《数值分析》第五章

- 矩阵基础知识
- Gauss 消去法
- 矩阵三角分解
- 向量与矩阵范数
- 误差分析

线性方程组直接解法

$$\boxed{Ax = b} \qquad \boxed{A \in R^{n\times n},\ \ b \in R^{n}}$$

- 自然科学和工程计算中, 很多问题最终都需要求解一个线性代数方程组

- 目前使用的数值解法:

  (1) **直接法:** 适合低阶方程组或某些特殊大型稀疏方程组

  (2) **迭代法:** 解大型稀疏方程组的主流算法

在本章中，我们总是假定 A 是 n 阶非奇异方阵

5.1 预备知识

- 向量与矩阵：定义，基本运算，行列式
- 特征值与特征向量，特征多项式，特征方程，矩阵相似

$$\boxed{Ax = \lambda x} \quad (\lambda \in C,\ x \in C^n,\ x \neq 0)$$

- 矩阵的谱：$\boxed{\sigma(A) = \{ A \text{ 的所有特征值} \}}$
- 矩阵的谱半径：$\boxed{\rho(A) = \max\limits_{\lambda \in \sigma(A)} \{|\lambda|\}}$
- 矩阵的迹：$\boxed{\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}}$

矩阵基本性质

- A^T 与 A 有相同的特征值，但特征向量不同
- A^{-1} 的特征值与特征向量

$$\boxed{Ax = \lambda x \Leftrightarrow A^{-1}x = \lambda^{-1}x}$$

- 矩阵的迹与特征值 $\boxed{\mathrm{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n}$

- 矩阵行列式与特征值 $\boxed{\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n}$

- 相似矩阵的特征值与特征向量

一些特殊矩阵

- 对角矩阵、三角矩阵、三对角矩阵
- 对称矩阵、Hermite对称矩阵、对称正定矩阵
- 正交矩阵、酉矩阵
- 初等置换阵、置换阵（排列阵）

- 上 Hessenberg 矩阵

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ & a_{32} & a_{33} & & a_{3n} \\ & & \ddots & \ddots & \vdots \\ & & & a_{n,n-1} & a_{nn} \end{bmatrix}$$

$$\boxed{a_{ij} = 0 \quad \text{for} \quad i > j+1}$$

一些重要性质

定理 1 （解的存在唯一性，教材 141 页）

定理 2 （对称正定矩阵的性质，教材 141 页）

定理 3 （对称矩阵正定的充分条件，教材 141 页）

定理 4 （Jordan 标准型，教材 142 页）

5.2 Gauss 消去法

例： 用直接法解线性方程组
$$\begin{cases} x_1 - 2x_2 + 2x_3 = -2 \\ 2x_1 - 3x_2 - 3x_3 = 4 \\ 4x_1 + x_2 + 6x_3 = 3 \end{cases}$$

解：

$$(A, b) = \left[\begin{array}{ccc|c} 1 & -2 & 2 & -2 \\ 2 & -3 & -3 & 4 \\ 4 & 1 & 6 & 3 \end{array}\right] \rightarrow \begin{bmatrix} 1 & -2 & 2 & -2 \\ 0 & 1 & -7 & 8 \\ 0 & 9 & -2 & 11 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & -2 & 2 & -2 \\ 0 & 1 & -7 & 8 \\ 0 & 0 & 61 & -61 \end{bmatrix} \rightarrow \begin{cases} x_3 = -1 \\ x_2 = 8 + 7x_3 = 1 \\ x_1 = -2 + 2x_2 - 2x_3 = 2 \end{cases}$$

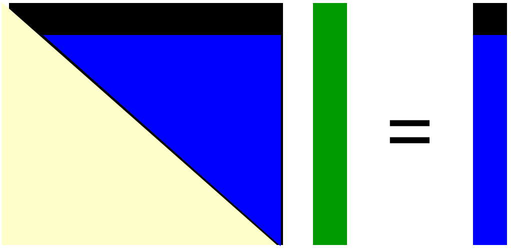
Gauss 消去法

考虑 n 阶线性方程组：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \xrightarrow{\text{矩阵形式}} Ax = b$$

- **高斯消去法的主要思路：**

将系数矩阵 A 化为上三角矩阵，然后回代求解



记 $A^{(1)}=(a_{ij}^{(1)})_{n\times n}=A,\ b^{(1)}=\begin{pmatrix} b_1^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix}=b$，即 $a_{ij}^{(1)}=a_{ij},\ b_i^{(1)}=b_i$

第 1 步：消第 1 列

设 $a_{11}^{(1)}\neq 0$，计算 $m_{i1}=a_{i1}^{(1)}/a_{11}^{(1)}\quad (i=2,\dots,n)$

依次将增广矩阵的 第 i 行 - $m_{i1}\times$ 第 1 行，得

$$A^{(2)}=\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ & \vdots & \ddots & \vdots & \vdots \\ & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right)$$

其中 $\boxed{\begin{aligned} a_{ij}^{(2)} &= a_{ij}^{(1)}-m_{i1}a_{1j}^{(1)} \\ b_i^{(2)} &= b_i^{(1)}-m_{i1}b_1^{(1)} \end{aligned}}$

$(i,j=2,\dots,n)$

第 2 步：消第 2 列

设 $a_{22}^{(2)} \neq 0$，计算 $m_{i2} = a_{i2}^{(2)} \big/ a_{22}^{(2)}$ $(i = 3, \ldots, n)$

