第 3 章 线性方程组的直接解法

线性方程组:

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

若 $m > n$，超定方程组，一般没有解，但可求出最小二乘解。方程个数多

若 $m < n$，一般有无穷多个解。方程个数少

$m = n$ 是常见的求解问题。(但不能互相矛盾)

矩阵形式 $Ax = b$。A 系数矩阵，x 解向量，b 右端向量或右端项。

矩阵形式 $Ax = b$。A 系数矩阵，x 解向量，b 右端向量或右端项。

线性方程组的求解分为直接解法和迭代解法。

直接解法：不计误差经过有限步计算能得到准确解的方法。

§1　基本概念与问题的敏感性

本节介绍向量与矩阵的范数，然后讨论线性方程组求解问题的敏感性。

一、有关矩阵的基本知识

$Ax=b$：当 A 为非奇异时，有唯一的解 $x=A^{-1}b$；当 A 为奇异时，无穷多个解或没有解。

$|A|=0$

特殊矩阵：对角矩阵；三对角矩阵；上三角矩阵；下三角矩阵；对称矩阵；

利用特殊矩阵求解.

对称正定矩阵：如果 A 对称，且对任意非零向量 $x \in \mathbb{R}^n$，二次型 $x^T A x > 0$。

判定：顺序主子式皆大于 0，

$\det(A_k) > 0$。其中

$$\underline{A_k} = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}$$

主子阵

二、向量范数和矩阵范数

定义 (向量范数)

记 $\|x\|$ 为向量的某个实值函数，若：(1) $\|x\| \geq 0$，当且仅当 $x = \mathbf{0}$ 时，$\|x\| = 0$;
(2) $\|ax\| = |a|\|x\|$，对于任意的 $a \in \mathbb{R}$;
(3) $\|x + y\| \leq \|x\| + \|y\|$，对于任意的 $x, y \in \mathbb{R}^n$。连续变化，不会跳跃
则称 $\|x\|$ 为 $\mathbb{R}^n$ 上的向量范数。

衡量近似解与准确解差多少

单位圆盘 $\Omega = \{x \mid \|x\| \leq 1\}$。

(3)' Ω 是凸的。

(1)，(2)，(3) 蕴含 (3)'

$$\|(1-t)x + ty\| \leq \|(1-t)x\| + \|ty\|。$$

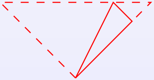
单位圆盘 $\Omega = \{x| \ \|x\| \le 1\}$。

(3)' Ω 是凸的。

(1)，(2)，(3) 蕴含 (3)'

$$\|(1-t)x + ty\| \le \|(1-t)x\| + \|ty\|。$$

(1)，(2)，(3)' 蕴含 (3)



等腰，腰长 $\|x\| + \|y\|$。

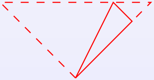
单位圆盘 $\Omega=\{x|\ \|x\|\le 1\}$。

(3)' Ω 是凸的。

(1)，(2)，(3) 蕴含 (3)'

$$\|(1-t)x+ty\|\le\|(1-t)x\|+\|ty\|。$$

(1)，(2)，(3)' 蕴含 (3)



等腰，腰长 $\|x\|+\|y\|$。

单位圆盘是凸的，有界，且关于原点中心对称。

常用的三种范数：

(1) 1-范数， $\|x\|_1 = \sum\limits_{i=1}^{n} |x_i|$;

(2) 2-范数， 一般用这个

$$\|x\|_2 = (\sum_{i=1}^{n} |x_i|^2)^{1/2} = (x^T x)^{1/2};$$

(3) ∞-范数， $\|x\|_\infty = \max\{|x_i|\}$。

$p(p \geq 1)$-范数， $\|x\|_p = (\sum\limits_{i=1}^{n} |x_i|^p)^{1/p}$。

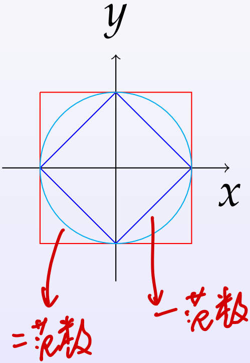
p不一定整数，开方不好办

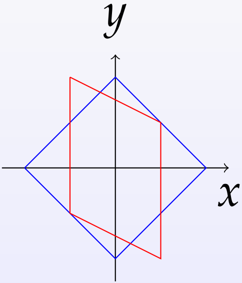
2-范数可利用内积讨论。内积的计算公式为 $\langle x,y\rangle=\sum\limits_{i=1}^{n}x_iy_i=x^Ty$。内积为零时，则称这两个向量正交。

内积范数为 $\|x\|=\sqrt{\langle x,x\rangle}$。

内积范数与 2-范数相同。

例：$\mathbb{R}^2$ 的 1 范数、2 范数和 ∞ 范数下的单位圆：





只是个标号，表示第k个，不需要算

定义

设 $\{x^{(k)}\}$ 为 $\mathbb{R}^n$ 中一向量序列，$x^* \in \mathbb{R}^n$。设 $x^{(k)} = [x_1^{(k)}, \ldots, x_n^{(k)}]^T$，分量收敛，序列收敛

