第二章　线性算子

注解 3.2 设 $\mathcal{A} \in \mathcal{L}(V)$. 则 $V = \ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$ 当且仅当 $\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$ 是直和.

证明. 设 $V = \ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$. 则

$$\begin{aligned}\dim(V) &= \dim(\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})) \\ &= \dim(\ker(\mathcal{A})) + \dim(\mathrm{im}(\mathcal{A})) - \dim(\ker(\mathcal{A}) \cap \mathrm{im}(\mathcal{A})) \\ &\quad \textit{(维数公式)} \\ &= \dim(V) - \dim(\ker(\mathcal{A}) \cap \mathrm{im}(\mathcal{A})) \quad \textit{(对偶定理)} \\ &\Rightarrow \dim(\ker(\mathcal{A}) \cap \mathrm{im}(\mathcal{A})) = 0.\end{aligned}$$

故 $\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$ 是直和.

设 $\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$ 是直和. 则

$$\begin{aligned}\dim(\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})) &= \dim(\ker(\mathcal{A})) + \dim(\mathrm{im}(\mathcal{A})) \\ &\quad \textit{(第一章命题 4.13)} \\ &= \dim(V). \quad \textit{(对偶定理)}\end{aligned}$$

因为 $\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A}) \subset V$, 所以 $\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A}) = V$. □

例 3.3 设 $\mathcal{A} \in \mathcal{L}(V)$ 满足 $\mathcal{A}^2 = \mathcal{A}$. 证明

$$\ker(\mathcal{A}) \oplus \mathrm{im}(\mathcal{A}) = V.$$

证明. 因为 $\mathcal{A}^2 = \mathcal{A}$, 所以 $\mathrm{rank}(\mathcal{A}^2) = \mathrm{rank}(\mathcal{A})$. 由上述核像分解定理可知结论成立. □

例 3.4 设 $\mathcal{D}$ 是 $\mathbb{R}[x]^{(n)}$ 上的导数算子. 则 $\ker(\mathcal{D}) = \mathbb{R}$ 且 $\mathrm{im}(\mathcal{D}) = \mathbb{R}[x]^{(n-1)}$. 因为 $\mathbb{R} \subset \mathbb{R}[x]^{(n-1)}$, 所以 $\ker(\mathcal{D}) + \mathrm{im}(\mathcal{D})$ 不是直和.

4 极小多项式

注解 4.1 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 4.2 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. 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. 为明确起见, 我们设极小多项式是首一的.

类似地, 对 $A \in \mathrm{M}_n(F)$, 我们有关于 A 的零化多项式和极小多项式的概念.

引理 4.3 *(极小多项式的整除判别法)* 设 $\mathcal{A} \in \mathcal{L}(V)$, $f(t) \in F[t]$, $p(t) \in F[t] \setminus \{0\}$ 零化 $\mathcal{A}$ 且首一. 则

$$p = \mu_{\mathcal{A}} \iff \text{对任意 } f \in F[t] \text{ 零化 } \mathcal{A},\quad p|f.$$

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 $p|f$. 反之, 因为 $\mu_{\mathcal{A}}$ 零化 $\mathcal{A}$, 所以 $p|\mu_{\mathcal{A}}$. 由此可知, $\deg(p) \leq \mu_{\mathcal{A}}$. 根据极小多项式的定义, $\deg(p) \geq \deg(\mu_{\mathcal{A}})$. 再利用首一性, $\mu_{\mathcal{A}} = p$. □

命题 4.4 *设 $\mathcal{A} \in \mathcal{L}(V)$. 则 $\mathcal{A}$ 的极小多项式存在且唯一. 极小多项式的次数不大于 n^2.*

证明. 因为 $\dim(\mathcal{L}(V)) = n^2$, 所以 $1, \mathcal{A}, \ldots, \mathcal{A}^{n^2}$ 在 F 上线性相关. 由此可知, $\mathcal{A}$ 有非零的次数不高于 n^2 的零化多项式. 于是, 极小多项式存在且次数不高于 n^2. 设 p, q 是 $\mathcal{A}$ 的两个极小多项式. 则 $\deg(p) = \deg(q)$. 由引理 4.3, $p|q$ 且 $q|p$. 于是 $p = cq$, 其中 $c \in F \setminus \{0\}$. 因为 p 和 q 都首一, 所以 $c = 1$. □

注解 4.5 *以上结论对 $A \in \mathrm{M}_n(F)$ 同样成立.*

记号. 设 $\mathcal{A} \in \mathcal{L}(V)$, $A \in \mathrm{M}_n(F)$. 它们的极小多项式分别记为 $\mu_{\mathcal{A}}$ 和 μ_A.

