第二章　矩阵

5.2　与基底和维数有关的性质

定理 5.14 *(线性映射基本定理)* 设 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 $\mathbb{R}^n$ 的一组基, $\mathbf{w}_1, \ldots, \mathbf{w}_n$ 是 $\mathbb{R}^m$ 中的任意给定的向量. 则存在唯一的线性映射 ϕ 使得 $\phi(\mathbf{v}_j) = \mathbf{w}_j$, $j = 1, 2, \ldots, n$.

证明. 定义

$$\begin{array}{rccc} \phi: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & \sum_{j=1}^n \alpha_j \mathbf{v}_j & \mapsto & \sum_{j=1}^n \alpha_j \mathbf{w}_j, \quad \alpha_1, \ldots, \alpha_n \in \mathbb{R} \end{array}.$$

因为 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是基底, 所以对任意 $\mathbf{x} \in \mathbb{R}^n$, 存在唯一的 $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ 使得 $\mathbf{x} = \sum_{j=1}^n \alpha_j \mathbf{v}_j$. 于是, ϕ 是良定义的. 显然 $\phi(\mathbf{v}_j) = \mathbf{w}_j$, $j = 1, 2, \ldots, n$. 下面验证 ϕ 是线性的. 设

$$\mathbf{x} = \sum_{j=1}^n \alpha_j \mathbf{v}_j,\ \mathbf{y} = \sum_{j=1}^n \beta_j \mathbf{v}_j,\ \alpha_j, \beta_j \in \mathbb{R}.$$

则

$$\begin{aligned} \phi(\mathbf{x}+\mathbf{y}) &= \phi\left(\sum_{j=1}^n \alpha_j \mathbf{v}_j + \sum_{j=1}^n \beta_j \mathbf{v}_j\right) = \phi\left(\sum_{j=1}^n (\alpha_j + \beta_j) \mathbf{v}_j\right) \\ &= \sum_{j=1}^n (\alpha_j + \beta_j) \mathbf{w}_j \quad (\phi \text{ 的定义}) \\ &= \sum_{j=1}^n \alpha_j \mathbf{w}_j + \sum_{j=1}^n \beta_j \mathbf{w}_j = \phi(\mathbf{x}) + \phi(\mathbf{y}) \quad (\phi \text{ 的定义}). \end{aligned}$$

设 $\lambda \in \mathbb{R}$. 则

$$\begin{aligned}\phi(\lambda \mathbf{x}) &= \phi\left(\sum_{j=1}^{n}(\lambda\alpha_j)\mathbf{v}_j\right) = \sum_{j=1}^{n}(\lambda\alpha_j)\mathbf{w}_j \quad (\phi \text{ 的定义}) \\ &= \lambda\sum_{j=1}^{n}\alpha_j\mathbf{w}_j = \lambda\phi(\mathbf{x}) \quad (\phi \text{ 的定义}).\end{aligned}$$

最后, 我们来验证唯一性. 设 $\psi : \mathbb{R}^n :\longrightarrow \mathbb{R}^m$ 是线性映射满足 $\psi(\mathbf{v}_j) = \mathbf{w}_j$, $j = 1, 2, \ldots, n$. 则

$$\psi(\mathbf{x}) = \psi\left(\sum_{j=1}^{n}\alpha_j\mathbf{v}_j\right) = \sum_{j=1}^{n}\alpha_j\psi(\mathbf{v}_j) = \sum_{j=1}^{n}\alpha_j\mathbf{w}_j = \phi(\mathbf{x}).$$

故 $\psi = \phi$. □

例 5.15 设 $\mathbf{e}_1, \mathbf{e}_2$ 是 $\mathbb{R}^2$ 的标准基, 且 $\theta \in [0, 2\pi)$. 设 $R_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ 的线性映射满足

$$R_\theta(\mathbf{e}_1) = \cos(\theta)\mathbf{e}_1 + \sin(\theta)\mathbf{e}_2, \ R_\theta(\mathbf{e}_2) = -\sin(\theta)\mathbf{e}_1 + \cos(\theta)\mathbf{e}_2.$$

设 $\mathbf{x} = (\rho\cos(\alpha), \rho\sin(\alpha))^t$. 可直接计算可得

$$R_\theta(\mathbf{x}) = (\rho\cos(\alpha+\theta), \rho(\sin(\alpha+\theta)).$$

我们称 R_θ 是 $\mathbb{R}^2$ 上的旋转 *(rotation)*.