$x^* = [x_1^*, \ldots, x_n^*]^T$。如果 $\lim\limits_{k\to\infty} x_i^{(k)} = x_i^*$，$(i = 1, 2, \ldots, n)$，则称序列 $\{x^{(k)}\}$ 收敛于 x^*，记为 $\lim\limits_{k\to\infty} x^{(k)} = x^*$。

定理

$\mathbb{R}^n$ 上的任一种向量范数 $\|x\|$ 都是关于 x 分量的连续函数。

定理

$\mathbb{R}^n$ 上的任一种向量范数 $\|x\|$ 都是关于 x 分量的连续函数。

定理 (范数的等价性) s,t 表示两种不同范数. 只是个符号

设 $\|x\|_s$ 和 $\|x\|_t$ 为 $\mathbb{R}^n$ 上的任意两种向量范数，则存在常数 $c_1, c_2 > 0$，使得对一切 $x \in \mathbb{R}^n$ 有 $c_1\|x\|_s \leq \|x\|_t \leq c_2\|x\|_s$。

定理

$\mathbb{R}^n$ 上的任一种向量范数 $\|x\|$ 都是关于 x 分量的连续函数。

定理 (范数的等价性)

设 $\|x\|_s$ 和 $\|x\|_t$ 为 $\mathbb{R}^n$ 上的任意两种向量范数，则存在常数 $c_1, c_2 > 0$，使得对一切 $x \in \mathbb{R}^n$ 有 $c_1\|x\|_s \leq \|x\|_t \leq c_2\|x\|_s$。

$\{x^{(k)}\}$ 按范数收敛到 x^*：

$\lim \|x^{(k)} - x^*\| = 0$。

在一个范数下收敛到0，在另一个范数下也收敛到0。

定理

$\lim x^{(k)} = x^*$ 的充要条件是 $\lim \|x^{(k)} - x^*\| = 0$，其中 $\|\ \|$ 可取任一种向量范数。

向量序列收敛等价于向量序列按范数收敛。

定理

$\lim x^{(k)} = x^*$ 的充要条件是 $\lim \|x^{(k)} - x^*\| = 0$，其中 $\|\ \|$ 可取任一种向量范数。

向量序列收敛等价于向量序列按范数收敛。

如：向量序列在某种范数意义下收敛到 0(或 ∞)，则按任意范数下都收敛到 0(或 ∞)。

定义 (矩阵范数)

记 $\|A\|$ 为矩阵 $A \in \mathbb{R}^{n\times n}$ 的某个实值函数，若它满足条件

(1) $\|A\| \geq 0$，且 $\|A\| = 0$ 当且仅当 $A = O$;

(2) $\|aA\| = |a| \cdot \|A\|$，对实数 a;

(3) 对于任意 $A, B \in \mathbb{R}^{n\times n}$，有

$\|A + B\| \leq \|A\| + \|B\|$,

$\|AB\| \leq \|A\| \, \|B\|$,

则称 $\|A\|$ 为 $\mathbb{R}^{n\times n}$ 的矩阵范数。

定义 (矩阵算子范数)

设 $x \in \mathbb{R}^n$，$A \in \mathbb{R}^{n\times n}$，对某种向量范数 $\|x\|_v$，矩阵算子范数为 由向量范数确定

$$\|A\|_v = \max_{x\neq 0} \frac{\|Ax\|_v}{\|x\|_v}。$$

定理

矩阵算子范数是满足相容性条件 $\|Ax\|_v \le \|A\|_v\|x\|_v$ 的矩阵范数。

$\|AB\| \le \|A\|\|B\|$，x视作B即可

定理

对应于向量的 1-范数，2-范数和 ∞-范数，矩阵算子范数分别为

(1) 1-范数：$\|A\|_1 = \max\limits_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$。按列算

(2) 2-范数：$\|A\|_2 = \sqrt{\lambda_{max}(A^T A)}$，其中 $\lambda_{max}()$ 表示取矩阵最大特征值的函数。

(3) ∞-范数：$\|A\|_\infty = \max\limits_{\|x\|_\infty=1} \{\|Ax\|_\infty\}$

$= \max\limits_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$。按行算

例：$A = \begin{pmatrix} 2 & -3 & 4 \\ 4 & 2 & 1 \\ -4 & -1 & -3 \end{pmatrix}$，

$$\begin{array}{c|ccc|} & 10 & 6 & 8 \\ 9 & 2 & -3 & 4 \\ 7 & 4 & 2 & 1 \\ 8 & -4 & -1 & -3 \end{array}, \quad \|A\|_1 = 10,$$

$\|A\|_\infty = 9$。

高维，执行方式更多
(矩阵中横着，竖着，斜着

三、问题的敏感性与矩阵条件数

定义 (矩阵的条件数)

