《线性代数》

第二章: 矩阵

曾鹏程
(zengpch@shanghaitech.edu.cn)

2025.10.10

教学安排

线性映射的定义

定义 (2.18)

一个映射 $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ 如果满足条件：

$$(i)\ \varphi(x+y) = \varphi(x) + \varphi(y),\ \ \forall x, y \in \mathbb{R}^n,$$
$$(ii)\ \varphi(\lambda x) = \lambda\varphi(x),\ \ \forall \lambda \in \mathbb{R}, x \in \mathbb{R}^n,$$

则称 φ 是（从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的）线性映射.

线性映射的定义

定义 (2.18)

一个映射 $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ 如果满足条件:

$$(i)\ \varphi(x+y) = \varphi(x) + \varphi(y),\ \ \forall x, y \in \mathbb{R}^n,$$
$$(ii)\ \varphi(\lambda x) = \lambda\varphi(x),\ \ \forall \lambda \in \mathbb{R}, x \in \mathbb{R}^n,$$

则称 φ 是（从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的）线性映射.

- $m = 1$, 称为线性函数

线性映射的定义

例 2.19: 设 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 分别是高 n 和高 m 的列向量空间.
$\forall\, \mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n \in \mathbb{R}^n$, 对 $\mathbf{x} = [x_1, x_2, \cdots, x_n] \in \mathbb{R}^n$,
$\varphi(\mathbf{x}) = \sum_{i=1}^n x_i \mathbf{a}_i \in \mathbb{R}^m$. 则映射

$$\varphi : \mathbb{R}^n \to \mathbb{R}^m,\ \mathbf{x} \to \varphi(\mathbf{x}).$$

是否为线性？

线性映射的定义

- 例 2.19 中的线性映射 φ 完全由向量组 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ 确定
- 这个向量组也完全由线性映射 φ 确定. 为什么？

线性映射的定义

- 例 2.19 中的线性映射 φ 完全由向量组 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ 确定
- 这个向量组也完全由线性映射 φ 确定. 为什么？
- 以 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ 为列向量的矩阵记作 A, 则线性映射 φ 完全由 A 确定，反之亦然; 记这个映射为 φ_A
- 不同的 $m \times n$ 矩阵给出不同的线性映射

线性映射的定义

- 例 2.19 中的线性映射 φ 完全由向量组 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ 确定
- 这个向量组也完全由线性映射 φ 确定. 为什么？
- 以 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ 为列向量的矩阵记作 A, 则线性映射 φ 完全由 A 确定，反之亦然; 记这个映射为 φ_A
- 不同的 $m \times n$ 矩阵给出不同的线性映射

定理 (2.20)

映射 $A \to \varphi_A$ 是从 $M_{m,n}(\mathbb{R})$ 到 $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ 的双射. 矩阵 A 称为线性映射 φ_A 的矩阵，映射 φ_A 称为 A 的线性映射.

练习一

判断下面的映射是否为线性？如果是，确定其矩阵:

- (1) $[a, b] \to [2a + 3b, a - 5b]$
- (2) $[a, b] \to [e^a, e^b]$

线性映射的运算

- 向量空间 $\mathbb{R}^m$ 有加法和数乘运算，因而可以对从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射定义加法和数乘运算

线性映射的运算

- 向量空间 $\mathbb{R}^m$ 有加法和数乘运算，因而可以对从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射定义加法和数乘运算
- 设 $\varphi, \psi : \mathbb{R}^n \to \mathbb{R}^m$ 是线性映射, $\alpha, \beta \in \mathbb{R}$, 定义

$$\theta = \alpha\varphi + \beta\psi : \mathbb{R}^n \to \mathbb{R}^m, \quad u \to \alpha\varphi(u) + \beta\psi(u), \quad \forall u \in \mathbb{R}^n.$$

