&\cdots&a_n^{n-1}\end{pmatrix}.$$

显然,$|\boldsymbol{A}|=\prod\limits_{1\le i<j\le n}(a_j-a_i)\neq 0$, 故 $\boldsymbol{A}$ 是满秩阵, 从而 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 也是 V 的一组基. □

3.3 线性同构和几何问题代数化

我们有一类特别重要的线性同构. 设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,$\{e_1, e_2, \cdots, e_n\}$ 是 V 的一组基并固定次序. 对任一向量 $\alpha \in V$, 设 $\alpha = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n$, 则映射 $\eta : V \to \mathbb{K}^n$ 定义为:$\eta(\alpha) = (\lambda_1, \lambda_2, \cdots, \lambda_n)'$, 即 η 将 V 中的向量映射到它在给定基下的坐标向量. 容易验证 $\eta : V \to \mathbb{K}^n$ 是一个线性同构. 因此, 通过这个线性同构, 我们可将抽象的线性空间 V 和具体的列向量空间 $\mathbb{K}^n$ 等同起来.

定理 3.5

(1) 同构关系是一种等价关系;

(2) 线性同构不仅将线性相关的向量组映射为线性相关的向量组, 而且将线性无关的向量组映射为线性无关的向量组;

(3) 同一个数域 $\mathbb{F}$ 上的线性空间同构的充要条件是它们具有相同的维数.

证明

□

定理 3.6

定理假设和记号同上, 设 $\alpha_1, \alpha_2, \cdots, \alpha_m, \beta$ 是 V 中向量, 它们在给定基下的坐标向量记为 $\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_m, \widetilde{\beta}$, 则

(1) $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性无关的充要条件是 $\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_m$ 线性无关.

(2) β 可以用 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性表示的充要条件是 $\widetilde{\beta}$ 可以用 $\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_m$ 线性表示, 并且线性表示的系数不变. 即

$$\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m \Leftrightarrow \widetilde{\beta} = c_1\widetilde{\alpha}_1 + c_2\widetilde{\alpha}_2 + \cdots + c_m\widetilde{\alpha}_m.$$

(3) $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 是向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 的极大无关组的充要条件是 $\widetilde{\alpha}_{i_1}, \widetilde{\alpha}_{i_2}, \cdots, \widetilde{\alpha}_{i_r}$ 是向量组 $\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_m$ 的极大无关组. 特别地, 我们有

$$\mathrm{r}(\alpha_1, \alpha_2, \cdots, \alpha_m) = \mathrm{r}(\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_m).$$

笔记 由上述定理, 我们可以将抽象线性空间 V 中向量组线性关系的判定和秩的计算, 转化为具体列向量空间 $\mathbb{K}^n$ 中由它们的坐标向量构成的列向量组线性关系的判定和秩的计算. 由于后者通常可以通过矩阵的方法来处理, 故上述过程被称为 “几何问题代数化” .

证明 将定理 3.5运用到线性同构 η上就能得到证明. □

命题 3.17

设 $\alpha_1, \alpha_2, \cdots, \alpha_k; \beta_1, \beta_2, \cdots, \beta_m$ 是向量空间 V 中的向量, 且满足:

$$\begin{cases} \beta_1 = c_{11}\alpha_1 + c_{12}\alpha_2 + \cdots + c_{1k}\alpha_k, \\ \beta_2 = c_{21}\alpha_1 + c_{22}\alpha_2 + \cdots + c_{2k}\alpha_k, \\ \cdots\cdots\cdots\cdots \\ \beta_m = c_{m1}\alpha_1 + c_{m2}\alpha_2 + \cdots + c_{mk}\alpha_k. \end{cases}$$

记上述表示式中的系数矩阵为 $C = (c_{ij})_{m\times k}$, 则

(1) 若 $\mathrm{r}(C) = k$, 则这两组向量等价.

(2) 若 $\mathrm{r}(C) = r$, 则向量组 $\beta_1, \beta_2, \cdots, \beta_m$ 的秩不超过 r.

证明

(1) 在 V 中取定一组基 $e_1, e_2, \cdots, e_n$, 假设在这组基下 α_i 的坐标向量是 $\widetilde{\alpha}_i(1 \le i \le k)$,$\beta_j$ 的坐标向量是 $\widetilde{\beta}_j(1 \le$

$j \le m$), 则

$$\begin{cases} \widetilde{\beta}_1 = c_{11}\widetilde{\alpha}_1 + c_{12}\widetilde{\alpha}_2 + \cdots + c_{1k}\widetilde{\alpha}_k, \\ \widetilde{\beta}_2 = c_{21}\widetilde{\alpha}_1 + c_{22}\widetilde{\alpha}_2 + \cdots + c_{2k}\widetilde{\alpha}_k, \\ \cdots\cdots\cdots\cdots \\ \widetilde{\beta}_m = c_{m1}\widetilde{\alpha}_1 + c_{m2}\widetilde{\alpha}_2 + \cdots + c_{mk}\widetilde{\alpha}_k, \end{cases}$$

写成矩阵形式为

$$(\widetilde{\beta}_1, \widetilde{\beta}_2, \cdots, \widetilde{\beta}_m) = (\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_k)\boldsymbol{C}'.$$

因为 $\boldsymbol{C}'$ 是一个行满秩 $k \times m$ 矩阵, 故由行满秩矩阵性质可知, 存在 $m \times k$ 矩阵 $\boldsymbol{T}$, 使得 $\boldsymbol{C}'\boldsymbol{T} = \boldsymbol{I}_k$, 于是

$$(\widetilde{\beta}_1, \widetilde{\beta}_2, \cdots, \widetilde{\beta}_m)\boldsymbol{T} = (\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_k).$$

这表明 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k$ 可用 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 来线性表示, 于是这两组向量等价.

(2) 在 V 中取定一组基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$, 假设在这组基下 $\boldsymbol{\alpha}_i$ 的坐标向量是 $\widetilde{\alpha}_i(1 \le i \le k)$,$\boldsymbol{\beta}_j$ 的坐标向量是 $\widetilde{\beta}_j(1 \le j \le m)$, 则

$$\begin{cases} \widetilde{\beta}_1 = c_{11}\widetilde{\alpha}_1 + c_{12}\widetilde{\alpha}_2 + \cdots + c_{1k}\widetilde{\alpha}_k, \\ \widetilde{\beta}_2 = c_{21}\widetilde{\alpha}_1 + c_{22}\widetilde{\alpha}_2 + \cdots + c_{2k}\widetilde{\alpha}_k, \\ \cdots\cdots\cdots\cdots \\ \widetilde{\beta}_m = c_{m1}\widetilde{\alpha}_1 + c_{m2}\widetilde{\alpha}_2 + \cdots + c_{mk}\widetilde{\alpha}_k, \end{cases}$$

写成矩阵形式为

$$(\widetilde{\beta}_1, \widetilde{\beta}_2, \cdots, \widetilde{\beta}_m) = (\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_k)\boldsymbol{C}'.$$

由于两个矩阵乘积的秩不超过每个矩阵的秩, 因此

$$\mathrm{r}\left(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\right) = \mathrm{r}\left(\widetilde{\beta}_1, \widetilde{\beta}_2, \cdots, \widetilde{\beta}_m\right) = \mathrm{r}\left(\left(\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_k\right)\boldsymbol{C}'\right) \leqslant \mathrm{r}\left(\boldsymbol{C}'\right) = \mathrm{r}\left(\boldsymbol{C}\right) = r.$$

□

命题 3.18

设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 中的 m 个向量, 且已知它们的秩等于 r. 求证: 全体满足 $x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_m\boldsymbol{\alpha}_m = \boldsymbol{0}$ 的列向量 $(x_1, x_2, \cdots, x_m)'(x_i \in \mathbb{F})$ 构成数域 $\mathbb{F}$ 上 m 维列向量空间 $\mathbb{F}^m$ 的 $m - r$ 维子空间.

证明 在 V 中引进基以后, 记 $\widetilde{\alpha}_i$ 是 $\boldsymbol{\alpha}_i$ 的坐标向量, 则 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_m\boldsymbol{\alpha}_m = \boldsymbol{0}$ 等价于 $x_1\widetilde{\alpha}_1+x_2\widetilde{\alpha}_2+\cdots+x_m\widetilde{\alpha}_m = \boldsymbol{0}$. 而后者是一个齐次线性方程组, 其系数矩阵的秩等于 r(将 x_i 视为未知数), 故其解构成 $\mathbb{F}^m$ 的 $m - r$ 维子空间. □

例题 3.6 设 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 是线性空间 V 的一组基, 问:$\{\boldsymbol{e}_1, \boldsymbol{e}_1 + \boldsymbol{e}_2, \cdots, \boldsymbol{e}_1 + \boldsymbol{e}_2 + \cdots + \boldsymbol{e}_n\}$ 是否也是 V 的基?

笔记 利用定理 3.5即可.

解 将 $\{\boldsymbol{e}_1, \boldsymbol{e}_1 + \boldsymbol{e}_2, \cdots, \boldsymbol{e}_1 + \boldsymbol{e}_2 + \cdots + \boldsymbol{e}_n\}$ 对应的坐标向量拼成如下矩阵:

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

显然 $|\boldsymbol{A}| = 1$, 从而 $\boldsymbol{A}$ 是满秩阵, 于是 $\{\boldsymbol{e}_1, \boldsymbol{e}_1 + \boldsymbol{e}_2, \cdots, \boldsymbol{e}_1 + \boldsymbol{e}_2 + \cdots + \boldsymbol{e}_n\}$ 也是 V 的一组基. □

例题 3.7 已知向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s\}(s > 1)$ 是线性空间 V 的一组基, 设 $\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2, \boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3, \cdots, \boldsymbol{\beta}_s = \boldsymbol{\alpha}_s+\boldsymbol{\alpha}_1$. 讨论向量 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 的线性相关性.

笔记 利用定理 3.5即可.

解 将 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 对应的坐标向量拼成如下矩阵:

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

经计算可得 $|\boldsymbol{A}|=1+(-1)^{s+1}$. 因此当 s 为偶数时,$|\boldsymbol{A}|=0$, 从而向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性相关; 当 s 为奇数时,$|\boldsymbol{A}|=2$, 从而向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性无关. □

例题 3.8 设 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4\}$ 是线性空间 V 的一组基, 已知

$$\begin{cases}\boldsymbol{\alpha}_1=\boldsymbol{e}_1+\boldsymbol{e}_2+\boldsymbol{e}_3+3\boldsymbol{e}_4,\\\boldsymbol{\alpha}_2=-\boldsymbol{e}_1-3\boldsymbol{e}_2+5\boldsymbol{e}_3+\boldsymbol{e}_4,\\\boldsymbol{\alpha}_3=3\boldsymbol{e}_1+2\boldsymbol{e}_2-\boldsymbol{e}_3+4\boldsymbol{e}_4,\\\boldsymbol{\alpha}_4=-2\boldsymbol{e}_1-6\boldsymbol{e}_2+10\boldsymbol{e}_3+2\boldsymbol{e}_4,\end{cases}$$

求 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的一个极大无关组.

笔记 利用定理 3.5即可.

解 将 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 对应的坐标向量拼成如下矩阵, 并用初等行变换将其化为阶梯形矩阵:

$$\boldsymbol{A}=\begin{pmatrix}1&-1&3&-2\\1&-3&2&-6\\1&5&-1&10\\3&1&4&2\end{pmatrix}\to\begin{pmatrix}1&-1&3&-2\\0&-2&-1&-4\\0&0&-7&0\\0&0&0&0\end{pmatrix}.$$

因此, 矩阵 $\boldsymbol{A}$ 的第一列、第二列和第三列是坐标向量组的极大无关组, 从而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的一个极大无关组. □

例题 3.9 设 $a_1,a_2,\cdots,a_n$ 是 n 个不同的数,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是线性空间 V 的一组基, 已知

$$\begin{cases}\boldsymbol{\alpha}_1=\boldsymbol{e}_1+a_1\boldsymbol{e}_2+\cdots+a_1^{n-1}\boldsymbol{e}_n,\\\boldsymbol{\alpha}_2=\boldsymbol{e}_1+a_2\boldsymbol{e}_2+\cdots+a_2^{n-1}\boldsymbol{e}_n,\\\cdots\cdots\cdots\cdots\\\boldsymbol{\alpha}_n=\boldsymbol{e}_1+a_n\boldsymbol{e}_2+\cdots+a_n^{n-1}\boldsymbol{e}_n,\end{cases}$$

求证:$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 也是 V 的一组基.

笔记 利用定理 3.5即可.

证明 将 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 对应的坐标向量拼成如下矩阵:

$$\boldsymbol{A}=\begin{pmatrix}1&1&\cdots&1\\a_1&a_2&\cdots&a_n\\\vdots&\vdots&&\vdots\\a_1^{n-1}&a_2^{n-1}&\cdots&a_n^{n-1}\end{pmatrix}.$$

显然,$|\boldsymbol{A}|=\prod\limits_{1\le i<j\le n}(a_j-a_i)\neq 0$, 故 $\boldsymbol{A}$ 是满秩阵, 从而 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 也是 V 的一组基. □

3.4 基变换与过渡矩阵

定义 3.1 (过渡矩阵)

设 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 和 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 是 n 维线性空间 V 的两组基, 若

$$\begin{cases} \boldsymbol{f}_1 = a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2 + \cdots + a_{n1}\boldsymbol{e}_n, \\ \boldsymbol{f}_2 = a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2 + \cdots + a_{n2}\boldsymbol{e}_n, \\ \cdots\cdots\cdots\cdots \\ \boldsymbol{f}_n = a_{1n}\boldsymbol{e}_1 + a_{2n}\boldsymbol{e}_2 + \cdots + a_{nn}\boldsymbol{e}_n, \end{cases}$$

则矩阵

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

称为从基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵. 并且 $(\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n) = (\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)\boldsymbol{A}$. ♣

定理 3.7 (同一向量在不同基下坐标向量的关系)

设 V 是数域 $\mathbb{F}$ 上 n 维线性空间, 从基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵为 $\boldsymbol{A}=(a_{ij})$. 若 V 中向量 $\boldsymbol{\alpha}$ 在基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下的坐标向量是 $(x_1,x_2,\cdots,x_n)'$, 在基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 下的坐标向量是 $(y_1,y_2,\cdots,y_n)'$, 则

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

♡

证明 由过渡矩阵定义可得

$$\boldsymbol{\alpha} = (\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = (\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)\boldsymbol{A}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

又因为 $(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)$ 可逆, 所以

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

□

定理 3.8

矩阵 $\boldsymbol{A}$ 是 n 维线性空间 V 的基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵, 则 $\boldsymbol{A}$ 是可逆矩阵且从基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 到基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 的过渡矩阵为 $\boldsymbol{A}^{-1}$. 又若 $\boldsymbol{B}$ 是从基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 到基 $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_n\}$ 的过渡矩阵, 则从基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到基 $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_n\}$ 的过渡矩阵为 $\boldsymbol{AB}$.

♡

例题 3.10 设 $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n\},\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\},\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 是向量空间 V 的 3 组基. 若从 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 到

$\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 的过渡矩阵是 $\boldsymbol{A}$, 从 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 到 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 的过渡矩阵是 $\boldsymbol{B}$, 求从 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 到 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 的过渡矩阵.

解 从 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 到 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 的过渡矩阵为 $\boldsymbol{A}^{-1}$, 故从 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 到 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 的过渡矩阵为 $\boldsymbol{A}^{-1}\boldsymbol{B}$.

□

例题 3.11 在四维行向量空间中求从基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 到 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 的过渡矩阵, 其中

$$\boldsymbol{e}_1=(1,1,0,1),\boldsymbol{e}_2=(2,1,2,0),\boldsymbol{e}_3=(1,1,0,0),\boldsymbol{e}_4=(0,1,-1,-1),$$
$$\boldsymbol{f}_1=(1,0,0,1),\boldsymbol{f}_2=(0,0,1,-1),\boldsymbol{f}_3=(2,1,0,3),\boldsymbol{f}_4=(-1,0,1,2).$$

笔记 这类题如用求解线性方程组的方法比较繁, 可采用下列方法.

解 设该向量空间的标准基为

$$\boldsymbol{u}_1=(1,0,0,0),\boldsymbol{u}_2=(0,1,0,0),\boldsymbol{u}_3=(0,0,1,0),\boldsymbol{u}_4=(0,0,0,1),$$

则由条件可知从 $\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4$ 到 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4$ 的过渡矩阵为

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

于是就有

$$(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4)=(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4)\boldsymbol{A}\Rightarrow(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4)=(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4)\boldsymbol{A}^{-1}. \tag{3.3}$$

又由条件可知从 $\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4$ 到 $\boldsymbol{f}_1,\boldsymbol{f}_2,\boldsymbol{f}_3,\boldsymbol{f}_4$ 的过渡矩阵为

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

于是就有

$$(\boldsymbol{f}_1,\boldsymbol{f}_2,\boldsymbol{f}_3,\boldsymbol{f}_4)=(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4)\boldsymbol{B}. \tag{3.4}$$

从而由(3.3)(3.4)式可得

$$(\boldsymbol{f}_1,\boldsymbol{f}_2,\boldsymbol{f}_3,\boldsymbol{f}_4)=(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4)\boldsymbol{B}=(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4)\boldsymbol{A}^{-1}\boldsymbol{B}.$$

故从基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4$ 到 $\boldsymbol{f}_1,\boldsymbol{f}_2,\boldsymbol{f}_3,\boldsymbol{f}_4$ 的过渡矩阵为 $\boldsymbol{A}^{-1}\boldsymbol{B}$. 它可以用初等变换和求逆矩阵类似的方法直接求得 (对矩阵 $(\boldsymbol{A}|\boldsymbol{B})$ 进行初等行变换, 将 $\boldsymbol{A}$ 化为单位矩阵, 则右边一块就化为了 $\boldsymbol{A}^{-1}\boldsymbol{B}$) 因此, 所求之过渡矩阵为

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

□

例题 3.12 设 a 为常数, 求向量 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)$ 在基 $\{\boldsymbol{f}_1=(a^{n-1},a^{n-2},\cdots,a,1),\boldsymbol{f}_2=(a^{n-2},a^{n-3},\cdots,1,0),\cdots,\boldsymbol{f}_n=(1,0,\cdots,0,0)\}$ 下的坐标.

证明 设 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是标准单位行向量, 则 $\boldsymbol{\alpha}$ 在 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下的坐标向量就是 $\boldsymbol{\alpha}'$, 并且从 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$

到 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵是

$$\boldsymbol{A}=\begin{pmatrix} a^{n-1} & a^{n-2} & \cdots & 1\\ a^{n-2} & a^{n-3} & \cdots & 0\\ \vdots & \vdots & & \vdots\\ a & 1 & \cdots & 0\\ 1 & 0 & \cdots & 0\end{pmatrix}.$$

设 $\boldsymbol{\alpha}$ 在 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 下的坐标向量为 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$. 则由同一向量在不同基下坐标向量的关系有 $\boldsymbol{A}\boldsymbol{x}'=\boldsymbol{\alpha}'$. 这是一个非齐次线性方程组, 可由初等行变换求出方程组的解:

$$\begin{pmatrix} a^{n-1} & a^{n-2} & \cdots & 1 & a_1\\ a^{n-2} & a^{n-3} & \cdots & 0 & a_2\\ \vdots & \vdots & & \vdots & \vdots\\ a & 1 & \cdots & 0 & a_{n-1}\\ 1 & 0 & \cdots & 0 & a_n\end{pmatrix}\to\begin{pmatrix} 0 & a^{n-2} & \cdots & 1 & a_1-a^{n-1}a_n\\ 0 & a^{n-3} & \cdots & 0 & a_2-a^{n-2}a_n\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 1 & \cdots & 0 & a_{n-1}-aa_n\\ 1 & 0 & \cdots & 0 & a_n\end{pmatrix}$$

$$\to\begin{pmatrix} 0 & 0 & \cdots & 1 & a_1-aa_2\\ 0 & 0 & \cdots & 0 & a_2-aa_3\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 1 & \cdots & 0 & a_{n-1}-aa_n\\ 1 & 0 & \cdots & 0 & a_n\end{pmatrix}\to\begin{pmatrix} 1 & 0 & \cdots & 0 & a_n\\ 0 & 1 & \cdots & 0 & a_{n-1}-aa_n\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & a_2-aa_3\\ 0 & 0 & \cdots & 1 & a_1-aa_2\end{pmatrix},$$

因此 $\boldsymbol{x}=(a_n,a_{n-1}-aa_n,\cdots,a_2-aa_3,a_1-aa_2)$. □

例题 3.13 设 V 是次数不超过 n 的实系数多项式全体组成的线性空间, 求从基 $\{1,x,x^2,\cdots,x^n\}$ 到基 $\{1,x-a,(x-a)^2,\cdots,(x-a)^n\}$ 的过渡矩阵, 并以此证明多项式的 Taylor 公式:

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n,$$

其中 $f^{(n)}(x)$ 表示 $f(x)$ 的 n 次导数.

解 从基 $\{1,x,x^2,\cdots,x^n\}$ 到基 $\{1,x-a,(x-a)^2,\cdots,(x-a)^n\}$ 的过渡矩阵 ($n+1$ 阶) 利用二项式定理容易求出为

$$\boldsymbol{P}=\begin{pmatrix} 1 & -a & a^2 & \cdots & (-1)^n a^n\\ 0 & 1 & -2a & \cdots & (-1)^{n-1}na^{n-1}\\ 0 & 0 & 1 & \cdots & (-1)^{n-2}\dfrac{n(n-1)}{2!}a^{n-2}\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}.$$

注意 $\boldsymbol{P}$ 的逆矩阵实际上就是从基 $\{1,x-a,(x-a)^2,\cdots,(x-a)^n\}$ 到基 $\{1,x,x^2,\cdots,x^n\}$ 的过渡矩阵, 结合 $x^n=[(x-a)+a]^n$, 再利用二项式定理可以马上得到 (不必用初等变换法求逆矩阵):

$$\boldsymbol{P}^{-1}=\begin{pmatrix} 1 & a & a^2 & \cdots & a^n\\ 0 & 1 & 2a & \cdots & na^{n-1}\\ 0 & 0 & 1 & \cdots & \dfrac{n(n-1)}{2!}a^{n-2}\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}.$$

设 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$, 则 $f(x)$ 在基 $\{1,x,x^2,\cdots,x^n\}$ 下的坐标向量为 $(a_0,a_1,\cdots,a_n)'$. 设 $f(x)$ 在基

$\{1, x-a, (x-a)^2, \cdots, (x-a)^n\}$ 下的坐标向量为 $(y_0, y_1, \cdots, y_n)'$. 则由同一向量在不同基下坐标向量的关系可知

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \boldsymbol{P} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

于是

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \boldsymbol{P}^{-1} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & a & a^2 & \cdots & a^n \\ 0 & 1 & 2a & \cdots & na^{n-1} \\ 0 & 0 & 1 & \cdots & \dfrac{n(n-1)}{2!}a^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f(a) \\ \dfrac{f'(a)}{1!} \\ \vdots \\ \dfrac{f^{(n)}(a)}{n!} \end{pmatrix}.$$

由此可得

$$f(x) = \left(1, x-a, (x-a)^2, \cdots, (x-a)^n\right) \begin{pmatrix} f(a) \\ \dfrac{f'(a)}{1!} \\ \vdots \\ \dfrac{f^{(n)}(a)}{n!} \end{pmatrix} = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

□

3.4.1 练习

练习 3.1 验证下列映射是线性同构:

(1) 一维实行向量空间 $\mathbb{R}$, 例题????中的实线性空间 $\mathbb{R}^+$, 映射 $\varphi: \mathbb{R} \to \mathbb{R}^+$ 定义为 $\varphi(x) = \mathrm{e}^x$;

(2) 二维实行向量空间 $\mathbb{R}_2$, 例题????中的实线性空间 V, 映射 $\varphi: \mathbb{R}_2 \to V$ 定义为 $\varphi(a, b) = (a, b + \dfrac{1}{2}a^2)$.

解

(1) φ 的逆映射是 $\psi: \mathbb{R}^+ \to \mathbb{R}, \psi(y) = \ln y$, 故 φ 是一一对应的. 根据加法和数乘的定义可得

$$\varphi(x+y) = \mathrm{e}^{x+y} = \mathrm{e}^x\mathrm{e}^y = \varphi(x) \oplus \varphi(y), \varphi(kx) = \mathrm{e}^{kx} = (\mathrm{e}^x)^k = k \circ \varphi(x),$$

因此 $\varphi: \mathbb{R} \to \mathbb{R}^+$ 是线性同构.

(2) φ 的逆映射是 $\psi: V \to \mathbb{R}_2, \psi(x, y) = (x, y - \dfrac{x^2}{2})$, 故 φ 是一一对应的. 根据具体的计算可得

$$\varphi(a_1 + a_2, b_1 + b_2) = \varphi(a_1, b_1) \oplus \varphi(a_2, b_2), \varphi(ka, kb) = k \circ \varphi(a, b),$$

因此 $\varphi: \mathbb{R}_2 \to V$ 是线性同构.

□

练习 3.2 构造下列线性空间之间的线性同构:

(1) V 是数域 $\mathbb{K}$ 上的 n 阶上三角矩阵构成的线性空间,U 是数域 $\mathbb{K}$ 上的 n 阶对称矩阵构成的线性空间 (例题????);

(2) V 是数域 $\mathbb{K}$ 上主对角元全为零的 n 阶上三角矩阵构成的线性空间,U 是数域 $\mathbb{K}$ 上的 n 阶反对称矩阵构成的线性空间 (例题????);

(3) V 是 n 阶 Hermite 矩阵构成的实线性空间,U 是 n 阶斜 Hermite 矩阵构成的实线性空间 (例题??).

解

(1) $\varphi: V \to U$ 定义为: 对任意的 $\boldsymbol{A} = (a_{ij}) \in V$, 当 $i \leq j$ 时, 矩阵 $\varphi(\boldsymbol{A})$ 的第 (i, j) 元素为 a_{ij}; 当 $i > j$ 时, 矩阵 $\varphi(\boldsymbol{A})$ 的第 (i, j) 元素为 a_{ji}. 容易验证 $\varphi: V \to U$ 是定义好的映射, 并且是数域 $\mathbb{K}$ 上的线性同构.

(2) $\varphi: V \to U$ 定义为: 对任意的 $\boldsymbol{A} = (a_{ij}) \in V, \varphi(\boldsymbol{A}) = \boldsymbol{A} - \boldsymbol{A}'$. 容易验证 $\varphi: V \to U$ 是定义好的映射, 并且是数域 $\mathbb{K}$ 上的线性同构.

(3) $\varphi: V \to U$ 定义为: 对任意的 $\boldsymbol{A} = (a_{ij}) \in V, \varphi(\boldsymbol{A}) = \mathrm{i}\boldsymbol{A}$. 容易验证 $\varphi: V \to U$ 是定义好的映射, 并且是实数域上的线性同构. 注意到 φ 的逆映射 $\psi: U \to V$ 为:$\psi(\boldsymbol{B}) = -\mathrm{i}\boldsymbol{B}$.

□

3.5 子空间、直和与商空间

定理 3.9 (基扩张定理)

设 V 是 n 维线性空间,$v_1, v_2, \cdots, v_m$ 是 V 中 $m(m < n)$ 个线性无关的向量, 又假设 $\{e_1, e_2, \cdots, e_n\}$ 是 V 的一组基, 则必可在 $\{e_1, e_2, \cdots, e_n\}$ 中选出 $n - m$ 个向量, 使之和 $v_1, v_2, \cdots, v_m$ 一起组成 V 的一组基.

基扩张定理还有几种等价形式:

(1)n 维线性空间 V 中任意 $m(m < n)$ 个线性无关的向量均可扩张为 V 的一组基.

(2)n 维线性空间 V 的任意一个子空间的基均可扩张为 V 的一组基.

♡

证明 将 $e_i(i = 1, \cdots, n)$ 依次放入 $\{v_1, v_2, \cdots, v_m\}$, 则必有一个 $e_{i'}$, 使 $v_1, v_2, \cdots, v_m, e_{i'}$ 线性无关. 这是因为若任一 e_i 加入 $v_1, v_2, \cdots, v_m$ 后线性相关, 则每个 e_i 可用 $v_1, v_2, \cdots, v_m$ 线性表示, 将和定理 3.2(1)的结论矛盾. 现不妨设 $i' = m + 1$. 若 $m + 1 < n$, 又可从 $e_1, e_2, \cdots, e_n$ 中找到一个向量, 加入 $\{v_1, v_2, \cdots, v_m, e_{m+1}\}$ 后仍线性无关. 不断这样做下去, 便可将 $v_1, v_2, \cdots, v_m$ 扩张成为 V 的一组基. □

定义 3.2 (直和)

设 $V_1, V_2, \cdots, V_k$ 是线性空间 V 的子空间, 若对任意的 $i(1 \leq i \leq k)$, 均有

$$V_i \cap (V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots + V_k) = 0,$$

则称和 $V_1 + V_2 + \cdots + V_k$ 是直接和, 简称直和, 记为 $V_1 \oplus V_2 \oplus \cdots \oplus V_k$.

♣

定理 3.10 (直和的等价条件)

设 $V_1, V_2, \cdots, V_k$ 是线性空间 V_0 的子空间,$V_0 = V_1 + V_2 + \cdots + V_k$, 则下列命题等价:

(1) $V_0 = V_1 \oplus V_2 \oplus \cdots \oplus V_k$;

(2) 对任意的 $2 \leq i \leq k$, 有 $V_i \cap (V_1 + V_2 + \cdots + V_{i-1}) = 0$;

(3) $\dim V_0 = \dim V_1 + \dim V_2 + \cdots + \dim V_k$;

(4) $V_1, V_2, \cdots, V_k$ 的一组基可以拼成 V_0 的一组基;

(5) V_0 中的向量表示为 $V_1, V_2, \cdots, V_k$ 中的向量之和时其表示唯一.

(6) 零向量表示唯一.

♡

证明

□

定理 3.11 (维数公式)

设 V_1, V_2 是线性空间 V 的两个子空间, 则

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

♡

证明

□

3.5.1 证明直和的方法

证明直和的方法大致有两种:

第一种: 先证和, 再证直和.

第二种: 对于给定的 V, V_1, V_2, 求证 $V = V_1 \oplus V_2$ 的题目, 如果“和”不好证明的话, 可以记 $W = V_1 + V_2$, 先证 $W = V_1 \oplus V_2$, 再证 $V = W$(证明 $V = W$ 通常会利用命题**??**). 具体例子见例题 3.16

命题 3.19

设 V 是数域 $\mathbb{F}$ 上 n 阶矩阵组成的向量空间,V_1 和 V_2 分别是 $\mathbb{F}$ 上对称矩阵和反对称矩阵组成的子集. 求证:V_1 和 V_2 都是 V 的子空间且 $V = V_1 \oplus V_2$.

笔记 要证明向量空间 V 是其子空间 V_1, V_2 的直和, 只需证明两件事: 一是证明 V 中任一向量均可表示为 V_1 与 V_2 中向量之和, 即 $V = V_1 + V_2$; 二是证明 V_1 与 V_2 的交等于零.

证明 由于对称矩阵之和仍是对称矩阵, 一个数乘以对称矩阵仍是对称矩阵, 因此 V_1 是 V 的子空间. 同理 V_2 也是 V 的子空间. 又由命题 2.12可知, 任一 n 阶矩阵都可以表示为一个对称矩阵和一个反对称矩阵之和, 故 $V = V_1 + V_2$. 若一个矩阵既是对称矩阵又是反对称矩阵, 则它一定是零矩阵. 这就是说 $V_1 \cap V_2 = \mathbf{0}$. 于是 $V = V_1 \oplus V_2$. □

例题 3.14 设 V_1, V_2 分别是数域 $\mathbb{F}$ 上的齐次线性方程组 $x_1 = x_2 = \cdots = x_n$ 与 $x_1 + x_2 + \cdots + x_n = 0$ 的解空间, 求证:$\mathbb{F}^n = V_1 \oplus V_2$.

笔记 要证明向量空间 V 是其子空间 V_1, V_2 的直和, 只需证明两件事: 一是证明 V 中任一向量均可表示为 V_1 与 V_2 中向量之和, 即 $V = V_1 + V_2$; 二是证明 V_1 与 V_2 的交等于零.

证明 由线性方程组解的定理知,V_1 的维数是 1,V_2 的维数是 $n-1$. 若列向量 $\boldsymbol{\alpha} \in V_1 \cap V_2$, 则 $\boldsymbol{\alpha}$ 既是第一个线性方程组的解, 也是第二个线性方程组的解, 不难看出 $\boldsymbol{\alpha}$ 只能等于零向量, 因此 $V_1 \cap V_2 = 0$. 又因为

$$\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2 = 1 + (n-1) = n = \dim \mathbb{F}^n,$$

故 $\mathbb{F}^n = V_1 \oplus V_2$. □

例题 3.15 设 U, V 是数域 $\mathbb{K}$ 上的两个线性空间,$W = U \times V$ 是 U 和 V 的积集合, 即 $W = \{(\boldsymbol{u}, \boldsymbol{v}) | \boldsymbol{u} \in U, \boldsymbol{v} \in V\}$. 现在 W 上定义加法和数乘:

$$(\boldsymbol{u}_1, \boldsymbol{v}_1) + (\boldsymbol{u}_2, \boldsymbol{v}_2) = (\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{v}_1 + \boldsymbol{v}_2), k(\boldsymbol{u}, \boldsymbol{v}) = (k\boldsymbol{u}, k\boldsymbol{v}).$$

验证:W 是 $\mathbb{K}$ 上的线性空间 (这个线性空间称为 U 和 V 的外直和).

又若设 $U' = \{(\boldsymbol{u}, \mathbf{0}) | \boldsymbol{u} \in U\}, V' = \{(\mathbf{0}, \boldsymbol{v}) | \boldsymbol{v} \in V\}$, 求证:$U', V'$ 是 W 的子空间,U' 和 U 同构,V' 和 V 同构, 并且 $W = U' \oplus V'$.

证明 易验证 W 在上述加法和数乘下满足线性空间的 8 条公理, 从而是 $\mathbb{K}$ 上的线性空间. 任取 $(\boldsymbol{u}_1, \mathbf{0}), (\boldsymbol{u}_2, \mathbf{0}) \in U', k \in \mathbb{K}$, 则 $(\boldsymbol{u}_1, \mathbf{0}) + (\boldsymbol{u}_2, \mathbf{0}) = (\boldsymbol{u}_1 + \boldsymbol{u}_2, \mathbf{0}) \in U', k(\boldsymbol{u}_1, \mathbf{0}) = (k\boldsymbol{u}_1, \mathbf{0}) \in U'$, 因此 U' 是 W 的子空间. 同理可证 V' 是 W 的子空间. 构造映射 $\varphi : U \to U', \varphi(\boldsymbol{u}) = (\boldsymbol{u}, \mathbf{0})$, 容易验证 φ 是一一对应并且保持加法和数乘运算, 所以 $\varphi : U \to U'$ 是一个线性同构. 构造映射 $\psi : V \to V', \psi(\boldsymbol{v}) = (\mathbf{0}, \boldsymbol{v})$, 同理可证 $\psi : V \to V'$ 是一个线性同构. 显然 $U' \cap V' = 0$, 又对 W 中任一向量 $(\boldsymbol{u}, \boldsymbol{v})$, 有 $(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \mathbf{0}) + (\mathbf{0}, \boldsymbol{v}) \in U' + V'$, 因此 $W = U' \oplus V'$. □

例题 3.16 给定数域 P, 设 $\boldsymbol{A}$ 是数域 P 上的一个 n 级可逆方阵,$\boldsymbol{A}$ 的前 r 个行向量组成的矩阵为 $\boldsymbol{B}$, 后 $n-r$ 个行向量组成的矩阵为 $\boldsymbol{C}$,n 元线性方程组 $\boldsymbol{BX} = 0$ 与 $\boldsymbol{CX} = 0$ 的解空间分别为 V_1, V_2, 证 $P^n = V_1 \oplus V_2$.

证明 先记 $W = V_1 + V_2$. 若 $\boldsymbol{\alpha} \in V_1 \cap V_2$, 则 $\boldsymbol{B\alpha} = \boldsymbol{C\alpha} = 0$, 所以

$$A\alpha = \begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{C} \end{pmatrix} \alpha = 0.$$

由于 $\boldsymbol{A}$ 可逆, 知 $\boldsymbol{\alpha} = 0$, 所以 $V_1 \cap V_1 = \{0\}$, 即 $W = V_1 \oplus V_2$.

最后说 $W = P^n$: 显然 $r(\boldsymbol{B}) = r, r(\boldsymbol{C}) = n-r$, 则 $\dim V_1 = n-r, \dim V_2 = n-(n-r) = r$. 所以

$$\dim W = \dim V_1 + \dim V_2 = n = \dim P^n.$$

又 $W = V_1 \oplus V_2 \subseteq P^n$, 从而 $W = P^n$, 即

$$P^n = V_1 \oplus V_2.$$

□

命题 3.20 (任意子空间一定存在相应的补空间)

设 U 是 V 的子空间, 则一定存在 V 的子空间 W, 使得 $V = U \oplus W$. 这样的子空间 W 称为子空间 U 在 V 中的**补空间**.

注 在这个命题中 $U \cap W = \{\mathbf{0}\}$, 而不是 $U \cap W = \varnothing$; 同时 $V = U + W$ 是子空间的和, 而不是 $V = U \cup W$. 因此, 补空间绝不是补集, 请读者务必注意! 一般来说, 补空间并不唯一. 例如下面证明中, 取 U 中不同的基, 再将基扩张得到的补空间也不相同. 还例如, 若 $\dim V - \dim U \geq 1$ 且 $\dim U \geq 1$, 则 U 有无限个补空间.

证明 取子空间 U 的一组基 $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m\}$, 由基扩张定理可将其扩张为 V 的一组基 $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \cdots, \boldsymbol{e}_n\}$. 令 $W = L(\boldsymbol{e}_{m+1}, \cdots, \boldsymbol{e}_n)$, 则 $V = U + W$. 由于 $\{\boldsymbol{e}_{m+1}, \cdots, \boldsymbol{e}_n\}$ 是 W 的一组基, 故 $\dim V = \dim U + \dim W$, 从而 $V = U \oplus W$.
□

命题 3.21

若 $V = U \oplus W$ 且 $U = U_1 \oplus U_2$, 求证:$V = U_1 \oplus U_2 \oplus W$.

证明 由 $U = U_1 \oplus U_2$ 可得 $U_1 \cap U_2 = 0$; 由 $V = U \oplus W$ 可得 $(U_1 + U_2) \cap W = U \cap W = 0$, 因此由定理 3.10(2)可得 $U_1 + U_2 + W$ 是直和, 从而 $V = U_1 + U_2 + W = U_1 \oplus U_2 \oplus W$. □

命题 3.22

每一个 n 维线性空间均可表示为 n 个一维子空间的直和.

证明 设 V 是 n 维线性空间, 取其一组基为 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$. 设 $V_i = L(\boldsymbol{e}_i)(1 \leq i \leq n)$, 则 V_i 是 V 的一维子空间. 任取 $\alpha \in V$, 存在唯一一组常数 $k_1, k_2, \cdots, k_n$, 使得 $\alpha = k_1\boldsymbol{e}_1 + k_2\boldsymbol{e}_2 + \cdots + k_n\boldsymbol{e}_n$, 而 $k_i\boldsymbol{e}_i \in V_i, i = 1, 2, \cdots, n$. 因此 $V = V_1 + V_2 + \cdots + V_n$. 注意到 $\dim V = n = \dim V_1 + \dim V_2 + \cdots + \dim V_n$, 故由定理 3.10(3)可知,$V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$.

(注意到 V_i 的基是 $\{\boldsymbol{e}_i\}$, 因此 $V_i(1 \leq i \leq n)$ 的基能拼成 V 的基, 故由定理 3.10(4)也可得到结论. 再注意到 V 中任一向量写成基向量 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 的线性组合时, 其表示是唯一的. 这就是说,V 中任一向量写成 V_i 中的向量之和时, 其表示是唯一的, 故由定理 3.10(5)同样可得结论.) □

命题 3.23

设 V_0 是数域 $\mathbb{F}$ 上 n 维向量空间 V 的真子空间, 则 V_0 至多包含 $n-1$ 个 V 中的基向量.

证明 反证法,若 V_0 包含 n 个 V 中的基向量,则 V_i 就包含了 V 的一组基. 不妨设 V_0 中的这组基向量为 $\{e_1, e_2, \cdots e_n\}$, 则 $\forall \alpha \in V$, 有 $\alpha = k_1e_1 + k_2e_2 + \cdots + k_ne_n \in V_0$, 其中 $k_i \in \mathbb{F}, i = 1, 2, \cdots, n$. 故 $V_0 \supset V$, 又 $V_0 \subset V$, 因此 $V_0 = V$. 这与 V_0 是 V 的真子空间矛盾. □

命题 3.24

设 $V_1, V_2, \cdots, V_m$ 是数域 $\mathbb{F}$ 上向量空间 V 的 m 个真子空间, 证明: 在 V 中必存在一个向量 $\boldsymbol{\alpha}$, 它不属于任何一个 V_i.

笔记 这个命题表明: **有限个真子空间不能覆盖全空间**.

证明 证法一: 对个数 m 进行归纳, 当 $m = 1$ 时结论显然成立. 设 $m = k$ 时结论成立, 现要证明 $m = k + 1$ 时结论也成立. 由归纳假设, 存在向量 $\boldsymbol{\alpha}$, 它不属于任何一个 $V_i(1 \leq i \leq k)$. 若 $\boldsymbol{\alpha}$ 也不属于 V_{k+1}, 则结论已成立, 因此可设 $\boldsymbol{\alpha} \in V_{k+1}$. 在 V_{k+1} 外选一个向量 $\boldsymbol{\beta}$, 作集合

$$M = \{t\boldsymbol{\alpha} + \boldsymbol{\beta} | t \in \mathbb{F}\}.$$

事实上, 我们可将 M 看成是通过 $\boldsymbol{\beta}$ 的终点且平行于 $\boldsymbol{\alpha}$ 的一根 "直线", 现要证明它和每个 V_i 最多只有一个交点. 首先,M 和 V_{k+1} 无交点, 因为若 $t\boldsymbol{\alpha}+\boldsymbol{\beta}\in V_{k+1}$, 则从 $t\boldsymbol{\alpha}\in V_{k+1}$ 可推出 $\boldsymbol{\beta}\in V_{k+1}$, 与假设矛盾. 又若对某个 $V_i(i<k+1)$, 存在 $t_1\neq t_2$, 使得 $t_1\boldsymbol{\alpha}+\boldsymbol{\beta}\in V_i$,$t_2\boldsymbol{\alpha}+\boldsymbol{\beta}\in V_i$, 则 $(t_1-t_2)\boldsymbol{\alpha}\in V_i$, 从而导致 $\boldsymbol{\alpha}\in V_i$, 与假设矛盾. 因此,$M$ 和每个 V_i 最多只有一个交点, 从而 M 中只有有限个向量属于 V_i 的并集, 而 t 有无穷多个选择, 由此即得结论.

证法二:任取 V 的一组基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$. 对任意的正整数 k, 构造 V 中向量 $\boldsymbol{\alpha}_k=\boldsymbol{e}_1+k\boldsymbol{e}_2+\cdots+k^{n-1}\boldsymbol{e}_n$, 设向量族 $S=\{\boldsymbol{\alpha}_k|k=1,2,\cdots\}$. 由例题 3.9可知,$S$ 中任意 n 个不同的向量都构成 V 的一组基. 因为 V_i 都是 V 的真子空间, 所以每个 V_i 至多包含 S 中 $n-1$ 个向量. 因此 $\bigcup\limits_{i=1}^{m}V_i$ 至多包含 S 中 $m(n-1)$ 个向量. 又由于 S 是无限集合, 故存在某个向量 $\boldsymbol{\alpha}_k$, 使得 $\boldsymbol{\alpha}_k$ 不属于任何一个 V_i. □

注 上述证明要用到任意一个数域都有无穷个元素这一事实. 因此, 对于有限域 (读者以后可能会学到) 上的向量空间, 上例结论不一定成立.

命题 3.25

设 $V_1,V_2,\cdots,V_m$ 是数域 $\mathbb{F}$ 上向量空间 V 的 m 个真子空间, 证明:V 中必有一组基, 使得每个基向量都不在诸 V_i 的并中.

证明 证法一: 由命题 3.24可知, 存在非零向量 $\boldsymbol{e}_1\in V$, 使得 $\boldsymbol{e}_1\notin\bigcup\limits_{i=1}^{m}V_i$. 定义 $V_{m+1}=L(\boldsymbol{e}_1)$, 再由命题 3.24可知, 存在向量 $\boldsymbol{e}_2\in V$, 使得 $\boldsymbol{e}_2\notin\bigcup\limits_{i=1}^{m+1}V_i$. 由推论 3.2可知,$\boldsymbol{e}_2\notin L(\boldsymbol{e}_1)$ 意味着 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 线性无关. 重新定义 $V_{m+1}=L(\boldsymbol{e}_1,\boldsymbol{e}_2)$, 再由命题 3.24可知, 存在向量 $\boldsymbol{e}_3\in V$, 使得 $\boldsymbol{e}_3\notin\bigcup\limits_{i=1}^{m+1}V_i$. 再由推论 3.2可知,$\boldsymbol{e}_3\notin L(\boldsymbol{e}_1,\boldsymbol{e}_2)$ 意味着 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 线性无关. 不断重复上述讨论, 即添加线性无关的向量重新定义 V_{m+1}, 并反复利用命题 3.24和推论 3.2的结论, 最后可以得到 n 个线性无关的向量 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$, 它们构成 V 的一组基, 且满足 $\boldsymbol{e}_j\notin\bigcup\limits_{i=1}^{m}V_i(1\leq j\leq n)$.

证法二:任取 V 的一组基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$. 对任意的正整数 k, 构造 V 中向量 $\boldsymbol{\alpha}_k=\boldsymbol{e}_1+k\boldsymbol{e}_2+\cdots+k^{n-1}\boldsymbol{e}_n$, 设向量族 $S=\{\boldsymbol{\alpha}_k|k=1,2,\cdots\}$. 由例题 3.9可知,$S$ 中任意 n 个不同的向量都构成 V 的一组基. 因为 V_i 都是 V 的真子空间, 所以每个 V_i 至多包含 S 中 $n-1$ 个向量. 因此 $\bigcup\limits_{i=1}^{m}V_i$ 至多包含 S 中 $m(n-1)$ 个向量. 又由于 S 是无限集合, 故存在某个向量 $\boldsymbol{\alpha}_k$, 使得 $\boldsymbol{\alpha}_k$ 不属于任何一个 V_i. 进一步, 在 S 中一定还存在 n 个不同的向量 $\boldsymbol{\alpha}_{k_1},\boldsymbol{\alpha}_{k_2},\cdots,\boldsymbol{\alpha}_{k_n}$, 使得每个 $\boldsymbol{\alpha}_{k_j}$ 都不属于任何一个 V_i, 此时 $\{\boldsymbol{\alpha}_{k_1},\boldsymbol{\alpha}_{k_2},\cdots,\boldsymbol{\alpha}_{k_n}\}$ 就构成了 V 的一组基. □

定义 3.3 ($U-$ 陪集与商空间)

设 V 是数域 $\mathbb{K}$ 上的线性空间,U 是 V 的子空间. 对任意的 $\boldsymbol{v}\in V$, 集合 $\boldsymbol{v}+U:=\{\boldsymbol{v}+\boldsymbol{u}|\boldsymbol{u}\in U\}$ 称为 $\boldsymbol{v}$ 的**$U-$陪集**. 在所有 $U-$ 陪集构成的集合 $S=\{\boldsymbol{v}+U|\boldsymbol{v}\in V\}$ 中, 定义加法和数乘如下, 其中 $\boldsymbol{v}_1,\boldsymbol{v}_2\in V,k\in\mathbb{K}$:

$$(\boldsymbol{v}_1+U)+(\boldsymbol{v}_2+U):=(\boldsymbol{v}_1+\boldsymbol{v}_2)+U,\ k\cdot(\boldsymbol{v}_1+U):=k\cdot\boldsymbol{v}_1+U.$$

S 在上述加法和数乘下成为数域 $\mathbb{K}$ 上的线性空间, 称为 V 关于子空间 U 的**商空间**, 记为 V/U.

笔记 容易验证 S 在上述加法和数乘下满足线性空间的 8 条公理, 因此商空间是良定义的. 故任意 V 的子空间 U 都存在相应的商空间.

注 商空间的向量是 $U-$ 陪集. 商空间的零向量就是 $\boldsymbol{0}+U=U$.

命题 3.26 ($U-$ 陪集的性质)

(1) $U-$ 陪集之间的关系是: 作为集合或者相等, 或者不相交;

(2) $\boldsymbol{v}_1+U=\boldsymbol{v}_2+U$(作为集合相等) 当且仅当 $\boldsymbol{v}_1-\boldsymbol{v}_2\in U$. 特别地,$\boldsymbol{v}+U$ 是 V 的子空间当且仅当 $\boldsymbol{v}\in U$;

(3) S 中的加法以及 $\mathbb{K}$ 关于 S 的数乘不依赖于代表元的选取, 即若 $\boldsymbol{v}_1+U=\boldsymbol{v}_1'+U$ 以及 $\boldsymbol{v}_2+U=\boldsymbol{v}_2'+U$,

则 $(\boldsymbol{v}_1+U)+(\boldsymbol{v}_2+U)=(\boldsymbol{v}_1'+U)+(\boldsymbol{v}_2'+U)$, 以及 $k\cdot(\boldsymbol{v}_1+U)=k\cdot(\boldsymbol{v}_1'+U)$;

证明

(1) 设 $(\boldsymbol{v}_1+U)\cap(\boldsymbol{v}_2+U)\neq\varnothing$, 即存在 $\boldsymbol{u}_1,\boldsymbol{u}_2\in U$, 使得 $\boldsymbol{v}_1+\boldsymbol{u}_1=\boldsymbol{v}_2+\boldsymbol{u}_2$, 从而 $\boldsymbol{v}_1-\boldsymbol{v}_2=\boldsymbol{u}_2-\boldsymbol{u}_1\in U$, 于是

$$\boldsymbol{v}_1+U=\boldsymbol{v}_2+(\boldsymbol{v}_1-\boldsymbol{v}_2)+U\subseteq\boldsymbol{v}_2+U,\ \boldsymbol{v}_2+U=\boldsymbol{v}_1+(\boldsymbol{v}_2-\boldsymbol{v}_1)+U\subseteq\boldsymbol{v}_1+U,$$

因此 $\boldsymbol{v}_1+U=\boldsymbol{v}_2+U$.

(2) 由 (1) 的证明过程即得. 特别地,$\boldsymbol{v}+U$ 是 V 的子空间 $\Rightarrow\boldsymbol{0}\in\boldsymbol{v}+U\Rightarrow$ 存在$\boldsymbol{u}\in U$, 使得$\boldsymbol{0}=\boldsymbol{v}+\boldsymbol{u}\Rightarrow\boldsymbol{v}=-\boldsymbol{u}\in U$. 若 $\boldsymbol{v}\in U$, 则一方面,$\forall\boldsymbol{\alpha}\in\boldsymbol{v}+U$, 存在 $\boldsymbol{u}'\in U$, 使得 $\boldsymbol{\alpha}=\boldsymbol{v}+\boldsymbol{u}'$. 又 $\boldsymbol{v}\in U$, 因此 $\boldsymbol{\alpha}=\boldsymbol{v}+\boldsymbol{u}'\in U$. 故 $\boldsymbol{v}+U\subset U$. 另一方面,$\forall\boldsymbol{\beta}\in U$, 有 $\boldsymbol{\beta}=\boldsymbol{v}+\boldsymbol{\beta}-\boldsymbol{v}$. 又由 $\boldsymbol{v}\in U$ 可知 $\boldsymbol{\beta}-\boldsymbol{v}\in U$, 于是 $\boldsymbol{\beta}=\boldsymbol{v}+\boldsymbol{\beta}-\boldsymbol{v}\in\boldsymbol{v}+U$. 故 $\boldsymbol{v}+U\supset U$. 因此 $\boldsymbol{v}+U=U$ 是 V 的子空间.

(实际上, 若 $\boldsymbol{v}\in U$, 则因为 $\boldsymbol{v}\in U$并且$\boldsymbol{v}\in\boldsymbol{v}+U$, 所以 $\boldsymbol{v}+U\cap U\neq\varnothing$. 故由 (1) 可知 $\boldsymbol{v}+U=U$ 是 V 的子空间. 这样也能得到证明.)

(3) 若 $\boldsymbol{v}_1+U=\boldsymbol{v}_1'+U$ 以及 $\boldsymbol{v}_2+U=\boldsymbol{v}_2'+U$, 则存在 $\boldsymbol{u}_1,\boldsymbol{u}_2\in U$, 使得 $\boldsymbol{v}_1-\boldsymbol{v}_1'=\boldsymbol{u}_1,\boldsymbol{v}_2-\boldsymbol{v}_2'=\boldsymbol{u}_2$, 从而 $(\boldsymbol{v}_1+\boldsymbol{v}_2)-(\boldsymbol{v}_1'+\boldsymbol{v}_2')=\boldsymbol{u}_1+\boldsymbol{u}_2\in U,k\cdot\boldsymbol{v}_1-k\cdot\boldsymbol{v}_1'=k\cdot\boldsymbol{u}_1\in U$, 于是由 (2) 可得

$$(\boldsymbol{v}_1+U)+(\boldsymbol{v}_2+U)=(\boldsymbol{v}_1+\boldsymbol{v}_2)+U=(\boldsymbol{v}_1'+\boldsymbol{v}_2')+U=(\boldsymbol{v}_1'+U)+(\boldsymbol{v}_2'+U),$$

$$k\cdot(\boldsymbol{v}_1+U)=k\cdot\boldsymbol{v}_1+U=k\cdot\boldsymbol{v}_1'+U=k\cdot(\boldsymbol{v}_1'+U).$$

□

注 若 $\boldsymbol{v}_1+U=\boldsymbol{v}_1'+U$ 以及 $\boldsymbol{v}_2+U=\boldsymbol{v}_2'+U$, 则 $\forall\boldsymbol{u}_1'\in U$, 有 $\boldsymbol{v}_1+\boldsymbol{u}_1'\in\boldsymbol{v}_1+U=\boldsymbol{v}_1'+U$. 从而存在 $\boldsymbol{u}_1''\in U$, 使得 $\boldsymbol{v}_1+\boldsymbol{u}_1'=\boldsymbol{v}_1'+\boldsymbol{u}_1''$. 于是 $\boldsymbol{v}_1-\boldsymbol{v}_1'=\boldsymbol{u}_1''-\boldsymbol{u}_1'$. 再令 $\boldsymbol{u}_1=\boldsymbol{u}_1''-\boldsymbol{u}_1'$, 则 $\boldsymbol{v}_1-\boldsymbol{v}_1'=\boldsymbol{u}_1\in U$. 同理可得, 存在 $u_2\in U$, 使得 $\boldsymbol{v}_2-\boldsymbol{v}_2'=\boldsymbol{u}_2\in U$.

命题 3.27 (商空间的维数公式和商空间与补空间同构)

设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,U 是 V 的子空间,W 是 U 的补空间, 证明:$\dim V/U=\dim V-\dim U$, 并且存在线性同构 $\varphi:W\to V/U$.

证明 取子空间 U 的一组基 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_m\}$, 补空间 W 的一组基 $\{\boldsymbol{e}_{m+1},\cdots,\boldsymbol{e}_n\}$, 则 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_m,\boldsymbol{e}_{m+1},\cdots,\boldsymbol{e}_n\}$ 是 V 的一组基. 我们断言 $\{\boldsymbol{e}_{m+1}+U,\cdots,\boldsymbol{e}_n+U\}$ 是商空间 V/U 的一组基. 一方面, 对任意的 $\boldsymbol{v}\in V$, 设 $\boldsymbol{v}=\sum\limits_{i=1}^n a_i\boldsymbol{e}_i$, 则

$$\boldsymbol{v}+U=\left(\sum_{i=1}^n a_i\boldsymbol{e}_i\right)+U=\left(\sum_{i=m+1}^n a_i\boldsymbol{e}_i\right)+U=\sum_{i=m+1}^n a_i(\boldsymbol{e}_i+U).$$

另一方面, 设 $a_{m+1},\cdots,a_n\in\mathbb{K}$, 使得 $\sum\limits_{i=m+1}^n a_i(\boldsymbol{e}_i+U)=\boldsymbol{0}+U$, 即 $\left(\sum\limits_{i=m+1}^n a_i\boldsymbol{e}_i\right)+U=U$, 从而 $\sum\limits_{i=m+1}^n a_i\boldsymbol{e}_i\in U$. 于是存在 $a_1,\cdots,a_m\in\mathbb{K}$, 使得 $\sum\limits_{i=m+1}^n a_i\boldsymbol{e}_i=-\sum\limits_{i=1}^m a_i\boldsymbol{e}_i$, 即 $\sum\limits_{i=1}^n a_i\boldsymbol{e}_i=\boldsymbol{0}$, 从而 $a_i=0(1\leq i\leq n)$. 于是 $\{\boldsymbol{e}_{m+1}+U,\cdots,\boldsymbol{e}_n+U\}$ 线性无关. 因此,$\dim V/U=n-m=\dim V-\dim U$.

对任意的 $\boldsymbol{w}\in W$, 设 $\boldsymbol{w}=\sum\limits_{i=m+1}^n a_i\boldsymbol{e}_i$, 定义映射 $\varphi:W\to V/U$ 为

$$\varphi(\boldsymbol{w})=\boldsymbol{w}+U=\sum_{i=m+1}^n a_i(\boldsymbol{e}_i+U).$$

容易验证 φ 保持加法和数乘, 并且是一一对应 (W 的基 e_i 映射过去得到 $\varphi(e_i)$ 仍是 V/U 的基,$i=m+1,\cdots,n$.), 从而是线性同构. □

3.5.2 练习

练习 3.3 设 $V = M_n(\mathbb{K})$ 是数域 $\mathbb{K}$ 上的 n 阶矩阵全体组成的线性空间,$\boldsymbol{A} \in V$, 求证: 与 $\boldsymbol{A}$ 乘法可交换的矩阵全体 $C(\boldsymbol{A})$ 组成 V 的子空间且其维数不为零. 又若 T 是 V 的非空子集, 求证: 与 T 中任一矩阵乘法可交换的矩阵全体 $C(T)$ 也构成 V 的子空间且其维数不为零.

证明 由于纯量阵 $c\boldsymbol{I}_n$ 与任一 n 阶矩阵 $\boldsymbol{A}$ 乘法可交换, 故 $L(\boldsymbol{I}_n) \subseteq C(\boldsymbol{A})$. 任取 $\boldsymbol{B},\boldsymbol{C} \in C(\boldsymbol{A}),k \in \mathbb{K}$, 容易验证 $\boldsymbol{B}+\boldsymbol{C} \in C(\boldsymbol{A}),k\boldsymbol{B} \in C(\boldsymbol{A})$, 故 $C(\boldsymbol{A})$ 是 $M_n(\mathbb{K})$ 的子空间且其维数不为零. $C(T)$ 的结论同理可证. □

练习 3.4 设 $\boldsymbol{\alpha}_1 = (1,0,-1,0),\boldsymbol{\alpha}_2 = (0,1,2,1),\boldsymbol{\alpha}_3 = (2,1,0,1)$ 是四维实行向量空间 V 中的向量, 它们生成的子空间为 V_1, 又向量 $\boldsymbol{\beta}_1 = (-1,1,1,1),\boldsymbol{\beta}_2 = (1,-1,-3,-1),\boldsymbol{\beta}_3 = (-1,1,-1,1)$ 生成的子空间为 V_2, 求子空间 V_1+V_2 和 $V_1 \cap V_2$ 的基.

解 解法一:V_1+V_2 是由 $\boldsymbol{\alpha}_i$ 和 $\boldsymbol{\beta}_i$ 生成的, 因此只要求出这 6 个向量的极大无关组即可. 将这 6 个向量按列分块方式拼成矩阵, 并用初等行变换将其化为阶梯形矩阵:

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

故可取 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1$ 为 V_1+V_2 的基 (不唯一).

再来求 $V_1 \cap V_2$ 的基. 首先注意到 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是 V_1 的基 (从上面的矩阵即可看出), 又不难验证 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是 V_2 的基,V_2 中的向量可以表示为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 的线性组合. 假设 $t_1\boldsymbol{\beta}_1+t_2\boldsymbol{\beta}_2$ 属于 V_1, 则向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,t_1\boldsymbol{\beta}_1+t_2\boldsymbol{\beta}_2$ 和向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 的秩相等 (因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是 V_1 的基). 因此, 我们可以用矩阵方法来求出参数 t_1,t_2. 注意到

$$\begin{pmatrix} 1 & 0 & -t_1+t_2 \\ 0 & 1 & t_1-t_2 \\ -1 & 2 & t_1-3t_2 \\ 0 & 1 & t_1-t_2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -t_1+t_2 \\ 0 & 1 & t_1-t_2 \\ 0 & 2 & -2t_2 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -t_1+t_2 \\ 0 & 1 & t_1-t_2 \\ 0 & 0 & -2t_1 \\ 0 & 0 & 0 \end{pmatrix},$$

故可得 $t_1 = 0$, 所以 $V_1 \cap V_2$ 的基可取为 $\boldsymbol{\beta}_2$.

解法二: 求 V_1+V_2 的基同解法 1, 现用解线性方程组的方法来求 $V_1 \cap V_2$ 的基. 因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是 V_1 的基,$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是 V_2 的基, 故对任一向量 $\boldsymbol{\gamma} \in V_1 \cap V_2,\boldsymbol{\gamma} = x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2 = (-x_3)\boldsymbol{\beta}_1+(-x_4)\boldsymbol{\beta}_2$. 因此, 求向量 $\boldsymbol{\gamma}$ 等价于求解线性方程组

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\beta}_1+x_4\boldsymbol{\beta}_2 = \boldsymbol{0}.$$

通过初等行变换将其系数矩阵 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)$ 进行化简:

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

故上述线性方程组的通解为 $(x_1,x_2,x_3,x_4) = k(-1,1,0,1)$, 从而 $\boldsymbol{\gamma} = -k(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2) = -k\boldsymbol{\beta}_2(k \in \mathbb{R})$, 于是 $\boldsymbol{\beta}_2$ 是 $V_1 \cap V_2$ 的基. □

3.6 矩阵的秩

3.6.1 初等变换法

矩阵的秩在初等变换或分块初等变换下不变.

想法: 遇到关于秩不等式的问题, 可以考虑构造分块矩阵, 对其做适当的初等变换, 再利用秩的基本公式.

定理 3.12

矩阵 $\boldsymbol{A}$ 的秩等于 r 的充要条件是 $\boldsymbol{A}$ 有一个 r 阶子式不等于零, 而 $\boldsymbol{A}$ 的所有 $r+1$ 阶子式都等于零.

命题 3.28 (矩阵秩的基本公式)

(1) 若 $k\neq 0$,$\mathrm{r}(k\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})$;

(2) $\mathrm{r}(\boldsymbol{AB})\leq\min\{\mathrm{r}(\boldsymbol{A}),\mathrm{r}(\boldsymbol{B})\}$;

(3) $\mathrm{r}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}=\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$;

(4) $\mathrm{r}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{C}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}\geq\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$,$\mathrm{r}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{D}&\boldsymbol{B}\end{pmatrix}\geq\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$;

(5) $\mathrm{r}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{B}\end{pmatrix}\leq\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$,$\mathrm{r}\begin{pmatrix}\boldsymbol{A}\\\boldsymbol{B}\end{pmatrix}\leq\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$;

(6) $\mathrm{r}(\boldsymbol{A}+\boldsymbol{B})\leq\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$,$\mathrm{r}(\boldsymbol{A}-\boldsymbol{B})\leq\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$;

(7) $\mathrm{r}(\boldsymbol{A}-\boldsymbol{B})\geq|\mathrm{r}(\boldsymbol{A})-\mathrm{r}(\boldsymbol{B})|$.

证明

(1) 由于 $k\boldsymbol{A}=\boldsymbol{P}_1(k)\boldsymbol{P}_2(k)\cdots\boldsymbol{P}_m(k)\boldsymbol{A}$, 故 $\mathrm{r}(k\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})$.

(2) 证法一:设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵,$\boldsymbol{B}$ 是 $n\times s$ 矩阵. 将矩阵 $\boldsymbol{B}$ 按列分块,$\boldsymbol{B}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s)$, 则 $\boldsymbol{AB}=(\boldsymbol{A\beta}_1,\boldsymbol{A\beta}_2,\cdots,\boldsymbol{A\beta}_s)$. 若 $\boldsymbol{B}$ 列向量的极大无关组为 $\{\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_r}\}$, 则 $\boldsymbol{B}$ 的任一列向量 $\boldsymbol{\beta}_j$ 均可用 $\{\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_r}\}$ 线性表示. 于是任一 $\boldsymbol{A\beta}_j$ 也可用 $\{\boldsymbol{A\beta}_{j_1},\boldsymbol{A\beta}_{j_2},\cdots,\boldsymbol{A\beta}_{j_r}\}$ 来线性表示. 因此, 向量组 $\{\boldsymbol{A\beta}_1,\boldsymbol{A\beta}_2,\cdots,\boldsymbol{A\beta}_s\}$ 的秩不超过 r, 即 $\mathrm{r}(\boldsymbol{AB})\leq\mathrm{r}(\boldsymbol{B})$. 同理, 对矩阵 $\boldsymbol{A}$ 用行分块的方法可以证明 $\mathrm{r}(\boldsymbol{AB})\leq\mathrm{r}(\boldsymbol{A})$.

证法二:见例题 4.15.

(3) 设 $\boldsymbol{A},\boldsymbol{B}$ 的秩分别为 r_1,r_2, 则存在非异阵 $\boldsymbol{P}_1,\boldsymbol{Q}_1$ 和非异阵 $\boldsymbol{P}_2,\boldsymbol{Q}_2$, 使得

$$\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1=\begin{pmatrix}\boldsymbol{I}_{r_1}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix},\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2=\begin{pmatrix}\boldsymbol{I}_{r_2}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}.$$

于是

$$\begin{pmatrix}\boldsymbol{P}_1&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{P}_2\end{pmatrix}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{Q}_1&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{Q}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_{r_1}&\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{I}_{r_2}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}.$$

因此,$\mathrm{r}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}=r_1+r_2=\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$.

(4) 证法一: 我们只证明第一个不等式, 第二个不等式同理可证. 设 $\boldsymbol{A},\boldsymbol{B}$ 的秩分别为 r_1,r_2, 则存在非异阵 $\boldsymbol{P}_1,\boldsymbol{Q}_1$ 和非异阵 $\boldsymbol{P}_2,\boldsymbol{Q}_2$, 使得

$$\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1=\begin{pmatrix}\boldsymbol{I}_{r_1}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix},\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2=\begin{pmatrix}\boldsymbol{I}_{r_2}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}.$$

于是

$$\begin{pmatrix}\boldsymbol{P}_1&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{P}_2\end{pmatrix}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{C}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{Q}_1&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{Q}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1&\boldsymbol{P}_1\boldsymbol{C}\boldsymbol{Q}_2\\\boldsymbol{O}&\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_{r_1}&\boldsymbol{O}&\boldsymbol{C}_{11}&\boldsymbol{C}_{12}\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{C}_{21}&\boldsymbol{C}_{22}\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{I}_{r_2}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}.$$

在上面的分块矩阵中实施第三类分块初等变换, 用 $\boldsymbol{I}_{r_1}$ 消去同行的矩阵; 用 $\boldsymbol{I}_{r_2}$ 消去同列的矩阵, 再将 $\boldsymbol{C}_{22}$ 对

换到第 (2,2) 位置:

$$\begin{pmatrix} I_{r_1} & O & C_{11} & C_{12} \\ O & O & C_{21} & C_{22} \\ O & O & I_{r_2} & O \\ O & O & O & O \end{pmatrix} \to \begin{pmatrix} I_{r_1} & O & O & O \\ O & O & O & C_{22} \\ O & O & I_{r_2} & O \\ O & O & O & O \end{pmatrix} \to \begin{pmatrix} I_{r_1} & O & O & O \\ O & C_{22} & O & O \\ O & O & I_{r_2} & O \\ O & O & O & O \end{pmatrix},$$

最后由 (3) 的结论可得

$$\mathrm{r}\begin{pmatrix} A & C \\ O & B \end{pmatrix} = \mathrm{r}(I_{r_1}) + \mathrm{r}(C_{22}) + \mathrm{r}(I_{r_2}) \geq r_1 + r_2 = \mathrm{r}(A) + \mathrm{r}(B).$$

证法二: 我们也可用子式法来证明. 设 $\mathrm{r}\begin{pmatrix} A & O \\ O & B \end{pmatrix} = r$, 则由定理 3.12可知, $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ 有一个 r 阶子式不为零, 不妨设为 $\begin{vmatrix} A_1 & O \\ O & B_1 \end{vmatrix}$, 其中 A_1, B_1 分别是 A, B 的子阵. 注意 A_1 或 B_1 允许是零阶矩阵, 这对应于该子式完全包含在 B 或 A 中, 但若 A_1, B_1 的阶数都大于零, 则通过该子式非零, 再结合由 Laplace 定理容易验证 A_1, B_1 都是方阵. 设在矩阵 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 中对应的 r 阶子式是 $\begin{vmatrix} A_1 & C_1 \\ O & B_1 \end{vmatrix}$, 则由 Laplace 定理可得 $\begin{vmatrix} A_1 & C_1 \\ O & B_1 \end{vmatrix} = |A_1||B_1| = \begin{vmatrix} A_1 & O \\ O & B_1 \end{vmatrix} \neq 0$, 再次由定理 3.12可得

$$\mathrm{r}\begin{pmatrix} A & C \\ O & B \end{pmatrix} \geq r = \mathrm{r}\begin{pmatrix} A & O \\ O & B \end{pmatrix} = \mathrm{r}(A) + \mathrm{r}(B).$$

证法三:设 $A = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ 是 A 的列分块,$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 是 A 的列向量的极大无关组; 设 $B = (\beta_1, \beta_2, \cdots, \beta_l)$,$C = (\gamma_1, \gamma_2, \cdots, \gamma_l)$ 是 B, C 的列分块,$\beta_{j_1}, \beta_{j_2}, \cdots, \beta_{j_s}$ 是 B 的列向量的极大无关组, 则 $\mathrm{r}(A) = r$ 且 $\mathrm{r}(B) = s$. 我们接下来证明: 作为 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 的列向量,$\begin{pmatrix} \alpha_{i_1} \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} \alpha_{i_r} \\ 0 \end{pmatrix}, \begin{pmatrix} \gamma_{j_1} \\ \beta_{j_1} \end{pmatrix}, \cdots, \begin{pmatrix} \gamma_{j_s} \\ \beta_{j_s} \end{pmatrix}$ 线性无关. 设

$$c_1\begin{pmatrix} \alpha_{i_1} \\ 0 \end{pmatrix} + \cdots + c_r\begin{pmatrix} \alpha_{i_r} \\ 0 \end{pmatrix} + d_1\begin{pmatrix} \gamma_{j_1} \\ \beta_{j_1} \end{pmatrix} + \cdots + d_s\begin{pmatrix} \gamma_{j_s} \\ \beta_{j_s} \end{pmatrix} = \mathbf{0},$$

即

$$c_1\alpha_{i_1} + \cdots + c_r\alpha_{i_r} + d_1\gamma_{j_1} + \cdots + d_s\gamma_{j_s} = \mathbf{0}, d_1\beta_{j_1} + \cdots + d_s\beta_{j_s} = \mathbf{0}.$$

由上面的假设即得 $c_1 = \cdots = c_r = d_1 = \cdots = d_s = 0$, 于是上述结论得证. 因为 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 的列向量中有 $r+s$ 个线性无关, 故 $\mathrm{r}\begin{pmatrix} A & C \\ O & B \end{pmatrix} \geq r + s = \mathrm{r}(A) + \mathrm{r}(B)$.

(5) 注意到

$$\begin{pmatrix} I & I \end{pmatrix}\begin{pmatrix} A & O \\ O & B \end{pmatrix} = \begin{pmatrix} A & B \end{pmatrix}, \begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} I \\ I \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}.$$

故由 (2) 和 (3) 可得

$$\mathrm{r}\begin{pmatrix} A & B \end{pmatrix} = \mathrm{r}\left(\begin{pmatrix} I & I \end{pmatrix}\begin{pmatrix} A & O \\ O & B \end{pmatrix}\right) \leqslant \mathrm{r}\begin{pmatrix} A & O \\ O & B \end{pmatrix} = \mathrm{r}(A) + \mathrm{r}(B),$$

$$\mathrm{r}\begin{pmatrix} A \\ B \end{pmatrix} = \mathrm{r}\left(\begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} I \\ I \end{pmatrix}\right) \leqslant \mathrm{r}\begin{pmatrix} A & O \\ O & B \end{pmatrix} = \mathrm{r}(A) + \mathrm{r}(B).$$

(6) 注意到

$$\begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} I \\ I \end{pmatrix} = A + B, \begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} I \\ -I \end{pmatrix} = A - B.$$

故由 (2) 和 (5) 可得

$$\mathrm{r}(A+B)=\mathrm{r}\left(\begin{pmatrix}A & B\end{pmatrix}\begin{pmatrix}I\\ I\end{pmatrix}\right)\leqslant \mathrm{r}\begin{pmatrix}A & B\end{pmatrix}\leqslant \mathrm{r}(A)+\mathrm{r}(B),$$

$$\mathrm{r}(A-B)=\mathrm{r}\left(\begin{pmatrix}A & B\end{pmatrix}\begin{pmatrix}I\\ -I\end{pmatrix}\right)\leqslant \mathrm{r}\begin{pmatrix}A & B\end{pmatrix}\leqslant \mathrm{r}(A)+\mathrm{r}(B).$$

(7) 由于 $\mathrm{r}(A-B)=\mathrm{r}(B-A)$, 故不妨设 $\mathrm{r}(A)\geq \mathrm{r}(B)$, 则由 (6) 可得 $\mathrm{r}(A-B)+\mathrm{r}(B)\geq \mathrm{r}(A-B+B)=\mathrm{r}(A)$, 即 $\mathrm{r}(A-B)\geq \mathrm{r}(A)-\mathrm{r}(B)$. □

例题 3.17 设 $A=(a_{ij}), B=(b_{ij})$ 是 $m\times n$ 矩阵, 且 $b_{ij}=(-1)^{i+j}a_{ij}$. 求证:$\mathrm{r}(A)=\mathrm{r}(B)$.

证明 将 A 的第 i 行乘以 $(-1)^i$, 又将第 j 列乘以 $(-1)^j$, 即得矩阵 B, 因此 A 和 B 相抵, 故结论成立. □

命题 3.29 (Sylvester 不等式)

设 A 是 $m\times n$ 矩阵,B 是 $n\times t$ 矩阵, 求证:

$$\mathrm{r}(AB)\geq \mathrm{r}(A)+\mathrm{r}(B)-n.$$

证明 证法一: 考虑下列矩阵的分块初等变换:

$$\begin{pmatrix}I_n & O\\ O & AB\end{pmatrix}\to\begin{pmatrix}I_n & O\\ A & AB\end{pmatrix}\to\begin{pmatrix}I_n & -B\\ A & O\end{pmatrix}\to\begin{pmatrix}B & I_n\\ O & A\end{pmatrix},$$

由矩阵秩的基本公式 (3)和矩阵秩的基本公式 (4)可得

$$\mathrm{r}(AB)+n=\mathrm{r}\begin{pmatrix}I_n & O\\ O & AB\end{pmatrix}=\mathrm{r}\begin{pmatrix}B & I_n\\ O & A\end{pmatrix}\geq \mathrm{r}(A)+\mathrm{r}(B),$$

即 $\mathrm{r}(AB)\geq \mathrm{r}(A)+\mathrm{r}(B)-n$.

证法二:见例题 4.15. □

推论 3.3

若 A 是 $m\times n$ 矩阵,B 是 $n\times t$ 矩阵且 $AB=O$, 则 $\mathrm{r}(A)+\mathrm{r}(B)\leq n$.

命题 3.30 (Sylvester 不等式的推广)

设 $A_1,A_2,\cdots,A_m$ 为 n 阶方阵, 求证:

$$\mathrm{r}(A_1)+\mathrm{r}(A_2)+\cdots+\mathrm{r}(A_m)\leq (m-1)n+\mathrm{r}(A_1A_2\cdots A_m).$$

特别地, 若 $A_1A_2\cdots A_m=O$, 则 $\mathrm{r}(A_1)+\mathrm{r}(A_2)+\cdots+\mathrm{r}(A_m)\leq (m-1)n$.

证明 反复利用Sylvester 不等式可得

$$\begin{aligned}&\mathrm{r}(A_1)+\mathrm{r}(A_2)+\mathrm{r}(A_3)+\cdots+\mathrm{r}(A_m)\\ \leq& n+\mathrm{r}(A_1A_2)+\mathrm{r}(A_3)+\cdots+\mathrm{r}(A_m)\\ \leq& 2n+\mathrm{r}(A_1A_2A_3)+\cdots+\mathrm{r}(A_m)\\ \leq&\cdots\leq (m-1)n+\mathrm{r}(A_1A_2\cdots A_m).\end{aligned}$$

□

例题 3.18 设 A,B 为 n 阶方阵, 满足 $AB=O$. 证明: 若 n 是奇数, 则 $AB'+A'B$ 必为奇异阵; 若 n 为偶数, 举例说明上述结论一般不成立.

证明 由推论 3.3可知,$\mathrm{r}(A)+\mathrm{r}(B)\leq n$. 若 n 为奇数, 则 $\mathrm{r}(A),\mathrm{r}(B)$ 中至少有一个小于等于 $\frac{n}{2}$, 从而小于等于 $\frac{n-1}{2}$. 不妨设 $\mathrm{r}(A)\leq\frac{n-1}{2}$, 于是

$$\mathrm{r}(AB'+A'B)\leq \mathrm{r}(AB')+\mathrm{r}(A'B)\leq \mathrm{r}(A)+\mathrm{r}(A')=2\mathrm{r}(A)\leq n-1,$$

从而 $\boldsymbol{AB}'+\boldsymbol{A}'\boldsymbol{B}$ 为奇异阵. 例如, 当 $n=2$ 时, 令 $\boldsymbol{A}=\boldsymbol{B}=\begin{pmatrix}0&1\\0&0\end{pmatrix}$, 则 $\boldsymbol{AB}=\boldsymbol{O}$, 但 $\boldsymbol{AB}'+\boldsymbol{A}'\boldsymbol{B}=\boldsymbol{I}_2$ 为非异阵. □

命题 3.31 (Frobenius 不等式)

证明:$\mathrm{r}(\boldsymbol{ABC})\geq\mathrm{r}(\boldsymbol{AB})+\mathrm{r}(\boldsymbol{BC})-\mathrm{r}(\boldsymbol{B})$.

证明 证法一: 考虑下列分块初等变换:

$$\begin{pmatrix}\boldsymbol{ABC}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{ABC}&\boldsymbol{AB}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{O}&\boldsymbol{AB}\\-\boldsymbol{BC}&\boldsymbol{B}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{AB}&\boldsymbol{O}\\\boldsymbol{B}&\boldsymbol{BC}\end{pmatrix}.$$

由矩阵秩的基本公式 (3)和矩阵秩的基本公式 (4)可得

$$\mathrm{r}(\boldsymbol{ABC})+\mathrm{r}(\boldsymbol{B})=\mathrm{r}\begin{pmatrix}\boldsymbol{ABC}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}=\mathrm{r}\begin{pmatrix}\boldsymbol{AB}&\boldsymbol{O}\\\boldsymbol{B}&\boldsymbol{BC}\end{pmatrix}\geq\mathrm{r}(\boldsymbol{AB})+\mathrm{r}(\boldsymbol{BC}),$$

由此即得结论.

证法二 (几何方法):将问题转化成几何的语言即为: 设 $\varphi:V_1\to V_2,\psi:V_2\to V_3,\theta:V_3\to V_4$ 是线性映射, 证明:$\mathrm{r}(\theta\psi\varphi)\geq\mathrm{r}(\theta\psi)+\mathrm{r}(\psi\varphi)-\mathrm{r}(\psi)$.

下面考虑通过定义域的限制得到的线性映射. 将 θ 的定义域限制在 $\mathrm{Im}\psi\varphi$ 上可得线性映射 $\theta_1:\mathrm{Im}\psi\varphi\to V_4$, 它的像空间是 $\mathrm{Im}\theta\psi\varphi$, 核空间是 $\mathrm{Ker}\theta\cap\mathrm{Im}\psi\varphi$; 将 θ 的定义域限制在 $\mathrm{Im}\psi$ 上可得线性映射 $\theta_2:\mathrm{Im}\psi\to V_4$, 它的像空间是 $\mathrm{Im}\theta\psi$, 核空间是 $\mathrm{Ker}\theta\cap\mathrm{Im}\psi$, 故由线性映射的维数公式可得

$$\dim(\mathrm{Im}\psi\varphi)=\dim(\mathrm{Ker}\theta\cap\mathrm{Im}\psi\varphi)+\dim(\mathrm{Im}\theta\psi\varphi),\tag{3.5}$$

$$\dim(\mathrm{Im}\psi)=\dim(\mathrm{Ker}\theta\cap\mathrm{Im}\psi)+\dim(\mathrm{Im}\theta\psi).\tag{3.6}$$

注意到 $\mathrm{Im}\psi\varphi\subseteq\mathrm{Im}\psi$, 故 $\dim(\mathrm{Ker}\theta\cap\mathrm{Im}\psi\varphi)\leq\dim(\mathrm{Ker}\theta\cap\mathrm{Im}\psi)$, 从而由(3.5)式和(3.6)式可得

$$\begin{aligned}\dim(\mathrm{Im}\psi\varphi)-\dim(\mathrm{Im}\theta\psi\varphi)&=\dim(\mathrm{Ker}\theta\cap\mathrm{Im}\psi\varphi)\\&\leqslant\dim(\mathrm{Ker}\theta\cap\mathrm{Im}\psi)=\dim(\mathrm{Im}\psi)-\dim(\mathrm{Im}\theta\psi).\end{aligned}$$

于是

$$\begin{aligned}&\dim(\mathrm{Im}\psi\varphi)-\dim(\mathrm{Im}\theta\psi\varphi)\leqslant\dim(\mathrm{Im}\psi)-\dim(\mathrm{Im}\theta\psi)\\&\Leftrightarrow\mathrm{r}(\psi\varphi)-\mathrm{r}(\theta\psi\varphi)\leq\mathrm{r}(\psi)-\mathrm{r}(\theta\psi),\end{aligned}$$

结论得证. □

命题 3.32 (幂等矩阵关于秩的判定准则)

求证:n 阶矩阵 $\boldsymbol{A}$ 是幂等矩阵 (即 $\boldsymbol{A}^2=\boldsymbol{A}$) 的充要条件是:

$$\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{I}_n-\boldsymbol{A})=n.$$

证明 在下列矩阵的分块初等变换中矩阵的秩保持不变:

$$\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{I}-\boldsymbol{A}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A}&\boldsymbol{A}\\\boldsymbol{O}&\boldsymbol{I}-\boldsymbol{A}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A}&\boldsymbol{A}\\\boldsymbol{A}&\boldsymbol{I}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A}-\boldsymbol{A}^2&\boldsymbol{A}\\\boldsymbol{O}&\boldsymbol{I}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A}-\boldsymbol{A}^2&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{I}\end{pmatrix}.$$

因此

$$\mathrm{r}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{I}-\boldsymbol{A}\end{pmatrix}=\mathrm{r}\begin{pmatrix}\boldsymbol{A}-\boldsymbol{A}^2&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{I}\end{pmatrix},$$

即 $\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{I}-\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}-\boldsymbol{A}^2)+n$, 由此即得结论. □

命题 3.33 (对合矩阵关于秩的判定准则)

求证:n 阶矩阵 $\boldsymbol{A}$ 是对合矩阵 (即 $\boldsymbol{A}^2=\boldsymbol{I}_n$) 的充要条件是:

$$\mathrm{r}(\boldsymbol{I}_n+\boldsymbol{A})+\mathrm{r}(\boldsymbol{I}_n-\boldsymbol{A})=n.$$

证明 在下列矩阵的分块初等变换中,矩阵的秩保持不变:

$$\begin{pmatrix}\boldsymbol{I}_n+\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{I}_n-\boldsymbol{A}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{I}_n+\boldsymbol{A} & \boldsymbol{I}_n+\boldsymbol{A}\\ \boldsymbol{O} & \boldsymbol{I}_n-\boldsymbol{A}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{I}_n+\boldsymbol{A} & \boldsymbol{I}_n+\boldsymbol{A}\\ \boldsymbol{I}_n+\boldsymbol{A} & 2\boldsymbol{I}_n\end{pmatrix}\to\begin{pmatrix}\frac{1}{2}(\boldsymbol{I}_n-\boldsymbol{A}^2) & \boldsymbol{I}_n+\boldsymbol{A}\\ \boldsymbol{O} & 2\boldsymbol{I}_n\end{pmatrix}\to\begin{pmatrix}\frac{1}{2}(\boldsymbol{I}_n-\boldsymbol{A}^2) & \boldsymbol{O}\\ \boldsymbol{O} & 2\boldsymbol{I}_n\end{pmatrix}.$$

因此

$$\mathrm{r}\begin{pmatrix}\boldsymbol{I}_n+\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{I}_n-\boldsymbol{A}\end{pmatrix}=\mathrm{r}\begin{pmatrix}\frac{1}{2}(\boldsymbol{I}_n-\boldsymbol{A}^2) & \boldsymbol{O}\\ \boldsymbol{O} & 2\boldsymbol{I}_n\end{pmatrix},$$

即 $\mathrm{r}(\boldsymbol{I}_n+\boldsymbol{A})+\mathrm{r}(\boldsymbol{I}_n-\boldsymbol{A})=\mathrm{r}(\boldsymbol{I}_n-\boldsymbol{A}^2)+n$, 由此即得结论. □

例题 3.19 设 $\boldsymbol{A}$ 是 n 阶矩阵, 求证:$\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{I}_n+\boldsymbol{A})\geq n$.

证明 证法一: 由下列分块初等变换即得结论

$$\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{I}+\boldsymbol{A}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A} & \boldsymbol{A}\\ \boldsymbol{O} & \boldsymbol{I}+\boldsymbol{A}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A} & \boldsymbol{A}\\ -\boldsymbol{A} & \boldsymbol{I}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A}+\boldsymbol{A}^2 & \boldsymbol{A}\\ \boldsymbol{O} & \boldsymbol{I}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A}+\boldsymbol{A}^2 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{I}\end{pmatrix}.$$

证法二: $\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{I}+\boldsymbol{A})=\mathrm{r}(-\boldsymbol{A})+\mathrm{r}(\boldsymbol{I}+\boldsymbol{A})\geq\mathrm{r}(-\boldsymbol{A}+\boldsymbol{I}+\boldsymbol{A})=\mathrm{r}(\boldsymbol{I})=n$. □

命题 3.34 (秩的降阶公式)

设有分块矩阵 $M=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}$, 证明:

(1) 若 $\boldsymbol{A}$ 可逆, 则 $\mathrm{r}(M)=\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})$;

(2) 若 $\boldsymbol{D}$ 可逆, 则 $\mathrm{r}(M)=\mathrm{r}(\boldsymbol{D})+\mathrm{r}(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})$;

(3) 若 $\boldsymbol{A},\boldsymbol{D}$ 都可逆, 则 $\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})=\mathrm{r}(\boldsymbol{D})+\mathrm{r}(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})$.

证明

(1) 由分块初等变换可得

$$\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\end{pmatrix},$$

由此即得结论.

(2) 同理可证明.

(3) 由 (1) 和 (2) 即得.

□

例题 3.20 设

$$M=\begin{pmatrix}a_1^2 & a_1a_2+1 & \cdots & a_1a_n+1\\ a_2a_1+1 & a_2^2 & \cdots & a_2a_n+1\\ \vdots & \vdots & & \vdots\\ a_na_1+1 & a_na_2+1 & \cdots & a_n^2\end{pmatrix},$$

证明:$\mathrm{r}(M)\geq n-1$, 等号成立当且仅当 $|M|=0$.

证明 若 $n=1$, 结论显然成立. 下设 $n\geq 2$. 取 $A=\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}$, 则

$$M=-I_n+\begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix} I_2^{-1}\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}=-I_n+A'I_2^{-1}A=-\left(I_n-A'I_2^{-1}A\right).$$

由秩的降阶公式可得

$$2+\mathrm{r}(M)=2+\mathrm{r}(-M)=\mathrm{r}(I_2)+\mathrm{r}\left(I_n-AI_2^{-1}A'\right)=\mathrm{r}(I_n)+\mathrm{r}\left(I_2-A'I_n^{-1}A\right)=n+\mathrm{r}\begin{pmatrix} \sum\limits_{i=1}^{n} a_i^2-1 & \sum\limits_{i=1}^{n} a_i \\ \sum\limits_{i=1}^{n} a_i & n-1 \end{pmatrix}.$$

而 $\mathrm{r}\begin{pmatrix} \sum\limits_{i=1}^{n} a_i^2-1 & \sum\limits_{i=1}^{n} a_i \\ \sum\limits_{i=1}^{n} a_i & n-1 \end{pmatrix}\geqslant 1$, 于是 $\mathrm{r}(M)=n-2+\mathrm{r}\begin{pmatrix} \sum\limits_{i=1}^{n} a_i^2-1 & \sum\limits_{i=1}^{n} a_i \\ \sum\limits_{i=1}^{n} a_i & n-1 \end{pmatrix}\geqslant n-1.$, 等号成立当且仅当 M 不满秩, 即 $|M|=0$. □

命题 3.35

设 A,B 都是数域 $\mathbb{K}$ 上的 n 阶矩阵且 $AB=BA$, 证明:

$$\mathrm{r}(A+B)\leq \mathrm{r}(A)+\mathrm{r}(B)-\mathrm{r}(AB).$$

♠

注 证法一:这里乘的不只是初等变换矩阵. 记住这个分块矩阵乘法和构造.

证法二:和证法三:思路分析: 将秩不等式转化为维数公式就能自然得到证明的想法.

证明 证法一: 考虑如下分块矩阵的乘法:

$$\begin{pmatrix} I & I \\ O & I \end{pmatrix}\begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} I & -B \\ I & A \end{pmatrix}=\begin{pmatrix} A+B & -AB+BA \\ B & BA \end{pmatrix}=\begin{pmatrix} A+B & O \\ B & AB \end{pmatrix}.$$

由矩阵秩的基本公式 (2)和矩阵秩的基本公式 (4)可得

$$\mathrm{r}(A)+\mathrm{r}(B)=\mathrm{r}\begin{pmatrix} A & O \\ O & B \end{pmatrix}\geq \mathrm{r}\begin{pmatrix} A+B & O \\ B & BA \end{pmatrix}\geq \mathrm{r}(A+B)+\mathrm{r}(AB),$$

由此即得结论.

证法二:设 V_A 是方程组 $Ax=0$ 的解空间,V_B,V_{AB},V_{A+B} 的意义同理. 若列向量 $\alpha\in V_A\cap V_B$, 即 α 满足 $A\alpha=0$ 且 $B\alpha=0$, 于是 $(A+B)\alpha=0$, 即 $\alpha\in V_{A+B}$, 从而 $V_A\cap V_B\subseteq V_{A+B}$. 同理可证 $V_A\subseteq V_{BA}$,$V_B\subseteq V_{AB}$. 因为 $AB=BA$, 所以 $V_{BA}=V_{AB}$, 从而 $V_A+V_B\subseteq V_{AB}$. 因此, 我们有

$$\dim(V_A\cap V_B)\leq \dim V_{A+B}=n-\mathrm{r}(A+B),\quad \dim(V_A+V_B)\leq \dim V_{AB}=n-\mathrm{r}(AB).$$

将上面两个不等式相加, 再由交和空间维数公式可得

$$\begin{aligned} n-\mathrm{r}(A+B)+n-\mathrm{r}(AB)&\geq \dim(V_A\cap V_B)+\dim(V_A+V_B) \\ &=\dim V_A+\dim V_B=n-\mathrm{r}(A)+n-\mathrm{r}(B), \end{aligned}$$

因此 $\mathrm{r}(A+B)+\mathrm{r}(AB)\leq \mathrm{r}(A)+\mathrm{r}(B)$, 结论得证.

证法三:设 $A=(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 为 A 的列分块,$B=(\beta_1,\beta_2,\cdots,\beta_n)$ 为 B 的列分块. 记 $U_A=L(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 为 A 的列向量生成的 $\mathbb{K}^n$ 的子空间,U_B,U_{AB},U_{A+B} 的意义同理. 因为向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的极大无关组就是 $L(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 的一组基, 故 $\mathrm{r}(A)=\dim U_A$, 关于 $B,AB,A+B$ 的等式同理可得. 显然, 我们有 $U_{A+B}\subseteq U_A+U_B$. 注意到 $AB=(A\beta_1,A\beta_2,\cdots,A\beta_n)$, 若设 $\beta_j=(b_{1j},b_{2j},\cdots,b_{nj})'$, 则 AB 的列向量 $A\beta_j=b_{1j}\alpha_1+b_{2j}\alpha_2+\cdots+b_{nj}\alpha_n\in U_A$, 从而 $U_{AB}\subseteq U_A$. 同理可得 $U_{BA}\subseteq U_B$. 又因为 $AB=BA$, 故 $U_{AB}\subseteq U_A\cap U_B$. 最后, 由上述包含关系

以及交和空间维数公式可得

$$\begin{aligned}\mathrm{r}(\boldsymbol{A}+\boldsymbol{B})+\mathrm{r}(\boldsymbol{A}\boldsymbol{B}) &= \dim U_{\boldsymbol{A}+\boldsymbol{B}}+\dim U_{\boldsymbol{A}\boldsymbol{B}} \le \dim(U_{\boldsymbol{A}}+U_{\boldsymbol{B}})+\dim(U_{\boldsymbol{A}}\cap U_{\boldsymbol{B}})\\ &= \dim U_{\boldsymbol{A}}+\dim U_{\boldsymbol{B}}=\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B}).\end{aligned}$$

□

3.6.2 利用线性方程组的求解理论讨论矩阵的秩

定理 3.13

设 $\boldsymbol{A}$ 是数域 $\mathbb{K}$ 上的 $m\times n$ 矩阵, 则齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解集 $V_{\boldsymbol{A}}$ 是 n 维列向量空间 $\mathbb{K}^n$ 的子空间. 根据线性方程组的求解理论, 我们有

$$\dim V_{\boldsymbol{A}}+\mathrm{r}(\boldsymbol{A})=n,$$

笔记 即齐次线性方程组解空间的维数与系数矩阵的秩之和等于未知数的个数. 根据上述公式, 由矩阵的秩可以讨论线性方程组解的性质; 反过来, 也可以由线性方程组解的性质讨论矩阵的秩.

推论 3.4

线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 只有零解的充要条件是 $\boldsymbol{A}$ 为列满秩阵. 特别地, 若 $\boldsymbol{A}$ 是方阵, 则线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 只有零解的充要条件是 $\boldsymbol{A}$ 为非异阵.

证明 由定理 3.13 中的公式$\dim V_{\boldsymbol{A}}+\mathrm{r}(\boldsymbol{A})=n$ 即可得到证明. □

命题 3.36

(1) 设 $\boldsymbol{A}$ 是 $m\times n$ 实矩阵, 求证:$\mathrm{r}(\boldsymbol{A}'\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}\boldsymbol{A}')=\mathrm{r}(\boldsymbol{A})$.

(2) 若 $\boldsymbol{A}$ 是 $m\times n$ 复矩阵, 则 $\mathrm{r}\left(\overline{\boldsymbol{A}}'\boldsymbol{A}\right)=\mathrm{r}\left(\boldsymbol{A}\overline{\boldsymbol{A}}'\right)=\mathrm{r}(\boldsymbol{A})$.

证明

(1) 首先证明 $\mathrm{r}(\boldsymbol{A}'\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})$, 为此我们将证明齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 和 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 同解. 显然 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解都是 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解. 反之, 任取方程组 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解 $\boldsymbol{\alpha}\in\mathbb{R}^n$, 则 $\boldsymbol{\alpha}'\boldsymbol{A}'\boldsymbol{A}\boldsymbol{\alpha}=0$, 即 $(\boldsymbol{A}\boldsymbol{\alpha})'(\boldsymbol{A}\boldsymbol{\alpha})=0$. 记 $\boldsymbol{A}\boldsymbol{\alpha}=(b_1,b_2,\cdots,b_m)'\in\mathbb{R}^m$, 则

$$b_1^2+b_2^2+\cdots+b_m^2=0.$$

因为 b_i 是实数, 故每个 $b_i=0$, 即 $\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{0}$, 也即 $\boldsymbol{\alpha}$ 是 $\boldsymbol{A}\boldsymbol{x}=0$ 的解. 这就证明了方程组 $\boldsymbol{A}\boldsymbol{x}=0$ 和 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=0$ 同解, 即 $V_{\boldsymbol{A}}=V_{\boldsymbol{A}'\boldsymbol{A}}$, 于是由定理 3.13可得 $\mathrm{r}(\boldsymbol{A}'\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})$. 在上述等式中用 $\boldsymbol{A}'$ 替代 $\boldsymbol{A}$ 可得 $\mathrm{r}(\boldsymbol{A}\boldsymbol{A}')=\mathrm{r}(\boldsymbol{A}')$, 又因为 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}')$, 故结论得证.

(2) 由 (1) 类似的方法可以证明.

□

例题 3.21 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是数域 $\mathbb{K}$ 上的 n 阶矩阵, 若线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 和 $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 同解, 且每个方程组的基础解系含 m 个线性无关的向量, 求证:$\mathrm{r}(\boldsymbol{A}-\boldsymbol{B})\le n-m$.

证明 由方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 和 $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 同解可知,$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解都是 $(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{x}=\boldsymbol{0}$ 的解, 即 $V_{\boldsymbol{A}}\subseteq V_{\boldsymbol{A}-\boldsymbol{B}}$, 从而 $\dim V_{\boldsymbol{A}-\boldsymbol{B}}\ge\dim V_{\boldsymbol{A}}=m$, 于是由定理 3.13可得 $\mathrm{r}(\boldsymbol{A}-\boldsymbol{B})=n-\dim V_{\boldsymbol{A}-\boldsymbol{B}}\le n-m$. □

命题 3.37

设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵,$\boldsymbol{B}$ 是 $n\times k$ 矩阵, 证明: 方程组 $\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 和方程组 $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 同解的充要条件是 $\mathrm{r}(\boldsymbol{A}\boldsymbol{B})=\mathrm{r}(\boldsymbol{B})$.

证明 显然方程组 $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 的解都是方程组 $\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 的解, 即 $V_{\boldsymbol{B}}\subseteq V_{\boldsymbol{A}\boldsymbol{B}}$, 于是两个线性方程组同解, 即 $V_{\boldsymbol{B}}=V_{\boldsymbol{A}\boldsymbol{B}}$ 的充要条件是 $\dim V_{\boldsymbol{B}}=\dim V_{\boldsymbol{A}\boldsymbol{B}}$. 又由定理 3.13可知 $\dim V_{\boldsymbol{B}}=k-\mathrm{r}(\boldsymbol{B})$,$\dim V_{\boldsymbol{A}\boldsymbol{B}}=k-\mathrm{r}(\boldsymbol{A}\boldsymbol{B})$, 因此上述两个方程组同解的充要条件是 $\mathrm{r}(\boldsymbol{A}\boldsymbol{B})=\mathrm{r}(\boldsymbol{B})$. □

命题 3.38

设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵,$\boldsymbol{B}$ 是 $n\times k$ 矩阵. 若 $\boldsymbol{AB}$ 和 $\boldsymbol{B}$ 有相同的秩, 求证: 对任意的 $k\times l$ 矩阵 $\boldsymbol{C}$, 矩阵 $\boldsymbol{ABC}$ 和矩阵 $\boldsymbol{BC}$ 也有相同的秩.

证明 证法一: 由假设和命题 3.37可知, 方程组 $\boldsymbol{ABx}=\boldsymbol{0}$ 和方程组 $\boldsymbol{Bx}=\boldsymbol{0}$ 同解. 要证明 $\mathrm{r}(\boldsymbol{ABC})=\mathrm{r}(\boldsymbol{BC})$, 我们只要证明方程组 $\boldsymbol{ABCx}=\boldsymbol{0}$ 和方程组 $\boldsymbol{BCx}=\boldsymbol{0}$ 同解即可. 显然方程组 $\boldsymbol{BCx}=\boldsymbol{0}$ 的解都是方程组 $\boldsymbol{ABCx}=\boldsymbol{0}$ 的解. 反之, 若列向量 $\boldsymbol{\alpha}$ 是方程组 $\boldsymbol{ABCx}=\boldsymbol{0}$ 的解, 则 $\boldsymbol{C\alpha}$ 是方程组 $\boldsymbol{ABx}=\boldsymbol{0}$ 的解, 因此 $\boldsymbol{C\alpha}$ 也是方程组 $\boldsymbol{Bx}=\boldsymbol{0}$ 的解, 即 $\boldsymbol{BC\alpha}=\boldsymbol{0}$, 于是 $\boldsymbol{\alpha}$ 也是方程组 $\boldsymbol{BCx}=\boldsymbol{0}$ 的解. 这就证明了方程组 $\boldsymbol{ABCx}=\boldsymbol{0}$ 和方程组 $\boldsymbol{BCx}=\boldsymbol{0}$ 同解, 从而结论得证.

证法二: 由Frobenius 不等式可得

$$\mathrm{r}(\boldsymbol{ABC})\geq \mathrm{r}(\boldsymbol{AB})+\mathrm{r}(\boldsymbol{BC})-\mathrm{r}(\boldsymbol{B})=\mathrm{r}(\boldsymbol{BC}),$$

又由矩阵秩的基本公式 (2)可知 $\mathrm{r}(\boldsymbol{ABC})\leq \mathrm{r}(\boldsymbol{BC})$, 故结论得证. □

命题 3.39

设数域 $\mathbb{K}$ 上的 n 阶矩阵 $\boldsymbol{A}=(a_{ij})$ 满足:$|\boldsymbol{A}|=0$ 且某个元素 a_{ij} 的代数余子式 $A_{ij}\neq 0$. 求证: 齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的所有解都可写为下列形式:

$$k\begin{pmatrix} A_{i1}\\ A_{i2}\\ \vdots\\ A_{in}\end{pmatrix}, k\in\mathbb{K}.$$

证明 由条件和定理 3.12可知 $\boldsymbol{A}$ 的秩等于 $n-1$, 因此线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系只含一个向量. 注意到 $|\boldsymbol{A}|=0$, 故 $\boldsymbol{AA}^*=|\boldsymbol{A}|\boldsymbol{I}_n=\boldsymbol{O}$, 于是伴随矩阵 $\boldsymbol{A}^*$ 的任一列向量都是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解. 又已知 $A_{ij}\neq 0$, 因此 $\boldsymbol{A}^*$ 的第 i 个列向量 $(A_{i1},A_{i2},\cdots,A_{in})'$(不是零向量) 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系. □

定义 3.4 (严格对角占优阵)

如果 n 阶实方阵 $\boldsymbol{A}=(a_{ij})$ 适合条件:

$$|a_{ii}|>\sum_{j=1,j\neq i}^{n}|a_{ij}|, 1\leq i\leq n,$$

则称 $\boldsymbol{A}$ 是**严格对角占优阵**.

命题 3.40 (严格对角占优阵必是非异阵)

如果 n 阶实方阵 $\boldsymbol{A}=(a_{ij})$ 是严格对角占优阵, 则 $\boldsymbol{A}$ 必是非异阵.

证明 证法一: 只需证明线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解. 若有非零解, 设为 $(c_1,c_2,\cdots,c_n)$, 假设 c_k 是其中绝对值最大者. 将解代入该方程组的第 k 个方程式, 得

$$a_{k1}c_1+\cdots+a_{kk}c_k+\cdots+a_{kn}c_n=0,$$

即有

$$-a_{kk}c_k=a_{k1}c_1+\cdots+a_{k,k-1}c_{k-1}+a_{k,k+1}c_{k+1}+\cdots+a_{kn}c_n.$$

上式两边取绝对值, 由三角不等式以及 c_k 是绝对值最大的假设可得

$$|a_{kk}||c_k|\leq|a_{k1}||c_1|+\cdots+|a_{k,k-1}||c_{k-1}|+|a_{k,k+1}||c_{k+1}|+\cdots+|a_{kn}||c_n|\leq\left(\sum_{j=1,j\neq k}^{n}|a_{kj}|\right)|c_k|,$$

从而有

$$|a_{kk}| \le \sum_{j=1,j\ne k}^{n} |a_{kj}|,$$

得到矛盾. 因此, 方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解.

证法二:由第一圆盘定理, A 的特征值落在下列戈氏圆盘中:

$$|z-a_{ii}| \le R_i = \sum_{j=1,j\ne i}^{n} |a_{ij}|,\ 1\le i\le n.$$

A 的严格对角占优条件保证了复平面的原点不落在这些戈氏圆盘中, 因此 A 的特征值全不为零, 从而 A 是非异阵. □

命题 3.41

若 n 阶实方阵 $\boldsymbol{A}=(a_{ij})$ 满足

$$a_{ii} > \sum_{j=1,j\ne i}^{n} |a_{ij}|, 1\le i\le n,$$

求证: $|\boldsymbol{A}|>0$.

证明 证法一: 考虑矩阵 $t\boldsymbol{I}_n+\boldsymbol{A}$, 当 $t\ge 0$ 时, 这是一个严格对角占优阵, 因此由上一个命题可知其行列式 $f(t)=|t\boldsymbol{I}_n+\boldsymbol{A}|$ 不为零. 又 $f(t)$ 是关于 t 的多项式且首项系数为 1, 所以当 t 充分大时,$f(t)>0$. 注意到 $f(t)$ 是 $[0,+\infty)$ 上处处不为零的连续函数, 并且当 t 充分大时取值为正, 因此 $f(t)$ 在 $[0,+\infty)$ 上取值恒为正 (原因见:命题 7.2). 特别地,$f(0)=|\boldsymbol{A}|>0$.

证法二:由第一圆盘定理, A 的特征值落在下列戈氏圆盘中:

$$|z-a_{ii}| \le R_i = \sum_{j=1,j\ne i}^{n} |a_{ij}|,\ 1\le i\le n.$$

由条件可知 $a_{ii}>R_i(1\le i\le n)$, 从而这些戈氏圆盘全部位于虚轴的右侧, 因此 A 的特征值 λ_i 或者是正实数, 或者是实部为正的共轭虚数, 从而 $|A|=\lambda_1\lambda_2\cdots\lambda_n>0$. □

命题 3.42

(1) 设 $\boldsymbol{A}$ 是 n 阶实对称阵, 则:$\boldsymbol{I}_n+\mathrm{i}\boldsymbol{A}$ 和 $\boldsymbol{I}_n-\mathrm{i}\boldsymbol{A}$ 都是非异阵.

(2) 设 $\boldsymbol{A}$ 是 n 阶实反对称阵, $\boldsymbol{D}=\mathrm{diag}\{d_1,d_2,\cdots,d_n\}$ 是同阶对角阵且主对角元素全大于零, 求证:$|\boldsymbol{A}+\boldsymbol{D}|>0$. 特别地, $|\boldsymbol{I}_n\pm\boldsymbol{A}|>0$, 从而 $\boldsymbol{I}_n\pm\boldsymbol{A}$ 都是非异阵.

笔记 (2) 的证明思路: 利用行列式构造连续的多项式函数, 再利用函数连续的性质证明.

证明

(1) 证法一:只需证明 $(\boldsymbol{I}_n+\mathrm{i}\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 只有零解. 由 $\overline{\boldsymbol{x}}'(\boldsymbol{I}_n+\mathrm{i}\boldsymbol{A})\boldsymbol{x}=0$ 共轭转置可得 $\overline{\boldsymbol{x}}'(\boldsymbol{I}_n-\mathrm{i}\boldsymbol{A})\boldsymbol{x}=0$. 上述两式相加, 可得 $\overline{\boldsymbol{x}}'\boldsymbol{I}_n\boldsymbol{x}=0$, 因此 $\boldsymbol{x}=\boldsymbol{0}$.

证法二:设 $\lambda_0\in\mathbb{C}$ 是 A 的任一特征值, $\alpha=(a_1,a_2,\ldots,a_n)'\in\mathbb{C}^n$ 是对应的特征向量, 即 $A\alpha=\lambda_0\alpha$. 此式两边同时左乘 $\overline{\alpha}'$, 则有

$$\overline{\alpha}'A\alpha=\lambda_0\overline{\alpha}'\alpha.$$

注意到 α 是非零向量, 故 $\overline{\alpha}'\alpha=\sum\limits_{i=1}^{n}|a_i|^2>0$. 注意到 A 为实对称矩阵, 故

$$\overline{(\overline{\alpha}'A\alpha)}'=\overline{\alpha}'A\alpha,$$

即 $\overline{\alpha}'A\alpha$ 是一个实数, 从而 $\lambda_0=\overline{\alpha}'A\alpha/\overline{\alpha}'\alpha$ 也是实数. 于是 $I_n\pm\mathrm{i}A$ 的任一特征值为 $1\pm\mathrm{i}\lambda_0\ne 0$. 因此由特征值与特征多项式系数的关系可知 $I_n\pm\mathrm{i}A$ 是非异阵.

(2) 证法一:先证明 $|\boldsymbol{A}+\boldsymbol{D}|\ne 0$, 只需证明 $(\boldsymbol{A}+\boldsymbol{D})\boldsymbol{x}=\boldsymbol{0}$ 只有零解. 因为 $\boldsymbol{x}'(\boldsymbol{A}+\boldsymbol{D})\boldsymbol{x}=0$, 转置可得 $\boldsymbol{x}'(-\boldsymbol{A}+\boldsymbol{D})\boldsymbol{x}=0$,

上述两式相加即得 $\boldsymbol{x}'\boldsymbol{D}\boldsymbol{x}=0$. 若设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)'$, 则有 $d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2=0$. 由于 d_i 都大于零并且 x_i 都是实数, 故只能是 $x_1=x_2=\cdots=x_n=0$, 即有 $\boldsymbol{x}=\boldsymbol{0}$.

再证明 (2) 的结论. 设 $f(t)=|t\boldsymbol{A}+\boldsymbol{D}|$, 则 $f(t)$ 是关于 t 的多项式, 从而是关于 t 的连续函数. 注意到对任意的实数 t,$t\boldsymbol{A}$ 仍是实反对称阵, 故由上面的讨论可得 $f(t)=|t\boldsymbol{A}+\boldsymbol{D}|\neq 0$, 即 $f(t)$ 是 $\mathbb{R}$ 上处处不为零的连续函数. 注意到当 $t=0$ 时,$f(0)=|\boldsymbol{D}|>0$, 因此 $f(t)$ 只能是 $\mathbb{R}$ 上取值恒为正数的连续函数 (原因见:命题 7.2). 特别地,$f(1)=|\boldsymbol{A}+\boldsymbol{D}|>0$.

证法二:设 $\lambda_0\in\mathbb{C}$ 是 A 的任一特征值, $\alpha=(a_1,a_2,\ldots,a_n)'\in\mathbb{C}^n$ 是对应的特征向量, 即 $A\alpha=\lambda_0\alpha$. 此式两边同时左乘 $\overline{\alpha}'$, 则有

$$\overline{\alpha}'A\alpha=\lambda_0\overline{\alpha}'\alpha.$$

注意到 α 是非零向量, 故 $\overline{\alpha}'\alpha=\sum\limits_{i=1}^{n}|a_i|^2>0$. 注意到 A 为实反称矩阵, 故

$$\overline{(\overline{\alpha}'A\alpha)}'=-\overline{\alpha}'A\alpha,$$

即 $\overline{\alpha}'A\alpha$ 是零或纯虚数, 从而 $\lambda_0=\overline{\alpha}'A\alpha/\overline{\alpha}'\alpha$ 也是零或纯虚数. 设 $\lambda=c\mathrm{i}$, 其中 $c\in\mathbb{C}$. 于是 $I_n\pm A$ 的任一特征值为 $1\pm\lambda_0=1\pm c\mathrm{i}\neq 0\neq 0$, 由特征值与特征多项式系数的关系可知 $I_n\pm A$ 是非异阵.

□

3.6.3 利用线性空间理论讨论矩阵的秩

按照最初的定义, 矩阵的秩就是矩阵的行 (列) 向量组的秩, 因此通过线性空间理论去讨论矩阵的秩是十分自然的事情.

命题 3.43

求证: 矩阵 $\boldsymbol{A}$ 的秩等于 r 的充要条件是 $\boldsymbol{A}$ 存在一个 r 阶子式 $|\boldsymbol{D}|$ 不等于零, 而 $|\boldsymbol{D}|$ 的所有 $r+1$ 阶加边子式全等于零.

证明 必要性由定理 3.12可直接得到, 只需证明充分性. 不失一般性, 我们可设 $|\boldsymbol{D}|$ 是由 $\boldsymbol{A}$ 的前 r 行和前 r 列构成的 r 阶子式. 设

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$$

为矩阵 $\boldsymbol{A}$ 的行分块和列分块, 记 $\tau_{\le r}\boldsymbol{\alpha}_i$ 为行向量 $\boldsymbol{\alpha}_i$ 关于前 r 列的缩短向量,$\tau_{\le r}\boldsymbol{\beta}_j$ 为列向量 $\boldsymbol{\beta}_j$ 关于前 r 行的缩短向量. 由 $|\boldsymbol{D}|\neq 0$ 可得 $\tau_{\le r}\boldsymbol{\alpha}_1,\cdots,\tau_{\le r}\boldsymbol{\alpha}_r$ 线性无关, 由命题 3.7可知 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性无关.

我们只要证明 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 是 $\boldsymbol{A}$ 的行向量的极大无关组即可得到 $\mathrm{r}(\boldsymbol{A})=r$. 用反证法证明, 若它们不是极大无关组, 则可以添加一个行向量, 不妨设为 $\boldsymbol{\alpha}_{r+1}$, 使得 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1}$ 线性无关. 设 $\boldsymbol{A}_1$ 是 $\boldsymbol{A}$ 的前 $r+1$ 行构成的矩阵, 则 $\boldsymbol{A}_1=(\tau_{\le r+1}\boldsymbol{\beta}_1,\tau_{\le r+1}\boldsymbol{\beta}_2,\cdots,\tau_{\le r+1}\boldsymbol{\beta}_n)$ 且 $\mathrm{r}(\boldsymbol{A}_1)=r+1$. 由 $|\boldsymbol{D}|\neq 0$ 可得 $\tau_{\le r}\boldsymbol{\beta}_1,\cdots,\tau_{\le r}\boldsymbol{\beta}_r$ 线性无关, 由命题 3.7可知 $\tau_{\le r+1}\boldsymbol{\beta}_1,\cdots,\tau_{\le r+1}\boldsymbol{\beta}_r$ 线性无关. 因为 $\mathrm{r}(\boldsymbol{A}_1)=r+1$, 故存在 $\boldsymbol{A}_1$ 的一个列向量, 不妨设为 $\tau_{\le r+1}\boldsymbol{\beta}_{r+1}$, 使得 $\tau_{\le r+1}\boldsymbol{\beta}_1,\cdots,\tau_{\le r+1}\boldsymbol{\beta}_r,\tau_{\le r+1}\boldsymbol{\beta}_{r+1}$ 线性无关. 设 $\boldsymbol{A}_2=(\tau_{\le r+1}\boldsymbol{\beta}_1,\cdots,\tau_{\le r+1}\boldsymbol{\beta}_r,\tau_{\le r+1}\boldsymbol{\beta}_{r+1})$, 即 $\boldsymbol{A}_2$ 是 $\boldsymbol{A}$ 的前 $r+1$ 行和前 $r+1$ 列构成的方阵, 则 $\mathrm{r}(\boldsymbol{A}_2)=r+1$. 因此,$|\boldsymbol{A}_2|\neq 0$ 是包含 $|\boldsymbol{D}|$ 的 $r+1$ 阶加边子式, 这与假设矛盾. □

命题 3.44

设 $m\times n$ 矩阵 $\boldsymbol{A}$ 的 m 个行向量为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$, 且 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 是其极大无关组, 又设 $\boldsymbol{A}$ 的 n 个列向量为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$, 且 $\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_r}$ 是其极大无关组. 证明:$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 和 $\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_r}$ 交叉点上的元素组成的子矩阵 $\boldsymbol{D}$ 的行列式 $|\boldsymbol{D}|\neq 0$.

证明 因为 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 是极大无关组, 故 $\boldsymbol{A}$ 的任一行向量 $\boldsymbol{\alpha}_s$ 均可表示为 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 的线性组合. 记 $\widetilde{\boldsymbol{\alpha}}_{i_1},\widetilde{\boldsymbol{\alpha}}_{i_2},\cdots,\widetilde{\boldsymbol{\alpha}}_{i_r},\widetilde{\boldsymbol{\alpha}}_s$ 分别是 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r},\boldsymbol{\alpha}_s$ 在 $j_1,j_2,\cdots,j_r$ 列处的缩短向量, 由命题 3.7可知,$\widetilde{\boldsymbol{\alpha}}_s$ 均可表示为 $\widetilde{\boldsymbol{\alpha}}_{i_1},\widetilde{\boldsymbol{\alpha}}_{i_2},\cdots,\widetilde{\boldsymbol{\alpha}}_{i_r}$ 的线性组合.

考虑由列向量 $\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_r}$ 组成的矩阵 $\boldsymbol{B}=(\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_r})$, 这是一个 $m\times r$ 矩阵且秩等于 r. 由于矩阵 $\boldsymbol{B}$ 的任一行向量 $\widetilde{\boldsymbol{\alpha}}_s$ 均可用 $\widetilde{\boldsymbol{\alpha}}_{i_1},\widetilde{\boldsymbol{\alpha}}_{i_2},\cdots,\widetilde{\boldsymbol{\alpha}}_{i_r}$ 线性表示, 并且 $\boldsymbol{B}$ 的行秩等于 r, 故由命题 3.11可知,$\widetilde{\boldsymbol{\alpha}}_{i_1},\widetilde{\boldsymbol{\alpha}}_{i_2},\cdots,\widetilde{\boldsymbol{\alpha}}_{i_r}$ 是 $\boldsymbol{B}$ 的行向量的极大无关组, 从而它们线性无关. 注意到 r 阶方阵 $\boldsymbol{D}$ 的行向量恰好是 $\widetilde{\boldsymbol{\alpha}}_{i_1},\widetilde{\boldsymbol{\alpha}}_{i_2},\cdots,\widetilde{\boldsymbol{\alpha}}_{i_r}$, 因此 $\boldsymbol{D}$ 是满秩阵, 从而 $|\boldsymbol{D}|\neq 0$. □

命题 3.45

设 $\boldsymbol{A}$ 是一个 n 阶方阵,$\boldsymbol{A}$ 的第 $i_1,\cdots,i_r$ 行和第 $i_1,\cdots,i_r$ 列交叉点上的元素组成的子式称为 $\boldsymbol{A}$ 的主子式. 若 $\boldsymbol{A}$ 是对称阵或反对称阵且秩等于 r, 求证:$\boldsymbol{A}$ 必有一个 r 阶主子式不等于零.

证明 由对称性或反对称性, 设 $\boldsymbol{A}$ 的第 $i_1,\cdots,i_r$ 行是 $\boldsymbol{A}$ 的行向量的极大无关组, 则由命题 3.12它的第 $i_1,\cdots,i_r$ 列也是 $\boldsymbol{A}$ 的列向量的极大无关组, 因此由命题 3.44可知, 它们交叉点上的元素组成的 r 阶主子式不等于零. □

命题 3.46 (反对称阵的秩必为偶数)

证明: 反对称阵的秩必为偶数.

证明 用反证法, 设反对称阵 $\boldsymbol{A}$ 的秩等于 $2r+1$, 则由命题 3.45可知,$\boldsymbol{A}$ 有一个 $2r+1$ 阶主子式 $|\boldsymbol{D}|$ 不等于零. 注意到反对称阵的主子式是反对称行列式, 而由命题 1.15奇数阶反对称行列式的值等于零, 从而 $|\boldsymbol{D}|=0$, 矛盾. □

3.7 相抵标准型及其应用

定理 3.14 (矩阵的相抵标准型)

对任意一个秩为 r 的 $m\times n$ 矩阵 $\boldsymbol{A}$, 总存在 m 阶非异阵 $\boldsymbol{P}$ 和 n 阶非异阵 $\boldsymbol{Q}$, 使得

$$\boldsymbol{PAQ}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}.$$

证明

□

命题 3.47 (矩阵的秩 1 分解)

求证: 秩等于 r 的矩阵可以表示为 r 个秩等于 1 的矩阵之和, 但不能表示为少于 r 个秩为 1 的矩阵之和.

证明 将 $\boldsymbol{A}$ 化为相抵标准型, 即存在非异矩阵 $\boldsymbol{P}$ 及 $\boldsymbol{Q}$, 使得

$$\begin{aligned}\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}&=\boldsymbol{P}(\boldsymbol{E}_{11}+\boldsymbol{E}_{22}+\cdots+\boldsymbol{E}_{rr})\boldsymbol{Q}\\&=\boldsymbol{PE}_{11}\boldsymbol{Q}+\boldsymbol{PE}_{22}\boldsymbol{Q}+\cdots+\boldsymbol{PE}_{rr}\boldsymbol{Q}.\end{aligned}$$

于是记 $\boldsymbol{A}_1=\boldsymbol{PE}_{11}\boldsymbol{Q},\boldsymbol{A}_2=\boldsymbol{PE}_{22}\boldsymbol{Q},\cdots,\boldsymbol{A}_r=\boldsymbol{PE}_{rr}\boldsymbol{Q}$, 则每个 $\boldsymbol{A}_i$ 的秩都等于 1. 故 $\boldsymbol{A}$ 可以化为 r 个秩等于 1 的矩阵之和.

若 $\boldsymbol{A}=\boldsymbol{B}_1+\boldsymbol{B}_2+\cdots+\boldsymbol{B}_k,k<r$, 且每个 $\boldsymbol{B}_i$ 的秩都等于 1, 则由命题 3.28(6)可知 $\mathrm{r}(\boldsymbol{A})\le\mathrm{r}(\boldsymbol{B}_1)+\mathrm{r}(\boldsymbol{B}_2)+\cdots+\mathrm{r}(\boldsymbol{B}_k)=k$, 这与 $\mathrm{r}(\boldsymbol{A})=r$ 矛盾, 故不可能. □

命题 3.48 (对称矩阵的秩 1 分解)

秩等于 r 的对称矩阵可以表成 r 个秩等于 1 的对称矩阵之和.

证明 设 A 是一个秩为 r 的对称矩阵, 则存在一个可逆矩阵 C, 使得

$$C^TAC=\begin{pmatrix}I_r & \\ & O\end{pmatrix}=E_{11}+E_{22}+\cdots+E_{rr}.$$

从而

$$\begin{aligned}A&=(C^T)^{-1}(E_{11}+E_{22}+\cdots+E_{rr})C^{-1}\\&=(C^T)^{-1}E_{11}C^{-1}+(C^T)^{-1}E_{22}C^{-1}+\cdots+(C^T)^{-1}E_{nn}C^{-1}.\end{aligned}$$

因为 E_{ii} 的秩为 1, 且 $(C^T)^{-1}$,C^{-1} 均可逆, 所以 $(C^T)^{-1}E_{ii}C^{-1}$ 的秩也为 1. 又由于

$$((C^T)^{-1}E_{ii}C^{-1})^T=(C^{-1})^TE_{ii}^TC^{-1}=(C^T)^{-1}E_{ii}C^{-1}.$$

因此 $(C^T)^{-1}E_{ii}C^{-1}$ 也是对称矩阵. 故 A 可以表成 r 个秩等于 1 的对称矩阵之和. □

例题 3.22 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 分别为 $m\times n,p\times q$ 和 $m\times q$ 矩阵,$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}$. 证明:$\mathrm{r}(\boldsymbol{M})=\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$ 成立的充要条件是矩阵方程 $\boldsymbol{AX}+\boldsymbol{YB}=\boldsymbol{C}$ 有解, 其中 $\boldsymbol{X},\boldsymbol{Y}$ 分别是 $n\times q$ 和 $m\times p$ 未知矩阵.

笔记 证明必要性时不妨设的原因: 假设当 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$,$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{I}_s & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$ 时, 结论成立. 则当 $\boldsymbol{A}\neq\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$,$\boldsymbol{B}\neq\begin{pmatrix}\boldsymbol{I}_s & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$ 时, 记 $\boldsymbol{A}_1=\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$,$\boldsymbol{B}_1=\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2=\begin{pmatrix}\boldsymbol{I}_s & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$,$\boldsymbol{C}_1=\boldsymbol{P}_1\boldsymbol{C}\boldsymbol{Q}_2$,$\boldsymbol{M}_1=\begin{pmatrix}\boldsymbol{A}_1 & \boldsymbol{C}_1\\ \boldsymbol{O} & \boldsymbol{B}_1\end{pmatrix}$.

由于矩阵乘可逆矩阵不改变其秩, 因此

$$\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1)=\mathrm{r}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}=\mathrm{r}(\boldsymbol{A}_1),$$

$$\mathrm{r}(\boldsymbol{B})=\mathrm{r}(\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2)=\mathrm{r}\begin{pmatrix}\boldsymbol{I}_s & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}=\mathrm{r}(\boldsymbol{B}_1),$$

$$\mathrm{r}(\boldsymbol{M})=\mathrm{r}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}=\mathrm{r}\left(\begin{pmatrix}\boldsymbol{P}_1 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{P}_2\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{Q}_1 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{Q}_2\end{pmatrix}\right)=\mathrm{r}\begin{pmatrix}\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1 & \boldsymbol{P}_1\boldsymbol{C}\boldsymbol{Q}_2\\ \boldsymbol{O} & \boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2\end{pmatrix}=\mathrm{r}(\boldsymbol{M}_1).$$

从而

$$\mathrm{r}(\boldsymbol{M})=\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})\Leftrightarrow\mathrm{r}(\boldsymbol{M}_1)=\mathrm{r}\begin{pmatrix}\boldsymbol{A}_1 & \boldsymbol{C}_1\\ \boldsymbol{O} & \boldsymbol{B}_1\end{pmatrix}=\mathrm{r}(\boldsymbol{A}_1)+\mathrm{r}(\boldsymbol{B}_1).$$

于是由假设可知 $\boldsymbol{A}_1\boldsymbol{X}_1+\boldsymbol{Y}_1\boldsymbol{B}_1=\boldsymbol{C}_1$ 有解 $\boldsymbol{X}_1,\boldsymbol{Y}_1$. 记 $\boldsymbol{X}=\boldsymbol{Q}_1\boldsymbol{X}_1\boldsymbol{Q}_2^{-1}$,$\boldsymbol{Y}=\boldsymbol{P}_1^{-1}\boldsymbol{Y}_1\boldsymbol{P}_2$, 则

$$\begin{aligned}&\boldsymbol{A}_1\boldsymbol{X}_1+\boldsymbol{Y}_1\boldsymbol{B}_1=\boldsymbol{C}_1\text{有解}\boldsymbol{X}_1,\boldsymbol{Y}_1\\ \Leftrightarrow&\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{X}_1+\boldsymbol{Y}_1\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2=\boldsymbol{P}_1\boldsymbol{C}\boldsymbol{Q}_2\text{有解}\boldsymbol{X}_1,\boldsymbol{Y}_1\\ \Leftrightarrow&\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{X}_1\boldsymbol{Q}_2^{-1}+\boldsymbol{P}_1^{-1}\boldsymbol{Y}_1\boldsymbol{P}_2\boldsymbol{B}=\boldsymbol{C}\text{有解}\boldsymbol{X}_1,\boldsymbol{Y}_1\\ \Leftrightarrow&\boldsymbol{A}\boldsymbol{X}+\boldsymbol{Y}\boldsymbol{B}=\boldsymbol{C}\text{有解}\boldsymbol{X},\boldsymbol{Y}\end{aligned}$$

故可以不妨设 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$,$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{I}_s & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$.

证明 先证充分性. 设 $\boldsymbol{X}=\boldsymbol{X}_0,\boldsymbol{Y}=\boldsymbol{Y}_0$ 是矩阵方程 $\boldsymbol{AX}+\boldsymbol{YB}=\boldsymbol{C}$ 的解, 则将 $\boldsymbol{M}$ 的第一分块列右乘 $-\boldsymbol{X}_0$ 加到第二分块列上, 再将第二分块行左乘 $-\boldsymbol{Y}_0$ 加到第一分块行上, 可得分块对角阵 $\begin{pmatrix}A & O\\ O & B\end{pmatrix}$, 于是 $\mathrm{r}(\boldsymbol{M})=\mathrm{r}\begin{pmatrix}A & O\\ O & B\end{pmatrix}=\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})$.

再证必要性. 设 $\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$,$\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2=\begin{pmatrix}\boldsymbol{I}_s & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$, 其中 $\boldsymbol{P}_1,\boldsymbol{Q}_1,\boldsymbol{P}_2,\boldsymbol{Q}_2$ 为非异阵,$r=\mathrm{r}(\boldsymbol{A}),s=\mathrm{r}(\boldsymbol{B})$. 注意到问题的条件和结论在相抵变换: $\boldsymbol{A}\mapsto\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1,\boldsymbol{B}\mapsto\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2,\boldsymbol{C}\mapsto\boldsymbol{P}_1\boldsymbol{C}\boldsymbol{Q}_2,\boldsymbol{X}\mapsto\boldsymbol{Q}_1^{-1}\boldsymbol{X}\boldsymbol{Q}_2,\boldsymbol{Y}\mapsto\boldsymbol{P}_1\boldsymbol{Y}\boldsymbol{P}_2^{-1}$ 下保持不变, 故不妨从一开始就假设 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$,$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{I}_s & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$ 都是相抵标准型. 设 $\boldsymbol{C}=\begin{pmatrix}\boldsymbol{C}_1 & \boldsymbol{C}_2\\ \boldsymbol{C}_3 & \boldsymbol{C}_4\end{pmatrix}$,$\boldsymbol{X}=$

$\begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}, Y = \begin{pmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{pmatrix}$ 为对应的分块. 考虑 M 的如下分块初等变换:

$$M = \begin{pmatrix} I_r & O & C_1 & C_2 \\ O & O & C_3 & C_4 \\ O & O & I_s & O \\ O & O & O & O \end{pmatrix} \to \begin{pmatrix} I_r & O & O & O \\ O & O & O & C_4 \\ O & O & I_s & O \\ O & O & O & O \end{pmatrix},$$

由于 $\mathrm{r}(M) = \mathrm{r}(A) + \mathrm{r}(B) = r + s$, 故 $C_4 = O$. 于是矩阵方程 $AX + YB = C$, 即

$$\begin{pmatrix} X_1 & X_2 \\ O & O \end{pmatrix} + \begin{pmatrix} Y_1 & O \\ Y_3 & O \end{pmatrix} = \begin{pmatrix} X_1 + Y_1 & X_2 \\ Y_3 & O \end{pmatrix} = \begin{pmatrix} C_1 & C_2 \\ C_3 & O \end{pmatrix}$$

有解, 例如 $X_1 = C_1, X_2 = C_2, Y_1 = O, Y_3 = C_3$, 其余分块取法任意. □

命题 3.49 (行/列满秩矩阵性质)

由矩阵的相抵标准型可设 A 是 $m \times n$ 矩阵, 则

(1) 若 $\mathrm{r}(A) = n$, 即 A 是列满秩阵, 则必存在秩等于 n 的 $n \times m$ 矩阵 B(行满秩), 使得 $BA = I_n$(这样的矩阵 B 称为 A 的左逆);

(2) 若 $\mathrm{r}(A) = m$, 即 A 是行满秩阵, 则必存在秩等于 m 的 $n \times m$ 矩阵 C(列满秩), 使得 $AC = I_m$(这样的矩阵 C 称为 A 的右逆).

证明

(1) 设 P 为 m 阶非异阵,Q 为 n 阶非异阵, 使得

$$PAQ = \begin{pmatrix} I_n \\ O \end{pmatrix},$$

因此 $(I_n, O)PAQ = I_n$, 即 $(I_n, O)PA = Q^{-1}$, 于是 $Q(I_n, O)PA = I_n$. 令 $B = Q(I_n, O)P$ 即可.

(2) 同理可证, 或者考虑 A' 并利用 (1) 的结论.

□

推论 3.5

列满秩矩阵适合左消去律, 即若 A 列满秩且 $AD = AE$, 则 $D = E$. 同理, 行满秩矩阵适合右消去律, 即若 A 行满秩且 $DA = EA$, 则 $D = E$.

命题 3.50 (满秩分解)

设 $m \times n$ 矩阵 A 的秩为 r, 证明:

(1) $A = BC$, 其中 B 是 $m \times r$(列满秩) 矩阵且 $\mathrm{r}(B) = r$,C 是 $r \times n$(行满秩) 矩阵且 $\mathrm{r}(C) = r$, 这种分解称为 A 的满秩分解;

(2) 若 A 有两个满秩分解 $A = B_1C_1 = B_2C_2$, 则存在 r 阶非异阵 P, 使得 $B_2 = B_1P, C_2 = P^{-1}C_1$.

证明

(1) 设 P 为 m 阶非异阵,Q 为 n 阶非异阵, 使得

$$A = P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}Q = P\begin{pmatrix} I_r \\ O \end{pmatrix}(I_r, O)Q.$$

令 $B = P\begin{pmatrix} I_r \\ O \end{pmatrix}$,$C = (I_r, O)Q$, 即得结论.

(2) 由行/列满秩矩阵性质可知, 存在 $r \times m$ 行满秩阵 S_2,$n \times r$ 列满秩阵 T_2, 使得 $S_2B_2 = I_r$,$C_2T_2 = I_r$, 于是

$$B_2 = B_2(C_2T_2) = (B_2C_2)T_2 = (B_1C_1)T_2 = B_1(C_1T_2),$$

$$C_2 = (S_2B_2)C_2 = S_2(B_2C_2) = S_2(B_1C_1) = (S_2B_1)C_1,$$
$$(S_2B_1)(C_1T_2) = S_2(B_1C_1)T_2 = S_2(B_2C_2)T_2 = (S_2B_2)(C_2T_2) = I_r.$$

令 $P = C_1T_2$, 即得结论.

□

命题 3.51

$A = BC$ 是满秩分解当且仅当 B 的 r 个列向量是 A 的 n 个列向量张成线性空间的一组基, 也当且仅当 C 的 r 个行向量是 A 的 m 个行向量张成线性空间的一组基.

证明

□

例题 3.23 设 A 为 $m\times n$ 矩阵, 证明: 存在 $n\times m$ 矩阵 B, 使得 $ABA = A$.

笔记 证法一的不妨设原因与例题 3.22类似.

证明 证法一: 设 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 其中 P 是 m 阶非异阵,Q 是 n 阶非异阵. 注意到问题的条件和结论在相抵变换:$A \mapsto PAQ, B \mapsto Q^{-1}BP^{-1}$ 下保持不变, 故不妨从一开始就假设 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 是相抵标准型. 设 $B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$ 为对应的分块, 由 $ABA = A$ 可得 $B_1 = I_r$, 其余分块取法任意.

证法二:设 $A = CD$ 为 A 的满秩分解,E 为列满秩阵 C 的左逆,F 是行满秩阵 D 的右逆. 令 $B = FE$, 则

$$ABA = (CD)(FE)(CD) = C(DF)(EC)D = CD = A.$$

□

例题 3.24 设 A, B 分别是 $3\times 2, 2\times 3$ 矩阵且满足

$$AB = \begin{pmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{pmatrix},$$

试求 BA.

证明 解法一: 通过简单的计算可得 $\mathrm{r}(AB) = 2$, 从而 $\mathrm{r}(A) \geq 2, \mathrm{r}(B) \geq 2$. 又因为矩阵的秩不超过行数和列数的最小值, 故 $\mathrm{r}(A) = \mathrm{r}(B) = 2$, 即 A 是列满秩阵,B 是行满秩阵. 又注意到 $(AB)^2 = 9AB$, 经整理可得 $A(BA - 9I_2)B = O$. 根据推论 3.5, 可以在上式的左边消去 A, 右边消去 B, 从而可得 $BA = 9I_2$.

解法二:由解法一中矩阵秩的计算可知,AB 是题中 3 阶矩阵 C 的满秩分解. 注意到 C 的后两列线性无关, 因此可取另一种满秩分解为

$$C = \begin{pmatrix} 2 & -2 \\ 5 & 4 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} = A_1B_1.$$

由矩阵的满秩分解 (2)可知, 存在可逆矩阵 P, 使得 $A_1 = AP, B_1 = P^{-1}B$. 于是 $B_1A_1 = P^{-1}BAP$, 故 BA 相似于 $B_1A_1 = 9I_2$, 从而 $BA = P^{-1}(9I_2)P = 9I_2$.

解法三: 经简单的计算可得 $|\lambda I_3 - AB| = \lambda(\lambda - 9)^2$, 且特征值 9 的几何重数也等于 2, 因此 AB 可对角化. 由特征值的降价公式可得 $|\lambda I_2 - BA| = (\lambda - 9)^2$, 从而 BA 的两个特征值都是 9, 于是 BA 是可逆矩阵 (特征值都非零). 因此由命题 6.38可知 BA 也可对角化, 于是 BA 相似于 $9I_2$, 即存在可逆矩阵 P, 使得 $BA = P^{-1}(9I_2)P = 9I_2$. □

命题 3.52 (幂等矩阵关于满秩分解的刻画)

设 A 是 n 阶方阵且 $\mathrm{r}(A) = r$, 求证:$A^2 = A$ 的充要条件是存在秩等于 r 的 $n\times r$ 矩阵 S 和秩等于 r 的 $r\times n$ 矩阵 T, 使得 $A = ST, TS = I_r$.

证明 充分性显然, 现证必要性. 设 $\boldsymbol{P},\boldsymbol{Q}$ 为 n 阶非异阵, 使得

$$\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}.$$

代入 $\boldsymbol{A}^2=\boldsymbol{A}$ 消去两侧的非异阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$, 可得

$$\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{QP}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}.$$

只需令

$$\boldsymbol{S}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_r\\ \boldsymbol{O}\end{pmatrix},\boldsymbol{T}=(\boldsymbol{I}_r,\boldsymbol{O})\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q},$$

则 $\boldsymbol{S}$ 列满秩,$\boldsymbol{I}$ 行满秩, 经简单计算即得结论. □

推论 3.6 (幂等矩阵的迹和秩相等)

设 $\boldsymbol{A}$ 为 n 阶幂等矩阵, 则 $\mathrm{tr}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})$.

证明 证法一:由命题 3.52可知,$\mathrm{tr}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{ST})=\mathrm{tr}(\boldsymbol{TS})=\mathrm{tr}(\boldsymbol{I}_r)=r=\mathrm{r}(\boldsymbol{A})$.

证法二 (相似标准型):事实上, 由 $\boldsymbol{A}^2=\boldsymbol{A}$ 可知, 存在可逆矩阵 $\boldsymbol{P}$, 使得

$$\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{P}^{-1}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r\\ \boldsymbol{O}\end{pmatrix}(\boldsymbol{I}_r,\boldsymbol{O})\boldsymbol{P}^{-1},$$

令 $\boldsymbol{S}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r\\ \boldsymbol{O}\end{pmatrix},\boldsymbol{T}=(\boldsymbol{I}_r,\boldsymbol{O})\boldsymbol{P}^{-1}$, 可得 $\mathrm{tr}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{ST})=\mathrm{tr}(\boldsymbol{TS})=\mathrm{tr}(\boldsymbol{I}_r)=r=\mathrm{r}(\boldsymbol{A})$. □

3.8 线性方程组的解及其应用

3.8.1 线性方程组的解的讨论

命题 3.53

线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 与 $\boldsymbol{Bx}=\boldsymbol{0}$ 同解当且仅当 $r(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})$.

