线性代数 Week 7

线性代数 (VII)

Lecturer: 杨启哲 Last modified: 2025 年 4 月 8 日

在这一章，我们将首先给出上一章所遗留的定理的证明:

定理 0.1

给定一个 $m \times n$ 的矩阵 A，则 row-rank(A) = column-rank(A).

然后我们将给出一般情况下 $Ax = b$ 的解的情况。

1 矩阵的秩

回顾之前的高斯消元法，其作用是将一个方阵 A 转换成了一个上三角矩阵 U，如下所示:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 9 \\ 7 & 8 & 16 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 7 & 8 & 16 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 0 & -6 & -5 \end{bmatrix} \Rightarrow U = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 1 \end{bmatrix}$$

我们将通过解方程的方式来证明定理0.1，为此我们需要在一般情况下 (方程个数和变元个数不一定相同) 时进一步讨论高斯消元法。注意到对任意的矩阵，上述变元的消去过程可能如下所示:

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 9 & 2 \\ 7 & 8 & 16 & 7 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -3 & 2 \\ 7 & 8 & 16 & 7 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -3 & 2 \\ 0 & -6 & -5 & 7 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -3 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 9 \\ 7 & 8 & 16 \\ 8 & 7 & 12 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 7 & 8 & 16 \\ 8 & 7 & 12 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 0 & -6 & -5 \\ 0 & -9 & -12 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

我们称这样最后形成的矩阵为**行阶梯形矩阵**:

定义 1.1 (行阶梯形矩阵)

对于一个 $m \times n$ 的矩阵 A:

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

对于其中每个 $i \in [m]$，我们定义：

$$j_i = \begin{cases} +\infty & \text{如果第 } i \text{ 行是零行} \\ \min\{j \in [n] : a_{ij} \neq 0\} & \text{o.w.} \end{cases}$$

即 $a_{j i}$ 是第 j 行中最左边的不为 0 的系数。如果存在 $0 \leqslant r \leqslant m$ 使得：

$$j_1 < j_2 < \cdots < j_r, \quad j_{r+1} = \cdots = j_m = +\infty$$

则称 A 为**行阶梯形矩阵**。

如果 $Ax = \mathbf{b}$ 的 A 是行阶梯形，则我们称其是**行阶梯形的**。

一些例子。 上三角矩阵都是行阶梯形：

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

下列矩阵也是行阶梯形矩阵：

$$\begin{bmatrix} 4 & -13 & 0 & 2 \\ 0 & 11 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 4 & -13 & 0 & 2 \\ 0 & 11 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

命题 1.2

令方程组 $Ax = \mathbf{b}$ 是行阶梯形的，并且令 r 满足：

$$j_1 < j_2 < \cdots < j_r, \quad j_{r+1} = \cdots = j_m = +\infty$$

则方程组有 r 个首元。

证明： 由首元和行阶梯形的定义即得。 □

现在我们可以给出高斯消元法在一般方程组下的更为精确的描述：

高斯消元法的一般描述。 对于方程组 $Ax = \mathbf{b}$，我们维护一个值 $r \leqslant m$ 使得：

R1 第一行到第 r 行已经满足行阶梯形的形式 (Row Echelon Form)，即：

$$j_1 < j_2 < \cdots < j_r \leqslant n$$

R2 对于矩阵中第 $r+1$ 行到第 m 行第 1 列到第 j_r 列是一个全零矩阵，即：

$$a_{ij} = 0, \quad \forall i \in [r+1, m], \forall j \in [1, j_r]$$

第一步: 初始 $r = 0$:

1. 选择最小的 $j \in [n]$ 使得第 j 列有非零的元素，如果这样的 j 不存在，意味着 $A = O$ 并且当 $r = 0$ 时 R1，R2 已经满足，算法结束。
2. 选择最大的$i \in [m]$ 使得第 $a_{ij} \neq 0$。
3. 将第 1 行与第 i 行进行交换, 交换后我们得到在新的第一行中 $a_{1j} \neq 0$, 并且对于所有的 $t \in [1, j-1]$ 都有 $a_{1t} = 0$，此时我们有：$j_1 = j$.
4. 对每个 $i' > 1$，我们将第 i' 行替换成:

$$(\text{row } i') - \frac{a_{i'j}}{a_{1j}}(\text{row } 1)$$

替换之后我们有对于所有的 $i' > 1$，$a_{i'j} = 0$。

显然此时对于 $r = 1$，R1 和 R2 都是满足的。

第 $i+1$ 步: 假设对于 $r = i$，R1 和 R2 都是满足的，我们继续进行如下操作:

1. 选择最小的 $j \in [j_i + 1, n]$ 使得存在 $i' \in [i+1, m]$ 使得 $a_{i'j} \neq 0$，如果这样的 j 不存在，算法结束。
2. 选择最大的$i' \in [i+1, m]$ 使得第 $a_{i'j} \neq 0$。
3. 将第 $i+1$ 行与第 i' 行进行交换，交换后我们得到在新的第 $i+1$ 行中 $a_{(i+1)j} \neq 0$, 并且对于所有的 $t \in [1, j-1]$ 都有 $a_{(i+1)t} = 0$，此时我们有：$j_{i+1} = j$.
4. 对每个 $i' > i+1$，我们将第 i' 行替换成:

$$(\text{row } i') - \frac{a_{i'j}}{a_{(i+1)j}}(\text{row } (i+1))$$

替换之后我们有对于所有的 $i' > i+1$，$a_{i'j} = 0$。

显然此时对于 $r = i+1$，R1 和 R2 都是满足的。注意到 $r \leqslant m$，每一步使得算法结束或者满足 R1 和 R2 的 r 增加 1，因此算法一定会结束。因此:

引理 1.3

给定一个 $m \times n$ 的矩阵 A，高斯消元法将 A 变成一个有 r 个首元的行阶梯形矩阵 U。

注 1.4

上述过程中，选择最大的是可以更改的，如课堂里我们是按最小的来选择的，这一步主要是为了挑选一个用来进行消元的行。

我们再用一个例子来说明上述过程:

例 1.5

考察这样一个矩阵:

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

<u>**第一步**</u> 的具体过程为:

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

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

<u>**第二步**</u> 的具体过程为:

$$\begin{bmatrix}0&2&4&2&0\\0&0&0&0&-3\\0&0&6&4&3\\0&0&3&2&2\end{bmatrix}\Rightarrow\begin{bmatrix}0&2&4&2&0\\0&0&0&0&-3\\0&0&6&4&3\\0&0&3&2&2\end{bmatrix}\Rightarrow\begin{bmatrix}0&2&4&2&0\\0&0&6&4&3\\0&0&0&0&-3\\0&0&3&2&2\end{bmatrix}$$

$$\Rightarrow\begin{bmatrix}0&2&4&2&0\\0&0&6&4&3\\0&0&0&0&-3\\0&0&3&2&2\end{bmatrix}\Rightarrow\begin{bmatrix}0&2&4&2&0\\0&0&6&4&3\\0&0&0&0&-3\\0&0&0&0&\frac{1}{2}\end{bmatrix}\Rightarrow\begin{bmatrix}0&2&4&2&0\\0&0&6&4&3\\0&0&0&0&-3\\0&0&0&0&\frac{1}{2}\end{bmatrix}$$

<u>**第三步**</u> 的具体过程为:

$$\begin{bmatrix}0&2&4&2&0\\0&0&6&4&3\\0&0&0&0&-3\\0&0&0&0&\frac{1}{2}\end{bmatrix}\Rightarrow\begin{bmatrix}0&2&4&2&0\\0&0&6&4&3\\0&0&0&0&-3\\0&0&0&0&\frac{1}{2}\end{bmatrix}\Rightarrow\begin{bmatrix}0&2&4&2&0\\0&0&6&4&3\\0&0&0&0&-3\\0&0&0&0&\frac{1}{2}\end{bmatrix}$$

$$\Rightarrow\begin{bmatrix}0&2&4&2&0\\0&0&6&4&3\\0&0&0&0&-3\\0&0&0&0&0\end{bmatrix}\Rightarrow\begin{bmatrix}0&2&4&2&0\\0&0&6&4&3\\0&0&0&0&-3\\0&0&0&0&0\end{bmatrix}$$

此时我们有:

$$j_1=2,\quad j_2=3,\quad j_3=5$$

显然我们有如下结论：

引理 1.6

给定一个 $m \times n$ 的矩阵 A，高斯消元法将 A 变成一个有 r 个首元的行阶梯形矩阵 U。

现在我们用**首元的个数**来定义**矩阵的秩**：

定义 1.7（矩阵的秩）

给定一个 $m \times n$ 的矩阵 A，其秩 (rank) 定义为：

$$\operatorname{rank}(A) = \text{矩阵 } A \text{ 的首元的个数} = r.$$

注 1.8（关于矩阵的秩）

关于矩阵的秩，有很多定义方式，比如同济大学数学科学学院编制的教材《工程数学线性代数》第 7 版中给出的定义为**矩阵 A 的最高阶非零子式**，，而梁鑫等人编著的《线性代数入门》则是直接定义为**矩阵 A 的列秩**。这些定义方式都是等价的，当然我会慢慢的证明，对于一个矩阵，这些定义下的秩都是一样的，从而在使用的时候可以进行混淆。

例 1.9（可逆矩阵的逆）

顾上一节，注意到一个 n 阶可逆矩阵自然会有 n 个首元，从而：

> 给定一个 n 阶可逆矩阵 A，则：
>
> $$\text{rank}(A) = \text{column-rank}(A) = \text{row-rank}(A) = n.$$

现在我们来考察上述的行阶梯形矩阵，如：

$$A = \begin{bmatrix} 4 & -13 & 0 & 2 \\ 0 & 11 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

不难发现：$\text{rank}(A) = \text{row-rank}(A) = 2$，所以我们有：

命题 1.11

令 A 是一个 $m \times n$ 的行阶梯形矩阵，则 $\text{rank}(A) = \text{row-rank}(A)$.

