第二章　线性算子

1　线性映射的矩阵表示

记号. 在本节中 F 是任意域, V 和 W 是 F 上的线性空间, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基, $\epsilon_1, \ldots, \epsilon_m$ 是 W 的一组基.

设 $\phi \in \mathrm{Hom}(V, W)$, $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, $M \in F^{k\times\ell}$. 定义

$$\phi(\mathbf{v}_1, \ldots, \mathbf{v}_k) := (\phi(\mathbf{v}_1), \ldots, \phi(\mathbf{v}_k)).$$

则

$$\phi((\mathbf{v}_1, \ldots, \mathbf{v}_k)M) = (\phi(\mathbf{v}_1), \ldots, \phi(\mathbf{v}_k))M.$$

验证如下. 设 $M = (m_{i,j})_{k\times\ell}$. 则

$$(\mathbf{v}_1, \ldots, \mathbf{v}_k)M = \left(\sum_{i=1}^{k} m_{i,1}\mathbf{v}_i, \ldots, \sum_{i=1}^{k} m_{i,\ell}\mathbf{v}_i\right).$$

于是

$$\begin{aligned}
\phi((\mathbf{v}_1, \ldots, \mathbf{v}_k)M) &= \phi\left(\sum_{i=1}^{k} m_{i,1}\mathbf{v}_i,\ \ldots,\ \sum_{i=1}^{k} m_{i,\ell}\mathbf{v}_i\right) \\
&= \left(\phi\left(\sum_{i=1}^{k} m_{i,1}\mathbf{v}_i\right),\ \ldots,\ \phi\left(\sum_{i=1}^{k} m_{i,\ell}\mathbf{v}_i\right)\right) \\
&= \left(\sum_{i=1}^{k} m_{i,1}\phi(\mathbf{v}_i),\ \ldots,\ \sum_{i=1}^{k} m_{i,\ell}\phi(\mathbf{v}_i)\right) \\
&= (\phi(\mathbf{v}_1), \ldots, \phi(\mathbf{v}_k))M. \quad \square
\end{aligned}$$

1.1 线性映射的矩阵

设 $\phi \in \mathrm{Hom}(V, W)$,

$$\phi(\mathbf{e}_j) = a_{1,j}\epsilon_1 + \cdots + a_{m,j}\epsilon_m = (\epsilon_1, \ldots, \epsilon_m)\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix},$$

$j = 1, 2, \ldots, n$. 称

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

是 ϕ 在基底 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 和 $\epsilon_1, \ldots, \epsilon_m$ 下的矩阵(表示). 当 V 和 W 的基底选定后 ϕ 的矩阵是唯一的. 我们有

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

设 $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$ 和 $\phi(\mathbf{x}) = y_1\epsilon_1 + \cdots + y_m\epsilon_m$. 则

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

例 1.1 设 $V = F^n$, $W = F^m$, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 和 $\epsilon_1, \ldots, \epsilon_m$ 都是标准基. 则 ϕ 的矩阵表示是

$$(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_n)).$$

例 1.2 设 $\phi: V \longrightarrow W$ 由公式 $\forall\, \mathbf{x} \in V,\ \phi(\mathbf{x}) = \mathbf{0}_W$ 给出. 则 ϕ 在 V 和 W 的任意基底下的矩阵都是 $O_{m\times n}$.

例 1.3 设 $\phi: V \longrightarrow V$ 由公式 $\forall\, \mathbf{x} \in V,\ \phi(\mathbf{x}) = \mathbf{x}$ 给出. 则 ϕ 在 V 的任意基底 $\mathbf{e}_1, \ldots, \mathbf{e}_n; \mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵都是 E_n.

例 1.4 设 V 是 W 的子空间, $\phi: V \longrightarrow W$ 由公式 $\forall\, \mathbf{x} \in V,\ \phi(\mathbf{x}) = \mathbf{x}$ 给出. 取 W 的基是 $\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{e}_{n+1}, \ldots, \mathbf{e}_m$. 则 ϕ 的矩阵是

$$\begin{pmatrix} E_n \\ O_{(m-n)\times n} \end{pmatrix}.$$

例 1.5 线性空间 V 上的线性函数. 求 f 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n; 1$ 下的矩阵.

