

### 第三章 解线性方程组的直接方法

- 一个线性方程组：

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

- 线性代数方程组的解法分为两种：
  - 直接法：对于给定的方程组，在没有舍入误差的假设下，能在预定的运算次数内求得精确解
  - 迭代法：是基于一定的逆推格式，产生逼近方程组精确解的近似解序列



# 第一节 Gauss（高斯）消去法

- Gauss消去法是线性代数方程组的最简单而实用的解法
- 设n阶线性代数方程组形如：

$$\bullet \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = a_{1\ n+1} \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = a_{2\ n+1} \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = a_{n\ n+1} \end{cases} \quad (1)$$

其中系数 $a_{ij}$ 是实数或复数，也可以写成矩阵形式： $Ax=b$

记  $a_{ij}^{(1)} = a_{ij}$ ,  $i=1,2,\dots,n, j=1,2,\dots,n+1$ , 则有 $A^{(1)}=A$ ,  $b^{(1)}=b$

方程组记为： $A^{(1)}x=b^{(1)}$



## 第1次消去

- 设 $a_{11}^{(1)} \neq 0$ , 在(1)式中利用第1个方程乘以数 $-a_{i1}^{(1)}/a_{11}^{(1)}$ 加至第i个方程, 从而消去第i个方程中的 $x_1$ ,  $i = 2, \dots, n$ , 方程组(1)变换为:

$$\left\{ \begin{array}{l} a_{11}^{(1)} x_1 + a_{12}^{(1)} x_2 + \cdots + a_{1n}^{(1)} x_n = a_{1n+1}^{(1)} \\ a_{22}^{(2)} x_2 + \cdots + a_{2n}^{(2)} x_n = a_{2n+1}^{(2)} \\ \vdots \\ a_{n2}^{(2)} x_2 + \cdots + a_{nn}^{(2)} x_n = a_{nn+1}^{(2)} \end{array} \right.$$



## 第1次消去

- 其中后n-1个方程的系数的计算公式为：

$$l_{i1} = a_{i1}^{(1)} / a_{11}^{(1)}, \quad a_{ij}^{(2)} = a_{ij}^{(1)} - l_{i1} a_{1j}^{(1)}, \quad i=2, \dots, n \quad j=2, \dots, n+1$$

这样，方程组写为：  $A^{(2)}x=b^{(2)}$

第1次消去完成。

继续这个过程。

下面看一般情况。



## 第k次消去

- 若已得到第k-1次消去后的结果：  $A^{(k)}x=b^{(k)}$

其形式是：

$$\left\{ \begin{array}{l} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots + a_{1k}^{(1)}x_k + \cdots + a_{1n}^{(1)}x_n = a_{1\ n+1}^{(1)} \\ a_{22}^{(2)}x_2 + \cdots + a_{2k}^{(2)}x_k + \cdots + a_{2n}^{(2)}x_n = a_{2\ n+1}^{(2)} \\ \vdots \\ a_{kk}^{(k)}x_k + \cdots + a_{kn}^{(k)}x_n = a_{k\ n+1}^{(k)} \\ \vdots \\ a_{nk}^{(k)}x_k + \cdots + a_{nn}^{(k)}x_n = a_{n\ n+1}^{(k)} \end{array} \right.$$



## 第k次消去

- 设 $a_{kk}^{(k)} \neq 0$ ，则可进行第k次消去。在前式中利用第k个方程乘以数 $-a_{ik}^{(k)}/a_{kk}^{(k)}$ 加至第i个方程，从而消去第i个方程中的 $x_k$ ， $i = k + 1, \dots, n$ ，得到方程组：

$$A^{(k+1)}x = b^{(k+1)}$$

其中 $A^{(k+1)}$ 与 $A^{(k)}$ 的前k行相同， $b^{(k+1)}$ 与 $b^{(k)}$ 的前k个元素相同。后n-k个新方程的系数需要计算，计算公式为：

$$l_{ik} = a_{ik}^{(k)} / a_{kk}^{(k)}, \quad a_{ij}^{(k+1)} = a_{ij}^{(k)} - l_{ik} a_{kj}^{(k)}, \quad i = k + 1, \dots, n, \quad j = k + 1, \dots, n + 1$$

第k次消去完成。



## 第n-1次消去

- 如果依次有 $a_{kk}^{(k)} \neq 0$ ,  $k=1,2,\dots,n-1$ , 则一直可继续到第n-1次消去, 得到与最初的方程组等价的方程组:  $A^{(n)}x=b^{(n)}$

其中, 系数矩阵 $A^{(n)}$ 是上三角形, 方程组的形式如下:

$$\begin{cases} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots + a_{1n}^{(1)}x_n = a_{1\ n+1}^{(1)} \\ a_{22}^{(2)}x_2 + \cdots + a_{2n}^{(2)}x_n = a_{2\ n+1}^{(2)} \\ \vdots \\ a_{nn}^{(n)}x_n = a_{n\ n+1}^{(n)} \end{cases}$$

前面所述的这个过程, 是一个递推过程, 称为消去过程。



# 回代过程

- 在得到上述方程组之后，可以采用另一个递推过程，求出方程组的解：

$$x_n = a_{n\ n+1}^{(n)} / a_{nn}^{(n)}$$

$$x_i = (a_{i\ n+1}^{(i)} - \sum_{j=i+1}^n a_{ij}^{(i)} x_j) / a_{ii}^{(i)} \quad i = n - 1, \dots, 2, 1$$

- 这个过程称为回代过程。





## 例题

- 例：使用Gauss消去法求解以下方程组：

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

## 使用条件

- Gauss消去法实现方程组求解的条件是：

$$a_{11}^{(1)} a_{22}^{(2)} \cdots a_{nn}^{(n)} \neq 0$$



- 定理1： 设方程组 $Ax=b$ 的系数矩阵 $A$ 的顺序主子式全不为零， 即

$$\Delta_i = \begin{bmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{bmatrix} \neq 0 \quad i = 1, 2, \dots, n$$

