第二章 向量和矩阵基础

第 2 讲 向量与矩阵的概念和运算

黄定江

DaSE @ ECNU

djhuang@dase.ecnu.edu.cn

1. 2.1 向量与矩阵的基本概念：数据表示的观点
2. 2.2 向量和矩阵的运算

2.1.1 表示文本

例 1

下面是纽约时报在 *2010* 年 *12* 月 *7* 日的四则新闻标题：

(a) Suit Over Targeted Killing in Terror Case Is Dismissed. A federal judge on Tuesday dismissed a lawsuit that sought to block the United States from attempting to kill an American citizen, Anwar Al-Awlaki, who has been accused of aiding Al Qaeda.

(b) In Tax Deal With G.O.P, a Portent for the Next 2 Years. President Obama made clear that he was willing to alienate his liberal base in the interest of compromise. Tax Deal suggests new path for Obama. President Obama agreed to a tentative deal to extend the Bush tax cuts, part of a package to keep jobless aid and cut pay roll taxes.

2.1.1 表示文本

(c) Obama Urges China to Check North Koreans. In a frank discussion, President Obama urged China's president to put the North Korean government on a tighter leash after a series of provocations.

(d) Top Test Scores From Shanghai Stun Educators. With China's debut in international standardized testing Shanghai students have surprised experts by outscoring counterparts in dozens of other countries.

2.1.1 表示文本

- 假设一本字典 V（dict），字典 V中的单词为 {aid, kill, deal, president, tax, china}，我们想知道每则新闻标题中的单词在字典中出现的频率。
- 非常容易可以看出，在新闻 (a) 中 aid 共出现了 1 次，kill 共出现了 2 次，而字典 V中的其它单词并没有出现，通过一个一维数组表示这一结果，即

$$\boldsymbol{a} = (1, 2, 0, 0, 0, 0).$$

- 将 $\boldsymbol{a}$ 中每个单词除以总共出现的次数，便可以得到这则新闻标题在字典 $\boldsymbol{V}$ 中出现的相对频率，即

$$\boldsymbol{a}^{'} = (\frac{1}{3}, \frac{2}{3}, 0, 0, 0, 0).$$

2.1.1 表示文本

- 将其它三则新闻也进行同样的处理，即分别为

$$
\begin{aligned}
\boldsymbol{b}' &= (\frac{1}{10}, 0, \frac{3}{10}, \frac{1}{5}, \frac{2}{5}, 0),\\
\boldsymbol{c}' &= (0, 0, 0, \frac{1}{2}, 0, \frac{1}{2}),\\
\boldsymbol{d}' &= (0, 0, 0, 0, 0, 1).
\end{aligned}
$$

2.1.2 向量的定义

定义 1

设 $\mathbb{K}$ 是由一些数组成的集合，如果 $\mathbb{K}$ 中包含 0 与 1，并且 $\mathbb{K}$ 中任意两个数的和，差，积，商（除数不为零）仍在 $\mathbb{K}$ 中，则称 $\mathbb{K}$ 是一个**数域**。

有理数集，实数集，复数集都是数域，他们分别称为有理数域 $\mathbb{Q}$，实数域 $\mathbb{R}$，复数域 $\mathbb{C}$.

定义 2

由数域 $\mathbb{K}$ 中的 n 个数组成的有序数组 $(a_1, a_2, \cdots, a_n)$ 称为 $\mathbb{K}$ 上的 n 维**向量**。即

$$\boldsymbol{a} = \left(a_1, a_2, a_3, \vdots, a_n\right)$$

其中第 i 个数 a_i 称为 $\boldsymbol{a}$ 的第 i 个分量.

