线性代数方程组的直接解法

数值计算方法课程组

中国石油大学-华东

（课堂专属 切勿网传）

历史发展

问题

求解线性方程组 $AX = b, \quad A \in R^{n\times n}, det(A) \neq 0.$

线性方程组的求解是一个古老的数学问题.

- 东汉时期的《九章算术》中介绍了消元法.
- 直到 19 世纪，西方才有了 Gauss 消去法.
- 20 世纪中叶，计算机的问世，大型线性方程组求解成为可能.

其数值方法可分两种：直接法和迭代法。

直接解法：在没有舍入误差的情况下，经过有限次运算可以得到方程组的精确解的方法。

迭代法：从解的某个近似值出发，通过构造一个无穷序列去逼近精确解的方法 (一般有限步内得不到精确解)。

Cramer 法则:

$$x_i = \frac{|A_i|}{|A|}, \quad i = 1, 2, \cdots, n.$$

- 所需乘除的运算量大约为 $(n+1)!+n$: 计算 $(n+1)$ 个 n 阶矩阵 $A, A_1, \ldots, A_n$ 行列式的 $(n+1)\times n!$ 乘法运算, 以及求解 X 的 n 个除法运算.
- 若 $n=26$, 目前最快的超算机日本的富岳（442 千万亿次/秒）, 要算大概 782 年
- Matlab 计算命令:

```
factorial(27)/4.42e+17/60/60/24/365=781.185(年)
```

高斯消去法是目前求解中小规模线性方程组的最常用方法.

目 录

目 录

§3.1.1 三角形方程组的解法

考虑如下 n 阶下三角形方程组:

$$Ly = b$$

这里 L 非奇异且 $l_{ii} \neq 0$,

$$\begin{aligned} b &= [b_1, b_2, \ldots, b_n]^T \in R^n, \\ y &= [y_1, y_2, \ldots, y_n]^T \in R^n, \end{aligned}$$

$$L = \begin{bmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & \cdots & 0 \\ l_{31} & l_{32} & l_{33} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{nn} \end{bmatrix}, \quad l_{ii} \neq 0, i = 1, 2, \cdots, n.$$

显然

$$
\begin{aligned}
y_1 &= \frac{b_1}{L_{11}}, \\
y_i &= \frac{1}{l_{ii}}(b_i - \sum_{j=1}^{i-1} l_{ij} y_j), \\
i &= 2, 3, \cdots, n.
\end{aligned}
$$

注

- 此方法称为前代法.
- 实际计算中可将得到的 y_i 存到 b_i 所在的存储单元以节省存储空间.

前代法算法

```
for j=1:n-1
    b(j)=b(j)/l(j,j)
    b(j+1:n)=b(j+1:n)-b(j)L(j+1:n,j)
end
b(n)=b(n)/L(n,n)
```

注

该算法所需加、减、乘、除运算次数为

$$\sum_{j=1}^{n-1}[1+(n-j)+(n-j)]+1=\sum_{j=1}^{n}(2j-1)+1=n^2+1.$$

对于 n 阶上三角形方程组

$$Uy = b$$

这里 U 非奇异且 $u_{ii} \neq 0, u_{ij} = 0(i > j)$. 显然

$$\begin{aligned} y_n &= \frac{b_n}{u_{nn}}, \\ y_i &= \frac{1}{u_{ii}}(b_i - \sum_{j=i+1}^{n} u_{ij} y_j), \\ i &= n-1, n-2, \cdots, 1. \end{aligned}$$

此方法称为回代法, 所需加减乘除运算次数同前代法一样均为 $n^2 + 1$.

回代法算法

```
for j=n:-1:2
    b(j)=b(j)/u(j,j)
    b(1:j-1)=b(1:j-1)-b(j)u(1:j-1,j)
end
b(1)=b(1)/u(1,1)
```

对于一般的线性方程组

$$Ax = b, \quad A \in R^{n\times n}, \quad b \in R^n.$$

若 $A = LU$,L, U 分别为下三角阵和上三角阵，则原方程的解 x 可由两步得到

$$\begin{aligned} (1) \quad Ly &= b \\ (2) \quad Ux &= y \end{aligned}$$

记

$$
\begin{aligned}
l_k &= [0, 0, \cdots, l_{k+1,k}, l_{k+2,k}, \cdots, l_{n,k}]^T, k = 1, 2, \cdots, n-1, \\
L_k &= I - l_k e_k^T \\
&= \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{n,k} & & & 1 \end{bmatrix}
\end{aligned}
$$

称此类型的初等下三角矩阵为 Gauss 变换，而向量 l_k 为 Gauss 向量.

Gauss 变换的举例

求 Gauss 变换矩阵 $\mathbf{L}$，使得 $\mathbf{L}x = \widetilde{x}$
(1)

$$x = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \ \widetilde{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$x, \widetilde{x}$ 第一个元素相等，则

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ -l_{21} & 1 & 0 \\ -l_{31} & 0 & 1 \end{bmatrix},$$

且 $l_{21} = \frac{x_2}{x_1} = 2, l_{31} = \frac{x_3}{x_1} = 5.$

Gauss 变换的举例

(2)

$$x = \begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix}, \ \widetilde{x} = \begin{bmatrix} 3 \\ 5 \\ 0 \end{bmatrix}$$

$x, \widetilde{x}$ 前 2 个元素相等，则

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -l_{32} & 1 \end{bmatrix},$$