- θ 是否为线性映射？

线性映射的运算

- 向量空间 $\mathbb{R}^m$ 有加法和数乘运算，因而可以对从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射定义加法和数乘运算
- 设 $\varphi, \psi : \mathbb{R}^n \to \mathbb{R}^m$ 是线性映射, $\alpha, \beta \in \mathbb{R}$, 定义

$$\theta = \alpha\varphi + \beta\psi : \mathbb{R}^n \to \mathbb{R}^m, \quad u \to \alpha\varphi(u) + \beta\psi(u), \quad \forall u \in \mathbb{R}^n.$$

- θ 是否为线性映射？
- 映射的合成给出线性映射的乘法. 设 $\varphi : \mathbb{R}^n \to \mathbb{R}^s$, $\psi : \mathbb{R}^s \to \mathbb{R}^m$ 是线性映射，定义它们的乘积 $\psi\varphi : \mathbb{R}^n \to \mathbb{R}^m$ 为:

$$(\psi\varphi)(u) = \varphi(\psi(u)), \quad \forall u \in \mathbb{R}^n.$$

- $\psi\varphi$ 是否为线性映射？

矩阵的运算：矩阵的加法和数乘

- 线性映射上的运算给出矩阵相应的运算
- 设 $A=(a_{ij}), B=(b_{ij})$ 是 $m\times n$ 矩阵，$\varphi_A, \varphi_B : \mathbb{R}^n \to \mathbb{R}^m$ 是相应的线性映射，则存在 $m\times n$ 矩阵 $C=(c_{ij})$, s.t. $\varphi_C=\alpha\varphi_A+\beta\varphi_B$. Why?
- C 的第 j 列是什么？

矩阵的运算：矩阵的加法和数乘

- 线性映射上的运算给出矩阵相应的运算
- 设 $A=(a_{ij}), B=(b_{ij})$ 是 $m\times n$ 矩阵，$\varphi_A, \varphi_B : \mathbb{R}^n \to \mathbb{R}^m$ 是相应的线性映射，则存在 $m\times n$ 矩阵 $C=(c_{ij})$, s.t. $\varphi_C = \alpha\varphi_A + \beta\varphi_B$. Why?
- C 的第 j 列是什么？
- $\mathbf{c}_j = [\alpha a_{1j}+\beta b_{1j}, \alpha a_{2j}+\beta b_{2j}, \cdots, \alpha a_{mj}+\beta b_{mj}]$
- C 可以定义为 $\alpha A+\beta B$, i.e., $\varphi_{\alpha A+\beta B} = \alpha\varphi_A + \beta\varphi_B$
- $\alpha=\beta=1$ 时，C 即为矩阵加法
- $\beta=0$ 时，C 即为矩阵与纯量的乘法
- $m\times n$(实数) 矩阵全体 $M_{m,n}(\mathbb{R})$ 是一个向量空间

- 设 $A=(a_{ik}), B=(b_{kj})$ 分别是 $m\times s, s\times n$ 矩阵，$\varphi_A:\mathbb{R}^s\to\mathbb{R}^m, \varphi_B:\mathbb{R}^n\to\mathbb{R}^s$ 是相应的线性映射，则存在 $m\times n$ 矩阵 $C=(c_{ij})$, s.t. $\varphi_C=\varphi_A\varphi_B:\mathbb{R}^n\to\mathbb{R}^m$
- C 的第 j 列是什么？

矩阵的运算：矩阵的乘法

- 设 $A = (a_{ik}), B = (b_{kj})$ 分别是 $m \times s, s \times n$ 矩阵，$\varphi_A : \mathbb{R}^s \to \mathbb{R}^m, \varphi_B : \mathbb{R}^n \to \mathbb{R}^s$ 是相应的线性映射，则存在 $m \times n$ 矩阵 $C = (c_{ij})$, s.t. $\varphi_C = \varphi_A \varphi_B : \mathbb{R}^n \to \mathbb{R}^m$
- C 的第 j 列是什么？
- $c_{ij} = A_i \mathbf{b}_j = \sum_{k=1}^{s} a_{ik} b_{kj}, \ i = 1, \cdots, m; \ j = 1, \cdots, n$
- C 可以定义为 AB, i.e., $\varphi_{AB} = \varphi_A \varphi_B$
- 矩阵乘积 AB 有意义当且仅当 A 的列数等于 B 的行数

