

线性代数

-7 课 求解 $Ax = 0$ ，主变量，特解

一、知识概要

记得上一节中我们讨论了列空间和零空间的相关问题，那么这一节我们从它们的定义过渡到它们的计算，即如何求解出这些空间的一般形式。给出一种可以解出 $Ax = 0$ 中的 x 构成的零空间的算法。

二、消元法求解零空间

记得之前在讲解使用消元法解方程组 $Ax = b$ (第2课) 时，我们对一种情况是无法处理的，那就是矩阵 A 不可逆的情况。之前对这种情况的解释是：求出的解不唯一。这正好对应了我们现在学到的“空间”概念。

我们首先从最简单的零空间 ($b = 0$) 的计算谈起。

2.1 消元法确定主变量与自由变量 (消元)

【例】设 $A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}$ ，求由 $Ax = 0$ 中的 x 构成的零空间。

$Ax = 0$ 其实就是一个方程组：

$$\begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0$$

还是之前学过的消元法来直接处理矩阵 A ：

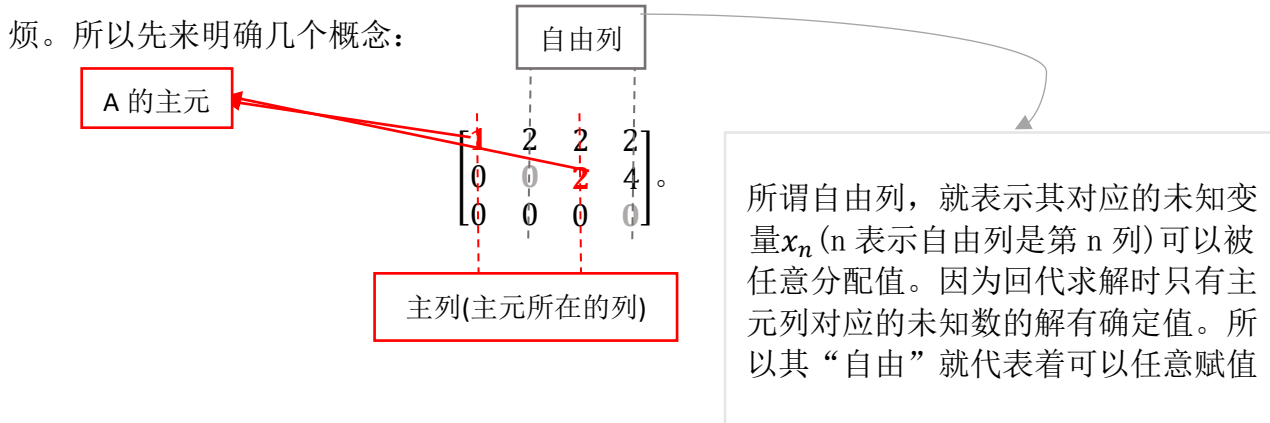
$$\begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

首先注意 A 矩阵消元之后只有两个主元：**1** 和 **2**，主元的个数被称为 **秩**

即 A 的秩 (r) 为：2

接下来应该是进行回代求解了，但是在这之前，由于消元得到的 U 不是一个

严格的上三角矩阵： $\begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ 。对角线上的 0 给我们造成了解不唯一的麻烦。所以先来明确几个概念：



所以这个U的主变量(主元)为 x_1, x_3 。自由变量为 x_2, x_4 。

2.2 对自由变量赋值覆盖零空间(回代)

1. 首先给自由变量 $\begin{bmatrix} x_2 \\ x_4 \end{bmatrix}$ 赋值为 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 。

回代入方程组:

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

当 $\begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 时, 解向量为: $\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

2. 再给自由变量 $\begin{bmatrix} x_2 \\ x_4 \end{bmatrix}$ 赋值为 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 。

再次回代入方程组:

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

这次当 $\begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 时, 解向量为 $\begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$

还有别的给 $\begin{bmatrix} x_2 \\ x_4 \end{bmatrix}$ 赋值的方法吗? 很明显其余的赋值方法都可以被 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 的线性组合所覆盖, 所以这两个解向量足够代表零空间的特征了, 我们称这两个解向量为: **特解**。其特殊之处便在于给自由变量赋值为了 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 。

