

# 线性代数 中国科学技术大学 2025 春

主讲: 杨金榜

数院新楼 610

助教: 陈鉴、申长硕、武熙川

2025 年 6 月 19 日

# 课程考核方式

- ① 考核成绩为平时成绩、期中成绩和期末成绩加权平均。具体权重会由线性代数课题组根据期中和期末考试的难易度来确定。
- ② 平时成绩包含作业成绩和上课考勤两部分。
- ③ 周二收发作业。迟交一天到一个星期，得分减半；迟交一个星期以上，零分。（诚信问题零容忍，一旦发现抄袭，将给予零分处理）收发作业务必准时。
- ④ 关于纪律：上课时出入请尽量从后门。如有事不能来上课，可以无理由请假，但必须课前告知助教。

平时表现 =  $5 \times$  考勤到课次数 - 请假次数 = 总评微调力度。

# 什么是线性代数 (linear algebra)?

线性代数是关于[向量空间](#)和[线性映射](#)的一个数学分支。

——[维基百科](#)

线性代数的方法广泛的用在[数学其他分支](#)、[物理化学](#)、[计算机科学](#)、[经济学](#)等学科中。例如

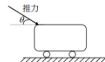
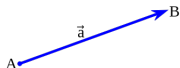
- ① 泛函分析 (研究函数组成的空间);
- ② 量子力学 (波函数, 密度泛函理论);
- ③ 科学计算 (天气预报);
- ④ 机器学习 (运动学正解);
- ⑤ 数据传输 (编码理论);
- ⑥ ...



# 第一章 向量

# 什么是向量

我们初中学习过一些物理量包括速度、位移、力等等.



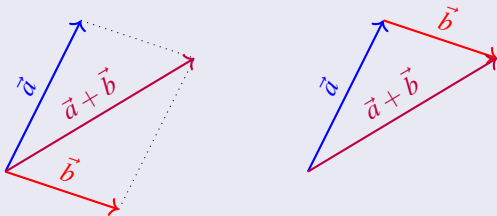
数学上的抽象总结:

向量 = 既有大小, 又有方向 的量.

# 向量加法

速度, 力的合成  $\xrightarrow{\text{用数学语言抽象化}}$  向量的加法

定义 (向量的加法—平行四边形法则或者三角形法则)



# 向量加法的基本性质

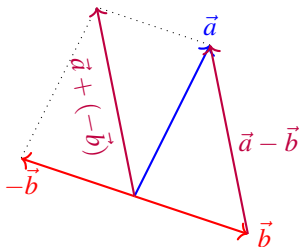
## 性质 (向量加法的基本性质)

- ① 加法交换律:  $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ ;
- ② 加法结合律:  $\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$ ;
- ③ 存在零元:  $\vec{a} + \vec{0} = \vec{a} = \vec{0} + \vec{a}$ ;
- ④ 存在负元:  $\vec{a} + (-\vec{a}) = \vec{0} = (-\vec{a}) + \vec{a}$ ;

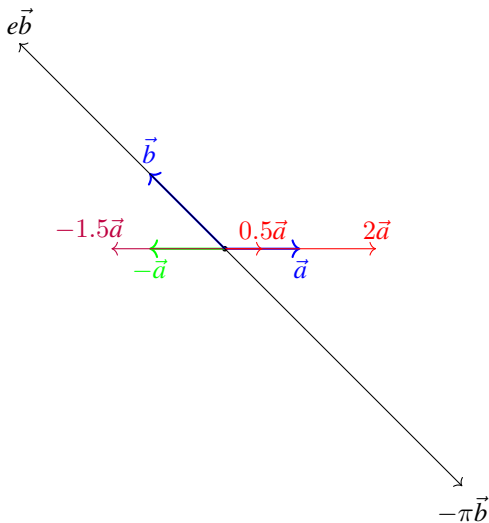


## 定义 (向量的减法)

$$\vec{a} - \vec{b} := \vec{a} + (-\vec{b}).$$



# 向量的数乘



## 定义 (向量的数乘)

令  $\vec{a}$  为一向量,  $\lambda$  为一实数.

- 若  $\lambda \geq 0$ , 则  $\lambda\vec{a}$  定义为长度为  $\lambda|\vec{a}|$  且方向与  $\vec{a}$  相同的向量.
- 若  $\lambda < 0$ , 则  $\lambda\vec{a}$  定义为长度为  $-\lambda|\vec{a}|$  且方向与  $\vec{a}$  相反的向量.

记号: 若  $\vec{a} \neq \vec{0}$ , 则记  $\vec{a}^0 := \frac{1}{|\vec{a}|}\vec{a}$ . 即,  $\vec{a}^0$  为方向与  $\vec{a}$  相同的单位向量.

注: 零向量:  $|\vec{a}| = 0$ . 规定任意方向都为零向量的方向.

## 性质 (向量数乘的基本性质)

- ⑤ 数乘单位元:  $1\vec{a} = \vec{a}$ ;
- ⑥ 数乘结合律:  $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$ ;
- ⑦ 左分配律:  $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$ ;
- ⑧ 右分配律:  $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$ ;

## 定义 (线性运算)

向量的**加法**和**数乘**运算统称为向量的**线性运算**.

## 性质 (向量线性运算的八条基本性质)

- ① 加法交换律:  $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ ;
- ② 加法结合律:  $\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$ ;
- ③ 存在零元:  $\vec{a} + \vec{0} = \vec{a} = \vec{0} + \vec{a}$ ;
- ④ 存在负元:  $\vec{a} + (-\vec{a}) = \vec{0} = (-\vec{a}) + \vec{a}$ ;
- ⑤ 数乘单位元:  $1\vec{a} = \vec{a}$ ;
- ⑥ 数乘结合律:  $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$ ;
- ⑦ 左分配律:  $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$ ;
- ⑧ 右分配律:  $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$ ;

注: 我们将通过这八条性质来公理化地定义一般的**线性空间**或**向量空间** (书本第五章, 课件第六章).

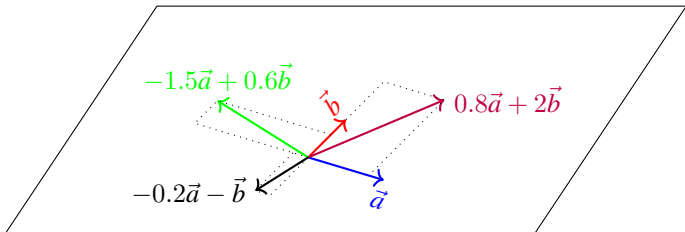
## 定义 (线性组合)

设  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  为一组向量,  $\lambda_1, \lambda_2, \dots, \lambda_m$  为一组实数. 称向量

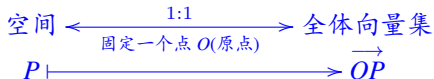
$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_m \vec{a}_m$$

为向量  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  的线性组合.

也就是说, 一组向量的线性组合就是从这组向量出发通过线性运算能够获得的向量.



# 线性组合的几何解释



- ① 设  $\vec{a}$  为一个非零向量。则

过原点与  $\vec{a}$  平行的直线  $\xleftrightarrow{1:1}$   $\vec{a}$  的全体线性组合。

- ② 设  $\vec{a}, \vec{b}$  不平行 (不共线)。则

过原点与  $\vec{a}$  和  $\vec{b}$  平行的平面  $\xleftrightarrow{1:1}$   $\vec{a}, \vec{b}$  的全体线性组合。

# 线性相关与线性无关

## 定义 (线性相关, 线性无关)

给定一组向量  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ .

- 如果存在一组不全为零的实数  $\lambda_1, \lambda_2, \dots, \lambda_m$  使得

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_m \vec{a}_m = \vec{0},$$

则称向量组  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性相关.

- 反之, 若对任意一组不全为零的实数  $\lambda_1, \lambda_2, \dots, \lambda_m$  都有

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_m \vec{a}_m \neq \vec{0},$$

则称向量组  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  的线性无关.

特别地, 设  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性无关. 若  $\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_m \vec{a}_m = \vec{0}$ , 则  $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$ .

## 例

证明向量组  $\vec{e}_1 = \vec{a} + \vec{b} + \vec{c}, \vec{e}_2 = \vec{a} - \vec{b} - \vec{c}, \vec{e}_3 = \vec{a} + 2\vec{b} + 2\vec{c}$  线性相关.

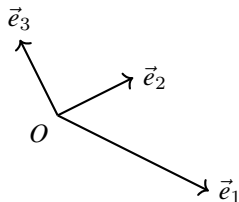
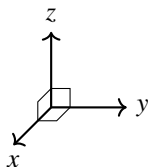
$$3\vec{e}_1 - \vec{e}_2 - 2\vec{e}_3 = \vec{0} \text{ 或 } \vec{e}_2 = 3\vec{e}_1 - 2\vec{e}_3.$$



# 线性相关的几何解释

- ① 一个向量  $\vec{a}$  线性相关  $\iff \vec{a} = \vec{0}$ ;
- ② 两个向量  $\vec{a}, \vec{b}$  线性相关  $\iff \vec{a}$  与  $\vec{b}$  平行 (共线);
- ③ 三个向量  $\vec{a}, \vec{b}, \vec{c}$  线性相关  $\iff \vec{a}, \vec{b}, \vec{c}$  共面;
- ④ 四个及四个以上的向量一定线性相关.

为了推广笛卡尔坐标系到坐标轴不相互垂直的情形,



我们需要引入向量的基本定理.

## 定理 (向量的基本定理)

设  $\vec{e}_1, \vec{e}_2, \vec{e}_3$  为空间中的三个不共面的向量, 则对每个向量  $\vec{a}$  都**存在唯一**的三元有序实数组  $(x_1, x_2, x_3)$ , 使得

$$\vec{a} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3.$$

## 定义 (基、坐标)

称不共面的三个向量  $\vec{e}_1, \vec{e}_2, \vec{e}_3$  为一组**基**. 若

$$\vec{a} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + x_3 \vec{e}_3,$$

则称  $(x_1, x_2, x_3)$  为向量  $\vec{a}$  在基  $\vec{e}_1, \vec{e}_2, \vec{e}_3$  下的**(仿射) 坐标**.

$$\text{仿射坐标系} = \text{点 } O + \text{基 } \vec{e}_1, \vec{e}_2, \vec{e}_3 \quad \text{记作 } [O; \vec{e}_1, \vec{e}_2, \vec{e}_3].$$

## 推论 (一一对应)

若给定仿射坐标系  $[O; \vec{e}_1, \vec{e}_2, \vec{e}_3]$ , 则有如下一一对应

$$\begin{array}{ccc} \text{空间} & \xleftrightarrow{1:1} & \text{全体向量集} \xleftrightarrow{1:1} \mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \\ P & \xrightarrow{\quad} & \overrightarrow{OP} \xrightarrow{\quad} \text{坐标}(x_1, x_2, x_3) \end{array}$$

给定仿射坐标系  $[O; \vec{e}_1, \vec{e}_2, \vec{e}_3]$ , 我们用  $(x_1, x_2, x_3)$  表示向量  $x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$ . 则我们有

## 性质

- $(x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$ ;
- $\lambda(x_1, x_2, x_3) = (\lambda x_1, \lambda x_2, \lambda x_3)$ .

# 坐标变换

给定两个仿射坐标系  $[O; \vec{e}_1, \vec{e}_2, \vec{e}_3]$  和  $[O'; \vec{e}'_1, \vec{e}'_2, \vec{e}'_3]$ .

## 问题

设空间中的点  $P$  在两个坐标系下的坐标分别为  $(x, y, z)$  和  $(x', y', z')$ . 求两个坐标之间的关系式?

为了回答这一问题, 我们需要给出两个坐标系之间的位置关系:

- 设  $O$  在  $[O'; \vec{e}'_1, \vec{e}'_2, \vec{e}'_3]$  下的坐标为  $(x'_0, y'_0, z'_0)$ . 即

$$\overrightarrow{O'O} = x_0 \vec{e}'_1 + y_0 \vec{e}'_2 + z_0 \vec{e}'_3.$$

- 设  $\vec{e}_j$  在基  $\vec{e}'_1, \vec{e}'_2, \vec{e}'_3$  下的坐标为  $(a_{1j}, a_{2j}, a_{3j})$ . 即

$$\vec{e}_j = a_{1j} \vec{e}'_1 + a_{2j} \vec{e}'_2 + a_{3j} \vec{e}'_3.$$

然后  $P$  在两个坐标系下的坐标之间的关系式可写为<sup>1</sup>:

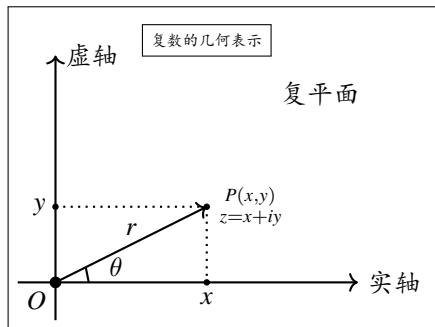
$$\begin{aligned} x' &= a_{11}x + a_{12}y + a_{13}z + x'_0 \\ y' &= a_{21}x + a_{22}y + a_{23}z + y'_0 \\ z' &= a_{31}x + a_{32}y + a_{33}z + z'_0 \end{aligned}$$

$$pf: \overrightarrow{O'P} = \overrightarrow{O'O} + \overrightarrow{OP}.$$

<sup>1</sup>后续, 通过矩阵乘法可简化为  $X' = AX + X'_0$ .

复数的代数表示

$$z = \underset{\substack{\uparrow \\ \text{实部} \\ \text{Re}z}}{\underline{x}} + \underset{\substack{\uparrow \\ \text{虚部} \\ \text{Im}z}}{\sqrt{-1} \underline{y}}$$



- **模长:**  
 $|z| := |\overrightarrow{OP}| = \sqrt{x^2 + y^2};$
- **辐角:** 实轴沿逆时针方向旋转到  $\overrightarrow{OP}$  的角度.
- **共轭:**  $\bar{z} := x - iy.$
- **三角表示:**  
 $z = r(\cos \theta + i \sin \theta).$

设  $\omega = a + ib = t(\cos \varphi + i \sin \varphi)$ 。任取复平面中的点  $z = x + iy = r(\cos \theta + i \sin \theta)$ ，则根据复数的乘法和积化和差公式

$$\omega \cdot z = (ax - by) + i(bx + ay) = rt(\cos(\theta + \varphi) + i \sin(\theta + \varphi)).$$

- 若用坐标表示复数，则乘复数  $\omega$  给出如下复平面的变换：

$$(x, y) \mapsto (ax - by, bx + ay).$$

- 若用极坐标表示复数，则乘复数  $\omega$  有如下几何解释：

$$z \xrightarrow{\text{伸缩 } t \text{ 倍}} tz \xrightarrow{\text{逆时针旋转 } \varphi \text{ 角度}} \omega z.$$

复数域的任意子集称为**数集**. 例如:  $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \dots$ .

## 定义

若数集  $\mathbb{F}$  至少包含两个元素, 且关于数的加减乘除封闭, 那么称  $\mathbb{F}$  为**数域**.

## 例 (本课程中最常用的域)

有理数集  $\mathbb{Q}$ , 实数集  $\mathbb{R}$ , 和复数集  $\mathbb{C}$  均为数域.

注:  $\mathbb{N}, \mathbb{Z}$  不为数域.

## 例

数集  $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$  为一个数域.

## 性质

有理数域  $\mathbb{Q}$  为最小的数域. 即, 若  $\mathbb{F}$  为数域, 则  $\mathbb{Q} \subseteq \mathbb{F}$ .



# 高维数组向量的定义

设  $\mathbb{F}$  为数域,  $n$  为正整数.

## 定义

我们称一个由  $n$  个  $\mathbb{F}$  上的数  $a_1, a_2, \dots, a_n$  组成的有序数组

$$\vec{a} := (a_1, \dots, a_n)$$

为数域  $\mathbb{F}$  上的一个  $n$  维数组向量. 其中  $a_i$  称为  $\vec{a}$  的第  $i$  个分量. 数域  $\mathbb{F}$  上  $n$  维数组向量全体记为  $\mathbb{F}^n$ .

记法:

$$\text{行向量: } \vec{a} = (a_1, \dots, a_n) \in \mathbb{F}^n \quad \text{列向量: } \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{F}^n$$

# 高维数组向量的线性运算

设  $\vec{a} = (a_1, \dots, a_n) \in \mathbb{F}^n$ ,  $\vec{b} = (b_1, \dots, b_n) \in \mathbb{F}^n$ ,  $\lambda \in \mathbb{F}$ .

● 加法:  $\vec{a} + \vec{b} := (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$ .

● 数乘:  $\lambda \vec{a} := (\lambda a_1, \lambda a_2, \dots, \lambda a_n)$ .

● 零向量:  $\vec{0} := (0, \dots, 0)$ .

● 负向量:  $-\vec{a} := (-a_1, \dots, -a_n)$ .

● 相等:  $\vec{a} = \vec{b} \iff a_i = b_i \quad i = 1, \dots, n$ .

## 八条基本性质:

① 加法交换律:  $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ ;

② 加法结合律:  $\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$ ;

③ 存在零元:  $\vec{a} + \vec{0} = \vec{a} = \vec{0} + \vec{a}$ ;

④ 存在负元:  $\vec{a} + (-\vec{a}) = \vec{0} = (-\vec{a}) + \vec{a}$ ;

⑤ 数乘单位元:  $1\vec{a} = \vec{a}$ ;

⑥ 数乘结合律:  $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$ ;

⑦ 左分配律:  $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$ ;

⑧ 右分配律:  $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$ ;

# 线性相关 (高维数组向量)

设  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n \in \mathbb{F}^n$ ,  $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{F}$ .

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_m \vec{a}_m \quad \longleftarrow \text{线性组合}$$

定义 (线性相关, 线性无关)

一组向量  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  称为**线性相关**, 若存在一组不全为零的实数  $\lambda_1, \lambda_2, \dots, \lambda_m$  使得

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_m \vec{a}_m = 0.$$

反之, 则称向量  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  的**线性无关**.

### 例

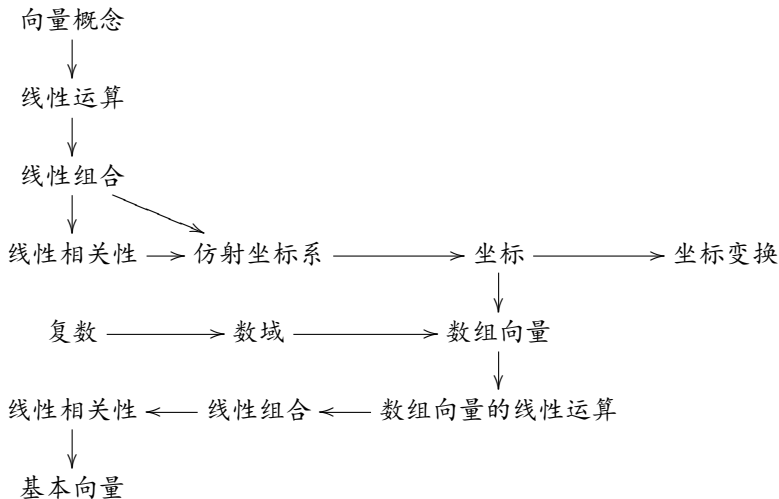
记  $\vec{e}_1 = (1, 0, \dots, 0)$ ,  $\vec{e}_2 = (0, 1, \dots, 0)$ ,  $\dots$ ,  $\vec{e}_n = (0, 0, \dots, 1)$ . 则  $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$  线性无关. 称这  $n$  个向量为基本向量.

### 事实

基本向量线性无关, 且任意  $n$  维数组向量都可以表示为基本向量的线性组合.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + \dots + x_n \vec{e}_n.$$

# 本章主要内容





## 第二章 线性方程组

## ① 小学 (鸡兔同笼)

今有雉兔同笼，  
上有三十五头，  
下有九十四足，  
问雉兔各几何？



## ② 初中 (二元一次方程组)

$$\begin{cases} x + y = 35 \\ 2x + 4y = 94 \end{cases}$$

## ③ 现在 (线性方程组: 即, 任意个变量及任意个方程构成的一次方程组.)



# 线性方程组

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

- ①  $x_j$ : 第  $j$  个变量
- ②  $a_{ij}$ : 第  $i$  个方程第  $j$  个变量  $x_j$  的系数;
- ③  $b_i$ : 第  $i$  个方程的常数项.
- ④ 如果  $b_1 = \cdots = b_m = 0$ , 则称方程组(\*)为齐次线性方程组。否则称之为非齐次线性方程组。

例

$$\begin{cases} x+y=0 \\ 2x+4y=0 \end{cases} \text{ (齐次)} \quad \begin{cases} x+y=35 \\ 2x+4y=94 \end{cases} \text{ (非齐次)}$$

- ① 令  $c_1, \dots, c_n$  为一组数。若将  $x_1 = c_1, \dots, x_n = c_n$  代入方程组(\*), 等式都成立, 则称  $(c_1, \dots, c_n)$  为一组解。称所有解组成的集合为解集。
- ② 如果解集非空, 我们称方程组(\*)是相容的, 否则称之为不相容的。

例

线性方程组  $\begin{cases} x + y = 35 \\ 2x + 4y = 94 \end{cases}$  是相容, 因为二元数组

$(23, 12)$  为它的一组解。通过初等计算验证, 可以得出这一方程组的解集正好为  $\{(23, 12)\}$ 。

# 关于线性方程组的几个基本问题

- ① 解是否存在，是否唯一？
- ② 如何求解？
- ③ 解集的公式表达。
- ④ 解集的几何结构。

注：

- 前两个问题可以用Gauss消元法解决(本次课的主要内容)。
- 后两个问题需要用矩阵、行列式和线性空间等基本工具与研究对象(书本第四五章的主要内容)。

## 例 (鸡兔同笼)

$$\text{求解 } \begin{cases} x + y = 35 & (1) \\ 2x + 4y = 94 & (2) \end{cases} .$$

解：第一步, 方程 (2) 减去两倍的方程 (1) 得

$$\xrightarrow{-2(1) \rightarrow (2)} \begin{cases} x + y = 35 & (3) \\ 2y = 24 & (4) \end{cases}$$

第二步, 方程 (4) 除以 2

$$\xrightarrow{\frac{1}{2}(4)} \begin{cases} x + y = 35 & (5) \\ y = 12 & (6) \end{cases}$$

第三步. 方程 (5) 减去方程 (6)

$$\xrightarrow{-(6) \rightarrow (5)} \begin{cases} x = 23 \\ y = 12 \end{cases}$$

例

$$\text{求解} \begin{cases} 3x_1 + 2x_2 - x_3 = 6 & (1) \\ x_1 + 3x_2 + 2x_3 = 9 & (2) \\ 2x_1 - x_2 - 3x_3 = 3 & (3) \end{cases}.$$

解：交换前两个方程的顺序

$$\xrightarrow{(1) \leftrightarrow (2)} \begin{cases} x_1 + 3x_2 + 2x_3 = 9 & (4) \\ \textcolor{red}{3}x_1 + 2x_2 - x_3 = 6 & (5) \\ \textcolor{red}{2}x_1 - x_2 - 3x_3 = 3 & (6) \end{cases}$$

$$\begin{array}{l} \xrightarrow{-3(4) \rightarrow (5)} \\ \xrightarrow{-2(4) \rightarrow (6)} \end{array} \begin{cases} x_1 + 3x_2 + 2x_3 = 9 & (7) \\ -7x_2 - 7x_3 = -21 & (8) \\ \textcolor{red}{-7}x_2 - x_3 = -15 & (9) \end{cases}$$

$$\xrightarrow{-(8) \rightarrow (9)} \begin{cases} x_1 + 3x_2 + 2x_3 = 9 & (10) \\ -7x_2 - 7x_3 = -21 & (11) \\ \textcolor{red}{6}x_3 = 6 & (12) \end{cases}$$

$$\xrightarrow{\frac{1}{6}(12)} \begin{cases} x_1 + 3x_2 + 2x_3 = 9 & (13) \\ -7x_2 - 7x_3 = -21 & (14) \\ x_3 = 1 & (15) \end{cases}$$

$$\xrightarrow{7(15) \rightarrow (14)} \begin{cases} x_1 + 3x_2 + 2x_3 = 9 & (16) \\ -7x_2 = -14 & (17) \\ x_3 = 1 & (18) \end{cases}$$

$$\xrightarrow{-\frac{1}{7}(14)} \begin{cases} x_1 + 3x_2 + 2x_3 = 9 & (19) \\ x_2 = 2 & (20) \\ x_3 = 1 & (21) \end{cases}$$

$$\xrightarrow{\begin{matrix} -3(20) \rightarrow (19) \\ -2(21) \rightarrow (19) \end{matrix}} \begin{cases} x_1 = 1 \\ x_2 = 2 \\ x_3 = 1 \end{cases} \quad (\text{代回验证, 成立}) \Rightarrow \text{解集为 } \{(1, 2, 1)\}.$$

# 方程的三种初等变换

	初等变换	记号
(1)	交换两个方程	$(i) \leftrightarrow (j)$
(2)	某个方程乘以一个非零常数	$\lambda(i)$
(3)	某个方程乘一个常数加到另一个方程上	$\lambda(i) \rightarrow (j)$

# 例 (存在不唯一)

$$\text{求解} \begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = -3 & (1) \\ x_1 + 2x_2 \quad \quad - 5x_4 = 1 & (2) \\ 2x_1 + 4x_2 - 3x_3 - 19x_4 = 6 & (3) \\ 3x_1 + 6x_2 - 3x_3 - 24x_4 = 7 & (4) \end{cases}$$

解:

$$\begin{array}{l} -1(1) \rightarrow (2) \\ -2(1) \rightarrow (3) \\ -3(1) \rightarrow (4) \end{array} \rightarrow \begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 1 & (5) \\ \quad \quad -3x_3 - 9x_4 = 4 & (6) \\ \quad \quad -9x_3 - 27x_4 = 12 & (7) \\ \quad \quad -12x_3 - 36x_4 = 16 & (8) \end{cases}$$

$$\begin{array}{l} (6) \rightarrow (5) \\ -3(6) \rightarrow (7) \\ -4(6) \rightarrow (8) \end{array} \rightarrow \begin{cases} x_1 + 2x_2 \quad \quad - 5x_4 = -3 & (9) \\ \quad \quad -3x_3 - 9x_4 = 4 & (10) \\ \quad \quad \quad \quad 0 = 0 & (11) \\ \quad \quad \quad \quad 0 = 0 & (12) \end{cases}$$

$$-\frac{1}{3}(10) \rightarrow \begin{cases} x_1 + 2x_2 \quad \quad - 5x_4 = 1 & (13) \\ \quad \quad x_3 + 3x_4 = -\frac{4}{3} & (14) \\ \quad \quad \quad 0 = 0 & (15) \\ \quad \quad \quad 0 = 0 & (16) \end{cases} \Rightarrow \begin{cases} x_1 = 1 - 2x_2 + 5x_4 \\ x_3 = -\frac{4}{3} - 3x_4 \end{cases}$$

因此  $x_2$  和  $x_4$  可取任意值,  $x_1$  和  $x_3$  由  $x_2$  和  $x_4$  的取值唯一确定。即,

$$\begin{cases} x_1 = 1 - 2t_1 + 5t_2 \\ x_2 = t_1 \\ x_3 = -\frac{4}{3} - 3t_2 \\ x_4 = t_2 \end{cases} \quad \text{其中 } t_1, t_2 \in \mathbb{F}.$$



### 例 (解存在且唯一)

$$\text{求解} \begin{cases} x + y = 35 & (1) \\ 2x + 4y = 94 & (2) \end{cases}.$$

$$\Rightarrow \begin{cases} x = 23 \\ y = 12 \end{cases}$$

### 例 (解存在但不唯一)

$$\text{求解} \begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = -3 & (1) \\ x_1 + 2x_2 - 5x_4 = 1 & (2) \\ 2x_1 + 4x_2 - 3x_3 - 19x_4 = 6 & (3) \\ 3x_1 + 6x_2 + 3x_3 - 24x_4 = 7 & (4) \end{cases}$$

$$\Rightarrow \begin{cases} x_1 = 1 - 2t_1 + 5t_2 \\ x_2 = t_1 \\ x_3 = -\frac{4}{3} - 3t_2 \\ x_4 = t_2 \end{cases} \quad \text{其中 } t_1, t_2 \in \mathbb{F}.$$

### 例 (无解)

$$\text{求解} \begin{cases} x_1 & -x_2 & +x_3 & = & 1 & (1) \\ x_1 & -x_2 & -x_3 & = & 3 & (2) \\ 2x_1 & -2x_2 & -x_3 & = & 3 & (3) \end{cases}$$

解:

$$\begin{array}{l} \xrightarrow{-1(1) \rightarrow (2)} \\ \xrightarrow{-2(1) \rightarrow (3)} \end{array} \begin{cases} x_1 & -x_2 & +x_3 & = & 1 & (4) \\ & & -2x_3 & = & 2 & (5) \\ & & -3x_3 & = & 1 & (6) \end{cases}$$

$$\begin{array}{l} \xrightarrow{\frac{1}{2}(5) \rightarrow (4)} \\ \xrightarrow{-\frac{3}{2}(5) \rightarrow (6)} \end{array} \begin{cases} x_1 & -x_2 & & = & 2 \\ & & -2x_3 & = & 2 \\ & & 0 & = & -2 \end{cases}$$

$\Rightarrow$  无解!!

我们称有相同的解集的两个线性方程组互为对方的同解方程组.

### 定理

三种初等变换将线性方程组变为同解线性方程组.

### 证明.

这里只给出第三种初等变换的证明, 其他两种情形证明类似。

$$\left\{ \begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \quad \ddots \quad \vdots \quad \vdots \\ a_{i1}x_1 + \cdots + a_{in}x_n = b_i \\ \vdots \quad \ddots \quad \vdots \quad \vdots \\ a_{j1}x_1 + \cdots + a_{jn}x_n = b_j \\ \vdots \quad \ddots \quad \vdots \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{array} \right. \xLeftrightarrow[\begin{array}{l} \lambda(i) \rightarrow (j) \\ -\lambda(i) \rightarrow (j) \end{array}] \left\{ \begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \quad \ddots \quad \vdots \quad \vdots \\ a_{i1}x_1 + \cdots + a_{in}x_n = b_i \\ \vdots \quad \ddots \quad \vdots \quad \vdots \\ (a_{j1} + \lambda a_{i1})x_1 + \cdots + (a_{jn} + \lambda a_{in})x_n = b_j + \lambda b_i \\ \vdots \quad \ddots \quad \vdots \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{array} \right.$$

□

# 矩阵

## 定义

我们称由若干行及若干列的数组成的阵列为**矩阵**.

例如: 线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad (*)$$

的系数和常数项组成一个矩阵

$$\left( \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right) \quad \leftarrow \text{方程组 } (*) \text{ 的增广矩阵}$$

注：由于变元不参与运算，因此可以用增广矩阵来表示方程组。

# Gauss 消元法的矩阵表示

例 (非齐次线性方程组)

$$\text{求解} \begin{cases} 3x_1 + 2x_2 - x_3 = 6 \\ x_1 + 3x_2 + 2x_3 = 9 \\ 2x_1 - x_2 - 3x_3 = 3 \end{cases}.$$

$$\text{解: } \begin{pmatrix} 3 & 2 & -1 & 6 \\ 1 & 3 & 2 & 9 \\ 2 & -1 & 3 & 3 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 3 & 2 & 9 \\ 3 & 2 & -1 & 6 \\ 2 & -1 & 3 & 3 \end{pmatrix}$$

$$\xrightarrow{\substack{-3r_1 \rightarrow r_2 \\ -2r_1 \rightarrow r_3}} \begin{pmatrix} 1 & 3 & 2 & 9 \\ 0 & -7 & -7 & -21 \\ 0 & -7 & -1 & -15 \end{pmatrix} \xrightarrow{-r_2 \rightarrow r_3} \begin{pmatrix} 1 & 3 & 2 & 9 \\ 0 & -7 & -7 & -21 \\ 0 & 0 & 6 & 6 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{6}r_3} \begin{pmatrix} 1 & 3 & 2 & 9 \\ 0 & -7 & -7 & -21 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\substack{7r_3 \rightarrow r_2 \\ -2r_3 \rightarrow r_1}} \begin{pmatrix} 1 & 3 & 0 & 7 \\ 0 & -7 & 0 & -14 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

$$\xrightarrow{\substack{-\frac{1}{7}r_2 \\ \frac{3}{7}r_2 \rightarrow r_1}} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix} \Rightarrow \begin{cases} x_1 = 1 \\ x_2 = 2 \\ x_3 = 1 \end{cases}$$

# Gauss 消元法的矩阵表示

例 (齐次线性方程组)

$$\text{求解} \begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 = 0 \\ \quad \quad x_2 + 2x_3 + 2x_4 + 6x_5 = 0 \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 = 0 \end{cases}.$$

$$\text{解: } \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 1 & -3 \\ 0 & 1 & 2 & 2 & 6 \\ 5 & 4 & 3 & 3 & -1 \end{pmatrix} \xrightarrow[-5r_1 \rightarrow r_3]{-3r_1 \rightarrow r_2} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -2 & -6 \\ 0 & 1 & 2 & 2 & 6 \\ 0 & -1 & -2 & -2 & -6 \end{pmatrix}$$

$$\xrightarrow[-r_2 \rightarrow r_4]{\begin{matrix} r_2 \rightarrow r_1 \\ r_2 \rightarrow r_3 \end{matrix}} \begin{pmatrix} 1 & 0 & -1 & -1 & -5 \\ 0 & 1 & 2 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \Rightarrow \begin{cases} x_1 = t_1 + t_2 + 5t_3 \\ x_2 = -2t_1 - 2t_2 - 6t_3 \\ x_3 = t_1 \\ x_4 = t_2 \\ x_5 = t_3 \end{cases}$$

其中  $t_1, t_2, t_3$  为参变量.