且 $l_{32} = \frac{x_3}{x_2} = -0.4.$

Gauss 变换的性质

性质 1

设向量 $x=(x_1,x_2,\cdots,x_n)^T$, $x_k\neq 0$, 则存在唯一的下三角阵 $L_k=I-l_ke_k^T$, 满足

$$L_kx=(x_1,x_2,\cdots,x_k,0,\cdots,0)^T.$$

证明：记 $y=(x_1,x_2,\cdots,x_k,0,\cdots,0)^T$. 由 Gauss 变换的定义可知

$$L_kx=(x_1,x_2,\cdots,x_k,x_{k+1}-x_kl_{k+1,k},\cdots,x_n-x_kl_{n,k})^T.$$

从而可知

$$x_j-x_kl_{j,k}=0,\Rightarrow l_{j,k}=\frac{x_j}{x_k},j=k+1,k+2,\cdots,n.$$

性质 2

$$L_k^{-1} = I + l_k e_k^T = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & l_{n,k} & & & 1 \end{bmatrix}.$$

且

$$L_k \cdot L_j = I - l_k e_k^T - l_j e_j^T, \quad (k < j).$$

性质 3

若记 $L_k^{-1} = I + l_k e_k^T$，则有

$$L_1^{-1} L_2^{-1} \cdots L_{n-1}^{-1} = L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix}$$

证明:

$$\begin{aligned} L &= L_1^{-1} L_2^{-1} \cdots L_{n-1}^{-1} \\ &= (I + l_1 e_1^T)(I + l_2 e_2^T) \cdots (I + l_{n-1} e_{n-1}^T) \\ &= I + l_1 e_1^T + \cdots + l_{n-1} e_{n-1}^T. \end{aligned}$$

即，一个单位下三角阵可以分解成一系列初等下三角矩阵的乘积。

§3.1.3 Gauss 消去法举例

高斯消去法求线性方程组:

$$\begin{bmatrix} 12 & -3 & 3 \\ -18 & 3 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 15 \\ -15 \\ 6 \end{bmatrix}$$

解: $[A, b] = \begin{bmatrix} 12 & -3 & 3 & 15 \\ -18 & 3 & -1 & -15 \\ 1 & 1 & 1 & 6 \end{bmatrix}$

$$\rightarrow \begin{bmatrix} 12 & -3 & 3 & 15 \\ 0 & -\frac{3}{2} & \frac{7}{2} & \frac{15}{2} \\ 0 & \frac{5}{4} & \frac{3}{4} & \frac{19}{4} \end{bmatrix} \rightarrow \begin{bmatrix} 12 & -3 & 3 & 15 \\ 0 & -\frac{3}{2} & \frac{7}{2} & \frac{15}{2} \\ 0 & 0 & \frac{11}{3} & 11 \end{bmatrix}$$ 上三角形方程组

从而 $x_3 = 3, x_2 = 2, x_1 = 1.$

上述高斯消去法求线性方程组的初等行变换可以用 Gauss 变换描述，如：

$$A^{(0)} = \begin{bmatrix} 12 & -3 & 3 & 15 \\ -18 & 3 & -1 & -15 \\ 1 & 1 & 1 & 6 \end{bmatrix} \to A^{(1)} = \begin{bmatrix} 12 & -3 & 3 & 15 \\ 0 & -\frac{3}{2} & \frac{7}{2} & \frac{15}{2} \\ 0 & \frac{5}{4} & \frac{3}{4} & \frac{19}{4} \end{bmatrix}$$

表示为

$$L_1 A^{(0)} = A^{(1)}, \text{其中} L_1 = \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ -\frac{1}{12} & 0 & 1 \end{bmatrix}$$

$$A^{(1)} = \begin{bmatrix} 12 & -3 & 3 & 15 \\ 0 & -\frac{3}{2} & \frac{7}{2} & \frac{15}{2} \\ 0 & \frac{5}{4} & \frac{3}{4} & \frac{19}{4} \end{bmatrix} \quad \to A^{(2)} = \begin{bmatrix} 12 & -3 & 3 & 15 \\ 0 & -\frac{3}{2} & \frac{7}{2} & \frac{15}{2} \\ 0 & 0 & \frac{11}{3} & 11 \end{bmatrix}$$

表示为

$$L_2 A^{(1)} = A^{(2)}, \text{其中} L_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{5}{6} & 1 \end{bmatrix}$$

综上:

$$L_2 L_1 A^{(0)} = A^{(2)},$$

上述例子，如果只看系数矩阵 A 的变换，则

$$L_2L_1A = U, \text{其中上三角形矩阵} U = \begin{bmatrix} 12 & -3 & 3 \\ 0 & -\frac{3}{2} & \frac{7}{2} \\ 0 & 0 & \frac{11}{3} \end{bmatrix}$$

从而

$$A = (L_2L_1)^{-1}U = L_1^{-1}L_2^{-1}U := LU$$

其中

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{2} & 1 & 0 \\ \frac{1}{12} & -\frac{5}{6} & 1 \end{bmatrix}$$

矩阵的三角分解

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 10 \end{bmatrix}$$

解：$L_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$, $\qquad L_1 A = \begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{bmatrix}$

$L_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$, $\qquad L2L_1 A = \begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{bmatrix}$ 记为 U

矩阵的三角分解

$$L = (L_2 L_1)^{-1} = L_1^{-1} L_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}$$

从而

$$A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{bmatrix}$$

矩阵的三角分解

上述过程可简写为

$$
\begin{aligned}
A &= \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 10 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

