第二章　矩阵

6.4　矩阵运算与秩

命题 6.10 设 $A, B \in \mathbb{R}^{m\times n}$. 证明*:*

$$\operatorname{rank}(A+B) \leq \operatorname{rank}(A) + \operatorname{rank}(B).$$

证明. 因为 $A + B = (\vec{A}^{(1)} + \vec{B}^{(1)}, \ldots, \vec{A}^{(n)} + \vec{B}^{(n)})$ 且

$$\vec{A}^{(1)} + \vec{B}^{(1)}, \ldots, \vec{A}^{(n)} + \vec{B}^{(n)} \in V_c(A) + V_c(B),$$

所以 $V_c(A+B) \subset V_c(A) + V_c(B)$. 故

$$\dim V_c(A+B) \leq \dim(V_c(A)+V_c(B)) \leq \dim V_c(A) + \dim V_c(B),$$

即 $\operatorname{rank}(A+B) \leq \operatorname{rank}(A) + \operatorname{rank}(B)$. □

定理 6.11 设 $A \in \mathbb{R}^{m\times s}, B \in \mathbb{R}^{s\times n}$. 则

(i) $\operatorname{rank}(AB) \leq \min(\operatorname{rank}(A), \operatorname{rank}(B))$;

(ii) $\operatorname{rank}(A) + \operatorname{rank}(B) - s \leq \operatorname{rank}(AB)$

(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:

$$\begin{aligned}\operatorname{rank}(AB) &= \dim(\operatorname{im}(\phi_{AB})) \quad (\text{命题 } 6.6\ (\text{i}))\\ &= \dim(\phi_A(\operatorname{im}(\phi_B)) \quad (\operatorname{im}(\phi_{AB}) = \phi_A(\operatorname{im}(\phi_B))\\ &\leq \dim(\operatorname{im}(\phi_B)) \quad (\text{命题 } 5.13\ (\text{i}))\\ &= \operatorname{rank}(B) \quad (\text{命题 } 6.6\ (\text{i})).\end{aligned}$$

类似地,

$$\begin{aligned}\operatorname{rank}(AB) &= \dim(\phi_A(\operatorname{im}(\phi_B)) \quad (\text{见上面的计算})\\ &\leq \dim(\phi_A(\mathbb{R}^s)) \quad (\operatorname{im}(\phi_B) \subset \mathbb{R}^s)\\ &= \dim(\operatorname{im}(\phi_A)) \quad (\phi_A(\mathbb{R}^s) = \operatorname{im}(\phi_A))\\ &= \operatorname{rank}(A) \quad (\text{命题 } 6.6\ (\text{i})).\end{aligned}$$

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

推论 6.13 *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 $\operatorname{rank}(B) = m$, 所以 $\operatorname{rank}(A) \leq \operatorname{rank}(BA) \leq \operatorname{rank}(A)$. 故 $\operatorname{rank}(BA) = \operatorname{rank}(A)$. 另一个等式可以类似地证明. □

例 6.14 *上述推论不用 Sylvester 秩不等式证明.*

注意到我们有三个线性映射: $\phi_A : \mathbb{R}^n \to \mathbb{R}^m$, $\phi_B : \mathbb{R}^m \to \mathbb{R}^m$, $\phi_C : \mathbb{R}^n \to \mathbb{R}^n$, *其中* ϕ_B *和* ϕ_C *都是双射 (上一讲命题 5.21).*

因为

$$\operatorname{im}(\phi_{AC}) = \phi_A(\operatorname{im}(\phi_C)) \overset{\phi_C\ \text{满}}{=} \phi_A(\mathbb{R}^n) = \operatorname{im}(\phi_A).$$

所以 $\dim(\operatorname{im}(\phi_{AC})) = \dim(\operatorname{im}(\phi_A))$. 故 $\operatorname{rank}(AC) = \operatorname{rank}(A)$.

