

线代笔记

mny

2024 年 1 月 12 日

目录

1	线性空间	2
1.1	实数中运算的性质	3
1.2	$(\mathbb{R}^m, +, *)$ 为一个线性空间	4
1.3	矩阵	4
1.4	矩阵的乘法的应用	7
1.4.1	逆矩阵的一些性质	7
1.4.2	线性组合的矩阵乘法表示	9
1.4.3	矩阵方程	9
2	矩阵的初等变换	10
2.1	初等变换的应用	11
2.2	用行约化阶梯形式求解线性方程组	14
3	线性方程组的解	16
3.1	齐次线性方程解空间的性质	16
3.2	一些概念	17
3.3	线性代数基本定理	21
4	正交投影	24
4.1	投影矩阵	24
4.1.1	投影矩阵的性质	25
4.2	正交投影的应用	26
4.2.1	最小二乘法	26
4.2.2	线性回归	27

1 线性空间	2
4.2.3 正交基	28
4.2.4 QR 分解	29
5 行列式	30
5.1 行列式的定义和唯一性	30
5.2 行列式的递归定义	33
5.3 行列式的应用	36
5.3.1 克拉默 (Cramer) 法则求解线性方程组	36
5.3.2 用行列式求逆的公式	37
6 特征值和特征向量	38
6.1 特征多项式的系数和特征值的关系	40
6.2 特征值的一些简单性质	40
6.3 特征向量的一些简单性质	41
6.4 相似对角化	41
6.4.1 对称矩阵	44
6.5 特征值的应用	46
6.5.1 求 A^n	46
6.5.2 求解矩阵线性微分方程	47
6.5.3 求解非矩阵二阶微分方程	48
6.5.4 奇异值分解	48
6.5.5 奇异值分解的应用	50
7 线性映射	52
7.1 对偶空间	54
7.2 内积的推广	55
7.3 不变子空间	55
7.4 复线性空间	56
8 一般域上的线性空间	56

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, $A = [\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n]$

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

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:

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

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

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

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,

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

- 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.

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

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:

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

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 .¹

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. 2×2 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

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

¹将在后面证明.

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

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

引入两个矩阵

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

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 $[\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n]$ 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

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

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 $P_{ij} = (\text{交换单位矩阵的 } i, j \text{ 列})$ 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

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

最终的增广矩阵右侧就是 A 的逆,

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

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.

$$A \rightarrow \left[\begin{array}{ccc|ccc} 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & & & \\ 0 & 0 & 0 & & & \\ 0 & 0 & 0 & & & \end{array} \begin{array}{c} \\ B \\ \end{array} \right] \tag{2.21}$$

- 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:

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

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2.3, 约化后的方程为

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

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.

1. 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

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

2. 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.

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

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,

$$v = x_1e_1 + x_2e_2 + \dots + x_me_m. \quad (3.5)$$

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

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

其中 (a_1, a_2, \cdots, a_m) 称为 v 在基 (e_1, e_2, \cdots, 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:

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

两式相减得到

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

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 $(e_1, e_2, \cdots, e_m), (f_1, f_2, \cdots, f_n)$, $n > m$.

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

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

把上述过程写成矩阵乘法的形式

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

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 $m \times m$ 矩阵. 如果 A 是不可逆的, 则 $Ax = 0$ 有非零解即

$$\begin{bmatrix} f_1, f_2, \dots, f_m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = 0 \quad (3.14)$$

这与 f_i 这组基的定义矛盾. □

矩阵的转置

定义 3.3. 给定一个矩阵 A ,

$$(A^T)_{ij} = A_{ji}. \quad (3.15)$$

A^T 把 A 的行变成列.

转置的一些性质

-

$$(A^T)^T = A. \quad (3.16)$$

-

$$(AB)^T = B^T A^T \quad (3.17)$$

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

$$(AB)_{ij} = \sum_{l=1}^k a_{il} b_{lj}. \quad (3.18)$$

根据转置的定义, 有

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

另一方面,

$$(B^T A^T)_{ij} = \sum_{l=1}^k (B^T)_{li} (A^T)_{jl} = \sum_{l=1}^k a_{jl} b_{li}. \quad (3.20)$$

□

•

$$(A^T)^{-1} = (A^{-1})^T. \quad (3.21)$$

证明. 因为

$$AA^{-1} = I, \quad (3.22)$$

两边取转置得到

$$(A^{-1})^T A^T = I \implies (A^T)^{-1} = (A^{-1})^T. \quad (3.23)$$

□

特殊矩阵

- 对称矩阵: $A^T = A$.

$$A_{ij} = A_{ji}. \quad (3.24)$$

例 3.4.

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

例 3.5.

$$\begin{bmatrix} 2 & 3 & 1 \\ 3 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}. \quad (3.26)$$

- 反对称矩阵: $A^T = -A$. 可知, 其对角线都为零.

例 3.6.

$$\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}. \quad (3.27)$$

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.

$$r_1 = r_2 = r = r'. \quad (3.28)$$

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

$$w' = (w_1, w_2, \dots, w_i + aw_j, \dots, w_m). \quad (3.29)$$

$V(w')$ 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

$$w' = x_1 w'_1 + \dots + x_m w'_m = x_1 w_1 + \dots + (x_j + ax_i)w_j + \dots + x_m w_m. \quad (3.30)$$

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

$$Ax = 0 \iff Bx = 0. \quad (3.31)$$

其中 B 为 A 初等行变换后的矩阵. 这也就是说,

$$x_1 v_1 + \dots + x_n v_n = 0 \iff x_1 v'_1 + \dots + x_n v'_n = 0. \quad (3.32)$$

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,

$$A = \begin{bmatrix} a_1 v_i, \dots, v_i, \dots, a_n v_i \end{bmatrix}. \quad (3.33)$$

(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).

$$\begin{aligned} \dim(N(A)) &= n - r \\ \dim(N(A^T)) &= m - r \\ \dim(C(A)) &= r \\ \dim(C(A^T)) &= r \end{aligned} \quad (3.34)$$

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

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^m v_i w_i. \quad (3.35)$$

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,

$$\vec{v} \cdot \vec{w} = v^T w = w^T v. \quad (3.36)$$

有了内积, 可以定义一些东西

- 向量 v 的长度

$$|v| = \sqrt{v \cdot v} = \sqrt{v^T v}. \quad (3.37)$$

- 两个向量垂直 $v \perp w$, 如果

$$v \cdot w = 0. \quad (3.38)$$

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

$$A' = [A \ b], \quad \text{rank}(A') = \text{rank } A \quad (3.39)$$

无解时,

$$\text{rank}(A') = \text{rank}(A) + 1. \quad (3.40)$$

考虑 $Ax = 0$ 齐次线性方程组的解,

$$Ax = \begin{bmatrix} \vec{w}_1 \cdot x \\ \vec{w}_2 \cdot x \\ \vdots \\ \vec{w}_n \cdot x \end{bmatrix} = 0 \quad (3.41)$$

这意味着 $Ax = 0$ 的解垂直于 A 的行向量空间

$$N(A) \perp C(A^T). \quad (3.42)$$

例 3.9. 对于两个矩阵 A, B , 有

$$\text{rank}(A + B) \leq \text{rank } A + \text{rank } B. \quad (3.43)$$

证明. 令 $A = [v_1, v_2, \dots, v_n]$, $B = [w_1, w_2, \dots, w_n]$, 则

$$A + B = [v_1 + w_1, v_2 + w_2, \dots, v_n + w_n]. \quad (3.44)$$

这个矩阵的列空间是 A 和 B 的列空间的直和, 即 $A \oplus B$.

