线代笔记

mny

2023 年 10 月 27 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 $a + b = 0$, $b = -a$
  $\Longrightarrow$ 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 $\begin{bmatrix} a_1, a_2, \cdots, a_m \end{bmatrix}$, 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.

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

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

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

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

$$x_1\vec{v}_1+x_2\vec{v}_2+\cdots+x_n\vec{v}_n \tag{1.3}$$

  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,

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

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, $A=\begin{bmatrix}\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n\end{bmatrix}$
- 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:

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

定义 1.6. 数乘, 任意一个实数 c, 一个矩阵 A, 得到

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

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:

$$(C)_{ij} = a_{i1}b_{1j} + a_{i_2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{m} a_{ik}b_{kj} \tag{1.7}$$

矩阵乘法的性质:

- 结合律:

$$(AB)\,C = A\,(BC) \tag{1.8}$$

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

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

同样

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

□

- 分配律:

$$A\,(B + C) = AB + AC \tag{1.11}$$

$$(A + B)\,C = AC + AB \tag{1.12}$$

- **矩阵乘法不满足交换律**

$$AB \overset{\text{不一定}}{\neq} BA \tag{1.13}$$

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

$$A = \left[\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\right] \tag{1.14}$$

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,

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

- 从矩阵 B 的行向量来看

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

  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.

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

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:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{1.18}$$

  对角线全为 1

- 上三角矩阵:

$$\begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{bmatrix} \tag{1.19}$$

- 下三角矩阵:

$$\begin{bmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{bmatrix} \tag{1.20}$$

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

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

命题 1.1.

$$I_{n\times n}A = AI_{n\times n} = A \tag{1.22}$$

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

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

左侧

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

□

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

$$AB = BA = AC = CA = I_{n\times n} \tag{1.25}$$

所以

$$B = B(AC) = (BA)C = C. \tag{1.26}$$

□

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

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

证明.

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

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

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

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

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

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

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

[1]将在后面证明.

引入两个矩阵

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

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 $\left[\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\right]$ 和向量 b 的关系.

我们把这个方程写成分量的形式

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

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

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

于是方程的增广矩阵变为

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

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.

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

  $S_i^{-1}(c)$ 是可逆的

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

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

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

  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

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

- 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,

$$P_{ij}^{-1} = P_{ij}. \tag{2.5}$$

2.1 初等变换的应用

LU 分解 把方矩阵分解成下列形式

$$A = LU, \tag{2.6}$$

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

$$E_1E_2\cdots E_nA=U \iff A=E_n^{-1}\cdots E_2^{-1}E_1^{-1}U \tag{2.7}$$

由于 $E_1,\cdots,E_n$ 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.

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

做操作

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

用矩阵乘法写起来

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

即

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

计算可得

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

如果 U 的对角线都不为零, 那么

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

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,

$$(E_p \cdots E_1)A = I. \tag{2.14}$$

A 的逆可以这样求解:

$$AB = I \tag{2.15}$$

在两边同时做行变换.

$$(E_p \cdots E_1)AB = (E_p \cdots E_1)I \tag{2.16}$$

即

$$B = IB = (E_p \cdots E_1)I \tag{2.17}$$

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.

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \tag{2.18}$$

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.

$$A \to \begin{bmatrix} 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & & & \\ 0 & 0 & 0 & & \boxed{B} & \\ 0 & 0 & 0 & & & \end{bmatrix} \tag{2.21}$$

- 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:*

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

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

$$Rx = x_1 Y_1 + x_2 Y_2 + \cdots + x_n Y_n, \tag{2.23}$$

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

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

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

$$[R|b'] \tag{2.25}$$

  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

$$x = x_p + x_n \tag{2.26}$$

  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.

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

     特解为

$$x_p = \begin{bmatrix}2\\3\\0\\0\end{bmatrix}. \tag{2.28}$$

  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,

$$x_n = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n. \tag{2.29}$$

例 2.5. 继续求解上例中的线性方程.

