Linear Algebra

邱泽辉

December 2024

摘要

由于线性代数前后交叉内容极多，相较于高数渐进式学习不同，线性代数是一种回环往复，螺旋上升，柳暗花明的感觉，只有综合理解各个知识点，才能掌握。故作此文。

1 序

或许由于教材的安排原因，学习线性代数就像被裹了层雾一样，学习矩阵时不知道向量，学习行列式时不知道矩阵，导致这些有机统一的东西，无法被准确的学习，非此即彼，没办法综合起来，以至于到后面一头雾水，茫然无措。因此，以一个俯瞰全书的视角来看待线性代数，是个不错的选择，看到后面的路，更能有信心走下去。

Calvin: You know, I don’t think math is a science, I think it’s a religion.

Hobbes: A religion?

Calvin:Yeah. All these equations are like miracles. You take two numbers and when you add them, they magically become one NEW number! No one can say how it happens. You either believe it or you don’t.

卡尔文：你知道吗，我觉得数学不是一门科学，而是一种宗教

霍布斯：一种宗教？

卡尔文：是啊。这些公式就像奇迹一样。你取出两个数字，把它们相加时，它们神奇地成为了一个全新的数！没人能说清楚这种事情到底是怎么发生的。你要么完全相信，要么完全不信。

2 线性方程组

线性代数最早就是为了解决线性方程组的问题，所以我们先给出线性方程组的定义

2.1 齐次线性方程组

Definition 2.1 齐次线性方程组

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = 0\\
&a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = 0\\
&\cdots\\
&a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = 0
\end{aligned}
$$

为 n 元齐次线性方程组

但是这样写太复杂了，根据初中的做法，我们加减消元，其实只需要对系数做文章就行，所以我们把系数专门拿出来，写成一个表格的形式

Definition 2.2 系数矩阵
将方程组中的系数写成一个数表，称为系数矩阵

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

没错，虽然矩阵在后面会有十分神奇的东西，但在此时，他就是个数表，就是一群数的表格

我们解线性方程组，就是为了得到所有未知数的值，所以我们把这些未知数的值也写成一个表格，不过这个表格只有一列，我们叫做解向量，虽然向量很高级，但在这里，他就是一列数的表格

Definition 2.3 解向量
将方程组中的未知数写成一个数表，称为解向量

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

所以，有了这两个定义，我们就可以简化齐次线性方程组的写法，学过后面课程的我们已经知道，利用矩阵乘法就行

Definition 2.4 齐次线性方程组的矩阵观点
$AX = O$

而且，我们也可以把线性方程组看成向量的线性组合

Definition 2.5 齐次线性方程组的向量观点
$x_1\alpha_1 + x_2\alpha_2 + ... + x_n\alpha_n = 0$

Theory 2.6 齐次方程组必有零解

显然 $x = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ 是齐次方程组的解，不过这个解与具体的齐次方程组无关，太平凡了 $trivial$, 因此称为平凡解

2.2 非齐次线性方程组

Definition 2.7 非齐次线性方程组

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

为 n 元非齐次线性方程组

但这时候右边不是零了，我们可以类比解向量，把右边的常数写成一列的形式，称为**增广列**

同时我们还可以定义增广矩阵

Definition 2.8 增广矩阵
系数矩阵加上增广列就是增广矩阵，记作 $\overline{A}$
$\overline{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{pmatrix}$

为了明显一点，也可以画一条虚线区分增广列，这在解方程的时候十分必要，有助于认识线性方程组。

$$\overline{A} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

与此同时也产生了一些新的概念

Definition 2.9 相容 如果非齐次线性方程组有解，称这个非齐次线性方程组相容

Definition 2.10 导出组 当右边常数都为 *0* 时，变为了一个齐次线性方程组，称为这个齐次线性方程组的导出组

2.3 线性方程组的解

在中学阶段，我们已经学会了解简单的二元，三元一次线性方程组，在此进一步讨论之前，先引出一些概念

Definition 2.11 主元 *(首元)*

解方程，我们最后实际上就想得到这样一个东西

$$\begin{gathered} x_1 = c_1 \\ x_2 = c_2 \\ \vdots \\ x_n = c_n \end{gathered}$$

这实际上也是一个方程，别看他十分简单，这既是解，又是方程，既然是方程，那么就对应了增广矩阵，我们把它写出来

$$\overline{A} = \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_n \end{array}\right)$$

我们想实现的效果就是，每行只有一个元素为 *1*，其他都是 *0*，这样的矩阵对应的就是解的方程组

但是，这个是最好的顺序，实际上解的顺序可以调换，比如

$$\begin{gathered} x_1 = c_1 \\ x_3 = c_3 \\ x_2 = c_2 \\ \vdots \\ x_n = c_n \end{gathered}$$

其对应的增广矩阵为：

$$\overline{A} = \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & c_1 \\ 0 & 0 & \cdots & 1 & c_3 \\ 0 & 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_n \end{array}\right)$$

总而言之，每行需要但也只需要一个不为 *0* 的 *1*，但是这个 *1* 出现在哪里，倒是无所谓

我们把每行第一个不为 *0* 的元素，**称为这行的首元。**

并且称此时的矩阵为行最简形，这个后面再说

显然，虽然这两个矩阵不同，但他们对应的方程组同解

为了引出高斯消元法，还需要介绍同解变换

Definition 2.12 同解变换 *(初等变换)*

1. 交换两个方程的位置。

2. 将一个方程乘以一个非零常数。

3. 将一个方程的倍数加到另一个方程上。

把这种概念移植到矩阵上，就成了矩阵的初等行变换

Definition 2.13 矩阵初等行变换

1. 交换矩阵两行的位置。

2. 将矩阵的一行乘以一个非零常数。

3. 将一行的倍数加到另一个行上。

前面的矩阵和他的行最简形同解，而行最简形是由原矩阵经过初等变换得来的，为了表述这种关系，引入矩阵等价的概念

Definition 2.14 矩阵等价 设 A 和 B 是两个 $m \times n$ 矩阵。我们称 A 和 B 是**等价**（或 **行等价**），如果存在一个有限的初等行变换序列可以将 A 变换为 B。记作 $A \cong B$

Theory 2.15 高斯消元法 关于高斯消元法，其实就是初中学的加减消元法，那为什么这里还要再提呢，那自然是有些不同了，首先高斯消元法的操作是

1. 交换两个方程的位置。

2. 将一个方程乘以一个非零常数。

3. 将一个方程的倍数加到另一个方程上。

$$\begin{cases} x+y+z &= 6 \\ 2x+3y+4z &= 20 \\ 4x+5y+6z &= 36 \end{cases}$$

对应的增广矩阵为：

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 2 & 3 & 4 & 20 \\ 4 & 5 & 6 & 36 \end{array}\right]$$

拿到这个方程，如何应对？前面说了，我们希望得到 $x_1 = ..., x_2 = ...$ 的形式，换而言之，就是每行一个未知数，告诉未知数的值，用前面的定义，就是每行只有一个首元

那么我们的想法是，按照顺序来，第一行的首元是 x_1，第二行是 x_2，这样，除了第一行，其他行都不能有 x_1，那么只要把第一列的系数都消成除了第一行都是 *0* 的形式

- $r_2 = r_2 - 2r_1$ （从第二个方程中减去第一个方程的两倍）
- $r_3 = r_3 - 4r_1$ （从第三个方程中减去第一个方程的四倍）

得到新的矩阵：

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 1 & 2 & 12 \end{array}\right]$$

这里有一个问题，如果第一行 x_1 的系数是 *0* 的话，就没法消其他的了，所以就要用交换操作，把系数不为 *0* 的调到第一行

这又有一个问题，如果都是 *0* 呢?，这不可能，因为这样方程里都没有 x_1，你是怎么写出这个矩阵的？

不过，如果只从矩阵形式，确实可能存在这种情况，这样的话，我们从 x_2 开始选主元就可以，

回到正题，这样其他行就都没有 x_1 了，现在我们让第二行只有 x_2

- $r_3 = r_3 - r_2$ （从第三个方程中减去第二个方程）

得到新的矩阵：

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 0 & 4 \end{array}\right]$$

诶，这时候巧合来了，第三行没有主元了，如果我们把最后一行翻译一下，最后一行表示了等式 $0x + 0y + 0z = 4$，这是不可能的，因此这个特定的方程组是不相容的（没有解）。这就得到了非齐次线性方程组有解的条件。

不过我们的讨论还得继续，这毕竟只是巧合，假设我们改一个数字让他出现主元，这样就相容了，例如：

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

此时你会发现第三行只有一个主元，那就是一个一元一次方程啊！这小学就会了，这样实际上形成了一种阶梯型的方程组，越上面的方程组越多未知数，越下面的越少，所以我们只要从下往上解就可以

