0.1 极小多项式与 Cayley-Hamilton 定理

命题 0.1

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 0.1 (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 0.1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 0.1 (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

$$\varphi(\alpha)=k_1\varphi(e_1)+k_2\varphi(e_2)+\cdots+k_n\varphi(e_n)=(\varphi(e_1),\varphi(e_2),\cdots,\varphi(e_n))\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}$$

$$=(e_1,e_2,\cdots,e_n)A\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=A\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=A(e_1,e_2,\cdots,e_n)\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=A\alpha.$$

因此 $f(\varphi)(\alpha)=f(A)(\alpha)=0$. 故由 α 的任意性可知 $f(\varphi)=O$.

□

0.1.1 极小多项式的性质

命题 0.2 (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 $f(x)$. 即极小多项式整除任意零化多项式.

(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**??**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 0.3

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 0.4

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 0.1

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 0.4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 0.1 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 0.1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 0.5 (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 $\lambda_i(1\le i\le n)$, 则 $f(x)=(x-\lambda_1)\cdots(x-\lambda_n)$. 由推论 0.1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 0.4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**??**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 0.6

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 0.1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 0.1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 0.7

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 $g(\lambda_i) \neq 0 (i = 1, 2, \cdots, m)$, 即 λ_i 都不是 $g(x)$ 的根. 而由推论 0.1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 0.8

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 0.1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. □

0.1.2 Cayley-Hamilton 定理的应用: 逆矩阵和伴随矩阵的多项式表示

命题 0.9

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 0.10

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 0.9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 $a_i(0) = a_i(1 \le i \le n)$. 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 $\mathrm{r}(A) \le n-2$, 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**??**可知结论也成立. □

0.1.3 Cayley - Hamilton 定理的应用:AX = XB 型矩阵方程的求解及其应用

命题 0.11

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 0.6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 $\mathrm{r}(\boldsymbol{C}) = r \ge 1$. 由例题**??**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 0.2 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 0.11可得 $\boldsymbol{A}-\boldsymbol{B} = \boldsymbol{O}$, 即 $\boldsymbol{A} = \boldsymbol{B}$. □

例题 0.3 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 0.11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 0.12

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 0.11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??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 $\lambda_i - \mu_j(1\le i\le m; 1\le j\le n)$, 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 $\lambda_i - \mu_j(1\le i\le m; 1\le j\le n)$ 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 0.4 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 $\boldsymbol{A}'\boldsymbol{B}+\boldsymbol{B}\boldsymbol{A}=\boldsymbol{C}$.

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 0.12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. □

0.1.4 Cayley-Hamilton 定理的应用: 特征多项式诱导的直和分解

例题 0.5 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: $\mathrm{Ker}\,\sigma\tau=\mathrm{Ker}\,\sigma\oplus\mathrm{Ker}\,\tau$.

笔记 $(f(x),g(x))=1$ 之后的证明类似命题??.

证明 由命题??可知 C 的特征多项式就是 $f(x)$. 由命题 0.7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

$$u(\varphi)f(\varphi)+v(\varphi)g(\varphi)=I_V. \tag{1}$$

于是对 $\forall\alpha\in\mathrm{Ker}\,\sigma\tau$, 由 (1)式可得

$$\alpha=u(\varphi)f(\varphi)(\alpha)+v(\varphi)g(\varphi)(\alpha).$$

又因为 $\alpha\in\mathrm{Ker}\,\sigma\tau$, 所以 $f(\varphi)g(\varphi)(\alpha)=g(\varphi)f(\varphi)(\alpha)=0$. 因此 $u(\varphi)f(\varphi)(\alpha)\in\mathrm{Ker}\,g(\varphi)$,$v(\varphi)g(\varphi)(\alpha)\in\mathrm{Ker}\,f(\varphi)$. 故有 $\mathrm{Ker}\,\sigma\tau=\mathrm{Ker}\,\sigma+\mathrm{Ker}\,\tau$. 任取 $\beta\in\mathrm{Ker}\,\sigma\cap\mathrm{Ker}\,\tau$, 则 $\sigma(\beta)=f(\varphi)(\beta)=0,\tau(\beta)=g(\varphi)(\beta)=0$. 由(1) 式可得

$$\beta=u(\varphi)f(\varphi)(\beta)+v(\varphi)g(\varphi)(\beta)=0.$$

故 $\mathrm{Ker}\,\sigma\cap\mathrm{Ker}\,\tau=0$, 因此 $\mathrm{Ker}\,\sigma\tau=\mathrm{Ker}\,\sigma\oplus\mathrm{Ker}\,\tau$. □