上述定义的是行向量。在实际使用中，我们通常使用列向量。

我们可以使用**转置**把列向量转换成行向量，或者把行向量转换成列向量。

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = \left(a_1, a_2, a_3, \ldots, a_n\right)^{\mathrm{T}}; \left(a_1, a_2, a_3, \ldots, a_n\right) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix}^{\mathrm{T}}$$

2.1.2 向量的定义

上面的几个数组用向量表示即为：

$$
\begin{aligned}
\boldsymbol{a}' &= (\frac{1}{3}, \frac{2}{3}, 0, 0, 0, 0)^{\mathrm{T}}. \\
\boldsymbol{b}' &= (\frac{1}{10}, 0, \frac{3}{10}, \frac{1}{5}, \frac{2}{5}, 0)^{\mathrm{T}}, \\
\boldsymbol{c}' &= (0, 0, 0, \frac{1}{2}, 0, \frac{1}{2})^{\mathrm{T}}, \\
\boldsymbol{d}' &= (0, 0, 0, 0, 0, 1)^{\mathrm{T}}.
\end{aligned}
$$

2.1.2 向量的定义

在科学和工程中遇到的向量可以分为以下三种：

(1) 物理向量：泛指既有幅值，又有方向的物理量，如速度、加速度、位移等。

(2) 几何向量：为了将物理向量可视化，常用带方向的（简称“有向”）线段表示之。这种有向线段称为几何向量。例如，$\boldsymbol{v} = \overrightarrow{\boldsymbol{AB}}$ 表示的有向线段，其起点为 $\boldsymbol{A}$，终点为 $\boldsymbol{B}$。

(3) 代数向量：几何向量可以用代数形式表示。例如，若平面上的几何向量 $\boldsymbol{v} = \overrightarrow{\boldsymbol{AB}}$ 的起点坐标 $A = (a_1, a_2)$，终点坐标 $\boldsymbol{B} = (b_1, b_2)$，则该几何向量可以表示为代数形式 $\boldsymbol{v} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix}$。这种用代数形式表示的几何向量称为代数向量。

2.1.2 向量的定义

下图归纳了向量的分类。

$$\text{向量}\begin{cases}\text{物理向量}\\ \text{代数向量}\begin{cases}\text{常数向量}\\ \text{函数向量}\\ \text{随机向量}\end{cases}\\ \text{几何向量}\end{cases}$$

2.1.2 向量的定义

根据元素取值种类的不同，代数向量又可分为以下三种：

(1) 常数向量：向量的元素全部为实常数或者复常数，如 $\boldsymbol{a}=[1,5,4]^{\mathrm{T}}$ 等。

(2) 函数向量：向量的元素包含了函数值，如 $\boldsymbol{x}=[1,x^2,\cdots,x^n]^{\mathrm{T}}$ 等。

(3) 随机向量：向量的元素为随机变量或随机过程，如 $\boldsymbol{x}(n)=[x_1(n),\cdots,x_m(n)]^{\mathrm{T}}$，其中 $x_1(n),\cdots,x_m(n)$ 是 m 个随机过程或随机信号。

2.1.3 表示文本向量集

例 2

在例1中，每则新闻标题都由一个 6 维向量所表示，那么这四则新闻标题组成的新闻集可以由 4 个这样的 6 维向量组成的向量集所表示，换言之，这个新闻集可以按列组成一个 6×4 的二维数组。即

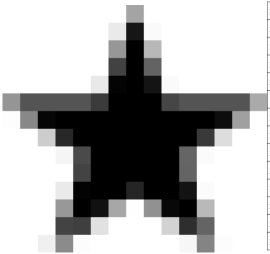
$$\boldsymbol{A}=\begin{pmatrix} \frac{1}{3} & \frac{1}{10} & 0 & 0 \\ \frac{2}{3} & 0 & 0 & 0 \\ 0 & \frac{3}{10} & 0 & 0 \\ 0 & \frac{1}{5} & \frac{1}{2} & 0 \\ 0 & \frac{2}{5} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{pmatrix}$$

2.1.4 表示灰度图像

例 3

在计算机中，如果把只保留图像的灰度，那么我们可以使用一个二维数组存储该图像，其中数组中的元素是图像中相应像素的强度值 (介于 0 至 255 之间”)。

下面左图，具有 15 个水平像素和 14 个垂直像素，二维数组如右图所示。



255	255	255	255	255	255	255	164	255	255	255	255	255	255	255
255	255	255	255	255	255	234	22	250	255	255	255	255	255	255
255	255	255	255	255	255	151	0	184	255	255	255	255	255	255
255	255	255	255	255	255	66	0	99	255	255	255	255	255	255
255	255	255	255	255	233	4	0	20	250	255	255	255	255	255
149	84	84	84	84	59	0	0	0	70	84	84	84	84	170
253	136	6	0	0	0	0	0	0	0	0	0	13	160	255
255	255	205	40	0	0	0	0	0	0	0	57	221	255	255
255	255	255	245	74	0	0	0	0	0	101	251	255	255	255
255	255	255	255	67	0	0	0	0	0	100	255	255	255	255
255	255	255	234	3	0	0	37	0	0	19	251	255	255	255
255	255	255	152	0	17	167	254	145	8	0	185	255	255	255
255	255	255	67	63	225	255	255	255	210	46	100	255	255	255
255	255	240	139	252	255	255	255	255	255	247	132	252	255	255

