《线性代数》

7-解线性方程组 (II)(Solving Linear Equations(II))

杨启哲

上海师范大学信机学院计算机系

2024 年 4 月 12 日

复习：矩阵的列秩与行秩 (I)

定义 1 **[列秩 (Column Rank)].**

给定一个 $m \times n$ 的矩阵 A，其列秩 (Column Rank) 定义为：

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

定义 2 **[行秩 (Row Rank)].**

矩阵 A 的行秩 (Row Rank) 定义为：

$$\text{row}-\text{rank}(A) = \dim(C(A^{\mathsf{T}}))$$

引理 3.

下面的叙述是等价的：

1. A 的行秩是 n.
2. A^T 的列秩是 n.
3. 存在一个矩阵 B，使得 $A^TB = I$。
4. 存在一个矩阵 B，使得 $BA = I$

定理 4.

给定一个 $m \times n$ 的矩阵，我们有 $\text{row–rank}(A) = \text{column–rank}(A)$.

复习：高斯若尔当消元法

1. 考察一个矩阵：

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

2. 注意到对于任意 $b \in \mathbb{R}^3$，方程 $Ax = b$ 都有解，这是因为通过高斯消元法可以发现 A 有3 个首元。从而我们通过高斯-若尔当消元法可以解出：

$$DE_{12}E_{13}E_{23}P_{23}E_{31}\begin{bmatrix} A & e_1 & e_2 & e_3 \end{bmatrix} = \begin{bmatrix} I & x_1 & x_2 & x_3 \end{bmatrix}$$

即存在 $B = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$ 使得 $AB = I$.

3. 另一方面，如果令 $C = DE_{12}E_{13}E_{23}P_{23}E_{31}$，我们也有 $CA = I$.
4. 最终我们可以发现 A 是可逆的，并且：

$$A^{-1} = B = C = DE_{12}E_{13}E_{23}P_{23}E_{31}$$

引理 5.

令 A 是一个 $n \times n$ 的矩阵, 下面的叙述是等价的:

1. A 是可逆的。
2. 方程 $Ax = b$ 对任意 $b \in \mathbb{R}^m$ 都有唯一解.
3. $\dim(C(A)) = n$.
4. $\dim(C(A^T)) = n$.
5. 存在矩阵 B 使得 $AB = I$。
6. 存在矩阵 C 使得 $CA = I$。
7. A 有 n 个首元。

主要内容

- 矩阵的秩
- $Ax = \mathbf{0}$ 的解
- $Ax = b$ 的解

矩阵的秩

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

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_{ji} 是第 j 行中最左边的不为 0 的系数。

行阶梯形 (Row Echelon Form)

这个方程组，或者说对应的系数矩阵，是行阶梯形的，如果存在 $0 \leqslant r \leqslant m$ 使得：

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

这也意味着该方程组具有 r 个首元 $a_{1j_1}, \ldots, a_{rj_r}$。

行阶梯形的例子

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 &= 5 \\ 1x_2 + 2x_3 &= 3 \\ 1x_3 &= 2 \end{aligned} \quad \Leftrightarrow \quad \begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}$$

$$\begin{aligned} 4x_1 - 13x_2 + 2x_4 &= 0 \\ 11x_2 + x_4 &= 6 \end{aligned} \quad \Leftrightarrow \quad \begin{bmatrix} 4 & -13 & 0 & 2 \\ 0 & 11 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$$

$$\begin{aligned} 4x_1 - 13x_2 + 2x_4 &= 0 \\ 11x_2 + x_4 &= 6 \\ 0 &= 7 \end{aligned} \quad \Leftrightarrow \quad \begin{bmatrix} 4 & -13 & 0 & 2 \\ 0 & 11 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ 7 \end{bmatrix}$$

高斯消元法的精确描述 (I)

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 1 步

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. 选择最小的 $i \in [m]$ 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

此时矩阵对于 $r=i$，R1 和 R2 都是满足的，我们继续进行下一步:

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. 选择最小的 $i'\in[i+1,m]$ 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。

一个例子 (I)

考察这样一个矩阵：

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

第一步的具体过程

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

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

上海师范大学 Shanghai Normal University

第二步的具体过程

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

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

一个例子 (III)

第三步的具体过程

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

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

此时我们有:

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

引理 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。

我们利用首元的个数来定义矩阵的秩。

定义 7 [矩阵的秩 (Rank)].

