第一章：线性映射和矩阵

本章主要内容

1.1 线性方程组问题

1.2 三种视角看待方程组问题与映射的基本概念

1.3 矩阵的定义

1.4 线性方程组有解之判定

1.5 矩阵的运算（加法、数乘和乘法）

1.6 矩阵的逆

1.7 矩阵的相抵标准型

1.8 分块矩阵

1.9 LU分解

通过线性方程问题引入线性映射和矩阵

矩阵在线性方程组问题的初步应用；第二章完全解决线性方程组问题

研究矩阵的性质；
1.5，1.6是重点

1.1 线性方程组问题

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

n元1次方程组（m个方程）

本章：判断方程组是否有解

线性方程组求解

$$\begin{cases} x - 2y = 1 & ① \\ 3x + 2y = 11 & ② \end{cases}$$

代数上看：

消去含x的项

方程② - 3×方程①：

$$8y = 8$$

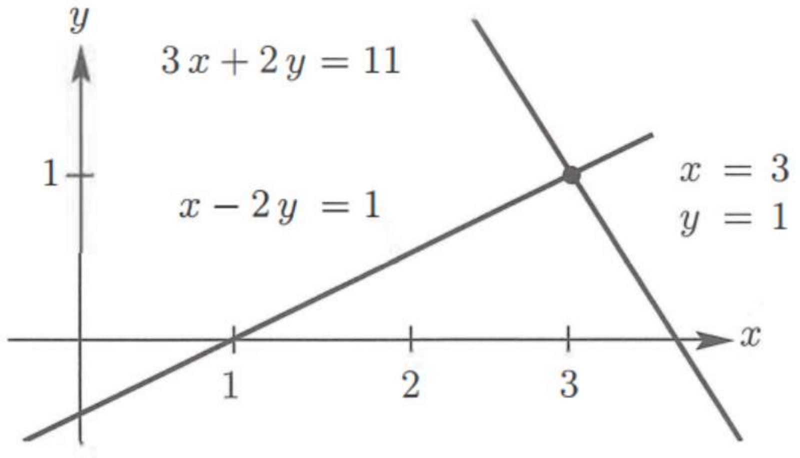
进一步得到：

$$y = 1$$
$$x = 3$$

线性方程组问题示例

$$\begin{cases} \boxed{x - 2y = 1} \\ \boxed{3x + 2y = 11} \end{cases}$$

几何上看：

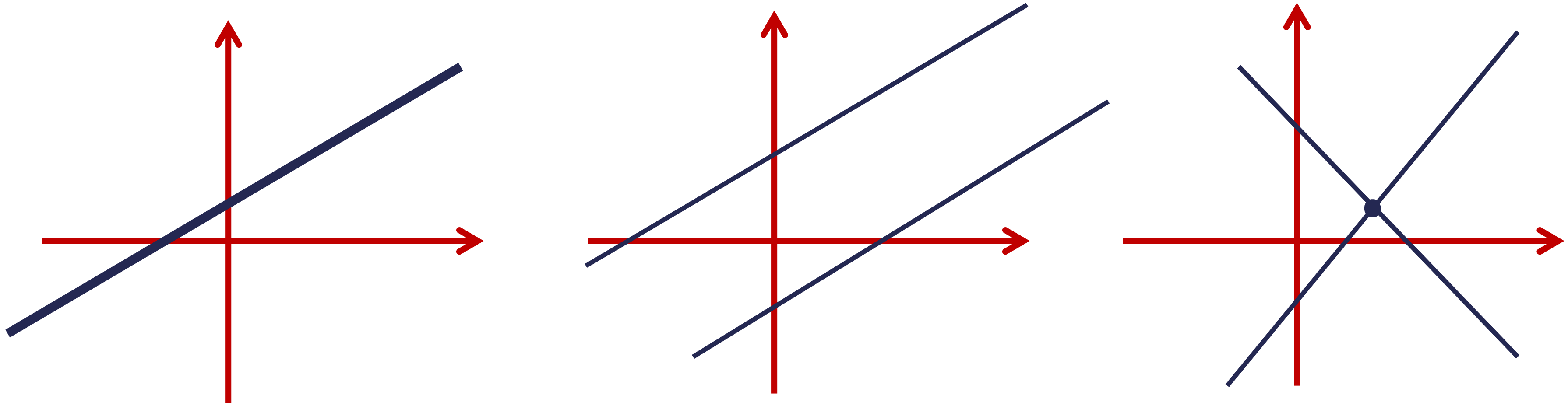


两条直线交于（3，1）点

线性方程组解的可能情况