三角分解

在计算机程序中，矩阵 A 的三角分解过程中 $A^{(k)}$ 与 L_k 可共用一个数组，所以三角分解过程可采用节省空间的写法：

$$A=\begin{bmatrix}1&4&7\\2&5&8\\3&6&10\end{bmatrix}\to\begin{bmatrix}1&4&7\\\color{red}{2}&-3&-6\\\color{red}{3}&-6&-11\end{bmatrix}\to\begin{bmatrix}1&4&7\\\color{red}{2}&-3&-6\\\color{red}{3}&\color{red}{2}&1\end{bmatrix}$$

所以 A 的三角分解式为

$$A=\begin{bmatrix}1&0&0\\2&1&0\\3&2&1\end{bmatrix}\begin{bmatrix}1&4&7\\0&-3&-6\\0&0&1\end{bmatrix}.$$

§3.1.3 三角分解法

一般地，设给定 n 阶矩阵 $A=(a_{ij})\in R^{n\times n}$, 记

$$A^{(0)}=(a_{ij}^{(0)})=(a_{ij})=A$$

取

$$\begin{aligned} L_1 &= I-l_1e_1^T, \\ l_1 &= (0,l_{21},l_{31},\cdots,,l_{n1})^T, \\ l_{i1} &= \frac{a_{i,1}^{(0)}}{a_{11}^{(0)}}, \quad i=2,3,\cdots,n. \end{aligned}$$

记

$$A^{(1)} = L_1 A^{(0)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix} = \begin{bmatrix} A_{11}^{(1)} & A_{12}^{(1)} \\ 0 & A_{22}^{(1)} \end{bmatrix}$$

这里

$$a_{ij}^{(1)} = a_{ij}^{(0)} - \frac{a_{i1}^{(0)} a_{1j}^{(0)}}{a_{11}^{(0)}}, \quad i, j = 2, 3, \cdots, n.$$

记

$$A^{(2)} = L_2 A^{(1)} = \begin{bmatrix} A_{11}^{(2)} & A_{12}^{(2)} \\ 0 & A_{22}^{(2)} \end{bmatrix}$$

这里 $A_{11}^{(2)}$ 为 2×2 的上三角阵，$A_{22}^{(2)}$ 为 $(n-2)\times(n-2)$ 的矩阵

$$
\begin{aligned}
L_2 &= I - l_2 e_2^T, \\
l_2 &= (0, 0, l_{32}, \cdots, l_{n2})^T, \\
l_{i2} &= \frac{a_{i,2}^{(1)}}{a_{22}^{(1)}}, \quad i = 3, \cdots, n. \\
a_{ij}^{(2)} &= a_{ij}^{(1)} - \frac{a_{i2}^{(1)} a_{2j}^{(1)}}{a_{22}^{(1)}}, \quad i, j = 3, 4, \cdots, n.
\end{aligned}
$$

依此类推，可得第 k 步消元过程的计算公式:

$$
A^{(k)} = L_k A^{(k-1)} = (a_{ij}^{(k)}), k = 1, 2, \cdots, n-1
$$

这里

$$
\begin{aligned}
L_k &= I - l_k e_k^T, \\
l_k &= (0, 0, \cdots, l_{(k+1)k}, \cdots, l_{nk})^T \\
l_{ik} &= a_{ik}^{(k-1)} / a_{kk}^{(k-1)}, \quad i = k+1, k+2, \cdots, n
\end{aligned}
$$

$$
a_{ij}^{(k)} = \begin{cases} a_{ij}^{(k-1)} & i = 1, 2, \cdots, k, \quad j = k+1, k+2, \cdots, n \\ a_{ij}^{(k-1)} - l_{ik} a_{kj}^{(k-1)} & i = k+1, k+2, \cdots, n, \quad j = k+1, k+2, \cdots, n \\ 0 & i = k+1, \cdots, n, \quad j = 1, 2, \cdots, k. \end{cases}
$$

从而可知

$$
A^{(n-1)} = L_{n-1} L_{n-2} \cdots L_1 A = U(\text{上三角阵})
$$

所以，有

$$A = L_1^{-1} L_2^{-1} \cdots L_{n-1}^{-1} U = LU,$$

其中 L 单位下三角阵，U 上三角阵：

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 & 0 \\ \vdots & \vdots & & \ddots & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix}, \quad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

经过 $n-1$ 次消元，并将 l_{ik} 存放在矩阵零元素位置.

三角分解的计算过程：先计算 U 的行，在计算 L 的列

$$\begin{array}{cccccc}
u_{11} & u_{12} & u_{13} & \cdots & u_{1n} & \text{第1步} \\
l_{21} & u_{22} & u_{23} & \cdots & u_{2n} & \text{第3步} \\
l_{31} & l_{32} & u_{33} & \cdots & u_{3n} & \text{第5步} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
l_{n1} & l_{n2} & l_{n3} & \cdots & u_{nn} & \text{第}2n-1\text{步} \\
\text{第2步} & \text{第4步} & \text{第6步} & \cdots & 0 & \text{第}2n-2\text{步}
\end{array}$$

注

对方程组求解，只要得到了系数矩阵的三角分解形式，再利用前代算法和回代算法解两个三角方程组即得.

