

第二章 矩阵

6 矩阵的运算

在本节中, 设 $\mathbf{e}_1, \dots, \mathbf{e}_n; \boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_m$ 分别是 \mathbb{R}^n 和 \mathbb{R}^m 的标准基.

6.1 线性映射在标准基下的矩阵表示

考虑线性映射 $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$. 对 $j = 1, 2, \dots, n$, 设

$$\phi(\mathbf{e}_j) = \sum_{i=1}^m a_{i,j} \boldsymbol{\epsilon}_i.$$

根据第二章第三讲定理 5.9, ϕ 由矩阵

$$A = (\phi(\mathbf{e}_1), \dots, \phi(\mathbf{e}_n)) = (a_{i,j})_{m \times n}$$

唯一确定. 我们称 A 是线性映射 ϕ 在标准基下的矩阵的表示, 简称 ϕ 的矩阵. 记为 A_ϕ .

设 $\mathbf{x} = (x_1, \dots, x_n)^t$. 则

$$\begin{aligned}\phi(\mathbf{x}) &= \phi\left(\sum_{j=1}^n x_j \mathbf{e}_j\right) \\ &= \sum_{j=1}^n x_j \phi(\mathbf{e}_j) \\ &= \sum_{j=1}^n x_j \vec{A}^{(j)} \quad (\text{向量版的公式}) \\ &= \begin{pmatrix} \sum_{j=1}^n a_{1,j} x_j \\ \vdots \\ \sum_{j=1}^n a_{m,j} x_j \end{pmatrix} \quad (\text{坐标版的公式})\end{aligned}$$

例 6.1 设 $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 是零映射, 即 $\phi(\mathbf{e}_j) = \mathbf{0}_m$, $j = 1, 2, \dots, n$. 则

$$A_\phi = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{m \times n}.$$

称之为 $m \times n$ 阶零矩阵, 记为 $O_{m \times n}$. 当 $m = n$ 时, 简称为 n 阶零方阵, 记为 O_n 或 O .

例 6.2 设 $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是数乘映射, 即 $\phi(\mathbf{e}_j) = \lambda \mathbf{e}_j$,

$j = 1, 2, \dots, n$, 其中 $\lambda \in \mathbb{R}$. 则

$$A_\phi = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

称之为 n 阶数乘(方)阵, 有时记为 $\text{diag}_n(\lambda)$. 记 $\text{diag}_n(1)$ 为 E_n . 称为 n 阶单位方阵. 它对应的线性映射是恒同映射.

例 6.3 设 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是线性函数, $f(\mathbf{e}_1) = \alpha_1, \dots, f(\mathbf{e}_n) = \alpha_n$. 则 $A_f = (\alpha_1, \dots, \alpha_n)$. 设 $\mathbf{x} = (x_1, \dots, x_n)^t$. 则

$$f(\mathbf{x}) = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

例 6.4 设 $T_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 是旋转, 即

$$T_\theta(\mathbf{e}_1) = \cos(\theta)\mathbf{e}_1 + \sin(\theta)\mathbf{e}_2 \quad \text{和} \quad T_\theta(\mathbf{e}_2) = -\sin(\theta)\mathbf{e}_1 + \cos(\theta)\mathbf{e}_2.$$

则 T_θ 的矩阵是

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

于是

$$T_\theta \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} + y \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta)x - \sin(\theta)y \\ \sin(\theta)x + \cos(\theta)y \end{pmatrix}.$$

定理 6.5 设 $\text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 是从 \mathbb{R}^n 到 \mathbb{R}^m 的线性映射的集合. 定义

$$\begin{aligned} \Phi : \text{Hom}(\mathbb{R}^n, \mathbb{R}^m) &\longrightarrow \mathbb{R}^{m \times n} \\ \phi &\longmapsto A_\phi \end{aligned} .$$

则 Φ 是双射且其逆是

$$\begin{aligned} \Psi : \mathbb{R}^{m \times n} &\longrightarrow \text{Hom}(\mathbb{R}^n, \mathbb{R}^m) \\ A &\longmapsto \phi_A \end{aligned} .$$

证明. 设 $\phi \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$. $\Psi \circ \Phi(\phi) = \Psi(A_\phi)$. 注意到

$$\Psi(A_\phi)(\mathbf{e}_j) = \vec{A}_\phi^{(j)} = \phi(\mathbf{e}_j), \quad j = 1, 2, \dots, n.$$

根据第二章第三讲定理 5.9, $\Psi(A_\phi) = \phi$. 于是,

$$\Psi \circ \Phi = \text{id}_{\text{Hom}(\mathbb{R}^n, \mathbb{R}^m)}.$$

设 $A \in \mathbb{R}^{m \times n}$. 则 $\Phi \circ \Psi(A) = \Phi(\phi_A)$. 注意到矩阵 $\Phi(\phi_A)$ 的第 j 列是 $\phi_A(\mathbf{e}_j) = \vec{A}^{(j)}$, $j = 1, 2, \dots, n$. 于是, $\Phi(\phi_A) = A$. 由此可知 $\Phi \circ \Psi = \text{id}_{\mathbb{R}^{m \times n}}$. \square

下面的命题是用矩阵来描述线性映射.

