解线性代数方程组的直接法

温丹苹

邮箱：dpwen@nju.edu.cn

办公室：工管院协鑫楼306

3 平方根法

4 追赶法

5 向量和矩阵的范数

6 线性方程组的性态

3.3 平方根法

* 在工程技术问题中，常常需要求解**系数矩阵为对称正定阵**的线性代数方程组。
* 这类问题，可利用**对称正定矩阵的三角分解**而得到的求解对称正定方程组的简洁高效的方法——**平方根法**。

3.3 平方根法

∗ 由线性代数知识可知，**当矩阵A为对称正定时，A的各阶主子式都大于零**。

由**定理**可知，存在唯一的杜立特尔分解$A = LU$。

利用A的对称性，将U再分解为对角阵D和单位上三角阵U_0，则：

$$U = DU_0 = \begin{pmatrix} u_{11} & & & & \\ & u_{22} & & & \\ & & \ddots & & \\ & & & u_{n-1,n-1} & \\ & & & & u_{nn} \end{pmatrix} \begin{pmatrix} 1 & \frac{u_{12}}{u_{11}} & \frac{u_{13}}{u_{11}} & \cdots & \frac{u_{1n}}{u_{11}} \\ & 1 & \frac{u_{23}}{u_{22}} & \cdots & \frac{u_{2n}}{u_{22}} \\ & & \ddots & \ddots & \vdots \\ & & & 1 & \frac{u_{n-1,n}}{u_{n-1,n-1}} \\ & & & & 1 \end{pmatrix}$$

◆ 因此，

$$A = LU = LDU_0$$

由于A是对称阵，则$A = A^T = (LDU_0)^T = U_0^T DL^T$，

此时，U_0^T为单位下三角矩阵，DL^T为上三角阵。

由LU分解的唯一性，即得$U_0^T = L$，即$A = LDL^T$。

* **定理（对称阵的三角分解定理）** 设A为n阶对称阵，且A的所有顺序主子式均不为零，则A可唯一分解为$A = LDL^T$，其中，L为单位下三角阵，D为对角阵。

3.3 平方根法

由于$u_{ii}>0$，则 $D=\begin{pmatrix} u_{11} & & \\ & \ddots & \\ & & u_{nn} \end{pmatrix}=\begin{pmatrix} \sqrt{u_{11}} & & \\ & \ddots & \\ & & \sqrt{u_{nn}} \end{pmatrix}\begin{pmatrix} \sqrt{u_{11}} & & \\ & \ddots & \\ & & \sqrt{u_{nn}} \end{pmatrix}=D^{\frac{1}{2}}D^{\frac{1}{2}}$,

由**对称阵的三角分解定理**得，

$$A=LDL^T=LD^{\frac{1}{2}}D^{\frac{1}{2}}L^T=\left(LD^{\frac{1}{2}}\right)\left(LD^{\frac{1}{2}}\right)^T=L_1L_1^T\text{，其中}L_1=LD^{\frac{1}{2}}\text{。}$$

定理 (Cholesky 分解) 设 A 对称正定, 则存在唯一的对角线元素全为正的下三角矩阵 L, 使得

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

该分解称为 **Cholesky 分解**.

* Cholesky (乔列斯基)分解仅针对**对称正定矩阵成立**.

◆如何计算**Cholesky分解**：待定系数法？

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

类似于 LU 分解, 直接比较等式两边的元素, 可得一般公式

$$\boxed{a_{ij} = \sum_{k=1}^{n} l_{ik} l_{jk} = \sum_{k=1}^{j-1} l_{ik} l_{jk} + l_{jj} l_{ij}}, \quad j = 1, 2, \ldots, n, \ \ i = j, j+1, \ldots, n.$$

根据这个计算公式即可得 l_{ij} 的表达式.

