

第二章 线性算子

验证相似是等价关系如下. 对任意 $A \in M_n(F)$, $A = E^{-1}AE$. 于是 $A \sim_s A$. 自反性成立. 设 $B \sim_s A$. 则存在 $P \in GL_n(F)$ 使得 $B = P^{-1}AP$. 于是 $A = PBP^{-1}$, $A \sim_s B$. 对称性成立. 设 $A \sim_s B$, $B \sim_s C$. 则存在 $P, Q \in GL_n(F)$ 使得 $B = P^{-1}AP$ 和 $C = Q^{-1}BQ$. 于是 $C = Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ)$, 即 $A \sim_s C$. 传递性成立.

两个矩阵相似当且仅当它们是同一个线性映射在不同基底下的矩阵. 于是我们的问题可等价地叙述为给定矩阵 $A \in M_n(F)$ 研究并计算与 A 相似的尽可能简单的矩阵.

命题 2.1 设 $A, B \in M_n(F)$. 如果 $A \sim_s B$, 则

$$\text{rank}(A) = \text{rank}(B), \quad \det(A) = \det(B), \quad \text{tr}(A) = \text{tr}(B).$$

证明. 设 $B = P^{-1}AP$, 其中 $P \in GL_n(F)$. 因为乘以可逆矩阵不改变矩阵的秩, 所以 $\text{rank}(B) = \text{rank}(A)$. 由行列式乘法定理可知, $\det(B) = \det(P^{-1}) \det(A) \det(P) = \det(A)$.

为了证明 $\text{tr}(A) = \text{tr}(B)$. 我们首先注意到迹是交换不变量, 即对任意 $M = (m_{i,j}), N = (n_{i,j}) \in M_n(F)$,

$$\text{tr}(MN) = \text{tr}(NM).$$

(见上学期第二章第 5 讲命题 7.8.)

我们由 $\text{tr}(B) = \text{tr}(P^{-1}AP) = \text{tr}(PP^{-1}A) = \text{tr}(A)$. \square

例 2.2 证明:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \not\sim_s \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

证明. 因为这两个矩阵的迹不同, 所以它们不相似.

例 2.3 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?

证明. 设 $P \in GL_2(F)$ 使得 $B = P^{-1}AP$. 则 $PB = AP$, 即 $P(E + C) = P$. 其中

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

于是 $PC = O$. 因为 P 可逆, 所以 $C = O$. 矛盾. 这两个矩阵不相似. \square

2.2 线性算子代数

我们已经知道 $(\mathcal{L}(V), +, \mathcal{O}, \text{数乘})$ 是 F 上的线性空间. 注意到任何两个 V 上的线性算子都可以复合, 且对任意 $A, B, C \in \mathcal{L}(V)$,

$$A \circ (B \circ C) = (A \circ B) \circ C \quad \text{和} \quad A \circ \mathcal{E} = \mathcal{E} \circ A = A.$$

再由上一讲命题 1.19 可知, $(\mathcal{L}(V), +, \mathcal{O}, \circ, \mathcal{E})$ 是一个环. 此外对任意 $\alpha \in F$,

$$\alpha(\mathcal{A} \circ \mathcal{B}) = (\alpha\mathcal{A}) \circ \mathcal{B} = \mathcal{A} \circ (\alpha\mathcal{B}).$$

这个性质使得我们称 $\mathcal{L}(V)$ 是 F 上的一个代数.

定理 2.4 设

$$\Phi: \mathcal{L}(V) \longrightarrow M_n(F)$$

$$\mathcal{A} \longmapsto A, \quad \mathcal{A} \text{ 在 } \mathbf{e}_1, \dots, \mathbf{e}_n \text{ 下的矩阵}$$

则 Φ 既是线性同构又是环同构(此时称 Φ 是代数同构).

证明. 根据定理 1.9, Φ 是线性同构. 设 $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$, 它们在 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 下的矩阵分别是 A, B . 根据命题 1.10, $\Phi(\mathcal{A} \circ \mathcal{B}) = AB = \Phi(\mathcal{A})\Phi(\mathcal{B})$. 可直接验证 $\Phi(\mathcal{E}) = E_n$. \square .

