高等代数作业答案 (编纂中)

章亦流 A24201011

2025 年 4 月 14 日

目录

4 第四章 线性映射

4.4 线性变换及其矩阵

4.4.1 判断下列变换是否为线性变换:

1. 线性空间 V 中 $\mathscr{A}\xi = \xi + \alpha, \alpha \in V$ 是一固定向量.
2. 线性空间 V 中 $\mathscr{A}\xi = \alpha, \alpha \in V$ 是一固定向量.
3. $M_n(\mathbb{F})$ 中 $\mathscr{A}X = BXC, B, C \in M_n(\mathbb{F})$ 是两个固定矩阵.

证明. 1. 不是; 2. 不是; 3. 是. □

4.4.2 $\mathscr{A},\mathscr{B}$ 为线性变换, 且 $\mathscr{A}\mathscr{B}-\mathscr{B}\mathscr{A}=\mathscr{I}$, 证明 $\mathscr{A}^k\mathscr{B}-\mathscr{B}\mathscr{A}^k=k\mathscr{A}^{k-1},\forall k\in\mathbb{N}$.

证明. 对 k 归纳,$k=1$ 时已证 ($\mathscr{A}^0=\mathscr{I}$). 假设 $<k$ 时等式均成立, 则

$$\begin{aligned}\mathscr{A}^k\mathscr{B}-\mathscr{B}\mathscr{A}^k&=\mathscr{A}\left(\mathscr{A}^{k-1}\mathscr{B}\right)-\mathscr{B}\mathscr{A}^k=\mathscr{A}\left((k-1)\mathscr{A}^{k-2}+\mathscr{B}\mathscr{A}^{k-1}\right)-\mathscr{B}\mathscr{A}^k\\&=(k-1)\mathscr{A}^{k-1}+(\mathscr{A}\mathscr{B}-\mathscr{B}\mathscr{A})\,\mathscr{A}^{k-1}=(k-1)\mathscr{A}^{k-1}+\mathscr{A}^{k-1}=k\mathscr{A}^{k-1}\end{aligned}$$

从而得证. □

注 1. *需要注意的是, 满足 $\mathscr{A}\mathscr{B}-\mathscr{B}\mathscr{A}=\mathscr{I}$ 的线性变换在有限维线性空间中并不存在, 因为 $\operatorname{tr}(AB-BA)=0\neq n=\operatorname{tr}I_n$, 但可以在无限维线性空间中存在, 如在 $C^\infty(\mathbb{R})$ 中取 $\mathscr{A}:f(x)\mapsto f'(x),\mathscr{B}:f(x)\mapsto xf(x)$. 这两个线性变换实际上即量子力学中的位置算符与动量算符 (差一个常数), 而 $\mathscr{A}\mathscr{B}-\mathscr{B}\mathscr{A}=\mathscr{I}$ 即两个算符的对易关系.*

从数学上来说, 我们能在大量广泛的结构中定义 Lie 代数和 Lie 括号, 在向量空间中通常定义 Lie 括号为线性变换的交换子 $[\mathscr{A},\mathscr{B}]=\mathscr{A}\mathscr{B}-\mathscr{B}\mathscr{A}$, 导子 $\operatorname{ad}_{\mathscr{B}}(\mathscr{X})=[\mathscr{B},\mathscr{X}]$. 而该命题即 (通过修改符号): 若 $\operatorname{ad}_{\mathscr{B}}(\mathscr{A})=\mathscr{I}$, 则 $\operatorname{ad}_{\mathscr{B}}(\mathscr{A}^k)=k\mathscr{A}^{k-1}$. 这与微分中的 $\dfrac{\mathrm{d}}{\mathrm{d}x}x^k=kx^{k-1}$ 具有形式上的相似性, 这也暗示了 Lie 代数中的导子与微积分乃至微分流形中的微分有联系.

4.4.3 $\mathscr{A},\mathscr{B}$ 为线性空间 V 中的线性变换, 且 $\mathscr{A}^2=\mathscr{A},\mathscr{B}^2=\mathscr{B}$, 证明:

1. $(\mathscr{A}+\mathscr{B})^2=\mathscr{A}+\mathscr{B}\iff\mathscr{A}\mathscr{B}+\mathscr{B}\mathscr{A}=\mathscr{O}$.
2. $(\mathscr{A}+\mathscr{B})^2=\mathscr{A}+\mathscr{B}\iff\mathscr{A}\mathscr{B}=\mathscr{O},\mathscr{B}\mathscr{A}=\mathscr{O}$.
3. 若 $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$, 则 $(\mathscr{A}+\mathscr{B}-\mathscr{A}\mathscr{B})^2=\mathscr{A}+\mathscr{B}-\mathscr{A}\mathscr{B}$.

证明. 1. $(\mathscr{A}+\mathscr{B})^2=\mathscr{A}^2+\mathscr{A}\mathscr{B}+\mathscr{B}\mathscr{A}+\mathscr{B}^2=\mathscr{A}+\mathscr{A}\mathscr{B}+\mathscr{B}\mathscr{A}+\mathscr{B}=\mathscr{A}+\mathscr{B}\iff\mathscr{A}\mathscr{B}+\mathscr{B}\mathscr{A}=\mathscr{O}$.

2.(证法一) 由 (1) 知 $\Longleftarrow$ 显然. $\Longrightarrow$: 对于幂等变换 $\mathscr{A}$ 的像中元素 $v\in\operatorname{im}\mathscr{A},\exists w\in V,\mathscr{A}w=v$, 因此 $\mathscr{A}v=\mathscr{A}^2w=\mathscr{A}w=v$, 即 $\forall v\in\operatorname{im}\mathscr{A},\mathscr{A}v=v$. 取 $\forall v\in\operatorname{im}\mathscr{A},\mathscr{A}\mathscr{B}v=-\mathscr{B}\mathscr{A}v=-\mathscr{B}v$, 而 $\mathscr{B}v\in\operatorname{im}\mathscr{A}$, 因此 $\mathscr{A}\mathscr{B}v=\mathscr{B}v$, 综上知 $\mathscr{B}v=0$. 故 $\forall x\in V,\mathscr{B}\mathscr{A}x=0,\mathscr{B}\mathscr{A}=\mathscr{O}$, 从而由 (1) 知 $\mathscr{A}\mathscr{B}=\mathscr{O}$.

2.(证法二) 由 $\mathscr{A}\mathscr{B}+\mathscr{B}\mathscr{A}=\mathscr{O}$ 知 $\mathscr{A}(\mathscr{A}\mathscr{B}+\mathscr{B}\mathscr{A})=\mathscr{A}\mathscr{B}+\mathscr{A}\mathscr{B}\mathscr{A}=\mathscr{O}$ 和 $(\mathscr{A}\mathscr{B}+\mathscr{B}\mathscr{A})\mathscr{A}=\mathscr{A}\mathscr{B}\mathscr{A}+\mathscr{B}\mathscr{A}=\mathscr{O}$, 故 $\mathscr{B}\mathscr{A}=-\mathscr{A}\mathscr{B}\mathscr{A}=\mathscr{A}\mathscr{B}$, 从而 $2\mathscr{A}\mathscr{B}=\mathscr{O},\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}=\mathscr{O}$.

3. $(\mathscr{A}+\mathscr{B}-\mathscr{A}\mathscr{B})^2=\mathscr{A}^2\mathscr{B}^2-2\mathscr{A}^2\mathscr{B}+\mathscr{A}^2-2\mathscr{A}\mathscr{B}^2+2\mathscr{A}\mathscr{B}+\mathscr{B}^2=\mathscr{A}+\mathscr{B}-\mathscr{A}\mathscr{B}$. □

4.4.4 取 $A,B,C,D\in M_n(\mathbb{F})$, 令 $\mathscr{A}(X)=AXB+CX+XD$, 证明: 若 $\mathscr{A}$ 是 $M_n(\mathbb{F})$ 中的线性变换, 且 $C=D=O$ 时,$\mathscr{A}$ 可逆 $\iff\det AB\neq0$.

证明. $\Longrightarrow$: $\mathscr{A}$ 可逆即有 $\mathscr{A}\mathscr{A}^{-1}(Y)=A\mathscr{A}^{-1}(Y)B=Y,A\mathscr{A}^{-1}(I_n)B=I_n$, 故 A,B 均满秩, 即 $\det AB\neq0$.

$\Longleftarrow$: AB 可逆则 A,B 可逆, 取 $\mathscr{A}^{-1}(Y)=A^{-1}YB^{-1}$ 即为 $\mathscr{A}$ 的逆. □

4.4.5 $\mathbb{R}^3$ 上的线性变换 $\mathscr{A}$ 将下列 $\alpha_1,\alpha_2,\alpha_3$ 变换为 β_1,β_2,β_3, 求 $\mathscr{A}$ 在 $\mathbb{R}^3$ 中标准基下的矩阵 A:

1. $\alpha_1=(2,3,5),\alpha_2=(0,1,2),\alpha_3=(1,0,0),\beta_1=(1,1,1),\beta_2=(1,1,-1),\beta_3=(2,1,2)$;
2. $\alpha_1=(2,0,3),\alpha_2=(4,1,5),\alpha_3=(3,1,2),\beta_1=(1,2,-1),\beta_2=(4,5,-2),\beta_3=(1,1,0)$.

证明. 设 $\mathscr{A}$ 在标准基 e_1,e_2,e_3 下的矩阵为 A, 则有 $\mathscr{A}(e_1,e_2,e_3)=(e_1,e_2,e_3)A$, 从而有

$$(\beta_1,\beta_2,\beta_3)=\mathscr{A}(\alpha_1,\alpha_2,\alpha_3)=\mathscr{A}(e_1,e_2,e_3)(\alpha_1^\mathsf{T},\alpha_2^\mathsf{T},\alpha_3^\mathsf{T})=(e_1,e_2,e_3)A(\alpha_1^\mathsf{T},\alpha_2^\mathsf{T},\alpha_3^\mathsf{T})$$

又因 $(\beta_1,\beta_2,\beta_3)=(e_1,e_2,e_3)(\beta_1^\mathsf{T},\beta_2^\mathsf{T},\beta_3^\mathsf{T})$, 故 $A=(\beta_1^\mathsf{T},\beta_2^\mathsf{T},\beta_3^\mathsf{T})(\alpha_1^\mathsf{T},\alpha_2^\mathsf{T},\alpha_3^\mathsf{T})^{-1}$.

1.
$$A=\begin{pmatrix}1&1&2\\1&1&1\\1&-1&2\end{pmatrix}\begin{pmatrix}2&0&1\\3&1&0\\5&2&0\end{pmatrix}^{-1}=\begin{pmatrix}1&1&2\\1&1&1\\1&-1&2\end{pmatrix}\begin{pmatrix}0&2&-1\\0&-5&3\\1&-4&2\end{pmatrix}=\begin{pmatrix}2&-11&6\\1&-7&4\\2&-1&0\end{pmatrix}$$

2.
$$A=\begin{pmatrix}1&4&1\\2&5&1\\-1&-2&0\end{pmatrix}\begin{pmatrix}2&4&3\\0&1&1\\3&5&2\end{pmatrix}^{-1}=\begin{pmatrix}1&4&1\\2&5&1\\-1&-2&0\end{pmatrix}\begin{pmatrix}1&-7/3&-1/3\\-1&5/3&2/3\\1&-2/3&-2/3\end{pmatrix}=\begin{pmatrix}-2&11/3&5/3\\-2&3&2\\1&-1&-1\end{pmatrix}$$

□

4.4.6 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 下的矩阵为

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

求其在 $\mathscr{A}$ 在下列基下的矩阵:

1. $\alpha_2,\alpha_4,\alpha_3,\alpha_1$;
2. $\alpha_1,\alpha_1+\alpha_2,\alpha_1+\alpha_2+\alpha_3,\alpha_1+\alpha_2+\alpha_3+\alpha_4$.

证明. 基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 到如上两组基的过渡矩阵分别为

$$P_1=\begin{pmatrix}0&0&0&1\\1&0&0&0\\0&0&1&0\\0&1&0&0\end{pmatrix},P_2=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{pmatrix},P_1^{-1}=\begin{pmatrix}0&1&0&0\\0&0&0&1\\0&0&1&0\\1&0&0&0\end{pmatrix},P_2^{-1}=\begin{pmatrix}1&-1&0&0\\0&1&-1&0\\0&0&1&-1\\0&0&0&1\end{pmatrix}$$

故 $\mathscr{A}$ 在两组基下的矩阵分别为

$$P_1^{-1}AP_1=\begin{pmatrix}0&2&-1&3\\2&3&1&1\\5&1&3&2\\2&1&0&1\end{pmatrix},P_2^{-1}AP_2=\begin{pmatrix}-2&0&1&0\\1&-4&-8&-7\\1&4&6&4\\1&3&4&7\end{pmatrix}.$$

□

4.4.7 V 为 $\mathbb{F}$ 上的 n 维线性空间, 证明: 与 V 上全体线性变换可交换的线性变换有且仅有数乘变换 $c\mathscr{I},c\in\mathbb{F}$.

证明. 在 V 中任取一基, 设线性变换 $\mathscr{A}$ 在该基下的矩阵为 A. 若 $\mathscr{A}$ 与 V 的任意线性变换可交换, 即 A 与任意 n 阶方阵可交换. 取 $D=\mathrm{diag}(1,2,\dots,n)$, 由 $AD=DA$ 知乘积的第 (i,j) 元为 $ja_{ij}=ia_{ij}$, 故 $a_{ij}=0(i\neq j),A=\mathrm{diag}(a_1,a_2,\dots,a_n)$. 再对任意 $i,j=1,2,\dots,n(i\neq j)$ 取矩阵 $P_{ij}=I_n+E_{ij}$. 由 $AP_{ij}=P_{ij}A$ 知 $a_iE_{ij}=a_jE_{ij}$, 故 $a_i=a_j$, 再由 i,j 任意性知 $a_1=\cdots=a_n=c,A=cI_n,c\in\mathbb{F}$. □

4.4.8 $\mathscr{A}$ 为 $\mathbb{F}$ 上 n 维线性空间 V 的一个线性变换, 证明: 若 $\mathscr{A}$ 在任意基下的矩阵相同, 则 $\mathscr{A}$ 为数乘变换.

证明. 在 V 中任取一基, 设线性变换 $\mathscr{A}$ 在该基下的矩阵为 A. 若 $\mathscr{A}$ 在任意基下的矩阵相同, 即等价于对任意可逆矩阵 P 有 $P^{-1}AP=A$, 即 $AP=PA$. 同上证法即得证. □

4.4.9 $\mathbb{F}$ 上的线性空间 V 有基 $\alpha_1,\alpha_2,\dots,\alpha_n$, 由 $\mathscr{A}\alpha_i=\alpha_{i+1}(i=1,2,\dots,n-1),\mathscr{A}\alpha_n=0$ 定义了线性变换 $\mathscr{A}$.

1. 求 $\mathscr{A}$ 在该基下的矩阵.
2. 证明 $\mathscr{A}^n=\mathscr{O},\mathscr{A}^{n-1}\neq\mathscr{O}$.
3. 若 V 上线性变换 $\mathscr{B}$ 同样满足 $\mathscr{B}^n=\mathscr{O},\mathscr{B}^{n-1}\neq\mathscr{O}$, 则存在 V 中的基使 $\mathscr{B}$ 在该基下的矩阵与 (1) 中 $\mathscr{A}$ 的矩阵相同.
4. 证明:$M,N\in M_n(\mathbb{F})$ 若满足 $M^n=N^n=O,M^{n-1}\neq O,N^{n-1}\neq O$, 则 M 与 N 相似.