设 A 为非奇异矩阵，称 $cond(A)_v = \|A^{-1}\|_v \|A\|_v$ 为矩阵的条件数，其中下标 v 用于标识某种矩阵的算子范数。

只是标识

系数矩阵的条件数是线性方程组求解问题的条件数的上限。如果系数矩阵的条件数很大，称之为病态矩阵，对应的线性方程组求解问题是敏感问题；如果系数矩阵的条件数很小，称之为良态矩阵，对应的线性方程组求解问题不太敏感。

例 (病态方程)：方程组

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 + 1.0001x_2 = 2 \end{cases}$$

解为 $(x_1, x_2)^T = (2, 0)^T$。而方程组

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 + 1.0001x_2 = 2.0001 \end{cases}$$

解为 $(x_1, x_2)^T = (1, 1)^T$。表示成 $A(x + \Delta x) = (b + \Delta b)$，这里 $\Delta b = (0, 0.0001)^T, \Delta x = (-1, 1)^T$。

∞-范数下，

$$cond = \frac{\|\Delta x\|/\|x\|}{\|\Delta b\|/\|b\|} = \frac{1/2}{0.0001/2} = 10000，而$$

$$cond(A)_\infty = \|A^{-1}\|_\infty \|A\|_\infty \approx 40000。$$

例 (良态方程)：方程组

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

解为 $(x_1, x_2)^T = (1, 1)^T$。而方程组

$$\begin{cases} x_1 - x_2 = 0 \\ x_1 + x_2 = 2.0001 \end{cases}$$

解为 $(x_1, x_2)^T = (1.00005, 1.00005)^T$。

表示成 $A(x + \Delta x) = (b + \Delta b)$，这里 $\Delta b = (0, 0.0001)^T, \Delta x = (0.00005, 0.00005)^T$。

∞-范数下，$cond = \frac{\|\Delta x\|/\|x\|}{\|\Delta b\|/\|b\|} = 1$，而 $cond(A)_\infty = \|A^{-1}\|_\infty \|A\|_\infty = 2$。

一般以10为界，但并不绝对

例：希尔伯特 (Hilbert) 矩阵

$$H_n = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{n} & \frac{1}{1+n} & \cdots & \frac{1}{2n-1} \end{bmatrix}$$

这类矩阵病态性严重

按 ∞-范数计算 H_3 和 H_4 的条件数。

解：(1)

$$H_3 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix},$$

$$H_3^{-1} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$$

$$cond(H_3)_\infty = \|H_3\|_\infty \|H_3^{-1}\|_\infty$$

$$= \frac{11}{6} \times 408 = 748。$$

(2)

$$H_4 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{bmatrix},$$

$$H_4^{-1} = \begin{bmatrix} 16 & -120 & 240 & -140 \\ -120 & 1200 & -2700 & -1680 \\ 240 & -2700 & 6480 & -4200 \\ -140 & -1680 & -4200 & 2800 \end{bmatrix}$$

$$cond(H_4)_\infty = \|H_4\|_\infty \|H_4^{-1}\|_\infty$$

$$= \frac{25}{12} \times 13620 = 28375。$$

希尔伯特矩阵是一种著名的病态矩阵，阶数 n 越大，其病态性越严重。

矩阵条件数的一些重要性质。

定理

在矩阵算子范数意义下，矩阵 A 的条件数满足 $cond(A) = \max\limits_{x\neq\mathbf{0}} \frac{\|Ax\|}{\|x\|} / \min\limits_{x\neq\mathbf{0}} \frac{\|Ax\|}{\|x\|}$。

以 $\mathbb{R}^2$ 的 2-范数为例，$\{Ax \mid \|x\|_2 = 1\}$ 为椭圆。$cond(A)_2$ 为椭圆的长轴与短轴的比。

§2 高斯消去法 进入正题，前面了解

一、基本的高斯消去法

例：用消去法解线性方程组

$$\begin{cases} 10x_1 - 7x_2 = 7 \\ -3x_1 + 2x_2 + 6x_3 = 4 \\ 5x_1 - x_2 + 5x_3 = 6 \end{cases}$$

将系数矩阵变为上三角

解：高斯消去法，分为消去和回代。

$$\begin{bmatrix} 10 & -7 & 0 & 7 \\ -3 & 2 & 6 & 4 \\ 5 & -1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -0.7 & 0 & 0.7 \\ -3 & 2 & 6 & 4 \\ 5 & -1 & 1 & 1 \end{bmatrix}$$

$$
\to \begin{bmatrix} 1 & -0.7 & 0 & 0.7 \\ 0 & -0.1 & 6 & 6.1 \\ 0 & 2.5 & 5 & 2.5 \end{bmatrix}
\to \begin{bmatrix} 1 & -0.7 & 0 & 0.7 \\ 0 & 1 & -60 & -61 \\ 0 & 2.5 & 5 & 2.5 \end{bmatrix}
\to \begin{bmatrix} 1 & -0.7 & 0 & 0.7 \\ 0 & 1 & -60 & -61 \\ 0 & 0 & 155 & 155 \end{bmatrix}
$$

