第二章　线性算子

推论 1.11 设 $\phi \in \mathrm{Hom}(V, W)$ 且 $\mathrm{rank}(\phi) = r$. 则存在 V 的一组基 $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ 和 W 的一组基 $\epsilon'_1, \ldots, \epsilon'_m$ 使得在该基下 ϕ 的矩阵是

$$M = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m \times n}$$

且 $r = \dim(\mathrm{im}(\phi))$.

证明. 设 $d = \ker(\phi)$, $\mathbf{e}'_{n-d+1}, \ldots, \mathbf{e}'_n$ 是 $\ker(\phi)$ 的一组基. 把它扩充为 V 的一组基

$$\mathbf{e}'_1, \ldots, \mathbf{e}'_{n-d}, \mathbf{e}'_{n-d+1}, \ldots, \mathbf{e}'_n.$$

因为 $\mathrm{im}(\phi) = \langle \phi(\mathbf{e}'_1), \ldots, \phi(\mathbf{e}'_n) \rangle$ 且 $\phi(\mathbf{e}'_{n-d+1}) = \cdots = \phi(\mathbf{e}'_n) = \mathbf{0}_W$, 所以

$$\mathrm{im}(\phi) = \langle \phi(\mathbf{e}'_1), \ldots, \phi(\mathbf{e}'_{n-d}) \rangle.$$

因为 $\dim(\mathrm{im}(\phi)) = n - d$, 所以 $\epsilon'_1 = \phi(\mathbf{e}'_1), \ldots, \epsilon'_{n-d} = \phi(\mathbf{e}'_{n-d})$ 是 $\mathrm{im}(\phi)$ 的一组基. 将其扩充为 W 的一组基

$$\epsilon'_1, \ldots, \epsilon'_{n-d}, \epsilon'_{n-d+1}, \ldots, \epsilon'_m.$$

则 ϕ 在上述基底下的矩阵等于 M. 特别地, $r = n - d$. □

注解 1.12 第一章第二讲命题 *4.14 (iii)* 可写为

$$\dim(\ker(\phi)) + \mathrm{rank}(\phi) = \dim(V).$$

例 1.13 计算例 *1.6* 中 ϕ 的秩. 设矩阵 B 由例 *1.6* 给出. 则 $\mathrm{rank}(\phi) = \mathrm{rank}(B) = k\mathrm{rank}(A)$.

例 1.14 设 $\phi : F^{m\times n} \longrightarrow F^{n\times m}$ 由公式 $\phi(X) = X^t$ 给出. 求 $\mathrm{rank}(\phi)$.

解. 因为 ϕ 是单射, 所以 $\dim(\ker(\phi)) = 0$. 于是 $\mathrm{rank}(\phi) = mn$.

推论 1.15 设 $\phi \in \mathrm{Hom}(V, W)$. 则

(i) ϕ 是单射当且仅当 $\mathrm{rank}(\phi) = \dim(V)$;

(ii) ϕ 是满射当且仅当 $\mathrm{rank}(\phi) = \dim(W)$.

(iii) 如果 ϕ 是双射, 则 $\dim(V) = \dim(W)$.

证明. (i) ϕ 单当且仅当 $\ker(\phi) = \{\mathbf{0}_V\}$ (第一章第一讲命题 2.3) 当且仅当 $\mathrm{rank}(\phi) = \dim(V)$ (上述注释).

(ii) ϕ 满当且仅当 $\dim(\mathrm{im}(\phi)) = \dim(W)$ 当且仅当 $\mathrm{rank}(\phi) = \dim(W)$ (推论 1.13).

(iii) 由 (i) 和 (ii) 可知,

$$\dim(V) = \mathrm{rank}(\phi) \quad 且 \quad \dim(W) = \mathrm{rank}(\phi). \quad \square$$

推论 1.16 设 $\phi \in \mathrm{Hom}(V, W)$ 且 $\dim(V) = \dim(W)$. 则以下断言等价

(i) ϕ *是单射;*

(ii) ϕ *是满射;*

(iii) ϕ *是双射.*

证明. 设 $n = \dim(V)$, $r = \mathrm{rank}(\phi)$.

(i) $\Longrightarrow$ (ii). 由上述推论 (i), $r = n$. 于是 $r = \dim(W)$. 从而, ϕ 是满射(上述推论 (ii)).

(ii) $\Longrightarrow$ (iii). 由上述推论 (ii), $r = n$. 于是 $r = \dim(V)$. 从而, ϕ 是单射(上述推论 (i)), 即 ϕ 是双射.