证明. 容易看出

$$A=\begin{pmatrix}0&&&&\\1&0&&&\\&1&0&&\\&&\ddots&\ddots&\\&&&1&0\end{pmatrix},A^2=\begin{pmatrix}0&&&&&\\0&0&&&&\\1&0&0&&&\\&\ddots&\ddots&\ddots&&\\&&&1&0&0\end{pmatrix},\cdots,A^{n-1}=\begin{pmatrix}0_{n-1\times1}&O_{n-1}\\1&0_{1\times n-1}\end{pmatrix}\neq O,A^n=O$$

而对于 V 上线性变换 $\mathscr{B}$, 若有 $\mathscr{B}^n=\mathscr{O},\mathscr{B}^{n-1}\neq\mathscr{O}$, 取 $\alpha\in V$ 使 $\mathscr{B}^{n-1}\alpha\neq0$, 下证 $\alpha,\mathscr{B}\alpha,\dots,\mathscr{B}^{n-1}\alpha$ 是 V 的一组基. 设有 $k_1,\dots,k_n\in\mathbb{F}$ 使得 $\alpha'=\sum\limits_{i=1}^{n}k_i\mathscr{B}^{i-1}\alpha=0$, 则 $\mathscr{B}^{n-1}\alpha'=k_1\mathscr{B}^{n-1}\alpha=0,k_1=0$, 故 $\mathscr{B}^{n-2}\alpha'=k_2\mathscr{B}^{n-1}\alpha=0,k_2=0$, 以此类推. 从而 $k_1=\cdots=k_n=0$, 即 $\alpha,\mathscr{B}\alpha,\dots,\mathscr{B}^{n-1}\alpha$ 线性无关, 是一组基, 且有 $\mathscr{B}(\mathscr{B}^i\alpha)=\mathscr{B}^{i+1}\alpha(i=0,1,\dots,n-1),\mathscr{B}(\mathscr{B}^{n-1})\alpha=0$, 从而由 (1) 知 $\mathscr{B}$ 在该基下的矩阵也为 A.

最后, 由上知 M,N 均相似于 A, 即 M 与 N 相似. □

4.5 不变子空间

4.5.1 $\mathbb{F}^4$ 中的线性变换 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 下的矩阵为

$$A=\begin{pmatrix}1&0&2&-1\\0&1&4&-2\\2&-1&0&1\\2&-1&-1&2\end{pmatrix}$$

令 $W=\operatorname{span}(\alpha_1+2\alpha_2,\alpha_2+\alpha_3+2\alpha_4)$, 证明 W 是 $\mathscr{A}$-子空间.

证明. 由于

$$\begin{aligned}\mathscr{A}(\alpha_1+2\alpha_2)&=(\alpha_1+2\alpha_3+2\alpha_4)+2(\alpha_2-\alpha_3-\alpha_4)=\alpha_1+2\alpha_2\in W,\\\mathscr{A}(\alpha_2+\alpha_3+2\alpha_4)&=(\alpha_2-\alpha_3-\alpha_4)+(2\alpha_1+4\alpha_2-\alpha_4)+2(\alpha_1-2\alpha_2+\alpha_3+2\alpha_4)\\&=\alpha_2+\alpha_3+2\alpha_4\in W\end{aligned}$$

故 W 是 $\mathscr{A}$-子空间. □

4.5.2 线性空间 V 有子空间 $W=\operatorname{span}(\alpha_1,\alpha_2,\dots,\alpha_k)$, 证明 W 是 $\mathscr{A}$-子空间 $\iff\mathscr{A}\alpha_i\in W(i=1,2,\dots,k)$.

证明. $\Longrightarrow$ 显然, $\Longleftarrow$: 由于 W 中任意向量 α 可看作 $\alpha_1,\alpha_2,\dots,\alpha_k$ 的线性组合 $\alpha=\sum\limits_{i=1}^{k}k_i\alpha_i$, 则 $\mathscr{A}\alpha=\sum\limits_{i=1}^{k}k_i\mathscr{A}\alpha_i\in W$, 故 W 是 $\mathscr{A}$ 的不变子空间. □

4.5.3 $\mathscr{A},\mathscr{B}$ 为线性空间 V 上的线性变换,U 为 $\mathscr{A},\mathscr{B}$ 的不变子空间, 证明 U 是 $\mathscr{A}+\mathscr{B},\mathscr{A}\mathscr{B}$ 的不变子空间. 若 $\mathscr{A}$ 可逆, 则 U 也是 $\mathscr{A}^{-1}$ 的不变子空间.

证明. 由于 $\forall\alpha\in U,\mathscr{A}\alpha,\mathscr{B}\alpha\in U$, 故 $\mathscr{A}\alpha+\mathscr{B}\alpha\in U,\mathscr{A}\mathscr{B}\alpha\in U$, 即 U 是 $\mathscr{A}+\mathscr{B},\mathscr{A}\mathscr{B}$ 的不变子空间. 而 $\mathscr{A}$ 可逆时,$\dim\mathscr{A}U\le\dim U=\dim(\mathscr{A}^{-1}\mathscr{A}U)\le\dim\mathscr{A}U$, 由 $\mathscr{A}U\subset U$ 知 $\mathscr{A}U=U$, 故 $\mathscr{A}^{-1}U=\mathscr{A}^{-1}\mathscr{A}U=U$, 从而 U 是 $\mathscr{A}^{-1}$ 的不变子空间. □

4.5.4 n 维线性空间 V 上的线性变换 $\mathscr{A}$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵为

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

1. 若 α_n 在 $\mathscr{A}$ 的不变子空间 U 中, 则 $U=V$;
2. α_1 属于 $\mathscr{A}$ 的任意非零不变子空间中.
3. V 不能被分解为两个非平凡 $\mathscr{A}$-子空间的直和.

证明. 1. 由于 $\mathscr{A}\alpha_1=\lambda_0\alpha_1,\mathscr{A}\alpha_k=\alpha_{k-1}+\lambda_0\alpha_k(k=2,\dots,n)$, 因此若 $\alpha_n\in U$ 则 $\alpha_{n-1}=\mathscr{A}\alpha_n-\lambda_0\alpha_n\in U,\dots,\alpha_1=\mathscr{A}\alpha_2-\lambda_0\alpha_2\in U$, 即 $\alpha_1,\dots,\alpha_n\in U,V=\operatorname{span}(\alpha_1,\dots,\alpha_n)\subset U$, 从而 $U=V$.

2. 任取 $\mathscr{A}$ 的非零不变子空间 U, 易知 U 也在 $\mathscr{B}=\mathscr{A}-\lambda_0\mathscr{I}$ 下不变, 且 $\mathscr{B}\alpha_1=0,\mathscr{B}\alpha_k=\alpha_{k-1}(k=2,\dots,n)$. 因此任取 U 中非零向量 $\alpha=\sum_{i=1}^{n}k_i\alpha_i=\sum_{i=1}^{m}k_i\alpha_i$, 其中 $k_{m+1}=\cdots=k_n=0,k_m\ne0$, 从而有

$$\mathscr{B}\alpha=\sum_{i=2}^{m}k_i\alpha_{i-1}\in U,\mathscr{B}^2\alpha=\sum_{i=3}^{m}k_i\alpha_{i-2}\in U,\dots,\mathscr{B}^{m-1}\alpha=k_m\alpha_1\in U$$

故 $\alpha_1\in U$.

3. 由于 α_1 在任意非平凡 $\mathscr{A}$-子空间中, 故非平凡 $\mathscr{A}$-子空间之间必有交, 从而无法作直和. □

4.5.5 $\mathscr{A}$ 是 $\mathbb{R}^2$ 上的线性变换, 其在基 α_1,α_2 下的矩阵为 $\begin{pmatrix}-1 & 2\\ 3 & -6\end{pmatrix}$. 求 $\mathscr{A}$ 的所有不变子空间.

证明. $\mathscr{A}$ 有平凡不变子空间 $\mathbb{R}^2$ 和零空间,$\mathscr{A}$ 的非平凡子空间 U 仅能 $\dim U=1$, 从而 $\forall\alpha\in U,\mathscr{A}\alpha\in U$. 设 $\alpha=k_1\alpha_1+k_2\alpha_2$, 即有方程 $k_1(-\alpha_1+3\alpha_2)+k_2(2\alpha_1-6\alpha_2)=(-k_1+2k_2)\alpha_1+(3k_1-6k_2)\alpha_2=\lambda(k_1\alpha_1+k_2\alpha_2)$. 解得 $k_1/k_2=3$ 或 $-1/2$, 从而可取 $U_1=\operatorname{span}(3\alpha_1+\alpha_2)$ 和 $U_2=\operatorname{span}(\alpha_1-2\alpha_2)$, 它们和两个平凡不变子空间构成 $\mathscr{A}$ 的所有不变子空间. □

5 第五章 多项式

5.1 一元多项式

5.1.1 $f,g\in\mathbb{F}[x]$, 证明 $fg=0\iff f$ 和 g 中至少一个是 0.

证明一. $\Longleftarrow$ 显然. $\Longrightarrow$: 若 f,g 均非零, 则两者的首项系数之积非零, 从而 $fg\neq 0$. □

证明二. $\Longrightarrow$: 按逐项系数递推, 记

$$f(x)=\sum_{i=0}^{m}a_ix^i,g(x)=\sum_{j=0}^{n}b_jx^j,f(x)g(x)=\sum_{t=0}^{m+n}\left(\sum_{i+j=t}a_ib_j\right)x^t=0$$

从而 $c_t=\sum\limits_{i+j=t}a_ib_j=0,\forall t=0,\dots,m+n$. 若 $f=0$ 则命题得证; 若 $f\neq 0$, 则 $a_m\neq 0$, 而 $c_{m+n}=a_mb_n=0,b_n=0$.

下证明 $b_{n-r}=0,r=0,\dots,n$. 对 r 归纳, $r=0$ 时已证. 若 $<r$ 的情形已证, 即

$$b_{n-0}=b_{n-1}=\cdots=b_{n-r+1}=0$$

则

$$c_{m+n-r}=a_mb_{n-r}+a_{m-1}b_{n-r+1}+\cdots+a_{m-r}b_n=a_mb_{n-r}=0$$

从而 $b_{n-r}=0$, 得证. □

5.1.2 $f,g,h\in\mathbb{F}[x]$, 若 $f\neq 0$, 则 $fg=fh\iff g=h$.

证明. $fg=fh\iff f(g-h)=0$, 而 $f\neq 0$, 由上题知 $g-h=0,g=h$. □

5.1.3 对于 $f\in\mathbb{R}[x],f\neq 0$ 满足 $f(x^2)=f^2(x)$, 求多项式 $f(x)$.

证明一. 记 $f(x)=\sum\limits_{i=0}^{n}a_ix^i,a_i\in\mathbb{R},a_n\neq 0$. 取 $m=\max\{k\mid a_k\neq 0,k=0,\dots,n-1\}$, 即除 a_nx^n 外最高非零项的次数, 则

$$f(x^2)=\sum_{i=0}^{n}a_ix^{2i}=a_nx^{2n}+a_mx^{2m}+\dots,f^2(x)=\sum_{t=0}^{2n}\left(\sum_{i+j=t}a_ia_j\right)x^t=a_n^2x^{2n}+2a_na_mx^{n+m}+\dots$$

比较系数,$a_n=a_n^2,a_n=1$. 而 $n+m>2m$, 故 x^{n+m} 项系数 $2a_na_m=0,a_m=0$. 这与 m 定义矛盾, 故 m 不存在, 即 $a_0=\cdots=a_{n-1}=0,f(x)=x^n$. □

证明二. 同上记号且易证 $a_n=1$. 对 $f(x^2),f^2(x)$ 展开有:

$$f(x^2)=\sum_{i=0}^{n}a_ix^{2i},f^2(x)=\sum_{t=0}^{2n}\left(\sum_{i+j=t}a_ia_j\right)x^t$$

逐项比较系数, 可得:

$$a_k=\sum_{i+j=2k}a_ia_j,\qquad 0=\sum_{i+j=2k+1}a_ia_j,\qquad \forall k=0,\dots,n$$

从而 $0=2a_na_{n-1},a_{n-1}=0$. 下证 $a_{n-r}=0,r=1,\dots,n$. 对 r 归纳, $r=1$ 时已证. 假设 $<r$ 的情形已证, 即 $a_{n-1}=\cdots=a_{n-r+1}=0$ 时: 若 r 为偶数, 则

$$0=a_{n-r/2}=\sum_{i+j=2n-r}a_ia_j=2a_na_{n-r}$$

若 r 为奇数, 则取 $k=n-(r+1)/2$,

$$0=\sum_{i+j=2n-r}a_ia_j=2a_na_{n-r}$$

可知总有 $a_{n-r}=0$, 从而得证, 即 $f(x)=x^n$. $\square$

5.1.4 $f,g,h\in\mathbb{R}[x]$, 证明若 $f^2(x)=xg^2(x)+xh^2(x)$, 则 $f=g=h=0$.

证明一. 若 $f\neq 0$ 则 $\deg f^2=2\deg f$ 为偶数, 且此时 $g^2+h^2\neq 0$. 由于 g^2 与 h^2 的首项系数均为正数, 故两者和也为正数, 故 $\deg(g^2+h^2)=\max(\deg g^2,\deg h^2)$, 从而有

$$2\deg f=\deg f^2=\deg\big(xg^2(x)+xh^2(x)\big)=2\max(\deg g,\deg h)+1$$

左端为偶数, 右端为奇数, 矛盾, 从而 $f=0,g^2+h^2=0,g=h=0$. $\square$

证明二. 若 g,h 中至少有一个非零, 取 $g\neq 0$, 则 $\exists c\in\mathbb{R},g(c)\neq 0$, 故 $g^2(c)+h^2(c)>0,g^2+h^2\neq 0$. 而 $\deg f^2$ 为偶数,$\deg\big(xg^2(x)+xh^2(x)\big)$ 为奇数, 矛盾. 故 $g=h=0,f=0$. $\square$

5.1.5 在 $\mathbb{C}[x]$ 中找一组不全为 0 的多项式 f,g,h 使得 $f^2(x)=xg^2(x)+xh^2(x)$.

证明. $f(x)=0,g(x)=\mathrm{i},h(x)=1$. $\square$

通解. 由于 $x\mid f^2(x)$, 则 $x\mid f(x)$. 记 $f(x)=xq(x)$, 有

$$xq^2(x)=g^2(x)+h^2(x)=(g(x)+\mathrm{i}h(x))(g(x)-\mathrm{i}h(x))$$

不失一般性地认为 g,h 互素, 因为上式等价于

$$x\left(\frac{q(x)}{(g(x),h(x))}\right)^2=\left(\frac{g(x)}{(g(x),h(x))}\right)^2+\left(\frac{h(x)}{(g(x),h(x))}\right)^2$$

另一方面,$x\mid(g(x)+\mathrm{i}h(x))(g(x)-\mathrm{i}h(x))$, 不失一般性地认为 $x\mid g(x)+\mathrm{i}h(x)$.

