线性代数 Week 2

线性代数 (II)

Lecturer: 杨启哲 Last modified: 2025 年 2 月 11 日

1 线性方程组

1.1 求解线性方程组

考察线性方程组

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

如图1所示，每个方程定义了 $\mathbb{R}^2$ 上的一条直线。因此方程组(1)的解就是两条直线的交点。这是所谓方程组的**行观点**。

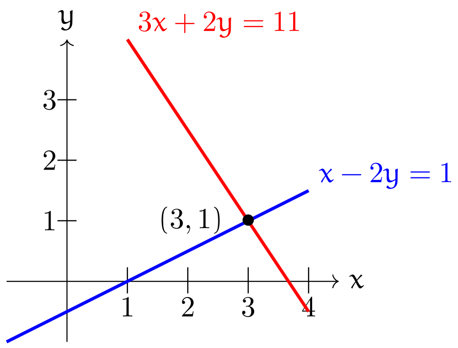


图 1　从行的观点理解线性方程组

我们还能从线性组合的观点来理解方程组求解。方程组(1)等价于

$$x\begin{bmatrix}1\\3\end{bmatrix} + y\begin{bmatrix}-2\\2\end{bmatrix} = \begin{bmatrix}1\\11\end{bmatrix}.$$

图2给出了向量 $(1,3)$ 和 $(-2,2)$ 的一个等于 $(1,11)$ 的线性组合，其系数就是方程组的解。这是方程组的**列观点**。

方程组的矩阵表示　我们可以将(1)写成**矩阵乘法**的形式：

$$\begin{bmatrix}1 & -2\\3 & 2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}1\\11\end{bmatrix} \tag{2}$$

矩阵 $\begin{bmatrix}1 & -2\\3 & 2\end{bmatrix}$ 称为线性方程组(1)的**系数矩阵**。因为(2)等价于(1)，所以

$$\begin{bmatrix}1 & -2\\3 & 2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x-2y\\3x+2y\end{bmatrix} = \begin{bmatrix}(1,-2)\cdot(x,y)\\(3,2)\cdot(x,y)\end{bmatrix}.$$

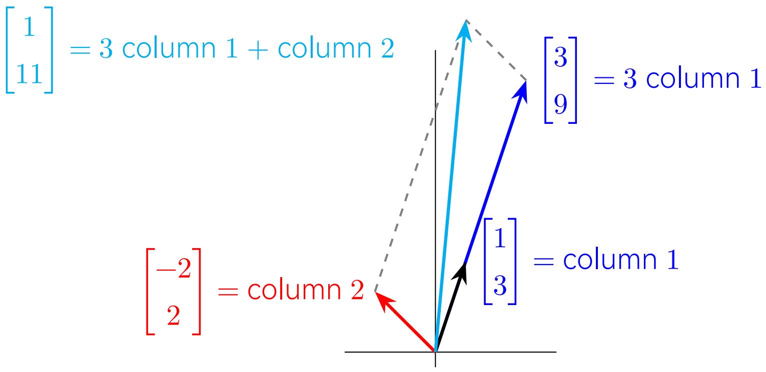


图 2　从列的观点理解线性方程组

我们后面会看到，一般的矩阵乘法都可由由向量的点积定义。在上面的例子中，当 $x = 3$ 且 $y = 1$ 时，我们就能得到

$$\begin{bmatrix}1 & -2\\3 & 2\end{bmatrix}\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}1\times 3 - 2\times 1\\3\times 3 + 2\times 1\end{bmatrix} = \begin{bmatrix}(1,-2)\cdot(3,1)\\(3,2)\cdot(3,1)\end{bmatrix} = \begin{bmatrix}1\\11\end{bmatrix}.$$

这恰好对应了 $x = 3$、$y = 1$ 是方程组(1)的解。

类似的我们可以讨论 $\mathbb{R}^3, \ldots,$ 至任意高维 $\mathbb{R}^n$ 中的线性方程组及其矩阵表示。例如对方程组

$$\begin{array}{rcrcrcr} 2x_1 & + & 4x_2 & - & 2x_3 & = & 2\\ 4x_1 & + & 9x_2 & - & 3x_3 & = & 8\\ -2x_1 & - & 3x_2 & + & 7x_3 & = & 10 \end{array} \tag{3}$$