矩阵的三角分解法算法

```
for k=1:n-1
    A(k+1:n,k)=A(k+1:n,k)/A(k,k)
    A(k+1:n,k+1:n)=A(k+1:n,k+1:n)-A(k+1:n,k)A(k,k+1:n)
end
L=tril(A, -1)+eye(n);    U=triu(A);
```

三角分解法工作量

$$\sum_{k=1}^{n-1}[(n-k)+2(n-k)^2]=\frac{2}{3}n^3+O(n^2)$$

三角分解法求解线性方程组 $Ax = b$

矩阵 A 的三角分解法算法 $\rightarrow$ 下三角形矩阵 L 和上三角形矩阵 U;

下三角形方程组算法 $Ly = b \rightarrow$ 解 y;

上三角形方程组算法 $Ux = y \rightarrow$ 解 x.

高斯消去法求解线性方程组 $Ax = b$

对增广矩阵 $a = [A, b]$ 进行三角分解（高斯消去）$\rightarrow$ 得到上三角形增广矩阵 $\widetilde{U} = [U, y]$;

上三角形线性方程组算法 $Ux = y \rightarrow$ 解 x.

两种方法等价

问题与思考

三角分解过程中 $a_{kk}^{(k-1)}$ 称为主元素. 显然 $a_{kk}^{(k-1)}(k=1,2,\cdots,n-1)$ 当且仅当都不为零时三角分解才能进行到底.
给定的矩阵 A 满足什么条件，才能保证 $a_{kk}^{(k-1)}(k=1,2,\cdots,n-1)$ 都不为零?

定理 (高斯消去法的实现条件)

主元素 $a_{ii}^{(i-1)}(i=1,,2,\cdots,n)$ 全不为零的充要条件是 A 的各顺序主子式都不等于零, 即

$$\Delta_i = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{21} & a_{22} & \cdots & a_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & l_{i2} & \cdots & a_{ii} \end{vmatrix} \neq 0, \quad i=1,2,\cdots k(\leq n).$$

证明：对 k 用归纳法证明.$k=1$ 时，$A_1=a_{11}^{(1)}$, 自然成立. 假定定理直到 $k-1$ 成立，下面只需要证明 $\Delta_1,\Delta_2,\cdots,\Delta_{k-1}$ 非零时，Δ_k 非零的充要条件是 $a_{kk}^{(k-1)}\neq 0$ 即可.

在归纳假设下，Gauss 消去法可进行到第 $k-1$ 步

$$A^{(k-1)} = L_{k-1}L_{k-2}\cdots L_1 A = \begin{bmatrix} A_{11}^{(k-1)} & A_{12}^{(k-1)} \\ 0 & A_{22}^{(k-1)} \end{bmatrix}$$

$A_{11}^{(k)}$ 为主对角元是 $a_{11}^{(1)}, a_{22}^{(2)}, \cdots, a_{k-1,k-1}^{(k-1)}$ 的上三角矩阵. $A^{(k-1)}$ 的 k 阶顺序主子阵 $A_k^{(k-1)}$ 是上三角的，

$$A_{k-1}^{(k-1)} = [L_{k-1}]_k[L_{k-2}]_k\cdots[L_1]_k A_k = \begin{bmatrix} A_{11}^{(k-1)} & \\ 0 & a_{kk}^{(k-1)} \end{bmatrix}$$

故

$$\Delta_k^{(k-1)} = a_{11}^{(1)}a_{22}^{(2)}\cdots a_{k,k}^{(k-1)} \neq 0 \Leftrightarrow a_{k,k}^{(k-1)} \neq 0.$$

因 $L_j, j=1,2,\cdots,k-1$ 均为单位下三角矩阵，其行列式

$$|L_j| = 1.$$

从而可知

$$\Delta_k \neq 0 \Leftrightarrow a_{k,k}^{(k-1)} \neq 0.$$

定理 (矩阵三角分解的充分条件)

若 $A \in R^{n\times n}$ 的顺序主子阵 $A_k \in R^{k\times k}(k-1,2,\cdots,n)$ 均非奇异，则存在唯一的单位下三角阵 $L \in R^{n\times n}$ 和上三角阵 $U \in R^{n\times n}$ 使得 $A = LU$.

§3.1.4 其他三角分解

定义

- 如果矩阵 $A \in R^{n\times n}$ 可以分解为一个单位下三角阵 $L \in R^{n\times n}$ 和一个上三角阵 $U \in R^{n\times n}$ 的乘积，即 $A = LU$，则称此分解为 *Doolittle 分解*;
- 如果矩阵 $A \in R^{n\times n}$ 可以分解一个下三角阵 $L \in R^{n\times n}$ 和单位上三角阵 $U \in R^{n\times n}$ 的乘积, 则称此分解为 *Crout 分解*.

定义

给定矩阵 $A \in R^{n\times n}$，如果满足:

$$a_{ij} = \begin{cases} 0, & i-j > p(i > j), \\ 0, & j-i > q(j > i). \end{cases}$$

则称 A 为上半带宽为 q，下半带宽为 p 的带状矩阵，$AX = b$ 称为带状方程组；如果 $p = q = t$，则称 t 为 A 的半带宽，$AX = b$ 为等带宽方程组 $2t+1$ 为 A 的总带宽.

定理 (保带状结构定理)

设 $A \in R^{n\times n}$ 为上半带宽为 q，下半带宽为 p 的带状矩阵，且顺序主子式非零，则 A 有唯一的三角分解 $A = LU$, 其中 L 是下半带宽为 p 单位下三角阵和 U 是上半带宽为 q 的上三角阵.