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases}$$

不会出现交于两个点或三个点的情况



几何 两直线重合，无穷多个交点
代数 方程组有无穷多解

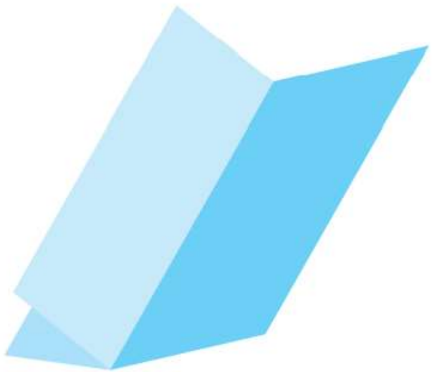
两直线平行，无交点
方程组无解

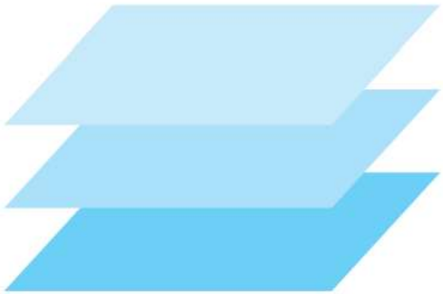
两直线交点于一点
方程组有唯一解

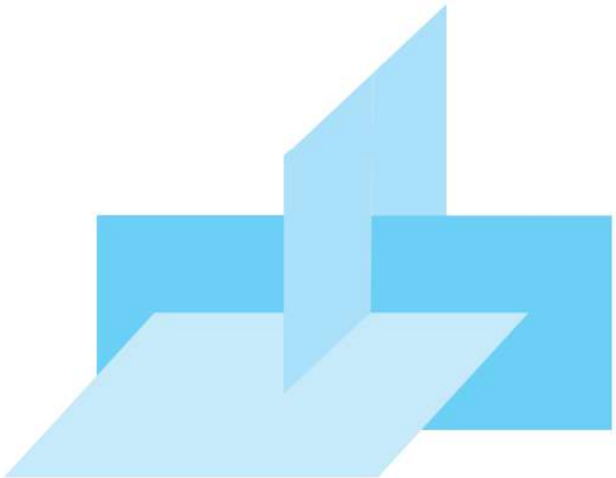
线性方程组解的可能情况

$$\begin{cases} \boxed{a_{11}x + a_{12}y + a_{13}z = b_1} \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases}$$

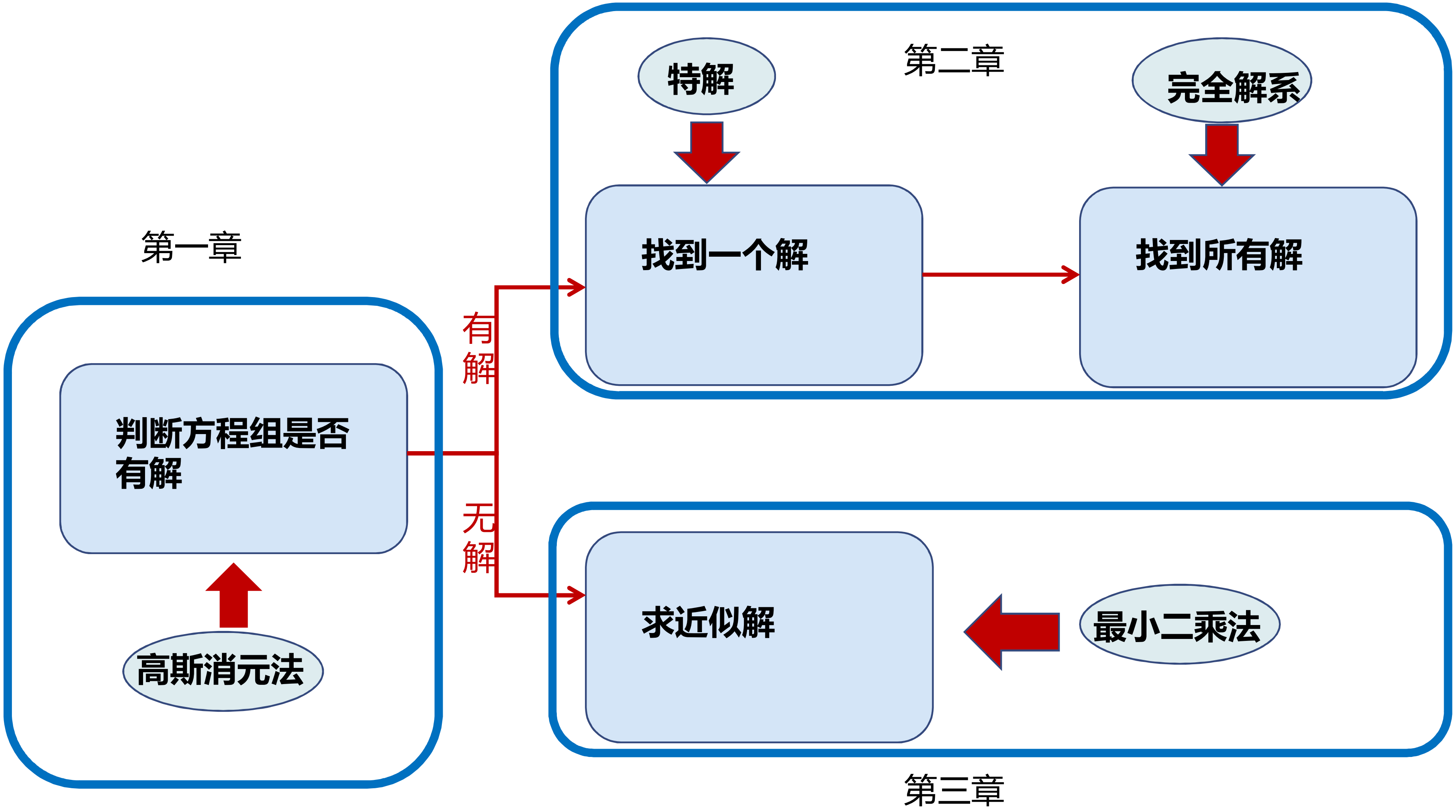
不会出现交于两个点，或交于一条直线与直线外一点的情况。
为什么？方程组的解集有什么性质？







几何	→	交于一条线或一个平面	无交点	交于一点
代数	→	方程组有无穷多解	方程组无解	方程组有唯一解


第一章
判断方程组是否有解
高斯消元法
有解
无解
第二章
特解
完全解系
找到一个解
找到所有解
求近似解
最小二乘法
第三章

π

1.2 三种视角看待方程组问题与映射的基本概念

(a) 行的视角

(b) 列的视角

(c) 线性映射的视角

— 将回顾映射的基本概念

还可以怎么理解?

