

# 线代笔记

mny

2023 年 10 月 18 日

## 目录

### 1 线性空间

线性空间  $\mathbb{R}^m$ ,  $m$  是一个自然数.

$m = 1$  时, 是实数. 有两个代数运算  $+$  和  $*$ , 有两个特殊元素  $0$  和  $1$ .

#### 1.1 实数中运算的性质

$+$  满足的性质:

- 交换的,  $a + b = b + a$
- 对于任意一个  $a$ , 存在  $b$ , 使得  $a + b = 0$ ,  $b = -a$   
 $\implies$  减法运算  $a - b = a + (-b)$
- 加法满足结合律  $a + (b + c) = (a + b) + c$

$*$  满足的性质:

- 交换的  $a * b = b * a$
- 对于一个非  $0$  元素  $a$ , 存在一个元素  $b$ , 使得  $a * b = 1$ ,  $b = a^{-1}$
- 结合律  $a * (b * c) = (a * b) * c$

$+$  和  $*$  满足分配律:  $a * (b + c) = a * b + a * c$

定义 1.1.  $\mathbb{R}^m$  中的元素为  $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$ , 其中  $a_1, \dots, a_m$  为任意实数.

$\mathbb{R}^m$  中的元素  $v = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$  称为列向量. 有时一个元素表示为  $[a_1, a_2, \dots, a_m]$ , 称作行向量.

$\mathbb{R}^m$  上定义两个运算  $+$  和  $*$  (用列向量来表示)

定义 1.2.  $+$  加法: 任意两个列向量  $a, b$  得到一个新的列向量.

$$v + w = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_m + b_m \end{bmatrix} \quad (1.1)$$

例 1.1. 在  $\mathbb{R}^2$  中,  $\begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$

定义 1.3.  $*$  数乘: 任意一个实数  $c$ , 以及一个列向量  $v$ , 得到一个新的列向量  $cv$

$$cv = c \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} ca_1 \\ ca_2 \\ \vdots \\ ca_m \end{bmatrix} \quad (1.2)$$

## 1.2 $(\mathbb{R}^m, +, *)$ 为一个线性空间

两种运算满足:

- 交换律  $v + w = w + v$
- 结合律  $c_1(c_2v) = (c_1c_2)v$
- 分配律  $c(v + w) = cv + cw$

- 通过加法可以定义减法运算  $v - w = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_1 - b_1 \\ a_2 - b_2 \\ \vdots \\ a_m - b_m \end{bmatrix}$ .

- 给定一组  $\mathbb{R}^m$  中的向量,  $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ , 和一组实数  $(x_1, x_2, \dots, x_n)$  可以构成新的向量

$$x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_n \vec{v}_n \quad (1.3)$$

这个新的向称为  $(v_1, v_2, \dots, v_n)$  的线性组合.

### 1.3 矩阵

定义 1.4.  $m \times n$  矩阵,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (1.4)$$

$a_{ij}$  为实数

从线性空间的角度, 矩阵可以有下列的理解:

- 从矩阵列的角度,  $A = [\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n]$

- 从矩阵行的角度,  $A = \begin{bmatrix} \vec{v}_1 \\ \vec{v}_2 \\ \vdots \\ \vec{v}_m \end{bmatrix}$

固定  $m$  和  $n$ , 矩阵空间上可以定义两个自然的运算

定义 1.5. 矩阵加法:

$$A = (a_{ij}), \quad B = (b_{ij}), \quad (A + B)_{ij} = (a_{ij} + b_{ij}) \quad (1.5)$$

定义 1.6. 数乘, 任意一个实数  $c$ , 一个矩阵  $A$ , 得到

$$(cA)_{ij} = (ca_{ij}) \quad (1.6)$$

**定义 1.7.** 矩阵乘法.

定义: 矩阵乘法是 把一个  $m \times n$  矩阵乘上一个  $n \times k$  矩阵, 得到一个  $m \times k$  矩阵.

运算规则:

$$(C)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj} \quad (1.7)$$

矩阵乘法的性质:

- 结合律:

$$(AB)C = A(BC) \quad (1.8)$$

证明. 设  $A: m \times n, B: n \times k, C: k \times l$ , 根据定义

$$[(AB)C]_{ij} = \sum_{u=1}^k (AB)_{iu}C_{uj} = \sum_{u=1}^k \sum_{v=1}^n a_{iv}b_{vu}C_{uj} \quad (1.9)$$

同样

$$[A(BC)]_{ij} = \sum_{v=1}^n A_{iv}(BC)_{vj} = \sum_{u=1}^k \sum_{v=1}^n a_{iv}b_{vu}C_{uj} \quad (1.10)$$

□

- 分配律:

$$A(B+C) = AB + AC \quad (1.11)$$

$$(A+B)C = AC + AB \quad (1.12)$$

- 矩阵乘法不满足交换律

$$AB \stackrel{\text{不一定}}{\neq} BA \quad (1.13)$$

不论交换有没有定义, 都不一定相等.

矩阵乘法的几种理解:

- $C = AB$ ,  $C_{ij}$  为把  $A$  的第  $i$  行和  $B$  的第  $j$  列乘起来.
- 从矩阵  $A$  的列向量的角度看

$$A = [\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n] \quad (1.14)$$

那么  $C$  的第  $j$  列为  $A$  的列向量的线性组合, 组合系数为  $B$  的第  $j$  列,

$$b_{1j}\vec{v}_1 + b_{2j}\vec{v}_2 + \cdots + b_{nj}\vec{v}_n \quad (1.15)$$

- 从矩阵  $B$  的行向量来看

$$B = \begin{bmatrix} \vec{w}_1 \\ \vec{w}_2 \\ \vdots \\ \vec{w}_m \end{bmatrix} \quad (1.16)$$

矩阵  $C$  的第  $i$  行为  $B$  的行向量的线性组合, 组合系数为矩阵  $A$  的第  $i$  行.

$$a_{i1}\vec{w}_1 + a_{i2}\vec{w}_2 + \cdots + a_{in}\vec{w}_n \quad (1.17)$$

几种特殊矩阵:

- 方阵: 行和列数目一致,  $n \times n$
- 零矩阵: 元素都为 0
- $n$  阶单位矩阵:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (1.18)$$

对角线全为 1

- 上三角矩阵:

$$\begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{bmatrix} \quad (1.19)$$

- 下三角矩阵:

$$\begin{bmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{bmatrix} \quad (1.20)$$

## 1.4 矩阵的乘法的应用

对于  $n \times n$  的方阵  $A$ , 可以利用矩阵的乘法来定义它的逆矩阵  $A^{-1}$

**定义 1.8.**  $A^{-1}$  称为  $A$  的逆矩阵, 如果  $A^{-1}$  满足

$$A^{-1}A = I_{n \times n}, \text{ 且 } AA^{-1} = I_{n \times n}. \quad (1.21)$$

**命题 1.1.**

$$I_{n \times n}A = AI_{n \times n} = A \quad (1.22)$$

证明. 根据乘法的定义, 不难证明. □

### 1.4.1 逆矩阵的一些性质

**命题 1.2.** 一个矩阵的逆矩阵不一定存在

**例 1.2.** 非平凡的例子  $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ .

证明. 假设存在  $A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ , 那么  $A^{-1}$  满足

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (1.23)$$

左侧

$$= \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix} \neq I_{2 \times 2} \quad (1.24)$$

□

**命题 1.3.** 如果逆矩阵存在, 它是唯一的.

证明. 假设  $A$  有两个逆矩阵  $B, C$ , 则

$$AB = BA = AC = CA = I_{n \times n} \quad (1.25)$$

所以

$$B = B(AC) = (BA)C = C. \quad (1.26)$$

□

**命题 1.4.** 若一个矩阵存在左逆  $L$ , 满足  $LA = I_{n \times n}$ , 那么矩阵  $A$  的逆矩阵存在, 且等于  $L$ .<sup>1</sup>

**命题 1.5.** 如果  $A$  的逆为  $A^{-1}$ ,  $B$  的逆为  $B^{-1}$ , 则

$$(AB)^{-1} = B^{-1}A^{-1} \quad (1.27)$$

证明.

$$\begin{aligned} (AB)B^{-1}A^{-1} &= A(BB^{-1})A^{-1} \\ &= AA^{-1} \\ &= I_{n \times n}. \end{aligned} \quad (1.28)$$

证明也可以推广到一般情况. □

**命题 1.6.**  $2 \times 2$  矩阵  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  的逆矩阵

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad (1.29)$$

**命题 1.7.** 对角矩阵  $D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$  的逆为

$$D^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix} \quad (1.30)$$

这意味着  $D^{-1}$  存在当且仅当对角元素都不为零!

### 1.4.2 线性组合的矩阵乘法表示

线性组合

$$x_1 \vec{v}_1 + x_2 \vec{v}_2 + \cdots + x_n \vec{v}_n, \quad (1.31)$$

---

<sup>1</sup>将在后面证明.

引入两个矩阵

$$A = [\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n], \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (1.32)$$

其中  $A$  为  $m \times n$  矩阵,  $X$  为  $n \times 1$  矩阵. 线性组合的矩阵表示为  $AX$ .

### 1.4.3 矩阵方程

方程为  $AX = b$ . 这个方程的解的性质取决于  $A$  中的向量  $[\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n]$  和向量  $b$  的关系.

我们把这个方程写成分量的形式

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \quad (1.33)$$

采用高斯 (Gauss) 消元法. 引入三种变换, 这些变换都不改变方程的解.

- 把某一个方程乘上一个系数  $a$ .
- 消元: 第  $i$  个方程  $+ a \times$  第  $j$  个方程.
- 换行: 把第  $i$  行和第  $j$  行交换.

这三种称为矩阵的初等变换.

把第 2 个方程到第  $m$  个方程中的  $x_1$  消掉. 把第  $i$  个方程变为

$$\text{方程}(i) - \frac{a_{i1}}{a_{11}} \times \text{方程}(1) \quad (1.34)$$

于是方程的增广矩阵变为

$$\left[ \begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ 0 & \cdots & a'_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & a'_{mn} & b_m \end{array} \right] \quad (1.35)$$



## 2 矩阵的初等变换

下面我们要将初等变换用矩阵乘法来实现.

- 倍乘变换: 矩阵  $A$  的第  $i$  行乘上  $c$ , 其他行不变,  $S_i(c)A = A'$ .  $S_i(c)$  为将单位矩阵的第  $i$  个元素换为  $c$ .

$$S_i(c) = \begin{bmatrix} \ddots & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (2.1)$$

$S_i^{-1}(c)$  是可逆的

$$S_i^{-1}(c) = \begin{bmatrix} \ddots & 0 & \cdots & 0 \\ 0 & \frac{1}{c} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (2.2)$$

- 消元变换: 把第  $i$  行换成 第  $i$  行  $+ a \times$  第  $j$  行. 它的形式为

$$E_{ij}(a) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ a & & & & \ddots \\ & & & & & 1 \end{bmatrix}. \quad (2.3)$$

其中的  $a$  位于  $E_{ij}(a)$  的第  $i$  行第  $j$  列, 对角线元素都为 1.

$E_{ij}(a)$  是可逆的

$$E_{ij}^{-1}(a) = E_{ij}(-a) \quad (2.4)$$

- 换行变换: 把第  $i$  行和第  $j$  行交换. 用一个矩阵  $P_{ij} =$  (交换单位矩阵的  $i, j$  列) 来表示. 它是可逆的,

$$P_{ij}^{-1} = P_{ij}. \quad (2.5)$$

### 2.1 初等变换的应用

**LU 分解** 把方矩阵分解成下列形式

$$A = LU, \quad (2.6)$$

$L$  为下三角矩阵,  $U$  为上三角矩阵.

方法: 通过初等变换, 把方矩阵变成一个上三角矩阵 (如果这个过程不涉及到换行的话) 那么有

$$E_1 E_2 \cdots E_n A = U \iff A = E_n^{-1} \cdots E_2^{-1} E_1^{-1} U \quad (2.7)$$

由于  $E_1, \dots, E_n$  都是下三角矩阵, 它们的乘积也是下三角矩阵.

**例 2.1.** 对矩阵  $A$  做  $LU$  分解.

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} \quad (2.8)$$

做操作

$$\begin{aligned} A &\xrightarrow{E_{41}(3)E_{31}(-2)E_{21}(1)} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 10 & 4 & -9 \\ 0 & -16 & -11 & 18 \end{bmatrix} \\ &\xrightarrow{E_{42}(-8)E_{32}(5)} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -3 & 2 \end{bmatrix} \\ &\xrightarrow{E_{43}(-3)} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (2.9)$$

用矩阵乘法写起来

$$E_{43}(-3)E_{42}(-8)E_{32}(5)E_{41}(3)E_{31}(-2)E_{21}(1)A = U \quad (2.10)$$

即

$$L = E_{21}(-1)E_{31}(2)E_{41}(-3)E_{32}(-5)E_{42}(8)E_{43}(-3) \quad (2.11)$$

计算可得

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix}. \quad (2.12)$$

如果  $U$  的对角线都不为零, 那么

$$U = DU', \quad (U' \text{ 对角线都为 } 1) \quad (2.13)$$

用初等变换求逆 (**Gauss-Jordan**) 原理: 先找到一组初等变换来使得矩阵  $A$  变为单位矩阵,

$$(E_p \cdots E_1)A = I. \quad (2.14)$$

$A$  的逆可以这样求解:

$$AB = I \quad (2.15)$$

在两边同时做行变换.

$$(E_p \cdots E_1)AB = (E_p \cdots E_1)I \quad (2.16)$$

即

$$B = IB = (E_p \cdots E_1)I \quad (2.17)$$

在实际操作时, 考虑增广矩阵  $[A|I]$ , 做初等变换, 变为  $[I|A^{-1}]$

**例 2.2.**

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \quad (2.18)$$

做初等变换

$$\begin{aligned} [A|I] &= \left[ \begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array} \right] \\ &\xrightarrow{E_{12}(\frac{2}{3})E_{23}(\frac{3}{4})E_{32}(\frac{2}{3})E_{21}(\frac{1}{2})} \left[ \begin{array}{ccc|ccc} 2 & 0 & 0 & \frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \frac{3}{2} & 0 & \frac{3}{4} & \frac{3}{2} & \frac{3}{4} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{array} \right] \\ &\xrightarrow{S_3(\frac{3}{4})S_2(\frac{2}{3})S_1(\frac{1}{2})} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{array} \right] \end{aligned} \quad (2.19)$$

最终的增广矩阵右侧就是  $A$  的逆,

$$A^{-1} = \begin{bmatrix} \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix} \quad (2.20)$$

## 行约化阶梯形及一般线性方程组的一般解

## 行约化阶梯形式

- 如果第  $i$  行都是零, 那么对于  $j > i$  行都是零.
- 如果第  $i$  行不都是零, 那么第一个非零元素为 1, 称为主元.
- 如果第  $(i + 1)$  行不都是零, 那么这一行的主元在第  $i$  行的主元右边.
- 主元上方的元素都为零.

## 找到一个矩阵行约化阶梯形的方法

- 找到第一个不为零的列, 然后, 如果需要的话, 做换行操作, 使得这一列的第一个元素不为零. 之后消元操作, 把下方的元素都变为零.

$$A \rightarrow \begin{bmatrix} 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & \boxed{B} \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & \end{bmatrix} \quad (2.21)$$

- 对子矩阵  $B$  做同样的操作.
- 把主元上方的元素变为零.

例 2.3. 行约化阶梯形式:

$$\begin{bmatrix} 1 & 1 & 2 & 4 \\ 1 & 2 & 2 & 5 \\ 1 & 3 & 2 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (2.22)$$

## 行约化阶梯形式的一些名词

- 主元列: 如果该列有一个主元, 该列只有一个非零元素.
- 自由列: 没有主元之列.
- 主元变量, 自由变量:

对于一个方程

$$Rx = x_1 Y_1 + x_2 Y_2 + \cdots + x_n Y_n, \quad (2.23)$$

如果  $Y_i$  为自由列, 则  $x_i$  为自由变量. 如果  $Y_i$  为主元列, 则  $x_i$  为主元变量.

对于例??, 约化后的方程为

$$\begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad (2.24)$$

其中的  $x_1, x_2$  为主元变量,  $x_3, x_4$  为自由变量.

## 2.2 用行约化阶梯形式求解线性方程组

用行约化阶梯形式求解方程  $Ax = b$ , 方法如下

- 考虑增广矩阵  $[A|b]$ , 做初等行变换, 把  $A$  变成行约化阶梯形式, 得到增广矩阵

$$[R|b'] \quad (2.25)$$

新的方程组  $Rx = b'$  的解空间和原来的方程一样.

- $Rx = b'$  的解 (如果存在) 为

$$x = x_p + x_n \quad (2.26)$$

其中  $x_p$  为  $Rx = b'$  的特解,  $x_n$  为对应的齐次线性方程组 ( $b' = 0$ ) 的所有解.

- $x_p$  可以求解如下: 取自由变量为零, 主元变量任意, 可以得到一个解.

**例 2.4.**

$$R = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad b' = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \quad (2.27)$$

特解为

$$x_p = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 0 \end{bmatrix}. \quad (2.28)$$

- 齐次线性方程的解可以这样求: 取某一个自由变量为 1, 其他自由变量为 0, 主元变量任意. 这样可以一共得到  $n - r$  个解, 记为  $s_i$ ,  $n$  是变量数目,  $r$  是主元数目. 则,

$$x_n = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n. \quad (2.29)$$

**例 2.5.** 继续求解上例中的线性方程.

第一个通解:

$$\begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \\ 0 \end{bmatrix} = 0 \quad (2.30)$$

得到  $x_1 = -2, x_2 = 0$ , 特解向量为  $s = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ . 另一个通解:

$$\begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 0 \\ 1 \end{bmatrix} = 0 \quad (2.31)$$

解向量为  $s_2 = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \end{bmatrix}$ .

于是原方程的所有解为

$$x = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 0 \end{bmatrix} + a_1 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \end{bmatrix}. \quad (2.32)$$

### 3 线性方程组的解

#### 3.1 齐次线性方程解空间的性质

**基本性质**

- 证明: 任何一个解都可以做上面的分解.

$x'$  为一个解,  $x_p$  为另一个解, 那么

$$\begin{cases} Ax' = b \\ Ax_p = b \end{cases} \quad (3.1)$$

两式相减得到

$$A(x' - x_p) = 0 \quad (3.2)$$

即,  $x' = x_p + x_n$  中的  $x_n$  是齐次线性方程的解.

- 反之, 对于任意的齐次线性方程的解  $x_n$ ,  $x_p + x_n$  都是方程  $Ax = b$  的解.

证明: 因为 
$$\begin{cases} Ax_p = b \\ Ax_n = 0 \end{cases} \implies A(x_p + x_n) = b.$$

为什么  $Ax = b$  的解可以写成这种形式

- 如果  $v_1$  为  $Ax = 0$  的解,  $v_2$  也为解, 那么  $v_1 + v_2$  也是方程的解.
- 如果  $v$  是一个解, 那么乘上一个系数  $c$ ,  $cv$  也是方程的解.

因为  $Av = 0$ , 那么  $A(cv) = c(Av) = 0$ .

这证明了  $Ax = 0$  的解空间  $N(A)$  在向量加法及数乘下是封闭的.

### 3.2 线性子空间, 线性无关, 基, 维数

**线性子空间**  $\mathbb{R}^m$  中的一个子空间  $V$ , 如果  $V$  在加法和数乘下面是封闭的, 那么这个子空间称为线性子空间.

**构造线性子空间的方法** 给定一组固定的向量  $(v_1, v_2, \dots, v_n)$ , 考虑所有的线性组合构成的空间

$$V = \{x_1v_1 + x_2v_2 + \dots + x_nv_n\} \quad (3.3)$$

$V$  是一个线性子空间, 称之为  $(v_1, v_2, \dots, v_n)$  张成的线性子空间.

**线性无关**

**定义 3.1.** 一组向量  $(v_1, v_2, \dots, v_n)$  称为线性无关的, 如果下列的方程

$$x_1v_1 + x_2v_2 + \dots + x_nv_n = 0 \quad (3.4)$$

只有 0 解, 即对应的齐次线性方程  $Ax = 0$  只有 0 解.

**例 3.1.**  $\mathbb{R}^2$  中的  $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ,  $v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  是线性无关的.

**例 3.2.** 如果 0 向量在这组向量中, 那么这组向量是线性相关的.

### 线性空间的基

**定义 3.2.** 一组线性无关的向量  $(e_1, e_2, \dots, e_m)$  称之为  $V$  的一组基, 如果  $V$  中任意一个向量都可以表示为这组向量的线性组合,

$$v = x_1 e_1 + x_2 e_2 + \dots + x_n e_n. \quad (3.5)$$

基中的向量个数称作维数.

**例 3.3.**  $\mathbb{R}^2$  中的  $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ,  $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  为一组基. 所以  $\mathbb{R}^2$  的维数为 2.

### 基的几个重要性质

- 坐标唯一性: 给定一组基  $(e_1, e_2, \dots, e_m)$ , 根据基的定义, 任意的向量都可以写成

$$(e_1, e_2, \dots, e_m)$$

的线性组合, 即

$$v = a_1 e_1 + a_2 e_2 + \dots + a_m e_m. \quad (3.6)$$

其中  $(a_1, a_2, \dots, a_m)$  称为  $v$  在基  $(e_1, e_2, \dots, e_m)$  下的坐标.

坐标是唯一的.

**证明.** 假设坐标不唯一,  $v$  可以有两种展开方式:

$$\begin{aligned} v &= a_1 e_1 + a_2 e_2 + \dots + a_m e_m \\ v &= b_1 e_1 + b_2 e_2 + \dots + b_m e_m \end{aligned} \quad (3.7)$$

两式相减得到

$$0 = (a_1 - b_1)e_1 + (a_2 - b_2)e_2 + \dots + (a_m - b_m)e_m \quad (3.8)$$

这与基的线性无关矛盾.  $\square$

- 基不唯一, 但维数定义的维数一样.

**证明.** 反证法. 假设有两组基  $(e_1, e_2, \dots, e_m), (f_1, f_2, \dots, f_n), n > m$ .

根据基的定义,  $(f_1, f_2, \dots, f_n)$  可以写成  $e_1, e_2, \dots, e_m$  的线性组合.

$$\begin{aligned} f_1 &= a_{11}e_1 + \dots + a_{m1}e_m \\ &\vdots \\ f_n &= a_{1n}e_1 + \dots + a_{mn}e_m \end{aligned} \quad (3.9)$$



把上述过程写成矩阵乘法的形式

$$F = [f_1, f_2, \dots, f_n], \quad E = [e_1, e_2, \dots, e_m] \quad (3.10)$$

并且

$$F = EA, \quad (3.11)$$

其中  $A = (a_{ij})$ ,  $A$  为一个  $m \times n$  的矩阵.

考虑  $Ax = 0$  的解, 利用之前齐次线性方程组的解的性质, 参数个数为  $(n - r)$ ,  $r$  为主元数目, 且  $r \leq m$ . 所以  $Ax = 0$  一定有非 0 的解 ( $m \neq n$ ).

利用方程  $F = EA$ , 如果  $Ax = 0$  有非零解, 那么

$$Fx = EAx = 0 \quad (3.12)$$

也有非零解, 和假设矛盾.  $\square$

- 基的变换矩阵  $A$  为可逆的.

证明. 有两组基  $(f_1, f_2, \dots, f_m)$ ,  $(e_1, e_2, \dots, e_m)$ ,

$$F = [f_1, f_2, \dots, f_m], \quad E = [e_1, e_2, \dots, e_m], \quad F = EA. \quad (3.13)$$

$A$  为  $m \times m$   $\square$

引入矩阵的转置.

**定义 3.3.** 给定一个矩阵  $A$ ,

$$(A^T)_{ij} = A_{ji}. \quad (3.14)$$

$A^T$  把  $A$  的行变成列.

转置的一些性质

•

$$(A^T)^T = A. \quad (3.15)$$

•

$$(AB)^T = B^T A^T \quad (3.16)$$

证明. 设  $A: m \times n, B: k \times n$ , 那么

$$(AB)_{ij} = \sum_{l=1}^k a_{il} b_{lj}. \quad (3.17)$$

根据转置的定义, 有

$$(AB)_{ij}^T = (AB)_{ji} = \sum_{l=1}^k a_{jl} b_{li}. \quad (3.18)$$

另一方面,

$$(B^T A^T)_{ij} = \sum_{l=1}^k (B^T)_{li} (A^T)_{jl} = \sum_{l=1}^k a_{jl} b_{li}. \quad (3.19)$$

□

•

$$(A^T)^{-1} = (A^{-1})^T. \quad (3.20)$$

证明. 因为

$$AA^{-1} = I, \quad (3.21)$$

两边取转置得到

$$(A^{-1})^T A^T = I \implies (A^T)^{-1} = (A^{-1})^T. \quad (3.22)$$

□

### 特殊矩阵

- 对称矩阵:  $A^T = A$ .

$$A_{ij} = A_{ji}. \quad (3.23)$$

例 3.4.

$$\begin{bmatrix} 2 & 4 \\ 4 & 3 \end{bmatrix}. \quad (3.24)$$

例 3.5.

$$\begin{bmatrix} 2 & 3 & 1 \\ 3 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}. \quad (3.25)$$

- 反对称矩阵:  $A^T = -A$ . 可知, 其对角线都为零.

例 3.6.

$$\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}. \quad (3.26)$$

下面我们回到方程  $Ax = b$ ,  $A$  可以定义四个线性子空间.

1.  $A$  的列向量张成的线性子空间  $C(A)$ , 它的维数称为  $A$  的列秩.

例 3.7. 对于

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

$$C(A) = \begin{bmatrix} x_1 \\ x_3 \\ 0 \end{bmatrix}, \text{ 是三维空间的 } x-y \text{ 平面.}$$

2.  $A$  的行向量张成的线性子空间  $C(A^T)$ , 它的维数称为  $A$  的行秩.
3.  $A$  的零空间  $N(A)$ . 线性方程组  $Ax = 0$  的所有解.  $N(A)$  的维数为  $n - r$ ,  $r$  为主元数.
4.  $A^T$  的零空间  $N(A^T)$ . 线性方程组  $A^T x = 0$  的所有解.

### 3.3 线性代数基本定理

定理 3.1.

$$r_1 = r_2 = r = r'. \quad (3.27)$$

命题 3.1. 初等行变换不改变行秩和列秩.

证明. 初等行变换对于行线性空间的影响

$$\text{设矩阵 } A = \begin{bmatrix} \vec{w}_1 \\ \vec{w}_2 \\ \vdots \\ \vec{w}_m \end{bmatrix}. \text{ 把行线性子空间记为 } V(w) \subset \mathbb{R}^n.$$

做倍加变换之后, 新的向量组为

$$w' = (w_1, w_2, \dots, w_i + aw_j, \dots, w_m). \quad (3.28)$$

$V(w')$  为另一个线性子空间, 但是  $V(w) = V(w')$ , 因为对于任意的一个向量  $w' \in V(w')$ , 有

$$w' = x_1 w'_1 + \cdots + x_m w'_m = x_1 w_1 + \cdots + (x_j + ax_i)w_j + \cdots + x_m w_m. \quad (3.29)$$

所以有  $V(w') \subset V(w)$ . 反之, 也有  $V(w) \subset V(w')$ .

可得  $V(w') = V(w)$ .

### 初等行变换对于列向量空间的影响

注意到, 初等行变换不改变齐次线性方程组的解, 也就是说

$$Ax = 0 \iff Bx = 0. \quad (3.30)$$

其中  $B$  为  $A$  初等行变换后的矩阵. 这也就是说,

$$x_1 v_1 + \cdots + x_n v_n = 0 \iff x_1 v'_1 + \cdots + x_n v'_n = 0. \quad (3.31)$$

假如  $x_1 \neq 0$ , 那么  $v_1$  可以用其他向量线性组合表示.

所以,  $v$  中线性独立的列向量数之和等于  $v'$  中线性独立的列向量数之和.

□

我们只需考虑行约化阶梯形式  $R$ , 通过观察  $R$  的形式, 可以发现

- $R$  的列秩等于行秩.
- $R$  的行向量空间及列向量空间的维数等于主元数目.

因为主元列是线性无关的, 自由列都可以用主元列的线性组合表示, 主元行是线性无关的, 而自由行是零.

**定义 3.4.** 矩阵的秩 (*rank*) 为列向量空间  $C(A)$  的维数. 秩在初等行变换下不变.

**例 3.8.** 秩为 1 的矩阵的形式: 从列向量的角度来看,

$$A = \begin{bmatrix} a_1 v_i, \cdots, v_i, \cdots, a_n v_i \end{bmatrix}. \quad (3.32)$$

(其中  $v_i$  是非零向量)

**定义 3.5.** 一个矩阵称为满秩的, 如果秩为最大可能值 (行数列数中较小的一个).

$$\begin{aligned} \dim(N(A)) &= n - r \\ \dim(N(A^T)) &= m - r \\ \dim(C(A)) &= r \\ \dim(C(A^T)) &= r \end{aligned} \quad (3.33)$$

我们给线性空间上附加一个新的结构: 内积

**定义 3.6.** 对于线性空间  $\mathbb{R}^m$  中的两个向量, 定义内积

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^m v_i w_i. \quad (3.34)$$

把  $v, w$  看成  $m \times 1$  的矩阵, 可以把内积写成矩阵乘法的形式,

$$\vec{v} \cdot \vec{w} = v^T w = w^T v. \quad (3.35)$$

有了内积, 可以定义一些东西

- 向量  $v$  的长度

$$|v| = \sqrt{v \cdot v} = \sqrt{v^T v}. \quad (3.36)$$

- 两个向量垂直  $v \perp w$ , 如果

$$v \cdot w = 0. \quad (3.37)$$

对于线性方程组  $Ax = b$ , 当  $b$  属于  $C(A)$  时, 有解. 此时

$$A' = [A \ b], \quad \text{rank}(A') = \text{rank } A \quad (3.38)$$

无解时,

$$\text{rank}(A') = \text{rank}(A) + 1. \quad (3.39)$$

考虑  $Ax = 0$  齐次线性方程组的解,

$$Ax = \begin{bmatrix} \vec{w}_1 \cdot x \\ \vec{w}_2 \cdot x \\ \vdots \\ \vec{w}_n \cdot x \end{bmatrix} = 0 \quad (3.40)$$

这意味着  $Ax = 0$  的解垂直于  $A$  的行向量子空间

$$N(A) \perp C(A^T). \quad (3.41)$$

## 4 正交投影

如果  $Ax = b$  无解,  $b \notin C(A)$ . 在这种情况下, 我们寻找一个最接近的  $b' \in C(A)$ , 此时  $e = \vec{b} - \vec{b}' \perp C(A)$ .

定义 4.1. 上述的  $b'$  称为  $b$  在空间  $C(A)$  中的正交投影.

例 4.1. 下面考虑一个矢量  $\vec{b}$  在另一个矢量  $\vec{a}$  上的投影  $\vec{p}$ ,

$$\vec{p} \parallel \vec{a}, \quad |\vec{p}| = |\vec{b}| \cos \theta \quad (4.1)$$

于是

$$\vec{p} = |\vec{b}| \cos \theta \frac{\vec{a}}{|\vec{a}|} = \frac{\vec{b} \cdot \vec{a}}{|\vec{a}|^2} \vec{a} = \frac{(a^T b)}{a^T a} a = \left( \frac{a a^T}{a^T a} \right) b \equiv P b. \quad (4.2)$$

上式中的  $P = \frac{a a^T}{a^T a}$  称为投影矩阵.

#### 4.1 投影矩阵

考虑一般情况, 有一组向量  $(v_1, v_2, \dots, v_n)$ , 这组向量张成一个线性子空间, 记为  $C(A)$ . 下面我们要将一个向量  $\vec{b}$  正交投影到这个空间, 投影后的向量记为  $\vec{p}$ . 正交投影意味着  $\vec{b} - \vec{p}$  垂直于  $C(A)$ .

回忆前面线性方程组的几何意义,  $(\vec{b} - \vec{p}) \perp C(A)$  等价于

$$A^T (b - p) = 0. \quad (4.3)$$

因为  $\vec{p}$  在  $C(A)$  中, 可以用  $(v_1, v_2, \dots, v_n)$  来线性表示, 表示系数记为  $\hat{x}$ , 具体来说,

$$p = \hat{x}_1 v_1 + \dots + \hat{x}_n v_n = \hat{x} A. \quad (4.4)$$

带入上面的垂直条件,

$$A^T b - A^T A \hat{x} = 0 \implies A^T A \hat{x} = A^T b. \quad (4.5)$$

如果  $A^T A$  可逆, 那么我们有

$$\hat{x} = (A^T A)^{-1} A^T b. \quad (4.6)$$

于是我们得到

$$p = A \hat{x} = \underbrace{\left[ A (A^T A)^{-1} A^T \right]}_{\text{投影矩阵 } P} b. \quad (4.7)$$

##### 4.1.1 投影矩阵的性质

- $P^2 = P$

$$P^2 = A (A^T A)^{-1} A^T A (A^T A)^{-1} A^T = A (A^T A)^{-1} A^T = P. \quad (4.8)$$

$$\bullet P^T = P$$

$$P^T = \left[ A (A^T A)^{-1} A^T \right]^T = A \left[ (A^T A)^{-1} \right]^T A^T = A (A^T A)^{-1} A^T. \quad (4.9)$$

例 4.2. 求投影矩阵

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (4.10)$$

计算投影矩阵  $P = A(AA^T)^{-1}A^T$ ,

$$A^T A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad (4.11)$$

$$(A^T A)^{-1} = \frac{1}{1 \times 2 - 1 \times 1} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}. \quad (4.12)$$

于是可得

$$\begin{aligned} P = A(A^T A)^{-1}A^T &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (4.13)$$

下面我们考虑什么时候  $A^T A$  可逆的问题.

**命题 4.1.**  $A^T A$  的零空间和  $A$  的零空间是一样的.

证明. 如果  $Ax = 0$ , 那么两边左乘  $A^T$ , 得到

$$A^T Ax = 0. \quad (4.14)$$

反之, 如果  $A^T Ax = 0$ , 两边左乘  $x^T$ , 得到

$$x^T A^T Ax = (Ax)^T Ax = 0. \quad (4.15)$$

也就是说,  $Ax$  的模长为零, 那它必然是零向量.  $\square$

根据上述命题, 我们可以发现: 假设  $A^T A$  可逆  $\iff A^T A x = 0$  只有零解  $\iff A x = 0$  只有零解.

这意味着:

- $A$  的列向量是线性无关的.
- $A$  的秩为  $r = n$ .

## 4.2 正交投影的应用

### 4.2.1 最小二乘法

在以后的学习中, 会经常遇到求以下函数的极小值

$$f(x) = |Ax - b| \quad (4.16)$$

其中的  $A$  是一个  $m \times n$  的矩阵,  $b$  是一个  $m \times 1$  的向量.

- 如果  $b \in C(A)$ , 那么  $f(x)$  的极小值为 0,  $x$  的解为  $Ax = b$  的解.
- 如果  $b \notin C(A)$ , 这时候极小值的  $x$  对应正交投影的坐标

$$\hat{x} = (A^T A)^{-1} A^T b. \quad (4.17)$$

### 4.2.2 线性回归

收集到一些数据,  $(y_1, t_1), (y_2, t_2), \dots, (y_m, t_m)$ . 假设  $y$  和  $t$  之间有一个线性关系  $y = Dt + C$ , 用数据去估计  $C$  和  $D$ .

估计方法: 考虑一个损失函数

$$L = \sum_{i=1}^m (y(t_i) - y_i)^2 = \sum_{i=1}^m (C + Dt_i - y_i)^2 = |Ax - b|^2 \quad (4.18)$$

其中

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}, \quad x = \begin{bmatrix} C \\ D \end{bmatrix}, \quad b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}. \quad (4.19)$$



我们做一些计算

$$A^T A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_m \end{bmatrix} \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} = \begin{bmatrix} m & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 \end{bmatrix}, \quad (4.20)$$

$$A^T b = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m y_i t_i \end{bmatrix}, \quad (4.21)$$

$$(A^T A)^{-1} = \frac{1}{m \sum_{i=1}^m t_i^2 - (\sum_{i=1}^m t_i)^2} \begin{bmatrix} \sum_{i=1}^m t_i^2 & -\sum_{i=1}^m t_i \\ -\sum_{i=1}^m t_i & m \end{bmatrix} \quad (4.22)$$

最后我们可以得到

$$C = \frac{\sum_{i=1}^m y_i \sum_{i=1}^m t_i^2 - \sum_{i=1}^m t_i \sum_{i=1}^m y_i t_i}{m \sum_{i=1}^m t_i^2 - (\sum_{i=1}^m t_i)^2}, \quad D = \frac{m \sum_{i=1}^m y_i t_i - \sum_{i=1}^m y_i \sum_{i=1}^m t_i}{m \sum_{i=1}^m t_i^2 - (\sum_{i=1}^m t_i)^2}. \quad (4.23)$$

#### 4.2.3 正交基

一组基需要满足条件:

- 线性无关.
- 任何向量都可以写成  $(v_1, v_2, \dots, v_m)$  的线性组合.

**定义 4.2.**  $(q_1, q_2, \dots, q_n)$  是一组基, 且满足

$$q_i \cdot q_j = \delta_{ij}. \quad (4.24)$$

则它们是正交归一基.

如果  $q_1, q_2, \dots, q_n$  是一组正交归一基, 那么对应的矩阵  $Q = [q_1, q_2, \dots, q_n]$  满足

$$Q^T Q = Q Q^T = I. \quad (4.25)$$

验证:

$$Q^T Q = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} [q_1^T, q_2^T, \dots, q_n^T] = \begin{bmatrix} \ddots & & & \\ & q_j^T q_i & & \\ & & \ddots & \\ & & & \ddots \end{bmatrix} = \begin{bmatrix} \ddots & & & \\ & \delta_{ij} & & \\ & & \ddots & \\ & & & \ddots \end{bmatrix} = I. \quad (4.26)$$

给定一组基, 可以构造一组正交归一基 (Gram-Schmit). 方法如下:

1. 选一个向量  $a$ , 令矩阵  $A = [a]$ .
2. 通过  $b$  构造一个向量  $B$ , 要求  $B$  垂直于  $A$  的列向量. 那么,

$$B = b - A(A^T A)^{-1} A^T b. \quad (4.27)$$

3. 通过  $c$  构造一个向量  $C$ ,

$$C = c - A(A^T A)^{-1} A^T c - B(B^T B)^{-1} B^T c, \quad (4.28)$$

下面验证它垂直于  $A$  和  $B$ : 令  $Q = [A \ B]$ , 则

$$Q^T Q = \begin{bmatrix} A^T \\ B^T \end{bmatrix} \begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} A^T A & 0 \\ 0 & B^T B \end{bmatrix} \quad (4.29)$$

所以

$$(Q^T Q)^{-1} = \begin{bmatrix} \frac{1}{A^T A} & 0 \\ 0 & \frac{1}{B^T B} \end{bmatrix} \quad (4.30)$$

得到

$$Q(Q^T Q)^{-1} Q^T = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} \frac{1}{A^T A} & 0 \\ 0 & \frac{1}{B^T B} \end{bmatrix} \begin{bmatrix} A^T \\ B^T \end{bmatrix} = A(A^T A)^{-1} A^T + B(B^T B)^{-1} B^T. \quad (4.31)$$

所以上面的  $C$  等价于将  $c$  减去  $A, B$  面内的投影, 自然  $C$  是垂直于  $A, B$  的.

4. 构造  $D$  垂直于  $A, B, C$ ,

$$D = d - A(A^T A)^{-1} A^T d - B(B^T B)^{-1} B^T d - C(C^T C)^{-1} C^T d. \quad (4.32)$$

$\vdots$

最终可以得到一组正交的基向量, 之后将它们归一化就得到了正交归一基向量.

#### 4.2.4 QR 分解

正交归一化的过程, 给出了一个可逆矩阵的  $QR$  分解,

$$A = QR \quad (4.33)$$

其中  $Q$  是正交矩阵,  $Q^{-1} = Q^T$ ,  $R$  是上三角矩阵.

正交归一化当中,

$$A = a \quad (4.34)$$

$$B = b - \frac{A^T b}{A^T A} A \quad (4.35)$$

$$C = c - \frac{A^T c}{A^T A} A - \frac{B^T c}{B^T B} B \quad (4.36)$$

整个过程可以用矩阵乘法表示. 正交归一基为

$$Q = \left[ \frac{A}{|A|}, \frac{B}{|B|}, \dots \right] \quad (4.37)$$

有

$$Q = AR \quad (4.38)$$

$R$  是一个上三角, 可以把原矩阵写为  $A = QR^{-1}$ .

## 5 行列式

我们之前讨论了秩  $r: M_{n \times n} \rightarrow \mathbb{Z}_+$ , 它在初等行变换下不变.

我们引入行列式

$$\delta: M_{n \times n} \rightarrow \mathbb{R} \quad (5.1)$$

是一个实数.

这个函数满足三个性质:

- 作用在单位阵上,  $\delta(I) = 1$
- 作用在行向量上是线性的

$$\delta \begin{bmatrix} \vdots \\ cA_i + c'B_i \\ \vdots \end{bmatrix} = c\delta \begin{bmatrix} \vdots \\ A_i \\ \vdots \end{bmatrix} + c'\delta \begin{bmatrix} \vdots \\ B_i \\ \vdots \end{bmatrix} \quad (5.2)$$

- 如果  $A$  有两行是一样的, 那么行列式为零.

下面研究行列式在初等变换下的性质

- 倍加变换:  $\delta(A') = \delta(A)$ .

证明. 用到了性质三

$$\delta(A') = \delta \begin{bmatrix} \vdots \\ w_j \\ \vdots \\ w_i + aw_j \\ \vdots \end{bmatrix} = \delta \begin{bmatrix} \vdots \\ w_j \\ \vdots \\ w_i \\ \vdots \end{bmatrix} + a\delta \begin{bmatrix} \vdots \\ w_j \\ \vdots \\ w_j \\ \vdots \end{bmatrix} = \delta(A) \quad (5.3)$$

□

- 换行变换:  $\delta(A') = -\delta(A)$ .

证明.

$$0 = \delta \begin{bmatrix} \vdots \\ w_j + w_i \\ \vdots \\ w_i + w_j \\ \vdots \end{bmatrix} = \delta \begin{bmatrix} \vdots \\ w_j \\ \vdots \\ w_i \\ \vdots \end{bmatrix} + \delta \begin{bmatrix} \vdots \\ w_i \\ \vdots \\ w_j \\ \vdots \end{bmatrix} = \delta(A') + \delta(A). \quad (5.4)$$

□

- 倍乘变换  $\delta(A') = c\delta(A)$ . 由线性性可得.

初等矩阵的行列式

- $\delta(E_{ij}(a)) = 1$

证明. 令  $I = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ , 则初等矩阵可以写为

$$E_{ij}(a) = \begin{bmatrix} \vdots \\ w_j \\ \vdots \\ w_i + aw_j \\ \vdots \end{bmatrix}. \quad (5.5)$$

于是

$$\delta(E) = \delta \begin{bmatrix} \vdots \\ w_j \\ \vdots \\ w_i + aw_j \\ \vdots \end{bmatrix} = \delta \begin{bmatrix} \vdots \\ w_j \\ \vdots \\ w_i \\ \vdots \end{bmatrix} + a\delta \begin{bmatrix} \vdots \\ w_j \\ \vdots \\ w_j \\ \vdots \end{bmatrix} = \delta(I) = 1. \quad (5.6)$$

□

- 换行  $\delta(P_{ij}) = -1$ . 证明同上, 把这个矩阵写成单位阵的换行即可.
- 倍乘  $\delta(S_i(c)) = c$ . 证明利用线性性.

总结上面的结论, 我们可以发现, 对于初等矩阵  $E$ ,

$$\delta(EA) = \delta(A). \quad (5.7)$$

并且, 如果一个矩阵某一行为零, 那么行列式为零.

**命题 5.1.** 满足行列式定义的三个性质的函数是唯一的.

证明. 对于任意矩阵  $A$ , 通过初等变换可以变为一个行约化阶梯形式. 这分为两种情况.

如果  $A'$  为单位矩阵, 那么

$$\delta(A) = \frac{1}{\delta(E_p) \cdots \delta(E_1)}. \quad (5.8)$$

如果  $A'$  不是单位矩阵, 那么  $A'$  的最后一行为零, 则

$$\delta(A) = 0. \quad (5.9)$$

□

行列式满足的一个重要性质:

$$\delta(AB) = \delta(A)\delta(B). \quad (5.10)$$

证明. 找到变换使得  $A$  变为行约化阶梯形式  $A'$

$$A' = (E_p \cdots E_1) A, \quad (5.11)$$

如果  $A'$  为单位矩阵, 那么  $\delta(A') = 1$

$$\delta(B) = \delta(A'B) = \delta(E_p \cdots E_1 AB) = \delta(E_p) \cdots \delta(E_1) \delta(AB), \quad (5.12)$$

所以

$$\delta(B) = \frac{\delta(AB)}{\delta(A)} \implies \delta(AB) = \delta(A)\delta(B). \quad (5.13)$$

如果  $A'$  不是单位矩阵, 那么  $\delta(A') = 0$ ,  $AB$  也不是满秩的, 所以

$$\delta(AB) = 0 = \delta(A)\delta(B). \quad (5.14)$$

□

### 5.0.1 行列式的递归定义

**定义 5.1.** 余矩阵:  $A_{ij}$ : 把  $A$  的第  $i$  行第  $j$  列去掉, 得到一个  $(n-1) \times (n-1)$  的矩阵.

**例 5.1.** 一个矩阵  $A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix}$ ,

$$A_{11} = \begin{bmatrix} 6 & 7 \\ 9 & 10 \end{bmatrix} \quad A_{21} = \begin{bmatrix} 3 & 4 \\ 9 & 10 \end{bmatrix} \quad A_{31} = \begin{bmatrix} 3 & 4 \\ 6 & 7 \end{bmatrix}. \quad (5.15)$$

行列式的递归定义:

$$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31} - \cdots + (-1)^{1+j} a_{1j} \det A_{1j}. \quad (5.16)$$

**例 5.2.**  $1 \times 1$  矩阵行列式  $\det[a] = a$ .

$2 \times 2$  矩阵行列式  $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$ .