例 5.16 设 $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 $\mathbb{R}^n$ 的标准基. 设 $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ 是满足 $f(\mathbf{e}_j) = \lambda_j, j = 1, 2, \ldots, n$ 的线性映

射. 则对任意 $\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n\in\mathbb{R}^n$,

$$f(\mathbf{x})=\lambda_1x_1+\cdots+\lambda_nx_n.$$

称 f 是 $\mathbb{R}^n$ 上的线性函数.

5.3 线性映射和矩阵的互相表示

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

命题 5.17 设 $A\in\mathbb{R}^{m\times n}$, 线性映射 $\phi_A:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ 是满足

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

则对任意 $\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n\in\mathbb{R}^n$,

$$\phi_A(\mathbf{x})=x_1\vec{A}^{(1)}+\cdots+x_n\vec{A}^{(n)}.$$

进而, $V_c(A)=\operatorname{im}(\phi_A)$ 和 $V_A=\ker(\phi_A)$, 其中 V_A 是以 A 为系数矩阵的齐次线性方程程组的解空间.

证明. 直接计算得

$$\phi_A(\mathbf{x})=\phi_A\left(\mathbf{e}_1\right)=\sum_{j=1}^{n}x_j\phi_A\left(\mathbf{e}_1\right)=\sum_{j=1}^{n}x_j\vec{A}^{(j)}.$$

设 $\mathbf{y}\in\mathbb{R}^m$. 则

$$\mathbf{y}\in\operatorname{im}(\phi_A)\Leftrightarrow\exists\,\mathbf{x}\in\mathbb{R}^n,\mathbf{y}=\phi_A(\mathbf{x})\Leftrightarrow\mathbf{y}=\sum_{j=1}^{n}x_j\vec{A}^{(j)}\Leftrightarrow\mathbf{y}\in V_c(A).$$

故 $\mathrm{im}(\phi_A) = V_c(A)$. 类似地,

$$\mathbf{x} \in \ker(\phi_A) \Leftrightarrow \sum_{j=1}^{n} x_j \vec{A}^{(j)} = \mathbf{0}_m \Leftrightarrow \mathbf{x} \in V_A.$$

故 $\ker(\phi_A) = V_A$. $\square$

我们称上述命题中的 ϕ_A 是由矩阵 A 在标准基下诱导的线性映射. 根据线性映射基本定理, ϕ_A 是唯一的. 上述命题还说明映射版的对偶定理蕴含方程版的对偶定理.

命题 5.18 设线性映射 $\phi_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. 则存在唯一的 $A \in \mathbb{R}^{m\times n}$ 满足 $\phi = \phi_A$, 其中 ϕ_A 是 A 在标准基下的诱导的线性映射. 进而, $\mathrm{im}(\phi) = V_c(A)$ 和 $\ker(\phi) = V_A$, 其中 V_A 是以 A 为系数矩阵的齐次线性方程程组的解空间.

证明. 对 $j = 1, 2, \ldots, n$, 设

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

令

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

则 $\phi = \phi_A$ (线性映射基本定理). 由此和命题 5.18 得 $\mathrm{im}(\phi) = V_c(A)$ 和 $\ker(\phi) = V_A$. $\square$

我们称上述命题中的 A 是线性映射在标准基下矩阵（表示）, 有时记为 A_ϕ. 上述命题还说明方程版的对偶定理蕴含映射的对偶定理.

例 5.19 设 $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ 是零映射, 即 $\phi(\mathbf{e}_j) = \mathbf{0}_m$, $j = 1, 2, \ldots, n$. 则 $A_\phi = O_{m\times n}$.

