《线性代数》

2-解线性方程组 (I)(Solving Linear Equations(I))

杨启哲

上海师范大学信机学院计算机系

2024 年 3 月 6 日

复习：向量的线性组合

给定三个向量：

$$u = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad w = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

其线性组合可以表示为 $x_1u + x_2v + x_3w$，也就是：

$$x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix}$$

矩阵表示

我们可以利用矩阵来表示上述行为：

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix}$$

主要内容

上海师范大学 Shanghai Normal University

- 解线性方程组
- 解线性方程组的矩阵表示
- 矩阵的运算

教材参考章节

上海师范大学 Shanghai Normal University

- 第 2 章 2.1, 2.2, 第 3 章 3.1

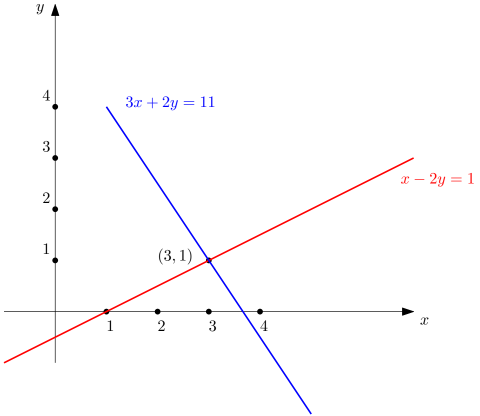
解线性方程组

线性方程组的行图像

考虑如下的线性方程组:

$$x - 2y = 1$$
$$3x + 2y = 11$$

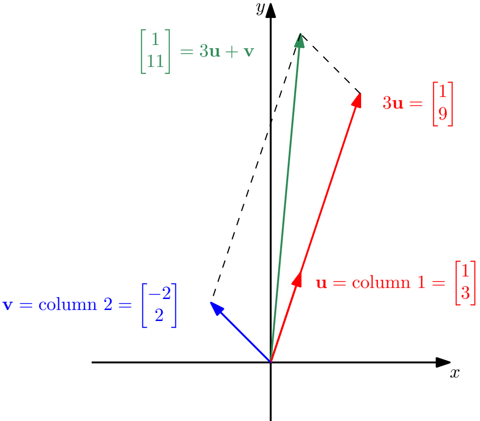
其行图像为:



该方程组也可表示为:

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

其列图像为:



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

上述矩阵 $\begin{bmatrix}1 & -2\\3 & 2\end{bmatrix}$ 被称作系数矩阵 (coefficient matrix)，我们也可以用点积的形式理解：

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

不止 2 维!

以上所有当然可以推广到 3 维，4 维，$\ldots$!

一个三元一次方程组的例子 (I)

考虑 3 个变元的例子：

$$\begin{cases} x + 2y + 3z = 6 \\ 2x + 5y + 2z = 4 \\ 6x - 3y + z = 2 \end{cases}$$

这个方程具有唯一的解 $(x, y, z) = (0, 0, 2)$。

几何视角

- 从行视角来看，这 3 个方程代表了 3 个平面，它们相交于一个点 $(0, 0, 2)$。
- 从列视角来看，这 3 个方程代表了 3 个列向量的线性组合，即：

$$0 \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 5 \\ -3 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$$

一个三元一次方程组的例子 (II)

同样的，我们有：

$$\begin{cases} x + 2y + 3z = 6 \\ 2x + 5y + 2z = 4 \\ 6x - 3y + z = 2 \end{cases} \iff Ax \overset{\text{记为}}{=} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 6 & -3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$$

其 $x = (x, y, z)$ 是未知数，对应的解是 $(0, 0, 2)$。从而 Ax 的目标便是能够有如下的成立：

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 6 & -3 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$$

对于:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

定义 Ax 的第 i 部分值为:

$$a_{i1}x_1 + \cdots + a_{in}x_n = \sum_{j=1}^{n} a_{ij}x_j = (a_{i1}, \ldots, a_{in}) \cdot (x_1, \ldots, x_n)$$

现在一个 n 元一次方程组已经可以自然的表示成 $Ax = b$ 了。我们来开始解这个方程组。

高斯消元法 (Gaussian Elimination)

来考虑下面的方程组:

