0 说明

笔记标题：MIT_LA_Lecture1-3

笔记版本：v1.2.1

对于文档的说明：

1. 你可以在我的Github[仓库]中下载本笔记的Markdwon源文档，并通过浏览目录进行更方便高效地浏览；也欢迎在[知乎文章]中进行浏览。
2. 本笔记参考的课程为MIT Linear Algebra（麻省理工线性代数），本课程在[网易公开课]、[Bilibili]和[youtube]等网站上都有视频资源，读者可以选择合适的平台观看。

3. 本笔记并未完全按照视频课的内容记录，添加了许多自己的理解、资料的补充和顺序的调整。
4. 本系列笔记在不断更新，已经发布的笔记也会偶尔进行内容更新，版本号可以在文件标题或说明的开头查看，你可以通过Github的commit信息来查看笔记更新内容。
5. 如果你对笔记内容有好的建议，请提出来，笔者在这里表示感谢。

对于内容的说明：

1. 小写字母表示的向量，比如$\vec{a}$，除非在特殊说明的情况下，都表示的是列向量。用$\vec{a}^T$来表示行向量。
2. 部分矩阵中`.`用来表示元素省略，并不表示元素为0。
3. 单位矩阵用I表示。

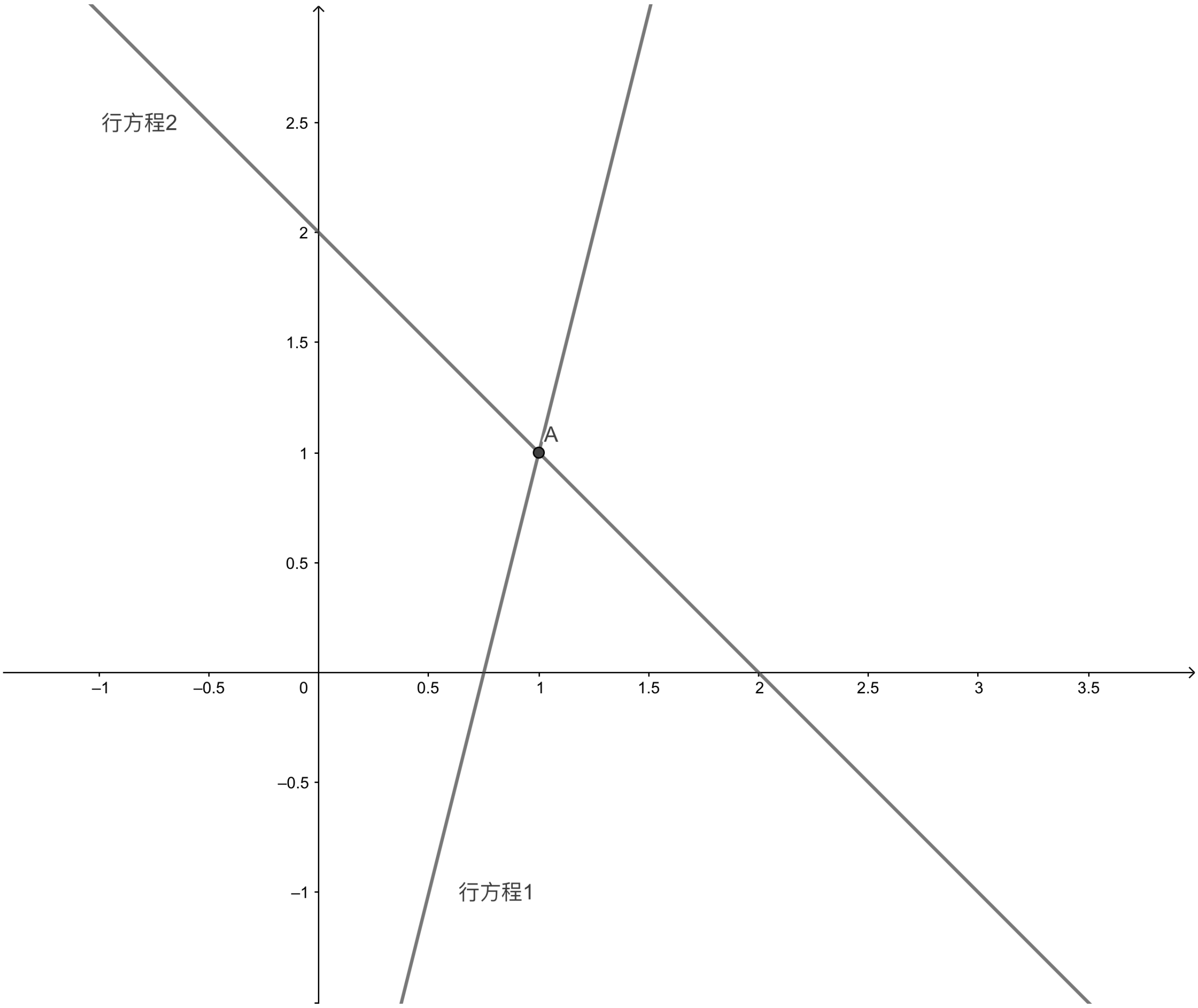
1 方程组的几何解释

二元线性方程组

我们考虑以下二元方程组以及其矩阵形式$Ax=b$：

$$\begin{aligned} 4x - y &= 3 \\ 2x + 2y &= 4 \end{aligned} \tag{1}$$

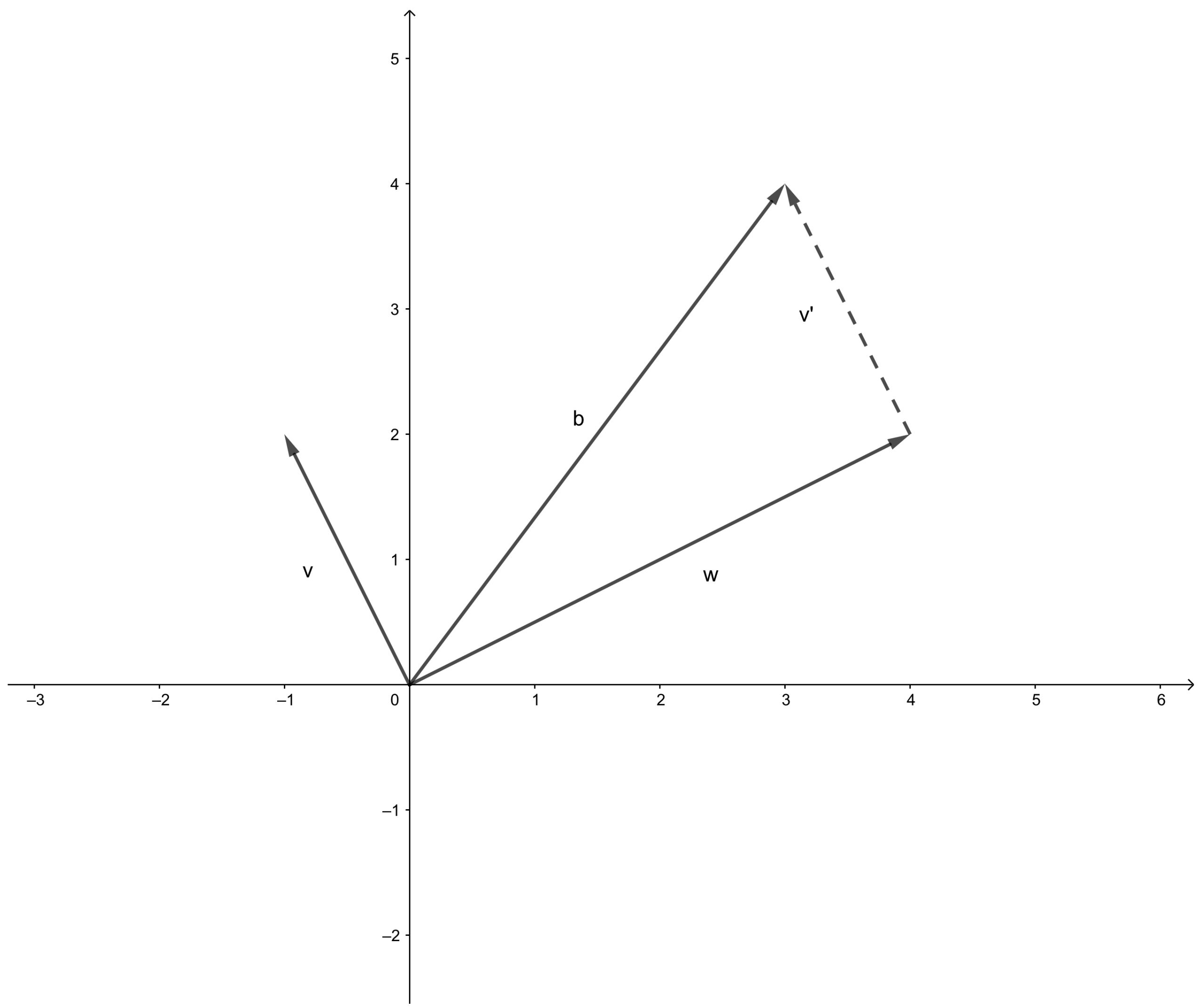
一般我们从**每一行**来认识这个方程，由于二元方程可以表示在XOY平面，我们可以画出(1)中两方程所表示的直线，我们也能很容易地求出满足以上方程的组合(x,y)=(1,1)，即图中的A点。