依次将 $A^{(2)}$ 的 第 i 行 $- m_{i2} \times$ 第 2 行，得

$$A^{(3)} = \left(\begin{array}{ccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ & & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ & & \vdots & \ddots & \vdots & \vdots \\ & & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} & b_n^{(3)} \end{array} \right)$$

其中 $\boxed{\begin{cases} a_{ij}^{(3)} = a_{ij}^{(2)} - m_{i2} a_{2j}^{(2)} \\ b_i^{(3)} = b_i^{(2)} - m_{i2} b_2^{(2)} \end{cases}}$ $(i, j = 3, \ldots, n)$

依此类推，第 k-1 步 后可得

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

第 k 步：消第 k 列

设 $a_{kk}^{(k)}\neq 0$

先计算：$\boxed{m_{ik}=a_{ik}^{(k)}/a_{kk}^{(k)}}$

$(\, i=k+1, \ldots, n \,)$

再计算：

$$\boxed{\begin{aligned} a_{ij}^{(k+1)} &= a_{ij}^{(k)}-m_{ik}a_{kj}^{(k)} \\ b_{i}^{(k+1)} &= b_{i}^{(k)}-m_{ik}b_{k}^{(k)} \end{aligned}}$$

$(\, i=k+1, \ldots, n \,)$

$\Longrightarrow A^{(k+1)}$

第 n-1 步后

$$\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{(n)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(n)} \end{bmatrix} \Leftrightarrow A^{(n)} x = b^{(n)}$$

回代求解：

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

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

$(i = n\text{-}1, \ldots, 1)$

几点注记

- **主元：** $a_{ii}^{(i)} \quad (i = 1, 2, \ ..., n)$
- **Gauss 消去法能进行到底的条件：主元全不为 0**

定理： $a_{ii}^{(i)} \neq 0$（i=1, 2, ..., n）的充要条件是 A 的顺序主子式不为零，即

$$D_1 = a_{11} \neq 0, \quad D_i = \begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{vmatrix} \neq 0, \quad i = 1, 2, \ldots, n$$

推论： $a_{11}^{(1)} = D_1, \quad a_{ii}^{(i)} = D_i / D_{i-1}, \quad i = 2, \ldots, n$

运算量

- **乘除运算** 的次数

第 k 步：消第 k 列

计算 $m_{ik} = a_{ik}^{(k)} / a_{kk}^{(k)}$ $(i = k+1, \ldots, n)$ — $n-k$ 次

计算 $a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)}$ $(i = k+1, \ldots, n)$ — $(n-k)^2$ 次

$b_i^{(k+1)} = b_i^{(k)} - m_{ik} b_k^{(k)}$ — $n-k$ 次

回代求解：

$$\begin{cases} x_n = b_n^{(n)} / a_{nn}^{(n)} \\ x_i = \left(b_i^{(i)} - \sum_{j=i+1}^{n} a_{ij}^{(i)} x_j\right) \Big/ a_{ii}^{(i)} \end{cases} \quad (i = n-1, \ldots, 2, 1)$$

— $n(n+1)/2$ 次

Gauss 消去法的乘除运算量为：$\boxed{\dfrac{n^3}{3} + n^2 - \dfrac{n}{3}}$

- **加减运算** 的次数 $\boxed{\dfrac{n^3}{3} + \dfrac{n^2}{2} - \dfrac{5n}{6}}$

5.2.2 LU 分解

- 换个角度看 Gauss 消去法

  ⇒ 矩阵的三角分解过程

矩阵分解，即将一个较复杂的矩阵分解成若干结构简单的矩阵的乘积，是矩阵计算中的一个很重要的技术

LU 分解

将 Gauss 消去过程中第 $k-1$ 步消元后的系数矩阵记为：

$$A^{(k)}=\begin{bmatrix} a_{11}^{(1)} & \cdots & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} \\ & \ddots & \vdots & & \vdots \\ & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ & & \vdots & \ddots & \vdots \\ & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} \end{bmatrix} \qquad (k=1,\ldots,n-1)$$

则 $A^{(k)}$ 与 $A^{(k+1)}$ 之间的关系式可以表示为：

$$\boxed{A^{(k+1)}=L_k A^{(k)}}$$

其中：

$$L_k=\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -m_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -m_{n,k} & & & 1 \end{bmatrix}$$

$$m_{ik}=a_{ik}^{(k)}\Big/a_{kk}^{(k)}$$

$$(\,i=k+1,\ldots,n\,)$$

于是有： $A^{(n)} = L_{n-1} \cdots L_2 L_1 A^{(1)}$

$$\Rightarrow \boxed{A = A^{(1)} = (L_{n-1} \cdots L_2 L_1)^{-1} A^{(n)}}$$

容易验证：

$$L_k^{-1} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & m_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & m_{n,k} & & & 1 \end{bmatrix} \quad (k = 1, \ldots, n-1)$$

且 $$L_1^{-1} L_2^{-1} \cdots L_{n-1}^{-1} = \begin{bmatrix} 1 & & & & & \\ m_{2,1} & 1 & & & & \\ m_{3,1} & m_{3,2} & 1 & & & \\ m_{4,1} & m_{4,2} & m_{4,3} & \ddots & & \\ \vdots & \vdots & \vdots & & 1 & \\ m_{n,1} & m_{n,2} & m_{n,3} & \cdots & m_{n,n-1} & 1 \end{bmatrix}$$