3.3 平方根法

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

◆如何反向推导出l_{jj}和l_{ij}？

（**自左向右逐列计算**）对于$j=1,2,\ldots,n$,

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

（2）计算l_{ij}:　$l_{ij}=(a_{ij}-\sum_{k=1}^{j-1} l_{ik}l_{jk})/l_{jj}(i=j+1,\cdots,n)$

3.3 平方根法

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

◆如何求解$\boldsymbol{Ax = b}$？

(1) 求解下三角方程：$\boldsymbol{Ly = b}$, 求$\boldsymbol{y}$;

(2)求解上三角方程：$\boldsymbol{L^T x = y}$，求$\boldsymbol{x}$。

具体步骤为：

（1）求解下三角方程：$\boldsymbol{Ly = b}$：

$$\begin{cases} y_1 = \frac{b_1}{l_{11}} \\ y_i = \dfrac{(b_i - \sum\limits_{k=1}^{i-1} l_{ik} y_k)}{l_{ii}} (i = 2, \cdots, n) \end{cases}$$

（2）求解上三角方程：$\boldsymbol{L^T x = y}$：

$$\begin{cases} x_n = \frac{b_n}{l_{nn}} \\ x_i = \dfrac{(y_i - \sum\limits_{k=i+1}^{n} l_{ki} x_k)}{l_{ii}} (i = n-1, \cdots, 1) \end{cases}$$

3.3 平方根法

例： 用平方根法解方程组

$$\begin{pmatrix} 4 & 2 & -2 \\ 2 & 2 & -3 \\ -2 & -3 & 14 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \\ 4 \end{pmatrix}$$

解： 经验证方程组的系数矩阵为对称正定阵。
由公式依次计算L的第1,2,3列元素，得

$l_{11} = 2, l_{21} = 1, l_{31} = -1;$
$l_{22} = 1, l_{32} = -1;$
$l_{33} = 3.$

则：

$$L = \begin{pmatrix} 2 & & \\ 1 & 1 & \\ -1 & -2 & 3 \end{pmatrix}$$

先求解方程组Ly=b，得

$y_1 = 5, y_2 = 0, y_3 = 3.$

再求解方程组$L^T x$=y，得

$x_3 = 1, x_2 = 2, x_1 = 2.$

3.3 平方根法

◆ 记不住公式如何计算L的第1,2,3列元素？

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} l_{11} & & \\ l_{21} & l_{22} & \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} l_{11} & l_{12} & l_{13} \\ & l_{22} & l_{23} \\ & & l_{33} \end{pmatrix}$$

$k=1$ 时： $a_{11} = l_{11}^2 \Rightarrow l_{11} = \sqrt{a_{11}}$；

$a_{21} = l_{21} l_{11} \Rightarrow l_{21} = \dfrac{a_{21}}{l_{11}}$； 同理得 $l_{31} = \dfrac{a_{31}}{l_{11}}$

$k=2$ 时： $a_{22} = l_{21}^2 + l_{22}^2 \Rightarrow l_{22} = \sqrt{a_{22} - l_{21}^2}$；

$a_{32} = l_{31} l_{21} + l_{32} l_{22} \Rightarrow l_{32} = \dfrac{a_{32} - l_{31} l_{21}}{l_{22}}$

$k=3$ 时： $a_{33} = l_{31}^2 + l_{32}^2 + l_{33}^2 \Rightarrow l_{33} = \sqrt{a_{33} - \sum_{i=1}^{2} l_{3i}^2}$

3.3 平方根法

◆**平方根法优势：**

1.平方根法乘、除法为$\frac{n}{6}(n^2+9n+2)$，比顺序消去法乘除法运算次数减少一半；

2.由于A为对称阵，计算机求解时只需存储 $\frac{n(n+1)}{2}$ 个元素。

◆**平方根法劣势：**

1.平方根法需完成n次开平方运算。

改进！！！

3.3.1 改进平方根法

◆改进的平方根法是针对开方运算，为避免对角阵$D^{\frac{1}{2}}$的元素而构造的，即$A = LDL^T$。

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

$$D=\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

3.3.1 改进平方根法

计算方法：由待定系数法

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

$$\Longrightarrow \boxed{a_{ij} = d_j l_{ij} + \sum_{k=1}^{j-1} l_{ik} d_k l_{jk}}, \quad j = 1, \ldots, n, \quad i = j+1, \ldots, n.$$

$$a_{ij} = \sum_{k=1}^{n} (LD)_{ik} (L^T)_{kj} = \sum_{k=1}^{n} l_{ik} d_k l_{jk} = \sum_{k=1}^{j-1} l_{ik} d_k l_{jk} + l_{ij} d_j l_{jj}, \quad \text{（其中，} l_{jj} = 1\text{，} l_{jk} = 0 (j < k)\text{）}$$