同时，我们也可以从每一列来理解这一方程组，即两向量的何种线性组合可以得到等号右端的向量：

$$x\begin{bmatrix}4\\2\end{bmatrix}+y\begin{bmatrix}-1\\2\end{bmatrix}=\begin{bmatrix}3\\4\end{bmatrix} \tag{2}$$
$$x\vec{w}+y\vec{v}=\vec{b}$$

我们也可以在XOY平面内画出$\vec{w}$和$\vec{v}$两向量，同理可以找出x=1，y=1：



多元线性方程组

我们将视角拉回多元情形，比如一般意义上的：

$$\begin{aligned} & A_{m\times n}\vec{x}_{n\times 1}=\vec{b}_{n\times 1} \\ & A=(\vec{a}_1,\vec{a}_2,\quad \ldots \quad \vec{a}_n) \\ & \vec{x}=(x_1,x_2,\quad \ldots \quad x_n)^T \end{aligned} \tag{3}$$

可以以列向量的线性表示的角度来看，即：

$$x_1\vec{a}_1+x_2\vec{a}_2+\ldots+x_n\vec{a}_n=\vec{b} \tag{4}$$

矩阵列的线性组合

于是我们可以将方程组的几何解释理解为，在m维空间中，系数矩阵A的列向量，是否可以通过某组系数来线性表示为$\vec{b}$。线性方程组或矩阵方程看成**矩阵列的线性组合**是理解线性方程组、矩阵方程的很好的方式。

当然，是否存在这样的一组数据也就是等价于这个方程组是否有解，也可以从诸多角度来理解，这里不做过多解释。

2 矩阵消元

主元，行阶梯型，简化行阶梯型

考虑以下线性方程组：

$$\begin{aligned} x+2y+z&=2 \\ 3x+8y+z&=12 \\ 4y+z&=2 \end{aligned} \tag{5}$$

我们可以通过对（6）中方程加减，从而得到一个等价的更容易求解的方程组，而实际中，我们是对其系数矩阵或增广矩阵进行行变换，得到理想的形式，从而方便求解。比如，我们将（5）中的增广矩阵进行如下行变换得到（6）：

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{array}\right] \overset{r_2-3r_1}{\sim} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{array}\right] \overset{r_3-2r_2}{\sim} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 0 & 5 & -10 \end{array}\right] \tag{6}$$

主元

我们接下来观察（6）中最后的增广矩阵的系数矩阵部分，其非零行首非零元被称为**主元**（pivot），比如这里的主元是1，2和5。寻找主元的过程被称为**pivoting**。随后把主元所在的行（或列）交换到固定位置，用于随后的计算。主元所在的列组成列空间的一个基。但实际的算法很少移动矩阵的行，因为这对于大矩阵（含有几千到几百万的行与列）将招致极大的时间花费；替代的办法是仅仅记录矩阵的行的交换信息。

行阶梯型矩阵

（6）中进行行变换的结果是得到了一个**行阶梯形矩阵**（Row Echelon Form），它满足以下条件:

- 所有非零行（矩阵的行至少有一个非零元素）在所有全零行的上面。即全零行都在矩阵的底部。
- 非零行的首项系数（leading coefficient），即主元严格地比上面行的首项系数更靠右。
- 首项系数所在列，在该首项系数下面的元素都是零（前两条的推论）。

简化行阶梯形矩阵

按理说找到行阶梯形矩阵就能够对方程组进行求解，同时也找到了系数列向量的基，但我们发现，比如第一行第二个位置不为0，这种主元列上还有其他非零元素是很讨厌的，比如从第一行我们可以看出，为了求解x，我们还必须先得到y和z，为了求解y还必须先求解z。如果每一行最多只有一个主元，那么我们就能更容易地用非主元列来表示出主元列了，为此我们可以进行如下行变换：

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 0 & 5 & -10 \end{array}\right] \overset{r_2+\frac{2}{5}r_3}{\sim} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 5 & -10 \end{array}\right] \overset{r_1-r_2}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 5 & -10 \end{array}\right] \overset{r_2-\frac{1}{5}r_3}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 5 & -10 \end{array}\right] \overset{\frac{1}{2}r_2,\ \frac{1}{5}r_3}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right] \tag{7}$$