证明

□

例题 3.25 设 $\boldsymbol{A}$ 是一个 $m\times n$ 矩阵, 记 $\boldsymbol{\alpha}_i$ 是 $\boldsymbol{A}$ 的第 i 个行向量,$\boldsymbol{\beta}=(b_1,b_2,\cdots,b_n)$. 求证: 若齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解全是方程 $b_1x_1+b_2x_2+\cdots+b_nx_n=0$ 的解, 则 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的线性组合.

证明 令 $\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{\beta}\end{pmatrix}$, 由已知, 方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 和方程组 $\boldsymbol{Bx}=\boldsymbol{0}$ 同解, 故 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})$, 从而 $\boldsymbol{A}$ 的行向量的极大无关组也是 $\boldsymbol{B}$ 的行向量的极大无关组. 因此,$\boldsymbol{\beta}$ 可表示为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的线性组合. □

例题 3.26 设 $\boldsymbol{Ax}=\boldsymbol{\beta}$ 是 m 个方程式 n 个未知数的线性方程组, 求证: 它有解的充要条件是方程组 $\boldsymbol{A}'\boldsymbol{y}=\boldsymbol{0}$ 的任一解 $\boldsymbol{\alpha}$ 均适合等式 $\boldsymbol{\alpha}'\boldsymbol{\beta}=0$.

证明 方程组 $\boldsymbol{Ax}=\boldsymbol{\beta}$ 有解当且仅当 $\mathrm{r}(\boldsymbol{A}\,|\,\boldsymbol{\beta})=\mathrm{r}(\boldsymbol{A})$, 当且仅当 $\mathrm{r}\begin{pmatrix}\boldsymbol{A}'\\ \boldsymbol{\beta}'\end{pmatrix}=\mathrm{r}(\boldsymbol{A}')$, 当且仅当方程组 $\begin{pmatrix}\boldsymbol{A}'\\ \boldsymbol{\beta}'\end{pmatrix}\boldsymbol{y}=\boldsymbol{0}$ 与

$A'y = 0$ 同解, 而这当且仅当 $A'y = 0$ 的任一解 α 均适合等式 $\beta'\alpha = 0$, 即 $\alpha'\beta = 0$. □

例题 3.27 设有两个线性方程组:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m; \end{cases} \tag{3.7}$$

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \cdots + a_{m1}x_m = 0, \\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{m2}x_m = 0, \\ \cdots\cdots\cdots\cdots \\ a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{mn}x_m = 0, \\ b_1x_1 + b_2x_2 + \cdots + b_mx_m = 1. \end{cases} \tag{3.8}$$

求证: 方程组(3.7)有解的充要条件是方程组(3.8)无解.

证明 设第一个线性方程组的系数矩阵为 A, 常数向量为 β, 则第二个线性方程组的系数矩阵和增广矩阵分别为

$$B = \begin{pmatrix} A' \\ \beta' \end{pmatrix}, \widetilde{B} = \begin{pmatrix} A' & O \\ \beta' & 1 \end{pmatrix}.$$

显然, 由矩阵初等变换可知, 我们有 $\mathrm{r}(\widetilde{B}) = \mathrm{r}(A') + 1 = \mathrm{r}(A) + 1$.

若方程组(3.7)有解, 则 $\mathrm{r}(A \mid \beta) = \mathrm{r}(A)$, 故 $\mathrm{r}(B) = \mathrm{r}(B') = \mathrm{r}(A \mid \beta) = \mathrm{r}(A) \neq \mathrm{r}(\widetilde{B})$. 因此, 方程组(3.8)无解.

反之, 若方程组(3.7)无解, 则 $\mathrm{r}(A \mid \beta) = \mathrm{r}(A)+1$, 故 $\mathrm{r}(B) = \mathrm{r}(B') = \mathrm{r}(A \mid \beta) = \mathrm{r}(A)+1 = \mathrm{r}(\widetilde{B})$. 因此, 方程组(3.8)有解. □

命题 3.54

设 A 是秩为 r 的 $m \times n$ 矩阵, 求证: 必存在秩为 $n-r$ 的 $n \times (n-r)$ 矩阵 B, 使得 $AB = O$.

证明 考虑线性方程组 $Ax = 0$, 它有 $n-r$ 个基础解系, 不妨设为 $\beta_1, \cdots, \beta_{n-r}$. 令 $B = (\beta_1, \cdots, \beta_{n-r})$, 则 $AB = (A\beta_1, \cdots, A\beta_{n-r}) = O$, 结论得证. □

例题 3.28 设

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} (m < n),$$

已知 $Ax = 0$ 的基础解系为 $\beta_i = (b_{i1}, b_{i2}, \cdots, b_{in})'(1 \le i \le n-m)$, 试求齐次线性方程组

$$\sum_{j=1}^{n} b_{ij}y_j = 0(i = 1, 2, \cdots, n-m)$$

的基础解系.

解 令 $B = (\beta_1, \beta_2, \cdots, \beta_{n-m})$, 则 $AB = O, B'A' = O$. 因为 $Ax = 0$ 有 $n-m$ 个基础解系, 所以 A 的秩为 m. 又由于 $\mathrm{r}(B) = \mathrm{r}(B) = n-m$, 因此 $B'y = 0$ 的基础解系有 m 个. 故 $B'y = 0$ 的基础解系为 A' 的全部列向量, 即 A 的所有行向量. □

命题 3.55

设 V_0 是数域 $\mathbb{K}$ 上 n 维列向量空间的真子空间, 求证: 必存在矩阵 A, 使得 V_0 是 n 元齐次线性方程组 $Ax = 0$ 的解空间.

证明 设 $\beta_1, \cdots, \beta_r$ 是子空间 V_0 的一组基. 令 $B = (\beta_1, \cdots, \beta_r)$, 这是一个 $n \times r$ 矩阵. 考虑齐次线性方程组 $B'x = 0$,

因为 $\boldsymbol{B}$ 的秩等于 r, 故其基础解系含 $n-r$ 个向量, 记为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r}$. 令 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r})'$, 这是个 $(n-r)\times n$ 矩阵且秩为 $n-r$. 由 $\boldsymbol{B}'\boldsymbol{A}'=\boldsymbol{O}$ 可得 $\boldsymbol{AB}=\boldsymbol{O}$, 因此齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系是 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$, 其解空间就是 V_0. □

注 设 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 是子空间 V_0 的一组基. 令 $\boldsymbol{B}=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)$, 这是一个 $n\times r$ 矩阵. 也可以由命题 3.54直接得到存在矩阵 $\boldsymbol{A}$, 使得 $\boldsymbol{AB}=\boldsymbol{O}$, 因此齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系是 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$, 其解空间就是 V_0..

例题 3.29 设 $\boldsymbol{A}$ 是秩为 r 的 $m\times n$ 矩阵,$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r}$ 与 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{n-r}$ 是齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的两个基础解系. 求证: 必存在 $n-r$ 阶可逆矩阵 $\boldsymbol{P}$, 使得

$$(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{n-r})=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r})\boldsymbol{P}.$$

证明 设 U 是齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解空间, 则向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r}$ 与 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{n-r}$ 是 U 的两组基. 令 $\boldsymbol{P}$ 是这两组基之间的过渡矩阵, 则

$$(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{n-r})=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r})\boldsymbol{P}.$$

□

定理 3.15

1. 设 $\boldsymbol{A},\boldsymbol{B}$ 为 $m\times n$ 和 $m\times p$ 矩阵,$\boldsymbol{X}$ 为 $n\times p$ 未知矩阵, 证明: 矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 有解的充要条件是 $\mathrm{r}(\boldsymbol{A}\ \ \boldsymbol{B})=\mathrm{r}(\boldsymbol{A})$.
2. 设 $\boldsymbol{A},\boldsymbol{B}$ 为 $m\times n$ 和 $p\times n$ 矩阵,$\boldsymbol{X}$ 为 $p\times m$ 未知矩阵, 证明: 矩阵方程 $\boldsymbol{XA}=\boldsymbol{B}$ 有解的充要条件是 $\mathrm{r}\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}=\mathrm{r}(\boldsymbol{A})$.

♡

证明

1. 设 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n),\boldsymbol{B}=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_p),\boldsymbol{X}=(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_p)$ 为对应的列分块. 设 $\mathrm{r}(\boldsymbol{A})=r$ 且 $\boldsymbol{\alpha}_{i_1},\cdots,\boldsymbol{\alpha}_{i_r}$ 是 $\boldsymbol{A}$ 的列向量的极大无关组. 注意到矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 有解当且仅当 p 个线性方程组 $\boldsymbol{Ax}_i=\boldsymbol{\beta}_i(1\le i\le p)$ 都有解. 因此, 若 $\boldsymbol{AX}=\boldsymbol{B}$ 有解, 则每个 $\boldsymbol{\beta}_i$ 都是 $\boldsymbol{A}$ 的列向量的线性组合, 从而是 $\boldsymbol{\alpha}_{i_1},\cdots,\boldsymbol{\alpha}_{i_r}$ 的线性组合, 于是 $\boldsymbol{\alpha}_{i_1},\cdots,\boldsymbol{\alpha}_{i_r}$ 是 $(\boldsymbol{A}\,|\,\boldsymbol{B})$ 的列向量的极大无关组, 故 $\mathrm{r}(\boldsymbol{A}\,|\,\boldsymbol{B})=r$.
   反之, 若 $\mathrm{r}(\boldsymbol{A}\,|\,\boldsymbol{B})=r$, 则由命题 3.11(1)可知,$\boldsymbol{\alpha}_{i_1},\cdots,\boldsymbol{\alpha}_{i_r}$ 是 $(\boldsymbol{A}\,|\,\boldsymbol{B})$ 的列向量的极大无关组, 于是每个 $\boldsymbol{\beta}_i$ 都是 $\boldsymbol{A}$ 的列向量的线性组合, 从而 $\boldsymbol{AX}=\boldsymbol{B}$ 有解.
2. 线性方程组 $\boldsymbol{XA}=\boldsymbol{B}$ 两边取转置可得 $\boldsymbol{A}'\boldsymbol{X}'=\boldsymbol{B}'$, 从而
$$\text{线性方程组}\boldsymbol{XA}=\boldsymbol{B}\text{有解}\Leftrightarrow\text{线性方程组}\boldsymbol{A}'\boldsymbol{Y}=\boldsymbol{B}'\text{有解}.$$
   又由第一问可知
$$\text{线性方程组}\boldsymbol{A}'\boldsymbol{Y}=\boldsymbol{B}'\text{有解}\Leftrightarrow\mathrm{r}\begin{pmatrix}\boldsymbol{A}' & \boldsymbol{B}'\end{pmatrix}=\mathrm{r}(\boldsymbol{A}').$$
   而 $\mathrm{r}(\boldsymbol{A}')=\mathrm{r}(\boldsymbol{A})$,$\mathrm{r}\begin{pmatrix}\boldsymbol{A}' & \boldsymbol{B}'\end{pmatrix}=\mathrm{r}\left(\begin{pmatrix}\boldsymbol{A}' & \boldsymbol{B}'\end{pmatrix}'\right)=\mathrm{r}\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}$, 故
$$\text{线性方程组}\boldsymbol{XA}=\boldsymbol{B}\text{有解}\Leftrightarrow\mathrm{r}\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}=\mathrm{r}(\boldsymbol{A}).$$

□

命题 3.56

矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 有解当且仅当 p 个线性方程组 $\boldsymbol{Ax}_i=\boldsymbol{\beta}_i(1\le i\le p)$ 都有解. 从而每个 $\boldsymbol{\beta}_i$ 都是 $\boldsymbol{A}$ 的列向量的线性组合.

♠

证明 证明是显然的. □

命题 3.57

设 $\boldsymbol{A},\boldsymbol{B}$ 为 $m\times n$ 和 $n\times p$ 矩阵, 证明: 存在 $p\times n$ 矩阵 $\boldsymbol{C}$, 使得 $\boldsymbol{ABC}=\boldsymbol{A}$ 的充要条件是 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{AB})$.

证明 必要性由秩的不等式 $\mathrm{r}(\boldsymbol{A})\geq \mathrm{r}(\boldsymbol{AB})\geq \mathrm{r}(\boldsymbol{ABC})=\mathrm{r}(\boldsymbol{A})$ 即得.

充分性由秩的不等式可知 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{AB})\leq \mathrm{r}(\boldsymbol{AB}\mid\boldsymbol{B})=\mathrm{r}(\boldsymbol{A}(\boldsymbol{B}\mid\boldsymbol{I}_n))\leq \mathrm{r}(\boldsymbol{A})$. 故 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{AB}\mid\boldsymbol{B})$. 于是由定理 3.15可知, 矩阵方程 $\boldsymbol{ABX}=\boldsymbol{A}$ 有解. 即存在 $p\times n$ 矩阵 $\boldsymbol{C}$, 使得 $\boldsymbol{ABC}=\boldsymbol{A}$ 的充要条件是 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{AB})$. □

3.8.2 线性方程组的公共解

对两个非齐次线性方程组, 若只已知它们的通解, 而不知道方程组本身, 要求它们的公共解, 我们可以这样来做: 设 $\boldsymbol{Ax}=\boldsymbol{\beta}_1,\boldsymbol{Bx}=\boldsymbol{\beta}_2$ 是两个含 n 个未知数的非齐次线性方程组. 方程组 $\boldsymbol{Ax}=\boldsymbol{\beta}_1$ 有特解 $\boldsymbol{\gamma}$ 且 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系为 $\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_{n-r}$. 方程组 $\boldsymbol{Bx}=\boldsymbol{\beta}_2$ 有特解 $\boldsymbol{\delta}$ 且 $\boldsymbol{Bx}=\boldsymbol{0}$ 的基础解系为 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_{n-s}$.

方法一: 假设它们的公共解为 $\boldsymbol{\gamma}+t_1\boldsymbol{\eta}_1+\cdots+t_{n-r}\boldsymbol{\eta}_{n-r}$, 则 $\boldsymbol{\gamma}+t_1\boldsymbol{\eta}_1+\cdots+t_{n-r}\boldsymbol{\eta}_{n-r}-\boldsymbol{\delta}$ 是 $\boldsymbol{Bx}=\boldsymbol{0}$ 的解, 因此可以表示为 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_{n-s}$ 的线性组合. 于是矩阵 $(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_{n-s},\boldsymbol{\gamma}+t_1\boldsymbol{\eta}_1+\cdots+t_{n-r}\boldsymbol{\eta}_{n-r}-\boldsymbol{\delta})$ 的秩等于 $n-s$. 由此可以求出 $t_1,\cdots,t_{n-r}$, 从而求出公共解.

方法二:假设它们的公共解为 $\boldsymbol{\zeta}$, 则

$$\boldsymbol{\zeta}=\boldsymbol{\gamma}+t_1\boldsymbol{\eta}_1+\cdots+t_{n-r}\boldsymbol{\eta}_{n-r}=\boldsymbol{\delta}+(-u_1)\boldsymbol{\xi}_1+\cdots+(-u_{n-s})\boldsymbol{\xi}_{n-s}.$$

要求公共解 $\boldsymbol{\zeta}$ 等价于求解下列关于未定元 $t_1,\cdots,t_{n-r};u_1,\cdots,u_{n-s}$ 的线性方程组:

$$t_1\boldsymbol{\eta}_1+\cdots+t_{n-r}\boldsymbol{\eta}_{n-r}+u_1\boldsymbol{\xi}_1+\cdots+u_{n-s}\boldsymbol{\xi}_{n-s}=\boldsymbol{\delta}-\boldsymbol{\gamma}.$$

例题 3.30 设有两个非齐次线性方程组 (I), (II), 它们的通解分别为

$$\boldsymbol{\gamma}+t_1\boldsymbol{\eta}_1+t_2\boldsymbol{\eta}_2;\boldsymbol{\delta}+k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2,$$

其中 $\boldsymbol{\gamma}=(5,-3,0,0)',\boldsymbol{\eta}_1=(-6,5,1,0)',\boldsymbol{\eta}_2=(-5,4,0,1)';\boldsymbol{\delta}=(-11,3,0,0)',\boldsymbol{\xi}_1=(8,-1,1,0)',\boldsymbol{\xi}_2=(10,-2,0,1)'$. 求这两个方程组的公共解.

证明 解法一: 设公共解为

$$\boldsymbol{\gamma}+t_1\boldsymbol{\eta}_1+t_2\boldsymbol{\eta}_2=\begin{pmatrix}5-6t_1-5t_2\\-3+5t_1+4t_2\\t_1\\t_2\end{pmatrix}.$$

注意矩阵 $(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\gamma}-\boldsymbol{\delta}+t_1\boldsymbol{\eta}_1+t_2\boldsymbol{\eta}_2)$ 的秩等于 2, 对此矩阵作初等行变换:

$$\begin{pmatrix}8&10&16-6t_1-5t_2\\-1&-2&-6+5t_1+4t_2\\1&0&t_1\\0&1&t_2\end{pmatrix}\to\begin{pmatrix}1&0&t_1\\0&1&t_2\\8&10&16-6t_1-5t_2\\-1&-2&-6+5t_1+4t_2\end{pmatrix}\to\begin{pmatrix}1&0&t_1\\0&1&t_2\\0&0&16-14t_1-15t_2\\0&0&-6+6t_1+6t_2\end{pmatrix}$$

可得关于 t_1,t_2 的方程组

$$\begin{cases}14t_1+15t_2=16,\\6t_1+6t_2=6.\end{cases}$$

解得 $t_1=-1,t_2=2$, 所以公共解为 (只有一个向量)$\boldsymbol{\gamma}-\boldsymbol{\eta}_1+2\boldsymbol{\eta}_2=(1,0,-1,2)'$.

解法二: 求公共解等价于求解下列线性方程组:

$$t_1\boldsymbol{\eta}_1+t_2\boldsymbol{\eta}_2+u_1\boldsymbol{\xi}_1+u_2\boldsymbol{\xi}_2=\boldsymbol{\delta}-\boldsymbol{\gamma}.$$

对其增广矩阵实施初等行变换, 可得

$$\begin{pmatrix} -6 & -5 & 8 & 10 & -16 \\ 5 & 4 & -1 & -2 & 6 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix},$$

故 (t_1, t_2, u_1, u_2) 只有唯一解 $(-1, 2, 1, -2)$. 因此, 公共解为 $\boldsymbol{\gamma} - \boldsymbol{\eta}_1 + 2\boldsymbol{\eta}_2 = \boldsymbol{\delta} - \boldsymbol{\xi}_1 + 2\boldsymbol{\xi}_2 = (1, 0, -1, 2)'$. □

例题 3.31 设有非齐次线性方程组 (I):

$$\begin{cases} 7x_1 - 6x_2 + 3x_3 = b, \\ 8x_1 - 9x_2 + ax_4 = 7. \end{cases}$$

又已知方程组 (II) 的通解为

$$(1, 1, 1, 0)' + t_1(1, 0, -1, 0)' + t_2(2, 3, 0, 1)'.$$

若这两个方程组有无穷多组公共解, 求出 a, b 的值并求出公共解.

证明 将 (II) 的通解写为 $(1 + t_1 + 2t_2, 1 + 3t_2, -t_1, t_2)'$, 代入方程组 (I) 化简得到

$$\begin{cases} 4t_1 - 4t_2 = b - 1, \\ 8t_1 + (a - 11)t_2 = 8. \end{cases}$$

要使这两个方程组有无穷多组公共解,t_1, t_2 必须有无穷多组解, 于是上面方程组的系数矩阵和增广矩阵的秩都应该等于 1, 从而 $a = 3, b = 5$. 解出方程组得到 $t_1 = t_2 + 1$, 因此方程组 (I), (II) 的公共解为

$$(1 + t_1 + 2t_2, 1 + 3t_2, -t_1, t_2)' = (2, 1, -1, 0)' + t_2(3, 3, -1, 1)',$$

其中 t_2 为任意数. □

3.8.3 在解析几何上的应用

命题 3.58

求平面上 n 个点 $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ 位于同一条直线上的充要条件.

证明 充要条件为第一个点和其余点代表的向量之差属于一个一维子空间, 即 $(x_i - x_1, y_i - y_1)$ 都成比例. 写成矩阵形式为

$$\mathrm{r}\begin{pmatrix} x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \\ y_2 - y_1 & y_3 - y_1 & \cdots & y_n - y_1 \end{pmatrix} \leq 1,$$

或

$$\mathrm{r}\begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ y_1 & y_2 & y_3 & \cdots & y_n \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \leq 2.$$

□

命题 3.59

求三维实空间中 4 点 $(x_i, y_i, z_i)(1 \leq i \leq 4)$ 共面的充要条件.

证明 设 4 点的向量为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4$, 则 4 点共面的充要条件是: 向量组 $\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_3 - \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_4 - \boldsymbol{\alpha}_1$ 的秩不超过 2. 不难将此写成矩阵形式:

$$\mathrm{r}\begin{pmatrix} x_2 - x_1 & x_3 - x_1 & x_4 - x_1 \\ y_2 - y_1 & y_3 - y_1 & y_4 - y_1 \\ z_2 - z_1 & z_3 - z_1 & z_4 - z_1 \end{pmatrix} \leq 2,$$

或

$$\mathrm{r}\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \leq 3.$$

□

例题 3.32 证明: 通过平面内不在一条直线上的 3 点 $(x_1,y_1),(x_2,y_2),(x_3,y_3)$ 的圆方程为

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

证明 圆方程可设为

$$u_1(x^2+y^2)+u_2x+u_3y+u_4=0,$$

于是得到未知数 u_1,u_2,u_3,u_4 的方程组为

$$\begin{cases} (x_1^2+y_1^2)u_1+x_1u_2+y_1u_3+u_4=0, \\ (x_2^2+y_2^2)u_1+x_2u_2+y_2u_3+u_4=0, \\ (x_3^2+y_3^2)u_1+x_3u_2+y_3u_3+u_4=0. \end{cases}$$

上述方程组加上原方程组成一个含 4 个未知数、4 个方程式的齐次线性方程组, 它有非零解的充要条件是系数行列式等于零, 即

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

由命题 3.58可知 3 点不在一条直线上意味着

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

故圆方程不退化. □

命题 3.60

求平面上不在一条直线上的 4 个点 $(x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4)$ 位于同一个圆上的充要条件.

证明 由例题 3.32可得充要条件为

$$\begin{vmatrix} x_1^2+y_1^2 & x_1 & y_1 & 1 \\ x_2^2+y_2^2 & x_2 & y_2 & 1 \\ x_3^2+y_3^2 & x_3 & y_3 & 1 \\ x_4^2+y_4^2 & x_4 & y_4 & 1 \end{vmatrix} = 0.$$

□

例题 3.33 已知平面上两条不同的二次曲线 $a_ix^2+b_ixy+c_iy^2+d_ix+e_iy+f_i=0(i=1,2)$ 交于 4 个不同的点 $(x_i,y_i)(1\leq i\leq 4)$. 求证: 过这 4 个点的二次曲线均可写为如下形状:

$$\lambda_1(a_1x^2+b_1xy+c_1y^2+d_1x+e_1y+f_1)+\lambda_2(a_2x^2+b_2xy+c_2y^2+d_2x+e_2y+f_2)=0.$$

证明 显然上述曲线过这 4 个交点. 现设 $ax^2+bxy+cy^2+dx+ey+f=0$ 是过这 4 个交点的二次曲线, 则有

$$\begin{cases} ax_1^2+bx_1y_1+cy_1^2+dx_1+ey_1+f=0, \\ ax_2^2+bx_2y_2+cy_2^2+dx_2+ey_2+f=0, \\ ax_3^2+bx_3y_3+cy_3^2+dx_3+ey_3+f=0, \\ ax_4^2+bx_4y_4+cy_4^2+dx_4+ey_4+f=0. \end{cases} \tag{3.9}$$

视 a,b,c,d,e,f 为未知数, 则线性方程组 (3.9) 有线性无关的解 $(a_1,b_1,c_1,d_1,e_1,f_1)',(a_2,b_2,c_2,d_2,e_2,f_2)'$. 如果能证明方程组 (3.9) 的系数矩阵的秩等于 4, 则这两个解就构成了基础解系, 从而即得结论.

容易验证 4 个交点中的任意 3 个点都不共线, 而且经过坐标轴适当的旋转, 可以假设这 4 个交点的横坐标 x_1,x_2,x_3,x_4 互不相同. 用反证法证明结论, 设方程组 (3.9) 系数矩阵 $\boldsymbol{A}$ 的秩小于 4. 由任意 3 个交点不共线以及命题 3.58知,$(x_1,x_2,x_3,x_4)',(y_1,y_2,y_3,y_4)',(1,1,1,1)'$ 线性无关, 从而它们是 $\boldsymbol{A}$ 的列向量的极大无关组, 于是 $(x_1^2,x_2^2,x_3^2,x_4^2)'$ 是它们的线性组合, 故可设 $x_i^2=rx_i+sy_i+t(1\leq i\leq 4)$, 其中 r,s,t 是实数. 由于 x_1,x_2,x_3,x_4 互不相同, 故 $s\neq 0$, 于是 $y_i=\dfrac{1}{s}x_i^2-\dfrac{r}{s}x_i-\dfrac{t}{s}(1\leq i\leq 4)$. 考虑 $\boldsymbol{A}$ 的第一列、第二列、第四列和第六列构成的四阶行列式 $|\boldsymbol{B}|$, 利用 Vander - monde 行列式容易算出 $|\boldsymbol{B}|=-\dfrac{1}{s}\prod\limits_{1\leq i<j\leq 4}(x_i-x_j)\neq 0$, 于是 $\boldsymbol{A}$ 的秩等于 4, 这与假设矛盾. 因此方程组 (3.9) 的系数矩阵的秩只能等于 4. □

第四章 线性映射

4.1 基本定理和命题

命题 4.1 (满射的复合仍是满射)

若 f,g 均为满射, 则 $f\circ g$ 也是满射.

注 单射的复合不一定是单射.

证明 设 $f:V_1\to V_2, g:V_2\to V_3$, 则对 $\forall\alpha\in V_3$, 由 g 为满射可知, 存在 $\beta\in V_2$, 使得 $\alpha=g(\beta)$. 又由 f 为满射可知, 存在 $\gamma\in V_1$, 使得 $\beta=f(\gamma)$. 从而 $(g\circ f)(\gamma)=g(f(\gamma))=g(\beta)=\alpha$. 故 $g\circ f$ 也是满射. □

定义 4.1 (线性映射的表示矩阵)

设 φ 是 $V\to U$ 的线性映射, 分别取 V 和 U 的基如下:

$$V:\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n;\ U:\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_m.$$

假设有

$$\begin{cases}\varphi(\boldsymbol{e}_1)=a_{11}\boldsymbol{f}_1+a_{21}\boldsymbol{f}_2+\cdots+a_{m1}\boldsymbol{f}_m,\\ \varphi(\boldsymbol{e}_2)=a_{12}\boldsymbol{f}_1+a_{22}\boldsymbol{f}_2+\cdots+a_{m2}\boldsymbol{f}_m,\\ \cdots\cdots\cdots\\ \varphi(\boldsymbol{e}_n)=a_{1n}\boldsymbol{f}_1+a_{2n}\boldsymbol{f}_2+\cdots+a_{mn}\boldsymbol{f}_m,\end{cases}$$

则矩阵

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

称为线性映射 φ 在基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 和 $\left\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_m\right\}$ 下的表示矩阵.

注 若 φ 是向量空间 V 上的线性变换, 则取 V 的一组基, 而不取两组基.

推论 4.1

设 φ 是 $V\to U$ 的线性映射, 分别取 V 和 U 的基如下:

$$V:\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n;\quad U:\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_m.$$

并且设 φ 在基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 和 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_m\}$ 下的表示矩阵为 $\boldsymbol{A}$. 则有

$$(\varphi(\boldsymbol{e}_1),\varphi(\boldsymbol{e}_2),\cdots,\varphi(\boldsymbol{e}_n))=(\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_m)\boldsymbol{A}.$$

$\forall\boldsymbol{\alpha}\in V$, 设 $\boldsymbol{\alpha}$ 在基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下的坐标向量为 $(x_1,x_2,\cdots,x_n)'$, 则

$$\boldsymbol{\alpha}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_n\boldsymbol{e}_n=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},$$

$$\varphi(\boldsymbol{\alpha}) = \varphi(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n) = (\varphi(\boldsymbol{e}_1), \varphi(\boldsymbol{e}_2), \cdots, \varphi(\boldsymbol{e}_n))\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_m)\boldsymbol{A}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

即 $\varphi(\boldsymbol{\alpha})$ 在基 $\{\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_m\}$ 下的坐标向量为 $\boldsymbol{A}(x_1, x_2, \cdots, x_n)'$.

定理 4.1

设 V 是数域 $\mathbb{F}$ 上的 n 维向量空间,φ 是 V 上的线性变换,$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 和 $\{\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n\}$ 是 V 的两组基, 从第一组基到第二组基的过渡矩阵为 $\boldsymbol{P}$. 假设 φ 在第一组基下的表示矩阵为 $\boldsymbol{A}$, 在第二组基下的表示矩阵为 $\boldsymbol{B}$, 则 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 即向量空间上同一个线性变换在不同基下的表示矩阵必相似.

4.2 线性映射及其运算

在许多问题中, 常常需要定义向量空间之间的线性映射 (或某一向量空间上的线性变换). 一般来说, 无须对向量空间中的每个元素进行定义, 我们可采用下列两种方法来简化定义: 第一, 只要对向量空间的基向量进行定义即可; 第二, 若向量空间可分解为两个 (或多个) 子空间的直和, 则只要对每个子空间进行定义即可. 这两点可由下面两个定理 (定理 4.2和定理定理 4.3) 得到.

定理 4.2 (线性扩张定理)

设 V 和 U 是数域 $\mathbb{F}$ 上的向量空间,$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是 V 的一组基,$\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 是 U 中 n 个向量, 求证: 存在唯一的 V 到 U 的线性映射 φ, 使得 $\varphi(\boldsymbol{e}_i) = \boldsymbol{u}_i$.

注 这个线性扩张定理表明只要选定 V 的一组基和 U 中 n 个向量, 则有且仅有一个线性映射将基向量映到对应的向量. 后面我们将经常采用线性扩张定理来构造线性映射以及判定两个线性映射是否相等.

笔记 这个线性扩张定理表明:

1. 在向量空间上定义线性映射只要对向量空间的基向量进行定义即可.
2. 判定两个线性映射是否相等只要判断这个两个线性映射原像空间的基向量的像是否相等即可.

证明 先证存在性. 对任意的 $\boldsymbol{\alpha} \in V$, 设 $\boldsymbol{\alpha} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n$, 则 $a_1, a_2, \cdots, a_n$ 被 $\boldsymbol{\alpha}$ 唯一确定. 令

$$\varphi(\boldsymbol{\alpha}) = a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + \cdots + a_n\boldsymbol{u}_n,$$

则 φ 是 V 到 U 的映射. 若另有 $\boldsymbol{\beta} = b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2 + \cdots + b_n\boldsymbol{e}_n$, 则

$$\varphi(\boldsymbol{\alpha} + \boldsymbol{\beta}) = (a_1 + b_1)\boldsymbol{u}_1 + (a_2 + b_2)\boldsymbol{u}_2 + \cdots + (a_n + b_n)\boldsymbol{u}_n = \varphi(\boldsymbol{\alpha}) + \varphi(\boldsymbol{\beta}).$$

又对 $\mathbb{F}$ 中的任意元素 k, 有

$$\varphi(k\boldsymbol{\alpha}) = ka_1\boldsymbol{u}_1 + ka_2\boldsymbol{u}_2 + \cdots + ka_n\boldsymbol{u}_n = k\varphi(\boldsymbol{\alpha}).$$

因此 φ 是线性映射, 显然它满足 $\varphi(\boldsymbol{e}_i) = \boldsymbol{u}_i$.

设另有 V 到 U 的线性映射 ψ 满足 $\psi(\boldsymbol{e}_i) = \boldsymbol{u}_i$, 则对任意的 $\boldsymbol{\alpha} \in V$, 有

$$\begin{aligned}\psi(\boldsymbol{\alpha}) &= \psi(a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n) \\ &= a_1\psi(\boldsymbol{e}_1) + a_2\psi(\boldsymbol{e}_2) + \cdots + a_n\psi(\boldsymbol{e}_n) \\ &= a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + \cdots + a_n\boldsymbol{u}_n = \varphi(\boldsymbol{\alpha}).\end{aligned}$$

因此 $\psi = \varphi$, 这就证明了唯一性. □

定理 4.3

设线性空间 $V = V_1 \oplus V_2$, 并且 φ_1 及 φ_2 分别是 V_1, V_2 到 U 的线性映射, 求证: 存在唯一的从 V 到 U 的线性映射 φ, 当 φ 限制在 V_i 上时等于 φ_i.

注 定理 4.3可以推广到多个子空间的情形: 设 $V = V_1 \oplus \cdots \oplus V_m$, 给定线性映射 $\varphi_i : V_i \to U(1 \le i \le m)$, 则存在唯一的线性映射 $\varphi : V \to U$, 使得 $\varphi|_{V_i} = \varphi_i(1 \le i \le m)$. 我们可以把这样的线性映射 φ 简记为 $\varphi_1 \oplus \cdots \oplus \varphi_m$.

笔记 这个定理表明: 在向量空间上定义线性映射, 若这个向量空间可分解为两个 (或多个) 子空间的直和, 则只要对每个子空间进行定义即可.

证明 因为 $V = V_1 \oplus V_2$, 故对任意的 $\alpha \in V$,α 可唯一地写为 $\alpha = \alpha_1 + \alpha_2$, 其中 $\alpha_1 \in V_1, \alpha_2 \in V_2$. 令 $\varphi(\alpha) = \varphi_1(\alpha_1) + \varphi_2(\alpha_2)$, 则 φ 是 V 到 U 的映射. 不难验证 φ 保持加法和数乘, 因此 φ 是线性映射. 若另有线性映射 ψ, 它在 V_i 上的限制等于 φ_i, 则

$$\psi(\alpha) = \psi(\alpha_1) + \psi(\alpha_2) = \varphi_1(\alpha_1) + \varphi_2(\alpha_2) = \varphi(\alpha).$$

因此 $\psi = \varphi$, 唯一性得证. □

例题 4.1 设 φ 是有限维线性空间 V 到 U 的线性映射, 求证: 必存在 U 到 V 的线性映射 ψ, 使得 $\varphi\psi\varphi = \varphi$.

笔记 可以直接定义 ψ 是 U 到 V 的线性映射的原因: 由线性扩张定理可知存在唯一的线性映射 ψ, 使得它在基上的作用为

$$\psi(\boldsymbol{f}_i) = \boldsymbol{e}_i, 1 \le i \le r; \psi(\boldsymbol{f}_j) = \boldsymbol{0}, r+1 \le j \le m.$$

以后这种利用线性扩张定理得到线性映射的存在性不再额外说明, 而是直接定义.

证明 证法一: 设 V 和 U 的维数分别是 n 和 m. 由命题 4.11可知, 存在 V 和 U 的基 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}, \{\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_m\}$, 使得 φ 在这两组基下的表示矩阵为

$$\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}.$$

这就是 $\varphi(\boldsymbol{e}_i) = \boldsymbol{f}_i, 1 \le i \le r; \varphi(\boldsymbol{e}_j) = \boldsymbol{0}, r+1 \le j \le n$.定义 ψ 是 U 到 V 的线性映射, 它在基上的作用为

$$\psi(\boldsymbol{f}_i) = \boldsymbol{e}_i, 1 \le i \le r; \psi(\boldsymbol{f}_j) = \boldsymbol{0}, r+1 \le j \le m,$$

则在 V 的基上, 有

$$\varphi\psi\varphi(\boldsymbol{e}_i) = \varphi\psi(\boldsymbol{f}_i) = \varphi(\boldsymbol{e}_i), 1 \le i \le r;$$
$$\varphi\psi\varphi(\boldsymbol{e}_j) = \varphi\psi(\boldsymbol{0}) = \boldsymbol{0} = \varphi(\boldsymbol{e}_j), r+1 \le j \le n.$$

于是 $\varphi\psi\varphi = \varphi$.

证法二 (代数方法):取定 V 和 U 的两组基, 设 φ 在这两组基下的表示矩阵为 $m \times n$ 矩阵 $\boldsymbol{A}$, 则由例题 3.23可知, 存在 $n \times m$ 矩阵 $\boldsymbol{B}$, 使得 $\boldsymbol{ABA} = \boldsymbol{A}$. 由矩阵 $\boldsymbol{B}$ 可定义从 U 到 V 的线性映射 ψ, 它适合 $\varphi\psi\varphi = \varphi$. □

命题 4.2

设有数域 $\mathbb{F}$ 上的有限维线性空间 V, V', 又 U 是 V 的子空间,φ 是 U 到 V' 的线性映射. 求证: 必存在 V 到 V' 的线性映射 ψ, 它在 U 上的限制就是 φ.

笔记 可以直接定义 ψ 为 V 到 V' 的线性映射的原因: 由定理 4.3可知, 存在唯一的从 V 到 V' 的线性映射 ψ, 使得它在 U 上的限制是 φ, 它在 W 上的限制是零线性映射.

以后这种利用定理 4.3得到线性映射的存在性不再额外说明, 而是直接定义.

证明 令 W 是子空间 U 在 V 中的补空间, 即 $V = U \oplus W$. 定义 ψ 为 V 到 V' 的线性映射, 它在 U 上的限制是 φ, 它在 W 上的限制是零线性映射, 这样的 ψ 即为所求. □

命题 4.3

设 V,U 是 $\mathbb{F}$ 上的有限维线性空间,φ 是 V 到 U 的线性映射, 求证:

(1) φ 是单映射的充要条件是存在 U 到 V 的线性映射 ψ, 使 $\psi\varphi=\mathrm{Id}_V$, 这里 Id_V 表示 V 上的恒等映射;

(2) φ 是满映射的充要条件是存在 U 到 V 的线性映射 η, 使 $\varphi\eta=\mathrm{Id}_U$, 这里 Id_U 表示 U 上的恒等映射.

证明 证法一:

(1) 若 $\psi\varphi=\mathrm{Id}_V$, 则对任意的 $\boldsymbol{v}\in\mathrm{Ker}\varphi,\boldsymbol{v}=\psi(\varphi(\boldsymbol{v}))=\mathbf{0}$, 即 $\mathrm{Ker}\varphi=0$, 从而 φ 是单映射.

反之, 若 φ 是单映射, 则定义映射 $\varphi_1:V\to\mathrm{Im}\varphi$, 它与 φ 有相同的映射法则, 但值域变为 $\mathrm{Im}\varphi$. 因为 φ_1 与 φ 有相同的映射法则, 所以对 $\forall\alpha\in\mathrm{Ker}\varphi_1$, 有 $\varphi(\alpha)=\varphi_1(\alpha)=0$. 于是 $\alpha\in\mathrm{Ker}\varphi$, 故 $\mathrm{Ker}\varphi_1\subset\mathrm{Ker}\varphi$. 又由 φ 是单射可知,$\mathrm{Ker}\varphi=0$. 因此 $\mathrm{Ker}\varphi_1=0$, 即 φ_1 也是单射. 因为 φ_1 与 φ 有相同的映射法则, 所以对 $\forall\beta\in\mathrm{Im}\varphi$, 存在 $b\in V$, 使得 $\beta=\varphi(b)=\varphi_1(b)\in\mathrm{Im}\varphi_1$. 故 $\mathrm{Im}\varphi\subset\mathrm{Im}\varphi_1$. 又根据 φ_1 的定义可知,$\mathrm{Im}\varphi_1\subset\mathrm{Im}\varphi$. 因此 $\mathrm{Im}\varphi=\mathrm{Im}\varphi_1$, 即 φ_1 是满射. 故 φ_1 是双射. 由于 φ_1 与 φ 有相同的映射法则, 因此容易验证 φ_1 是线性映射. 综上,φ_1 是线性同构.

设 U_0 是 $\mathrm{Im}\varphi$ 在 U 中的补空间, 即 $U=\mathrm{Im}\varphi\oplus U_0$. 定义 ψ 为 U 到 V 的线性映射, 它在 $\mathrm{Im}\varphi$ 上的限制为 φ_1^{-1}, 它在 U_0 上的限制是零线性映射, 则容易验证 $\psi\varphi=\mathrm{Id}_V$ 成立.

(2) 若 $\varphi\eta=\mathrm{Id}_U$, 则对任意的 $\boldsymbol{u}\in U,\boldsymbol{u}=\varphi(\eta(\boldsymbol{u}))$, 从而 φ 是满映射.

反之, 若 φ 是满映射, 则可取 U 的一组基 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_m$, 一定存在 V 中的向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_m$, 使得 $\varphi(\boldsymbol{v}_i)=\boldsymbol{f}_i(1\le i\le m)$. 定义 η 为 U 到 V 的线性映射, 它在基上的作用为 $\eta(\boldsymbol{f}_i)=\boldsymbol{v}_i(1\le i\le m)$, 则容易验证 $\varphi\eta=\mathrm{Id}_U$ 成立.

证法二 (代数方法):

充分性同证法一, 现只证必要性. 取定 V 和 U 的两组基, 设 φ 在这两组基下的表示矩阵为 $m\times n$ 矩阵 $\boldsymbol{A}$.

(1) 若 φ 是单映射, 则由命题 4.13可知 $\boldsymbol{A}$ 是列满秩矩阵. 再由命题 3.49(1)可知, 存在 $n\times m$ 矩阵 $\boldsymbol{B}$, 使得 $\boldsymbol{BA}=\boldsymbol{I}_n$. 由矩阵 $\boldsymbol{B}$ 可定义从 U 到 V 的线性映射 ψ, 它适合 $\psi\varphi=\mathrm{Id}_V$.

(2) 若 φ 是满映射, 则由命题 4.13可知 $\boldsymbol{A}$ 是行满秩矩阵. 再由命题 3.49(2)可知, 存在 $n\times m$ 矩阵 $\boldsymbol{C}$, 使得 $\boldsymbol{AC}=\boldsymbol{I}_m$. 由矩阵 $\boldsymbol{C}$ 可定义从 U 到 V 的线性映射 η, 它适合 $\varphi\eta=\mathrm{Id}_U$.

□

命题 4.4

(1) 设 V,U 是数域 $\mathbb{K}$ 上的有限维线性空间,$\varphi,\psi:V\to U$ 是两个线性映射, 证明: 存在 U 上的线性变换 ξ, 使得 $\psi=\xi\varphi$ 成立的充要条件是 $\mathrm{Ker}\varphi\subseteq\mathrm{Ker}\psi$.

(2) 设 V,U 是数域 $\mathbb{K}$ 上的有限维线性空间,$\varphi,\psi:V\to U$ 是两个线性映射, 证明: 存在 V 上的线性变换 ξ, 使得 $\psi=\varphi\xi$ 成立的充要条件是 $\mathrm{Im}\psi\subseteq\mathrm{Im}\varphi$.

证明

(1) 先证必要性: 任取 $\boldsymbol{v}\in\mathrm{Ker}\varphi$, 则 $\psi(\boldsymbol{v})=\xi\varphi(\boldsymbol{v})=\mathbf{0}$, 即有 $\boldsymbol{v}\in\mathrm{Ker}\psi$, 从而 $\mathrm{Ker}\varphi\subseteq\mathrm{Ker}\psi$.

再证充分性: 设 $\dim V=n,\dim U=m,\dim\mathrm{Ker}\varphi=n-r$. 取 $\mathrm{Ker}\varphi$ 的一组基 $\boldsymbol{e}_{r+1},\cdots,\boldsymbol{e}_n$, 扩张为 V 的一组基 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_r,\boldsymbol{e}_{r+1},\cdots,\boldsymbol{e}_n$. 由推论 4.4可知,$\varphi(\boldsymbol{e}_1),\cdots,\varphi(\boldsymbol{e}_r)$ 是 $\mathrm{Im}\varphi$ 的一组基, 将其扩张为 U 的一组基 $\varphi(\boldsymbol{e}_1),\cdots,\varphi(\boldsymbol{e}_r),\boldsymbol{g}_{r+1},\cdots,\boldsymbol{g}_m$. 定义 ξ 为 U 上的线性变换, 它在基上的作用为:$\xi(\varphi(\boldsymbol{e}_i))=\psi(\boldsymbol{e}_i)(1\le i\le r),\xi(\boldsymbol{g}_j)=\mathbf{0}(r+1\le j\le m)$. 由于 $\mathrm{Ker}\varphi\subseteq\mathrm{Ker}\psi$, 故容易验证 $\psi(\boldsymbol{e}_i)=\xi\varphi(\boldsymbol{e}_i)(1\le i\le n)$ 成立, 从而 $\psi=\xi\varphi$.

(2) 先证必要性: 任取 $\boldsymbol{v}\in V$, 则 $\psi(\boldsymbol{v})=\varphi(\xi(\boldsymbol{v}))\in\mathrm{Im}\varphi$, 从而 $\mathrm{Im}\psi\subseteq\mathrm{Im}\varphi$.

再证充分性: 取 V 的一组基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$, 则 $\psi(\boldsymbol{e}_i)\in\mathrm{Im}\psi\subseteq\mathrm{Im}\varphi$, 从而存在 $\boldsymbol{f}_i\in V$, 使得 $\varphi(\boldsymbol{f}_i)=\psi(\boldsymbol{e}_i)(1\le i\le n)$. 定义 ξ 为 V 上的线性变换, 它在基上的作用为:$\xi(\boldsymbol{e}_i)=\boldsymbol{f}_i(1\le i\le n)$. 容易验证 $\psi(\boldsymbol{e}_i)=\varphi\xi(\boldsymbol{e}_i)(1\le i\le n)$ 成立, 从而 $\psi=\varphi\xi$.

□

命题 4.5

设 φ 是 n 维线性空间 V 上的线性变换,$\alpha \in V$. 若 $\varphi^{m-1}(\alpha) \neq \mathbf{0}$, 而 $\varphi^m(\alpha) = \mathbf{0}$, 求证:$\alpha, \varphi(\alpha), \varphi^2(\alpha), \cdots, \varphi^{m-1}(\alpha)$ 线性无关.

证明 设有 m 个数 $a_0, a_1, \cdots, a_{m-1}$, 使得

$$a_0\alpha + a_1\varphi(\alpha) + \cdots + a_{m-1}\varphi^{m-1}(\alpha) = \mathbf{0}.$$

上式两边同时作用 φ^{m-1}, 则有 $a_0\varphi^{m-1}(\alpha) = \mathbf{0}$, 由于 $\varphi^{m-1}(\alpha) \neq \mathbf{0}$, 故 $a_0 = 0$. 上式两边同时作用 φ^{m-2}, 则有 $a_1\varphi^{m-1}(\alpha) = \mathbf{0}$, 由于 $\varphi^{m-1}(\alpha) \neq \mathbf{0}$, 故 $a_1 = 0$. 不断这样做下去, 最后可得 $a_0 = a_1 = \cdots = a_{m-1} = 0$, 于是 $\alpha, \varphi(\alpha), \varphi^2(\alpha), \cdots, \varphi^{m-1}(\alpha)$ 线性无关. □

推论 4.2

设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,φ 是 V 上的幂零线性变换, 满足 $\mathrm{r}(\varphi) = n - 1$. 求证: 存在 V 的一组基, 使得 φ 在这组基下的表示矩阵为

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

证明 由假设存在正整数 m, 使得 $\varphi^m = 0, \varphi^{m-1} \neq 0$, 从而存在 $\alpha \in V$, 使得 $\varphi^m(\alpha) = 0, \varphi^{m-1}(\alpha) \neq 0$. 由命题 4.5可知,$\alpha, \varphi(\alpha), \cdots, \varphi^{m-1}(\alpha)$ 线性无关, 于是 $m \leq \dim V = n$. 另一方面, 由Sylvester 不等式以及 $\mathrm{r}(\varphi) = n - 1$ 可知,$\mathrm{r}(\varphi^2) \geq 2\mathrm{r}(\varphi) - n = n - 2$. 不断这样讨论下去, 最终可得 $0 = \mathrm{r}(\varphi^m) \geq n - m$, 即有 $m \geq n$, 从而 $m = n$. 于是 $\alpha, \varphi(\alpha), \cdots, \varphi^{n-1}(\alpha)$ 是 V 的一组基,φ 在这组基下的表示矩阵为 $\boldsymbol{A}$. □

4.3 线性同构

线性同构刻画了不同线性空间之间的相同本质, 即同构的线性空间具有相同的线性结构 (或从线性结构的观点来看没有任何区别). 要证明线性映射 $\varphi : V \to U$ 是线性同构, 通常一方面需要验证 φ 是单映射 (或等价地验证 $\mathrm{Ker}\varphi = 0$), 另一方面需要验证 φ 是满映射 (或等价地验证 $\mathrm{Im}\varphi = U$). 但若已知前后两个线性空间的维数相等, 则由线性映射的维数公式容易证明,φ 是线性同构当且仅当 φ 是单映射, 也当且仅当 φ 是满映射, 从而只需验证 φ 是单映射或满映射即可得到 φ 是线性同构.

命题 4.6

设 V, U 为两个线性空间, 若 $\dim V = \dim U$, 则线性映射 $\varphi : V \to U$ 是线性同构当且仅当 φ 是单映射, 也当且仅当 φ 是满映射.

证明 由线性映射的维数公式容易证明. □

推论 4.3

设 V 为线性空间, 则线性变换 $\varphi : V \to V$ 是自同构当且仅当 φ 是单映射, 也当且仅当 φ 是满映射.

证明 由命题 4.6立得. □

引理 4.1

设 $a_0, a_1, \cdots, a_n$ 是数域 $\mathbb{F}$ 中 $n+1$ 个不同的数,V 是 $\mathbb{F}$ 上次数不超过 n 的多项式全体组成的线性空间. 设 φ 是 V 到 $n+1$ 维行向量空间 U 的映射:

$$\varphi(f) = (f(a_0), f(a_1), \cdots, f(a_n)),$$

求证:φ 是线性同构.

证明 不难验证 φ 是一个线性映射. 若 $f(x) \in \text{Ker}\varphi$, 则 $f(a_i) = 0 (0 \le i \le n)$. 因为 $f(x)$ 的次数不超过 n, 故由多项式根的有限性可知 $f(x) = 0$, 即 $\text{Ker}\varphi = 0$, 这证明了映射 φ 是单映射. 注意到线性空间 V 和 U 的维数都等于 $n+1$, 因此由命题 4.6可知 φ 是线性同构. □

定理 4.4 (Lagrange 插值公式)

设 $a_0, a_1, \cdots, a_n$ 是数域 $\mathbb{F}$ 中 $n+1$ 个不同的数,$b_0, b_1, \cdots, b_n$ 是 $\mathbb{F}$ 中任意 $n+1$ 个数, 求证: 必存在 $\mathbb{F}$ 上次数不超过 n 的多项式 $f(x)$, 使得 $f(a_i) = b_i (0 \le i \le n)$, 并将 $f(x)$ 构造出来.

证明 由引理 4.1可知映射 φ 是映上的 (满射), 因此存在性已经证明. 现来构造 $f(x)$. 设 $\boldsymbol{e}_i = (0, \cdots, 1, \cdots, 0) (1 \le i \le n+1)$ 是 $\mathbb{F}$ 上的 $n+1$ 维标准单位行向量. 对任意的 $0 \le i \le n$, 令

$$f_i(x) = \frac{(x-a_0)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_n)}{(a_i-a_0)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_n)},$$

则 $f_i(a_i) = 1, f_i(a_j) = 0 (j \ne i)$, 于是 $\varphi(f_i) = \boldsymbol{e}_{i+1} (0 \le i \le n)$. 再令

$$f(x) = b_0 f_0(x) + b_1 f_1(x) + \cdots + b_n f_n(x),$$

则容易验证 $\varphi(f) = (b_0, b_1, \cdots, b_n)$, 即 $f(a_i) = b_i (0 \le i \le n)$ 成立. □

4.3.1 证明线性变换可逆的方法

要证明某个有限维线性空间 V 上的线性变换 φ 是自同构 (可逆线性变换), 通常有 3 种方法. 一是可尝试直接构造出 φ 的逆变换. 二是证明 φ 是单映射或者 φ 是满映射 (两者只需其一)(命题 4.6). 三是用矩阵方法, 即选取 V 的一组基, 设 φ 在这组基下的表示矩阵为 $\boldsymbol{A}$, 设法证明 $\boldsymbol{A}$ 是可逆矩阵.

对于无限维线性空间之间的线性映射, 我们并没有定义表示矩阵这一概念, 也没有维数公式等结论, 因此研究线性映射或线性变换, 无限维线性空间的情形远比有限维线性空间的情形难得多, 也常出现对有限维线性空间成立的结论在无限维线性空间却不成立的情况. 例如, 要证明无限维线性空间上的线性变换是自同构, 只能按照定义证明它既是单映射又是满映射, 而不能像有限维线性空间上的线性变换那样, 只验证它是单映射或满映射即可.

例题 4.2 设 φ 是数域 $\mathbb{F}$ 上线性空间 V 上的线性变换, 若存在正整数 n 以及 $a_1, a_2, \cdots, a_n \in \mathbb{F}$, 使得

$$\varphi^n + a_1\varphi^{n-1} + \cdots + a_{n-1}\varphi + a_n I_V = 0,$$

其中 I_V 表示恒等变换并且 $a_n \ne 0$, 求证:φ 是 V 上的自同构.

证明 由条件可得

$$\varphi^n + a_1\varphi^{n-1} + \cdots + a_{n-1}\varphi = -a_n I_V,$$

从而

$$\varphi\left(-\frac{1}{a_n}(\varphi^{n-1} + \cdots + a_{n-1}I_V)\right) = I_V,$$

于是

$$\varphi^{-1} = -\frac{1}{a_n}(\varphi^{n-1} + \cdots + a_{n-1}I_V).$$

□

命题 4.7

设 φ 是 n 维线性空间 V 上的线性变换, 证明:φ 是可逆变换的充要条件是 φ 将 V 的基变为基.

证明 若 φ 是可逆变换, 则显然 φ 将 V 的基变为基.

反之,证法一: 若 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 和 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 是 V 的两组基, 使得 $\varphi(\boldsymbol{e}_i)=\boldsymbol{f}_i(1\leq i\leq n)$, 则对任意 $\boldsymbol{\alpha}\in V,\boldsymbol{\alpha}=\lambda_1\boldsymbol{f}_1+\lambda_2\boldsymbol{f}_2+\cdots+\lambda_n\boldsymbol{f}_n$, 有 $\varphi(\lambda_1\boldsymbol{e}_1+\lambda_2\boldsymbol{e}_2+\cdots+\lambda_n\boldsymbol{e}_n)=\boldsymbol{\alpha}$, 即 φ 是满映射, 从而是自同构.(我们也可以证明 φ 是单映射, 从而是自同构.)

证法二:若 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 和 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 是 V 的两组基, 使得 $\varphi(\boldsymbol{e}_i)=\boldsymbol{f}_i(1\leq i\leq n)$. 设从基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 到基 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 的过渡矩阵为 $\boldsymbol{P}$, 则 φ 在基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 下的表示矩阵就是 $\boldsymbol{P}$, 这是一个可逆矩阵, 从而 φ 是可逆变换. □

命题 4.8

设 U_1,U_2 是 n 维线性空间 V 的子空间, 假设它们维数相同. 求证: 存在 V 上的可逆线性变换 φ, 使得 $U_2=\varphi(U_1)$.

证明 取 U_1 的一组基 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_m$, 并扩张为 V 的一组基 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_m,\boldsymbol{e}_{m+1},\cdots,\boldsymbol{e}_n$; 取 U_2 的一组基 $\boldsymbol{f}_1,\cdots,\boldsymbol{f}_m$, 并扩张为 V 的一组基 $\boldsymbol{f}_1,\cdots,\boldsymbol{f}_m,\boldsymbol{f}_{m+1},\cdots,\boldsymbol{f}_n$. 定义 φ 为 V 上的线性变换, 它在基上的作用为:$\varphi(\boldsymbol{e}_i)=\boldsymbol{f}_i(1\leq i\leq n)$, 则由命题 4.7可知,$\varphi$ 是可逆线性变换, 再由定义容易验证 $\varphi(U_1)=U_2$ 成立. □

例题 4.3 设 φ 是 n 维线性空间 V 上的线性变换, 若对 V 中任一向量 $\boldsymbol{\alpha}$, 总存在正整数 m(m 可能和 $\boldsymbol{\alpha}$ 有关), 使得 $\varphi^m(\boldsymbol{\alpha})=0$. 求证:$I_V-\varphi$ 是自同构.

证明 证法一: 首先证明线性变换 φ 是幂零的. 设 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是线性空间 V 的一组基. 对每个 $\boldsymbol{e}_i$, 都有 m_i, 使得 $\varphi^{m_i}(\boldsymbol{e}_i)=0$, 令 m 为诸 m_i 中最大者. 对 V 中任一向量 $\boldsymbol{v}$, 设 $\boldsymbol{v}=a_1\boldsymbol{e}_1+a_2\boldsymbol{e}_2+\cdots+a_n\boldsymbol{e}_n$, 则有

$$\varphi^m(\boldsymbol{v})=a_1\varphi^m(\boldsymbol{e}_1)+a_2\varphi^m(\boldsymbol{e}_2)+\cdots+a_n\varphi^m(\boldsymbol{e}_n)=\boldsymbol{0}.$$

因此 $\varphi^m=0$.

注意到下列等式:

$$(I_V-\varphi)(I_V+\varphi+\varphi^2+\cdots+\varphi^{m-1})=I_V-\varphi^m=I_V.$$

由此即知 $I_V-\varphi$ 是自同构.

证法二: 只要证明 $I_V-\varphi$ 是单映射即可. 任取 $\boldsymbol{\alpha}\in\operatorname{Ker}(I_V-\varphi)$, 即 $(I_V-\varphi)(\boldsymbol{\alpha})=0$, 则 $\varphi(\boldsymbol{\alpha})=\boldsymbol{\alpha}$. 设 m 为正整数, 使得 $\varphi^m(\boldsymbol{\alpha})=0$, 则 $0=\varphi^m(\boldsymbol{\alpha})=\varphi^{m-1}(\boldsymbol{\alpha})=\cdots=\varphi(\boldsymbol{\alpha})=\boldsymbol{\alpha}$, 故 $\operatorname{Ker}(I_V-\varphi)=0$, 即 $I_V-\varphi$ 是单映射. □

例题 4.4 设 $V=M_n(\mathbb{F})$ 是 $\mathbb{F}$ 上 n 阶矩阵全体组成的线性空间,$\boldsymbol{A},\boldsymbol{B}$ 是两个 n 阶矩阵, 定义 V 上的变换:$\varphi(\boldsymbol{X})=\boldsymbol{AXB}$. 求证:$\varphi$ 是 V 上的线性变换,φ 是可逆变换的充要条件是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是可逆矩阵.

注 用命题 4.9的结论来看这个例题, 就能发现 D 之所以不是可逆变换, 是因为它的右逆变换除了 S 之外, 还有无穷多个.

证明 容易验证 φ 是线性变换. 若 $\boldsymbol{A},\boldsymbol{B}$ 都是可逆矩阵, 则 $\psi(\boldsymbol{X})=\boldsymbol{A}^{-1}\boldsymbol{X}\boldsymbol{B}^{-1}$ 是 φ 的逆线性变换. 下面用两种方法来证明必要性.

证法一: 若 $\boldsymbol{A}$ 是不可逆矩阵, 则我们可证明 φ 不是单映射, 即存在 $\boldsymbol{X}\neq\boldsymbol{O}$, 使得 $\varphi(\boldsymbol{X})=\boldsymbol{AXB}=\boldsymbol{O}$, 从而 φ 不是可逆变换. 事实上, 若 $\boldsymbol{A}$ 的秩等于 $r<n$, 则存在可逆矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$, 使得 $\boldsymbol{PAQ}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$. 令 $\boldsymbol{C}=\begin{pmatrix}\boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{I}_{n-r}\end{pmatrix}$, 则 $\boldsymbol{PAQC}=\boldsymbol{O}$, 而 $\boldsymbol{P}$ 是可逆矩阵, 故 $\boldsymbol{AQC}=\boldsymbol{O}$, 再令 $\boldsymbol{X}=\boldsymbol{QC}$ 即可. 同理, 若 $\boldsymbol{B}$ 的秩小于 n, 也可以证明 φ 不是可逆变换.

证法二: 若 $\boldsymbol{A}$ 是不可逆矩阵, 则对任意的 n 阶矩阵 $\boldsymbol{X}$,$\varphi(\boldsymbol{X})=\boldsymbol{AXB}$ 总是不可逆矩阵 (行列式一定为零). 因此 φ 不可能是映上的 (可逆矩阵不在值域里但是在像空间中). 同理, 若 $\boldsymbol{B}$ 是不可逆矩阵,φ 也不是映上的. □

例题 4.5 设 V 是实系数多项式全体构成的实线性空间, 定义 V 上的变换 D,S 如下:

$$D(f(x))=\frac{\mathrm{d}}{\mathrm{d}x}f(x),\ S(f(x))=\int_0^x f(t)\mathrm{d}t.$$

证明:D,S 均为 V 上的线性变换且 $DS=I_V$, 但 $SD\neq I_V$.

证明 简单验证即得结论. 由 $DS=I_V$ 可知,S 是单线性映射,D 是满线性映射. 又容易看出 S 不是满映射 (值域不包含常数),D 不是单映射, 从而它们都不是自同构. □

命题 4.9

设 V 是 $\mathbb{K}$ 上的无限维线性空间,φ,ψ 是 V 上的线性变换.

(1) *证明:φ 和 ψ 都是可逆变换的充要条件是 $\varphi\psi$ 和 $\psi\varphi$ 都是可逆变换;*

(2) *若 $\psi\varphi=I_V$, 则称 ψ 是 φ 的左逆变换,φ 是 ψ 的右逆变换. 证明:φ 是可逆变换的充要条件是 φ 有且仅有一个左逆变换 (右逆变换).*

笔记 这个命题对有限维空间仍成立.

证明

(1) 若 φ 和 ψ 都是可逆变换, 则 $(\psi^{-1}\varphi^{-1})(\varphi\psi)=(\varphi\psi)(\psi^{-1}\varphi^{-1})=I_V,(\varphi^{-1}\psi^{-1})(\psi\varphi)=(\psi\varphi)(\varphi^{-1}\psi^{-1})=I_V$, 因此 $\varphi\psi$ 和 $\psi\varphi$ 都是可逆变换. 反之, 若 $\varphi\psi$ 和 $\psi\varphi$ 都是可逆变换, 则存在 V 上的线性变换 ξ,η, 使得 $\varphi\psi\xi=\xi\varphi\psi=I_V,\psi\varphi\eta=\eta\psi\varphi=I_V$. 由 $\varphi\psi\xi=I_V$ 及命题 4.12(2)可得 φ 是满映射, 由 $\eta\psi\varphi=I_V$ 及命题 4.12(1)可得 φ 是单映射, 从而 φ 是可逆变换. 同理可证 ψ 也是可逆变换.

(2) 若 φ 是可逆变换, 任取 φ 的一个左逆变换 ψ, 则

$$\psi=\psi I_V=\psi\varphi\varphi^{-1}=I_V\varphi^{-1}=\varphi^{-1},$$

即 φ 的任一左逆变换都是逆变换 φ^{-1}. 由逆变换的唯一性可知,φ 有且仅有一个左逆变换. 反之, 若 φ 有且仅有一个左逆变换 ψ, 则 $\psi\varphi=I_V$, 且有

$$(\psi+\varphi\psi-I_V)\varphi=\psi\varphi+\varphi\psi\varphi-\varphi=I_V+\varphi-\varphi=I_V,$$

即 $\psi+\varphi\psi-I_V$ 也是 φ 的左逆变换, 从而 $\psi+\varphi\psi-I_V=\psi$, 即 $\varphi\psi=I_V$. 因此 ψ 也是 φ 的右逆变换, 从而 φ 是可逆变换. 同理可证关于右逆变换的结论.

□

例题 4.6 试构造无限维线性空间 V 以及 V 上的线性变换 φ,ψ, 使得 $\varphi\psi-\psi\varphi=I_V$.

解 设 V 是实系数多项式全体构成的实线性空间, 线性变换 φ,ψ 定义为: 对任一 $f(x)\in V,\varphi(f(x))=f'(x),\psi(f(x))=xf(x)$. 容易验证 $\varphi\psi-\psi\varphi=I_V$ 成立. □

注 事实上, 满足上述性质的线性变换 φ,ψ 绝不可能存在于有限维线性空间 V 上. 若存在, 取 V 的一组基并设 φ,ψ 的表示矩阵为 $\boldsymbol{A},\boldsymbol{B}$, 则有 $\boldsymbol{AB}-\boldsymbol{BA}=\boldsymbol{I}$ 成立. 上式两边同时取迹, 可得

$$0=\mathrm{tr}(\boldsymbol{AB}-\boldsymbol{BA})=\mathrm{tr}(\boldsymbol{I})=\dim V,$$

导出矛盾.

4.4 线性映射与矩阵

线性映射与矩阵的关系是这一章的核心. 线性映射是一个几何概念, 矩阵是一个代数概念, 它们之间的关系需要掌握以下几点:

(1) 记数域 $\mathbb{F}$ 上 n 维向量空间 V 到 m 维向量空间 U 的线性映射全体为 $\mathcal{L}(V,U)$,$\mathbb{F}$ 上 $m\times n$ 矩阵全体为 $M_{m\times n}(\mathbb{F})$. 各自取定 V 和 U 的一组基, 设 $\varphi\in\mathcal{L}(V,U)$ 在给定基下的表示矩阵为 $\boldsymbol{A}$, 则 $\varphi\mapsto\boldsymbol{A}$ 定义了从 $\mathcal{L}(V,U)$ 到 $M_{m\times n}(\mathbb{F})$ 的一一对应, 这个对应还是一个线性同构. 若 $m=n$, 则在这个对应下, 线性同构 (可逆线性映射) 对应于可逆矩阵. 特别地, 若 $V=U$, 上述对应还定义了一个代数同构, 即除了保持加法与数乘外, 还保持乘法. 因此, 两个向量空间之间线性映射的运算完全可以归结为矩阵的运算.

(2) 设线性映射 φ 在给定基下的表示矩阵为 $\boldsymbol{A}$, 则 $\mathrm{Ker}\varphi$ 和齐次线性方程组 $\boldsymbol{Ax}=0$ 的解空间同构,$\mathrm{Im}\varphi$ 和 $\boldsymbol{A}$ 的全体列向量张成的向量空间同构. 因此 $\dim\mathrm{Im}\varphi=\mathrm{r}(\varphi)=\mathrm{r}(\boldsymbol{A}),\dim\mathrm{Ker}\varphi=n-\mathrm{r}(\varphi)=n-\mathrm{r}(\boldsymbol{A})$.

这两点由定理 4.5的结论即得.

定理 4.5

设 φ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 到 m 维线性空间 U 的线性映射. 令 $\mathbb{F}^n$ 和 $\mathbb{F}^m$ 分别是 $\mathbb{F}$ 上 n 维和 m 维列向量空间. 又设 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 和 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_m$ 分别是 V 和 U 的基,φ 在给定基下的表示矩阵为 $\boldsymbol{A}$. 记 $\eta_1: V\to\mathbb{F}^n$ 为 V 中向量映射到它在基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 下的坐标向量的线性同构,$\eta_2: U\to\mathbb{F}^m$ 为 U 中向量映射到它在基 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_m$ 下的坐标向量的线性同构,$\boldsymbol{A}:\mathbb{F}^n\to\mathbb{F}^m$ 为矩阵乘法诱导的线性映射, 即 $\boldsymbol{A}(\alpha)=\boldsymbol{A}\alpha$. 求证:$\eta_2\varphi=\boldsymbol{A}\eta_1$, 即下列图交换, 并且 $\eta_1:\mathrm{Ker}\varphi\to\mathrm{Ker}\boldsymbol{A}$, $\eta_2:\mathrm{Im}\varphi\to\mathrm{Im}\boldsymbol{A}$ 都是线性同构.

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "\eta_1"'] & U \arrow[d, "\eta_2"] \\ \mathbb{F}^n \arrow[r, "\boldsymbol{A}"] & \mathbb{F}^m \end{tikzcd}$$

证明

□

命题 4.10

设 φ 是线性空间 V 到 U 的线性映射,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 和 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 是 V 的两组基,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵为 $\boldsymbol{P}$. $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$ 和 $\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$ 是 U 的两组基,$\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$ 到 $\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$ 的过渡矩阵为 $\boldsymbol{Q}$. 又设 φ 在基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 和基 $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$ 下的表示矩阵为 $\boldsymbol{A}$, 在基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 和基 $\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$ 下的表示矩阵为 $\boldsymbol{B}$. 求证:$\boldsymbol{B}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{P}$.

特别地, 若 φ 是线性空间 V 上的线性变换,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 和 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 是 V 的两组基,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵为 $\boldsymbol{P}$, 又设 φ 在基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下的表示矩阵为 $\boldsymbol{A}$, 则 φ 在基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 下的表示矩阵 $B=P^{-1}AP$.

证明 任取 $\boldsymbol{v}\in V$, 设它在基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下的坐标向量为 $(x_1,x_2,\cdots,x_n)'$, 则它在基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 下的坐标向量为 $\boldsymbol{P}^{-1}(x_1,x_2,\cdots,x_n)'$. $\varphi(\boldsymbol{v})$ 在基 $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$ 下的坐标向量为 $\boldsymbol{A}(x_1,x_2,\cdots,x_n)'$, 在基 $\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$ 下的坐标向量为 $\boldsymbol{B}\boldsymbol{P}^{-1}(x_1,x_2,\cdots,x_n)'$. 由于从 $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$ 到 $\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$ 的过渡矩阵为 $\boldsymbol{Q}$, 故

$$\boldsymbol{A}(x_1,x_2,\cdots,x_n)'=\boldsymbol{Q}\boldsymbol{B}\boldsymbol{P}^{-1}(x_1,x_2,\cdots,x_n)'.$$

因为 $(x_1,x_2,\cdots,x_n)'$ 是任意的, 故 $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{B}\boldsymbol{P}^{-1}$, 即 $\boldsymbol{B}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{P}$. □

命题 4.11

设 φ 是有限维线性空间 V 到 U 的线性映射, 求证: 必存在 V 和 U 的两组基, 使线性映射 φ 在两组基下的表示矩阵为 $\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$.

证明 设 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是 V 的一组基,$\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$ 是 U 的一组基,φ 在这两组基下的表示矩阵为 $\boldsymbol{A}$. 由相抵标准型理论可知, 存在 m 阶非异阵 $\boldsymbol{Q}$,n 阶非异阵 $\boldsymbol{P}$, 使得 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$. 设 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 是 V 的一组新基, 使得从 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵为 $\boldsymbol{P}$; 设 $\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$ 是 U 的一组新基, 使得从 $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$ 到 $\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$ 的过渡矩阵为 $\boldsymbol{Q}$, 则由命题 4.10可知,φ 在两组新基下的表示矩阵为 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$. □

注 利用这个命题可以得到 $\mathrm{Ker}\varphi=L(\boldsymbol{f}_{r+1},\cdots,\boldsymbol{f}_n)$, $\mathrm{Im}\varphi=L(\boldsymbol{h}_1,\cdots,\boldsymbol{h}_r)$, 由此即得线性映射的维数公式.

命题 4.12 (线性映射维数公式)

设 V,U 是数域 $\mathbb{K}$ 上的有限维线性空间,$\varphi:V\to U$ 是线性映射,$\varphi:V\to U$ 为线性映射, 求证:

$$\dim\mathrm{Ker}\varphi+\dim\mathrm{Im}\varphi=\dim V.$$

证明 证法一: 设 $\dim V=n,\dim\mathrm{Ker}\varphi=k$, 我们只要证明 $\dim\mathrm{Im}\varphi=n-k$ 即可. 取 $\mathrm{Ker}\varphi$ 的一组基 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_k$, 并将其扩张为 V 的一组基 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_k,\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n$. 任取 $\boldsymbol{\alpha}\in V$, 设 $\boldsymbol{\alpha}=c_1\boldsymbol{e}_1+\cdots+c_k\boldsymbol{e}_k+c_{k+1}\boldsymbol{e}_{k+1}+\cdots+c_n\boldsymbol{e}_n$, 则 $\varphi(\boldsymbol{\alpha})=c_{k+1}\varphi(\boldsymbol{e}_{k+1})+\cdots+c_n\varphi(\boldsymbol{e}_n)$, 即 $\mathrm{Im}\varphi$ 中任一向量都是 $\varphi(\boldsymbol{e}_{k+1}),\cdots,\varphi(\boldsymbol{e}_n)$ 的线性组合. 下证 $\varphi(\boldsymbol{e}_{k+1}),\cdots,\varphi(\boldsymbol{e}_n)$ 线性无关. 设 $\lambda_{k+1}\varphi(\boldsymbol{e}_{k+1})+\cdots+\lambda_n\varphi(\boldsymbol{e}_n)=\boldsymbol{0}$, 则 $\varphi(\lambda_{k+1}\boldsymbol{e}_{k+1}+\cdots+\lambda_n\boldsymbol{e}_n)=\boldsymbol{0}$, 即 $\lambda_{k+1}\boldsymbol{e}_{k+1}+\cdots+\lambda_n\boldsymbol{e}_n\in\mathrm{Ker}\varphi$, 故可设 $\lambda_{k+1}\boldsymbol{e}_{k+1}+\cdots+\lambda_n\boldsymbol{e}_n=\lambda_1\boldsymbol{e}_1+\cdots+\lambda_k\boldsymbol{e}_k$, 再由 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_k,\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n$ 线性无关可知 $\lambda_1=\cdots=\lambda_k=\lambda_{k+1}=\cdots=\lambda_n=0$. 因此 $\varphi(\boldsymbol{e}_{k+1}),\cdots,\varphi(\boldsymbol{e}_n)$ 是 $\mathrm{Im}\varphi$ 的一组基, 从而 $\dim\mathrm{Im}\varphi=n-k$, 结论得证.

证法二 (从商空间的角度): 设由 φ 诱导的线性映射 $\overline{\varphi}:V/\mathrm{Ker}\varphi\to\mathrm{Im}\varphi,\overline{\varphi}(\boldsymbol{v}+\mathrm{Ker}\varphi)=\varphi(\boldsymbol{v})$. 先证是 $\overline{\varphi}$ 线性同构的.

首先,$\overline{\varphi}$ 的定义不依赖于 $\mathrm{Ker}\varphi-$ 陪集代表元的选取. 事实上, 若 $\boldsymbol{v}_1+\mathrm{Ker}\varphi=\boldsymbol{v}_2+\mathrm{Ker}\varphi$, 即 $\boldsymbol{v}_1-\boldsymbol{v}_2\in\mathrm{Ker}\varphi$, 则 $0=\varphi(\boldsymbol{v}_1-\boldsymbol{v}_2)=\varphi(\boldsymbol{v}_1)-\varphi(\boldsymbol{v}_2)$, 即 $\varphi(\boldsymbol{v}_1)=\varphi(\boldsymbol{v}_2)$. 其次, 容易验证 $\overline{\varphi}$ 是一个线性映射. 再次, 由 $\overline{\varphi}$ 的定义不难看出它是满射. 最后, 由 $\overline{\varphi}$ 的定义可知 $\mathrm{Ker}\overline{\varphi}=\{\boldsymbol{0}+\mathrm{Ker}\varphi\}$ 是商空间 $V/\mathrm{Ker}\varphi$ 的零子空间, 故为单射, 从而 $\overline{\varphi}:V/\mathrm{Ker}\varphi\to\mathrm{Im}\varphi$ 是线性同构. 由商空间的维数公式可得

$$\dim\mathrm{Im}\varphi=\dim(V/\mathrm{Ker}\varphi)=\dim V-\dim\mathrm{Ker}\varphi,$$

由此即得线性映射的维数公式. □

推论 4.4

设 $\varphi:V\to U$ 为线性映射,$\mathrm{Ker}\varphi$ 的一组基为 $\boldsymbol{e}_{r+1}\cdots,\boldsymbol{e}_n$, 并将其扩张为 V 的一组基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$. 则 $\varphi(\boldsymbol{e}_1),\cdots,\varphi(\boldsymbol{e}_r)$ 一定是 $\mathrm{Im}\varphi$ 的一组基.

证明 由命题 4.12的证法一立得.

□

命题 4.13

设 φ 是 n 维线性空间 V 到 m 维线性空间 U 的线性映射,φ 在给定基下的表示矩阵为 $\boldsymbol{A}_{m\times n}$. 求证:$\varphi$ 是满映射的充要条件是 $\mathrm{r}(\boldsymbol{A})=m(\boldsymbol{A}$ 行满秩),φ 是单映射的充要条件是 $\mathrm{r}(\boldsymbol{A})=n(\boldsymbol{A}$ 列满秩).

笔记 $\dim\mathrm{Im}\varphi=\mathrm{r}(\boldsymbol{A})$ 和 $\dim\mathrm{Ker}\varphi=n-\mathrm{r}(\boldsymbol{A})$ 的原因见线性映射与矩阵基本结论 (2).

证明 注意到 $\dim\mathrm{Im}\varphi=\mathrm{r}(\boldsymbol{A})$, 并且 φ 是满映射的充要条件是 $\mathrm{Im}\varphi=U$, 这也等价于 $\dim\mathrm{Im}\varphi=\dim U=m$, 故第一个结论成立.

注意到 $\dim\mathrm{Ker}\varphi=n-\mathrm{r}(\boldsymbol{A})$, 并且 φ 是单映射的充要条件是 $\mathrm{Ker}\varphi=0$, 这也等价于 $\dim\mathrm{Ker}\varphi=0$, 故第二个结论成立. □

命题 4.14

设 $\varphi:V\to U$ 为线性映射且 φ 的秩为 r, 证明: 存在 r 个秩为 1 的线性映射 $\varphi_i:V\to U(1\le i\le r)$, 使得 $\varphi=\varphi_1+\cdots+\varphi_r$.

证明 取定 V 和 U 的两组基, 设 φ 在这两组基下的表示矩阵为 $\boldsymbol{A}$, 则 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\varphi)=r$. 由矩阵的秩 1 分解可知, 存在 r 个秩为 1 的矩阵 $\boldsymbol{A}_i(1\le i\le r)$, 使得 $\boldsymbol{A}=\boldsymbol{A}_1+\cdots+\boldsymbol{A}_r$. 由于线性映射和表示矩阵之间一一对应, 故存在线性映射 $\varphi_i:V\to U(1\le i\le r)$, 使得 $\varphi=\varphi_1+\cdots+\varphi_r$, 且 $\mathrm{r}(\varphi_i)=\mathrm{r}(\boldsymbol{A}_i)=1$. □

命题 4.15

设 φ 是线性空间 V 上的线性变换, 若它在 V 的任一组基下的表示矩阵都相同, 求证:φ 是纯量变换, 即存在常数 k, 使得 $\varphi(\alpha)=k\alpha$ 对一切 $\alpha\in V$ 都成立.

证明 取定 V 的一组基, 设 φ 在这组基下的表示矩阵是 $\boldsymbol{A}$. 由已知条件可知, 对任意一个同阶可逆矩阵 $\boldsymbol{P}$,$\boldsymbol{A}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 即 $\boldsymbol{P}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{P}$. 因此矩阵 $\boldsymbol{A}$ 和任意一个可逆矩阵乘法可交换, 于是由命题 2.8可知 $\boldsymbol{A}=k\boldsymbol{I}_n$, 由此即知 φ 是纯量变换. □

4.4.1 将矩阵问题转化为线性映射问题

我们将线性映射的问题转化为矩阵问题来处理. 反之, 我们也可将矩阵问题转化为线性映射 (线性变换) 问题来处理. 一般的处理方式如下:

设 $\boldsymbol{A}$ 是数域 $\mathbb{F}$ 上的 $m\times n$ 矩阵, 定义列向量空间 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的线性映射:$\varphi(\boldsymbol{\alpha})=\boldsymbol{A\alpha}$, 容易验证在 $\mathbb{F}^n$ 和 $\mathbb{F}^m$ 的标准单位列向量构成的基下,φ 的表示矩阵就是 $\boldsymbol{A}$. 同理, 若 $\boldsymbol{A}$ 是 $\mathbb{F}$ 上的 n 阶矩阵, 定义 $\mathbb{F}^n$ 上的线性变换:$\varphi(\boldsymbol{\alpha})=\boldsymbol{A\alpha}$, 容易验证在 $\mathbb{F}^n$ 的标准单位列向量构成的基下,φ 的表示矩阵就是 $\boldsymbol{A}$.

因此, 我们有时就把这个线性映射 (线性变换) 写为 $\boldsymbol{A}$. 上述把代数问题转化成几何问题的语言表述, 在后面的章节中一直会用到. 某些矩阵问题采用这种方式转化为线性映射 (线性变换) 问题后, 往往变得比较容易解决或者可以充分利用几何直观去得到解题思路.

例题 4.7 设 $\boldsymbol{A},\boldsymbol{B}$ 都是数域 $\mathbb{F}$ 上的 $m\times n$ 矩阵, 求证: 方程组 $\boldsymbol{Ax}=0,\boldsymbol{Bx}=0$ 同解的充要条件是存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{B}=\boldsymbol{PA}$.

证明 因为 $\boldsymbol{P}$ 是可逆矩阵, 充分性是显然的. 现通过两种方法来证明必要性.

证法一 (代数方法): 由条件可得方程组 $\boldsymbol{Ax}=0,\boldsymbol{Bx}=0,\begin{pmatrix}\boldsymbol{A}\\\boldsymbol{B}\end{pmatrix}\boldsymbol{x}=0$ 都同解, 从而有

$$\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})=\mathrm{r}\begin{pmatrix}\boldsymbol{A}\\\boldsymbol{B}\end{pmatrix}.$$

注意到结论 $\boldsymbol{B}=\boldsymbol{PA}$ 就是说 $\boldsymbol{A},\boldsymbol{B}$ 可以通过初等行变换相互转化, 因此在证明的过程中, 对 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 实施初等行变换不影响结论的证明. 设

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_m\end{pmatrix},\boldsymbol{B}=\begin{pmatrix}\boldsymbol{\beta}_1\\\boldsymbol{\beta}_2\\\vdots\\\boldsymbol{\beta}_m\end{pmatrix}$$

分别为 $\boldsymbol{A},\boldsymbol{B}$ 的行分块. 不妨对 $\boldsymbol{A},\boldsymbol{B}$ 都进行行对换, 故可设 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 是 $\boldsymbol{A}$ 的行向量的极大无关组,$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 是 $\boldsymbol{B}$ 的行向量的极大无关组. 由于 $\mathrm{r}\begin{pmatrix}\boldsymbol{A}\\\boldsymbol{B}\end{pmatrix}=r$, 故由命题 3.10可知,$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 和 $\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$ 是向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 的两组极大无关组. 设 $\boldsymbol{\beta}_i=\sum\limits_{j=1}^{r}c_{ij}\boldsymbol{\alpha}_j(1\le i\le r)$, 则容易验证 r 阶方阵 $\boldsymbol{C}=(c_{ij})$ 是非异阵. 设 $\boldsymbol{\beta}_i-\boldsymbol{\alpha}_i=\sum\limits_{j=1}^{r}d_{ij}\boldsymbol{\alpha}_j(r+1\le i\le m)$,$\boldsymbol{D}=(d_{ij})$ 是 $(m-r)\times r$ 矩阵, 则容易验证 $\boldsymbol{P}=\begin{pmatrix}\boldsymbol{C}&\boldsymbol{O}\\\boldsymbol{D}&\boldsymbol{I}_{m-r}\end{pmatrix}$ 是 m 阶非异阵, 并且满足 $\boldsymbol{B}=\boldsymbol{PA}$.

证法二 (几何方法): 将问题转化成几何的语言即为: 设 V 是 $\mathbb{F}$ 上的 n 维线性空间,U 是 $\mathbb{F}$ 上的 m 维线性空间,$\varphi,\psi:V\to U$ 是两个线性映射. 求证: 若 $\mathrm{Ker}\varphi=\mathrm{Ker}\psi$, 则存在 U 上的自同构 σ, 使得 $\psi=\sigma\varphi$.

设 $\mathrm{r}(\varphi)=r$, 则 $\dim\mathrm{Ker}\varphi=\dim\mathrm{Ker}\psi=n-r$. 取 $\mathrm{Ker}\varphi=\mathrm{Ker}\psi$ 的一组基 $\boldsymbol{e}_{r+1},\cdots,\boldsymbol{e}_n$, 并将其扩张为 V 的一组基 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_r,\boldsymbol{e}_{r+1},\cdots,\boldsymbol{e}_n$. 根据推论 4.4可知,$\varphi(\boldsymbol{e}_1),\cdots,\varphi(\boldsymbol{e}_r)$ 是 $\mathrm{Im}\varphi$ 的一组基, 故可将其扩张为 U 的一组基 $\varphi(\boldsymbol{e}_1),\cdots,\varphi(\boldsymbol{e}_r),\boldsymbol{f}_{r+1},\cdots,\boldsymbol{f}_m$. 同理可知,$\psi(\boldsymbol{e}_1),\cdots,\psi(\boldsymbol{e}_r)$ 是 $\mathrm{Im}\psi$ 的一组基, 故可将其扩张为 U 的一组基 $\psi(\boldsymbol{e}_1),\cdots,\psi(\boldsymbol{e}_r),\boldsymbol{g}_{r+1},\cdots,\boldsymbol{g}_m$. 定义 U 上的线性变换 σ 如下:

$$\sigma(\varphi(\boldsymbol{e}_i))=\psi(\boldsymbol{e}_i),1\le i\le r;\ \sigma(\boldsymbol{f}_j)=\boldsymbol{g}_j,r+1\le j\le m.$$

因为 σ 把 U 的一组基映射为 U 的另一组基, 故 σ 是 U 的自同构. 又对 $r+1\le j\le n$,$\sigma(\varphi(\boldsymbol{e}_j))=0=\psi(\boldsymbol{e}_j)$, 故 $\sigma\varphi=\psi$ 成立. □

命题 4.16

若数域 $\mathbb{F}$ 上的 n 阶方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似, 求证: 它们可以看成是某个线性空间上同一个线性变换在不同基下的表示矩阵.

笔记 由下面的证明可知这个线性变换 φ 就是由矩阵 $\boldsymbol{A}$ 的乘法诱导的线性变换. 两组不同的基就是标准基与可逆矩阵 $\boldsymbol{P}$ 的列向量.

证明 令 $V=\mathbb{F}^n$ 是 n 维列向量空间,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是由 n 维标准单位列向量构成的基,φ 是由矩阵 $\boldsymbol{A}$ 的乘法诱导的线性变换, 容易验证 φ 在基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下的表示矩阵就是 $\boldsymbol{A}$. 已知 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似, 即存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. 令 $\boldsymbol{P}=(\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n)$ 为其列分块, 由于 $\boldsymbol{P}$ 可逆, 故 $\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$ 线性无关, 从而是 V 的一组基. 注意到从基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵就是 $\boldsymbol{P}$, 因此线性变换 φ 在基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 下的表示矩阵为 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$. □

例题 4.8 设 V 是数域 $\mathbb{F}$ 上 n 阶矩阵全体构成的线性空间,φ 是 V 上的线性变换:$\varphi(\boldsymbol{A})=\boldsymbol{A}'$. 证明: 存在 V 的一组基, 使得 φ 在这组基下的表示矩阵是一个对角矩阵且主对角元素全是 1 或 -1, 并求出 1 和 -1 的个数.

证明 设 V_1 是由 n 阶对称矩阵组成的子空间,V_2 是由反对称矩阵组成的子空间, 则由命题 3.19可得

$$V=V_1\oplus V_2.$$

取 V_1 的一组基和 V_2 的一组基拼成 V 的一组基, 则 φ 在这组基下的表示矩阵是对角矩阵且主对角元素或为 1 或为 -1. 因为 $\dim V_1=\dfrac{1}{2}n(n+1)$,$\dim V_2=\dfrac{1}{2}n(n-1)$, 故 1 的个数为 $\dfrac{1}{2}n(n+1)$,-1 的个数为 $\dfrac{1}{2}n(n-1)$. □

例题 4.9 设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,φ,ψ 是 V 上的线性变换且 $\varphi^2=0$,$\psi^2=0$,$\varphi\psi+\psi\varphi=\boldsymbol{I}$,$\boldsymbol{I}$ 是 V 上的恒等变换. 求证:

(1) $V=\mathrm{Ker}\varphi\oplus\mathrm{Ker}\psi$;

(2) 若 V 是二维空间, 则存在 V 的基 $\boldsymbol{e}_1,\boldsymbol{e}_2$, 使得 φ,ψ 在这组基下的表示矩阵分别为

$$\boldsymbol{A}=\begin{pmatrix}0&0\\1&0\end{pmatrix},\boldsymbol{B}=\begin{pmatrix}0&1\\0&0\end{pmatrix};$$

(3) V 必是偶数维空间且若 V 是 $2k$ 维空间, 则存在 V 的一组基, 使得 φ,ψ 在这组基下的表示矩阵分别为下列分块对角矩阵:

$$\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}&\cdots&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{A}&\cdots&\boldsymbol{O}\\\vdots&\vdots&&\vdots\\\boldsymbol{O}&\boldsymbol{O}&\cdots&\boldsymbol{A}\end{pmatrix},\quad\begin{pmatrix}\boldsymbol{B}&\boldsymbol{O}&\cdots&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}&\cdots&\boldsymbol{O}\\\vdots&\vdots&&\vdots\\\boldsymbol{O}&\boldsymbol{O}&\cdots&\boldsymbol{B}\end{pmatrix},$$

其中主对角线上分别有 k 个 $\boldsymbol{A}$ 和 k 个 $\boldsymbol{B}$.

证明

(1) 任取 $\boldsymbol{\alpha}\in V$, 则由 $\boldsymbol{I}=\varphi\psi+\psi\varphi$ 得到 $\boldsymbol{\alpha}=\varphi\psi(\boldsymbol{\alpha})+\psi\varphi(\boldsymbol{\alpha})$. 注意到 $\varphi\psi(\boldsymbol{\alpha})\in\mathrm{Ker}\varphi$,$\psi\varphi(\boldsymbol{\alpha})\in\mathrm{Ker}\psi$, 因此 $V=\mathrm{Ker}\varphi+\mathrm{Ker}\psi$. 又若 $\boldsymbol{\beta}\in\mathrm{Ker}\varphi\cap\mathrm{Ker}\psi$, 则 $\boldsymbol{\beta}=\varphi\psi(\boldsymbol{\beta})+\psi\varphi(\boldsymbol{\beta})=0$, 即 $\mathrm{Ker}\varphi\cap\mathrm{Ker}\psi=0$. 于是 $V=\mathrm{Ker}\varphi\oplus\mathrm{Ker}\psi$.

(2) 取 $\boldsymbol{0}\neq\boldsymbol{e}_1\in\mathrm{Ker}\psi$,$\boldsymbol{e}_2=\varphi(\boldsymbol{e}_1)$, 则 $\varphi(\boldsymbol{e}_2)=\varphi^2(\boldsymbol{e}_1)=0$, 即 $\boldsymbol{e}_2\in\mathrm{Ker}\varphi$. 又若 $\boldsymbol{e}_2=\boldsymbol{0}$, 则 $\boldsymbol{e}_1\in\mathrm{Ker}\varphi\cap\mathrm{Ker}\psi=0$, 和假设矛盾, 于是 $\boldsymbol{e}_2\neq\boldsymbol{0}$. 因此 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 组成 V 的一组基, 不难验证在这组基下,φ,ψ 的表示矩阵符合要求 $(\psi(\boldsymbol{e}_2)=\psi(\varphi(\boldsymbol{e}_1))=\psi\varphi(\boldsymbol{e}_1)=\boldsymbol{I}(\boldsymbol{e}_1)-\varphi\psi(\boldsymbol{e}_1)=\boldsymbol{e}_1)$.

(3) 设 $\dim\mathrm{Ker}\psi=k$, 并取 $\mathrm{Ker}\psi$ 的一组基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k$. 令 $\boldsymbol{e}_{k+1}=\varphi(\boldsymbol{e}_1)$,$\boldsymbol{e}_{k+2}=\varphi(\boldsymbol{e}_2),\cdots,\boldsymbol{e}_{2k}=\varphi(\boldsymbol{e}_k)$, 则由 $\varphi^2=0$ 可得 $\boldsymbol{e}_{k+1},\boldsymbol{e}_{k+2},\cdots,\boldsymbol{e}_{2k}$ 都属于 $\mathrm{Ker}\varphi$. 我们先证明向量组 $\boldsymbol{e}_{k+1},\boldsymbol{e}_{k+2},\cdots,\boldsymbol{e}_{2k}$ 是线性无关的. 设有

$$c_1\boldsymbol{e}_{k+1}+c_2\boldsymbol{e}_{k+2}+\cdots+c_k\boldsymbol{e}_{2k}=\boldsymbol{0},$$

两边作用 ψ, 可得

$$c_1\psi(\boldsymbol{e}_{k+1})+c_2\psi(\boldsymbol{e}_{k+2})+\cdots+c_k\psi(\boldsymbol{e}_{2k})=\boldsymbol{0}.$$

注意到 $\boldsymbol{e}_1=\varphi\psi(\boldsymbol{e}_1)+\psi\varphi(\boldsymbol{e}_1)=\psi(\boldsymbol{e}_{k+1})$, 同理 $\boldsymbol{e}_2=\psi(\boldsymbol{e}_{k+2}),\cdots,\boldsymbol{e}_k=\psi(\boldsymbol{e}_{2k})$. 因此上式就是

$$c_1\boldsymbol{e}_1+c_2\boldsymbol{e}_2+\cdots+c_k\boldsymbol{e}_k=\boldsymbol{0}.$$

而 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k$ 线性无关, 故 $c_1=c_2=\cdots=c_k=0$, 即向量组 $\boldsymbol{e}_{k+1},\boldsymbol{e}_{k+2},\cdots,\boldsymbol{e}_{2k}$ 线性无关. 特别地, 我们有 $\dim\mathrm{Ker}\varphi\geq k=\dim\mathrm{Ker}\psi$. 由于 φ,ψ 的地位是对称的, 故同理可证 $\dim\mathrm{Ker}\psi\geq\dim\mathrm{Ker}\varphi$, 从而 $\dim\mathrm{Ker}\varphi=\dim\mathrm{Ker}\psi=k$, 并且 $\boldsymbol{e}_{k+1},\boldsymbol{e}_{k+2},\cdots,\boldsymbol{e}_{2k}$ 是 $\mathrm{Ker}\varphi$ 的一组基. 因为 $V=\mathrm{Ker}\varphi\oplus\mathrm{Ker}\psi$, 故 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_k,\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_{2k}$

组成 V 的一组基. 现将基向量排列如下:

$$\boldsymbol{e}_1, \boldsymbol{e}_{k+1}, \boldsymbol{e}_2, \boldsymbol{e}_{k+2}, \cdots, \boldsymbol{e}_k, \boldsymbol{e}_{2k}.$$

不难验证, 在这组基下 φ, ψ 的表示矩阵即为所求.

□

4.5 像空间和核空间

命题 4.17

设 φ 是向量空间 V 上的线性变换, 则

$$V \supseteq \mathrm{Im}\varphi \supseteq \mathrm{Im}\varphi^2 \supseteq \cdots \supseteq \mathrm{Im}\varphi^n \supseteq \mathrm{Im}\varphi^{n+1} \supseteq \cdots,$$

$$\mathrm{Ker}\varphi \subseteq \mathrm{Ker}\varphi^2 \subseteq \cdots \subseteq \mathrm{Ker}\varphi^n \subseteq \mathrm{Ker}\varphi^{n+1} \subseteq \cdots \subseteq V.$$

证明 由像空间和核空间的定义易证. □

例题 4.10 设线性空间 V 上的线性变换 φ 在基 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4\}$ 下的表示矩阵为

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

求 φ 的核空间与像空间 (用基的线性组合来表示).

证明 像空间通过坐标向量同构于 $\boldsymbol{A}$ 的列向量生成的子空间, 通过计算可得 $\boldsymbol{A}$ 的秩等于 2, 且 $\boldsymbol{A}$ 的第一、第二列向量线性无关, 于是 $\mathrm{Im}\varphi$ 的基的坐标向量为 $(1,-1,1,2)', (0,2,2,-2)'$, 从而 $\mathrm{Im}\varphi = k_1(\boldsymbol{e}_1 - \boldsymbol{e}_2 + \boldsymbol{e}_3 + 2\boldsymbol{e}_4) + k_2(2\boldsymbol{e}_2 + 2\boldsymbol{e}_3 - 2\boldsymbol{e}_4)$. 核空间通过坐标向量同构于齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = 0$ 的解空间, 通过计算可得该方程组的基础解系为 $(-4,-3,2,0)', (-1,-2,0,1)'$, 此即 $\mathrm{Ker}\varphi$ 的基的坐标向量, 于是 $\mathrm{Ker}\varphi = k_1(-4\boldsymbol{e}_1 - 3\boldsymbol{e}_2 + 2\boldsymbol{e}_3) + k_2(-\boldsymbol{e}_1 - 2\boldsymbol{e}_2 + \boldsymbol{e}_4)$. □

命题 4.18

设 V 是数域 $\mathbb{F}$ 上的线性空间,$\varphi_1, \varphi_2, \cdots, \varphi_k$ 是 V 上的非零线性变换. 求证: 存在 $\boldsymbol{\alpha} \in V$, 使得 $\varphi_i(\boldsymbol{\alpha}) \neq 0 (1 \leq i \leq k)$.

证明 因为 $\varphi_i \neq 0$, 所以 $\mathrm{Ker}\varphi_i$ 是 V 的真子空间. 由命题 3.24可知, 有限个真子空间 $\mathrm{Ker}\varphi_i$ 不能覆盖全空间 V, 故必存在 $\boldsymbol{\alpha} \in V$, 使得 $\boldsymbol{\alpha}$ 不属于任意一个 $\mathrm{Ker}\varphi_i$, 从而结论得证. □

命题 4.19

设 V 是数域 $\mathbb{F}$ 上的线性空间,$\varphi_1, \varphi_2, \cdots, \varphi_k$ 是 V 上互不相同的线性变换. 求证: 存在 $\boldsymbol{\alpha} \in V$, 使得 $\varphi_1(\boldsymbol{\alpha}), \varphi_2(\boldsymbol{\alpha}), \cdots, \varphi_k(\boldsymbol{\alpha})$ 互不相同.

证明 令 $\varphi_{ij} = \varphi_i - \varphi_j (1 \leq i < j \leq k)$, 则 φ_{ij} 是 V 上的非零线性变换. 由命题 4.18可知, 存在 $\boldsymbol{\alpha} \in V$, 使得 $\varphi_{ij}(\boldsymbol{\alpha}) \neq 0$, 即 $\varphi_i(\boldsymbol{\alpha}) \neq \varphi_j(\boldsymbol{\alpha}) (1 \leq i < j \leq k)$, 从而结论得证. □

命题 4.20

设 $\boldsymbol{A}$ 是 n 阶方阵, 求证:$\mathrm{r}(\boldsymbol{A}^n) = \mathrm{r}(\boldsymbol{A}^{n+1}) = \mathrm{r}(\boldsymbol{A}^{n+2}) = \cdots$.

证明 证法一 (代数方法): 由秩的不等式可得

$$n = \mathrm{r}(\boldsymbol{I}_n) \geq \mathrm{r}(\boldsymbol{A}) \geq \mathrm{r}(\boldsymbol{A}^2) \geq \cdots \geq \mathrm{r}(\boldsymbol{A}^n) \geq \mathrm{r}(\boldsymbol{A}^{n+1}) \geq 0.$$

上述 $n+2$ 个整数都在 $[0,n]$ 之间, 故由抽屉原理可知, 存在某个整数 $m \in [0,n]$, 使得 $\mathrm{r}(\boldsymbol{A}^m) = \mathrm{r}(\boldsymbol{A}^{m+1})$. 对任意的

$k \geq m$, 由矩阵秩的 Frobenius 不等式可得

$$\mathrm{r}(\boldsymbol{A}^{k+1}) = \mathrm{r}(\boldsymbol{A}^{k-m}\boldsymbol{A}^m\boldsymbol{A}) \geq \mathrm{r}(\boldsymbol{A}^{k-m}\boldsymbol{A}^m) + \mathrm{r}(\boldsymbol{A}^m\boldsymbol{A}) - \mathrm{r}(\boldsymbol{A}^m) = \mathrm{r}(\boldsymbol{A}^k),$$

又 $\mathrm{r}(\boldsymbol{A}^{k+1}) \leq \mathrm{r}(\boldsymbol{A}^k)$, 故 $\mathrm{r}(\boldsymbol{A}^{k+1}) = \mathrm{r}(\boldsymbol{A}^k)$ 对任意的 $k \geq m$ 成立, 结论得证.

证法二 (几何方法): 令 φ 为在 n 维列向量空间上由矩阵 $\boldsymbol{A}$ 乘法诱导的线性变换, 则 φ 在标准基下的表示矩阵就是 $\boldsymbol{A}$, 并且不难发现对 $\forall k \in \mathbb{N}, \varphi^k$ 在标准基下的表示矩阵是 $\boldsymbol{A}^k$. 因此 $\mathrm{r}(\boldsymbol{A}^k) = \dim \mathrm{Im}\varphi^k$, 故原命题等价于 $\dim \mathrm{Im}\varphi^n = \dim \mathrm{Im}\varphi^{n+1} = \dim \mathrm{Im}\varphi^{n+2} = \cdots$. 注意下列子空间链:

$$V \supseteq \mathrm{Im}\varphi \supseteq \mathrm{Im}\varphi^2 \supseteq \cdots \supseteq \mathrm{Im}\varphi^n \supseteq \mathrm{Im}\varphi^{n+1}.$$

上述 $n+2$ 个子空间的维数都在 $[0,n]$ 之间, 故由抽屉原理可知, 存在某个整数 $m \in [0,n]$, 使得 $\mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{m+1}$. 现要证明对任意的 $k \geq m, \mathrm{Im}\varphi^k = \mathrm{Im}\varphi^{k+1}$. 一方面,$\mathrm{Im}\varphi^{k+1} \subseteq \mathrm{Im}\varphi^k$ 是显然的. 另一方面, 任取 $\boldsymbol{\alpha} \in \mathrm{Im}\varphi^k$, 则存在 $\boldsymbol{\beta} \in V$, 使得 $\boldsymbol{\alpha} = \varphi^k(\boldsymbol{\beta})$. 由于 $\varphi^m(\boldsymbol{\beta}) \in \mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{m+1}$, 故存在 $\boldsymbol{\gamma} \in V$, 使得 $\varphi^m(\boldsymbol{\beta}) = \varphi^{m+1}(\boldsymbol{\gamma})$, 从而

$$\boldsymbol{\alpha} = \varphi^k(\boldsymbol{\beta}) = \varphi^{k-m}(\varphi^m(\boldsymbol{\beta})) = \varphi^{k-m}(\varphi^{m+1}(\boldsymbol{\gamma})) = \varphi^{k+1}(\boldsymbol{\gamma}) \in \mathrm{Im}\varphi^{k+1},$$

故 $\mathrm{Im}\varphi^k = \mathrm{Im}\varphi^{k+1}$ 对任意的 $k \geq m$ 成立, 取维数后即得结论. □

推论 4.5

(1) 设 $\boldsymbol{A}$ 是 n 阶方阵, 则一定存在整数 $m \in [0,n]$, 使得 $\mathrm{r}(\boldsymbol{A}^m) = \mathrm{r}(\boldsymbol{A}^{m+1}) = \mathrm{r}(\boldsymbol{A}^{m+2}) = \cdots$.

(2) 设 φ 是 n 维线性空间 V 上的线性变换, 则必存在整数 $m \in [0,n]$, 使得

$$\dim \mathrm{Im}\varphi^m = \dim \mathrm{Im}\varphi^{m+1} = \dim \mathrm{Im}\varphi^{m+2} = \cdots.$$

证明

(1) 证法一 (代数方法): 由秩的不等式可得

$$n = \mathrm{r}(\boldsymbol{I}_n) \geq \mathrm{r}(\boldsymbol{A}) \geq \mathrm{r}(\boldsymbol{A}^2) \geq \cdots \geq \mathrm{r}(\boldsymbol{A}^n) \geq \mathrm{r}(\boldsymbol{A}^{n+1}) \geq 0.$$

上述 $n+2$ 个整数都在 $[0,n]$ 之间, 故由抽屉原理可知, 存在某个整数 $m \in [0,n]$, 使得 $\mathrm{r}(\boldsymbol{A}^m) = \mathrm{r}(\boldsymbol{A}^{m+1})$. 对任意的 $k \geq m$, 由矩阵秩的 Frobenius 不等式可得

$$\mathrm{r}(\boldsymbol{A}^{k+1}) = \mathrm{r}(\boldsymbol{A}^{k-m}\boldsymbol{A}^m\boldsymbol{A}) \geq \mathrm{r}(\boldsymbol{A}^{k-m}\boldsymbol{A}^m) + \mathrm{r}(\boldsymbol{A}^m\boldsymbol{A}) - \mathrm{r}(\boldsymbol{A}^m) = \mathrm{r}(\boldsymbol{A}^k),$$

又 $\mathrm{r}(\boldsymbol{A}^{k+1}) \leq \mathrm{r}(\boldsymbol{A}^k)$, 故 $\mathrm{r}(\boldsymbol{A}^{k+1}) = \mathrm{r}(\boldsymbol{A}^k)$ 对任意的 $k \geq m$ 成立, 结论得证.

证法二 (几何方法): 令 φ 为在 n 维列向量空间上由矩阵 $\boldsymbol{A}$ 乘法诱导的线性变换, 则 φ 在标准基下的表示矩阵就是 $\boldsymbol{A}$, 并且不难发现对 $\forall k \in \mathbb{N}, \varphi^k$ 在标准基下的表示矩阵是 $\boldsymbol{A}^k$. 因此 $\mathrm{r}(\boldsymbol{A}^k) = \dim \mathrm{Im}\varphi^k$, 故原命题等价于 $\dim \mathrm{Im}\varphi^n = \dim \mathrm{Im}\varphi^{n+1} = \dim \mathrm{Im}\varphi^{n+2} = \cdots$. 注意下列子空间链:

$$V \supseteq \mathrm{Im}\varphi \supseteq \mathrm{Im}\varphi^2 \supseteq \cdots \supseteq \mathrm{Im}\varphi^n \supseteq \mathrm{Im}\varphi^{n+1}.$$

上述 $n+2$ 个子空间的维数都在 $[0,n]$ 之间, 故由抽屉原理可知, 存在某个整数 $m \in [0,n]$, 使得 $\mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{m+1}$. 现要证明对任意的 $k \geq m, \mathrm{Im}\varphi^k = \mathrm{Im}\varphi^{k+1}$. 一方面,$\mathrm{Im}\varphi^{k+1} \subseteq \mathrm{Im}\varphi^k$ 是显然的. 另一方面, 任取 $\boldsymbol{\alpha} \in \mathrm{Im}\varphi^k$, 则存在 $\boldsymbol{\beta} \in V$, 使得 $\boldsymbol{\alpha} = \varphi^k(\boldsymbol{\beta})$. 由于 $\varphi^m(\boldsymbol{\beta}) \in \mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{m+1}$, 故存在 $\boldsymbol{\gamma} \in V$, 使得 $\varphi^m(\boldsymbol{\beta}) = \varphi^{m+1}(\boldsymbol{\gamma})$, 从而

$$\boldsymbol{\alpha} = \varphi^k(\boldsymbol{\beta}) = \varphi^{k-m}(\varphi^m(\boldsymbol{\beta})) = \varphi^{k-m}(\varphi^{m+1}(\boldsymbol{\gamma})) = \varphi^{k+1}(\boldsymbol{\gamma}) \in \mathrm{Im}\varphi^{k+1},$$

故 $\mathrm{Im}\varphi^k = \mathrm{Im}\varphi^{k+1}$ 对任意的 $k \geq m$ 成立, 取维数后即得结论.

(2) 注意下列子空间链:

$$V \supseteq \mathrm{Im}\varphi \supseteq \mathrm{Im}\varphi^2 \supseteq \cdots \supseteq \mathrm{Im}\varphi^n \supseteq \mathrm{Im}\varphi^{n+1}.$$

上述 $n+2$ 个子空间的维数都在 $[0,n]$ 之间, 故由抽屉原理可知, 存在某个整数 $m \in [0,n]$, 使得 $\mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{m+1}$. 现要证明对任意的 $k \geq m, \mathrm{Im}\varphi^k = \mathrm{Im}\varphi^{k+1}$. 一方面,$\mathrm{Im}\varphi^{k+1} \subseteq \mathrm{Im}\varphi^k$ 是显然的. 另一方面, 任取 $\boldsymbol{\alpha} \in \mathrm{Im}\varphi^k$, 则存在 $\boldsymbol{\beta} \in V$, 使得 $\boldsymbol{\alpha} = \varphi^k(\boldsymbol{\beta})$. 由于 $\varphi^m(\boldsymbol{\beta}) \in \mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{m+1}$, 故存在 $\boldsymbol{\gamma} \in V$, 使得 $\varphi^m(\boldsymbol{\beta}) = \varphi^{m+1}(\boldsymbol{\gamma})$,

从而

$$\alpha = \varphi^k(\beta) = \varphi^{k-m}(\varphi^m(\beta)) = \varphi^{k-m}(\varphi^{m+1}(\gamma)) = \varphi^{k+1}(\gamma) \in \mathrm{Im}\varphi^{k+1},$$

故 $\mathrm{Im}\varphi^k = \mathrm{Im}\varphi^{k+1}$ 对任意的 $k \geq m$ 成立, 取维数后即得结论.

□

命题 4.21

设 φ 是 n 维线性空间 V 上的线性变换, 求证: 必存在整数 $m \in [0, n]$, 使得

$$\mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{m+1} = \mathrm{Im}\varphi^{m+2} = \cdots, \mathrm{Ker}\varphi^m = \mathrm{Ker}\varphi^{m+1} = \mathrm{Ker}\varphi^{m+2} = \cdots, V = \mathrm{Im}\varphi^m \oplus \mathrm{Ker}\varphi^m.$$

证明 根据推论 4.5(2)可知, 存在整数 $m \in [0, n]$, 使得

$$\mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{m+1} = \mathrm{Im}\varphi^{m+2} = \cdots.$$

注意到对任意的正整数 i,$\mathrm{Ker}\varphi^i \subseteq \mathrm{Ker}\varphi^{i+1}$. 再由维数公式可知, 对任意的 $i \geq m$,$\dim \mathrm{Ker}\varphi^i = \dim V - \dim \mathrm{Im}\varphi^i = n - \dim \mathrm{Im}\varphi^m$ 是一个不依赖于 i 的常数, 因此由命题??可得

$$\mathrm{Ker}\varphi^m = \mathrm{Ker}\varphi^{m+1} = \mathrm{Ker}\varphi^{m+2} = \cdots.$$

若 $\alpha \in \mathrm{Im}\varphi^m \cap \mathrm{Ker}\varphi^m$, 则 $\alpha = \varphi^m(\beta)$,$\varphi^m(\alpha) = 0$. 于是 $0 = \varphi^m(\alpha) = \varphi^{2m}(\beta)$, 即 $\beta \in \mathrm{Ker}\varphi^{2m} = \mathrm{Ker}\varphi^m$, 从而 $\alpha = \varphi^m(\beta) = 0$, 这证明了 $\mathrm{Im}\varphi^m \cap \mathrm{Ker}\varphi^m = 0$. 又对 V 中任一向量 α, 因为 $\varphi^m(\alpha) \in \mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{2m}$, 所以 $\varphi^m(\alpha) = \varphi^{2m}(\beta)$, 其中 $\beta \in V$. 我们有分解式

$$\alpha = \varphi^m(\beta) + (\alpha - \varphi^m(\beta)).$$

注意到 $\varphi^m(\alpha - \varphi^m(\beta)) = 0$, 即 $\alpha - \varphi^m(\beta) \in \mathrm{Ker}\varphi^m$, 这就证明了 $V = \mathrm{Im}\varphi^m + \mathrm{Ker}\varphi^m$. 因此

$$V = \mathrm{Im}\varphi^m \oplus \mathrm{Ker}\varphi^m.$$

□

注 也可不证明 $\mathrm{Im}\varphi^m \cap \mathrm{Ker}\varphi^m = 0$, 改由线性映射维数公式$\dim \mathrm{Im}\varphi^m + \dim \mathrm{Ker}\varphi^m = n$ 直接得到 $V = \mathrm{Im}\varphi^m \oplus \mathrm{Ker}\varphi^m$.

命题 4.22

设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,φ 是 V 上的线性变换, 证明以下 9 个结论等价:

(1) $V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$;

(2) $V = \mathrm{Ker}\varphi + \mathrm{Im}\varphi$;

(3) $\mathrm{Ker}\varphi \cap \mathrm{Im}\varphi = 0$;

(4) $\mathrm{Ker}\varphi = \mathrm{Ker}\varphi^2$, 或等价地,$\dim \mathrm{Ker}\varphi = \dim \mathrm{Ker}\varphi^2$;

(5) $\mathrm{Ker}\varphi = \mathrm{Ker}\varphi^2 = \mathrm{Ker}\varphi^3 = \cdots$, 或等价地,$\dim \mathrm{Ker}\varphi = \dim \mathrm{Ker}\varphi^2 = \dim \mathrm{Ker}\varphi^3 = \cdots$;

(6) $\mathrm{Im}\varphi = \mathrm{Im}\varphi^2$, 或等价地,$\mathrm{r}(\varphi) = \mathrm{r}(\varphi^2)$;

(7) $\mathrm{Im}\varphi = \mathrm{Im}\varphi^2 = \mathrm{Im}\varphi^3 = \cdots$, 或等价地,$\mathrm{r}(\varphi) = \mathrm{r}(\varphi^2) = \mathrm{r}(\varphi^3) = \cdots$;

(8) $\mathrm{Ker}\varphi$ 存在 φ-不变补空间, 即存在 φ-不变子空间 U, 使得 $V = \mathrm{Ker}\varphi \oplus U$(实际上,$U = \mathrm{Im}\varphi$);

(9) $\mathrm{Im}\varphi$ 存在 φ-不变补空间, 即存在 φ-不变子空间 W, 使得 $V = \mathrm{Im}\varphi \oplus W$(实际上,$W = \mathrm{Ker}\varphi$).

证明 由直和的定义可知 (1) $\Leftrightarrow$ (2) + (3), 于是 (1) $\Rightarrow$ (2) 和 (1) $\Rightarrow$ (3) 都是显然的. 根据交空间维数公式和线性映射维数公式可知

$$\begin{aligned}\dim(\mathrm{Ker}\varphi + \mathrm{Im}\varphi) &= \dim \mathrm{Ker}\varphi + \dim \mathrm{Im}\varphi - \dim(\mathrm{Ker}\varphi \cap \mathrm{Im}\varphi)\\ &= \dim V - \dim(\mathrm{Ker}\varphi \cap \mathrm{Im}\varphi),\end{aligned}$$

于是 (2) $\Leftrightarrow$ (3) 成立, 从而前 3 个结论两两等价.

(3) $\Rightarrow$ (4): 显然 $\mathrm{Ker}\varphi \subseteq \mathrm{Ker}\varphi^2$ 成立. 任取 $\alpha \in \mathrm{Ker}\varphi^2$, 则 $\varphi(\alpha) \in \mathrm{Ker}\varphi \cap \mathrm{Im}\varphi = 0$, 于是 $\varphi(\alpha) = 0$, 即 $\alpha \in \mathrm{Ker}\varphi$, 从而 $\mathrm{Ker}\varphi^2 \subseteq \mathrm{Ker}\varphi$ 也成立, 故 (4) 成立.

(4) $\Rightarrow$ (3): 任取 $\alpha \in \mathrm{Ker}\varphi \cap \mathrm{Im}\varphi$, 则存在 $\beta \in V$, 使得 $\alpha = \varphi(\beta)$, 于是 $0 = \varphi(\alpha) = \varphi^2(\beta)$, 即 $\beta \in \mathrm{Ker}\varphi^2 = \mathrm{Ker}\varphi$,

从而 $\boldsymbol{\alpha} = \varphi(\boldsymbol{\beta}) = 0$,(3) 成立.

(5) $\Rightarrow$ (4) 是显然的, 下证 (4) $\Rightarrow$ (5): 设 $\mathrm{Ker}\varphi^k = \mathrm{Ker}\varphi^{k+1}$ 已对正整数 k 成立, 先证 $\mathrm{Ker}\varphi^{k+1} = \mathrm{Ker}\varphi^{k+2}$ 也成立, 然后用归纳法即得结论. $\mathrm{Ker}\varphi^{k+1} \subseteq \mathrm{Ker}\varphi^{k+2}$ 是显然的. 任取 $\boldsymbol{\alpha} \in \mathrm{Ker}\varphi^{k+2}$, 即 $0 = \varphi^{k+2}(\boldsymbol{\alpha}) = \varphi^{k+1}(\varphi(\boldsymbol{\alpha}))$, 于是 $\varphi(\boldsymbol{\alpha}) \in \mathrm{Ker}\varphi^{k+1} = \mathrm{Ker}\varphi^k$, 从而 $\varphi^{k+1}(\boldsymbol{\alpha}) = \varphi^k(\varphi(\boldsymbol{\alpha})) = 0$, 即 $\boldsymbol{\alpha} \in \mathrm{Ker}\varphi^{k+1}$, 于是 $\mathrm{Ker}\varphi^{k+2} \subseteq \mathrm{Ker}\varphi^{k+1}$ 也成立.

(3) $\Leftrightarrow$ (6): 考虑 φ 在不变子空间 $\mathrm{Im}\varphi$ 上的限制变换 $\varphi|_{\mathrm{Im}\varphi} : \mathrm{Im}\varphi \to \mathrm{Im}\varphi$, 由限制的定义可知它的核等于 $\mathrm{Ker}\varphi \cap \mathrm{Im}\varphi$, 它的像等于 $\mathrm{Im}\varphi^2$. 由于有限维线性空间上的线性变换是单射当且仅当它是满射, 当且仅当它是同构, 故 (3) $\Leftrightarrow$ (6) 成立.

(7) $\Rightarrow$ (6) 是显然的, 下证 (6) $\Rightarrow$ (7): 设 $\mathrm{Im}\varphi^k = \mathrm{Im}\varphi^{k+1}$ 已对正整数 k 成立, 先证 $\mathrm{Im}\varphi^{k+1} = \mathrm{Im}\varphi^{k+2}$ 也成立, 然后用归纳法即得结论. $\mathrm{Im}\varphi^{k+2} \subseteq \mathrm{Im}\varphi^{k+1}$ 是显然的. 任取 $\boldsymbol{\alpha} \in \mathrm{Im}\varphi^{k+1}$, 即存在 $\boldsymbol{\beta} \in V$, 使得 $\boldsymbol{\alpha} = \varphi^{k+1}(\boldsymbol{\beta})$. 由于 $\varphi^k(\boldsymbol{\beta}) \in \mathrm{Im}\varphi^k = \mathrm{Im}\varphi^{k+1}$, 故存在 $\boldsymbol{\gamma} \in V$, 使得 $\varphi^k(\boldsymbol{\beta}) = \varphi^{k+1}(\boldsymbol{\gamma})$, 于是 $\boldsymbol{\alpha} = \varphi^{k+1}(\boldsymbol{\beta}) = \varphi(\varphi^k(\boldsymbol{\beta})) = \varphi(\varphi^{k+1}(\boldsymbol{\gamma})) = \varphi^{k+2}(\boldsymbol{\gamma}) \in \mathrm{Im}\varphi^{k+2}$, 从而 $\mathrm{Im}\varphi^{k+1} \subseteq \mathrm{Im}\varphi^{k+2}$ 也成立.

(1) $\Rightarrow$ (8) 是显然的, 下证 (8) $\Rightarrow$ (1). 我们先证 $\mathrm{Im}\varphi \subseteq U$: 任取 $\varphi(\boldsymbol{v}) \in \mathrm{Im}\varphi$, 由直和分解可设 $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{u}$, 其中 $\boldsymbol{v}_1 \in \mathrm{Ker}\varphi, \boldsymbol{u} \in U$, 则由 U 的 $\varphi-$ 不变性可得 $\varphi(\boldsymbol{v}) = \varphi(\boldsymbol{v}_1) + \varphi(\boldsymbol{u}) = \varphi(\boldsymbol{u}) \in U$. 考虑不等式

$$\dim V = \dim(\mathrm{Ker}\varphi \oplus U) = \dim \mathrm{Ker}\varphi + \dim U \geq \dim \mathrm{Ker}\varphi + \dim \mathrm{Im}\varphi = \dim V,$$

从而只能是 $U = \mathrm{Im}\varphi$, 于是 (1) 成立.

(1) $\Rightarrow$ (9) 是显然的, 下证 (9) $\Rightarrow$ (1). 我们先证 $W \subseteq \mathrm{Ker}\varphi$: 任取 $\boldsymbol{w} \in W$, 则由 W 的 $\varphi-$ 不变性可得 $\varphi(\boldsymbol{w}) \in \mathrm{Im}\varphi \cap W = 0$, 即有 $\boldsymbol{w} \in \mathrm{Ker}\varphi$. 考虑不等式

$$\dim V = \dim(\mathrm{Im}\varphi \oplus W) = \dim \mathrm{Im}\varphi + \dim W \leq \dim \mathrm{Im}\varphi + \dim \mathrm{Ker}\varphi = \dim V,$$

从而只能是 $W = \mathrm{Ker}\varphi$, 于是 (1) 成立. □

命题 4.23

设 U, W 是 n 维线性空间 V 的子空间且 $\dim U + \dim W = \dim V$. 求证: 存在 V 上的线性变换 φ, 使得 $\mathrm{Ker}\varphi = U, \mathrm{Im}\varphi = W$.

证明 取 U 的一组基 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m$, 并将其扩张为 V 的一组基 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \cdots, \boldsymbol{e}_n$, 再取 W 的一组基 $\boldsymbol{f}_{m+1}, \cdots, \boldsymbol{f}_n$. 定义 φ 为 V 上的线性变换, 它在基上的作用为:$\varphi(\boldsymbol{e}_i) = \boldsymbol{0}(1 \leq i \leq m)$,$\varphi(\boldsymbol{e}_j) = \boldsymbol{f}_j(m+1 \leq j \leq n)$. 注意到 $\boldsymbol{f}_{m+1}, \cdots, \boldsymbol{f}_n$ 是 W 的一组基, 故通过简单的验证可得 $\mathrm{Ker}\varphi = U, \mathrm{Im}\varphi = W$. □

例题 4.11 设 $V = M_n(\mathbb{F})$ 是 $\mathbb{F}$ 上 n 阶矩阵全体构成的线性空间,$\varphi : V \to \mathbb{F}$ 是迹函数, 即对任意的 $\boldsymbol{A} = (a_{ij}) \in V$,

$$\varphi(\boldsymbol{A}) = a_{11} + a_{22} + \cdots + a_{nn}.$$

求证:φ 是 V 到一维空间 $\mathbb{F}$ 上的线性映射, 并求 $\mathrm{Ker}\varphi$ 的维数及其一组基.

证明 容易验证 φ 是线性映射且是映上的. 注意到 V 是 n^2 维线性空间, 由线性映射的维数公式可知

$$\dim\mathrm{Ker}\varphi = \dim V - \dim \mathrm{Im}\varphi = \dim V - \dim \mathbb{F} = n^2 - 1.$$

记 $\boldsymbol{E}_{ij}$ 为 n 阶基础矩阵, 即第 (i, j) 元素为 1, 其余元素为 0 的矩阵. 容易验证下列 $n^2 - 1$ 个矩阵迹为零且线性无关, 因此它们组成了 $\mathrm{Ker}\varphi$ 的一组基:

$$\boldsymbol{E}_{ij}(i \neq j), \boldsymbol{E}_{11} - \boldsymbol{E}_{22}, \boldsymbol{E}_{22} - \boldsymbol{E}_{33}, \cdots, \boldsymbol{E}_{n-1,n-1} - \boldsymbol{E}_{nn}.$$

□

例题 4.12 设 φ 是有限维线性空间 V 到 U 的线性映射, 且 V 的维数大于 U 的维数, 求证:$\mathrm{Ker}\varphi \neq 0$.

证明 由线性映射的维数公式

$$\dim V = \dim \mathrm{Im}\varphi + \dim \mathrm{Ker}\varphi,$$

以及 $\dim \mathrm{Im}\varphi \leq \dim U < \dim V$ 可得 $\dim \mathrm{Ker}\varphi > 0$, 即 $\mathrm{Ker}\varphi \neq 0$. □

例题 4.13 设 φ 是有限维线性空间 V 到 U 的满线性映射, 求证: 必存在 V 的子空间 W, 使得 $V = W \oplus \mathrm{Ker}\varphi$, 且 φ 在 W 上的限制是 W 到 U 上的线性同构.

证明 证法一:取 $\mathrm{Ker}\varphi$ 的一组基 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_k$, 并将其扩张为 V 的一组基 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_k,\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n$. 令 $W=L(\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n)$, 则显然 $V=W\oplus\mathrm{Ker}\varphi$. 由推论 4.4可知,$\varphi(\boldsymbol{e}_{k+1}),\cdots,\varphi(\boldsymbol{e}_n)$ 是 $\mathrm{Im}\varphi=U$ 的一组基, 故 φ 在 W 上的限制将 W 的一组基 $\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n$ 映射为 U 的一组基 $\varphi(\boldsymbol{e}_{k+1}),\cdots,\varphi(\boldsymbol{e}_n)$, 从而由命题 4.7可知 $\varphi|_W$ 必为线性同构.

证法二: 取 W 为 $\mathrm{Ker}\varphi$ 在 V 中的补空间. 对任意的 $\boldsymbol{u}\in U$, 由于 φ 是映上的, 故存在 $\boldsymbol{v}=\boldsymbol{w}+\boldsymbol{v}_1$, 其中 $\boldsymbol{w}\in W,\boldsymbol{v}_1\in\mathrm{Ker}\varphi$, 使得 $\boldsymbol{u}=\varphi(\boldsymbol{v})=\varphi(\boldsymbol{w})$, 于是 φ 在 W 上的限制也是映上的, 故 $\dim U=\dim\mathrm{Im}\varphi|_W$. 另一方面, 由维数公式可知,$\dim W=\dim V-\dim\mathrm{Ker}\varphi=\dim\mathrm{Im}\varphi=\dim U$. 再对 φ 在 W 上的限制用线性映射维数公式可知,$\dim\mathrm{Ker}\varphi|_W=\dim W-\dim\mathrm{Im}\varphi|_W=\dim U-\dim\mathrm{Im}\varphi|_W=0$. 从而它必是单映射, 于是 φ 在 W 上的限制是 W 到 U 上的线性同构. □

例题 4.14 设 φ 是有限维线性空间 V 到 V' 的线性映射,U 是 V' 的子空间且 $U\subseteq\mathrm{Im}\varphi$, 求证:$\varphi^{-1}(U)=\{\boldsymbol{v}\in V|\varphi(\boldsymbol{v})\in U\}$ 是 V 的子空间, 且

$$\dim U+\dim\mathrm{Ker}\varphi=\dim\varphi^{-1}(U).$$

笔记 注意对线性映射做限制这个操作.

证明 容易验证 $\varphi^{-1}(U)$ 是 V 的子空间. 将 φ 限制在 $\varphi^{-1}(U)$ 上, 它是到 U 上的线性映射. 因为 $\mathbf{0}\in U$, 故 $\mathrm{Ker}\varphi\subseteq\varphi^{-1}(U)$. 从而 $\mathrm{Ker}\varphi|_{\varphi^{-1}(U)}=\mathrm{Ker}\varphi\cap\varphi^{-1}(U)=\mathrm{Ker}\varphi$, 又显然 $\mathrm{Im}\varphi|_{\varphi^{-1}(U)}=U$. 再对 φ 在 $\varphi^{-1}(U)$ 上的限制用线性映射维数公式即得

$$\dim\varphi^{-1}(U)=\dim\mathrm{Im}\varphi|_{\varphi^{-1}(U)}+\dim\mathrm{Ker}\varphi|_{\varphi^{-1}(U)}=\dim U+\dim\mathrm{Ker}\varphi.$$

□

注 $\varphi^{-1}(U)=\{\boldsymbol{v}\in V|\varphi(\boldsymbol{v})\in U\}$ 称为 U 在线性映射 φ 下的原像集.

命题 4.24

设 U 是有限维线性空间 V 的子空间,φ 是 V 上的线性变换, 求证:

(1) $\dim U-\dim\mathrm{Ker}\varphi\le\dim\varphi(U)\le\dim U$;

(2) $\dim\varphi^{-1}(U)\le\dim U+\dim\mathrm{Ker}\varphi$.

证明

(1) 注意到当 φ 限制在 U 上时,$\mathrm{Ker}(\varphi|_U)=U\cap\mathrm{Ker}\varphi$, 故由线性映射的维数公式可得

$$\dim U=\dim(U\cap\mathrm{Ker}\varphi)+\dim\varphi(U)\leqslant\dim\mathrm{Ker}\varphi+\dim\varphi(U).$$

于是

$$\dim U-\dim\mathrm{Ker}\varphi\le\dim\varphi(U),$$

而由线性映射维数公式, 可知 $\dim U=\dim\mathrm{Ker}\varphi|_U+\dim\varphi(U)$. 进而 $\dim\varphi(U)\le\dim U$.

(2) 设 $\overline{\varphi}$ 是线性变换 φ 在子空间 $\varphi^{-1}(U)$ 上的限制, 则 $\mathrm{Im}\overline{\varphi}=U\cap\mathrm{Im}\varphi$. 因为 $\mathbf{0}\in U$, 故 $\mathrm{Ker}\varphi\subseteq\varphi^{-1}(U)$. 从而 $\mathrm{Ker}\overline{\varphi}=\mathrm{Ker}\varphi\cap\varphi^{-1}(U)=\mathrm{Ker}\varphi$. 由线性映射的维数公式可得

$$\dim\varphi^{-1}(U)=\dim(U\cap\mathrm{Im}\varphi)+\dim\mathrm{Ker}\varphi.$$

显然, 由 $\dim(U\cap\mathrm{Im}\varphi)\le\dim U$ 可推出

$$\dim\varphi^{-1}(U)\le\dim U+\dim\mathrm{Ker}\varphi.$$

□

例题 4.15 证明: 若 $\boldsymbol{A},\boldsymbol{B}$ 是数域 $\mathbb{F}$ 上两个 n 阶方阵, 则

$$\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})-n\le\mathrm{r}(\boldsymbol{AB})\le\min\{\mathrm{r}(\boldsymbol{A}),\mathrm{r}(\boldsymbol{B})\}.$$

证明 令 V 是 $\mathbb{F}$ 上 n 维列向量空间, 则将 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 看成是 V 上由矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 乘法诱导的线性变换. 又令 $U=\boldsymbol{B}(V)$, 注意到 $\boldsymbol{A}(U)=\boldsymbol{AB}(V)$, 故 $\dim\boldsymbol{A}(U)=\dim\boldsymbol{AB}(V)=\mathrm{r}(\boldsymbol{AB})$,$\dim\mathrm{Ker}\boldsymbol{A}=n-\mathrm{r}(\boldsymbol{A})$, 即线性方程组 $\boldsymbol{Ax}=\mathbf{0}$ 的解空间维数. 而 $\dim U=\dim\boldsymbol{B}(V)=\mathrm{r}(\boldsymbol{B})$, 由命题 4.24(1)的结论, 可得

$$\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})-n=\dim U-\dim\mathrm{Ker}\boldsymbol{A}\leqslant\dim\boldsymbol{A}(U)=\mathrm{r}(\boldsymbol{AB})\leqslant\dim U=\mathrm{r}(\boldsymbol{B}).$$

又显然有 $\dim A(U) \le \dim A(V)$, 故得 $\mathrm{r}(AB) \le \mathrm{r}(A)$. □

4.6 不变子空间

例题 4.16 设线性空间 V 上的线性变换 φ 在基 $\{e_1, e_2, e_3, e_4\}$ 下的表示矩阵为

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

求证:$U = L(e_1 + 2e_2, e_3 + e_4, e_1 + e_2)$ 和 $W = L(e_2 + e_3 + 2e_4)$ 都是 φ 的不变子空间.

证明 要证明由若干个向量生成的子空间是某个线性变换的不变子空间, 通常只需证明这些向量在线性变换的作用下仍在这个子空间中即可. 因此只需证明这些子空间的一组基在线性变换的作用下仍在这个子空间中即可. 注意到 $\varphi(e_1 + 2e_2)$ 的坐标向量为

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

即 $\varphi(e_1 + 2e_2) = e_1 + 2e_2 \in U$. 同理可计算出

$$\varphi(e_3 + e_4) = (e_1 + 2e_2) + (e_3 + e_4) \in U,$$
$$\varphi(e_1 + e_2) = (e_1 + e_2) + (e_3 + e_4) \in U,$$
$$\varphi(e_2 + e_3 + 2e_4) = e_2 + e_3 + 2e_4 \in W,$$

因此结论成立. □

命题 4.25

设 V_1, V_2 是 V 上线性变换 φ 的不变子空间, 任取 $V_0 \subset V$, 求证:,V_0,$V_1 \cap V_2$,$V_1 + V_2$ 也是 φ 的不变子空间.

证明 V_0 是 $\varphi-$ 不变子空间是显然的.

任取 $v \in V_1 \cap V_2$, 则由 $v \in V_i$ 可得 $\varphi(v) \in V_i (i = 1, 2)$, 于是 $\varphi(v) \in V_1 \cap V_2$, 从而 $V_1 \cap V_2$ 是 $\varphi-$ 不变子空间.

任取 $v \in V_1 + V_2$, 则 $v = v_1 + v_2$, 其中 $v_i \in V_i$, 故 $\varphi(v_i) \in V_i (i = 1, 2)$, 于是 $\varphi(v) = \varphi(v_1) + \varphi(v_2) \in V_1 + V_2$, 从而 $V_1 + V_2$ 是 $\varphi-$ 不变子空间. □

命题 4.26

设 φ 是 $n (n \ge 2)$ 维线性空间 V 上的线性变换, 证明以下 $n + 1$ 个结论等价:

(1) V 的任一 1 维子空间都是 $\varphi-$ 不变子空间;

......

(r) V 的任一 r 维子空间都是 $\varphi-$ 不变子空间;

......

(n - 1) V 的任一 $n - 1$ 维子空间都是 $\varphi-$ 不变子空间;

(n)V 本身就是 $\varphi-$ 不变子空间;

(n+1) φ 是纯量变换.

证明 (n+1) $\Rightarrow$ (n) 是显然的. 注意到当 $1 \le i \le n - 2$ 时, 任一 i 维子空间 V_0 都可表示为两个 $i + 1$ 维子空间 V_1, V_2 的交; 而 V 的任意 $n - 1$ 维子空间都是 V 的子空间. 于是由命题 4.25可知:(n) $\Rightarrow$ (n - 1) $\Rightarrow$ (n - 2) $\Rightarrow \cdots \Rightarrow$ (1) 显然成立, 剩下只要证明 (1) $\Rightarrow$ (n+1) 即可. 取 V 的一组基 $\{e_1, e_2, \cdots, e_n\}$, 由 (1) 可知 $L(e_1), L(e_2), \cdots, L(e_n)$ 都是 $\varphi-$ 不变子空间, 设 $\varphi(e_i) = \lambda_i e_i (1 \le i \le n)$. 只要证明 $\lambda_1 = \lambda_2 = \cdots = \lambda_n$ 即可得到 φ 为纯量变换. 用反证法, 不妨设

$\lambda_1 \neq \lambda_2$, 则由 $L(\boldsymbol{e}_1+\boldsymbol{e}_2)$ 也是 $\varphi-$ 不变子空间可设 $\varphi(\boldsymbol{e}_1+\boldsymbol{e}_2)=\lambda_0(\boldsymbol{e}_1+\boldsymbol{e}_2)$, 于是 $(\lambda_1-\lambda_0)\boldsymbol{e}_1+(\lambda_2-\lambda_0)\boldsymbol{e}_2=\boldsymbol{0}$, 从而由 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 线性无关可知 $\lambda_1=\lambda_2=\lambda_0$, 矛盾. □

命题 4.27

设 φ,ψ 是线性空间 V 上的线性变换且 $\varphi\psi=\psi\varphi$, 求证:Imφ 及 Kerφ 都是 ψ 的不变子空间. 同理,Imψ 及 Kerψ 也都是 φ 的不变子空间.

笔记 显然 Imφ 及 Kerφ 都是 φ 自身的不变子空间,Imψ 及 Kerψ 也都是 ψ 自身的不变子空间.

证明 任取 $\boldsymbol{v}\in\mathrm{Im}\varphi$, 即 $\boldsymbol{v}=\varphi(\boldsymbol{u})$, 则 $\psi(\boldsymbol{v})=\psi\varphi(\boldsymbol{u})=\varphi\psi(\boldsymbol{u})\in\mathrm{Im}\varphi$, 即 Im$\varphi$ 是 ψ 的不变子空间.

任取 $\boldsymbol{v}\in\mathrm{Ker}\varphi$, 即 $\varphi(\boldsymbol{v})=0$, 则 $\varphi\psi(\boldsymbol{v})=\psi\varphi(\boldsymbol{v})=0$. 因此,$\psi(\boldsymbol{v})\in\mathrm{Ker}\varphi$, 即 Ker$\varphi$ 是 ψ 的不变子空间. □

命题 4.28

设 φ 是 n 维线性空间 V 上的线性变换,W 为 $\varphi-$ 不变子空间,φ 在 W 上的限制为 $\varphi|_W$, 则 $\varphi|_W$ 的像集与原像集相同且均为 W, 对 $\forall k\in\mathbb{N}$,$(\varphi|_W)^k$ 有意义并且 $\varphi^k|_W=(\varphi|_W)^k$.

证明 因为 W 为 $\varphi-$ 不变子空间, 所以 $\varphi|_W$ 的像集与原像集相同且均为 W 是显然的. 并且对 $\forall\alpha\in W$, 有 $\varphi|_W(\alpha)\in W$. 因此对 $\forall k\in\mathbb{N}$, 有 $(\varphi|_W)^k(\alpha)\in W$. 故 $(\varphi|_W)^k$ 有意义. 下证 $\varphi^k|_W=(\varphi|_W)^k$.

对 $\forall k\in\mathbb{N}$, 显然 $\varphi^k|_W$ 和 $(\varphi|_W)^k$ 的定义域都是 W. 从而对 $\forall a\in W$, 有

$$\varphi^k|_W(a)=\varphi^k(a),$$
$$(\varphi|_W)^k(a)=(\varphi|_W)^{k-1}\varphi|_W(a)=(\varphi|_W)^{k-1}\varphi(a)=\cdots=\varphi^k(a).$$

因此 $\varphi^k|_W(a)=(\varphi|_W)^k(a)=\varphi^k(a)$. 故 $\varphi^k|_W=(\varphi|_W)^k$. □

命题 4.29

设 φ 是 n 维线性空间 V 上的自同构,W 为 $\varphi-$ 不变子空间,φ 在 W 上的限制为 $\varphi|_W$, 则 $\varphi|_W$ 的像集与原像集相同且均为 W,$\varphi|_W$ 是 W 上的自同构并且 $(\varphi|_W)^{-1}=\varphi^{-1}|_W$.

证明

□

例题 4.17 设 $\boldsymbol{A}$ 为数域 $\mathbb{K}$ 上的 n 阶幂零阵,$\boldsymbol{B}$ 为 n 阶方阵, 满足 $\boldsymbol{AB}=\boldsymbol{BA}$ 且 $\mathrm{r}(\boldsymbol{AB})=\mathrm{r}(\boldsymbol{B})$. 求证:$\boldsymbol{B}=\boldsymbol{O}$.

注 因为 Im$\boldsymbol{B}$ 是 $\boldsymbol{A}-$ 不变子空间, 所以 $\boldsymbol{A}|_{\mathrm{Im}\boldsymbol{B}}(\mathrm{Im}\boldsymbol{B})\in\mathrm{Im}\boldsymbol{B}$. 因此 $(\boldsymbol{A}|_{\mathrm{Im}\boldsymbol{B}})^k$ 有意义. 对于一般的限制 W,$(\boldsymbol{A}|_W)^k$ 不一定有意义. 见命题 4.28.