从几何上看我们有

行观点	三个平面相交于一点
列观点	三个列向量的线性组合得到向量 $(2, 8, 10)$

而(3)的矩阵表示为

$$\underset{A}{\begin{bmatrix}2 & 4 & -2\\4 & 9 & -3\\-2 & -3 & 7\end{bmatrix}}\ \underset{\boldsymbol{x}}{\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}} = \underset{\mathbf{b}}{\begin{bmatrix}2\\8\\10\end{bmatrix}}$$

$A\boldsymbol{x} = \mathbf{b}$ 中的变元是 $\boldsymbol{x} = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$。很容易验证 $\begin{bmatrix}-1\\2\\2\end{bmatrix}$，即 $x = -1$，$y = 2$ 和 $z = 2$ 满足方程组(3)。所以我们需要定义 $A\boldsymbol{x}$ 使得

$$\begin{bmatrix}2 & 4 & -2\\4 & 9 & -3\\-2 & -3 & 7\end{bmatrix}\begin{bmatrix}-1\\2\\2\end{bmatrix} = \begin{bmatrix}2\\8\\10\end{bmatrix}.$$

矩阵和向量的乘法. 为了能将

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

写成

$$\underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}}_{\mathbf{b}},$$

对任意 $i \in [m]$ 我们令 $A\boldsymbol{x}$ 的第 i **个坐标**为

$$a_{i1}x_1 + \cdots + a_{in}x_n = \sum_{j \in [n]} a_{ij}x_j = (a_{i1}, \ldots, a_{in}) \cdot (x_1, \ldots, x_n).$$

那么

$$A\boldsymbol{x} = \begin{bmatrix} (a_{11}, \ldots, a_{1n}) \cdot (x_1, \ldots, x_n) \\ (a_{21}, \ldots, a_{2n}) \cdot (x_1, \ldots, x_n) \\ \vdots \\ (a_{m1}, \ldots, a_{mn}) \cdot (x_1, \ldots, x_n) \end{bmatrix}. \tag{4}$$

有了这个定义后，就可以将我们熟悉的**高斯消元**通过矩阵乘法来理解。

高斯消元的乘法解释. 考察如下的线性方程组

$$\begin{aligned} x - 2y &= 1 \\ 3x + 2y &= 11. \end{aligned} \tag{5}$$

通过一步简单消元就能得到

$$\begin{aligned} x - 2y &= 1 \quad (\text{将(5)中的第一个等式乘以 } 3) \\ 8y &= 8 \quad (\text{(5)中的第二个等式减去上面的结果就可以消去 } 3x). \end{aligned}$$

我们称这个新的方程组是**上三角**的，而我们能很容易地解出这样的上三角方程组。

图3给出了这个过程的几何解释。

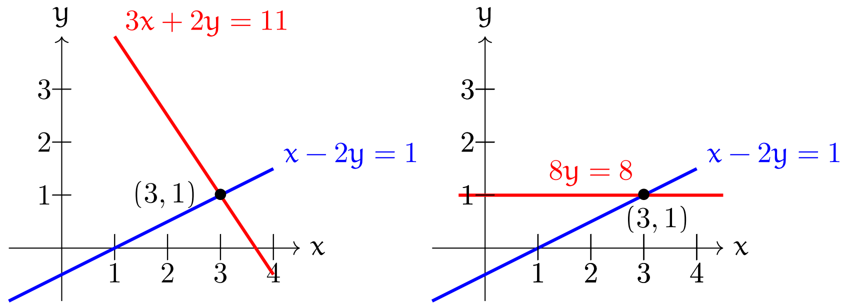
类似的，对方程组

$$\begin{aligned} 4x - 8y &= 4 \\ 3x + 2y &= 11 \end{aligned}$$

我们可以得到

$$\begin{aligned} 4x - 8y &= 4 \quad (\text{将第一个等式乘以 } 3/4) \\ 8y &= 8 \quad (\text{第二个等式减去上面的结果就可以消去 } 3x). \end{aligned} \tag{6}$$

