2023线性代数1讲义

目录

1 第一章：线性方程组与矩阵

1.1. 线性方程组入门，Introduction to systems of linear equations

Definition 1.1. 令n为任意自然数，$a_1, \ldots, a_n$，b为常数。我们称n元**一次**方程

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

为一个(含有n个未知数的)**线性**方程。

也就是说，在本课程中，**线性**与**一次**是同一个意思。比如方程$2x_1^2 + \sin x_2 - 5x_3 = 0$不是一个线性方程，因为$x_1$的次数为2，并且$\sin x_2$也不能写成$x_2$的一次方形式。

Definition 1.2. 令m与n为任意自然数， $a_{11}, \ldots, a_{1n}$，$a_{21}, \ldots, a_{2n}$，$\ldots a_{m1}, \ldots, a_{mn}$，

$b_1, \ldots, b_m$为常数，我们称以下m个n元一次方程组成的方程组

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

为一个**线性方程组**。特别地，若$b_1 = \ldots = b_m = 0$，则称该线性方程组为**齐次**线性方程组。显然，a_{ij}指代第i个方程里第j个未知数x_j的系数。这里i的取值范围为$1, 2, \ldots, m$, j的取值范围为$1, 2, \ldots, n$。

给定任何一个(由m个n元一次方程组成的)线性方程组

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

我们将由未知数系数$a_{ij}, i = 1, \ldots, m, j = 1, \ldots, n$组成的矩形点阵

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

为该线性方程组的**系数矩阵**(coefficient matrix)。显然a_{ij}位于该矩阵第i行与第j列相交的位置上。如果我们把方程组最右侧的常数$b_1, \ldots, b_m$作为**列**添加到系数矩阵的最右边，那么所得到的矩阵

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{bmatrix}$$

被称为该线性方程组的**增广矩阵**(augmented matrix)。
反之，若给定一个由m行和n列组成的矩阵A

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

(这里a_{ij}指代该矩阵第i行与第j列相交的位置上出现的常数)以及一个由m个坐标组成的向量$\boldsymbol{b} = (b_1, \ldots, b_m)$, 那么$A$与$b$可以确定一个(由$m$个$n$元一次方程组成的)线性方程组

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

显然，该线性方程组对应的系数矩阵为A，对应的增广矩阵为$(A, \boldsymbol{b})$。注意我们经常用记号$(A, \boldsymbol{b})$指代以下矩阵

$$
\begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\
a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
a_{m1} & a_{m2} & \ldots & a_{mn} & b_m
\end{bmatrix}
$$

Remark 1.3. *1.* *若一个矩阵由m行和n列所构成，我们称其为一个$m \times n$-矩阵。因此一个(由m个n元一次方程组成的)线性方程组的系数矩阵为一个$m \times n$-矩阵，它的增广矩阵则为一个$m \times (n+1)$-矩阵。*

2. *一个由m个坐标组成的向量(简称为m-向量)$\boldsymbol{b} = (b_1, \ldots, b_m)$可以被视为一个行向量，即一个$1 \times m$-矩阵；也可以被视为一个列向量，即一个$m \times 1$-矩阵, 且在此时记为*

$$
\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}
$$

比如在记号$(A, \boldsymbol{b})$里我们就把$\boldsymbol{b}$看作一个列向量。究竟是把一个向量视为行向量还是列向量取决于具体的应用环境，我们在后面的课程学习中会逐渐了解。

3. *诚然每一个(由m个n元一次方程组成的)线性方程组都可以被它的增广矩阵唯一表示，但是矩阵作为线性代数里的一个独立概念并**不是必须**要与线性方程组联系起来(尽管目前我们使用矩阵的目的是为了求解线性方程组)。更多关于矩阵本身的知识与应用将在后续课程中逐渐展示。*

例子：

1. 线性方程组

$$\begin{aligned} &x-y=1 \\ &2x+y=6 \end{aligned}$$

的系数矩阵为

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

增广矩阵则为

$$\begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & 6 \end{bmatrix}.$$

2. 线性方程组

$$\begin{aligned} &x+y=4 \\ &3x+3y=6 \end{aligned}$$

的系数矩阵为

$$\begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix},$$

增广矩阵则为

$$\begin{bmatrix} 1 & 1 & 4 \\ 3 & 3 & 6 \end{bmatrix}.$$

3. 矩阵$A=$

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

与向量$\boldsymbol{b}=(5,10,15)$所对应的线性方程组为

$$\begin{aligned} &x-y+2z=5 \\ &2x-2y+4z=10 \\ &3x-3y+6z=15. \end{aligned}$$

即，该方程组的系数矩阵为A, 增广矩阵为$(A,\boldsymbol{b})$。

现在我们来解上面的第一个方程组

$$\begin{aligned} &x-y=1 \\ &2x+y=6. \end{aligned}$$

注意它的增广矩阵为$B_1 =$

$$\begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & 6 \end{bmatrix}.$$

求解的思路当然是**消元**。具体步骤如下：

1. 首先我们将**第一个方程乘以-2, 然后加到第二个方程上去**来消掉第二个方程里的未知数x。我们所得到的新方程组为

$$\begin{aligned} x - y &= 1 \\ 3y &= 4. \end{aligned}$$

   注意这个方程组的增广矩阵变成了$B_2 =$

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

   显然，B_2是通过**将矩阵B_1的第一行乘以-2后加到第二行，同时保留第一行不变**所得到的。

2. 对于方程组

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

   我们将第二个方程等号两边都乘以$1/3$, 得到另一个方程组

$$\begin{aligned} x - y &= 1 \\ y &= \frac{4}{3} \end{aligned}$$

   对应的增广矩阵为$B_3 =$

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & \frac{4}{3} \end{bmatrix}.$$

   显然，B_3是通过**将矩阵B_2的第二行乘以$1/3$，同时保留第一行不变**所得到的。

3. 现在我们把方程组

$$\begin{aligned} x - y &= 1 \\ y &= \frac{4}{3} \end{aligned}$$

   的**第二个方程加到第一个方程上面**，得到新方程组

$$\begin{aligned} x &= \frac{7}{3} \\ y &= \frac{4}{3} \end{aligned}$$

显然这个最终得到的方程组给出了原方程组的解，其对应的增广矩阵为$B_4 =$

$$\begin{bmatrix} 1 & 0 & \frac{7}{3} \\ 0 & 1 & \frac{4}{3} \end{bmatrix}$$

并且系数矩阵为

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

注意B_4是通过**将矩阵B_3的第二行加到第一行，同时保留第二行不变**所得到的。

由以上观察我们可以得到以下结论：

- 对以上(由2个二元一次方程组成的)线性方程组进行**消元求解**的过程实际上是对该方程组所对应的**增广矩阵**的行进行一系列操作使其最终变为

$$\begin{bmatrix} 1 & 0 & s_1 \\ 0 & 1 & s_2 \end{bmatrix}$$

这种形式的过程，且此时$x = s_1, y = s_2$即为该方程组**唯一的一组解**。
- 在此过程中矩阵**列与列之间不发生任何关系**。
- 在此过程中矩阵的行与行之间所进行的操作包括:
  1. 将某一行乘以某个**非零**的常数，其他行保持不变；
  2. 将某一行乘以某个常数后加到另外一行上去，且除了被加的行之外其他的行保持不变。

基于此，我们引入一下重要概念:

Definition 1.4 (初等行变换，elementary row operations)**.** *假设C为任意$m \times p$-矩阵。以下三种操作被称为(对矩阵C的)初等行变换:*

1. *将C的第i行 ($i = 1, \ldots, m$) 乘以某个**非零**的常数c，其他行保持不变 (可记为: i-th row $\times c$);*
2. *将C的第i行 ($i = 1, \ldots, m$)乘以某个常数c后加到第k行上去 ($k = 1, \ldots, m$且$k \neq i$)，且除了第k行之外其他的行保持不变 (可记为: i-th row $\times c + k$-th row);*
3. *将C的第i行 ($i = 1, \ldots, m$)与第k行 ($k = 1, \ldots, m$且$k \neq i$) 互换 (可记为: i-th row $\leftrightarrow k$-th row)，其他行保持不变。*

考虑线性方程组

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

及其对应的增广矩阵为$B = (A, \boldsymbol{b})$ (A为其系数矩阵)

$$
\begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\
a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
a_{m1} & a_{m2} & \ldots & a_{mn} & b_m
\end{bmatrix}.
$$

那么由之前的讨论我们知道

1. “将第i个方程的等号两边乘以$c \neq 0$” 对应 “将B的第i行乘以某个**非零**的常数c，其他行保持不变”，即第一类初等行变换；
2. “将第i个方程乘以c再加到第k个方程上去”对应 “将B的第i行乘以某个常数c后加到第k行上去 ($k \neq i$)，且除了第k行之外其他的行保持不变”，即第二类初等行变换；
3. “将第i个方程与第k个方程互换”对应 “将A的第i行与第k行 ($k \neq i$) 互换，其他行保持不变”，即第三类初等行变换。
4. 若$m = n$, 即方程个数与未知数个数相等，且通过对增广矩阵$B = (A, \boldsymbol{b})$进行一系列初等行变换操作后得到以下形状的(增广)矩阵$\tilde{B} =$

$$
\begin{bmatrix}
1 & 0 & \ldots & 0 & s_1 \\
0 & 1 & \ldots & 0 & s_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & 1 & s_n
\end{bmatrix},
$$

   那么该方程组具有**唯一的一组解**$x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$。这里我们把$\tilde{B}$记为$\tilde{B} = (I_n, \boldsymbol{s})$, I_n**被称为**$n \times n$**–单位矩阵，它的对角线上所有元素均为**1**，其他不在对角线上的元素均为**0; 列向量$\boldsymbol{s} = (s_1, \ldots, s_n)$是方程的解。

因此，从现在开始，我们可以**把对线性方程组** ($m = n$) **的消元求解过程**转化为如何通过一系列初等行变换将其对应的增广矩阵$B = (A, \boldsymbol{b})$转化为$\tilde{B} = (I_n, \boldsymbol{s})$的过程，此时$\boldsymbol{s}$即为方程组的唯一解。特别地，**我们现在操作的对象由方程组变为矩阵**。

在结束本节之前，我们需要注意到另一个非常重要的问题: **并非所有线性方程组对应的增广矩阵都可以通过初等行变换转为$\tilde{B}=(I_n,\boldsymbol{s})$的形式!** 比如当方程个数$m$与未知数个数$n$不相等时，原增广矩阵$B$无法变成$\tilde{B}=(I_n,\boldsymbol{s})$的形式 (为什么？)；同时，即使$m=n$，一些线性方程组的增广矩阵依然不能通过初等行变换转为$\tilde{B}=(I_n,\boldsymbol{s})$的形式，比如方程组

$$
\begin{aligned}
&x-y+2z=5\\
&2x-2y+4z=10\\
&3x-3y+6z=15
\end{aligned}
$$

对应的增广矩阵$B=$

$$
\begin{bmatrix}
1 & -1 & 2 & 5\\
2 & -2 & 4 & 10\\
3 & -3 & 6 & 15
\end{bmatrix}
$$