解. $f(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n))$. 矩阵是

$$(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)). \quad \square$$

例 1.6 设 $A \in F^{m\times n}$, $\phi: F^{n\times k} \longrightarrow F^{m\times k}$ 由公式 $\forall\, X \in F^{n\times k},\ \phi(X) = AX$ 给出. 求 ϕ 在标准基下的矩阵.

解. 对 $j = 1, 2, \ldots, k$, $\overrightarrow{\phi(X)}^{(j)} = A\vec{X}^{(j)}$. 于是,

$$\begin{pmatrix} \overrightarrow{\phi(X)}^{(1)} \\ \vdots \\ \overrightarrow{\phi(X)}^{(k)} \end{pmatrix} = \underbrace{\begin{pmatrix} A & O & \cdots & O \\ O & A & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A \end{pmatrix}_{km\times kn}}_{B} \begin{pmatrix} \vec{X}^{(1)} \\ \vdots \\ \vec{X}^{(k)} \end{pmatrix}.$$

矩阵 B 是 ϕ 在标准基下的矩阵. 此时标准基的顺序如下. 设 $E_{i.j}\in F^{n\times k}$, $i\in\{1,2,\ldots,n\}$, $j\in\{1,2,\ldots,k\}$; $L_{i,j}\in F^{m\times k}$, $i\in\{1,2,\ldots,m\}, j\in\{1,2,\ldots,k\}$ 分别是两个矩阵空间的标准基. 则 $F^{n\times k}$ 的标准基排列是

$$E_{1,1},E_{2,1},\ldots,E_{n,1},\ldots,E_{1,k},\ldots,E_{n,k}.$$

类似地, $F^{m\times k}$ 的标准基排列是

$$L_{1,1},L_{2,1},\ldots,L_{m,1},\ldots,L_{1,k},\ldots,L_{m,k}.$$

例 1.7 设 $\mathcal{D}:\mathbb{R}[x]^{(n)}\to\mathbb{R}[x]^{(n)}$ 由 $\mathcal{D}(f)=f'$ 定义. 求 $\mathcal{D}$ 在 $1,x,\ldots,x^{n-1};1,x,\ldots,x^{n-1}$ 下的矩阵.

解. 直接计算得:

$$\begin{aligned}\mathcal{D}(1,x,x^2,\ldots,x^{n-1})&=(0,1,2x,\ldots,(n-1)x^{n-2})\\&=(1,x,2x,\ldots,x^{n-1})\begin{pmatrix}0&1&0&0&\cdots&0\\0&0&2&0&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&0&\cdots&n-1\\0&0&0&0&\cdots&0\end{pmatrix}.\end{aligned}$$

1.2 线性映射在不同基底下的矩阵

定理 1.8 设 $\phi\in\mathrm{Hom}(V,W)$. 再设 $\mathbf{e}'_1,\ldots,\mathbf{e}'_n$ 是 V 的另一组基, $\epsilon'_1,\ldots,\epsilon'_m$ 是 W 的另一组基, 且

$$(\mathbf{e}'_1,\ldots,\mathbf{e}'_n)=(\mathbf{e}_1,\ldots,\mathbf{e}_n)P\quad \text{和}\quad(\epsilon'_1,\ldots,\epsilon'_m)=(\epsilon_1,\ldots,\epsilon_m)Q,$$

其中 $P \in \mathrm{GL}_n(F)$ 和 $Q \in \mathrm{GL}_m(F)$. 如果 ϕ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$; $\epsilon_1, \ldots, \epsilon_m$ 下的矩阵是 A, 则 ϕ 在 $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$; $\epsilon'_1, \ldots, \epsilon'_m$ 下的矩阵是 $Q^{-1}AP$.

证明. 由 $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)P$ 得

$$(\phi(\mathbf{e}'_1), \ldots, \phi(\mathbf{e}'_n)) = (\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_n))P.$$

于是

$$(\phi(\mathbf{e}'_1), \ldots, \phi(\mathbf{e}'_n)) = (\epsilon_1, \ldots, \epsilon_m)AP = (\epsilon'_1, \ldots, \epsilon'_m)Q^{-1}AP. \quad \square$$

上述定理说明线性映射在不同基底下的矩阵有秩相同.

定理 1.9 设 $\phi \in \mathrm{Hom}(V, W)$, $A \in F^{m\times n}$ 是 ϕ 在符号约定基底下的矩阵. 设 $B \in F^{m\times n}$. 则 B 是 ϕ 在 V 和 W 某组基下的矩阵当且仅当 $A \sim_e B$, 即 $\mathrm{rank}(A) = \mathrm{rank}(B)$.