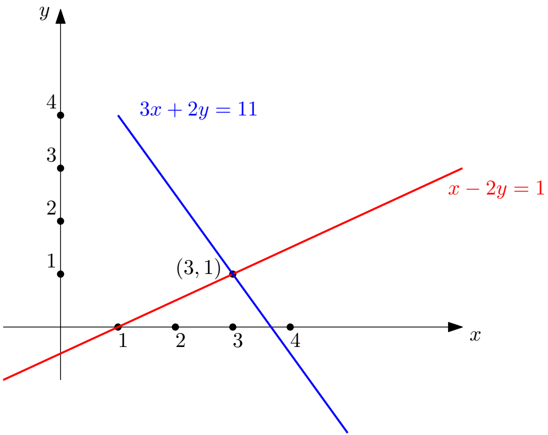
$$\begin{aligned} x - 2y &= 1 \\ 3x + 2y &= 11 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} x - 2y &= 1 (\text{将第一个式子乘以 } 3) \\ 8y &= 8 (\text{减去上面的式子消去 } x) \end{aligned}$$

新的方程是上三角的 (upper triangular)，类似的，有:

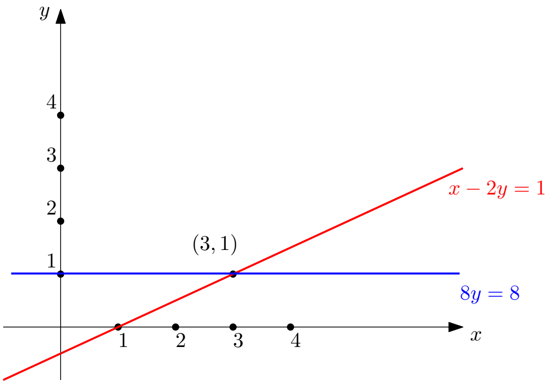
$$\begin{aligned} 4x - 8y &= 4 \\ 3x + 2y &= 11 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} {\color{red}4}x - 8y &= 4 (\text{将第一个式子乘以} \tfrac{3}{4}) \\ {\color{red}8}y &= 8 (\text{减去上面的式子消去 } x) \end{aligned}$$

- 首元 (pivot)：行中第一个做其他行消去的非零元素。在完成消去后，首元在对角线上。
- 倍数 (multiplier)：用来消去其他行的倍数。

几何视角



消去前



消去后

另外一个例子 (I)

考察方程组:

$$\begin{cases} 2x + 4y - 2z = 2 \\ 4x + 9y - 3z = 8 \\ -2x - 3y + 7z = 10 \end{cases}$$

1 第二个等式减去 2 倍的第一个式子，得到 $y + z = 4$。

2 第三个等式减去 -1 倍的第一个式子，得到 $y + 5z = -8$。

此时 x 已经被消去了，剩下:

$$\begin{cases} 1y + z = 4 \\ y + 5z = 12 \end{cases}$$

另外一个例子 (II)

3 将第二个等式乘以 -1，然后加上第三个等式，得到 $4z = 8$。

最终我们得到了:

$$
\begin{aligned}
2x + 4y - 2z &= 2 \\
4x + 9y - 3z &= 8 \\
-2x - 3y + 7z &= 10
\end{aligned}
\quad \Longrightarrow \quad
\begin{aligned}
2x + 4y - 2z &= 2 \\
1y + z &= 4 \\
4z &= 8
\end{aligned}
$$

回代 (Back Substitution)(I)

接下来我们只要从最底下的方程组依次往上便可求出对应的解:

- $$\begin{aligned} x - 2y &= 1 \\ 3x + 2y &= 11 \end{aligned} \Longrightarrow \begin{aligned} x - 2y &= 1 \\ 8y &= 8 \end{aligned}$$

  由 $8y = 8$ 自然而然有 $y = 1$，再代入 $x - 2y = 1$ 便有 $x = 3$。

- $$\begin{aligned} x - 2y &= 1 \\ 3x - 6y &= 3 \end{aligned} \Longrightarrow \begin{aligned} x - 2y &= 1 \\ 0y &= 0 \end{aligned}$$

  由 $0y = 0$ 从而有无数多个解。

- $$\begin{aligned} x - 2y &= 1 \\ 3x - 6y &= 11 \end{aligned} \implies \begin{aligned} x - 2y &= 1 \\ 0y &= 8 \end{aligned}$$

