# 计算方法

## Ch06 线性方程组的直接法

程勇

buctcourse@163.com

http://www.buct.edu.cn

Dept. of Computer
Beijing University of Chemical Technology

March 1, 2017

# 提纲

# 本章提纲

# 线性方程组求解问题

在工程技术、自然科学和社会科学中，经常遇到的许多问题最终都可归结为解线性方程组，如电学中网络问题、用最小二乘法求实验数据的曲线拟合问题，工程中的三次样条函数的插值问题，经济运行中的投入产出问题以及大地测量、机械与建筑结构的设计计算问题等等，都归结为求解线性方程组或非线性方程组的数学问题。因此线性方程组的求解对于实际问题是极其重要的。

实际问题中的线性方程组分类：

- 按系数矩阵中零元素的个数 (稠密线性方程组和稀疏线性方程组);
- 按未知量的个数 (高阶线性方程组和低阶线性方程组);
- 按系数矩阵的形状 (对称正定方程组、三角形方程组和三对角方程组);

# $n$ 阶线性方程组表示

一般形式的 $n$ 阶线性方程组为：

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

写成矩阵形式为：

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

其中

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

当系数矩阵 $\boldsymbol{A}$ 非奇异 (即 $det(\boldsymbol{A}) \neq 0$) 时，方程组有惟一解。

# 解线性方程组的直接法

线性方程组的直接解法一般包括:

- Gauss 消元法
  1. Gauss 消元法;
  2. Gauss 列主元消元法;
- 矩阵分解法
  1. 杜利特尔 (Doolittle) 分解;
  2. 克劳特 (Crout) 分解;
  3. 乔累斯基 (Cholesky) 分解;
- 追赶法

# 高斯消元法

高斯消元法将原方程组的增广矩阵通过初等变换，即：

$$\bar{\boldsymbol{A}} = (\boldsymbol{A}, \boldsymbol{b}) \Longrightarrow \left(\boldsymbol{A}^{(1)}, \boldsymbol{b}^{(1)}\right) \Longrightarrow \cdots \Longrightarrow \left(\boldsymbol{A}^{(n)}, \boldsymbol{b}^{(n)}\right)$$

其中 $\boldsymbol{A}^n$ 为上三角矩阵。
不难得到，方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

和

$$\boldsymbol{A}^n\boldsymbol{x} = \boldsymbol{b}^n$$

同解，以上求解线性方程组的方法称为 Gauss 消元法。

# 下三角形线性方程组求解

下三角形线性方程组求解

$$\boldsymbol{Lx} = \boldsymbol{b} \Longrightarrow \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

通过从上到下代入，可以得到其解为：

$$\begin{cases} x_1 = \frac{b_1}{l_{11}} \\ x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij} x_j}{l_{ii}} \qquad i = 2, 3, 4, \cdots, n \end{cases}$$

# 上三角形线性方程组求解

上三角形线性方程组求解

$$\boldsymbol{U}\boldsymbol{x} = \boldsymbol{b} \Longrightarrow \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

通过从下到上代入，可以得到其解为：

$$\begin{cases} x_n = \frac{b_n}{u_{nn}} \\ x_i = \frac{b_i - \sum_{j=i+1}^{n} u_{ij}x_j}{u_{ii}} & i = n-1, n-2, n-3, \cdots, 2, 1 \end{cases}$$

## Gauss 消元法

对于线性方程组

$$\boldsymbol{Ax} = \boldsymbol{b}$$

如果 $det(\boldsymbol{A}) \neq 0$，对其增广矩阵进行初等变换:

$$\bar{\boldsymbol{A}} = (\boldsymbol{A}, \boldsymbol{b}) = \left(\boldsymbol{A}^{(1)}, \boldsymbol{b}^{(1)}\right) = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{pmatrix}$$

如果 $a_{11}^{(1)} \neq 0$，定义行乘子

$$m_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \quad i = 2, 3, 4, \cdots, n$$

# Gauss 消元法 (续)

然后第 $i$ 行减去第 1 行 $\times m_{i1}$，有：

$$
\begin{aligned}
a_{ij}^{(2)} &= a_{ij}^{(1)} - m_{i1}\,a_{1j}^{(1)} \quad & i = 2, 3, \cdots, n, j = 1, 2, 3, \cdots, n \\
b_{i}^{(2)} &= b_{i}^{(1)} - m_{i1}\,b_{1}^{(1)} \quad & i = 2, 3, \cdots, n
\end{aligned}
$$

得到变换后的矩阵为：

$$
\left(\boldsymbol{A}^{(1)}, \boldsymbol{b}^{(1)}\right) \Longrightarrow \left(\boldsymbol{A}^{(2)}, \boldsymbol{b}^{(2)}\right) = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_{1}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_{2}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_{n}^{(2)} \end{pmatrix}
$$

## Gauss 消元法 (续)

如果 $a_{11}^{(1)}=0$，由于 $det(\boldsymbol{A})\neq 0$，因此第一列中至少有一个元素不为零。如果有 $a_{i_1 1}^{(1)}\neq 0$，则将增广矩阵 $\left(\boldsymbol{A}^{(1)},\boldsymbol{b}^{(1)}\right)$ 的第 1 行和第 $i_1$ 行进行交换后再进行消元。

依次类推，经过 $n-1$ 步后，$\left(\boldsymbol{A}^{(1)},\boldsymbol{b}^{(1)}\right)$ 将变换为如下形式：

$$\left(\boldsymbol{A}^{(1)},\boldsymbol{b}^{(1)}\right)\Longrightarrow\left(\boldsymbol{A}^{(n)},\boldsymbol{b}^{(n)}\right)=\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(n)} & b_n^{(n)} \end{pmatrix}$$

由于 $det(\boldsymbol{A})\neq 0$，因此，上三角矩阵 $\boldsymbol{A}^{(n)}\boldsymbol{x}=\boldsymbol{b}^{(n)}$ 有唯一解。

# Gauss 消元法 (续)

因此可得到线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

的解。即：

$$\begin{cases} x_n = \frac{b_n^{(n)}}{a_{nn}^{(n)}} \\ x_i = \frac{b_i^{(i)} - \sum_{j=i+1}^{n} a_{ij}^{(i)} x_j}{a_{ii}^{(i)}} & i = n-1, n-2, n-3, \cdots, 2, 1 \end{cases}$$

## Gauss 消元法例子

例题：用 Gauss 消元法求解方程组：

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 6 \\ 2x_1 + 3x_2 + 4x_3 = 9 \\ x_1 + 3x_2 + 2x_3 = 6 \end{cases}$$

解：$n = 3, a_{11} = 1 \neq 0,$

$$m_{21} = a_{21}/a_{11} = 2/1 = 2$$
$$m_{31} = a_{31}/a_{11} = 1/1 = 1$$

将 $i(i = 2, 3)$ 个方程减去 $m_{i1}\times$ 第 1 个方程，完成第一步消元，得到：

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 6 \\ -x_2 - 2x_3 = -3 \\ x_2 - x_3 = 0 \end{cases}$$

## Gauss 消元法例子 (续)

此时 $a_{22}^{(1)} = -1 \neq 0, m_{32} = a_{32}^{(1)} / a_{22}^{(1)} = 1/-1 = -1$,
将 $i(i=3)$ 个方程减去 $m_{32}\times$ 第 2 个方程，完成第二步消元，得到：

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 6 \\ -x_2 - 2x_3 = -3 \\ -3x_3 = -3 \end{cases}$$

通过以下回代可求得方程的解：

$$\begin{aligned} x_3 &= -3/-3 = 1 \\ x_2 &= -(-3 + 2x_3) = -(-3 + 2\times 1) = 1 \\ x_1 &= 6 - 2x_2 - 3x_3 = 6 - 2\times 1 - 3\times 1 = 1 \end{aligned}$$

故求得的方程解为：$x_1 = x_2 = x_3 = 1$。

# Gauss 列主元消去法

例题：用 Gauss 消去法解线性方程组 (用 3 位十进制浮点数计算)。

$$\begin{pmatrix} 10^{-8} & 2 & 3 \\ -1 & 3.712 & 4.623 \\ -2 & 1.072 & 5.643 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

解: 上述方程组和例 1 一样若用 Gauss 消去法计算会有小数作除数的现象，若采用换行的技巧，则可避免。

$$\bar{\boldsymbol{A}} = (\boldsymbol{A}, \boldsymbol{b}) = \begin{pmatrix} 10^{-8} & 2 & 3 & 1 \\ -1 & 3.712 & 4.623 & 2 \\ -2 & 1.072 & 5.643 & 3 \end{pmatrix}$$

因为 $10^{-8}$ 很小，绝对值最大的列元素为 $a_{13} = -2$，因此可以将 1,3 行交换，变换后结果为：

# Gauss 列主元消去法 (续)

$$\left(\boldsymbol{A}^{(1)}, \boldsymbol{b}^{(1)}\right) = \begin{pmatrix} -2 & 1.072 & 5.643 & 3 \\ -1 & 3.712 & 4.623 & 2 \\ 10^{-8} & 2 & 3 & 1 \end{pmatrix}$$

求得 $m_{21} = 0.5, m_{31} = -0.5 * 10^{-8}$，然后第一步消元，结果如下：

$$\left(\boldsymbol{A}^{(2)}, \boldsymbol{b}^{(2)}\right) = \begin{pmatrix} -2 & 1.072 & 5.643 & 3 \\ 0 & 3.176 & 1.8015 & 0.5 \\ 0 & 2 & 3 & 1 \end{pmatrix}$$

这时 3.176 绝对值最大，不需要换行。取 $m_{32} = 0.62972292$，第二步消元，结果如下：

## Gauss 列主元消去法 (续)

$$\left(\boldsymbol{A}^{(3)}, \boldsymbol{b}^{(3)}\right) = \begin{pmatrix} -2 & 1.072 & 5.643 & 3 \\ 0 & 3.176 & 1.8015 & 0.5 \\ 0 & 0 & 1.8655541 & 0.68513854 \end{pmatrix}$$

再回代，计算方程的解为：

$$x_3 = \frac{b_3^{(3)}}{a_{33}^{(3)}} = \frac{0.68513854}{1.8655541} = 0.36725739$$

$$x_2 = \frac{b_2^{(2)} - a_{23}^{(2)} x_3}{a_{22}^{(2)}} = \frac{0.5 - 1.8015 \times x_3}{3.176} = -0.05088607$$

$$x_1 = \frac{b_1^{(1)} - a_{12}^{(1)} x_2 - a_{13}^{(1)} x_3}{a_{11}^{(1)}} = -0.49105820$$

方程的准确解为：$(-0.491058227, -0.050886075, 0.367257384)^\top$

## 矩阵分解法

如果将线性方程组

$$\boldsymbol{Ax} = \boldsymbol{b}$$

的系数矩阵 $\boldsymbol{A}$ 分解为两个三角形矩阵 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 的乘积，即：

$$\boldsymbol{A} = \boldsymbol{LU}$$

则

$$\boldsymbol{Ax} = \boldsymbol{b}$$
$$\boldsymbol{LUx} = \boldsymbol{b}$$

$$\boldsymbol{Ly} = \boldsymbol{b} \qquad \boldsymbol{Ux} = \boldsymbol{y}$$

因为 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 都是三角形矩阵，因此可以通过回代直接求解，这种方法我们称为求解线性方程组的三角形分解法，简称矩阵分解法。

# 矩阵分解

**定义**

若方阵 $\boldsymbol{A}$ 可以分解为一个下三角矩阵 $\boldsymbol{L}$ 和一个上三角矩阵 $\boldsymbol{U}$ 的乘积，即 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$，则这种分解称为 $\boldsymbol{A}$ 的一种三角分解或 LU 分解。如果 $\boldsymbol{L}$ 为单位下三角矩阵，则称为杜利特尔 (Doolittle) 分解；若 $\boldsymbol{U}$ 为单位上三角矩阵，则称为克劳特 (Crout) 分解。

**矩阵的 LU 分解定理**

设 $\boldsymbol{A}$ 为 $n$ 阶方阵，如果 $\boldsymbol{A}$ 的顺序主子矩阵 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_{n-1}$ 均非奇异，则 $\boldsymbol{A}$ 可分解为一个单位下三角矩阵 $\boldsymbol{L}$ 和一个上三角矩阵 $\boldsymbol{U}$ 的乘积，即 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$，且这种分解是唯一的。

# 杜利特尔 (Doolittle) 分解

根据上述定理，如果 $n$ 阶方阵 $\boldsymbol{A}$ 的各阶顺序主子式不为 0，则存在唯一的 LU 分解。现考察 4 阶方阵，有：

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

右边两个矩阵相乘，可得到如下等式：

$$a_{11} = u_{11}, a_{12} = u_{12}, a_{13} = u_{13}, a_{14} = u_{14}$$

$$a_{21} = l_{21}u_{11}, a_{22} = l_{21}u_{12} + u_{22}, a_{23} = l_{21}u_{13} + u_{23}, a_{24} = l_{21}u_{14} + u_{24}$$

$$a_{31} = l_{31}u_{11}, a_{32} = l_{31}u_{12} + l_{32}u_{22}, a_{33} = l_{31}u_{13} + l_{32}u_{23} + u_{33}$$

$$a_{34} = l_{31}u_{14} + l_{32}u_{24} + u_{34}, a_{41} = l_{41}u_{11}, a_{42} = l_{41}u_{12} + l_{42}u_{22}$$

$$a_{43} = l_{41}u_{13} + l_{42}u_{23} + l_{43}u_{33}, a_{44} = l_{41}u_{14} + l_{42}u_{24} + l_{43}u_{34} + u_{44}$$

## 杜利特尔 (Doolittle) 分解 (续)

对于上述等式，我们可以按行设定计算顺序：

$$u_{11}, u_{12}, u_{13}, u_{14} \to l_{21} \to u_{22}, u_{23}, u_{24}$$

$$\to l_{31}, l_{32} \to u_{33}, u_{34} \to l_{41}, l_{42}, l_{43} \to u_{44}$$

很容易导出显示的计算公式：

$$\begin{aligned}
u_{11} &= a_{11}, u_{12} = a_{12}, u_{13} = a_{13}, u_{14} = a_{14}, \\
l_{21} &= a_{21}/u_{11}, u_{22} = a_{22} - l_{21}u_{12}, \\
u_{23} &= a_{23} - l_{21}u_{13}, u_{24} = a_{24} - l_{21}u_{14} \\
l_{31} &= a_{31}/u_{11}, l_{32} = (a_{32} - l_{31}u_{12})/u_{22}, u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} \\
u_{34} &= a_{34} - l_{31}u_{14} - l_{32}u_{24}, l_{41} = a_{41}/u_{11}, l_{42} = (a_{42} - l_{41}u_{12})/u_{22} \\
l_{43} &= (a_{43} - l_{41}u_{13} - l_{42}u_{23})/u_{33}, \\
u_{44} &= a_{44} - l_{41}u_{14} - l_{42}u_{24} - l_{43}u_{34}
\end{aligned}$$

# 杜利特尔 (Doolittle) 分解 (续)

可以进一步将上述结果推广到一般的 $n$ 阶方阵，将 $\boldsymbol{A}$ 进行杜利特尔 (Doolittle) 分解，可以得到一个单位下三角矩阵 $\boldsymbol{L}$ 和一个上三角矩阵 $\boldsymbol{U}$ 的乘积。

其中 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 的计算公式为：

$$\begin{cases} u_{1j} = a_{1j}, \quad j = 1, 2, 3, \cdots, n \\ l_{i1} = \frac{a_{i1}}{u_{11}}, \quad i = 2, 3, 4, \cdots, n \\ u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj}, \quad j = i, i+1, \cdots, n \\ l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj}}{u_{jj}}, \quad j = 1, 2, \cdots, i-1 \end{cases}$$

矩阵分解法有数值不稳定性。当 $u_{kk} = 0$ 时，三角分解计算中断；当 $u_{kk} \neq 0$，但 $|u_{kk}|$ 很小时，计算 $l_{ik}$ 可能引起较大的舍入误差，使得计算结果不可靠。此时可采用与 Gauss 列主元消去法类似的方法，先选取列主元再进行三角分解。实际上，Doolittle 列主元分解法和 Gauss 列主元消去法在理论上是等价的。

## Doolittle 分解法例子

例题：用 Doolittle 分解求解下列线性方程组。

$$\begin{pmatrix} 2 & 10 & 0 & -3 \\ -3 & -4 & -12 & 13 \\ 1 & 2 & 3 & -4 \\ 4 & 14 & 9 & -13 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \\ -2 \\ 7 \end{pmatrix}$$

解，由 Doolittle 分解公式，可得到：

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1.5 & 1 & 0 & 0 \\ 0.5 & -\frac{3}{11} & 1 & 0 \\ 2 & -\frac{6}{11} & -9 & 1 \end{pmatrix} \begin{pmatrix} 2 & 10 & 0 & -3 \\ 0 & 11 & -12 & 8.5 \\ 0 & 0 & -\frac{3}{11} & -\frac{2}{11} \\ 0 & 0 & 0 & -4 \end{pmatrix}$$

# Doolittle 分解法例子 (续)

因此有：$\boldsymbol{L}\boldsymbol{y}=\boldsymbol{b}$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1.5 & 1 & 0 & 0 \\ 0.5 & -\frac{3}{11} & 1 & 0 \\ 2 & -\frac{6}{11} & -9 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \\ -2 \\ 7 \end{pmatrix}$$

得到 $\boldsymbol{y}=(10 \quad 20 \quad -\frac{17}{11} \quad -16)^\top$。
又有 $\boldsymbol{U}\boldsymbol{x}=\boldsymbol{y}$,

$$\begin{pmatrix} 2 & 10 & 0 & -3 \\ 0 & 11 & -12 & 8.5 \\ 0 & 0 & -3/11 & -2/11 \\ 0 & 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 10 \\ 20 \\ -17/11 \\ -16 \end{pmatrix}$$

得到 $\boldsymbol{x}=(1 \quad 2 \quad 3 \quad 4)^\top$。

# 克劳特 (Crout) 分解

根据上述定理，如果 $n$ 阶方阵 $\boldsymbol{A}$ 的各阶顺序主子式不为 0，则存在唯一的 LU 分解。现考察 4 阶方阵，有：

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix} \begin{pmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

右边两个矩阵相乘，可得到如下等式：

$$a_{11} = l_{11}, a_{12} = l_{11}u_{12}, a_{13} = l_{11}u_{13}, a_{14} = l_{11}u_{14}$$

$$a_{21} = l_{21}, a_{22} = l_{21}u_{12} + l_{22}, a_{23} = l_{21}u_{13} + l_{22}u_{23}, a_{24} = l_{21}u_{14} + l_{22}u_{24}$$

$$a_{31} = l_{31}, a_{32} = l_{31}u_{12} + l_{32}, a_{33} = l_{31}u_{13} + l_{32}u_{23} + l_{33}$$

$$a_{34} = l_{31}u_{14} + l_{32}u_{24} + l_{33}u_{34}, a_{41} = l_{41}, a_{42} = l_{41}u_{12} + l_{42}$$

$$a_{43} = l_{41}u_{13} + l_{42}u_{23} + l_{43}, a_{44} = l_{41}u_{14} + l_{42}u_{24} + l_{43}u_{34} + l_{44}$$

# 克劳特 (Crout) 分解 (续)

对于上述等式，我们可以按列设定计算顺序：

$$l_{11}, l_{21}, l_{31}, l_{41} \to u_{12}, u_{13}, u_{14} \to l_{22}, l_{32}, l_{42}$$

$$\to u_{23}, u_{24} \to l_{33}, l_{43} \to u_{34} \to l_{44}$$

很容易导出显示的计算公式：

$$\begin{aligned}
&l_{11} = a_{11}, l_{21} = a_{21}, l_{31} = a_{31}, l_{41} = a_{41}, u_{12} = a_{12}/l_{11},\\
&u_{13} = a_{13}/l_{11}, u_{14} = a_{14}/l_{11}, l_{22} = a_{22} - l_{21}u_{12},\\
&l_{32} = a_{32} - l_{31}u_{12}, l_{42} = a_{42} - l_{41}u_{12},\\
&u_{23} = (a_{23} - l_{21}u_{13})/l_{22}, u_{24} = (a_{24} - l_{21}u_{14})/l_{22},\\
&l_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23}, l_{43} = a_{43} - l_{41}u_{13} - l_{42}u_{23},\\
&u_{34} = (a_{34} - l_{31}u_{14} - l_{32}u_{24})/l_{33},\\
&l_{44} = a_{44} - l_{41}u_{14} - l_{42}u_{24} - l_{43}u_{34}
\end{aligned}$$

# 克劳特 (Crout) 分解 (续)

可以进一步将上述结果推广到一般的 $n$ 阶方阵，将 $\boldsymbol{A}$ 进行克劳特 (Crout) 分解，可以得到一个下三角矩阵 $\boldsymbol{L}$ 和一个单位上三角矩阵 $\boldsymbol{U}$ 的乘积。

其中 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 的计算公式为：

$$\begin{cases} l_{i1} = a_{i1}, \quad i = 1, 2, 3, \cdots, n \\ u_{1j} = \frac{a_{1j}}{l_{11}}, \quad j = 2, 3, 4, \cdots, n \\ l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj}, \quad j = 1, 2, \cdots, i \\ u_{ij} = \frac{a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}}{l_{ii}}, \quad j = i+1, i+2, \cdots, n \end{cases}$$

## Crout 分解法例子

例题：用 Crout 分解对下面的矩阵进行分解。

$$\begin{pmatrix} 2 & 10 & 0 & -3 \\ -3 & -4 & -12 & 13 \\ 1 & 2 & 3 & -4 \\ 4 & 14 & 9 & -13 \end{pmatrix}$$

解，由 Crout 分解公式，可得到：

$$l_{11} = 2, l_{21} = -3, l_{31} = 1, l_{41} = 4,$$

$$u_{12} = 10/2, u_{13} = 0/2 = 0, u_{14} = -\frac{3}{2},$$

$$l_{22} = -4 - (-3 \cdot 5) = 11, l_{32} = 2 - 1 \cdot 5 = -3,$$

$$l_{42} = 14 - 4 \cdot 5 = -6, u_{23} = \frac{-12 - (-3 \cdot 0)}{11} = -\frac{12}{11},$$

$$u_{24} = \frac{13 - (-3 \cdot (-1.5))}{11} = \frac{17}{22},$$

# Crout 分解法例子 (续)

$$l_{33} = 3 - 1 \cdot 0 - (-3) \cdot (-\frac{12}{11}) = -\frac{3}{11},$$

$$l_{43} = 9 - 4 \cdot 0 - (-6) \cdot (-\frac{12}{11}) = \frac{27}{11},$$

$$u_{34} = \frac{-4 - 1 \cdot (-1.5) - (-3) \cdot \frac{17}{22}}{-\frac{3}{11}} = \frac{2}{3},$$

$$l_{44} = -13 - 4 \cdot (-1.5) - (-6) \cdot \frac{17}{22} - \frac{27}{11} \cdot \frac{2}{3} = -4$$

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ -3 & 11 & 0 & 0 \\ 1 & -3 & -\frac{3}{11} & 0 \\ 4 & -6 & \frac{27}{11} & -4 \end{pmatrix} \begin{pmatrix} 1 & 5 & 0 & -\frac{3}{2} \\ 0 & 1 & -\frac{12}{11} & \frac{17}{22} \\ 0 & 0 & 1 & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

# 平方根法分解

在工程计算中，如应用有限元法解结构力学问题、应用差分方法解椭圆型偏微分方程问题等，最后都归结为求解系数矩阵为对称正定线性方程组的问题。对于这类矩阵有更好的解决方案。

**定理**

设 $\boldsymbol{A}$ 为对称正定矩阵，则存在三角矩阵 $\boldsymbol{L}$，使得 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^{\top}$ 成立，如限定 $\boldsymbol{L}$ 的主对角线元素取正值，则这种分解是唯一的。

此时，按照 $l_{11}\to l_{21}\to l_{22}\to l_{31}\to l_{32}\to l_{33},\cdots$ 逐行求出分解矩阵 $\boldsymbol{L}$ 的元素，计算公式为：

$$
\begin{cases}
l_{ij}=\frac{a_{ij}-\sum_{k=1}^{j-1}l_{ik}l_{jk}}{l_{jj}}, & j=1,2,3,\cdots,i-1\\
l_{ii}=\left(a_{ii}-\sum_{k=1}^{i-1}l_{ik}^2\right)^{\frac{1}{2}}, & i=1,2,3,4,\cdots,n
\end{cases}
$$