为了简洁, $\mathcal{A} \circ \mathcal{B}$ 也写成 $\mathcal{A}\mathcal{B}$. 设 $\mathcal{A} \in \mathcal{L}(V)$. 如果 \mathcal{A} 可逆, 则称 \mathcal{A} 是可逆算子. 如果存在 $\lambda \in F$ 使得对任意 $\mathbf{x} \in V$, $\mathcal{A}(\mathbf{x}) = \lambda\mathbf{x}$, 则称 \mathcal{A} 是数乘算子. 此时 $\mathcal{A} = \lambda\mathcal{E}$. 如果存在 $k \in \mathbb{Z}^+$ 使得 $\mathcal{A}^k = \mathcal{O}$, 则称 \mathcal{A} 是幂零算子. 如果 $\mathcal{A}^2 = \mathcal{A}$, 则称 \mathcal{A} 是幂等算子.

由上述定理可知, \mathcal{A} 是可逆(数乘, 幂零, 幂等)算子当且仅当 $\Phi(\mathcal{A})$ 是(数乘, 幂零, 幂等)矩阵.

例 2.5 设 $\mathcal{D}: \mathbb{R}[x]^{(n)} \longrightarrow \mathbb{R}[x]^{(n)}$ 由公式 $\mathcal{D}(f) = f'$ 定义. 则 $\mathcal{D}^n = \mathcal{O}$. 该算子在 $1, x, \dots, x^{n-1}$ 下的矩阵见第一章第三讲例 5.7.

例 2.6 设 U_1, U_2 是 V 的子空间满足 $V = U_1 \oplus U_2$. 设 π_i 是 V 关于上述直和到 U_i 的投影, $i = 1, 2$. 因为关于直和的投影具有等方性(见第一章第七讲命题 12.2(ii)), 所以 π_1 和 π_2 都是幂等的. 设 $\mathbf{e}_1, \dots, \mathbf{e}_d$ 是 U_1 的基, $\mathbf{e}_{d+1}, \dots, \mathbf{e}_n$ 是 U_2 的基. 则 $\mathbf{e}_1, \dots, \mathbf{e}_d, \mathbf{e}_{d+1}, \dots, \mathbf{e}_n$ 是 V 的基. 在该基下 π_1 和 π_2 的矩阵分别是

$$\begin{pmatrix} E_d & O \\ O & O \end{pmatrix} \quad \text{和} \quad \begin{pmatrix} O & O \\ O & E_{n-d} \end{pmatrix}.$$

3 单个算子生成的子环

设 $\mathcal{A} \in \mathcal{L}(V)$. 令 $F[\mathcal{A}] = \langle \{\mathcal{A}^k \mid k \in \mathbb{N}\} \rangle$. 则 $F[\mathcal{A}]$ 是 $\{\alpha_k \mathcal{A}^k + \alpha_{k-1} \mathcal{A}^{k-1} + \dots + \alpha_1 \mathcal{A} + \alpha_0 \mathcal{E} \mid k \in \mathbb{N}, \alpha_k, \alpha_{k-1}, \dots, \alpha_0 \in F\}$.

注意到 $F[\mathcal{A}] \subset \mathcal{L}(V)$. 对任意 $G, H \in F[\mathcal{A}]$, $GH \in F[\mathcal{A}]$. 而且 $\mathcal{O}, \mathcal{E} \in F[\mathcal{A}]$. 于是 $F[\mathcal{A}]$ 是子环. 直接验证可得 $GH = HG$. 于是 $F[\mathcal{A}]$ 是交换环.

我们还可以从另一个角度看出 $F[\mathcal{A}]$ 是交换环. 设 A 是 \mathcal{A} 在 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 下的矩阵. 则由定理 2.1, 代数同构

$$\Phi^{-1} : M_n(F) \longrightarrow \mathcal{L}(V)$$

把交换环 $F[A]$ 映到 $F[\mathcal{A}]$. 因为 $F[A]$ 是交换环(见上学期第四章第 3.5 节), 所以 $F[\mathcal{A}]$ 是交换环.