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

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

定理 8.

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

行最简形

为了证明定理8，我们考虑行最简形矩阵。

行最简形 (Reduced Row Echelon Form)

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。

行最简形的例子

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 & 3 & 4 & 0 \\ 0 & 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

定理 9.

高斯若尔当消元法会将矩阵 A 变成一个行最简形矩阵 R。

定理 10.

我们有:

1. $\mathrm{rank}(A) = \mathrm{rank}(R)$.
2. $\mathrm{rank}(R) = \mathrm{column\text{-}rank}(R) = \mathrm{row\text{-}rank}(R)$.

行变换

我们知道高斯若尔当消元法是通过一系列的行变换来实现的，其中一共有三种行变换:

1. 行加法（Row Addition）。
2. 行交换（Row Exchange）。
3. 行乘法（Row Multiplication）。

我们称其为初等行变换 (Row Elementary Operations)。

回顾：行加法

上海师范大学 Shanghai Normal University

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

引理 11.

$E_{ij}^{-1}(-k)=E_{ij}(k)$

回顾：行交换

上海师范大学 Shanghai Normal University

$$P_{ij}A = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix}$$

引理 12.

$$P_{ij}^{-1} = P_{ij}$$

回顾：行乘法

$$D_i(k)A = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & k & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ ka_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{bmatrix}$$

引理 13.

$D_i(k)^{-1} = D_i(\frac{1}{k})$

定理 14.

初等行变换不改变矩阵的行秩和列秩。

证明. 我们先来证明行变换不改变矩阵的行秩。事实上，令 $A = \begin{bmatrix} a_1^T & a_i^T & \cdots & a_m^T \end{bmatrix}^T$，即令其写成行向量的形式，这里 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：

$$A' = \begin{bmatrix} a_1' \\ \vdots \\ a_i' \\ \vdots \\ a_m' \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ ka_i \\ \vdots \\ a_m \end{bmatrix} \text{或者} \begin{bmatrix} a_1 \\ \vdots \\ a_i - ka_j \\ \vdots \\ a_m \end{bmatrix} \text{或者} \begin{bmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_m \end{bmatrix}$$

行变换的性质 (II)

$$A' = \begin{bmatrix} a'_1 \\ \vdots \\ a'_i \\ \vdots \\ a'_m \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ ka_i \\ \vdots \\ a_m \end{bmatrix} \text{或者} \begin{bmatrix} a_1 \\ \vdots \\ a_i - ka_j \\ \vdots \\ a_m \end{bmatrix} \text{或者} \begin{bmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_m \end{bmatrix}$$

我们有:

$$\begin{aligned} C(A^T) &= \operatorname{span}(\{a_1, \ldots, a_i, \ldots, a_j, \ldots, a_m\}) \\ &= \operatorname{span}(\{a_1, \ldots, a_j, \ldots, a_i, \ldots, a_m\}) \\ &= \operatorname{span}(\{a_1, \ldots, ka_i, \ldots, a_j, \ldots, a_m\}) \\ &= \operatorname{span}(\{a_1, \ldots, a_i - ka_j, \ldots, a_j, \ldots, a_m\}) \end{aligned}$$

即:

$$\text{row—rank}(A) = \dim(C(A^T)) = \dim(C(A'^T)) = \text{row—rank}(A')$$

我们现在来考虑列秩。

- 事实上，由于我们已经证明列秩和行秩是相等的，所以我们已经可以得到列秩不改变的结论。
- 但我们依旧想直接证明一下。我们只在这里考虑列交换的情况，其他的情况大家可以自行练习。