则使用Gauss消去法能实现方程组 $Ax=b$ 的求解。

- 推论： 若定理1的条件成立， 则

$$a_{11}^{(1)} = \Delta_1$$

$$a_{ii}^{(i)} = \Delta_i / \Delta_{i-1} \quad i=2, \dots, n$$



- 定理2： 设方程组 $Ax=b$ 满足定理1中的条件， 使用Gauss消去法求解时， 共需乘法次数为：

$$n^3/3+n^2-n/3$$

加减法次数为：

$$n(n-1)(2n+5)/6$$



# 矩阵的三角分解

• 令:

$$L_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1\ k} & 1 & & \\ & & -l_{k+2\ k} & & 1 & \\ & & \vdots & & & \ddots \\ & & -l_{n\ k} & & & & 1 \end{bmatrix} \quad k = 1, 2, \dots, n-1$$

其中,  $l_{ik} = a_{ik}^{(k)} / a_{kk}^{(k)}$



- $L_k$  是如下  $n-k$  个初等矩阵之积:

$$\begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & \vdots & \ddots & & & \\ & -l_{i\ k} & \dots & 1 & & \\ & & & \ddots & & \\ & & & & 1 & \end{bmatrix} \quad i = k + 1, \dots, n$$



- 可以验证以下结论：

第1次消去等价于在矩阵形式的方程 $A^{(1)}x=b^{(1)}$ 的两端左乘 $L_1$ ,

得到 $A^{(2)}x=b^{(2)}$  , 即  $A^{(2)}=L_1A^{(1)}$  ,  $b^{(2)}=L_1b^{(1)}$

一般地, 第 $k$ 次消去等价于在矩阵方程 $A^{(k)}x=b^{(k)}$ 的两端左乘 $L_k$ ,

得到 $A^{(k+1)}x=b^{(k+1)}$  , 即  $A^{(k+1)}=L_kA^{(k)}$  ,  $b^{(k+1)}=L_kb^{(k)}$

看矩阵 $A$ 的变化过程:  $L_{n-1} \dots L_2 L_1 A^{(1)} = A^{(n)}$



## 三角分解

- 定理3： 设矩阵A满足定理1的条件（即顺序主子式全不为零），则A存在唯一的分解  $A=LU$ ，其中L是单位下三角矩阵（对角线全为1的下三角矩阵），U是上三角矩阵。





# 简单说明

- 由线性代数知识可知:

$$L_{n-1} \dots L_2 L_1 A^{(1)} = A^{(n)}$$

$$L_{n-1} \dots L_2 L_1 b^{(1)} = b^{(n)}$$

$$U = A^{(n)} = L_{n-1} \dots L_2 L_1 A^{(1)}$$

- 在第一个式子中消去左侧的 $L_k$ :

$$L_1^{-1} L_2^{-1} \dots L_{n-1}^{-1} L_{n-1} \dots L_2 L_1 A^{(1)} = L_1^{-1} L_2^{-1} \dots L_{n-1}^{-1} A^{(n)}$$

则有:  $A^{(1)} = L_1^{-1} L_2^{-1} \dots L_{n-1}^{-1} A^{(n)}$

- 实际上, 可得到:

$$L = L_1^{-1} L_2^{-1} \dots L_{n-1}^{-1} = \begin{bmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ & \dots & & \ddots & \\ l_{n1} & l_{n2} & \dots & l_{nn-1} & 1 \end{bmatrix}$$



## 例题

- 例：将以下方程组的系数矩阵进行三角分解：

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

## 第二节 Gauss主元素消去法

- Gauss消去法第k次消去时，需要满足的条件是 $a_{kk}^{(k)} \neq 0$ ，这个元素是实现第k次消去时的关键元素，称为第k次消去的主元素。
- 特点：不进行行交换、列交换。
- 存在的问题：实际求解过程中，如果主元素为零，或者非常小，都可能使消去过程中断（不能得到解），或求得的解的误差非常大。  
(大家思考原因是什么？)
- 改进的办法：采用主元素消去法  
分为两种：完全主元素消去法、列主元素消去法



# 完全主元素消去法

- 记 $A^{(1)}=A$ ,  $b^{(1)}=b$ , 方程组形式如下:

$$[A^{(1)}|b^{(1)}] = \left[ \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & a_{1n+1} \\ a_{21} & a_{22} & \cdots & a_{2n} & a_{2n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & a_{nn+1} \end{array} \right]$$

- 在每一步消去之前, 先全面选择主元素, 然后进行行交换、列交换。之后, 再进行Gauss消去。



# 完全主元素消去法

- 看一般情形。设经过 $k-1$  ( $1 \leq k \leq n-1$ ) 次选主元、行交换、列交换及消去后，已经将矩阵 $[A^{(1)}|b^{(1)}]$  约化为如下形式：

$$[A^{(k)}|b^{(k)}] = \left[ \begin{array}{cccccc|c} a_{11}^{(k)} & a_{12}^{(k)} & \cdots & a_{1k}^{(k)} & \cdots & a_{1n}^{(k)} & a_{1\ n+1}^{(k)} \\ & a_{22}^{(k)} & \cdots & a_{2k}^{(k)} & \cdots & a_{2n}^{(k)} & a_{2\ n+1}^{(k)} \\ & & \cdots & \vdots & \cdots & \vdots & \vdots \\ & & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & a_{k\ n+1}^{(k)} \\ & & & \vdots & & \vdots & \vdots \\ & & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & a_{n\ n+1}^{(k)} \end{array} \right]$$

- 下面看第 $k$ 次消去。



# 完全主元素消去法第k次消去

- 首先，全面选主元素。

- 在矩阵 $[A^{(k)}|b^{(k)}]$ 中的红色方框内选取绝对值最大的主元素 $a_{ikjk}^{(k)}$ ，即

$$|a_{ikjk}^{(k)}| = \max_{k \leq i, j \leq n} |a_{ij}^{(k)}|$$

- 然后，交换 $[A^{(k)}|b^{(k)}]$ 中第k行与第 $i_k$ 行、第k列与第 $j_k$ 列

- 结果：元素 $a_{ikjk}^{(k)}$ 放在主元素的位置上

- 含义：

- 第k行与第 $i_k$ 行：在原方程组中交换两个方程的次序（对解没有任何影响）
  - 第k列与第 $j_k$ 列：在原方程组中交换对应于第k个及第 $j_k$ 个分量，之后需要还原