显然 $\ker(\phi_A) \subset \ker(\phi_{BA})$. 设 $\mathbf{x} \in \ker(\phi_{BA})$. 则

$$\phi_B(\phi_A(\mathbf{x})) = \mathbf{0}_m.$$

因为 ϕ_B 单, 所以 $\phi_A(\mathbf{x}) = \mathbf{0}_m$. 故 $\mathbf{x} \in \ker(\phi_A)$. 于是, $\ker(A) = \ker(\phi_{BA})$.

根据对偶定理: $\dim(\ker(\phi_A)) + \dim(\operatorname{im}(\phi_A)) = n$ 和 $\dim(\ker(\phi_{BA})) + \dim(\operatorname{im}(\phi_{BA})) = n$. 所以

$$\dim(\operatorname{im}(\phi_A)) = \dim(\operatorname{im}(\phi_{BA})).$$

故 $\operatorname{rank}(A) = \operatorname{rank}(BA)$. $\square$

例 6.15 设 $A \in \mathbb{R}^{m\times n}$. 则 $\operatorname{rank}(A) = 1$当且仅当存在 $B \in \mathbb{R}^{m\times 1}$ 和 $C \in \mathbb{R}^{1\times n}$ 都非零, 使得 $A = BC$.

证明. 设 $A = BC$. 根据定理 *6.11 (i)*, $\operatorname{rank}(A) \leq 1$. 在根据 *(ii)*, $\operatorname{rank}(A) \geq 1 + 1 - 1 = 1$. 故 $\operatorname{rank}(A) = 1$. 反之, 设 $\operatorname{rank}(A) = 1$. 不妨设 $\vec{A}^{(1)} \neq \mathbf{0}_m$. 则存在 $\lambda_2, \ldots, \lambda_m \in \mathbb{R}$ 使得

$$\vec{A}^{(j)} = \lambda_j \vec{A}^{(1)}.$$

则

$$A = \vec{A}^{(1)}(1, \lambda_2, \ldots, \lambda_n). \quad \square$$

6.5 对应定理

定理 6.16 设 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 是从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射的集合. 定义

$$\begin{array}{rccc} \Phi: & \mathbb{R}^{m\times n} & \longrightarrow & \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \\ & A & \mapsto & \phi_A \end{array},$$

其中 ϕ_A 代表由 $\mathbf{e}_j \mapsto \vec{A}^{(j)}$ 确定的线性映射. 则 Φ 是双射且对任意 $A, B \in \mathbb{R}^{m\times n}$ 和 $\lambda \in \mathbb{R}$,

$$\Phi(A+B) = \Phi(A) + \Phi(B) \quad \text{和} \quad \Phi(\lambda A) = \lambda\Phi(A).$$

证明. 由线性映射基本定理可知, Φ 是单射. 因为每个线性映射都有在标准基下的矩阵, 所以 Φ 是满射. 注意到

$$\Phi(A+B) \overset{\Phi\text{ 的定义}}{=} \phi_{A+B} \overset{\text{加法的定义}}{=} \phi_A+\phi_B \overset{\Phi\text{ 的定义}}{=} \Phi(A)+\Phi(B)$$

和

$$\Phi(\lambda A) \overset{\Phi\text{ 的定义}}{=} \phi_{\lambda A} \overset{\text{数乘的定义}}{=} \lambda\phi_A \overset{\Phi\text{ 的定义}}{=} \lambda\Phi(A). \quad \square$$

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 $A^k A^\ell = A^{k+\ell}$ 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

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

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

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

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

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 7.6 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 7.7 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

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

于是,

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

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

例 7.8 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 7.7). 但 $\mathrm{tr}(E) = n > 0$. 故 $AB - BA \neq E$. □

定义 7.9 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 7.10 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 7.11 (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 $A \in \mathbb{R}^{m \times n}$,

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

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 7.10 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 $i, j \in \{1, \ldots, n\}$, $E_{i,j}^{(n)}C = CE_{i,j}^{(n)}$. 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 7.12 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.