通过**特解**的任意倍的线性组合可以构造出整个零空间。

即 $\begin{cases} Ax = 0 \text{ 的所有解} \\ Ax = 0 \text{ 中的 } x \text{ 构成的零空间} \end{cases}$ 为:

$$x = c \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + d \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

2.3 算法总结

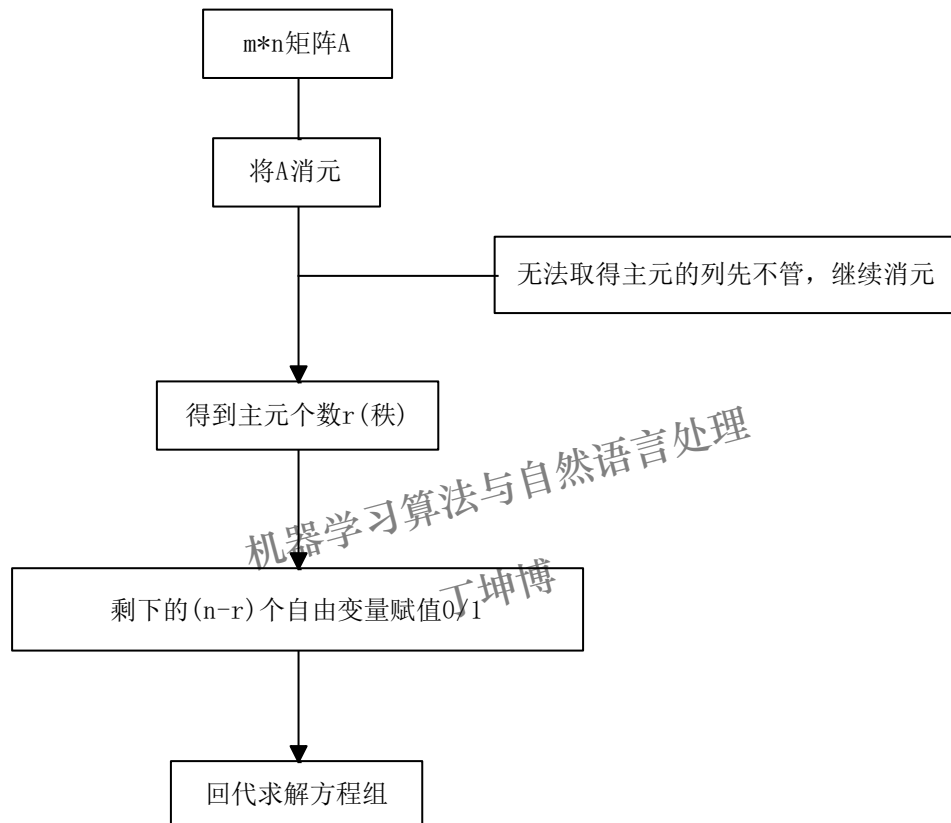
对于一个 $m \times n$ 的矩阵 A, 若其秩(R)为 r, 那么就意味着其主变量为 r 个, 而

自由变量为 $n-r$ 个。也就是只有 r 个方程起作用，而一共有 n 个变量 x ，我们将

其中的 $n-r$ 个自由变量依次赋值为 $\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ 。接下来解方程求特解，将特解

的任意倍进行线性组合就可以了。

整体流程：



三. 简化行阶梯形式

上面的消元法看上去已经很完美了，但是最后一步解方程还有化简的余地，最后得到的 U 矩阵还可以被进一步化简。

拿上面【例】中的 U 矩阵 $\begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ 为例，继续化简：

- 首先向上消元，使主元列除主元之外都是 0: $\begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

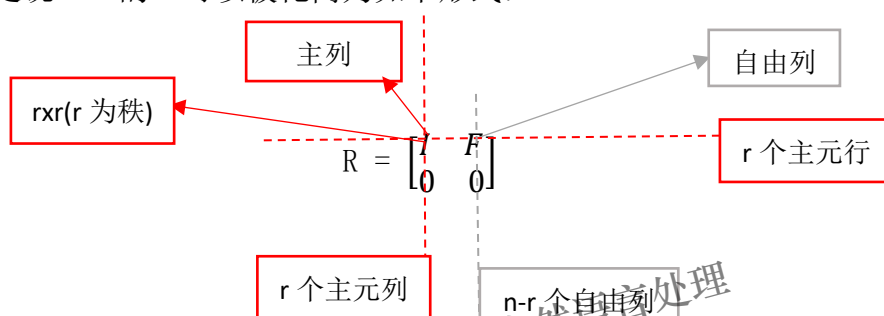
- 提出一列元素公倍数，使主元均为 1: $2 \begin{bmatrix} \mathbf{1} & 2 & 0 & -2 \\ 0 & 0 & \mathbf{1} & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

