矩阵与高斯消元

> 参考资料：
>
> 《LINEAR ALGEBRA AND ITS APPLICATIONS》 -- Gilbert Strang
>
> MIT 18.06 线性代数笔记：linear-algebra-notes
>
> MIT 18.06：麻省理工学院 - MIT - 线性代数（我愿称之为线性代数教程天花板）

线性方程组的几何表示

二维空间

给定以下一个简单的线性方程组，我们可以分别从行图像（row picture）或列图像（column picture）的角度进行求解：

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

在行图像的观点中，通过在二维平面上单独绘制式(1-1)的两个方程，其两条直线的交点即为解，如图1-1所示。

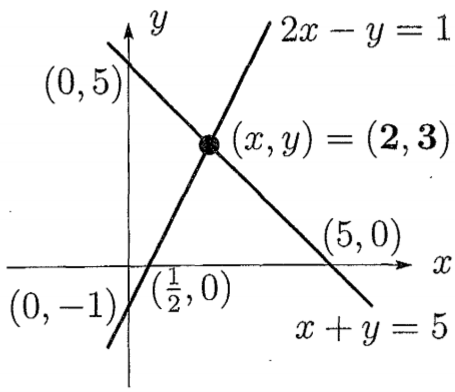

图1-1 行图像 row picture

在列图像的观点中，把式(1-1)的系数用向量表示，可得到如下的向量形式：

$$x\begin{bmatrix}2\\1\end{bmatrix} + y\begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}1\\5\end{bmatrix}. \tag{1-2}$$

此时将求解问题转化成寻找一种线性组合，使得方程左边的向量产生右边的向量。显然，其线性组合方式为2倍的列1加上3倍的列2，如图1-2所示。

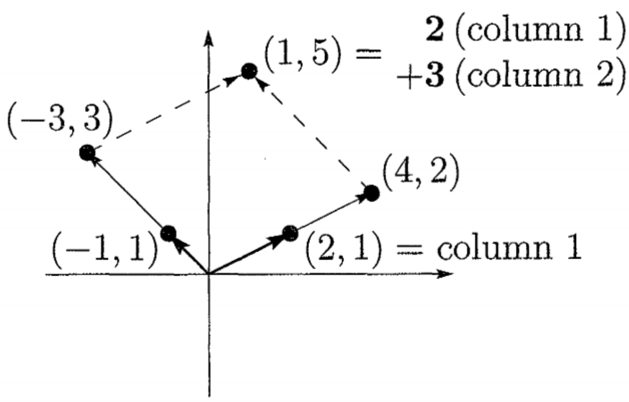

图1-2 列图像 column picture

三维空间

现考虑三维的情况，其更具有一般性：

$$\begin{aligned} 2u+v+w&=5\\ 4u-6v&=-2\\ -2u+7v+2w&=9. \end{aligned} \tag{1-3}$$

式(1-3)的每个方程都描述了三维空间中的一个平面，第一个平面$2u+v+w=5$，如图1-3所示，其包含了点$\left(\frac{5}{2},0,0\right)$、$(0,5,0)$和$(0,0,5)$，它由其任意三个点决定，前提是三点不共线。

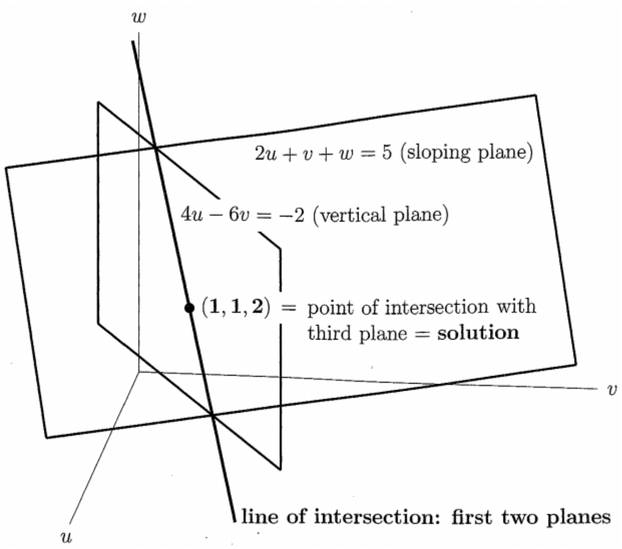

图1-3 行图像：三个方程的三个相交平面

若将第一个方程的右边常数修改为10，即$2u+v+w=10$，那么该方程绘制的平面将和原平面平行，且平面上的点$(5,0,0)$、$(0,10,0)$、$(0,0,10)$ 到原点的距离是原平面上对应点的两倍。这说明改变方程右边的常数项会使平面平行移动。

