第二章　矩阵

6.4　矩阵的乘法

命题 6.15 设 $\psi \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^s), \phi \in \mathrm{Hom}(\mathbb{R}^s, \mathbb{R}^m)$. 即

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "\psi"] \arrow[dr, "\phi\circ\psi"] & \mathbb{R}^s \arrow[d, "\phi"] \\ & \mathbb{R}^m. \end{tikzcd}$$

则

$$\phi \circ \psi \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m).$$

证明. 设 $\alpha, \beta \in \mathbb{R}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. 我们计算

$$\phi\circ\psi(\alpha\mathbf{x}+\beta\mathbf{y}) = \phi(\alpha\psi(\mathbf{x})+\beta\psi(\mathbf{y})) = \alpha\phi\circ\psi(\mathbf{x})+\beta\phi\circ\psi(\mathbf{y}). \quad \square$$

我们来推导上述 $\psi \circ \phi$ 的矩阵. 为此, 再设 $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_s$ 是 $\mathbb{R}^s$ 的标准基. 设 $A \in \mathbb{R}^{m\times s}$ 是 ϕ 在 $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_s$; $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_m$ 下的矩阵; $B \in \mathbb{R}^{s\times n}$ 是 ψ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$; $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_s$ 下的矩阵. 令 $A = (a_{i,k})_{m\times s}$ 和 $B = (b_{k,j})_{s\times n}$.

设 $C = (c_{i,j})_{m\times n} \in \mathbb{R}^{m\times n}$ 是 $\phi\circ\psi$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$; $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_m$

下的矩阵. 我们计算

$$
\begin{aligned}
C &= (\phi\circ\psi(\mathbf{e}_1),\dots,\phi\circ\psi(\mathbf{e}_n)) \quad (\text{矩阵表示的定义})\\
&= \left(\phi(\vec{B}^{(1)}),\dots,\phi(\vec{B}^{(n)})\right) \quad (\text{矩阵表示的定义})\\
&= \left(\sum_{k=1}^{s} b_{k,1}\vec{A}^{(k)},\dots,\sum_{k=1}^{s} b_{k,n}\vec{A}^{(k)}\right) \quad (\text{向量形式})\\
&= \left(\begin{pmatrix}\sum_{k=1}^{s} b_{k,1}a_{1,k}\\ \vdots\\ \sum_{k=1}^{s} b_{k,1}a_{m,k}\end{pmatrix},\dots,\begin{pmatrix}\sum_{k=1}^{s} b_{k,n}a_{1,k}\\ \vdots\\ \sum_{k=1}^{s} b_{k,n}a_{m,k}\end{pmatrix}\right) \quad (\text{坐标形式})\\
&= \begin{pmatrix}\sum_{k=1}^{s} b_{k,1}a_{1,k} & \cdots & \sum_{k=1}^{s} b_{k,n}a_{1,k}\\ \vdots & \ddots & \vdots\\ \sum_{k=1}^{s} b_{k,1}a_{m,k} & \cdots & \sum_{k=1}^{s} b_{k,n}a_{m,k}\end{pmatrix}\\
&= \left(\sum_{k=1}^{s} a_{i,k}b_{k,j}\right)_{m\times n}.
\end{aligned}
$$

故

$$c_{i,j} = \sum_{k=1}^{s} a_{i,k}b_{k,j}, \quad i=1,2,\dots,m, j=1,2,\dots,n. \qquad (1)$$

定义 6.16 设 $A=(a_{i,k})_{m\times s}\in\mathbb{R}^{m\times s}$, $B=(b_{k,j})_{s\times n}\in\mathbb{R}^{s\times n}$. 则 A 和 B 的乘积 $C=(c_{i,j})\in\mathbb{R}^{m\times n}$ 由 (1) 给出, 其中 $i=1,\dots,m$, $j=1,\dots,n$. 我们把 C 记为 AB.