另一方面,

$$\text{rank}(A \oplus B) \leq \text{rank } A + \text{rank } B. \quad (3.45)$$

□

例 3.10. 一个矩阵 A , $A^2 = A$ 当且仅当

$$\text{rank } A + \text{rank}(I - A) = n. \quad (3.46)$$

证明. 因为 $A^2 = A$, 即 $A(I - A) = 0$, 所以 $C(I - A) \subset N(A)$, 即

$$\text{rank } A + \text{rank } (I - A) \leq n. \quad (3.47)$$

另一方面, 因为

$$\text{rank } I \geq \text{rank } (I - A) + \text{rank } A, \quad (3.48)$$

所以 $\text{rank } A + \text{rank } (I - A) = n$. \square

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

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

上式中的 $P = \frac{aa^T}{a^T a}$ 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

- $P^2 = P$

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

- $P^T = P$

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

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 $P = A (A A^T)^{-1} A^T$,

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

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

于是可得

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

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

$$A^T Ax = 0. \tag{4.14}$$

反之, 如果 $A^T Ax = 0$, 两边左乘 x^T , 得到

$$x^T A^T Ax = (Ax)^T Ax = 0. \tag{4.15}$$

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 $\iff A^T Ax = 0$ 只有零解 $\iff Ax = 0$ 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

$$f(x) = |Ax - b| \tag{4.16}$$

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:

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

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

5.1 行列式的定义和唯一性

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

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

□

5.2 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. 1×1 矩阵行列式 $\det[a] = a$. 2×2 矩阵行列式 $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$.

行列式的几个性质:

- $\delta(A) \neq 0$, 当且仅当:
 1. A 是可逆的
 2. A 满秩
 3. A 的列向量线性无关
 4. $Ax = 0$ 只有零解
 5. A 可以行约化为单位矩阵
- $\delta(A) = 0$, 当且仅当:
 1. A 不可逆
 2. A 不满秩
 3. A 的列向量线性相关
 4. $Ax = 0$ 有非零解
 5. A 不能行约化为单位矩阵

命题 5.2. 上文定义的行列式满足 $\det I = 1$.

证明. 可以用递归定义验证. □

命题 5.3. 上文定义的行列式作用在行向量上是线性的.

$$\text{设 } D = \begin{bmatrix} \vdots \\ ca_k + c'b_k \\ \vdots \end{bmatrix}, A = \begin{bmatrix} \vdots \\ a_k \\ \vdots \end{bmatrix}, B = \begin{bmatrix} \vdots \\ b_k \\ \vdots \end{bmatrix}, \text{ 那么}$$

$$\det D = c \det A + c' \det B. \quad (5.17)$$

证明. 根据上面的递归定义, 我们把 D, A, B 的行列式展开为

$$\det D = \sum (-)^{\mu+1} d_{\mu 1} \det D_{\mu 1}, \quad (5.18)$$

$$\det A = \sum (-)^{\mu+1} a_{\mu 1} \det A_{\mu 1}, \quad (5.19)$$

$$\det B = \sum (-)^{\mu+1} b_{\mu 1} \det B_{\mu 1}. \quad (5.20)$$

我们需要证明

$$d_{\mu 1} \det D_{\mu 1} = ca_{\mu 1} \det A_{\mu 1} + c'b_{\mu 1} \det B_{\mu 1}. \quad (5.21)$$

分情况讨论, 如果 $\mu = k$, 那么三个余子式是一样的, 而前面的系数满足 $d_{k1} = ca_{k1} + c'b_{k1}$ 所以等式成立.

如果 $\mu \neq k$, 那么 $d_{\mu 1} = a_{\mu 1} = b_{\mu 1}$, 有递归假设

$$\det D_{\mu 1} = c \det A_{\mu 1} + c' \det B_{\mu 1}, \quad (5.22)$$

那么

$$d_{\mu 1} \det D_{\mu 1} = ca_{\mu 1} \det A_{\mu 1} + c'b_{\mu 1} \det B_{\mu 1}. \quad (5.23)$$

□

命题 5.4. 如上定义的行列式, 如果 A 有两行是一样的, 那么 $\det A = 0$.

证明. 不妨设 A 的第 k 行和第 $k+1$ 行是一样的. 那么有

$$a_{k1} = a_{k+1,1}, \quad \det A_{k1} = \det A_{k+1,1}. \quad (5.24)$$

由递归假设, $\det A_{i1} = 0, i \neq k, k+1$.

$$\det A = (-)^{k+1} a_{k1} \det A_{k1} + (-)^{k+2} a_{k+1,1} \det A_{k+1,1} = 0. \quad (5.25)$$

□

一些特殊矩阵的行列式:

- 对角矩阵的行列式等于对角线上元素的乘积.

$$\det \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} = d_1 d_2 \cdots d_n. \quad (5.26)$$

- 上三角矩阵的行列式等于对角线上元素的乘积.

$$\det \begin{bmatrix} d_1 & * & * & * \\ & d_2 & * & * \\ & & \ddots & * \\ & & & d_n \end{bmatrix} = d_1 d_2 \cdots d_n. \quad (5.27)$$

- 下三角矩阵的行列式等于对角线上元素的乘积.

$$\det \begin{bmatrix} d_1 & & & \\ * & d_2 & & \\ * & * & \ddots & \\ * & * & * & d_n \end{bmatrix} = d_1 d_2 \cdots d_n. \quad (5.28)$$

命题 5.5. 行列式可以用任意一行或者一列展开.