我们将 A 写成列向量的形式：

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

行变换的性质 (IV)

交换第 i 行和第 j 行后：

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

- 我们比较一下 a_i 和 a_j：

$$a_i = \begin{bmatrix} a_{i1} & \dots & a_{ik} & \dots & a_{jk} & \dots & a_{in} \end{bmatrix}^T$$

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

我们现在需要证明:

$$\dim(\mathrm{span}(\{a_1,\ldots,a_n\})) = \dim(\mathrm{span}(\{a'_1,\ldots,a'_n\}))$$

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 $c_1,\ldots,c_t \in \mathbb{R}$, 我们有:

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

也就是说:

- $a_{k_1},\ldots a_{k_t}$ 是线性相关的当且仅当 $a'_{k_1},\ldots a'_{k_t}$ 是线性相关的。
- $a_{k_1},\ldots a_{k_t}$ 是 $\mathrm{span}(a_1,\ldots,a_n)$ 的一组基当且仅当 $a'_{k_1},\ldots a'_{k_t}$ 是 $\mathrm{span}(a'_1,\ldots,a'_n)$ 的一组基。

从而:

$$\dim(C(A)) = \dim(\mathrm{span}(\{a_1,\ldots,a_n\})) = \dim(\mathrm{span}(\{a'_1,\ldots,a'_n\})) = \dim(C(A'))$$

$\square$

定理8的证明

让我们回到定理8。

定理 8.

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

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 $\mathrm{rank}(A) = r$。
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. $\mathrm{rank}(R) = \mathrm{column{-}rank}(R) = \mathrm{row{-}rank}(R) = r = \mathrm{rank}(A)$.
4. 注意到我们只使用了初等行变换，从而：

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

即：

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

上海师范大学 Shanghai Normal University

- 矩阵的秩 $\mathrm{rank}(A) =$ 矩阵的首元个数。
- $\mathrm{rank}(A) = \mathrm{row{-}rank}(A) = \mathrm{column{-}rank}(A)$.

引理 9.

令 A 是一个 $n \times n$ 的矩阵, 下面的叙述是等价的:

1. A 是可逆的。
2. 方程 $Ax = b$ 对任意 $b \in \mathbb{R}^m$ 都有唯一解.
3. $\mathrm{rank}(A) = n$.
4. $\mathrm{column{-}rank}(A) = n$.
5. $\mathrm{row{-}rank}(A) = n$.
6. 存在矩阵 B 使得 $AB = I$。
7. 存在矩阵 C 使得 $CA = I$。

当 A 是可逆矩阵的时候，我们已经足够清楚 $Ax = b$ 的解了。

问题 10.

那对于任意的 A，比如 A 不是可逆的，或者说 A 不是方阵的情况那?

$A\mathbf{x} = \mathbf{0}$ 的解

定义 11 **[零空间 (Null Space)].**

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

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

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

定理 12.

零空间 $N(A)$ 是 $\mathbb{R}^n$ 的一个子空间。

引理 13.

$N(A) \subseteq \mathbb{R}^n$，从而 $\dim(N(A)) \leqslant n$.

引理 14.

$\dim(C(A)) = n$ 当且仅当 $\dim(N(A)) = 0$。

说明

事实上，这是线性代数基本定理的一个特殊情况:

$$\dim(C(A)) + \dim(N(A)) = n$$

证明. 令

$$A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$$

则我们有:

$$\begin{aligned} \dim(C(A)) = \dim(\mathrm{span}(\{a_1, \ldots, a_n\})) = n \\ \Leftrightarrow a_1, \ldots, a_n \text{是 } C(A) \text{ 的一组基} \\ \Leftrightarrow a_1, \ldots, a_n \text{线性无关} \\ \Leftrightarrow N(A) = Z \\ \Leftrightarrow \dim(N(A)) = 0 \end{aligned}$$

$\square$

问题 15.