这个过程称为回代

$$\begin{aligned} z &= 4 \\ y + 2z = 8 &\Rightarrow y = 8 - 2(4) = 0 \\ x + y + z = 6 &\Rightarrow x = 6 - 0 - 4 = 2 \end{aligned}$$

所以解为 $(x, y, z) = (2, 0, 4)$。

对此应高度重视，这是线性代数的开始。

从例子中我们也可以看到

如果非齐次线性方程组相容，那么不应该存在 $[0\,0\,0\,0...b]$ 这种东西，也就是化为最简形之后，左边全 0，右边却不为 0

Theory 2.16 非齐次线性方程组有解的充要条件

增广矩阵的最右列不是主元列 (最后一行没有主元),

即不存在 $\begin{pmatrix} 0 & 0 & ... & 0 & b \end{pmatrix}$

下面总结一下高斯消元法的算法

高斯消元法算法

1、选择主元：从当前处理的列中选择一个非零元素作为“主元”，

2、消除下方元素：使用选定的主元，通过适当的行操作，将主元下面的所有元素变成 0。

3、重复上述步骤，直到所有可能的列都处理完毕，或者遇到全零行

经过高斯消元法后，矩阵会变成一种一半都是 0，一半不全是 0 的结构，我们称为是上三角矩阵

这未免有点巧合，我们再看一个例子

考虑如下线性方程组：

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 4 \\ 2x_1 + 4x_2 - 2x_3 + 2x_4 = 8 \\ -x_1 - 2x_2 + x_3 - x_4 = -4 \end{cases}$$

其增广矩阵表示为：

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 4 \\ 2 & 4 & -2 & 2 & 8 \\ -1 & -2 & 1 & -1 & -4 \end{array}\right]$$

通过高斯消元法，我们将第二个方程减去第一个方程的两倍，得到：

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

我们发现，下面全是 0 了！这和前面可不一样，这说明后面三个变量没有限制条件，可以随意取值，都满足方程组，因为这个方程组相当于只有第一个方程，只有 x_1 受限制，其他变量都是自由的，我们称为是**自由变量**

此时，x_2, x_3, 和 x_4 是自由变量，而 x_1 可以用其他变量表示。

从第一个方程中解出 x_1：

$$x_1 = 4 - 2x_2 + x_3 - x_4$$

所以该线性方程组的通解可以写成：

$$\begin{cases} x_1 = 4 - 2x_2 + x_3 - x_4 \\ x_2, x_3, x_4 \text{ 为自由变量} \end{cases}$$

这完全颠覆了我们中学的认知，原来方程组的解并非一定是一个确定的值，而可以是一个集合

3 矩阵代数

前面研究的是线性方程组，我们对矩阵的认识，仅仅停留在方程组的系数上面，但是，这种数据的结构，也同样值得研究。

矩阵可以看成是一群数的集合，那么，矩阵的性质，自然应该是一群数的性质，而非一个数的性质，不过，研究一群数，还是得从一个数开始，这就要求我们选出代表元

所以我们可以完善矩阵的定义

3.1 矩阵定义、加法、数乘、乘法

Definition 3.1 a_{ij} 形成的数表称为矩阵，特别地，如果 $m=n$，则称为方阵

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

记作 $A=(a_{ij})_{m\times n}$

Definition 3.2 矩阵加法

设 $A=(a_{ij})$ 和 $B=(b_{ij})$ 是两个 $m\times n$ 矩阵。矩阵加法 $C=A+B$ 的结果也是一个 $m\times n$ 矩阵 $C=(c_{ij})$，其中每个元素定义为：

$$c_{ij}=a_{ij}+b_{ij},\quad \forall i\in\{1,2,\ldots,m\},\forall j\in\{1,2,\ldots,n\}$$

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

Definition 3.3 矩阵数乘

设 k 是一个标量，$A=(a_{ij})$ 是一个 $m\times n$ 矩阵。数乘 $B=kA$ 的结果是一个 $m\times n$ 矩阵 $B=(b_{ij})$，其中每个元素定义为：

$$b_{ij}=ka_{ij},\quad \forall i\in\{1,2,\ldots,m\},\forall j\in\{1,2,\ldots,n\}$$

$$B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}=kA=\begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}$$

减法可由加法和数乘定义，这里不做解释

尽管加法减法，数乘的定义十分简单，不过，对于矩阵的乘法，我们一筹莫展，如何定义矩阵的乘法？

我们想到，一元一次方程的形式就是 ax=b, 我们前面已经定义了系数矩阵，解向量，增广列向量

那么很自然就可以把线性方程组写成

$Ax = b$ 的这种形式，其中 x 和 b 都是向量，A 是一个矩阵

也就是说，我们可以定义矩阵和向量的乘法，乘完之后是一个新的向量 b

我们知道，在线性方程组中，**系数矩阵的行数应该就等于增广列的行数，就是有多少个方程，列数就等于未知数的个数**，如果有 m 个 n 元方程，那么系数矩阵就是 $m \times n$ 的矩阵

Definition 3.4 矩阵和向量乘法

设 $A = (a_{ij})$ 是一个 $m \times n$ 矩阵，$\mathbf{x} = (x_j)$ 是一个 $n \times 1$ 的列向量。矩阵 A 与向量 $\mathbf{x}$ 的乘积 $\mathbf{b} = A\mathbf{x}$ 是一个 $m \times 1$ 的列向量，其元素 b_i 定义为：

$$b_i = \sum_{j=1}^{n} a_{ij}x_j, \quad \forall i \in \{1, 2, \ldots, m\}$$

用矩阵形式表示：

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

则

$$\mathbf{b} = A\mathbf{x} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{pmatrix}$$

你可能会问，为什么不把矩阵乘法定义为两个矩阵的元素对应相乘呢？类似加法和数乘那样？

不过定义总是需要一些合理性，我们看看为什么这样定义式不合理的？

由于矩阵不应该脱离或者说暂时不应该脱离方程组的语境

我们假设 $Ax = b$，如果做乘法应该是什么样的？

PAx, 因为 $Ax = b$, 所以就是 $PAx = Pb$

b 和 x 没有本质区别，也是一个向量，所以 $Pb = c$

这对应一个方程组

$$
\begin{aligned}
&p_{11}b_1 + p_{12}b_2 + ... + p_{1m}b_n = c_1\\
&p_{21}b_1 + p_{22}b_2 + ... + p_{2m}b_m = c_2\\
&\cdots\\
&p_{m1}b_1 + p_{m2}b_2 + ... + p_{mm}b_m = c_m
\end{aligned}
$$

不管怎么说，PA 应该是一个矩阵，所以也可以写成

$(PA)x = c$, 假设 $PA = d_{ij}$

而

$$
\begin{aligned}
&p_{11}b_1 + p_{12}b_2 + ... + p_{1m}b_m =\\
&p_{11}(a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n) + p_{12}(a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n) + ...+\\
&p_{1m}(a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n) = c_1
\end{aligned}
$$

整理一下就会发现

$$\sum_{i=1}^{n}(p_{1i}a_{i1})x_1 + \sum_{i=1}^{m}(p_{1i}a_{i2})x_2 + ... \sum_{i=1}^{m}(p_{1i}a_{in})x_m = c_1$$

x 前面的系数，就是矩阵 PA 的元素

所以显然矩阵乘法不是直接两个元素相乘，这同前面矩阵的引出是不吻合的，简单来说就是没有意义

我们发现矩阵乘法 PA 其实就是，我们知道 b 和 c 的关系，而 b 和 x 又有关系，我们想要建立 x 到 c 的关系，也就一种变换，这种思想非常重要，如果我们想用一组变量，表示另一组变量，其实就是乘以一个变换矩阵乘以这组变量

因此矩阵乘法可以定义为

Definition 3.5 矩阵乘法

如果 $A_{m\times n} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}$ $B_{n\times t} = (b_{ij}) = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1t}\\ b_{21} & b_{22} & \cdots & b_{2t}\\ \vdots & \vdots & \ddots & \vdots\\ b_{n1} & b_{n2} & \cdots & b_{nt}\end{pmatrix}$