命题 0.13

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 $V_1=\mathrm{Ker}\,f_1(\varphi),V_2=\mathrm{Ker}\,f_2(\varphi)$, 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) $\varphi|_{V_1}$ 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??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

$$f_1(\lambda)=(\lambda-x_1)^{i_1}\cdots(\lambda-x_s)^{i_s},\quad f_2(\lambda)=(\lambda-y_1)^{j_1}\cdots(\lambda-y_l)^{j_l}, \tag{2}$$

$$g_1(\lambda)=(\lambda-x_1)^{i_1'}\cdots(\lambda-x_s)^{i_s'},\quad g_2(\lambda)=(\lambda-y_1)^{j_1'}\cdots(\lambda-y_l)^{j_l'}. \tag{3}$$

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(2)和(3)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**??**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**??**可得

$$f(\lambda)=f_1(\lambda)f_2(\lambda)=g_1(\lambda)g_2(\lambda). \tag{4}$$

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(4) 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 0.14

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

$$m(\lambda)=m_1(\lambda)m_2(\lambda), V_1=\operatorname{Ker} m_1(\varphi), V_2=\operatorname{Ker} m_2(\varphi),$$

则 $\varphi|_{V_i}$ 的极小多项式就是 $m_i(\lambda)$.

注 这个命题是命题 0.13的推广.

证明 由命题 0.13完全类似的讨论可证. □

0.1.5 Cayley-Hamilton 定理的其他应用

例题 0.6 设 $\boldsymbol{A}$ 为 n 阶矩阵,$\boldsymbol{C}$ 为 $k\times n$ 矩阵, 且对任意的 $\lambda\in\mathbb{C}$,$\begin{pmatrix}\boldsymbol{A}-\lambda\boldsymbol{I}_n\\ \boldsymbol{C}\end{pmatrix}$ 均为列满秩阵. 证明: 对任意的 $\lambda\in\mathbb{C}$,$\begin{pmatrix}\boldsymbol{C}\\ \boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{I}_n)\\ \boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{I}_n)^2\\ \vdots\\ \boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{I}_n)^{n-1}\end{pmatrix}$ 均为列满秩阵.

证明 由推论??可知, 对任意的 $\lambda\in\mathbb{C}$, 下列线性方程组只有零解:

$$\begin{cases}(\boldsymbol{A}-\lambda\boldsymbol{I}_n)\boldsymbol{x}=\boldsymbol{0},\\ \boldsymbol{C}\boldsymbol{x}=\boldsymbol{0}.\end{cases} \tag{5}$$

而要证明结论, 根据推论??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:

$$\begin{cases}\boldsymbol{C}\boldsymbol{x}=\boldsymbol{0},\\ \boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{I}_n)\boldsymbol{x}=\boldsymbol{0},\\ \boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{I}_n)^2\boldsymbol{x}=\boldsymbol{0},\\ \cdots\cdots\cdots\\ \boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{I}_n)^{n-1}\boldsymbol{x}=\boldsymbol{0}.\end{cases} \tag{6}$$

任取 $\lambda_0\in\mathbb{C}$ 以及对应线性方程组 (6)的任一解 $\boldsymbol{x}_0$, 则由线性方程组 (6)可得 $\boldsymbol{C}\boldsymbol{x}_0=\boldsymbol{0}$,$\boldsymbol{C}\boldsymbol{A}\boldsymbol{x}_0=\boldsymbol{0}$,$\cdots$,$\boldsymbol{C}\boldsymbol{A}^{n-1}\boldsymbol{x}_0=\boldsymbol{0}$, 因此对任意次数小于 n 的多项式 $g(x)$, 均有 $\boldsymbol{C}g(\boldsymbol{A})\boldsymbol{x}_0=\boldsymbol{0}$. 设

$$f(\lambda)=|\lambda\boldsymbol{I}_n-\boldsymbol{A}|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

