第二章　矩阵代数

2021 年 10 月 20 日

内容

名词与记号

- 矩阵的维数：指的是其行数和列数
- 矩阵元素的标记：$A = [a_{ij}]$; $(A)_{ij} = a_{ij}$.
- 方阵
- 零矩阵，对角矩阵和单位矩阵

  单位矩阵的矩阵元：$(I_n)_{ij} = \delta_{ij}$, $i, j = 1, \cdots, n$. 其中

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

  对角矩阵的矩阵元：$(D)_{ij} = d_i \delta_{ij}$.

矩阵的和与标量乘法

- 矩阵的和: 只能在相同维数的矩阵之间进行
- 矩阵与数的乘法: 标量乘法
- 矩阵的和与标量乘法的运算满足关于向量空间加法与数乘的 8 个性质 (参看教材 P26)

同一数域上的所有 $m \times n$ 的矩阵亦构成了一个线性空间!

同理, $\mathbb{R}^n \to \mathbb{R}^m$ 上的所有线性变换也构成了一个线性空间.

矩阵乘法

引入:

有两个映射 $T_1: M \to N$ 和 $T_2: L \to M$, 定义复合映射 $T_1 \circ T_2$ (简写为 $T_1 T_2$): $L \to N$ 如下:

$$\vec{x}\,(\in L) \xrightarrow{T_2} T_2(\vec{x})\ (\in M) \xrightarrow{T_1} T_1(T_2(\vec{x}))\ (\in N)$$

即

$$(T_1 T_2)(\vec{x}) := T_1((T_2(\vec{x})))$$

从定义出发, 我们可以证明两个线性变换的复合依然是线性变换.

复合线性变换的矩阵表达式

下面我们讨论两个复合线性变换 T_1T_2 对应的矩阵变换 AB 的表达式，即假设 $T_1(\vec{y}) = A\vec{y}$，$T_2(\vec{x}) = B\vec{x}$，从 $T_1T_2(\vec{x}) := (AB)\vec{x}$ 出发确定矩阵乘积 AB.

首先根据定义,

$$T_1 T_2(\vec{x}) = T_1(T_2(\vec{x})) = T_1(B\vec{x}) = A(B\vec{x}).$$

而

$$B\vec{x} = x_1\vec{b}_1 + \cdots + x_p\vec{b}_p, \quad \Rightarrow$$

$$\begin{aligned} A(B\vec{x}) &= A(x_1\vec{b}_1) + \cdots + A(x_p\vec{b}_p) \\ &= x_1 A(\vec{b}_1) + \cdots + x_p A(\vec{b}_p) \\ &= \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_p \end{bmatrix} \vec{x} := (AB)\vec{x}, \end{aligned}$$

即

$$AB = A\begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} = \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_p \end{bmatrix}.$$

这就是矩阵乘法的定义.

矩阵乘法的规则

- AB 的每一列都是 A 的各列的线性组合, 以 B 的对应列的元素为权.
- AB 的每一行都是 B 的各行的线性组合, 以 A 的对应行的元素为权.
- 行列法则

$$(AB)_{ij} = \sum_l a_{il} b_{lj}$$

注意: A 的列数和 B 的行数要匹配.

矩阵乘法的性质

- $A(BC) = (AB)C$ 乘法的结合律
- $A(B+C) = AB + AC$ 乘法的左分配律
- $(B+C)A = BA + CA$ 乘法的右分配律
- $r(AB) = (rA)B = A(rB)$, r 是数
- $I_m A = A = A I_n$ 矩阵乘法的恒等式

示例

- 证明 $A(BC) = (AB)C$.

  证明一: (略, 见教材)

  证明二:

$$\begin{aligned}(A(BC))_{ij} &= \sum_l a_{il}(BC)_{lj} = \sum_l a_{il}\left(\sum_k b_{lk}c_{kj}\right)\\ &= \sum_k \left(\sum_l a_{il}b_{lk}\right) c_{kj} = \sum_k (AB)_{ik}c_{kj}\\ &= ((AB)C)_{ij}.\end{aligned}$$

- 证明 $I_mA = A$.