那么 AB 是有意义的，并且我们定义 $C_{m\times t} = (c_{ij}) = \sum_{u=1}^{n} a_{iu}b_{uj}$

根据上面的分析，B 的列数是 x 的个数，B 的行数是 b 的个数，A 的列数是 b 的个数，也就说 A 的列数是 B 的行数

Theory 3.6 矩阵相乘条件

矩阵 A 和 B 能够相乘的条件是 A 的列数 $=B$ 的行数

定义完了运算，就可以研究这些运算的规律了，

3.2 矩阵的加法交换结合律，乘法结合律

由于代表元素是数，具有加法交换律、结合律，而矩阵的加法就是对应元素相加，则矩阵加法具有交换律，结合律

Theory 3.7 矩阵加法交换律、结合律

$$A + B = B + A$$
$$A + B + C = A + (B + C) = (A + C) + B$$

但是，矩阵乘法并不是简单的两个矩阵对应元素相乘，所以未必满足交换律和结合律，我们需要验证

例如

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

$B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$

$C = \begin{pmatrix} 9 & 10 \\ 11 & 12 \end{pmatrix}$

$AB = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$

$(AB)C = \begin{pmatrix} 413 & 454 \\ 937 & 1030 \end{pmatrix}$

$A(BC) = \begin{pmatrix} 413 & 454 \\ 937 & 1030 \end{pmatrix}$

$BA = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$

显然 $AB \neq BA, A(BC) = (AB)C$

因此矩阵乘法没有交换律, 可能有结合律

如果矩阵没有交换律，那么方向就很重要了，左边乘法叫做“左乘”，右边乘法叫做“右乘”

Theory 3.8 矩阵乘法结合律 $A(BC) = (AB)C$

这个证明非常困难，不过思想很简单，就是利用矩阵定义，把矩阵的性质变为和式的性质

Prove 3.9

设$A = (a_{ij})_{m\times n}, B = (b_{ij})_{n\times s}, C = (c_{ij})_{s\times t}$

$D = BC = (d_{ij})_{n\times t}, F = AB = (e_{ij})_{m\times s},$

下证$AD = FC$

$d_{ij} = \sum_{u_1=1}^{s} b_{iu_1}c_{u_1j}$

$e_{ij} = \sum_{u_2=1}^{n} a_{iu_2}b_{u_2j}$

$AD(i,j) = \sum_{u=1}^{n} a_{iu}d_{uj} = \sum_{u=1}^{n} a_{iu}\sum_{u_1=1}^{n} b_{uu_1}c_{u_1j} = \sum_{u=1}^{n}\sum_{u_1=1}^{s} a_{iu}b_{uu_1}c_{u_1j}$

$FC(i,j) = \sum_{u_3=1}^{s} e_{iu_3}c_{u_3j} = \sum_{u_3=1}^{s}\sum_{u_2=1}^{n} a_{iu_2}b_{u_2u_3}c_{u_3j}$

由于u_1与u_3遍历过程一样 由于u_2与u遍历过程一样，则

$FC(i,j) = \sum_{u_1=1}^{s}\sum_{u=1}^{n} a_{iu}b_{uu_1}c_{u_1j}$

$AD = FC$

证毕

上述式子需要一个引理

Corollary 3.10 双重求和可以交换顺序

Prove 3.11 考虑二重求和：

$$S = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}$$

我们要证明它可以与如下形式互换：

$$T = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}$$

$$S = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}$$

$$S = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\right)$$

$$T = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} \right)$$

注意到在 S 中，每个 a_{ij} 都会出现在第 i 行的求和中，然后在所有行的求和中；而在 T 中，每个 a_{ij} 都会出现在第 j 列的求和中，然后在所有列的求和中。因此，无论是先按行求和再按列求和，还是先按列求和再按行求和，我们实际上都在遍历所有的 a_{ij} 并将它们相加。

事实上如果把 a_{ij} 写成一个矩阵，这两种方法都是在求矩阵的所有元素的和，因此可以交换顺序

结合律看似简单，实际上很重要，这表明我们可以规定矩阵乘法的次序，无论是从左到右，还是从右往左都可以，只要不改变矩阵的顺序，最后结果是一样的

不过，就算矩阵不满足交换律，难道就没有矩阵能够交换吗？我们来寻找一下

Question 3.12 (可交换矩阵) 是否能找到一对矩阵 P，A，满足交换律，使得 $PA = AP$，

怎么找到这样的 P 呢，还是利用矩阵乘法定义根据矩阵乘法，PA 的行数是 P 的行数，AP 的行数是 P 的列数，则 P 的行数等于列数等于 A 的列数，P 为方阵，同理 A 也是方阵，且 A，P 大小一样

$$A = (a_{ij})_{n\times n}, P = (p_{ij})_{n\times n},$$
$$PA(i,j) = \sum_{u=1}^{n} p_{iu}a_{uj} = \sum_{u=1}^{n} a_{uj}p_{iu}$$
$$AP(i,j) = \sum_{u=1}^{n} a_{iu}p_{uj}$$

注意到如果 P 中所有元素为 0，则显然成立

我们知道乘法是有单位元和零元的，就是说存在一个数，乘以任何数都是任何数本身，也存在一个数，乘以任何数都是这个数，这就是 1 和 0，那么矩阵中有没有这样的东西呢？

显然我们上面就发现了零元，就是元素全为 0 的矩阵

Definition 3.13 (零矩阵) 元素全为 0 的矩阵，记作 O

Property 3.14 $A_{m\times n}O_{n\times p} = O_{m\times p}$

如果为方阵，则 $AO = OA = O$

前面的问题还没完，还有没有其他矩阵可以交换呢？我们还是考虑方程组，我们前面说了高斯消元法最后就是想得到这样的方程

$$\begin{aligned} x_1 &= c_1 \\ x_2 &= c_2 \\ &\vdots \\ x_n &= c_n \end{aligned}$$

这对应的系数矩阵是

$$P = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

我们发现，如果 $PX = B$, 那么 $X = B = PX$, 换而言之如果 $AB = C$，那么 $APX = C, PAX = PAB = PC = C = APX$

这种 P 是可以交换的，而且，我们通过计算发现 $PA = AP = A$

这是非常好的性质，我们确认这就是单位元

于是定义

Definition 3.15 单位矩阵

对角线上全为 *1*，其他位置均为 *0* 的方阵为单位矩阵，形如，

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

记作 E

你以为就结束了吗？显然我们给单位矩阵乘上一个系数，也仍然符合题意，那么我们可以定义数量矩阵

Definition 3.16 数量矩阵

对角线上全为 k，其他位置均为 *0* 的方阵为数量矩阵，形如，

$$\begin{pmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k \end{pmatrix}$$

这三种矩阵都符合题意，我们发现这些矩阵都是对角线上有非零元素，而其他地方都是 0 的矩阵

我们继续给出定义，

Definition 3.17 对角矩阵

除了主对角线上元素，其他位置均为 0 的方阵为对角矩阵，形如，

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

那么，对角矩阵是否可以交换呢？如果不是数量矩阵就不行

举个例子即可

Question 3.18 证明：与任意方阵可交换的是同阶的数量矩阵

$$\begin{aligned} &A=(a_{ij})_{n\times n}, P=(p_{ij})_{n\times n}, \\ &PA(i,j)=\sum_{u=1}^{n} p_{iu}a_{uj}=\sum_{u=1}^{n} a_{uj}p_{iu} \\ &AP(i,j)=\sum_{u=1}^{n} a_{iu}p_{uj} \end{aligned}$$

先来点好算的，A 是个对角矩阵的话，只要 a 角标两个不相等，那肯定就是 0，只会剩下一项

$$\begin{aligned} PA(i,j)&=\sum_{u=1}^{n} p_{iu}a_{uj}=a_{jj}p_{ij} \\ AP(i,j)&=\sum_{u=1}^{n} a_{iu}p_{uj}=a_{ii}p_{ij} \end{aligned}$$

如果 A 对角线上元素都不相等，那说明什么，对于任意 $i \neq j, p_{ij}=0$ 所以 P 就只能是对角矩阵了，这时候假设 A 是其他的，我们再看看

$$\begin{aligned} PA(i,j)&=\sum_{u=1}^{n} p_{iu}a_{uj}=a_{ij}p_{ii} \\ AP(i,j)&=\sum_{u=1}^{n} a_{iu}p_{uj}=a_{ij}p_{jj} \end{aligned}$$

假设 A 是一个全不为 0 的矩阵，那么 $p_{11}=p_{22}....$ 所以 P 就只能是数量矩阵，证毕

你以为这就结束了吗？

3.3 特殊矩阵

定义完了矩阵的乘法之后，就提供了很多新鲜的视角，我们说矩阵乘法其实就是线性替换，把一个矩阵变成另一个矩阵，而向量可以看成是只有一列的矩阵，

$A\boldsymbol{x}=\boldsymbol{b}$ 这就是说 x 左乘 A 之后变成了 $\boldsymbol{b}$

注意变成这两个字，这说明矩阵就像一个机器，你输入一个东西，按照一定的规则，出来另一个东西，这让你联想到了什么？

函数，是的，矩阵可以对应一个函数，不过这是后话了

不过矩阵暂时还得依赖方程组语境，我们在解方程组的时候，不是用到了一个高斯消元法吗？**解方程最后其实就把系数矩阵变成一个单位矩阵**, 这个思想很重要，那么这也是变化的一个过程，能不能用矩阵乘法来代替呢？