# 选主元素的可能性

- 如果出现某个主元素为零，则方程组系数矩阵A必为奇异矩阵，方程组无解。
- 否则，经过n-1次消去后，化为如下的形式：

$$\begin{bmatrix} a_{11}^{(n)} & a_{12}^{(n)} & \cdots & a_{1n}^{(n)} \\ & a_{22}^{(n)} & \cdots & a_{2n}^{(n)} \\ & & \cdots & \vdots \\ & & & a_{nn}^{(n)} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{1\ n+1}^{(n)} \\ a_{2\ n+1}^{(n)} \\ \vdots \\ a_{n\ n+1}^{(n)} \end{bmatrix}$$

- 通过回代过程，可以求出 $y_1, y_2, \cdots, y_n$ ，这里， $y_1, y_2, \cdots, y_n$ 是 $x_1, x_2, \cdots, x_n$ 的某种排列，通过还原即可得到 $x_1, x_2, \cdots, x_n$ 的解。



## 完全主元素方法的特点

- 因为是在剩余待选区域内选择绝对值最大的元素，而在计算机上目前只能通过比较操作进行，所以时间开销比较大。
- 如果选出的主元素为零，可以立即判定原方程组无解。
- 替代的方法：列主元素消去法。





# 列主元素消去法

- 完全主元素的时间开销较大，可以采用另一种替代方法：列主元素消去法。
- 列主元素消去法是在局部范围内选取主元素，这个局部范围即是当前候选区域中的第1列（红色框内）。

$$[A^{(k)}|b^{(k)}] = \left[ \begin{array}{cccccc|c} a_{11}^{(k)} & a_{12}^{(k)} & \cdots & a_{1k}^{(k)} & \cdots & a_{1n}^{(k)} & a_{1\ n+1}^{(k)} \\ & a_{22}^{(k)} & \cdots & a_{2k}^{(k)} & \cdots & a_{2n}^{(k)} & a_{2\ n+1}^{(k)} \\ & & \cdots & \vdots & \cdots & \vdots & \vdots \\ & & & \cdots & a_{kk}^{(k)} & \cdots & a_{k\ n+1}^{(k)} \\ & & & & \vdots & \vdots & \vdots \\ & & & \cdots & a_{nk}^{(k)} & \cdots & a_{n\ n+1}^{(k)} \end{array} \right]$$



# 列主元素消去法

- 因为是在 $a_{kk}^{(k)} \ a_{k+1k}^{(k)} \ \dots \ a_{nk}^{(k)}$ 之间选择主元素，所以有以下两个特点：
  - 候选元素个数少，元素之间的比较次数就少，时间开销较小
  - 因为只涉及到这一列，所以只相当于是改变了方程的次序，求解后不需要还原
  - 主元素为零时，并不能得出方程组无解的结论
  - 列主元素消去法得到的解，可能误差较大



# 列主元素消去法的矩阵变换

- 基于线性代数的知识，矩阵进行两行互换，相当于对矩阵左乘初等矩阵；进行两列互换，相当于对矩阵右乘初等矩阵。

- 设 $L_k$ 的定义如下：

$$L_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1\ k} & 1 & & \\ & & -l_{k+2\ k} & & 1 & \\ & & \vdots & & & \ddots \\ & & -l_{n\ k} & & & & 1 \end{bmatrix} \quad k = 1, 2, \dots, n-1$$

- $I_{ij}$ 是初等排列矩阵，即将单位矩阵互换第 $i$ 行和第 $j$ 行之后的矩阵。使用 $I_{ij}$ 左乘矩阵 $A$ ，相当于互换 $A$ 的第 $i$ 行和第 $j$ 行。
  - $I_{ij} I_{ij} = I$  （存在逆矩阵，且逆矩阵等于自己）

- 这样，我们有

$$L_1 I_{1\ i1} A^{(1)} = A^{(2)} , \quad L_1 I_{1\ i1} b^{(1)} = b^{(2)}$$

- 一般情况可表示为：

$$L_k I_{k\ ik} A^{(k)} = A^{(k+1)} , \quad L_k I_{k\ ik} b^{(k)} = b^{(k+1)}$$

- 经过所有n-1步消去的矩阵变换如下：

$$L_{n-1} I_{n-1\ in-1} \cdots L_2 I_{2\ i2} L_1 I_{1\ i1} A^{(1)} = A^{(n)} = U$$

$$\text{令： } \tilde{P} = L_{n-1} I_{n-1\ in-1} \cdots L_2 I_{2\ i2} L_1 I_{1\ i1}$$



- 接下来，要简化矩阵A左侧的矩阵串 $\tilde{P}$ 。以n=4为例。

$$\tilde{P} = \mathbf{L}_3 \mathbf{I}_{3 \ i3} \mathbf{L}_2 \mathbf{I}_{2 \ i2} \mathbf{L}_1 \mathbf{I}_{1 \ i1}$$

$$= \mathbf{L}_3 (\mathbf{I}_{3 \ i3} \mathbf{L}_2 \mathbf{I}_{3 \ i3}) (\mathbf{I}_{3 \ i3} \mathbf{I}_{2 \ i2} \mathbf{L}_1 \mathbf{I}_{2 \ i2} \mathbf{I}_{3 \ i3}) (\mathbf{I}_{3 \ i3} \mathbf{I}_{2 \ i2} \mathbf{I}_{1 \ i1})$$

$$= \tilde{L}_3 \tilde{L}_2 \tilde{L}_1 P$$

$$\tilde{L}_3 = \mathbf{L}_3 \quad (\text{单位下三角阵, 元素绝对值} \leq 1)$$

$$\tilde{L}_2 = \mathbf{I}_{3 \ i3} \mathbf{L}_2 \mathbf{I}_{3 \ i3} \quad (\text{单位下三角阵, 元素绝对值} \leq 1)$$

$$\tilde{L}_1 = \mathbf{I}_{3 \ i3} \mathbf{I}_{2 \ i2} \mathbf{L}_1 \mathbf{I}_{2 \ i2} \mathbf{I}_{3 \ i3} \quad (\text{单位下三角阵, 元素绝对值} \leq 1)$$

$$P = \mathbf{I}_{3 \ i3} \mathbf{I}_{2 \ i2} \mathbf{I}_{1 \ i1} \quad (\text{排列阵})$$



- 得到:  $\tilde{L}_3\tilde{L}_2\tilde{L}_1\mathbf{P}\mathbf{A}=\mathbf{U}$

记  $\tilde{L}_3\tilde{L}_2\tilde{L}_1=\mathbf{L}^{-1}$ , 则有  $\mathbf{P}\mathbf{A}=\mathbf{L}\mathbf{U}$