（a）行的视角

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

$\mathbb{R}^n$中m个超平面的交集

$$\widetilde{\boldsymbol{a}_1} = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix}, \widetilde{\boldsymbol{a}_2} = \begin{bmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{bmatrix}, \ldots, \widetilde{\boldsymbol{a}_m} = \begin{bmatrix} a_{m1} \\ a_{m2} \\ \vdots \\ a_{mn} \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

如果$b_1 = b_2 = \cdots = b_m = 0$，$\boldsymbol{x} \in \mathbb{R}^n$为方程组的解当且仅当$\boldsymbol{x}$与$\widetilde{\boldsymbol{a}_1}, \ldots, \widetilde{\boldsymbol{a}_m}$均正交

(b) 列的视角

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

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

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

（b）列的视角

令$\boldsymbol{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$，$\boldsymbol{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots,$ $\boldsymbol{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m$

方程组有解当且仅当$\boldsymbol{b} \in Span(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots \boldsymbol{a}_n)$

一个解即是一个将$\boldsymbol{b}$写成$\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$的线性组合方式

（c）线性映射的视角

需要以下三方面的准备：

1. 映射的基础知识
2. 矩阵的概念
3. 线性映射与矩阵

我们在下面几节分别介绍.

映射的基本概念

X, Y为两个非空集合，一个X到Y的**映射**是一个法则：

将X中每个元素x对应到Y中唯一一个元素$f(x)$.

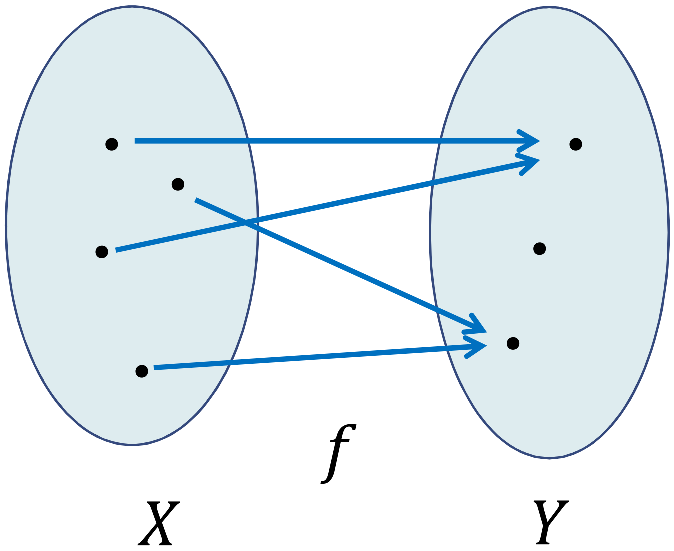
$$f: X \to Y, \qquad x \mapsto f(x).$$

定义域　陪域　　原象　　象

X称为f的**定义域**, Y称为f的**陪域**

$y = f(x)$称为（在f下）的**象**, x称为y (在f下）的**原象**

Y的子集$f(X) = \{f(x) \mid x \in X\}$称为$f$的**值域**

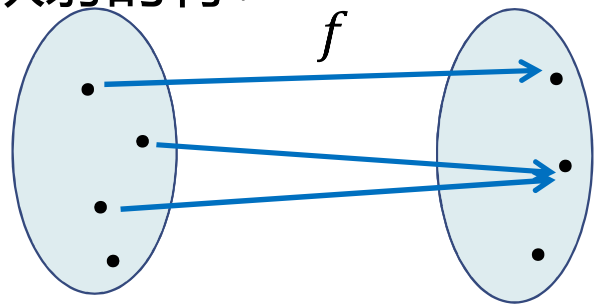


多选题 1分

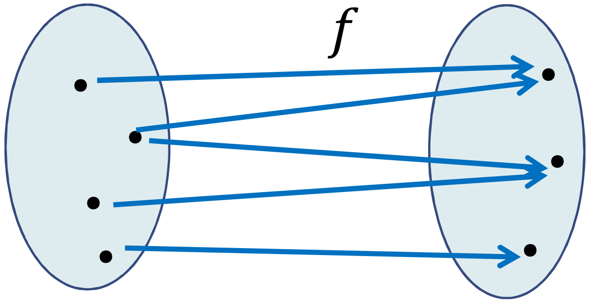
设置

下列**不是**映射的有：

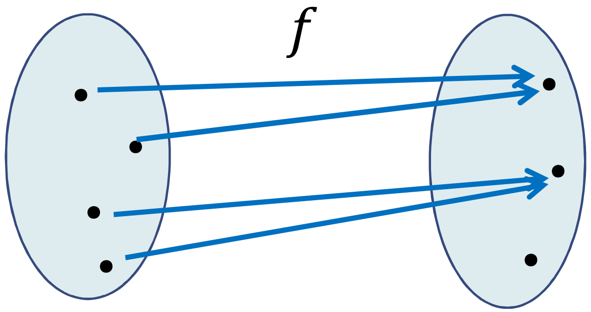
A



B

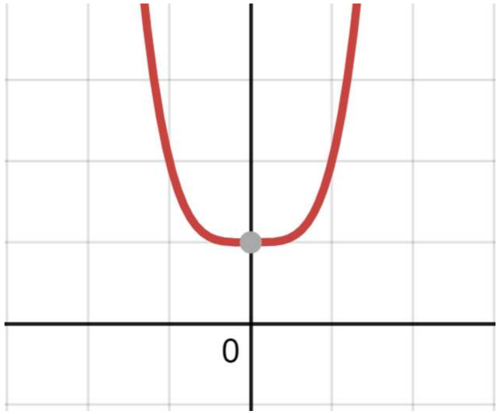


C



提交

例:



(1) 函数$f(a) = a^4 + 1: \mathbb{R} \to \mathbb{R}$给出一个映射

f的定义域=$\mathbb{R}$, 陪域=$\mathbb{R}$, 值域=$\mathbb{R}_{\geq 1}$.

(2) 函数集上的函数（映射）：

令$X = \{f: \mathbb{R} \to \mathbb{R}\}, Y = \mathbb{R}$. 固定一个$a \in \mathbb{R}$,

定义一个映射

在a处取值

$$a: X \to Y, f \mapsto a(f) = f(a)$$

映射a的定义域=X，值域=Y.

例：线性函数

令$\boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n, \mathbb{R}^n$上的内积给出一个函数

$$f_{\boldsymbol{v}}: \mathbb{R}^n \to \mathbb{R}, \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto f_{\boldsymbol{v}}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = v_1 x_1 + \cdots + v_n x_n.$$

也就是说$f_{\boldsymbol{v}}(\boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{w}$.

由内积的性质，$f_{\boldsymbol{v}}$满足两条容易验证的重要性质

$$f_{\boldsymbol{v}}(\boldsymbol{w}_1 + \boldsymbol{w}_2) = f_{\boldsymbol{v}}(\boldsymbol{w}_1) + f_{\boldsymbol{v}}(\boldsymbol{w}_2)$$ （分配律）

$$f_{\boldsymbol{v}}(c\boldsymbol{w}) = c f_{\boldsymbol{v}}(\boldsymbol{w}), c \in \mathbb{R}$$ （数乘交换）

线性

因此，$f_{\boldsymbol{v}}$也称为一个**线性函数**

线性映射

重点研究对象

映射$f: \mathbb{R}^n \to \mathbb{R}^m$称为线性映射, 如果$f$满足:

$$f(\boldsymbol{v}_1 + \boldsymbol{v}_2) = f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2), \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^n,$$

$$f(c\boldsymbol{v}) = cf(\boldsymbol{v}), \forall c \in \mathbb{R}, \forall \boldsymbol{v} \in \mathbb{R}^n.$$

因此，上例中线性函数$f_{\boldsymbol{v}}$是$\mathbb{R}^n$到$\mathbb{R}$的线性映射

定义线性映射的两个条件可以等价地写成

$$f(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = c_1 f(\boldsymbol{v}_1) + c_2 f(\boldsymbol{v}_2), \forall c_1, c_2 \in \mathbb{R}, \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^n.$$

若$f: \mathbb{R}^n \to \mathbb{R}^m$为线性映射，则$f(\mathbf{0}) =?$ $\quad f(\mathbf{0}_{\mathbb{R}^n}) = \mathbf{0}_{\mathbb{R}^m}$

更多的关于映射的定义

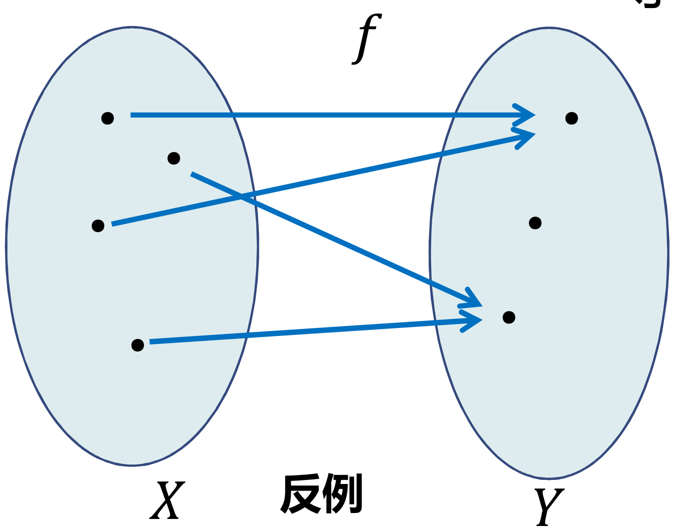
由于 f 不一定可逆，
单独的f^{-1}无意义

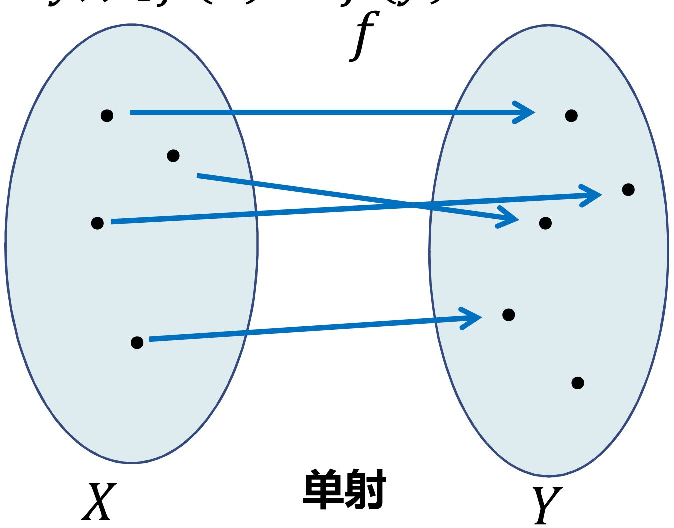
$f: X \to Y$为一映射.