通过消元求解所对应的初等行变换后最终所得到的形式是

$$
\begin{bmatrix}
1 & -1 & 2 & 5\\
0 & 0 & 0 & 0\\
0 & 0 & 0 & 0
\end{bmatrix}
$$

并不满足$(I_3,\boldsymbol{s})$这样的形式。注意此时可以看出该方程组具有**无穷多个解**$x=5+r-2s, y=r, z=s$，这里r和s可以为任何数。

作为总结，我们在这一节中的收获为:

- 我们可以**把对线性方程组 ($m=n$) 的消元求解过程**转化为如何通过一系列初等行变换将其对应的增广矩阵$B=(A,\boldsymbol{b})$转化为$\tilde{B}=(I_n,\boldsymbol{s})$的过程。此时向量$\boldsymbol{s}$给出方程的**唯一解**。
- 并非所有方程组的增广矩阵B都可以通过一系列初等行变换转化为$\tilde{B}=(I_n,\boldsymbol{s})$的形式。
- **一个线性方程组的增广矩阵B通过一系列初等行变换所转化成的某种"最终形态"决定了该方程组是否有解，以及解是否唯一**。

在1.2节我们将学习如何将一个线性方程组的增广矩阵B通过一系列初等行变换转化成的一种特殊的"最终形态"，以及如何通过这个"最终形态"确定该方程组是否有解，解是否唯一，以及求出具体解 (在有解的情况下)。

1.2. 高斯消元法，Gaussian Elimination

在上一节我们提到，可以通过一系列初等行变换，将一个线性方程组的增广矩阵转化为一种特殊的"最终形态"，这种"最终形态"能够帮助我们确定该方程组

是否有解，解是否唯一，以及求出具体解(在有解的情况下)。在本节我们将会学习到这种“最终形态”就是所谓的阶梯型 (row echelon form) 或者简化阶梯型 (reduced row echelon form)。首先我们给出定义：

Definition 1.5 (阶梯型与简化阶梯型)**.** *若一个矩阵满足以下三个条件:*

1. *该矩阵中任何不为0的行里所包含的第一个非零的项都是1。我们称这样的项为***主**1 *(leading* 1*);*
2. *该矩阵中任何为0的行一定在不为0的行的下方;*
3. *对于该矩阵中任意两个***相邻***且***不为***0的行，***上面***的行里所包含的***主***1一定位于***下面***的行里所包含的主1的***左侧***;*

*那么该矩阵被称为***阶梯型***。若一个***阶梯型***额外满足以下条件*

- *该矩阵中任何不为0的行里所包含的***主***1是其所在的***列里唯一一个不为***0的项，*

*那么该矩阵被称为***简化阶梯型***。*

例子：

1. 矩阵
$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
不是阶梯型，因为第二行中第一个非零的项是2，不等于1。
2. 矩阵
$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
不是阶梯型，因为不为0的第三行在为0的第二行下面。
3. 矩阵
$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$
不是阶梯型，因为第二行的主1在第三行主1的右边。
4. 矩阵
$$\begin{bmatrix} 1 & 4 & -3 & 7 \\ 0 & 0 & 6 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

是阶梯型，但不是简化阶梯型，因为第二行的主1不是它所在列中唯一的非零项，同样，第三行的主1不是它所在列中唯一的非零项。

5. 矩阵

$$\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

是简化阶梯型。

下面我们介绍如果将任意给定的一个矩阵通过一系列初等行变化将其转化为一个阶梯型或者简化阶梯型。大家可以参照教材第14页到17页的两个例子的具体计算过程。我们这里使用教材第14页上的矩阵作为示例：考虑矩阵$B =$

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}.$$

1. 首先我们知道原矩阵第一行的第一个非零项必然在第一列之后 (实际上通过观察第二列的情况可知在第二列之后)，而原矩阵第二行与第三行的第一个非零项都在第一列里，这意味着目前第一行所包含的主1必然在目前第二行和第三行所包含的主1右边，与阶梯型要求的条件3不符。因此我们需要将第一行与第二行或者与第三行进行交换 (来保证阶梯型的条件3成立)。这里**第一行与第二行交换**，所得到的矩阵为$B_1 =$

$$\begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}.$$

2. 现在我们将**第一行乘以**$1/2$使得第一行第一个非零项为1，从而得到第一行的主1: $B_2 =$

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

3. 现在我们注意到B_2里第三行的第一个非零项2仍在第一列里，这意味着B_2的第三行的主1与第一行的主1都在第一列里，不满足阶梯型的条件3。因此我们必须把第三行第一列上的项消掉，即，**将第一行乘以**-2**加到第三行上**，得到$B_3 =$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}.$$

4. 现在我们注意到B_3第二行的第一个非零项为-2位于第三列上。通过对**第二行乘以**$-1/2$我们得到第二行的主1: $B_4 =$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}.$$

5. 注意到B_4里第三行的第一个非零项为5，它与第二行的主1处于同一列上，不符合阶梯型条件3，因此我们仿照第三步，通过将**第二行乘以**-5**加到第三行**，将该项消掉，得到$B_5 =$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}.$$

6. 现在只需对**第三行乘以**2去得到第三行的主1: $B_6 =$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

显然，根据定义，矩阵B_6是一个阶梯型。如果我们想将其继续转化为简化阶梯型，则需要将第三列 (即第二行的主1所在的列) 和第五列 (即第三行的主1所在的列) 通过初等行变换分别转化为

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

与

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

最终得到的简化阶梯型为$R =$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

在实际操作中，大家可以根据自己的个人喜好，经验，以及矩阵的具体样子来决定如何进行初等行变换将其转化为阶梯型或简化阶梯型。比如可以遵循下面的几个核心步骤：

- 首先浏览矩阵的所有行，找出所有不为0的行，并将所有为0的行全部移动到矩阵的下方，确保所有不为0的行都在所有为0的行的上方。
- 在每个不为0的行里都找出其包含的第一个不为0的项所在的列。假设第i行所包含的第一个不为0的项所在的列是其中**最靠左**的，那么将第i行与第一行进行交换。
- 在交换过后得到的矩阵里，第一行所包含的第一个不为0的项假设为常数a，将第一行乘以$1/a$，得到第一行里的主1。
- 假设现在所得到的矩阵里第一行的主1所在的列为第j列，那么通过初等行变换将该列里所有其他的项 (即首行主1下方的所有项)均消为0。注意这一步的意图是为了确保首行所包含的主1处于最左边。
- 现在可以暂时忘记第一行和第一列，我们观察剩下的$m-1$行和$n-1$列 (假设原矩阵为$m\times n$-矩阵)，然后对其**重复以上步骤**。直到将其化为阶梯型为止。
- 将该阶梯型里所有行包含的主1上方和下方的非零项通过初等行变换全部消为0，即可得到简化阶梯型。

现在我们将原矩阵$B=$

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

视为方程组

$$\begin{aligned} -2x_3+7x_5&=12\\ 2x_1+4x_2-10x_3+6x_4+12x_5&=28\\ 2x_1+4x_2-5x_3+6x_4-5x_5&=-1 \end{aligned}$$

的增广矩阵，那么该方程组的解与其阶梯型$B_6=$

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

所对应的方程组

$$\begin{aligned} x_1+2x_2-5x_3+3x_4+6x_5&=14\\ x_3-\frac{7}{2}x_5&=-6\\ x_5&=2 \end{aligned}$$

具有同样的解；类似的，与其简化阶梯型$R=$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

所对应的方程组

$$
\begin{aligned}
x_1 + 2x_2 + 3x_4 &= 7 \\
x_3 &= 1 \\
x_5 &= 2
\end{aligned}
$$

具有同样的解。显然不论是方程组

$$
\begin{aligned}
x_1 + 2x_2 - 5x_3 + 3x_4 + 6x_5 &= 14 \\
x_3 - \frac{7}{2}x_5 &= -6 \\
x_5 &= 2
\end{aligned}
$$

还是方程组

$$
\begin{aligned}
x_1 + 2x_2 + 3x_4 &= 7 \\
x_3 &= 1 \\
x_5 &= 2
\end{aligned}
$$

都能容易的给出原方程组的解：

$$x_1 = 7 - 3r - 2s, x_2 = s, x_3 = 1, x_4 = r, x_5 = 2$$

这里r, s可以为任意常数。

我们将把一个线性方程组对应的**增广矩阵**通过一系列**初等行变换**转化为阶梯型的过程称之为高斯消元法 (Gauss elimination)，把一个线性方程组对应的**增广矩阵**通过一系列**初等行变换**转化为简化阶梯型的过程称之为高斯-若尔当消元法 (Gauss–Jordan elimination)。

Definition 1.6. *假设B是一个线性方程组的增广矩阵，其对应的阶梯型$\tilde{B}$或者简化阶梯型R里不为0的行里所包含的主1所对应的未知数被称为该方程组的主元 (leading variable)，主元之外的所有未知数被称为该方程组的自由元 (free variable)。*

比如在上面的例子里，由于简化阶梯型为$R =$

$$
\begin{bmatrix}
1 & 2 & 0 & 3 & 0 & 7 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 2
\end{bmatrix},
$$

我们知道x_1, x_3, x_5为方程组的主元， x_2, x_4为方程组的自由元。

由以上的计算容易看出，若方程组有解，那么该方程组的自由元所对应的解可以取任意值 (即自由取值，这是自由元名字里“自由”一词的由来)；而当我们给

自由元赋予特殊的值之后，主元所对应的解则被唯一确定。比如上面的例子中，我们已求出方程组

$$\begin{aligned} -2x_3+7x_5&=12\\ 2x_1+4x_2-10x_3+6x_4+12x_5&=28\\ 2x_1+4x_2-5x_3+6x_4-5x_5&=-1 \end{aligned}$$

的解为

$$x_1=7-3r-2s, x_2=s, x_3=1, x_4=r, x_5=2$$

这里r, s可以为任意常数。若我们选择$r=1$, $s=1$, 那么在对应的解里面，主元x_1必然只能取值为$x_1=7-3-2=2$，此外$x_3=1$和$x_5=2$显然已经被唯一确定。

我们将

$$x_1=7-3r-2s, x_2=s, x_3=1, x_4=r, x_5=2; \quad r,s\text{ 为任意常数}$$

这样的表示称为方程组的**通解 (general solution)**，将

$$x_1=2, x_2=1, x_3=1, x_4=1, x_5=2$$

(选择$r=s=1$后得到的解)这样的写法称为方程组的**特解**。显然**特解是通过对通解里的自由元赋予具体数值后得到**。

通过本节内的计算与观察，我们可以看出来一个线性方程组的增广矩阵经过初等行变换所得到的阶梯型或简化阶梯型的样子能够决定这个方程组是否有解，以及有解时是否有唯一解。总结如下: 设某个(由m个n元一次方程组成的)方程组的增广矩阵$B=(A,\boldsymbol{b})$ (A为系数矩阵)，其对应的简化阶梯型为$R=(U,\boldsymbol{s})$ (这里$\boldsymbol{s}$为矩阵R最右边的列)。