其中,  $\mathbf{L}$ 是单位下三角矩阵,  $\mathbf{U}$ 是上三角矩阵,  $\mathbf{P}$ 是排列矩阵

- 定义: 如果矩阵 $\mathbf{P}$ 可由单位矩阵经若干次行交换而得到, 即 $\mathbf{P}$ 等于若干个 $\mathbf{I}_{ij}$ 类型的初等矩阵之积, 则称 $\mathbf{P}$ 为排列矩阵。
- 当用一个排列矩阵左乘某个矩阵时, 将实现矩阵行的重排。对于矩阵的列也有类似的结论, 即用一个排列矩阵右乘某个矩阵时, 将实现矩阵列的重排。



- 列主元素消去法中，依次按列选主元素，实际上就是确定排列矩阵 $P$ ，使得方程组重排后得到的等价方程组 $PAx=Pb$ ，能使用Gauss消去法求解。
- $PA=LU$ 式子的含义：（满足条件的）系数矩阵 $A$ 经过某些行的互换之后，可以分解为一个单位下三角矩阵和一个上三角矩阵的乘积。
- 定理1：设 $A$ 是非奇异矩阵，则存在排列矩阵 $P$ 及单位下三角矩阵 $L$ 和上三角矩阵 $U$ ，使得  $PA=LU$  （列主元素三角分解）
- 定理2：设 $A$ 是非奇异矩阵，则存在排列矩阵 $P$ 和 $Q$ ，及单位下三角矩阵 $L$ 和上三角矩阵 $U$ ，使得  $PAQ=LU$  （完全主元素三角分解）



# Gauss-Jordan消去法

- Gauss消去法在消去时，仅消去对角线下方的元素，然后通过回代得出方程组的解。G-J消去法对Gauss消去法做了修改：同时消去对角线下方及上方的元素，这样，可以取消回代过程。
- G-J消去后，将 $[A|b]$ 约化为如下的形式：

$$[A^{(n)}|b^{(n)}] = \left[ \begin{array}{cccc|c} 1 & 0 & \cdots & 0 & a_{1\ n+1}^{(n)} \\ 0 & 1 & \cdots & 0 & a_{2\ n+1}^{(n)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{n\ n+1}^{(n)} \end{array} \right]$$

- A约化为单位阵，方程组的解为： $x_i = a_{i\ n+1}^{(n)} \quad i = 1, 2, \dots, n$



## G-J消去法的优势

- G-J消去法没有回代过程，但并没有因此而在此时间开销上有所降低，因为回代过程的工作分摊在前面的消去过程中了。求解线性方程组时，其乘除法的次数及加减法的次数，均多于使用Gauss消去法，也就是说，效率不如Gauss消去法。
- G-J消去法的优势：主要用于求逆矩阵。



## 使用G-J消去法求逆矩阵

- 设 $A=(a_{ij})_n$ 为非奇异矩阵， $I$ 为 $n$ 阶单位矩阵，将增广矩阵 $[A|I]$ 约化为 $[I|B]$ 的形式，则 $A^{-1}=B$ 。

- 例 $A = \begin{bmatrix} 2 & -2 & 5 \\ 2 & 3 & 4 \\ 3 & -1 & 3 \end{bmatrix}$ ，求 $A$ 的逆矩阵

$$[A|I] = \begin{bmatrix} 2 & -2 & 5 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & -1 & 3 & 0 & 0 & 1 \end{bmatrix}$$

## 示例

$$= \begin{bmatrix} 1 & 0 & 0 & -13/41 & -1/41 & 23/41 \\ 0 & 1 & 0 & -6/41 & 9/41 & -2/41 \\ 0 & 0 & 1 & 11/41 & 4/41 & -10/41 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} -13/41 & -1/41 & 23/41 \\ -6/41 & 9/41 & -2/41 \\ 11/41 & 4/41 & -10/41 \end{bmatrix}$$

### 第三节 Gauss消去法的变形

- 若矩阵 $A=(a_{ij})_n$ 的顺序主子式均不为零，则 $A$ 可分解为单位下三角矩阵 $L$ 与上三角矩阵 $U$ 的乘积： $A=LU$ ，称这种分解为矩阵 $A$ 的LU分解。
- 借助于LU分解，则求解 $Ax=b$ 时，可将求解过程归结为利用递推方式相继求解两个三角形方程组：

$$Ly=b \text{ 求 } y$$

$$Ux=y \text{ 求 } x$$



# Doolittle分解法

- 求解LU分解时，并不像之前那样通过Gauss消去法。而是直接对矩阵A进行分解。

- 设矩阵 $A=(a_{ij})_n$ 有LU分解，记
$$L = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \dots & 1 \end{bmatrix} \quad U = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ & u_{22} & \dots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{bmatrix}$$

- 利用矩阵乘法，可以推出： $a_{1i} = u_{1i} (i = 1, 2, \dots, n)$

$$a_{i1} = l_{i1}u_{11} \quad l_{i1} = a_{i1}/u_{11} \quad (i = 2, 3, \dots, n)$$

- 这样可得出U的第一行元素和L的第一列元素。

## Doolittle分解法

- 设已得到U的第1行到第r-1行元素，L的第1列到第r-1列元素，下面求U的第r行及L的第r列元素：
- 由矩阵乘法有：
$$a_{ri} = \sum_{k=1}^n l_{rk} u_{ki} = \sum_{k=1}^{r-1} l_{rk} u_{ki} + u_{ri}$$
- 故：
$$u_{ri} = a_{ri} - \sum_{k=1}^{r-1} l_{rk} u_{ki} \quad (i = r, r+1, \dots, n)$$
- 又有：
$$a_{ir} = \sum_{k=1}^n l_{ik} u_{kr} = \sum_{k=1}^{r-1} l_{ik} u_{kr} + l_{ir} u_{rr}$$

## 计算公式

- 由此得到以下计算公式：

$$(1) \ u_{1i} = a_{1i} \quad (i = 1, 2, \dots, n)$$

$$l_{i1} = a_{i1}/u_{11} \quad (i = 2, 3, \dots, n)$$

- 计算U的第r行、L的第r列元素， $r=2, 3, \dots, n$

$$(2) \ u_{ri} = a_{ri} - \sum_{k=1}^{r-1} l_{rk} u_{ki} \quad (i = r, r+1, \dots, n)$$

$$(3) \ l_{ir} = (a_{ir} - \sum_{k=1}^{r-1} l_{ik} u_{kr})/u_{rr} \quad (i = r+1, r+2, \dots, n, r \neq n)$$