这种矩阵分解公式由于含有开方运算而被称作平方根分解。

# 平方根法分解例子

例题：用平方根法对下面的对称正定矩阵进行分解。

$$\begin{pmatrix} 6 & 7 & 5 \\ 7 & 13 & 8 \\ 5 & 8 & 6 \end{pmatrix}$$

解：由平方根法分解公式，可得到：

$$\boldsymbol{L} = \begin{pmatrix} \sqrt{6} & 0 & 0 \\ \frac{7}{\sqrt{6}} & \sqrt{\frac{29}{6}} & 0 \\ \frac{5}{\sqrt{6}} & \frac{13}{\sqrt{174}} & \sqrt{\frac{25}{29}} \end{pmatrix}$$

因此

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^{\top}$$

## Cholesky 分解

**定理**

设对称正定矩阵 $\boldsymbol{A}$ 可分解成 $\boldsymbol{A}=\boldsymbol{LDL}^\top$ 的形式，其中 $\boldsymbol{D}$ 为对角阵，而 $\boldsymbol{L}$ 是单位下三角矩阵。

此时，可按 $d_1 \to l_{21} \to d_2 \to l_{31} \to l_{32} \to d_3 \to \cdots$ 逐行来计算，具体分解公式如下：

$$\begin{cases} l_{ij} = \frac{a_{ij}-\sum_{k=1}^{j-1} d_k l_{ik} l_{jk}}{d_j}, & j = 1,2,3,\cdots i-1 \\ d_i = a_{ii} - \sum_{k=1}^{i-1} d_k l_{ik}^2, & i = 1,2,3,4,\cdots, n \end{cases}$$

这种关于对称正定矩阵的分解称作改进的平方根法，或乔累斯基 (Cholesky) 分解。这种方法具有较高的计算效率，总的计算量约为 $n^3/6$。

# 三对角矩阵

在科学计算中，如三次样条插值 (教材 pp.35) 中，非零元素集中在三条对角线中，如下所示。对于这种三对角矩阵的方程组，可以使用追赶法进行求解。

$$\begin{cases} b_1x_1 + c_1x_2 = f_1 \\ a_2x_1 + b_2x_2 + c_2x_3 = f_2 \\ \cdots \\ a_{n-1}x_{n-2} + b_{n-1}x_{n-1} + c_{n-1}x_n = f_{n-1} \\ a_nx_{n-1} + b_nx_n = f_n \end{cases}$$

可以将其系数记为：

$$\boldsymbol{A} = \begin{bmatrix} b_1 & c_1 & 0 & \cdots & 0 & 0 & 0 \\ a_2 & b_2 & c_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & a_n & b_n \end{bmatrix}$$

# 追赶法

**定理**

设上述三对角矩阵为对角占优矩阵，即满足如下条件：

$$\begin{cases} |b_1| > |c_1| \\ |b_i| > |a_i| + |c_i| \quad i = 2, 3, 4, \cdots, n-1 \\ |b_n| > |a_n| \end{cases}$$

则它是非奇异的，这时方程有唯一解。

对于上述三对角矩阵，可以先从方程 (2) 中消去 $x_1$，方程 (3) 中消去 $x_2$，方程 (4) 中消去 $x_3$，以此类推，从方程 (n) 中消去 $x_{n-1}$，因此得到如下形式的方程组。

# 追赶法 (续)

(一) 消元过程

$$\begin{cases} x_1 + u_1 x_2 = g_1 \\ x_2 + u_2 x_3 = g_2 \\ \cdots \\ x_{n-1} + u_{n-1} x_n = g_{n-1} \\ x_n = g_n \end{cases}$$

其中系数为,

$$\begin{cases} u_1 = c_1/b_1, g_1 = f_1/b_1 \\ u_i = \frac{c_i}{b_i - u_{i-1} a_i} \quad i = 2, 3, \cdots, n-1 \\ g_i = \frac{f_i - g_{i-1} a_i}{b_i - u_{i-1} a_i} \quad i = 2, 3, \cdots, n \end{cases}$$

## 追赶法 (续)

(二) 回代过程
使用下面的递推式来求得方程的解：

$$\begin{cases} x_n = g_n \\ x_i = g_i - u_i x_{i+1} \quad i = n-1, n-2, \cdots, 1 \end{cases}$$

综上所述，解三对角方程组的追赶法分为“追”和“赶”两个环节：

1. 追的过程 (消元过程)。计算系数 $u_1, u_2 \to \cdots \to u_{n-1}$ 和 $g_1, g_2 \to \cdots \to g_n$；
2. 赶的过程 (回代过程)。按逆序求得解 $x_n \to x_{n-1} \to \cdots \to x_1$；

# 追赶法例子

例题：用追赶法求解下列线性方程组。

$$\begin{pmatrix} 3 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

首先进行“追”的过程，得到如下方程组：

$$u_1 = c_1/b_1 = \frac{1}{3} \qquad g_1 = \frac{1}{3}$$

$$u_2 = f_2 - \frac{a_2}{u_1} c_1 = \frac{3}{7} \qquad g_2 = -\frac{2}{7}$$

$$u_3 = f_3 - \frac{a_3}{u_2} c_2 = \frac{15}{7} \qquad g_3 = \frac{11}{7}$$

$$u_4 = f_4 - \frac{a_4}{u_3} c_3 = \frac{38}{15} \qquad g_4 = -\frac{11}{15}$$

# 追赶法例子

$$\begin{pmatrix} 1 & 1/3 & 0 & 0 \\ 0 & 7/3 & 1 & 0 \\ 0 & 0 & 15/7 & 1 \\ 0 & 0 & 0 & 38/15 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -2/3 \\ 11/7 \\ -11/15 \end{pmatrix}$$

通过“赶的”回代过程，可得到方程的解：

$$\boldsymbol{x} = \begin{pmatrix} \frac{21}{38} \\ -\frac{25}{38} \\ \frac{33}{38} \\ -\frac{11}{38} \end{pmatrix}$$

# 向量和矩阵的范数

为了研究线性方程组近似解的误差估计和迭代法的收敛性，有必要对向量及矩阵的“大小”引进某种度量，即范数的概念。向量范数是用来度量向量长度的，它可以看成是二、三维解析几何中向量长度概念的推广。用 $R^n$ 表示 $n$ 维实向量空间。

**定义**

对任一向量 $\boldsymbol{X} \in R^n$，按照一定规则确定一个实数与它对应，该实数记为 $\|\boldsymbol{X}\|$。若 $\|\boldsymbol{X}\|$ 满足下面三个性质：

1. $\|\boldsymbol{X}\| \geq 0$，$\|\boldsymbol{X}\| = 0$ 当且仅当 $\boldsymbol{X} = 0$；
2. 对任意实数 $\lambda$，$\|\lambda \boldsymbol{X}\| = |\lambda| \|\boldsymbol{X}\|$；
3. 对任意向量 $\boldsymbol{Y} \in R^n$，$\|\boldsymbol{X} + \boldsymbol{Y}\| \leq \|\boldsymbol{X}\| + \|\boldsymbol{Y}\|$；

则称该实数 $\|\boldsymbol{X}\|$ 为向量 $\boldsymbol{X}$ 的范数。

# 常见的范数

在 $R^n$ 中，常见的范数有以下几种：

1. $\|\boldsymbol{X}\|_1 = |x_1| + |x_2| + \cdots + |x_n| = \sum_{i=1}^n |x_i|$;
2. $\|\boldsymbol{X}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$;
3. $\|\boldsymbol{X}\|_\infty = \max\{|x_1|, |x_2|, \cdots, |x_n|\} = \max_{1\le i\le n}\{|x_i|\}$;

其中 $x_1, x_2, \cdots, x_n$ 分别是 $\boldsymbol{X}$ 的 $n$ 个分量。以上定义的范数分别称为 1-范数，2-范数和 $\infty$-范数。可以验证它们都是满足范数性质的，其中 $\|x\|_2$ 是由内积导出的向量范数。

上述范数都是 $p$ 范数

$$\|\boldsymbol{X}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$$

当不需要指明使用哪一种向量范数时，就用记号 $\|\cdot\|$ 泛指任何一种向量范数。

## $\infty$-范数

**定理**

对任意向量 $\boldsymbol{x} \in R^n$，有 $\lim_{p\to\infty} ||\boldsymbol{x}||_p = ||\boldsymbol{x}||_\infty$。

证明：由于 $||\boldsymbol{x}||_\infty = \max_{1\leq i\leq n} |x_i|$
所以

$$||\boldsymbol{x}||_\infty = \left(\max_{1\leq i\leq n} |x_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \leq \left(n \max_{1\leq i\leq n} |x_i|^p\right)^{\frac{1}{p}}$$

即：

$$||\boldsymbol{x}||_\infty \leq ||\boldsymbol{x}||_p \leq n^{\frac{1}{p}} ||\boldsymbol{x}||_\infty$$

当 $p \to \infty$，$n^{\frac{1}{p}} \to 1$。
所以：

$$\lim_{p\to\infty} ||\boldsymbol{x}||_p = ||\boldsymbol{x}||_\infty$$

# 范数计算的例子

例题：设 $\boldsymbol{x}=(1,0,-1,2)^\top$，试计算 1-范数，2-范数和 $\infty$-范数。

解:

$$
\begin{aligned}
\|\boldsymbol{x}\|_1 &= 1+0+|-1|+2=4 \\
\|\boldsymbol{x}\|_2 &= \sqrt{1^2+0^2+(-1)^2+2^2} \\
&= \sqrt{6} \\
\|\boldsymbol{x}\|_\infty &= \max\{1,0,|-1|,2\} \\
&= 2
\end{aligned}
$$

# 矩阵的范数

**定义**

如果矩阵 $\boldsymbol{A} \in R^{n\times n}$ 的某个非负实值函数 $N(\boldsymbol{A}) = \|\boldsymbol{A}\|$，满足：

1. $\|\boldsymbol{A}\| \geq 0$，当且仅当 $\boldsymbol{A} = 0$ 时，$\|\boldsymbol{A}\| = 0$;
2. 对任意实数 $\lambda$，$\|\lambda\boldsymbol{A}\| = |\lambda|\|\boldsymbol{A}\|$;
3. 对任意矩阵 $\boldsymbol{A}, \boldsymbol{B} \in R^{n\times n}$，$\|\boldsymbol{A}+\boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$;
4. $\|\boldsymbol{AB}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\|$

则称 $N(\boldsymbol{A})$ 是 $R^{n\times n}$ 的一个矩阵范数 (或模)。

对 $n$ 阶方阵 $\boldsymbol{A} = (a_{ij})_n$，其 1,2,$\infty$ 范数计算如下。其中，$\lambda_{\max}(\boldsymbol{A}^\top\boldsymbol{A})$ 表示 $\boldsymbol{A}^\top\boldsymbol{A}$ 的最大特征值，即满足

$$f(\lambda) = |\lambda \boldsymbol{I}_n - \boldsymbol{A}^\top\boldsymbol{A}| = 0$$

# 矩阵范数的计算

## 定义

1. $\|\boldsymbol{A}\|_\infty$ 范数 (称为 $\boldsymbol{A}$ 的行范数):

$$\|\boldsymbol{A}\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

2. $\|\boldsymbol{A}\|_1$ 范数 (称为 $\boldsymbol{A}$ 的列范数):

$$\|\boldsymbol{A}\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$$

3. $\|\boldsymbol{A}\|_2$ 范数 (称为 $\boldsymbol{A}$ 的 2 范数或谱范数):

$$\|\boldsymbol{A}\|_2 = \sqrt{\lambda_{\max}(\boldsymbol{A}^\top \boldsymbol{A})}$$

# 矩阵范数计算例子

例题：计算方阵

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & -2 & 4 \end{pmatrix}$$

的三种常用范数。

解：

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}| = \max\{1,4,8\} = 8$$

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}| = \max\{1,6,6\} = 6$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)} = \sqrt{32} = 4\sqrt{2}$$

# 矩阵的谱半径

**定义**

设 $\boldsymbol{A} \in R^{n\times n}$ 的特征值为 $\lambda_i (i = 1, 2, 3, \cdots, n)$，称

$$\rho(\boldsymbol{A}) = \max_{1\leq i\leq n} |\lambda_i|$$

为 $\boldsymbol{A}$ 的谱半径。

**定理**

设 $\boldsymbol{A} \in R^{n\times n}$ 为 $n$ 阶方阵，则对任意矩阵范数 $\|\cdot\|$ 都有 $\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|$。

# 迭代法收敛性定理

## 迭代法收敛性定理

设有线性方程组 $\boldsymbol{X} = \boldsymbol{B}\boldsymbol{X} + \boldsymbol{b}$，则对于任意的初始向量 $\boldsymbol{X}^{(0)}$，迭代法

$$\boldsymbol{X}^{(k+1)} = \boldsymbol{B}\boldsymbol{X}^{(k)} + \boldsymbol{b}$$

收敛的充分必要条件是

$$\rho(\boldsymbol{B}) < 1$$

上述定理表明，线性方程组迭代法收敛与否与 $\boldsymbol{X}^{(0)}$ 和 $\boldsymbol{b}$ 无关，而只与迭代矩阵 $\boldsymbol{B}$ 的性质有关。

# 本章小结

本章主要介绍了解线性方程组的直接法。主要包括高斯消元法、矩阵分解法和追赶法。直接法是一种计算量小而精度高的方法 (克莱姆算法也是一种直接法)。最具有代表性的算法是高斯 (Gauss) 消去法，其它算法大都是它的变型，这类方法是解具有稠密矩阵或非结构矩阵 (零元分布无规律) 方程组的有效方法。

选主元的算法有很好的数值稳定性。从计算简单出发实际中多选用列主元法。解三对角矩阵方程组 ($\boldsymbol{A}$ 的对角元占优) 的追赶法，解对称正定矩阵方程组的平方根法都是三角分解法，且都是数值稳定的方法，这些方法不选主元素，也具有较高的精度。

在实际应用中如何选择算法是一个重要问题，往往从解的精度、计算量和所需内存等三个方面考虑：但这些条件相互间是矛盾而不能兼顾的，因此实际计算时应根据问题的特点和要求及所用计算机的性能来选择算法。一般说，系数矩阵为中、小型满矩阵，用直接法较好；当系数矩阵为大型、稀疏矩阵时，有效的解法是前章讨论的迭代法。

# 练习题

1. 编程实现 Doolittle 矩阵分解算法。
2. 编程实现高斯消元法。
3. 编程实现 Cholesky 算法。

# 谢谢!

AUTHOR: Cheng Yong
ADDRESS: Dept. of Computer
Beijing University of Chemical Technology
Beijing, 100029, China
EMAIL: buctcourse@163.com