(iii) $\Longrightarrow$ (i). 显然. □

推论 1.17 *设* $\phi \in \mathrm{Hom}(V, W)$, A *是* ϕ *在基底* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *和* $\epsilon_1, \ldots, \epsilon_m$ *下的矩阵. 则*

(i) ϕ *是单射当且仅当* A *列满秩, 即* $\mathrm{rank}(A) = n$*;*

(ii) ϕ *是满射当且仅当* A *行满秩, 即* $\mathrm{rank}(A) = m$*.*

(iii) ϕ *是双射, 则* A *是可逆方阵.*

证明. 注意到 $A \in F^{m\times n}$.

(i) ϕ 单当且仅当 $\mathrm{rank}(\phi) = n$, 即 $\mathrm{rank}(A) = n$. (推论 1.15 (i))

(ii) ϕ 满当且仅当 $\mathrm{rank}(\phi) = m$, 即 $\mathrm{rank}(A) = m$. (推论 1.15 (ii))

(iii) 由 (i) 和 (ii) 直接得出. □

1.3 线性同构

定理 1.18 映射*:*

$$\begin{array}{rccc} \rho: & \mathrm{Hom}(V,W) & \longrightarrow & F^{m\times n} \\ & \phi & \mapsto & A_\phi, \end{array}$$

是线性同构, 其中 A_ϕ 是 ϕ 在基底 $\mathbf{e}_1,\ldots,\mathbf{e}_n$; $\epsilon_1,\ldots,\epsilon_m$ 下的矩阵.

证明*:* 设 $\phi,\psi\in\mathrm{Hom}(V,W)$. 则

$$(\phi(\mathbf{e}_1),\ldots,\phi(\mathbf{e}_n))=(\epsilon_1.\ldots.\epsilon_m)A_\phi$$

和

$$(\psi(\mathbf{e}_1),\ldots,\psi(\mathbf{e}_n))=(\epsilon_1.\ldots.\epsilon_m)A_\psi.$$

于是,

$$((\phi+\psi)(\mathbf{e}_1),\ldots,(\phi+\psi)(\mathbf{e}_n))=(\epsilon_1.\ldots.\epsilon_m)\left(A_\phi+A_\psi\right).$$

由此可知, $A_{\phi+\psi}=A_\phi+A_\psi$. 换言之, $\rho(\phi+\psi)=\rho(\phi)+\rho(\psi)$.

再设 $\alpha\in F$. 则

$$(\alpha\phi(\mathbf{e}_1),\ldots,\alpha\phi(\mathbf{e}_n))=(\epsilon_1.\ldots.\epsilon_m)(\alpha A_\phi).$$

故 $\alpha A_\phi=A_{\alpha\phi}$. 换言之, $\rho(\alpha\phi)=\alpha\rho(\phi)$.

我们得到 ρ 是线性映射.

设 $A \in F^{m\times n}$. 设 ϕ_A 是由 $\phi_A(\mathbf{e}_j) = (\epsilon_1, \ldots, \epsilon_m)\vec{A}^{(j)}$, $j = 1, \ldots, n$, 确定的线性映射. 令

$$\begin{array}{ccc} \theta: & F^{m\times n} & \longrightarrow & \mathrm{Hom}(V, W) \\ & A & \mapsto & \phi_A. \end{array}$$

由 ϕ_A 的定义可知, 它在给定基底下的矩阵是 A.

注意到, $\rho \circ \theta(A) = \rho(\phi_A) = A$ 和 $\theta \circ \rho(\phi) = \theta(A_\phi) = \phi$ (θ 的定义). 故 $\theta = \rho^{-1}$. □

1.4 复合与矩阵乘法

定理 1.19 设 Z 是 F 上的线性空间 $\delta_1, \ldots, \delta_s$ 是 Z 的一组基. 设 $\phi: V \to Z$ 是线性的, 它在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$; $\delta_1, \ldots, \delta_s$ 下的矩阵是 $A \in F^{s\times n}$, 而 $\psi: Z \to W$ 是线性的, 它在 $\delta_1, \ldots, \delta_s$, $\epsilon_1, \ldots, \epsilon_m$ 下的矩阵是 $B \in F^{m\times s}$. 则 $\psi \circ \phi$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$; $\epsilon_1, \ldots, \epsilon_m$ 下的矩阵是 BA.