记：$L = L_1^{-1} L_2^{-1} \cdots L_n^{-1}$，$U = A^{(n)}$，则

$\boxed{A = LU}$ $\longrightarrow$ LU 分解

其中：L --- 单位下三角矩阵，U --- 非奇异上三角矩阵

注：LU分解中要求 L 是单位下三角，U 是非奇异上三角！

LU 分解存在唯一性

LU 分解存在 ⟺ 高斯消去法不被中断

⇕

所有顺序主子式不为零 ⟺ $a_{kk}^{(k)} \neq 0$

定理： *A* 存在唯一的 LU 分解的充要条件是 *A* 的所有顺序主子式都不为零。

5.2.3 列主元 Gauss 消去法

- **为什么要选主元？**

Gauss 消去法有效的条件是 主元全不为零！

例： 解线性方程组 $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

- **列主元 Gauss 消去法**

在第 k 步消元时，在第 k 列的剩余部分选取主元

① 先选取列主元：$|a_{i_k k}^{(k)}| = \max\limits_{k\le i\le n}\left\{|a_{ik}^{(k)}|\right\} \neq 0$

② if $i_k \neq k$ then 交换第 k 行和第 i_k 行

③ 消元

列主元 Gauss 消去法

算法 (列主元Gauss消去法)

for k=1 **to** n-1

$$|a_{i_k k}^{(k)}| = \max_{k\le i\le n}\left\{|a_{ik}^{(k)}|\right\} \neq 0$$

if $a_{i_k k}^{(k)}$ =0 **then stop**

if $i_k \neq k$ **then swap** k**-th and** i_k**-th row (including** b**)**

for i=k+1 **to** n

$$m_{ik} = a_{ik} \big/ a_{kk}$$

$$a_{ij} = a_{ij} - m_{ik}a_{kj}, \quad j = k+1, k+2, \ldots, n$$

$$b_i = b_i - m_{ik}b_k$$

end

end

$$x_n = b_n \big/ a_n, \quad x_i = \left(b_i - \sum_{j=i+1}^{n} a_{ij}x_j\right) \Big/ a_{ii}, \quad i = n-1, \ldots, 2, 1$$

PLU 分解

- 列主元 Gauss 消去法对应的矩阵分解为 **PLU 分解**

> **定理**：若 A 非奇异，则存在排列矩阵 P，使得
>
> $$PA = LU$$
>
> 其中 L 为单位下三角矩阵，U 为上三角矩阵

- 列主元 Gauss 消去法比普通 Gauss 消去法要多一些比较运算，但比普通高斯消去法稳定
- 列主元 Gauss 消去法是目前直接法的首选算法

全主元Gauss消去法

- 全主元高斯消去法

第 k 步消元时，在剩余的 $n-k$ 阶子矩阵中选取主元

① 先选取全主元：$|a_{i_k j_k}^{(k)}| = \max\limits_{k \le i, j \le n} \left\{ |a_{ij}^{(k)}| \right\} \neq 0$

② if $i_k \neq k$ then 交换第 k 行和第 i_k 行

if $j_k \neq k$ then 交换第 k 列和第 j_k 列

③ 消元

- 列交换改变了 x_i 的顺序，须记录交换次序，解完后再换回来
- 全主元高斯消去法具有更好的稳定性，但很费时，在实际计算中很少使用

5.3 LU 三角分解

- 什么是矩阵的三角分解
  - 将一个矩阵分解成结构简单的三角矩阵的乘积

- Gauss 消去法对应的矩阵三角分解

$\boxed{\text{矩阵的 } LU \text{ 分解}}$ $$A = LU$$

$\boxed{\text{矩阵的 } LDR \text{ 分解}}$ $$A = LDR$$

$\boxed{\text{克洛脱 (Crout) 分解}}$ $$A = \tilde{L}\tilde{U}$$

LU 分解

● **利用矩阵乘法直接计算 LU 分解 —— 待定系数法**

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

$$L \quad \times \quad U \quad = \quad A$$

□比较等式两边的第一行得：$u_{1j} = a_{1j}$ $(j = 1, \ldots, n$ ← U 的第一行

比较等式两边的第一列得：$l_{i1} = a_{i1} / u_{11}$ $(i = 2, \ldots$ ← L 的第一列

□比较等式两边的第二行得：$u_{2j} = a_{2j} - l_{21} u_{1j}$ $(j =$ ← U 的第二行

比较等式两边的第二列得：$l_{i2} = \left(a_{i2} - l_{i1} u_{12} \right) / u_{22}$ $($ ← L 的第二列

计算 LU 分解

第 k 步：此时 U 的前 k-1 行和 L 的前 k-1 列已经求出

比较等式两边的第 k 行得：

$$u_{kj} = a_{kj} - \left(l_{k1}u_{1j} + \cdots + l_{k,k-1}u_{k-1,j}\right) = a_{kj} - \sum_{i=1}^{k-1} l_{ki}u_{ij}$$

$(\, j = k, \ldots, n \,)$

比较等式两边的第 k 列得：

$$l_{ik} = \left(a_{ik} - l_{i1}u_{1k} - \cdots - l_{i,k-1}u_{k-1,k}\right) \Big/ u_{kk} = \left(a_{ik} - \sum_{j=1}^{i-1} l_{ij}u_{jk}\right) \Big/ u_{kk}$$