将 $q(x)$ 分解为不可约多项式的乘积, 即 $q=p_1p_2\dots p_m$, 则

$$g(x)+\mathrm{i}h(x)=xp_1^2(x)\dots p_s^2(x)p_{s+1}(x)\dots p_t(x),g(x)-\mathrm{i}h(x)=p_{s+1}(x)\dots p_t(x)p_{t+1}^2(x)\dots p_m^2(x)$$

记

$$a(x)=p_1(x)\dots p_s(x),b(x)=p_{t+1}(x)\dots p_m(x),d(x)=p_{s+1}(x)\dots p_t(x)$$

则 $q=abd,g+\mathrm{i}h=xa^2d,g-\mathrm{i}h=db^2,(g+\mathrm{i}h,g-\mathrm{i}h)=d(a,b)^2$. 而 $(g+\mathrm{i}h,g-\mathrm{i}h)=(g+\mathrm{i}h,2g)=(g,h)=1$, 因此 $d=(a,b)=1,g+\mathrm{i}h=xa^2,g-\mathrm{i}h=b^2$, 解得:

$$f=xq=xab,g=\frac{xa^2+b^2}{2},h=\frac{xa^2-b^2}{2\mathrm{i}}$$

最后回代 g,h 不互素的情况, 得到通解: 对于 $\forall a,b\in\mathbb{C}[x]$, 上式为通解. $\square$

5.2 整除

5.2.1 求下列 $f(x)$ 除以 $g(x)$ 的商式 $q(x)$ 与余式 $r(x)$:

1. $f(x)=5x^4+3x^3+2x^2+x-1,g(x)=x^2+2x-2$;
2. $f(x)=6x+3x^4-4x^3,g(x)=x+2$.

证明.

1. $q(x)=5x^2-7x+26,r(x)=-65x+51$.
2. $q(x)=3x^3-10x^2+20x-34,r(x)=68$.

$\square$

5.2.2 求 $f(x)$ 按 $x-c$ 幂的展开式, 即写成 $f(x)=\sum_{k=0}^{n} a_k(x-c)^k$ 的形式:

1. $f(x)=x^5, c=1$;
2. $f(x)=x^3-10x^2+13, c=-2$.

证明. 1. $(x-1)^5+5(x-1)^4+10(x-1)^3+10(x-1)^2+5(x-1)+1$.

2. $(x+2)^3-16(x+2)^2+52(x+2)-35$.

□

5.2.3 问参数 m,n,p 满足什么条件时有

1. $x^2-2x+1 \mid x^4-5x^3+11x^2+mx+n$;
2. $x^2-2mx+2 \mid x^4+3x^2+mx+n$;
3. $x^2+m-1 \mid x^3+nx+p$;
4. $x^2+mx+1 \mid x^4+nx^2+p$.

证明. 1. 要求除法余式 $r(x)=(m+11)x+(n-4)=0$, 即 $m=-11, n=4$.

2. 要求除法余式 $r(x)=(8m^3-m)x-8m^2+n-2=0$, 解得 $m=0, n=2$ 或 $m=\pm\dfrac{\sqrt{2}}{4}, n=3$.
3. 要求除法余式 $r(x)=(1-m+n)x+p=0$, 即 $m=n+1, p=0$.
4. 要求除法余式 $r(x)=(-m^3-mn+2m)x-m^2-n+p+1=0$, 解得 $m=0, n=p+1$ 或 $m^2+n=2, p=1$.

□

5.2.4 求 $u(x), v(x)$ 使得 $uf+vg=(f,g)$.

1. $f(x)=x^4+3x^3-x^2-4x-3, g(x)=3x^3+10x^2+2x-3$;
2. $f(x)=x^4-10x^2+1, g(x)=x^4-4\sqrt{2}x^3+6x^2+4\sqrt{2}x+1$;
3. $f(x)=x^4-x^3-4x^2+4x+1, g(x)=x^2-x-1$.

证明. 1. $u(x)=\dfrac{3}{5}x-1, v(x)=-\dfrac{1}{5}x^2+\dfrac{2}{5}x$.

2. $u(x)=-\dfrac{\sqrt{2}}{8}x+\dfrac{1}{2}, v(x)=\dfrac{\sqrt{2}}{8}x+\dfrac{1}{2}$.
3. $u(x)=-x-1, v(x)=x^3+x^2-3x-2$.

(答案均不唯一.)

□

5.2.5 设 $f(x)=x^3+(t+1)x^2+2x+2u$ 与 $g(x)=x^3+tx^2+u$ 的最大公因式为二次多项式, 求 t,u.

证明. 考虑带余除法 $f=qg+r$, 比较次数与系数可知 $q(x)=1$, 故 $r(x)=f(x)-g(x)=x^2+2x+u$. 继续辗转相除得到 $g=q_1r+r_1$, 其中 $\deg r_1<\deg r=2$, 而 $(f,g)\mid r_1$, 因此 $r_1=0, g=q_1r$. 比较系数知 q_1 为首项系数为 1 的一次多项式 $(x-a)$, 因此有

$$g(x)=x^3+tx^2+u=(x-a)(x^2+2x+u)=x^3+(2-a)x^2+(u-2a)x-au$$

比较系数可得

$$t=2-a, \quad 0=u-2a, \quad u=-au$$

解得 $t=2, u=0, a=0$ 或 $t=3, u=-2, a=-1$.

□

5.2.6 对于多项式 f,g,d, 若 $d\mid f, d\mid g$ 且存在多项式 u,v 使得 $d=uf+vg$, 证明 $d=(f,g)$.

证明. 由 $d\mid f, d\mid g$ 知 $d\mid (f,g)$, 而 $(f,g)\mid uf+vg=d$, 因此 d 与 (f,g) 间差一个非零常数, 即 d 是 f,g 的一个最大公因数.

□

5.2.7 设 $f,g\in\mathbb{F}[x]$, 证明:

1. 若 $a,b,c,d\in\mathbb{F}$ 满足 $ad-bc\neq 0$, 则 $(af+bg,cf+dg)=(f,g)$;
2. $(f^2,g^2)=(f,g)^2$;
3. $(f,f+g)=1 \iff (f,g)=1$.

证明. 首先证明引理: 对于任意多项式 $q\in\mathbb{F}[x],(f,g)=(f+qg,g)$. 证: 由于 (f,g) 整除 $f+qg$ 和 g, 因此 $(f,g)\mid(f+qg,g)$, 同理 $(f+qg,g)\mid(f+qg-qg,g)=(f,g)$, 从而两者相等.

1.
$$(af+bg,cf+dg)=\left(af+bg,cf+dg-\frac{c}{a}(af+bg)\right)=\left(af+bg,\frac{ad-bc}{a}g\right)=(f,g)$$

2. 记 $d=(f,g)$, 有 $f=df_1,g=dg_1,(f_1,g_1)=1,(f^2,g^2)=d^2(f_1^2,g_1^2)$. 而 $(f_1,g_1)=1 \iff (f_1^2,g_1^2)=1$(书上推论 5.2.12, 或由 Bézout 定理), 从而得证.

3. 由 1 或引理显然. □

5.2.8 $f,g\in\mathbb{F}[x]$ 不全为 0, 且 $uf+vg=(f,g)$, 证明 $(u,v)=1$.

证明. 记 $f=(f,g)f_1,g=(f,g)g_1$, 其中 $(f_1,g_1)=1$. 从而有 $(f,g)=uf+vg=(f,g)(uf_1+vg_1)$, 因此 $uf_1+vg_1=1$, 这等价于 $(u,v)=1$. □

5.2.9 设 $f_1,\dots,f_m,g_1,\dots,g_n\in\mathbb{F}[x]$ 且 $(f_i,g_j)=1(\forall i\in[m],j\in[n])$, 证明 $(f_1\dots f_m,g_1\dots,g_n)=1$.

证明. 首先证明 $n=1$ 的情形, 即 $\forall i=1,\dots,m,(f_i,g)=1$ 则有 $(f_1\dots f_m,g)=1$. 对 m 归纳,$m=1$ 时已证, 下设 $<m$ 的情形已得证, 而 $(f_1\dots f_{m-1},g)=(f_m,g)=1 \iff (f_1\dots f_m,g)=1$(书上推论 5.2.12), 从而得证.

再对原命题考虑, 记 $f=f_1\dots f_m$, 由上知 $(f,g_1)=\cdots=(f,g_n)=1$, 从而又有 $(f,g_1\dots g_n)=1$. □

5.2.10 证明定理 5.2.16

定理 5.2.16 设 $f_1,\dots,f_k\in\mathbb{F}[x]$ 不全为 0, 则 $(f_1,\dots,f_k)$ 唯一存在, 且

$$(f_1,\dots,f_k)=((f_1,\dots,f_{k-1}),f_k)$$

从而 $\exists u_i\in\mathbb{F}[x],i\in[k]$ 使得

$$(f_1,\dots,f_k)=\sum_{i=1}^{k}u_if_i$$

证明. 对 k 归纳,$k=2$ 时已得证. 下设 $k\geq 3,<k$ 的情形已证. 设 $d_1=(f_1,\dots,f_{k-1})$, 由归纳假设知其唯一确定, 且有 $v_1f_1+\cdots+v_{k-1}f_{k-1}=d_1$.

首先证明 $d=(d_1,f_k)$ 为 $f_1,\dots,f_k$ 的最大公因式, 从而证明存在性. 显然 $d\mid d_1\mid f_i(i=1,\dots,k-1)$ 且 $d\mid f_k$. 又对于 $f_1,\dots,f_k$ 的任意公因式 g, 均有 (由归纳假设)

$$g\mid v_1f_1+\cdots+v_{k-1}f_{k-1}=d_1,g\mid f_k$$

从而 $g\mid(d_1,f_k)=d$, 即 d 为最大公因式.

再证明唯一性: 若有多项式 d,d' 均为 $f_1,\dots,f_k$ 的最大公因式, 则 $d'\mid d,d\mid d'$, 从而相同 (差一个非零常数而首项系数均为 1).

最后, 由归纳假设有 $v_1f_1+\cdots+v_{k-1}f_{k-1}=d_1$, 又有 $ud_1+vf_k=d$, 从而

$$u(v_1f_1+\cdots+v_{k-1}f_{k-1})+vf_k=uv_1f_1+\cdots+uv_{k-1}f_{k-1}+vf_k=d$$

综上得证. □

5.2.11 称多项式 $m(x)$ 为多项式 $f(x),g(x)$ 的最小公倍式, 若 $f \mid m, g \mid m$ 且 f,g 的任意公倍式是 m 的倍式. 记 $m=[f,g]$, 证明若 f,g 首项系数为 1, 则 $[f,g]=\dfrac{fg}{(f,g)}$.

证明. 记 $d=(f,g), m=fg/d, f=df_1, g=dg_1, (f_1,g_1)=1$. 从而 $m=f_1g=fg_1$, 故 $f \mid m, g \mid m$.

再设 f,g 的任意公倍式 $h=h_1f=h_2g$, 有 $h=dh_1f_1=dh_2g_1$, 从而 $h_1f_1=h_2g_1$. 而 $(f_1,g_1)=1$, 因此 $f_1 \mid h_2, m=df_1g_1 \mid dh_2g_1=h$. 综上,$m$ 满足最小公倍式的所有条件, 即 $m=[f,g]$. □

思考题 1 对于 $f(x)=\sum\limits_{i=0}^{m}a_ix^i, g(x)=\sum\limits_{i=0}^{n}b_ix^i, h(x)=\sum\limits_{i=0}^{m-n}c_ix^i$, 若有 $f(x)=g(x)h(x)$, 显式表达出 c_i.

证明. 考虑 $g(x)$ 的最低非零次数 $r=\min\{i|b_i\neq 0, i=0,1,\ldots,n\}$, 则 $g(x)=\sum\limits_{i=r}^{n}b_ix^i$. 又由 $a_k=\sum\limits_{i+j=k}b_ic_j$ 有:

$$a_{r+k}=\sum_{i+j=r+k}b_ic_j=b_rc_k+\sum_{i=0}^{k-1}b_{r+k-i}c_i, (k=0,\ldots,m-r)$$

因此有

$$c_k=\frac{1}{b_r}\left(a_{r+k}-\sum_{i=0}^{k-1}b_{r+k-i}c_i\right)$$

其在 $r=0$, 即 $b_0\neq 0$ 时化为

$$c_k=\frac{1}{b_0}\left(a_k-\sum_{i=0}^{k-1}b_{k-i}c_i\right)$$

□

5.3 因式分解定理

5.3.1 x^2+1 在 $\mathbb{Q}$ 上不可约.

证明. 由于 x^2+1 在 $\mathbb{R}$ 上的唯一分解式为 $(x-\mathrm{i})(x+\mathrm{i})$, 故其不能被分解为 $\mathbb{Q}[x]$ 中的一次多项式之积, 故在 $\mathbb{Q}$ 上不可约. □

5.3.2 判别下列多项式是否有重因式:

1. $f(x)=x^4+x^3+2x^2+x+1$,
2. $f(x)=x^6-3x^5+6x^3-3x^2-3x+2$.

证明. 1. 辗转相除可得 $(f,f')=1$ 从而无重因式. 但辗转相除太过麻烦, 有其他方法:

- 注意到 $(f,g)=(f+qg,g), \forall q\in\mathbb{F}[x]$, 故对第一小问有

$$\begin{aligned}(f,f')&=(x^4+x^3+2x^2+x+1, 4x^3+3x^2+4x+1)=\left(x^3+4x^2+3x+4, 4x^3+3x^2+4x+1\right)\\&=\left(x^3+4x^2+3x+4, 13x^2+8x+15\right)=\left(11x^2+6x+13, 13x^2+8x+15\right)\\&=\left(11x^2+6x+13, 10x-4\right)=(4x+5, 10x-4)=1\end{aligned}$$

  但该方法对第二小问太麻烦.

- 由于该题为四次多项式, 故可设

$$x^4+x^3+2x^2+x+1=(x^2+ax+b)(x^2+cx+d)$$

  展开后比较系数可得

$$a+c=1,\quad ac+b+d=2,\quad ad+bc=1,\quad bd=1$$

尝试带入 $b=d=\pm 1$ 发现 $b=d=1, a=0, c=1$ 时方程成立, 即

$$f(x)=(x^2+1)(x^2+x+1)$$

从而无重因式.

- 注意到方程系数 $(1,1,2,1,1)$ 是对称的, 因此可令 $z=x+1/x$ 换元, 即

$$\begin{aligned} f(x) &= x^2\left(x^2+x+2+\frac{1}{x}+\frac{1}{x^2}\right)=x^2\left(\left(x+\frac{1}{x}\right)^2+\left(x+\frac{1}{x}\right)\right) \\ &= x^2\left(x+\frac{1}{x}+1\right)\left(x+\frac{1}{x}\right)=(x^2+1)(x^2+x+1) \end{aligned}$$

从而无重因式.

2. 辗转相除可得 $(f,f')=x^3-x^2-x+1$ 从而有重因式. 也可直接试根: 注意到 $f(x)$ 的有理根 $x_0=r/s$ 总有 $r\mid 2, s\mid 1$, 故 x_0 仅可能为 $\pm 1, \pm 2$, 故带入验算发现 $1,-1,2$ 均为根, 相除得到

$$\frac{f(x)}{(x-2)(x-1)(x+1)}=x^3-x^2-x+1=(x-1)^2(x+1)$$

从而 $f(x)=(x-1)^3(x+1)^2(x-2)$, 其有重根.

□

5.3.3 求 A,B 使得 $(x-1)^2\mid Ax^4+Bx^2+1$.

证明一. 设 $f(x)=Ax^4+Bx^2+1$, 由题知 $(x-1)\mid f'(x)=4Ax^3+2Bx$, 从而 $f(1)=f'(1)=0$, 即 $A+B+1=4A+2B=0$, 解得 $A=1, B=-2$. □

证明二. 设 $f(x)=Ax^4+Bx^2+1$, 注意到 $(x-1)\mid(f,f')=(Ax^4+Bx^2+1, 4Ax^3+2Bx)=(Bx^2/2+1, 4Ax^3+2Bx)$, 而 $(x-1)\mid \dfrac{B}{2}x^2+1$ 要求 $B=-2$, 以及 $(x-1)\mid x(4Ax^2-4)$ 要求 $A=1$. □

5.3.4 设 $f(x)=x^5-3x^4+2x^3+2x^2-3x+1$, 在 $\mathbb{Q}[x]$ 中求一个没有重因式的多项式 g, 使其与 f 有完全相同的不可约多项式 (不计重数).