因为w的系数为0，则第二个方程$4u-6v=-2$表示的第二个平面是垂直的。第一个平面和第二个平面相交，其交线是一条直线，在三维空间中确定一条直线需要两个方程，在n维空间中确定一条直线需要 $n-1$个方程。

第三个平面$-2u+7v+2w=9$与前面两个平面相交所得的直线相交于一点$(1,1,2)$，这个点就是线性方程组的解。

在n维空间中，第一个方程产生一个$(n-1)$维平面，在理想情况下第二个平面与它相交得到一个更小的$n-2$维集合。在理想情况下，每引入一个新平面，即一个新方程，维度就降低一维。当考虑完所有n个平面后，它们的交集维度为零，也就是一个点，这个点在所有平面上，其坐标满足所有n个方程，即为方程组的解。

现在从列图像的角度解决这个三维问题，方程组(1-3)可以用列向量的形式表示：

$$u\begin{bmatrix}2\\4\\-2\end{bmatrix}+v\begin{bmatrix}1\\-6\\7\end{bmatrix}+w\begin{bmatrix}1\\0\\2\end{bmatrix}=\begin{bmatrix}5\\-2\\9\end{bmatrix}=b. \tag{1-4}$$

我们很容易可以找到一种线性组合，即1倍的列1加上1倍的列2再加上2倍列3，即可得到向量b，如图1-4所示。

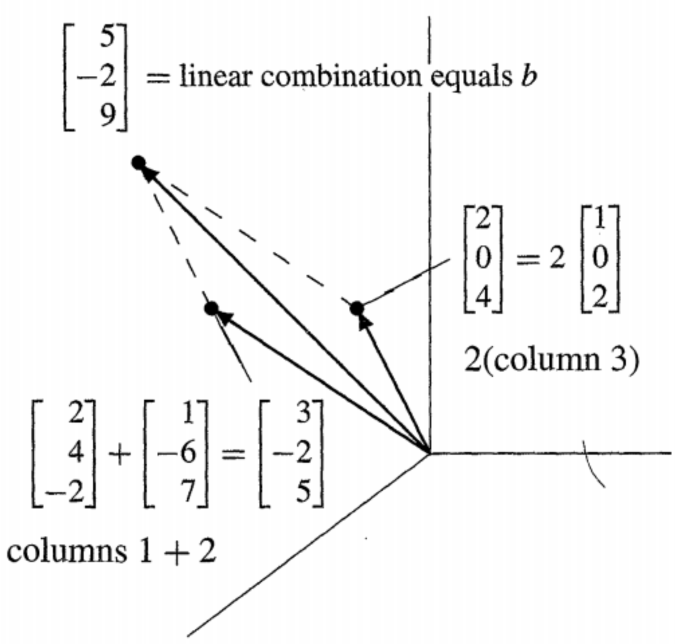

图1-4 列图像：三个向量的线性组合

对于n个未知数的n个方程，在行图像中有n个平面，在列图像中有n个向量以及右边的一个向量b。方程组的本质是要找到n个列向量的线性组合等于向量b。但对于某些方程组，不可能找到这样的线性组合。综上所述，行图像是平面的相交，列图像是列向量的组合。

在这门课程中，理解"有解"这种好情况的方法是研究"无解"这种坏情况，所以要研究在奇异（singular）情况下，即几何关系失效时的情况。

奇异系统

在这一小节中，将分别从行图像和列图像的角度讨论在三维空间无解的情况。

行图像的角度

第一种情况是平行平面导致无解。在三维空间中，若行图像里三个平面不相交，比如存在两个平面平行，将导致无解。例如方程$2u+v+w=5$和$4u+2v+2w=11$表示的两个平面平行没有交点，因此对应方程组无解，如图1-5(a)所示。值得一提的是，之所以这两个方程代表的平面是平行的，是因为它们的法向量平行。