$(\, i = k+1, \ldots, n \,)$

直到第 n 步，便可求出矩阵 L 和 U 的所有元素。

- **这种计算 LU 分解的方法也称为 Doolittle 分解**

LU 分解

算法：(LU 分解)

for $k = 1$ **to** n

$$a_{kj} \leftarrow u_{kj} = a_{kj} - \sum_{i=1}^{k-1} l_{ki} u_{ij} \ , \quad j = k, \ldots, n$$

$$a_{ik} \leftarrow l_{ik} = \left(a_{ik} - \sum_{j=1}^{k-1} l_{ij} u_{jk} \right) \Big/ u_{kk} \ , \quad i = k+1, \ldots, n$$

end

运算量：$(n^3 - n)/3$

为了节省存储空间，通常用 A 的绝对下三角部分来存放 L (对角线元素无需存储)，用 A 的上三角部分来存放 U

LU 分解算法

算法：(LU 分解求解方程组)

% 先计算 LU 分解

for $k = 1$ **to** n

$$a_{kj} = a_{kj} - \sum_{i=1}^{k-1} a_{ki} a_{ij} \ , \quad j = k, k+1, \ldots, n$$

$$a_{ik} = \frac{1}{a_{kk}} \left(a_{ik} - \sum_{j=1}^{k-1} a_{ij} a_{jk} \right), \quad i = k+1, k+2, \ldots, n$$

end

% 解三角方程组 $Ly = b$ 和 $Ux = y$

$$y_1 = b_1 \ , \qquad y_i = b_i - \sum_{k=1}^{i-1} a_{ik} y_k \ , \quad i = 2, 3, \ldots, n$$

$$x_n = y_n \ , \qquad x_i = \frac{1}{a_{ii}} \left(y_i - \sum_{k=i+1}^{n} a_{ik} x_k \right) , \quad i = n-1, n-2, \ldots, 1$$

PLU 分解

- 列主元 Gauss 消去法 —— PLU 分解 $\boxed{PA = LU}$

for $k = 1$ to n

$$a_{ik} = a_{ik} - \sum_{j=1}^{k-1} a_{ij} a_{jk}, \quad i = k, k+1, \ldots, n$$

$$a_{i_k k} = \max_{k \le i \le n} |a_{ik}|, \quad \mathrm{Ip}(k) \leftrightarrow \mathrm{Ip}(i_k)$$

$$a_{kj} \leftrightarrow a_{i_k j}, \quad j = 1, 2, \ldots, n$$

$$a_{ik} = a_{ik} / a_{kk}, \quad i = k+1, \ldots, n$$

$$a_{kj} = a_{kj} - \sum_{i=1}^{k-1} a_{ki} a_{ij}, \quad j = k+1, \ldots, n$$

end

- 此算法可以用于计算矩阵的行列式和逆

PLU 分解算法

算法：(PLU 分解求解方程组)

for $k = 1$ to n

$$a_{ik} = a_{ik} - \sum_{j=1}^{k-1} a_{ij} a_{jk}\ ,\quad i = k, k+1, \ldots, n$$

$$a_{i_k k} = \max_{k \le i \le n} |a_{ik}|\ ,\quad \mathrm{Ip}(k) \leftrightarrow \mathrm{Ip}(i_k)$$

$$a_{kj} \leftrightarrow a_{i_k j}\ ,\quad j = 1, 2, \ldots, n$$

$$a_{ik} = a_{ik} / a_{kk}\ ,\quad i = k, k+1, \ldots, n$$

$$a_{kj} = a_{kj} - \sum_{i=1}^{k-1} a_{ki} a_{ij}\ ,\quad j = k, k+1, \ldots, n$$

end

% 解三角方程组 $Ly = Pb$ 和 $Ux = y$

$$y_1 = b_{\mathrm{Ip}(1)}\ ,\quad y_i = b_{\mathrm{Ip}(i)} - \sum_{k=1}^{i-1} a_{ik} y_k\ ,\quad i = 2, 3, \ldots, n$$

$$x_n = y_n\ ,\quad x_i = \frac{1}{a_{ii}} \left(y_i - \sum_{k=i+1}^{n} a_{ik} x_k \right),\quad i = n-1, n-2, \ldots, 1$$

5.3.2 Cholesky 分解

- 对称正定矩阵的三角分解 —— Cholesky 分解

定理：（Cholesky分解）若 A 对称正定，则 A 可唯一分解为

$$A = LL^T$$

其中 L 为下三角实矩阵，且对角元素都大于 0

定理：设 A 是对称矩阵，若 A 的所有顺序主子式都不为 0，则 A 可唯一分解为

$$A = LDL^T$$

其中 L 为单位下三角阵，D 为对角矩阵

计算 Cholesky 分解

- **如何计算 Cholesky 分解 —— 待定系数法**

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

$$\Rightarrow \boxed{a_{ij} = \sum_{k=1}^{n} l_{ik} l_{jk} = \sum_{k=1}^{j-1} l_{ik} l_{jk} + l_{jj} l_{ij}} \quad (i \geq j)$$

$$j = 1, 2, \ldots, n, \quad i = j, j+1, \ldots, n$$

Cholesky 分解算法

算法：(Cholesky 分解)

for $j = 1$ to n

$$l_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2 \right)^{\frac{1}{2}} ,$$

$$l_{ij} = \frac{1}{l_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk} \right), \qquad i = j+1, \ldots, n$$

end

平方根法

- 用 Cholesky 分解求解线性方程组 —— 平方根法

$$Ax = b$$

A 对称正定

算法：(解对称正定线性方程组的平方根法)