注解 6.17 设 $A=(a_{i,k})_{m\times s}\in\mathbb{R}^{m\times s}$, $B=(b_{k,j})_{s\times n}\in\mathbb{R}^{s\times n}$. 乘积 AB 有不同的等价表述方式如下:

(i) (映射版) 设 $\phi_B : \mathbb{R}^n \longrightarrow \mathbb{R}^s$ 和 $\phi_A : \mathbb{R}^s \longrightarrow \mathbb{R}^m$ 分别是 A 和 B 对应的线性映射. 由上述计算可知 AB 是线性映射 $\phi_A \circ \phi_B$ 对应的矩阵. 即

$$\phi_A \circ \phi_B = \phi_{AB}.$$

(ii) (列向量版) 设

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_s \end{pmatrix} \Longrightarrow A\mathbf{v} = \sum_{k=1}^{s} v_k \vec{A}^{(k)} = \begin{pmatrix} \sum_{k=1}^{s} a_{1,k} v_k \\ \vdots \\ \sum_{k=1}^{s} a_{m,k} v_k \end{pmatrix}.$$

故

$$AB = \left(A\vec{B}^{(1)}, \ldots, A\vec{B}^{(n)} \right).$$

验证如下: 等式右侧第 i 行第 j 列处的元素是 $A\vec{B}^{(j)}$ 中的第 i 个元素. 即 $\sum_{k=1}^{s} a_{i,k} b_{k,j}$.

(iii) (行向量版) 设

$$\begin{aligned} \mathbf{w} &= (w_1, \ldots, w_s) \Longrightarrow \mathbf{w}B &&= \textstyle\sum_{k=1}^{s} w_k \vec{B}_k \\ & && = \left(\textstyle\sum_{k=1}^{s} w_k b_{k,1}, \ldots, \sum_{k=1}^{s} w_k b_{k,n} \right). \end{aligned}$$

故

$$AB = \begin{pmatrix} \vec{A}_1 B \\ \vdots \\ \vec{A}_m B \end{pmatrix}.$$

验证如下: 等式右侧第 i 行第 j 列处的元素是 $\vec{A}_i \vec{B}$ 中的第 j 个元素. 即 $\sum_{k=1}^{s} a_{i,k} b_{k,j}$.

(iv) (行-列向量版) 设:

$$\mathbf{u} = (u_1, \dots, u_k), \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_s \end{pmatrix} \Longrightarrow \mathbf{uv} = u_1 v_1 + \cdots + u_s v_s.$$

故

$$AB = \begin{pmatrix} \vec{A}_1 \vec{B}^{(1)} & \cdots & \vec{A}_1 \vec{B}^{(n)} \\ \vdots & \ddots & \vdots \\ A_m \vec{B}^{(1)} & \cdots & \vec{A}_m \vec{B}^{(n)} \end{pmatrix} = \left(\vec{A}_i \vec{B}^{(j)} \right)_{m \times n}.$$

验证如下*:* 等式右侧第 i 行第 j 列处的元素是

$$\vec{A}_i \vec{B}^{(j)} = (a_{i,1}, \dots, a_{i,s}) \begin{pmatrix} b_{1,j} \\ \vdots \\ b_{s,j} \end{pmatrix} = \sum_{k=1}^{s} a_{i,k} b_{k,j}.$$

例 6.18 设

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

计算 AB.

解.

$$AB = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 6 & -1 \\ 4 & 15 & -1 \end{pmatrix}.$$

注意到 BA 没有定义.

例 6.19 设

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

计算 AB 和 BA.

解.

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

而

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

注意到 $AB \neq BA$.

例 6.20 设 $A = (a_{i,j}) \in \mathbb{R}^{m\times n}$. 证明:

$$\mathrm{diag}(\lambda_1, \ldots, \lambda_m)A = \begin{pmatrix} \lambda_1 \vec{A}_1 \\ \vdots \\ \lambda_m \vec{A}_m \end{pmatrix}$$

和

$$A\mathrm{diag}(\lambda_1, \ldots, \lambda_n) = (\lambda_1 \vec{A}^{(1)}, \ldots, \lambda_n \vec{A}^{(n)}).$$

证明. 矩阵 $B = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)A$ 中第 i 行第 j 列处的元素等于

$$(0, \ldots, 0, \lambda_i, 0, \ldots 0)\vec{A}^{(j)} = \lambda_i a_{i,j} \Longrightarrow \vec{B}_i = \lambda_i \vec{A}_i.$$

矩阵 $C = A\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ 中第 i 行第 j 列处的元素等于

$$\vec{A}_i \begin{pmatrix} 0 \\ \vdots \\ \lambda_j \\ \vdots \\ 0 \end{pmatrix} = \lambda_j a_{i,j} \Longrightarrow \vec{C}^{(j)} = \lambda_j \vec{A}^{(j)}.$$