# 计算公式

- 求解 $Ly=b$  及  $Ux=y$ 的计算公式

$$y_1 = b_1$$

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

$$x_n = y_n / u_{nn}$$

$$x_i = (y_i - \sum_{k=i+1}^n u_{ik} x_k) / u_{ii} \quad i = n-1, \dots, 1$$





## 实现细节

- 这种直接三角分解法称为**Doolittle分解**。
- 计算效率：大约需要 $n^3/3$ 次乘除法。
- 空间存储上，可利用原存储空间保存L和U，如下所示：

$u_{11}$	$u_{12}$	$u_{13}$	---	$u_{1n}$	第一步
$l_{21}$	$u_{22}$	$u_{23}$	---	$u_{2n}$	第二步
$l_{31}$	$l_{32}$				
$\vdots$	$\vdots$				
$l_{n1}$	$l_{n2}$			$u_{nn}$	第n步



## 示例

- 例：用直接三角分解（Doolittle）法解

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 14 \\ 18 \\ 20 \end{pmatrix}$$

- 矩阵的LU分解为：

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -5 & 1 \end{pmatrix} \quad U = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -4 \\ 0 & 0 & -24 \end{pmatrix} \quad y = \begin{pmatrix} 14 \\ -10 \\ -72 \end{pmatrix} \quad x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

## 列主元素三角分解法

- 求解线性方程组时，需要满足一定的条件。
- 三角分解也不例外。当不满足条件时（请思考条件是什么？），也有对应的主元素三角分解，比如列主元素三角分解法。

## 平方根法（cholesky分解法）

- 当矩阵A是正定矩阵时，对A的三角分解方法可以有显著的改进。
- 定理1：若A是正定矩阵，则存在唯一的对角元素均为正数的下三角矩阵L，使

$$A=LL^T$$



# 计算公式

- A的 $LL^T$ 分解过程

对  $j=1,2,\dots,n$

(1) 
$$l_{jj} = (a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2)^{1/2}$$

(2) 对  $i=j+1, \dots, n$

$$l_{ij} = (a_{ij} - \sum_{k=1}^{j-1} l_{ik}l_{jk})/l_{jj}$$



## 计算公式

• 解  $Ly=b$  与  $L^Tx=y$

(1) 对  $i=1,2,\dots,n$

$$y_i = (b_i - \sum_{k=1}^{i-1} l_{ik}y_k)/l_{ii}$$

(2) 对  $i=n, n-1,\dots,1$

$$x_i = (y_i - \sum_{k=i+1}^n l_{ki}x_k)/l_{ii}$$

- 定理2: 用cholesky分解法求解n阶正定方程组共需

n 次开平方

$(n^3+9n^2+2n)/6$  次乘除运算

$(n^3+6n^2-7n)/6$  次加减运算

- cholesky分解法所需的计算量大约为Gauss消去法的一半
- 正定性保证了不需选主元

• 例：求解方程组

$$4x_1 - x_2 + x_3 = 6$$

$$-x_1 + 4.25x_2 + 2.75x_3 = -0.5$$

$$x_1 + 2.75x_2 + 3.5x_3 = 1.25$$

系数矩阵的顺序主子式：

$\Delta_1=4$     $\Delta_2=16$     $\Delta_3=16$  均大于零，故系数矩阵为正定阵



## 求解过程

- j=1:

$$l_{11}=(a_{11})^{1/2}=4^{1/2}=2 \quad l_{21}=a_{21}/l_{11}=-1/2=-0.5 \quad l_{31}=a_{31}/l_{11}=1/2=0.5$$

- j=2:

$$l_{22}=(a_{22}-l_{21}^*l_{21})^{1/2}=(4.25-(-0.5)^2)^{1/2}=2$$

$$l_{32}=(a_{32}-l_{31}^*l_{21})/l_{22}=(2.75-0.5^*(-0.5))/2=1.5$$

- j=3:

$$l_{33}=(a_{33}-l_{31}^*l_{31}-l_{32}^*l_{32})^{1/2}=(3.5-0.5^2-1.5^2)^{1/2}=1$$

- 得到矩阵L，故有：

$$\begin{bmatrix} 4 & -1 & 1 \\ -1 & 4.25 & 2.75 \\ 1 & 2.75 & 3.5 \end{bmatrix} = \begin{bmatrix} 2 & & \\ -0.5 & 2 & \\ 0.5 & 1.5 & 1 \end{bmatrix} \begin{bmatrix} 2 & -0.5 & 0.5 \\ & 2 & 1.5 \\ & & 1 \end{bmatrix}$$

## 求解两个三角方程组

- 解  $Ly=b$

$$y_1=b_1/l_{11}=6/2=3$$

$$y_2=(b_2-l_{21}y_1)/l_{22}=(-0.5+0.5*3)/2=0.5$$

$$y_3=(b_3-l_{31}y_1-l_{32}y_2)/l_{33}=(1.25-0.5*3-1.5*0.5)/1=-1$$

- 解  $L^T x=y$

$$x_3=y_3/l_{33}=-1/1=-1$$

$$x_2=(y_2-l_{32}x_3)/l_{22}=(0.5+1.5)/2=1$$

$$x_1=(y_1-l_{21}x_2-l_{31}x_3)/l_{11}=(3-(-0.5)*1-0.5*(-1))/2=2$$

# 改进的平方根法

- 求解系数矩阵是正定矩阵的线性方程组时，可以采用chelesky分解法求解，但其中不能避免开平方运算
- 现在改进这个方法，从而避免开方（开方不能必须的）
- 如果正定矩阵A采用形如LDL<sup>T</sup>这样的分解形式，则可以避免开方运算：

$$A=LDL^T$$

$$A = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \dots & 1 \end{bmatrix} \begin{matrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{matrix} \begin{bmatrix} 1 & l_{21} & \dots & l_{n1} \\ & 1 & \dots & l_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$