# Gauss 消元法的矩阵表示

例 (无解)

$$\text{求解} \begin{cases} 3x_1 - 2x_2 + 5x_3 + 4x_4 = 2 \\ 6x_1 - 7x_2 + 4x_3 + 3x_4 = 3 \\ 9x_1 - 9x_2 + 9x_3 + 7x_4 = -1 \end{cases}.$$

$$\text{解: } \begin{pmatrix} 3 & -2 & 5 & 4 & 2 \\ 6 & -7 & 4 & 3 & 3 \\ 9 & -9 & 9 & 7 & -1 \end{pmatrix} \xrightarrow[-3r_1 \rightarrow r_3]{-2r_1 \rightarrow r_2} \begin{pmatrix} 3 & -2 & 5 & 4 & 2 \\ 0 & -3 & -6 & -5 & -1 \\ 0 & -3 & -6 & -5 & -7 \end{pmatrix}$$

$$\xrightarrow{-r_2 \rightarrow r_3} \begin{pmatrix} 3 & -2 & 5 & 4 & 2 \\ 0 & -3 & -6 & -5 & -1 \\ 0 & 0 & 0 & 0 & -6 \end{pmatrix} \Rightarrow \text{原方程组无解!!}$$

例

求解下列方程组

$$\begin{cases} 2x_1 + 4x_2 - 6x_3 + x_4 = 2 \\ x_1 - x_2 + 4x_3 + x_4 = 1 \\ -x_1 + x_2 - x_3 + x_4 = 0 \end{cases}$$

$$\begin{cases} x_1 + x_2 - 3x_4 - x_5 = 0 \\ x_1 - x_2 - 2x_3 - x_4 = 0 \\ 4x_1 - 2x_2 + 6x_3 + 3x_4 - 4x_5 = 0 \\ 2x_1 + 4x_2 - 2x_3 + 4x_4 - 7x_5 = 0 \end{cases}$$



# 一般线性方程组的 Gauss 消元法 (算法)

按照  $a_{11}, a_{21}, \dots, a_{m1}, a_{12}, \dots, a_{m2}, \dots, a_{mn}, b_1, \dots, b_m$  的顺序找到第一个非零元. 不妨设其为  $a_{i_1 j_1}$ . 即, 矩阵前  $j_1 - 1$  列全为零, 第  $j_1$  列的前  $i_1 - 1$  个元素全为零且  $a_{i_1 j_1} \neq 0$ .

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix} = \begin{pmatrix} 0 \cdots 0 & 0 & a_{1j_1+1} & \cdots & a_{1n} & b_1 \\ 0 \cdots 0 & 0 & a_{2j_1+1} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 \cdots 0 & 0 & a_{i_1-1j_1+1} & \cdots & a_{i_1-1n} & b_{i_1-1} \\ 0 \cdots 0 & a_{i_1 j_1} & a_{i_1 j_1+1} & \cdots & a_{i_1 n} & b_{i_1} \\ 0 \cdots 0 & a_{i_1+1 j_1} & a_{i_1+1 j_1+1} & \cdots & a_{i_1+1 n} & b_{i_1+1} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 \cdots 0 & a_{mj_1} & a_{mj_1+1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

- ① 将第  $i_1 + 1$  至  $m$  行减去第  $i_1$  行的合适倍数, 使得第  $j_1$  列除第  $i_1$  行外全为零;
- ② 将矩阵的第 1 行与第  $i_1$  行交换. 从而我们得到如下形式的矩阵

$$\left( \begin{array}{ccc|cccc} 0 & \cdots & a_{1,j_1}^{(1)} & a_{1,j_1+1}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ \hline & & & a_{2,j_1+1}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ & & & a_{3,j_1+1}^{(1)} & \cdots & a_{3n}^{(1)} & b_3^{(1)} \\ & & & \vdots & \ddots & \vdots & \vdots \\ & & & a_{m,j_1+1}^{(1)} & \cdots & a_{mn}^{(1)} & b_m^{(1)} \end{array} \right)$$

③ 对 2 中得到的矩阵的右下角矩阵重复 1 和 2 得

$$\left( \begin{array}{cccc|cccc} 0 & \cdots & a_{1,j_1}^{(1)} & \cdots & a_{1,j_2}^{(1)} & a_{1,j_2+1}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ & & & & a_{2,j_2}^{(2)} & a_{2,j_2+1}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \hline & & & & & a_{3,j_1+1}^{(2)} & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\ & & & & & \vdots & & \vdots & \vdots \\ & & & & & a_{m,j_1+1}^{(2)} & \cdots & a_{mn}^{(2)} & b_m^{(2)} \end{array} \right)$$

④ 继续重复上述步骤最终得到矩阵 (最简形式、标准形式) 如下 ( $r \leq n$ )

$$\left( \begin{array}{cccccccccccc} 0 & \cdots & c_{1,j_1} & \cdots & c_{1,j_2} & \cdots & \cdots & \cdots & c_{1,j_r} & \cdots & c_{1n} & d_1 \\ & & & & c_{2,j_2} & \cdots & \cdots & \cdots & c_{2,j_r} & \cdots & c_{2n} & d_2 \\ & & & & & & & & \vdots & & \vdots & \vdots \\ & & & & & & & & c_{r,j_r} & \cdots & c_{rn} & d_r \\ & & & & & & & & & & & d_{r+1} \\ & & & & & & & & & & & \vdots \\ & & & & & & & & & & & 0 \end{array} \right)$$

### 定理 (解集大小判定)

方程组(\*)的解有如下性质:

- ① 若  $d_{r+1} \neq 0$ , 则无解;
- ② 若  $d_{r+1} = 0$  且  $r = n$ , 则解唯一;
- ③ 若  $d_{r+1} = 0$  且  $r < n$ , 则有多解.

显然  $(0, \dots, 0)$  为齐次线性方程的一组解, 我们称之为零解或平凡解. 齐次线性方程组的非零解称为非平凡解.

### 推论

- ① 齐次线性方程组有非平凡解当且仅当  $r < n$ . 特别地, 若齐次方程个数小于变量个数, 即  $m < n$ , 则其一定有非平凡解.

注: 数量  $r$  代表独立方程的个数, 它决定了解集的“大小”.

# 证明

• 若  $d_{r+1} \neq 0$ , 则最后一个非零行代表  $0 = d_{r+1}$ 。显然矛盾。因此原方程无解。

• 若  $d_{r+1} = 0$  且  $r = n$ , 则

$$\text{标准形式} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1,n} & d_1 \\ & c_{22} & \cdots & c_{2,n} & d_2 \\ & & \ddots & \vdots & \vdots \\ & & & c_{n,n} & d_n \\ & & & 0 & 0 \end{pmatrix} \xrightarrow{\text{初等变换}} \begin{pmatrix} c_{11} & & & & \tilde{d}_1 \\ & c_{22} & & & \tilde{d}_2 \\ & & \ddots & & \vdots \\ & & & c_{n,n} & \tilde{d}_n \\ & & & 0 & 0 \end{pmatrix}$$

$\Rightarrow \left( \frac{\tilde{d}_1}{c_{11}}, \dots, \frac{\tilde{d}_n}{c_{nn}} \right)$  为原方程的唯一解。

• 若  $d_{r+1} = 0$  且  $r < n$ , 则标准形式可通过初等变换化为

$$\begin{pmatrix} 0 \cdots 0 & 1 & e_{1,j_1+1} \cdots e_{1,j_2-1} & 0 & e_{1,j_2+1} \cdots e_{1,j_r-1} & 0 & e_{1,j_r+1} \cdots e_{1n} & f_1 \\ & & & 1 & e_{2,j_2+1} \cdots e_{2,j_r-1} & 0 & e_{2,j_r+1} \cdots e_{2n} & f_2 \\ & & & & \ddots & \vdots & \vdots & \vdots \\ & & & & & 1 & e_{r,j_r+1} \cdots e_{rn} & f_r \end{pmatrix}$$

因此, 我们有

$$\begin{cases} x_{j_1} &= f_1 - e_{1,j_1+1}x_{j_1+1} - \cdots - e_{1,j_2-1}x_{j_2-1} - e_{1,j_2+1}x_{j_2+1} - \cdots - e_{1,n}x_n \\ x_{j_2} &= f_2 - e_{2,j_2+1}x_{j_2+2} - \cdots - e_{2,j_3-1}x_{j_3-1} - e_{2,j_3+1}x_{j_3+1} - \cdots - e_{2,n}x_n \\ &\vdots \\ x_{j_r} &= f_r - e_{r,j_r+1}x_{j_r+1} - \cdots - e_{r,n}x_n \end{cases}$$

显然, 其中  $n-r$  变元  $x_1, \cdots, x_{j_1-1}, x_{j_1+1}, \cdots, x_{j_r-1}, x_{j_r+1}, \cdots, x_n$  可取任意值, 且任意给定的取值可以唯一确定剩下的变元  $x_{j_1}, \cdots, x_{j_r}$  的取值。

从而我们可以将方程组(\*)的解集表示为

$$\begin{cases} x_1 &= \alpha_{11}t_1 + \cdots + \alpha_{1,n-r}t_{n-r} + \beta_1, \\ x_2 &= \alpha_{21}t_1 + \cdots + \alpha_{2,n-r}t_{n-r} + \beta_2, \\ \cdots &\quad \cdots \quad \quad \quad \cdots \quad \quad \cdots \\ x_n &= \alpha_{n1}t_1 + \cdots + \alpha_{n,n-r}t_{n-r} + \beta_n, \end{cases} \quad (**)$$

其中  $t_1, \cdots, t_{n-r} \in \mathbb{F}$ . 我们称(\*\*)为方程组(\*)的**通解**. 当参数  $t_1, \cdots, t_{n-r}$  取一组值带入(\*\*)得到的解成为一个**特解**.

# 例

$$\text{求解} \begin{cases} 2x_1 - x_2 + x_3 + x_4 = 1 \\ x_1 + 2x_2 - x_3 + 4x_4 = 2 \\ x_1 + 7x_2 - 4x_3 + 11x_4 = \lambda. \end{cases}$$

$$\text{解: } \begin{pmatrix} 2 & -1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 4 & 2 \\ 1 & 7 & -4 & 11 & \lambda \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 2 & -1 & 4 & 2 \\ 2 & -1 & 1 & 1 & 1 \\ 1 & 7 & -4 & 11 & \lambda \end{pmatrix}$$

$$\xrightarrow[\begin{smallmatrix} -2r_1 \rightarrow r_2 \\ -r_1 \rightarrow r_3 \end{smallmatrix}]{\begin{smallmatrix} \frac{2}{5}r_2 \rightarrow r_1 \\ -\frac{1}{5}r_2 \end{smallmatrix}} \begin{pmatrix} 1 & 2 & -1 & 4 & 2 \\ 0 & -5 & 3 & -7 & -3 \\ 0 & 5 & -3 & 7 & \lambda - 2 \end{pmatrix} \xrightarrow[r_2 \rightarrow r_3]{\begin{smallmatrix} \frac{1}{5}r_2 \rightarrow r_1 \\ \frac{6}{5}r_2 \rightarrow r_3 \end{smallmatrix}} \begin{pmatrix} 1 & 0 & \frac{1}{5} & \frac{6}{5} & \frac{4}{5} \\ 0 & 1 & -\frac{3}{5} & \frac{7}{5} & \frac{3}{5} \\ 0 & 0 & 0 & 0 & \lambda - 5 \end{pmatrix}.$$

因此，原方程有解当且仅当  $\lambda = 5$ . 当  $\lambda = 5$  时，原方程通解为

$$\begin{cases} x_1 = -\frac{1}{5}t_1 - \frac{6}{5}t_2 + \frac{4}{5} \\ x_2 = \frac{3}{5}t_1 - \frac{7}{5}t_2 + \frac{3}{5} \\ x_3 = t_1 \\ x_4 = t_2 \end{cases} \quad t_1, t_2 \in \mathbb{F}.$$

例

当  $a$  为何值时, 下列线性方程组有解? 有解时求出它的通解.

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 2 \\ x_1 - x_2 - 2x_3 = -3 \\ ax_1 - 2x_2 + 2x_3 = 6 \end{cases}$$



# 高斯消元法的计算效率 \*

解  $n$  元线性方程组需约  $\frac{2}{3}n^3$  次运算, 算术复杂度  $O(n^3)$

- 除法:  $n(n+1)/2$  次
- 乘法:  $\frac{2n^3+3n^2-5n}{6}$  次
- 减法:  $\frac{2n^3+3n^2-5n}{6}$  次

系数类型影响

- 浮点数: 算术复杂度主导 (算术运算时间相近)
- 整数/有理数: 中间值指数增长  $\rightarrow$  空间复杂度  $O(n^5)$  (Bareiss 算法)

应用范围

- 适用: 数千未知数的系统
- 瓶颈: 百万级系统耗时剧增 (如 i9-12900K 大约需 50 天)
- 解决方案: 大规模系统采用迭代法, 或者特殊矩阵采用特殊办法

# 数值不稳定性问题 \*

根源：主元趋零导致误差放大

消元时需除以极小主元  $\rightarrow$  计算误差呈指数级增长

稳定性条件

- 天然稳定：对角占优矩阵、正定矩阵
- 条件稳定：一般矩阵需配合部分选主元
- 例外情况：存在特殊稳定矩阵仍可能失稳

工程应对

- 强制选主元策略（列/全主元）
- 条件数预检测： $\kappa(A) > 10^{10}$  时预警（病态矩阵）
- 混合方法：结合迭代修正（如 Gauss-Jordan 迭代精炼）

感兴趣的同学，自行查阅相关文献.

# 2024 春 (本次课程内容)

## 问题 (期中, 填空 (5'))

对于矩阵  $A = \begin{pmatrix} 1 & -1 & 2 & -1 \\ 2 & 3 & 4 & 1 \\ 1 & -6 & 2 & a \end{pmatrix}$ , 若存在列向量  $b \in \mathbb{R}^3$  使得线性方程组  $Ax = b$  无解, 则  $a =$  \_\_\_\_\_.

## 问题 (期中, 大题 (12'))

当  $a$  为何值时, 如下的线性方程组有解? 当有解时, 求出它的所有解.

$$\begin{cases} ax_1 + x_2 + x_3 + x_4 = 1 \\ x_1 + ax_2 + x_3 + x_4 = a \\ x_1 + x_2 + ax_3 + x_4 = a^2 \\ x_1 + x_2 + x_3 + ax_4 = a^3 \end{cases}$$

问题 (期中, 填空 (5'))

方程  $X \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$  的解为  $X =$  \_\_\_\_\_.

问题 (期中, 判断题 (5'))

设  $A$  为  $n \times n$  方阵,  $b$  为  $n$  维列向量. 若线性方程组  $Ax = b$  只有唯一解, 则线性方程组  $A^T x = b$  也只有唯一解.

## 问题 (期中, 大题 (10'))

给定方程组

$$\begin{cases} x + 2y + z = 1 \\ x + y - 3z = 1 \\ x + 5y + \lambda z = \mu. \end{cases}$$

求当  $\lambda, \mu$  取何值时, 方程组无解、有多解、有唯一解; 当有唯一解时, 求出该解。

解:  $\left( \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & -3 & 1 & 1 \\ 1 & 5 & \lambda & 1 & \mu \end{array} \right) \xrightarrow{\substack{-r_1 \rightarrow r_2 \\ -r_1 \rightarrow r_3}} \left( \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -4 & 0 & 0 \\ 0 & 3 & \lambda-1 & 0 & \mu-1 \end{array} \right)$

$\xrightarrow{3r_2 - r_3} \left( \begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -4 & 0 & 0 \\ 0 & 0 & \lambda-13 & 0 & \mu-4 \end{array} \right) \dots \dots \dots 4'$

故  $\lambda=13, \mu \neq 4$  时无解;  $\dots \dots \dots 5'$

$\lambda=13, \mu=4$  时有无穷解;  $\dots \dots \dots 6'$

$\lambda \neq 13$  时有唯一解:  $\left\{ \begin{array}{l} x = \frac{7\mu + \lambda - 20}{\lambda - 13} \\ y = \frac{-4\mu + 4}{\lambda - 13} \\ z = \frac{\mu - 1}{\lambda - 13} \end{array} \right. \dots \dots \dots 10'$

## 问题 (期中, 判断题 (5'))

设  $A, B$  为  $n$  阶方阵, 若方程组  $AX = 0$  与  $BX = 0$  同解, 则方程组  $\begin{pmatrix} A & B \\ 0 & B \end{pmatrix} Y = 0$  与  $\begin{pmatrix} B & A \\ 0 & A \end{pmatrix} Y = 0$  同解.

## 第三章 矩阵 (一)

## 定义 (矩阵)

一个  $m \times n$  矩阵为由  $m \times n$  个数排成的  $m$  行  $n$  列的阵列

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

记作  $A = (a_{ij})_{m \times n}$ . 称  $a_{ij}$  为矩阵  $A$  的第  $(i, j)$  元素; 当  $i = j$  时,  $a_{ii}$  也称为  $A$  的对角元.

若两个矩阵  $A$  和  $B$  的行数和列数相同且对应位置的元素都相同, 则称  $A$  与  $B$  相等, 记为  $A = B$ . 否则称  $A$  与  $B$  不相等, 记为  $A \neq B$ . 例如:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$



# 特殊矩阵命名

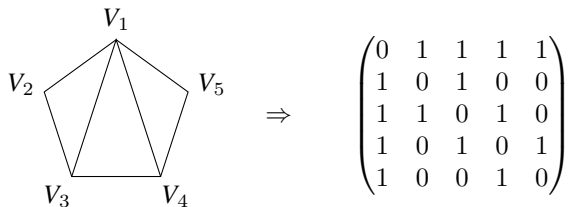
按照矩阵行列数大小, 非零元分布或元素取值范围等, 我们有如下特殊矩阵命名:

- ①  $n$  维行向量  $:= 1 \times n$  矩阵;
- ②  $n$  维列向量  $:= n \times 1$  矩阵;
- ③ 零矩阵  $O$ , 或  $0$ ;
- ④  $n$  阶方阵;
- ⑤ 单位阵  $I$ ;
- ⑥ 数量矩阵  $aI$ ;
- ⑦ 对角矩阵  $\text{diag}(a_1, \dots, a_n)$ ;
- ⑧ 上(下)三角矩阵;
- ⑨ (反)对称矩阵;
- ⑩ 整数矩阵; 有理数矩阵; 实矩阵; 复矩阵;
- ⑪ 数域  $\mathbb{F}$  上的矩阵;
- ⑫ 多项式矩阵.

## 例

- $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  为整系数对称 2 阶方阵.
- $\begin{pmatrix} 0 & \sqrt{-1} & -1 \\ -\sqrt{-1} & 0 & \pi \\ 1 & -\pi & 0 \end{pmatrix}$  为复系数反对称 3 阶方阵.
- $\begin{pmatrix} e^{\frac{i\pi}{3}} & 0 \\ 0 & e^{\frac{i\pi}{3}} \end{pmatrix}$  为复系数 2 阶数量矩阵.

- 图的邻接矩阵（更一般地，可加入箭向和长度）：



- 数字图像 3 个  $m \times n$  取值在  $0, 1, \dots, 255$  的矩阵.

# 矩阵的线性运算

## 定义 (加法与数乘)

设  $\lambda$  为数域  $\mathbb{F}$  中的数. 取两个  $\mathbb{F}$  系数的  $m \times n$  矩阵  $A = (a_{ij})_{m \times n} \in \mathbb{F}^{m \times n}$ ,  $B = (b_{ij})_{m \times n} \in \mathbb{F}^{m \times n}$ . 定义

### ● 加法

$$A + B := (a_{ij} + b_{ij})_{m \times n} := \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix},$$

### ● 数乘

$$\lambda A := (\lambda a_{ij})_{m \times n} := \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

类似地定义负矩阵和矩阵的减法运算:

$$-A := (-a_{ij})_{m \times n}, \quad A - B := A + (-B) = (a_{ij} - b_{ij})_{m \times n}.$$

注: 只有行数和列数相同的矩阵才可以相加减. 否则无意义.

## 例

①  $A = \begin{pmatrix} 1 & 3 & 5 \\ -1 & 2 & 4 \end{pmatrix}, B = \begin{pmatrix} 2 & -1 & 3 \\ 7 & 3 & 1 \end{pmatrix}$ , 求  $A + B$  和  $A - B$ .

②  $A = \begin{pmatrix} 1 & 3 & 5 \\ -1 & 2 & 4 \end{pmatrix}$ , 求  $2A$ .

③ 设  $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, B = \begin{pmatrix} 3 & 5 \\ 7 & 6 \end{pmatrix}, C = \begin{pmatrix} 0 & -1 \\ 9 & 3 \end{pmatrix}$ .

- 求  $3A + 2B - 4C$ .
- 若  $5A + 3X = B$ . 求  $X$ .

类似于向量的线性运算, 矩阵的线性运算也满足如下八条基本性质:

## 定理

- ① 加法交换律:  $A + B = B + A$ ;
- ② 加法结合律:  $A + (B + C) = (A + B) + C$ ;
- ③ 有零矩阵:  $A + 0 = A = 0 + A$ ;
- ④ 有负矩阵:  $A + (-A) = 0 = (-A) + A$ ;
- ⑤ 数乘单位元:  $1A = A$ ;
- ⑥ 数乘结合律:  $\lambda(\mu A) = (\lambda\mu)A$ ;
- ⑦ 左分配律:  $(\lambda + \mu)A = \lambda A + \mu A$ ;
- ⑧ 右分配律:  $\lambda(A + B) = \lambda A + \lambda B$ ;

其中  $A, B, C$  为使得运算有有意的矩阵,  $\lambda, \mu$  为数.

$$E_{ij} := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \leftarrow \begin{matrix} \text{第 } i \text{ 行} \\ \uparrow \\ \text{第 } j \text{ 列} \end{matrix}$$

## 引理

任意  $m \times n$  矩阵  $A = (a_{ij})_{m \times n}$  都可以唯一的表示为基本矩阵的线性组合

$$A = \sum_{i=1}^m \sum_{j=1}^n a_{ij} E_{ij}.$$

# 数组向量空间之间的线性映射

## 定义 (线性映射)

给定一个数组向量空间之间的映射  $\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^m$ . 若  $\mathcal{A}$  满足

$$\begin{aligned}(\text{保持加法}) \quad & \mathcal{A}(\vec{a} + \vec{b}) = \mathcal{A}(\vec{a}) + \mathcal{A}(\vec{b}); \\(\text{保持数乘}) \quad & \mathcal{A}(\lambda \vec{a}) = \lambda \mathcal{A}(\vec{a})\end{aligned} \tag{*}$$

则  $\mathcal{A}$  称为线性映射.

线性映射保持原点并将直线映为直线的映射。(选做习题)

实际上, 在其他课程中有许多算子也具备这种性质。例如, 积分和求导运算均保持加法和数乘性质:

$$\begin{aligned}\int_0^x (f(t) + g(t))dt &= \int_0^x f(t)dt + \int_0^x g(t)dt, & \int_0^x \lambda f(t)dt &= \lambda \int_0^x f(t)dt; \\(f + g)' &= f' + g', & (\lambda f)' &= \lambda f'.\end{aligned}$$

为了统一研究这类对象, 有必要定义一般的线性空间及其上的线性映射。这是线性代数课程中最核心的两个概念。量子力学中, 可观测量在数学上使用一类特殊的线性映射来描述的。



# 矩阵与线性映射

- 给定线性映射  $\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^m$ . 考虑基本向量  $\vec{e}_j$  在这个映射下的像

$$\mathcal{A}(\vec{e}_j) = \mathcal{A} \left( \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \right) =: \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

因此, 通过线性映射  $\mathcal{A}$ , 可以得到一个矩阵  $A = (a_{ij})_{m \times n}$ .

- 根据  $\mathcal{A}$  为线性映射, 对于任意  $y_1, \dots, y_n \in \mathbb{F}$ , 我们有

$$\mathcal{A} \left( \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right) = \begin{pmatrix} a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ \vdots \\ a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n \end{pmatrix}. \quad (*)$$

**注:** 线性映射  $\mathcal{A}$  由其在基本向量处的值唯一确定。

# 矩阵与线性映射

反之, 给定矩阵  $A$  我们可以通过  $(*)$ , 确定一个线性映射  $\mathcal{A}$ .

$$\mathbb{F}^{m \times n} \xleftrightarrow{1:1} \text{从 } \mathbb{F}^n \text{ 到 } \mathbb{F}^m \text{ 的全体线性映射.}$$

**注：** 一个从  $\mathbb{F}^m$  到  $\mathbb{F}^n$  的映射是线性的, 当且仅当其每个分量都为一次齐次函数.

**思考：** 前面定义的各种矩阵在几何中对应哪些线性映射? 例如: 零矩阵、 $n$  阶方阵、单位阵、数量矩阵、对角矩阵.

类似于函数的加法和数乘; 线性映射也可考虑加法和数乘.

**思考：** 用矩阵的语言, 线性映射的加法和数乘是如何描述的?

## 引理

线性映射的合成仍然是线性映射.

**思考：** 用矩阵的语言, 线性映射的合成是如何描述的?

# 线性映射的合成与矩阵的乘法

设矩阵  $A = (a_{ik})_{m \times n} \in \mathbb{F}^{m \times n}$ ,  $B = (b_{kj})_{n \times p} \in \mathbb{F}^{n \times p}$  对应线性映射  $\mathcal{A}$  和  $\mathcal{B}$ . 即,

$$\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^m, \quad (y_1, \dots, y_n)^T \mapsto \left( \sum_{k=1}^n a_{1k} y_k, \sum_{k=1}^n a_{2k} y_k, \dots, \sum_{k=1}^n a_{mk} y_k \right)^T.$$

$$\mathcal{B}: \mathbb{F}^p \rightarrow \mathbb{F}^n, \quad (z_1, \dots, z_p)^T \mapsto \left( \sum_{j=1}^p b_{1j} z_j, \sum_{j=1}^p b_{2j} z_j, \dots, \sum_{j=1}^p b_{nj} z_j \right)^T.$$

因此

$$\mathcal{A} \circ \mathcal{B}(z_1, \dots, z_p)^T = \left( \sum_{j=1}^p \left( \sum_{k=1}^n a_{1k} b_{kj} \right) z_j, \dots, \sum_{j=1}^p \left( \sum_{k=1}^n a_{mk} b_{kj} \right) z_j \right)^T.$$

记  $C = (c_{ij})_{m \times p} \in \mathbb{F}^{m \times p}$ , 其中  $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ . 则  $\mathcal{A} \circ \mathcal{B}$  为矩阵  $C$  对应的线性变换.

**注:** 下面我们将这个矩阵  $C$  定义为矩阵  $A$  和  $B$  的乘积.

# 矩阵的乘法

## 定义 (矩阵乘法)

设  $A = (a_{ik})_{m \times n} \in \mathbb{F}^{m \times n}$ ,  $B = (b_{kj})_{n \times p} \in \mathbb{F}^{n \times p}$ . 定义

$$AB := \left( \sum_{k=1}^n a_{ik} b_{kj} \right)_{m \times p} \in \mathbb{F}^{m \times p}.$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$
$$:= \begin{pmatrix} \boxed{a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}} & \cdots & \boxed{a_{11}b_{1p} + a_{12}b_{2p} + \cdots + a_{1n}b_{np}} \\ \vdots & & \vdots \\ \boxed{a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1}} & \cdots & \boxed{a_{m1}b_{1p} + a_{m2}b_{2p} + \cdots + a_{mn}b_{np}} \end{pmatrix}$$

**注:** 并非任意两矩阵都可以相乘. 只有在  $A$  的列数等  $B$  的行数时,  $A$  与  $B$  才可以相乘.

特别地, 若  $A$  为方阵, 我们可以定义  $A$  的方幂. 规定  $A^0 := I$  并定义 ( $k > 0$ ):

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

例

已知  $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ ,  $B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ ,  $C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  以及  $D = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ .  
求  $AB, BA, CD, DC, C^2, D^2$ .

从这个例子, 你发现了什么现象?

例

已知  $A = (a_{ij})_{m \times n}$ ,  $B = \text{diag}(b_1, \cdots, b_m)$ ,  $C = \text{diag}(c_1, \cdots, c_n)$ . 求  $BA$  和  $AC$ .

你发现了什么规律?

# 矩阵乘法运算的基本性质

## 定理

矩阵乘法运算的基本性质:

- ① 乘法结合律:  $(AB)C = A(BC)$ ;
- ② 乘法单位元:  $IA = A = AI$ ;
- ③ 左分配律:  $(A + B)C = AC + BC$ ;
- ④ 右分配律:  $A(B + C) = AB + AC$ ;
- ⑤ 数乘结合律:  $\lambda(AB) = (\lambda A)B = A(\lambda B)$ .

其中  $A, B, C$  是使得运算有意义的矩阵,  $\lambda$  为常数.

证明思路: 计算并比较两边矩阵的相同位置的元素.

第一条的几何证明:  $(\mathcal{A} \circ \mathcal{B}) \circ \mathcal{C} = \mathcal{A} \circ (\mathcal{B} \circ \mathcal{C})$ .

# 矩阵多项式

设  $A$  为  $\mathbb{F}$  上的  $n$  阶方阵,  $f(x) = c_0 + c_1x + \cdots + c_kx^k$  为  $\mathbb{F}$  系数的多项式, 定义矩阵多项式

$$f(A) := c_0I_n + c_1A + \cdots + c_kA^k.$$

例

记  $n$  阶方阵  $J_n = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$ . 计算  $J_n^k$  和  $(aI_n + bJ_n)^k$ .

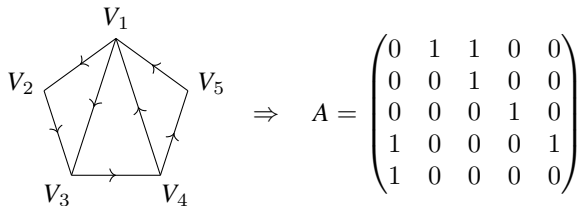
以下关系式是否成立?

- ①  $(A + B)^2 = A^2 + 2AB + B^2$ ?
- ②  $(I + A)^m = \sum_{k=0}^m \binom{m}{k} A^k$ ?  $(A + B)^m = ?$
- ③  $f(A)g(A) = g(A)f(A)$ ?

例

求 Fibonacci 数列  $F_1 = F_2 = 1, F_n = F_{n-1} + F_{n-2}, (n \geq 3)$  的通项公式.

答:  $F_n = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$



$$A^2 = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

$$A^3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

$$A^5 = \begin{pmatrix} 1 & 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 2 & 1 \\ 3 & 1 & 1 & 1 & 2 \\ 2 & 1 & 1 & 0 & 1 \end{pmatrix}$$

$$A^{10} = \begin{pmatrix} 7 & 5 & 11 & 10 & 5 \\ 5 & 2 & 5 & 6 & 4 \\ 10 & 5 & 7 & 5 & 6 \\ 11 & 10 & 15 & 7 & 5 \\ 5 & 6 & 10 & 5 & 2 \end{pmatrix}$$



# 复数与二阶实数矩阵

我们定义如下映射

$$f: \mathbb{C} \longrightarrow \mathbb{R}^{2 \times 2}$$
$$a + bi \longmapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

## 引理

映射  $f$  保持两边的加法和乘法. 即, 对任意  $z_1 = x_1 + iy_1, z_2 = x_2 + iy_2$

- $f(z_1 + z_2) = f(z_1) + f(z_2)$ ;
- $f(z_1 z_2) = f(z_1) f(z_2)$ .

## 例

计算  $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}^k$  和  $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}^k$ .



线性映射与矩阵有如下一一对应

$$\mathbb{F}^{m \times n} \xleftrightarrow{1:1} \text{从 } \mathbb{F}^n \text{ 到 } \mathbb{F}^m \text{ 的全体线性映射.}$$

若线性映射  $\mathcal{A}$  与矩阵  $A$  对应, 则对于任意  $\mathbf{x} \in \mathbb{F}^n$ ,

$$\mathcal{A}(\mathbf{x}) = \mathbf{Ax}.$$

# 坐标变换与矩阵乘法.

给定两个仿射坐标系  $[O; \vec{e}_1, \vec{e}_2, \vec{e}_3]$  和  $[O'; \vec{e}'_1, \vec{e}'_2, \vec{e}'_3]$ . 设  $O$  在

$[O'; \vec{e}'_1, \vec{e}'_2, \vec{e}'_3]$  下的坐标为  $X'_0 = \begin{pmatrix} x'_0 \\ y'_0 \\ z'_0 \end{pmatrix}$ , 以及  $e_j$  在基  $e'_1, e'_2, e'_3$  下的坐标

为  $\begin{pmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \end{pmatrix}$ . 若空间中的点  $P$  在两个坐标系下的坐标分别为  $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$  和

$$X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \text{ 则 } \begin{cases} x' = a_{11}x + a_{12}y + a_{13}z + x'_0 \\ y' = a_{21}x + a_{22}y + a_{23}z + y'_0 \\ z' = a_{31}x + a_{32}y + a_{33}z + z'_0 \end{cases}$$

用矩阵的乘法则表示:

$$X' = AX + X'_0$$

其中  $A = (a_{ij})_{3 \times 3}$ .

解一元一次方程.

$$ax = b \Rightarrow a^{-1}(ax) = a^{-1}b$$

$$\xrightarrow{\text{结合律}} (a^{-1}a)x = a^{-1}b$$

$$\xrightarrow{a^{-1}a=1} 1x = a^{-1}b$$

$$\xrightarrow{1x=x} x = a^{-1}b$$

$$\text{验证: } \xrightarrow{aa^{-1}=1} a(a^{-1}b) = b$$

解一般线性方程组

$$Ax = b \Rightarrow A^{-1}(Ax) = A^{-1}b$$

$$\xrightarrow{\text{结合律}} (A^{-1}A)x = A^{-1}b$$

$$\xrightarrow{A^{-1}A=I} Ix = A^{-1}b$$

$$\xrightarrow{Ix=x} x = A^{-1}b$$

$$\text{验证: } \xrightarrow{AA^{-1}=I} A(A^{-1}b) = b$$

要让上述过程有意义, 仅需要存在矩阵  $A^{-1}$  满足  $A^{-1}A = I = AA^{-1}$ .