事实上，我们可以证明不仅是针对行阶梯形矩阵，对于任意的矩阵 A，其**列秩、行秩和秩**都是一致的：

定理 1.12

给定一个 $m \times n$ 的矩阵 A，则

$$\text{rank}(A) = \text{row-rank}(A) = \text{column-rank}(A)$$

这也直接证明了上一章所遗留的定理0.1。证明的关键是考虑如下的一类矩阵 A:

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

其是一个行阶梯形矩阵，不仅如此其首元都是 1，并且**首元的上方也都为** 0，从而我们有:

$$\text{rank}(R) = \text{row-rank}(R) = \text{column-rank}(R) = 3$$

我们称这样的矩阵为**行最简形**:

定义 1.13

回顾一个 $m \times n$ 矩阵 A，定义:

$$j_i = \begin{cases} +\infty & \text{如果第 } i \text{ 行是零行} \\ \min\{j \in [n] : a_{ij} \neq 0\} & \text{o.w.} \end{cases}$$

则称 A 是**行最简形的 (Reduced Row Echelon Form)**，如果:

1. 其是行阶梯形的，即存在 $0 \leqslant r \leqslant m$ 使得:
$$j_1 < j_2 < \cdots < j_r, \quad j_{r+1} = \cdots = j_m = +\infty$$
2. $a_{1j_1} = \cdots = a_{rj_r} = 1$，即所有的首元都是 1。
3. 对于所有的 $l \in [1, r], i \in [1, l-1] \cup [l+1, m]$，我们都有 $a_{ij_l} = 0$，即在首元的那一列中，除了首元之外的所有元素都是 0。

首先注意到，对于一个行最简形矩阵 R, 我们有:

命题 1.14

令 R 是一个 $m \times n$ 的行最简形矩阵，则 $\text{rank}(R) = \text{row-rank}(R) = \text{column-rank}(R)$.

其次对于一个行阶梯形矩阵 A，我们只需要:

1. 选择最大的 j_r 满足 $j_r \neq \infty$，并且存在 $k < j_r$ 满足 $A(j_r, k) \neq 0$，如果这样的 j_r 不存在，算法结束。
2. 对每个 $k < j_r$ 满足 $A(j_r, k) \neq 0$，我们将第 k 行替换成:
$$(\text{row } k) - \frac{A(j_r, k)}{A(j_r, j_r)} (\text{row } j_r)$$

就能将 A 变成一个行最简形矩阵 R，而这恰恰就是**高斯-若尔当消元法**:

引理 1.15

给定一个 $m \times n$ 的矩阵 A，高斯-若尔当消元法将 A 变成一个行最简形矩阵 R，并且有：

$$\text{rank}(A) = \text{rank}(R) = \text{column-rank}(R) = \text{row-rank}(R)$$

现在我们回到定理1.12，我们将证明如下引理：

引理 1.16

初等行变换不改变矩阵的行秩。

引理 1.17

初等行变换不改变矩阵的列秩。

在引理1.16和引理1.17的基础上，我们可以证明定理1.12：

证明：[定理1.12的证明]

1. 我们将 A 变成行阶梯形矩阵 U，并且假设其有 r 个首元，则 $\text{rank}(A) = r$。
2. 进一步我们将 U 变换成行最简形，得到 R，其还是有 r 个首元。
3. $\text{rank}(R) = \text{column-rank}(R) = \text{row-rank}(R) = \text{rank}(A)$.
4. 注意到我们只使用了初等行变换，从而：

$$\text{column-rank}(A) = \text{column-rank}(R), \quad \text{row-rank}(A) = \text{row-rank}(R)$$

即：

$$\text{rank}(A) = \text{row-rank}(A) = \text{column-rank}(A)$$

□

下面我们完成引理1.16和引理1.17的证明，注意到高斯若尔当消元法是基于如下的三种行变换来进行的：

1. 行加法：将一行的倍数加到另一行，矩阵如下所示：

$$E_{ij}(-k)A = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 1 & \cdots & -k & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_i - ka_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{bmatrix}$$

2. 行交换：交换两行的位置，矩阵如下所示：

$$P_{ij}A = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{bmatrix}$$

3. 行乘法：将一行乘以一个非零常数，矩阵如下所示：

$$D_i(k)A = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & k & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ k\mathbf{a}_i \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_n \end{bmatrix}$$

