《计算物理学》讲义

第二讲　线性代数方程组的直接求解

彭良友

北京大学物理学院现代光学所

Email: liangyou.peng@pku.edu.cn

2019年版

通知

- IEEE-754中的特殊值，例如：有符号的零、NaN、无穷大(∞)。请同学们自行查询IEEE-754标准中的相关规定。
- 成绩考评：网上已经更新为：平时大作业70% + 期末闭卷笔试30%。
- 习题课安排，隔周一次，从下周起的**周**二，5-6 节; 地点：二教109 。具体安排如下(5.1前是单周，5.1后是双周)：
  3.5日（刘）、3.19 日（刘）、4.2 日（刘）、4.16日（刘）、5.7 日（刘、赵）、5.21日（赵）、6.4 日（赵）。
- **学号以14、15开头的同学，务必严肃认真以大量精力投入到本课程学习中。**

Notes:

本讲主要内容

- 引言
- 三角形方程组的解法
- 高斯消元法
- 矩阵的直接LU分解
- 三对角方程组的追赶法
- 列主元消元法
- 高斯－约当消元法
- 列主元LU分解
- 对称正定矩阵的Cholesky分解
- 一般对称正定带状系数矩阵方程组的算法
- 向量和矩阵的范数、方程组的性态和条件数

Notes:

本讲的目的是讨论如下n阶线性方程组的直接求解方法：

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

或写为$\sum\limits_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \cdots, n$。写成矩阵形式为：

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{2}$$

其中，

$$\mathbf{A} = [a_{ij}]_{n\times n}, \ \mathbf{x} = (x_1, \cdots, x_n)^T, \ \mathbf{b} = (b_1, \cdots, b_n)^T, \tag{3}$$

分别称作**系数矩阵**、**解向量**和**右端向量**。本讲义中，我们仅讨论$\mathbf{A}$和$\mathbf{b}$均为实数向量的情况，此时$\mathbf{x}$也为实向量。

Notes:

线性方程组求解的重要性

- **线性方程组的求解**是一个基本而十分重要的的问题。它与**矩阵特征值问题**构成矩阵计算的两大类问题。实际上，很多理论和实际问题都直接或间接的与线性方程组的求解问题相关。在自然科学和社会科学中，很多物理量之间的关系就是线性的，如：牛顿定理$\mathbf{F}=m\mathbf{a}$；欧姆定理$V=IR$；胡克定律$\mathbf{F}=-k\mathbf{x}$。欧姆定理和基尔霍夫定理导出的大型电路方程属于线性方程组，IBM 的大规模集成电路方程可达几十万。
- 在数值计算中，线性方程组的求解显得异常重要，很多计算问题最后归结为线性方程组的求解。譬如某些问题在局部做线性关系近似；微分和积分方程这类非代数问题，离散化后可以用线性代数方程组来近似。可以说，**线性方程组的求解构成了求解各种计算问题数值方法的基础。**
- 很多实际问题的需求也为线性方程组的数值求解提出了新的挑战，例如方程规模庞大，变量众多($>10^6$)、存储问题、算法的稳定性和效率问题、舍入误差传递问题、并行算法的设计等等。

Notes:

线性代数的基本概念回顾

- **线性相关：** 设$\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_m} \in \mathcal{R}^n$，若存在不全为零的数$\alpha_1, \alpha_2, \cdots, \alpha_m \in \mathcal{R}$，使得$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_m \mathbf{x}_m = \mathbf{0}$，则称向量$\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_m}$是线性相关的。若只有$\alpha_i$全为0 时才成立，则称它们是线性无关的。
- **非奇异矩阵：** 是指满足以下等价条件之一的矩阵$\mathbf{A}_{n\times n}$：(1) 存在逆矩阵$\mathbf{A}^{-1}$，满足$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$，$\mathbf{I}$为单位矩阵；(2) 矩阵的行列式$\det(\mathbf{A}) \neq 0$；(3) $\mathbf{A}$ 的秩$\mathrm{rank}(\mathbf{A}) = \mathrm{n}$ (矩阵的秩是其包含的线性无关的行或列的最多个数)；(4) 对任意的非零向量$\mathbf{z} \neq \mathbf{0}$，有$\mathbf{A}\mathbf{z} \neq \mathbf{0}$。
- **方程组$\mathbf{A}\mathbf{x} = \mathbf{b}$解的存在性和唯一性：** (1) 若$\mathbf{A}$为非奇异矩阵，则有唯一的解：$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ (对$\mathbf{b}$没有要求)；(2) 若$\mathbf{A}$ 为 奇异矩阵，且$\mathbf{b} \in \mathrm{span}(\mathbf{A})$，其中集合$\mathrm{span}(\mathbf{A})$表示$\mathbf{A}$ 的各个列向量张成的线性空间，则方程组有无穷多个解；(3) 若$\mathbf{A}$为奇异矩阵，且$\mathbf{b} \notin \mathrm{span}(\mathbf{A})$，则方程组无解。

Notes:

奇异和非奇异矩阵的例子

对方程组$\mathbf{Ax} = \mathbf{b}$:

- 若$A = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}$，则无论$b$取任何值，方程组总有唯一解。
- 但是若$A = \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}$，$A$是奇异矩阵。

  其解的情况依赖于b的取值，但无论如何都不可能有唯一解。

  $b = [4 \quad 7]^T$时，方程组无解；

  而当$b = [4 \quad 8]^T$时，$x = [\gamma \quad \frac{4-2\gamma}{3}]$($\gamma$ 为任意实数)都是方程组的解。

Notes:

如何数值求解线性方程组$\mathbf{Ax} = \mathbf{b}$?

- 如果线性方程组的系数行列式不为零，即$\det(\mathbf{A}) \neq 0$，则该方程组有唯一的解。由克莱姆(Cramer) 法则，其解为：

$$x_i = \frac{\det(A_i)}{\det(A),} \quad (i = 1, 2, \cdots, n). \tag{4}$$

上式中，A_i是被列向量$\mathbf{b}$取代了$\mathbf{A}$的第i列的列向量后得到的矩阵。

- **计算复杂度：** 如果按照这种方法直接计算方程的解，则需要计算$n+1$个n阶行列式，并作n次除法。每个n阶行列式计算需要$(n-1)\times n!$次乘法，计算量将十分可怕。例如，取$n = 30$，将有$\sim 10^{35}$次乘法。因此使用克莱姆法则求线性方程组的解的算法时间复杂度太高，一般没有实际计算价值。
- $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ 需要求矩阵的逆，再做矩阵与矢量乘积，计算复杂度低很多；但是我们要寻求更好的算法。

Notes:

几种特殊矩阵

设$\mathbf{A}=(a_{ij})_{n\times n}\in\mathcal{R}^{n\times n}$，则矩阵$\mathbf{A}$ 为：

- 对角矩阵 (diagonal matrix)：当$i\neq j$ 时，$a_{ij}=0$；
- 三对角矩阵 (tridiagonal matrix)：当$|i-j|>1$时，$a_{ij}=0$；
- 上三角(形)矩阵 (upper triangle matrix)：当$i>j$时，$a_{ij}=0$；
- 下三角(形)矩阵 (lower triangle matrix)：当$i<j$时，$a_{ij}=0$；
- 对称矩阵 (symmetric matrix)：$\mathbf{A}^T=\mathbf{A}$，其中$\mathbf{A}^T$为$\mathbf{A}$的转置；
- 对称正定矩阵 (symmetric positive definite matrix)：$\mathbf{A}^T=\mathbf{A}$，且对任意非零向量$\mathbf{x}\in\mathcal{R}^n$，二次型$\mathbf{x}^T\mathbf{A}\mathbf{x}>0$；
- 对称半正定矩阵 (symmetric positive semidefinite matrix)：$\mathbf{A}^T=\mathbf{A}$，且对任意非零向量$\mathbf{x}\in\mathcal{R}^n$，二次型$\mathbf{x}^T\mathbf{A}\mathbf{x}\geq 0$；
- 正交矩阵 (orthogonal matrix)：$\mathbf{A}^T=\mathbf{A}^{-1}$，其中$\mathbf{A}^{-1}$为$\mathbf{A}$的逆矩阵。

Notes:

几种特殊矩阵

容易看出：

- 对角阵的逆矩阵以及多个对角阵的乘积仍为对角阵；
- 上(下)三角矩阵的逆矩阵、多个上(下)三角矩阵的积仍为上三角矩阵；
- 常称对角元素均为1的上(下)三角矩阵为单位上(下)三角矩阵(U_0, L_0)；
- 正交矩阵的转置也是正交矩阵；且它的行向量、列向量各自构成n 维向量空间中的一组单位正交向量。

对于稀疏矩阵，常用“×”标记非零元素，这种矩阵图称为威尔金森图(Wilkinson graph)。储存和计算时，尽量只考虑非零元素。如下图：

$$\begin{bmatrix} \times & & & \\ & \times & & \\ & & \times & \\ & & & \times \end{bmatrix} \quad \begin{bmatrix} \times & \times & & \\ \times & \times & \times & \\ & \times & \times & \times \\ & & \times & \times \end{bmatrix} \quad \begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{bmatrix} \quad \begin{bmatrix} \times & & & \\ \times & \times & & \\ \times & \times & \times & \\ \times & \times & \times & \times \end{bmatrix}$$

对角阵　三对角阵　上三角阵　下三角阵

Notes:

本讲主要内容

- 引言
- **三角形方程组的解法**
- 高斯消元法
- 矩阵的直接LU分解
- 三对角方程组的追赶法
- 列主元消元法
- 高斯—约当消元法
- 列主元LU分解
- 对称正定矩阵的Cholesky分解
- 一般对称正定带状系数矩阵方程组的算法
- 向量和矩阵的范数、方程组的性态和条件数

Notes:

三角形方程组的求解

最简单的系数矩阵莫过于对角矩阵$\mathbf{Dx} = \mathbf{b}$，方程组的解可以直接写出$x_i = b_i/d_{ii} \quad (i = 1, 2, ..., n)$。三角形方程组的求解也非常简便:

- 下三角方程组：$\mathbf{Lx} = \mathbf{b}$:

$$\begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix}. \tag{5}$$

- 上述方程很容易利用**前代算法**直接求解：(设$l_{ii} \neq 0$，$i = 1, 2, ..., n$)

$$x_1 = b_1/l_{11}, \tag{6}$$

$$x_i = \left(b_i - \sum_{j=1}^{i-1} l_{ij} x_j \right) / l_{ii}, \quad i = 2, 3, ..., n. \tag{7}$$

Notes:

三角形方程组的求解

- 上三角方程组：$\mathbf{Ux} = \mathbf{b}$：

$$\begin{pmatrix} u_{11} & u_{21} & \cdots & u_{n1} \\ & u_{22} & \cdots & u_{n2} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix}. \tag{8}$$

- 上述方程很容易利用**回代算法**直接求解：

$$x_n = b_n/u_{nn}, \tag{9}$$

$$x_i = \left(b_i - \sum_{j=i+1}^{n} u_{ij}x_j\right)/u_{ii}, \quad i = n-1, n-2, ..., 1. \tag{10}$$

上述算法能顺利进行到底的前提是$u_{ii} \neq 0$，$i = 1, 2, ..., n$。

Notes:

三角方程的算法

下三角方程的前代算法

输入：$\mathbf{L}, n, \mathbf{b}$；　输出：$\mathbf{x}$:

For $i = 1, 2, \cdots, n$

　If $l_{ii} = 0$ **then STOP;**

　$x_i := b_i$;

　For $j = 1, 2, \cdots, i-1$

　　$x_i := x_i - l_{ij} x_j$;

　End

　$x_i := x_i / l_{ii}$;

End

上三角方程的回代算法

输入：$\mathbf{U}, n, \mathbf{b}$；　输出：$\mathbf{x}$:

For $i = n, n-1, \cdots, 1$

　If $u_{ii} = 0$ **then STOP;**

　$x_i := b_i$;

　For $j = n, n-1, \cdots, i+1$

　　$x_i := x_i - u_{ij} x_j$;

　End

　$x_i := x_i / u_{ii}$;

End

Notes:

本讲主要内容

- 引言
- 三角形方程组的解法
- **高斯消元法**
- 矩阵的直接LU分解
- 三对角方程组的追赶法
- 列主元消元法
- 高斯—约当消元法
- 列主元LU分解
- 对称正定矩阵的Cholesky分解
- 一般对称正定带状系数矩阵方程组的算法
- 向量和矩阵的范数、方程组的性态和条件数

Notes:

线性方程组的三种基本等价变换

假定给线性方程组的每一个方程标上号：$E_1, E_2, ..., E_n$：

$$\left\{\begin{array}{ll} E_1: & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ E_2: & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ & \cdots \\ E_n: & a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{array}\right. \tag{11}$$

有三种基本等价变换(非奇异线性变换)可用来简化所要求解的方程组：

- $(\lambda E_i) \longrightarrow (E_i)$：即任一方程$E_i$可以与任何非零常数$\lambda$相乘，所得方程用以代替$E_i$；
- $(E_i + \lambda E_j) \longrightarrow (E_i)$：即方程$E_j$可以与任何常数$\lambda$相乘并与方程$E_i$相加，所得方程用以代替$E_i$；
- $(E_i) \longleftrightarrow (E_j)$：即任意两个方程可以交换次序。

Notes:

等价线性变换：转化为同解的、易求的方程组

- **等价线性变换**是贯穿矩阵计算的基本数学思想之一，是指通过等价变换，将矩阵$\mathbf{A}$变为较简单、特殊的矩阵$\mathbf{B}$，使得关于矩阵$\mathbf{B}$的求解变得直接和容易。**所谓等价，是指变换前后的两问题同解**。
- 考虑线性方程组$\mathbf{Ax}=\mathbf{b}$的求解。对方程组做非奇异线性变换，保证变换前后的两方程组同解。设$\mathbf{P}$ 为非奇异矩阵，用数学表示如下：$\mathbf{PAx}=\mathbf{Pb}$，令$\mathbf{PA}=\mathbf{U}$、$\tilde{\mathbf{b}}=\mathbf{Pb}$，则新方程组$\mathbf{Ux}=\tilde{\mathbf{b}}$ 与原方程组同解。**对角矩阵和三角形矩阵使得方程的求解变得直接和容易**。因此研究如何实现$\mathbf{A}$的非奇异上三角化是一般方程求解的核心。
- 再考虑矩阵的本征值问题：$\mathbf{Ax}=\lambda\mathbf{x}$。$\mathbf{A}$ 的阶数很高时，用线性代数方法直接求解特征值方程效率很低。但若$\mathbf{A}$相似于$\mathbf{B}$ $(\mathbf{B}=\mathbf{P}^{-1}\mathbf{AP})$，则两矩阵的特征值相同。若$\mathbf{B}$是对角矩阵或三角形矩阵，其特征值就是对角线上的全部元素。**研究从$\mathbf{A}$到$\mathbf{B}$ 的相似变换是特征值问题的核心**。

Notes:

线性方程组的求解的两大类方法

- **直接解法：** 假定计算过程没有舍入误差的情况下，经过有限步算术运算后能求得线性方程组精确解的方法。但在实际计算中由于舍入误差的影响，也只能求得近似解(但准确度高！)。直接法的核心是高斯消元。
- **迭代解法：** 构造适当的向量序列，用某种极限过程去逐步逼近精确解。例如：雅可比迭代法、高斯-赛德尔迭代法、逐次超松弛迭代法、最速下降法、共轭梯度法等。**迭代法的优点是所需计算机储存单元少，便于编写计算机程序。** 另外，迭代的极限过程一般不可能进行到底，因此只能得到满足一定精度要求的近似解。迭代法的收敛性依赖于系数矩阵的某些特性，为提高效率，往往要与预条件技术相结合。迭代法是高效求解大型稀疏系数矩阵线性方程组，尤其是微分方程离散后得到的大型方程组的一种重要方法。

本讲我们讨论直接法，下一讲，我们讨论迭代法。

Notes:

考虑一个二维向量$\mathbf{a}=[a_1,a_2]^T$，如果$a_1\neq 0$，则

$$\mathbf{M_2a}\equiv\begin{pmatrix}1 & 0\\ -a_2/a_1 & 1\end{pmatrix}\begin{pmatrix}a_1\\ a_2\end{pmatrix}=\begin{pmatrix}a_1\\ 0\end{pmatrix}. \tag{12}$$

更一般，给定n维向量$\mathbf{a}$，下面的变换可以消去第k个元素以后的所有元素：

$$\mathbf{M_ka}=\begin{pmatrix}1 & \cdots & 0 & 0 & \cdots & 0\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & \cdots & 1 & 0 & \cdots & 0\\ 0 & \cdots & -m_{k+1} & 1 & \cdots & 0\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & \cdots & -m_n & 0 & \cdots & 1\end{pmatrix}\begin{pmatrix}a_1\\ \vdots\\ a_k\\ a_{k+1}\\ \vdots\\ a_n\end{pmatrix}=\begin{pmatrix}a_1\\ \vdots\\ a_k\\ 0\\ \vdots\\ 0\end{pmatrix}, \tag{13}$$

其中$m_i=a_i/a_k,(i=k+1,\cdots,n)$，除数$a_k$称为主元。

Notes:

初等消去阵

$\mathbf{M_k}$这类矩阵称为**初等消去阵或高斯变换**。初等消去阵对向量的作用是将第k行后面的行依次加上第k行的某个倍数m_i $(i=k+1,\cdots,n)$，使得这些行的元素变为零。初等消去阵具有如下性质：

- $\mathbf{M_k}$是一个单位下三角矩阵，因而它一定是非奇异的。
- $\mathbf{M_k}=\mathbf{I}-\mathbf{m}\mathbf{e_k^T}$，其中$\mathbf{m}=[0,\cdots,0,m_{k+1},\cdots,m_n]^T$，$\mathbf{e_k}$为单位阵的第$k$列。【注：$\mathbf{m}$为$n\times 1$，列向量；$e_k^T$为$1\times n$，行向量.】
- $\mathbf{M_k^{-1}}=\mathbf{I}+\mathbf{m}\mathbf{e_k^T}\equiv\mathbf{L_k}$, $\mathbf{L_k}$与$\mathbf{M_k}$ 的差别只是乘数的符号相反。
- 若$\mathbf{M_j}(j>k)$为另一消去阵，对应的乘数向量为$\mathbf{t}$。因为$\mathbf{e_k^T}\mathbf{t}=\mathbf{0}$，故：

$$\mathbf{M_k}\mathbf{M_j}=\mathbf{I}-\mathbf{m}\mathbf{e_k^T}-\mathbf{t}\mathbf{e_j^T}+\mathbf{m}\mathbf{e_k^T}\mathbf{t}\mathbf{e_j^T}=\mathbf{I}-\mathbf{m}\mathbf{e_k^T}-\mathbf{t}\mathbf{e_j^T}. \tag{14}$$

  这说明，两个消去阵的乘积相当于它们的“并”。注意乘积的次序，反次序的乘积不一定正确。另外，其逆矩阵的乘积$\mathbf{L_k}\mathbf{L_j}$也有类似结果。

Notes:

高斯消元和回代过程

将n阶线性方程组$\mathbf{Ax}=\mathbf{b}$:
$$\begin{cases} a_{11}\,x_1+a_{12}\,x_2+\cdots+a_{1n}\,x_n=b_1 \\ a_{21}\,x_1+a_{22}\,x_2+\cdots+a_{2n}\,x_n=b_2 \\ \cdots \\ a_{n1}\,x_1+a_{n2}\,x_2+\cdots+a_{nn}\,x_n=b_n \end{cases} \tag{15}$$

变换为同解的三角形方程组$\mathbf{ux}=\mathbf{b}'$：【通过倍乘和倍加变换】

$$\begin{aligned} u_{11}x_1+u_{12}x_2+\cdots+u_{1n}x_n &= b_1' \\ u_{22}x_2+\cdots+u_{2n}x_n &= b_2' \\ \cdots &\quad \cdots \\ u_{nn}x_n &= b_n' \end{aligned} \tag{16}$$

的过程称为**消元过程**；逐次计算出$x_n,x_{n-1},\cdots,x_1$的过程称为**回代过程**。

Notes:

高斯消元法的步骤：第一步

- 原始矩阵的元素用上标$^{(1)}$标记，即：$a_{ij} \to a_{ij}^{(1)}, b_i \to b_i^{(1)}$ 。
- 执行：(第i行) − (第1行) $\times\, a_{i1}^{(1)} \Big/ a_{11}^{(1)} \Longrightarrow$ (新的第i 行)，$(i = 2, \cdots n)$。
- 除第一个方程外，系数和常数项都得到新的值，方程组变为：

$$
\left\{
\begin{array}{l}
a_{11}^{(1)} x_1 + a_{12}^{(1)} x_2 + a_{13}^{(1)} x_3 + \cdots + a_{1n}^{(1)} x_n = b_1^{(1)} \\
\qquad\qquad a_{22}^{(2)} x_2 + a_{23}^{(2)} x_3 + \cdots + a_{2n}^{(2)} x_n = b_2^{(2)} \\
\qquad\qquad a_{32}^{(2)} x_2 + a_{33}^{(2)} x_3 + \cdots + a_{3n}^{(2)} x_n = b_3^{(2)} \\
\qquad\qquad \cdots\cdots \\
\qquad\qquad a_{n2}^{(2)} x_2 + a_{n3}^{(2)} x_3 + \cdots + a_{nn}^{(2)} x_n = b_n^{(2)}
\end{array}
\right. \tag{17}
$$

用矩阵表示为：$\mathbf{A}^{(2)}\mathbf{x} = \mathbf{b}^{(2)}$. 这相当于用初等消去阵$\mathbf{M}_2$作用于系数矩阵的第一个列矢量。

Notes:

高斯消元法的步骤：第二步

- 第一步已经消掉了第一列，现进行第二步消元。对除了第一行和第一列的**子阵**，执行与第一步完全类似的操作(设$a_{22}^{(2)} \neq 0$)，第$(i = 3, \cdots n)$行的与第2 行的相加系数分别是：$-a_{i2}^{(2)}/a_{22}^{(2)}$。【第二行不变】
- 由此得到同解方程组$\mathbf{A}^{(3)}\mathbf{x} = \mathbf{b}^{(3)}$，其中

$$A^{(3)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} \end{bmatrix}, \qquad b^{(3)} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ b_3^{(3)} \\ \vdots \\ b_n^{(3)} \end{bmatrix} \tag{18}$$

- 这就消掉了第二列$a_{22}^{(2)}$以下的元素。得到方程：$\mathbf{A}^{(3)}\mathbf{x} = \mathbf{b}^{(3)}$. 这相当于用初等消去阵$\mathbf{M}_3$作用于第一步之后的系数矩阵的第二个列矢量。

Notes:

- 如此反复直至最后一行，最终得到同解方程组$\mathbf{A}^{(n)}\mathbf{x}=\mathbf{b}^{(n)}$，即

$$\left\{\begin{aligned} a_{11}^{(1)}x_1+a_{12}^{(1)}x_2+a_{13}^{(1)}x_3+\cdots+a_{1n}^{(1)}x_n&=b_1^{(1)}\\ a_{22}^{(2)}x_2+a_{23}^{(2)}x_3+\cdots+a_{2n}^{(2)}x_n&=b_2^{(2)}\\ a_{33}^{(3)}x_3+\cdots+a_{3n}^{(3)}x_n&=b_3^{(3)}\\ \cdots\cdots&\\ a_{nn}^{(n)}x_n&=b_n^{(n)} \end{aligned}\right. \tag{19}$$

- 上述上三角方程组$\mathbf{U}\mathbf{x}=\mathbf{c}$，容易利用前面讨论的回代过程求解：

$$x_n=b_n^{(n)}\Big/a_{nn}^{(n)}, \tag{20}$$

$$x_i=\left(b_i^{(i)}-\sum_{j=i+1}^{n}a_{ij}^{(i)}x_j\right)\Big/a_{ii}^{(i)},\ (i=n-1,n-2,\cdots,1) \tag{21}$$

Notes:

高斯消元法算法

高斯消元法采用的是原地工作的储存方式，不需要额外的储存空间：

输入： $\mathbf{A}, n, \mathbf{b}$;　　　**输出：** $\mathbf{A}, \mathbf{b}$

For $k = 1, 2, \cdots, n-1$

　　If $a_{kk} = 0$ **then STOP;**

　　For $i = k+1, k+2, \cdots, n$

　　　　$c := -a_{ik}/a_{kk}$; **(计算倍乘因子)**

　　　　For $j = k+1, k+2, \cdots, n$

　　　　　　$a_{ij} := a_{ij} + c a_{kj}$; **(更新矩阵元素)**

　　　　End

　　　　$b_i := b_i + c b_k$; **(更新常数项)**

　　End

End

这之后，利用前面讲述的上三角方程的**回代算法**，即可求解$\mathbf{Ax} = \mathbf{b}$的解。

Notes:

高斯消元法的计算复杂度：$O(n^3)$

- 第k步消去过程中，计算倍乘系数需$(n-k)$次乘法；计算$a_{ij}^{(k+1)}$需要乘法和加法均为$(n-k)^2$次；计算$b_i^{(k+1)}$需要乘法和加法均为$n-k$次。消去过程中的乘除法总数为：
$\sum_{k=1}^{n-1}[(n-k)^2+2(n-k)]=\frac{1}{3}n^3+\frac{1}{2}n^2-\frac{5}{6}n$；
加减法总数为：$\sum_{k=1}^{n-1}[(n-k)^2+(n-k)]=\frac{1}{3}n^3-\frac{1}{3}n$.
- 回代过程中，乘除法总数：$1+\sum_{i=1}^{n-1}((n-i)+1)=\frac{n^2+n}{2}$；加减法总数为：$\sum_{i=1}^{n-1}((n-i-1)+1)=\frac{n^2-n}{2}$.
- **求解过程的总操作次数为：**$\frac{2}{3}n^3+\frac{3}{2}n^2-\frac{7}{6}n$，**时间复杂度为**$O(n^3)$。

Notes:

课堂练习：回代过程

请求解如下上三角方程组：

$$\begin{aligned} 4x_1 + 5x_2 + 6x_3 &= 10 \\ 2x_2 + 3x_3 &= 3 \\ 7x_3 &= 7 \end{aligned} \tag{22}$$

其增广矩阵为：

$$\begin{bmatrix} 4 & 5 & 6 & 10 \\ 0 & 2 & 3 & 3 \\ 0 & 0 & 7 & 7 \end{bmatrix} \tag{23}$$

答案是：$x_1 = 1, x_2 = 0, x_3 = 1.$

Notes:

课堂练习：整个高斯消元和回代过程

请求解下述方程组：

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ x_1 + 3x_2 + 2x_3 &= 6 \\ x_1 + 2x_2 + 2x_3 &= 5 \end{aligned} \tag{24}$$

其增广矩阵为：

$$\begin{bmatrix} 2 & 1 & 1 & 4 \\ 1 & 3 & 2 & 6 \\ 1 & 2 & 2 & 5 \end{bmatrix} \tag{25}$$