每一步高斯消元都涉及两个重要的量



消元前　　　　消元后

图 3　高斯消元对应的几何变换

主元 (pivot)	用于消元的方程中第一个非零项
乘的系数	以待消去项的系数除以主元的系数

我们重新考察方程组(3)，即

$$\begin{aligned} 2x_1 + 4x_2 - 2x_3 &= 2 \\ 4x_1 + 9x_2 - 3x_3 &= 8 \\ -2x_1 - 3x_2 + 7x_3 &= 10 \end{aligned} \tag{7}$$

写成矩阵乘法 $Ax = \mathbf{b}$ 的形式：

$$\underbrace{\begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix}}_{\mathbf{b}}. \tag{8}$$

作为消元的第一步我们将(7)中的第二个等式减去 2 倍的第一个等式，得到

$$\begin{aligned} 2x_1 + 4x_2 - 2x_3 &= 2 \\ x_2 + x_3 &= 4 \\ -2x_1 - 3x_2 + 7x_3 &= 10 \end{aligned}. \tag{9}$$

其矩阵表示为

$$\begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ -2 & -3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix}. \tag{10}$$

因此从矩阵的角度看，我们将矩阵等式(8)转换成了矩阵等式(10)。而这个过程会由一个所谓的**消去矩阵**

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

刻画。如果我们将该矩阵乘以(8)的右侧，根据(4)就能得到(10)的右侧：

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix}.$$

而根据我们将引入的矩阵乘法（定义2.2），将消去矩阵乘以(8)的左侧系数矩阵，就能得到(10)的左侧系数矩阵：

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ -2 & -3 & -7 \end{bmatrix}.$$

作为第二步我们将(9)中的第一个等式加到第三个等式上，得到

$$\begin{aligned} 2x_1 + 4x_2 - 2x_3 &= 2 \\ x_2 + x_3 &= 4 \\ x_2 + 5x_3 &= 12 \end{aligned}. \tag{11}$$

其矩阵表示为

$$\begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 12 \end{bmatrix}. \tag{12}$$

这样变元 x_1 在第二、第三个方程中就被完全消去。而这步消元所对应的消去矩阵为

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

将其分别乘以等式(10)的两侧，就能得到等式(12)的两侧：

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 12 \end{bmatrix} \quad \text{和} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ -2 & -3 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 1 & 5 \end{bmatrix}.$$

在第三步，我们将(11)中的第三个等式减去第二个等式，那么就有

$$\begin{aligned} 2x_1 + 4x_2 - 2x_3 &= 2 \\ x_2 + x_3 &= 4 \\ 4x_3 &= 8 \end{aligned}. \tag{13}$$

写成矩阵形式：

$$\begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}. \tag{14}$$

这里变元 x_2 在第三个方程中被消去。我们得到了一个上三角方程组，也就很容易得到（原）方程组的解。这最后一步消元对应的消去矩阵是

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

将其分别乘以等式(12)的两侧，就能得到等式(14)的两侧：

$$\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}\begin{bmatrix}2\\4\\12\end{bmatrix}=\begin{bmatrix}2\\4\\8\end{bmatrix}\quad \text{和} \quad \begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}\begin{bmatrix}2&4&-2\\0&1&1\\0&1&5\end{bmatrix}=\begin{bmatrix}2&4&-2\\0&1&1\\0&0&4\end{bmatrix}.$$

我们可以将上述三步消元过程对原始方程组(7)的系数矩阵，即(8)等式的左侧，的作用写成如下的矩阵连乘：

$$\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}\left(\begin{bmatrix}1&0&0\\0&1&0\\1&0&1\end{bmatrix}\left(\begin{bmatrix}1&0&0\\-2&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}2&4&-2\\4&9&-3\\-2&-3&7\end{bmatrix}\right)\right) \tag{15}$$