比如我们要实现交换两个方程，也就是交换矩阵的两行，怎么做到？

由于引入的时候用的是左乘

我们假设 P 这个矩阵可以交换 A 矩阵的两行，那么 P 矩阵是什么？

我们知道，方程组最后的解，就是按 x_1,x_2 这样排序的，交换 A 矩阵的两行，不会改变 A 对应方程组的解，但是会改变的解的顺序，比如本来解是 x_1,x_2,x_3，现在变成了 x_1,x_3,x_2

思考，经过 A 变换后 x_1,x_2,x_3 变成了 b_1,b_2,b_3，那如果我想交换两行，变成 x_1,x_2,x_3 对应 b_1,b_3,b_2, 这简单，先让 x_1,x_2,x_3 变成了 b_1,b_2,b_3(A), 再让 b_1,b_2,b_3 变成 x_1,x_3,x_2(C) 就行

所以 C 对应的方程

$$
\begin{aligned}
b_1 &= x_1 \\
b_3 &= x_2 \\
&\vdots \\
b_2 &= x_3
\end{aligned}
$$

写出对应矩阵就是 $C=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}$ 我们发现这个矩阵和单位矩阵很像，就是单位矩阵交换了第二行和第三行，而我们知道 EA=A, 乘以 E 相当于什么都没做，那么交换 E 的两行，就得到了 C 矩阵。我们称这样的矩阵为**置换矩阵**

同样的方法我们也能得到其他两个初等变换对应的矩阵，并且这些矩阵只需要单位矩阵进行一次变换

Definition 3.19 *初等矩阵 单位矩阵经过一次初等变换得到的矩阵*

左乘以初等矩阵，就相当于矩阵做了和单位矩阵一样的初等行变换

3.4 矩阵的转置，多项式，幂

矩阵的转置内容很深，这里仅做表层不过，我倾向于认为是先射箭再画靶，转置实际上就是为了对称矩阵这个醋包的饺子。怎么说？

中学的时候我们学过，关于 y=x 直线对称的点具有什么特征？x，y 坐标互换！也就是横纵坐标互换，横纵其实就是行列

那关于方阵对角线对称的元素，自然就是行列指标互换了

所以我们有了转置的概念

Definition 3.20 (矩阵转置) 设 A 是一个 $m \times n$ 的矩阵，其元素记为 a_{ij}，其中 i 是行索引，j 是列索引。那么 A 的转置矩阵 A^T（或 A'）是一个 $n \times m$ 的矩阵，其元素 b_{ji} 满足：

$$b_{ji} = a_{ij}$$

换句话说，矩阵 A 的转置 A^T 是通过将 A 的行变成列，列变成行而得到的矩阵。即第 i 行第 j 列的元素在转置后变为第 j 行第 i 列的元素。

例如，如果有一个 2×3 的矩阵 A：

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

那么它的转置矩阵 A^T 是一个 3×2 的矩阵：

$$A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

Definition 3.21 (对称矩阵) 设 A 是一个 $n \times n$ 的方阵。如果 A 满足以下条件，则称 A 是一个**对称矩阵**：

$$A = A^T$$

即对于所有 i, j，有：

$$a_{ij} = a_{ji}$$

换句话说，对称矩阵是指沿主对角线对称的矩阵，即矩阵中任意元素与其关于主对角线对称位置的元素相等。

例如，考虑一个 3×3 的对称矩阵 B：

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$$

在这个例子中，矩阵 B 是对称的，因为 $b_{12} = b_{21}$，$b_{13} = b_{31}$，以及 $b_{23} = b_{32}$。

3.5 转置的性质

设 A 和 B 是两个适当维度的矩阵，c 是一个标量。矩阵转置有以下性质：

Property 3.22 (双重转置)

$$(A^T)^T = A$$

矩阵转置两次后等于原矩阵。

Property 3.23 (加法转置)

$$(A+B)^T = A^T + B^T$$

两个矩阵相加后的转置等于它们各自转置后再相加。

Property 3.24 (数乘转置)

$$(cA)^T = cA^T$$

标量与矩阵相乘后的转置等于标量乘以矩阵的转置。

Property 3.25 (乘法转置（注意顺序反转）)

$$(AB)^T = B^T A^T$$

两个矩阵相乘后的转置等于第二个矩阵的转置乘以第一个矩阵的转置。请注意，这里顺序是反转的。

Property 3.26 (逆矩阵转置（如果存在逆矩阵）)

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

矩阵的逆矩阵的转置等于矩阵的转置的逆矩阵。这个性质也表明了可逆矩阵的转置仍然是可逆的。

Property 3.27 (迹的性质)

$$tr(A^T) = tr(A)$$

矩阵的迹（即主对角线元素之和）等于其转置矩阵的迹。

Property 3.28 (行列式的性质（对于方阵）)

$$\det(A^T) = \det(A)$$

方阵的行列式等于其转置矩阵的行列式。

Property 3.29 (内积（点积）的性质) 如果 A 和 B 都是列向量，则

$$A^T B = (B^T A)^T$$

即两个向量的内积是一个标量，并且等于其转置。

Property 3.30 (正交矩阵的性质) 对于正交矩阵 Q，有

$$Q^T Q = QQ^T = I$$

其中 I 是单位矩阵。这表示正交矩阵的转置等于它的逆矩阵。

转置的性质很多，但大部分都是显而易见的，我们只证明 3.24 **证明：**

$$\text{设} A = (a_{ij})_{m\times n}, B = (b_{ij})_{n\times s} AB(i,j) = \sum_{u=1}^{n} a_{iu} b_{uj}$$

3.6 矩阵的逆———Jordan 消元法

矩阵代表的是线性变换，将一个向量映射到另一个向量上，那么，肯定也有一个矩阵，把向量还原到映射前的向量，这个矩阵和前面的矩阵互为逆向，当这两个矩阵同时作用时，向量不动，所以乘积为单位矩阵

Definition 3.31 可逆矩阵 $n \times n$ 方阵 A，如果存在 $n \times n$ 方阵 B，满足 $AB = BA = E$，则称 B 为 A 的逆矩阵，记作 A^{-1}，并称矩阵 A 为可逆矩阵，非奇异矩阵，如果 A 不可逆，称为奇异矩阵

不过，得论证一下逆矩阵是唯一的，不然如何说 B 就是 A 的逆矩阵呢？假设 $AC = CA = AB = BA = E$，则 $(BA)C = EC = C = B(AC) = BE = B$，则 $B = C$

这里就是乘法结合律的重要应用

说到乘法，那就不得不谈谈逆的乘法性质，如果 A，B 都可逆，那么 AB 可逆吗？

Theory 3.32 可逆的乘法性质

如果 n 阶方阵 A，B 均可逆，则 $(AB)^{-1} = B^{-1}A^{-1}$

证明：

$B^{-1}A^{-1}AB = B^{-1}EB = E$

上面的定理可以用归纳法推广到任意有限个数矩阵

显然，逆矩阵可以用来解方程组，方程组我们知道的就是 x 经过 A 线性变换得到的 b，那么只需要两边同时乘以 A 的逆，将 b 映射到 x，就能求出 x

Theory 3.33 逆矩阵与方程组

若 A 为 $n \times n$ 可逆矩阵，则对于任意 $\boldsymbol{b} \in R^n$,$Ax = b$ 有唯一解，$x = A^{-1}\boldsymbol{b}$

必要性证明：两边同时乘以 A^{-1}

$$\begin{aligned} A^{-1}Ax &= A^{-1}b \\ x &= A^{-1}b \end{aligned}$$

充分性证明：

$$\begin{gathered} \text{若方程组} Ax = b, \text{有解} x = Cb, \text{求证:} AC = CA = E \\ \text{由于 } b \text{ 是任意的,不妨取} \boldsymbol{b} = \boldsymbol{e_1}, \boldsymbol{e_2}, .., \boldsymbol{e_n}, Ax_i = e_i, \\ \text{令} C = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}, \text{根据矩阵乘法}, AC = E \end{gathered}$$

不过，这里就卡住了，如果 AC=E，怎么证明 CA=E 呢？

还是从方程组的角度出发，如果方程组是有解的，那么系数矩阵每一行都应该有主元！否则就出现 000b 这种东西！

如果方程组有解，那意味着系数矩阵每一行都有主元，而由高斯消元法，我们可以得到一个上三角矩阵，但这还不够，我们其实是想把系数矩阵变成一个单位矩阵，这样增广列直接就是解