设 $BA = AB = E$ 且 $CA = AC = E$. 则

$$CAB = C \implies (CA)B = C \implies EB = C \implies B = C.$$

故上述定义中 B 是唯一的. 我们把 B 称为可逆矩阵的逆矩阵, 记为 A^{-1}.

定理 7.13 设 $A \in \mathrm{M}_n(\mathbb{R})$. 则 A 可逆当且仅当 A 满秩.

证明. 设 A 可逆, 则 $AA^{-1}=E$. 根据定理 6.11 (i), $\mathrm{rank}(A) \geq \mathrm{rank}(AA^{-1}) = \mathrm{rank}(E) = n$. 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, $\mathrm{rank}(A^t) = n$. 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 7.14 *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. 由第九周作业题 *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 7.15 *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 7.16 *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 7.17 *设 $A, B \in \mathrm{M}_n(\mathbb{R})$ 都可逆. 则*

(i) AB 可逆且它的逆是 $B^{-1}A^{-1}$;

(ii) A^{-1} 可逆且它的逆是 $(A^{-1})^{-1} = A$;

(iii) A^t 可逆且其逆是 $(A^{-1})^t$.

证明. (i) $AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = E$. 类似可验证 $(B^{-1}A^{-1})AB = E$.

(ii) 这是因为 $AA^{-1} = A^{-1}A = E$.

(iii) 由第二章第四讲命题 6.23 可知，

$$A^t(A^{-1})^t = (A^{-1}A)^t = E^t = E.$$

类似地, $(A^{-1})^tA^t = E$. $\square$

注解 7.18 上述命题中第一个结论实际上是逆映射的穿衣脱衣规则*(第一章第二讲命题 4.14 (ii))*, 而第二个结论对应着第一章第二讲命题 *4.14 (i)*.

例 7.19 设 $A \in \mathrm{M}_n(\mathbb{R})$. 则线性方程组 $A\mathbf{x} = \mathbf{b}$ 确定当且仅当 A 可逆. 此时它的唯一解是 $\mathbf{x} = A^{-1}\mathbf{b}$.

例 7.20 设 $A \in \mathrm{M}_n(\mathbb{R})$ 幂零. 证明 $E - A$ 可逆.

证明. 设 $A^k = O$, 其中 $k > 0$. 则

$$E = E - A^k = (E - A)(E + A + \cdots + A^{k-1}).$$

故 $(E - A)^{-1} = E + A + \cdots + A^{k-1}$. $\square$

例 7.21 设

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}).$$

证明: 对 $m \in \mathbb{Z}^+$,

$$A^m = \begin{pmatrix} f_{m-1} & f_m \\ f_m & f_{m+1} \end{pmatrix},$$

其中 $f_0 = 0, f_1 = 1, f_2 = 1, f_{m+1} = f_m + f_{m-1}$.

证明. $m=1$ 时, 结论显然. 设 $m>1$ 且 $m-1$ 时结论成立. 则

$$A^m = A^{m-1}A = \begin{pmatrix} f_{m-2} & f_{m-1} \\ f_{m-1} & f_m \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} f_{m-1} & f_{m-1}+f_{m-2} \\ f_m & f_m+f_{m-1} \end{pmatrix} = \begin{pmatrix} f_{m-1} & f_m \\ f_m & f_{m+1} \end{pmatrix}. \quad \square$$

设

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}$$

和

$$B = \begin{pmatrix} -\frac{\lambda_2}{5} & \frac{1}{5} \\ -\sqrt{5}\lambda_1 & \sqrt{5} \end{pmatrix}.$$

不难计算

$$B^{-1} = \begin{pmatrix} \sqrt{5} & -\frac{1}{5} \\ \sqrt{5}\lambda_1 & -\frac{\lambda_2}{5} \end{pmatrix} \quad 和 \quad A = B^{-1}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}B.$$