第二种情况是特殊相交导致无解。如图1-5b所示，从端视角看起来三个平面形成一个三角形，每对平面相交于一条直线，且这些直线平行。第三个平面虽不与其他平面平行，但与它们的交线平行。比如奇异系统：$u+v+w=2$，$2u+3w=5$，$3u+v+4w=6$，前两个方程左边部分相加等于第三个方程的左边部分，但它们的右边部分却不成立（$2+5\neq 6$），即方程1加方程2减方程3得到不可能的情况$0=1$，所以无解。值得一提的是，平面1的法向量为$\vec{v}_1=(1,1,1)$，平面2的法向量为$\vec{v}_2=(2,0,3)$，向量叉乘结果为$\vec{s}=(3,-1,-2)$，平面3的法向量为$\vec{v}_3=(3,1,4)$，$\vec{v}_3$与$\vec{s}$的向量点积结果为0，说明平面3与前两个平面的交线平行。

第三种情况是平面共线导致无穷多解。当上述奇异系统中最后一个方程的6变为7时，三个方程组合得到$0=0$，此时第三个方程是前两个方程的和，三个平面有一条公共直线，方程组有无穷多解，如图1-5c所示。值得一提的是，前两个平面的交线上一点为$(1,0,1)$，结合$\vec{s}=(3,-1,-2)$，可以写出交线的参数方程：

$$\begin{cases}u=1+3t\\ v=-t\\ w=1-2t.\end{cases} \tag{1-5}$$

代入第三个平面方程$3u+v+4w=7$的左边：

$$
\begin{aligned}
&3(1+3t)+(-t)+4(1-2t)\\
=&3+9t-t+4-8t\\
=&7.
\end{aligned}
\tag{1-6}
$$

结果等于方程右边，这说明前两个平面交线上的点都在第三个平面上。由于第三个平面与前两个平面存在线性关系，且前两个平面的交线上的点都在第三个平面上，所以这三个平面的交线是同一条。

第四种情况是极端情况，比如三个平面平行，如图1-5d所示。对于大多数右边向量b，方程组无解；但对于特殊的b如$b=(0,0,0)$，有一整个平面的解，因为可以把这三个平面看作是同一个平面经过平移后得到的不同位置的平面，当把它们都平移到重合时，就变成了同一个平面，那么这个平面上的所有点都满足这三个方程，所以就有一整个平面的解。

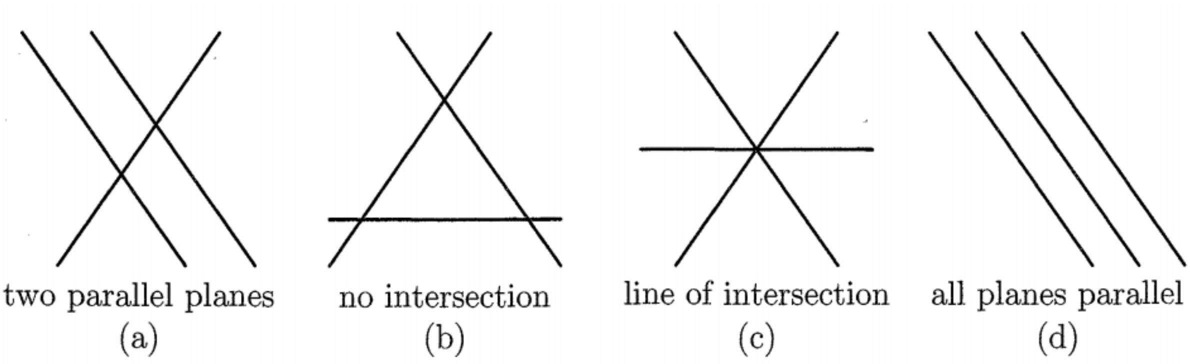

图1-5 行图像理解奇异情况：(a)(b)(d)没有解，(c)有无限解

列图像的角度

奇异系统的列图像也会出现问题，即列向量共面导致无解或无穷多解。现在考虑如下的奇异系统：

$$
u\begin{bmatrix}1\\2\\3\end{bmatrix}+v\begin{bmatrix}1\\0\\1\end{bmatrix}+w\begin{bmatrix}1\\3\\4\end{bmatrix}=b. \tag{1-7}
$$