- 保带状结构定理说明：矩阵的三角分解中，L 和 U 带外元素为零，因此不必计算，且不必参加求和运算.

若 $t=1$, 则 $AX=d$ 为三对角方程组,

$$A=\begin{bmatrix} b_1 & c_1 & 0 & 0 & \cdots & 0 \\ a_2 & b_2 & c_2 & 0 & \cdots & 0 \\ 0 & a_3 & b_3 & c_3 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & 0 & \cdots & 0 & a_n & b_n \end{bmatrix}, \quad d=\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix}.$$

根据保带状结构定理，系数矩阵可作如下三角分解:

$$A=LU=\begin{bmatrix} 1 & 0 & & 0 \\ l_2 & 1 & & \\ & \ddots & \ddots & \\ 0 & & l_n & 1 \end{bmatrix} \cdot \begin{bmatrix} u_1 & c_1 & & 0 \\ 0 & u_2 & \ddots & \\ & & \ddots & c_{n-1} \\ 0 & & 0 & u_n \end{bmatrix}$$

三对角方程组求解的计算公式:

- LU 分解:

$$u_1 = b_1, l_i = \frac{a_i}{u_{i-1}}, u_i = b_i - l_i c_{i-1}, i = 2, 3, \cdots, n.$$

- 解方程组 $Ly = d$(追的过程)

$$y_1 = d_1, y_i = d_i - l_i y_{i-1}, i = 2, \cdots, n.$$

- 解方程组 $Ux = y$(赶的过程)

$$x_n = \frac{y_n}{u_n}, x_i = \frac{y_i - c_i x_{i+1}}{u_i}, i = n-1, \cdots, 1.$$

显然其运算工作量为 $O(n)$，为一种具有最优运算量的算法.

目 录

选主元三角分解的思想

三角分解过程中存在的问题

- Gauss 消元法完成的条件是矩阵的各阶顺序主子式均不为零.
- 三角分解过程中的除法运算要求分母不能太小, 否则将可能产生不稳定情况.

选主元的目的就是为了完成消元且避免不稳定情况的发生.

例:

在 8 位制计算机上解方程组

$$\begin{aligned} 10^{-9}x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2 \end{aligned}$$

要求用三角分解方法计算.

解：注意到，小主元可能导致计算失败。

$$l_{21} = a_{21}/a_{11} = 10^9, a_{22} = 1 - l_{21} \times 1 = 0.00000001 \times 10^9 - 10^9 = -10^9$$

$$L = \begin{bmatrix} 1 & 0 \\ 10^9 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 10^{(-9)} & 1 \\ 0 & -10^9 \end{bmatrix}$$

$$Ly = b \Rightarrow y = \begin{bmatrix} 1 \\ -10^9 \end{bmatrix},$$

$$Ux = y \Rightarrow x = \begin{bmatrix} {\color{red}0} \\ 1 \end{bmatrix}$$

交换方程组的两行

$$
\begin{aligned}
x_1 + x_2 &= 2 \\
10^{-9} x_1 + x_2 &= 1
\end{aligned}
$$

$$l_{21} = a_{21}/a_{11} = 10^{-9}, a_{22} = 1 - l_{21} \times 1 = 1 - 0.000000001 = 1$$

$$L = \begin{bmatrix} 1 & 0 \\ 10^{-9} & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$Ly = b \Rightarrow y = \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

$$Ux = y \Rightarrow x = \begin{bmatrix} \color{red}{1} \\ 1 \end{bmatrix}$$

Gauss 列主元三角分解法

Gauss 列主元三角分解法与全主元三角分解法的区别就是在消元过程中只作行变换, 这样即可以减少选择主元时的逻辑计算量, 又可以避免记录交换信息.

与全主元 Gauss 消去法的差别仅在于，在 k 步只在 $A_{22}^{(k-1)}$ 的第 k 列上寻找绝对值最大元，即

$$|a_{p,k}^{(k-1)}| = \max\{|a_{i,k}^{(k-1)}| : k \le i \le n\}$$

这样只进行行交换，不进行列交换，即 $P_k = I_{p,k}$, $Q_k = I$. 用矩阵表示就是

$$L_k P_{k-1} \cdots L_1 P_1 A = A^{(k)} = (a_{i,j}^{(k)})$$

则有

$$PA = LU$$

Gauss 列主元三角分解法算法

for k=1:n-1

确定 $p(k \le p \le n)$，使得

$$A(p,k) = \max\{|A(i,k)| : i = k : n\}$$

$$A(k,1:n) \leftrightarrow A(p,1:n) \quad (\text{交换}k\text{行和}p\text{行})$$

$$u(k) = p \quad (\text{记录置换矩阵}P_k)$$

if $A(k,k) \neq 0$

$A(k+1:n,k) = A(k+1:n,k)/A(k,k)$

$A(k+1:n,k+1:n) = A(k+1:n,k+1:n) - A(k+1:n,k)A(k,k+1:n)$

else

stop(矩阵奇异)

end

end

列主元三角分解法计算工作量:

- 选主元

$$\sum_{k=1}^{n-1}(n-k+1)=\frac{n(n+1)}{2}-1$$

- 消元

$$\sum_{k=1}^{n-1}[(n-k)+2(n-k)^2]=\frac{2n^3}{3}+O(n^2)$$

- 总运算工作量

$$\sum_{k=1}^{n-1}[(n-k)+2(n-k)^2+(n-k+1)]=\frac{2}{3}n^3+O(n^2)$$

例题

用 Gauss 列主元法求解下列方程组

$$
\begin{aligned}
2x_1 + 2x_2 + 3x_3 &= 3 \\
4x_1 + 7x_2 + 7x_3 &= 1 \\
-2x_1 + 4x_2 + 5x_3 &= -7
\end{aligned}
$$