可直接验证映射

$$\begin{aligned}\phi : F &\longrightarrow F[\mathcal{A}] \\ \alpha &\mapsto \alpha\mathcal{E}\end{aligned}$$

是环同态. 由多项式赋值同态定理, ϕ 可以扩展为一个从 $F[t]$ 到 $F[\mathcal{A}]$ 的环同态 $\phi_{\mathcal{A}}$ 满足 $\phi(t) = \mathcal{A}$. 通过赋值同态得到对任意 $f(t) = f_k t^k + f_{k-1} t^{k-1} + \cdots + f_1 t + f_0 \in F[t]$, 其中 $f_k, f_{k-1}, \dots, f_1, f_0 \in F$ 得到

$$\phi_{\mathcal{A}}(f) = f_k \mathcal{A}^k + f_{k-1} \mathcal{A}^{k-1} + \cdots + f_1 \mathcal{A} + f_0 \mathcal{E} = f(\mathcal{A}).$$

且对任意 $p, q \in F[t]$, 我们有

$$(p + q)(\mathcal{A}) = p(\mathcal{A}) + q(\mathcal{A}) \quad \text{和} \quad (pq)(\mathcal{A}) = p(\mathcal{A})q(\mathcal{A}).$$

事实上, 上述赋值同态也可由赋值同态 $\phi_{\mathcal{A}} : F[t] \longrightarrow F[\mathcal{A}]$ 与 $\Phi^{-1} : F[\mathcal{A}] \longrightarrow F[A]$ 得到. 赋值同态 $\phi_{\mathcal{A}}$ 的构造见上学期第五章讲义一定理 1.10.

由上一讲定理 1.9 可知, $\dim(\mathcal{L}(V)) = n^2$. 于是 $\mathcal{E}, \mathcal{A}, \dots, \mathcal{A}^{n^2}$ 在 F 上必然线性相关. 换言之, 存在 $\alpha_0, \alpha_1, \dots, \alpha_{n^2} \in F$, 不全为零, 使得

$$\alpha_0 \mathcal{E} + \alpha_1 \mathcal{A} + \cdots + \alpha_{n^2-1} \mathcal{A}^{n^2-1} + \alpha_{n^2} \mathcal{A}^{n^2} = \mathcal{O}.$$

于是 \mathcal{A} 不可能是未定元. 对 $G \in F[\mathcal{A}]$, 我们不能直接定义 \mathcal{A} 的“次数”和“系数”.

例 3.1 设 \mathcal{A} 是数乘算子 $\lambda\mathcal{E}$, $f(t) = t^2 - 3t - \lambda^2$. 则

$$f(\mathcal{A}) = \mathcal{A}^2 - 3\mathcal{A} - \lambda^2\mathcal{E} = -3\mathcal{A} = -2\lambda\mathcal{E}. \quad \square$$

注解 3.2 赋值同态 $f(t) \mapsto f(\mathcal{A})$ 是交换环 $F[t]$ 到 $F[\mathcal{A}]$ 的满同态. 于是

$$F[\mathcal{A}] = \{p(\mathcal{A}) \mid p \in F[t]\} \quad \text{且} \quad F[\mathcal{A}] = \{p(\mathcal{A}) \mid p \in F[t]\}.$$

定理 3.3 (核核分解) 设 $\mathcal{A} \in \mathcal{L}(V)$, $p, q \in F[t]$ 互素. 如果 $(pq)(\mathcal{A}) = \mathcal{O}$, 则

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

证明. 由 Bezout 关系, 存在 $u, v \in F[t]$ 使得

$$u(t)p(t) + v(t)q(t) = 1.$$

于是,

$$u(\mathcal{A})p(\mathcal{A}) + v(\mathcal{A})q(\mathcal{A}) = \mathcal{E}. \quad (1)$$