命题 4.6 *设 $\mathcal{A} \in \mathcal{L}(V)$ 且 $A \in \mathrm{M}_n(F)$ 是 $\mathcal{A}$ 的某个矩阵表示. 则 $\mu_{\mathcal{A}} = \mu_A$.*

证明. 设 $\Phi : F[\mathcal{A}] \longrightarrow F[A]$ 是线性同构和环同构, 其中 $\Phi(\mathcal{A}) = A$ (见定理 2.5). 则对任意 $f \in F[t]$,

$$\Phi(f(\mathcal{A})) = f(\Phi(\mathcal{A})) = f(A) \quad \text{且} \quad \Phi^{-1}(f(A)) = f(\Phi^{-1}(A)) = f(\mathcal{A}).$$

故 $f(\mathcal{A}) = \mathcal{O}$ 当且仅当 $f(A) = O$. 于是, $\mu_{\mathcal{A}}(A) = O$ 且 $\mu_A(\mathcal{A}) = \mathcal{O}$. 根据引理 4.3,

$$\mu_A(t)|\mu_{\mathcal{A}}(t) \quad \text{且} \quad \mu_{\mathcal{A}}(t)|\mu_A(t).$$

再由 $\mu_A(t)$ 和 $\mu_{\mathcal{A}}(t)$ 都首一得出 $\mu_A(t) = \mu_{\mathcal{A}}(t)$. □

例 4.7 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, $\mu_{\mathcal{O}} = t$, $\mu_{\mathcal{E}} = t - 1$.

例 4.8 设 $\mathcal{A} \in \mathcal{L}(V)$ 是幂零算子. 证明 $\mu_{\mathcal{A}}$ 是 t 的幂次.

证明. 设 $\mathcal{A}^k = \mathcal{O}$. 则 t^k 零化 $\mathcal{A}$. 由引理 *4.3*, $\mu_{\mathcal{A}}|t^k$. 于是 $\mu_{\mathcal{A}}$ 是 t 的幂次. □

命题 4.9 设 $A, B \in \mathrm{M}_n(F)$. 如果 $A \sim_s B$, 则 $\mu_A = \mu_B$.

证明. 由注释 4.1 和 $\mu_A(A) = O$ 可知, $\mu_A(B) = O$. 于是 $\mu_B|\mu_A$ (引理 4.3). 同理 $\mu_A|\mu_B$. 因为 μ_A 和 μ_B 都首一, 所以 $\mu_A = \mu_B$. □

例 4.10 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 4.7). 于是, $\mu_A \neq \mu_B$. 故 $A \nsim_s B$. □

命题 4.11 设 $\mathcal{A} \in \mathcal{L}(V)$. 则 $\dim(F[\mathcal{A}]) = \deg(\mu_{\mathcal{A}})$ 且 $\mathcal{A}$ 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

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

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. □

注解 4.12 由 (1) 可知, 当 $\mathcal{A}$ 可逆时, $\mathcal{A}^{-1} \in F[\mathcal{A}]$.

5 不变子空间

定义 5.1 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 5.2 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 5.3 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 4.3, $\mu_B|\mu_A$, $\mu_D|\mu_A$, 且 $\mu_{\mathcal{A}_U}|\mu_{\mathcal{A}}$. □

给定 $\mathcal{A} \in \mathcal{L}(V)$, $\{\mathbf{0}\}$ 和 V 是 $\mathcal{A}$ 的平凡的不变子空间. 下面的引理指出如何寻找 $\mathcal{A}$ 的非平凡的子空间.

引理 5.4 *设 $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$ 满足 $\mathcal{AB} = \mathcal{BA}$. 则 $\ker(\mathcal{B})$ 和 $\mathrm{im}(\mathcal{B})$ 是 $\mathcal{A}$ 的不变子空间.*

证明. 设 $\mathbf{x} \in \ker(\mathcal{B})$. 则

$$\mathcal{B}(\mathcal{A}(\mathbf{x})) = (\mathcal{BA})(\mathbf{x}) = (\mathcal{AB})(\mathbf{x}) = \mathcal{A}(\mathcal{B}(\mathbf{x})) = \mathcal{A}(\mathbf{0}) = \mathbf{0}.$$

于是 $\mathcal{A}(\mathbf{x}) \in \ker(\mathcal{B})$. 即 $\ker(\mathcal{B})$ 是 $\mathcal{A}$ 不变的. 设 $\mathbf{x} \in \mathrm{im}(\mathcal{B})$. 则存在 $\mathbf{y} \in V$ 使得 $\mathbf{x} = \mathcal{B}(\mathbf{y})$. 于是

$$\mathcal{A}(\mathbf{x}) = \mathcal{A}(\mathcal{B}(\mathbf{y})) = \mathcal{B}(\mathcal{A}(\mathbf{y})) \in \mathrm{im}(\mathcal{B}). \quad \square$$

命题 5.5 设 $\mathcal{A} \in \mathcal{L}(V)$, $f \in F[t]$. 则 $\ker(f(\mathcal{A}))$ 和 $\mathrm{im}(f(\mathcal{A}))$ 都是 $\mathcal{A}$ 的不变子空间.