先计算 A 的 Cholesky 分解

然后解方程：$Ly = b$ 和 $L^T x = y$

$$y_1 = b_1 / l_{11}, \quad y_i = \left(b_i - \sum_{k=1}^{i-1} l_{ik} y_k \right) \Big/ l_{ii} \quad , \quad i = 2, 3, \ldots, n$$

$$x_n = y_n / l_{nn} \quad , \quad x_i = \left(y_i - \sum_{k=i+1}^{n} l_{ki} x_k \right) \Big/ l_{ii}$$

$$i = n\text{-}1, \ldots, 2, 1$$

改进的 Cholesky 分解

- 改进的 Cholesky 分解

$$A = LDL^T = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{bmatrix} \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \begin{bmatrix} 1 & l_{21} & \cdots & l_{n1} \\ & 1 & & l_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

$$\Rightarrow \boxed{a_{ij} = \sum_{k=1}^{n} l_{ik} d_k l_{jk} = \sum_{k=1}^{j-1} l_{ik} d_k l_{jk} + l_{ij} d_j} \quad (i \geq j)$$

$$j = 1, 2, \ldots, n, \quad i = j, j+1, \ldots, n$$

改进的 Cholesky 分解

算法：(改进的 Cholesky 分解)

for $j = 1$ to n

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

$$l_{ij} = \frac{1}{d_j}\left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} d_k l_{jk}\right), \quad i = j+1, \ldots, n$$

end

- 优点：避免开方运算

改进的平方根法

- **改进的平方根法**

$$\boldsymbol{Ax = b}$$

A 对称正定

算法：(解对称正定线性方程组的改进的平方根法)

先计算 改进的 Cholesky 分解

然后解方程组： $Ly = b$ **和** $DL^T x = y$

$$y_1 = b_1 \ , \quad y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k \ , \qquad i = 2, 3, \ldots, n$$

$$x_n = y_n / d_n \ , \quad x_i = y_i / d_i - \sum_{k=i+1}^{n} l_{ki} x_k \ ,$$

$$i = n\text{-}1, \ldots, 2, 1$$

5.3.3 三对角矩阵 LU 分解

- 对角占优的**不可约**三对角矩阵的 Crout 分解

$$\begin{bmatrix} b_1 & c_1 & & \\ a_1 & \ddots & \ddots & \\ & \ddots & \ddots & c_{n-1} \\ & & a_{n-1} & b_n \end{bmatrix} = \begin{bmatrix} \alpha_1 & & & \\ a_1 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & a_{n-1} & \alpha_n \end{bmatrix} \begin{bmatrix} 1 & \beta_1 & & \\ & 1 & \ddots & \\ & & \ddots & \beta_{n-1} \\ & & & 1 \end{bmatrix}$$

- 计算过程

(1) 第一步： $\alpha_1 = b_1, \quad \beta_1 = c_1/\alpha_1 = c_1/b_1$

(2) 第二步： $\alpha_i = b_i - a_{i-1}\beta_{i-1}\,, \quad \beta_i = c_i/\alpha_i$ $\quad i = 2, 3, \ldots, n\text{-}1$

(3) 第三步： $\alpha_n = b_n - a_{n-1}\beta_{n-1}$

- 为了编程方便，这里 a 的下标从 1 开始（教材是从 2 开始）
- 关于对角占优、不可约等性质，会在第六章详细讲述

追赶法

- 三对角线性方程组的求解 —— 追赶法

$$\boldsymbol{Ax} = \boldsymbol{f}$$

A 三对角矩阵（对角占优不可约）

算法：(追赶法)

$$\beta_1 = c_1 / b_1 \ , \quad \beta_i = c_i / \left(b_i - a_{i-1}\beta_{i-1}\right)$$ 追

$i = 2, 3, \ldots, n$-1

$$y_1 = f_1 / b_1 \ , \quad y_i = \left(f_i - a_{i-1}y_{i-1}\right) / \left(b_i - a_{i-1}\beta_{i-1}\right)$$ 追

$i = 2, 3, \ldots, n$

$$x_n = y_n \ , \quad x_i = y_i - \beta_i x_{i+1}$$ 赶

$i = n$-1, …, 2, 1

- 运算量：约 5n+3n

实际计算中$Ax=f$的阶数往往很高,应注意A的存贮技术.

已知数据只用4个一维数组就可存完.即$\{a_i\}$,$\{b_i\}$,$\{c_i\}$,$\{f_i\}$各占一个一维数组, $\{\alpha_i\}$和$\{\beta_i\}$可存放在$\{b_i\}$,$\{c_i\}$的位置,$\{y_i\}$ 和$\{x_i\}$则可放在$\{f_i\}$的位置,整个运算可在4个一维数组中运行.

追赶法的计算量很小,只是$5(n-1)$次乘除法.追赶法的计算也不要选主元素.

5.4向量和矩阵范数

- 向量内积（数量积）
  - 定义与性质、Cauchy-Schwarz不等式
  - 导出范数（欧氏范数）

- 向量范数

定义： 设函数 $f: R^n \to R$，若 f 满足

(1) $f(x) \geq 0$，$\forall\, x \in R^n$，等号当且仅当 $x = 0$ 时成立 （正定性）

(2) $f(\alpha x) = |\alpha| \cdot f(x)$， $\forall\, x \in R^n, \forall\, \alpha \in R$ （齐次性）