等价的上三角方程组

$$\begin{cases} x_1 & -0.7x_2 & & = 0.7 \\ & x_2 & -60x_3 & = -61 \\ & & 155x_3 & = 155 \end{cases}$$

回代求解：$x_3 = 1$，$x_2 = -1$，$x_1 = 0$。

高斯消去过程对增广矩阵执行两种初等行变换：

(1) 某一行乘以非零常数 c。

(2) 将某一行乘以非零常数 c 后加到另一行。

回代过程依次解出 $x_n, x_{n-1}, \ldots, x_1$。

进行一步释放一步，只存储当前的。

设计算法时，采用了“原地工作”的存储方式。
消去过程要求所有主元 $a_{kk}^{(k)} \neq 0$。
高斯消去法的时间复杂度，浮点数乘除法和加减法的次数约为 $n^3/3$。

§3 矩阵的 LU 分解

一、高斯消去过程的矩阵形式

高斯消去过程是通过一系列初等行变换将系数矩阵变为上三角矩阵。初等行变换等同左乘初等变换矩阵。

三种初等变换矩阵：倍乘变换、倍加变换、交换两行变换

$$\begin{bmatrix} \ddots & & \\ & c & \\ & & \ddots \end{bmatrix},$$

$$
\begin{bmatrix}
\ddots & & & & \\
& 1 & & & \\
& & \ddots & & \\
& c & & 1 & \\
& & & & \ddots
\end{bmatrix},
\quad
\begin{bmatrix}
\ddots & & & & \\
& 0 & & 1 & \\
& & \ddots & & \\
& 1 & & 0 & \\
& & & & \ddots
\end{bmatrix},
$$

第i行乘c倍加到第j行

高斯消去的矩阵形式 $A = LU$，其中 L 下三角矩阵，记录消去过程，U 上三角矩阵，为消去法的结果。

高斯消去的矩阵形式 $A = LU$，其中 L 下三角矩阵，记录消去过程，U 上三角矩阵，为消去法的结果。

n^2 n^2+n

矩阵分解 $A = LU$ 称为三角分解。它有无穷多个方法。 多n个未知. 但n个未知可指定

最具有代表性：杜利特尔 (Doolittle) 分解，L 为单位下三角矩阵 (对角元素全为 1)。克劳特 (Crout) 分解，U 为单位上三角矩阵 (对角元素全为 1)。

→ 高斯消去，其是自然的

本节仅考虑杜利特尔分解:

$$L = \begin{bmatrix} 1 & & & \\ -m_{21} & \ddots & & \\ \vdots & \ddots & 1 & \\ -m_{n1} & \cdots & -m_{n,n-1} & 1 \end{bmatrix},$$

m_{ij} 称为乘数，在高斯消去过程中第 i 行加上第 j 行的 m_{ij} 倍。

例：用高斯消去法求矩阵 $A = \begin{bmatrix} 1 & 2 & 2 \\ 4 & 4 & 2 \\ 4 & 6 & 4 \end{bmatrix}$ 的 LU 分解。

解：$m_{21} = -4$，$m_{31} = -4$，

$$A^{(2)} = \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -6 \\ 0 & -2 & -4 \end{bmatrix}$$

$$m_{32} = -1/2, \quad A^{(3)} = \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -6 \\ 0 & 0 & -1 \end{bmatrix}$$

$A^{(3)}$ 即为矩阵 U，将乘数的相反数填入单位阵的下三角部分得到 L:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 4 & 1/2 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -6 \\ 0 & 0 & -1 \end{bmatrix},$$

U 的对角元素是主元。

二、矩阵的直接 LU 分解算法

"原地存储"思想.

例：将矩阵 $A = \begin{bmatrix} 1 & 2 & 2 \\ 4 & 4 & 2 \\ 4 & 6 & 4 \end{bmatrix}$ 分解为

$A = LU$ 的形式，其中 L 为单位下三角矩阵，U 为上三角矩阵。

方程法求解的话

解：

9个未知数　　9个方程

$$\begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ 4 & 4 & 2 \\ 4 & 6 & 4 \end{bmatrix}$$

第一行：$u_{11}=1$，$u_{12}=2$，$u_{13}=2$。

第一行：$u_{11}=1$，$u_{12}=2$，$u_{13}=2$。
第二行：$l_{21}u_{11}=4$，$l_{21}u_{12}+u_{22}=4$，
$l_{21}u_{13}+u_{23}=2$，解出 $l_{21}=4$，
$u_{22}=-4$，$u_{23}=-6$。