例 (鸡兔同笼)

$$\begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 35 \\ 94 \end{pmatrix}.$$

解: 因为

$$\begin{pmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix},$$

所以

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 35 \\ 94 \end{pmatrix} = \begin{pmatrix} 23 \\ 12 \end{pmatrix}.$$

## 定义 (逆矩阵)

设  $A$  为数域  $\mathbb{F}$  上的  $n$  阶方阵. 如果存在  $n$  阶方阵  $X$  满足

$$XA = I = AX$$

则称  $A$  可逆, 并称  $X$  为  $A$  的逆矩阵. 记做  $A^{-1}$ .

奇异方阵, 非奇异方阵

注: 若  $A \in \mathbb{F}^{m \times n}$  不为方阵, 则一定不存在矩阵  $X \in \mathbb{F}^{n \times m}$  使得

$$AX = I_m \quad \text{且} \quad XA = I_n.$$

这也是我们为什么不考虑非方阵的逆.

## 引理 (存在则唯一)

若  $X$  和  $Y$  都为  $A$  的逆矩阵, 则  $X = Y$ .

## 性质 (逆矩阵的基本性质)

设  $A$  和  $B$  为同阶可逆方阵,  $\lambda$  为非零常数. 则

- ❶  $(A^{-1})^{-1} = A$ ;
- ❷  $(\lambda A)^{-1} = \lambda^{-1} A^{-1}$ ;
- ❸  $(AB)^{-1} = B^{-1} A^{-1}$  (穿脱原理).



例

若  $ad \neq bc$ , 则  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  可逆. 此时

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

例

若  $a_1, a_2, \dots, a_n$  全不为零, 则

$$\text{diag}(a_1, a_2, \dots, a_n)^{-1} = \text{diag}\left(\frac{1}{a_1}, \frac{1}{a_2}, \dots, \frac{1}{a_n}\right).$$

## 例

- ① 证明若  $f(A) = 0$  且  $f(a) \neq 0$ , 则  $A - aI$  可逆.
- ② 证明  $I + J_n$  可逆.
- ③ 已知  $2A^2 - 3A + 4I = 0$ , 求  $A$  和  $A - I$  的逆.
- ④ 已知  $3A^3 - 2A^2 + 5A + I = 0$ , 求  $A$  和  $A + I$  的逆.

## 例

证明若  $A, B$  和  $A + B$  均可逆, 则  $A^{-1} + B^{-1}$  也可逆.

问题: 如何求一般可逆矩阵的逆?

## 定义

- ① 设  $A = (a_{ij})_{m \times n}$  为数域  $\mathbb{F}$  上的矩阵. 我们称矩阵

$$\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

为  $A$  的**转置矩阵**, 记为  $A^T$ ;

- ② 设  $A = (a_{ij})_{m \times n}$  为复数域  $\mathbb{C}$  上的矩阵. 我们称矩阵

$$\begin{pmatrix} \overline{a_{11}} & \overline{a_{12}} & \cdots & \overline{a_{1n}} \\ \overline{a_{21}} & \overline{a_{22}} & \cdots & \overline{a_{2n}} \\ \cdots & \cdots & \cdots & \cdots \\ \overline{a_{m1}} & \overline{a_{m2}} & \cdots & \overline{a_{mn}} \end{pmatrix}$$

为  $A$  的**共轭矩阵**, 记为  $\bar{A}$ ;

- ③ 设  $A = (a_{ij})_{n \times n}$  为数域  $\mathbb{F}$  上的方阵. 我们称对角线元素之和为矩阵  $A$  的**迹**, 记为

$$\text{tr}(A) := a_{11} + \cdots + a_{nn}.$$

## 例

已知  $A = \begin{pmatrix} 1+i & 2-i \\ 3-i & 4+i \end{pmatrix}$ . 求  $A^T, \bar{A}$  和  $\text{tr}(A)$ .

# 转置, 共轭与迹的基本性质

## 引理 (转置和共轭的基本性质)

- $(A + B)^T = A^T + B^T$ ;
- $(\lambda A)^T = \lambda A^T$ ;
- $(AB)^T = B^T A^T$ ;
- $(A^{-1})^T = (A^T)^{-1}$ .
- $\overline{A + B} = \overline{A} + \overline{B}$ ;
- $\overline{\lambda A} = \overline{\lambda} \overline{A}$ ;
- $\overline{A \cdot B} = \overline{A} \cdot \overline{B}$ ;
- $\overline{A^{-1}} = \overline{A}^{-1}$ .

## 引理 (迹的基本性质)

- $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$ ;
- $\text{tr}(\lambda A) = \lambda \text{tr}(A)$ ;
- $\text{tr}(A^T) = \text{tr}(A)$ ,  $\text{tr}(\overline{A}) = \overline{\text{tr}(A)}$ ;
- $\text{tr}(AB) = \text{tr}(BA)$ .

## 例

设  $A$  为复矩阵. 证明若  $\text{tr}(A\overline{A}^T) = 0$ , 则  $A = 0$ .

# 分块矩阵

## 定义 (分块矩阵)

对于一个矩阵  $A$ , 我们将它的行和列分为几个部分后得到的由小矩阵  $A_{ij}$  组成的矩阵

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix} = (A_{ij})_{r \times s}$$

称为 **分块矩阵**, 其中每个  $A_{ij}$  称为 **子块**.

## 例

我们将矩阵的第  $i$  行记为  $\beta_i$  第  $j$  列记为  $\alpha_j$ , 则

$$A = (\alpha_1, \cdots, \alpha_n) = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}.$$

# 分块矩阵例子

$$A = \begin{pmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{pmatrix} = \left( \begin{array}{ccc|c} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{array} \right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

- 行的分隔方式为  $3 = 1 + 2$ ,
- 列的分隔方式为  $4 = 3 + 1$ 。

$$A = \begin{pmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{pmatrix} = \left( \begin{array}{c|c|c|c} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{array} \right) = ( \beta_1 \quad \beta_2 \quad \beta_3 \quad \beta_4 )$$

- 行的分隔方式为  $3 = 3$ ,
- 列的分隔方式为  $4 = 1 + 1 + 1 + 1$ 。

# 几类特殊分块矩阵

准对角矩阵, 准上三角矩阵, 准下三角矩阵.

$$\begin{pmatrix} * & * & & \\ & * & & \\ & & * & \\ & & * & * \end{pmatrix}$$

注: 依赖于分块方式.

## 引理 (分块矩阵基本性质)

- $(A_{ij})_{r \times s} + (B_{ij})_{r \times s} = (A_{ij} + B_{ij})_{r \times s}$ ;
- $\lambda(A_{ij})_{r \times s} = (\lambda A_{ij})_{r \times s}$ ;
- $((A_{ij})_{r \times s})^T = (B_{ij})_{s \times r}$ , 其中  $B_{ij} = A_{ji}^T$ ;
- $\overline{(A_{ij})_{r \times s}} = (\overline{A_{ij}})_{r \times s}$ ;
- 若  $A = (A_{ij})_{r \times r}$  为方阵且行列拆分方式相同, 则

$$\text{tr}(A) = \sum_{i=1}^r \text{tr}(A_{ii});$$

- 若  $A_1, \dots, A_r$  均可逆, 则  $\text{diag}(A_1, A_2, \dots, A_r)$  也可逆, 且逆矩阵为  $\text{diag}(A_1, A_2, \dots, A_r)^{-1} = \text{diag}(A_1^{-1}, A_2^{-1}, \dots, A_r^{-1})$ .



# 分块矩阵的乘法

两个分块矩阵可以分块相乘的要求是:

- ① 第一个矩阵的列的数目等于第二个矩阵的行的数目;
- ② 第一个矩阵的列的分隔方式等于第二个矩阵的行的分隔方式.

分块相乘的方法与矩阵的乘法法则一致.

## 定理

若  $A = (A_{ij})_{r \times s}$  的列拆分方式与  $B = (B_{jk})_{s \times t}$  的行拆分方式相同, 则

$$AB = \left( \sum_{j=1}^s A_{ij} B_{jk} \right)_{r \times t}.$$

证明: 任取不大于  $A$  行数的正整数  $p$  和不大于  $B$  列数的正整数  $q$ . 不妨设  $A$  的  $(p, 1)$  元素处在  $A_{i1}$  的第  $(u, 1)$  位置,  $B$  的  $(1, q)$  元素处在  $B_{1k}$  的第  $(1, w)$  位置. 我们需要验证:  $(AB)_{(pq)} = \left( \sum_{j=1}^s A_{ij} B_{jk} \right)_{(uw)}$ . 设  $A$  的列数拆分方式为

$n = n_1 + n_2 + \cdots + n_s$ . 则

$$\left( \sum_{j=1}^s A_{ij} B_{jk} \right)_{(uw)} = \sum_{j=1}^s (A_{ij} B_{jk})_{(uw)} = \sum_{j=1}^s \sum_{v=1}^{n_j} (A_{ij})_{(uv)} (B_{jk})_{(vw)} = (AB)_{(pq)}.$$

# 分块矩阵的乘法

例

选择矩阵  $A$  和  $B$  的合适分块并计算它们的乘积  $AB$ , 其中

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 3 & 2 & 0 \\ 5 & 2 & 0 & 2 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 & 2 & 1 \\ 0 & -1 & 3 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

例

已知  $A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ . 求  $A^k$ .

解: 归纳  $\Rightarrow A^k = \begin{pmatrix} B^k & kB^{k-1} \\ 0 & B^k \end{pmatrix}, B^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}.$

例

若  $A$  与  $B$  可逆, 求  $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$  和  $\begin{pmatrix} A & 0 \\ D & B \end{pmatrix}$  的逆.

# 相关考试题

问题 (24 春期中, 填空, 每题 5 分)

1). 方程  $X \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$  的解为  $X =$  \_\_\_\_\_.

2).  $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}^n =$  \_\_\_\_\_.

问题 (24 春期中, 判断题 (5'))

若  $n$  阶方阵  $A$  满足  $A^2 = I_n$ , 则  $A = I_n$  或  $A = -I_n$ .

问题 (24 秋期中, 填空, 每题 5 分)

1). 设  $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ ,  $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$   $Q = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ , 则

$P^2 A Q^2 =$  \_\_\_\_\_

2). 设  $A, B$  均为可逆矩阵, 则  $\begin{pmatrix} & A \\ B & \end{pmatrix}^{-1} =$  \_\_\_\_\_.



## 第四章 行列式

# 行列式的引入

考察线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \quad (*)$$

当  $n = 2$  时, 若  $a_{11}a_{22} - a_{12}a_{21} \neq 0$ , 则通过消元法可证明方程组有唯一解

$$\begin{cases} x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \\ x_2 = \frac{b_2a_{11} - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}} \end{cases}$$

对任意四个数  $a, b, c, d$ , 引入记号

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - bc,$$

则方程的公式解可简洁地表示为

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

当  $n = 3$  时, 方程组为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

若方程组有唯一解, 通过消元法 (见课本 p64-65) 可得

$$\begin{cases} x_1 = \frac{b_1 \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - b_2 \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + b_3 \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}}{a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}} \\ x_2 = \frac{-b_1 \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + b_2 \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - b_3 \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}}{a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}} \\ x_3 = \frac{b_1 \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} - b_2 \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + b_3 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}{a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}} \end{cases}$$

引入记号

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} := \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ := a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

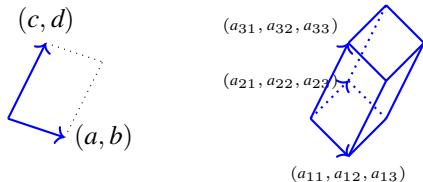
则方程的公式解可简化为

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

利用归纳法, 我们可以引入  $n$ -阶行列式, 将其定义为  $n$  个  $n-1$  行列式的某个组合. 从而将一般的  $n$ -元线性方程组的公式解表示为两个  $n$ -阶行列式的比值.



# 行列式的引入



- 给定两个二维数组向量  $(a, b)$  和  $(c, d)$ . 如图我们有平行四边形. 这时平行四边形的 (有向) 面积为  $S = ad - bc =: \begin{vmatrix} a & b \\ c & d \end{vmatrix}$ .
- 给定三个三维数组向量  $(a_{11}, a_{12}, a_{13})$ ,  $(a_{21}, a_{22}, a_{23})$  和  $(a_{31}, a_{32}, a_{33})$ . 如图我们可构造一个平行六面体. 其 (有向) 体积正好为

$$\Delta = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

- 若给定  $n$  个  $n$  维数组向量呢?

# 行列式的定义(高维平行多面体的有向体积)

## 定义(行列式)

方阵  $A = (a_{ij})_{n \times n}$  的行列式记为

$$\det(A) \quad \text{或} \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

当  $n = 1$  时,

$$\det(A) := a_{11}.$$

当  $n \geq 2$  时,  $\det(A)$  递归地定义为

$$\det(A) := \sum_{k=1}^n (-1)^{k+1} a_{1k} \cdot \begin{vmatrix} a_{12} & a_{13} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k-1,2} & a_{k-1,3} & \cdots & a_{k-1,n} \\ a_{k+1,2} & a_{k+1,3} & \cdots & a_{k+1,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n2} & a_{n,3} & \cdots & a_{nn} \end{vmatrix}.$$

例

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_1 b_3 c_2 - a_2 b_1 c_2 - a_3 b_2 c_1.$$

问题

用定义计算  $n$  阶行列式需要计算多少个乘法？

$$10 \text{ 阶行列式} \rightsquigarrow 10! \times 9 \text{ 个乘法} \rightsquigarrow \frac{10! \times 9}{24 \times 60 \times 60} = 378 \text{ 天}.$$

## 定义

由矩阵  $A = (a_{ij})_{m \times n}$  的第  $i_1, i_2, \dots, i_r$  行和第  $j_1, j_2, \dots, j_s$  列上的元素依次排列组成的  $r \times s$  矩阵称为  $A$  的**子矩阵**, 记作

$$A \begin{pmatrix} i_1 i_2 \cdots i_r \\ j_1 j_2 \cdots j_s \end{pmatrix} = \begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_s} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_s} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_r j_1} & a_{i_r j_2} & \cdots & a_{i_r j_s} \end{pmatrix}$$

## 例

$$\text{若 } A = \begin{pmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{pmatrix}, \text{ 则 } A \begin{pmatrix} 13 \\ 13 \end{pmatrix} = \begin{pmatrix} 11 & 13 \\ 31 & 33 \end{pmatrix};$$

$$A \begin{pmatrix} 412 \\ 13 \end{pmatrix} = \begin{pmatrix} 41 & 43 \\ 11 & 13 \\ 21 & 23 \end{pmatrix};$$

## 定义

给定一个  $n$  阶矩阵  $A = (a_{ij})_{n \times n}$ .

- ① 称  $M_{ij} := \det \left( A \begin{pmatrix} 1, \dots, i-1, i+1, \dots, n \\ 1, \dots, j-1, j+1, \dots, n \end{pmatrix} \right)$  为元素  $a_{ij}$  的余子式.
- ② 称  $A_{ij} := (-1)^{i+j} M_{ij}$  为  $a_{ij}$  的代数余子式.
- ③ 更一般地, 称  $A$  的某个  $k$  阶子矩阵的行列式为  $A$  的  $k$  阶子式, 删去子式所在的行列所得的矩阵的行列式称为该子式的余子式.
- ④ 特别地, 我们称  $\det \left( A \begin{pmatrix} 1, 2, \dots, k \\ 1, 2, \dots, k \end{pmatrix} \right)$  为  $A$  的  $k$  阶主子式.

根据这个定义, 我们可以将行列式的定义式简写为

$$\det(A) := \sum_{i=1}^n a_{i1} A_{i1} = \sum_{i=1}^n (-1)^{i+1} a_{i1} M_{i1}.$$

例

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \ddots & \vdots \\ & & \ddots & a_{n2} \\ & & & a_{nn} \end{vmatrix} = a_{11}a_{22} \cdots a_{nn}.$$

例

$$\begin{vmatrix} & & & a_{1n} \\ & & \ddots & a_{2n} \\ & & & \vdots \\ a_{n-1,2} & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} a_{n1}a_{n-1,2} \cdots a_{1n}.$$

# 行列式列展开

## 定理 (行列式列展开)

设  $A = (a_{ij})_{n \times n}$  为  $n$  阶方阵. 则对任意  $1 \leq j \leq n$

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

思路: 根据对行列式阶数的归纳, 将行列式

$$\det(A) := \sum_{i'=1}^n (-1)^{i'+1} a_{i'1} M_{i'1}$$

中的  $M_{i'1}$  按第  $j-1$  列 (即, 原行列式的第  $j$  行) 再次展开. 同时右式

$$\sum_{i=1}^n (-1)^{i+j} a_{ij} M_{ij}$$

中的  $M_{ij}$  按第一列展开. 然后比较两式是否相等即可.

记  $M \begin{bmatrix} i_1 & i_2 \\ j_1 & j_2 \end{bmatrix}$  为矩阵  $A$  去掉第  $i_1, i_2$  行和  $j_1, j_2$  列后的行列式。即

$$M \begin{bmatrix} i_1 & i_2 \\ j_1 & j_2 \end{bmatrix} := \det \left( A \begin{pmatrix} 1, \dots, i_1 - 1, i_1 + 1, \dots, i_2 - 1, i_2 + 1, \dots, n \\ 1, \dots, j_1 - 1, j_1 + 1, \dots, j_2 - 1, j_2 + 1, \dots, n \end{pmatrix} \right).$$

对行列式的阶数  $n$  进行归纳。当  $n = 1$ , 结论显然成立。假设结论对  $n-1$  阶行列式成立。则

$$\begin{aligned} \det(A) &= \sum_{i'=1}^n (-1)^{i'+1} a_{i'1} M_{i'1} \\ &= \sum_{i'=1}^n (-1)^{i'+1} a_{i'1} \left( \sum_{i=1}^{i'-1} (-1)^{(j-1)+i} a_{ij} \cdot M \begin{bmatrix} i' & i \\ 1 & j \end{bmatrix} + \sum_{i=i'+1}^n (-1)^{(i-1)+(j-1)} a_{ij} \cdot M \begin{bmatrix} i' & i \\ 1 & j \end{bmatrix} \right) \\ &= \sum_{i=1}^n (-1)^{i+j} a_{ij} \left( \sum_{i'=1}^{i-1} (-1)^{i'+1} a_{i'1} M \begin{bmatrix} i' & i \\ 1 & j \end{bmatrix} + \sum_{i'=i+1}^n (-1)^{(i'-1)+1} a_{i'1} \cdot M \begin{bmatrix} i' & i \\ 1 & j \end{bmatrix} \right) \\ &= \sum_{i=1}^n (-1)^{i+j} a_{ij} M_{ij} \\ &= \sum_{i=1}^n a_{ij} A_{ij} \end{aligned}$$

其中第二个等式为对  $n-1$  阶行列式  $M_{i'1}$  的第  $j-1$  列进行展开, 第三个等式为交换求和号, 第四个等式为对  $n-1$  阶行列式  $M_{ij}$  的第 1 列进行展开。



例

$$\begin{vmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \ddots & \ddots & \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} \end{vmatrix} = a_{11}a_{22} \cdots a_{nn}.$$

例

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & \ddots & \\ \vdots & & & \\ a_{n1} & a_{n-1,2} & & \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} a_{n1}a_{n-1,2} \cdots a_{1n}.$$

# 行列式展开

## 定理 (行列式行展开)

设  $A = (a_{ij})_{n \times n}$  为  $n$  阶方阵. 则对任意  $1 \leq i \leq n$

$$\det(A) = \sum_{j=1}^n a_{ij} A_{kj} = \sum_{j=1}^n (-1)^{k+j} a_{kj} M_{kj}.$$

思路: 不妨设  $k = 1$ . 根据对行列式阶数的归纳, 将行列式

$$\det(A) := \sum_{i=1}^n (-1)^{i+1} a_{i1} M_{i1}$$

中除  $M_{11}$  以外的  $M_{i1}$  按第一行展开. 同时右式

$$\sum_{i=1}^n (-1)^{i+j} a_{ij} M_{ij}$$

中除  $M_{11}$  以外的  $M_{1j}$  按第一列展开. 然后比较两式是否相等即可.

$$M \begin{bmatrix} i_1 & i_2 \\ j_1 & j_2 \end{bmatrix} := \det \left( A \begin{pmatrix} 1, \dots, i_1 - 1, i_1 + 1, \dots, i_2 - 1, i_2 + 1, \dots, n \\ 1, \dots, j_1 - 1, j_1 + 1, \dots, j_2 - 1, j_2 + 1, \dots, n \end{pmatrix} \right).$$

对行列式的阶数  $n$  进行归纳。当  $n = 1$ , 结论显然成立。假设结论对  $n-1$  阶行列式成立。则

$$\begin{aligned} \det(A) &= \sum_{i=1}^n (-1)^{i+1} a_{i1} M_{i1} \\ &= a_{11} M_{11} + \sum_{i=2}^n (-1)^{i+1} a_{i1} \left( \sum_{j=2}^n (-1)^{1+(j-1)} a_{1j} \cdot M \begin{bmatrix} 1i \\ 1j \end{bmatrix} \right) \\ &= a_{11} M_{11} + \sum_{j=2}^n (-1)^{1+j} a_{1j} \left( \sum_{i=2}^n (-1)^{(i-1)+1} a_{i1} \cdot M \begin{bmatrix} 1i \\ 1j \end{bmatrix} \right) \\ &= \sum_{j=1}^n (-1)^{1+j} a_{1j} M_{1j} = \sum_{j=1}^n a_{1j} A_{1j} \end{aligned}$$

其中第二个等式为对  $n-1$  阶行列式  $M_{i1}$  的第 1 行展开, 第三个等式为交换求和号, 第四个等式为对  $n-1$  阶行列式  $M_{1j}$  的第 1 列展开。

## 定理

方阵  $A$  的行列式具有下列性质:

- ❶ 交换  $A$  某两行 (列) 得  $B$ , 则  $\det(B) = -\det(A)$ .
- ❷ 将  $A$  的某一行 (列) 乘  $\lambda$  得  $B$ , 则  $\det(B) = \lambda \det(A)$ .
- ❸ 若  $A$  的某一行 (列) 是两个向量之和, 则  $\det(A)$  可拆成两个行列式之和.

证明思路: 将  $\det(A)$  按第  $p, q$  两行展开得

$$\det(A) = \sum_{i=1}^n \sum_{j=1}^{i-1} (-1)^{p+q+i+j-1} (a_{pi}a_{qj} - a_{pj}a_{qi}) M \begin{bmatrix} p, q \\ i, j \end{bmatrix}.$$

几何解释?

## 推论

- ④ 若  $A$  的某两行(列)成比例, 则  $\det(A) = 0$ .
- ⑤ 将  $A$  的某一行(列) $\lambda$  倍加到另一行(列)得  $B$ , 则  $\det(B) = \det(A)$ .

## 几何解释?

设  $A = (a_{ij})_{n \times n} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ . 则  $\det$  可以看成  $\alpha_1, \dots, \alpha_n$  的函数. 这一函数

满足如下性质:

- ① 反对称性:  $\det(\dots, \alpha_i, \dots, \alpha_j, \dots) = -\det(\dots, \alpha_j, \dots, \alpha_i, \dots)$ ;
- ② 多重线性:  
 $\det(\dots, \lambda\alpha_i + \mu\beta_i, \dots) = \lambda\det(\dots, \alpha_i, \dots) + \mu\det(\dots, \beta_i, \dots)$ ;
- ③ 规范性:  $\det(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = 1$ .

注: 这三条性质唯一地确定了函数  $\det$ . (几何意义!!)

例:  $\det(\alpha + \beta, \beta + \gamma, \gamma + \alpha) = 2\det(\alpha, \beta, \gamma)$ .

$$\begin{aligned}
 \det(A) &= \det \left( \sum_{j_1=1}^n a_{1j_1} \vec{e}_{j_1}, \sum_{j_2=1}^n a_{2j_2} \vec{e}_{j_2}, \dots, \sum_{j_n=1}^n a_{nj_n} \vec{e}_{j_n} \right) \\
 &= \sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} a_{2j_2} \cdots a_{nj_n} \det(\vec{e}_{j_1}, \vec{e}_{j_2}, \dots, \vec{e}_{j_n}) \\
 &= \sum_{(j_1 j_2 \cdots j_n) \in S_n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \underbrace{\det(\vec{e}_{j_1}, \vec{e}_{j_2}, \dots, \vec{e}_{j_n})}_{=??}
 \end{aligned}$$

其中  $S_n$  为集合  $\{1, \dots, n\}$  的全体排列集.

## 定义

- ① 由  $n$  个两两不同的正数组成的有序数组  $(j_1, j_2, \dots, j_n)$  称为一个  $n$  元排列. 由  $1, 2, \dots, n$  组成的排列总数为  $n!$ .
- ② 称  $(1, 2, \dots, n)$  为标准排列.
- ③ 给定排列  $(j_1, j_2, \dots, j_n)$ . 若  $1 \leq p < q \leq n$  且  $j_p > j_q$ , 则称二元组  $(j_p, j_q)$  为排列  $(j_1, j_2, \dots, j_n)$  的一个逆序. 例如:  $(1, 4, 3, 2)$  的有三个逆序  $(4, 3), (4, 2), (3, 2)$ .
- ④ 排列  $(j_1, j_2, \dots, j_n)$  的逆序总数记为  $\tau(j_1, j_2, \dots, j_n)$ . 例如:  
 $\tau(1, 4, 3, 2) = 3$ .
- ⑤ 若  $\tau(j_1, j_2, \dots, j_n)$  为奇数, 则称  $(j_1, j_2, \dots, j_n)$  为奇排列. 否则称之为偶排列.
- ⑥ 将一个排列中的两个元素互换位置, 其余元素位置不变. 这个过程称为一个对换.

## 引理

- ① 对换改变奇偶性.
- ② 排列  $(j_1, j_2, \dots, j_n)$  可经过  $\tau(j_1, j_2, \dots, j_n)$  次相邻位置的对换变为标准排列.

## 定理 (行列式显示表示)

设  $A = (a_{ij})_{n \times n}$  为  $n$  阶方阵. 则

$$\det(A) = \sum_{(j_1 j_2 \cdots j_n) \in S_n} (-1)^{\tau(j_1 j_2, \cdots, j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$



## 二三阶行列式示意图

例

①  $S_2 = \{(12), (21)\}, \tau(12) = 0, \tau(21) = 1$

$$\Rightarrow \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

②  $S_3 = \{(123), (231), (312), (132), (213), (321)\}, \tau(123) = 0,$   
 $\tau(231) = \tau(312) = 2, \tau(213) = \tau(132) = 1, \tau(321) = 3,$

$$\begin{aligned} & \bullet \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = +a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & \quad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \\ & \bullet \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = +a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & \quad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \end{aligned}$$

例

$$\begin{vmatrix} & & & a_1 \\ & & \ddots & \\ & a_{n-1} & & \\ a_n & & & \end{vmatrix} = (-1)^{\frac{(n)(n-1)}{2}} a_1 a_2 \cdots a_n.$$

例

求下列三阶行列式:  $\begin{vmatrix} 1 & 2 & 3 \\ 8 & 0 & 4 \\ 7 & 6 & 5 \end{vmatrix}$  和  $\begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 2 & 7 & 8 \end{vmatrix}.$

# 三角分块矩阵行列式

## 性质

设  $A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ & A_{22} & \cdots & A_{2k} \\ & & \ddots & \vdots \\ & & & A_{kk} \end{pmatrix}$  为分块矩阵, 其中  $A_{ii}$  均为方阵. 则

$$\det(A) = \prod_{i=1}^k \det(A_{ii}).$$

证明思路: 只需考虑  $k = 2$ . 此时, 不妨设拆分方式为  $n = r + (n - r)$ . 则

$$\begin{aligned} \det(A) &= \sum_{(j_1, j_2, \dots, j_n) \in S_n} (-1)^{\tau(j_1, \dots, j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \\ &= \sum_{\substack{(j_1, j_2, \dots, j_r) \in S_r \\ (j_{r+1}, \dots, j_n) \in S_{n-r}(r+1, \dots, n)}} \left( (-1)^{\tau(j_1, \dots, j_r)} a_{1j_1} \cdots a_{rj_r} \right) \cdot \left( (-1)^{\tau(j_{r+1}, \dots, j_n)} a_{r+1, j_{r+1}} \cdots a_{nj_n} \right) \\ &= \det(A_{11}) \det(A_{22}) \end{aligned}$$

# 行列式基本性质

## 定理

$$\det(A) = \det(A^T).$$

证明：比较两边显示展开。(或，对阶数归纳，并利用展开式。)

## 定理

$$\det(AB) = \det(A) \det(B).$$

几何解释？

## 例

设  $m > n$ ,  $A \in F^{m \times n}$ ,  $B \in F^{n \times m}$ . 则

$$\det(AB) = 0.$$

## 例

设  $A \in F^{n \times n}$ ,  $B \in F^{n \times m}$ ,  $C \in F^{m \times n}$ ,  $D \in F^{m \times m}$ . 若  $A$  可逆, 则

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A) \cdot \det(D - CA^{-1}B).$$

例

$$\text{求} \begin{vmatrix} 3 & x+y+z & x^2+y^2+z^2 \\ x+y+z & x^2+y^2+z^2 & x^3+y^3+z^3 \\ x^2+y^2+z^2 & x^3+y^3+z^3 & x^4+y^4+z^4 \end{vmatrix}$$

$$\begin{pmatrix} 3 & x+y+z & x^2+y^2+z^2 \\ x+y+z & x^2+y^2+z^2 & x^3+y^3+z^3 \\ x^2+y^2+z^2 & x^3+y^3+z^3 & x^4+y^4+z^4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{pmatrix} \begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{pmatrix}$$

# 矩阵可逆的判定条件

## 定义 (伴随矩阵)

设  $A = (a_{ij})_{n \times n}$  为  $n$  阶方阵. 称矩阵

$$A^* := \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} = (A_{ij})_{n \times n}^T$$

为  $A$  的伴随矩阵.

## 定理

设  $A = (a_{ij})_{n \times n}$  为  $n$  阶方阵. 则

$$A^*A = AA^* = \det(A)I_n.$$

注: 定理等价于  $\sum_{k=1}^n a_{ik}A_{jk} = \det(A)\delta_{ij}$  以及  $\sum_{k=1}^n a_{ki}A_{kj} = \det(A)\delta_{ij}$ , 其中  $\delta_{ij}$  为 Kronecker 记号, 满足  $(\delta_{ij})_{n \times n} = I_n$ .

## 定理

设  $A$  为  $n$  阶方阵, 则以下几条等价

- ①  $A$  可逆 (i.e. 存在  $X$  使得  $AX = I = XA$ .);
- ②  $\det(A) \neq 0$ ;
- ③ 存在  $X$  使得  $I = XA$ ;
- ④ 存在  $X$  使得  $AX = I$ .

若以上几条成立, 则  $A^{-1} = \frac{1}{\det(A)} A^*$ .

## 定理 (Cramer 法则)

若  $n$  阶方阵  $A = (a_{ij})_{n \times n} = (\alpha_1, \dots, \alpha_n)$  可逆. 则一般线性方程组

$$AX = b$$

有唯一解

$$(x_1, \dots, x_n) = \left( \frac{\Delta_1}{\Delta}, \frac{\Delta_2}{\Delta}, \dots, \frac{\Delta_n}{\Delta} \right),$$

其中  $\Delta = \det(A)$ ,  $\Delta_i = \det(\alpha_1, \dots, \alpha_{i-1}, b, \alpha_{i+1}, \dots, \alpha_n)$ .

证明思路:  $X = A^{-1}b = \frac{1}{\det(A)}A^*b$ . 几何解释?

理论层面上, 这种方法既有用又美观, 但在实际应用中, 由于效率过低, 通常不会被采用来解方程。

## 例

使用 Cramer 法则求解线性方程组:

$$\begin{cases} \quad \quad +x_2 & +x_3 & +x_4 & = 1 \\ x_1 & & +x_3 & +x_4 & = 2 \\ x_1 & +x_2 & & +x_4 & = 3 \\ x_1 & +x_2 & +x_3 & & = 4 \end{cases} \Rightarrow (x_1, x_2, x_3, x_4) = (7/3, 4/3, 1/3, -2/3).$$



例

$$\begin{vmatrix} 0 & 5 & -4 & 5 \\ -3 & -1 & -5 & 3 \\ 3 & 1 & -2 & -3 \\ -1 & 4 & -5 & -1 \end{vmatrix} = 665.$$

例

$$\begin{vmatrix} x & 1 & \cdots & 1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & x & 1 \\ 1 & \cdots & 1 & x \end{vmatrix} = (x + n - 1)(x - 1)^{n-1}.$$

例

$$\begin{vmatrix} x & & & a_0 \\ -1 & \ddots & & \vdots \\ & \ddots & x & a_{n-2} \\ & & -1 & x + a_{n-1} \end{vmatrix} = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0.$$

例

$$\begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

例

$$\Delta_n := \begin{vmatrix} 2 & 1 & & \\ 1 & 2 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 2 \end{vmatrix} = n + 1.$$

## 第五章 矩阵 (二)

- 初等行变换
  - 交换两行
  - 某行乘以一个非零常数
  - 某行乘以一个常数加到另一行上
- 初等列变换
  - 交换两列
  - 某列乘以一个非零常数
  - 某列乘以一个常数加到另一列上

初等变换(6个) = 初等行变换(3个) + 初等列变换(3个)

问题: 对单位矩阵做初等变换, 我们会得到什么矩阵?