  $0y = 8$ 没有解。

首元的个数

可以看到，消元会失败在当 n 个方程没有 n 个首元的时候：

- ”$0 \neq 0$“的方程：没有解。
- ”$0 = 0$“的方程：有无数多个解。

阶段总结

- 线性方程组的几何视角，行图像和列图像。
- 线性方程组的矩阵表示，矩阵与向量的乘积。
- 高斯消元法 (Gaussian Elimination)。

解线性方程组的矩阵表示

矩阵形式

现在我们用矩阵的形式研究高斯消元法:

$$\begin{aligned} 2x + 4y - 2z &= 2 \\ 4x + 9y - 3z &= 8 \\ -2x - 3y + 7z &= 10 \end{aligned} \qquad \Longrightarrow \qquad \begin{aligned} {\color{red}2}x + 4y - 2z &= 2 \\ {\color{red}1}y + z &= 4 \\ {\color{red}4}z &= 8 \end{aligned}$$

矩阵形式:

$$\begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ 2 & -2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 3 \end{bmatrix} \qquad \Longrightarrow \qquad \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}$$

记左边形式为 $Ax = b$，右边的形式为 $Ux = c$。矩阵 A 怎么变到 U, b 怎么变到 c?

我们来考察其一个步骤：

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

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

第二行减去 2 倍的第一行，得到：

$$\begin{aligned} 2x_1 + 4x_2 - 2x_3 &= 2 \\ 1x_2 + 1x_3 &= 4 \\ -2x_1 - 3x_2 + 7x_3 &= 10 \end{aligned}$$

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

一步消去步骤的矩阵表示 (II)

我们同样希望能够用矩阵来表示这个过程，即：

$$\begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix} \qquad \Rightarrow \qquad \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ -2 & -3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix}$$

注意到：

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

一步消去步骤的矩阵表示 (III)

自然我们希望:

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

我们称上述的矩阵为初等矩阵 (elementary matrix)或者消元矩阵 (elimination matrix).

- 如何去定义?　　点积。

矩阵的乘法视角

$A\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ 的列视角:

$$A\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=x_1(\text{colomn 1 of }A)+x_2(\text{colomn 2 of }A)+x_3(\text{colomn 3 of }A)$$

$\begin{bmatrix}x_1 & x_2 & x_3\end{bmatrix}A$ 的行视角:

$$\begin{bmatrix}x_1 & x_2 & x_3\end{bmatrix}A=x_1(\text{row 1 of }A)+x_2(\text{row 2 of }A)+x_3(\text{row 3 of }A)$$

矩阵的乘法 AB 理解

- $AB=A\begin{bmatrix}b_1 & b_2 & b_3\end{bmatrix}=\begin{bmatrix}Ab_1 & Ab_2 & Ab_3\end{bmatrix}$
- $AB=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}B=\begin{bmatrix}a_1B\\a_2B\\a_3B\end{bmatrix}$

我们现在再回过头来看消元矩阵，我们从如下的单位矩阵 (identity matrix)开始:

$$
\mathrm{I}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}
$$

对于任意 b 有:

$$
\mathrm{Ib}=\mathrm{b}
$$

当我们将其中某一项 0 的位置替换成了一个非零的数 $-k$，则我们有：

$$E_{21}=\begin{bmatrix}1&0&0\\-k&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}=\begin{bmatrix}b_1\\b_2-kb_1\\b_3\end{bmatrix}$$

- 我们将这样子将第 i 行第 j 列的 0 替换成一个非零的数 $-k$ 的矩阵称为 E_{ij}.
- $E_{ij}b$ 的作用实际上是将 b 的第 j 行的 $-k$ 倍加到了第 i 行后得到的

定义 1 **[消元矩阵].**

消元矩阵 E_{ij} 是将单位矩阵 I 的第 i 行第 j 列的 0 替换成一个非零的数 $-k$ 得到的矩阵。

符号说明

同济的教材里使用了 $E(ij(-k))$ 的符号表示。

有的时候我们需要将两行进行交换，这样的矩阵称为置换矩阵 (permutation matrix)。

- 如何将 b 的第一行和第二行交换？

$$P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad P_{12} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_2 \\ b_1 \\ b_3 \end{bmatrix}$$

定义 2 **[置换矩阵].**

置换矩阵 P_{ij} 是将单位矩阵 I 的第 i 行和第 j 行进行交换得到的矩阵。

符号说明

同济的教材里使用了 $E(i,j)$ 的符号表示。

让我们再来看刚刚的例子:

1

$$\begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix} \qquad \Rightarrow \qquad \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ -2 & -3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix}$$