解：首先写出增广矩阵

$$
[A|b] = \begin{bmatrix} 2 & 2 & 3 & 3 \\ 4 & 7 & 7 & 1 \\ -2 & 4 & 5 & -7 \end{bmatrix}
$$

$$
\Rightarrow \begin{bmatrix} 4 & 7 & 7 & 1 \\ 2 & 2 & 3 & 3 \\ -2 & 4 & 5 & -7 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 7 & 7 & 1 \\ 0 & -3/2 & -1/2 & 5/2 \\ 0 & 15/2 & 17/2 & -13/2 \end{bmatrix}
$$

$$
\Rightarrow \begin{bmatrix} 4 & 7 & 7 & 1 \\ 0 & 15/2 & 17/2 & -13/2 \\ 0 & -3/2 & -1/2 & 5/2 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 7 & 7 & 1 \\ 0 & 15/2 & 17/2 & -13/2 \\ 0 & 0 & 6/5 & 6/5 \end{bmatrix}
$$

从而可得

$$
x = [2 \quad -2 \quad 1]^T
$$

列主元三角分解法

$$[A]=\begin{bmatrix}2&2&3\\4&7&7\\-2&4&5\end{bmatrix}\Rightarrow \text{交换行}\begin{bmatrix}4&7&7\\2&2&3\\-2&4&5\end{bmatrix}\Rightarrow\begin{bmatrix}4&7&7\\1/2&-3/2&-1/2\\-1/2&15/2&17/2\end{bmatrix}$$

$$\Rightarrow \text{交换行}\begin{bmatrix}4&7&7\\-1/2&15/2&17/2\\1/2&-3/2&-1/2\end{bmatrix}\Rightarrow\begin{bmatrix}4&7&7\\-1/2&15/2&17/2\\1/2&-1/5&6/5\end{bmatrix}$$

$$b=\begin{bmatrix}3\\1\\-7\end{bmatrix}\Rightarrow \text{交换行}\begin{bmatrix}1\\3\\-7\end{bmatrix}\Rightarrow \text{交换行}\begin{bmatrix}1\\-7\\3\end{bmatrix}$$

列主元三角分解法

$$[L] = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 1/2 & -1/5 & 1 \end{bmatrix}, \quad [U] = \begin{bmatrix} 4 & 7 & 7 \\ 0 & 15/2 & 17/2 \\ 0 & 0 & 6/5 \end{bmatrix}$$

$PA = LU$, 求解 $Ax = b \Rightarrow PAx = Pb \Rightarrow LUx = Pb \Rightarrow$

$$Ly = Pb = \begin{bmatrix} 1 \\ -7 \\ 3 \end{bmatrix} \quad \Rightarrow y = \begin{bmatrix} 1 \\ -13/2 \\ 6/5 \end{bmatrix}$$

$$Ux = y \quad \Rightarrow x = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

目 录

平方根法

平方根法又称为 Cholesky(乔列斯基) 分解方法, 是求解对称正定线性方程组的常用方法.

实对称正定矩阵 $A \in R^{n\times n}$ 的性质

- A^{-1} 对称正定，且 $a_{ii} > 0$
- A 的顺序主子阵 A_k 对称正定
- A 的特征值大于零
- A 的全部顺序主子式大于零（充要条件）

定理 (Cholesky 分解定理)

如果 $A \in R^{n\times n}$ 是对称正定矩阵，则存在一个对角元素为正数的下三角矩阵 $L \in R^{n\times n}$，使得 $A = LL^T$。L 称为 A 的 *Cholesky* 因子.

证明：由于 $A \in R^{n\times n}$ 是对称正定矩阵，则 A 的全部顺序主子阵均正定，因此可知，存在一个单位下三角阵 $\widetilde{L}$ 和一个上三角阵 U 使得 $A = \widetilde{L}U$. 令

$$D = \mathrm{diag}(u_{11}, u_{22}, \cdots, u_{nn}), \quad \widetilde{U} = D^{-1}U,$$

则

$$\widetilde{U}^T D \widetilde{L}^T = A^T = A = \widetilde{L} D \widetilde{U}$$

从而

$$\widetilde{L}^T \widetilde{U}^{-1} = D^{-1} \widetilde{U}^{-T} \widetilde{L} D$$

上式左边为单位上三角矩阵，右边是一个下三角矩阵，故两边均为单位矩阵. 于是，

$$\widetilde{U} = \widetilde{L}^T \Longrightarrow A = \widetilde{L} D \widetilde{L}^T$$

由此可知 D 的对角元均为正数，令 $L = \widetilde{L}\mathrm{diag}(\sqrt{u_{11}}, \cdots, \sqrt{u_{nn}})$, 可得 $A = LL^T$, 且 L 的对角元 $l_{ii} = \sqrt{u_{ii}} > 0 (i = 1, \cdots, n)$.

Cholesky 分解

$$Ax = b \Leftrightarrow LL^Tx = b \Leftrightarrow \begin{cases} Ly = b \\ L^Tx = y \end{cases}$$

平方根法求解对称正定线性方程组 $Ax = b$ 的计算步骤:

1. 求 A 的 Cholesky 分解 $A = LL^T$;
2. 求解下三角形方程组 $Ly = b$;
3. 求解下三角形方程组 $L^Tx = y$.