验证:

$$\begin{pmatrix} \sqrt{5} & -\frac{1}{5} \\ \sqrt{5}\lambda_1 & -\frac{\lambda_2}{5} \end{pmatrix}\begin{pmatrix} -\frac{\lambda_2}{5} & \frac{1}{5} \\ -\sqrt{5}\lambda_1 & \sqrt{5} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{5}}{5}(\lambda_1-\lambda_2) & 0 \\ 0 & \frac{\sqrt{5}}{5}(\lambda_1-\lambda_2) \end{pmatrix} = E_2$$

和

$$
\begin{aligned}
&\begin{pmatrix} \sqrt{5} & -\frac{1}{5} \\ \sqrt{5}\lambda_1 & -\frac{\lambda_2}{5} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} -\frac{\lambda_2}{5} & \frac{1}{5} \\ -\sqrt{5}\lambda_1 & \sqrt{5} \end{pmatrix} \\
=& \begin{pmatrix} \sqrt{5}\lambda_1 & -\frac{1}{5}\lambda_2 \\ \sqrt{5}\lambda_1^2 & -\frac{\lambda_2^2}{5} \end{pmatrix} \begin{pmatrix} -\frac{\lambda_2}{5} & \frac{1}{5} \\ -\sqrt{5}\lambda_1 & \sqrt{5} \end{pmatrix} \\
=& \begin{pmatrix} 0 & \frac{\sqrt{5}}{5}(\lambda_1 - \lambda_2) \\ \frac{\sqrt{5}}{5}\lambda_1\lambda_2(\lambda_2 - \lambda_1) & \frac{\sqrt{5}}{5}(\lambda_1^2 - \lambda_2^2) \end{pmatrix} = A.
\end{aligned}
$$

注意到: 对任意 $C \in \mathrm{M}_n(\mathbb{R})$,

$$\left(B^{-1}CB\right)^m = B^{-1}C^mB.$$

故

$$A^m = B^{-1}\mathrm{diag}(\lambda_1^m, \lambda_2^m)B.$$

于是,

$$f_m = \left(\sqrt{5}\lambda_1^m, -\frac{1}{5}\lambda_2^m\right)\vec{B}^{(2)} = \frac{\sqrt{5}}{5}(\lambda_1^m - \lambda_2^m).$$

于是, *Fibonacci* 序列的闭形式是

$$f_m = \frac{\sqrt{5}}{5}\left(\left(\frac{1+\sqrt{5}}{2}\right)^m - \left(\frac{1-\sqrt{5}}{2}\right)^m\right).$$

注意到 $f_0 = 0, f_1 = 1, f_2 = 1, f_3 = 2, f_4 = 3, \ldots, f_{10} = 55$.
但 $f_{50} = 12586269025$. 渐近公式是

$$\lim_{m\to\infty} \frac{f_m}{\left(\frac{1+\sqrt{5}}{2}\right)^m} = \frac{\sqrt{5}}{5} \Longrightarrow f_m \sim \left(\frac{1+\sqrt{5}}{2}\right)^m.$$

小结. *Fibonacci* 序列:

$$\underbrace{f_0 = 0,\ f_1 = 1,\ f_{m+1} = f_m + f_{m-1}}_{\text{递归公式}},$$

$$\underbrace{f_m = \frac{\sqrt{5}}{5}\left(\left(\frac{1+\sqrt{5}}{2}\right)^m - \left(\frac{1-\sqrt{5}}{2}\right)^m\right)}_{\text{闭形式}},$$

和

$$\underbrace{f_m \sim \left(\frac{1+\sqrt{5}}{2}\right)^m}_{\text{渐进公式}}.$$

8 矩阵的初等等价

定义 8.1 设 $A, B \in \mathbb{R}^{m\times n}$. 如果存在 m 阶可逆方阵 P 和 n 阶可逆方阵 Q 使得 $A = PBQ$. 则称 A 与 B 初等等价. 记为 $A \sim_e B$.

我们来验证 $\sim_e$ 是等价关系. 对任意 $A \in \mathbb{R}^{m\times n}$, 我们有 $A = E_m A E_n$. 于是, $A \sim_e A$. 自反性成立.