答案是：$x_1 = 1, x_2 = 1, x_3 = 1.$

Notes:

本讲主要内容

- 引言
- 三角形方程组的解法
- 高斯消元法
- **矩阵的直接LU分解**
- 三对角方程组的追赶法
- 列主元消元法
- 高斯一约当消元法
- 列主元LU分解
- 对称正定矩阵的Cholesky分解
- 一般对称正定带状系数矩阵方程组的算法
- 向量和矩阵的范数、方程组的性态和条件数

Notes:

矩阵的LU分解

- 回顾一下顺序高斯消元的过程，其每一步(即每一列)消去过程相当于左乘一个初等变换矩阵$\mathbf{L_k}$【亦即我们在前面用$\mathbf{M_k}$表示的初等消去阵】。
- 譬如，第一步$\mathbf{A}^{(2)} = \mathbf{L_1}\,\mathbf{A}^{(1)}$，$\mathbf{b}^{(2)} = \mathbf{L_1}\,\mathbf{b}^{(1)}$，其中

$$\mathrm{L}_1 = \begin{pmatrix} 1 & & & & \\ -l_{21} & 1 & & & \\ -l_{31} & 0 & 1 & & \\ & \cdots & \cdots & & \\ -l_{n1} & 0 & 0 & \cdots & 1 \end{pmatrix}, \qquad \begin{array}{l} l_{i1} = a_{i1}^{(1)} \Big/ a_{11}^{(1)}, \\ i = 2, 3, \cdots, n. \end{array} \tag{26}$$

Notes:

- 第二步也类似，$\mathbf{A}^{(3)} = \mathbf{L_2}\,\mathbf{A}^{(2)}$，$\mathbf{b}^{(3)} = \mathbf{L_2}\,\mathbf{b}^{(2)}$，其中

$$L_2 = \begin{pmatrix} 1 & & & & \\ 0 & 1 & & & \\ 0 & -l_{32} & 1 & & \\ & \cdots & \cdots & & \\ 0 & -l_{n2} & 0 & \cdots & 1 \end{pmatrix}, \qquad \begin{array}{l} l_{i2} = a_{i2}^{(2)} \Big/ a_{22}^{(2)}, \\ i = 3, 4, \cdots, n. \end{array} \tag{27}$$

- 依此继续直到将$\mathbf{A}$变成一个上三角矩阵$\mathbf{U}$，右端矢量$\mathbf{b}$也做相应变换：

$$\begin{array}{l} \mathbf{A}^{(n)} = L_{n-1}\,L_{n-2}\cdots L_1\,\mathbf{A}^{(1)} \\ \mathbf{b}^{(n)} = L_{n-1}\,L_{n-2}\cdots L_1\,\mathbf{b}^{(1)} \end{array} \Rightarrow \begin{array}{l} \mathbf{A}^{(1)} = L_1^{-1}\,L_2^{-1}\cdots L_{n-2}^{-1}\,L_{n-1}^{-1}\,\mathbf{A}^{(n)} \\ \mathbf{b}^{(1)} = L_1^{-1}\,L_2^{-1}\cdots L_{n-2}^{-1}\,L_{n-1}^{-1}\,\mathbf{b}^{(n)} \end{array} \tag{28}$$

因此，**高斯消元的过程中，我们将矩阵分解成了一个下三角和上三角矩阵的乘积：** $\mathbf{A} = \mathbf{L}^{-1}\mathbf{U}$，$\mathbf{L}^{-1} = L_1^{-1}\,L_2^{-1}\cdots L_{n-2}^{-1}\,L_{n-1}^{-1}$.

Notes:

- 我们了解过，多个下三角矩阵的乘积仍为下三角矩阵。而且，如果下三角矩阵可逆，则其逆矩阵仍然是下三角矩阵。重要的是，下三角矩阵的乘积和逆矩阵很容易求得。
- 前面提到过，初等消去阵的逆L_i^{-1}与L_i的区别，仅仅是乘数的符号相反，它也是一个第i类的初等消去矩阵：

$$L_i = \begin{pmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ \vdots & & 1 & & & \\ 0 & \cdots & -l_{i+1,i} & 1 & & \\ \vdots & & \vdots & \vdots & \ddots & \\ 0 & & -l_{ni} & & & 1 \end{pmatrix}, \; L_{\mathrm{i}}^{-1} = \begin{pmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ \vdots & & \ddots & & & \\ \vdots & & & 1 & & \\ 0 & & & l_{i+1,i} & 1 & \\ \vdots & & & \vdots & & \ddots & \\ 0 & & & l_{ni} & & & 1 \end{pmatrix}.$$

Notes:

L与其逆矩阵L^{-1}

- 多个**按照顺序相乘**的消去矩阵，其乘积结果的非零元素是各个i因子矩阵非零元素的并集。由高斯消元法得到的矩阵L**是单位下三角矩阵**。
- 因此，若记$\mathbf{L}^{-1} = L_1^{-1} L_2^{-1} \cdots L_{n-2}^{-1} L_{n-1}^{-1}$，我们有：

$$\mathbf{L}^{-1} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix} \tag{30}$$

注意，$\mathbf{L} = L_{n-1} L_{n-2} \cdots L_1$ 在形式上没有这么简单。

- 课堂练习：

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

Notes:

- 高斯消元求解方程组的过程，也是矩阵**A**的LU分解过程。
- 实际上，我们也可以根据矩阵乘法直接进行三角(LU)分解。可以证明：**如果A的各阶顺序主子式均不为零(或者等价的说，各顺序主子阵$\mathbf{A}_k$均非奇异)**，则它可以唯一的分解成一个单位下三角矩阵$\mathbf{L}_0$和一个非奇异的三角矩阵**U**的乘积。这种分解称为**杜利特尔(Doolittle)分解**。如果上三角矩阵是$\mathbf{U}_0$，则**L**是一般下三角矩阵，称为**克劳特(Crout)分解**。

  但是，如果**L**、**U**均为一般三角矩阵，LU分解不是唯一的，例如：

$$\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 0 & -5 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} -2 & -3 \\ 0 & -5/3 \end{bmatrix} \tag{31}$$

Notes:

杜利特尔分解和克洛特分解用矩阵分别表示为:

$$
\mathbf{A} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{bmatrix} = \mathbf{L_0 U}
$$

$$
= \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & \cdots & u_{1n} \\ & 1 & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix} = \mathbf{L U_0}
$$

Notes:

注意，非奇异矩阵不一定存在LU分解，例如：$\mathbf{A}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$。实际上，我们有下面的定理：

【定理】对方程线性方程组$\mathbf{Ax}=\mathbf{b}$，其中$\mathbf{A}\in R^{n\times n}$，则矩阵$\mathbf{A}$存在唯一的LU分解的充分必要条件是执行高斯消元的整个过程中，始终有主元$a_{kk}^{(k)}\neq 0$，$(k=1,2,\cdots,n-1)$。

Notes:

矩阵的直接*LU*分解的用途

- 矩阵分解后($\mathbf{A} = \mathbf{LU}$)，求解方程$\mathbf{Ax} = \mathbf{b}$相当于：$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ $= \mathbf{U}^{-1}\mathbf{L}^{-1}\mathbf{b}$. 可以分解为两步：$\mathbf{Ly} = \mathbf{b}$及$\mathbf{Ux} = \mathbf{y}$，它两分别通过前代和回代过程非常快捷的完成求解。LU 分解可以非常高效的求解多右端方程组问题。例如，变更m 次$\mathbf{b}$，总计算量约为$n^3/3 + mn^2$。
- **可用来求解A的逆矩阵。** 可以直接先求三角矩阵L和U的逆矩阵，然后代入：$\mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}$求出$\mathbf{A}$的逆矩阵。也可以将单位矩阵$\mathbf{I}$分解成$n$个列向量作为右端向量，然后用前面讨论过的求解线性方程的方法解出逆矩阵的列向量，然后拼起来。后者的复杂度是$O(n^2)$，比高斯法优越。
- **可用来计算矩阵的行列式。** 矩阵L和U可以用来快速地计算矩阵A的行列式，因为$\det(\mathbf{A}) = \det(\mathbf{L})\det(\mathbf{U})$，而三角矩阵的行列式就是对角线元素的乘积。若$\mathbf{L}$ 是单位三角矩阵，则$\det(\mathbf{A}) = \det(\mathbf{L})\det(\mathbf{U})$
$$= \prod_{i=1}^{n} u_{ii}.$$

Notes:

基于高斯消去法的直接LU分解算法

输入：$\mathbf{A}, n$；输出：$\mathbf{A}$　【L和U覆盖原来的$\mathbf{A}$】

- DO $k = 1, 2, \cdots, n-1$
- 　IF $a_{kk} = 0$ THEN STOP
- 　DO $i = k+1, k+2, \cdots, n$　(计算L的第k列)
- 　　DO $j = 1, 2, \cdots, k-1$
- 　　　$a_{ik} := a_{ik} - a_{jk}a_{ij}$;
- 　　ENDDO
- 　　$a_{ik} := a_{ik}/a_{kk}$;
- 　ENDDO
- 　DO $j = k+1, k+2, \cdots, n$　(计算U的第$k+1$列)
- 　　DO $i = 1, 2, \cdots, k$
- 　　　$a_{k+1,j} := a_{k+1,j} - a_{k+1,i}a_{ij}$;
- 　　ENDDO
- 　ENDDO
- ENDDO

Notes:

矩阵的直接LU分解：两个例子

- 例一：$\mathbf{A}=\begin{bmatrix}1&2&2\\4&4&2\\4&6&4\end{bmatrix}=\begin{bmatrix}1&0&0\\4&1&0\\4&1/2&1\end{bmatrix}\begin{bmatrix}1&2&2\\0&-4&-6\\0&0&-1\end{bmatrix}=\mathbf{LU}$
- 例二：求方程组的解：$\begin{bmatrix}1&2&3\\2&5&2\\3&1&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}14\\18\\20\end{bmatrix}$

A的LU分解为：$\mathbf{A}=\begin{bmatrix}1&0&0\\2&1&0\\3&-5&1\end{bmatrix}\begin{bmatrix}1&2&3\\0&1&-4\\0&0&-24\end{bmatrix}=\mathbf{LU}.$

解两个三角方程组：
$$\mathbf{Ly}=(14,18,20)^T \implies \mathbf{y}=(14,-10,-72)^T,$$
$$\mathbf{Ux}=(14,-10,-72)^T \implies \mathbf{x}=(1,2,3)^T.$$

Notes:

对角占优矩阵

对于n阶方阵$\mathbf{A}=(a_{ij})_{n\times n}$，若

- $|a_{ii}|\geq\sum\limits_{j=1,j\neq i}^{n}|a_{ij}|\ (i=1,2,...,n)$，且至少有一个不等式严格成立，则矩阵A按行对角占优；
- $|a_{jj}|\geq\sum\limits_{i=1,i\neq j}^{n}|a_{ij}|\ (j=1,2,...,n)$，且至少有一个不等式严格成立，则矩阵$\mathbf{A}$ 按列对角占优；
- $|a_{ii}|>\sum\limits_{j=1,j\neq i}^{n}|a_{ij}|\ (i=1,2,...,n)$，则矩阵$\mathbf{A}$按行严格对角占优；
- $|a_{jj}|>\sum\limits_{i=1,i\neq j}^{n}|a_{ij}|\ (j=1,2,...,n)$，则矩阵$\mathbf{A}$按列严格对角占优。

前两者都称为**对角占优矩阵(或弱对角占优矩阵)**，后两者一般称为**严格对角占优矩阵**。

Notes:

对角占优矩阵在高斯消去算法中稳定

若n阶方阵$\mathbf{A} = (a_{ij})_{n\times n}$为对角占优矩阵，则：

- 如果$\mathbf{A}$是严格对角占优，则$\mathbf{A}$必为非奇异矩阵。而且，此时对其进行直接顺序LU分解是数值稳定的。
- 如果$\mathbf{A}$是非奇异矩阵，则对其进行直接的LU分解是数值稳定的。

最后说明一点，对带状矩阵A做LU分解时，结果矩阵的L和U的非零元素仍然分布在A的原始带宽之内，因此分解和回代过程可以得到高效的算法。但是，带状矩阵的逆矩阵一般不能保留原始矩阵的稀疏特点，通常$\mathbf{A}, \mathbf{L}, \mathbf{U}$的逆矩阵均为稠密矩阵。这再次说明，我们在求解线性方程组时，要避免求逆矩阵，不能按照$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$来求解方程组。

Notes:

本讲主要内容

- 引言
- 三角形方程组的解法
- 高斯消元法
- 矩阵的直接LU分解
- **三对角方程组的追赶法**
- 列主元消元法
- 高斯－约当消元法
- 列主元LU分解
- 对称正定矩阵的Cholesky分解
- 一般对称正定带状系数矩阵方程组的算法
- 向量和矩阵的范数、方程组的性态和条件数

Notes:

追赶法(托马斯法）：解三对角线性方程组

三对角线性方程组: $\mathbf{Ax} = \mathbf{f}$，矩阵形式为：

$$\begin{bmatrix} b_1 & c_1 & & & & \\ a_2 & b_2 & c_2 & & & \\ & a_3 & b_3 & c_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ \vdots \\ f_{n-1} \\ f_n \end{bmatrix} \tag{32}$$

系数矩阵为三对角矩阵，非零元素分布在主对角线及其相邻两条次对角线上。储存时，可以用三个向量$\mathbf{a}, \mathbf{b}, \mathbf{c}$ 分别表示非零元素：