# 改进的平方根法算法

- A的LDL<sup>T</sup>分解过程

对  $i=1, 2, \dots, n$

(1) 对  $j=1, 2, \dots, i-1$

$$t_{ij} = a_{ij} - \sum_{k=1}^{j-1} t_{ik} l_{jk}$$

(2) 对  $j=1, 2, \dots, i-1$

$$l_{ij} = t_{ij} / d_j$$

(3)  $d_i = a_{ii} - \sum_{k=1}^{i-1} t_{ik} l_{ik}$



## 改进的平方根法算法

- 解  $Ly=b$  与  $DL^Tx=y$

- (1) 对  $i=1, 2, \dots, n$

$$y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k$$

- (2) 对  $i=n, n-1, \dots, 1$

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

# 追赶法

- 如果线性方程组的系数矩阵是三对角矩阵，则求解过程可进一步简化。
- 所谓三对角矩阵是如下这样的：

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



# LU分解

- 三对角矩阵也可以进行LU分解，且L是单位下三角矩阵，U是上三角矩阵。
- 可以验证，将三对角矩阵进行LU分解时，L及U仍保持与A的一致性，这意味着，L和U中都仅有两条对角线上存在非零元素，且L的主对角线元素均为1，具体形式如下：

$$L = \begin{bmatrix} 1 & & & & \\ l_2 & 1 & & & \\ & l_3 & 1 & & \\ & & \ddots & \ddots & \\ & & & l_n & 1 \end{bmatrix} \quad U = \begin{bmatrix} u_1 & v_1 & & & \\ & u_2 & v_2 & & \\ & & u_3 & \ddots & \\ & & & \ddots & v_{n-1} \\ & & & & u_n \end{bmatrix}$$





## 求解公式

$$v_i = c_i \quad i=1,2,\dots,n-1$$

$$u_1 = b_1 \quad l_i = a_i / u_{i-1} \quad u_i = b_i - l_i v_{i-1} \quad i=2,\dots,n$$

- 得到L和U后，后续的求解过程与前述一致。

# 追赶法算法

- (1)  $u_1=b_1$
- (2) 对  $i=2, 3, \dots, n$

$$l_i=a_i/u_{i-1} \quad u_i=b_i-l_i c_{i-1}$$

- (3) 解  $Ly=d$

$$y_1=d_1$$

$$\text{对 } i=2, 3, \dots, n \quad y_i=d_i-l_i y_{i-1}$$

- (4) 解  $Ux=y$

$$x_n=y_n/u_n$$

$$\text{对 } i=n-1, n-2, \dots, 1 \quad x_i=(y_i-c_i x_{i+1})/u_i$$

## “追赶”的含义

- 计算过程是：

$$u_1 \rightarrow l_2 \rightarrow u_2 \rightarrow l_3 \rightarrow \dots \rightarrow u_n \rightarrow l_n$$

$$y_1 \rightarrow y_2 \rightarrow \dots \rightarrow y_n$$

这个过程是“追”的过程

然后，计算：

$$x_n \rightarrow x_{n-1} \rightarrow \dots \rightarrow x_1$$

这个过程是“赶”的过程

## 例题

- 利用追赶法求解三对角线性方程组

$$\begin{bmatrix} 2 & -1 & & \\ -1 & 3 & -2 & \\ & -2 & 4 & -3 \\ & & -3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \\ -2 \\ 1 \end{bmatrix}$$

## 使用条件

- 定理： 设A是三对角矩阵，如下：

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

且  $|b_1| > |c_1|$  ,  $|b_n| > |a_n|$  ,  $|b_i| > |a_i| + |c_i|$ ,  $a_i c_i \neq 0$ ,  $i=2,3,\dots,n-1$

则A非奇异，并能满足使用追赶法算法求解方程组 $Ax=d$ 的条件。



## 第四节 向量和矩阵的范数

- 导引

记号： $\mathbb{R}^n$ ：表示n维实向量空间

$\mathbb{C}^n$ ：表示n维复向量空间

设  $x=(x_1, x_2, \dots, x_n)^T$  是 $\mathbb{R}^n$ 中的任一向量，在线性代数中，记x的**长度**为 $|x|$ ，定义如下：

$$|x| = (x_1^2 + x_2^2 + \dots + x_n^2)^{1/2}$$

- 它有如下的**性质**：

- $|x| \geq 0$ ，当且仅当 $x=0$ 时， $|x|=0$
- $|\alpha x| = |\alpha| |x|$
- $|x+y| \leq |x| + |y|$



## 向量范数

• 设  $x=(x_1,x_2,\dots,x_n)^T$  ,  $y=(y_1,y_2,\dots,y_n)^T$  ,  $x,y\in\mathbb{R}^n$  (或 $\mathbb{C}^n$ ) , 将实数  $(x,y)=y^Tx=\sum_{i=1}^n x_i y_i$  (或复数  $(x,y)=y^Hx=\sum_{i=1}^n x_i \bar{y}_i$ ) 称为向量 $x$ 、 $y$ 的数量积, 或内积。

• 它有如下的性质:

- $(x,y)=(y,x)$  (或  $(x,y)=\overline{(y,x)}$ )
- $(\alpha x,y)=\alpha(x,y)$   $\alpha$ 是实数 (或是  $(x,\alpha y)=\bar{\alpha}(x,y)$   $\alpha$ 是复数 )
- $(x,x)\geq 0$ , 当且仅当 $x=0$ 时,  $(x,x)=0$
- $(x_1+x_2,y)=(x_1,y)+(x_2,y)$



## 向量范数的定义

- 定义：如果向量  $\mathbf{x} \in \mathbb{R}^n$ （或  $\mathbb{C}^n$ ）的某个实值函数  $N(\mathbf{x}) = \|\mathbf{x}\|$ ，满足条件：
  - 非负性：  $\|\mathbf{x}\| \geq 0$ ，当且仅当  $\mathbf{x} = \mathbf{0}$  时，  $\|\mathbf{x}\| = 0$
  - 齐次性：对任一数  $\alpha \in \mathbb{R}$ （或  $\mathbb{C}$ ），有  $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$
  - 成立三角不等式：  $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

则称  $N(\mathbf{x})$  是  $\mathbb{R}^n$ （或  $\mathbb{C}^n$ ）上的一个向量范数（或称做模）。





## 典型的向量范数

- 定理：对 $\mathbf{R}^n$ 中的任一向量  $\mathbf{x}=(x_1, x_2, \dots, x_n)^T$

记：  $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$

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

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

则， $\|\cdot\|_1$ 、 $\|\cdot\|_2$ 和 $\|\cdot\|_\infty$ 都是向量范数，分别称为向量的1-范数、2-范数和 $\infty$ -范数。