3.3.1 改进平方根法

◆改进平方根法步骤：

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

► 计算L的元素及D的元素：
对于$i=1,2,\dots,n$,

步1：$d_1=a_{11}$

步2：$l_{ij}=(a_{ij}-\sum_{k=1}^{j-1}l_{ik}d_k l_{jk})/d_j \quad (j=1,2,\cdots,i\text{-}1)$

步3：$d_i=a_{ii}-\sum_{k=1}^{i-1}l_{ik}^2 d_k$

► 求解$Ly=b$及$DL^T x=y$：

步4：$\begin{cases} y_1=b_1 \\ y_i=b_i-\sum_{k=1}^{i-1}l_{ik}y_k\,(i=2,\cdots,n) \end{cases}$

步5：$\begin{cases} x_n=\frac{y_n}{d_n} \\ x_i=\frac{y_i}{d_i}-\sum_{k=i+1}^{n}l_{ki}x_k\,(i=n-1,\cdots,1) \end{cases}$

3.3 平方根法

例： 用改进平方根法求解方程组

$$\begin{pmatrix} 4 & 2 & -2 \\ 2 & 2 & -3 \\ -2 & -3 & 14 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \\ 4 \end{pmatrix}$$

解：根据改进平方根法公式得：

$$d_1 = a_{11} = 4,\quad l_{21} = \frac{a_{21}}{d_1} = \frac{1}{2},$$

$$d_2 = a_{22} - l_{21}^2 d_1 = 1,$$

$$l_{31} = \frac{a_{31}}{d_1} = -\frac{1}{2},\quad l_{32} = \frac{a_{32} - l_{31}l_{21}d_1}{d_2} = -2$$

$$d_3 = a_{33} - l_{31}^2 d_1 - l_{32}^2 d_2 = 9$$

故矩阵 D 和 L 分别为

$$L = \begin{pmatrix} 1 & & \\ \frac{1}{2} & 1 & \\ -\frac{1}{2} & -2 & 1 \end{pmatrix},\quad D = \begin{pmatrix} 4 & & \\ & 1 & \\ & & 9 \end{pmatrix}$$

求解$Ly=b$及$DL^T x = y$得:

$$y_1 = 10,\quad y_2 = 0,\quad y_3 = 9;$$

$$x_3 = 1,\quad x_2 = 2,\quad x_1 = 2$$

4 追赶法：解三对角方程组

在一些实际问题中，如解常微分方程边值问题，解船体数学放样中建立三次样条函数等中，要解系数矩阵为对角占优的三对角方程组：

$$\begin{bmatrix} b_1 & c_1 & & & \\ a_1 & \ddots & \ddots & & \\ & \ddots & \ddots & c_{n-1} \\ & & a_{n-1} & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}.$$

简记：$Ax = f$

4 追赶法：解三对角方程组

$$\begin{bmatrix} b_1 & c_1 & & & \\ a_1 & \ddots & \ddots & & \\ & \ddots & \ddots & c_{n-1} \\ & & a_{n-1} & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}.$$

我们假定

$$|b_1| > |c_1| > 0, \quad |b_n| > |a_{n-1}| > 0, \tag{2.1}$$

且

$$|b_i| \geq |a_{i-1}| + |c_i|, \quad a_i c_i \neq 0, \quad i = 2, \dots, n-1. \tag{2.2}$$

即 A 是不可约 (行) 弱对角占优的. (可以证明此时 A 是非奇异的)

4 追赶法：解三对角方程组

计算 A 的 LU 分解:

$$A=\begin{bmatrix} b_1 & c_1 & & \\ a_1 & \ddots & \ddots & \\ & \ddots & \ddots & c_{n-1} \\ & & a_{n-1} & b_n \end{bmatrix}=\begin{bmatrix} \alpha_1 & & & \\ a_1 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & a_{n-1} & \alpha_n \end{bmatrix}\begin{bmatrix} 1 & \beta_1 & & \\ & 1 & \ddots & \\ & & \ddots & \beta_{n-1} \\ & & & 1 \end{bmatrix}\triangleq LU. \tag{2.3}$$

