

第二章 线性算子

例 3.7 设 $\mathcal{A} \in \mathcal{L}(V)$ 是幂等算子. 证明 $\mu_{\mathcal{A}}$ 或者等于 \mathcal{O} , 或者等于 \mathcal{E} , 或者 $\mu_{\mathcal{A}} = t^2 - t$.

证明. 设 $f(t) = t^2 - t$. 则 $f(\mathcal{A}) = \mathcal{A}^2 - \mathcal{A} = \mathcal{O}$. 故 $\mu_{\mathcal{A}}(t) | t^2 - t$ (\because 引理 3.3). 故 $\mu_{\mathcal{A}}(t) = t$ 或 $\mu_{\mathcal{A}}(t) = t - 1$ 或 $\mu_{\mathcal{A}}(t) = t^2 - t$. 于是, $\mu_{\mathcal{A}}$ 或者等于 \mathcal{O} , 或者等于 \mathcal{E} , 或者 $\mu_{\mathcal{A}} = t^2 - t$. \square

例 3.8 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 3.3, $\mu_{\mathcal{A}} | t^k$. 于是 $\mu_{\mathcal{A}}$ 是 t 的幂次. \square

命题 3.9 设 $\mathcal{A} \in \mathcal{L}(V)$ 且 $A \in M_n(F)$ 是 \mathcal{A} 的某个矩阵表示. 则 $\mu_{\mathcal{A}} = \mu_A$.

证明. 设 $\Phi: \mathcal{L}(\mathcal{A}) \rightarrow M_n(F)$ 是线性同构和环同构, 其中 $\Phi(\mathcal{A}) = A$ (见定理 2.6). 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 $f(A) = \mathcal{O}$. 于是, $\mu_{\mathcal{A}}(A) = \mathcal{O}$ 且 $\mu_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$. 根据引理 3.3,

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

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

命题 3.10 设 $A, B \in M_n(F)$. 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 3.3). 同理 $\mu_A | \mu_B$. 因为 μ_A 和 μ_B 都首一, 所以 $\mu_A = \mu_B$. \square

例 3.11 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 3.6). 于是, $\mu_A \neq \mu_B$. 故 $A \not\sim_s B$. \square

命题 3.12 (i) 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 $\mu_{\mathcal{A}}(0) \neq 0$.

(ii) 设 $A \in M_n(F)$. 则 $\dim(F[A]) = \deg(\mu_A)$ 且 A 可逆当且仅当 $\mu_A(0) \neq 0$.

证明. (i) 设 $d = \deg_t(\mu_{\mathcal{A}})$. 我们来证明 $\mathcal{E}, \mathcal{A}, \dots, \mathcal{A}^{d-1}$ 是 $F[\mathcal{A}]$ 的一组基.

设 $\alpha_0, \alpha_1, \dots, \alpha_{d-1} \in F$ 使得

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

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 $\mathcal{E}, \mathcal{A}, \dots, \mathcal{A}^{d-1}$ 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 $\mathcal{E}, \mathcal{A}, \dots, \mathcal{A}^{d-1}$ 在 F 上的线性组合. 于是 $\mathcal{E}, \mathcal{A}, \dots, \mathcal{A}^{d-1}$ 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 $\beta_0, \beta_1, \dots, \beta_{d-1} \in F$.

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

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

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.

(ii) 类似. \square

注解 3.13 由 (1) 可知, 当 \mathcal{A} 可逆时, $\mathcal{A}^{-1} \in F[\mathcal{A}]$. 类似地, 当 \mathcal{A} 可逆时, $A^{-1} \in F[A]$.

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, 简称 \mathcal{A} -子空间.

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

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

两个平凡的 \mathcal{A} -子空间是 $\{\mathbf{0}\}$ 和 V .

例 4.2 设 $\mathcal{A} \in \mathcal{L}(V)$. 证明: \mathcal{A} 是数乘算子当且仅当 V 中的任何子空间都是 \mathcal{A} -子空间.