Cholesky 分解计算, 以 4 阶对阵矩阵为例

设 $\mathbf{A}=\begin{bmatrix} a_{11} & * & * & * \\ a_{21} & a_{22} & * & * \\ a_{31} & a_{32} & a_{33} & * \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$, $\mathbf{L}=\begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ l_{31} & l_{32} & l_{33} & \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix}$,

$$\mathbf{L}\mathbf{L}^T=\begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ l_{31} & l_{32} & l_{33} & \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix}\begin{bmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ & l_{22} & l_{32} & l_{42} \\ & & l_{33} & l_{43} \\ & & & l_{44} \end{bmatrix}$$

比较 $\mathbf{A}=\mathbf{L}\mathbf{L}^T$ 两边对应的元素，得

$$a_{11}=l_{11}^2 \quad \to \quad l_{11}=\sqrt{a_{11}}$$

$$a_{i1}=l_{11}l_{i1} \quad \to \quad l_{i1}=a_{i1}/l_{11}, \qquad i=2:n.$$

$$a_{22}=l_{21}^2+l_{22}^2 \quad \to \quad l_{22}=\sqrt{a_{22}-l_{21}^2},$$

Cholesky 分解计算

$$a_{i2}=l_{21}l_{i1}+l_{22}l_{i2}\ \rightarrow l_{i2}=\left(a_{i2}-l_{21}l_{i1}\right)/l_{22}\text{，}\quad i=3:n$$

$$a_{33}=l_{31}^2+l_{32}^2+l_{33}^3\ \rightarrow\ l_{33}=\sqrt{a_{33}-l_{31}^2-l_{32}^2}\text{，}$$

$$a_{i3}=l_{31}l_{i1}+l_{32}l_{i2}+l_{33}l_{i3}\ \rightarrow l_{i3}=\left(a_{i3}-l_{31}l_{i1}-l_{32}l_{i2}\right)/l_{33}\text{，}\quad i=4:n$$

一般地，由公式 $A=LL^T$ 可知

$$a_{ij}=\sum_{r=1}^{j}l_{ir}l_{jr}=l_{ij}l_{jj}+\sum_{r=1}^{j-1}l_{ir}l_{jr},\quad 1\leq j\leq i\leq n$$

从而可知

$$l_{11}=\sqrt{a_{11}}$$

$$l_{jj} = (a_{jj} - \sum_{r=1}^{j-1} l_{jr}^2)^{1/2}, \quad j = 2, 3, \cdots, n.$$

$$l_{ij} = \frac{a_{ij} - \sum_{r=1}^{j-1} l_{ir} l_{jr}}{l_{jj}}, \quad j = 1, 2, 3, \cdots, n, \quad i = j+1, \cdots, n.$$

从上可知，矩阵 L 是逐列进行相关计算的. 由于计算过程中 A 的元素 a_{ij} 被用来计算出 l_{ij} 以后不再使用，所以 L 的元素可以存储在 A 的对应位置

$$\begin{bmatrix} l_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ l_{21} & l_{22} & a_{23} & \cdots & a_{2n} \\ l_{31} & l_{32} & l_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{nn} \end{bmatrix}$$

Cholesky 分解算法

for k=1:n

$$A(k,k)=\sqrt{A(k,k)}$$
$$A(k+1:n,k)=A(k+1:n,k)/A(k,k)$$

for j=k+1:n

$$A(j:n,j)=A(j:n,j)-A(j:n,k)A(j,k)$$

end

end

Cholesky 分解法运算工作量为

$$\sum_{k=1}^{n}[1+(n-k)+\sum_{j=k+1}^{n}2(n-j+1)]=\frac{1}{3}n^3+O(n^2)$$

例题

运用 Cholesky 分解法求解下列方程组

$$
\begin{aligned}
4x_1 - x_2 + x_3 &= 6 \\
-x_1 + 4.25x_2 + 2.75x_3 &= -0.5 \\
x_1 + 2.75x_2 + 3.5x_3 &= 1.25
\end{aligned}
$$

解： 系数矩阵为

$$
A = \begin{bmatrix} 4 & & \\ -1 & 4.25 & \\ 1 & 2.75 & 3.5 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & & \\ -1/2 & 4.25 & \\ 1/2 & 2.75 & 3.5 \end{bmatrix}
$$

$$\Rightarrow \begin{bmatrix} 2 & & \\ -1/2 & 2 & \\ 1/2 & 1.5 & 3.5 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & & \\ -1/2 & 2 & \\ 1/2 & 1.5 & 1 \end{bmatrix} = L$$

$$b = [6 \quad -1/2 \quad 1.25]^T$$

求

$$\begin{aligned} Ly &= b \\ \Rightarrow y &= [3 \quad 0.5 \quad -1]^T \\ L^T x &= y \\ \Rightarrow x &= [2 \quad 1 \quad -1]^T \end{aligned}$$

Cholesky 分解法求解方程组中需说明的几个问题

- 工作量：约为 Gauss 消去法的一半
- 稳定性：是数值稳定的 $|l_{ik}| \leq \sqrt{a_{ii}} (i \geq k)$
- 不必选主元：A 的正定性和算法的稳定性
- 缺陷：存在开平方运算
- 改进方法：LDL^T 分解, 即

$$A = LDL^T, \quad L = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix}, \quad D = \mathrm{diag}(d_1, \cdots, d_n).$$