左边有三个列向量，要尝试组合它们得到b。当三个列向量在同一平面时，只有当b在这个平面上方程组才有解，例如对于$b=(2,5,7)$有解，但对于$b=(2,5,6)$无解。因为这三个列向量在一个平面内，其线性组合也都在这个过原点的平面内，若b不在此平面则无解，如图1-6a所示。若b在这个平面内，则有无数种方式组合三个列向量得到b，即有无穷多解，如图1-6b所示，对应行图像的图1-5c。

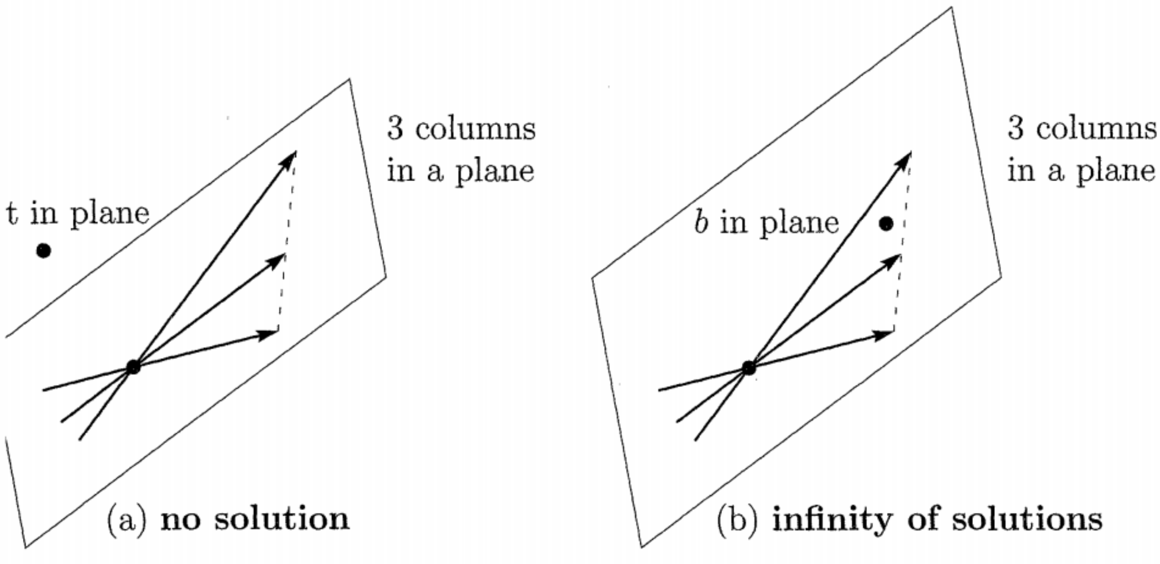

图1-6 列图像理解奇异情况：平面内有三列或b在平面外

判断三个列向量是否共面的一种方法是找到它们的线性组合为零向量。对于上述举例的奇异系统，当$u=3,v=-1,w=-2$时，三个列向量相加为零向量，即三倍列1等于列2加上2倍列3。这说明，第一列在第二列和第三列确定的平面上，只有两列是独立的。向量$b=(2,5,7)$在列向量所在平面内（列1加列3），故$(1,0,1)$是一个特解，还可加上使$b=0$的组合，即$(3,-1,-2)$的任意倍数，则可以得到一整条直线的通解。

行图像的平面相交情况和列图像的列向量的共面情况之间的紧密联系，在n维空间中，如果由n个方程所代表的n个平面没有公共点，即方程组无解，或者有无数个公共点，即方程组有无穷多解，那么对应的n个列向量会位于同一个平面。在维空间中更准确地说是位于一个$n-1$维子空间，但这里用 “平面” 表述便于理解。

高斯消元的一个例子

现考虑如下的线性方程组：

$$\begin{aligned} 2u+v+w&=5 \\ 4u-6v&=-2 \\ -2u+7v+2w&=9. \end{aligned} \tag{1-8}$$

对于高斯消元来说，首先对方程2减去2倍的方程1，接着对方程3减去-1倍的方程1，可得到如下消元形式：

$$\begin{aligned} 2u+v+w&=5 \\ -8v-2w&=-12 \\ 8v+3w&=14. \end{aligned} \tag{1-9}$$

在式(1-9)中，系数2是第一个主元。消元过程是不断地用主元除去其下方的数字，以找出合适的乘数。第二阶段消元的主元是-8，此时我们会忽略第一个方程，将从剩余方程中减去第二个方程的某个倍数，以便消去v。