上例说明 $\mathrm{diag}_m(\lambda)A = A\mathrm{diag}_n(\lambda) = \lambda A$. 特别地,

$$O_m A = AO_n = O_{m\times n} \quad \text{和} \quad E_m A = AE_n = A.$$

矩阵的乘法大大化简了我们对线性方程组和线性映射的表示. 设 $A = (a_{i,j}) \in \mathbb{R}^{m\times n}$. 则齐次线性方程组

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = 0 \\ \qquad\qquad \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = 0 \end{cases}$$

除了可以写为向量形式 $\sum_{j=1}^n x_j \vec{A}^{(j)} = \mathbf{0}_m$ 外, 还可以写为乘法形式

$$A\mathbf{x} = \mathbf{0}_m,$$

其中 $\mathbf{x} = (x_1, \ldots, x_n)^t$. 类似地, 设 $\mathbf{b} \in \mathbb{R}^m$. 则以 $(A|\mathbf{b})$ 为增广矩阵的线性方程组, 除了可以写为向量形式 $\sum_{j=1}^n x_j \vec{A}^{(j)} = \mathbf{b}$ 外, 还可以写为乘法形式

$$A\mathbf{x} = \mathbf{b}.$$

设 $\phi:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ 是线性映射, 它在标准基下的矩阵是 $A\in\mathbb{R}^{m\times n}$. 则对任意 $\mathbf{x}\in\mathbb{R}^n$,

$$\phi(\mathbf{x})=A\mathbf{x}.$$

命题 6.21 设 $A\in\mathbb{R}^{m\times s}, B\in\mathbb{R}^{s\times k}, C\in\mathbb{R}^{k\times n}$. 则

$$(AB)C=A(BC).$$

证明. 考虑线性映射:

$$\phi_A:\mathbb{R}^s\longrightarrow\mathbb{R}^m,\ \ \phi_B:\mathbb{R}^k\longrightarrow\mathbb{R}^s,\ \ \phi_C:\mathbb{R}^n\longrightarrow\mathbb{R}^k.$$

我们有交换图:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "\phi_C"] \arrow[dr, "\phi_B\circ\phi_C=\phi_{BC}"'] & \mathbb{R}^k \arrow[d, "\phi_B"] \arrow[dr, "\phi_A\circ\phi_B=\phi_{AB}"] & \\
& \mathbb{R}^s \arrow[r, "\phi_A"] & \mathbb{R}^m.
\end{tikzcd}$$

根据第一章第二讲定理 4.11, $(\phi_A\circ\phi_B)\circ\phi_C=\phi_A\circ(\phi_B\circ\phi_C)$. 由矩阵乘法的定义可知

$$(\phi_A\circ\phi_B)\circ\phi_C=\phi_{AB}\circ\phi_C=\phi_{(AB)C}$$

和

$$\phi_A\circ(\phi_B\circ\phi_C)=\phi_A\circ\phi_{BC}=\phi_{A(BC)}.$$

于是, $\phi_{(AB)C}=\phi_{A(BC)}$. 根据上一讲定理 6.5, $(AB)C=A(BC)$.

□

命题 6.22 设 $A, B\in\mathbb{R}^{m\times n}, C\in\mathbb{R}^{n\times k}, D\in\mathbb{R}^{k\times m}$, $\lambda, \mu\in\mathbb{R}$. 则 $(\lambda A+\mu B)C = \lambda(AC)+\mu(BC)$ 和 $D(\lambda A+\mu B) = \lambda(DA)+\mu(DB)$.

证明. 考虑线性映射:

$$\phi_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m,\ \phi_B : \mathbb{R}^n \longrightarrow \mathbb{R}^m,\ \phi_C : \mathbb{R}^k \longrightarrow \mathbb{R}^n.$$

根据上一讲命题 6.10,

$$(\lambda\phi_A + \mu\phi_B) \circ \phi_C = (\lambda\phi_A) \circ \phi_C + (\mu\phi_B) \circ \phi_C.$$

即 $\phi_{(\lambda A+\mu B)C} = \phi_{\lambda AC+\mu BC}$. 根据上一讲定理 6.5,

$$(\lambda A + \mu B)C = \lambda(AC) + \mu(BC).$$

再考虑线性映射 $\phi_D : \mathbb{R}^m \longrightarrow \mathbb{R}^k$. 根据上一讲命题 6.10,

$$\phi_D \circ (\lambda\phi_A + \mu\phi_B) = \lambda(\phi_D \circ \phi_A) + \mu(\phi_D \circ \phi_B).$$