所以我们只需要证明对于这三种初等行变换，其不改变矩阵的列秩和行秩即可：

证明： [引理1.16的证明] 令 $A = \begin{bmatrix} \mathbf{a}_1^T & \mathbf{a}_i^T & \cdots & \mathbf{a}_m^T \end{bmatrix}^T$，即令其写成行向量的形式，这里 $\mathbf{a}_i$ 是 $1 \times n$ 的行向量。则经过一次初等行变换后，矩阵会变成如下的形式：

$$A' = \begin{bmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_i' \\ \vdots \\ \mathbf{a}_m' \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ k\mathbf{a}_i \\ \vdots \\ \mathbf{a}_m \end{bmatrix} \text{或者} \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i - k\mathbf{a}_j \\ \vdots \\ \mathbf{a}_m \end{bmatrix} \text{或者} \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$$

从而我们有：

$$\begin{aligned} \mathbf{C}(A^T) &= \mathrm{span}(\{\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_m\}) \\ &= \mathrm{span}(\{\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_m\}) \\ &= \mathrm{span}(\{\mathbf{a}_1, \ldots, k\mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_m\}) \\ &= \mathrm{span}(\{\mathbf{a}_1, \ldots, \mathbf{a}_i - k\mathbf{a}_j, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_m\}) \end{aligned}$$

即：

$$\mathrm{row{-}rank}(A) = \dim(C(A^T)) = \dim(C(A'^T)) = \mathrm{row{-}rank}(A')$$

$\square$

列秩不改变的证明则略微麻烦一些。

证明： [引理1.17的证明] 我们这里仅考虑**行交换** 的情形，**其他两种情绪请同学们自行验证**。令：

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_i & \cdots & \mathbf{a}_j & \cdots & \mathbf{a}_n \end{bmatrix}$$

交换第 i 行和第 j 行后：

$$A' = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{a}'_1 & \cdots & \mathbf{a}'_i & \cdots & \mathbf{a}'_j & \cdots & \mathbf{a}'_n \end{bmatrix}$$

我们比较一下 $\mathbf{a}_i$ 和 $\mathbf{a}_j$：

$$\mathbf{a}_i = \begin{bmatrix} a_{i1} & \dots & a_{ik} & \dots & a_{jk} & \dots & a_{in} \end{bmatrix}^T$$
$$\mathbf{a}'_i = \begin{bmatrix} a_{i1} & \dots & a_{jk} & \dots & a_{ik} & \dots & a_{in} \end{bmatrix}^T$$

我们需要证明：

$$\dim(\mathrm{span}(\{\mathbf{a}_1, \dots, \mathbf{a}_n\})) = \dim(\mathrm{span}(\{\mathbf{a}'_1, \dots, \mathbf{a}'_n\})).$$

我们可以发现，对于任意的 $t \in [n]$，$1 \leqslant k_1 \leqslant k_2 \leqslant \cdots \leqslant k_t \leqslant n$ 和任意的 $c_1, \dots, c_t \in \mathbb{R}$，我们有：

$$\sum_{l \in [t]} c_l \mathbf{a}_{k_l} = 0 \quad \Leftrightarrow \quad \sum_{l \in [t]} c_l \mathbf{a}'_{k_l} = 0$$

也就是说：

- $\mathbf{a}_{k_1}, \dots \mathbf{a}_{k_t}$ 是线性相关的当且仅当 $\mathbf{a}'_{k_1}, \dots \mathbf{a}'_{k_t}$ 是线性相关的。
- $\mathbf{a}_{k_1}, \dots \mathbf{a}_{k_t}$ 是 $\mathrm{span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$ 的一组基当且仅当 $\mathbf{a}'_{k_1}, \dots \mathbf{a}'_{k_t}$ 是 $\mathrm{span}(\mathbf{a}'_1, \dots, \mathbf{a}'_n)$ 的一组基。

从而：

$$\dim(\mathbf{C}(A)) = \dim(\mathrm{span}(\{\mathbf{a}_1, \dots, \mathbf{a}_n\})) = \dim(\mathrm{span}(\{\mathbf{a}'_1, \dots, \mathbf{a}'_n\})) = \dim(\mathbf{C}(A')).$$

