第一章 线性方程组的直接解法

—高斯消去法，LU 分解法等

杜磊

dulei@dlut.edu.cn

大连理工大学 数学科学学院

创新园大厦 B1207

2019 年 9 月 20 日

内容提要

1 简介

2 三角形方程组和三角分解

3 选主元三角分解

4 平方根法

5 分块三角分解

6 Others

1 简介

2 三角形方程组和三角分解

3 选主元三角分解

4 平方根法

5 分块三角分解

6 Others

线性方程组

线性方程组的一般形式为:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad \vdots \qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

其中 a_{ij}, b_i 是已知常数, x_j 是待求未知量, $i = 1, \cdots, m, j = 1, \cdots, n$.

线性方程组

线性方程组的一般形式为:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad \vdots \qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

其中 a_{ij}, b_i 是已知常数, x_j 是待求未知量, $i = 1, \cdots, m, j = 1, \cdots, n$. 对应的**矩阵表达式**为 $Ax = b$:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

线性方程组的一般形式为:

$$
\begin{cases}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\
\qquad \vdots \qquad\qquad\qquad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m,
\end{cases}
$$

其中 a_{ij}, b_i 是已知常数, x_j 是待求未知量, $i = 1, \cdots, m, j = 1, \cdots, n$. 对应的矩阵表达式为 $Ax = b$:

$$
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\ x_2 \\ \vdots \\ x_n
\end{bmatrix}
\begin{bmatrix}
b_1 \\ b_2 \\ \vdots \\ b_m
\end{bmatrix}.
$$

本章我们仅考虑 $m = n$ 且矩阵 A 非奇异.

《九章算术》中的线性方程组问题

《九章算术》是中国最古老的数学典籍之一, 成书约于公元前 100 年. 全书以问题集形式编撰, 共收录 246 个问题. 在一个或数个问题之后, 然后列出答案. 这些问题按性质与解法分为九大类: 方田、粟米、衰分、少廣、商功、均輸、盈不足、方程、勾股. 今天我们使用的 "方程" 一词即源自卷八章名. 方程章专门讨论线性方程组, 共 18 道题.

《九章算术》中的线性方程组问题

《九章算术》是中国最古老的数学典籍之一, 成书约于公元前 100 年. 全书以问题集形式编撰, 共收录 246 个问题. 在一个或数个问题之后, 然后列出答案. 这些问题按性质与解法分为九大类: 方田、粟米、衰分、少廣、商功、均輸、盈不足、方程、勾股. 今天我们使用的 “方程” 一词即源自卷八章名. 方程章专门讨论线性方程组, 共 18 道题.
其中第一题如下:

方程：今有上禾三秉，中禾二秉，下禾一秉，實三十九斗；上禾二秉, 中禾三秉，下禾一秉，實三十四斗；上禾一秉，中禾二秉，下禾三秉，實二十六斗。問上、中、下禾實一秉各幾何?

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。

《九章算术》中的线性方程组问题

方程：今有上禾三秉，中禾二秉，下禾一秉，實三十九斗；上禾二秉, 中禾三秉，下禾一秉，實三十四斗；上禾一秉，中禾二秉，下禾三秉，實二十六斗。問上、中、下禾實一秉各幾何?

方程：今有上禾三秉，中禾二秉，下禾一秉，實三十九斗；上禾二秉, 中禾三秉，下禾一秉，實三十四斗；上禾一秉，中禾二秉，下禾三秉，實二十六斗。問上、中、下禾實一秉各幾何?

若令 x, y, z 分别代表上等稻、中等稻和下等稻各一捆所能得到的稻米斗数, 则对应的方程组为:

$$\begin{cases} 3x+2y+z=39, \\ 2x+3y+z=34, \\ x+2y+3z=26. \end{cases}$$

方程：今有上禾三秉，中禾二秉，下禾一秉，實三十九斗；上禾二秉, 中禾三秉，下禾一秉，實三十四斗；上禾一秉，中禾二秉，下禾三秉，實二十六斗。問上、中、下禾實一秉各幾何?

若令 x, y, z 分别代表上等稻、中等稻和下等稻各一捆所能得到的稻米斗数, 则对应的方程组为:

$$\begin{cases} 3x + 2y + z = 39, \\ 2x + 3y + z = 34, \\ x + 2y + 3z = 26. \end{cases}$$

術曰: 置上禾三秉, 中禾二秉, 下禾一秉, 實三十九斗, 於右方. 中、左禾列如右方. 以右行上禾遍乘中行而以直除. 又乘其次, 亦以直除. 然以中行中禾不盡者遍乘左行而以直除. 左方下禾不盡者, 上為法, 下為實. 實即下禾之實. 求中禾, 以法乘中行下實, 而除下禾之實. 餘如中禾秉數而一, 即中禾之實. 求上禾亦以法乘右行下實, 而除下禾、中禾之實. 餘如上禾秉數而一, 即上禾之實. 實皆如法, 各得一斗.

数值解法

- 直接法
  - Gaussian 消去法 (上/下三角方程组, (选主元)LU 分解等)
  - Cholesky 分解法
  - 解结构线性方程组的其它方法

数值解法

- 直接法
  - Gaussian 消去法 (上/下三角方程组, (选主元)LU 分解等)
  - Cholesky 分解法
  - 解结构线性方程组的其它方法
- 迭代法
  - 古典迭代法 (Jacobi, Gauss-Seidel, SOR 等)
  - 共轭梯度法 t
  - 其它 Krylov 子空间法 (GMRES, BiCGSTAB 等)

1 简介

2 三角形方程组和三角分解

3 选主元三角分解

4 平方根法

5 分块三角分解

6 Others

下三角形方程组的解法

先考虑下三角形方程组 $Ly = b$, $l_{ii} \neq 0\ (i = 1, \cdots, n)$,

$$\begin{bmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & & \\ l_{31} & l_{32} & l_{33} & & \\ \vdots & \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

下三角形方程组的解法

先考虑下三角形方程组 $Ly = b$, $l_{ii} \neq 0$ $(i = 1, \cdots, n)$,

$$\begin{bmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & & \\ l_{31} & l_{32} & l_{33} & & \\ \vdots & \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

$$\Longleftrightarrow \begin{cases} l_{11} y_1 & = b_1 \\ l_{21} y_1 + l_{22} y_2 & = b_2 \\ l_{31} y_1 + l_{32} y_2 + l_{33} y_3 & = b_3 \\ \qquad \vdots & \quad \vdots \\ l_{n1} y_1 + l_{n2} y_2 + \cdots + l_{nn} y_n & = b_n \end{cases}$$

下三角形方程组: 前代法一

$$
\begin{cases}
l_{11}y_1 & = b_1 \\
l_{21}y_1 + l_{22}y_2 & = b_2 \\
\qquad\vdots & \;\;\vdots \\
l_{n1}y_1 + l_{n2}y_2 + \cdots + l_{nn}y_n & = b_n
\end{cases}
$$

下三角形方程组: 前代法一

$$
\left\{\begin{array}{lr}
l_{11}y_1 & = b_1 \\
l_{21}y_1 + l_{22}y_2 & = b_2 \\
\vdots & \vdots \\
l_{n1}y_1 + l_{n2}y_2 + \cdots + l_{nn}y_n & = b_n
\end{array}\right.
\longrightarrow
\left\{\begin{array}{l}
y_1 = b_1/l_{11} \\
y_2 = (b_2 - l_{21}y_1)/l_{22} \\
\quad\vdots \\
y_n = (b_n - \sum_{i=1}^{n-1} l_{ni}y_i)/l_{nn}
\end{array}\right.
$$

$$
\begin{cases}
l_{11}y_1 & = b_1 \\
l_{21}y_1 + l_{22}y_2 & = b_2 \\
\quad\vdots & \vdots \\
l_{n1}y_1 + l_{n2}y_2 + \cdots + l_{nn}y_n & = b_n
\end{cases}
\longrightarrow
\begin{cases}
y_1 = b_1/l_{11} \\
y_2 = (b_2 - l_{21}y_1)/l_{22} \\
\quad\vdots \\
y_n = (b_n - \sum\limits_{i=1}^{n-1} l_{ni}y_i)/l_{nn}
\end{cases}
$$

Algorithm 3 前代法一

```
b(1) = b(1)/L(1,1)
for j = 2 to n do
    for i = 1 to j − 1 do
        b(j) = b(j) − L(j,i) * b(i)
    end for
    b(j) = b(j)/L(j,j)
end for
```

下三角形方程组: 前代法二

$$\begin{cases} l_{11}y_1 & = b_1 \\ l_{21}y_1 + l_{22}y_2 & = b_2 \\ \qquad \vdots & \quad \vdots \\ l_{n1}y_1 + l_{n2}y_2 + \cdots + l_{nn}y_n & = b_n \end{cases}$$

下三角形方程组: 前代法二

$$\begin{cases} l_{11}y_1 & = b_1 \\ l_{21}y_1 + l_{22}y_2 & = b_2 \\ \qquad \vdots & \quad \vdots \\ l_{n1}y_1 + l_{n2}y_2 + \cdots + l_{nn}y_n & = b_n \end{cases}$$

$$\xrightarrow[b(2:n)=b(2:n)-y(1)L(2:n,1)]{y_1=b_1/l_{11}} \begin{cases} l_{22}y_2 & = b_2 \\ l_{32}y_2 + l_{33}y_3 & = b_3 \\ \qquad \vdots & \quad \vdots \\ l_{n2}y_2 + l_{n3}y_3 + \cdots + l_{nn}y_n & = b_n \end{cases}$$