由待定系数法, 我们可以得到递推公式:

$$\alpha_1=b_1,\quad \beta_1=c_1/\alpha_1=c_1/b_1,$$

$$\begin{cases} \alpha_i=b_i-a_{i-1}\beta_{i-1}, \\ \beta_i=c_i/\alpha_i=c_i/(b_i-a_{i-1}\beta_{i-1}), \quad i=2,3,\ldots,n-1 \end{cases}$$

$$\alpha_n=b_n-a_{n-1}\beta_{n-1}.$$

4 追赶法：解三对角方程组

求解$Ax=f$等价于求解两个三角方程组$Ly=f$ 与$Ux=y$，先后求出y与x，进而获得求解三对角方程组的追赶法公式：

步1：求解$Ly=f$: $y_1=\frac{f_1}{b_1}, y_i=\frac{(f_i-a_i y_{i-1})}{(b_i-a_i\beta_{i-1})}, (i=2,3,\dots,n-1)$

步2：求解$Ux=y$: $x_n=y_n, x_i=y_i-\beta_i x_{i+1}, (i=n-1,n-2,\dots,2,1)$

- 计算系数$\beta_1\to\beta_2\to\cdots\to\beta_n$及$y_1\to y_2\to\cdots\to y_n$的过程称为“**追**”的过程；
- 计算方程组的解$x_n\to x_{n-1}\to\cdots\to x_1$的过程称为“**赶**”的过程；

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

4 追赶法：解三对角方程组

算法 追赶法 (矩阵分解与方程求解交叉进行)

$\alpha_1 = b_1$

$\beta_1 = c_1 / b_1$

$y_1 = f_1 / b_1$

for $i = 2$ to $n-1$ **do**

  $\alpha_i = b_i - a_{i-1}\beta_{i-1}$

  $\beta_i = c_i / \alpha_i$

  $y_i = (f_i - a_{i-1} y_{i-1}) / \alpha_i$ % 求解 $Ly = f$

end for

$\alpha_n = b_n - a_{n-1}\beta_{n-1}$

$y_n = (f_n - a_{n-1} y_{n-1}) / \alpha_n$

$x_n = y_n$

for $i = n-1$ to 1 **do**

  $x_i = y_i - \beta_i x_{i+1}$

end for

4 追赶法：解三对角方程组

为了使得算法能够顺利进行下去，我们需要证明 $\alpha_i \neq 0$.

定理 设三对角矩阵 A 满足条件 (2.1) 和 (2.2). 则 A 非奇异，且

(1) $0 < |\beta_i| < 1,\ i = 1, 2, \ldots, n-1;$

(2) $0 < |c_i| \leq |b_i| - |a_{i-1}| < |\alpha_i| < |b_i| + |a_{i-1}|, \quad i = 2, 3, \ldots, n;$

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

5 向量和矩阵的范数

对于用计算机求解方程组 $Ax=b$ 假设原始数据无误差和数值运算无舍入误差是不可能的。

通常都是在系数矩阵 A 有误差 δA，或右端项 b 有误差 δb 的情形下，求出方程组 $(A+\delta A)x=(b+\delta b)$ 的近似解 $\hat{x}$

需要讨论方程组本身在 A 和 b 有误差时对解的影响，即方程组的性态。

需要对向量和矩阵的大小进行度量，即引进向量范数和矩阵范数的概念。

5.1 向量的范数

定义： 设向量 $x \in \Re^n$，称对应于 x 且满足下列三个条件的实数为 x 的范数（或模），记为 $\boxed{\|x\|}$：

(1) 当 $x \neq 0$ 时，$\|x\| > 0$；当且仅当 $x = 0$ 时，$\|x\| = 0$ 　　正定

(2) 对任意实数 c 及实向量 x 有 $\|cx\| = |c|\|x\|$ 　　齐次

(3) 对任意实向量 x, y 有

$$\|x+y\| \le \|x\| + \|y\|$$

三角不等式

5.1 向量的范数

对于 $x \in \Re^n$，由向量的分量定义的以下三个非负实数是向量范数：

$$\|x\|_\infty = \max\{|x_1|, |x_2|, \cdots, |x_n|\}$$ （无穷范数）

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$ （1-范数）

$$\|x\|_2 = (x_1^2 + x_2^2 + \cdots + x_n^2)^{\frac{1}{2}}$$ （2-范数）

上面三种范数是向量空间 $\Re^n$ 中的常用范数，因此成为 $\Re^n$ 中的**基本范数**。

向量的 p-范数：$\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}$，其中 $p \in [1, \infty]$.