那就需要介绍 Jordan 消元法了

还是之前那个例子

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

我们已经做到了每一行都有主元，也就是化成了阶梯型矩阵

不过，第二行有两个元，第一行有三个未知数，这还是不方便，我们得一个一个解，这时候我们就想，能不能让每一行只剩主元，这当然可以，只要把高斯消元法反过来再来一次就行

从最后一行开始，第三行主元上面的元素都消成 0，

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

这时候你发现第一行还有两个，那么就把第二行主元上面的元素消成 0

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

那么这样就出现了一个单位矩阵，我们的解就是右边的数字

当然也有可能得到这样的矩阵

$$\left[\begin{array}{ccc|c} 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

这种我们只要把第一行除以 2 就行

当然也未必都是单位矩阵

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

这里 x_4 不是主元，我们得到的不是单位矩阵，不过包含主元的部分就是单位矩阵

我们把这种经过 Gauss-Jordan 消元法得到的矩阵称为**行最简形矩阵**

Definition 3.34 行最简形 *每个非零行的第一个非零元素（即主元）是 1。*

每个主元所在的列中，除了该主元外，其他元素都是 0。

主元所在的列数随着行数增加而严格递增。

所有全零行（如果有）位于矩阵的底部。

满足这样条件的矩阵定义为行最简形

如果矩阵可逆，那么线性方程组就有解，那么，说明系数矩阵的每一行都有主元，那么，系数矩阵就可以经过初等变换变为单位矩阵

不过，反过来也可以说，单位矩阵可以经过初等变换变为 A，只需要取完全相反的动作就行

这说明初等矩阵都可逆，而且逆就是相反的变换过程

Theory 3.35 *可逆矩阵行等价于单位矩阵 $A \cong E \Leftrightarrow A$ 可逆*

充分性证明： *如果 A 可逆，则 $A\boldsymbol{x} = \boldsymbol{b}$ 有解，则 A 的每一行都有主元，又 A 为方阵，所以主元都在对角线上，所以 $A \cong E$*

必要性证明： *若 $A \cong E$，则存在 $E_pE_{p-1}...E_1A = E$，由于初等矩阵都可逆，$A = (E_pE_{p-1}...E_1)^{-1}E = (E_pE_{p-1}...E_1)^{-1}$*

则 $AE_pE_{p-1}...E_1 = (E_pE_{p-1}...E_1)^{-1}E_pE_{p-1}...E_1 = E$ 因此 A 可逆

现在我们可以完成上面定理的证明了，不过，先来证一下引理

Theory 3.36 可逆矩阵可交换 *设 A，B 为 $n \times n$ 矩阵，若 $AB = E$，那么，$BA = E$，则 A 和 B 都可逆*

证明：

如果 AB=E，考虑 Bx=0，由于 ABx=x=0，则只有平凡解，这说明 B 有 n 个主元，则 $B \cong E$，则 B 可逆，$ABB^{-1} = B^{-1}$，则 $A = B^{-1}$，则 BA=E

于是马上有

若 A 为 $n \times n$ 可逆矩阵，则对于任意 $\boldsymbol{b} \in R^n$,$Ax = b$ 有唯一解，$x = A^{-1}\boldsymbol{b}$

必要性证明： 两边同时乘以 A^{-1}

$$\begin{aligned} A^{-1}Ax &= A^{-1}b \\ x &= A^{-1}b \end{aligned}$$

充分性证明：

$$\begin{aligned} &\text{若方程组}Ax = b,\text{有解}x = Cb,\text{求证:}AC = CA = E \\ &\text{由于 b 是任意的,不妨取}\boldsymbol{b} = \boldsymbol{e_1}, \boldsymbol{e_2}, .., \boldsymbol{e_n}, Ax_i = e_i, \\ &\text{令}C = \begin{pmatrix} x_1 & x_2 & ... & x_n \end{pmatrix}, \text{根据矩阵乘法}, AC = E, \text{则}CA = E \end{aligned}$$

QED

把上面的综合起来，就是可逆矩阵定理 (IMT)

Theory 3.37 *IMT*

以下所有说法等价

1.A 是可逆矩阵

2.$A \cong E_n$

3.A 有 n 个主元位置

4.$Ax = 0$ 只有零解

5.A 的各列线性无关

6. 线性变换 $T(x) = Ax$ 为满射

7. 对 R^n 中任意 b, 方程 $Ax = b$ 至少一个解

8.A 的各列生成 R^n

9.$T(x) = Ax$ 为 $R^n \to R^n$

10. 存在 $n \times n$ 矩阵 C 使得 $CA = E$

11. 存在 $n \times n$ 矩阵 D 使得 $AD = E$

12.A^T 可逆

13.A 的列构成 R^n 的一个基

14.$ColA = R^n$

15.$rankA = n$

16.$dimNulA = 0$

17.$NulA = \{0\}$

18.A 没有 0 特征值

19.A 可以写成一系列初等矩阵的乘积

20. |A|$\neq$ 0

21. AA^T 正定

22. A^TA 正定

这个定理贯穿了整个线性代数，我们最后证明

4 向量

前面基本介绍了线性方程组，现在我们来研究其中的一部分，向量

4.1 向量的定义及运算

Definition 4.1 n 个有序的数 $a_1, a_2, ...a_n$ 所组成的数组 $(a_1, a_2, ..., a_n)$ 称为 n 维向量

向量可以横着写，也可以竖着写

Definition 4.2 列向量

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

Definition 4.3 行向量

$$\alpha = (a_1, a_2, ..., a_n)$$

接着就可以定义向量的运算了

Definition 4.4 加法 设 $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ 和 $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ 是两个 n-维向量。向量加法 $\mathbf{w} = \mathbf{u} + \mathbf{v}$ 的结果也是一个 n-维向量 $\mathbf{w} = (w_1, w_2, \ldots, w_n)$，其元素为：

$$w_i = u_i + v_i, \quad \forall i \in \{1, 2, \ldots, n\}$$

Definition 4.5 减法 设 $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ 和 $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ 是两个 n-维向量。向量减法 $\mathbf{w} = \mathbf{u} - \mathbf{v}$ 的结果也是一个 n-维向量 $\mathbf{w} = (w_1, w_2, \ldots, w_n)$，其元素为：

$$w_i = u_i - v_i, \quad \forall i \in \{1, 2, \ldots, n\}$$

Definition 4.6 数乘 设 k 是一个标量，$\mathbf{u} = (u_1, u_2, \ldots, u_n)$ 是一个 n-维向量。数乘 $\mathbf{v} = k\mathbf{u}$ 的结果也是一个 n-维向量 $\mathbf{v} = (v_1, v_2, \ldots, v_n)$，其元素为：

$$v_i = ku_i, \quad \forall i \in \{1, 2, \ldots, n\}$$

Definition 4.7 (线性组合) 设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 是向量空间 V 中的一组向量，$x_1, x_2, \ldots, x_n$ 是一组标量（即实数或复数）。这些向量与标量的线性组合是指形式为：

$$\mathbf{b} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n$$

的向量。换句话说，线性组合是通过将每个向量 $\boldsymbol{\alpha}_i$ 乘以一个标量 x_i，然后将结果相加得到的新向量 $\mathbf{b}$。此时我们称 $\mathbf{b}$ 可以由这些向量线性表示。

4.2 向量方程、矩阵乘法的列行展开

使用上面的概念，我们就可以以一种新的角度来看到线性方程组了

注意到

$$\mathbf{b} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n$$

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

这说明线性方程组可以看成是列向量的线性组合，可以理解成向量方程

这说明矩阵可以看成列向量组

$A = \begin{pmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \ldots & \boldsymbol{\alpha}_n \end{pmatrix}$

$A\boldsymbol{x} = \begin{pmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \ldots & \boldsymbol{\alpha}_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n$

这可以称为是矩阵乘法的列行展开

4.3 线性变换

你以为我要讲行列式了吗？不不不，还早。前面说了矩阵乘法就是一种替换，把一个向量变成另一个向量，把一个矩阵变成另一个矩阵

所以矩阵乘法可以看成是一个函数

比如说我们有一个向量 $a = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

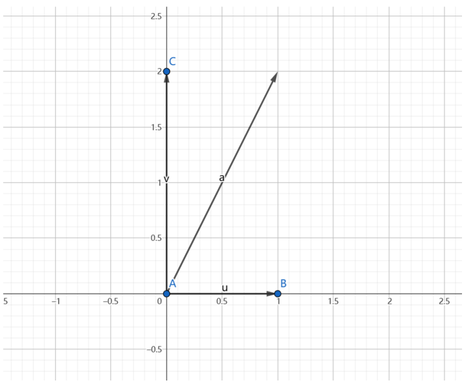

图 1: 一开始的向量

在平面直角坐标系下可以写成线性组合 $a = u + 2v$

而 u,v 是什么呢？$u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

这其实可以看成是向量方程

如果把 u 和 v 放在一起的话，就能组成一个矩阵即 $A = \begin{pmatrix} u & v \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 那么就可以写成 $A \begin{pmatrix} 1 \\ 2 \end{pmatrix} = a$ 所以这个矩阵乘法，其实就是把 (1,2) 用 u，v 为基底表示出来，那么，如果换一个矩阵呢？

