第二章　线性算子

引理 4.7 设 $\mathcal{A} \in \mathcal{L}(V)$, U_1, U_2 是非平凡 $\mathcal{A}$-子空间, 且 $V = U_1 \oplus U_2$. 设 $\epsilon_1, \ldots, \epsilon_{d_1}$ 是 U_1 的基, $\delta_1, \ldots, \delta_{d_2}$ 是 U_2 的基. 则在 V 的基底 $\epsilon_1, \ldots, \epsilon_{d_1}, \delta_1, \ldots, \delta_{d_2}$ 下 $\mathcal{A}$ 的矩阵是

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

其中 $A_i \in \mathrm{M}_{d_i}(F)$ 是 $\mathcal{A}_{U_i}$ 在对应基下的矩阵, $i = 1, 2$. 进而 $\mu_{\mathcal{A}} = \mathrm{lcm}(\mu_{\mathcal{A}_{U_1}}, \mu_{\mathcal{A}_{U_2}})$ (取首一的最小公倍式).

证明. 注意到 $V = U_1 \oplus U_2$ 蕴含 $d_1 + d_2 = n (= \dim(V))$ 且 $\epsilon_1, \ldots, \epsilon_{d_1}, \delta_1, \ldots, \delta_{d_2}$ 线性无关. 所以 $\epsilon_1, \ldots, \epsilon_{d_1}, \delta_1, \ldots, \delta_{d_2}$ 是 V 的一组基. 对 $i \in \{1, 2, \ldots, d_1\}$, $\mathcal{A}(\epsilon_i) \in U_1$, $\mathcal{A}(\epsilon_i)$ 是 $\epsilon_1, \ldots, \epsilon_{d_1}$ 的线性组合, 它关于 $\delta_1, \ldots, \delta_{d_2}$ 的坐标都是零. 于是, 存在 $A_1 \in \mathrm{M}_{d_1}(F)$ 使得

$$(\mathcal{A}(\epsilon_1), \ldots, \mathcal{A}(\epsilon_{d_1})) = (\epsilon_1, \ldots, \epsilon_{d_1}) A_1.$$

类似地, 存在 $A_2 \in \mathrm{M}_{d_1}(F)$ 使得

$$(\mathcal{A}(\delta_1), \ldots, \mathcal{A}(\delta_{d_1})) = (\delta_1, \ldots, \delta_{d_2}) A_2.$$

于是 A_i 是 $\mathcal{A}_{U_i}$ 在对应基底下的矩阵, $i = 1, 2$. 进而, $\mathcal{A}$ 在 V 的基底 $\epsilon_1, \ldots, \epsilon_{d_1}, \delta_1, \ldots, \delta_{d_2}$ 下的矩阵等于 A.

设 $p=\operatorname{lcm}(\mu_{\mathcal{A}_{U_1}},\mu_{\mathcal{A}_{U_2}})$. 由上一讲命题 4.3,

$$\mu_{\mathcal{A}_{U_1}}|\mu_{\mathcal{A}} \quad 和 \quad \mu_{\mathcal{A}_{U_2}}|\mu_{\mathcal{A}}.$$

于是 $p|\mu_{\mathcal{A}}$. 又因为 $\mu_{\mathcal{A}_{U_1}}|p, \mu_{\mathcal{A}_{U_2}}|p$, 所以 $p(\mathcal{A}_{U_1})=\mathcal{O}$ 和 $p(\mathcal{A}_{U_2})=\mathcal{O}$. 于是

$$p(A)=\begin{pmatrix} p(A_1) & O \\ O & p(A_2) \end{pmatrix}=\begin{pmatrix} O & O \\ O & O \end{pmatrix}.$$

由此和上一讲引理 3.5, $\mu_{\mathcal{A}}|p$. 再由首一性可得 $p=\mu_{\mathcal{A}}$. □

例 4.8 设

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

计算 μ_A.

解. 由上述引理

$$\mu_A=\operatorname{lcm}(\mu_{(1)},\mu_{(0)})=\operatorname{lcm}(t-1,t)=(t-1)t. \quad \square$$