验证： 实数$\|x\|_2$满足向量范数的三个条件：

证： 当$x \neq 0$时，数$\|x\|_2 > 0$；当且仅当$x = 0$时，数$\|x\|_2 = 0$，即数$\|x\|_2$满足条件（1）。

用实数c乘以向量x有$cx = (cx_1, cx_2, \cdots, cx_n)^T$，

故有

$$\|cx\|_2 = \left[(cx_1)^2 + (cx_2)^2 + \cdots + (cx_n)^2\right]^{\frac{1}{2}} = |c|\|x\|_2,$$

满足条件（2）。

证（续）：设 $x \in \Re^n$，$y \in \Re^n$，由线性代数中的柯西-许瓦兹不等式 $\langle x, y\rangle^2 \le \langle x, x\rangle \cdot \langle y, y\rangle$，

有 $\|x+y\|_2^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle$

符号 $\langle x, y\rangle$ 表示 x 和 y 的内积

$$\le \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2$$

$$= \left(\|x\|_2 + \|y\|_2\right)^2$$

即 $\|x+y\|_2 \le \|x\|_2 + \|y\|_2$ 满足条件（3）。

所以非负实数 $\|x\|_2$ 为对应于向量 x 的一种范数，称为2-范数。

另外两种范数的验证类似。

5.1 向量的范数

例： 计算向量$\boldsymbol{x}=(1,-2,3)^T$的1-、2-和∞范数。

5.1 向量的范数

定义：设$\|x\|_t$，$\|x\|_s$ 是$\Re^n$ 中向量x 的任意两种范数。若存在**正数**m 和M 使得对一切非零向量 x 恒有 $m\|x\|_s \le \|x\|_t \le M\|x\|_s$ 则称$\|x\|_t$ 与 $\|x\|_s$ **等价**。

三种基本向量范数满足以下关系式：

$$\boxed{\begin{aligned}&\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2,\\ &\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty,\\ &\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty\end{aligned}}$$

求证： $\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2$

证： 设 $x=(x_1,x_2,\cdots,x_n)^T$，$y=(1,1,\cdots,1)^T$，

$z=(|x_1|,|x_2|,\cdots,|x_n|)^T$，

由柯西-许瓦兹不等式，有 $\langle y,z\rangle \le \langle y,y\rangle^{\frac{1}{2}}\cdot\langle z,z\rangle^{\frac{1}{2}}$，

由内积的定义，有 $1\cdot|x_1|+1\cdot|x_2|+\cdots 1\cdot|x_n| \le$

$$\Big(\underbrace{1+1+\cdots+1}_{n\text{个}1}\Big)^{\frac{1}{2}}\cdot\left(|x_1|^2+|x_2|^2+\cdots|x_n|^2\right)^{\frac{1}{2}},$$

即 $\|x\|_1 \le \sqrt{n}\|x\|_2$

证（续）：另外，由于

$$\left(\sum_{i=1}^{n}|x_i|\right)^2 \geq \sum_{i=1}^{n}|x_i|^2$$

所以又有 $\|x\|_2 \leq \|x\|_1$

由定义 可知，右侧关系式表明$\Re^n$中的三种基本范数是等价的。

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2,$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty,$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

向量空间 $\Re^n$ 中的任意两种范数都是等价的。

5.2 矩阵的范数

定义 设 A 为 n 阶方阵，$\vec{x} \in \Re^n$，由

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

所定义的实数称为 A 的范数。

由 上述 定义的矩阵范数，是向量空间 $\Re^n$ 中两个向量范数 $\|Ax\|$ 与 $\|x\|$ 之比，当 $\vec{x} \in \Re^n$ 时取最大值所得到的实数。因此也称为由向量范数导出的矩阵范数，且 上述定义 可等价地表示为 $\|A\| = \max\limits_{\|\vec{x}\|=1} \|Ax\|$