用行展开, 用 A 的第 i 行展开, 有

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}. \quad (5.29)$$

用列展开, 用 A 的第 j 列展开, 有

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}. \quad (5.30)$$

例 5.3. 计算

$$\det A = \det \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 1 & 0 & 2 \end{bmatrix} \quad (5.31)$$

我们对第 1 列展开,

$$\det A = 1 \det \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} - 2 \det \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} + 1 \det \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} = 1 \cdot 6 - 2 \cdot 4 + 1 \cdot (-7) = -9. \quad (5.32)$$

对第 1 行展开

$$\det A = 1 \cdot 6 - 2 \cdot 3 + 3 \cdot (-3) = -9. \quad (5.33)$$

对第 2 行展开

$$\det A = -2 \cdot 4 + 3 \cdot (-1) - 1 \cdot (-2) = -9. \quad (5.34)$$

行列式的置换定义:

$$\det A = \sum_p \operatorname{sgn}(p) a_{1, p_1} a_{2, p_2} \cdots a_{n, p_n}, \quad (5.35)$$

其中的 p 为一个 n 阶置换 $p: \{1, 2, \cdots, n\} \rightarrow \{1, 2, \cdots, n\}$ 的一一映射, $\operatorname{sgn}(p)$ 为置换 p 的符号.

例 5.4. 三阶置换群的群元:

$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 1 & 2 & 3 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 2 & 1 & 3 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 3 & 2 & 1 \\ \hline \end{array}$
$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 1 & 3 & 2 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 2 & 3 & 1 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 3 & 1 & 2 \\ \hline \end{array}$

例 5.5. 用置换的方法计算三阶行列式.

$$\begin{aligned} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} &= \sum_p \operatorname{sgn}(p) a_{1, p_1} a_{2, p_2} a_{3, p_3} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \\ &\quad + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned} \quad (5.36)$$

5.3 行列式的应用

5.3.1 克拉默 (Cramer) 法则求解线性方程组

定理 5.1. 设 A 是一个 n 阶方阵, $\det(A) \neq 0$, b 是一个 n 维列向量, 那么线性方程组

$$Ax = b \quad (5.37)$$

有唯一解

$$x_i = \frac{\det(B_i)}{\det(A)}, \quad i = 1, 2, \cdots, n, \quad (5.38)$$

其中 B_i 是把 A 的第 i 列换成 b 得到的矩阵,

$$A = [v_1, v_2, \dots, v_n], B_i = [v_1, v_2, \dots, b, \dots, v_n]. \quad (5.39)$$

证明. 对于矩阵 $A = [v_1, v_2, \dots, v_n]$, 考虑下面的矩阵方程

$$A [E_1, \dots, x, \dots, E_n] = [v_1, v_2, \dots, b, \dots, v_n]. \quad (5.40)$$

两边取行列式, 有

$$\det(A) \det([E_1, \dots, x, \dots, E_n]) = \det([v_1, v_2, \dots, b, \dots, v_n]). \quad (5.41)$$

我们需要计算上式左侧的行列式, 不难发现,

$$\det \begin{bmatrix} 1 & 0 & \cdots & x_1 & \cdots & 0 \\ 0 & 1 & \cdots & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n & \cdots & 1 \end{bmatrix} = x_i. \quad (5.42)$$

所以

$$(\det A) x_i = \det(B_i). \quad (5.43)$$

即

$$x_i = \frac{\det(B_i)}{\det(A)}. \quad (5.44)$$

□

5.3.2 用行列式求逆的公式

定理 5.2. 我们构造一个代数余子式矩阵

$$M_{ij} = (-1)^{i+j} \det A_{ij}, \quad (5.45)$$

那么 A 的逆矩阵为

$$A^{-1} = \frac{1}{\det A} M^T. \quad (5.46)$$

其中的 M^T 称为 A 的伴随矩阵, 记为 A^* .

例 5.6. 求 A 的逆矩阵,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 3 & 1 \\ 2 & 1 & 0 \end{bmatrix}. \quad (5.47)$$

$$M = \begin{bmatrix} -1 & 2 & -2 \\ 3 & -6 & 3 \\ -7 & 11 & -5 \end{bmatrix}, \quad (5.48)$$

$$A^{-1} = -\frac{1}{3} \begin{bmatrix} -1 & 3 & -7 \\ 2 & -6 & 11 \\ -2 & 3 & -5 \end{bmatrix}. \quad (5.49)$$

证明逆矩阵公式. 求 A 的逆, 假设 $A^{-1} = [w_1, w_2, \dots, w_n]$, 那么由于 $AA^{-1} = I$,

$$A[w_1, w_2, \dots, w_n] = I \implies [Aw_1, Aw_2, \dots, Aw_n] = [E_1, E_2, \dots, E_n]. \quad (5.50)$$

需要求解线性方程组,

$$Aw_i = E_i. \quad (5.51)$$

由克拉默法则,

$$w_j = \frac{\det(B_j)}{\det(A)}. \quad (5.52)$$

由于 B_i 是把 A 的第 i 列换成 E_i 得到的矩阵, 所以

$$\det(B_j) = (-1)^{i+j} \det(A_{ij}) = M_{ij}. \quad (5.53)$$

□

6 特征值和特征向量

考虑下列线性方程组 (特征方程),

$$A\vec{x} = \lambda\vec{x}. \quad (6.1)$$

\vec{x} 是一个非零的向量, 称为 A 的特征向量, 常数 λ 称为 A 的特征值.

例 6.1. 求矩阵的特征向量和特征值 $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

解 特征方程可以写为 $(A - \lambda I)\vec{x} = 0$, 即

$$\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (6.2)$$

得到

$$\begin{cases} (1-\lambda)x_1 + x_2 = 0 \\ (1-\lambda)x_2 = 0 \end{cases} \quad (6.3)$$

所以只有 $\lambda = 1$ 时才有非零解. 此时, $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

例 6.2. 求矩阵的特征向量和特征值 $A = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}$

解 特征方程为 $(A - \lambda I)\vec{x} = 0$, 即

$$\begin{bmatrix} 1-\lambda & 3 \\ 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (6.4)$$

得到

$$\begin{cases} (1-\lambda)x_1 + 3x_2 = 0 \\ x_1 + (2-\lambda)x_2 = 0 \end{cases} \quad (6.5)$$

把第二个方程带入第一个方程, 有

$$[(1-\lambda)(2-\lambda) - 3]x_2 = 0. \quad (6.6)$$

所以 λ 有两个解.

考虑一般的特征方程,

$$(A - \lambda I)\vec{x} = 0. \quad (6.7)$$

所以 $A - \lambda I$ 的零空间维数大于等于一. 那么特征方程有非零解的充要条件就是

$$\det(A - \lambda I) = 0. \quad (6.8)$$

计算 $|\lambda I - A|$,

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}. \quad (6.9)$$

这个行列式的值是一个关于 λ 的多项式, 称为特征多项式, 记为 $p(\lambda)$. 我们可以发现, λ 的最高次幂和次高次幂都来自于对角元的乘积.