命题 6.6 设 $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ 是线性映射, 它在标准基下的矩阵是 $A \in \mathbb{R}^{m \times n}$.

(i) $\text{im}(\phi) = V_c(A)$, 从而 $\dim(\text{im}(\phi)) = \text{rank}(A)$. 特别地, ϕ 是满射当且仅当 A 行满秩.

(ii) $\ker(\phi)$ 是 A 对应的齐次线性方程组的解空间, 从而 $\dim(\ker(\phi)) = n - \text{rank}(A)$. 特别地, ϕ 是单射当且仅当 A 列满秩.

(iii) ϕ 是双射当且仅当 $m = n$ 且 A 满秩.

证明. 根据定理 6.5, $\phi = \phi_A$.

(i) 由第二章第三讲例 5.12 可知, $\text{im}(\phi) = V_c(A)$. 故

$$\dim(V_c(A)) = \text{rank}(A).$$

注意到 ϕ 满当且仅当 $V_c(A) = \mathbb{R}^m$, 即 $V_c(A) = \mathbb{R}^m$ 当且仅当 $\dim V_c(A) = m$.

(ii) 由第二章第三讲例 5.12 可知, $\ker(\phi)$ 是 A 对应的齐次线性方程组的解空间. 根据对偶定理和 (i), 我们有

$$\dim(\ker(\phi)) = n - \text{rank}(A).$$

根据第二章第三讲命题 5.8, ϕ 单当且仅当 $\ker(\phi) = \{\mathbf{0}_n\}$, 即上述解空间是 $\mathbf{0}_n$. 根据第二章第三讲引理 4.2, $\text{rank}(A) = n$.

(iii) 是 (i) 和 (ii) 的直接推论. \square

例 6.7 设 $\mathbf{e}_1, \dots, \mathbf{e}_4$ 是 \mathbb{R}^4 的标准基, $\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \boldsymbol{\epsilon}_3$ 是 \mathbb{R}^3 的一

组基. 线性映射 $\phi: \mathbb{R}^4 \rightarrow \mathbb{R}^3$ 由

$$\begin{cases} \phi(\mathbf{e}_1) = 2\boldsymbol{\epsilon}_1 - \boldsymbol{\epsilon}_3 \\ \phi(\mathbf{e}_2) = \boldsymbol{\epsilon}_2 - \boldsymbol{\epsilon}_3 \\ \phi(\mathbf{e}_3) = 4\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2 - 3\boldsymbol{\epsilon}_3 \\ \phi(\mathbf{e}_4) = 2\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2 - 2\boldsymbol{\epsilon}_3 \end{cases}$$

计算

(i) 计算 ϕ 在 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4; \boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \boldsymbol{\epsilon}_3$ 下的矩阵;

(ii) 计算 $\ker(\phi)$ 和 $\text{im}(\phi)$ 的维数;

(iii) 分别计算 $\ker(\phi)$ 和 $\text{im}(\phi)$ 的一组基底.

解. (i) 由定义可知:

$$A_\phi = (\phi(\mathbf{e}_1), \phi(\mathbf{e}_2), \phi(\mathbf{e}_3), \phi(\mathbf{e}_4)) = \begin{pmatrix} 2 & 0 & 4 & 2 \\ 0 & 1 & 1 & 1 \\ -1 & -1 & -3 & -2 \end{pmatrix}.$$

(ii) 利用初等行变换得

$$A_\phi \longrightarrow \begin{pmatrix} 1 & 1 & 3 & 2 \\ 2 & 0 & 4 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 3 & 2 \\ 0 & -2 & -2 & -2 \\ 0 & 1 & 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 3 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

因为 $\text{rank}(A_\phi) = 2$, 所以 $\dim(\text{im}(\phi)) = 2$. 由对偶定理可知, $\dim(\ker(\phi)) = 4 - 2 = 2$.