(3) $f(x+y) \leq f(x) + f(y)$ （三角不等式）

则称 f 为 R^n 上的（向量）范数，通常记为 $\|\cdot\|$

常见向量范数

- R^n 空间上常见的向量范数

- **1-范数：** $$\|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|$$

- **2-范数：** $$\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

- **p-范数：** $$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$$

- **∞-范数（有时也称最大范数）：** $$\|x\|_\infty = \max_{1 \le i \le n} |x_i|$$

范数性质

- **范数的性质**

(1) 连续性

定理： 设 f 是 R^n 上的任一向量范数，则 f 关于 x 的每个分量连续。

(2) 等价性

定理： 设 $\|\cdot\|_s$ 和 $\|\cdot\|_t$ 是 R^n 上的任意两个范数，则存在常数 c_1 和 c_2，使得对任意的 $x\in R^n$ 有

$$c_1\|x\|_s \le \|x\|_t \le c_2\|x\|_s$$

定理 （范数的等价性）对任何向量$x\in R^n$,有

（1） $$\|X\|_2 \le \|X\|_1 \le \sqrt{n}\|X\|_2$$

（2） $$\|X\|_\infty \le \|X\|_1 \le n\|X\|_\infty$$

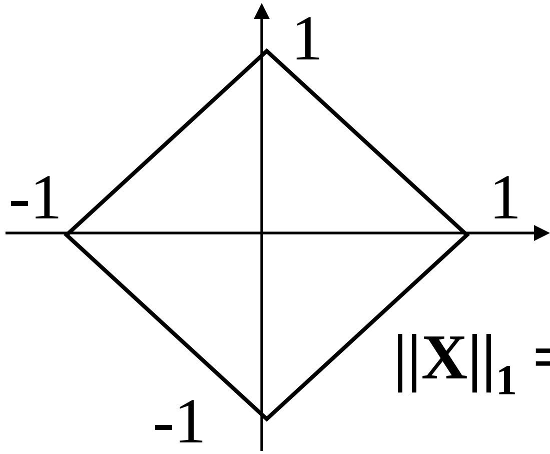
（3） $$\|X\|_\infty \le \|X\|_2 \le \sqrt{n}\|X\|_\infty$$

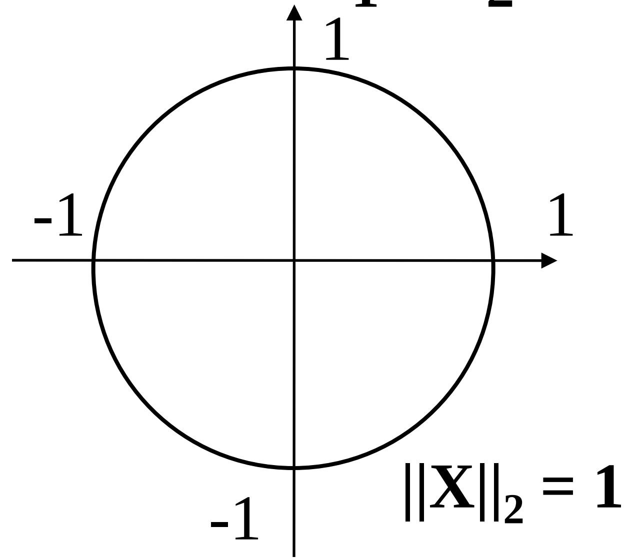
证明（2）： $$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

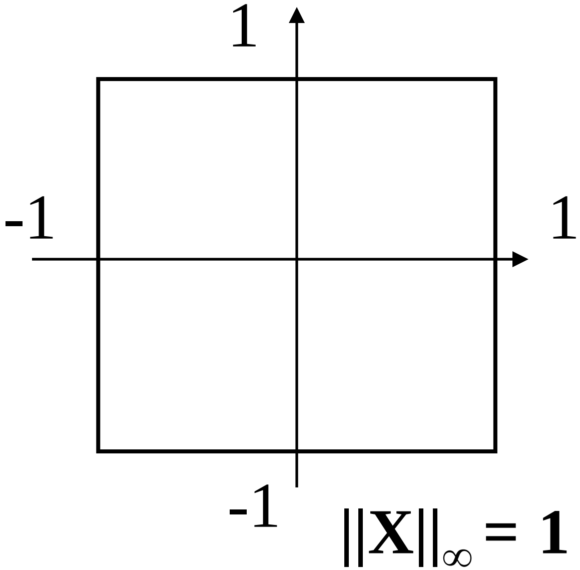
$$\max_{1\le k\le n}|x_k| \le \|x\|_1 \le n\times\max_{1\le k\le n}|x_k|$$

所以 $$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$$

例. 范数意义下的单位向量: X=[x_1, x_2]^T







$$\| X \|_1 = | x_1 | + | x_2 |$$

$$\| X \|_2 = \sqrt{x_1^2 + x_2^2}$$

$$\| X \|_\infty = \max\{| x_1 |, | x_2 |\}$$

范数性质

(3) Cauchy-Schwarz 不等式

定理： $|(x,y)| \le \|x\|_2 \cdot \|y\|_2$

(4) 向量序列的收敛性

定义： 设 $\{x^{(k)}\}$ 是 R^n 中的一个向量序列，其中

$$x^{(k)} = \left[x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}\right]^T$$

如果 $\lim\limits_{k\to\infty} x_i^{(k)} = x_i$ ，则称 $x^{(k)}$ 收敛到 x ，记为 $\lim\limits_{k\to\infty} x^{(k)} = x$

定理： 设 $\|\cdot\|$ 是 R^n 上的任意一个向量范数，则

$$\lim_{k\to\infty} x^{(k)} = \tilde{x} \iff \lim_{k\to\infty} \left\| x^{(k)} - \tilde{x} \right\| = 0$$

矩阵范数

矩阵范数

定义： 设函数 $f: R^{n\times n} \to R$，若 f 满足

(1) $f(A) \geq 0$，$\forall A \in R^{n\times n}$，且 $f(A) = 0 \Leftrightarrow A = 0$ （正定性）