## 定义 (初等矩阵)

- ① 交换单位阵的第  $i, j$  行 (或  $i, j$  列) 得

$$S_{ij} := \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & & 1 \\ & & & \ddots & \\ & 1 & & & 0 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

- ② 将单位阵第  $i$  行 (或  $i$  列) 乘以非零常数  $\lambda$  得

$$D_i(\lambda) := \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \lambda & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

## 定义 (初等矩阵)

- ③ 将单位阵的第  $j$  行的  $\lambda$  倍加到第  $i$  行 (或第  $i$  列的  $\lambda$  倍加到第  $j$  列) 得

$$T_{ij}(\lambda) := \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & \lambda & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

我们称上面定义三类矩阵为初等矩阵.

# 初等矩阵基本性质

## 定理

- ① 对矩阵做初等行变换相当于矩阵左乘一个相应的初等矩阵.
- ② 对矩阵做初等列变换相当于矩阵右乘一个相应的初等矩阵.

证明思路: 将矩阵写成由向量组成的分块矩阵, 然后直接验证.

性质 (初等矩阵总是可逆的, 且其逆仍然是初等矩阵)

- ①  $S_{ij}$  对称且  $S_{ij}^{-1} = S_{ij}$ ;
- ②  $D_i(\lambda)$  为对角阵且  $D_i(\lambda)^{-1} = D_i(\lambda^{-1})$ ;
- ③  $T_{ij}(\lambda)$  为三角阵, 且  $T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda)$ ;

特别地, 由于可逆矩阵的乘积仍然可逆, 我们有任意有限多个初等矩阵的乘积也是可逆的.

# 高斯消元法与可逆阵的 $LU$ 分解.

由高斯消元法, 对任意矩阵  $A$  可通过初等行变换将其变为阶梯矩阵. 用初等矩阵, 这过程可描述为: 存在初等方阵  $P_1, P_2, \dots, P_s$  使得

$$P_s \cdots P_2 P_1 A =: J$$

为阶梯矩阵.

- 若  $A$  可逆, 则  $J$  为可逆上三角矩阵.
- 若  $P_1, P_2, \dots, P_s$  仅出现下三角初等矩阵  $T_{ij}(\lambda) (j < i)$ . 则记  $L := P_1^{-1} P_2^{-1} \cdots P_s^{-1}$  以及  $U := J$ , 称  $A = LU^2$  为  $A$  的  $LU$  分解.
- 若可逆阵  $A$  的  $LU$  分解存在, 则将解方程组  $Ax = b$  可转化为两个简单方程组  $Ly = b, Ux = y$  (这两个方程都可通过带入法求解).

---

<sup>2</sup>注:  $L$  的  $(i, j)$  元素正好为依次将  $A$  的  $(i, j)$  元素化为 0 所需减去的  $(j, j)$  元素的倍数.



## 例 (用 LU 分解求解线性方程组)

$$\begin{cases} 2x_1 + 3x_2 + 4x_3 = 1 \\ 4x_1 + 11x_2 + 14x_3 = 3 \\ 6x_1 + 19x_2 + 26x_3 = 5. \end{cases}$$

- 第一步：对系数矩阵进行 LU 分解得 (Gauss 消元法):

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 0 & 0 & 2 \end{pmatrix}.$$

- 第二步：求解  $Ly = b$ . 这一步通过自上而下带入完成.

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \Rightarrow y = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

- 第三步：求解  $Ux = y$ . 通过前向替换 (forward substitution)

$$\begin{pmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \Rightarrow x = \begin{pmatrix} \frac{1}{5} \\ \frac{1}{5} \\ 0 \end{pmatrix}.$$

## 定理

对于任意矩阵  $A \in F^{m \times n}$ , 存在  $m$  阶初等方阵  $P_1, P_2, \dots, P_s$  和  $n$  阶初等方阵  $Q_1, Q_2, \dots, Q_t$ , 以及非负整数  $r$ , 使得

$$P_s \cdots P_2 P_1 A Q_1 Q_2 \cdots Q_t = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

## 推论

- ① 对任意矩阵  $A \in \mathbb{F}^{m \times n}$ , 存在  $m$  阶可逆阵  $P$  和  $n$  阶可逆阵  $Q$  使得

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

- ② 设  $A$  为  $n$  阶方阵. 则  $A$  可逆当且仅当  $A$  可分解为初等方阵的乘积.

## 推论

设  $A$  为  $n$  阶可逆方阵, 则

- ① 可对  $A$  做一系列初等行变换变为最简形式  $I$ .
- ② 可对  $A$  做一系列初等列变换变为最简形式  $I$ .

问题: 如何求逆矩阵?

- **算法原理:** 通过初等变换我们可以找到初等矩阵  $P_1, P_2, \dots, P_s$  使得  $P_s \cdots P_2 P_1 A = I$ . 从而, 我们有

$$A^{-1} = P_s \cdots P_2 P_1.$$

- **具体实现过程:** 对矩阵  $(A, I)$  做行初等变换, 将第一个子矩阵变为单位阵. 则第二个子矩阵自动变为  $A$  的逆.

$$(A, I) \xrightarrow{\text{行初等变换}} P_s \cdots P_2 P_1 (A, I) = (I, A^{-1}).$$

例

求矩阵  $\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$  的逆.

$$\frac{1}{4} \begin{pmatrix} 3 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 3 \end{pmatrix}.$$

思考: 若  $A$  可逆, 如何求解  $AX = B$ ?

$$(A, B) \xrightarrow{\text{行初等变换}} P_s \cdots P_2 P_1 (A, B) = (I, A^{-1}B).$$

例

若  $\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} X = \begin{pmatrix} 3 & 1 \\ 4 & 0 \\ 3 & 1 \end{pmatrix}$ , 求  $X$ .

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

思考: 若  $A$  可逆, 如何求解  $XA = B$ ?

$$\begin{pmatrix} A \\ B \end{pmatrix} \xrightarrow{\text{列初等变换}} \begin{pmatrix} A \\ B \end{pmatrix} Q_1 Q_2 \cdots Q_t = \begin{pmatrix} I \\ BA^{-1} \end{pmatrix}.$$

例

若  $X \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 3 \\ 1 & 0 & 1 \end{pmatrix}$ , 求  $X$ .

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

# 分块矩阵的行列变换

设  $A$  为  $m \times n$  矩阵. 我们将其按  $m = m_1 + m_2 + \cdots + m_r$  和  $n = n_1 + n_2 + \cdots + n_s$  的拆分方式, 得到如下分块形式

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix} = (A_{ij})_{r \times s}.$$

类似于矩阵的初等变换, 我们可以对分块矩阵  $A$  做相似的操作:

- ① 交换分块矩阵的第  $i, j$  行 (或  $i, j$  列).
- ② 将分块矩阵的第  $i$  行左乘一个  $n_i$  阶 (或, 第  $i$  列右乘一个  $m_i$  阶) 可逆矩阵.
- ③ 将分块矩阵的第  $i$  行左乘一个  $n_j \times n_i$  矩阵加到第  $j$  行 (或, 第  $i$  列右乘一个  $m_i \times m_j$  矩阵加到第  $j$  列).

例

若  $A$  可逆, 则

$$\begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}$$

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & D - CA^{-1}B \end{pmatrix}$$

$$\begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{pmatrix}$$

例

设  $A, B, I$  为  $n$  阶方阵满足  $BA = 0$ . 则

$$\begin{pmatrix} I & A \\ B & I \end{pmatrix}^{-1} = \begin{pmatrix} I + AB & -A \\ -B & I \end{pmatrix}$$

$$\begin{pmatrix} I & A & I & 0 \\ B & I & 0 & I \end{pmatrix} \xrightarrow[\text{第一行左乘 } B]{\text{第二行减去}} \begin{pmatrix} I & A & I & 0 \\ 0 & I & -B & I \end{pmatrix}$$

$$\xrightarrow[\text{第二行左乘 } A]{\text{第一行减去}} \begin{pmatrix} I & 0 & I + AB & -A \\ 0 & I & -B & I \end{pmatrix}$$

## 定义 (相抵)

称两矩阵  $A, B$  **相抵**, 若存在可逆矩阵  $P$  和  $Q$  使得

$$B = PAQ.$$

注: 因可逆矩阵为方阵, 因此要使  $B = PAQ$  成立,  $B$  和  $A$  的行列数必须一致.

## 性质 (相抵的基本性质)

- ① 任意矩阵与自身相抵;
- ② 若  $A$  与  $B$  相抵, 则  $B$  与  $A$  也相抵;
- ③ 若  $A$  与  $B$  相抵且  $B$  与  $C$  相抵, 则  $A$  与  $C$  相抵.

满足上述三条性质的关系称为**等价关系**。换句话说, 上述性质表明相抵为等价关系。



# 等价关系在集合论中的严谨定义

## 定义 (等价关系严谨定义)

给定集合  $X$ , 我们称  $X \times X$  的一个子集为  $X$  上的一个关系. 设  $R$  为  $X$  上的一个关系, 即

$$R \subseteq X \times X.$$

我们称  $R$  是等价关系, 若  $R$  满足如下三条性质

- (自反性) 对任意  $x \in X$ , 我们有  $(x, x) \in R$ ;
- (对称性) 对任意  $x, y \in X$ ,  $(x, y) \in R$  当且仅当  $(y, x) \in R$ ;
- (传递性) 对任意  $x, y, z \in X$ , 若  $(x, y) \in R$ ,  $(y, z) \in R$  则  $(x, z) \in R$ .

此时, 对任意  $x, y \in X$ , 若  $(x, y) \in R$ , 则称  $x$  等价于  $y$ ; 否则, 称  $x$  不等价于  $y$ .

## 例

例子: 老乡; 同号; 同余等价关系;

反例: 朋友; 父子; 异号

# 集合上等价关系与拆分

对等价关系, 我们有如下通俗理解:

## 定理

给定集合某上的一个等价关系等价于给定这集合上的一个拆分.

证明思路:

- 给定  $X$  上一个等价关系  $R$ . 我们称  $X$  的子集

$$[x] := \{y \in X \mid (x, y) \in R\} \subseteq X$$

为  $x$  在等价关系  $R$  下的等价类.

- 对任意  $x, y \in X$ ,  $[x] \cap [y] \neq \emptyset$  当且仅当  $[x] = [y]$ .
- 集合  $X$  为一些等价类的无交并.
- 反之, 给定集合  $X$  的一个拆分  $X$ . 我们如下定义一个等价关系  $R$ : 对于任意  $x, y \in X$ ,  $(x, y) \in R$  当且仅当  $x, y$  落入同一个拆分子集中.

按照相抵关系, 我们将全体  $m \times n$  矩阵分成若干个两两不交的等价类.

基本问题:

- ① 如何判定两个矩阵是否相抵?
- ② 在每个等价类中有没有最简单的表达形式?

## 定理

对于任意矩阵  $A \in F^{m \times n}$ , 存在  $m$  阶可逆矩阵  $P$  和  $n$  阶可逆矩阵  $Q$  使得

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

其中非负整数  $r$  不依赖于  $P$  和  $Q$  的选取.

假若不唯一, 不妨设  $s > r$  且  $P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}$ . 对  $P$  和  $Q$  进行分块:

$$P = \begin{pmatrix} (P_{11})_{s \times r} & (P_{12})_{s \times (m-r)} \\ (P_{21})_{(m-s) \times r} & (P_{22})_{(m-s) \times (m-r)} \end{pmatrix}_{m \times m}$$
$$Q = \begin{pmatrix} (Q_{11})_{s \times r} & (Q_{12})_{s \times (n-r)} \\ (Q_{21})_{(n-s) \times r} & (Q_{22})_{(n-s) \times (n-r)} \end{pmatrix}_{n \times n}.$$

则  $P_{11}Q_{11} = I_s$  (矛盾!).

# 相抵判定

## 定义

定理中的  $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$  称为  $A$  的**相抵标准形**. 非负整数  $r$  称为  $A$  的**秩**, 记为  $\text{rank}(A)$  或  $r(A)$ . 若  $r = m$ , 则称  $A$  为**行满秩**. 若  $r = n$ , 则称  $A$  为**列满秩**.

## 推论 (相抵判定)

给定两矩阵  $A, B \in F^{m \times n}$ , 则  $A$  与  $B$  相抵当且仅当  $\text{rank}(A) = \text{rank}(B)$ .

## 推论

若  $P, Q$  可逆, 则  $\text{rank}(PAQ) = \text{rank}(A)$ .

## 例

$$\text{rank}(A^T) = \text{rank}(A).$$

$$\text{rank} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \text{rank}(A) + \text{rank}(B).$$

## 性质

设  $A \in F^{m \times n}$ ,  $P, Q$  分别为  $m, n$  阶初等方阵. 若  $A$  的  $k$  阶子式全为零, 则  $PA$  和  $AQ$  的所有  $k$  阶子式全为零.

## 性质 (秩的内蕴定义)

若矩阵  $A$  至少有一个  $r$  阶非零子式, 且  $A$  的所有  $r+1$  阶子式都为零, 则  $\text{rank}(A) = r$ .

例

求  $n$  阶方阵  $\begin{pmatrix} x & 1 & \cdots & 1 \\ 1 & x & \ddots & \vdots \\ \vdots & \ddots & x & 1 \\ 1 & \cdots & 1 & x \end{pmatrix}$  的秩.

例

求  $n$  阶方阵  $\begin{pmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ 1 & & & 1 \end{pmatrix}$  的秩.

例

每个非零矩阵都为 一个列满秩矩阵和一个行满秩矩阵的乘积.

例

每个秩为  $r$  的矩阵都可以写成  $r$  个秩为 1 的矩阵的和.

例

若  $A$  为列满秩  $m \times n$  矩阵, 则  $A$  为某个  $m$  阶可逆阵的前  $n$  列.

若  $A$  为行满秩  $m \times n$  矩阵, 则  $A$  为某个  $n$  阶可逆阵的前  $m$  列.



例

设矩阵  $A$  的列数和矩阵  $B$  的行数为  $n$ , 则

$$\text{rank}(A) + \text{rank}(B) - n \leq \text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B)).$$

例

若  $A$  为  $n$  阶方阵满足  $A^2 = A$ , 则  $\text{rank}(A) + \text{rank}(I - A) = n$ .

例

设  $A \in F^{m \times n}$ ,  $B \in F^{n \times m}$ . 则

$$m + \text{rank}(I_n - BA) = \text{rank} \begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix} = n + \text{rank}(I_m - AB).$$



## 第六章 线性空间

# $n$ 维数组空间, 线性相关性

$$\text{空间} \xrightarrow[1:1]{\text{坐标系}} \mathbb{R}^3 \xrightarrow{\text{推广}} \mathbb{F}^n = \mathbb{F} \times \mathbb{F} \times \cdots \times \mathbb{F}.$$

## 定义

设  $\mathbb{F}$  为数域. 带线性运算的  $n$  维数组向量全体

$$\{(a_1, a_2, \cdots, a_n) \mid a_i \in \mathbb{F}\}$$

称为  $n$  维数组空间. 记为  $\mathbb{F}^n$ .

线性组合 (组合系数, 线性表示)

## 定义 (线性相关, 线性无关)

一组向量  $\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_m$  称为 **线性相关**, 若存在一组不全为零的实数  $\lambda_1, \lambda_2, \cdots, \lambda_m$  使得

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \cdots + \lambda_m \vec{a}_m = 0.$$

反之, 则称向量  $\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_m$  的 **线性无关**.

给定空间中的  $m$  个向量  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ . 考虑这  $m$  个向量的全体线性组合

$$V := \langle \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \rangle := \{ \lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_m \vec{a}_m \mid \lambda_1, \dots, \lambda_m \in \mathbb{F} \}$$

这是  $\mathbb{F}^n$  的一个非空子集满足

- ① 任取  $\vec{b}_1, \vec{b}_2 \in V$ , 都有  $\vec{b}_1 + \vec{b}_2 \in V$ ;
- ② 任取  $\vec{b} \in V$  以及  $\lambda \in \mathbb{F}$ , 都有  $\lambda \vec{b} \in V$ .

上面两条也等价于

- ③ 任取  $\vec{b}_1, \dots, \vec{b}_\ell \in V$  以及  $\mu_1, \dots, \mu_\ell$ , 都有  $\mu_1 \vec{b}_1 + \dots + \mu_\ell \vec{b}_\ell \in V$ .

## 定义

设  $V$  为  $\mathbb{F}^n$  的一个非空子集. 若  $V$  满足上述前两条或第三条, 则称  $V$  为  $\mathbb{F}^n$  的**子空间**.

子空间可看成是**三维空间中过原点的线和平面在高维时的推广**.

# 子空间的例子

## 例 (平凡子空间)

$V = \{0\}$  或  $V = \mathbb{F}^n$ .

## 例 (生成子空间)

$\langle \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \rangle$  由  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  生成的子空间.

特别地, 给定一个阶矩阵  $A \in \mathbb{F}^{m \times n}$ , 则其所有列向量可生成  $\mathbb{F}^m$  的一个子空间, 其所有行向量可生成  $\mathbb{F}^n$  的一个子空间.

## 例 (用子空间的观点来理解线性方程组)

设  $A = (\vec{a}_1, \dots, \vec{a}_m)$ . 则  $Ax = \vec{b}$  有解当且仅当  $\vec{b} \in \langle \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \rangle$ .

## 例 (齐次线性方程组的解空间)

$n$  个变元的齐次线性方程组的解空间为  $\mathbb{F}^n$  的一个子空间.

## 例 (线性映射的像 (image))

设  $\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^m$  为一个线性映射. 则

$$\text{im}(\mathcal{A}) := \{\mathcal{A}(\vec{x}) \mid \vec{x} \in \mathbb{F}^n\} \subseteq \mathbb{F}^m$$

为  $\mathbb{F}^m$  的一个子空间。称为线性映射  $\mathcal{A}$  的像。

## 例 (线性映射的核 (kernel))

设  $\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^m$  为一个线性映射. 则

$$\ker(\mathcal{A}) := \{\vec{x} \in \mathbb{F}^n \mid \mathcal{A}(\vec{x}) = 0\} \subseteq \mathbb{F}^n$$

为  $\mathbb{F}^n$  的一个子空间。称为线性映射  $\mathcal{A}$  的核。

# 线性相关等价刻画

给定一组 (列) 向量  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ . 则以下几条相互等价:

- ①  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性相关;
- ② 存在  $i \in \{1, \dots, m\}$  以及  $\lambda_1 \cdots \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_m \in \mathbb{F}$  使得
$$\vec{a}_i = \lambda_1 \vec{a}_1 + \cdots + \lambda_{i-1} \vec{a}_{i-1} + \lambda_{i+1} \vec{a}_{i+1} + \cdots + \lambda_m \vec{a}_m;$$
- ③ 存在  $i \in \{1, \dots, m\}$  使得
$$\vec{a}_i \in \langle \vec{a}_1, \dots, \vec{a}_{i-1}, \vec{a}_{i+1}, \dots, \vec{a}_m \rangle;$$
- ④ 存在  $i \in \{1, \dots, m\}$  使得
$$\langle \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \rangle = \langle \vec{a}_1, \dots, \vec{a}_{i-1}, \vec{a}_{i+1}, \dots, \vec{a}_m \rangle;$$
- ⑤ 线性方程组  $AX = 0$  有非零解, 其中  $A = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m)$ ;
- ⑥  $A$  不是列满秩;
- ⑦  $\det(A) = 0$  (当  $m = n$  时).



# 线性无关等价刻画

给定一组 (列) 向量  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ . 则以下几条相互等价:

- ①  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性无关;
- ② 任取  $i \in \{1, \dots, m\}$  以及  $\lambda_1 \cdots \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_m \in \mathbb{F}$ ,  
 $\vec{a}_i \neq \lambda_1 \vec{a}_1 + \cdots + \lambda_{i-1} \vec{a}_{i-1} + \lambda_{i+1} \vec{a}_{i+1} + \cdots + \lambda_m \vec{a}_m$ ;
- ③ 任意  $i \in \{1, \dots, m\}$ ,  
 $\vec{a}_i \notin \langle \vec{a}_1, \dots, \vec{a}_{i-1}, \vec{a}_{i+1}, \dots, \vec{a}_m \rangle$ ;
- ④ 任意  $i \in \{1, \dots, m\}$ ,  
 $\langle \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \rangle \neq \langle \vec{a}_1, \dots, \vec{a}_{i-1}, \vec{a}_{i+1}, \dots, \vec{a}_m \rangle$ ;
- ⑤ 线性方程组  $AX = 0$  无非平凡解, 其中  $A = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m)$ ;
- ⑥  $A$  为列满秩;
- ⑦  $\det(A) \neq 0$  (当  $m = n$  时).

## 例

包含零向量的任意向量组线性相关.

## 例

设  $S_1, S$  为  $\mathbb{F}^n$  的两个有限子集满足  $S_1 \subseteq S$ . 则

- ①  $S_1$  线性相关  $\Rightarrow S$  线性相关;
- ②  $S$  线性无关  $\Rightarrow S_1$  线性无关;

## 例 (判定下列向量组的线性相关性)

- ①  $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$  单位坐标向量.
- ②  $\vec{a}_1 = (1, 0, 0, \dots, 0)^T, \vec{a}_2 = (1, 1, 0, \dots, 0)^T, \dots,$   
 $\vec{a}_n = (1, 1, 1, \dots, 1)^T.$
- ③  $\vec{a}_1 + \vec{a}_2, \vec{a}_2 + \vec{a}_3, \vec{a}_3 + \vec{a}_1$ , 其中  $\vec{a}_1, \vec{a}_2, \vec{a}_3$  线性无关.
- ④  $\vec{a}_1 = (3, 4, -2, 5)^T, \vec{a}_2 = (2, -5, 0, -3)^T, \vec{a}_3 = (5, 0, -1, 2)^T,$   
 $\vec{a}_4 = (3, 3, -3, 5)^T.$

# 线性映射与线性相关性

## 性质

任意给定一个线性映射  $\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^m$ . 对  $\mathbb{F}^n$  中任意有限个向量  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ , 都有

- ①  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性相关  $\Rightarrow \mathcal{A}(\vec{a}_1), \mathcal{A}(\vec{a}_2), \dots, \mathcal{A}(\vec{a}_m)$  线性相关;
- ②  $\mathcal{A}(\vec{a}_1), \mathcal{A}(\vec{a}_2), \dots, \mathcal{A}(\vec{a}_m)$  线性无关  $\Rightarrow \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性无关;

## 例 (投影)

给定  $m$  个  $r$  维数组向量  $\vec{a}_i = (a_{i1}, \dots, a_{ir}) \in \mathbb{F}^r$  ( $i = 1, \dots, m$ ). 将每个向量都扩充为一个  $n$  维向量  $\vec{b}_i = (a_{i1}, \dots, a_{ir}, a_{ir+1}, \dots, a_{in}) \in \mathbb{F}^n$ . 则

- ①  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性无关  $\Rightarrow \vec{b}_1, \vec{b}_2, \dots, \vec{b}_m$  线性无关;
- ②  $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_m$  线性相关  $\Rightarrow \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性相关;

# 极大无关组

## 定义 (极大无关组)

设  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  为一组向量. 若子向量组  $\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_r}$  线性无关, 且任加另一个向量  $\vec{a}_{i_{r+1}}$  后, 向量组  $\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_{r+1}}$  线性相关, 则称  $\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_r}$  为  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  的**极大无关组**.

## 性质 (通过生成子空间来判定极大无关组)

给定向量组  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  的一个子向量组  $\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_r}$ . 则  $\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_r}$  为  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  的极大无关组当且仅当

- $\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_r}$  线性无关, 且
- $\langle \vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_r} \rangle = \langle \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \rangle$ .

13. 若向量组  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in F^n$  线性无关, 而  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{b}$  线性相关, 则  $\mathbf{b}$  可以表示成  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$  的线性组合, 且表示唯一。
14. 证明向量表示基本定理: 设  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in F^n$  线性无关, 则任意向量  $\mathbf{b} \in F^n$  可以表示为  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$  的线性组合, 且表示唯一。
15. 证明: 非零向量组  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_s$  线性无关的充要条件是, 每个  $\mathbf{a}_i (1 \leq i \leq s)$  都不能用它前面的向量线性表示。
16. 设向量组  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_s$  线性无关,  $\mathbf{b} = \lambda_1 \mathbf{a}_1 + \dots + \lambda_s \mathbf{a}_s$ 。如  $\lambda_i \neq 0$ , 则用  $\mathbf{b}$  代替  $\mathbf{a}_i$  后, 向量组  $\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_s$  线性无关。
17. 设向量组  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$  线性无关, 且  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$  可以由向量组  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_r$  线性表示, 则  $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_r$  也线性无关。

# 极大无关组例子

例

证明  $\vec{a}_1 = (2, -1, 3, 1)$ ,  $\vec{a}_2 = (4, -2, 5, 4)$ ,  $\vec{a}_3 = (2, -1, 4, -1)$  中的任两个向量组成极大无关组.

$$3\vec{a}_1 = \vec{a}_2 + \vec{a}_3.$$

注: 不唯一! 如何寻找极大无关组? 一个一个去掉? 比较麻烦! 下面介绍简单方法.

# 寻找极大无关组的理论工具

原理: 初等行变换不改变列向量的线性相关性.

## 定理

设  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  为一组列向量. 对矩阵

$$A = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m) \in \mathbb{F}^{n \times m}$$

做一系列初等变换得到

$$B = (\vec{b}_1, \vec{b}_2, \dots, \vec{b}_m) \in \mathbb{F}^{n \times m}.$$

则对于任意  $i_1, i_2, \dots, i_r \in \{1, 2, \dots, m\}$ ,

- ①  $\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_r}$  线性相关 (无关)  $\Leftrightarrow \vec{b}_{i_1}, \vec{b}_{i_2}, \dots, \vec{b}_{i_r}$  线性相关 (无关);
- ②  $\vec{a}_{i_1}, \vec{a}_{i_2}, \dots, \vec{a}_{i_r}$  极大无关  $\Leftrightarrow \vec{b}_{i_1}, \vec{b}_{i_2}, \dots, \vec{b}_{i_r}$  极大无关;

注: 这个定理保证了行变换不改变列秩!

$$\begin{aligned}
 (1) \quad LHS &\Leftrightarrow AX = 0 \text{ 有非零解} \\
 &\xLeftrightarrow[P \text{ 可逆}]{B=PA} BX = 0 \text{ 有非零解} \\
 &\Leftrightarrow RHS.
 \end{aligned}$$

$$\begin{aligned}
 (2) \quad LHS &\Leftrightarrow \begin{cases} a_{i_1}, \dots, a_{i_r} \text{ 线性无关} \\ a_{i_1}, \dots, a_{i_r}, a_j \text{ 线性相关 } \forall j \end{cases} \\
 &\xLeftrightarrow{(1)} \begin{cases} b_{i_1}, \dots, b_{i_r} \text{ 线性无关} \\ b_{i_1}, \dots, b_{i_r}, b_j \text{ 线性相关 } \forall j \end{cases} \\
 &\Leftrightarrow RHS.
 \end{aligned}$$



# 寻找极大无关组例子

例

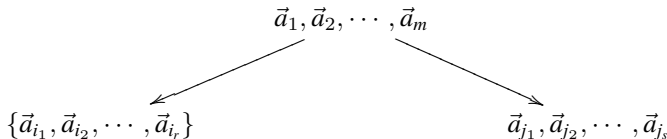
求向量组  $\vec{a}_1 = (-1, 5, 3, -2)$ ,  $\vec{a}_2 = (4, 1, -2, 9)$ ,  $\vec{a}_3 = (2, 0, -1, 4)$ ,  $\vec{a}_4 = (0, 3, 4, -5)$

$$\begin{pmatrix} -1 & 4 & 2 & 0 \\ 5 & 1 & 0 & 3 \\ 3 & -2 & -1 & 4 \\ -2 & 9 & 4 & -5 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & 4 & 2 & 0 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 5 & 24 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\xrightarrow{\text{行变换}} \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}, \{\vec{b}_1, \vec{b}_2, \vec{b}_4\}$  极大无关.

$\Rightarrow \{\vec{a}_1, \vec{a}_2, \vec{a}_3\}, \{\vec{a}_1, \vec{a}_2, \vec{a}_4\}$  极大无关.

问题: 两个极大无关组的个数是否相等?



# 等价向量组

## 定义 (等价)

称两向量组  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  和  $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_\ell$  等价, 若

- ① 任意  $i \in \{1, \dots, m\}$ ,  $\vec{a}_i$  可由  $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_\ell$  线性表示;
- ② 任意  $i \in \{1, \dots, \ell\}$ ,  $\vec{b}_i$  可由  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性表示;

此时记为  $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m\} \sim \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_\ell\}$ .

## 定理 (通过生成子空间来判定是否等价)

$$\{\vec{a}_1, \dots, \vec{a}_m\} \sim \{\vec{b}_1, \dots, \vec{b}_\ell\} \Leftrightarrow \langle \vec{a}_1, \dots, \vec{a}_m \rangle = \langle \vec{b}_1, \dots, \vec{b}_\ell \rangle.$$

注:  $\sim$  为等价关系.

## 推论

- ① 一个向量组与它的任一极大无关组等价;
- ② 任两极大无关组等价.

# 极大无关组的基本性质与秩的定义

## 定理

设  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r$  和  $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_s$  为两线性无关的向量组. 若它们相互等价, 则  $r = s$ .

证明思路:  $\begin{cases} (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r) = (\vec{b}_1, \vec{b}_2, \dots, \vec{b}_s) A_{s \times r} \\ (\vec{b}_1, \vec{b}_2, \dots, \vec{b}_s) = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r) B_{r \times s} \end{cases} \Rightarrow \begin{cases} AB = I_s \\ BA = I_r \end{cases} \Rightarrow r = s.$

## 推论

向量组  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  的任两个极大无关组中的向量个数相同. 这个数称为向量组的秩. 记为  $\text{rank}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m)$  或者  $r(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m)$ .

## 性质 (用秩判定相关性)

- ①  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性无关  $\Leftrightarrow \text{rank}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m) = m$ ;
- ②  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$  线性相关  $\Leftrightarrow \text{rank}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m) < m$ ;

# 秩与线性相关性

## 定理

若  $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_s$  可由  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r$  线性表示, 则

$$\text{rank}(\vec{b}_1, \vec{b}_2, \dots, \vec{b}_s) \leq \text{rank}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r).$$

证明思路: 不妨设  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_\ell$  和  $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_k$  为极大无关组. 则  $(\vec{b}_1, \dots, \vec{b}_k) = (\vec{a}_1, \dots, \vec{a}_\ell)A$ , 其中  $A \in \mathbb{F}^{\ell \times k}$  列满秩. 因此  $k \leq \ell$ .

## 推论

- ①  $\mathbb{F}^n$  中任意  $n+1$  个向量一定线性相关.
- ② 若  $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_s$  可由  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r$  线性表示, 则  $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r\} \sim \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_s\} \Leftrightarrow \text{rank}(\vec{b}_1, \vec{b}_2, \dots, \vec{b}_s) = \text{rank}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r)$ ;

## 推论 (用秩来判定线性方程组是否有解)

$\vec{b}$  为  $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r$  的线性组合  $\Leftrightarrow$

$$\text{rank}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r) = \text{rank}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r, \vec{b}).$$

# 向量组的秩与矩阵的秩

$$A = (a_{ij})_{n \times n} = \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{pmatrix} = (\vec{b}_1 \quad \cdots \quad \vec{b}_n)$$

我们有如下三种秩:

- ❶  $\text{rank}(A)$                       矩阵  $A$  的秩;
- ❷  $\text{rank}(\vec{a}_1, \dots, \vec{a}_m)$         矩阵  $A$  的行秩;
- ❸  $\text{rank}(\vec{b}_1, \dots, \vec{b}_n)$         矩阵  $A$  的列秩;

## 定理

秩 = 行秩 = 列秩.

证明思路: 初等变换不改变三者且对于标准形矩阵三者一致.

# 矩阵秩的一些性质

## 推论

设  $A$  为  $n$  阶方阵, 则

- ①  $A$  可逆  $\Leftrightarrow \text{rank}(A) = n \Leftrightarrow$  行向量线性无关  $\Leftrightarrow$  列向量线性无关.
- ②  $\text{rank}(A) = r \Rightarrow$  不为零的  $r$  阶子式所在的行 (列) 构成的  $A$  的行 (列) 向量的极大无关组.

## 例

$$\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B)).$$

## 定理

向量空间  $\mathbb{F}^n$  的任意子空间都可以由有限个向量生成.

证明思路: 反证. 假若子空间  $V$  不能由有限个向量生成, 则存在一系列向量

$$a_1, a_2, a_3, \cdots$$

使得  $a_i \in V \setminus \langle a_1, \cdots, a_{i-1} \rangle$ . 特别地, (由习题 15 知)

$$a_1, \cdots, a_{n+1}$$

线性无关. 矛盾!

## 推论

对于任意  $\mathbb{F}^n$  的子空间  $V$ , 存在一组线性无关的向量  $\vec{a}_1, \cdots, \vec{a}_r$  使得

$$V = \langle \vec{a}_1, \cdots, \vec{a}_r \rangle.$$

## 定义 (基)

设  $V$  为  $\mathbb{F}^n$  的子空间. 若向量组  $\vec{a}_1, \dots, \vec{a}_r$

- 线性无关, 且
- 生成子空间  $V$ ,

则称  $\vec{a}_1, \dots, \vec{a}_r$  为  $V$  的一组基. 称基中向量的个数  $r$  为子空间  $V$  的维数.

## 性质 (坐标)

设向量组  $\vec{a}_1, \dots, \vec{a}_r$  为  $V$  的一组基. 则任意  $\vec{a} \in V$  可唯一地表示为  $\vec{a}_1, \dots, \vec{a}_r$  的线性组合. 即, 存在唯一的一组数  $\lambda_1, \dots, \lambda_r \in \mathbb{F}$  使得

$$\vec{a} = \sum_{i=1}^r \lambda_i \vec{a}_i =: (\vec{a}_1, \dots, \vec{a}_r) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix}.$$

称  $(\lambda_1, \dots, \lambda_r)$  为  $\vec{a}$  在基  $\{\vec{a}_1, \dots, \vec{a}_r\}$  下的坐标.