图 1: 图像的表示

2.1.5 矩阵的定义

定义 3

由数域 $\mathbb{K}$ 中的 $m \times n$ 个数 $a_{ij}(i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$ 排成 m 行、n 列的表

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

称为 $\mathbb{K}$ 上的 $m \times n$ **矩阵**, 记为 $\boldsymbol{A} = (a_{ij})_{m\times n}$ 或 $\boldsymbol{A}_{m\times n}$, 表中的每一个数都称为矩阵 $\boldsymbol{A}$ 的一个元素. 若 $m = n$, 则 $n \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})_{m\times n}$ 也称为 **n 阶方阵**.

2.1.5 矩阵的定义

- 从定义上可以看出，矩阵是一个二维数组，而向量是矩阵的一种特殊形式，因为它的 $m = 1$ 或者 $n = 1$
- 其中 $1 \times n$ 的矩阵称之为行向量，而 $m \times 1$ 的矩阵称之为列向量。

2.1.6 例子

例 4

某公司生产产品 $N_1, N_2, \ldots, N_n$，其需要的资源分别为 $R_1, R_2, \ldots, R_m$。为了生产一单位的 N_j 需要 a_{ij} 单位的 R_i，其中 $i = 1, \ldots, m; j = 1, \ldots, n$. 目的是找到最佳的生产计划。例如如果有 b_i 单位的 R_i 可供使用，那么应该生产多少 (x_j) 单位的 N_j 使得恰好用尽资源。

如果我们生产 $x_1, x_2, \ldots, x_n$ 单位的对应产品，我们一共需要 $a_{i1}x_1 + \ldots + a_{in}x_n$ 单位资源 R_i。最优生产计划 $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$，因此它必须满足方程组

$$
\begin{aligned}
a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + \ldots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m
\end{aligned}
$$

2.1.6 例子

上述方程组可以写出矩阵的形式

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{pmatrix}$$

用一个紧凑形式表示即为

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

2.2.1 向量的线性运算

向量之间的基本代数运算有两种：加法和数乘，统称为向量的线性运算。

定义 4

设向量 $\boldsymbol{a}=(a_1,a_2,\cdots,a_n)$,$\boldsymbol{b}=(b_1,b_2,\cdots,b_n)$，则向量 $(a_1+b_1,a_2+b_2,\cdots,a_n+b_n)$ 称为向量 $\boldsymbol{a}$ 与向量 $\boldsymbol{b}$ 的加法，记作 $\boldsymbol{a}+\boldsymbol{b}$，即

$$\boldsymbol{a}+\boldsymbol{b}=(a_1+b_1,a_2+b_2,\cdots,a_n+b_n).$$

2.2.1 向量的线性运算

向量之间的基本代数运算有两种：加法和数乘，统称为向量的线性运算。

定义 5

设向量 $\boldsymbol{a}=(a_1,a_2,\cdots,a_n)$，$k$ 是数域 $\mathbb{K}$ 中的数，则向量 $(ka_1,ka_2,\cdots,ka_n)$ 称为数 k 与向量 $\boldsymbol{a}$ 的乘积，简称**数乘**，记作 $k\boldsymbol{a}$，即

$$k\boldsymbol{a}=(ka_1,ka_2,\cdots,ka_n).$$

定理 1

向量的加法和数乘运算满足交换律、结合律和分配律

1. $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$,
2. $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$,
3. $\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a}$,
4. $\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}$,
5. $1\boldsymbol{a}=\boldsymbol{a}$,
6. $\lambda(\mu\boldsymbol{a})=(\lambda\mu)\boldsymbol{a}$,
7. $(\lambda+\mu)\boldsymbol{a}=\lambda\boldsymbol{a}+\mu\boldsymbol{a}$,
8. $\lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}$.