- 假如存在某个$i\in\{1,\ldots,m\}$使得矩阵U的第i行为0，而列向量$\boldsymbol{s}$的第i个坐标$s_i\neq 0$，那么该方程组无解。反之该方程组至少有一个解。
- 如果矩阵$U=I_n$为$n\times n$-单位矩阵，那么该方程组有且只有一组解，这个解就是向量$\boldsymbol{s}$。
- 若方程有解，且$U\neq I_n$，那么该方程组有无穷多组解。

我们将在后续的学习中给出以上结论的严格证明。这里我们注意到以上结论已经回答了第1.1节最后提出的猜测。此外，我们很容易看出

- 假设某个(含有n个未知数的)线性方程组的增广矩阵为B，$\tilde{B}$与R分别为由B转化的阶梯型和简化阶梯型。那么该方程组主元的个数$k=\tilde{B}$或者R里主1的个数 = $\tilde{B}$或者R里非零行的个数；该方程组自由元的个数$=n-k$。

在本节的最后我们讨论一类特殊的线性方程组：齐次线性方程组

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

显然，任何齐次线性方程组都至少有一个解：

$$x_1 = 0, x_2 = 0, \ldots, x_n = 0.$$

这样的解被称为**平凡解 (trivial solution)**。如果齐次线性方程组有区别于平凡解的其他解，那么这些解被称为非平凡解 (nontrivial solutions)。注意：**若一个齐次线性方程组有一个非平凡解，那么它必然有无穷多个非平凡解** (请思考为什么)。

1.3. 矩阵与矩阵运算，Matrices and Matrix Operations

在本节中大家不需要把矩阵与线性方程组联系起来。我们单独研究矩阵这个对象。

首先我们引入一下记号和定义：

- 所有$m \times n$-矩阵的集合记为$M_{m\times n}$。因此我们经常用$A \in M_{m\times n}$指代A是一个$m \times n$-矩阵。
- 若$m = n$, 即矩阵的行数等于列数，那么我们也称$n\times n$-矩阵为n阶**方阵 (square matrix)**。
- 对于一个$n \times n$-方阵$A =$

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix},$$

  我们称向量$(a_{11}, a_{22}, \ldots, a_{nn})$为$A$的**对角线**。注意对角线是只有**方阵**才有的概念。
- 对于$m \times n$-矩阵$A =$

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

  有时候我们也用A_{ij}指代a_{ij}, 用$[a_{ij}]$表示A。

- 1×1-矩阵就是一个常数a。
- 对于一个向量$\boldsymbol{b}=(b_1,b_2,\ldots,b_n)$，我们可以将其视为一个行向量，即$1\times n$-矩阵，此时将其记为
$$\boldsymbol{b}=\begin{bmatrix} b_1 & b_2 & \ldots & b_n \end{bmatrix},$$
也可以将其视为一个列向量，即$n\times 1$-矩阵，此时记为
$$\boldsymbol{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$
- 对于$m\times n$-矩阵$A=$
$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$
我们将其第i行记为$\boldsymbol{r}_i=\begin{bmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{bmatrix}, i=1,\ldots,m$，那么$A$可以被表示为 $A=$
$$\begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_m \end{bmatrix}.$$
我们称其为矩阵A的行向量表示。
- 类似的，对于$m\times n$-矩阵$A=$
$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$
我们将其第j列记为$\boldsymbol{c}_j=\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$，那么$A$可以被表示为 $A=$
$$\begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \ldots & \boldsymbol{c}_n \end{bmatrix}.$$
我们称其为矩阵A的列向量表示。

本章的主要内容是定义矩阵的运算法则：

1. 矩阵$A \in M_{m_1\times n_1}$与矩阵$B \in M_{m_2\times n_2}$**相等**，即$A = B$，当且仅当$m_1 = m_2$，$n_1 = n_2$，并且$A_{ij} = B_{ij}$对一切$i = 1,\dots,m_1 = m_2$和$j = 1,\dots,n_1 = n_2$成立。
2. 对于矩阵$A \in M_{m\times n}$, 常数c, 矩阵A与c的**标量积(scalar product)**，记为cA，定义为 $(cA)_{ij} = cA_{ij}$。显然$cA \in M_{m\times n}$。
3. 若矩阵$A \in M_{m\times n}$与矩阵$B \in M_{m\times n}$，那么A与B的**和**(sum)定义为 $(A+B)_{ij} = A_{ij} + B_{ij}$。类似的，$A$与$B$的**差**(difference)定义为 $(A-B)_{ij} = A_{ij} - B_{ij}$。显然$A + B \in M_{m\times n}$，$A - B \in M_{m\times n}$。注意，若$A \in M_{m_1\times n_1}$，$B \in M_{m_2\times n_2}$**不满足**$m_1 = m_2$**,** $n_1 = n_2$，那么$A + B$与$A - B$**无法被定义**。
4. 对于$A_1 \in M_{m\times n}$, $A_2 \in M_{m\times n}$, ..., $A_r \in M_{m\times n}$，$c_1, c_2 \dots, c_r$为常数，定义$c_1A_1 + c_2A_2 + \dots + c_rA_r$为$m \times n$-矩阵，满足

$$(c_1A_1 + c_2A_2 + \dots + c_rA_r)_{ij} = c_1(A_1)_{ij} + c_2(A_2)_{ij} + \dots + c_r(A_r)_{ij}.$$

   矩阵$c_1A_1+c_2A_2+\dots+c_rA_r$被称为$A_1, A_2, \dots, A_r$关于系数$c_1, c_2, \dots, c_r$的**线性组合(linear combination of** $A_1, A_2, \dots, A_r$ **with coefficients** $c_1, c_2, \dots, c_r$**)**。
5. 现在我们来定义矩阵的**乘法**运算法则。首先我们定义**行向量与列向量之间的乘积**:令$\boldsymbol{r} = \begin{bmatrix} r_1 & r_2 & \dots & r_n \end{bmatrix}$为一个具有$n$个坐标的行向量，即$1 \times n$-矩阵；令$\boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$为一个具有$n$个坐标的列向量，即$n \times 1$-矩阵，那么它们的**乘积**$\boldsymbol{rc}$为一个常数，即$1 \times 1$-矩阵 $\boldsymbol{rc} = r_1c_1 + r_2c_2 + \dots r_nc_n$。
   利用这个规则，我们可以对更多的矩阵定义乘法运算。具体来说，若A为$m\times r$-矩阵，B为$r \times n$-矩阵，那么A与B的**乘积**AB为一个$m \times n$-矩阵，它满足：将A用行向量表示：$A =$

$$\begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_m \end{bmatrix},$$

   $\boldsymbol{r}_i = \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{ir} \end{bmatrix}, i = 1,\dots,m$; 将$B$用列向量表示：$B =$

$$\begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \dots & \boldsymbol{c}_n \end{bmatrix},$$

$\boldsymbol{c}_j = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{rj} \end{bmatrix}, j = 1, \ldots, n$，那么乘积$AB$在第$i$行第$j$列上的项定义为

$$(AB)_{ij} = \boldsymbol{r}_i \boldsymbol{c}_j = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ir}b_{rj},$$

即，$(AB)_{ij}$为矩阵A第i行与矩阵B第j列的**乘积**，对任何$i = 1, \ldots, m$，$j = 1, \ldots, n$。可将AB写为矩阵的形式：$AB =$

$$\begin{bmatrix} \boldsymbol{r}_1\boldsymbol{c}_1 & \boldsymbol{r}_1\boldsymbol{c}_2 & \ldots & \boldsymbol{r}_1\boldsymbol{c}_n \\ \boldsymbol{r}_2\boldsymbol{c}_1 & \boldsymbol{r}_2\boldsymbol{c}_2 & \ldots & \boldsymbol{r}_2\boldsymbol{c}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{r}_m\boldsymbol{c}_1 & \boldsymbol{r}_m\boldsymbol{c}_2 & \ldots & \boldsymbol{r}_m\boldsymbol{c}_n. \end{bmatrix}$$

- **例子**：假设$A =$

  $$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix},$$

  $B =$

  $$\begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}$$

  那么由于A为2×3-矩阵，B为3×4-矩阵，根据矩阵乘法的定义，AB为一个2×4-矩阵，且有

  $$(AB)_{11} = \underbrace{\begin{bmatrix} 1 & 2 & 4 \end{bmatrix}}_{A\text{的第1行}} \underbrace{\begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}}_{B\text{的第1列}} = 1 \times 4 + 2 \times 0 + 4 \times 2 = 12;$$

  $$(AB)_{12} = \underbrace{\begin{bmatrix} 1 & 2 & 4 \end{bmatrix}}_{A\text{的第1行}} \underbrace{\begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix}}_{B\text{的第2列}} = 1 \times 1 + 2 \times (-1) + 4 \times 7 = 27;$$

  $$(AB)_{13} = \underbrace{\begin{bmatrix} 1 & 2 & 4 \end{bmatrix}}_{A\text{的第1行}} \underbrace{\begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}}_{B\text{的第3列}} = 1 \times 4 + 2 \times 3 + 4 \times 5 = 30;$$

$$(AB)_{13} = \underbrace{\begin{bmatrix} 1 & 2 & 4 \end{bmatrix}}_{A\text{的第1行}} \underbrace{\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}}_{B\text{的第4列}} = 1 \times 3 + 2 \times 1 + 4 \times 2 = 13;$$

$$(AB)_{21} = \underbrace{\begin{bmatrix} 2 & 6 & 0 \end{bmatrix}}_{A\text{的第2行}} \underbrace{\begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}}_{B\text{的第1列}} = 2 \times 4 + 6 \times 0 + 0 \times 2 = 8;$$

$$(AB)_{22} = \underbrace{\begin{bmatrix} 2 & 6 & 0 \end{bmatrix}}_{A\text{的第2行}} \underbrace{\begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix}}_{B\text{的第2列}} = 2 \times 1 + 6 \times (-1) + 0 \times 2 = -4;$$

$$(AB)_{23} = \underbrace{\begin{bmatrix} 2 & 6 & 0 \end{bmatrix}}_{A\text{的第2行}} \underbrace{\begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}}_{B\text{的第3列}} = 2 \times 4 + 6 \times 3 + 0 \times 5 = 26;$$

$$(AB)_{24} = \underbrace{\begin{bmatrix} 2 & 6 & 0 \end{bmatrix}}_{A\text{的第2行}} \underbrace{\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}}_{B\text{的第4列}} = 2 \times 3 + 6 \times 1 + 0 \times 2 = 12.$$

因此$AB =$

$$\begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}.$$

- 假设A为一个$m \times r$-矩阵，$\boldsymbol{c}$为一个$r \times 1$-矩阵，即一个含有r个坐标的列向量，那么根据矩阵乘法定义，$A\boldsymbol{c}$为一个$m \times 1$-矩阵，即一个含有m个坐标的列向量。因此，若B为一个$r \times n$-矩阵，将B用列向量表示为

$$\begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \ldots & \boldsymbol{c}_n \end{bmatrix},$$

每个$\boldsymbol{c}_j$为$r \times 1$-矩阵， $j = 1, \ldots, n$，那么容易验证

$$AB = A \begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \ldots & \boldsymbol{c}_n \end{bmatrix} = \begin{bmatrix} A\boldsymbol{c}_1 & A\boldsymbol{c}_2 & \ldots & A\boldsymbol{c}_n \end{bmatrix},$$