按顺序解什么麻烦都没有

第一行：$u_{11}=1$，$u_{12}=2$，$u_{13}=2$。
第二行：$l_{21}u_{11}=4$，$l_{21}u_{12}+u_{22}=4$，
$l_{21}u_{13}+u_{23}=2$，解出 $l_{21}=4$，
$u_{22}=-4$，$u_{23}=-6$。
第三行：$l_{31}u_{11}=4$，$l_{31}u_{12}+l_{32}u_{22}=6$，
$l_{31}u_{13}+l_{32}u_{23}+u_{33}=4$，解出 $l_{31}=4$，
$u_{32}=1/2$，$u_{33}=-1$。

A 的 LU 分解为

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 4 & 1/2 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 2 & 2 \\ 0 & -4 & -6 \\ 0 & 0 & -1 \end{bmatrix}$$

例：A 是 3 阶矩阵，设计 $A = LU$ 算法。

按 A 的行列出如下方程：

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

$l_{21}u_{11} = a_{21}$，$l_{21}u_{12} + u_{22} = a_{22}$，计算到此，就没有存储 a_{11} 了.

$l_{21}u_{13} + u_{23} = a_{23}$；

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

算法:

$u_{11}=a_{11}$，$u_{12}=a_{12}$，$u_{13}=a_{13}$；
$l_{21}=a_{21}/u_{11}$，$u_{22}=a_{22}-l_{21}u_{12}$，
$u_{23}=a_{23}-l_{21}u_{13}$；
$l_{31}=a_{31}/u_{11}$，$l_{32}=(a_{32}-l_{31}u_{12})/u_{22}$，
$u_{33}=a_{33}-l_{31}u_{13}-l_{32}u_{23}$。

对于 n 阶矩阵 A，按 A 的行列出方程组
第一行：$u_{1j} = a_{1j}$，$j = 1, 2, \ldots, n$。
假设 L 和 U 的前 $k-1$ 行元素已知，
第 k 行前 $k-1$ 个元素：
$\sum\limits_{i=1}^{j-1} l_{ki}u_{ij} + l_{kj}u_{jj} = a_{kj}$，解出

$$l_{kj} = (a_{kj} - \sum_{i=1}^{j-1} l_{ki}u_{ij})/u_{jj}\text{。}$$

第 k 行后 $n-k+1$ 个元素：

$$\sum_{i=1}^{k-1} l_{ki}u_{ij} + u_{kj} = a_{kj},$$

解出 $u_{kj} = a_{kj} - \sum\limits_{i=1}^{k-1} l_{ki}u_{ij}$。

上述过程一直重复，直至 $k=n$。

三、LU 分解的用途

假设已知 $A = LU$，则 $Ax = b$，先以前代求解 $Ly = b$，再以回代求解 $Ux = y$。右端项发生改变时，LU 分解优于直接用高斯消去求解。

批量解方程有优势（如A一样，b不一样）

↓

L与U相同不变

§4 选主元技术与算法稳定性

一、为什么要选主元

U的前 $n-1$ 个对角线元素是主元

定理

对于矩阵 $A \in \mathbb{R}^{n\times n}$，实施 LU 分解过程中不会出现零主元的充要条件是，A 的前 $n-1$ 个顺序主子式均不为零。

§4　选主元技术与算法稳定性

一、为什么要选主元

L下三角未知数，U上三角变量

定理

对于矩阵 $A \in \mathbb{R}^{n\times n}$，实施 LU 分解过程中不会出现零主元的充要条件是，A 的前 $n-1$ 个顺序主子式均不为零。

定理的条件是 A 存在唯一 LU 分解的充分条件。

例：奇异矩阵

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

存在唯一的 LU 分解。

例：奇异矩阵

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

存在唯一的 LU 分解。

例：矩阵 $A=\begin{bmatrix}0&2\\3&1\end{bmatrix}$ 不能进行 LU 分解。

例：奇异矩阵

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

存在唯一的 LU 分解。

消不为0，第二行第一个3.

例：矩阵 $A = \begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix}$ 不能进行 LU 分解。 $3-0a=3$

例： $A = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 0 & 1-2a \end{bmatrix},$

其中 a 为任意实数。 A 存在无穷多种 LU 分解。

第二行减去第一行任何倍都是上三角

解决零主元的问题，假设主元 $a_{kk}^{(k)} = 0$，可将第 k 行和它后面的第 $i_k(i_k > k)$ 行交换，只要 $a_{i_k,k}^{(k)} \neq 0$，交换后当前主元就不为零，可继续高斯消去过程。

对于非奇异线性方程组，可采用选主元的办法保证高斯消元过程的顺利进行，同时减少计算误差。

行列式不为0

0主元会出错，小主元也不好，主元起除数作用.

例 (小主元的情况)：解线性方程组

$$\begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix}, \quad \varepsilon = 2^{-54}。$$

出现大数吃小数的情况.

解：准确解 $x_2^* = 1 + \varepsilon \approx 1$，$x_1 = -1$。

双精度浮点计算

$$\begin{bmatrix} \varepsilon & 1 & 1 \\ 0 & 1 - 1/\varepsilon & \varepsilon - 1/\varepsilon \end{bmatrix}$$

$$\to \begin{bmatrix} \varepsilon & 1 & 1 \\ 0 & -1/\varepsilon & -1/\varepsilon \end{bmatrix},$$

大数吃小数.