根据我们马上要引入的矩阵乘法，可以得到(15)等于

$$\left(\left(\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\1&0&1\end{bmatrix}\right)\begin{bmatrix}1&0&0\\-2&1&0\\0&0&1\end{bmatrix}\right)\begin{bmatrix}2&4&-2\\4&9&-3\\-2&-3&7\end{bmatrix}$$

$$=\begin{bmatrix}1&0&0\\-2&1&0\\3&-1&1\end{bmatrix}\begin{bmatrix}2&4&-2\\4&9&-3\\-2&-3&7\end{bmatrix}=\begin{bmatrix}2&4&-2\\0&1&1\\0&0&4\end{bmatrix}.$$

事实上，矩阵

$$\begin{bmatrix}2&4&-2\\0&1&1\\0&0&4\end{bmatrix}$$

在整体上刻画了上述的高斯消元过程：

$$\begin{aligned}2x_1+4x_2-2x_3&=2\\4x_1+9x_2-3x_3&=8\\-2x_1-3x_2+7x_3&=10\end{aligned}\quad\Longrightarrow\quad\begin{aligned}2x_1+4x_2-2x_3&=2\\x_2+x_3&=4\\4x_3&=8\end{aligned}$$

即从矩阵的观点看，我们有如下的两个等式：

$$\begin{bmatrix}1&0&0\\-2&1&0\\3&-1&1\end{bmatrix}\begin{bmatrix}2&4&-2\\4&9&-3\\-2&-3&7\end{bmatrix}=\begin{bmatrix}2&4&-2\\0&1&1\\0&0&4\end{bmatrix}$$

和

$$\begin{bmatrix}1&0&0\\-2&1&0\\3&-1&1\end{bmatrix}\begin{bmatrix}2\\8\\10\end{bmatrix}=\begin{bmatrix}2\\4\\8\end{bmatrix}.$$

2 矩阵和矩阵上的运算

定义 2.1

令 $m, n \geqslant 1$。一个 m **行** n **列矩阵**（即 $m \times n$ **矩阵**）A 有如下的形式:

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

其中每一 $a_{ij} \in \mathbb{R}$，我们也称其为**第 i 行第 j 列项（或元素））**，记为

$$A(i, j).$$

一些常用的特殊矩阵

- **方阵**：行数和列数相等 $(m = n)$ 的矩阵。
- **全零矩阵**：所有元素都为 0 的矩阵，记为 $O_{m \times n}$，在不引起误会的前提下，也会直接用 O。
- **单位矩阵**：所有元素都为 0 的方阵，记为 $E_{n \times n}$，在不引起误会的前提下，也会直接用 E。
- **对角矩阵**：只有对角线 $A(i, i)$ 上非零的方阵，记为 $\mathrm{diag}(a_1, \ldots, a_n)$。

有了矩阵的定义后，我们最重要的目标之一是将线性方程组

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \ddots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{array}$$

转换为

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{m \times n\ \text{矩阵}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{n \times 1\ \text{矩阵}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{m \times 1\ \text{矩阵}}$$

的形式。但除了这里所需的矩阵乘法，我们也会看到矩阵加法及标量乘法也有重要的意义及应用。

矩阵加法. 对两个行数和列数相同的矩阵，我们可以定义它们的加法：

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{m1} & a'_{m2} & \cdots & a'_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+a'_{11} & a_{12}+a'_{12} & \cdots & a_{1n}+a'_{1n} \\ a_{21}+a'_{21} & a_{22}+a'_{22} & \cdots & a_{2n}+a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+a'_{m1} & a_{m1}+a'_{m2} & \cdots & a_{mn}+a'_{mn} \end{bmatrix}$$

直观上，它对应了将两个线性方程组相加：由

$$\begin{array}{rcl} a_{11}x_1+\ldots+a_{1n}x_n & = & b_1 \\ \ddots & & \vdots \\ a_{m1}x_1+\ldots+a_{mn}x_n & = & b_m \end{array} \qquad \begin{array}{rcl} a'_{11}x_1+\ldots+a'_{1n}x_n & = & b'_1 \\ \ddots & & \vdots \\ a'_{m1}x_1+\ldots+a'_{mn}x_n & = & b'_m \end{array}$$

我们得到

$$\begin{array}{rcl} (a_{11}+a'_{11})x_1+\ldots+(a_{1n}+a'_{1n})x_n & = & b_1+b'_1 \\ \ddots & & \vdots \\ (a_{m1}+a'_{m1})x_1+\ldots+(a_{mn}+a'_{mn})x_n & = & b_m+b'_m \end{array}.$$

标量乘法.