下三角形方程组: 前代法二

$$\begin{cases} l_{11}y_1 & = b_1 \\ l_{21}y_1 + l_{22}y_2 & = b_2 \\ \qquad \vdots & \vdots \\ l_{n1}y_1 + l_{n2}y_2 + \cdots + l_{nn}y_n & = b_n \end{cases}$$

$$\xrightarrow[b(2:n)=b(2:n)-y(1)L(2:n,1)]{y_1=b_1/l_{11}} \begin{cases} l_{22}y_2 & = b_2 \\ l_{32}y_2 + l_{33}y_3 & = b_3 \\ \qquad \vdots & \vdots \\ l_{n2}y_2 + l_{n3}y_3 + \cdots + l_{nn}y_n & = b_n \end{cases}$$

$$\xrightarrow[b(3:n)=b(3:n)-y(2)L(3:n,2)]{y_2=b_2/l_{22}} \begin{cases} l_{33}y_3 & = b_3 \\ l_{43}y_3 + l_{44}y_4 & = b_4 \\ \qquad \vdots & \vdots \\ l_{n3}y_3 + l_{n4}y_4 + \cdots + l_{nn}y_n & = b_n \end{cases}$$

下三角形方程组: 前代法二

Algorithm 4 前代法二

1: **for** $j = 1$ to $n-1$ **do**
2: $\quad b(j) = b(j)/L(j,j)$
3: $\quad$ **for** $i = j+1$ to n **do**
4: $\quad\quad b(i) = b(i) - b(j) * L(i,j)$
5: $\quad$ **end for**
6: **end for**
7: $b(n) = b(n)/L(n,n)$

该算法所需要的加、减、乘、除运算的次数为 $\sum_{i=1}^{n}(2i-1) = n^2$.

元素的存储量为 $n + \sum_{i=1}^{n} i = \frac{n^2+3n}{2}$.

上三角形方程组的解法

再考虑上三角形方程组 $Ux = y$, $u_{ii} \neq 0$ $(i = 1, \cdots, n)$,

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ & u_{22} & u_{23} & \cdots & u_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & u_{n-1,n-1} & u_{n-1,n} \\ & & & & u_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}$$

上三角形方程组的解法

再考虑上三角形方程组 $Ux = y,\ u_{ii} \neq 0\ (i = 1, \cdots, n)$,

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ & u_{22} & u_{23} & \cdots & u_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & u_{n-1,n-1} & u_{n-1,n} \\ & & & & u_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}$$

$$\Longleftrightarrow \left\{ \begin{aligned} u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n &= y_1 \\ u_{22}x_2 + \cdots + u_{2n}x_n &= y_2 \\ \vdots \qquad\qquad & \vdots \\ u_{n-1,n-1}x_{n-1} + u_{n-1,n}x_n &= y_{n-1} \\ u_{nn}x_n &= y_n \end{aligned} \right.$$

解一般线性方程组的前/后代法

对于一般的线性方程组

$$Ax = b,$$

如果能够将 A 分解为 $A = LU$, 则方程组的解 x 便可由下面两步得到:

(i) 前代法解 $Ly = b$;

(ii) 后代法解 $Ux = y$.

解一般线性方程组的前/后代法

对于一般的线性方程组

$$Ax = b,$$

如果能够将 A 分解为 $A = LU$, 则方程组的解 x 便可由下面两步得到:

(i) 前代法解 $Ly = b$;

(ii) 后代法解 $Ux = y$.

Algorithm 6 解方程组 $Ax = b$ 的直接法

1: 计算矩阵分解 $A = LU$ (主要讨论此内容)
2: 求下三角方程组 $Ly = b$
3: 求上三角方程组 $Ux = y$

Gauss 变换

设 $l_k = (0, \cdots, 0, l_{k+1,k}, \cdots, l_{nk})^{\mathrm{T}}$, 称如下形式的初等下三角阵$L_k$ 为 Gauss 变换, 向量l_k 为 Gauss 向量,

$$L_k = I - l_k e_k^{\mathrm{T}} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{nk} & & & 1 \end{bmatrix}.$$

Gauss 变换

设 $l_k = (0, \cdots, 0, l_{k+1,k}, \cdots, l_{nk})^{\mathrm{T}}$, 称如下形式的初等下三角阵$L_k$ 为 Gauss 变换, 向量l_k 为 Gauss 向量,

$$L_k = I - l_k e_k^{\mathrm{T}} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{nk} & & & 1 \end{bmatrix}.$$

Gauss 变换 L_k 的性质:

- L_k 是单位矩阵 I 的一个秩 1 修正;

Gauss 变换

设 $l_k = (0, \cdots, 0, l_{k+1,k}, \cdots, l_{nk})^{\mathrm{T}}$, 称如下形式的初等下三角阵$L_k$ 为 Gauss 变换, 向量l_k 为 Gauss 向量,

$$L_k = I - l_k e_k^{\mathrm{T}} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{nk} & & & 1 \end{bmatrix}.$$

Gauss 变换 L_k 的性质:

- L_k 是单位矩阵 I 的一个秩 1 修正; $\det(L_k) =?$

Gauss 变换

设 $l_k = (0, \cdots, 0, l_{k+1,k}, \cdots, l_{nk})^{\mathrm{T}}$, 称如下形式的初等下三角阵$L_k$ 为 Gauss 变换, 向量l_k 为 Gauss 向量,

$$L_k = I - l_k e_k^{\mathrm{T}} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{nk} & & & 1 \end{bmatrix}.$$

Gauss 变换 L_k 的性质:

- L_k 是单位矩阵 I 的一个秩 1 修正; $\det(L_k) = ?$
- $L_k^{-1} = I + l_k e_k^{\mathrm{T}}$;

Gauss 变换

设 $l_k = (0, \cdots, 0, l_{k+1,k}, \cdots, l_{nk})^{\mathrm{T}}$, 称如下形式的初等下三角阵$L_k$ 为 Gauss 变换, 向量l_k 为 Gauss 向量,

$$L_k = I - l_k e_k^{\mathrm{T}} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{nk} & & & 1 \end{bmatrix}.$$

Gauss 变换 L_k 的性质:

- L_k 是单位矩阵 I 的一个秩 1 修正; $\det(L_k) = ?$
- $L_k^{-1} = I + l_k e_k^{\mathrm{T}}$;
  设 $A \in R^{n\times n}$ 可逆, $U, V \in R^{n\times k}$ 且 $I + V^{\mathrm{T}} A^{-1} U$ 可逆,
  则$(A + UV^{\mathrm{T}})^{-1} = A^{-1} - A^{-1} U(I + V^{\mathrm{T}} A^{-1} U)^{-1} V^{\mathrm{T}} A^{-1}$.
  (Sherman-Morrison-Woodbury 公式)

设 $x=(x_1,\cdots,x_n)^{\mathrm{T}}$, 则有

$$L_k x=(I-l_k e_k^{\mathrm{T}})x=x-x_k l_k=\begin{bmatrix} x_1 \\ \vdots \\ x_k \\ x_{k+1}-x_k l_{k+1,k} \\ \vdots \\ x_n-x_k l_{nk} \end{bmatrix}.$$

设 $x = (x_1, \cdots, x_n)^{\mathrm{T}}$, 则有

$$L_k x = (I - l_k e_k^{\mathrm{T}})x = x - x_k l_k = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ x_{k+1} - x_k l_{k+1,k} \\ \vdots \\ x_n - x_k l_{nk} \end{bmatrix}.$$

因此, 存在 Gauss 向量 l_k 使得 $L_k x = (x_1, \cdots, x_k, 0, \cdots, 0)^{\mathrm{T}}$.

Gauss 消去法

设 $x = (x_1, \cdots, x_n)^{\mathrm{T}}$, 则有

$$L_k x = (I - l_k e_k^{\mathrm{T}})x = x - x_k l_k = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ x_{k+1} - x_k l_{k+1,k} \\ \vdots \\ x_n - x_k l_{nk} \end{bmatrix}.$$

因此, 存在 Gauss 向量 l_k 使得 $L_k x = (x_1, \cdots, x_k, 0, \cdots, 0)^{\mathrm{T}}$.
从而, 可尝试构造一系列 Gauss 变换 $L_1, \cdots, L_{n-1}$ 将 n 阶方阵 A 约化为上三角矩阵 U, 即

$$L_{n-1} \cdots L_1 A = U,$$

进而得

$$A = (L_{n-1} \cdots L_1)^{-1} U = L_1^{-1} \cdots L_{n-1}^{-1} U.$$

令 $L = L_1^{-1} \cdots L_{n-1}^{-1}$, 显然 L 为下三角阵.
即得 $A = LU$ 为下三角矩阵和上三角矩阵乘积.