我们把特征方程展开,

$$p(\lambda) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_n = 0. \quad (6.10)$$

通过观察可以发现,

$$c_1 = -(a_{11} + a_{22} + \cdots + a_{nn}) = -\operatorname{tr} A. \quad (6.11)$$

$$c_n = |-A| = (-1)^n \det A. \quad (6.12)$$

特征多项式有唯一分解:

$$p(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n). \quad (6.13)$$

- λ_i 可能是复数.
- λ_i 可能重合, 这一个特征值的代数重数为 λ_i 在上述分解中出现的次数.
- 特征值必定存在, 至少一个.
- 所有特征值的代数重数之和等于 n .

6.1 特征多项式的系数和特征值的关系

$$\begin{cases} c_1 = -(a_{11} + a_{22} + \cdots + a_{nn}) = -\operatorname{tr} A = -(\lambda_1 + \lambda_2 + \cdots + \lambda_n) \\ c_n = |-A| = (-1)^n \det A = (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n \end{cases} \quad (6.14)$$

所以,

$$\sum_i \lambda_i = \operatorname{tr} A, \quad \prod_i \lambda_i = \det A. \quad (6.15)$$

6.2 特征值的一些简单性质

- 如果 λ 为 A 的特征值, 那么 λ^k 也是 A^k 的特征值, 因为

$$A^k \vec{x} = \lambda^k \vec{x}. \quad (6.16)$$

- 如果 A 可逆, 则 $\lambda \neq 0$, 因为

$$Ax = \lambda x \implies A^{-1}x = \frac{1}{\lambda}x. \quad (6.17)$$

6.3 特征向量的一些简单性质

- 固定一个特征值, 所有对应的特征向量和零向量张成一个线性空间, 称为特征向量子空间, 记为 $V(\lambda)$.

1. 加法下封闭: $\vec{x}_1, \vec{x}_2 \in V(\lambda)$

$$A(\vec{x}_1 + \vec{x}_2) = \lambda(\vec{x}_1 + \vec{x}_2) \implies \vec{x}_1 + \vec{x}_2 \in V(\lambda) \quad (6.18)$$

2. 数乘下封闭: $\vec{x} \in V(\lambda)$

$$A(c\vec{x}) = c\lambda\vec{x} \implies c\vec{x} \in V(\lambda) \quad (6.19)$$

- 对于一个代数重数为 p 的特征值, 对应的特征向量子空间的维数称作几何重数, 满足几何重数 \leq 代数重数.
- 不同特征值对应的特征向量是线性无关的.

证明. 只考虑两个特征值的情况, 设 λ_1, λ_2 为特征值, x, y 为对应的特征向量, 我们需要证明方程 $c_1x + c_2y = 0$ 只有零解.

把 A 作用到这个方程, 得到

$$c_1\lambda_1x + c_2\lambda_2y = 0. \quad (6.20)$$

这时候得到了两个方程, 消去 x ,

$$c_2(\lambda_2 - \lambda_1)y = 0. \quad (6.21)$$

因为 $\lambda_1 \neq \lambda_2$, 所以 $c_2 = 0$, 同理 $c_1 = 0$. \square

6.4 相似对角化

我们之前讨论的把矩阵变为“标准形式”有

- 行变换 \longrightarrow 行约化阶梯形.
- 行变换, 对于可逆矩阵 $A = LDU$.
- 正交对角化, 对于可逆矩阵 $A = QR$.

一般来说, 行变换是改变特征值的, 对于 $Ax = \lambda x$, 它的初等变换

$$A' = EA \quad (6.22)$$

原本的 λ 不是 A' 的特征值.

我们引进一种保持特征值的变换: 相似变换.

定义 6.1. 对于一个矩阵 A , 它的相似变换为 $C = P^{-1}AP$, 其中 P 为可逆矩阵.

命题 6.1. 在相似变换下, 矩阵的特征值不变.

证明. 假设 λ 是 A 的特征值, 特征向量为 x , 有 $Ax = \lambda x$. 对于 $C = P^{-1}AP$, 有

$$C(P^{-1}x) = (P^{-1}AP)P^{-1}x = P^{-1}Ax = P^{-1}\lambda x = \lambda(P^{-1}x). \quad (6.23)$$

并且 $P^{-1}x$ 为非零向量.

反之, 对于 C 的一个特征值 λ 及对应的特征向量 x , $Cx = \lambda x$, 有

$$A(Px) = PCP^{-1}(Px) = PCx = P\lambda x = \lambda(Px). \quad (6.24)$$

□

命题 6.2. A 的特征多项式和 C 的特征多项式是一样的.

证明. 我们有 $P_A(\lambda) = \det(\lambda I - A)$.

对于 $C = P^{-1}AP$, 有

$$\begin{aligned} P_C(\lambda) &= \det(\lambda I - P^{-1}AP) = \det(\lambda P^{-1}P - P^{-1}AP) \\ &= \det(P^{-1}(\lambda I - A)P) \\ &= \det(P^{-1}) \det(\lambda I - A) \det(P) \\ &= \det(\lambda I - A) = P_A(\lambda). \end{aligned} \quad (6.25)$$

□

相似变换对角化: 如果矩阵 $A_{m \times n}$ 有 n 个线性无关的特征向量,

$$Ax_i = \lambda_i x_i, \quad i = 1, 2, \dots, n. \quad (6.26)$$

那么存在一个相似矩阵 P , 使得相似变换后的矩阵为对角矩阵 $\Lambda = P^{-1}AP$, Λ 的对角元为特征值.

证明. 把 $Ax_i = \lambda_i x_i$ 写成矩阵乘法的形式, 令 $P = [x_1, x_2, \dots, x_n]$,

$$A \underbrace{\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}}_P = \underbrace{\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}}_P \underbrace{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}}_{\Lambda}. \quad (6.27)$$

也就是说 $\Lambda = P^{-1}AP$.

反之, 如果 A 可以相似对角化, 那么 A 有 n 个线性无关的特征向量, 特征值为矩阵的对角元素. \square

例 6.3. 对于矩阵 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, 它不能相似对角化.

定理 6.1. 一个特征值的几何重数小于等于它的代数重数.

证明. 假设 λ 的特征向量的几何重数为 m , 特征向量为 x_1, x_2, \dots, x_m , 且满足 $Ax_i = \lambda x_i$, $i = 1, 2, \dots, m$.

添加向量构造 \mathbb{R}^n 中的一组基 $(x_1, x_2, \dots, x_m, e_1, e_2, \dots, e_{n-m})$.