例 4.9 设 $\mathcal{A}\in\mathcal{L}(V)$ 满足 $\ker(\mathcal{A})\oplus\operatorname{im}(\mathcal{A})=V$. 设 $\mathbf{e}_1,\ldots,\mathbf{e}_r$ 是 $\operatorname{im}(\mathcal{A})$ 的一组基, $\mathbf{e}_{r+1},\ldots,\mathbf{e}_n$ 是 $\ker(\mathcal{A})$ 的一组基. 则 $\mathbf{e}_1,\ldots,\mathbf{e}_r,\mathbf{e}_{r+1},\ldots,\mathbf{e}_n$ 是 V 的一组基. 因为 $\operatorname{im}(\mathcal{A})$ 和 $\ker(\mathcal{A})$ 都是 $\mathcal{A}$-子空间, 且 $\mathcal{A}(\mathbf{e}_j)=\mathbf{0}$, $j=r+1,r+2,\ldots,n$, 所以 $\mathcal{A}$ 在该基底下的矩阵是

$$A=\begin{pmatrix} B & O \\ O & O \end{pmatrix},$$

其中 $B\in \mathrm{M}_r(F)$ 满秩. 当 $r=n$ 时, $B=A$. 否则, $\mu_A=\operatorname{lcm}(\mu_B,t)$.

定理 4.10 设 $\mathcal{A}\in\mathcal{L}(V)$, $U_1,\dots,U_k$ 是非平凡 $\mathcal{A}$-子空间满足 $V=U_1\oplus\cdots\oplus U_k$. 设 Z_i 是 U_i 的一组基, $i=1,\dots,k$. 则 $\mathcal{A}$ 在 V 的基底 $Z_1\cup\cdots\cup Z_k$ 下的矩阵

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

其中 A_i 是 $\mathcal{A}_{U_i}$ 在 Z_i 下的矩阵, $i=1,2,\dots,k$. 进而,

$$\mu_{\mathcal{A}}=\operatorname{lcm}(\mu_{\mathcal{A}_{U_1}},\dots,\mu_{\mathcal{A}_{U_k}}).$$

证明. 对 k 归纳. 当 $k=1$ 时, 定理显然成立. 设 $k>1$ 且 $k-1$ 时定理成立. 设 $W=U_1\oplus\cdots\oplus U_{k-1}$. 则 $V=W\oplus U_k$, $Y=Z_1\cup\cdots\cup Z_{k-1}$ 是 W 的基. 由引理 4.7, $\mathcal{A}$ 在基底 $W\cup Z_k$ 下的矩阵是

$$A=\begin{pmatrix} B & O \\ O & A_k \end{pmatrix},$$

其中 B 是 $\mathcal{A}_W$ 在 Y 下的矩阵, A_k 是 $\mathcal{A}_{U_k}$ 在 Z_k 下的矩阵, 且 $\mu_{\mathcal{A}}=\operatorname{lcm}(\mu_{\mathcal{A}_W},\mu_{A_{U_k}})$.

对 $\mathcal{A}_W, W, U_1, \ldots, U_{k-1}$ 用归纳假设得

$$B = \begin{pmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A_{k-1} \end{pmatrix},$$

其中 A_i 是 $\mathcal{A}_{U_i}$ 在 Z_i 下的矩阵, $i = 1, 2, \ldots, k-1$. 进而, $\mu_{\mathcal{A}_W} = \mathrm{lcm}(\mu_{\mathcal{A}_{U_1}}, \ldots, \mu_{\mathcal{A}_{U_{k-1}}})$. 于是, A 是所要求的形式. 注意到

$$\begin{aligned} &\mathrm{lcm}\left(\mu_{\mathcal{A}_{U_1}}, \ldots, \mu_{\mathcal{A}_{U_k}}\right) \\ &= \mathrm{lcm}\left(\mathrm{lcm}\left(\mu_{\mathcal{A}_{U_1}}, \ldots, \mu_{\mathcal{A}_{U_{k-1}}}\right), \mu_{\mathcal{A}_{U_k}}\right) \\ &= \mathrm{lcm}(\mu_{\mathcal{A}_W}, \mu_{\mathcal{A}_{U_k}}) = \mu_{\mathcal{A}}. \quad \square \end{aligned}$$

5 不可分子空间

定义 5.1 设 $\mathcal{A} \in \mathcal{L}(V)$, U 是 $\mathcal{A}$-子空间. 如果 U 不能写成两个非零的 $\mathcal{A}$-子空间的直和, 则称 U 是 $\mathcal{A}$-不可分的.

命题 5.2 设 $\mathcal{A} \in \mathcal{L}(V)$. 则 V 是有限个 $\mathcal{A}$-不可分子空间的直和.