矩阵的运算：矩阵的乘法

- 设 $A=(a_{ik}), B=(b_{kj})$ 分别是 $m\times s, s\times n$ 矩阵，$\varphi_A:\mathbb{R}^s\to\mathbb{R}^m, \varphi_B:\mathbb{R}^n\to\mathbb{R}^s$ 是相应的线性映射，则存在 $m\times n$ 矩阵 $C=(c_{ij})$, s.t. $\varphi_C=\varphi_A\varphi_B:\mathbb{R}^n\to\mathbb{R}^m$
- C 的第 j 列是什么？
- $c_{ij}=A_i\mathbf{b}_j=\sum_{k=1}^s a_{ik}b_{kj},\ i=1,\cdots,m;\ j=1,\cdots,n$
- C 可以定义为 AB, i.e., $\varphi_{AB}=\varphi_A\varphi_B$
- 矩阵乘积 AB 有意义当且仅当 A 的列数等于 B 的行数

命题 (2.22)

矩阵的乘法满足结合律，对加法的分配律，即有
(1)$(AB)C=A(BC)$
(2)$(A+B)C=AC+BC,\ D(A+B)=DA+DB.$

矩阵的转置

- 把原矩阵的行按列排、列按行排得到的矩阵

定义 (2.23)

矩阵 $B=(b_{ij})$ 称为矩阵 $A=(a_{ji})$ 的转置，记作 $B={}^{t}A$，如果 $b_{ij}=a_{ji}$

- 行向量的转置是列向量，列向量的转置是行向量

矩阵的转置

- 把原矩阵的行按列排、列按行排得到的矩阵

定义 (2.23)

矩阵 $B=(b_{ij})$ 称为矩阵 $A=(a_{ji})$ 的转置，记作 $B={}^tA$，如果 $b_{ij}=a_{ji}$

- 行向量的转置是列向量，列向量的转置是行向量

命题 (2.24)

(i) ${}^t({}^tA)=A,\quad {}^t(A+B)={}^tA+{}^tB,\quad {}^t(\lambda A)=\lambda\,{}^tA.$
(ii) ${}^t(AB)={}^tB\,{}^tA$
(iii) $\operatorname{rank} A=\operatorname{rank}{}^tA$

矩阵乘积的秩

定理 (2.25)

设 A 为 $m \times s$ 矩阵，B 为 $s \times n$ 矩阵，那么
$rank\, AB \leq min\{rank\, A, rank\, B\}$

- 矩阵加法和乘法可以通过分块进行
- 假设两个 $m \times s$ 矩阵 X 和 X' 被纵横线划分成若干小的长方块:

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{l1} & X_{l2} & \cdots & X_{lk} \end{pmatrix}, \quad X' = \begin{pmatrix} X'_{11} & X'_{12} & \cdots & X'_{1k} \\ X'_{21} & X'_{22} & \cdots & X'_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ X'_{l1} & X'_{l2} & \cdots & X'_{lk} \end{pmatrix}$$

X_{ij} 和 X'_{ij} 都是 $m_i \times s_j$ 矩阵，$m = \sum_{i=1}^{l} m_i, \quad s = \sum_{i=1}^{k} s_i$，则 $X + X'$ 可以分块计算，$X + X' = ?$

矩阵的分块

- 如果 $s \times n$ 矩阵 $Y = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1r} \\ Y_{21} & Y_{22} & \cdots & Y_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{k1} & Y_{k2} & \cdots & Y_{kr} \end{pmatrix}$, Y_{ij} 是 $s_i \times n_j$ 矩阵 $n = \sum_{i=1}^{r} n_i$, 那么 $Z = XY = ?$

列向量空间之间的线性映射 φ_A

定理 (2.26)

