

矩阵计算及其理论

潘建瑜

华东师范大学 数学科学学院

<http://math.ecnu.edu.cn/~jypan/Teaching/MCT/index.html>

第二讲 矩阵与线性变换

2.1 线性空间

2.2 线性变换

2.3 内积空间

2.4 投影变换



2.2 | 线性变换

在本节中, 我们总是假定 $\mathbb{U}$ 和 $\mathbb{V}$ 是同一个数域 $\mathbb{F}$ 上的两个线性空间.

2.2 线性变换

2.2.1 线性变换的定义

2.2.2 线性变换与矩阵

2.2.3 过渡矩阵

2.2.4 线性变换在不同基下的矩阵



2.2.1 线性变换的定义

定义 设 $f: \mathbb{U} \rightarrow \mathbb{V}$ 是一个映射, 如果 f 满足:

$$(1) f(x+y) = f(x) + f(y), \quad \forall x, y \in \mathbb{U}; \quad (\text{可加性})$$

$$(2) f(\alpha x) = \alpha f(x), \quad \forall x \in \mathbb{U}, \alpha \in \mathbb{F}. \quad (\text{齐次性})$$

则称 f 是一个**线性映射**或**线性变换**.

 以上两个性质 (可加性和齐次性) 可统一写为

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y), \quad \forall x, y \in \mathbb{U}, \alpha, \beta \in \mathbb{F}.$$

 有教材中当 $\mathbb{V} = \mathbb{U}$ 时才将 f 称为**线性变换**或**线性算子**, 本讲义不做区分.



例 设 $U = V = \mathbb{R}^n$, 线性变换 $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

- **零变换**: $f(x) = 0, \forall x \in \mathbb{R}^n$;
- **数乘变换**: $f(x) = \alpha x, \forall x \in \mathbb{R}^n$, 其中 α 为给定的非零实数;
- **恒等变换**: $f(x) = x, \forall x \in \mathbb{R}^n$.



线性变换基本性质

设 $f: \mathbb{U} \rightarrow \mathbb{V}$ 是线性变换, 则

- (1) $f(0) = 0$;
- (2) $f(-x) = -f(x), \forall x \in \mathbb{U}$;
- (3) $f(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k) = \alpha_1 f(x_1) + \alpha_2 f(x_2) + \cdots + \alpha_k f(x_k), \forall x_i \in \mathbb{U}, \alpha_i \in \mathbb{F}$.

 设 f 是线性变换, 若 $x_1, x_2, \dots, x_k$ 线性相关, 则 $f(x_1), f(x_2), \dots, f(x_k)$ 也线性相关.



线性变换的和, 线性变换的乘积

设 f 和 g 是从 $\mathbb{U}$ 到 $\mathbb{V}$ 的两个线性变换, 则它们的和与乘积定义为

$$(f + g)(x) = f(x) + g(x), \quad \forall x \in \mathbb{U},$$

$$(fg)(x) = f(g(x)), \quad \forall x \in \mathbb{U}.$$

定理 设 f 和 g 是从 $\mathbb{U}$ 到 $\mathbb{V}$ 的两个线性变换, 则 $f + g$ 和 fg 也是线性变换.

(留作课外练习)



2.2.2 线性变换与矩阵

定理 (线性变换的构造) 设 $u_1, u_2, \dots, u_n$ 是线性空间 $\mathbb{U}$ 的一组基, $y_1, y_2, \dots, y_n$ 是 $\mathbb{V}$ 中任意一组向量, 则存在唯一的线性变换 f , 使得

$$f(u_i) = y_i, \quad i = 1, 2, \dots, n.$$

(板书)

 由该定理可知, 要确定一个线性变换, 只需知道其在某一组基下的像即可.



