选主元的高斯-约当 (Gauss-Jordan) 消元法解线性方程组 / 求逆矩阵

learnhard (codelast.com)

2011 年 03 月 13 日

Abstract

Tagged on: Gauss-Jordan optimization 最优化 消元 线性方程组 选主元 高斯-约当
Category: Algorithm, 原创. 14 Comments

选主元的高斯-约当 (Gauss-Jordan) 消元法在很多地方都会用到, 例如求一个矩阵的逆矩阵/解线性方程组 （插一句: LM 算法求解的一个步骤）, 等等。 它的速度不是最快的, 但是它非常稳定[1], 同时它的求解过程也比较清晰明了, 因而人们使用较多。 下面我就用一个例子来告诉你 Gauss-Jordan 法的求解过程吧。 顺便再提及一些注意事项以及扩展话题。

对本文中所提到的“主元”等概念的解释， 可以参考此链接。
假设有如下的方程组:

$$\begin{cases} x_1 + 3x_2 + x_3 = 11 \\ 2x_1 + x_2 + x_3 = 8 \\ 2x_1 + 2x_2 + x_3 = 10 \end{cases} \tag{1}$$

写成矩阵形式就是: $\mathbf{AX} = \mathbf{B}$， 其中:

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 11 \\ 8 \\ 10 \end{pmatrix} \tag{2}$$

且 $X = (X_1, X_2, X_3)^T$。
现对矩阵 $\mathbf{A}$ 作初等行变换， 同时矩阵 $\mathbf{B}$ 也作同样的初等变换， 则当 $\mathbf{A}$ 化为单位矩阵的时候， 有:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{X} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} \tag{3}$$

显而易见， 我们得到了方程组的解 $\mathbf{X} = (1, 2, 4)^T$。

所以， 我们要以一定的策略， 对 $\mathbf{A}$ 和 $\mathbf{B}$ 施以一系列的初等变换[2]， 当 $\mathbf{A}$ 化为单位矩阵的时候, $\mathbf{B}$ 就为方程组的解。

选主元的 G-J 消元法通过这样的方法来进行初等变换： 在每一个循环过程中， 先寻找到主元， 并将主元通过行变换 （无需列变换） 移动到矩阵的主对角线上， 然后将主元所在的行内的所有元素除以主元， 使得主元化为 1； 然后观察主元所在的列上的其他元素， 将它们所在的行减去主元所在的行乘以一定的倍数， 使得主元所在的列内、除主元外的其他元素化为 0， 这样就使得主元所在的列化为了单位矩阵的形式。 这就是一个循环内做的工作。 然后， 在第二轮循环的过程中， 不考虑上一轮计算过程中主元所在的行和列内的元素， 在剩下的矩阵范围内寻找主元， 然后（如果其不在主对角线上的话） 将其移动到主对角线上， 并再次进行列的处理， 将列化为单位矩阵的形式。 余下的步骤依此类推。 具体的计算过程的一个例子， 请看下面我举的求逆矩阵的过程。

如果要解系数矩阵相同、 右端向量不同的 n 个方程组， 在设计程序的时候， 没有必要“解 n 次方程组”， 我们完全可以在程序中， 将所有的右端向量以矩阵的数据结构（类似于二维数组） 来表示， 在系数矩阵作行变换的时候， 矩阵里的每一个右端向量也做同样的变换， 这样， 我们在一次求解运算的过程中， 实际上就是同时在解 n 个方程组了， 这是要注意的地方。

---

[1] 来自网上的定义: 一个计算方法, 如果在使用此方法的计算过程中, 舍入误差得到控制, 对计算结果影响较小, 称此方法为数值稳定的

[2] 过程如下:

1 交换矩阵的两行或列

2 用一个不为零的数乘矩阵的某一行或列

3 用一个数乘矩阵某一行或列加到另一行或列上

那么，G-J 法为什么可以用来求逆矩阵?

假设 $\mathbf{AX}=\mathbf{E}$，其中，$\mathbf{A}$ 为 n 阶系数矩阵（与上面的解线性方程组对照）；$\mathbf{E}$ 为单位矩阵，即 $\mathbf{E}=(e_1,e_2,\cdots,e_n)$，其中 $e_i(i=1,2,\cdots,n)$ 为单位列向量；$\mathbf{X}$ 为 n 个列向量构成的矩阵，即$\mathbf{X}=(x_1,x_2,\cdots,x_n)$，其中 $x_i(i=1,2,\cdots,n)$ 为列向量 。