(可以用换行解决)

解出 $x_2 = 1$，$x_1 = (1 - 1)/\varepsilon = 0$。

也可用其他办法

二、使用部分主元技术的 LU 分解

部分主元消去法：在高斯消去第 k 步选择 $i_k(i_k \geq k)$ 得 $|a_{i_k k}^{(k)}| = \max_{k \leq i \leq n} |a_{ik}^{(k)}|$。若 $i_k \neq k$，交换 A 的第 k 行和第 i_k 行，实现选主元。

同一列选绝对值最大的做主元

有误差干扰下选主元比不选主元好

数值分析总是考虑误差.

二、使用部分主元技术的 LU 分解 (换行)

部分主元消去法：在高斯消去第 k 步选择 $i_k(i_k \geq k)$ 得 $|a_{i_k k}^{(k)}| = \max\limits_{k\leq i\leq n} |a_{ik}^{(k)}|$。若 $i_k \neq k$，交换 A 的第 k 行和第 i_k 行，实现选主元。 原方程次序不好

矩阵形式：$PA = LU$，其中 L 为单位下三角，U 为上三角，P 是排列阵。

存储上：可用长度为 n 的数组表示 P，其第 i 个元素为第 i 行的 1 所在的列。

例：使用部分主元消去法将矩阵

$$A=\begin{bmatrix}1&2&2\\4&4&2\\4&6&4\end{bmatrix}$$ 进行 LU 分解，求相应的矩阵 L，U 和 P。

解：将 A 存储为二维数组 $\mathbb{A}$，$p=[1,2,3]$。交换第一、第二行得到

$$\mathbb{A}=\begin{bmatrix}4&4&2\\1&2&2\\4&6&4\end{bmatrix},\quad p=[2,1,3]。$$

第一个位置和第二个位置上的交换

$$
\mathbb{A} = \begin{bmatrix} 4 & 4 & 2 \\ 0 & 1 & 3/2 \\ 0 & 2 & 2 \end{bmatrix},
$$

$m_{21} = -1/4$，$m_{31} = -1$，以其相反数填入 $\mathbb{A}$，得

$$
\mathbb{A} = \begin{bmatrix} 4 & 4 & 2 \\ 1/4 & 1 & 3/2 \\ 1 & 2 & 2 \end{bmatrix},
$$

存储很珍贵，不浪费，这两个位置直接存储减去多少倍

交换第二、第三行，得到

$$\mathbb{A} = \begin{bmatrix} 4 & 4 & 2 \\ 1 & 2 & 2 \\ 1/4 & 1 & 3/2 \end{bmatrix}, \quad p = [2,3,1]。$$

乘数

$$\mathbb{A} = \begin{bmatrix} 4 & 4 & 2 \\ 1 & 2 & 2 \\ 1/4 & 0 & 1/2 \end{bmatrix},$$

(不处理乘数，乘数算完就不用了)。

$m_{32} = -1/2$，以其相反数填入 $\mathbb{A}$，得

$$\mathbb{A} = \begin{bmatrix} 4 & 4 & 2 \\ 1 & 2 & 2 \\ 1/4 & 1/2 & 1/2 \end{bmatrix},$$

乘数（L与U写一起了）

$$L=\begin{bmatrix}1&0&0\\1&1&0\\1/4&1/2&1\end{bmatrix},\quad U=\begin{bmatrix}4&4&2\\0&2&2\\0&0&1/2\end{bmatrix},$$

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

(单位阵根据刚才交换 2,3,1).

$P(2,3,1)$ 第一行第二列是1

行：方程次序不妨

列：变元次序不妨

三、其它选主元技术

全主元，在未消去的子矩阵所有元素中选一个绝对值最大的。

通过行、列交换到当前对角的位置。

对应形式为 $PAQ = LU$ 的 LU 分解，其中，P 和 Q 分别为对 A 的行和列进行重排的排列矩阵。

MATLAB 采用部分主元技术？

四、算法的稳定性

线性方程组的直接解法，计算中的误差主要是舍入误差。

1960 年左右，威尔金森 (Wilkinson) 提出“向后误差分析法”。 解x会让系数矩阵A发生一个变化

向后误差大体上满足：

$$\frac{\|\Delta A\|_\infty}{\|A\|_\infty} \le \rho \times n \times eps,$$

其中 ρ 被称为增长因子，n 为矩阵的阶数，eps 为机器精度。在实际应用中，也可将 ρ 看成是矩阵 U 最大元素与矩阵 A 最大元素的比值。

不选主元，ρ 可以任意大，算法不稳定。

部分主元，$\rho \leq 2^{n-1}$。例如：$\rho = \frac{4}{1} = 4 = 2^{3-1}$

$$A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix},$$

$$U = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{bmatrix}$$

最例寄的事例

平均增长因子量级：$n^{2/3}$。

一般认为部分主元高斯消去法是稳定的。

LU分解的变形

§5 对称正定矩阵与三对角矩阵的解法

一、对称正定矩阵的楚列斯基分解

对称矩阵 A，可节省将近一半的存储空间。

1. 对称矩阵的 LDL^T 分解 (本质还是LU，利用对称性，拿出了，因此有了D)