即A与B的乘积AB的第j列为矩阵A与B的第j列$\boldsymbol{c}_j$ (作为$r \times 1$-矩阵)的乘积$A\boldsymbol{c}_j$，也即AB的列向量表示为：

$$\begin{bmatrix} A\boldsymbol{c}_1 & A\boldsymbol{c}_2 & \ldots & A\boldsymbol{c}_n \end{bmatrix}.$$

- 假设若B为一个$r \times n$-矩阵，$\boldsymbol{r}$为一个$1 \times r$-矩阵，即一个含有r个坐标的行向量，那么根据矩阵乘法定义，$\boldsymbol{r}B$为一个$1 \times n$-矩阵，即一个含有n个坐标的行向量。因此，若A为一个$m \times r$-矩阵，将A用行向量表示为

$$\begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_m \end{bmatrix},$$

每个$\boldsymbol{r}_i$为$1 \times r$-矩阵，$i = 1, \ldots, m$，那么容易验证

$$AB = \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_m \end{bmatrix} B = \begin{bmatrix} \boldsymbol{r}_1 B \\ \boldsymbol{r}_2 B \\ \vdots \\ \boldsymbol{r}_m B, \end{bmatrix}$$

即A与B的乘积AB的第i行为矩阵A的第i行$\boldsymbol{r}_i$ (作为$1 \times r$-矩阵)与矩阵B的乘积$\boldsymbol{r}_i B$，也即AB的行向量表示为

$$\begin{bmatrix} \boldsymbol{r}_1 B \\ \boldsymbol{r}_2 B \\ \vdots \\ \boldsymbol{r}_m B \end{bmatrix}.$$

- 对任何$m \times n$-矩阵A与含有n个坐标的向量$\boldsymbol{s} = (s_1, \ldots, s_n)$，将$\boldsymbol{s}$视为列向量，即$n \times 1$-矩阵，那么由矩阵乘法定义可知$A\boldsymbol{s}$是一个$m \times 1$-矩阵，即一个包含$m$个坐标的列向量：

$$A\boldsymbol{s} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} = \begin{bmatrix} a_{11}s_1 + a_{12}s_2 + \ldots + a_{1n}s_n \\ a_{21}s_1 + a_{22}s_2 + \ldots + a_{2n}s_n \\ \vdots \\ a_{m1}s_1 + a_{m2}s_2 + \ldots + a_{mn}s_n \end{bmatrix}.$$

因此，$\boldsymbol{s} = (s_1, \ldots, s_n)$是线性方程组

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

的一个解**当且仅当**$A\boldsymbol{s}=\boldsymbol{b}$，这里$A$为方程组的系数矩阵，$\boldsymbol{s}$视为列向量，$\boldsymbol{b}=$

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

也视为列向量。

- 将$m\times n$-矩阵A用列向量表示：

$$A=\begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \dots & \boldsymbol{c}_n \end{bmatrix},$$

那么显然有

$$A\boldsymbol{s}=\begin{bmatrix} a_{11}s_1+a_{12}s_2+\ldots+a_{1n}s_n \\ a_{21}s_1+a_{22}s_2+\ldots+a_{2n}s_n \\ \vdots \\ a_{m1}s_1+a_{m2}s_2+\ldots+a_{mn}s_n \end{bmatrix}=s_1\boldsymbol{c}_1+s_2\boldsymbol{c}_2+\ldots+s_n\boldsymbol{c}_n.$$

即$A\boldsymbol{s}$是A的列向量$\boldsymbol{c}_1,\ldots,\boldsymbol{c}_n$(视为$m\times 1$-矩阵)关于系数$s_1,\ldots,s_n$的**线性组合**。

6. 若A为一个$m\times n$-矩阵，那么它的**转置**(transpose)，记为$A^\top$，是一个$n\times m$-矩阵，满足

$$(A^\top)_{ij}=(A)_{ji}$$

对所有$i=1,\ldots,n$，$j=1,\ldots,m$。注意，若将A用**列向量**表示：

$$A=\begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \dots & \boldsymbol{c}_n \end{bmatrix},$$

那么$A^\top$的**行向量**表示为：

$$A^\top=\begin{bmatrix} \boldsymbol{c}_1^\top \\ \boldsymbol{c}_2^\top \\ \vdots \\ \boldsymbol{c}_n^\top \end{bmatrix},$$

这里$\boldsymbol{c}_i=\begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}$，$\boldsymbol{c}_i^\top=\begin{bmatrix} a_{1i} & a_{2i} & \dots & a_{mi} \end{bmatrix}$。例如矩阵

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

的转置为

$$B^\top = \begin{bmatrix} 2 & 1 & 5 \\ 3 & 4 & 6 \end{bmatrix}.$$

7. 若A为$n \times n$-**方阵**，那么它对角线上所有项的和被称为A的**迹**(trace)，记为

$$\text{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}.$$

例如，矩阵

$$A = \begin{bmatrix} \color{red}{-1} & 2 & 7 & 0 \\ 3 & \color{red}{5} & -8 & 4 \\ 1 & 2 & \color{red}{7} & -3 \\ 4 & -2 & 1 & \color{red}{0} \end{bmatrix}$$

的迹为$\text{tr}(A) = \color{red}{-1 + 5 + 7 + 0} = 11$。

8. 矩阵可以被划分为若干个子矩阵(submatrix)，比如

$$A = \left(\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

$$A_{11} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}, A_{12} = \begin{bmatrix} a_{14} \\ a_{24} \end{bmatrix}, A_{21} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \end{bmatrix}, A_{22} = \begin{bmatrix} a_{34} \end{bmatrix}$$

就是A的子矩阵。像$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$这样的表示也被叫做**分块矩阵**(partitioned matrix)。显然矩阵的行向量表示与列向量表示都是将A表示为分块矩阵。

以上所定义的关于矩阵的加，减，乘，转置的运算所遵循的规则请阅读英文教材第1.4节，尤其是**Theorem 1.4.1**，**Theorem 1.4.2**， **Theorem 1.4.8**。这里我们特别注意到矩阵的乘法运算法则与我们高中所熟悉的实数的乘法运算法则在有些地方是有重要区别的：

- 我们知道，若a与b为两个常数，那么必然有$ab = ba$成立，即实数的乘法是满足交换律的(the commutative law)。然而，**乘法的交换律**对于一般的矩阵A与B**并不一定成立**，这是因为
  1. 有可能AB的乘积可以定义，但BA的乘积无法被定义。比如A为2×3-矩阵，B为3×4-矩阵，此时AB可以被定义且为一个2×4-矩阵，但BA无法被定义。
  2. 有可能AB与BA都可以被定义，但它们的尺寸不同。比如A为2×3-矩阵，B为3×2-矩阵.此时AB为2×2-矩阵，BA为3×3-矩阵。

3. 有可能AB与BA都可以被定义并且它们的尺寸相同，但仍然有$AB \neq BA$。比如

$$A = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix},$$

那么$AB = \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix}$，$BA = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}$，显然$AB \neq BA$。

- 若一个矩阵所有项均为0，那么我们称其为一个零矩阵(zero matrix)。这里我们把零矩阵也记作0（实际上，若零矩阵为$m \times n$-矩阵，那么我们最好将其记作$0_{m\times n}$。但在不影响理解的情况下我们可以将其简单写作0）。在处理加法运算时，我们可以认为零矩阵扮演实数中的0的角色，比如对于任何矩阵A，$A + 0 = 0 + A = A$。然而在乘法运算中，我们需要小心以下牵扯到零矩阵的情况：

  1. 在实数乘法运算中(a, b, c均为常数)，若$ab = ac$且$a \neq 0$，那么必有$b = c$，即实数乘法满足消去律(the cancellation law)。该消去律无法照搬到矩阵乘法情况：对于一般的矩阵A, B, C，即使$AB = AC$且$A \neq 0$为一个非零矩阵，那么也不一定会有$B = C$。比如

  $$A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}, C = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix},$$

  此时满足$A \neq 0$，$B \neq C$，$AB = AC = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$，但$B \neq C$。

  2. 对于常数a, b，我们知道若$ab = 0$，那么有$a = 0$或$b = 0$。该法则对矩阵乘法不成立：

  $$A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 3 & 4 \\ 0 & 0 \end{bmatrix},$$

  显然$A \neq 0$，$B \neq 0$，且$AB = 0$。

  因此**在矩阵乘法当中，零矩阵不能视为实数中的**0。

由矩阵乘法的定义可知，对于$A, B \in M_{n\times n}$，我们有$AB \in M_{n\times n}$，即两个尺寸相同的方阵之间的乘积还是同尺寸的方阵。而n阶单位矩阵$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$ (除了对角线上的所有项为1，其余不在对角线上的项均为0) 则在n阶方阵乘法运算中扮演实数乘法中1的角色：

$$AI_n = I_n A = A$$

对任何n阶方阵A成立。

以下定理非常重要，我们会在后续的学习中持续使用该定理。

Theorem 1.7. *令R为一个n阶简化阶梯型。那么只可能有如下两种情况：*

1. *R包含至少一个为0的行；*
2. *$R = I_n$为n阶单位矩阵。*

证明. 我们只需证明若一个n阶简化阶梯型R所有行均非0，那么它一定是一个单位矩阵。在这种情况下，由简化阶梯型的定义可知，R矩阵里的每一行都包含一个主1。现在我们只需要证明第i行的主1位于第i列上对任何$i = 1, \ldots, n$成立。(若该性质成立，那么由简化阶梯型定义的第四个条件可知R必为单位矩阵)。
现在我们用**反证法与数学归纳法**证明该性质。假设R的第一行里的主1不在第一列上，那么第一行的主1所在的第j_1列必然满足$j \geq 2$。那么，由阶梯型定义的第三个条件可知，第二行主1的位置必须在第一行主1的右边，即，第二行主1所在的第j_2列必然满足$j_2 > j_1 \geq 2$，特别地，我们必然有$j_2 \geq 3$。以此类推，我们可推出第n行的主1所在的第j_n列必然满足$j_n \geq n+1$。然而，R只包含n列，因此第n行的主1所在的第j_n列也必须满足$j_n \leq n$。这意味着我们现在得到了$n+1 \leq j \leq n$，矛盾！因此第一行的主1必须位于第一列上。
现在假设我们对某个$i < n$证明了第i行的主1位于第i列上。再次利用阶梯型定义的第三个条件可知，第$i+1$行主1的位置必须在第i行主1的右边。那么由归纳假设，可知第$i+1$行主1所在的第j_{i+1}列必须满足$j_{i+1} \geq i+1$。再次利用反证法，假设$j_{i+1} \neq i+1$，即$j_{i+1} \geq i+2$，依照证明第一行主1在第一列上的方法，可推得此时第$i+2$行主1所在的第j_{i+2}列必须满足$j_{i+2} \geq i+3$，...，第n行的主1所在的第j_n列必须满足$j_n \geq n+1$，这意味着我们现在又一次得到了$n+1 \leq j \leq n$，矛盾！因此第i+1行的主1必须位于第i+1列上。因此由归纳法可知该性质对所有$i = 1, \ldots, n$都成立。 □