S是Y的子集,记S的原象集为$f^{-1}(S) = \{x \in X \mid f(x) \in S\}$.

f称为**单射**，如果f满足若$f(x) = f(y)$,则$x = y$.

等价地，若$x \neq y$,则$f(x) \neq f(y)$.



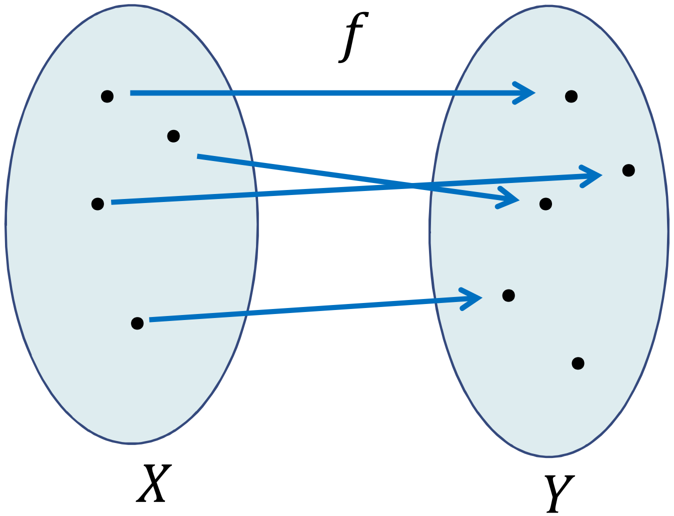


π

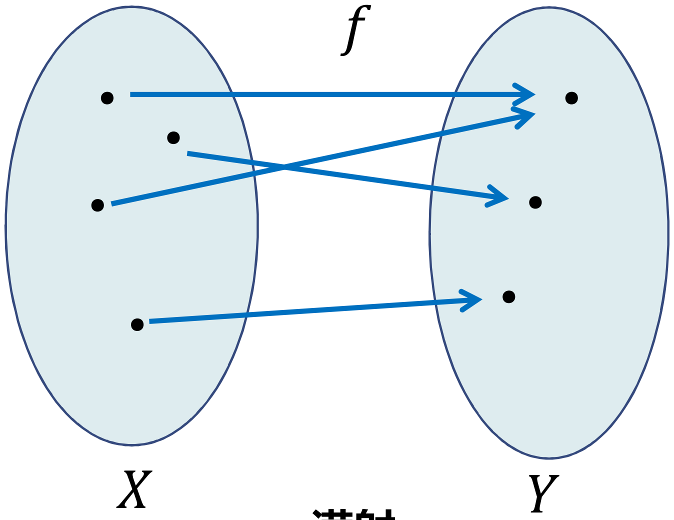
更多的关于映射的定义

$f: X \to Y$为一映射.

f称为**满射**，如果f满足$f(X) = Y$,即值域等于陪域。



反例

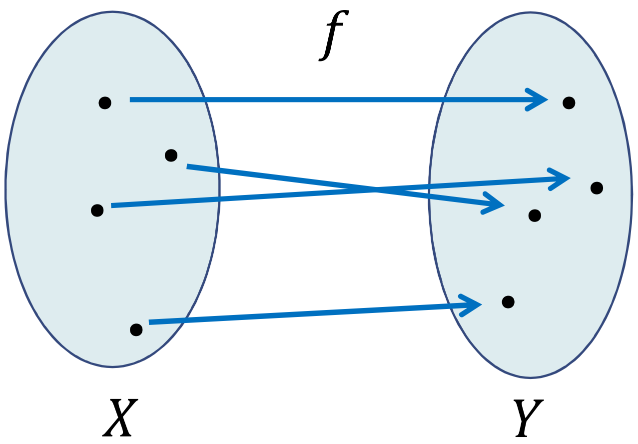


满射

更多的关于映射的定义

$f: X \to Y$为一映射.