线性变换与矩阵

设 $u_1, u_2, \dots, u_n$ 和 $v_1, v_2, \dots, v_m$ 分别是 $\mathbb{U}$ 和 $\mathbb{V}$ 的一组基, f 是从 $\mathbb{U}$ 到 $\mathbb{V}$ 的线性变换, 且

$$f(u_1) = a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_m,$$

$$\dots \quad \dots \quad \dots$$

$$f(u_n) = a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{mn}v_m,$$

写成矩阵形式为

$$[f(u_1), \dots, f(u_n)] = [v_1, \dots, v_m]A, \quad \text{其中 } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

我们称 A 为线性变换 f 关于基 $u_1, u_2, \dots, u_n$ 和 $v_1, v_2, \dots, v_m$ 的矩阵.



线性变换与矩阵

由上面的结论可知, 取定 $\mathbb{U}$ 和 $\mathbb{V}$ 的一组基后, 任意线性变换 f 都存在**唯一的**矩阵 A 与之对应, 因此我们可以直接用矩阵 A 来表示线性变换 f .

为了便于理解, 教材中常常会用花写字体表示线性变换, 直立字体表示对应的矩阵. **在后面的表述中, 我们就采用这种写法**, 即

$$[\mathcal{A}(u_1), \mathcal{A}(u_2), \dots, \mathcal{A}(u_n)] = [v_1, v_2, \dots, v_m]A, \quad (2.1)$$

或

$$\mathcal{A}(u_1, u_2, \dots, u_n) = [v_1, v_2, \dots, v_m]A.$$

我们称 A 为线性变换 $\mathcal{A}$ 在基 $u_1, u_2, \dots, u_n$ 和 $v_1, v_2, \dots, v_m$ 下的矩阵.

若 $\mathbb{V} = \mathbb{U}$ 且 $v_i = u_i, i = 1, \dots, n$, 则称 A 为 $\mathcal{A}$ 在基 $u_1, u_2, \dots, u_n$ 下的矩阵.



坐标

设 $u_1, u_2, \dots, u_n$ 是 $\mathbb{U}$ 的一组基, 对任意 $x \in \mathbb{U}$, 我们可将其表示为

$$x = x_1 u_1 + x_2 u_2 + \cdots + x_n u_n,$$

我们称 $x_1, x_2, \dots, x_n$ 为 x 在基 $u_1, u_2, \dots, u_n$ 下的坐标.

为方便起见, 我们通常记为

$$x = [x_1, x_2, \dots, x_n]^T.$$



矩阵与向量乘法的几何意义

设 $v_1, v_2, \dots, v_m$ 是 $\mathbb{V}$ 的一组基, $\mathcal{A}$ 是从 $\mathbb{U}$ 到 $\mathbb{V}$ 的线性变换, 则

$$\mathcal{A}(x) = [\mathcal{A}(u_1), \dots, \mathcal{A}(u_n)] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [v_1, \dots, v_m] Ax,$$

因此 $\mathcal{A}(x)$ 在 $v_1, v_2, \dots, v_m$ 下的坐标为 Ax . 我们可以简写为

$$\mathcal{A}(x) = Ax$$

这就是矩阵乘法的几何意义.

👉 需要指出的是, x 和 $\mathcal{A}(x)$ 在不同基下面的坐标可能是不一样的.

👉 若 $\mathbb{V} = \mathbb{U}$, 则 x 和 Ax 就是同一个元素在不同基下的坐标.



推论 设 $u_1, u_2, \dots, u_n$ 和 $v_1, v_2, \dots, v_m$ 分别是 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 一组基, 则所有从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性变换与所有 $\mathbb{R}^{m \times n}$ 中的矩阵一一对应.

定理 在相同的基下面, 线性变换的和对应矩阵的和, 线性变换的乘积对应矩阵的乘积, 即

$$(\mathcal{A} + \mathcal{B})(u_1, u_2, \dots, u_n) = [v_1, v_2, \dots, v_m](A + B),$$

$$(\mathcal{A}\mathcal{B})(u_1, u_2, \dots, u_n) = [v_1, v_2, \dots, v_m](AB).$$



2.2.3 过渡矩阵

设 $x \in \mathbb{U}$, 下面我们讨论 x 在不同基下的坐标之间的关系.