证明. 因为 $\mathcal{A}f(\mathcal{A}) = f(\mathcal{A})\mathcal{A}$, 所以 $\ker(f(\mathcal{A}))$ 和 $\mathrm{im}(f(\mathcal{A}))$ 都是 $\mathcal{A}$ 的不变子空间(引理 5.4). □

为了简单起见, 当 U 是 $\mathcal{A}$ 的不变子空间时, 我们说 U 是 $\mathcal{A}$-不变的或许 $\mathcal{A}$-子空间.

命题 5.6 设 $\mathcal{A} \in \mathcal{L}(V)$, U_1, U_2 是 $\mathcal{A}$-子空间. 则 $U_1 + U_2$ 和 $U_1 \cap U_2$ 都是 $\mathcal{A}$-子空间.

证明. 设 $\mathbf{x} \in U_1 + U_2$. 则存在 $\mathbf{x}_1 \in U_1$, $\mathbf{x}_2 \in U_2$ 使得 $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$. 于是,

$$\mathcal{A}(\mathbf{x}) = \mathcal{A}(\mathbf{x}_1) + \mathcal{A}(\mathbf{x}_2) \in U_1 + U_2.$$

设 $\mathbf{x} \in U_1 \cap U_2$, 则 $\mathcal{A}(\mathbf{x}) \in U_1$ 且 $\mathcal{A}(\mathbf{x}) \in U_2$. 由此可知, $\mathcal{A}(\mathbf{x}) \in U_1 \cap U_2$. □

引理 5.7 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 $p = \mathrm{lcm}(\mu_{\mathcal{A}_{U_1}}, \mu_{\mathcal{A}_{U_2}})$. 由引理 5.3, $\mu_{\mathcal{A}_{U_1}} | \mu_{\mathcal{A}}, \mu_{\mathcal{A}_{U_2}} | \mu_{\mathcal{A}}$. 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 $p(\mathcal{A}_{U_2}) = \mathcal{O}$ (引理 4.3). 于是

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

由此和引理 4.3, $\mu_{\mathcal{A}} | p$. 再利用首一性得出 $p = \mu_{\mathcal{A}}$. □

例 5.8 设

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

计算 μ_A.

解. 由上述引理 $\mu_A = \mathrm{lcm}(\mu_{(1)}, \mu_{(0)}) = \mathrm{lcm}(t-1, t) = (t-1)t$. □

例 5.9 设 $\mathcal{A} \in \mathcal{L}(V)$ 满足 $\ker(\mathcal{A}) \oplus \mathrm{im}(\mathcal{A}) = V$. 设 $\mathbf{e}_1, \ldots, \mathbf{e}_r$ 是 $\mathrm{im}(\mathcal{A})$ 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 $\mathrm{im}(\mathcal{A})$ 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, $\mu_A = \mathrm{lcm}(\mu_B, t)$.

定理 5.10 设 $\mathcal{A} \in \mathcal{L}(V)$, $U_1, \ldots, U_k$ 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, $i = 1, \ldots, k$. 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, $i = 1, 2, \ldots, k$. 进而, $\mu_{\mathcal{A}} = \mathrm{lcm}(\mu_{\mathcal{A}_{U_1}}, \ldots, \mu_{\mathcal{A}_{U_k}})$.

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 5.7, $\mathcal{A}$ 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, $\mu_{\mathcal{A}_W}=\operatorname{lcm}(\mu_{\mathcal{A}_{U_1}},\ldots,\mu_{\mathcal{A}_{U_{k-1}}})$. 于是, A 是所要求的形式. 注意到

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

定理 5.11 *(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. 因为 $p_1^{m_1}, \ldots, p_s^{m_s}$ 两个互素, 所以第一章第二讲定理 2.13(扩展的核核分解定理)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 $1 \leq k_i \leq m_i$. 由定理 5.10 可知,

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

根据第二周讲义命题 5.26 及其证明可知,

$$\mu_{\mathcal{A}} = p_1^{k_1} \cdots p_s^{k_s}.$$

由多项式不可约分解的唯一性得出 $k_i = m_i$, $i = 1, \ldots, s$.

定义 5.12 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 5.13 *(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 5.11,

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

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 $\mu_{\mathcal{A}} = \mu_{\mathcal{A},\mathbf{v}}$. □