如果f即单且满，则称f为**双射**



双射

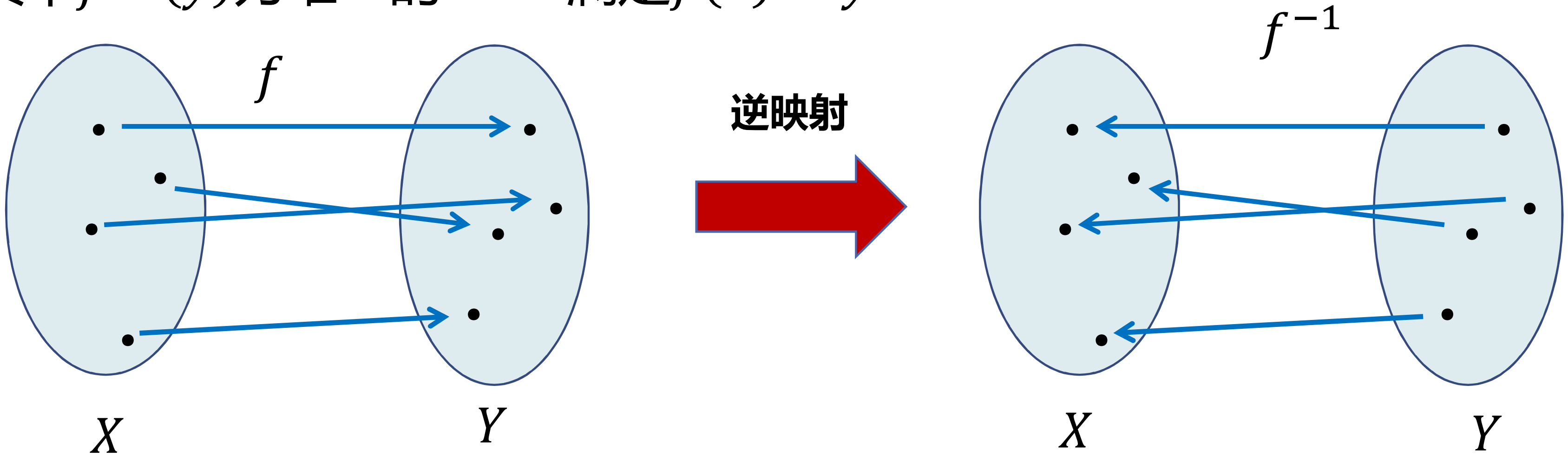
更多的关于映射的定义

如果f为双射，我们定义f的**逆映射**

$$f^{-1}: Y \to X,\ y \mapsto f^{-1}(y),$$

$f^{-1}(y)$为y在f下的原象

其中$f^{-1}(y)$为唯一的$x \in X$满足$f(x) = y$.

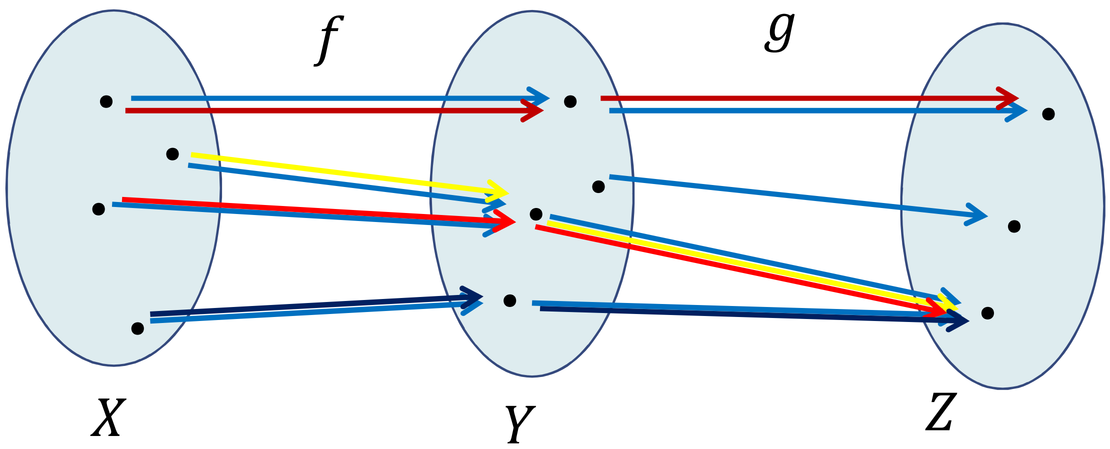


更多的关于映射的定义

$f,g: X \to Y$为两个映射，称f与g相等如果$f(x) = g(x), \forall x \in X$.

设$f: X \to Y, g: Y \to Z$,定义**复合映射**（g复合f):