令 $L = L_1^{-1} \cdots L_{n-1}^{-1}$, 显然 L 为下三角阵.
即得 $A = LU$ 为下三角矩阵和上三角矩阵乘积.
由定义 $L_k = I - l_k e_k^{\mathrm{T}}$, 易证若 $i < j$, 则 $L_i L_j = I - l_i e_i^{\mathrm{T}} - l_j e_j^{\mathrm{T}}$.

$$
L_i L_j = \begin{bmatrix}
1 & & & & & & & \\
& \ddots & & & & & & \\
& & 1 & & & & & \\
& & -l_{i+1,i} & \ddots & & & & \\
& & \vdots & & 1 & & & \\
& & \vdots & & -l_{j+1,j} & \ddots & & \\
& & \vdots & & \vdots & & \ddots & \\
& & -l_{ni} & & -l_{nj} & & & 1
\end{bmatrix}.
$$

类似, 可证 $L = L_1^{-1} \cdots L_{n-1}^{-1} = I + l_1 e_1^{\mathrm{T}} + l_2 e_2^{\mathrm{T}} + \cdots + l_{n-1} e_{n-1}^{\mathrm{T}}$.

$$L = \begin{bmatrix} 1 & & & & & \\ l_{21} & 1 & & & & \\ l_{31} & l_{32} & 1 & & & \\ l_{41} & l_{42} & l_{43} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{n,n-1} & 1 \end{bmatrix}.$$

进一步讨论 L_k 的计算与存储, 记

$$A^{(0)}=A=\begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ a_{21}^{(0)} & a_{22}^{(0)} & \cdots & a_{2n}^{(0)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}^{(0)} & a_{n2}^{(0)} & \cdots & a_{nn}^{(0)} \end{bmatrix}=\begin{bmatrix} a_{11}^{(0)} & \bar{a}_1^{\mathrm{T}} \\ \hat{a}_1 & A_{22}^{(0)} \end{bmatrix},$$ 可构造 $L_1=I-l_1e_1^{\mathrm{T}}$,

使得

进一步讨论 L_k 的计算与存储, 记

$$A^{(0)} = A = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ a_{21}^{(0)} & a_{22}^{(0)} & \cdots & a_{2n}^{(0)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}^{(0)} & a_{n2}^{(0)} & \cdots & a_{nn}^{(0)} \end{bmatrix} = \begin{bmatrix} a_{11}^{(0)} & \bar{a}_1^{\mathrm{T}} \\ \hat{a}_1 & A_{22}^{(0)} \end{bmatrix},$$

可构造 $L_1 = I - l_1 e_1^{\mathrm{T}}$,

使得

$$A^{(1)} = L_1 A^{(0)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix} = \begin{bmatrix} a_{11}^{(0)} & \bar{a}_1^{\mathrm{T}} \\ 0 & A_{22}^{(1)} \end{bmatrix}.$$

进一步讨论 L_k 的计算与存储，记

$$A^{(0)} = A = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ a_{21}^{(0)} & a_{22}^{(0)} & \cdots & a_{2n}^{(0)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}^{(0)} & a_{n2}^{(0)} & \cdots & a_{nn}^{(0)} \end{bmatrix} = \begin{bmatrix} a_{11}^{(0)} & \bar{a}_1^{\mathrm{T}} \\ \hat{a}_1 & A_{22}^{(0)} \end{bmatrix},$$ 可构造 $L_1 = I - l_1 e_1^{\mathrm{T}}$,

使得 $$A^{(1)} = L_1 A^{(0)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix} = \begin{bmatrix} a_{11}^{(0)} & \bar{a}_1^{\mathrm{T}} \\ 0 & A_{22}^{(1)} \end{bmatrix}.$$

$\longrightarrow l_1 = \begin{bmatrix} 0 \\ \hat{a}_1 \end{bmatrix} \Big/ a_{11}^{(0)}$. 若记 $\hat{l}_1 = \begin{bmatrix} l_{21} \\ \vdots \\ l_{n1} \end{bmatrix}$，则有 $A_{22}^{(1)} = A_{22}^{(0)} - \hat{l}_1 \bar{a}_1^{\mathrm{T}}$.

Algorithm 7 LU 分解

1: **for** $k = 1$ to $n-1$ **do**

2: $\quad A(k+1:n,k) = A(k+1:n,k)/A(k,k)$

3: $\quad A(k+1:n,k+1:n) = A(k+1:n,k+1:n) - A(k+1:n,k)A(k,k+1:n)$

4: **end for**

该算法所需要的加、减、乘、除运算的次数为
$\sum_{i=1}^{n-1}((n-k)+2(n-k)^2) = \frac{2}{3}n^3 + \mathcal{O}(n^2)$.
元素的存储量为 n^2.
通常称 Gauss 消去过程中的 $a_{kk}^{(k-1)}$ 为主元.

LU 分解存在的充分条件

定理

主元 $a_{ii}^{(i-1)}(i=1,\cdots,k)$ 均不为零的充分必要条件是 A 的 i 阶顺序主子阵 $A_i\,(i=1,\cdots,k)$ 都是非奇异的.

LU 分解存在的充分条件

定理

主元 $a_{ii}^{(i-1)}(i=1,\cdots,k)$ 均不为零的充分必要条件是 A 的 i 阶顺序主子阵 $A_i\,(i=1,\cdots,k)$ 都是非奇异的.

证明.

使用归纳法证明.

$$A^{(k-1)} = L_{k-1}\cdots L_1 A = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & \cdots & a_{1n}^{(0)} \\ & \ddots & \cdots & \cdots & \vdots \\ & & a_{kk}^{(k-1)} & \cdots & a_{kn}^{(k-1)} \\ & & \vdots & \ddots & \vdots \\ & & a_{nk}^{(k-1)} & \cdots & a_{nn}^{(k-1)} \end{bmatrix}.$$

$\square$

LU 分解存在的充分条件

定理

若 $A\in\mathbb{R}^{n\times n}$ 的顺序主子阵 $A_k\in\mathbb{R}^{k\times k}\,(k=1,\cdots,n-1)$ 均非奇异, 则存在唯一的单位下三角阵 $L\in\mathbb{R}^{n\times n}$ 和上三角阵 $U\in\mathbb{R}^{n\times n}$ 使得 $A=LU$.

LU 分解存在的充分条件

定理

若 $A \in \mathbb{R}^{n\times n}$ 的顺序主子阵 $A_k \in \mathbb{R}^{k\times k}\,(k = 1, \cdots, n-1)$ 均非奇异, 则存在唯一的单位下三角阵 $L \in \mathbb{R}^{n\times n}$ 和上三角阵 $U \in \mathbb{R}^{n\times n}$ 使得 $A = LU$.

证明.

使用反证法. 假设有 $A = L_1 U_1 = L_2 U_2$. 由条件知 L_1, L_2 非奇异, 所以 $U_1 = L_1^{-1} L_2 U_2$. 进而可证 $L_1^{-1} L_2 = I$, 即证 $L_1 = L_2, U_1 = U_2$. □

Gauss 消去法: LU 分解

如果 A 存在 LU 分解, 可将矩阵 A, L, U 划分为如下形式:

$$A = \begin{bmatrix} a_{11} & \bar{a}_1^{\mathrm{T}} \\ \hat{a}_1 & A_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ l_1 & L_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_1^{\mathrm{T}} \\ 0 & U_{22} \end{bmatrix},$$

其中 l_1, u_{11}, u_1 未知待求.

Gauss 消去法: LU 分解

如果 A 存在 LU 分解, 可将矩阵 A, L, U 划分为如下形式:

$$A = \begin{bmatrix} a_{11} & \bar{a}_1^{\mathrm{T}} \\ \hat{a}_1 & A_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ l_1 & L_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_1^{\mathrm{T}} \\ 0 & U_{22} \end{bmatrix},$$

其中 l_1, u_{11}, u_1 未知待求. 等价于

1. $u_{11} = a_{11}$
2. $u_1 = \bar{a}_1$
3. $l_1 = \hat{a}_1 / a_{11}$
4. $L_{22} U_{22} = A_{22} - l_1 u_1^{\mathrm{T}}$ (更新等式右端)

其它方式计算 LU 分解: I

如果 A 存在 LU 分解, 假设矩阵 A, L, U 的 k 阶顺序主子阵有如下划分形式:

$$\begin{bmatrix} A_{11} & \bar{a}_k \\ \hat{a}_k^{\mathrm{T}} & a_{kk} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ l_k^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} U_{11} & u_k \\ 0 & u_{kk} \end{bmatrix},$$

其中 l_k, u_k, u_{kk} 未知待求.