(2) $f(\alpha A) = |\alpha| \cdot f(A)$，$\forall A \in R^{n}, \forall \alpha \in R$ （齐次性）

(3) $f(A+B) \leq f(A) + f(B)$ （三角不等式）

(4) $f(AB) \leq f(A)f(B)$ （相容性）

则称 f 为 $R^{n\times n}$ 上的（矩阵）范数，通常记为 $\|\cdot\|$

常见矩阵范数

- **常见的矩阵范数**

(1) **F-范数 (Frobenious 范数):** $$\|A\|_F = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}$$

(2) **算子范数 (从属范数、诱导范数)**

$$\|A\| = \sup_{\substack{x\in R^n\\ x\neq 0}} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

其中 $\|\cdot\|$ 是 R^n 上的任意一个向量范数

注：教材上的定义不太严谨 $\|A\| = \max_{\|x\|\neq 0} \frac{\|Ax\|}{\|x\|}$

算子范数

- **常见的算子范数**

① 1-范数（列范数） ⟺ $\|A\|_1 = \max\limits_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$

② 2-范数（谱范数） ⟺ $\|A\|_2 = \sqrt{\rho(A^T A)}$

③ ∞-范数（行范数） ⟺ $\|A\|_\infty = \max\limits_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$

算子范数举例

例： 设 $A=\begin{pmatrix} 1 & -2 \\ -3 & 4 \end{pmatrix}$ 计算 $\|A\|_1, \|A\|_2, \|A\|_\infty, \|A\|_F$

矩阵范数性质

- **矩阵范数的性质**

(1) 连续性：设 f 是 $R^{n\times n}$ 上的任一矩阵范数，则 f 关于 A 的每个分量是连续的。

(2) 等价性：设 $\|\cdot\|_s$ 和 $\|\cdot\|_t$ 是 $R^{n\times n}$ 上的任意两个矩阵范数，则存在常数 c_1 和 c_2，使得对任意的 $A\in R^{n\times n}$ 有

$$c_1\|A\|_s \le \|A\|_t \le c_2\|A\|_s$$

(3) 若 A 是对称矩阵，则 $\|A\|_2 = \rho(A)$

算子范数性质

- **算子范数的性质**

定理：设 $\|\cdot\|$ 是任一算子范数，则 $\rho(A) \le \|A\|$

注：该性质对 F-范数也成立。

定理：对任意 $\varepsilon > 0$, 总存在一算子范数 $\|\cdot\|_{\varepsilon}$，使得

$$\|A\|_{\varepsilon} \le \rho(A) + \varepsilon$$

算子范数性质

- **算子范数的性质**

定理： 设 $\|\cdot\|$ 是 R^n 上的任一向量范数，其对应的算子范数也记为 $\|\cdot\|$，则有

$$\|Ax\| \le \|A\| \cdot \|x\|$$

证明：直接由算子范数定义可得

⇒ 该性质就是矩阵范数与向量范数的**相容性**

定理： 设 $\|\cdot\|$ 是任一算子范数，若 $\|B\|<1$，则 $I \pm B$ 非奇异，且

$$\left\|(I \pm B)^{-1}\right\| \le \frac{1}{1-\|B\|}$$

5.5 误差分析

引例. Hilbert矩阵的病态性

$$A=\begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{bmatrix} \qquad b=\begin{bmatrix} 1 \\ 2 \\ 1.41 \\ 2 \end{bmatrix}$$

$$b_1=\begin{bmatrix} 1 \\ 2 \\ 1.42 \\ 2 \end{bmatrix}$$

方程组 $Ax=b$ 的解为 x

方程组 $Ax=b_1$ 的解为x_1

数据计算结果

$$x-x_1=[\ -2.4 \quad 27.0 \quad -64.8 \quad 42.0\ ]^T$$

● 什么是病态矩阵

定义：考虑线性方程组 $Ax=b$，如果 A 或 b 的微小变化会导致解的巨大变化，则称此线性方程组是病态的，并称矩阵 A 是病态的，反之则是良态的。

例：
$$\begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 2 \\ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 2 \\ 2.0001 \end{bmatrix} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

结论：（1）的常数项b的第二个分量只有1/1000的微小变化，方程组的解变化却很大。

定义　若矩阵A或常数项b的微小变化引起方程组$Ax=b$的解的巨大变化，则称此方程组为病态方程组，A为病态矩阵（相对方程组而言）；否则称方程组为良态方程组，A为良态矩阵。

研究方程组中A或b的微小误差对解的影响的分析称“扰动分析”。

设$Ax=b$的扰动方程组为$(A+\delta A)(x+\delta x)=b+\delta b$，

其中，δA叫A的扰动矩阵，

δx和δb叫x和b的扰动向量。

设$Ax=b$的扰动方程组为$(A+\delta A)(x+\delta x)=b+\delta b$，下面进行扰动分析：

(1) $\delta A=0$，则$A(x+\delta x)=b+\delta b$，减去$Ax=b$，得

$A\delta x=\delta b$，故 $\delta x=A^{-1}\delta b$，即$\|\delta x\|\le\|A^{-1}\|\|\delta b\|$，又由 $Ax=b$，有$\|b\|\le\|A\|\|x\|$，所以

$$\frac{\|\delta x\|}{\|x\|}\le\|A^{-1}\|\cdot\|A\|\cdot\frac{\|\delta b\|}{\|b\|}$$

