第二章　线性算子

3　极小多项式

注解 3.1 设 $A, B{\in}\mathrm{M}_n(F)$ 且 $A{\sim_s}B$. 则存在 $P{\in}\mathrm{GL}_n(F)$ 使得 $A = P^{-1}BP$. 则

$$\forall\ k \in \mathbb{N}, A^k = P^{-1}B^kP.$$

进而,

$$\forall\ f \in F[t], f(A) = P^{-1}f(B)P.$$

特别地, $f(A) \sim_s f(B)$.

定义 3.2 设 $f \in F[t]$, $\mathcal{A} \in \mathcal{L}(V)$. 如果 $f(\mathcal{A}) = \mathcal{O}$, 则称 f 是关于 $\mathcal{A}$ 的零化多项式. 类似地, 对 $A \in \mathrm{M}_n(F)$, 我们有关于 A 的零化多项式的概念.

注解 3.3 由线性同构 $\Phi : \mathcal{L}(V) \longrightarrow \mathrm{M}_n(F)$ 可知

$$\dim(\mathcal{L}(V)) = n^2.$$

故 $1, \mathcal{A}, \ldots, \mathcal{A}^{n^2}$ 在 F 上线性相关. 进而, $\mathcal{A}$ 有非零的次数不高于 n^2 的多项零化 A.

定义 3.4 关于 $\mathcal{A}$ 的非零的零化多项式中次数最小的称为 $\mathcal{A}$ 的极小多项式. 类似地, 对 $A \in \mathrm{M}_n(F)$, 我们有关于 A 的极小多项式的概念.

引理 3.5 *(极小多项式的整除判别法)* 设 $\mathcal{A} \in \mathcal{L}(V)$ 和 $p(t) \in F[t] \setminus \{0\}$ 零化 $\mathcal{A}$. 则

$$p \text{ 是 } \mathcal{A} \text{ 的极小多项式} \iff \text{对任意 } f \in F[t] \text{ 零化 } \mathcal{A}, \quad p \mid f.$$

证明. 由多项式除法可知

$$f(t) = q(t)p(t) + r(t),$$

其中 $q, r \in F[t]$ 且 $\deg(r) < \deg(p)$. 由赋值同态定理

$$f(\mathcal{A}) = q(\mathcal{A})p(\mathcal{A}) + r(\mathcal{A}).$$

因为 $p(\mathcal{A}) = \mathcal{O}$, 所以 $f(\mathcal{A}) = r(\mathcal{A})$.

如果 $f(\mathcal{A}) = \mathcal{O}$, 则 $r(\mathcal{A}) = \mathcal{O}$. 由极小多项式的定义可知, $r(t) = 0$. 即 $p \mid f$. 反之, 设 $f(t) \in F[t]$ 是 $\mathcal{A}$ 的极小多项式. 则 $f(\mathcal{A}) = \mathcal{O}$. 故 $p \mid f$. 由极小多项式的定义可知, $\deg(p) = \deg(f)$. 于是, p 也是 $\mathcal{A}$ 的极小多项式. □

由上述判别法可知, $\mathcal{A}$ 的两个极小多项式必然在 $F[t]$ 中相伴. 我们把关于 $\mathcal{A}$ 的首一的极小多项式记为 $\mu_{\mathcal{A}}$. 注意到上述判别法对方阵也成立. 于是, 我们把关于 $A \in \mathrm{M}_n(F)$ 的首一的极小多项式记为 μ_A.

例 3.6 域 F 中的非零元素不可能是任何算子和矩阵的极小多项式.

例 3.7 设 $\mathcal{A}\in\mathcal{L}(V)$. 证明 $\deg(\mu_{\mathcal{A}})=1$ 当且仅当 $\mathcal{A}$ 是数乘算子.

证明. 设 $\mathcal{A}=\lambda\mathcal{E}$, $\lambda\in F$. 则 $\mu_{\mathcal{A}}=t-\lambda$. 反之, 设 $\mu_{\mathcal{A}}=t-\lambda$. 则 $\mathcal{O}=\mu_{\mathcal{A}}(\mathcal{A})=\mathcal{A}-\lambda\mathcal{E}$. 于是, $\mathcal{A}=\lambda\mathcal{E}$. □

特别地, $\mu_{\mathcal{O}}=t$, $\mu_{\mathcal{E}}=t-1$.

类似地, $A\in\mathrm{M}_n(F)$ 是数乘矩阵当且仅当 $\deg\mu_A=1$.