为 $\boldsymbol{A}$ 的特征多项式, 则由 Cayley - Hamilton 定理可得

$$(\boldsymbol{A}-\lambda_1\boldsymbol{I}_n)(\boldsymbol{A}-\lambda_2\boldsymbol{I}_n)\cdots(\boldsymbol{A}-\lambda_n\boldsymbol{I}_n)=\boldsymbol{O}.$$

因此 $\boldsymbol{y}=(\boldsymbol{A}-\lambda_2\boldsymbol{I}_n)\cdots(\boldsymbol{A}-\lambda_n\boldsymbol{I}_n)\boldsymbol{x}_0$ 既满足 $(\boldsymbol{A}-\lambda_1\boldsymbol{I}_n)\boldsymbol{y}=\boldsymbol{0}$, 又满足 $\boldsymbol{C}\boldsymbol{y}=\boldsymbol{0}$, 故由线性方程组(5)只有零解可得 $\boldsymbol{y}=(\boldsymbol{A}-\lambda_2\boldsymbol{I}_n)\cdots(\boldsymbol{A}-\lambda_n\boldsymbol{I}_n)\boldsymbol{x}_0=\boldsymbol{0}$. 不断重复上述论证, 最后可得 $\boldsymbol{x}_0=\boldsymbol{0}$, 结论得证. □

例题 0.7 设 $\boldsymbol{A}$ 是 n 阶矩阵,$\boldsymbol{B}$ 是 $n\times m$ 矩阵, 分块矩阵 $(\boldsymbol{B},\boldsymbol{A}\boldsymbol{B},\cdots,\boldsymbol{A}^{n-2}\boldsymbol{B},\boldsymbol{A}^{n-1}\boldsymbol{B})$ 的秩为 r. 证明: 存在 n 阶可逆矩阵 $\boldsymbol{P}$, 使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\boldsymbol{A}_{11}&\boldsymbol{A}_{12}\\ \boldsymbol{O}&\boldsymbol{A}_{22}\end{pmatrix},\quad \boldsymbol{P}^{-1}\boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_1\\ \boldsymbol{O}\end{pmatrix},$$

其中 $\boldsymbol{A}_{11}$ 是 r 阶矩阵,$\boldsymbol{B}_1$ 是 $r\times m$ 矩阵.

注 $\boldsymbol{A}\boldsymbol{\alpha}_i(1\le i\le r)$ 都是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 的线性组合的原因: 因为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 是 $(\boldsymbol{B},\boldsymbol{A}\boldsymbol{B},\cdots,\boldsymbol{A}^{n-2}\boldsymbol{B},\boldsymbol{A}^{n-1}\boldsymbol{B})$ 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 $\boldsymbol{\alpha}_i$ 是 $\boldsymbol{A}^k\boldsymbol{B}$ 的某一列向量.

当 $\boldsymbol{\alpha}_i$ 是 $\boldsymbol{A}^k\boldsymbol{B}(0\leqslant k\leqslant n-2)$ 的某一列向量时, 则 $\boldsymbol{A}\boldsymbol{\alpha}_i$ 一定是 $\boldsymbol{A}^{k+1}\boldsymbol{B}$ 的某一列向量, 又由于 $1\leqslant k+1\leqslant n-1$, 因此 $\boldsymbol{A}\boldsymbol{\alpha}_i$ 仍是 $(\boldsymbol{B},\boldsymbol{A}\boldsymbol{B},\cdots,\boldsymbol{A}^{n-2}\boldsymbol{B},\boldsymbol{A}^{n-1}\boldsymbol{B})$ 的某一列向量, 从而 $\boldsymbol{A}\boldsymbol{\alpha}_i$ 可由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$ 线性表出.

当 $\boldsymbol{\alpha}_i$ 是 $\boldsymbol{A}^{n-1}\boldsymbol{B}$ 的某一列向量时, 则 $\boldsymbol{A}\boldsymbol{\alpha}_i$ 一定是 $\boldsymbol{A}^n\boldsymbol{B}$ 的某一列向量. 由(7)式可知

$$\boldsymbol{A}^n\boldsymbol{B} = -a_1\boldsymbol{A}^{n-1}\boldsymbol{B} - \cdots - a_{n-1}\boldsymbol{A}\boldsymbol{B} - a_n\boldsymbol{B}.$$

而上式右边的每一个列向量都可以由 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r$ 线性表出, 于是 $\boldsymbol{A}^n\boldsymbol{B}$ 的每一个列向量都可以由 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r$ 线性表出. 故 $\boldsymbol{A}\boldsymbol{\alpha}_i$ 也可以由 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r$ 线性表出.