$\square$

注 1.18

可以看到，对于初等行变换来说：

1. 其不会改变矩阵的行空间。
2. 其会改变矩阵的列空间，但是不会改变列空间的维数。

至此，我们完成了定理0.1的证明，

定理0.1的一个技巧性证明 ：我们再给出一个充满技巧性地简洁证明：

我们只需证明对任意的矩阵 A，其有 row-rank(A) $\leqslant$ column-rank(A) 即可。

设 row-rank(A) $= r$，column-rank(A) $= c$。则存在 A 的 c 个列向量 $\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_c}$ 是线性无关的，并且对于 A 中的每个列向量 $\mathbf{a}_j (j \leqslant n)$，存在 $p_{j1}, \dots, p_{jc}$ 使得：

$$\mathbf{a}_j = p_{j1}\mathbf{a}_{i_1} + \cdots + p_{jc}\mathbf{a}_{i_c}$$

从而我们有：

$$A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{i_1} & \cdots & \mathbf{a}_{i_c} \end{bmatrix} \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{c1} & \cdots & p_{cn} \end{bmatrix}$$

记等式最右面的矩阵为 P,P 是一个 $c \times n$ 的矩阵，则有：

$$\text{row-rank}(A) \leqslant row-rank(P) \leqslant c = \text{column-rank}(A)$$

□

2 $Ax = 0$ 的解

这一节开始，我们来正式讨论方程组的解的情况。首先来关注 $A\mathbf{x} = \mathbf{0}$，回顾 A 的零空间 $\mathbf{N}(A)$ 的定义：

定义 (零空间 (Null Space))

给定一个 $m \times n$ 的矩阵 A，定义其零空间 $\mathbf{N}(A)$ 为：

$$\mathbf{N}(A) = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{0}\}$$

即 $\mathbf{N}(A)$ 是所有满足 $A\mathbf{x} = \mathbf{0}$ 的 $\mathbf{x}$ 的集合。

显然 $\mathbf{N}(A)$ 便是 $A\mathbf{x} = \mathbf{0}$ 的解的全体集合，注意到 $\mathbf{N}(A)$ 是 $\mathbb{R}^n$ 的一个子空间，所以 $A\mathbf{x} = \mathbf{0}$ 的求解实际上等价于求出 $\mathbf{N}(A)$ 的一组基。首先我们关注一个特殊情况：

引理 2.1

令 A 是一个 $m \times n$ 的矩阵，则 $\dim(\mathbf{C}(A)) = n$ 当且仅当 $\dim(\mathbf{N}(A)) = 0$。

注 2.2

事实上，这是线性代数基本定理的一个特殊情况：

$$\dim(\mathbf{C}(A)) + \dim(\mathbf{N}(A)) = n$$

我们在后续的内容中会详细讨论这个定理。

证明： 令

$$A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$$

则我们有：

$$\begin{aligned}
\dim(\mathbf{C}(A)) = \dim(\text{span}(\{\mathbf{a}_1, \dots, \mathbf{a}_n\})) = n \\
&\Leftrightarrow \mathbf{a}_1, \dots, \mathbf{a}_n \text{是 } \mathbf{C}(A) \text{ 的一组基} \\
&\Leftrightarrow \mathbf{a}_1, \dots, \mathbf{a}_n \text{线性无关} \\
&\Leftrightarrow \mathbf{N}(A) = \{\mathbf{0}\} \\
&\Leftrightarrow \dim(\mathbf{N}(A)) = 0
\end{aligned}$$

□

引理2.1实际上描述了 $A\mathbf{x} = \mathbf{0}$ 的一种特殊情况，即当 $\mathbf{C}(A)$ 的维数等于 n 时，$\mathbf{N}(A)$ 的维数为 0，此时对任意的 $\mathbf{b} \in \mathbf{C}(A)$，$A\mathbf{x} = \mathbf{b}$ 都只有唯一解，自然 $A\mathbf{x} = \mathbf{0}$ 只有平凡解 $\mathbf{0}$。下面我们在看几个例子。

例 2.3（第一个例子）

考察如下的方程组：

$$\begin{aligned} x + 2y = 0 \\ 2x + 4y = 0 \end{aligned} (\text{记作}A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}) \implies \begin{aligned} x + 2y = 0 \\ 0 = 0 \end{aligned}$$

方程组只有一个首元，我们称剩下对应的变元是自由变量 (free variable)，这里 y 是自由变量，注意到任取一个 y 的值，我们都可以构造该方程组的一个解，特别的，**我们令** $y = 1$，此时产生的解 $\mathbf{s}$：

$$\mathbf{s} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

称之为特解 (special solution)，不难验证：

$$\mathbf{N}(A) = \text{span}(\{\mathbf{s}\}).$$

例 2.4（第二个例子）

再考虑如下的方程组：

$$x+2y+3z=0,\quad \text{i.e.}\ \begin{bmatrix}1&2&3\end{bmatrix}(\text{记作}A)\begin{bmatrix}x\\y\\z\end{bmatrix}=0$$

这个方程组中，y,z 都是自由变量，**分别令 $y=1$ 和 $z=1$** 可得两个特解 $\mathbf{s}_1,\mathbf{s}_2$：

$$\mathbf{s}_1=\begin{bmatrix}-2\\1\\0\end{bmatrix},\quad \mathbf{s}_2=\begin{bmatrix}-3\\0\\1\end{bmatrix}$$