(2) $\delta b=0$，则$(A+\delta A)(x+\delta x)=b$，同理可得

$$\frac{\|\delta x\|}{\|x+\delta x\|}\le\|A^{-1}\|\cdot\|A\|\cdot\frac{\|\delta A\|}{\|A\|}$$

矩阵条件数

- **如何判别矩阵是否病态 —— 矩阵的条件数**

定义： 设 A 非奇异，则称

$$\mathrm{Cond}(A)_v = \|A^{-1}\|_v \|A\|_v$$

为 A 的**条件数**，其中 v 是 1，2 或 ∞ 。

矩阵条件数

定理：考虑线性方程组 $Ax=b$，设 b 是精确的，A 有微小的变化 δA，此时的解为 $x+\delta x$ 。假定 $\|A^{-1}\|\|\delta A\|<1$ ，则

$$\frac{\|\delta x\|}{\|x\|} \le \frac{\|A^{-1}\|\|A\|\dfrac{\|\delta A\|}{\|A\|}}{1-\|A^{-1}\|\|A\|\dfrac{\|\delta A\|}{\|A\|}}$$

- 当 δA 充分小时，不等式右端约为 $\|A^{-1}\|\|A\|\dfrac{\|\delta A\|}{\|A\|}$

- 通常，当 A 的条件数较大时，就称 A 就是病态的
- 一般来说，条件数越大，病态越严重，此时就越难用一般方法求得线性方程组的比较精确的解。

矩阵条件数

- **条件数与范数有关，常用的有无穷范数和2-范数**

$$\mathrm{Cond}(A)_\infty = \left\|A^{-1}\right\|_\infty \left\|A\right\|_\infty$$

$$\mathrm{Cond}(A)_2 = \left\|A^{-1}\right\|_2 \left\|A\right\|_2$$

注：$\mathrm{Cond}(A)_2$ 称为谱条件数，当 A 对称时有

$$\mathrm{Cond}(A)_2 = \frac{\max\limits_{1\le i\le n}|\lambda_i|}{\min\limits_{1\le i\le n}|\lambda_i|}$$

条件数性质

条件数的性质

- $\text{Cond}(A) \geq 1$
- $\text{Cond}(\alpha A) = \text{Cond}(A)$，其中 α 为任意非零实数
- 若 R 是正交矩阵，则 $\text{Cond}(R)_2=1$
- 若 R 是正交矩阵，则对任意非奇异矩阵 A，有 $\text{Cond}(AR)_2=\text{Cond}(RA)_2=\text{Cond}(A)_2$
- 条件数小，扰动引起的解的相对误差一定小；条件数大，扰动引起的解的误差可能很大（条件数与所取的范数有关，最常用的是$\|\cdot\|_\infty$和$\|\cdot\|_2$）

注： 一般判断矩阵是否病态，并不计算A^{-1}，而由经验得出。

- 行列式很大或很小（如某些行、列近似相关）；
- 元素间相差大数量级，且无规则；
- 主元消去过程中出现小主元；
- 特征值相差大数量级。

例 $A=\begin{pmatrix} 1 & 0.99 \\ 0.99 & 0.98 \end{pmatrix}$，$\vec{b}=\begin{pmatrix} 1.99 \\ 1.97 \end{pmatrix}$ 精确解为 $\vec{x}=\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

计算$cond\,(A)_2$。

$$A^{-1}=\begin{pmatrix} -9800 & 9900 \\ 9900 & -10000 \end{pmatrix}$$

举例

例： $A=\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix}$ 计算 $\mathrm{Cond}(A)_\infty$ 和 $\mathrm{Cond}(A)_2$

解：

$$A^{-1}=\begin{bmatrix}10001 & -10000\\ -10000 & 10000\end{bmatrix}$$

$$\mathrm{Cond}(A)_\infty=\|A^{-1}\|_\infty\,\|A\|_\infty\approx 4\times10^4$$

A 对称，且 $\lambda(A)=\dfrac{2.0001\pm\sqrt{2.0001^2-0.0004}}{2}>0$

$$\mathrm{Cond}(A)_2=\lambda_{max}/\lambda_{min}\approx 4\times10^4$$

举例

例： 计算 Cond$(H_k)_\infty$ 其中 H_k 为 k 阶 Hilbert 矩阵

解： k=1 时， Cond$(H_1)_\infty$=1

k=2 时， $H_2=\begin{bmatrix}1 & 1/2\\ 1/2 & 1/3\end{bmatrix}$， $H_2^{-1}=\begin{bmatrix}4 & -6\\ -6 & 12\end{bmatrix}$

Cond$(H_2)_\infty$=27

k=3 时， $H_3=\begin{bmatrix}1 & 1/2 & 1/3\\ 1/2 & 1/3 & 1/4\\ 1/3 & 1/4 & 1/5\end{bmatrix}$， $H_2^{-1}=\begin{bmatrix}9 & -36 & 30\\ -36 & 192 & -180\\ 30 & -180 & 180\end{bmatrix}$

Cond$(H_3)_\infty$=748

$\boxed{\text{Cond}(H_4)_\infty=28375,\ \text{Cond}(H_{10})_\infty=3.5\times10^{13}}$

解的改善（了解）

（1）：使用更高精度的数进行计算

（2）：使用迭代法提高数值解的精度

$$\boxed{x^{new} = x^{old} + \Delta x}$$

其中 Δx 是 $A\Delta x = b - Ax^{old}$ 的解。

总结

- LU分解与Gauss消去法区别
- 对称正定矩阵的Cholesky分解
- 向量与矩阵范数
- 误差分析（条件数）