这样我们就直接解出了结果，我们将这种化简后的行阶梯形矩阵称为**简化行阶梯形矩阵**（reduced row echelon form），它需要满足额外的条件：每个首项系数是1，且是其所在列的唯一的非零元素。我们再举一个例子比如：

$$\left[\begin{array}{cccc|c} 1 & 0 & a_1 & 0 & 2 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right] \tag{8}$$

两种消元法

高斯消元法

数学上，**高斯消元法**（Gaussian Elimination），是线性代数中的一个算法，可用来为线性方程组，求出矩阵的秩，以及求出可逆方阵的逆矩阵。当用于一个矩阵时，高斯消元法会产生出一个行阶梯形矩阵。

也就是（6）中的实现过程，就是高斯消元的过程。

高斯-若当消元法

高斯－若尔当消元法（Gauss-Jordan Elimination），是高斯消元法的另一个版本。它在线性代数中用来找出线性方程组的解，其方法与高斯消元法相同。唯一相异之处就是这算法产生出来的矩阵是一个**简化行梯阵式**，而不是高斯消元法中的**行梯阵式**。

相比起高斯消元法，此算法的效率比较低，却可把方程组的解用矩阵一次过表示出来。

矩阵行的线性组合

在1中，我们将**矩阵乘列向量**理解为矩阵的列向量以列向量为系数进行线性表示，即**矩阵列的线性组合** (a combination of the columns of the matrix)。那么，如何理解矩阵的**行表示**呢？也就是如何来表示**行变换**呢：

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vec{a}_3^T \end{bmatrix} = 1\vec{a}_1^T + 2\vec{a}_2^T + 3\vec{a}_3^T \tag{9}$$

也就是说，我们可以将**行向量×矩阵**理解为矩阵的行向量以行向量为系数进行线性表示，即**矩阵行的线性组合**，最终得到一个4维行向量。那么更进一步地理解，多个行向量×矩阵也就可以表示如下，最终得到一个3×4的矩阵，即3个4维行向量：

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vec{a}_3^T \end{bmatrix} = \begin{bmatrix} 1\vec{a}_1^T + 2\vec{a}_2^T + 3\vec{a}_3 \\ 2\vec{a}_1^T + 3\vec{a}_2^T + 1\vec{a}_3 \\ 3\vec{a}_1^T + 1\vec{a}_2^T + 2\vec{a}_3 \end{bmatrix} \tag{10}$$

矩阵的初等变换

有了矩阵行表示之后，我们就可以很容易地理解(6)总矩阵的行变换，通过对矩阵左乘（右乘）可逆方阵来实现矩阵的行列变换，这就是矩阵的**初等变换**。

初等行变换

左乘初等矩阵可以实现矩阵的行变换。比如，我们是通过第2行-3倍的第1行得到(6)中间的矩阵，用矩阵变换也就是：

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix} \tag{11}$$

在此基础上，我们是通过第3行-2倍的第2行得到(6)右边的矩阵，用矩阵变换也就是：

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

(10)也可以表示为$E_{32}\left(E_{21}A\right)=U$，虽然矩阵的顺序不能变，但矩阵计算具有**结合律**(associate law)，所以也可以写成$\left(E_{32}E_{21}\right)A=U$，即A可以通过某种行变换一次性变为U。

初等列变换

右乘初等矩阵，也就是1中我们理解的那样，我们便可以实现矩阵的列变换。比如我们想交换两列的顺序：

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} c & a \\ d & b \end{bmatrix} \tag{13}$$

3 矩阵乘法和逆

矩阵乘法的解释

我们考虑以下情况的矩阵乘法结果中的某一项c_{34}：

$$\begin{bmatrix} . & . & . \\ a_{31} & a_{32} & . \\ . & . & . \end{bmatrix} \begin{bmatrix} . & b_{14} & . \\ . & b_{24} & . \\ . & . & . \end{bmatrix} = \begin{bmatrix} . & . & . \\ . & c_{34} & . \\ . & . & . \end{bmatrix} \tag{14}$$

我们有以下几种理解矩阵乘法的方式。

代数表达式

c_{34}即A矩阵的第3行与B矩阵的第4列的点积：

$$c_{34}=(row3 \quad of \quad A)\cdot(col4 \quad of \quad B)=a_{31}b_{14}+a_{32}b_{24}+\ldots=\sum_{k=1}^{n}a_{3k}b_{k4} \tag{15}$$

