第 7 课 求解 $Ax=0$：主变量，特解

> 本章主要讲解**如何求解** $Ax=0$。其中涉及到了诸多步骤和概念，包括**消元（从中可以看出一些线性相关的信息）、秩、主列、主变量、自由列、自由变量以及特解**。注意到，在整个课程的后半部分都在探讨 RREF，这有助于我们更加多角度深入地了解 $Ax=0$ 的求解，当然，**RREF 在求解零空间时也有其重要的用途**。

首先我们给出一个矩阵：

$$A=\begin{bmatrix}1&2&2&2\\2&4&6&8\\3&6&8&10\end{bmatrix}$$

观察该矩阵的行和列，容易发现，列二是列一的倍数，而行三也等于行一加上行二，**这些线性相关的情况都会在消元的过程中表现出来。**

- **在消元的过程中，零空间不会改变。**这一点非常重要。**零空间实际上就是由满足 $Ax=0$ 的所有 x 所组成。在消元过程中，解是不会改变的，因此零空间也不会改变。实际上，消元过程改变的是列空间（行变换改变列空间）。**
  - 行变换改变的是列空间这一点应该不难理解。**随着行变换的进行，A 的列向量本身发生了改变，导致 A 的各个列向量的线性组合也发生了变化，从而 A 的列空间发生了变化。**我们就拿上面给出的矩阵 A 来具体说明：

    $$\text{对矩阵}A\text{进行消元，其中}\color{red}{\text{红色}}\text{标记的为主元}$$
    $$\begin{bmatrix}\color{red}{1}&2&2&2\\2&4&6&8\\3&6&8&10\end{bmatrix}\xrightarrow[row3=row3-3\times row1]{row2=row2-2\times row1}\begin{bmatrix}\color{red}{1}&2&2&2\\0&0&\color{red}{2}&4\\0&0&2&4\end{bmatrix}\xrightarrow{row3=row3-row2}\begin{bmatrix}\color{red}{1}&2&2&2\\0&0&\color{red}{2}&4\\0&0&0&0\end{bmatrix}$$

    首先，**矩阵 A 的列空间必定包括其列向量**，而注意到，**消元后得到的 U 中其所有列向量的第三行都为 0，显然 U 的列向量无论怎么进行线性组合都无法得到 A 中的任何一个列向量。**这说明，列空间随着行变换发生了改变。

    但就上面这个例子来看，我们也应该不难理解，**对有些矩阵而言，行变换并不改变其列空间，这样的矩阵往往是在每一行都有主元存在。**
  - 行变换不改变其零空间这一点同样很好理解，上面可能说的比较笼统，消元过程中，解是不会改变的，所以零空间不会改变，但为什么解是不会改变的呢？这里简单地用 3×3 的例子说明一下：

    $$\text{设 }A=\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}\text{，且存在 }x=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\text{ 使得 }Ax=0$$
    $$\text{那么此时显然有}\begin{cases}ax_1+bx_2+cx_3=0\\dx_1+ex_2+fx_3=0\\gx_1+hx_2+ix_3=0\end{cases}$$
    $$\text{简单地，将}A\text{中}k\text{倍的行}2\text{加到行}1\text{上，}A\text{的行}1\text{变为}\begin{bmatrix}a+kd&b+ke&c+kf\end{bmatrix}\text{。此时仍有：}$$
    $$(a+kd)x_1+(b+ke)x_2+(c+kf)x_3=(ax_1+bx_2+cx_3)+k(dx_1+ex_2+fx_3)=0$$

在最开始我们提到，消元过程中行之间和列之间线性相关的情况会表现出来。我们重新列出 A 的消元过程：

$$\text{对矩阵}A\text{进行消元，其中}\color{red}{\text{红色}}\text{标记的为主元}$$
$$\begin{bmatrix}\color{red}{1}&2&2&2\\2&4&6&8\\3&6&8&10\end{bmatrix}\xrightarrow[row3=row3-3\times row1]{row2=row2-2\times row1}\begin{bmatrix}\color{red}{1}&2&2&2\\0&0&\color{red}{2}&4\\0&0&2&4\end{bmatrix}\xrightarrow{row3=row3-row2}\begin{bmatrix}\color{red}{1}&2&2&2\\0&0&\color{red}{2}&4\\0&0&0&0\end{bmatrix}$$

- **列之间的线性相关情况：**
  - **在进行到第二步时，第二列没有主元，这说明第二列是其前面的列的一个线性组合，显然第二列是第一列的两倍。**
  - **在进行到第三步时，第四列没有主元，这说明第四列是其前面的列的一个线性组合，第四列等于两倍的第三列减去第二列。**
- **行之间的线性相关情况：**
  - **在进行到第三步时，第三行没有主元，这说明第三行是其前面的行的一个线性组合，显然第三行等于第一行加上第二行。**

最后我们**通过消元将矩阵化简成阶梯形式（非零元素以一种阶梯形式出现），一般记为** U**。**

- **主元的个数对于矩阵而言意义非凡，我们将矩阵中主元的个数称为秩。**

化简完成以后，我们要开始正式对 $Ax=0$ 进行求解了，只不过我们现在具体是对 $Ux=0$ 进行求解，但不管是哪个式子，解（也即零空间）都是不变的。

- 首先我们需要找出**主变量和自由变量，主变量也即主列对应的变量，自由变量也即自由列对应的变量。这里自由变量的"自由"也即我们可以对其分配任意的数值。**注意到，在课程中所提到的分配数值的策略是：**分别地对其中一个自由变量分配** 1**，其他所有自由变量分配** 0**。**

$$\begin{cases} x_1+2x_2+2x_3+2x_4=0 \\ 2x_3+4x_4=0 \end{cases} \rightarrow \begin{cases} x_1=2x_4-2x_2 \\ x_3=-2x_4 \end{cases}$$