最后我们注意，对于任何$m \times n$-矩阵A，以下等式成立：

$$AI_n = A, \quad I_m A = A$$

这里I_n为n阶单位矩阵，，I_m为m阶单位矩阵。

1.4. 矩阵的逆， Inverses of matrices

由上一节的讨论我们知道，n阶单位矩阵在n阶方阵的乘法运算中扮演了与实数乘法中1一样的角色，即

$$AI_n = I_n A = A$$

对任何n阶方阵A成立。那么我们很自然地会问，**一个n阶方阵A的“倒数”是什么？** 对于一个实数a我们知道，只要$a \neq 0$那么a的倒数$\frac{1}{a}$一定存在，即满足$a\frac{1}{a} = \frac{1}{a}a = 1$。基于上一节的相关讨论我们已经知道在矩阵乘法中零矩阵并不能直接视为实数乘法中的0，因此在这里我们也会怀疑并非任何非零矩阵都拥有“倒数”。事实也确实如此。首先我们先给出矩阵“倒数”的定义。

Definition 1.8. *设A为一个n阶**方阵**。我们说A是**可逆的**(invertible)*
或者非奇异的(nonsingular)当且仅当存在一个n阶方阵B使得

$$AB = BA = I_n.$$

*若存在这样的n阶方阵B，则称B为A的**逆矩阵**或简称为A的**逆**(inverse)，并记为$B = A^{-1}$。*
*若这样的B**不存在**，那么则称A为**奇异的**(singular)。*

显然从以上定义可以看出若B满足

$$AB = BA = I_n,$$

那么$B = A^{-1}$就是A关于矩阵乘法的“倒数”。

Remark 1.9.
- *可逆性这个概念只对**方阵**有效。也可以说任何$m\times n$-矩阵在$m \neq n$时都是奇异的（不可逆的）。*
- *一个非零的n阶方阵不一定可逆，这与实数的情况不同。比如2阶方阵*

  $$A = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}$$

  不是零矩阵，但它不可逆，因为对任何2阶矩阵B，我们有

  $$AB = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ b_{11}+2b_{21} & b_{12}+2b_{22} \end{bmatrix}.$$

  因此对任何2阶方阵B，AB都不可能等于I_2。
- *显然若A的逆矩阵为$B = A^{-1}$，那么由定义可知矩阵B可逆且它的逆为A，即$B^{-1} = (A^{-1})^{-1} = A$。*
- *在1.5节我们会学习如何判断一个方阵是否可逆以及如何求出它的逆矩阵(在可逆的情况下)。*

首先我们容易证明矩阵逆的唯一性：

Theorem 1.10 (英文教材Theorem 1.4.4)**.** *假设B和C都是方阵A的逆，即$BA = AB = I_n$，$CA = AC = I_n$，那么必然有$B = C$。*

证明. 由于B是A的逆，我们有$AB = BA = I_n$。对该等式两边同时**右乘**矩阵C可得

$$(BA)C = I_n C = C.$$

另一方面，由矩阵乘法的结合律以及$AC = I_n$，我们还有

$$(BA)C = B(AC) = BI_n = B.$$

因此$B = C$。 □

根据逆矩阵的定义，若B为A的逆，那么需要满足$AB = BA = I_n$。事实上若要验证B为A的逆，我们只需要证明$AB = I_n$或者$BA = I_n$成立即可。

Theorem 1.11. *对一个n阶方阵A，若存在一个矩阵B使得$BA = I_n$成立，那么矩阵A可逆且$A^{-1} = B$，即若$BA = I_n$成立那么必有$AB = I_n$。类似的，若存在一个矩阵B使得$AB = I_n$成立，那么矩阵A可逆且$A^{-1} = B$，即若$AB = I_n$成立那么必有$BA = I_n$。*

我们目前还没有足够的工具对以上定理给出一个简单的证明。所以现在大家只需要记住这个结论，具体证明在后续学习中会给出。以下定理则很容易被证明。

Theorem 1.12 (英文教材Theorem 1.4.6)**.** *若A与B均为可逆n阶方阵，那么AB也可逆，且有$(AB)^{-1} = B^{-1}A^{-1}$。*

证明. 利用矩阵乘法的结合律，易得

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

因此由Theorem 1.11可得AB可逆，且$(AB)^{-1} = B^{-1}A^{-1}$。事实上我们也可以直接利用矩阵乘法的结合律验证$(B^{-1}A^{-1})(AB) = I_n$。 □

下面我们介绍矩阵的幂。

Definition 1.13. *设A为n阶方阵。我们定义矩阵A的幂(power of matrix)为*

$$A^0 = I_n, A^k = \underbrace{AA\dots A}_{k\text{个}}, \quad k = 1, 2, \dots$$

若A可逆，那么它的负幂定义为

$$A^{-k} = (A^{-1})^k = \underbrace{A^{-1}A^{-1}\dots A^{-1}}_{k\text{个}}, \quad k = 1, 2, \dots$$

容易验证以下等式成立：对任何自然数r,s，有

$$A^{r+s}=A^rA^s,\quad (A^r)^s=A^{rs},$$

若A可逆(因此A的负幂也存在)，那么以上等式对任何整数r,s成立。特别地我们有以下定理。

Theorem 1.14 (英文教材Theorem 1.4.7)**.** *设A为可逆n阶方阵，k为任意自然数。那么*

1. *A^{-1}可逆且$(A^{-1})^{-1}=A$，*
2. *A^k可逆且$(A^k)^{-1}=A^{-k}=(A^{-1})^k$，*
3. *对任何非0的常数c，cA可逆且$(cA)^{-1}=c^{-1}A^{-1}$。*

该定理的证明较为简单，请参考英文教材。

Remark 1.15. *对于实数a,b我们有二项式定理:*

$$(a+b)^k=a^kb^0+\binom{k}{1}a^{k-1}b+\binom{k}{2}a^{k-2}b^2+\ldots+\binom{k}{k-1}ab^{k-1}+a^0b^k,$$

但该公式对矩阵的情况不适用。比如存在矩阵A,B使得

$$(A+B)^2\neq A^2+2AB+B^2.$$

这是因为$(A+B)^2=A^2+AB+BA+B^2$，因此若$AB\neq BA$，那么就有

$$A^2+AB+BA+B^2\neq A^2+2AB+B^2.$$

*换句话说，**矩阵乘法的非交换性导致了二项式定理对矩阵的失效**。*

利用矩阵的幂我们可以定义矩阵多项式(matrix polynomial)。

Definition 1.16. *设A为一个n阶矩阵，令$p(x)=a_0+a_1x+a_2x^2+\ldots+a_mx^m$为一个多项式，$a_0,a_1,\ldots,a_m$为常数。我们定义矩阵多项式$p(A)$为*

$$p(A)=a_0I_n+a_1A+a_2A^2+\ldots+a_mA^m.$$

显然根据矩阵的运算法则可知$p(A)$为一个n阶矩阵。

例子：设$A=\begin{bmatrix}-1 & 2\\ 0 & 3\end{bmatrix}$，$p(x)=x^2-2x-3$，那么

$$\begin{aligned}p(A)&=A^2-2A-3I_2\\&=\begin{bmatrix}-1 & 2\\ 0 & 3\end{bmatrix}^2-2\begin{bmatrix}-1 & 2\\ 0 & 3\end{bmatrix}-3\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\\&=\begin{bmatrix}1 & 4\\ 0 & 9\end{bmatrix}-\begin{bmatrix}-2 & 4\\ 0 & 6\end{bmatrix}-\begin{bmatrix}3 & 0\\ 0 & 3\end{bmatrix}\\&=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.\end{aligned}$$

以下定理容易被证明：

Theorem 1.17 (英文教材Theorem 1.4.9)**.** 设A为n阶方阵，若A可逆，那么它的转置$A^\top$也可逆，且有

$$(A^\top)^{-1}=(A^{-1})^\top.$$

证明. 由于$(A^{-1}A)^\top=A^\top(A^{-1})^\top$ (这里利用运算法则$(AB)^\top=B^\top A^\top$，见英文教材Theorem 1.4.8(e)或者第二次作业Problem A)，以及显然的事实$I_n^\top=I_n$，我们得到

$$A^\top(A^{-1})^\top=(A^{-1}A)^\top=I_n^\top=I_n,$$

类似可证

$$(A^{-1})^\top A^\top=(AA^{-1})^\top=I_n^\top=I_n.$$

这意味着矩阵$B=(A^{-1})^\top$满足$AB=BA=I_n$。因此由定义Definition 1.8可知$A^\top$可逆且有

$$(A^\top)^{-1}=(A^{-1})^\top.$$

□

在本节的最后，我们介绍一种判断2 × 2-方阵是否可逆且求出其逆矩阵的方法。

Theorem 1.18 (英文教材Theorem 1.4.5)**.** 断2×2-方阵$A=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$可逆当且仅当$ad-bc\neq 0$。若该条件满足，那么$A^{-1}=\frac{1}{ad-bc}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$.

该定理的严格证明将在第二章给出。目前大家只需要记住该定理的内容以及能够使用即可。

例子： 考虑2×2-方阵阵$A=\begin{bmatrix}6 & 1\\ 5 & 2\end{bmatrix}$。计算可得

$$ad-bc=6\times 2-5\times 1=7\neq 0$$

因此由Theorem1.18可知该矩阵可逆，且有

$$A^{-1}=B=\frac{1}{7}\begin{bmatrix}2 & -1\\ -5 & 6\end{bmatrix}=\begin{bmatrix}\frac{2}{7} & \frac{-1}{7}\\ \frac{-5}{7} & \frac{6}{7}\end{bmatrix}.$$

经简单计算可快速验证$AB=BA=I_2$。
在1.5节我们将学习如何判断一般的n阶方阵是否可逆，并在可逆的情况下求出它的逆。

1.5. 利用初等矩阵求矩阵的逆, Elementary matrices and a method for finding A^{-1}

在本节我们将学习到如何判断一般的n阶方阵是否可逆，并在可逆的情况下求出它的逆。要做到这一点我们需要首先回顾初等行变换这个概念。根据Definition 1.4，我们有如下三种初等行变换：(设A为$m\times n$-矩阵)

1. 将A的第i行 ($i=1,\ldots,m$) 乘以某个**非零**的常数c，其他行保持不变 (可记为: i-th row $\times c$);
2. 将A的第i行 ($i=1,\ldots,m$)乘以某个常数c后加到第k行上去 ($k=1,\ldots,m$且$k\neq i$)，且除了第k行之外其他的行保持不变 (可记为: i-th row $\times c+k$-th row);
3. 将A的第i行 ($i=1,\ldots,m$)与第k行 ($k=1,\ldots,m$且$k\neq i$) 互换 (可记为: i-th row $\leftrightarrow k$-th row)，其他行保持不变。

以下这个与初等行变换相关的概念非常重要(期中考试重要考点之一):

Definition 1.19 (英文教材52页Definition 2)**.** *若一个m阶方阵E是通过**对**m**阶单位矩阵**I_m进行一次**初等行变换**得到的，那么E被称为一个初等矩阵(elementary matrix)*。

例子：

- $B=\begin{bmatrix}1 & 0\\ 0 & -3\end{bmatrix}$是一个初等矩阵，因为它显然是通过对2阶单位矩阵的第二行乘以-3得到的。