例 5.20 设 $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ 是数乘映射, 即 $\phi(\mathbf{e}_j) = \lambda\mathbf{e}_j$, $j = 1, 2, \ldots, n$, 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 $\mathrm{diag}_n(\lambda)$. 记 $\mathrm{diag}_n(1)$ 为 E_n. 称为 n 阶单位方阵. 它对应的线性映射是恒同映射.

命题 5.21 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) $\dim(\mathrm{im}(\phi)) = \mathrm{rank}(A)$. 特别地, ϕ 是满射当且仅当 A 行满秩.

(ii) $\dim(\ker(\phi)) = n - \mathrm{rank}(A)$. 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. 根据命题 5.18, $\dim(V_c(A)) = \mathrm{rank}(A)$. 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. 故 (i) 成立. 根据对偶定理, $\dim(\ker(\phi)) = n - \mathrm{rank}(A)$. 特别地, ϕ 是单射当且仅当 A 列满秩. 故 (ii) 成立. (iii) 是 (i) 和 (ii) 的直接推论. □

例 5.22 设 $\mathbf{e}_1, \ldots, \mathbf{e}_4$ 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 $\phi : \mathbb{R}^4 :\longrightarrow \mathbb{R}^3$ 由

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

确定. 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 $\mathrm{im}(\phi)$ 的维数;

(iii) 分别计算 $\ker(\phi)$ 和 $\mathrm{im}(\phi)$ 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 $\mathrm{rank}(A_\phi) = 2$, 所以 $\dim(\mathrm{im}(\phi)) = 2$. 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

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

于是, $\ker(\phi)$ 的一组基是

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

像空间 $\mathrm{im}(\phi)$ 的基是 A_ϕ 中的极大线性无关组. 因为

$$\dim(\mathrm{im}(\phi)) = 2,$$

所以取 A_ϕ 中任意两个线性无关的列向量即可. 例如 $\mathrm{im}(\phi)$ 的一组基是 $\vec{A}_\phi^{(1)}, \vec{A}_\phi^{(2)}$.

6 矩阵的运算

6.1 通过线性映射的运算定义矩阵的运算

定义 6.1 设 ϕ 和 ψ 是从集合 S 到 $\mathbb{R}^m$ 的映射, $\lambda \in \mathbb{R}$. 令:

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

分别称为映射的加法与映射的数乘.

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 6.2 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,

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

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

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

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

由此我们定义两个矩阵 A 和 B 的和

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

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

设 $\lambda\in\mathbb{R}$. 则

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

由此我们定义矩阵的数乘

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

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

$$\phi_A + \phi_B = \phi_{A+B} \quad \text{和} \quad (\lambda\phi)_A = \phi_{\lambda A}.$$

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 6.3 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.

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

6.2 乘法

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

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

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

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

故

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

定义 6.4 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 $i=1,\ldots,m$, $j=1,\ldots,n$. 我们把 C 记为 AB.

注解 6.5 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

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

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:

$$\mathbf{u} = (u_1, \ldots, u_k), \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_s \end{pmatrix} \Longrightarrow \mathbf{uv} = u_1v_1 + \cdots + u_sv_s.$$

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

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

例 6.6 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 6.7 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 6.8 设 $A = (a_{i,j}) \in \mathbb{R}^{m \times n}$. 证明:

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

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

$$(0,\ldots,0,\lambda_i,0,\ldots 0)\vec{A}^{(j)}=\lambda_i a_{i,j}\xRightarrow{j=1,2,\ldots,n}\vec{B}_i=\lambda_i\vec{A}_i.$$

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

$$\vec{A}_i\begin{pmatrix}0\\ \vdots\\ \lambda_j\\ \vdots\\ 0\end{pmatrix}=\lambda_j a_{i,j}\xRightarrow{i=1,2,\ldots,m}\vec{C}^{(j)}=\lambda_j\vec{A}^{(j)}.$$

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

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

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 $\mathbf{x} \in \mathbb{R}^n$,

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

设 A 是 n 阶方阵. $A^0 := E$. 对 $k \in \mathbb{Z}^+$, $A^k := \underbrace{A\cdots A}_{k}$.