从第三个方程中减去第二个方程的-1倍，可得到如下消元形式：

$$\begin{aligned} 2u+v+w&=5 \\ -8v-2w&=-12 \\ 1w&=2. \end{aligned} \tag{1-10}$$

式(1-10)即为高斯消元的结果，只需要从第三个方程往回代，即可求解出u、v和w。

记录正向消元步骤的一个好方法是将等式右边作为额外的一列，最后求得三角方程组，可用于回代（back - substitution）求解：

$$\begin{bmatrix} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix} \to \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{bmatrix} \to \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{bmatrix}. \tag{1-11}$$

在奇异状态下，消元过程肯定会崩溃，然而，在非奇异状态下也有可能出现问题。这是因为如果在主元位置出现了零，消元就必须停止。比如左上角的第一个系数为零，就不可能从其他方程中消去u。在每一个中间阶段，情况都是如此。需要注意的是，即使原始系数不为零，但零可能会在消元过程中出现在主元位置。

在非奇异的情况下，可以通过交换方程的位置进行修正；在奇异的情况下，方程组无解或无穷多解。

如下的例子即为非奇异情况，可以通过交换方程2和方程3进行修正：

$$\begin{aligned} u+v+w &= \underline{\qquad} \\ 2u+2v+5w &= \underline{\qquad} \\ 4u+6v+8w &= \underline{\qquad} \end{aligned} \rightarrow \begin{aligned} u+v+w &= \underline{\qquad} \\ 3w &= \underline{\qquad} \\ 2v+4w &= \underline{\qquad} \end{aligned} \rightarrow \begin{aligned} u+v+w &= \underline{\qquad} \\ 2v+4w &= \underline{\qquad} \\ 3w &= \underline{\qquad} \end{aligned}. \tag{1-12}$$

如下的例子即为奇异情况，我们只能解出$w=2$，而不能确定u和v。

$$\begin{aligned} u+v+w &= \underline{\qquad} \\ 2u+2v+5w &= \underline{\qquad} \\ 4u+4v+8w &= \underline{\qquad} \end{aligned} \rightarrow \begin{aligned} u+v+w &= \underline{\qquad} \\ -3w &= \underline{\qquad} \\ 4w &= \underline{\qquad} \end{aligned}. \tag{1-13}$$

矩阵表示和矩阵乘法

对于式(1-8)，我们可以把九个系数写成3×3的矩阵，并用A表示，三个未知变量写成3×1的向量，用x表示，三个等式右边的值写成3×1的向量，用向量b表示：

$$A=\begin{bmatrix}2&1&1\\4&-6&0\\-2&7&2\end{bmatrix}, x=\begin{bmatrix}u\\v\\w\end{bmatrix}, b=\begin{bmatrix}5\\-2\\9\end{bmatrix}. \tag{1-14}$$

A是一个方阵，因为方程的数量等于未知数的数量。如果有n个方程和n个未知数，我们就有$n\times n$矩阵。更一般地，我们可能有m个方程和n个未知数。那么A是一个长方形矩阵，有m行和n列，它被称为$m\times n$矩阵。 矩阵可以像向量一样逐个元素相加，或者乘以数值常数。实际上，我们把向量看作是矩阵的特殊情况，因为它们只有一列。和向量一样，只有当两个矩阵具有相同的形状时才能相加，比如：

$$A+B=\begin{bmatrix}2&1\\3&0\\0&4\end{bmatrix}+\begin{bmatrix}1&2\\-3&1\\1&2\end{bmatrix}=\begin{bmatrix}3&3\\0&1\\1&6\end{bmatrix}. \tag{1-15}$$

$$2A=2\begin{bmatrix}2&1\\3&0\\0&4\end{bmatrix}=\begin{bmatrix}4&2\\6&0\\0&8\end{bmatrix}. \tag{1-16}$$

如果希望把式(1-8)用式(1-14)的矩阵形式来表示，写成矩阵形式$Ax=b$就是：

$$\begin{bmatrix}2&1&1\\4&-6&0\\-2&7&2\end{bmatrix}\begin{bmatrix}u\\v\\w\end{bmatrix}=\begin{bmatrix}5\\-2\\9\end{bmatrix}. \tag{1-17}$$