证明 将 $\boldsymbol{A},\boldsymbol{B}$ 都看成是 $\mathbb{K}^n$ 上 (由矩阵 $\boldsymbol{A},\boldsymbol{B}$ 乘法诱导) 的线性变换, 设 $\boldsymbol{A}^k=\boldsymbol{O}$, 其中 k 为正整数. 由 $\boldsymbol{AB}=\boldsymbol{BA}$ 以及命题 4.27可知 Im$\boldsymbol{B}$ 是 $\boldsymbol{A}-$ 不变子空间. 考虑 $\boldsymbol{A}$ 在 Im$\boldsymbol{B}$ 上的限制 $\boldsymbol{A}|_{\mathrm{Im}\boldsymbol{B}}$, 其像空间的维数 $\dim\boldsymbol{AB}(\mathbb{K}^n)=\mathrm{r}(\boldsymbol{AB})=\mathrm{r}(\boldsymbol{B})=\dim\mathrm{Im}\boldsymbol{B}$, 故 $\boldsymbol{A}|_{\mathrm{Im}\boldsymbol{B}}$ 是 Im$\boldsymbol{B}$ 上的满线性变换. 于是由命题 4.28和满射的复合仍是满射可知 $(\boldsymbol{A}|_{\mathrm{Im}\boldsymbol{B}})^k=\boldsymbol{A}^k|_{\mathrm{Im}\boldsymbol{B}}=\boldsymbol{O}|_{\mathrm{Im}\boldsymbol{B}}$ 也是 Im$\boldsymbol{B}$ 上的满线性变换, 从而只能是 Im$\boldsymbol{B}=0$, 即 $\boldsymbol{B}=\boldsymbol{O}$. □

命题 4.30

设 φ 是 n 维线性空间 V 上的自同构, 若 W 是 φ 的不变子空间, 求证:W 也是 φ^{-1} 的不变子空间.

证明 将 φ 限制在 W 上, 得到 $\varphi: W\to W$. 它是 W 上的线性变换. 由于 φ 是单映射, 故它在 W 上的限制也是单映射, 从而由推论 4.3可知,φ 在 W 上的限制也是满映射, 即它是 W 上的自同构, 于是结合命题 4.29可知 $W=\varphi|_W(W)=\varphi(W)$, 对其两边同时取 φ^{-1} 可得 $\varphi^{-1}(W)=W$. 故结论得证. □

注 如果 V 是无限维线性空间, 则这个命题的结论一般并不成立. 例如,$V=\mathbb{K}[x^{-1},x]$ 是由数域 $\mathbb{K}$ 上的 Laurent 多项式 $f(x)=\sum\limits_{i=-m}^{n}a_ix^i(m,n\in\mathbb{N})$ 构成的线性空间,V 上的线性变换 φ,ψ 定义为 $\varphi(f(x))=xf(x)$,$\psi(f(x))=x^{-1}f(x)$. 显然,φ,ψ 互为逆映射, 从而都是自同构. 注意到 $W=\mathbb{K}[x]$ 是 V 的 $\varphi-$ 不变子空间, 但 W 显然不是 $\varphi^{-1}-$ 不变子空间.

例题 4.18 设 V 是次数小于 n 的实系数多项式组成的线性空间,D 是 V 上的求导变换. 求证:D 的任一 $k(k\geq 1)$ 维不变子空间必是由 $\{1,x,\cdots,x^{k-1}\}$ 生成的子空间. 特别地, 向量 1 包含在 D 的任一非零不变子空间中.

证明 任取 D 的一个 $k(k\geq 1)$ 维不变子空间 V_0, 再取出 V_0 中次数最高的一个多项式 (不唯一)$f(x)=a_lx^l+a_{l-1}x^{l-1}+\cdots+a_1x+a_0$, 其中 $a_l\neq 0$. 注意到 V_0 是 $D-$ 不变子空间, 由 $D^lf(x)=a_ll!\in V_0$ 可得 $1\in V_0$; 由 $D^{l-1}f(x)=a_ll!x+a_{l-1}(l-1)!\in V_0$ 可得 $x\in V_0$;$\cdots$; 由 $Df(x)=a_llx^{l-1}+a_{l-1}(l-1)x^{l-2}+\cdots+a_1\in V_0$ 可得 $x^{l-1}\in V_0$; 最后由 $f(x)\in V_0$ 可得 $x^l\in V_0$. 因为 V_0 中所有多项式的次数都小于等于 l, 所以 $\{1,x,\cdots,x^l\}$ 构成了 V_0 的一组基, 于是 $k=\dim V_0=l+1$, 即 $l=k-1$, 从而结论得证. □

例题 4.19 设 φ 是 n 维线性空间 V 上的线性变换,φ 在 V 的一组基下的表示矩阵为对角阵且主对角线上的元素互不相同, 求 φ 的所有不变子空间.

证明 证法一: 设 φ 在基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 下的表示矩阵为 $\mathrm{diag}\{d_1,d_2,\cdots,d_n\}$, 其中 $d_1,d_2,\cdots,d_n$ 互不相同, 则 $\varphi(\boldsymbol{e}_i)=d_i\boldsymbol{e}_i$. 对任意的指标集 $1\leq i_1<i_2<\cdots<i_r\leq n$, 容易验证 $U=L(\boldsymbol{e}_{i_1},\boldsymbol{e}_{i_2},\cdots,\boldsymbol{e}_{i_r})$ 是 φ 的不变子空间. 注意到 $1,2,\cdots,n$ 的子集共有 2^n 个 (空集对应于零子空间), 故上述形式的 $\varphi-$ 不变子空间共有 2^n 个. 下面我们证明 φ 的任一不变子空间都是上述不变子空间之一.

任取 φ 的非零不变子空间 U, 设指标集

$$I=\{i\in[1,n]|\text{存在某个}\boldsymbol{\alpha}\in U,\text{使得}\boldsymbol{\alpha}=c\boldsymbol{e}_i+\cdots,\text{其中}c\neq 0\}.$$

因为 $U\neq 0$, 故 $I\neq\varnothing$, 不妨设 $I=\{i_1,i_2,\cdots,i_r\}$. 由指标集 I 的定义可知,$U\subseteq L(\boldsymbol{e}_{i_1},\boldsymbol{e}_{i_2},\cdots,\boldsymbol{e}_{i_r})$. 下面我们证明 $\boldsymbol{e}_{i_j}\in U(j=1,2,\cdots,r)$ 成立. 不失一般性, 我们只需证明 $\boldsymbol{e}_{i_1}\in U$ 即可. 由指标集 I 的定义可知, 存在 $\boldsymbol{\alpha}\in U$, 使得

$$\boldsymbol{\alpha}=c_1\boldsymbol{e}_{i_1}+c_2\boldsymbol{e}_{i_2}+\cdots+c_k\boldsymbol{e}_{i_k},$$

其中 $c_1,c_2,\cdots,c_k$ 都是非零常数. 将上式作用 φ^l, 可得

$$\varphi^l(\boldsymbol{\alpha})=c_1d_{i_1}^l\boldsymbol{e}_{i_1}+c_2d_{i_2}^l\boldsymbol{e}_{i_2}+\cdots+c_kd_{i_k}^l\boldsymbol{e}_{i_k},l=1,2,\cdots,k-1.$$

因此, 我们有

$$(\boldsymbol{\alpha},\varphi(\boldsymbol{\alpha}),\cdots,\varphi^{k-1}(\boldsymbol{\alpha}))=(\boldsymbol{e}_{i_1},\boldsymbol{e}_{i_2},\cdots,\boldsymbol{e}_{i_k})\begin{pmatrix}c_1&c_1d_{i_1}&\cdots&c_1d_{i_1}^{k-1}\\c_2&c_2d_{i_2}&\cdots&c_2d_{i_2}^{k-1}\\\vdots&\vdots&&\vdots\\c_k&c_kd_{i_k}&\cdots&c_kd_{i_k}^{k-1}\end{pmatrix}.$$

上式右边的矩阵记为 $\boldsymbol{A}$, 由于 $|\boldsymbol{A}|=c_1c_2\cdots c_k\prod\limits_{1\leq r<s\leq k}(d_{i_s}-d_{i_r})\neq 0$, 故 $\boldsymbol{A}$ 为可逆矩阵, 从而

$$(\boldsymbol{e}_{i_1},\boldsymbol{e}_{i_2},\cdots,\boldsymbol{e}_{i_k})=(\boldsymbol{\alpha},\varphi(\boldsymbol{\alpha}),\cdots,\varphi^{k-1}(\boldsymbol{\alpha}))\boldsymbol{A}^{-1},$$

特别地,$\boldsymbol{e}_{i_1}$ 可以表示为 $\boldsymbol{\alpha},\varphi(\boldsymbol{\alpha}),\cdots,\varphi^{k-1}(\boldsymbol{\alpha})$ 的线性组合. 因为 U 是 φ 的不变子空间, 故 $\boldsymbol{\alpha},\varphi(\boldsymbol{\alpha}),\cdots,\varphi^{k-1}(\boldsymbol{\alpha})$ 都是 U 中的向量, 从而 $\boldsymbol{e}_{i_1}\in U$, 因此 $U=L(\boldsymbol{e}_{i_1},\boldsymbol{e}_{i_2},\cdots,\boldsymbol{e}_{i_r})$.

又因为任取 $j_1,j_2,\cdots,j_m\in\{1,2,\cdots,n\}$, 都有

$$L(\boldsymbol{e}_{j_1},\boldsymbol{e}_{j_2},\cdots,\boldsymbol{e}_{j_r})=V_{j_1}\oplus V_{j_2}\oplus\cdots\oplus V_{j_m}.$$

而特征子空间 $V_{j_k}(k=1,2,\cdots,m)$ 都是 φ 的不变子空间,φ 的不变子空间的直和仍是不变子空间, 所以 $L(\boldsymbol{e}_{j_1},\boldsymbol{e}_{j_2},\cdots,\boldsymbol{e}_{j_r})$ 也是 φ 的不变子空间.

综上所述,φ 的不变子空间共有 2^n 个.

证法二: 设线性变换 $\boldsymbol{\varphi}$ 在 V 的一组基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下的表示矩阵是对角矩阵 $\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$, 且 λ_i 互不相同, 因此 $\boldsymbol{\varphi}$ 可对角化, $\boldsymbol{\varphi}$ 有 n 个不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 且 $\boldsymbol{\varphi}(\boldsymbol{e}_i)=\lambda_i\boldsymbol{e}_i(1\leq i\leq n)$. 此时, 特征值 λ_i 的特征子空间 $V_i=L(\boldsymbol{e}_i)$, 并且 $V=V_1\oplus V_2\oplus\cdots\oplus V_n$.

任取 $\boldsymbol{\varphi}$ 的非零不变子空间 U 以及 U 的一组基, 并将这组基扩张为 V 的一组基, 则 $\boldsymbol{\varphi}$ 在这组基下的表示矩阵是分块上三角矩阵 $\begin{pmatrix}\boldsymbol{A}&\boldsymbol{C}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}$, 其中 $\boldsymbol{A}$ 是 $\boldsymbol{\varphi}|_U$ 的表示矩阵, 不妨设为 r 阶矩阵. 考虑到

$$|\lambda\boldsymbol{I}_V-\boldsymbol{\varphi}|=|\lambda\boldsymbol{I}-\boldsymbol{A}||\lambda\boldsymbol{I}-\boldsymbol{B}|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n),$$

故 $\boldsymbol{A}$ 或 $\boldsymbol{\varphi}|_U$ 有 r 个不同的特征值, 设为 $\lambda_{i_1},\lambda_{i_2},\cdots,\lambda_{i_r}$. 考虑 $\boldsymbol{\varphi}|_U$ 关于特征值 λ_{i_j} 的特征子空间 $U_{i_j}=\{\boldsymbol{u}\in U|\boldsymbol{\varphi}(\boldsymbol{u})=\lambda_{i_j}\boldsymbol{u}\}$, 由于 $U_{i_j}=U\cap V_{i_j}$ 且 $\dim V_{i_j}=1$, 故只能是 $U_{i_j}=V_{i_j}=L(\boldsymbol{e}_{i_j})(1\leq j\leq r)$. 因为 $\boldsymbol{\varphi}|_U$ 有 r 个不同的特征值, 所以 $\boldsymbol{\varphi}|_U$ 可对角化, 于是

$$U=U_{i_1}\oplus U_{i_2}\oplus\cdots\oplus U_{i_r}=L(\boldsymbol{e}_{i_1},\boldsymbol{e}_{i_2},\cdots,\boldsymbol{e}_{i_r}).$$

又因为任取 $j_1,j_2,\cdots,j_m\in\{1,2,\cdots,n\}$, 都有

$$L(\boldsymbol{e}_{j_1},\boldsymbol{e}_{j_2},\cdots,\boldsymbol{e}_{j_r})=V_{j_1}\oplus V_{j_2}\oplus\cdots\oplus V_{j_m}.$$

而特征子空间 $V_{j_k}(k=1,2,\cdots,m)$ 都是 φ 的不变子空间,φ 的不变子空间的直和仍是不变子空间, 所以 $L(\boldsymbol{e}_{j_1},\boldsymbol{e}_{j_2},\cdots,\boldsymbol{e}_{j_r})$ 也是 φ 的不变子空间.

综上所述,φ 的不变子空间共有 2^n 个. □

定理 4.6

设 φ 是数域 $\mathbb{F}$ 上向量空间 V 上的线性变换,W 是 φ 的不变子空间. 若取 W 的一组基 $\{e_1,\cdots,e_r\}$, 再扩张为 V 的一组基 $\{e_1,\cdots,e_r,e_{r+1},\cdots,e_n\}$, 则 φ 在这组基下的表示矩阵具有下列分块上三角矩阵的形状:

$$\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{O} & \boldsymbol{A}_{22}\end{pmatrix},$$

其中 $\boldsymbol{A}_{11}$ 是一个 r 阶矩阵. ♡

定理 4.7

设 φ 是数域 $\mathbb{F}$ 上向量空间 V 上的线性变换,$V_1,V_2,\cdots,V_m$ 是 φ 的不变子空间且 $V=V_1\oplus V_2\oplus\cdots\oplus V_m$. 若取 V_i 的基拼成 V 的一组基 $\{e_1,e_2,\cdots,e_n\}$, 则 φ 在这组基下的表示矩阵具有下列分块对角矩阵的形状:

$$\begin{pmatrix}\boldsymbol{A}_{11} & & & \\ & \boldsymbol{A}_{22} & & \\ & & \ddots & \\ & & & \boldsymbol{A}_{mm}\end{pmatrix}.$$

♡

命题 4.31

设 φ 是 n 维线性空间 V 上的线性变换,U 是 r 维 $\varphi-$ 不变子空间. 取 U 的一组基 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_r\}$, 并扩张为 V 的一组基 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_r,\boldsymbol{e}_{r+1},\cdots,\boldsymbol{e}_n\}$. 设 φ 在这组基下的表示矩阵 $\boldsymbol{A}=(a_{ij})=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{O} & \boldsymbol{A}_{22}\end{pmatrix}$ 为分块上三角阵, 其中 $\boldsymbol{A}_{11}$ 是 φ 在不变子空间 U 上的限制 $\varphi|_U$ 在基 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_r\}$ 下的表示矩阵. 证明:φ 诱导的变换 $\overline{\varphi}(\boldsymbol{v}+U)=\varphi(\boldsymbol{v})+U$ 是商空间 V/U 上的线性变换, 并且在 V/U 的一组基 $\{\boldsymbol{e}_{r+1}+U,\cdots,\boldsymbol{e}_n+U\}$ 下的表示矩阵为 $\boldsymbol{A}_{22}$. ♠

证明 由 U 是 $\varphi-$ 不变子空间容易验证 $\overline{\varphi}$ 的定义不依赖于 $U-$ 陪集代表元的选取, 从而是定义好的变换. $\overline{\varphi}$ 的线性由 φ 的线性即得. 由 φ 的表示矩阵为 $\boldsymbol{A}$, 再结合 $\boldsymbol{e}_1,\boldsymbol{e}_2\cdots,\boldsymbol{e}_r\in U$ 及U-陪集的性质 (2)和商空间的加法和数乘的定义可得

$$\begin{aligned}\overline{\varphi}(\boldsymbol{e}_{r+1}+U)&=\varphi(\boldsymbol{e}_{r+1})+U=a_{1,r+1}\boldsymbol{e}_1+\cdots+a_{r,r+1}\boldsymbol{e}_r+a_{r+1,r+1}\boldsymbol{e}_{r+1}+\cdots+a_{n,r+1}\boldsymbol{e}_n+U\\&=a_{r+1,r+1}\boldsymbol{e}_{r+1}+\cdots+a_{n,r+1}\boldsymbol{e}_n+U=a_{r+1,r+1}(\boldsymbol{e}_{r+1}+U)+\cdots+a_{n,r+1}(\boldsymbol{e}_n+U),\\&\cdots\cdots\cdots\cdots\\ \overline{\varphi}(\boldsymbol{e}_n+U)&=a_{r+1,n}(\boldsymbol{e}_{r+1}+U)+\cdots+a_{n,n}(\boldsymbol{e}_n+U),\end{aligned}$$

故 $\overline{\varphi}$ 在基 $\{\boldsymbol{e}_{r+1}+U,\cdots,\boldsymbol{e}_n+U\}$ 下的表示矩阵为 $\boldsymbol{A}_{22}$. □

4.7 幂等变换

定义 4.2 (幂等变换)

线性变换 φ 若满足 $\varphi^2=\varphi$, 则称为**幂等变换**.

定义 4.3 (投影变换)

设 $V=V_1\oplus V_2\oplus\cdots\oplus V_m$ 为线性空间 V 关于子空间 $V_i(1\le i\le m)$ 的直和分解, 则 V 中任一向量 $\boldsymbol{v}$ 可唯一地分解为 $\boldsymbol{v}=\boldsymbol{v}_1+\boldsymbol{v}_2+\cdots+\boldsymbol{v}_m$, 其中 $\boldsymbol{v}_i\in V_i$. 定义 $\varphi_i:V\to V,\varphi_i(\boldsymbol{v})=\boldsymbol{v}_i(1\le i\le m)$, 容易验证 φ_i 是 V 上的线性变换, 称为 V 到 V_i 上的**投影变换**.

命题 4.32 (投影变换的性质)

设 $V=V_1\oplus V_2\oplus\cdots\oplus V_m$ 为线性空间 V 关于子空间 $V_i(1\le i\le m)$ 的直和分解,φ_i 为 V 到 V_i 上的投影变换. 投影变换 φ_i 满足如下性质:

(1) $\varphi_i^2=\varphi_i,\varphi_i\varphi_j=\boldsymbol{0}(i\neq j),\boldsymbol{I}_V=\varphi_1+\varphi_2+\cdots+\varphi_m$;

(2) $\mathrm{Im}\varphi_i=V_i,\mathrm{Ker}\varphi_i=\bigoplus\limits_{j\neq i}V_j,V=\mathrm{Im}\varphi_i\oplus\mathrm{Ker}\varphi_i$.

(3) 投影变换 φ_i 都是幂等变换;

(4) 若取 V_i 的一组基拼成 V 的一组基, 则 φ_i 在这组基下的表示矩阵为 $\mathrm{diag}\{0,\cdots,0,1,\cdots,1,0,\cdots,0\}$, 其中有 $\dim V_i$ 个 1;

(5) $V=\mathrm{Im}\varphi_1\oplus\mathrm{Im}\varphi_2\oplus\cdots\oplus\mathrm{Im}\varphi_m$;

(6) $\mathrm{Ker}\varphi_1\cap\mathrm{Ker}\varphi_2\cap\cdots\cap\mathrm{Ker}\varphi_m=0$.

笔记 提示: 两个集合相等等价于这两个集合相互包含.

证明

(1) 证明是显然的.

(2) 证明是显然的.

(3) 由 (1) 易得.

(4) 由 (2) 易得.

(5) 由 (2) 易知 $V=\mathrm{Im}\varphi_1\oplus\mathrm{Im}\varphi_2\oplus\cdots\oplus\mathrm{Im}\varphi_m$.

(6) 任取 $\alpha\in\mathrm{Ker}\varphi_1\cap\mathrm{Ker}\varphi_2\cap\cdots\cap\mathrm{Ker}\varphi_m$. 由投影变换性质 (2), 可知 $\mathrm{Ker}\varphi_i=\bigoplus\limits_{j\neq i}V_j$. 于是对任意整数 $i,j\in[1,m]$ 且 $i\neq j$, 有 $\alpha\in\mathrm{Ker}\varphi_i\cap\mathrm{Ker}\varphi_j=\bigoplus\limits_{k\neq i}V_k\cap\bigoplus\limits_{k\neq j}V_k$. 从而

$$\alpha=v_1+\cdots+v_{i-1}+v_{i+1}+\cdots+v_m=u_1+\cdots+u_{j-1}+u_{j+1}+\cdots+u_m,$$

其中 $v_k,u_k\in V_k$. 上式经整理可得 $v_j-u_i=\sum\limits_{k\neq i,j}(u_k-v_k)\in\bigoplus\limits_{k\neq i,j}V_k$. 又 $v_j-u_i\in V_i\oplus V_j$. 故 $v_j-u_i\in(V_i\oplus V_j)\cap\bigoplus\limits_{k\neq i,j}V_k$. 而由于 $V=\bigoplus\limits_{k=1}^{m}V_k$, 因此 $(V_i\oplus V_j)\cap\bigoplus\limits_{k\neq i,j}V_k=\boldsymbol{0}$. 故 $v_j-u_i=\boldsymbol{0}$, 从而 $v_j=u_i\in V_i\cap V_j$. 又因为 $V_i\oplus V_j$, 所以 $V_i\cap V_j=\boldsymbol{0}$. 故 $v_j=u_i=\boldsymbol{0}$. 再由 i,j 的任意性可知,$v_i=\boldsymbol{0},i=1,2,\cdots,m$. 因此 $\alpha=\sum\limits_{k\neq i}v_k=\boldsymbol{0}$. 故 $\mathrm{Ker}\varphi_1\cap\mathrm{Ker}\varphi_2\cap\cdots\cap\mathrm{Ker}\varphi_m=\boldsymbol{0}$.

□

命题 4.33

设 φ 是 n 维线性空间 V 上的幂等变换, 证明:$V=U\oplus W$, 其中 $U=\mathrm{Im}\varphi=\mathrm{Ker}(\boldsymbol{I}_V-\varphi),W=\mathrm{Im}(\boldsymbol{I}_V-\varphi)=\mathrm{Ker}\varphi$, 且 φ 就是 V 到 U 上的投影变换.

笔记 由上述命题可知n **维线性空间 V 上的幂等变换 φ 也是 V 到 $\mathrm{Im}\varphi$ 上的投影变换**. 于是由命题 4.33和命题 4.32可知, 投影变换可以看作幂等变换, 幂等变换也可以看作投影变换.(即幂等变换和投影变换等价)

证明 因为 $\varphi^2=\varphi$, 故 $\mathrm{Im}\varphi\subseteq\mathrm{Ker}(\boldsymbol{I}-\varphi)$,$\mathrm{Im}(\boldsymbol{I}-\varphi)\subseteq\mathrm{Ker}\varphi$. 对任意的 $\boldsymbol{\alpha}\in V$,$\varphi(\boldsymbol{\alpha})\in\mathrm{Ker}(\boldsymbol{I}-\varphi)$,$(\boldsymbol{I}-\varphi)(\boldsymbol{\alpha})\in\mathrm{Ker}\varphi$, 于是 $\boldsymbol{\alpha}=(\boldsymbol{I}-\varphi)(\boldsymbol{\alpha})+\varphi(\boldsymbol{\alpha})\in\mathrm{Ker}\varphi+\mathrm{Ker}(\boldsymbol{I}-\varphi)$, 从而 $V=\mathrm{Ker}\varphi+\mathrm{Ker}(\boldsymbol{I}-\varphi)$. 任取 $\boldsymbol{\beta}\in\mathrm{Ker}\varphi\cap\mathrm{Ker}(\boldsymbol{I}-\varphi)$, 则 $\boldsymbol{\beta}=(\boldsymbol{I}-\varphi)(\boldsymbol{\beta})+\varphi(\boldsymbol{\beta})=\mathbf{0}$, 即 $\mathrm{Ker}\varphi\cap\mathrm{Ker}(\boldsymbol{I}-\varphi)=0$. 因此,$V=\mathrm{Ker}\varphi\oplus\mathrm{Ker}(\boldsymbol{I}-\varphi)$. 特别地, 由维数公式可得 $\dim\mathrm{Im}\varphi=\dim\mathrm{Ker}(\boldsymbol{I}-\varphi)$,$\dim\mathrm{Im}(\boldsymbol{I}-\varphi)=\dim\mathrm{Ker}\varphi$, 从而 $\mathrm{Im}\varphi=\mathrm{Ker}(\boldsymbol{I}-\varphi)$,$\mathrm{Im}(\boldsymbol{I}-\varphi)=\mathrm{Ker}\varphi$.

令 $U=\mathrm{Im}\varphi=\mathrm{Ker}(\boldsymbol{I}-\varphi)$,$W=\mathrm{Im}(\boldsymbol{I}-\varphi)=\mathrm{Ker}\varphi$, 则 $V=U\oplus W$. 注意到对任意的 $\boldsymbol{\alpha}\in V$,$\boldsymbol{\alpha}=\varphi(\boldsymbol{\alpha})+(\boldsymbol{I}-\varphi)(\boldsymbol{\alpha})$, 其中 $\varphi(\boldsymbol{\alpha})\in U$,$(\boldsymbol{I}-\varphi)(\boldsymbol{\alpha})\in W$, 故 φ 就是 V 到 U 上的投影变换. □

推论 4.6

对线性空间 V 上的幂等变换 φ, 总存在 V 的一组基 (它由 $\mathrm{Im}\varphi$ 的基和 $\mathrm{Ker}\varphi$ 的基拼成), 使得 φ 在这组基下的表示矩阵为下列对角矩阵:

$$\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix},$$

其中 $\boldsymbol{I}_r$ 为 r 阶单位矩阵,r 等于 $\dim\mathrm{Im}\varphi$, 即 φ 的像空间的维数.

证明 由这个命题 4.33和投影变换的性质容易证明. □

命题 4.34

设 $\boldsymbol{A}$ 是数域 $\mathbb{F}$ 上的 n 阶幂等矩阵, 求证:

(1) 存在 n 阶非异阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}$, 其中 $r=\mathrm{r}(\boldsymbol{A})$;

(2) $\mathrm{r}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{A})$.

证明 将 $\boldsymbol{A}$ 看成是 n 维列向量空间 $\mathbb{F}^n$ 上 (由矩阵 $\boldsymbol{A}$ 乘法诱导) 的线性变换, 则它是幂等变换, 因此由推论 4.6即得 (1).

注意到 $\mathrm{tr}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{P}^{-1}\boldsymbol{AP})=\mathrm{tr}\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}=r=\mathrm{r}(\boldsymbol{A})$, 故 (2) 也成立. □

例题 4.20 设 $\boldsymbol{A},\boldsymbol{B}$ 是数域 $\mathbb{F}$ 上的 n 阶幂等矩阵, 且 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的秩相同, 求证: 必存在 $\mathbb{F}$ 上的 n 阶可逆矩阵 $\boldsymbol{C}$, 使得 $\boldsymbol{CB}=\boldsymbol{AC}$.

证明 由命题 4.34可知,$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 均相似于矩阵 $\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}$, 于是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似, 即存在可逆矩阵 $\boldsymbol{C}$, 使得 $\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{AC}$, 即 $\boldsymbol{CB}=\boldsymbol{AC}$. □

命题 4.35

设 φ,ψ 是 n 维线性空间 V 上的幂等线性变换, 求证:

(1) $\mathrm{Im}\varphi=\mathrm{Im}\psi$ 的充要条件是 $\varphi\psi=\psi$,$\psi\varphi=\varphi$;

(2) $\mathrm{Ker}\varphi=\mathrm{Ker}\psi$ 的充要条件是 $\varphi\psi=\varphi$,$\psi\varphi=\psi$.

笔记 也可由幂等变换等价于投影变换来给出直观的几何证明.

证明 (1) 由 $\psi=\varphi\psi$ 可得 $\mathrm{Im}\psi\subseteq\mathrm{Im}\varphi$. 同理由 $\varphi=\psi\varphi$ 可得 $\mathrm{Im}\varphi\subseteq\mathrm{Im}\psi$. 因此 $\mathrm{Im}\varphi=\mathrm{Im}\psi$.

反之, 若 $\mathrm{Im}\varphi=\mathrm{Im}\psi$, 则对任意的 $\boldsymbol{\alpha}\in V$,$\psi(\boldsymbol{\alpha})\in\mathrm{Im}\psi=\mathrm{Im}\varphi$, 故存在 $\boldsymbol{\beta}\in V$, 使得 $\psi(\boldsymbol{\alpha})=\varphi(\boldsymbol{\beta})$. 注意到 $\varphi^2=\varphi$, 故 $\varphi\psi(\boldsymbol{\alpha})=\varphi^2(\boldsymbol{\beta})=\varphi(\boldsymbol{\beta})=\psi(\boldsymbol{\alpha})$, 于是 $\varphi\psi=\psi$. 同理可证 $\psi\varphi=\varphi$.

(2) 设 $\varphi\psi=\varphi$,$\psi\varphi=\psi$. 对任意的 $\boldsymbol{\alpha}\in\mathrm{Ker}\varphi$, 即 $\varphi(\boldsymbol{\alpha})=\mathbf{0}$, 有 $\psi(\boldsymbol{\alpha})=\psi\varphi(\boldsymbol{\alpha})=\mathbf{0}$, 即 $\boldsymbol{\alpha}\in\mathrm{Ker}\psi$, 于是 $\mathrm{Ker}\varphi\subseteq\mathrm{Ker}\psi$. 同理可证 $\mathrm{Ker}\psi\subseteq\mathrm{Ker}\varphi$, 因此 $\mathrm{Ker}\varphi=\mathrm{Ker}\psi$.

反之, 设 $\mathrm{Ker}\varphi=\mathrm{Ker}\psi$. 对任意的 $\boldsymbol{\alpha}\in V$, 有 $\psi(\boldsymbol{\alpha}-\psi(\boldsymbol{\alpha}))=\psi(\boldsymbol{\alpha})-\psi^2(\boldsymbol{\alpha})=\mathbf{0}$, 因此 $\boldsymbol{\alpha}-\psi(\boldsymbol{\alpha})\in\mathrm{Ker}\psi=\mathrm{Ker}\varphi$, 从而 $\varphi(\boldsymbol{\alpha}-\psi(\boldsymbol{\alpha}))=\mathbf{0}$, 即 $\varphi(\boldsymbol{\alpha})=\varphi\psi(\boldsymbol{\alpha})$, 于是 $\varphi=\varphi\psi$. 同理可证 $\psi\varphi=\psi$. □

命题 4.36

设 φ,ψ 是 n 维线性空间 V 上的幂等线性变换, 求证:

(1) $\varphi+\psi$ 是幂等变换的充要条件是 $\varphi\psi=\psi\varphi=0$;

(2) $\varphi-\psi$ 是幂等变换的充要条件是 $\varphi\psi=\psi\varphi=\psi$.

笔记 也可由幂等变换等价于投影变换来给出直观的几何证明.

证明 充分性容易验证, 下面证明必要性.

(1) 若 $(\varphi+\psi)^2=\varphi+\psi$, 则 $\varphi\psi+\psi\varphi=0$, 即 $\varphi\psi=-\psi\varphi$. 将上式两边分别左乘及右乘 φ, 可得 $\varphi\psi\varphi=-\varphi\psi=-\psi\varphi$. 因此 $\varphi\psi=\psi\varphi=0$.

(2) 若 $(\varphi-\psi)^2=\varphi-\psi$, 则 $\varphi\psi+\psi\varphi=2\psi$. 将上式两边分别左乘及右乘 φ, 可得 $\varphi\psi\varphi=\varphi\psi=\psi\varphi$. 因此 $\varphi\psi=\psi\varphi=\psi$. □

命题 4.37

设 $\varphi_1,\cdots,\varphi_m$ 是 n 维线性空间 V 上的线性变换, 且适合条件:

$$\varphi_i^2=\varphi_i,\ \varphi_i\varphi_j=0\ (i\neq j),\ \mathrm{Ker}\varphi_1\cap\cdots\cap\mathrm{Ker}\varphi_m=0.$$

求证:$V=\mathrm{Im}\varphi_1\oplus\mathrm{Im}\varphi_2\oplus\cdots\oplus\mathrm{Im}\varphi_m$.

证明 任取 $\boldsymbol{\alpha}\in\mathrm{Im}\varphi_i\cap(\sum\limits_{j\neq i}\mathrm{Im}\varphi_j)$, 设 $\boldsymbol{\alpha}=\varphi_i(\boldsymbol{\beta})$, 其中 $\boldsymbol{\beta}\in V$, 则 $\varphi_i(\boldsymbol{\alpha})=\varphi_i^2(\boldsymbol{\beta})=\varphi_i(\boldsymbol{\beta})=\boldsymbol{\alpha}$. 又可设

$$\boldsymbol{\alpha}=\varphi_1(\boldsymbol{\alpha}_1)+\cdots+\varphi_{i-1}(\boldsymbol{\alpha}_{i-1})+\varphi_{i+1}(\boldsymbol{\alpha}_{i+1})+\cdots+\varphi_m(\boldsymbol{\alpha}_m),$$

于是

$$\boldsymbol{\alpha}=\varphi_i(\boldsymbol{\alpha})=\varphi_i(\varphi_1(\boldsymbol{\alpha}_1)+\cdots+\varphi_{i-1}(\boldsymbol{\alpha}_{i-1})+\varphi_{i+1}(\boldsymbol{\alpha}_{i+1})+\cdots+\varphi_m(\boldsymbol{\alpha}_m))=0.$$

因此 $\mathrm{Im}\varphi_i\cap(\sum\limits_{j\neq i}\mathrm{Im}\varphi_j)=0$.

对 V 中任一向量 $\boldsymbol{\alpha}$ 以及任意的 i, 有

$$\varphi_i(\boldsymbol{\alpha}-(\varphi_1(\boldsymbol{\alpha})+\cdots+\varphi_m(\boldsymbol{\alpha})))=\varphi_i(\boldsymbol{\alpha})-\varphi_i^2(\boldsymbol{\alpha})=0,$$

因此

$$\boldsymbol{\alpha}-(\varphi_1(\boldsymbol{\alpha})+\cdots+\varphi_m(\boldsymbol{\alpha}))\in\mathrm{Ker}\varphi_1\cap\cdots\cap\mathrm{Ker}\varphi_m=0,$$

从而 $\boldsymbol{\alpha}-(\varphi_1(\boldsymbol{\alpha})+\cdots+\varphi_m(\boldsymbol{\alpha}))=0$, 即 $\boldsymbol{\alpha}=\varphi_1(\boldsymbol{\alpha})+\cdots+\varphi_m(\boldsymbol{\alpha})$, 于是 $V=\mathrm{Im}\varphi_1+\cdots+\mathrm{Im}\varphi_m$. 这就证明了 V 是 $\mathrm{Im}\varphi_1,\cdots,\mathrm{Im}\varphi_m$ 的直和. □

命题 4.38

设 $\varphi,\varphi_1,\cdots,\varphi_m$ 是 n 维线性空间 V 上的线性变换, 满足:$\varphi^2=\varphi$ 且 $\varphi=\varphi_1+\varphi_2+\cdots+\varphi_m$. 求证:$\mathrm{r}(\varphi)=\mathrm{r}(\varphi_1)+\mathrm{r}(\varphi_2)+\cdots+\mathrm{r}(\varphi_m)$ 成立的充要条件是 $\varphi_i^2=\varphi_i,\varphi_i\varphi_j=0\ (i\neq j)$.

笔记 $\mathrm{r}(\varphi)=\mathrm{r}(\varphi_1)+\mathrm{r}(\varphi_2)+\cdots+\mathrm{r}(\varphi_m)$ 等价于 $\dim\mathrm{Im}\varphi=\dim\mathrm{Im}\varphi_1+\dim\mathrm{Im}\varphi_2+\cdots+\dim\mathrm{Im}\varphi_m$.

证明 证法一 (几何方法): 令 $V_0=\mathrm{Im}\varphi,V_i=\mathrm{Im}\varphi_i$, 则由 $\varphi=\varphi_1+\varphi_2+\cdots+\varphi_m$ 可得 $V_0\subseteq V_1+V_2+\cdots+V_m$.

先证充分性. 由 $\varphi_i^2=\varphi_i,\varphi_i\varphi_j=0\ (i\neq j)$ 可得 $\varphi_i=(\varphi_1+\varphi_2+\cdots+\varphi_m)\varphi_i=\varphi\varphi_i$, 故 $V_i\subseteq V_0$, 从而 $V_0=V_1+V_2+\cdots+V_m$. 要证上述和为直和, 只要证明零向量表示唯一即可. 设

$$\mathbf{0}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_i=\varphi_i(\boldsymbol{v}_i)\in V_i(1\leq i\leq m),$$

则 $\mathbf{0}=\varphi_i(\varphi_1(\boldsymbol{v}_1))+\varphi_i(\varphi_2(\boldsymbol{v}_2))+\cdots+\varphi_i(\varphi_m(\boldsymbol{v}_m))=\varphi_i^2(\boldsymbol{v}_i)=\varphi_i(\boldsymbol{v}_i)=\boldsymbol{\alpha}_i$. 因此 $V_0=V_1\oplus V_2\oplus\cdots\oplus V_m$. 两边同取维数即得 $\mathrm{r}(\varphi)=\mathrm{r}(\varphi_1)+\mathrm{r}(\varphi_2)+\cdots+\mathrm{r}(\varphi_m)$.

再证必要性. 由于 $V_0 \subseteq V_1 + V_2 + \cdots + V_m$, 于是

$$\dim V_0 \leq \dim(V_1 + V_2 + \cdots + V_m) \leq \dim V_1 + \dim V_2 + \cdots + \dim V_m,$$

故由 $\mathrm{r}(\varphi) = \mathrm{r}(\varphi_1) + \mathrm{r}(\varphi_2) + \cdots + \mathrm{r}(\varphi_m)$ 可得 $\dim V_0 = \dim V_1 + \dim V_2 + \cdots + \dim V_m$, 从而上式中的不等号只能取等号. 由命题**??**及直和的充要条件可知,$V_1 + V_2 + \cdots + V_m$ 是直和, 并且

$$V_0 = V_1 \oplus V_2 \oplus \cdots \oplus V_m.$$

因为 $\mathrm{Im}\varphi_i = V_i \subseteq V_0 = \mathrm{Im}\varphi$, 故对 V 中任一向量 $\boldsymbol{\alpha}$, 存在 $\boldsymbol{\beta} \in V$, 使得 $\varphi_i(\boldsymbol{\alpha}) = \varphi(\boldsymbol{\beta})$, 从而

$$\begin{aligned}
\varphi_i(\boldsymbol{\alpha}) = \varphi(\boldsymbol{\beta}) = \varphi^2(\boldsymbol{\beta}) &= (\varphi_1 + \varphi_2 + \cdots + \varphi_m)\varphi(\boldsymbol{\beta}) \\
&= (\varphi_1 + \varphi_2 + \cdots + \varphi_m)\varphi_i(\boldsymbol{\alpha}) \\
&= \varphi_1\varphi_i(\boldsymbol{\alpha}) + \varphi_2\varphi_i(\boldsymbol{\alpha}) + \cdots + \varphi_m\varphi_i(\boldsymbol{\alpha}).
\end{aligned}$$

由直和表示的唯一性可知

$$\varphi_i^2(\boldsymbol{\alpha}) = \varphi_i(\boldsymbol{\alpha}), \varphi_j\varphi_i(\boldsymbol{\alpha}) = 0\ (j \neq i),$$

于是 $\varphi_i^2 = \varphi_i, \varphi_i\varphi_j = 0\ (i \neq j)$.

证法二 (代数方法):把问题转换成代数的语言: 设 $\boldsymbol{A}, \boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m$ 是 n 阶矩阵, 满足 $\boldsymbol{A}^2 = \boldsymbol{A}$ 且 $\boldsymbol{A} = \boldsymbol{A}_1 + \boldsymbol{A}_2 + \cdots + \boldsymbol{A}_m$, 求证:$\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\boldsymbol{A}_1) + \mathrm{r}(\boldsymbol{A}_2) + \cdots + \mathrm{r}(\boldsymbol{A}_m)$ 成立的充要条件是 $\boldsymbol{A}_i^2 = \boldsymbol{A}_i, \boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{O}\ (i \neq j)$.

先证充分性. 若 $\boldsymbol{A}_i^2 = \boldsymbol{A}_i$, 则由命题 4.34可知 $\mathrm{r}(\boldsymbol{A}_i) = \mathrm{tr}(\boldsymbol{A}_i)$, 从而

$$\begin{aligned}
\mathrm{r}(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A}) &= \mathrm{tr}(\boldsymbol{A}_1 + \boldsymbol{A}_2 + \cdots + \boldsymbol{A}_m) \\
&= \mathrm{tr}(\boldsymbol{A}_1) + \mathrm{tr}(\boldsymbol{A}_2) + \cdots + \mathrm{tr}(\boldsymbol{A}_m) = \mathrm{r}(\boldsymbol{A}_1) + \mathrm{r}(\boldsymbol{A}_2) + \cdots + \mathrm{r}(\boldsymbol{A}_m).
\end{aligned}$$

再证必要性. 因为 $\boldsymbol{A}$ 是幂等矩阵, 故由命题 3.32可得 $n = \mathrm{r}(\boldsymbol{I}_n - \boldsymbol{A}) + \mathrm{r}(\boldsymbol{A})$, 从而 $n = \mathrm{r}(\boldsymbol{I}_n - \boldsymbol{A}) + \mathrm{r}(\boldsymbol{A}_1) + \mathrm{r}(\boldsymbol{A}_2) + \cdots + \mathrm{r}(\boldsymbol{A}_m)$. 构造如下分块对角矩阵并对其实施分块初等变换, 可得

$$\begin{pmatrix}
\boldsymbol{I}_n - \boldsymbol{A} & & & & \\
 & \boldsymbol{A}_1 & & & \\
 & & \boldsymbol{A}_2 & & \\
 & & & \ddots & \\
 & & & & \boldsymbol{A}_m
\end{pmatrix} \to \begin{pmatrix}
\boldsymbol{I}_n & & & & \\
\boldsymbol{A}_1 & \boldsymbol{A}_1 & & & \\
\boldsymbol{A}_2 & & \boldsymbol{A}_2 & & \\
\vdots & & & \ddots & \\
\boldsymbol{A}_m & & & & \boldsymbol{A}_m
\end{pmatrix} \to$$

$$\begin{pmatrix}
\boldsymbol{I}_n & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_m \\
\boldsymbol{A}_1 & \boldsymbol{A}_1 & & & \\
\boldsymbol{A}_2 & & \boldsymbol{A}_2 & & \\
\vdots & & & \ddots & \\
\boldsymbol{A}_m & & & & \boldsymbol{A}_m
\end{pmatrix} \to \begin{pmatrix}
\boldsymbol{I}_n & \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{O} \\
\boldsymbol{O} & \boldsymbol{A}_1 - \boldsymbol{A}_1^2 & -\boldsymbol{A}_1\boldsymbol{A}_2 & \cdots & -\boldsymbol{A}_1\boldsymbol{A}_m \\
\boldsymbol{O} & -\boldsymbol{A}_2\boldsymbol{A}_1 & \boldsymbol{A}_2 - \boldsymbol{A}_2^2 & \cdots & -\boldsymbol{A}_2\boldsymbol{A}_m \\
\vdots & \vdots & \vdots & & \vdots \\
\boldsymbol{O} & -\boldsymbol{A}_m\boldsymbol{A}_1 & -\boldsymbol{A}_m\boldsymbol{A}_2 & \cdots & \boldsymbol{A}_m - \boldsymbol{A}_m^2
\end{pmatrix}.$$

由 $n = \mathrm{r}(\boldsymbol{I}_n - \boldsymbol{A}) + \mathrm{r}(\boldsymbol{A}_1) + \mathrm{r}(\boldsymbol{A}_2) + \cdots + \mathrm{r}(\boldsymbol{A}_m)$ 可得最后一个矩阵的右下角部分必为零矩阵, 从而 $\boldsymbol{A}_i^2 = \boldsymbol{A}_i, \boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{O}\ (i \neq j)$. □

推论 4.7

若取 $\boldsymbol{I}_V$ 为 n 维线性空间 V 上的恒等变换, 并且此时线性变换 φ_i 满足 $\varphi_1 + \varphi_2 + \cdots + \varphi_m = \boldsymbol{I}_n$. 如果下列条件之一成立:

(1) $\dim V = \dim \mathrm{Im}\varphi_1 + \dim \mathrm{Im}\varphi_2 + \cdots + \dim \mathrm{Im}\varphi_m$;

(2) $\varphi_i^2 = \varphi_i, \varphi_i\varphi_j = 0\ (i \neq j)$,

则 $V = \mathrm{Im}\varphi_1 \oplus \mathrm{Im}\varphi_2 \oplus \cdots \oplus \mathrm{Im}\varphi_m$, 并且 φ_i 就是 V 到 $\mathrm{Im}\varphi_i$ 上的投影变换.

证明 由命题 4.38可知条件 (1)(2) 等价, 并且由命题 4.38 证法一的必要性证明过程可直接由条件 (1) 推出 $V = \mathrm{Im}\varphi_1 \oplus \mathrm{Im}\varphi_2 \oplus \cdots \oplus \mathrm{Im}\varphi_m$(直和的证明也可由条件 (2) 及投影变换的性质直接得到). 又因为幂等变换和投影变换

等价, 故由条件 (2) 可直接得到 φ_i 就是 V 到 $\mathrm{Im}\varphi_i$ 上的投影变换. 因此结论得证. □

第五章 多项式

5.1 基本定理和命题

命题 5.1 (多项式次数的性质)

设 $f(x), g(x) \in \mathbb{K}[x]$. 则

1. $\deg(cf(x)) = \deg f(x), 0 \neq 0 \in \mathbb{K}$.
2. $\deg(f(x) + g(x)) \leqslant \max\{\deg f(x), \deg g(x)\}$.
3. $\deg(f(x)g(x)) = \deg f(x) + \deg g(x)$.

5.2 整除与带余除法

定义 5.1 (整除的定义)

设 $f(x), g(x)$ 是 $\mathbb{F}$ 上的多项式, 若存在 $\mathbb{F}$ 上的多项式 $h(x)$, 使得

$$f(x) = g(x)h(x),$$

则称 $g(x)$ 是 $f(x)$ 的因式, 或称 $g(x)$ 可整除 $f(x)$(也称 $f(x)$ 可被 $g(x)$ 整除), 记为 $g(x) \mid f(x)$.

命题 5.2 (整除的基本性质)

设 $f(x), g(x), h(x) \in \mathbb{K}[x], 0 \neq c \in \mathbb{K}$, 则

(1) 若 $f(x) \mid g(x)$, 则 $cf(x) \mid g(x)$, 因此非零常数多项式 c 是任一非零多项式的因式;

(2) $f(x) \mid f(x)$;

(3) 若 $f(x) \mid g(x), g(x) \mid h(x)$, 则 $f(x) \mid h(x)$;

(4) 若 $f(x) \mid g(x), f(x) \mid h(x)$, 则对任意的多项式 $u(x), v(x)$, 有

$$f(x) \mid g(x)u(x) + h(x)v(x);$$

(5) 设 $f(x) \mid g(x), g(x) \mid f(x)$ 且 $f(x), g(x)$ 都是非零多项式, 则存在 $\mathbb{K}$ 中非零元 c, 使

$$f(x) = cg(x).$$

(6) 若 $g_1(x) \mid f(x), g_2(x) \mid f(x)$, 则 $g_1(x)g_2(x) \mid f^2(x)$.

证明

(1) 若 $g(x) = f(x)p(x)$, 则

$$g(x) = (cf(x))(c^{-1}p(x)).$$

此即 $cf(x) \mid g(x)$.

特别地, 任取 $a \in \mathbb{K}$, 令 $g(x) = a$, 则 $a \mid a$, 从而 $ca \mid a$, 故 c 是 a 的因式.

(2) 显然.

(3) 若 $g(x) = f(x)p(x), h(x) = g(x)q(x)$, 则

$$h(x) = (f(x)p(x))q(x) = f(x)(p(x)q(x)).$$

(4) 若 $g(x) = f(x)p(x), h(x) = f(x)q(x)$, 则

$$g(x)u(x) + h(x)v(x) = f(x)(p(x)u(x) + q(x)v(x)).$$

(5) 设 $g(x) = f(x)p(x), f(x) = g(x)q(x)$, 则

$$f(x) = f(x)(p(x)q(x)).$$

由此即得

$$\deg f(x) = \deg f(x) + \deg(p(x)q(x)),$$

从而

$$\deg(p(x)q(x)) = 0,$$

于是

$$\deg p(x) = \deg q(x) = 0.$$

因此 $p(x)$ 及 $q(x)$ 均为非零常数多项式, 即 $f(x)$ 和 $g(x)$ 相差一个非零常数倍.

(6) 由 $g_1(x), g_2(x) \mid f(x)$ 可知, 存在多项式 $h_1(x), h_2(x)$, 使得

$$f(x) = g_1(x)h_1(x) = g_2(x)h_2(x).$$

从而 $f^2(x) = g_1(x)g_2(x)h_1(x)h_2(x)$, 故 $g_1(x)g_2(x) \mid f^2(x)$.

□

定义 5.2 (相伴多项式)

若 $f(x) \mid g(x), g(x) \mid f(x)$ 且 $f(x), g(x)$ 都是非零多项式, 则 $f(x), g(x)$(即可以互相整除的两个多项式) 称为**相伴多项式**, 记为 $f(x) \sim g(x)$.

♣

笔记 由整除的基本性质 (5)可知, 相伴的多项式只相差一个非零常数倍.

命题 5.3 (相伴多项式的基本性质)

若 $f(x) \sim g(x)$, 则任意的多项式 $u(x)$ 都有 $f(x)u(x) \sim g(x)u(x)$.

♠

证明 由 $f(x) \sim g(x)$ 及整除的基本性质 (4)可知, 任意的多项式 $u(x)$ 都有 $f(x)u(x) \mid g(x)u(x)$,$g(x)u(x) \mid f(x)u(x)$. 故 $f(x)u(x) \sim g(x)u(x)$. □

定理 5.1 (多项式的带余除法)

设 $f(x), g(x) \in \mathbb{F}[x], g(x) \neq 0$, 则必存在唯一的 $q(x), r(x) \in \mathbb{F}[x]$, 使得

$$f(x) = g(x)q(x) + r(x),$$

且 $\deg\ r(x) < \deg\ g(x)$.

♡

证明 若 $\deg f(x) < \deg g(x)$, 只需令 $q(x) = 0, r(x) = f(x)$ 即可. 现设 $\deg f(x) \geq \deg g(x)$, 对 $f(x)$ 的次数用数学归纳法. 若 $\deg f(x) = 0$, 则 $\deg g(x) = 0$. 因此可设 $f(x) = a, g(x) = b(a \neq 0, b \neq 0)$. 这时令 $q(x) = ab^{-1}, r(x) = 0$ 即可. 作为归纳假设, 我们设结论对小于 n 次的多项式均成立. 设

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, a_n \neq 0,$$
$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0, b_m \neq 0,$$

由于 $n \geq m$, 可令

$$f_1(x) = f(x) - a_n b_m^{-1} x^{n-m} g(x),$$

则 $\deg f_1(x) < n$. 由归纳假设, 有

$$f_1(x) = g(x)q_1(x) + r(x),$$

且 $\deg r(x) < \deg g(x)$, 于是

$$f(x) - a_n b_m^{-1} x^{n-m} g(x) = g(x)q_1(x) + r(x).$$

因此

$$f(x) = g(x)(a_n b_m^{-1} x^{n-m} + q_1(x)) + r(x).$$

令

$$q(x) = a_n b_m^{-1} x^{n-m} + q_1(x),$$

即得 $f(x) = g(x)q(x) + r(x)$.

再证明唯一性. 设另有 $p(x), t(x)$, 使

$$f(x) = g(x)p(x) + t(x),$$

且 $\deg t(x) < \deg g(x)$, 则

$$g(x)(q(x) - p(x)) = t(x) - r(x).$$

注意上式左边若 $q(x) - p(x) \neq 0$, 便有

$$\deg g(x)(q(x) - p(x)) \geq \deg g(x) > \deg (t(x) - r(x)),$$

引出矛盾. 因此只可能 $p(x) = q(x), t(x) = r(x)$. □

推论 5.1

设 $f(x), g(x) \in \mathbb{F}[x], g(x) \neq 0$, 必存在唯一的 $q(x), r(x) \in \mathbb{F}[x]$, 使得 $f(x) = g(x)q(x) + r(x)$. 则 $g(x) \mid f(x)$ 的充要条件是 $r(x) = 0$.

例题 5.1 设 $g(x) = ax + b \in \mathbb{F}[x]$ 且 $a \neq 0$, 又 $f(x) \in \mathbb{F}[x]$, 求证:$g(x) \mid f(x)^2$ 的充要条件是 $g(x) \mid f(x)$.

证明 充分性显然, 只需证明必要性.

证法一: 设 $f(x) = g(x)q(x) + r$, 则

$$f(x)^2 = g(x)^2 q(x)^2 + 2rg(x)q(x) + r^2.$$

由 $g(x) \mid f(x)^2$ 可得 $g(x) \mid r^2$, 故 $r^2 = 0$, 即 $r = 0$, 从而 $g(x) \mid f(x)$.

证法二: 由余数定理,$f\left(-\dfrac{b}{a}\right)^2 = 0$, 故 $f\left(-\dfrac{b}{a}\right) = 0$, 从而 $g(x) \mid f(x)$. □

例题 5.2 设 $g(x) = ax^2 + bx + c(abc \neq 0)$, $f(x) = x^3 + px^2 + qx + r$, 满足 $g(x) \mid f(x)$, 求证:

$$\frac{ap - b}{a} = \frac{aq - c}{b} = \frac{ar}{c}.$$

证明 用待定系数法, 设

$$x^3 + px^2 + qx + r = (ax^2 + bx + c)(mx + n) = amx^3 + (an + bm)x^2 + (bn + cm)x + cn.$$

比较系数得

$$am = 1,\ an + bm = p,\ bn + cm = q,\ cn = r.$$

由此即可得到所需等式. □

5.2.1 凑项法

“凑项法”是指在要证明的等式中添加若干项再减去若干项来证明结论的方法.

命题 5.4

$(x^d - a^d) \mid (x^n - a^n)$ 的充要条件是 $d \mid n$, 其中 $a \neq 0$.

证明 $(\Leftarrow)$: 由 $d|n$ 可设 $n = kd, k \in \mathbb{N}_+$. 从而

$$x^n - a^n = (x^d)^k - (a^d)^k = (x^d - a^d)(x^{d(k-1)} + x^{d(k-2)}a^d + \cdots + a^{d(k-1)}).$$

故 $(x^d - a^d)|(x^n - a^n)$.

(⇒): 假设 $d \nmid n$, 则由带余除法可知, 存在 $q, r \in \mathbb{N}_+$ 且 $0 \leqslant r < d$, 使得 $n = qd + r$. 于是

$$x^n - a^n = x^{dq+r} - a^{dq+r} = (x^{dq} - a^{dq})x^r + x^r a^{dq} - a^{dq+r} = (x^{dq} - a^{dq})x^r + a^{dq}(x^r - a^r).$$

注意到 $(x^{dq} - a^{dq}) | (x^d - a^d)$, 但由 $0 \leqslant r < d$ 可知,$(x^d - a^d) \nmid (x^r - a^r)$. 故 $(x^d - a^d) \nmid (x^n - a^n)$ 矛盾! □

例题 5.3 设 $f(x) = x^{3m} + x^{3n+1} + x^{3p+2}$, 其中 m, n, p 为自然数, 又 $g(x) = x^2 + x + 1$, 求证: $g(x) \mid f(x)$.

证明 由命题 5.4可知,$(x^3 - 1) | (x^{3k} - 1), \forall k \in \mathbb{N}_+$. 又因为 $(x^2 + x + 1) | (x^3 - 1)$, 所以 $(x^2 + x + 1) | (x^{3k} - 1), \forall k \in \mathbb{N}_+$. 注意到

$$x^{3m} + x^{3n+1} + x^{3p+2} = (x^{3m} - 1) + x(x^{3n} - 1) + x^2(x^{3p} - 1) + (x^2 + x + 1).$$

再结合 $(x^2 + x + 1) | (x^{3m} - 1), (x^{3n} - 1), (x^{3p} - 1)$ 可得 $g(x) | f(x)$. □

5.3 最大公因式与互素多项式

定义 5.3 (最大公因式和互素)

设 $f(x), g(x)$ 是 $\mathbb{F}$ 上的多项式,$d(x)$ 是 $\mathbb{F}$ 上的首 1 多项式, 若 $d(x)$ 满足
(i) $d(x)$ 是 $f(x), g(x)$ 的公因式,
(ii) 对 $f(x), g(x)$ 的任一公因式 $h(x)$, 都有 $h(x) \mid d(x)$,
则称 $d(x)$ 为 $f(x), g(x)$ 的**最大公因式**(或称 $d(x)$ 为 $f(x), g(x)$ 的 g.c.d.), 记为 $d(x) = (f(x), g(x))$. 特别地, 若 $d(x) = 1$, 则称 $f(x), g(x)$ 互素.

♣

命题 5.5

(1) 若 $\mathbb{F}$ 上的多项式 $d_0(x)$(但不一定是首 1 多项式) 也满足
 (i) $d_0(x)$ 是 $f(x), g(x)$ 的公因式,
 (ii) 对 $f(x), g(x)$ 的任一公因式 $h(x)$, 都有 $h(x) \mid d_0(x)$,
 则 $d_0(x) \sim d(x)$, 即 $d_0(x)$ 与 $d(x)$ 相差一个非零常数倍.
(2) 对 $\forall a, b \in \mathbb{F}, (af(x), bg(x)) = (f(x), g(x)) = d(x)$.

♠

证明

(1) 证明是显然的.
(2) 证明是显然的.

□

定义 5.4 (最小公倍式)

设 $f(x), g(x)$ 是 $\mathbb{F}$ 上的多项式,$m(x)$ 是 $\mathbb{F}$ 上的首 1 多项式, 若 $d(x)$ 满足
(i) $m(x)$ 是 $f(x)$ 与 $g(x)$ 的公倍式,
(ii) 对 $f(x)$ 与 $g(x)$ 的任一公倍式 $l(x)$ 均有 $m(x) \mid l(x)$,
则称 $m(x)$ 为 $f(x)$ 与 $g(x)$ 的**最小公倍式**(或称 $m(x)$ 为 $f(x), g(x)$ 的 l.c.m.), 记为 $m(x) = [f(x), g(x)]$.

♣

命题 5.6

(1) 若 $\mathbb{F}$ 上的多项式 $m_0(x)$(但不一定是首 1 多项式) 也满足
 (i) $m(x)$ 是 $f(x)$ 与 $g(x)$ 的公倍式,
 (ii) 对 $f(x)$ 与 $g(x)$ 的任一公倍式 $l(x)$ 均有 $m(x) \mid l(x)$,
 则 $m_0(x) \sim m(x)$, 即 $m_0(x)$ 与 $m(x)$ 相差一个非零常数倍.
(2) 对 $\forall a, b \in \mathbb{F}, [af(x), bg(x)] = [f(x), g(x)] = m(x)$.

♠

证明

(1) 证明是显然的.
(2) 证明是显然的.

□

定理 5.2 (最大公因式的必要条件)

设 $f(x), g(x)$ 是 $\mathbb{F}$ 上的多项式,$d(x)$ 是它们的最大公因式, 则必存在 $\mathbb{F}$ 上的多项式 $u(x), v(x)$, 使得

$$f(x)u(x) + g(x)v(x) = d(x).$$

注 设 $d(x) = f(x)u(x) + g(x)v(x)$, 则 $d(x)$ 不一定是 $f(x)$ 和 $g(x)$ 的最大公因式.

证明 若 $f(x) = 0$, 则显然 $(f(x), g(x)) = g(x)$; 若 $g(x) = 0$, 则 $(f(x), g(x)) = f(x)$. 故不妨设 $f(x) \neq 0, g(x) \neq 0$. 由带余除法, 我们有下列等式:

$$f(x) = g(x)q_1(x) + r_1(x),$$
$$g(x) = r_1(x)q_2(x) + r_2(x),$$
$$r_1(x) = r_2(x)q_3(x) + r_3(x),$$
$$\cdots\cdots\cdots$$
$$r_{s-2}(x) = r_{s-1}(x)q_s(x) + r_s(x),$$
$$\cdots\cdots\cdots$$

余式的次数是严格递减的, 因此经过有限步后, 必有一个等式其余式为零. 不妨设 $r_{s+1}(x) = 0$, 于是

$$r_{s-1}(x) = r_s(x)q_{s+1}(x). \tag{5.1}$$

现在要证明 $r_s(x)$ 即为 $f(x)$ 与 $g(x)$ 的最大公因式. 由上式知 $r_s(x) \mid r_{s-1}(x)$, 但

$$r_{s-2}(x) = r_{s-1}(x)q_s(x) + r_s(x), \tag{5.2}$$

因此 $r_s(x) \mid r_{s-2}(x)$. 这样可一直推下去, 得到 $r_s(x) \mid g(x), r_s(x) \mid f(x)$. 这表明 $r_s(x)$ 是 $f(x)$ 与 $g(x)$ 的公因式. 又设 $h(x)$ 是 $f(x)$ 与 $g(x)$ 的公因式, 则 $h(x) \mid r_1(x)$, 于是 $h(x) \mid r_2(x)$, 不断往下推, 容易看出有 $h(x) \mid r_s(x)$. 因此 $r_s(x)$ 是最大公因式.

再证明(5.1)式. 从(5.2)式得

$$r_s(x) = r_{s-2}(x) - r_{s-1}(x)q_s(x), \tag{5.3}$$

但我们有

$$r_{s-3}(x) = r_{s-2}(x)q_{s-1}(x) + r_{s-1}(x),) \tag{5.4}$$

从(5.4)式中解出 $r_{s-1}(x)$ 代入(5.3)式, 得

$$r_s(x) = r_{s-2}(x)(1 + q_{s-1}(x)q_s(x)) - r_{s-3}(x)q_s(x).$$

用类似的方法逐步将 $r_i(x)$ 用 $r_{i-1}(x), r_{i-2}(x)$ 代入, 最后得到

$$r_s(x) = f(x)u(x) + g(x)v(x).$$

显然 $u(x), v(x) \in \mathbb{K}[x]$. □

定理 5.3 (最大公因式的充分条件)

设 $f(x), g(x), d(x)$ 是 $\mathbb{F}$ 上的多项式. 若 $d(x) \mid f(x), d(x) \mid g(x)$ 并且存在 $\mathbb{F}$ 上的多项式 $u(x), v(X)$, 使得 $d(x) = f(x)u(x) + g(x)v(x)$, 则 $d(x)$ 必是 $f(x)$ 和 $g(x)$ 的最大公因式.

证明 如果同时 $d(x) \mid f(x), d(x) \mid g(x)$, 则 $d(x)$ 是 $f(x)$ 和 $g(x)$ 的公因式. 若 $h(x)$ 也是 $f(x), g(x)$ 的公因式, 则由

$h(x) \mid f(x), h(x) \mid g(x)$ 可推出 $h(x) \mid (f(x)u(x) + g(x)v(x)) = d(x)$, 因此 $d(x)$ 是最大公因式. □

例题 5.4 设 $(f(x), g(x)) = d(x)$, 求证: 对任意的正整数 n,

$$(f(x)^n, f(x)^{n-1}g(x), \cdots, g(x)^n) = d(x)^n.$$

证明 显然 $d(x)^n$ 是 $f(x)^{n-k}g(x)^k(0 \le k \le n)$ 的公因式. 又假设

$$f(x)u(x) + g(x)v(x) = d(x),$$

两边同时 n 次方得到

$$f^n(x)u^n(x) + f^{n-1}(x)g(x)u^{n-1}(x)v(x) + \cdots + g^n(x)v^n(x) = d^n(x).$$

于是由最大公因式的充分条件可知 $d(x)^n$ 是 $f(x)^{n-k}g(x)^k(0 \le k \le n)$ 的最大公因式. □

推论 5.2 (次数不小于 1 的多项式互素的充要条件)

设 $f(x), g(x)$ 是次数不小于 1 的多项式互素的充要条件是必唯一地存在两个多项式 $u(x), v(x)$, 使得

$$f(x)u(x) + g(x)v(x) = 1,$$

且 $\deg u(x) < \deg g(x), \deg v(x) < \deg f(x)$.

证明 充分性由多项式互素的充要条件可直接得到. 下面证明必要性.

先证存在性. 因为 $(f(x), g(x)) = 1$ 且 $\deg f(x), \deg g(x) > 1$, 所以由多项式互素的充要条件可知, 必存在非零多项式 $h(x), k(x)$, 使得

$$f(x)h(x) + g(x)k(x) = 1. \tag{5.5}$$

由带余除法可知, 存在 $q(x), u(x)$, 使得

$$h(x) = g(x)q(x) + u(x), \quad \deg u(x) < \deg g(x).$$

代入(5.5)式可得

$$f(x)[g(x)q(x) + u(x)] + g(x)k(x) = 1.$$

即有

$$f(x)u(x) + g(x)[f(x)q(x) + k(x)] = 1. \tag{5.6}$$

令 $v(x) = f(x)q(x) + k(x)$, 则 $\deg v(x) < \deg f(x)$. 否则, 若 $\deg v(x) \geqslant \deg f(x)$, 则由(5.6)式可知

$$f(x)u(x) + g(x)v(x) = 1. \tag{5.7}$$

从而由 $\deg v(x) \geqslant \deg f(x)$ 及 $\deg u(x) < \deg g(x)$ 可得

$$\deg(f(x)u(x)) = \deg f(x) + \deg u(x) < \deg v(x) + \deg g(x) = \deg(g(x)v(x)).$$

而由(5.7)式可知 $\deg(f(x)u(x)) = \deg(1 - g(x)v(x)) = \deg(g(x)v(x))$ 矛盾!

再证唯一性, 设另有 $u_1(x), v_1(x)$ 适合条件, 即

$$f(x)u_1(x) + g(x)v_1(x) = f(x)u(x) + g(x)v(x) = 1.$$

从而

$$f(x)(u(x) - u_1(x)) = g(x)(v(x) - v_1(x)).$$

上式表明 $g(x) \mid f(x)(u(x) - u_1(x))$, 又由于 $(f(x), g(x)) = 1$, 因此 $g(x) \mid (u(x) - u_1(x))$. 而 $\deg(u(x) - u_1(x)) < \deg g(x)$, 故 $u(x) - u_1(x) = 0$, 即 $u(x) = u_1(x)$. 同理可得 $v(x) = v_1(x)$. □

定理 5.4 (多项式互素的充要条件)

设 $f(x), g(x)$ 是 $\mathbb{F}$ 上的多项式. 则

(1) $(f(x), g(x)) = 1$ 的充要条件是存在 $\mathbb{F}$ 上的多项式 $u(x), v(x)$, 使得 $f(x)u(x) + g(x)v(x) = 1$.

(2) $(f(x), g(x)) = 1$ 的充要条件是对任意给定的正整数 m, n, $(f(x)^m, g(x)^n) = 1$.

证明

1. 必要性: 由最大公因式的必要条件立得.

   充分性: 设 $(f(x), g(x)) = d(x)$, 则由 $f(x)u(x) + g(x)v(x) = 1$ 可知,$d(x) \mid 1$, 因此 $d(x) = 1$.
2. 必要性由命题 5.8即得. 反过来, 若 $d(x) \neq 1$ 是 $f(x)$ 和 $g(x)$ 的公因式, 则它也是 $f(x)^m$ 和 $g(x)^n$ 的公因式, 因此 $f(x)^m$ 和 $g(x)^n$ 不可能互素.

□

命题 5.7 (互素多项式和最大公因式的基本性质)

设 $f(x), g(x), f_1(x), f_2(x) \in \mathbb{K}[x]$, 则

(1) 若 $f_1(x) \mid g(x), f_2(x) \mid g(x)$, 且 $(f_1(x), f_2(x)) = 1$, 则 $f_1(x)f_2(x) \mid g(x)$.

(2) 若 $(f(x), g(x)) = 1$, 且 $f(x) \mid g(x)h(x)$, 则 $f(x) \mid h(x)$.

(3) 若 $(f(x), g(x)) = d(x), f(x) = f_1(x)d(x), g(x) = g_1(x)d(x)$, 则 $(f_1(x), g_1(x)) = 1$.

(4) 若 $(f(x), g(x)) = d(x)$, 则 $(t(x)f(x), t(x)g(x)) = t(x)d(x)$.

(5) 若 $(f_1(x), g(x)) = 1, (f_2(x), g(x)) = 1$, 则 $(f_1(x)f_2(x), g(x)) = 1$.

(6) 若 $(f(x), g(x)) = 1$, 则 $(f(x^m), g(x^m)) = 1$, 其中 m 为任一正整数.

(7) 若 $(f(x), g(x)) = 1$, 则 $(f(x)g(x), f(x) + g(x)) = 1$.

证明

(1) 由多项式互素的充要条件 (1)可知, 存在 $u(x), v(x) \in \mathbb{K}[x]$, 使

$$f_1(x)u(x) + f_2(x)v(x) = 1.$$

设 $g(x) = f_1(x)s(x) = f_2(x)t(x)$, 则

$$\begin{aligned} g(x) &= g(x)(f_1(x)u(x) + f_2(x)v(x)) \\ &= f_2(x)t(x)f_1(x)u(x) + f_1(x)s(x)f_2(x)v(x) \\ &= f_1(x)f_2(x)(t(x)u(x) + s(x)v(x)), \end{aligned}$$

即 $f_1(x)f_2(x) \mid g(x)$.

(2) 由多项式互素的充要条件可知, 存在 $u(x), v(x) \in \mathbb{K}[x]$, 使

$$f(x)u(x) + g(x)v(x) = 1,$$

则

$$f(x)u(x)h(x) + g(x)v(x)h(x) = h(x).$$

因上式左边可被 $f(x)$ 整除, 故 $f(x) \mid h(x)$.

(3) 由多项式互素的充要条件可知, 存在 $u(x), v(x) \in \mathbb{K}[x]$, 使

$$f(x)u(x) + g(x)v(x) = d(x),$$

即

$$f_1(x)d(x)u(x) + g_1(x)d(x)v(x) = d(x),$$

两边消去 $d(x)$ 即得

$$f_1(x)u(x) + g_1(x)v(x) = 1,$$

因此 $f_1(x), g_1(x)$ 互素.

(4) $u(x), v(x) \in \mathbb{K}[x]$, 使

$$f(x)u(x) + g(x)v(x) = d(x),$$

则
$$t(x)f(x)u(x)+t(x)g(x)v(x)=t(x)d(x).$$
因此, 若 $h(x)\mid t(x)f(x), h(x)\mid t(x)g(x)$, 则必有 $h(x)\mid t(x)d(x)$. 又 $t(x)d(x)$ 是 $t(x)f(x), t(x)g(x)$ 的公因式, 因此 $t(x)d(x)$ 是 $t(x)f(x)$ 与 $t(x)g(x)$ 的最大公因式.

(5) 由多项式互素的充要条件可知, 存在 $u_1(x), v_1(x), u_2(x), v_2(x)\in\mathbb{K}[x]$, 使
$$f_1(x)u_1(x)+g(x)v_1(x)=1,$$
$$f_2(x)u_2(x)+g(x)v_2(x)=1,$$
将上两式两边分别相乘得
$$(f_1(x)f_2(x))u_1(x)u_2(x)+g(x)(v_1(x)f_2(x)u_2(x)+g(x)v_1(x)v_2(x)+v_2(x)f_1(x)u_1(x))=1.$$
这就是说 $f_1(x)f_2(x)$ 和 $g(x)$ 互素.

(6) 因为 $f(x)$ 和 $g(x)$ 互素, 故存在多项式 $u(x), v(x)$, 使得
$$f(x)u(x)+g(x)v(x)=1,$$
从而有
$$f(x^m)u(x^m)+g(x^m)v(x^m)=1,$$
于是 $f(x^m)$ 和 $g(x^m)$ 互素.

(7) 由互素多项式的充要条件 (1)可知, 存在 $u(x), v(x)$, 使得
$$f(x)u(x)+g(x)v(x)=1.$$
从而
$$f(x)[u(x)-v(x)]+[f(x)+g(x)]v(x)=1.$$
故由互素多项式的充要条件 (1)可知,$(f(x), f(x)+g(x))=1$. 同理可得,$(f(x), f(x)+g(x))=1$. 再由互素多项式和最大公因式的基本性质 (5)即得 $(f(x)g(x), f(x)+g(x))=1$.

□

定理 5.5 (多个多项式的最大公因式的必要条件)

设 $d(x)$ 是 $f_1(x), f_2(x),\cdots,f_m(x)$ 的最大公因式, 求证: 必存在多项式 $g_1(x), g_2(x),\cdots,g_m(x)$, 使得
$$f_1(x)g_1(x)+f_2(x)g_2(x)+\cdots+f_m(x)g_m(x)=d(x).$$

证明 用数学归纳法. 对 $m=2$, 结论已成立. 设结论对 $m-1$ 成立. 设 $h(x)$ 是 $f_1(x), f_2(x),\cdots,f_{m-1}(x)$ 的最大公因式, 则有 $g_1(x), g_2(x),\cdots,g_{m-1}(x)$, 使得
$$f_1(x)g_1(x)+f_2(x)g_2(x)+\cdots+f_{m-1}(x)g_{m-1}(x)=h(x).$$
结合上式由条件可知 $d(x)$ 是 $h(x)$ 和 $f_m(x)$ 的最大公因式, 故存在 $u(x), v(x)$, 使得
$$h(x)u(x)+f_m(x)v(x)=d(x).$$
将 $h(x)$ 代入可得
$$f_1(x)g_1(x)u(x)+f_2(x)g_2(x)u(x)+\cdots+f_{m-1}(x)g_{m-1}(x)u(x)+f_m(x)v(x)=d(x),$$
即知结论成立.

□

推论 5.3 (多个多项式互素的充要条件)

数域 $\mathbb{F}$ 上的多项式 $f_1(x), f_2(x),\cdots,f_m(x)$ 互素的充要条件是存在 $\mathbb{F}$ 上的多项式 $g_1(x), g_2(x),\cdots,g_m(x)$, 使得
$$f_1(x)g_1(x)+f_2(x)g_2(x)+\cdots+f_m(x)g_m(x)=1.$$

证明 必要性: 由多个多项式的最大公因式的必要条件立即得到.

充分性: 设存在 $\mathbb{F}$ 上的多项式 $g_1(x), g_2(x), \cdots, g_m(x)$, 使得

$$f_1(x)g_1(x) + f_2(x)g_2(x) + \cdots + f_m(x)g_m(x) = 1.$$

设 $d(x)$ 是 $f_1(x), f_2(x), \cdots, f_m(x)$ 的最大公因式, 则由上式可知,$d(x) \mid 1$, 从而 $d(x) = 1$. □

命题 5.8 (两两互素的多项式组的乘积也互素)

设 $f_1(x), \cdots, f_m(x), g_1(x), \cdots, g_n(x)$ 为多项式, 且

$$(f_i(x), g_j(x)) = 1, 1 \le i \le m; 1 \le j \le n,$$

求证:

$$(f_1(x)f_2(x)\cdots f_m(x), g_1(x)g_2(x)\cdots g_n(x)) = 1.$$

证明 利用互素多项式和最大公因式的基本性质 (5)以及数学归纳法即得结论. □

推论 5.4

设 $f(x), g(x)$ 为多项式, 若 $(f(x), g(x)) = 1$, 则 $(f(x), g^n(x)) = 1$.

证明 在命题 5.8(上一个命题)中取 $f_1(x) = f(x), f_i(x) = 1(i = 2, 3\cdots, n), g_j(x) = g(x)(j = 1, 2, \cdots, n)$ 即可得到结论.
□

定理 5.6 (中国剩余定理)

设 $g_1(x),\cdots,g_n(x)$ 是两两互素的多项式,$r_1(x),\cdots,r_n(x)$ 是 n 个多项式, 则存在多项式 $f(x),q_1(x),\cdots,q_n(x)$, 使

$$f(x) = g_i(x)q_i(x) + r_i(x), i = 1, \cdots, n.$$

证明 先证明存在多项式 $f_i(x)$, 使对任意的 i, 有

$$f_i(x) = g_i(x)p_i(x) + 1,\ g_j(x) \mid f_i(x)(j \ne i).$$

一旦得证, 只需令 $f(x) = r_1(x)f_1(x) + \cdots + r_n(x)f_n(x)$ 即可. 现构造 $f_1(x)$ 如下. 因为 $g_1(x)$ 和 $g_j(x)(j \ne 1)$ 互素, 故存在 $u_j(x), v_j(x)$, 使 $g_1(x)u_j(x) + g_j(x)v_j(x) = 1$. 令

$$f_1(x) = g_2(x)v_2(x)\cdots g_n(x)v_n(x) = (1 - g_1(x)u_2(x))\cdots(1 - g_1(x)u_n(x)),$$

显然 $f_1(x)$ 符合要求. 同理可构造 $f_i(x)$. □

命题 5.9 (两个多项式的乘积与其最大公因式和最小公倍式的乘积相伴)

设 $f(x), g(x)$ 是非零多项式, 则

$$f(x)g(x) \sim (f(x), g(x))[f(x), g(x)].$$

证明 证法一: 设 $d(x) = (f(x), g(x))$ 且 $f(x) = f_0(x)d(x), g(x) = g_0(x)d(x)$, 则由互素多项式和最大公因式的基本性质 (3)可知 $f_0(x), g_0(x)$ 互素. 设 $l(x)$ 是 $f(x), g(x)$ 的公倍式且

$$l(x) = f(x)u(x) = g(x)v(x),$$

则 $f_0(x)d(x)u(x) = g_0(x)d(x)v(x)$, 消去 $d(x)$ 得

$$f_0(x)u(x) = g_0(x)v(x).$$

上式表明 $f_0(x) \mid g_0(x)v(x), g_0(x) \mid f_0(x)u(x)$, 又因为 $f_0(x), g_0(x)$ 互素, 所以由互素多项式和最大公因式的基本性质 (2)可知,$f_0(x) \mid v(x), g_0(x) \mid u(x)$. 设 $u(x) = g_0(x)p(x)$, 则

$$l(x) = f_0(x)d(x)g_0(x)p(x),$$

即 $f_0(x)d(x)g_0(x) \mid l(x)$. 显然 $f_0(x)d(x)g_0(x)$ 是 $f(x),g(x)$ 的公倍式, 因此由命题 5.5(1)可知

$$\frac{f(x)g(x)}{d(x)} = f_0(x)d(x)g_0(x) \sim [f(x),g(x)].$$

故由相伴多项式的基本性质可知

$$f(x)g(x) \sim (f(x),g(x))[f(x),g(x)].$$

证法二: 设 $f(x),g(x)$ 的公共标准分解为

$$f(x) = c_1p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_k(x)^{e_k}, \quad g(x) = c_2p_1(x)^{f_1}p_2(x)^{f_2}\cdots p_k(x)^{f_k},$$

其中 $p_i(x)$ 为互不相同的首一不可约多项式,c,d 是非零常数, 则

$$d(x) = p_1(x)^{r_1}p_2(x)^{r_2}\cdots p_k(x)^{r_k}, \quad h(x) = p_1(x)^{s_1}p_2(x)^{s_2}\cdots p_k(x)^{s_k},$$

其中 $r_i = \min\{e_i,f_i\}, s_i = \max\{e_i,f_i\}$. 注意到

$$f(x)g(x) = c_1c_2p_1(x)^{e_1+f_1}p_2(x)^{e_2+f_2}\cdots p_k(x)^{e_k+f_k},$$

并且

$$(f(x),g(x)) = p_1(x)^{r_1}p_2(x)^{r_2}\cdots p_k(x)^{r_k}, \quad [f(x),g(x)] = p_1(x)^{s_1}p_2(x)^{s_2}\cdots p_k(x)^{s_k},$$

其中 $r_i = \min\{e_i,f_i\}, s_i = \max\{e_i,f_i\}$. 令 $c = c_1c_2$, 则有

$$f(x)g(x) = cd(x)h(x).$$

□

命题 5.10 (最大公因式与最小公倍式在开方下不变)

设 $(f(x),g(x)) = d(x), [f(x),g(x)] = h(x)$, 求证:

$$(f(x)^n,g(x)^n) = d(x)^n,\ [f(x)^n,g(x)^n] = h(x)^n.$$

注 不妨设 $f(x),g(x)$ 都是首一多项式的原因: 若 $f(x),g(x)$ 的首项系数分别为 a,b, 则用 $\dfrac{f(x)}{a},\dfrac{g(x)}{b}$ 代替, 再结合命题 5.5(2)和命题 5.6(2)即可得到结论.

证明 证法一: 不妨设 $f(x),g(x)$ 都是首 1 多项式, $f(x) = f_1(x)d(x), g(x) = g_1(x)d(x)$, 则 $f_1(x),g_1(x),d(x)$ 都是首 1 多项式. 由互素多项式和最大公因式的基本性质 (3)可知 $(f_1(x),g_1(x)) = 1$. 由命题 5.9可知 $h(x) \sim f_1(x)g_1(x)d(x)$, 又因为 $h(x),f_1(x),g_1(x),d(x)$ 均为首 1 多项式, 所以 $h(x) = f_1(x)g_1(x)d(x)$. 由命题 5.8可知,$(f_1(x)^n,g_1(x)^n) = 1$, 从而由互素多项式和最大公因式的基本性质 (4)可知

$$(f(x)^n,g(x)^n) = (f_1(x)^nd(x)^n,g_1(x)^nd(x)^n) = d(x)^n.$$

由命题 5.9可知 $f(x)^ng(x)^n \sim (f(x)^n,g(x)^n)[f(x)^n,g(x)^n]$, 又因为 $f(x),g(x)$ 都是首 1 多项式, 所以 $f(x)^ng(x)^n = (f(x)^n,g(x)^n)[f(x)^n,g(x)^n] = d(x)^n[f(x)^n,g(x)^n]$. 于是可得

$$[f(x)^n,g(x)^n] = f_1(x)^ng_1(x)^nd(x)^n = h(x)^n.$$

证法二:设 $f(x),g(x)$ 的公共标准分解为

$$f(x) = cp_1(x)^{e_1}p_2(x)^{e_2}\cdots p_k(x)^{e_k}, \quad g(x) = dp_1(x)^{f_1}p_2(x)^{f_2}\cdots p_k(x)^{f_k},$$

其中 $p_i(x)$ 为互不相同的首一不可约多项式,c,d 是非零常数, 则

$$d(x) = p_1(x)^{r_1}p_2(x)^{r_2}\cdots p_k(x)^{r_k}, \quad h(x) = p_1(x)^{s_1}p_2(x)^{s_2}\cdots p_k(x)^{s_k},$$

其中 $r_i = \min\{e_i,f_i\}, s_i = \max\{e_i,f_i\}$. 注意到

$$f(x)^n = c^np_1(x)^{ne_1}p_2(x)^{ne_2}\cdots p_k(x)^{ne_k}, \quad g(x)^n = d^np_1(x)^{nf_1}p_2(x)^{nf_2}\cdots p_k(x)^{nf_k},$$

并且 $\min\{ne_i,nf_i\} = nr_i, \max\{ne_i,nf_i\} = ns_i$, 因此

$$(f(x)^n,g(x)^n) = p_1(x)^{nr_1}p_2(x)^{nr_2}\cdots p_k(x)^{nr_k} = d(x)^n,$$

$$[f(x)^n, g(x)^n] = p_1(x)^{ns_1} p_2(x)^{ns_2} \cdots p_k(x)^{ns_k} = h(x)^n.$$

□

命题 5.11

设 $f(x) = x^m - 1, g(x) = x^n - 1$, 求证: $(f(x), g(x)) = x^d - 1$, 其中 d 是 m, n 的最大公因子.

证明 证法一: 不妨设 $m \geq n, m = nq + r$, 先证明 $(x^m - 1, x^n - 1) = (x^r - 1, x^n - 1)$. 假设 $d_1(x) = (x^m - 1, x^n - 1), d_2(x) = (x^r - 1, x^n - 1)$. 注意到

$$x^m - 1 = x^{nq+r} - 1 = x^r(x^{nq} - 1) + (x^r - 1),$$

$(x^n - 1) \mid (x^{nq} - 1)$, 故 $d_1(x) \mid (x^r - 1)$, 从而 $d_1(x) \mid d_2(x)$. 从上式也可以看出 $d_2(x) \mid (x^m - 1)$, 从而 $d_2(x) \mid d_1(x)$, 因此 $d_1(x) = d_2(x)$. 又设 $n = q_1 r + r_1$, 则 $(x^m - 1, x^n - 1) = (x^n - 1, x^r - 1) = (x^r - 1, x^{r_1} - 1)$. 再由辗转相除, 有某个 $r_{s-1} = q_{s+1} r_s$, 其中 $r_s = d$ 是 m, n 的最大公因子, 于是 $(x^m - 1, x^n - 1) = (x^{r_{s-1}} - 1, x^{r_s} - 1) = x^d - 1$.

证法二:只需求出 $f(x), g(x)$ 的公根. $f(x)$ 的根为

$$\cos\frac{2k\pi}{m} + \mathrm{i}\sin\frac{2k\pi}{m}, 1 \leq k \leq m,$$

$g(x)$ 的根为

$$\cos\frac{2k\pi}{n} + \mathrm{i}\sin\frac{2k\pi}{n}, 1 \leq k \leq n,$$

则公根为

$$\cos\frac{2k\pi}{d} + \mathrm{i}\sin\frac{2k\pi}{d}, 1 \leq k \leq d.$$

这就是 $x^d - 1$ 的全部根, 于是结论成立. □

5.4 不可约多项式与因式分解

定义 5.5 (不可约多项式的定义)

设 $f(x)$ 是数域 $\mathbb{F}$ 上的多项式, 若 $f(x)$ 可以分解为两个次数小于 $f(x)$ 的 $\mathbb{F}$ 上多项式之积, 则称 $f(x)$ 是 $\mathbb{F}$ 上的可约多项式, 否则称 $f(x)$ 为 $\mathbb{F}$ 上的不可约多项式.

命题 5.12 (不可约多项式的基本性质)

(1) 设 $p(x)$ 是数域 $\mathbb{K}$ 上的不可约多项式, 则对 $\mathbb{K}$ 上任一多项式 $f(x)$, 或者 $p(x) \mid f(x)$, 或者 $(p(x), f(x)) = 1$.

(2) 设 $p(x)$ 是数域 $\mathbb{F}$ 上的不可约多项式, $f(x)$ 是 $\mathbb{F}$ 上的多项式. 证明: 若 $p(x)$ 的某个复根 a 也是 $f(x)$ 的根, 则 $p(x) \mid f(x)$. 特别地, $p(x)$ 的任一复根都是 $f(x)$ 的根.

笔记 不可约多项式的基本性质 (2) 表明: 不可约多项式也满足极小多项式的基本性质.

证明

(1) 设 $d(x) = (p(x), f(x))$. 因为 $p(x)$ 不可约, 故 $f(x)$ 的因式只能是非零常数多项式或 $cp(x)(c \neq 0)$, 从而或者 $d(x) = 1$ 或者 $d(x) = cp(x)$ (首一多项式), 故得结论.

(2) 若 $(p(x), f(x)) = 1$, 则存在 $\mathbb{F}$ 上的多项式 $u(x), v(x)$, 使得 $p(x)u(x) + f(x)v(x) = 1$. 令 $x = a$ 可得 $1 = p(a)u(a) + f(a)v(a) = 0$, 矛盾. 因此 $p(x)$ 与 $f(x)$ 不互素, 从而只能是 $p(x) \mid f(x)$, 结论得证.

□

定理 5.7 (不可约多项式的“素性”)

设 $p(x)$ 是数域 $\mathbb{F}$ 上的非常数多项式, 则 $p(x)$ 为 $\mathbb{F}$ 上不可约多项式的充要条件是对 $\mathbb{F}$ 上任意适合 $p(x) \mid f(x)g(x)$ 的多项式 $f(x)$ 与 $g(x)$, 或者 $p(x) \mid f(x)$, 或者 $p(x) \mid g(x)$.

证明 必要性: 设 $p(x)$ 是 $\mathbb{F}[x]$ 中的不可约多项式, 且 $p(x) \mid f(x)g(x)$. 若 $p(x) \mid f(x)$, 则结论成立. 若 $p(x) \nmid f(x)$, 则由定理可知 $(p(x), f(x)) = 1$, 从而由互素多项式与最大公因式的基本性质可知 $p(x) \mid g(x)$.

充分性:(反证法) 假设 $p(x)$ 可约, 则必存在次数小于 $\deg(p(x))$ 的多项式 $f(x)$,$g(x)$, 使得 $p(x) = f(x)g(x)$. 从而 $p(x) \mid f(x)g(x)$, 于是由条件可知 $p(x) \mid f(x)$ 或 $p(x) \mid g(x)$. 因此 $\deg(p(x)) \leq \deg(f(x))$ 或 $\deg(g(x))$. 这与 $\deg(p(x)) > \deg(f(x))$,$\deg(g(x))$ 矛盾. □

推论 5.5

设 $p(x)$ 为不可约多项式且
$$p(x) \mid f_1(x)f_2(x)\cdots f_m(x),$$
则 $p(x)$ 必可整除其中某个 $f_i(x)$.

证明 由不可约多项式的“素性”归纳可得. □

命题 5.13

设 $f(x)$ 是数域 $\mathbb{F}$ 上的非常数多项式, 求证: $f(x)$ 等于某个不可约多项式的幂的充要条件是对任意的非常数多项式 $g(x)$, 或者 $f(x)$ 和 $g(x)$ 互素, 或者 $f(x)$ 可以整除 $g(x)$ 的某个幂.

证明 设 $f(x) = p(x)^k$, $p(x)$ 在 $\mathbb{F}$ 上不可约, 且 $f(x)$ 和 $g(x)$ 不互素, 则 $p(x)$ 是 $f(x)$ 和 $g(x)$ 的公因式, 故 $f(x)$ 可以整除 $g(x)^k$.

反之, 由因式分解定理, 可设 $f(x) = p(x)^m h(x)$, $p(x)$ 在 $\mathbb{F}$ 上不可约, $\deg h(x) > 0$, 且 $p(x)$ 不能整除 $h(x)$, 则 $f(x) \mid h(x)$, 故 $f(x)$ 不和 $h(x)$ 互素. 由于 $\deg h(x) < \deg f(x)$, 因此 $f(x)$ 也不能整除 $h(x)$, 矛盾! □

定义 5.6 (代数数)

设 u 是复数域中某个数, 若 u 适合某个非零有理系数多项式 (或整系数多项式) $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, 则称 u 是一个**代数数**.

定义 5.7 (极小多项式 (最小多项式))

对任一代数数 u, 存在唯一一个 u 适合的首一有理系数多项式 $g(x)$, 使得 $g(x)$ 是 u 适合的所有非零有理系数多项式中次数最小者. 这样的 $g(x)$ 称为 u 的**极小多项式**或**最小多项式**.

证明 现在证明这个定义是良定义的, 只须证明对任一代数数所对应的极小多项式的存在性和唯一性.

先证存在性. 在 u 适合的所有非零有理系数多项式构成的集合中 (由假设这个集合非空, 否则 u 就不是一个代数数), 由良序公理可知, 存在一个次数最小的多项式, 然后将其首一化, 即可得到 u 的极小多项式 $g(x)$.

再证唯一性. 为了证明极小多项式的唯一性, 我们先证明极小多项式的一个基本性质, 即极小多项式可以整除 u 适合的任一多项式 $f(x)$. 假设
$$f(x) = g(x)q(x) + r(x), \deg r(x) < \deg g(x),$$
则由 $f(u) = g(u) = 0$ 可知 $r(u) = 0$. 若 $r(x) \neq 0$, 则 u 适合一个比 $g(x)$ 的次数更小的多项式 $r(x)$, 这和 $g(x)$ 是极小多项式矛盾. 因此 $r(x) = 0$, 即 $g(x) \mid f(x)$. 设 $h(x)$ 也是 u 的极小多项式, 则由上述性质可得 $g(x) \mid h(x)$, $h(x) \mid g(x)$, 从而 $g(x)$ 和 $h(x)$ 只差一个非零常数, 又它们都是首一的, 故只能相等, 唯一性得证. □

命题 5.14 (极小多项式的基本性质)

(1) 设 $g(x)$ 为 u 的极小多项式, 则 $g(x)$ 一定整除 u 适合的任一多项式 $f(x)$.

证明

(1) 假设
$$f(x) = g(x)q(x) + r(x), \deg r(x) < \deg g(x),$$

则由 $f(u)=g(u)=0$ 可知 $r(u)=0$. 若 $r(x)\neq 0$, 则 u 适合一个比 $g(x)$ 的次数更小的多项式 $r(x)$, 这和 $g(x)$ 是极小多项式矛盾. 因此 $r(x)=0$, 即 $g(x)\mid f(x)$.

□

命题 5.15 (极小多项式式的充要条件)

设 $g(x)$ 是一个 u 适合的首一有理系数多项式, 则 $g(x)$ 是 u 的极小多项式的充要条件是 $g(x)$ 是有理数域上的不可约多项式.

证明 先证必要性. 若极小多项式 $g(x)$ 在有理数域上可约, 则 $g(x)=g_1(x)g_2(x)$ 可分解为两个比 $g(x)$ 的次数更小的多项式的乘积. 由 $0=g(u)=g_1(u)g_2(u)$ 可知 $g_1(u)$ 和 $g_2(u)$ 中至少有一个等于零. 不妨设 $g_1(u)=0$, 则 u 适合一个比 $g(x)$ 的次数更小的多项式 $g_1(x)$, 这和 $g(x)$ 是极小多项式矛盾.

再证充分性. 设 $g(x)$ 是 u 适合的有理数域上的首一不可约多项式, $h(x)$ 是 u 的极小多项式. 由极小多项式的基本性质 (1)可知 $h(x)\mid g(x)$. 因为 $g(u)=h(u)=0$, 所以 $g(x)$ 和 $h(x)$ 有公共根, 从而 $x-u$ 一定是 $g(x),h(x)$ 的公因式, 于是 $g(x)$ 和 $h(x)$ 不互素. 又 $g(x)$ 是不可约多项式, 因此 $g(x)\mid h(x)$. 于是 $g(x)\sim h(x)$, 即 $g(x)$ 和 $h(x)$ 只差一个非零常数, 而它们又都是首一的, 故只能相等. 因此 $g(x)$ 就是 u 的极小多项式. □

5.4.1 多项式的标准分解

多项式的标准分解是证明某些问题的有力工具.

定理 5.8 (因式分解定理)

设 $f(x)$ 是数域 $\mathbb{K}$ 上的多项式且 $\deg f(x)\geq 1$, 则

(1) $f(x)$ 可分解为有限个 $\mathbb{K}$ 上的不可约多项式之积;

(2) 若

$$f(x)=p_1(x)p_2(x)\cdots p_s(x)=q_1(x)q_2(x)\cdots q_t(x). \tag{5.8}$$

是 $f(x)$ 的两个不可约分解, 即 $p_i(x),q_j(x)$ 都是 $\mathbb{K}$ 上的次数大于零的不可约多项式, 则 $s=t$, 且经过适当调换因式的次序以后, 有

$$q_i(x)\sim p_i(x),\ i=1,2,\cdots,s.$$

笔记

1. 这个定理表明, 任一多项式可唯一地分解为若干个不可约多项式之积. 这里唯一是在相伴意义下的唯一, 即相应的多项式可以差一个常数因子. 如果把分解式中相同或仅差一个常数的因式合并在一起, 就得到了一个 **标准分解式**:
$$f(x)=cp_1(x)^{e_1}p_2(x)^{e_2}\cdots p_m(x)^{e_m}, \tag{5.9}$$
其中 $c\neq 0, p_i(x)$ 是互异的首一不可约多项式, $e_i\geq 1(i=1,2,\cdots,m)$.
若 $e_i>1$ $(e_i=1)$, 我们称(5.9)式中的因式 $p_i(x)$ 为 $f(x)$ 的e_i **重因式 (单因式)**. 显然这时 $p_i(x)^{e_i}\mid f(x)$, 但 $p_i(x)^{e_i+1}$ 不能整除 $f(x)$.
2. 设 $f(x),g(x)$ 是 $\mathbb{K}$ 上的两个多项式, 在它们的标准分解式中适当添加零次项, 就能得到公共的标准分解. 故对 $\mathbb{K}$ 上任意的两个多项式 $f(x),g(X)$, 都可以不妨设它们有如下的**公共的标准分解式**:
$$f(x)=c_1p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_n(x)^{e_n};$$
$$g(x)=c_2p_1(x)^{f_1}p_2(x)^{f_2}\cdots p_n(x)^{f_n},$$
其中 $e_i\geq 0, f_i\geq 0(i=1,2,\cdots,n)$.

证明 (1) 对多项式 $f(x)$ 的次数用数学归纳法. 若 $\deg f(x)=1$, 结论显然成立. 设次数小于 n 的多项式都可以分解

为 $\mathbb{K}$ 上的不可约多项式之积而 $\deg f(x) = n$. 若 $f(x)$ 不可约, 结论自然成立. 若 $f(x)$ 可约, 则

$$f(x) = f_1(x)f_2(x),$$

其中 $f_1(x), f_2(x)$ 的次数小于 n, 由归纳假设它们可以分解为有限个 $\mathbb{K}$ 上的不可约多项式之积. 所有这些多项式之积就是 $f(x)$.

(2) 对(5.8)式中的 s 用数学归纳法. 若 $s = 1$, 则 $f(x) = p_1(x)$, 因此 $f(x)$ 是不可约多项式, 于是 $t = 1, q_1(x) = p_1(x)$. 现假设对不可约因式个数小于 s 的多项式结论正确. 由(5.8)式, 有

$$p_1(x) \mid q_1(x)q_2(x)\cdots q_t(x),$$

由推论 5.5可知, 必存在某个 i, 不妨设 $i = 1$, 使

$$p_1(x) \mid q_1(x).$$

但是 $p_1(x), q_1(x)$ 都是不可约多项式, 因此存在 $0 \neq c_1 \in \mathbb{K}$, 使

$$q_1(x) = c_1 p_1(x),$$

此即 $p_1(x) \sim q_1(x)$. 将上式代入(5.8)式并消去 $p_1(x)$, 得到

$$p_2(x)\cdots p_s(x) = c_1 q_2(x)\cdots q_t(x).$$

这时左边为 $s-1$ 个不可约多项式之积, 由归纳假设, $s-1 = t-1$, 即 $s = t$. 另一方面, 存在 $0 \neq c_i \in \mathbb{K}$, 使 $q_i(x) = c_i p_i(x)$. □

推论 5.6

设 $f(x), g(x)$ 是 $\mathbb{K}$ 上的两个多项式, 不妨设它们有如下的公共的标准分解式:

$$f(x) = c_1 p_1(x)^{e_1} p_2(x)^{e_2} \cdots p_n(x)^{e_n};$$
$$g(x) = c_2 p_1(x)^{f_1} p_2(x)^{f_2} \cdots p_n(x)^{f_n},$$

其中 $e_i \geq 0, f_i \geq 0 (i = 1, 2, \cdots, n)$, 则 $f(x), g(x)$ 的最大公因式

$$(f(x), g(x)) = p_1(x)^{k_1} p_2(x)^{k_2} \cdots p_n(x)^{k_n},$$

其中 $k_i = \min\{e_i, f_i\} (i = 1, 2, \cdots, n)$.

类似地, $f(x), g(x)$ 的最小公倍式

$$[f(x), g(x)] = p_1(x)^{h_1} p_2(x)^{h_2} \cdots p_n(x)^{h_n},$$

其中 $h_i = \max\{e_i, f_i\} (i = 1, 2, \cdots, n)$.

♡

证明 利用最大公因式和最小公倍式的定义容易证明. □

命题 5.16 (整除关系在平方下不变)

证明:$g(x)^2 \mid f(x)^2$ 的充要条件是 $g(x) \mid f(x)$.

♠

证明 充分性是显然的, 只需证明必要性. 设 $f(x), g(x)$ 的公共标准分解为

$$f(x) = cp_1(x)^{e_1} p_2(x)^{e_2} \cdots p_k(x)^{e_k}, \quad g(x) = dp_1(x)^{f_1} p_2(x)^{f_2} \cdots p_k(x)^{f_k},$$

其中 $p_i(x)$ 为互不相同的首一不可约多项式,c, d 是非零常数, 则

$$f(x)^2 = c^2 p_1(x)^{2e_1} p_2(x)^{2e_2} \cdots p_k(x)^{2e_k}, \quad g(x)^2 = d^2 p_1(x)^{2f_1} p_2(x)^{2f_2} \cdots p_k(x)^{2f_k}.$$

若 $g(x)^2 \mid f(x)^2$, 则 $2f_i \leq 2e_i$, 从而 $f_i \leq e_i (1 \leq i \leq k)$. 因此 $g(x) \mid f(x)$. □

5.5 多项式函数与根

定义 5.8 (多项式的重根)

设 $f(x) \in \mathbb{K}[x]$,$b \in \mathbb{K}$, 若存在正整数 k, 使 $(x-b)^k \mid f(x)$, 但 $(x-b)^{k+1}$ 不能整除 $f(x)$, 则称 b 是 $f(x)$ 的一个 **k 重根**. 若 $k=1$, 则称 b 为**单根**.

定理 5.9 (多项式没有重因式的充要条件)

数域 $\mathbb{K}$ 上的多项式 $f(x)$ 没有重因式的充分必要条件是 $f(x)$ 与 $f'(x)$ 互素.

证明 设多项式 $p(x)$ 是 $f(x)$ 的 $m(m>1)$ 重因式, 则 $f(x)=p(x)^m g(x)$, 故

$$f'(x)=mp(x)^{m-1}p'(x)g(x)+p(x)^m g'(x).$$

于是 $p(x)^{m-1} \mid f'(x)$, 这表明 $f(x)$ 与 $f'(x)$ 有公因式 $p(x)^{m-1}$. 反之, 若不可约多项式 $p(x)$ 是 $f(x)$ 的单因式, 可设 $f(x)=p(x)g(x)$, $p(x)$ 不能整除 $g(x)$. 于是

$$f'(x)=p'(x)g(x)+p(x)g'(x).$$

若 $p(x)$ 是 $f'(x)$ 的因式, 则 $p(x) \mid p'(x)g(x)$. 但 $p(x)$ 不能整除 $g(x)$ 且 $p(x)$ 不可约, 故 $p(x) \mid p'(x)$. 而 $p'(x) \neq 0$ 且 $\deg p'(x) < \deg p(x)$, 这是不可能的. 若 $f(x)$ 无重因式, 则在 $f(x)$ 的标准分解式(5.9)中, $e_i=1$ 对一切 $i=1,2,\cdots,m$ 成立, 于是 $p_i(x)$ 都不能整除 $f'(x)$. 由于 $p_i(x)$ 为不可约多项式, 故 $(p_i(x),f'(x))=1$, 由互素多项式和最大公因式的基本性质 (5)可知

$$(p_1(x)p_2(x)\cdots p_m(x), f'(x))=1,$$

即 $(f(x),f'(x))=1$. □

定理 5.10

设 $d(x)=(f(x),f'(x))$, 则 $f(x)/d(x)$ 是一个没有重因式的多项式, 且这个多项式的不可约因式与 $f(x)$ 的不可约因式相同 (不计重数).

证明 设 $f(x)$ 有如(5.9)式的标准分解式, 则

$$\begin{aligned} f'(x) &= ce_1 p_1(x)^{e_1-1}p_2(x)^{e_2}\cdots p_s(x)^{e_s}p_1'(x)\\ &\quad + ce_2 p_1(x)^{e_1}p_2(x)^{e_2-1}\cdots p_s(x)^{e_s}p_2'(x)\\ &\quad +\cdots\\ &\quad + ce_s p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_s(x)^{e_s-1}p_s'(x). \end{aligned}$$

因此 $p_1(x)^{e_1-1}p_2(x)^{e_2-1}\cdots p_s(x)^{e_s-1}$ 是 $f(x)$ 与 $f'(x)$ 的公因式. 注意到 $f(x)$ 的因式一定具有 $p_1(x)^{k_1}p_2(x)^{k_2}\cdots p_s(x)^{k_s}$ 的形状. 不妨设 $h(x)$ 是 $f(x),f'(x)$ 的公因式. 注意到 $p_1(x)^{e_1}$ 可以整除(??)式中右边除第一项外的所有项, 但不能整除第一项, 因此 $p_1(x)^{e_1}$ 不能整除 $f'(x)$. 同理, $p_i(x)^{e_i}$ 不能整除 $f'(x)$. 由此我们不难看出

$$h(x) \mid p_1(x)^{e_1-1}p_2(x)^{e_2-1}\cdots p_s(x)^{e_s-1},$$

即 $p_1(x)^{e_1-1}p_2(x)^{e_2-1}\cdots p_s(x)^{e_s-1}=d(x)$. 显然 $f(x)/d(x)$ 没有重因式且与 $f(x)$ 含有相同的不可约因式. □

命题 5.17 (多项式有 k 重根的充要条件)

求证:a 是多项式 $f(x)$ 的 k 重根的充要条件是:

$$f(a)=f'(a)=\cdots=f^{(k-1)}(a)=0, \quad f^{(k)}(a)\neq 0.$$

证明 若 a 是 $f(x)$ 的 k 重根, 可设 $f(x)=(x-a)^k g(x)$, $g(x)$ 不含因式 $x-a$. 通过对 $f(x)$ 求导可发现, $x-a$ 可整除

$f^{(j)}(x)\ (1 \le j \le k-1)$. 因此

$$f(a) = f'(a) = \cdots = f^{(k-1)}(a) = 0.$$

而 $f^{(k)}(a) = k!g(a) \neq 0$, 故必要性得证.

反之, 若 a 是 $f(x)$ 的 m 重根, 若 $m > k$, 则由必要性的证明可知, 将有 $f^{(k)}(a) = 0$, 这与已知矛盾. 同样, 若 $m < k$, 则由必要性的证明可知, 将有 $f^{(m)}(a) \neq 0$, 这也与已知矛盾, 于是只能 $m = k$. □

命题 5.18

设 $\deg f(x) = n \ge 1$, 若 $f'(x) \mid f(x)$, 证明:$f(x)$ 有 n 重根.

证明 证法一: 设 $f(x) = \frac{1}{n}(x-a)f'(x)$, 现证明 a 是 $f(x)$ 的 n 重根. 假设 a 是 $f(x)$ 的 k 重根,$f(x) = (x-a)^k g(x)$,$k < n$ 且 $g(x)$ 不含因式 $x-a$, 则

$$f'(x) = k(x-a)^{k-1}g(x) + (x-a)^k g'(x) = n(x-a)^{k-1}g(x).$$

于是 $g(x) \mid (x-a)g'(x)$, 而 $g(x)$ 与 $x-a$ 互素, 故将有 $g(x) \mid g'(x)$. 引出矛盾.

证法二: 设 $f(x) = \frac{1}{n}(x-a)f'(x)$, 则

$$\frac{f(x)}{(f(x), f'(x))} = b(x-a), \quad b \neq 0.$$

由定理 5.10可知,$x-a$ 是 $f(x)$ 唯一的不可约因式, 因此 $f(x) = b(x-a)^n$. □

命题 5.19

数域 $\mathbb{F}$ 上任意一个不可约多项式在复数域 $\mathbb{C}$ 中无重根.

证明 设 $f(x)$ 是 $\mathbb{F}$ 上的不可约多项式, 则 $\deg f(x) <\deg f'(x)$. 从而 $f(x) \nmid f'(x)$, 于是 $(f(x), f'(x)) = 1$. 故由多项式没有重因式的充要条件可知 $f(x)$ 复数域 $\mathbb{C}$ 中无重根. □

引理 5.1 (次数不为 1 得到不可约多项式没有根)

设 $f(x)$ 是数域 $\mathbb{K}$ 上的不可约多项式且 $\deg f(x) \ge 2$, 则 $f(x)$ 在 $\mathbb{K}$ 中没有根.

证明 用反证法, 设 $b \in \mathbb{K}$ 是 $f(x)$ 的根, 由余数定理 知 $(x-b) \mid f(x)$, 即 $f(x) = (x-b)g(x)$ 可分解为两个低次多项式之积, 这与 $f(x)$ 不可约矛盾. □

定理 5.11 (多项式根的有限性)

设 $f(x)$ 是数域 $\mathbb{F}$ 上的 n 次多项式, 则 $f(x)$ 在 $\mathbb{F}$ 中最多只有 n 个根.

笔记 由命题 1.14可知, 若一个 n 次多项式的根超过 n 个, 则这个多项式一定恒为零.

证明 将 $f(x)$ 作标准因式分解, 则由次数不为 1 得到不可约多项式没有根知 $f(x)$ 在 $\mathbb{K}$ 中根的个数等于该分解式中一次因式的个数, 它不会超过 n. □

推论 5.7 (两个多项式相等的判定准则)

设 $f(x)$ 与 $g(x)$ 是 $\mathbb{K}$ 上的次数不超过 n 的两个多项式, 若存在 $\mathbb{K}$ 上 $n+1$ 个不同的数 $b_1, b_2, \ldots, b_{n+1}$, 使

$$f(b_i) = g(b_i), \quad i = 1, 2, \ldots, n+1,$$

则 $f(x) = g(x)$.

证明 作 $h(x) = f(x) - g(x)$, 显然 $h(x)$ 次数不超过 n. 但它有 $n+1$ 个不同的根, 因此只可能 $h(x) = 0$, 即 $f(x) = g(x)$.
□

例题 5.5 求证:$f(x) = \sin x$ 在实数域内不能表示为 x 的多项式.

证明 注意到 $f(x) = \sin x$ 在实数域内有无穷多个根, 而任一非零多项式只能有有限个根, 因此 $f(x) = \sin x$ 在实数

域内不能表示为 x 的多项式. □

例题 5.6 设 $f(x)$ 是数域 $\mathbb{F}$ 上的多项式, 若对 $\mathbb{F}$ 中某个非零常数 a, 有 $f(x+a)=f(x)$, 求证:$f(x)$ 必是常数多项式.

证明 假设 $f(x)$ 不是常数多项式, 则 $f(x)-f(a)$ 也不是常数多项式, 但由 $f(x+a)=f(x)$ 可知,ka $(k\in\mathbb{Z})$ 是 $f(x)-f(a)$ 的无穷多个根, 矛盾. □

例题 5.7 设 $f(x)$ 是非常数多项式且 $f(x)$ 可以整除 $f(x^m)$ $(m\in\mathbb{N}_+)$, 求证:$f(x)$ 的根只能是 0 或 1 的某个方根.

证明 将 $f(x)$ 看成复数域上的多项式, 则 $f(x^m)=f(x)g(x)$. 假设 c 是 $f(x)$ 的一个复根, 即 $f(c)=0$, 则 $f(c^m)=0$, 即 c^m 也是 $f(x)$ 的根. 由此可知 $c^m, c^{m^2}, c^{m^3}, \ldots$ 也都是 $f(x)$ 的根. 由于 $f(x)$ 只有有限个不同的复根, 故存在正整数 $k>t$, 使得 $c^{m^k}=c^{m^t}$. 因此若 $c\neq 0$, 取 $n=m^k-m^t\in\mathbb{N}_+$, 则有 $c^n=1$. □

定理 5.12 (余数定理)

设 $f(x)\in\mathbb{F}[x], b\in\mathbb{F}$, 则存在 $\mathbb{F}$ 上的多项式 $g(x)$, 使得

$$f(x)=(x-b)g(x)+f(b).$$

特别地,b 是 $f(x)$ 的根的充要条件是 $(x-b)\mid f(x)$.

笔记 利用余数定理可以实现求根与判断整除性之间的相互转换.

证明 由带余除法知

$$f(x)=(x-b)g(x)+r(x),$$

其中 $\deg r(x)<1$, 因此 $r(x)$ 为常数多项式. 在上式中用 b 代替 x, 即得 $r(x)=f(b)$. □

例题 5.8 设 n 是奇数, 求证:$(x+y)(y+z)(x+z)$ 可整除 $(x+y+z)^n-x^n-y^n-z^n$.

证明 将多项式 $(x+y+z)^n-x^n-y^n-z^n$ 看成是未定元 x 的多项式. 当 $x=-y$ 时,$(x+y+z)^n-x^n-y^n-z^n=0$, 因此由余数定理可知 $x+y$ 是 $(x+y+z)^n-x^n-y^n-z^n$ 的因式. 同理 $x+z, y+z$ 也是因式. 又这 3 个因式互素, 故 $(x+y)(y+z)(x+z)$ 可整除 $(x+y+z)^n-x^n-y^n-z^n$. □

例题 5.9 设 $f(x)$ 是一个 n 次多项式, 若当 $k=0,1,\ldots,n$ 时有 $f(k)=\dfrac{k}{k+1}$, 求 $f(n+1)$.

证明 解令 $g(x)=(x+1)f(x)-x$, 则 $0,1,\ldots,n$ 是 $g(x)$ 的根, 因此

$$g(x)=cx(x-1)(x-2)\cdots(x-n),$$

即

$$(x+1)f(x)-x=cx(x-1)(x-2)\cdots(x-n),$$

其中 c 是一个常数. 令 $x=-1$, 可求出 $c=\dfrac{(-1)^{n+1}}{(n+1)!}$. 从而

$$f(x)=\frac{1}{x+1}\left(\frac{(-1)^{n+1}x(x-1)\cdots(x-n)}{(n+1)!}+x\right),$$

故

$$f(n+1)=\frac{1}{n+2}\left(\frac{(-1)^{n+1}(n+1)!}{(n+1)!}+n+1\right).$$

当 n 是奇数时,$f(n+1)=1$; 当 n 是偶数时,$f(n+1)=\dfrac{n}{n+2}$. □

例题 5.10 设 $(x^4+x^3+x^2+x+1)\mid(x^3f_1(x^5)+x^2f_2(x^5)+xf_3(x^5)+f_4(x^5))$, 这里 $f_i(x)$ $(1\le i\le 4)$ 都是实系数多项式, 求证:$f_i(1)=0$ $(1\le i\le 4)$.

证明 设 ε_i $(1\le i\le 4)$ 是 1 的五次虚根, 则 ε_i $(1\le i\le 4)$ 都适合 x^5-1, 从而由余数定理可知

$$x^5-1=(x-1)(x-\varepsilon_1)(x-\varepsilon_2)(x-\varepsilon_3)(x-\varepsilon_4)=(x-1)(x^4+x^3+x^2+x+1).$$

故

$$x^4+x^3+x^2+x+1=(x-\varepsilon_1)(x-\varepsilon_2)(x-\varepsilon_3)(x-\varepsilon_4).$$

因此 ε_i $(1 \le i \le 4)$ 都是 $x^4+x^3+x^2+x+1$ 的根. 由条件可得

$$\varepsilon_i^3 f_1(1)+\varepsilon_i^2 f_2(1)+\varepsilon_i f_3(1)+f_4(1)=0 \quad (1 \le i \le 4).$$

这是一个由 4 个未知数、4 个方程式组成的线性方程组（将 $f_i(1)$ 看成是未知数），其系数行列式是一个 Vandermonde 行列式, 显然其值不等于零, 因此 $f_i(1)=0$. □

5.6 复系数多项式

定理 5.13 (代数基本定理)

次数大于零的复数域上的一元多项式至少有一个复数根. ♡

证明 设复数域上的 n 次多项式为

$$f(z)=a_nz^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0. \tag{5.10}$$

首先证明, 必存在一个复数 z_0, 使对一切复数 z, 有

$$|f(z)| \ge |f(z_0)|.$$

令 $z=x+iy$, 其中 x,y 是实变量. 展开 $f(x+iy)$ 并分开实部和虚部, 则

$$f(z)=u(x,y)+iv(x,y),$$

其中 $u(x,y)$ 及 $v(x,y)$ 为实系数二元多项式函数. 又

$$|f(z)|=\sqrt{u(x,y)^2+v(x,y)^2}$$

是一个二元连续函数, 但

$$\begin{aligned}|f(z)|&=|a_nz^n+a_{n-1}z^{n-1}+\cdots+a_0|\\&\ge |a_nz^n|-|a_{n-1}z^{n-1}+\cdots+a_0|\\&\ge |z|^n\left(|a_n|-\left|\frac{a_{n-1}}{z}\right|-\left|\frac{a_{n-2}}{z^2}\right|+\cdots+\left|\frac{a_0}{z^n}\right|\right),\end{aligned}$$

因此当 $|z|\to\infty$ 时,$|f(z)|\to\infty$. 于是必存在一个实常数 R, 当 $|z|\ge R$ 时,$|f(z)|$ 充分大, 因此 $|f(z)|$ 的最小值必含于圆圈 $|z|\le R$ 中. 但这是平面上的一个闭区域, 因此必存在 z_0 使 $|f(z_0)|$ 为最小.

接下来要证明 $f(z_0)=0$. 用反证法, 即若 $f(z_0)\ne 0$, 则必可找到 z_1, 使 $|f(z_1)|<|f(z_0)|$, 这样就与 $|f(z_0)|$ 是最小值相矛盾. 将 $z=z_0+h$ 代入 (5.10)式便可得到一个关于 h 的 n 次多项式:

$$f(z_0+h)=b_nh^n+b_{n-1}h^{n-1}+\cdots+b_1h+b_0. \tag{5.11}$$

当 $h=0$ 时,$f(z_0)=b_0$, 由假设 $f(z_0)\ne 0$, 故

$$\frac{f(z_0+h)}{f(z_0)}=\frac{b_n}{f(z_0)}h^n+\frac{b_{n-1}}{f(z_0)}h^{n-1}+\cdots+\frac{b_1}{f(z_0)}h+1.$$

$b_1,b_2,\ldots,b_n$ 中有些可能为零, 但绝不全为零. 设 b_k 是第一个不为零的复数, 则

$$\frac{f(z_0+h)}{f(z_0)}=1+c_kh^k+c_{k+1}h^{k+1}+\cdots+c_nh^n, \tag{5.12}$$

其中 $c_j=\dfrac{b_j}{f(z_0)}$. 令 $d=\sqrt[k]{\dfrac{1}{|c_k|}}$,$h=ed$ 代入(5.12)式得

$$\frac{f(z_0+h)}{f(z_0)}=1-e^k+e^{k+1}(c_{k+1}d^{k+1}+c_{k+2}d^{k+2}e+\cdots).$$

取充分小的正实数 e（至少小于 1）, 使

$$e(|c_{k+1}d^{k+1}|+|c_{k+2}d^{k+2}|+\cdots)<\frac{1}{2},$$

于是

$$\begin{aligned}\left|\frac{f(z_0+h)}{f(z_0)}\right| &\le |1-e^k|+|e^{k+1}(c_{k+1}d^{k+1}+c_{k+2}d^{k+2}e+\cdots)| \\ &\le 1-e^k+e^{k+1}(|c_{k+1}d^{k+1}|+|c_{k+2}d^{k+2}|+\cdots) \\ &< 1-e^k+\frac{1}{2}e^k \\ &= 1-\frac{1}{2}e^k < 1.\end{aligned}$$

将这样的 e 代入 $h=ed$, 得

$$|f(z_0+ed)|<|f(z_0)|.$$

这就推出了矛盾. □

推论 5.8

1. 复数域上的一元 n 次多项式恰有 n 个复根（包括重根）.
2. 复数域上的不可约多项式都是一次多项式.
3. 复数域上的一元 n 次多项式必可分解为一次因式的乘积.

定理 5.14 (Vieta 定理)

若数域 $\mathbb{F}$ 上的多项式 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$ 在 $\mathbb{F}$ 中有 n 个根 $x_1,x_2,\cdots,x_n$, 则

$$\begin{aligned}&\sum_{i=1}^{n}x_i=x_1+x_2+\cdots+x_n=-\frac{a_1}{a_0}, \\ &\sum_{1\le i<j\le n}x_ix_j=x_1x_2+\cdots+x_1x_n+x_2x_3+\cdots+x_{n-1}x_n=\frac{a_2}{a_0}, \\ &\cdots\cdots\cdots\cdots \\ &x_1x_2\cdots x_n=(-1)^n\frac{a_n}{a_0}.\end{aligned}$$

证明 $f(x)=a_0(x-x_1)(x-x_2)\cdots(x-x_n)$, 将这个式子的右边展开与 $f(x)$ 比较系数即得结论. □

例题 5.11

(1) 设三次方程 $x^3+px^2+qx+r=0$ 的 3 个根成等差数列, 求证:

$$2p^3-9pq+27r=0.$$

(2) 设三次方程 $x^3+px^2+qx+r=0\ (r\ne 0)$ 的 3 个根成等比数列, 求证:

$$rp^3=q^3.$$

(3) 设多项式 x^3+3x^2+mx+n 的 3 个根成等差数列, 多项式 $x^3-(m-2)x^2+(n-3)x+8$ 的 3 个根成等比数列, 求 m 和 n.

证明

(1) 设方程的 3 个根为 $c-d,c,c+d$, 则由 Vieta 定理可得

$$\begin{cases}3c=-p, \\ 3c^2-d^2=q, \\ c(c^2-d^2)=-r.\end{cases}$$

由此可得 $2p^3-9pq+27r=0$.

(2) 设方程的 3 个根为 $\frac{c}{d}, c, cd$, 则由 Vieta 定理可得

$$\begin{cases} \frac{c}{d}+c+cd=-p, \\ \frac{c^2}{d}+c^2+c^2d=q, \\ \frac{c^3}{d}=-r. \end{cases}$$

由此可得 $rp^3=q^3$.

(3) 由 (1)(2) 可知,m,n 应满足如下关系:

$$\begin{cases} m=n+2, \\ -8(m-2)^3=(n-3)^3. \end{cases}$$

若 $n-3=-2(m-2)$, 则可联立求得 $m=3, n=1$.

若 $n-3=-2\omega(m-2)$, 其中 $\omega=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$, 则可联立求得 $m=2-\sqrt{3}i, n=-\sqrt{3}i$.

若 $n-3=-2\omega^2(m-2)$, 则可联立求得 $m=2+\sqrt{3}i, n=\sqrt{3}i$.

□

例题 5.12 设 x_1,x_2,x_3 是三次方程 $x^3+px^2+qx+r=0\ (r\neq 0)$ 的 3 个根, 求这 3 个根倒数的平方和.

证明 由 Vieta 定理可得

$$\frac{1}{x_1^2}+\frac{1}{x_2^2}+\frac{1}{x_3^2}=\frac{(x_1x_2+x_1x_3+x_2x_3)^2-2x_1x_2x_3(x_1+x_2+x_3)}{x_1^2x_2^2x_3^2}=\frac{q^2-2pr}{r^2}.$$

□

例题 5.13 已知方程 $x^3+px^2+qx+r=0$ 的 3 个根为 x_1,x_2,x_3, 求一个三次方程使其根为 x_1^3,x_2^3,x_3^3.

笔记 利用代数恒等式:$a^3+b^3+c^3=(a+b+c)^3-3(a+b+c)(ab+bc+ac)+3abc$ 得到

$$x_1^3+x_2^3+x_3^3=(x_1+x_2+x_3)^3-3(x_1+x_2+x_3)(x_1x_2+x_2x_3+x_1x_3)+3x_1x_2x_3.$$

$$x_1^3x_2^3+x_1^3x_3^3+x_2^3x_3^3=(x_1x_2+x_2x_3+x_1x_3)^3-3x_1x_2x_3(x_1+x_2+x_3)(x_1x_2+x_2x_3+x_1x_3)+3x_1^3x_3^3x_2^3x_3^3.$$

即可由 Vieta 定理得到结果.

证明 由 Vieta 定理经计算可得

$$\begin{cases} x_1^3+x_2^3+x_3^3=-p^3+3pq-3r, \\ x_1^3x_2^3+x_1^3x_3^3+x_2^3x_3^3=q^3-3pqr+3r^2, \\ x_1^3x_2^3x_3^3=-r^3. \end{cases}$$

因此, 以 x_1^3,x_2^3,x_3^3 为根的三次方程为

$$x^3+(p^3-3pq+3r)x^2+(q^3-3pqr+3r^2)x+r^3=0.$$

□

例题 5.14 设多项式 x^3+px^2+qx+r 的 3 个根都是实数, 求证:$p^2\geq 3q$.

证明 设多项式的 3 个根为 x_1,x_2,x_3, 由已知条件可知:

$$(x_1-x_2)^2+(x_2-x_3)^2+(x_1-x_3)^2\geq 0.$$

用 Vieta 定理可计算出

$$\begin{aligned}(x_1-x_2)^2+(x_2-x_3)^2+(x_1-x_3)^2&=2\left(x_1^2+x_2^2+x_3^2\right)-2(x_1x_2+x_2x_3+x_1x_3)\\&=2(x_1+x_2+x_3)^2-6(x_1x_2+x_2x_3+x_1x_3)=2(p^2-3q).\end{aligned}$$

因此结论为真. □

例题 5.15 设 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 的 n 个根 $x_1,x_2,\cdots,x_n$ 皆不等于零, 求以 $\frac{1}{x_1},\frac{1}{x_2},\cdots,\frac{1}{x_n}$ 为根的多项式.

证明 令

$$g(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n,$$

则

$$x^n g\left(\frac{1}{x_i}\right) = a_0 + a_1x_i + \cdots + a_{n-1}x_i^{n-1} + a_nx_i^n = f(x_i) = 0.$$

因为 $x_i \neq 0$, 故 $g\left(\frac{1}{x_i}\right) = 0$, 即 $g(x)$ 的根为 $f(x)$ 根之倒数. □

例题 5.16 设 $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ $(a_n \neq 0)$ 是数域 $\mathbb{F}$ 上的可约多项式, 求证: 多项式 $g(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ 在 $\mathbb{F}$ 上也可约.

证明 设 $f(x) = p(x)q(x)$, 其中 $\deg p(x) = m, \deg q(x) = n - m, 0 < m < n$, 则

$$g(x) = x^n f\left(\frac{1}{x}\right) = x^n p\left(\frac{1}{x}\right) q\left(\frac{1}{x}\right) = \left(x^m p\left(\frac{1}{x}\right)\right)\left(x^{n-m} q\left(\frac{1}{x}\right)\right),$$

因此 $g(x)$ 也可约. □

5.7 实系数多项式

定理 5.15 (实系数多项式的复根成对出现)

设 $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ 是实系数多项式, 若复数 $a + bi$ $(b \neq 0)$ 是其根, 则 $a - bi$ 也是它的根.

证明 令 $z = a + bi$, 其共轭复数为 $\overline{z} = a - bi$, 则

$$f(\overline{z}) = a_n\overline{z}^n + a_{n-1}\overline{z}^{n-1} + \cdots + a_1\overline{z} + a_0 = \overline{a_nz^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0} = 0.$$

由此即得结论. □

推论 5.9

实数域上的不可约多项式为一次多项式或下列二次多项式:

$$ax^2 + bx + c, \quad 其中\ b^2 - 4ac < 0.$$

证明 一次多项式显然为不可约. 当 $b^2 - 4ac < 0$ 时,$ax^2 + bx + c$ 没有实根, 故不可约.

反过来, 任一高于二次的实系数多项式 $f(x)$ 如有实根, 则 $f(x)$ 可约; 如有一复根 $a + bi$ $(b \neq 0)$, 则 $a - bi$ 也是它的根, 从而

$$(x - (a + bi))(x - (a - bi)) = x^2 - 2ax + (a^2 + b^2)$$

是 $f(x)$ 的因式, 故任一高于二次的实系数多项式 $f(x)$ 都可约. 从而我们只需考虑一次和二次多项式的情况.

(i) 当 $f(x)$ 为一次实系数多项式时, 显然 $f(x)$ 一定不可约.

(ii) 当 $f(x)$ 为二次多项式时, 设 $f(x) = ax^2 + bx + c$, 则当 $f(x)$ 有实根, 即 $b^2 - 4ac \geq 0$ 时, 设 $f(x)$ 的两个实根分别为 x_1, x_2, 则 $f(x) = (x - x_1)(x - x_2)$, 此时 $f(x)$ 可约. 当 $f(x)$ 无实根, 即 $b^2 - 4ac < 0$ 时, 此时 $f(x)$ 在实数域上不可约. □

推论 5.10

实数域上的多项式 $f(x)$ 必可分解为有限个一次因式及不可约二次因式的乘积.

命题 5.20

设 $f(x)$ 是复数域上的多项式, 若对任意的实数 c,$f(c)$ 总是实数, 求证:$f(x)$ 是实系数多项式.

证明 设 $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, 分别令 $x = 0, 1, 2, \cdots, n$, 得到一个以 $a_n, a_{n-1}, \cdots, a_1, a_0$ 为未知

数, 由 $n+1$ 个方程式组成的实系数线性方程组. 该方程组的系数行列式是一个非零的 Vandermonde 行列式, 故方程组必有唯一解, 且解为实数. 因此 $f(x)$ 是实系数多项式. □

例题 5.17 证明: 奇数次实系数多项式必有实数根.

证明 实系数多项式的虚根总是成对出现的, 因此奇数次实系数多项式必有实数根. □

命题 5.21

设 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 是实系数多项式, 求证:

(1) 若 $a_i\ (0\le i\le n)$ 全是正数或全是负数, 则 $f(x)$ 没有非负实根, 即只有负实根.

(2) 若 $(-1)^ia_i\ (0\le i\le n)$ 全是正数或全是负数, 则 $f(x)$ 没有非正实根, 即只有正实根.

(3) 若 $a_n>0$ 且 $(-1)^{n-i}a_i>0\ (0\le i\le n-1)$, 则 $f(x)$ 没有非正实根, 即只有正实根.; 若 $a_n>0$ 且 $(-1)^{n-i}a_i\ge 0\ (0\le i\le n-1)$, 则 $f(x)$ 没有负实根,, 即只有正实数和零能作为根.

证明 (1) 若 a_i 全是正数且 $f(x)$ 有非负实根 $c\ge 0$, 代入后可得

$$f(c)=a_nc^n+a_{n-1}c^{n-1}+\cdots+a_1c+a_0\ge a_0>0,$$

这和 c 是根矛盾, 因此 $f(x)$ 没有非负实根. 同理可证 a_i 全是负数的情形.

(2) 和 (3) 同理可证. □

例题 5.18 求证: 实系数方程 $x^3+px^2+qx+r=0$ 的根的实部全是负数的充要条件是

$$p>0,\quad r>0,\quad pq>r.$$

证明 先证必要性: 设原方程的 3 个根为 x_1,x_2,x_3, 其中 x_1 是实数根,$x_1<0$. 另假设 $x_2=a+bi,x_3=a-bi,a<0$, 则由 Vieta 定理可得

$$\begin{aligned}p&=-(x_1+x_2+x_3)=-(x_1+2a)>0,\\ r&=-x_1x_2x_3=-x_1(a^2+b^2)>0.\end{aligned}$$

$$\begin{aligned}pq-r&=-(x_1+2a)(x_1x_2+x_1x_3+x_2x_3)+x_1(a^2+b^2)\\ &=-(x_1+2a)(2x_1a+a^2+b^2)+x_1(a^2+b^2)\\ &=-2a((x_1+a)^2+b^2)>0.\end{aligned}$$

又假设 x_1,x_2,x_3 全是负实数, 则显然 $p>0,q>0,r>0$, 而

$$\begin{aligned}pq-r&=-(x_1+x_2+x_3)(x_1x_2+x_1x_3+x_2x_3)+x_1x_2x_3\\ &=-((x_1+x_2+x_3)(x_1x_2+x_1x_3+x_2x_3)+x_1x_2x_3)>0.\end{aligned}$$

再证充分性: 由 $p>0,r>0,pq-r>0$ 可知 $q>0$, 若方程的根是实数, 则由命题 5.21(1)可知, 此根必是负数. 现假设方程有根 $x_1<0,x_2=a+bi,x_3=a-bi$, 因为

$$pq-r=-2a((x_1+a)^2+b^2)>0,$$

故得 $a<0$, 结论得证. □

例题 5.19 设 ε 是 1 的 n 次根:

$$\varepsilon=\cos\frac{2\pi}{n}+i\sin\frac{2\pi}{n},$$

求证:$\varepsilon^{mi}\ (1\le i\le n)$ 是 $x^n-1=0$ 的全部根的充要条件是 $(m,n)=1$.

证明 若 $(m,n)=1$, 只要证明 $\varepsilon^{mi}\ (1\le i\le n)$ 互不相同即可. 若不然, 有 $\varepsilon^{ms}=\varepsilon^{mt}\ (1\le s<t\le n)$, 便有 $\varepsilon^{m(t-s)}=1$,$n\mid m(t-s)$. 因为 m,n 互素, 故 $n\mid(t-s)$, 而 $0<t-s<n$, 矛盾.

反之, 若 $(m,n)=d>1$, 则 $\varepsilon^{m\frac{n}{d}}=\varepsilon^{n\frac{m}{d}}=1$, 而 $\frac{n}{d}\in\mathbb{N}_+$ 且 $\frac{n}{d}<n$, 故 $\varepsilon^{mj}=\varepsilon^{m(j-\frac{n}{d})+m\frac{n}{d}}=\varepsilon^{m(j-\frac{n}{d})}$, $j=\frac{n}{d}+1,\cdots,n$. 于是 $\varepsilon^{mi}\ (i=1,2,\cdots,n)$ 只生成了 $\frac{n}{d}<n$ 个不同根. 而 x^n-1 有 n 个不同根, 故 $\varepsilon^{mi}\ (i=1,2,\cdots,n)$ 无法覆盖所有 x^n-1 的所有 n 个不同根. 从而 $\varepsilon^{mi}\ (1\le i\le n)$ 不可能是 $x^n-1=0$ 的全部根. □

命题 5.22

设 $f(x)$ 是实系数首一多项式且无实数根, 求证:$f(x)$ 可以表示为两个实系数多项式的平方和.

证明 因为实系数多项式的虚根成对出现, 故 $f(x)$ 是偶数次多项式, 不妨设它的根为

$$x_1, x_2, \cdots, x_n; \overline{x_1}, \overline{x_2}, \cdots, \overline{x_n}.$$

令

$$u(x) = (x - x_1)(x - x_2)\cdots(x - x_n); \quad v(x) = (x - \overline{x_1})(x - \overline{x_2})\cdots(x - \overline{x_n}),$$

则 $v(x) = \overline{u(x)}, f(x) = u(x)v(x)$. 又将 $u(x), v(x)$ 的实部和虚部分开, 可设

$$u(x) = g(x) + ih(x), \quad v(x) = g(x) - ih(x),$$

即有

$$f(x) = g(x)^2 + h(x)^2.$$

□

5.8 有理系数多项式

定理 5.16 (整数系数多项式有有理根的必要条件)

设有 n 次整系数多项式

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \tag{5.13}$$

则有理数 $\dfrac{q}{p}$ 是 $f(x)$ 的根的必要条件是 $p \mid a_n, q \mid a_0$, 其中 p, q 是互素的整数.

证明 将 $\dfrac{q}{p}$ 代入(5.13)式得

$$a_n\left(\frac{q}{p}\right)^n + a_{n-1}\left(\frac{q}{p}\right)^{n-1} + \cdots + a_1\left(\frac{q}{p}\right) + a_0 = 0,$$

将上式两边乘以 p^n 得

$$a_n q^n + a_{n-1}q^{n-1}p + \cdots + a_1 q p^{n-1} + a_0 p^n = 0.$$

从而

$$q\left(a_n q^{n-1} + a_{n-1}q^{n-2}p + \cdots + a_1 p^{n-1}\right) = -a_0 p^n.$$

于是 $q \mid a_0 p^n$, 又因为 $(q, p) = 1$, 所以 $q \mid a_0$. 同理可得 $p \mid a_n$. □

定义 5.9 (本原多项式)

设多项式

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

是整系数多项式, 若 $a_n, a_{n-1}, \cdots, a_1, a_0$ 的最大公约数等于 1, 则称 $f(x)$ 为**本原多项式**.

引理 5.2 (Gauss 引理)

两个本原多项式之积仍是本原多项式.

证明 设

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$$

是两个本原多项式. 若

$$f(x)g(x) = c_{m+n}x^{m+n} + c_{m+n-1}x^{m+n-1} + \cdots + c_1 x + c_0$$

不是本原多项式, 则 $c_0, c_1, \cdots, c_{m+n}$ 必有一个公约素因子 p. 因为 $f(x)$ 是本原多项式, 故 p 不能整除 $f(x)$ 的所有系数, 可设 $p \mid a_0, p \mid a_1, \cdots, p \mid a_{i-1}$, 但 p 不能整除 a_i. 同理, 可设 $p \mid b_0, p \mid b_1, \cdots, p \mid b_{j-1}$, 但 p 不能整除 b_j. 注意到

$$c_{i+j} = \cdots + a_{i-2}b_{j+2} + a_{i-1}b_{j+1} + a_i b_j + a_{i+1}b_{j-1} + \cdots,$$

p 可整除 c_{i+j},p 也能整除右式除 $a_i b_j$ 以外的所有项. 但 p 不能整除 a_i 和 b_j, 故 p 不能整除 $a_i b_j$, 引出矛盾. □

定理 5.17

若整系数多项式 $f(x)$ 在有理数域上可约, 则它必可分解为两个次数较低的整系数多项式之积.

证明 假设整系数多项式 $f(x)$ 可以分解为两个次数较低的有理系数多项式之积:

$$f(x) = g(x)h(x),$$

$g(x)$ 的各项系数为有理数, 必有一个公分母记为 c, 于是 $g(x) = \frac{1}{c}(cg(x))$, 其中 $cg(x)$ 为整系数多项式. 若把 $cg(x)$ 中所有系数的最大公因数 d 提出来, 则

$$g(x) = \frac{d}{c}\left(\frac{c}{d}g(x)\right),$$

$\frac{c}{d}g(x)$ 是一个本原多项式. 这表明 $g(x) = ag_1(x)$,a 为有理数,$g_1(x)$ 为本原多项式. 同理,$h(x) = bh_1(x)$, 其中 b 为有理数,$h_1(x)$ 为本原多项式. 于是我们得到

$$f(x) = g(x)h(x) = abg_1(x)h_1(x).$$

由Gauss 引理知,$g_1(x)h_1(x)$ 是本原多项式. 若 ab 不是一个整数, 则 $abg_1(x)h_1(x)$ 将不是整系数多项式, 这与 $f(x)$ 是整系数多项式相矛盾. 因此 ab 必须是整数, 于是 $f(x)$ 可以分解为两个次数较小的整系数多项式之积. □

定义 5.10 (整系数多项式在整数环上可约)

我们通常称一个整系数多项式 $f(x)$ **在整数环上可约**, 若它可以分解为两个次数较低的整系数多项式之积.

命题 5.23

整系数多项式 $f(x)$ 若在整数环上不可约, 则在有理数域上也不可约.

证明 由定理 5.17即得. □

例题 5.20 $f(x)$ 是次数大于零的首一整系数多项式, 若 $f(0), f(1)$ 都是奇数, 求证:$f(x)$ 没有有理根.

证明 若 c 是偶数, 则上述左边为奇数, 不可能等于零. 若 c 是奇数, 令 $c = 2b + 1$, 其中 b 是整数, 可得

$$(2b+1)^n + a_{n-1}(2b+1)^{n-1} + \cdots + a_1(2b+1) + a_0 = 0.$$

用二项式定理展开后将看到, 上式左边是一个偶数加上 $1 + a_{n-1} + \cdots + a_1 + a_0$, 故必是奇数, 也不可能等于零. 因此 $f(x)$ 没有有理根. □

命题 5.24

设 $f(x)$ 是实系数多项式, 若对任意的有理数 c,$f(c)$ 总是有理数, 求证:$f(x)$ 是有理系数多项式.