$\mathbf{a} = (a_2, a_3, ..., a_n)$； $\mathbf{b} = (b_1, b_2, b_3, ..., b_n)$； $\mathbf{c} = (c_1, c_2, c_3, ..., c_{n-1})$。

我们要寻求三对角方程组的最高效的算法。

Notes:

对系数矩阵$\mathbf{A}$进行克劳特分解$\mathbf{A} = \mathbf{LU_0}$:

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

[L:下二对角矩阵,U_0:单位上二对角矩阵]

$$= \begin{bmatrix} l_1 & & & & & \\ \alpha_2 & l_2 & & & & \\ & \alpha_3 & l_3 & & & \\ & & \ddots & \ddots & & \\ & & & \alpha_{n-1} & l_{n-1} & \\ & & & & \alpha_n & l_n \end{bmatrix} \begin{bmatrix} 1 & u_1 & & & & \\ & 1 & u_2 & & & \\ & & 1 & u_3 & & \\ & & & \ddots & \ddots & \\ & & & & 1 & u_{n-1} \\ & & & & & 1 \end{bmatrix}.$$

Notes:

求解各个元素

利用矩阵乘法，可得：$a_i = \alpha_i \ (i = 1, 2, \cdots, n)$，$b_1 = l_1$，以及

$$\begin{cases} c_1 = l_1 u_1 \\ c_2 = l_2 u_2 \\ \cdots \\ c_{n-1} = l_{n-1} u_{n-1} \end{cases} \quad (34) \qquad \begin{cases} b_2 = a_2 u_1 + l_2 \\ b_3 = a_3 u_2 + l_3 \\ \cdots \\ b_n = a_n u_{n-1} + l_n \end{cases} \quad (35)$$

也就是说对$i = 1, 2, \cdots, n-1$，有：

$$\begin{cases} b_1 = l_1 \\ c_i = l_i u_i, \\ b_{i+1} = a_{i+1} u_i + l_{i+1}, \end{cases} \Longrightarrow \begin{cases} l_1 = b_1 \\ u_i = \frac{c_i}{l_i}, i = 1, 2, \cdots, n-1 \\ l_{i+1} = b_{i+1} - a_{i+1} u_i, \end{cases} \quad (36)$$

依次算出：$l_1 \to u_1 \to l_2 \to u_2 \to \cdots l_n$。前提是$l_i \neq 0$，则LU$_0$ 唯一确定。

Notes:

追赶法顺利实施的条件

- 系数行列式$\det \mathbf{A} = \det \mathbf{L} \det \mathbf{U}_0 = l_1 l_2 \cdots l_n$。进行克劳特分解的条件是三角形矩阵的各阶顺序主子式不为零。但对三对角矩阵分解的条件是$l_i \neq 0,\ i = 1, 2, \cdots, n$。当系数矩阵的三对角元素非零，首行和末行严格对角占优，其余行弱对角占优，则$l_i \neq 0,\ i = 1, 2, \cdots, n$；或系数矩阵为严格对角占优时$l_i \neq 0,\ i = 1, 2, \cdots, n$。
- **定理：** 若三对角矩阵满足条件

$$\begin{cases} |b_1| > |c_1| > 0 \\ |b_k| \geq |a_k| + |c_k| > 0, \quad a_k c_k \neq 0 \ (k = 2, 3, \cdots, n-1), \\ |b_n| > |a_n| > 0 \end{cases} \tag{37}$$

则其所有顺序主子矩阵均非奇异，即能满足$l_i \neq 0,\ i = 1, 2, \cdots, n$。
- **由实际问题产生的带状方程组，其系数矩阵常常是对角占优或者是对称正定的，因此追赶法一般能达到数值稳定性的要求。**

Notes:

三对角方程组的追赶算法

前面的讨论是基于$\mathbf{A}=\mathbf{L}\mathbf{U}_0$的，而下面给出的基于$\mathbf{A}=\mathbf{L}_0\mathbf{U}$的真实算法：

输入： $n,\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{f}$； **输出：** $\mathbf{x}$:

For $i=2,3,\cdots,n$

$m_i:=a_i/b_{i-1};$

$b_i:=b_i-m_ic_{i-1};$

$f_i:=f_i-m_if_{i-1};$

END

$x_n:=f_n/b_n;$

For $i=n-1,n-2,\cdots,1$

$x_i:=(f_i-c_ix_{i+1})/b_i;$

End

Notes:

本讲主要内容

- 引言
- 三角形方程组的解法
- 高斯消元法
- 矩阵的直接LU分解
- 三对角方程组的追赶法
- **列主元消元法**
- 高斯－约当消元法
- 列主元LU分解
- 对称正定矩阵的Cholesky分解
- 一般对称正定带状系数矩阵方程组的算法
- 向量和矩阵的范数、方程组的性态和条件数

Notes:

列主元

- 前面讨论的高斯消元法是按照方程给定的自然顺序逐一进行的，当方程组系数矩阵的顺序主子式$\mathbf{D}_k$ $(k = 1, 2, ..., n-1)$ 都不为零时，该算法能顺利进行。
- 但在消元过程中，可能会出现$a_{kk}^{(k)} = 0$的情况让消元无法继续进行；另外，即使$a_{kk}^{(k)} \neq 0$ 但非常小，我们将其做除数，也会导致其它元素数量级的严重增长和舍入误差的快速扩散和传播。
- 下面的方程无法直接进行高斯消元，系数矩阵$\mathbf{A}$也无法进行$\mathbf{LU}$分解：

$$2x_2 = 2; \quad 3x_1 + x_2 = 4。$$

- 设矩阵$\mathbf{A}$是n阶方阵，如果$\mathbf{A}$的1到$n-1$阶主子式都非零，那么矩阵$\mathbf{A}$存在$\mathbf{LU}$分解；如果矩阵$\mathbf{A}$存在$\mathbf{LU}$分解且$\mathbf{A}$ 非奇异，那么$\mathbf{LU}$分解唯一。

Notes:

列主元

为避免高斯消元法中断或者误差的快速传播，可以通过交换方程的次序，选取绝对值大的元素作为主元，这就是**选主消去法**。

为避免$l_{ik}=a_{ik}^{(k)}/a_{kk}^{(k)}$过大，可有两种选择：

- $a_{kk}^{(k)}=\max\left\{a_{ik}^{(k)},i=k,k+1,\cdots,n\right\}$，称为**部分主元或列主元消去法**，即选取当前列里未消去部分的最大元素作为主元，交换那一行与当前行（$E_j\leftrightarrow E_k$）。
- $a_{kk}^{(k)}=\max\left\{a_{ij}^{(k)},i,j=k,k+1,\cdots,n\right\}$，称为**全主元消去法**，即在未消去的子矩阵中所有元素里中选取最大的一个作为主元，将其通过行、列交换到当前对角元的位置。注意，列交换后需要相应交换解向量中的对应分量，以保证解不变。全主元消去算法很稳定，但计算量要比列主元法大很多。

Notes:

求解三阶方程组：$\begin{bmatrix} 0.001 & 2.000 & 3.000 \\ -1.000 & 3.712 & 4.623 \\ -2.000 & 1.072 & 5.643 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1.000 \\ 2.000 \\ 3.000 \end{bmatrix}$。其四位有效数字的精确解为：$\mathbf{x} = (-0.4904, -0.05104, 0.3675)^T$.

【**解**：】(1) 顺序高斯消去法，考虑增广矩阵：

$$(A|b) = \begin{bmatrix} 0.001 & 2.000 & 3.000 & 1.000 \\ -1.000 & 3.712 & 4.623 & 2.000 \\ -2.000 & 1.072 & 5.643 & 3.000 \end{bmatrix} \Longrightarrow$$

$$\begin{bmatrix} 0.001 & 2.000 & 3.000 & 1.000 \\ 0 & 2004 & 3005 & 1002 \\ 0 & 4001 & 6006 & 2003 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0.001 & 2.000 & 3.000 & 1.000 \\ 0 & 2004 & 3005 & 1002 \\ 0 & 0 & 5.000 & 2.000 \end{bmatrix}$$

回代得到：$\bar{\mathbf{x}} = (-0.4000, -0.09989, 0.4000)^T$.

Notes:

【解：】(2) 列主元消去法：

$$(A|b) = \begin{bmatrix} -2.000 & 1.072 & 5.643 & 3.000 \\ -1.000 & 3.712 & 4.623 & 2.000 \\ 0.001 & 2.000 & 3.000 & 1.000 \end{bmatrix} \Longrightarrow$$

$$\begin{bmatrix} -2.000 & 1.072 & 5.643 & 3.000 \\ 0 & 3.712 & 1.801 & 0.500 \\ 0 & 2.001 & 3.003 & 1.002 \end{bmatrix} \Longrightarrow \begin{bmatrix} -2.000 & 1.072 & 5.643 & 3.000 \\ 0 & 3.712 & 1.801 & 0.500 \\ 0 & 0 & 1.868 & 0.687 \end{bmatrix}$$

回代得到：$\overline{\mathbf{x}}' = (-0.4900, -0.05113, 0.3678)^T$.

与四位有效数字的精确解$\mathbf{x} = (-0.4904, -0.05104, 0.3675)^T$很接近。

此例充分说明了列主消元法的必要性和有效性。

Notes:

列主元法：课堂作业

试分别利用列主元法和顺序消元法求解三阶方程组：

$$\begin{aligned}
0.50x_1 + 1.1x_2 + 3.1x_3 &= 6.0 \\
2.0x_1 + 4.5x_2 + 0.36x_3 &= 0.020 \\
5.0x_1 + 0.96x_2 + 6.5x_3 &= 0.96
\end{aligned}$$

【为减小误差，计算过程中请保留3位有效数字】。

其两位有效数字的精确解分别为：

$$\begin{aligned}
\mathbf{x} &= (-2.60, 1.00, 2.00)^T, \\
\mathbf{x}^* &= (-5.80, 2.40, 2.02)^T.
\end{aligned}$$

Notes:

可以利用列主元法来求矩阵的行列式。设有如下矩阵**A**：

$$\begin{pmatrix} a_{11} \cdots a_{1n} \\ \cdots\cdots\cdots \\ a_{n1} \cdots a_{nn} \end{pmatrix} \tag{38}$$

通过m**次行交换**的列主元消去法将其变成上三角矩阵：

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn}^{(n)} \end{pmatrix} \tag{39}$$

则**A**的行列式为：

$$\det \mathbf{A} = (-1)^m a_{11}^{(1)} a_{22}^{(2)} \cdots a_{nn}^{(n)}. \tag{40}$$

Notes:

列主元消元法算法

用列主元消元法算求解n阶线性方程组$\mathbf{A}\mathbf{x}=\mathbf{b}$。我们可以顺便算出矩阵$\mathbf{A}$的行列式值，存为det。

- $\det := 1$ 【行列式赋初值1】
- DO $k = 1, 2, \cdots, n-1$ 【对第二行及以后行消元】
  - (1) $|a_{i_k k}| = \max\limits_{k\le i\le n} |a_{ik}|.$ 【按列选主元】
  - (2) IF $a_{i_k k} = 0$ THEN $\det := 0$ STOP
  - (3) IF $i_k = k$ GOTO (4);

    ELSE $a_{kj} \Leftrightarrow a_{i_k j}$ $(j = k, k+1, \cdots, n)$ $b_k \Leftrightarrow b_{i_k}$ $\det := -\det$ ENDIF
  - (4) $a_{ik} \leftarrow l_{ik} = a_{ik}/a_{kk}$ $(i = k+1, \cdots, n)$ $(|l_{ik}| \le 1)$ 【计算倍乘数】
  - (5) $a_{ij} := a_{ij} - l_{ik}a_{kj} \quad (i, j = k+1, \cdots, n)$

    $b_i := b_i - l_{ik}b_k \quad (i = k+1, \cdots, n)$ 【消元】
  - (6) $\det := a_{kk}\det$. 【行列式计算】

Notes:

- ENDDO
- 若$a_{nn}=0$，则STOP【因为此时$\det(A)=0$】
- 回代求解：
$$b_n := b_n/a_{nn};$$
$$b_i := (b_i - \sum_{j=i+1}^{n} a_{ij})/a_{ii}\ \ (i=n-1,n-2,\cdots,1).$$
- $\det := a_{nn}\det.$ 【行列式乘上最后一行的对角元】

注意，在上述算法汇总，为了节省储存空间，消元结果冲掉$\mathbf{A}$，倍乘因子l_{ij}冲掉a_{ij}，最后的解向量$\mathbf{x}$冲掉右端矢量$\mathbf{b}$。

Notes:

列主元法求矩阵的行列式：例子

请用列主元法求矩阵的行列式：

$$\begin{pmatrix} 12 & -3 & 3 \\ -18 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix} \tag{41}$$

答案为：$(-1)^2 \times (-18) \times \frac{7}{6} \times \frac{22}{7} = -66.$

Notes:

请用四位有效算术求解方程组：

$$30.00x_1 + 591400x_2 = 591700, \quad (42)$$

$$5.291x_1 - 6.130x_2 = 46.78 \quad (43)$$

你们的答案：$x_1 = -10.00, x_2 = 1.001$.

正确答案：$x_1 = 10.00, x_2 = 1.000$.