一般情况下怎么去计算 $N(A)$ 的一组基。特别的 $\dim(N(A))$?

一些例子 (I)

$$\begin{aligned} x + 2y &= 0 \\ 2x + 4y &= 0 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} x + 2y &= 0 \\ 0 &= 0 \end{aligned}$$

我们称 y 是自由的 (free)。所以:

$$N(A) = \{(-2y, y) \mid y \in \mathbb{R}\}$$

我们称下列的解是特解 (special solution):

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

- s 是将 y 设为 1 得到的解。
- $N(A)$ 由所有 s 的线性组合构成，即:

$$N(A) = span(\{s\})$$

一些例子 (II)

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

在这个例子中 y 和 z 都是自由变量。所以我们可以选择两个特解:

$$s_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad s_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

同样我们有:

$$N(A) = \text{span}(\{s_1, s_2\})$$

求零空间的思路

1. 将矩阵 A 变成行阶梯形矩阵 U(Row echelon form)。
2. 寻找到 U 的特解。

例 16.

让我们从这几个例子再思考一下。

$$A=\begin{bmatrix}1&2\\3&8\end{bmatrix},\quad B=\begin{bmatrix}A\\2A\end{bmatrix}=\begin{bmatrix}1&2\\3&8\\2&4\\6&16\end{bmatrix},\quad C=\begin{bmatrix}A&2A\end{bmatrix}=\begin{bmatrix}1&2&2&4\\3&8&6&16\end{bmatrix}$$

矩阵 A 和矩阵 B

矩阵 A

$$Ax = \begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

从而我们有:

$$N(A) = Z = \{\mathbf{0}\}$$

矩阵 B

$$Bx = \begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 4 \\ 6 & 16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \implies \begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \implies \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

从而我们有:

$$N(B) = Z = \{\mathbf{0}\}$$

上海师范大学 Shanghai Normal University

$$Cx = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 3 & 8 & 6 & 16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad \begin{bmatrix} \color{red}{1} & 2 & 2 & 4 \\ 0 & \color{red}{2} & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

这里存在两个自由变量:

$$\{x_3, x_4\}$$

我们从而可以选择两个特解:

$$s_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad s_2 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

使得:

$$N(C) = span(\{s_1, s_2\})$$

首元列和自由列

$$Cx = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 3 & 8 & 6 & 16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 2 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

在其行阶梯形中:

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

- 我们称首元 (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)
- 剩余的列则称为自由列 (free columns)，对应的变量 x_3, x_4 称为自由变量 (free variables)

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

在行最简形 (Reduced Row Echelon Form) 下，所有的主元列构成了一个 $r \times r$ 的单位矩阵，其中 $r = \mathrm{rank}(A)$.

问题 17.

通过不同的初等行变换，我们是否可以得到不同的首元列和自由列?

行最简形下的首元列视角 (II)

上海师范大学 Shanghai Normal University

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

3 个首元列 p

2 个自由列 f

$\operatorname{rank} r = 3$

从而 $Ax = \mathbf{0}$ 等价于:

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

主变量和自由变量

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

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:

$$x_1 = -ax_3 - cx_5,$$

$$x_2 = -bx_3 - dx_5,$$

$$x_4 = -ex_5$$

$$x_1 = -ax_3 - cx_5,$$
$$x_2 = -bx_3 - dx_5,$$
$$x_4 = -ex_5$$

通过选择自由解 $(x_3, x_5) \in \{(1,0),(0,1)\}$ 我们得到了两个特殊解:

$$s_1 = \begin{bmatrix} -a \\ -b \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad s_2 = \begin{bmatrix} -c \\ -d \\ 0 \\ -e \\ 1 \end{bmatrix}$$

并且所有 $Ax = \mathbf{0}$ 的解都可以表示成 s_1, s_2 的线性组合（为什么？），即:

$$N(A) = N(R) = span(\{s_1, s_2\})$$

关于 $Ax = \mathbf{0}$ 一般的描述 (I)