  证明:

$$(I_mA)_{ij} = \sum_k (I_m)_{ik}a_{kj} = \sum_k \delta_{ik}a_{kj} = a_{ij} = (A)_{ij}.$$

- 矩阵的乘幂

$$A^k := \underbrace{A \cdots A}_{k}$$

- 矩阵的转置

$$\left(A^T\right)_{ij} := (A)_{ji} \equiv a_{ji}$$

- 对称矩阵与反对称矩阵
  - 对称矩阵: $A^T = A$, 即 $a_{ij} = a_{ji}$;
  - 反对称矩阵: $A^T = -A$, 即 $a_{ij} = -a_{ji}$.

- 矩阵转置的性质 (定理 3)
  1. $(A^T)^T = A$
  2. $(A+B)^T = A^T + B^T$
  3. $(rA)^T = rA^T$
  4. $(AB)^T = B^T A^T$

示例

- 证明 $(A^T)^T = A$.

  证明:
  $$\left((A^T)^T\right)_{ij} = (A^T)_{ji} = a_{ij} = (A)_{ij}.$$

- 证明 $(AB)^T = B^T A^T$.

  证明:
  $$\begin{aligned}\left((AB)^T\right)_{ij} &= (AB)_{ji} = \sum_k a_{jk} b_{ki} \\ &= \sum_k \left(A^T\right)_{kj} \left(B^T\right)_{ik} = \left(B^T A^T\right)_{ij}.\end{aligned}$$

内容

矩阵可逆的定义

初始版:

定义

对于一个 $m \times n$ 的矩阵 A, 如果存在 $n \times m$ 的矩阵 C 和 D, 使得

$$CA = I_n \quad \text{及} \quad AD = I_m,$$

我们就说 A 是可逆的.

我们可以证明, 根据上面的定义, 对于有限维的矩阵 A, 如果其可逆, 则 A 应当是方阵, 且有 $C = D$.

简化版:

定义

对于一个$n \times n$的矩阵 A, 如果存在一个 $n \times n$ 的矩阵 C, 使得

$$CA = I_n \quad 及 \quad AC = I_n,$$

我们就说 A 是可逆的. 其可逆矩阵记为 A^{-1}.

实际上, 我们可以证明, 对于 $n \times n$ 的矩阵 A, 如果存在一个 $n \times n$ 的矩阵 C, 使得 $CA = I_n$, 则必有 $AC = I_n$ (参看教材 2.3 节可逆矩阵的特征).

可逆变换与可逆矩阵

定义

我们称一个变换 $T: A \to B$ 是可逆的, 若当存在变换 $G: B \to A$, 能够使得 $GT = \mathbf{1}_A$ 以及 $TG = \mathbf{1}_B$.

注: $\mathbf{1}_A$ 和 $\mathbf{1}_B$ 分别是集合 A 和 B 上的恒等变换.

- 变换 T 可逆的充要条件是 T 是双射, 即 T 既是满射又是单射.
- 若 T 是有限维空间上的线性变换, $T(\vec{x}) = A\vec{x}$, 则 T 可逆当且仅当矩阵 A 可逆.

内容

可逆矩阵的一些等价陈述

一个 $n \times n$ 的矩阵 A 可逆，当且仅当

- A 的各行各列都有主元.
- A 有 n 个主元.
- A 的各列线性无关.
- $A\vec{x} = \vec{0}$ 只有平凡解.
- 对于 $\forall \vec{b} \in \mathbb{R}^n$, $A\vec{x} = \vec{b}$ 都有唯一解.
- A 等价于 $n \times n$ 的单位矩阵 I_n.
- A 的各列生成 $\mathbb{R}^n$.
- 存在 $n \times n$ 矩阵 C 使 $CA = I$.
- 存在 $n \times n$ 矩阵 D 使 $AD = I$.
- $\cdots$

可逆矩阵的运算性质

若 A (和 B) 可逆,

- $\left(A^{-1}\right)^{-1} = A$
- $(AB)^{-1} = B^{-1}A^{-1}$ (可逆矩阵的乘积也是可逆的)
- A^T 可逆, 且 $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

示例

- 2×2 矩阵的逆矩阵

  设 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, 则 A 可逆当且仅当 $ad - bc \neq 0$, 且

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

- 上 (下) 三角矩阵可逆的条件

内容

初等矩阵

定义

把单位矩阵进行一次初等行变换后所得到的矩阵称为初等矩阵.