即 $\phi_{D(\lambda A+\mu B)} = \phi_{\lambda DA+\mu DB}$. 根据定理 **??**,

$$D(\lambda A + \mu B) = \lambda(DA) + \mu(DB). \quad \square$$

命题 6.23 设 $A \in \mathbb{R}^{m\times s}, B \in \mathbb{R}^{s\times n}$. 则 $(AB)^t = B^tA^t$.

证明. 设 $A = (a_{i,k})_{m\times s}$, $B = (b_{k,j})_{s\times n}$, $C = (c_{i,j})_{m\times n} = AB$. 再设 $A^t = (a'_{k,i})_{s\times m}$ 和 $B^t = (b'_{j,k})_{n\times s}$. 则 $a'_{k,i} = a_{i,k}$ 和 $b'_{j,k} = b_{k,j}$. 令 $D = (d_{j,i})_{n\times m} = B^tA^t$. 我们计算

$$d_{j,i} = \sum_{k=1}^{s} b'_{j,k}a'_{k,i} = \sum_{k=1}^{s} a_{i,k}b_{k,j} = c_{i,j}.$$

故 $C^t = D$. $\square$

定理 6.24 设 $A \in \mathbb{R}^{m\times s}, B \in \mathbb{R}^{s\times n}$. 则

(i) $\mathrm{rank}(AB) \le \min(\mathrm{rank}(A), \mathrm{rank}(B))$;

(ii) $\mathrm{rank}(A) + \mathrm{rank}(B) - s \le \mathrm{rank}(AB)$ *(Sylvester 不等式)*.

证明. 考虑线性映射 $\phi_A : \mathbb{R}^s \longrightarrow \mathbb{R}^m$ 和 $\phi_B : \mathbb{R}^n \longrightarrow \mathbb{R}^s$. 我们有下列交换图:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "\phi_B"] \arrow[dr, "\phi_{AB}=\phi_A\circ\phi_B"'] & \mathbb{R}^s \arrow[d, "\phi_A"] \\
& \mathbb{R}^m.
\end{tikzcd}$$

(i) 直接计算得:

$$\begin{aligned}
\mathrm{rank}(AB) &= \dim(\mathrm{im}(\phi_{AB})) \quad (\text{第二章第四讲命题 } 6.6\ (\mathrm{i})) \\
&= \dim(\phi_A(\mathrm{im}(\phi_B)) \quad (\mathrm{im}(\phi_{AB}) = \phi_A(\mathrm{im}(\phi_B)) \\
&\le \dim(\mathrm{im}(\phi_B)) \quad (\text{第二章第四讲命题 } 5.17\ (\mathrm{i})) \\
&= \mathrm{rank}(B) \quad (\text{第二章第四讲命题 } 6.6\ (\mathrm{i})).
\end{aligned}$$

类似地,

$$\begin{aligned}
\mathrm{rank}(AB) &= \dim(\phi_A(\mathrm{im}(\phi_B)) \quad (\text{见上面的计算}) \\
&\le \dim(\phi_A(\mathbb{R}^s)) \quad (\mathrm{im}(\phi_B) \subset \mathbb{R}^s) \\
&= \dim(\mathrm{im}(\phi_A)) \quad (\phi_A(\mathbb{R}^s) = \mathrm{im}(\phi_A)) \\
&= \mathrm{rank}(A) \quad (\text{第二章第四讲命题 } 6.6\ (\mathrm{i})).
\end{aligned}$$

(i) 成立.

(ii) 设 $\mathbf{v}_1, \ldots, \mathbf{v}_d$ 是 $\ker(\phi_A) \cap \operatorname{im}(\phi_B)$ 的一组基, $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_d$ 是 $\operatorname{im}(\phi_B)$ 的一组基. 则 $\operatorname{rank}(B) = d + k$.

断言. $\operatorname{rank}(AB) \geq k$.