- $B=\begin{bmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{bmatrix}$是一个初等矩阵，因为它显然是通过对4阶单位矩阵交换第2行与第4行得到的。
- $B=\begin{bmatrix}1&0&3\\0&1&0\\0&0&1\end{bmatrix}$是一个初等矩阵，因为它显然是通过对3阶单位矩阵的第三行乘以3再加到第一行上得到的。
- $B=\begin{bmatrix}1&0&0\\0&0&0\\0&0&1\end{bmatrix}$不是一个初等矩阵，因为无论对3阶单位矩阵进行哪种初等行变换都不会得到这种形状的矩阵。
- $B=\begin{bmatrix}1&0&3\\0&0&1\\0&1&0\end{bmatrix}$不是一个初等矩阵，因为它是通过对3阶单位矩阵进行了两次初等行变化得到的：先对3阶单位矩阵的第三行乘以3再加到第一行上，再交换第三行与第二行。

以下定理告诉我们，任何矩阵的初等行变换(注意，这里指对矩阵做一次初等行变换)都可以通过对该矩阵左乘一个初等矩阵得到。

Theorem 1.20 (英文教材Theorem 1.5.1)**.** *设A为$m\times n$-矩阵。若E是一个m阶初等矩阵，即E通过对m阶单位矩阵进行一次初等行变换得到，那么EA等于对矩阵A做与E相同的初等行变换所得到的矩阵。具体来讲，*

- *若E是通过将I_m的第i行 $(i=1,\dots,m)$ 乘以某个非零的常数c得到，那么EA等于对A的第i行乘以这个非零的常数c所得到的的矩阵。*
- *若E是通过将I_m的第i行 $(i=1,\dots,m)$乘以某个常数c后加到第k行上去 $(k=1,\dots,m$且$k\neq i)$得到的，那么EA等于对A第i行乘以常数c后加到第k行上去所得到的矩阵。*
- *若E是通过将I_m的第i行 $(i=1,\dots,m)$与第k行 $(k=1,\dots,m$且$k\neq i)$ 互换得到的，那么EA等于对A第i行与第k行互换所得到的矩阵。*

证明. **注意：该证明不要求一定掌握(但建议理解证明过程)，但务必记住该定理的内容(期中考试考点之一)。**

我们首先把m阶单位矩阵I_m用行向量来表示：

$$I_m = \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_m \end{bmatrix}.$$

显然，由单位矩阵的定义可知，对任何$i = 1, \ldots, m$，$\boldsymbol{e}_i = \begin{bmatrix} 0 & \ldots & 0 & \underbrace{1}_{\text{第i列}} & 0 & \ldots & 0 \end{bmatrix}$，即$\boldsymbol{e}_i$除了第$i$列上的项为1之外，其他项均为0。

对于$m \times n$-矩阵A，将其表示为列向量形式$A = \begin{bmatrix} \boldsymbol{c}_1 & \ldots & \boldsymbol{c}_n \end{bmatrix}$，由矩阵乘法的列向量表示(见讲义第19页)可知，

$$\boldsymbol{e}_i A = \begin{bmatrix} \boldsymbol{e}_i \boldsymbol{c}_1 & \boldsymbol{e}_i \boldsymbol{c}_2 & \ldots \boldsymbol{e}_i \boldsymbol{c}_n \end{bmatrix}.$$

由$\boldsymbol{e}_i$的特殊形式，容易计算出

$$\begin{aligned} \boldsymbol{e}_i \boldsymbol{c}_1 &= 0 \times a_{11} + \ldots + 1 \times a_{i1} + \ldots 0 \times a_{m1} = a_{i1} \\ \boldsymbol{e}_i \boldsymbol{c}_2 &= 0 \times a_{12} + \ldots + 1 \times a_{i2} + \ldots 0 \times a_{m2} = a_{i2} \\ &\vdots \\ \boldsymbol{e}_i \boldsymbol{c}_n &= 0 \times a_{1n} + \ldots + 1 \times a_{in} + \ldots 0 \times a_{mn} = a_{in}. \end{aligned}$$

因此，我们得到

$$\boldsymbol{e}_i A = \begin{bmatrix} \boldsymbol{e}_i \boldsymbol{c}_1 & \boldsymbol{e}_i \boldsymbol{c}_2 & \ldots \boldsymbol{e}_i \boldsymbol{c}_n \end{bmatrix} = \begin{bmatrix} a_{i1} & a_{i2} & \ldots a_{in} \end{bmatrix} = \boldsymbol{r}_i = \text{A的第i行}. \tag{1.1}$$

现在我们依次对三种不同的初等行变换进行验证。

- 若E是通过将I_m的第i行 $(i = 1, \ldots, m)$ 乘以某个非零的常数c得到，那么显然E的行向量表示为

$$E = \begin{bmatrix} \boldsymbol{e}_1 \\ \vdots \\ c\boldsymbol{e}_i \\ \vdots \\ \boldsymbol{e}_m \end{bmatrix},$$

  由讲义第19页-20页关于矩阵乘积的行向量表示，我们知道EA的行向量表示为

$$EA = \begin{bmatrix} \boldsymbol{e}_1 A \\ \vdots \\ c\boldsymbol{e}_i A \\ \vdots \\ \boldsymbol{e}_m A \end{bmatrix},$$

由等式(1.1)可知，

$$EA = \begin{bmatrix} \boldsymbol{e}_1 A \\ \vdots \\ c\boldsymbol{e}_i A \\ \vdots \\ \boldsymbol{e}_m A \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_1 \\ \vdots \\ c\boldsymbol{r}_i \\ \vdots \\ \boldsymbol{r}_m \end{bmatrix},$$

这里$\boldsymbol{r}_1, \ldots, \boldsymbol{r}_m$为$A$的行向量。因此我们得到$EA$等于对$A$的第$i$行乘以这个非零的常数$c$所得到的的矩阵。

- 若E是通过将I_m的第i行 $(i = 1, \ldots, m)$乘以某个常数c后加到第k行上去 $(k = 1, \ldots, m$且$k \neq i)$得到的，那么显然E的行向量表示为

$$E = \begin{bmatrix} \boldsymbol{e}_1 \\ \vdots \\ c\boldsymbol{e}_i + \boldsymbol{e}_k(\text{第k行}) \\ \vdots \\ \boldsymbol{e}_m \end{bmatrix},$$

像第一种情况一样，EA的行向量表示为

$$EA = \begin{bmatrix} \boldsymbol{e}_1 A \\ \vdots \\ (c\boldsymbol{e}_i + \boldsymbol{e}_k)A \\ \vdots \\ \boldsymbol{e}_m A \end{bmatrix},$$

由等式(1.1)可知，

$$EA = \begin{bmatrix} \boldsymbol{e}_1 A \\ \vdots \\ (c\boldsymbol{e}_i + \boldsymbol{e}_k)A \\ \vdots \\ \boldsymbol{e}_m A \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_1 \\ \vdots \\ c\boldsymbol{r}_i + \boldsymbol{r}_k \\ \vdots \\ \boldsymbol{r}_m \end{bmatrix},$$

因此EA等于对A第i行乘以常数c后加到第k行上去所得到的矩阵。

- 与前两种情况类似，若E是通过将I_m的第i行 $(i = 1, \ldots, m)$与第k行 $(k =$

$1,\dots,m$且$k \neq i$) 互换得到的，那么显然E的行向量表示为

$$E = \begin{bmatrix} \boldsymbol{e}_1 \\ \vdots \\ \boldsymbol{e}_k(\text{第i行}) \\ \vdots \\ \boldsymbol{e}_i(\text{第k行}) \\ \boldsymbol{e}_m \end{bmatrix},$$

那么同样的推导可以证明EA等于对A第i行与第k行互换所得到的矩阵。具体细节留作练习。

□

例子：矩阵$A = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix}$，初等矩阵$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$是通过对$I_3$第1行乘以常数3后加到第3行上得到的，那么通过计算可得

$$EA = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 9 \end{bmatrix},$$

显然等于对矩阵A的第1行乘以常数3后加到第3行后得到的矩阵。

由初等行变换的定义可知，**每个初等矩阵E都可以通过一次初等行变换变成单位矩阵**。具体来讲，我们有：

1. 若E是通过将I_m的第i行 ($i = 1,\dots,m$) 乘以某个非零的常数c得到，那么将E的第i行乘以$1/c$可以得到单位矩阵I_m。
2. 若E是通过将I_m的第i行 ($i = 1,\dots,m$)乘以某个常数c后加到第k行上去 ($k = 1,\dots,m$且$k \neq i$)得到的，那么将E的第i行乘以$-c$加到第k行可以得到单位矩阵I_m。
3. 若E是通过将I_m的第i行 ($i = 1,\dots,m$)与第k行 ($k = 1,\dots,m$且$k \neq i$) 互换得到的，那么将E的第i行与第k行互换可以得到单位矩阵I_m。

以上观察暗示我们**任何初等矩阵都是可逆的，且它的逆也是一个初等矩阵**。

Theorem 1.21. *任何初等矩阵都是可逆的，且它的逆也是一个初等矩阵。具体来讲，*

- *若E是通过将I_m的第i行 ($i = 1,\dots,m$) 乘以某个非零的常数c得到，那么E^{-1}等于对I_m的第i行乘以$1/c$所得到的的矩阵。*

- 若E是通过将I_m的第i行 $(i=1,\dots,m)$乘以某个常数c后加到第k行上去 $(k=1,\dots,m$且$k\neq i)$得到的，那么E^{-1}等于对I_m第i行乘以常数$-c$后加到第k行上去所得到的矩阵。
- 若E是通过将I_m的第i行 $(i=1,\dots,m)$与第k行 $(k=1,\dots,m$且$k\neq i)$ 互换得到的，那么E^{-1}等于对I_m第i行与第k行互换所得到的矩阵。

证明.
- 若E是通过将I_m的第i行 $(i=1,\dots,m)$ 乘以某个非零的常数c得到，令E_0为通过对I_m的第i行乘以$1/c$所得到的的矩阵。由初等矩阵的定义可得E_0为初等矩阵。再利用Theorem 1.20我们知道E_0E等于对E的第i行乘以$1/c$所得到的的矩阵。基于我们之前的讨论可知，对E的第i行乘以$1/c$所得到的的矩阵为单位矩阵，即$E_0E=I_m$。由Theorem 1.11可断定E可逆且$E_0=E^{-1}$。
- 若E是通过将I_m的第i行 $(i=1,\dots,m)$乘以某个常数c后加到第k行上去 $(k=1,\dots,m$且$k\neq i)$得到的，令E_0为对I_m第i行乘以常数$-c$后加到第k行上去所得到的矩阵。同样的推理可得到$E_0E=I_m$，即这样的E可逆且$E_0=E^{-1}$。
- 若E是通过将I_m的第i行 $(i=1,\dots,m)$与第k行 $(k=1,\dots,m$且$k\neq i)$ 互换得到的，令E_0为对I_m第i行与第k行互换所得到的矩阵。同样的推理可得到$E_0E=I_m$，即这样的E可逆且$E_0=E^{-1}$。