显然，对于式(1-17)涉及到列和行的计算。

假设现在有一个矩阵A：

$$A=\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}. \tag{1-18}$$

那么对其进行列的线性组合是右乘一个列向量：

$$\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}\begin{bmatrix}3\\4\\5\end{bmatrix}=3\times\begin{bmatrix}a\\d\\g\end{bmatrix}+4\times\begin{bmatrix}b\\e\\h\end{bmatrix}+5\times\begin{bmatrix}c\\f\\i\end{bmatrix}. \tag{1-19}$$

对其进行行的线性组合是左乘一个行向量：

$$\begin{bmatrix}1&2&3\end{bmatrix}\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}=1\times\begin{bmatrix}a&b&c\end{bmatrix}+2\times\begin{bmatrix}d&e&f\end{bmatrix}+3\times\begin{bmatrix}g&h&i\end{bmatrix}. \tag{1-20}$$

一个消元步骤的矩阵形式

到目前为止，我们有一个方便的简写$Ax = b$来表示原始的方程组，即式(1-17)。在我们的例子中，第一步是从第二个方程中减去2倍的第一个方程，同样的结果可以通过将向量b乘以这个初等矩阵（或消元矩阵）来实现：

$$Eb = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} = \begin{bmatrix} 5 \\ -12 \\ 9 \end{bmatrix}. \tag{1-21}$$

矩阵E的作用是保持第一个和第三个方程不变，在第二个方程减去2倍的第一个方程。向量b的第一个和第三个分量保持不变，因为E在第一行和第三行是单位矩阵。观察其右边的向量，第二个分量-12是在第一次消元步骤后得到的。

更一般的，对于初等矩阵E_{ij}，表示从第i行减去ℓ倍的第j行，这个E_{ij}在第i行第j列包含$-\ell$。假如有如下矩阵：

$$E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\ell & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{1-22}$$

初等矩阵E_{31}对向量b的操作如下所示：

$$E_{31}b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 - \ell b_1 \end{bmatrix}. \tag{1-23}$$

矩阵乘法的五种方式

第一种方式是从元素的角度看，矩阵AB相乘结果矩阵C的元素可用如下方法计算：

$$c_{ij} = (\text{Row}_i \text{ of } A) \cdot (\text{Col}_j \text{ of } B) = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}. \tag{1-24}$$

第二种方式是从列的角度看，矩阵AB相乘结果矩阵C的各列是A中各列的一个线性组合，组合方式由B决定。比如下面两个矩阵相乘：

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & 8 \\ 9 & 10 \end{bmatrix}, \tag{1-25}$$

C的第一列为：

$$C_{col1} = 7 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + 9 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 25 \\ 57 \\ 89 \end{bmatrix}, \tag{1-26}$$

C的第二列为：

$$C_{col2} = 8 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + 10 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 28 \\ 24 \\ 100 \end{bmatrix}. \tag{1-26}$$

因此，矩阵AB相乘的结果为：

$$C=\begin{bmatrix}25 & 28\\57 & 64\\89 & 100\end{bmatrix}. \tag{1-27}$$

第三种角度上从行的角度看，矩阵AB相乘结果矩阵C的各行是B中各行的一个线性组合，组合方式由A决定，C的各行分别为：

$$\begin{array}{l}C_{row1}=1\begin{bmatrix}7 & 8\end{bmatrix}+2\begin{bmatrix}9 & 10\end{bmatrix}=\begin{bmatrix}25 & 28\end{bmatrix}\\C_{row2}=3\begin{bmatrix}7 & 8\end{bmatrix}+4\begin{bmatrix}9 & 10\end{bmatrix}=\begin{bmatrix}57 & 64\end{bmatrix}\\C_{row3}=5\begin{bmatrix}7 & 8\end{bmatrix}+6\begin{bmatrix}9 & 10\end{bmatrix}=\begin{bmatrix}89 & 100\end{bmatrix}\end{array}. \tag{1-28}$$

可见，从行的角度和列的角度，计算结果都是一致的。

第四种是从矩阵的角度看，矩阵AB相乘的结果可表示为：

$$C=\sum_{i=1}^{n}(\mathrm{Col}_i\text{ of }A)\cdot(\mathrm{Row}_i\text{ of }B), \tag{1-29}$$