## 定理 (扩充基)

设  $V$  为  $\mathbb{F}^n$  的  $r$  维子空间. 设  $\vec{a}_1, \dots, \vec{a}_s$  为  $V$  中一组线性无关向量. 则  $s \leq r$  且存在  $\vec{a}_{s+1}, \dots, \vec{a}_r \in V$  使得  $\vec{a}_1, \dots, \vec{a}_r$  构成  $V$  的一组基. 称  $\vec{a}_1, \dots, \vec{a}_r$  为  $\vec{a}_1, \dots, \vec{a}_s$  的一组**扩充基**.

## 性质

设  $U$  和  $V$  为  $\mathbb{F}^n$  的两个子空间.

- ① 若  $\dim(V) = r$ , 则  $V$  中的任意  $r+1$  个向量线性相关;
- ② 若  $\dim(V) = r$  且  $\vec{a}_1, \dots, \vec{a}_r \in V$  线性无关, 则  $\{\vec{a}_1, \dots, \vec{a}_r\}$  为  $V$  的一组基.
- ③ 若  $U \subseteq V$ , 则  $\dim U \leq \dim V$ .
- ④ 若  $U \subseteq V$  且  $\dim U = \dim V$ , 则  $U = V$ .

## 例 (自然基)

称  $\vec{e}_1, \dots, \vec{e}_n$  为  $\mathbb{F}^n$  的自然基. 任意向量在自然基下的向量为自身, 即

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = (\vec{e}_1, \dots, \vec{e}_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

$$\text{空间} \xrightarrow[\text{1:1}]{\text{坐标系}} \mathbb{R}^3$$

$$\text{点} \longmapsto \text{点在坐标系下的坐标}$$

推广

$$\mathbb{F}^n \text{ 的 } r \text{ 维子空间} \xrightarrow[\text{1:1}]{\text{基}} \mathbb{F}^r$$

$$\text{向量} \longmapsto \text{向量在基下的坐标}$$

设  $\vec{a}_1, \dots, \vec{a}_r$  和  $\vec{b}_1, \dots, \vec{b}_r$  为  $V$  的两组基. 设向量  $v \in V$  在两组基下的坐

标分别为  $X = \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix}$  和  $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_r \end{pmatrix}$ . 即,

$$v = (\vec{a}_1, \dots, \vec{a}_r)X = (\vec{b}_1, \dots, \vec{b}_r)Y.$$

**问题:** 如何确定  $X$  和  $Y$  之间的关系?

## 性质

设  $\vec{a}_1, \dots, \vec{a}_r$  和  $\vec{b}_1, \dots, \vec{b}_r$  为  $V$  的两组基. 则

- ① 存在唯一  $r$  阶方阵  $T$  使得

$$(\vec{b}_1, \dots, \vec{b}_r) = (\vec{a}_1, \dots, \vec{a}_r)T.$$

矩阵  $T$  称为从基  $\vec{a}_1, \dots, \vec{a}_r$  到基  $\vec{b}_1, \dots, \vec{b}_r$  的过渡矩阵.

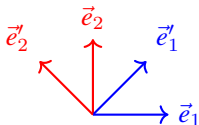
- ② 设向量  $v \in V$  在两组基下的坐标分别为  $X = (x_1, \dots, x_r)^T$  和  $Y = (y_1, \dots, y_r)^T$ . 则

$$X = TY. \quad (\text{坐标变换公式})$$

注: 过渡矩阵  $T$  总是可逆的.

例

逆时针旋转平面直角坐标系  $\theta$  角.



$$\Rightarrow (e'_1, e'_2) = (e_1, e_2) \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

$$\Rightarrow \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

例

证明  $V = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$  为  $\mathbb{R}^3$  的子空间. 求  $V$  的维数并找出其一组基.

$$\text{通解: } X = t_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

例

证明  $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0, 2x_1 - x_2 + x_3 = 3\}$  不是  $\mathbb{R}^3$  的子空间.

设  $A \in \mathbb{F}^{m \times n}$ ,  $b \in \mathbb{F}^m$  为非零向量. 则

- ①  $V := \{X \in \mathbb{F}^n \mid AX = 0\}$  为子空间;
- ②  $W := \{X \in \mathbb{F}^n \mid AX = b\}$  不是子空间.

接下来学习  $V$  和  $W$  更进一步地性质.

# 有(唯一)解的判定

## 定理

设  $A \in \mathbb{F}^{m \times n}$ ,  $b \in \mathbb{F}^m$ . 则

- ①  $AX = b$  有解  $\Leftrightarrow \text{rank}(A) = \text{rank}(A, b)$ .
- ②  $AX = b$  有唯一解  $\Leftrightarrow \text{rank}(A) = \text{rank}(A, b) = n$ .

## 例

设  $A \in \mathbb{F}^{m \times n}$ . 则

- ①  $AX = 0$  一定有解;
- ②  $AX = 0$  有非零解  $\Leftrightarrow \text{rank}(A) < n \xLeftrightarrow{\text{若 } A \text{ 为方阵}} \det(A) = 0$ .

# 齐次线性方程组的解空间

## 定义 (基础解系)

设  $A \in \mathbb{F}^{m \times n}$ . 称解空间<sup>a</sup>

$$V = \{X \in \mathbb{F}^n \mid AX = 0\}$$

的一组基为齐次线性方程组  $AX = 0$  的一个基础解系.

---

<sup>a</sup> $V$  也称为矩阵  $A$  的零空间, 记为  $N(A)$ .

## 定理 (解空间大小)

$$\dim(V) = n - \text{rank}(A).$$



# 齐次线性方程组的解空间与线性映射的核\*

给定矩阵  $A \in \mathbb{F}^{m \times n}$ . 通过  $\mathcal{A}(\vec{x}) := A\vec{x}$ , 我们可定义一个线性映射

$$\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^m.$$

## 性质

线性映射  $\mathcal{A}$  的核正好为齐次线性方程组  $AX = 0$  的解空间. 即

$$\ker \mathcal{A} = \{X \in \mathbb{F}^n \mid AX = 0\}.$$

此外, 我们还有

## 性质

- 线性映射  $\mathcal{A}$  的像由  $A$  的全体列向量生成;<sup>a</sup>
- $\dim(\ker \mathcal{A}) + \dim(\operatorname{im} \mathcal{A}) = n$ .

<sup>a</sup>称  $A$  的列向量生成的子空间为  $A$  的**列空间**. 类似地, 称  $A$  的行向量生成的子空间为  $A$  的**行空间**.

# 非齐次线性方程组的解空间

如何描述非齐次线性方程组的解空间?

$$W := \{X \in \mathbb{F}^n \mid AX = b\}$$

$W$  和  $V$  有什么联系? 基本事实:

- $\forall \alpha, \beta \in W \Rightarrow \alpha - \beta \in V$ .
- $\forall \alpha \in W, \gamma \in V \Rightarrow \alpha + \gamma \in W$ .

## 定理

若  $W$  非空, 任取  $\gamma_0 \in W$ , 记  $\gamma_0 + V := \{\gamma_0 + \alpha \mid \alpha \in V\}$ . 则

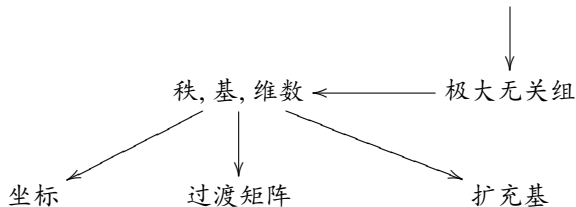
$$W = \gamma_0 + V.$$

几何解释?

# 一般线性空间

$n$  维数组空间 = 赋予了加法和数乘运算的  $n$  维数组向量组成的集合.

加法, 数乘  $\longrightarrow$  线性组合  $\longrightarrow$  线性相关 (无关), 等价



问题: 能否赋予其他集合 “+” 和 “ $\cdot$ ”, 使得其满足类似的性质和结构?

例

$\mathbb{F}_n[x] := \{a_0 + a_1x + \cdots + a_nx^n \mid a_i \in \mathbb{F} \text{ 对所有的 } i = 1, \cdots, n\}.$

● 加法

$$(a_0 + \cdots + a_nx^n) + (b_0 + \cdots + b_nx^n) := (a_0 + b_0) + \cdots + (a_n + b_n)x^n;$$

● 数乘  $\lambda(a_0 + \cdots + a_nx^n) := (\lambda a_0) + \cdots + (\lambda a_n)x^n.$

$\Rightarrow$  多项式的线性相关, 线性无关, 极大无关组, 秩, 基.

e.g.  $x^2 + 1, x^2, 1$  线性相关,  $1, x, \cdots, x^n$  为一组基.

## 例

$\mathbb{F}^{m \times n}$  数域  $\mathbb{F}$  上的全体  $m \times n$  矩阵.

- 矩阵的加法;
- 矩阵的数乘.

$\Rightarrow$  矩阵的线性相关, 线性无关, 极大无关组, 秩, 基.

e.g.  $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$  构成  $\mathbb{F}^{2 \times 2}$  的一组基.

## 例

$E_n := \{a_1x_1 + \cdots + a_nx_n = b \mid a_i, b \in \mathbb{F}\}.$

- 加法  $(a_1x_1 + \cdots + a_nx_n = b) + (a'_1x_1 + \cdots + a'_nx_n = b') := ((a_1 + a'_1)x_1 + \cdots + (a_n + a'_n)x_n = (b + b'))$ ;

- 数乘  $\lambda(a_1x_1 + \cdots + a_nx_n = b) = (\lambda a_1x_1 + \cdots + \lambda a_nx_n = \lambda b)$

$\Rightarrow$  线性方程组的线性相关, 线性无关, 极大无关组 (等价的独立方程组), 秩 (独立方程的个数), 基.

# 一般线性空间定义

设  $V$  为一个非空集,  $\mathbb{F}$  为一个数域. 若  $V$  上存在两个运算

- ① **加法**: 任意  $V$  中的有序对  $(\alpha, \beta)$ , 存在唯一的  $\gamma \in V$  与之对应. 记为  $\alpha + \beta$ . 即,  $V \times V \rightarrow V \quad (\alpha, \beta) \mapsto \alpha + \beta$ .
- ② **数乘**: 任意  $\lambda \in \mathbb{F}$ , 任意  $\alpha \in V$ , 存在唯一的  $\gamma \in V$  与之对应. 记为  $\lambda\alpha$ . 即,  $\mathbb{F} \times V \rightarrow V \quad (\lambda, \alpha) \mapsto \lambda\alpha$ .

满足如下规律 (八大公理):

- ① A1):  $\alpha + \beta = \beta + \alpha, \quad (\forall \alpha, \beta);$
- ② A2):  $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma, \quad (\forall \alpha, \beta, \gamma);$
- ③ A3):  $\exists \theta \in V \quad s.t. \quad \alpha + \theta = \alpha = \theta + \alpha, \quad (\forall \alpha),$  这个  $\theta$  也常记为  $0$ ;
- ④ A4):  $\forall \alpha \quad \exists \beta \in V \quad s.t. \quad \alpha + (\beta) = \theta = (\beta) + \alpha,$  称  $\beta$  为  $\alpha$  的负元, 记为  $-\alpha$ , 定义减法为:  $\gamma - \alpha = \gamma + (-\alpha);$
- ⑤ D1):  $(\lambda + \mu)\alpha = \lambda\alpha + \mu\alpha, \quad (\forall \lambda, \mu, \alpha);$
- ⑥ D2):  $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta, \quad (\forall \lambda, \alpha, \beta);$
- ⑦ M1):  $\lambda(\mu\alpha) = (\lambda\mu)\alpha, \quad (\forall \lambda, \mu, \alpha);$
- ⑧ M2):  $1\alpha = \alpha, \quad (\forall \alpha);$

则称  $V$  为  $\mathbb{F}$  上的线性空间.  $V$  中的元素称为向量.

八条公理保证了抽象的线性空间有好的性质和结构.

例

设  $\alpha_1, \dots, \alpha_m$  为  $\mathbb{F}$  上的线性空间  $V$  上的一组向量. 若存在一组不全为零的常数  $\lambda_1, \dots, \lambda_m$  使得

$$\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_m \alpha_m = \theta,$$

则存在  $i$  使得

$$\alpha_i = -\frac{\lambda_1}{\lambda_i} \alpha_1 - \frac{\lambda_{i-1}}{\lambda_i} \alpha_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} \alpha_{i+1} - \frac{\lambda_m}{\lambda_i} \alpha_m.$$

# 线性空间的一种判定方法

## 性质

设集合  $V$  上带有两个运算

$$\oplus: V \times V \rightarrow V \quad \text{和} \quad \odot: \mathbb{F} \times V \rightarrow V.$$

若存在正整数  $n$  和一个双射

$$\varphi: V \rightarrow \mathbb{F}^n$$

满足以下两条

①  $\varphi(a \oplus b) = \varphi(a) + \varphi(b) \quad (\forall a, b \in V);$

②  $\varphi(\lambda \odot a) = \lambda \varphi(a) \quad (\forall a \in V, \forall \lambda \in \mathbb{F},$

则  $V$  为  $\mathbb{F}$  上的线性空间.

## 例

$$\mathbb{F} = \mathbb{R}, V := \mathbb{R}^+ := \{r \in \mathbb{R} \mid r \geq 0\}, r_1 \oplus r_2 := r_1 r_2, \lambda \odot \alpha := \alpha^\lambda.$$

## 性质

- ① 零向量  $\theta$  唯一;
- ② 负向量唯一;
- ③  $0\alpha = \theta, (-1)\alpha = -\alpha, \lambda\theta = \theta$ ;
- ④  $\lambda\alpha = \theta \Rightarrow \lambda = 0$  或  $\alpha = \theta$ .

## 例

- ①  $\mathbb{F}^n$ ;
- ②  $E_n$ :  $n$  元线性方程全体;
- ③  $\mathbb{F}_n[x]$ :  $\mathbb{F}$  上次数不超过  $n$  的多项式全体;
- ④  $\mathbb{F}^{m \times n}$ ;
- ⑤  $\mathbb{F} = \mathbb{R}, V = \mathbb{C}$ ;
- ⑥  $\mathbb{F} = \mathbb{R}, V := \mathbb{R}^+ := \{r \in \mathbb{R} \mid r \geq 0\}, r_1 \oplus r_2 := r_1 r_2, \lambda \circ \alpha := \alpha^\lambda$ .
- ⑦  $C_n := \{a_0 + a_1 \cos \theta + b_1 \sin \theta + \cdots + a_n \cos(n\theta) + b_n \sin(n\theta) \mid a_i, b_i \in \mathbb{F}\}$ ;
- ⑧  $C[a, b]$  区间  $[a, b]$  上的连续函数全体;



## 例 (反例)

- ❶  $\mathbb{F} = \mathbb{R}, V = \mathbb{R}^+, \text{通常} + \times$
- ❷  $\mathbb{F} = \mathbb{R}, v_0 \in \mathbb{F}^n, V := \{v \in \mathbb{F}^n \mid v \nparallel v_0\}, \text{通常} + \times;$
- ❸  $\mathbb{F} = \mathbb{C}, V = \mathbb{R}_n[x] \text{通常} + \times;$

注: 一般线性空间没有长度和夹角等几何概念.

## 定义

设  $V$  为  $\mathbb{F}$ -线性空间,  $W$  为  $V$  的非空子集, 若

- ① 任取  $\alpha, \beta \in W$ , 都有  $\alpha + \beta \in W$ ;
- ② 任取  $\alpha \in W$  以及  $\lambda \in \mathbb{F}$ , 都有  $\lambda\alpha \in W$ .

则称  $W$  为  $V$  的**子空间**.

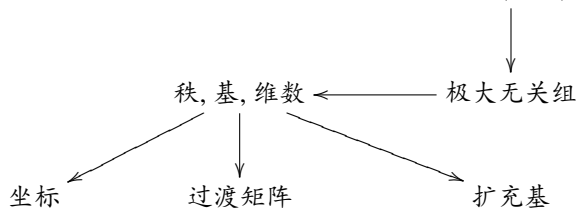
## 例

- ① 平凡子空间:  $W = \{0\}$  或者  $W = V$ .
- ② 生成子空间: 任意子集  $S \subseteq V$ ,  
 $\langle S \rangle := \{\lambda_1\alpha_1 + \cdots + \lambda_n\alpha_n \mid \lambda_1, \cdots, \lambda_n \in \mathbb{F}, \alpha_1, \cdots, \alpha_n \in S\}$ .
- ③  $\mathbb{F}_n[x], C_n \subset C[a, b]$
- ④  $W = \{P \in \mathbb{F}_n[x] \mid P(x) = P(-x)\} \subset \mathbb{F}_n[x]$ ;
- ⑤ 任意  $A \in \mathbb{F}^{n \times n}$ ,  $W := \{X \in \mathbb{F}^{n \times n} \mid AX = 0\} \subseteq \mathbb{F}^{n \times n}$ .

# 数组向量空间上的概念以及性质推广到一般线性空间

前面关于数组空间的绝大部分结论都可以平行的推广到一般线性空间上(只需要,证明过程中只涉及加法和数乘).

加法, 数乘  $\longrightarrow$  线性组合  $\longrightarrow$  线性相关(无关), 等价



# 线性相关性

## 定理 (线性相关等价刻画)

给定向量空间  $V$  上的一组向量  $\alpha_1, \alpha_2, \dots, \alpha_m$ . 则以下几条相互等价:

- ①  $\alpha_1, \alpha_2, \dots, \alpha_m$  线性相关. 即, 存在不全为零的  $\lambda_1, \lambda_2, \dots, \lambda_m$  使得  $\sum_{i=1}^m \lambda_i \alpha_i = \theta$ .
- ② 存在  $i \in \{1, \dots, m\}$  以及  $\lambda_1 \dots \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_m \in \mathbb{F}$  使得  $\alpha_i = \sum_{j \neq i} \lambda_j \alpha_j$ .
- ③ 存在  $i \in \{1, \dots, m\}$  使得  $\alpha_i \in \langle \alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_m \rangle$ .
- ④ 存在  $i \in \{1, \dots, m\}$  使得  $\langle \alpha_1, \alpha_2, \dots, \alpha_m \rangle = \langle \alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_m \rangle$ .

注: 这时候没有与“ $AX = 0$  有非零解”的等价形式.

## 定理

若  $S_1 \subseteq S \subseteq V$ , 则

- ①  $S_1$  线性相关  $\Rightarrow S$  线性相关;
- ②  $S$  线性无关  $\Rightarrow S_1$  线性无关;

# 极大无关组

## 定义 (极大线性无关组)

设  $\alpha_1, \alpha_2, \dots, \alpha_m$  为向量空间  $V$  上的一组向量. 若

- 子向量组  $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_r}$  线性无关, 且
  - 任加另一个向量  $\alpha_{i_{r+1}}$  后, 向量组  $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_{r+1}}$  线性相关,
- 则称  $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_r}$  为  $\alpha_1, \alpha_2, \dots, \alpha_m$  的极大无关组.

## 定理

若  $S_1 \subseteq S \subseteq V$ , 则以下几条等价:

- ①  $S_1$  为  $S$  的极大无关组;
- ②  $\begin{cases} S_1 \text{ 线性无关,} \\ S \text{ 可由 } S_1 \text{ 线性表示} \end{cases}$
- ③  $\begin{cases} S_1 \text{ 线性无关,} \\ S \text{ 与 } S_1 \text{ 等价} \end{cases}$
- ④  $\begin{cases} S_1 \text{ 线性无关,} \\ \langle S \rangle = \langle S_1 \rangle \end{cases}$

# 向量组的等价与秩

## 定义 (向量组等价)

两个向量组称为**等价**, 若它们可以相互线性表示.

## 定理

两个等价线性无关的向量组个数相同.

## 定义 (秩)

一个向量组的**秩**定义为其某个极大无关组中向量的个数.

## 定理

设  $S, T \subseteq V$ . 则

- ①  $S$  线性无关  $\Leftrightarrow \text{rank}(S) = \#S$ ;
- ②  $S$  线性相关  $\Leftrightarrow \text{rank}(S) < \#S$ ;
- ③  $T$  可由  $S$  线性表示  $\Rightarrow \text{rank}(T) \leq \text{rank}(S)$ ;
- ④  $S \sim T \Leftrightarrow \text{rank}(T) = \text{rank}(S)$ ;
- ⑤  $T$  可由  $S$  线性表示且  $T$  的线性无关  $\Rightarrow \#T \leq \#S$ .

## 定义

- ① 设  $V$  为线性空间. 若  $S \subseteq V$  线性无关, 且  $V = \langle S \rangle$ , 则称  $S$  为  $V$  的一组基.
- ② 设  $S$  为一组基. 若  $S$  为有限集, 则称  $V$  为有限维空间, 否则称之为无限维线性空间. 称  $S$  的个数为  $V$  的维数, 记作  $\dim_{\mathbb{F}} V$ .
- ③ 若  $S = \{\alpha_1, \dots, \alpha_n\}$  为  $V$  的一组基, 则对任意  $\alpha \in V$ , 存在唯一的  $n$  维数组向量  $(\lambda_1, \dots, \lambda_n) \in \mathbb{F}^n$  使得  $\alpha = \sum_{i=1}^n \lambda_i \alpha_i$ . 称向量  $(\lambda_1, \dots, \lambda_n)$  为  $\alpha$  在  $S$  下的坐标.

## 定理

- ①  $n$  维空间中的任意  $n+1$  个向量线性相关.
- ②  $n$  维空间中的任意  $n$  个线性无关的向量组为一组基.
- ③  $n$  维空间中的任意线性无关的向量组可扩充为一组基.

例

$\{1, \cos \theta, \sin \theta, \cos(2\theta), \sin(2\theta)\}$  为  
 $\{1, \cos \theta, \sin \theta, \cos^2(x), \sin^2(x), \cos(2\theta), \sin(2\theta)\}$  的一个极大无关组.

例

设  $T \subseteq S \subseteq V$ . 则

$$\text{rank}(S) - \text{rank}(T) \leq \#S - \#T.$$

证明: 设  $S_1$  为  $S$  的极大无关组, 则  $S_1 \cap T$  线性无关. 因此

$$\text{rank}(T) \geq \text{rank}(S_1 \cap T) = \#S_1 + \#T - \#(S_1 \cup T) \geq \text{rank}(S) + \#T - \#S.$$

例

- ①  $\dim_{\mathbb{R}} \mathbb{C} = 2$ ;  $\dim_{\mathbb{F}} \mathbb{F}[x] = \infty$ ;  $\dim_{\mathbb{F}} \mathbb{F}^{m \times n} = mn$ .
- ② 记  $W = \{f(x) \in \mathbb{F}[x] \mid \deg f(x) \leq n, f(2) = 0\}$ . 则  $W$  是  $\mathbb{F}[x]$  的子空间. 则  $W$  有一组基\_\_\_\_\_.



例

求线性空间  $\mathbb{F}_{n-1}[x]$  从基  $\{1, x, \dots, x^{n-1}\}$  到基  $\{1, x+1, \dots, (x+1)^{n-1}\}$  的过渡矩阵.

例

矩阵  $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$  构成线性空间  $\mathbb{F}^{2 \times 2}$  的一组基. 这组基到  $\mathbb{F}^{2 \times 2}$  的自然基  $E_{11}, E_{12}, E_{21}, E_{22}$  的过渡矩阵为\_\_\_\_\_.

如何从已有的子空间构造新的子空间? 子空间的交还是子空间.

## 定理

设  $W_i (i \in I)$  为  $V$  的子空间, 则  $\cap_{i \in I} W_i$  也为  $V$  的子空间. 特别地, 设  $W_1, W_2$  为  $V$  的子空间, 则  $W_1 \cap W_2$  也为  $V$  的子空间.

并集呢? 几何解释? 包含并集的最小子空间?

## 定理

设  $W_1, W_2$  为  $V$  的子空间, 则  $W_1$  与  $W_2$  的和

$$W_1 + W_2 := \{\alpha_1 + \alpha_2 \mid \alpha_1 \in W_1, \alpha_2 \in W_2\}$$

构成  $V$  的子空间, 并且是包含  $W_1 \cup W_2$  的最小子空间.

## 引理

$$\langle \alpha_1, \dots, \alpha_r \rangle + \langle \beta_1, \dots, \beta_s \rangle = \langle \alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s \rangle.$$

## 定理 (维数公式)

设  $W_1, W_2$  为  $V$  的子空间, 则

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

证明思路: 扩充  $W_1 \cap W_2$  的一组基.

记  $r = \dim W_1, s = \dim W_2$  以及  $t = \dim W_1 \cap W_2$ . 将  $W_1 \cap W_2$  一组基  $\alpha_1, \dots, \alpha_t$  扩充为

$$\begin{cases} W_1 \text{ 的一组基: } \alpha_1, \dots, \alpha_t, \beta_1, \dots, \beta_{r-t} \\ W_2 \text{ 的一组基: } \alpha_1, \dots, \alpha_t, \gamma_1, \dots, \gamma_{s-t} \end{cases}$$

则  $W_1 + W_2 = \langle \alpha_1, \dots, \alpha_t, \beta_1, \dots, \beta_{r-t}, \gamma_1, \dots, \gamma_{s-t} \rangle$ . 下面仅需证明这  $r + s - t$  个向量线性无关.

## 推论

设  $W_1, W_2$  为  $V$  的子空间. 则

- ❶  $\dim(W_1 + W_2) \leq \dim(W_1) + \dim(W_2)$ ;
- ❷  $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) \Leftrightarrow W_1 \cap W_2 = \{0\}$ ;
- ❸  $\dim(W_1 \cap W_2) \geq \dim(W_1) + \dim(W_2) - \dim(V)$ . 特别地, 若  $\dim(W_1) + \dim(W_2) > \dim(V)$ , 则  $W_1 \cap W_2 \neq \{0\}$ .

## 定义(直和)

设  $W_1, W_2$  为  $V$  的子空间. 若任意  $\alpha \in W_1 + W_2$  可**唯一地**写成

$$\alpha = \alpha_1 + \alpha_2 \quad (\alpha_1 \in W_1 \text{ \& } \alpha_2 \in W_2),$$

则称  $W_1 + W_2$  为**直和**, 记为  $W_1 \oplus W_2$ . 若  $V = W_1 \oplus W_2$ , 则称  $W_1$  为  $W_2$  的**补空间**.

注: 补空间不唯一(例子). 如何构造补空间? 扩充基!

## 定理

设  $W_1, W_2$  为  $V$  的子空间. 则以下几条等价:

- ①  $W_1 + W_2$  为直和;
- ②  $W_1 \cap W_2 = \{0\}$ ;
- ③  $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2)$ ;
- ④ 任取  $W_1$  的一组基  $\alpha_1, \dots, \alpha_r$  和  $W_2$  的一组基  $\beta_1, \dots, \beta_s$ , 则  $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s$  构成  $W_1 + W_2$  的一组基.

例

$$\mathbb{F}^{n \times n} = \{\text{对称矩阵}\} \oplus \{\text{反对称矩阵}\}.$$

例

令  $W_1 = \{(x_1, \dots, x_n) \in \mathbb{F}^n \mid x_1 + \dots + x_n = 0\}$  和  
 $W_2 = \{(x_1, \dots, x_n) \in \mathbb{F}^n \mid x_1 = x_2 = \dots = x_n\}$ , 则  
$$\mathbb{F}^n = W_1 \oplus W_2.$$

例

$$\{\text{函数}\} = \{\text{奇函数}\} \oplus \{\text{偶函数}\}.$$

## 第七章 线性映射

# 数组向量空间之间的线性变换与矩阵

给定一个数组向量空间之间的映射  $\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^m$ . 若  $\mathcal{A}$  满足

$$\begin{aligned} (\text{保持加法}) \quad & \mathcal{A}(\vec{a} + \vec{b}) = \mathcal{A}(\vec{a}) + \mathcal{A}(\vec{b}); \\ (\text{保持数乘}) \quad & \mathcal{A}(\lambda \vec{a}) = \lambda \mathcal{A}(\vec{a}) \end{aligned} \tag{*}$$

则  $\mathcal{A}$  称为从  $\mathbb{F}^n$  到  $\mathbb{F}^m$  的一个线性映射.

我们有如下对应:

从 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的全体线性映射	$\xleftrightarrow[\mathcal{A}(\vec{x})=A\vec{x}]{1:1} \mathbb{F}^{m \times n}$
---	--



## 例 (伸缩变换)

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} k_1 x \\ k_2 y \end{pmatrix} = \begin{pmatrix} k_1 & \\ & k_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

## 例 (旋转变换)

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

## 例

- $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \leftrightarrow$  恒等变换;
- $\begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} \leftrightarrow$  伸缩变换;
- $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \leftrightarrow$  反射变换;
- $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix} \leftrightarrow$  零变换;
- $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \leftrightarrow$  旋转变换;
- $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \leftrightarrow$  投射变换;

# 一般线性空间之间的线性变换

接下来将线性映射推广到一般的线性空间上.

## 定义 (线性映射, 线性变换和同构)

设  $V$  和  $W$  为两个  $\mathbb{F}$ -线性空间.(注意: 这里  $V$  和  $W$  不再要求是数组向量空间, 而只是一般的线性空间.)

① 若映射  $\mathcal{A}: V \rightarrow W$  满足

- $\mathcal{A}(P_1 + P_2) = \mathcal{A}(P_1) + \mathcal{A}(P_2)$ ;
- $\mathcal{A}(\lambda P_1) = \lambda \mathcal{A}(P_1)$ .

则称  $\mathcal{A}$  为从  $V$  到  $W$  的线性映射.

② 若  $V = W$ , 则称线性映射  $\mathcal{A}$  为  $V$  上的一个线性变换.

③ 若线性映射  $\mathcal{A}$  为双射, 则称  $\mathcal{A}$  为线性空间  $V$  到  $W$  的一个同构映射. 若两个线性空间之间存在同构映射, 则称它们互相同构.

## 性质 (同构基本性质)

- 同构映射的逆映射也是同构映射.
- 同构为等价关系.
- (同一域上的) 有限维线性空间同构当且仅当它们维数相同.

## 例

❶ 单位变换 (恒等变换):  $\varepsilon: V \rightarrow V, x \mapsto x$ ;

❷ 零映射:  $\varepsilon: V \rightarrow W, x \mapsto 0$ ;

❸ 微分变换:

$$\mathcal{A}: \mathbb{F}_n[x] \rightarrow \mathbb{F}_n[x], \quad p(x) \mapsto \frac{d}{dx}p(x).$$

❹ 积分变换:

$$\mathcal{A}: C[a, b] \rightarrow C[a, b], \quad f \mapsto \int_a^b K(x, t)f(t)dt$$

其中  $K(x, t)$  为  $[a, b] \times [a, b]$  上的实值连续函数.

❺ 线性化变换: 记  $V$  为  $[a, b]$  上的函数集合.

$$\mathcal{A}: V \rightarrow V, \quad f \mapsto (1-x)f(a) + xf(b).$$

❻ 矩阵对应的线性映射:  $A \in \mathbb{F}^{m \times n}$ .

$$\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^m, \quad x \mapsto Ax(\text{列向量}).$$

## 例

- ⑦ 投影映射:  $\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^r, (x_1, \dots, x_n) \mapsto (x_1, \dots, x_r);$
- ⑧ 嵌入映射:  $\mathcal{A}: \mathbb{F}^r \rightarrow \mathbb{F}^n, (x_1, \dots, x_r) \mapsto (x_1, \dots, x_r, 0, \dots, 0);$
- ⑨  $\mathcal{A}: \mathbb{F}^n \rightarrow \mathbb{F}^n, (x_1, x_2, \dots, x_n) \mapsto (x_n, x_{n-1}, \dots, x_1);$
- ⑩ 设  $A \in \mathbb{F}^{p \times m}$  以及  $B \in \mathbb{F}^{n \times q}$  为固定矩阵,  
$$\mathcal{A}: \mathbb{F}^{m \times n} \rightarrow \mathbb{F}^{p \times q}, \quad X \mapsto AXB.$$
- ⑪ 设  $V$  为  $\mathbb{F}$  上的线性空间,  $\alpha_1, \dots, \alpha_n$  为  $V$  中的任意  $n$  个向量.  
$$\mathcal{A}: \mathbb{F}^n \rightarrow V, \quad (x_1, \dots, x_n) \mapsto x_1\alpha_1 + \dots + x_n\alpha_n.$$

## 例 (反例)

- ①  $\mathcal{A}: \mathbb{C}^3 \rightarrow \mathbb{C}^3, (x, y, z) \mapsto (x^2, xy, z^2)$  不是  $\mathbb{C}$ -线性的.
- ②  $\mathcal{A}: \mathbb{C} \rightarrow \mathbb{C}, z \mapsto \bar{z}$  是  $\mathbb{R}$ -线性的但不是  $\mathbb{C}$ -线性的.

## 性质 (线性映射的基本性质)

设  $\mathcal{A}$  为从  $\mathbb{F}$ -线性空间  $V$  到  $\mathbb{F}$ -线性空间  $W$  的线性映射. 则

- ①  $\mathcal{A}(0) = 0$ ;
- ②  $\mathcal{A}(-\alpha) = -\mathcal{A}(\alpha), (\forall \alpha \in V)$ ;
- ③ 设  $\alpha_1, \dots, \alpha_n$  为  $V$  的一组基. 若  $\alpha = \lambda_1\alpha_1 + \dots + \lambda_n\alpha_n$ , 则

$$\mathcal{A}(\alpha) = \lambda_1\mathcal{A}(\alpha_1) + \dots + \lambda_n\mathcal{A}(\alpha_n).$$

即, 线性映射  $\mathcal{A}$  由  $\mathcal{A}(\alpha_1), \dots, \mathcal{A}(\alpha_n)$  这  $n$  个向量唯一确定.

- ④ 若  $\beta_1, \dots, \beta_m$  线性相关, 则  $\mathcal{A}\beta_1, \dots, \mathcal{A}\beta_m$  也线性相关.
- ⑤ 若  $\mathcal{A}\beta_1, \dots, \mathcal{A}\beta_m$  线性无关, 则  $\beta_1, \dots, \beta_m$  也线性无关.