设 $u_1, u_2, \dots, u_n$ 和 $\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n$ 是 $\mathbb{U}$ 的两组基, 并设

$$\begin{aligned}\tilde{u}_1 &= s_{11}u_1 + s_{21}u_2 + \cdots + s_{n1}u_n, \\ &\quad \dots \quad \dots \quad \dots \\ \tilde{u}_n &= s_{1n}u_1 + s_{2n}u_2 + \cdots + s_{nn}u_n,\end{aligned}$$

即

$$[\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n] = [u_1, u_2, \dots, u_n]S \quad \text{其中} \quad S = \begin{bmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{bmatrix},$$

或

$$[\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n]S^{-1} = [u_1, u_2, \dots, u_n].$$

我们称 S 为从基 $u_1, u_2, \dots, u_n$ 到基 $\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n$ 的过渡矩阵.



设 x 在基 $u_1, u_2, \dots, u_n$ 和 $\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n$ 下的坐标分别为 $[x_1, x_2, \dots, x_n]^\top$ 和 $[\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n]^\top$, 则

$$x = [u_1, u_2, \dots, u_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n] \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{bmatrix} = [u_1, u_2, \dots, u_n] S \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{bmatrix}.$$

所以有

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = S \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{bmatrix},$$

即 x 在不同基下的坐标相差一个过渡矩阵.



设 $\mathcal{A}$ 是 $\mathbb{U}$ 到 $\mathbb{V}$ 的线性变换, 则

$$\mathcal{A}(\tilde{u}_i) = s_{11}\mathcal{A}(u_1) + s_{21}\mathcal{A}(u_2) + \cdots + s_{n1}\mathcal{A}(u_n), \quad i = 1, 2, \dots, n.$$

所以

$$\mathcal{A}(\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n) = \mathcal{A}(u_1, u_2, \dots, u_n)S. \quad (2.2)$$

即 $\mathcal{A}(\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)$ 与 $\mathcal{A}(u_1, u_2, \dots, u_n)$ 也相差一个过渡矩阵.



2.2.4 线性变换在不同基下的矩阵

设 $u_1, u_2, \dots, u_n$ 和 $\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n$ 是 $\mathbb{U}$ 的两组基, 从 $u_1, u_2, \dots, u_n$ 到 $\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n$ 的过渡矩阵为 $S \in \mathbb{R}^{n \times n}$, 即

$$[\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n] = [u_1, u_2, \dots, u_n]S \quad (2.3)$$

设 $v_1, v_2, \dots, v_m$ 和 $\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_m$ 是 $\mathbb{V}$ 的两组基, 从 $v_1, v_2, \dots, v_m$ 到 $\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_m$ 的过渡矩阵为 $T \in \mathbb{R}^{m \times m}$, 即

$$[\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_m] = [v_1, v_2, \dots, v_m]T \text{ 或 } [\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_m]T^{-1} = [v_1, v_2, \dots, v_m] \quad (2.4)$$



设线性变换 $\mathcal{A}$ 在基 $u_1, u_2, \dots, u_n$ 和 $v_1, v_2, \dots, v_m$ 下的矩阵为 A , 在基 $\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n$ 和 $\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_m$ 下的矩阵为 $\tilde{A}$, 即

$$\mathcal{A}(u_1, u_2, \dots, u_n) = [v_1, v_2, \dots, v_m]A, \quad \mathcal{A}(\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n) = [\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_m]\tilde{A} \quad (2.5)$$

由 (2.2)-(2.5) 可知

$$\begin{aligned}\mathcal{A}(\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n) &= \mathcal{A}(u_1, u_2, \dots, u_n)S \\ &= [v_1, v_2, \dots, v_m]AS \\ &= [\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_m]T^{-1}AS.\end{aligned}$$

由线性变换与矩阵之间对应关系的唯一性可知

$$\tilde{A} = T^{-1}AS.$$



矩阵相似

如果 $\mathbb{U} = \mathbb{V} = \mathbb{R}^n$, 并取 $v_i = u_i, \tilde{v}_i = \tilde{u}_i, i = 1, 2, \dots, n$, 则 $T = S$, 且

$$\tilde{A} = S^{-1}AS,$$

即 A 在两组不同的基下的矩阵相似.