证明. 设 B 是 ϕ 在 V 和 W 某组基下的矩阵. 根据定理 1.8, 存在 $P \in \mathrm{GL}_n(F)$ 和 $Q \in \mathrm{GL}_m(F)$ 使得 $B = Q^{-1}AP$. 于是, $A \sim_e B$.

反之, 设 $A \sim_e B$. 则存在 $M \in \mathrm{GL}_m(F)$ 和 $N \in \mathrm{GL}_n(F)$ 使得 $B = MAN$. 令基变换:

$$(\mathbf{e}'_1, \ldots, \mathbf{e}'_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)N, \quad (\epsilon'_1, \ldots, \epsilon'_n) = (\epsilon_1, \ldots, \epsilon_n)M^{-1}.$$

则 B 是 ϕ 在上述新基底下的矩阵. $\square$

定义 1.10 设 $\phi \in \mathrm{Hom}(V,W)$, A 是 ϕ 在 V 的一组基和 W 的一组基下的矩阵. 则 $\mathrm{rank}(A)$ 称为 ϕ 的秩, 记为 $\mathrm{rank}(\phi)$.

例 1.11 计算例 *1.6* 中 ϕ 的秩. 设矩阵 B 由例 *1.6* 给出. 则 $\mathrm{rank}(\phi) = \mathrm{rank}(B) = k\mathrm{rank}(A)$.

推论 1.12 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 ϕ 在基底 $\mathbf{e}_1, \ldots, \mathbf{e}_n; \epsilon_1, \ldots, \epsilon_m$ 下得矩阵是 A. 由打洞引理, 存在 $P \in \mathrm{GL}_m(F)$ 和 $Q \in \mathrm{GL}_n(F)$ 使得 $M = PAQ$. 令

$$(\mathbf{e}'_1, \ldots, \mathbf{e}'_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)Q$$

和

$$(\epsilon'_1, \ldots, \epsilon'_m) = (\epsilon_1, \ldots, \epsilon_m)P^{-1}.$$

则 ϕ 在新的基底下的矩阵是 M.

注意到:

$$\mathrm{im}(\phi) = \langle \phi(\mathbf{e}'_1), \ldots, \phi(\mathbf{e}'_n)\rangle = \langle \epsilon'_1, \ldots, \epsilon'_r\rangle.$$

故 $\dim(\mathrm{im}(\phi)) = r$. □

注解 1.13 根据推论 *1.12* 和对偶定理,

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

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 单, 所以 $\dim(\ker(\phi)) = 0$. 故 $\mathrm{rank}(\phi) = mn$.

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 $\mathrm{rank}(\phi) = \dim(W)$.

证明. (i) ϕ 单当且仅当 $\ker(\phi) = \{\mathbf{0}_V\}$ (第一章第一讲命题 2.2) 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 1.12). □

推论 1.16 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 $A \in F^{m\times n}$. 故 ϕ 单当且仅当 $\operatorname{rank}(\phi) = n$, 即 $\operatorname{rank}(A) = n$. (推论 1.15 (i)); ϕ 满当且仅当 $\operatorname{rank}(\phi) = m$, 即 $\operatorname{rank}(A) = m$. (推论 1.15 (ii)) 由 (i) 和 (ii) 直接得出 (iii). □

1.3 线性映射的复合

命题 1.17 设 $\phi \in \operatorname{Hom}(V, W)$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n; \epsilon_1, \ldots, \epsilon_m$ 下的矩阵是 A. 设 Z 是 F 上的线性空间 $\delta_1, \ldots, \delta_k$ 是 Z 的一组基, $\psi \in \operatorname{Hom}(W, Z)$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_m$; $\delta_1, \ldots, \delta_k$ 下的矩阵是 $B \in F^{k\times m}$. 则 $\psi \circ \phi \in \operatorname{Hom}(V, Z)$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$; $\delta_1, \ldots, \delta_k$ 下的矩阵是 $BA \in F^{k\times n}$.

$$\begin{tikzcd}
V \arrow[r, "\phi"] \arrow[dr, "\psi\circ\phi"'] & W \arrow[d, "\psi"] \\
 & Z
\end{tikzcd}$$