设 $\mathbf{x} \in V$. 则

$$\begin{aligned} \mathbf{x} &= \mathcal{E}(\mathbf{x}) = (u(\mathcal{A})p(\mathcal{A}) + v(\mathcal{A})q(\mathcal{A}))(\mathbf{x}) \quad (\because (1)) \\ &= (u(\mathcal{A})p(\mathcal{A}))(\mathbf{x}) + (v(\mathcal{A})q(\mathcal{A}))(\mathbf{x}) \quad (\text{映射加法的定义}) \\ &= (p(\mathcal{A})u(\mathcal{A}))(\mathbf{x}) + (q(\mathcal{A})v(\mathcal{A}))(\mathbf{x}) \quad (F[\mathcal{A}] \text{ 是交换环}) \\ &= p(\mathcal{A})\underbrace{(u(\mathcal{A}))(\mathbf{x}))}_{\mathbf{y}} + q(\mathcal{A})\underbrace{(v(\mathcal{A}))(\mathbf{x}))}_{\mathbf{z}}. \quad (\text{乘法即复合}) \\ &= p(\mathcal{A})(\mathbf{y}) + q(\mathcal{A})(\mathbf{z}). \end{aligned}$$

因为 $(pq)(\mathcal{A}) = \mathcal{O}$, 所以 $p(\mathcal{A})q(\mathcal{A}) = \mathcal{O}$. 于是 $q(\mathcal{A})(p(\mathcal{A})(\mathbf{y})) = \mathbf{0}$, 即 $p(\mathcal{A})(\mathbf{y}) \in \ker(q(\mathcal{A}))$. 类似可知 $q(\mathcal{A})(\mathbf{z}) \in \ker(p(\mathcal{A}))$. 我们得到 $\mathbf{x} \in \ker(p(\mathcal{A})) + \ker(q(\mathcal{A}))$. 由 \mathbf{x} 的任意性推出 $V = \ker(p(\mathcal{A})) + \ker(q(\mathcal{A}))$.

再设 $\mathbf{x} \in \ker(p(\mathcal{A})) \cap \ker(q(\mathcal{A}))$. 则由 (1) 得出

$$\mathbf{x} = \mathcal{E}(\mathbf{x}) = u(\mathcal{A})p(\mathcal{A})(\mathbf{x}) + v(\mathcal{A})q(\mathcal{A})(\mathbf{x}) = \mathbf{0}.$$

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

例 3.4 设 $\mathcal{A} \in \mathcal{L}(V)$ 满足 $\mathcal{A}^2 = \mathcal{E}$. 证明: 当 F 的特征不等于 2 时,

$$\text{rank}(\mathcal{A} - \mathcal{E}) + \text{rank}(\mathcal{A} + \mathcal{E}) = \dim(V).$$

证明. 由核像版的对偶公式 (第一章第二讲命题 4.14 (iii)), 我们只要证明

$$\dim(\ker(\mathcal{A} - \mathcal{E})) + \dim(\ker(\mathcal{A} + \mathcal{E})) = \dim(V).$$

设 $f(t) = t^2 - 1$. 则 $f(\mathcal{A}) = \mathcal{A}^2 - \mathcal{E} = \mathcal{O}$. 设 $p = (t - 1)$, $q = (t + 1)$. 因为 $pq = f$, 所以 $(pq)(\mathcal{A}) = \mathcal{O}$. 因为 F 的特征不等于 2, 所以 $\text{gcd}(p, q) = 1$. 由核分解定理可知

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

又因为 $p(\mathcal{A}) = \mathcal{A} - \mathcal{E}$ 和 $q(\mathcal{A}) = \mathcal{A} + \mathcal{E}$, 所以

$$\dim(V) = \dim(\ker(\mathcal{A} - \mathcal{E})) + \dim(\ker(\mathcal{A} + \mathcal{E}))$$

(直和维数的基本性质—第一章第二讲命题 4.15). \square

满足 $\mathcal{A}^2 = \mathcal{E}$ 的算子称为对合算子. 典型例子是矩阵

$$\begin{pmatrix} E_k & O \\ O & -E_\ell \end{pmatrix}.$$

定理 3.5 (核像分解 I)¹ 设 $\mathcal{A} \in \mathcal{L}(V)$. 则

$$V = \ker(\mathcal{A}) \oplus \operatorname{im}(\mathcal{A}) \iff \operatorname{rank}(\mathcal{A}) = \operatorname{rank}(\mathcal{A}^2).$$