证明. 设 $n = \dim(V)$. 我们对 n 归纳. 当 $n = 1$ 时, V 本身是 $\mathcal{A}$ 不可分的. 定理显然成立. 设 $n > 1$ 且当空间维

数小于 n 时定理成立. 如果 V 是 $\mathcal{A}$ 不可分的, 则定理成立. 否则存在两个非零 $\mathcal{A}$-子空间 U, W 使得 $V = U \oplus W$. 则 $\dim(U)$ 和 $\dim(W)$ 的维数都小于 n. 由归纳假设, $U = U_1 \oplus \cdots \oplus U_k$, 其中 U_i 是 A_U 不可分的, 从而也是 $\mathcal{A}$ 不可分的. 同样, $W = W_1 \oplus \cdots \oplus W_\ell$, 其中 W_j 是 $\mathcal{A}_W$ 不可分的, 从而也是 $\mathcal{A}$ 不可分的. 于是

$$V = U \oplus W = U_1 \oplus \cdots \oplus U_k \oplus W_1 \oplus \cdots \oplus W_\ell. \quad \square$$

6 核核分解

定理 6.1 设 $\mathcal{A}{\in}\mathrm{Hom}(V,V)$, $f{\in}F[x]$ 且 $f(\mathcal{A}){=}\mathcal{O}$. 如果

$$f = pq$$

其中 $p, q \in F[x]$ 互素, 则

$$V = \ker(p(\mathcal{A})) \oplus \ker(q(\mathcal{A})).$$

证明. 见本学期第一讲定理 5.17. □

定理 6.2 设 $\mathcal{A}{\in}\mathrm{Hom}(V,V)$, $f{\in}F[x]$ 且 $f(\mathcal{A}){=}\mathcal{O}$. 如果

$$f = p_1 \cdots p_m,$$

其中 $p_1, \ldots p_m \in F[x]$ 两两互素, 则

$$V = K_1 \oplus \cdots \oplus K_m,$$

其中 $K_i = \ker(p_i(\mathcal{A}))$, $i = 1, \ldots, m$.

证明. 对 m 归纳. 当 $m=1$ 时结论是平凡的. 设 $m>1$ 且结论对 $m-1$ 成立. 令 $q=p_1\cdots p_{m-1}$. 因为 $\gcd(p_i,p_m)=1$, $i=1,\ldots,m$, 所以 $\gcd(q,p_m)=1$. 根据上述定理,

$$V=W\oplus K_m,\quad \text{其中 } W=\ker(q(\mathcal{A})). \tag{1}$$

设 $\mathcal{B}=\mathcal{A}|_W$. 对任意 $\mathbf{w}\in W$, $\mathcal{B}(\mathbf{w})=\mathcal{A}(\mathbf{w})$. 故

$$q(\mathcal{A})(\mathcal{B}(\mathbf{w}))=q(\mathcal{A})\mathcal{A}(\mathbf{w})=\mathcal{A}q(\mathcal{A})(\mathbf{w})=\mathcal{A}(\mathbf{0})=\mathbf{0}.$$

于是, $\mathcal{B}(\mathbf{w})\in W$. 即 $\mathcal{B}\in\mathrm{Hom}(W,W)$. 对任意 $\mathbf{w}\in W$,

$$q(\mathcal{B})(\mathbf{w})=q(\mathcal{A})(\mathbf{w})=\mathbf{0}.$$

由此可知, $q(\mathcal{B})=\mathcal{O}$. 根据归纳假设, 我们有

$$W=\ker(p_1(\mathcal{B}))\oplus\cdots\oplus\ker(p_{m-1}(\mathcal{B})). \tag{2}$$

下面验证:

$$\ker(p_i(\mathcal{B}))=K_i,\quad i=1,\ldots,m-1. \tag{3}$$

设 $\mathbf{w}\in\ker(p_i(\mathcal{B}))$. 则 $\mathbf{0}=p_i(\mathcal{B})(\mathbf{w})=p_i(\mathcal{A})(\mathbf{w})$. 故 $\mathbf{w}\in K_i$. 故 $\ker(p_i(\mathcal{B}))\subset K_i$. 设 $\mathbf{v}\in K_i$. 则 $q(\mathcal{A})(\mathbf{v})=\mathbf{0}$. 故 $\mathbf{v}\in W$. 由此得出, $\mathbf{0}=p_i(\mathcal{A})(\mathbf{v})=p_i(\mathcal{B})(\mathbf{v})$. 从而, $\mathbf{v}\in\ker(p_i(\mathcal{B}))$. 故 $K_i\subset\ker(p_i(\mathcal{B}))$. 等式 (3) 成立.