第一个通解:

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

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:

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

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

于是原方程的所有解为

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

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

$$\begin{cases}Ax' = b\\ Ax_p = b\end{cases} \tag{3.1}$$

两式相减得到

$$A(x' - x_p) = 0 \tag{3.2}$$

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 $(v_1, v_2, \cdots, v_n)$, 考虑所有的线性组合构成的空间

$$V = \{x_1 v_1 + x_2 v_2 + \cdots + x_n v_n\} \tag{3.3}$$

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

线性无关

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

$$x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = 0 \tag{3.4}$$

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

$$v = x_1e_1 + x_2e_2 + \cdots + x_ne_n. \tag{3.5}$$

基中的向量个数称作维数.

例 3.3. $\mathbb{R}^2$ 中的 $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 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 $(e_1, e_2, \cdots, e_m)$, 根据基的定义, 任意的向量都可以写成

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

  的线性组合, 即

  $$v = a_1e_1 + a_2e_2 + \cdots + a_me_m. \tag{3.6}$$

  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:

  $$\begin{aligned} v &= a_1e_1 + a_2e_2 + \cdots + a_me_m \\ v &= b_1e_1 + b_2e_2 + \cdots + b_me_m \end{aligned} \tag{3.7}$$

  两式相减得到

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

  这与基的线性无关矛盾. □

- 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.

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

把上述过程写成矩阵乘法的形式

$$F = \begin{bmatrix} f_1, f_2, \cdots, f_n \end{bmatrix}, \quad E = \begin{bmatrix} e_1, e_2, \cdots, e_m \end{bmatrix} \tag{3.10}$$

并且

$$F = EA, \tag{3.11}$$

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

$$Fx = EAx = 0 \tag{3.12}$$

也有非零解, 和假设矛盾. □

- 基的变换矩阵 A 为可逆的.

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

$$F = [f_1, f_2, \cdots, f_m], \quad E = [e_1, e_2, \cdots, e_m], \quad F = EA. \tag{3.13}$$

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

$$\begin{bmatrix} f_1, f_2, \cdots, f_m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = 0 \tag{3.14}$$

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,

$$\left(A^{\mathrm{T}}\right)_{ij} = A_{ji}. \tag{3.15}$$

A^{T} 把 A 的行变成列.

转置的一些性质

- $$\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A. \tag{3.16}$$

- $$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}} \tag{3.17}$$

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

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

  根据转置的定义, 有

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

  另一方面,

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

  □

- $$\left(A^{\mathrm{T}}\right)^{-1} = \left(A^{-1}\right)^{\mathrm{T}}. \tag{3.21}$$

  证明. 因为

  $$AA^{-1} = I, \tag{3.22}$$

  两边取转置得到

  $$\left(A^{-1}\right)^{\mathrm{T}}A^{\mathrm{T}} = I \implies \left(A^{\mathrm{T}}\right)^{-1} = \left(A^{-1}\right)^{\mathrm{T}}. \tag{3.23}$$

  □

特殊矩阵

- 对称矩阵: $A^{\mathrm{T}} = A$.

  $$A_{ij} = A_{ji}. \tag{3.24}$$

  例 3.4.

  $$\begin{bmatrix} 2 & 4 \\ 4 & 3 \end{bmatrix}. \tag{3.25}$$

例 3.5.

$$\begin{bmatrix} 2 & 3 & 1 \\ 3 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}. \tag{3.26}$$

- 反对称矩阵: $A^{\mathrm{T}} = -A$. 可知, 其对角线都为零.

例 3.6.

$$\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}. \tag{3.27}$$

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

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

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

2. A 的行向量张成的线性子空间 $C(A^{\mathrm{T}})$, 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 $N(A^{\mathrm{T}})$. 线性方程组 $A^{\mathrm{T}}x = 0$ 的所有解.