证明. 断言. 对任意 $\mathcal{A} \in \mathcal{L}(V)$,

$$\ker(\mathcal{A}) \subset \ker(\mathcal{A}^2), \quad \operatorname{im}(\mathcal{A}) \supset \operatorname{im}(\mathcal{A}^2).$$

断言的证明. 设 $\mathbf{v} \in \ker(\mathcal{A})$. 则

$$\mathcal{A}^2(\mathbf{v}) = \mathcal{A}(\mathcal{A}(\mathbf{v})) = \mathcal{A}(\mathbf{0}) = \mathbf{0}.$$

设 $\mathbf{y} \in \operatorname{im}(\mathcal{A}^2)$. 则存在 $\mathbf{z} \in V$ 使得 $\mathbf{y} = \mathcal{A}^2(\mathbf{z})$. 于是 $\mathbf{y} = \mathcal{A}(\mathcal{A}(\mathbf{z})) \in \operatorname{im}(\mathcal{A})$. 断言成立.

(\Leftarrow) 因为 $\operatorname{rank}(\mathcal{A}) = \operatorname{rank}(\mathcal{A}^2)$, 所以

$$\dim(\ker(\mathcal{A})) = \dim(\ker(\mathcal{A}^2)).$$

这是因为 $\dim(\ker(\mathcal{A})) + \operatorname{rank}(\mathcal{A}) = \dim(\ker(\mathcal{A}^2)) + \operatorname{rank}(\mathcal{A}^2)$ (上一讲注释 1.16). 由断言可知 $\ker(\mathcal{A}) = \ker(\mathcal{A}^2)$. 设

¹袁力, 沈洁. 常州工学院学报 27 卷第二期, 2014 年 4 月.

$\mathbf{x} \in \ker(\mathcal{A}) \cap \text{im}(\mathcal{A})$. 则存在 $\mathbf{y} \in V$ 使得 $\mathbf{x} = \mathcal{A}(\mathbf{y})$ 且 $\mathcal{A}(\mathbf{x}) = \mathbf{0}$. 于是 $\mathcal{A}^2(\mathbf{y}) = \mathbf{0}$. 因为 $\ker(\mathcal{A}) = \ker(\mathcal{A}^2)$, 所以 $\mathbf{y} \in \ker(\mathcal{A})$. 于是 $\mathbf{x} = \mathbf{0}$. 即 $\ker(\mathcal{A}) + \text{im}(\mathcal{A})$ 是直和. 于是 $\dim(\ker(\mathcal{A}) + \text{im}(\mathcal{A})) = \dim(\ker(\mathcal{A})) + \dim(\text{im}(\mathcal{A})) = \dim(V)$. 我们得出 $V = \ker(\mathcal{A}) \oplus \text{im}(\mathcal{A})$.

(\implies) 由断言和推论 1,14 可知, 我们只要证明

$$\text{im}(\mathcal{A}) \subset \text{im}(\mathcal{A}^2)$$

即可. 设 $\mathbf{x} \in \text{im}(\mathcal{A})$. 则存在 $\mathbf{y} \in V$ 使得 $\mathbf{x} = \mathcal{A}(\mathbf{y})$. 因为 $V = \ker(\mathcal{A}) + \text{im}(\mathcal{A})$, 所以存在 $\mathbf{u} \in \ker(\mathcal{A})$, $\mathbf{v} \in \text{im}(\mathcal{A})$ 使得 $\mathbf{y} = \mathbf{u} + \mathbf{v}$ 且 $\mathbf{v} = \mathcal{A}(\mathbf{w})$, 其中 \mathbf{w} 是 V 中某个向量. 于是 $\mathbf{x} = \mathbf{u} + \mathcal{A}(\mathbf{w})$, 从而

$$\mathbf{x} = \mathcal{A}(\mathbf{y}) = \mathcal{A}(\mathbf{u}) + \mathcal{A}^2(\mathbf{w}) = \mathcal{A}^2(\mathbf{w}) \in \text{im}(\mathcal{A}^2).$$

我们有 $\text{im}(\mathcal{A}) \subset \text{im}(\mathcal{A}^2)$. \square

注解 3.6 由上述定理和证明中的断言可知, 以下结论是彼此等价的.