令 $P = [x_1, x_2, \dots, x_m, e_1, e_2, \dots, e_{n-m}]$, 把 $Ax_i = \lambda x_i$ 写成矩阵乘法的形式

$$A \underbrace{\begin{bmatrix} x_1, x_2, \dots, x_m, e_1, e_2, \dots, e_{n-m} \end{bmatrix}}_P = \underbrace{\begin{bmatrix} x_1, x_2, \dots, x_m, e_1, e_2, \dots, e_{n-m} \end{bmatrix}}_P \underbrace{\begin{bmatrix} \lambda & 0 & \cdots & 0 & * & * \\ 0 & \lambda & \ddots & 0 & * & * \\ 0 & 0 & \ddots & 0 & * & * \\ \vdots & \ddots & 0 & \lambda & * & * \\ 0 & \ddots & & 0 & * & * \\ 0 & 0 & \ddots & 0 & * & * \\ 0 & 0 & \cdots & 0 & * & * \end{bmatrix}}_C. \quad (6.28)$$

因为 C 和 A 的特征多项式相同 (由相似变换性质), 又因为 C 至少有 m 个相同的特征值 λ (可以通过计算 C 的特征多项式很容易地看出来) $\implies A$ 至少有 m 个为 λ 的特征值. \square

不是每一个矩阵都可以相似对角化的, 但是每一个矩阵都可以相似上三角化. 即, 存在可逆矩阵 P , 使得 $P^{-1}AP$ 为上三角矩阵, 且对角元为特征值.

证明. A 一定有一个特征值 λ_1 和一个特征向量 v_1 , $Av_1 = \lambda_1 v_1$.

下一步添加向量构成一组基 $(v_1, e_1, e_2, \dots, e_{n-1})$. 令 $P = [v_1, e_1, e_2, \dots, e_{n-1}]$, 把特征方

程写成矩阵乘法的形式,

$$A \begin{bmatrix} v_1, e_1, e_2, \dots, e_{n-1} \end{bmatrix} = \begin{bmatrix} v_1, e_1, e_2, \dots, e_{n-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & \vdots & B_1 \\ 0 & \ddots & \vdots \\ \vdots & & A_1 \\ 0 & \vdots & \vdots \end{bmatrix} \quad (6.29)$$

之后再对子矩阵 A_1 做相同的操作. \square

命题 6.3. 如果 A 有 n 个不同的特征值, 那么 A 可以对角化.

证明. 不同特征值对应的特征向量是线性无关的, 所以可以找到 n 个线性无关的特征向量 $\implies A$ 可以对角化. \square

问题二: 矩阵的对角化是否唯一? 不唯一.

问题三: 如果一个矩阵可以相似对角化, 那么 P 是否可以选为正交矩阵? 不可以.

所有矩阵都可以用相似上三角化成若当标准型.

定义 6.2. 一个 m 阶若当块 $J_m(\lambda)$ 为对角元为 λ , 对角元上方为 1, $J_m(\lambda) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$.

所有矩阵都可以化为

$$P^{-1}AP = \begin{bmatrix} J_{n_1}(\lambda_1) & & \\ & J_{n_2}(\lambda_2) & \\ & & \ddots \\ & & & J_{n_s}(\lambda_s) \end{bmatrix} \quad (6.30)$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_s$ 为特征值, 并且 $\sum_{i=1}^s n_i = n$.

6.4.1 对称矩阵: 永远可以对角化

定理 6.2. 对称矩阵的特征值都是实数

证明. 考虑矩阵方程 $Sx = \lambda x$, 取复共轭 $\bar{S}\bar{x} = \bar{\lambda}\bar{x}$. 第一个方程左乘 \bar{x}^T , 第二个方程左乘 x^T , 得到

$$\bar{x}^T Sx = \lambda \bar{x}^T x, \quad x^T S\bar{x} = \bar{\lambda} x^T \bar{x}. \quad (6.31)$$

注意到

$$\bar{x}^T Sx = (Sx)^T \bar{x} = x^T S^T \bar{x} = x^T S\bar{x}, \quad \bar{x}^T x = x^T \bar{x}. \quad (6.32)$$

并且由于 $x \neq 0$, 得到 $\bar{\lambda} = \lambda$. \square

定理 6.3. 对称矩阵的不同特征值的特征向量正交.

证明. 考虑两个特征值对应的特征向量 x, y , 有

$$x^T S y = \lambda_1 x^T y = x^T \lambda_2 y = \lambda_2 x^T y. \quad (6.33)$$

于是

$$(\lambda_1 - \lambda_2) x^T y = 0. \quad (6.34)$$

由于 $\lambda_1 \neq \lambda_2$, 所以 $x^T y = 0$. \square

定理 6.4. 任何对称矩阵都可以相似对角化.

证明. 取一个特征向量 x_1 和特征值 λ_1 , 并且规定 x_1 的长度为 1. 添加向量构成一组正交归一基 $X = [x_1, e_1, e_2, \dots, e_{n-1}]$, 满足

$$X^T X = I, \quad (6.35)$$

则

$$AX = X \begin{bmatrix} \lambda_1 & & a \\ 0 & & \\ \vdots & & B \\ 0 & & \end{bmatrix}. \quad (6.36)$$

由 X 的正交性,

$$X^T AX = \begin{bmatrix} \lambda_1 & & a \\ 0 & & \\ \vdots & & B \\ 0 & & \end{bmatrix} \quad (6.37)$$

两边取转置, $(X^T AX)^T = X^T A^T X = X^T AX$, 这说明上式分块矩阵中的 $a = 0$, $B = B^T$.

之后再如上操作即可“正交相似对角化”. \square

定义 6.3. 实正定矩阵满足

- 对称矩阵
- 特征值都是正的 $\lambda_i > 0$

对于正定矩阵, 下面的不等式成立:

$$x^T S x \geq 0 \quad (6.38)$$

且取等当且仅当 x 为零向量.

证明. 利用对称矩阵可以正交对角化, 即存在一个正定矩阵 P , 使得 $\Lambda = P^T S P$, $P^T P = I$. 有,

$$x^T S x = x^T (P P^T) S (P P^T) x = (P^T x)^T \Lambda (P^T x) \quad (6.39)$$

上式是 $P^T x$ 的每个分量的二次函数. \square

反之, 如果一个矩阵 S 定义的二次型 $x^T S x$ 满足 $x^T S x \geq 0$ 且当且仅当 $x = 0$ 时取等, 则这个矩阵一定是正交矩阵.

证明. 需要证明 S 的每个特征值都是正的. 由于 S 是对称矩阵, 所以可以正交对角化, 即存在一个正交矩阵 P , 使得 $P^T S P = \Lambda$, $P^T P = I$. 有 $(P^T x)^T \Lambda (P^T x) \geq 0$. 取特殊的向量 $P^T x$, 可得 S 的每个特征值都是正的. \square

命题 6.4. 如果 A 是一个可逆矩阵, 那么 $S = A^T A$ 是一个正定矩阵.

证明. S 是对称矩阵, 且 $x^T S x = |Ax|^2$. \square

命题 6.5. 任意正定矩阵都可以写成 $S = A^T A$ 的形式, 且 A 为可逆矩阵.