# 证明

- 证明：1-范数、 $\infty$ -范数证明略。只证2-范数。

根据定理中给出的公式， $\|\cdot\|_2$ 非负，满足定义中的条件(1)。

对任一数 $\alpha \in \mathbb{R}$ ，有

$$\|\alpha x\|_2 = \left( \sum_{i=1}^n (\alpha x_i)^2 \right)^{1/2} = |\alpha| \left( \sum_{i=1}^n x_i^2 \right)^{1/2} = |\alpha| \|x\|_2$$

故满足条件(2)。

由 $\|\cdot\|_2$ 的公式知，可用内积表示 $\|\cdot\|_2$ ： $\|x\|_2 = (x^T x)^{1/2}$

任取向量 $y \in \mathbb{R}^n$ ，则有： $\|x + y\|_2^2 = (x + y)^T (x + y) = \|x\|_2^2 + 2x^T y + \|y\|_2^2$

## 证明

- 根据Cauchy-Schwarz不等式：  $(x, y)^2 \leq (x, x)(y, y)$ （向量形式）

对应到这里是：  $(x^T y)^2 \leq (x^T x) (y^T y)$

则有：  $\|x + y\|_2^2 \leq \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2$

因为非负性，则有：  $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$

故满足条件(3)。



## p-范数

- 2-范数又称为向量x的欧氏范数。
- 推广到一般情况：

定义向量的p-范数： $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$   $p \in [1, \infty)$

- 定理：设 $\|\cdot\|_\alpha$   $\|\cdot\|_\beta$ 是 $\mathbb{R}^n$ 上的任二种范数，则存在与x无关的常数m和M（ $0 < m \leq M$ ），使得

$$m\|x\|_\alpha \leq \|x\|_\beta \leq M\|x\|_\alpha \quad \forall x \in \mathbb{R}^n$$

- 含义：向量x的某一种范数可以任意小（大）时，该向量的其他任何一种范数也会任意小（大）

# 矩阵范数

- 以 $\mathbf{R}^{n \times n}$ 表示 $n \times n$ 矩阵集合
- 定义：称定义在 $\mathbf{R}^{n \times n}$ 上的实值函数 $\|\cdot\|$ 为矩阵范数，如果对于 $\mathbf{R}^{n \times n}$ 中的任意矩阵A和B，它满足
  - 非负性： $\|A\| \geq 0$ ，当且仅当 $A=0$ 时， $\|A\| = 0$
  - 齐次性：对任一数 $k \in \mathbf{R}$ ，有 $\|kA\| = |k| \|A\|$
  - 成立三角不等式： $\|A+B\| \leq \|A\| + \|B\|$
  - 成立三角不等式： $\|AB\| \leq \|A\| \|B\|$



# 相容

- 矩阵与向量通常会出现在一起，所以常常要求矩阵范数与向量范数之间存在某种关系，即相容关系。
- 定义：对于给定的向量范数 $\|\cdot\|$ 和矩阵范数 $\|\cdot\|$ ，如果对任一个 $\mathbf{x} \in \mathbf{R}^n$ 和任一个 $\mathbf{A} \in \mathbf{R}^{n \times n}$ ，满足： $\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|$

则称所给的矩阵范数与向量范数是相容的。

- 当定义一种矩阵范数时，应当使它能与某种向量范数相容。
- 在同一个问题中，要同时使用矩阵范数和向量范数时，这两种范数应当是相容的。

- **定理**: 设在 $\mathbb{R}^n$ 中给定了一种向量范数, 对任一矩阵 $A \in \mathbb{R}^{n \times n}$ , 令

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

则, 上面定义的 $\|\cdot\|$ 是一种矩阵范数, 并且它与所给定的向量范数相容。

证明: **先证明相容性, 再证明满足定义的4个条件。**

对任意的矩阵 $A \in \mathbb{R}^{n \times n}$ , 和任意的非零向量 $y \in \mathbb{R}^n$

$$\text{由于 } \max_{\|x\|=1} \|Ax\| \geq \left\| A \frac{y}{\|y\|} \right\| = \frac{1}{\|y\|} \|Ay\|$$

$$\text{故有: } \|Ay\| \leq \|y\| \max_{\|x\|=1} \|Ax\| = \|A\| \|y\|$$

这个等式对于 $y=0$ 显然也成立。——**满足相容性定义**

(1) 当 $A=0$ 时,  $\|A\|=0$ , 当 $A\neq 0$ 时, 必有 $\|A\|>0$

(2) 对任一数 $k\in\mathbb{R}$ , 有

$$\|kA\| = \max_{\|x\|=1} \|kAx\| = |k| \max_{\|x\|=1} \|Ax\| = |k| \|A\|$$

(3) 对 $\forall A, B \in \mathbb{R}^{n \times n}$ , 有

$$\|A+B\| = \max_{\|x\|=1} \|(A+B)x\| = \max_{\|x\|=1} \|Ax + Bx\| \leq \max_{\|x\|=1} (\|Ax\| + \|Bx\|)$$

$$\leq \max_{\|x\|=1} \|Ax\| + \max_{\|x\|=1} \|Bx\| = \|A\| + \|B\|$$



$$(4) \|AB\| = \max_{\|x\|=1} \|(AB)x\| \leq \max_{\|x\|=1} \|A\| \|Bx\| = \|A\| \max_{\|x\|=1} \|Bx\| = \|A\| \|B\|$$



- 定理中定义的范数为从属于所给定向量范数的矩阵范数。
- 设给定向量范数为 $\|\cdot\|_p$ ，则从属于向量范数 $\|\cdot\|_p$ 的矩阵范数仍记为 $\|\cdot\|_p$ ，即

$$\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p$$

其中， $A \in \mathbb{R}^{n \times n}$ ， $x \in \mathbb{R}^n$ ， $\|A\|_p$ 也称为矩阵A的p-范数。



• 定理： 设 $A=(a_{ij}) \in \mathbb{R}^{n \times n}$ ， 则

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \quad \text{列范数}$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)} \quad \text{2-范数}$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad \text{行范数}$$