(iii) $\ker(\phi)$ 对应的齐次线性方程组是

$$\begin{cases} x_1 + x_2 + 3x_3 + 2x_4 = 0 \\ x_2 + x_3 + x_4 = 0. \end{cases} \iff \begin{cases} x_1 + x_2 = -3x_3 - 2x_4 \\ x_2 = -x_3 - x_4. \end{cases}$$

于是, $\ker(\phi)$ 的一组基是

$$\begin{pmatrix} -2 \\ -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{和} \quad \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

像空间 $\text{im}(\phi)$ 的基是 A_ϕ 中的极大线性无关组. 因为 $\dim(\text{im}(\phi)) = 2$, 所以取 A_ϕ 中任意两个线性无关的列向量即可. 例如 $\text{im}(\phi)$ 的一组基是 $\vec{A}_\phi^{(1)}, \vec{A}_\phi^{(2)}$.

例 6.8 利用方程版的对偶定理证明映射版的对偶定理.

证明. 设 $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 是线性映射, 它在标准基下的矩阵是 $A \in \mathbb{R}^{m \times n}$. 再设 H_A 是以 A 为系数矩阵的齐次线性方程组. 根据方程版对偶定理(第二章第三讲定理 4.6)

$$\dim(\text{sol}(H_A)) + \text{rank}(A) = n.$$

根据命题 6.6 (i) 和 (ii),

$$\dim(\ker(\phi)) + \dim(\text{im}(\phi)) = n.$$

即映射版对偶定理(第二章第三讲定理 5.13)成立.

6.2 线性映射的运算

定义 6.9 设 ϕ 和 ψ 是从集合 S 到 \mathbb{R}^m 的映射, $\lambda \in \mathbb{R}$. 令:

$$\begin{array}{ccc} \phi + \psi : S \longrightarrow \mathbb{R}^m & & \lambda\phi : S \longrightarrow \mathbb{R}^m \\ s \mapsto \phi(s) + \psi(s) & \text{和} & s \mapsto \lambda\phi(s). \end{array}$$

分别称为映射的加法与映射的数乘运算.

根据 \mathbb{R}^m 中线性运算的性质, 我们可直接验证映射的加法和数乘满足交换律和结合律且这两个运算满足分配律.

命题 6.10 (i) 设 $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 是线性映射, $\psi_1, \psi_2 : S \rightarrow \mathbb{R}^n$ 是映射, $\lambda_1, \lambda_2 \in \mathbb{R}$. 则

$$\phi \circ (\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1(\phi \circ \psi_1) + \lambda_2(\phi \circ \psi_2).$$

(ii) 设 $\phi_1, \phi_2 : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\psi : S \rightarrow \mathbb{R}^n$ 是映射, $\lambda_1, \lambda_2 \in \mathbb{R}$. 则

$$(\lambda_1\phi_1 + \lambda_2\phi_2) \circ \psi = \lambda_1(\phi_1 \circ \psi) + \lambda_2(\phi_2 \circ \psi).$$

证明. (i) 设 $s \in S$. 则

$$\begin{aligned} \phi \circ (\lambda_1\psi_1 + \lambda_2\psi_2)(s) &= \phi((\lambda_1\psi_1 + \lambda_2\psi_2)(s)) \quad (\text{复合的定义}) \\ &= \phi(\lambda_1\psi_1(s) + \lambda_2\psi_2(s)) \quad (\text{定义 6.9}) \\ &= \lambda_1\phi \circ \psi_1(s) + \lambda_2\phi \circ \psi_2(s) \quad (\phi \text{ 线性}) \\ &= (\lambda_1\phi \circ \psi_1 + \lambda_2\phi \circ \psi_2)(s) \quad (\text{定义 6.9}). \end{aligned}$$

故 $\phi \circ (\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1(\phi \circ \psi_1) + \lambda_2(\phi \circ \psi_2)$.

(ii) 设 $s \in S$. 则

$$\begin{aligned}(\lambda_1\phi_1 + \lambda_2\phi_2) \circ \psi(s) &= (\lambda_1\phi_1 + \lambda_2\phi_2)(\psi(s)) \quad (\text{复合的定义}) \\ &= \lambda_1\phi_1 \circ \psi(s) + \lambda_2\phi_2 \circ \psi(s) \quad (\text{定义 6.9}) \\ &= (\lambda_1\phi_1 \circ \psi + \lambda_2\phi_2 \circ \psi)(s) \quad (\text{定义 6.9}).\end{aligned}$$

故 $(\lambda_1\phi_1 + \lambda_2\phi_2) \circ \psi = \lambda_1(\phi_1 \circ \psi) + \lambda_2(\phi_2 \circ \psi)$. \square

命题 6.11 设 ϕ 和 ψ 是从 \mathbb{R}^n 到 \mathbb{R}^m 的线性映射, $\lambda \in \mathbb{R}$. 则 $\phi + \psi$ 和 $\lambda\phi$ 也是线性映射.