3.3 线性代数基本定理

定理 3.1.

$$r_1 = r_2 = r = r'. \tag{3.28}$$

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

设矩阵 $A = \begin{bmatrix} \vec{w}_1 \\ \vec{w}_2 \\ \vdots \\ \vec{w}_m \end{bmatrix}$. 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

$$w' = (w_1, w_2, \cdots, w_i + aw_j, \cdots, w_m). \tag{3.29}$$

$V(w')$ 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

$$w' = x_1 w_1' + \cdots + x_m w_m' = x_1 w_1 + \cdots + (x_j + ax_i) w_j + \cdots + x_m w_m. \tag{3.30}$$

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子空间的影响

注意到, 初等行变换不改变齐次线性方程组的解, 也就是说

$$Ax = 0 \iff Bx = 0. \tag{3.31}$$

其中 B 为 A 初等行变换后的矩阵. 这也就是说,

$$x_1 v_1 + \cdots x_n v_n = 0 \iff x_1 v_1' + \cdots + x_n v_n' = 0. \tag{3.32}$$

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 的行向量子空间及列向量子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) 为列向量子空间 $C(A)$ 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,

$$A = \left[a_1 v_i, \cdots, v_i, \cdots, a_n v_i\right]. \tag{3.33}$$

(其中 v_1 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).

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

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

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

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,

$$\vec{v} \cdot \vec{w} = v^{\mathrm{T}} w = w^{\mathrm{T}} v. \tag{3.36}$$

有了内积, 可以定义一些东西

- 向量 v 的长度
$$|v| = \sqrt{v \cdot v} = \sqrt{v^{\mathrm{T}} v}. \tag{3.37}$$
- 两个向量垂直 $v \perp w$, 如果
$$v \cdot w = 0. \tag{3.38}$$

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

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

无解时,

$$\operatorname{rank}(A') = \operatorname{rank}(A) + 1. \tag{3.40}$$

考虑 $Ax=0$ 齐次线性方程组的解,

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

这意味着 $Ax=0$ 的解垂直于 A 的行向量子空间

$$N(A)\perp C(A^{\mathrm{T}}). \tag{3.42}$$

例 3.9. 对于两个矩阵 A, B, 有

$$\operatorname{rank}(A+B)\le\operatorname{rank}A+\operatorname{rank}B. \tag{3.43}$$

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

$$A+B=\left[v_1+w_1,v_2+w_2,\cdots,v_n+w_n\right]. \tag{3.44}$$

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,

$$\operatorname{rank}(A\oplus B)\le\operatorname{rank}A+\operatorname{rank}B. \tag{3.45}$$

□

例 3.10. 一个矩阵 A, $A^2=A$ 当且仅当

$$\operatorname{rank}A+\operatorname{rank}(I-A)=n. \tag{3.46}$$

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

$$\operatorname{rank}A+\operatorname{rank}(I-A)\le n. \tag{3.47}$$

另一方面, 因为

$$\operatorname{rank}I\ge\operatorname{rank}(I-A)+\operatorname{rank}A, \tag{3.48}$$

所以 $\operatorname{rank}A+\operatorname{rank}(I-A)=n$. □

4 正交投影

如果 $Ax = b$ 无解, $b \notin C(A)$. 在这种情况, 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 $\vec{p}$,*

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

于是

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

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

4.1 投影矩阵

考虑一般情况, 有一组向量 $(v_1, v_2, \cdots, v_n)$, 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

$$A^{\mathrm{T}} (b - p) = 0. \tag{4.3}$$

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

$$p = \hat{x}_1 v_1 + \cdots + \hat{x}_n v_n = \hat{x} A. \tag{4.4}$$

带入上面的垂直条件,

$$A^{\mathrm{T}} b - A^{\mathrm{T}} A \hat{x} = 0 \implies A^{\mathrm{T}} A \hat{x} = A^{\mathrm{T}} b. \tag{4.5}$$

如果 $A^{\mathrm{T}} A$ 可逆, 那么我们有

$$\hat{x} = \left(A^{\mathrm{T}} A\right)^{-1} A^{\mathrm{T}} b. \tag{4.6}$$

于是我们得到

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

4.1.1 投影矩阵的性质

- $P^2 = P$

$$P^2 = A\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}A\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}} = A\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}} = P. \tag{4.8}$$

- $P^{\mathrm{T}} = P$

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

例 4.2. 求投影矩阵

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{4.10}$$

计算投影矩阵 $P = A\left(AA^{\mathrm{T}}\right)^{-1}A^{\mathrm{T}}$,

$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \tag{4.11}$$

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

于是可得

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

下面我们考虑什么时候 $A^{\mathrm{T}}A$ 可逆的问题.