再设 $B \in \mathbb{R}^{m\times n}$ 满足 $A \sim_e B$. 则存在 m 阶可逆方阵 B 和 n 阶可逆方阵 Q 使得 $A = PBQ$. 则 $B = P^{-1}AQ^{-1}$. 故 $B \sim_e A$. 对称性成立.

设 $C \in \mathbb{R}^{m\times n}$ 且 $A \sim_e B$ 和 $B \sim_e C$. 则存在 m 阶可逆方阵 P 和 n 阶可逆方阵 Q 使得 $A = PBQ$ 和 $B = SCT$.

故 $A=(PS)C(TQ)$. 因为可逆方阵的积仍可逆, 所以 $A\sim_e C$. 传递性成立.

我们将证明

定理 8.2 设 $A,B\in\mathbb{R}^{m\times n}$. 则 $A\sim_e B\Longleftrightarrow \operatorname{rank}(A)=\operatorname{rank}(B)$.

为此, 我们需要用矩阵乘法来解释初等变换.

定义 8.3 设 $i,j\in\{1,2,\ldots,n\}$.

(i) $F_{i,j}^{(n)}$ 是把 E_n 中第 i 行和第 j 行互换的得到的矩阵. 称之为第一类初等矩阵.

(ii) 设 $\alpha\in\mathbb{R}$ 且 $i\neq j$. 则 $F_{i,j}^{(n)}(\alpha)$ 是把 E_n 中第 j 行通乘 α 加到第 i 行得到的矩阵. 称之为第二类初等矩阵.

(iii) 设 $\lambda\in\mathbb{R}$ 且 $\lambda\neq 0$. 则 $F_{i}^{(n)}(\lambda)$ 是把 E_n 中第 i 行通乘 λ 得到的矩阵. 称之为第三类类初等矩阵.

这三类矩阵统称为 n 阶初等矩阵.

注解 8.4 可直接验证 $(F_{i,j}^{(n)})^2=E_n$, $F_{i,j}^{(n)}(\alpha)F_{i,j}^{(n)}(-\alpha)=E_n$, 和 $F_{i}^{(n)}(\lambda)F_{i}^{(n)}(\lambda^{-1})=E_n$. 故初等矩阵都是可逆的, 且它们的逆也是初等矩阵.

我们可以通过搬运工引理(上一讲引理 7.12)中定义矩阵 $E_{i,j}^{(k)}$ 表示初等矩阵如下:

$$F_{i,j}^{(n)}=E_n-E_{i,i}^{(n)}-E_{j,j}^{(n)}+E_{i,j}^{(n)}+E_{j,i}^{(n)}, \tag{1}$$

$$F_{i,j}^{(n)}(\alpha) = E_n + \alpha E_{i,j}^{(n)}, \tag{2}$$

$$F_i^{(n)}(\lambda) = E_n + (\lambda - 1)E_{i,i}^{(n)}. \tag{3}$$

设 $A \in \mathbb{R}^{m\times n}$. 根据 (1),

$$F_{i,j}^{(m)}A = A - E_{i,i}^{(m)}A - E_{j,j}^{(m)}A + E_{i,j}^{(m)}A + E_{j,i}^{(m)}A.$$

根据搬运工引理, $F_{i,j}^{(m)}A$ 是把 A 中 i, j 两行对换得到的矩阵. 根据 (2),

$$F_{i,j}^{(m)}(\alpha)A = A + \alpha E_{i,j}^{(m)}A.$$

同理, $F_{i,j}^{(m)}(\alpha)A$ 是把 A 中第j 行乘以 α 后加到第 i 行得到的矩阵. 根据 (3),

$$F_i^{(m)}(\lambda)A = A + (\lambda - 1)E_{i,i}^{(m)}A.$$

从而得到把 A 中第 i 行通乘 λ 的矩阵.

类似地, $AF_{i,j}^{(n)}$, $AF_{i,j}^{(n)}(\alpha)$ 和 $AF_i^{(n)}(\lambda)$ 分别是把 A 中 i, j 两行对换, 把 A 中第 i 列乘以 α 后加到第 j 列, 和把 A 中第 i 列通乘 λ 后得到的矩阵.