$A = \begin{pmatrix} u' & v' \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

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

这说明经过 A 作用后，a 变成了 2a

再比如 $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$

$A \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \end{pmatrix}$

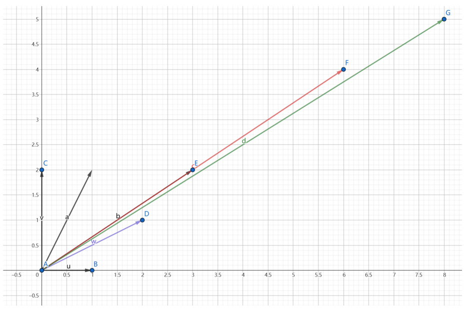

图 2: 之后的向量

我们发现矩阵向量乘法其实就是把矩阵的列向量以向量的分量线性组合。

线性组合完之后，得到一个新的向量，我们把这种变换称为线性变换

设 V 和 W 是两个定义在相同域 $\mathbb{F}$ 上的向量空间。一个函数 $T : V \to W$ 称为从 V 到 W 的 **线性变换**，如果它满足以下两个条件：

1. **加法不变性（Additivity）**：对于所有 $\mathbf{u}, \mathbf{v} \in V$，有

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$$

2. **齐次性（Homogeneity of degree 1）**：对于所有 $\mathbf{v} \in V$ 和所有标量 $c \in \mathbb{F}$，有

$$T(c\mathbf{v}) = cT(\mathbf{v})$$

这两个条件可以合并为一个等价条件：
线性组合不变性：对于所有 $\mathbf{u}, \mathbf{v} \in V$ 和所有标量 $a, b \in \mathbb{F}$，有

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$$

示例
考虑从 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的一个线性变换 T，其定义为：

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 2x + y \\ x - y \end{pmatrix}$$

我们可以验证 T 是否满足线性变换的条件：

1. **加法不变性**：设 $\mathbf{u} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ 和 $\mathbf{v} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$，则

$$T(\mathbf{u} + \mathbf{v}) = T\left(\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}\right) = \begin{pmatrix} 2(x_1 + x_2) + (y_1 + y_2) \\ (x_1 + x_2) - (y_1 + y_2) \end{pmatrix}$$

同时，

$$T(\mathbf{u}) + T(\mathbf{v}) = \begin{pmatrix} 2x_1 + y_1 \\ x_1 - y_1 \end{pmatrix} + \begin{pmatrix} 2x_2 + y_2 \\ x_2 - y_2 \end{pmatrix} = \begin{pmatrix} 2(x_1 + x_2) + (y_1 + y_2) \\ (x_1 + x_2) - (y_1 + y_2) \end{pmatrix}$$

因此，$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$。

2. **齐次性**：设 $c \in \mathbb{R}$ 和 $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$，则

$$T(c\mathbf{v}) = T\left(\begin{pmatrix} cx \\ cy \end{pmatrix}\right) = \begin{pmatrix} 2(cx) + (cy) \\ (cx) - (cy) \end{pmatrix} = \begin{pmatrix} c(2x + y) \\ c(x - y) \end{pmatrix} = cT\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)$$

因此，这个例子中的 T 满足线性变换的条件。

4.4 线性关系

根据前面的向量方程，我们考虑齐次线性方程，也就是 $\mathbf{b} = \mathbf{0}$, 那么

$$x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n = \mathbf{0}$$

我们知道齐次线性方程组必有平凡解，所以 $x_1 = x_2 = ... = x_n = 0$ 肯定是一组解，不过这种解太过平凡，有没有其他的解呢？显然齐次线性方程组是可能有其他解的，如果有其他的解，那么就存在至少一个不为 0 的 x_i, 假设是 x_1, 那么移项得到

$$x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n = x_1\boldsymbol{\alpha}_1$$

由于 $x_1 \neq 0$, 则

$$\frac{x_2}{x_1}\boldsymbol{\alpha}_2 + \cdots + \frac{x_n}{x_1}\boldsymbol{\alpha}_n = \boldsymbol{\alpha}_1$$

这说明 $\boldsymbol{\alpha}_1$ 可由其他向量线性表示，我们把这种关系，称为线性相关。

Theory 4.8 齐次线性方程组只有平凡解 $\Leftrightarrow$ 系数矩阵的列向量组线性无关

Definition 4.9 线性相关，无关

存在一组不全为 *0* 的数 $k_1, k_2, ...k_n$ 使得 $k_1\alpha_1 + k_2\alpha_2 + ... + k_n\alpha_n = 0$ 那么我们称这 n 个向量线性相关，否则线性无关

那如果是非齐次线性方程组呢，如果向量方程组有解，那么 **b** 是 A 列向量的线性组合，也就是说可以由 A 的列向量线性表示，我们将 b 移项过来，其实就是

$$x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n - \mathbf{b} = \mathbf{0}$$

但是，求解这个方程也未必要等于 **b**, 可以等于它的数乘，比如

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_n\boldsymbol{\alpha}_n = k_{n+1}\mathbf{b}$$

即

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_n\boldsymbol{\alpha}_n - k_{n+1}\mathbf{b} = \mathbf{0}$$

总而言之，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\ \boldsymbol{b}$ 线性相关

Theory 4.10 如果 $\alpha_1, \alpha_2, ..\alpha_n$ 线性相关，那么至少有一个向量可以被其他向量线性表示

证明：根据前面分析马上得到，复制一下，那么就存在至少一个不为 *0* 的 x_i, 假设是 x_1, 那么移项得到

$$x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n = x_1\boldsymbol{\alpha}_1$$

由于 $x_1 \neq 0$, 则

$$\frac{x_2}{x_1}\boldsymbol{\alpha}_2 + \cdots + \frac{x_n}{x_1}\boldsymbol{\alpha}_n = \boldsymbol{\alpha}_1$$

这说明 $\boldsymbol{\alpha}_1$ 可由其他向量线性表示

Theory 4.11 如果 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..\boldsymbol{\alpha}_n$ 线性无关，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..\boldsymbol{\alpha}_n, \beta$ 线性相关，那么 β 可由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..\boldsymbol{\alpha}_n$ 线性表示，且表法唯一

证明：$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..\boldsymbol{\alpha}_n$ 线性无关，说明导出组只有平凡解，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..\boldsymbol{\alpha}_n, \beta$ 线性相关，说明非齐次线性方程组有解

设

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_n\boldsymbol{\alpha}_n + k_{n+1}\boldsymbol{\beta} = \mathbf{0}$$

若 $k_{n+1}=0$, 则至少存在一个 $k_i(i=1,2...n)\neq 0$, 则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,..\boldsymbol{\alpha}_n$ 线性相关，矛盾则 $k_{n+1}\neq 0$, 移项得到

$$-\frac{k_1}{k_{n+1}}\boldsymbol{\alpha}_1-\frac{k_2}{k_{n+1}}\boldsymbol{\alpha}_2-\cdots+\frac{k_n}{k_{n+1}}\boldsymbol{\alpha}_n=\boldsymbol{\beta}$$

因此可以线性表出，现在证明唯一性

假设有两种表出方式

$$\boldsymbol{\beta}=l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_n\boldsymbol{\alpha}_n$$
$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n$$

则相减得

$$(l_1-k_1)\boldsymbol{\alpha}_1+(l_2-k_2)\boldsymbol{\alpha}_2+\cdots+(l_n-k_n)\boldsymbol{\alpha}_n$$

由于线性相关，那么 $k_1=l_1,k_2=l_2...$

因此表法唯一

Theory 4.12 如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,..\boldsymbol{\alpha}_n$ 线性相关, 那么 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,..\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_{n+1},\cdots\boldsymbol{\alpha}_s$ 也线性相关

Corollary 4.13 如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,..\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_{n+1},\cdots\boldsymbol{\alpha}_s$ 线性无关, 那么 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,..\boldsymbol{\alpha}_n$ 也线性无关

从几何上很好理解，如果前面 n 个向量线性相关，说明有一个向量可以由其他 *n-1* 个线性表示，那么，根本无需其他向量，新增的向量可以全部系数设置为 *0*,

如果任意一个向量都不能被其他 *s-1* 个向量都不能线性表示，那么更少的向量就更不行了

从物理上来说，如果 *n-1* 个合力已经足够抵消外力，那么再加上其他分力，也足够，如果 *s-1* 个都不够，更遑论更少的力

证明: 假设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$ 线性相关。那么存在一组不全为零的标量 $c_1,c_2,\ldots,c_n$ 使得:

$$c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_n\boldsymbol{\alpha}_n=\mathbf{0}.$$