A列向量的线性组合

在了解了初等变换后，我们可以将AB=C这个等式看做，矩阵A的各列向量通过乘B的各列来得到C的各列，即C的各列是矩阵A各列的线性组合。

Columns of C are the combinations of columns of A.

$$\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \vec{a}_3\end{bmatrix}\begin{bmatrix}\begin{bmatrix}b_{11}\\b_{21}\\b_{31}\end{bmatrix} & \begin{bmatrix}.\\.\\.\end{bmatrix} & \begin{bmatrix}.\\.\\.\end{bmatrix}\end{bmatrix}=\begin{bmatrix}\vec{c}_1 & . & .\end{bmatrix} \tag{16}$$

B行向量的线性组合

我们还可以将AB=C这个等式看做，矩阵B的各行向量通过乘A的各行来得到C的各行，即C的各行是矩阵B的各行的线性组合。

Rows of C are the combinations of rows of B.

$$\begin{bmatrix}\begin{bmatrix}a_{11} & a_{12} & a_{13}\end{bmatrix}\\\begin{bmatrix}. & . & .\end{bmatrix}\\\begin{bmatrix}. & . & .\end{bmatrix}\end{bmatrix}\begin{bmatrix}\vec{b}_1^T\\\vec{b}_2^T\\\vec{b}_3^T\end{bmatrix}=\begin{bmatrix}\vec{c}_1^T\\.\\.\end{bmatrix} \tag{17}$$

A的列×B的行

$$\begin{bmatrix}2\\3\\4\end{bmatrix}\begin{bmatrix}1 & 6\end{bmatrix}=\begin{bmatrix}2 & 12\\3 & 18\\4 & 24\end{bmatrix} \tag{18}$$

以上有两种理解方法，即左边这个列向量分别乘1和6得到右边这两个列向量，或右边这个行向量分别乘2，3，4得到右边3个行向量。而AB其实就是A和B分别以n进行列划分和行划分后，A的列和B的行的乘积的和：

$$\begin{bmatrix}2 & 7\\3 & 8\\4 & 9\end{bmatrix}\begin{bmatrix}1 & 6\\0 & 0\end{bmatrix}=\begin{bmatrix}2\\3\\4\end{bmatrix}[1 \quad 6]+\begin{bmatrix}7\\8\\9\end{bmatrix}[0 \quad 0]=\begin{bmatrix}2 & 12\\3 & 18\\4 & 24\end{bmatrix} \tag{19}$$

分块矩阵

有时候，我们也可以通过对矩阵分块，得到一些为0矩阵或单位矩阵的项，使计算变得更加容易。以下给出了矩阵分块的一种划分，当然矩阵分块不一定局限于这种格式：

$$\begin{bmatrix}A_1 & A_2\\A_3 & A_4\end{bmatrix}\begin{bmatrix}B_1 & B_2\\B_3 & B_4\end{bmatrix}=\begin{bmatrix}A_1B_1+A_2B_2 & .\\. & .\end{bmatrix} \tag{20}$$

逆矩阵

对于方阵A，如果有$A^{-1}A=AA^{-1}=I$，则称A是可逆矩阵(invertible)或非奇异矩阵(non-sigular)，反之则称方阵A为不可逆矩阵或奇异矩阵。

如何理解或判定矩阵是否可逆

行列式是否为0

如果$|A| = 0$，则A不可逆，反之可逆。

秩是否为n

如果$rank\left(A\right) = n$，则A可逆，否则A不可逆。

列向量是否线性相关

以列向量分析，在n维空间内，如果列向量线性相关，即线性无关的列向量个数小于n，比如是k，那么由这n个列向量表示出的向量一定在k维平面内，而等式右端的I各列共有n维向量，所以I无法由这些列向量线性表出，即不存在与A相乘积为I的矩阵。

Ax=0

如果能找到一个非零向量$\vec{x}$，使得$A\vec{x} = \vec{0}$成立，则A不可逆。因为如果A^{-1}存在，将其左乘于等式，便能得到$\vec{x} = \vec{0}$，这与非零向量的假设矛盾了。

求逆矩阵

高斯-若当消元

求逆矩阵的本质是同时对n个方程组进行高斯消元或进行初等变换。我们可以通过对其增广矩阵进行行变换，使A变为I，与此同时原本的I就变成了A^{-1}，其原理如下：

$$\begin{aligned}
&E_n \ldots E_2 E_1 A = I = A^{-1} A \\
&E_n \ldots E_2 E_1 I = E_n \ldots E_2 E_1 = A^{-1}
\end{aligned} \tag{21}$$