命题 4.1. $A^{\mathrm{T}}A$ 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

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

反之, 如果 $A^{\mathrm{T}}Ax=0$, 两边左乘 x^{T}, 得到

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

也就是说, Ax 的模长为零, 那它必然是零向量. □

根据上述命题, 我们可以发现: 假设 $A^{\mathrm{T}}A$ 可逆 $\iff$ $A^{\mathrm{T}}Ax=0$ 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

$$\hat{x}=\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}b. \tag{4.17}$$

4.2.2 线性回归

收集到一些数据, $(y_1,t_1),(y_2,t_2),\cdots,(y_m,t_m)$. 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

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

其中

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

我们做一些计算

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

$$A^{\mathrm{T}}b=\begin{bmatrix}\sum_{i=1}^m y_i\\ \sum_{i=1}^m y_it_i\end{bmatrix}, \tag{4.21}$$

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

最后我们可以得到

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

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 $(v_1, v_2, \cdots, v_m)$ 的线性组合.

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

$$q_i\cdot q_j=\delta_{ij}. \tag{4.24}$$

则它们是正交归一基.

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

$$Q^{\mathrm{T}}Q=QQ^{\mathrm{T}}=I. \tag{4.25}$$

验证:

$$Q^{\mathrm{T}}Q=\begin{bmatrix} q_1^{\mathrm{T}} \\ q_2^{\mathrm{T}} \\ \vdots \\ q_n^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} q_1^{\mathrm{T}}, q_2^{\mathrm{T}}, \cdots, q_n^{\mathrm{T}} \end{bmatrix}=\begin{bmatrix} \ddots & & \\ & q_j^{\mathrm{T}}q_i & \\ & & \ddots \end{bmatrix}=\begin{bmatrix} \ddots & & \\ & \delta_{ij} & \\ & & \ddots \end{bmatrix}=I. \tag{4.26}$$

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,

$$B=b-A\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}b. \tag{4.27}$$

3. 通过 c 构造一个向量 C,

$$C=c-A\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}c-B\left(B^{\mathrm{T}}B\right)^{-1}B^{\mathrm{T}}c, \tag{4.28}$$

下面验证它垂直于 A 和 B: 令 $Q=[A\ B]$, 则

$$Q^{\mathrm{T}}Q=\begin{bmatrix} A^{\mathrm{T}} \\ B^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} A & B \end{bmatrix}=\begin{bmatrix} A^{\mathrm{T}}A & 0 \\ 0 & B^{\mathrm{T}}B \end{bmatrix} \tag{4.29}$$

所以

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

得到

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

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,

$$D=d-A\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}d-B\left(B^{\mathrm{T}}B\right)^{-1}B^{\mathrm{T}}d-C\left(C^{\mathrm{T}}C\right)^{-1}C^{\mathrm{T}}d. \tag{4.32}$$

$$\vdots$$

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,

$$A = QR \tag{4.33}$$

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

正交归一化当中,

$$A = a \tag{4.34}$$

$$B = b - \frac{A^{\mathrm{T}} b}{A^{\mathrm{T}} A} A \tag{4.35}$$

$$C = c - \frac{A^{\mathrm{T}} c}{A^{\mathrm{T}} A} A - \frac{B^{\mathrm{T}} c}{B^{\mathrm{T}} B} B \tag{4.36}$$

整个过程可以用矩阵乘法表示. 正交归一基为

$$Q = \left[\frac{A}{|A|}, \frac{B}{|B|}, \cdots \right] \tag{4.37}$$

有

$$Q = AR \tag{4.38}$$

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 $r \colon M_{n\times n} \to \mathbb{Z}_+$, 它在初等行变换下不变.

我们引入行列式

$$\delta \colon M_{n\times n} \to \mathbb{R} \tag{5.1}$$

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

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

- 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

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

□

- 换行变换: $\delta(A') = -\delta(A)$.