对于扩展后的向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_{n+1},\ldots,\boldsymbol{\alpha}_s$，我们可以简单地将额外向量的系数设置为零，即 $c_{n+1}=c_{n+2}=\ldots=c_s=0$。于是有:

$$c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\cdots+c_n\boldsymbol{\alpha}_n+0\cdot\boldsymbol{\alpha}_{n+1}+\cdots+0\cdot\boldsymbol{\alpha}_s=\mathbf{0},$$

这表明扩展后的向量组也是线性相关的，因为我们找到了一组非全零的系数，使得它们的线性组合等于零向量。■

推论证明： 假设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_{n+1}, \ldots, \boldsymbol{\alpha}_s$ 线性无关。这意味着只有当所有系数都为零时，才有：

$$c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_n\boldsymbol{\alpha}_n + c_{n+1}\boldsymbol{\alpha}_{n+1} + \cdots + c_s\boldsymbol{\alpha}_s = \mathbf{0}.$$

如果我们只考虑原始的向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 并且设 $c_{n+1} = c_{n+2} = \ldots = c_s = 0$，则上述条件简化为：

$$c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_n\boldsymbol{\alpha}_n = \mathbf{0}.$$

由于整个向量组是线性无关的，因此必须有 $c_1 = c_2 = \ldots = c_n = 0$，不然整个向量组就线性相关了，所以，原始的向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 也是线性无关的。■

不过，有没有什么很快判断线性关系的方法？暂时先给出一个

Theory 4.14 m 个 n 维向量

$$\boldsymbol{\alpha_i} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$$

如果 $m>n$, 则一定线性相关

证明： 将向量摆成一个矩阵,$A = \begin{pmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \ldots & \boldsymbol{\alpha}_m \end{pmatrix}$

经过 *Gauss* 消元法后, 每行有一个主元，至多有 n 个主元，而一共有 m 个未知数，主元数不够，因此肯定有自由变量，所以齐次线性方程组有非零解，因此线性相关

4.5 向量空间

前面我们定义了向量的线性关系和线性运算，那么，我们就可以引出线性代数的研究对象了————向量空间，我们终于有机会脱离线性方程组了，我们知道，几个向量通过线性组合就可以获得无数个向量，在几何意义来说，构成了一个空间，我们说这是这几个向量张成的空间

Definition 4.15 向量空间

设 V 是非空的向量集合，如果对于任意 $\boldsymbol{u}, \boldsymbol{v} \in V$ 及任意 $k \in R$ 都有 $u + v \in V, ku \in V$，则称 V 为向量空间

所以向量空间就是向量组合出的向量的集合

如果只研究向量空间的一部分，也就是向量空间的子集，就需要定义子空间

Definition 4.16 子空间 设有向量空间 W, V，若 $W \subseteq V$，则称 W 为 V 的子空间，若 $W \subset V$ 则称为真子空间

如果我们只关心部分向量，那么就得到了这部分向量张成的子空间

Definition 4.17 张成子空间

$\boldsymbol{v_1}, \boldsymbol{v_2}..., \boldsymbol{v_n}$ 的线性组合组成的子空间，记作 $span\{\boldsymbol{v_1}, \boldsymbol{v_2}..., \boldsymbol{v_n}\}$

我们知道矩阵其实就是一组列向量

那么矩阵中也有向量空间

Definition 4.18 列空间

矩阵列向量张成的空间, 记作 $\mathrm{Col}A$

而矩阵也可以看成行向量组，于是就构成行空间

Definition 4.19 行空间

矩阵行向量张成的空间, 记作 $\mathrm{Row}A$

我们知道，矩阵还对应着齐次线性方程组，那么齐次线性方程组的解其实也构成向量空间

Definition 4.20 零空间 *(*解空间*)*

$Ax=0$ 的 x 张成的空间, 记作 $\mathrm{Nul}A$

4.6 向量组的秩

有了前面的基础，我们可以谈论向量组的秩了

还是那个道理，如果一个向量可以被其他向量表示，那么这个向量就没有什么代表性，也就是说没有什么“缺一不可”的价值，没有核心竞争力，因此即使拿掉也无妨，时间照样流逝，空间继续延展

那么值得研究的就是，向量组中那些老不死的，不可取代的，它们决定了这个向量组的规模

那么怎么筛选这些老不死的呢？很简单，一个一个来

从第一个向量开始分析，如果可以被其他向量线性表示，那么踢出去，一个一个分析，那么留下来的，自然都是重量级，我们把这些重量级叫做，极大线性无关组

踢出去的也可以组成向量组，还有剩下的也组成一个向量组，我们先介绍一下向量组的线性关系

Definition 4.21 *向量组的线性表示 如果我们有一个向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$，并且对于每一个 $\boldsymbol{\beta}_i$（$i = 1, 2, \ldots, n$），都存在一组标量 $c_{i1}, c_{i2}, \ldots, c_{im}$ 使得*

$$\boldsymbol{\beta}_i = c_{i1}\boldsymbol{\alpha}_1 + c_{i2}\boldsymbol{\alpha}_2 + \cdots + c_{im}\boldsymbol{\alpha}_m,$$

那么我们说向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$ 可以由向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$ **线性表示**。

Definition 4.22 (极大线性无关组) *设向量组 $A : \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 如果从 A 中能选出 r 个向量 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r\}$，满足以下两个条件：*

1. *向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r\}$ 线性无关；*
2. *向量组 A 中任意 $r+1$ 个向量（若有的话）都线性相关，(或者说向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r\}$ 可由 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r\}$ 线性表示)*

*则称向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r\}$ 是向量组 A 的一个***极大线性无关组***（简称为***极大无关组**）。

为什么说极大，而不是最大呢？这倒是有些说法了，“极”其实有一种局部的意思，极大线性无关组是这个向量组中最大的无关组，但是未必是整个向量空间的最大无关组（这个叫做基）。当然，极大线性无关组在英语中就是 maximal，翻译成极大，我猜测是上面这个原因。

那么问题来了，极大线性无关组是唯一的吗？这显然不是，我们前面的算法其实和顺序有关系，所以极大线性无关组应该是不唯一的，但是，位置应该是一样多的

极大线性无关组的个数是否相同？

思考，比如说两个极大线性无关组，各有 s，t 个向量如果 $s < t$, 可能吗？因为他们都是极大线性无关组，所以肯定可以彼此线性表示，少的可以表示多

的，这可能吗？如果说线性相关，多的里面有一些水货，倒是还有可能，但是这些全都是线性无关，个个都很强，那么以少胜多，或未易量

也就是说，如果想以少胜多，那么多的里面应该有水货

于是我们来证明这个定理

Theory 4.23 若向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r$ 可由，$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_s$ 线性表示且 $r > s$，则 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r$ 线性相关

怎么证明呢？那么这就需要理解线性表示了上面说明左边的每一个向量，都可以由后面的向量线性表示，写出方程

$$\begin{aligned}
&\boldsymbol{\alpha}_1 = k_{11}\boldsymbol{\beta}_1 + k_{12}\boldsymbol{\beta}_2 + ... + k_{1s}\boldsymbol{\beta}_s \\
&\boldsymbol{\alpha}_2 = k_{21}\boldsymbol{\beta}_1 + k_{22}\boldsymbol{\beta}_2 + ... + k_{2s}\boldsymbol{\beta}_s \\
&\vdots \\
&\boldsymbol{\alpha}_r = k_{r1}\boldsymbol{\beta}_1 + k_{r2}\boldsymbol{\beta}_2 + ... + k_{rs}\boldsymbol{\beta}_s
\end{aligned}$$

$$A = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_r \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1s} \\ k_{21} & k_{22} & \cdots & k_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ k_{r1} & k_{r2} & \cdots & k_{rs} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_s \end{pmatrix} = KB$$

$$\text{rank}A = \text{rank}KB \leq s < r$$

注：用了秩不等式 rankKB$\leq$ rankK $\leq$ s

如果不用秩不等式呢？那就得从方程组角度了，并且需要使用列摆法根据方程设 $B = \begin{pmatrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & ...\boldsymbol{\beta}_s \end{pmatrix}$

$\boldsymbol{\alpha}_i = k_{i1}\boldsymbol{\beta}_1 + k_{i2}\boldsymbol{\beta}_2 + ... + k_{is}\boldsymbol{\beta}_s = Bc_i$