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

矩阵乘法. 下面我们定义最一般的矩阵乘法。

定义 2.2

假设 $m,n,p \geqslant 1$。给定 $m\times n$ 矩阵 A 和 $n\times p$ 矩阵 B，即

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad 且 \quad B=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}.$$

那么 AB 是形式为

$$AB=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix},$$

的 $m\times p$ 矩阵，其中每个 c_{ij} 是 A 的第 i 个行向量与 B 的第 j 个列向量的内积定义：

$$c_{ij}=(a_{i1},\ldots,a_{in})\cdot(b_{1j},\ldots,b_{nj})=a_{i1}b_{1j}+\ldots+a_{in}b_{nj}=\sum_{k\in[n]}a_{ik}b_{kj}.$$

当 $p=1$，即矩阵 B 为一列向量，上面的定义恰好是我们已经讨论过的(4)。

矩阵乘法可以由线性方程组的代入来解释。考察如下的两个线性方程组:

$$\begin{aligned} a_{11}x_1+\cdots+a_{1n}x_n &= y_1 \\ \ddots \quad &\ \ \vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n &= y_m \end{aligned} \quad \text{即} \quad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

和

$$\begin{aligned} b_{11}z_1+\cdots+b_{1p}z_p &= x_1 \\ \ddots \quad &\ \ \vdots \\ b_{n1}z_1+\cdots+b_{np}z_p &= x_n \end{aligned} \quad \text{即} \quad \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_p \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

将 $z_1,\ldots,z_p$ 代入 $x_1,\ldots,x_n$，那么对于任意 $i\in[m]$

$$\begin{aligned} y_i &= a_{i1}x_1+\ldots+a_{in}x_n = \sum_{k\in[n]} a_{ik}x_k \\ &= \sum_{k\in[n]} a_{ik}\left(\sum_{j\in[p]} b_{kj}z_j\right) = \sum_{k\in[n]}\sum_{j\in[p]} a_{ik}b_{kj}z_j = \sum_{j\in[p]}\left(\sum_{k\in[n]} a_{ik}b_{kj}\right) z_j \\ &= \sum_{j\in[p]} (a_{i1},\ldots,a_{in})\cdot(b_{1j},\ldots,b_{nj})z_j. \end{aligned}$$

我们可以将上面的过程写成矩阵形式:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \quad \text{和} \quad \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_p \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

那么

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \left(\begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_p \end{bmatrix} \right)$$

因为每一 $y_i = \sum_{j\in[p]}(a_{i1},\ldots,a_{in})\cdot(b_{1j},\ldots,b_{nj})z_j$，这完美符合我们的矩阵乘法定义:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \left(\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix} \right) \begin{bmatrix} z_1 \\ \vdots \\ z_p \end{bmatrix}.$$

除了线性方程组，我们还能通过别的一些角度来理解矩阵乘法。

- AB **的每一列都是** A **的列向量的线性组合**：将 B 写成

  $$\begin{bmatrix} \mathbf{b}_1 \cdots \mathbf{b}_p \end{bmatrix}.$$

  那么

  $$AB = \begin{bmatrix} A\mathbf{b}_1 \cdots A\mathbf{b}_p \end{bmatrix}.$$

- AB **的每一行都是** B **的行向量的线性组合：** 将 A 写成：

$$\begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}.$$

  那么

$$AB = \begin{bmatrix} \mathbf{a}_1 B \\ \vdots \\ \mathbf{a}_m B \end{bmatrix}.$$

- 将 A 写成由列向量构成的矩阵

$$\begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix},$$

  将 B 写成行向量构成的矩阵

$$\begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{bmatrix}.$$

  那么

$$AB = \mathbf{a}_1 \mathbf{b}_1 + \ldots + \mathbf{a}_n \mathbf{b}_n.$$