证明. 关于 $\psi \circ \phi$ 是线性映射的验证见上学期第二章第四讲命题 6.15. 计算

$$\psi \circ \phi(\mathbf{e}_j) = (\psi(\epsilon_1), \ldots, \psi(\epsilon_n))\vec{A}^{(j)} = (\delta_1, \ldots, \delta_k)B\vec{A}^{(j)},$$

$j = 1, 2, \ldots, n$. 于是

$$\begin{aligned}
(\psi \circ \phi(\mathbf{e}_1), \ldots, \psi \circ \phi(\mathbf{e}_n)) &= (\delta_1, \ldots, \delta_k)B(\vec{A}^{(1)}, \ldots, \vec{A}^{(n)}) \\
&= (\delta_1, \ldots, \delta_k)BA.
\end{aligned}$$

映射 $\psi \circ \phi$ 在选定基底下的矩阵是 BA. □

1.4 线性同构

定理 1.18 设

$$\begin{array}{rccl} \Phi: & \mathrm{Hom}(V,W) & \longrightarrow & F^{m\times n} \\ & \phi & \mapsto & A, \quad \phi \text{ 在选定基底下的矩阵} \end{array}.$$

则 Φ 是线性同构.

证明. 设 $\phi, \psi \in \mathrm{Hom}(V,W)$, $A=\Phi(\phi)$ 和 $B=\Phi(\psi)$. 再设 $\alpha, \beta \in F$. 对 $j=1,2,\ldots,n$,

$$\begin{aligned} (\alpha\phi+\beta\psi)(\mathbf{e}_j) &= \alpha\phi(\mathbf{e}_j)+\beta\psi(\mathbf{e}_j) \\ &= \alpha(\epsilon_1,\ldots\epsilon_m)\vec{A}^{(j)}+\beta(\epsilon_1,\ldots\epsilon_m)\vec{B}^{(j)} \\ &= (\epsilon_1,\ldots\epsilon_m)(\alpha\vec{A}^{(j)}+\beta\vec{B}^{(j)}) \\ &= (\epsilon_1,\ldots\epsilon_m)\vec{C}^{(j)}, \quad \text{其中 } C=\alpha A+\beta B. \end{aligned}$$

于是, $\Phi(\alpha\phi+\beta\psi)=C=\alpha A+\beta B=\alpha\Phi(\phi)+\beta\Phi(\psi)$. 映射 Φ 是线性的.

设 $A\in F^{m\times n}$, $\phi_A \in \mathrm{Hom}(V,W)$ 满足

$$\phi(\mathbf{e}_j)=(\epsilon_1,\ldots,\epsilon_m)\vec{A}^{(j)},$$

$j=1,2,\ldots,n$. 由线性映射基本定理 II 可知 ϕ 存在. 由矩阵表示的定义可知, A 是 ϕ 在选定基底下的矩阵.

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

根据定义直接验证得: $\Theta\circ\Phi(\phi)=\phi$ 且 $\Phi\circ\Theta(A)=A$. $\square$

由上述定理可知, $\dim(\mathrm{Hom}(V,W))=mn$. $\square$

1.5 对偶映射

设 $\phi\in\mathrm{Hom}(V,W)$. 则

$$\begin{aligned}\phi^*:\ W^* &\longrightarrow V^*\\ f &\mapsto f\circ\phi\end{aligned}$$

是从 W^* 到 V^* 的映射.

$$\begin{tikzcd} V \arrow[r, "\phi"] \arrow[dr, "f\circ\phi"'] & W \arrow[d, "f"] \\ & F \end{tikzcd}$$

下面我们来验证 $\phi^*\in\mathrm{Hom}(W^*,V^*)$. 设 $\alpha,\beta\in F$, $f,g\in W^*$. 根据上学期第二章第四讲命题 6.10, 我们有

$$\phi^*(\alpha f+\beta g)=(\alpha f+\beta g)\circ\phi=\alpha f\circ\phi+\beta g\circ\phi=\alpha\phi^*(f)+\beta\phi^*(g).$$

验证完毕. 我们称 ϕ^* 是 ϕ 的对偶映射.

命题 1.19 设 $\phi\in\mathrm{Hom}(V,W)$ 在基底 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\epsilon_1,\ldots,\epsilon_m$ 下的矩阵为 $A=(a_{i,j})_{m\times n}$. 则对偶映射 ϕ^* 在对偶基 $\epsilon_1^*,\ldots,\epsilon_m^*$ 和 $\mathbf{e}_1^*,\ldots,\mathbf{e}_n^*$ 下的矩阵是 A^t.