以式(1-25)为例，取A的第一列与B的第一行进行外积：

$$\begin{bmatrix}1\\3\\5\end{bmatrix}\begin{bmatrix}7 & 8\end{bmatrix}=\begin{bmatrix}7 & 8\\21 & 24\\35 & 40\end{bmatrix}. \tag{1-30}$$

取A的第二列与B的第二行进行外积：

$$\begin{bmatrix}2\\4\\6\end{bmatrix}\begin{bmatrix}9 & 10\end{bmatrix}=\begin{bmatrix}18 & 20\\36 & 40\\54 & 60\end{bmatrix}. \tag{1-31}$$

将这两个结果矩阵相加得到最终的矩阵C：

$$C=\begin{bmatrix}7+18 & 8+20\\21+36 & 24+40\\35+54 & 40+60\end{bmatrix}=\begin{bmatrix}25 & 28\\57 & 64\\89 & 100\end{bmatrix}. \tag{1-32}$$

第五种方式是分块的角度，其中$C_1=A_1B_1+A_2B_3$：

$$\begin{bmatrix}A_1 & A_2\\A_3 & A_4\end{bmatrix}\begin{bmatrix}B_1 & B_2\\B_3 & B_4\end{bmatrix}=\begin{bmatrix}C_1 & C_2\\C_3 & C_4\end{bmatrix}. \tag{1-33}$$

三角因子和行变换

在这一小节，将重新讨论矩阵乘法的例子$Ax=b$：

$$Ax=\begin{bmatrix}2 & 1 & 1\\4 & -6 & 0\\-2 & 7 & 2\end{bmatrix}\begin{bmatrix}u\\v\\w\end{bmatrix}=\begin{bmatrix}5\\-2\\9\end{bmatrix}=b. \tag{1-34}$$

在"高斯消元的一个例子"章节中，进行两次消元操作：

- 步骤1：从第二个方程中减去第一个方程的2倍。
- 步骤2：从第三个方程中减去第一个方程的-1倍。
- 步骤3：从第三个方程中减去第二个方程的-1倍。

这样做的结果是得到了一个上三角矩阵$Ux=c$，其中新的系数矩阵U为：

$$Ux = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 2 \end{bmatrix} = c. \tag{1-35}$$

这个矩阵U是上三角矩阵，即所有主对角线下方的元素都为零。右边的向量项c是通过对原始向量b进行相同的操作得到的。

这里我们重点关注矩阵A和U之间的联系。用于步骤1、2、3的矩阵E被称为初等矩阵。

初等矩阵的工作方式是，若要从其他方程中减去方程j的倍数来得到方程i，只需将数字$-\ell$放入(i,j)位置，否则保持单位矩阵的形式，同时，对角线上为1，其他位置为0。最后，通过矩阵乘法执行了行操作。

所有三个步骤的结果是$GFEA = U$。需要注意的是，矩阵乘法有左乘和右乘的区别，E是第一个与A相乘的矩阵，然后是F，最后是G。我们可以将GFE相乘得到一个单一的矩阵，即看成是一个整体，该矩阵将A变换为U，同时也将b变换为c。

把GFE看成是一个整体，那这个矩阵是下三角矩阵，对角线上为1，下方的元素为0：

$$GFE = \begin{bmatrix} 1 & & \\ & 1 & \\ & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ 1 & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ -2 & 1 & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -2 & 1 & \\ -1 & 1 & 1 \end{bmatrix}. \tag{1-36}$$

现在的问题是，如何从U回到A？如何撤销高斯消元的步骤？ 要撤销步骤1并不难，加上两倍的第一行到第二行就行。实际上，减去然后加上同一行的操作相互抵消。如果初等矩阵E在(i,j)位置有数字$-\ell$，那么它的逆矩阵E^{-1}在该位置有$+\ell$。因此$E^{-1}E = I$，如下式所示：

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1-37}$$

我们可以对消元的每一步进行逆操作，使用E^{-1}、F^{-1}和G^{-1}。现在我们考虑一次性撤销整个过程，看看哪个矩阵能将U变回A。 由于步骤3是在从A到U的过程中最后执行的，所以在反向过程中，矩阵G必须首先被反转。第二个反向步骤是F^{-1}，最后是E^{-1}：

$$E^{-1}F^{-1}G^{-1}U = A, \tag{1-38}$$

或者，我们可以简洁地写成如下形式：

$$LU = A,$$

$$E^{-1}F^{-1}G^{-1} = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ -1 & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ -1 & -1 & 1 \end{bmatrix} = L. \tag{1-39}$$