# 线性映射的矩阵 \*

设  $\mathcal{A}$  为从  $n$  维  $\mathbb{F}$ -向量空间  $V$  到  $m$  维  $\mathbb{F}$ -向量空间的一个线性变换. 分别取定  $V$  和  $W$  的一组基  $\alpha_1, \dots, \alpha_n$  和  $\beta_1, \dots, \beta_m$ . 对任意  $j = 1, 2, \dots, n$ , 设  $\mathcal{A}(\alpha_j) \in W$  在基  $\beta_1, \dots, \beta_m$  下的坐标为  $(a_{1j}, \dots, a_{mj})^T$ , 即

$$\mathcal{A}(\alpha_j) = a_{1j}\beta_1 + \dots + a_{mj}\beta_m.$$

则这一式子可改写为

$$\mathcal{A}(\alpha_1, \dots, \alpha_n) := (\mathcal{A}(\alpha_1), \dots, \mathcal{A}(\alpha_n)) = (\beta_1, \dots, \beta_m)A,$$

其中  $A := (a_{ij})_{m \times n} \in F^{m \times n}$ . 称  $A$  为线性映射  $\mathcal{A}$  在基  $\alpha_1, \dots, \alpha_n$  和  $\beta_1, \dots, \beta_m$  下的矩阵.

## 例

设  $A \in \mathbb{F}^{m \times n}$ . 定义从  $\mathbb{F}^n$  到  $\mathbb{F}^m$  的线性变换

$$\mathcal{A}(X) = AX.$$

则  $\mathcal{A}$  在自然基下的矩阵为  $A$ .

证: 记  $e_1, \dots, e_n$  为  $\mathbb{F}^n$  的自然基,  $e'_1, \dots, e'_m$  为  $\mathbb{F}^m$  的自然基. 则

$$\begin{aligned}\mathcal{A}(e_1, \dots, e_n) &:= (\mathcal{A}e_1, \dots, \mathcal{A}e_n) = (Ae_1, \dots, Ae_n) = A(e_1, \dots, e_n) \\ &= AI_n = A = I_m A = (e'_1, \dots, e'_m)A.\end{aligned}$$

## 例 (从矩阵出发构造线性映射)

分别取定  $\mathbb{F}$ -线性空间  $V$  和  $W$  的一组基  $\alpha_1, \dots, \alpha_n$  和  $\beta_1, \dots, \beta_m$ . 对于给定矩阵  $B = (b_{ij})_{m \times n} \in \mathbb{F}^{m \times n}$ , 如下给出一个从  $V$  到  $W$  的映射  $\mathcal{B}$ . 对任意的  $\lambda_1 \alpha_1 + \dots + \lambda_n \alpha_n \in V$ , 定义<sup>a</sup>

$$\mathcal{B} \left( \sum_{j=1}^n \lambda_j \alpha_j \right) := \sum_{i=1}^m \left( \sum_{j=1}^n b_{ij} \lambda_j \right) \beta_i.$$

则  $\mathcal{B}$  为从  $V$  到  $W$  的一个线性映射, 且其在基  $\alpha_1, \dots, \alpha_n$  和  $\beta_1, \dots, \beta_m$  下的矩阵正好为  $B$ .

$$^a \text{这一定义也可改写为 } \mathcal{B} \left( (\alpha_1, \dots, \alpha_n) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \right) = (\beta_1, \dots, \beta_m) B \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

从而, 我们得到如下对应:

$$\left\{ \begin{array}{l} \text{从 } V \text{ 到 } W \text{ 的} \\ \text{全体线性映射} \end{array} \right\} \begin{array}{c} \xleftarrow[\mathcal{A}(\alpha_1, \dots, \alpha_n) = (\beta_1, \dots, \beta_m)A]{\begin{array}{l} \text{基 } \alpha_1, \dots, \alpha_n \\ \text{基 } \beta_1, \dots, \beta_m \\ 1:1 \end{array}} \mathbb{F}^{m \times n} \end{array}$$

# 线性映射的坐标表示 \*

## 定理

设  $A$  为线性映射  $\mathcal{A}: V \rightarrow W$  在基  $\alpha_1, \dots, \alpha_n$  和  $\beta_1, \dots, \beta_m$  下的矩阵. 若  $X$  为向量  $v \in V$  在基  $\alpha_1, \dots, \alpha_n$  下的坐标, 以及  $Y$  为  $\mathcal{A}(v)$  在基  $\beta_1, \dots, \beta_m$  下的坐标, 则

$$Y = AX.$$

换言之, 线性映射的作用可以通过对坐标左乘矩阵  $A$  实现. 即, 下图交换

$$\begin{array}{ccc} V & \xrightarrow[\text{1:1}]{\text{基 } \alpha_1, \dots, \alpha_n} & \mathbb{F}^n \\ \downarrow \scriptstyle v \mapsto \mathcal{A}(v) \quad \mathcal{A} & \mathcal{A}(\alpha_1, \dots, \alpha_n) = (\beta_1, \dots, \beta_m)A & \downarrow \scriptstyle A \quad X \mapsto AX \\ W & \xrightarrow[\text{1:1}]{\text{基 } \beta_1, \dots, \beta_m} & \mathbb{F}^m \end{array}$$

证明思路:  $\mathcal{A}(v) = \mathcal{A}((\alpha_1, \dots, \alpha_n)X) = \mathcal{A}(\alpha_1, \dots, \alpha_n) \cdot X = (\beta_1, \dots, \beta_m)A \cdot X = (\beta_1, \dots, \beta_m) \cdot AX.$



# 线性代数的核心 \*

几何 $\xrightleftharpoons[\text{基}]{\text{坐标系}}$ 代数	
线性空间	数组空间
向量	坐标
线性映射	基下矩阵
内积	矩阵
二次型	矩阵
张量	由数组组成的高维阵列
...	...

# 线性变换的矩阵

现在, 我们考虑  $W$  等于  $V$  的情形.

设  $\mathcal{A}$  为  $V$  上的线性变换. 固定  $V$  的一组基  $\alpha_1, \dots, \alpha_n$ . 则存在矩阵  $A := (a_{ij})_{n \times n}$  使得

$$\mathcal{A}(\alpha_1, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_n)A.$$

称矩阵  $A$  为线性变换  $\mathcal{A}$  在基  $(\alpha_1, \dots, \alpha_n)$  下的矩阵.

例

设  $V = \mathbb{F}^{2 \times 2}$ . 记  $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in V$ . 定义  $V$  上的线性变换  $\mathcal{A}$

$$\mathcal{A}(M) := AM.$$

求  $\mathcal{A}$  在  $e_{11}, e_{12}, e_{21}, e_{22}$  下的矩阵.

$$\left\{ \begin{array}{l} \text{从 } V \text{ 到 } V \text{ 的} \\ \text{全体线性变换} \end{array} \right\} \begin{array}{c} \xleftarrow{\text{基 } \alpha_1, \dots, \alpha_n} \\ \xrightarrow[1:1]{\mathcal{A}(\alpha_1, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_n)A} \end{array} \mathbb{F}^{n \times n}$$

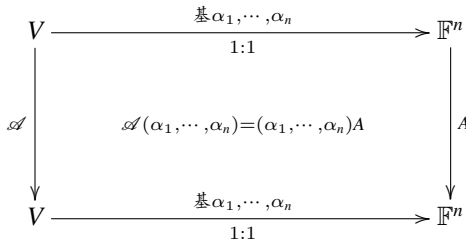
# 线性变换的坐标表示

向量  $x$  和  $\mathcal{A}x$  在同一组基下坐标之间的关系.

## 定理

设  $\mathcal{A}: V \rightarrow V$  在基  $\alpha_1, \dots, \alpha_n$  下的矩阵为  $A$ . 若向量  $x \in V$  在基  $\alpha_1, \dots, \alpha_n$  下的坐标  $X \in \mathbb{F}^n$ , 向量  $\mathcal{A}x$  在基  $\alpha_1, \dots, \alpha_n$  下的坐标为  $Y \in \mathbb{F}^n$ , 则

$$Y = AX.$$



例

记  $\alpha_1 = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$ ,  $\alpha_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$ ,  $\alpha_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ,  $\beta_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ ,  $\beta_2 = \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}$ ,  
 $\beta_3 = \begin{pmatrix} 3 \\ 0 \\ 5 \end{pmatrix}$ . 设  $\mathcal{A}$  为  $\mathbb{F}^3$  上线性变换满足  $\mathcal{A}(\alpha_i) = \beta_i$  ( $i = 1, 2, 3$ ). 求

- ①  $\mathcal{A}$  在  $\alpha_1, \alpha_2, \alpha_3$  下的矩阵;
- ②  $\mathcal{A}$  在自然基下的矩阵.

解答:

- ①  $A = (\alpha_1, \alpha_2, \alpha_3)^{-1}(\beta_1, \beta_2, \beta_3) = \begin{pmatrix} 4 & 9 & -5 \\ -10 & -23 & 15 \\ -7 & -16 & 13 \end{pmatrix};$
- ②  $B = (\beta_1, \beta_2, \beta_3)(\alpha_1, \alpha_2, \alpha_3)^{-1} = \begin{pmatrix} 3 & -20 & 11 \\ 0 & -16 & 10 \\ 5 & -15 & 7 \end{pmatrix}.$

# 矩阵对应的线性变换

设  $\alpha_1, \dots, \alpha_n$  为  $n$  维  $\mathbb{F}$ -线性空间  $V$  的一组基.

$$\begin{array}{ccc} V \text{ 上的线性变换} & \xrightarrow[\quad 1:1 \quad]{\alpha_1, \dots, \alpha_n} & \mathbb{F}^{n \times n} \\ \mathcal{A} \mapsto & \longrightarrow & \mathcal{A} \text{ 在 } \alpha_1, \dots, \alpha_n \text{ 下的矩阵} \\ ?? & \longleftarrow & B \end{array}$$

反之, 给定  $n$  阶矩阵  $B = (b_{ij})_{n \times n}$ , 我们可以如下定义一个映射 (对于任意  $\alpha = x_1\alpha_1 + \dots + x_n\alpha_n \in V$ )

$$\mathcal{B}(\alpha) := \sum_{i=1}^n \left( \sum_{j=1}^n b_{ij} x_j \right) \alpha_i = (\alpha_1, \dots, \alpha_n) B \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

## 性质

映射  $\mathcal{B}$  为  $V$  上的线性变换, 且其在  $\alpha_1, \dots, \alpha_n$  下的矩阵为  $B$ .

# 线性映射的运算 \*

设  $U, V$  和  $W$  为有限维  $\mathbb{F}$ -线性空间.

- ① (加法) 设  $\mathcal{A}$  和  $\mathcal{B}$  为从  $V$  到  $W$  的两线性映射. 对任意  $v \in V$ , 定义

$$(\mathcal{A} + \mathcal{B})(v) := \mathcal{A}(v) + \mathcal{B}(v);$$

- ② (数乘) 设  $\mathcal{A}$  为从  $V$  到  $W$  的线性映射. 对于任意  $v \in V$ , 定义

$$(\lambda \mathcal{A})(v) := \lambda \mathcal{A}(v);$$

- ③ (合成) 设  $\mathcal{A}$  为从  $V$  到  $W$  的线性映射,  $\mathcal{B}$  为从  $U$  到  $V$  的线性映射. 对于任意  $v \in V$ , 定义

$$(\mathcal{A} \circ \mathcal{B})(u) := \mathcal{A}(\mathcal{B}(u)).$$

## 性质

- ① 以上三者均为线性映射的运算. 即,  $\mathcal{A} + \mathcal{B}$ ,  $\lambda \mathcal{A}$  和  $\mathcal{A} \circ \mathcal{B}$  仍为线性变换.
- ② 若各自取定  $U, V$  和  $W$  的一组基, 设  $\mathcal{A}, \mathcal{B}$  在这些基下的矩阵分别为  $A, B$ , 则  $\mathcal{A} + \mathcal{B}$ ,  $\lambda \mathcal{A}$  和  $\mathcal{A} \circ \mathcal{B}$  在这些基下的矩阵分别为  $A + B$ ,  $\lambda A$  和  $AB$ .
- ③ 从  $V$  到  $W$  上的全体线性映射组成的集合, 记为  $\text{Hom}(V, W)$ , 在线性映射的加法和数乘下构成  $\mathbb{F}$ -线性空间.

# 多项式在线性变换处取值

设  $\mathcal{A}$  和  $\mathcal{B}$  为  $\mathbb{F}$ -线性空间  $V$  上的线性变换. 记  $\mathcal{A}^0 := \text{id} (= \varepsilon)$ . 对任意正整数  $k$ , 定义

$$\mathcal{A}^k := \underbrace{\mathcal{A} \circ \mathcal{A} \circ \cdots \circ \mathcal{A}}_{k\text{次}}.$$

对于  $f(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{F}[x]$ , 定义

$$f(\mathcal{A}) := a_0 \cdot \varepsilon + a_1\mathcal{A} + \cdots + a_n\mathcal{A}^n.$$

例

若  $\mathcal{A}$  在  $V$  的某组基  $e_1, \dots, e_n$  下的矩阵为  $A$ . 则  $f(\mathcal{A})$  在基  $e_1, \dots, e_n$  下的矩阵为  $f(A)$ .

更一般地, 记  $\exp(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$ . 定义

$$\exp(\mathcal{A}) = \varepsilon + \frac{\mathcal{A}}{1!} + \frac{\mathcal{A}^2}{2!} + \frac{\mathcal{A}^3}{3!} + \cdots.$$

性质

若  $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$ , 则  $\exp(\mathcal{A} + \mathcal{B}) = \exp(\mathcal{A})\exp(\mathcal{B})$ .

我们考虑  $W = \mathbb{F}$  的情形.

称从  $V$  到  $\mathbb{F}$  的线性映射为  $V$  上的线性函数. 记

$$V^* = \text{Hom}(V, \mathbb{F}) = \{f: V \rightarrow \mathbb{F} \mid f \text{ 为线性函数}\}.$$

为  $V$  上全体线性函数组成的集合. 由线性映射的加法和数乘, 我们有线性函数的加法和数乘

$$(f + g)(v) := f(v) + g(v)$$

$$(\lambda f)(v) := \lambda f(v)$$

## 性质

在线性函数的加法和数乘下,  $V^*$  构成  $\mathbb{F}$ -线性空间. 称之为  $V$  的对偶空间.



# 对偶空间基本性质 \*

## 性质 (对偶空间的维数与基)

- 对偶空间  $V^*$  的维数与  $V$  的维数相同.
- 若  $e_1, \dots, e_n$  为  $V$  的一组基, 则  $V^*$  存在唯一的一组基  $f_1, \dots, f_n$  满足

$$f_j(e_i) = \delta_{ij} = \begin{cases} 1, & \text{若 } i = j; \\ 0, & \text{若 } i \neq j. \end{cases}$$

称  $f_1, \dots, f_n$  为  $e_1, \dots, e_n$  的**对偶基**.

我们有如下自然的**取值映射 (evaluation map)**:

$$ev: V^* \times V \rightarrow \mathbb{F}; \quad (f, v) \mapsto f(v).$$

## 性质

$$V \cong V^{**} \quad v \mapsto ev(-, v)$$

注: 若我们不要求  $V$  为有限维的, 则映射  $V \rightarrow V^{**}$  仅为单射.

# 向量的逆变性与对偶向量的协变性 \*

## 性质 (向量的逆变性与对偶向量的协变性)

设  $e_1, \dots, e_n$  以及  $e'_1, \dots, e'_n$  为  $V$  为有限维  $\mathbb{F}$ -线性空间的两组基. 设过渡矩阵为  $A$ , 即

$$(e'_1, \dots, e'_n) = (e_1, \dots, e_n)A.$$

记这两组基的对偶基分别为  $f_1, \dots, f_n$  和  $f'_1, \dots, f'_n$ .

① (向量的逆变性) 设  $v \in V$  在两组基下坐标为  $X$  和  $X'$ . 即,

$$v = (e_1, \dots, e_n)X = (e'_1, \dots, e'_n)X'.$$

则

$$X' = A^{-1}X \quad \text{或} \quad (x'_1, \dots, x'_n) = (x_1, \dots, x_n)(A^{-1})^T.$$

② (对偶向量协变性) 设对偶向量  $f$  在两组对偶基下坐标为  $Y$  和  $Y'$ . 即,

$$f = (f_1, \dots, f_n)Y = (f'_1, \dots, f'_n)Y'.$$

则

$$Y' = A^T Y \quad \text{或} \quad (y'_1, \dots, y'_n) = (y_1, \dots, y_n)A.$$

证明思路:

$$(f'_1, \dots, f'_n) = (f_1, \dots, f_n)(A^{-1})^T.$$

# 线性映射在不同基下的矩阵与相抵关系 \*

设  $\mathcal{A}$  为从  $n$  维  $\mathbb{F}$ -向量空间  $V$  到  $m$  维  $\mathbb{F}$ -向量空间的一个线性映射. 设线性映射  $\mathcal{A}$  在基  $\alpha_1, \dots, \alpha_n$  和  $\beta_1, \dots, \beta_m$  下矩阵为  $A$ . 即

$$\mathcal{A}(\alpha_1, \dots, \alpha_n) = (\beta_1, \dots, \beta_m)A.$$

设线性映射  $\mathcal{A}$  在另外两组基  $\alpha'_1, \dots, \alpha'_n$  和  $\beta'_1, \dots, \beta'_m$  下矩阵为  $B$ . 即

$$\mathcal{A}(\alpha'_1, \dots, \alpha'_n) = (\beta'_1, \dots, \beta'_m)B.$$

## 定理

记从线性空间  $V$  的基  $\alpha_1, \dots, \alpha_n$  到基  $\alpha'_1, \dots, \alpha'_n$  的过渡矩阵为  $Q$ , 以及从线性空间  $W$  的基  $\beta_1, \dots, \beta_m$  到基  $\beta'_1, \dots, \beta'_m$  的过渡矩阵为  $P$ . 则

$$B = P^{-1}AQ.$$

$$\mathcal{A}(\alpha'_1, \dots, \alpha'_n) = \mathcal{A}(\alpha_1, \dots, \alpha_n)Q = (\beta_1, \dots, \beta_m)AQ = (\beta'_1, \dots, \beta'_m)P^{-1}AQ.$$

## 推论

线性映射在不同基下的矩阵之间相抵.

注: 反之亦然. 即, 相抵的矩阵是同一个线性映射在不同基下的矩阵.  
相抵标准形的几何意义?

# 线性变换在不同基下的矩阵与相似关系

## 定理

设线性空间  $V$  上的线性变换  $\mathcal{A}$  在两组基  $\alpha_1, \dots, \alpha_n$  和  $\beta_1, \dots, \beta_n$  下的矩阵为  $A$  和  $B$ . 设  $\alpha_1, \dots, \alpha_n$  到  $\beta_1, \dots, \beta_n$  的过渡矩阵为  $T$ . 则

$$B = T^{-1}AT.$$

## 例

设线性变换  $\mathcal{A}: \mathbb{F}^3 \rightarrow \mathbb{F}^3$  满足

$$\mathcal{A} \left( \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \left( \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) \begin{pmatrix} 4 & 9 & -5 \\ -1 & -23 & 15 \\ -7 & -16 & 13 \end{pmatrix}.$$

求  $\mathcal{A}$  在基  $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ -1 \end{pmatrix}$  下的矩阵.

$$\text{解: } T = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & 0 \\ 5 & 2 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 & -2 \\ 0 & 1 & 0 \\ 1 & 2 & -1 \end{pmatrix}, C = T^{-1} \begin{pmatrix} 4 & 9 & -5 \\ -1 & -23 & 15 \\ -7 & -16 & 13 \end{pmatrix} T = \begin{pmatrix} -10 & 2 & 3 \\ 10 & 4 & 10 \\ -2 & 1 & 0 \end{pmatrix}.$$

# 矩阵的相似

## 定义 (相似)

设  $A, B \in \mathbb{F}^{n \times n}$  为两个  $n$  阶方阵. 若存在可逆阵  $T \in \mathbb{F}^{n \times n}$  使得  $B = T^{-1}AT$ , 则称  $A$  与  $B$  相似. 记为  $A \sim B$ .

## 性质

相似为等价关系.

根据相似关系, 将全体  $n$  阶矩阵分为若干类. 相似类, 代表元.

## 定理

一个线性变换在不同基下矩阵相似. 反之任意属于该相似类的矩阵, 均为该线性变换在某组基下的矩阵.

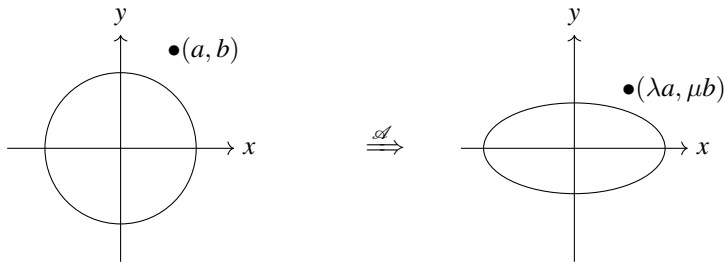
证明思路: 设  $\mathcal{A}(\alpha_1, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_n)A$ . 若  $B = T^{-1}AT$ , 记  $(\beta_1, \dots, \beta_n) = (\alpha_1, \dots, \alpha_n)T$ , 则  $\mathcal{A}(\beta_1, \dots, \beta_n) = (\beta_1, \dots, \beta_n)B$ .  
相似不变量.e.g. 行列式, 秩.

类比于相抵关系, 对相似关系我们有如下基本问题:

- ① 两个矩阵相似充要条件
- ② 相似等价类中的最简代表元

# 伸缩变换与特征向量

设线性变换  $\mathcal{A}$  在基  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  下的矩阵为  $\begin{pmatrix} \lambda & \\ & \mu \end{pmatrix}$ .



问题:

- 1 是否任意线性变换都由某组基下的各个方向伸缩给出?(答: 错误.)
- 2 哪些线性变换可以由某组基下的伸缩给出? 即, 哪些线性变换在合适的基下为对角阵?

为了研究回答这些问题, 我们需要引入特征值和特征向量等概念.

# 线性变换的特征值, 特征向量以及特征子空间

定义 (线性变换的特征值, 特征向量以及特征子空间)

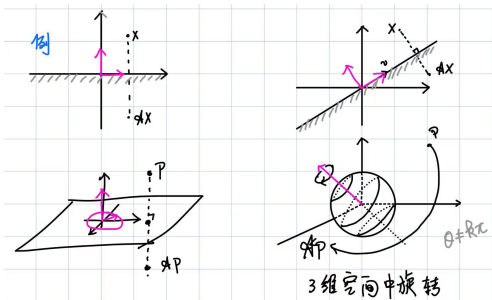
设  $\mathcal{A}$  为  $n$  维  $\mathbb{F}$ -向量空间  $V$  上的线性变换. 若存在  $\lambda \in F$  以及非零向量  $\alpha \in V$ , 使得

$$\mathcal{A}\alpha = \lambda\alpha,$$

则称  $\lambda$  为  $\mathcal{A}$  的一个特征值, 称  $\alpha$  为属于特征值  $\lambda$  的一个特征向量. 称

$$V_{\mathcal{A}}(\lambda) := \ker(\mathcal{A} - \lambda\varepsilon) = \{\alpha \in V \mid \mathcal{A}\alpha = \lambda\alpha\}$$

为属于  $\lambda$  的特征子空间.



# 如何求特征值和特征向量

**思路:** 通过代数 (矩阵) 来解决几何问题 (特征值, 特征向量).

设  $\alpha_1, \dots, \alpha_n$  为  $\mathbb{F}$ -线性空间  $V$  的一组基.

$$\boxed{\begin{array}{ccc} \left\{ \begin{array}{l} V \text{ 上的全体} \\ \text{线性变换} \end{array} \right\} & \begin{array}{c} \xleftarrow{\text{基 } \alpha_1, \dots, \alpha_n} \\ \xrightarrow[1:1]{\mathcal{A}(\alpha_1, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_n)A} \end{array} & \mathbb{F}^{n \times n} \end{array}}$$

**定义 (矩阵的特征值, 特征向量以及特征子空间)**

设  $A$  为  $\mathbb{F}$  系数的  $n$  阶方阵. 若存在  $\lambda \in F$  及 **非零** 数组向量  $x \in \mathbb{F}^n$ , 使得

$$Ax = \lambda x,$$

则称  $\lambda$  为方阵  $A$  的**特征值**, 称  $x$  为属于特征值  $\lambda$  的一个**特征向量**. 称

$$V_A(\lambda) := \{x \in \mathbb{F}^n \mid Ax = \lambda x\}$$

为属于  $\lambda$  的**特征子空间**.



# 如何求特征值和特征向量

求  $\mathcal{A}$  的特征值和特征向量可以转换为求  $A$  的特征值和特征向量.

## 性质

设  $\alpha_1, \dots, \alpha_n$  为  $\mathbb{F}$ -线性空间  $V$  的一组基. 设  $V$  上的线性变换在基  $\alpha_1, \dots, \alpha_n$  下矩阵为  $A$ . 则

- ①  $\mathcal{A}$  与  $A$  有相同的特征值;
- ② 若  $\lambda$  为  $\mathcal{A}$  的特征值, 则  $V_{\mathcal{A}}(\lambda) = \{(\alpha_1, \dots, \alpha_n)x \mid x \in V_A(\lambda)\}.$

证明思路:  $x \in V_A(\lambda) \Leftrightarrow Ax = \lambda x \Leftrightarrow \mathcal{A}((\alpha_1, \dots, \alpha_n)x) = \lambda(\alpha_1, \dots, \alpha_n)x \Leftrightarrow (\alpha_1, \dots, \alpha_n)x \in V_{\mathcal{A}}(\lambda).$

## 性质

设  $\lambda$  为方阵  $A$  的特征值. 则

- ①  $f(\lambda)$  为  $f(A)$  的特征值; 特别地,  $\lambda^k$  为方阵  $A^k$  的特征值;
- ②  $\lambda$  为方阵  $A^T$  的特征值;
- ③ 若  $\lambda \neq 0$ , 则  $\frac{1}{\lambda} \det(A)$  为  $A^*$  的特征值. 特别地, 若  $A$  可逆, 则  $\frac{1}{\lambda}$  为  $A^{-1}$  的特征值.
- ④ 若  $A$  为实矩阵且  $AA^T = 1$ , 则  $|\lambda| = 1$ .

注: 实矩阵的特征值不一定还是实数!

# 矩阵特征值和特征向量一般求解方法

例

求矩阵  $A = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$  的特征值和特征向量.

解:  $A \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \begin{pmatrix} 1-\lambda & 1 \\ 1 & 3-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \Rightarrow \det \begin{pmatrix} 1-\lambda & 1 \\ 1 & 3-\lambda \end{pmatrix} = 0 \Rightarrow \lambda = 2 \pm \sqrt{2}$ .

解方程组:  $A \begin{pmatrix} x \\ y \end{pmatrix} = 2 \pm \sqrt{2} \begin{pmatrix} x \\ y \end{pmatrix}$  得  $\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 1 \\ 1 \pm \sqrt{2} \end{pmatrix}$ .

对于方阵  $A$  的特征值的计算, 我们有如下一般方法:

$\lambda_0$  为  $A$  的特征值  $\Leftrightarrow \exists x \neq 0 \quad s.t. Ax = \lambda_0 x$

$\Leftrightarrow$  线性方程组  $(\lambda_0 I_n - A)X = 0$  有非零解

$\Leftrightarrow \det(\lambda_0 I_n - A) = 0$

即,  $\lambda_0$  为  $A$  的特征值当且仅当  $\lambda_0$  为多项式  $\det(\lambda I_n - A)$  的根. 称行列式  $\det(\lambda I - A)$  为  $A$  的特征多项式. 记为  $P_A(\lambda)$ .

注:

- ① 数域  $\mathbb{F}$  上的多项式在  $\mathbb{F}$  上不一定有根!
- ② (代数基本定理) 复数域  $\mathbb{C}$  上的非常值多项式在  $\mathbb{C}$  上一定有根.

在计算特征值和特征向量时, 总是假设  $\mathbb{F} = \mathbb{C}$ .

# 矩阵特征值和特征向量一般求解方法

设  $A$  为  $n$  阶复矩阵. 求解过程分为如下两步:

- ① 求  $P_A(\lambda) = \det(\lambda I_n - A)$ , 并做分解

$$P_A(\lambda) = (\lambda - \lambda_1)^{n_1} \cdot (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}$$

其中  $\lambda_i \in \mathbb{C}$ ,  $n_i \geq 1$ ,  $n_1 + n_2 + \cdots + n_s = n$ . 我们称  $n_i$  为特征值  $\lambda_i$  的**重数**.

- ② 给定  $i = 1, \dots, s$ . 解齐次线性方程组  $(\lambda_i I_n - A)X = 0$  得解空间  $V_A(\lambda_i)$  或者基础解系.

例

求矩阵  $A = \begin{pmatrix} 3 & -1 & 2 \\ 2 & 0 & -2 \\ 2 & -1 & -1 \end{pmatrix}$  的特征值和特征向量.

答案:  $P_A(\lambda) = \lambda(\lambda - 1)^2$ .

- $(0 \cdot I - A)\vec{x} = 0 \Rightarrow \vec{x} = c_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix};$
- $(1 \cdot I - A)\vec{x} = 0 \Rightarrow \vec{x} = c_2 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}.$

# 矩阵特征值和特征向量一般求解方法

例

设  $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ . 设  $\mathcal{A}$  为  $V = \mathbb{F}^{2 \times 2}$  上的线性变换满足  $\mathcal{A}(M) = AM$ . 求  $\mathcal{A}$  的特征值和特征向量.

答案:  $\mathcal{A}(E_{11}, E_{12}, E_{21}, E_{22}) = (E_{11}, E_{12}, E_{21}, E_{22})B$ , 其中

$$B = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}. \Rightarrow P_{\mathcal{A}}(\lambda) = P_B(\lambda) = (\lambda - 1)^2(\lambda - 2)^2.$$

$$\bullet (I_4 - B)\vec{x} = 0 \Rightarrow \text{基础解系 } \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \text{特征子空间 } \langle E_{11}, E_{12} \rangle$$

$$\bullet (2I_4 - B)\vec{x} = 0 \Rightarrow \text{基础解系 } \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \text{特征子空间}$$

$$\langle E_{11} + E_{21}, E_{12} + E_{22} \rangle.$$

# 相似不变量

可以用矩阵来定义线性变换的哪些不变量 (即, 不依赖于基的选取的量).

## 性质

相似的矩阵具有相同的特征多项式和特征值.

证明: 设  $B = T^{-1}AT$ . 则

$$P_B(\lambda) = \det(\lambda I - T^{-1}AT) = \det(T^{-1}) \det(\lambda I - A) \det(T) = P_A(\lambda).$$

即, 可以用矩阵来定义线性变换的特征多项式和特征值. 有没有其它相似不变量呢?

## 定理

设  $A$  为  $n$  阶复矩阵的  $n$  个特征值分别为  $\lambda_1, \dots, \lambda_n$  (可能有重复的). 则

- ①  $\operatorname{tr}(A) = \sum_{i=1}^n \lambda_i$ ;
- ②  $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$ .

证明思路: 展开  $P_A(\lambda)$  并使用根与系数之间的关系.

## 推论

一个  $n$  阶方阵可逆当且仅当其  $n$  个特征值均不为零.

# 相似不变量的应用

## 例

设  $n$  阶方阵  $A$  的全体特征值为  $\lambda_1, \lambda_2, \dots, \lambda_n$ .

- 求  $I + A$  的全体特征值和行列式;
- 更一般地, 求矩阵  $f(A)$  的全体特征值.

解题思路:  $P_{I+A}(\lambda) = P_A(\lambda - 1)$ . 设  $\lambda - f(x) = a_0 \prod_{i=1}^d (\alpha_i - x)$ . 则

$$\begin{aligned}\det(\lambda I - f(A)) &= \det\left(a_0 \prod_{i=1}^d (\alpha_i I - A)\right) = a_0^n \prod_{i=1}^d \det(\alpha_i I - A) \\ &= a_0^n \prod_{i=1}^d \prod_{j=1}^n (\alpha_i - \lambda_j) = \prod_{j=1}^n \left(a_0 \prod_{i=1}^d (\alpha_i - \lambda_j)\right) = \prod_{j=1}^n (\lambda - f(\lambda_j))\end{aligned}$$

## 例

若  $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & x \end{pmatrix}$  和  $B = \begin{pmatrix} y & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$  相似. 求  $x, y$ .

提示: 相似不变量.

# 矩阵的相似对角化

现在我们回到原始问题: 哪些线性变换由伸缩给出?

## 定义 (可对角化)

相似于对角矩阵的方阵称为**可对角化**的. 我们称一个线性变换为**可对角化**的, 若它在某组基下的矩阵可对角化.

**问题:** 如何判断一个线性变换是否可对角化.

## 例 (不是所有的线性变换和矩阵都可对角化)

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

# 可对角化的判定准则

若  $A = T \text{diag}(\lambda_1, \dots, \lambda_n) T^{-1}$ . 记  $T = (\eta_1, \dots, \eta_n)$ . 则

- ①  $\eta_1, \dots, \eta_n$  线性无关,
- ②  $A\eta_i = \lambda_i \eta_i$ .

定理 (利用特征向量来判定)

方阵  $A$  可对角化  $\Leftrightarrow A$  有  $n$  个线性无关的特征向量.

有没有简单一点的办法来判断呢?

有时候, 可以用相似不变量来判断.(不要求特征向量!)

引理

属于不同特征值的特征向量线性无关. 即, 若  $\lambda_1, \dots, \lambda_k$  为方阵  $A$  的两两不同的特征值,  $\eta_i$  为属于  $\lambda_i$  的特征向量, 则  $\eta_1, \dots, \eta_k$  线性无关.

证明思路: 范德姆行列式不为零!

推论 (利用特征值来判定)

若  $n$  阶方阵有  $n$  个两两不同的特征值, 则  $A$  可对角化.