其对应的消元矩阵

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} : \qquad \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 2 \\ 0 & 1 & 1 \\ -2 & -3 & 7 \end{bmatrix}$$

高斯消元法的矩阵表示 (II)

上海师范大学 Shanghai Normal University

2

$$\begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ -2 & -3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 12 \end{bmatrix}$$

其对应的消元矩阵

$$E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} : \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ -2 & -3 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 1 & 5 \end{bmatrix}$$

高斯消元法的矩阵表示 (III)

上海师范大学 Shanghai Normal University

3

$$\begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 12 \end{bmatrix}$$

其对应的消元矩阵

$$E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}: \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

我们将上述消去的过程合起来，可以表示成：

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

矩阵乘法的运算规则

我们先不加证明的列举出矩阵乘法的运算性质：

1. 满足结合律 (Associate Law)：$(AB)C = A(BC)$。
2. 不满足交换律 (Commutative Law)：$AB \neq BA$。

高斯消元法的矩阵表示-总结 (II)

上海师范大学 Shanghai Normal University

通过矩阵乘法的运算规则可知:

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

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

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

回顾线性方程组的表示:

$$\begin{aligned} 2x + 4y - 2z &= 2 \\ 4x + 9y - 3z &= 8 \\ -2x - 3y + 7z &= 10 \end{aligned} \qquad \Longrightarrow \qquad \begin{aligned} 2x + 4y - 2z &= 2 \\ 1y + z &= 4 \\ 4z &= 8 \end{aligned}$$

相应的消元过程可以用矩阵表示为:

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

和

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

增广矩阵 (Augmented Matrix)

注意到消元对应的矩阵是同样作用在系数矩阵 A 和 b 上的，我们可以将其合并成一个矩阵看待：

- 增广矩阵 (Augmented Matrix)：$\begin{bmatrix} A & b \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 & 2 \\ 4 & 9 & -3 & 8 \\ -2 & -3 & 7 & 3 \end{bmatrix}$

则上述过程可以一并表示为：

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 & 2 \\ 4 & 9 & -3 & 8 \\ -2 & -3 & 7 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 & 2 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 4 & 8 \end{bmatrix}$$

阶段总结

上海师范大学 Shanghai Normal University

- 高斯消元法的矩阵表示
- 消元矩阵，置换矩阵。
- 矩阵的乘法视角。

矩阵的运算

定义 3 **[矩阵 (Matrix)].**

令 $m, n \geqslant 1$. 一个 $m \times n$ 的矩阵 A 具有如下的形式:

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

特别的，我们用 $A(i,j)$ 表示矩阵 A 的第 i 行第 j 列的元素。

一般情况下线性方程组的矩阵表示

Shanghai Normal University

线性方程组:

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

可以转换成如下的矩阵形式:

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

定义 4 **[矩阵加法 (Addition)].**

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

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

矩阵加法 (II)

将两个线性方程组加起来:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \qquad \begin{cases} a'_{11}x_1 + a'_{12}x_2 + \cdots + a'_{1n}x_n = b'_1 \\ a'_{21}x_1 + a'_{22}x_2 + \cdots + a'_{2n}x_n = b'_2 \\ \qquad\qquad\vdots \\ a'_{m1}x_1 + a'_{m2}x_2 + \cdots + a'_{mn}x_n = b'_m \end{cases}$$

得到:

$$\begin{cases} (a_{11} + a'_{11})x_1 + (a_{12} + a'_{12})x_2 + \cdots + (a_{1n} + a'_{1n}x_n = b_1 + b'_1 \\ (a_{21} + a'_{21})x_1 + (a_{22} + a'_{22})x_2 + \cdots + (a_{2n} + a'_{2n})x_n = b_2 + b'_2 \\ \qquad\qquad\vdots \\ (a_{m1} + a'_{m1})x_1 + (a_{m2} + a'_{m2})x_2 + \cdots + (a_{mn} + a'_{mn})x_n = b_m + b'_m \end{cases}$$

定义 5 [矩阵数乘 (scalar multiplication)].