证明. 由于 S 是对称矩阵, 所以可以正交对角化, 即存在一个正交矩阵 P , 使得 $P^T S P = \Lambda$, $P^T P = I$. 令 $A = P \sqrt{\Lambda} P^T$, 则 $S = A^T A$. \square

顺序主子式:

定义 6.4. 顺序主子阵: 从矩阵的第一行和第一列开始, 依次取第 i 行和第 i 列, 得到的子阵.

命题 6.6. 正交矩阵 \iff 所有的顺序主子阵的行列式大于零.

6.5 特征值的应用

6.5.1 求 A^n

若 A 可以对角化, $\Lambda = P^{-1} A P$, 则

$$A^n = P \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m^n \end{bmatrix} P^{-1}. \quad (6.40)$$

例 6.4. 求斐波那契数列的第 100 项. ($F_1 = F_2 = 0, F_n = F_{n-1} + F_{n-2}$)

解 令 $u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$, 则有

$$u_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} u_k \equiv A u_k. \quad (6.41)$$

求 A 的特征值, 特征向量, 可以得到

$$A^n = P \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix} P^{-1}, \quad (6.42)$$

$$P = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}, \quad P^{-1} = \frac{\sqrt{5}}{5} \begin{bmatrix} 1 & \frac{\sqrt{5}-1}{2} \\ -1 & \frac{\sqrt{5}+1}{2} \end{bmatrix} \quad (6.43)$$

于是

$$A^n = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & \frac{\sqrt{5}-1}{2} \\ -1 & \frac{\sqrt{5}+1}{2} \end{bmatrix} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \quad (6.44)$$

$$= \frac{\sqrt{5}}{5} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} & \left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \\ \left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n & \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1} \end{bmatrix}$$

得到

$$F_n = \frac{\sqrt{5}}{5} \left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \right] \quad (6.45)$$

6.5.2 求解矩阵线性微分方程

方程 $\frac{du}{dt} = Au$, 将 A 对角化之后可以直接求解, 设 $\Lambda = P^{-1}AP$, 则

$$\frac{d(P^{-1}u)}{dt} = \Lambda P^{-1}u, \quad (6.46)$$

于是解为

$$u = P e^{\Lambda t} P^{-1} u_0 = P \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix} P^{-1} u_0. \quad (6.47)$$

6.5.3 求解非矩阵二阶微分方程

对于方程 $u'' + Bu' + Cu$, 其中 B, C 都是数, $u = u(t)$. 它可以写为下面的形式

$$\frac{d}{dt} \begin{bmatrix} u \\ u' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -C & -B \end{bmatrix} \begin{bmatrix} u \\ u' \end{bmatrix}. \quad (6.48)$$

之后就可以用对角化的方法求解.

6.5.4 奇异值分解

对于任意矩阵都有下列的分解

$$A = U\Sigma V^T, \quad (6.49)$$

其中 U 为 $m \times m$ 矩阵, V 为 $n \times n$ 的正交矩阵, Σ 为 $m \times n$ 矩阵, Σ 的对角线上的元素为奇异值, 且奇异值按照从大到小排列,

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix} \quad (6.50)$$

证明. 我们知道, $A^T A$ 和 A 的秩是一样的, 我们考察 $A^T A$ 的特征向量的性质.

- $A^T A$ 是对称矩阵, 特征值大于等于零, 特征向量正交.
- 用这 n 个特征向量得到一组正交标准化基 $\{v_i\}$, $i = 1, 2, \dots, n$.
- 令 $u_i = \frac{1}{\|Av_i\|} Av_i = \frac{1}{\sqrt{\lambda_i}} Av_i$, $i = 1, 2, \dots, r$, 它们互相正交且模长为一, 扩充它们得到一组正交标准化基 $\{u_i\}$.

于是有

$$\begin{aligned} A[v_1, v_2, \dots, v_r, v_{r+1}, v_{r+2}, \dots, v_n] &= [\sqrt{\lambda_1}u_1, \sqrt{\lambda_2}u_2, \dots, \sqrt{\lambda_r}u_r, 0, \dots, 0] \\ &= [u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_n] \begin{bmatrix} \sqrt{\lambda_1} & & & & \\ & \ddots & & & \\ & & \sqrt{\lambda_r} & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix} \end{aligned} \quad (6.51)$$

即,

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T \quad (6.52)$$

□

例 6.5. 计算 $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \\ 0 & 0 \end{bmatrix}$ 的奇异值分解.

解

$$A^T A = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 4 & 5 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} \quad (6.53)$$

求它的特征值可以得到

$$\det(\lambda I - A) = (\lambda - 25)^2 - 400 = 0, \quad (6.54)$$

解得 $\lambda_1 = 45, \lambda_2 = 5$, 奇异值分别为 $\sigma_1 = 3\sqrt{5}, \lambda_2 = 5$, 对应的特征向量分别为

$$v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \quad (6.55)$$

下面计算 U ,

$$\begin{aligned} u_1 &= \frac{Av_1}{\sigma_1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 3 & 0 \\ 4 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} \\ u_2 &= \frac{Av_2}{\sigma_2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 3 & 0 \\ 4 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} \end{aligned} \quad (6.56)$$

于是

$$U = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \\ 0 & 0 \end{bmatrix} \quad (6.57)$$

但是我们需要添加一个向量 $v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ 使 U 的列向量构成正交归一基,

奇异值分解的一个几何理解

$$\begin{cases} AA^T u_i = \sigma_i^2 v_i, & i = 1, 2, \dots, r \\ AA^T u_i = 0, & i = r + 1, \dots, n \end{cases} \quad (6.58)$$

有下列事实

- $u_i, i = 1, 2, \dots, r \in \mathbb{R}^m$ 且在 $C(A)$ 中, 是 $C(A)$ 的一组正交归一基.
- $u_i, i = r + 1, \dots, n$ 是在 $N(A^T)$ 中的一组正交归一基, 是 $C(A)$ 的正交补的一组正交归一基.

同样地,

- $v_i, i = 1, 2, \dots, r$ 是 $C(A^T)$ 的一组正交归一基.
- $v_i, i = r + 1, \dots, n$ 是 $C(A^T)$ 的正交补的一组基.

6.5.5 奇异值分解的应用

$$A = U\Sigma V^T = \sum_{i=1}^r \sigma_i u_i v_i^T \quad (6.59)$$

其中 U, V 为正交矩阵, Σ 为对角矩阵, 对角线上的元素为奇异值 σ_i .

矩阵近似 我们定义矩阵的最大奇异值为其范数, $\|A\| = \max_{x \neq 0} \frac{|Ax|}{|x|} = \sigma_1$ 在所有秩为 k 的矩阵当中, 有

$$\|A - B\| \geq \|A - A_k\|, \quad A_k = \sum_{i=1}^k \sigma_i u_i v_i^T. \quad (6.60)$$

并且有 $\|A - A_k\| = \sigma_{k+1}$.