优点: 不要求求解特征向量; 缺点: 不是充要条件, 有重根咋办?



一般地, 我们可以用特征值的重数来判定是否可对角化.

## 定义 (代数重数, 几何重数)

设  $A$  为  $n$  阶复矩阵的特征多项式有分解

$$P_A(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}.$$

称  $n_i$  为  $\lambda_i$  的**代数重数**, 称子空间  $V_A(\lambda_i)$  的维数为  $\lambda_i$  的**几何重数**.

## 定理 (利用特征值的重数来判定)

- ①  $1 \leq \text{几何重数} \leq \text{代数重数}$ .
- ②  $A$  可对角化  $\Leftrightarrow$  每个特征值的几何重数与代数重数相等.

**证明思路:** 将  $V_A(\lambda_i)$  的一组基扩充为整个空间的一组基.

例

什么时候  $A = \begin{pmatrix} 0 & 0 & x \\ 1 & 1 & y \\ 1 & 0 & 0 \end{pmatrix}$  可对角化?

解题思路:  $P_A(\lambda) = (\lambda - 1)(\lambda^2 - x)$ .

- $x \neq 0, 1 \Rightarrow$  三个不同特征值.
- $x = 0 \Rightarrow 0$  的代数重数为 2 而几何重数为 1.
- $x = 1 \Rightarrow -1$  的重数为 1. 1 的代数重数为 2, 几何重数为 2 ( $y = -1$ ) 或者 1 ( $y \neq -1$ ).

# 极小多项式\*

定理 (哈密尔顿-凯莱定理)

$$P_A(A) = 0.$$

特别地, 任意方阵  $A$  都存在零化多项式. 称次数最小的首一零化多项式为  $A$  的**极小多项式**. (注: 极小多项式为相似不变量.)

性质

- ① 极小多项式整除其他任意零化多项式. 特别地, 极小多项式为特征多项式的因子.
- ② 一个数为特征值当且仅当其为极小多项式的根.

定理 (利用极小多项式)

方阵可对角化当且仅当其极小多项式无重根.

推论 (利用零化多项式)

方阵可对角化当且仅当其存在无重根的零化多项式.

例: 设方阵  $A$  满足  $A^2 = A$ ,  $A^2 = I$  或  $A^3 = A$  等. 则  $A$  可对角化.

# 相似上三角化

不是每个矩阵都可以对角化, 但总是可上三角化!

## 定理

- ① 任意  $n$  阶复方阵  $A$  都相似于一个上三角形矩阵  $J$ , 且
- ②  $J$  的对角线上的元素为  $A$  的全体特征值.

证明思路: 将  $A$  的一个特征向量  $\xi_1$  扩充为  $\mathbb{C}^n$  的一组基  $\xi_1, \dots, \xi_n$ , 并记组成的可逆矩阵  $(\xi_1, \dots, \xi_n)$  为  $P$ . 则

$$AP = P \begin{pmatrix} \lambda & * \\ & A_{22} \end{pmatrix}.$$

对矩阵阶数归纳, 不妨设  $A_{22} = Q \begin{pmatrix} \lambda_2 & * \\ & \ddots \\ & & \lambda_n \end{pmatrix} Q^{-1}$ . 则

$$A = P \begin{pmatrix} 1 & \\ & Q \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & * \\ & \ddots \\ & & \lambda_n \end{pmatrix} \left( P \begin{pmatrix} 1 & \\ & Q \end{pmatrix} \right)^{-1}.$$

注:  $J$  的取法不唯一. (例:  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  和  $\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ .)

例: 若复二阶矩阵  $A$  满足  $A^2 = 0$ , 则  $A = 0$  或者  $A$  相似于  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ .

相似等价类中的最简代表元是什么？为了回答这一问题，我们需要引进：

定义 (若当块, 若当矩阵)

① 给定任意正整数  $m$  和复数  $\lambda$ . 称矩阵  $J_m(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$

为若当块.

② 称由若当块组成的对角分块矩阵

$$J = \begin{pmatrix} J_{m_1}(\lambda_1) & & \\ & J_{m_2}(\lambda_2) & \\ & & \ddots \\ & & & J_{m_s}(\lambda_s) \end{pmatrix} \text{ 为若当矩阵.}$$

# 相似等价类中的最简代表元(若当标准形)\*

## 定理

- ① 任意复方阵都与某若当矩阵相似, 并且
- ② 不计若当块排序下, 这个若当矩阵取法唯一. 称之为给矩阵的**若当标准形**

## 例

设  $\lambda$  为 5 阶复矩阵  $A$  的 5 重特征值. 请分析  $A$  的若当标准形.

## 例

求矩阵  $A = \begin{pmatrix} 3 & -2 & 1 \\ 2 & -2 & 2 \\ 3 & -6 & 5 \end{pmatrix}$  的若当标准形.

$$T^{-1}AT = \begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 2 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 0 \\ 3 & 0 & -1 \end{pmatrix}.$$

# 若当标准形在求解常微分方程组上的一个应用 \*

例

设  $x, y, z$  都是  $t$  的函数, 求解常微分方程组

$$\begin{cases} \frac{dx}{dt} = 3x - 2y + z \\ \frac{dy}{dt} = 2x - 2y + 2z \\ \frac{dz}{dt} = 3x - 6y + 5z \end{cases}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 0 \\ 3 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix}.$$





## 第八章 内积空间及变换

三维空间中有距离, 夹角等概念  $\xrightarrow{\text{坐标系}}$  三维数组向量空间  $\mathbb{R}^n$  上向量的模长和夹角.

比较容易地推广到  $\mathbb{R}^n$ :  $|(x_1, \dots, x_n)| := \sqrt{x_1^2 + \dots + x_n^2}$ .

问题: 如何推广到抽象的线性空间上呢?

线性空间	$\xleftrightarrow[\text{1:1}]{\text{基 } \alpha_1, \dots, \alpha_n}$	$\mathbb{R}^n$
向量 $v$	$\longleftrightarrow$	坐标
线性变换 $\mathcal{A}$	$\longleftrightarrow$	$A$
??	$\longleftrightarrow$	长度, 夹角
$\mathcal{A}$ (保持长度)	$\longleftrightarrow$	??
$\mathcal{A}$ (??)	$\longleftrightarrow$	转置矩阵
$\mathcal{A}$ (??)	$\longleftrightarrow$	对称矩阵

回顾  $\mathbb{R}^3$  上的模长与夹角公式:

$$|a| = \sqrt{(a, a)}$$

$$\theta = \arccos \frac{(a, b)}{\sqrt{(a, a)(b, b)}}$$

容易看出模长和夹角可以由内积完全确定. 反之, 内积可由模长唯一确定

$$(a, b) = \frac{|a+b|^2 - |a|^2 - |b|^2}{2}.$$

总结  $\mathbb{R}^3$  上内积有如下基本性质:

- ① 对称性
- ② 线性性
- ③ 正定性

为了在一般的  $\mathbb{R}$ -线性空间上定义度量, 我们需要引入一般空间上的内积.

# 欧氏空间的定义

## 定义 (内积, 欧氏空间)

设  $V$  为有限维  $\mathbb{R}$ -线性空间. 若存在一个映射  $(-, -): V \times V \rightarrow \mathbb{R}$  满足:

- ① 对称性:  $(a, b) = (b, a)$ ;
- ② 双线性:  $(\lambda a, b) = \lambda(b, a)$ ,  $(a + b, c) = (a, c) + (b, c)$ ;
- ③ 正定性:  $(a, a) \geq 0$ , 且等号成立当且仅当  $a = 0$ ,

则称  $(a, b)$  为  $a$  和  $b$  的**内积**. 带内积的  $\mathbb{R}$ -线性空间称为**欧氏空间**(或,**欧几里得空间**).

欧式空间 =  $\mathbb{R}$ -线性空间 + 内积;

$\mathbb{R}$ -线性空间 = 集合 + 加法 + ( $\mathbb{R}$ ) 数乘

注: 内积  $(-, -)$  对第二个分量也是线性的.

## 性质 (内积的基本性质)

- ①  $\left(\sum_{i=1}^k \lambda_i a_i, \sum_{j=1}^m \mu_j b_j\right) = \sum_{i=1}^k \sum_{j=1}^m \lambda_i \mu_j (a_i, b_j)$ ;
- ②  $(a, 0) = 0$ ;
- ③ 内积结构由一组基之间的内积完全确定.

# 模长与夹角

为了定义夹角, 我们需要引入:

## 定理 (Cauchy-Schwarz 不等式)

设  $V$  为欧式空间. 对于任意  $a, b \in V$ , 我们有

$$|(a, b)| \leq \sqrt{(a, a) \cdot (b, b)}.$$

证明思路: 考虑二次多项式  $(xa + b, xa + b)$  的判别式.  
可以定义夹角了!(不确定原理.)

## 定义 (长度, 夹角)

设  $V$  为欧氏空间. 对于任意向量  $a, b \in V$ ,

- ① 称  $|a| := \sqrt{(a, a)}$  为  $a$  的**长度**(或,**模长**);
- ② 称  $|a - b|$  为  $a$  和  $b$  之间的**距离**. 记为  $d(a, b)$ .
- ③ 若  $a, b$  不为零, 定义  $a$  和  $b$  之间的**夹角**为  $\theta = \arccos \frac{(a, b)}{|a| \cdot |b|}$ . 特别地, 当  $(a, b) = 0$  时, 称  $a$  与  $b$  **正交**或**垂直**, 记为  $a \perp b$ .
- ④ 称  $a$  为**单位向量**, 若  $|a| = 1$ . 若  $b \neq 0$ , 称  $\frac{1}{|b|}b$  为  $b$  的**单位化**.

## 性质

- ① 对称性:  $d(a, b) = d(b, a)$ ;
- ② 正定性:  $d(a, b) \geq 0$ , 等号成立当且仅当  $a = b$ ;
- ③ 三角不等式:  $d(a, c) \leq d(a, b) + d(b, c)$ ;
- ④ 勾股定理:  $a \perp b$  当且仅当  $|a|^2 + |b|^2 = |a + b|^2$ ;
- ⑤ 菱形定理: 若  $|a| = |b|$ , 则  $(a + b) \perp (a - b)$ .

# 内积的基本例子

例

设  $V = \mathbb{R}^n$ , 对于任意两向量  $a = (a_1, \dots, a_n)$  和  $b = (b_1, \dots, b_n)$ , 定义

$$(a, b) = a_1 b_1 + \dots + a_n b_n.$$

则  $(-, -)$  为  $V$  上的一个内积. 这一内积给出通常的模长和夹角. 此时 Cauchy-Schwarz 公式为

$$\left( \sum_{i=1}^n a_i b_i \right)^2 \leq \left( \sum_{i=1}^n a_i^2 \right) \cdot \left( \sum_{i=1}^n b_i^2 \right).$$

例

类似的  $(a, b) = a_1 b_1 + 2a_2 b_2 + \dots + na_n b_n$  也可以定义  $V = \mathbb{R}^n$  上的一个内积.

# 内积的基本例子

构造一个内积使得给定的一组基均为单位向量且两两垂直.

例

设  $\alpha_1, \dots, \alpha_n$  为  $n$  维  $\mathbb{R}$ -线性空间  $V$  上的一组基. 对于任意向量  $\alpha, \beta \in V$ , 定义

$$(\alpha, \beta) = a_1 b_1 + a_2 b_2 + \dots + a_n b_n,$$

其中  $(a_1, a_2, \dots, a_n)$  和  $(b_1, b_2, \dots, b_n)$  分别为  $\alpha$  和  $\beta$  在基  $\alpha_1, \dots, \alpha_n$  下的坐标. 则  $(-, -)$  为  $V$  上的内积. 在此内积下  $\alpha_1, \dots, \alpha_n$  均为单位向量且两两垂直. 即,

$$(\alpha_i, \alpha_j) = \delta_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

任意有限维实线性空间一定存在内积. 注: 任意内积均为这一形式!



在无穷维空间上也可类似的定义内积.

## 例 (无限维空间上的内积)

对应任意  $[a, b]$  上的连续函数  $f, g \in C[a, b]$ , 定义

$$(f, g) := \int_a^b f(x)g(x)dx.$$

类似地, 定义连续函数的模长, 夹角和正交等概念. 此时 Cauchy-Schwarz 公式为

$$\left( \int_a^b f(x)g(x)dx \right)^2 \leq \left( \int_a^b f(x)^2 dx \right) \cdot \left( \int_a^b g(x)^2 dx \right).$$

当  $a = -\pi, b = \pi$  时, 三角函数  $1, \cos x, \sin x, \cos 2x, \sin 2x, \dots$  在这一内积下两两正交.

我们称带内积完备的  $\mathbb{R}$ -线性空间为 Hilbert 空间.

# 内积的矩阵表示

设  $\alpha_1, \dots, \alpha_n$  为欧氏空间  $V$  的一组基. 任取  $\alpha = a_1\alpha_1 + \dots + a_n\alpha_n$ ,  $\beta = b_1\alpha_1 + \dots + b_n\alpha_n \in V$ , 则

$$(\alpha, \beta) = \sum_{i=1}^n \sum_{j=1}^n a_i b_j (\alpha_i, \alpha_j). \quad (1)$$

因此内积  $(-, -)$  由值  $g_{ij} := (\alpha_i, \alpha_j) (1 \leq i, j \leq n)$  唯一确定. 记

$$G := (g_{ij})_{n \times n}.$$

称  $G$  为内积  $(-, -)$  在基  $\alpha_1, \dots, \alpha_n$  下的度量矩阵. 记

$$x = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, y = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

则(1)可简写为

$$(\alpha, \beta) = x^T G y.$$

通过度量矩阵我们有如下映射

$$\begin{aligned} V \text{ 上的内积} &\xrightarrow{\text{基 } \alpha_1, \dots, \alpha_n} \mathbb{R} \text{ 上的 } n \text{ 阶矩阵} \\ (-, -) &\longmapsto G = \left( (\alpha_i, \alpha_j) \right)_{n \times n} \end{aligned}$$

问题: 哪些矩阵落在这个映射的像集里面? 即, 哪些矩阵能够成为某个内积的度量矩阵?

### 性质 (度量矩阵的基本性质)

- ① 设  $G$  为  $V$  上某内积  $(-, -)$  在基  $\alpha_1, \dots, \alpha_n$  下的度量矩阵. 则
  - $G$  为实对称矩阵;
  - 对任意  $x \in \mathbb{R}^n$ , 都有  $x^T G x \geq 0$ , 且等号成立当且仅当  $x = 0$ .称满足如上性质的矩阵为**正定矩阵**. 因此内积的度量矩阵为正定矩阵.
- ② 反之, 对于任意给定的正定矩阵  $G$ , 通过
$$(\alpha, \beta) := x^T G y$$
可以构造出  $V$  上的一个内积, 其中  $x, y$  为  $\alpha$  和  $\beta$  的坐标.

问题:

- ① 不同基下度量矩阵之间的关系?
- ② 度量矩阵的最简形式, 在哪组基下度量矩阵为最简单形式?

性质

设  $P$  为欧氏空间的两组基  $\alpha_1, \dots, \alpha_n$  和  $\eta_1, \dots, \eta_n$  为之间的过渡矩阵

$$(\eta_1, \dots, \eta_n) = (\alpha_1, \dots, \alpha_n)P.$$

设内积在  $\alpha_1, \dots, \alpha_n$  和  $\eta_1, \dots, \eta_n$  下的度量矩阵为  $G$  和  $\overline{G}$ . 即,

$$G = \left( (\alpha_i, \alpha_j) \right)_{n \times n}, \quad \overline{G} = \left( (\eta_i, \eta_j) \right)_{n \times n}.$$

则

$$\overline{G} = P^T G P.$$

证明思路:

$$\overline{G}_{(i,j)} = (\eta_i, \eta_j) = \left( \sum_{s=1}^n p_{si} \alpha_s, \sum_{t=1}^n p_{tj} \alpha_t \right) = \sum_{s=1}^n \sum_{t=1}^n p_{si} G_{(s,t)} p_{tj} = (P^T G P)_{(i,j)}.$$

# 相合关系及其基本性质

## 定义 (相合)

称两个矩阵  $G$  和  $\overline{G}$  相合, 若存在可逆阵  $P$  使得

$$\overline{G} = P^T G P.$$

## 性质

- ① 内积在不同基下的度量矩阵相合;
- ② 相合为等价关系.

实对称矩阵的相合分类, 以及相合标准形 (第八章).

# 标准正交基

度量矩阵的最简形式? 为了回答这一问题, 我们需要引入标准正交基.

## 定义 (标准正交基)

设  $V$  为  $n$  维欧氏空间.

- ① 称由一组两两正交的非零向量为 **正交向量组**;
- ② 称由正交向量组构成的基为 **正交基**;
- ③ 称由单位向量组成的正交基为 **标准正交基**.

## 例

设  $(-, -)$  为  $\mathbb{R}^n$  上的一内积.

- ① 若  $((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sum_{i=1}^n x_i y_i$ . 则自然基  
$$e_1, \dots, e_n$$

为一组标准正交基.

- ② 若  $((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sum_{i=1}^n i x_i y_i$ . 则  
$$e_1, \frac{e_2}{\sqrt{2}} \cdots, \frac{e_n}{\sqrt{n}}$$

为一组标准正交基.

# 标准正交基

度量矩阵的最简形式? 为了回答这一问题, 我们需要引入标准正交基.

## 定义 (标准正交基)

设  $V$  为  $n$  维欧氏空间.

- ① 称由一组两两正交的非零向量为**正交向量组**;
- ② 称由正交向量组构成的基为**正交基**;
- ③ 称由单位向量组成的正交基为**标准正交基**.

## 性质

正交向量组线性无关.

## 定理 (标准正交基使度量矩阵最简)

设  $V$  为  $n$  维欧氏空间. 设  $G$  为内积在基  $\alpha_1, \dots, \alpha_n$  下的度量矩阵. 则

$$\alpha_1, \dots, \alpha_n \text{ 为标准正交基} \Leftrightarrow G = I_n.$$

下面将讨论标准正交基的存在性.

## 定理 (Schmidt 正交化)

给定欧氏空间的任意一组基  $\alpha_1, \dots, \alpha_n$ , 则存在一组标准正交基  $e_1, \dots, e_n$  使得 (对所有的  $i = 1, 2, \dots, n$ )

$$\langle \alpha_1, \dots, \alpha_i \rangle = \langle e_1, \dots, e_i \rangle.^a$$

<sup>a</sup>也等价于  $(\alpha_1, \dots, \alpha_i) = (e_1, \dots, e_i)R$ , 其中  $R$  为可逆上三角矩阵.

几何解释? 证明思路: 递归地定义  $\beta_k = \alpha_k - \sum_{i=1}^{k-1} (\alpha_k, e_i) e_i \neq 0$  以及  $e_k = \frac{\beta_k}{|\beta_k|}$ .

## 推论 (实可逆阵的 QR 分解)

对任意可逆实矩阵  $A$ , 存在方阵  $Q$  和上三角矩阵  $R$  使得  $A = QR$ , 其中  $Q$  的列向量构成  $\mathbb{R}^n$  的标准正交基.



例

将  $\alpha_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ ,  $\alpha_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ ,  $\alpha_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ ,  $\alpha_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix}$  标准正交化.

推论

对任意正定矩阵  $A$ , 存在可逆实矩阵  $P$  使得  $A = P^T P$ .

# 正交变换

$n$ 维欧氏空间	$\xleftrightarrow[\text{1:1}]{\text{标准正交基基}\alpha_1, \dots, \alpha_n}$	$\mathbb{R}^n$ (带标准内积)
线性变换 $\mathcal{A}$	$\xleftrightarrow[\text{1:1}]{\mathcal{A}(\alpha_1, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_n)A}$	矩阵 $A$
$\mathcal{A}$ (保持内积)	$\longleftrightarrow$	??

例

空间 (平面) 的旋转和镜面反射等.

定义 (正交变换)

设  $\mathcal{A}$  为欧氏空间  $V$  上的线性变换. 若  $\mathcal{A}$  保持内积, 即 (对任意  $a, b \in V$ ,

$$(\mathcal{A}a, \mathcal{A}b) = (a, b),$$

则称  $\mathcal{A}$  为 正交变换.

# 正交变换的等价刻画

## 定理

设  $\mathcal{A}$  为欧氏空间  $V$  上的线性变换. 则以下几条等价

- 1  $\mathcal{A}$  正交;
- 2  $\mathcal{A}$  保持向量长度;
- 3  $\mathcal{A}$  将标准正交基变为标准正交基.

证明思路: (1)  $\Leftrightarrow$  (2) & (1)  $\Leftrightarrow$  (3)

$$(1) \Rightarrow (2): |\mathcal{A}(\alpha)| = \sqrt{(\mathcal{A}(\alpha), \mathcal{A}(\alpha))} = \sqrt{(\alpha, \alpha)} = |\alpha|;$$

$$(2) \Rightarrow (1): (\mathcal{A}(\alpha), \mathcal{A}(\beta)) = \frac{|\mathcal{A}(\alpha+\beta)|^2 - |\mathcal{A}(\alpha)|^2 - |\mathcal{A}(\beta)|^2}{2} \\ = \frac{|\alpha+\beta|^2 - |\alpha|^2 - |\beta|^2}{2} = (\alpha, \beta)$$

$$(1) \Rightarrow (3): e_1, \dots, e_n = \text{标准正交基} \Rightarrow (\mathcal{A}(e_i), \mathcal{A}(e_j)) = (e_i, e_j) = \delta_{ij}$$

$$\Rightarrow \mathcal{A}(e_1), \dots, \mathcal{A}(e_n) = \text{标准正交基}.$$

(3)  $\Rightarrow$  (1): 设  $\mathcal{A}$  将标准正交基  $e_1, \dots, e_n$  映为另一组标准正交基  $\mathcal{A}(e_1), \dots, \mathcal{A}(e_n)$ . 任取

$$\alpha = \sum_{i=1}^n a_i e_i \quad \text{和} \quad \beta = \sum_{i=1}^n b_i e_i.$$

则

$$(\mathcal{A}(\alpha), \mathcal{A}(\beta)) = \left( \sum_{i=1}^n a_i \mathcal{A}(e_i), \sum_{j=1}^n b_j \mathcal{A}(e_j) \right) = \sum_{i=1}^n \sum_{j=1}^n a_i b_j (\mathcal{A}(e_i), \mathcal{A}(e_j)) \\ = \sum_{i=1}^n \sum_{j=1}^n a_i b_j \delta_{ij} = \sum_{i=1}^n \sum_{j=1}^n a_i b_j (e_i, e_j) = (\alpha, \beta)$$

# 正交变换基本性质

定理 (全体正交变换在复合运算下构成群)

设  $V$  为欧氏空间, 则

- ① 单位变换为正交变换;
- ② 正交变换的复合仍然为正交变换;
- ③ 正交变换可逆且其逆也为正交变换.

证明思路: 保持长度.

注: 正交变换群描述了欧氏空间的对称性.

# 正交变换在标准正交基下的矩阵

设正交变换  $\mathcal{A}$  在标准正交基  $e_1, \dots, e_n$  下的矩阵为  $A$ .

这个矩阵  $A$  满足什么特别的性质?

$\mathcal{A}$  正交  $\Leftrightarrow \mathcal{A}e_1, \dots, \mathcal{A}e_n$  为标准正交基

$$\Leftrightarrow (\mathcal{A}e_i, \mathcal{A}e_j) = \delta_{ij} \text{ (对任意的 } 1 \leq i, j \leq n)$$

$$\Leftrightarrow \left( \sum_{\ell=1}^n a_{\ell i} e_{\ell}, \sum_{k=1}^n a_{k j} e_k \right) = \delta_{ij} \text{ (对任意的 } 1 \leq i, j \leq n)$$

$$\Leftrightarrow \sum_{k=1}^n a_{ki} a_{kj} = \delta_{ij} \text{ (对任意的 } 1 \leq i, j \leq n)$$

$$\Leftrightarrow A^T A = I_n$$

## 定义 (正交矩阵)

若  $n$  阶实方阵  $A$  满足  $A^T A = I_n$  (或  $A^{-1} = A^T$ ), 则称  $A$  为 **正交矩阵**.

注: 正交矩阵的行向量 (或列向量) 构成  $\mathbb{R}^n$  的标准正交基 (在标准内积下).

# 正交变换在标准正交基下的矩阵

## 定理

设  $\mathcal{A}$  为欧氏空间  $V$  上的线性变换, 则  $\mathcal{A}$  为正交变换当且仅当  $\mathcal{A}$  在标准正交基下的矩阵  $A$  为正交矩阵.

$$\text{正交变换 } \mathcal{A} \xrightleftharpoons[1:1]{\text{标准正交基}} \text{正交矩阵 } A$$

## 定义 (第一类变换, 第二类变换)

设  $A$  为正交矩阵. 则  $A^T A = I$ . 因此  $\det(A) = \pm 1$ .

- ① 若  $\det(A) = 1$ , 则称  $\mathcal{A}$  为 **第一类变换**;
- ② 若  $\det(A) = -1$ , 则称  $\mathcal{A}$  为 **第二类变换**.

## 例

三维空间或在二维空间的旋转为第一类变换, 而镜面反射为第二类变换.

## 性质

设  $\mathcal{A}$  为欧氏空间  $V$  上的正交变换.

- ① 则  $\mathcal{A}$  的特征值模长都为 1. 特别地, 实特征值只可能为  $\pm 1$ .
- ② 若  $V$  的维数为奇数且  $\mathcal{A}$  为第一类正交变换, 则 1 为  $\mathcal{A}$  的特征值.

证明思路: (1). 设  $A$  为  $\mathcal{A}$  在某组标准正交基下的矩阵. 设  $A(\xi) = \lambda\xi$ , 则

$$\bar{\xi}^T \xi = \bar{\xi}^T \bar{A}^T A \xi = \bar{\lambda} \lambda \bar{\xi}^T \xi.$$

因此  $|\lambda| = 1$ .

(2). 多项式  $\varphi_{\mathcal{A}}(\lambda)$  为实系数, 其复根共轭成对出现. 而实根仅取  $\pm 1$ . 若 1 不为特征值, 则  $-1$  出现奇数次, 因此行列式为  $-1$ . 矛盾!

## 推论

三维空间中的第一类正交变换保持一个向量不变, 从而一定为旋转变换.

# 转置与伴随变换

$n$ 维欧氏空间	$\xleftrightarrow[\text{1:1}]{\text{标准正交基基}\alpha_1, \dots, \alpha_n}$	$\mathbb{R}^n$ (带标准内积)
线性变换 $\mathcal{A}$	$\xleftrightarrow[\text{1:1}]{\mathcal{A}(\alpha_1, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_n)A}$	矩阵 $A$
正交变换 $\mathcal{A}$	$\xleftrightarrow[\text{1:1}]{\mathcal{A}(\alpha_1, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_n)A}$	正交矩阵 $A$
??	$\longleftrightarrow$	转置矩阵 $A^T$

## 定理

设  $\mathcal{A}$  为欧氏空间  $V$  上的线性变换. 设  $\mathcal{A}$  在标准正交基  $e_1, \dots, e_n$  下的矩阵为  $A$ . 则

① 存在唯一的  $V$  上的线性变换, 记为  $\mathcal{A}^*$ , 满足

$$(\mathcal{A}\alpha, \beta) = (\alpha, \mathcal{A}^*\beta), \quad (\forall \alpha, \beta \in V).$$

② 线性变换  $\mathcal{A}^*$  在基  $e_1, \dots, e_n$  下的矩阵为  $A^T$ .

称  $\mathcal{A}^*$  为  $\mathcal{A}$  的伴随变换.

证明思路: 存在性: 证明矩阵  $A^T$  对应的线性变换满足性质. 唯一性: 若  $\mathcal{B}$  也满足条件, 则仅需证明  $|(\mathcal{A}^* - \mathcal{B})(\beta)| = 0$  对所有的  $\beta$  成立.



# 伴随变换的基本性质

## 性质

- ❶  $(\mathcal{A}^*)^* = \mathcal{A}$ ;
- ❷  $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$ ;
- ❸  $(\lambda \mathcal{A})^* = \lambda \mathcal{A}^*$ ;
- ❹  $(\mathcal{A} \circ \mathcal{B})^* = \mathcal{B}^* \circ \mathcal{A}^*$ ;

## 性质

设  $\mathcal{A}$  为欧氏空间上的线性变换, 则

$$\mathcal{A} \text{ 正交} \Leftrightarrow \mathcal{A}^* \mathcal{A} = \varepsilon \Leftrightarrow \mathcal{A}^* \text{ 正交}.$$

# 对称变换与对称矩阵

定义 (对称变换, 自伴随变换)

设  $\mathcal{A}$  为欧氏空间  $V$  上的线性变换. 若  $\mathcal{A} = \mathcal{A}^*$ , 则称  $\mathcal{A}$  为  $V$  上的 **对称变换** (或 **自伴随变换**).

注:  $\mathcal{A}$  对称  $\Leftrightarrow (\mathcal{A}a, b) = (a, \mathcal{A}b) \forall a, b \in V$ .

定理

设  $A$  为某欧氏空间上的线性变换  $\mathcal{A}$  在某组标准正交基下的矩阵. 则

$\mathcal{A}$  为对称变换  $\Leftrightarrow A$  为实对称矩阵.

证明:  $\mathcal{A}$  对称  $\Leftrightarrow \mathcal{A}^* = \mathcal{A} \Leftrightarrow A^T = A \Leftrightarrow A$  对称.

定理

对称变换的不同特征值对应的特征向量正交.

证明思路:  $(\mathcal{A}(\xi_1), \xi_2) = (\xi_1, \mathcal{A}(\xi_2)) \Rightarrow (\lambda_1 - \lambda_2) \cdot (\xi_1, \xi_2) = 0$ .

推论

实对称矩阵  $A$  的属于不同特征值的特征向量正交.

# 实对称矩阵的对角化

下面将证明实对称矩阵总是可对角化的.

## 性质

实对称矩阵的特征值都为实数.

证明思路: 设  $A\xi = \lambda\xi (\xi \neq 0)$ . 由于  $A$  为实矩阵,  $A\bar{\xi} = \overline{A\xi} = \overline{\lambda \cdot \xi} = \bar{\lambda} \cdot \bar{\xi}$ . 由于  $A$  对称, 我们有  $(A\bar{\xi})^T \xi = \bar{\xi}^T (A\xi)$ . 因此  $\bar{\lambda} = \lambda$ .

## 定理

任意  $n$  阶实对称矩阵  $A$ , 存在  $n$  阶正交矩阵  $T$  使得  $T^{-1}AT$  为对角矩阵.

证明思路: 将一个单位特征向量扩充为一组标准正交基, 并得正交阵  $T_n$ . 则  $T_n^{-1}AT_n = \begin{pmatrix} \lambda_1 & & \\ & A_{n-1} & \\ & & \end{pmatrix}$ , 其中  $A_{n-1}$  仍然为实对称. 归纳即可得证.

## 例

设  $A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$ . 求正交矩阵  $T$  使得  $T^{-1}AT$  为对角矩阵.

提示:  $P_A(\lambda) = (\lambda - 5)(\lambda + 1)^2$ .

## 定义 (正交)

设  $V_1, V_2$  为欧氏空间  $V$  的两子空间. 若对任意  $a_1 \in V_1, a_2 \in V_2$  都有  $(a_1, a_2) = 0$ , 则称  $V_1$  和  $V_2$  **相互正交**, 记为  $V_1 \perp V_2$ . 若一个向量  $a$  满足  $\langle a \rangle \perp V_1$ , 则称  $a$  与  $V_1$  **正交**, 记为  $a \perp V_1$ .

## 定理

- ① 若  $V_1 \perp V_2$ , 则  $V_1 + V_2$  为直和;
- ② 若  $V_1, V_2, \dots, V_r$  两两正交, 则  $V_1 + V_2 + \dots + V_r$  为直和.

## 定义 (正交补)

若  $V_1 \perp V_2$  且  $V = V_1 + V_2$ , 则称  $V_1, V_2$  互为 **正交补 (空间)**.

## 定理

欧氏空间的任意子空间的正交补存在且唯一.

# 正交投影

## 定义 (正交投影)

设  $W$  是欧氏空间  $\mathbb{R}^n$  的子空间,  $\mathbf{x} \in \mathbb{R}^n$ . 若  $\mathbf{y} \in W$  满足  $\mathbf{x} - \mathbf{y} \perp W$ , 则称  $\mathbf{y}$  为  $\mathbf{x}$  在  $W$  上的**正交投影**, 记为  $\mathbf{y} = P\mathbf{x}$ .

## 定理 (存在、唯一)

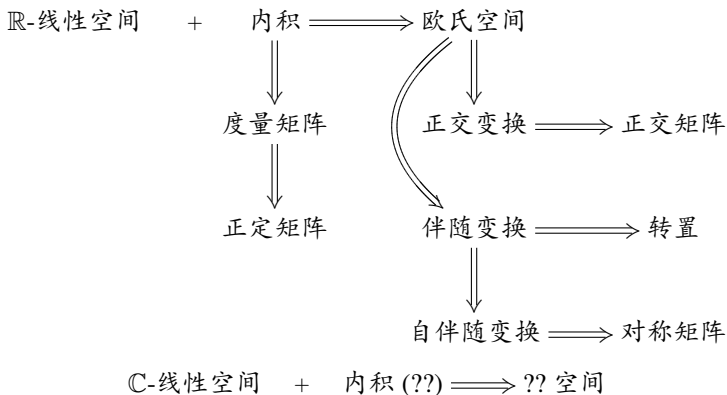
设  $W$  为欧氏空间  $\mathbb{R}^n$  的子空间. 设  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$  是  $W$  的一组标准正交基. 则

- ① 对任意  $\mathbf{x} \in \mathbb{R}^n$  在  $W$  上的正交投影存在唯一, 且
$$P\mathbf{x} = (\mathbf{x}, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{x}, \mathbf{u}_2)\mathbf{u}_2 + \dots + (\mathbf{x}, \mathbf{u}_r)\mathbf{u}_r.$$
- ② 对任意  $\mathbf{y} \in W$ ,  $|\mathbf{x} - \mathbf{y}| \geq |\mathbf{x} - P\mathbf{x}|$ . 等号成立当且仅当  $\mathbf{y} = P\mathbf{x}$ .