断言的证明. 设 $\alpha_1\phi_A(\mathbf{u}_1) + \cdots + \alpha_k\phi_A(\mathbf{u}_k) = \mathbf{0}_m$, 其中 $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$. 则 $\phi_A(\mathbf{u}) = \mathbf{0}_m$, 其中 $\mathbf{u} = \alpha_1\mathbf{u}_1 + \cdots + \alpha_k\mathbf{u}_k$. 故 $\mathbf{u} \in \operatorname{im}(\phi_B) \cap \ker(\phi_A)$. 于是, $\mathbf{u} \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_d \rangle$. 因为 $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_d$ 线性无关, 所以 $\alpha_1 = \cdots = \alpha_k = 0$. 于是, $\phi_A(\mathbf{u}_1), \ldots, \phi_A(\mathbf{u}_k)$ 是 $\operatorname{im}(\phi_{AB})$ 中的一组线性无关向量. 故 $\dim(\operatorname{im}(\phi_{AB})) \geq k$. 从而, $\operatorname{rank}(AB) \geq k$. 断言成立.

由对偶定理(线性映射版)可知,

$$\operatorname{rank}(A) = s - \dim(\ker(\phi_A)) \leq s - d.$$

于是

$$\operatorname{rank}(B) + \operatorname{rank}(A) - s \leq d + k + s - d - s = k.$$

根据上述断言, (ii) 成立. □

注解 6.25 上述证明中的断言可以加强为 $\operatorname{rank}(AB) = k$.

推论 6.26 设 $A \in \mathbb{R}^{m\times n}$, B 是 m 阶满秩方阵, C 是 n 阶满秩方阵. 则 $\operatorname{rank}(A) = \operatorname{rank}(BA) = \operatorname{rank}(AC)$.

证明. 由上述定理的两个不等式可知

$$\operatorname{rank}(A) + \operatorname{rank}(B) - m \leq \operatorname{rank}(BA) \leq \min(\operatorname{rank}(B), \operatorname{rank}(A)).$$

因为 $\mathrm{rank}(B) = m$, 所以 $\mathrm{rank}(A) \le \mathrm{rank}(BA) \le \mathrm{rank}(A)$. 故 $\mathrm{rank}(BA) = \mathrm{rank}(A)$. 另一个等式可以类似地证明. □

7 方阵

实数上所有方阵的集合记为 $\mathrm{M}_n(\mathbb{R})$.

7.1 $\mathrm{M}_n(\mathbb{R})$ 上的运算

对于任意 $A, B, C \in \mathrm{M}_n(\mathbb{R})$, $\lambda, \mu \in \mathbb{R}$, 我们有

(A1) $A + B = B + A$;

(A2) $(A + B) + C = A + (B + C)$;

(A3) $A + O = A$, 其中 O 是 n 阶零矩阵;

(A4) $A + (-A) = O$.

(M1) $(AB)C = A(BC)$;

(M2) $AE = EA = A$, 其中 E 是 n 阶单位矩阵.

(S1) $(\lambda\mu)A = \lambda(\mu A)$;

(S2) $1A = A$.

(AM) $A(B + C) = AB + AC$; $(A + B)C = AC + BC$.

(AS) $(\lambda+\mu)A=\lambda A+\mu A$; $\lambda(A+B)=\lambda A+\lambda B$.

(MS) $(\lambda A)(\mu B)=\lambda(A(\mu B))=(\lambda\mu)(AB)$.

我们把 $\mathrm{M}_n(\mathbb{R})$ 称为 $\mathbb{R}$ 上的 n 阶矩阵代数.

记号. 设 $k\in\mathbb{Z}^+$, $A\in\mathrm{M}_n(\mathbb{R})$. 根据乘法结合律, 我们定义:

$$A^k:=\underbrace{A\cdots A}_{k}.$$

此外 $A^0:=E$. 可直接验证 $A^kA^\ell=A^{k+\ell}$ 和 $A^{k\ell}=(A^k)^\ell$.

例 7.1 设 $A,B\in\mathrm{M}_n(\mathbb{R})$. 展开 $(A+B)^2$ 和 $(A+B)(A-B)$.

解. $(A+B)^2=(A+B)(A+B)=A^2+AB+BA+B^2$.
$(A+B)(A-B)=A^2+BA-AB-B^2$.