注 证明与命题 5.20

证明 设 $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, 分别令 $x = 0, 1, 2, \cdots, n$, 得到一个以 $a_n, a_{n-1}, \cdots, a_1, a_0$ 为未知数, 由 $n+1$ 个方程组成的实系数线性方程组. 该方程组的系数行列式是一个非零的 Vandermonde 行列式, 故方

程组必有唯一解,且解为有理数. 因此 $f(x)$ 是有理系数多项式. □

例题 5.21 设 $f(x)$ 是有理系数多项式,a,b,c 是有理数, 但 $\sqrt{c}$ 是无理数. 求证: 若 $a+b\sqrt{c}$ 是 $f(x)$ 的根, 则 $a-b\sqrt{c}$ 也是 $f(x)$ 的根.

证明 设 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$, 则

$$f(a+b\sqrt{c})=a_n(a+b\sqrt{c})^n+a_{n-1}(a+b\sqrt{c})^{n-1}+\cdots+a_1(a+b\sqrt{c})+a_0=0.$$

将 $(a+b\sqrt{c})^k$ 用二项式定理展开, 可设

$$f(a+b\sqrt{c})=A+B\sqrt{c}=0,$$

其中 A,B 都是有理数. 因为 $\sqrt{c}$ 是无理数, 故 $A=B=0$. 因此

$$f(a-b\sqrt{c})=A-B\sqrt{c}=0,$$

即 $a-b\sqrt{c}$ 也是 $f(x)$ 的根. □

例题 5.22 设 $f(x)$ 是有理系数多项式,a,b,c,d 是有理数, 但 $\sqrt{c},\sqrt{d},\sqrt{cd}$ 都是无理数. 求证: 若 $a\sqrt{c}+b\sqrt{d}$ 是 $f(x)$ 的根, 则下列数也是 $f(x)$ 的根:

$$a\sqrt{c}-b\sqrt{d},-a\sqrt{c}+b\sqrt{d},-a\sqrt{c}-b\sqrt{d}.$$

证明 令

$$g(x)=(x-(a\sqrt{c}+b\sqrt{d}))(x-(a\sqrt{c}-b\sqrt{d}))(x-(-a\sqrt{c}+b\sqrt{d}))(x-(-a\sqrt{c}-b\sqrt{d})),$$

则经计算可得

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

注意到 $g(x)$ 是一个有理数首一多项式,只要证明它不可约,便可由极小多项式的充要条件得到 $g(x)$ 是 $a\sqrt{c}+b\sqrt{d}$ 的极小多项式,从而由极小多项式的基本性质可知 $g(x)\mid f(x)$, 于是结论成立. 显然 $g(x)$ 没有有理系数的一次因式, 只要证明它没有有理系数的二次因式即可. 经过简单的计算可知, 在 $g(x)$ 的一个一次因式中任取一个一次因式相乘都不是有理系数多项式, 因此 $g(x)$ 没有有理系数的二次因式. □

例题 5.23 求以 $\sqrt{2}+\sqrt[3]{3}$ 为根的次数最小的首一有理系数多项式.

注 确定 f(x) 的 6 个根的方法: 原方程 $x-\sqrt{2}=\sqrt[3]{3}$ 的解为 $x=\sqrt{2}+\sqrt[3]{3}$. 但三次方程 $y^3=3$ 的所有根为 $y=\sqrt[3]{3},\sqrt[3]{3}\omega,\sqrt[3]{3\omega^2}$（其中 $\omega=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$ 是三次单位根）, 因此原方程对应三个解:

$$x=\sqrt{2}+\sqrt[3]{3},\quad \sqrt{2}+\sqrt[3]{3}\omega,\quad \sqrt{2}+\sqrt[3]{3\omega^2}.$$

在消去 $\sqrt{2}$ 的平方步骤中, 方程 $x^3+6x-3=(3x^2+2)\sqrt{2}$ 的两边平方后, 原方程中的 $\sqrt{2}$ 可以被替换为 $-\sqrt{2}$, 从而产生另一组解:

$$x=-\sqrt{2}+\sqrt[3]{3},\quad -\sqrt{2}+\sqrt[3]{3}\omega,\quad -\sqrt{2}+\sqrt[3]{3\omega^2}.$$

解 本题即求 $\sqrt{2}+\sqrt[3]{3}$ 的极小多项式. 令 $x-\sqrt{2}=\sqrt[3]{3}$, 两边立方得到 $(x-\sqrt{2})^3=3$. 整理可得 $x^3+6x-3=(3x^2+2)\sqrt{2}$, 再两边平方可得,$\sqrt{2}+\sqrt[3]{3}$ 适合下列多项式:

$$f(x)=x^6-6x^4-6x^3+12x^2-36x+1.$$

由 $f(x)$ 的构造过程,不难看出 $f(x)$ 的 6 个根分别为 $\pm\sqrt{2}+\sqrt[3]{3},\pm\sqrt{2}+\sqrt[3]{3}\omega,\pm\sqrt{2}+\sqrt[3]{3}\omega^2$. 其中 $\omega=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$. 因此, 我们有

$$f(x)=\left(x-\sqrt{2}-\sqrt[3]{3}\right)\left(x+\sqrt{2}-\sqrt[3]{3}\right)\left(x-\sqrt{2}-\sqrt[3]{3}\omega\right)\left(x+\sqrt{2}-\sqrt[3]{3}\omega\right)\left(x-\sqrt{2}-\sqrt[3]{3}\omega^2\right)\left(x+\sqrt{2}-\sqrt[3]{3}\omega^2\right).$$

通过简单的验证可知, 任取 $f(x)$ 的 2 个一次因式相乘都不是有理系数多项式; 任取 $f(x)$ 的 3 个一次因式相乘也都不是有理系数多项式, 因此 $f(x)$ 是有理数域上的不可约多项式, 从而由极小多项式式的充要条件可知,$f(x)$ 是 $\sqrt{2}+\sqrt[3]{3}$ 的极小多项式. □

例题 5.24 求证: 有理系数多项式 x^4+px^2+q 在有理数域上可约的充要条件是或者 $p^2-4q=k^2$, 其中 k 是一个有

理数; 或者 q 是某个有理数的平方, 且 $\pm 2\sqrt{q}-p$ 也是有理数的平方.

证明 必要性: 若多项式 x^4+px^2+q 在有理数域上可约, 考虑下列两种情况:

(1) x^4+px^2+q 有有理数根 t, 这时 t^2 是 x^2+px+q 的有理根, 因此其判别式 p^2-4q 必是一个有理数的完全平方.

(2) x^4+px^2+q 无有理数根, 则 x^4+px^2+q 在有理数域上可分解为两个二次多项式的积. 设 $x^4+px^2+q=(x^2+ax+b)(x^2+cx+d)$, 展开后比较系数可得

$$\begin{cases} a+c=0, \\ ad+bc=0. \end{cases}$$

若 $a=0$, 则 $c=0$, 这时将有 $p=b+d$, $q=bd$, 因此 $p^2-4q=(b-d)^2$. 若 $a\neq 0$, 则 $b=d$, 比较系数后可知 $p=2b-a^2$, $q=b^2$, 因此 $\pm 2\sqrt{q}-p=a^2$.

充分性: 若 $p^2-4q=k^2$, 则

$$x^4+px^2+q=x^4+px^2+\frac{1}{4}(p+k)(p-k)=\left(x^2+\frac{1}{2}(p+k)\right)\left(x^2+\frac{1}{2}(p-k)\right).$$

因此多项式可约.

若 $q=b^2$, $\pm 2\sqrt{q}-p=\pm 2b-p=a^2$, 则 $p=-a^2\pm 2b$. 于是

$$x^4+px^2+q=x^4+(-a^2\pm 2b)x^2+b^2=(x^2\pm b)^2-a^2x^2$$

也可约. □

定理 5.18 (Eisenstein 判别法)

设多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

是整系数多项式,$a_n\neq 0, n\geq 1, p$ 是一个素数. 若 $p\mid a_i(i=0,1,\cdots,n-1)$, 但 $p\nmid a_n$ 且 $p^2\nmid a_0$, 则 $f(x)$ 在有理数域上不可约.

证明 只需证明 $f(x)$ 在整数环上不可约即可. 设 $f(x)$ 可分解为两个次数较低的整系数多项式之积:

$$f(x)=(b_mx^m+b_{m-1}x^{m-1}+\cdots+b_0)(c_tx^t+c_{t-1}x^{t-1}+\cdots+c_0),$$

其中 $m+t=n$. 显然 $a_0=b_0c_0, a_n=b_mc_t$. 由假设 $p\mid a_0$, 故 $p\mid b_0$ 或 $p\mid c_0$. 又 $p^2\nmid a_0$, 故 p 不能同时整除 b_0 及 c_0. 不妨设 $p\mid b_0$ 但 $p\nmid c_0$. 又由假设,p 不能整除 $a_n=b_mc_t$, 故 p 既不能整除 b_m 又不能整除 c_t. 因此不妨设 $p\mid b_0, p\mid b_1,\cdots,p\mid b_{j-1}$ 但 p 不能整除 b_j, 其中 $0<j\leq m<n$. 而

$$a_j=b_jc_0+b_{j-1}c_1+\cdots+b_0c_j,$$

根据假设,$p\mid a_j$, 又 p 可整除上述右端除 b_jc_0 外的其余项, 而不能整除 b_jc_0 这一项, 引出矛盾. □

例题 5.25 设 $p_1,\cdots,p_m$ 是 m 个互不相同的素数, 求证: 对任意的 $n\geq 1$, 下列多项式在有理数域上不可约:

$$f(x)=x^n-p_1\cdots p_m.$$

证明 用 Eisenstein 判别法即可证明.(取 $p=p_i$ 即可) □

例题 5.26 证明:x^8+1 在有理数域上不可约.

证明 作代换 $x=y+1$, 得

$$x^8+1=(y+1)^8+1=y^8+8y^7+28y^6+56y^5+70y^4+56y^3+28y^2+8y+2.$$

显然 2 可整除除第一项外的所有系数, 但 4 不能整除常数项. 用 Eisenstein 判别法可知 $(y+1)^8+1$ 不可约, 故 x^8+1 也不可约. □

例题 5.27 设 $f(x)$ 是有理系数多项式, 已知 $\sqrt{2}$ 是 $f(x)$ 的根, 证明:$\sqrt{2}\varepsilon,\sqrt{2}\varepsilon^2,\cdots,\sqrt{2}\varepsilon^{n-1}$ 也是 $f(x)$ 的根, 其中

$\varepsilon = \cos\dfrac{2\pi}{n} + i\sin\dfrac{2\pi}{n}$ 是 1 的 n 次根.

证明 显然 $\sqrt{2}$ 适合多项式 $x^n - 2$, 由 Eisenstein 判别法可知,$x^n - 2$ 在有理数域上不可约, 因此它是 $\sqrt{2}$ 的极小多项式. 最后由极小多项式的基本性质可得 $(x^n - 2) \mid f(x)$, 从而结论得证. □

例题 5.28 设 $f(x)$ 是次数大于 1 的奇数次有理系数不可约多项式, 求证: 若 x_1, x_2 是 $f(x)$ 在复数域内两个不同的根, 则 $x_1 + x_2$ 必不是有理数.

证明 不妨设 $f(x)$ 为首一多项式, 我们用反证法来证明结论. 设 $x_1 + x_2 = r$ 为有理数, 则有理系数多项式 $f(x)$ 与 $f(r-x)$ 有公共根 x_1. 因为 $f(x)$ 在有理数域上不可约, 故 $f(x)$ 是 x_1 的极小多项式, 从而由极小多项式的基本性质可得 $f(x) \mid f(r-x)$. 注意到 $f(x)$ 与 $f(r-x)$ 次数相同, 首项系数相同, 从而有 $f(r-x) = -f(x)$. 令 $x = \dfrac{r}{2}$, 则可得 $f\left(\dfrac{r}{2}\right) = 0$, 即 $\dfrac{r}{2}$ 是 $f(x)$ 的一个有理根, 这与 $f(x)$ 在有理数域上不可约相矛盾. □

例题 5.29 设 $f(x) = (x-a_1)(x-a_2)\cdots(x-a_n) - 1$, 其中 $a_1, a_2, \cdots, a_n$ 是 n 个不同的整数, 求证:$f(x)$ 在有理数域上不可约.

证明 由命题 5.23可知, 只要证明 $f(x)$ 在整数环上不可约即可. 用反证法, 设 $f(x) = g(x)h(x)$, 其中 $g(x), h(x)$ 都是次数小于 n 的首一整数系数多项式. 注意到

$$g(a_i)h(a_i) = -1,$$

因为 $g(x), h(x)$ 是整数系数多项式, 故 $g(a_i) = 1, h(a_i) = -1$ 或 $g(a_i) = -1, h(a_i) = 1$. 无论是哪种情况, 都有

$$g(a_i) + h(a_i) = 0, \quad 1 \le i \le n,$$

即次数小于 n 的多项式 $g(x) + h(x)$ 有 n 个不同的根, 故 $g(x) + h(x) = 0$. 因此 $f(x) = -g(x)^2$, 但 $f(x)$ 是首一多项式, 而 $-g(x)^2$ 的首项系数为 -1, 矛盾. □

例题 5.30 设 $f(x) = (x-a_1)^2(x-a_2)^2\cdots(x-a_n)^2 + 1$, 其中 $a_1, a_2, \cdots, a_n$ 是 n 个不同的整数, 求证:$f(x)$ 在有理数域上不可约.

证明 由命题 5.23可知, 只要证明 $f(x)$ 在整数环上不可约即可. 用反证法, 设 $f(x) = u(x)v(x)$, 其中 $u(x), v(x)$ 都是次数小于 $2n$ 的首一整数系数多项式. 注意到 $f(x)$ 没有实根, 故 $u(x), v(x)$ 也都没有实根, 从而由实系数多项式虚根成对可知,$u(x), v(x)$ 作为实数域上的函数都恒大于零. 由于 $f(x)$ 是 $2n$ 次多项式, 故 $u(x)$ 和 $v(x)$ 的次数至少有一个不超过 n, 不妨设 $u(x)$ 的次数不超过 n.

若 $u(x)$ 的次数小于 n, 则由 $f(a_i) = 1$ 可得 $u(a_i)v(a_i) = 1$, 因此 $u(a_i) = 1$. 考虑非零多项式 $u(x) - 1$, 由上面的分析可知它有 n 个不同的根 $a_1, a_2, \cdots, a_n$, 这与它的次数小于 n 矛盾.

因此 $u(x)$ 只能是 n 次首一多项式, 于是 $v(x)$ 也是 n 次首一多项式. 另一方面, 由于 $u(a_i)v(a_i) = 1$, 故 $u(a_i) = v(a_i) = \pm 1 (1 \le i \le n)$. 注意到 $u(x) - v(x)$ 的次数小于 n 并且它有 n 个不同的根 $a_1, a_2, \cdots, a_n$, 因此 $u(x) = v(x)$ 或 $u(x) = -v(x)$. 今设 $u(x) = v(x)$, 则 $f(x) = u(x)^2 + 1$, 即

$$(u(x) + h(x))(u(x) - h(x)) = 1.$$

因为 $u(x), h(x)$ 都是整数系数多项式, 故或者 $u(x) + h(x) = 1, u(x) - h(x) = 1$; 或者 $u(x) + h(x) = -1, u(x) - h(x) = -1$, 于是作差可得 $h(x) = 0$, 矛盾. 因此结论得证. □

5.9 多元多项式

因式分解定理对多元多项式仍成立.(证明见抽象代数内容)

引理 5.3

若 $f(x_1, x_2, \cdots, x_n)$ 及 $g(x_1, x_2, \cdots, x_n)$ 都是 K 上非零的 n 元多项式, 则按字典排列法排列后乘积的首项等于 f 的首项与 g 的首项之积.

证明 设 $ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$ 和 $bx_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$ 分别是 f 和 g 的首项(按字典排列法), 它们的乘积为 $abx_1^{i_1+j_1}x_2^{i_2+j_2}\cdots x_n^{i_n+j_n}$. 其他任意两个单项式 $cx_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 和 $dx_1^{r_1}x_2^{r_2}\cdots x_n^{r_n}$ 之积为 $cdx_1^{k_1+r_1}x_2^{k_2+r_2}\cdots x_n^{k_n+r_n}$. 设 $i_1 = k_1, \cdots, i_{t-1} = k_{t-1}, i_t >$

$k_t; j_1 = r_1, \cdots, j_{s-1} = r_{s-1}, j_s > r_s$. 不妨设 $t \le s$, 显然

$$i_1 + j_1 = k_1 + r_1, \cdots, i_{t-1} + j_{t-1} = k_{t-1} + r_{t-1}, i_t + j_t > k_t + r_t.$$

因此 $abx_1^{i_1+j_1} x_2^{i_2+j_2} \cdots x_n^{i_n+j_n}$ 先于 $cdx_1^{k_1+r_1} x_2^{k_2+r_2} \cdots x_n^{k_n+r_n}$.

同理可证明:$abx_1^{i_1+j_1} x_2^{i_2+j_2} \cdots x_n^{i_n+j_n}$ 先于 $adx_1^{i_1+r_1} x_2^{i_2+r_2} \cdots x_n^{i_n+r_n}$ 和 $cbx_1^{k_1+j_1} x_2^{k_2+j_2} \cdots x_n^{k_n+j_n}$. 因此它确是 fg 的首项. □

命题 5.25 (多元多项式的整性)

若 $f(x_1, x_2, \cdots, x_n) \neq 0, g(x_1, x_2, \cdots, x_n) \neq 0$, 则

$$f(x_1, x_2, \cdots, x_n) g(x_1, x_2, \cdots, x_n) \neq 0.$$

证明 由 f 和 g 的首项不为零及引理 5.3可知 fg 的首项不为零, 于是 $fg \neq 0$. □

推论 5.11

若 $h(x_1, x_2, \cdots, x_n) \neq 0$, 且

$$f(x_1, x_2, \cdots, x_n) h(x_1, x_2, \cdots, x_n) = g(x_1, x_2, \cdots, x_n) h(x_1, x_2, \cdots, x_n),$$

则

$$f(x_1, x_2, \cdots, x_n) = g(x_1, x_2, \cdots, x_n).$$

证明 由条件可得

$$\left[f(x_1, x_2, \cdots, x_n) - g(x_1, x_2, \cdots, x_n)\right] h(x_1, x_2, \cdots, x_n) = 0.$$

又因为 $h(x_1, x_2, \cdots, x_n) \neq 0, f(x_1, x_2, \cdots, x_n) - g(x_1, x_2, \cdots, x_n) \neq 0$, 则由命题 5.25 可知

$$\left[f(x_1, x_2, \cdots, x_n) - g(x_1, x_2, \cdots, x_n)\right] h(x_1, x_2, \cdots, x_n) \neq 0$$

矛盾! 故 $f(x_1, x_2, \cdots, x_n) - g(x_1, x_2, \cdots, x_n) = 0$, 即

$$f(x_1, x_2, \cdots, x_n) = g(x_1, x_2, \cdots, x_n).$$

□

定义 5.11 (齐次多项式)

若一个多项式 $f(x_1, x_2, \cdots, x_n)$ 的每个单项式都是 k 次式, 则**称之为 k 次齐次多项式或 k 次型**.

命题 5.26 (齐次多项式的基本性质)

(1) 两个次数相同的齐次多项式之和若不为零, 则必仍是同次齐次多项式. 任意两个齐次多项式之积仍为齐次多项式.

(2) 任一 n 元多项式均可表示为若干个齐次多项式之和,

证明

(1) 显然.

(2) 这只需要将各次数相等的项放在一起即可.

□

引理 5.4 (多元多项式的非零性)

设 $f(x_1, x_2, \cdots, x_n)$ 是 K 上非零的 n 元多项式, 则必存在 K 中的数 $a_1, a_2, \cdots, a_n$, 使 $f(a_1, a_2, \cdots, a_n) \neq 0$.

证明 对未定元个数 n 用数学归纳法. 当 $n = 1$ 时, 多项式 $f(x)$ 只有有限个零点, 故总有 $a \in K$ 使 $f(a) \neq 0$. 现设对

有 $n-1$ 个未定元的多项式结论成立, 将 $f(x_1,x_2,\cdots,x_n)$ 写成未定元 x_n 的多项式:

$$f(x_1,x_2,\cdots,x_n)=b_0+b_1x_n+\cdots+b_mx_n^m,$$

其中 $b_i=b_i(x_1,x_2,\cdots,x_{n-1})$ 是 $n-1$ 元多项式. 因为 $f(x_1,x_2,\cdots,x_n)\neq 0$, 故可设 $b_m\neq 0$. 由归纳假设, 存在 $a_1,\cdots,a_{n-1}\in K$, 使

$$b_m(a_1,\cdots,a_{n-1})\neq 0.$$

因而

$$f(a_1,\cdots,a_{n-1},x_n)=b_0(a_1,\cdots,a_{n-1})+b_1(a_1,\cdots,a_{n-1})x_n+\cdots+b_m(a_1,\cdots,a_{n-1})x_n^m$$

是一个非零的以 x_n 为未定元的一元多项式, 故存在 $a_n\in K$, 使

$$f(a_1,a_2,\cdots,a_n)\neq 0.$$

□

命题 5.27 (多元多项式相等的充要条件)

数域 K 上的两个 n 元多项式 $f(x_1,x_2,\cdots,x_n)$ 与 $g(x_1,x_2,\cdots,x_n)$ 相等的充分必要条件是对一切 $a_1,a_2,\cdots,a_n\in K$, 均有

$$f(a_1,a_2,\cdots,a_n)=g(a_1,a_2,\cdots,a_n).$$

证明 只需证明充分性. 作

$$h(x_1,x_2,\cdots,x_n)=f(x_1,x_2,\cdots,x_n)-g(x_1,x_2,\cdots,x_n).$$

若 $h(x_1,x_2,\cdots,x_n)\neq 0$, 则由多元多项式的非零性非零可知必有 $a_1,a_2,\cdots,a_n\in K$, 使

$$h(a_1,a_2,\cdots,a_n)\neq 0,$$

这与假设矛盾. □

例题 5.31 设 $f(x_1,\cdots,x_n),g(x_1,\cdots,x_n)\neq 0$ 是 K 上的多元多项式. 假设对一切使 $g(a_1,\cdots,a_n)\neq 0$ 的 $a_1,\cdots,a_n\in K$, 均有 $f(a_1,\cdots,a_n)=0$, 求证:

$$f(x_1,\cdots,x_n)=0.$$

证明 用反证法, 假设 $f(x_1,\cdots,x_n)\neq 0$, 则由多元多项式的整性可知

$$h(x_1,\cdots,x_n)=f(x_1,\cdots,x_n)g(x_1,\cdots,x_n)\neq 0,$$

于是由元多项式的非零性存在 $a_1,\cdots,a_n\in K$, 使得 $h(a_1,\cdots,a_n)\neq 0$, 从而 $f(a_1,\cdots,a_n)\neq 0$ 并且 $g(a_1,\cdots,a_n)\neq 0$, 这与条件矛盾. □

命题 5.28

设 $A(x_1,x_2,\cdots,x_m)=(a_{ij})$ 为 n 阶方阵, 其元素 $a_{ij}=a_{ij}(x_1,x_2,\cdots,x_m)$ 都是 $\mathbb{K}$ 上的多元多项式. 设 $g(x_1,x_2,\cdots,x_m),h_i(x_1,x_2,\cdots,x_m)\neq 0\ (1\leq i\leq k)$ 都是 $\mathbb{K}$ 上的多元多项式,

$$U=\{(a_1,a_2,\cdots,a_m)\in\mathbb{K}^m\mid h_i(a_1,a_2,\cdots,a_m)\neq 0\ (1\leq i\leq k)\}.$$

若对所有的 $(a_1,a_2,\cdots,a_m)\in U$, 都成立

$$|A(a_1,a_2,\cdots,a_m)|=g(a_1,a_2,\cdots,a_m),$$

证明:$|A(x_1,x_2,\cdots,x_m)|=g(x_1,x_2,\cdots,x_m)$.

笔记 这个命题告诉我们: 在元素为多元多项式的文字行列式的求值过程中, 在假设某些非零条件下成立的情形下得到的结果, 其实就是所求行列式的值. 因此, 在求行列式的过程中, **可以暂不考虑未定元取特殊值的情形, 而把主要精力放在一般的情形进行计算即可**.

证明 用反证法, 设 $(|A|-g)(x_1,x_2,\cdots,x_m)\neq 0$, 则由多元多项式的整性可知

$$(|A|-g)h_1\cdots h_k(x_1,x_2,\cdots,x_m)\neq 0,$$

于是存在 $a_1,a_2,\cdots,a_m\in\mathbb{K}$, 使得 $(|A|-g)h_1\cdots h_k(a_1,a_2,\cdots,a_m)\neq 0$, 从而 $h_i(a_1,a_2,\cdots,a_m)\neq 0\ (1\leq i\leq k)$, 即 $(a_1,a_2,\cdots,a_m)\in U$, 并且 $(|A|-g)(a_1,a_2,\cdots,a_m)\neq 0$, 即 $|A(a_1,a_2,\cdots,a_m)|\neq g(a_1,a_2,\cdots,a_m)$, 这与条件矛盾. □

定理 5.19 (行列式的求根法)

设 $A(x_1,x_2,\cdots,x_m)=(a_{ij})$ 为 n 阶方阵, 其元素 $a_{ij}=a_{ij}(x_1,\cdots,x_m)$ 都是 $\mathbb{K}$ 上的多元多项式, 于是 $|A|$ 也是 $\mathbb{K}$ 上的多元多项式. 若把 x_1 看成未定元, 则可将 $|A|$ 整理成关于 x_1 的一元多项式:

$$|A|=c_0(x_2,\cdots,x_m)x_1^d+c_1(x_2,\cdots,x_m)x_1^{d-1}+\cdots+c_d(x_2,\cdots,x_m),\tag{5.14}$$

其中 $c_0(x_2,\cdots,x_m)\neq 0, d\geq 1$ 为次数. 假设存在互异的多项式 $g_1(x_2,\cdots,x_m),\cdots,g_d(x_2,\cdots,x_m)$, 使得当 $x_1=g_i(x_2,\cdots,x_m)(1\leq i\leq d)$ 时 $|A|=0$, 证明:

$$|A|=c_0(x_2,\cdots,x_m)(x_1-g_1(x_2,\cdots,x_m))\cdots(x_1-g_d(x_2,\cdots,x_m)).$$

证明 由假设 $0=c_0(x_2,\cdots,x_m)g_1^d+\cdots+c_{d-1}(x_2,\cdots,x_m)g_1+c_d(x_2,\cdots,x_m)$, 故有

$$|A|=c_0(x_2,\cdots,x_m)(x_1^d-g_1^d)+\cdots+c_{d-1}(x_2,\cdots,x_m)(x_1-g_1)R_1(x_1,x_2,\cdots,x_m).$$

在上述式中令 $x_1=g_2(x_2,\cdots,x_m)$, 则有 $0=(g_2-g_1)R_1(g_2,x_2,\cdots,x_m)$. 注意到 $g_2-g_1\neq 0$, 故由多元多项式的整性可得 $R_1(g_2,x_2,\cdots,x_m)=0$. 再由相同的讨论可知,$x_1-g_2$ 是 $R_1(x_1,x_2,\cdots,x_m)$ 的因式, 从而

$$|A|=(x_1-g_1)(x_1-g_2)R_2(x_1,x_2,\cdots,x_m).$$

不断地这样做下去, 可得

$$|A|=(x_1-g_1)\cdots(x_1-g_d)R_d(x_1,x_2,\cdots,x_m).$$

最后与(5.14)式比较 x_1 的首项系数可得 $R_d(x_1,x_2,\cdots,x_m)=c_0(x_2,\cdots,x_m)$. □

5.10 互素多项式的应用

命题 5.29

设 $f(x),g(x)$ 是数域 $\mathbb{K}$ 上的互素多项式,A 是 $\mathbb{K}$ 上的 n 阶方阵, 满足 $f(A)=O$, 证明:$g(A)$ 是可逆矩阵.

证明 根据假设, 存在 $\mathbb{K}$ 上的多项式 $u(x),v(x)$, 使得

$$f(x)u(x)+g(x)v(x)=1.$$

在上述式中代入 $x=A$, 可得恒等式

$$f(A)u(A)+g(A)v(A)=I_n.$$

因为 $f(A)=O$, 故有 $g(A)v(A)=I_n$, 从而 $g(A)$ 是非零矩阵且 $g(A)^{-1}=v(A)$. □

命题 5.30

设 $f(x),g(x)$ 是数域 $\mathbb{K}$ 上的互素多项式,A 是 $\mathbb{K}$ 上的 n 阶方阵, 证明:$f(A)g(A)=O$ 的充要条件是 $r(f(A))+r(g(A))=n$.

证明 根据假设, 存在 $\mathbb{K}$ 上的多项式 $u(x),v(x)$, 使得

$$f(x)u(x)+g(x)v(x)=1.$$

在上述式中代入 $x=A$, 可得恒等式

$$f(A)u(A)+g(A)v(A)=I_n.$$

考虑如下分块矩阵的初等变换:

$$\begin{pmatrix} f(A) & O \\ O & g(A) \end{pmatrix} \to \begin{pmatrix} f(A) & f(A)u(A) \\ O & g(A) \end{pmatrix} \to \begin{pmatrix} f(A) & I_n \\ O & g(A) \end{pmatrix} \to \begin{pmatrix} f(A) & I_n \\ -f(A)g(A) & O \end{pmatrix} \to \begin{pmatrix} O & I_n \\ -f(A)g(A) & O \end{pmatrix},$$

故有 $r(f(A))+r(g(A))=r(f(A)g(A))+n$, 从而结论得证. □

命题 5.31

设 $f(x),g(x)$ 是数域 $\mathbb{K}$ 上的互素多项式,φ 是 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换, 满足 $f(\varphi)g(\varphi)=0$, 证明:$V=V_1\oplus V_2$, 其中 $V_1=\mathrm{Ker}f(\varphi),V_2=\mathrm{Ker}g(\varphi)$.

笔记 这个命题告诉我们: 多项式的互素因式分解可以诱导出空间的直和分解, 从几何层面上看, 这就是相似标准型理论原始的除法点.

证明 根据假设, 存在 $\mathbb{K}$ 上的多项式 $u(x),v(x)$, 使得

$$f(x)u(x)+g(x)v(x)=1.$$

在上述式中代入 $x=\varphi$, 可得恒等式

$$f(\varphi)u(\varphi)+g(\varphi)v(\varphi)=I_V.$$

对任意的 $\alpha\in V$, 由上述可得

$$\alpha=f(\varphi)u(\varphi)(\alpha)+g(\varphi)v(\varphi)(\alpha),$$

注意到

$$g(\varphi)(f(\varphi)u(\varphi)(\alpha))=g(\varphi)f(\varphi)u(\varphi)(\alpha)=u(\varphi)f(\varphi)g(\varphi)(\alpha)=0,$$
$$f(\varphi)(g(\varphi)u(\varphi)(\alpha))=f(\varphi)g(\varphi)u(\varphi)(\alpha)=u(\varphi)f(\varphi)g(\varphi)(\alpha)=0.$$

于是 $f(\varphi)u(\varphi)(\alpha)\in\mathrm{Ker}g(\varphi),g(\varphi)v(\varphi)(\alpha)\in\mathrm{Ker}f(\varphi)$, 故有 $V=V_1+V_2$. 任取 $\beta\in V_1\cap V_2$, 由上述可得

$$\beta=u(\varphi)f(\varphi)(\beta)+v(\varphi)g(\varphi)(\beta)=0,$$

故有 $V_1\cap V_2=0$, 因此 $V=V_1\oplus V_2$. □

例题 5.32 设 $\mathbb{Q}(\sqrt[n]{2})=\{a_0+a_1\sqrt[n]{2}+a_2\sqrt[n]{4}+\cdots+a_{n-1}\sqrt[n]{2^{n-1}}\mid a_i\in\mathbb{Q},0\le i\le n-1\}$, 证明:$\mathbb{Q}(\sqrt[n]{2})$ 是一个数域, 并求 $\mathbb{Q}(\sqrt[n]{2})$ 作为 $\mathbb{Q}$ 上线性空间的一组基.

证明 设 $f(x)=x^n-2$, 由 Eisenstein 判别法可知 $f(x)$ 在 $\mathbb{Q}$ 上不可约, 从而 $f(x)$ 是 $\sqrt[n]{2}$ 的极小多项式. 我们先证明:$a_0+a_1\sqrt[n]{2}+\cdots+a_{n-1}\sqrt[n]{2^{n-1}}=0$ 的充要条件是 $a_0=a_1=\cdots=a_{n-1}=0$. 充分性是显然的, 现证必要性: 令 $g(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$, 则 $g(\sqrt[n]{2})=0$, 由极小多项式的基本性质可得 $f(x)\mid g(x)$. 因为 $g(x)$ 的次数小于 n, 故只能是 $g(x)=0$, 即 $a_0=a_1=\cdots=a_{n-1}=0$.

利用 $\sqrt[n]{2}^n=2$ 容易验证,$\mathbb{Q}(\sqrt[n]{2})$ 中任意两个数的加法、减法和乘法都是封闭的. 要证明 $\mathbb{Q}(\sqrt[n]{2})$ 是数域, 只要证明除法或者取倒数封闭即可. 任取 $\mathbb{Q}(\sqrt[n]{2})$ 中的非零数 $\alpha=a_0+a_1\sqrt[n]{2}+\cdots+a_{n-1}\sqrt[n]{2^{n-1}}\neq0$, 由上面的讨论可知 $a_0,a_1,\cdots,a_{n-1}$ 不全为零. 令 $g(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$, 则 $g(\sqrt[n]{2})\neq0$. 因为 $f(x)$ 不可约且 $g(x)\neq0$ 的次数小于 n, 故 $f(x)$ 与 $g(x)$ 互素, 由多项式互素的充要条件可知, 存在有理系数多项式 $u(x),v(x)$, 使得

$$f(x)u(x)+g(x)v(x)=1,$$

在上述中代入 $x=\sqrt[n]{2}$, 可得 $\sqrt[n]{2}v(\sqrt[n]{2})=1$, 于是 $\alpha^{-1}=v(\sqrt[n]{2})\in\mathbb{Q}(\sqrt[n]{2})$. 因此,$\mathbb{Q}(\sqrt[n]{2})$ 是数域.

由 $\mathbb{Q}(\sqrt[n]{2})$ 的定义可知,$\mathbb{Q}(\sqrt[n]{2})$ 中任一元都是 $1,\sqrt[n]{2},\cdots,\sqrt[n]{2^{n-1}}$ 的 $\mathbb{Q}$-线性组合; 又由开始的讨论可知,$1,\sqrt[n]{2},\cdots,\sqrt[n]{2^{n-1}}$ 是 $\mathbb{Q}$-线性无关的, 因此它们构成了 $\mathbb{Q}(\sqrt[n]{2})$ 作为 $\mathbb{Q}$ 上线性空间的一组基. 特别地,$\dim_{\mathbb{Q}}\mathbb{Q}(\sqrt[n]{2})=n$. □

命题 5.32

设 $f(x)=x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$ 是数域 $\mathbb{K}$ 上的不可约多项式,φ 是 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换,$\alpha_1\neq 0,\alpha_2,\cdots,\alpha_n$ 是 V 中的向量, 满足

$$\varphi(\alpha_1)=\alpha_2,\varphi(\alpha_2)=\alpha_3,\cdots,\varphi(\alpha_{n-1})=\alpha_n,\varphi(\alpha_n)=-a_n\alpha_1-a_{n-1}\alpha_2-\cdots-a_1\alpha_n.$$

证明:$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 是 V 的一组基.

证明 我们只要证明 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关即可. 用反证法, 设存在不全为零的 n 个数 $c_1,c_2,\cdots,c_n$, 使得

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n=0,$$

则有

$$(c_1I_V+c_2\varphi+\cdots+c_n\varphi^{n-1})(\alpha_1)=0.$$

令 $g(x)=c_1+c_2x+\cdots+c_nx^{n-1}$, 则 $g(x)\neq 0$ 且 $g(\varphi)(\alpha_1)=0$. 另一方面, 由假设容易验证 $f(\varphi)(\alpha_1)=0$. 因为 $f(x)$ 不可约且 $g(x)$ 的次数小于 n, 故 $f(x)$ 与 $g(x)$ 互素, 从而存在 $\mathbb{K}$ 上的多项式 $u(x),v(x)$, 使得

$$f(x)u(x)+g(x)v(x)=1.$$

在上述中代入 $x=\varphi$, 可得恒等式

$$f(\varphi)u(\varphi)+g(\varphi)v(\varphi)=I_V.$$

上式两边同时作用 α_1 可得

$$\alpha_1=u(\varphi)f(\varphi)(\alpha_1)+v(\varphi)g(\varphi)(\alpha_1)=0,$$

这与条件 $\alpha_1\neq 0$ 矛盾, 从而结论得证. □

第六章 特征值

注 代数基本定理保证了任一 $n(n \geq 1)$ 阶复矩阵 A 或 n 维复线性空间 V 上的线性变换 φ 至少有一个复特征值 λ_0, 线性方程组的求解理论保证了 λ_0 至少有一个复特征向量. 如果是在数域 $\mathbb{F}$ 上, 则需要 A 或 φ 的特征值 λ_0 属于 $\mathbb{F}$, 然后线性方程组的求解理论才能保证 λ_0 在 $\mathbb{F}^n$ 或 V 中有对应的特征向量. 因此, **后面如无特殊说明, 总是假设在复数域 $\mathbb{C}$ 上考虑问题**.

6.1 特征值与特征向量

定义 6.1 (线性变换的特征值和特征向量)

设 φ 是数域 $\mathbb{K}$ 上线性空间 V 上的线性变换, 若 $\lambda_0 \in \mathbb{K}, x \in V$ 且 $x \neq 0$, 使

$$\varphi(x) = \lambda_0 x,$$

则称 λ_0 是线性变换 φ 的一个**特征值**, 向量 x 称为 φ 关于特征值 λ_0 的**特征向量**.

♣

笔记 显然 φ 的关于特征值 λ_0 的全体特征向量加上零向量构成 V 的子空间.

定义 6.2 (线性变换的特征子空间)

设 λ_0 是线性空间 V 上的线性变换 φ 的特征值, 令

$$V_{\lambda_0} = \{\alpha \in V \mid \varphi(\alpha) = \lambda_0 \alpha\} = \{\alpha \in V \mid \alpha\text{是}\varphi\text{的属于}\lambda_0\text{的特征向量}\} \cup \{0\},$$

则显然 V_{λ_0} 是 V 的子空间, 称为 φ 的属于特征值 λ_0 的**特征子空间**.

♣

笔记 显然 V_{λ_0} 是 φ 的不变子空间.

定义 6.3 (矩阵的特征值和特征向量)

设 A 是数域 $\mathbb{F}$ 上的 n 阶方阵, 若存在 $\lambda_0 \in \mathbb{F}$ 及 n 维非零列向量 α, 使

$$A\alpha = \lambda_0 \alpha$$

式成立, 则称 λ_0 为矩阵 A 的一个**特征值**,α 为 A 关于特征值 λ_0 的**特征向量**.

♣

定义 6.4 (矩阵的特征子空间)

设 λ_0 是 $\mathbb{F}$ 上的 n 阶矩阵 A 的特征值, 令

$$V_{\lambda_0} = \{x \in \mathbb{F}^n \mid Ax = \lambda_0 x\} = \{x \in \mathbb{F}^n \mid x\text{是}A\text{的属于}\lambda_0\text{的特征向量}\} \cup \{0\},$$

则 V_{λ_0} 是线性方程组 $(\lambda_0 I_n - A)x = 0$ 的解空间, 从而是 $\mathbb{F}^n$ 的子空间, 称为 A 的属于特征值 λ_0 的**特征子空间**.

♣

定义 6.5 (特征多项式)

设 A 是 n 阶方阵, 称 $|\lambda I_n - A|$ 为 A 的**特征多项式**.

♣

定理 6.1 (特征值的和与积)

矩阵 A 的 n 个特征值的和与积分别为

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(A),$$

$$\lambda_1 \lambda_2 \cdots \lambda_n = |A|.$$

♡

证明 设

$$
\begin{aligned}
|\lambda I_n - A| &= \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n \\
&= (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).
\end{aligned}
$$

由 Vieta 定理知 $\lambda_1+\lambda_2+\cdots+\lambda_n = -a_1$,$\lambda_1\lambda_2\cdots\lambda_n = (-1)^n a_n$. 由例题 1.30可知 $a_1 = -(a_{11}+a_{22}+\cdots+a_{nn}) = -\mathrm{tr}(A)$,$a_n = (-1)^n|A|$. 因此矩阵 A 的 n 个特征值的和与积分别为

$$
\begin{aligned}
\lambda_1+\lambda_2+\cdots+\lambda_n &= \mathrm{tr}(A), \\
\lambda_1\lambda_2\cdots\lambda_n &= |A|.
\end{aligned}
$$

□

定义 6.6 (特征多项式)

设 φ 是线性空间 V 上的线性变换,φ 在 V 的某组基下的表示矩阵为 A, 由相似矩阵有相同特征值知 $|\lambda I_n - A|$ 与基或表示矩阵的选取无关, 称 $|\lambda I_n - A|$ 为 φ 的**特征多项式**, 记为 $|\lambda I_V - \varphi|$.

♣

定理 6.2 (复方阵必相似于上三角阵)

任何复方阵必相似于一个上三角阵, 并且对角元素都是其特征值.

♡

注 一般数域 $\mathbb{K}$ 上的矩阵未必相似于上三角阵.

证明 设 A 是 n 阶复方阵, 现对 n 用数学归纳法. 当 $n=1$ 时结论显然成立. 假设对 $n-1$ 阶矩阵结论成立, 现对 n 阶矩阵 A 来证明. 设 λ_1 是 A 的一个特征值, 则存在非零列向量 α_1, 使

$$
A\alpha_1 = \lambda_1\alpha_1.
$$

将 α_1 作为 C_n 的一个基向量, 并扩展为 C_n 的一组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 将这些基向量按照列分块方式拼成矩阵 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 则 P 为 n 阶非异阵, 且

$$
\begin{aligned}
AP &= A(\alpha_1,\alpha_2,\cdots,\alpha_n) = (A\alpha_1,A\alpha_2,\cdots,A\alpha_n) \\
&= (\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix},
\end{aligned}
$$

其中 A_1 是一个 $n-1$ 阶方阵. 注意到 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 非异, 上式即为

$$
P^{-1}AP = \begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}.
$$

因为 A_1 是一个 $n-1$ 阶方阵, 所以由归纳假设可知, 存在 $n-1$ 阶非异阵 Q, 使 $Q^{-1}A_1Q$ 是一个上三角阵. 令

$$
R = \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix},
$$

则 R 是 n 阶非异阵, 且

$$
\begin{aligned}
R^{-1}P^{-1}APR &= \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix}^{-1}\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}\begin{pmatrix} 1 & O \\ O & Q \end{pmatrix} \\
&= \begin{pmatrix} 1 & O \\ O & Q^{-1} \end{pmatrix}\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}\begin{pmatrix} 1 & O \\ O & Q \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & * \\ O & Q^{-1}A_1Q \end{pmatrix}.
\end{aligned}
$$

这是一个上三角阵, 它与 A 相似, 并且对角元素都是其特征值. □

推论 6.1

若数域 $\mathbb{K}$ 上的 n 阶方阵 A 的特征值全在 $\mathbb{K}$ 中, 则存在 $\mathbb{K}$ 上的非异阵 P, 使 $P^{-1}AP$ 是一个上三角阵.

证明 由复方阵必相似于上三角阵的证明类似可得. □

命题 6.1

1. 设 φ 为 n 维线性空间 V 上的线性变换, 则 φ 在 V 上至少存在一个特征值 $\lambda_0 \in \mathbb{C}$ 及其特征向量 $\alpha_0 \in V$.
2. 设 A 为 n 阶复矩阵, 则 A 在复数域上至少存在一个特征值 $\lambda_0 \in \mathbb{C}$ 及其特征向量 $\alpha_0 \in \mathbb{C}^n$.

证明

1. 任取 V 的一组基 $\{e_1, e_2, \cdots, e_n\}$, 设 φ 在这组基下的表示矩阵为 A, 由代数学基本定理可知, 特征多项式 $|\lambda I_V - \varphi| = |\lambda I_n - A|$ 在复数域上至少有一个根 $\lambda_0 \in \mathbb{C}$. 又由线性方程组理论可知, $(\lambda_0 I_n - A)x = 0$ 一定有非零解 $(x_1, x_2, \cdots, x_n)^T$, 即 $\lambda_0 I_n \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. 记 $\alpha_0 = (e_1, e_2, \cdots, e_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, 则
$$\varphi(\alpha_0) = \varphi(e_1, e_2, \cdots, e_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (e_1, e_2, \cdots, e_n)A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$
$$= (e_1, e_2, \cdots, e_n) \cdot \lambda_0 I_n \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda_0 (e_1, e_2, \cdots, e_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda_0 \alpha_0.$$
故 φ 在 V 上至少存在一个特征值 $\lambda_0 \in \mathbb{C}$ 及其特征向量 $\alpha_0 \in V$.
2. 由代数学基本定理可知, 特征多项式 $|\lambda I_n - A|$ 在复数域上至少有一个根 $\lambda_0 \in \mathbb{C}$. 又由线性方程组理论可知, $(\lambda_0 I_n - A)x = 0$ 一定有非零解 $\alpha_0 \in \mathbb{C}^n$. 故 A 在复数域上至少存在一个特征值 $\lambda_0 \in \mathbb{C}$ 及其特征向量 $\alpha_0 \in \mathbb{C}^n$.

□

6.1.1 直接利用定义计算和证明

例题 6.1 设 V 是 n 阶矩阵全体组成的线性空间, φ 是 V 上的线性变换: $\varphi(X) = AX$, 其中 A 是一个 n 阶矩阵. 求证: φ 和 A 具有相同的特征值 (重数可能不同).

证明 设 λ_0 是 A 的特征值, x_0 是对应的特征向量, 即 $Ax_0 = \lambda_0 x_0$. 令 $X = (x_0, 0, \ldots, 0)$, 则 $\varphi(X) = AX = \lambda_0 X$ 且 $X \neq 0$, 因此 λ_0 也是 φ 的特征值.

反之, 设 λ_0 是 φ 的特征值, X 是对应的特征向量, 即 $\varphi(X) = AX = \lambda_0 X$. 令 $X = (x_1, x_2, \ldots, x_n)$ 为列分块, 设第 i 个列向量 $x_i \neq 0$, 则 $Ax_i = \lambda_0 x_i$, 因此 λ_0 也是 A 的特征值. □

例题 6.2 设 λ_1, λ_2 是矩阵 A 的两个不同的特征值, α_1, α_2 分别是 λ_1, λ_2 的特征向量, 求证: $\alpha_1 + \alpha_2$ 必不是 A 的特征向量.

证明 用反证法, 设 $A(\alpha_1 + \alpha_2) = \mu(\alpha_1 + \alpha_2)$, 又
$$A(\alpha_1 + \alpha_2) = A\alpha_1 + A\alpha_2 = \lambda_1\alpha_1 + \lambda_2\alpha_2,$$
于是 $(\lambda_1 - \mu)\alpha_1 + (\lambda_2 - \mu)\alpha_2 = 0$. 由于属于不同特征值的特征向量线性无关, 故有 $\lambda_1 = \mu, \lambda_2 = \mu$, 从而 $\lambda_1 = \lambda_2$, 引出矛盾. □

命题 6.2

设 φ 是线性空间 V 上的线性变换, V 有一个直和分解:

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_m,$$

其中 V_i 都是 φ-不变子空间.

(1) 设 φ 限制在 V_i 上的特征多项式为 $f_i(\lambda)$, 求证: φ 的特征多项式

$$f(\lambda) = f_1(\lambda)f_2(\lambda)\cdots f_m(\lambda).$$

(2) 设 λ_0 是 φ 的特征值, $V_0 = \{v \in V \mid \varphi(v) = \lambda_0 v\}$ 为特征子空间, $V_{i,0} = V_i \cap V_0 = \{v \in V_i \mid \varphi(v) = \lambda_0 v\}$, 求证:

$$V_0 = V_{1,0} \oplus V_{2,0} \oplus \cdots \oplus V_{m,0}.$$

证明

(1) 取 V_i 的一组基, 将它们拼成 V 的一组基. 记 A_i 是 φ 在 V_i 上的限制在 V_i 所取基下的表示矩阵, 则由定理 4.6可知 φ 在 V 的这组基下的表示矩阵为分块对角矩阵 $A = \mathrm{diag}(A_1, A_2, \cdots, A_m)$, 于是

$$f(\lambda) = |\lambda I_n - A| = |\lambda I - A_1||\lambda I - A_2|\cdots|\lambda I - A_m|,$$

即 $f(\lambda) = f_1(\lambda)f_2(\lambda)\cdots f_m(\lambda)$.

(2) 任取 $\alpha \in V_0$, 设 $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m$, 其中 $\alpha_i \in V_i$, 则

$$\varphi(\alpha_1) + \varphi(\alpha_2) + \cdots + \varphi(\alpha_m) = \varphi(\alpha) = \lambda_0\alpha = \lambda_0\alpha_1 + \lambda_0\alpha_2 + \cdots + \lambda_0\alpha_m.$$

注意到 $\varphi(\alpha_i) \in V_i, \varphi(\alpha) = \lambda_0\alpha \in V$, 故由直和的等价条件 (5)可得 $\varphi(\alpha_i) = \lambda_0\alpha_i$, 即 $\alpha_i \in V_{i,0}$, 从而 $V_0 = V_{1,0} \oplus V_{2,0} \oplus \cdots \oplus V_{m,0}$. 注意到 $V_{i,0} \subseteq V_i$, 故

$$V_{i,0} \cap (V_{1,0} + \cdots + V_{i-1,0}) \subseteq V_i \cap (V_1 + \cdots + V_{i-1}) = \{0\}, \quad 2 \le i \le m,$$

于是由直和的等价条件 (2)可知上述为直和.

□

推论 6.2

对分块对角矩阵 $A = \mathrm{diag}\{A_1, A_2, \cdots, A_m\}$ 的任一特征值 λ_0, 其代数重数等于每个分块的代数重数之和, 其几何重数等于每个分块的几何重数之和.

证明 将命题 6.2的条件和结论代数化之后, 即可得到结论. □

命题 6.3 (特征向量的延拓)

设 n 阶分块对角矩阵 $A = \mathrm{diag}\{A_1, A_2, \cdots, A_m\}$, 其中 A_i 是 n_i 阶矩阵.

(1) 任取 A_i 的特征值 λ_i 及其特征向量 $x_i \in \mathbb{C}^{n_i}$, 求证: 可在 x_i 的上下添加适当多的零, 得到非零向量 $\tilde{x}_i \in \mathbb{C}^n$, 使得 $A\tilde{x}_i = \lambda_i\tilde{x}_i$, 即 $\tilde{x}_i$ 是 A 关于特征值 λ_i 的特征向量, 称为 x_i 的**延拓**.

(2) 任取 A 的特征值 λ_0, 并设 λ_0 是 $A_{i_1}, \cdots, A_{i_r}$ 的特征值, 但不是其他 $A_j\ (1 \le j \le m, j \ne i_1, \cdots, i_r)$ 的特征值, 求证: A 关于特征值 λ_0 的特征子空间的一组基可取为 $A_{i_1}(1 \le k \le r)$ 关于特征值 λ_0 的特征子空间的一组基的延拓的并集.

证明

(1) 令 $\tilde{x}_i = (0, \cdots, 0, x_i, 0, \cdots, 0)'$, 即 $\tilde{x}_i$ 的第 i 块为 x_i, 其余块均为 0, 显然 $\tilde{x}_i \ne 0$. 容易验证 $A\tilde{x}_i = \lambda_i\tilde{x}_i$, 故结论成立.

(2) 由命题 6.2(2)以及直和的等价条件 (5)即得.

□

例题 6.3 设 A 是 n 阶整数矩阵, p, q 为互素的整数且 $q > 1$. 求证: 矩阵方程 $Ax = \dfrac{p}{q}x$ 必无非零解.

证明 用反证法. 设上述矩阵方程有非零解, 则 $\dfrac{p}{q}$ 为 A 的特征值, 即为特征多项式 $f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$ 的根. 由于 A 是整数矩阵, 故 $f(\lambda)$ 为整数系数多项式. 由整数系数多项式有有理根的必要条件可知 $q\mid 1$, 从而 $q=\pm1$, 于是 $q\mid p$, 这与 p,q 互素矛盾. □

例题 6.4 求下列 n 阶矩阵的特征值:

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

解 若 $a=0$ 或 $b=0$, 则 A 是主对角元全为零的下三角或上三角矩阵, 故 A 的特征值全为零. 设 $a\neq0, b\neq0$, 则由命题 1.7可知: 若 $a\neq b$, 则 $|\lambda I_n-A|=\dfrac{a(\lambda+b)^n-b(\lambda+a)^n}{a-b}$. 设 $\dfrac{b}{a}$ 的 n 次方根为 $\omega_i(1\leq i\leq n)$, 则

$$|\lambda I_n-A|=\frac{a(\lambda+b)^n-b(\lambda+a)^n}{a-b}=0\Rightarrow\left(\frac{\lambda+b}{\lambda+a}\right)^n=\frac{b}{a}$$
$$\Rightarrow\frac{\lambda+b}{\lambda+a}=w_i\,(1\leq i\leq n)\Rightarrow\lambda=\frac{a\omega_i-b}{1-\omega_i}(1\leq i\leq n).$$

从而 A 的特征值为 $\dfrac{a\omega_i-b}{1-\omega_i}(1\leq i\leq n)$.

若 $a=b$, 则 $|\lambda I_n-A|=(\lambda-(n-1)a)(\lambda+a)^{n-1}$, 从而 A 的特征值为 $(n-1)a$ (1 重), $-a$ $(n-1$ 重).

综上, 容易验证当 $a=b=0$ 或 $ab\neq0$ 时,A 有完全的特征向量系或有 n 个不同的特征值, 从而此时 A 可对角化. 若 A 可对角化也不难得到 $a=b=0$ 或 $ab\neq0$. 故 A 可对角化的充分必要条件是 $a=b=0$ 或 $ab\neq0$. □

6.1.2 正向利用矩阵的多项式

定义 6.7 (矩阵多项式)

若 A 是一个 n 阶矩阵,$f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$ 是一个多项式, 记

$$f(A)=a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0I_n.$$

♣

命题 6.4

设 n 阶矩阵 A 的全部特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, $f(x)$ 是一个多项式, 则 $f(A)$ 的全部特征值为 $f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$.

♠

注 这个命题告诉我们: 如果能够将一个复杂矩阵写成一个简单矩阵的多项式, 那么就可以由简单矩阵的特征值得到复杂矩阵的特征值.

证明 因为任一 n 阶矩阵均复相似于上三角阵, 可设

$$P^{-1}AP=\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

因为上三角阵的和、数乘及乘方仍是上三角阵, 经计算不难得到

$$P^{-1}f(A)P=f(P^{-1}AP)=\begin{pmatrix} f(\lambda_1) & * & \cdots & * \\ 0 & f(\lambda_2) & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix}.$$

因此 $f(A)$ 的全部特征值为 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$. □

例题 6.5 设 n 阶矩阵 A 的全体特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 求 $2n$ 阶矩阵

$$\begin{pmatrix} A^2 & A \\ A^2 & A \end{pmatrix}$$

的全体特征值.

证明 由命题 2.40可知

$$\left|\lambda I_{2n} - \begin{pmatrix} A^2 & A \\ A^2 & A \end{pmatrix}\right| = \left|\begin{pmatrix} \lambda I_n - A & -A^2 \\ -A^2 & \lambda I_n - A \end{pmatrix}\right| = |\lambda I_n - A - A^2||\lambda I_n - A + A^2|.$$

由命题 6.4可知 $A + A^2$ 的全体特征值为 $\lambda_i + \lambda_i^2 (1 \leq i \leq n)$, $A - A^2$ 的全体特征值为 $\lambda_i - \lambda_i^2 (1 \leq i \leq n)$, 因此所求矩阵的全体特征值为

$$\lambda_1 + \lambda_1^2, \lambda_1 - \lambda_1^2, \lambda_2 + \lambda_2^2, \lambda_2 - \lambda_2^2, \cdots, \lambda_n + \lambda_n^2, \lambda_n - \lambda_n^2.$$

□

命题 6.5 (循环矩阵的特征值)

求下列循环矩阵的特征值:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}.$$

♠

解 设 $J = \begin{pmatrix} O & I_{n-1} \\ 1 & O \end{pmatrix}$, $f(x) = a_1 + a_2 x + a_3 x^2 + \cdots + a_n x^{n-1}$, 则由循环矩阵的性质 2可知 $A = f(J)$. 经简单计算可得

$$|\lambda I_n - J| = \begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & -1 \\ -1 & 0 & 0 & \cdots & 0 & \lambda \end{vmatrix} \xlongequal{\text{按第一列展开}} \lambda \begin{vmatrix} \lambda & -1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & -1 \\ 0 & 0 & \cdots & 0 & \lambda \end{vmatrix} + (-1)^{n+1}(-1) \begin{vmatrix} -1 & 0 & \cdots & 0 & 0 \\ \lambda & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & \cdots & \lambda & -1 \end{vmatrix}$$

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

于是 J 的特征值为

$$\omega_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, \quad 0 \leq k \leq n-1.$$

因此 A 的特征值为 $f(1), f(\omega_1), \cdots, f(\omega_{n-1})$. □

定义 6.8 (友矩阵)

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{n-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

称为多项式 $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ 的 **友矩阵**.

♣

命题 6.6 (友矩阵的特征多项式及特征值)

设首一多项式 $f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$, $f(x)$ 的友矩阵

$$C=\begin{pmatrix}0&0&\cdots&0&-a_0\\1&0&\cdots&0&-a_1\\0&1&\cdots&0&-a_2\\\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&\cdots&1&-a_{n-1}\end{pmatrix}.$$

(1) 求证: 矩阵 C 的特征多项式就是 $f(\lambda)$.

(2) 设 $f(x)$ 的根为 $\lambda_1,\lambda_2,\cdots,\lambda_n$,$g(x)$ 为任一多项式, 求以 $g(\lambda_1),g(\lambda_2),\cdots,g(\lambda_n)$ 为根的 n 次多项式.

证明

(1)

$$|xE-\boldsymbol{A}|=\begin{vmatrix}x&0&0&\cdots&0&a_0\\-1&x&0&\cdots&0&a_1\\0&-1&x&\cdots&0&a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&x&a_{n-2}\\0&0&0&\cdots&-1&x+a_{n-1}\end{vmatrix}\xlongequal[i=n,n-1,\cdots,2]{xr_i+r_{i-1}}\begin{vmatrix}0&0&0&\cdots&0&x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\\-1&0&0&\cdots&0&x^{n-1}+a_{n-1}x^{n-2}+\cdots+a_2x+a_1\\0&-1&0&\cdots&0&x^{n-2}+a_{n-1}x^{n-3}+\cdots+a_3x+a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&x^2+a_{n-1}x+a_{n-2}\\0&0&0&\cdots&-1&x+a_{n-1}\end{vmatrix}$$

$$\xlongequal{\text{按第一行展开}}\left(x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a\right)(-1)^{n+1}\begin{vmatrix}-1&0&0&\cdots&0\\0&-1&0&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&0\\0&0&0&\cdots&-1\end{vmatrix}$$

$$=\left(x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a\right)(-1)^{n+1}(-1)^{n-1}$$

$$=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a=f(x).$$

(2) 由假设及 (1) 的结论可知 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 C 的全体特征值, 故由命题 6.4可知 $g(\lambda_1),g(\lambda_2),\cdots,g(\lambda_n)$ 是 $g(C)$ 的全体特征值, 从而 $h(x)=|xI_n-g(C)|$ 即为所求的多项式.

□

6.1.3 反向利用矩阵的多项式

命题 6.7

设 n 阶矩阵 A 适合一个多项式 $g(x)$, 即 $g(A)=O$, 则 A 的任一特征值 λ_0 也必适合 $g(x)$, 即 $g(\lambda_0)=0$.

证明 证法一: 设 α 是 A 关于特征值 λ_0 的特征向量, 经简单计算得

$$g(\lambda_0)\alpha=g(A)\alpha=0.$$

而 $\alpha\neq 0$, 因此 $g(\lambda_0)=0$.

证法二: 设 A 的极小多项式为 $m(x)$, 则 $m(x)\mid g(x)$, 由极小多项式的性质 (5)及整除的传递性可知 $(x-\lambda_0)\mid g(x)$, 故 $g(\lambda_0)=0$. □

命题 6.8 (幂零矩阵关于特征值的充要条件)

求证:n 阶矩阵 A 为幂零矩阵的充要条件是 A 的特征值全为零.

证明 若 A 为幂零矩阵, 即存在正整数 k, 使得 $A^k = O$, 则由命题 6.7可知 A 的任一特征值 λ_0 也适合 x^k, 于是 $\lambda_0 = 0$.

反之, 证法一:若 A 的特征值全为零, 则存在可逆矩阵 P, 使得 $P^{-1}AP = B$ 为上三角矩阵且主对角元素全为零. 由上三角阵性质 (1)可知 $B^n = O$, 于是 $A^n = (PBP^{-1})^n = PB^nP^{-1} = O$, 即 A 为幂零矩阵.

证法二:也可以利用 Cayley-Hamilton 定理来证明, 由于 A 的特征值全为零, 故其特征多项式为 λ^n, 从而 $A^n = O$. □

例题 6.6 设 V 是数域 $\mathbb{F}$ 上的 n 阶方阵全体构成的线性空间, n 阶方阵

$$P = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & & \vdots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix},$$

V 上的线性变换 η 定义为 $\eta(X) = PX'P$. 试求 η 的全体特征值及其特征向量.

笔记 任意 n 阶矩阵 A 左乘 P 相当于行倒排, 右乘 P 矩阵相当于列倒排.

解 由 $P = P', P^2 = I_n$ 容易验证 $\eta^2(X) = P(PX'P)P = X$, 即 $\eta^2 = I_V$, 于是 η 的特征值也适合多项式 $x^2 - 1$, 从而特征值只能是 ± 1.

设 $\eta(X_0) = PX_0'P = \pm X_0$, 这等价于 $(PX_0')P = \pm PX_0$, 即 PX_0 为对称矩阵或反对称矩阵.

令 $PX_0 = E_{ii}, E_{ij} + E_{ji}$ (对称矩阵空间的基向量), 容易证明 η 关于特征值 1 的线性无关的特征向量为 $X_0 = PE_{ii}\ (1 \le i \le n), P(E_{ij} + E_{ji})\ (1 \le i < j \le n)$.

令 $PX_0 = E_{ij} - E_{ji}$ (反对称矩阵空间的基向量), 容易证明 η 关于特征值 -1 的线性无关的特征向量为 $X_0 = P(E_{ij} - E_{ji})\ (1 \le i < j \le n)$. 注意到这些特征向量恰好构成 V 的一组基, 故 η 的特征值为 1 $\dfrac{n(n+1)}{2}$ 重, -1 $\dfrac{n(n-1)}{2}$ 重. □

例题 6.7 设 n 阶方阵 A 的每行每列只有一个元素非零, 并且那些非零元素为 1 或 -1, 证明: A 的特征值都是单位根.

证明 设 S 为由每行每列只有一个元素非零, 并且那些非零元素为 1 或 -1 的所有 n 阶方阵构成的集合, 由排列组合可得 $\overline{\overline{S}} = 2^n n!$, 即 S 是一个有限集合. 注意到矩阵 $M \in S$ 当且仅当 $M = P_1P_2\cdots P_r$, 其中 P_k 是初等矩阵 P_{ij} 或 $P_i(-1)$, 因此对任意的 $M, N \in S$, $MN \in S$. 特别地, 由 $A \in S$ 可知 $A^k \in S\ (k \ge 1)$, 即 $\{A, A^2, A^3, \cdots\} \subseteq S$, 于是存在正整数 $k > l$, 使得 $A^k = A^l$. 注意到 $|A| = \pm 1$, 故 A 可逆, 于是 $A^{k-l} = I_n$, 从而 A 的特征值适合多项式 $x^{k-l} - 1$, 即为单位根. □

例题 6.8 设 A 是 n 阶实方阵, 又 $I_n - A$ 的特征值的模长都小于 1, 求证: $0 < |A| < 2^n$.

证明 设 A 的特征值为 $\lambda_1, \cdots, \lambda_n$, 则 $I_n - A$ 的特征值为 $1 - \lambda_1, \cdots, 1 - \lambda_n$. 由假设 $|1 - \lambda_i| < 1$, 若 λ_i 是实数, 则 $0 < \lambda_i < 2$; 若 λ_i 是虚数, 则 $\overline{\lambda_i}$ 也是 A 的特征值, 此时 $1 - \overline{\lambda_i}$ 也是 $I_n - A$ 的特征值. 从而 $|1 - \lambda_i| < 1, |1 - \overline{\lambda_i}| < 1$, 于是

$$|1 - \lambda_i^2| = |(1 - \lambda_i)(1 - \overline{\lambda_i})| = |1 - \lambda_i||1 - \overline{\lambda_i}| < 1.$$

因此 $0 < \lambda_i^2 < 2$, 故此时 $0 < \lambda_i < \sqrt{2}$.

综上, 无论 λ_i 是实数还是虚数, 都有 $0 < |\lambda_i| < 2$. 由于 $|A|$ 等于所有特征值之积, 故 $0 < |A| < 2^n$. □

命题 6.9 (逆矩阵的特征值)

设 n 阶矩阵 A 是可逆矩阵, 且 A 的全部特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则 A^{-1} 的全部特征值为 $\lambda_1^{-1}, \lambda_2^{-1}, \cdots, \lambda_n^{-1}$.

证明 首先注意到 A 是可逆矩阵, $\lambda_1\lambda_2\cdots\lambda_n = |A| \ne 0$, 因此每个 $\lambda_i \ne 0$ (事实上, A 可逆的充分必要条件是它的特

征值全不为零). 由复方阵必相似于上三角阵可设

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

因为上三角矩阵的逆矩阵仍然是上三角矩阵, 经过计算不难得到

$$P^{-1}A^{-1}P=(P^{-1}AP)^{-1}=\begin{pmatrix}\lambda_1^{-1} & * & \cdots & * \\ 0 & \lambda_2^{-1} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^{-1}\end{pmatrix}.$$

因此 A^{-1} 的全部特征值为 $\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$. □

命题 6.10 (伴随矩阵的特征值)

设 n 阶矩阵 A 的全体特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 求证:$A^*$ 的全体特征值为

$$\prod_{i\neq 1}\lambda_i,\prod_{i\neq 2}\lambda_i,\cdots,\prod_{i\neq n}\lambda_i.$$

证明 因为任一 n 阶矩阵均复相似于上三角矩阵, 故可设

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

注意到上三角矩阵的伴随矩阵仍是上三角矩阵, 经计算可得

$$P^{-1}A^*P=P^*A^*(P^{-1})^*=(P^{-1}AP)^*=\begin{pmatrix}\prod\limits_{i\neq 1}\lambda_i & * & \cdots & * \\ 0 & \prod\limits_{i\neq 2}\lambda_i & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \prod\limits_{i\neq n}\lambda_i\end{pmatrix}.$$

因此 A^* 的全部特征值为

$$\prod_{i\neq 1}\lambda_i,\prod_{i\neq 2}\lambda_i,\cdots,\prod_{i\neq n}\lambda_i.$$

□

6.1.4 特征值的降价公式

定理 6.3 (特征值的降价公式)

设 A 是 $m\times n$ 矩阵,B 是 $n\times m$ 矩阵, 且 $m\geq n$. 求证:

$$|\lambda I_m-AB|=\lambda^{m-n}|\lambda I_n-BA|.$$

特别地, 若 A,B 都是 n 阶矩阵, 则 AB 与 BA 有相同的特征多项式.

笔记 本质上就是打洞原理.

证明 证法一 (打洞原理): 当 $\lambda \neq 0$ 时, 考虑下列分块矩阵:

$$\begin{pmatrix} \lambda I_m & A \\ B & I_n \end{pmatrix},$$

因为 $\lambda I_m, I_n$ 都是可逆矩阵, 故由行列式的降阶公式可得

$$|I_n| \cdot |\lambda I_m - A(I_n)^{-1}B| = |\lambda I_m| \cdot |I_n - B(\lambda I_m)^{-1}A|,$$

即有

$$|\lambda I_m - AB| = \lambda^{m-n}|\lambda I_n - BA|$$

成立.

当 $\lambda = 0$ 时, 若 $m > n$, 则 $r(AB) \leq \min\{r(A), r(B)\} \leq \min\{m, n\} = n < m$, 故 $|-AB| = 0$, 结论成立; 若 $m = n$, 则 $|-AB| = (-1)^n|A||B| = |-BA|$, 结论也成立.

事实上, $\lambda = 0$ 的情形也可以用 Cauchy-Binet 公式来处理, 还可以通过摄动法由 $\lambda \neq 0$ 的情形来得到.

证法二 (相抵标准型): 设 A 的秩等于 r, 则存在 m 阶可逆矩阵 P 和 n 阶可逆矩阵 Q, 使得

$$PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

令

$$Q^{-1}BP^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

其中 B_{11} 是 $r \times r$ 矩阵, 则

$$PABP^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix}, \quad Q^{-1}BAQ = \begin{pmatrix} B_{11} & O \\ B_{21} & O \end{pmatrix}.$$

因此

$$|\lambda I_m - AB| = \left| \begin{pmatrix} \lambda I_r - B_{11} & -B_{12} \\ O & \lambda I_{m-r} \end{pmatrix} \right| = \lambda^{m-r}|\lambda I_r - B_{11}|,$$

同理

$$|\lambda I_n - BA| = \left| \begin{pmatrix} \lambda I_r - B_{11} & O \\ -B_{21} & \lambda I_{n-r} \end{pmatrix} \right| = \lambda^{n-r}|\lambda I_r - B_{11}|.$$

比较上面两个式子即可得出结论.

证法三 (摄动法): 先证明 $m = n$ 的情形. 若 A 可逆, 则 $BA = A^{-1}(AB)A$, 即 AB 和 BA 相似, 因此它们的特征多项式相等. 对于一般的方阵 A, 可取到一列有理数 $t_k \to 0$, 使得 $t_kI_n + A$ 是可逆矩阵. 由可逆情形的证明可得

$$|\lambda I_n - (t_kI_n + A)B| = |\lambda I_n - B(t_kI_n + A)|.$$

注意到上述两边的行列式都是 t_k 的多项式, 从而关于 t_k 连续. 上式两边同时取极限, 令 $t_k \to 0$, 即有 $|\lambda I_n - AB| = |\lambda I_n - BA|$ 成立.

再证明 $m > n$ 的情形. 令

$$C = \begin{pmatrix} A & O \end{pmatrix}, \quad D = \begin{pmatrix} B \\ O \end{pmatrix},$$

其中 C, D 均为 $m \times m$ 分块矩阵, 则

$$CD = AB, \quad DC = \begin{pmatrix} BA & O \\ O & O \end{pmatrix}.$$

因此由方阵的情形可得

$$|\lambda I_m - AB| = |\lambda I_m - CD| = |\lambda I_m - DC| = \lambda^{m-n}|\lambda I_n - BA|.$$

□

例题 6.9 设 α 是 n 维实列向量且 $\alpha'\alpha = 1$, 试求矩阵 $I_n - 2\alpha\alpha'$ 的特征值.

解 设 $A = I_n - 2\alpha\alpha'$, 则由特征值的降价公式可得

$$|\lambda I_n - A| = |(\lambda - 1)I_n + 2\alpha\alpha'| = (\lambda - 1)^{n-1}(\lambda - 1 + 2\alpha'\alpha) = (\lambda - 1)^{n-1}(\lambda + 1).$$

因此, 矩阵 A 的特征值为 $1(n-1)$ 重,-1(1 重). 进一步, 容易验证 A 有完全的特征向量系 ($|-I_n - A|$ 为零, 但其 $n-1$ 阶子式不为零), 于是 A 可对角化. □

例题 6.10 设 A 为 n 阶方阵, α, β 为 n 维列向量, 试求矩阵 $A\alpha\beta'$ 的特征值.

解 设 $B = A\alpha\beta'$, 则由特征值的降价公式可得

$$|\lambda I_n - B| = |I_n - (A\alpha)\beta'| = \lambda^{n-1}(\lambda - \beta' A\alpha).$$

若 $\beta' A\alpha \neq 0$, 则 B 的特征值为 $0\,(n-1$ 重$)$, $\beta' A\alpha$ (1 重). 进一步, 容易验证此时 B 有完全的特征向量系, 从而可对角化. 若 $\beta' A\alpha = 0$, 则 B 的特征值为 $0\,(n$ 重$)$.

综上, 容易验证 $B = \boldsymbol{A\alpha\beta'}$ 可对角化的充要条件是 $\boldsymbol{\beta' A\alpha} \neq 0$ 或 $\boldsymbol{A\alpha\beta'} = \boldsymbol{O}$. □

例题 6.11 设 $a_i\ (1 \leq i \leq n)$ 都是实数, 且 $a_1 + a_2 + \cdots + a_n = 0$, 试求下列矩阵的特征值:

$$A = \begin{pmatrix} a_1^2 & a_1a_2 + 1 & \cdots & a_1a_n + 1 \\ a_2a_1 + 1 & a_2^2 & \cdots & a_2a_n + 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_na_1 + 1 & a_na_2 + 1 & \cdots & a_n^2 \end{pmatrix}.$$

解 矩阵 A 可以分解为 $A = -I_n + BC$, 其中

$$B = \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}, \quad C = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

由特征值的降价公式得

$$\begin{aligned} |\lambda I_n - A| &= |(\lambda + 1)I_n - BC| \\ &= (\lambda + 1)^{n-2}|(\lambda + 1)I_2 - CB|. \end{aligned}$$

注意到 $a_1 + a_2 + \cdots + a_n = 0$, 故有

$$CB = \begin{pmatrix} a_1^2 + a_2^2 + \cdots + a_n^2 & 0 \\ 0 & n \end{pmatrix}.$$

因此 A 的特征值为 $-1\,(n-2$ 重$)$, $n-1$, $a_1^2 + a_2^2 + \cdots + a_n^2 - 1$. 进一步, 若 a_i 全部为零, 则特征值 -1 和 $n-1$ 都有完全的特征向量系. 若 $\sum\limits_{i=1}^{n} a_i^2 = n$, 利用秩的降阶公式可得特征值 -1 和 $n-1$ 都有完全的特征向量系. 在剩余情况, 利用秩的降阶公式可得 3 个特征值都有完全的特征向量系. 因此, $\boldsymbol{A}$ 可对角化. 事实上, 即使去掉 $a_1 + a_2 + \cdots + a_n = 0$ 的条件, 也可以计算出 $\boldsymbol{A}$ 的全体特征值的代数重数和几何重数, 从而得到 $\boldsymbol{A}$ 可对角化. 这一结论的深层次背景是: $\boldsymbol{A}$ 是实对称矩阵, 从而可正交对角化. □

例题 6.12 设 A, B, C 分别是 $m \times m, n \times n, m \times n$ 矩阵, 满足: $AC = CB$, $\mathrm{r}(C) = r$. 求证: A 和 B 至少有 r 个相同的特征值.

注 不妨设 $C = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 的原因: 假设当 $C = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 时, 结论已经成立, 则对于一般的满足条件的矩阵 C, 由条件我们有

$$r(C) = r, \quad AC = BC.$$

由 $r(C)=r$ 可知, 存在可逆矩阵 P,Q, 使得

$$PCQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

从而对 $AC=BC$ 两边同时左乘 P, 右乘 Q 得到

$$(PAP^{-1})(PCQ)=(PCQ)(Q^{-1}BQ).$$

于是由假设可知 PAP^{-1} 和 $Q^{-1}BQ$ 都至少有 r 个相同的特征值. 又因为相似矩阵有相同的特征值, 所以 A,B 也至少有 r 个相同的特征值. 故不妨设成立.

证明 设 P 为 m 阶非异阵, Q 为 n 阶非异阵, 使得

$$PCQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

注意到问题的条件和结论在相抵变换 $C\mapsto PCQ$, $A\mapsto PAP^{-1}$, $B\mapsto Q^{-1}BQ$ 下保持不变, 故不妨从一开始就假设 $C=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 是相抵标准型. 设

$$A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},\quad B=\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

为对应的分块, 则

$$AC=\begin{pmatrix} A_{11} & O \\ A_{21} & O \end{pmatrix},\quad CB=\begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix}.$$

由 $AC=CB$ 可得 $A_{11}=B_{11}, A_{21}=0, B_{12}=0$. 于是

$$|\lambda I_m-A|=|\lambda I_r-A_{11}|\cdot|\lambda I_{m-r}-A_{22}|,$$
$$|\lambda I_n-B|=|\lambda I_r-B_{11}|\cdot|\lambda I_{n-r}-B_{22}|.$$

从而 A,B 至少有 r 个相同的特征值 (即 $A_{11}=B_{11}$ 的特征值). □

6.1.5 特征值与特征多项式系数的关系

命题 6.11 (特征值与特征多项式系数的关系)

设 n 阶矩阵 A 的特征多项式为

$$f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n.$$

求证: a_r 等于 $(-1)^r$ 乘以 A 的所有 r 阶主子式之和, 即

$$a_r=(-1)^r\sum_{1\le i_1<i_2<\cdots<i_r\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix},\quad 1\le r\le n.$$

进一步, 若设 A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则

$$\sum_{1\le i_1<i_2<\cdots<i_r\le n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}=\sum_{1\le i_1<i_2<\cdots<i_r\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix},\quad 1\le r\le n.$$

注 上述结论中最常用的是 $r=1$ 和 $r=n$ 的情形:

$$\lambda_1+\lambda_2+\cdots+\lambda_n=\mathrm{tr}(A),\quad \lambda_1\lambda_2\cdots\lambda_n=|A|.$$

特别地, **A 是非异阵的充要条件是 A 的特征值全不为零.** 因此, 特征值的计算是判断矩阵是否非异阵的重要依据.

证明 第一种结论是推论 1.5.

由Vieta 定理可得

$$\sum_{1\le i_1<i_2<\cdots<i_r\le n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}=(-1)^r a_r=(-1)^r(-1)^r\sum_{1\le i_1<i_2<\cdots<i_r\le n}A\begin{pmatrix}i_1&i_2&\cdots&i_r\\i_1&i_2&\cdots&i_r\end{pmatrix}$$
$$=\sum_{1\le i_1<i_2<\cdots<i_r\le n}A\begin{pmatrix}i_1&i_2&\cdots&i_r\\i_1&i_2&\cdots&i_r\end{pmatrix},1\le r\le n.$$

因此第二种结论也成立. □

例题 6.13 设 n 阶方阵 A 满足

$$A^2-A-3I_n=O,$$

求证:$A-2I_n$ 是非奇异阵.

笔记 用特征值判断矩阵非异性.

证明 用反证法. 设 $A-2I_n$ 为奇异阵, 则 2 是 A 的特征值. 注意到 A 适合

$$f(x)=x^2-x-3,$$

但特征值 2 却不适合 $f(x)$, 这与命题 6.7矛盾. □

例题 6.14 设 P 是可逆矩阵, $B=PAP^{-1}-P^{-1}AP$, 求证: B 的特征值之和为零.

证明 由特征值与特征多项式系数的关系可知, 只要证 $\mathrm{tr}(B)=0$ 即可. 由迹的线性和交换性即得

$$\mathrm{tr}(B)=\mathrm{tr}(PAP^{-1})-\mathrm{tr}(P^{-1}AP)=\mathrm{tr}(A)-\mathrm{tr}(A)=0.$$

□

例题 6.15 设 n 阶实方阵 A 的特征值全是实数, 且 A 的一阶主子式之和与二阶主子式之和都等于零. 求证: A 是零矩阵.

证明 设 A 的特征值为 $\lambda_1,\lambda_2,\ldots,\lambda_n$, 由条件和特征值与特征多项式系数的关系可知

$$\sum_{i=1}^n\lambda_i=\lambda_1+\lambda_2+\cdots+\lambda_n=0,$$
$$\sum_{1\le i<j\le n}\lambda_i\lambda_j=\lambda_1\lambda_2+\lambda_1\lambda_3+\cdots+\lambda_{n-1}\lambda_n=0.$$

则

$$\sum_{i=1}^n\lambda_i^2=\left(\sum_{i=1}^n\lambda_i\right)^2-2\sum_{1\le i<j\le n}\lambda_i\lambda_j=0.$$

由于 λ_i 都是实数, 故 $\lambda_i=0\,(1\le i\le n)$ 成立, 再由命题 6.8可知 A 为零矩阵. □

例题 6.16 设 $n(n\ge 3)$ 阶非异实方阵 A 的特征值都是实数, 且 A 的 $n-1$ 阶主子式之和等于零. 证明: 存在 A 的一个 $n-2$ 阶主子式, 其符号与 $|A|$ 的符号相反.

证明 设 A 的特征值为 $\lambda_1,\lambda_2,\ldots,\lambda_n$, 由 A 非异可知它们都是非零实数. 再由条件和例 6.24 可知

$$\sum_{1\le i_1<i_2<\cdots<i_{n-1}\le n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_{n-1}}=0. \tag{6.1}$$

将(6.1)式左边除以 $|A|=\lambda_1\lambda_2\cdots\lambda_n$ 可得

$$\sum_{i=1}^n\frac{1}{\lambda_i}=0, \tag{6.2}$$

将(6.2) 式左边平方, 并将平方项移到等式的右边可得

$$\sum_{1\le i<j\le n}\frac{1}{\lambda_i\lambda_j}=-\frac{1}{2}\left(\sum_{i=1}^n\frac{1}{\lambda_i}\right)^2<0, \tag{6.3}$$

将(6.3)式两边同时乘以 $|A| = \lambda_1\lambda_2\cdots\lambda_n$ 可得

$$\sum_{1\le i_1<i_2<\cdots<i_{n-2}\le n} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_{n-2}} = -\frac{1}{2}\left(\sum_{i=1}^{n}\frac{1}{\lambda_i^2}\right)|A|. \tag{6.4}$$

由 (6.4)式和特征值与特征多项式系数的关系可得

$$\sum_{1\le i_1<i_2<\cdots<i_{n-2}\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_{n-2} \\ i_1 & i_2 & \cdots & i_{n-2}\end{pmatrix} = -\frac{1}{2}\left(\sum_{i=1}^{n}\frac{1}{\lambda_i^2}\right)|A|,$$

于是 A 的 $n-2$ 阶主子式之和与 $|A|$ 的符号相反, 从而至少存在 A 的一个 $n-2$ 阶主子式, 其符号与 $|A|$ 的符号相反. □

结论 设 n 阶方阵 A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则对任意的正整数 k,A^k 的特征值为 $\lambda_1^k,\lambda_2^k,\cdots,\lambda_n^k$, 于是特征值的 k 次幂和

$$s_k = \lambda_1^k+\lambda_2^k+\cdots+\lambda_n^k = \mathrm{tr}(A^k),\quad k\ge 1.$$

若已知 n 阶方阵 A 的迹 $\mathrm{tr}(A^k)(1\le k\le n)$, 则由 Newton 公式可以计算出特征值的初等对称多项式

$$\sigma_r = \sum_{1\le i_1<i_2<\cdots<i_r\le n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r},\quad 1\le r\le n,$$

从而可以确定特征多项式的系数, 最后便可计算出 A 的所有特征值.

例题 6.17 设 A 是 n 阶对合矩阵, 即 $A^2=I_n$, 证明: $n-\mathrm{tr}(A)$ 为偶数, 并且 $\mathrm{tr}(A)=n$ 的充要条件是 $A=I_n$.

证明 由 $A^2=I_n$ 可知 A 的特征值也适合 x^2-1, 从而只能是 ±1. 设 A 的特征值为 $1(p$ 重$)$, $-1(q$ 重$)$, 则 $p+q=n$. 且 $\mathrm{tr}(A)=p-q$, 于是 $n-\mathrm{tr}(A)=2q$ 为偶数. 若 $A=I_n$, 则 $\mathrm{tr}(A)=n$. 反之, 若 $\mathrm{tr}(A)=n$, 则由上述讨论可知 $p=n$, $q=0$, 从而 -1 不是 A 的特征值, 即 $A+I_n$ 是非奇异阵. 最后由 $A^2=I_n$ 可得

$$(A-I_n)(A+I_n)=O \Rightarrow A-I_n=O \Rightarrow A=I_n.$$

□

例题 6.18 设 4 阶方阵 A 满足: $\mathrm{tr}(A^k)=k\ (1\le k\le 4)$, 试求 A 的行列式.

证明 题目条件即为 $s_k=k\ (1\le k\le 4)$, 要求 $|A|=\sigma_4$. 根据 Newton 公式 (白皮书这一部分还没看)

$$s_k - s_{k-1}\sigma_1+\cdots+(-1)^k k\sigma_k = 0\quad (1\le k\le 4)$$

可依次算出 $\sigma_1=1,\sigma_2=-\dfrac{1}{2},\sigma_3=\dfrac{1}{6},\sigma_4=\dfrac{1}{24}$. 故 $|A|=\dfrac{1}{24}$. 也可以直接利用例 5.64(白皮书这一部分还没看) 来计算 σ_4. □

命题 6.12 (幂零矩阵关于迹的充要条件)

求证: n 阶矩阵 A 是零矩阵的充要条件是 $\mathrm{tr}(A^k)=0\ (1\le k\le n)$.

证明 设 A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 若 A 是零矩阵, 则 A 的特征值全为零, 从而 $\mathrm{tr}(A^k)=s_k=0\ (k\ge 1)$. 若 $s_k=\mathrm{tr}(A^k)=0\ (1\le k\le n)$, 则由 Newton 公式 (白皮书这一部分还没看) 或直接利用例 5.64(白皮书这一部分还没看) 可计算出 $\sigma_r=0\ (1\le r\le n)$, 于是 A 的特征多项式为 λ^n, 从而 A 的特征值全为零, 再由幂零矩阵关于特征值的充要条件可知 A 为零矩阵. □

定义 6.9 (线性变换的迹)

线性变换的迹定义为它在任一组基下的表示矩阵的迹.

笔记 因为矩阵的迹在相似变换下保持不变, 并且同一线性变换在不同基下的表示矩阵必相似, 所以同一线性变换在任意一组基下的表示矩阵的迹都相同, 故线性变换的迹是良定义的.

命题 6.13

设 φ 为 n 维线性空间,λ_0 为 φ 的一个特征值,V_0 为特征值 λ_0 的特征子空间, 则存在 V_0 上的一组基, 使得 φ 在 V_0 上的限制 $\varphi|_{V_0}$ 在这组基下的表示矩阵为 $\dim V_0$ 阶的对角阵 $\mathrm{diag}\{\lambda_0,\lambda_0,\cdots,\lambda_0\}$, 从而 $\mathrm{tr}\left(\varphi|_{V_0}\right)=\lambda_0\dim V_0$.

注 因为线性变换的特征子空间一定是不变子空间, 所以线性变换在其特征子空间上做限制后仍是线性变换, 因此线性变换在其特征子空间上的限制是良定义的.

证明 设 x_1 是 φ 属于 λ_0 的特征向量, 将其扩充成 V_0 的一组基 $\{x_1,x_2,\cdots,x_r\}$, 则 $r=\dim V_0$. 注意到 $x_1,x_2,\cdots,x_r$ 也是 $\varphi|_{V_0}$ 属于 λ_0 的特征向量, 从而

$$\varphi|_{V_0}(x_1,x_2,\cdots,x_r)=(\lambda_0x_1,\lambda_0x_2,\cdots,\lambda_0x_r)=(x_1,x_2,\cdots,x_r)\begin{pmatrix}\lambda_0&&&\\&\lambda_0&&\\&&\ddots&\\&&&\lambda_0\end{pmatrix}.$$

故 $\varphi|_{V_0}$ 在 $\{x_1,x_2,\cdots,x_r\}$ 下的表示矩阵为 $\dim V_0$ 阶对角阵 $\mathrm{diag}\{\lambda_0,\lambda_0,\cdots,\lambda_0\}$. □

命题 6.14

设 A,B,C 是 n 阶矩阵, 其中 $C=AB-BA$. 若它们满足条件 $AC=CA, BC=CB$, 求证: C 的特征值全为零.

注 若将条件减弱为 $ABC=CAB, BAC=CBA$, 则上述结论不再成立. 原因如下:

如将条件减弱为如题所述, 则结论不再成立, 可参考下面的反例:

$$A=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad B=\begin{pmatrix}0&\frac{1}{2}\\\frac{1}{2}&0\end{pmatrix},\quad C=\begin{pmatrix}0&1\\1&0\end{pmatrix}.$$

由计算可得 $C=AB-BA, ABC=CAB, CBA=BAC$, 但 C 的特征值为 1 和 -1.

笔记 从下述的证法一不难看出: 只需要 $AC=CA$ 和 $BC=CB$ 这两个条件中的一个就能证明本题的结论.

证明 证法一: 由 $AC=CA$ 可知, 对任意的正整数 k,

$$C^k=C^{k-1}AB-C^{k-1}BA=A(C^{k-1}B)-(C^{k-1}B)A.$$

由迹的线性和交换性可得 $\mathrm{tr}(C^k)=0\ (k\geq 1)$, 再由幂零矩阵关于迹的充要条件可知 C 为幂零矩阵, 从而 C 的特征值全为零.

证法二:将 A,B,C 看成是 n 维复列向量空间 V 上的线性变换. 任取 C 的特征值 λ_0 及其特征子空间 V_0, 由 $AC=CA$, $BC=CB$ 以及命题 6.16可知, V_0 是 $A-$ 不变子空间, 也是 $B-$ 不变子空间. 将等式 $C=AB-BA$ 两边的线性变换同时限制在 V_0 上, 可得 V_0 上线性变换的等式 $C|_{V_0}=A|_{V_0}B|_{V_0}-B|_{V_0}A|_{V_0}$. 两边同时取迹, 由迹的线性和交换性及命题 6.13可知

$$\lambda_0\dim V_0=\mathrm{tr}(C|_{V_0})=\mathrm{tr}(A|_{V_0}B|_{V_0})-\mathrm{tr}(B|_{V_0}A|_{V_0})=0,$$

从而 $\lambda_0=0$, 结论得证.

证法三:注意到问题的条件和结论在同时相似变换:$A\mapsto P^{-1}AP$, $B\mapsto P^{-1}BP$, $C\mapsto P^{-1}CP$ 下不改变, 故不妨从一开始就假设 C 为 Jordan 标准型. 设 $C=\mathrm{diag}\{J_1,J_2,\cdots,J_k\}$, 其中 $\lambda_1,\lambda_2\cdots,\lambda_k$ 是 C 的全体不同特征值,J_i 是属于特征值 λ_i 的所有 Jordan 块拼成的根子空间分块. 由于 J_i 的特征值为 λ_i, 它们互不相同, 又 $AC=CA$, $BC=CB$, 故由例题 6.40可知,$A=\mathrm{diag}\{A_1,A_2,\cdots,A_k\}$, $B=\mathrm{diag}\{B_1,B_2,\cdots,B_k\}$ 和 C 一样也是分块对角矩阵. 于是我们有 $J_i=A_iB_i-B_iA_i$, 两边同取迹可得

$$n_i\lambda_i=\mathrm{tr}(J_i)=\mathrm{tr}(A_iB_i-B_iA_i)=\mathrm{tr}(A_iB_i)-\mathrm{tr}(B_iA_i)=0,$$

从而 $k=1$ 且 C 的特征值全为零. □

注 上述证法二中,A,B,C 在不变子空间 V_0 上的限制只能理解成线性变换在不变子空间上的限制, 而不是矩阵在

不变子空间上的限制.

推论 6.3

设 A, B, C 是 n 阶矩阵, 其中 $C = AB - BA$. 若它们满足条件 $AC = CA, BC = CB$, 求证: A, B, C 可同时上三角化.

证明 对阶数进行归纳. 由命题 6.14 证法二可知, C 的特征值全为 0, 其特征子空间 V_0 满足

$$A|_{V_0}B|_{V_0} - B|_{V_0}A|_{V_0} = C|_{V_0} = 0,$$

即 $A|_{V_0}, B|_{V_0}$ 乘法可交换. 由命题 6.19可知 $A|_{V_0}, B|_{V_0}$ 有公共的特征向量, 即存在 $0 \neq e_1 \in V_0$, 使得

$$Ae_1 = A|_{V_0}(e_1) = \lambda_1 e_1,\ Be_1 = B|_{V_0}(e_1) = \mu_1 e_1,\ Ce_1 = 0.$$

余下的证明完全类似于命题 6.20的证明, 请读者自行补充相关的细节. □

6.1.6 特征值的估计

定理 6.4 (第一圆盘定理)

设 $A = (a_{ij})$ 是 n 阶矩阵, 则 A 的特征值在复平面的下列圆盘中:

$$|z - a_{ii}| \leq R_i,\ 1 \leq i \leq n,$$

其中 $R_i = |a_{i1}| + \cdots + |a_{i,i-1}| + |a_{i,i+1}| + \cdots + |a_{in}|$.

注 该定理又称为 Gerschgorin 圆盘第一定理, 即戈氏圆盘第一定理. 上述圆盘称为戈氏圆盘.

证明

□

定理 6.5 (第二圆盘定理)

若 n 阶矩阵 A 的 n 个戈氏圆盘分成若干个连通区域, 其中某个连通区域恰含 k 个戈氏圆盘, 则有且仅有 k 个特征值落在该连通区域内 (若两个圆盘重合应计算重数, 若特征值为重根也要计算重数).

证明

□

例题 6.19 如果圆盘定理中有一个连通分支由两个圆盘外切组成, 证明: 每个圆盘除去切点的区域不可能同时包含两个特征值.

证明 设 $A = (a_{ij})$ 为 n 阶矩阵, $D_i : |z - a_{ii}| \leq R_i (1 \leq i \leq n)$ 是 A 的 n 个戈氏圆盘. 不妨设 A 的两个戈氏圆盘 D_1, D_2 外切并组成一个连通分支. 令

$$A(t) = \begin{pmatrix} a_{11} & ta_{12} & \cdots & ta_{1n} \\ ta_{21} & a_{22} & \cdots & ta_{2n} \\ \vdots & \vdots & & \vdots \\ ta_{n1} & ta_{n2} & \cdots & a_{nn} \end{pmatrix},$$

由第一圆盘定理, $A(t)$ 的特征值落在下列圆盘中:

$$tD_i : |z - a_{ii}| \leq tR_i,\ 1 \leq i \leq n.$$

由于当 $0 \leq t < 1$ 时, $A(t)$ 的特征值是关于 t 的连续函数, 故 $A(t)$ 的特征值 $\lambda_i(t)$ 从 D_i 的圆心开始, 始终在圆盘 $tD_i(1 \leq i \leq n)$ 中连续变动. 注意此时 tD_1, tD_2 不相交, 它们是两个连通分支, 于是特征值 $\lambda_i(t)$ 落在 $tD_i(i = 1, 2)$ 中. 最后当 $t = 1$ 时, A 的特征值 $\lambda_1 = \lambda_1(1)$ 落在 D_1 中, 特征值 $\lambda_2 = \lambda_2(1)$ 落在 D_2 中. 因此, λ_1, λ_2 不可能同时落

在 D_1 或 D_2 除去切点的区域中. □

例题 6.20 设 $A=(a_{ij})$ 为 n 阶矩阵, 证明: 存在正数 δ, 使得对任意的 $s\in(0,\delta)$, 下列矩阵均有 n 个不同的特征值:

$$A(s)=\begin{pmatrix} a_{11}+s & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}+s^2 & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}+s^n \end{pmatrix}.$$

证明 先证当 s 充分大时, $A(s)$ 有 n 个不同的特征值. 由第一圆盘定理, $A(s)$ 的特征值落在下列戈氏圆盘中:

$$D_i:|z-a_{ii}-s^i|\le R_i=\sum_{j=1,j\ne i}^{n}|a_{ij}|,\ 1\le i\le n.$$

取 s 充分大, 使得 $s^n\gg s^{n-1}\gg\cdots\gg s$. 注意到 R_i 的值固定, 故 D_i 的圆心之间的距离大于半径 R_i, 从而 D_i 互不相交, 各自构成了一个连通分支. 再由第二圆盘定理, 每个连通分支 D_i 中有且仅有一个特征值, 于是 $A(s)$ 有 n 个不同的特征值.

设 $f_s(\lambda)=|\lambda I_n-A(s)|$ 是 $A(s)$ 的特征多项式, 则其判别式 $\Delta(f_s(\lambda))$ 是关于 s 的多项式. 由前面的讨论可知, 当 s 充分大时, $f_s(\lambda)$ 无重根, 从而 $\Delta(f_s(\lambda))\ne 0$, 即 $\Delta(f_s(\lambda))$ 是关于 s 的非零多项式. 若 $\Delta(f_s(\lambda))$ 的所有复根都是零, 则任取一个正数 δ; 若 $\Delta(f_s(\lambda))$ 的复根不全为零, 则可取 δ 为 $\Delta(f_s(\lambda))$ 的非零复根的模长的最小值. 于是对任意的 $s\in(0,\delta)$, s 都不是 $\Delta(f_s(\lambda))$ 的根, 即 $\Delta(f_s(\lambda))\ne 0$, 从而 $f_s(\lambda)$ 都无重根, 即 $A(s)$ 都有 n 个不同的特征值. □

6.2 由乘法交换性诱导的同时性质

命题 6.15 (矩阵乘法可交换的基本性质)

若两个矩阵或线性变换 A,B 乘法可交换, 即 $AB=BA$, 则有 $(AB)^m=A^mB^m$, $f(A)g(B)=g(B)f(A)$ 以及二项式定理

$$(A+B)^m=A^m+\mathrm{C}_m^1A^{m-1}B+\cdots+\mathrm{C}_m^{m-1}AB^{m-1}+B^m$$

等成立, 其中 $m\ge 1$, $f(x),g(x)$ 为多项式.

特别地, 一个矩阵或线性变换 A 一定与其自身可交换, 从而也满足 $f(A)g(A)=g(A)f(A)$, 其中 $f(x),g(x)$ 为多项式.

证明 证明是显然的. □

6.2.1 特征子空间互为不变子空间

命题 6.16 (特征子空间互为不变子空间)

1. 设 φ,ψ 是复线性空间 V 上乘法可交换的线性变换, 即 $\varphi\psi=\psi\varphi$, 求证: φ 的特征子空间是 ψ 的不变子空间, ψ 的特征子空间是 φ 的不变子空间.
2. 若 n 阶复矩阵 A,B 乘法可交换, 即 $AB=BA$, 则 A,B 的特征子空间互为不变子空间.

注 这个命题的结论对一般的数域是不成立的. 例如, $A=I_2$, $B=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$, 显然 A,B 乘法可交换, 但它们在有理数域或实数域上没有公共的特征向量. 事实上, B 在有理数域或实数域上都没有特征值 (它的特征值是 $\pm\mathrm{i}$), 从而也没有特征向量, 所以更谈不上公共的特征向量了. 为了这个命题的结论推广到数域 $\mathbb{F}$ 上, 我们必须假设 A,B 的特征值都在 $\mathbb{F}$ 中.

证明

1. 由代数基本定理以及线性方程组的求解理论可知, $n(n \geq 1)$ 维复线性空间上的线性变换或 n 阶复矩阵至少有一个特征值和特征向量. 任取线性变换 φ 的一个特征值 λ_0, 设 V_0 是特征值 λ_0 的特征子空间, 则对任意的 $\alpha \in V_0$, 有
$$\varphi\psi(\alpha) = \psi\varphi(\alpha) = \psi(\lambda_0\alpha) = \lambda_0\psi(\alpha),$$
即 $\psi(\alpha) \in V_0$, 因此 V_0 是 ψ 的不变子空间. 同理可证 ψ 的特征子空间是 φ 的不变子空间.
2.

□

命题 6.17

设 V 为 n 维复线性空间, S 是 $\mathcal{L}(V)$ 的非空子集, 满足: S 中的全体线性变换没有非平凡的公共不变子空间. 设线性变换 φ 与 S 中任一线性变换乘法均可交换, 证明: φ 是纯量变换.

证明 任取 φ 的特征值 λ_0 及其特征子空间 V_0. 任取 $\psi \in S$, 则 $\varphi\psi = \psi\varphi$, 由命题 6.16可知 V_0 是 $\psi-$ 不变子空间, 从而是 S 中全体线性变换的公共不变子空间. 又 $V_0 \neq 0$(特征向量均非零), 故 $V_0 = V$, 从而 $\varphi = \lambda_0 I_V$ 为纯量变换. □

6.2.2 有公共的特征向量

命题 6.18

1. 设 φ, ψ 是复线性空间 V 上乘法可交换的线性变换, 求证: φ, ψ 至少有一个公共的 (复) 特征向量.
2. 若 n 阶复矩阵 A, B 乘法可交换, 即 $AB = BA$, 则 A, B 至少有一个公共的 (复) 特征向量.

注 这个命题的结论对一般的数域是不成立的. 例如, $A = I_2$, $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, 显然 A, B 乘法可交换, 但它们在有理数域或实数域上没有公共的特征向量. 事实上, B 在有理数域或实数域上都没有特征值 (它的特征值是 $\pm\mathrm{i}$), 从而也没有特征向量, 所以更谈不上公共的特征向量了. 为了这个命题的结论推广到数域 $\mathbb{F}$ 上, 我们必须假设 A, B 的特征值都在 $\mathbb{F}$ 中.

证明

1. 任取 φ 的特征值 λ_0 及其特征子空间 V_0, 由命题 6.16可知, V_0 是 $\psi-$ 不变子空间. 将线性变换 ψ 限制在 V_0 上, 由于 V_0 是维数大于零的复线性空间, 故由命题 6.1可知 $\psi|_{V_0}$ 至少有一个特征值 μ_0 及其特征向量 $\alpha \in V_0$, 从而 $\varphi(\alpha) = \lambda_0\alpha$, $\psi(\alpha) = \mu_0\alpha$, 于是 α 就是 φ, ψ 的公共特征向量.
2.

□

命题 6.19

1. 设 φ, ψ 是数域 $\mathbb{F}$ 上线性空间 V 上的乘法可交换的线性变换, 且 φ, ψ 的特征值都在 $\mathbb{F}$ 中, 求证: φ, ψ 的特征子空间互为不变子空间, 并且 φ, ψ 至少有一个公共的特征向量.
2. 若数域 $\mathbb{F}$ 上的 n 阶矩阵 A, B 乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 则 A, B 的特征子空间互为不变子空间, 并且 A, B 在 $\mathbb{F}^n$ 中至少有一个公共的特征向量.

证明

1. 由线性方程组的求解理论可知, 若数域 $\mathbb{F}$ 上的线性变换或 $\mathbb{F}$ 上的矩阵在 $\mathbb{F}$ 中有一个特征值, 则在 $\mathbb{F}$ 上的线性空间或 $\mathbb{F}$ 上的列向量空间中必存在对应的特征向量. 任取线性变换 φ 的一个特征值 $\lambda_0 \in \mathbb{F}$, 设 V_0 是特征值 λ_0 的特征子空间, 则对任意的 $\alpha \in V_0$, 有
$$\varphi\psi(\alpha) = \psi\varphi(\alpha) = \psi(\lambda_0\alpha) = \lambda_0\psi(\alpha),$$
即 $\psi(\alpha) \in V_0$, 因此 V_0 是 $\psi-$ 不变子空间. 取 V_0 的一组基并扩张为 V 的一组基, 则 ψ 在这组基下的表示矩阵

为分块对角矩阵 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$, 其中 A 是 $\psi|_{V_0}$ 在给定基下的表示矩阵, 于是 $|\lambda I_V - \psi| = |\lambda I - A||\lambda I - B|$. 因为 ψ 的特征值都在 $\mathbb{F}$ 中, 故 A 的特征值都在 $\mathbb{F}$ 中, 于是 $\psi|_{V_0}$ 的特征值都在 $\mathbb{F}$ 中. 任取 $\psi|_{V_0}$ 的一个特征值 $\mu_0 \in \mathbb{F}$ 及其特征向量 $\alpha \in V_0$, 则 $\varphi(\alpha) = \lambda_0\alpha, \psi(\alpha) = \mu_0\alpha$, 于是 α 就是 φ, ψ 的公共特征向量.

2.

□

6.2.3 可同时上三角化

命题 6.20 (矩阵的上三角化)

1. 设数域 $\mathbb{F}$ 上的 n 阶矩阵 A 的特征值都在 $\mathbb{F}$ 中, 求证: A 在 $\mathbb{F}$ 上可上三角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}AP$ 是上三角矩阵.
2. 设数域 $\mathbb{F}$ 上线性空间 V 上的线性变换 φ 的特征值都在 $\mathbb{F}$ 中, 则存在 V 的一组基, 使得 φ 在这组基下的表示矩阵是上三角矩阵.

♠

证明

1. 对阶数进行归纳. 当 $n = 1$ 时结论显然成立, 设对 $n-1$ 阶矩阵结论成立, 现对 n 阶矩阵 A 进行证明. 设 $\lambda_1 \in \mathbb{F}$ 是 A 的一个特征值, 则由线性方程组的求解理论可知, 存在特征向量 $e_1 \in \mathbb{F}^n$, 使得 $Ae_1 = \lambda_1 e_1$. 由基扩张定理, 可将 e_1 扩张为 $\mathbb{F}^n$ 的一组基 $\{e_1, e_2, \cdots, e_n\}$, 于是

$$(Ae_1, Ae_2, \cdots, Ae_n) = (e_1, e_2, \cdots, e_n)\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix},$$

其中 A_1 是 $\mathbb{F}$ 上的 $n-1$ 阶矩阵. 令 $P = (e_1, e_2, \cdots, e_n)$, 则 P 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 且由上式可得 $AP = P\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}$, 即 $P^{-1}AP = \begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}$. 由此可得 $|\lambda I_n - A| = (\lambda - \lambda_1)|\lambda I_{n-1} - A_1|$, 又 A 的特征值全在 $\mathbb{F}$ 中, 从而 A_1 的特征值也全在 $\mathbb{F}$ 中, 故由归纳假设, 存在 $\mathbb{F}$ 上的 $n-1$ 阶可逆矩阵 Q, 使得 $Q^{-1}A_1Q$ 是上三角矩阵. 令

$$R = P\begin{pmatrix} 1 & O \\ O & Q \end{pmatrix},$$

则 R 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 且

$$R^{-1}AR = \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix}^{-1}\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}\begin{pmatrix} 1 & O \\ O & Q \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \\ O & Q^{-1}A_1Q \end{pmatrix}$$

是上三角矩阵.

2.

□

命题 6.21

1. 设 A, B 是数域 $\mathbb{F}$ 上的 n 阶矩阵, 满足: $AB = BA$ 且 A, B 的特征值都在 $\mathbb{F}$ 中, 求证: A, B 在 $\mathbb{F}$ 上可同时上三角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是上三角矩阵.
2. 设数域 $\mathbb{F}$ 上线性空间 V 上的线性变换 φ, ψ 乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 则存在 V 的一组基, 使得 φ, ψ 在这组基下的表示矩阵都是上三角矩阵.

♠

证明

1. 对阶数进行归纳. 当 $n = 1$ 时结论显然成立, 设对 $n-1$ 阶矩阵结论成立, 现对 n 阶矩阵进行证明. 因为 $AB = BA$ 且 A, B 的特征值都在 $\mathbb{F}$ 中, 故由命题 6.19可知, A, B 有公共的特征向量 $e_1 \in \mathbb{F}^n$, 不妨设

$$Ae_1 = \lambda_1 e_1, \ Be_1 = \mu_1 e_1,$$

其中 $\lambda_1,\mu_1 \in \mathbb{F}$ 分别是 A,B 的特征值. 由基扩张定理, 可将 e_1 扩张为 $\mathbb{F}^n$ 的一组基 $\{e_1,e_2,\cdots,e_n\}$. 令 $P=(e_1,e_2,\cdots,e_n)$, 则 P 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 从而有

$$\begin{aligned} A(e_1,e_2,\cdots,e_n)=(e_1,e_2,\cdots,e_n)\begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix} \Leftrightarrow AP=P\begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix} \Leftrightarrow P^{-1}AP=\begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix}, \\ B(e_1,e_2,\cdots,e_n)=(e_1,e_2,\cdots,e_n)\begin{pmatrix}\lambda_1 & * \\ O & B_1\end{pmatrix} \Leftrightarrow BP=P\begin{pmatrix}\lambda_1 & * \\ O & B_1\end{pmatrix} \Leftrightarrow P^{-1}BP=\begin{pmatrix}\lambda_1 & * \\ O & B_1\end{pmatrix}. \end{aligned} \tag{6.5}$$

其中 A_1,B_1 是 $\mathbb{F}$ 上的 $n-1$ 阶矩阵. 由 $AB=BA$ 及(6.5)式可得到

$$\begin{aligned} &\left(P^{-1}AP\right)\left(P^{-1}BP\right)=P^{-1}ABP=P^{-1}BAP=\left(P^{-1}BP\right)\left(P^{-1}AP\right) \\ \Leftrightarrow &\begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix}\begin{pmatrix}\lambda_1 & * \\ O & B_1\end{pmatrix}=\begin{pmatrix}\lambda_1 & * \\ O & B_1\end{pmatrix}\begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix} \\ \Leftrightarrow &\begin{pmatrix}\lambda_1 & * \\ O & A_1B_1\end{pmatrix}=\begin{pmatrix}\lambda_1 & * \\ O & B_1A_1\end{pmatrix} \end{aligned}$$

从而 $A_1B_1=B_1A_1$. 又由(6.5)式可得

$$|\lambda I_n-A|=|\lambda-\lambda_1|\,|\lambda I_{n-1}-A_1|,\quad |\lambda I_n-B|=|\lambda-\lambda_1|\,|\lambda I_{n-1}-B_1|.$$

因此 A_1,B_1 的特征值也是 A,B 的特征值. 又由于 A,B 的特征值都在 $\mathbb{F}$ 中, 故 A_1,B_1 的特征值都在 $\mathbb{F}$ 中. 故由归纳假设, 存在 $\mathbb{F}$ 上的 $n-1$ 阶可逆矩阵 Q, 使得 $Q^{-1}A_1Q$ 和 $Q^{-1}B_1Q$ 都是上三角矩阵. 令

$$R=P\begin{pmatrix}1 & O \\ O & Q\end{pmatrix},$$

则 R 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 且

$$\begin{aligned} R^{-1}AR&=\begin{pmatrix}1 & O \\ O & Q\end{pmatrix}^{-1}\begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix}\begin{pmatrix}1 & O \\ O & Q\end{pmatrix}=\begin{pmatrix}\lambda_1 & * \\ O & Q^{-1}A_1Q\end{pmatrix}, \\ R^{-1}BR&=\begin{pmatrix}1 & O \\ O & Q\end{pmatrix}^{-1}\begin{pmatrix}\mu_1 & * \\ O & B_1\end{pmatrix}\begin{pmatrix}1 & O \\ O & Q\end{pmatrix}=\begin{pmatrix}\mu_1 & * \\ O & Q^{-1}B_1Q\end{pmatrix} \end{aligned}$$

都是上三角矩阵.

2.

□

6.2.4 可同时对角化

命题 6.22

1. 设 φ,ψ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 上的线性变换, 满足: $\varphi\psi=\psi\varphi$ 且 φ,ψ 都可对角化, 求证: φ,ψ 可同时对角化, 即存在 V 的一组基, 使得 φ,ψ 在这组基下的表示矩阵都是对角矩阵.
2. 设 A,B 是数域 $\mathbb{F}$ 上的 n 阶矩阵, 满足: $AB=BA$ 且 A,B 都在 $\mathbb{F}$ 上可对角化, 则 A,B 在 $\mathbb{F}$ 上可同时对角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是对角矩阵.

证明

1. 对空间维数进行归纳. 当 $n=1$ 时结论显然成立, 设对维数小于 n 的线性空间结论成立, 现对 n 维线性空间进行证明. 设 φ 的全体不同特征值为 $\lambda_1,\cdots,\lambda_s\in\mathbb{F}$, 对应的特征子空间分别为 $V_1,\cdots,V_s$, 则由 φ 可对角化可知

$$V=V_1\oplus\cdots\oplus V_s.$$

若 $s=1$, 则 $\varphi=\lambda_1 I_V$ 为纯量变换, 此时只要取 V 的一组基, 使得 ψ 在这组基下的表示矩阵为对角矩阵, 则 φ 在这组基下的表示矩阵为 $\lambda_1 I_n$, 结论成立. 若 $s>1$, 则 $\dim V_i<n$. 注意到 $\varphi\psi=\psi\varphi$ 且 φ,ψ 的特征值都

在 $\mathbb{F}$ 中, 由命题 6.16可知 V_i 都是 $\psi-$ 不变子空间. 考虑线性变换的限制 $\varphi|_{V_i}, \psi|_{V_i}$: 它们乘法可交换, 且由可对角化线性变换的性质可知它们都可对角化, 故由归纳假设可知, $\varphi|_{V_i}, \psi|_{V_i}$ 可同时对角化, 即存在 V_i 的一组基, 使得 $\varphi|_{V_i}, \psi|_{V_i}$ 在这组基下的表示矩阵都是对角矩阵. 将 V_i 的基拼成 V 的一组基, 则 φ, ψ 在这组基下的表示矩阵都是对角矩阵, 即 φ, ψ 可同时对角化.

2.

□

6.2.5 个数的推广

命题 6.23

设数域 $\mathbb{F}$ 上的 n 阶矩阵 $A_1, A_2, \cdots, A_m$ 两两乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 求证: 它们在 $\mathbb{F}^n$ 中至少有一个公共的特征向量.

证明 对 m 进行归纳, $m = 2$ 时就是命题 6.18. 设矩阵个数小于 m 时结论成立, 现证 m 个矩阵的情形. 将所有的 A_i 都看成是列向量空间 $\mathbb{F}^n$ 上的线性变换, 任取 A_1 的一个特征值 $\lambda_1 \in \mathbb{F}$ 及其特征子空间 $V_1 \subseteq \mathbb{F}^n$. 注意到 $A_1A_i = A_iA_1$, 故由命题 6.16可知, V_1 是 $A_2, \cdots, A_m$ 的不变子空间. 将 $A_2, \cdots, A_m$ 限制在 V_1 上, 它们仍然两两乘法可交换且特征值都在 $\mathbb{F}$ 中, 故由归纳假设可得 $A_2|_{V_1}, \cdots, A_m|_{V_1}$ 有公共的特征向量 $\alpha \in V_1$. 注意到 α 也是 A_1 的特征向量, 于是 α 是 $A_1, A_2, \cdots, A_m$ 的公共特征向量. □

命题 6.24

设数域 $\mathbb{F}$ 上的 n 阶矩阵 $A_1, A_2, \cdots, A_m$ 两两乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 求证: 它们在 $\mathbb{F}$ 上可同时上三角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}A_iP (1 \le i \le m)$ 都是上三角矩阵.

证明 完全类似于命题 6.21的证明, 其中利用命题 6.23得到 $A_1, A_2, \cdots, A_m$ 的公共特征向量, 请读者自行补充相关的细节. □

命题 6.25

设数域 $\mathbb{F}$ 上的 n 阶矩阵 $A_1, A_2, \cdots, A_m$ 两两乘法可交换, 且它们都在 $\mathbb{F}$ 上可对角化, 求证: 它们在 $\mathbb{F}$ 上可同时对角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}A_iP (1 \le i \le m)$ 都是对角矩阵.

证明 若 A_i 都是纯量矩阵, 则结论显然成立. 以下不妨设 A_1 不是纯量矩阵, 余下的证明完全类似于命题 6.22的证明, 请读者自行补充相关的细节. □

例题 6.21 设 A, B 都是 n 阶矩阵且 $AB = BA$. 若 A 是幂零矩阵, 求证: $|A + B| = |B|$.

证明 证法一: 由命题 6.21可知, A, B 可同时上三角化, 即存在可逆矩阵 P, 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是上三角矩阵. 因为上三角矩阵的主对角元是矩阵的特征值, 而幂零矩阵的特征值全为零, 所以 $|P^{-1}AP + P^{-1}BP| = |P^{-1}BP|$, 即有 $|A + B| = |B|$.

证法二: 先假设 B 是可逆矩阵, 则 $|A + B| = |I_n + AB^{-1}||B|$, 只要证明 $|I_n + AB^{-1}| = 1$ 即可. 由 $AB = BA$ 可知 $AB^{-1} = B^{-1}A$, 再由 A 是幂零矩阵容易验证 AB^{-1} 也是幂零矩阵, 从而其特征值全为零. 因此 $I_n + AB^{-1}$ 的特征值全为 1, 故 $|I_n + AB^{-1}| = 1$.

对于一般的矩阵 B, 可取到一列有理数 $t_k \to 0$, 使得 $t_kI_n + B$ 是可逆矩阵. 由可逆情形的证明可得 $|A + t_kI_n + B| = |t_kI_n + B|$. 注意到上式两边都是 t_k 的多项式, 从而关于 t_k 连续. 将上式两边同时取极限, 令 $t_k \to 0$, 即得结论. □

6.3 矩阵相似和可对角化的计算

6.3.1 相似初等变换及其应用

命题 6.26 (相似初等变换)

设 A 为 n 阶矩阵, 容易验证以下 3 种变换都是相似变换, 称为**相似初等变换**:

1. 对换 A 的第 i 行与第 j 行, 再对换第 i 列与第 j 列;
2. 将 A 的第 i 行乘以非零常数 c, 再将第 i 列乘以 c^{-1};
3. 将 A 的第 i 行乘以常数 c 加到第 j 行上, 再将第 j 列乘以 $-c$ 加到第 i 列上.

设 A 是具有相同行列分块方式的分块矩阵, 容易验证以下 3 种变换都是相似变换, 称为**相似分块初等变换**:

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;
2. 将 A 的第 i 分块行左乘非异阵 M, 再将第 i 分块列右乘 M^{-1};
3. 将 A 的第 i 分块行左乘矩阵 M 加到第 j 分块行上, 再将第 j 分块列右乘 $-M$ 加到第 i 分块列上.

容易验证: **任一相似变换都是若干次相似初等变换的复合**.

推论 6.4

设 $A = \mathrm{diag}\{A_1, A_2, \cdots, A_m\}$ 是分块对角矩阵, 其中 A_i 都是方阵, 求证: $\mathrm{diag}\{A_1, A_2, \cdots, A_m\}$ 相似于 $\mathrm{diag}\{A_{i_1}, A_{i_2}, \cdots, A_{i_m}\}$, 其中 $A_{i_1}, A_{i_2}, \cdots, A_{i_m}$ 是 $A_1, A_2, \cdots, A_m$ 的一个排列.

证明 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. 这是一个相似变换, 变换的结果是将 A 的 (i,i) 分块和第 (j,j) 分块对换了位置. 又任一排列都可以通过若干次对换来实现, 因此 $\mathrm{diag}\{A_1, A_2, \cdots, A_m\}$ 和 $\mathrm{diag}\{A_{i_1}, A_{i_2}, \cdots, A_{i_m}\}$ 相似. □

命题 6.27

若分块矩阵 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 满足 $\mathrm{r}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \mathrm{r}(A)$, 则 $\mathrm{r}\begin{pmatrix} A & B \end{pmatrix} = \mathrm{r}\begin{pmatrix} A \\ C \end{pmatrix} = \mathrm{r}(A)$.

证明 由条件可得

$$\mathrm{r}(A) \leqslant \mathrm{r}\begin{pmatrix} A & B \end{pmatrix} = \mathrm{r}\begin{pmatrix} A \\ C \end{pmatrix} \leqslant \mathrm{r}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \mathrm{r}(A).$$

故 $\mathrm{r}\begin{pmatrix} A & B \end{pmatrix} = \mathrm{r}\begin{pmatrix} A \\ C \end{pmatrix} = \mathrm{r}(A)$. □

例题 6.22 设 n 阶方阵 A, B 满足 $\mathrm{r}(ABA) = \mathrm{r}(B)$, 求证: AB 与 BA 相似.

注 不妨设 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 的原因: 假设当 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 时, 结论已经成立. 则对于一般的满足条件的矩阵 A, 设 $\mathrm{r}(A) = r$, 则存在可逆矩阵 P, Q, 使得 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$. 再根据条件可知

$$\mathrm{r}((PAQ)(Q^{-1}BP^{-1})(PAQ)) = \mathrm{r}(PABAQ) = \mathrm{r}(ABA) = \mathrm{r}(B) = \mathrm{r}(Q^{-1}BP^{-1}).$$

从而由假设可知,$(PAQ)(Q^{-1}BP^{-1})$ 与 $(Q^{-1}BP^{-1})(PAQ)$ 相似, 即 $PABP^{-1}$ 与 $Q^{-1}BAQ$ 相似. 又注意到 $PABP^{-1}$ 与 AB 相似,$Q^{-1}BAQ$ 与 BA 相似, 因此 AB 与 BA 也相似. 故不妨设成立.

证明 设 P, Q 为 n 阶非异阵, 使得 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 其中 $r = \mathrm{r}(A)$. 注意到问题的条件和结论在相抵变换: $A \mapsto PAQ, B \mapsto Q^{-1}BP^{-1}$ 下保持不变, 故不妨从一开始就假设 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 是相抵标准型. 设 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ 为

对应的分块, 则由 $\mathrm{r}(\boldsymbol{ABA}) = \mathrm{r}(\boldsymbol{B})$ 可得 $\mathrm{r}\begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix} = \mathrm{r}(\boldsymbol{B}_{11})$. 从而由命题 6.27可得 $\mathrm{r}\begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \end{pmatrix} = \mathrm{r}(\boldsymbol{B}_{11})$ 以及 $\mathrm{r}\begin{pmatrix} \boldsymbol{B}_{11} \\ \boldsymbol{B}_{21} \end{pmatrix} = \mathrm{r}(\boldsymbol{B}_{11})$, 再由定理 3.15可知存在矩阵 $\boldsymbol{M}, \boldsymbol{N}$, 使得 $\boldsymbol{B}_{11}\boldsymbol{N} = \boldsymbol{B}_{12}, \boldsymbol{M}\boldsymbol{B}_{11} = \boldsymbol{B}_{21}$.

将 $\boldsymbol{AB} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$ 的第二分块行左乘 $\boldsymbol{N}$ 加到第一分块行, 再将第一分块列右乘 $-\boldsymbol{N}$ 加到第二分块列, 于是 $\boldsymbol{AB}$ 相似于 $\begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$. 将 $\boldsymbol{BA} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{O} \\ \boldsymbol{B}_{21} & \boldsymbol{O} \end{pmatrix}$ 的第一分块行左乘 $-\boldsymbol{M}$ 加到第二分块行, 再将第二分块列右乘 $\boldsymbol{M}$ 加到第一分块列, 于是 $\boldsymbol{BA}$ 相似于 $\begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$. 因此, $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 相似. □

6.3.2 利用相似不变量来判定矩阵不相似

定义 6.10 (相似不变量)

相似的矩阵具有相同的**迹**、**行列式**、**特征多项式**和**极小多项式**等, 故它们被称为矩阵相似关系下的不变量, 也称为**相似不变量**. ♣

注 若两个矩阵的相似不变量不相同, 则它们必不相似. 利用这种方法来判断两个矩阵不相似是很简便的.

笔记 由矩阵迹、行列式、特征多项式的性质易知, 相似的矩阵具有相同的迹、行列式、特征多项式. 由命题??可知, 相似的矩阵具有相同的极小多项式. 因此上述定义是良定义的.

例题 6.23 设 $\boldsymbol{A}, \boldsymbol{B}$ 为 n 阶方阵, 求证: $\boldsymbol{AB} - \boldsymbol{BA}$ 必不相似于 $k\boldsymbol{I}_n$, 其中 k 是非零常数.

证明 注意到 $\mathrm{tr}(\boldsymbol{AB} - \boldsymbol{BA}) = 0, \mathrm{tr}(k\boldsymbol{I}_n) = nk \neq 0$, 又矩阵的迹是相似不变量, 因此 $\boldsymbol{AB} - \boldsymbol{BA}$ 和 $k\boldsymbol{I}_n$ 必不相似. □

例题 6.24 设 $\boldsymbol{A}, \boldsymbol{B}$ 为 n 阶正交矩阵, 且线性方程组 $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{x} = 0$ 的解空间维数是奇数, 求证: $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 必不相似.

证明 由假设可知 $n - \mathrm{r}(\boldsymbol{A} + \boldsymbol{B})$ 为奇数, 再由命题??可知 $|\boldsymbol{A}| = -|\boldsymbol{B}| \neq 0$, 又矩阵的行列式是相似不变量, 因此 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 必不相似. □

6.3.3 过渡矩阵 P 的计算

首先, 我们介绍一下当矩阵相似于对角矩阵时求过渡矩阵的方法. 设 n 阶矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 可逆矩阵 $\boldsymbol{P} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)$ 为其列分块, 且

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

则

$$\boldsymbol{AP} = \boldsymbol{P}\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

即

$$(\boldsymbol{A\alpha}_1, \boldsymbol{A\alpha}_2, \cdots, \boldsymbol{A\alpha}_n) = (\lambda_1\boldsymbol{\alpha}_1, \lambda_2\boldsymbol{\alpha}_2, \cdots, \lambda_n\boldsymbol{\alpha}_n),$$

于是 $\boldsymbol{A\alpha}_i = \lambda_i\boldsymbol{\alpha}_i$, 这表明 $\boldsymbol{\alpha}_i$ 就是属于特征值 λ_i 的特征向量. 因此 $\boldsymbol{P}$ 的 n 个列向量就是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量. 注意: 因为特征向量不唯一, 所以过渡矩阵 $\boldsymbol{P}$ 也不唯一. 另外, $\boldsymbol{P}$ 的第 i 个列向量对应于 $\boldsymbol{A}$ 的第 i 个特征值.

例题 6.25 设三阶矩阵 $\boldsymbol{A}$ 的特征值为 $1, 1, 4$, 对应的特征向量依次为 $(2, 1, 0)', (-1, 0, 1)', (0, 1, 1)'$, 试求矩阵 $\boldsymbol{A}$.

解 容易验证 $\boldsymbol{A}$ 的这 3 个特征向量线性无关, 故 $\boldsymbol{A}$ 必相似于对角矩阵, 即有

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

根据上面的分析, 有

$$\boldsymbol{P}=\begin{pmatrix}2&-1&0\\1&0&1\\0&1&1\end{pmatrix},$$

于是

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

□

6.3.4 可对角化判定的计算

例题 6.26 已知矩阵 $\boldsymbol{A}=\begin{pmatrix}1&-1&1\\2&x&-2\\-3&-3&y\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}2&0&0\\0&2&0\\0&0&z\end{pmatrix}$ 相似.

(1) 求 x,y,z 的值;

(2) 求一个满足 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$ 的可逆矩阵 $\boldsymbol{P}$.

解 (1) 显然 $z\neq 2$, 否则由 $\boldsymbol{A}$ 相似于 $2\boldsymbol{I}_3$ 可知 $\boldsymbol{A}=2\boldsymbol{I}_3$, 矛盾. 于是 $\boldsymbol{A}$ 的特征值为 2 (2 重), z (1 重). 因为 $\boldsymbol{A}$ 可对角化, 所以特征值 2 的几何重数也等于 2, 即线性方程组 $(A-2I_3)x=O$ 的解空间维数也是 2, 也即 $3-\mathrm{r}(\boldsymbol{A}-2\boldsymbol{I}_3)=2$. 故有

$$\mathrm{r}(\boldsymbol{A}-2\boldsymbol{I}_3)=\mathrm{r}\begin{pmatrix}-1&-1&1\\2&x-2&-2\\-3&-3&y-2\end{pmatrix}=1,$$

由此可得 $x=4,y=5$. 再由矩阵的迹等于特征值之和可得 $10=\mathrm{tr}(\boldsymbol{A})=4+z$, 故 $z=6$.

(2) 通过求解线性方程组 $(\lambda I_3-A)x=O(\lambda=2,6)$ 计算可得: 特征值 2 的两个线性无关的特征向量为 $\boldsymbol{\alpha}_1=(-1,1,0)'$, $\boldsymbol{\alpha}_2=(1,0,1)'$; 特征值 6 的特征向量为 $\boldsymbol{\alpha}_3=(1,-2,3)'$. 因此

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

□

例题 6.27 设 $\boldsymbol{A}=\begin{pmatrix}3&2&-2\\-k&-1&k\\4&2&-3\end{pmatrix}$, 当 k 为何值时, 存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是对角矩阵? 求出 $\boldsymbol{P}$ 和对角矩阵.

解 经计算可得 $|\lambda\boldsymbol{I}_3-\boldsymbol{A}|=(\lambda-1)(\lambda+1)^2$, 因此 $\boldsymbol{A}$ 的特征值为 1 (1 重), -1 (2 重). 对单特征值 1, 其几何重数与代数重数必相等 (因为几何重数总小于代数重数且一个特征值必存在一个对应的特征向量); 因此要使 $\boldsymbol{A}$ 可对角化, 特征值 -1 的几何重数必须等于 2 才行, 故有

$$\mathrm{r}(\boldsymbol{A}+\boldsymbol{I}_3)=\mathrm{r}\begin{pmatrix}4&2&-2\\-k&0&k\\4&2&-2\end{pmatrix}=1,$$

于是 $k=0$. 通过计算可得: 特征值 1 的特征向量为 $(1,0,1)'$; 特征值 -1 的两个线性无关的特征向量为 $(-1,2,0)',(1,0,2)'$.

因此

$$P=\begin{pmatrix}1&-1&1\\0&2&0\\1&0&2\end{pmatrix},P^{-1}AP=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix}.$$

□

6.3.5 可对角化矩阵的应用

例题 6.28 设矩阵 $A=\begin{pmatrix}1&-1&1\\2&4&-2\\-3&-3&5\end{pmatrix}$, 求 A^n.

笔记 本题中的矩阵是一个可对角化矩阵, 因此可以使用下列方法: 先求出可逆矩阵 P, 使得 $P^{-1}AP=B$ 是对角矩阵. 因为对角矩阵的幂很容易求出, 故由 $A^n=PB^nP^{-1}$ 即可得到结果.

解 经计算可得 $|\lambda I_3-A|=(\lambda-2)^2(\lambda-6)$, 因此 A 的特征值为 2 (2 重), 6 (1 重). 通过计算可得: 特征值 2 有两个线性无关的特征向量 $(-1,1,0)',(1,0,1)'$; 特征值 6 有特征向量 $(1,-2,3)'$, 于是 A 有完全的特征向量系, 从而可对角化. 注意到

$$P=\begin{pmatrix}-1&1&1\\1&0&-2\\0&1&3\end{pmatrix},B=P^{-1}AP=\begin{pmatrix}2&0&0\\0&2&0\\0&0&6\end{pmatrix},$$

故

$$A^n=PB^nP^{-1}=\frac{1}{4}\begin{pmatrix}5\cdot2^n-6^n&2^n-6^n&6^n-2^n\\2\cdot6^n-2^{n+1}&2\cdot6^n+2^{n+1}&2^{n+1}-2\cdot6^n\\3\cdot2^n-3\cdot6^n&3\cdot2^n-3\cdot6^n&2^n+3\cdot6^n\end{pmatrix}.$$

□

例题 6.29 下列数列称为 Fibonacci 数列: $a_0=0,a_1=1,a_2=1,a_3=2,a_4=3,a_5=5,a_6=8,\cdots$, 通用递推式来表示为 $a_{n+2}=a_{n+1}+a_n$, 试求 Fibonacci 数列通项的显式表达式.

注 本例为用递推式定义的数列的通项提供了一个一般性的方法. 在三对角行列式中, 我们用技巧性相当高的方法求出了用下列递推式定义的数列的通项:

$$D_n=aD_{n-1}-bcD_{n-2},D_1=a,D_2=a^2-bc.$$

现在请读者自己用上面的方法来求出通项表达式.

证明 首先用矩阵来表示递推式:

$$\begin{pmatrix}a_{n+1}\\a_n\end{pmatrix}=\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}a_n\\a_{n-1}\end{pmatrix}.$$

令 $A=\begin{pmatrix}1&1\\1&0\end{pmatrix}$, 则

$$\begin{pmatrix}a_{n+1}\\a_n\end{pmatrix}=A\begin{pmatrix}a_n\\a_{n-1}\end{pmatrix}=A^2\begin{pmatrix}a_{n-1}\\a_{n-2}\end{pmatrix}=\cdots=A^n\begin{pmatrix}a_1\\a_0\end{pmatrix}=A^n\begin{pmatrix}1\\0\end{pmatrix}.$$

只要求出 A^n 就可以算出 a_n 来. A 的特征多项式为 $|\lambda I_2-A|=\lambda^2-\lambda-1$, 解得

$$\lambda_1=\frac{1+\sqrt5}{2},\lambda_2=\frac{1-\sqrt5}{2}.$$

对应于特征值 λ_1,λ_2 的特征向量分别是:

$$\alpha_1=\begin{pmatrix}\frac{1+\sqrt5}{2}\\1\end{pmatrix},\alpha_2=\begin{pmatrix}\frac{1-\sqrt5}{2}\\1\end{pmatrix}.$$

若记

$$P = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix},$$

则

$$P^{-1}AP = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}.$$

因此

$$A = P \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} P^{-1}, A^n = P \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}^n P^{-1}.$$

经计算可得

$$A^n = \frac{1}{\sqrt{5}} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} & \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} & \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1} \end{pmatrix},$$

由此可得

$$a_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \right).$$

□

6.4 可对角化的判定

6.4.1 可对角化的基本知识

定义 6.11 (可对角化线性变换)

若 n 维线性空间 V 上的线性变换 φ 在某组基 $\{e_1, e_2, \cdots, e_n\}$ 下的表示矩阵为对角阵:

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

则称 φ 为**可对角化线性变换**.

♣

定理 6.6 (线性变换可对角化的充要条件)

设 φ 是 n 维线性空间 V 上的线性变换, 则 φ 可对角化的充分必要条件是 φ 有 n 个线性无关的特征向量.

♡

证明 若 φ 是 V 上可对角化线性变换, 则可设 φ 在某组基 $\{e_1, e_2, \cdots, e_n\}$ 下的表示矩阵为对角阵:

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

此时 $\varphi(e_i) = \lambda_i e_i$, 即 $e_1, e_2, \cdots, e_n$ 是 φ 的特征向量, 于是 φ 有 n 个线性无关的特征向量.

反过来, 若 n 维线性空间 V 上的线性变换 φ 有 n 个线性无关的特征向量 $e_1, e_2, \cdots, e_n$, 则这组向量构成了 V 的一组基, 且 φ 在这组基下的表示矩阵显然是一个对角阵. □

定义 6.12 (可对角化矩阵)

设 A 是 n 阶矩阵, 若 A 相似于对角阵, 即存在可逆阵 P, 使 $P^{-1}AP$ 为对角阵, 则称 A 为**可对角化矩阵**. ♣

引理 6.1

设 A 是 n 阶矩阵,φ 是线性空间 V 上由矩阵 A 乘法诱导的线性变换, 即 $\varphi(\alpha) = A\alpha, \forall \alpha \in V$. 设 $\{e_1, e_2, \cdots, e_n\}$ 是 V 的一组标准基, 则 φ 在这组基下的矩阵就是 A. 证明:

(1) 矩阵 A 与线性变换 φ 的特征值相同;

(2) 矩阵 A 可对角化等价于线性变换 φ 可对角化. ♡

证明

(1) 若 λ 为矩阵 A 的特征值, 则存在 $\xi \in V$, 使得 $\varphi(\xi) = A\xi = \lambda\xi$, 因此矩阵 A 的特征值也是线性变换 φ 的特征值.

若 λ 为线性变换 φ 的特征值, 则存在 $\eta \in V$, 使得 $\varphi(\eta) = A\eta = \lambda\eta$, 因此线性变换 φ 的特征值也是矩阵 A 的特征值.

故矩阵 A 与线性变换 φ 的特征值相同.

(2) 若矩阵 A 可对角化, 则存在可逆阵 P, 使得 $P^{-1}AP$ 为对角矩阵.

从而 $(e_1, e_2, \cdots, e_n)P$ 的列向量也是 V 的一组基, 于是由命题 4.10 可知 φ 在这组基下的矩阵为 $P^{-1}AP$ 是对角矩阵, 故 φ 也可对角化.

若线性变换 φ 可对角化, 则存在 V 的一组基 $\{f_1, f_2, \cdots, f_n\}$, 使得 φ 在这组基下的矩阵 B 为对角矩阵. 设基 $\{e_1, e_2, \cdots, e_n\}$ 到基 $\{f_1, f_2, \cdots, f_n\}$ 的过渡矩阵为 G, 则由命题 4.10可知 $B = G^{-1}AG$. 因此矩阵 A 也可对角化.

故矩阵 A 可对角化等价于线性变换 φ 可对角化.

□

定理 6.7 (矩阵可对角化的充要条件)

设 A 是 n 阶矩阵, 则 A 可对角化的充分必要条件是 A 有 n 个线性无关的特征向量. ♡

证明 设 φ 是线性空间 V 上由矩阵 A 乘法诱导的线性变换.

若矩阵 A 有 n 个线性无关的特征值, 则由引理 6.1(1)可知线性变换 φ 也有相同的 n 个线性无关的特征值, 于是由定理 6.6可知线性变换 φ 可对角化, 从而再由引理 6.1(2) 可知矩阵 A 也可对角化.

若矩阵 A 可对角化, 则由引理 6.1(2)可知线性变换 φ 也可对角化, 从而由定理 6.6可知 φ 有 n 个线性无关的特征值, 于是由引理 6.1(1)可知矩阵 A 也有相同的 n 个线性无关的特征值. □

定理 6.8

若 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 为 n 维线性空间 V 上的线性变换 φ 的不同的特征值, 记 λ_i 的特征子空间为 $V_i (1 \leq i \leq k)$, 则

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

♡

证明 对 k 用数学归纳法. 若 $k = 1$, 结论显然成立. 现设对 $k-1$ 个不同的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_{k-1}$, 它们相应的特征子空间 $V_1, V_2, \cdots, V_{k-1}$ 之和是直和. 我们要证明 $V_1, V_2, \cdots, V_{k-1}, V_k$ 之和为直和, 这只需证明:

$$V_k \cap (V_1 + V_2 + \cdots + V_{k-1}) = 0. \tag{6.6}$$

即可. 设 $v \in V_k \cap (V_1 + V_2 + \cdots + V_{k-1})$, 则

$$v = v_1 + v_2 + \cdots + v_{k-1}, \tag{6.7}$$

其中 $v_i \in V_i(i=1,2,\cdots,k-1)$. 在(6.7)式两边作用 φ, 得

$$\varphi(v)=\varphi(v_1)+\varphi(v_2)+\cdots+\varphi(v_{k-1}). \tag{6.8}$$

但 $v,v_1,v_2,\cdots,v_{k-1}$ 都是 φ 的特征向量或零向量, 因此

$$\lambda_k v=\lambda_1 v_1+\lambda_2 v_2+\cdots+\lambda_{k-1}v_{k-1}. \tag{6.9}$$

在(6.8)式两边乘以 λ_k 减去(6.9)式得

$$0=(\lambda_k-\lambda_1)v_1+(\lambda_k-\lambda_2)v_2+\cdots+(\lambda_k-\lambda_{k-1})v_{k-1}.$$

由于 $v_1,v_2,\cdots,v_{k-1}$ 是直和, 因此 $(\lambda_k-\lambda_i)v_i=0$, 而 $\lambda_k-\lambda_i\neq 0$, 从而 $v_i=0(i=1,2,\cdots,k-1)$. 这就证明了(6.6)式.
□

推论 6.5

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 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 是线性变换 φ 的 k 个不同特征值, 由定理 6.8可知 $V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}$. 于是任取 $\alpha_i\in V_{\lambda_i}\ (1\leqslant i\leqslant k)$ 且 $\alpha_i\neq 0$, 假设 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 线性无关, 则存在一组不全为零的数 $b_1,b_2,\cdots,b_k$, 使得

$$b_1\alpha_1+b_2\alpha_2+\cdots+b_k\alpha_k=0.$$

不妨设 $b_1\neq 0$, 则

$$\alpha_1=\frac{b_2}{b_1}\alpha_2+\frac{b_3}{b_1}\alpha_3+\cdots+\frac{b_k}{b_1}\alpha_k\in V_{\lambda_1}\cap\left(V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}\right).$$

又由 $V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}$ 及直和的等价条件可知,

$$V_{\lambda_1}\cap\left(V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}\right)=\{0\},$$

从而 $\alpha_1=0$, 这与 $\alpha_i\neq 0\ (1\leqslant i\leqslant k)$ 矛盾! □

推论 6.6

若 n 维线性空间 V 上的线性变换 φ 有 n 个不同的特征值, 则 φ 必可对角化.