$$g \circ f: X \to Z, x \mapsto g\big(f(x)\big).$$

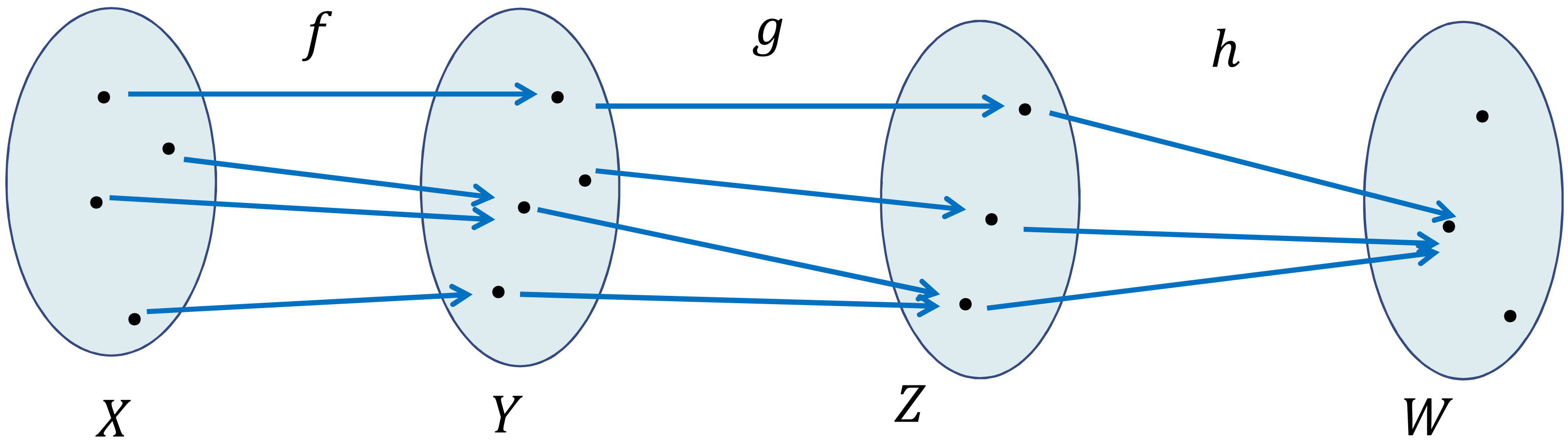


f的陪域等于g的定义域

容易验证，映射的复合满足：

如果$f: X \to Y, g: Y \to Z, h: Z \to W$,则

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

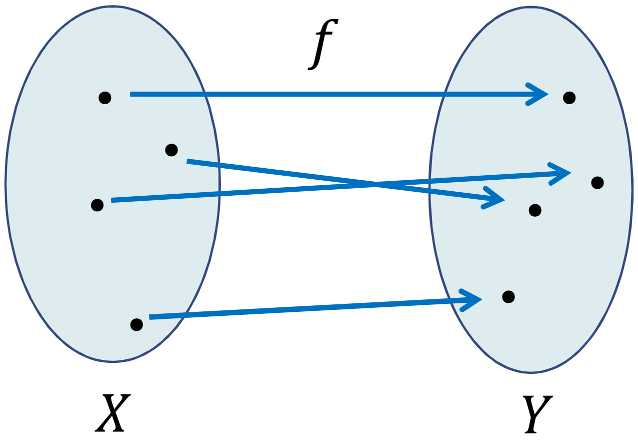


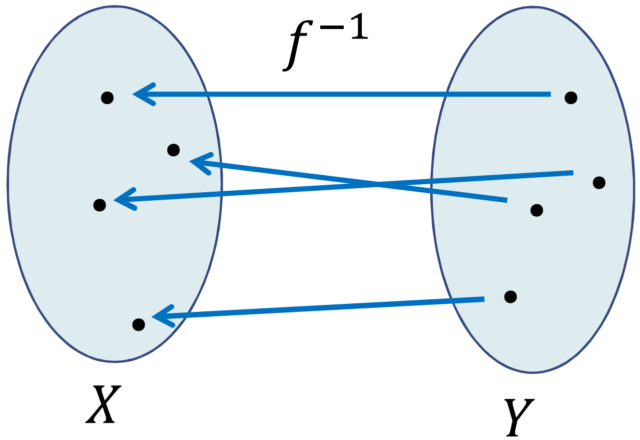
$$\big(h \circ (g \circ f)\big)(x) = \big((h \circ g) \circ f\big)(x) = h(g(f(x)))$$

由定义容易得到

如果$f: X \to Y$为双射, $f^{-1}: Y \to X$为f的逆，则

$$f^{-1} \circ f(x) = x, \forall x \in X, \qquad f \circ f^{-1}(y) = y, \forall y \in Y.$$





1.3 矩阵的定义（★）

主要内容：

（a）矩阵的定义

（b）矩阵在列向量上的作用

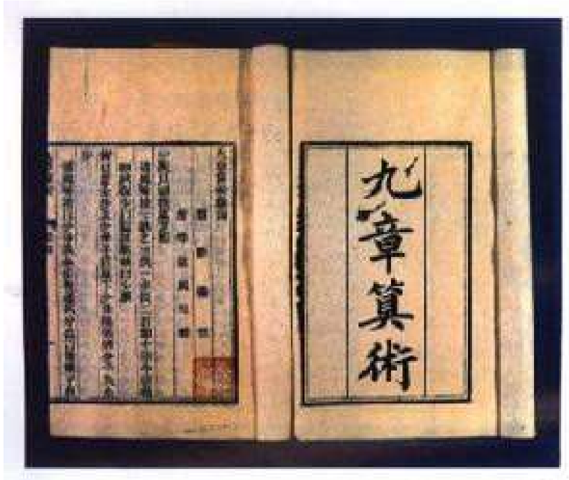
（c）线性映射与矩阵的关系

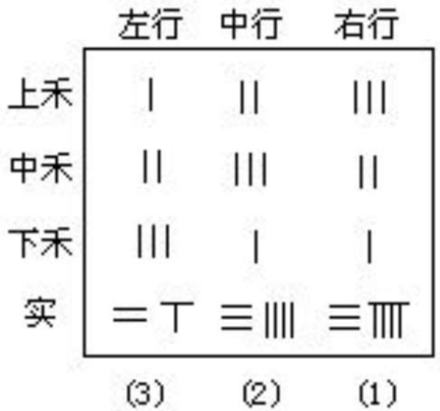
（d）矩阵在列向量上的作用举例

(a) 矩阵的定义

$$\begin{cases} 3x + 2y + z = 39 \\ 2x + 3y + z = 34 \\ x + 2y + 3z = 26 \end{cases}$$

3	2	1	39
2	3	1	34
1	2	3	26





《九章算术》算筹图

矩阵的定义:

$m, n \in \mathbb{Z}_{\geq 1}$,一个$m \times n$阶矩阵即是将$mn$个数$a_{11}, \dots, a_{mn}$按如下方式排列

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

如果$m = n$,称A为m阶方阵

$m \times n$阶矩阵的全体记为$M_{m \times n}(\mathbb{R})$

例

当$n=1$时，$m\times 1$阶矩阵

$$\boldsymbol{a}=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

列向量

为$\mathbb{R}^m$中的一个向量。

不加说明，向量默认为**列向量**

例:

$m=1$, $1\times n$阶矩阵

令$\boldsymbol{a}=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}\in M_{n\times 1}(\mathbb{R})$