证明. 设 $\mathcal{A} = \lambda\mathcal{E}$, 其中 $\lambda \in F$. 设 $U \subset V$ 是子空间. 则对任意 $\mathbf{u} \in U$, $\mathcal{A}(\mathbf{u}) = \lambda\mathbf{u} \in U$. 故 U 是 \mathcal{A} -子空间.

反之, 设 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 是 V 的一组基. 因为 $\langle \mathbf{e}_i \rangle$ 是 \mathcal{A} -子空间, 所以 $\mathcal{A}(\mathbf{e}_i) \in \langle \mathbf{e}_i \rangle$. 故存在 $\lambda_i \in F$ 使得 $\mathcal{A}(\mathbf{e}_i) = \lambda_i \mathbf{e}_i$, $i = 1, \dots, n$. 设 $1 \leq i < j \leq n$ 且 $U = \langle \mathbf{e}_i + \mathbf{e}_j \rangle$. 因为 U 是 \mathcal{A} -子空间, 所以存在 $\alpha \in F$ 使得

$$\mathcal{A}(\mathbf{e}_i + \mathbf{e}_j) = \alpha(\mathbf{e}_i + \mathbf{e}_j).$$

另一方面,

$$\mathcal{A}(\mathbf{e}_i + \mathbf{e}_j) = \lambda_i \mathbf{e}_i + \lambda_j \mathbf{e}_j.$$

故

$$(\alpha - \lambda_i)\mathbf{e}_i + (\alpha - \lambda_j)\mathbf{e}_j = \mathbf{0}.$$

因为 $\mathbf{e}_i, \mathbf{e}_j$ 线性无关, 所以 $\lambda_i = \alpha$ 和 $\lambda_j = \alpha$. 于是, $\lambda_i = \lambda_j$. 由此可知

$$\lambda_1 = \dots = \lambda_n =: \lambda.$$

故 $\mathcal{A}(\mathbf{e}_i) = \lambda\mathbf{e}_i$, $i = 1, \dots, n$. 故 \mathcal{A} 在 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 下的矩阵是 λE . 从而, $\mathcal{A} = \lambda\mathcal{A}$. \square

例 4.3 设 U_1 和 U_2 是 V 的非平凡子空间且 $V = U_1 \oplus U_2$. 再设 \mathcal{A} 是从 V 到 U_1 的投影. 证明:

(i) U_1 和 U_2 都是 \mathcal{A} -子空间;

(ii) 设 $\mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2$ 都非零. 则 $\langle \mathbf{u}_1 + \mathbf{u}_2 \rangle$ 不是 \mathcal{A} -子空间.

证明. (i) 设 $\mathbf{v}_1 \in U_1$. 则 $\mathcal{A}(\mathbf{v}_1) = \mathbf{v}_1$. 故 U_1 是 \mathcal{A} -子空间. 设 $\mathbf{v}_2 \in U_2$. 则 $\mathcal{A}(\mathbf{v}_2) = \mathbf{0}$. 故 U_2 也是 \mathcal{A} -子空间.

(ii) 假设 $\langle \mathbf{u}_1 + \mathbf{u}_2 \rangle$ 是 \mathcal{A} -子空间. 则 $\mathcal{A}(\mathbf{u}_1 + \mathbf{u}_2) = \lambda(\mathbf{u}_1 + \mathbf{u}_2)$, 其中 $\lambda \in F$. 另一方面, $\mathcal{A}(\mathbf{u}_1 + \mathbf{u}_2) = \mathbf{u}_1$. 故 $(\lambda - 1)\mathbf{u}_1 + \mathbf{u}_2 = \mathbf{0}$. 根据第一章定理 1.18 (ii), $\mathbf{u}_2 = \mathbf{0}$. 矛盾. \square

命题 4.4 设 $\mathcal{A} \in \mathcal{L}(V)$, U 是 d 维 \mathcal{A} -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 $B \in M_d(F)$ 是 \mathcal{A}_U 的某个矩阵表示. 进而 $\mu_{\mathcal{A}_U} | \mu_{\mathcal{A}}$, $\mu_B | \mu_{\mathcal{A}}$, $\mu_D | \mu_{\mathcal{A}}$.