2.2.2 矩阵的加法

定义 6

设 $\boldsymbol{A}=(a_{ij})_{m\times n}, \boldsymbol{B}=(b_{ij})_{m\times n}$。令 $\boldsymbol{C}=(c_{ij})_{m\times n}$，其中 $c_{ij}=a_{ij}+b_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$，则称 $\boldsymbol{C}$ 为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和，记作 $\boldsymbol{C}=\boldsymbol{A}+\boldsymbol{B}$。

2.2.2 矩阵的加法

设两个 $m \times n$ 矩阵

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

$$\boldsymbol{A} + \boldsymbol{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

2.2.3 矩阵的数乘与转置

定义 7

$\boldsymbol{A}=(a_{ij})_{m\times n}, \lambda\in\mathbb{K}$，则 λ 与 $\boldsymbol{A}$ 的**数量乘积**或**标量乘积**定义为 $\lambda\boldsymbol{A}=(\lambda a_{ij})_{m\times n}$.

定义 8

设 $\boldsymbol{A}=(a_{ij})_{m\times n}$，把 $\boldsymbol{A}$ 的行、列互换所得到的矩阵称为 $\boldsymbol{A}$ 的**转置**，记作 $\boldsymbol{A}^{\mathrm{T}}$，或 $\boldsymbol{A}^{'}$，即 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{'}=(a_{ji})_{n\times m}$.

2.2.3 矩阵的数乘与转置

$$\lambda \boldsymbol{A} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{pmatrix} \boldsymbol{A}^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

2.2.4 矩阵的乘法

定义 9

$\boldsymbol{A}=(a_{ij})_{m\times r}, \boldsymbol{B}=(b_{ij})_{r\times n}$，令 $\boldsymbol{C}=(c_{ij})_{m\times n}$，其中 $c_{ij}=\sum_{k=1}^{r} a_{ik}b_{kj}(i=1,2,\cdots,m;j=1,2,\cdots,n)$，则 $\boldsymbol{C}$ 称为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积，记作 $\boldsymbol{C}=\boldsymbol{A}\boldsymbol{B}$。

2.2.4 矩阵的乘法

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mr} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rm} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nm} \end{pmatrix}$$

其中

$$c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{ir} b_{rj} = \sum_{k=1}^{r} a_{ik} b_{kj}$$

2.2.4 矩阵的乘法

定义 10

$\boldsymbol{A}=(a_{ij})_{m\times n}, \boldsymbol{x}=(x_1, x_2, ..., x_n)^{\mathrm{T}}$，令 $\boldsymbol{b}=(b_1, b_2, ..., b_m)$，其中 $b_i=\sum_{k=1}^{n} a_{ik}x_k(i=1,2,\cdots,m;)$，则 $\boldsymbol{b}$ 称为 $\boldsymbol{A}$ 与 $\boldsymbol{x}$ 的乘积，记作 $\boldsymbol{b}=\boldsymbol{A}\boldsymbol{x}$。

2.2.4 矩阵的乘法

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

其中

$$b_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = \sum_{k=1}^{n} a_{ik}x_k$$

2.2.5 分块矩阵的乘法

定义 11

将高阶矩阵划分为若干个低阶矩阵块，每个块称为**分块矩阵**。处理高阶矩阵的运算时，可以转换为低阶矩阵的运算。

例： 设矩阵 $\boldsymbol{A}$

$$\boldsymbol{A}=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ -1 & 2 & 1 & 0\\ 1 & 1 & 0 & 1\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_2 & \boldsymbol{O}\\ \boldsymbol{A}_1 & \boldsymbol{I}_2\end{pmatrix}$$

其中，$\boldsymbol{I}_2$ 表示 2×2 的单位矩阵，$\boldsymbol{O}$ 表示零矩阵，而

$$\boldsymbol{A}_1=\begin{pmatrix}-1 & 2\\ 1 & 1\end{pmatrix},$$

2.2.5 分块矩阵的乘法

有另一矩阵 $\boldsymbol{B}$

$$\boldsymbol{B}=\begin{pmatrix}1 & 0 & 3 & 2\\ -1 & 2 & 0 & 1\\ 1 & 0 & 4 & 1\\ -1 & -1 & 2 & 0\end{pmatrix}=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{pmatrix}$$