于是，可以把等式 $\mathbf{AX}=\mathbf{E}$ 看成是求解 n 个线性方程组 $Ax_i=e_i(i=1,2,\cdots,n)$，求出了所有的 x_i 之后，也即得到了矩阵 $\mathbf{X}$。而由 $\mathbf{AX}=\mathbf{E}$ 可知，矩阵 $\mathbf{X}$ 是 $\mathbf{A}$ 的逆矩阵，即 $\mathbf{X}=\mathbf{A}-1$。这样，就求出了 $\mathbf{A}$ 的逆矩阵了 。于是，求逆矩阵的过程被化成了解线性方程组的过程，因此我们可以用 Gauss-Jordan 消元法来求逆矩阵 。

求逆矩阵时，系数矩阵 $\mathbf{A}$ 和单位矩阵 $\mathbf{E}$ 可以共用一块存储区，在每一次约化过程中，系数矩阵逐渐被其逆矩阵替代 。

在这里，我用一个实际的例子来说明 G-J 法求逆矩阵的过程:

有如下的方程组:

$$\begin{cases} x_1+3x_2+x_3=11 \\ 2x_1+x_2+x_3=8 \\ 2x_1+2x_2+x_3=10 \end{cases} \tag{4}$$

显而易见，该方程组对应的系数矩阵 $\mathbf{A}$ 和右端向量矩阵 $\mathbf{B}$（此处只有一个右端向量）分别为:

$$A=\begin{pmatrix} 1 & 3 & 1 \\ 2 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}, B=\begin{pmatrix} 11 \\ 8 \\ 10 \end{pmatrix} \tag{5}$$

其实在求逆矩阵的过程中，矩阵 $\mathbf{B}$ 无关紧要，可以忽略，不过此处还是把它写出来了 。下面，把单位矩阵 $\mathbf{E}$ 附在 $\mathbf{A}$ 的右边，构成另一个矩阵 $(\mathbf{A}|\mathbf{E})$:

$$(\mathbf{A}|\mathbf{E})=\left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \tag{6}$$

下面，我们就通过矩阵的初等变换 $elementarytransformation$，将 $\mathbf{A}$ 化为单位矩阵 $\mathbf{E}$，而 $\mathbf{E}$ 则化为了 $\mathbf{A}$ 的逆矩阵 。以下是转化步骤:

1 主元选为 3，所以将 $Row1$（第一行）与 $Row2$（第二行）交换:

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 3 & 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \tag{7}$$

2 主元所在行的所有元素除以主元:

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 0 & 1 & 0 \\ {}^{1}/_{3} & {}^{3}/_{3} & {}^{1}/_{3} & {}^{1}/_{3} & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \tag{8}$$

3 $Row1-Row2$，$Row3-(2\times Row2)$：

$$\left(\begin{array}{ccc|ccc} 2-{}^{1}/_{3} & 1-{}^{3}/_{3} & 1-{}^{1}/_{3} & 0-{}^{1}/_{3} & 1 & 0 \\ {}^{1}/_{3} & {}^{3}/_{3} & {}^{1}/_{3} & {}^{1}/_{3} & 0 & 0 \\ 2-{}^{1\times2}/_{3} & 2-{}^{3\times2}/_{3} & 1-{}^{1\times2}/_{3} & 0-{}^{1\times2}/_{3} & 0 & 1 \end{array}\right) \tag{9}$$

9 式最终被计算为:

$$\left(\begin{array}{ccc|ccc} 1\frac{2}{3} & 0 & {}^{2}/_{3} & -{}^{1}/_{3} & 1 & 0 \\ {}^{1}/_{3} & 1 & {}^{1}/_{3} & {}^{1}/_{3} & 0 & 0 \\ 1\frac{1}{3} & 0 & {}^{1}/_{3} & -{}^{2}/_{3} & 0 & 1 \end{array}\right) \tag{10}$$

（现在，原来的矩阵 $\mathbf{A}$ 有一列被化为了单位阵的形式）

4 重新选主元，这一次主元选为 $5/3$，于是 $Row1 \div {}^{5}/_{3}$（主元所在行的所有元素除以主元）：

$$\left(\begin{array}{ccc|ccc} {}^{5}/_{3} \div {}^{5}/_{3} & 0 \div {}^{5}/_{3} & {}^{2}/_{3} \div {}^{5}/_{3} & -{}^{1}/_{3} \div {}^{5}/_{3} & 1 \div {}^{5}/_{3} & 0 \div {}^{5}/_{3} \\ {}^{1}/_{3} & 1 & {}^{1}/_{3} & {}^{1}/_{3} & 0 & 0 \\ 1\frac{1}{3} & 0 & {}^{1}/_{3} & -{}^{2}/_{3} & 0 & 1 \end{array}\right) \tag{11}$$

11 式最终被计算为:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & {}^{2}/_{5} & -{}^{1}/_{5} & {}^{5}/_{3} & 0 \\ {}^{1}/_{3} & 1 & {}^{1}/_{3} & {}^{1}/_{3} & 0 & 0 \\ 1\frac{1}{3} & 0 & {}^{1}/_{3} & -{}^{2}/_{3} & 0 & 1 \end{array}\right) \tag{12}$$

5 $Row2 - ({}^{1}/_{3} \times Row1)$，$Row3 - ({}^{4}/_{3} \times Row1)$：

$$\left(\begin{array}{ccc|ccc} 1 & 0 & {}^{2}/_{5} & -{}^{1}/_{5} & {}^{5}/_{3} & 0 \\ {}^{1}/_{3} - (1 \times {}^{1}/_{3}) & 1 - (0 \times {}^{1}/_{3}) & {}^{1}/_{3} - ({}^{2}/_{5} \times {}^{1}/_{3}) & {}^{1}/_{3} - (-{}^{1}/_{5} \times {}^{1}/_{3}) & 0 - ({}^{5}/_{3} \times {}^{1}/_{3}) & 0 - (0 \times {}^{1}/_{3}) \\ 1\frac{1}{3} & 0 & {}^{1}/_{3} & -{}^{2}/_{3} & 0 & 1 \end{array}\right) \tag{13}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & {}^{2}/_{5} & -{}^{1}/_{5} & {}^{5}/_{3} & 0 \\ {}^{1}/_{3} - (1 \times {}^{1}/_{3}) & 1 - (0 \times {}^{1}/_{3}) & {}^{1}/_{3} - ({}^{2}/_{5} \times {}^{1}/_{3}) & {}^{1}/_{3} - (-{}^{1}/_{5} \times {}^{1}/_{3}) & 0 - ({}^{5}/_{3} \times {}^{1}/_{3}) & 0 - (0 \times {}^{1}/_{3}) \\ {}^{4}/_{3} - (1 \times {}^{4}/_{3}) & 0 - (0 \times {}^{4}/_{3}) & {}^{1}/_{3} - ({}^{2}/_{5} \times {}^{4}/_{3}) & -{}^{2}/_{3} - (-{}^{1}/_{5} \times {}^{4}/_{3}) & 0 - ({}^{5}/_{3} \times {}^{4}/_{3}) & 1 - (0 \times {}^{4}/_{3}) \end{array}\right) \tag{14}$$

式子 14 可以被简化为:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & \frac{2}{5} & -\frac{1}{5} & \frac{3}{5} & 0 \\ 0 & 1 & \frac{1}{5} & \frac{2}{5} & -\frac{1}{5} & 0 \\ 0 & 0 & -\frac{1}{5} & -\frac{2}{5} & -\frac{4}{5} & 1 \end{array}\right) \tag{15}$$

（现在，原来的矩阵 $\mathbf{A}$ 又有一列被化为了单位阵的形式）

6 重新选主元，这一次主元选为 $-1/5$，于是 $Row3 \div (-1/5)$（主元所在行的所有元素除以主元）：

$$\left(\begin{array}{ccc|ccc} 1 & 0 & \frac{2}{5} & -\frac{1}{5} & \frac{3}{5} & 0 \\ 0 & 1 & \frac{1}{5} & \frac{2}{5} & -\frac{1}{5} & 0 \\ 0 & 0 & 1 & 2 & 4 & -5 \end{array}\right) \tag{16}$$

7 $Row1 - (2/5) \times Row3$，$Row2 - (1/5) \times Row3$：

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -1 & 2 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 2 & 4 & -5 \end{array}\right) \tag{17}$$

现在，原来的矩阵 $\mathbf{A}$ 的所有列都被化为了单位阵的形式。

可见，以上过程非常适合于计算机编程求解。

至此，我们完成了从 $\mathbf{A}$ 到 $\mathbf{E}$ 的转换，这个过程中使用了选主元的方法，但没有使用列交换。于是，原来的单位矩阵 $\mathbf{E}$ 就变成了 $\mathbf{A}-1$，即:

$$A^{-1} = \begin{pmatrix} -1 & -1 & 2 \\ 0 & -1 & 1 \\ 2 & 4 & 5 \end{pmatrix} \tag{18}$$

有人说，在进行转化的过程中，如果某一步发现选中的主元为 0，怎么办？当然，这种情况就进行不下去了（矩阵是奇异的）。