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

$$\begin{aligned}(\phi + \psi)(\mathbf{x} + \mathbf{y}) &= \phi(\mathbf{x} + \mathbf{y}) + \psi(\mathbf{x} + \mathbf{y}) \quad (\text{映射加法的定义}) \\ &= \phi(\mathbf{x}) + \phi(\mathbf{y}) + \psi(\mathbf{x}) + \psi(\mathbf{y}) \quad (\text{线性映射的定义}) \\ &= (\phi + \psi)(\mathbf{x}) + (\phi + \psi)(\mathbf{y}) \quad (\text{映射加法的定义}).\end{aligned}$$

设 $\alpha \in \mathbb{R}$. 则

$$\begin{aligned}(\phi + \psi)(\alpha\mathbf{x}) &= \phi(\alpha\mathbf{x}) + \psi(\alpha\mathbf{x}) \quad (\text{映射加法的定义}) \\ &= \alpha\phi(\mathbf{x}) + \alpha\psi(\mathbf{x}) \quad (\text{线性映射的定义}) \\ &= \alpha(\phi + \psi)(\mathbf{x}) \quad (\text{映射加法的定义}).\end{aligned}$$

于是, $\phi + \psi$ 是线性映射. 类似地,

$$\begin{aligned}(\lambda\phi)(\mathbf{x} + \mathbf{y}) &= \lambda\phi(\mathbf{x} + \mathbf{y}) \quad (\text{映射数乘的定义}) \\ &= \lambda(\phi(\mathbf{x}) + \phi(\mathbf{y})) \quad (\text{线性映射的定义}) \\ &= (\lambda\phi)(\mathbf{x}) + (\lambda\phi)(\mathbf{y}) \quad (\text{映射数乘的定义}).\end{aligned}$$

$$\begin{aligned}(\lambda\phi)(\alpha\mathbf{x}) &= \lambda\phi(\alpha\mathbf{x}) \quad (\text{映射数乘的定义}) \\ &= \lambda\alpha\phi(\mathbf{x}) \quad (\text{线性映射的定义}) \\ &= \alpha(\lambda\phi)(\mathbf{x}) \quad (\text{映射数乘的定义}). \quad \square\end{aligned}$$

6.3 矩阵的线性运算

设 $\phi, \psi \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, 它们在标准基下的表示分别是 $A = (a_{i,j})_{m \times n}$ 和 $B = (b_{i,j})_{m \times n}$. 则

$$(\phi + \psi)(\mathbf{e}_j) = \phi(\mathbf{e}_j) + \psi(\mathbf{e}_j) = \vec{A}^{(j)} + \vec{B}^{(j)}, \quad j = 1, 2, \dots, n.$$

由此我们定义两个矩阵 A 和 B 的和

$$A + B = (\vec{A}^{(1)} + \vec{B}^{(1)}, \dots, \vec{A}^{(n)} + \vec{B}^{(n)}).$$

等价地, $A + B = (a_{i,j} + b_{i,j})_{m \times n}$.

设 $\lambda \in \mathbb{R}$. 则

$$(\lambda\phi)(\mathbf{e}_j) = \lambda\phi(\mathbf{e}_j) = \lambda\vec{A}^{(j)}, \quad j = 1, 2, \dots, n.$$

由此我们定义矩阵的数乘

$$\lambda A = (\lambda \vec{A}^{(1)}, \dots, \lambda \vec{A}^{(n)}).$$

等价地, $\lambda A = (\lambda a_{i,j})_{m \times n}$.

由矩阵加法和数乘的定义与定理 6.5 可知: 对任意 $\lambda \in \mathbb{R}$, $A, B \in \mathbb{R}^{m \times n}$, 我们有 $\phi_A + \phi_B = \phi_{A+B}$ 和 $\lambda \phi_A = \phi_{\lambda A}$.

可直接验证矩阵的加法满足交换律和结合律, 且对于任意 $A, B \in \mathbb{R}^{m \times n}$, $A + O_{m \times n} = A$ 和 $A + (-A) = O_{m \times n}$. 进而, 对于任意 $\lambda, \mu \in \mathbb{R}$, $(\lambda\mu)A = \lambda(\mu A)$. 分配律:

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

例 6.12 设

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \end{pmatrix} \quad \text{和} \quad B = \begin{pmatrix} 2 & 1 & -3 \\ 2 & 2 & 1 \end{pmatrix}.$$

计算 $3A - 2B$.