这需要利用**比例因子列选主元**技术：

- 在每一个方程的系数里，找到最大元素：$s_i = \max\limits_{1\le k\le n} a_{ik}$;
- 若对某个i，有$s_i = 0$，则方程组没有唯一解。否则，我们来选取把哪个方程放在第一位：$\frac{|a_{p1}|}{s_p} = \max\limits_{1\le k\le n} \frac{|a_{k1}|}{s_k}$，然后进行行交换$1 \Leftrightarrow p$。
- 对剩下的$n-1$个方程做类似的操作。

Notes:

比例因子列选主元：例子

请用三位舍入算术求解方程组：

$$2.11x_1 - 4.21x_2 + 0.921x_3 = 2.01 \tag{44}$$

$$4.01x_1 + 10.2x_2 - 1.12x_3 = 3.09 \tag{45}$$

$$1.09x_1 + 0.987x_2 + 0.832x_3 = 4.21 \tag{46}$$

正确答案：$x_1 = 0.431, x_2 = 0.430, x_3 = 5.12.$

Notes:

本讲主要内容

- 引言
- 三角形方程组的解法
- 高斯消元法
- 矩阵的直接LU分解
- 三对角方程组的追赶法
- 列主元消元法
- **高斯－约当消元法**
- 列主元LU分解
- 对称正定矩阵的Cholesky分解
- 一般对称正定带状系数矩阵方程组的算法
- 向量和矩阵的范数、方程组的性态和条件数

Notes:

- 它是高斯消元法的一种修正，其基本方法与高斯消元法相同，区别是同时消去矩阵对角线**下方**和**上方**的元素，并将矩阵**A约化为单位矩阵**。也称为**无回代消去法**，直接给出方程组的解。
- 高斯一约当消元算法的计算量：乘除法总数为：$\frac{n^3}{2}+n^2-\frac{n}{2}$，加减法总数为：$\frac{n^3}{2}-\frac{n}{2}$，总计算量为：$n^3+n^2-n$。相比于高斯消元法的运算总次数$\frac{2}{3}n^3+\frac{3}{2}n^2-\frac{7}{6}n$要大很多。

Notes:

列主元法的G-J算法

DO $k=1,2,\cdots,n$

- (1). 选取i_k，使得$|a_{i_k k}| = \max\limits_{k\le i\le n} |a_{ik}|$.
- (2). IF $\quad a_{i_k k}=0 \quad$ STOP

  IF $i_k=k$ GOTO (3) ELSE

  $a_{kj}\leftrightarrow a_{i_k j},\quad j=k,k+1,\cdots,n;\quad b_k\leftrightarrow b_{i_k}.$

  ENDIF
- (3). 计算乘数：$a_{ik}\leftarrow l_{ik}=-\frac{a_{ik}}{a_{kk}}$, $i=1,2,\cdots,n$, $i\neq k$;

  $$a_{kk}\leftarrow l_{kk}=\frac{1}{a_{kk}}.$$
- (4). 消元：$a_{ij}\leftarrow a_{ij}+l_{ik}a_{kj}$, $i=1,2,\cdots,n$ 且$i\neq k$, $j=k+1,\cdots,n$

  $$b_i\leftarrow b_i+l_{ik}b_k,\ i=1,2,\cdots,n,\ 且 i\neq k$$
- (5). 计算主行(第k行)：$a_{kj}\leftarrow a_{kj}l_{kk}$, $j=k,k+1,\cdots,n$;

  $b_k\leftarrow b_k l_{kk}$.

ENDDO 【最终的b即为方程的解。】

Notes:

高斯一约当(*Gauss-Jordan*)消元法：例子

$$\begin{cases} 4x_1+3x_2+x_3=11 \\ 2x_1+x_2+2x_3=6 \\ 6x_1+x_2+5x_3=13 \end{cases} \text{选主} \Longrightarrow \begin{cases} \color{red}{6}x_1+x_2+5x_3=13 \\ 2x_1+x_2+2x_3=6 \\ 4x_1+3x_2+x_3=11 \end{cases} \text{归一、消元} \Longrightarrow$$

$$\begin{cases} x_1+\frac{1}{6}x_2+\frac{5}{6}x_3=\frac{13}{6} \\ 0x_1+\frac{2}{3}x_2+\frac{1}{3}x_3=\frac{5}{3} \\ 0x_1+\frac{7}{3}x_2-\frac{7}{3}x_3=\frac{7}{3} \end{cases} \text{选主} \Longrightarrow \begin{cases} x_1+\frac{1}{6}x_2+\frac{5}{6}x_3=\frac{13}{6} \\ 0x_1+\color{red}{\frac{7}{3}}x_2-\frac{7}{3}x_3=\frac{7}{3} \\ 0x_1+\frac{2}{3}x_2+\frac{1}{3}x_3=\frac{5}{3} \end{cases} \text{归一} \Longrightarrow$$

$$\begin{cases} x_1+\frac{1}{6}x_2+\frac{5}{6}x_3=\frac{13}{6} \\ 0x_1+x_2-x_3=1 \\ 0x_1+\frac{2}{3}x_2+\frac{1}{3}x_3=\frac{5}{3} \end{cases} \text{消元} \Longrightarrow \begin{cases} x_1+0x_2+x_3=2 \\ 0x_1+x_2-x_3=1 \\ 0x_1+0x_2+\color{red}{1}\cdot x_3=1 \end{cases} \text{消元} \Longrightarrow$$

$$\begin{cases} x_1+0x_2+0x_3=1 \\ 0x_1+x_2+0x_3=2 \\ 0x_1+0x_2+x_3=1 \end{cases}$$

Notes:

G-J消元法求逆矩阵

- G-J法的主要用途是求解逆矩阵$\mathbf{A}^{-1}$. 【例】求$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$的$\mathbf{A}^{-1}$。

$$(A|I_n)=\begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 4 & 5 & 0 & 1 & 0 \\ 3 & 5 & 6 & 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & 5 & 6 & 0 & 0 & 1 \\ 2 & 4 & 5 & 0 & 1 & 0 \\ 1 & 2 & 3 & 1 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 5/3 & 2 & 0 & 0 & 1/3 \\ 0 & 2/3 & 1 & 0 & 1 & -2/3 \\ 0 & 1/3 & 1 & 1 & 0 & -1/3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1/2 & 0 & -5/2 & 2 \\ 0 & 1 & 3/2 & 0 & 3/2 & -1 \\ 0 & 0 & 1/2 & 1 & -1/2 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & -3 & 2 \\ 0 & 1 & 0 & -3 & 3 & -1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{bmatrix} = \left(I_n|A^{-1}\right).$$

Notes:

本讲主要内容

Notes:

部分主元高斯消去法进行*LU*分解

- 为了算法的稳定性，我们通常采用部分主元高斯消去法求解方程组。在LU 分解时，我们同样可以采用部分主元法。由于我们有行的交换，因此，我们最终是将$\mathbf{PA}$分解了，即$\mathbf{PA}=\mathbf{LU}$，其中$\mathbf{P}$是一系列初等交换阵的乘积，是对单位矩阵做行交换得到的矩阵，通常称为**排列阵或者置换阵**。容易证明，$\mathbf{P}$ 是正交矩阵，即有$\mathbf{P}^{-1}=\mathbf{P}^{\mathbf{T}}$。由于其每行只有一个非零元素1，我们通常用一个长度为n的整型数组$\mathbf{p}$来表示.
- 采用原地工作的储存方式，L、U覆盖原矩阵$\mathbf{A}$，$\mathbf{p}$记录选主元时行交换的结果，亦即交换矩阵的第k和s行，相当于交换$\mathbf{p}$的第k 和s个储存单元的值。最终，$\mathbf{p}$代表了一种新的行排列顺序，以它调整单位阵即得到矩阵$\mathbf{P}$。

Notes:

部分主元高斯消去法进行LU分解的算法

输入：$\mathbf{A}, n$；输出：$\mathbf{A}$，一维数组p　【L和U覆盖原来的$\mathbf{A}$】

- $p = [1, 2, 3, \cdots, n]$
- DO $k = 1, 2, \cdots, n-1$
- 　　确定满足$|a_{sk}| = \max_{k \le i \le n} |a_{ik}|$的$s$值；
- 　　IF $s \neq k$ THEN
- 　　　　交换矩阵$\mathbf{A}$的第k和s行；
- 　　　　交换p的第k和s个元素；
- 　　ENDIF
- 　　DO $i = k+1, k+2, \cdots, n$
- 　　　　$a_{ik} := a_{ik}/a_{kk}$;　　【即l_{ik}】
- 　　　　DO $j = k+1, k+2, \cdots, n$
- 　　　　　$a_{ij} := a_{ij} - a_{ik}a_{kj}$;
- 　　　　ENDDO
- 　　ENDDO
- ENDDO

Notes:

简单例子：$A = \begin{bmatrix} 1 & 2 & 2 \\ 4 & 4 & 2 \\ 4 & 6 & 4 \end{bmatrix}$.

【解：】初始化$p=[1,2,3]$. 第一列中，最大元素为4，因此交换第1、2行。得到$A = \begin{bmatrix} 4 & 4 & 2 \\ 1 & 2 & 2 \\ 4 & 6 & 4 \end{bmatrix}$，同时，$p=[2,1,3]$. 消去矩阵第一列对角线以下的元素得到$A = \begin{bmatrix} 4 & 4 & 2 \\ 0 & 1 & 3/2 \\ 0 & 2 & 2 \end{bmatrix}$,同时将倍乘数$l_{21}=-1/4, l_{31}=-1$的相反数填入到刚刚消去的位置$\begin{bmatrix} 4 & 4 & 2 \\ 1/4 & 1 & 3/2 \\ 1 & 2 & 2 \end{bmatrix}$.

Notes:

部分主元高斯消去法进行LU分解：例子

【解：】第二列及以下最大元素为2，因此交换第2、3行，得到：

$$\begin{bmatrix} 4 & 4 & 2 \\ 1 & 2 & 2 \\ 1/4 & 1 & 3/2 \end{bmatrix}$$

，同时，$p = [2, 3, 1]$。消去第二列对角线下方的元素，同时将倍乘数$l_{32} = -1/2$的相反数填入到相应位置，得到：

$$\begin{bmatrix} 4 & 4 & 2 \\ 1 & 2 & 2 \\ 1/4 & 1/2 & 1/2 \end{bmatrix}$$

。此二维数组中，分别储存了矩阵L和U，分别是:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \qquad U = \begin{bmatrix} 4 & 4 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 1/2 \end{bmatrix}. \tag{47}$$

Notes:

【解：】需要注意的是，我们并不是LU分解的原矩阵**A**，而是排列阵**P**作用后的**PA**。排列阵**P**可以通过数组**p**对单位阵的**行重排列**得到。这个例子中，最终我们有$\mathbf{p} = [2, 3, 1]$，因此，

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix},$$

容易验证：

$$\mathbf{PA} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 4 & 4 & 2 \\ 4 & 6 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 4 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 1/2 \end{bmatrix} = LU.$$

我们再次看到，采用部分选主元技术时，我们得到的L矩阵对角线下方的元素(即倍乘因子)的绝对值都不超过1，避免了把较小的数做为除数带来的舍入误差扩大。

Notes:

部分主元高斯消去法进行LU分解的计算量和储存

- 从上面的讨论看出，与直接的LU分解算法相比，我们增加了$n-1$次求最大值的运算，以及若干次的矩阵的行交换。
- 与直接的LU分解算法相比，储存量只增加了一个长度为n的整型数组。
- 实际上，矩阵行的交换也可以不显式的进行，也就是说不交换二维数组的元素，而是通过数组$\mathbf{p}$跟踪新的行次序，将取矩阵第i行、j列元素的操作，由$A(i,j)$变为$A(\mathbf{p}[i],j)$即可。最后，利用数组$\mathbf{p}$得到所需要的L和U。不显式进行矩阵的行交换使得数据的移动量最小化，在某些情况下，可能使得算法的执行效率更高。
- 利用部分主元的LU分解，也可以方便的求解线性方程组$\mathbf{Ax}=\mathbf{b}$: $\mathbf{x}=\mathbf{A}^{-1}\mathbf{b}=\mathbf{U}^{-1}\mathbf{L}^{-1}\mathbf{Pb}$。也就是，用$\mathbf{P}$将$\mathbf{b}$重排后再利用前代和回代过程求解。这种选列主元的算法一般不会出现算法的中断，同时有效减小舍入误差，也适合多右端【如更换m次$\mathbf{b}$】方程组的问题。

Notes:

全主元LU分解

- 全主元技术在第k步消去过程中，在未消去的子矩阵所有元素中选取最大的一个，通过行以及列的交换将其交换到当前的对角元素位置$a_{kk}^{(k)}$。与列选主元技术相比，为了保证解的不变性，我们需要在系数矩阵的列交换同时，相应交换解向量的对应分量。全主元法对应的LU分解形式为：$\mathbf{PAQ} = \mathbf{LU}$, 其中，$\mathbf{P}$和$\mathbf{Q}$分别是对$\mathbf{A}$的行和列进行重排的排列阵。在列主元分解的算法上，增加一个记录列交换的数组$\mathbf{q}$，就能得到全主元消去法的LU分解算法。
- 相应的，求解方程$\mathbf{Ax} = \mathbf{b}$变为：$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \mathbf{QU}^{-1}\mathbf{L}^{-1}\mathbf{Pb}$.
- 全主元法虽然理论上极为稳定，数值误差更小，但是与部分主元比，寻找主元的计算量大大增加，因此在实际应用中，主要还是使用部分主元(列主元)消去法求解一般矩阵的线性方程组。一般情况下，部分主元已足够稳定，数值误差也在可控范围内。

Notes:

关于LU分解过程中行列交换合法性的证明

在处理完第k行后，我们有$\mathbf{P}_k\mathbf{A} = \mathbf{L}_k\mathbf{U}_K$，其中根据LU分解的特性，$\mathbf{L}_k$和$\mathbf{U}_k$有以下形式：

$$\mathbf{L}_k = \begin{pmatrix} \mathbf{L}_k^{(0)} & \\ \mathbf{L}_{k1} & \mathbf{I}_{n-k} \end{pmatrix} \tag{1}$$

$$\mathbf{U}_k = \begin{pmatrix} \mathbf{U}_k^{(0)} & \mathbf{U}_{k1} \\ & \mathbf{U}_{k2} \end{pmatrix} \tag{2}$$

其中，$\mathbf{L}_k^{(0)}$为$k \times k$单位下三角矩阵，$\mathbf{U}_k^0$为$k \times k$上三角矩阵。

关于LU分解过程中行列交换合法性的证明

直接相乘得到

$$\mathbf{P}_k\mathbf{A} = \mathbf{L}_k\mathbf{U}_k = \begin{pmatrix} \mathbf{L}_k^{(0)}\mathbf{U}_k^{(0)} & \mathbf{L}_k^{(0)}\mathbf{U}_{k1} \\ \mathbf{L}_{k1}\mathbf{U}_{k1} & \mathbf{U}_{k2} \end{pmatrix} \tag{3}$$

所谓“将**L**和**U**拼成一个矩阵”得到的是

$$\mathbf{B}_k = \begin{pmatrix} \mathbf{L}_k^{(0)} - \mathbf{I}_k + \mathbf{U}_k^{(0)} & \mathbf{U}_{k1} \\ \mathbf{L}_{k1} & \mathbf{U}_{k2} \end{pmatrix} \tag{4}$$

我们要证明的是：对$\mathbf{B}_k$做行交换$k+1 \leftrightarrow p \quad (p > k+1)$，其得到的**L**和**U**乘积即为$\mathbf{P}_k\mathbf{A}$做了相同行交换的结果

关于LU分解过程中行列交换合法性的证明

将行变换$k+1 \leftrightarrow p \quad (p > k+1)$对应的的初等矩阵记为

$$\mathbf{P}'_{k+1} = \begin{pmatrix} \mathbf{I}_k & \\ & \mathbf{P}_{k+1} \end{pmatrix} \tag{5}$$

则作用后有

$$\mathbf{B}'_k = \mathbf{P}'_{k+1}\mathbf{B}_k = \begin{pmatrix} \mathbf{L}_k^{(0)} - \mathbf{I}_k + \mathbf{U}_k^{(0)} & \mathbf{U}_{k1} \\ \mathbf{P}_{k+1}\mathbf{L}_{k1} & \mathbf{P}_{k+1}\mathbf{U}_{k2} \end{pmatrix} \tag{6}$$

关于LU分解过程中行列交换合法性的证明

对应有

$$\mathbf{L}'_k = \begin{pmatrix} \mathbf{L}_k^{(0)} & \\ \mathbf{P}_{k+1}\mathbf{L}_{k1} & \mathbf{I}_{n-k} \end{pmatrix} \tag{7}$$

$$\mathbf{U}'_k = \begin{pmatrix} \mathbf{U}_k^{(0)} & \mathbf{U}_{k1} \\ & \mathbf{P}_{k+1}\mathbf{U}_{k2} \end{pmatrix} \tag{8}$$

将二者相乘得到

$$\mathbf{A}' = \mathbf{L}'_k\mathbf{U}'_k = \begin{pmatrix} \mathbf{L}_k^{(0)}\mathbf{U}_k^{(0)} & \mathbf{L}_k^{(0)}\mathbf{U}_{k1} \\ \mathbf{P}_{k+1}\mathbf{L}_{k1}\mathbf{U}_{k1} & \mathbf{P}_{k+1}\mathbf{U}_{k2} \end{pmatrix} \tag{9}$$

比较(3)与(9)，发现确有

$$\mathbf{A}' = \mathbf{P}'_{k+1}\mathbf{P}_k\mathbf{A} \tag{10}$$

命题证毕。

关于LU分解过程中行列交换合法性的证明

这一结论对于普遍情况下的矩阵相乘显然是不成立的。成立来源于：

- $\mathbf{L}$和$\mathbf{U}$各占据$\mathbf{B}$后$n-k$行的一个矩形部分
- 乘法中在左的$\mathbf{L}$直接对应成立
- 乘法中在右的$\mathbf{U}$所对应的矩阵块与一个单位矩阵相乘，从而也成立

对于将$\mathbf{B}$右乘列交换矩阵$\mathbf{Q}$，直接由(4)可以看出，这一操作相当于对$\mathbf{U}_{k1}$和$\mathbf{U}_{k2}$做了对应的列变换，再代入(3)即证得。

本讲主要内容

- 引言
- 三角形方程组的解法
- 高斯消元法
- 矩阵的直接LU分解
- 三对角方程组的追赶法
- 列主元消元法
- 高斯—约当消元法
- 列主元LU分解
- **对称正定矩阵的Cholesky分解**
- 一般对称正定带状系数矩阵方程组的算法
- 向量和矩阵的范数、方程组的性态和条件数

Notes:

对称矩阵的LU分解

对称矩阵$\mathbf{A}$，有$a_{ji}=a_{ij}$，因此在存储上节约近一半的空间。

- 对于对称矩阵，将其分解为LU后，其中的U矩阵可以写为对角阵乘以一个单位上三角矩阵的形式：

$$U=\begin{bmatrix} u_{11} & & & \\ & u_{22} & & \\ & & \cdots & \\ & & & u_{nn} \end{bmatrix}\begin{bmatrix} 1 & u_{12}/u_{11} & \cdots & u_{1n}/u_{11} \\ & 1 & \cdots & u_{2n}/u_{22} \\ & & \cdots & \cdots \\ & & & 1 \end{bmatrix}=DU_0 \quad (48)$$

其中U_0为单位上三角阵，D为对角阵，因此我们有$A=LDU_0$。因A为对称阵，故$A=A^T=U_0^TDL^T$，U_0^T为单位下三角阵，DL^T为上三角阵。根据LU分解的唯一性，我们得到$U_0^T=L$，因此$\mathbf{A}=\mathbf{LDL^T}$。

Notes:

对称正定矩阵的LU分解

- **定理：**若$\mathbf{A}$为n阶对称阵，且其顺序主子式$\mathbf{D}_k \neq 0, k = 1, ..., n-1$，则$\mathbf{A}$可以唯一的分解为

$$\mathbf{A} = \mathbf{L_0 D L_0^T}. \tag{49}$$

其中，$\mathbf{L}_0$为单位下三角阵，$\mathbf{D}$为对角阵。

注意，当矩阵$\mathbf{A}$为奇异时，$\mathbf{D}$的最后一个对角元素为零。

- 若$\mathbf{A}$为**对称正定矩阵**，易知对角阵$\mathbf{D}$的对角元素$u_{ii} > 0$，$(i = 1, ..., n)$。这是因为对任意的$\mathbf{x} \neq 0$，设$\mathbf{x} = \mathbf{L_0^T y}$，由$\mathbf{L_0^T}$的非奇异性，可推知$\mathbf{y} \neq 0$，故$\mathbf{x^T D x} = \mathbf{y^T L_0 D L_0^T y} = \mathbf{y^T A y} > \mathbf{0}$，即$\mathbf{D}$为对称正定矩阵，其对角元素一定大于0。

Notes:

对称正定矩阵的LU分解

- 设

$$\mathbf{D}^{1/2} = \begin{bmatrix} \sqrt{u_{11}} & & & \\ & \sqrt{u_{22}} & & \\ & & \cdots & \\ & & & \sqrt{u_{nn}} \end{bmatrix}, \tag{50}$$

则$\mathbf{A} = \mathbf{L_0 D^{1/2} D^{1/2} L_0^T} = L_1 L_1^T$，$\mathbf{L_1}$为下三角矩阵，且对角元素为$\sqrt{u_{11}}, \cdots, \sqrt{u_{nn}}$，均大于零。【任意$\mathbf{L_0 D}$必为下三角矩阵。】

- **对称正定矩阵的Cholesky分解：** 若$\mathbf{A}$为实对称正定矩阵，则存在非奇异下三角阵$\mathbf{L}$，使得$\mathbf{A} = \mathbf{LL^T}$，其中$\mathbf{L}$的对角元素均大于零。满足这种条件的分解是唯一的。当然，$\mathbf{A}$也可以唯一分解为$\mathbf{A} = \mathbf{R^T R}$，$\mathbf{R}$为上三角矩阵。这称为对称正定矩阵的Cholesky分解(乔里斯基、楚列斯基)。

Notes:

对称正定矩阵的性质

古尔维兹定理: 实对称矩阵正定的充要条件是它的各阶顺序主子式为正。

设$\mathbf{A} \in \mathcal{R}^{n\times n}$为对称正定矩阵，则有如下性质：

1. $\mathbf{A}$非奇异，且$\mathbf{A}^{-1}$ 也为对称正定矩阵。
2. $\mathbf{A}$的顺序主子阵$\mathbf{A}_k$ $(k = 1, ..., n)$也是对称正定矩阵，其中

$$\mathbf{A}_k = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}.$$

3. $\mathbf{A}$的特征值λ_i为大于零的实数。
4. $\mathbf{A}$的顺序主子式都大于零，即$\det \mathbf{A}_k > 0$, $k = 1, 2, ..., n$.
5. $\mathbf{A}$的对角元素$a_{ii} > 0$, $i = 1, \cdots, n$。

Notes:

对称正定矩阵的LU分解

对称正定矩阵的Cholesky分解的算法：平方根法。将$A = LL^T$展开：

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

其中$a_{ij} = a_{ji}$。按照矩阵乘法，对$i = 1, 2, \cdots, n$，有

$$a_{ij} = a_{ji} = \sum_{k=1}^{n} l_{jk} l_{ik} = \sum_{k=1}^{i} l_{jk} l_{ik} = \sum_{k=1}^{i-1} l_{jk} l_{ik} + l_{ji} l_{ii}, j = i, i+1, \cdots, n \quad (51)$$

故有算法：$l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2}, l_{ji} = \left(a_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik} \right) / l_{ii}, j = i, \cdots, n$

Notes:

输入： $\mathbf{A}, n$；　　　　**输出：** $\mathbf{A}$:

For $j = 1, 2, \cdots, n$

　For $k = 1, 2, \cdots, j-1$

　　$a_{ij} := a_{jj} - a_{jk}^2$

　END

　$a_{jj} := \sqrt{a_{jj}}$;　　$\left\{ \text{注：}\ l_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2} \right\}$

　For $i = j+1, j+2, \cdots, n$

　　For $k = 1, 2, \cdots, j-1$

　　　$a_{ij} := a_{ij} - a_{ik}a_{jk}$

　　END

　　$a_{ij} := a_{ij}/a_{jj}$;　　$\left\{ \text{注：}\ l_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik}l_{jk} \right) /l_{jj} \right\}$

　END

END

Notes:

对称正定矩阵的*Cholesky*分解的稳定性及应用

- 对于对称正定矩阵，上面讨论的Cholesky分解算法是稳定的，不会中断，也不必像高斯消去法那样必须考虑选主元的问题。可以证明，对称正定矩阵的直接LU分解和Cholesky分解都是数值稳定的。相应分析请参见有关书籍，例如喻文健的《数值分析与算法》，p95。
- 上面的算法本身很容易判断一个对称矩阵是否是正定的：如果算法发生中断（被开方数小于或者等于0），则$\mathbf{A}$一定不是正定的；如果算法执行顺利，直至最终$a_{nn} \neq 0$，则对称矩阵$\mathbf{A}$一定是正定的。上面算法的储存量和计算量，都大约是一般矩阵LU分解算法的一半。
- 对称正定矩阵的Cholesky分解后，$\mathbf{A} = \mathbf{L}\mathbf{L}^\mathbf{T}$，求解$\mathbf{A}\mathbf{x} = \mathbf{b}$时，只需要依次执行前面所讲的前代$(\mathbf{L}\mathbf{y} = \mathbf{b})$和回代$(\mathbf{L}^\mathbf{T}\mathbf{x} = \mathbf{y})$过程即可求得方程组的解。

Notes:

改进平方根法

将对称正定矩阵$\mathbf{A}$，进行$\mathbf{A}=\mathbf{LDL}^T$分解，可避免开方运算，其中$\mathbf{D}=\mathrm{diag}(d_i)$，且$d_i>0$，$L$ 为单位下三角矩阵，有

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

$$= \begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix} \begin{pmatrix} 1 & l_{21} & \cdots & l_{n1} \\ & 1 & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix} \quad (52)$$

Notes:

由矩阵乘法，当$i \geq j$ 时，有

$$a_{ij} = \sum_{k=1}^{j} l_{ik} d_k l_{kj} = \sum_{k=1}^{j-1} l_{ik} d_k l_{jk} + l_{ij} d_j, \qquad l_{jj} = 1 \tag{53}$$

于是，对$i = 1, 2, ..., n$，有

$$l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik} d_k l_{jk} \frac{1}{d_j} \; (j = 1, 2, ..., i-1) \tag{54}$$

$$d_i = a_{ii} - \sum_{k=1}^{j-1} l_{ik}^2 d_k. \tag{55}$$

上述算法，虽然避免了开方运算，但在计算每个元素时多了相乘的因子，因此乘积计算量比LL^T分解约大了一倍，但分析计算L和D元素的公式可以看出，许多计算是重复的。

Notes:

改进平方根法

为了避免重复计算，作如下的变换$A=LDL^{\mathrm{T}}=TL^{\mathrm{T}}$。

$$
\begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{pmatrix}
\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}
\begin{pmatrix} 1 & l_{21} & \cdots & l_{n1} \\ & 1 & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}
$$

$$
= \begin{pmatrix} d_1 & & & \\ t_{21} & d_2 & & \\ \vdots & \vdots & \ddots & \\ t_{n1} & t_{n2} & \cdots & d_n \end{pmatrix}
\begin{pmatrix} 1 & l_{21} & \cdots & l_{n1} \\ & 1 & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}. \tag{56}
$$

Notes:

引入辅助变量$t_{ij}=l_{ij}\cdot d_j$，故LDL^T分解计算公式为：

$$t_{ij}=a_{ij}-\sum_{k=1}^{j-1}t_{ik}l_{jk},\qquad l_{ij}=\frac{t_{ij}}{d_j},\qquad d_i=a_{ii}-\sum_{k=1}^{i-1}t_{ik}l_{ik}, \tag{57}$$

其中，$i=1,2,\cdots,n$，$j=1,2,\cdots,i-1$.