只有当 $AB=BA$ 时, 我们才有

$$(A+B)^2=A^2+2AB+B^2 \quad 和 \quad (A+B)(A-B)=A^2-B^2.$$

定义 7.2 设 $A\in\mathrm{M}_n(\mathbb{R})$. 如果 $A^t=A$, 则称 A 是对称矩阵. 如果 $A^t=-A$, 则称 A 是斜对称矩阵. 如果存在 $k\in\mathbb{Z}^+$ 使得 $A^k=O$. 则称 A 是幂零矩阵. 如果 $A^2=A$, 则称 A 是幂等矩阵.

注解 7.3 设 $A=(a_{i,j})\in\mathrm{M}_n(\mathbb{R})$. 则 A 对称当且仅当 $a_{i,j}=a_{j,i}$, $i,j\in\{1,2,\ldots,n\}$. 而 A 斜对称当且仅当

$a_{i.i} = 0$ 且 $a_{i,j} = -a_{j,i}$, 其中 $i, j \in \{1, 2, \ldots, n\}$ 且 $i \neq j$. 我们来验证关于斜对称的必要充分条件如下

$$
\begin{aligned}
& A = -A^t \Longleftrightarrow A + A^t = O \\
& \Longleftrightarrow a_{i,j} + a_{j,i} = 0,\ i, j \in \{1, 2, \ldots, n\} \\
& \Longleftrightarrow 2a_{i,i} = 0, a_{i,j} + a_{j,i} = 0,\ i, j \in \{1, 2, \ldots, n\}, i \neq j \\
& \Longleftrightarrow a_{i,i} = 0, a_{i,j} + a_{j,i} = 0,\ i, j \in \{1, 2, \ldots, n\}, i \neq j \quad (2 \neq 0).
\end{aligned}
$$

例 7.4 设

$$
A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{和} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & {}^{‘}0 \\ 0 & 9 & 0 \end{pmatrix}.
$$

可直接验证 A 是幂零的和 B 是幂等的.

例 7.5 设 $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. 则 $A^k = \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)$.

例 7.6 设

$$
A = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix}.
$$

对任意 $n \in \mathbb{N}$, 计算 A^n.

解. 显然有 $A^0 = E_2$ 和 $A^1 = A$. 我们计算

$$
A^2 = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} = \begin{pmatrix} a^2 & c(a+b) \\ 0 & b^2 \end{pmatrix}.
$$

和

$$A^3 = \begin{pmatrix} a^2 & c(a+b) \\ 0 & b^2 \end{pmatrix} \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} = \begin{pmatrix} a^3 & c(a^2+ab+b^2) \\ 0 & b^3 \end{pmatrix}.$$

下面我们证明: 当 $n>0$ 时,

$$A^n = \begin{pmatrix} a^n & cd_n \\ 0 & b^n \end{pmatrix},$$

其中 $d_n = \sum_{i=0}^{n-1} a^i b^{n-1-i}$.

当 $n=1,2,3$ 时结论成立. 设 $n>3$ 且 $n-1$ 时结论成立. 则

$$A^n = A^{n-1}A = \begin{pmatrix} a^{n-1} & cd_{n-1} \\ 0 & b^{n-1} \end{pmatrix} \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} = \begin{pmatrix} a^n & ca^{n-1}+cbd_{n-1} \\ 0 & b^n \end{pmatrix}.$$

而

$$ca^{n-1}+cbd_{n-1} = c(a^{n-1}+b\sum_{i=0}^{n-2} a^i b^{n-2-i}) = c\sum_{i=0}^{n-1} a^i b^{n-1-i} = cd_n.$$

结论成立.

我们可以更简洁地把结果写成:

$$A^n = \begin{pmatrix} a^n & c\frac{a^n-b^n}{a-b} \\ 0 & b^n \end{pmatrix}, a \neq b, \quad \text{和} \quad A^n = \begin{pmatrix} a^n & nca^{n-1} \\ 0 & a^n \end{pmatrix}, a=b.$$

事实上, 这也包括了 $n=1$ 的情形.

7.2 交换不变量与中心元

由第二章第四讲例 6.19 可知, 对 $A, B \in \mathrm{M}_n(\mathbb{R})$, $\mathrm{rank}(AB)$ 可能不等于 $\mathrm{rank}(BA)$. 换言之, 秩关于矩阵乘法不是交换不变量.

定义 7.7 设 $A = (a_{i,j}) \in \mathrm{M}_n(\mathbb{R})$. 我们称 $\sum_{i=1}^n a_{i,i}$ 是 A 的迹 *(trace)*. 记为 $\mathrm{tr}(A)$.

可直接验证: 对任意 $\alpha, \beta \in \mathbb{R}, A, B \in \mathrm{M}_n(\mathbb{R})$,

$$\mathrm{tr}(\alpha A + \beta B) = \alpha \mathrm{tr}(A) + \beta \mathrm{tr}(B).$$

命题 7.8 设 $A = (a_{i,j}), B = (b_{i,j}) \in \mathrm{M}_n(\mathbb{R})$. 则

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

证明. 设 $C = (c_{ij}) = AB$ 和 $D = (d_{i,j}) = BA$. 则

$$c_{i,i} = \sum_{k=1}^n a_{i,k} b_{k,i} \quad \text{和} \quad d_{i,i} = \sum_{k=1}^n b_{i,k} a_{k,i}.$$

于是,

$$\mathrm{tr}(C) = \sum_{i=1}^n \sum_{k=1}^n a_{i,k} b_{k,i} \quad \text{和} \quad \mathrm{tr}(D) = \sum_{i=1}^n \sum_{k=1}^n b_{i,k} a_{k,i}.$$

故 $\mathrm{tr}(C) = \mathrm{tr}(D)$. □

例 7.9 证明*: 对任意 $A, B \in \mathrm{M}_n(\mathbb{R})$, $AB - BA \neq E$.*

证明. 我们计算 $\mathrm{tr}(AB - BA) = \mathrm{tr}(AB) - \mathrm{tr}(BA) = 0$ *(命题 7.8)*. 但 $\mathrm{tr}(E) = n > 0$. 故 $AB - BA \neq E$. □

定义 7.10 设 $C \in \mathrm{M}_n(\mathbb{R})$. 如果对任意 $A \in \mathrm{M}_n(\mathbb{R})$, 我们有 $AC = CA$. 则称 C 是中心元.

定理 7.11 设 $C \in \mathrm{M}_n(\mathbb{R})$. 则 C 是中心元当且仅当 C 是数乘矩阵.

为了证明每个中心元是数乘矩阵, 我们需要下述引理.

引理 7.12 *(搬运工引理)* 对任意 $i, j \in \{1, \ldots, k\}$, 设 $E_{i,j}^{(k)}$ 是在 i 行 j 列处的元素等于 1, 而其它元素都等于零的 k 阶方阵. 则对于 $A \in \mathbb{R}^{m\times n}$,

$$E_{i,j}^{(m)} A = \begin{pmatrix} O_{(i-1)\times n} \\ \vec{A}_j \\ O_{(m-i)\times n} \end{pmatrix} \quad \text{和} \quad AE_{i,j}^{(n)} = (O_{m\times(j-1)}, \vec{A}^{(i)}, O_{m\times(n-j)}).$$

证明. 根据列向量乘法公式, $AE_{i,j}^{(n)}$ 中除第 j 列外都是 $\mathbf{0}_n$. 而第 j 列是 $A\mathbf{e}_i{=}\vec{A}^{(i)}$. 故第二个等式成立. 由此可知

$$(E_{i,j}^{(m)} A)^t = A^t E_{j,i}^{(m)} = (O_{n\times(i-1)}, \vec{A}_j^t, O_{n\times(m-i)}).$$

对上式再次转置得到

$$(E_{i,j}^{(m)} A) = (O_{n\times(i-1)}, \vec{A}_j^t, O_{n\times(m-i)})^t = \begin{pmatrix} O_{(i-1)\times n} \\ \vec{A}_j \\ O_{(m-i)\times n} \end{pmatrix}.$$

第一个等式成立. □

定理 7.11 的证明. 设 $C = \lambda E$. 则对任意 $A \in \mathrm{M}_n(\mathbb{R})$, $CA = \lambda A$ 且 $AC = \lambda A$. 故 $CA = AC$.