注：这就是简化行阶梯形式，将原来的行阶梯型矩阵简化，得到主列，自由列的最简单形式。

- 列交换，使左上角变为单位阵 I: $2 \begin{bmatrix} \mathbf{1} & 0 & 2 & -2 \\ 0 & \mathbf{1} & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

显然，倍数 2 可以略去不看，不影响我们的解。此时的 $\begin{bmatrix} \mathbf{1} & 0 & 2 & -2 \\ 0 & \mathbf{1} & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R$ 。

也就是说 $m \times n$ 的 A 可以被化简为如下形式：



对比上面的 $R = \begin{bmatrix} \mathbf{1} & 0 & 2 & -2 \\ 0 & \mathbf{1} & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ ，很好理解 R 的构成。左上角 I 是消元化简

之后得到的 $r \times r$ 大小的单位阵，代表着主列。右上角的 F 代表着自由列经过化简剩余的形式。

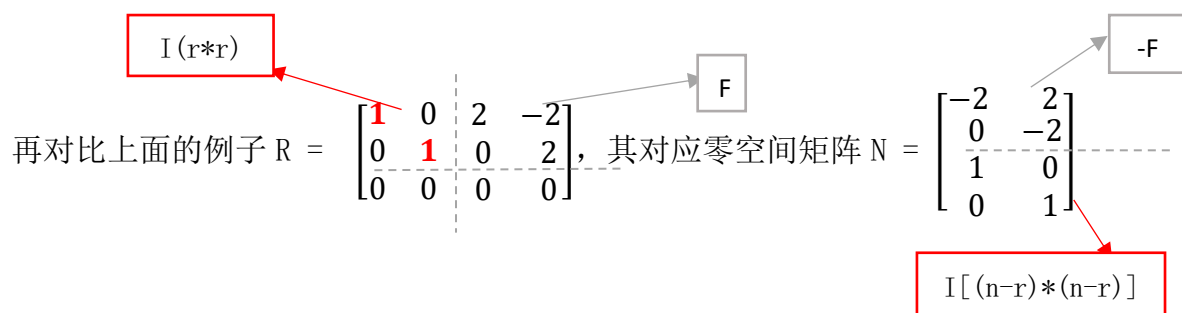
现在假设有一个零空间矩阵：即零空间矩阵各列由特解组成，记 N 为零空间矩阵。联系之前学习矩阵乘法时学到的分块乘法，不难得到：

$$Ax = Rx = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{\text{主元}} \\ x_{\text{自由变量}} \end{bmatrix} = RN = 0$$

得到：

$$\text{零空间矩阵 } N = \begin{bmatrix} -F \\ I \end{bmatrix}$$

$I[(n-r) \times (n-r)]$ 这个单位阵是将 $(n-r)$ 个自由变量分别赋值 0/1 得到的



交换二三行，取 N 的列向量，即为特解 $\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$ 。所以 R 可以直接求解零空间。

接下来我们用一道例题熟悉一下这个流程：

【例】 $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 6 & 8 \\ 2 & 8 & 10 \end{bmatrix}$ ，求解 $Ax = 0$ 中 x 构成的零空间。

解：

(1) A 消元为 U: $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 6 & 8 \\ 2 & 8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = U$ (秩 r 为 2)

(2) U 化简为 R: $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = R$ (提一行的公倍数不影响解)

(3) 分块，写出零空间矩阵：

机器学习算法与自然语言处理

$I (r \times r = 2 \times 2)$

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

F

$-F$

$N = \begin{bmatrix} -F \\ I \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$

$I (n-r) \times (n-r) = 1 \times 1$

五. 学习感悟

这节学习的是计算 $Ax = 0$ 中的 x 构成的零空间的方法，即：消元，找主变量与自由变量，为自由变量赋值，得到特解，特解线性组合得到零空间。后面又介绍了化简 U 变为 R，直接利用 R 的结构得到零空间矩阵 N 的方法。这节重在计算流程，需要加以练习才可以熟练掌握。