证明.

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

□

- 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

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

于是

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

□

- 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,

$$\delta(EA) = \delta(A). \tag{5.7}$$

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

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

如果 A' 不是单位矩阵, 那么 A' 的最后一行为零, 则

$$\delta(A) = 0. \tag{5.9}$$

□

行列式满足的一个重要性质:

$$\delta(AB)=\delta(A)\delta(B). \tag{5.10}$$

证明. 找到变换使得 A 变为行约化阶梯形式 A'

$$A'=(E_p\cdots E_1)\,A, \tag{5.11}$$

如果 A' 为单位矩阵, 那么 $\delta(A')=1$

$$\delta(B)=\delta(A'B)=\delta(E_p\cdots E_1AB)=\delta(E_p)\cdots\delta(E_1)\delta(AB), \tag{5.12}$$

所以

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

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

$$\delta(AB)=0=\delta(A)\delta(B). \tag{5.14}$$

□

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 $A=\begin{bmatrix}2&3&4\\5&6&7\\8&9&10\end{bmatrix}$,

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

行列式的递归定义:

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

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:

- $\delta\,(A)\neq0$, 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.

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

$$\det D = c \det A + c' \det B. \tag{5.17}$$

证明. 根据上面的递归定义, 我们把 D, A, B 的行列式展开为

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

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

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

我们需要证明

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

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

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

那么

$$d_{\mu 1} \det D_{\mu 1} = c a_{\mu 1} \det A_{\mu 1} + c' b_{\mu 1} \det B_{\mu 1}. \tag{5.23}$$

□

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

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

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

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

□

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.

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

- 上三角矩阵的行列式等于对角线上元素的乘积.

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

- 下三角矩阵的行列式等于对角线上元素的乘积.

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

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

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

用列展开, 用 A 的第 j 列展开, 有

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

例 5.3. 计算

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

我们对第 1 列展开,

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

对第 1 行展开

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

对第 2 行展开

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

行列式的置换定义:

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

其中的 p 为一个 n 阶置换 p: $\{1, 2, \cdots, n\} \to \{1, 2, \cdots, n\}$ 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:

1	*2*	*3*
1	*2*	*3*

1	*2*	*3*
2	*1*	*3*

1	*2*	*3*
3	*2*	*1*

1	*2*	*3*
1	*3*	*2*

1	*2*	*3*
2	*3*	*1*

1	*2*	*3*
3	*1*	*2*

例 5.5. 用置换的方法计算三阶行列式.

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

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

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

$$Ax = b \tag{5.37}$$

有唯一解

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

其中 B_i 是把 A 的第 i 列换成 b 得到的矩阵,

$$A = \left[v_1, v_2, \cdots, v_n\right], B_i = \left[v_1, v_2, \cdots, b, \cdots, v_n\right]. \tag{5.39}$$

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

$$A\left[E_1, \cdots, x, \cdots, E_n\right] = \left[v_1, v_2, \cdots, b, \cdots, v_n\right]. \tag{5.40}$$

两边取行列式, 有

$$\det(A)\det\left(\left[E_1, \cdots, x, \cdots, E_n\right]\right) = \det\left(\left[v_1, v_2, \cdots, b, \cdots, v_n\right]\right). \tag{5.41}$$

我们需要计算上式左侧的行列式, 不难发现,

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

所以

$$(\det A)\, x_i = \det(B_i)\,. \tag{5.43}$$

即

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

□

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

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

那么 A 的逆矩阵为

$$A^{-1} = \frac{1}{\det A} M^{\mathrm{T}}. \tag{5.46}$$

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,

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

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

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

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

$$A\left[w_1, w_2, \cdots, w_n\right] = I \implies \left[Aw_1, Aw_2, \cdots, Aw_n\right] = \left[E_1, E_2, \cdots, E_n\right]. \tag{5.50}$$