笔记 注意这个推论只是可对角化的充分条件而非必要条件, 比如说纯量变换 $\varphi=cI_V$ 当然可对角化, 但 φ 的 n 个特征值都是 c.

证明 设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是线性变换 φ 的 n 个不同特征值, 则任取 $\alpha_i\in V_{\lambda_i}\ (1\leqslant i\leqslant n)$, 由推论 6.5可知 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 一定线性无关. 从而由定理 6.6可知,φ 一定可对角化. □

定理 6.9 (线性变换可对角化的充要条件)

设 φ 是 n 维线性空间 V 上的线性变换, $\lambda_1,\lambda_2,\cdots,\lambda_k$ 是 φ 的全部不同的特征值, $V_i(i=1,2,\cdots,k)$ 是特征值 λ_i 的特征子空间, 则 φ 可对角化的充要条件是

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

证明 先证充分性. 设

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

分别取 V_i 的一组基 $\{e_{i1},e_{i2},\cdots,e_{it_i}\}(i=1,2,\cdots,k)$, 则由直和的等价条件 (4) 知这些向量拼成了 V 的一组基, 并且它们都是 φ 的特征向量. 因此 φ 有 n 个线性无关的特征向量, 从而定理 6.6可知 φ 可对角化.

再证必要性. 设 φ 可对角化, 则由定理 6.6可知 φ 有 n 个线性无关的特征向量 $\{e_1,e_2,\cdots,e_n\}$, 它们构成了 V 的一组基. 不失一般性, 可设这组基中前 t_1 个是关于特征值 λ_1 的特征向量; 接下去的 t_2 个是关于特征值 λ_2 的特征向量; $\cdots$; 最后 t_k 个是关于特征值 λ_k 的特征向量. 对任一 $\alpha\in V$, 设 $\alpha=a_1e_1+a_2e_2+\cdots+a_ne_n$, 则 α 可写成

$V_1, V_2, \cdots, V_k$ 中向量之和, 因此由定理 6.8可知

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

□

定义 6.13 (线性变换的几何重数与代数重数)

设 λ_0 是 n 维线性空间 V 上的线性变换 φ 的一个特征值, V_0 是属于 λ_0 的特征子空间, 称 $\dim V_0$ 为 λ_0 的**度数或几何重数**. λ_0 作为 φ 的特征多项式根的重数称为 λ_0 的**重数或代数重数**.

笔记 由线性映射的维数公式可知, 特征值 λ_0 的度数 $\dim V_0 = \dim \operatorname{Ker}(\lambda_0 I_V - \varphi) = n - r(\lambda_0 I_V - \varphi)$, 而特征值 λ_0 的重数则由特征多项式 $|\lambda I_V - \varphi|$ 的因式分解决定.

定义 6.14 (矩阵的几何重数与代数重数)

设 λ_0 是 n 阶方阵的 A 的一个特征值, V_0 是属于 λ_0 的特征子空间, 称 $\dim V_0$ 为 λ_0 的**度数或几何重数**.λ_0 作为 A 的特征多项式根的重数称为 λ_0 的**重数或代数重数**.

笔记 由线性方程组的理论可知, 特征值 λ_0 的度数 $\dim V_0 = n - r(\lambda_0 I_n - A)$, 若将 A 看作由矩阵 A 乘法诱导的 V 上的线性变换, 则由线性变换的维数公式可知 $\dim V_0 = \dim \operatorname{Ker}(\lambda_0 I_V - A) = n - r(\lambda_0 I_V - A)$. 而特征值 λ_0 的重数则由特征多项式 $|\lambda I_n - A|$ 的因式分解决定.

引理 6.2 (特征值的几何重数数总小于代数重数)

设 φ 是 n 维线性空间 V 上的线性变换, λ_0 是 φ 的一个特征值, 则 λ_0 的度数总是小于等于 λ_0 的重数.

证明 设特征值 λ_0 的重数为 m, 度数为 t, 又 V_0 是属于 λ_0 的特征子空间, 则 $\dim V_0 = t$. 设 $\{e_1, \cdots, e_t\}$ 是 V_0 的一组基. 由于 V_0 中的非零向量都是 φ 关于 λ_0 的特征向量, 故

$$\varphi(e_i) = \lambda_0 e_i, \quad i = 1, \cdots, t.$$

将 $\{e_1, \cdots, e_t\}$ 扩充为 V 的一组基, 记为 $\{e_1, \cdots, e_t, e_{t+1}, \cdots, e_n\}$, 则 φ 在这组基下的表示矩阵为

$$A = \begin{pmatrix} \lambda_0 I_t & * \\ O & B \end{pmatrix},$$

其中 B 是一个 $n-t$ 阶方阵. 因此, 线性变换 φ 的特征多项式具有如下形式:

$$|\lambda I_V - \varphi| = |\lambda I_n - A| = (\lambda - \lambda_0)^t |\lambda I_{n-t} - B|,$$

这表明 λ_0 的重数至少为 t, 即 $t \le m$. □

定义 6.15 (完全的特征向量系)

设 λ_0 是 φ (或 A) 的 m 重特征值, 即它是 φ (或 A) 的特征多项式的 m 重根. 此时若有 $m = \dim V_{\lambda_0}$, 即 λ_0 的代数重数和几何重数相等, 则称 λ_0 **有完全的特征向量系**. 若对 φ (或 A) 的任一特征值, 其代数重数和几何重数都相等, 则称 φ **(或 A) 有完全的特征向量系**.

定理 6.10 (线性变换可对角化的充要条件)

设 φ 是 n 维线性空间 V 上的线性变换, 则 φ 可对角化的充分必要条件是 φ 有完全的特征向量系.

证明 设 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 φ 的全部不同的特征值, 它们对应的特征子空间、重数和度数分别记为 $V_i, m_i, t_i (i = 1, 2, \cdots, k)$. 由重数的定义以及引理 6.2可知 $m_1 + m_2 + \cdots + m_k = n, t_i \le m_i, i = 1, 2, \cdots, k$.

由定理 6.9可知, 我们只要证明 φ 有完全的特征向量系当且仅当 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$.

若 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$, 则

$$n = \dim V = \dim(V_1 \oplus V_2 \oplus \cdots \oplus V_k)$$

$$= \dim V_1 + \dim V_2 + \cdots + \dim V_k$$
$$= \sum_{i=1}^{k} t_i \le \sum_{i=1}^{k} m_i = n,$$

因此 $t_i = m_i(i = 1, 2, \cdots, k)$, 即 φ 有完全的特征向量系. 反过来, 若 φ 有完全的特征向量系, 则

$$\dim(V_1 \oplus V_2 \oplus \cdots \oplus V_k) = \sum_{i=1}^{k} t_i = \sum_{i=1}^{k} m_i = n = \dim V,$$

又 $V_1 \oplus V_2 \oplus \cdots \oplus V_k \subset V$, 故 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ 成立. □

定理 6.11

设 A 为 n 阶复矩阵, 其全体特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_r$, 并且 $\lambda_i\,(1 \le i \le r)$ 的代数重数为 n_i, 则 $\sum_{i=1}^{r} n_i = n$. 若 A 可对角化, 则 A 一定相似于 $\operatorname{diag}\{A_1, A_2, \cdots, A_r\}$, 其中 $A_i = \operatorname{diag}\{\lambda_i, \lambda_i, \cdots, \lambda_i\}\,(1 \le i \le r)$ 并且阶数为 n_i. ♡

证明 由于 A 可对角化, 因此其特征值的代数重数等于几何重数. 记 V_i 为 λ_i 的特征子空间, 则任取 V_i 中一组基 $\{e_i1, e_i2, \cdots, e_{i,n_i}\}$. 由可对角化的判定条件 (3)及直和的等价条件可知,$\{e_i1, e_i2, \cdots, e_{i,n_i}\}\,(1 \le i \le r)$ 可以拼成 $\mathbb{C}^n$ 的一组基. 于是记 $P = (e_{11}, \cdots, e_{1,n_1}, \cdots, e_{r1}, \cdots, e_{r,n_r})$, 则 P 可逆, 并且

$$\begin{aligned} AP &= A\left(e_{11}, \cdots, e_{1,n_1}, \cdots, e_{r1}, \cdots, e_{r,n_r}\right) = \left(\lambda_1 e_{11}, \cdots, \lambda_1 e_{1,n_1}, \cdots, \lambda_r e_{r1}, \cdots, \lambda_r e_{r,n_r}\right) \\ &= \left(e_{11}, \cdots, e_{1,n_1}, \cdots, e_{r1}, \cdots, e_{r,n_r}\right) \operatorname{diag}\{A_1, A_2, \cdots, A_r\} = P\operatorname{diag}\{A_1, A_2, \cdots, A_r\}. \end{aligned}$$

故 $P^{-1}AP = \operatorname{diag}\{A_1, A_2, \cdots, A_r\}$. 结论得证. □

定理 6.12 (可对角化的判定条件)

判定 n 阶复矩阵 $\boldsymbol{A}$ (或 n 维复线性空间 V 上的线性变换 φ) 是否可对角化, 通常有以下 7 种方法:

(1) $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量;

(2) 若 $\boldsymbol{A}$ 有 n 个不同的特征值, 则 $\boldsymbol{A}$ 可对角化;

(3) $\boldsymbol{A}$ 可对角化的充要条件是 $\mathbb{C}^n$ 是 $\boldsymbol{A}$ 的特征子空间的直和;

(4) $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 有完全的特征向量系, 即对 $\boldsymbol{A}$ 的任一特征值, 其几何重数等于其代数重数;

(5) $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 的极小多项式无重根;

(6) $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 的 Jordan 块都是一阶的 (或 $\boldsymbol{A}$ 的初等因子都是一次多项式);

(7) 若 $\boldsymbol{A}$ 相似于实对称矩阵或复正规矩阵, 则 $\boldsymbol{A}$ 可对角化. ♡

注 上述第五、第六种方法将放在 §7.5 进行探讨, 另外命题 7.39也是可对角化判定准则的一个补充; 第七种方法将放在 §9.7.4 进行探讨; 本节主要阐述可对角化判定的前 4 种方法.

证明 **几何形式**:(即 n 维复线性空间 V 上的线性变换 φ 可对角化的条件)

(1) 证明见定理 6.6.

(2) 证明见定理 6.9.

(3)

(4) 证明见定理 6.10.

(5)

(6)

(7)

代数形式:(即 n 阶复矩阵可对角化的条件) 由上述几何形式的结论及引理 6.1立即得到证明. □

注 若要考虑数域 $\mathbb{F}$ 上的 n 阶矩阵 $\boldsymbol{A}$ (或 $\mathbb{F}$ 上 n 维线性空间 V 上的线性变换 φ) 在 $\mathbb{F}$ 上的可对角化问题, 那么首先需要验证 $\boldsymbol{A}$ (或 φ) 的特征值都在 $\mathbb{F}$ 中, 否则由可对角化的定义可知, $\boldsymbol{A}$ (或 φ) 在 $\mathbb{F}$ 上必不可对角化. 若假设 $\boldsymbol{A}$ (或 φ) 的特征值都在 $\mathbb{F}$ 中, 则 $\boldsymbol{A}$ (或 φ) 在 $\mathbb{F}$ 上的可对角化判定准则也是上述前 6 种方法. 因此, 为了突出重点, 本节总

是在复数域 $\mathbb{C}$ 上考虑可对角化问题. 请读者自行将某些例题推广到数域 $\mathbb{F}$ 的情形.

6.4.2 有 n 个线性无关的特征向量

寻找 $\boldsymbol{A}$ 的 n 个线性无关的特征向量, 等价于寻找 n 阶可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为对角矩阵.

命题 6.28 (循环矩阵一定可对角化)

求证: 复数域上 n 阶循环矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}$$

可对角化, 并求出它相似的对角矩阵及过渡矩阵.

笔记 这个命题实际上就是命题 2.3.

证明 设 $f(x)=a_1+a_2x+\cdots+a_nx^{n-1}, \omega_k=\cos\dfrac{2k\pi}{n}+\mathrm{i}\sin\dfrac{2k\pi}{n}(0\leq k\leq n-1)$, 则

$$\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}\begin{pmatrix} 1 \\ \omega_k \\ \vdots \\ \omega_k^{n-1} \end{pmatrix}=f(\omega_k)\begin{pmatrix} 1 \\ \omega_k \\ \vdots \\ \omega_k^{n-1} \end{pmatrix}.$$

这表明 $(1,\omega_k,\cdots,\omega_k^{n-1})'$ 是 $\boldsymbol{A}$ 的属于特征值 $f(\omega_k)$ 的特征向量. 令

$$\boldsymbol{P}=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega_1 & \cdots & \omega_{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \omega_1^{n-1} & \cdots & \omega_{n-1}^{n-1} \end{pmatrix},$$

由 Vandermonde 行列式可知 $|\boldsymbol{P}|\neq 0$, 从而这 n 个特征向量线性无关, 因此 $\boldsymbol{A}$ 可对角化, 且有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\operatorname{diag}\{f(1),f(\omega_1),\cdots,f(\omega_{n-1})\}.$$

□

例题 6.30 设 n 阶复矩阵 $\boldsymbol{A}$ 可对角化, 证明: 矩阵 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{A}^2 \\ \boldsymbol{A}^2 & \boldsymbol{A} \end{pmatrix}$ 也可对角化.

证明 证法一: 因为 $\boldsymbol{A}$ 可对角化, 故可设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量, 满足 $\boldsymbol{A}\boldsymbol{\alpha}_i=\lambda_i\boldsymbol{\alpha}_i(1\leq i\leq n)$. 注意到

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{A}^2 \\ \boldsymbol{A}^2 & \boldsymbol{A} \end{pmatrix}\begin{pmatrix} \boldsymbol{\alpha}_i \\ \boldsymbol{\alpha}_i \end{pmatrix}=(\lambda_i+\lambda_i^2)\begin{pmatrix} \boldsymbol{\alpha}_i \\ \boldsymbol{\alpha}_i \end{pmatrix}, \begin{pmatrix} \boldsymbol{A} & \boldsymbol{A}^2 \\ \boldsymbol{A}^2 & \boldsymbol{A} \end{pmatrix}\begin{pmatrix} \boldsymbol{\alpha}_i \\ -\boldsymbol{\alpha}_i \end{pmatrix}=(\lambda_i-\lambda_i^2)\begin{pmatrix} \boldsymbol{\alpha}_i \\ -\boldsymbol{\alpha}_i \end{pmatrix}.$$

通过定义不难验证 $\begin{pmatrix} \boldsymbol{\alpha}_i \\ \boldsymbol{\alpha}_i \end{pmatrix}, \begin{pmatrix} \boldsymbol{\alpha}_i \\ -\boldsymbol{\alpha}_i \end{pmatrix}(1\leq i\leq n)$ 是线性无关的, 因此 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{A}^2 \\ \boldsymbol{A}^2 & \boldsymbol{A} \end{pmatrix}$ 有 $2n$ 个线性无关的特征向量, 从而可对角化.

证法二: 容易验证 $\begin{pmatrix} I_n & I_n \\ I_n & -I_n \end{pmatrix}$ 的逆矩阵为 $\dfrac{1}{2}\begin{pmatrix} I_n & I_n \\ I_n & -I_n \end{pmatrix}$. 考虑如下相似变换:

$$\frac{1}{2}\begin{pmatrix} I_n & I_n \\ I_n & -I_n \end{pmatrix}\begin{pmatrix} A & A^2 \\ A^2 & A \end{pmatrix}\begin{pmatrix} I_n & I_n \\ I_n & -I_n \end{pmatrix}=\begin{pmatrix} A+A^2 & O \\ O & A-A^2 \end{pmatrix}.$$

由命题 6.31可知,$A+A^2, A-A^2$ 作为 A 的多项式也可对角化, 故原矩阵可对角化. 具体地, 设 P 为可逆矩阵, 使得

$P^{-1}AP=\Lambda$ 为对角矩阵，则

$$\begin{pmatrix} P^{-1} & O \\ O & P^{-1} \end{pmatrix}\begin{pmatrix} A+A^2 & O \\ O & A-A^2 \end{pmatrix}\begin{pmatrix} P & O \\ O & P \end{pmatrix}=\begin{pmatrix} \Lambda+\Lambda^2 & O \\ O & \Lambda-\Lambda^2 \end{pmatrix}$$

为对角矩阵，因此原矩阵可对角化. □

例题 6.31

1. 设 V 为 n 阶矩阵全体构成的线性空间，V 上的线性变换 φ 定义为 $\varphi(\boldsymbol{X})=\boldsymbol{AXA}$，其中 $\boldsymbol{A}\in V$. 证明：若 $\boldsymbol{A}$ 可对角化，则 φ 也可对角化.
2. 设 V 为 n 阶矩阵全体构成的线性空间，V 上的线性变换 φ 定义为 $\varphi(X)=AXA$，其中 $A\in V$. 证明：φ 可对角化的充要条件是 A 可对角化.

注 第 2 问是第 1 问的延拓.

证明

1. 证法一：设 $\boldsymbol{P}$ 为 n 阶可逆矩阵，使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$，则 $\boldsymbol{P}'\boldsymbol{A}'(\boldsymbol{P}')^{-1}=\boldsymbol{\Lambda}$，即 $\boldsymbol{A}'$ 也可对角化. 设

   $$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n),(\boldsymbol{P}')^{-1}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$$

   分别为两个矩阵的列分块，则

   $$\boldsymbol{A\alpha}_i=\lambda_i\boldsymbol{\alpha}_i,\boldsymbol{A}'\boldsymbol{\beta}_j=\lambda_j\boldsymbol{\beta}_j,1\le i,j\le n,$$

   且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 线性无关. 由命题 3.14可知，$\{\boldsymbol{\alpha}_i\boldsymbol{\beta}_j',1\le i,j\le n\}$ 是 V 中 n^2 个线性无关的矩阵. 注意到

   $$\varphi(\boldsymbol{\alpha}_i\boldsymbol{\beta}_j')=\boldsymbol{A\alpha}_i\boldsymbol{\beta}_j'\boldsymbol{A}=(\boldsymbol{A\alpha}_i)(\boldsymbol{A}'\boldsymbol{\beta}_j)'=\lambda_i\lambda_j\boldsymbol{\alpha}_i\boldsymbol{\beta}_j',$$

   故 φ 有 n^2 个线性无关的特征向量，从而可对角化.

   证法二：由于 A 可对角化，故存在可逆矩阵 P，使得 $P^{-1}AP=\Lambda$ 为对角矩阵. 由命题 6.55可知，φ 在基础矩阵这组基下的表示矩阵为 $A\otimes A'$，于是

   $$(P\otimes(P')^{-1})^{-1}(A\otimes A')(P\otimes(P')^{-1})=\Lambda\otimes\Lambda$$

   由矩阵的 Kronecker 积的基本性质 (11)可知 $\Lambda\otimes\Lambda$ 为对角矩阵，即 $A\otimes A'$ 可对角化，从而 φ 可对角化.
2. 充分性就是第 1 问，下证必要性. 用反证法，设 A 不可对角化，则存在可逆矩阵 P,Q，使得

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

   为 Jordan 标准型，其中 $r_1>1$. 设 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n),Q=(\beta_1,\beta_2,\cdots,\beta_n)$ 分别为两个矩阵的列分块，令 $U=L(\alpha_i\beta_j',1\le i,j\le r_1)$，则由命题 3.14可知 $\{\alpha_i\beta_j',1\le i,j\le r_1\}$ 是 U 的一组基. 经简单计算可得

   $$\begin{aligned}
   &\varphi(\alpha_1\beta_1')=\lambda_1^2\alpha_1\beta_1';\\
   &\varphi(\alpha_1\beta_j')=\lambda_1\alpha_1\beta_{j-1}'+\lambda_1^2\alpha_1\beta_j',\ 2\le j\le r_1;\\
   &\varphi(\alpha_i\beta_1')=\lambda_1\alpha_{i-1}\beta_1'+\lambda_1^2\alpha_i\beta_1',\ 2\le i\le r_1;\\
   &\varphi(\alpha_i\beta_j')=\alpha_{i-1}\beta_{j-1}'+\lambda_1\alpha_{i-1}\beta_j'+\lambda_1\alpha_i\beta_{j-1}'+\lambda_1^2\alpha_i\beta_j',\ 2\le i,j\le r_1,
   \end{aligned}\tag{6.10}$$

   于是 U 是 φ-不变子空间. 由于 φ 可对角化，故由命题 7.33可知 $\varphi|_U$ 也可对角化，但(6.10)式告诉我们 $\varphi|_U$ 在基 $\{\alpha_1\beta_1',\cdots,\alpha_1\beta_{r_1}';\cdots;\alpha_{r_1}\beta_1',\cdots,\alpha_{r_1}\beta_{r_1}'\}$ 下的表示矩阵是一个上三角矩阵，主对角元全为 λ_1^2，主对角线上方至少有一个非零元素 1（其实是 Kronecker 积 $J_{r_1}(\lambda_1)\otimes J_{r_1}(\lambda_1)$），由例题 6.38可知这个矩阵不可对角化，矛盾！

□

例题 6.32

1. 设 V 为 n 阶矩阵全体构成的线性空间，V 上的线性变换 φ 定义为 $\varphi(\boldsymbol{X})=\boldsymbol{AX}-\boldsymbol{XA}$，其中 $\boldsymbol{A}\in V$. 证明：若 $\boldsymbol{A}$ 可对角化，则 φ 也可对角化.
2. 设 V 为 n 阶矩阵全体构成的线性空间，V 上的线性变换 φ 定义为 $\varphi(X)=AX-XA$，其中 $A\in V$. 证明：φ 可对

角化的充要条件是 A 可对角化.

注 第 2 问是第 1 问的延拓.

证明

1. 证法一: 设 $\boldsymbol{P}$ 为 n 阶可逆矩阵, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$, 则 $\boldsymbol{P}'\boldsymbol{A}'(\boldsymbol{P}')^{-1}=\boldsymbol{\Lambda}$, 即 $\boldsymbol{A}'$ 也可对角化. 设
$$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n),(\boldsymbol{P}')^{-1}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$$
分别为两个矩阵的列分块, 则
$$\boldsymbol{A}\boldsymbol{\alpha}_i=\lambda_i\boldsymbol{\alpha}_i,\boldsymbol{A}'\boldsymbol{\beta}_j=\lambda_j\boldsymbol{\beta}_j,1\le i,j\le n,$$
且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关, $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 线性无关. 由命题 3.14可知, $\{\boldsymbol{\alpha}_i\boldsymbol{\beta}_j',1\le i,j\le n\}$ 是 V 中 n^2 个线性无关的矩阵. 注意到
$$\varphi(\boldsymbol{\alpha}_i\boldsymbol{\beta}_j')=\boldsymbol{A}\boldsymbol{\alpha}_i\boldsymbol{\beta}_j'-\boldsymbol{\alpha}_i\boldsymbol{\beta}_j'\boldsymbol{A}=(\boldsymbol{A}\boldsymbol{\alpha}_i)\boldsymbol{\beta}_j'-\boldsymbol{\alpha}_i(\boldsymbol{A}'\boldsymbol{\beta}_j)'=(\lambda_i-\lambda_j)\boldsymbol{\alpha}_i\boldsymbol{\beta}_j',$$
故 φ 有 n^2 个线性无关的特征向量, 从而可对角化.

证法二: 由于 A 可对角化, 故存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ 为对角矩阵. 由命题 6.56可知,φ 在基础矩阵这组基下的表示矩阵为 $\boldsymbol{A}\otimes\boldsymbol{I}_n-\boldsymbol{I}_n\otimes\boldsymbol{A}'$, 于是
$$(\boldsymbol{P}\otimes(\boldsymbol{P}')^{-1})^{-1}(\boldsymbol{A}\otimes\boldsymbol{I}_n-\boldsymbol{I}_n\otimes\boldsymbol{A}')(\boldsymbol{P}\otimes(\boldsymbol{P}')^{-1})=\boldsymbol{\Lambda}\otimes\boldsymbol{I}_n-\boldsymbol{I}_n\otimes\boldsymbol{\Lambda}$$
为对角矩阵, 即 $\boldsymbol{A}\otimes\boldsymbol{I}_n-\boldsymbol{I}_n\otimes\boldsymbol{A}'$ 可对角化, 从而 φ 可对角化.

2. 充分性就是第 1 问, 下证必要性. 用反证法, 设 A 不可对角化, 则存在可逆矩阵 P, 使得 $P^{-1}AP=J=\mathrm{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_k}(\lambda_k)\}$ 为 Jordan 标准型, 其中 $r_1>1$. 设 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 为列分块, 任取 A' 的特征值 λ_0 及其特征向量 β, 即 $A'\beta=\lambda_0\beta$. 令 $U=L(\alpha_i\beta',1\le i\le r_1)$, 则由第 3 章的解答题 3 可知 $\{\alpha_i\beta',1\le i\le r_1\}$ 是 U 的一组基. 经简单计算可得
$$\varphi(\alpha_1\beta')=(\lambda_1-\lambda_0)\alpha_1\beta',\ \varphi(\alpha_2\beta')=\alpha_1\beta'+(\lambda_1-\lambda_0)\alpha_2\beta',\cdots,\ \varphi(\alpha_{r_1}\beta')=\alpha_{r_1-1}\beta'+(\lambda_1-\lambda_0)\alpha_{r_1}\beta',\tag{6.11}$$
于是 U 是 φ-不变子空间. 由于 φ 可对角化, 故由命题 7.33可知 $\varphi|_U$ 也可对角化, 但(6.11)式告诉我们 $\varphi|_U$ 在基 $\{\alpha_i\beta',1\le i\le r_1\}$ 下的表示矩阵为 $J_{r_1}(\lambda_1-\lambda_0)$, 这个矩阵不可对角化, 矛盾!

□

6.4.3 有 n 个不同特征值

由于属于不同特征值的特征向量线性无关, 故若 $\boldsymbol{A}$ 有 n 个不同的特征值, 则 $\boldsymbol{A}$ 必有 n 个线性无关的特征向量, 从而可对角化. 请注意 $\boldsymbol{A}$ 有 n 个不同的特征值只是可对角化的充分条件, 而并非必要条件.

例题 6.33 设 $\boldsymbol{A}$ 是实二阶矩阵且 $|\boldsymbol{A}|<0$, 求证: $\boldsymbol{A}$ 实相似于对角矩阵.

证明 设
$$\boldsymbol{A}=\begin{pmatrix}a&b\\c&d\end{pmatrix},$$
由 $|\boldsymbol{A}|<0$ 可得 $ad-bc<0$. 又 $\boldsymbol{A}$ 的特征多项式
$$|\lambda\boldsymbol{I}_2-\boldsymbol{A}|=\lambda^2-(a+d)\lambda+(ad-bc),$$
上述关于 λ 的二次方程其判别式大于零, 从而 $\boldsymbol{A}$ 有两个不相等的实特征值, 因此 $\boldsymbol{A}$ 实相似于对角矩阵. □

例题 6.34 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 都是 n 阶矩阵, $\boldsymbol{A},\boldsymbol{B}$ 各有 n 个不同的特征值, 又 $f(\lambda)$ 是 $\boldsymbol{A}$ 的特征多项式, 且 $f(\boldsymbol{B})$ 是可逆矩阵. 求证: 矩阵
$$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{C}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}$$
相似于对角矩阵.

证明 任取 $\boldsymbol{B}$ 的一个特征值 μ_0, 则 $f(\mu_0)$ 是 $f(\boldsymbol{B})$ 的特征值. 由于 $f(\boldsymbol{B})$ 可逆, 故 $f(\boldsymbol{B})$ 的特征值非零, 从而 $f(\mu_0)\neq0$,

即 μ_0 不是 $\boldsymbol{A}$ 的特征值, 于是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的特征值互不相同. 注意到

$$|\lambda \boldsymbol{I}_{2n} - \boldsymbol{M}| = \begin{vmatrix} \lambda \boldsymbol{I}_n - \boldsymbol{A} & -\boldsymbol{C} \\ \boldsymbol{O} & \lambda \boldsymbol{I}_n - \boldsymbol{B} \end{vmatrix} = |\lambda \boldsymbol{I}_n - \boldsymbol{A}||\lambda \boldsymbol{I}_n - \boldsymbol{B}|,$$

故矩阵 $\boldsymbol{M}$ 有 $2n$ 个不同的特征值, 从而相似于对角矩阵. □

例题 6.35 设 n 阶矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 有相同的特征值, 且这 n 个特征值互不相等. 求证: 存在 n 阶矩阵 $\boldsymbol{P}, \boldsymbol{Q}$, 使得 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{Q}, \boldsymbol{B} = \boldsymbol{Q}\boldsymbol{P}$.

证明 由假设以及定理 6.11和例题 6.4可知, 矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 相似于同一个对角矩阵, 因此 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似. 不妨设 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 令 $\boldsymbol{Q} = \boldsymbol{P}^{-1}\boldsymbol{A}$, 则 $\boldsymbol{P}\boldsymbol{Q} = \boldsymbol{A}, \boldsymbol{Q}\boldsymbol{P} = \boldsymbol{B}$. □

命题 6.29

设 A, B 是 n 阶矩阵,A 有 n 个不同的特征值, 并且 $AB = BA$, 求证:B 相似于对角矩阵, 并且 A 与 B 可同时对角化.

证明 证法一 (几何方法): 因为 A 有 n 个不同的特征值, 故 A 可对角化. 令 V 是 n 维复列向量空间, 将 A, B 看成是 V 上的线性变换. 又设 A 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 对应的特征向量为 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 则 λ_i 的特征子空间 $V_i = L(\alpha_i)(1 \le i \le n)$, 且

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

注意到 $AB = BA$, 故由命题 6.16可知,A 的特征子空间 V_i 是 B 的不变子空间. 将 B 限制在 V_i 上, 这是一维线性空间 V_i 上的线性变换, 从而只能是纯量变换, 即存在 μ_i, 使得 $B\alpha_i = \mu_i\alpha_i(1 \le i \le n)$, 于是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 也是 B 的特征向量. 因此,B 有 n 个线性无关的特征向量, 从而 B 可对角化. 事实上, 我们得到了一个更强的结果:A 和 B 可同时对角化, 即存在可逆矩阵 $P = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, 使得 $P^{-1}AP = \operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ 和 $P^{-1}BP = \operatorname{diag}\{\mu_1, \mu_2, \cdots, \mu_n\}$ 都是对角矩阵.

证法二 (代数方法): 因为 A 有 n 个不同的特征值, 故 A 可对角化, 即存在可逆矩阵 P, 使得 $P^{-1}AP = \operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$. 注意到问题的条件和结论在同时相似变换:$A \mapsto P^{-1}AP, B \mapsto P^{-1}BP$ 下保持不变, 故不妨从一开始就假设 $A = \operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ 为对角矩阵. 设 $B = (b_{ij})$, 则

$$AB = \begin{pmatrix} \lambda_1 b_{11} & \lambda_1 b_{12} & \cdots & \lambda_1 b_{1n} \\ \lambda_2 b_{21} & \lambda_2 b_{22} & \cdots & \lambda_2 b_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_n b_{n1} & \lambda_n b_{n2} & \cdots & \lambda_n b_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 b_{11} & \lambda_2 b_{12} & \cdots & \lambda_n b_{1n} \\ \lambda_1 b_{21} & \lambda_2 b_{22} & \cdots & \lambda_n b_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 b_{n1} & \lambda_2 b_{n2} & \cdots & \lambda_n b_{nn} \end{pmatrix} = BA,$$

比较元素可得 $\lambda_i b_{ij} = \lambda_j b_{ij}$. 注意到 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 互不相同, 故 $b_{ij} = 0(i \ne j)$, 即 B 为对角矩阵. □

命题 6.30

设 A, B 是 n 阶矩阵,A 有 n 个不同的特征值, 并且 $AB = BA$, 求证: 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $B = f(A)$.

证明 证法一: 由命题 6.29可知 A 和 B 可以同时对角化, 即存在可逆矩阵 P, 使得

$$P^{-1}AP = \operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}, P^{-1}BP = \operatorname{diag}\{\mu_1, \mu_2, \cdots, \mu_n\},$$

其中 λ_i, μ_i 分别是 A, B 的特征值. 因为 λ_i 互不相同, 故由Lagrange 插值公式可知, 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $f(\lambda_i) = \mu_i(1 \le i \le n)$. 于是

$$P^{-1}BP = \operatorname{diag}\{f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)\} = f(P^{-1}AP) = P^{-1}f(A)P,$$

从而 $B = f(A)$.

证法二: 由定理 7.12和命题 7.21可知,$\mathbb{C}^n$ 是关于 $\mathbb{A}$ 的循环空间, 再由定理 7.25即得结论. □

命题 6.31

若 A 可对角化, 则对任意的多项式 $f(x)$,$f(A)$ 也可对角化.

笔记 这一结论提醒我们: 在处理可对角化问题时, 如能将矩阵写成可对角化矩阵的多项式, 则往往讨论起来更加方便.

证明 事实上, 设 P 为可逆矩阵, 使得 $P^{-1}AP = \mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$ 为对角矩阵, 则 $P^{-1}f(A)P = f(P^{-1}AP) = \mathrm{diag}\{f(\lambda_1), f(\lambda_2),\cdots, f(\lambda_n)\}$ 也为对角矩阵. □

命题 6.32

设 A 是 n 阶复矩阵且有 n 个不同的特征值, 求证:n 阶复矩阵 B 可对角化的充要条件是存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 B 相似于 $f(A)$.

证明 先证充分性. 由于 A 有 n 个不同的特征值, 故 A 可对角化, 从而由命题 6.31$f(A)$ 也可对角化, 又 B 相似于 $f(A)$, 于是 B 也可对角化.

再证必要性. 设 P,Q 为可逆矩阵, 使得

$$P^{-1}AP = \mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}, Q^{-1}BQ = \mathrm{diag}\{\mu_1,\mu_2,\cdots,\mu_n\},$$

其中 λ_i,μ_i 分别是 A,B 的特征值. 因为 λ_i 互不相同, 故由Lagrange 插值公式可知, 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $f(\lambda_i)=\mu_i(1\le i\le n)$. 于是

$$Q^{-1}BQ = \mathrm{diag}\{f(\lambda_1), f(\lambda_2),\cdots, f(\lambda_n)\} = f(P^{-1}AP) = P^{-1}f(A)P,$$

即有 $B = (PQ^{-1})^{-1}f(A)(PQ^{-1})$, 从而 B 相似于 $f(A)$. □

推论 6.7

n 阶复方阵 B 可对角化的充要条件是 B 相似于某个循环矩阵.

证明 设 $J = \begin{pmatrix} O & I_{n-1} \\ 1 & O \end{pmatrix}$, 经简单计算可得 $|\lambda I_n - J| = \lambda^n - 1$, 于是 J 有 n 个不同的特征值. 对任一循环矩阵 C, 由循环矩阵的性质 2可知, 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $C = f(J)$, 故由命题 6.32即得本推论. □

命题 6.33

设 a,b,c 为复数且 $bc\ne 0$, 证明下列 n 阶矩阵 A 可对角化:

$$A = \begin{pmatrix} a & b & & & & \\ c & a & b & & & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & & & c & a & b \\ & & & & c & a \end{pmatrix}$$

笔记 当 $(\lambda-a)^2 = 4bc$ 时, 利用摄动法, 设 $t>0$, 则当 $(\lambda-a)^2-4bc=0$ 时,$(\lambda+t-a)^2-4bc>0$, 由下述证明可知,$\lambda+t$ 有 n 个不同取值, 从而令 $t\to 0$, 则此时 $(\lambda-a)^2-4bc=0$, 并且 λ 仍有 n 个不同取值.

证明 我们先来计算 A 的特征多项式 $|\lambda I_n - A|$. 设 x_1,x_2 是二次方程 $x^2-(\lambda-a)x+bc=0$ 的两个根, 则当 $(\lambda-a)^2\ne 4bc$ 时, 由推论 1.2可得

$$|\lambda I_n - A| = \frac{x_1^{n+1}-x_2^{n+1}}{x_1-x_2}.$$

注意到 x_1,x_2 都是关于 λ 的连续函数, 要求 A 的特征值 λ, 即是求 λ 的值, 使得 $|\lambda I_n - A| = 0$, 而这也等价于 $x_1^{n+1}=x_2^{n+1}$. 令 $\omega = \cos\dfrac{2\pi}{n+1}+\mathrm{i}\sin\dfrac{2\pi}{n+1}$ 为 1 的 $n+1$ 次方根, 则由 $x_1^{n+1}=x_2^{n+1}$ 可得 $x_1 = x_2\omega^k(1\le k\le n)$. 由Vieta

定理可得 $x_1x_2 = bc$, 在选定 bc 的某一平方根 $\sqrt{bc}$ 之后, 可解出

$$x_1 = \sqrt{bc}\left(\cos\frac{k\pi}{n+1} + \mathrm{i}\sin\frac{k\pi}{n+1}\right), x_2 = \sqrt{bc}\left(\cos\frac{k\pi}{n+1} - \mathrm{i}\sin\frac{k\pi}{n+1}\right), 1 \le k \le n.$$

再次由Vieta 定理可得 $\lambda - a = x_1 + x_2 = 2\sqrt{bc}\cos\frac{k\pi}{n+1}$, 即

$$\lambda = a + 2\sqrt{bc}\cos\frac{k\pi}{n+1}, 1 \le k \le n.$$

容易验证上述 n 个数的确是 A 的 n 个不同的特征值, 从而 A 可对角化. □

6.4.4 全空间等于特征子空间的直和

矩阵或线性变换可对角化当且仅当全空间等于特征子空间的直和这一判定准则, 不仅给了我们很多几何想象的空间, 而且与矩阵或线性变换适合的多项式密切相关.

命题 6.34

设 n 阶矩阵 $\boldsymbol{A}$ 适合首一多项式 $g(x)$, 并且 $g(x)$ 在复数域中无重根, 证明:$\boldsymbol{A}$ 可对角化.

证明 证法一: 设 $g(x) = (x-a_1)(x-a_2)\cdots(x-a_m)$ 是复数域上的因式分解, 其中 $a_1, a_2, \cdots, a_m$ 是互异的复数. 我们先来证明:

$$\mathbb{C}^n = \mathrm{Ker}(\boldsymbol{A} - a_1\boldsymbol{I}_n) \oplus \mathrm{Ker}(\boldsymbol{A} - a_2\boldsymbol{I}_n) \oplus \cdots \oplus \mathrm{Ker}(\boldsymbol{A} - a_m\boldsymbol{I}_n). \tag{6.12}$$

设 $g_i(x) = \prod\limits_{j\ne i}(x-a_j)$, 则 $(g_1(x), g_2(x), \cdots, g_m(x)) = 1$, 故存在 $u_i(x)(1 \le i \le m)$, 使得

$$g_1(x)u_1(x) + g_2(x)u_2(x) + \cdots + g_m(x)u_m(x) = 1.$$

代入 $x = \boldsymbol{A}$, 可得恒等式

$$g_1(\boldsymbol{A})u_1(\boldsymbol{A}) + g_2(\boldsymbol{A})u_2(\boldsymbol{A}) + \cdots + g_m(\boldsymbol{A})u_m(\boldsymbol{A}) = \boldsymbol{I}_n. \tag{6.13}$$

对任一 $\boldsymbol{\alpha} \in \mathbb{C}^n$, 由上式可知 $\boldsymbol{\alpha} = \sum\limits_{i=1}^{m} g_i(\boldsymbol{A})u_i(\boldsymbol{A})\boldsymbol{\alpha}$. 注意到 $(\boldsymbol{A}-a_i\boldsymbol{I}_n)g_i(\boldsymbol{A})u_i(\boldsymbol{A})\boldsymbol{\alpha} = g(\boldsymbol{A})u_i(\boldsymbol{A})\boldsymbol{\alpha} = \boldsymbol{0}$, 故 $g_i(\boldsymbol{A})u_i(\boldsymbol{A})\boldsymbol{\alpha} \in \mathrm{Ker}(\boldsymbol{A} - a_i\boldsymbol{I}_n)$, 于是

$$\mathbb{C}^n = \mathrm{Ker}(\boldsymbol{A} - a_1\boldsymbol{I}_n) + \mathrm{Ker}(\boldsymbol{A} - a_2\boldsymbol{I}_n) + \cdots + \mathrm{Ker}(\boldsymbol{A} - a_m\boldsymbol{I}_n). \tag{6.14}$$

任取 $\boldsymbol{\alpha} \in \mathrm{Ker}(\boldsymbol{A}-a_1\boldsymbol{I}_n)\cap(\mathrm{Ker}(\boldsymbol{A}-a_2\boldsymbol{I}_n)+\cdots+\mathrm{Ker}(\boldsymbol{A}-a_m\boldsymbol{I}_n))$, 则 $\boldsymbol{\alpha} = \boldsymbol{\alpha}_2 + \cdots + \boldsymbol{\alpha}_m$, 其中 $\boldsymbol{\alpha}_i \in \mathrm{Ker}(\boldsymbol{A} - a_i\boldsymbol{I}_n)(i \ge 2)$. 由(6.13)式可知

$$\boldsymbol{\alpha} = u_1(\boldsymbol{A})g_1(\boldsymbol{A})(\boldsymbol{\alpha}_2 + \cdots + \boldsymbol{\alpha}_m) + u_2(\boldsymbol{A})g_2(\boldsymbol{A})\boldsymbol{\alpha} + \cdots + u_m(\boldsymbol{A})g_m(\boldsymbol{A})\boldsymbol{\alpha} = \boldsymbol{0}.$$

注意到下指标可任意选, 故(6.14)式是直和.

由于 $\boldsymbol{A}$ 适合 $g(x)$, 故 $\boldsymbol{A}$ 的特征值也适合 $g(x)$, 从而只可能是 $a_1, a_2, \cdots, a_m$ 中的一部分. 不妨设 a_1 不是 $\boldsymbol{A}$ 的特征值, 则对 $\forall\alpha \ne 0$, 都有 $(\boldsymbol{A} - a_1\boldsymbol{I_n})\alpha \ne 0$, 故 $\mathrm{Ker}(\boldsymbol{A} - a_1\boldsymbol{I}_n) = 0$. 于是在(6.12)式中剔除等于零的直和分量, 这就证明了全空间等于特征子空间的直和, 从而 $\boldsymbol{A}$ 可对角化.

证法二:设 $m(x)$ 为 A 的极小多项式, 则 $m(x) \mid g(x)$. 由于 $g(x)$ 无重根, 故 $m(x)$ 也无重根, 从而 A 可对角化. □

例题 6.36 求证:

(1) 若 n 阶矩阵 $\boldsymbol{A}$ 适合 $\boldsymbol{A}^2 = \boldsymbol{I}_n$, 则 $\boldsymbol{A}$ 必可对角化;

(2) 若 n 阶矩阵 $\boldsymbol{A}$ 适合 $\boldsymbol{A}^2 = \boldsymbol{A}$, 则 $\boldsymbol{A}$ 必可对角化.

(3) $M_n(\mathbb{F})$ 上的线性变换 $\boldsymbol{\eta}$ 满足 $\boldsymbol{\eta}^2 = \boldsymbol{I}_V$, 则 $\boldsymbol{\eta}$ 可对角化.

证明

(1) 对合矩阵 $\boldsymbol{A}$ 适合多项式 $x^2 - 1$, 它在复数域中无重根, 故由命题 6.34即得结论.

(2) 幂等矩阵 $\boldsymbol{A}$ 适合多项式 $x^2 - x$, 它在复数域中无重根, 故由命题 6.34即得结论.

(3) 线性变换 $\boldsymbol{\eta}$ 适合多项式 $x^2 - 1$, 它在复数域中无重根, 故由命题 6.34即得结论.

□

6.4.5 有完全的特征向量系

矩阵或线性变换有完全的特征向量系, 即任一特征值的代数重数等于其几何重数, 也就是特征值与线性无关的特征向量完全一一对应.

命题 6.35

若矩阵 $\boldsymbol{A},\boldsymbol{B}$ 有完全的特征向量系, 求证: $\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}$ 也有完全的特征向量系.

证明 因为 $\boldsymbol{A},\boldsymbol{B}$ 有完全的特征向量系, 故相似于对角矩阵. 设 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 和 $\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}$ 是对角矩阵, 则

$$\begin{pmatrix}\boldsymbol{P} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{Q}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{P} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{Q}\end{pmatrix}=\begin{pmatrix}\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}\end{pmatrix}$$

是对角矩阵. 因此 $\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}$ 有完全的特征向量系. □

例题 6.37 设 n 阶矩阵 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{B}\\ \boldsymbol{O} & -\boldsymbol{I}_{n-r}\end{pmatrix}$, 求证: $\boldsymbol{A}$ 可对角化.

证明 证法一: 显然 $\boldsymbol{A}$ 有特征值 1 (r 重) 与 -1 ($n-r$ 重). 注意到矩阵 $\boldsymbol{I}_n-\boldsymbol{A}=\begin{pmatrix}\boldsymbol{O} & -\boldsymbol{B}\\ \boldsymbol{O} & 2\boldsymbol{I}_{n-r}\end{pmatrix}$ 的秩等于 $n-r$, 因此特征值 1 的几何重数等于 $n-\mathrm{r}(\boldsymbol{I}_n-\boldsymbol{A})=r$, 与其代数重数相等. 同理可证特征值 -1 的几何重数为 $n-r$, 与其代数重数相同. 因此 $\boldsymbol{A}$ 可对角化, 且相似于对角矩阵 $\mathrm{diag}\{\boldsymbol{I}_r,-\boldsymbol{I}_{n-r}\}$.

证法二: 容易算出 $\boldsymbol{A}^2=\boldsymbol{I}_n$, 由例题 6.36(1)可知 $\boldsymbol{A}$ 可对角化.

证法三: 做第三种初等相似变换, 由 $\begin{pmatrix}\boldsymbol{I}_r & \frac{1}{2}\boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{I}_{n-r}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{B}\\ \boldsymbol{O} & -\boldsymbol{I}_{n-r}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_r & -\frac{1}{2}\boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{I}_{n-r}\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & -\boldsymbol{I}_{n-r}\end{pmatrix}$ 即得. □

命题 6.36

设 m 阶矩阵 $\boldsymbol{A}$ 与 n 阶矩阵 $\boldsymbol{B}$ 没有公共的特征值, 且 $\boldsymbol{A},\boldsymbol{B}$ 均可对角化, 又 $\boldsymbol{C}$ 为 $m\times n$ 矩阵, 求证: $\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}$ 也可对角化.

证明 证法一: 任取 $\boldsymbol{A}$ 的特征值 λ_0, 记其代数重数为 $m_{\boldsymbol{A}}(\lambda_0)$, 几何重数为 $t_{\boldsymbol{A}}(\lambda_0)$. 首先注意到 $\boldsymbol{A},\boldsymbol{B}$ 没有公共的特征值, 故 λ_0 不是 $\boldsymbol{B}$ 的特征值, 又 $|\lambda\boldsymbol{I}-\boldsymbol{M}|=|\lambda\boldsymbol{I}-\boldsymbol{A}||\lambda\boldsymbol{I}-\boldsymbol{B}|$, 从而 $m_{\boldsymbol{M}}(\lambda_0)=m_{\boldsymbol{A}}(\lambda_0)$. 由于 $\lambda_0\boldsymbol{I}-\boldsymbol{B}$ 是非异阵, 故有如下分块矩阵的初等变换:

$$\lambda_0\boldsymbol{I}-\boldsymbol{M}=\begin{pmatrix}\lambda_0\boldsymbol{I}-\boldsymbol{A} & -\boldsymbol{C}\\ \boldsymbol{O} & \lambda_0\boldsymbol{I}-\boldsymbol{B}\end{pmatrix}\to\begin{pmatrix}\lambda_0\boldsymbol{I}-\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \lambda_0\boldsymbol{I}-\boldsymbol{B}\end{pmatrix}.$$

因为矩阵的秩在分块初等变换下不变, 故由矩阵秩的等式可得

$$\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})=\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})+\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{B})=\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})+n,$$

于是 $t_{\boldsymbol{M}}(\lambda_0)=(m+n)-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})=m-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})=t_{\boldsymbol{A}}(\lambda_0)$. 因为 $\boldsymbol{A}$ 可对角化, 所以 $\boldsymbol{A}$ 有完全的特征向量系, 从而 $m_{\boldsymbol{A}}(\lambda_0)=t_{\boldsymbol{A}}(\lambda_0)$, 于是 $m_{\boldsymbol{M}}(\lambda_0)=t_{\boldsymbol{M}}(\lambda_0)$. 同理可证, 对 $\boldsymbol{B}$ 的任一特征值 μ_0, 成立 $m_{\boldsymbol{M}}(\mu_0)=t_{\boldsymbol{M}}(\mu_0)$. 因此 $\boldsymbol{M}$ 有完全的特征向量系, 从而可对角化.

证法二: 由命题 6.51可知, 矩阵方程 $\boldsymbol{AX}-\boldsymbol{XB}=\boldsymbol{C}$ 存在唯一解 $\boldsymbol{X}=\boldsymbol{X}_0$. 考虑如下相似变换:

$$\begin{pmatrix}\boldsymbol{I}_m & \boldsymbol{X}_0\\ \boldsymbol{O} & \boldsymbol{I}_n\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_m & -\boldsymbol{X}_0\\ \boldsymbol{O} & \boldsymbol{I}_n\end{pmatrix}=\begin{pmatrix}\boldsymbol{A} & -\boldsymbol{A}\boldsymbol{X}_0+\boldsymbol{X}_0\boldsymbol{B}+\boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix},$$

由命题 6.35可知上式最右边的分块对角矩阵可对角化, 于是原矩阵也可对角化. $\square$

命题 6.37

设 $\boldsymbol{A}$ 为 m 阶矩阵, $\boldsymbol{B}$ 为 n 阶矩阵, $\boldsymbol{C}$ 为 $m\times n$ 矩阵, $\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}$, 求证: 若 $\boldsymbol{M}$ 可对角化, 则 $\boldsymbol{A},\boldsymbol{B}$ 均可对角化.

注 这个命题的几何版本 (见命题 7.33) 是: 设 φ 是复线性空间 V 上的线性变换, U 是 φ-不变子空间, 若 φ 可对角化, 则 φ 在不变子空间 U 上的限制变换 $\varphi|_U$ 以及 φ 在商空间 V/U 上的诱导变换 $\overline{\varphi}$ 均可对角化.

证明 任取 $\boldsymbol{M}$ 的特征值 λ_0, 记其代数重数为 $m_{\boldsymbol{A}}(\lambda_0)$, 几何重数为 $t_{\boldsymbol{A}}(\lambda_0)$. 由 $|\lambda\boldsymbol{I}-\boldsymbol{M}|=|\lambda\boldsymbol{I}-\boldsymbol{A}||\lambda\boldsymbol{I}-\boldsymbol{B}|$ 可得 $m_{\boldsymbol{M}}(\lambda_0)=m_{\boldsymbol{A}}(\lambda_0)+m_{\boldsymbol{B}}(\lambda_0)$. 考虑如下分块矩阵:

$$\lambda_0\boldsymbol{I}-\boldsymbol{M}=\begin{pmatrix}\lambda_0\boldsymbol{I}-\boldsymbol{A} & -\boldsymbol{C}\\ \boldsymbol{O} & \lambda_0\boldsymbol{I}-\boldsymbol{B}\end{pmatrix},$$

由矩阵秩的基本公式 (4)可得

$$\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})\geq \mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})+\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{B}),$$

于是 $t_{\boldsymbol{M}}(\lambda_0)=(m+n)-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})\leq(m-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A}))+(n-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{B}))=t_{\boldsymbol{A}}(\lambda_0)+t_{\boldsymbol{B}}(\lambda_0)$. 由于几何重数总是小于等于代数重数, 故有

$$t_{\boldsymbol{M}}(\lambda_0)\leq t_{\boldsymbol{A}}(\lambda_0)+t_{\boldsymbol{B}}(\lambda_0)\leq m_{\boldsymbol{A}}(\lambda_0)+m_{\boldsymbol{B}}(\lambda_0)=m_{\boldsymbol{M}}(\lambda_0).$$

因为 $\boldsymbol{M}$ 可对角化, 所以 $\boldsymbol{M}$ 有完全的特征向量系, 从而 $t_{\boldsymbol{M}}(\lambda_0)=m_{\boldsymbol{M}}(\lambda_0)$, 再由上述不等式可得 $t_{\boldsymbol{A}}(\lambda_0)=m_{\boldsymbol{A}}(\lambda_0)$, $t_{\boldsymbol{B}}(\lambda_0)=m_{\boldsymbol{B}}(\lambda_0)$. 由 λ_0 的任意性即知, $\boldsymbol{A},\boldsymbol{B}$ 均有完全的特征向量系, 从而均可对角化. $\square$

命题 6.38

设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵, $\boldsymbol{B}$ 为 $n\times m$ 矩阵, 又 $|\boldsymbol{BA}|\neq 0$, 求证: $\boldsymbol{AB}$ 可对角化的充要条件是 $\boldsymbol{BA}$ 可对角化.

证明 证法一: 记其代数重数为 $m_{\boldsymbol{A}}(\lambda_0)$, 几何重数为 $t_{\boldsymbol{A}}(\lambda_0)$(其他记号同理). 由特征值的降价公式可得 $|\lambda\boldsymbol{I}_m-\boldsymbol{AB}|=\lambda^{m-n}|\lambda\boldsymbol{I}_n-\boldsymbol{BA}|$, 因此 $\boldsymbol{AB}$ 的特征值为 $\boldsymbol{BA}$ 的特征值以及 0. 由于 $\boldsymbol{BA}$ 非异, 故其特征值全部非零, 从而 0 作为 $\boldsymbol{AB}$ 的特征值, 其代数重数为 $m-n$. 另一方面, 我们有

$$n=\mathrm{r}(\boldsymbol{BA})\leq\min\{\mathrm{r}(\boldsymbol{A}),\mathrm{r}(\boldsymbol{B})\}\leq\max\{\mathrm{r}(\boldsymbol{A}),\mathrm{r}(\boldsymbol{B})\}\leq\min\{m,n\}=n,$$

从而 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})=n$. 再由Sylvester 不等式可得

$$n=\mathrm{r}(\boldsymbol{A})+\mathrm{r}(\boldsymbol{B})-n\leq\mathrm{r}(\boldsymbol{AB})\leq\min\{\mathrm{r}(\boldsymbol{A}),\mathrm{r}(\boldsymbol{B})\}=n,$$

从而 $\mathrm{r}(\boldsymbol{AB})=n$. 因此 0 作为 $\boldsymbol{AB}$ 的特征值, 其几何重数为 $m-\mathrm{r}(\boldsymbol{AB})=m-n$, 即特征值 0 的代数重数等于几何重数. 任取 $\boldsymbol{BA}$ 的特征值 λ_0, 它也是 $\boldsymbol{AB}$ 的非零特征值, 显然 $m_{\boldsymbol{AB}}(\lambda_0)=m_{\boldsymbol{BA}}(\lambda_0)$. 考虑分块矩阵 $\begin{pmatrix}\boldsymbol{I}_m & \boldsymbol{A}\\ \boldsymbol{B} & \lambda_0\boldsymbol{I}_n\end{pmatrix}$, 由秩的降阶公式可得

$$m+\mathrm{r}(\lambda_0\boldsymbol{I}_n-\boldsymbol{BA})=n+\mathrm{r}(\boldsymbol{I}_m-\frac{1}{\lambda_0}\boldsymbol{AB})=n+\mathrm{r}(\lambda_0\boldsymbol{I}_m-\boldsymbol{AB}),$$

于是 $t_{\boldsymbol{AB}}(\lambda_0)=m-\mathrm{r}(\lambda_0\boldsymbol{I}_m-\boldsymbol{AB})=n-\mathrm{r}(\lambda_0\boldsymbol{I}_n-\boldsymbol{BA})=t_{\boldsymbol{BA}}(\lambda_0)$. 由 λ_0 的任意性即知, $\boldsymbol{AB}$ 有完全的特征向量系当且仅当 $\boldsymbol{BA}$ 有完全的特征向量系, 从而 $\boldsymbol{AB}$ 可对角化当且仅当 $\boldsymbol{BA}$ 可对角化.

证法二:设 AB 的极小多项式为 $g(\lambda)$,BA 的极小多项式为 $h(\lambda)$. 因为 BA 是可逆矩阵, 故 0 不是 BA 的特征值, 从而 0 也不是 $h(\lambda)$ 的根（极小多项式和特征多项式有相同的根 (不计重数)）. 注意到

$$(AB)^m=A(BA)^{m-1}B,\quad (BA)^m=B(AB)^{m-1}A,\quad m\geq 1$$

故不难验证 $g(BA)BA=Bg(AB)A=O$,$h(AB)AB=Ah(BA)B=O$, 从而由极小多项式的基本性质可知,$h(\lambda)\mid g(\lambda)\lambda$,$g(\lambda)\mid h(\lambda)\lambda$. 若 AB 可对角化, 则 $g(\lambda)$ 无重根, 从而 $g(\lambda)\lambda$ 无非零重根, 于是 $h(\lambda)$ 无重根, 故 BA 也可对角化. 反之, 若 BA 可对角化, 则 $h(\lambda)$ 无重根, 从而 $h(\lambda)\lambda$ 也无重根, 于是 $g(\lambda)$ 无重根, 故 AB 也可对角化. $\square$

6.4.6 利用反证法证明不可对角化

例题 6.38 求证:

(1) 若 n 阶矩阵 $\boldsymbol{A}$ 的特征值都是 λ_0, 但 $\boldsymbol{A}$ 不是纯量矩阵, 则 $\boldsymbol{A}$ 不可对角化. 特别地, 非零的幂零矩阵不可对角化.

(2) 若 n 阶实矩阵 $\boldsymbol{A}$ 适合 $\boldsymbol{A}^2+\boldsymbol{A}+\boldsymbol{I}_n=\boldsymbol{O}$, 则 $\boldsymbol{A}$ 在实数域上不可对角化.

证明

(1) 用反证法, 设 $\boldsymbol{A}$ 可对角化, 则存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ 为对角矩阵. 由假设 $\boldsymbol{\Lambda}$ 的主对角元素全为 λ_0, 故 $\boldsymbol{\Lambda}=\lambda_0\boldsymbol{I}_n$, 于是 $\boldsymbol{A}=\boldsymbol{P}(\lambda_0\boldsymbol{I}_n)\boldsymbol{P}^{-1}=\lambda_0\boldsymbol{I}_n$, 这与假设矛盾. 特别地, 幂零矩阵的特征值都是零 (特征值适合 $x^k=0$), 因此也不可对角化.

(2) 用反证法, 设 $\boldsymbol{A}$ 在实数域上可对角化, 则 $\boldsymbol{A}$ 的特征值都是实数. 因为 $\boldsymbol{A}$ 适合多项式 x^2+x+1, 故由命题 6.7可知, $\boldsymbol{A}$ 的特征值也适合 x^2+x+1, 从而不可能是实数, 矛盾.

□

引理 6.3 (秩 1 矩阵的列向量分解)

设 $n(n>1)$ 阶矩阵 $\boldsymbol{A}$ 的秩为 1, 则存在非零列向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$, 使得 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\beta}'$.

证明 由于 $\mathrm{r}(\boldsymbol{A})=1$, 因此 $\boldsymbol{A}$ 的列向量成比例, 从而存在非零列向量 $\boldsymbol{\alpha}$, 使得 $\boldsymbol{A}$ 的列分块为

$$\boldsymbol{A}=(\boldsymbol{k_1}\boldsymbol{\alpha},\boldsymbol{k_2}\boldsymbol{\alpha},\cdots,\boldsymbol{k_n}\boldsymbol{\alpha}).$$

其中 $\boldsymbol{k_1},\boldsymbol{k_2},\cdots,\boldsymbol{k_n}$ 为不全为零的实数. 否则,$\boldsymbol{A}=\boldsymbol{O}$ 矛盾! 于是令 $\boldsymbol{\beta}=(\boldsymbol{k_1},\boldsymbol{k_2},\cdots,\boldsymbol{k_n})'\neq\boldsymbol{0}$, 则

$$\boldsymbol{A}=(\boldsymbol{k}_1\boldsymbol{\alpha},\boldsymbol{k}_2\boldsymbol{\alpha},\cdots,\boldsymbol{k_n}\boldsymbol{\alpha})=\boldsymbol{\alpha}\boldsymbol{\beta}'.$$

□

命题 6.39

设 $n(n>1)$ 阶矩阵 $\boldsymbol{A}$ 的秩为 1, 求证: $\boldsymbol{A}$ 可对角化的充要条件是 $\mathrm{tr}(\boldsymbol{A})\neq 0$.

笔记 这个命题告诉我们命题 6.38的条件 $|BA|\neq 0$ 是必要的.

证明 由 $\mathrm{r}(\boldsymbol{A})=1$ 及秩 1 矩阵的列向量分解可知, 存在非零列向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$, 使得 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\beta}'$, 于是由迹的交换性可得 $\mathrm{tr}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{\alpha}\boldsymbol{\beta}')=\mathrm{tr}(\boldsymbol{\beta}'\boldsymbol{\alpha})=\boldsymbol{\beta}'\boldsymbol{\alpha}$.

证法一: 由例题 6.10及其可对角化的讨论可知, 令例题 6.10条件中的 $\boldsymbol{A}=\boldsymbol{I_n}$ 即可得到 A 可对角化的充要条件是 $\mathrm{tr}(\boldsymbol{A})=\boldsymbol{\beta}'\boldsymbol{\alpha}\neq 0$ 或 $A=O$, 而 $\boldsymbol{A}=\boldsymbol{O}$ 与 $\mathrm{r}(\boldsymbol{A})=1$ 矛盾, 故 A 可对角化的充要条件是 $\mathrm{tr}(\boldsymbol{A})=\neq 0$.

证法二: 注意到 $\boldsymbol{A}^2=(\boldsymbol{\alpha}\boldsymbol{\beta}')(\boldsymbol{\alpha}\boldsymbol{\beta}')=\boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{\alpha})\boldsymbol{\beta}'=(\boldsymbol{\beta}'\boldsymbol{\alpha})\boldsymbol{\alpha}\boldsymbol{\beta}'=\mathrm{tr}(\boldsymbol{A})\boldsymbol{A}$, 故 $\boldsymbol{A}$ 适合多项式 $x^2-\mathrm{tr}(\boldsymbol{A})x$. 若 $\mathrm{tr}(\boldsymbol{A})\neq 0$, 则由命题 6.34可知 $\boldsymbol{A}$ 可对角化; 若 $\mathrm{tr}(\boldsymbol{A})=0$, 则 $\boldsymbol{A}$ 是幂零矩阵, 又 $\boldsymbol{A}\neq\boldsymbol{O}$, 故由例题 6.38(1)可知 $\boldsymbol{A}$ 不可对角化. □

6.5 极小多项式与 Cayley-Hamilton 定理

命题 6.40

数域 $\mathbb{K}$ 上的 n 阶矩阵 A 一定适合数域 $\mathbb{K}$ 上的一个非零多项式.

证明 我们已经知道, 数域 $\mathbb{K}$ 上的 n 阶矩阵全体组成了 $\mathbb{K}$ 上的线性空间, 其维数等于 n^2. 因此对任一 n 阶矩阵 A, 下列 n^2+1 个矩阵必线性相关: $A^{n^2},A^{n^2-1},\cdots,A,I_n$.

也就是说, 存在 $\mathbb{K}$ 中不全为零的数 $c_i(i=0,1,2,\cdots,c_{n^2})$, 使

$$c_{n^2}A^{n^2}+c_{n^2-1}A^{n^2-1}+\cdots+c_1A+c_0I_n=O.$$

这表明矩阵 A 适合数域 $\mathbb{K}$ 上的一个非零多项式.

□

定义 6.16 (矩阵的极小多项式)

若 n 阶矩阵 A (或 n 维线性空间 V 上的线性变换 φ) 适合一个非零首一多项式 $m(x)$, 且 $m(x)$ 是 A (或 φ) 所适合的非零多项式中次数最小者, 则称 $m(x)$ 是 A (或 φ) 的一个**极小多项式**或**最小多项式**.

注 由命题 6.40可知矩阵 A 的极小多项式 $m(x)$ 一定存在, 故极小多项式是良定义的.

定理 6.13 (Cayley-Hamilton 定理)

1. **代数形式:** 设 A 是数域 $\mathbb{K}$ 上的 n 阶矩阵, $f(x)$ 是 A 的特征多项式, 则 $f(A) = O$.
2. **几何形式:** 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 $f(\varphi) = O$.

证明

1. **代数形式:** 因为复数域是最大数域, 所以可将 A 看作一个复矩阵. 由复方阵必相似于上三角阵知 A 复相似于一个上三角阵, 也就是说存在的可逆矩阵 P, 使 $P^{-1}AP = B$ 是一个上三角阵, 其中 P 与 B 都是复矩阵, 由相似矩阵有相同特征多项式可知 A 与 B 有相同的特征多项式 $f(x)$. 记
$$f(x) = x^n + a_1x^{n-1} + \cdots + a_n,$$
则 $f(B) = O$. 而
$$\begin{aligned} f(A) &= A^n + a_1A^{n-1} + \cdots + a_nI_n \\ &= (PBP^{-1})^n + a_1(PBP^{-1})^{n-1} + \cdots + a_nI_n \\ &= PB^nP^{-1} + a_1PB^{n-1}P^{-1} + \cdots + a_nI_n \\ &= P(B^n + a_1B^{n-1} + \cdots + a_nI_n)P^{-1} \\ &= Pf(B)P^{-1} = O. \end{aligned}$$
2. **几何形式:** 设 $\{e_1, e_2, \cdots, e_n\}$ 是 V 的一组标准基,φ 在这组基下的矩阵为 A, 则由 $f(x)$ 是 φ 的特征多项式可知,$f(x)$ 也是 A 的特征多项式. 从而由代数形式的结论可知 $f(A) = 0$. 于是对 $\forall \alpha \in V$, 都存在 $k_1, k_2, \cdots, k_n$, 使得
$$\alpha = k_1e_1 + k_2e_2 + \cdots + k_ne_n = (e_1, e_2, \cdots, e_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}.$$
两边同时作用 φ 得到
$$\begin{aligned} \varphi(\alpha) &= k_1\varphi(e_1) + k_2\varphi(e_2) + \cdots + k_n\varphi(e_n) = (\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n))\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} \\ &= (e_1, e_2, \cdots, e_n)A\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = A\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = A(e_1, e_2, \cdots, e_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = A\alpha. \end{aligned}$$
因此 $f(\varphi)(\alpha) = f(A)(\alpha) = 0$. 故由 α 的任意性可知 $f(\varphi) = O$.

□

6.5.1 极小多项式的性质

命题 6.41 (极小多项式的性质)

(1) 若 $f(x)$ 是 A 适合的一个多项式, 则 A 的极小多项式 $m(x)$ 整除 $f(x)$.

(2) 任一 n 阶矩阵的极小多项式必唯一.

(3) 相似的矩阵具有相同的极小多项式.

(4) 矩阵及其转置有相同的极小多项式.

(5) 设 $m(x)$ 是 n 阶矩阵 A 的极小多项式, λ_0 是 A 的特征值, 则 $(x-\lambda_0) \mid m(x)$.

(6) 设 A 是一个分块对角阵

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

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.

笔记 性质 (5) 告诉我们: **矩阵的特征值一定是其极小多项式的根**.

证明

(1) 由多项式的带余除法知道

$$f(x)=m(x)q(x)+r(x),$$

且 $\deg r(x)<\deg m(x)$. 将 $x=A$ 代入上式得 $r(A)=O$, 若 $r(x)\neq 0$, 则 A 适合一个比 $m(x)$ 次数更小的非零多项式, 矛盾. 故 $r(x)=0$, 即 $m(x)\mid f(x)$.

(2) 若 $m(x),g(x)$ 都是矩阵 A 的极小多项式, 则由矩阵极小多项式的性质 (1)知道 $m(x)$ 能够整除 $g(x)$, $g(x)$ 也能够整除 $m(x)$. 因此 $m(x)$ 与 $g(x)$ 只差一个常数因子, 又极小多项式必须首项系数为 1, 故 $g(x)=m(x)$.

(3) 设矩阵 A 和 B 相似, 即存在可逆矩阵 P, 使 $B=P^{-1}AP$. 设 A,B 的极小多项式分别为 $m(x),g(x)$, 注意到

$$m(B)=m(P^{-1}AP)=P^{-1}m(A)P=O,$$

因此 $g(x)\mid m(x)$. 同理, $m(x)\mid g(x)$, 故 $m(x)=g(x)$.

(4) 设 $\boldsymbol{A}$ 的极小多项式是 $m(x)$, 转置 $\boldsymbol{A}'$ 的极小多项式是 $n(x)$. 将 $m(\boldsymbol{A})=\boldsymbol{O}$ 转置可得 $m(\boldsymbol{A}')=\boldsymbol{O}$, 因此 $n(x)\mid m(x)$. 同理可证 $m(x)\mid n(x)$, 故 $m(x)=n(x)$.

(5) 由 $m(A)=O$ 及 命题 6.7可得 $m(\lambda_0)=0$, 再由余数定理得 $(x-\lambda_0)\mid m(x)$.

(6) 设 A 的极小多项式为 $m(x)$, A_i 的极小多项式为 $m_i(x)$, 诸 $m_i(x)$ 的最小公倍式为 $g(x)$, 则 $g(A_i)=O$, 于是

$$g(A)=\begin{pmatrix} g(A_1) & & & \\ & g(A_2) & & \\ & & \ddots & \\ & & & g(A_k) \end{pmatrix}=O,$$

从而 $m(x)\mid g(x)$. 又因为

$$m(A)=\begin{pmatrix} m(A_1) & & & \\ & m(A_2) & & \\ & & \ddots & \\ & & & m(A_k) \end{pmatrix}=O,$$

所以对每个 i 有 $m(A_i)=O$, 从而 $m_i(x)\mid m(x)$, 即 $m(x)$ 是 $m_i(x)$ 的公倍式. 又 $g(x)$ 是诸 $m_i(x)$ 的最小公倍式, 故 $g(x)\mid m(x)$. 综上所述,$m(x)=g(x)$.

□

命题 6.42

设数域 $\mathbb{F}$ 上的 n 阶矩阵 $\boldsymbol{A}$ 的极小多项式为 $m(x)$, 求证: $\mathbb{F}[\boldsymbol{A}]=\{f(\boldsymbol{A})|f(x)\in\mathbb{F}[x]\}$ 是 $M_n(\mathbb{F})$ 的子空间, 且 $\dim\mathbb{F}[\boldsymbol{A}]=\deg m(x)$.

证明 容易验证 $\mathbb{F}[\boldsymbol{A}]$ 在矩阵的加法和数乘下封闭, 从而是 $M_n(\mathbb{F})$ 的子空间. 对任一 $f(x)\in\mathbb{F}[x]$, 由多项式的带余除法可知, 存在 $q(x),r(x)\in\mathbb{F}[x]$, 使得 $f(x)=m(x)q(x)+r(x)$, 其中 $\deg r(x)<\deg m(x)=d$, 于是 $f(\boldsymbol{A})=m(\boldsymbol{A})q(\boldsymbol{A})+r(\boldsymbol{A})=r(\boldsymbol{A})$ 是 $\boldsymbol{I}_n,\boldsymbol{A},\cdots,\boldsymbol{A}^{d-1}$ 的线性组合. 另一方面, 若设 $c_0,c_1,\cdots,c_{d-1}\in\mathbb{F}$, 使得

$$c_0\boldsymbol{I}_n+c_1\boldsymbol{A}+\cdots+c_{d-1}\boldsymbol{A}^{d-1}=\boldsymbol{O},$$

则 $\boldsymbol{A}$ 适合多项式 $g(x)=c_{d-1}x^{d-1}+\cdots+c_1x+c_0$, 由矩阵极小多项式的性质 (1)可知 $m(x)\mid g(x)$, 又因为 $d-1=\deg g(x)<\deg m(x)=d$, 所以 $g(x)=0$, 即 $c_0=c_1=\cdots=c_{d-1}=0$, 于是 $\boldsymbol{I}_n,\boldsymbol{A},\cdots,\boldsymbol{A}^{d-1}$ 在 $\mathbb{F}$ 上线性无关. 因此, $\{\boldsymbol{I}_n,\boldsymbol{A},\cdots,\boldsymbol{A}^{d-1}\}$ 是 $\mathbb{F}[\boldsymbol{A}]$ 的一组基, 特别地, $\dim\mathbb{F}[\boldsymbol{A}]=d=\deg m(x)$. □

命题 6.43

n 阶矩阵 A 的极小多项式是其特征多项式的因式. 特别, A 的极小多项式的次数不超过 n.

证明 设 A 的极小多项式和特征多项式分别为 $m(x)$ 和 $f(x)$, 则由Cayley-Hamilton 定理可知 $f(A)=O$, 于是再由矩阵极小多项式的基本性质 (1)可知 $m(x)\mid f(x)$. 又因为特征多项式 $f(x)$ 一定不是零多项式, 所以 $\deg m(x)\leq deg f(x)=n$. □

推论 6.8

n 阶矩阵 A 的极小多项式和特征多项式有相同的根 (不计重数).

证明 设 $m(x)$ 和 $f(x)$ 分别是 n 阶矩阵 $\boldsymbol{A}$ 的极小多项式和特征多项式, 由极小多项式的性质 (5)可知, $f(x)$ 的根 (即特征值) 都是 $m(x)$ 的根. 又由推论 6.43可知, $m(x)\mid f(x)$, 从而 $m(x)$ 的根也都是 $f(x)$ 的根. 因此若不计重数, $m(x)$ 和 $f(x)$ 有相同的根. □

例题 6.39 设 $m(x)$ 和 $f(x)$ 分别是 n 阶矩阵 $\boldsymbol{A}$ 的极小多项式和特征多项式, 求证: $f(x)\mid m(x)^n$.

证明 由于 n 阶矩阵 $\boldsymbol{A}$ 的特征值最多是 n 重的, 因此设 n 阶矩阵 $\boldsymbol{A}$ 的特征值为 $x_i(1\leq i\leq n)$, 即 $f(x)$ 为 $x_i(1\leq i\leq n)$, 并且

$$f(x)=(x-x_1)(x-x_2)\cdots(x-x_n).$$

又由推论 6.8可知 $x_i(1\leq i\leq n)$ 也都是 $m(x)$ 的根. 从而由余数定理可知 $(x-x_i)\mid m(x),i=1,2,\cdots,n$. 于是由整除的基本性质 (6)归纳可得

$$(x-x_1)(x-x_2)\cdots(x-x_n)\mid m^n(x).$$

即 $f(x)\mid m^n(x)$. □

命题 6.44 (常见矩阵的极小多项式)

(1) 若 n 阶矩阵 $\boldsymbol{A}$ 有 n 个不同的特征值, 则极小多项式等于特征多项式. 特别地,n 阶基础循环矩阵的极小多项式等于 x^n-1.

(2) 设 n 阶矩阵 $\boldsymbol{A}$ 可对角化, $\lambda_1,\lambda_2,\cdots,\lambda_k$ 是 $\boldsymbol{A}$ 的全体不同的特征值, 则 $\boldsymbol{A}$ 的极小多项式为 $(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_k)$.

(3) n 阶幂零 Jordan 块的极小多项式是 x^n.

(4) 设 $n(n>1)$ 阶矩阵 $\boldsymbol{A}$ 的秩为 1, 求证: $\boldsymbol{A}$ 的极小多项式为 $x^2-\mathrm{tr}(\boldsymbol{A})x$.

证明

(1) 设 A 的极小多项式和特征多项式分别为 $m(x)$ 和 $f(x)$,A 的 n 个不同的特征值为 $\lambda_i(1\leq i\leq n)$, 则 $f(x)=$

$(x-\lambda_1)\cdots(x-\lambda_n)$. 由推论 6.8可知,$\lambda_i(1\le i\le n)$ 也是 $m(x)$ 的根. 从而

$$(x-\lambda_1)\cdots(x-\lambda_n)\mid m(x).$$

即 $f(x)\mid m(x)$, 又由推论 6.43可知 $m(x)\mid f(x)$, 故 $m(x)=f(x)$.

(2) 设 $\boldsymbol{A}$ 的极小多项式为 $m(x)$. 由 $\boldsymbol{A}$ 可对角化知存在可逆矩阵 $\boldsymbol{P}$, 使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}=\mathrm{diag}\{\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_k\},$$

其中 $\boldsymbol{B}_i=\lambda_i\boldsymbol{I}$ 为纯量矩阵. 显然 $\boldsymbol{B}_i$ 的极小多项式为 $x-\lambda_i$, 故由极小多项式的性质 (3) 和 (6)可得

$$m(x)=m(\boldsymbol{B})=[x-\lambda_1,x-\lambda_2,\cdots,x-\lambda_k]=(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_k).$$

(3) 设 n 阶幂零 Jordan 块为 A, 则由命题??可知 $A^k\ne O(k=1,2,\cdots,n-1)$, 但 $A^n=O$. 故 n 阶幂零 Jordan 块 A 的极小多项式为 x^n.

(4) 由命题 6.39 证法二可知, $\boldsymbol{A}$ 适合多项式 $x^2-\mathrm{tr}(\boldsymbol{A})x$. 显然 $\boldsymbol{A}$ 不可能适合多项式 x. 若 $\boldsymbol{A}$ 适合多项式 $x-\mathrm{tr}(\boldsymbol{A})$, 则 $\boldsymbol{A}=\mathrm{tr}(\boldsymbol{A})\boldsymbol{I}_n$ 为纯量矩阵, 其秩等于 0 或 n, 这与 $\mathrm{r}(\boldsymbol{A})=1$ 矛盾. 因此, $\boldsymbol{A}$ 的极小多项式为 $x^2-\mathrm{tr}(\boldsymbol{A})x$.

□

命题 6.45

设 $f(x)$ 和 $m(x)$ 分别是 m 阶矩阵 $\boldsymbol{A}$ 的特征多项式和极小多项式, $g(x)$ 和 $n(x)$ 分别是 n 阶矩阵 $\boldsymbol{B}$ 的特征多项式和极小多项式, 证明以下结论等价:

(1) $\boldsymbol{A},\boldsymbol{B}$ 没有公共的特征值;

(2) $(f(x),g(x))=1$ 或 $(f(x),n(x))=1$ 或 $(m(x),g(x))=1$ 或 $(m(x),n(x))=1$;

(3) $f(\boldsymbol{B})$ 或 $m(\boldsymbol{B})$ 或 $g(\boldsymbol{A})$ 或 $n(\boldsymbol{A})$ 是可逆矩阵.

证明 (1) $\Leftrightarrow$ (2): 由推论 6.8可知, (2) 中所有的条件都等价. 显然 (1) 与 $(f(x),g(x))=1$ 等价, 故 (1) 与 (2) 等价.

(2) $\Rightarrow$ (3): 例如, 若 $(f(x),n(x))=1$, 则存在 $u(x),v(x)$, 使得 $f(x)u(x)+n(x)v(x)=1$. 将 $x=\boldsymbol{B}$ 代入上式并注意到 $n(\boldsymbol{B})=\boldsymbol{O}$, 故可得 $f(\boldsymbol{B})u(\boldsymbol{B})=\boldsymbol{I}_n$, 这表明 $f(\boldsymbol{B})$ 是可逆矩阵. 将 $x=\boldsymbol{A}$ 代入上式并注意到 $f(\boldsymbol{A})=\boldsymbol{O}$(Cayley-Hamilton 定理), 故可得 $n(\boldsymbol{A})v(\boldsymbol{A})=\boldsymbol{I}_n$, 这表明 $n(\boldsymbol{A})$ 是可逆矩阵. 同理可证其他的情形.

(3) $\Rightarrow$ (1): 设 $\lambda_1,\cdots,\lambda_m$ 是 $\boldsymbol{A}$ 的特征值, 则 $n(\lambda_1),\cdots,n(\lambda_m)$ 是 $n(\boldsymbol{A})$ 的特征值. 例如, 若 $n(\boldsymbol{A})$ 是可逆矩阵, 则 $n(\lambda_i)\ne 0$, 即 λ_i 都不是 $n(x)$ 的根. 由推论 6.8可知,λ_i 都不是 $g(x)$ 的根, 即 $\lambda_1,\cdots,\lambda_m$ 都不是 $\boldsymbol{B}$ 的特征值, 从而 $\boldsymbol{A},\boldsymbol{B}$ 没有公共的特征值. 同理可证其他的情形. □

命题 6.46

设 $f(x)$ 和 $m(x)$ 分别是 n 阶矩阵 $\boldsymbol{A}$ 的特征多项式和极小多项式,$g(x)$ 是一个多项式, 求证:$g(\boldsymbol{A})$ 是可逆矩阵的充要条件是 $(f(x),g(x))=1$ 或 $(m(x),g(x))=1$.

证明 先证充分性, 若 $(f(x),g(x))=1$, 则存在多项式 $u(x),v(x)$, 使得

$$u(x)f(x)+v(x)g(x)=1.$$

又由Cayley-Hamilton 定理可知,$f(\boldsymbol{A})=\boldsymbol{O}$. 从而将 $x=\boldsymbol{A}$ 代入上式得 $v(\boldsymbol{A})g(\boldsymbol{A})=\boldsymbol{I}_n$, 故 $g(\boldsymbol{A})$ 可逆.

若 $(m(x),g(x))=1$, 则存在多项式 $u(x),v(x)$, 使得

$$u(x)m(x)+v(x)g(x)=1.$$

又注意到 $m(\boldsymbol{A})=\boldsymbol{O}$. 从而将 $x=\boldsymbol{A}$ 代入上式得 $v(\boldsymbol{A})g(\boldsymbol{A})=\boldsymbol{I}_n$, 故 $g(\boldsymbol{A})$ 可逆.

再证必要性, 设 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 为 $\boldsymbol{A}$ 的所有特征值, 则 $g(\lambda_1),g(\lambda_2),\cdots,g(\lambda_m)$ 为 $g(\boldsymbol{A})$ 的所有特征值. 又因为 $g(\boldsymbol{A})$ 可逆, 所以其特征值 $g(\lambda_i)\ne 0$（$i=1,2,\cdots,m$）, 即 λ_i 都不是 $g(x)$ 的根. 而由推论 6.8可知,λ_i 是 $f(x),m(x)$ 的全部根. 因此 $f(x),m(x)$ 与 $g(x)$ 没有公共根, 故 $(f(x),g(x))=1,(m(x),g(x))=1$. □

命题 6.47

证明:n 阶方阵 $\boldsymbol{A}$ 为可逆矩阵的充要条件是 $\boldsymbol{A}$ 的极小多项式的常数项不为零.

笔记 也可利用推论 6.8和Vieta 定理来证明.

证明 设 $f(x)$ 和 $m(x)$ 分别是 $\boldsymbol{A}$ 的特征多项式和极小多项式, 则 $m(x) \mid f(x)$. 若 $\boldsymbol{A}$ 可逆, 则 $f(x)$ 的常数项 $(-1)^n|\boldsymbol{A}|$ 不等于零, 因此 $m(x)$ 的常数项也不为零.

反之, 设 $m(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_0$, 其中 $b_0 \neq 0$, 则

$$m(\boldsymbol{A}) = \boldsymbol{A}^m + b_{m-1}\boldsymbol{A}^{m-1} + \cdots + b_0\boldsymbol{I}_n = \boldsymbol{O},$$

于是

$$\boldsymbol{A}(\boldsymbol{A}^{m-1} + b_{m-1}\boldsymbol{A}^{m-2} + \cdots + b_1\boldsymbol{I}_n) = -b_0\boldsymbol{I}_n.$$

由 $b_0 \neq 0$ 即知 $\boldsymbol{A}$ 可逆. □

6.5.2 Cayley-Hamilton 定理的应用: 逆矩阵和伴随矩阵的多项式表示

命题 6.48

设 $\boldsymbol{A}$ 是 n 阶可逆矩阵, 求证:$\boldsymbol{A}^{-1} = g(\boldsymbol{A})$, 其中 $g(x)$ 是一个 $n-1$ 次多项式.

证明 设 $f(x) = x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ 是 $\boldsymbol{A}$ 的特征多项式, 因为 $\boldsymbol{A}$ 可逆, 故 $a_n = (-1)^n|\boldsymbol{A}| \neq 0$. 由Cayley-Hamilton 定理可得 $f(\boldsymbol{A}) = \boldsymbol{O}$, 于是

$$\boldsymbol{A}\left(-\frac{1}{a_n}(\boldsymbol{A}^{n-1} + a_1\boldsymbol{A}^{n-2} + \cdots + a_{n-1}\boldsymbol{I}_n)\right) = \boldsymbol{I}_n.$$

因此

$$\boldsymbol{A}^{-1} = -\frac{1}{a_n}(\boldsymbol{A}^{n-1} + a_1\boldsymbol{A}^{n-2} + \cdots + a_{n-1}\boldsymbol{I}_n).$$

□

命题 6.49

设 $\boldsymbol{A}$ 是 n 阶矩阵, 求证: 伴随矩阵 $\boldsymbol{A}^* = h(\boldsymbol{A})$, 其中 $h(x)$ 是一个 $n-1$ 次多项式.

证明 证法一: 我们用摄动法来证明结论. 设 $f(x) = x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ 是 $\boldsymbol{A}$ 的特征多项式, 其中 $a_n = (-1)^n|\boldsymbol{A}|$. 若 $\boldsymbol{A}$ 是可逆矩阵, 则由命题 6.48可得

$$\boldsymbol{A}^* = |\boldsymbol{A}|\boldsymbol{A}^{-1} = (-1)^{n-1}(\boldsymbol{A}^{n-1} + a_1\boldsymbol{A}^{n-2} + \cdots + a_{n-1}\boldsymbol{I}_n).$$

令 $h(x) = (-1)^{n-1}(x^{n-1} + a_1x^{n-2} + \cdots + a_{n-1})$, 则 $\boldsymbol{A}^* = h(\boldsymbol{A})$, 并且 $h(x)$ 的系数由特征多项式 $f(x)$ 的系数唯一确定.

对于一般的方阵 $\boldsymbol{A}$, 可取到一列有理数 $t_k \to 0$, 使得 $t_k\boldsymbol{I}_n + \boldsymbol{A}$ 为可逆矩阵. 设

$$f_{t_k}(x) = |x\boldsymbol{I}_n - (t_k\boldsymbol{I}_n + \boldsymbol{A})| = x^n + a_1(t_k)x^{n-1} + \cdots + a_{n-1}(t_k)x + a_n(t_k)$$

为 $t_k\boldsymbol{I}_n + \boldsymbol{A}$ 的特征多项式, 则 $a_i(t_k)$ 都是 t_k 的多项式且 $a_i(0) = a_i$ （$1 \leq i \leq n$）. 由可逆矩阵情形的证明可得

$$(t_k\boldsymbol{I}_n + \boldsymbol{A})^* = (-1)^{n-1}\left((t_k\boldsymbol{I}_n + \boldsymbol{A})^{n-1} + a_1(t_k)(t_k\boldsymbol{I}_n + \boldsymbol{A})^{n-2} + \cdots + a_{n-1}(t_k)\boldsymbol{I}_n\right).$$

注意到上式两边的矩阵中的元素都是 t_k 的多项式, 从而关于 t_k 连续. 上式两边同时取极限, 令 $t_k \to 0$, 即得

$$\boldsymbol{A}^* = (-1)^{n-1}(\boldsymbol{A}^{n-1} + a_1\boldsymbol{A}^{n-2} + \cdots + a_{n-1}\boldsymbol{I}_n).$$

因此无论 $\boldsymbol{A}$ 是否可逆, 我们都有 $\boldsymbol{A}^* = h(\boldsymbol{A})$ 成立.

证法二: 若 $\mathrm{r}(A) = n$, 则由 Cayley - Hamilton 定理易证结论成立; 若 $\mathrm{r}(A) \leq n-2$, 则 $A^* = O$, 结论显然成立; 若 $\mathrm{r}(A) = n-1$, 则 $A^* \neq O$ 且 $AA^* = A^*A = O$, 由命题 7.31可知结论也成立. □

6.5.3 Cayley - Hamilton 定理的应用:AX = XB 型矩阵方程的求解及其应用

命题 6.50

设 $\boldsymbol{A}$ 为 m 阶矩阵,$\boldsymbol{B}$ 为 n 阶矩阵, 求证: 若 $\boldsymbol{A},\boldsymbol{B}$ 没有公共的特征值, 则矩阵方程 $\boldsymbol{AX}=\boldsymbol{XB}$ 只有零解 $\boldsymbol{X}=\boldsymbol{O}$.

证明 证法一: 设 $f(\lambda)=|\lambda\boldsymbol{I}_m-\boldsymbol{A}|$ 为 $\boldsymbol{A}$ 的特征多项式, 则由Cayley-Hamilton 定理可知 $f(\boldsymbol{A})=\boldsymbol{O}$, 再由 $\boldsymbol{AX}=\boldsymbol{XB}$ 可得

$$\boldsymbol{O}=f(\boldsymbol{A})\boldsymbol{X}=\boldsymbol{X}f(\boldsymbol{B}).$$

因为 $\boldsymbol{A},\boldsymbol{B}$ 没有公共的特征值, 故由命题 6.45可知,$f(\boldsymbol{B})$ 是可逆矩阵, 从而由上式即得 $\boldsymbol{X}=\boldsymbol{O}$.

证法二: 任取矩阵方程的一个解 $\boldsymbol{X}=\boldsymbol{C}$, 若 $\boldsymbol{C}\neq\boldsymbol{O}$, 则 $\mathrm{r}(\boldsymbol{C})=r\geq 1$. 由例题 6.12可知,$\boldsymbol{A},\boldsymbol{B}$ 至少有 r 个相同的特征值, 这与 $\boldsymbol{A},\boldsymbol{B}$ 没有公共的特征值相矛盾. 因此 $\boldsymbol{C}=\boldsymbol{O}$, 即矩阵方程只有零解.□ □

例题 6.40 设 n 阶方阵 $\boldsymbol{A},\boldsymbol{B}$ 的特征值全部大于零且满足 $\boldsymbol{A}^2=\boldsymbol{B}^2$, 求证:$\boldsymbol{A}=\boldsymbol{B}$.

证明 由 $\boldsymbol{A}^2=\boldsymbol{B}^2$ 可得 $\boldsymbol{A}(\boldsymbol{A}-\boldsymbol{B})=(\boldsymbol{A}-\boldsymbol{B})(-\boldsymbol{B})$, 即 $\boldsymbol{A}-\boldsymbol{B}$ 是矩阵方程 $\boldsymbol{AX}=\boldsymbol{X}(-\boldsymbol{B})$ 的解. 注意到 $\boldsymbol{A}$ 的特征值全部大于零,$-\boldsymbol{B}$ 的特征值全部小于零, 故它们没有公共的特征值, 由命题 6.50可得 $\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{O}$, 即 $\boldsymbol{A}=\boldsymbol{B}$. □

例题 6.41 设 $\boldsymbol{A}=\mathrm{diag}\{\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m\}$ 为 n 阶分块对角矩阵, 其中 $\boldsymbol{A}_i$ 是 n_i 阶矩阵且两两没有公共的特征值. 设 $\boldsymbol{B}$ 是 n 阶矩阵, 满足 $\boldsymbol{AB}=\boldsymbol{BA}$, 求证:$\boldsymbol{B}=\mathrm{diag}\{\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_m\}$, 其中 $\boldsymbol{B}_i$ 也是 n_i 阶矩阵.

证明 按照 $\boldsymbol{A}$ 的分块方式对 $\boldsymbol{B}$ 进行分块, 可设 $\boldsymbol{B}=(\boldsymbol{B}_{ij})$, 其中 $\boldsymbol{B}_{ij}$ 是 $n_i\times n_j$ 矩阵. 由 $\boldsymbol{AB}=\boldsymbol{BA}$ 可知, 对任意的 i,j, 有 $\boldsymbol{A}_i\boldsymbol{B}_{ij}=\boldsymbol{B}_{ij}\boldsymbol{A}_j$. 因为 $\boldsymbol{A}_i,\boldsymbol{A}_j$（$i\neq j$）没有公共的特征值, 故由命题 6.50可得 $\boldsymbol{B}_{ij}=\boldsymbol{O}$（$i\neq j$）, 从而 $\boldsymbol{B}=\mathrm{diag}\{\boldsymbol{B}_{11},\boldsymbol{B}_{22},\cdots,\boldsymbol{B}_{mm}\}$ 也是分块对角矩阵. □

命题 6.51

设 $\boldsymbol{A},\boldsymbol{B}$ 分别为 m,n 阶矩阵,V 为 $m\times n$ 矩阵全体构成的线性空间,V 上的线性变换 φ 定义为:$\varphi(\boldsymbol{X})=\boldsymbol{AX}-\boldsymbol{XB}$. 求证:$\varphi$ 是线性自同构的充要条件是 $\boldsymbol{A},\boldsymbol{B}$ 没有公共的特征值. 此时, 对任一 $m\times n$ 矩阵 $\boldsymbol{C}$, 矩阵方程 $\boldsymbol{AX}-\boldsymbol{XB}=\boldsymbol{C}$ 存在唯一解.

注 由证法二不难看出, 这个命题的结论在数域 $\mathbb{F}$ 上也成立.

证明 证法一: 若 $\boldsymbol{A},\boldsymbol{B}$ 没有公共的特征值, 则由命题 6.50可知,$\varphi(X)=AX-XB=0$ 只有零解, 即 $\mathrm{Ker}\varphi=0$. 从而 φ 是 V 上的单映射, 从而是线性自同构. 若 $\boldsymbol{A},\boldsymbol{B}$ 有公共的特征值 λ_0, 则 λ_0 也是 $\boldsymbol{B}'$ 的特征值. 设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 为对应的特征向量, 即 $\boldsymbol{A\alpha}=\lambda_0\boldsymbol{\alpha},\boldsymbol{B}'\boldsymbol{\beta}=\lambda_0\boldsymbol{\beta}$, 则 $\boldsymbol{\alpha\beta}'\neq\boldsymbol{O}$ 且

$$\varphi(\boldsymbol{\alpha\beta}')=(\boldsymbol{A\alpha})\boldsymbol{\beta}'-\boldsymbol{\alpha}(\boldsymbol{B}'\boldsymbol{\beta})'=\lambda_0\boldsymbol{\alpha\beta}'-\lambda_0\boldsymbol{\alpha\beta}'=\boldsymbol{O},$$

于是 $\mathrm{Ker}\varphi\neq 0$, 从而 φ 不是线性自同构.

证法二: 由命题 6.56知,φ 的表示矩阵为 $\boldsymbol{A}\otimes\boldsymbol{I}_n-\boldsymbol{I}_m\otimes\boldsymbol{B}'$, 其特征值为 $\lambda_i-\mu_j(1\leq i\leq m;1\leq j\leq n)$, 其中 λ_i,μ_j 分别为 A,B 的特征值. 因此 φ 是 V 上的线性自同构当且仅当其表示矩阵 $\boldsymbol{A}\otimes\boldsymbol{I}_n-\boldsymbol{I}_m\otimes\boldsymbol{B}'$ 是可逆矩阵, 这当且仅当 φ 的特征值 $\lambda_i-\mu_j(1\leq i\leq m;1\leq j\leq n)$ 全都非零. 这也当且仅当 A,B 在复数域中没有公共的特征值. □

例题 6.42 设 n 阶实矩阵 $\boldsymbol{A}$ 的所有特征值都是正实数, 证明: 对任一实对称矩阵 $\boldsymbol{C}$, 存在唯一的实对称矩阵 $\boldsymbol{B}$, 满足 $\boldsymbol{A}'\boldsymbol{B}+\boldsymbol{BA}=\boldsymbol{C}$.

证明 考虑矩阵方程 $\boldsymbol{A}'\boldsymbol{X}-\boldsymbol{X}(-\boldsymbol{A})=\boldsymbol{C}$, 注意到 $\boldsymbol{A}'$ 的特征值全部大于零,$-\boldsymbol{A}$ 的特征值全部小于零, 它们没有公共的特征值, 故由命题 6.51可得上述矩阵方程存在唯一解 $\boldsymbol{X}=\boldsymbol{B}$. 容易验证 $\boldsymbol{X}=\overline{\boldsymbol{B}},\boldsymbol{B}'$ 也都是上述矩阵方程的解, 故由解的唯一性可知 $\boldsymbol{B}=\overline{\boldsymbol{B}}$ 且 $\boldsymbol{B}=\boldsymbol{B}'$, 即 $\boldsymbol{B}$ 为实对称矩阵, 结论得证. □

6.5.4 Cayley-Hamilton 定理的应用: 特征多项式诱导的直和分解

例题 6.43 设 φ 是复线性空间 V 上的线性变换, 又有两个复系数多项式:

$$f(x)=x^m+a_1x^{m-1}+\cdots+a_m,\quad g(x)=x^n+b_1x^{n-1}+\cdots+b_n.$$

设 $\sigma = f(\varphi), \tau = g(\varphi)$, 矩阵 C 是 $f(x)$ 的友阵, 即

$$C = \begin{pmatrix} 0 & 0 & 0 & \cdots & -a_m \\ 1 & 0 & 0 & \cdots & -a_{m-1} \\ 0 & 1 & 0 & \cdots & -a_{m-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & -a_1 \end{pmatrix}.$$

若 $g(C)$ 是可逆矩阵, 求证: $\operatorname{Ker}\sigma\tau = \operatorname{Ker}\sigma \oplus \operatorname{Ker}\tau$.

笔记 $(f(x), g(x)) = 1$ 之后的证明类似命题 5.31.

证明 由命题 6.6可知 C 的特征多项式就是 $f(x)$. 由命题 6.46可知 $(f(x), g(x)) = 1$. 由 $(f(x), g(x)) = 1$ 可知, 存在多项式 $u(x), v(x)$, 使得

$$u(x)f(x) + v(x)g(x) = 1.$$

从而

$$u(\varphi)f(\varphi) + v(\varphi)g(\varphi) = I_V. \tag{6.15}$$

于是对 $\forall \alpha \in \operatorname{Ker}\sigma\tau$, 由 (6.15)式可得

$$\alpha = u(\varphi)f(\varphi)(\alpha) + v(\varphi)g(\varphi)(\alpha).$$

又因为 $\alpha \in \operatorname{Ker}\sigma\tau$, 所以 $f(\varphi)g(\varphi)(\alpha) = g(\varphi)f(\varphi)(\alpha) = 0$. 因此 $u(\varphi)f(\varphi)(\alpha) \in \operatorname{Ker} g(\varphi)$,$v(\varphi)g(\varphi)(\alpha) \in \operatorname{Ker} f(\varphi)$. 故有 $\operatorname{Ker}\sigma\tau = \operatorname{Ker}\sigma + \operatorname{Ker}\tau$. 任取 $\beta \in \operatorname{Ker}\sigma \cap \operatorname{Ker}\tau$, 则 $\sigma(\beta) = f(\varphi)(\beta) = 0$,$\tau(\beta) = g(\varphi)(\beta) = 0$. 由(6.15) 式可得

$$\beta = u(\varphi)f(\varphi)(\beta) + v(\varphi)g(\varphi)(\beta) = 0.$$

故 $\operatorname{Ker}\sigma \cap \operatorname{Ker}\tau = 0$, 因此 $\operatorname{Ker}\sigma\tau = \operatorname{Ker}\sigma \oplus \operatorname{Ker}\tau$. □

命题 6.52

设 φ 是数域 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换, 其特征多项式是 $f(\lambda)$ 且 $f(\lambda) = f_1(\lambda)f_2(\lambda)$, 其中 $f_1(\lambda), f_2(\lambda)$ 是互素的首一多项式. 令 $V_1 = \operatorname{Ker} f_1(\varphi)$,$V_2 = \operatorname{Ker} f_2(\varphi)$, 求证:

(1) V_1, V_2 是 φ-不变子空间且 $V = V_1 \oplus V_2$;

(2) $V_1 = \operatorname{Im} f_2(\varphi)$,$V_2 = \operatorname{Im} f_1(\varphi)$;

(3) $\varphi|_{V_1}$ 的特征多项式是 $f_1(\lambda)$,$\varphi|_{V_2}$ 的特征多项式是 $f_2(\lambda)$.

笔记 这个命题是命题 5.31的推广.

注 (3) 中 $f_1(\lambda) = g_1(\lambda), f_2(\lambda) = g_2(\lambda)$ 的原因: 由于 $f_i(\lambda)$ 与 $g_i(\lambda)$ 的根相同, 且 $f_1(\lambda)$ 与 $f_2(\lambda)$ 没有公共根, 因此不妨设

$$f_1(\lambda) = (\lambda - x_1)^{i_1} \cdots (\lambda - x_s)^{i_s}, \quad f_2(\lambda) = (\lambda - y_1)^{j_1} \cdots (\lambda - y_l)^{j_l}, \tag{6.16}$$

$$g_1(\lambda) = (\lambda - x_1)^{i_1'} \cdots (\lambda - x_s)^{i_s'}, \quad g_2(\lambda) = (\lambda - y_1)^{j_1'} \cdots (\lambda - y_l)^{j_l'}. \tag{6.17}$$

其中 $x_1, \cdots, x_s, y_1, \cdots, y_l$ 互不相同. 则

$$f_1(\lambda)f_2(\lambda) = [(\lambda - x_1)^{i_1} \cdots (\lambda - x_s)^{i_s}][(\lambda - y_1)^{j_1} \cdots (\lambda - y_l)^{j_l}],$$
$$g_1(\lambda)g_2(\lambda) = [(\lambda - x_1)^{i_1'} \cdots (\lambda - x_s)^{i_s'}][(\lambda - y_1)^{j_1'} \cdots (\lambda - y_l)^{j_l'}].$$

又由 $f(\lambda) = f_1(\lambda)f_2(\lambda) = g_1(\lambda)g_2(\lambda)$ 可得

$$[(\lambda - x_1)^{i_1} \cdots (\lambda - x_s)^{i_s}][(\lambda - y_1)^{j_1} \cdots (\lambda - y_l)^{j_l}] = [(\lambda - x_1)^{i_1'} \cdots (\lambda - x_s)^{i_s'}][(\lambda - y_1)^{j_1'} \cdots (\lambda - y_l)^{j_l'}].$$

比较上式两边的常数项可得

$$x_1^{i_1} \cdots x_s^{i_s} y_1^{j_1} \cdots y_l^{j_l} = x_1^{i_1'} \cdots x_s^{i_s'} y_1^{j_1'} \cdots y_l^{j_l'}.$$

又因为 $x_1,\cdots,x_s,y_1,\cdots,y_l$ 互不相同, 所以

$$i_1=i_1',\cdots,i_s=i_s',j_1=j_1',\cdots,j_l=j_l'.$$

再由(6.16)和(6.17)式可知 $f_1(\lambda)=g_1(\lambda),f_2(\lambda)=g_2(\lambda)$.

证明

(1) 对 $\forall\alpha\in V_1$, 都有 $f_1(\varphi)(\alpha)=0$. 从而

$$f_1(\varphi)(\varphi(\alpha))=(f_1(\varphi)\varphi)(\alpha)=(\varphi f_1(\varphi))(\alpha)=\varphi(f_1(\varphi)(\alpha))=\varphi(0)=0.$$

故 V_1 是 φ-不变子空间, 同理可得 V_2 也是 φ-不变子空间. 由 Cayley - Hamilton 定理可得 $f(\varphi)=f_1(\varphi)f_2(\varphi)=\mathbf{0}$, 故由命题 5.31可知 $V=V_1\oplus V_2$.

(2) 由 $f_1(\varphi)f_2(\varphi)=\mathbf{0}$ 可得 $\mathrm{Im}\,f_2(\varphi)\subseteq\mathrm{Ker}\,f_1(\varphi)=V_1$,$\mathrm{Im}\,f_1(\varphi)\subseteq\mathrm{Ker}\,f_2(\varphi)=V_2$. 因为 $V=V_1\oplus V_2$, 故由维数公式可得

$$\dim\mathrm{Im}\,f_2(\varphi)=\dim V-\dim\mathrm{Ker}\,f_2(\varphi)=\dim V-\dim V_2=\dim V_1,$$

$$\dim\mathrm{Im}\,f_1(\varphi)=\dim V-\dim\mathrm{Ker}\,f_1(\varphi)=\dim V-\dim V_1=\dim V_2,$$

从而 $V_1=\mathrm{Im}\,f_2(\varphi)$,$V_2=\mathrm{Im}\,f_1(\varphi)$.

(3) 设 $\varphi|_{V_i}$ 的特征多项式为 $g_i(\lambda)(i=1,2)$, 则由命题 6.2可得

$$f(\lambda)=f_1(\lambda)f_2(\lambda)=g_1(\lambda)g_2(\lambda). \tag{6.18}$$

注意到 $f_i(\varphi|_{V_i})=f_i(\varphi)|_{V_i}=\mathbf{0}$, 即 $\varphi|_{V_i}$ 适合多项式 $f_i(\lambda)$, 因此 $\varphi|_{V_i}$ 的特征值也适合 $f_i(\lambda)$, 即 $g_i(\lambda)$ 的根都是 $f_i(\lambda)$ 的根. 因为 $(f_1(\lambda),f_2(\lambda))=1$, 故 $f_1(\lambda)$ 与 $f_2(\lambda)$ 没有公共根, 从而由 $f_i(\lambda)$ 的首一性和(6.18) 式即得 $f_1(\lambda)=g_1(\lambda),f_2(\lambda)=g_2(\lambda)$.

□

这个命题的结论还可以进一步推广, 例如不限定 $f(\lambda)$ 是 φ 的特征多项式, 而只要求 φ 适合它 (比如 φ 的极小多项式 $m(\lambda)$), 则由完全相同的讨论可以证明此时对这个命题的 (1) 和 (2) 都成立.

命题 6.53

设 φ 是数域 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换, 若其适合多项式 $f(\lambda)$ 且 $f(\lambda)=f_1(\lambda)f_2(\lambda)$, 其中 $f_1(\lambda),f_2(\lambda)$ 是互素的首一多项式. 令 $V_1=\mathrm{Ker}\,f_1(\varphi)$,$V_2=\mathrm{Ker}\,f_2(\varphi)$, 求证:

(1) V_1,V_2 是 φ-不变子空间且 $V=V_1\oplus V_2$;

(2) $V_1=\mathrm{Im}\,f_2(\varphi)$,$V_2=\mathrm{Im}\,f_1(\varphi)$;

(3) $\varphi|_{V_1}$ 适合多项式 $f_1(\lambda)$,$\varphi|_{V_2}$ 适合多项式 $f_2(\lambda)$.

设 φ 的极小多项式为 $m(\lambda)$, 特别地, 如果 $f(\lambda)=m(\lambda)$, 并且考虑极小多项式的首一互素因式分解

$$m(\lambda)=m_1(\lambda)m_2(\lambda),V_1=\mathrm{Ker}\,m_1(\varphi),V_2=\mathrm{Ker}\,m_2(\varphi),$$

则 $\varphi|_{V_i}$ 的极小多项式就是 $m_i(\lambda)$.

注 这个命题是命题 6.52的推广.

证明 由命题 6.52完全类似的讨论可证. □

6.5.5 Cayley-Hamilton 定理的其他应用

例题 6.44 设 A 为 n 阶矩阵,C 为 $k\times n$ 矩阵, 且对任意的 $\lambda\in\mathbb{C}$, $\begin{pmatrix}A-\lambda I_n\\ C\end{pmatrix}$ 均为列满秩阵. 证明: 对任意的 $\lambda\in\mathbb{C}$, $\begin{pmatrix}C\\ C(A-\lambda I_n)\\ C(A-\lambda I_n)^2\\ \vdots\\ C(A-\lambda I_n)^{n-1}\end{pmatrix}$ 均为列满秩阵.

证明 由推论 3.4可知, 对任意的 $\lambda\in\mathbb{C}$, 下列线性方程组只有零解:

$$\begin{cases}(A-\lambda I_n)x=0,\\ Cx=0.\end{cases}\tag{6.19}$$

而要证明结论, 根据推论 3.4可知, 只要证明对任意的 $\lambda\in\mathbb{C}$, 下列线性方程组只有零解即可:

$$\begin{cases}Cx=0,\\ C(A-\lambda I_n)x=0,\\ C(A-\lambda I_n)^2x=0,\\ \cdots\cdots\cdots\\ C(A-\lambda I_n)^{n-1}x=0.\end{cases}\tag{6.20}$$

任取 $\lambda_0\in\mathbb{C}$ 以及对应线性方程组 (6.20)的任一解 x_0, 则由线性方程组 (6.20)可得 $Cx_0=0$,$CAx_0=0$,$\cdots$,$CA^{n-1}x_0=0$, 因此对任意次数小于 n 的多项式 $g(x)$, 均有 $Cg(A)x_0=0$. 设

$$f(\lambda)=|\lambda I_n-A|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

为 A 的特征多项式, 则由 Cayley - Hamilton 定理可得

$$(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n)=O.$$

因此 $y=(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n)x_0$ 既满足 $(A-\lambda_1 I_n)y=0$, 又满足 $Cy=0$, 故由线性方程组(6.19)只有零解可得 $y=(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n)x_0=0$. 不断重复上述论证, 最后可得 $x_0=0$, 结论得证. □

例题 6.45 设 A 是 n 阶矩阵,B 是 $n\times m$ 矩阵, 分块矩阵 $(B,AB,\cdots,A^{n-2}B,A^{n-1}B)$ 的秩为 r. 证明: 存在 n 阶可逆矩阵 P, 使得

$$P^{-1}AP=\begin{pmatrix}A_{11}&A_{12}\\ O&A_{22}\end{pmatrix},\quad P^{-1}B=\begin{pmatrix}B_1\\ O\end{pmatrix},$$

其中 A_{11} 是 r 阶矩阵,B_1 是 $r\times m$ 矩阵.

注 $A\alpha_i(1\le i\le r)$ 都是 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的线性组合的原因: 因为 $\alpha_1,\cdots,\alpha_r$ 是 $(B,AB,\cdots,A^{n-2}B,A^{n-1}B)$ 列向量的极大无关组, 所以对 $\forall i\in\{1,2,\cdots,r\}$, 都存在 $k\in\{0,1,\cdots,n-1\}$, 使得 α_i 是 A^kB 的某一列向量.

当 α_i 是 $A^kB(0\leqslant k\leqslant n-2)$ 的某一列向量时, 则 $A\alpha_i$ 一定是 $A^{k+1}B$ 的某一列向量, 又由于 $1\leqslant k+1\leqslant n-1$, 因此 $A\alpha_i$ 仍是 $(B,AB,\cdots,A^{n-2}B,A^{n-1}B)$ 的某一列向量, 从而 $A\alpha_i$ 可由 $\alpha_1,\cdots,\alpha_r$ 线性表出.

当 α_i 是 $A^{n-1}B$ 的某一列向量时, 则 $A\alpha_i$ 一定是 A^nB 的某一列向量. 由(6.21)式可知

$$A^nB=-a_1A^{n-1}B-\cdots-a_{n-1}AB-a_nB.$$

而上式右边的每一个列向量都可以由 $\alpha_1,\cdots,\alpha_r$ 线性表出, 于是 A^nB 的每一个列向量都可以由 $\alpha_1,\cdots,\alpha_r$ 线性表出. 故 $A\alpha_i$ 也可以由 $\alpha_1,\cdots,\alpha_r$ 线性表出.

证明 设 $(B,AB,\cdots,A^{n-2}B,A^{n-1}B)$ 列向量的极大无关组为 $\alpha_1,\alpha_2,\cdots,\alpha_r$, 由基扩张定理可将其扩张为 $\mathbb{F}^n$ 的一组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 令 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 则 P 为可逆矩阵. 设 A 的特征多项式为 $f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+$

$a_{n-1}\lambda + a_n$, 则由Cayley - Hamilton 定理可得

$$f(A) = A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI_n = O,$$

从而

$$A^nB = -a_1A^{n-1}B - \cdots - a_{n-1}AB - a_nB. \tag{6.21}$$

由上式容易验证 $A\alpha_i (1 \le i \le r)$ 都是 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 的线性组合, 于是 $AP = P\begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$, 即有 $P^{-1}AP = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$. 又 B 的列向量都是 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 的线性组合, 于是 $B = P\begin{pmatrix} B_1 \\ O \end{pmatrix}$, 即有 $P^{-1}B = \begin{pmatrix} B_1 \\ O \end{pmatrix}$. □

例题 6.46 设 A 是数域 $\mathbb{F}$ 上的 n 阶矩阵, 递归地定义矩阵序列 $\{A_k\}_{k=1}^{\infty}$:

$$A_1 = A, \quad p_k = -\frac{1}{k}\mathrm{tr}(A_k), \quad A_{k+1} = A(A_k + p_kI_n), \quad k = 1, 2, \cdots.$$

求证: $A_{n+1} = O$.

证明 设 A 的全体特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 它们的幂和记为 $s_k = \sum_{i=1}^{n} \lambda_i^k = \mathrm{tr}(A^k)$, 它们的初等对称多项式记为 σ_k, 则 A 的特征多项式为

$$f(\lambda) = \lambda^n - \sigma_1\lambda^{n-1} + \cdots + (-1)^{n-1}\sigma_{n-1}\lambda + (-1)^n\sigma_n.$$

下面用归纳法证明: $p_k = (-1)^k\sigma_k (1 \le k \le n)$. $p_1 = -\mathrm{tr}(A) = -\sigma_1$, 结论成立. 假设小于等于 k 时结论成立, 则 $A_{k+1} = A^{k+1} - \sigma_1A^k + \cdots + (-1)^k\sigma_kA$. 由 Newton 公式可得

$$p_{k+1} = -\frac{1}{k+1}\mathrm{tr}(A_{k+1}) = -\frac{1}{k+1}(s_{k+1} - s_k\sigma_1 + \cdots + (-1)^k s_1\sigma_k) = (-1)^{k+1}\sigma_{k+1},$$

结论得证. 最后, 由Cayley - Hamilton 定理可得

$$A_{n+1} = A^{n+1} - \sigma_1A^n + \cdots + (-1)^n\sigma_nA = f(A)A = O.$$

□

6.6 矩阵的 Kronecker 积

定义 6.17 (矩阵的 Kronecker 积)

设 $A = (a_{ij})$ 和 $B = (b_{ij})$ 分别是数域 $\mathbb{F}$ 上的 $m \times n$ 和 $k \times l$ 矩阵, 它们的 **Kronecker 积** $A \otimes B$ 是 $\mathbb{F}$ 上的 $mk \times nl$ 矩阵:

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

♣

定理 6.14 (矩阵的 Kronecker 积的基本性质)

证明矩阵的 Kronecker 积满足下列性质 (假设以下的矩阵加法和乘法都有意义):

(1) $(A + B) \otimes C = A \otimes C + B \otimes C$, $A \otimes (B + C) = A \otimes B + A \otimes C$;

(2) $(kA) \otimes B = k(A \otimes B) = A \otimes (kB)$;

(3) $(A \otimes C)(B \otimes D) = (AB) \otimes (CD)$;

(4) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$;

(5) $I_m \otimes I_n = I_{mn}$;

(6) $(A \otimes B)' = A' \otimes B'$;

(7) 若 A, B 都是可逆矩阵, 则 $A \otimes B$ 也是可逆矩阵, 并且

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1};$$

(8) 若 A 是 m 阶矩阵,B 是 n 阶矩阵, 则 $|A \otimes B| = |A|^n|B|^m$;

(9) 若 A 是 m 阶矩阵,B 是 n 阶矩阵, 则 $\mathrm{tr}(A \otimes B) = \mathrm{tr}(A) \cdot \mathrm{tr}(B)$.

(10) 设 A, B 均为上三角阵, 且 A, B 的主对角元素分别依次为 $a_1, \cdots, a_n$ 和 $b_1, \cdots, b_m$, 则 $A \otimes B$ 仍是上三角阵, 且 $A \otimes B$ 的主对角元素依次为 $a_1b_1, \cdots, a_1b_m, a_2b_1, \cdots, a_2b_m, \cdots, a_nb_1, \cdots, a_nb_m$.

(11) 设 A, B 均为对角阵, 且 A, B 的主对角元素分别依次为 $a_1, \cdots, a_n$ 和 $b_1, \cdots, b_m$, 则 $A \otimes B$ 仍是对角阵, 且 $A \otimes B$ 的主对角元素依次为 $a_1b_1, \cdots, a_1b_m, a_2b_1, \cdots, a_2b_m, \cdots, a_nb_1, \cdots, a_nb_m$. ♡

证明

(1) 由 Kronecker 积的定义经简单计算即可验证.

(2) 由 Kronecker 积的定义经简单计算即可验证.

(3) 设 $A = (a_{ij})$ 是 $m \times p$ 矩阵,$B = (b_{ij})$ 是 $p \times n$ 矩阵,$C = (c_{ij})$ 是 $k \times q$ 矩阵,$D = (d_{ij})$ 是 $q \times l$ 矩阵. 由 Kronecker 积的定义以及分块矩阵的乘法可得

$$\begin{aligned}(A \otimes C)(B \otimes D) &= \begin{pmatrix} a_{11}C & a_{12}C & \cdots & a_{1p}C \\ a_{21}C & a_{22}C & \cdots & a_{2p}C \\ \vdots & \vdots & & \vdots \\ a_{m1}C & a_{m2}C & \cdots & a_{mp}C \end{pmatrix} \begin{pmatrix} b_{11}D & b_{12}D & \cdots & b_{1n}D \\ b_{21}D & b_{22}D & \cdots & b_{2n}D \\ \vdots & \vdots & & \vdots \\ b_{p1}D & b_{p2}D & \cdots & b_{pn}D \end{pmatrix} \\ &= \begin{pmatrix} \sum\limits_{j=1}^{p} a_{1j}b_{j1}CD & \sum\limits_{j=1}^{p} a_{1j}b_{j2}CD & \cdots & \sum\limits_{j=1}^{p} a_{1j}b_{jn}CD \\ \sum\limits_{j=1}^{p} a_{2j}b_{j1}CD & \sum\limits_{j=1}^{p} a_{2j}b_{j2}CD & \cdots & \sum\limits_{j=1}^{p} a_{2j}b_{jn}CD \\ \vdots & \vdots & & \vdots \\ \sum\limits_{j=1}^{p} a_{mj}b_{j1}CD & \sum\limits_{j=1}^{p} a_{mj}b_{j2}CD & \cdots & \sum\limits_{j=1}^{p} a_{mj}b_{jn}CD \end{pmatrix} \\ &= (AB) \otimes (CD).\end{aligned}$$

(4) 设 $A = (a_{ij})$, $B = (b_{ij})$ 和 $C = (c_{ij})$ 分别是 $m \times n$, $k \times l$ 和 $p \times q$ 矩阵, 则经计算即可发现 $(A \otimes B) \otimes C$ 和 $A \otimes (B \otimes C)$ 都等于下面的 $mkp \times nlq$ 矩阵:

$$\begin{pmatrix} a_{11}b_{11}C & \cdots & a_{11}b_{1l}C & \cdots & a_{1n}b_{11}C & \cdots & a_{1n}b_{1l}C \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{11}b_{k1}C & \cdots & a_{11}b_{kl}C & \cdots & a_{1n}b_{k1}C & \cdots & a_{1n}b_{kl}C \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}b_{11}C & \cdots & a_{m1}b_{1l}C & \cdots & a_{mn}b_{11}C & \cdots & a_{mn}b_{1l}C \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}b_{k1}C & \cdots & a_{m1}b_{kl}C & \cdots & a_{mn}b_{k1}C & \cdots & a_{mn}b_{kl}C \end{pmatrix}.$$

(5) 由 Kronecker 积的定义经简单计算即可验证.

(6) 由 Kronecker 积的定义经简单计算即可验证.

(7) 由 (3) 和 (5) 可得

$$(A \otimes B)(A^{-1} \otimes B^{-1}) = (AA^{-1}) \otimes (BB^{-1}) = I_m \otimes I_n = I_{mn}.$$

(8) 由 Laplace 定理容易证明:

$$|A\otimes I_n| = |A|^n,\quad |I_m\otimes B| = |B|^m;$$

再由 (3) 以及矩阵乘积的行列式等于行列式的乘积可得

$$|A\otimes B| = |(A\otimes I_n)(I_m\otimes B)| = |A\otimes I_n||I_m\otimes B| = |A|^n|B|^m.$$

(9) 由 Kronecker 积的定义经简单计算即可验证.

(10) 由 Kronecker 积的定义经简单计算即可验证.

(11) 由 Kronecker 积的定义经简单计算即可验证.

□

命题 6.54 (矩阵的 Kronecker 积的秩)

设 A, B 分别为 $m\times n, k\times l$ 矩阵, 求证: $\mathrm{r}(A\otimes B) = \mathrm{r}(A)\cdot\mathrm{r}(B)$.

证明 设 $\mathrm{r}(A) = r, \mathrm{r}(B) = s, P, Q, R, S$ 为可逆矩阵, 使得

$$PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix},\quad RBS = \begin{pmatrix} I_s & O \\ O & O \end{pmatrix},$$

则由性质 (7) 可知 $P\otimes R, Q\otimes S$ 均非异, 再由性质 (3)可得

$$(P\otimes R)(A\otimes B)(Q\otimes S) = (PAQ)\otimes(RBS) \sim \begin{pmatrix} I_{rs} & O \\ O & O \end{pmatrix},$$

于是 $\mathrm{r}(A\otimes B) = rs = \mathrm{r}(A)\cdot\mathrm{r}(B)$. □

推论 6.9

设 A, B 分别为 $m\times n, k\times l$ 矩阵, 求证: $A\otimes B$ 是行满秩阵 (列满秩阵) 的充要条件是 A, B 均为行满秩阵 (列满秩阵).

证明 由矩阵的 Kronecker 积的秩可知

$$\mathrm{r}(A\otimes B) = \mathrm{r}(A)\cdot\mathrm{r}(B).$$

于是立得结论. □

命题 6.55 (矩阵的 Kronecker 积的特征值)

设 A, B 分别是 m, n 阶矩阵,A 的特征值为 $\lambda_i\ (1\le i\le m)$,B 的特征值为 $\mu_j\ (1\le j\le n)$, 求证: $A\otimes B$ 的特征值为 $\lambda_i\mu_j\ (1\le i\le m; 1\le j\le n)$.

证明 由命题 6.20可知, 存在 m 阶可逆矩阵 P 以及 n 阶可逆矩阵 Q, 使得

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * & * & * \\ & \lambda_2 & * & * \\ & & \ddots & \vdots \\ & & & \lambda_m \end{pmatrix},\quad Q^{-1}BQ = \begin{pmatrix} \mu_1 & * & * & * \\ & \mu_2 & * & * \\ & & \ddots & \vdots \\ & & & \mu_n \end{pmatrix}.$$

由性质 (10)可知 $(P^{-1}AP)\otimes(Q^{-1}BQ)$ 仍是上三角矩阵且 $(P^{-1}AP)\otimes(Q^{-1}BQ)$ 的主对角元素依次为

$$\lambda_1\mu_1,\cdots,\lambda_1\mu_n,\lambda_2\mu_1,\cdots,\lambda_2\mu_n,\cdots,\lambda_m\mu_1\cdots,\lambda_m\mu_n.$$

注意到 $(P^{-1}AP)\otimes(Q^{-1}BQ) = (P\otimes Q)^{-1}(A\otimes B)(P\otimes Q)$, 因此 $(P^{-1}AP)\otimes(Q^{-1}BQ)$ 和 $A\otimes B$ 相似, 又相似矩阵特征值相同, 故结论得证. □

命题 6.56

设 A, B 分别为 m, n 阶矩阵,V 为 $m \times n$ 矩阵全体构成的线性空间,V 上的线性变换 φ 定义为: $\varphi(\boldsymbol{X}) = \boldsymbol{AXB}$. 设 $\boldsymbol{A}$ 的特征值为 $\lambda_i\,(1 \le i \le m)$,$\boldsymbol{B}$ 的特征值为 $\mu_j\,(1 \le j \le n)$. 求证: 线性变换 φ 的特征值为 $\lambda_i\mu_j\,(1 \le i \le m; 1 \le j \le n)$.

注 本题是例题 6.1的推广.

证明 取 V 的一组基为 $m \times n$ 基础矩阵:

$$E_{11}, \cdots, E_{1n}, E_{21}, \cdots, E_{2n}, \cdots, E_{m1}, \cdots, E_{mn},$$

我们首先证明 φ 在这组基下的表示矩阵为 $A \otimes B'$. 事实上,

$$\varphi(E_{ij}) = AE_{ij}B = Ae_if_j'B = \sum_{k=1}^{m}\sum_{l=1}^{n} a_{ki}b_{jl}E_{kl},$$

其中 e_i, f_j 分别是 m, n 维标准单位列向量, 故 φ 的表示矩阵为

$$\begin{pmatrix} a_{11}B' & a_{12}B' & \cdots & a_{1m}B' \\ a_{21}B' & a_{22}B' & \cdots & a_{2m}B' \\ \vdots & \vdots & & \vdots \\ a_{m1}B' & a_{m2}B' & \cdots & a_{mm}B' \end{pmatrix} = A \otimes B'.$$

注意到 B' 与 B 有相同的特征值, 故由矩阵的 Kronecker 积的特征值可知, φ 的特征值为 $\lambda_i\mu_j$. □

例题 6.47 设 A, B 分别为 m, n 阶矩阵,V 为 $m\times n$ 矩阵全体构成的线性空间,V 上的线性变换 φ 定义为: $\varphi(\boldsymbol{X}) = \boldsymbol{AXB}$. 证明: φ 是线性自同构的充要条件是 A, B 都是可逆矩阵.

注 例 4.16 作为本题的特例, 我们已经给出了两种证法, 其中证法 1 仍然可以适用于本题, 证法 2 则需改用例 6.99 进行讨论, 当然也可用第 4 章解题 13 进行统一的处理, 请读者自行补充细节. 下面再给出两种证法.(这里的题目与题号都是指白皮书上的)

证明 证法三:由命题 6.55的证明过程可知,φ 在基础矩阵这组基下的表示矩阵为 $A\otimes B'$, 再由性质 (8) 可知 $|A\otimes B'| = |A|^n|B|^m$, 故 φ 是自同构当且仅当表示矩阵 $A \otimes B'$ 是可逆矩阵, 这也当且仅当 A, B 都是可逆矩阵.

证法四:由命题 6.55可知,φ 是自同构当且仅当 φ 所有的特征值 $\lambda_i\mu_j \ne 0$, 这当且仅当所有的 $\lambda_i \ne 0$ 以及所有的 $\mu_j \ne 0$, 这也当且仅当 A, B 都是可逆矩阵. □

例题 6.48 设 A, B 分别为 m, n 阶矩阵,V 为 $m\times n$ 矩阵全体构成的线性空间,V 上的线性变换 φ 定义为: $\varphi(\boldsymbol{X}) = \boldsymbol{AXB}$. 证明: φ 是幂零线性变换的充要条件是 A, B 至少有一个是幂零矩阵.

证明 先证充分性. 不妨设 A 是幂零矩阵, 即存在正整数 k, 使得 $A^k = O$, 则 $\varphi^k(X) = A^kXB^k = O$, 即 $\varphi^k = 0$, 于是 φ 是幂零线性变换.

再证必要性. 我们考虑必要性的逆否命题. 设 A, B 都不是幂零矩阵, 即对任意给定的正整数 k,$A^k \ne O$,$B^k \ne O$, 只要证明 $\varphi^k \ne 0$ 即可. 我们给出以下 4 种证法.

证法一: 不妨设 A^k 的第 i 列非零,B^k 的第 j 行非零, 即有列向量 $A^ke_i \ne \mathbf{0}$, 行向量 $f_j'B^k \ne \mathbf{0}$, 其中 e_i, f_j 分别是 m, n 维标准单位列向量, 于是

$$\varphi^k(E_{ij}) = A^kE_{ij}B^k = A^ke_if_j'B^k = (A^ke_i)(f_j'B^k) \ne O.$$

证法二:设 P_i, Q_i 为可逆矩阵, 使得 $P_1A^kQ_1 = \mathrm{diag}\{I_r, O\}$,$P_2B^kQ_2 = \mathrm{diag}\{I_s, O\}$, 不妨设 $r \ge s \ge 1$, 于是

$$\varphi^k(Q_1P_2) = P_1^{-1}\mathrm{diag}\{I_r, O\}\mathrm{diag}\{I_s, O\}Q_2^{-1} = P_1^{-1}\mathrm{diag}\{I_s, O\}Q_2^{-1} \ne O.$$

证法三: 由命题 6.55的证明过程可知,φ^k 在基础矩阵这组基下的表示矩阵为 $A^k \otimes (B^k)'$, 再由 Kronecker 积的定义可知 $A^k \otimes (B^k)' \ne O$, 于是 $\varphi^k \ne 0$.

证法四:由命题 6.8可知,φ 是幂零线性变换当且仅当 φ 的所有特征值都等于零. 由于 A, B 都不是幂零矩阵, 故 A 的特征值 λ_i 不全为零,B 的特征值 μ_j 不全为零. 再由命题 6.55可知,φ 的特征值 $\lambda_i\mu_j$ 也不全为零, 从而 φ 不是幂零线性变换. □

命题 6.57

设 A, B 分别为 m, n 阶矩阵,V 为 $m\times n$ 矩阵全体构成的线性空间,V 上的线性变换 φ 定义为:$\varphi(X) = \boldsymbol{AX}-\boldsymbol{XB}$. 设 A 的特征值为 $\lambda_i(1 \le i \le m)$,B 的特征值为 $\mu_j(1 \le j \le n)$. 求证: 线性变换 φ 的特征值为 $\lambda_i - \mu_j(1 \le i \le m; 1 \le j \le n)$.

证明 取 V 的一组基为 $m\times n$ 基础矩阵: $E_{11},\cdots,E_{1n},E_{21},\cdots,E_{2n},\cdots,E_{m1},\cdots,E_{mn}$, 类似命题 6.8的讨论可得,$\varphi$ 在上述基下的表示矩阵为 $\boldsymbol{A}\otimes\boldsymbol{I}_n - \boldsymbol{I}_m\otimes\boldsymbol{B}'$. 由命题 6.20可知, 存在 m 阶可逆矩阵 $\boldsymbol{P}$ 以及 n 阶可逆矩阵 $\boldsymbol{Q}$, 使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} \lambda_1 & * & * & * \\ & \lambda_2 & * & * \\ & & \ddots & \vdots \\ & & & \lambda_m \end{pmatrix}, \quad \boldsymbol{Q}^{-1}\boldsymbol{B}'\boldsymbol{Q} = \begin{pmatrix} \mu_1 & * & * & * \\ & \mu_2 & * & * \\ & & \ddots & \vdots \\ & & & \mu_n \end{pmatrix}.$$

注意到

$$(\boldsymbol{P}\otimes\boldsymbol{Q})^{-1}(\boldsymbol{A}\otimes\boldsymbol{I}_n - \boldsymbol{I}_m\otimes\boldsymbol{B}')(\boldsymbol{P}\otimes\boldsymbol{Q}) = (\boldsymbol{P}^{-1}\boldsymbol{AP})\otimes\boldsymbol{I}_n - \boldsymbol{I}_m\otimes(\boldsymbol{Q}^{-1}\boldsymbol{B}'\boldsymbol{Q})$$

是一个上三角矩阵, 其主对角元素依次为 $\lambda_1-\mu_1,\cdots,\lambda_1-\mu_n,\lambda_2-\mu_1,\cdots,\lambda_2-\mu_n,\cdots,\lambda_m-\mu_1,\cdots,\lambda_m-\mu_n$, 由此即得结论. □

例题 6.49 设 A, B 分别为 m, n 阶矩阵,V 为 $m\times n$ 矩阵全体构成的线性空间,V 上的线性变换 φ 定义为:$\varphi(\boldsymbol{X}) = \boldsymbol{AX}-\boldsymbol{XB}$. 证明: 若 A, B 都是幂零矩阵, 则 φ 是幂零线性变换.

证明 因为 A, B 都是幂零矩阵, 所以它们的特征值都为零. 由命题 6.56可知,φ 的特征值也都为零, 于是 φ 是幂零线性变换.(也可由矩阵的运算直接证明本题.) □

例题 6.50 设 $\boldsymbol{A}=(a_{ij})$ 是 n 阶矩阵,$g(\lambda) = |\lambda\boldsymbol{I}_n + \boldsymbol{A}|$. 求证:$n^2$ 阶矩阵

$$\boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{I}_n+\boldsymbol{A} & a_{12}\boldsymbol{I}_n & \cdots & a_{1n}\boldsymbol{I}_n \\ a_{21}\boldsymbol{I}_n & a_{22}\boldsymbol{I}_n+\boldsymbol{A} & \cdots & a_{2n}\boldsymbol{I}_n \\ \vdots & \vdots & & \vdots \\ a_{n1}\boldsymbol{I}_n & a_{n2}\boldsymbol{I}_n & \cdots & a_{nn}\boldsymbol{I}_n+\boldsymbol{A} \end{pmatrix}$$

是可逆矩阵的充要条件是 $g(\boldsymbol{A})$ 是可逆矩阵.

证明 显然 $\boldsymbol{B} = \boldsymbol{A}\otimes\boldsymbol{I}_n + \boldsymbol{I}_n\otimes\boldsymbol{A}$. 设 $\boldsymbol{A}$ 的全体特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则 $g(\lambda) = (\lambda+\lambda_1)(\lambda+\lambda_2)\cdots(\lambda+\lambda_n)$. 由命题 6.20可知, 存在 n 阶可逆矩阵 $\boldsymbol{P}$, 使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} \lambda_1 & * & * & * \\ & \lambda_2 & * & * \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix}.$$

注意到

$$(\boldsymbol{P}\otimes\boldsymbol{P})^{-1}\boldsymbol{B}(\boldsymbol{P}\otimes\boldsymbol{P}) = (\boldsymbol{P}^{-1}\boldsymbol{AP})\otimes\boldsymbol{I}_n + \boldsymbol{I}_n\otimes(\boldsymbol{P}^{-1}\boldsymbol{AP})$$

是一个上三角矩阵, 其主对角元素为 $\lambda_i+\lambda_j(1\le i,j\le n)$, 故

$$|\boldsymbol{B}| = \prod_{i,j=1}^{n}(\lambda_i+\lambda_j) = \prod_{i=1}^{n} g(\lambda_i).$$

因为 $g(\boldsymbol{A})$ 的特征值为 $g(\lambda_1),g(\lambda_2),\cdots,g(\lambda_n)$, 所以 $|\boldsymbol{B}| = |g(\boldsymbol{A})|$, 从而 $\boldsymbol{B}$ 是可逆矩阵等价于 $g(\boldsymbol{A})$ 是可逆矩阵. □

第七章 相似标准型

7.1 多项式矩阵

定义 7.1 (λ-矩阵)

一般地, 下列形式的矩阵:

$$\boldsymbol{A}(\lambda)=\begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \cdots & a_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \cdots & a_{mn}(\lambda) \end{pmatrix},$$

其中 $a_{ij}(\lambda)$ 是以 λ 为未定元的数域 $\mathbb{K}$ 上的多项式, 称为**多项式矩阵**, 或 **λ-矩阵**. λ-矩阵的加法、数乘及乘法与数域上的矩阵运算一样, 只需在运算过程中将数的运算代之于多项式即可.

定义 7.2 (λ-矩阵的初等变换)

对 λ-矩阵 $A(\lambda)$ 施行的下列 3 种变换称为 λ-矩阵的初等行变换:

(1) 将 $A(\lambda)$ 的两行对换;

(2) 将 $A(\lambda)$ 的第 i 行乘以 $\mathbb{K}$ 中的非零常数 c;

(3) 将 $A(\lambda)$ 的第 i 行乘以 $\mathbb{K}$ 上的多项式 $f(\lambda)$ 后加到第 j 行上去.

同理我们可以定义 3 种 λ-矩阵的初等列变换.

定义 7.3 (λ-矩阵的相抵)

若 $A(\lambda), B(\lambda)$ 是同阶 λ-矩阵且 $A(\lambda)$ 经过 λ-矩阵的初等变换后可变为 $B(\lambda)$, 则称 $A(\lambda)$ 与 $B(\lambda)$ **相抵**.

与数字矩阵一样, λ-矩阵的相抵关系也是一种等价关系, 即

(1) $A(\lambda)$ 与自身相抵;

(2) 若 $A(\lambda)$ 与 $B(\lambda)$ 相抵, 则 $B(\lambda)$ 与 $A(\lambda)$ 相抵;

(3) 若 $A(\lambda)$ 与 $B(\lambda)$ 相抵, $B(\lambda)$ 与 $C(\lambda)$ 相抵, 则 $A(\lambda)$ 与 $C(\lambda)$ 相抵.

笔记 λ-矩阵的相抵关系也是一种等价关系的证明与数域上相同, 类似易证.

定义 7.4 (初等 λ-矩阵)

下列 3 种矩阵称为初等 λ-矩阵:

(1) 将 n 阶单位阵的第 i 行与第 j 行对换, 记为 $\boldsymbol{P}_{ij}$;

(2) 将 n 阶单位阵的第 i 行乘以非零常数 c, 记为 $\boldsymbol{P}_i(c)$;

(3) 将 n 阶单位阵的第 i 行乘以多项式 $f(\lambda)$ 后加到第 j 行上去得到的矩阵, 记为 $\boldsymbol{T}_{ij}(f(\lambda))$.

注 第一类与第二类初等 λ-矩阵与数域上的第一类与第二类初等矩阵没有什么区别. 第三类初等 λ-矩阵的形状如

下:

$$T_{ij}(f(\lambda)) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & f(\lambda) & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

定理 7.1

对 λ-矩阵 $A(\lambda)$ 施行第 k $(k = 1, 2, 3)$ 类初等行 (列) 变换等于用第 k 类初等 λ-矩阵左 (右) 乘以 $A(\lambda)$.

注 下列 λ-矩阵的变换不是 λ-矩阵的初等变换:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} \lambda & \lambda \\ 0 & 1 \end{pmatrix}.$$

这是因为前面一个矩阵的第一行乘以 λ 不是 λ-矩阵的初等变换. 同理下面的变换需第一行乘以 λ^{-1}, 因此也不是 λ-矩阵的初等变换:

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

证明 证明是显然的. □

定义 7.5 (可逆 λ-矩阵)

若 $A(\lambda), B(\lambda)$ 都是 n 阶 λ-矩阵, 且

$$A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I_n,$$

则称 $B(\lambda)$ 是 $A(\lambda)$ 的逆 λ-矩阵. 这时称 $A(\lambda)$ 为**可逆 λ-矩阵**, 在不引起混淆的情形下, 有时简称为**可逆阵**.

笔记 容易证明, **有限个可逆 λ-矩阵之积仍是可逆 λ-矩阵**, 而**初等 λ-矩阵都是可逆 λ-矩阵**, 因此**有限个初等 λ-矩阵之积也是可逆的**. λ-矩阵.

注 注意不要将数字矩阵中的一些结论随意搬到 λ-矩阵上. 比如下面的 λ-矩阵的行列式不为零, 但它不是可逆 λ-矩阵:

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}.$$

这是因为矩阵

$$\begin{pmatrix} \lambda^{-1} & 0 \\ 0 & 1 \end{pmatrix}$$

不是 λ-矩阵之故.

引理 7.1

设 $M(\lambda)$ 是一个 n 阶 λ-矩阵, 则 $M(\lambda)$ 可以化为如下形状:

$$M(\lambda) = M_m\lambda^m + M_{m-1}\lambda^{m-1} + \cdots + M_0,$$

其中 M_i 为数域 $\mathbb{K}$ 上的 n 阶数字矩阵. 因此, 一个多项式矩阵可以化为系数为矩阵的多项式, 反之亦然.