计算顺序为：$d_1,l_{21},d_2,l_{31},l_{32},d_3,l_{41},l_{42},l_{43},\cdots$ 改进算法的计算量与平方根法相当，且不需要开方。矩阵的行列式：$|A|=|L|\,|D|\,\left|L^{\mathrm{T}}\right|=|D|=d_1d_2\cdots d_n$. 求解方程$Ax=b$变为求解：

$$\begin{cases} Ly=b \\ DL^{\mathrm{T}}x=y \end{cases} \Longrightarrow \begin{cases} y_1=b_1 \\ y_i=b_i-\sum\limits_{k=1}^{i-1}l_{ik}y_k \end{cases} \quad (i=1,2,...,n).$$

$$\begin{cases} x_n=\frac{y_n}{d_n} \\ x_i=\frac{y_i}{d_i}-\sum\limits_{k=i+1}^{n}l_{ki}x_k \end{cases} \quad (i=n-1,...,2,1)$$

Notes:

改进平方根法：例子

利用改进的平方根法求解：

$$\begin{bmatrix} 3 & 3 & 5 \\ 3 & 5 & 9 \\ 5 & 9 & 17 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ -4 \end{bmatrix} \tag{58}$$

【解：】$\mathbf{A}$是正定对称的，$\mathbf{L} = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ \frac{5}{3} & 2 & 1 \end{bmatrix}$，$\mathbf{D} = \begin{bmatrix} 3 & & \\ & 2 & \\ & & \frac{2}{3} \end{bmatrix}$，因此：

$$\begin{cases} y_1 = b_1 = 0 \\ y_2 = b_2 - l_{21} y_1 = -2 \\ y_3 = 0 \end{cases} \qquad \begin{cases} x_3 = \frac{y_3}{d_3} = 0 \\ x_2 = \frac{y_2}{d_2} - l_{32} x_3 = -1 \\ x_1 = \frac{y_1}{d_1} - l_{21} x_2 - l_{31} x_3 = 1 \end{cases}$$

Notes:

本讲主要内容

- 引言
- 三角形方程组的解法
- 高斯消元法
- 矩阵的直接LU分解
- 三对角方程组的追赶法
- 列主元消元法
- 高斯－约当消元法
- 列主元LU分解
- 对称正定矩阵的Cholesky分解
- **一般对称正定带状系数矩阵方程组的算法**
- 向量和矩阵的范数、方程组的性态和条件数

Notes:

带状矩阵

定义：对于n阶方阵$A=(a_{ij})_{n\times n}$，若当$|i-j|>m$时，$a_{ij}=0$，且至少有一个k值，使得$a_{k,k-m}\neq 0$或$a_{k,k+m}\neq 0$成立，则**称矩阵A为带状矩阵**，$2m+1$为其带宽，m 为半宽度。带状矩阵的威尔金森图如下：

$|\iff|$ 半带宽m

$$\begin{bmatrix}
x & x & & x & & & & & \\
x & x & x & & x & & & & \\
& x & x & & & x & & & \\
x & & & x & x & & x & & \\
& x & & x & x & & & x & \\
& & x & & & x & x & & x \\
& & & x & & x & x & x & \\
& & & & x & & x & x & x \\
& & & & & & & x & x
\end{bmatrix} \tag{59}$$

Notes:

一般对称正定带状系数矩阵方程组的算法

设$\mathbf{A}$是n阶**对称正定带状矩阵**，带宽为$2m+1$，则求方程组$\mathbf{Ax}=\mathbf{b}$的解时，$\mathbf{A}$可稳定的直接分解为$\mathbf{A}=\mathbf{LDL^T}$，其中$\mathbf{L}$为下半带宽为$m$的带状阵：

$$L=\begin{bmatrix} l_{11} & & & & & \\ l_{21} & l_{22} & & & & \\ \vdots & \ddots & \ddots & & & \\ l_{m+1,1} & \cdots & l_{m+1,m} & l_{m+1,m+1} & & \\ & \ddots & & \ddots & \ddots & \\ & & l_{n,n-m} & \cdots & l_{n,n-1} & l_{nn} \end{bmatrix} \tag{60}$$

即当$i<j$或$i-j>m$时，$l_{ij}=0$。$\mathbf{D}$为对角阵，对角元素为：$d_{ii}=\frac{1}{l_{ii}}$，$i=1,2,...,n$。

Notes:

利用$\mathbf{A}=\mathbf{L}\mathbf{D}\mathbf{L}^{\mathbf{T}}$，可以直接利用矩阵乘法，导出计算$\mathbf{L}$元素的公式：

$$l_{ij}=a_{ij}-\sum_{k=r}^{j-1} l_{ik}l_{jk}/l_{kk},\ i=1,2,...,n. \tag{61}$$

其中

$$r=\begin{cases}1, & i\le m+1,\\ i-m, & i>m+1.\end{cases} \tag{62}$$

同时，我们对方程组的右端项$\mathbf{b}$也作分解：$\mathbf{b}=\mathbf{L}\mathbf{D}\tilde{\mathbf{b}}$，直接利用矩阵和向量乘法，可得：

$$\tilde{b}_i=b_i-\sum_{j=r}^{i-1} l_{ij}\tilde{b}_j/l_{jj},\ i=1,2,...,n. \tag{63}$$

由$\mathbf{A}=\mathbf{L}\mathbf{D}\mathbf{L}^{\mathbf{T}}$及$\mathbf{b}=\mathbf{L}\mathbf{D}\tilde{\mathbf{b}}$可得等价方程组：$\mathbf{L}^{\mathbf{T}}\mathbf{x}=\tilde{\mathbf{b}}$。

Notes:

故原方程组的解，可利用上三角方程组的回代算法计算，公式为：

$$x_i = \left(\tilde{b}_i - \sum_{j=i+1}^{t} l_{ji}x_j\right)/l_{ii},\ i = n, n-1, ..., 1, \tag{64}$$

其中

$$t = \begin{cases} n, & i > n-m-1, \\ i+m, & i \le n-m-1. \end{cases} \tag{65}$$

由于$\mathbf{A}$对称，只需要存放其下半带区(包括对角线元素)于一个矩形数组$\mathbf{C}(1:n, 1:m+1)$。以六阶对称五对角矩阵为例，形式如下：

Notes:

储存

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & a_{11} \\ 0 & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} \\ a_{42} & a_{43} & a_{44} \\ a_{53} & a_{54} & a_{55} \\ a_{64} & a_{65} & a_{66} \end{bmatrix}. \quad (66)$$

其中，**C**的每列是**A**的每条对角线，且左上角添零。a_{ij} 在C的位置如下：

$$c_{i,j-i+m+1} = a_{ij}, \quad \begin{cases} 当 i \leq m+1 时, & j = 1, 2, ..., i; \\ 当 i > m+1 时, & j = i-m, ..., i. \end{cases} \quad (67)$$

请同学们自行写出以上算法的伪代码，并进行程序实现。

Notes:

本讲主要内容

- 引言
- 三角形方程组的解法
- 高斯消元法
- 矩阵的直接LU分解
- 三对角方程组的追赶法
- 列主元消元法
- 高斯－约当消元法
- 列主元LU分解
- 对称正定矩阵的Cholesky分解
- 一般对称正定带状系数矩阵方程组的算法
- **向量和矩阵的范数、方程组的性态和条件数**

Notes:

向量和矩阵的范数

三维空间中向量的长度也称为向量的模或范数，将其进行推广，我们对一般的n维向量也可以定义范数的概念。向量和矩阵的范数的概念，对于分析线性方程组求解问题的误差非常重要。因此我们简要介绍一下它们。

- **定义：** 对两个n维实向量$\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$, $\mathbf{y} = (y_1, y_2, \cdots, y_n)^T$，称$(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{x}^{\mathbf{T}}\mathbf{y}$为向量$\mathbf{x}$与$\mathbf{y}$ 的**内积**；称非负实数$||\mathbf{x}||_2 = (\mathbf{x}, \mathbf{x})^{1/2} = \sqrt{\sum_{i=1}^{n} x_i^2}$为向量$\mathbf{x}$的欧氏范数。
- **内积和欧氏范数的性质：**
  - 正定性：$(\mathbf{x}, \mathbf{x}) \geq 0$，等号仅当$\mathbf{x} = 0$时成立
  - 对称性：$(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$
  - 双线性：$(a\mathbf{x}, \mathbf{y}) = (\mathbf{x}, a\mathbf{y}) = a(\mathbf{x}, \mathbf{y})$；$(\mathbf{x} + \mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z})$
  - 满足Cauchy-Schwartz不等式：$|(\mathbf{x}, \mathbf{y})| \leq ||\mathbf{x}||_2||\mathbf{y}||_2$，等号仅当两者线性相关时成立。
  - 三角不等式：$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}||_2 + ||\mathbf{y}||_2$

Notes:

向量范数的定义

设对任意向量$\mathbf{x} \in R^n$，按一定的规则有一个实数与之对应，记为$||\mathbf{x}||$，若$||\mathbf{x}||$满足以下三个条件：

- 正定性：$||\mathbf{x}|| \geq 0$，等号仅当$\mathbf{x} = 0$时成立；
- 齐次性：$||a\mathbf{x}|| = |a| \cdot ||\mathbf{x}||$，$a$为任意实数；
- 三角不等式：对任意的$\mathbf{x}, \mathbf{y} \in R^n$，总有$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$，

则**称$||\mathbf{x}||$为向量$\mathbf{x}$的范数**。由第三个条件可以推论出，对任意的$\mathbf{x}, \mathbf{y} \in R^n$，总有$|\,||\mathbf{x}|| - ||\mathbf{y}||\,| \leq ||\mathbf{x} - \mathbf{y}||$。常见的**三种向量的范数**：

- 1-范数：$\|x\|_1 = |x_1| + \cdots + |x_n|$；
- 2-范数：$\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$【欧式范数】；
- ∞-范数：$\|x\|_\infty = \max\{|x_1|, \cdots, |x_n|\}$。

三者均是p-范数的特例：$||\mathbf{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$。

例子：$\mathbf{x} = (1, -2, 3)^T$，则：$\|x\|_1 = 6$，$\|x\|_\infty = 3$，$\|x\|_2 = \sqrt{14}$.

Notes:

向量序列的收敛

- **定义：** 设$\{\mathbf{x}^{(k)}\}$为R^n中的一向量序列，$\mathbf{x}^* \in R^n$，记$\{\mathbf{x}^{(k)}\} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^T$，$\mathbf{x}^* = (x_1^*, x_2^*, \cdots, x_n^*)^T$。如果$\lim\limits_{k\to\infty} x_i^{(k)} = x_i^*$ $(i = 1, 2, \ldots, n)$，则称$\{\mathbf{x}^{(k)}\}$收敛于$\mathbf{x}^*$，记为$\lim\limits_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^*$。
- 如果R^n中两个范数$||\bullet||$和$||\bullet||'$，存在实数$m, M > 0$，使得对任意n维向量x，都有$m||\mathbf{x}|| \le ||\mathbf{x}||' \le M||\mathbf{x}||$，则称这两个范数是等价的。对两个等价的范数而言，同一向量序列有相同的极限。
- 【定理】$\lim\limits_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^*$的充要条件是$\lim\limits_{k\to\infty} ||\mathbf{x}^{(k)} - \mathbf{x}^*|| = 0$，其中$||\bullet||$是向量的任意一种范数。若不做特别指明，本讲义中，$||\bullet||$是指任意一种向量范数。

Notes:

矩阵的范数

- **定义：** 对于任意n阶方阵$\mathbf{A}$，按一定的规则有一实数与之对应，记为$||\mathbf{A}||$。若$||\mathbf{A}||$满足：
  - 正定性：$||\mathbf{A}|| \geq 0$，等号仅当$\mathbf{A} = 0$时成立
  - 齐次性：$||\alpha\mathbf{A}|| = |\alpha| \cdot ||\mathbf{A}||$，$\alpha$为任意实数
  - 三角不等式：$||\mathbf{A} + \mathbf{B}|| \leq ||\mathbf{A}|| + ||\mathbf{B}||$，$\mathbf{B}$也为任意$n$阶方阵
  - 相容性条件：$||\mathbf{AB}|| \leq ||\mathbf{A}|| \cdot ||\mathbf{B}||$

  则$||\mathbf{A}||$**称为矩阵的范数**。
- 由于在很多与误差估计有关的问题中，矩阵和向量会同时加以讨论，因此需要引入一种矩阵范数，**与向量范数相容**，即

$$||\mathbf{Ax}|| \leq ||\mathbf{A}|| \cdot ||\mathbf{x}||, \tag{68}$$

对任意的向量$\mathbf{x} \in \mathcal{R}^n$和$n$阶方阵$\mathbf{A}$都成立。

Notes:

矩阵的范数

与前面的三种向量范数相容的三种矩阵范数为：

- **列范数：** $\|\mathbf{A}\|_1 = \max\limits_{\|\mathbf{x}\|_1=1} \|\mathbf{A}\mathbf{x}\|_1 = \max\limits_{1\le j\le n} \sum\limits_{i=1}^{n} |a_{ij}|$，为矩阵的列向量的1-范数的最大值。
- **谱范数：** $\|\mathbf{A}\|_2 = \max\limits_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2 = \sqrt{\lambda_1}$，其中$\lambda_1$是$\mathbf{A}^{\mathbf{T}}\mathbf{A}$ 的最大特征值，这里$\mathbf{A}$为方阵。
- **行范数：** $\|\mathbf{A}\|_\infty = \max\limits_{\|\mathbf{x}\|_\infty=1} \|\mathbf{A}\mathbf{x}\|_\infty = \max\limits_{1\le i\le n} \sum\limits_{j=1}^{n} |a_{ij}|$，为矩阵的行向量1-范数的最大值。

Notes:

求解线性方程组时的误差分析

- 我们说过，一个实际问题转化为数学问题时，初始数据往往会有观测误差和舍入误差，亦即扰动，从而使得最终计算结果产生误差。
- 向量的误差可以用向量的范数表示：设$\mathbf{x}^*$是$\mathbf{x}$的近似向量，$||\mathbf{x}-\mathbf{x}^*||$、$||\mathbf{x}-\mathbf{x}^*||/||\mathbf{x}^*||$分别称为$\mathbf{x}^*$的关于范数$||\bullet||$的绝对误差与相对误差。
- 矩阵的误差可以用矩阵算子范数表示：设$\mathbf{A}^*$是$\mathbf{A}$的近似矩阵，$||\mathbf{A}-\mathbf{A}^*||$、$||\mathbf{A}-\mathbf{A}^*||/||\mathbf{A}^*||$分别称为$\mathbf{A}^*$的关于范数$||\bullet||$的绝对误差与相对误差。
- 由于范数等价，用何种向量范数都是合理的，关键在于实际过程中容易计算。对理论分析，谱范数是非常有效的。但是在计算上，行范数和列范数更加方便。

Notes:

方程组的性态与条件数

比较下面两个方程：

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 + 1.00001x_2 = 2 \end{cases} \Rightarrow \begin{cases} x_1 = 2 \\ x_2 = 0 \end{cases} \tag{69}$$

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 + 1.00001x_2 = 2.00001 \end{cases} \Rightarrow \begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases} \tag{70}$$

显然，只是右端项有很小的差别，最大相对误差仅为$\frac{1}{2} \times 10^{-5}$，但是它们的解截然不同，解的分量的相对误差至少为$\frac{1}{2}$。

对于以上系数矩阵的方程组，输入数据的误差对解的结果影响巨大。

Notes:

线性方程组的性态

- **定义：** 如果矩阵$\mathbf{A}$或者右端项$\mathbf{b}$的微小变化，可以引起方程组$\mathbf{Ax}=\mathbf{b}$解的巨大变化，则称此方程组为“**病态**”**方程组**，矩阵$\mathbf{A}$相对于该方程组而言称为**病态矩阵**；否则称方程组为“**良态**”**方程组**，$\mathbf{A}$称为**良态矩阵**。矩阵的病态性质是矩阵本身的特性。
- 为了定量刻画方程组$\mathbf{Ax}=\mathbf{b}$的病态程度，可分别对方程组的系数矩阵或右端项有扰动时的两种情形进行讨论。

Notes:

右端项b的扰动对解的影响

设$\mathbf{b}$有扰动$\delta\mathbf{b}$，相应的解$\mathbf{x}$的扰动记为$\delta\mathbf{x}$，即:

$$\mathbf{A}(\mathbf{x}+\delta\mathbf{x}) = \mathbf{b}+\delta\mathbf{b}. \tag{71}$$

由$\mathbf{A}\mathbf{x}=\mathbf{b} \quad \Rightarrow \quad \mathbf{A}\delta\mathbf{x}=\delta\mathbf{b} \quad \Rightarrow \quad \delta\mathbf{x}=\mathbf{A}^{-1}\delta\mathbf{b}.$

两边取范数得到：$\|\delta\mathbf{x}\| = \left\|\mathbf{A}^{-1}\delta\mathbf{b}\right\| \le \left\|\mathbf{A}^{-1}\right\| \|\delta\mathbf{b}\|$. 又因为

$$\|\mathbf{x}\| \ge \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{A}\|} = \frac{\|\mathbf{b}\|}{\|\mathbf{A}\|}, \qquad (\text{因为}\,\|\mathbf{A}\mathbf{x}\| \le \|\mathbf{A}\|\,\|\mathbf{x}\|) \tag{72}$$

故

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \frac{\left\|\mathbf{A}^{-1}\right\| \|\delta\mathbf{b}\|}{\frac{\|\mathbf{b}\|}{\|\mathbf{A}\|}} = \frac{\|\mathbf{A}\| \left\|\mathbf{A}^{-1}\right\| \|\delta\mathbf{b}\|}{\|\mathbf{b}\|}, \tag{73}$$

上式表明，当右端项有扰动时，解的相对误差不超过右端项的相对误差的$\|\mathbf{A}\| \left\|\mathbf{A}^{-1}\right\|$倍。

Notes:

系数矩阵A的扰动对解的影响

如果右端项没有扰动，系数矩阵$\mathbf{A}$有扰动$\delta\mathbf{A}$，相应的解的扰动仍然记为$\delta\mathbf{x}$，则

$$
\begin{aligned}
&(\mathbf{A}+\delta\mathbf{A})(\mathbf{x}+\delta\mathbf{x})=\mathbf{b} \quad \Rightarrow \quad \mathbf{A}\delta\mathbf{x}+\delta\mathbf{A}(\mathbf{x}+\delta\mathbf{x})=0 \\
&\Longrightarrow \quad \|\delta\mathbf{x}\|=\left\|\mathbf{A}^{-1}\delta\mathbf{A}(\mathbf{x}+\delta\mathbf{x})\right\| \le \left\|\mathbf{A}^{-1}\right\| \|\delta\mathbf{A}\| \left(\|\mathbf{x}\|+\|\delta\mathbf{x}\|\right)
\end{aligned} \tag{74}
$$

如果$\delta\mathbf{A}$充分小，使得$\left\|\mathbf{A}^{-1}\right\| \|\delta\mathbf{A}\| < \mathbf{1}$，则由上式得

$$
\begin{aligned}
&\|\delta\mathbf{x}\|\left(\mathbf{1}-\left\|\mathbf{A}^{-1}\right\| \|\delta\mathbf{A}\|\right) \le \left\|\mathbf{A}^{-1}\right\| \|\delta\mathbf{A}\| \|\mathbf{x}\| \\
&\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \frac{\left\|\mathbf{A}^{-1}\right\| \|\delta\mathbf{A}\|}{\mathbf{1}-\left\|\mathbf{A}^{-1}\right\| \|\delta\mathbf{A}\|} = \frac{\|\mathbf{A}\| \left\|\mathbf{A}^{-1}\right\| \frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}}{\mathbf{1}-\|\mathbf{A}\| \left\|\mathbf{A}^{-1}\right\| \frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}}
\end{aligned} \tag{75}
$$

此式表明，当系数矩阵有扰动时，解的扰动仍然与$\|\mathbf{A}\| \left\|\mathbf{A}^{-1}\right\|$有关。一般的，$\|\mathbf{A}\| \left\|\mathbf{A}^{-1}\right\|$越大，所导致的解的扰动也越大。

Notes:

条件数的定义

- 综上分析可知，$\|\mathbf{A}\|\,\|\mathbf{A}^{-1}\|$**实际上刻画了解对原始数据变化的敏感程度，即刻画了方程组的"病态"程度。**
- **定义：** 设$\mathbf{A}$为非奇异矩阵，称数$\text{cond}(\mathbf{A})_\nu = \|\mathbf{A}\|_\nu \, \|\mathbf{A}^{-1}\|_\nu$ $(\nu = 1, 2, \infty)$为**矩阵的条件数**。常用的条件数有：
  - $\text{cond}(\mathbf{A})_\infty = \|\mathbf{A}^{-1}\|_\infty \, \|\mathbf{A}\|_\infty$;
  - $\mathbf{A}$的谱条件数：$\text{cond}(\mathbf{A})_2 = \|\mathbf{A}^{-1}\|_2 \, \|\mathbf{A}\|_2 = \sqrt{\frac{\lambda_{\max}(\mathbf{A}^T\mathbf{A})}{\lambda_{\min}(\mathbf{A}^T\mathbf{A})}}$.
- 当$\mathbf{A}$为对称矩阵时，

$$\text{cond}(\mathbf{A})_2 = \frac{|\lambda_1|}{|\lambda_n|}, \tag{76}$$

其中，λ_1和λ_n为$\mathbf{A}$的绝对值最大和绝对值最小的特征值。

Notes:

条件数的性质

- 对任意非奇异矩阵$\mathbf{A}$，都有

$$\text{cond}(\mathbf{A})_\nu \geq 1. \tag{77}$$

这是因为：

$$\text{cond}(\mathbf{A})_\nu = \left\|\mathbf{A}^{-1}\right\|_\nu \|\mathbf{A}\|_\nu \geq \left\|\mathbf{A}^{-1}\mathbf{A}\right\|_\nu = \|\mathbf{I}\| = 1. \tag{78}$$

- 设$\mathbf{A}$为非奇异矩阵而且c为不为零的常数，则：

$$\text{cond}(c\mathbf{A})_\nu = \text{cond}(\mathbf{A})_\nu;$$

- 如果$\mathbf{A}$为正交矩阵，则$\text{cond}(\mathbf{A})_2 = 1$；
- 如果$\mathbf{A}$为非奇异矩阵，$\mathbf{R}$为正交矩阵，对谱条件数有：

$$\text{cond}(\mathbf{RA})_2 = \text{cond}(\mathbf{AR})_2 = \text{cond}(\mathbf{A})_2. \tag{79}$$

Notes:

试计算Hilbert矩阵

$$\mathbf{H}_n = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{n} & \frac{1}{1+n} & \cdots & \frac{1}{2n-1} \end{bmatrix} \tag{80}$$

$n = 3$时，H_3的条件数。

$$\mathbf{H}_3 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}, \quad \mathbf{H}_3{}^{-1} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix} \tag{81}$$

Notes:

条件数的例子

通过计算可得：

$$\mathrm{cond}(\mathbf{H}_3)_\infty = \|\mathbf{H}_3\|_\infty \|\mathbf{H}_3^{-1}\|_\infty = \frac{11}{6} \times 408 = 748. \tag{82}$$

同理可计算得到

$$\mathrm{cond}(\mathbf{H}_6)_\infty = 2.9 \times 10^6. \tag{83}$$

事实上，希尔伯特矩阵是一种著名的病态矩阵，阶数n越大，其条件数越大、病态性越严重。

Notes:

考虑一个例子：

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{11}{6} \\ \frac{13}{12} \\ \frac{47}{60} \end{bmatrix} \Rightarrow \begin{cases} x_1 = 1 \\ x_2 = 1 \\ x_3 = 1 \end{cases} \tag{84}$$

假设$\mathbf{H}_3$和$\mathbf{b}$有微小的误差，$(\mathbf{H}_3 + \delta\mathbf{H}_3)(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$。这里，我们取3位有效数字，则

$$\begin{bmatrix} 1.00 & 0.500 & 0.333 \\ 0.500 & 0.333 & 0.250 \\ 0.333 & 0.250 & 0.200 \end{bmatrix} \begin{bmatrix} x_1 + \delta x_1 \\ x_2 + \delta x_2 \\ x_3 + \delta x_3 \end{bmatrix} = \begin{bmatrix} 1.83 \\ 1.08 \\ 0.783 \end{bmatrix} \tag{85}$$

其解为$x + \delta x = (1.0895, 0.4880, 1.491)^T$，故$\delta x = (0.0895, -0.5120, 0.4910)^T$.

Notes:

条件数的性质

$$
\begin{aligned}
\frac{\|\delta \mathbf{H}_3\|_\infty}{\|\mathbf{H}_3\|_\infty} &\approx 0.18 \times 10^{-3} < 0.02\% \\
\frac{\|\delta \mathbf{b}\|_\infty}{\|\mathbf{b}\|_\infty} &\approx 0.182\% \\
\frac{\|\delta \mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} &\approx \frac{0.5120}{1} = 51.2\%
\end{aligned} \tag{86}
$$

这表明，矩阵和右端项相对误差不超过0.2%，可是引起解的相对误差超过了50%！

Notes:

“病态”方程的经验判断

计算条件数需要求矩阵的逆矩阵，因而代价比较大。根据数值经验，在下列情况下，方程组常常是病态的：

- 在使用主元消去法时出现小主元
- 如果A的最大特征值和最小特征值绝对值之比是大的，则$\mathbf{A}$是病态的
- 系数矩阵中有行或列近似线性相关，或系数行列式的值近似为零。但这不是绝对的，例如当$\mathbf{A}=\varepsilon\mathbf{I}$，$\varepsilon$是很小的数，有$\det(\mathbf{A})=\varepsilon^n\approx 0$，但是$\mathrm{cond}(\mathbf{A})=\mathrm{cond}(\mathbf{I})=1$，方程组是良态的。
- 系数矩阵$\mathbf{A}$元素间数量级相差很大，并且没有一定的规则，则$\mathbf{A}$可能是病态的。

Notes: