计算方法及MATLAB实现

郑勋烨　编著

主审：　高世臣　褚宝增　王祖朝　王翠香

副审：　李少琪　李明霞　赵琳琳　赵俊芳

国防工业出版社

§5.1 Gauss 消去法

5.1.1 引言：高斯消去法的源流和背景

一、禾实问题 —— 中国古代的方程术

【问题 5.1.1. 禾实问题】

成书于秦汉时期的《九章算术》之第八章名为《方程》，其中给出了一道“禾实问题”并附标准答案。全文如下：

今有上禾三秉，中禾二秉，下禾一秉，实三十九斗。上禾二秉，中禾三秉，下禾一秉，实三十四斗。上禾一秉，中禾二秉，下禾三秉，实二十六斗。问：上、中、下三禾，一秉各几何？

答曰：上禾一秉，九斗四分斗之一。中禾一秉，四斗四分斗之一。下禾一秉，二斗四分斗之三。

译成白话文就是说：

现在有上等水稻（比如袁隆平院士研制的超级稻）3 捆，中等水稻 2 捆，下等水稻（比如北方粳稻）1 捆，共打出大米 39 斗；有上等水稻 2 捆，中等水稻 3 捆，下等水稻 1 捆，共打出大米 34 斗；上等水稻 1 捆，中等水稻 2 捆，下等水稻 3 捆，共打出大米 26 斗；问：上、中、下三种水稻，每捆分别可以打出几斗大米？

答：上等水稻每捆可以打出 $9\frac{1}{4}$ 斗大米，中等水稻每捆可以打出 $4\frac{1}{4}$ 斗大米，下等水稻每捆可以打出 $2\frac{3}{4}$ 斗大米。

【问题 1 的求解】

若用现代数学的理论，问题可以这样解决：

设上、中、下三种水稻，每捆分别可以打出 x_1, x_2, x_3 斗大米，则由题目可知，成立如下 3 元一次线性代数方程组：

$$\begin{cases} 3x_1 + 2x_2 + x_3 & = 39 \\ 2x_1 + 3x_2 + x_3 & = 34 \\ x_1 + 2x_2 + 3x_3 & = 26 \end{cases}$$

其惟一解为 $x_1 = 9\frac{1}{4}, x_2 = 4\frac{1}{4}, x_3 = 2\frac{3}{4}$.

而在没有阿拉伯数字、小数点和行列式等等现代工具的古代中国，智慧的先民是如何解决此问题并给出了精确答案的呢？

依照《九章算术》记载，禾实问题是用“方程术”求解的。具体是摆算筹（类似于冰棍棒子），其大意转化为现代矩阵的语言描述，就是如下的 通过初等变换法消元化简为三角形方程组 的过程：

$$\begin{pmatrix} 3 & 2 & 1 & 39 \\ 2 & 3 & 1 & 34 \\ 1 & 2 & 3 & 26 \end{pmatrix} \underline{3r_2 - 2r_1} \rightarrow$$

$$\begin{pmatrix} 3 & 2 & 1 & 39 \\ 0 & 5 & 1 & 24 \\ 1 & 2 & 3 & 26 \end{pmatrix} \underline{3r_3 - r_1} \rightarrow$$

$$\begin{pmatrix} 3 & 2 & 1 & 39 \\ 0 & 5 & 1 & 24 \\ 0 & 4 & 8 & 39 \end{pmatrix} \underline{5r_3 - 4r_2} \rightarrow$$

$$\begin{pmatrix} 3 & 2 & 1 & 39 \\ 0 & 5 & 1 & 24 \\ 0 & 0 & 36 & 99 \end{pmatrix}.$$

这样我们获得了一个以上三角形矩阵作为系数矩阵的方程组

$$\begin{cases} 3x_1 + 2x_2 & +x_3 = 39 \\ \quad 5x_2 & +x_3 = 24 \\ & 36x_3 = 99 \end{cases}$$