解.

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

例 6.13 设 $A, B \in \mathbb{R}^{m \times n}$. 证明:

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

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

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

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

推论 6.14 设 $\text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 从 \mathbb{R}^n 到 \mathbb{R}^m 的线性映射的集合. 定义

$$\begin{aligned} \Phi : \text{Hom}(\mathbb{R}^n, \mathbb{R}^m) &\longrightarrow \mathbb{R}^{m \times n} \\ \phi &\longmapsto A_\phi \end{aligned} .$$

则 Φ 是双射且其逆是

$$\begin{aligned} \Psi : \mathbb{R}^{m \times n} &\longrightarrow \text{Hom}(\mathbb{R}^n, \mathbb{R}^m) \\ A &\longmapsto \phi_A \end{aligned} .$$

进而, 对任意 $\lambda, \mu \in \mathbb{R}$, $\phi, \psi \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, $A, B \in \mathbb{R}^{m \times n}$, 我们有:

$$\Phi(\lambda\phi + \mu\psi) = \lambda\Phi(\phi) + \mu\Phi(\psi)$$

和

$$\Psi(\lambda A + \mu B) = \lambda\Psi(A) + \mu\Psi(B).$$

证明. 根据定理 6.5, 我们只要验证 Φ 保持线性运算. 为此, 我们计算

$$\begin{aligned}
 \Phi(\lambda\phi + \mu\psi) &= A_{\lambda\phi + \mu\psi} \quad (\Phi \text{ 的定义}) \\
 &= A_{\lambda\phi} + A_{\mu\psi} \quad (\text{矩阵加法的定义}) \\
 &= \lambda A_{\phi} + \mu A_{\psi} \quad (\text{矩阵数乘的定义}) \\
 &= \lambda\Phi(\phi) + \mu\Phi(\psi) \quad (\Phi \text{ 的定义}).
 \end{aligned}$$

类似地,

$$\begin{aligned}
 \Psi(\lambda A + \mu B) &= \phi_{\lambda A + \mu B} \quad (\Psi \text{ 的定义}) \\
 &= \phi_{\lambda A} + \phi_{\mu B} \quad (\text{矩阵加法的定义}) \\
 &= \lambda\phi_A + \mu\phi_B \quad (\text{矩阵数乘的定义}) \\
 &= \lambda\Psi(A) + \mu\Psi(B) \quad (\Psi \text{ 的定义}).
 \end{aligned}$$

6.4 矩阵的乘法

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

$$\begin{array}{ccc}
 \mathbb{R}^n & \xrightarrow{\psi} & \mathbb{R}^s \\
 & \searrow \phi \circ \psi & \downarrow \phi \\
 & & \mathbb{R}^m.
 \end{array}$$

则

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

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 $\boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_s$ 是 \mathbb{R}^s 的标准基. 设 $A \in \mathbb{R}^{m \times s}$ 是 ϕ 在 $\boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_s; \boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_m$ 下的矩阵; $B \in \mathbb{R}^{s \times n}$ 是 ψ 在 $\mathbf{e}_1, \dots, \mathbf{e}_n; \boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_s$ 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 $\mathbf{e}_1, \dots, \mathbf{e}_n; \boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_m$ 下的矩阵. 我们计算

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

故

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

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 $\phi_B : \mathbb{R}^n \rightarrow \mathbb{R}^s$ 和 $\phi_A : \mathbb{R}^s \rightarrow \mathbb{R}^m$ 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

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

故

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

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

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

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:

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

故

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

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:

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

和

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

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

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

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

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

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

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

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

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

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 $\mathbf{x} = (x_1, \dots, x_n)^t$. 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 $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 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:

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

我们有交换图:

$$\begin{array}{ccc}
 \mathbb{R}^n & \xrightarrow{\phi_C} & \mathbb{R}^k \\
 & \searrow \phi_{B \circ \phi_C} = \phi_{BC} & \downarrow \phi_B \\
 & & \mathbb{R}^s & \xrightarrow{\phi_A} & \mathbb{R}^m.
 \end{array}$$

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

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

$$(\lambda A + \mu B)C = \lambda(AC) + \mu(BC) \text{ 和 } D(\lambda A + \mu B) = \lambda(DA) + \mu(DB).$$

证明. 考虑线性映射:

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

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 $\phi_D : \mathbb{R}^m \rightarrow \mathbb{R}^k$. 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 6.5,

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

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

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 $D = (d_{j,i})_{n \times m} = B^t A^t$. 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