证明 证明是显然的. □

引理 7.2

设 $\boldsymbol{M}(\lambda)$ 与 $\boldsymbol{N}(\lambda)$ 是两个 n 阶 λ-矩阵且都不等于零. 又设 $\boldsymbol{B}$ 为 n 阶数字矩阵, 则必存在 λ-矩阵 $\boldsymbol{Q}(\lambda)$ 及 $\boldsymbol{S}(\lambda)$ 和数字矩阵 $\boldsymbol{R}$ 及 $\boldsymbol{T}$, 使下式成立:

$$\boldsymbol{M}(\lambda) = (\lambda\boldsymbol{I} - \boldsymbol{B})\boldsymbol{Q}(\lambda) + \boldsymbol{R}, \tag{7.1}$$

$$\boldsymbol{N}(\lambda) = \boldsymbol{S}(\lambda)(\lambda\boldsymbol{I} - \boldsymbol{B}) + \boldsymbol{T}. \tag{7.2}$$

♡

证明 将 $\boldsymbol{M}(\lambda)$ 写为

$$\boldsymbol{M}(\lambda) = \boldsymbol{M}_m\lambda^m + \boldsymbol{M}_{m-1}\lambda^{m-1} + \cdots + \boldsymbol{M}_0,$$

其中 $\boldsymbol{M}_m \neq \boldsymbol{O}$. 可对 m 用归纳法, 若 $m = 0$, 则已适合要求 (取 $\boldsymbol{Q}(\lambda) = \boldsymbol{O}$). 现设对小于 m 次的矩阵多项式,(7.1)式成立. 令

$$\boldsymbol{Q}_1(\lambda) = \boldsymbol{M}_m\lambda^{m-1},$$

则

$$\boldsymbol{M}(\lambda) - (\lambda\boldsymbol{I} - \boldsymbol{B})\boldsymbol{Q}_1(\lambda) = (\boldsymbol{B}\boldsymbol{M}_m + \boldsymbol{M}_{m-1})\lambda^{m-1} + \cdots + \boldsymbol{M}_0. \tag{7.3}$$

上式是一个次数小于 m 的矩阵多项式, 由归纳假设得

$$\boldsymbol{M}(\lambda) - (\lambda\boldsymbol{I} - \boldsymbol{B})\boldsymbol{Q}_1(\lambda) = (\lambda\boldsymbol{I} - \boldsymbol{B})\boldsymbol{Q}_2(\lambda) + \boldsymbol{R}.$$

于是

$$\boldsymbol{M}(\lambda) = (\lambda\boldsymbol{I} - \boldsymbol{B})[\boldsymbol{Q}_1(\lambda) + \boldsymbol{Q}_2(\lambda)] + \boldsymbol{R}.$$

令 $\boldsymbol{Q}(\lambda) = \boldsymbol{Q}_1(\lambda) + \boldsymbol{Q}_2(\lambda)$ 即得(7.1)式. 同理可证 (7.2)式. □

定理 7.2

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

♡

证明 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则存在 $\mathbb{K}$ 上的非异阵 $\boldsymbol{P}$, 使 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 于是

$$\boldsymbol{P}^{-1}(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{P} = \lambda\boldsymbol{I} - \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \lambda\boldsymbol{I} - \boldsymbol{B}.$$

把 $\boldsymbol{P}$ 看成是常数 λ-矩阵, 上式表明 $\lambda\boldsymbol{I} - \boldsymbol{A}$ 与 $\lambda\boldsymbol{I} - \boldsymbol{B}$ 相抵.

反过来, 若 $\lambda\boldsymbol{I} - \boldsymbol{A}$ 与 $\lambda\boldsymbol{I} - \boldsymbol{B}$ 相抵, 即存在 $\boldsymbol{M}(\lambda)$ 及 $\boldsymbol{N}(\lambda)$, 使

$$\boldsymbol{M}(\lambda)(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{N}(\lambda) = \lambda\boldsymbol{I} - \boldsymbol{B}, \tag{7.4}$$

其中 $\boldsymbol{M}(\lambda)$ 与 $\boldsymbol{N}(\lambda)$ 都是有限个初等矩阵之积, 因而都是可逆阵. 因此可将 (7.4)式写为

$$\boldsymbol{M}(\lambda)(\lambda\boldsymbol{I} - \boldsymbol{A}) = (\lambda\boldsymbol{I} - \boldsymbol{B})\boldsymbol{N}(\lambda)^{-1}, \tag{7.5}$$

由引理 7.2可设

$$\boldsymbol{M}(\lambda) = (\lambda\boldsymbol{I} - \boldsymbol{B})\boldsymbol{Q}(\lambda) + \boldsymbol{R},$$

代入 (7.5)式经整理得

$$\boldsymbol{R}(\lambda\boldsymbol{I} - \boldsymbol{A}) = (\lambda\boldsymbol{I} - \boldsymbol{B})[\boldsymbol{N}(\lambda)^{-1} - \boldsymbol{Q}(\lambda)(\lambda\boldsymbol{I} - \boldsymbol{A})].$$

上式左边是次数小于等于 1 的矩阵多项式, 因此上式右边中括号内的矩阵多项式的次数必须小于等于零, 也即必是一个常数矩阵, 设为 $\boldsymbol{P}$. 于是

$$\boldsymbol{R}(\lambda\boldsymbol{I} - \boldsymbol{A}) = (\lambda\boldsymbol{I} - \boldsymbol{B})\boldsymbol{P}. \tag{7.6}$$

(7.6)式又可整理为

$$(\boldsymbol{R} - \boldsymbol{P})\lambda = \boldsymbol{R}\boldsymbol{A} - \boldsymbol{B}\boldsymbol{P}.$$

再次比较次数得 $\boldsymbol{R}=\boldsymbol{P},\boldsymbol{RA}=\boldsymbol{BP}$. 现只需证明 $\boldsymbol{P}$ 是一个非异阵即可. 由假设

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

将上式两边右乘 $\boldsymbol{N}(\lambda)$ 并移项得

$$\boldsymbol{PN}(\lambda)+\boldsymbol{Q}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{N}(\lambda)=\boldsymbol{I}.$$

但由(7.4)式可得

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

因此

$$\boldsymbol{PN}(\lambda)+\boldsymbol{Q}(\lambda)\boldsymbol{M}(\lambda)^{-1}(\lambda\boldsymbol{I}-\boldsymbol{B})=\boldsymbol{I}. \tag{7.7}$$

再由引理 7.2可设

$$\boldsymbol{N}(\lambda)=\boldsymbol{S}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{B})+\boldsymbol{T},$$

代入(7.7)式并整理得

$$[\boldsymbol{PS}(\lambda)+\boldsymbol{Q}(\lambda)\boldsymbol{M}(\lambda)^{-1}](\lambda\boldsymbol{I}-\boldsymbol{B})=\boldsymbol{I}-\boldsymbol{PT}.$$

上式右边是次数小于等于零的矩阵多项式, 因此上式左边中括号内的矩阵多项式必须为零, 从而 $\boldsymbol{PT}=\boldsymbol{I}$, 即 $\boldsymbol{P}$ 是非异阵. □

7.2 矩阵的法式

引理 7.3

设 $A(\lambda)=(a_{ij}(\lambda))_{m\times n}$ 是任一非零 λ-矩阵, 则 $A(\lambda)$ 必相抵于这样的一个 λ-矩阵 $B(\lambda)=(b_{ij}(\lambda))_{m\times n}$, 其中 $b_{11}(\lambda)\neq 0$ 且 $b_{11}(\lambda)$ 可整除 $B(\lambda)$ 中的任一元素 $b_{ij}(\lambda)$.

证明 设 $k=\min\{\deg a_{ij}(\lambda)\mid a_{ij}(\lambda)\neq 0,1\leq i\leq m;1\leq j\leq n\}$, 我们对 k 用数学归纳法. 首先, 经行对换及列对换可将 $A(\lambda)$ 的第 $(1,1)$ 元素变成次数最低的非零多项式, 因此不妨设 $a_{11}(\lambda)\neq 0$ 且 $\deg a_{11}(\lambda)=k$. 若 $k=0$, 则 $a_{11}(\lambda)$ 是一个非零常数, 结论显然成立. 假设对非零元素次数的最小值小于 k 的任一 λ-矩阵, 引理的结论成立, 现考虑非零元素次数的最小值等于 k 的 λ-矩阵 $A(\lambda)$. 若 $a_{11}(\lambda)$ 可整除所有的 $a_{ij}(\lambda)$, 则结论已成立. 若否, 设在第一列中有元素 $a_{i1}(\lambda)$ 不能被 $a_{11}(\lambda)$ 整除, 作带余除法:

$$a_{i1}(\lambda)=a_{11}(\lambda)q(\lambda)+r(\lambda).$$

用 $-q(\lambda)$ 乘以第一行加到第 i 行上, 第 $(i,1)$ 元素就变为 $r(\lambda)$. 注意到 $r(\lambda)\neq 0$ 且 $\deg r(\lambda)<\deg a_{11}(\lambda)=k$, 由归纳假设即知结论成立.

同样的方法可施于第一行. 因此我们不妨设 $a_{11}(\lambda)$ 可整除第一行及第一列. 这时, 设 $a_{21}(\lambda)=a_{11}(\lambda)g(\lambda)$. 将第一行乘以 $-g(\lambda)$ 加到第二行上, 则第 $(2,1)$ 元素变为零. 用同样的方法可消去第一行、第一列除 $a_{11}(\lambda)$ 以外的所有元素, 于是 $A(\lambda)$ 经初等变换后变成下列形状:

$$\begin{pmatrix} a_{11}(\lambda) & 0 & \cdots & 0\\ 0 & a'_{22}(\lambda) & \cdots & a'_{2n}(\lambda)\\ \vdots & \vdots & \ddots & \vdots\\ 0 & a'_{m2}(\lambda) & \cdots & a'_{mn}(\lambda)\end{pmatrix}.$$

这时, 若 $a_{11}(\lambda)$ 可整除所有其他元素, 则结论已成立. 若否, 比如 $a_{11}(\lambda)$ 不能整除 $a'_{ij}(\lambda)$, 则将第 i 行加到第一行上去, 这时在第一行又出现了一元素 $a'_{ij}(\lambda)$, 它不能被 $a_{11}(\lambda)$ 整除. 重复上面的做法, 通过归纳假设即可得到结论. □

定理 7.3

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

$$\operatorname{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda), 0, \cdots, 0\}, \tag{7.8}$$

其中 $d_i(\lambda)$ 是非零首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ $(i = 1, 2, \cdots, r-1)$. 我们称上式中的对角 λ-矩阵为 $A(\lambda)$ 的**法式**或**相抵标准型**或 **Smith 标准型**.

证明 对 n 用数学归纳法, 当 $n=1$ 时结论显然, 现设 $A(\lambda)$ 是 n 阶 λ-矩阵. 由引理 7.3可知 $A(\lambda)$ 相抵于 n 阶 λ-矩阵 $B(\lambda) = (b_{ij}(\lambda))$, 其中 $b_{11}(\lambda) \mid b_{ij}(\lambda)$ 对一切 i, j 成立. 因此, 将 $B(\lambda)$ 的第一行乘以 λ 的某个多项式加到第二行上去便可消去 $b_{21}(\lambda)$. 同理可依次消去第一列除 $b_{11}(\lambda)$ 以外的所有元素. 再用类似方法消去第一行其余元素. 这样便得到了一个矩阵:

$$\begin{pmatrix} b_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & b'_{22}(\lambda) & \cdots & b'_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b'_{n2}(\lambda) & \cdots & b'_{nn}(\lambda) \end{pmatrix}.$$

不难看出, 这时 $b_{11}(\lambda)$ 仍可整除所有的 $b'_{ij}(\lambda)$. 设 c 为 $b_{11}(\lambda)$ 的首项系数, 记 $d_1(\lambda) = c^{-1}b_{11}(\lambda)$, 设 $\overline{B}(\lambda)$ 为上面的矩阵中右下方的 $n-1$ 阶 λ-矩阵, 则由归纳假设可知存在 $P(\lambda)$ 及 $Q(\lambda)$, 使

$$P(\lambda)\overline{B}(\lambda)Q(\lambda) = \operatorname{diag}\{d_2(\lambda), \cdots, d_r(\lambda); 0, \cdots, 0\},$$

且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ $(i = 2, \cdots, r-1)$, 其中 $P(\lambda)$ 与 $Q(\lambda)$ 可写成为有限个 $n-1$ 阶初等 λ-矩阵之积. 因此

$$\begin{pmatrix} 1 & O \\ O & P(\lambda) \end{pmatrix} \begin{pmatrix} d_1(\lambda) & O \\ O & \overline{B}(\lambda) \end{pmatrix} \begin{pmatrix} 1 & O \\ O & Q(\lambda) \end{pmatrix} = \operatorname{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda); 0, \cdots, 0\},$$

且

$$\begin{pmatrix} 1 & O \\ O & P(\lambda) \end{pmatrix}, \begin{pmatrix} 1 & O \\ O & Q(\lambda) \end{pmatrix}$$

可写成有限个 n 阶初等 λ-矩阵之积. 于是只需证明 $d_1(\lambda) \mid d_2(\lambda)$ 即可. 但这点很容易看出, 事实上由于 $\overline{B}(\lambda)$ 中的任一元素均可被 $d_1(\lambda)$ 整除, 因此 $P(\lambda)\overline{B}(\lambda)Q(\lambda)$ 中的任一元素也可被 $d_1(\lambda)$ 整除, 这就证明了定理. □

注 我们上面对 n 阶 λ-矩阵证明了它必相抵于一个对角阵. 事实上, 对长方 λ-矩阵, 结论也同样成立, 证明也类似.(7.8)式中的 r 通常称为 $A(\lambda)$ 的秩. 但要注意即使某个 n 阶 λ-矩阵的秩等于 n, 它也未必是可逆 λ-矩阵.

推论 7.1

任一 n 阶可逆 λ-矩阵都可表示为有限个初等 λ-矩阵之积.

证明 由定理 7.3, 存在 $P(\lambda), Q(\lambda)$, 使可逆阵 $A(\lambda)$ 适合

$$P(\lambda)A(\lambda)Q(\lambda) = \operatorname{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda); 0, \cdots, 0\},$$

其中 $P(\lambda), Q(\lambda)$ 为有限个初等 λ-矩阵之积. 因为上式左边是个可逆阵, 故右边的矩阵也可逆, 从而 $r = n$. 注意一个对角 λ-矩阵要可逆必须 $d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)$ 皆为非零常数, 又它们都是首一多项式, 故只能是 $d_1(\lambda) = d_2(\lambda) = \cdots = d_n(\lambda) = 1$, 于是

$$A(\lambda) = P(\lambda)^{-1}Q(\lambda)^{-1}.$$

因为初等 λ-矩阵的逆仍是初等 λ-矩阵, 故 $P(\lambda)^{-1}$ 与 $Q(\lambda)^{-1}$ 都是有限个初等 λ-矩阵之积, 从而 $A(\lambda)$ 也是有限个初等 λ-矩阵之积. □

推论 7.2

设 A 是数域 $\mathbb{K}$ 上的 n 阶矩阵, 则 A 的特征矩阵 $\lambda I_n - A$ 必相抵于

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

其中 $d_i(\lambda) \mid d_{i+1}(\lambda)$ $(i = 1, 2, \cdots, m-1)$. 我们称上式中的对角 λ-矩阵为 $A(\lambda)$ 的特征矩阵 $\lambda I_n - A$ 的**法式**或**相抵标准型**.

证明 由定理 7.3, 存在 $P(\lambda), Q(\lambda)$, 使

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

其中 $P(\lambda), Q(\lambda)$ 为有限个初等 λ-矩阵之积. 根据 λ-矩阵初等变换的定义以及行列式的性质可得, 上式左边的行列式等于 $c|\lambda I_n - A|$, 其中 c 是一个非零常数, 从而上式右边的行列式不为零, 故 $r = n$. 把 $d_i(\lambda)$ 中的常数多项式写出来（因是首一多项式, 故为常数 1）, 即得结论. □

例题 7.1 求 $\lambda I - A$ 的法式, 其中

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

解

$$\lambda I - A = \begin{pmatrix} \lambda & -1 & 1 \\ -3 & \lambda+2 & 0 \\ 1 & -1 & \lambda+1 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{pmatrix} 1 & -1 & \lambda+1 \\ -3 & \lambda+2 & 0 \\ \lambda & -1 & 1 \end{pmatrix}$$

$$\xrightarrow{3r_1+r_2, -\lambda r_1+r_3} \begin{pmatrix} 1 & -1 & \lambda+1 \\ 0 & \lambda-1 & 3\lambda+3 \\ 0 & \lambda-1 & -\lambda^2-\lambda+1 \end{pmatrix} \xrightarrow{j_1+j_2, -(\lambda+1)j_1+j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-1 & 3\lambda+3 \\ 0 & \lambda-1 & -\lambda^2-\lambda+1 \end{pmatrix}$$

$$\xrightarrow{-3j_2+j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-1 & 6 \\ 0 & \lambda-1 & -\lambda^2-4\lambda+4 \end{pmatrix} \xrightarrow{j_2 \leftrightarrow j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & \lambda-1 \\ 0 & -\lambda^2-4\lambda+4 & \lambda-1 \end{pmatrix}$$

$$\xrightarrow{6j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 6(\lambda-1) \\ 0 & -\lambda^2-4\lambda+4 & 6(\lambda-1) \end{pmatrix} \xrightarrow{-(\lambda-1)j_2+j_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & -\lambda^2-4\lambda+4 & (\lambda-1)(\lambda^2+4\lambda+2) \end{pmatrix}$$

$$\xrightarrow{\frac{1}{6}j_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\lambda^2-4\lambda+4 & (\lambda-1)(\lambda^2+4\lambda+2) \end{pmatrix} \xrightarrow{-(-\lambda^2-4\lambda+4)r_2+r_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda-1)(\lambda^2+4\lambda+2) \end{pmatrix}.$$

□

7.3 不变因子

定义 7.6 (k 阶行列式因子)

设 $A(\lambda)$ 是 n 阶 λ-矩阵,k 是小于等于 n 的正整数. 如果 $A(\lambda)$ 有一个 k 阶子式不为零, 则定义 $A(\lambda)$ 的 k **阶行列式因子** $D_k(\lambda)$ 为 $A(\lambda)$ 的所有 k 阶子式的最大公因式（首一多项式）. 如果 $A(\lambda)$ 的所有 k 阶子式都等于零, 则定义 $A(\lambda)$ 的 k 阶行列式因子 $D_k(\lambda)$ 为零.

引理 7.4

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

$$D_i(\lambda) \mid D_{i+1}(\lambda), \quad i = 1, 2, \cdots, r-1.$$

证明 设 A_{i+1} 是 $A(\lambda)$ 的任一 $i+1$ 阶子式, 即在 $A(\lambda)$ 中任意取出 $i+1$ 行及 $i+1$ 列组成的行列式. 将这个行列式按某一行展开, 则它的每一个展开项都是一个多项式与一个 i 阶子式的乘积. 由于 $D_i(\lambda)$ 是所有 i 阶子式的公因子, 因此 $D_i(\lambda) \mid A_{i+1}$. 而 $D_{i+1}(\lambda)$ 是所有 $i+1$ 阶子式的最大公因子, 因此 $D_i(\lambda) \mid D_{i+1}(\lambda)$ 对一切 $i = 1, 2, \cdots, r-1$ 成立. □

定义 7.7 (不变因子)

设 $D_1(\lambda), D_2(\lambda), \cdots, D_r(\lambda)$ 是 λ-矩阵 $A(\lambda)$ 的非零行列式因子, 则

$$\begin{aligned} g_1(\lambda) &= D_1(\lambda), \\ g_2(\lambda) &= D_2(\lambda)/D_1(\lambda), \\ &\cdots \\ g_r(\lambda) &= D_r(\lambda)/D_{r-1}(\lambda) \end{aligned}$$

称为 $A(\lambda)$ 的**不变因子**.

笔记 由不变因子和行列式因子的定义可知, 不变因子和行列式因子相互唯一确定.

注 以后特征矩阵 $\lambda I - A$ 的行列式因子和不变因子均简称为 A 的行列式因子和不变因子.

命题 7.1

求下列矩阵的行列式因子和不变因子:

$$A(\lambda) = \begin{pmatrix} d_1(\lambda) & & & & & \\ & \ddots & & & & \\ & & d_r(\lambda) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

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

解 $A(\lambda)$ 的非零行列式因子为

$$\begin{aligned} D_1(\lambda) &= d_1(\lambda), \\ D_2(\lambda) &= d_1(\lambda)d_2(\lambda), \\ &\cdots \\ D_r(\lambda) &= d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda). \end{aligned}$$

根据不变因子的定义可知 $A(\lambda)$ 的不变因子分别为:$d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)$. □

定理 7.4

相抵的 λ-矩阵有相同的行列式因子, 从而有相同的不变因子.

证明 我们只需证明行列式因子在三类初等变换下不改变就可以了. 对第一类初等变换, 交换 λ-矩阵 $A(\lambda)$ 的任意两行（列）, 显然 $A(\lambda)$ 的 i 阶子式最多改变一个符号, 因此行列式因子不改变.

对第二类初等变换,$A(\lambda)$ 的 i 阶子式与变换后矩阵的 i 阶子式最多差一个非零常数, 因此行列式因子也不改

变.

对第三类初等变换, 记变换后的矩阵为 $B(\lambda)$, 则 $B(\lambda)$ 与 $A(\lambda)$ 的 i 阶子式可能出现以下 3 种情形: 子式完全相同;$B(\lambda)$ 子式中的某一行（列）等于 $A(\lambda)$ 中相应子式的同一行（列）加上该子式中某一行（列）与某个多项式之积;$B(\lambda)$ 子式中的某一行（列）等于 $A(\lambda)$ 中相应子式的同一行（列）加上不在该子式中的某一行（列）与某个多项式之积. 在前面两种情形, 行列式的值不改变, 因此不影响行列式因子. 现在来讨论第三种情形. 设 B_i 为 $B(\lambda)$ 的 i 阶子式, 相应的 $A(\lambda)$ 的 i 阶子式记为 A_i, 则由行列式的性质得

$$B_i = A_i + f(\lambda)\widetilde{A}_i,$$

其中 $\widetilde{A}_i$ 由 $A(\lambda)$ 中的 i 行与 i 列组成, 因此它与 $A(\lambda)$ 的某个 i 阶子式最多差一个符号.$f(\lambda)$ 是乘以某一行（列）的那个多项式, 于是 $A(\lambda)$ 的行列式因子 $D_i(\lambda) \mid A_i$,$D_i(\lambda) \mid \widetilde{A}_i$, 故 $D_i(\lambda) \mid B_i$. 这说明,$D_i(\lambda)$ 可整除 $B(\lambda)$ 的所有 i 阶子式, 因此 $D_i(\lambda)$ 可整除 $B(\lambda)$ 的 i 阶行列式因子 $\widetilde{D}_i(\lambda)$. 但 $B(\lambda)$ 也可用第三类初等变换变成 $A(\lambda)$, 于是 $\widetilde{D}_i(\lambda) \mid D_i(\lambda)$. 由于 $D_i(\lambda)$ 及 $\widetilde{D}_i(\lambda)$ 都是首一多项式, 因此必有 $D_i(\lambda) = \widetilde{D}_i(\lambda)$. □

推论 7.3

设 n 阶 λ-矩阵 $A(\lambda)$ 的法式为

$$\Lambda = \operatorname{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda); 0, \cdots, 0\},$$

其中 $d_i(\lambda)$ 是非零首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ $(i = 1, 2, \cdots, r-1)$, 则 $A(\lambda)$ 的不变因子为 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 特别地, **法式和不变因子之间相互唯一确定**.

证明 首先, 由定理 7.4可知,$A(\lambda)$ 与 Λ 有相同的不变因子. 再由命题 7.1可知,Λ 的不变因子为 $d_1(\lambda)$, $d_2(\lambda)$, $\cdots$, $d_r(\lambda)$, 从而它们也是 $A(\lambda)$ 的不变因子. 故 $A(\lambda)$ 的法式可以唯一确定其不变因子.

接着, 设 $A(\lambda)$ 的不变因子为 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$, 由定理 7.3, 可设 $A(\lambda)$ 相抵于对角阵

$$B(\lambda) = \operatorname{diag}\{d_1'(\lambda), d_2'(\lambda), \cdots, d_r'(\lambda); 0, \cdots, 0\}, \tag{7.9}$$

其中 $d_i'(\lambda)$ 是非零首一多项式且 $d_i'(\lambda) \mid d_{i+1}'(\lambda)$ $(i = 1, 2, \cdots, r-1)$. 再由命题 7.1可知,$B(\lambda)$ 的不变因子为 $d_1'(\lambda)$, $d_2'(\lambda), \cdots, d_r'(\lambda)$. 由定理 7.4可知,$A(\lambda)$ 和 $B(\lambda)$ 的不变因子相同, 故根据 $d_i(\lambda), d_i'(\lambda)$ 的整除关系, 我们就有

$$\begin{aligned} d_1(\lambda) &= d_1'(\lambda), \\ d_2(\lambda) &= d_2'(\lambda), \\ &\cdots\cdots\cdots\cdots \\ d_r(\lambda) &= d_r'(\lambda). \end{aligned}$$

因此 $A(\lambda)$ 的相抵于对角阵

$$\Lambda = \operatorname{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda); 0, \cdots, 0\},$$

其中 $d_i(\lambda)$ 是非零首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ $(i = 1, 2, \cdots, r-1)$. 上式也就是 $A(\lambda)$ 的法式. 故 $A(\lambda)$ 的不变因子可以唯一确定其法式. □

推论 7.4

设 $A(\lambda), B(\lambda)$ 为 n 阶 λ-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 $A(\lambda)$ 与 $B(\lambda)$ 有相同的法式, 显然它们相抵. 若 $A(\lambda)$ 与 $B(\lambda)$ 相抵, 由定理 7.4 知 $A(\lambda)$ 与 $B(\lambda)$ 有相同的不变因子, 从而由推论 7.3可知, $A(\lambda)$ 与 $B(\lambda)$ 有相同的法式. □

推论 7.5

n 阶 λ-矩阵 $A(\lambda)$ 的法式与初等变换的选取无关.

证明 设 Λ_1, Λ_2 是 $A(\lambda)$ 通过不同的初等变换得到的两个法式, 则 Λ_1 与 Λ_2 相抵, 由推论 7.4可得 $\Lambda_1 = \Lambda_2$. □

定理 7.5

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 $\lambda I-A$ 与 $\lambda I-B$ 具有相同的行列式因子或不变因子.

证明 显然不变因子与行列式因子之间相互唯一确定. 再由定理 7.2、推论 7.4 及推论 7.3即得结论. □

推论 7.6

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它们在 $\mathbb{K}$ 上相似.

笔记 这个推论告诉我们: **矩阵的相似关系在基域扩张下不变**. 事实上, 这个推论的证明过程也说明: **矩阵的不变因子在基域扩张下也不变**. 此即即**矩阵的相似关系与数域无关.**

证明 若 A 与 B 在 $\mathbb{F}$ 上相似, 由于 $\mathbb{F}\subseteq\mathbb{K}$, 它们当然在 $\mathbb{K}$ 上也相似. 反之, 若 A 与 B 在 $\mathbb{K}$ 上相似, 则 $\lambda I-A$ 与 $\lambda I-B$ 在 $\mathbb{K}$ 上有相同的不变因子, 也就是说它们有相同的法式. 由推论 7.5可知, 求法式与初等变换的选取无关. 注意到 $\lambda I-A$ 与 $\lambda I-B$ 是数域 $\mathbb{F}$ 上的 λ-矩阵, 故可用 $\mathbb{F}$ 上 λ-矩阵的初等变换就能将它们变成法式, 其中只涉及 $\mathbb{F}$ 中数的加、减、乘、除运算以及 $\mathbb{F}$ 上的多项式的加、减、乘、数乘运算, 最后得到法式中的不变因子 $d_i(\lambda)$ 仍是 $\mathbb{F}$ 上的多项式. 这就是说存在 $\mathbb{F}$ 上的可逆 λ-矩阵 $P(\lambda),Q(\lambda),M(\lambda),N(\lambda)$, 使

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

从而

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

即 $\lambda I-A$ 与 $\lambda I-B$ 在 $\mathbb{F}$ 上相抵, 由定理 7.2 可得 A 与 B 在 $\mathbb{F}$ 上相似. □

7.4 有理标准型

命题 7.2

设矩阵 A 的特征矩阵 $\lambda I-A$ 的法式为

$$\operatorname{diag}\{1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)\},$$

其中 $d_i(\lambda)$ 为非常数首一多项式且 $d_i(\lambda)\mid d_{i+1}(\lambda)\ (i=1,2,\cdots,k-1)$, 则 A 的不变因子就是

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

证明 由推论 7.2可知, 矩阵 A 的特征矩阵 $\lambda I-A$ 的法式为

$$\operatorname{diag}\{1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)\},$$

其中 $d_i(\lambda)$ 为非常数首一多项式且 $d_i(\lambda)\mid d_{i+1}(\lambda)\ (i=1,2,\cdots,k-1)$, 则根据不变因子的定义可知,$A$ 的不变因子就是

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

□

引理 7.5 (Frobenius 块的基本性质)

设 r 阶矩阵

$$F=F(f(\lambda))=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&1\\-a_r&-a_{r-1}&-a_{r-2}&\cdots&-a_1\end{pmatrix},$$

其中 $f(\lambda)=\lambda^r+a_1\lambda^{r-1}+\cdots+a_r$, 则

(1) $|F(f(\lambda))|=(-1)^{r+2}a_r=(-1)^{r+2}f(0)$.

(2) $F=F(f(\lambda))$ 的特征多项式 $|\lambda I-F|$ 为 $f(\lambda)$.

(3) F 的行列式因子为

$$1,\cdots,1,f(\lambda), \tag{7.10}$$

其中共有 $r-1$ 个 $1,f(\lambda)=\lambda^r+a_1\lambda^{r-1}+\cdots+a_r,F$ 的不变因子也由(7.10)式给出,F 的不变因子分别为:

$$1,\cdots,1,f(\lambda).$$

进而,$\lambda I-F$ 相抵于 $\operatorname{diag}\{1,\cdots,1,f(\lambda)\}$.

(4) F 的极小多项式等于 $f(\lambda)$.

证明

(1) 注意到 $f(0)=a_r$, 于是就有

$$|F(f(\lambda))|=\begin{vmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&1\\-a_r&-a_{r-1}&-a_{r-2}&\cdots&-a_1\end{vmatrix}\xlongequal{\text{按第一列展开}}(-1)^{r+2}a_r=(-1)^{r+2}f(0).$$

(2) 由命题 6.6(1)同理可得得

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

(3) F 的 r 阶行列式因子就是它的特征多项式, 由命题 6.6(1)同理可得得

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

对任一 $k<r,\lambda I-F$ 总有一个 k 阶子式其值等于 $(-1)^k$, 故 $D_k(\lambda)=1$. 又由推论 7.3可知,$\lambda I-F$ 的法式为 $\operatorname{diag}\{1,\cdots,1,f(\lambda)\}$. 故 $\lambda I-F$ 相抵于 $\operatorname{diag}\{1,\cdots,1,f(\lambda)\}$.

(4) 因为 F 的特征多项式为 $f(\lambda)$, 所以 F 适合多项式 $f(\lambda)$. 设 $e_i\ (i=1,2,\cdots,r)$ 是 r 维标准单位行向量, 则不难算出:

$$e_1F=e_2,\quad e_1F^2=e_2F=e_3,\quad\cdots,\quad e_1F^{r-1}=e_{r-1}F=e_r.$$

显然,$e_1,e_1F,\cdots,e_1F^{r-1}$ 是一组线性无关的向量, 从而任取 $g(x)\in P_{r-1}[x]$ 且 $g(x)$ 非零, 则存在一组不全为零的数 $a_1,a_2,\cdots,a_r$, 使得

$$g(x)=a_1x^{r-1}+a_2x^{r-2}+\cdots+a_r.$$

于是将 F 代入上式, 再在等式两边同乘 e_1 得到

$$e_1g(F)=a_1e_1F^{r-1}+a_2e_1F^{r-2}+\cdots+a_re_1F.$$

又因为 $e_1,e_1F,\cdots,e_1F^{r-1}$ 是一组线性无关的向量, 且 $a_1,a_2,\cdots,a_r$ 不全为零, 所以 $e_1g(F)\neq0$. 即 $g(F)$ 的第

一行不为零, 故 $g(F) \neq O$. 因此 F 不可能适合一个次数不超过 $r-1$ 的非零多项式, 从而 F 的极小多项式就是 $f(\lambda)$.

□

引理 7.6

设 λ-矩阵 $A(\lambda)$ 相抵于对角 λ-矩阵

$$\mathrm{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)\}, \tag{7.11}$$

λ-矩阵 $B(\lambda)$ 相抵于对角 λ-矩阵

$$\mathrm{diag}\{d_1'(\lambda), d_2'(\lambda), \cdots, d_n'(\lambda)\}, \tag{7.12}$$

且 $d_1'(\lambda), d_2'(\lambda), \cdots, d_n'(\lambda)$ 是 $d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)$ 的一个置换（即若不计次序, 这两组多项式完全相同）, 则 $A(\lambda)$ 相抵于 $B(\lambda)$.

证明 利用初等行对换及初等列对换即可将(7.11)式变成(7.12)式, 因此(7.11) 式所示的矩阵与 (7.12)式所示的矩阵相抵, 从而 $A(\lambda)$ 与 $B(\lambda)$ 相抵. □

定理 7.6 (有理标准型/Frobenius 标准型)

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 $\deg d_i(\lambda) = m_i \geq 1$, 则 A 相似于下列分块对角阵:

$$F = \begin{pmatrix} F_1 & & & \\ & F_2 & & \\ & & \ddots & \\ & & & F_k \end{pmatrix}, \tag{7.13}$$

其中 F_i 的阶等于 m_i, 且 F_i 是形如引理 7.5 中的矩阵,F_i 的最后一行由 $d_i(\lambda)$ 的系数（除首项系数之外）的负值组成. 此即, 设 $d_i = \lambda^{m_i} + a_{1i}\lambda^{m_i-1} + \cdots + a_{m_i,i}$, 则

$$F_i = F(d_i(\lambda)) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_{m_i,i} & -a_{m_i-1,i} & -a_{m_i-2,i} & \cdots & -a_{1i} \end{pmatrix}.$$

(7.13)式称为矩阵 A 的**有理标准型**或 **Frobenius 标准型**, 每个 F_i 称为 **Frobenius 块**.

进而,A 也相似于下列分块对角阵:

$$C = \begin{pmatrix} C_1 & & & \\ & C_2 & & \\ & & \ddots & \\ & & & C_k \end{pmatrix},$$

其中 C_i 的阶等于 m_i,C_i 就是上述 F_i 的转置, 即

$$C_i = C(d_i(\lambda)) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_{m_i,i} \\ 1 & 0 & \cdots & 0 & -a_{m_i-1,i} \\ 0 & 1 & \cdots & 0 & -a_{m_i-2,i} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{1i} \end{pmatrix}.$$

证明 注意到 $\lambda I - A$ 的第 n 个行列式因子就是 A 的特征多项式 $|\lambda I - A|$, 再由不变因子的定义可知:

$$|\lambda I - A| = d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda).$$

而 $|\lambda I - A|$ 是一个 n 次多项式, 因此 $m_1 + m_2 + \cdots + m_k = n$. 一方面,$\lambda I - A$ 的法式为

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

其中有 $n-k$ 个 1. 另一方面, 对 $\lambda I - F$ 的每个分块都施以 λ-矩阵的初等变换, 由引理 7.5可知,$\lambda I - F$ 相抵于如下对角阵:

$$\operatorname{diag}\{1, \cdots, 1, d_1(\lambda); 1, \cdots, 1, d_2(\lambda); \cdots; 1, \cdots, 1, d_k(\lambda)\}, \tag{7.14}$$

其中每个 $d_i(\lambda)$ 前各有 $m_i - 1$ 个 1, 从而共有 $\sum_{i=1}^{k}(m_i - 1) = n - k$ 个 1. 因此 (7.14) 式所示的矩阵与 $\lambda I - A$ 的法式只相差主对角线上元素的置换, 由引理 7.6可得 $\lambda I - A$ 与 $\lambda I - F$ 相抵, 从而 A 与 F 相似.

又因为矩阵与其自身的转置相似, 所以矩阵 A 也相似于 F 的转置, 即 C. □

例题 7.2 设 6 阶矩阵 A 的不变因子为

$$1, 1, 1, \lambda - 1, (\lambda - 1)^2, (\lambda - 1)^2(\lambda + 1),$$

则 A 的有理标准型为

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

定理 7.7

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 的特征多项式为 $d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda)$, 极小多项式为 $m(\lambda) = d_k(\lambda)$. ♡

证明 首先证明特征多项式, 根据不变因子和行列式因子的定义可知

$$1 \cdot 1 \cdot d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda) = D_n(\lambda).$$

其中 $D_n(\lambda)$ 为 $\lambda I_n - A$ 的 n 阶行列式因子, 即 A 的特征多项式 $|\lambda I_n - A|$. 因此

$$|\lambda I_n - A| = d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda).$$

然后证明极小多项式, 设 A 的有理标准型为

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

因为相似矩阵有相同的极小多项式, 故只需证明 F 的极小多项式是 $d_k(\lambda)$ 即可. 但 F 是分块对角阵, 由极小多项式的性质 (6)知 F 的极小多项式是诸 F_i 极小多项式的最小公倍式. 又由引理 7.5 知 F_i 的极小多项式为 $d_i(\lambda)$. 因为

$d_i(\lambda) \mid d_{i+1}(\lambda)$, 故诸 $d_i(\lambda)$ 的最小公倍式等于 $d_k(\lambda)$. □

例题 7.3 下面两个 4 阶矩阵

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

的不变因子分别为 A:$1, \lambda, \lambda, \lambda^2$ 和 B:$1, 1, \lambda^2, \lambda^2$. 它们的特征多项式和极小多项式分别相等, 但它们不相似.

定义 7.8 (循环子空间)

设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,φ 是 V 上的线性变换. 设 $\mathbf{0} \neq \alpha \in V$, 则 $U = L(\alpha, \varphi(\alpha), \varphi^2(\alpha), \cdots)$ 称为 V 的**循环子空间**, 记为 $U = C(\varphi, \alpha)$,α 称为 U 的**循环向量**. 若 $U = V$, 则称 V 为**循环空间**. ♣

定理 7.8 (循环子空间的基本性质)

设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,φ 是 V 上的线性变换,$\mathbf{0} \neq \alpha \in V$, $U = C(\varphi, \alpha)$ 为循环子空间, 则循环子空间 U 是 V 的 φ-不变子空间, 并且是包含 α 的最小 φ-不变子空间. ♡

证明 U 是 V 的 φ-不变子空间是显然的. 下证 U 是包含 α 的最小 φ-不变子空间.

设 $\alpha \in W$, 且 W 为 φ-不变子空间, 则由数学归纳法易知

$$\alpha, \varphi^k(\alpha) \in W, \forall k \in \mathbb{N}_1.$$

于是

$$U = L(\alpha, \varphi(\alpha), \cdots) \in W.$$

□

定理 7.9

设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,φ 是 V 上的线性变换,$\mathbf{0} \neq \alpha \in V$, $U = C(\varphi, \alpha)$ 为循环子空间, 若 $\dim U = r$, 求证:$\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 是 U 的一组基. ♡

证明 设 $m = \max\{k \in \mathbb{Z}^+ \mid \alpha, \varphi(\alpha), \cdots, \varphi^{k-1}(\alpha)$ 线性无关$\}$, 则显然 $1 \leq m \leq r$, 故 m 是良定义的. 于是由命题 3.3和数学归纳法容易验证: 对任意的 $k \geq m$,$\varphi^k(\alpha)$ 都是 $\alpha, \varphi(\alpha), \cdots, \varphi^{m-1}(\alpha)$ 的线性组合, 于是 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{m-1}(\alpha)\}$ 是 U 的一组基, 从而 $m = \dim U = r$. □

定理 7.10

设 U 是 V 的 φ-不变子空间, 求证:U 为循环子空间的充要条件是 $\varphi|_U$ 在 U 的某组基下的表示矩阵为某个首一多项式的友阵. ♡

证明 **充分性**: 设 $\varphi|_U$ 在 U 的一组基 $\{e_1, e_2, \cdots, e_r\}$ 下的表示矩阵是友阵 $C(d(\lambda))$, 其中 $d(\lambda) = \lambda^r + a_1\lambda^{r-1} + \cdots + a_{r-1}\lambda + a_r$, 则由友阵的定义可知 $\varphi(e_i) = e_{i+1} (1 \leq i \leq r-1)$,$\varphi(e_r) = -\sum\limits_{i=1}^{r} a_{r-i+1}e_i$. 因此 $e_i = \varphi^{i-1}(e_1) (2 \leq i \leq r)$,$U = L(e_1, e_2, \cdots, e_r) = C(\varphi, e_1)$ 为循环子空间.

必要性: 设 $U = C(\varphi, \alpha)$ 是 r 维循环子空间, 则由定理 7.9可知,$\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 是 U 的一组基. 设

$$\varphi^r(\alpha) = -a_r\alpha - a_{r-1}\varphi(\alpha) - \cdots - a_1\varphi^{r-1}(\alpha)$$

令 $d(\lambda) = \lambda^r + a_1\lambda^{r-1} + \cdots + a_{r-1}\lambda + a_r$, 容易验证:$\varphi|_U$ 在基 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 下的表示矩阵就是友阵 $C(d(\lambda))$. □

定理 7.11 (有理标准型的几何意义)

设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,φ 是 V 上的线性变换, 且 φ 的不变因子组是 $1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)$, 其中 $d_i(\lambda)$ 是非常数首一多项式,$d_i(\lambda)\mid d_{i+1}(\lambda)\ (1\leq i\leq k-1)$, 则 V 存在一个循环子空间的直和分解:

$$V=C(\varphi,\alpha_1)\oplus C(\varphi,\alpha_2)\oplus\cdots\oplus C(\varphi,\alpha_k) \tag{7.15}$$

使得 $\varphi|_{C(\varphi,\alpha_i)}$ 在基 $\{\alpha_i,\varphi(\alpha_i),\cdots,\varphi^{r_i-1}(\alpha_i)\}$ 下的表示矩阵就是友阵 $C(d_i(\lambda))$, 其中 $r_i=\dim C(\varphi,\alpha_i)$.

笔记 线性变换 φ 的有理标准型诱导的 V 的上述循环子空间直和分解 (7.15)就是有理标准型的几何意义.

证明 由定理 7.6可知, 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

$$C=\operatorname{diag}\{C(d_1(\lambda)),C(d_2(\lambda)),\cdots,C(d_k(\lambda))\}$$

其中 $\varphi|_{L(e_{i1},\cdots,e_{ir_i})}$ 的表示矩阵就是友阵 $C(d_i(\lambda)),i=1,2,\cdots,k$. 再结合定理 7.10的讨论可知,$L(e_{i1},\cdots,e_{ir_i})$ 就是一个循环子空间. 于是任取 $\alpha_i\in L(e_{i1},\cdots,e_{ir_i})$ 作为循环向量, 则

$$C(\varphi,\alpha_i)=L(e_{i1},\cdots,e_{ir_i})=L(\alpha_i,\varphi(\alpha_i),\cdots,\varphi^{r_i-1}(\alpha_i))$$

其中 $\dim C(\varphi,\alpha_i)=r_i$.

综上可知, 此时 V 存在一个循环子空间的直和分解:

$$V=C(\varphi,\alpha_1)\oplus C(\varphi,\alpha_2)\oplus\cdots\oplus C(\varphi,\alpha_k)$$

使得 $\varphi|_{C(\varphi,\alpha_i)}$ 在基 $\{\alpha_i,\varphi(\alpha_i),\cdots,\varphi^{r_i-1}(\alpha_i)\}$ 下的表示矩阵就是友阵 $C(d_i(\lambda))$, 其中 $r_i=\dim C(\varphi,\alpha_i)$. □

定理 7.12 (循环子空间的刻画)

设 φ 是数域 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换,φ 的特征多项式和极小多项式分别为 $f(\lambda)$ 和 $m(\lambda)$, 证明以下 4 个结论等价:

(1) φ 的行列式因子组或不变因子组为 $1,\cdots,1,f(\lambda)$;

(2) φ 的初等因子组为 $P_1(\lambda)^{r_1},P_2(\lambda)^{r_2},\cdots,P_k(\lambda)^{r_k}$, 其中 $P_i(\lambda)$ 是 $\mathbb{K}$ 上互异的首一不可约多项式,$r_i\geq 1,1\leq i\leq k$;

(3) φ 的极小多项式 $m(\lambda)$ 等于特征多项式 $f(\lambda)$;

(4) V 是关于线性变换 φ 的循环空间.

证明 (1) $\Leftrightarrow$ (2): 由不变因子和初等因子之间的相互转换即得.

(1) $\Leftrightarrow$ (3): 由极小多项式等于最大的不变因子, 以及所有不变因子的乘积等于特征多项式即得.

(1) $\Leftrightarrow$ (4): 若 V 是循环空间, 则由定理 7.10可知,φ 在某组基下的表示矩阵是友阵 $C(g(\lambda))$, 再由友阵的性质 (引理 7.5) 可知,φ 的行列式因子组和不变因子组均为 $1,\cdots,1,g(\lambda)=f(\lambda)$. 若 φ 的不变因子组为 $1,\cdots,1,f(\lambda)$, 则由有理标准型的几何意义 (定理 7.11) 可知,V 是循环空间. □

7.5 初等因子

定义 7.9 (初等因子)

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, 由于 $d_i(\lambda)\mid d_{i+1}(\lambda)$, 因此可以在 $\mathbb{K}$ 上把 $d_i(\lambda)$ 分解成不可约因式之积:

$$\begin{aligned}
d_1(\lambda)&=p_1(\lambda)^{e_{11}}p_2(\lambda)^{e_{12}}\cdots p_t(\lambda)^{e_{1t}},\\
d_2(\lambda)&=p_1(\lambda)^{e_{21}}p_2(\lambda)^{e_{22}}\cdots p_t(\lambda)^{e_{2t}},\\
&\cdots\cdots\cdots\cdots\\
d_k(\lambda)&=p_1(\lambda)^{e_{k1}}p_2(\lambda)^{e_{k2}}\cdots p_t(\lambda)^{e_{kt}},
\end{aligned} \tag{7.16}$$

其中 e_{ij} 是非负整数 (注意 e_{ij} 可以为零!), 并且

$$e_{1j} \le e_{2j} \le \cdots \le e_{kj}, \quad j = 1, 2, \cdots, t.$$

若(7.16)式中的 $e_{ij} > 0$, 则称 $p_j(\lambda)^{e_{ij}}$ 为 A 的一个**初等因子**,A 的全体初等因子称为 A 的**初等因子组**.

命题 7.3

矩阵 A 的初等因子组与不变因子组相互唯一确定.

证明 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 $p_j(\lambda)^{e_{ij}}$, 适当增加一些 1 (表示为 $p_j(\lambda)^{e_{ij}}$, 其中 $e_{ij} = 0$), 则可将这组初等因子按不可约因式的降幂排列如下:

$$\begin{aligned} &p_1(\lambda)^{e_{k1}}, p_1(\lambda)^{e_{k-1,1}}, \cdots, p_1(\lambda)^{e_{11}}, \\ &p_2(\lambda)^{e_{k2}}, p_2(\lambda)^{e_{k-1,2}}, \cdots, p_2(\lambda)^{e_{12}}, \\ &\quad\cdots\cdots\cdots\cdots \\ &p_t(\lambda)^{e_{kt}}, p_t(\lambda)^{e_{k-1,t}}, \cdots, p_t(\lambda)^{e_{1t}}, \end{aligned} \tag{7.17}$$

令

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

则 $d_i(\lambda) \mid d_{i+1}(\lambda)$ $(i = 1, \cdots, k-1)$, 且 $d_1(\lambda), \cdots, d_k(\lambda)$ 的初等因子组就如 (7.17)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. □

定理 7.13

数域 $\mathbb{K}$ 上的两个矩阵 A 与 B 相似的充分必要条件是它们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.

证明 由定理 7.5可知, 矩阵 A 和 B 相似等价于 A 和 B 有相同的不变因子. 又由命题 7.3可知,A 和 B 有相同的不变因子等价于它们有相同的初等因子组. 故矩阵 A 与 B 相似的充分必要条件是它们有相同的初等因子组. □

例题 7.4 设 9 阶矩阵 A 的不变因子组为

$$1, \cdots, 1, (\lambda-1)(\lambda^2+1), (\lambda-1)^2(\lambda^2+1)(\lambda^2-2),$$

试分别在有理数域、实数域和复数域上求 A 的初等因子组.

解 A 在有理数域上的初等因子组为

$$\lambda-1, (\lambda-1)^2, \lambda^2+1, \lambda^2+1, \lambda^2-2.$$

A 在实数域上的初等因子组为

$$\lambda-1, (\lambda-1)^2, \lambda^2+1, \lambda^2+1, \lambda+\sqrt{2}, \lambda-\sqrt{2}.$$

A 在复数域上的初等因子组为

$$\lambda-1, (\lambda-1)^2, \lambda+\mathrm{i}, \lambda+\mathrm{i}, \lambda-\mathrm{i}, \lambda-\mathrm{i}, \lambda+\sqrt{2}, \lambda-\sqrt{2}.$$

□

例题 7.5 设 A 是一个 10 阶矩阵, 它的初等因子组为

$$\lambda-1, \lambda-1, (\lambda-1)^2, (\lambda+1)^2, (\lambda+1)^3, \lambda-2.$$

求 A 的不变因子组.

解 将上述多项式按不可约因式的降幂排列:

$$\begin{array}{ccc}
(\lambda-1)^2, & \lambda-1, & \lambda-1; \\
(\lambda+1)^3, & (\lambda+1)^2, & 1; \\
\lambda-2, & 1, & 1.
\end{array}$$

于是

$$d_3(\lambda)=(\lambda-1)^2(\lambda+1)^3(\lambda-2), \quad d_2(\lambda)=(\lambda-1)(\lambda+1)^2, \quad d_1(\lambda)=\lambda-1.$$

从而 A 的不变因子组为

$$1, \cdots, 1, \lambda-1, (\lambda-1)(\lambda+1)^2, (\lambda-1)^2(\lambda+1)^3(\lambda-2),$$

其中有 7 个 1. □

7.6 Jordan 标准型

引理 7.7

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 显然 J 的特征多项式为 $(\lambda-\lambda_0)^r$. 对任一小于 r 的正整数 k,$\lambda I-J$ 总有一个 k 阶子式, 其值等于 $(-1)^k$, 因此 J 的行列式因子为

$$1, \cdots, 1, (\lambda-\lambda_0)^r. \tag{7.18}$$

(7.18)式也是 J 的不变因子组, 故 J 的初等因子组只有一个多项式 $(\lambda-\lambda_0)^r$. □

引理 7.8

设特征矩阵 $\lambda I-A$ 经过初等变换化为下列对角阵:

$$\begin{pmatrix}
f_1(\lambda) & & & \\
& f_2(\lambda) & & \\
& & \ddots & \\
& & & f_n(\lambda)
\end{pmatrix}, \tag{7.19}$$

其中 $f_i(\lambda)$ $(i=1,\cdots,n)$ 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**, 所有 $f_i(\lambda)$ 的准素因子称为 A 的**准素因子组**, 则矩阵 A 的初等因子组等于所有 $f_i(\lambda)$ 的准素因子组.

注 这个引理给出了求矩阵初等因子组的另外一个方法, 它可以不必先求不变因子组而直接用初等变换把特征矩阵化为对角阵, 再分解主对角线上的多项式即可. 另外, 这个引理的结论及其证明在一般的数域 $\mathbb{K}$ 上也成立.

证明 第一步, 先证明下列事实:

若 $f_i(\lambda), f_j(\lambda)$ $(i \neq j)$ 的最大公因式和最小公倍式分别为 $g(\lambda), h(\lambda)$, 则

$$\mathrm{diag}\{f_1(\lambda), \cdots, f_i(\lambda), \cdots, f_j(\lambda), \cdots, f_n(\lambda)\}$$

经过初等变换可以变为

$$\mathrm{diag}\{f_1(\lambda), \cdots, g(\lambda), \cdots, h(\lambda), \cdots, f_n(\lambda)\},$$

且这两个对角阵具有相同的准素因子组.

不失一般性, 令 $i = 1, j = 2$. 因为 $(f_1(\lambda), f_2(\lambda)) = g(\lambda)$, 所以存在 $u(\lambda), v(\lambda)$, 使

$$f_1(\lambda)u(\lambda) + f_2(\lambda)v(\lambda) = g(\lambda).$$

又令 $f_1(\lambda) = g(\lambda)q(\lambda), f_2(\lambda) = g(\lambda)q'(\lambda)$. 则 $h(\lambda) = g(\lambda)q(\lambda)q'(\lambda) = f_2(\lambda)q(\lambda)$. 对(7.19)式作下列初等变换:

$$\begin{pmatrix} f_1(\lambda) & & & \\ & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow{u(\lambda)\cdot r_1 + r_2} \begin{pmatrix} f_1(\lambda) & 0 & & \\ f_1(\lambda)u(\lambda) & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow{v(\lambda) j_2 + j_1}$$

$$\begin{pmatrix} f_1(\lambda) & & & \\ g(\lambda) & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow{-q(\lambda) r_2 + r_1} \begin{pmatrix} 0 & -h(\lambda) & & \\ g(\lambda) & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow[(-1)\cdot r_2]{r_2 \longleftrightarrow r_1}$$

$$\begin{pmatrix} g(\lambda) & f_2(\lambda) & & \\ & h(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix} \xrightarrow{-q'(\lambda) j_1 + j_2} \begin{pmatrix} g(\lambda) & & & \\ & h(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix}.$$

现来考察 $g(\lambda)$ 与 $h(\lambda)$ 的准素因子. 将 $f_1(\lambda), f_2(\lambda)$ 作标准因式分解, 其分解式不妨设为

$$f_1(\lambda) = (\lambda - \lambda_1)^{c_1}(\lambda - \lambda_2)^{c_2} \cdots (\lambda - \lambda_t)^{c_t},$$
$$f_2(\lambda) = (\lambda - \lambda_1)^{d_1}(\lambda - \lambda_2)^{d_2} \cdots (\lambda - \lambda_t)^{d_t},$$

其中 c_i, d_i 为非负整数. 令

$$e_i = \max\{c_i, d_i\}, \quad k_i = \min\{c_i, d_i\},$$

则

$$g(\lambda) = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_t)^{k_t},$$
$$h(\lambda) = (\lambda - \lambda_1)^{e_1}(\lambda - \lambda_2)^{e_2} \cdots (\lambda - \lambda_t)^{e_t}.$$

不难看出 $g(\lambda), h(\lambda)$ 的准素因子组与 $f_1(\lambda), f_2(\lambda)$ 的准素因子组相同.

第二步证明 (7.19)式所示矩阵的法式可通过上述变换得到.

先将第 $(1,1)$ 位置的元素依次和第 $(2,2)$ 位置,$\cdots$, 第 (n,n) 位置的元素进行上述变换, 此时第 $(1,1)$ 元素的所有一次因式的幂都是最小的; 再将第 $(2,2)$ 位置的元素依次和第 $(3,3)$ 位置,$\cdots$, 第 (n,n) 位置的元素进行上述变换;$\cdots$; 最后将第 $(n-1,n-1)$ 位置的元素和第 (n,n) 位置的元素进行上述变换. 可以看出, 最后得到的对角阵就是(7.19)式所示矩阵的法式. 注意到在每一次变换的过程中, 准素因子组都保持不变, 这就证明了结论. □

例题 7.6 设 $\lambda I - A$ 经过初等变换后化为下列对角阵:

$$\begin{pmatrix} 1 & & & & \\ & (\lambda-1)^2(\lambda+2) & & & \\ & & \lambda+2 & & \\ & & & 1 & \\ & & & & \lambda-1 \end{pmatrix},$$

求 A 的初等因子组.

解 由引理 7.8 知,A 的初等因子组为 $\lambda-1,(\lambda-1)^2,\lambda+2,\lambda+2$. □

引理 7.9

设 J 是分块对角阵:

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix},$$

其中每个 J_i 都是形如引理 7.7中的矩阵,J_i 的初等因子组为 $(\lambda-\lambda_i)^{r_i}$, 则 J 的初等因子组为

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

证明 $\lambda I-J$ 是一个分块对角 λ-矩阵. 由于对分块对角阵中某一块施行初等变换时其余各块保持不变, 故由引理 7.7及命题 7.2知,$\lambda I-J$ 相抵于下列分块对角阵:

$$H=\begin{pmatrix} H_1 & & & \\ & H_2 & & \\ & & \ddots & \\ & & & H_k \end{pmatrix},$$

其中 $H_i=\mathrm{diag}\{1,\cdots,1,(\lambda-\lambda_i)^{r_i}\}$. 再由引理 7.8即得结论. □

定理 7.14 (Jordan 标准型)

设 A 是复数域上的矩阵且 A 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 相似于分块对角阵:

$$J=\begin{pmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_k(\lambda_k) \end{pmatrix}, \tag{7.20}$$

其中 J_i 为 r_i 阶矩阵, 且

$$J_i=\begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}.$$

(7.20)式中的矩阵 J 称为 A 的 **Jordan 标准型**, 每个 $J_i(\lambda_i)$ 称为 A 的一个 **Jordan 块**.

注 由引理 7.7可以看出, 若交换任意两个 Jordan 块的位置, 得到的矩阵与原来的矩阵仍有相同的初等因子组, 它们仍相似. 因此矩阵 A 的 Jordan 标准型中 Jordan 块的排列可以是任意的. 但是, 由于每个初等因子唯一确定了一个 Jordan 块, 故若不计 Jordan 块的排列次序, 则矩阵的 Jordan 标准型是唯一确定的.

证明 由定理 7.13知,A 与 J 有相同的初等因子组, 因此 A 与 J 相似. □

命题 7.4 (Jordan 块的性质)

(1) $J_n(\lambda_0)$ 的行列式因子组和不变因子组都是

$$1, \cdots, 1, (\lambda - \lambda_0)^n.$$

进而初等因子组为 $(\lambda-\lambda_0)^n$, 极小多项式等于特征多项式等于 $(\lambda-\lambda_0)^n$.

(2) 设 $J_0 = J_n(0)$, 则 J_0 是基础幂零阵, 且 $J_0^n = 0$.

(3) 设 $J = J_n(\lambda_0)$ 是特征值为 λ_0 的 n 阶 Jordan 块, 则和 J 乘法可交换的 n 阶矩阵必可表示为 J 的次数不超过 $n-1$ 的多项式.

证明

(1) 由引理 7.7和定理 7.7即得.

(2) 由 Jordan 块的定义可直接得到

$$J_0 = J_n(0) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix}.$$

从而结论显然成立.

(3) 证法一 (几何方法):注意到 $J - \lambda_0 I_n = J_n(0)$ 为基础幂零阵, 于是

$$\varphi(e_n) = e_{n-1}, \varphi^2(e_n) = \varphi(e_n - 1) = e_{n-2}, \cdots, \varphi^{n-1}(e_n) = e_1.$$

故 $\mathbb{C}^n = L(\varphi^{n-1}(e_n), \varphi^{n-2}(e_n), \cdots, e_n) = C(J - \lambda_0 I_n, \boldsymbol{e}_n)$ 是关于线性变换 $J - \lambda_0 I_n$ 的循环空间 (也可由Jordan 标准型的几何意义直接得到), 循环向量是标准单位列向量中的最后一个 $\boldsymbol{e}_n = (0, \cdots, 0, 1)'$, 又由 (1) 可知,$J - \lambda_0 I_n = J_n(0)$ 的特征多项式与极小多项式相等都是 λ^n. 于是再由定理 7.25即得结论.

证法二 (代数方法): 设 A 和 J 可交换, 注意到 $J = \lambda_0 I_n + J_0$, 其中 $J_0 = J_n(0)$ 是特征值为零的 Jordan 块, 故 A,J 乘法可交换当且仅当 A,J_0 乘法可交换. 经计算得到 A 必为下列形状的上三角矩阵:

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ & a_1 & \ddots & \vdots \\ & & \ddots & a_2 \\ & & & a_1 \end{pmatrix}$$

于是

$$A = a_1 I_n + a_2 J_0 + \cdots + a_n J_0^{n-1} = a_1 I_n + a_2(J - \lambda_0 I_n) + \cdots + a_n (J - \lambda_0 I_n)^{n-1}$$

可表示为 J 的次数不超过 $n-1$ 的多项式.

□

定理 7.15

设 φ 是复数域上线性空间 V 上的线性变换, 则必存在 V 的一组基, 使得 φ 在这组基下的表示矩阵为(7.20)式所示的 Jordan 标准型.

证明 由定理 7.14立得. □

推论 7.7

设 A 是 n 阶复矩阵, 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) $\Rightarrow$ (2): 由可对角化的判定条件 (5)的结论即得.

(2) $\Rightarrow$ (3): 设 A 的极小多项式 $m(\lambda)$ 无重根. 由于 $m(\lambda)$ 是 A 的最后一个不变因子, 故 A 的所有不变因子都无重根, 从而 A 的初等因子都是一次多项式.

(3) $\Rightarrow$ (1): 设 A 的初等因子组为 $\lambda-\lambda_1, \lambda-\lambda_2, \cdots, \lambda-\lambda_n$, 则由定理 7.14知,$A$ 相似于对角阵 $\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 即 A 可对角化. □

推论 7.8

设 φ 是复线性空间 V 上的线性变换, 则 φ 可对角化当且仅当

1. φ 的极小多项式无重根;
2. φ 的初等因子都是一次多项式;
3. φ 的 Jordan 块都是一阶矩阵.

证明

□

推论 7.9

设 φ 是复线性空间 V 上的线性变换,V_0 是 φ 的不变子空间. 若 φ 可对角化, 则 φ 在 V_0 上的限制也可对角化.

证明 设 $\varphi, \varphi|_{V_0}$ 的极小多项式分别为 $g(\lambda), h(\lambda)$, 则由推论 7.8知,$g(\lambda)$ 无重根. 又 $g(\varphi|_{V_0}) = g(\varphi)|_{V_0} = \mathbf{0}$, 故 $h(\lambda) \mid g(\lambda)$, 于是 $h(\lambda)$ 也无重根, 再次由推论 7.8知,$\varphi|_{V_0}$ 可对角化. □

推论 7.10

设 φ 是复线性空间 V 上的线性变换, 且 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$, 其中每个 V_i 都是 φ 的不变子空间, 则 φ 可对角化的充分必要条件是 φ 在每个 V_i 上的限制都可对角化.

证明 必要性由推论 7.9即得, 下证充分性. 若 φ 在每个 V_i 上的限制都可对角化, 则由定义存在 V_i 的一组基, 使得 $\varphi|_{V_i}$ 在这组基下的表示矩阵是对角阵. 再由定理**??**知 V_i 的一组基可以拼成 V 的一组基, 因此 φ 在这组基下的表示阵是对角阵, 即 φ 可对角化. □

推论 7.11

设 A 是数域 $\mathbb{K}$ 上的矩阵, 如果 A 的特征值全在 $\mathbb{K}$ 中, 则 A 在 $\mathbb{K}$ 上相似于其 Jordan 标准型.

证明 由于 A 的特征值全在 $\mathbb{K}$ 中, 故 A 的 Jordan 标准型 J 实际上是 $\mathbb{K}$ 上的矩阵. 因为 A 在复数域上相似于 J, 由推论 7.10 知,A 在 $\mathbb{K}$ 上也相似于 J. □

例题 7.7 设 A 是 7 阶矩阵, 其初等因子组为

$$\lambda - 1, (\lambda - 1)^3, (\lambda + 1)^2, \lambda - 2,$$

求其 Jordan 标准型.

解 A 的 Jordan 标准型为

$$J = \begin{pmatrix} 1 & & & & & & \\ & 1 & 1 & 0 & & & \\ & 0 & 1 & 1 & & & \\ & 0 & 0 & 1 & & & \\ & & & & -1 & 1 & \\ & & & & 0 & -1 & \\ & & & & & & 2 \end{pmatrix},$$

J 含有 4 个 Jordan 块. □

例题 7.8 设复数域上的四维线性空间 V 上的线性变换 φ 在一组基 $\{e_1, e_2, e_3, e_4\}$ 下的表示矩阵为

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

求 V 的一组基, 使 φ 在这组基下的表示矩阵为 Jordan 标准型, 并求出从原来的基到新基的过渡矩阵.

解 用初等变换把 $\lambda I - A$ 化为对角 λ-矩阵并求出它的初等因子组为

$$(\lambda-1)^2, (\lambda-1)^2.$$

因此,A 的 Jordan 标准型为

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

设矩阵 P 是从 $\{e_1, e_2, e_3, e_4\}$ 到新基的过渡矩阵, 则

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

此即

$$AP = PJ. \tag{7.21}$$

设 $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 其中 α_i 是四维列向量, 代入(7.21)式得

$$(A\alpha_1, A\alpha_2, A\alpha_3, A\alpha_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)\begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix},$$

化成方程组为

$$(A-I)\alpha_1 = \mathbf{0},$$
$$(A-I)\alpha_2 = \alpha_1,$$
$$(A-I)\alpha_3 = \mathbf{0},$$
$$(A-I)\alpha_4 = \alpha_3.$$

由于 α_1, α_3 都是 A 的属于特征值 1 的特征向量, 故 α_2, α_4 称为属于特征值 1 的广义特征向量. 我们可取方程组 $(A-I)x = \mathbf{0}$ 的两个线性无关的解分别作为 α_1, α_3 (注意不能取线性相关的两个解, 因为 P 是非异阵), 然后再分别

求出 α_2,α_4 (注意诸 α_i 的解可能不唯一, 只需取比较简单的一组解) 即可. 经计算可得

$$\alpha_1=\begin{pmatrix}1\\-2\\1\\-5\end{pmatrix},\quad \alpha_2=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\quad \alpha_3=\begin{pmatrix}0\\0\\1\\-1\end{pmatrix},\quad \alpha_4=\begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

于是

$$P=\begin{pmatrix}1&0&0&0\\-2&1&0&0\\1&0&1&0\\-5&0&-1&1\end{pmatrix}.$$

因此新基为

$$\{e_1-2e_2+e_3-5e_4,e_2,e_3-e_4,e_4\}.$$

□

7.7 Jordan 标准型的进一步讨论和应用

定理 7.16

线性变换 φ 的特征值 λ_1 的度数等于 φ 的 Jordan 标准型中属于特征值 λ_1 的 Jordan 块的个数,λ_1 的重数等于所有属于特征值 λ_1 的 Jordan 块的阶数之和.

♡

证明 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

$$(\lambda-\lambda_1)^{r_1},(\lambda-\lambda_2)^{r_2},\cdots,(\lambda-\lambda_k)^{r_k}, \tag{7.22}$$

定理 7.15告诉我们, 存在 V 的一组基 $\{e_{11},e_{12},\cdots,e_{1r_1};e_{21},e_{22},\cdots,e_{2r_2};\cdots;e_{k1},e_{k2},\cdots,e_{kr_k}\}$, 使得 φ 在这组基下的表示矩阵为

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

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 是由基向量 $e_{i1},e_{i2},\cdots,e_{ir_i}$ 生成的子空间, 则

$$\begin{aligned}\varphi(e_{i1})&=\lambda_ie_{i1},\\ \varphi(e_{i2})&=e_{i1}+\lambda_ie_{i2},\\ &\cdots\cdots\cdots\cdots\\ \varphi(e_{ir_i})&=e_{i,r_i-1}+\lambda_ie_{ir_i}.\end{aligned} \tag{7.23}$$

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. 显然我们有

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

线性变换 φ 限制在 V_1 上 (仍记为 φ) 便成为 V_1 上的线性变换. 这个线性变换在基 $\{e_{11},e_{12},\cdots,e_{1r_1}\}$ 下的表

示矩阵为

$$J_1=\begin{pmatrix}\lambda_1 & 1 & & & \\ & \lambda_1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_1\end{pmatrix}.$$

注意到 $\boldsymbol{J}_1$ 的特征值全为 λ_1, 并且 $\lambda_1\boldsymbol{I}-\boldsymbol{J}_1$ 的秩等于 r_1-1, 故 $\boldsymbol{J}_1$ 只有一个线性无关的特征向量, 不妨选为 e_{11}. 显然 e_{11} 也是 φ 作为 V 上线性变换关于特征值 λ_1 的特征向量. 不失一般性, 不妨设在 φ 的初等因子组即(7.22) 式中

$$\lambda_1=\lambda_2=\cdots=\lambda_s,\quad \lambda_i\neq\lambda_1(i=s+1,\cdots,k),$$

则 $\boldsymbol{J}_1,\cdots,\boldsymbol{J}_s$ 都以 λ_1 为特征值, 且**相应于每一块有且只有一个线性无关的特征向量**. 相应的特征向量可取为

$$e_{11},e_{21},\cdots,e_{s1},\tag{7.24}$$

显然这是 s 个线性无关的特征向量. 如果 $\lambda_i\neq\lambda_1$, 则容易看出 $\mathrm{r}(\lambda_1\boldsymbol{I}-\boldsymbol{J}_i)=r_i$, 于是

$$\mathrm{r}(\lambda_1\boldsymbol{I}-\boldsymbol{J})=\sum_{i=1}^{k}\mathrm{r}(\lambda_1\boldsymbol{I}-\boldsymbol{J}_i)=(r_1-1)+\cdots+(r_s-1)+r_{s+1}+\cdots+r_k=n-s.$$

因此 φ 关于特征值 λ_1 的特征子空间 V_{λ_1} 的维数等于 $n-\mathrm{r}(\lambda_1\boldsymbol{I}-\boldsymbol{J})=s$, 从而特征子空间 V_{λ_1} 以(7.24)式中的向量为一组基. 又 λ_1 是 φ 的 $r_1+r_2+\cdots+r_s$ 重特征值, 因此 λ_1 的重数与度数之差等于

$$(r_1+r_2+\cdots+r_s)-s.$$

□

推论 7.12

线性变化 (矩阵) 的 Jordan 标准型中属于特征值 λ_0 的每一个 Jordan 块都有且仅有一个线性无关的特征向量.

♡

证明 由上述定理 7.16的证明立得. □

定义 7.10 (根子空间)

设 λ_0 是 n 维复线性空间 V 上线性变换 φ 的特征值, 则

$$R(\lambda_0)=\{\boldsymbol{v}\in V\mid(\varphi-\lambda_0\boldsymbol{I})^n(\boldsymbol{v})=\boldsymbol{0}\}$$

构成了 V 的一个子空间, 称为属于特征值 λ_0 的**根子空间**.

♣

定理 7.17

设 φ 是 n 维复线性空间 V 上的线性变换.

(1) 若 φ 的初等因子组为

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

则 V 可分解为 k 个不变子空间的直和:

$$V=V_1\oplus V_2\oplus\cdots\oplus V_k,\tag{7.25}$$

其中 V_i 是维数等于 r_i 的关于 $\varphi-\lambda_i\boldsymbol{I}$ 的循环子空间;

(2) 若 $\lambda_1,\cdots,\lambda_s$ 是 φ 的全体不同特征值, 则 V 可分解为 s 个不变子空间的直和:

$$V=R(\lambda_1)\oplus R(\lambda_2)\oplus\cdots\oplus R(\lambda_s),$$

其中 $R(\lambda_i)$ 是 λ_i 的根子空间,$R(\lambda_i)$ 的维数等于 λ_i 的重数, 且每个 $R(\lambda_i)$ 又可分解为(7.25) 式中若干个 V_j 的直和.

♡

证明 在定理 7.16的证明的基础上, 现在再来看 $\boldsymbol{J}_1$ 所对应的子空间 V_1, 由 (7.23)式中诸等式可知

$$(\varphi-\lambda_1\boldsymbol{I})(e_{1r_1})=e_{1,r_1-1},\cdots,(\varphi-\lambda_1\boldsymbol{I})(e_{12})=e_{11},(\varphi-\lambda_1\boldsymbol{I})(e_{11})=\boldsymbol{0},$$

因此, 若记 $\boldsymbol{\alpha}=e_{1r_1},\psi=\varphi-\lambda_1\boldsymbol{I}$, 则

$$\psi(\boldsymbol{\alpha})=e_{1,r_1-1},\psi^2(\boldsymbol{\alpha})=e_{1,r_1-2},\cdots,\psi^{r_1-1}(\boldsymbol{\alpha})=e_{11},\psi^{r_1}(\boldsymbol{\alpha})=\boldsymbol{0}.$$

也就是说

$$\{\boldsymbol{\alpha},\psi(\boldsymbol{\alpha}),\psi^2(\boldsymbol{\alpha}),\cdots,\psi^{r_1-1}(\boldsymbol{\alpha})\}$$

构成了 V_1 的一组基.

上面的事实说明, 每个 Jordan 块 $\boldsymbol{J}_i$ 对应的子空间 V_i 是一个循环子空间. 把属于同一个特征值, 比如属于 λ_1 的所有循环子空间加起来构成 V 的一个子空间:

$$R(\lambda_1)=V_1\oplus\cdots\oplus V_s.$$

若 $\boldsymbol{v}\in R(\lambda_1)$, 则不难算出 $(\varphi-\lambda_1\boldsymbol{I})^s(\boldsymbol{v})=\boldsymbol{0}$, 其中

$$s=\dim R(\lambda_1)=r_1+\cdots+r_s.$$

事实上, 我们可以证明

$$R(\lambda_1)=\{\boldsymbol{v}\in V\mid(\varphi-\lambda_1\boldsymbol{I})^n(\boldsymbol{v})=\boldsymbol{0}\}.\tag{7.26}$$

为证明 (7.26)式成立, 设 $U=\{\boldsymbol{v}\in V\mid(\varphi-\lambda_1\boldsymbol{I})^n(\boldsymbol{v})=\boldsymbol{0}\}$, 则由上面的分析知道,$R(\lambda_1)\subseteq U$. 另一方面, 任取 $\boldsymbol{v}\in U$, 设 $\boldsymbol{v}=\boldsymbol{v}_1+\boldsymbol{v}_2$, 其中 $\boldsymbol{v}_1\in R(\lambda_1)$,$\boldsymbol{v}_2\in V_{s+1}\oplus\cdots\oplus V_k$. 因为 $(\lambda-\lambda_1)^n$ 与 $(\lambda-\lambda_{s+1})^n\cdots(\lambda-\lambda_k)^n$ 互素, 故存在多项式 $p(\lambda),q(\lambda)$, 使

$$(\lambda-\lambda_1)^np(\lambda)+(\lambda-\lambda_{s+1})^n\cdots(\lambda-\lambda_k)^nq(\lambda)=1.$$

将 $\lambda=\varphi$ 代入上式并作用在 $\boldsymbol{v}$ 上可得

$$\begin{aligned}\boldsymbol{v}&=p(\varphi)(\varphi-\lambda_1\boldsymbol{I})^n(\boldsymbol{v})+q(\varphi)(\varphi-\lambda_{s+1}\boldsymbol{I})^n\cdots(\varphi-\lambda_k\boldsymbol{I})^n(\boldsymbol{v})\\&=q(\varphi)(\varphi-\lambda_{s+1}\boldsymbol{I})^n\cdots(\varphi-\lambda_k\boldsymbol{I})^n(\boldsymbol{v}_1)+q(\varphi)(\varphi-\lambda_{s+1}\boldsymbol{I})^n\cdots(\varphi-\lambda_k\boldsymbol{I})^n(\boldsymbol{v}_2)\\&=q(\varphi)(\varphi-\lambda_{s+1}\boldsymbol{I})^n\cdots(\varphi-\lambda_k\boldsymbol{I})^n(\boldsymbol{v}_1)\in R(\lambda_1).\end{aligned}$$

这就证明了 (7.26)式.

上面的结果表明: 特征值 λ_0 的根子空间可表示为若干个循环子空间的直和, 每个循环子空间对应于一个 Jordan 块. 虽然我们前面的讨论是对特征值 λ_1 进行的, 其实对任一特征值 λ_i 均适用. □

命题 7.5

证明: 复数域上的方阵 $\boldsymbol{A}$ 必可分解为两个对称阵的乘积.

证明 设 $\boldsymbol{P}$ 是非异阵且使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}$ 为 $\boldsymbol{A}$ 的 Jordan 标准型, 于是 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}$. 设 $\boldsymbol{J}_i$ 是 $\boldsymbol{J}$ 的第 i 个 Jordan 块, 则

$$\boldsymbol{J}_i=\begin{pmatrix}\lambda_i&1&&&\\&\lambda_i&1&&\\&&\ddots&\ddots&\\&&&\ddots&1\\&&&&\lambda_i\end{pmatrix}=\begin{pmatrix}&&&1&\lambda_i\\&&1&\lambda_i&\\&\iddots&\iddots&&\\1&\iddots&&&\\\lambda_i&&&&\end{pmatrix}\begin{pmatrix}&&&&1\\&&&1&\\&&\iddots&&\\&1&&&\\1&&&&\end{pmatrix},$$

即 $\boldsymbol{J}_i$ 可分解为两个对称阵之积. 因此 $\boldsymbol{J}$ 也可以分解为两个对称阵之积, 记为 $\boldsymbol{S}_1,\boldsymbol{S}_2$, 于是

$$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}=\boldsymbol{P}\boldsymbol{S}_1\boldsymbol{S}_2\boldsymbol{P}^{-1}=(\boldsymbol{P}\boldsymbol{S}_1\boldsymbol{P}')(\boldsymbol{P}^{-1})'\boldsymbol{S}_2\boldsymbol{P}^{-1}.$$

显然 $\boldsymbol{P}\boldsymbol{S}_1\boldsymbol{P}'$ 和 $(\boldsymbol{P}^{-1})'\boldsymbol{S}_2\boldsymbol{P}^{-1}$ 都是对称矩阵, 故 $\boldsymbol{A}$ 必可分解为两个对称阵的乘积. □

例题 7.9 已知

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

计算 A^k.

解 用初等变换把 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 化为对角 λ-矩阵并求出它的初等因子组为

$$(\lambda-1)^2,(\lambda-1)^2.$$

因此,$\boldsymbol{A}$ 的 Jordan 标准型为

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

因为

$$\boldsymbol{P}^{-1}\boldsymbol{A}^k\boldsymbol{P}=(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^k=\boldsymbol{J}^k,$$

故先计算 $\boldsymbol{J}^k$. 注意 $\boldsymbol{J}$ 是分块对角阵, 它的 k 次方等于将各对角块 k 次方, 因此

$$\boldsymbol{J}^k=\begin{pmatrix}1&1&&\\0&1&&\\&&1&1\\&&0&1\end{pmatrix}^k=\begin{pmatrix}1&k&&\\0&1&&\\&&1&k\\&&0&1\end{pmatrix},$$

$$\begin{aligned}\boldsymbol{A}^k=\boldsymbol{P}\boldsymbol{J}^k\boldsymbol{P}^{-1}&=\begin{pmatrix}1&0&0&0\\-2&1&0&0\\1&0&1&0\\-5&0&-1&1\end{pmatrix}\begin{pmatrix}1&k&&\\0&1&&\\&&1&k\\&&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\-2&1&0&0\\1&0&1&0\\-5&0&-1&1\end{pmatrix}^{-1}\\&=\begin{pmatrix}2k+1&k&0&0\\-4k&-2k+1&0&0\\6k&k&k+1&k\\-14k&-5k&-k&-k+1\end{pmatrix}.\end{aligned}$$

□

定理 7.18 (Jordan-Chevalley 分解)

设 $\boldsymbol{A}$ 是 n 阶复矩阵, 则 $\boldsymbol{A}$ 可分解为 $\boldsymbol{A}=\boldsymbol{B}+\boldsymbol{C}$, 其中 $\boldsymbol{B},\boldsymbol{C}$ 适合下面条件:

(1) $\boldsymbol{B}$ 是一个可对角化矩阵;

(2) $\boldsymbol{C}$ 是一个幂零阵;

(3) $\boldsymbol{B}\boldsymbol{C}=\boldsymbol{C}\boldsymbol{B}$;

(4) $\boldsymbol{B},\boldsymbol{C}$ 均可表示为 $\boldsymbol{A}$ 的多项式.

不仅如此, 上述满足条件 (1)(3) 的分解是唯一的 (即只要满足条件 (1)(3) 的分解就是唯一的). 进而, 上述满足条件 (1)(2)(3)(4) 的分解也是唯一的.

证明 先对 $\boldsymbol{A}$ 的 Jordan 标准型 $\boldsymbol{J}$ 证明结论. 设 $\boldsymbol{A}$ 的全体不同特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 且

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s \end{pmatrix},$$

其中 $\boldsymbol{J}_i$ 是属于特征值 λ_i 的根子空间对应的块, 其阶设为 m_i. 显然对每个 i 均有 $\boldsymbol{J}_i = \boldsymbol{M}_i + \boldsymbol{N}_i$, 其中 $\boldsymbol{M}_i = \lambda_i \boldsymbol{I}$ 是对角阵,$\boldsymbol{N}_i$ 是幂零阵且 $\boldsymbol{M}_i\boldsymbol{N}_i = \boldsymbol{N}_i\boldsymbol{M}_i$. 令

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{M}_1 & & & \\ & \boldsymbol{M}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{M}_s \end{pmatrix}, \quad \boldsymbol{N} = \begin{pmatrix} \boldsymbol{N}_1 & & & \\ & \boldsymbol{N}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{N}_s \end{pmatrix},$$

则 $\boldsymbol{J} = \boldsymbol{M} + \boldsymbol{N}, \boldsymbol{M}\boldsymbol{N} = \boldsymbol{N}\boldsymbol{M}, \boldsymbol{M}$ 是对角阵,$\boldsymbol{N}$ 是幂零阵.

因为 $(\boldsymbol{J}_i - \lambda_i \boldsymbol{I})^{m_i} = \boldsymbol{O}$, 所以 $\boldsymbol{J}_i$ 适合多项式 $(\lambda - \lambda_i)^{m_i}$. 而 λ_i 互不相同, 因此多项式 $(\lambda-\lambda_1)^{m_1}, (\lambda-\lambda_2)^{m_2}, \cdots, (\lambda-\lambda_s)^{m_s}$ 两两互素. 由中国剩余定理, 存在多项式 $g(\lambda)$ 满足条件

$$g(\lambda) = h_i(\lambda)(\lambda - \lambda_i)^{m_i} + \lambda_i,$$

对所有 $i = 1, 2, \cdots, s$ 成立 (这里 $h_i(\lambda)$ 也是多项式). 代入 $\boldsymbol{J}_i$ 得到

$$g(\boldsymbol{J}_i) = h_i(\boldsymbol{J}_i)(\boldsymbol{J}_i - \lambda_i \boldsymbol{I})^{m_i} + \lambda_i \boldsymbol{I} = \lambda_i \boldsymbol{I} = \boldsymbol{M}_i.$$

于是

$$g(\boldsymbol{J}) = \begin{pmatrix} g(\boldsymbol{J}_1) & & & \\ & g(\boldsymbol{J}_2) & & \\ & & \ddots & \\ & & & g(\boldsymbol{J}_s) \end{pmatrix} = \begin{pmatrix} \boldsymbol{M}_1 & & & \\ & \boldsymbol{M}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{M}_s \end{pmatrix} = \boldsymbol{M}.$$

又因为 $\boldsymbol{N} = \boldsymbol{J} - \boldsymbol{M} = \boldsymbol{J} - g(\boldsymbol{J})$, 所以 $\boldsymbol{N}$ 也是 $\boldsymbol{J}$ 的多项式.

现考虑一般情形, 设 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J}$, 则 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1} = \boldsymbol{P}(\boldsymbol{M}+\boldsymbol{N})\boldsymbol{P}^{-1}$. 令 $\boldsymbol{B} = \boldsymbol{P}\boldsymbol{M}\boldsymbol{P}^{-1}, \boldsymbol{C} = \boldsymbol{P}\boldsymbol{N}\boldsymbol{P}^{-1}$, 则 $\boldsymbol{B}$ 是可对角化矩阵,$\boldsymbol{C}$ 是幂零阵,$\boldsymbol{B}\boldsymbol{C} = \boldsymbol{C}\boldsymbol{B}$ 并且

$$g(\boldsymbol{A}) = g(\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}) = \boldsymbol{P}g(\boldsymbol{J})\boldsymbol{P}^{-1} = \boldsymbol{P}\boldsymbol{M}\boldsymbol{P}^{-1} = \boldsymbol{B},$$

从而 $\boldsymbol{C} = \boldsymbol{A} - g(\boldsymbol{A})$.

最后证明唯一性. 假设 $\boldsymbol{A}$ 有另一满足条件 (1) (3) 的分解 $\boldsymbol{A} = \boldsymbol{B}_1 + \boldsymbol{C}_1$, 则 $\boldsymbol{B} - \boldsymbol{B}_1 = \boldsymbol{C}_1 - \boldsymbol{C}$. 由 $\boldsymbol{B}_1\boldsymbol{C}_1 = \boldsymbol{C}_1\boldsymbol{B}_1$ 不难验证 $\boldsymbol{A}\boldsymbol{B}_1 = \boldsymbol{B}_1\boldsymbol{A}, \boldsymbol{A}\boldsymbol{C}_1 = \boldsymbol{C}_1\boldsymbol{A}$. 因为 $\boldsymbol{B} = g(\boldsymbol{A})$, 故 $\boldsymbol{B}\boldsymbol{B}_1 = \boldsymbol{B}_1\boldsymbol{B}$. 同理 $\boldsymbol{C}\boldsymbol{C}_1 = \boldsymbol{C}_1\boldsymbol{C}$. 设 $\boldsymbol{C}^r = \boldsymbol{O}, \boldsymbol{C}_1^t = \boldsymbol{O}$, 用二项式定理即知 $(\boldsymbol{C}_1 - \boldsymbol{C})^{r+t} = \boldsymbol{O}$. 于是

$$(\boldsymbol{B} - \boldsymbol{B}_1)^{r+t} = (\boldsymbol{C}_1 - \boldsymbol{C})^{r+t} = \boldsymbol{O}.$$

因为 $\boldsymbol{B}\boldsymbol{B}_1 = \boldsymbol{B}_1\boldsymbol{B}$, 它们都是可对角化矩阵, 由命题 6.22知它们可同时对角化, 即存在可逆阵 $\boldsymbol{Q}$, 使 $\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}$ 和 $\boldsymbol{Q}^{-1}\boldsymbol{B}_1\boldsymbol{Q}$ 都是对角阵. 注意到

$$(\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q} - \boldsymbol{Q}^{-1}\boldsymbol{B}_1\boldsymbol{Q})^{r+t} = \left(\boldsymbol{Q}^{-1}(\boldsymbol{B} - \boldsymbol{B}_1)\boldsymbol{Q}\right)^{r+t} = \boldsymbol{Q}^{-1}(\boldsymbol{B} - \boldsymbol{B}_1)^{r+t}\boldsymbol{Q} = \boldsymbol{O},$$

两个对角阵之差仍是一个对角阵, 这个差的幂要等于零矩阵, 则这两个矩阵必相等, 由此即得 $\boldsymbol{B} = \boldsymbol{B}_1$, 从而 $\boldsymbol{C} = \boldsymbol{C}_1$.
□

7.8 矩阵函数

引理 7.10

设 $A \in M_n(\mathbb{C})$, 则存在 P 是非异阵, 使

$$P^{-1}AP = J = \operatorname{diag}\{J_1, J_2, \cdots, J_k\}$$

是 A 的 Jordan 标准型, 其中 J_i 是 A 的特征值 λ_i 的 r 阶 Jordan 块. 若 $f(x) = a_0 + a_1 x + \cdots + a_p x^p$, 则

$$f(A) = P^{-1} f(J) P = P\operatorname{diag}\left\{f\left(J_1\right), f\left(J_2\right), \cdots, f\left(J_k\right)\right\} P^{-1},$$

其中

$$f(J_i) = \begin{pmatrix} f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \frac{1}{2!}f^{(2)}(\lambda_i) & \cdots & \frac{1}{(r-1)!}f^{(r-1)}(\lambda_i) \\ & f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \cdots & \frac{1}{(r-2)!}f^{(r-2)}(\lambda_i) \\ & & f(\lambda_i) & \cdots & \frac{1}{(r-3)!}f^{(r-3)}(\lambda_i) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_i) \end{pmatrix}.$$

证明 注意到

$$J^m = \operatorname{diag}\{J_1^m, J_2^m, \cdots, J_k^m\}.$$

又

$$A^m = (PJP^{-1})^m = PJ^mP^{-1},$$

因此要计算 $f(A)$, 只需计算出 J_i^m 即可. 利用二项式定理和数学归纳法不难证明

$$J_i^m = \left[\lambda_i I_r + \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix}\right]^m = \begin{pmatrix} \lambda_i^m & \mathrm{C}_m^1\lambda_i^{m-1} & \mathrm{C}_m^2\lambda_i^{m-2} & \cdots & \cdots \\ & \lambda_i^m & \mathrm{C}_m^1\lambda_i^{m-1} & \cdots & \cdots \\ & & \lambda_i^m & \cdots & \cdots \\ & & & \ddots & \vdots \\ & & & & \lambda_i^m \end{pmatrix}.$$

则不难算出

$$f(J_i) = \begin{pmatrix} f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \frac{1}{2!}f^{(2)}(\lambda_i) & \cdots & \frac{1}{(r-1)!}f^{(r-1)}(\lambda_i) \\ & f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \cdots & \frac{1}{(r-2)!}f^{(r-2)}(\lambda_i) \\ & & f(\lambda_i) & \cdots & \frac{1}{(r-3)!}f^{(r-3)}(\lambda_i) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_i) \end{pmatrix}.$$

再由

$$\begin{aligned} f(A) &= f(PJP^{-1}) = Pf(J)P^{-1} \\ &= Pf(\operatorname{diag}\{J_1, J_2, \cdots, J_k\})P^{-1} \\ &= P\operatorname{diag}\{f(J_1), f(J_2), \cdots, f(J_k)\}P^{-1}, \end{aligned}$$

即可计算出 $f(A)$. □

定义 7.11 (复方阵幂级数)

设有 n 阶复方阵序列 $\{A_p\}$:

$$A_p = \begin{pmatrix} a_{11}^{(p)} & \cdots & a_{1n}^{(p)} \\ \vdots & & \vdots \\ a_{n1}^{(p)} & \cdots & a_{nn}^{(p)} \end{pmatrix},$$

$B = (b_{ij})$ 是一个同阶方阵, 若对每个 (i, j), 序列 $\{a_{ij}^{(p)}\}$ 均收敛于 b_{ij}, 即

$$\lim_{p\to\infty} a_{ij}^{(p)} = b_{ij},$$

则称矩阵序列 $\{A_p\}$ 收敛于 B, 记为

$$\lim_{p\to\infty} A_p = B.$$

否则称 $\{A_p\}$ 发散.

设

$$f(z) = a_0 + a_1 z + a_2 z^2 + \cdots$$

是一个幂级数, 记

$$f_p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_p z^p$$

是其部分和. 若矩阵序列 $\{f_p(A)\}$ 收敛于 B, 则称矩阵级数

$$f(A) = a_0 I + a_1 A + a_2 A^2 + \cdots.$$

收敛, 极限为 B, 记为 $f(A) = B$. 否则称 $f(A)$ 发散. 用变量矩阵 X 代替 A, 便可定义矩阵幂级数

$$f(X) = a_0 I + a_1 X + a_2 X^2 + \cdots.$$

♣

命题 7.6

设 A 是 n 阶矩阵, 证明: $\lim\limits_{k\to\infty} A^k$ 存在的充要条件是 A 的特征值的模长小于 1, 或者特征值等于 1 并且 A 关于特征值 1 的 Jordan 块都是一阶的. 此时, 极限矩阵为

$$\lim_{k\to\infty} A^k = P\mathrm{diag}\{1, \cdots, 1, 0, \cdots, 0\}P^{-1},$$

其中 1 的个数等于 A 的特征值 1 的代数重数.

♠

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_s}(\lambda_s)\}$ 为 A 的 Jordan 标准型, 则

$$A^k = PJ^kP^{-1} = P\mathrm{diag}\{J_{r_1}(\lambda_1)^k, J_{r_2}(\lambda_2)^k, \cdots, J_{r_s}(\lambda_s)^k\}P^{-1},$$

因此 $\lim\limits_{k\to\infty} A^k$ 存在当且仅当 $\lim\limits_{k\to\infty} J_{r_i}(\lambda_i)^k (1 \leqslant i \leqslant s)$ 都存在. 不妨取 $k > n$, 由引理 7.10计算可得 Jordan 块 $J_{r_i}(\lambda_i)$ 的 k 次幂为

$$J_{r_i}(\lambda_i)^k = \begin{pmatrix} \lambda_i^k & \mathrm{C}_k^1\lambda_i^{k-1} & \mathrm{C}_k^2\lambda_i^{k-2} & \cdots & \mathrm{C}_k^{r_i-1}\lambda_i^{k-r_i+1} \\ & \lambda_i^k & \mathrm{C}_k^1\lambda_i^{k-1} & \cdots & \mathrm{C}_k^{r_i-2}\lambda_i^{k-r_i+2} \\ & & \lambda_i^k & \cdots & \mathrm{C}_k^{r_i-3}\lambda_i^{k-r_i+3} \\ & & & \ddots & \vdots \\ & & & & \lambda_i^k \end{pmatrix},$$

故当 $|\lambda_i| \geqslant 1$ 且 $\lambda_i \neq 1$ 时, $\lim\limits_{k\to\infty} \lambda_i^k$ 发散; 当 $\lambda_i = 1$ 且 $r_i \geqslant 2$ 时, $\lim\limits_{k\to\infty} \mathrm{C}_k^1\lambda_i^{k-1}$ 发散; 当 $\lambda_i = 1$ 且 $r_1 = 1$ 时, $\lim\limits_{k\to\infty} J_{r_i}(\lambda_i)^k = J_1(1)$; 当 $|\lambda_i| < 1$ 时, $\lim\limits_{k\to\infty} J_{r_i}(\lambda_i)^k = O$. 因此, $\lim\limits_{k\to\infty} A^k$ 存在的充要条件是 A 的特征值的模长小于 1, 或者特征值等于 1 并且 A 关于特征值 1 的 Jordan 块都是一阶的. 此时, 极限矩阵 $\lim\limits_{k\to\infty} A^k = P\,\mathrm{diag}\,\{1, \cdots, 1, 0, \cdots, 0\}\,P^{-1}$, 其中 1 的个数等于 A 的特征值 1 的代数重数. □

定理 7.19

设 $f(z)=\sum_{i=0}^{\infty}a_iz^i$ 是复幂级数, 则

(1) 方阵幂级数 $f(X)$ 收敛的充分必要条件是对任一非异阵 $P,f(P^{-1}XP)$ 都收敛, 这时

$$f(P^{-1}XP)=P^{-1}f(X)P;$$

(2) 若 $X=\mathrm{diag}\{X_1,\cdots,X_k\}$, 则 $f(X)$ 收敛的充分必要条件是 $f(X_1),\cdots,f(X_k)$ 都收敛, 这时

$$f(X)=\mathrm{diag}\{f(X_1),\cdots,f(X_k)\};$$

(3) 若 $f(z)$ 的收敛半径为 r,J_0 是特征值为 λ_0 的 n 阶 Jordan 块

$$J_0=\begin{pmatrix}\lambda_0 & 1 & & & \\ & \lambda_0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_0\end{pmatrix},$$

则当 $|\lambda_0|<r$ 时 $f(J_0)$ 收敛, 且

$$f(J_0)=\begin{pmatrix}f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \frac{1}{2!}f^{(2)}(\lambda_0) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda_0) \\ & f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda_0) \\ & & f(\lambda_0) & \cdots & \frac{1}{(n-3)!}f^{(n-3)}(\lambda_0) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_0)\end{pmatrix}. \tag{7.27}$$

证明 设 $f_p(z)=a_0+a_1z+a_2z^2+\cdots+a_pz^p$ 是 $f(z)$ 前 $p+1$ 项的部分和.

(1) 注意到 $f_p(z)$ 是多项式, 从而有

$$f_p(P^{-1}XP)=P^{-1}f_p(X)P.$$

由于 n 阶矩阵序列的收敛等价于 n^2 个数值序列的收敛, 故

$$\begin{aligned}f(P^{-1}XP)&=\lim_{p\to\infty}f_p(P^{-1}XP)=\lim_{p\to\infty}P^{-1}f_p(X)P\\&=P^{-1}(\lim_{p\to\infty}f_p(X))P=P^{-1}f(X)P.\end{aligned}$$

(2) 注意到 $f_p(z)$ 是多项式, 从而有

$$f_p(X)=f_p(\mathrm{diag}\{X_1,\cdots,X_k\})=\mathrm{diag}\{f_p(X_1),\cdots,f_p(X_k)\}.$$

由于分块矩阵序列的收敛等价于每个分块的矩阵序列的收敛, 故

$$\begin{aligned}f(X)&=\lim_{p\to\infty}f_p(X)=\lim_{p\to\infty}\mathrm{diag}\{f_p(X_1),\cdots,f_p(X_k)\}\\&=\mathrm{diag}\{\lim_{p\to\infty}f_p(X_1),\cdots,\lim_{p\to\infty}f_p(X_k)\}=\mathrm{diag}\{f(X_1),\cdots,f(X_k)\}.\end{aligned}$$

(3) 由引理 7.10可知

$$f_p(J_0)=\begin{pmatrix} f_p(\lambda_0) & \frac{1}{1!}f_p'(\lambda_0) & \frac{1}{2!}f_p^{(2)}(\lambda_0) & \cdots & \frac{1}{(n-1)!}f_p^{(n-1)}(\lambda_0) \\ & f_p(\lambda_0) & \frac{1}{1!}f_p'(\lambda_0) & \cdots & \frac{1}{(n-2)!}f_p^{(n-2)}(\lambda_0) \\ & & f_p(\lambda_0) & \cdots & \frac{1}{(n-3)!}f_p^{(n-3)}(\lambda_0) \\ & & & \ddots & \vdots \\ & & & & f_p(\lambda_0) \end{pmatrix}.$$

令 $p\to\infty$, 由矩阵序列收敛与 n^2 个数值序列收敛的等价性和幂级数的相关性质即得结论.

□

定理 7.20

设 $f(z)$ 是复幂级数, 收敛半径为 r. 设 A 是 n 阶复方阵, 特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 定义 A 的**谱半径**

$$\rho(A)=\max_{1\le i\le n}|\lambda_i|.$$

(1) 若 $\rho(A)<r$, 则 $f(A)$ 收敛;
(2) 若 $\rho(A)>r$, 则 $f(A)$ 发散;
(3) 若 $\rho(A)=r$, 则 $f(A)$ 收敛的充分必要条件是: 对每一模长等于 r 的特征值 λ_j, 若 A 的属于 λ_j 的初等因子中最高幂为 n_j 次, 则 n_j 个数值级数

$$f(\lambda_j),f'(\lambda_j),\cdots,f^{(n_j-1)}(\lambda_j). \tag{7.28}$$

都收敛;
(4) 若 $f(A)$ 收敛, 则 $f(A)$ 的特征值为

$$f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n).$$

♡

证明

(1) 设 A 的 Jordan 标准型为 $J=\mathrm{diag}\{J_1,J_2,\cdots,J_k\}$. 显然 $f(A)$ 的收敛性等价于所有 $f(J_i)$ （$i=1,\cdots,k$）的收敛性. 由定理 7.19即知 (1) 成立.
(2) 若某一个 $|\lambda_j|>r$, 则 $f(\lambda_j)$ 发散, 因此 $f(J_j)$ 发散, 故 $f(A)$ 发散, 这就证明了 (2).
(3) 当 $\rho(A)=r$ 时, 对 $|\lambda_i|<r$ 的 J_i,$f(J_i)$ 收敛. 对 $|\lambda_j|=r$ 的特征值 λ_j, 注意到 $f(z)$ 的任意次导数的收敛半径仍为 r, 又初等因子 $(\lambda-\lambda_j)^{n_j}$ 对应的 Jordan 块为 n_j 阶, 从(7.27)式即可知道 $f(J_j)$ 的收敛性等价于(7.28)式中 n_j 个级数的收敛性.
(4) 最后若 $f(A)$ 收敛, 则 $f(A)$ 与 $f(J)$ 有相同的特征值, 即为 $f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$.

□

定义 7.12

于是对一切方阵, 定义

$$\mathrm{e}^{\boldsymbol{A}}=\boldsymbol{I}+\frac{1}{1!}\boldsymbol{A}+\frac{1}{2!}\boldsymbol{A}^2+\frac{1}{3!}\boldsymbol{A}^3+\cdots,$$
$$\sin\boldsymbol{A}=\boldsymbol{A}-\frac{1}{3!}\boldsymbol{A}^3+\frac{1}{5!}\boldsymbol{A}^5-\frac{1}{7!}\boldsymbol{A}^7+\cdots,$$
$$\cos\boldsymbol{A}=\boldsymbol{I}-\frac{1}{2!}\boldsymbol{A}^2+\frac{1}{4!}\boldsymbol{A}^4-\frac{1}{6!}\boldsymbol{A}^6+\cdots$$

都有意义. 若 $\boldsymbol{A}$ 所有特征值的模长都小于 1, 则

$$\ln(\boldsymbol{I}+\boldsymbol{A})=\boldsymbol{A}-\frac{1}{2}\boldsymbol{A}^2+\frac{1}{3}\boldsymbol{A}^3-\frac{1}{4}\boldsymbol{A}^4+\cdots$$

也有意义. 同理还可以定义幂函数、双曲函数等.

♣

注 由复分析知道:

$$\begin{aligned}
e^z &= 1+\frac{1}{1!}z+\frac{1}{2!}z^2+\frac{1}{3!}z^3+\cdots,\\
\sin z &= z-\frac{1}{3!}z^3+\frac{1}{5!}z^5-\frac{1}{7!}z^7+\cdots,\\
\cos z &= 1-\frac{1}{2!}z^2+\frac{1}{4!}z^4-\frac{1}{6!}z^6+\cdots,\\
\ln(1+z) &= z-\frac{1}{2}z^2+\frac{1}{3}z^3-\frac{1}{4}z^4+\cdots.
\end{aligned}$$

前 3 个级数在整个复平面上收敛, 而 $\ln(1+z)$ 的收敛半径为 1. 于是由定理 7.20可知 $e^{\boldsymbol{A}}, \sin\boldsymbol{A}, \cos\boldsymbol{A}, \ln\boldsymbol{A}$ 都收敛, 从而都有意义. 故上述定义是良定义的.

定理 7.21

证明:$\cos A=\dfrac{e^{\mathrm{i}A}+e^{-\mathrm{i}A}}{2},\quad \sin A=\dfrac{e^{\mathrm{i}A}-e^{-\mathrm{i}A}}{2\mathrm{i}}.$

证明 由定义 7.12可知

$$\begin{aligned}
e^{\mathrm{i}A} &= I+\frac{1}{1!}\mathrm{i}A-\frac{1}{2!}A^2-\frac{1}{3!}\mathrm{i}A^3-\cdots+\frac{(-1)^k}{2k!}A^{2k}+\frac{(-1)^k}{(2k+1)!}\mathrm{i}A^{2k+1}+\cdots,\\
e^{-\mathrm{i}A} &= I-\frac{1}{1!}\mathrm{i}A-\frac{1}{2!}A^2+\frac{1}{3!}\mathrm{i}A^3-\cdots+\frac{(-1)^k}{2k!}A^{2k}+\frac{(-1)^{k+1}}{(2k+1)!}\mathrm{i}A^{2k+1}+\cdots.
\end{aligned}$$

从而

$$\begin{aligned}
\frac{e^{\mathrm{i}A}+e^{-\mathrm{i}A}}{2} &= I-\frac{1}{2!}A^2+\cdots+\frac{(-1)^k}{2k!}A^{2k}+\cdots=\cos A,\\
\frac{e^{\mathrm{i}A}-e^{-\mathrm{i}A}}{2\mathrm{i}} &= \frac{1}{1!}A-\frac{1}{3!}A^3+\cdots+\frac{(-1)^k}{(2k+1)!}A^{2k+1}\cdots=\sin A.
\end{aligned}$$

□

命题 7.7

求证: 若 n 阶矩阵 A, B 乘法可交换, 则 $e^A\cdot e^B=e^B\cdot e^A$.

注 对一般来说对矩阵 $\boldsymbol{A}, \boldsymbol{B}$, 下面的等式并不一定成立:

$$e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}}=e^{\boldsymbol{A}+\boldsymbol{B}}=e^{\boldsymbol{B}}\cdot e^{\boldsymbol{A}}.$$

如对

$$\boldsymbol{A}=\begin{pmatrix}1&1\\0&0\end{pmatrix}, \boldsymbol{B}=\begin{pmatrix}0&0\\1&1\end{pmatrix},$$

通过计算不难验证 $AB=A, BA=B$, 并且

$$e^A=\begin{pmatrix}e&e-1\\0&1\end{pmatrix},\quad e^B=\begin{pmatrix}1&0\\e-1&e\end{pmatrix},\quad e^{A+B}=\begin{pmatrix}\dfrac{e^2+1}{2}&\dfrac{e^2-1}{2}\\ \dfrac{e^2-1}{2}&\dfrac{e^2+1}{2}\end{pmatrix}$$

故 $e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}}\neq e^{\boldsymbol{A}+\boldsymbol{B}}$.

证明 设 $f(z)=e^z$, 并且 $f_p(z)=1+\frac{1}{1!}z+\frac{1}{2!}z^2+\cdots+\frac{1}{p!}z^p$ 为 $f(z)$ 的部分和, 因为 $f(z)$ 的收敛半径为 $+\infty$, 所以对任一矩阵 A, $\lim\limits_{p\to\infty}f_p(A)=f(A)=e^A$. 由于 $AB=BA$, 故对任意的正整数 p, q, 成立 $f_p(A)f_q(B)=f_q(B)f_p(A)$. 先固定 p, 令 $q\to\infty$, 则可得

$$\begin{aligned}
f_p(A)f(B) &= f_p(A)\left(\lim_{q\to\infty}f_q(B)\right)=\lim_{q\to\infty}\left(f_p(A)f_q(B)\right)=\lim_{q\to\infty}\left(f_q(B)f_p(A)\right)\\
&= \left(\lim_{q\to\infty}f_q(B)\right)f_p(A)=f(B)f_p(A)
\end{aligned}$$

同理, 再对上式令 $p\to\infty$, 则可得 $f(A)f(B)=f(B)f(A)$, 即结论成立. □

推论 7.13

若 $f(z), g(z)$ 是两个收敛半径都是 $+\infty$ 的复幂级数, 则对任意乘法可交换的 A, B, 均有 $f(A)g(B) = g(B)f(A)$. ♡

证明 由命题 7.7类似的讨论可证明. □

定义 7.13 (矩阵的范数)

设 $A = (a_{ij})$ 是 n 阶复矩阵, 定义 A 的**范数**为其所有元素模长的平方和的算术平方根, 即 $\|A\| = \sqrt{\sum\limits_{i,j=1}^{n} |a_{ij}|^2}$. ♣

命题 7.8 (矩阵的范数的基本性质)

设 $A = (a_{ij})$ 是 n 阶复矩阵,$B = (b_{ij})$ 也是 n 阶复矩阵, 求证:

(1) $\|A\| \geq 0$, 等号成立当且仅当 $A = O$;

(2) $\|A + B\| \leq \|A\| + \|B\|$;

(3) $\|AB\| \leq \|A\| \cdot \|B\|$. ♠

证明 (1) 显然成立.

(2) 注意到

$$\begin{aligned}\|A+B\|^2 &= \sum_{i,j=1}^{n} \left|a_{ij}+b_{ij}\right|^2 = \sum_{i,j=1}^{n} \left(\left|a_{ij}\right|^2 + \left|b_{ij}\right|^2 + 2\left|a_{ij}b_{ij}\right|\right) \\ &= \sum_{i,j=1}^{n} \left|a_{ij}\right|^2 + \sum_{i,j=1}^{n} \left|b_{ij}\right|^2 + 2\sum_{i,j=1}^{n} \left|a_{ij}b_{ij}\right|\end{aligned}$$

于是

$$\begin{aligned}(\|A\|+\|B\|)^2 &= \|A\|^2 + \|B\|^2 + 2\|A\|\cdot\|B\| \\ &= \sum_{i,j=1}^{n} \left|a_{ij}\right|^2 + \sum_{i,j=1}^{n} \left|b_{ij}\right|^2 + 2\sqrt{\left(\sum_{i,j=1}^{n} \left|a_{ij}\right|^2\right)\left(\sum_{i,j=1}^{n} \left|b_{ij}\right|^2\right)} \\ &\overset{\text{Cauchy-Schwarz 不等式}}{\geqslant} \sum_{i,j=1}^{n} \left|a_{ij}\right|^2 + \sum_{i,j=1}^{n} \left|b_{ij}\right|^2 + 2\sqrt{\left(\sum_{i,j=1}^{n} \left|a_{ij}b_{ij}\right|\right)^2} \\ &= \sum_{i,j=1}^{n} \left|a_{ij}\right|^2 + \sum_{i,j=1}^{n} \left|b_{ij}\right|^2 + 2\sum_{i,j=1}^{n} \left|a_{ij}b_{ij}\right| \\ &= \|A+B\|^2 .\end{aligned}$$

故结论得证.

(3) 注意到 $\|AB\|^2 = \sum\limits_{i,j=1}^{n} \left|\sum\limits_{k=1}^{n} a_{ik}b_{kj}\right|^2$,$\|A\|^2 \cdot \|B\|^2 = \left(\sum\limits_{i,k=1}^{n} |a_{ik}|^2\right) \cdot \left(\sum\limits_{k,j=1}^{n} |b_{kj}|^2\right)$, 任取 $i, j \in \{1, 2, \cdots, n\}$, 固定 i 和 j, 由 Cauchy-Schwarz 不等式可知

$$\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2 \leqslant \left(\sum_{k=1}^{n} a_{ik}^2\right) \cdot \left(\sum_{k=1}^{n} b_{kj}^2\right).$$

从而先对 i 求和可得

$$\sum_{i=1}^{n} \left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2 \leqslant \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik}^2\right) \cdot \left(\sum_{k=1}^{n} b_{kj}^2\right) = \left(\sum_{k=1}^{n} b_{kj}^2\right) \left(\sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik}^2\right)\right).$$

再对 j 求和可得

$$\begin{aligned}\sum_{i,j=1}^{n}\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right|^{2} &\leqslant \sum_{j=1}^{n}\left[\left(\sum_{k=1}^{n}b_{kj}^{2}\right)\left(\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{ik}^{2}\right)\right)\right]\\ &=\left(\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{ik}^{2}\right)\right)\left(\sum_{j=1}^{n}\left(\sum_{k=1}^{n}b_{kj}^{2}\right)\right)\\ &=\left(\sum_{i,k=1}^{n}|a_{ik}|^{2}\right)\cdot\left(\sum_{k,j=1}^{n}|b_{kj}|^{2}\right).\end{aligned}$$

由此即得结论. □

命题 7.9

如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 乘法可交换, 即 $\boldsymbol{AB}=\boldsymbol{BA}$, 则 $\mathrm{e}^{\boldsymbol{A}}\cdot\mathrm{e}^{\boldsymbol{B}}=\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$ 必成立.

注 利用这个命题也可给出命题 7.7的另一证明.

证明 设 $f(z)=\mathrm{e}^{z}$, 并且 $f_p(z)=1+\frac{1}{1!}z+\frac{1}{2!}z^2+\cdots+\frac{1}{p!}z^p$ 为 $f(z)$ 的部分和. 注意到 $AB=BA$, 经简单的计算可知,$f_p(A)f_p(B)$ 展开后的单项包含 $f_p(A+B)$ 展开后的所有单项, 且剩余单项可表示为 $\frac{A^i}{i!}\frac{B^j}{j!}$ 的形式, 其中 $i+j>p$, 故由矩阵的范数的基本性质 (3)可得

$$\begin{aligned}\|f_p(A)f_p(B)-f_p(A+B)\| &\le \sum_{k>p}\left(\sum_{i+j=k}\frac{\|A\|^i}{i!}\frac{\|B\|^j}{j!}\right)=\sum_{k>p}\left(\sum_{i=0}^{k}\frac{\|A\|^i}{i!}\frac{\|B\|^{k-i}}{(k-i)!}\right)\\ &=\sum_{k>p}\left(\sum_{i=0}^{k}\frac{k!}{i!\,(k-i)!}\frac{\|A\|^i\|B\|^{k-i}}{k!}\right)=\sum_{k>p}\left(\sum_{i=0}^{k}\mathrm{C}_k^i\frac{\|A\|^i\|B\|^{k-i}}{k!}\right)\\ &=\sum_{k>p}\frac{(\|A\|+\|B\|)^k}{k!}.\end{aligned}$$

由于数项级数 $\sum\limits_{k=0}^{\infty}\frac{1}{k!}(\|A\|+\|B\|)^k$ 收敛到 $\mathrm{e}^{\|A\|+\|B\|}$, 故当 p 充分大时, 上式右边趋于零. 令 $p\to\infty$, 则由上式即得 $\|f(A)f(B)-f(A+B)\|=0$, 再次由矩阵的范数的基本性质 (1)可得 $\mathrm{e}^{A}\cdot\mathrm{e}^{B}=\mathrm{e}^{A+B}$. □

推论 7.14

若矩阵幂级数 e^{A} 绝对收敛, 则矩阵级数的 Cauchy 乘积

$$\mathrm{e}^{A+B}=\sum_{p=0}^{\infty}\left(\sum_{i+j=p}\frac{A^i}{i!}\frac{B^j}{j!}\right)$$

收敛到

$$\left(\sum_{i=0}^{\infty}\frac{A^i}{i!}\right)\cdot\left(\sum_{j=0}^{\infty}\frac{B^j}{j!}\right)=\mathrm{e}^{A}\cdot\mathrm{e}^{B}.$$

注 注意矩阵级数的 Cauchy 积有

$$\begin{aligned}\mathrm{e}^{A+B}&=\sum_{p=0}^{\infty}\frac{(A+B)^p}{p!}=\sum_{p=0}^{\infty}\left(\sum_{i=0}^{p}\mathrm{C}_p^i\frac{A^iB^{p-i}}{p!}\right)\\ &=\sum_{p=0}^{\infty}\left(\sum_{i=0}^{p}\frac{p!}{i!\,(p-i)!}\frac{A^iB^{p-i}}{p!}\right)=\sum_{p=0}^{\infty}\left(\sum_{i=0}^{p}\frac{A^i}{i!}\frac{B^{p-i}}{(p-i)}\right)\end{aligned}$$

$$= \sum_{p=0}^{\infty}\left(\sum_{i+j=p}\frac{A^i}{i!}\frac{B^j}{j!}\right).$$

证明 由命题 7.9立得. □

命题 7.10

设 t 是一个数值变量,$\boldsymbol{A}$ 是一个 n 阶复方阵. $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}=\operatorname{diag}\{\boldsymbol{J}_1,\cdots,\boldsymbol{J}_k\}$ 是 $\boldsymbol{A}$ 的 Jordan 标准型,$\boldsymbol{J}_i$ 是特征值为 λ_i 的 r 阶 Jordan 块, 则

$$\mathrm{e}^{t\boldsymbol{A}}=\boldsymbol{P}\mathrm{e}^{t\boldsymbol{J}}\boldsymbol{P}^{-1},$$

其中

$$\mathrm{e}^{t\boldsymbol{J}}=\begin{pmatrix}\mathrm{e}^{t\boldsymbol{J}_1} & & \\ & \ddots & \\ & & \mathrm{e}^{t\boldsymbol{J}_k}\end{pmatrix},\quad \mathrm{e}^{t\boldsymbol{J}_i}=\mathrm{e}^{t\lambda_i}\begin{pmatrix}1 & t & \frac{1}{2!}t^2 & \frac{1}{3!}t^3 & \cdots & \frac{1}{(r-1)!}t^{r-1}\\ & 1 & t & \frac{1}{2!}t^2 & \cdots & \frac{1}{(r-2)!}t^{r-2}\\ & & 1 & t & \cdots & \frac{1}{(r-3)!}t^{r-3}\\ & & & \ddots & \ddots & \vdots\\ & & & & \ddots & t\\ & & & & & 1\end{pmatrix}.$$

证明 证法一: 若令 $f(z)=\mathrm{e}^{tz}$, 则由定理 7.19即得 $f(\boldsymbol{A})=\mathrm{e}^{t\boldsymbol{A}}$ 的计算结果.

证法二: 注意到

$$\boldsymbol{J}_i=\lambda_i\boldsymbol{I}+\boldsymbol{N},$$

其中 $\boldsymbol{N}$ 是 r 阶基础幂零阵, 即

$$\boldsymbol{N}^r=\boldsymbol{O},\quad \boldsymbol{N}=\begin{pmatrix}0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ & & & & 0\end{pmatrix}.$$

于是

$$\mathrm{e}^{\boldsymbol{N}}=\boldsymbol{I}+\boldsymbol{N}+\frac{1}{2!}\boldsymbol{N}^2+\frac{1}{3!}\boldsymbol{N}^3+\cdots+\frac{1}{(r-1)!}\boldsymbol{N}^{r-1}.$$

因为 $(\lambda_i\boldsymbol{I})\boldsymbol{N}=\boldsymbol{N}(\lambda_i\boldsymbol{I})$, 故由命题 7.9可知

$$\begin{aligned}\mathrm{e}^{\boldsymbol{J}_i}&=\mathrm{e}^{\lambda_i\boldsymbol{I}+\boldsymbol{N}}=\mathrm{e}^{\lambda_i\boldsymbol{I}}\cdot\mathrm{e}^{\boldsymbol{N}}=\mathrm{e}^{\lambda_i}\cdot\mathrm{e}^{\boldsymbol{N}}\\&=\mathrm{e}^{\lambda_i}\boldsymbol{I}+\mathrm{e}^{\lambda_i}\boldsymbol{N}+\frac{1}{2!}\mathrm{e}^{\lambda_i}\boldsymbol{N}^2+\cdots+\frac{1}{(r-1)!}\mathrm{e}^{\lambda_i}\boldsymbol{N}^{r-1}.\end{aligned}$$

同理

$$\mathrm{e}^{t\boldsymbol{J}_i}=\mathrm{e}^{t\lambda_i}\cdot\mathrm{e}^{t\boldsymbol{N}}=\mathrm{e}^{t\lambda_i}\left[t\boldsymbol{I}+t\boldsymbol{N}+\frac{t}{2!}\boldsymbol{N}^2+\frac{t}{3!}\boldsymbol{N}^3+\cdots+\frac{t}{(r-1)!}\boldsymbol{N}^{r-1}\right]$$

$$
= \mathrm{e}^{t\lambda_i} \begin{pmatrix} 1 & t & \frac{1}{2!}t^2 & \frac{1}{3!}t^3 & \cdots & \frac{1}{(r-1)!}t^{r-1} \\ & 1 & t & \frac{1}{2!}t^2 & \cdots & \frac{1}{(r-2)!}t^{r-2} \\ & & 1 & t & \cdots & \frac{1}{(r-3)!}t^{r-3} \\ & & & \ddots & \ddots & \vdots \\ & & & & \ddots & t \\ & & & & & 1 \end{pmatrix}.
$$

又注意到

$$
\mathrm{e}^{tA} = \mathrm{e}^{\boldsymbol{P}(t\boldsymbol{J})\boldsymbol{P}^{-1}} = \boldsymbol{P}\mathrm{e}^{t\boldsymbol{J}}\boldsymbol{P}^{-1},
$$

$$
\mathrm{e}^{t\boldsymbol{J}} = \begin{pmatrix} \mathrm{e}^{t\boldsymbol{J}_1} & & \\ & \ddots & \\ & & \mathrm{e}^{t\boldsymbol{J}_k} \end{pmatrix}.
$$

于是将 $\mathrm{e}^{t\boldsymbol{J}_i}$ 的式子代入上面的式子即可求出 e^{tA}. □

命题 7.11 (矩阵三角函数的性质)

设 A 是 n 阶矩阵, 求证:

(1) $\sin^2 A + \cos^2 A = I_n$.

(2) $\sin 2A = 2\sin A \cos A$.

证明 由定理 7.21可知

$$
\cos A = \frac{1}{2}(e^{\mathrm{i}A} + e^{-\mathrm{i}A}), \quad \sin A = \frac{1}{2\mathrm{i}}(e^{\mathrm{i}A} - e^{-\mathrm{i}A}). \tag{7.29}
$$

由命题 7.7和命题 7.9可知

$$
e^{\mathrm{i}A}e^{-\mathrm{i}A} = e^{-\mathrm{i}A}e^{\mathrm{i}A} = e^{\mathrm{i}A-\mathrm{i}A} = I_n, (e^{\mathrm{i}A})^2 = e^{2\mathrm{i}A}, (e^{-\mathrm{i}A})^2 = e^{-2\mathrm{i}A}. \tag{7.30}
$$

(1) 由(7.29)和(7.30)式可得

$$
\sin^2 A + \cos^2 A = \frac{1}{4}(e^{2\mathrm{i}A} + 2I_n + e^{-2\mathrm{i}A}) - \frac{1}{4}(e^{2\mathrm{i}A} - 2I_n + e^{-2\mathrm{i}A}) = I_n.
$$

(2) 由(7.29)和(7.30)式可得

$$
2\sin A\cos A = \frac{1}{2}\left(e^{\mathrm{i}A} - e^{-\mathrm{i}A}\right)\left(e^{\mathrm{i}A} + e^{-\mathrm{i}A}\right) = \frac{1}{2}\left(e^{\mathrm{i}2A} - e^{-\mathrm{i}2A}\right) = \sin 2A.
$$

□

例题 7.10 计算 $\sin(e^{cI})$ 及 $\cos(e^{cI})$, 其中 c 是非零常数.

解 由指数矩阵函数的定义可得

$$
\begin{aligned}
e^{cI} &= I + \frac{1}{1!}(cI) + \frac{1}{2!}(cI)^2 + \frac{1}{3!}(cI)^3 + \cdots \\
&= \left(1 + \frac{1}{1!}c + \frac{1}{2!}c^2 + \frac{1}{3!}c^3 + \cdots\right) I = e^c I.
\end{aligned}
$$

因此

$$
\begin{aligned}
\sin(e^{cI}) = \sin(e^c I) &= e^c I - \frac{1}{3!}(e^c I)^3 + \frac{1}{5!}(e^c I)^5 - \frac{1}{7!}(e^c I)^7 + \cdots \\
&= \left(e^c - \frac{1}{3!}(e^c)^3 + \frac{1}{5!}(e^c)^5 - \frac{1}{7!}(e^c)^7 + \cdots\right) I = (\sin e^c)I, \\
\cos(e^{cI}) = \cos(e^c I) &= I - \frac{1}{2!}(e^c I)^2 + \frac{1}{4!}(e^c I)^4 - \frac{1}{6!}(e^c I)^6 + \cdots \\
&= \left(1 - \frac{1}{2!}(e^c)^2 + \frac{1}{4!}(e^c)^4 - \frac{1}{6!}(e^c)^6 + \cdots\right) I = (\cos e^c)I.
\end{aligned}
$$

□

命题 7.12

设 A 是 n 阶方阵, 证明:e^A 的行列式为 $e^{\mathrm{tr}(A)}$.

证明 设 A 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则由命题 6.4可知 e^A 的特征值为 $e^{\lambda_1}, e^{\lambda_2}, \cdots, e^{\lambda_n}$, 因此

$$|e^A| = e^{\lambda_1} e^{\lambda_2} \cdots e^{\lambda_n} = e^{\lambda_1+\lambda_2+\cdots+\lambda_n} = e^{\mathrm{tr}(A)}.$$

□

推论 7.15

求证: 对任一 n 阶方阵 A,e^A 总是非异阵.

证明 由命题 7.12可知 $|e^A| = e^{\mathrm{tr}(A)} \neq 0$, 从而 e^A 非异. 也可由命题 7.9得到

$$e^A e^{-A} = e^{A-A} = I_n,$$

于是 e^A 非异且 $(e^A)^{-1} = e^{-A}$. □ □

7.9 矩阵相似的全系不变量

7.9.1 矩阵相似的判定准则之一: 特征矩阵相抵

回顾定理 7.2中矩阵相似的充要条件.

命题 7.13

设 A, B 是数域 $\mathbb{F}$ 上的 n 阶矩阵,$\lambda I_n - A$ 相抵于 $\mathrm{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$,$\lambda I_n - B$ 相抵于 $\mathrm{diag}\{\ f_{i_1}(\lambda), f_{i_2}(\lambda), \cdots, f_{i_n}(\lambda)\ \}$, 其中 $f_{i_1}(\lambda)$, $f_{i_2}(\lambda), \cdots, f_{i_n}(\lambda)$ 是 $f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)$ 的一个排列. 求证:A 与 B 相似.

证明 对换 λ-矩阵 $\mathrm{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$ 的第 i, j 行, 再对换第 i, j 列, 可将 $f_i(\lambda)$ 与 $f_j(\lambda)$ 互换位置. 由于任一排列都可由若干次对换实现, 故 $\mathrm{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$ 相抵于 $\mathrm{diag}\{f_{i_1}(\lambda), f_{i_2}(\lambda), \cdots, f_{i_n}(\lambda)\}$, 于是 $\lambda I_n - A$ 相抵于 $\lambda I_n - B$, 从而 A 与 B 相似. □

例题 7.11 设 n 阶方阵 A, B, C, D 中 A, C 可逆, 求证: 存在可逆矩阵 P, Q, 使得 $A = PCQ$,$B = PDQ$ 的充要条件是 $\lambda A - B$ 与 $\lambda C - D$ 相抵.

证明 必要性由 $\lambda A - B = P(\lambda C - D)Q$ 即得. 下证充分性.

设 $\lambda A - B$ 与 $\lambda C - D$ 相抵, 则由 A, C 可逆知,$\lambda I_n - A^{-1}B$ 与 $\lambda I_n - C^{-1}D$ 相抵, 于是 $A^{-1}B$ 与 $C^{-1}D$ 相似. 设 Q 为可逆矩阵, 使得 $A^{-1}B = Q^{-1}(C^{-1}D)Q$, 令 $P = AQ^{-1}C^{-1}$, 则 P 可逆且 $A = PCQ$,$B = PDQ$. □

7.9.2 矩阵相似的判定准则二: 有相同的行列式因子组

回顾定理 7.5中矩阵相似的充要条件和 λ-矩阵的行列式因子相关定义和性质.

命题 7.14 (矩阵必与其转置相似)

求证: 任一 n 阶矩阵 A 都与它的转置 A' 相似.

证明 注意到 $(\lambda I_n - A)' = \lambda I_n - A'$, 并且行列式的值在转置下不改变, 故 $\lambda I_n - A$ 和 $\lambda I_n - A'$ 有相同的行列式因子组, 从而 A 和 A' 相似. □

命题 7.15

求证: 对任意的 $b \neq 0$,n 阶方阵 $\boldsymbol{A}(a,b)$ 均相互相似:

$$\boldsymbol{A}(a,b) = \begin{pmatrix} a & b & \cdots & b & b \\ & a & \ddots & \ddots & b \\ & & \ddots & \ddots & \vdots \\ & & & a & b \\ & & & & a \end{pmatrix}$$

证明 只要证明对任意的 $b \neq 0$,$\boldsymbol{A}(a,b)$ 的行列式因子组都一样即可. 显然 $D_n(\lambda) = (\lambda - a)^n$.$\lambda\boldsymbol{I}_n - \boldsymbol{A}(a,b)$ 的前 $n-1$ 行、前 $n-1$ 列构成的子式, 其值为 $(\lambda - a)^{n-1}$;$\lambda\boldsymbol{I}_n - \boldsymbol{A}(a,b)$ 的前 $n-1$ 行、后 $n-1$ 列构成的子式, 其值设为 $g(\lambda)$. 注意到 $g(a)$ 是 $n-1$ 阶上三角行列式, 主对角元素全为 $-b$, 从而 $g(a) = (-b)^{n-1} \neq 0$. 因此 $(\lambda - a)^{n-1}$ 与 $g(\lambda)$ 没有公共根, 故 $((\lambda - a)^{n-1}, g(\lambda)) = 1$, 于是 $D_{n-1}(\lambda) = 1$, 从而 $\boldsymbol{A}(a,b)$ 的行列式因子组为 $1, \cdots, 1, (\lambda - a)^n$, 结论得证. □

注

(1) 在上（下）三角矩阵（如 Jordan 块）或类上（下）三角矩阵（如友阵或 Frobenius 块）中, 若上（下）次对角线上的元素全部非零, 可以尝试计算行列式因子组. 对一般的矩阵（如数字矩阵）, 不建议计算行列式因子组, 推荐使用 λ-矩阵的初等变换计算法式, 得到不变因子组.

(2) 注意到 $\boldsymbol{A}(a,0) = a\boldsymbol{I}_n$ 的行列式因子组为 $D_i(\lambda) = (\lambda - a)^i (1 \leq i \leq n)$. 因此, 在求相似标准型的过程中, 注意千万不能使用摄动法!

7.9.3 矩阵相似的判定准则三: 有相同的不变因子组

回顾定理 7.7可知, 所有不变因子的乘积等于特征多项式, 整除关系下最大的那个不变因子等于极小多项式. 因此, 确定特征多项式和极小多项式可帮助确定不变因子组.

命题 7.16 (同阶幂零阵必相似)

设 A 是 n 阶 n 次幂零矩阵, 即 $A^n = O$ 但 $A^{n-1} \neq O$. 若 B 也是 n 阶 n 次幂零矩阵, 求证:A 相似于 B.

证明 显然 A 的极小多项式为 λ^n, 故 A 的不变因子组是 $1, \cdots, 1, \lambda^n$. 同理 B 的不变因子组也是 $1, \cdots, 1, \lambda^n$, 因此 A 和 B 相似. □

命题 7.17

设 A 为 n 阶矩阵, 证明以下 3 个结论等价:

(1) $A = cI_n$, 其中 c 为常数;

(2) A 的 $n-1$ 阶行列式因子是一个 $n-1$ 次多项式;

(3) A 的不变因子组中无常数.

证明 (1) $\Rightarrow$ (2): 显然成立.

(2) $\Rightarrow$ (3): 由于 A 的 n 阶行列式因子 $D_n(\lambda)$ 是一个 n 次多项式, 故 A 的最后一个不变因子 $d_n(\lambda) = D_n(\lambda)/D_{n-1}(\lambda)$ 是一个一次多项式, 设为 $\lambda - c$. 因为其他不变因子都要整除 $d_n(\lambda)$, 并且所有不变因子的乘积等于 n 阶行列式因子 $D_n(\lambda)$, 故 A 的不变因子组只能是 $\lambda - c, \lambda - c, \cdots, \lambda - c$.

(3) $\Rightarrow$ (1): 设 A 的不变因子组为 $d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)$, 则 $\deg d_i(\lambda) \geq 1$. 注意到 $d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda) = D_n(\lambda)$ 的次数为 n, 因此 $\deg d_i(\lambda) = 1$. 又 $d_i(\lambda) \mid d_n(\lambda)$, 故只能是 $d_1(\lambda) = d_2(\lambda) = \cdots = d_n(\lambda) = \lambda - c$. 因此 A 与 cI_n 有相同不变因子组, 从而它们相似, 即存在可逆矩阵 P, 使得 $A = P^{-1}(cI_n)P = cI_n$. 另外, 也可以利用 A 的极小多项式等于 $\lambda - c$ 或 A 的 Jordan 标准型来证明. □

命题 7.18

设 n 阶矩阵 A 的特征值全为 1, 求证: 对任意的正整数 k,A^k 与 A 相似.

注 证法一是用"三段论法"和极小多项式来证明的 (当然用行列式因子和几何重数替代也可以); 后面利用命题 7.45给出了第二种证法; 而命题 7.46 (当 $a = \pm 1$ 时) 给出了第三种证法.

证明 证法一: 由 A 的特征值全为 1 可知 A^k 的特征值也全为 1. 设 P 为可逆矩阵, 使得 $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(1),\cdots,J_{r_s}(1)\}$ 为 Jordan 标准型. 由于 $P^{-1}A^kP = (P^{-1}AP)^k = J^k$, 故只要证明 J^k 与 J 相似即可. 又因为 $J^k = \mathrm{diag}\{J_{r_1}(1)^k,\cdots,J_{r_s}(1)^k\}$, 故问题可进一步归结到每个 Jordan 块, 即只要证明 $J_{r_i}(1)^k$ 与 $J_{r_i}(1)$ 相似即可. 因此不妨设 $J = J_n(1)$ 只有一个 Jordan 块, 则 $J = I_n + J_0$, 其中 $J_0 = J_n(0)$ 是特征值为 0 的 n 阶 Jordan 块. 注意到

$$J^k = (I_n + J_0)^k = I_n + \mathrm{C}_k^1 J_0 + \mathrm{C}_k^2 J_0^2 + \cdots + J_0^k.$$

故 J^k 是一个上三角矩阵, 其主对角线上的元素全为 1, 上次对角线上的元素全为 k, 从而它的特征多项式为 $(\lambda-1)^n$. 为了确定它的极小多项式, 我们可进行如下计算:

$$(J^k - I_n)^{n-1} = (\mathrm{C}_k^1 J_0 + \mathrm{C}_k^2 J_0^2 + \cdots + J_0^k)^{n-1} = k^{n-1}J_0^{n-1} \neq O.$$

于是 J^k 的极小多项式为 $(\lambda-1)^n$, 其不变因子组为 $1,\cdots,1,(\lambda-1)^n$. 因此 J^k 与 J 有相同的不变因子, 从而 J^k 与 J 相似.

证法二: 显然 A^k 的特征值也全为 1. 注意到

$$(A^k - I_n)^l = (A - I_n)^l (A^{k-1} + A^{k-2} + \cdots + I_n)^l, \quad l \geq 1.$$

由于 $A^{k-1} + A^{k-2} + \cdots + I_n$ 的特征值全为 k, 故为可逆矩阵, 从而 $\mathrm{r}((A^k - I_n)^l) = \mathrm{r}((A - I_n)^l)$ 对任意的正整数 l 都成立. 由命题 7.45可知, A^k 与 A 相似. □

命题 7.19

设 n 阶矩阵 A 的特征值全为 1 或 -1, 求证:A^{-1} 与 A 相似.

注 证法一是用"三段论法"和极小多项式来证明的 (当然用行列式因子和几何重数替代也可以); 后面利用命题 7.45给出了第二种证法; 而命题 7.46 (当 $a = \pm 1$ 时) 给出了第三种证法.

证明 证法一: 设 P 为可逆矩阵, 使得 $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_s}(\lambda_s)\}$ 为 Jordan 标准型, 其中 $\lambda_i = \pm 1$. 由于 $P^{-1}A^{-1}P = (P^{-1}AP)^{-1} = J^{-1}$, 故只要证明 J^{-1} 与 J 相似即可. 又因为 $J^{-1} = \mathrm{diag}\{J_{r_1}(\lambda_1)^{-1},\cdots,J_{r_s}(\lambda_s)^{-1}\}$, 故问题可进一步归结到每个 Jordan 块, 即只要证明 $J_{r_i}(\lambda_i)^{-1}$ 与 $J_{r_i}(\lambda_i)$ 相似即可. 因此不妨设 $J = J_n(\lambda_0)$ 只有一个 Jordan 块, 则 $J = \lambda_0 I_n + J_0$, 其中 $\lambda_0 = \pm 1$,$J_0 = J_n(0)$ 是特征值为 0 的 n 阶 Jordan 块. 注意到

$$\lambda_0^n I_n = (\lambda_0 I_n)^n - (-J_0)^n = (\lambda_0 I_n + J_0)(\lambda_0^{n-1} I_n - \lambda_0^{n-2} J_0 + \cdots + (-1)^{n-1} J_0^{n-1}).$$

以及 $\lambda_0^{-1} = \lambda_0$, 故可得

$$J^{-1} = (\lambda_0 I_n + J_0)^{-1} = \lambda_0 I_n - \lambda_0^2 J_0 + \cdots + (-1)^{n-1}\lambda_0^n J_0^{n-1}.$$

因此 J^{-1} 是一个上三角矩阵, 其主对角线上的元素全为 λ_0, 上次对角线上的元素全为 $-\lambda_0^2$, 从而它的特征多项式为 $(\lambda - \lambda_0)^n$. 为了确定它的极小多项式, 我们可进行如下计算:

$$(J^{-1} - \lambda_0 I)^{n-1} = (-\lambda_0^2 J_0 + \cdots + (-1)^{n-1}\lambda_0^n J_0^{n-1})^{n-1} = (-1)^{n-1} J_0^{n-1} \neq O$$

于是 J^{-1} 的极小多项式为 $(\lambda - \lambda_0)^n$, 其不变因子组为 $1,\cdots,1,(\lambda - \lambda_0)^n$. 因此 J^{-1} 与 J 有相同的不变因子组, 从而 J^{-1} 与 J 相似.

证法二: 显然 A^{-1} 的特征值也全为 1 或 -1. 设 $\lambda_0 = \pm 1$, 则由 A 可逆以及 $(A^{-1} - \lambda_0 I_n)^l = (-\lambda_0)^l A^{-l}(A - \lambda_0 I_n)^l$ 可得 $\mathrm{r}((A^{-1} - \lambda_0 I_n)^l) = \mathrm{r}((A - \lambda_0 I_n)^l)$ 对任意的正整数 l 都成立. 由命题 7.45可知, A^{-1} 与 A 相似. □

7.9.4 矩阵相似的判定准则四: 有相同的初等因子组

定义 7.14 (准素因子)

设 $f(\lambda)$ 为数域 $\mathbb{K}$ 上的多项式,$p(\lambda)$ 是 $\mathbb{K}$ 上的首一不可约多项式, 若存在正整数 k, 使得 $p(\lambda)^k \mid f(\lambda)$, 但 $p(\lambda)^{k+1} \nmid f(\lambda)$, 则称 $p(\lambda)^k$ 为 $f(\lambda)$ 的一个**准素因子**. 所有 $f(\lambda)$ 的准素因子称为 $f(\lambda)$ 的**准素因子组**.

事实上, 若设 $f(\lambda)$ 在 $\mathbb{K}$ 上的标准因式分解为

$$f(\lambda) = cP_1(\lambda)^{e_1}P_2(\lambda)^{e_2}\cdots P_t(\lambda)^{e_t}$$

其中 c 为非零常数,$P_i(\lambda)$ 为互异的首一不可约多项式,$e_i > 0(1 \le i \le t)$, 则 $f(\lambda)$ 的所有准素因子为 $P_1(\lambda)^{e_1}, P_2(\lambda)^{e_2}, \cdots, P_t(\lambda)^{e_t}$.

定理 7.22 (λ-矩阵和初等因子的基本性质)

(1) 设 $f(\lambda), g(\lambda)$ 是数域 $\mathbb{K}$ 上的首一多项式,$d(\lambda) = (f(\lambda), g(\lambda))$,$m(\lambda) = [f(\lambda), g(\lambda)]$ 分别是 $f(\lambda)$ 和 $g(\lambda)$ 的最大公因式和最小公倍式, 证明下列 λ-矩阵相抵:

$$\begin{pmatrix} f(\lambda) & 0 \\ 0 & g(\lambda) \end{pmatrix}, \begin{pmatrix} g(\lambda) & 0 \\ 0 & f(\lambda) \end{pmatrix}, \begin{pmatrix} d(\lambda) & 0 \\ 0 & m(\lambda) \end{pmatrix}$$

(2) 设 A 是数域 $\mathbb{K}$ 上的 n 阶矩阵, 其特征矩阵 $\lambda I_n - A$ 经过初等变换可化为对角矩阵 $\mathrm{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$, 其中 $f_i(\lambda)$ 是 $\mathbb{K}$ 上的首一多项式. 求证: 矩阵 A 的初等因子组等于所有 $f_i(\lambda)$ 的准素因子组.

笔记 由 (2) 可知, 矩阵 A 的初等因子组就是 A 的所有不变因子的准素因子组. 实际上,(2) 就是引理 7.8的一个推广.