证明. 对于一个任意的秩为 k 的矩阵, 有分解

$$B = XY^T \quad (6.61)$$

其中, X 是一个 $m \times k$ 矩阵, Y 是一个 $n \times k$ 矩阵.

因为对于矩阵 A , 存在一个可逆矩阵 P , 使得 $A = PR$, R 为行约化阶梯形式, 因为 A 的秩为 k , 所以 R 的后 $m - k$ 行全为零, 所以有

$$R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}. \quad (6.62)$$

把可逆矩阵 P 写为 $P = \begin{bmatrix} P_1 & P_2 \end{bmatrix}$, 这样有

$$A = PR = P_1 R_1. \quad (6.63)$$

这样令 $X = P_1$, $Y^T = R_1$. 这样就得到了分解 $A = XY^T$.

可以找到一个特殊向量,

$$w = \sum_{i=1}^{k+1} y_i v_i, \quad (6.64)$$

使得 $Y^T w = 0$. (不妨令 $y_1^2 + y_2^2 + \cdots + y_{k+1}^2 = 1$) 这样就有

$$\begin{aligned} \|A - B\|^2 &\geq |(A - B)w|^2 = |Aw|^2 = w^T A^T A w \\ &= y_1^2 \sigma_1^2 + \cdots + y_{k+1}^2 \sigma_{k+1}^2 \\ &\geq \sigma_{k+1}^2 \end{aligned} \quad (6.65)$$

□

QS 分解 用奇异值分解可以得到一个 QS 分解

$$\begin{aligned} A &= U \Sigma V^T = UV^T V \Sigma V^T \\ &= (UV^T) (V \Sigma V^T) \\ &\equiv QS \end{aligned} \quad (6.66)$$

其中 S 为一个对称矩阵, 且为一个半正定矩阵, Q 为一个正交矩阵.

验证 S 为半正定矩阵.

$$x^T S x = x^T V \Sigma V^T x = (V^T x)^T \Sigma (V^T x) \geq 0 \quad (6.67)$$

□

广义逆 通过奇异值分解可以定义广义逆

$$A^+ = V \Sigma^+ U^T, \quad (6.68)$$

其中

$$\Sigma^+ = \begin{bmatrix} \sigma_1^{-1} & & & & \\ & \sigma_2^{-1} & & & \\ & & \ddots & & \\ & & & \sigma_r^{-1} & \\ & & & & 0 \end{bmatrix} \quad (6.69)$$

满足

$$AA^+A = A, \quad A^+AA^+ = A^+. \quad (6.70)$$

对于一个不可逆的 A , $Ax = b$ 有多个解, $x = A^+b$ 是最佳解, 且是满足投影方程 $A^T Ax = A^T b$ 的最短的向量.

7 线性映射

线性映射的性质

- 把零向量映射到零向量
- 所有映射到零向量的在 x 中的向量称为核, 记为 $\ker f$
- f 在 Y 中的像称为值域, 记为 $\operatorname{Im} f$

定义 7.1. 线性映射 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m, v \mapsto f(v)$, 满足

$$f(c_1v_1 + c_2v_2 + \cdots + c_nv_n) = c_1f(v_1) + c_2f(v_2) + \cdots + c_nf(v_n). \quad (7.1)$$

根据定义, 线性映射满足

- $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$, 根据定义, $f(v - v) = f(v) - f(v) = 0_{\mathbb{R}^m}$.
- 因为线性映射保持线性性, 且 \mathbb{R}^n 中任何一个向量都可以用一组基的线性组合表示 $v = c_1v_1 + c_2v_2 + \cdots + c_nv_n$, 故

$$f(v) = c_1f(v_1) + c_2f(v_2) + \cdots + c_nf(v_n). \quad (7.2)$$

所以线性映射的可以完全由它对于基的映射决定.

线性映射的矩阵 下面在 \mathbb{R}^m 和 \mathbb{R}^n 中分别取一组基

$$\begin{aligned} (v_1, v_2, \cdots, v_n) &\subset \mathbb{R}^n, \\ (w_1, w_2, \cdots, w_m) &\subset \mathbb{R}^m. \end{aligned} \quad (7.3)$$

线性映射的像和原像都可以用各自空间的基的线性组合表示出来. 将 f 作用于 \mathbb{R}^n 中的基, 并分解在 \mathbb{R}^m 中的基上,

$$f(v_i) = w_1a_{i1} + w_2a_{i2} + \cdots + w_ma_{im}, \quad (7.4)$$

线性映射 f 在基 $(v_1, v_2, \cdots, v_n)(w_1, w_2, \cdots, w_m)$ 中的表示矩阵为

$$(A)_{ij} = a_{ij} \quad (7.5)$$

线性映射和矩阵的一些性质给定一个线性映射 f 和 \mathbb{R}^n 和 \mathbb{R}^m 中的两组基 (v_1, v_2, \dots, v_n) , (w_1, w_2, \dots, w_m) , 这给出了 f 在这组基下的表示矩阵.

考虑 f 中的任意向量 \vec{x} ,

$$\vec{x} = x_1 v_1 + x_2 v_2 + \dots + x_n v_n, \quad (7.6)$$

那么

$$\begin{aligned} f(\vec{x}) &= x_1 f(v_1) + x_2 f(v_2) + \dots + x_n f(v_n) \\ &= \sum_{j=1}^n w_j \sum_{i=1}^m a_{ij} x_i. \end{aligned} \quad (7.7)$$

所以 $f(\vec{x})$ 在 $\{w_i\}$ 上的展开系数为 Ax .

换基公式 下面我们换一组基,

$$v'_i = \sum_{j=1}^n P_{ji} v_j, \quad w'_i = \sum_{j=1}^m Q_{ji} w_j. \quad (7.8)$$

P, Q 为可逆矩阵, 把上述换基的形式写为矩阵乘法

$$V' = VP, \quad W' = WQ. \quad (7.9)$$

在新的基下的表示矩阵为

$$f(v'_i) = f\left(\sum_{j=1}^n v_j P_{ji}\right) = \sum_{j=1}^n f(v_j) P_{ji} = \sum_{j=1}^n \sum_{k=1}^m w_k a_{jk} P_{ji}. \quad (7.10)$$

再利用 $W = W'Q^{-1}$, 得到

$$f(v'_i) = \sum_{j=1}^n \sum_{k=1}^m w_k a_{jk} P_{ji} = \sum_{j=1}^n \sum_{k=1}^m \sum_{l=1}^m w'_l Q_{lk}^{-1} a_{kj} P_{ji} = \sum_{l=1}^m w'_l \sum_{k=1}^m \sum_{j=1}^n Q_{lk}^{-1} a_{kj} P_{ji}. \quad (7.11)$$