$\square$

以下定理给出了判断一个方阵A是否可逆的标准，并间接给出了计算A^{-1}的一个算法。因此该定理也是期中考试考点之一。

Theorem 1.22 (英文教材Theorem 1.5.3, Equivalent Theorem)**.** *设A为n阶方阵。那么以下说法等价(equivalent):*

1. *A可逆。*
2. *齐次线性方程组$A\boldsymbol{x}=\mathbf{0}$只有一个平凡解(trivial solution)。注意这里$\mathbf{0}=\begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix}$为$n\times 1$的零列向量，$\boldsymbol{x}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$为$n\times 1$的未知数列向量，那么由矩阵乘法定义可知$A\boldsymbol{x}=\mathbf{0}$是在描述齐次线性方程组*

$$\begin{aligned}
a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=0\\
a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=0\\
&\vdots\\
a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n&=0.
\end{aligned}$$

3. A的简化阶梯型为单位矩阵I_n。

4. A可以被表示为一系列初等矩阵的乘积。

证明. 我们证明1. $\Rightarrow$ 2. $\Rightarrow$ 3. $\Rightarrow$ 4. $\Rightarrow$ 1.。

1. $\Rightarrow$ 2.: 假设A可逆。令$\boldsymbol{s}$为齐次线性方程组$A\boldsymbol{x}=\mathbf{0}$的一个解。由矩阵乘法定义(见讲义第20页的讨论)可知$\boldsymbol{s}$满足

$$A\boldsymbol{s}=\mathbf{0}.$$

由于A可逆，A^{-1}存在，因此对以上等式两边同时左乘A^{-1}可得

$$A^{-1}(A\boldsymbol{s})=A^{-1}\mathbf{0}=\mathbf{0},$$

而由矩阵乘法的结合律可知$A^{-1}(A\boldsymbol{s})=(A^{-1}A)\boldsymbol{s}=I_n\boldsymbol{s}=\boldsymbol{s}$。因此我们实际上已经得到$\boldsymbol{s}=\mathbf{0}$。这意味着齐次线性方程组$A\boldsymbol{x}=\mathbf{0}$的任何一个解都等于零向量$\mathbf{0}$，即，该方程组仅有一个平凡解。

2. $\Rightarrow$ 3.: 现在假设齐次线性方程组$A\boldsymbol{x}=\mathbf{0}$仅有一个平凡解。令$R$标记$A$对应的简化阶梯型。那么$R$必然不可能包含一个0行。这是因为$A$和$R$为$n\times n$-矩阵，因此若$R$包含有一个0行意味着$R$最多只能有$n-1$个非零行，也就是说此时$R$最多只能有$n-1$个主1，也即方程组$A\boldsymbol{x}=\mathbf{0}$最多只能有$n-1$个主元，也就是说，方程组$A\boldsymbol{x}=\mathbf{0}$至少有1个自由元(自由元的个数等于未知数个数-主元个数，这里我们有n个未知数)。显然自由元的存在意味着方程组$A\boldsymbol{x}=\mathbf{0}$不可能只有一个平凡解而是有无穷多个解，与假设2.矛盾！因此那么R必然不可能包含一个0行。

现在，利用Theorem 1.7可知，不包含0行的n阶简化阶梯型R必然为单位矩阵I_n。

3. $\Rightarrow$ 4.: 假设A对应的简化阶梯型R为单位矩阵I_n。我们知道$R=I_n$是通过对原系数矩阵A进行一系列初等行变换得到的。因此假设我们一共进行了k次初等行变换。**而由Theorem 1.20我们又知道每次初等行变换可以通过对矩阵左乘一个初等矩阵实现。** 这意味着存在k个初等矩阵$E_1,\dots E_k$使得

$$E_kE_{k-1}\dots E_1A=I_n.$$

(注意这里我们用E_1标记第一次初等行变换对应的初等矩阵，E_2代表第二次初等行变换对应的初等矩阵，因此E_2需要左乘于E_1A，这解释了为何顺序是$E_kE_{k-1}\dots E_1A$。) 显然，以上等式配合Theorem 1.11告诉我们A可逆，且$A^{-1}=E_kE_{k-1}\dots E_1$。另外，以上等式同时告诉我们

$$A=(A^{-1})^{-1}=(E_kE_{k-1}\dots E_1)^{-1}=E_1^{-1}E_2^{-1}\dots E_k^{-1}.$$

而由Theorem 1.21 我们知道初等矩阵的逆也为初等矩阵，因此以上所有$E_1^{-1},E_2^{-1},\dots,E_k^{-1}$均为初等矩阵，即$A=E_1^{-1}E_2^{-1}\dots E_k^{-1}$ 可以表示为若干个初等矩阵的乘积。

4. $\Rightarrow$ 1.: 假设$A=P_1P_2\dots P_k$可以表示为k个初等矩阵$P_1,P_2,\dots,P_k$的乘积，那么首

先因为Theorem 1.21可知所有$P_1, P_2, \dots, P_k$均可逆，再次利用Theorem 1.12可知作为若干个可逆矩阵的乘积，A也可逆，且有

$$A^{-1} = (P_1P_2 \dots P_k)^{-1} = P_k^{-1} \dots P_2^{-1}P_1^{-1}.$$

□

例子(2021年线性代数期中考试题)：Find an invertible matrix P such that $PA = B$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad B = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} - a_{21} & a_{32} - a_{22} & a_{33} - a_{23} \end{bmatrix}.$$

答案：显然，B是通过先对A的第二行乘以-1加到第三行上，再通过对A的第一行与第二行互换得到的。由于第一次初等行变换对应的初等矩阵为$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$，第二次初等行变化对应的初等矩阵为$E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$，我们得到 (**注意$E_1$与$E_2$相乘的顺序！**)

$$P = E_2E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

由于P是两个初等矩阵的乘积且初等矩阵皆可逆，可知P也可逆，符合题意要求。

Theorem 1.22 告诉我们可以通过观察A的简化阶梯型R的形状来判断A是否可逆，**即若R包含有一个0行，那么A必然不可逆；若R是一个单位矩阵，那么A可逆**。另外，Theorem 1.22 3. $\Rightarrow$ 4. 的证明过程实际上告诉我们，若我们通过一系列初等行变换将A转化为简化阶梯型$R = I_n$，即$E_kE_{k-1} \dots E_1A = I_n$，那么利用关系$A^{-1} = E_kE_{k-1} \dots E_1 = E_kE_{k-1} \dots E_1I_n$可知**同样的这一系列初等行变换作用在单位矩阵$I_n$上即可得到$A^{-1}$**。因此我们得到了求矩阵$A$的逆的一个算法 (inversion algorithm)。我们结合下面这个简单例子 (英文教材56页Example 4) 详细阐述这个求逆算法的具体步骤：

考虑矩阵$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$。我们现在想要求出它的逆矩阵$A^{-1}$。由以上讨论可知，

我们的核心思路是通过一系列初等行变换将A变为它的简化阶梯型$R = I_3$，那么通过将同样的这一系列初等行变换作用在单位矩阵I_3上即可得到A^{-1}。具体操作如下：

第一步：将矩阵A与单位矩阵I_3写在一起，A在左边，I_3在右边，中间可以用一条竖线隔开以避免将两者混淆。

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right)$$

第二步：在这一步里我们将左边的矩阵A使用初等行变换化为其简化阶梯型，即单位矩阵，同时在右边记录每一次初等行变换对I_3作用后所得到的矩阵。

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right)$$

左边：A的第一行乘以-2加到第二行，第一行乘以-1加到第三行；右边：I_3的第一行乘以-2加到第二行，第一行乘以-1加到第三行。

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right)$$

左边：A的第二行乘以2加到第三行；右边：I_3的第二行乘以2加到第三行。

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right)$$

左边：A的第三行乘以-1；右边：I_3的第三行乘以-1。

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right)$$

左边：A的第三行乘以3加到第二行，第三行乘以-3加到第一行；右边：I_3的第三行乘以3加到第二行，第三行乘以-3加到第一行。

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right)$$

左边：A的第二行乘以-2加到第一行；右边：I_3的第二行乘以-2加到第一行。

第三步：现在，左边的矩阵A被转化为了它的简化阶梯型$R = I_3$，那么现在出现在右边的矩阵即是A^{-1}:

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}.$$

例子(2022年线性代数期中考试填空题)：The inverse of the matrix $A = \begin{bmatrix} 1 & -2 \\ 3 & 5 \end{bmatrix}$ is $\underline{?}$

该题可以通过以上算法求出答案。当然，由于题目中的矩阵为2阶方阵，也可以直接利用Theorem 1.18 进行求解。

在第二章我们将学习到其他判定一个矩阵是否可逆以及求逆的方法。

1.6. 线性方程组与可逆矩阵, More on linear system and invertible matrices

在本节我们利用已经学到的知识对线性方程组解的情况进行严格的研究与证明。

Theorem 1.23 (英文教材Theorem 1.6.1)**.** *一个线性方程组的解只有三种情况：无解，有且仅有一个解，以及有无穷多个解。*

证明. 显然我们只需要证明若一个线性方程组有两个不同的解，那么实际上它就有无穷多个解。首先还是注意到一个线性方程组

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ &\vdots \\ &a_{m1}x_1 + a_{n2}x_2 + \ldots + a_{mn}x_n = b_m. \end{aligned}$$

可以被表示为$A\boldsymbol{x} = \boldsymbol{b}$，$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$为$n \times 1$的未知数列向量，$\boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$为$m \times 1$的列向量， $A \in M_{m\times n}$为系数矩阵。

实际上在之前的课堂上我已经简单提到过这个定理的证明思路。这里我给出详细的证明细节。假设$\boldsymbol{s} = \begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix}$与$\boldsymbol{t} = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}$为方程组$A\boldsymbol{x} = \boldsymbol{b}$的两个不同的解，即$\boldsymbol{s} \neq \boldsymbol{t}$，但

$$A\boldsymbol{s} = \boldsymbol{b} = A\boldsymbol{t}.$$

现在定义列向量$\boldsymbol{u}=\boldsymbol{s}-\boldsymbol{t}$，由于$\boldsymbol{s}\neq\boldsymbol{t}$，可知$\boldsymbol{u}\neq\boldsymbol{0}$不是零列向量。那么这意味着对任何不同的两个常数$c\neq d$，我们必然有$c\boldsymbol{u}\neq d\boldsymbol{u}$。另一方面，因为$A\boldsymbol{s}=\boldsymbol{b}=A\boldsymbol{t}$，我们还有$A\boldsymbol{u}=A\boldsymbol{s}-A\boldsymbol{t}=\boldsymbol{b}-\boldsymbol{b}=\boldsymbol{0}$，即，$\boldsymbol{u}=\boldsymbol{s}-\boldsymbol{t}$作为同一个线性方程组的两个解的差是该方程组系数矩阵A诱导出的齐次线性方程组$A\boldsymbol{x}=\boldsymbol{0}$的一个解。
现在，我们对每一个常数c定义一个新的列向量$\boldsymbol{s}+c\boldsymbol{u}$。那么由以上讨论可知

$$A(\boldsymbol{s}+c\boldsymbol{u})=A\boldsymbol{s}+cA\boldsymbol{u}=\boldsymbol{b}+\boldsymbol{0}=\boldsymbol{b},$$