称为$\boldsymbol{a}$的转置
(transpose)

定义$\boldsymbol{a}^T=[a_1,a_2,\dots,a_n]$ 行向量

$\boldsymbol{a}^T$为$1\times n$阶矩阵

$\boldsymbol{a},\boldsymbol{b},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ ...表示列向量 $\boldsymbol{a}^T,\boldsymbol{b}^T,\boldsymbol{u}^T,\boldsymbol{v}^T,\boldsymbol{w}^T$ ...表示行向量

例：线性方程组给出的矩阵

$$\begin{cases} 3x + 2y + z = 39 \\ 2x + 3y + z = 34 \\ x + 2y + 3z = 26 \end{cases}$$

$A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 2 & 3 \end{bmatrix}$称为方程组的**系数矩阵**

$\left[A \,\middle|\, \boldsymbol{b}\right] = \left[\begin{array}{ccc|c} 3 & 2 & 1 & 39 \\ 2 & 3 & 1 & 34 \\ 1 & 2 & 3 & 26 \end{array}\right]$称为方程组的**增广矩阵**

例：线性方程组给出的矩阵

$$\begin{cases} 1x_1 + 2x_2 + 3x_3 + 4x_4 = 2 \\ 5x_1 + 6x_2 + 7x_3 + 8x_4 = 3 \\ 9x_1 + 1x_2 + 2x_3 + 3x_4 = 3 \end{cases}$$

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 1 & 2 & 3 \end{bmatrix}$$ 系数矩阵（3×4阶）

$$\left[A \mid \boldsymbol{b} \right] = \left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 2 \\ 5 & 6 & 7 & 8 & 3 \\ 9 & 1 & 2 & 3 & 3 \end{array}\right]$$ 增广矩阵

一些特殊的矩阵：

零矩阵$O_{m\times n}$是系数全为零的$m \times n$阶矩阵

$$I_n = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \in M_{n\times n}(\mathbb{R})$$ 称为n阶**恒等矩阵**或**单位矩阵**

$$D(a_1, \dots, a_n) = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix} \in M_{n\times n}(\mathbb{R})$$ 称为对角矩阵

上(下)三角阵:

n阶上三角方阵$U = \begin{bmatrix} a_1 & * & * \\ & \ddots & * \\ & & a_n \end{bmatrix}$ n阶下三角方阵$L = \begin{bmatrix} a_1 & & \\ * & \ddots & \\ * & * & a_n \end{bmatrix}$

Upper triangular Lower triangular

例如:$\begin{bmatrix} 1 & 2 & 2 \\ & 2 & 1 \\ & & 2 \end{bmatrix}$, $\begin{bmatrix} 2 & & \\ 1 & 2 & \\ 1 & 1 & 1 \end{bmatrix}$

(b) 矩阵乘以列向量：

$$\begin{cases} 1x_1 + 2x_2 + 3x_3 + 4x_4 = 2 \\ 5x_1 + 6x_2 + 7x_3 + 8x_4 = 3 \\ 9x_1 + 1x_2 + 2x_3 + 3x_4 = 3 \end{cases}$$

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 1 & 2 & 3 \end{bmatrix}}_{\text{4列}} \qquad \left.\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right\} \text{4行}$$

系数矩阵 A $\qquad$ $\boldsymbol{x}$

定义:

令$A=(a_{ij})_{1\le i\le m,1\le j\le n}$为一个$m\times n$阶矩阵，$\boldsymbol{x}\in\mathbb{R}^n$，定义

$$\widetilde{\boldsymbol{a}_1}\cdot\boldsymbol{x}$$

$$\underset{A}{\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mn} \end{bmatrix}} \cdot \underset{\boldsymbol{x}}{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}} = \begin{bmatrix} \boxed{a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n} \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_x + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

这是$m\times n$阶矩阵A与$n\times 1$阶矩阵$\boldsymbol{x}$的乘积，

也称一个$m\times n$阶A**作用**在一个$n\times 1$阶矩阵$\boldsymbol{x}$上得到一个$m\times 1$阶矩阵

线性方程组的矩阵写法：

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mn} \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_{\boldsymbol{x}} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

线性方程组问题的左侧

线性方程组的矩阵写法：

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

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mn} \end{bmatrix} \qquad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m$$

线性方程组写成为 $\boxed{A\boldsymbol{x} = \boldsymbol{b}.}$

例:

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \in \mathbb{R}^m = M_{m\times 1}(\mathbb{R}), \quad c \in \mathbb{R} = M_{1\times 1}(\mathbb{R}).$$

$$\boldsymbol{a}c = \begin{bmatrix} a_1 c \\ a_2 c \\ \vdots \\ a_n c \end{bmatrix} = c\boldsymbol{a}$$

矩阵乘法

向量的数乘

例：$(1 \times n)$乘以$(n \times 1)$

$$\tilde{\boldsymbol{a}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n \quad \tilde{\boldsymbol{a}}^T = [a_1, a_2, \dots, a_n] \in M_{1\times n}(\mathbb{R}) \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in M_{n\times 1}(\mathbb{R})$$

$\tilde{\boldsymbol{a}}$与$\boldsymbol{x}$的点积

$$\tilde{\boldsymbol{a}}^T \boldsymbol{x} = [a_1, a_2, \dots, a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1 + a_2x_2 + \cdots + a_nx_n \in \mathbb{R}$$

从今以后，我们使用一个行向量乘以一个列向量的形式，避免使用两个列向量的点积