反之, 设 $C = (c_{i,j}) \in \mathrm{M}_n(\mathbb{R})$ 是中心元. 对任意 $i, j \in \{1, \ldots, n\}$, $E_{i,j}^{(n)} C = C E_{i,j}^{(n)}$. 于是

$$\begin{pmatrix} O_{(i-1)\times n} \\ \vec{C}_j \\ O_{(n-i)\times n} \end{pmatrix} = (O_{n\times(j-1)}, \vec{C}^{(i)}, O_{n\times(n-j)}).$$

比较等式两侧矩阵的第 i 行可知, 当 $k \neq j$ 时, $c_{j,k} = 0$; 当 $k = j$ 时, $c_{j,j} = c_{i,i}$. 于是, C 是数乘矩阵. □

7.3 可逆元

定义 7.13 设 $A \in \mathrm{M}_n(\mathbb{R})$. 如果存在 $B \in \mathrm{M}_n(\mathbb{R})$ 使得 $AB = BA = E$, 则称 A 是可逆矩阵. 此时, B 称为 A 的一个逆矩阵.

定理 7.14 *设 $A \in \mathrm{M}_n(\mathbb{R})$. 则 A 可逆当且仅当 A 满秩.*

证明. 设 $B \in \mathrm{M}_n(\mathbb{R})$ 使得 $AB = E$. 根据定理 6.24 (i), $\mathrm{rank}(A) \geq \mathrm{rank}(E) = n$. 故 $\mathrm{rank}(A) = n$.

反之, 设 A 满秩. 则对任意 $\mathbf{b} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ 有解(第二章第三讲推论 4.3). 设 $\mathbf{v}_j$ 是方程组 $A\mathbf{x} = \mathbf{e}_j$ 的解, $j = 1, 2, \ldots, n$. 令 $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. 则

$$AB = (A\mathbf{v}_1, \ldots, A\mathbf{v}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) = E.$$

由第二章第三讲例 3.15, $\text{rank}(A^t) = n$. 由上述证明可知存在 $C \in \mathrm{M}_n(\mathbb{R})$ 使得 $A^tC = E$. 故 $C^tA = E$ (见第二章第四讲命题 6.23). 由此得出, $C^tAB = B$. 从而,

$$C^t(AB) = C^tE = C^t = B.$$

我们得到 $BA = E$. 于是, A 可逆. □

注解 7.15 *上述定理也可以用线性映射证明如下: 如果 A 可逆, 则存在 $B \in \mathrm{M}_n(\mathbb{R})$ 使得 $BA = AB = E$. 注意到 ϕ_A, ϕ_B 都是从 $\mathbb{R}^n$ 到自身的映射. 于是, $\phi_A \circ \phi_B = \phi_E$ 和 $\phi_B \circ \phi_A = \phi_E$. 因为 ϕ_E 是恒同映射, 所以 ϕ_A 是双射. 故 A 满秩. (第二章第四讲命题 6.10 (iii)).*

反之, 同样的命题蕴含 ϕ_A 是双射. 由第七周作业题 4 可知, ϕ_A^{-1} 是线性映射. 故存在 $B \in \mathrm{M}_n(\mathbb{R})$, 使得 $\phi_A^{-1} = \phi_B$. 我们得出 $\phi_A \circ \phi_B = \phi_E$ 和 $\phi_B \circ \phi_A = \phi_E$. 于是, $AB = BA = E$.

推论 7.16 *矩阵 $A \in \mathrm{M}_n(\mathbb{R})$ 可逆当且仅当存在 $B \in \mathrm{M}_n(\mathbb{R})$ 使得 $AB = E$ 或 $BA = E$.*

证明. 这是因为 $AB = E$ 或 $BA = E$ 都可推出 A 满秩. □

注解 7.17 *上述推论也可以根据第二章第三讲推论 5.15 推出. 这是因为 $\phi_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ 是单射当且仅当它是满射当且仅当它是双射.*

命题 7.18 设 $A \in \mathrm{M}_m(\mathbb{R})$. 如果 A 可逆, 则存在唯一的矩阵 $B \in \mathrm{M}_n(\mathbb{R})$ 使得 $AB = E$ 或 $BA = E$.

证明. 设 $B, C \in \mathrm{M}_n(\mathbb{R})$ 使得 $AB = E$ 和 $AC = E$. 由上述推论可知, 存在 $D \in \mathrm{M}_n(\mathbb{R})$ 使得 $DA = E$. 故 $DAB = D$. 从而 $B = D$. 类似有 $C = D$. 我们得到 $B = C$. 另一半结论类似可证. □

由上述推论可知, 可逆矩阵的逆是唯一的. 我们把可逆矩阵 A 的逆矩阵记为 A^{-1}. 可逆矩阵的逆的唯一性也可以由双射逆的唯一性直接推出(第一章第二讲命题 4.13).