这意味着所有这些$\boldsymbol{s}+c\boldsymbol{u}$都是方程组$A\boldsymbol{x}=\boldsymbol{b}$的解。又因为我们已经知道对于任何不同的两个常数$c\neq d$，我们必然有$c\boldsymbol{u}\neq d\boldsymbol{u}$，即$\boldsymbol{s}+c\boldsymbol{u}\neq\boldsymbol{s}+d\boldsymbol{u}$。这意味着该方程组实际上有无穷多个解(因为实数$c$有无穷多个)。证毕。 □

若一个方程组的方程个数与未知数个数相等，且系数矩阵A可逆，那么该方程组的唯一解可以直接通过对方程组最右边的常数列向量$\boldsymbol{b}$左乘A^{-1}得到。

Theorem 1.24 (英文教材Theorem 1.6.2)**.** *若A为可逆n阶方阵，那么对任意$n\times 1$-矩阵(即列向量)$\boldsymbol{b}$，以$(A,\boldsymbol{b})$作为增广矩阵的线性方程组$A\boldsymbol{x}=\boldsymbol{b}$的唯一解为*

$$\boldsymbol{s}=A^{-1}\boldsymbol{b}.$$

证明. 首先容易验证对于列向量$\boldsymbol{s}=A^{-1}\boldsymbol{b}$我们有$A\boldsymbol{s}=A(A^{-1}\boldsymbol{b})=(AA^{-1})\boldsymbol{b}=\boldsymbol{b}$，这意味着$\boldsymbol{s}=A^{-1}\boldsymbol{b}$是该方程组的一个解。
现在我们证明$\boldsymbol{s}=A^{-1}\boldsymbol{b}$是该方程组唯一的一个解。那么假设$\boldsymbol{t}$是方程组的另一个解，即$A\boldsymbol{t}=\boldsymbol{b}$。对该等式两边同左乘$A^{-1}$可得

$$A^{-1}A\boldsymbol{t}=A^{-1}\boldsymbol{b}$$

从而得到$\boldsymbol{t}=A^{-1}\boldsymbol{b}=\boldsymbol{s}$。因此该方程组只有$\boldsymbol{s}$一个解。 □

利用以上定理我们可以将Theorem 1.22 扩展到以下形式。

Theorem 1.25 (英文教材Theorem 1.6.4)**.** *设A为n阶方阵。那么以下说法等价(equivalent):*

1. A可逆。
2. 齐次线性方程组$A\boldsymbol{x}=\boldsymbol{0}$只有一个平凡解(*trivial solution*)。
3. A的简化阶梯型为单位矩阵I_n。
4. A可以被表示为一系列初等矩阵的乘积。
5. 对任何$n\times 1$-矩阵(即列向量)$\boldsymbol{b}$，线性方程组$A\boldsymbol{x}=\boldsymbol{b}$有解。
6. 对任何$n\times 1$-矩阵(即列向量)$\boldsymbol{b}$，线性方程组$A\boldsymbol{x}=\boldsymbol{b}$有且只有一个解。

证明. 1. 到 4. 的等价关系我们已经在Theorem 1.22中证明。我们这里证明1. $\Rightarrow$ 6. $\Rightarrow$ 5. $\Rightarrow$ 1.。
1. $\Rightarrow$ 6.: 见Theorem 1.24的证明。
6. $\Rightarrow$ 5.: 显然。
5. $\Rightarrow$ 1.: 由假设，对任何$n\times 1$-矩阵(即列向量)$\boldsymbol{b}$，线性方程组$A\boldsymbol{x}=\boldsymbol{b}$都有解。特别地，考虑$i=1,\ldots,n$， $\boldsymbol{e}_i=\begin{bmatrix}0 & \ldots & 0 & \underbrace{1}_{\text{第i列}} & 0 & \ldots & 0\end{bmatrix}$(即$\boldsymbol{e}_i$除了第$i$列上的项为1之外，其他项均为0)，令$\boldsymbol{b}$依次取$\boldsymbol{b}=\boldsymbol{e}_i^\top$， $i=1,\ldots,n$，由假设我们可以找到$n\times 1$-列向量$\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n$使得$A\boldsymbol{s}_i=\boldsymbol{e}_i^\top$，即

$$A\boldsymbol{s}_1=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}, A\boldsymbol{s}_2=\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix}, \ldots, A\boldsymbol{s}_n=\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}.$$

令矩阵$C=\begin{bmatrix}\boldsymbol{s}_1 & \boldsymbol{s}_2 & \ldots & \boldsymbol{s}_n\end{bmatrix}$，即$C$的列向量为$\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n$。那么由矩阵乘法的列向量表示，见讲义第19页，我们得到

$$AC=\begin{bmatrix}A\boldsymbol{s}_1 & A\boldsymbol{s}_2 & \ldots & A\boldsymbol{s}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{e}_1^\top & \boldsymbol{e}_2^\top & \ldots & \boldsymbol{e}_n^\top\end{bmatrix}=\begin{bmatrix}1 & 0 & \ldots & 0\\0 & 1 & \ldots & 0\\\vdots & \vdots & \ldots & \vdots\\0 & 0 & \ldots & 1\end{bmatrix}=I_n.$$

这意味着A可逆，且$A^{-1}=C$。 □

在Theorem 1.12 里我们知道若A,B为n阶可逆方阵，那么AB也可逆，且有$(AB)^{-1}=B^{-1}A^{-1}$。下面这个定理证明反之亦成立。

Theorem 1.26 (英文教材Theorem 1.6.5)**.** 设A,B为n阶方阵，且AB可逆。那么A与B都可逆。

证明. 首先证明B可逆。假设$\boldsymbol{s}$是齐次线性方程组$B\boldsymbol{x}=\boldsymbol{0}$的一个解，即$B\boldsymbol{s}=\boldsymbol{0}$，那么我们有$AB\boldsymbol{s}=A(B\boldsymbol{s})=\boldsymbol{0}$，即$\boldsymbol{s}$也是齐次线性方程组$(AB)\boldsymbol{x}=\boldsymbol{0}$的一个解。由假设，$AB$可逆，那么由Theorem 1.22 1. $\Rightarrow$ 2.，可知$(AB)\boldsymbol{x}=\boldsymbol{0}$仅有一个平凡解，因此必然有$\boldsymbol{s}=\boldsymbol{0}$。这意味着齐次线性方程组$B\boldsymbol{x}=\boldsymbol{0}$只有平凡解。因此利用Theorem 1.22 2. $\Rightarrow$ 1.，我们可以推出B可逆。
由于B可逆，我们有$A=AI_n=A(BB^{-1})=(AB)B^{-1}$。那么$A$作为可逆矩阵$AB$与$B^{-1}$的乘积也必然可逆(利用Theorem 1.12)。证毕。 □

1.7. 对角矩阵，三角矩阵与对称矩阵, Diagonal, triangular and symmetric matrices

这一节主要介绍三种特殊的矩阵，即对角矩阵，三角矩阵与对称矩阵。由于该节内容较简单，讲义里将只简单提及一些相对重要的知识点，建议大家在读完讲义后再仔细阅读英文教材的第1.7节。

注意本节出现的所有矩阵都是方阵。

首先介绍对角矩阵。这个概念我们已经很熟悉，即若一个**方阵**所有非对角线上的项均为0，那么该矩阵就是一个对角矩阵(diagonal matrix)。比如单位矩阵就是一个典型的对角矩阵。显然，对一个对角矩阵D而言，我们只需要知道它对角线上的项。通常我们将对角矩阵写为

$$D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix},$$

即$D_{11} = d_1, D_{22} = d_2, \dots, D_{nn} = d_n$。关于对角矩阵我们需要掌握以下两点：

- 对于任何自然数k，我们有
$$D^k = \begin{bmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^k \end{bmatrix}.$$
  这个特点让我们可以轻易计算对角矩阵的幂。
- **D可逆当且仅当它对角线上的所有项均不为零，或者说当且仅当**$d_1 d_2 \dots d_n \neq 0$。

第二个需要介绍的特殊矩阵是三角矩阵(triangular matrix)。三角矩阵又分为上三角矩阵(upper triangular matrix)与下三角矩阵(lower triangular matrix)。具体来讲

- 若一个方阵对角线**右上方**的所有项均为零，那么该方阵是一个**下三角矩阵**。即，方阵A是一个下三角矩阵当且仅当$A_{ij} = 0$对于所有$i < j$成立。
- 若一个方阵对角线**左下方**的所有项均为零，那么该方阵是一个**上三角矩阵**。即，方阵A是一个上三角矩阵当且仅当$A_{ij} = 0$对于所有$i > j$成立。

关于三角矩阵的特点请仔细阅读英文教材Theorem 1.7.1。这里我们简单提及该定理(b)部分的证明，即证明：若A, B为n阶下三角矩阵，那么AB也为下三角矩阵。(上三角的情况也同样成立。)

假设A, B为n阶下三角矩阵，即$A_{ij} = 0, B_{ij} = 0$对所有$i < j$成立。令$C = AB$。要

证明C是下三角矩阵，我们只需证明$C_{ij}=0$对所有$i<j$成立。由矩阵乘法定义可知，

$$C_{ij}=A_{i1}B_{1j}+A_{i2}B_{2j}+\ldots+A_{in}B_{nj}.$$

那么，对$i<j$的情况，我们可以将以上求和拆成两部分：

$$C_{ij}=A_{i1}B_{1j}+A_{i2}B_{2j}+\ldots+A_{i(j-1)}B_{(j-1)j}+A_{ij}B_{jj}+\ldots+A_{in}B_{nj}.$$

注意在红色部分中出现的$B_{1j},\ldots,B_{(j-1)j}$都为0 (由于$1<j,\ldots,j-1<j$，所以$B_{1j},\ldots,B_{(j-1)j}$都位于B对角线的右上方，故而等于零，因为B为下三角矩阵)，因此红色部分等于$A_{i1}\times 0+A_{i2}\times 0+\ldots+A_{i(j-1)}\times 0=0$；注意蓝色部分出现的$A_{ij},\ldots,A_{in}$都为0 (由于$i<j,i<j+1,\ldots,i<n$，$A_{ij},\ldots,A_{in}$都位于矩阵$A$右上方，故而等于零，因为$A$为下三角矩阵)，因此蓝色部分等于$0\times B_{jj}+\ldots+0\times B_{nj}=0$。这意味着$C_{ij}=0$。证毕。

最后我们介绍对称矩阵(symmetric matrix)。一个方阵A是对称的当且仅当$A^\top=A$，即它的转置矩阵等于它本身。显然这意味着$A_{ij}=A_{ji}$对所有i,j成立。

关于对称矩阵的性质请仔细阅读英文教材Theorem 1.7.2, 1.7.3, 1.7.4, 1.7.5。这里值得一提的是若矩阵$B\in M_{m\times n}$，那么$BB^\top$为m阶对称矩阵，$B^\top B$为n阶对称矩阵：

$$(BB^\top)^\top=(B^\top)^\top B^\top=BB^\top,$$

$$(B^\top B)^\top=B^\top(B^\top)^\top=B^\top B.$$