证明. 观察多项式系数可知其有理根仅可能有 ± 1, 验算可知均为根, 从而有

$$\frac{f(x)}{(x-1)(x+1)}=x^3-3x^2+3x-1=(x-1)^3$$

从而取 $g(x)=(x-1)(x+1)=x^2-1$ 即可. □

5.3.5 证明多项式 $f(x)=x^4+2x^3-15x^2+4x+20$ 有重根, 并求其所有根.

证明. 辗转相除可得 $(f,f')=x-2$, 从而知 $(x-2)^2\mid f(x)$,

$$\frac{f(x)}{(x-2)^2}=x^2+6x+5=(x+1)(x+5)$$

从而 $2,-1,-5$ 为其所有根. □

5.3.6 证明: 不可约多项式 p 是多项式 f 的 k 重因式 $\iff p \mid f, p \mid f', \ldots, p \mid f^{(k-1)}$ 但 $p \nmid f^{(k)}$.

证明一. 容易看出, 该命题等价于: $p^k \mid f \iff p$ 整除 $f, f', \ldots, f^{(k-1)}$. 下对 k 归纳,$k=1$ 时显然, 下设 $k \geq 2, < k$ 时命题成立.

$\Longrightarrow$: 显然 $p^{k-1} \mid f$, 故由归纳假设,p 整除 $f, f', \ldots, f^{(k-2)}$, 下证 $p \mid f^{(k-1)}$. 由于有 $f = p^k g$, 即 $f' = p^{k-1}(kp'g + pg')$, 故 $p^{(k-1)} \mid f'$. 从而由归纳假设,$p \mid (f')^{(k-2)} = f^{(k-1)}$.

$\Longleftarrow$: 由于 p 整除 $f', (f')', \ldots, (f')^{(k-2)}$, 故由归纳假设知 $p^{k-1} \mid f'$, 其等价于 $p^k \mid f$. □

证明二. $k=1$ 时已证,$k>1$ 时: p 是 f 的 k 重因式 $\iff p$ 是 f' 的 $k-1$ 重因式 $\iff \cdots \iff p$ 是 $f^{(k-1)}$ 的 2 重因式 $\Longrightarrow p$ 是 $f^{(k-2)}$ 的 1 重因式 $\iff p \nmid (f^{(k-1)}, f^{(k)})$, 故 $p \nmid f^{(k)}$.

另一方面,$p \nmid f^{(k)}, p \mid f^{(k-1)}$, 故 p 不为 $f^{(k-1)}$ 的重因式; 而 p 整除 $f^{(k-1)}, f^{(k-2)}$, 故 p 是 $f^{(k-2)}$ 的重因式. 综上,p 是 $f^{(k-1)}$ 的 2 重因式, 其余同上, 从而得证. □

注 2. 该结果只对 $\operatorname{char}\mathbb{F} > k$ 或 $\operatorname{char}\mathbb{F} = 0$ 的数域 $\mathbb{F}$ 上的多项式成立.

5.3.7 举例否定 "若 α 是 f' 的 m 重根, 则 α 是 f 的 $m+1$ 重根".

证明. 取 $f(x) = x^{m+1} + 1, f'(x) = (m+1)x^m, 0$ 为 f' 的 m 重根但不是 f 的 $m+1$ 重根. □

注 3. 该命题若加上条件 "α 是 f 的根" 即正确.

证明: 由题知 $(x-\alpha) \mid f, (x-\alpha)^m \mid f', (x-\alpha)^{m+1} \nmid f'$. *由 5.3.6 知* $(x-\alpha)$ *整除* $f', f'', \ldots, (f')^{(m-1)} = f^{(m)}$ *但* $(x-\alpha) \nmid f^{(m+1)}$, *加上题设* $(x-\alpha) \mid f$ *再由 5.3.6 知* $(x-\alpha)$ *是* f *的* $m+1$ *重因式.*

5.3.8 证明: 若 $(x-1) \mid f(x^n)$ 则 $(x^n - 1) \mid f(x^n)$.

证明一. 显然 $f(1) = 0$, 故 $(x-1) \mid f(x), f(x) = (x-1)g(x)$, 从而 $f(x^n) = (x^n-1)g(x^n), (x^n-1) \mid f(x^n)$. □

证明二. 显然 $f(1) = 0$. 考虑 1 的任意 n 次单位根 $\omega_k = \mathrm{e}^{\frac{2k\pi\mathrm{i}}{n}}$, 有 $f(\omega_k^n) = f(1) = 0$, 故 $(x - \omega_k) \mid f(x^n)$, 从而

$$\prod_{k=0}^{n-1}(x - \omega_k) = (x^n - 1) \mid f(x^n).$$

□

5.3.9 $p \in \mathbb{F}[x], \deg p > 0$. 若对于 $\forall f \in \mathbb{F}[x]$ 均有 $p \mid f$ 或 $(p, f) = 1$, 则 p 在 $\mathbb{F}$ 中不可约.

证明. 若 p 可被分解为次数小于 $\deg p$ 的多项式 q, r 之积, 则必有其中一个多项式次数非零, 设其为 q. 从而取 $f = q, (p, f) \neq 1, p \nmid f$, 矛盾. □

5.3.10 $p \in \mathbb{F}[x], \deg p > 0$. 若对于 $\forall f, g \in \mathbb{F}[x], p \mid fg \Longrightarrow p \mid f$ 或 $p \mid g$, 则 p 在 $\mathbb{F}$ 中不可约.

证明. 若 p 可被分解为次数小于 $\deg p$ 的多项式 q, r 之积, 则 $p \mid qr = p$ 但 $p \nmid q, p \nmid r$, 矛盾. □

思考题 2 $x^2 - 2$ 在 $\mathbb{Q}$ 上不可约而在 $\mathbb{R}$ 上可约.

证明一. 在 $\mathbb{R}$ 上有 $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ 从而可约. 而该多项式在 $\mathbb{Q}$ 上若有根 $a = p/q$, 则 $q \mid 1, p \mid (-2)$, 即 a 仅可能为 $\pm 1, \pm 2$, 而这些均不为根, 从而无根, 即不可约. □

证明二. 若在 $\mathbb{Q}$ 上有唯一分解 $x^2 - 2 = (x-a)(x-b)$, 即 $a + b = 0, ab = -2$, 即 $a^2 = 2$. 对 $\sqrt{2}$ 的无理性证明导出 $x^2 - 2$ 在 $\mathbb{Q}$ 上不可约. □

证明三. 书上例 5.3.1. □

思考题 3 设 $f = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_s^{\alpha_s}, g = p_1^{\beta_1} p_2^{\beta_2} \dots p_s^{\beta_s}$, 其中 p_i 均为不可约多项式. 证明 $(f,g) = p_1^{\gamma_1} p_2^{\gamma_2} \dots p_s^{\gamma_s}$, 其中 $\gamma_i = \min(\alpha_i, \beta_i), i = 1, \dots, s$.

证明. 设 $d = p_1^{\gamma_1} p_2^{\gamma_2} \dots p_s^{\gamma_s}$, 显然 $d \mid f, d \mid g$. 若有 f, g 的公因式 $d' = p_1^{\delta_1} p_2^{\delta_2} \dots p_s^{\delta_s}$, 则 $\forall i = 1, 2, \dots, s, \delta_i \le \alpha_i$ 且 $\delta_i \le \beta_i$, 故 $\delta_i \le \gamma_i$, 从而 $d' \mid d$, 故 $d = (f, g)$. □

思考题 4 $f_1, \dots, f_s \in \mathbb{F}[x]$ 之间两两互素, 记 $f = f_1 \dots f_s, g_i = f/f_i$, 证明 $(g_1, g_2, \dots, g_s) = 1$.

证明. 对 s 归纳,$s = 2$ 时 $(g_1, g_2) = (f_2, f_1) = 1$ 从而成立. 设 $< s$ 时命题成立, 考虑两两互素的多项式 $f_1, f_2, \dots, f_s$, 如上定义 f, g_i, 则有

$$d = (g_1, g_2, \dots, g_s) = ((g_1, g_2, \dots, g_{s-1}), g_s)$$

而

$$(g_1, g_2, \dots, g_{s-1}) = \left(\frac{f_1 \cdots f_s}{f_1}, \frac{f_1 \cdots f_s}{f_2}, \dots, \frac{f_1 \cdots f_s}{f_{s-1}} \right) = f_s \left(\frac{f_1 \cdots f_{s-1}}{f_1}, \frac{f_1 \cdots f_{s-1}}{f_2}, \dots, \frac{f_1 \cdots f_{s-1}}{f_{s-1}} \right)$$

由归纳假设知右端项为 f_s, 从而 $d = (f_s, f_1 \dots f_{s-1}) = 1$, 从而得证. □

5.4 复系数与实系数多项式的因式分解

5.4.1 求多项式 $x^5 - 1$ 在 $\mathbb{C}$ 和 $\mathbb{R}$ 上的因式分解.

证明. 在 $\mathbb{C}$ 上显然有

$$x^5 - 1 = (x-1)(x-\omega)(x-\omega^2)(x-\omega^3)(x-\omega^4)$$

其中 $\omega = \mathrm{e}^{\frac{2\pi \mathrm{i}}{5}} = \cos\dfrac{2\pi}{5} + \mathrm{i}\sin\dfrac{2\pi}{5}$. 而由于复数根成对, 故在 $\mathbb{R}$ 上有

$$\begin{aligned} x^5 - 1 &= (x-1)\left[(x-\omega)(x-\omega^4)\right]\left[(x-\omega^2)(x-\omega^3)\right] \\ &= (x-1)\left(x^2 - 2\cos\frac{2\pi}{5}x + 1\right)\left(x^2 - 2\cos\frac{4\pi}{5}x + 1\right) \\ &= (x-1)\left(x^2 + \frac{1-\sqrt{5}}{2}x + 1\right)\left(x^2 + \frac{1+\sqrt{5}}{2}x + 1\right) \end{aligned}$$

□

5.4.2 $f \in \mathbb{R}[x], \deg f = n$ 且 f 有 ℓ 个实根 (计重数), 证明 $n - \ell$ 是偶数.

证明. 将 f 分解为不可约多项式的乘积, 即

$$f = a p_1^{\alpha_1} \dots p_s^{\alpha_s} q_1^{\beta_1} \dots q_t^{\beta_t}$$

其中 p_i 均为一次多项式,q_i 均为二次多项式, 则

$$n = \sum_{i=1}^{s} \alpha_i + \sum_{i=1}^{t} 2\beta_i, \qquad \ell = \sum_{i=1}^{s} \alpha_i, \qquad n - \ell = 2\sum_{i=1}^{t} \beta_i$$

从而 $n - \ell$ 显然为偶数. □

5.4.3 求 x^4+1 在 $\mathbb{C}$ 和 $\mathbb{R}$ 上的标准分解.

证明. 在 $\mathbb{C}$ 上 x^4+1 有根 $(-1)^{1/4}=\mathrm{e}^{\frac{\pi\mathrm{i}}{4}},(-1)^{3/4}=\mathrm{e}^{\frac{3\pi\mathrm{i}}{4}},(-1)^{5/4}=\mathrm{e}^{\frac{5\pi\mathrm{i}}{4}},(-1)^{7/4}=\mathrm{e}^{\frac{7\pi\mathrm{i}}{4}}$, 因此有

$$\begin{aligned}
x^4+1&=(x-\mathrm{e}^{\frac{\pi\mathrm{i}}{4}})(x-\mathrm{e}^{\frac{3\pi\mathrm{i}}{4}})(x-\mathrm{e}^{\frac{5\pi\mathrm{i}}{4}})(x-\mathrm{e}^{\frac{7\pi\mathrm{i}}{4}})\\
&=\left[(x-\mathrm{e}^{\frac{\pi\mathrm{i}}{4}})(x-\mathrm{e}^{\frac{7\pi\mathrm{i}}{4}})\right]\left[(x-\mathrm{e}^{\frac{3\pi\mathrm{i}}{4}})(x-\mathrm{e}^{\frac{5\pi\mathrm{i}}{4}})\right]\\
&=\left(x^2-2\cos\frac{\pi}{4}x+1\right)\left(x^2-2\cos\frac{3\pi}{4}x+1\right)\\
&=\left(x^2-\sqrt{2}x+1\right)\left(x^2+\sqrt{2}x+1\right)
\end{aligned}$$

$\square$

5.4.4 $f\in\mathbb{R}[x]$ 的首项系数 $a_n>0$, 若 f 无实根, 则存在 $g,h\in\mathbb{R}[x]$ 使得 $f=g^2+h^2$.

证明. 由于实系数多项式 f 无实根, 故其不可约分解中均为二次不可约多项式, 即在 $\mathbb{C}$ 中有分解

$$f(x)=q_1(x)\dots q_m(x)=\prod_{i=1}^{m}\left[(x-\lambda_i)(x-\bar{\lambda}_i)\right]=\left(\prod_{i=1}^{m}(x-\lambda_i)\right)\left(\prod_{i=1}^{m}(x-\bar{\lambda}_i)\right)=p(x)q(x)$$

其中 $q_i(x)=(x-\lambda_i)(x-\bar{\lambda}_i)$ 均为在 $\mathbb{R}$ 上不可约的二次多项式,$\lambda_i\in\mathbb{C}$. 将 $p(x)$ 按系数的实部和虚部分为两个实系数多项式, 即

$$p(x)=g(x)+\mathrm{i}h(x),\qquad g,h\in\mathbb{R}[x]$$

再对 $p(x)$ 的系数取共轭, 有

$$g(x)-\mathrm{i}h(x)=\bar{p}(x)=\prod_{i=1}^{m}(x-\bar{\lambda}_i)=q(x)$$

从而 $f=(g+\mathrm{i}h)(g-\mathrm{i}h)=g^2+h^2$. $\square$

5.4.5 设 $p,f\in\mathbb{R}[x]$ 且 p 在 $\mathbb{R}$ 上不可约, 证明: 若 $\exists\alpha\in\mathbb{C},p(\alpha)=f(\alpha)=0$ 则 $p\mid f$.

证明. 显然 $\deg p=1$ 或 2. 若 $\deg p=1$ 则 $\alpha\in\mathbb{R},p(x)=a(x-\alpha)\mid f(x)$. 若 $\deg p=2$ 则 $\alpha\notin\mathbb{R},p(x)=a(x-\alpha)(x-\bar{\alpha})$, 而 $f(\alpha)=f(\bar{\alpha})=0$, 从而 $p(x)=a(x-\alpha)(x-\bar{\alpha})\mid f(x)$. $\square$

思考题 5 找出 x^n-1 在 $\mathbb{C}$ 上的所有 n 次本原单位根.

证明. 取任一 n 次本原单位根 $\omega=\mathrm{e}^{\frac{2k\pi\mathrm{i}}{n}}$, 则其幂次 $\omega^0,\omega^1,\omega^2,\dots,\omega^{n-1}$ 遍历所有 n 次单位根. 设 $d=\gcd(n,k)$, 即 $n=dm,k=d\ell$, 则 $\omega^m=\mathrm{e}^{\frac{2mk\pi\mathrm{i}}{n}}=\mathrm{e}^{\frac{2\ell n\pi\mathrm{i}}{n}}=1$, 仅在 $m=n,d=1$ 时能遍历所有 n 次单位根, 故所有 n 次本原单位根即 $\mathrm{e}^{\frac{2k\pi\mathrm{i}}{n}},\gcd(n,k)=1$. $\square$

5.5 有理系数多项式

5.5.1 求下列多项式的有理根:

1. $2x^4-x^3+2x-3$;
2. $4x^4-7x^2-5x-1$;
3. $x^4+6x^3+12x^2+11x+6$.