若对称矩阵 A 前 $n-1$ 个顺序主子式不为零，则可分解为 $A = LDL^T$，其中 L 为单位下三角矩阵，D 为对角矩阵。

2. 楚列斯基分解

设 A 为实对称正定矩阵，则可分解为 LDL^T 的形式，D 的对角线元素为正。

这个L是单位下三角，与下面的L不同

相当是LU，且$l_{kk}=u_{kk}$

2. 楚列斯基分解

设 A 为实对称正定矩阵，则可分解为 LDL^T 的形式，D 的对角线元素为正。

对角线元不相同

定理

如果 A 为 n 阶实对称正定矩阵，则存在非奇异下三角矩阵 L，使得 $A = LL^T$，其中，L 的对角元素均大于 0，这种分解称为楚列斯基分解，且满足上述条件的分解是唯一的。

将上面D的对角线元素开方，乘在两边L都放

3. 楚列斯基分解算法

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

对称正定：起码的：

对称，且对角线元素>0

解的过程中出差了，说明矩阵非对称正定。

第一列，$l_{11} = \sqrt{a_{11}}$，$l_{i1} = a_{i1}/l_{11} (i > 1)$。

记解方程组规律，实质就是一直接解方程法

假设 L 的前 $j-1$ 列都已求出，即 $l_{ik}(k \leq j-1)$ 都已知，则

$$l_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2},$$

$$l_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}\right) / l_{jj} \quad (i > j)。$$

例：计算对称正定矩阵

$$A=\begin{bmatrix}5 & -1 & -1\\ -1 & 3 & -1\\ -1 & -1 & 5\end{bmatrix}$$ 的楚列斯基分解。

解：按列列出矩阵下三角部分的变化情况。

$$\begin{bmatrix}2.2361 & 0 & 0\\ -0.4472 & 3 & 0\\ -0.4472 & -1 & 5\end{bmatrix},$$

$$\begin{bmatrix} 2.2361 & 0 & 0 \\ -0.4472 & 1.6733 & 0 \\ -0.4472 & -0.7171 & 5 \end{bmatrix},$$

$$\begin{bmatrix} 2.2361 & 0 & 0 \\ -0.4472 & 1.6733 & 0 \\ -0.4472 & -0.7171 & 2.0702 \end{bmatrix}。$$

$$\begin{bmatrix} 2.2361 & 0 & 0 \\ -0.4472 & 1.6733 & 0 \\ -0.4472 & -0.7171 & 5 \end{bmatrix},$$

$$\begin{bmatrix} 2.2361 & 0 & 0 \\ -0.4472 & 1.6733 & 0 \\ -0.4472 & -0.7171 & 2.0702 \end{bmatrix}。$$

4. 算法稳定性与楚列斯基分解的应用:

楚列斯基分解不必考虑选主元，推导可知，分解算法的稳定性。增长因子 ρ 的上限为 1，算法是稳定的。

二、三对角方程组的解法 样条插值会用

A 存储为三个向量，设

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

样条中，
b大于a+c很多，
此时没必要
选主元
即对角占优

$$L = \begin{bmatrix} 1 & & & & \\ m_2 & 1 & & & \\ & \ddots & \ddots & & \\ & & m_{n-1} & 1 & \\ & & & m_n & 1 \end{bmatrix},$$

m与d是系数

$$U = \begin{bmatrix} d_1 & c_1 & & & \\ & d_2 & c_2 & & \\ & & \ddots & \ddots & \\ & & & d_{n-1} & c_{n-1} \\ & & & & d_n \end{bmatrix}。$$

一行一行推就能推出来

计算公式：$d_1 = b_1$，$m_i = a_i / d_{i-1}$，
$d_i = b_i - m_i c_{i-1}$。
算法时间和空间复杂度都是 $O(n)$。
不选主元，稳定性和可靠性存在问题。
采用部分主元技术，需要预留存储空间。
即便如此，选主元的求解方法仍具有很高的效率。

(为下一章准备)　尽管对角占优，也可能出现，有的就叫弱对角占优，此时可分块讨论 (下章会有)

定义

对于 n 阶方阵 $A=(a_{ij})_{n\times n}$，若

比同行非对角元加起来还要大，

(1) $|a_{ii}| \geq \sum\limits_{j=1,j\neq i}^{n} |a_{ij}|\ (i=1,2,\cdots,n)$ 且

不可以都取等号

至少有一个不等式严格成立，则称矩阵 A 按行对角占优。

(2) $|a_{jj}| \geq \sum\limits_{i=1,i\neq j}^{n} |a_{ij}|\ (j=1,2,\cdots,n)$ 且

至少有一个不等式严格成立，则称矩阵 A 按列对角占优。

定义 (续)