同样我们有：

$$\mathbf{N}(A)=\text{span}(\{\mathbf{s}_1,\mathbf{s}_2\}).$$

上述例子给了我们求解 $A\boldsymbol{x}=\boldsymbol{0}$ 的思路，即：

1. 将 A 变成行阶梯形 U 或者行最简形 R。
2. 找到 U 或者 R 中的自由变量，分别令其为 1 求得得到特解。
3. 将特解组合起来，得到 $\mathbf{N}(A)$ 的一组基。

特别的，我们称：

1. 首元 (pivot) 所在的列为首元列 (pivot column)，对应的变量 x_1,x_2 称为主变量 (pivot variables).
2. 剩余列则称为自由列 (free columns)，对应的变量 x_3,x_4 称为自由变量 (free variables).

下面则是一个帮助理解的例子：

$$C\boldsymbol{x}=\begin{bmatrix}1&2&2&4\\3&8&6&16\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}\implies\begin{bmatrix}1&2&2&4\\0&2&0&4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

在其行阶梯形中，用红色来表示首元列，蓝色表示自由列：

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

到此为止，似乎**行阶梯形**已经够我们处理 $A\boldsymbol{x}=\boldsymbol{0}$ 了，再让我们再来看下**行最简形** 的好处。我们继续将上述的矩阵变成行最简形：

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

可以发现：

1. 所有的首元列构成了一个 $r \times r$ 的单位矩阵，这里 r 是首元的个数。
2. 每列自由列都表明了该方程组令该列对应的变元为 1 的一组特解，即：

$$\mathbf{s}_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{s}_2 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

我们再用一个例子来具体描述一下：令：

$$A = \begin{bmatrix} \color{red}{p} & \color{red}{p} & f & \color{red}{p} & f \\ | & | & | & | & | \\ | & | & | & | & | \\ | & | & | & | & | \end{bmatrix} \Longrightarrow R = \begin{bmatrix} 1 & 0 & a & 0 & c \\ 0 & 1 & b & 0 & d \\ 0 & 0 & 0 & 1 & e \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3 个首元列 p　　　　$\text{rank}(R) = \text{rank}(A) = r = 3$

2 个自由列 f

从而 $A\boldsymbol{x} = \boldsymbol{0}$ 等价于：

$$\begin{bmatrix} 1 & 0 & \color{red}{a} & 0 & \color{blue}{c} \\ 0 & 1 & \color{red}{b} & 0 & \color{blue}{d} \\ 0 & 0 & \color{red}{0} & 1 & \color{blue}{e} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \boldsymbol{0}, \quad \text{i.e.} \quad \begin{aligned} x_1 + ax_3 + cx_5 &= 0 \\ x_2 + bx_3 + dx_5 &= 0 \\ x_4 + ex_5 &= 0 \end{aligned}$$

我们有：

首元	自由变量
x_1, x_2, x_4	x_3, x_5

也即，x_3, x_5 是可以任意选择的，而 x_1, x_2, x_4 则由 x_3, x_5 决定：

$$\begin{aligned} x_1 &= -ax_3 - cx_5, \\ x_2 &= -bx_3 - dx_5, \\ x_4 &= -ex_5 \end{aligned}$$

特别的，分别令 x_3 和 x_5 为 1 我们可以得到如下两组特解：

$$\mathbf{s}_1 = \begin{bmatrix} -\color{red}{a} \\ -\color{red}{b} \\ 1 \\ \color{red}{0} \\ 0 \end{bmatrix}, \quad \mathbf{s}_2 = \begin{bmatrix} -\color{blue}{c} \\ -\color{blue}{d} \\ 0 \\ -\color{blue}{e} \\ 1 \end{bmatrix}$$

也就是说，行最简形上自由列的每一列其实都<u>**直接反映出了一个特解上首元的取值！**</u> 同时我们有：

$$\mathbf{N}(A) = \mathbf{N}(R) = \text{span}(\{\mathbf{s}_1, \mathbf{s}_2\}).$$

$Ax=0$ 的一般描述. 现在我们给出 $Ax=0$ 的解的一般描述: 令 A 是 $m\times n$ 的矩阵, 我们考虑 $Ax=0$ 的解。我们先使用高斯若尔当消元法将 A 转化成行最简形 R, 即:

$$Ax=0 \quad \Longleftrightarrow \quad \begin{bmatrix} 0 & \cdots & b_{1j_1} & \cdots & b_{1n} \\ 0 & \cdots & 0 & \cdots & b_{2n} \\ 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{j1} \\ \vdots \\ x_n \end{bmatrix} = 0$$