根据等式 (1), (2) 和 (3), 定理成立. □

定理 6.3 *(核核分解定理之极小多项式版)* 设 $\mathcal{A} \in \mathcal{L}(V)$, $\mu_{\mathcal{A}} = p_1^{m_1} \cdots p_s^{m_s}$, 其中 $p_1, \ldots, p_s \in F[t] \setminus F$, 不可约且两两互素, $m_1, \ldots, m_s \in \mathbb{Z}^+$. 令

$$K_i = \ker(p_i^{m_i}(\mathcal{A})), \quad i = 1, 2, \ldots, s.$$

则

$$V = K_1 \oplus \cdots \oplus K_s,$$

且 $\mathcal{A}|_{K_i}$ 的极小多项式是 $p_i^{m_i}$, $i = 1, \ldots, s$.

证明. 上述定理蕴含 $V = K_1 \oplus \cdots \oplus K_s$.

设 $\mathcal{A}_i = \mathcal{A}|_{K_i}$. 因为对任意 $\mathbf{v} \in K_i$, $p_i^{m_i}(\mathbf{v}_i) = \mathbf{0}$. 所以 $\mathcal{A}_i$ 的极小多项式 μ_i 整除 $p_i^{m_i}$ (引理 3.2). 因为 p_i 不可约, 所以 $\mu_i = p_i^{k_i}$, 其中 $1 \le k_i \le m_i$. 由定理 4.10 可知,

$$\mu_{\mathcal{A}} = \operatorname{lcm}\left(p_1^{k_1}, \ldots, p_s^{k_s}\right).$$

因为 $p_1, \ldots, p_s$ 两两互素, 所以 $\mu_{\mathcal{A}} = p_1^{k_1} \cdots p_s^{k_s}$. 由多项式不可约分解的唯一性得出 $k_i = m_i$, $i = 1, \ldots, s$.

例 6.4 设 $\mathcal{A}$ 是数乘算子. 则 $\mu_{\mathcal{A}}(t) = t - \lambda$, 其中 $\lambda \in F$. 故 $V = \ker(\mathcal{A} - \lambda\mathcal{E})$. 故 $\mathcal{A} = \lambda\mathcal{E}$. 于是, $\mathcal{A}$ 在任何基底下的矩阵都是 λE.

例 6.5 设 $\mathcal{A}$ 是幂等算子, $\mathcal{A} \ne \mathcal{O}$, $\mathcal{A} \ne \mathcal{E}$. 则 $\mu_{\mathcal{A}}(t) = t^2 - t = t(t-1)$. 故 $V = \ker(\mathcal{A}) \oplus \ker(\mathcal{A} - \mathcal{E})$. 设 $\ker(\mathcal{A})$ 的

一组基是 $\mathbf{e}_1, \ldots, \mathbf{e}_d$, $\ker(\mathcal{A}-\mathcal{E})$ 的一组基是 $\mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$. 则 $\mathcal{A}(\mathbf{e}_i) = \mathbf{0}$, $i = 1, \ldots, d$, $\mathcal{A}(\mathbf{e}_j) = \mathbf{e}_j$, $j = d+1, \ldots, n$. 故 $\mathcal{A}$ 在 V 的基底下的矩阵是

$$\begin{pmatrix} O & O \\ O & E_{n-d} \end{pmatrix}.$$

例 6.6 设 $\mathcal{A} \in \mathcal{L}(V)$ 是幂零算子. 则 $\mu_{\mathcal{A}}(t) = t^k$, 其中 k 是某个正整数. 则 $V = \ker(\mathcal{A}^k)$. 不提供任何关于 $\mathcal{A}$ 的矩阵表示的信息.

定义 6.7 设 $\mathcal{A}\in\mathcal{L}(V), \mathbf{v}\in V, f(t)\in F[t]$. 如果 $f(\mathcal{A})(\mathbf{v})=\mathbf{0}$, 则称 $f(t)$ 是通过 $\mathcal{A}$ 零化 $\mathbf{v}$ 的多项式. 非零、次数最小的通过 $\mathcal{A}$ 零化 $\mathbf{v}$ 的多项式称为通过 $\mathcal{A}$ 零化 $\mathbf{v}$ 的极小多项式. 该极小多项式记为 $\mu_{\mathcal{A},\mathbf{v}}$, 它通常是首一的.

注意到 $\mu_{\mathcal{A}}(\mathcal{A})(\mathbf{v}) = \mathcal{O}(\mathbf{v}) = \mathbf{0}$. 于是, $\mu_{\mathcal{A},\mathbf{v}}$ 存在. 设 $f(\mathcal{A})(\mathbf{v}) = \mathbf{0}$. 由多项式带余除法可知

$$f(t) = q(t)\mu_{\mathcal{A},\mathbf{v}}(t) + r(t),$$

其中 $q, r \in F[t], \deg(r) < \deg(\mu_{\mathcal{A},\mathbf{v}})$. 带入 $\mathcal{A}$ 得

$$\mathcal{O} = q(\mathcal{A})\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A}) + r(\mathcal{A}).$$

两侧同时作用在 $\mathbf{v}$ 上得到

$$\mathbf{0} = q(\mathcal{A})\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v}) + r(\mathcal{A})(\mathbf{v}) = \mathbf{0} + r(\mathcal{A})(\mathbf{v}).$$

于是, $r(\mathcal{A})(\mathbf{v})=\mathbf{0}$. 因为 $\deg(r)<\deg(\mu_{\mathcal{A},\mathbf{v}})$, 所以 $r(t)=0$. 由此得出 $\mu_{\mathcal{A},\mathbf{v}}|f$. 特别地, $\mu_{\mathcal{A},\mathbf{v}}|\mu_{\mathcal{A}}$.

命题 6.8 *(科斯特利金第二卷第56页习题 9)* 设 $\mathcal{A}\in\mathcal{L}(V)$. 则 存在 $\mathbf{v}\in V$ 使得 $\mu_{\mathcal{A},\mathbf{v}}=\mu_{\mathcal{A}}$.

证明. 先设 $\mu_{\mathcal{A}}=p^k$, 其中 $p\in F[t]$ 不可约和首一. 因为 $\mu_{\mathcal{A},\mathbf{v}}|\mu_{A}$ 且 p 不可约, 所以 $\mu_{\mathcal{A},\mathbf{v}}=p^{m_{\mathbf{v}}}$, 其中 $1\le m_{\mathbf{v}}\le k$. 假设不存在 $\mathbf{v}$ 使得 $m_{\mathbf{v}}=k$. 则对任意 $\mathbf{v}\in V$, $m_{\mathbf{v}}\le k-1$. 于是 $p^{k-1}=q_{\mathbf{v}}\mu_{\mathcal{A},\mathbf{v}}$, 其中 $q_{\mathbf{v}}\in F[t]$. 我们有

$$
\begin{aligned}
p^{k-1}(\mathcal{A})= \quad & q_{\mathbf{v}}(\mathcal{A})\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A}) \\
\Rightarrow \quad & p^{k-1}(\mathcal{A})(\mathbf{v})=q_{\mathbf{v}}(\mathcal{A})\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v}) \\
& =q_{\mathbf{v}}(\mathcal{A})\left(\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v})\right)=\mathbf{0}.
\end{aligned}
$$

由 $\mathbf{v}$ 的任意性得出 $p^{k-1}(\mathcal{A})=\mathcal{O}$. 矛盾. 故当 $\mu_{\mathcal{A}}=p^k$ 时, 结论成立.

下面考虑一般情形. 设 $\mu_{\mathcal{A}}=p_1^{m_1}\cdots p_s^{m_s}$, 其中 $p_1,\ldots,p_s\in F[t]\setminus F$, 不可约且两两互素, $m_1,\ldots,m_s\in\mathbb{Z}^+$. 令

$$
K_i=\ker(p_i^{m_i}(\mathcal{A})),\mathcal{A}_i=\mathcal{A}|_{k_i},\mu_i=\mu_{\mathcal{A}_i},\quad i=1,2,\ldots,s.
$$

由定理 6.3,

$$
V=K_1\oplus\cdots\oplus K_s
$$

且 $\mu_i=p_i^{m_i}$. 由上述证明可知存在 $\mathbf{v}_i\in K_i$ 使得 $\mu_{\mathcal{A}_i,\mathbf{v}_i}=\mu_i$, $i=1,2,\ldots,s$.

令 $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_s$. 则,

$$\mathbf{0} = \mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v}) = \mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v}_1) + \cdots + \mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v}_s).$$

因为 $V = K_1 \oplus \cdots \oplus K_s$, 且每个 K_i 都是 $\mathcal{A}$ 不变的, 所以 $\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v}_i) \in K_i$. 由直和的基本性质(见第一章第一讲定理 1.11 (ii)), $\mu_{\mathcal{A},\mathbf{v}}(\mathcal{A})(\mathbf{v}_i) = \mathbf{0}$. 于是, $\mu_{\mathcal{A}_i,\mathbf{v}_i} | \mu_{\mathcal{A},\mathbf{v}}$. 由此可知, $\mu_{\mathcal{A}_i} | \mu_{\mathcal{A},\mathbf{v}}$, $i = 1, 2, \ldots, s$. 从而 $\mu_{\mathcal{A}} = \mathrm{lcm}(\mu_{\mathcal{A}_1}, \ldots, \mu_{\mathcal{A}_s}) | \mu_{\mathcal{A},\mathbf{v}}$. 又因为 $\mu_{\mathcal{A},\mathbf{v}} | \mu_{\mathcal{A}}$. 我们有 $\mu_{\mathcal{A}} = \mu_{\mathcal{A},\mathbf{v}}$. □

例 6.9 *设 $\mathcal{A}$ 是数乘算子. 则 $\mu_{\mathcal{A}}(t) = t - \lambda$, 其中 $\lambda \in F$. 设 $\mathbf{v} \in V$ 且 $\mathbf{v} \neq \mathbf{0}$. 则 $\mu_{\mathcal{A},\mathbf{v}}(t) = t - \lambda$.*

例 6.10 *设 $\mathcal{A}$ 是幂等算子, $\mathcal{A} \neq \mathcal{O}$, $\mathcal{A} \neq \mathcal{E}$. 则 $\mu_{\mathcal{A}}(t) = t^2 - t = t(t-1)$. 设 $\mathbf{v} = \mathbf{x} + \mathbf{y}$, 其中 $\mathbf{x} \in \ker(\mathcal{A})$, $\mathbf{y} \in \ker(\mathcal{A} - \mathcal{E})$ (见例 6.5) 再设 $\mathbf{x} \neq \mathbf{0}$ 且 $\mathbf{y} \neq \mathbf{0}$. 则 $\mathcal{A}(\mathbf{v}) \neq \mathbf{0}$ 且 $(\mathcal{A} - \mathcal{E})(\mathbf{v}) \neq \mathbf{0}$. 故 $\mu_{\mathcal{A},\mathbf{v}}(t) = t(t-1)$.*

例 6.11 *设 $\mathcal{A} \in \mathcal{L}(V)$ 是幂零算子. 则 $\mu_{\mathcal{A}}(t) = t^k$, 其中 k 是某个正整数. 设 $\mathbf{v} \notin \ker(\mathcal{A}^{k-1}$. 则 $\mu_{\mathcal{A},\mathbf{v}} = t^k$.*

7 特征向量和特征值

在本节中, V 是域 F 上的有限维线性空间且 $\dim(V)$>0.

7.1 特征向量

定义 7.1 *设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. 如果 $\langle \mathbf{v} \rangle$ 是 $\mathcal{A}$ 子空间, 则称 $\mathbf{v}$ 是 $\mathcal{A}$ 的一个特征向量 (eigenvector).*

命题 7.2 *设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. 则下列结论等价:*

(i) $\mathbf{v}$ 是 $\mathcal{A}$ 的特征向量;

(ii) $\mathcal{A}(\mathbf{v}) \in \langle \mathbf{v} \rangle$;

(iii) 存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$.

证明. (i) $\Longrightarrow$ (ii) 显然.

(ii) $\Longrightarrow$ (iii) 因为 $\mathcal{A}(\mathbf{v}) \in \langle \mathbf{v} \rangle$, 所以存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$.

(iii) $\Longrightarrow$ (i) 设 $\mathbf{x} \in \langle \mathbf{v} \rangle$. 则存在 $\alpha \in F$ 使得 $\mathbf{x} = \alpha \mathbf{v}$. 于是,

$$\mathcal{A}(\mathbf{x}) = \mathcal{A}(\alpha \mathbf{v}) = \alpha \mathcal{A}(\mathbf{v}) = \alpha \lambda \mathbf{v} \in \langle \mathbf{v} \rangle \implies \langle \mathbf{v} \rangle \text{ 是 } \mathcal{A} \text{ 不变的.} \quad \square$$

根据上述命题非零 $\mathbf{v}$ 是 $\mathcal{A}$ 特征向量当且仅当存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$. 我们称 λ 是关于特征向量 $\mathbf{v}$ 的特征值 (eigenvalue). 简称 $\mathcal{A}$ 的特征值. 反之, 设 $\lambda \in F$ 是 $\mathcal{A}$ 的特征值. 令

$$V^{\lambda} = \{\mathbf{x} \in V \mid \mathcal{A}(\mathbf{x}) = \lambda \mathbf{x}\}$$

称为 $\mathcal{A}$ 关于 λ 的特征子空间(eigenspace). 下面我们来验证 V^λ 是 $\mathcal{A}$-子空间.

设 $\alpha, \beta \in F, \mathbf{x}, \mathbf{y} \in V^\lambda$. 则

$$\mathcal{A}(\alpha\mathbf{x}+\beta\mathbf{y}) = \alpha\mathcal{A}(\mathbf{x})+\beta\mathcal{A}(\mathbf{y}) = \alpha\lambda\mathbf{x}+\beta\lambda\mathbf{y} = \lambda(\alpha\mathbf{x}+\beta\mathbf{y}).$$

由此可知 $\alpha\mathbf{x}+\beta\mathbf{y} \in V^\lambda$. 即 V^λ 是子空间. 因为

$$\mathcal{A}(\mathbf{x}) = \lambda\mathbf{x} \in V^\lambda,$$

所以 V^λ 是 $\mathcal{A}$ 不变的.

例 7.3 设 $\mathcal{D}$ 是 $\mathbb{R}[x]^{(n)}$ 中的导数算子. 求 $\mathcal{D}$ 所有特征值和特征向量.

解. 设 $f = f_d x^d + \cdots + f_1 x + f_0$, 其中 $f_d, \ldots, f_1, f_0 \in \mathbb{R}$ 且 $f_d \neq 0$. 如果 $\mathcal{D}(f) = \lambda f$, 其中 $\lambda \in \mathbb{R}$. 则

$$d f_d x^{d-1} + \cdots + f_1 = \lambda(f_d x^d + \cdots + f_1 x + f_0).$$

上式成立当且仅当 $\lambda = 0$ 且 $d = 0$. 于是, f 是 $\mathcal{D}$ 的特征向量当且仅当 $f \in \mathbb{R} \setminus \{0\}$. 这些特征向量对应的特征值都等于 0. 而 $V^0 = \mathbb{R}$. □

7.2 特征多项式

当我们把矩阵 $A \in \mathrm{M}_n(F)$ 看成 $\mathcal{L}(F^n)$ 中由 $A(\mathbf{x}) = A\mathbf{x}$ 定义得线性算子时, 我们同样有矩阵 A 的特征向量, 特征值和特征子空间的概念.

设 $A \in \mathrm{M}_n(F)$, $\mathbf{x} = (x_1, \ldots, x_n)^t \in F^n \setminus \{\mathbf{0}\}$. 则存在 $\lambda \in F$ 使得 $A\mathbf{x} = \lambda\mathbf{x}$, 即 $\mathbf{x}$ 是 A的特征向量, 当且仅当

$$\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \iff (\lambda E - A) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

由此推出 $\mathbf{x}$ 是 $\mathcal{A}$ 的特征向量蕴含 $\det(\lambda E - A) = 0$.

定义 7.4 设 $A \in \mathrm{M}_n(F)$, t 是 F 上的未定元. 多项式

$$\det(tE - A) \in F[t]$$

称为 A 的特征多项式, 记为 $\chi_A(t)$.

定义 7.5 设 $A \in \mathrm{M}_n(F)$. 特征多项式 χ_A 在 F 中的根称为 A 的特征根 *(eigenroots)*. 这些特征根的集合记为 $\mathrm{spec}_F(A)$, 称为 A 在 F 中的谱 *(spectrum)*.

矩阵的特征根就是矩阵的特征值.

例 7.6 设实二阶矩阵

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

求 A 和 B 的所有特征根和特征向量.

解. 直接计算得

$$\chi_A(t) = \det(tE - A) = \det \begin{pmatrix} t & -1 \\ -1 & t \end{pmatrix} = t^2 - 1$$

和

$$\chi_B(t) = \det(tE - B) = \det \begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix} = t^2 + 1.$$

于是, $\mathrm{spec}_{\mathbb{R}}(A) = \{1, -1\}$, $\mathrm{spec}_{\mathbb{R}}(B) = \emptyset$. 从而 B 没有实特征根，从而没有特征向量和特征子空间.

设特征根 $\lambda_1 = 1$. 它对应的特征子空间是方程组

$$(\lambda_1 E - A) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的解空间. 即方程组

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的解空间. 解方程组得 $V^{\lambda_1} = \langle (1,1)^t \rangle$. 类似地, 特征根 $\lambda_2 = -1$ 对应的特征子空间是 $V^{\lambda_2} = \langle (1,-1)^t \rangle$.

例 7.7 设上例中矩阵 B 是复矩阵. 求它的特征值和特征向量.

解. 由上例可知,

$$\chi_B(t) = t^2 + 1.$$

于是, $\mathrm{spec}_{\mathbb{C}}(B) = \{\sqrt{-1}, -\sqrt{-1}\}$. 设特征根 $\lambda_1 = \sqrt{-1}$. 它对应的特征子空间是方程组

$$(\lambda_1 E - A)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的解空间. 即方程组

$$\begin{pmatrix} \sqrt{-1} & 1 \\ -1 & \sqrt{-1} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的解空间. 解方程组得 $V^{\lambda_1} = \langle (1, -\sqrt{-1})^t \rangle$. 类似地, 特征根 $\lambda_2 = -\sqrt{-1}$ 对应的特征子空间是 $V^{\lambda_2} = \langle (1, \sqrt{-1})^t \rangle$.

设 $\chi_A(t) = \det(tE - A) \in F[t]$. 则 $\mathbf{x}$ 是 $\mathcal{A}$ 的特征向量蕴含着它对应的特征值 λ 是 $\chi_A(t)$ 的根. 反之, 设 $\lambda \in F$ 是 $\chi_A(t)$ 的根. 则方程组

$$(\lambda E - A)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

由非零解 $(\alpha_1, \ldots, \alpha_n)$. 于是, $\mathbf{v} = \alpha_1 \mathbf{e}_1 + \cdots + \alpha_n \mathbf{e}_n$ 满足 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$. 由此推出 $\lambda \in F$ 是 $\chi_A(t)$ 的根当且仅当 λ 是 $\mathcal{A}$ 的特征值.

定义 7.8 设 $A \in \mathrm{M}_n(F)$, t 是 F 上的未定元. 多项式

$$\det(tE - A) \in F[t]$$

称为 A 的特征多项式, 记为 $\chi_A(t)$.

命题 7.9 矩阵特征多项式是相似不变量.

证明. 设 $A, B \in \mathrm{M}_n(F)$ 且 $A \sim_s B$. 则存在 $P \in \mathrm{GL}_n(F)$ 使得 $B = P^{-1}AP$. 直接计算得

$$\det(tE - B) = \det(tE - P^{-1}AP) = \det(P^{-1}(tE - A)P) = \det(tE - A).$$

即 $\chi_A = \chi_B$. □

定义 7.10 设 $A \in \mathrm{M}_n(F)$. 特征多项式 χ_A 在 F 中的根称为 A 的特征根 *(eigenroots)*. 这些特征根的集合记为 $\mathrm{spec}_F(A)$, 称为 A 在 F 中的谱 *(spectrum)*.

矩阵的特征根就是矩阵的特征值. 由命题 7.9 可知, $\mathrm{spec}_F(A)$ 也是相似不变量.

命题 7.11 设 V 是复数域 $\mathbb{C}$ 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$. 则 $\mathcal{A}$ 一定有特征向量.

证明. 因为 $\chi_{\mathcal{A}} \in \mathbb{C}[t] \setminus \mathbb{C}$, 所以 $\chi_{\mathcal{A}}$ 在 $\mathbb{C}$ 中至少有一个根 λ (代数学基本定理). 即 $\mathcal{A}$ 有特征根. 于是有特征向量. □

例 7.12 设 $A \in \mathrm{M}_n(\mathbb{C})$. 证明: A 相似于一个上三角矩阵.

证明. 对 n 归纳. 当 $n = 1$ 时, 结论显然成立. 设 $n > 1$ 且 $n - 1$ 时结论成立.

考虑 $A \in \mathrm{M}_n(\mathbb{C})$. 把 A 看成 $\mathbb{C}^n$ 上在标准基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下矩阵等于 A 的线性算子. 有上例可知, A 有一个 1 维 $\mathcal{A}$ 子空间 $\langle \mathbf{u} \rangle$. 根据第二章第二讲命题 *5.3*,

$$A \sim_s \begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & B \end{pmatrix},$$

其中 $\lambda \in \mathbb{C}$, $B \in \mathrm{M}_{n-1}(\mathbb{C})$. 根据归纳假设. 存在 $P \in \mathrm{GL}_{n-1}(\mathbb{C})$ 使得 $P^{-1}BP$ 是上三角的. 令

$$Q = \begin{pmatrix} 1 & O_{1\times n-1} \\ O_{(n-1)\times 1} & P \end{pmatrix}.$$

则 P 可逆且

$$\begin{aligned} & Q^{-1} \begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & B \end{pmatrix} Q \\ = & \begin{pmatrix} 1 & O_{1\times(n-1)} \\ O_{(n-1)\times 1} & P^{-1} \end{pmatrix} \begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & B \end{pmatrix} \begin{pmatrix} 1 & O_{1\times(n-1)} \\ O_{(n-1)\times 1} & P \end{pmatrix} \\ = & \underbrace{\begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & P^{-1}BP \end{pmatrix}}_{T}. \end{aligned}$$

因为 $P^{-1}BP$ 已经是上三角矩阵, 所以 T 也是上三角矩阵. 显然, $A \sim_s T$. $\square$