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

定义 6 **[矩阵乘法 (Matrix Multiplication)].**

令 $m, n, p \geqslant 1$，A 是一个 $m \times n$ 的矩阵，B 是一个 $n \times p$ 的矩阵：

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

则矩阵的乘积 AB 是一个 $m \times p$ 的矩阵：

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

其中 c_{ij} 定义为：

$$c_{ij} = (a_{i1}, \ldots, a_{in}) \cdot (b_{1j}, \ldots, b_{nj}) = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

说明

当 $p=1$ 的时候便是我们已经定义过的矩阵与向量的乘法。

例 7.

$$\begin{bmatrix}1 & 2 & 3\\0 & 1 & 2\\3 & 2 & 4\end{bmatrix}\begin{bmatrix}2 & 0 & 2\\1 & 1 & 1\\0 & 1 & 1\end{bmatrix}=\begin{bmatrix}4 & 5 & 7\\1 & 3 & 3\\8 & 6 & 12\end{bmatrix}$$

矩阵乘法视角-替换 (I)

Shanghai Normal University

考察如下的两个线性方程组:

$$(\mathrm{I})\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=y_1\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=y_2\\ \qquad\qquad\vdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=y_m \end{cases}$$

矩阵形式

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n \end{bmatrix}=\begin{bmatrix} y_1\\ y_2\\ \vdots\\ y_m \end{bmatrix}$$

$$(\mathrm{II})\begin{cases} b_{11}z_1+b_{12}z_2+\cdots+b_{1p}z_p=x_1\\ b_{21}z_1+b_{22}y_2+\cdots+b_{2p}z_p=x_2\\ \qquad\qquad\vdots\\ b_{n1}y_1+b_{n2}z_2+\cdots+b_{np}z_p=x_n \end{cases}$$

矩阵形式

$$\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p}\\ b_{21} & b_{22} & \cdots & a_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}\begin{bmatrix} z_1\\ z_2\\ \vdots\\ z_p \end{bmatrix}=\begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n \end{bmatrix}$$

矩阵乘法视角-替换 (II)

如果我们希望用 $z_1, \ldots z_p$ 来表示 $y_1, \ldots, y_m$，则有:

$$
\begin{aligned}
y_i &= a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \\
&= a_{i1}(b_{11}z_1 + b_{12}z_2 + \cdots + b_{1p}z_p) + \cdots + a_{in}(b_{n1}z_1 + b_{n2}z_2 + \cdots + b_{np}z_p) \\
&= (a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1})z_1 + \cdots + (a_{i1}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np})z_p \\
&= \sum_{j=1}^{p} (a_{j1}, \ldots, a_{jn}) \cdot (b_{1j}, \ldots, b_{nj})z_j
\end{aligned}
$$

用矩阵的形式描述便是由:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \text{和} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_p \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \left(\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_p \end{bmatrix} \right)$$

注意到: $y_i = \sum_{j=1}^{p} (a_{i1}, \ldots, a_{in}) \cdot (b_{1j}, \ldots, b_{nj}) z_j$, 这就是:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \right) \begin{bmatrix} z_1 \\ \vdots \\ z_p \end{bmatrix}$$

矩阵乘法的列视角

当我们将 B 看成若干列向量组合而成的矩阵时，即：

$$B = \begin{bmatrix} b_1 & b_2 & \cdots & b_p \end{bmatrix}$$

我们有：

$$AB = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_p \end{bmatrix}$$

即：

矩阵 AB 的每一列都是 A 中列向量的线性组合。

当我们将 A 看成若干行向量组合而成的矩阵时，即:

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

我们有:

$$AB = \begin{bmatrix} a_1 B \\ a_2 B \\ \vdots \\ a_m B \end{bmatrix}$$

即:

矩阵 AB 的每一行都是 B 中行向量的线性组合。

矩阵乘法的第四种视角

将 A 写成列向量的形式:

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

将 B 写成行向量的形式:

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

则我们有:

$$AB = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

矩阵运算的性质 (I)

下面我们给出矩阵运算的一些性质，我们不妨假定里面的矩阵总是符合运算的要求的：

引理 8.

矩阵加法和数乘满足：

1. 交换律 (Commutative Law)：$A + B = B + A$
2. 分配律 (Distributive Law)：$c(A + B) = cA + cB$
3. 结合律 (Associative Law)：$(A + B) + C = A + (B + C)$

引理 9.

矩阵乘法满足:

1. 结合律 (不需要括号): $(AB)C = A(BC)$
2. 分配律 (左分配律): $(A + B)C = AC + BC$
3. 分配律 (右分配律): $A(B + C) = AB + AC$

注意

我们再次强调，矩阵乘法不满足交换律，即一般情况下 $AB \neq BA$。

阶段总结

- 矩阵以及矩阵运算的定义。
- 矩阵乘法的四种理解方式。
- 矩阵运算的性质。