证明 设 $(\boldsymbol{B}, \boldsymbol{A}\boldsymbol{B}, \cdots, \boldsymbol{A}^{n-2}\boldsymbol{B}, \boldsymbol{A}^{n-1}\boldsymbol{B})$ 列向量的极大无关组为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$, 由基扩张定理可将其扩张为 $\mathbb{F}^n$ 的一组基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$. 令 $\boldsymbol{P} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)$, 则 $\boldsymbol{P}$ 为可逆矩阵. 设 $\boldsymbol{A}$ 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

$$f(\boldsymbol{A}) = \boldsymbol{A}^n + a_1\boldsymbol{A}^{n-1} + \cdots + a_{n-1}\boldsymbol{A} + a_n\boldsymbol{I}_n = \boldsymbol{O},$$

从而

$$\boldsymbol{A}^n\boldsymbol{B} = -a_1\boldsymbol{A}^{n-1}\boldsymbol{B} - \cdots - a_{n-1}\boldsymbol{A}\boldsymbol{B} - a_n\boldsymbol{B}. \tag{7}$$

由上式容易验证 $\boldsymbol{A}\boldsymbol{\alpha}_i(1 \le i \le r)$ 都是 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 的线性组合, 于是 $\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{O} & \boldsymbol{A}_{22} \end{pmatrix}$, 即有 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{O} & \boldsymbol{A}_{22} \end{pmatrix}$. 又 $\boldsymbol{B}$ 的列向量都是 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 的线性组合, 于是 $\boldsymbol{B} = \boldsymbol{P}\begin{pmatrix} \boldsymbol{B}_1 \\ \boldsymbol{O} \end{pmatrix}$, 即有 $\boldsymbol{P}^{-1}\boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_1 \\ \boldsymbol{O} \end{pmatrix}$. □

例题 0.8 设 $\boldsymbol{A}$ 是数域 $\mathbb{F}$ 上的 n 阶矩阵, 递归地定义矩阵序列 $\{\boldsymbol{A}_k\}_{k=1}^{\infty}$:

$$\boldsymbol{A}_1 = \boldsymbol{A}, \quad p_k = -\frac{1}{k}\mathrm{tr}(\boldsymbol{A}_k), \quad \boldsymbol{A}_{k+1} = \boldsymbol{A}(\boldsymbol{A}_k + p_k\boldsymbol{I}_n), \quad k = 1, 2, \cdots.$$

求证: $\boldsymbol{A}_{n+1} = \boldsymbol{O}$.

证明 设 $\boldsymbol{A}$ 的全体特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 它们的幂和记为 $s_k = \sum\limits_{i=1}^{n}\lambda_i^k = \mathrm{tr}(\boldsymbol{A}^k)$, 它们的初等对称多项式记为 σ_k, 则 $\boldsymbol{A}$ 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: $p_k = (-1)^k\sigma_k(1 \le k \le n)$. $p_1 = -\mathrm{tr}(\boldsymbol{A}) = -\sigma_1$, 结论成立. 假设小于等于 k 时结论成立, 则 $\boldsymbol{A}_{k+1} = \boldsymbol{A}^{k+1} - \sigma_1\boldsymbol{A}^k + \cdots + (-1)^k\sigma_k\boldsymbol{A}$. 由 Newton 公式可得

$$p_{k+1} = -\frac{1}{k+1}\mathrm{tr}(\boldsymbol{A}_{k+1}) = -\frac{1}{k+1}(s_{k+1} - s_k\sigma_1 + \cdots + (-1)^k s_1\sigma_k) = (-1)^{k+1}\sigma_{k+1},$$

结论得证. 最后, 由Cayley - Hamilton 定理可得

$$\boldsymbol{A}_{n+1} = \boldsymbol{A}^{n+1} - \sigma_1\boldsymbol{A}^n + \cdots + (-1)^n\sigma_n\boldsymbol{A} = f(\boldsymbol{A})\boldsymbol{A} = \boldsymbol{O}.$$

□