证明. (1)1; (2)$-1/2$ (2 重); (3)$-2,-3$. $\square$

5.5.2 判别下列多项式在 $\mathbb{Q}$ 上是否可约:

1. x^6+x^3+1;
2. x^p+px+1, p 是奇素数;
3. x^4+4;
4. $x^4+4kx+1, k\in\mathbb{Z}$.

证明. 1. 代换 $x=t+1$, 故原式 $=(t+1)^6+(t+1)^3+1=t^6+6t^5+15t^4+21t^3+18t^2+9t+3$. 用 Eisenstein 判别法 (取 $p=3$) 知其在 $\mathbb{Q}$ 上不可约.

2. 令 $x=t-1$, 则原式 $=(t-1)^p+pt+1-p=t^p+\sum_{k=2}^{p-1}\binom{p}{k}(-1)^kt^k+2pt+p$, 从而由 Eisenstein 判别法知其在 $\mathbb{Q}$ 上不可约.

3. $x^4+4=(x^2-2x+2)(x^2+2x+2)$, 故可约.

4. 若原式 $f(x)$ 在 $\mathbb{Q}$ 上可约, 则也在 $\mathbb{Z}$ 上可约. 显然 $f(x)$ 的有理根仅可能有 ±1, 但 $f(\pm1)\neq 0, \pm1$ 均不是根, 从而 $f(x)$ 在 $\mathbb{Z}$ 上没有一次 (和三次) 因式. 若原式在 $\mathbb{Z}$ 上有二次因式, 即设

$$x^4+4kx+1=(x^2+ax+b)(x^2+cx+d),\qquad a,b,c,d\in\mathbb{Z}$$

比较系数可得 $a+c=0, ac+b+d=0, ad+bc=4k, bd=1$, 从而 $b=d=\pm1$,$ac=-a^2=\mp2$, 矛盾于 $a\in\mathbb{Z}$, 故原式在 $\mathbb{Z}$ 上也没有二次因式, 故在 $\mathbb{Z}$ 和 $\mathbb{Q}$ 上不可约.

□

5.5.3 p 为素数, 证明 $f(x)=x^p-px+(2p-1)$ 在 $\mathbb{Q}$ 上不可约.

证明. 令 $x=t+1$, 则

$$f(t+1)=(t+1)^p-p(t+1)+(2p-1)=t^p+\sum_{k=2}^{p-1}\binom{p}{k}t^k+p$$

从而由 Eisenstein 判别法知其在 $\mathbb{Q}$ 上不可约. □

5.5.4 设 $p_i(i=1,2,\dots,t)$ 为 t 个互异素数, 证明 $f(x)=x^n-p_1\dots p_t$ 在 $\mathbb{Q}$ 上不可约.

证明. 取素数 p 为任一 p_i, 由 Eisenstein 判别法知 $f(x)$ 在 $\mathbb{Q}$ 上不可约. □

5.5.5 f 是首一整系数多项式, 若 $f(0)$ 和 $f(1)$ 均为奇数, 则 f 没有有理根.

证明一. 设 $f(x)=\sum_{k=0}^{n}a_kx^k$, 若其有有理根 r/s 则 $s\mid a_n=1, r\mid a_0=f(0)$, 从而有理根仅可能为 $c\in\mathbb{Z}, c\mid f(0)$. 而 $f(0)$ 为奇数, 故 c 为奇数. 由于 $f(c)=0$, 故

$$-f(1)=f(c)-f(1)=\sum_{k=0}^{n}a_k(c^k-1)$$

而对 $\forall k\in\mathbb{N}, c^k-1$ 为偶数, 故等式右端为偶数, 但左端为奇数, 矛盾, 从而 f 无有理根. □

证明二. 若 f 有有理根 r/s, 则 $s\mid a_n=1, r\mid a_0=f(0)$, 即 $s=\pm1, r$ 为奇数. 而 $(r-ms)\mid f(m)$, 故 $(r\pm1)\mid f(1)$, 但 $r\pm1$ 为偶数,$f(1)$ 为奇数, 矛盾, 故无有理根. □

证明三. 若 f 有有理根 c, 同上知 c 为奇数, 从而有 $f(x)=(x-c)g(x), g\in\mathbb{Z}[x]$. 分别代入 0,1 知 $f(0)=-cg(0), f(1)=(1-c)g(1)$, 由 $f(0), f(1)$ 均为奇数知 $-c, 1-c$ 均为奇数, 矛盾. □

注 4. 若无首一条件, 可取 $f(x)=2x-1, f(0)=1, f(1)=3$ 但有有理根 $1/2$.

注 5. 命题可作简单推广:f 是首一整系数多项式, p 是素数, 若 $f(0)\neq 0, f(1)\neq 0 \pmod p$, 则 f 没有有理根.

思考题 6 设 $a_1,\dots,a_n$ 是互不相同的整数, 证明 $f(x)=\prod_{i=1}^{n}(x-a_i)-1$ 在 $\mathbb{Q}$ 上不可约.

证明. 若 f 在 $\mathbb{Q}$ 上可约, 则在 $\mathbb{Z}$ 上可约, 设 $f(x)=g(x)h(x), g,h\in\mathbb{Z}[x]$, 则 $\forall i=1,\dots,n, f(a_i)=g(a_i)h(a_i)=-1$, 从而 $g(a_i)=\pm1, h(a_i)=\mp1$, 故 $g(a_i)+h(a_i)=0$, 即多项式 $g(x)+h(x)$ 有 n 个不同的根. 但 $\deg(g+h)\le\max(\deg g,\deg h)<\deg f=n$, 从而由代数基本定理得到矛盾. $\square$

6 第六章 相似标准形

6.1 特征值与特征向量

6.1.1 求下列矩阵的特征值与特征向量:

$$(1)\begin{pmatrix} 3 & 2 & -1 \\ -2 & -2 & 2 \\ 3 & 6 & -1 \end{pmatrix}, \qquad (2)\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

证明. (1) $\det(\lambda I - A) = \lambda^3 - 12\lambda + 16 = (\lambda-2)^2(\lambda+4)$, 从而特征值为 2 (1 重) 和 -4 (2 重). 可解得 2 对应的特征向量为 $(1,0,1)^\mathsf{T}, (2,-1,0)^\mathsf{T}$; -4 对应的特征向量为 $(1,-2,3)^\mathsf{T}$.

(2) $\det(\lambda I - A) = \lambda^3 - \lambda^2 - \lambda + 1 = (\lambda-1)^2(\lambda+1)$, 从而特征值为 1 (2 重) 和 -1 (1 重). 可解得 1 对应的特征向量为 $(1,0,1)^\mathsf{T}, (0,1,0)^\mathsf{T}$; -1 对应的特征向量为 $(1,0,-1)^\mathsf{T}$.x □

6.1.2 若 $A = \begin{pmatrix} -2 & 0 & 0 \\ 2 & x & 2 \\ 3 & 1 & 1 \end{pmatrix}$ 与 $B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & y \end{pmatrix}$ 相似, 求 x, y.

证明. A 的特征多项式为 $\phi_A(\lambda) = \lambda^3 + (-x+1)\lambda^2 + (-x-4)\lambda + (2x-4)$, B 的特征多项式为 $\phi_B(\lambda) = \lambda^3 + (-y-1)\lambda^2 + (y-2)\lambda + 2y$. 对比系数可得方程组

$$1 - x = -y - 1, -x - 4 = y - 2, 2y = 2x - 4$$

解得 $x = 0, y = -2$. □

6.1.3 设 $A = \begin{pmatrix} a & 1 & 1 \\ 1 & a & -1 \\ 1 & -1 & a \end{pmatrix}$, 其中 $a \in \mathbb{R}$, 求 $A - I$ 的特征值.

证明. $\det(\lambda I - A) = (\lambda - a)^3 - 3(\lambda - a) - 2 = (\lambda - a - 1)^2(\lambda - a + 2)$, 因此 A 的特征值为 $a+1$ (2 重) 和 $a-2$ (1 重), 从而 $A - I$ 的特征值为 a (2 重) 和 $a-3$ (1 重), $\det(A - I) = a^2(a-3)$. □

6.1.4 设 $A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & -1 & 2 \\ 1 & 2 & 1 \end{pmatrix}$, $k \in \mathbb{N}_+$, 求 A^k 的特征值与特征向量.

证明. A 的特征多项式 $\phi_A(\lambda) = \det(\lambda I - A) = \lambda^3 - 2\lambda^2 - 6\lambda + 9 = (\lambda - 3)(\lambda^2 + \lambda - 3)$, 因此特征值为

$$\lambda_1 = 3, \lambda_2 = \frac{-1+\sqrt{13}}{2}, \lambda_3 = \frac{-1-\sqrt{13}}{2}$$

其分别对应特征向量

$$x_1 = (2,1,2)^\mathsf{T}, x_2 = (3+\sqrt{13}, -4, -1-\sqrt{13})^\mathsf{T}, x_3 = (3-\sqrt{13}, -4, -1+\sqrt{13})^\mathsf{T}$$

□

6.1.5 线性空间 V 上的线性变换 $\mathscr{A}$ 有特征值 λ_0, 证明: (1) 对任意 $k\in\mathbb{N},\lambda_0^k$ 为 $\mathscr{A}^k$ 的特征值. (2) 若 $\mathscr{A}$ 可逆, 则 $\lambda_0\neq 0$ 且 λ_0^{-1} 为 $\mathscr{A}^{-1}$ 的特征值.

证明. (1) 取对应于 λ_0 的特征向量 x, 有 $\mathscr{A}x=\lambda_0 x$, 因此 $\mathscr{A}^2x=\lambda_0\mathscr{A}x=\lambda_0^2x,\dots,\mathscr{A}^kx=\lambda_0^kx$, 从而 λ_0^k 为 $\mathscr{A}^k$ 的特征值.

(2) 若 $\mathscr{A}$ 可逆, 则其行列式非零, 而行列式为全体特征值之积, 故 $\mathscr{A}$ 的所有特征值均非零. 取对应于 λ_0 的特征向量 x, 有 $\mathscr{A}x=\lambda_0x$, 故 $x=\lambda_0\mathscr{A}^{-1}x$, 即 $\mathscr{A}^{-1}x=\lambda_0^{-1}x$, 即 λ_0^{-1} 是 $\mathscr{A}^{-1}$ 的特征值. □

6.1.6 $A\in M_n(\mathbb{F})$, 若 $A^2=A$, 则 A 的特征值仅能为 0 或 1.

证明. 对 A 的任意特征值 λ 及其对应的特征向量 x, $\lambda x=Ax=A^2x=\lambda^2x$, 从而 $\lambda=\lambda^2$, 即 $\lambda=0$ 或 1. □

6.1.7 $A,B\in M_n(\mathbb{F})$, 证明 AB 与 BA 的特征多项式相等.

证明一. 由于 $\forall A,B\in M_n(\mathbb{F}),\det(I-AB)=\det(I-BA)$, 因此 $\lambda\neq 0$ 时有 $\phi_{AB}(\lambda)=\det(\lambda I-AB)=\lambda^n\det(I-AB/\lambda)=\lambda^n\det(I-BA/\lambda)=\det(\lambda-BA)=\phi_{BA}(\lambda)$, 而 $\lambda=0$ 时自然有 $\phi_{AB}(0)=\det(AB)=\det(BA)=\phi_{BA}(0)$, 从而两特征多项式相等. □

证明二. 对于 AB 的任意特征值 $\lambda\in\mathbb{C}$ 及其特征向量 $x\in\mathbb{F}^n$, 有 $BABx=B(\lambda x)=\lambda Bx$, 从而 λ 是 BA 的特征值. 同理知 BA 的特征值也是 AB 的特征值, 因此 AB 与 BA 的特征值相同, 故两者的特征多项式相同. □

6.1.8 $A\in M_n(\mathbb{F}),M\in M_m(\mathbb{F}),P\in\mathbb{F}^{n\times m}$ 是列满秩方阵. 若 $AP=PM$, 证明 M 的特征值也是 A 的特征值.

证明. 对 M 的任一特征值 λ 及其对应的特征向量 $x\in\mathbb{F}^m$, 有 $APx=PMx=\lambda Px$. 由于 P 列满秩, 故由 Sylvester 不等式知 $\operatorname{rank}(Px)\geq\operatorname{rank}P+\operatorname{rank}x-m=1$, 因此 Px 不为零向量, 从而 λ 为 A 的特征值, 其特征向量为 Px.

也可设 $P=(\alpha_1,\dots,\alpha_m),x=(x_1,\dots,x_m)^{\mathsf{T}}$, 从而 $Px=x_1\alpha_1+\cdots+x_m\alpha_m$, 由 P 列满秩知 α_i 之间线性无关, 故 $Px\neq 0$. □

6.1.9 求如下方阵的极小多项式:

$$(1)\begin{pmatrix}3&1&-1\\2&2&-1\\2&2&0\end{pmatrix},\qquad(2)\begin{pmatrix}1&1&0&0\\-1&-1&0&0\\-2&-2&2&1\\1&1&-1&0\end{pmatrix}$$

证明. (1) 特征多项式 $\phi_A(\lambda)=\det(\lambda I-A)=\lambda^3-5\lambda^2+8\lambda-4=(\lambda-1)(\lambda-2)^2$, 因此其极小多项式 $m(\lambda)$ 仅能为 $(\lambda-1)(\lambda-2)$ 或 $\phi_A(\lambda)$. 取前者可得

$$m(A)=(A-I)(A-2I)=\begin{pmatrix}2&0&-1\\2&0&-1\\4&0&-2\end{pmatrix}\neq O$$

因此其极小多项式 $m(\lambda)=\phi_A(\lambda)=(\lambda-1)(\lambda-2)^2$.

(2) 特征多项式 $\phi_A(\lambda)=\lambda^2(\lambda-1)^2$, 因此其极小多项式 $m(\lambda)=\lambda^i(\lambda-1)^j(i,j=1,2)$. 分别计算:

$$A(A-I)=\begin{pmatrix}-1&-1&0&0\\1&1&0&0\\-1&-1&1&1\\1&1&-1&-1\end{pmatrix},A^2(A-I)=\begin{pmatrix}0&0&0&0\\0&0&0&0\\-1&-1&1&1\\1&1&-1&-1\end{pmatrix},A(A-I)^2=\begin{pmatrix}1&1&0&0\\-1&-1&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

均不为 O, 因此 $m(\lambda)=\phi_A(\lambda)=\lambda^2(\lambda-1)^2$. □

6.1.10 设 $A = \mathrm{diag}(A_1, A_2)$ 为准对角方阵, 若 A_1, A_2 的极小多项式分别为 $m_1(\lambda), m_2(\lambda)$, 则 A 的极小多项式 $m(\lambda) = [m_1(\lambda), m_2(\lambda)]$.

证明. 对于 A 的任一零化多项式 f, 都有 $f(A) = \mathrm{diag}(f(A_1), f(A_2)) = O$, 即 $f(A_1) = O, f(A_2) = O$, 因此 $m_1 \mid f, m_2 \mid f$, 故 $[m_1, m_2] \mid f$, 特别地 $[m_1, m_2] \mid m$. 易知 $[m_1, m_2]$ 也是 A 的零化多项式, 因此 A 的极小多项式 $m \mid [m_1, m_2]$. 综上 $m = [m_1, m_2]$. □

6.2 特征子空间与根子空间

6.2.1 V 是 $\mathbb{C}$ 上的 n 维线性空间, V 上有线性变换 $\mathscr{A}, \mathscr{B}$, 其满足 $\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A}$. 证明: (1) 若 λ_0 是 $\mathscr{A}$ 的特征值, 则 V_{λ_0} 是 $\mathscr{B}$ 的不变子空间;(2)$\mathscr{A}, \mathscr{B}$ 至少有一个公共特征向量.