经过简单的“自下而上”的回代运算，先由最后一个方程求得 $x_3 = 2\frac{3}{4}$，代入倒数第二个方程获得 $x_2 = 4\frac{1}{4}$，最后代入第一个方程求得 $x_1 = 9\frac{1}{4}$，易知其惟一解为

$x_1 = 9\frac{1}{4}, x_2 = 4\frac{1}{4}, x_3 = 2\frac{3}{4}.$

这就是中国先哲发现的“方程术” —— 在地上摆算筹给出来的线性代数方程组的基本求解思想。事实上，它就是西方称为“ Gauss 消去法”的消元术，却比西方早出现了两千余年！对于西欧数学世界，高斯 (C.F.Gauss) 仿佛一个给出了“第一推动力”的全能的上帝，而在中国，这样的上帝却是无数收割稻谷的普通农民。

【 证毕】

显然，方程组的系数矩阵是一个 对称矩阵 ，称为网络的 环路电阻矩阵 。整个环路电流的获得现在完全由此线性代数方程组的求解决定。

二、低阶稠密阵和高阶稀疏阵问题

【 **定义 1. 低阶稠密阵和高阶稀疏阵** 】通俗地定义，一个 低阶稠密矩阵 就是阶数不太高（这是相对的，比如 50 阶的矩阵，相对工程应用出现的上万阶的矩阵，显然是小巫见大巫，可以认为是低阶的）而且又有相当多的非零元素。一个 高阶稀疏矩阵 则是阶数比较高而且又有非常多的零元素的矩阵，在工程应用中经常出现。

【问题 1. 低阶稠密阵问题】

对于一个系数矩阵是低阶稠密阵的 n 元一次线性代数方程组，我们早在线性代数的专门课程中便知道可以用 Cramer 法则求解。形如下式的 n 元一次线性代数方程组：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = b_n \end{cases}$$

n 阶系数方阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

非奇异时，其行列式

$$Det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0$$

从而此 n 元一次线性代数方程组存在惟一解向量

$$x_1 = \frac{|A_1|}{|A|}, x_2 = \frac{|A_2|}{|A|}, \cdots\cdots, x_n = \frac{|A_n|}{|A|}$$

其中相应于每个 x_j，分式上的分子是系数行列式 A 的第 j 列向量被相应的非齐次列向量 $\overrightarrow{b} = (b_1, b_2, \cdots, b_n)^T$ 代替后得到的行列式 $|A_j|$，即

$$|A_1| = \begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}, |A_2| = \begin{vmatrix} a_{11} & b_1 & \cdots & a_{1n} \\ a_{21} & b_2 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & b_n & \cdots & a_{nn} \end{vmatrix},$$

$$\cdots\cdots, |A_n| = \begin{vmatrix} a_{11} & a_{12} & \cdots & b_1 \\ a_{21} & a_{22} & \cdots & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & b_n \end{vmatrix}.$$

问题看起来已经大功告成。要解一个正方形的方程组，就来算算行列式再求求商就好了。然而，一切都是“站着说话不腰疼”！我们知道，一个 n 阶行列式的展开计算式是 $n!$ 项的加和：

$$Det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \sum (-1)^{\tau(j_1, j_2, \cdots, j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

其中 $\tau(j_1, j_2, \cdots, j_n)$ 作为排列 $j_1, j_2, \cdots, j_n$ 的逆序数取值为 1 或 -1. 而每一项含有 n 个元素的乘积，要做 $n-1$ 次乘法，这样，计算一个行列式要做 $n!(n-1)$ 次乘法，而完全算出所有需要的 $n+1$ 个行列式 - - 包括一个系数行列式 $|A|$ 和 n 个作为分子的行列式 $|A_j|$—— 就要做 $(n+1)n!(n-1)$ 次乘法，还要做 n 次除法。倘若只算乘法的工作量，求解一个 20 阶的方程组，计算量是

$$(n+1)n!(n-1) = 20! \times 19 \times 20 \approx 9.7 \times 10^{20}$$

即使用每秒 10 亿次的“银河”巨型计算机，也要算 3 万多年！

所以即使对于低阶的稠密矩阵，理论上完美无瑕的 Cramer 法则也是几乎不可行的（甚至还不如大理国 段誉 先生使用的时灵时不灵的“六脉神剑”）。即此，必须从“算法”上寻找根本的解决之道。

【 问题 2. 高阶稀疏阵问题 】

利用固支边界条件（ Clamped Boundary ）建立样条插值函数 $S(x)$, 得到关于 矩 为未知元的 $n+1$ 阶线性代数方程组：其矩阵形式为

$$
\begin{pmatrix}
2 & \lambda_0 & & & \\
\mu_1 & 2 & \lambda_1 & & \\
 & \ddots & \ddots & \ddots & \\
 & & \ddots & 2 & \lambda_{n-1} \\
 & & & \mu_n & 2
\end{pmatrix}
\begin{pmatrix}
M_0 \\ M_1 \\ \vdots \\ M_n
\end{pmatrix}
=
\begin{pmatrix}
d_0 \\ d_1 \\ \vdots \\ d_n
\end{pmatrix}
$$

其系数矩阵是所谓 三对角矩阵, 即只有主对角线及其腋下和肩上的斜对角线上的元素非零，而其他所有位置的元素均为 0. 显然，这是一个 高阶稀疏矩阵 ，根据系数的定义可知它还是 严格对角优势矩阵, 即各行的主对角线元素严格大于其左右两边的元素之和：

$$2 > 1 = \lambda_0 = \mu_n, \quad 2 > 1 = \mu_j + \lambda_j, \quad j = 1, 2, \cdots, n-1$$

此类型方程组存在唯一解，并且可以用 解三对角矩阵方程的追赶法 来求．事实上，对于一般的对角占优的三对角矩阵 A 我们有下述 分解定理 ：

【定理 1】

对角占优的三对角矩阵 A 满足

(1)$|b_1| > |c_1| > 0$;

(2)$|b_i| \geq |a_i| + |c_i|$;

(3)$|b_n| > |a_n| > 0$;

则可唯一分解为二对角阵 $A = LU$ 形式．其中 L 为下二对角方阵， D 为单位上二对角方阵．即

$$
\begin{pmatrix}
b_1 & c_1 & & & & & \\
a_2 & b_2 & c_2 & & & & \\
& a_3 & b_3 & c_3 & & & \\
& & \ddots & \ddots & \ddots & & \\
& & & \ddots & \ddots & \ddots & \\
& & & & \ddots & \ddots & \ddots \\
& & & & a_{n-1} & b_{n-1} & c_{n-1} \\
& & & & & a_n & b_n
\end{pmatrix}
$$

$$
= \begin{pmatrix}
\alpha_1 & & & & & & \\
r_2 & \alpha_2 & & & & & \\
& r_3 & \alpha_3 & & & & \\
& & \ddots & \ddots & & & \\
& & & \ddots & \ddots & & \\
& & & & r_{n-1} & \alpha_{n-1} & \\
& & & & & r_n & \alpha_n
\end{pmatrix}
$$

$$
\begin{pmatrix}
1 & \beta_1 & & & & & \\
 & 1 & \beta_2 & & & & \\
 & & 1 & \beta_3 & & & \\
 & & & \ddots & \ddots & & \\
 & & & & \ddots & \ddots & \\
 & & & & & 1 & \beta_{n-1} \\
 & & & & & & 1
\end{pmatrix}
$$

追赶法的本质是作对角占优的三对角矩阵的 Crout 分解 $A = LU$(参见本章矩阵分解的理论). 其他类型的 高阶稀疏矩阵 ，也有更好的算法解决。比如，分解为更加稀疏的矩阵, 相应于更加简单的方程组。使得许多解能够成为“秃子头上的虱子”，一目了然，那是我们期望的情形。