需要求解线性方程组,

$$Aw_i = E_i. \tag{5.51}$$

由克拉默法则,

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

由于 B_i 是把 A 的第 i 列换成 E_i 得到的矩阵, 所以

$$\det\left(B_j\right) = (-)^{i+j} \det\left(A_{ij}\right) = M_{ij}. \tag{5.53}$$

□

6 特征值和特征向量

考虑下列线性方程组 (特征方程),

$$A\vec{x} = \lambda\vec{x}. \tag{6.1}$$

$\vec{x}$ 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

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

得到

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

所以只有 $\lambda = 0$ 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

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

得到

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

把第二个方程带入第一个方程, 有

$$\left[(1-\lambda)\,(2-\lambda) - 3\right] x_2 = 0. \tag{6.6}$$

所以 λ 有两个解.

考虑一般的特征方程,

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

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

$$\det(A - \lambda I) = 0. \tag{6.8}$$

计算 $|\lambda I - A|$,

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

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,

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

通过观察可以发现,

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

$$c_n = |-A| = (-1)^n \det A. \tag{6.12}$$

特征多项式有唯一分解:

$$p(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n). \tag{6.13}$$

- 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

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

所以,

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

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

$$A^k\vec{x} = \lambda^k\vec{x}. \tag{6.16}$$

- 如果 A 可逆, 则 $\lambda \neq 0$, 因为

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

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: $\vec{x}_1, \vec{x}_2 \in V(\lambda)$

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

  2. 数乘下封闭: $\vec{x} \in V(\lambda)$

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

- 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

$$c_1\lambda_1x + c_2\lambda_2y = 0. \tag{6.20}$$

这时候得到了两个方程, 消去 x,

$$c_2(\lambda_2 - \lambda_1)y = 0. \tag{6.21}$$

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

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

$$A' = EA \tag{6.22}$$

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

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

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

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

□

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

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

□

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,

$$Ax_i = \lambda_i x_i, \quad i = 1, 2, \cdots, n. \tag{6.26}$$

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 $P = \left[x_1, x_2, \cdots, x_n\right]$,

$$A \underbrace{\left[x_1, x_2, \cdots, x_n\right]}_{P} = \underbrace{\left[x_1, x_2, \cdots, x_n\right]}_{P} \underbrace{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}}_{\Lambda}. \tag{6.27}$$

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. □

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 $x_1, x_2, \cdots, x_m$, 且满足 $Ax_i = \lambda x_i$, $i = 1, 2, \cdots, m$.

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

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

$$
\begin{aligned}
&A\underbrace{\left[x_1,x_2,\cdots,x_m,e_1,e_2,\cdots,e_{n-m}\right]}_{P}\\
&=\underbrace{\left[x_1,x_2,\cdots,x_m,e_1,e_2,\cdots,e_{n-m}\right]}_{P}\underbrace{\left[\begin{array}{cccc|cc}
\lambda_1 & 0 & \cdots & 0 & * & * \\
0 & \lambda_2 & \ddots & 0 & * & * \\
0 & 0 & \ddots & 0 & * & * \\
\vdots & \ddots & 0 & \lambda_m & * & * \\
\hline
0 & \ddots & & 0 & * & * \\
0 & 0 & \ddots & 0 & * & * \\
0 & 0 & \cdots & 0 & * & *
\end{array}\right]}_{C}.
\end{aligned}
\tag{6.28}
$$

因为 C 和 A 的特征多项式相同 (由相似变换性质), 所以 C 至少有 m 个特征值 $\implies$ A 至少有 m 个关于 λ 的特征值. □

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 $(v_1,e_1,e_2,\cdots,e_{n-1})$. 令 $P=\left[v_1,e_1,e_2,\cdots,e_{n-1}\right]$, 把特征方程写成矩阵乘法的形式,

$$
A\left[v_1,e_1,e_2,\cdots,e_{n-1}\right]=\left[v_1,e_1,e_2,\cdots,e_{n-1}\right]\left[\begin{array}{c:c}
\lambda_1 & B_1 \\
\hdashline
0 & \\
\vdots & A_1 \\
0 &
\end{array}\right]
\tag{6.29}
$$

之后再对子矩阵 A_1 做相同的操作. □