证明. (1)$\forall x \in V_{\lambda_0}, \mathscr{A}(\mathscr{B}x) = \mathscr{B}\mathscr{A}x = \lambda_0 \mathscr{B}x$, 因此 $\mathscr{B}x \in V_{\lambda_0}$, 从而 V_{λ_0} 是 $\mathscr{B}$ 的不变子空间.

(2) 取限制映射 $\mathscr{B}|_{V_{\lambda_0}}$, 则其在 V_{λ_0} 上有特征值 μ 及特征向量 $v \in V_{\lambda_0}$, 满足 $\mathscr{B}|_{V_{\lambda_0}} v = \mathscr{B}v = \mu v$. 而 $v \in V_{\lambda_0}$ 可知 v 也是 $\mathscr{A}$ 的特征向量, 从而得证. □

6.2.2 V 是 $\mathbb{C}$ 上的 n 维线性空间,V 上的线性变换 $\mathscr{A}$ 在基 $\varepsilon_1, \ldots, \varepsilon_n$ 下的矩阵为 A, 求 $\mathscr{A}$ 的所有特征子空间和根子空间.

$$(1)A = \begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}, \qquad (2)A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

证明. (1)$\det(\lambda I - A) = \lambda^3 - \lambda^2 - 5\lambda - 3 = (\lambda - 3)(\lambda + 1)^2$, 从而有特征值 $\lambda_1 = 3, \lambda_2 = -1$ (2 重). 计算得 $\lambda_1 = 3$ 的特征向量为 $\alpha_1 = (1, 2, 2)^{\mathsf{T}}$, 对于 $\lambda_2 = -1$, 有

$$-I - A = \begin{pmatrix} -2 & 3 & -4 \\ -4 & 6 & -8 \\ -6 & 7 & -8 \end{pmatrix}, (-I - A)^2 = 16 \begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \\ 2 & -2 & 2 \end{pmatrix},$$

从而 $(-I - A)x = 0$ 的基础解系为 $\alpha_2 = (1, 2, 1)^{\mathsf{T}}, (-I - A)^2 x = 0$ 的基础解系为 $\alpha_2, \alpha_3 = (1, 0, -1)^{\mathsf{T}}$, 从而 $V_3 = \mathrm{span}(\alpha_1), W_{-1} = \mathrm{span}(\alpha_2, \alpha_3)$.

(2)$\det(\lambda I - A) = \lambda^4 - 5\lambda^3 + 6\lambda^2 + 4\lambda - 8 = (\lambda - 2)^3(\lambda + 1)$, 从而有特征值 $\lambda_1 = 2$ (3 重) 和 $\lambda_2 = -1$. 对于 $\lambda_1 = 2$, 有

$$2I - A = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, (2I - A)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}, (2I - A)^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 27 \end{pmatrix}$$

从而 $(2I - A)x = 0$ 的基础解系为 $\alpha_1 = (1, 0, 0, 0)^{\mathsf{T}}, (2I - A)^2 x = 0$ 的基础解系为 $\alpha_1, \alpha_2 = (0, 1, 0, 0)^{\mathsf{T}}, (2I - A)^3 x = 0$ 的基础解系为 $\alpha_1, \alpha_2, \alpha_3 = (0, 0, 1, 0)^{\mathsf{T}}$, 因此 $W_2 = \mathrm{span}(\alpha_1, \alpha_2, \alpha_3)$. 对于 $\lambda_2 = -1$, 可解得其特征向量为 $\alpha_4 = (0, 0, 0, 1)^{\mathsf{T}}$, 即 $V_{-1} = \mathrm{span}(\alpha_4)$. □

6.2.3 V 是 $\mathbb{C}$ 上的 n 维线性空间,V 上的线性变换 $\mathscr{A}$ 的极小多项式 $m(\lambda) = (\lambda - \lambda_1)^{m_1} \ldots (\lambda - \lambda_s)^{m_s}$, 其中 $\lambda_1, \ldots, \lambda_s$ 互不相同. 证明根子空间 $W_{\lambda_i} = \ker(\mathscr{A} - \lambda_i \mathscr{I})^{m_i}, i = 1, \ldots, s$.

证明. 显然 $\ker(\mathscr{A}-\lambda_i\mathscr{I})^{m_i}\subset W_{\lambda_i}$. 由于有准素分解 $V=W_{\lambda_1}\oplus\cdots\oplus W_{\lambda_s}$, 从而仅需证明

$$V=\ker(\mathscr{A}-\lambda_1\mathscr{I})^{m_1}\oplus\cdots\oplus\ker(\mathscr{A}-\lambda_s\mathscr{I})^{m_s}.$$

设 $f_i(\lambda)=m(\lambda)/(\lambda-\lambda_i)^{m_i}$, 则显然 $(f_1,\ldots,f_s)=1$, 即有多项式 $u_1,\ldots,u_s$ 满足

$$u_1f_1+\cdots+u_sf_s=1$$

从而 $\forall v\in V$,

$$u_1(\mathscr{A})f_1(\mathscr{A})v+\cdots+u_s(\mathscr{A})f_s(\mathscr{A})v=v$$

记 $v_i=u_i(\mathscr{A})f_i(\mathscr{A})v,(\mathscr{A}-\lambda_i\mathscr{I})^{m_i}v_i=u_i(\mathscr{A})m(\mathscr{A})v=0$, 从而 $v_i\in\ker(\mathscr{A}-\lambda_i\mathscr{I})^{m_i}$, 从而得到

$$V=\ker(\mathscr{A}-\lambda_1\mathscr{I})^{m_1}+\cdots+\ker(\mathscr{A}-\lambda_s\mathscr{I})^{m_s}.$$

而由 W_{λ_i} 的和为直和得到上式的和为直和, 从而得证. □

6.2.4 线性空间 V 上的线性变换 $\mathscr{A}$ 有至少两个不同的特征值, 证明 $\mathscr{A}$ 的全体特征向量并上零向量不构成 V 的子空间.

证明. 设 $U=\{v\in V|\exists\lambda\in\mathbb{F},\mathscr{A}v=\lambda v\}$. 若其为线性子空间, 由于 $\mathscr{A}$ 有两个不同的特征值 λ_1,λ_2, 分别对应两个特征向量 $v_1,v_2\in U$, 则 $\mathscr{A}(v_1+v_2)=\lambda_1v_1+\lambda_2v_2=\lambda(v_1+v_2)$, 即 $(\lambda_1-\lambda)v_1+(\lambda_2-\lambda)v_2=0$. 由于 v_1,v_2 线性无关, 则该式仅有零解, 即 $\lambda=\lambda_1=\lambda_2$, 矛盾. 从而 U 不是线性子空间. □

6.2.5 V 是 $\mathbb{F}$ 上的线性空间,V 上的线性变换 $\mathscr{A}$ 有零化多项式 $f\in\mathbb{F}[\lambda]$. 设 $f=f_1\cdots f_k$, 其中 $f_1,\ldots,f_k$ 之间两两互素. 证明 $V=\ker f_1(\mathscr{A})\oplus\cdots\oplus\ker f_k(\mathscr{A})$.

证明. 先证 $V=\ker f_1(\mathscr{A})+\cdots+\ker f_k(\mathscr{A})$. 设 $g_i=f/f_i$, 则 $(g_1,g_2,\ldots,g_k)=1$(思考题 5.4), 从而有多项式 $u_1,\ldots,u_k$ 满足 $u_1g_1+\cdots+u_kg_k=1$, 因此

$$u_1(\mathscr{A})g_1(\mathscr{A})+\cdots+u_k(\mathscr{A})g_k(\mathscr{A})=\mathscr{I},$$

从而 $\forall v\in V$,

$$u_1(\mathscr{A})g_1(\mathscr{A})v+\cdots+u_k(\mathscr{A})g_k(\mathscr{A})v=v.$$

设 $v_i=u_i(\mathscr{A})g_i(\mathscr{A})v$, 则 $f_i(\mathscr{A})v_i=u_i(\mathscr{A})f(\mathscr{A})v=0$, 即 $v_i\in\ker f_i(\mathscr{A})$, 且有

$$v=v_1+\cdots+v_k,$$

从而得证.

再证直和, 即证

$$\widehat{W_i}=\ker f_i(\mathscr{A})\cap\sum_{j\neq i}\ker f_j(\mathscr{A})=\{0\}$$

任取 $w\in\widehat{W_i}$, 则 w 可写成 $w_j\in\ker f_j(\mathscr{A})(j\neq i)$ 的和的形式. 而 $f_i(\mathscr{A})w=0$, 因此 $g_j(\mathscr{A})w=0(j\neq i)$, 从而

$$w=\sum_{j\neq i}w_j=\sum_{j\neq i}u_j(\mathscr{A})g_j(\mathscr{A})w=0$$

从而得证. □

6.2.6 $\mathscr{A}$ 是线性空间 V 上的线性变换,W 是 $\mathscr{A}$ 的不变子空间,$\lambda_1,\dots,\lambda_k$ 是 $\mathscr{A}$ 互不相同的特征值,$\alpha_1,\dots,\alpha_k$ 分别是 $\mathscr{A}$ 属于 $\lambda_1,\dots,\lambda_k$ 的根向量. 若 $\alpha=\alpha_1+\cdots+\alpha_k\in W$, 证明 $\alpha_i\in W, i=1,\dots,k$.

证明. 由准素分解知 $\alpha\in\bigoplus_{i=1}^{k}W_{\lambda_i}$, 因此 $\alpha\in\left(\bigoplus_{i=1}^{k}W_{\lambda_i}\right)\cap W=\bigoplus_{i=1}^{k}(W_{\lambda_i}\cap W)$. 由于 $\alpha_i\in W_{\lambda_i}$, 因此 $\alpha_i\in W_{\lambda_i}\cap W$, 从而得证. □

6.2.7 证明推论 6.2.10.

> **推论 6.2.10.** λ_0 是线性变换 $\mathscr{A}$ 的特征值, 其代数重数为 r_0, W_{λ_0} 为 $\mathscr{A}$ 属于特征值 λ_0 的根子空间, 则 $W_{\lambda_0}=\ker(\mathscr{A}-\lambda_0\mathscr{I})^{r_0}$ 且 $\dim W_{\lambda_0}=r_0$.

证明. 由推论 6.2.9 知 $\mathscr{A}|_{W_{\lambda_0}}$ 的特征多项式为 $(\lambda-\lambda_0)^{r_0}$, 因此 $\dim W_{\lambda_0}=r_0$ 且 $(\mathscr{A}|_{W_{\lambda_0}}-\lambda_0\mathscr{I}_{W_{\lambda_0}})^{r_0}=0$, 从而 $\forall\alpha\in W_{\lambda_0},(\mathscr{A}-\lambda_0\mathscr{I})^{r_0}\alpha=0$, 即 $W_{\lambda_0}\subset\ker(\mathscr{A}-\lambda_0\mathscr{I})^{r_0}$. 反之显然 $W_{\lambda_0}\supset\ker(\mathscr{A}-\lambda_0\mathscr{I})^{r_0}$, 从而得证. □

6.3 对角化

6.3.1 设下列的 A 为复矩阵, 求可逆矩阵 P 使得 $P^{-1}AP$ 为对角阵.

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

证明.

$$(1)P=\begin{pmatrix}-1&0&-2\\1&0&1\\1&1&0\end{pmatrix},P^{-1}AP=\operatorname{diag}(-2,1,1).\qquad (2)P=\begin{pmatrix}0&0&-1&-3\\0&0&-1&2\\0&1&0&10\\1&0&5&0\end{pmatrix},P^{-1}AP=\operatorname{diag}(2,2,1,1)$$

□

6.3.2 V 是数域 $\mathbb{F}$ 上的 4 维向量空间, 线性变换 $\mathscr{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的矩阵为

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

求 V 的一个基, 使 $\mathscr{A}$ 在其下的矩阵为对角矩阵, 并写出该对角阵.

证明. 可取可逆矩阵

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

使得 $P^{-1}AP=\operatorname{diag}(1,1,0,0)$. 从而可知在基 $\varepsilon_4,\varepsilon_1+\varepsilon_3,\varepsilon_3,\varepsilon_2$ 下 $\mathscr{A}$ 的矩阵为对角矩阵. □

6.3.3 若线性变换 $\mathscr{A}$ 满足 $\mathscr{A}^2=\mathscr{I}$, 证明 $\mathscr{A}$ 可对角化.

证明. $\mathscr{A}$ 的零化多项式 $\lambda^2-1=(\lambda-1)(\lambda+1)$ 无重根, 从而极小多项式无重根, 因此可知 $\mathscr{A}$ 可对角化. □

6.3.4 若存在 $m\in\mathbb{N}_+$ 使 n 阶矩阵 A 满足 $A^m=I_n$, 证明 A 可对角化.

证明. $\mathscr{A}$ 的零化多项式 $\lambda^m-1=\prod\limits_{k=0}^{m-1}(\lambda-\mathrm{e}^{\frac{2k\pi\mathrm{i}}{m}})$ 无重根, 从而极小多项式无重根, 因此可知 $\mathscr{A}$ 可对角化. □

6.3.5 $\mathscr{A}$ 为 $\mathbb{C}$ 上线性空间 V 上的线性变换, 证明 $\mathscr{A}$ 可对角化 $\iff$ 对 $\mathscr{A}$ 的任一不变子空间 V_1 都存在另一不变子空间 V_2, 使得 $V=V_1\oplus V_2$.

证明. $\Longrightarrow$: 由于 $\mathscr{A}$ 可对角化, 故其极小多项式 $m(\lambda)$ 无重根. 而 $m(\lambda)$ 是 $\mathscr{A}|_{V_1}$ 的零化多项式, 从而 $\mathscr{A}|_{V_1}$ 的极小多项式无重根, 即 $\mathscr{A}|_{V_1}$ 可对角化. 因此可在 V_1 中取一组基 $\alpha_1,\dots,\alpha_r$, 使每个基向量都是 $\mathscr{A}|_{V_1}$ 的特征向量, 从而也是 $\mathscr{A}$ 的特征向量. 因此可将其扩充为 V 上的一组基 $\alpha_1,\dots,\alpha_r,\alpha_{r+1},\dots,\alpha_n$, 其中每个基向量是 $\mathscr{A}$ 的特征向量. 因此取 $V_2=\operatorname{span}(\alpha_{r+1},\dots,\alpha_n)$ 即可.

$\Longleftarrow$: 任取 $\mathscr{A}$ 的特征值 λ, 取特征子空间 V_λ, 由题设有不变子空间 V_2 使得 $V=V_\lambda\oplus V_2$. 从而根子空间 W_λ 与 V_2 的交 $W_\lambda\cap V_2$ 也是不变子空间. 任取 $v\in W_\lambda\cap V_2, v\neq 0$, 有 $k\in\mathbb{N}_+$ 使得 $(\mathscr{A}-\lambda\mathscr{I})^k v=0,(\mathscr{A}-\lambda\mathscr{I})^{k-1}v\neq 0$, 故 $(\mathscr{A}-\lambda\mathscr{I})^{k-1}v\in V_\lambda$, 但该向量也在 V_2 中, 故也在 $V_2\cap V_\lambda=\{0\}$ 中, 矛盾, 从而 $W_\lambda\cap V_2=\{0\}, W_\lambda=V_\lambda$. 由 λ 的任意性可知, $V=\bigoplus\limits_{i=1}^{k}W_{\lambda_k}=\bigoplus\limits_{i=1}^{k}V_{\lambda_k}$, 其中 $\lambda_1,\dots,\lambda_k$ 是 $\mathscr{A}$ 的全体互不相同特征值, 从而可知 $\mathscr{A}$ 可对角化.