- 初等矩阵都是可逆的.
- 对矩阵 A 进行某种初等行变换可以写成 EA 两个矩阵相乘的形式, 其中 E 是对应的初等矩阵.

初等行变换求逆

如果 $n \times n$ 的矩阵 A 是可逆的, 则把矩阵 A 变为单位矩阵 I_n 的一系列初等行变换将把 I_n 变成 A^{-1}.

因为如果 $(E_p \cdots E_1) A = I_n$, 则有 $(E_p \cdots E_1) I_n = A^{-1}$.

算法:

把增广矩阵$[A \; I]$ 进行行化简. 若 A 可逆, 则 $[A \; I] \to [I \; A^{-1}]$.

逆矩阵与求解线性方程组

设矩阵 A 可逆, $A^{-1} = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_n \end{bmatrix}$, 则

$$AA^{-1} = \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_n \end{bmatrix} = I_n = \begin{bmatrix} \vec{e}_1 & \cdots & \vec{e}_n \end{bmatrix}.$$

即逆矩阵 A^{-1} 的第 j 列满足方程 $A\vec{x} = \vec{e}_j$.

逆矩阵与求解线性方程组

设矩阵 A 可逆, $A^{-1} = \begin{bmatrix} \vec{b}_1 & \cdots \vec{b}_n \end{bmatrix}$, 则

$$AA^{-1} = \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_n \end{bmatrix} = I_n = [\vec{e}_1 \quad \cdots \quad \vec{e}_n] .$$

即逆矩阵 A^{-1} 的第 j 列满足方程 $A\vec{x} = \vec{e}_j$.

问题: 设 A 为可逆 $n \times n$ 的矩阵, B 为 $n \times p$ 矩阵, 求 $A^{-1}B$.

解答:

$$[A \ B] \quad \to \quad [I \ X],$$

则 $X = A^{-1}B$.

内容

- 加法和标量乘法
- 乘法

分块矩阵的乘法，如 AB 可以按照通常的行列法则进行，只要A 的列的分法和B 的行的分法一致．如

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix}.$$

列行法则

矩阵乘法的列行展开法则:

若 A 是 $m \times n$ 矩阵, B 是 $n \times p$ 矩阵, 则

$$\begin{aligned} AB &= \begin{bmatrix} \mathrm{col}_1(A) & \mathrm{col}_2(A) & \cdots & \mathrm{col}_n(A) \end{bmatrix} \begin{bmatrix} \mathrm{row}_1(B) \\ \mathrm{row}_2(B) \\ \vdots \\ \mathrm{row}_n(B) \end{bmatrix} \\ &= \mathrm{col}_1(A)\mathrm{row}_1(B) + \cdots + \mathrm{col}_n(A)\mathrm{row}_n(B). \end{aligned}$$

内容

$\mathbb{R}^n$ 的子空间

定义

集合 $H \subset \mathbb{R}^n$, H 是 $\mathbb{R}^n$ 的一个线性子空间, 要求

1. $\vec{0} \in H$;
2. H 中的向量对加法和数乘运算封闭. 即 $\forall\, \vec{u},\ \vec{v} \in H$, $\vec{u} + \vec{v} \in H$; $\forall\, \vec{u} \in H$ 和数 c, $c\vec{u} \in H$.

$\mathbb{R}^n$ 的子空间

定义

集合 $H \subset \mathbb{R}^n$, H 是 $\mathbb{R}^n$ 的一个线性子空间, 要求

1. $\vec{0} \in H$;
2. H 中的向量对加法和数乘运算封闭. 即 $\forall\, \vec{u},\ \vec{v} \in H,\ \vec{u} + \vec{v} \in H$; $\forall\, \vec{u} \in H$ 和数 c, $c\vec{u} \in H$.