按照课程中的分配数值的策略，首先给自由变量 $\begin{bmatrix} x_2 \\ x_4 \end{bmatrix}$ 赋值为 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$，代入方程组得到解向量为 $\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$。再给自由变量 $\begin{bmatrix} x_2 \\ x_4 \end{bmatrix}$ 赋值为 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$，代入方程组得到解向量为 $\begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$。

我们一般称这两个解向量为**特解，其特殊之处就在于我们给自由变量的赋值策略。**正如视频中所说的，"**这样的分配是很有讲究的**"，那到底讲究在哪呢？我认为其讲究之处就在于，**这种** 0 **和** 1 **的取值对应着一种考虑：即当只考虑某个自由变量时可以取到什么（特）解。这种考虑的好处是，能够确保所得到的特解之间是线性无关的！这一点非常重要也很容易理解，也正是这样，我们才能保证我们所得到的 k（自由变量个数）个特解是能够完整正确地表示整个零空间。**

此外，我们也可以像下面这样把 x **写成以自然变量为权重的两个向量的线性组合的形式，自然而然地得到特解。**

$$x=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}=\begin{bmatrix} 2x_4-2x_2 \\ x_2 \\ -2x_4 \\ x_4 \end{bmatrix}=x_4\begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}+x_2\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

到这里结论就显而易见了：**零空间所包含的正好是特解的线性组合，每个自由变量对应着有一个特解。而自由变量的个数实际上就由矩阵的列数（设为** n**）减去矩阵的主列个数（也即秩，设为** r**）从而得到（**$n-r$**）**

阶梯型矩阵 U 还可以进一步化简为**简化行阶梯形式 RREF（主元所在列除主元外其他所有元素都为** 0**，且主元为** 1**）。**

$$U=\begin{bmatrix} \color{red}{1} & 2 & 2 & 2 \\ 0 & 0 & \color{red}{2} & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{row1=row1-row2} \begin{bmatrix} \color{red}{1} & 2 & 0 & -2 \\ 0 & 0 & \color{red}{2} & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{row2=row2/2} \begin{bmatrix} \color{red}{1} & 2 & 0 & -2 \\ 0 & 0 & \color{red}{1} & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}=R$$

课程后半部分花费大量介绍简化行阶梯形式（RREF），其对于我们进一步理解 $Ax=0$ 的求解具有非常重要的作用。

- **观察我们最终得到的 R，容易发现它还包含了一个单位阵，位于主元行与主元列的交汇处**，现在我们将 R 中的主元列与自由列分开，即交换一下列，使得主元列在一起，自由列在一起， 从而有：

$$\begin{bmatrix} \color{red}{1} & 2 & 0 & -2 \\ 0 & 0 & \color{red}{1} & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{exchange\ col2=col3} \begin{bmatrix} \color{red}{1} & 0 & 2 & -2 \\ 0 & \color{red}{1} & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\text{将矩阵中的} \begin{bmatrix} 2 & -2 \\ 0 & 2 \end{bmatrix} \text{部分记为} F, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{部分记为} I$$

$$\text{仔细观察前面得到的两个特解：} \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

$$\text{注意到前面我们对} R \text{交换了列，所以特解中的分量也需要进行交换（} x_2 \text{和} x_3 \text{）}$$

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

$$\text{仔细观察可发现，这两个解若作为列向量组成一个矩阵，恰好为} \begin{bmatrix} -F \\ I \end{bmatrix}$$

  如何解释上面的现象呢？下面给出上面结论的推导过程：

  现假设矩阵已经消元和简化到 RREF 形式为 R，其列数为 n，其秩为 r，且列顺序已经调整好，也即 R 矩阵主列都在前，自由列都在后，R 的下面是零行，故$R = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}$。那么 $Rx = 0$ 的特解是什么呢？

  我们可以构造一个零空间矩阵 $N = \begin{bmatrix} x_1 & \cdots & x_{n-r} \end{bmatrix} = \begin{bmatrix} X_{pivot} \\ X_{free} \end{bmatrix}$，该矩阵的各列由特解（$x_1, \cdots, x_{n-r}$）组成，故显然有 $RN = 0$（0矩阵）。此时 $\begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_{pivot} \\ X_{free} \end{bmatrix} = IX_{pivot} + FX_{free} = 0$，其中 I 为 $r \times r$，F 为 $r \times (n-r)$，X_{pivot} 为 $r \times (n-r)$，X_{free} 为 $(n-r) \times (n-r)$。

  如果把单位矩阵赋值给 N 的 X_{free} 部分使得 $X_{free} = I$（所有特解的自由变量部分组成的一个矩阵），那么显然每个特解对应于单位矩阵的一列，也即最上面提到的分配策略：给其中一个自由变量分配 1，其他所有自由变量分配 0。这样一来，$X_{pivot} = -F$，也即 $N = \begin{bmatrix} -F \\ I \end{bmatrix}$。

- **矩阵的主元的个数与其转置的主元的个数相同。**

  在本课的最开头实际上已经提到，求解 $Ax = 0$ 时需要进行的消元步骤告诉我们一些行和列的线性相关的信息。**虽然消元后我们得到的是列主元，但列主元本身既能够反映列的线性相关的信息，也能够反映行的线性相关信息，换句话说，列主元个数（秩）= 矩阵线性无关列的个数 = 矩阵线性无关行的个数 = 行主元个数（秩）。**