(证明二) $\Longleftarrow$: 记 $\mathscr{A}$ 的全体特征子空间的直和为 V_1. 若 $\mathscr{A}$ 不可对角化, 则 $V_1\subsetneq V$, 从而有不变子空间 $V_2, V=V_1\oplus V_2$. 但 $\mathscr{A}|_{V_2}$ 在 $\mathbb{C}$ 上必然有特征向量 $v\in V_1\cap V_2=\{0\}$, 矛盾. □

6.3.6 设 A 为数域 $\mathbb{F}$ 上的幂零矩阵, 若 A 可对角化, 证明 $A=O$.

证明. A 可对角化即存在可逆矩阵 P 使得 $PAP^{-1}=D=\operatorname{diag}(\lambda_1,\dots,\lambda_n)$, 而 A 幂零即存在 $m\in\mathbb{N}_+$ 使 $A^m=O$, 从而 $D^m=PA^mP^{-1}=O$, 即 $\lambda_i^m=0,\lambda_i=0$, 故 $D=O,A=O$. □

6.3.7 n 阶矩阵 A 有 k 个不同的特征值 $\lambda_1,\dots,\lambda_k$, 证明: 若 A 可对角化, 则存在 n 阶矩阵 $A_1,\dots,A_k$ 使得

$$(1)A_iA_j=\delta_{ij}A_i, (2)\sum_{i=1}^{k}A_i=I_n, (3)A=\sum_{i=1}^{k}\lambda_iA_i$$

证明. A 可对角化即存在可逆矩阵 P 使得 $PAP^{-1}=D=\operatorname{diag}(\lambda_1I_{r_1},\dots,\lambda_kI_{r_k})$, 其中 r_i 为 λ_i 的重数. 设 $D_i=\operatorname{diag}(O_{r_1},\dots,I_{r_i},\dots,O_{r_k})$, 记 $A_i=P^{-1}D_iP$, 下证 $A_1,\dots,A_k$ 满足题设三条性质.

1. $A_iA_j=P^{-1}D_iD_jP$, 若 $i\neq j$ 则 $D_iD_j=O,A_iA_j=O$; 若 $i=j$ 则 $D_iD_j=D_i,A_iA_j=P^{-1}D_iP=A_i$.
2. $\sum\limits_{i=1}^{k}A_i=P^{-1}\left(\sum\limits_{i=1}^{k}D_i\right)P=P^{-1}I_nP=I_n$.
3. $\sum\limits_{i=1}^{k}\lambda_iA_i=P^{-1}\left(\sum\limits_{i=1}^{k}\lambda_iD_i\right)P=P^{-1}DP=A$.

从而得证. □

6.3.8 设 $\mathscr{A}$ 的所有互不相同特征值为 $\lambda_1,\dots,\lambda_s,V_{\lambda_i},W_{\lambda_i}$ 分别为对应于 λ_i 的特征子空间和根子空间, 证明 $\mathscr{A}$ 可对角化 $\iff V_{\lambda_i}=W_{\lambda_i}, i=1,\dots,s$.

证明. $\mathscr{A}$ 可对角化 $\iff$ 每个特征值 λ_i 的几何重数和代数重数相等 $\iff \dim V_{\lambda_i}=\dim W_{\lambda_i} \iff V_{\lambda_i}=W_{\lambda_i}$. 最后一个等价关系是由 $V_{\lambda_i}\subset W_{\lambda_i}$ 得到的. □

6.4 λ-矩阵

6.4.1 判断下列 λ-矩阵是否可逆, 若可逆则求其逆矩阵.

$$(1)A(\lambda)=\begin{pmatrix}1-2\lambda^2 & \lambda^2-\lambda & -\lambda\\ \lambda^2-\lambda & -\lambda^2-1 & \lambda\\ \lambda & -\lambda & 1\end{pmatrix},\qquad (2)A(\lambda)=\begin{pmatrix}1-\lambda & \lambda^2 & \lambda & 1\\ \lambda & \lambda & -\lambda & \lambda+1\\ \lambda^2+1 & \lambda^2 & -\lambda^2 & 0\\ 1 & -1 & \lambda & -\lambda\end{pmatrix}$$

证明. (1)$\det A(\lambda)=-1\in\mathbb{C}$, 因此可逆, 且

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

(2)$\det A(\lambda)=-\lambda^6-3\lambda^5-2\lambda^4-\lambda^3-\lambda^2-2\lambda$, 因此不可逆. 事实上 $\det A(0)=0$, 因此 $\det A(\lambda)$ 不为非零常数, 从而不可逆. □

6.4.2 求下列 λ-矩阵的标准型.

$$(1)\begin{pmatrix}\lambda^3-\lambda & 2\lambda^2\\ \lambda^2+5\lambda & 3\lambda\end{pmatrix},\qquad (2)\begin{pmatrix}1-\lambda & \lambda^2 & \lambda\\ \lambda & \lambda & -\lambda\\ \lambda^2+1 & \lambda^2 & -\lambda^2\end{pmatrix}$$

证明. (1) $\mathrm{diag}(\lambda,\lambda(\lambda^2-10\lambda-3))$, (2) $\mathrm{diag}(1,\lambda,\lambda(\lambda+1))$. □

6.4.3 n 阶 λ-矩阵 $A(\lambda)$ 可逆 $\iff\forall c\in\mathbb{C},A(c)$ 可逆.

证明. $A(\lambda)$ 可逆 $\iff\det A(\lambda)=a\neq 0,a\in\mathbb{F}\iff\forall c\in\mathbb{C},\det A(c)=a\neq 0,a\in\mathbb{F}\iff\forall c\in\mathbb{C},A(c)$ 可逆. □

6.4.4 数域 $\mathbb{F}$ 任一 $m\times n$ 阶 λ-矩阵 $A(\lambda)$ 都可以写成 $A(\lambda)=\lambda^kA_k+\cdots+\lambda A_1+A_0$ 的形式, 其中 $A_0,A_1,\ldots,A_k\in\mathbb{F}^{m\times n},k\in\mathbb{N}$.

证明. $A(\lambda)$ 的每个元素都是 $\mathbb{F}$ 上的多项式, 记 $A(\lambda)$ 的第 (i,j) 元 $a_{ij}(\lambda)=\sum\limits_{\ell=0}^{k}a_{ij\ell}\lambda^\ell$, 其中 $k=\max\limits_{i,j}\deg a_{ij}(\lambda)$. 再对 $\ell=0,1,\ldots,k$ 取矩阵 $A_\ell=(a_{ij\ell})_{i,j}\in\mathbb{F}^{m\times n}$, 则 $A(\lambda)=\sum\limits_{\ell=0}^{k}A_\ell\lambda^\ell$. □

6.4.5 证明: 任意满秩 λ-方阵 $A(\lambda)$ 都可以写成 $P(\lambda),Q(\lambda)$ 的乘积, 其中 $P(\lambda)$ 为可逆 λ-方阵,$Q(\lambda)$ 是上三角方阵, 其对角元均为首一多项式, 对角线以上的元素的次数都小于同列对角元的次数.

证明. 命题等价于证明通过初等行变换将任意满秩 λ-方阵 $A(\lambda)$ 变换为上三角方阵 $Q(\lambda)$, 且 $Q(\lambda)$ 的元素满足 $\deg q_{ij}(\lambda)<q_{jj}(\lambda),1\leq i<j$. 下设 $a_{11}(\lambda)$ 为第一列次数最低的非零多项式, 因为总能用初等行变换做到这一点.

首先证明, 通过初等行变换能将 $A(\lambda)$ 变换为

$$B(\lambda)=\begin{pmatrix}b_{11}(\lambda) & b_{12}(\lambda) & \cdots & b_{1n}(\lambda)\\ 0 & b_{22}(\lambda) & \cdots & b_{2n}(\lambda)\\ \vdots & \vdots & & \vdots\\ 0 & b_{n2}(\lambda) & \cdots & b_{nn}(\lambda)\end{pmatrix}$$

的形式. 若 $a_{11}(\lambda) \mid a_{i1}(\lambda)(2 \le i \le n)$ 则结论显然成立, 否则对 $\deg a_{11}(\lambda)$ 归纳.

$\deg a_{11}(\lambda) = 0$ 时总有 $a_{11}(\lambda) \mid a_{i1}(\lambda)$, 从而结论成立, 下设 $\deg a_{11}(\lambda) \le k-1$ 时结论成立. 若 $\deg a_{11}(\lambda) = k$, 对每个 $2 \le i \le n$ 作带余除法 $a_{i1}(\lambda) = q_{i1}(\lambda)a_{11}(\lambda) + r_{i1}(\lambda)$, 有 $\deg r_{i1}(\lambda) < \deg a_{11}(\lambda) = k$, 因此可作初等行变换将第 $(i,1)$ 元变为 $r_{i1}(\lambda)$. 最后将次数最小的非零 $r_{i1}(\lambda)$ 通过行变换换到第 $(1,1)$ 元, 其满足归纳假设, 结论成立.

依次对 $B(\lambda)$ 的右下角子矩阵作上述操作, 从而可变为

$$C(\lambda) = \begin{pmatrix} c_{11}(\lambda) & c_{12}(\lambda) & \cdots & c_{1n}(\lambda) \\ & c_{22}(\lambda) & \cdots & c_{2n}(\lambda) \\ & & \ddots & \vdots \\ & & & c_{nn}(\lambda) \end{pmatrix}$$

再将对角线上元素用同列对角元除, 可得到余式, 因此可依次作初等行变换使得对角线上元素变为余式, 其次数总小于同列对角元, 从而可得到 $Q(\lambda)$. □

6.5 行列式因子、不变因子与初等因子

6.5.1 求下列 λ-矩阵的不变因子和初等因子.

$$(1)\begin{pmatrix} 1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ \lambda^2+1 & \lambda^2 & -\lambda^2 \end{pmatrix}, \qquad (2)\begin{pmatrix} \lambda-4 & -10 & 19 & -4 \\ -1 & \lambda-6 & 8 & -3 \\ -1 & -4 & \lambda+6 & -2 \\ 0 & 1 & -1 & \lambda \end{pmatrix}.$$

证明. (1) 不变因子为 $1, \lambda, \lambda(\lambda+1)$, 故初等因子为 $\lambda; \lambda, \lambda+1$.

(2) 不变因子为 $1, 1, (\lambda-1)^2, (\lambda-1)^2$, 故初等因子为 $(\lambda-1)^2; (\lambda-1)^2$. □

6.5.2 $f, g \in \mathbb{F}[\lambda], (f,g) = 1$, 证明下列 λ-矩阵等价:

$$\begin{pmatrix} f(\lambda) & 0 \\ 0 & g(\lambda) \end{pmatrix}, \qquad \begin{pmatrix} g(\lambda) & 0 \\ 0 & f(\lambda) \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ 0 & f(\lambda)g(\lambda) \end{pmatrix}.$$

证明. 通过互换行列容易看出前两个矩阵等价, 下证明第一个矩阵与第三个矩阵等价. 由 $(f,g) = 1$ 知存在多项式 u, v 使得 $uf + vg = 1$, 故作如下初等变换:

$$\begin{pmatrix} f & 0 \\ 0 & g \end{pmatrix} \longrightarrow \begin{pmatrix} f & uf \\ 0 & g \end{pmatrix} \longrightarrow \begin{pmatrix} f & uf+vg \\ 0 & g \end{pmatrix} = \begin{pmatrix} f & 1 \\ 0 & g \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & f \\ g & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & f \\ 0 & -fg \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ 0 & fg \end{pmatrix}$$

从而三个矩阵等价. □

6.5.3 $A(\lambda)$ 为满秩 12 阶 λ-矩阵, 若其初等因子为 $(\lambda-1)^2, (\lambda-1)^2, (\lambda-1)^2, (\lambda+1), (\lambda+1), (\lambda-\mathrm{i})^2, (\lambda+\mathrm{i})^2$, 求 $A(\lambda)$ 的不变因子和行列式因子.

证明. 将初等因子排序为:

$$\begin{matrix} (\lambda-1)^2 & \lambda+1 & (\lambda-\mathrm{i})^2 & (\lambda+\mathrm{i})^2 \\ (\lambda-1)^2 & \lambda+1 & & \\ (\lambda-1)^2 & & & \end{matrix}$$

从而不变因子为 $d_{12} = (\lambda-1)^2(\lambda+1)(\lambda^2+1)^2, d_{11} = (\lambda-1)^2(\lambda+1), d_{10} = (\lambda-1)^2, d_9 = \cdots = d_1 = 1$. 行列式因子为 $D_1 = \cdots = D_9 = 1, D_{10} = (\lambda-1)^2, D_{11} = (\lambda-1)^4(\lambda+1), D_{12} = (\lambda+1)^6(\lambda-1)^2(\lambda^2+1)^2$. □

6.5.4 证明

$$A(\lambda)=\begin{pmatrix}\lambda & 0 & \cdots & 0 & a_n\\ -1 & \lambda & \cdots & 0 & a_{n-1}\\ & -1 & \ddots & \vdots & \vdots\\ & & \ddots & \lambda & a_2\\ & & & -1 & \lambda+a_1\end{pmatrix}$$

的不变因子为 $1,\dots,1,f(\lambda)$, 其中 $f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$.

证明. 由于 $\det A(\lambda)=f(\lambda)$, 且 $A(\lambda)$ 的第 $2,3,\dots,n$ 行与第 $1,2,\dots,n-1$ 列构成的子式 $=(-1)^{n-1}$, 从而由行列式因子的定义知 $D_{n-1}=\cdots=D_1=1, D_n=f(\lambda)$, 从而可得不变因子为 $1,\dots,1,f(\lambda)$. □

6.5.5 $A(\lambda)$ 为 n 阶 λ-矩阵, 证明 $A(\lambda)$ 与 $A^{\mathsf{T}}(\lambda)$ 等价.

证明. 由于 $A(\lambda)$ 与 $A^{\mathsf{T}}(\lambda)$ 的行列式因子等价, 因此两者的不变因子等价, 即两者等价. □

6.6 Jordan 标准形

6.6.1 设

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

求其 Jordan 标准型 J 及可逆方阵 P, 使 $P^{-1}AP=J$.

证明.

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

□

6.6.2 $A\in M_n(\mathbb{F})$, 证明 A 与 A^{T} 相似.

证明. 由题 6.5.5 知 $\lambda I-A$ 与 $\lambda I-A^{\mathsf{T}}=(\lambda I-A)^{\mathsf{T}}$ 等价, 因此 A 与 A^{T} 相似. □

6.6.3 设方阵 A 的非常数不变因子为 $(\lambda-1),(\lambda-1)(\lambda+1),(\lambda-1)^2(\lambda+1)^2$, 求 A 的 Jordan 标准型 J.

证明. 其初等因子为 $\lambda-1;\lambda-1,\lambda+1;(\lambda-1)^2,(\lambda+1)^2$, 故可得 $J=\operatorname{diag}\left(1,1,\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix},-1,\begin{pmatrix}-1 & 1\\ 0 & -1\end{pmatrix}\right)$. □

6.6.4 证明方阵 A 可对角化 $\iff$ 对于任意 A 的特征值 λ 都有 $\operatorname{rank}(\lambda I-A)^2=\operatorname{rank}(\lambda I-A)$.

证明一. 该等式即等价于

$$\dim V_\lambda=n-\operatorname{rank}(\lambda I-A)=n-\operatorname{rank}(\lambda I-A)^2=\dim\ker(\lambda I-A)^2=\dim W_\lambda^{(2)}$$

又由 $V_\lambda\subset W_\lambda^{(2)}$ 知 $V_\lambda=W_\lambda^{(2)}$. 若 A 可对角化, 即 $V_\lambda=W_\lambda$, 故有 $V_\lambda=W_\lambda^{(2)}$, 从而必要性得证.