## 例

- ① 求  $(2, 1, 3)$  在  $\langle (1, 1, 0), (1, 0, -1) \rangle$  上的正交投影.
- ② 设  $A \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$ , 且  $A$  列满秩. 求线性方程组  $A\mathbf{x} = \mathbf{b}$  的**最小二乘解**, 即求  $\mathbf{x} \in \mathbb{R}^n$  使得  $|A\mathbf{x} - \mathbf{b}|$  达到最小.

考虑  $A$  的列空间  $C(A)$ .  $\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}$ .



# 复数模长和复数组向量模长 \*

$$\textcircled{1} \quad x \in \mathbb{R} \Rightarrow |x| = \sqrt{x \cdot x};$$

$$\textcircled{2} \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n \xrightarrow{\text{标准内积下}} |x| = \sqrt{x_1 x_1 + \dots + x_n x_n};$$

$$\textcircled{3} \quad z \in \mathbb{C} \Rightarrow |z| = \sqrt{\bar{z} \cdot z};$$

很自然地, 我们可以如下定义复数组向量的长度:

$$z = (z_1, \dots, z_n) \in \mathbb{C}^n \Rightarrow |z| = \sqrt{\bar{z}_1 z_1 + \dots + \bar{z}_n z_n};$$

对于任意复矩阵  $A$ , 记

$$A^H := \bar{A}^T$$

称其为  $A$  的 **共轭转置**. 则复数组 (列) 向量的长度公式可写为

$$|z|^2 = z^H \cdot z.$$

注: 共轭转置保持加法, 但只是在共轭意义下保持数乘.

# 复向量空间上的内积定义\*

## 定义 (酉空间)

设  $V$  为复线性空间. 若映射

$$(-, -): V \times V \longrightarrow \mathbb{C}$$

满足

- ① 共轭对称性  $(a, b) = \overline{(b, a)}$ ;
- ② 线性性  $(a, \lambda b + \mu c) = \lambda(a, b) + \mu(a, c)$ ;
- ③ 正定性  $(a, a) \geq 0$ , 且等号成立当且仅当  $a = 0$ ,

则称  $(a, b)$  为  $a$  和  $b$  的**内积**. 定义内积的复线性空间称为**酉空间**.

注:  $(\lambda a, b) = \overline{\lambda}(a, b)$ .



## 定义

设  $V$  为  $n$  维酉空间,  $G$  为  $n$  阶复矩阵.

① 称  $G$  为 **正定复矩阵**, 若

①  $G^H = G$ ;

② 任意给定列向量  $z \in \mathbb{C}^n$  都有  $z^H G z \geq 0$ , 并且等号成立当且仅当  $z = 0$ .

② 向量  $a$  的**长度 (或模长)**定义为  $|a| := \sqrt{(a, a)}$ .

③ **垂直 (正交)**, **标准正交基**, **Schmit 正交化**, **正交补**, **Schwarz 不等式**, ...

# 共轭变换(伴随变换)\*

## 定义(共轭变换, 伴随变换)

设  $\mathcal{A}$  为酉空间  $V$  上的一个线性变换. 则

- 存在唯一的  $V$  上的线性变换  $\mathcal{A}^*$  满足对任意  $a, b \in V$

$$(\mathcal{A}a, b) = (a, \mathcal{A}^*b).$$

称  $\mathcal{A}^*$  是  $\mathcal{A}$  的共轭变换(或伴随变换).

- 若  $\mathcal{A}$  在标准正交基下矩阵为  $A$ , 则  $\mathcal{A}^*$  在同一组标准正交基下矩阵为  $A^H$ .

设  $\mathcal{U}$  为酉空间  $V$  上的线性变换. 则

$\mathcal{U}$  为酉变换  $\xLeftrightarrow{\text{定义}} \mathcal{U}$  保持内积

$\Leftrightarrow \mathcal{U}$  保持向量长度

$\Leftrightarrow \mathcal{U}$  将标准正交基变为标准正交基

$\Leftrightarrow \mathcal{U}^* \mathcal{U} = \text{id}_V$

$\Leftrightarrow \mathcal{U}$  在标准正交基下的矩阵  $U$  为酉矩阵(即,  $U^H U = I$ )

## 定理

酉空间  $V$  上的全体酉变换组成的集合  $U(V)$  在复合的作用下构成群.

设酉空间  $V$  上的线性变换  $\mathcal{A}$  在某组标准正交基下矩阵为  $A$ . 则<sup>3</sup>

$\mathcal{A}$  为Hermite 变换  $\xLeftrightarrow{\text{定义}} \mathcal{A}^* = \mathcal{A}$  (自伴随)

$$\Leftrightarrow (\mathcal{A}a, b) = (a, \mathcal{A}b), \quad (\forall a, b \in V)$$

$$\Leftrightarrow A \text{ 为Hermite 矩阵 (即, } A^H = A \text{)}.$$

---

<sup>3</sup>在物理中,Hermite 变换代表可观测量.

# 规范变换\*

设  $\mathcal{A}$  为酉空间  $V$  上的线性变换  $\mathcal{A}$ .

$\mathcal{A}$  为规范变换  $\xLeftrightarrow{\text{定义}} \mathcal{A}^* \mathcal{A} = \mathcal{A} \mathcal{A}^*$

$\Leftrightarrow A$  酉相似于对角阵

$\Leftrightarrow A$  在某组标准正交基下为对角阵.

## 推论

设  $\mathcal{U}$  为酉变换. 则

- ①  $\mathcal{U}$  的特征值  $\lambda$  模为 1. 即, 存在  $\theta \in [0, 2\pi)$  使得  $\lambda = e^{i\theta}$ .
- ②  $\mathcal{U}$  在某组标准正交基下矩阵为对角阵  $\text{diag}(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_n})$ .

证明思路:  $D^H D = I \Rightarrow |\lambda_i| = 1$ .

## 推论

设  $\mathcal{A}$  为 Hermite 变换. 则

- ①  $\mathcal{A}$  的特征值全为实数;
- ②  $\mathcal{A}$  在某组标准正交基下的矩阵为实对角矩阵.

证明思路:  $D^H = D \Rightarrow \lambda_i \in \mathbb{R}$ .

问题: 反 Hermite 变换?



## 第九章 实二次型

# 实对称矩阵的正交相合标准形

遗留问题: 相合等价关系的标准形.

下面我们仅考虑实对称矩阵的相合标准形<sup>4</sup>.

由于任意实对称矩阵正交相似于对角阵, 因此我们有如下结论.

## 性质 (正交相合标准形)

设  $n$  阶实对称矩阵  $A$  的全体特征值为  $\lambda_1, \dots, \lambda_n$ . 则存在正交矩阵  $P$  使得

$$P^T A P = \text{diag}(\lambda_1, \dots, \lambda_n).$$

从而矩阵  $A$  相合于对角矩阵  $\text{diag}(\lambda_1, \dots, \lambda_n)$ . 称这个对角阵为  $A$  的 **正交相合标准形**.

<sup>4</sup>注: 若矩阵  $A, B$  相合, 则  $A$  对称当且仅当  $B$  对称.



# 实对称矩阵的相合规范形

实对称矩阵的特征值分为三类: 正, 负, 零.

## 定理 (相合规范形)

设  $A$  为实对称矩阵, 则存在可逆阵  $P$  使得

$$P^T A P = \text{diag}(I_r, -I_s, 0). \quad (A \text{ 的相合规范形})$$

其中  $r, s$  由  $A$  唯一确定不依赖于  $P$  的选取, 且满足

$$r + s = \text{rank}(A) \leq n.$$

称  $r$  为  $A$  的**正惯性指数**, 称  $s$  为  $A$  的**负惯性指数**, 称  $r - s$  为  $A$  的**符号差**.

证明思路: 存在性, 显然. 反证唯一性: 若  $\text{diag}(I_r, -I_s, 0)$  与  $\text{diag}(I_{r'}, -I_{s'}, 0)$  相合, 设

$P^T \text{diag}(I_{r'}, -I_{s'}, 0) P = \text{diag}(I_r, -I_s, 0)$ , 则  $r + s = r' + s'$ . 不妨设  $r < r'$  (则  $s > s'$ ). 记  $y = Px$ . 则齐次方程组  $x_1 = \cdots = x_r = y_{r'+1} = \cdots = y_n = 0$  有非零解  $x_0 = (0, \cdots, 0, a_{r+1}, \cdots, a_n)^T \neq 0$ . 令  $y_0 = Px_0 = (b_1, \cdots, b_{r'}, 0, \cdots, 0)^T$ . 则

$$\begin{aligned} 0 < b_1^2 + \cdots + b_{r'}^2 &= (Px_0)^T \text{diag}(I_{r'}, -I_{s'}, 0) (Px_0) \\ &= x_0^T \text{diag}(I_r, -I_s, 0) x_0 = -a_{r+1}^2 - \cdots - a_{r+s}^2 \leq 0. \end{aligned}$$

注: 规范形唯一. 其中  $r, s$  分别为  $A$  的正特征值和负特征值的个数.

## 推论

$n$  阶实对称矩阵  $A$  正定当且仅当其正惯性指数为  $n$ .

在各个领域我们会碰到很多二次齐次多项式. 例如, 三维空间和四维时空中点到原点的距离.

例

❶ 空间距离公式

$$d^2 = x^2 + y^2 + z^2 = (x, y, z) \begin{pmatrix} x \\ y \\ z \end{pmatrix};$$

❷ 时空 (Lorentz 空间) 中的距离公式

$$d^2 = x^2 + y^2 + z^2 - c^2 t^2 = (x, y, z, t) \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -c^2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix};$$

注: 保持时空距离 (Lorentz 内积) 的变换称为 Lorentz 变换.

# 二次型的定义

## 定义 ((实) 二次型)

称实数域<sup>a</sup>上的二次齐次多项式为(实)二次型.

<sup>a</sup>这门课程仅考虑实数域上的二次型. 更一般地, 可类似的定义任意数域上的二次型.

## 例

- ①  $Q(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_3x_1$ ;
- ②  $Q(x_1, x_2, \dots, x_n) = \lambda_1x_1^2 + \lambda_2x_2^2 + \dots + \lambda_nx_n^2$ ;
- ③ 欧氏度量  $Q(x, y, z) = x^2 + y^2 + z^2$ ;
- ④ 时空度量  $Q(x, y, z, t) = x^2 + y^2 + z^2 - c^2t^2$ .

# 二次型的矩阵

由于  $n$  个变元  $x_1, x_2, \dots, x_n$  的二次单项式全体集为

$$x_1^2, x_1x_2, \dots, x_1x_n, \quad x_2^2, \dots, x_2x_n, \quad x_3^2, \dots, \quad x_n^2,$$

因此一般的二次型可表示为

$$\begin{aligned} Q(x_1, \dots, x_n) = & a_{11}x_1^2 + 2a_{12}x_1x_2 + \dots + 2a_{1n}x_1x_n \\ & + a_{22}x_2^2 + \dots + 2a_{2n}x_2x_n \\ & \qquad \qquad \qquad \ddots \qquad \qquad \qquad \vdots \\ & \qquad \qquad \qquad \qquad \qquad \qquad + a_{nn}x_n^2 \end{aligned}$$

其中  $a_{ij}$  全部为实数. 根据  $x_ix_j = x_jx_i$ , 我们也可将上式改写为

$$\begin{aligned} Q(x_1, \dots, x_n) = & a_{11}x_1x_1 + 2a_{12}x_1x_2 + \dots + 2a_{1n}x_1x_n \\ & + 2a_{12}x_2x_1 + a_{22}x_2x_2 + \dots + 2a_{2n}x_2x_n \\ & \qquad \qquad \qquad \vdots \qquad \qquad \qquad \ddots \qquad \qquad \qquad \vdots \\ & + 2a_{1n}x_nx_1 + 2a_{2n}x_nx_2 + \dots + a_{nn}x_nx_n \end{aligned}$$

# 二次型的矩阵

利用矩阵的乘法,可进一步的将二次型改写为

$$Q(x_1, \cdots, x_n) = x^T A x, \quad (2)$$

其中  $A = (a_{ij})_{n \times n}$  实对称,  $x = (x_1, x_2, \cdots, x_n)^T$  为变元组成的列向量.

## 性质 (二次型的矩阵)

- ① 式(2)中的实对称矩阵  $A$  由二次型  $Q(x_1, \cdots, x_n)$  唯一确定, 称  $A$  为二次型  $Q$  的矩阵. 此外, 称  $A$  的秩为二次型  $Q$  的秩, 称  $A$  的特征值为二次型  $Q$  的特征值.
- ② 反之, 对于任意实对称矩阵, 通过式(2), 可以构造一个二次型.

简言之, 我们有如下对应:

$$\text{二次型} \xleftrightarrow[1:1]{Q=x^T A x} \text{实对称矩阵}^5$$

<sup>5</sup>若我们不要求式(2)中的矩阵  $A$  对称, 则  $A$  的选取不唯一.

例如, 给定任意反对称矩阵  $A$ , 二次型  $x^T A x$  都为零二次型.

例

求下列二次型对应的矩阵.

- ❶  $Q(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_3x_1$ ;
- ❷  $Q(x_1, x_2, \dots, x_n) = \lambda_1x_1^2 + \lambda_2x_2^2 + \dots + \lambda_nx_n^2$ ;
- ❸ 欧氏度量  $Q(x, y, z) = x^2 + y^2 + z^2$ ;
- ❹ 时空度量  $Q(x, y, z, t) = x^2 + y^2 + z^2 - c^2t^2$ ;
- ❺  $Q(x_1, \dots, x_n) = x^T Ax$ , 其中  $A$  为  $n$  阶实方阵.

# 二次型变元的可逆线性替换

给定二次型

$$Q = Q(x_1, \cdots, x_n) = x^T A x. \quad (A^T = A),$$

对变量  $x_1, \cdots, x_n$  做一个可逆 (非退化) 的线性变换

$$\begin{cases} x_1 = p_{11}y_1 + p_{12}y_2 + \cdots + p_{1n}y_n, \\ x_2 = p_{21}y_1 + p_{22}y_2 + \cdots + p_{2n}y_n, \\ \quad \quad \quad \cdots \cdots \cdots \\ x_n = p_{n1}y_1 + p_{n2}y_2 + \cdots + p_{nn}y_n, \end{cases}$$

其中系数矩阵  $P = (p_{ij})_{n \times n}$  可逆. 则

$$Q = x^T A x = (y^T P^T) A (P y) = y^T (P^T A P) y,$$

其中  $x = (x_1, \cdots, x_n)^T, y = (y_1, \cdots, y_n)^T$ . 因此, 二次型  $Q$  关于新变元  $y_1, \cdots, y_n$  的矩阵为  $P^T A P$ .

# 二次型的标准形

通过变元的线性替换, 实对称矩阵的相合标准形可以用来化简二次型.

## 定义 ((有理) 标准形)

若二次型经过非退化的线性变换  $x = Py$  化为

$$\tilde{Q}(y_1, \dots, y_n) = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2,$$

则称  $\tilde{Q}$  为  $Q$  的一个(有理)标准形.

下面通过对称矩阵的正交合同标准形和规范形, 给出两个特殊的 (有理) 标准形:

## 定理 (正交标准形)

任意给定实二次型  $Q = x^T A x$ , 存在正交变换  $x = Py$  将  $Q$  化为

$$\tilde{Q}(y_1, \dots, y_n) = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2,$$

这里的  $\lambda_1, \dots, \lambda_n$  为  $A$  的特征值. 称  $\tilde{Q}$  为  $Q$  的正交标准形.



## 定理 (规范形 (实标准形))

任意给定二次型  $Q$ , 存在线性变换  $x = Py$  使得

$$Q = y_1^2 + \cdots + y_r^2 - y_{r+1}^2 - \cdots - y_{r+s}^2,$$

其中  $r$  和  $s$  不依赖于  $P$  的选取. 我们称

- $y_1^2 + \cdots + y_r^2 - y_{r+1}^2 - \cdots - y_{r+s}^2$  为  $Q$  的**规范形**(或**实标准形**);
- $r$  和  $s$  为  $Q$  的**正负惯性指数**.

问题:

- ① 如何寻找二次型的标准形?
- ② 如何寻找二次型的规范形?
- ③ 如何寻找二次型的正交标准形?

# 配方法 (求标准形)

例 (含平方项)

$$Q(x_1, x_2, x_3) = 2x_1^2 + x_1x_2 + x_3^2.$$

例 (无平方项)

$$Q(x_1, x_2, x_3) = 2x_1x_2 - 6x_2x_3 + 2x_1x_3.$$

定理

任意二次型都可由配方法化为标准形.

证明思路: 情形一:  $a_{11} \neq 0$ ; 情形二:  $a_{11} = 0$  但对某个指标  $i$  有  $a_{ii} \neq 0$ ; 情形三: 对角元  $a_{ii}$  全为零, 但  $a_{ij} \neq 0$ ; 情形四:  $A = 0$ .

类似于求解线性方程组的 Gauss 消元法可以用矩阵的初等变换表示, 在这里二次型的配方法也可以用矩阵的初等变换来表示.

定理 (配方法能用矩阵实现的理论依据)

对于任意给定的实对称矩阵  $A$  都存在初等矩阵  $P_1, \dots, P_r$  使得

$$P_r^T \cdots P_1^T A P_1 \cdots P_r = \text{diag}(\mu_1, \mu_2, \dots, \mu_n).$$

# 初等变换法 (具体算法步骤)

第一步: ① 若  $a_{11} \neq 0$ ,

- 将第 1 列的  $-a_{i1}a_{11}^{-1}$  倍加到第  $i$  列;
- 将第 1 行的  $-a_{1i}a_{11}^{-1}$  倍加到第  $i$  行.

$$\Rightarrow \text{可逆阵 } P_1 \text{ 使得 } P_1^T A P_1 = \begin{pmatrix} a_{11} & & \\ & A_{n-1} & \end{pmatrix}.$$

② 若  $a_{11} = 0$ , 且存在对角元  $a_{ii} \neq 0$ .

- 交换第 1 列与第  $i$  列;
- 交换第 1 行与第  $i$  行;
- 回到第一步.

③ 若对角线全为零且第一行 (第一列) 不为零. 即, 存在  $i = 2, \dots, n$  使得  $a_{i1} = a_{1i} \neq 0$ .

- 将第  $i$  行加到第一行;
- 将第  $i$  列加到第一列;
- 回到第一步.

④ 若对角线全为零且第一行和第一列全为零. 则

$$A = \begin{pmatrix} 0 & & \\ & A_{n-1} & \end{pmatrix}.$$

第二步: 递归下去, 对  $n-1$  阶矩阵  $A_{n-1}$  重复第一步.

# 初等变换法 (具体算法步骤)

第一步: ① 若  $a_{11} \neq 0$ ,

- 将第 1 列的  $-a_{i1}a_{11}^{-1}$  倍加到第  $i$  列;
- 将第 1 行的  $-a_{1i}a_{11}^{-1}$  倍加到第  $i$  行.

$$\Rightarrow \text{可逆阵 } P_1 \text{ 使得 } P_1^T A P_1 = \begin{pmatrix} a_{11} & & \\ & A_{n-1} & \end{pmatrix}.$$

② 若  $a_{11} = 0$ , 且存在对角元  $a_{ii} \neq 0$ .

- 交换第 1 列与第  $i$  列;
- 交换第 1 行与第  $i$  行;
- 回到第一步.

③ 若对角线全为零且第一行 (第一列) 不为零. 即, 存在  $i = 2, \dots, n$  使得  $a_{i1} = a_{1i} \neq 0$ .

- 将第  $i$  行加到第一行;
- 将第  $i$  列加到第一列;
- 回到第一步.

④ 若对角线全为零且第一行和第一列全为零. 则

$$A = \begin{pmatrix} 0 & & \\ & A_{n-1} & \end{pmatrix}.$$

第二步: 递归下去, 对  $n-1$  阶矩阵  $A_{n-1}$  重复第一步.

$$\begin{pmatrix} A \\ I \end{pmatrix} \xrightarrow[\text{对 } I \text{ 做初等列变换}]{\text{对 } A \text{ 做成对的行列初等变换}} \begin{pmatrix} P^T A P \\ P \end{pmatrix}$$

类似的

$$(A, I) \xrightarrow[\text{对 } I \text{ 做初等行变换}]{\text{对 } A \text{ 做成对的行列初等变换}} (P^T A P, P^T)$$

例

$$Q(x_1, x_2, x_3) = x_1^2 - x_2^2 + x_3^2 + 4x_1x_3 + 2x_2x_3.$$

$$\begin{aligned} \text{解: } \begin{pmatrix} A \\ I \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} &\xrightarrow{-2C_1 \rightarrow C_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 2 & 1 & -3 \\ 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{-2r_1 \rightarrow r_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -3 \\ 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &\xrightarrow{C_2 \rightarrow C_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -2 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_2 \rightarrow r_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

$$\text{从而得标准形 } \tilde{Q}(y_1, y_2, y_3) = y_1^2 - y_2^2 - 2y_3^2, \text{ 其中 } P = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

例

$$Q(x_1, x_2, x_3) = 2x_1x_2 + 4x_1x_3.$$

解:  $(A, I) = \begin{pmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

$$\xrightarrow{r_2 \rightarrow r_1} \begin{pmatrix} 1 & 1 & 2 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{c_2 \rightarrow c_1} \begin{pmatrix} 2 & 1 & 2 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[-c_1 \rightarrow c_3]{-\frac{1}{2}c_1 \rightarrow c_2} \begin{pmatrix} 2 & 0 & 0 & 1 & 1 & 0 \\ 1 & -\frac{1}{2} & -1 & 0 & 1 & 0 \\ 2 & -1 & -2 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[-r_1 \rightarrow r_3]{-\frac{1}{2}r_1 \rightarrow r_2} \begin{pmatrix} 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & -1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -1 & -2 & -1 & -1 & 1 \end{pmatrix}$$

$$\xrightarrow{-2r_1 \rightarrow r_3} \begin{pmatrix} 2 & 0 & 0 & 1 & 1 & 0 \\ 1 & -\frac{1}{2} & -1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & -2 & 1 \end{pmatrix} \xrightarrow{-2c_1 \rightarrow c_3} \begin{pmatrix} 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & -2 & 1 \end{pmatrix}$$

从而得到标准形

$$\tilde{Q}(y_1, y_2, y_3) = 2y_1^2 - \frac{1}{2}y_2^2.$$

其中

$$P = \begin{pmatrix} 1 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -2 & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

## 定义 (正定二次型)

称一个  $n$  元二次型  $Q(x_1, \dots, x_n) = x^T A x (A^T = A)$  为 **正定二次型**, 如果对于任意向量  $x \in \mathbb{R}^n \setminus \{0\}$  都有  $Q(x_1, \dots, x_n) > 0$  成立.

## 定理

$Q$  正定  $\Leftrightarrow A$  正定  $\Leftrightarrow A$  相合于单位阵  
 $\Leftrightarrow Q$  的 (或  $A$  的) 正惯性指数为  $n$

# 正定性判定

## 性质

设  $A$  为  $n$  阶实对称方阵.

- ① 若  $P$  可逆, 则  $P^TAP > 0$  当且仅当  $A > 0$ . 相合不变性
- ②  $A > 0$  当且仅当存在可逆阵  $P$  使得  $A = P^TP$ . 相合标准形
- ③ 若  $A > 0$ , 则  $\det(A) > 0$ .
- ④  $A > 0$  当且仅当  $A$  的特征值全为正.

## 定理

设  $A = (a_{ij})_{n \times n}$  为  $n$  阶实对称方阵. 则  $A$  正定当且仅当  $A$  的各阶顺序主子式均大于零. 即,  $A$  正定当且仅当

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \dots, \quad \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \det A > 0,$$

证明思路: 对阶数归纳. 由归纳假设  $A$  可写为  $A = \begin{pmatrix} (P^{-1})^T P^{-1} & C \\ C^T & a_{nn} \end{pmatrix}$ . 记

$$R = \begin{pmatrix} P & -PP^TC \\ 0 & 1 \end{pmatrix}. \text{ 则 } R^TAR = \begin{pmatrix} I_{n-1} & \\ & a_{nn} - C^T P P^T C \end{pmatrix}.$$



## 例

- 证明: 对任意实数  $x, y, f(x, y) = x^2 + 2xy + 2y^2 + 2x + 6y + 6 > 0$ .
- 给定实二次型  $Q(x, y, z) = \lambda(x^2 + y^2 + z^2) - 2xy - 2yz - 2zx$ .  $\lambda$  为何值时,  $Q(x, y, z)$  正定?
- 设向量组  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m \in \mathbb{R}^n$ . 则方阵  $G = ((\mathbf{a}_i, \mathbf{a}_j)_{m \times m})$  正定当且仅当  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$  线性无关.

设  $A$  为实对称矩阵, 记  $Q(x_1, \dots, x_n) = x^T A x$  为  $A$  对应的二次型.

## 定义 (半正定二次型)

称  $Q(x_1, \dots, x_n)$  为半正定二次型, 如果对于任意向量  $x \in \mathbb{R}^n \setminus \{0\}$  都有  $Q(x_1, \dots, x_n) \geq 0$  成立. 此时, 称矩阵  $A$  为半正定矩阵, 简记为  $A \geq 0$ .

## 性质

$Q$  半正定  $\Leftrightarrow A$  半正定

$\Leftrightarrow A$  相合于  $\text{diag}(I_r, 0)$  (相合不变性)

$\Leftrightarrow$  存在实矩阵  $P$  使得  $A = P^T P$

$\Leftrightarrow A$  的负惯性指数为 0

$\Leftrightarrow A$  的特征值均非负 ( $\Rightarrow \det(A) \geq 0$ )

$\Leftrightarrow A$  的各阶主子式均非负

## 例

- 设  $A, B, C$  为三角形  $ABC$  的三个内角. 证明: 对任意实数  $x, y, z$ ,  
$$x^2 + y^2 + z^2 \geq 2xy \cos(A) + 2yz \cos(B) + 2zx \cos(C).$$

提示:  $\det = 0$ .

- 三维欧氏空间  $\mathbb{R}^3$  中, 是否存在向量  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  满足

$$\begin{aligned} (\mathbf{a}_1, \mathbf{a}_1) &= 4, & (\mathbf{a}_2, \mathbf{a}_2) &= 9, & (\mathbf{a}_3, \mathbf{a}_3) &= 4, \\ (\mathbf{a}_1, \mathbf{a}_2) &= \frac{9}{2}, & (\mathbf{a}_1, \mathbf{a}_3) &= -\frac{1}{2}, & (\mathbf{a}_2, \mathbf{a}_3) &= \frac{9}{2}. \end{aligned}$$

提示:  $\det = -81/2$ .

## 定理 (矩阵的奇异值分解)

设  $A \in \mathbb{R}^{m \times n}$ , 则存在  $m$  阶正交阵  $U$  及  $n$  阶正交阵  $V$  使得

$$A = U \begin{pmatrix} \Lambda & O \\ O & O \end{pmatrix} V$$

其中  $\Lambda = \text{diag}(\sigma_1, \dots, \sigma_r)$ ,  $\sigma_1 \geq \dots \geq \sigma_r > 0$  称为矩阵  $A$  的奇异值.

取正交阵  $U$  使得  $AA^T = U \text{diag}(\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0) U^T$ . 再将  $B = U^{-1}A$  的非零行向量施密特正交化后扩充得正交矩阵  $V$ .

## 定理 (多元函数的极值)

设  $\Omega \subset \mathbb{R}^n$  是有界区域,  $\mathbf{x}^0$  是  $\Omega$  内一点,  $f(\mathbf{x}) \in C^2(\Omega)$ , 且  $\frac{\partial f}{\partial x_i}(\mathbf{x}^0) = 0$ ,

$i = 1, \dots, n$ ,  $H(\mathbf{x}^0) = \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}^0) \right)_{n \times n} > 0$ . 则  $f(\mathbf{x})$  在  $\mathbf{x}^0$  取极小值.

设  $\lambda > 0$  为  $H(\mathbf{x}^0)$  的最小特征值. 则

$$\begin{aligned} f(x) &= f(x^0) + \frac{1}{2}(x - x^0)^T H(x^0)(x - x^0) + o(\|x - x^0\|^2) \\ &\geq f(x^0) + \frac{\lambda}{2}\|x - x^0\|^2 + o(\|x - x^0\|^2) \end{aligned}$$

# 根据正负惯性指数命名二次型

① 正定  $\Leftrightarrow r = n (\Rightarrow s = 0)$ ;

② 半正定  $\Leftrightarrow s = 0$ ;

③ 负定  $\Leftrightarrow s = n (\Rightarrow r = 0)$ ;

④ 半负定  $\Leftrightarrow r = 0$ ;

⑤ 不定  $\Leftrightarrow r \geq 1$  且  $s \geq 1$ .

部分关于正定的结论可以平移到半正定, 负定, 半负定的二次型上.

# 二次曲线与二次曲面的分类

二次曲线的标准形式

$$\left\{ \begin{array}{l} \text{椭圆: } \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \\ \text{双曲线: } \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \\ \text{抛物线: } y = ax^2 \\ \text{退化: } x^2 = \frac{y^2}{a^2} \text{ 或 } \frac{x^2}{a^2} = 1 \text{ 或 } x^2 = 0 \end{array} \right.$$

平面二次曲线方程:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c = 0.$$

定理

任意平面二次曲线均可经过选定合适的直角坐标系变为标准形式.

证明思路: 通过正交变换化简为 (其中  $\lambda_1, \lambda_2$  不全为零)

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2b_1 x' + 2b_2 y' + c' = 0.$$

- ①  $\lambda_1 \lambda_2 > 0$  椭圆型
- ②  $\lambda_1 \lambda_2 < 0$  双曲型或退化
- ③  $\lambda_1 \lambda_2 = 0$  抛物型或退化

# 二次曲面标准形式

## ① 椭圆面

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1;$$

## ② 单叶双曲面

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1;$$

## ③ 双叶双曲面

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1;$$

## ④ 二次锥面

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2};$$

5 椭圆抛物面

$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2};$$

6 双曲抛物面 (马鞍面)

$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2};$$

7 二次柱面 (方程中不含第三个变量)

- 椭圆柱面  $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ ;
- 双曲柱面  $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ ;
- 抛物柱面  $y = ax^2$ ;

8 退化二次曲面

- 两个相交平面  $x^2 = \frac{y^2}{a^2}$ ;
- 两个平行平面  $\frac{x^2}{a^2} = 1$ ;
- 两个重合平面  $x^2 = 0$ .



二次曲面方程:

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz \\ + 2b_1x + 2b_2y + 2b_3z + c = 0$$

### 定理

任意二次曲面可经过选择合适的直角坐标系变为标准形式.

证明思路: 二次曲面通过正交变换可化简为 (其中  $\lambda_1, \lambda_2, \lambda_3$  不全为零)

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + 2b'_1 x' + 2b'_2 y' + 2b'_3 z' + c' = 0.$$

①  $\lambda_1 \lambda_2 \lambda_3 \neq 0 \xrightarrow{\text{化简}} \lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + \lambda_3 \tilde{z}^2 = \lambda_4$

- $\lambda_4 \neq 0 \Rightarrow$  椭圆型或双曲型
- $\lambda_4 = 0 \Rightarrow$  二次锥面

②  $\lambda_1, \lambda_2, \lambda_3$  中两个非零 (不妨设  $\lambda_1 \lambda_2 \neq 0$  且  $\lambda_3 = 0$ )  $\xrightarrow{\text{化简}}$

$$\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 = \tilde{b}_3 \tilde{z} + \tilde{c}.$$

- $\tilde{b}_3 \neq 0$  (可设  $\tilde{c} = 0$ )  $\Rightarrow$  抛物面
- $\tilde{b}_3 = 0$  且  $\tilde{c} \neq 0 \Rightarrow$  椭圆柱面或双曲柱面
- $\tilde{b}_3 = 0 = \tilde{c} \Rightarrow$  两相交平面

③  $\lambda_1, \lambda_2, \lambda_3$  中仅一个非零 (不妨设  $\lambda_1 \neq 0$  且  $\lambda_2 = 0 = \lambda_3$ )  $\xrightarrow{\text{化简}}$

$$\lambda_1 \tilde{x}^2 = \tilde{b}_2 \tilde{y} + \tilde{c}.$$

- $\tilde{b}_2 \neq 0$  (可设  $\tilde{c} = 0$ )  $\Rightarrow$
- $\tilde{b}_2 = 0$  且  $\tilde{c} \neq 0 \Rightarrow$  两平行的平面
- $\tilde{b}_3 = 0 = \tilde{c} \Rightarrow$  两重合的平面

例

$$x^2 + 4y^2 + z^2 - 4xy - 8xz - 4yz - 1.$$

答: 正交方阵  $P = \begin{pmatrix} \frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} & \frac{2}{3} \\ -\frac{2\sqrt{5}}{5} & \frac{2\sqrt{5}}{15} & \frac{1}{3} \\ 0 & -\frac{\sqrt{5}}{3} & \frac{2}{3} \end{pmatrix}$

$$P^T A P = \text{diag}(5, 5, -4).$$

$$(x, y, z)^T = P(x', y', z')^T \Rightarrow 5x'^2 + 5y'^2 - 4z'^2 = 1 \Rightarrow \text{单叶双曲面}.$$

方法二 (配方):

$$\Rightarrow (x - 2y - 4z)^2 - 15z^2 - 20yz - 1 = 0$$

$$\Rightarrow (x - 2y - 4z)^2 - 15(z + \frac{2}{3}y)^2 + \frac{20}{3}y^2 = 1$$

$$\Rightarrow \tilde{x}^2 + \tilde{y}^2 - \tilde{z}^2 = 1$$

$\Rightarrow$  单叶双曲面.