设 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 分别是 n 维列向量空间和 m 维列向量空间. 那么
(1) 任何 $m \times n$*(实)* 矩阵 A 都定义了一个线性映射

$$\varphi_A : \mathbb{R}^n \to \mathbb{R}^m, \quad X \to AX.$$

(2) 如果 $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ 是线性映射，则有唯一的 $m \times n$*(实)* 矩阵 A 使得 $\varphi = \varphi_A$.
(3) 如果 B 是 $m \times n$ 矩阵且 $A \neq B$, 那么 $\varphi_A \neq \varphi_B$.
(4) 特别，$\varphi_A = 0$, 即对任意的 $X \in \mathbb{R}^n$, 有 $\varphi_A(X) = AX = 0$, 当且仅当 $A = 0$, 即 A 是 $m \times n$ 零矩阵.

习题 2.3

1. 判断下面的映射哪些是线性映射. 如果是, 确定线性映射的矩阵.

(1) $[a,b] \to [b,a]$, (2) $[a,b] \to [a,-b]$,

(3) $[a,b] \to [b,0]$, (4) $[a,b] \to [a,a]$,

(5) $[a,b] \to [a^2,b^2]$, (6) $[a,b] \to [e^a,e^b]$,

(7) $[a,b] \to [a,1]$, (8) $[a,b] \to [a+1,b-1]$,

(9) $[a,b] \to [a-b,a+b]$, (10) $[a,b] \to [2a+3b,a-5b]$,

(11) $[a_1, a_2, \ldots, a_{n-1}, a_n] \to [a_2, a_3, \ldots, a_n, a_1]$;

(12) $[a_1, a_2, \ldots, a_{n-1}, a_n] \to [a_1, a_1a_2, \ldots, a_1a_2\cdots a_{n-1}, a_1a_2\cdots a_n]$;

(13) $[a_1, a_2, \ldots, a_{n-1}, a_n] \to [a_1, a_1+a_2, \ldots, a_1+a_2\cdots+a_{n-1}, a_1+a_2+\cdots+a_n]$.

2. 计算:

(1) $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$; (2) $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$;

(3) $\begin{pmatrix} 2 & -1 & 3 & 0 \\ 1 & -2 & 3 & 1 \\ 2 & 3 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & -2 \\ 3 & 1 & 2 \\ 5 & -1 & -3 \\ 2 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 4 & 0 & -3 \\ 0 & 2 & -5 \\ 4 & -1 & -7 \end{pmatrix}$;

(4) $\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}^k$, 其中 k 是任意正整数;

习题 2.3

(5) $\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ & & \cdots & \cdots & \cdots & & \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}^k$，其中 k 是任意正整数，

(6) $\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ & & \cdots & \cdots & \cdots & & \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}^n$，其中 n 是该矩阵的阶数.

3\. 设 A 和 B 是 $m \times n$ 矩阵. 证明:

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

4\. 对任意的 $m \times s$ 矩阵 A 和 $s \times n$ 矩阵 B, 证明:

$$\operatorname{rank} A + \operatorname{rank} B - s \leqslant \operatorname{rank}(AB).$$

5\. 证明: 如果三个 n 阶方阵的乘积为 0, 那么它们的秩的和不超过 $2n$.

6\. 证明: 如果 $m \times n$ 矩阵 A 的秩为 1, 则存在高为 m 的列向量($m \times 1$ 矩阵) B 和长为 n 的行向量($1 \times n$ 矩阵) C 使得 $A = BC$.

7\. 证明小节六中的矩阵加法和乘法的分块计算的正确性.

8\. 设 A 是 $m \times n$ 矩阵, T 是 $s \times m$ 矩阵, S 是 $n \times t$ 矩阵. 利用矩阵的分块乘法证明: TA 的行向量都是 A 的行向量的线性组合, AS 的列向量都是 A 的列向量的线性组合. 由此得到定理2.25的证明的一个简单表述.