充分性: 若有 $V_\lambda=W_\lambda^{(2)}$, 则 $\forall v\in W_\lambda^{(3)},(\lambda I-A)v\in W_\lambda^{(2)}=V_\lambda$, 故 $(\lambda I-A)^2v=0$, 即 $v\in W_\lambda^{(2)}$, 因此 $W_\lambda^{(2)}=W_\lambda^{(3)}$. 以此类推有 $V_\lambda=W_\lambda^{(2)}=W_\lambda^{(3)}=\cdots=W_\lambda^{(n)}=W_\lambda$, 由 λ 任意性知 A 可对角化. □

证明二. $\Longrightarrow$: 设 A 相似于对角阵 $\operatorname{diag}(\lambda_1 I_{r_1},\dots,\lambda_k I_{r_k})$, 任取特征值 λ_i, 有相似关系

$$\lambda_i I_n - A \sim \operatorname{diag}((\lambda_i-\lambda_1)I_{r_1},\dots,(\lambda_i-\lambda_k)I_{r_k}),$$
$$(\lambda_i I_n - A)^2 \sim \operatorname{diag}((\lambda_i-\lambda_1)^2 I_{r_1},\dots,(\lambda_i-\lambda_k)^2 I_{r_k}),$$

两者的第 i 个对角分块矩阵均为 O_{r_i}. 又由于 $i\neq j$ 时 $\lambda_i-\lambda_j\neq 0$, 故 $\operatorname{rank}(\lambda_i I_n - A)=\operatorname{rank}(\lambda_i I_n - A)^2 = n-r_i$, 从而得证.

$\Longleftarrow$: 设 A 的 Jordan 标准型

$$J=\operatorname{diag}(J(\lambda_1,m_{11}),\dots,J(\lambda_1,m_{1t_1}),J(\lambda_2,m_{21}),\dots,J(\lambda_s,m_{st_s}))$$

其中 $\displaystyle\sum_{j=1}^{t_i} m_{ij}=r_i$ 为特征值 λ_i 的重数. 任取特征值 λ_k, 则有相似关系

$$\lambda_k I_n - A \sim \operatorname{diag}(J(\lambda_k-\lambda_1,m_{11}),\dots,J(\lambda_k-\lambda_s,m_{st_s})),$$
$$(\lambda_k I_n - A)^2 \sim \operatorname{diag}(J(\lambda_k-\lambda_1,m_{11})^2,\dots,J(\lambda_k-\lambda_s,m_{st_s})^2),$$

而对于每个 Jordan 块 $J_{ij}=J(\lambda_k-\lambda_i,m_{ij})$, 由 $i\neq k$ 时 $\lambda_i\neq\lambda_k$ 知 $\operatorname{rank} J_{ij}^2=\operatorname{rank} J_{ij}$. 而 $i=k$ 时 $J_{ij}=J(0,m_{ij})$, 若 $m_{kj}\geq 2$ 则 $\operatorname{rank} J(0,m_{ij})^2=\operatorname{rank} J(0,m_{ij})-1$. 因此由题设 $\operatorname{rank}(\lambda_k I-A)^2=\operatorname{rank}(\lambda_k I - A)$ 知,$m_{kj}=1$. 又由 k 的任意性知所有 Jordan 块的尺寸 $m_{ij}=1$, 即 J 为对角阵, 即 A 可对角化. □

6.6.5 若方阵 A 的特征值全为 0, 则 A 是幂零矩阵.

证明. A 的 Jordan 标准型 $J=\operatorname{diag}(J(0,s_1),\dots,J(0,s_k))$, 从而取 $m=\max\limits_{1\leq i\leq k} s_k$, $J^m=O$, 因此存在可逆矩阵 P 使得 $A=PJP^{-1}$, $A^m=PJ^mP^{-1}=O$, 即 A 是幂零矩阵. □

6.6.6 设 λ 是 n 阶矩阵 A 的 k 重特征值, 证明 $\operatorname{rank}(\lambda I-A)^k=n-k$.

证明一. 即 $\dim W_\lambda=\dim\ker(\lambda I-A)^k=k$, 从而 $\operatorname{rank}(\lambda I-A)^k=n-k$. □

证明二. 设 A 的 Jordan 标准型 $J=\operatorname{diag}(J(\lambda,n_1),\dots,J(\lambda,n_t),J(\lambda_1,n_{11}),\dots,J(\lambda_s,n_{sm_s}))$, 其中 $\displaystyle\sum_{i=1}^{t} n_i=k, 1\leq n_i\leq k$, 因此 $J(0,n_i)^m=O$, 从而 $(\lambda I-A)^k$ 相似于 $\operatorname{diag}(O_k,J(\lambda-\lambda_1,n_{11})^k,\dots,J(\lambda-\lambda_s,n_{sm_s})^k)$, 即 $\operatorname{rank}(\lambda I-A)^k=n-k$. □

复习题 6

6.1 设 n 维线性空间 V 的线性变换 $\mathscr{A}$ 有 n 个互不相同的特征值, 证明: 线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 可交换 $\Longleftrightarrow$ $\mathscr{B}$ 是 $\mathscr{I},\mathscr{A},\dots,\mathscr{A}^{n-1}$ 的线性组合.

证明. 若 $\mathscr{B}$ 可被表为 $\mathscr{A}$ 的多项式, 则自然可与 $\mathscr{A}$ 交换. 反之, 由题设知 $\mathscr{A}$ 可对角化, 则设 $\mathscr{A}$ 在 V 的某基下的矩阵为 $A=\operatorname{diag}(\lambda_1,\dots,\lambda_n)$, 在此基下 $\mathscr{B}$ 的矩阵为 B. 由 $AB=BA$ 可解得 $b_{ij}=0(i\neq j)$, 故可设 $B=\operatorname{diag}(b_1,\dots,b_n)$. 考虑方程 $B=x_0I+x_1A+\cdots+x_{n-1}A^{n-1}$, 其等价于线性方程组

$$\begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}=\begin{pmatrix}1 & \lambda_1 & \cdots & \lambda_1^{n-1}\\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & \lambda_n & \cdots & \lambda_n^{n-1}\end{pmatrix}\begin{pmatrix} x_0\\ x_1\\ \vdots\\ x_{n-1}\end{pmatrix}$$

其系数矩阵的行列式为 Vandermonde 行列式 $\displaystyle\prod_{1\leq i<j\leq n}(\lambda_i-\lambda_j)\neq 0$ (由 A 的特征值互不相同), 因此方程有唯一解, 从而 $\mathscr{B}$ 可表为 $\mathscr{I},\mathscr{A},\dots,\mathscr{A}^{n-1}$ 的线性组合. □

6.2 $A\in\mathbb{F}^{m\times n},B\in\mathbb{F}^{n\times m}$, 证明 $\phi_{AB}(\lambda)$ 与 $\phi_{BA}(\lambda)$ 差一个 λ^{n-m}.

证明. 在有理函数域 $\mathbb{F}(\lambda)$ 上有

$$\det(\lambda I_m-AB)=\lambda^m\det(I_m-AB/\lambda)=\lambda^m\det(I_n-BA/\lambda)=\lambda^{m-n}\det(\lambda I_n-BA),$$

从而两多项式差一个 λ^{n-m}. 或者考虑 $\lambda\neq 0$ 时上式依然成立, 而 $\lambda=0$ 时 $\phi_{AB}(0)=\det(-AB)=\det(-BA)=\phi_{BA}(0)$, 综上得证. □

6.3 $\alpha=(a_1,\dots,a_n),\beta=(b_1,\dots,b_n)$, 求 $I+\alpha^{\mathsf T}\beta$ 的特征值.

证明. 其特征多项式为 $\det(\lambda I-I-\alpha^{\mathsf T}\beta)=(\lambda-1)^n\det\left(I-\dfrac{\alpha^{\mathsf T}\beta}{\lambda-1}\right)=(\lambda-1)^n\det\left(I_1-\dfrac{\beta\alpha^{\mathsf T}}{\lambda-1}\right)=(\lambda-1)^{n-1}\big(\lambda-1-\beta\alpha^{\mathsf T}\big)$, 因此其特征值为 $1(n-1$ 重) 和 $1+\beta\alpha^{\mathsf T}=1+\displaystyle\sum_{i=1}^n a_ib_i$. □

6.4 A,B 是复方阵, 记 $C=AB-BA$, 若 $AC=CA$, 证明 C 是幂零矩阵.

证明. C 幂零即所有特征值为 0, 为此仅需证明 $\forall k\geq 1,\operatorname{tr}C^k=\displaystyle\sum_{\lambda\in\operatorname{Spec}C}\lambda^k=0$. 任取 $k\geq 0$, 有

$$AC^k=CAC^{k-1}=C^2AC^{k-2}=\cdots=C^{k-1}AC=C^kA$$

因此 $\operatorname{tr}C^{k+1}=\operatorname{tr}\Big(C^kAB-C^kBA\Big)=\operatorname{tr}\Big(A(C^kB)\Big)-\operatorname{tr}\Big((C^kB)A\Big)=0$, 从而得证. □

6.5 设 $B=\begin{pmatrix}O&A\\A&O\end{pmatrix}$, 其中对称矩阵 $A\in M_n(\mathbb{R})$ 有特征值 $\lambda_1,\dots,\lambda_n$, 求 B 的特征值.

证明.

$$\phi_B(\lambda)=\begin{vmatrix}\lambda I_n&-A\\-A&\lambda I_n\end{vmatrix}=\begin{vmatrix}\lambda_n I_n&-A\\O&\lambda I_n-A^2/\lambda\end{vmatrix}=\lambda^n\det(\lambda I_n-A^2/\lambda)=\det(\lambda^2I_n-A^2)$$

而 A^2 特征值为 $\lambda_1^2,\dots,\lambda_n^2$, 故 $\phi_B(\lambda)=\displaystyle\prod_{i=1}^n(\lambda^2-\lambda_i^2)$, 因此 B 的特征值为 $\pm\lambda_1,\dots,\pm\lambda_n$. □

6.6 若 $\mathscr{A}^m=\mathscr{I}(m\geq 2)$, 证明 $\mathscr{A}$ 可对角化.

证明. $\mathscr{A}$ 零化多项式有 $f(\lambda)=\lambda^m-1$, 其无重根, 故极小多项式无重根, 即可对角化. □

6.7 $A\in M_n(\mathbb{R})$ 的特征多项式为 $f(\lambda)$, 记 $g=f/(f,f')$, 证明 A 可对角化 $\iff g(A)=O$.

证明. 记 A 的极小多项式为 $m(\lambda)$. 设 $f(\lambda)$ 在 $\mathbb{C}$ 上有不可约分解 $f(\lambda)=\displaystyle\prod_{i=1}^s(\lambda-\lambda_i)^{r_i}$, 则 $g(\lambda)=\displaystyle\prod_{i=1}^s(\lambda-\lambda_i)$, 显然 $g(\lambda)\mid m(\lambda)$. 因此 A 可对角化 $\iff m(A)$ 无重根 $\iff m(\lambda)=g(\lambda),g(A)=O$. □

6.8 $A, B \in M_n(\mathbb{C})$ 可交换, 证明存在 n 阶可逆矩阵 Q 使得 $Q^{-1}AQ, Q^{-1}BQ$ 均为上三角矩阵.

证明. 对 n 归纳,$n=1$ 时自然成立, 下设 $<n$ 时命题已成立. 视 A, B 为 $\mathbb{C}^n$ 上线性变换 $\mathscr{A}, \mathscr{B}$ 在标准基下的矩阵, 故可由题 6.2.1 知 A, B 有公共特征向量 v_1. 扩充其为 $\mathbb{C}^n$ 上的一组基. 从而在此基下 $\mathscr{A}, \mathscr{B}$ 的矩阵为

$$A_1 = \begin{pmatrix} \lambda & * \\ & A' \end{pmatrix}, B_1 = \begin{pmatrix} \mu & * \\ & B' \end{pmatrix}$$

其中 λ, μ 分别为 A, B 的特征值. 从而存在可逆矩阵 Q_1 使得 $Q_1^{-1}AQ_1 = A_1, Q_1^{-1}BQ_1 = B_1$, 且由 $AB = BA$ 知 $A'B' = B'A'$. 对 A', B' 运用归纳假设, 则存在 $n-1$ 阶可逆矩阵 Q_2 使 $Q_2^{-1}A'Q_2, Q_2^{-1}B'Q_2$ 均为上三角矩阵, 从而可取 $Q = Q_1 \operatorname{diag}(1, Q_2)$, 使得

$$Q^{-1}AQ = \begin{pmatrix} \lambda & * \\ & Q_2^{-1}A'Q_2 \end{pmatrix}, Q^{-1}BQ = \begin{pmatrix} \lambda & * \\ & Q_2^{-1}B'Q_2 \end{pmatrix}$$

均为上三角矩阵. □

注 6. *该命题的推广为 Lie 定理, 其在 $M_n(\mathbb{C})$(或者说,$\mathfrak{gl}(n, \mathbb{C})$) 上的一个特例是: 若矩阵 A, B 生成的 Lie 代数 $\mathfrak{g} = \operatorname{span}(A, B, [A, B], [A, [A, B]], \dots)$ 可解 (即导出列终于 0), 则其中元素可同时上三角化. 在该题中 $[A, B] = O$, 则 $\mathfrak{g}^{(1)} = [\mathfrak{g}, \mathfrak{g}] = \operatorname{span}\{\}[X, Y] | X, Y \in \mathfrak{g} = 0$, 故导出列 $\mathfrak{g} > \mathfrak{g}^{(1)} = 0$, 从而 $\mathfrak{g} = \operatorname{span}(A, B)$ 中元素均可同时上三角化.*

6.9 $A, B \in M_n(\mathbb{F}), \operatorname{rank} A + \operatorname{rank} B < n$, 证明两者有公共特征向量.

证明. 由题知 $\dim \ker A + \dim \ker B > n$, 故有 $v \in \ker A \cap \ker B, v \neq 0$, 其为 A, B 关于特征值 0 的特征向量. □

6.10 设 $\mathscr{A}$ 为 n 维线性空间 V 上的线性变换,$\mathscr{A}$ 有 n 个不同的特征值, 证明 $\mathscr{A}$ 有 2^n 个不变子空间.

证明. 由题知 $\mathscr{A}$ 有 n 个线性无关特征向量 $v_1, \dots, v_n$, 从而其任意子集 S 张成的线性子空间均为 $\mathscr{A}$ 的不变子空间, 共有 2^n 个. 而对于任意 $\mathscr{A}$ 的不变子空间 U, 由准素分解知

$$U = U \cap V = U \cap \left(\bigoplus_{i=1}^{n} V_{\lambda_i} \right) = \bigoplus_{i=1}^{n} (U \cap V_{\lambda_i})$$

故 U 有基底 $u_1, \dots, u_k$, 其中每个向量都在 $U \cap V_\lambda$ 中, 故 U 也有特征向量张成, 即在上述 2^n 个不变子空间中, 从而得证. □

注 7. *可改条件为 A 可对角化, 证明同上.*

6.11 A, B 是 n 阶矩阵,A 有 n 个不同特征值, 证明以下三者等价:

1. $AB = BA$;
2. 存在可逆矩阵 P 使得 $P^{-1}AP, P^{-1}BP$ 均为对角矩阵;
3. A, B 有相同的 n 个线性无关特征向量.

证明. (1) $\Longrightarrow$ (2): 证明同题 6.8, 对 n 归纳. □