证明. 我们有:

$$(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_n)) = (\delta_1, \ldots, \delta_s)A$$

和

$$(\psi(\delta_1), \ldots, \psi(\delta_s)) = (\epsilon_1, \ldots, \epsilon_m)B.$$

根据上一讲引理 1.7,

$$\begin{aligned}(\psi\circ\phi(\mathbf{e}_1),\dots,\psi\circ\phi(\mathbf{e}_n)) &= (\psi(\delta_1),\dots,\psi(\delta_s))A\\ &= (\epsilon_1,\dots,\epsilon_m)BA\end{aligned}$$

故 $\psi\circ\phi$ 在 $\mathbf{e}_1,\dots,\mathbf{e}_n$; $\epsilon_1,\dots,\epsilon_m$ 下的矩阵是 BA. □

2 线性算子

设 $\mathcal{A}: V\to V$ 是线性映射. 称 $\mathcal{A}$ 是 V 上的线性算子*(*线性变换*)*.

2.1 矩阵的相似

设 $\epsilon_1,\dots,\epsilon_n$ 是 V 的另一组基, 且

$$(\epsilon_1,\dots,\epsilon_n) = (\mathbf{e}_1,\dots,\mathbf{e}_n)P,$$

其中 $P\in \mathrm{GL}_n(F)$. 设算子 $\mathcal{A}\in\mathcal{L}(V)$ 在 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵等于 A. 根据定理 1.11, $\mathcal{A}$ 在 $\epsilon_1,\dots,\epsilon_n$ 下的矩阵等于 $P^{-1}AP$. 我们的问题是如何选取 V 的一组基使得 $\mathcal{A}$ 在该基下的矩阵尽可能简单(零元素尽可能多,非零元出现的尽可能有规律).

定义 2.1 设 $A,B\in \mathrm{M}_n(F)$. 如果存在 $P\in\mathrm{GL}_n(F)$ 使得 $B=P^{-1}AP$, 则称 B 与 A 相似. 记为 $B\sim_s A$.

验证相似是等价关系如下. 对任意 $A \in \mathrm{M}_n(F)$, $A = E^{-1}AE$. 于是 $A \sim_s A$. 自反性成立. 设 $B \sim_s A$. 则存在 $P \in \mathrm{GL}_n(F)$ 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 $P, Q \in \mathrm{GL}_n(F)$ 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矩阵 $A \in \mathrm{M}_n(F)$ 研究并计算与 A 相似的尽可能简单的矩阵.

命题 2.2 设 $A, B \in \mathrm{M}_n(F)$. 如果 $A \sim_s B$, 则

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

证明. 设 $B = P^{-1}AP$, 其中 $P \in \mathrm{GL}_n(F)$. 因为乘以可逆矩阵不改变矩阵的秩, 所以 $\mathrm{rank}(B) = \mathrm{rank}(A)$. 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 $\mathrm{tr}(A) = \mathrm{tr}(B)$. 我们首先注意到迹是交换不变量, 即对任意 $M = (m_{i,j}), N = (n_{i,j}) \in \mathrm{M}_n(F)$,

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

(见上学期第二章第 5 讲命题 7.8.)

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

例 2.3 证明:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \nsim_s \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

证明. 因为这两个矩阵的迹不同, 所以它们不相似.

例 2.4 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 $P \in \mathrm{GL}_2(F)$ 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. □

例 2.5 设 $A, B \in \mathrm{SM}_n(\mathbb{R})$ 都正定. 证明: AB 相似于一个正定矩阵.

证明. 因为 A 和 B 正定, 所以存在 $P, Q \in \mathrm{GL}_n(\mathbb{R})$ 使得 $A = P^tP$ 和 $B = Q^tQ$ (第一章定理 9.16). 故

$$AB = P^tPQ^tQ \sim_s Q\left(P^tPQ^tQ\right)Q^{-1} = (QP^t)(PQ^t) = \underbrace{(PQ^t)^t(PQ^t)}_{M}.$$

因为 PQ^t 可逆, 所以 M 正定 (第一章定理 9.16). □

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 $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathcal{L}(V)$,

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

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 2.6 设

$$\begin{aligned} \Phi: \ \mathcal{L}(V) &\longrightarrow \mathrm{M}_n(F) \\ \mathcal{A} &\mapsto A, \quad \mathcal{A} \text{ 在 } \mathbf{e}_1, \ldots, \mathbf{e}_n \text{ 下的矩阵} \end{aligned}$$

则 Φ 既是线性同构又是环同构, 且对任意 $\alpha, \beta \in F$, $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$,

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

(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 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵分别是 A, B. 根据定理 1.19,

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

可直接验证 $\Phi(\mathcal{E}) = E_n$.