这个矩阵的元素是在消元过程中作为乘数出现的。当矩阵相乘时，这些元素恰好能正确地组合在一起。如果消元过程没有行交换，那么任何方阵A都可以分解为$A = LU$，其中L是下三角矩阵，主对角线上为1，U是上三角矩阵。L的元素是在消元过程中用于减去主元行倍数的乘数，它们位于主对角线下方。U的对角线上的元素是主元。

行交换和置换矩阵

现在我们考虑在式(1-12)中遇到的通过交换两行解决主元为0的情况，一个简单的例子是：

$$\begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{1-40}$$

这个问题很明显，没有任何倍数的第一个方程能够消除系数3。 解决方法同样简单。交换这两个方程，将系数3移动到主元位置。在这个例子中，矩阵将变为上三角矩阵：

$$\begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix}, \tag{1-41}$$

为了用矩阵形式表示这个操作，我们需要置换矩阵P，它是通过交换单位矩阵I的行得到的：

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad PA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix}. \tag{1-42}$$

P对向量b也有相同的作用，即交换b_1和b_2。新的系统是$PAx = Pb$。

在这个过程中，未知数u和v不会被交换。 置换矩阵P具有与单位矩阵相同的行，但顺序不同，每一行和每一列都有且只有一个1。最常见的置换矩阵是$P = I$，它不进行任何交换。两个置换矩阵的乘积仍然是一个置换矩阵，它对单位矩阵的行进行了重新排序。 在知道了逆矩阵和转置的概念后，我们发现一个重要的事实：P^{-1}总是等于P^T。

因此，我们可以直接将矩阵A乘以置换矩阵P，这样矩阵A的行就处于正确的顺序，任何非奇异矩阵都可以准备进行消元。

逆和转置

一个$n \times n$矩阵的逆是另一个$n \times n$矩阵，矩阵A的逆写作A^{-1}，其基本性质很简单：如果先乘以A，然后再乘以A^{-1}，就会回到起始状态。若$b = Ax$，那么 $A^{-1}b = x$。 因此$A^{-1}Ax = x$。矩阵A^{-1}乘以A得到单位矩阵。但并非所有矩阵都有逆。下面列举几个关于逆矩阵的性质：

- 当且仅当消元过程产生n个主元时，逆矩阵存在。
- 矩阵A不可能有两个不同的逆。
- 如果A是可逆的，那么$Ax = b$的唯一解是$x = A^{-1}b$。
- 假设存在一个非零向量x使得$Ax = 0$，不可能有逆。
- 一个对角矩阵只要对角元素都不为零就有逆。

下面证明可逆矩阵的乘积AB的逆是$B^{-1}A^{-1}$，我们将它们相乘，并使用结合律去掉括号：

$$(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I. \tag{1-43}$$

对于逆矩阵的计算，可以使用高斯-若尔当方法，具体而言，将A和I矩阵放到一起形成一个增广矩阵，E矩阵即总的消元矩阵，其能够将A行变换为I，即$EA = I$：

$$E\,[A \quad I] = [I \quad A^{-1}]. \tag{1-44}$$

由$EA = I$可知E是A的逆矩阵，即$E = A^{-1}$，又因为$EI = E$，所以变换后的增广矩阵右侧自然就得到了逆矩阵。 同时，高斯-若尔当思想也告诉我们，如果矩阵A可以行变换化简成I，那么A是一个可逆矩阵。

对于矩阵A，若另一个矩阵的列直接取自A的行，那么我们称这个矩阵为A的转置矩阵，记为A^T。转置矩阵也具有一些性质：

- $(AB)^T = B^T A^T$
- $(A^{-1})^T = (A^T)^{-1}$

如果矩阵A转置以后没有发生任何变化，那么这个矩阵被称为对称矩阵。对于如何构建一个对称矩阵，只需要将这个矩阵的转置与这个矩阵相乘即可，即$(A^TA)^T = A^T(A^T)^T = A^TA$。

假设$A = A^T$且可以在不进行行交换的情况下分解为$A = LDU$，其中D是对角矩阵，U则是L的转置，对称分解就变为$A = LDL^T$，其例子如下所示：

$$\begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = LDL^T. \tag{1-45}$$