5.2 矩阵的范数

由矩阵范数的定义，可得到n阶方阵A的范数具有下列性质：

性质1：当$A \neq O$时，$\|A\| > 0$；当且仅当$A = O$时，$\|A\| = 0$

性质2：设c为实数，则$\|cA\| = |c|\|A\|$

性质3：对于$A, B \in \Re^{n\times n}$，有$\|A+B\| \leq \|A\| + \|B\|$

性质4：对于$x \in \Re^{n}$，有$\|Ax\| \leq \|A\|\|x\|$

性质5：对于$A, B \in \Re^{n\times n}$，有$\|AB\| \leq \|A\|\|B\|$

下面仅证性质4和性质5

证：由矩阵范数定义 有

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

移除作乘得 $\|A\|\,\|x\| \geq \|Ax\|$，证明了性质4。

由矩阵范数定义 并利用性质4，有

$$\|AB\| = \max_{x \neq 0} \frac{\|ABx\|}{\|x\|} \leq \max_{x \neq 0} \frac{\|A\|\,\|Bx\|}{\|x\|} = \|A\|\,\|B\|$$

证明了性质5。

5.2 矩阵的范数

◆ 由**矩阵范数的定义**可以看到，用矩阵范数的定义求矩阵的范数几乎是不可能的。那么如何求矩阵的范数？

◆ 下面**定理**给出了用矩阵的元素直接计算矩阵范数的方法。

定理 设 $\vec{x} \in \Re^n$，A 为 n 阶方阵，则有

(1) $\|A\|_\infty = \max\limits_{1 \le i \le n} \sum\limits_{j=1}^{n} |a_{ij}|$ 行范数

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

(3) $\|A\|_2 = \left(\text{方阵} A^T A \text{的最大特征值}\right)^{\frac{1}{2}}$ 谱范数

定义 设 λ_i 是方阵 B 的任一特征值，实数

$$\rho(B)=\max_{1\le i\le n}|\lambda_i|$$

称为方阵 B 的谱半径。

由定义，有 $\|A\|_2=\sqrt{\rho(A^TA)}$，既然 $\|A\|_2$ 为 A^TA 的谱半径的开方，通常称 $\|A\|_2$ 为矩阵 A 的谱范数。

又由于 $\|A\|_\infty(\|A\|_1)$ 等于 A 的每一行（列）元素的绝对值之和中的最大者，故称 $\|A\|_\infty$ 为矩阵 A 的行范数，$\|A\|_1$ 称为矩阵 A 的列范数。

定理： 如果$\|B\|<1$，则$I\pm B$为非奇异矩阵，且$\|(I\pm B)^{-1}\|\le\dfrac{1}{1-\|B\|}$.

证明：（反证法）

若$\det(\boldsymbol{I}-\boldsymbol{B})=0$, 则$(\boldsymbol{I}-\boldsymbol{B})\boldsymbol{x}=0$有非零解，即存在$\boldsymbol{x}_0\neq 0$使$\boldsymbol{B}\boldsymbol{x}_0=\boldsymbol{x}_0$。

即$\dfrac{\|\boldsymbol{B}\boldsymbol{x}_0\|}{\|\boldsymbol{x}_0\|}$=1。又因为$\|\boldsymbol{B}\boldsymbol{x}_0\|\le\|\boldsymbol{B}\|\|\boldsymbol{x}_0\|$，因此，$\|\boldsymbol{B}\|\ge \mathbf{1}$，与假设矛盾。

又因为：$(\boldsymbol{I}\pm\boldsymbol{B})(\boldsymbol{I}\pm\boldsymbol{B})^{-1}=\boldsymbol{I}$，展开化简可得，

$$(\boldsymbol{I}\pm\boldsymbol{B})^{-1}=\boldsymbol{I}\mp\boldsymbol{B}(\boldsymbol{I}\pm\boldsymbol{B})^{-1}。$$

进而，$\|(\boldsymbol{I}\pm\boldsymbol{B})^{-1}\|\le\|\boldsymbol{I}\|+\|\boldsymbol{B}\|\|(\boldsymbol{I}\pm\boldsymbol{B})^{-1}\|$，

化简可得，$\|(\boldsymbol{I}\pm\boldsymbol{B})^{-1}\|\le\dfrac{1}{1-\|\boldsymbol{B}\|}$. 证毕。