根据 $(\alpha\mathcal{A}) \circ (\beta\mathcal{B}) = \alpha\beta(\mathcal{A} \circ \mathcal{B})$ 可知,

$$\Phi((\alpha\mathcal{A}) \circ (\beta\mathcal{B})) = \alpha\beta\Phi(\mathcal{A})\Phi(\mathcal{B}). \quad \square$$

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 2.7 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 $1, x, \ldots, x^{n-1}$ 下的矩阵是:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & n-1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

例 2.8 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, $i = 1, 2$. 设 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 是 U_1 的基, $\mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$ 是 U_2 的基. 则 $\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$

是 V 的基. 在该基下 π_1 和 π_2 的矩阵分别是

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

2.3 核像分解

定理 2.9 *(核像分解)* 设 $\mathcal{A} \in \mathcal{L}(V)$. 则以下四个断言等价:

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

(ii) $\mathrm{im}(\mathcal{A}) = \mathrm{im}(\mathcal{A}^2)$,

(iii) $\mathrm{rank}(\mathcal{A}) = \mathrm{rank}(\mathcal{A}^2)$,

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

证明. (i) $\Longrightarrow$ (ii). 设 $\mathbf{x} \in \mathrm{im}(\mathcal{A})$. 则存在 $\mathbf{y} \in V$ 使得 $\mathbf{x} = \mathcal{A}(\mathbf{y})$. 因为 $V = \ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$, 所以存在 $\mathbf{u} \in \ker(\mathcal{A})$, $\mathbf{v} \in \mathrm{im}(\mathcal{A})$ 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

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

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

(ii) $\Longrightarrow$ (iii). 因为 $\dim(\mathrm{im}(\mathcal{A})) = \dim(\mathrm{im}(\mathcal{A}^2))$, 所以 $\mathrm{rank}(\mathcal{A}) = \mathrm{rank}(\mathcal{A}^2)$.

(iii) $\Longrightarrow$ (iv). 根据对偶定理, 我们有

$$\dim(\ker(\mathcal{A}))+\mathrm{rank}(\mathcal{A}) = n \quad \text{and} \quad \dim(\ker(\mathcal{A}^2))+\mathrm{rank}(\mathcal{A}^2) = n.$$

于是, $\dim(\ker(\mathcal{A})) = \dim(\ker(\mathcal{A}^2))$. 因为 $\ker(\mathcal{A}) \subset \ker(\mathcal{A}^2)$ (直接验证), 所以 $\ker(\mathcal{A}) = \ker(\mathcal{A}^2)$.

(iv) $\Longrightarrow$ (i). 设 $\mathbf{x} \in \ker(\mathcal{A}) \cap \mathrm{im}(\mathcal{A})$. 则存在 $\mathbf{y} \in V$ 使得 $\mathbf{x} = \mathcal{A}(\mathbf{y})$. 我们有

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

故 $\mathbf{y} \in \ker(\mathcal{A}^2)$. 从而, $\mathbf{y} \in \ker(\mathcal{A})$. 由此得出, $\mathcal{A}(\mathbf{y}) = \mathbf{0}$, 即 $\mathbf{x} = \mathbf{0}$. 于是, $\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$ 是直和且

$$\dim(\ker(\mathcal{A})+\mathrm{im}(\mathcal{A})) = \dim(\ker(\mathcal{A}))+\dim(\mathrm{im}(\mathcal{A})) = n \text{ (对偶定理).}$$

由此得到 (i). □

例 2.10 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 2.11 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.

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. 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 3.3 *(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 3.4 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 3.3, $p|q$ 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 3.5 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.

例 3.6 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, $\mathcal{A} = \lambda\mathcal{E}$. □

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

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 $\mu_{\mathcal{A}}(t) | t^2 - t$ (∵ 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 $\mu_{\mathcal{A}} = t^2 - t$. □

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. □

命题 3.9 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 $\Phi : \mathcal{L}(\mathcal{A}) \longrightarrow \mathrm{M}_n(F)$ 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 $f(A) = O$. 于是, $\mu_{\mathcal{A}}(A) = O$ 且 $\mu_A(\mathcal{A}) = \mathcal{O}$. 根据引理 3.3,

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

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

命题 3.10 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 3.3). 同理 $\mu_A|\mu_B$. 因为 μ_A 和 μ_B 都首一, 所以 $\mu_A = \mu_B$. □

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 $A \nsim_s B$. □

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 $A \in \mathrm{M}_n(F)$. 则 $\dim(F[A]) = \deg(\mu_A)$ 且 A 可逆当且仅当 $\mu_{\mathcal{A}}(0) \neq 0$.

证明. (i) 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.

(ii) 类似. □

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, $A^{-1} \in F[A]$.