(i) $V = \ker(\mathcal{A}) \oplus \text{im}(\mathcal{A})$;

(ii) $\text{rank}(\mathcal{A}) = \text{rank}(\mathcal{A}^2)$;

(iii) $\text{im}(\mathcal{A}) = \text{im}(\mathcal{A}^2)$;

(iv) $\ker(\mathcal{A}) = \ker(\mathcal{A}^2)$;

(v) $\dim(\ker(\mathcal{A})) = \dim(\ker(\mathcal{A}^2))$.

例 3.7 设 $\mathcal{A} \in \mathcal{L}(V)$ 满足 $\mathcal{A}^2 = \mathcal{A}$. 证明

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

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

例 3.8 设 \mathcal{D} 是 $\mathbb{R}[x]_n$ 上的导数算子. 则 $\ker(\mathcal{D}) = \mathbb{R}$ 且 $\operatorname{im}(\mathcal{D}) = \mathbb{R}[x]^{(n-1)}$. 因为 $\mathbb{R} \subset \mathbb{R}[x]^{(n-1)}$, 所以 $\ker(\mathcal{D}) + \operatorname{im}(\mathcal{D})$ 不是直和.

4 算子和矩阵的极小多项式

定义 4.1 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 $A \in M_n(F)$, 我们有关于 A 的零化多项式和极小多项式的概念.

引理 4.2 设 $\mathcal{A} \in \mathcal{L}(V)$, $f(t) \in F[t]$, $p(t)$ 是 \mathcal{A} 的极小多项式. 则

$$f(\mathcal{A}) = \mathcal{O} \iff p|f.$$

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, $r(t) = 0$. 如果 $r(t) = 0$, 则 $r(\mathcal{A}) = \mathcal{O}$. 于是, $f(\mathcal{A}) = \mathcal{O}$. \square

命题 4.3 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 $1, \mathcal{A}, \dots, \mathcal{A}^{n^2}$ 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 4.2, $p|q$ 且 $q|p$. 于是 $p = cq$, 其中 $c \in F \setminus \{0\}$. 因为 p 和 q 都首一, 所以 $c = 1$. \square

注解 4.4 以上结论对 $A \in M_n(F)$ 同样成立.

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

命题 4.5 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 $\Phi : F[\mathcal{A}] \rightarrow F[A]$ 是线性同构和环同构, 其中 $\Phi(\mathcal{A}) = A$ (见第 3 节开始的段落). 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 $\mu_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$. 根据引理 4.2,

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

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

例 4.6 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, $\mathcal{A} = \lambda\mathcal{E}$. \square

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

例 4.7 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 4.2, $\mu_{\mathcal{A}} | t^k$. 于是 $\mu_{\mathcal{A}}$ 是 t 的幂次. \square

引理 4.8 设 $A, B \in M_n(F)$, $P \in GL_n(F)$ 使得 $B = P^{-1}AP$. 设 $f \in F[t]$. 则

$$f(B) = P^{-1}f(A)P.$$

特别地, $A \sim_s B \implies f(A) \sim_s f(B)$.

证明. 直接计算得对任意 $i \in \mathbb{N}$,

$$B^i = \underbrace{(P^{-1}AP)(P^{-1}AP) \cdots (P^{-1}AP)(P^{-1}AP)}_i = P^{-1}A^iP.$$

设 $f(t) = f_k t^k + f_{k-1} t^{k-1} + \cdots + f_1 t + f_0$. 则

$$\begin{aligned} f(B) &= f_k P^{-1}A^kP + f_{k-1} P^{-1}A^{k-1}P + \cdots + f_1 P^{-1}AP + f_0 P^{-1}EP \\ &= P^{-1}f(A)P. \quad \square \end{aligned}$$

命题 4.9 设 $A, B \in M_n(F)$. 如果 $A \sim_s B$, 则 $\mu_A = \mu_B$.