习题 2.3

Quiz 1:

1. (10 points) 姓名: 学号:
2. (5 points) 选择 λ 的值，使得下面线性方程组有解:

$$\begin{aligned} 2x_1 - x_2 + x_3 + x_4 &= 1, \\ x_1 + 2x_2 - x_3 + 4x_4 &= 2, \\ x_1 + 7x_2 - 4x_3 + 11x_4 &= \lambda. \end{aligned}$$

3. (5 points) 证明: 在四维实行向量空间 $\mathbb{R}^4$ 中，向量 $\boldsymbol{\alpha}_1 = (1,1,0,0), \boldsymbol{\alpha}_2 = (0,0,1,1), \boldsymbol{\alpha}_3 = (1,0,0,4), \boldsymbol{\alpha}_4 = (0,0,0,2)$ 构成一组基.

方阵

- 行数与列数相同的矩阵称为方阵，这个数称为方阵的阶（或级）
- 同阶方阵之间可以相加、相乘、与纯量相乘，满足结合律和分配律
- 方阵转置后仍是同阶
- 全体 n 阶实（数）方阵的集合记作 $M_n(\mathbb{R})$

单位矩阵和纯量矩阵

- 恒等映射 $\mathrm{Id}_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n, x \to x$ 所对应的 n 阶方阵是什么？
- 该方阵称为 n 阶单位矩阵，记作 E 或 I 或 I_n
- $EA = AE = A, \quad A \in M_n(\mathbb{R})$
- 对任意 $\lambda \in \mathbb{R}$ 和 $A \in M_n(\mathbb{R})$,

$$(\lambda E)A = A(\lambda E)$$

- λE 称为纯量矩阵，记作 $\mathrm{diag}_n(\lambda)$

定理 (2.27)

在 $M_n(\mathbb{R})$ 中与所有矩阵可交换的矩阵是纯量矩阵

可逆矩阵

- 简记恒等映射为 $\mathcal{E}$, 则对任何映射 $\theta: \mathbb{R}^n \to \mathbb{R}^n$, 都有 $\mathcal{E}\theta = \theta\mathcal{E} = \mathcal{E}$
- 映射 $\varphi: \mathbb{R}^n \to \mathbb{R}^n$ 称为可逆的，如果存在映射 $\psi: \mathbb{R}^n \to \mathbb{R}^n$ 使得 $\varphi\psi = \psi\varphi = \mathcal{E}$. 此时，$\varphi$ 和 ψ 互为逆映射
- 映射 φ 的逆映射如果存在，则唯一
- 如果 φ 是线性映射且可逆，ψ 是其逆映射，那么 ψ 也是线性映射. Why?
- 可逆的线性映射对应的矩阵称为可逆矩阵

可逆矩阵

- 设 $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ 是可逆线性映射，则存在 n 阶方阵 B，使得 $\varphi_B = \varphi_A^{-1}$. 于是

$$\varphi_{AB} = \varphi_A \varphi_B = \mathcal{E} = \varphi_E = \varphi_B \varphi_A = \varphi_{BA}$$

可逆矩阵

- 设 $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ 是可逆线性映射，则存在 n 阶方阵 B，使得 $\varphi_B = \varphi_A^{-1}$. 于是

$$\varphi_{AB} = \varphi_A \varphi_B = \mathcal{E} = \varphi_E = \varphi_B \varphi_A = \varphi_{BA}$$

- 称 $A \in M_n(\mathbb{R})$ 为可逆矩阵如果存在 $B \in M_n(\mathbb{R})$ 使得

$$AB = BA = E,$$

称 B 为 A 的逆矩阵，记作 A^{-1}

可逆矩阵

定义 (2.28)

$A \in M_n(\mathbb{R})$ 称为非退化的，如果 $\mathrm{rank} A = n$. 如果 $\mathrm{rank} A < n$, 则称 A 为退化的.

可逆矩阵

定义 (2.28)

$A \in M_n(\mathbb{R})$ 称为非退化的，如果 $\mathrm{rank} A = n$. 如果 $\mathrm{rank} A < n$, 则称 A 为退化的.

定理 (2.29)

设 A 是 n 阶方阵，那么以下陈述等价:
(1) A 是可逆的.
(2) A 是非退化的.
(3) A 的秩为 n.
(4) A 的行向量（或等价地，列向量）线性无关.