其中，

$$\boldsymbol{B}_{11}=\begin{pmatrix}1 & 0\\ -1 & 2\end{pmatrix},\boldsymbol{B}_{12}=\begin{pmatrix}3 & 2\\ 0 & 1\end{pmatrix}$$

$$\boldsymbol{B}_{21}=\begin{pmatrix}1 & 0\\ -1 & -1\end{pmatrix},\boldsymbol{B}_{22}=\begin{pmatrix}4 & 1\\ 2 & 0\end{pmatrix}$$

2.2.5 分块矩阵的乘法

在计算 $\boldsymbol{AB}$ 时，把 $\boldsymbol{A},\boldsymbol{B}$ 都看成是由这些小矩阵组成的，即按 2 维矩阵来计算，于是有

$$\boldsymbol{AB}=\begin{pmatrix}\boldsymbol{I}_2 & \boldsymbol{O}\\ \boldsymbol{A}_1 & \boldsymbol{I}_2\end{pmatrix}\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{pmatrix}$$

$$=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12}\\ \boldsymbol{A}_1\boldsymbol{B}_{11}+\boldsymbol{B}_{21} & \boldsymbol{A}_1\boldsymbol{B}_{12}+\boldsymbol{B}_{22}\end{pmatrix}=\begin{pmatrix}1 & 0 & 3 & 2\\ -1 & 2 & 0 & 1\\ -2 & 4 & 1 & 1\\ -1 & 1 & 5 & 3\end{pmatrix}$$

2.2.6 逆

设 $\boldsymbol{A}$ 是数域 $\mathbb{K}$ 上的 $n \times n$ 矩阵，如果存在 $\mathbb{K}$ 上的 $n \times n$ 矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$, 则称 $\boldsymbol{A}$ 为**可逆矩阵**，简称 $\boldsymbol{A}$ 可逆，而 $\boldsymbol{B}$ 则称为 $\boldsymbol{A}$ 的**逆矩阵**，记作 $\boldsymbol{A}^{-1}$，即 $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$。

2.2.6 逆

值得注意的是，如果 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{-1}$ 也可逆，且 $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$。进而可知，若 $\boldsymbol{A}$ 可逆，则其逆矩阵是唯一的。

2.2.6 逆

考虑一个 2×2 矩阵：

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

则

$$(\boldsymbol{A}\boldsymbol{I}) = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

即

$$\boldsymbol{A}^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

2.2.7 广义逆

不是所有的矩阵都有逆矩阵，只有满秩的**方阵**才有逆矩阵，即 $\mathrm{rank}(\boldsymbol{A}) = n$，那么对于不是方阵，不满秩矩阵的情况，数学中也引申出类似于逆矩阵的定义：

定理 2

设 $\boldsymbol{A} = (a_{ij})_{m\times n}$ 是数域 $\mathbb{R}$ 上的矩阵，则方程组

$$\begin{cases} \boldsymbol{AXA} = \boldsymbol{A}, \\ \boldsymbol{XAX} = \boldsymbol{X}, \\ \boldsymbol{AX} = (\boldsymbol{AX})^{\mathrm{T}}, \\ \boldsymbol{XA} = (\boldsymbol{XA})^{\mathrm{T}} \end{cases} \tag{2}$$

有唯一解.

2.2.7 广义逆

定理中方程组 (2) 的唯一解称为矩阵 $\boldsymbol{A}$ 的**广义逆矩阵**, 记作 $\boldsymbol{A}^{+}$.

若 $\boldsymbol{A}$ 是非奇异方阵，则 $\boldsymbol{A}^{-1}$ 满足方程组，故 $\boldsymbol{A}^{+}=\boldsymbol{A}^{-1}$. 因此, 矩阵 $\boldsymbol{A}$ 的广义逆矩阵 $\boldsymbol{A}^{+}$ 可视为非奇异方阵 $\boldsymbol{A}$ 的逆 $\boldsymbol{A}^{-1}$ 的推广

通常情况下有 $\boldsymbol{A}^{+}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}$.