证明. 由引理 4.8 和 $\mu_A(A) = O$ 可知, $\mu_A(B) = O$. 于是 $\mu_B | \mu_A$ (引理 4.2). 同理 $\mu_A | \mu_B$. 因为 μ_A 和 μ_B 都首一, 所以 $\mu_A = \mu_B$. \square

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 4.6). 于是, $\mu_A \neq \mu_B$. 故 $A \not\sim_s B$. \square

例 4.11 设 $A, B \in M_n(F)$, A 是数乘矩阵, B 是幂零矩阵. 则

$$A \sim_s B \iff A = B = O.$$

证明. 由例 4.6 和例 4.7 可知, $\mu_A = t - \lambda$, $\mu_B = t^k$, 其中 $\lambda \in F, k \in \mathbb{Z}^+$. 设 $A \sim_s B$. 则 $\mu_A = \mu_B$ (命题 4.9). 于是 $\lambda = 0$ 且 $k = 1$. 由此得出 $A = O$ 和 $B = O$. 另一个方向是平凡的. \square

命题 4.12 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 $\mathcal{E}, \mathcal{A}, \dots, \mathcal{A}^{d-1}$ 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 $p(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_{d-1} t^{d-1} \in F[t]$. 则 $p(\mathcal{A}) = \mathcal{O}$. 因为 $\deg_t(p) < d$, 所以 $p = 0$. 于是, $\alpha_0 = \alpha_1 = \dots = \alpha_{d-1} = 0$. 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

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

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 4.13 由 (2) 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, $k = 1, 2, \dots, n$. 但 $\langle x^k \rangle$ 不是, $k = 0, 1, \dots, n-1$.

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 $B \in M_d(F)$ 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 4.2, $\mu_B | \mu_A$, $\mu_D | \mu_A$, 且 $\mu_{A_U} | \mu_A$. \square

引理 5.4 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 $\epsilon_1, \dots, \epsilon_{d_1}$ 是 U_1 的基, $\delta_1, \dots, \delta_{d_2}$ 是 U_2 的基. 则在 V 的基底 $\epsilon_1, \dots, \epsilon_{d_1}, \delta_1, \dots, \delta_{d_2}$ 下 \mathcal{A} 的矩阵是

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

其中 $A_i \in M_{d_i}(F)$ 是 \mathcal{A}_{U_i} 在对应基下的矩阵, $i = 1, 2$. 进而 $\mu_A = \text{lcm}(\mu_{\mathcal{A}_{U_1}}, \mu_{\mathcal{A}_{U_2}})$ (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 $\epsilon_1, \dots, \epsilon_{d_1}, \delta_1, \dots, \delta_{d_2}$ 线性无关. 所以 $\epsilon_1, \dots, \epsilon_{d_1}, \delta_1, \dots, \delta_{d_2}$ 是 V 的一组基. 对 $i \in \{1, 2, \dots, d_1\}$, $\mathcal{A}(\epsilon_i) \in U_1$, $\mathcal{A}(\epsilon_i)$ 是 $\epsilon_1, \dots, \epsilon_{d_1}$ 的线性组合, 它关于 $\delta_1, \dots, \delta_{d_2}$ 的坐标都是零. 于是, 存在 $A_1 \in M_{d_1}(F)$ 使得

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

类似地, 存在 $A_2 \in M_{d_2}(F)$ 使得

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

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 $\epsilon_1, \dots, \epsilon_{d_1}, \delta_1, \dots, \delta_{d_2}$ 下的矩阵等于 A .

设 $p = \text{lcm}(\mu_{\mathcal{A}_{U_1}}, \mu_{\mathcal{A}_{U_2}})$. 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 4.2). 于是

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

由此和引理 4.2, $\mu_{\mathcal{A}} | p$. 再利用首一性得出 $p = \mu_{\mathcal{A}}$. \square