一般来说，令 A 是 $m \times n$ 的矩阵，我们考虑 $Ax = \mathbf{0}$ 的解。我们先使用 Gauss–Jordan 消元法将 A 转化成行最简形 R，即：

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

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:

$$
s_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, s_{j_1-1} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, s_{j_1+1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -b_{1,j_1+1} \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots,
$$

这意味着:

$$N(A) = N(R) = \mathrm{span}(\{s_1, \cdots, s_{j_1-1}, s_{j_1+1}, \cdots, s_{j_2-1}, \cdots, s_{j_r+1}, \cdots, s_n\})$$

定理 18.

对于任意的 $m \times n$ 的矩阵 A，我们有:

$$\mathrm{rank}(A) + \dim(N(A)) = n$$

定理 19 [Fundamental Theorem of Linear Algebra, Part I].

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

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

阶段总结

- 零空间的基的计算。
- 线性代数的基本定理，第一部分

所以我们目前对于 $A\mathbf{x} = \mathbf{0}$ 的解已经有了一个比较清晰的认识。

问题 20.

那对于 $A\mathbf{x} = \mathbf{b}$ 的解呢？

$\mathrm{Ax} = \mathrm{b}$ 的解

三个问题

1. 什么时候 $Ax = b$ 有解?
2. 如果 $Ax = b$ 有解，其解的结构是什么?
3. 怎么计算 $Ax = b$ 的解?

回顾

当 $b = 0$ 的时候，我们已经有了一个比较清晰的认识。其解的结构就是 $N(A)$，一个维度为 $n - \mathrm{rank}(A)$ 的 $\mathbb{R}^n$ 的子空间。

第一个问题 (I)

定理 21.

$Ax = b$ 有解当且仅当 $b \in C(A)$

定理 22.

$Ax = b$ 有解当且仅当

$$\operatorname{rank}(A) = \operatorname{rank}(\begin{bmatrix} A & b \end{bmatrix})$$

Shanghai Normal University

证明. 将 A 写成列向量的形式，即：

$$A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$$

则我们有：

$$\begin{aligned} & Ax = b \text{ 有解} \\ & \iff b \in C(A) = \mathrm{span}(\{a_1, \cdots, a_n\}) \\ & \iff \mathrm{span}(\{a_1, \cdots, a_n\}) = \mathrm{span}(\{a_1, \cdots, a_n, b\}) \\ & \iff \dim(\mathrm{span}(\{a_1, \cdots, a_n\})) = \dim(\mathrm{span}(\{a_1, \cdots, a_n, b\})) \\ & \iff \text{column}-\text{rank}(A) = \text{column}-\text{rank}(\begin{bmatrix} A & b \end{bmatrix}) \\ & \iff \mathrm{rank}(A) = \mathrm{rank}(\begin{bmatrix} A & b \end{bmatrix}) \end{aligned}$$

□

现在我们来考察 $Ax = b$ 的解的结构，假设 x_p 是其一个解，即 $Ax_p = b$，我们也称其为特解 (particular solution)。

定理 23.

$$Ax = b \quad \Longleftrightarrow \quad x - x_p \in N(A)$$

证明. 只需注意到:

$$A(x - x_p) = Ax - Ax_p = b - b = \mathbf{0}$$

□

事实上，令 x_p 为 $Ax = b$ 的任一特解，令：

$$s_1, \ldots, s_l$$

是 $Ax = \mathbf{0}$ 的一组特殊解，即其零空间 $N(A)$ 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：

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

即一个特解 + 一个齐次解 ($Ax = \mathbf{0}$) 的形式。

$Ax = b$ 的通解 (II)

注意到 A 的秩 r 应当满足:

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

从而我们有:

m	n	$\dim(N(A))$	$Ax = b$ 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 ∞

怎么计算 $Ax = b$ 的解

对其增广矩阵使用 Gauss–Jordan:

$$\begin{bmatrix} A & b \end{bmatrix}$$