例　设$A=\begin{pmatrix}1 & -2\\ -3 & 4\end{pmatrix}$，求 A 的三种范数。

解：$\|A\|_\infty=\max\{1+2,3+4\}=7$，以及

$$\|A\|_1=\max\{1+3,2+4\}=6$$

由于$A^TA=\begin{pmatrix}10 & -14\\ -14 & 20\end{pmatrix}$，故有

$$\left|\lambda E-A^TA\right|=\begin{vmatrix}\lambda-10 & 14\\ 14 & \lambda-20\end{vmatrix}=\lambda^2-30\lambda+4=0$$

$$\lambda_{1,2}=15\pm\sqrt{221},\quad \rho\left(A^TA\right)=15+\sqrt{221},$$

所以

$$\|A\|_2=\sqrt{15+\sqrt{221}}\approx 5.46$$

6 线性方程组的性态

◆ 在实际应用中，所给的数据可能是通过实验、测量或观察得来的，因此通常会带有一定的误差，这些误差不可避免地会对问题的解产生影响.

◆ 定义：若矩阵A或常数项b的微小变化，引起方程组$Ax=b$解的巨大变化，则称方程组为**病态方程组**，矩阵A为**病态矩阵**；否则称方程组为**良态方程组**，矩阵A为**良态矩阵**。

6 线性方程组的性态

例 考虑线性方程组 $Ax=b$，其中 $A=\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix}$，$b=\begin{bmatrix}2\\2\end{bmatrix}$.

$\Longrightarrow$ 解为 $x=\begin{bmatrix}2\\0\end{bmatrix}$.

若 b 的第二个元素出现小扰动，变为 $b=\begin{bmatrix}2\\2.0001\end{bmatrix}$，则解变为 $x=\begin{bmatrix}1\\1\end{bmatrix}$.

当右端项出现细微变化时，解会出现很大的变化，因此该线性方程组是病态的.

◆ **怎样判断一个线性方程组是否病态？**

1. 对于线性方程组而言，问题是否病态主要取决于**系数矩阵是否病态**.
2. 判断一个矩阵是否病态的一个重要指标是**矩阵条件数**.

◆矩阵条件数（A 精确，$\mathbf{b}$ 有微小变化$\Delta\mathbf{b}$）

由$A(x+\Delta x)=b+\Delta b$ 和$Ax=b$可得，$A\Delta x=\Delta b$，

即：$\Delta x=A^{-1}\Delta b$，进而，$\|\Delta x\|=\|A^{-1}\Delta b\|\le\|A^{-1}\|\|\Delta b\|$

又因为：$Ax=b$，进而，$\|b\|=\|Ax\|\le\|A\|\|x\|$,

假设$\|b\|\ne 0$，则$\frac{1}{\|x\|}\le\frac{\|A\|}{\|b\|}$。

因此，$\underline{\frac{\|\Delta x\|}{\|x\|}}\le\frac{\|A\|\|A^{-1}\|\|\Delta b\|}{\|b\|}=\|A\|\|A^{-1}\|\boxed{\frac{\|\Delta b\|}{\|b\|}}$.

6 线性方程组的性态

在范数定义的基础上，考虑线性方程组 $Ax = b$，解是 x^*

在 A 是精确的情况下，若**对 b 的观测有微小变化Δb**

此时方程组 $A(x+\Delta x) = b+\Delta b$,解为 $x^*+\Delta x^*$，则

$$\frac{\|\Delta x^*\|}{\|x^*\|} \le \|A^{-1}\|\|\mathrm{A}\|\frac{\|\Delta b\|}{\|b\|}$$

这里向量和矩阵的范数统一。

＊ 实际上是给出方程组 $Ax = b$ **解的相对误差的上界**。

◆矩阵条件数（A有微小变化ΔA，b精确）

$(A+\Delta A)(x+\Delta x)=b$，解为 $x^*+\Delta x^*$。展开，考虑$Ax=b$，化简得，$(A+\Delta A)\Delta x=-(\Delta A)x$。若$\Delta A$无限制，则$A+\Delta A$可能奇异。由于$A+\Delta A=A(I+A^{-1}\Delta A)$，联想定理可知，若$\|A^{-1}\Delta A\|<1$，则$I+A^{-1}\Delta A$为非奇异矩阵，$(I+A^{-1}\Delta A)^{-1}$存在。

因此，$\Delta x=-(A(I+A^{-1}\Delta A))^{-1}(\Delta A)x$

$$=-(I+A^{-1}\Delta A))^{-1}A^{-1}(\Delta A)x$$，即$\|\Delta x\|\le\dfrac{\|A^{-1}\|\|\Delta A\|\|x\|}{1-\|A^{-1}\Delta A\|}$。