其 $\text{rank}(A)=r$ 意味着存在 r 个首元:

$$b_{1j_1} = b_{2j_2} = \cdots = b_{rj_r} = 1$$

也就是

首元	自由变量
$x_{j_1}, x_{j_2}, \cdots, x_{j_r}$	$x_1, \cdots, x_{j_1-1}, x_{j_1+1}, \cdots, x_{j_2-1}, \cdots, x_{j_r+1}, \cdots, x_n$

$Rx=0$ 对应的方程组为:

$$\begin{aligned} x_{j_1} + b_{1,j_1+1}x_{j_1+1} + \cdots + b_{1,j_2-1}x_{j_2-1} + b_{1,j_2+1}x_{j_2+1} + \cdots + b_{1n}x_n &= 0 \\ x_{j_2} + b_{2,j_2+1}x_{j_2+1} + \cdots + b_{2n}x_n &= 0 \\ &\vdots \\ x_{j_r} + \cdots + b_{rn}x_n &= 0 \end{aligned}$$

从而我们可以构造出$n-r$ 个特殊解:

$$\mathbf{s}_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \mathbf{s}_{j_1-1} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \mathbf{s}_{j_1+1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -b_{1,j_1+1} \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots,$$

这意味着:

$$\mathbf{N}(A) = \mathbf{N}(R) = \text{span}(\{\mathbf{s}_1, \cdots, \mathbf{s}_{j_1-1}, \mathbf{s}_{j_1+1}, \cdots, \mathbf{s}_{j_2-1}, \cdots, \mathbf{s}_{j_r+1}, \cdots, \mathbf{s}_n\}).$$

即:

$$\dim(\mathbf{N}(A)) + \text{rank}(A) = n.$$

这就是**线性代数基本定理的第一部分**:

定理 2.5 (Fundamental Theorem of Linear Algebra, Part I)

令 A 是一个 $m \times n$ 的矩阵并且 rank$(A) = r$，则:

1. $\dim(\mathbf{C}(A)) = \dim(\mathbf{C}(A^T)) = r$。
2. $\dim(\mathbf{N}(A)) = n - r$, $\dim(\mathbf{N}(A^T)) = m - r$。

3 $Ax = \mathbf{b}$ 的解

最后我们来讨论对任意的 $\mathbf{b}$，$Ax = \mathbf{b}$ 的解的情况。我们需要搞清如下的三个问题:

1. 什么时候 $Ax = \mathbf{b}$ 有解?
2. $Ax = \mathbf{b}$ 的解的结构是怎么样的?
3. 如何计算 $Ax = \mathbf{b}$ 的解?

事实上，当我们搞清楚 $Ax = \mathbf{0}$ 的解的情况时，上述问题都会有很清晰的解答。

第一个问题。 注意到 Ax 实际上是 A 的列向量的线性组合，从而:

定理 3.1

$Ax = \mathbf{b}$ 有解当且仅当 $\mathbf{b} \in \mathbf{C}(A)$

上述定理也等价于下列的表述方式:

定理 3.2

$Ax = \mathbf{b}$ 有解当且仅当

$$\text{rank}(A) = \text{rank}(\begin{bmatrix} A & \mathbf{b} \end{bmatrix}).$$

证明: 将 A 写成列向量的形式，即:

$$A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$$

则我们有:

$$\begin{aligned}
& Ax = \mathbf{b} \text{ 有解} \\
& \iff \mathbf{b} \in \mathbf{C}(A) = \text{span}(\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}) \\
& \iff \text{span}(\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}) = \text{span}(\{\mathbf{a}_1, \cdots, \mathbf{a}_n, \mathbf{b}\}) \\
& \iff \dim(\text{span}(\{\mathbf{a}_1, \cdots, \mathbf{a}_n\})) = \dim(\text{span}(\{\mathbf{a}_1, \cdots, \mathbf{a}_n, \mathbf{b}\})) \\
& \iff \text{column-rank}(A) = \text{column-rank}(\begin{bmatrix} A & \mathbf{b} \end{bmatrix}) \\
& \iff \text{rank}(A) = \text{rank}(\begin{bmatrix} A & \mathbf{b} \end{bmatrix})
\end{aligned}$$

□

第二个问题。 假设 x_{p_1}, x_{p_2} 是 $Ax = b$ 的两个解，则:

$$Ax_{p_1} - Ax_{p_2} = 0 \implies A(x_{p_1} - x_{p_2}) = 0$$

即：$x_{p_1} - x_{p_2}$ 是 $Ax = 0$ 的一个解。从而我们有:

定理 3.3

令 A 是一个 $m \times n$ 的矩阵，$x_p \in \mathbb{R}^b$ 是方程 $Ax = b$ 的一个解，则对任意的 $x \in \mathbb{R}^n$:

$$Ax = b \iff x - x_p \in \mathbf{N}(A).$$

定理3.3告诉我们，$Ax = b$ 的解可以分解成如下的形式:

$$Ax = b\text{的一个解}x_p + Ax = 0\text{的解}x$$

特别的，我们称x_p **为 $Ax = b$ 的一个特解**。从而令

$$s_1, \dots, s_l$$

是 $Ax = 0$ 的一组特殊解，即其零空间 $\mathbf{N}(A)$ 的一组基，从而 $l = n - \text{rank}(A)$。则任何一个 $Ax = b$ 的解都可以表示为:

$$x = x_p + c_1 s_1 + \cdots + c_l s_l.$$

注意到 A 的秩 r 应该满足:

$$r = \text{rank}(A) = \text{column-rank}(A) = \text{row-rank}(A) \leqslant \min\{m, n\}$$

从而我们有:

m	n	$\dim(\mathbf{N}(A))$	$Ax = b$ 的解的个数
$= r$	$= r$	0	1
$= r$	$> r$	$\geqslant 1$	∞
$> r$	$= r$	0	0 or 1
$> r$	$> r$	$\geqslant 1$	0 or ∞

表 1　不同情况下 $Ax = b$ 的解的情况

第三个问题。 由前面的讨论可知，我们可以通过**高斯若尔当消元法**来求得 $Ax = b$ 的解:

1. 对增广矩阵 $\begin{bmatrix} A & b \end{bmatrix}$ 使用高斯若尔当消元法，将其变成行最简形 $\begin{bmatrix} R & c \end{bmatrix}$。
2. 根据 $\begin{bmatrix} R & c \end{bmatrix}$ 令所有的自由变量为 0，求得 $Ax = b$ 的一组特解 x_p，如果不存在特解，则 $Ax = b$ 无解。
3. 根据 R 求得 $Ax = 0$ 的一组基 $s_1, \dots, s_l$。

4. 最后 $A\boldsymbol{x}=\mathbf{b}$ 的解可以表示为:

$$\boldsymbol{x}=\boldsymbol{x}_p+c_1\mathbf{s}_1+\cdots+c_l\mathbf{s}_l,$$

这里 $c_1,\ldots,c_l\in\mathbb{R}$.

最后我们再通过一个例子来理解一下上述过程:

例 3.4

我们来求解下列的方程组 $A\boldsymbol{x}=\mathbf{b}$，其中:

$$A=\begin{bmatrix}1&2&3&5\\2&4&8&12\\3&6&7&13\end{bmatrix},\mathbf{b}=\begin{bmatrix}0\\6\\-6\end{bmatrix}\quad 即:\begin{cases}x_1+2x_2+3x_3+5x_4=0\\2x_1+4x_2+8x_3+12x_4=6\\3x_1+6x_2+7x_3+13x_4=-6\end{cases}$$

解: 对 $\begin{bmatrix}A&\mathbf{b}\end{bmatrix}$ 使用高斯若尔当消元法变成行最简形得:

$$\begin{bmatrix}1&2&3&5&0\\2&4&8&12&6\\3&6&7&13&-6\end{bmatrix}\Longrightarrow\begin{bmatrix}1&2&0&2&-9\\0&0&1&1&3\\0&0&0&0&0\end{bmatrix}(记为\begin{bmatrix}R&\mathbf{c}\end{bmatrix})$$

1. x_1,x_3 是首元，令自由变元 $x_2=x_4=0$ 可求得 $R\boldsymbol{x}=\mathbf{c}$(即 $A\boldsymbol{x}=\mathbf{b}$) 的一个解 $\boldsymbol{x}_p$:

$$\boldsymbol{x}_p=\begin{bmatrix}-9\\0\\3\\0\end{bmatrix}$$

2. 考察方程组 $R\boldsymbol{x}=\mathbf{0}$，分别令 $x_2=1$ 和 $x_4=1$ 可以求得其解得一组基 $\mathbf{s}_1,\mathbf{s}_2$:

$$\mathbf{s}_1=\begin{bmatrix}-2\\1\\0\\0\end{bmatrix},\quad\mathbf{s}_2=\begin{bmatrix}-2\\0\\-1\\1\end{bmatrix}$$

3. 从而 $A\boldsymbol{x}=\mathbf{b}$ 的解可以表示为 $\boldsymbol{x}$:

$$\boldsymbol{x}=\boldsymbol{x}_p+c_1\mathbf{s}_1+c_2\mathbf{s}_2=\begin{bmatrix}-9\\0\\3\\0\end{bmatrix}+c_1\begin{bmatrix}-2\\1\\0\\0\end{bmatrix}+c_2\begin{bmatrix}-2\\0\\-1\\1\end{bmatrix}=\begin{bmatrix}-9-2c_1-2c_2\\c_1\\3-c_2\\c_2\end{bmatrix}$$

□