(3) $|a_{ii}| \underset{0}{>} \sum\limits_{j=1,j\neq i}^{n} |a_{ij}| \ (i = 1, 2, \cdots, n)$，则称矩阵 A 按行严格对角占优。

(4) $|a_{jj}| \underset{0}{>} \sum\limits_{i=1,i\neq j}^{n} |a_{ij}| \ (j = 1, 2, \cdots, n)$，则称矩阵 A 按列严格对角占优。

$$\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

→行列式0，奇异，退化的矩阵

→上述按行与按列条件都取等号，不算占优.

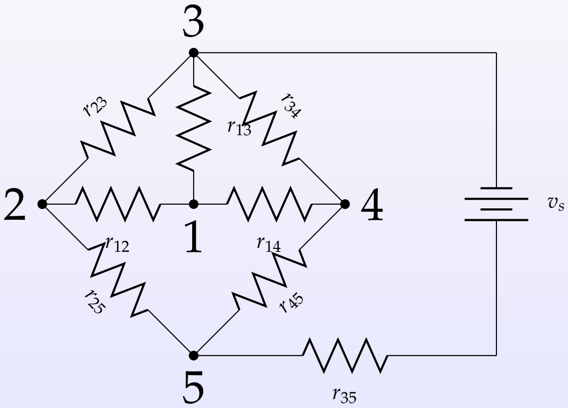
定理

设 n 阶方阵 A 为按列严格对角占优矩阵，则对 A 做 LU 分解不需要选主元，即对其不选主元的 LU 分解是稳定的。

三、应用实例：稳态电路的求解

1. 问题背景

电路图，计算各节点的电压，以及各支路上的电流。



节点分析法得到方程：

$$\begin{bmatrix} g_{12}+g_{13}+g_{14} & -g_{12} & -g_{13} & -g_{14} \\ -g_{12} & g_{12}+g_{23}+g_{25} & -g_{23} & 0 \\ -g_{13} & -g_{23} & g_{12}+g_{13}+g_{14} & -g_{34} \\ -g_{14} & 0 & -g_{34} & g_{14}+g_{34}+g_{45} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ g_{35}v_s \\ 0 \end{bmatrix}$$

其中 v_s 是电压源的电压。

方程的矩阵形式为 $Gv = c$，G 是对称矩阵，并且对角占优，因此不需要选主元。

事实上，可以证明 v^TGv 为电路的总功率，因此矩阵 G 是对称正定的。

一般不通过计算主子式看是否对称正定

节点数目不太多时，采用 Cholesky 分解算法。节点数目很多 (例如大于 10000) 时，利用 G 的稀疏性，采用针对大型稀疏矩阵的直接解法或迭代解法进行求解。

在 MATLAB 中，将多种直接解法集成在一起，提供了一个简单的反斜杠运算符 “\” 实现线性方程组的求解。

解方程 $Ax = b$，执行命令

```
>> x=A\b;
```

即可。

习题：

4. 采用部分主元高斯消去法对矩阵

$$A=\begin{bmatrix}1&2&3\\2&4&5\\3&5&6\end{bmatrix},\quad B=\begin{bmatrix}2&1&1&2\\2&2&2&3\\4&2&4&3\\0&0&6&-1\end{bmatrix}$$

进行 LU 分解，写出矩阵 L、U 和 P。

6. 下列矩阵能否分解为 LU(其中 L 为单位下三角矩阵，U 为上三角矩阵)？若能分解，那么分解是否唯一？

$$A=\begin{bmatrix}1&2&3\\2&4&1\\4&6&7\end{bmatrix},\quad B=\begin{bmatrix}1&1&1\\2&2&1\\3&3&1\end{bmatrix}。$$

上机题

希尔伯特矩阵，高度病态 $\begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & & \cdots & \frac{1}{n+1} \\ \vdots & & & \vdots \\ \frac{1}{n} & & \cdots & \frac{1}{2n-1} \end{pmatrix}$

1. 编写程序生成 Hilbert 矩阵 H_n，以及 n 维向量 $b = H_n x$，其中，x 为所有分量都是 1 的向量。用 Cholesky 分解算法求解方程 $H_n x = b$，得到近似解 $\hat{x}$，计算残差 $r = b - H_n \hat{x}$ 和 $\Delta x = \hat{x} - x$ 的 ∞-范数。

楚列斯基分解

(1) 设 $n = 10$，计算 $\|r\|_\infty$、$\|\Delta x\|_\infty$。

(2) 在右端项上施加 10^{-7} 的扰动然后解方程组，观察残差和误差的变化情况。

(3) 改变 n 的值为 8 和 12，求解相应的方程，观察 $\|r\|_\infty$、$\|\Delta x\|_\infty$ 的变化情况。通过这个实验说明了什么问题?