其它方式计算 LU 分解: I

如果 A 存在 LU 分解, 假设矩阵 A, L, U 的 k 阶顺序主子阵有如下划分形式:

$$\begin{bmatrix} A_{11} & \bar{a}_k \\ \hat{a}_k^{\mathrm{T}} & a_{kk} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ l_k^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} U_{11} & u_k \\ 0 & u_{kk} \end{bmatrix},$$

其中 l_k, u_k, u_{kk} 未知待求. 等价于

1. $A_{11} = L_{11}U_{11}$ $(\surd)$
2. $\bar{a}_k = L_{11}u_k$, 计算 $\bar{a}_k$
3. $\hat{a}_k^{\mathrm{T}} = l_k^{\mathrm{T}}U_{11}$, 计算 l_k
4. $a_{kk} = u_{kk} + l_k^{\mathrm{T}}u_k$

其它方式计算 LU 分解: I

如果 A 存在 LU 分解, 假设矩阵 A, L, U 的 k 阶顺序主子阵有如下划分形式:

$$\begin{bmatrix} A_{11} & \bar{a}_k \\ \hat{a}_k^{\mathrm{T}} & a_{kk} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ l_k^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} U_{11} & u_k \\ 0 & u_{kk} \end{bmatrix},$$

其中 l_k, u_k, u_{kk} 未知待求. 等价于

1. $A_{11} = L_{11} U_{11}$ $(\surd)$
2. $\bar{a}_k = L_{11} u_k$, 计算 $\bar{a}_k$
3. $\hat{a}_k^{\mathrm{T}} = l_k^{\mathrm{T}} U_{11}$, 计算 l_k
4. $a_{kk} = u_{kk} + l_k^{\mathrm{T}} u_k$

缺点:

1. 无法选主元
2. 计算过程多次解三角方程组影响计算性能

其它方式计算 LU 分解: II

如果 A 存在 LU 分解, 假设矩阵 A 的前 k 列有如下形式的 LU 分解 (A_{11} 为 $k-1$ 阶方阵):

$$\begin{bmatrix} A_{11} & \bar{a}_k \\ A_{21} & \hat{a}_k \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & l_k \end{bmatrix} \begin{bmatrix} U_{11} & u_k \\ 0 & u_{kk} \end{bmatrix},$$

其中 l_k, u_k, u_{kk} 未知待求.

其它方式计算 LU 分解: II

如果 A 存在 LU 分解, 假设矩阵 A 的前 k 列有如下形式的 LU 分解 (A_{11} 为 $k-1$ 阶方阵):

$$\begin{bmatrix} A_{11} & \bar{a}_k \\ A_{21} & \hat{a}_k \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & l_k \end{bmatrix} \begin{bmatrix} U_{11} & u_k \\ 0 & u_{kk} \end{bmatrix},$$

其中 l_k, u_k, u_{kk} 未知待求. 等价于

1. $A_{11} = L_{11} U_{11}$ ($\surd$)
2. $A_{21} = L_{21} U_{11}$ ($\surd$)
3. $\bar{a}_k = L_{11} u_k$, 计算 u_k.
4. $\hat{a}_k = L_{21} u_k + u_{kk} l_k$, 选取 u_{kk} 使 l_k 首元素为 1.

其它方式计算 LU 分解: II

如果 A 存在 LU 分解, 假设矩阵 A 的前 k 列有如下形式的 LU 分解 (A_{11} 为 $k-1$ 阶方阵):

$$\begin{bmatrix} A_{11} & \bar{a}_k \\ A_{21} & \hat{a}_k \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & l_k \end{bmatrix} \begin{bmatrix} U_{11} & u_k \\ 0 & u_{kk} \end{bmatrix},$$

其中 l_k, u_k, u_{kk} 未知待求. 等价于

1. $A_{11} = L_{11}U_{11}$ ($\surd$)
2. $A_{21} = L_{21}U_{11}$ ($\surd$)
3. $\bar{a}_k = L_{11}u_k$, 计算 u_k.
4. $\hat{a}_k = L_{21}u_k + u_{kk}l_k$, 选取 u_{kk} 使 l_k 首元素为 1.

类似的, 可按 A 的前 k 行计算对应的 LU 分解.

其它方式计算 LU 分解: III

如果 A 存在 LU 分解, 假设矩阵 A 的前 k 列有如下形式的 LU 分解 (A_{11} 为 $k-1$ 阶方阵):

$$\begin{bmatrix} A_{11} & \bar{a}_k \\ A_{21} & \hat{a}_k \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & l_k \end{bmatrix} \begin{bmatrix} U_{11} & u_k \\ 0 & u_{kk} \end{bmatrix},$$

其中 l_k 和 u_{kk} 未知待求.

其它方式计算 LU 分解: III

如果 A 存在 LU 分解, 假设矩阵 A 的前 k 列有如下形式的 LU 分解 (A_{11} 为 $k-1$ 阶方阵):

$$\begin{bmatrix} A_{11} & \bar{a}_k \\ A_{21} & \hat{a}_k \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & l_k \end{bmatrix} \begin{bmatrix} U_{11} & u_k \\ 0 & u_{kk} \end{bmatrix},$$

其中 l_k 和 u_{kk} 未知待求. 等价于

1. $A_{11} = L_{11} U_{11}$ ($\surd$)
2. $A_{21} = L_{21} U_{11}$ ($\surd$)
3. $\bar{a}_k = L_{11} u_k$ ($\surd$)
4. $\hat{a}_k = L_{21} u_k + u_{kk} l_k$, 选取 u_{kk} 使 l_k 首元素为 1.

其它方式计算 LU 分解: III

再假设矩阵 A 的前 k 行有如下形式的 LU 分解 (A_{11} 为 $k-1$ 阶方阵):

$$\begin{bmatrix} A_{11} & \bar{a}_k & A_{13} \\ \tilde{a}_k^{\mathrm{T}} & a_{kk} & a_k^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ \hat{l}_k^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} U_{11} & u_k & U_{13} \\ 0 & u_{kk} & \hat{u}_k^{\mathrm{T}} \end{bmatrix},$$

其中仅 $\hat{u}_k$ 未知待求.

其它方式计算 LU 分解: III

再假设矩阵 A 的前 k 行有如下形式的 LU 分解 (A_{11} 为 $k-1$ 阶方阵):

$$\begin{bmatrix} A_{11} & \bar{a}_k & A_{13} \\ \tilde{a}_k^{\mathrm{T}} & a_{kk} & a_k^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ \hat{l}_k^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} U_{11} & u_k & U_{13} \\ 0 & u_{kk} & \hat{u}_k^{\mathrm{T}} \end{bmatrix},$$

其中仅 $\hat{u}_k$ 未知待求. 等价于

1. $A_{11} = L_{11}U_{11}$ ($\surd$)
2. $\bar{a}_k = L_{11}u_k$ ($\surd$)
3. $A_{13} = L_{11}U_{13}$ ($\surd$)
4. $\tilde{a}_k^{\mathrm{T}} = \hat{l}_k U_{11}$ ($\surd$)
5. $a_{kk} = \tilde{a}_k^{\mathrm{T}} u_k + u_{kk}$ ($\surd$)
6. $a_k^{\mathrm{T}} = \hat{l}_k^{\mathrm{T}} U_{13} + \hat{u}_k^{\mathrm{T}}$, 计算 $\hat{u}_k$.

与 Gauss 消去法一样, 第 k 步计算得到 L 的第 k 列, U 的第 k 行, 但避免了右下角矩阵块的更新.

1 简介

2 三角形方程组和三角分解

3 选主元三角分解

4 平方根法

5 分块三角分解

6 Others

为什么选主元?

主要有以下原因:

- 主元值为 0. 例如:

$$\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.00 \\ 2.00 \end{bmatrix}.$$

为什么选主元?

主要有以下原因:

- 主元值为 0. 例如:

$$\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.00 \\ 2.00 \end{bmatrix}.$$

- 主元值较小. 例如:

$$\begin{bmatrix} 0.001 & 1.00 \\ 1.00 & 2.00 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.00 \\ 3.00 \end{bmatrix}.$$

为什么选主元?

主要有以下原因:

- 主元值为 0. 例如:

$$\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.00 \\ 2.00 \end{bmatrix}.$$

- 主元值较小. 例如:

$$\begin{bmatrix} 0.001 & 1.00 \\ 1.00 & 2.00 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.00 \\ 3.00 \end{bmatrix}.$$

假设在 3 位 10 进制的浮点数系下解该方程组, 用 Gauss 消去法得

$$\hat{L} = \begin{bmatrix} 1 & 0 \\ 1000 & 1 \end{bmatrix}, \quad \hat{U} = \begin{bmatrix} 0.001 & 1.00 \\ 0 & -1000 \end{bmatrix}.$$

解得 $\hat{x} = (0, 1)^{\mathrm{T}}$, 与精确解 $x = (1.002\cdots, 0.998\cdots)^{\mathrm{T}}$ 相差甚远.

为什么选主元?

如果交换方程的顺序:

- 主元值为 0. 例如:

$$\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.00 \\ 2.00 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.00 \\ 2.00 \end{bmatrix}.$$

为什么选主元?

如果交换方程的顺序:

- 主元值为 0. 例如:

$$\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.00 \\ 2.00 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.00 \\ 2.00 \end{bmatrix}.$$

- 主元值较小. 例如:

$$\begin{bmatrix} 0.001 & 1.00 \\ 1.00 & 2.00 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.00 \\ 3.00 \end{bmatrix} \longrightarrow \begin{bmatrix} 1.00 & 2.00 \\ 0.001 & 1.00 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3.00 \\ 1.00 \end{bmatrix}.$$

为什么选主元?

如果交换方程的顺序:

- 主元值为 0. 例如:

$$\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.00 \\ 2.00 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.00 \\ 2.00 \end{bmatrix}.$$

- 主元值较小. 例如:

$$\begin{bmatrix} 0.001 & 1.00 \\ 1.00 & 2.00 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.00 \\ 3.00 \end{bmatrix} \longrightarrow \begin{bmatrix} 1.00 & 2.00 \\ 0.001 & 1.00 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3.00 \\ 1.00 \end{bmatrix}.$$

假设在 3 位 10 进制的浮点数系下解该方程组, 用 Gauss 消去法得

$$\hat{L} = \begin{bmatrix} 1 & 0 \\ 0.001 & 1 \end{bmatrix}, \quad \hat{U} = \begin{bmatrix} 1 & 2.00 \\ 0 & 1.00 \end{bmatrix}.$$

解得 $\hat{x} = (1.00, 1.00)^{\mathrm{T}}$, 与精确解 $x = (1.002\cdots, 0.998\cdots)^{\mathrm{T}}$ 相当接近了.

初等置换矩阵

将单位矩阵 I 的第 p, q 两列交换得到的矩阵记为初等置换矩阵 I_{pq}, 即

$$I_{pq} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0_{pp} & & 1_{pq} & & \\ & & & \ddots & & & \\ & & 1_{qp} & & 0_{qq} & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}.$$

性质:

- $I_{pq} = I_{pq}^{\mathrm{T}} = I_{pq}^{-1}$
- $\det(I_{pq}) = -1$

全主元 Gauss 消去法

设全主元 Gauss 消去法进行到 r 步终止, 得到初等变换阵 P_k, Q_k 和初等下三角阵 $L_k\,(k=1,\cdots,r)$, 使得

$$L_rP_r\cdots L_1P_1AQ_1\cdots Q_r = U$$

为上三角阵, 令

$$\begin{aligned} Q &= Q_1\cdots Q_r \\ P &= P_r\cdots P_1 \\ L &= P(L_rP_r\cdots L_1P_1)^{-1}, \end{aligned}$$

则有

$$PAQ = LU.$$

全主元 Gauss 消去法

设全主元 Gauss 消去法进行到 r 步终止, 得到初等变换阵 P_k, Q_k 和初等下三角阵 $L_k\,(k=1,\cdots,r)$, 使得

$$L_rP_r\cdots L_1P_1AQ_1\cdots Q_r = U$$

为上三角阵, 令

$$\begin{aligned} Q &= Q_1\cdots Q_r \\ P &= P_r\cdots P_1 \\ L &= P(L_rP_r\cdots L_1P_1)^{-1}, \end{aligned}$$

则有

$$PAQ = LU.$$

定理

设 $A\in\mathbb{R}^{n\times n}$, 则存在排列矩阵 $P, Q\in\mathbb{R}^{n\times n}$, 以及单位下三角矩阵 $L\in\mathbb{R}^{n\times n}$ 和上三角矩阵 $U\in\mathbb{R}^{n\times n}$, 使得

$$PAQ = LU,$$

且 L 的所有元素均满足 $|a_{ij}|<1$, U 的非零对角元的个数恰好等于矩阵 A 的秩.

全主元 Gauss 消去法

Algorithm 8 全主元 Gauss 消去法

for $k = 1$ to $n-1$ **do**
　确定 $p, q\,(k \le p, q \le n)$, 使得 $|A(p,q)| = \max\{|A(i,j)| : i = k:n, j = k:n\}$
　$A(k, 1:n) \leftrightarrow A(p, 1:n)$ (交换第 k 行和第 p 行)
　$A(1:n, k) \leftrightarrow A(1:n, q)$ (交换第 k 列和第 q 列)
　$u(k) = p$ (记录置换矩阵 P_k)
　$v(k) = q$ (记录置换矩阵 Q_k)
　if $A(k,k) \neq 0$ **then**
　　$A(k+1:n, k) = A(k+1:n, k)/A(k,k)$
　　$A(k+1:n, k+1:n) = A(k+1:n, k+1:n) - A(k+1:n, k)A(k, k+1:n)$
　else
　　Stop (矩阵奇异)
　end if
end for

列主元 Gauss 消去法

若 A 非奇异, 全选主元必须进行 $\sum_{k=1}^{n-1}(n-k+1)^2 = \frac{1}{3}n^3 + O(n^3)$ 次两两元素之间的比较和相应的逻辑判断, 为减少比较次数, 提出了列主元 Gauss 消去法.

若 A 非奇异, 全选主元必须进行 $\sum\limits_{k=1}^{n-1}(n-k+1)^2 = \frac{1}{3}n^3 + O(n^3)$ 次两两元素之间的比较和相应的逻辑判断, 为减少比较次数, 提出了列主元 Gauss 消去法.

Algorithm 10 列主元 Gauss 消去法

for $k = 1$ to $n-1$ **do**
- 确定 p, $\cancel{q}$ $(k \le p, \cancel{q} \le n)$, s.t. $|A(p,k)| = \max\{|A(i,k)| : i = k:n, \cancel{j = k:n}\}$
- $A(k,1:n) \leftrightarrow A(p,1:n)$ (交换第 k 行和第 p 行)
- $\cancel{A(1:n,k) \leftrightarrow A(1:n,q)}$ (交换第 k 列和第 q 列)
- $u(k) = p$ (记录置换矩阵 P_k)
- $\cancel{v(k) = q}$ (记录置换矩阵 Q_k)
- **if** $A(k,k) \ne 0$ **then**
  - $A(k+1:n,k) = A(k+1:n,k)/A(k,k)$
  - $A(k+1:n,k+1:n) = A(k+1:n,k+1:n) - A(k+1:n,k)A(k,k+1:n)$
- **else**
  - Stop (矩阵奇异)
- **end if**

end for

例

考虑线性方程组 $Ax=b$, 其中右端项 $b=\mathrm{randn}(n,1)$, 系数矩阵 A 为:

$$A=\begin{bmatrix} 1 & & & & 1 \\ -1 & 1 & & & 1 \\ -1 & -1 & 1 & & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -1 & -1 & \cdots & -1 & 1 \end{bmatrix}.$$

选主元后矩阵 A 的 LU 分解为:

$$L=\begin{bmatrix} 1 & & & & 1 \\ -1 & 1 & & & \\ -1 & -1 & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ -1 & -1 & \cdots & -1 & 1 \end{bmatrix}, U=\begin{bmatrix} 1 & & & & 1 \\ & 1 & & & 2 \\ & & \ddots & & \vdots \\ & & & 1 & 2^{n-2} \\ & & & & 2^{n-1} \end{bmatrix}.$$

计算实例 (反例)

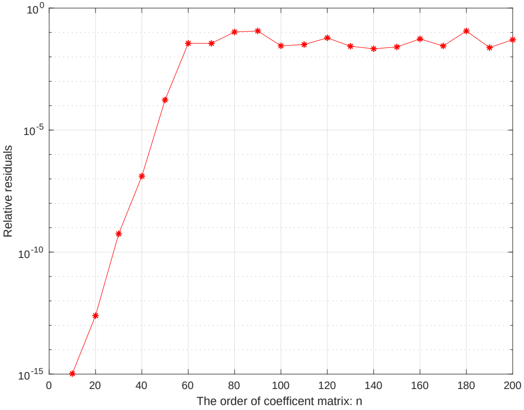


图: 横轴表示问题规模, 纵轴表示相对残差 $\frac{||b-Ax||}{||A||||x||}$

1 简介

2 三角形方程组和三角分解

3 选主元三角分解

4 平方根法

5 分块三角分解

6 Others

对称正定矩阵

定义

设矩阵 A 为 n 阶实对称矩阵, 若对于所有的非零向量 x, 都有 $x^{\mathrm{T}}Ax > 0$, 则称 A 为正定矩阵.

定义

设矩阵 A 为 n 阶实对称矩阵, 若对于所有的非零向量 x, 都有 $x^{\mathrm{T}}Ax > 0$, 则称 A 为正定矩阵.

相关性质及结论:

- A 是正定矩阵 $\Leftrightarrow$ A 的所有顺序主子式均为正

对称正定矩阵

定义

设矩阵 A 为 n 阶实对称矩阵, 若对于所有的非零向量 x, 都有 $x^{\mathrm{T}}Ax > 0$, 则称 A 为正定矩阵.

相关性质及结论:

- A 是正定矩阵 $\Leftrightarrow$ A 的所有顺序主子式均为正
- A 是正定矩阵 $\Leftrightarrow$ A 的所有主子式均为正

对称正定矩阵

定义

设矩阵 A 为 n 阶实对称矩阵, 若对于所有的非零向量 x, 都有 $x^{\mathrm{T}}Ax > 0$, 则称 A 为正定矩阵.

相关性质及结论:

- A 是正定矩阵 $\Leftrightarrow$ A 的所有顺序主子式均为正
- A 是正定矩阵 $\Leftrightarrow$ A 的所有主子式均为正
- A 是正定矩阵 $\Leftrightarrow$ A 的特征值均为正

对称正定矩阵

定义

设矩阵 A 为 n 阶实对称矩阵, 若对于所有的非零向量 x, 都有 $x^{\mathrm{T}}Ax > 0$, 则称 A 为正定矩阵.

相关性质及结论:

- A 是正定矩阵 $\Leftrightarrow$ A 的所有顺序主子式均为正
- A 是正定矩阵 $\Leftrightarrow$ A 的所有主子式均为正
- A 是正定矩阵 $\Leftrightarrow$ A 的特征值均为正
- A 是正定矩阵 $\Leftrightarrow$ 存在主对角线元素全为正的三角矩阵 R, 使 $A = RR^{\mathrm{T}}$

对称正定矩阵

定义

设矩阵 A 为 n 阶实对称矩阵, 若对于所有的非零向量 x, 都有 $x^{\mathrm{T}}Ax > 0$, 则称 A 为正定矩阵.

相关性质及结论:

- A 是正定矩阵 $\Leftrightarrow$ A 的所有顺序主子式均为正
- A 是正定矩阵 $\Leftrightarrow$ A 的所有主子式均为正
- A 是正定矩阵 $\Leftrightarrow$ A 的特征值均为正
- A 是正定矩阵 $\Leftrightarrow$ 存在主对角线元素全为正的三角矩阵 R, 使 $A = RR^{\mathrm{T}}$
- A 是正定矩阵 $\Leftrightarrow$ 存在可逆矩阵 C, 使 $A = C^{\mathrm{T}}C$

对称正定矩阵

定义

设矩阵 A 为 n 阶实对称矩阵, 若对于所有的非零向量 x, 都有 $x^{\mathrm{T}}Ax > 0$, 则称 A 为正定矩阵.

相关性质及结论:

- A 是正定矩阵 $\Leftrightarrow$ A 的所有顺序主子式均为正
- A 是正定矩阵 $\Leftrightarrow$ A 的所有主子式均为正
- A 是正定矩阵 $\Leftrightarrow$ A 的特征值均为正
- A 是正定矩阵 $\Leftrightarrow$ 存在主对角线元素全为正的三角矩阵 R, 使 $A = RR^{\mathrm{T}}$
- A 是正定矩阵 $\Leftrightarrow$ 存在可逆矩阵 C, 使 $A = C^{\mathrm{T}}C$
- A 是正定矩阵 $\Rightarrow$ A 的所有对角元都是正数, 且 $\max\limits_{i\neq j}\{|a_{ij}|\} < \max\limits_{i}\{a_{ii}\}$

Cholesky 分解定理

定理 (Cholesky 分解定理)

若 $A \in \mathbb{R}^{n\times n}$ 对称正定, 则存在一个对角元均为正数的下三角阵 $L \in \mathbb{R}^{n\times n}$, 使得

$$A = LL^{\mathrm{T}}.$$

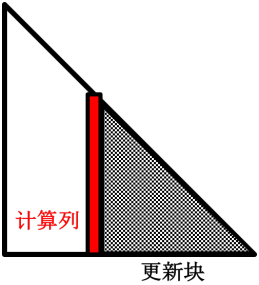
上式称为 Cholesky 分解, 其中 L 称为 A 的 Cholesky 因子.

证明.

略. □

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ & l_{22} & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & l_{nn} \end{bmatrix}.$$

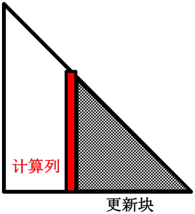
Cholesky 分解: 方式一 (按行计算)



若将矩阵 A 分成 2×2 块, 对其可有如下分解:

$$A = \begin{bmatrix} a_{11} & u^{\mathrm{T}} \\ u & A_1 \end{bmatrix} = \begin{bmatrix} l_{11} & \\ \frac{u}{l_{11}} & I \end{bmatrix} \begin{bmatrix} 1 & \\ & A_1 - \frac{uu^{\mathrm{T}}}{l_{11}^{\mathrm{T}}} \end{bmatrix} \begin{bmatrix} l_{11} & \frac{u^{\ rmT}}{l_{11}} \\ & I \end{bmatrix}.$$

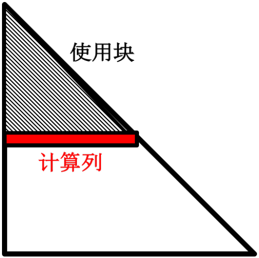
Cholesky 分解: 方式一 (按行计算)



Algorithm 11 平方根法

for $k = 1$ to $n-1$ **do**
  $A(k,k) = \sqrt{A(k,k)}$
  $A(k+1:n,k) = A(k+1:n,k)/A(k,k)$
  for $j = k+1$ to n **do**
    $A(j:n,j) = A(j:n,j) - A(j:n,k)A(j,k)$
  end for
end for
$A(n,n) = \sqrt{A(n,n)}$

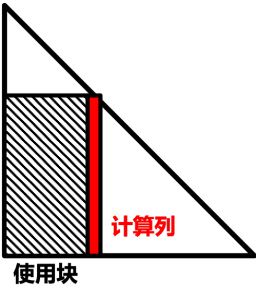
Cholesky 分解: 方式二 (按行计算)



若将矩阵 A 分成 2×2 块, 对其可有如下分解:

$$A_k = \begin{bmatrix} A_{k-1} & u \\ u^{\mathrm{T}} & a_{kk} \end{bmatrix} = \begin{bmatrix} L_{k-1} & \\ v^{\mathrm{T}} & l_{kk} \end{bmatrix} \begin{bmatrix} L_{k-1}^{\mathrm{T}} & v \\ & l_{kk} \end{bmatrix}.$$

Cholesky 分解: 方式三 (按列计算)



课后练习: 编程实现该情形.

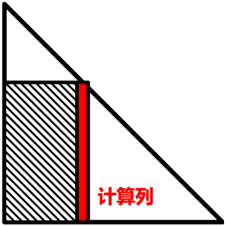
改进的平方根法: LDL^{T} 分解

为避免开放运算, 我们可求 A 如下形式的分解

$$A = LDL^{\mathrm{T}},$$

其中 L 是单位下三角矩阵, D 是对交元素为正数的对角矩阵. 这一分解称作 LDL^{T} 分解.

$$A = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{bmatrix} \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \begin{bmatrix} 1 & l_{21} & \cdots & l_{n1} \\ & 1 & \ddots & \vdots \\ & & \ddots & l_{n,n-1} \\ & & & 1 \end{bmatrix}.$$



Algorithm 12 LDL^{T} 分解

1: **for** $j = 1$ to n **do**
2:     **for** $i = 1$ to $j-1$ **do**
3:         $v(i) = A(j,i)A(i,i)$
4:     **end for**
5:     $A(j,j) = A(j,j) - A(j,1:j-1)v(1:j-1)$
6:     $A(j+1:n,j) = \Big(A(j+1:n,j) - A(j+1:n,j-1)\Big)A(j,j)$
7: **end for**

1 简介

2 三角形方程组和三角分解

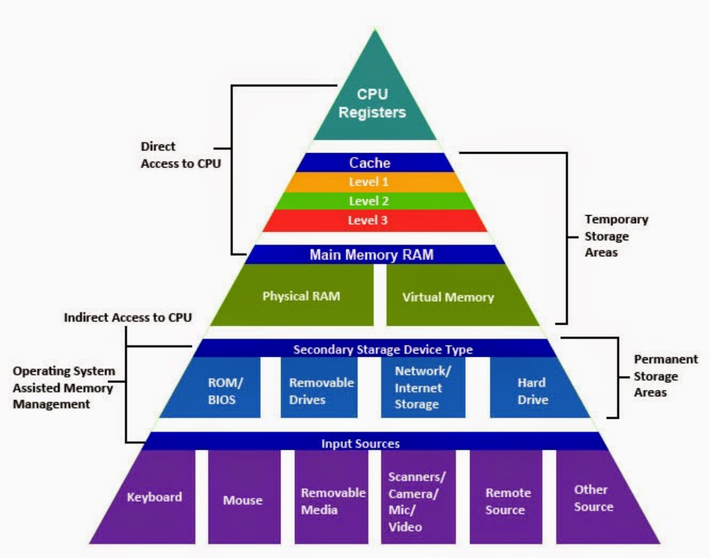
3 选主元三角分解

4 平方根法

5 分块三角分解

6 Others

计算机多级存储结构



CPU
Registers
Direct
Access to CPU
Cache
Level 1
Level 2
Level 3
Temporary
Storage
Areas
Main Memory RAM
Physical RAM
Virtual Memory
Indirect Access to CPU
Secondary Starage Device Type
Operating System
Assisted Memory
Management
ROM/
BIOS
Removable
Drives
Network/
Internet
Storage
Hard
Drive
Permanent
Storage
Areas
Input Sources
Keyboard
Mouse
Removable
Media
Scanners/
Camera/
Mic/
Video
Remote
Source
Other
Source

计算时间模型

例 (简易模型)

假设完成某计算任务的运算量为 f 浮点数, 数据的存取次数为 m 次, 则平均每次可完成 $q = \frac{f}{m}$ 运算. 总的计算时间为

$$f \cdot t_{\text{arith}} + m \cdot t_{\text{mem}} = f \cdot t_{\text{arith}} \left(1 + \frac{t_{\text{mem}}}{q_{\text{arith}}} \right),$$

其中 t_{arith} 表示做一次运算所需要时间, t_{mem} 表示存取一次所需要的时间.

例 (简易模型)

假设完成某计算任务的运算量为 f 浮点数, 数据的存取次数为 m 次, 则平均每次可完成 $q = \frac{f}{m}$ 运算. 总的计算时间为

$$f \cdot t_{\text{arith}} + m \cdot t_{\text{mem}} = f \cdot t_{\text{arith}} \left(1 + \frac{t_{\text{mem}}}{q_{\text{arith}}}\right),$$

其中 t_{arith} 表示做一次运算所需要时间, t_{mem} 表示存取一次所需要的时间.

典型运算	f	m	$q = f/m$
$y = \alpha x + y$	$2n$	$3n$	$2/3$
$y = \alpha Ax + \beta y$	$2n^2$	n^2	2
$C = \alpha AB + \beta C$	$2n^3$	$4n^2$	$n/2$

表: BLAS 三种典型元算的运算次数, 数据存取次数及它们的比值.

Level 1 BLAS

dim scalar vector vector scalars 5-element array		prefixes
SUBROUTINE xROTG (A, B, C, S)	Generate plane rotation	S, D
SUBROUTINE xROTMG(D1, D2, A, B, PARAM)	Generate modified plane rotation	S, D
SUBROUTINE xROT (N, X, INCX, Y, INCY, C, S)	Apply plane rotation	S, D
SUBROUTINE xROTM (N, X, INCX, Y, INCY, PARAM)	Apply modified plane rotation	S, D
SUBROUTINE xSWAP (N, X, INCX, Y, INCY)	$x \leftrightarrow y$	S, D, C, Z
SUBROUTINE xSCAL (N, ALPHA, X, INCX)	$x \leftarrow \alpha x$	S, D, C, Z, CS, ZD
SUBROUTINE xCOPY (N, X, INCX, Y, INCY)	$y \leftarrow x$	S, D, C, Z
SUBROUTINE xAXPY (N, ALPHA, X, INCX, Y, INCY)	$y \leftarrow \alpha x + y$	S, D, C, Z
FUNCTION xDOT (N, X, INCX, Y, INCY)	$dot \leftarrow x^T y$	S, D, DS
FUNCTION xDOTU (N, X, INCX, Y, INCY)	$dot \leftarrow x^T y$	C, Z
FUNCTION xDOTC (N, X, INCX, Y, INCY)	$dot \leftarrow x^H y$	C, Z
FUNCTION xxDOT (N, X, INCX, Y, INCY)	$dot \leftarrow \alpha + x^T y$	SDS
FUNCTION xNRM2 (N, X, INCX)	$nrm2 \leftarrow \|\|x\|\|_2$	S, D, SC, DZ
FUNCTION xASUM (N, X, INCX)	$asum \leftarrow \|\|re(x)\|\|_1 + \|\|im(x)\|\|_1$	S, D, SC, DZ
FUNCTION IxAMAX(N, X, INCX)	$amax \leftarrow 1^{st} k \ni \|re(x_k)\| + \|im(x_k)\| = max(\|re(x_i)\| + \|im(x_i)\|)$	S, D, C, Z

Level 2 BLAS

options dim b-width scalar matrix vector scalar vector		
xGEMV (TRANS, M, N, ALPHA, A, LDA, X, INCX, BETA, Y, INCY)	$y \leftarrow \alpha Ax + \beta y, y \leftarrow \alpha A^T x + \beta y, y \leftarrow \alpha A^H x + \beta y, A - m \times n$	S, D, C, Z
xGBMV (TRANS, M, N, KL, KU, ALPHA, A, LDA, X, INCX, BETA, Y, INCY)	$y \leftarrow \alpha Ax + \beta y, y \leftarrow \alpha A^T x + \beta y, y \leftarrow \alpha A^H x + \beta y, A - m \times n$	S, D, C, Z
xHEMV (UPLO, N, ALPHA, A, LDA, X, INCX, BETA, Y, INCY)	$y \leftarrow \alpha Ax + \beta y$	C, Z
xHBMV (UPLO, N, K, ALPHA, A, LDA, X, INCX, BETA, Y, INCY)	$y \leftarrow \alpha Ax + \beta y$	C, Z
xHPMV (UPLO, N, ALPHA, AP, X, INCX, BETA, Y, INCY)	$y \leftarrow \alpha Ax + \beta y$	C, Z
xSYMV (UPLO, N, ALPHA, A, LDA, X, INCX, BETA, Y, INCY)	$y \leftarrow \alpha Ax + \beta y$	S, D
xSBMV (UPLO, N, K, ALPHA, A, LDA, X, INCX, BETA, Y, INCY)	$y \leftarrow \alpha Ax + \beta y$	S, D
xSPMV (UPLO, N, ALPHA, AP, X, INCX, BETA, Y, INCY)	$y \leftarrow \alpha Ax + \beta y$	S, D
xTRMV (UPLO, TRANS, DIAG, N, A, LDA, X, INCX)	$x \leftarrow Ax, x \leftarrow A^T x, x \leftarrow A^H x$	S, D, C, Z
xTBMV (UPLO, TRANS, DIAG, N, K, A, LDA, X, INCX)	$x \leftarrow Ax, x \leftarrow A^T x, x \leftarrow A^H x$	S, D, C, Z
xTPMV (UPLO, TRANS, DIAG, N, AP, X, INCX)	$x \leftarrow Ax, x \leftarrow A^T x, x \leftarrow A^H x$	S, D, C, Z
xTRSV (UPLO, TRANS, DIAG, N, A, LDA, X, INCX)	$x \leftarrow A^{-1}x, x \leftarrow A^{-T}x, x \leftarrow A^{-H}x$	S, D, C, Z
xTBSV (UPLO, TRANS, DIAG, N, K, A, LDA, X, INCX)	$x \leftarrow A^{-1}x, x \leftarrow A^{-T}x, x \leftarrow A^{-H}x$	S, D, C, Z
xTPSV (UPLO, TRANS, DIAG, N, AP, X, INCX)	$x \leftarrow A^{-1}x, x \leftarrow A^{-T}x, x \leftarrow A^{-H}x$	S, D, C, Z
options dim scalar vector vector matrix		
xGER (M, N, ALPHA, X, INCX, Y, INCY, A, LDA)	$A \leftarrow \alpha xy^T + A, A - m \times n$	S, D
xGERU (M, N, ALPHA, X, INCX, Y, INCY, A, LDA)	$A \leftarrow \alpha xy^T + A, A - m \times n$	C, Z
xGERC (M, N, ALPHA, X, INCX, Y, INCY, A, LDA)	$A \leftarrow \alpha xy^H + A, A - m \times n$	C, Z
xHER (UPLO, N, ALPHA, X, INCX, A, LDA)	$A \leftarrow \alpha xx^H + A$	C, Z
xHPR (UPLO, N, ALPHA, X, INCX, AP)	$A \leftarrow \alpha xx^H + A$	C, Z
xHER2 (UPLO, N, ALPHA, X, INCX, Y, INCY, A, LDA)	$A \leftarrow \alpha xy^H + y(\alpha x)^H + A$	C, Z
xHPR2 (UPLO, N, ALPHA, X, INCX, Y, INCY, AP)	$A \leftarrow \alpha xy^H + y(\alpha x)^H + A$	C, Z
xSYR (UPLO, N, ALPHA, X, INCX, A, LDA)	$A \leftarrow \alpha xx^T + A$	S, D
xSPR (UPLO, N, ALPHA, X, INCX, AP)	$A \leftarrow \alpha xx^T + A$	S, D
xSYR2 (UPLO, N, ALPHA, X, INCX, Y, INCY, A, LDA)	$A \leftarrow \alpha xy^T + \alpha yx^T + A$	S, D
xSPR2 (UPLO, N, ALPHA, X, INCX, Y, INCY, AP)	$A \leftarrow \alpha xy^T + \alpha yx^T + A$	S, D

Level 3 BLAS

options dim scalar matrix matrix scalar matrix		
xGEMM (TRANSA, TRANSB, M, N, K, ALPHA, A, LDA, B, LDB, BETA, C, LDC)	$C \leftarrow \alpha op(A)op(B) + \beta C, op(X) = X, X^T, X^H, C - m \times n$	S, D, C, Z
xSYMM (SIDE, UPLO, M, N, ALPHA, A, LDA, B, LDB, BETA, C, LDC)	$C \leftarrow \alpha AB + \beta C, C \leftarrow \alpha BA + \beta C, C - m \times n, A = A^T$	S, D, C, Z
xHEMM (SIDE, UPLO, M, N, ALPHA, A, LDA, B, LDB, BETA, C, LDC)	$C \leftarrow \alpha AB + \beta C, C \leftarrow \alpha BA + \beta C, C - m \times n, A = A^H$	C, Z
xSYRK (UPLO, TRANS, N, K, ALPHA, A, LDA, BETA, C, LDC)	$C \leftarrow \alpha AA^T + \beta C, C \leftarrow \alpha A^T A + \beta C, C - n \times n$	S, D, C, Z
xHERK (UPLO, TRANS, N, K, ALPHA, A, LDA, BETA, C, LDC)	$C \leftarrow \alpha AA^H + \beta C, C \leftarrow \alpha A^H A + \beta C, C - n \times n$	C, Z
xSYR2K(UPLO, TRANS, N, K, ALPHA, A, LDA, B, LDB, BETA, C, LDC)	$C \leftarrow \alpha AB^T + \bar{\alpha} BA^T + \beta C, C \leftarrow \alpha A^T B + \bar{\alpha} B^T A + \beta C, C - n \times n$	S, D, C, Z
xHER2K(UPLO, TRANS, N, K, ALPHA, A, LDA, B, LDB, BETA, C, LDC)	$C \leftarrow \alpha AB^H + \bar{\alpha} BA^H + \beta C, C \leftarrow \alpha A^H B + \bar{\alpha} B^H A + \beta C, C - n \times n$	C, Z
xTRMM (SIDE, UPLO, TRANSA, DIAG, M, N, ALPHA, A, LDA, B, LDB)	$B \leftarrow \alpha op(A)B, B \leftarrow \alpha Bop(A), op(A) = A, A^T, A^H, B - m \times n$	S, D, C, Z
xTRSM (SIDE, UPLO, TRANSA, DIAG, M, N, ALPHA, A, LDA, B, LDB)	$B \leftarrow \alpha op(A^{-1})B, B \leftarrow \alpha Bop(A^{-1}), op(A) = A, A^T, A^H, B - m \times n$	S, D, C, Z

块 LU 分解 (Block LU Factorization)

设方程组 $Ax = b$ 可划分如下形式:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

其中 A_{11}, A_{22} 为方阵, 且 A_{11} 非奇异.