证明. 我们先计算 $\phi^*(\epsilon_i^*)$ 在 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 下的矩阵, $i=1,2,\ldots,m$. 对任意 $j\in\{1,2,\ldots,n\}$,

$$\phi^*(\epsilon_i^*)(\mathbf{e}_j)=\epsilon_i^*(\phi(\mathbf{e}_j))=\epsilon_i^*(a_{1,j}\epsilon_1+\cdots+a_{m,j}\epsilon_m)=a_{i,j}.$$

则对于 $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$,

$$\begin{aligned}\phi^*(\epsilon_i^*)(\mathbf{x}) &= a_{i,1}x_1 + \cdots + a_{i,n}x_n \\ &= a_{i,1}\mathbf{e}_1^*(\mathbf{x}) + \cdots + a_{i,n}\mathbf{e}_n^*(\mathbf{x}) \\ &= (a_{i,1}\mathbf{e}_1^* + \cdots + a_{i,n}\mathbf{e}_n^*)(\mathbf{x}).\end{aligned}$$

于是,

$$\phi^*(\epsilon_i^*) = a_{i,1}\mathbf{e}_1^* + \cdots + a_{i,n}\mathbf{e}_n^* = (\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*)\begin{pmatrix} a_{i,1} \\ \vdots \\ a_{i,n} \end{pmatrix} = (\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*)\vec{A^t}^{(i)}.$$

由此得出 A^t 是 ϕ^* 在 $\epsilon_1^*, \ldots, \epsilon_m^*$ 和 $\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*$ 下的矩阵. □

该命题的一个直接推论是 ϕ 和 ϕ^* 的秩相等.

2　线性算子代数和矩阵相似

在本节中 V 是域 F 上的 n 维线性空间, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基. 集合 $\mathrm{Hom}(V, V)$ 记为 $\mathcal{L}(V)$, 其中的元素称为线性算子. 线性算子通常用 $\mathcal{A}, \mathcal{B}, \ldots,$ 表示. 特别地 $\mathcal{O}$ 代表 V 上的零算子(映射), $\mathcal{E}$ 代表 V 上的恒等算子(映射).

设 $\mathcal{A} \in \mathcal{L}(V)$. 算子 $\mathcal{A}$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n; \mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵，简称 $\mathcal{A}$ 为在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵.

2.1 矩阵的相似

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

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

其中 $P \in \mathrm{GL}_n(F)$. 设算子 $\mathcal{A} \in \mathcal{L}(V)$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵等于 A. 根据定理 1.8, $\mathcal{A}$ 在 $\epsilon_1, \ldots, \epsilon_n$ 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给定矩阵 $A \in \mathrm{M}_n(F)$ 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*:*

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

证明. *因为这两个矩阵的迹不同, 所以它们不相似.*

例 2.4 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. $\square$

2.2 线性算子代数

我们已经知道 $(\mathcal{L}(V),+,\mathcal{O},$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 2.5 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 (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 $\mathcal{AB}$. 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 2.6 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 $1,x,\ldots,x^{n-1}$ 下的矩阵见例 *1.7*.

例 2.7 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 $\mathbf{e}_1,\ldots,\mathbf{e}_d$ 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

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

2.3 核像分解

定理 2.8 *(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.

($\Longleftarrow$) 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 $\mathbf{x} \in \ker(\mathcal{A}) \cap \operatorname{im}(\mathcal{A})$. 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 $\ker(\mathcal{A}) + \operatorname{im}(\mathcal{A})$ 是直和. 于是 $\dim(\ker(\mathcal{A}) + \operatorname{im}(\mathcal{A})) = \dim(\ker(\mathcal{A})) + \dim(\operatorname{im}(\mathcal{A})) = \dim(V)$. 我们得出 $V = \ker(\mathcal{A}) \oplus \operatorname{im}(\mathcal{A})$.

[1]袁力, 沈洁. 常州工学院学报 27 卷第二期, 2014 年 4 月.

($\Longrightarrow$) 由断言和推论 1,14 可知, 我们只要证明

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

即可. 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 $\mathrm{im}(\mathcal{A}) \subset \mathrm{im}(\mathcal{A}^2)$. $\square$

注解 2.9 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 $\ker(\mathcal{A})+\operatorname{im}(\mathcal{A})$ 是直和. 则

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

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

例 2.10 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. □

例 2.11 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.