假设$\|A^{-1}\|\|\Delta A\|<1$，则，$\|\Delta x\|\le\dfrac{\|A^{-1}\|\|\Delta A\|\|x\|}{1-\|A^{-1}\|\|\Delta A\|}=\dfrac{\|A^{-1}\|\|A\|\frac{\|\Delta A\|}{\|A\|}\|x\|}{1-\|A^{-1}\|\|A\|\frac{\|\Delta A\|}{\|A\|}}$，

移项可得，$\dfrac{\|\Delta x^*\|}{\|x^*\|}\le\dfrac{\|A^{-1}\|\|A\|\boxed{\frac{\|\Delta A\|}{\|A\|}}}{1-\|A^{-1}\|\|A\|\frac{\|\Delta A\|}{\|A\|}}$。

定理：如果$\|B\|<1$，则$I\pm B$为非奇异矩阵，且$\|(I\pm B)^{-1}\|<\dfrac{1}{1-\|B\|}$.

◆矩阵条件数

在 b 是精确的情况下，若**对 A 的观测有微小变化ΔA**

此时方程组 $(A+\Delta A)(x+\Delta x)=b$，解为 $x^*+\Delta x^*$，若

$\|A^{-1}\|\|\Delta A\|<1$，则

$$\frac{\|\Delta x^*\|}{\|x^*\|}\le\frac{\|A^{-1}\|\|A\|\frac{\|\Delta A\|}{\|A\|}}{1-\|A^{-1}\|\|A\|\frac{\|\Delta A\|}{\|A\|}}$$

这里向量和矩阵的范数统一。

* 实际上是给出方程组 $Ax=b$ **解的相对误差的上界**。

◆矩阵条件数

- 设 A 为非奇异矩阵，则称

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

为A的条件数，其中$\|\cdot\|_v$是1-范数，2-范数或∞-范数。

- 一般来说，当 A 的条件数较大时，A 就是病态的。
- 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比较精确的解。

◆矩阵条件数性质

性质1： Cond (A) ≥1

性质2： 对任意常数$c \neq 0$，有Cond (cA)= Cond (A)

性质3： A的谱条件数

$$Cond(A)_2 = \|A\|_2\,\|A^{-1}\|_2 = \sqrt{\frac{\lambda_{max}(A^TA)}{\lambda_{min}(AA^T)}}$$

*　　若A为对称矩阵，则

$Cond(A)_2 = \dfrac{\lambda_{max}}{\lambda_{min}}$, λ_{max}和λ_{min}分别为A的绝对值最大和绝对值最小的特征值。

6 线性方程组的性态

例 考虑线性方程组 $Ax=b$, 其中 $A=\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix}$, $b=\begin{bmatrix}2\\ 2\end{bmatrix}$.

◆ 计算系数矩阵A的条件数。

$$A=\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix},\ A^{-1}=\frac{1}{1.0001-1}\begin{bmatrix}1.0001 & -1\\ -1 & 1\end{bmatrix}$$

则： $Cond(A)_\infty=\|A\|_\infty\,\|A^{-1}\|_\infty$

=2. 0001*10000*2. 0001

=40004. 0001

◆ 由此可知，上述方程组是病态的。这就是为什么当右端向量的分量有微小波动时，会引起解有大误差的原因。

＊ 此类方程组不能用本章介绍的直接法求解。

◆ 总结：

介绍了计算机上求解线性代数方程组的几种常用的直接法。

1. **顺序消去法**：解线性方程组直接法的基础。
2. **列主元消去法**：数值运算较稳定的常用方法。因为在消元过程中引进了选主元的技巧，减少了误差对计算结果的影响。
3. **杜立特尔分解法**：可以直接计算$A = LU$分解中L和U的元素值。
4. 平方根法：**仅适用于**系数矩阵为对称正定的情形。
5. 追赶法：**仅适用于**系数矩阵为对角占优的三对角方程组。
6. 引进了**向量范数和矩阵范数**的概念。

分析了方程组的初始数据在用直接法求解时，**误差对方程组解的影响**。

介绍了**条件数的概念**以及**条件数在判断方程组性态时的应用**。

- Q & A
- 谢谢