例 3.8 $\mathcal{A}\in\mathcal{L}(V)$ 是幂零算子当且仅当 $\mu_{\mathcal{A}}$ 是 t 的幂次.

证明. 设 $\mathcal{A}^k=\mathcal{O}$. 则 t^k 零化 $\mathcal{A}$. 由引理 3.5, $\mu_{\mathcal{A}}|t^k$. 于是 $\mu_{\mathcal{A}}$ 是 t 的幂次. 另一个方向显然. □

微分算子 $\mathcal{D}:\mathbb{R}[x]^{(n)}\longrightarrow\mathbb{R}[x]^{(n)}$ 的极小多项式是 t^n.

类似地, $A\in\mathrm{M}_n(F)$ 是幂零矩阵当且仅当 $\deg\mu_A=t^k$, 其中 k 是某个正整数.

命题 3.9 设 $A,B\in\mathrm{M}_n(F)$. 如果 $A\sim_s B$, 则 $\mu_A=\mu_B$.

证明. 由注释 3.1 和 $\mu_A(A)=O$ 可知, $\mu_A(B)=O$. 于是 $\mu_B|\mu_A$ (引理 3.5). 同理 $\mu_A|\mu_B$. 因为 μ_A 和 μ_B 都首一, 所以 $\mu_A=\mu_B$. □

例 3.10 设

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

问 A 和 B 是否相似?

解. 注意到 $\mu_A = t-1$. 因为 B 不是数乘矩阵，所以 $\deg(\mu_B) > 1$ (例 3.7). 于是, $\mu_A \neq \mu_B$. 故 $A \not\sim_s B$. $\square$

命题 3.11 设 $\mathcal{A} \in \mathcal{L}(V)$. 则

(i) $\dim(F[\mathcal{A}]) = \deg(\mu_{\mathcal{A}})$,

(ii) $\mathcal{A}$ 可逆当且仅当 $\mu_{\mathcal{A}}(0) \neq 0$.

证明. 设 $d = \deg_t(\mu_{\mathcal{A}})$. 我们来证明 $\mathcal{E}, \mathcal{A}, \ldots, \mathcal{A}^{d-1}$ 是 $F[\mathcal{A}]$ 的一组基. 设 $\alpha_0, \alpha_1, \ldots, \alpha_{d-1} \in F$ 使得

$$\alpha_0\mathcal{E} + \alpha_1\mathcal{A} + \cdots + \alpha_{d-1}\mathcal{A}^{d-1} = \mathcal{O}.$$

令 $p(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_{d-1}t^{d-1} \in F[t]$. 则 $p(\mathcal{A}) = \mathcal{O}$. 因为 $\deg_t(p) < d$, 所以 $p = 0$. 于是, $\alpha_0 = \alpha_1 = \cdots = \alpha_{d-1} = 0$. 我们推出 $\mathcal{E}, \mathcal{A}, \ldots, \mathcal{A}^{d-1}$ 线性无关.

设 $G \in F[\mathcal{A}]$. 则存在 $g \in F[t]$ 使得 $G = g(\mathcal{A})$. 由多项式带余除法可知, 存在 $q, r \in F[t]$, $\deg_t(r) < d$ 使得

$$g(t) = q(t)\mu_{\mathcal{A}}(t) + r(t).$$

于是

$$G = g(\mathcal{A}) = q(\mathcal{A})\mu_{\mathcal{A}}(\mathcal{A}) + r(\mathcal{A}) = r(\mathcal{A}).$$

即 G 是 $\mathcal{E}, \mathcal{A}, \ldots, \mathcal{A}^{d-1}$ 在 F 上的线性组合. 于是 $\mathcal{E}, \mathcal{A}, \ldots, \mathcal{A}^{d-1}$ 是 $F[\mathcal{A}]$ 的一组基. 特别地, $\dim(F[\mathcal{A}]) = d$.

设 $\mu_{\mathcal{A}} = \beta_0 + \beta_1 t + \cdots + \beta_{d-1} t^{d-1} + t^d$, 其中 $\beta_0, \beta_1, \ldots, \beta_{d-1} \in F$. 则

$$\mathcal{O} = \beta_0 \mathcal{E} + \beta_1 \mathcal{A} + \cdots + \beta_{d-1} \mathcal{A}^{d-1} + \mathcal{A}^d.$$

如果 $\mu_{\mathcal{A}}(0) \neq 0$, 则 $\beta_0 \neq 0$. 于是

$$\mathcal{A} \underbrace{(-\beta_1 \mathcal{E} - \cdots + \beta_{d-1} \mathcal{A}^{d-2} - \mathcal{A}^{d-1}) \beta_0^{-1}}_{\mathcal{A}^{-1}} = \mathcal{E}. \tag{1}$$

即 $\mathcal{A}$ 可逆. 设 $\mathcal{A}$ 可逆. 如果 $\mu_{\mathcal{A}}(0) = 0$, 则 $\beta_0 = 0$. 于是

$$\mu_{\mathcal{A}}(t) = t(\beta_1 + \beta_2 t + \cdots + \beta_{n-1} t^{n-2} + t^{n-1}).$$

于是

$$\mathcal{O} = \mathcal{A}(\beta_1 \mathcal{E} + \beta_2 \mathcal{A} + \cdots + \beta_{d-1} \mathcal{A}^{d-2} + \mathcal{A}^{d-1}).$$

把上述等式两边同乘以 $\mathcal{A}^{-1}$. 则

$$\mathcal{O} = \beta_1 \mathcal{E} + \beta_2 \mathcal{A} + \cdots + \beta_{d-1} \mathcal{A}^{d-2} + \mathcal{A}^{d-1}.$$

我们看到非零多项式 $\beta_1 + \beta_2 t + \cdots + \beta_{d-1} t^{d-2} + t^{d-1}$ 零化 $\mathcal{A}$. 矛盾. $\square$

注解 3.12 由 (1) 可知, 当 $\mathcal{A}$ 可逆时, $\mathcal{A}^{-1} \in F[\mathcal{A}]$.

4 不变子空间

定义 4.1 设 $\mathcal{A} \in \mathcal{L}(V)$, U 是 V 的子空间. 如果 $\mathcal{A}(U) \subset U$, 即 $\forall \mathbf{u} \in U, \mathcal{A}(\mathbf{u}) \in U$, 则称 U 是 $\mathcal{A}$ 的不变子空间.

设 U 是 $\mathcal{A}$ 的不变子空间. 则 $A|_U$ 可以看做 U 上的线性算子. 为简明起见, 记限制映射 $A|_U$ 为 $\mathcal{A}_U$. 注意到

$$\mathcal{A}_U \in \mathcal{L}(U).$$

例 4.2 设 $\mathcal{D}$ 是 $\mathbb{R}[x]^{(n)}$ 上的导数算子. 则 $\mathbb{R}[x]^{(k)}$ 是 $\mathcal{D}$ 的不变子空间, $k = 1, 2, \ldots, n$. 但 $\langle x^k \rangle$ 不是, $k = 0, 1, \ldots, n-1$.

设 $\lambda \in F$, 则 V 的每个子空间都是关于 $\lambda\mathcal{E}$ 的不变的.

命题 4.3 设 $\mathcal{A} \in \mathcal{L}(V)$, U 是 $\mathcal{A}$ 的 d 维不变子空间, $0 < d < n$. 则存在 V 的一组基使得 $\mathcal{A}$ 在该基下的矩阵为

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

其中 $B \in \mathrm{M}_d(F)$ 是 $\mathcal{A}_U$ 的某个矩阵表示. 进而 $\mu_{\mathcal{A}_U} | \mu_{\mathcal{A}}$, $\mu_B | \mu_A$, $\mu_D | \mu_A$.

证明. 设 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 是 U 的一组基. 把它扩充为 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$. 因为 U 是 $\mathcal{A}$ 的不变子空间, 所以当 $j \in \{1, 2, \ldots, d\}$ 时, $\mathcal{A}(\mathbf{e}_j)$ 是 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 的线性组合, 即 $\mathcal{A}(\mathbf{e}_j)$ 关于 $\mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$ 的坐标都等于零. 于是 $\mathcal{A}$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$ 下的矩阵如命题所述形式, 且 B 是 $\mathcal{A}_U$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 下的矩阵.

直接计算可验证对任意 $k \in \mathbb{N}$

$$A^k = \begin{pmatrix} B^k & * \\ O & D^k \end{pmatrix},$$

其中 $*$ 是某个 $d\times(n-d)$ 阶的矩阵. 于是, 对任意 $f\in F[t]$.

$$f(A)=\begin{pmatrix} f(B) & * \\ O & f(D) \end{pmatrix}.$$

因为 $\mu_A(A)=O_{n\times n}$, 所以 $\mu_A(B)=O_{d\times d}$, $\mu_A(D)=O_{(n-d)\times(n-d)}$. 由引理 3.5, $\mu_B|\mu_A$, $\mu_D|\mu_A$, 且 $\mu_{\mathcal{A}_U}|\mu_{\mathcal{A}}$. $\square$