子空间示例: 矩阵的列空间与零空间

对于矩阵 $A_{m\times n} = [\vec{a}_1\ \cdots\ \vec{a}_n]$,

- A 的列空间 $\mathrm{Col\,A} := \mathrm{Span}\,\{\vec{a}_1, \cdots, \vec{a}_n\} \subset \mathbb{R}^m$.
- A 的零空间 $\mathrm{Nul\,A} := \left\{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\right\} \subset \mathbb{R}^n$.

动机: 研究能够生成线性空间 H 所必需的最小的向量集合.

每一个这样的向量组我们称为 H 的一个基.

对于给定的 H, 我们有

- 这个向量组必定是线性无关的.
- 基的选取不是唯一的.
- 每个基所包含的向量个数是一样的.

引理

有向量组 $\{\vec{a}_1, \cdots, \vec{a}_q\}$ (记为 α-向量组) 和 $\{\vec{b}_1, \cdots, \vec{b}_p\}$ (记为 β-向量组), 若 $\alpha \subset \mathrm{Span}\{\beta\}$, 即 α-向量组能被 β-向量组线性表出, 且 $q > p$, 则 α-向量组必线性相关.

证明:

$$\begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_q \end{bmatrix} = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} \begin{bmatrix} c_{11} & \cdots & c_{1q} \\ c_{21} & \cdots & c_{2q} \\ \vdots & \cdots & \vdots \\ c_{p1} & \cdots & c_{pq} \end{bmatrix} := \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} C_{p\times q}$$

因为 $q > p$, 则 $\exists \vec{u} (\neq \vec{0}) \in \mathbb{R}^q$, 使得 $C\vec{u} = \vec{0}$. 因此有 $\begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_q \end{bmatrix} \vec{u} = \vec{0}$, 即 α 线性相关.

线性空间的维数

从引理出发, 我们可以很容易推出, 对于线性空间 H 的任意两个基, 它们所包含的向量数目是一样的. 这个数目我们就定义为线性空间 H 的维数, 记为$\dim H$.

示例

- 列空间 $\mathrm{Col}\,A$:

  矩阵 A 的主元列构成了 $\mathrm{Col}\,A$ 的一个基;

  $\dim \mathrm{Col}\,A$ (:= rank A) = $\{A$ 的主元列的数目$\}$. (提示: 两个矩阵 A, B (行) 等价 (记为 $A \sim B$) 则意味着 A 和 B 的列具有相同的线性关系)

- 零空间 $\mathrm{Nul}\,A$:

  $\dim \mathrm{Nul}\,A = \{A$的自由变量的数目$\}$.

一些定理

- 秩定理: 对于 $m \times n$ 的矩阵 A, $\operatorname{rank} A + \dim \operatorname{Nul} A = n$

  即 基本变量数目 + 自由变量数目 = (系数) 矩阵的列数

- 若 $\dim H = p$, 有向量组 $\{\vec{a}_1, \cdots, \vec{a}_q\} \in H$ 且 $q > p$, 则该向量组必线性相关. (回忆前述引理及第一章中的定理 8)

- 基定理: 若 $\dim H = p$, 则 H 中任意一个由 p 个线性无关的向量组成的向量组便构成了 H 的一个基.

关于可逆矩阵 (续)

有矩阵 $A_{n\times n}$, A 可逆当且仅当

- $\operatorname{rank} A (= \dim \operatorname{Col} A) = n$
- $\dim \operatorname{Nul} A = 0$
- $\operatorname{Col} A = \mathbb{R}^n$
- $\cdots$

- 设有向量组 $\alpha = \{\vec{a}_i, i = 1, \cdots, 5\}$，其中 $\vec{a}_1 = (1, -1, 2, 4)$，$\vec{a}_2 = (0, 3, 1, 2)$，$\vec{a}_3 = (3, 0, 7, 14)$，$\vec{a}_4 = (1, -2, 2, 0)$，$\vec{a}_5 = (2, 1, 5, 10)$. 求该向量组的一个极大线性无关组，并将其余向量表示为该极大线性无关组的线性组合.

  答: 考虑矩阵 $A = [\vec{a}_1 \ \cdots \ \vec{a}_5]$，则 $\mathrm{Col}\, A$ 的一组基为 $\{\vec{a}_1, \vec{a}_2, \vec{a}_4\}$ (此即为 α-向量组的一个极大线性无关组)，而 $\vec{a}_3 = 3\vec{a}_1 + \vec{a}_2$，$\vec{a}_5 = 2\vec{a}_1 + \vec{a}_2$.

示例 II

- 设 A, B 分别为 $m \times n$, $n \times p$ 矩阵，且 $AB = 0$. 则 $\operatorname{rank} A + \operatorname{rank} B$?

  答: $\operatorname{rank} A + \operatorname{rank} B \leq n$ (提示: 秩定理)

内容

向量用基展开

有线性空间 H, $\dim H = p$, 设 $\beta = \{\vec{b}_1, \cdots, \vec{b}_p\}$ 是 H 的一个基, 则 $\forall \vec{x} \in H$ 都可以用 β-基展开:

$$\vec{x} = c_1\vec{b}_1 + \cdots + c_p\vec{b}_p = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix} = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} \vec{c}.$$

可以证明, 对于给定的基展开式是唯一的.

向量用基展开

有线性空间 H, $\dim H = p$, 设 $\beta = \{\vec{b}_1, \cdots, \vec{b}_p\}$ 是 H 的一个基, 则 $\forall \vec{x} \in H$ 都可以用 β-基展开:

$$\vec{x} = c_1\vec{b}_1 + \cdots + c_p\vec{b}_p = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix} = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} \vec{c}.$$

可以证明, 对于给定的基展开式是唯一的.

定义

我们称 $\vec{c} \in \mathbb{R}^p$ 是 H 中的 $\vec{x}$ 相对于 β-基的坐标向量, 或 $\vec{x}$ 的 β-坐标向量, 记为 $[\vec{x}]_\beta$.

上面给出的 $\vec{x} \mapsto [\vec{x}]_\beta$ 的对应, 即 $\vec{x} = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{bmatrix} [\vec{x}]_\beta$, 定义了一个映射 $\beta : H \to \mathbb{R}^p$.

- β 是双射.
- β 是线性变换. 因为

$$\beta(\vec{x} + \vec{y}) = [\vec{x} + \vec{y}]_\beta = [\vec{x}]_\beta + [\vec{y}]_\beta = \beta(\vec{x}) + \beta(\vec{y}),$$
$$\beta(c\vec{x}) = [c\vec{x}]_\beta = c[\vec{x}]_\beta = c\beta(\vec{x}).$$

我们称这样的映射为线性空间的同构映射.

在这里, 我们可以看到 $H \cong \mathbb{R}^p$, 即 H 与 $\mathbb{R}^p$ 同构.

事实上我们有

定理

在实数域上的任意一个 n 维线性空间均同构于 $\mathbb{R}^n$.

示例 I

有向量 $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$,

- 考虑基 $\beta_1 = \{\vec{e}_1, \vec{e}_2\}$, 则 $[\vec{x}]_{\beta_1} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.
- 考虑基 $\beta_2 = \{\vec{a}_1, \vec{a}_2\}$, 其中 $\vec{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{a}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. 求 $[\vec{x}]_{\beta_2}$?

设向量 $\vec{x}$ 属于一个基为 $\beta = \{\vec{b}_1, \vec{b}_2\}$ 的子空间 H, 其中 $\vec{b}_1 = \begin{bmatrix} 1 \\ 4 \\ -3 \end{bmatrix}$, $\vec{b}_2 = \begin{bmatrix} -2 \\ -7 \\ 5 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} 2 \\ 9 \\ -7 \end{bmatrix}$.

- $\dim H = 2$.
- $[\vec{x}]_\beta = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

本章作业

- 2.1 节: 3, 7, 11, 15, 21, 29
- 2.2 节: 5, 9, 15, 23, 37
- 2.3 节: 9, 17, 23, 27, 35
- 2.4 节: 5, 11, 17, 21
- 2.8 节: 3, 9, 17, 33
- 2.9 节: 3, 11, 15, 27

课堂练习

- 2.1 节: 5, 23, 33
- 2.2 节: 3, 11, 35
- 2.3 节: 7, 37, 39
- 2.4 节: 7, 23, 25
- 2.8 节: 7, 21, 25
- 2.9 节: 7, 17, 19