我们令 $C = \begin{pmatrix} c_1 & c_2 & ... & c_r \end{pmatrix}$

由于 C 只有 s 维，但是有 r 个向量，所以 C 一定线性相关，而 $A = BC$ 存在非零向量 $\boldsymbol{x}$,

$A\boldsymbol{x} = BC\boldsymbol{x} = 0$ 则 A 的列向量线性相关

Question 4.24 秩不等式

证明：rankA+rankB-n$\leq$rank$AB \leq$ min{rankA,rankB}

那么，根据逆否命题马上有

Theory 4.25 若向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r$ 可由，$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_s$ 线性表示且，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_r$ 线性无关，则 $s \leq t$

由于两个极大线性无关组可以相互线性表示，因此两个极大线性无关组的向量个数相等

所以极大线性无关组是一个不变量，我们就把极大线性无关组的向量个数称为向量组的秩

Definition 4.26 向量组的秩、矩阵的秩

极大线性无关组中包含的向量个数定义为向量组的秩，
记作 $rank(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n)$

现在我们可以定义矩阵的秩了

矩阵可以看成列向量组，我们定义

Definition 4.27 矩阵的列秩

矩阵的列秩为列空间的秩，即 rank(ColA)

但是矩阵也可以看成行向量组啊

Definition 4.28 矩阵的行秩

矩阵的行秩为行空间的秩，即 rank(RowA)

那么矩阵的列秩和矩阵的行秩是否相等呢？

这个问题确实很困难，矩阵的行空间好像没什么特殊的意义，

不过我们有一点可以明确，
因矩阵的行空间转置后就是列空间，所以 rank(RowA) = rank(ColA^T)

这样我们就可以把行空间转化成列空间来研究

只需要证明 A 和 A 转置的列空间秩相等就可以

了吗？

实际上引入转置更加困难，我们其实并不怎么了解转置

回想一下，我们真的不了解行空间吗？虽然列空间可以形象的描述方程，但是，实际上我们解方程的时候，都是进行行变换，也就是说，我们其实更了解行空间

解方程的时候，我们可以通过 Gauss-Jordan 消元法，获得一个行最简形矩阵，它可以通过交换方程，得到一个一边是包含主元的单位矩阵，一边是其他元素的分块矩阵

也就是 $A \cong \begin{pmatrix} E & B \\ O & O \end{pmatrix}$

如果把 A 看成是一个行向量组，那么意味着什么？

$$A = \begin{pmatrix} \boldsymbol{v_1} \\ \boldsymbol{v_2} \\ \vdots \\ \boldsymbol{v_n} \end{pmatrix} \to \begin{pmatrix} \boldsymbol{e_1} & \boldsymbol{b_1} \\ \boldsymbol{e_2} & \boldsymbol{b_2} \\ \vdots & \vdots \\ \boldsymbol{e_k} & \boldsymbol{b_k} \\ O & O \end{pmatrix}$$

因为 $\boldsymbol{e_1}, \boldsymbol{e_2}$... 这些肯定线性无关连带着右边也线性无关，k 是主元的个数，下面的向量之所以消为 0，正是因为能被上面 k 个向量线性表示，所以，这 k

个向量就是行向量组的极大线性无关组，因此，矩阵的行秩就等于主元数，把 k 改成 r 吧

其实，这 r 个列向量也是线性无关的，为什么？

还是把这个矩阵写出来更直观

$$I_A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1,n-r} \\ 0 & 1 & 0 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2,n-r} \\ 0 & 0 & 1 & \cdots & 0 & b_{31} & b_{32} & \cdots & b_{3,n-r} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & b_{r1} & b_{r2} & \cdots & b_{r,n-r} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} E_r & B_{r\times(n-r)} \\ O & O \end{pmatrix}.$$

在前 r 个向量里面，第一个向量有第一个分量，其他向量没有，以此类推，前 r 个向量肯定线性无关，而他们又可以表示出后面的任意一列向量，比如第 r+1 列就是 $b_{11}e_1 + b_{22}e_2...$ 以此类推

所以主元个数也是矩阵的列秩

所以矩阵的列秩等于矩阵的行秩，**由此我们发现主元数是个非常重要的量，我们把这个量称为矩阵的秩**

Theory 4.29 秩

矩阵的列秩等于矩阵的行秩

Definition 4.30 矩阵的秩

矩阵的秩就是矩阵列空间和行空间的秩，等于矩阵的主元数，记作 $rankA$

为什么不能这样定义？**矩阵的秩（不是定义）**

经过高斯消元法后，主元的个数称为矩阵的秩, 记作 $\mathrm{rank}A$

这不是定义，只是矩阵的秩的表现，为什么这不能是定义呢？（值得思考）从方程组角度来说，这可以说是系数矩阵秩的定义，但如果我们把矩阵视为一种研究的客体，显然，经过高斯消元法后，矩阵不再是原来的矩阵，变换后主元的个数，应当是属于新的矩阵的特性，而不是原来矩阵自身的性质，因此把这个作为矩阵的秩的定义不合适？

还有一个国内教材的定义，为什么合理？

Question 4.31 *为什么用存在不为 0 的 k 阶子式的最大阶数来定义矩阵的秩可以？（后面谈）*

我们先谈谈行列式吧

5 行列式

终于到了看似无用，但实际上很有用的行列式了，作为国内教材的集大成者，一开始放上一个行列式，让人一头雾水，不过，在我编写本作之时，确实发现了其中的用意，如果没有行列式，会增添很多理解的困难，比如上面所讲的矩阵的逆，我们后面会重新理解

不过，虽然行列式前置有合理性，不过我仍然认为不应该上来就什么逆序数，什么代数定义

我们还是先从线性方程组开始，那个梦开始的地方

首先，你希望行列式是什么？Cramer 法则？计算工具？线性函数？伸缩尺度？这都可以是

Definition 5.1 逆序
在一个排列中，如果一对数的前后位置与大小顺序相反，即前面的数大于后面的数，则他们构成一对逆序

Definition 5.2 逆序数
排列中逆序的总数, 记作 τ

Definition 5.3 行列式的几何定义
线性变换对体积的拉伸程度
二维：$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ 表示 $span\{(a,c)^T, (b,d)^T\}$ 张成的平行四边形面积

Definition 5.4 行列式的代数定义
由 n^2 个数组成的取自所有不同行不同列的 n 个元素乘积的代数和
方阵 $A_{n\times n} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$
$|A| = \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} ... a_{nj_n}$

Corollary 5.5 对角矩阵的行列式 $|A| = \begin{vmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_n \end{vmatrix} = \lambda_1 \lambda_2 .... \lambda_n$

Prove 5.6 由于除了对角线上的元素都为 0, 那么对于任意 a_{ij}，当且仅当 $i \neq j$ 时，$a_{ij} \neq 0$ 则对于 $|A| = \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} ... a_{nj_n}$

当且仅当 $j_1=1,j_2=2...j_i=i$ 时，乘积不为 *0*
则 $|A|=(-1)^{\tau(1,2,3...,n)}\lambda_1\lambda_2....\lambda_n=(-1)^0\lambda_1\lambda_2....\lambda_n=\lambda_1\lambda_2....\lambda_n$

Definition 5.7 余子式与代数余子式 选定 a_{ij} 划掉第 i 行和第 j 行，剩下的元素组成的行列式称为元素 a_{ij} 的余子式，记作 M_{ij}
记 $A_{ij}=(-1)^{i+j}M_{ij}$ 称为代数余子式

Definition 5.8 伴随矩阵 将各元素替换为代数余子式，再转置，得到伴随矩阵，记作 A^*
$$A^*=\begin{pmatrix}A_{11}&A_{21}&\cdots&A_{n1}\\A_{12}&A_{22}&\cdots&A_{n2}\\\vdots&\vdots&\ddots&\vdots\\A_{1n}&A_{2n}&\cdots&A_{nn}\end{pmatrix}$$
Theory 5.9 $AA^*=A^*A=|A|E$

Property 5.10 行列式的性质 *1*
行列式的行和列地位平等
$|A|=|A^T|$

Property 5.11 行列式的性质 *2*
互换行列式的两行，行列式变号

$$\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{i1}&a_{i2}&\cdots&a_{in}\\\vdots&\vdots&\ddots&\vdots\\a_{j1}&a_{j2}&\cdots&a_{jn}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}=-\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{j1}&a_{j2}&\cdots&a_{jn}\\\vdots&\vdots&\ddots&\vdots\\a_{i1}&a_{i2}&\cdots&a_{in}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}$$

Corollary 5.12 若行列式有两行元素相等，则此行列式值为 *0*

Corollary 5.13 如果代数余子式和元素不对应，所展开行列式为 *0*

Theory 5.14 */AB/=/A//B/*

Definition 5.15 奇异方阵
如果 $|A|\neq 0$ 那么称 A 是非奇异的，否则是奇异的
这种翻译并不好，应翻译成不可退化方阵，可退化方阵
非奇异方阵满秩，满秩方阵可逆，这三个概念完全等价