块 LU 分解 (Block LU Factorization)

设方程组 $Ax = b$ 可划分如下形式:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

其中 A_{11}, A_{22} 为方阵, 且 A_{11} 非奇异.
利用块 Gauss 消去法, 可将矩阵 A 约化为块上三角阵.

$$\begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & S_{22} \end{bmatrix}.$$

块 LU 分解 (Block LU Factorization)

设方程组 $Ax = b$ 可划分如下形式:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

其中 A_{11}, A_{22} 为方阵, 且 A_{11} 非奇异.
利用块 Gauss 消去法, 可将矩阵 A 约化为块上三角阵.

$$\begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & S_{22} \end{bmatrix}.$$

对应于矩阵 A 有如下分解:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & S_{22} \end{bmatrix},$$

其中 $S_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12}$, 称之为 A_{11} 在 A 中的 Schur 余阵 (或 Schur 补).

分割 LU 分解 (Partitioned LU Factorization)

设

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

其中 A_{11}, A_{22} 为方阵, 且 A_{11} 非奇异, 则有如下分解:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix},$$

其中 L_{11}, L_{22} 为单位下三角阵, U_{11}, U_{22} 是上三角阵.

分割 LU 分解 (Partitioned LU Factorization)

设

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

其中 A_{11}, A_{22} 为方阵, 且 A_{11} 非奇异, 则有如下分解:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix},$$

其中 L_{11}, L_{22} 为单位下三角阵, U_{11}, U_{22} 是上三角阵.
可通过以下步骤计算分割 LU 分解:

1. 计算 LU 分解: $A_{11} = L_{11} U_{11}$
2. 解方程组: $U_{11}^{\mathrm{T}} L_{21}^{\mathrm{T}} = A_{21}^{\mathrm{T}}$, $L_{11} U_{12} = A_{12}$
3. 计算 Schur 补: $S_{22} = A_{22} - L_{21} U_{12}$
4. 计算 LU 分解: $S_{22} = L_{22} U_{22}$

1 简介

2 三角形方程组和三角分解

3 选主元三角分解

4 平方根法

5 分块三角分解

6 Others

Strassen's 算法

计算两个矩阵相乘的高效算法. 为简化讨论, 假设矩阵 A, B 为 $2n$ 阶方阵, 为计算 $C = AB$, 将所有矩阵均划分为 2×2 块矩阵, 每个子矩阵均为 n 阶. 即

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

计算两个矩阵相乘的高效算法. 为简化讨论, 假设矩阵 A, B 为 $2n$ 阶方阵, 为计算 $C = AB$, 将所有矩阵均划分为 2×2 块矩阵, 每个子矩阵均为 n 阶. 即

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Strassen's 算法通过以下方式计算矩阵乘积:

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} P_1 + P_4 - P_5 + P_7 & P_3 + P_5 \\ P_2 + P_4 & P_1 + P_3 - P_2 + P_6 \end{bmatrix},$$

Strassen's 算法

计算两个矩阵相乘的高效算法. 为简化讨论, 假设矩阵 A, B 为 $2n$ 阶方阵, 为计算 $C = AB$, 将所有矩阵均划分为 2×2 块矩阵, 每个子矩阵均为 n 阶. 即

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Strassen's 算法通过以下方式计算矩阵乘积:

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} P_1 + P_4 - P_5 + P_7 & P_3 + P_5 \\ P_2 + P_4 & P_1 + P_3 - P_2 + P_6 \end{bmatrix},$$

其中

$$\begin{aligned} P_1 &= (A_{11} + A_{22})(B_{11} + B_{22}), \quad P_2 = (A_{21} + A_{22})B_{11}, \\ P_3 &= A_{11}(B_{12} - B_{22}), \quad P_4 = A_{22}(B_{21} - B_{11}), \\ P_5 &= (A_{11} + A_{12})B_{22}, \quad P_6 = (A_{21} - A_{11})(B_{11} + B_{12}), \\ P_7 &= (A_{12} - A_{22})(B_{21} + B_{22}). \end{aligned}$$

计算两个矩阵相乘的高效算法. 为简化讨论, 假设矩阵 A, B 为 $2n$ 阶方阵, 为计算 $C = AB$, 将所有矩阵均划分为 2×2 块矩阵, 每个子矩阵均为 n 阶. 即

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Strassen's 算法通过以下方式计算矩阵乘积:

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} P_1 + P_4 - P_5 + P_7 & P_3 + P_5 \\ P_2 + P_4 & P_1 + P_3 - P_2 + P_6 \end{bmatrix},$$

其中

$$\begin{aligned}
P_1 &= (A_{11} + A_{22})(B_{11} + B_{22}), \quad P_2 = (A_{21} + A_{22})B_{11}, \\
P_3 &= A_{11}(B_{12} - B_{22}), \quad P_4 = A_{22}(B_{21} - B_{11}), \\
P_5 &= (A_{11} + A_{12})B_{22}, \quad P_6 = (A_{21} - A_{11})(B_{11} + B_{12}), \\
P_7 &= (A_{12} - A_{22})(B_{21} + B_{22}).
\end{aligned}$$

QR 分解定理

定理

设 $A \in \mathbb{C}^{m\times n}(m \geq n)$, 则 A 有 QR 分解

$$A = Q\begin{bmatrix} R \\ 0 \end{bmatrix},$$

其中 $Q \in \mathbb{C}^{m\times m}$ 是正交矩阵, $R \in \mathbb{C}^{n\times n}$ 是具有非负对角元的上三角矩阵; 而且当 $m = n$ 且 A 非奇异时, 上述的分解还是唯一的.

QR 分解定理

定理

设 $A \in \mathbb{C}^{m\times n}(m \geq n)$, 则 A 有 QR 分解

$$A = Q\begin{bmatrix} R \\ 0 \end{bmatrix},$$

其中 $Q \in \mathbb{C}^{m\times m}$ 是正交矩阵, $R \in \mathbb{C}^{n\times n}$ 是具有非负对角元的上三角矩阵; 而且当 $m = n$ 且 A 非奇异时, 上述的分解还是唯一的.

实现方式:

- (Modified) Gram-Schmidt 正交化过程
- Householder 变换
- Givens 变换

带状线性方程组

定义 (带状矩阵)

若方阵 A 的元素 a_{ij} 满足如下条件:

$$a_{ij}=0,\ \ i>j+r,\ \ a_{ij}=0,\ \ j>i+s,$$

其中 r,s 为满足条件的最小整数, 则称 A 为带状矩阵, 其下带宽为 r, 上带宽为 s, 带宽为 $r+s+1$.

带状线性方程组

定义 (带状矩阵)

若方阵 A 的元素 a_{ij} 满足如下条件:

$$a_{ij} = 0, \;\; i > j + r, \;\; a_{ij} = 0, \;\; j > i + s,$$

其中 r, s 为满足条件的最小整数, 则称 A 为带状矩阵, 其下带宽为 r, 上带宽为 s, 带宽为 $r + s + 1$.

例

$$A = \begin{bmatrix} a_{11} & a_{12} & & & \\ a_{21} & a_{22} & a_{23} & & \\ a_{31} & a_{32} & a_{33} & a_{34} & \\ & a_{42} & a_{43} & a_{44} & a_{45} \\ & & a_{53} & a_{54} & a_{55} \end{bmatrix}. \; r = 2, s = 1.$$

带状矩阵的 LU 分解

定理

设带状矩阵 A 的下带宽为 r, 上带宽为 s. 若 A 可进行 *LU* 分解, 即 $A = LU$, 则下三角矩阵 L 的下带宽为 r, 上三角矩阵 U 的上带宽为 s.

带状矩阵的 LU 分解

定理

设带状矩阵 A 的下带宽为 r, 上带宽为 s. 若 A 可进行 LU 分解, 即 $A = LU$, 则下三角矩阵 L 的下带宽为 r, 上三角矩阵 U 的上带宽为 s.

例

设 $A = \begin{bmatrix} a_{11} & a_{12} & & & \\ a_{21} & a_{22} & a_{23} & & \\ a_{31} & a_{32} & a_{33} & a_{34} & \\ & a_{42} & a_{43} & a_{44} & a_{45} \\ & & a_{53} & a_{54} & a_{55} \end{bmatrix}$. 若 A 的 LU 分解存在, 则有

$$A = \begin{bmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ & l_{42} & l_{43} & 1 & \\ & & l_{53} & l_{54} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & & & \\ & u_{22} & u_{23} & & \\ & & u_{33} & u_{34} & \\ & & & u_{44} & u_{45} \\ & & & & u_{55} \end{bmatrix}.$$