证明

(1) 由已知, 存在多项式 $u(\lambda), v(\lambda)$, 使得 $f(\lambda)u(\lambda) + g(\lambda)v(\lambda) = d(\lambda)$. 设 $f(\lambda) = d(\lambda)h(\lambda)$, 则 $m(\lambda) = g(\lambda)h(\lambda)$. 作下列 λ-矩阵的初等变换:

$$\begin{aligned}\begin{pmatrix} f(\lambda) & 0 \\ 0 & g(\lambda) \end{pmatrix} &\to \begin{pmatrix} f(\lambda) & 0 \\ f(\lambda)u(\lambda) & g(\lambda) \end{pmatrix} \to \begin{pmatrix} f(\lambda) & 0 \\ f(\lambda)u(\lambda) + g(\lambda)v(\lambda) & g(\lambda) \end{pmatrix} = \begin{pmatrix} f(\lambda) & 0 \\ d(\lambda) & g(\lambda) \end{pmatrix} \\ &\to \begin{pmatrix} 0 & -g(\lambda)h(\lambda) \\ d(\lambda) & g(\lambda) \end{pmatrix} \to \begin{pmatrix} 0 & g(\lambda)h(\lambda) \\ d(\lambda) & 0 \end{pmatrix} \to \begin{pmatrix} d(\lambda) & 0 \\ 0 & m(\lambda) \end{pmatrix}.\end{aligned}$$

另一结论同理可得.

(2) 对任意的 $i < j$, 以下操作记为 $O(i,j)$: 设 $d(\lambda) = (f_i(\lambda), f_j(\lambda))$,$m(\lambda) = [f_i(\lambda), f_j(\lambda)]$ 分别是 $f_i(\lambda)$ 和 $f_j(\lambda)$ 的最大公因式和最小公倍式, 则用 $d(\lambda)$ 替代 $f_i(\lambda)$, 用 $m(\lambda)$ 替代 $f_j(\lambda)$. 我们先证明, 操作 $O(i,j)$ 可通过 λ-矩阵的初等变换来实现, 并且前后两个对角矩阵, 即 diag $\{ f_1(\lambda),\cdots,f_i(\lambda),\cdots,f_j(\lambda),\cdots,f_n(\lambda) \}$ 与 diag $\{ f_1(\lambda),\cdots, d(\lambda), \cdots, m(\lambda), \cdots, f_n(\lambda)\}$ 有相同的准素因子组.

由 (1) 即知 $O(i,j)$ 是 λ-矩阵的相抵变换. 设 $f_i(\lambda), f_j(\lambda)$ 的公共因式分解为

$$f_i(\lambda) = P_1(\lambda)^{e_{i1}}P_2(\lambda)^{e_{i2}}\cdots P_t(\lambda)^{e_{it}}, \quad f_j(\lambda) = P_1(\lambda)^{e_{j1}}P_2(\lambda)^{e_{j2}}\cdots P_t(\lambda)^{e_{jt}}$$

其中 $P_i(\lambda)$ 为互异的首一不可约多项式,$e_{ik} \ge 0$,$e_{jk} \ge 0(1 \le k \le t)$, 令 $r_k = \min\{e_{ik}, e_{jk}\}$,$s_k = \max\{e_{ik}, e_{jk}\}$, 则有

$$d(\lambda) = P_1(\lambda)^{r_1}P_2(\lambda)^{r_2}\cdots P_t(\lambda)^{r_t}, \quad m(\lambda) = P_1(\lambda)^{s_1}P_2(\lambda)^{s_2}\cdots P_t(\lambda)^{s_t}$$

显然 $\{f_i(\lambda), f_j(\lambda)\}$ 和 $\{d(\lambda), m(\lambda)\}$ 有相同的准素因子组, 因此 $O(i,j)$ 操作前后的两个对角矩阵也有相同的准素因子组.

对对角矩阵 $\mathrm{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$ 依次实施操作 $O(1,j)(2 \le j \le n)$, 则得到对角矩阵的第 $(1,1)$ 元素的所有不可约因式的幂在主对角元素中都是最小的; 然后依次操作 $O(2,j)(3 \le j \le n)$;$\cdots$; 最后操作 $O(n-1,n)$,

可得一个对角矩阵 $\Lambda = \text{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)\}$. 由操作的性质可知,$\Lambda$ 满足 $d_i(\lambda) \mid d_{i+1}(\lambda)(1 \le i \le n-1)$, 因此 Λ 就是矩阵 A 的法式. 又因为对角矩阵 $\text{diag}\{f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)\}$ 与法式有相同的准素因子组, 故所有 $f_i(\lambda)$ 的准素因子组就是矩阵 A 的初等因子组.

□

命题 7.20

设 $A = \text{diag}\{A_1, A_2, \cdots, A_k\}$ 为分块对角矩阵, 求证:A 的初等因子组等于 $A_i(1 \le i \le k)$ 的初等因子组的无交并集. 又若交换各块的位置, 则所得的矩阵仍和 A 相似.

证明 □

显然 $\lambda I - A$ 也是一个分块对角矩阵, 用 λ-矩阵的初等变换将每一块化为法式, 则由λ-矩阵和初等因子的基本性质 (2)可知,A 的初等因子组就是所有各块的初等因子组的无交并集. 又交换 A 的各块并不改变 A 的初等因子组, 因此所得之矩阵仍和 A 相似.

7.10 有理标准型的几何与应用

回顾有理标准型和循环子空间相关理论.

命题 7.21

设 n 阶矩阵 A 有 n 个不同的特征值, 求证:A 的特征多项式和极小多项式相等.

证明 证法一: 设 A 的 n 个不同的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则由推论 6.8可知, 特征多项式 $f(\lambda)$ 和极小多项式 $m(\lambda)$ 有相同的根（不计重数）, 因此

$$f(\lambda) = m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$$

证法二: 由于 A 有 n 个不同的特征值, 故 A 相似于对角矩阵. 又因为相似矩阵有相同的特征多项式和极小多项式, 所以只要对对角矩阵证明此结论即可. 设 $A = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 则 $\lambda I_n - A = \text{diag}\{\lambda - \lambda_1, \lambda - \lambda_2, \cdots, \lambda - \lambda_n\}$, 这是一个主对角元素两两互素的对角矩阵, 由$\lambda$-矩阵和初等因子的基本性质 (1)以及数学归纳法可知其法式为 $\text{diag}\{1, \cdots, 1, (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)\}$. 因此,$A$ 的特征多项式和极小多项式相等. □

命题 7.22

设 n 阶矩阵 A 有 n 个不同的特征值, 且特征值 λ_i 对应的特征向量为 α_i, 由推论 6.5可知 $\{\alpha_1, \cdots, \alpha_n\}$ 为 $\mathbb{C}^n$ 的一组基. 则 $\alpha = \alpha_1 + \cdots + \alpha_n$ 是 A 的循环空间 $\mathbb{C}^n$ 的循环向量, 即 $\mathbb{C}^n = L(\alpha, A\alpha, \cdots, A^{n-1}\alpha) = C(A, \alpha)$ 为循环空间,α 是循环向量.

证明 事实上, 由 $A^k\alpha = \lambda_1^k\alpha_1 + \cdots + \lambda_n^k\alpha_n$, 利用 Vandermonde 行列式容易证明 $\{\alpha, A\alpha, \cdots, A^{n-1}\alpha\}$ 是 $\mathbb{C}^n$ 的一组基, 从而 $\mathbb{C}^n = L(\alpha, A\alpha, \cdots, A^{n-1}\alpha) = C(A, \alpha)$ 为循环空间,α 是循环向量. □

命题 7.23

设数域 $\mathbb{K}$ 上的 n 阶矩阵 A 的特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda)$, 其中 $P_i(\lambda)(1 \le i \le k)$ 是 $\mathbb{K}$ 上互异的首一不可约多项式. 求证:A 的有理标准型只有一个 Frobenius 块, 并且 A 在复数域上可对角化.

注 我们也可以利用定理 7.12和初等因子证明这个命题. 若利用不变因子在基域扩张下的不变性, 则这个命题也可由命题 7.21得到.

证明 设 A 的不变因子组为 $d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)$, 其中 $d_i(\lambda) \mid d_{i+1}(\lambda), i = 1, 2, \cdots, n-1$, 则有

$$f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda) = d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda) \tag{7.31}$$

由于 $P_i(\lambda)$ 是不可约多项式, 故存在某个 j, 使得 $P_i(\lambda) \mid d_j(\lambda)$, 否则, 由不可约多项式的基本性质 (1)可知 $(P_i(\lambda), d_j(\lambda)) = 1, j = 1, 2, \cdots, n$. 再由互素多项式和最大公因式的基本性质 (5)可知 $(P_i(\lambda), d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda)) = 1$, 这

与(7.31)矛盾! 从而 $P_i(\lambda) \mid d_n(\lambda)(1 \le i \le k)$. 由互素多项式和最大公因式的基本性质 (1)可知,$P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda) \mid d_n(\lambda)$, 因此只能是 $d_1(\lambda) = \cdots = d_{n-1}(\lambda) = 1, d_n(\lambda) = f(\lambda)$, 从而 A 的有理标准型只有一个 Frobenius 块. 由于特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda)$ 在 $\mathbb{K}$ 上无重因式, 故 $(f(\lambda), f'(\lambda)) = 1$, 从而 $f(\lambda)$ 在复数域上无重根, 即 A 有 n 个不同的特征值, 于是 A 在复数域上可对角化. □

推论 7.16

设数域 $\mathbb{K}$ 上的 n 阶矩阵 A 的特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda)$, 其中 $P_i(\lambda)(1 \le i \le k)$ 是 $\mathbb{K}$ 上互异的首一不可约多项式. 并且 α_i 为线性方程组 $P_i(A)x = 0$ 的非零解, 则 $\alpha = \alpha_1 + \cdots + \alpha_k$ 是 A 的循环空间 $\mathbb{K}^n$ 的循环向量.

证明 由命题 7.23及定理 7.10可知 $\mathbb{K}^n$ 就是一个循环空间.(未完成证明) □

命题 7.24

设 φ 是数域 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换,φ 的特征多项式为 $f(\lambda)$, 证明以下 3 个结论等价:

(1) V 只有平凡的 φ-不变子空间;

(2) V 中任一非零向量都是循环向量, 使 V 成为循环空间;

(3) $f(\lambda)$ 是 $\mathbb{K}$ 上的不可约多项式.

证明 (1) $\Rightarrow$ (2): 任取 V 中非零向量 α, 则循环子空间 $C(\varphi, \alpha)$ 是非零 φ-不变子空间. 由于 V 只有平凡的 φ-不变子空间, 故 $C(\varphi, \alpha) = V$, 即 V 中任一非零向量都是循环向量, 使 V 成为循环空间.

(2) $\Rightarrow$ (3): 用反证法, 假设 $f(\lambda) = g(\lambda)h(\lambda)$, 其中 $g(\lambda), h(\lambda)$ 是 $\mathbb{K}$ 上次数小于 n 的首一多项式. 由 Cayley - Hamilton 定理可知 $\mathbf{0} = f(\varphi) = g(\varphi)h(\varphi)$, 故由命题 4.9(1)的逆否命题可知 $g(\varphi), h(\varphi)$ 中至少有一个是奇异 (不可逆/非双射) 线性变换, 不妨设为 $g(\varphi)$, 由推论 4.6可知 $\operatorname{Ker} g(\varphi) \neq 0$. 任取 $\operatorname{Ker} g(\varphi)$ 中的非零向量 α, 设 $\deg g(\lambda) = r$, 则不妨设

$$g(\varphi) = a_r\varphi^r + a_{r-1}\varphi^{r-1} + \cdots + a_1, \quad \text{其中 } a_r \neq 0.$$

由 $\alpha \in \ker g(\varphi)$ 可知

$$g(\varphi)(\alpha) = a_r\varphi^r(\alpha) + a_{r-1}\varphi^{r-1}(\alpha) + \cdots + a_1\alpha = 0.$$

于是

$$\varphi^r(\alpha) = -\frac{a_{r-1}}{a_r}\varphi^{r-1}(\alpha) - \cdots - \frac{a_1}{a_r}\alpha. \tag{7.32}$$

假设对 $k \ge r$ 且 $k \in \mathbb{N}$, 成立 $\varphi^k(\alpha)$ 可由 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 线性表示, 则对(7.32)式两边同时作用 φ^{k-r+1} 可得

$$\varphi^{k+1}(\alpha) = -\frac{a_{r-1}}{a_r}\varphi^k(\alpha) - \cdots - \frac{a_1}{a_r}\varphi^{k-r+1}(\alpha).$$

于是由归纳假设可知,$\varphi^{k+1}(\alpha)$ 可由 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 线性表示. 故由数学归纳法可得, 对 $\forall k \ge r$ 且 $k \in \mathbb{N}$, 都有 $\varphi^k(\alpha)$ 可由 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 线性表示. 因此 $C(\varphi, \alpha) = L(\alpha, \varphi(\alpha), \cdots\cdots) = L(\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha))$, 其维数 $\le r < n$, 故 $C(\varphi, \alpha) \neq V$, 这与 V 中任一非零向量都是循环向量矛盾!

(3) $\Rightarrow$ (1): 用反证法, 假设存在非平凡的 φ-不变子空间 U,$\dim U = r$, 则 φ 在一组基下的表示矩阵为分块上三角矩阵 $M = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}$, 其中 $\boldsymbol{A}$ 是 $\varphi|_U$ 的表示矩阵. 于是特征多项式

$$f(\lambda) = |\lambda \boldsymbol{I}_V - \varphi| = |\lambda \boldsymbol{I}_n - M| = |\lambda \boldsymbol{I}_r - \boldsymbol{A}| \cdot |\lambda \boldsymbol{I}_{n-r} - \boldsymbol{B}|.$$

是两个低次多项式的乘积, 这与 $f(\lambda)$ 的不可约性矛盾! □

命题 7.25

设 φ 是数域 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换,φ 的极小多项式为 $m(\lambda)$. 证明:$m(\lambda)$ 是 $\mathbb{K}$ 上的不可约多项式的充要条件是 V 的任一非零 φ-不变子空间 U 必为如下形式:

$$U = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \cdots \oplus C(\varphi, \alpha_k)$$

并且 $\varphi|_{C(\varphi,\alpha_i)}$ 的极小多项式都是 $m(\lambda)$. 此时,$\varphi|_U$ 的极小多项式也是 $m(\lambda)$.

证明 **必要性**: 设 $\varphi|_U$ 的极小多项式为 $n(\lambda)$, 则 $m(\varphi|_U) = m(\varphi)|_U = \mathbf{0}$, 从而 $n(\lambda) \mid m(\lambda)$. 因为 $m(\lambda)$ 不可约, 所以 $n(\lambda) = m(\lambda)$. 又由于 $\varphi|_U$ 的所有不变因子都要整除 $m(\lambda)$ 且 $m(\lambda)$ 不可约, 故所有的非常数不变因子都等于 $m(\lambda)$. 最后, 由有理标准型的几何意义即得

$$U = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \cdots \oplus C(\varphi, \alpha_k).$$

并且 $\varphi|_{C(\varphi,\alpha_i)}$ 在基 $\{\alpha_i, \varphi(\alpha_i), \cdots, \varphi^{r_i-1}(\alpha_i)\}$ 下的表示矩阵就是友阵 $C(m(\lambda))$, 其中 $r_i = \dim C(\varphi, \alpha_i)$. 于是 $\varphi|_{C(\varphi,\alpha_i)}$ 的极小多项式就是其表示矩阵 $C(m(\lambda))$ 的极小多项式. 又由引理 7.5可知,$C(m(\lambda))$ 的极小多项式就是 $m(\lambda)$, 并且 $n(\lambda) = m(\lambda)$, 故结论得证.

充分性: 用反证法, 设 $m(\lambda) = g(\lambda)h(\lambda)$, 其中 $g(\lambda), h(\lambda)$ 是 $\mathbb{K}$ 上次数小于 $m(\lambda)$ 次数的首一多项式, 则 $\mathbf{0} = m(\varphi) = g(\varphi)h(\varphi)$, 故由命题 4.9(1)的逆否命题可知 $g(\varphi), h(\varphi)$ 中至少有一个是奇异线性变换, 不妨设为 $g(\varphi)$, 于是由推论 4.6可知 $\mathrm{Ker}\, g(\varphi) \neq 0$. 任取 $\mathrm{Ker}\, g(\varphi)$ 中的非零向量 α, 得到循环子空间 $U = C(\varphi, \alpha)$,

由 $g(\varphi)(\alpha) = \mathbf{0}$ 可知, 对 $\forall k \in \mathbb{N}$, 都有 $\varphi^k(g(\varphi)(\alpha)) = 0$. 从而对 $\forall \beta \in U = C(\varphi, \alpha)$, 存在不全为零的 a_i 使得

$$\begin{aligned}
g(\varphi)(\beta) &= g\left(a_1\alpha + a_2\varphi(\alpha) + a_3\varphi^2(\alpha) + \cdots\right) \\
&= a_1 g(\alpha) + a_2 g(\varphi(\alpha)) + a_3 g\left(\varphi^2(\alpha)\right) + \cdots \\
&= a_1 g(\alpha) + a_2 \varphi(g(\alpha)) + a_3 \varphi^2(g(\alpha)) + \cdots \\
&= 0 + 0 + 0 + \cdots = 0.
\end{aligned}$$

因此 $g(\varphi|_U) = g(\varphi)|_U = \mathbf{0}$, 于是 $\varphi|_U$ 的极小多项式 $m(\lambda)$ 整除 $g(\lambda)$, 从而其次数 $\leq \deg g(\lambda) < \deg m(\lambda)$, 这与条件矛盾! □

定理 7.23 (基于初等因子组的有理标准型)

设数域 $\mathbb{K}$ 上的 n 阶矩阵 A 的初等因子组为 $P_1(\lambda)^{r_1}, P_2(\lambda)^{r_2}, \cdots, P_k(\lambda)^{r_k}$, 证明:$A$ 相似于分块对角矩阵

$$\widetilde{\boldsymbol{F}} = \mathrm{diag}\{\boldsymbol{F}(P_1(\lambda)^{r_1}), \boldsymbol{F}(P_2(\lambda)^{r_2}), \cdots, \boldsymbol{F}(P_k(\lambda)^{r_k})\}$$

$$\widetilde{\boldsymbol{C}} = \mathrm{diag}\{\boldsymbol{C}(P_1(\lambda)^{r_1}), \boldsymbol{C}(P_2(\lambda)^{r_2}), \cdots, \boldsymbol{C}(P_k(\lambda)^{r_k})\}$$

称为 A 的**基于初等因子组的有理标准型**.

证明 由 Frobenius 块和友阵的性质可知,$\lambda \boldsymbol{I}_n - \widetilde{\boldsymbol{F}}$ 和 $\lambda \boldsymbol{I}_n - \widetilde{\boldsymbol{C}}$ 都相抵于

$$\mathrm{diag}\{1, \cdots, 1, P_1(\lambda)^{r_1}; 1, \cdots, 1, P_2(\lambda)^{r_2}; \cdots; 1, \cdots, 1, P_k(\lambda)^{r_k}\}$$

再由λ-矩阵和初等因子的基本性质 (2)可知,$\widetilde{\boldsymbol{F}}, \widetilde{\boldsymbol{C}}$ 与 A 有相同的初等因子组, 从而它们相似. □

定理 7.24

设 φ 是数域 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换,φ 的初等因子组为 $P_1(\lambda)^{r_1}, P_2(\lambda)^{r_2}, \cdots, P_k(\lambda)^{r_k}$. 证明: 存在 $\alpha_1, \alpha_2, \cdots, \alpha_k \in V$, 使得

$$V = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \cdots \oplus C(\varphi, \alpha_k)$$

证明 由基于初等因子组的有理标准型和定理 7.10即得. □

例题 7.12 求证: 存在 n 阶实方阵 A, 满足 $A^2 + 2A + 5I_n = O$ 的充要条件是 n 为偶数. 当 $n \geq 4$ 时, 验证满足上述条件的矩阵 A 有无限个不变子空间.

证明 **必要性**: 注意到 A 适合多项式 $g(\lambda)=\lambda^2+2\lambda+5$, 故 A 的极小多项式 $m(\lambda)\mid g(\lambda)$, 又因为 $g(\lambda)$ 在实数域上不可约, 故只能是 $m(\lambda)=g(\lambda)$. 同理可证 A 所有的非常数不变因子都等于 $g(\lambda)$, 从而 A 的不变因子组为 $1,\cdots,1,g(\lambda),\cdots,g(\lambda)$（$k$ 个 $g(\lambda)$）. 因此 A 的特征多项式 $f(\lambda)=g(\lambda)^k$, 于是 $n=\deg f(\lambda)=2k$ 为偶数.

充分性: 设 $n=2k$ 为偶数, 则由必要性的证明可知,A 的不变因子组为 $1,\cdots,1,g(\lambda),\cdots,g(\lambda)$（$k$ 个 $g(\lambda)$）. 可用有理标准型构造满足条件的矩阵:

$$A=\operatorname{diag}\left\{\begin{pmatrix}0&-5\\1&-2\end{pmatrix},\cdots,\begin{pmatrix}0&-5\\1&-2\end{pmatrix}\right\}\text{（}k\text{ 个二阶方阵）}\tag{7.33}$$

当 $n\geq 4$ 时, 设 $\{e_1,e_2,e_3,e_4\}$ 是前 4 个标准单位列向量, 则容易验证(7.33)式中的矩阵 A 满足 $Ae_1=e_2,Ae_3=e_4$, 于是构造循环子空间 $\{C_l:=C(A,e_1+le_3)=L(e_1+le_3,e_2+le_4),l\in\mathbb{R}\}$, 进一步容易验证循环子空间 $\{C_l:=C(A,e_1+le_3)=L(e_1+le_3,e_2+le_4),l\in\mathbb{R}\}$ 是两两互异的 A-不变子空间, 故 A 有无限个不变子空间. □

命题 7.26

设 A 是数域 $\mathbb{K}$ 上的 n 阶方阵, 求证:A 的极小多项式的次数小于等于 $\mathrm{r}(A)+1$.

证明 证法一: 设 A 的不变因子组为 $1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)$, 则极小多项式 $m(\lambda)=d_k(\lambda)$, 并且由定理 7.6可知 A 相似于 $\boldsymbol{F}=\operatorname{diag}\{\boldsymbol{F}(d_1(\lambda)),\cdots,\boldsymbol{F}(d_k(\lambda))\}$. 设 $\deg d_k(\lambda)=r$, 若 $d_k(0)\neq 0$, 则由Frobenius 块的基本性质 (1)可知 $\boldsymbol{F}(d_k(\lambda))$ 非异; 若 $d_k(0)=0$, 则由Frobenius 块的基本性质 (1)可知 $\boldsymbol{F}(d_k(\lambda))$ 奇异且右上角的 $r-1$ 阶子式非零, 从而秩为 $r-1$. 因此,$\mathrm{r}(A)=\mathrm{r}(\boldsymbol{F})\geq\mathrm{r}(\boldsymbol{F}(d_k(\lambda)))\geq r-1=\deg d_k(\lambda)-1$.

证法二: 从 A 的极小多项式 $m(\lambda)$ 分离出来的初等因子中, 形如 λ^r 的初等因子至多只有 1 个, 对应于零特征值的 Jordan 块 $J_r(0)$, 其余的初等因子对应于非零特征值的 Jordan 块. 因此 $\mathrm{r}(A)$ 大于等于这些 Jordan 块秩的和, 后者等于 $\deg m(\lambda)-1$ 或 $\deg m(\lambda)$. □

命题 7.27

设数域 $\mathbb{K}$ 上的 n 阶矩阵 A 的不变因子组是 $1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)$, 其中 $d_i(\lambda)$ 是非常数首一多项式,$d_i(\lambda)\mid d_{i+1}(\lambda)(1\leq i\leq k-1)$. 求证: 对 A 的任一特征值 λ_0,

$$\mathrm{r}(\lambda_0\boldsymbol{I}_n-A)=n-\sum_{i=1}^{k}\delta_{d_i(\lambda_0),0}$$

其中记号 $\delta_{a,b}$ 表示: 若 $a=b$, 取值为 1; 若 $a\neq b$, 取值为 0.

证明 证法一: 设 $\deg d_i(\lambda)=r_i$, 则由定理 7.6可知 A 相似于 $\boldsymbol{F}=\operatorname{diag}\{\boldsymbol{F}(d_1(\lambda)),\cdots,\boldsymbol{F}(d_k(\lambda))\}$, 而相似矩阵有相同的特征多项式, 故 $\lambda_0I_n-A=\lambda_0I_n-F$. 由Frobenius 块的基本性质 (2)可知 $|\lambda_0\boldsymbol{I}_{r_i}-\boldsymbol{F}(d_i(\lambda))|=d_i(\lambda_0)$. 若 $d_i(\lambda_0)\neq 0$, 则 $\lambda_0\boldsymbol{I}_{r_i}-\boldsymbol{F}(d_i(\lambda))$ 非异; 若 $d_i(\lambda_0)=0$, 则 $\lambda_0\boldsymbol{I}_{r_i}-\boldsymbol{F}(d_i(\lambda))$ 奇异且右上角的 r_i-1 阶子式非零, 从而秩为 r_i-1. 因此,

$$\begin{aligned}\mathrm{r}(\lambda_0\boldsymbol{I}_n-A)=\mathrm{r}(\lambda_0\boldsymbol{I}_n-\boldsymbol{F})&=\sum_{i=1}^{k}\mathrm{r}(\lambda_0\boldsymbol{I}_{r_i}-\boldsymbol{F}(d_i(\lambda)))\\&=\sum_{i=1}^{k}(r_i-\delta_{d_i(\lambda_0),0})=n-\sum_{i=1}^{k}\delta_{d_i(\lambda_0),0}\end{aligned}$$

证法二: 由定理 7.6可知存在可逆 λ-矩阵 $\boldsymbol{P}(\lambda),\boldsymbol{Q}(\lambda)$, 使得

$$\boldsymbol{P}(\lambda)(\lambda\boldsymbol{I}_n-A)\boldsymbol{Q}(\lambda)=\operatorname{diag}\{1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)\}$$

在上式中令 $\lambda=\lambda_0$, 注意到 $\boldsymbol{P}(\lambda_0),\boldsymbol{Q}(\lambda_0)$ 是 $\mathbb{K}$ 上的可逆矩阵, 故 $\lambda_0\boldsymbol{I}_n-A$ 相抵于 $\operatorname{diag}\{1,\cdots,1,d_1(\lambda_0),\cdots,d_k(\lambda_0)\}$, 于是 $\mathrm{r}(\lambda_0\boldsymbol{I}_n-A)$ 等于 n 减去等于零的 $d_i(\lambda_0)$ 的个数, 从而结论得证. □

命题 7.28

设 A 是数域 $\mathbb{K}$ 上的 n 阶矩阵, 求证: 若 $\mathrm{tr}(A)=0$, 则 A 相似于一个 $\mathbb{K}$ 上主对角元全为零的矩阵.

证明 对阶数进行归纳. 当 $n=1$ 时,$A=O$, 结论显然成立. 设阶数小于 n 时结论成立, 现证 n 阶的情形. 由于题目的

条件和结论在相似关系下不改变，故不妨从一开始就假设 A 是有理标准型

$$\boldsymbol{F} = \mathrm{diag}\{\boldsymbol{F}(d_1(\lambda)), \cdots, \boldsymbol{F}(d_k(\lambda))\}$$

其中 $d_i(\lambda)$ 是 A 的非常数不变因子,$d_i(\lambda) \mid d_{i+1}(\lambda)(1 \le i \le k-1)$,$\deg d_i(\lambda) = r_i$. 若 r_i 都为 1, 则 $d_1(\lambda) = \cdots = d_n(\lambda) = \lambda - c$, 从而 $A = c\boldsymbol{I}_n$. 又 $\mathrm{tr}(A) = 0$, 故 $c = 0$, 从而 $A = O$, 结论成立. 以下假设存在某个 $r_i > 1$, 将第 $(1,1)$ 分块与第 (i,i) 分块对换, 这是一个相似变换, 此时矩阵的第 $(1,1)$ 元为零, 故不妨设 A 的第 $(1,1)$ 元为零. 注意到矩阵 $A = \begin{pmatrix} 0 & \boldsymbol{\alpha}' \\ \boldsymbol{\beta} & \boldsymbol{B} \end{pmatrix}$, 其中 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{K}^{n-1}$,$\boldsymbol{B} \in M_{n-1}(\mathbb{K})$,$\mathrm{tr}(\boldsymbol{B}) = 0$. 由归纳假设, 存在 $\mathbb{K}$ 上的 $n-1$ 阶非异阵 $\boldsymbol{Q}$, 使得 $\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}$ 的主对角元全为零, 令 $\boldsymbol{P} = \begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Q} \end{pmatrix}$ 为 $\mathbb{K}$ 上的 n 阶非异阵, 则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 0 & \boldsymbol{\alpha}'\boldsymbol{Q} \\ \boldsymbol{Q}^{-1}\boldsymbol{\beta} & \boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q} \end{pmatrix}$ 的主对角元全为零, 结论得证. □

命题 7.29

设 C 是数域 $\mathbb{K}$ 上的 n 阶矩阵, 求证: 存在 $\mathbb{K}$ 上的 n 阶矩阵 A, B, 使得 $AB - BA = C$ 的充要条件是 $\mathrm{tr}(C) = 0$. ♠

证明 必要性由矩阵迹的线性和交换性即得, 下证充分性. 由于题目的条件和结论在同时相似变换 $A \mapsto P^{-1}AP$,$B \mapsto P^{-1}BP$,$C \mapsto P^{-1}CP$ 下不改变, 故由命题 7.28不妨从一开始就假设 $C = (c_{ij})$ 的主对角元 $c_{ii} = 0(1 \le i \le n)$. 取定 $A = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ 为 $\mathbb{K}$ 上的主对角元互异的对角矩阵. 设 $B = (x_{ij})$, 则 $AB - BA = C$ 等价于方程 $\lambda_i x_{ij} - \lambda_j x_{ij} = c_{ij}$. 当 $i = j$ 时, 上式恒成立, 故 x_{ii} 可任取. 当 $i \ne j$ 时,$x_{ij} = \dfrac{c_{ij}}{\lambda_i - \lambda_j}$ 被唯一确定. 因此, 一定存在 $\mathbb{K}$ 上的矩阵 A, B, 使得 $AB - BA = C$ 成立. □

7.11 乘法交换性诱导的多项式表示

定理 7.25

设 φ 是数域 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换, 则对 V 上任一与 φ 乘法可交换的线性变换 ψ, 都存在不超过 $n-1$ 次的多项式 $g(x) \in \mathbb{K}[x]$, 使得 $\psi = g(\varphi)$ 成立的充要条件是 φ 的极小多项式等于其特征多项式. ♡

注 本题充分性证明的关键点是:$V = C(\varphi, e_1)$ 是一个循环空间, 循环向量 e_1 经过 φ 的 $n-1$ 次作用, 生成了 V 的一组基 $\{e_1, e_2, \cdots, e_n\}$. 因此, 只要验证了 ψ 和 $g(\varphi)$ 在循环向量 e_1 上的取值相同, 那么由 φ, ψ 的乘法交换性可知 ψ 和 $g(\varphi)$ 在上述基上的取值也相同, 从而它们必相等.

证明 **充分性**: 设 φ 的极小多项式等于其特征多项式 $f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$, 则 φ 只有一个非常数不变因子. 由有理标准型理论, 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友阵

$$C(f(\lambda)) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}$$

即有

$$\varphi(e_1) = e_2, \varphi(e_2) = e_3, \cdots, \varphi(e_{n-1}) = e_n, \varphi(e_n) = -a_n e_1 - a_{n-1}e_2 - \cdots - a_1 e_n$$

任取 V 上满足 $\varphi\psi = \psi\varphi$ 的线性变换 ψ, 设

$$\psi(e_1) = b_n e_1 + b_{n-1}e_2 + \cdots + b_1 e_n \tag{7.34}$$

令 $g(x) = b_1 x^{n-1} + \cdots + b_{n-1}x + b_n$, 我们来证明:$\psi = g(\varphi)$. 首先由 $e_k = \varphi^{k-1}(e_1)(k \ge 2)$ 以及 (7.34) 式可知 $\psi(e_1) = g(\varphi)(e_1)$ 成立. 其次由 φ, ψ 乘法可交换, 故对任意的 $e_k(k \ge 2)$ 有

$$\psi(e_k) = \psi(\varphi^{k-1}(e_1)) = \varphi^{k-1}(\psi(e_1)) = \varphi^{k-1}(g(\varphi)(e_1))$$

$$= g(\varphi)(\varphi^{k-1}(e_1)) = g(\varphi)(e_k)$$

最后, 注意到 ψ 与 $g(\varphi)$ 在基向量 $\{e_1, e_2, \cdots, e_n\}$ 上的取值都相等, 故由线性扩张定理可知 $\psi = g(\varphi)$ 成立.

必要性: 设 φ 的不变因子组为 $1, \cdots, 1, d_1(\lambda), \cdots, d_k(\lambda)$, 其中 $d_i(\lambda)$ 为非常数首一多项式,$d_i(\lambda) \mid d_{i+1}(\lambda)(1 \le i \le k-1)$, 则 φ 的有理标准型 $\boldsymbol{F} = \mathrm{diag}\{\boldsymbol{F}_1, \boldsymbol{F}_2, \cdots, \boldsymbol{F}_k\}$, 其中 $\boldsymbol{F}_i = \boldsymbol{F}(d_i(\lambda))$ 为 n_i 阶矩阵. 若 φ 的极小多项式不等于其特征多项式, 则 $k \ge 2$. 构造分块对角矩阵

$$\boldsymbol{B} = \mathrm{diag}\{\boldsymbol{I}_{n_1}, \boldsymbol{O}_{n_2}, \cdots, \boldsymbol{O}_{n_k}\}$$

显然 $\boldsymbol{BF} = \boldsymbol{FB}$. 用反证法, 若存在多项式 $g(x)$, 使得 $\boldsymbol{B} = g(\boldsymbol{F})$, 即

$$\boldsymbol{B} = \mathrm{diag}\{g(\boldsymbol{F}_1), g(\boldsymbol{F}_2), \cdots, g(\boldsymbol{F}_k)\}$$

则 $g(\boldsymbol{F}_1) = \boldsymbol{I}_{n_1}$,$g(\boldsymbol{F}_i) = \boldsymbol{O}(i \ge 2)$. 由引理 7.5可知 $d_k(\lambda)$ 是 $\boldsymbol{F}_k$ 的极小多项式（也是特征多项式）, 故 $d_k(\lambda) \mid g(\lambda)$, 从而 $d_1(\lambda) \mid g(\lambda)$, 于是 $g(\boldsymbol{F}_1) = \boldsymbol{O}$, 矛盾! 因此 $\boldsymbol{B}$ 不能表示为 $\boldsymbol{F}$ 的多项式, 从而由 $\boldsymbol{B}$ 定义的线性变换 ψ 与 φ 可交换, 但是不能表示为 φ 的多项式, 矛盾! □

推论 7.17

设 φ 是数域 $\mathbb{K}$ 上 n 维线性空间 V 上的线性变换,$\mathbb{K}[\varphi] = \{f(\varphi) \mid f(x) \in \mathbb{K}[x]\}$,$C(\varphi) = \{\psi \in \mathcal{L}(V) \mid \varphi\psi = \psi\varphi\}$, 则 V 是关于 φ 的循环空间的充要条件是 $C(\varphi) = \mathbb{K}[\varphi]$. 此时,$C(\varphi)$ 的一组基为 $\{\boldsymbol{I}_V, \varphi, \cdots, \varphi^{n-1}\}$.

笔记 定理 7.12证明了: 线性变换 φ 的极小多项式等于其特征多项式当且仅当 V 是关于 φ 的循环空间. 因此, 作为定理 7.25的推论, 我们给出了循环空间的另一刻画.

证明 由定理 7.25及 $f(\varphi)$ 全都与 φ 可交换以及定理 7.12立得. □

例题 7.13 设数域 $\mathbb{K}$ 上的 n 阶矩阵 A 的特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda)\cdots P_k(\lambda)$, 其中 $P_i(\lambda)(1 \le i \le k)$ 是 $\mathbb{K}$ 上互异的首一不可约多项式. 设 $\mathbb{K}$ 上的 n 阶矩阵 B 满足 $AB = BA$, 求证: 存在 $\mathbb{K}$ 上次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $B = f(A)$.

证明 由定理 7.12可知,$\mathbb{K}^n$ 是关于 A 的循环空间, 再由定理 7.25即得结论. □

例题 7.14 设 A 是数域 $\mathbb{K}$ 上的 2 阶矩阵, 试求 $C(A) = \{X \in M_2(\mathbb{K}) \mid AX = XA\}$.

证明 若 A 的极小多项式等于特征多项式, 则由定理 7.25可知 $C(A) = \mathbb{K}[A]$. 若极小多项式不等于特征多项式, 则极小多项式必为一次多项式 $x - c$, 从而 $A = c\boldsymbol{I}_2$, 于是 $C(A) = M_2(\mathbb{K})$. □

例题 7.15 设数域 $\mathbb{K}$ 上的 n 阶矩阵

$$A = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & * & & \ddots & b_{n-1} \\ & & & & a_n \end{pmatrix}$$

其中 $b_1, \cdots, b_{n-1}$ 均不为零. 记 $C(A) = \{X \in M_n(\mathbb{K}) \mid AX = XA\}$, 证明: 线性空间 $C(A)$ 的一组基为 $\{\boldsymbol{I}_n, A, \cdots, A^{n-1}\}$.

证明 题目中的 A 是类下三角矩阵, 上次对角元全部非零, 比如 Frobenius 块、Jordan 块和三对角矩阵都满足这样的特点. 考虑特征矩阵 $\lambda\boldsymbol{I}_n - A$ 的前 $n-1$ 行、后 $n-1$ 列构成的下三角行列式, 其值为 $(-1)^{n-1}b_1\cdots b_{n-1} \ne 0$, 于是 A 的第 $n-1$ 个行列式因子为 $d_{n-1} = 1$, 故 A 的行列式因子组为 $1, \cdots, 1, f(\lambda)$, 从而 $\mathbb{K}^n$ 是关于 A 的循环空间, 再由定理 7.25即得结论. □

命题 7.30

设有 n 阶分块对角矩阵

$$A=\begin{pmatrix}\boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_k\end{pmatrix},\quad B=\begin{pmatrix}\boldsymbol{B}_1 & & \\ & \ddots & \\ & & \boldsymbol{B}_k\end{pmatrix}$$

其中 $\boldsymbol{A}_i$ 和 $\boldsymbol{B}_i$ 是同阶方阵. 设 $\boldsymbol{A}_i$ 适合非零多项式 $g_i(x)$, 且 $g_i(x)(1\le i\le k)$ 两两互素. 求证: 若对每个 i, 存在多项式 $f_i(x)$, 使得 $\boldsymbol{B}_i=f_i(\boldsymbol{A}_i)$, 则必存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $B=f(A)$.

笔记 这个命题告诉我们, 在什么条件下可以将分块对角矩阵的多项式表示问题归结为对每一分块的讨论.

证明 因为 $g_i(x)$ 两两互素, 故由中国剩余定理可知, 存在多项式 $h(x)$ 满足 $h(x)=g_i(x)q_i(x)+f_i(x)$. 将 $x=\boldsymbol{A}_i$ 代入上式, 可得 $h(\boldsymbol{A}_i)=f_i(\boldsymbol{A}_i)=\boldsymbol{B}_i$, 从而

$$h(A)=\operatorname{diag}\{h(\boldsymbol{A}_1),\cdots,h(\boldsymbol{A}_k)\}=\operatorname{diag}\{\boldsymbol{B}_1,\cdots,\boldsymbol{B}_k\}=B$$

设 A 的特征多项式为 $g(x)$, 作带余除法 $h(x)=g(x)q(x)+f(x)$, 其中 $\deg f(x)<n$. 将 $x=A$ 代入上式, 则由 Cayley - Hamilton 定理可得 $B=h(A)=f(A)$. □

引理 7.11

设 α,β 均为列向量, 若 $\alpha\beta'=O$, 则 $\alpha=0$ 或 $\beta=0$.

证明 证明是显然的 (反证或者直接写出 $\alpha\beta'$ 的矩阵都可以). □

命题 7.31

设 n 阶矩阵 A 的秩等于 $n-1$,B 是同阶非零矩阵且 $AB=BA=O$, 求证: 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $B=f(A)$.

注 对适合 $AB=BA$ 的矩阵, 由于 $AB=BA$ 当且仅当 $(P^{-1}AP)(P^{-1}BP)=(P^{-1}BP)(P^{-1}AP)$, 因此我们可以通过同时相似变换, 把问题归结为其中一个矩阵是相似标准型（或分块对角型矩阵）的情形来证明.

证明 证法一: 由于题目的条件和结论在同时相似变换:$A\mapsto P^{-1}AP,B\mapsto P^{-1}BP$ 下保持不变, 故不妨从一开始就假设 A 为 Jordan 标准型. 因为 $\mathrm{r}(A)=n-1$, 从而线性方程 $Ax=O$ 的解空间维数为 $n-r(A)=1$, 故 A 关于特征值 0 的几何重数为 1, 从而属于特征值 0 的 Jordan 块只有一个, 记为 J_0; 将属于其他非零特征值的 Jordan 块合在一起, 记为 J_1, 于是 $A=\operatorname{diag}\{J_0,J_1\}$. 设 $B=\begin{pmatrix}B_{11}&B_{12}\\B_{21}&B_{22}\end{pmatrix}$ 为相应的分块, 则由 $AB=BA=O$ 可得 B_{12},B_{21},B_{22} 都是零矩阵, 于是 $B=\operatorname{diag}\{B_{11},O\}$ 且 $J_0B_{11}=B_{11}J_0$. 由于 J_0 是幂零矩阵而 J_1 是可逆矩阵, 并且 J_0 的特征值只有零,J_1 的特征值全非零, 故 J_0 的特征多项式 $g_0(x)$ 和 J_1 的特征多项式 $g_1(x)$ 互素; 又由Jordan 块的性质 (5)可知, 存在多项式 $f_0(x)$, 使得 $B_{11}=f_0(J_0)$; 再取 $f_1(x)=0$, 则 $O=f_1(J_1)$; 最后由命题 7.30即得结论.

证法二: 也可以用线性方程组的求解理论和极小多项式来做. 由于 $\mathrm{r}(A)=n-1$, 故线性方程组 $Ax=0$ 解空间的维数为 1, 再由 $AB=O$ 可知,B 的列向量都是解空间的向量, 从而它们成比例, 于是 $\mathrm{r}(B)=1$. 由引理 6.3可设 $B=\alpha\beta'$, 其中 α,β 为 n 维非零列向量, 由 $AB=O$ 可推出 $A\alpha=0$, 即 α 是线性方程组 $Ax=0$ 的基础解系. 同理, 由 $BA=O$ 可推出 β 是 $A'x=0$ 的基础解系. 设 $m(x)$ 是 A 的极小多项式, 由于 A 不是可逆矩阵, 故由命题 6.47可知 $m(x)$ 的常数项等于零, 从而存在次数比 $m(x)$ 低一次的多项式 $g(x)$, 使得 $m(x)=xg(x)$, 于是 $Ag(A)=O$ 但 $g(A)\neq O$, 否则, 若 $g(A)=O$, 则这与 $m(x)$ 是 A 的极小多项式矛盾! 由类似于矩阵 B 的讨论可得,$g(A)$ 也可以写为 $g(A)=\eta\xi'$, 其中 η 是 $Ax=0$ 的解, 故 $\eta=k\alpha$. 同理可得 $\xi=t\beta$, 于是 $g(A)=ktB$, 从而 $B=\dfrac{1}{kt}g(A)$ 可表示为 A 的次数不超过 $n-1$ 的多项式. □

7.12 可对角化的判断 (二)

7.12.1 极小多项式无重根

命题 7.32

证明:

(1) 若 $A^2 = A$, 则 A 可对角化, 并且 A 的 Jordan 标准型为 $\mathrm{diag}\{1, \cdots, 1, 0 \cdots, 0\}$;

(2) 若 $A^k = I_n$, 则 A 可对角化, 并且 A 的 Jordan 标准型为 $\mathrm{diag}\{\omega_1, \omega_2, \cdots, \omega_n\}$, 其中 $\omega_i^k = 1(1 \leq i \leq n)$.

证明 (1) 矩阵 A 适合 $g(x) = x^2 - x$ 且 $g(x)$ 无重根, 故由命题 6.34可知 A 可对角化, 并且由命题 6.7可知 A 的特征值也适合 $g(x)$, 故只能是 $0, 1$. 因此,A 的 Jordan 标准型为 $\mathrm{diag}\{1, \cdots, 1, 0 \cdots, 0\}$, 其中有 $\mathrm{r}(A)$ 个 1.

(2) 矩阵 A 适合 $g(x) = x^k - 1$ 且 $g(x)$ 无重根, 故由命题 6.34可知 A 可对角化, 并且由命题 6.7可知 A 的特征值也适合 $g(x)$, 故只能是 1 的 k 次方根. 因此,A 的 Jordan 标准型为 $\mathrm{diag}\{\omega_1, \omega_2, \cdots, \omega_n\}$, 其中 $\omega_i^k = 1(1 \leq i \leq n)$. □

例题 7.16 设 A 是有理数域上的 n 阶矩阵, 其特征多项式的所有不可约因式为 $\lambda^2 + \lambda + 1, \lambda^2 - 2$. 又 A 的极小多项式是四次多项式, 求证:A 在复数域上可对角化.

证明 因为 A 的极小多项式 $m(\lambda)$ 和特征多项式 $f(\lambda)$ 有相同的根（不计重数）, 且 $\deg m(\lambda) = 4$, 所以 $m(\lambda) = (\lambda^2 + \lambda + 1)(\lambda^2 - 2)$. 注意到 $m(\lambda)$ 在复数域内无重根, 故 A 在复数域上可对角化. □

命题 7.33

设 φ 是复线性空间 V 上的线性变换,V_0 是 φ 的不变子空间. 求证: 若 φ 可对角化, 则 φ 在 V_0 上的限制变换和 φ 在 V/V_0 上的诱导变换都可对角化.

证明 证法一:由命题 6.37的几何版本可知, 限制变换 $\varphi|_{V_0}$ 和诱导变换 $\overline{\varphi}$ 都有完全的特征向量系, 从而可对角化.

证法二:设线性变换 φ、限制变换 $\varphi|_{V_0}$ 和诱导变换 $\overline{\varphi}$ 的极小多项式分别为 $m(\lambda), g(\lambda)$ 和 $h(\lambda)$, 则由由命题 6.37容易验证 $\varphi|_{V_0}$ 和 $\overline{\varphi}$ 的表示矩阵都适合多项式 $m(x)$, 于是 $\varphi|_{V_0}$ 和 $\overline{\varphi}$ 都适合多项式 $m(\lambda)$, 从而 $g(\lambda) \mid m(\lambda)$ 且 $h(\lambda) \mid m(\lambda)$. 由于 φ 可对角化, 故 $m(\lambda)$ 无重根, 从而 $g(\lambda), h(\lambda)$ 也无重根, 于是 $\varphi|_{V_0}$ 和 $\overline{\varphi}$ 都可对角化. □

命题 7.34

设 φ 是 n 维复线性空间 V 上的线性变换, 求证:φ 可对角化的充要条件是对任一 φ-不变子空间 U, 均存在 φ-不变子空间 W, 使得 $V = U \oplus W$. 这样的 W 称为 U 的 **φ-不变补空间**.

证明 **先证充分性**: 假设 φ 不能对角化, 则 φ 只有 m 个线性无关的特征向量, 其中 $1 \leq m < n$. 设由这些特征向量张成的子空间为 U, 由条件可知,U 存在非零的 φ-不变补空间 W. 考虑限制变换 $\varphi|_W$, 它在 W 上必存在特征值和特征向量, 这些也是 φ 的特征值和特征向量, 于是 φ 有多于 m 个线性无关的特征向量, 矛盾!

再证必要性: 设 φ 可对角化,U 是 φ-不变子空间, 则由命题 7.33可知,$\varphi|_U$ 仍可对角化, 故存在 U 的一组基 $\alpha_1, \cdots, \alpha_r$, 它们是 $\varphi|_U$, 也是 φ 的线性无关的特征向量. 又因为 φ 可对角化, 故存在 n 个线性无关的特征向量 $\{e_1, e_2, \cdots, e_n\}$, 再由基扩张定理可知, 可从这组基中取出 $n - r$ 个向量和 $\alpha_1, \cdots, \alpha_r$ 一起组成 V 的一组基. 设这 $n - r$ 个向量张成的子空间为 W, 则 W 是 U 的 φ-不变补空间. □

命题 7.35

设 n 阶矩阵 A 的极小多项式 $m(\lambda)$ 的次数为 s,$B = (b_{ij})$ 为 s 阶矩阵, 其中 $b_{ij} = \mathrm{tr}(A^{i+j-2})$（约定 $b_{11} = n$）, 求证:A 可对角化的充要条件是 B 为可逆矩阵.

注 本题主要利用的方法是: 设矩阵 A 的全体不同特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_k$, 定义

$$g(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_k)$$

若 A 可对角化, 则 A 的极小多项式就是 $g(\lambda)$（参考常见矩阵的极小多项式 (2)）. 反之, 若 A 适合多项式 $g(\lambda)$, 则

由极小多项式的性质可知,$g(\lambda)$ 就是 A 的极小多项式. 特别地, 由于 $g(\lambda)$ 无重根, 故 A 可对角化.

证明 设 A 的全体不同特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_k$, 其代数重数分别为 $m_1,m_2,\cdots,m_k$, 则由命题 6.4可知 A^i 的全体特征值为 $\lambda_1^i,\lambda_2^i,\cdots,\lambda_k^i$, 其代数重数分别为 $m_1,m_2,\cdots,m_k$. 从而 $\operatorname{tr}(A^i)=m_1\lambda_1^i+m_2\lambda_2^i+\cdots+m_k\lambda_k^i$. 定义 $g(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$, 则 $g(\lambda)\mid m(\lambda)$, 从而 $s\geq k$. 若 A 可对角化, 则 $m(\lambda)=g(\lambda)$, 从而 $s=k$. 若 A 不可对角化, 则 $m(\lambda)$ 有重根, 从而 $s>k$. 考虑矩阵 B 的如下分解:

$$B=\begin{pmatrix} m_1 & m_2 & \cdots & m_k \\ m_1\lambda_1 & m_2\lambda_2 & \cdots & m_k\lambda_k \\ \vdots & \vdots & & \vdots \\ m_1\lambda_1^{s-1} & m_2\lambda_2^{s-1} & \cdots & m_k\lambda_k^{s-1} \end{pmatrix}\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{s-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{s-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_k & \cdots & \lambda_k^{s-1} \end{pmatrix}$$

其中上式右边第一个矩阵是 $s\times k$ 矩阵, 第二个矩阵是 $k\times s$ 矩阵.

必要性: 若 A 可对角化, 则由上述分析可知 $s=k$, 则由 Vandermonde 行列式可知

$$|B|=m_1m_2\cdots m_k\prod_{1\leq i<j\leq k}(\lambda_i-\lambda_j)^2\neq 0$$

即 B 是可逆矩阵.

充分性: 若 B 可逆, 反证, 设 A 不可对角化, 则由上述分析可知 $s>k$, 则由 Cauchy - Binet 公式可得 $|B|=0$, 即 B 不可逆, 矛盾! □

7.12.2 初等因子都是一次多项式, 或 Jordan 块都是一阶矩阵

回顾推论 7.8中可对角化的充要条件.

命题 7.36

设 n 阶复方阵 A 的特征多项式为 $f(\lambda)$, 复系数多项式 $g(\lambda)$ 满足 $(f(\lambda),g'(\lambda))=1$. 证明:$A$ 可对角化的充要条件是 $g(A)$ 可对角化.

证明 必要性显然成立, 下证充分性. 用反证法, 设 A 不可对角化, 则存在可逆矩阵 P, 使得

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

为 Jordan 标准型, 其中 $r_1>1$. 注意到

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

由引理 7.10可知

$$g(J_{r_1}(\lambda_1))=\begin{pmatrix} g(\lambda_1) & g'(\lambda_1) & \cdots & * \\ & g(\lambda_1) & \ddots & \vdots \\ & & \ddots & g'(\lambda_1) \\ & & & g(\lambda_1) \end{pmatrix}$$

由 $(f(\lambda),g'(\lambda))=1$ 可知 $g'(\lambda_1)\neq 0$, 于是 $g(J_{r_1}(\lambda_1))$ 的特征值全为 $g(\lambda_1)$, 其几何重数为 $r_1-\mathrm{r}(g(J_{r_1}(\lambda_1))-g(\lambda_1)I_{r_1})=1$, 因此 $g(J_{r_1}(\lambda_1))$ 的 Jordan 标准型为 $J_{r_1}(g(\lambda_1))$, 其阶数 $r_1>1$. 由于 $J_{r_1}(g(\lambda_1))$ 也是 $g(A)$ 的一个 Jordan 块, 故 $g(A)$ 不可对角化, 矛盾! □

命题 7.37

设 φ 是 n 维复线性空间 V 上的线性变换, 求证:φ 可对角化的充要条件是对 φ 的任一特征值 λ_0, 总有 $\operatorname{Ker}(\varphi-\lambda_0I_V)\cap\operatorname{Im}(\varphi-\lambda_0I_V)=0$.

注 这个命题 7.37是这个命题 7.39的特例.

证明 **先证必要性**: 若 φ 可对角化, 则存在一组基 $\{e_1,e_2,\cdots,e_n\}$, 使得 φ 在这组基下的表示矩阵为 diag $\{\lambda_1,\lambda_2,$

$\cdots, \lambda_n\}$. 适当调整基向量的顺序, 不妨设 $\lambda_0 = \lambda_1 = \cdots = \lambda_r, \lambda_0 \neq \lambda_j (j > r)$, 则 $\varphi - \lambda_0 I_V$ 在基 $\{e_1, e_2, \cdots, e_n\}$ 下的表示矩阵为 $\mathrm{diag}\{0, \cdots, 0, \lambda_{r+1}, \cdots, \lambda_n\}$.

于是对 $\forall v \in V$, 都存在非零列向量 $(a_1, a_2, \cdots, a_n)'$, 使得

$$v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} (e_1, e_2, \cdots, e_n).$$

从而

$$(\varphi - \lambda_0 I_V)(v) = \mathrm{diag}\{0, \cdots, 0, \lambda_{r+1}, \cdots, \lambda_n\} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} (e_1, e_2, \cdots, e_n) = \sum_{i=r+1}^{n} a_i e_i \in L(e_{r+1}, \cdots, e_n).$$

故 $\mathrm{Im}(\varphi - \lambda_0 I_V) \subset L(e_{r+1}, \cdots, e_n)$. 再任取 $\alpha \in L(e_{r+1}, \cdots, e_n)$, 存在非零列向量 $(0, \cdots, 0, b_{r+1}, \cdots, b_n)'$, 使得

$$\alpha = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{r+1} \\ \vdots \\ b_n \end{pmatrix} (e_1, \cdots, e_r, e_{r+1}, \cdots, e_n).$$

从而

$$(\varphi - \lambda_0 I_V)(\alpha) = \mathrm{diag}\{0, \cdots, 0, \lambda_{r+1}, \cdots, \lambda_n\} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{r+1} \\ \vdots \\ b_n \end{pmatrix} (e_1, \cdots, e_r, e_{r+1}, \cdots, e_n) = \sum_{i=r+1}^{n} b_i e_i = \alpha.$$

故 $L(e_{r+1}, \cdots, e_n) \subset \mathrm{Im}(\varphi - \lambda_0 I_V)$. 综上,$\mathrm{Im}(\varphi - \lambda_0 I_V) = L(e_{r+1}, \cdots, e_n)$.

任取 $u \in L(e_1, \cdots, e_r)$, 则存在非零列向量 $(u_1, \cdots, u_r, 0, \cdots, 0)'$, 使得

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} (e_1, \cdots, e_r, e_{r+1}, \cdots, e_n).$$

从而

$$(\varphi - \lambda_0 I_V)(u) = \mathrm{diag}\{0, \cdots, 0, \lambda_{r+1}, \cdots, \lambda_n\} \begin{pmatrix} u_1 \\ \vdots \\ u_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} (e_1, \cdots, e_r, e_{r+1}, \cdots, e_n) = 0.$$

于是 $L(e_1,\cdots,e_r)\subset \mathrm{Ker}(\varphi-\lambda_0 I_V)$. 再任取 $y\in \mathrm{Ker}(\varphi-\lambda_0 I_V)\subset V$, 则存在非零列向量 $(y_1,y_2,\cdots,y_n)'$, 使得

$$y=\begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}(e_1,e_2,\cdots,e_n)=\sum_{i=1}^n y_i e_i.$$

从而

$$(\varphi-\lambda_0 I_V)(y)=\mathrm{diag}\{0,\cdots,0,\lambda_{r+1},\cdots,\lambda_n\}\begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}(e_1,e_2,\cdots,e_n)=\sum_{i=r+1}^n y_i e_i=0.$$

因此 $y_{r+1}=y_{r+2}=\cdots=y_n=0$, 故 $y=\sum_{i=1}^n y_i e_i=\sum_{i=1}^r y_i e_i\in L(e_1,\cdots,e_r)$. 进而 $\mathrm{Ker}(\varphi-\lambda_0 I_V)\subset L(e_1,\cdots,e_r)$. 综上,$\mathrm{Ker}(\varphi-\lambda_0 I_V)=L(e_1,\cdots,e_r)$. 由此可知 $\mathrm{Ker}(\varphi-\lambda_0 I_V)=L(e_1,\cdots,e_r)$,$\mathrm{Im}(\varphi-\lambda_0 I_V)=L(e_{r+1},\cdots,e_n)$, 从而 $\mathrm{Ker}(\varphi-\lambda_0 I_V)\cap \mathrm{Im}(\varphi-\lambda_0 I_V)=0$.

再证充分性: 用反证法, 设 φ 不可对角化, 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 $J=\mathrm{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_k}(\lambda_k)\}$, 其中 $r_1>1$. 由表示矩阵的定义可得 $\varphi(e_1)=\lambda_1 e_1$,$\varphi(e_2)=e_1+\lambda_1 e_2$, 于是 $(\varphi-\lambda_1 I_V)(e_1)=0$,$(\varphi-\lambda_1 I_V)(e_2)=e_1$, 从而 $\mathbf{0}\neq e_1\in \mathrm{Ker}(\varphi-\lambda_1 I_V)\cap \mathrm{Im}(\varphi-\lambda_1 I_V)$, 这与假设矛盾. □

命题 7.38

求证:n 阶复矩阵 A 可对角化的充要条件是对 A 的任一特征值 λ_0,$(\lambda_0 I_n-A)^2$ 和 $\lambda_0 I_n-A$ 的秩相同.

注 这个命题 7.38是这个命题 7.39的特例.

证明 **先证必要性**: 若 A 可对角化, 则存在可逆矩阵 P, 使得 $P^{-1}AP=\Lambda=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$. 适当调整 P 的列向量的顺序, 不妨设 $\lambda_0=\lambda_1=\cdots=\lambda_r$,$\lambda_0\neq\lambda_j(j>r)$, 则由于相似矩阵的特征多项式相同可知,$\lambda_0 I_n-A=\lambda_0 I_n-\Lambda$. 从而 $\mathrm{r}(\lambda_0 I_n-A)=\mathrm{r}(\lambda_0 I_n-\Lambda)=n-r$,$\mathrm{r}((\lambda_0 I_n-A)^2)=\mathrm{r}((\lambda_0 I_n-\Lambda)^2)=n-r$, 于是结论成立.

再证充分性: 用反证法, 若 A 不可对角化, 则存在可逆矩阵 P, 使得 $P^{-1}AP=J=\mathrm{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_k}(\lambda_k)\}$ 为 Jordan 标准型, 其中不妨设 $r_1>1$. 由于相似矩阵的特征多项式相等, 因此 $\lambda_1 I_n-A=\lambda_0 I_n-J$. 从而注意到

$$\mathrm{r}((\lambda_1 I_n-A)^j)=\mathrm{r}((\lambda_1 I_n-J)^j)=\sum_{i=1}^k \mathrm{r}((\lambda_1 I_{r_i}-J_{r_i}(\lambda_i))^j),\ j\geq 1$$

又注意到 $\lambda_1 I_{r_1}-J_{r_1}(\lambda_1)$ 是 r_1 阶基础幂零阵, 故 $\mathrm{r}(\lambda_1 I_{r_1}-J_{r_1}(\lambda_1))=r_1-1$,$\mathrm{r}((\lambda_1 I_{r_1}-J_{r_1}(\lambda_1))^2)=r_1-2$, 因此 $\mathrm{r}((\lambda_1 I_n-A)^2)<\mathrm{r}(\lambda_1 I_n-A)$, 这与假设矛盾. □

命题 7.39

设 φ 是 n 维复线性空间 V 上的线性变换, 求证: φ 可对角化的充要条件是对 φ 的任一特征值 λ_0, 下列条件之一成立:

(1) $V=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)+\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)$;

(2) $V=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)\oplus\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)$;

(3) $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)\cap\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)=0$;

(4) $\dim\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)=\dim\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^2$;

(5) $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^2=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^3=\cdots$;

(6) $\mathrm{r}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{r}((\varphi-\lambda_0\boldsymbol{I}_V)^2)$;

(7) $\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)^2=\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)^3=\cdots$;

(8) $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)$ 存在 φ-不变补空间, 即存在 φ-不变子空间 U, 使得 $V=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)\oplus U$;

(9) $\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)$ 存在 φ-不变补空间, 即存在 φ-不变子空间 W, 使得 $V=\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)\oplus W$.

注 命题 7.37与命题 7.38都是这个命题 7.39的特例.

笔记 由命题 4.22可知条件 (1)~(9) 是相互等价的, 因此本题的结论由命题 7.37(与条件 (3) 对应) 或命题 7.38(与条件 (6) 对应) 即得. 事实上, 对充分性而言, 我们还可以从其他条件出发来证明 φ 可对角化, 下面是 3 种证法.

证明 证法一:对任一特征值 λ_0, 由 $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^2=\cdots=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^n$, 取维数之后可得特征值 λ_0 的几何重数等于代数重数, 从而 φ 有完全的特征向量系, 于是 φ 可对角化.

证法二: 对任一特征值 λ_0, 由 $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^2=\cdots=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^n$ 可知, 特征子空间等于根子空间, 再由根子空间的直和分解可知, 全空间等于特征子空间的直和, 从而 φ 可对角化.

证法三: 设 φ 的全体不同特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_k$, 特征多项式 $f(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_k)^{m_k}$, 则对任意的 $\boldsymbol{\alpha}\in V$, 由 Cayley - Hamilton 定理可得

$$(\varphi-\lambda_1\boldsymbol{I}_V)^{m_1}(\varphi-\lambda_2\boldsymbol{I}_V)^{m_2}\cdots(\varphi-\lambda_k\boldsymbol{I}_V)^{m_k}(\boldsymbol{\alpha})=\boldsymbol{0},$$

即有 $(\varphi-\lambda_2\boldsymbol{I}_V)^{m_2}\cdots(\varphi-\lambda_k\boldsymbol{I}_V)^{m_k}(\boldsymbol{\alpha})\in\mathrm{Ker}(\varphi-\lambda_1\boldsymbol{I}_V)^{m_1}=\mathrm{Ker}(\varphi-\lambda_1\boldsymbol{I}_V)$, 从而

$$(\varphi-\lambda_1\boldsymbol{I}_V)(\varphi-\lambda_2\boldsymbol{I}_V)^{m_2}\cdots(\varphi-\lambda_k\boldsymbol{I}_V)^{m_k}(\boldsymbol{\alpha})=\boldsymbol{0}.$$

不断这样做下去, 最终可得对任意的 $\boldsymbol{\alpha}\in V$, 总有

$$(\varphi-\lambda_1\boldsymbol{I}_V)(\varphi-\lambda_2\boldsymbol{I}_V)\cdots(\varphi-\lambda_k\boldsymbol{I}_V)(\boldsymbol{\alpha})=\boldsymbol{0},$$

即 φ 适合多项式 $g(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$, 从而 φ 可对角化. □

例题 7.17 若 $n(n\geq 2)$ 阶矩阵 B 相似于 $R=\mathrm{diag}\left\{\begin{pmatrix}0&1\\1&0\end{pmatrix},I_{n-2}\right\}$, 则称 B 为反射矩阵. 证明: 任一对合矩阵 A（即 $A^2=I_n$）均可分解为至多 n 个两两乘法可交换的反射矩阵的乘积.

证明 由命题 7.32(2)可知, 对合矩阵 A 可对角化, 即存在可逆矩阵 P, 使得 $P^{-1}AP=\mathrm{diag}\{-I_r,I_{n-r}\}$, 其中 $0\leq r\leq n$. 当 $r=0$ 时,$A=I_n=R^2$, 结论成立. 当 $r\geq 1$ 时, 设 $B_i=P\mathrm{diag}\{1,\cdots,1,-1,1,\cdots,1\}P^{-1}$, 其中 -1 在主对角线上的第 i 个位置, 则 $B_i(1\leq i\leq r)$ 两两乘法可交换, 并且 $A=B_1B_2\cdots B_r$. 由于 $\begin{pmatrix}0&1\\1&0\end{pmatrix}$ 的特征值是 $-1,1$, 故其相似于 $\mathrm{diag}\{-1,1\}$, 因此矩阵 B 是反射矩阵当且仅当 B 相似于 $\mathrm{diag}\{-1,1,\cdots,1\}$. 因为对角矩阵的两个主对角元素对换是一个相似变换, 所以上述 B_i 都是反射矩阵, 于是 A 可以分解为 r 个两两乘法可交换的反射矩阵的乘积. □

7.13 Jordan 标准型的求法

分析矩阵结构的方法:

(1) 计算行列式因子对于某些具有简单结构的矩阵 (如上 (下) 三角矩阵、类上 (下) 三角矩阵), 可以通过选取适当的子式, 计算出行列式因子, 再得到不变因子和初等因子. 比如, Frobenius 块和 Jordan 块就是利用这种方法的典型例子.

(2) 计算极小多项式因为矩阵的极小多项式是整除关系下最大的不变因子, 所以极小多项式确定了最大 Jordan 块的阶数.

(3) 计算特征值的几何重数因为特征值的几何重数等于其 Jordan 块的个数, 所以计算几何重数有助于 Jordan 标准型的确定.

命题 7.40

设 n 阶矩阵 $\boldsymbol{A}$ 的不变因子组为 $d_1(\lambda),d_2(\lambda),\cdots,d_n(\lambda)$, 其中 $d_i(\lambda)\mid d_{i+1}(\lambda)\ (1\leq i\leq n-1)$, 又 λ_0 是 $\boldsymbol{A}$ 的特征值. 求证: $\mathrm{r}(\lambda_0\boldsymbol{I}_n-\boldsymbol{A})=r$ 的充要条件是 $(\lambda-\lambda_0)\nmid d_r(\lambda)$ 但 $(\lambda-\lambda_0)\mid d_{r+1}(\lambda)$.

证明 证法一:$\mathrm{r}(\lambda_0\boldsymbol{I}_n-\boldsymbol{A})=r$ 当且仅当特征值 λ_0 的几何重数为 $n-r$; 这当且仅当特征值 λ_0 的 Jordan 块有 $n-r$ 个; 由不变因子之间的整除关系可知, 这当且仅当后 $n-r$ 个不变因子能被 $\lambda-\lambda_0$ 整除, 而前 r 个不变因子不能被 $\lambda-\lambda_0$ 整除.

证法二:由命题 7.27可知, $\mathrm{r}(\lambda_0 \boldsymbol{I}_n - \boldsymbol{A}) = r$ 当且仅当 $\sum_{i=1}^{n} \delta_{d_i(\lambda_0),0} = n - r$; 由不变因子之间的整除关系可知, 这当且仅当 $d_i(\lambda_0) \neq 0\ (1 \leq i \leq r)$ 且 $d_i(\lambda_0) = 0\ (r+1 \leq i \leq n)$; 最后由余数定理即得结论. □

命题 7.41

设 φ 是 n 维线性空间 V 上的线性变换,U 是 V 的非零 $\varphi-$ 不变子空间. 设 λ_0 是限制变换 $\varphi|_U$ 的特征值, 证明:$\varphi|_U$ 的属于特征值 λ_0 的 Jordan 块的个数不超过 φ 的属于特征值 λ_0 的 Jordan 块的个数.

证明 Jordan 块的个数等于特征值的几何重数, 即线性无关的特征向量的个数. 设 $\varphi|_U$ 的属于特征值 λ_0 的 Jordan 块的个数为 r, 则 $\varphi|_U$ 关于特征值 λ_0 有 r 个线性无关的特征向量, 它们也都是 φ 关于特征值 λ_0 的线性无关的特征向量, 从而 φ 的属于特征值 λ_0 的 Jordan 块至少有 r 个.

也可用纯代数的方法（矩阵的秩）进行证明, 请读者自行思考完成. □

例题 7.18 求下列 n 阶矩阵的 Jordan 标准型, 其中 $a \neq 0$:

$$\boldsymbol{A} = \begin{pmatrix} a & a & a & \cdots & a \\ & a & a & \cdots & a \\ & & a & \cdots & a \\ & & & \ddots & \vdots \\ & & & & a \end{pmatrix}$$

注 本题同时利用分析矩阵结构的三种方法计算其 Jordan 标准型.

解 解法一:由命题 7.15可知,$\boldsymbol{A}$ 的行列式因子组为 $1, \cdots, 1, (\lambda - a)^n$, 这也是 $\boldsymbol{A}$ 的不变因子组, 从而 $\boldsymbol{A}$ 的 Jordan 标准型为 $\boldsymbol{J}_n(a)$.

解法二:显然 $\boldsymbol{A}$ 的特征多项式为 $(\lambda - a)^n$, 故 $\boldsymbol{A}$ 的极小多项式是 $\lambda - a$ 的某个幂. 设 $\boldsymbol{N} = \boldsymbol{J}_n(0)$, 即特征值为 0 的 n 阶 Jordan 块, 它满足 $\boldsymbol{N}^{n-1} \neq \boldsymbol{O}$ 但 $\boldsymbol{N}^n = \boldsymbol{O}$, 则 $\boldsymbol{A} = a(\boldsymbol{I}_n + \boldsymbol{N} + \boldsymbol{N}^2 + \cdots + \boldsymbol{N}^{n-1})$. 注意到

$$(\boldsymbol{A} - a\boldsymbol{I}_n)^{n-1} = a^{n-1}(\boldsymbol{N} + \boldsymbol{N}^2 + \cdots + \boldsymbol{N}^{n-1})^{n-1} = a^{n-1}\boldsymbol{N}^{n-1} \neq \boldsymbol{O}$$

故 $\boldsymbol{A}$ 不适合多项式 $(\lambda - a)^{n-1}$, 于是 $\boldsymbol{A}$ 的极小多项式只能是 $(\lambda - a)^n$. 因此 $\boldsymbol{A}$ 的不变因子组是 $1, \cdots, 1, (\lambda - a)^n$, 从而 $\boldsymbol{A}$ 的 Jordan 标准型为 $\boldsymbol{J}_n(a)$.

解法三:显然 $\boldsymbol{A}$ 的特征值全为 a, 我们来计算它的几何重数. 注意到 $\mathrm{r}(a\boldsymbol{I}_n - \boldsymbol{A}) = n - 1$, 故特征值 a 的几何重数为 $n - \mathrm{r}(a\boldsymbol{I}_n - \boldsymbol{A}) = 1$, 于是 $\boldsymbol{A}$ 的 Jordan 标准型中关于特征值 a 的 Jordan 块只有一个, 因此 $\boldsymbol{A}$ 的 Jordan 标准型为 $\boldsymbol{J}_n(a)$. □

命题 7.42 (秩一阵的 Jordan 标准型)

设 $n(n > 1)$ 阶矩阵 $\boldsymbol{A}$ 的秩为 1, 证明:

若 $\mathrm{tr}(\boldsymbol{A}) \neq 0$, 则 $\boldsymbol{A}$ 的 Jordan 标准型为 $\mathrm{diag}\{0, \cdots, 0, \mathrm{tr}(\boldsymbol{A})\}$.

若 $\mathrm{tr}(\boldsymbol{A}) = 0$, 则 $\boldsymbol{A}$ 的 Jordan 标准型为 $\mathrm{diag}\{0, \cdots, 0, \boldsymbol{J}_2(0)\}$.

证明 解法一:由 $\mathrm{r}(\boldsymbol{A}) = 1$ 及引理 6.3可知, 存在非零列向量 $\boldsymbol{\alpha}, \boldsymbol{\beta}$, 使得 $\boldsymbol{A} = \boldsymbol{\alpha}\boldsymbol{\beta}'$. 由特征值的降价公式可得 $|\lambda \boldsymbol{I}_n - \boldsymbol{A}| = \lambda^{n-1}(\lambda - \boldsymbol{\beta}'\boldsymbol{\alpha})$, 再由所有特征值之和等于矩阵的迹可得 $\mathrm{tr}(\boldsymbol{A}) = \boldsymbol{\beta}'\boldsymbol{\alpha}$. 若 $\mathrm{tr}(\boldsymbol{A}) \neq 0$, 则特征值 $\mathrm{tr}(\boldsymbol{A})$ 的几何重数等于 1, 特征值 0 的几何重数等于 $n - \mathrm{r}(\boldsymbol{A}) = n - 1$, 因此 $\boldsymbol{A}$ 的 Jordan 标准型为 $\mathrm{diag}\{0, \cdots, 0, \mathrm{tr}(\boldsymbol{A})\}$. 若 $\mathrm{tr}(\boldsymbol{A}) = 0$, 则特征值 0 的代数重数是 n, 几何重数是 $n - 1$, 因此 $\boldsymbol{A}$ 的 Jordan 标准型为 $\mathrm{diag}\{0, \cdots, 0, \boldsymbol{J}_2(0)\}$.

解法二:特征多项式的计算同解法一, 又由常见矩阵的极小多项式 (4)可知,$\boldsymbol{A}$ 的极小多项式 $m(\lambda) = \lambda(\lambda - \mathrm{tr}(\boldsymbol{A}))$, 于是 $\boldsymbol{A}$ 的不变因子组为 $1, \lambda, \cdots, \lambda, m(\lambda)$. 若 $\mathrm{tr}(\boldsymbol{A}) \neq 0$, 则 $\boldsymbol{A}$ 的 Jordan 标准型为 $\mathrm{diag}\{0, \cdots, 0, \mathrm{tr}(\boldsymbol{A})\}$. 若 $\mathrm{tr}(\boldsymbol{A}) = 0$, 则 $\boldsymbol{A}$ 的 Jordan 标准型为 $\mathrm{diag}\{0, \cdots, 0, \boldsymbol{J}_2(0)\}$.

解法三:直接利用 Jordan 标准型来解最为简单. 特征多项式和特征值的计算同解法一, 由于还不知道每个特征值的几何重数, 故设 $\boldsymbol{A}$ 的 Jordan 标准型 $\boldsymbol{J} = \mathrm{diag}\left\{ \boldsymbol{J}_{r_1}(0), \cdots, \boldsymbol{J}_{r_k}(0), \boldsymbol{J}_{s_1}(\lambda_1), \cdots, \boldsymbol{J}_{s_l}(\lambda_l) \right\}$(因为 Jordan 块对应的特征值为 0 时秩会减 1, 所以单独把特征值为 0 的 Jordan 块分出来), 其中 $\lambda_j \neq 0\ (1 \leq j \leq l)$. 由于相似关系不改变

矩阵的秩, 故 $\boldsymbol{J}$ 的秩也为 1, 即有 $(r_1-1)+\cdots+(r_k-1)+s_1+\cdots+s_l=1$. 于是只有以下两种情况成立: 第一种情况是 $l=1, s_1=1, \lambda_1=\mathrm{tr}(\boldsymbol{A})\neq 0$, 且所有的 $r_i=1$, 此时 $\boldsymbol{A}$ 的 Jordan 标准型为 $\mathrm{diag}\{0,\cdots,0,\mathrm{tr}(\boldsymbol{A})\}$. 第二种情况是某个 $r_i=2$, 其余的 $r_i=1$ 且 $l=0$, 此时 $\boldsymbol{A}$ 的 Jordan 标准型为 $\mathrm{diag}\{0,\cdots,0,\boldsymbol{J}_2(0)\}$. □

命题 7.43

设 $n(n>1)$ 阶矩阵 A 的秩为 1, 求证:

(1) A 是幂等矩阵的充要条件是 $\mathrm{tr}\,(A)=1$,

(2) A 是幂零矩阵的充要条件是 $\mathrm{tr}\,(A)=0$.

证明 由命题 7.42的结论即得. □

例题 7.19 设 $A=\begin{pmatrix} 1 & 0 & 0 & 0 \\ a+2 & 1 & 0 & 0 \\ 5 & 3 & 1 & 0 \\ 7 & 6 & b+4 & 1 \end{pmatrix}$, 求 A 的 Jordan 标准型.

解 显然 A 的特征值全为 1, 首先我们来计算特征值 1 的几何重数. 考虑矩阵

$$A-I_4=\begin{pmatrix} 0 & 0 & 0 & 0 \\ a+2 & 0 & 0 & 0 \\ 5 & 3 & 0 & 0 \\ 7 & 6 & b+4 & 0 \end{pmatrix}.$$

(1) 当 $a+2\neq 0$ 且 $b+4\neq 0$ 时, $\mathrm{r}(A-I_4)=3$, 于是特征值 1 的几何重数等于 1, 从而只有一个 Jordan 块, 因此 A 的 Jordan 标准型是 $J_4(1)$.

(2) 当 $a+2=0$ 或 $b+4=0$ 时, $\mathrm{r}(A-I_4)=2$, 于是特征值 1 的几何重数等于 2, 从而有两个 Jordan 块. 进一步我们来计算 A 的极小多项式.

(2.1) 若 $a+2=0$ 和 $b+4=0$ 中只有一个成立, 容易验证 $(A-I_4)^2\neq O$, 但 $(A-I_4)^3=O$, 于是 A 的极小多项式是 $(\lambda-1)^3$, 从而不变因子组为 $1,1,\lambda-1,(\lambda-1)^3$, 因此 A 的 Jordan 标准型为 $\mathrm{diag}\{1,J_3(1)\}$.

(2.2) 若 $a+2=0$ 和 $b+4=0$ 都成立, 容易验证 $(A-I_4)^2=O$, 于是 A 的极小多项式是 $(\lambda-1)^2$, 从而不变因子组为 $1,1,(\lambda-1)^2,(\lambda-1)^2$, 因此 A 的 Jordan 标准型为 $\mathrm{diag}\{J_2(1),J_2(1)\}$. □

例题 7.20 设 $J=J_n(0)$ 是特征值为零的 $n(n\geq 2)$ 阶 Jordan 块, 求 J^2 的 Jordan 标准型.

解 显然 J^2 的特征值全为 0 且 $\mathrm{r}(J^2)=n-2$, 于是特征值 0 的几何重数等于 2, 从而有两个 Jordan 块. 接下去计算 J^2 的极小多项式, 注意到 $J^n=O, J^{n-1}\neq O$.

(1) 当 $n=2m$ 时, λ^m 是 J^2 的极小多项式, 于是 J^2 的不变因子组为 $1,\cdots,1,\lambda^m,\lambda^m$, 因此 J^2 的 Jordan 标准型为 $\mathrm{diag}\{J_m(0),J_m(0)\}$.

(2) 当 $n=2m+1$ 时, λ^{m+1} 是 J^2 的极小多项式, 于是 J^2 的不变因子组为 $1,\cdots,1,\lambda^m,\lambda^{m+1}$, 因此 J^2 的 Jordan 标准型为 $\mathrm{diag}\{J_m(0),J_{m+1}(0)\}$.

另外, 也可以用行列式因子的讨论来替代几何重数的讨论. 注意到 λI_n-J^2 的右上角有一个 $n-2$ 阶子式等于 $(-1)^{n-2}$, 故 J^2 的 $n-2$ 阶行列式因子为 1, 从而前 $n-2$ 个不变因子都是 1, 后面再用极小多项式的讨论即可得到结论. □

例题 7.21 求下列 $n(n\geq 2)$ 阶矩阵的 Jordan 标准型:

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

解 利用例题 7.20的记号和结论, 显然 $A = cI_n + J^2$. 设 P 是可逆矩阵, 使得 $P^{-1}J^2P$ 是 J^2 的 Jordan 标准型, 则 $P^{-1}AP = cI_n + P^{-1}J^2P$ 就是 A 的 Jordan 标准型. 具体地, 当 $n = 2m$ 时, A 的 Jordan 标准型是 $\mathrm{diag}\{J_m(c), J_m(c)\}$; 当 $n = 2m + 1$ 时, A 的 Jordan 标准型是 $\mathrm{diag}\{J_m(c), J_{m+1}(c)\}$. □

我们可以自然地考虑如下问题: 如果已知 n 阶矩阵 A 的 Jordan 标准型, 那么对任意的正整数 m, A^m 的 Jordan 标准型应该有怎样的形状呢?(后续的命题 7.46和命题 7.47完美解决了这个问题) 首先, 我们可以把这个问题化约到 Jordan 块的情形. 设 A 的 Jordan 标准型为 $J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_s}(\lambda_s)\}$, 则 A^m 相似于 $J^m = \mathrm{diag}\{J_{r_1}(\lambda_1)^m, J_{r_2}(\lambda_2)^m, \cdots, J_{r_s}(\lambda_s)^m\}$, 因此要求 A^m 的 Jordan 标准型, 只要求每一个 $J_{r_i}(\lambda_i)^m$ 的 Jordan 标准型即可. 若 $\lambda_i \neq 0$, 则由例题 7.18类似的讨论可知, $J_{r_i}(\lambda_i)^m$ 的 Jordan 标准型为 $J_{r_i}(\lambda_i^m)$. 若 $\lambda_i = 0$, 则例题 7.20处理了 $m = 2$ 的情形, 不过类似的讨论很难推广到 $m \geq 3$ 的情形, 换言之, 只依靠几何重数和极小多项式还不能完全确定 $J_{r_i}(0)^m$ 的 Jordan 标准型. 解决这个问题可以有代数和几何两种方法, 几何方法 (利用 Jordan 标准型的几何意义) , 而代数方法 (利用矩阵的秩) 则需要下面的命题 7.44.

命题 7.44

设 λ_0 是 n 阶矩阵 A 的特征值, 证明: 对任意的正整数 k, 特征值为 λ_0 的 k 阶 Jordan 块 $J_k(\lambda_0)$ 在 A 的 Jordan 标准型 J 中出现的个数为

$$\mathrm{r}((A - \lambda_0 I_n)^{k-1}) + \mathrm{r}((A - \lambda_0 I_n)^{k+1}) - 2\mathrm{r}((A - \lambda_0 I_n)^k),$$

其中约定 $\mathrm{r}((A - \lambda_0 I_n)^0) = n$.

注 这个命题 7.44告诉我们, n 阶矩阵 A 的 Jordan 标准型被若干个非负整数, 即 $\{\mathrm{r}((A - \lambda_i I_n)^j) \mid \lambda_i$ 为 A 的特征值, $1 \leq j \leq n\}$ 完全决定. 因此从理论上说, 我们可以不计算矩阵 A 的不变因子或初等因子, 改为计算上述若干个矩阵的秩, 也可以求出 A 的 Jordan 标准型. 进一步, 我们还可以得到如下矩阵相似的判定准则.

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_s}(\lambda_s)\}$ 为 A 的 Jordan 标准型. 注意到

$$(A - \lambda_0 I_n)^k = P\mathrm{diag}\{J_{r_1}(\lambda_1 - \lambda_0)^k, J_{r_2}(\lambda_2 - \lambda_0)^k, \cdots, J_{r_s}(\lambda_s - \lambda_0)^k\}P^{-1},$$

故 $\mathrm{r}((A - \lambda_0 I_n)^k) = \sum\limits_{i=1}^{s} \mathrm{r}(J_{r_i}(\lambda_i - \lambda_0)^k)$.

(i) 当 $\lambda_i \neq \lambda_0$ 时, $\mathrm{r}(J_{r_i}(\lambda_i - \lambda_0)^k) = r_i$. 当 $\lambda_i = \lambda_0$ 时, 若 $r_i < k$, 则 $\mathrm{r}(J_{r_i}(\lambda_i - \lambda_0)^k) = 0$; 若 $r_i \geq k$, 则 $\mathrm{r}(J_{r_i}(\lambda_i - \lambda_0)^k) = r_i - k$.

(ii) 当 $\lambda_i \neq \lambda_0$ 时, $\mathrm{r}(J_{r_i}(\lambda_i - \lambda_0)^{k-1}) = r_i$. 当 $\lambda_i = \lambda_0$ 时, 若 $r_i < k - 1$, 则 $\mathrm{r}(J_{r_i}(\lambda_i - \lambda_0)^{k-1}) = 0$; 若 $r_i \geq k - 1$, 则 $\mathrm{r}(J_{r_i}(\lambda_i - \lambda_0)^{k-1}) = r_i - k + 1$.

于是当 $\lambda_i \neq \lambda_0$ 时,$\mathrm{r}((A - \lambda_0 I_n)^{k-1}) - \mathrm{r}((A - \lambda_0 I_n)^k) = r_i - r_i = 0$;

当 $\lambda_i = \lambda_0$ 时, 若 $r_i < k - 1$, 则 $\mathrm{r}((A - \lambda_0 I_n)^{k-1}) - \mathrm{r}((A - \lambda_0 I_n)^k) = 0 - 0 = 0$;

若 $r_i = k - 1$, 则 $\mathrm{r}((A - \lambda_0 I_n)^{k-1}) - \mathrm{r}((A - \lambda_0 I_n)^k) = 0 - (k - 1 - k + 1) = 0$;

若 $r_i \geq k$, 则 $\mathrm{r}((A - \lambda_0 I_n)^{k-1}) - \mathrm{r}((A - \lambda_0 I_n)^k) = (r_i - k) - (r_i - k + 1) = 1$.

因此 $\mathrm{r}((A - \lambda_0 I_n)^{k-1}) - \mathrm{r}((A - \lambda_0 I_n)^k)$ 等于特征值为 λ_0 且阶数大于等于 k 的 Jordan 块的个数. 同理, $\mathrm{r}((A - \lambda_0 I_n)^k) - \mathrm{r}((A - \lambda_0 I_n)^{k+1})$ 等于特征值为 λ_0 且阶数大于等于 $k + 1$ 的 Jordan 块的个数, 从而特征值为 λ_0 的 k 阶 Jordan 块 $J_k(\lambda_0)$ 在 A 的 Jordan 标准型 J 中出现的个数为

$$\begin{aligned}&(\mathrm{r}((A - \lambda_0 I_n)^{k-1}) - \mathrm{r}((A - \lambda_0 I_n)^k)) - (\mathrm{r}((A - \lambda_0 I_n)^k) - \mathrm{r}((A - \lambda_0 I_n)^{k+1}))\\ =&\mathrm{r}((A - \lambda_0 I_n)^{k-1}) + \mathrm{r}((A - \lambda_0 I_n)^{k+1}) - 2\mathrm{r}((A - \lambda_0 I_n)^k).\end{aligned}$$

□

命题 7.45

设 A, B 为 n 阶矩阵, 证明: 它们相似的充要条件是对 A 或 B 的任一特征值 λ_0 以及任意的 $1 \leq k \leq n$, 有 $\mathrm{r}((A - \lambda_0 I_n)^k) = \mathrm{r}((B - \lambda_0 I_n)^k)$.

证明 必要性由相似矩阵特征多项式相同显然,现证充分性. 由已知条件及命题 4.20可知,

$$\mathrm{r}((A-\lambda_0 I_n)^{n+1})=\mathrm{r}((A-\lambda_0 I_n)^n)=\mathrm{r}((B-\lambda_0 I_n)^n)=\mathrm{r}((B-\lambda_0 I_n)^{n+1}).$$

因此由命题 7.44可知, 特征值为 λ_0 的 k 阶 Jordan 块 $J_k(\lambda_0)$ 在 A,B 的 Jordan 标准型中出现的个数相同, 从而 A,B 有相同的 Jordan 标准型, 于是它们相似. □

命题 7.46

设 $J=J_n(a)$ 是特征值为 $a\neq 0$ 的 n 阶 Jordan 块, 证明:J^m 的 Jordan 标准型为 $J_n(a^m)$, 其中 m 为非零整数.

证明 先处理 $m\geq 1$ 的情形, 采用几何重数的方法来做, 行列式因子和极小多项式的方法也可以做, 请读者自行补充完成. 显然 J^m 的所有特征值都为 a^m. 作分解 $J=aI_n+N$, 其中 $N=J_n(0)$, 则有

$$J^m=(aI_n+N)^m=a^mI_n+\mathrm{C}_m^1a^{m-1}N+\cdots+N^m,$$

于是 $\mathrm{r}(J^m-a^mI_n)=\mathrm{r}(\mathrm{C}_m^1a^{m-1}N+\cdots+N^m)=n-1$, 从而特征值 a^m 的几何重数等于 1, 因此 J^m 的 Jordan 标准型中只有一个 Jordan 块, 即 J^m 的 Jordan 标准型为 $J_n(a^m)$.

再处理 $m=-1$ 的情形. 显然 J^{-1} 的所有特征值都为 a^{-1}. 注意到

$$J^{-1}=(aI_n+N)^{-1}=a^{-1}I_n-a^{-2}N+\cdots+(-1)^{n-1}a^{-n}N^{n-1},$$

故 $\mathrm{r}(J^{-1}-a^{-1}I_n)=\mathrm{r}(-a^{-2}N+\cdots+(-1)^{n-1}a^{-n}N^{n-1})=n-1$, 从而特征值 a^{-1} 的几何重数等于 1, 因此 J^{-1} 的 Jordan 标准型中只有一个 Jordan 块, 即 J^{-1} 的 Jordan 标准型为 $J_n(a^{-1})$.

最后处理 $m\leq -1$ 的情形. 注意到 $J^m=(J^{-1})^{-m}$, 故由前面两个结论即得 J^m 的 Jordan 标准型为 $J_n((a^{-1})^{-m})=J_n(a^m)$. □

命题 7.47

设 $J=J_n(0)$ 是特征值为零的 n 阶 Jordan 块, 证明: $J^m(m\geq 1)$ 的 Jordan 标准型为 $\mathrm{diag}\{J_q(0),\cdots,J_q(0),J_{q+1}(0),\cdots,J_{q+1}(0)\}$, 其中有 $m-r$ 个 $J_q(0)$, r 个 $J_{q+1}(0)$.

注 这个命题是例题 7.20的推广.

解 若 $m\geq n$, 则 $J^m=O$, 这就是它的 Jordan 标准型. 下设 $m<n$, 并作带余除法: $n=mq+r$, 其中 $0\leq r<m$. 我们先来计算 J^m 的幂的秩, 再利用命题 7.44来计算 Jordan 块的个数. 注意到

$$\mathrm{r}((J^m)^k)=n-mk,\ 0\leq k\leq q;\ \mathrm{r}((J^m)^k)=0,\ k\geq q+1.$$

(1) 当 $1\leq k<q$ 时, $J_k(0)$ 的个数为 $\mathrm{r}((J^m)^{k-1})+\mathrm{r}((J^m)^{k+1})-2\mathrm{r}((J^m)^k)=(n-m(k-1))+(n-m(k+1))-2(n-mk)=0$;

(2) $J_q(0)$ 的个数为 $\mathrm{r}((J^m)^{q-1})+\mathrm{r}((J^m)^{q+1})-2\mathrm{r}((J^m)^q)=(n-m(q-1))+0-2(n-mq)=m-r$;

(3) $J_{q+1}(0)$ 的个数为 $\mathrm{r}((J^m)^q)+\mathrm{r}((J^m)^{q+2})-2\mathrm{r}((J^m)^{q+1})=(n-mq)+0-0=r$;

(4) 当 $k>q+1$ 时, $J_k(0)$ 的个数为 0.

因此 J^m 的 Jordan 标准型为 $\mathrm{diag}\{J_q(0),\cdots,J_q(0),J_{q+1}(0),\cdots,J_{q+1}(0)\}$, 其中有 $m-r$ 个 $J_q(0)$, r 个 $J_{q+1}(0)$. □

命题 7.48

设 m 阶矩阵 A 与 n 阶矩阵 B 没有公共的特征值, 且 A,B 的 Jordan 标准型分别为 J_1,J_2, 又 C 为 $m\times n$ 矩阵, 求证: $M=\begin{pmatrix}A&C\\O&B\end{pmatrix}$ 的 Jordan 标准型为 $\mathrm{diag}\{J_1,J_2\}$.

注 这个命题可用来化简矩阵, 消去其非主对角块, 使其剩下低阶的主对角块.

证明 证法一: 设 $P_1(\lambda),P_2(\lambda),Q_1(\lambda),Q_2(\lambda)$ 是可逆 λ-矩阵, 使得

$$P_1(\lambda)(\lambda I_m-A)Q_1(\lambda)=\Lambda_1=\mathrm{diag}\{f_1(\lambda),f_2(\lambda),\cdots,f_m(\lambda)\},$$

$$P_2(\lambda)(\lambda I_n-B)Q_2(\lambda)=\Lambda_2=\mathrm{diag}\{g_1(\lambda),g_2(\lambda),\cdots,g_n(\lambda)\}$$

分别是 A,B 的法式. 考虑如下 λ-矩阵的初等变换:

$$\begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix}\begin{pmatrix} \lambda I_m - A & -C \\ O & \lambda I_n - B \end{pmatrix}\begin{pmatrix} Q_1 & O \\ O & Q_2 \end{pmatrix} = \begin{pmatrix} \Lambda_1 & D \\ O & \Lambda_2 \end{pmatrix},$$

其中 $D = -P_1CQ_2 = (d_{ij}(\lambda))$ 是 $m\times n$ λ-矩阵. 由于 A,B 没有公共的特征值, 故对任意的 $1 \le i \le m, 1 \le j \le n$, $(f_i(\lambda), g_j(\lambda)) = 1$, 从而存在 $u_{ij}(\lambda), v_{ij}(\lambda)$, 使得 $f_i(\lambda)u_{ij}(\lambda) + g_j(\lambda)v_{ij}(\lambda) = 1$. 将 λ-矩阵 $\begin{pmatrix} \Lambda_1 & D \\ O & \Lambda_2 \end{pmatrix}$ 的第 i 列乘以 $-u_{ij}(\lambda)d_{ij}(\lambda)$ 加到第 $m+j$ 列上, 再将第 $m+j$ 行乘以 $-v_{ij}(\lambda)d_{ij}(\lambda)$ 加到第 i 行上, 则可以消去 D 的第 (i,j) 元素, 因此 M 的特征矩阵相抵于对角矩阵 $\mathrm{diag}\{\Lambda_1,\Lambda_2\}$. 再由$\lambda$-矩阵和初等因子的基本性质 (2)可知, M 的初等因子组是 $f_1(\lambda),\cdots,f_m(\lambda),g_1(\lambda),\cdots,g_n(\lambda)$ 的准素因子组, 而 $f_1(\lambda),\cdots,f_m(\lambda)$ 的准素因子组是 A 的初等因子组, $g_1(\lambda),\cdots,g_n(\lambda)$ 的准素因子组是 B 的初等因子组, 因此 M 的初等因子组是 A,B 的初等因子组的无交并集, 于是 M 的 Jordan 标准型为 $\mathrm{diag}\{J_1,J_2\}$.

证法二: 由命题 6.51可知, 矩阵方程 $AX - XB = C$ 存在唯一解 $X = X_0$. 考虑如下相似变换:

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

因此 M 的 Jordan 标准型为 $\mathrm{diag}\{J_1,J_2\}$. □

例题 7.22 设 $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ b & a+1 & 0 & 0 \\ 3 & b & 2 & 0 \\ 5 & 4 & a & 2 \end{pmatrix}$, 求 A 的 Jordan 标准型.

解 显然, A 的特征值为 $1, a+1, 2, 2$. 对 A 进行分块 $A = \begin{pmatrix} A_{11} & O \\ A_{21} & A_{22} \end{pmatrix}$, 其中所有的分块都是二阶方阵. 下面按 $a+1$ 是否等于 $1,2$ 进行分类讨论.

(1) 若 $a \ne 0$ 及 $a \ne 1$, 则可有两种方法来处理. **方法一** (几何重数): 经计算可知特征值 2 的几何重数等于 1, 因此 A 的 Jordan 标准型为 $\mathrm{diag}\{1, a+1, J_2(2)\}$. **方法二** (命题 7.48): 显然 A_{11} 可对角化, A_{22} 不可对角化, 但 A_{22} 此时就是 Jordan 块 $J_2(2)$, 且 A_{11}, A_{22} 无公共特征值, 故可消去 A_{21}, 因此 A 的 Jordan 标准型为 $\mathrm{diag}\{1, a+1, J_2(2)\}$.

(2) 若 $a = 0$ 及 $b \ne 0$, 则利用**方法二** (命题 7.48) 可得, A 的 Jordan 标准型为 $\mathrm{diag}\{J_2(1), 2, 2\}$.

(3) 若 $a = 0$ 及 $b = 0$, 则利用**方法二** (命题 7.48) 可得, A 的 Jordan 标准型为 $\mathrm{diag}\{1, 1, 2, 2\}$.

(4) 若 $a = 1$ 及 $b \ne 0$, 则利用**方法一** (几何重数) 可得, A 的 Jordan 标准型为 $\mathrm{diag}\{1, J_3(2)\}$.

(5) 若 $a = 1$ 及 $b = 0$, 则利用**方法一** (几何重数) 可得, A 的 Jordan 标准型为 $\mathrm{diag}\{1, 2, J_2(2)\}$. □

7.14 过渡矩阵的求法

7.14.1 方法 1: 计算特征矩阵之间的相抵变换

推论 7.18

设 A 是 n 阶数字矩阵, $P(\lambda)$ 及 $Q(\lambda)$ 是同阶可逆 λ-矩阵, 且

$$Q(\lambda)(\lambda I_n - A)P(\lambda) = \lambda I_n - J,$$

其中 J 是 A 的 Jordan 标准型. 又由引理 7.2可知, 存在 λ 矩阵 $T(\lambda)$ 和数字矩阵 P, 使得

$$P(\lambda) = T(\lambda)(\lambda I_n - J) + P,$$

其中 P 是数字矩阵, 求证: $P^{-1}AP = J$.

注 这个结论就是定理 7.2的推论. 因此我们可以先计算出特征矩阵之间的相抵变换的过渡矩阵, 再由引理 7.2算出我们需要的数字矩阵的相似变换的过渡矩阵.

证明 由已知可得 $(\lambda I_n - A)P(\lambda) = Q(\lambda)^{-1}(\lambda I_n - J)$. 代入 $P(\lambda)$, 可得

$$(\lambda I_n - A)(T(\lambda)(\lambda I_n - J) + P) = Q(\lambda)^{-1}(\lambda I_n - J).$$

整理可得

$$(\lambda I_n - A)P = \left(Q(\lambda)^{-1} - (\lambda I_n - A)T(\lambda)\right)(\lambda I_n - J).$$

比较 λ 的次数可知, $Q(\lambda)^{-1} - (\lambda I_n - A)T(\lambda)$ 必须是数字矩阵, 记之为 R, 于是

$$(\lambda I_n - A)P = R(\lambda I_n - J).$$

去括号再次比较次数可得 $P = R, AP = RJ$. 若可证明 P 是可逆矩阵, 即有 $P^{-1}AP = J$. 由 $Q(\lambda)^{-1} - (\lambda I_n - A)T(\lambda) = R$ 可得

$$I_n = Q(\lambda)(\lambda I_n - A)T(\lambda) + Q(\lambda)R.$$

注意到 $Q(\lambda)(\lambda I_n - A) = (\lambda I_n - J)P(\lambda)^{-1}$, 故

$$I_n = (\lambda I_n - J)P(\lambda)^{-1}T(\lambda) + Q(\lambda)R.$$

由引理 7.2可知, 存在 λ 矩阵 $M(\lambda)$ 和数字矩阵 N, 使得

$$Q(\lambda) = (\lambda I_n - J)M(\lambda) + N,$$

于是

$$I_n = (\lambda I_n - J)\left(P(\lambda)^{-1}T(\lambda) + M(\lambda)R\right) + NR.$$

比较次数可得 $NR = I_n$, 即 R 可逆, 也即 P 可逆. □

7.14.2 方法 2: 计算特征向量和广义特征向量

在例题 7.23中, 任取 $(A - I)x = 0$ 的两个线性无关的解作为特征向量 α_1, α_3, 都可以解出对应的广义特征向量 α_2, α_4, 即线性方程组 $(A - I)x = \alpha_1$ 和 $(A - I)x = \alpha_3$ 的可解性不依赖于 α_1, α_3 的选取（请读者自行思考其中的原因）, 但这并非是普遍的情形. 一般来说, 我们总可以取到 $(A - \lambda_0 I)x = 0$ 的一个非零解 α_1（即特征值 λ_0 的特征向量）, 但若 α_1 选取不当, 线性方程组 $(A - \lambda_0 I)x = \alpha_1$ 有可能是无解的（即求不出对应的广义特征向量）. 因此在选取特征向量时, 需要我们仔细观察或设立参数, 这样才能保证最终得到正确的结果. 让我们来看下面两个例题中的具体分析.

例题 7.23 设复四维空间上的线性变换 φ 在基 $\{e_1, e_2, e_3, e_4\}$ 下的表示矩阵为

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

求一组新基, 使 φ 在这组新基下的表示矩阵是 A 的 Jordan 标准型, 并求过渡矩阵.

解 解法一: 通过计算可知 $\lambda I_4 - A$ 的法式为 $\mathrm{diag}\{1, 1, (\lambda-1)^2, (\lambda-1)^2\}$, 故 A 的初等因子组为 $(\lambda-1)^2, (\lambda-1)^2$, 从而 A 的 Jordan 标准型为

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

设过渡矩阵为 $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 则 $P^{-1}AP = J$, 即

$$AP = (A\alpha_1, A\alpha_2, A\alpha_3, A\alpha_4) = PJ = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)J$$

从而得到线性方程组:

$$(A - I)\alpha_1 = 0,\ (A - I)\alpha_2 = \alpha_1,\ (A - I)\alpha_3 = 0,\ (A - I)\alpha_4 = \alpha_3$$

求解 $(A-I)x=0$ 得到两个线性无关的解, 将它们分别作为 α_1 和 α_3:

$$\alpha_1=(1,3,0,0)',\ \alpha_3=(5,0,6,3)'$$

再求解方程组 $(A-I)x=\alpha_1,(A-I)x=\alpha_3$, 得到

$$\alpha_2=(\frac{1}{3},0,0,0)',\ \alpha_4=(\frac{7}{6},0,\frac{3}{2},0)'$$

因此过渡矩阵

$$P=\begin{pmatrix}1&\frac{1}{3}&5&\frac{7}{6}\\3&0&0&0\\0&0&6&\frac{3}{2}\\0&0&3&0\end{pmatrix}$$

新基为 $(f_1,f_2,f_3,f_4)=(e_1,e_2,e_3,e_4)P$, 即

$$f_1=e_1+3e_2,\ f_2=\frac{1}{3}e_1,\ f_3=5e_1+6e_3+3e_4,\ f_4=\frac{7}{6}e_1+\frac{3}{2}e_3$$

解法二: A 的初等因子组的计算同解法一, 可得 A 的 Jordan 标准型 $J=\mathrm{diag}\{J_2(1),J_2(1)\}$. 注意到 $(A-I_4)^2=O$ 且 $\mathrm{r}(A-I_4)=2$, 故可取 $A-I_4$ 的第 1 列和第 3 列作为其列向量的极大无关组. 因此 $e_1=(1,0,0,0)',e_3=(0,0,1,0)'$ 为广义特征向量, 使得 $(A-I_4)e_1=(3,9,0,0)',(A-I_4)e_3=(1,-7,4,2)'$ 为线性无关的特征向量, 则过渡矩阵 $P=((A-I_4)e_1,e_1,(A-I_4)e_3,e_3)$ 满足 $P^{-1}AP=J$. □

例题 7.24 设

$$A=\begin{pmatrix}2&6&-15\\1&1&-5\\1&2&-6\end{pmatrix}$$

求非异阵 P, 使 $P^{-1}AP$ 为 Jordan 标准型.

解 解法一: 通过计算可知 λI_3-A 的法式为 $\mathrm{diag}\{1,\lambda+1,(\lambda+1)^2\}$, 故 A 的初等因子组为 $\lambda+1,(\lambda+1)^2$, 从而 A 的 Jordan 标准型为

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

设非异阵 $P=(\alpha_1,\alpha_2,\alpha_3)$, 使 $P^{-1}AP=J$, 则 $AP=(A\alpha_1,A\alpha_2,A\alpha_3)=PJ=(\alpha_1,\alpha_2,\alpha_3)J$, 从而得到线性方程组:

$$(A+I_3)\alpha_1=0,\ (A+I_3)\alpha_2=0,\ (A+I_3)\alpha_3=\alpha_2$$

求解 $(A+I_3)x=0$ 得到两个线性无关的解 $\beta_1=(-2,1,0)'$ 和 $\beta_2=(5,0,1)'$. 注意到 $(A+I_3)x=\beta_i\ (i=1,2)$ 都是无解的, 故不能将 β_1 或 β_2 直接作为 α_2 来求广义特征向量 α_3. 一般地, 可设 $\alpha_2=k_1\beta_1+k_2\beta_2=(-2k_1+5k_2,k_1,k_2)'$, 代入 $(A+I_3)x=\alpha_2$ 中, 利用 $\mathrm{r}(A+I_3,\alpha_2)=\mathrm{r}(A+I_3)$ 可得 $k_1=k_2$. 因此, 可取 $\alpha_1=\beta_1=(-2,1,0)',\alpha_2=\beta_1+\beta_2=(3,1,1)'$, 此时可解出 $\alpha_3=(1,0,0)'$, 于是

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

解法二:A 的初等因子组的计算同解法一, 可得 A 的 Jordan 标准型 $J=\mathrm{diag}\{-1,J_2(-1)\}$. 注意到 $(A+I_3)^2=O$ 且 $\mathrm{r}(A+I_3)=1$, 故可取 $A+I_3$ 的第 1 列作为其列向量的极大无关组. 因此 $e_1=(1,0,0)'$ 为循环向量（即广义特征向量）, 使得 $e_1,(A+I_3)e_1=(3,1,1)'$ 构成了 $J_2(-1)$ 的循环轨道. 再取线性无关的特征向量 $\xi_1=(-2,1,0)'$, 则过渡矩阵 $P=(\xi_1,(A+I_3)e_1,e_1)$ 满足 $P^{-1}AP=J$. □

例题 7.25 设

$$A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

求非异阵 P, 使 $P^{-1}AP$ 为 Jordan 标准型.

解 解法一: 通过计算可知 $\lambda I_4 - A$ 的法式为 $\mathrm{diag}\{1, 1, \lambda - 1, (\lambda - 1)^3\}$, 故 A 的初等因子组为 $\lambda - 1, (\lambda - 1)^3$, 从而 A 的 Jordan 标准型为 $J = \mathrm{diag}\{1, J_3(1)\}$. 设非异阵 $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, 使 $P^{-1}AP = J$, 则 $AP = (A\alpha_1, A\alpha_2, A\alpha_3, A\alpha_4) = PJ = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)J$, 从而得到线性方程组:

$$(A - I_4)\alpha_1 = 0,\ (A - I_4)\alpha_2 = 0,\ (A - I_4)\alpha_3 = \alpha_2,\ (A - I_4)\alpha_4 = \alpha_3.$$

求解 $(A - I_4)x = 0$ 得到两个线性无关的解 $\beta_1 = (-1, 0, 1, 0)'$ 和 $\beta_2 = (0, 1, 0, 1)'$. 设 $\alpha_2 = k_1\beta_1 + k_2\beta_2$, 代入 $(A - I_4)x = \alpha_2$ 中, 利用 $\mathrm{r}(A - I_4, \alpha_2) = \mathrm{r}(A - I_4)$ 可得 $k_1 = 0$. 于是可取 $\alpha_2 = k_2\beta_2$, 解出 $\alpha_3 = k_2e_1 + k_3\beta_1 + k_4\beta_2$, 其中 $e_1 = (1, 0, 0, 0)'$. 再代入 $(A - I_4)x = \alpha_3$ 中, 利用 $\mathrm{r}(A - I_4, \alpha_3) = \mathrm{r}(A - I_4)$ 可得 $k_2 = 2k_3$. 于是可取 $k_2 = 2, k_3 = 1, k_4 = 0$, 最终得到特征向量 $\alpha_1 = \beta_1 = (-1, 0, 1, 0)', \alpha_2 = 2\beta_2 = (0, 2, 0, 2)'$, 1 级广义特征向量 $\alpha_3 = 2e_1 + \beta_1 = (1, 0, 1, 0)'$, 2 级广义特征向量 $\alpha_4 = (0, 0, 0, 1)'$, 从而

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

解法二: A 的初等因子组的计算同解法一, 可得 A 的 Jordan 标准型 $J = \mathrm{diag}\{1, J_3(1)\}$. 注意到 $(A - I_4)^3 = O$ 且 $\mathrm{r}((A - I_4)^2) = 1$, 故可取 $(A - I_4)^2$ 的第 4 列作为其列向量的极大无关组. 因此 $e_4 = (0, 0, 0, 1)'$ 为循环向量（即 2 级广义特征向量）, 使得 $e_4, (A - I_4)e_4 = (1, 0, 1, 0)', (A - I_4)^2e_4 = (0, 2, 0, 2)'$ 构成了 $J_3(1)$ 的循环轨道. 再取线性无关的特征向量 $\xi_1 = (-1, 0, 1, 0)'$, 则过渡矩阵 $P = (\xi_1, (A - I_4)^2e_4, (A - I_4)e_4, e_4)$ 满足 $P^{-1}AP = J$. □

7.14.3 方法 3: 计算循环子空间的循环向量

根据 Jordan 标准型的几何意义, 全空间可分解为不同特征值的根子空间的直和, 每个根子空间可分解为若干个循环子空间的直和, 每个循环子空间对应于一条循环轨道, 这条轨道由循环向量（即最高级的广义特征向量）生成. 下面以幂零根子空间为例, 说明如何确定所有的循环向量, 从而确定所有的基向量（等价于求过渡矩阵 P）.

例题 7.26 设 9 阶幂零矩阵 A 的 Jordan 标准型 $J = \mathrm{diag}\{0, J_2(0), J_3(0), J_3(0)\}$, 求非异阵 P, 使 $P^{-1}AP = J$.

注 这个例题采用的方法可以推广到一般的情形, 其原理是: 设 n 阶幂零矩阵 A 的极小多项式为 λ^k, 则依次选取第 i 级广义特征向量 ξ_i $(i = k - 1, \cdots, 0)$, 使得所有的 $A^i\xi_i$ $(i = k - 1, \cdots, 0)$ 在 $\mathrm{Ker}\,A$ 中线性无关即可.

解 由已知条件 $A^3 = O, \mathrm{r}(A^2) = 2$ 且 $\mathrm{r}(A) = 5$, 可设 $A^2x = 0$ 的基础解系为 $\{\eta_i, 1 \le i \le 7\}$. 由于 A^2 的列秩为 2, 故不妨设 A^2 的第 1 列和第 2 列是 A^2 列向量的极大无关组, 即 A^2e_1, A^2e_2 线性无关, 其中 e_1, e_2 是 9 维标准单位列向量的前两个. 考虑限制映射 $A|_{\mathrm{Ker}A^2} : \mathrm{Ker}A^2 \to \mathrm{Ker}A$, 容易验证 $\mathrm{Ker}(A|_{\mathrm{Ker}A^2}) = \mathrm{Ker}A, \mathrm{Im}(A|_{\mathrm{Ker}A^2}) = \mathrm{Ker}A \cap \mathrm{Im}A$. 由 $\dim \mathrm{Ker}A^2 = 7, \dim \mathrm{Ker}A = 4$ 及线性映射的维数公式可知 $\dim(\mathrm{Ker}A \cap \mathrm{Im}A) = 3$, 且 $\mathrm{Ker}A \cap \mathrm{Im}A = L(A\eta_i, 1 \le i \le 7)$. 注意到 A^2e_1, A^2e_2 是 $\mathrm{Ker}A \cap \mathrm{Im}A$ 中两个线性无关的向量, 故可从其生成元中取出一个向量, 不妨设为 $A\eta_1$, 使得 $A^2e_1, A^2e_2, A\eta_1$ 线性无关. 再次注意到 $\dim \mathrm{Ker}A = 4$, 且 $A^2e_1, A^2e_2, A\eta_1$ 是 $\mathrm{Ker}A$ 中 3 个线性无关的向量, 故可从其一组基（即 $Ax = 0$ 的基础解系）中取出一个向量 ξ_1, 使得 $A^2e_1, A^2e_2, A\eta_1, \xi_1$ 线性无关.

下面证明: $\{e_1, Ae_1, A^2e_1, e_2, Ae_2, A^2e_2, \eta_1, A\eta_1, \xi_1\}$ 构成 $\mathbb{C}^9$ 的一组基. 只要证明它们线性无关即可. 设 $c_1, \cdots, c_9 \in \mathbb{C}$, 使得

$$c_1e_1 + c_2Ae_1 + c_3A^2e_1 + c_4e_2 + c_5Ae_2 + c_6A^2e_2 + c_7\eta_1 + c_8A\eta_1 + c_9\xi_1 = 0 \tag{7.35}$$

将(7.35)式作用 A^2 可得

$$c_1A^2e_1+c_4A^2e_2=0$$

由 A^2e_1,A^2e_2 线性无关可知 $c_1=c_4=0$. 将(7.35)式作用 A 可得

$$c_2A^2e_1+c_5A^2e_2+c_7A\eta_1=0$$

由 $A^2e_1,A^2e_2,A\eta_1$ 线性无关可知 $c_2=c_5=c_7=0$.(7.35)式最后变成

$$c_3A^2e_1+c_6A^2e_2+c_8A\eta_1+c_9\xi_1=0$$

由 $A^2e_1,A^2e_2,A\eta_1,\xi_1$ 线性无关可知 $c_3=c_6=c_8=c_9=0$. 有了上面这组基, 我们可以把 4 个循环子空间的循环轨道全部确定如下:

$$\begin{array}{ccccccccc}
\text{轨道1} & & \text{轨道2} & & \text{轨道3} & & \text{轨道4} & & \\
 & & & & e_1 & ------ & e_2 & ------ & \text{2级广义特征向量}\\
 & & & & \downarrow & & \downarrow & & \\
 & & \eta_1 & ------ & Ae_1 & ------ & Ae_2 & ------ & \text{1级广义特征向量}\\
 & & \downarrow & & \downarrow & & \downarrow & & \\
\xi_1 & ------ & A\eta_1 & ------ & A^2e_1 & ------ & A^2e_2 & ------ & \text{0级广义特征向量}\\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
0 & & 0 & & 0 & & 0 & &
\end{array}$$

最后, 令 $P=(\xi_1,A\eta_1,\eta_1,A^2e_1,Ae_1,e_1,A^2e_2,Ae_2,e_2)$ 即为所求. □

例题 7.27 设

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

求非异阵 P, 使 $P^{-1}AP$ 为 Jordan 标准型.

注 下面的例题也与过渡矩阵有关, 它告诉我们: 满足基础矩阵乘法性质的矩阵类与基础矩阵类之间存在着一个相似变换. 利用这一结论可以证明:n 阶矩阵环 $M_n(\mathbb{K})$ **的任一自同构都是内自同构.**

解 经计算可知 A 的初等因子组为 $(\lambda+1)^2,(\lambda-1)^2$, 于是 A 的 Jordan 标准型为 $J=\mathrm{diag}\{J_2(-1),J_2(1)\}$. 由命题 6.52可知,$\mathbb{C}^4=\mathrm{Ker}(A+I_4)^2\oplus\mathrm{Ker}(A-I_4)^2$, 且 $\mathrm{Ker}(A+I_4)^2=\mathrm{Im}(A-I_4)^2$,$\mathrm{Ker}(A-I_4)^2=\mathrm{Im}(A+I_4)^2$. 经计算可取 $(A-I_4)^2$ 的第二列 $\alpha=(A-I_4)^2e_2=(16,20,0,0)'$ 作为根子空间 $\mathrm{Ker}(A+I_4)^2$ 中的循环向量（即广义特征向量）, 于是 $\alpha,(A+I_4)\alpha=(-16,-16,0,0)'$ 构成根子空间 $\mathrm{Ker}(A+I_4)^2$ 中的循环轨道. 经计算可取 $(A+I_4)^2$ 的第三列 $\beta=(A+I_4)^2e_3=(12,8,12,8)'$ 作为根子空间 $\mathrm{Ker}(A-I_4)^2$ 中的循环向量（即广义特征向量）, 于是 $\beta,(A-I_4)\beta=(8,8,8,8)'$ 构成根子空间 $\mathrm{Ker}(A-I_4)^2$ 中的循环轨道. 因此, 过渡矩阵 $P=((A+I_4)\alpha,\alpha,(A-I_4)\beta,\beta)$ 满足 $P^{-1}AP=J$. □

定理 7.26

设有 n^2 个 n 阶非零矩阵 $A_{ij}\ (1\le i,j\le n)$, 适合

$$A_{ij}A_{jk}=A_{ik},\ A_{ij}A_{lk}=O\ (j\neq l).$$

求证: 存在可逆矩阵 P, 使得对任意的 i,j,$P^{-1}A_{ij}P=E_{ij}$, 其中 E_{ij} 是基础矩阵.

证明 因为 $A_{11}\neq O$, 故存在 α, 使得 $A_{11}\alpha\neq 0$. 令 $\alpha_1=A_{11}\alpha$, 由 $A_{11}A_{11}=A_{11}$ 可得 $A_{11}\alpha_1=\alpha_1$. 再令 $\alpha_i=A_{i1}\alpha_1$, 由 $A_{i1}A_{i1}=A_{11}$ 可知 $\alpha_i\neq 0$. 我们得到了 n 个非零向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 由已知条件容易验证这 n 个向量适合下列性质:

$$A_{ij}\alpha_j=\alpha_i,\ A_{ij}\alpha_k=0\,(j\neq k)$$

由此不难证明这 n 个向量线性无关. 令 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 则 P 是可逆矩阵, 且

$$A_{ij}P=(A_{ij}\alpha_1,A_{ij}\alpha_2,\cdots,A_{ij}\alpha_n)=(0,\cdots,0,\alpha_i,0,\cdots,0).$$

其中上式中的 α_i 在第 j 列. 另一方面, 有

$$PE_{ij}=(\alpha_1,\alpha_2,\cdots,\alpha_n)E_{ij}=(0,\cdots,0,\alpha_i,0,\cdots,0).$$

因此, 对任意的 i,j,$A_{ij}P=PE_{ij}$, 即 $P^{-1}A_{ij}P=E_{ij}$. □

7.15 Jordan 标准型的应用

7.15.1 利用 Jordan 标准型研究矩阵的性质

命题 7.49

设 A 是 n 阶复矩阵, 求证:A 相似于分块对角矩阵 $\mathrm{diag}\{B,C\}$, 其中 B 是幂零矩阵,C 是可逆矩阵.

注 这个命题告诉我们: 在相似的意义下, 对复方阵的研究可归结为对幂零矩阵和可逆矩阵这两类特殊矩阵的研究, 它们的刻画分别是: 特征值全为零以及特征值全不为零. 这也是前面很多例题都处理这两类矩阵的深层次原因.

证明 我们发现 A 的初等因子分离开了零特征值和非零特征值, 从而 A 的 Jordan 标准型满足题目要求. 此时, 可将零特征值的 Jordan 块 $J_r(0)$(幂零矩阵)放入 B 中, 将非零特征值的 Jordan 块 $J_r(\lambda_0)$(可逆矩阵)放入 C 中, 即得结论. □

定理 7.27

设 λ_0 是 n 阶矩阵 A 的特征值, 其代数重数为 m. 设属于特征值 λ_0 的最大 Jordan 块的阶数为 k, 求证:

$$\mathrm{r}(A-\lambda_0I_n)>\cdots>\mathrm{r}((A-\lambda_0I_n)^k)=\mathrm{r}((A-\lambda_0I_n)^{k+1})=\cdots=n-m$$

证明 设 P 为可逆矩阵, 使 $P^{-1}AP=J=\mathrm{diag}\{J_{r_1}(\lambda_1),J_{r_2}(\lambda_2),\cdots,J_{r_s}(\lambda_s)\}$ 为 Jordan 标准型, 则对任意的正整数 j,

$$\mathrm{r}((A-\lambda_0I_n)^j)=\mathrm{r}(P^{-1}(A-\lambda_0I_n)^jP)=\mathrm{r}((J-\lambda_0I_n)^j)=\sum_{i=1}^{s}\mathrm{r}(J_{r_i}(\lambda_i-\lambda_0)^j)$$

若 $\lambda_i\neq\lambda_0$, 则 $\mathrm{r}(J_{r_i}(\lambda_i-\lambda_0)^j)=r_i$. 若 $\lambda_i=\lambda_0$, 则当 $1\leq j\leq r_i$ 时,$\mathrm{r}(J_{r_i}(0)^j)=r_i-j$; 当 $j\geq r_i$ 时,$\mathrm{r}(J_{r_i}(0)^j)=0$. 注意到 A 至少有一个 Jordan 块 $J_k(\lambda_0)$, 并且属于特征值 λ_0 的所有 Jordan 块阶数之和等于 m, 故当 $1\leq j\leq k$ 时,$\mathrm{r}((A-\lambda_0I_n)^j)$ 严格递减; 当 $j\geq k$ 时,$\mathrm{r}((A-\lambda_0I_n)^j)=n-m$. □

命题 7.50

设 λ_0 是 n 阶复矩阵 A 的特征值, 并且属于 λ_0 的初等因子都是次数大于等于 2 的多项式. 求证: 特征值 λ_0 的任一特征向量 α 均可表示为 $A-\lambda_0I_n$ 的列向量的线性组合.

注 **若特征值 λ_0 有一个初等因子为一次多项式, 则必存在特征向量 α, 它不能表示为 $A-\lambda_0I_n$ 的列向量的线性组合.** 证明的细节留给读者完成. 一个极端的例子就是 $A=I_n$, 其特征值 1 的初等因子都是一次的, 并且任一特征向量都不是 $A-I_n=O$ 的列向量的线性组合.命题 7.50与命题 7.37(可对角化的判定)有着密切的联系, 请读者思考两者之间的关系.

证明 由推论 7.12可知, 属于特征值 λ_0 的每个 Jordan 块的线性无关的特征向量均只有一个, 并且特征值 λ_0 的任一特征向量都是这些特征向量的线性组合, 因此我们只要证明 A 的 Jordan 标准型只含一个 Jordan 块 $J_n(\lambda_0)$ 的情形即可. 设 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 为非异阵, 使得 $P^{-1}AP=J_n(\lambda_0)$, 即 $AP=PJ_n(\lambda_0)$, 利用分块矩阵的乘法可得

$$A\alpha_1=\lambda_0\alpha_1,$$
$$A\alpha_2=\alpha_1+\lambda_0\alpha_2,$$
$$\cdots$$

$$A\alpha_n = \alpha_{n-1} + \lambda_0\alpha_n$$

注意到 $n \geq 2$, 故有 $\alpha_1 = (A - \lambda_0 I_n)\alpha_2$, 从而特征向量 α_1 可表示为 $A - \lambda_0 I_n$ 的列向量的线性组合. □

7.15.2 运用 Jordan 标准型进行相似问题的化简

命题 7.51

设 A, B 为 n 阶矩阵, 满足 $AB = BA = O$,$\mathrm{r}(A) = \mathrm{r}(A^2)$, 求证:$\mathrm{r}(A + B) = \mathrm{r}(A) + \mathrm{r}(B)$.

证明 注意到问题的条件和结论在同时相似变换:$A \mapsto P^{-1}AP$,$B \mapsto P^{-1}BP$ 下不改变, 故不妨从一开始就假设 A 为 Jordan 标准型. 设 $A = \mathrm{diag}\{A_0, A_1\}$, 其中 A_0 由零特征值的 Jordan 块构成,A_1 由非零特征值的 Jordan 块构成. 由 $\mathrm{r}(A) = \mathrm{r}(A^2)$ 可知, 零特征值的 Jordan 块都是一阶的否则,$\mathrm{r}(A_0^2) < \mathrm{r}(A_0)$. 即 $A_0 = O$. 将 B 进行对应的分块 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$, 则由 $AB = BA = O$ 以及 A_1 非异可知,B_{12},B_{21} 和 B_{22} 都是零矩阵. 于是

$$\mathrm{r}(A + B) = \mathrm{r}\begin{pmatrix} B_{11} & O \\ O & A_1 \end{pmatrix} = \mathrm{r}(B_{11}) + \mathrm{r}(A_1) = \mathrm{r}(B) + \mathrm{r}(A).$$

□

命题 7.52

设 A, B 分别是 m, n 阶矩阵, 求证: 矩阵方程 $AX = XB$ 只有零解的充要条件是 A, B 无公共的特征值.

证明 先做两步化简. 注意到问题的条件和结论在矩阵变换:$B \mapsto P^{-1}BP$, $X \mapsto XP$ 下不改变, 故不妨从一开始就假设 B 为 Jordan 标准型. 设 $X = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ 为列分块, 则有

$$AX = (A\alpha_1, A\alpha_2, \cdots, A\alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)B = XB. \tag{7.36}$$

若 B 有 k 个 Jordan 块, 则方程组(7.36) 可分解为 k 个独立方程组. 注意到:

(i) 方程组(7.36)只有零解当且仅当这 k 个独立方程组都只有零解;

(ii) 方程组(7.36)有非零解当且仅当这 k 个独立方程组中至少有一个有非零解.

因此, 不妨进一步假设 $B = J_n(\lambda_0)$ 为 Jordan 块. 此时, 方程组(7.36)等价于下列方程组:

$$A\alpha_1 = \lambda_0\alpha_1,\ A\alpha_2 = \alpha_1 + \lambda_0\alpha_2,\ \cdots,\ A\alpha_n = \alpha_{n-1} + \lambda_0\alpha_n.$$

充分性: 假设 A, B 没有公共的特征值, 则 λ_0 不是 A 的特征值, 从而由 $A\alpha_1 = \lambda_0\alpha_1$ 只能得到 $\alpha_1 = \mathbf{0}$. 代入第二个方程可得 $A\alpha_2 = \lambda_0\alpha_2$, 相同的理由可推出 $\alpha_2 = \mathbf{0}$. 不断这样做下去, 最后可得 $\alpha_i = \mathbf{0}(1 \leq i \leq n)$, 即 $X = O$, 从而矩阵方程 $AX = XB$ 只有零解.

必要性: 假设 A 和 $B = J_n(\lambda_0)$ 有公共的特征值 λ_0, 在上述方程组中令 $\alpha_1 = \cdots = \alpha_{n-1} = \mathbf{0}$. 因为 λ_0 也是 A 的特征值, 所以 $A\alpha_n = \lambda_0\alpha_n$ 有非零解 $\alpha_n = \alpha$, 于是 $X_0 = (\mathbf{0}, \cdots, \mathbf{0}, \alpha)$ 是上述方程组的非零解, 从而矩阵方程 $AX = XB$ 有非零解, 矛盾! □

命题 7.53

设 A, B 分别是 m, n 阶矩阵,C 是 $m \times n$ 矩阵, 求证: 矩阵方程 $AX - XB = C$ 存在唯一解的充要条件是 A, B 无公共的特征值.

证明 先做两步化简. 注意到问题的条件和结论在矩阵变换:$B \mapsto P^{-1}BP$, $C \mapsto CP$, $X \mapsto XP$ 下不改变, 故不妨从一开始就假设 B 为 Jordan 标准型. 设 $X = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, $C = (\beta_1, \beta_2, \cdots, \beta_n)$ 为列分块, 则 $AX - XB = C$ 即为:

$$(A\alpha_1, A\alpha_2, \cdots, A\alpha_n) - (\alpha_1, \alpha_2, \cdots, \alpha_n)B = (\beta_1, \beta_2, \cdots, \beta_n). \tag{7.37}$$

若 B 有 k 个 Jordan 块, 则方程组 (7.37) 可分解为 k 个独立方程组. 注意到:

(i) 方程组 (7.37) 无解当且仅当这 k 个独立方程组中至少有一个无解;

(ii) 方程组 (7.37) 有唯一解当且仅当这 k 个独立方程组都只有唯一解;

(iii) 方程组 (7.37) 有无穷个解当且仅当这 k 个独立方程组都有解, 且至少有一个有无穷个解.

进一步, 若假设 $B = J_n(\lambda_0)$ 为 Jordan 块, 则方程组 (7.37) 等价于下列方程组: $(A-\lambda_0 I_n)\alpha_1 = \beta_1, (A-\lambda_0 I_n)\alpha_2 = \alpha_1+\beta_2, \cdots, (A-\lambda_0 I_n)\alpha_n = \alpha_{n-1}+\beta_n$.

充分性: 假设 A, B 没有公共的特征值, 则 λ_0 不是 A 的特征值, 从而 $A-\lambda_0 I_n$ 是可逆矩阵. 从第一个方程可解得 $\alpha_1 = (A-\lambda_0 I_n)^{-1}\beta_1$, 代入第二个方程可解得 $\alpha_2 = (A-\lambda_0 I_n)^{-1}(\alpha_1+\beta_2),\cdots$, 代入最后一个方程可解得 $\alpha_n = (A-\lambda_0 I_n)^{-1}(\alpha_{n-1}+\beta_n)$, 从而上述方程组有唯一解, 因此矩阵方程 $AX - XB = C$ 也有唯一解.

必要性: 假设 A, B 有公共的特征值 λ_0, 若这 k 个独立方程组中有一个无解, 则矩阵方程 $AX - XB = C$ 无解, 从而结论成立. 若这 k 个独立方程组都有解, 则不妨设 $B = J_n(\lambda_0)$ 为 Jordan 块. 由于 λ_0 是 A 的特征值, 故 $(A-\lambda_0 I_n)x = 0$ 有无穷个解. 注意到, 若 $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ 是上述方程组的一个解, 则对 $(A-\lambda_0 I_n)x = 0$ 的任一解 α_0,$(\alpha_1, \alpha_2, \cdots, \alpha_n+\alpha_0)$ 也是上述方程组的解, 因此矩阵方程 $AX - XB = C$ 有无穷个解, 矛盾! □

命题 7.54

设 A, B 分别是 m, n 阶矩阵,$M_{m\times n}(\mathbb{C})$ 上的线性变换 φ 定义为 $\varphi(X) = AX - XB$, 则下列 3 个结论等价:

(1) φ 是单映射;

(2) φ 是自同构;

(3) 对某个给定的 $m\times n$ 矩阵 C, 存在唯一的 X_0, 使得 $\varphi(X_0) = C$.

注 由这个命题可知,命题 7.52和命题 7.53都等价于命题 6.51, 并且命题 7.52和命题 7.53都给出了它们的 Jordan 标准型证法.

证明 事实上,(1) $\Rightarrow$ (2) 由推论 4.3立得. (2) $\Rightarrow$ (3) 显然成立. 用反证法来证明 (3) $\Rightarrow$ (1): 若 $\mathrm{Ker}\varphi \neq 0$, 则 $\mathrm{Ker}\varphi$ 中任一非零元 X_1 都满足 $\varphi(X_0+X_1) = C$, 这与唯一性矛盾. □

7.15.3 应用 Jordan 标准型的三段论法

如果矩阵问题的条件和结论在相似关系下不改变, 则可以先证明结论对 Jordan 块成立, 再证明对 Jordan 标准型成立, 最后证明对一般的矩阵也成立, 这就是所谓的“三段论法”. 事实上, 我们已经利用三段论法证明过例题 7.18和例题 7.19, 下面再来看一些典型的例题.

首先, 我们来看计算矩阵乘幂的问题. 设 A 为 n 阶矩阵,P 为 n 阶可逆矩阵, 使得

$$P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_k}(\lambda_k)\}$$

为 Jordan 标准型. 注意任一 Jordan 块 $J_{r_i}(\lambda_i)$ 都有分解 $J_{r_i}(\lambda_i) = \lambda_i I_{r_i} + N$, 其中 $N = J_{r_i}(0)$ 是特征值为零的 r_i 阶 Jordan 块, 故对任意的正整数 m,

$$J_{r_i}(\lambda_i)^m = (\lambda_i I_{r_i} + N)^m = \lambda_i^m I_{r_i} + \mathrm{C}_m^1\lambda_i^{m-1}N + \cdots + \mathrm{C}_m^{m-1}\lambda_i N^{m-1} + N^m.$$

于是 $J^m = \mathrm{diag}\{J_{r_1}(\lambda_1)^m, J_{r_2}(\lambda_2)^m, \cdots, J_{r_k}(\lambda_k)^m\}$, 从而 $A^m = (PJP^{-1})^m = PJ^mP^{-1}$ 便可计算出来了.

例题 7.28 设 $A = \begin{pmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{pmatrix}$, 求 $A^m (m \geq 1)$.

注 本题是例题 7.24的延拓.

解 由例题 7.24, 我们已经计算出过渡矩阵 P, 使得 $P^{-1}AP = J = \mathrm{diag}\{-1, J_2(-1)\}$, 于是可进一步计算出

$$A^m = PJ^mP^{-1} = \begin{pmatrix} -2 & 3 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} (-1)^m & 0 & 0 \\ 0 & (-1)^m & (-1)^{m-1}m \\ 0 & 0 & (-1)^m \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & 2 & -5 \end{pmatrix}$$

$$= (-1)^{m-1}\begin{pmatrix} 3m-1 & 6m & -15m \\ m & 2m-1 & -5m \\ m & 2m & -5m-1 \end{pmatrix}.$$

□

例题 7.29 求矩阵 B, 使得 $A = B^2$, 其中 $A = \begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix}$.

解 利用过渡矩阵的求法的方法, 可求出过渡矩阵 $P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, 使得 $P^{-1}AP = J = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$ 为 Jordan 标准型. 用待定元素法不难求得 $C = \pm \begin{pmatrix} 2 & \frac{1}{4} \\ 0 & 2 \end{pmatrix}$, 使得 $C^2 = J$.

注意到 $(PCP^{-1})^2 = PC^2P^{-1} = PJP^{-1} = A$, 故可取 $B = PCP^{-1}$. 经计算可得

$$B = \pm \begin{pmatrix} \frac{7}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{9}{4} \end{pmatrix}.$$

□

定义 7.15 (非异阵的 k 次方根)

设 A 为 n 阶非异复矩阵, 称 B 为 A 的 ***k* 次方根**, 若 B 满足 $A = B^k$, 其中 $k \in \mathbb{N}_1$.

定理 7.28 (非异阵存在任意次的方根)

设 A 为 n 阶非异复矩阵, 证明: 对任一正整数 m, 存在 n 阶复矩阵 B, 使得 $A = B^m$.

注 定理 7.28的结论对奇异矩阵一般并不成立. 例如, 设 $A = J_n(0)^{m-1}$, 其中 $n = mq - r$, $m \geq 2$ 且 $0 \leq r < m$, 则不存在 B, 使得 $A = B^m$. 我们用反证法来证明这个结论. 若存在满足条件的 B, 则 B 的特征值全为零, 从而 B 也是幂零矩阵, 即有 $B^n = O$. 于是 $O = B^{n+r} = (B^m)^q = A^q = J_n(0)^{(m-1)q} \neq O$, 这就导出了矛盾.

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_k}(\lambda_k)\}$ 为 A 的 Jordan 标准型. 由于 A 非异, 故 A 的所有特征值都非零. 对 A 的任一 Jordan 块 $J_{r_i}(\lambda_i)$, 取定 λ_i 的某个 m 次方根 μ_i, 即 $\mu_i^m = \lambda_i$, 则由命题 7.46可知,$J_{r_i}(\mu_i)^m$ 相似于 $J_{r_i}(\lambda_i)$, 即存在非异阵 Q_i, 使得 $J_{r_i}(\lambda_i) = Q_i^{-1}J_{r_i}(\mu_i)^m Q_i = (Q_i^{-1}J_{r_i}(\mu_i)Q_i)^m$, 即结论对 Jordan 块成立. 令

$$C = \mathrm{diag}\{Q_1^{-1}J_{r_1}(\mu_1)Q_1, Q_2^{-1}J_{r_2}(\mu_2)Q_2, \cdots, Q_k^{-1}J_{r_k}(\mu_k)Q_k\},$$

则 $J = C^m$, 即结论对 Jordan 标准型也成立. 最后,

$$A = PJP^{-1} = PC^mP^{-1} = (PCP^{-1})^m.$$

令 $B = PCP^{-1}$, 则有 $A = B^m$, 即结论对一般的矩阵也成立. □

引理 7.12 (Jordan 块的常见分解)

$$J_n(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix} = \begin{pmatrix} & & & 1 & \lambda_i \\ & & \iddots & \lambda_i & \\ & 1 & \iddots & & \\ 1 & \lambda_i & & & \\ \lambda_i & & & & \end{pmatrix} \begin{pmatrix} & & & & 1 \\ & & & \iddots & \\ & & \iddots & & \\ & 1 & & & \\ 1 & & & & \end{pmatrix}$$

$$= \begin{pmatrix} & & & & 1 \\ & & & \iddots & \\ & & \iddots & & \\ & 1 & & & \\ 1 & & & & \end{pmatrix} \begin{pmatrix} & & & & \lambda_i \\ & & & \lambda_i & 1 \\ & & \iddots & \iddots & \\ & \lambda_i & 1 & & \\ \lambda_i & 1 & & & \end{pmatrix}$$

命题 7.55

设 A 为 n 阶复矩阵, 证明: 存在 n 阶复对称矩阵 B, C, 使得 $A = BC$, 并且可以指定 B, C 中任何一个为可逆矩阵.

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_k}(\lambda_k)\}$ 为 A 的 Jordan 标准型. 考虑 Jordan 块 $J_{r_i}(\lambda_i)$ 的如下两种分解 (列倒排和行倒排):

$$\begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix} = \begin{pmatrix} & & & 1 & \lambda_i \\ & & \iddots & \lambda_i & \\ & 1 & \iddots & & \\ 1 & \lambda_i & & & \\ \lambda_i & & & & \end{pmatrix} \begin{pmatrix} & & & & 1 \\ & & & \iddots & \\ & & \iddots & & \\ & 1 & & & \\ 1 & & & & \end{pmatrix} \tag{7.38}$$

$$= \begin{pmatrix} & & & & 1 \\ & & & \iddots & \\ & & \iddots & & \\ & 1 & & & \\ 1 & & & & \end{pmatrix} \begin{pmatrix} & & & & \lambda_i \\ & & & \lambda_i & 1 \\ & & \iddots & \iddots & \\ & \lambda_i & 1 & & \\ \lambda_i & 1 & & & \end{pmatrix} \tag{7.39}$$

我们将 (7.38) 式的分解记为 $J_{r_i}(\lambda_i) = R_i S_i$,(7.39) 式的分解记为 $J_{r_i}(\lambda_i) = S_i T_i$, 注意到 R_i, S_i, T_i 都是对称矩阵, 并且 S_i 是可逆矩阵. 如果一开始选定 B 为可逆矩阵, 则利用 (7.39) 式的分解; 如果一开始选定 C 为可逆矩阵, 则利用 (7.38) 式的分解. 以下不妨设定 B 为可逆矩阵, 令

$$S = \mathrm{diag}\{S_1, S_2, \cdots, S_k\}, \quad T = \mathrm{diag}\{T_1, T_2, \cdots, T_k\}$$

则有 $J = ST$, 其中 S, T 都是对称矩阵, 并且 S 是可逆矩阵. 因此, 我们有

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

令 $B = PSP'$, $C = (P^{-1})'TP^{-1}$, 则 $A = BC$ 即为所求分解. □

命题 7.56

设 A 为 n 阶复矩阵, 证明: 存在 n 阶非异复对称矩阵 Q, 使得 $Q^{-1}AQ = A'$.