证明. 设 $\mathbf{e}_1, \dots, \mathbf{e}_d$ 是 U 的一组基. 把它扩充为 V 的一组基 $\mathbf{e}_1, \dots, \mathbf{e}_d, \mathbf{e}_{d+1}, \dots, \mathbf{e}_n$. 因为 U 是 \mathcal{A} 的不变子空间, 所以当 $j \in \{1, 2, \dots, d\}$ 时, $\mathcal{A}(\mathbf{e}_j)$ 是 $\mathbf{e}_1, \dots, \mathbf{e}_d$ 的线性组合, 即 $\mathcal{A}(\mathbf{e}_j)$ 关于 $\mathbf{e}_{d+1}, \dots, \mathbf{e}_n$ 的坐标都等于零. 于是 \mathcal{A} 在 $\mathbf{e}_1, \dots, \mathbf{e}_d, \mathbf{e}_{d+1}, \dots, \mathbf{e}_n$ 下的矩阵如命题所述形式, 且 B 是 \mathcal{A}_U 在 $\mathbf{e}_1, \dots, \mathbf{e}_d$ 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 3.3, $\mu_B | \mu_A$, $\mu_D | \mu_A$, 且 $\mu_{A_U} | \mu_A$. \square

给定 $\mathcal{A} \in \mathcal{L}(V)$, $\{\mathbf{0}\}$ 和 V 平凡的 \mathcal{A} -子空间. 下面的引理指出如何寻找非平凡的 \mathcal{A} -子空间.

引理 4.5 设 $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$ 满足 $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$. 则 $\ker(\mathcal{B})$ 和 $\text{im}(\mathcal{B})$ 是 \mathcal{A} 的不变子空间.

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

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

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

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

命题 4.6 设 $\mathcal{A} \in \mathcal{L}(V)$, $f \in F[t]$. 则 $\ker(f(\mathcal{A}))$ 和 $\text{im}(f(\mathcal{A}))$ 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 $\text{im}(f(\mathcal{A}))$ 都是 \mathcal{A} 的不变子空间(引理 4.5). \square

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 4.7 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, $\mathcal{A}(\mathbf{x}) \in U_1 \cap U_2$. \square

关于多项式最小公倍式的注记.

1. 设 $f_1, \dots, f_k \in F[t] \setminus \{0\}$. 如果 $h \in F[t]$ 满足 $f_i | h$, $i = 1, \dots, k$, 则称 h 是 f_1, \dots, f_k 的公倍式. 如果 l 是 f_1, \dots, f_k 的非零公倍式, 且 f_1, \dots, f_k 的公倍式都是 l 的倍式, 则称 l 是 f_1, \dots, f_k 的最小公倍式, 记为 $\text{lcm}(f_1, \dots, f_k)$. 它在相伴意义下是唯一的 (\because 相伴等价于相互整除, 见第一周讲义中命题 5.7).
2. 设 $f_1, f_2 \in F[t]$. 则 $\text{lcm}(f_1, f_2) \text{gcd}(f_1, f_2) = f_1 f_2$ (见第一周讲义中例 5.16). 特别地, 当 $\text{gcd}(f_1, f_2) = 1$ 时, $\text{lcm}(f_1, f_2) = f_1 f_2$.

3. 设 $k > 2$. 则

$$\text{lcm}(f_1, f_2, \dots, f_k) = \text{lcm}(f_1, \text{lcm}(f_2, \dots, f_k))$$

(见第一周讲义中推论 5.10).

特别地, 如果 $\text{gcd}(f_i, f_j) = 1, 1 \leq i < j \leq k$, 则
 $\text{lcm}(f_1, f_2, \dots, f_k) = f_1 f_2 \cdots f_k$.