引理 8.5 *(打洞引理) 设 $A \in \mathbb{R}^{m\times n}$. 则存在可逆矩阵 $P \in \mathrm{M}_m(\mathbb{R})$ 和 $Q \in \mathrm{M}_n(\mathbb{R})$, 其中 P 和 Q 都是初等矩阵的乘积, 使得*

$$PAQ = \begin{pmatrix} E_r & O_{r\times(n-r)} \\ O_{(m-r)\times r} & O_{(m-r)\times(n-r)} \end{pmatrix},$$

且 $\mathrm{rank}(A) = r$.

证明. 根据第一章第一讲命题 2.3 和第三类初等行变换, 存在若干个 m 阶初等矩阵, 使得它们的积 P 满足

$$PA=\begin{pmatrix} 0\cdots 0 & 1 & * & \cdots & * & * & * & \cdots & * & * & *\cdots * \\ 0\cdots 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & * & *\cdots * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0\cdots 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & *\cdots * \\ 0\cdots 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0\cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0\cdots 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0\cdots 0 \end{pmatrix},$$

其中共有 r 行非零. 则存在若干个 n 阶第一类初等矩阵, 使得它们的积 Q_1 满足

$$PAQ_1=\begin{pmatrix} 1 & * & \cdots & * & * & * & \cdots & * & * & *\cdots * \\ 0 & 1 & \cdots & * & * & * & \cdots & * & * & *\cdots * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & * & *\cdots * \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0\cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0\cdots 0 \end{pmatrix},$$

进而存在若干个 n 阶第一类初等矩阵, 使得它们的积 Q_2

满足

$$PAQ_1Q_2 = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 0 \cdots 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 0 \cdots 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 0 \cdots 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 0 \cdots 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 0 \cdots 0 \end{pmatrix}.$$

再令 $Q = Q_1Q_2$ 即可, 这是因为初等矩阵之积必然可逆. 根据第二章第五讲推论 6.27, $\mathrm{rank}(A) = r$. □

注解 8.6 打洞引理也可写成:

设 $A \in \mathbb{R}^{m\times n}$ 和 $\mathrm{rank}(A) = r$. 则

$$A = P \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q,$$

其中 $P \in \mathrm{M}_m(\mathbb{R})$ 可逆, $Q \in \mathrm{M}_n(\mathbb{R})$ 可逆.

例 8.7 设 $A \in \mathbb{R}^{m\times n}$ 且 $\mathrm{rank}(A)=r>0$. 则存在 $B \in \mathbb{R}^{m\times r}$ 和 $C \in \mathbb{R}^{r\times n}$ 满足 $A = BC$ 和 $\mathrm{rank}(B) = \mathrm{rank}(C) = r$.

证明. 根据打洞引理, 存在 m 阶和 n 阶可逆矩阵 P 和 Q 使得

$$A = P \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q.$$

直接验证得

$$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n} = \begin{pmatrix} E_r \\ O \end{pmatrix}_{m\times r} (E_r, O)_{r\times n}.$$

令

$$B = P \begin{pmatrix} E_r \\ O \end{pmatrix}_{m\times r} \quad 和 \quad C = (E_r, O)_{r\times n} Q$$

即可.

推论 8.8 可逆矩阵是若干初等矩阵之积.

证明. 设 A 可逆. 则 A 满秩. 由打洞引理的证明可知, 存在初等矩阵 $P_1 \ldots, P_k, Q_1, \ldots, Q_\ell$ 使得

$$P_k \cdots P_1 A Q_1 \cdots Q_\ell = E.$$

故

$$A = (P_k \cdots P_1)^{-1}(Q_1 \cdots Q_\ell)^{-1} = P_1^{-1} \cdots P_k^{-1} Q_\ell^{-1} \cdots Q_1^{-1}.$$

因为初等矩阵的逆还是初等矩阵, 所以推论成立. □