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \mathrm{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_k}(\lambda_k)\}$ 为 A 的 Jordan 标准型. 记 r_i 阶矩阵

$$S_i = \begin{pmatrix} & & & & 1 \\ & & & \iddots & \\ & & \iddots & & \\ & 1 & & & \\ 1 & & & & \end{pmatrix},$$

注意到 S_i 是非异对称矩阵, 且 $S_i^2 = I$, 即 $S_i^{-1} = S_i$, 我们来考虑 Jordan 块 $J_{r_i}(\lambda_i)$ 的如下相似关系:

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

即有 $J_{r_i}(\lambda_i) = S_i J_{r_i}(\lambda_i)' S_i$. 令 $S = \mathrm{diag}\{S_1, S_2, \cdots, S_k\}$, 则 S 是非异对称矩阵,$S^2 = I_n$, 且 $J = SJ'S$. 因此, 我们有

$$A = PJP^{-1} = PSJ'SP^{-1} = PSP'A'(P^{-1})'SP^{-1} = (PSP')A'(PSP')^{-1}$$

令 $Q=PSP'$, 则 Q 为非异复对称矩阵, 使得 $Q^{-1}AQ=A'$. □

命题 7.57

设 A 为 n 阶幂零矩阵, 证明:e^A 与 I_n+A 相似.

证明 证法一:由 A 是幂零矩阵及命题 6.8可知,A 的特征值全为零. 设 P 为非异阵, 使得 $P^{-1}AP=J=\mathrm{diag}\,\{J_{r_1}(0)\,,J_{r_2}(0)\,,\cdots,J_{r_k}(0)\,\}$ 为 A 的 Jordan 标准型. 先对 Jordan 块 $J_{r_i}(0)$ 进行证明. 注意到

$$\begin{aligned}e^{J_{r_i}(0)}&=I_{r_i}+\frac{1}{1!}J_{r_i}(0)+\frac{1}{2!}J_{r_i}(0)^2+\cdots+\frac{1}{(r_i-1)!}J_{r_i}(0)^{r_i-1}\\&=\begin{pmatrix}1&1&\cdots&*\\&1&\ddots&\vdots\\&&\ddots&1\\&&&1\end{pmatrix}\end{aligned}$$

故 $e^{J_{r_i}(0)}$ 的特征值全为 1, 其几何重数等于 $r_i-\mathrm{r}(e^{J_{r_i}(0)}-I_{r_i})=r_i-(r_i-1)=1$. 因此 $e^{J_{r_i}(0)}$ 只有一个 Jordan 块, 其 Jordan 标准型为 $J_{r_i}(1)=I_{r_i}+J_{r_i}(0)$, 即存在非异阵 Q_i, 使得 $e^{J_{r_i}(0)}=Q_i(I_{r_i}+J_{r_i}(0))Q_i^{-1}$ $(1\le i\le k)$. 再对 Jordan 标准型 J 进行证明. 令 $Q=\mathrm{diag}\{Q_1,Q_2,\cdots,Q_k\}$, 则 Q 为非异阵, 满足

$$e^J=\mathrm{diag}\{e^{J_{r_1}(0)},e^{J_{r_2}(0)},\cdots,e^{J_{r_k}(0)}\}=Q(I_n+J)Q^{-1}$$

最后对一般的矩阵 A 进行证明. 由前两步可得

$$\begin{aligned}e^A&=e^{PJP^{-1}}=Pe^JP^{-1}=PQ(I_n+J)Q^{-1}P^{-1}\\&=PQ(I_n+P^{-1}AP)Q^{-1}P^{-1}=(PQP^{-1})(I_n+A)(PQP^{-1})^{-1}\end{aligned}$$

即 e^A 与 I_n+A 相似.

证法二: 由 A 是幂零矩阵及命题 6.8可知,A 的特征值全为零, 从而 I_n+A 和 $e^A=I_n+A+\cdots+\dfrac{1}{(n-1)!}A^{n-1}$ 的特征值全为 1. 容易验证 $\mathrm{r}((e^A-I_n)^k)=\mathrm{r}(A^k)$ $(k\ge 1)$ 成立, 故由命题 7.45即得结论. □

7.15.4 采用 Jordan 块作为测试矩阵

在矩阵问题中, 如果需要构造满足某种性质的矩阵, 则可以采用 Jordan 块作为测试矩阵进行探索和讨论. 比如在定理 7.28的证明中, 为了构造 Jordan 块 $J_{r_i}(\lambda_i)$ 的 m 次方根, 我们采用了 Jordan 块 $J_{r_i}(\mu_i)$ 作为测试矩阵, 并最终得到了正确的答案.

例题 7.30 证明: 存在 71 阶实方阵 A, 使得

$$A^{70}+A^{69}+\cdots+A+I_{71}=\begin{pmatrix}2019&2018&\cdots&1949\\&2019&\ddots&\vdots\\&&\ddots&2018\\&&&2019\end{pmatrix}$$

证明 记 $f(x)=x^{70}+x^{69}+\cdots+x+1$, 上述等式右边的矩阵为 B. 注意到 $f(1)<2019$ 和 $f(2)>2019$, 故由连续函数的性质可知,$f(x)=2019$ 在开区间 $(1,2)$ 中必有一实根 λ_0. 将 Jordan 块 $J_{71}(\lambda_0)$ 代入 $f(x)$ 中, 由引理 7.10计算可得

$$f(J_{71}(\lambda_0))=\begin{pmatrix}f(\lambda_0)&f'(\lambda_0)&\cdots&*\\&f(\lambda_0)&\ddots&\vdots\\&&\ddots&f'(\lambda_0)\\&&&f(\lambda_0)\end{pmatrix}$$

这是一个上三角矩阵, 主对角元全为 $f(\lambda_0)=2019$, 上次对角元全为 $f'(\lambda_0)>0$, 从而 $f(J_{71}(\lambda_0))$ 的特征值全为 2019, 其几何重数为 $71-\mathrm{r}(f(J_{71}(\lambda_0))-2019I_{71})=1$. 因此,$f(J_{71}(\lambda_0))$ 的 Jordan 标准型中只有一个 Jordan 块 $J_{71}(2019)$, 即

$f(J_{71}(\lambda_0))$ 相似于 $J_{71}(2019)$. 另一方面, 矩阵 B 也是一个上三角矩阵, 主对角元全为 2019, 上次对角元全为 2018, 从而 B 的特征值全为 2019, 其几何重数为 $71 - \mathrm{r}(B - 2019I_{71}) = 1$. 因此,$B$ 的 Jordan 标准型中只有一个 Jordan 块 $J_{71}(2019)$, 即 B 也相似于 $J_{71}(2019)$. 由于矩阵的相似在基域扩张下不改变, 故 $f(J_{71}(\lambda_0))$ 和 B 在实数域上相似, 即存在非异实矩阵 P, 使得 $B = P^{-1}f(J_{71}(\lambda_0))P = f(P^{-1}J_{71}(\lambda_0)P)$. 令 $A = P^{-1}J_{71}(\lambda_0)P$, 则 A 是实矩阵, 且满足 $f(A) = B$. □

例题 7.31 设 a, b 都是实数, 其中 $b \neq 0$, 证明: 对任意的正整数 m, 存在四阶实方阵 A, 使得

$$A^m = B = \begin{pmatrix} a & b & 2 & 0 \\ -b & a & 2 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & -b & a \end{pmatrix}$$

注 这个例题采用了广义 Jordan 块作为测试矩阵.

证明 显然,B 的特征多项式 $f(\lambda) = ((\lambda - a)^2 + b^2)^2$. 我们可用 3 种方法求出 B 的 Jordan 标准型 (section:Jordan 标准型的求法). 第一种方法是计算行列式因子:

$$\lambda I_4 - B = \begin{pmatrix} \lambda - a & -b & -2 & 0 \\ b & \lambda - a & -2 & 0 \\ 0 & 0 & \lambda - a & -b \\ 0 & 0 & b & \lambda - a \end{pmatrix}$$

经计算可知

$$(\lambda I_4 - B)\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \end{pmatrix} = -b((\lambda - a)^2 + b^2), \quad (\lambda I_4 - B)\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} = -2b(\lambda - a + b)$$

显然这两个三阶子式互素, 故三阶行列式因子 $D_3(\lambda) = 1$, 于是 B 的行列式因子组和不变因子组均为 $1, 1, 1, ((\lambda - a)^2 + b^2)^2$, 从而初等因子组为 $(\lambda - a - bi)^2, (\lambda - a + bi)^2$, 因此 B 的 Jordan 标准型 $J = \mathrm{diag}\{J_2(a + bi), J_2(a - bi)\}$.

第二种方法是计算极小多项式: 由于 B 是实方阵, 故其极小多项式 $m(\lambda)$ 是实系数多项式, 又 $m(\lambda)$ 整除 $f(\lambda)$, 从而只能是 $m(\lambda) = (\lambda - a)^2 + b^2$ 或 $m(\lambda) = ((\lambda - a)^2 + b^2)^2$. 通过简单的计算可知 B 不适合多项式 $(\lambda - a)^2 + b^2$, 于是 $m(\lambda) = f(\lambda) = ((\lambda - a)^2 + b^2)^2$, 剩余的讨论同第一种方法.

第三种方法是计算特征值的几何重数:B 的全体特征值为 $a + bi$（2 重）,$a - bi$（2 重）, 通过简单的计算可知 $\mathrm{r}(B - (a + bi)I_4) = 3$ 以及 $\mathrm{r}(B - (a - bi)I_4) = 3$, 于是 $a \pm bi$ 的几何重数都等于 1, 从而分别只有一个二阶 Jordan 块, 因此 B 的 Jordan 标准型 $J = \mathrm{diag}\{J_2(a + bi), J_2(a - bi)\}$.

取 $a + bi$ 的 m 次方根 $c + di$ $(c, d \in \mathbb{R})$, 即满足 $(c + di)^m = a + bi$（取定一个即可）. 构造实方阵（取法不唯一）:

$$C = \begin{pmatrix} c & d & 2 & 0 \\ -d & c & 2 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & -d & c \end{pmatrix}, \text{或} C = \begin{pmatrix} c & d & 1 & 0 \\ -d & c & 0 & 1 \\ 0 & 0 & c & d \\ 0 & 0 & -d & c \end{pmatrix}, \text{或} C = \begin{pmatrix} c & d & 0 & 0 \\ -d & c & 1 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & -d & c \end{pmatrix}$$

注意到 $d \neq 0$, 故由开始处完全类似的讨论可知,C 的 Jordan 标准型为 $\mathrm{diag}\{J_2(c + di), J_2(c - di)\}$. 由命题 7.46可知,$J_2(c \pm di)^m$ 的 Jordan 标准型为 $J_2(a \pm bi)$, 从而 C^m 的 Jordan 标准型为 $\mathrm{diag}\{J_2(a + bi), J_2(a - bi)\}$, 于是 B 与 C^m 有相同的 Jordan 标准型, 故它们在复数域上相似. 注意到 B 与 C^m 都是实矩阵, 故由矩阵的相似在基域扩张下不改变可知, 它们在实数域上也相似, 即存在非异阵 $P \in M_4(\mathbb{R})$, 使得 $B = P^{-1}C^mP = (P^{-1}CP)^m$. 令 $A = P^{-1}CP$, 则 A 为实方阵, 满足 $A^m = B$. □

7.16 Jordan 标准型的几何

7.16.1 Jordan 标准型的几何构造

7.17 一般数域上的 Jordan 标准型

遇到数域 $\mathbb{K}$ 上的相似问题, 一般来说, 可以有 3 种处理方法.

第一种方法是先将问题转化成几何语言, 再利用线性变换理论进行研究;

第二种方法是先将问题转化成代数语言, 再把数域 $\mathbb{K}$ 上的矩阵自然地看成是复矩阵进行研究, 最后利用高等代数中若干概念在基域扩张下的不变性 (例如,相似在基域扩张下的不变性) 将所得结果返回到数域 $\mathbb{K}$ 上;

第三种方法是利用一般数域上基于初等因子的相似标准型理论对问题进行研究.

命题 7.58

设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,φ 是 V 上秩小于 n 的线性变换, 求证:$V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$ 的充要条件是 0 是 φ 的极小多项式的单根.

笔记 分析: 当 $\mathbb{K} = \mathbb{C}$ 时, 可以利用 Jordan 标准型理论进行证明. 若特征值 0 是 φ 的极小多项式的单根, 则可设 φ 的初等因子组为 $\lambda, \cdots, \lambda, (\lambda-\lambda_1)^{r_1}, \cdots, (\lambda-\lambda_s)^{r_s}$, 其中 $\lambda_1, \cdots, \lambda_s$ 是非零特征值, 且有 k 个 λ. 因此, 存在 V 的一组基 $e_1, \cdots, e_k, e_{k+1}, \cdots, e_n$, 使得 φ 在这组基下的表示矩阵为 $\mathrm{diag}\{0, \cdots, 0, J_{r_1}(\lambda_1), \cdots, J_{r_s}(\lambda_s)\}$. 容易验证 $\mathrm{Ker}\varphi = L(e_1, \cdots, e_k)$,$\mathrm{Im}\varphi = L(e_{k+1}, \cdots, e_n)$, 于是 $V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$. 反之, 若 $V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$, 则 $\mathrm{Ker}\varphi \cap \mathrm{Im}\varphi = 0$, 由命题 7.37的充分性的证明可知,$\varphi$ 关于特征值 0 的 Jordan 块都是一阶的, 因此 0 是 φ 的极小多项式的单根. 然而, 当 $\mathbb{K} \neq \mathbb{C}$ 时, 上述讨论就不再适用了, 并且本题的结论也不能简单地延拓到复数域上, 因为 $V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$ 是数域 $\mathbb{K}$ 上线性空间的直和分解, 一般并不能看成是复数域上线性空间的直和分解. 接下去让我们来看前两种方法是如何巧妙地解决问题的.

证明 证法一:若 $V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$, 则由命题 4.22可知, $\mathrm{Ker}\varphi = \mathrm{Ker}\varphi^2 = \cdots$. 设 φ 的极小多项式 $m(\lambda) = \lambda^k g(\lambda)$, 其中 $g(0) \neq 0$, 我们来证明 $k = 1$. 用反证法, 假设 $k \geqslant 2$, 则对任意的 $\alpha \in V$, 有 $\varphi^k g(\varphi)(\alpha) = 0$, 从而 $g(\varphi)(\alpha) \in \mathrm{Ker}\varphi^k = \mathrm{Ker}\varphi$, 于是 $\varphi g(\varphi)(\alpha) = 0$ 对任意的 $\alpha \in V$ 成立, 即 $\varphi g(\varphi) = 0$, 因此 φ 适合多项式 $\lambda g(\lambda)$, 其次数比极小多项式的次数还小, 这就导出了矛盾. 反之, 设 φ 的极小多项式 $m(\lambda) = \lambda g(\lambda)$, 其中 $g(0) \neq 0$, 则由命题??可知,$V = V_1 \oplus V_2$, 其中 $V_1 = \mathrm{Ker}\varphi = \mathrm{Im}g(\varphi)$,$V_2 = \mathrm{Ker}g(\varphi) = \mathrm{Im}\varphi$, 于是 $V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$.

证法二:由命题 4.22可知,$V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$ 当且仅当 $\mathrm{r}(\varphi) = \mathrm{r}(\varphi^2)$, 因此我们只要证明: $\mathrm{r}(\varphi) = \mathrm{r}(\varphi^2)$ 当且仅当 0 是 φ 的极小多项式的单根. 任取 φ 在某组基下的表示矩阵 A, 则上述问题的代数版本是: $\mathrm{r}(A) = \mathrm{r}(A^2)$ 当且仅当 0 是 A 的极小多项式的单根. 注意到数域 $\mathbb{K}$ 上的矩阵可自然地看成是复矩阵, 并且矩阵的秩和极小多项式在基域扩张下不改变, 因此我们可以把 A 当作复矩阵进行证明 (即本题分析中的讨论, 其中用命题 7.38替代命题 7.37的引用), 具体细节请读者自行完成.

证法三: 设 φ 在 $\mathbb{K}$ 上的初等因子为 $\lambda^{r_1}, \cdots, \lambda^{r_k}, P_1(\lambda)^{e_1}, \cdots, P_t(\lambda)^{e_t}$, 其中 $P_1(\lambda), \cdots, P_t(\lambda)$ 是 $\mathbb{K}$ 上常数项非零的不可约多项式, 则由定理 7.23或命题??可知, 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

$$\mathrm{diag}\{F(\lambda^{r_1}), \cdots, F(\lambda^{r_k}), F(P_1(\lambda)^{e_1}), \cdots, F(P_t(\lambda)^{e_t})\} \text{ 或}$$
$$\mathrm{diag}\{J_{r_1}(0), \cdots, J_{r_k}(0), J_{e_1}(P_1(\lambda)), \cdots, J_{e_t}(P_t(\lambda))\}.$$

若特征值 0 是 φ 的极小多项式的单根, 则 $r_1 = \cdots = r_k = 1$, 容易验证 $\mathrm{Ker}\varphi = L(e_1, \cdots, e_k)$,$\mathrm{Im}\varphi = L(e_{k+1}, \cdots, e_n)$, 从而 $V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$. 反之, 若 $V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$, 则 $\mathrm{Ker}\varphi \cap \mathrm{Im}\varphi = 0$. 若存在某个 $r_i > 1$, 比如说 $r_1 > 1$, 则由命题 7.37的充分性完全类似的证明可知,$0 \neq e_1 \in \mathrm{Ker}\varphi \cap \mathrm{Im}\varphi$, 这就推出了矛盾. 因此,$r_1 = \cdots = r_k = 1$, 从而 0 是 φ 的极小多项式的单根. □

命题 7.59

设 A 是数域 $\mathbb{K}$ 上的 n 阶矩阵, 求证: A 相似于 $\mathrm{diag}\{B, C\}$, 其中 B 是 $\mathbb{K}$ 上的幂零矩阵,C 是 $\mathbb{K}$ 上的可逆矩阵.

注 分析: 本题是命题 7.49的推广, 即将复数域上的结论推广到数域 $\mathbb{K}$ 上. 不过,命题 7.49的证明利用了 Jordan 标准型理论, 显然在数域 $\mathbb{K}$ 上不再适用. 通常当我们考虑线性变换的问题时, 数域都是事先给定的, 从而在讨论的过程中不会涉及数域的问题. 因此我们可用第一种方法来处理本题, 即把代数问题转化成几何问题, 然后再用线性变换理论加以解决. 本题的几何版本为: 设 V 是数域 $\mathbb{K}$ 上的 n 维线性空间,φ 是 V 上的线性变换, 证明: $V = V_1 \oplus V_2$, 其中 V_1, V_2 都是 φ-不变子空间, 且 $\varphi|_{V_1}$ 是幂零线性变换,$\varphi|_{V_2}$ 是可逆线性变换. 我们可用两种几何方法来证明这一结论.

证明 证法一:设 φ 的特征多项式为 $f(\lambda) = \lambda^k g(\lambda)$, 其中 $0 \leqslant k \leqslant n$,$g(0) \neq 0$. 注意到 $(\lambda^k, g(\lambda)) = 1$, 故由命题 6.52可知,$V = V_1 \oplus V_2$, 其中 $V_1 = \mathrm{Ker}\varphi^k$,$V_2 = \mathrm{Ker}g(\varphi)$, 并且 $\varphi|_{V_1}$ 的特征多项式是 λ^k,$\varphi|_{V_2}$ 的特征多项式是 $g(\lambda)$. 因此,$\varphi|_{V_1}$ 是幂零线性变换, 且由 $\varphi|_{V_2}$ 的行列式值为 $(-1)^{n-k}g(0) \neq 0$ 可知,$\varphi|_{V_2}$ 是可逆线性变换.

证法二:由命题 4.21可知, 存在整数 $m \in [0, n]$, 使得

$$V = \mathrm{Ker}\varphi^m \oplus \mathrm{Im}\varphi^m, \quad \mathrm{Ker}\varphi^m = \mathrm{Ker}\varphi^{m+1} = \cdots, \quad \mathrm{Im}\varphi^m = \mathrm{Im}\varphi^{m+1} = \cdots.$$

令 $V_1 = \mathrm{Ker}\varphi^m$,$V_2 = \mathrm{Im}\varphi^m$, 则 $V = V_1 \oplus V_2$. 因为 $V_1 = \mathrm{Ker}\varphi^m$, 所以 $\varphi|_{V_1}$ 适合多项式 λ^m, 从而它是幂零线性变换. 因为 $\varphi|_{V_2}$ 的像空间是 $\varphi(\mathrm{Im}\varphi^m) = \mathrm{Im}\varphi^{m+1} = \mathrm{Im}\varphi^m$, 所以 $\varphi|_{V_2}$ 是满映射, 从而它是可逆线性变换.

证法三: 设 A 在 $\mathbb{K}$ 上的初等因子为 $\lambda^{r_1}, \cdots, \lambda^{r_k}, P_1(\lambda)^{e_1}, \cdots, P_t(\lambda)^{e_t}$, 其中 $P_1(\lambda), \cdots, P_t(\lambda)$ 是 $\mathbb{K}$ 上常数项非零的不可约多项式, 则由定理 7.23或命题**??**可知,A 在 $\mathbb{K}$ 上相似于分块对角矩阵

$$\mathrm{diag}\{F(\lambda^{r_1}), \cdots, F(\lambda^{r_k}), F(P_1(\lambda)^{e_1}), \cdots, F(P_t(\lambda)^{e_t})\} \text{ 或}$$
$$\mathrm{diag}\{J_{r_1}(0), \cdots, J_{r_k}(0), J_{e_1}(P_1(\lambda)), \cdots, J_{e_t}(P_t(\lambda))\}.$$

令 $B = \mathrm{diag}\{F(\lambda^{r_1}), \cdots, F(\lambda^{r_k})\}$ 或 $\mathrm{diag}\{J_{r_1}(0), \cdots, J_{r_k}(0)\}$,$C = \mathrm{diag}\{F(P_1(\lambda)^{e_1}), \cdots, F(P_t(\lambda)^{e_t})\}$ 或 $\mathrm{diag}\{J_{e_1}(P_1(\lambda)), \cdots, J_{e_t}(P_t(\lambda))\}$, 则由每个 $F(\lambda^{r_i})$ 或 $J_{r_i}(0)$ 都幂零可知 B 是幂零矩阵, 由每个 $F(P_j(\lambda)^{e_j})$ 或 $J_{e_j}(P_j(\lambda))$ 的行列式的绝对值为 $P_j(0)^{e_j} \neq 0$ 可知 C 是可逆矩阵, 因此结论成立. □

上面只是比较简单的两道例题, 如果希望能更一般地处理数域 $\mathbb{K}$ 上的相似问题, 那么我们可以运用数域 $\mathbb{K}$ 上基于初等因子的相似标准型理论. 事实上,定理 7.23已经给出**数域 $\mathbb{K}$ 上基于初等因子的有理标准型**, 接下去我们将给出数域 $\mathbb{K}$ 上基于初等因子的 Jordan 标准型. 这一理论跟之前阐述的数域 $\mathbb{K}$ 上基于不变因子的有理标准型理论和复数域上的 Jordan 标准型理论之间有着密切的联系, 无论是从引入的方法, 还是从最终的结论来看, 这一理论都是前面两种理论的自然延续和推广, 因此不妨称之为**广义 Jordan 标准型理论**.

定义 7.16

用 C_m 表示第 $(m, 1)$ 元素为 1, 其他元素全为零的 m 阶矩阵:

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

命题 7.60

设 $P(\lambda) = \lambda^m + a_1\lambda^{m-1} + \cdots + a_{m-1}\lambda + a_m$ 是 $\mathbb{K}$ 上的首一不可约多项式,e 是正整数, 证明下列矩阵的不变因子组均为 $1, \cdots, 1, P(\lambda)^e$:

$$(1)\ J_e(P(\lambda)) = \begin{pmatrix} F(P(\lambda)) & I_m & O & \cdots & O & O \\ O & F(P(\lambda)) & I_m & \cdots & O & O \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ O & O & O & \cdots & F(P(\lambda)) & I_m \\ O & O & O & \cdots & O & F(P(\lambda)) \end{pmatrix}$$

$$(2)\ \widetilde{J}_e(P(\lambda)) = \begin{pmatrix} F(P(\lambda)) & C_m & O & \cdots & O & O \\ O & F(P(\lambda)) & C_m & \cdots & O & O \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ O & O & O & \cdots & F(P(\lambda)) & C_m \\ O & O & O & \cdots & O & F(P(\lambda)) \end{pmatrix}.$$

证明 (1) 由有理标准型理论可知, $F(P(\lambda))$ 的特征多项式和极小多项式都是 $P(\lambda)$, 故 $J_e(P(\lambda))$ 的特征多项式为 $P(\lambda)^e$, 从而 $J_e(P(\lambda))$ 的极小多项式为 $P(\lambda)^l$, 其中 $1 \leqslant l \leqslant e$. 下面验证 $J_e(P(\lambda))$ 不适合 $P(\lambda)^{e-1}$, 从而 $J_e(P(\lambda))$ 的极小多项式必为 $P(\lambda)^e$. 以下简记 $g(\lambda) = P(\lambda)^{e-1}$,$F = F(P(\lambda))$, 则通过分块矩阵的计算可得

$$g(J_e(P(\lambda))) = \begin{pmatrix} g(F) & \frac{1}{1!}g'(F) & \frac{1}{2!}g^{(2)}(F) & \cdots & \frac{1}{(e-1)!}g^{(e-1)}(F) \\ & g(F) & \frac{1}{1!}g'(F) & \cdots & \frac{1}{(e-2)!}g^{(e-2)}(F) \\ & & g(F) & \cdots & \frac{1}{(e-3)!}g^{(e-3)}(F) \\ & & & \ddots & \vdots \\ & & & & g(F) \end{pmatrix}.$$

由 Cayley - Hamilton 定理可得 $P(F) = O$, 从而 $g^{(i)}(F) = O(0 \leqslant i \leqslant e-2)$, 但 $g^{(e-1)}(F) = (e-1)!P'(F)^{e-1}$. 由于 $P(\lambda)$ 是不可约多项式, 故 $(P(\lambda), P'(\lambda)) = 1$, 进一步有 $(P(\lambda), P'(\lambda)^{e-1}) = 1$, 从而由命题 6.46可知,$P'(F)^{e-1}$ 是可逆矩阵, 于是 $\frac{1}{(e-1)!}g^{(e-1)}(F) = P'(F)^{e-1} \neq O$, 即有 $g(J_e(P(\lambda))) \neq O$. 因此 $J_e(P(\lambda))$ 的极小多项式为 $P(\lambda)^e$, 其不变因子组为 $1, \cdots, 1, P(\lambda)^e$.

(2) 我们来计算 $\widetilde{J}_e(P(\lambda))$ 的 $me-1$ 阶行列式因子, 注意到特征矩阵 $\lambda I - \widetilde{J}_e(P(\lambda))$ 的前 $me-1$ 行、后 $me-1$ 列构成的 $me-1$ 阶子式是一个主对角元全为 -1 的下三角行列式, 其值为 $(-1)^{me-1}$, 故 $\widetilde{J}_e(P(\lambda))$ 的 $me-1$ 阶行列式因子为 1. 又 $\widetilde{J}_e(P(\lambda))$ 的 me 阶行列式因子为 $P(\lambda)^e$, 故其行列式因子组为 $1, \cdots, 1, P(\lambda)^e$, 从而不变因子组也为 $1, \cdots, 1, P(\lambda)^e$. □

命题 7.61

设 A 是 $\mathbb{K}$ 上的 n 阶矩阵, 它在 $\mathbb{K}$ 上的初等因子组为 $P_1(\lambda)^{e_1}, P_2(\lambda)^{e_2}, \cdots, P_t(\lambda)^{e_t}$, 其中 $P_i(\lambda)$ 是 $\mathbb{K}$ 上的首一不可约多项式,$e_i \geqslant 1, 1 \leqslant i \leqslant t$, 证明 A 在 $\mathbb{K}$ 上相似于下列分块对角矩阵:

(1) $J = \mathrm{diag}\{J_{e_1}(P_1(\lambda)), J_{e_2}(P_2(\lambda)), \cdots, J_{e_t}(P_t(\lambda))\}$;

(2) $\widetilde{J} = \mathrm{diag}\{\widetilde{J}_{e_1}(P_1(\lambda)), \widetilde{J}_{e_2}(P_2(\lambda)), \cdots, \widetilde{J}_{e_t}(P_t(\lambda))\}$.

注 这个命题??中的 J 和 $\widetilde{J}$ 均称为数域 $\mathbb{K}$ 上基于初等因子的广义 Jordan 标准型. 当 $\mathbb{K} = \mathbb{C}$ 时, 注意到不可约多项式都是一次的, 故可设 $P(\lambda) = \lambda - \lambda_0$, 则命题??中的广义 Jordan 块 $J_e(P(\lambda))$ 和 $\widetilde{J}_e(P(\lambda))$ 都变成了复数域上的 Jordan 块 $J_e(\lambda_0)$, 广义 Jordan 标准型 J 和 $\widetilde{J}$ 都变成了复数域上的 Jordan 标准型. J 和 $\widetilde{J}$ 之间的区别只是形式上的, 即对每个广义 Jordan 块而言, 其上次对角线上的矩阵一个是单位矩阵 I_m, 一个是矩阵 C_m. 从本质上看, 这两种广义 Jordan 标准型其实是一致的, 只不过在一些具体问题的讨论中, 各有各的用途而已.

证明 将 $\lambda I - J$ 和 $\lambda I - \widetilde{J}$ 按照每个分块依次进行 λ-矩阵的初等变换, 由λ-矩阵和初等因子的基本性质可知, 上述两个矩阵都相抵于

$$\mathrm{diag}\{1, \cdots, 1, P_1(\lambda)^{e_1}; 1, \cdots, 1, P_2(\lambda)^{e_2}; \cdots; 1, \cdots, 1, P_t(\lambda)^{e_t}\}.$$

由λ-矩阵和初等因子的基本性质 (2)可知,J 和 $\widetilde{J}$ 的初等因子组都是 $P_1(\lambda)^{e_1}, P_2(\lambda)^{e_2}, \cdots, P_t(\lambda)^{e_t}$, 即它们与 A 在 $\mathbb{K}$ 上有相同的初等因子组, 因此它们与 A 在 $\mathbb{K}$ 上相似. □

命题 7.62 (实数域上的广义 Jordan 标准型)

设 A 是实数域上的 n 阶矩阵, 证明 A 在实数域上相似于下列分块对角矩阵:

(1) $J = \mathrm{diag}\{J_{r_1}(\lambda_1), \cdots, J_{r_k}(\lambda_k), J_{s_1}(a_1, b_1), \cdots, J_{s_l}(a_l, b_l)\}$;

(2) $\widetilde{J} = \mathrm{diag}\{J_{r_1}(\lambda_1), \cdots, J_{r_k}(\lambda_k), \widetilde{J}_{s_1}(a_1, b_1), \cdots, \widetilde{J}_{s_l}(a_l, b_l)\}$,

其中 $\lambda_1, \cdots, \lambda_k, a_1, b_1, \cdots, a_l, b_l$ 都是实数,$b_1, \cdots, b_l$ 都非零, 并且 $a_j \pm \mathrm{i} b_j$ 都是 A 的复特征值,$J_{r_i}(\lambda_i)$ 表示以 λ_i 为特征值的通常意义下的 Jordan 块,$R_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}$,$C_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, 且

$$J_{s_j}(a_j, b_j) = \begin{pmatrix} R_j & I_2 & & & \\ & R_j & I_2 & & \\ & & \ddots & \ddots & \\ & & & R_j & I_2 \\ & & & & R_j \end{pmatrix}, \quad \widetilde{J}_{s_j}(a_j, b_j) = \begin{pmatrix} R_j & C_2 & & & \\ & R_j & C_2 & & \\ & & \ddots & \ddots & \\ & & & R_j & C_2 \\ & & & & R_j \end{pmatrix}$$

证明 注意到实数域上的不可约多项式是一次多项式或者是判别式小于零的二次多项式, 故可设 A 的初等因子组为 $(\lambda - \lambda_1)^{r_1}, \cdots, (\lambda - \lambda_k)^{r_k}, ((\lambda - a_1)^2 + b_1^2)^{s_1}, \cdots, ((\lambda - a_l)^2 + b_l^2)^{s_l}$, 其中 $\lambda_1, \cdots, \lambda_k, a_1, b_1, \cdots, a_l, b_l$ 都是实数, 且 $b_1, \cdots, b_l$ 都非零.

(1) 由命题**??**(1) 可知,A 实相似于 $\mathrm{diag}\{J_{r_1}(\lambda_1), \cdots, J_{r_k}(\lambda_k), J_{s_1}((\lambda - a_1)^2 + b_1^2), \cdots, J_{s_l}((\lambda - a_l)^2 + b_l^2)\}$, 注意到 $F((\lambda - a_j)^2 + b_j^2) = \begin{pmatrix} 0 & 1 \\ -(a_j^2 + b_j^2) & 2a_j \end{pmatrix}$ 与 $R_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}$ 有相同的特征值 $a_j \pm \mathrm{i} b_j$, 故它们在复数域上, 从而也在实数域上相似. 因为 $J_{s_j}((\lambda - a_j)^2 + b_j^2)$ 的上次对角线都是 I_2, 所以不难把这种相似关系扩张到整个广义 Jordan 块上, 从而 $J_{s_j}((\lambda - a_j)^2 + b_j^2)$ 实相似于 $J_{s_j}(a_j, b_j)$, 于是 A 实相似于 J.

(2) 因为 $\widetilde{J}_{s_j}((\lambda - a_j)^2 + b_j^2)$ 的上次对角线都是 C_2, 所以用命题**??**(2) 很难推出第二个结论, 这里我们采用直接计算 $\widetilde{J}_{s_j}(a_j, b_j)$ 的不变因子组的方法来证明. 注意到 $\lambda I - \widetilde{J}_{s_j}(a_j, b_j)$ 右上方的 $2s_j - 1$ 阶子式等于 $(-1)^{2s_j - 1} b_j^{s_j} \neq 0$, 故 $\widetilde{J}_{s_j}(a_j, b_j)$ 的 $2s_j - 1$ 阶行列式因子为 1, 于是其行列式因子组和不变因子组均为 $1, \cdots, 1, ((\lambda - a_j)^2 + b_j^2)^{s_j}$. 由 λ-矩阵的初等变换以及λ-矩阵和初等因子的基本性质 (2)可知,A 和 $\widetilde{J}$ 在实数域上有相同的初等因子组, 从而它们在实数域上相似. □

定理 7.29 (数域 $\mathbb{K}$ 上的 Jordan-Chevalley 分解)

设 A 是数域 $\mathbb{K}$ 上的 n 阶矩阵, 证明存在 $\mathbb{K}$ 上的 n 阶矩阵 B, C, 使得 $A = B + C$, 且满足:

(1) B 在复数域上可对角化;

(2) C 是幂零矩阵;

(3) $BC = CB$, 并且满足上述条件的分解一定是唯一的.

注 类似于复数域上的 Jordan–Chevalley 分解, 我们还可以证明对于上述分解 $A = B + C$, 存在 $\mathbb{K}$ 上的多项式 $f(x)$, 使得 $B = f(A)$. 不过, 由于此证明涉及抽象代数中域的扩张等相关知识点, 故在这里就不作详细的展开了.

证明 设 A 在 $\mathbb{K}$ 上的初等因子组为 $P_1(\lambda)^{e_1}, P_2(\lambda)^{e_2}, \cdots, P_t(\lambda)^{e_t}$, 其中 $P_i(\lambda)$ 是 $\mathbb{K}$ 上的首一不可约多项式,$e_i \geqslant 1, 1 \leqslant i \leqslant t$. 由命题**??**可知, 存在 $\mathbb{K}$ 上的可逆矩阵 P, 使得

$$P^{-1}AP = J = \mathrm{diag}\{J_{e_1}(P_1(\lambda)), J_{e_2}(P_2(\lambda)), \cdots, J_{e_t}(P_t(\lambda))\}.$$

我们先对广义 Jordan 块 $J_{e_i}(P_i(\lambda))$ 来证明结论, 为方便起见, 记 $F_i = F(P_i(\lambda))$. 由于 $P_i(\lambda)$ 在 $\mathbb{K}$ 上不可约, 故

$(P_i(\lambda), P_i'(\lambda)) = 1$, 从而 $P_i(\lambda)$ 在复数域上无重根, 于是 F_i 在复数域上可对角化. 令

$$M_i = \begin{pmatrix} F_i & O & O & \cdots & O & O \\ O & F_i & O & \cdots & O & O \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ O & O & O & \cdots & F_i & O \\ O & O & O & \cdots & O & F_i \end{pmatrix}, \quad N_i = \begin{pmatrix} O & I & O & \cdots & O & O \\ O & O & I & \cdots & O & O \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ O & O & O & \cdots & O & I \\ O & O & O & \cdots & O & O \end{pmatrix},$$

则容易验证 $J_{e_i}(P_i(\lambda)) = M_i + N_i$,$M_i$ 复可对角化,N_i 幂零,$M_iN_i = N_iM_i$. 再令 $M = \mathrm{diag}\{M_1, \cdots, M_t\}$,$N = \mathrm{diag}\{N_1,\cdots,N_t\}$, 则 $J = M + N$,M 复可对角化,N 幂零,$MN = NM$. 最后令 $B = PMP^{-1}$,$C = PNP^{-1}$, 则 B, C 是 $\mathbb{K}$ 上的矩阵,$A = B + C$,B 复可对角化,C 幂零,$BC = CB$. 我们也可将 A, B, C 看成是复数域上的矩阵, 由复数域上的 Jordan - Chevalley 分解定理的唯一性可知, 满足上述条件的分解一定是唯一的. $\square$

第八章 二次型

8.1 二次型的化简和矩阵的合同

定义 8.1 (二次型)

设 f 是数域 $\mathbb{K}$ 上的 n 元二次齐次多项式:

$$f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n \tag{8.1}$$

$$+a_{22}x_2^2+\cdots+2a_{2n}x_2x_n+\cdots+a_{nn}x_n^2, \tag{8.2}$$

称 f 为数域 $\mathbb{K}$ 上的 n 元二次型, 简称**二次型**.

♣

定义 8.2 (二次型与矩阵的相伴)

用矩阵的乘法我们可以把(8.1) 式写成矩阵相乘的形式:

$$f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}, \tag{8.3}$$

其中

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}, \boldsymbol{x}=\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}.$$

在矩阵 $\boldsymbol{A}$ 中,$a_{ij}=a_{ji}$ 对一切 i,j 成立, 也就是说矩阵 $\boldsymbol{A}$ 是一个对称阵. 由此可知, 给定数域 $\mathbb{K}$ 上的一个 n 元二次型, 我们就得到了 $\mathbb{K}$ 上的一个 n 阶对称阵 $\boldsymbol{A}$, 称为该二次型的**相伴矩阵或系数矩阵**.

反过来, 若给定 $\mathbb{K}$ 上的一个 n 阶对称阵 $\boldsymbol{A}$, 则由 (8.3) 式, 我们可以得到 $\mathbb{K}$ 上的一个二次型, 称为对称阵 $\boldsymbol{A}$ 的**相伴二次型**.

♣

定理 8.1

证明: 二次型与其相伴矩阵一一对应. 此即

(1) (一个对称矩阵对应唯一一个二次型) 设 $\boldsymbol{A}=\boldsymbol{B}$ 都是对称矩阵, 则 $f=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}$.

(2) (一个二次型对应唯一一个系数矩阵) 设 $f=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}$, 其中 $\boldsymbol{A},\boldsymbol{B}$ 都是对称矩阵, 则有 $\boldsymbol{A}=\boldsymbol{B}$.

♡

注 事实上, 如果我们不限制矩阵是对称阵, 则系数矩阵将不唯一, 这样会给用矩阵方法研究二次型带来困难.

证明

(1) 由二次型的定义显然得证.

(2) 由 $f=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}$ 可知 $\boldsymbol{x}'\boldsymbol{A}-\boldsymbol{B}\boldsymbol{x}=\boldsymbol{O}$. 于是只需证 $\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{O}$ 即可. 又 $\boldsymbol{A}-\boldsymbol{B}$ 仍是对称阵. 这等价于证明下面的结论: 设 $\boldsymbol{A}=(a_{ij})$ 是 n 阶对称阵, 若 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}=0$ 对一切 $\boldsymbol{\alpha}$ 成立, 则 $\boldsymbol{A}=\boldsymbol{O}$. 令 $\boldsymbol{\alpha}=\boldsymbol{e}_i$ 是 n 维标准单位列向量, 则 $a_{ii}=\boldsymbol{e}_i'\boldsymbol{A}\boldsymbol{e}_i=0$. 再令 $\boldsymbol{\alpha}=\boldsymbol{e}_i+\boldsymbol{e}_j(i\neq j)$, 则

$$0=(\boldsymbol{e}_i+\boldsymbol{e}_j)'\boldsymbol{A}(\boldsymbol{e}_i+\boldsymbol{e}_j)=\boldsymbol{e}_i'\boldsymbol{A}\boldsymbol{e}_i+\boldsymbol{e}_j'\boldsymbol{A}\boldsymbol{e}_j+\boldsymbol{e}_i'\boldsymbol{A}\boldsymbol{e}_j+\boldsymbol{e}_j'\boldsymbol{A}\boldsymbol{e}_i=a_{ij}+a_{ji},$$

因为 $a_{ij}=a_{ji}$, 故 $a_{ij}=0(i\neq j)$, 于是 $\boldsymbol{A}=\boldsymbol{O}$. 这表明用对称阵来表示二次型时, 系数矩阵是唯一的.

□

定义 8.3 (矩阵的合同关系)

设 $\boldsymbol{A},\boldsymbol{B}$ 是数域 $\mathbb{K}$ 上的 n 阶矩阵, 若存在 n 阶非异阵 $\boldsymbol{C}$, 使

$$\boldsymbol{B}=\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C},$$

则称 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 是**合同的**, 或称 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 具有**合同关系**.

定理 8.2

矩阵的合同关系是一个等价关系.

证明

1. 任一矩阵 $\boldsymbol{A}$ 与自己合同, 因为 $\boldsymbol{A}=\boldsymbol{I}'\boldsymbol{A}\boldsymbol{I}$;
2. 若 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 合同, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同. 这是因为若 $\boldsymbol{B}=\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$, 则 $\boldsymbol{A}=(\boldsymbol{C}')^{-1}\boldsymbol{B}\boldsymbol{C}^{-1}=(\boldsymbol{C}^{-1})'\boldsymbol{B}\boldsymbol{C}^{-1}$;
3. 若 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 合同,$\boldsymbol{D}$ 与 $\boldsymbol{B}$ 合同, 则 $\boldsymbol{D}$ 与 $\boldsymbol{A}$ 合同. 事实上, 若 $\boldsymbol{B}=\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$,$\boldsymbol{D}=\boldsymbol{H}'\boldsymbol{B}\boldsymbol{H}$, 则 $\boldsymbol{D}=\boldsymbol{H}'\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}\boldsymbol{H}=(\boldsymbol{C}\boldsymbol{H})'\boldsymbol{A}(\boldsymbol{C}\boldsymbol{H})$.

□

引理 8.1 (初等合同变换)

对称阵 $\boldsymbol{A}$ 的下列变换都是合同变换:

1. 对换 $\boldsymbol{A}$ 的第 i 行与第 j 行, 再对换第 i 列与第 j 列;
2. 将非零常数 k 乘以 $\boldsymbol{A}$ 的第 i 行, 再将 k 乘以第 i 列;
3. 将 $\boldsymbol{A}$ 的第 i 行乘以 k 加到第 j 行上, 再将第 i 列乘以 k 加到第 j 列上.

证明 上述变换相当于将一个初等矩阵左乘以 $\boldsymbol{A}$ 后再将这个初等矩阵的转置右乘之, 因此是合同变换. 此即

$$\text{对换}\boldsymbol{A}\text{的第}i\text{行与第}j\text{行, 再对换第}i\text{列与第}j\text{列} \iff A\to P_{ij}AP'_{ij};$$
$$\text{将非零常数}k\text{乘以}\boldsymbol{A}\text{的第}i\text{行, 再将}k\text{乘以第}i\text{列} \iff A\to P_i(k)AP'_i(k);$$
$$\text{将}\boldsymbol{A}\text{的第}i\text{行乘以}k\text{加到第}j\text{行上, 再将第}i\text{列乘以}k\text{加到第}j\text{列上} \iff A\to P_{ij}(k)AP'_{ij}(k).$$

□

引理 8.2

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

证明 若 $a_{11}=0$, 而 $a_{ii}\neq 0$, 则将 $\boldsymbol{A}$ 的第一行与第 i 行对换, 再将第一列与第 i 列对换, 得到的矩阵的第 $(1,1)$ 元素不为零. 根据上述引理, 这样得到的矩阵和原矩阵合同.

若所有的 $a_{ii}=0(i=1,2,\cdots,n)$, 设 $a_{ij}\neq 0(i\neq j)$, 将 $\boldsymbol{A}$ 的第 j 行加到第 i 行上, 再将第 j 列加到第 i 列上. 因为 $\boldsymbol{A}$ 是对称阵,$a_{ij}=a_{ji}\neq 0$, 于是第 (i,i) 元素是 $2a_{ij}\neq 0$, 再用前面的办法使第 $(1,1)$ 元素不等于零. 根据上述引理, 这样得到的矩阵和原矩阵仍合同, 这就证明了结论. □

定理 8.3 (对称阵必合同于对角阵)

设 $\boldsymbol{A}$ 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 $\boldsymbol{C}$, 使 $\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$ 为对角阵. 进而

$$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}=\operatorname{diag}\{d_1,d_2,\cdots,d_r,0,\cdots,0\}.$$

其中 $r=\mathrm{r}\left(\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}\right)=\mathrm{r}(\boldsymbol{A})$.,$d_i\neq 0, i=1,2,\cdots,r$. 即秩 r 是矩阵合同关系下的一个不变量.

证明 由引理 8.2, 不妨设 $\boldsymbol{A}=(a_{ij})$ 中 $a_{11}\neq 0$. 若 $a_{i1}\neq 0$, 则可将第一行乘以 $-a_{11}^{-1}a_{i1}$ 加到第 i 行上, 再将第一列乘以 $-a_{11}^{-1}a_{i1}$ 加到第 i 列上. 由于 $a_{i1}=a_{1i}$, 故得到的矩阵的第 $(1,i)$ 元素及第 $(i,1)$ 元素均等于零. 由初等合同变换可知, 新得到的矩阵与 $\boldsymbol{A}$ 是合同的. 依次这样做下去, 可把 $\boldsymbol{A}$ 的第一行与第一列除 a_{11} 外的元素都消去, 于是 $\boldsymbol{A}$

合同于下列矩阵:

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & b_{23} & \cdots & b_{2n} \\ 0 & b_{32} & b_{33} & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & b_{n2} & b_{n3} & \cdots & b_{nn} \end{pmatrix}.$$

上式右下角是一个 $n-1$ 阶对称阵, 记为 $\boldsymbol{A}_1$. 因此由归纳假设, 存在 $n-1$ 阶非异阵 $\boldsymbol{D}$, 使 $\boldsymbol{D}'\boldsymbol{A}_1\boldsymbol{D}$ 为对角阵, 于是

$$\begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{D}' \end{pmatrix} \begin{pmatrix} a_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_1 \end{pmatrix} \begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{D} \end{pmatrix} = \begin{pmatrix} a_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{D}'\boldsymbol{A}_1\boldsymbol{D} \end{pmatrix}$$

是一个对角阵. 显然

$$\begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{D}' \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{D} \end{pmatrix}',$$

因此 $\boldsymbol{A}$ 合同于对角阵. □

8.2 二次型的化简

8.2.1 配方法

引理 8.3

$$\begin{aligned}(x_1+x_2+\cdots+x_n)^2 = x_1^2+x_2^2+\cdots+x_n^2 \\ &+2x_1x_2+2x_1x_3+\cdots+2x_1x_n \\ &+2x_2x_3+\cdots+2x_2x_n \\ &+\cdots \\ &+2x_{n-1}x_n.\end{aligned}$$

通过下面这个例子介绍**配方法**.

例题 8.1 将下列二次型化成对角型:

$$f(x_1,x_2,x_3)=x_1^2+2x_1x_2-4x_1x_3-3x_2^2-6x_2x_3+x_3^2.$$

解 先将含有 x_1 的项放在一起凑成完全平方再减去必要的项:

$$\begin{aligned} f(x_1,x_2,x_3) &= (x_1^2+2x_1x_2-4x_1x_3)-3x_2^2-6x_2x_3+x_3^2 \\ &= ((x_1+x_2-2x_3)^2-x_2^2-4x_3^2+4x_2x_3)-3x_2^2-6x_2x_3+x_3^2 \\ &= (x_1+x_2-2x_3)^2-4x_2^2-2x_2x_3-3x_3^2. \end{aligned}$$

再对后面那些项配方:

$$\begin{aligned} -4x_2^2-2x_2x_3-3x_3^2 &= -\left((2x_2+\frac{1}{2}x_3)^2-\frac{1}{4}x_3^2\right)-3x_3^2 \\ &= -(2x_2+\frac{1}{2}x_3)^2-\frac{11}{4}x_3^2. \end{aligned}$$

于是

$$f(x_1,x_2,x_3)=(x_1+x_2-2x_3)^2-(2x_2+\frac{1}{2}x_3)^2-\frac{11}{4}x_3^2.$$

令

$$\begin{cases} y_1 = x_1 + x_2 - 2x_3, \\ y_2 = 2x_2 + \dfrac{1}{2}x_3, \\ y_3 = x_3, \end{cases}$$

则

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 2 & \dfrac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

因此 $f = y_1^2 - y_2^2 - \dfrac{11}{4}y_3^2$, 其变换矩阵为

$$C = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 2 & \dfrac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -\dfrac{1}{2} & \dfrac{9}{4} \\ 0 & \dfrac{1}{2} & -\dfrac{1}{4} \\ 0 & 0 & 1 \end{pmatrix}.$$

□

注 在用配方法化二次型为只含平方项的标准型的过程中, 必须保证变换矩阵 C 是非异阵. 如果我们按照上面例题的方法, 将含 x_1 的项放在一起配成一个完全平方, 接下来将含 x_2 的项放在一起再配方, 如此不断做下去. 最后得到的变换矩阵 C 是一个主对角元全不为零的上三角阵, 因此是一个非异阵. 有时我们用看似简单的方法得到的结果未必正确. 比如用观察法即可得到下列配方:

$$\begin{aligned} f &= 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 + 2x_1x_3 + 2x_2x_3 \\ &= (x_1 - x_2)^2 + (x_1 + x_3)^2 + (x_2 + x_3)^2. \end{aligned}$$

若令 $y_1 = x_1 - x_2$, $y_2 = x_1 + x_3$, $y_3 = x_2 + x_3$, 则 $f = y_1^2 + y_2^2 + y_3^2$. 由于矩阵

$$\begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

不是非异阵, 因此上述配方不是我们所需要的结论.

如果已知的二次型中没有平方项, 我们可以采用下面例子中的方法.

例题 8.2 将二次型

$$f(x_1, x_2, x_3, x_4) = 2x_1x_2 - x_1x_3 + x_1x_4 - x_2x_3 + x_2x_4 - 2x_3x_4$$

化成对角型.

解 这个二次型缺少了 x_i^2 项, 因此无法用例题 8.1的方法配方, 但我们可作如下变换:

$$\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = y_3, \\ x_4 = y_4. \end{cases}$$

代入原二次型得

$$f = 2y_1^2 - 2y_2^2 - 2y_1y_3 + 2y_1y_4 - 2y_3y_4.$$

这时 y_1^2 项不为零, 于是

$$f = (2y_1^2 - 2y_1y_3 + 2y_1y_4) - 2y_2^2 - 2y_3y_4$$

$$
\begin{aligned}
&= 2\left((y_1 - \frac{1}{2}y_3 + \frac{1}{2}y_4)^2 - \frac{1}{4}y_3^2 - \frac{1}{4}y_4^2 + \frac{1}{2}y_3y_4\right) - 2y_2^2 - 2y_3y_4 \\
&= 2(y_1 - \frac{1}{2}y_3 + \frac{1}{2}y_4)^2 - 2y_2^2 - \frac{1}{2}y_3^2 - y_3y_4 - \frac{1}{2}y_4^2 \\
&= 2(y_1 - \frac{1}{2}y_3 + \frac{1}{2}y_4)^2 - 2y_2^2 - \frac{1}{2}(y_3 + y_4)^2.
\end{aligned}
$$

令

$$
\begin{cases}
z_1 = y_1 - \frac{1}{2}y_3 + \frac{1}{2}y_4, \\
z_2 = y_2, \\
z_3 = y_3 + y_4, \\
z_4 = y_4,
\end{cases}
$$

于是

$$
f = 2z_1^2 - 2z_2^2 - \frac{1}{2}z_3^2,
$$

其中 z_4^2 的系数为零, 故未写出.

为求变换矩阵 $\boldsymbol{C}$, 可从上面 z_i 的表示式中解出 y_i:

$$
\begin{cases}
y_1 = z_1 + \frac{1}{2}z_3 - z_4, \\
y_2 = z_2, \\
y_3 = z_3 - z_4, \\
y_4 = z_4,
\end{cases}
$$

再将 x_i 求出:

$$
\begin{cases}
x_1 = z_1 + z_2 + \frac{1}{2}z_3 - z_4, \\
x_2 = z_1 - z_2 + \frac{1}{2}z_3 - z_4, \\
x_3 = z_3 - z_4, \\
x_4 = z_4,
\end{cases}
$$

于是

$$
\boldsymbol{C} = \begin{pmatrix} 1 & 1 & \frac{1}{2} & -1 \\ 1 & -1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
$$

□

8.2.2 初等变换法

用配方法化简二次型有时比较麻烦, 求非异阵 $\boldsymbol{C}$ 也比较麻烦. 我们常常用初等变换法来化简二次型, 初等变换法的依据是初等合同变换.

这种方法可总结如下: 作 $n \times 2n$ 矩阵 $(\boldsymbol{A} \vdots \boldsymbol{I}_n)$, 对这个矩阵实施初等行变换, 同时施以同样的初等列变换, 将它左半边化为对角阵, 则这个对角阵就是已化简的二次型的相伴矩阵, 右半边的转置便是变换矩阵 $\boldsymbol{C}$.

如碰到第 $(1,1)$ 元素是零的矩阵, 可先设法将第 $(1,1)$ 元素化成非零, 再进行上述过程.

下面我们通过例子来说明这种方法.

例题 8.3 将下列二次型化为对角型:

$$f(x_1,x_2,x_3)=x_1^2-3x_2^2-2x_1x_2+2x_1x_3-6x_2x_3.$$

解 记与 f 相伴的对称阵为 $\boldsymbol{A}$, 写出 $(\boldsymbol{A}\vdots\boldsymbol{I}_3)$ 并作初等变换:

$$(\boldsymbol{A}\mid\boldsymbol{I}_3)=\begin{pmatrix}1&-1&1&1&0&0\\-1&-3&-3&0&1&0\\1&-3&0&0&0&1\end{pmatrix}\xrightarrow{r_1+r_2}$$

$$\begin{pmatrix}1&-1&1&1&0&0\\0&-4&-2&1&1&0\\1&-3&0&0&0&1\end{pmatrix}\xrightarrow{j_1+j_2}\begin{pmatrix}1&0&1&1&0&0\\0&-4&-2&1&1&0\\1&-2&0&0&0&1\end{pmatrix}\xrightarrow{-r_1+r_3}$$

$$\begin{pmatrix}1&0&1&1&0&0\\0&-4&-2&1&1&0\\0&-2&-1&-1&0&1\end{pmatrix}\xrightarrow{-j_1+j_3}\begin{pmatrix}1&0&0&1&0&0\\0&-4&-2&1&1&0\\0&-2&-1&-1&0&1\end{pmatrix}\xrightarrow{-\frac{1}{2}r_2+r_3}$$

$$\begin{pmatrix}1&0&0&1&0&0\\0&-4&-2&1&1&0\\0&0&0&-\frac{3}{2}&-\frac{1}{2}&1\end{pmatrix}\xrightarrow{-\frac{1}{2}j_2+j_3}\begin{pmatrix}1&0&0&1&0&0\\0&-4&0&1&1&0\\0&0&0&-\frac{3}{2}&-\frac{1}{2}&1\end{pmatrix}.$$

于是 f 可化简为

$$y_1^2-4y_2^2,$$

$$\boldsymbol{C}=\begin{pmatrix}1&0&0\\1&1&0\\-\frac{3}{2}&-\frac{1}{2}&1\end{pmatrix}'=\begin{pmatrix}1&1&-\frac{3}{2}\\0&1&-\frac{1}{2}\\0&0&1\end{pmatrix}.$$

□

例题 8.4 将二次型 $f(x_1,x_2,x_3)=2x_1x_2+4x_1x_3-4x_2x_3$ 化成对角型.

解 写出与 f 相伴的对称阵 $\boldsymbol{A}$, 作 $(\boldsymbol{A}\vdots\boldsymbol{I}_3)$ 并将它的第二行加到第一行上, 再将第二列加到第一列上:

$$(\boldsymbol{A}\vdots\boldsymbol{I}_3)=\begin{pmatrix}0&1&2&1&0&0\\1&0&-2&0&1&0\\2&-2&0&0&0&1\end{pmatrix}\to\begin{pmatrix}2&1&0&1&1&0\\1&0&-2&0&1&0\\0&-2&0&0&0&1\end{pmatrix}.$$

同例题 8.3一样, 对上述矩阵进行初等变换得到

$$\begin{pmatrix}2&0&0&1&1&0\\0&-\frac{1}{2}&0&-\frac{1}{2}&\frac{1}{2}&0\\0&0&8&2&-2&1\end{pmatrix}.$$

因此 f 化简为

$$2y_1^2-\frac{1}{2}y_2^2+8y_3^2,$$

$$\boldsymbol{C}=\begin{pmatrix}1&-\frac{1}{2}&2\\1&\frac{1}{2}&-2\\0&0&1\end{pmatrix}.$$

□

8.3 惯性定理

8.3.1 实二次型

定理 8.4 (实二次型的规范标准型)

设实二次型

$$f(x_1,x_2,\cdots,x_n)=d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2,$$

这个实二次型对应的系数矩阵为 $\boldsymbol{A}$ 是 n 阶实对称阵, 则 $\boldsymbol{A}$ 一定合同于下列对角阵:

$$\mathrm{diag}\{1,\cdots,1;-1,\cdots,-1;0,\cdots,0\}, \tag{8.4}$$

其中有 p 个 1,q 个 -1,$n-r$ 个零. 进而,f 一定可作变量替换得到

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2. \tag{8.5}$$

我们将(8.5)式中的二次型称为 f 的**规范标准型**.

证明 由定理 8.3, 任意一个实对称阵 $\boldsymbol{A}$ 必合同于一个对角阵:

$$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}=\mathrm{diag}\{d_1,d_2,\cdots,d_r,0,\cdots,0\},$$

其中 $d_i\neq 0(i=1,\cdots,r)$. 注意到 $\boldsymbol{C}$ 是可逆阵, 故 $r=\mathrm{r}(\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C})=\mathrm{r}(\boldsymbol{A})$, 即秩 r 是矩阵合同关系下的一个不变量. 于是我们不妨设实对称阵已具有下列对角阵的形状:

$$\boldsymbol{A}=\mathrm{diag}\{d_1,d_2,\cdots,d_r,0,\cdots,0\}.$$

由初等合同变换不难知道, 任意调换 $\boldsymbol{A}$ 的主对角线上的元素得到的矩阵仍与 $\boldsymbol{A}$ 合同. 因此我们可把零放在一起, 把正项与负项放在一起, 即可设 $d_1>0,\cdots,d_p>0$; $d_{p+1}<0,\cdots,d_r<0$. $\boldsymbol{A}$ 所代表的二次型为

$$f(x_1,x_2,\cdots,x_n)=d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2. \tag{8.6}$$

再对上述二次型作变量替换, 令

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

则 (8.6) 式变为

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

这一事实等价于说 $\boldsymbol{A}$ 合同于下列对角阵:

$$\mathrm{diag}\{1,\cdots,1;-1,\cdots,-1;0,\cdots,0\}.$$

其中有 p 个 1,q 个 -1,$n-r$ 个零. □

定理 8.5 (惯性定理)

证明(8.5)式中的数 p 及 $q=r-p$ 是两个合同不变量. 这等价于证明下面的结论.

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:

$$c_1y_1^2+\cdots+c_py_p^2-c_{p+1}y_{p+1}^2-\cdots-c_ry_r^2,$$
$$d_1z_1^2+\cdots+d_kz_k^2-d_{k+1}z_{k+1}^2-\cdots-d_rz_r^2,$$

其中 $c_i>0$, $d_i>0$, 则必有 $p=k$.

证明 用反证法, 设 $p>k$. 由前面的说明不妨设 c_i 及 d_i 均为 1, 因此

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2=z_1^2+\cdots+z_k^2-z_{k+1}^2-\cdots-z_r^2. \tag{8.7}$$

又设

$$\boldsymbol{x}=\boldsymbol{B}\boldsymbol{y},\ \boldsymbol{x}=\boldsymbol{C}\boldsymbol{z},$$

其中

$$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\ \boldsymbol{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},\ \boldsymbol{z}=\begin{pmatrix}z_1\\z_2\\\vdots\\z_n\end{pmatrix},$$

于是 $\boldsymbol{z}=\boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{y}$. 令

$$\boldsymbol{C}^{-1}\boldsymbol{B}=\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\vdots&\vdots&&\vdots\\c_{n1}&c_{n2}&\cdots&c_{nn}\end{pmatrix},$$

则

$$\begin{cases}z_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n,\\z_2=c_{21}y_1+c_{22}y_2+\cdots+c_{2n}y_n,\\\cdots\cdots\cdots\\z_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n.\end{cases}$$

因为 $p>k$, 故齐次线性方程组

$$\begin{cases}c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n=0,\\\cdots\cdots\cdots\\c_{k1}y_1+c_{k2}y_2+\cdots+c_{kn}y_n=0,\\y_{p+1}=0,\\\cdots\cdots\cdots\\y_n=0\end{cases}$$

必有非零解 (n 个未知数,$n-(p-k)$ 个方程式). 令其中一个非零解为 $y_1=a_1,\cdots,y_p=a_p,y_{p+1}=0,\cdots,y_n=0$, 把这组解代入 (8.7) 式左边得到

$$a_1^2+\cdots+a_p^2>0.$$

但这时 $z_1=\cdots=z_k=0$, 故 (8.7) 式右边将小于等于零, 引出了矛盾. 同理可证 $p<k$ 也不可能. □

定义 8.4

设 $f(x_1,x_2,\cdots,x_n)$ 是一个实二次型, 若它能化为形如 (8.5)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**. ♣

注 显然, 若已知秩 r 与符号差 s, 则 $p=\dfrac{1}{2}(r+s)$, $q=\dfrac{1}{2}(r-s)$. 事实上, 在 p,q,r,s 中只需知道其中两个数, 其余两个数也就知道了. 由于实对称阵与实二次型之间的等价关系, 我们将实二次型的秩、惯性指数及符号差也称为相应的实对称阵的秩、惯性指数及符号差.

定理 8.6

秩与符号差 (或正负惯性指数) 是实对称阵在合同关系下的全系不变量. ♡

证明 由惯性定理知道, 秩 r 与符号差 s 是实对称阵合同关系的不变量. 反之, 若 n 阶实对称阵 $\boldsymbol{A},\boldsymbol{B}$ 的秩都为 r, 符

号差都是 s, 则它们都合同于

$$\mathrm{diag}\{1,\cdots,1;-1,\cdots,-1;0,\cdots,0\},$$

其中有 $p=\frac{1}{2}(r+s)$ 个 $1,q=\frac{1}{2}(r-s)$ 个 -1 及 $n-r$ 个零, 因此 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同. 对正负惯性指数的结论也同样成立.
□

8.3.2 复二次型

定理 8.7 (复二次型的规范标准型)

设复二次型

$$f(x_1,x_2,\cdots,x_n)=d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2,$$

这个复二次型对应的系数矩阵为 $\boldsymbol{A}$ 是 n 阶复对称阵, 则 $\boldsymbol{A}$ 一定合同于下列对角阵:

$$\mathrm{diag}\{1,\cdots,1;0,\cdots,0\},\tag{8.8}$$

其中有 r 个 1, 并且 $r=\mathrm{r}(A)$. 进而,f 一定可作变量替换得到

$$z_1^2+z_2^2+\cdots+z_r^2.\tag{8.9}$$

我们将(8.9)式中的二次型称为 f 的**规范标准型**.

♡

证明 因为复二次型

$$f(x_1,x_2,\cdots,x_n)=d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2$$

必可化为

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

其中 $z_i=\sqrt{d_i}x_i(i=1,2,\cdots,r)$, $z_j=x_j(j=r+1,\cdots,n)$. 所以结论得证. 故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

定义 8.5

设 $f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ 是 n 元实二次型,$\boldsymbol{A}$ 是相伴矩阵.

(1) 若对任意 n 维非零列向量 $\boldsymbol{\alpha}$ 均有 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}>0$, 则称 f 是**正定二次** (简称**正定型**), 矩阵 $\boldsymbol{A}$ 称为**正定矩阵** (简称**正定阵**);

(2) 若对任意 n 维非零列向量 $\boldsymbol{\alpha}$ 均有 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}<0$, 则称 f 是**负定二次型** (简称**负定型**), 矩阵 $\boldsymbol{A}$ 称为**负定矩阵** (简称**负定阵**);

(3) 若对任意 n 维非零列向量 $\boldsymbol{\alpha}$ 均有 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}\geq 0$, 则称 f 是**半正定二次型** (简称**半正定型**), 矩阵 $\boldsymbol{A}$ 称为**半正定矩阵** (简称**半正定阵**);

(4) 若对任意 n 维非零列向量 $\boldsymbol{\alpha}$ 均有 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}\leq 0$, 则称 f 是**半负定二次型** (简称**半负定型**), 矩阵 $\boldsymbol{A}$ 称为**半负定矩阵** (简称**半负定阵**);

(5) 若存在 $\boldsymbol{\alpha}$, 使 $\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}>0$; 又存在 $\boldsymbol{\beta}$, 使 $\boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta}<0$, 则称 f 是**不定型**.

♣

注 显然

$$f(x_1,x_2,\cdots,x_n)=x_1^2+x_2^2+\cdots+x_n^2$$

是正定型, 而

$$f(x_1,x_2,\cdots,x_n)=-x_1^2-x_2^2-\cdots-x_n^2$$

是负定型.

定理 8.8

设 $f(x_1,x_2,\cdots,x_n)$ 是 n 元实二次型, 则

(1) f 是正定型的充分必要条件是 f 的正惯性指数等于 n;

(2) f 是负定型的充分必要条件是 f 的负惯性指数等于 n;

(3) f 是半正定型的充分必要条件是 f 的正惯性指数等于 f 的秩 r;

(4) f 是半负定型的充分必要条件是 f 的负惯性指数等于 f 的秩 r.

♡

证明 只证明 (1), 其余结论的证明类似.

若 f 的正惯性指数等于 n, 则 f 可化为下列标准型:

$$f = y_1^2 + y_2^2 + \cdots + y_n^2,$$

显然 f 是正定型. 反之, 若 f 是正定型, 如果 f 的正惯性指数 $p < n$, 则 f 可化为如下标准型:

$$f = y_1^2 + \cdots + y_p^2 - c_{p+1}y_{p+1}^2 - \cdots - c_n y_n^2, \tag{8.10}$$

其中 $c_j \geq 0(j = p+1,\cdots,n)$. 这时令 $b_1 = \cdots = b_p = 0$, $b_{p+1} = \cdots = b_n = 1$, 则 $b_1,b_2,\cdots,b_n$ 不全为零. 假设这时 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$, 其中

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

$\boldsymbol{C}$ 是非异阵, 则从 $y_i = b_i(i = 1,\cdots,n)$ 可得 $x_i = a_i(i = 1,\cdots,n)$ 是一组不全为零的实数. 于是

$$f(a_1,a_2,\cdots,a_n) \leq 0,$$

这与 f 是正定型矛盾. □

定理 8.9

(1) n 阶实对称阵 $\boldsymbol{A}$ 是正定阵当且仅当它合同于单位阵 $\boldsymbol{I}_n$;

(2) $\boldsymbol{A}$ 是负定阵当且仅当它合同于 $-\boldsymbol{I}_n$;

(3) $\boldsymbol{A}$ 是半正定阵当且仅当 $\boldsymbol{A}$ 合同于下列对角阵:

$$\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix};$$

(4) $\boldsymbol{A}$ 是半负定阵当且仅当 $\boldsymbol{A}$ 合同于下列对角阵:

$$\begin{pmatrix} -\boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}.$$

♡

证明 由定理 8.8可立即得到证明. □

定义 8.6 (顺序主子式)

设 $\boldsymbol{A} = (a_{ij})$ 是 n 阶矩阵,$\boldsymbol{A}$ 的 n 个子式:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix} \quad (k = 1,2,\cdots,n)$$

称为 $\boldsymbol{A}$ 的**顺序主子式**.

♣

定理 8.10

n 阶实对称阵 $\boldsymbol{A}$ 是正定阵的充分必要条件是它的 n 个顺序主子式全大于零.

证明 先证必要性. 设 n 阶实对称阵 $\boldsymbol{A}=(a_{ij})$ 为正定阵, 则对应的实二次型

$$f(x_1,x_2,\cdots,x_n)=\sum_{j=1}^{n}\sum_{i=1}^{n}a_{ij}x_ix_j$$

为正定型. 令

$$f_k(x_1,x_2,\cdots,x_k)=\sum_{j=1}^{k}\sum_{i=1}^{k}a_{ij}x_ix_j,$$

则对任意一组不全为零的实数 $c_1,c_2,\cdots,c_k$, 有

$$f_k(c_1,c_2,\cdots,c_k)=f(c_1,c_2,\cdots,c_k,0,\cdots,0)>0,$$

因此 f_k 是一个正定二次型, 从而它的相伴矩阵 $\boldsymbol{A}_k$ (由 $\boldsymbol{A}$ 的前 k 行及前 k 列组成) 是一个正定阵. 由于 $\boldsymbol{A}_k$ 合同于 $\boldsymbol{I}_k$, 故存在 k 阶非异阵 $\boldsymbol{B}$, 使

$$\boldsymbol{B}'\boldsymbol{A}_k\boldsymbol{B}=\boldsymbol{I}_k,$$

于是

$$\det(\boldsymbol{B}'\boldsymbol{A}_k\boldsymbol{B})=\det(\boldsymbol{B})^2\det(\boldsymbol{A}_k)=1,$$

即有 $\det(\boldsymbol{A}_k)>0(k=1,2,\cdots,n)$.

再证充分性. 对 $\boldsymbol{A}$ 的阶数进行归纳. 当 $n=1$ 时,$\boldsymbol{A}=(a)$,$a>0$, 于是 $f=ax_1^2$ 是正定型, 从而 $\boldsymbol{A}$ 是正定阵. 设结论对 $n-1$ 成立, 现证明对 n 阶实对称阵 $\boldsymbol{A}$, 若它的 n 个顺序主子式全大于零, 则 $\boldsymbol{A}$ 必是正定阵. 记 $\boldsymbol{A}_{n-1}$ 是 $\boldsymbol{A}$ 的 $n-1$ 阶顺序主子式所在的矩阵, 则 $\boldsymbol{A}$ 可写为

$$\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{\alpha}\\ \boldsymbol{\alpha}' & a_{nn}\end{pmatrix}.$$

因为 $\boldsymbol{A}$ 的顺序主子式全大于零, 故 $\boldsymbol{A}_{n-1}$ 的顺序主子式也全大于零, 由归纳假设,$\boldsymbol{A}_{n-1}$ 是正定阵. 于是 $\boldsymbol{A}_{n-1}$ 合同于 $n-1$ 阶单位阵, 即存在 $n-1$ 阶非异阵 $\boldsymbol{B}$, 使

$$\boldsymbol{B}'\boldsymbol{A}_{n-1}\boldsymbol{B}=\boldsymbol{I}_{n-1}.$$

令 $\boldsymbol{C}$ 是下列分块矩阵:

$$\boldsymbol{C}=\begin{pmatrix}\boldsymbol{B} & \boldsymbol{O}\\ \boldsymbol{O} & 1\end{pmatrix},$$

则

$$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix}\boldsymbol{B}' & \boldsymbol{O}\\ \boldsymbol{O} & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{\alpha}\\ \boldsymbol{\alpha}' & a_{nn}\end{pmatrix}\begin{pmatrix}\boldsymbol{B} & \boldsymbol{O}\\ \boldsymbol{O} & 1\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_{n-1} & \boldsymbol{B}'\boldsymbol{\alpha}\\ \boldsymbol{\alpha}'\boldsymbol{B} & a_{nn}\end{pmatrix}.$$

这是一个实对称阵, 其形式为

$$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix}1 & \cdots & 0 & c_1\\ \vdots & & \vdots & \vdots\\ 0 & \cdots & 1 & c_{n-1}\\ c_1 & \cdots & c_{n-1} & a_{nn}\end{pmatrix},$$

用第三类初等行及列变换可将上述矩阵化为对角阵. 这相当于对 $\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$ 右乘一个非异阵 $\boldsymbol{Q}$ 后, 再左乘 $\boldsymbol{Q}'$ 得到一个对角阵, 亦即 $\boldsymbol{Q}'\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}\boldsymbol{Q}$ 等于

$$\mathrm{diag}\{1,\cdots,1,c\}.$$

由于 $|\boldsymbol{A}|>0$, 故 $c>0$, 这就证明了 $\boldsymbol{A}$ 是一个正定阵. □

命题 8.1

若 $\boldsymbol{A}$ 是正定阵, 证明:

(1) $\boldsymbol{A}$ 的任一 k 阶主子阵, 即由 $\boldsymbol{A}$ 的第 $i_1, i_2, \cdots, i_k$ 行及 $\boldsymbol{A}$ 的第 $i_1, i_2, \cdots, i_k$ 列交点上元素组成的矩阵, 必是正定阵;

(2) $\boldsymbol{A}$ 的所有主子式全大于零, 特别,$\boldsymbol{A}$ 的主对角元素全大于零;

(3) $\boldsymbol{A}$ 中绝对值最大的元素仅在主对角线上.

证明

(1) 设 $\boldsymbol{A}_k$ 是矩阵 $\boldsymbol{A}$ 的第 k 个顺序主子式所在的矩阵, 则 $\boldsymbol{A}_k$ 是实对称阵且其顺序主子式都大于零, 因此 $\boldsymbol{A}_k$ 是正定阵.

经过若干次行对换以及相同的列对换, 我们不难将 $\boldsymbol{A}$ 的第 $i_1, i_2, \cdots, i_k$ 行及 $i_1, i_2, \cdots, i_k$ 列分别换成第 $1, 2, \cdots, k$ 行和第 $1, 2, \cdots, k$ 列. 利用上面的结论即知 (1) 成立.

(2) 由 (1) 的结论,$\boldsymbol{A}$ 的所有主子式都是正定阵. 又 $\boldsymbol{A}$ 的每个主子式都是其自身的顺序主子式, 故由定理 8.10可知 $\boldsymbol{A}$ 的所有主子式都大于零. 因此 (2) 成立.

(3) 用反证法. 假设 $a_{ij}(i \neq j)$ 是 $\boldsymbol{A}$ 的绝对值最大的元素. 根据 (1), 我们只需证明由第 i, j 行与第 i, j 列交点上元素组成的矩阵不是正定阵即可. 考虑矩阵 (由 $\boldsymbol{A}$ 的对称性不妨设 $i < j$)

$$\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix},$$

注意到 $a_{ij} = a_{ji}$ 且 $|a_{ij}| \geq a_{ii}, |a_{ij}| \geq a_{jj}$, 上述矩阵的行列式值 $a_{ii}a_{jj} - a_{ij}^2 \leq 0$, 所以这个矩阵一定不是正定阵.

□

8.5 Hermite 型

第九章 内积空间

9.1 内积空间的基本概念

定义 9.1 (Euclid 空间)

设 V 是实数域上的线性空间, 若存在某种规则, 使对 V 中任意一组有序向量 $\{\alpha,\beta\}$, 都唯一地对应一个实数, 记为 (α,β), 且适合如下规则:

(1) $(\beta,\alpha)=(\alpha,\beta)$;

(2) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$;

(3) $(c\alpha,\beta)=c(\alpha,\beta)$, c 为任一实数;

(4) $(\alpha,\alpha)\geq 0$ 且等号成立当且仅当 $\alpha=0$,

则称在 V 上定义了一个内积. 实数 (α,β) 称为 α 与 β 的内积. 线性空间 V 称为**实内积空间**. 有限维实内积空间称为 **Euclid 空间**, 简称为**欧氏空间**.

♣

定义 9.2 (酉空间)

设 V 是复数域上的线性空间, 若存在某种规则, 使对 V 中任意一组有序向量 $\{\alpha,\beta\}$, 都唯一地对应一个复数, 记为 (α,β), 且适合如下规则:

(1) $(\beta,\alpha)=\overline{(\alpha,\beta)}$;

(2) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$;

(3) $(c\alpha,\beta)=c(\alpha,\beta)$, c 为任一复数;

(4) $(\alpha,\alpha)\geq 0$ 且等号成立当且仅当 $\alpha=0$,

则称在 V 上定义了一个内积. 复数 (α,β) 称为 α 与 β 的内积. 线性空间 V 称为**复内积空间**. 有限维复内积空间称为**酉空间**.

♣

注 实内积空间的定义与复内积空间的定义是相容的. 事实上, 对一个实数 a,$\overline{a}=a$, 故定义 9.1 中的 (1) 与定义 9.2中的 (1) 是一致的. 因此, 我们经常将这两种空间统称为内积空间, 在某些定理的叙述及证明中也不区分它们, 而统一作为复内积空间来处理. 但是, 需要注意的是对复内积空间,定义 9.2中的 (1), (3) 意味着:

$$(\alpha,c\beta)=\overline{c}(\alpha,\beta).$$

定义 9.3 (标准内积)

1. 设 $\mathbb{R}^n$ 是 n 维实列向量空间,$\alpha=(x_1,x_2,\cdots,x_n)'$,$\beta=(y_1,y_2,\cdots,y_n)'$, 定义
$$(\alpha,\beta)=x_1y_1+x_2y_2+\cdots+x_ny_n,$$
则在此定义下 $\mathbb{R}^n$ 成为一个欧氏空间, 上述内积称为 $\mathbb{R}^n$ 的标准内积.
2. 设 $\mathbb{C}^n$ 是 n 维复列向量空间,$\alpha=(x_1,x_2,\cdots,x_n)'$,$\beta=(y_1,y_2,\cdots,y_n)'$, 定义
$$(\alpha,\beta)=x_1\overline{y_1}+x_2\overline{y_2}+\cdots+x_n\overline{y_n},$$
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.

♣

注 对 n 维实或复行向量空间, 我们也可同样定义标准内积.

例题 9.1

1. 设 V 是由 $[a,b]$ 区间上连续函数全体构成的实线性空间, 设 $f(t),g(t)\in V$, 定义
$$(f,g)=\int_a^b f(t)g(t)\mathrm{d}t,$$
则不难验证这是一个内积, 于是 V 成为内积空间. 这是一个无限维实内积空间.
2. (1) 设 V 是 n 维实列向量空间,G 是 n 阶正定实对称阵, 对 $\alpha,\beta\in V$, 定义
$$(\alpha,\beta)=\alpha'G\beta,$$
则这是一个内积, 并且 V 在上式的定义下成为欧氏空间.
(2) 设 U 是 n 维复列向量空间, 若有正定 Hermite 矩阵 H, 对 $\alpha,\beta\in U$, 定义:
$$(\alpha,\beta)=\alpha'H\overline{\beta}.$$
则这个 U 上的一个内积, 并且 U 在上式的定义下成为酉空间.

证明

1. 由内积空间的定义不难验证.
2. (1)定义 9.1中的 (2), (3) 显然成立. 对 (1), 注意到 $\alpha'G\beta$ 是实数, 其转置仍是它自己, 而 G 是对称阵, 故
$$(\alpha,\beta)=\alpha'G\beta=(\alpha'G\beta)'=\beta'G'\alpha=\beta'G\alpha=(\beta,\alpha).$$
又从 G 是正定阵即可知道 (4) 成立.
(2) 根据内积和酉空间的定义不难验证.

□

注 当 $G=I_n$ 为单位阵时,V 上内积就是标准内积. 对实列向量空间, 标准内积可用矩阵乘法表示为
$$(\alpha,\beta)=\alpha'\beta.$$
对实行向量空间, 标准内积也可表示为
$$(\alpha,\beta)=\alpha\beta'.$$
当 $H=I_n$ 为单位阵时,U 上内积就是标准内积. 对复列向量空间, 标准内积可用矩阵乘法表示为
$$(\alpha,\beta)=\alpha'\overline{\beta}.$$
对复行向量空间, 标准内积也可表示为
$$(\alpha,\beta)=\alpha\overline{\beta}'.$$

定义 9.4 (范数)

设 V 是实或复的内积空间,α 是 V 中的向量, 定义 α 的**长度**（或**范数**）为
$$\|\alpha\|=(\alpha,\alpha)^{\frac{1}{2}},$$
即实数 (α,α) 的算术平方根. ♣

注 注意由定义 9.1和定义 9.2中的规则 (4) 可知,(α,α) 总是非负实数. 从长度的定义知,$\|\alpha\|=0$ 当且仅当 $\alpha=0$. 当 $V=\mathbb{R}^n$ 且内积为标准内积时, 若 $\alpha=(x_1,x_2,\cdots,x_n)$, 则
$$\|\alpha\|=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}.$$

定义 9.5 (两个向量的距离)

定义内积空间中两个向量的距离. 设 $\alpha,\beta\in V$, 定义 α 与 β 的**距离**为
$$d(\alpha,\beta)=\|\alpha-\beta\|.$$
显然 $d(\alpha,\beta)=d(\beta,\alpha)$. ♣

定理 9.1 (范数的基本性质)

设 V 是实或复的内积空间,$\alpha,\beta\in V$,c 是任一常数（实数或复数）, 则

(1) $\|c\alpha\|=|c|\|\alpha\|$;

(2)(Cauchy – Schwarz不等式) $|(\alpha,\beta)|\leq\|\alpha\|\cdot\|\beta\|$;

(3)(三角不等式) $\|\alpha+\beta\|\leq\|\alpha\|+\|\beta\|$.

证明

(1) $\|c\alpha\|^2=(c\alpha,c\alpha)=c\bar{c}(\alpha,\alpha)=|c|^2\|\alpha\|^2$, 故 $\|c\alpha\|=|c|\|\alpha\|$.

(2) 若 $\alpha=0$, 则 $(0,\beta)=(0+0,\beta)=2(0,\beta)$, 故 $(0,\beta)=0$, 因此 (2) 成立. 若 $\alpha\neq 0$, 令

$$v=\beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha,$$

则 $(v,\alpha)=0$, 且

$$\begin{aligned}0\leq\|v\|^2&=\left(\beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha,\beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha\right)\\&=(\beta,\beta)-\frac{(\beta,\alpha)}{\|\alpha\|^2}(\alpha,\beta)\\&=\|\beta\|^2-\frac{|(\alpha,\beta)|^2}{\|\alpha\|^2},\end{aligned}$$

由此即可得 (2).

(3) 我们有

$$\begin{aligned}\|\alpha+\beta\|^2&=(\alpha+\beta,\alpha+\beta)\\&=\|\alpha\|^2+(\alpha,\beta)+(\beta,\alpha)+\|\beta\|^2\\&=\|\alpha\|^2+\|\beta\|^2+(\alpha,\beta)+\overline{(\alpha,\beta)}.\end{aligned}$$

由 (2) 得 $|(\alpha,\beta)|\leq\|\alpha\|\|\beta\|$, $\overline{|(\alpha,\beta)|}\leq\|\alpha\|\|\beta\|$, 故 $\|\alpha+\beta\|^2\leq\|\alpha\|^2+\|\beta\|^2+2\|\alpha\|\|\beta\|=(\|\alpha\|+\|\beta\|)^2$. □

定义 9.6 (向量的夹角)

当 V 是实内积空间时, 定义非零向量 α,β 的夹角 θ 之余弦为

$$\cos\theta=\frac{(\alpha,\beta)}{\|\alpha\|\|\beta\|}.\tag{9.1}$$

当 V 是复内积空间时, 定义非零向量 α,β 的夹角 θ 之余弦为

$$\cos\theta=\frac{|(\alpha,\beta)|}{\|\alpha\|\|\beta\|}.$$

内积空间中两个向量 α,β 若适合 $(\alpha,\beta)=0$, 则称 α 与 β **垂直**或**正交**, 我们用记号 $\alpha\perp\beta$ 来表示. 显然, 我们有以下结论:

1. 零向量和任何向量都正交;
2. 若 α 与 β 正交, 则 β 也与 α 正交;
3. 两个非零向量 α,β 正交时夹角为 90°.

注 (9.1)式中要使 θ 有意义, 必须保证 $|\cos\theta|\leq 1$, 而这就是范数的基本性质中的 (2). 因此上述定义的 θ 都是良定义的.

推论 9.1

1. (勾股定理) 在范数的基本性质 (3)的证明中我们可看出: 若 α 与 β 正交, 则 $(\alpha,\beta)=(\beta,\alpha)=0$, 因此

$$\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2.$$

上式通常称为勾股定理, 它是平面几何中勾股定理的推广.

2. (Cauchy 不等式) 设 V 是 n 维实行向量空间, 内积取标准内积, 从范数的基本性质 (2)立即可得到下列 Cauchy 不等式:

$$(x_1y_1+x_2y_2+\cdots+x_ny_n)^2 \le (x_1^2+x_2^2+\cdots+x_n^2)(y_1^2+y_2^2+\cdots+y_n^2).$$

3. (Schwarz 不等式) 设 V 是由 $[a,b]$ 区间上连续函数全体构成的实线性空间, 内积如例题 9.1(1), 则从范数的基本性质 (2)可得下列 Schwarz 不等式:

$$\left(\int_a^b f(t)g(t)\mathrm{d}t\right)^2 \le \int_a^b f(t)^2\mathrm{d}t\int_a^b g(t)^2\mathrm{d}t.$$

9.2 内积的表示和正交基

定义 9.7 (Gram 矩阵和度量矩阵)

设 $\beta_1,\beta_2,\cdots,\beta_s$ 是内积空间的一个向量组, 矩阵

$$G=\begin{pmatrix}(\beta_1,\beta_1) & \cdots & (\beta_1,\beta_n)\\ \vdots & \ddots & \vdots\\ (\beta_n,\beta_1) & \cdots & (\beta_n,\beta_n)\end{pmatrix}$$

称为向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 的 **Gram 矩阵**. 若 $\beta_1,\beta_2,\cdots,\beta_s$ 是一组基, 则将 Gram 矩阵称为该基的**度量矩阵**.

定理 9.2

1. 若 V 是一个 n 维欧式空间,$\alpha_1,\alpha_2,\cdots,\alpha_n$ 是它的一组基, 对 V 中任意向量 α,β, 其中

$$\alpha = x_1\alpha_1+\cdots+x_n\alpha_n,$$
$$\beta = y_1\alpha_1+\cdots+y_n\alpha_n.$$

则

$$(\alpha,\beta)=X^TGY, \tag{9.2}$$

其中

$$G=\begin{pmatrix}(\alpha_1,\alpha_1) & \cdots & (\alpha_1,\alpha_n)\\ \vdots & \ddots & \vdots\\ (\alpha_n,\alpha_1) & \cdots & (\alpha_n,\alpha_n)\end{pmatrix}, X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}, Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}.$$

此时,G 就是基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的度量矩阵. 并且 G 是一个 n 阶正定实对称阵.

由此可知, 若给定了 n 维实线性空间 V 的一组基, 则 V 上的内积结构和 n 阶正定实对称阵之间存在着一个一一对应.

2. 若 V 是一个 n 维酉空间,$\alpha_1,\alpha_2,\cdots,\alpha_n$ 是它的一组基, 对 V 中任意向量 α,β, 其中

$$\alpha = x_1\alpha_1+\cdots+x_n\alpha_n,$$
$$\beta = y_1\alpha_1+\cdots+y_n\alpha_n.$$

则

$$(\alpha,\beta)=X^TH\overline{Y},$$

其中

$$H=\begin{pmatrix}(\alpha_1,\alpha_1)&\cdots&(\alpha_1,\alpha_n)\\ \vdots&\ddots&\vdots\\ (\alpha_n,\alpha_1)&\cdots&(\alpha_n,\alpha_n)\end{pmatrix},X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}.$$

此时,H 就是基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的度量矩阵. 并且 H 是一个 n 阶正定 Hermite 阵.

由此可知, 若给定了 n 维复线性空间 V 的一组基, 则 V 上的内积结构和 n 阶正定 Hermite 阵之间存在着一个一一对应.

证明

1. 利用内积的线性性容易得到 $(\alpha,\beta)=X^TGY$.
   再来看矩阵 G. 因为 $(\alpha_i,\alpha_j)=(\alpha_j,\alpha_i)$, 所以 G 是实对称阵. 又因为对任意的非零向量 α, 总有 $(\alpha,\alpha)>0$, 所以 $x'Gx>0$ 对一切 n 维非零实列向量 x 成立. 这表明 G 是一个正定阵.
   反之, 若给定 n 阶正定实对称阵 G, 利用 (9.2)式也可以定义 V 上的内积 (参考例题 2.(1)). 由此我们可以看出, 若给定了 n 维实线性空间 V 的一组基, 则 V 上的内积结构和 n 阶正定实对称阵之间存在着一个一一对应.
2. 由 1 类似可证.

□

定义 9.8

设 $\{e_1,e_2,\cdots,e_n\}$ 是 n 维内积空间 V 的一组基. 若 $e_i\perp e_j$ 对一切 $i\neq j$ 成立, 则称这组基是 V 的一组**正交基**. 又若 V 的一组正交基中每个基向量的长度都等于 1, 则称这组正交基为**标准正交基**.

显然在标准正交基下, 度量矩阵就是单位矩阵.

引理 9.1

内积空间 V 中的任意一组两两正交的非零向量必线性无关.

证明 设 $v_1,v_2,\cdots,v_m$ 是 V 中两两正交的非零向量, 若

$$k_1v_1+k_2v_2+\cdots+k_mv_m=\mathbf{0},$$

则对任一 $1\le i\le m$, 有

$$(k_1v_1+k_2v_2+\cdots+k_mv_m,v_i)=0.$$

由于 $v_i\perp v_j(i\neq j)$, 故由上式可得 $k_i(v_i,v_i)=0$, 又 $v_i\neq\mathbf{0}$, 从而 $k_i=0$. □

引理 9.2

设向量 α 和 $\beta_1,\beta_2,\cdots,\beta_k$ 都正交, 则 α 和 $L(\beta_1,\beta_2,\cdots,\beta_k)$ 中的每个向量都正交.

证明 任取 $\beta=b_1\beta_1+b_2\beta_2+\cdots+b_k\beta_k\in L(\beta_1,\beta_2,\cdots,\beta_k)$, 则

$$(\beta,\alpha)=(b_1\beta_1+b_2\beta_2+\cdots+b_k\beta_k,\alpha)=\sum_{i=1}^k b_i(\beta_i,\alpha)=0,$$

结论得证. □

推论 9.2

n 维内积空间中任意一个正交非零向量组的向量个数不超过 n.

证明 假设 n 维内积空间 V 中有 $n+1$ 个正交非零的向量, 则由引理 9.1可知, 这 $n+1$ 个正交非零的向量一定线性无关, 这与 $\dim V=n$ 矛盾! □

定理 9.3 (Gram-Schmidt 正交化)

设 V 是内积空间,$u_1, u_2, \cdots, u_m$ 是 V 中 m 个线性无关的向量, 则在 V 中存在 m 个两两正交的非零向量 $v_1, v_2, \cdots, v_m$, 使由 $v_1, v_2, \cdots, v_m$ 张成的子空间恰好为由 $u_1, u_2, \cdots, u_m$ 张成的子空间, 即 $v_1, v_2, \cdots, v_m$ 是该子空间的一组正交基.

证明 设 $v_1 = u_1$, 其余 v_i 可用数学归纳法定义如下: 假设 $v_1, \cdots, v_k (k < m)$ 已定义好, 这时 $v_1, \cdots, v_k$ 两两正交非零且 $L(v_1, \cdots, v_k) = L(u_1, \cdots, u_k)$. 令

$$v_{k+1} = u_{k+1} - \sum_{j=1}^{k} \frac{(u_{k+1}, v_j)}{\|v_j\|^2} v_j. \tag{9.3}$$

注意 $v_{k+1} \neq \mathbf{0}$, 否则 u_{k+1} 将是 $v_1, \cdots, v_k$ 的线性组合, 从而也是 $u_1, \cdots, u_k$ 的线性组合, 此与 $u_1, u_2, \cdots, u_m$ 线性无关矛盾. 又对任意的 $1 \leq i \leq k$, 有

$$\begin{aligned}(v_{k+1}, v_i) &= (u_{k+1}, v_i) - \sum_{j=1}^{k} \frac{(u_{k+1}, v_j)}{\|v_j\|^2}(v_j, v_i) \\ &= (u_{k+1}, v_i) - (u_{k+1}, v_i) = 0,\end{aligned}$$

因此 $v_1, \cdots, v_k, v_{k+1}$ 两两正交. 由(9.3)式可知

$$u_{k+1} \in L(v_1, \cdots, v_k, v_{k+1}) \text{及} v_{k+1} \in L(v_1, \cdots, v_k) + L(u_{k+1}) = L(u_1, \cdots, u_k) + L(u_{k+1}) = L(u_1, \cdots, u_k, u_{k+1}),$$

于是 $L(v_1, \cdots, v_k, v_{k+1}) = L(u_1, \cdots, u_k, u_{k+1})$, 这就证明了结论. □

注 上述定理证明中的正交化过程通常称为 Gram - Schmidt (格列姆–施密特) 方法.

推论 9.3

任一有限维内积空间均有标准正交基.

定义 9.9 (正交和)

设 V 是 n 维内积空间,$V_1, V_2, \cdots, V_k$ 是 V 的子空间. 如果对任意的 $\alpha \in V_i$ 和任意的 $\beta \in V_j$ 均有 $(\alpha, \beta) = 0$, 则称子空间 V_i 和 V_j 正交. 若 $V = V_1 + V_2 + \cdots + V_k$ 且 V_i 两两正交, 则称 V 是 $V_1, V_2, \cdots, V_k$ 的正交和, 记为

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

注 由于引理 9.3, 正交和通常也称为正交直和.

引理 9.3

正交和必为直和且任一 V_i 和其余子空间的和正交.

证明 对任意的 $v_i \in V_i$ 和 $\sum\limits_{j \neq i} v_j (v_j \in V_j)$, 有

$$(v_i, \sum_{j \neq i} v_j) = \sum_{j \neq i} (v_i, v_j) = 0,$$

因此后一个结论成立. 任取 $v \in V_i \cap (\sum\limits_{j \neq i} V_j)$, 则由上述论证可得 $(v, v) = 0$, 故 $v = \mathbf{0}$, 从而 $V_i \cap (\sum\limits_{j \neq i} V_j) = 0$, 即正交和必为直和. □

定义 9.10 (正交补空间)

设 U 是内积空间 V 的子空间, 令

$$U^{\perp} = \{v \in V | (v, U) = 0\},$$

这里 $(v, U) = 0$ 表示对一切 $u \in U$, 均有 $(v, u) = 0$. 容易验证 $U^{\perp}$ 是 V 的子空间, 称为 U 的**正交补空间**.

定理 9.4

设 V 是 n 维内积空间,U 是 V 的子空间, 则

(1) $V = U \oplus U^{\perp} = U \perp U^{\perp}$;

(2) U 的任一组标准正交基均可扩张为 V 的一组标准正交基.

证明 (1) 若 $x \in U \cap U^{\perp}$, 则 $(x,x) = 0$, 因此 $x = \mathbf{0}$, 即 $U \cap U^{\perp} = 0$. 另一方面, 由推论 9.3可知, 存在 U 的一组标准正交基 $\{e_1, e_2, \cdots, e_m\}$. 对任意的 $v \in V$, 令

$$u = (v, e_1)e_1 + (v, e_2)e_2 + \cdots + (v, e_m)e_m,$$

则 $u \in U$. 又令 $w = v - u$, 则对任一 $e_i (i = 1, 2, \cdots, m)$, 有

$$(w, e_i) = (v, e_i) - (u, e_i) = (v, e_i) - (v, e_i) = 0.$$

因此 $w \in U^{\perp}$, 又 $v = u + w$, 这就证明了 $V = U \oplus U^{\perp}$.

(2) 设 $\{e_1, e_2, \cdots, e_m\}$ 是 U 的任一组标准正交基,$\{e_{m+1}, \cdots, e_n\}$ 是 $U^{\perp}$ 的任一组标准正交基, 则显然 $\{e_1, e_2, \cdots, e_n\}$ 是 V 的一组标准正交基. □

定义 9.11 (正交投影)

设 $V = V_1 \perp V_2 \perp \cdots \perp V_k$, 定义 V 上的线性变换 $E_i (i = 1, 2, \cdots, k)$ 如下: 若 $v = v_1 + \cdots + v_i + \cdots + v_k (v_i \in V_i)$, 令 $E_i(v) = v_i$. 容易验证 E_i 是 V 上的线性变换, 且满足

$$E_i^2 = E_i,\ E_i E_j = 0\,(i \neq j),\ E_1 + E_2 + \cdots + E_k = I_V.$$

线性变换 E_i 称为 V 到 V_i 上的**正交投影**（简称投影）.

命题 9.1

设 U 是内积空间 V 的子空间,$V = U \perp U^{\perp}$. 设 E 是 V 到 U 上的正交投影, 则对任意的 $\alpha, \beta \in V$, 有

$$(E(\alpha), \beta) = (\alpha, E(\beta)).$$

证明 设 $\alpha = u_1 + w_1$,$\beta = u_2 + w_2$, 其中 $u_1, u_2 \in U$,$w_1, w_2 \in U^{\perp}$, 则 $E(\alpha) = u_1$,$E(\beta) = u_2$, 于是

$$(E(\alpha), \beta) = (u_1, u_2 + w_2) = (u_1, u_2) + (u_1, w_2) = (u_1, u_2),$$

$$(\alpha, E(\beta)) = (u_1 + w_1, u_2) = (u_1, u_2) + (w_1, u_2) = (u_1, u_2).$$

由此即得结论. □

命题 9.2 (Bessel (贝塞尔) 不等式)

设 $v_1, v_2, \cdots, v_m$ 是内积空间 V 中的正交非零向量组,y 是 V 中任一向量, 则

$$\sum_{k=1}^{m} \frac{|(y, v_k)|^2}{\|v_k\|^2} \leq \|y\|^2,$$

且等号成立的充分必要条件是 y 属于由 $\{v_1, v_2, \cdots, v_m\}$ 张成的子空间.

注 Bessel (贝塞尔) 不等式是“斜边大于直角边”这一几何命题在内积空间中的推广.

证明 令

$$x = \sum_{k=1}^{m} \frac{(y, v_k)}{\|v_k\|^2} v_k,$$

则 x 属于由 $\{v_1, v_2, \cdots, v_m\}$ 张成的子空间. 容易验证

$$(y - x, v_k) = 0,\ k = 1, 2, \cdots, m,$$

因此 $(y-x,x)=0$. 由勾股定理可得

$$\|y\|^2=\|y-x\|^2+\|x\|^2,$$

故

$$\|x\|^2\le\|y\|^2.$$

又由 $v_1,v_2,\cdots,v_m$ 两两正交不难算出

$$\|x\|^2=\sum_{k=1}^m\frac{|(y,v_k)|^2}{\|v_k\|^2}.$$

若 y 属于由 $\{v_1,v_2,\cdots,v_m\}$ 张成的子空间, 则 $y=x$, 故等号成立. 反之, 若等号成立, 则 $\|y-x\|^2=0$, 故 $y=x$, 即 y 属于由 $\{v_1,v_2,\cdots,v_m\}$ 张成的子空间. □

9.3 伴随

定义 9.12 (伴随)

设 φ 是内积空间 V 上的线性算子, 若存在 V 上的线性算子 φ^*, 使等式

$$(\varphi(\alpha),\beta)=(\alpha,\varphi^*(\beta))$$

对一切 $\alpha,\beta\in V$ 成立, 则称 φ^* 是 φ 的伴随算子, 简称为 φ 的**伴随**.

定理 9.5

设 V 是 n 维内积空间,φ 是 V 上的线性变换, 则存在 V 上唯一的线性变换 φ^*, 使对一切 $\alpha,\beta\in V$, 成立

$$(\varphi(\alpha),\beta)=(\alpha,\varphi^*(\beta)).$$

笔记 这个定理表明: 对有限维内积空间 V 上的任一线性算子, 它的伴随必存在且唯一.

证明 只需证明唯一性. 若 $\varphi^\sharp$ 是 V 上的线性变换且

$$(\varphi(\alpha),\beta)=(\alpha,\varphi^\sharp(\beta))$$

对一切 $\alpha,\beta\in V$ 成立, 则 $(\alpha,\varphi^\sharp(\beta))=(\alpha,\varphi^*(\beta))$ 对一切 $\alpha\in V$ 成立, 即 $(\alpha,\varphi^\sharp(\beta)-\varphi^*(\beta))=0$ 对一切 $\alpha\in V$ 成立, 特别, 对 $\alpha=\varphi^\sharp(\beta)-\varphi^*(\beta)$ 也成立. 由内积定义即知 $\varphi^\sharp(\beta)-\varphi^*(\beta)=0$, 即 $\varphi^\sharp(\beta)=\varphi^*(\beta)$. 而 β 是任意的, 故有 $\varphi^\sharp=\varphi^*$.
□

定理 9.6

设 V 是 n 维内积空间,$\{e_1,e_2,\cdots,e_n\}$ 是 V 的一组标准正交基. 若 V 上的线性算子 φ 在这组基下的表示矩阵为 A, 则

(1) 当 V 是酉空间时,φ^* 在同一组基下的表示矩阵为 $\overline{A}'$, 即 A 的共轭转置;

(2) 当 V 是欧氏空间时,φ^* 的表示矩阵为 A', 即 A 的转置.

证明 由伴随的唯一性知道本节一开始由 $\overline{A}'$ 定义的线性变换 ψ 就是 φ 的伴随, 而 ψ 的表示矩阵就是 $\overline{A}'$. □

定理 9.7 (伴随算子的性质)

设 V 是有限维内积空间, 若 φ 及 ψ 是 V 上的线性变换,c 为常数, 则

(1) $(\varphi+\psi)^*=\varphi^*+\psi^*$;

(2) $(c\varphi)^*=\overline{c}\varphi^*$;

(3) $(\varphi\psi)^*=\psi^*\varphi^*$;

(4) $(\varphi^*)^*=\varphi$.

证明 由矩阵和线性变换的一一对应关系及矩阵共轭转置的性质即得. □

命题 9.3

设 V 是 n 维内积空间,φ 是 V 上的线性算子.

(1) 若 U 是 φ 的不变子空间, 则 $U^{\perp}$ 是 φ^* 的不变子空间;

(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 $\alpha\in U,\beta\in U^{\perp}$, 因为

$$(\alpha,\varphi^*(\beta))=(\varphi(\alpha),\beta)=0,$$

所以 $U^{\perp}$ 是 φ^* 的不变子空间.

(2) 取 V 的一组标准正交基, 设 φ 在这组基下的表示矩阵为 A, 则无论 V 是酉空间还是欧氏空间,φ^* 的表示矩阵总可写为 $\overline{A}'$. 由假设

$$|\lambda I_n-A|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n),$$

则容易验证

$$|\lambda I_n-\overline{A}'|=(\lambda-\overline{\lambda}_1)(\lambda-\overline{\lambda}_2)\cdots(\lambda-\overline{\lambda}_n),$$

故结论成立.

□

9.4 内积空间的同构、正交变换和酉变换

定义 9.13 (保积同构)

设 V 与 U 是域 $\mathbb{K}$ 上的内积空间,$\mathbb{K}$ 是实数域或复数域,φ 是 $V\to U$ 的线性映射. 若对任意的 $x,y\in V$, 有

$$(\varphi(x),\varphi(y))=(x,y),$$

则称 φ 是 $V\to U$ 的保持内积的线性映射. 又若 φ 作为线性映射是同构, 则称 φ 是内积空间 V 到 U 上的**保积同构**.

注

1. 在不引起误解的情况下, 我们常把内积空间的保积同构就称为同构.
2. 保持内积的线性映射一定是单映射, 这是因为任取 $x\in\mathrm{Ker}\varphi,\|\varphi(x)\|=\|x\|$, 从 $\|\varphi(x)\|=0$ 得到 $\|x\|=0$, 故 $x=0$. 因此 $\mathrm{Ker}\varphi=\{0\}$.
3. 容易证明保持内积的同构关系是一个等价关系.

定理 9.8

若 φ 是内积空间 V 到内积空间 U 的保持范数的线性映射, 则 φ 保持内积.

注 显然线性映射保持内积一定保持范数, 再结合上述定理可知, 保持内积与保持范数的等价性, 因此保持内积的同构也称为**保范同构**或**保距同构**.

证明 向量的范数可以用内积表示, 反过来内积也可以用范数来表示. 设 x,y 是 V 中的任意两个向量, 则

$$\|x+y\|^2=(x+y,x+y)=(x,x)+(y,y)+(x,y)+(y,x),$$
$$\|x-y\|^2=(x-y,x-y)=(x,x)+(y,y)-(x,y)-(y,x),$$

故

$$\|x+y\|^2-\|x-y\|^2=2(x,y)+2\overline{(x,y)}.\tag{9.4}$$

另一方面,

$$\|x+\mathrm{i}y\|^2=(x+\mathrm{i}y,x+\mathrm{i}y)=(x,x)+(y,y)+\mathrm{i}(y,x)-\mathrm{i}(x,y),$$
$$\|x-\mathrm{i}y\|^2=(x-\mathrm{i}y,x-\mathrm{i}y)=(x,x)+(y,y)-\mathrm{i}(y,x)+\mathrm{i}(x,y),$$

故

$$\|x+\mathrm{i}y\|^2-\|x-\mathrm{i}y\|^2=-2\mathrm{i}(x,y)+2\mathrm{i}\overline{(x,y)}. \tag{9.5}$$

由(9.4)式和(9.5)式得

$$(x,y)=\frac{1}{4}\|x+y\|^2-\frac{1}{4}\|x-y\|^2+\frac{\mathrm{i}}{4}\|x+\mathrm{i}y\|^2-\frac{\mathrm{i}}{4}\|x-\mathrm{i}y\|^2,$$

由此即可得到结论. □

注 我们仅对复空间进行了讨论, 事实上对实空间, 可得下列等式:

$$(x,y)=\frac{1}{4}\|x+y\|^2-\frac{1}{4}\|x-y\|^2,$$

因此对实空间结论也成立.

定理 9.9

设 V 与 U 都是 n 维内积空间 (同为实空间或同为复空间), 若 φ 是 $V\to U$ 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变成 U 的一组标准正交基;

(4) φ 将 V 的某一组标准正交基变成 U 的一组标准正交基.

证明 (1) $\Rightarrow$ (2):φ 保持内积, 因此 φ 为单映射. 由线性映射的维数公式可得

$$\dim\operatorname{Im}\varphi=\dim V=\dim U=n,$$

因此 $\operatorname{Im}\varphi=U$, 即 φ 是映上的, 故为同构.

(2) $\Rightarrow$ (3): 设 $\{e_1,e_2,\cdots,e_n\}$ 是 V 的任意一组标准正交基. 由于 φ 保持内积, 故对 $i\neq j$, 有

$$(\varphi(e_i),\varphi(e_j))=(e_i,e_j)=0,$$

又

$$(\varphi(e_i),\varphi(e_i))=(e_i,e_i)=1.$$

这表明 $\{\varphi(e_1),\varphi(e_2),\cdots,\varphi(e_n)\}$ 是 U 的标准正交基.

(3) $\Rightarrow$ (4): 显然.

(4) $\Rightarrow$ (1): 设 $\{e_1,e_2,\cdots,e_n\}$ 是 V 的标准正交基且 $\{\varphi(e_1),\varphi(e_2),\cdots,\varphi(e_n)\}$ 是 U 的标准正交基. 假设

$$u=\sum_{i=1}^n a_ie_i,\quad v=\sum_{i=1}^n b_ie_i,$$

则

$$\varphi(u)=\sum_{i=1}^n a_i\varphi(e_i),\varphi(v)=\sum_{i=1}^n b_i\varphi(e_i),$$

于是

$$(\varphi(u),\varphi(v))=\left(\sum_{i=1}^n a_i\varphi(e_i),\sum_{i=1}^n b_i\varphi(e_i)\right)=a_1\overline{b_1}+a_2\overline{b_2}+\cdots+a_n\overline{b_n}=(u,v).$$

□

推论 9.4

两个有限维内积空间 V 与 U (同为实空间或同为复空间) 同构的充分必要条件是它们有相同的维数.

证明 只需证明充分性. 设 $\{e_1, e_2, \cdots, e_n\}$ 是 V 的标准正交基,$\{f_1, f_2, \cdots, f_n\}$ 是 U 的标准正交基, 令 φ 是 $V \to U$ 的线性映射:

$$\varphi(e_i) = f_i, \quad i = 1, 2, \cdots, n,$$

则 φ 将标准正交基变为标准正交基, 由定理 9.9可知 φ 是保积同构. □

定义 9.14

设 φ 是内积空间 V 上保持内积的线性变换,
1. 若 V 是欧氏空间, 则称 φ 为**正交变换**或**正交算子**;
2. 若 V 是酉空间, 则称 φ 为**酉变换**或**酉算子**.

♣

定理 9.10

1. 正交变换及酉变换都是可逆线性变换.
2. (1) 正交变换等价于把欧氏空间中一组标准正交基变成标准正交基的线性变换.
   (2) 酉变换等价于把酉空间中一组标准正交基变成标准正交基的线性变换.

♡

证明 由定理 9.9可直接得到证明. □

定理 9.11

设 φ 是欧氏空间或酉空间上的线性变换, 则 φ 是正交变换或酉变换的充分必要条件是 φ 非异, 且

$$\varphi^* = \varphi^{-1}.$$

♡

证明 设 φ 是欧氏空间 V 上的正交变换, 则对 V 中的任意向量 α, β, 有

$$(\varphi(\alpha), \beta) = (\varphi(\alpha), \varphi(\varphi^{-1}(\beta))) = (\alpha, \varphi^{-1}(\beta)),$$

此即 $\varphi^* = \varphi^{-1}$.

反过来, 若 $\varphi^* = \varphi^{-1}$, 则

$$(\varphi(\alpha), \varphi(\beta)) = (\alpha, \varphi^*\varphi(\beta)) = (\alpha, \beta),$$

即 φ 保持内积, 故 φ 是正交变换.

对酉变换可类似证明. □

定义 9.15 (正交矩阵和酉矩阵)

设 A 是 n 阶实方阵, 若 $A' = A^{-1}$, 则称 A 是**正交矩阵**.
设 C 是 n 阶复方阵, 若 $\overline{C}' = C^{-1}$, 则称 C 是**酉矩阵**.

♣

注 由正交矩阵与酉矩阵的定义可知正交矩阵适合条件 $AA' = A'A = I_n$, 酉矩阵适合条件 $A\overline{A}' = \overline{A}'A = I_n$. 由此可知, 正交矩阵和酉矩阵都是可逆矩阵.

定理 9.12

设 $A = (a_{ij})$ 是 n 阶实矩阵, 则 A 是正交矩阵的充分必要条件是:

$$a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{in}a_{jn} = 0, \quad i \neq j,$$
$$a_{i1}^2 + a_{i2}^2 + \cdots + a_{in}^2 = 1,$$

或

$$a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = 0, \quad i \neq j,$$
$$a_{1i}^2 + a_{2i}^2 + \cdots + a_{ni}^2 = 1.$$

也就是说,A 为正交矩阵的充分必要条件是它的 n 个行向量是 n 维实行向量空间 (取标准内积) 的标准正交基, 或它的 n 个列向量是 n 维实列向量空间 (取标准内积) 的标准正交基.

证明 由 $AA' = I_n$ 得到第一个结论, 由 $A'A = I_n$ 得到第二个结论. □

定理 9.13

设 $A = (a_{ij})$ 是 n 阶复矩阵, 则 A 是酉矩阵的充分必要条件是:

$$a_{i1}\overline{a}_{j1} + a_{i2}\overline{a}_{j2} + \cdots + a_{in}\overline{a}_{jn} = 0, \quad i \neq j,$$
$$|a_{i1}|^2 + |a_{i2}|^2 + \cdots + |a_{in}|^2 = 1,$$

或

$$a_{1i}\overline{a}_{1j} + a_{2i}\overline{a}_{2j} + \cdots + a_{ni}\overline{a}_{nj} = 0, \quad i \neq j,$$
$$|a_{1i}|^2 + |a_{2i}|^2 + \cdots + |a_{ni}|^2 = 1.$$

也就是说,A 为酉矩阵的充分必要条件是它的 n 个行向量是 n 维复行向量空间 (取标准内积) 的标准正交基, 或它的 n 个列向量是 n 维复列向量空间 (取标准内积) 的标准正交基.

证明 类似定理 9.12的证明. □

定理 9.14

若 n 阶实矩阵 A 是正交矩阵, 则
1. A 的行列式值等于 1 或 -1;
2. A 的特征值的模长等于 1.

证明
1. 由 $AA' = I_n$, 取行列式即得结论.
2. 设 λ 是 A 的特征值,x 是属于 λ 的特征向量, 则 $Ax = \lambda x$, 于是 $\overline{x}'A' = \overline{\lambda}\overline{x}'$. 因此
$$\overline{x}'A'Ax = \overline{\lambda}\overline{x}'\lambda x,$$
即
$$\overline{x}'x = \overline{\lambda}\lambda(\overline{x}'x),$$
从而 $\overline{\lambda}\lambda = 1$, 即 $|\lambda| = 1$.

□

定理 9.15

若 n 阶复矩阵 A 是酉矩阵, 则
1. A 的行列式值的模长等于 1;
2. A 的特征值的模长等于 1.

证明 类似定理 9.14的证明. □

命题 9.4

(1) 单位阵是正交矩阵也是酉矩阵;
(2) 对角阵是正交矩阵的充分必要条件是主对角线上的元素为 1 或 -1.

证明 证明是显然的. □

定理 9.16

设 φ 是欧氏空间 (酉空间) V 上的线性变换, 则 φ 是正交变换 (酉变换) 的充分必要条件是在 V 的任一组标准正交基下,φ 的表示矩阵是正交矩阵 (酉矩阵).

证明 必要性: 由定理 9.11, 当 φ 是正交变换时, 若 φ 在 V 的一组标准正交基下的表示矩阵为 A, 则 φ^* 在同一组基下的表示矩阵为 A', 由 $\varphi^* = \varphi^{-1}$ 得 $A' = A^{-1}$, 即 A 是正交矩阵. 同理, 当 φ 是酉变换时,φ 在一组标准正交基下的表示矩阵 A 应适合 $\overline{A}' = A^{-1}$, 即 A 是酉矩阵.

充分性: 这由线性变换与其表示矩阵的关系即得. □

定理 9.17 (矩阵的 QR 分解)

设 A 是 n 阶实 (复) 矩阵, 则 A 可分解为

$$A = QR,$$

其中 Q 是正交 (酉) 矩阵,R 是一个主对角线上的元素均大于等于零的上三角阵, 并且若 A 是非异阵, 则这样的分解必唯一.

证明 设 A 是 n 阶实矩阵,$A = (u_1, u_2, \cdots, u_n)$ 是 A 的列分块. 考虑 n 维实列向量空间 $\mathbb{R}^n$, 并取其标准内积, 我们先通过类似于 Gram - Schmidt 方法的正交化过程, 把 $\{u_1, u_2, \cdots, u_n\}$ 变成一组两两正交的向量 $\{w_1, w_2, \cdots, w_n\}$, 并且 w_k 或者是零向量或者是单位向量.

我们用数学归纳法来定义上述向量. 假设 $w_1, \cdots, w_{k-1}$ 已经定义好, 现来定义 w_k. 令

$$v_k = u_k - \sum_{j=1}^{k-1} (u_k, w_j) w_j.$$

若 $v_k = 0$, 则令 $w_k = 0$; 若 $v_k \neq 0$, 则令 $w_k = \dfrac{v_k}{\|v_k\|}$. 容易验证 $\{w_1, w_2, \cdots, w_n\}$ 是一组两两正交的向量,w_k 或者是零向量或者是单位向量, 并且满足

$$u_k = \sum_{j=1}^{k-1} (u_k, w_j) w_j + \|v_k\| w_k, \quad k = 1, 2, \cdots, n. \tag{9.6}$$

由(9.6)式可得

$$A = (u_1, u_2, \cdots, u_n) = (w_1, w_2, \cdots, w_n) R, \tag{9.7}$$

其中 R 是一个上三角阵且主对角线上的元素依次为 $\|v_1\|, \|v_2\|, \cdots, \|v_n\|$, 均大于等于零, 并且由(9.7)式知, 如果 $w_k = 0$, 则 R 的第 k 行元素全为零.

假设 $w_{i_1}, w_{i_2}, \cdots, w_{i_r}$ 是其中的非零向量全体, 由定理 9.11可将它们扩张为 $\mathbb{R}^n$ 的一组标准正交基 $\{\widetilde{w}_1, \widetilde{w}_2, \cdots, \widetilde{w}_n\}$, 其中 $\widetilde{w}_j = w_j, j = i_1, i_2, \cdots, i_r$. 令 $Q = (\widetilde{w}_1, \widetilde{w}_2, \cdots, \widetilde{w}_n)$, 由定理 9.12知 Q 是正交矩阵. 注意到若 $w_k = 0$, 则 R 的第 k 行元素全为零, 此时用 $\widetilde{w}_k$ 代替 w_k 仍然可使(9.7)式成立, 因此

$$A = (u_1, u_2, \cdots, u_n) = (\widetilde{w}_1, \widetilde{w}_2, \cdots, \widetilde{w}_n) R = QR,$$

从而得到了 A 的 QR 分解.

复矩阵情形的证明完全类似. 至于非异阵 QR 分解的唯一性, 利用摄动法不难证明. □

例题 9.2 求下列矩阵的 QR 分解:

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

解 采用与定理 9.17证明中相同的记号, 经过计算可得:

$$v_1 = u_1 = (1, 1, 0)', w_1 = \frac{1}{\sqrt{2}}(1, 1, 0)';$$

$$v_2 = u_2 - \sqrt{2}w_1 = (1,-1,1)', w_2 = \frac{1}{\sqrt{3}}(1,-1,1)';$$
$$v_3 = u_3 - 3\sqrt{2}w_1 - 2\sqrt{3}w_2 = (0,0,0)', w_3 = (0,0,0)',$$

从而有

$$A = (u_1,u_2,u_3) = (w_1,w_2,w_3)\begin{pmatrix} \sqrt{2} & \sqrt{2} & 3\sqrt{2} \\ 0 & \sqrt{3} & 2\sqrt{3} \\ 0 & 0 & 0 \end{pmatrix}.$$

用 $\widetilde{w}_3 = \frac{1}{\sqrt{6}}(-1,1,2)'$ 代替 w_3 可得 A 的 QR 分解为

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{pmatrix}, \quad R = \begin{pmatrix} \sqrt{2} & \sqrt{2} & 3\sqrt{2} \\ 0 & \sqrt{3} & 2\sqrt{3} \\ 0 & 0 & 0 \end{pmatrix}.$$

□

9.5 自伴随算子

引理 9.4

欧氏空间中两组标准正交基之间的过渡矩阵是正交矩阵;

酉空间中两组标准正交基之间的过渡矩阵是酉矩阵.

♡

证明 设 V 是 n 维欧氏空间,$\{e_1,e_2,\cdots,e_n\}$ 和 $\{f_1,f_2,\cdots,f_n\}$ 是 V 的两组标准正交基且

$$\begin{cases} f_1 = a_{11}e_1 + a_{21}e_2 + \cdots + a_{n1}e_n, \\ f_2 = a_{12}e_1 + a_{22}e_2 + \cdots + a_{n2}e_n, \\ \cdots\cdots\cdots\cdots \\ f_n = a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{nn}e_n. \end{cases}$$

因为 $(f_i,f_i) = 1$, 故

$$a_{1i}^2 + a_{2i}^2 + \cdots + a_{ni}^2 = 1. \tag{9.8}$$

又若 $i \neq j$, 则 $(f_i,f_j) = 0$, 故

$$a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = 0. \tag{9.9}$$

(9.8)式、(9.9)式和定理 9.12表明过渡矩阵

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

是正交矩阵. 同理可证明酉空间中两组标准正交基之间的过渡矩阵是酉矩阵. □

定义 9.16 (正交相似和酉相似)

1. 设 A,B 是 n 阶实矩阵, 若存在正交矩阵 P, 使

$$B = P'AP = P^{-1}AP.$$

则称 B 和 A **正交相似**.

2. 设 A, B 是 n 阶复矩阵, 若存在酉矩阵 P, 使

$$B = \overline{P}' AP = P^{-1}AP.$$

则称 B 和 A **酉相似**.

注 和矩阵的相似关系一样, 我们不难证明正交 (酉) 相似关系是等价关系, 即:

1. n 阶矩阵 A 和自己正交 (酉) 相似;
2. 若 B 和 A 正交 (酉) 相似, 则 A 和 B 也正交 (酉) 相似;
3. 若 B 和 A 正交 (酉) 相似,C 和 B 正交 (酉) 相似, 则 C 和 A 也正交 (酉) 相似.

定义 9.17 (自伴随算子)

设 φ 是内积空间 V 上的线性变换,φ^* 是 φ 的伴随, 若 $\varphi^* = \varphi$, 则称 φ 是**自伴随算子**.
当 V 是欧氏空间时,φ 也称为**对称算子**或**对称变换**;
当 V 是酉空间时,φ 也称为 **Hermite 算子**或 **Hermite 变换**.

命题 9.5

设 V 是 n 维内积空间,V_0 是 V 的子空间,$V = V_0 \oplus V_0^{\perp}$. 令 E 是 V 到 V_0 上的正交投影, 则 E 是自伴随算子.

证明 由命题 9.1立得. □

定理 9.18

1. 设 V 是 n 维欧氏空间,φ 是 V 上的线性变换, 则 φ 是自伴随算子的充要条件是其在任一组标准正交基下的表示矩阵都是实对称阵.
2. 设 V 是 n 维酉空间,φ 是 V 上的线性变换, 则 φ 是自伴随算子的充要条件是其在任一组标准正交基下的表示矩阵都是实对称阵.

证明

1. 设 $\{e_1, e_2, \cdots, e_n\}$ 是 V 中的一组标准正交基, 若 φ 在这组基下的表示矩阵为 A, 则 φ^* 在同一组基下的表示矩阵为 A'. 若 $\varphi^* = \varphi$, 则 $A' = A$. 也就是说欧氏空间上的自伴随算子在任一组标准正交基下的表示矩阵都是实对称阵.
   反之亦容易看出, 若欧氏空间上的线性变换 φ 在某一组标准正交基下的表示矩阵是实对称阵, 则 $\varphi^* = \varphi$.
2. 同理可证明对酉空间也有类似结论.

□

定理 9.19

设 V 是 n 维酉空间,φ 是 V 上的自伴随算子, 则 φ 的特征值全是实数且属于不同特征值的特征向量互相正交.

证明 设 λ 是 φ 的特征值,x 是属于 λ 的特征向量, 则

$$\begin{aligned}\lambda(x,x) &= (\lambda x, x) = (\varphi(x), x) = (x, \varphi^*(x)) \\ &= (x, \varphi(x)) = (x, \lambda x) = \overline{\lambda}(x,x).\end{aligned}$$

因为 $(x,x) \neq 0$, 故 $\overline{\lambda} = \lambda$, 即 λ 是实数. 又若设 μ 是 φ 的另一个特征值,y 是属于 μ 的特征向量, 注意到 λ, μ 都是实数, 故有

$$\begin{aligned}\lambda(x,y) &= (\lambda x, y) = (\varphi(x), y) = (x, \varphi^*(y)) \\ &= (x, \varphi(y)) = (x, \mu y) = \mu(x,y).\end{aligned}$$

由于 $\lambda \neq \mu$, 故 $(x, y) = 0$, 即 $x \perp y$. □

推论 9.5

Hermite 矩阵的特征值全是实数, 实对称阵的特征值也全是实数. 这两种矩阵属于不同特征值的特征向量互相正交.

证明 Hermite 矩阵的结论是定理的显然推论, 而实对称阵也是 Hermite 矩阵, 因此结论成立. □

定理 9.20

设 V 是 n 维内积空间,φ 是 V 上的自伴随算子, 则存在 V 的一组标准正交基, 使 φ 在这组基下的表示矩阵为实对角阵, 且这组基恰为 φ 的 n 个线性无关的特征向量.

证明 首先需要说明的是, 若 V 是欧氏空间, 则由于自伴随算子 φ 的特征值都是实数, 故有实的特征向量. 不妨设 u 是 φ 的特征向量, 令 $v_1 = \dfrac{u}{\|u\|}$, 则 v_1 是 φ 的长度等于 1 的特征向量. 我们对维数 n 用归纳法.

若 $\dim V = 1$, 结论已成立. 设对小于 n 维的内积空间结论成立. 令 W 为由 v_1 张成的子空间,$W^{\perp}$ 为 W 的正交补空间, 则 W 是 φ 的不变子空间且

$$V = W \oplus W^{\perp}, \quad \dim W^{\perp} = n - 1.$$

由命题 9.3.1 可知 $W^{\perp}$ 是 $\varphi^* = \varphi$ 的不变子空间. 将 φ 限制在 $W^{\perp}$ 上仍是自伴随算子. 由归纳假设, 存在 $W^{\perp}$ 的一组标准正交基 $\{v_2, \cdots, v_n\}$, 使 φ 在这组基下的表示矩阵为实对角阵, 且 $\{v_2, \cdots, v_n\}$ 是其特征向量. 因此,$\{v_1, v_2, \cdots, v_n\}$ 构成了 V 的一组标准正交基,φ 在这组基下的表示矩阵为实对角阵, 且 $\{v_1, v_2, \cdots, v_n\}$ 为 φ 的 n 个线性无关的特征向量. □

定理 9.21

1. 设 A 是 n 阶实对称阵, 则存在正交矩阵 P, 使 $P'AP$ 为对角阵, 且 P 的 n 个列向量恰为 A 的 n 个两两正交的单位特征向量.
2. 设 A 是 n 阶 Hermite 矩阵, 则存在酉矩阵 P, 使 $\overline{P}'AP$ 为实对角阵, 且 P 的 n 个列向量恰为 A 的 n 个两两正交的单位特征向量.

证明 由定理 9.20即知实对称阵正交相似于对角阵,Hermite 矩阵酉相似于实对角阵. 从对角化相关理论知道 P 的列向量都是 A 的特征向量, 又 P 是正交 (酉) 矩阵, 故这些列向量两两正交且长度等于 1. □

推论 9.6

实对称阵的全体特征值是实对称阵在正交相似关系下的全系不变量,Hermite 矩阵的全体特征值是 Hermite 矩阵在酉相似关系下的全系不变量.

证明 只证明实的情形. 显然正交相似的矩阵有相同的特征值. 另一方面, 由定理 9.21知道只需对对角阵证明, 若它们的特征值相同, 则必正交相似即可. 设

$$B = \operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}, \quad D = \operatorname{diag}\{\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_n}\},$$

其中 $\{\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_n}\}$ 是 $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ 的一个排列. 由于任一排列可通过若干次对换来实现, 因此只要证明对 B 的第 (i, i) 元素和第 (j, j) 元素对换后得到的矩阵与 B 正交相似即可. 设 P_{ij} 是第一类初等矩阵, 则 P_{ij} 是正交矩阵且 $P_{ij}' = P_{ij}$, 因此 $P_{ij}'BP_{ij}$ 和 B 正交相似. 这就是我们要证明的结论. □

定理 9.22

设 $f(x) = x'Ax$ 是 n 元实二次型, 系数矩阵 A 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则 f 经过正交变换 $x = Py$ 可以化为下列标准型:

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

因此,f 的正惯性指数等于 A 的正特征值的个数, 负惯性指数等于 A 的负特征值的个数,f 的秩等于 A 的非零特征值的个数.

证明 注意到正交相似既是相似又是合同, 故由定理 9.21即得结论. □

推论 9.7

设 $f(x)=x'Ax$ 是 n 元实二次型, 则

1. f 是正定型当且仅当系数矩阵 A 的特征值全是正数,
2. f 是负定型当且仅当 A 的特征值全是负数,
3. f 是半正定型当且仅当 A 的特征值全非负,
4. f 是半负定型当且仅当 A 的特征值全非正.

下面我们通过具体例子来说明对实对称阵 A, 如何求正交矩阵 P, 使 $P'AP$ 是对角阵. 我们的方法和 §6.2 类似, 只是增加了特征向量的标准正交化过程.

例题 9.3 求正交矩阵 P, 使 $P'AP$ 为对角阵, 其中

$$A=\begin{pmatrix}4&2&2\\2&4&2\\2&2&4\end{pmatrix}.$$

解 先求特征值

$$|\lambda I-A|=\begin{vmatrix}\lambda-4&-2&-2\\-2&\lambda-4&-2\\-2&-2&\lambda-4\end{vmatrix}=(\lambda-8)(\lambda-2)^2.$$

因此,A 的特征值为 $\lambda_1=8,\lambda_2=\lambda_3=2$. 当 $\lambda=8$ 时, 求解齐次线性方程组 $(\lambda I-A)x=0$, 得到基础解系 (只有一个向量):

$$\eta_1=(1,1,1)'.$$

当 $\lambda=2$ 时, 求解齐次线性方程组 $(\lambda I-A)x=0$, 得到基础解系 (有两个向量):

$$\eta_2=(-1,1,0)',\quad \eta_3=(-1,0,1)'.$$

由于实对称阵属于不同特征值的特征向量必正交, 因此只要对上面两个向量正交化, 用 Gram - Schmidt 方法将 η_2,η_3 正交化得到

$$\xi_2=(-1,1,0)',\quad \xi_3=(-\frac{1}{2},-\frac{1}{2},1)'.$$

再将 η_1,ξ_2,ξ_3 化为单位向量得到

$$v_1=(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}})',\quad v_2=(-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0)',\quad v_3=(-\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}})'.$$

令

$$P=(v_1,v_2,v_3)=\begin{pmatrix}\frac{1}{\sqrt{3}}&-\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{3}}&0&\frac{2}{\sqrt{6}}\end{pmatrix},$$

于是

$$P'AP=\begin{pmatrix}8&0&0\\0&2&0\\0&0&2\end{pmatrix}.$$

□

注 上述方法也适用于 Hermite 矩阵, 即对 Hermite 矩阵 A, 求酉矩阵 P, 使 $\overline{P}'AP$ 为对角阵.

例题 9.4 设 A 是三阶实对称阵,A 的特征值为 $0,3,3$. 已知属于特征值 0 的特征向量为 $v_1 = (1,1,1)'$, 又向量 $v_2 = (-1,1,0)'$ 是属于特征值 3 的特征向量, 求矩阵 A.

解 设 A 属于特征值 3 的另一特征向量 $v_3 = (x_1,x_2,x_3)'$ 和 v_1, v_2 都正交, 则

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

求出一个非零解 $v_3 = (1,1,-2)'$. 因为 v_1, v_2, v_3 已经两两正交, 故只需将 v_1, v_2, v_3 标准化, 得到

$$\xi_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \quad \xi_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \xi_3 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{pmatrix}.$$

令

$$P = (\xi_1, \xi_2, \xi_3) = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

则

$$A = PBP' = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

□

例题 9.5 设 A 是 n 阶实对称阵且 $A^3 = I_n$, 证明: $A = I_n$.

解 设 A 的特征值为 λ, 则 $\lambda^3 = 1$. 因为 λ 是实数, 故 $\lambda = 1$, 这就是说 A 的特征值全是 1. 由定理 9.21, 存在正交矩阵 P, 使 $P'AP = I_n$, 于是 $A = PI_nP' = I_n$. □

9.6 复正规算子

定义 9.18 (正规算子和正规矩阵)

设 φ 是内积空间 V 上的线性变换,φ^* 是其伴随, 若 $\varphi\varphi^* = \varphi^*\varphi$, 则称 φ 是 V 上的**正规算子**.

为了不引起混淆, 我们也称酉空间 (欧氏空间) V 上的正规算子 φ 为**复正规算子** (**实正规算子**).

复矩阵 A 若适合 $\overline{A}'A = A\overline{A}'$, 则称其为**复正规矩阵**.

实矩阵 A 若适合 $A'A = AA'$, 则称其为**实正规矩阵**.

♣

命题 9.6

1. 酉算子 (酉矩阵) 和 Hermite 算子 (Hermite 矩阵) 都是复正规算子 (矩阵).
2. 正交变换 (正交矩阵) 和对称变换 (实对称矩阵) 都是实正规算子 (矩阵).

♠

证明 证明都是显然的. □

定理 9.23

酉空间 (欧氏空间) V 上的线性变换 φ 是复 (实) 正规算子的充分必要条件是 φ 在 V 的某一组或任一组标准正交基下的表示矩阵都是复 (实) 正规矩阵. 因此, 复 (实) 矩阵的正规性在酉 (正交) 相似下是不变的.

证明 证明都是显然的. □

引理 9.5

设 φ 是内积空间 V 上的正规算子, 则对任意的 $\alpha \in V$, 成立

$$\|\varphi(\alpha)\| = \|\varphi^*(\alpha)\|.$$

证明 由 φ 的正规性, 有

$$\begin{aligned}\|\varphi(\alpha)\|^2 &= (\varphi(\alpha), \varphi(\alpha)) = (\alpha, \varphi^*\varphi(\alpha)) \\ &= (\alpha, \varphi\varphi^*(\alpha)) = (\varphi^*(\alpha), \varphi^*(\alpha)) \\ &= \|\varphi^*(\alpha)\|^2.\end{aligned}$$

□

命题 9.7

设 V 是 n 维酉空间,φ 是 V 上的正规算子.

(1) 向量 u 是 φ 属于特征值 λ 的特征向量的充分必要条件为 u 是 φ^* 属于特征值 $\overline{\lambda}$ 的特征向量;

(2) 属于 φ 不同特征值的特征向量必正交.

证明

(1) 若 λ 是任一数, 则 $(\lambda I - \varphi)^* = \overline{\lambda} I - \varphi^*$, 且

$$(\lambda I - \varphi)(\overline{\lambda} I - \varphi^*) = (\overline{\lambda} I - \varphi^*)(\lambda I - \varphi),$$

即 $\lambda I - \varphi$ 也是正规算子. 于是由引理 9.5,

$$\|(\lambda I - \varphi)(\alpha)\| = \|(\overline{\lambda} I - \varphi^*)(\alpha)\|$$

对一切 $\alpha \in V$ 成立, 故 $(\lambda I - \varphi)(u) = 0$ 当且仅当 $(\overline{\lambda} I - \varphi^*)(u) = 0$ 成立.

(2) 设 $\varphi(u) = \lambda u, \varphi(v) = \mu v$ 且 $\lambda \neq \mu$, 则由 (1) 知 $\varphi^*(v) = \overline{\mu} v$, 于是

$$\lambda(u, v) = (\lambda u, v) = (\varphi(u), v) = (u, \varphi^*(v)) = (u, \overline{\mu} v) = \mu(u, v).$$

因为 $\lambda \neq \mu$, 故 $(u, v) = 0$.

□

引理 9.6

设 V 是 n 维酉空间,φ 是 V 上的线性变换, 又 $\{e_1, e_2, \cdots, e_n\}$ 是 V 的一组标准正交基. 设 φ 在这组基下的表示矩阵 A 是一个上三角阵, 则 φ 是正规算子的充分必要条件是 A 为对角阵.

证明 若 A 是对角阵, 则 $A\overline{A}' = \overline{A}'A$, 故 $\varphi\varphi^* = \varphi^*\varphi$, 即 φ 是正规算子. 反之, 设 φ 是正规算子. 由于 A 是上三角阵, 可记 $A = (a_{ij}), a_{ij} = 0 (i > j)$. 于是 $\varphi(e_1) = a_{11}e_1$, 再由上面的命题可知 $\varphi^*(e_1) = \overline{a}_{11}e_1$. 另一方面, 有

$$\varphi^*(e_1) = \overline{a}_{11}e_1 + \overline{a}_{12}e_2 + \cdots + \overline{a}_{1n}e_n.$$

因此 $a_{1j} = 0$ 对一切 $j > 1$ 成立. 又因为 A 是上三角阵, 所以

$$\varphi(e_2) = a_{22}e_2,$$

故又有 $\varphi^*(e_2) = \overline{a}_{22}e_2$ 及 $a_{2j} = 0 (j > 2)$. 不断这样做下去即得 A 是对角阵. □

定理 9.24 (Schur(舒尔) 定理)

设 V 是 n 维酉空间,φ 是 V 上的线性算子, 则存在 V 的一组标准正交基, 使 φ 在这组基下的表示矩阵为上三角阵.

证明 对 V 的维数 n 用数学归纳法. 当 $n=1$ 时结论显然成立. 设对 $n-1$ 维酉空间结论成立, 现证 n 维酉空间的情形. 由于 V 是复线性空间, 故 φ^* 总存在特征值与特征向量, 即有

$$\varphi^*(e)=\lambda e.$$

设 W 是由 e 张成的一维子空间的正交补空间, 由命题 9.3知 W 是 $(\varphi^*)^*=\varphi$ 的不变子空间, 将 φ 限制在 W 上得到 W 上的一个线性变换. 注意到 $\dim W=n-1$, 故由归纳假设, 存在 W 的一组标准正交基 $\{e_1,e_2,\cdots,e_{n-1}\}$, 使 $\varphi|_W$ 在这组基下的表示矩阵为上三角阵. 令 $e_n=\dfrac{e}{\|e\|}$, 则 $\{e_1,e_2,\cdots,e_n\}$ 成为 V 的一组标准正交基, 使 φ 在这组基下的表示矩阵为上三角阵. □

推论 9.8 (Schur 定理)

任一 n 阶复矩阵均酉相似于一个上三角阵.

定理 9.25

设 V 是 n 维酉空间,φ 是 V 上的线性算子, 则 φ 为正规算子的充分必要条件是存在 V 的一组标准正交基, 使 φ 在这组基下的表示矩阵是对角阵. 特别, 这组基恰为 φ 的 n 个线性无关的特征向量.

证明 利用引理 9.6和Schur 定理, 我们立即得到证明. □

定理 9.26

复矩阵 A 为复正规矩阵的充分必要条件是 A 酉相似于对角阵.

证明 利用引理 9.6和Schur 定理, 我们立即得到证明. □

定理 9.27

复正规矩阵的特征值就是复正规矩阵在酉相似关系下的全系不变量, 即两个复正规矩阵酉相似的充分必要条件是它们具有相同的特征值.

证明

□

命题 9.8

设 φ 是 n 维酉空间 V 上的线性算子,$\lambda_1,\lambda_2,\cdots,\lambda_k$ 是 φ 的全体不同特征值,$V_1,V_2,\cdots,V_k$ 是对应的特征子空间, 则 φ 是正规算子的充分必要条件是

$$V=V_1\perp V_2\perp\cdots\perp V_k. \tag{9.10}$$

证明 设 φ 是正规算子, 则它是一个可对角化线性变换, 因此

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

又从命题 9.7知道, 若 $i\neq j$, 则 $V_i\perp V_j$, 所以 (9.10)式成立.

反之, 若(9.10)式成立, 则在每个 V_i 中取一组标准正交基, 将这些基向量组成 V 的一组标准正交基. 因为每个 V_i 都是 φ 的特征子空间, 即 $\varphi(\alpha)=\lambda_i\alpha$ 对一切 $\alpha\in V_i$ 成立, 故 φ 在这组基下的表示矩阵是对角阵, 因此 φ 是正规算子. □

定理 9.28

任一 n 阶酉矩阵必酉相似于下列对角阵:

$$\mathrm{diag}\{c_1, c_2, \cdots, c_n\},$$

其中 c_i 为模长等于 1 的复数.

证明 由命题 9.6及定理 9.26知酉矩阵酉相似于 $\mathrm{diag}\{c_1, c_2, \cdots, c_n\}$. 由于与酉矩阵酉相似的矩阵仍是酉矩阵, 故 $\mathrm{diag}\{c_1,c_2,\cdots,c_n\}$ 是酉矩阵, 因此 $|c_i| = 1$. □