其中 $\lambda_{\max}(A^T A)$ 表示矩阵 $A^T A$ 的最大特征值。



## 示例

例 设  $A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 2 & -1 \\ 0 & 1 & 1 \end{bmatrix}$ , 求矩阵A的范数

解:

$$\|A\|_1 = \max \{|1| + |-1| + |0|, |2| + |2| + |-1|, |0| + |-1| + |1|\} = 5$$

$$\|A\|_\infty = \max \{|1| + |2| + |0|, |-1| + |2| + |-1|, |0| + |1| + |1|\} = 4$$

$$A^T A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 2 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ -1 & 2 & -1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 9 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$



## 示例

- $A^T A$ 的特征多项式为:

$$\det(\lambda I - A^T A) = \lambda^3 - 13\lambda^2 + 38\lambda - 25$$

它的三个根如下:

$$\lambda_1 = 9.1428 \quad \lambda_2 = 2.921125 \quad \lambda_3 = 0.936075$$

$$\|A\|_2 = (\lambda_{\max}(A^T A))^{1/2} = (\max(\lambda_1, \lambda_2, \lambda_3))^{1/2} = 3.0237$$

## 示例

例 设  $\mathbf{x}=(3,-5,1)^T$ ,  $A = \begin{bmatrix} 1 & 5 & -2 \\ -2 & 1 & 0 \\ 3 & -8 & 2 \end{bmatrix}$ , 求向量  $\mathbf{x}$ 、矩阵  $A$  的范数

解:

$$\|\mathbf{x}\|_1=9, \quad \|\mathbf{x}\|_2=\sqrt{35}, \quad \|\mathbf{x}\|_\infty=5$$

$$\|A\|_1=\max\{6, 14, 4\}=14$$

$$\|A\|_\infty=\max\{8, 3, 13\}=13$$

$$A^T A = \begin{bmatrix} 1 & -2 & 3 \\ 5 & 1 & -8 \\ -2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 5 & -2 \\ -2 & 1 & 0 \\ 3 & -8 & 2 \end{bmatrix} = \begin{bmatrix} 14 & -21 & 4 \\ -21 & 90 & -26 \\ 4 & -26 & 8 \end{bmatrix}$$

## 示例

- $A^T A$ 的特征多项式为:

$$\det(\lambda I - A^T A) = \lambda^3 - 112\lambda^2 + 959\lambda - 16 = 0$$

最大根  $\lambda_1 \approx 102.66$

$$\|A\|_2 = (\lambda_{\max}(A^T A))^{1/2} = \sqrt{\lambda_1} \approx 10.132$$

## 第五节 误差分析

- 考虑线性方程组  $Ax=b$

其中设 $A$ 为非奇异矩阵， $x$ 为方程组的精确解。

由于 $A$ 或 $b$ 是测量或计算的结果，其中必有观测误差或是舍入误差。因此处理的实际矩阵是 $(A+\delta A)$ 或 $(b+\delta b)$ ，则我们需要研究 $A$ 或 $b$ 的微小误差对解的影响。



- 看一个示例：

$$\begin{pmatrix} 1 & 1 \\ 1 & 1.0001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

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

- 两个方程组的解相差很大：常数项的微小变化，对解的影响非常大





- 定义：如果矩阵A或常数项b的微小变化，引起方程组 $Ax=b$ 解的巨大变化，则称此方程组为“病态”方程组，矩阵A称为“病态”矩阵，否则称方程组为“良态”方程组，矩阵A称为“良态”矩阵。
- 现设A是精确的，b有误差 $\delta b$ ，解为 $x+\delta x$ ，则

$$A(x+\delta x)=b+\delta b$$

$$\delta \mathbf{x} = \mathbf{A}^{-1} \delta \mathbf{b}$$

$$\|\delta \mathbf{x}\| \leq \|A^{-1}\| \|\delta \mathbf{b}\|$$

$$\|b\| \leq \|A\| \|x\| \quad \text{即:} \quad \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|} \quad (\text{设 } b \neq 0)$$

- 定理： 设A是非奇异阵，  $Ax=b \neq 0$ ， 且

$$A(x+\delta x)=b+\delta b$$

$$\text{则 } \frac{\|\delta x\|}{\|x\|} \leq \|A^{-1}\| \|A\| \frac{\|\delta b\|}{\|b\|}$$

- 含义： 解的相对误差的上界。 即： 常数项b的相对误差在解中可能放大 $\|A^{-1}\| \|A\|$ 倍。



- 定义：设A为非奇异阵，称数

$$\text{cond}(A)_v = \|A^{-1}\|_v \|A\|_v \quad (v = 1, 2, \infty)$$

为矩阵A的条件数。

- 当A的条件数相对的大，即 $\text{cond}(A)_v \gg 1$ 时，则 $Ax=b$ 是“病态”的。条件数越大，病态越严重。当A的条件数相对的小，则 $Ax=b$ 是“良态”的。



## 常用的条件数

- 通常使用的条件数有：

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

$$(2) A \text{ 的谱条件数 } \text{cond}(A)_2 = \|A^{-1}\|_2 \|A\|_2 = \sqrt{\frac{\lambda_{\max}(A^T A)}{\lambda_{\min}(A^T A)}}$$

当A为对称矩阵时， $\text{cond}(A)_2 = \frac{\lambda_1}{\lambda_n}$ ，其中 $\lambda_1$ 、 $\lambda_n$ 为A的绝对值最大、最小的特征值。



## 条件数的性质

(1) 对任何的非奇异矩阵A，都有 $\text{cond}(A)_v \geq 1$

$$\text{cond}(A)_v = \|A^{-1}\|_v \|A\|_v \geq \|A^{-1}A\|_v = 1$$

(2) 设A为非奇异矩阵，且 $c \neq 0$ （常数），则

$$\text{cond}(cA)_v = \text{cond}(A)_v$$

(3) 如果A为正交矩阵，则 $\text{cond}(A)_2 = 1$ 。如果A为非奇异矩阵，R为正交矩阵，则 $\text{cond}(RA)_2 = \text{cond}(AR)_2 = \text{cond}(A)_2$



# 近似的方法

- 实际计算中，计算 $A^{-1}$ 时时间开销较大，故检测病态方程组时采用下述近似的方法：
  - $A$ 的三角约化时，出现小主元
  - 系数矩阵的行列式值相对很小，或系数矩阵某些行近似线性相关
  - 系数矩阵 $A$ 元素间数量级相差很大，并且无规则