写成矩阵乘法的形式

$$A' = Q^{-1}AP. \quad (7.12)$$

命题 7.1. 选一组基后, $\text{Im}(f) = \{Ax, x \text{ 为任意的 } n \times 1 \text{ 矩阵}\}$,

$$\dim \ker(f) + \dim \text{Im}(f) = n. \quad (7.13)$$

证明. 选一组 $\ker(f)$ 中的基 v_1, v_2, \dots, v_s , 添加向量构成 \mathbb{R}^n 的一组基, 考虑集合

$$\text{span}\{f(v_1), f(v_2), \dots, f(v_s), f(v_{s+1}), \dots, f(v_n)\} \quad (7.14)$$

于是

$$\text{Im}(f) = \text{span}\{f(v_1), \dots, f(v_n)\} = \text{span}\{0, f(v_{s+1}), \dots, f(v_n)\}. \quad (7.15)$$

下面证明 $\text{span}\{f(v_{s+1}), \dots, f(v_n)\}$ 中的向量线性无关.

用方程 $x_{s+1}f(v_{s+1}) + \dots + s_nf(v_n) = 0$ 只有零解的性质易证. \square

线性映射和矩阵是 *isomorphic* 的, 映射的复合, 映射的逆, 都对应矩阵的乘法, 矩阵的逆.

可以定义线性变换的行列式, 行列式的值是和基的选取无关的.

对于线性空间 \mathbb{R}^m 和 \mathbb{R}^n , 线性映射 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, 这些线性映射构成一个空间 $M_{m \times n}(\mathbb{R})$, 这个空间是一个线性空间.

证明. 如果 f_1, f_2 为线性映射, 则有

$$(f_1 + f_2)(c_1v_1 + c_2v_2 + \dots + c_nv_n) = c_1(f_1 + f_2)(v_1) + \dots + c_n(f_1 + f_2)(v_n). \quad (7.16)$$

所以映射在加法也是线性映射.

$$(cf)(v) = cf(v) \quad (7.17)$$

在乘法下也是线性映射. \square

7.1 对偶空间

定义 7.2. V 的对偶空间 V' 为 $f: V \rightarrow \mathbb{R}$ 的线性映射构成的空间.

在 V 中取一组基, 对偶空间 V^* 的一组基为

$$(e^{*1}, e^{*2}, \dots, e^{*n}), \quad e^{*i}e_j = \delta_j^i \quad (7.18)$$

其中 δ_i^j 为 Kröneckel 符号.

V^* 中的一个线性函数可以展开

$$f = f_1e^{*1} + f_2e^{*2} + \dots + f_ne^{*n} \quad (7.19)$$

坐标变换时, V 和 V^* 的基分别按照协变和逆变来变化. V 中的基换为

$$e'_i = \sum_{j=1}^3 e_j P^j_i. \quad (7.20)$$

V 中向量的坐标变为

$$x' = P^{-1}x. \quad (7.21)$$

根据定义 V^* 中的对偶基仍然有

$$e'^{*i}(e'_j) = \delta_j^i \quad (7.22)$$

于是

$$e'^{*i} = e^{*i} (P^{-1})^T \quad (7.23)$$

指标全部缩并的式子在坐标变换下不变.

定义 7.3. 两个线性空间同构, 如果两个空间存在可逆线性映射.

命题 7.2. V 和 V^{**} 有一个自然的同构. 其中 V^{**} 中的元素 $v^{**}(f) = f(v)$.

7.2 内积的推广

定义一个对称双线性函数 $g: V \times V \rightarrow \mathbb{R}$, 我们在 (e_1, e_2, \dots, e_n) 这组基上面可以得到一个矩阵

$$g_{ij} \equiv g(e_i, e_j) \quad (7.24)$$

7.3 不变子空间

定义 7.4. 一个线性映射的一个不变子空间 V , 如果对于任何属于 V 的向量 v , 都有 $f(v) \in V$.

例 7.1. f 的一个特征值 λ 的特征向量子空间是一个不变子空间.

例 7.2. $\ker(f)$ 和 $\text{Im}(f)$ 也是 f 的不变子空间.

定义 7.5. 两个线性子空间的交 $V_1 \cap V_2 = \{v \in V_1 \text{ 且 } v \in V_2\}$, 这还是个线性子空间, 在加法和数乘下封闭.

定义 7.6. 给定一个线性空间 V , 以及它的一个线性子空间 U , 定义他的商空间 $W = V/U$, 其中的元素为一个等价类 $[v]$, 等价关系为 $v_1 \sim v_2$, 若 $v_1 = v_2 + u$, $u \in U$.

不难验证, 商空间的等价类在加法和数乘下保持不变. 商空间的维数为 $\dim(V/U) = \dim V - \dim U$.

线性空间的直和分解 (略).

定理 7.1 (维数公式). $\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$

7.4 复线性空间

复线性空间的内积满足

- 线性的.
- $h(w, v) = \overline{h(v, w)}$
- 正定的, $h(v, v) \geq 0$.

定义 7.7. 群是一个集合 G 和一个运算 \circ , 满足三个性质

- 结合律.
- 存在一个单位元.
- 每个元素都存在一个逆元.

群的基本分类: 交换群, 有限群.

例 7.3. 所有可逆矩阵的子集, 所有正交矩阵的集合是一个子群.

定义 7.8. 环是一个集合 R 加上两个运算 $(R, +, *)$, 满足性质

- R 和加法构成一个交换群, 加法单位元记为 0 .
- R 和乘法构成一个含幺半群, 单位元记为 1 .
- 满足左右分配律.

例 7.4. 整数 \mathbb{Z} 和通常的加法和乘法构成整数环.

例 7.5. 多项式环, 每个元素为一个多项式, 系数为实数或复数,

$$P_n(x) = a_n x^n + \cdots + a_0, \quad (7.25)$$

加法运算和乘法运算如通常定义.

定义 7.9. 环之间的同态, 单位元映射到单位元, 保持加法和乘法关系.

定义 7.10. 环的子集构成一个环, 称为子环.

环的子集 I 在加法下封闭, 且 $\forall s \in R, \forall r \in I, sr \in I$, 这样的 I 称为一个理想.

定义 7.11. 域: 一个环满足 $F \setminus \{0\}$ 关于乘法是一个交换群, 则这个环是一个域.

8 一般域上的线性空间

一个线性空间 (\mathcal{F}, V) , 其中 \mathcal{F} 是一个域 $(+, \times, 0_{\mathcal{F}}, 1)$, V 是一个集合 $(+, 0_V)$. 有两个运算, V 上的加法, 数乘 $c \in \mathcal{F}, v \in V \rightarrow cv \in V$.