例 6.9 计算 $\begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}^k$, 其中 k 是任意正整数.

解. 设 $A = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}$. 则

$$A^2 = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E.$$

断言. 如果 $k = 2m$, 其中 $m \in \mathbb{Z}^+$, 则 $A^k = E$.

断言的证明. 当 $m=1$ 时, 断言成立. 设 $m>1$ 且 $m-1$ 时, 断言成立. 则

$$A^{2m}=A^{2(m-1)}A^2=EE=E.$$

故断言成立.

当 $k=2m+1$. $A^k=A^{2m}A=EA=A$, 其中 $m\in\mathbb{N}$.

综上所述

$$A^k=\begin{cases}A & \text{如果 } k \text{ 是奇数},\\ E & \text{如果 } k \text{ 是偶数}.\end{cases}$$

6.3 矩阵运算的规律

加法. 设 $A,B,C\in\mathbb{R}^{m\times n}$.

(A1) 交换律: $A+B=B+A$.

(A2) 结合律: $(A+B)+C=A+(B+C)$.

(A3) 加法单位元: $A+O_{m\times n}=A$.

(A4) 加法逆元: $A+(-A)=O_{m\times n}$.

数乘. 设 $\alpha,\beta\in\mathbb{R}$ 和 $A\in\mathbb{R}^{m\times n}$.

(S1) 结合律: $(\alpha\beta)A=\alpha(\beta A)$.

(S2) 数乘单位元: $1A=A$.

以上规律由实数的运算规律可直接推出.

命题 6.10 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, $(AB)C=A(BC)$.
□

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

(M1) 结合律: $(AB)C = A(BC)$.

(M2) 数乘单位元: $E_m A = A E_s = A$.

加法与数乘的分配律. 设 $\alpha, \beta \in \mathbb{R}$ 和 $A, B \in \mathbb{R}^{m\times n}$.

(AS1) $(\alpha+\beta)A = \alpha A + \beta A$.

(AS2) $\alpha(A+B) = \alpha A + \alpha B$.

数乘与乘法的分配律. 设 $\alpha, \beta \in \mathbb{R}$, $A \in \mathbb{R}^{m\times s}, B \in \mathbb{R}^{s\times n}$.

(SM) $(\alpha A)(\beta B) = (\alpha\beta)AB$.

加法与乘法的分配律. 设 $A, B{\in}\mathbb{R}^{m\times n}$, $C{\in}\mathbb{R}^{k,m}$, $D{\in}\mathbb{R}^{n\times \ell}$.

(AM1) 左分配律: $C(A+B) = CA + CB$.

(AM2) 右分配律: $(A+B)D = AD + BD$.

我们来验证 (AM1). 其它验证或者是直接的或者与下述验证类似.

设 $A = (a_{i,j})_{m\times n}$, $B = (b_{i,j})_{m\times n}$, $C = (c_{p,i})_{p\times m}$ 和 $G = C(A+B) = (g_{p,j})_{k\times n}$. 则

$$g_{p,j} = \sum_{i=1}^{m} c_{p,i}(a_{i,j}{+}b_{i,j}) = \underbrace{\sum_{i=1}^{m} c_{p,i}a_{i,j}}_{CA\text{ 中 } i \text{ 行 } j \text{ 列处元素}} + \underbrace{\sum_{i=1}^{m} c_{p,i}b_{i,j}}_{CB\text{ 中 } i \text{ 行 } j \text{ 列处元素}} .$$

因为 $p = 1, \ldots, m$ 和 $j = 1, \ldots, n$, 所以 $G = CA + CB$.

转置的计算规律

(AT) 设 $A, B \in \mathbb{R}^{m\times n}$. 则 $(A+B)^t = A^t + B^t$.

(ST) 设 $\alpha \in \mathbb{R}$ 和 $A \in \mathbb{R}^{m\times n}$. 则 $(\alpha A)^t = \alpha A^t$

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

规律 (AT) 和 (ST) 可直接验证, (MT) 验证如下:

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 $C^t = D$. $\square$