改进的平方根法，以4阶矩阵 A 为例

$$L=\begin{bmatrix}1&&&\\l_{21}&1&&\\l_{31}&l_{32}&1&\\l_{41}&l_{42}&l_{43}&1\end{bmatrix},\quad D=\begin{bmatrix}d_1&&&\\&d_2&&\\&&d_3&\\&&&d_4\end{bmatrix},$$

$$LD=\begin{bmatrix}d_1&&&\\d_1l_{21}&d_2&&\\d_1l_{31}&d_2l_{32}&d_3&\\d_1l_{41}&d_2l_{42}&d_3l_{43}&d_4\end{bmatrix},\quad L^T=\begin{bmatrix}1&l_{21}&l_{31}&l_{41}\\&1&l_{32}&l_{42}\\&&1&l_{43}\\&&&1\end{bmatrix}$$

比较 $\mathbf{A=LDL^T}$ 两边对应元素。

第1列： $a_{11}=d_1\quad a_{21}=d_1l_{21}$，……，$a_{n1}=d_1l_{n1}$

由此得： $d_1=a_{11}$，$l_{i1}=a_{i1}/d_1, i=2:n$

改进的平方根法

第2列： $a_{22}=d_1l_{21}^2+d_2$， $a_{i2}=d_1l_{21}l_{i1}+d_2l_{i2}, i=3:n$

由此得： $d_2=a_{22}-d_1l_{21}^2$， $l_{i2}=\left(a_{i2}-d_1l_{21}l_{i1}\right)/d_2$， $i=3:n$

第3列：

$$d_3=a_{33}-d_1l_{31}^2-d_2l_{32}^2，\quad l_{i3}=\left(a_{i3}-d_1l_{31}l_{i1}-d_2l_{32}l_{i2}\right)/d_3，\quad i=4:n$$

记 $v_1=d_1l_{31}$， $v_2=d_2l_{32}$， **即** $v_k=d_kl_{3k}, k=1:2$

则上面公式变为

$$d_3=a_{33}-v_1l_{31}-v_2l_{32}=a_{33}-\sum_{k=1}^{2}v_kl_{jk}，$$

$$l_{i3}=\left(a_{i3}-v_1l_{i1}-v_2l_{i2}\right)/d_3=\left(a_{i3}-\sum_{k=1}^{2}v_kl_{ik}\right)\Big/d_3，\quad i=4:n$$

一般情况

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

$$l_{ij} d_j = a_{ij} - \sum_{k=1}^{j-1} l_{ik} d_k l_{jk}, \quad 1 \le j \le i \le n$$

由此可得计算 l_{ij} 和 d_j 的公式

$$\begin{aligned} v_k &= d_k l_{jk}, \quad k = 1, \cdots, j-1. \\ d_j &= a_{jj} - \sum_{k=1}^{j-1} l_{jk} v_k \\ l_{ij} &= (a_{ij} - \sum_{k=1}^{j-1} l_{ik} d_k l_{jk}) / d_j, \quad i = j+1, \cdots, n \end{aligned}$$

此法称为改进的平方根法. 其算法如下:

```
for j=1:n
    for i=1:j-1
        v(i)=A(j,i)A(i,i)
    end
```

$$A(j,j) = A(j,j) - A(j,1:j-1)v(1:j-1);$$
$$A(j+1:n,j) = (A(j+1:n,j) - A(j+1:n,j+1:n)v(1:j-1))/A(j,j);$$

```
end
```

此方法的运算量与平方根一样，都是 $n^3/3$, 优点就是不需要开方运算.

例题

运用改进的 Cholesky 分解法求解下列方程组

$$\begin{aligned} 4x_1 - x_2 + x_3 &= 6 \\ -x_1 + 4.25x_2 + 2.75x_3 &= -0.5 \\ x_1 + 2.75x_2 + 3.5x_3 &= 1.25 \end{aligned}$$

解： 系数矩阵为

$$A = \begin{bmatrix} 4 & & \\ -1 & 4.25 & \\ 1 & 2.75 & 3.5 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & & \\ -1/4 & 4.25 & \\ 1/4 & 2.75 & 3.5 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 4 & & \\ -1/4 & 4 & \\ 1/4 & 0.75 & 3.5 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & & \\ -1/4 & 4 & \\ 1/4 & 0.75 & 1 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & & \\ -1/4 & 1 & \\ 1/4 & 0.75 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & & \\ & 4 & \\ & & 1 \end{bmatrix}, \quad b = [6 \quad -1/2 \quad 1.25]^T$$

求

$$\begin{aligned} Ly &= b \\ \Rightarrow y &= [6 \quad 1 \quad -1]^T \\ L^T x &= D^{-1} y \\ \Rightarrow x &= [2 \quad 1 \quad -1]^T \end{aligned}$$

直接法小结

方法	问题类型	运算量
前代法	$Lx = b$,L 下三角阵	n^2
回代法	$Ux = b$, U 上三角阵	n^2
Gauss 消去法	$Ax = b$, $A = LU$ 三角分解	$\frac{2}{3}n^3$
追赶法	$Ax = b$, A 三对角阵	n
全主元消去法	$Ax = b, PAQ = LU$	n^3
列主元消去法	$Ax = b, PA = LU$	$\frac{2}{3}n^3$
平方根法	$Ax = b$, $A = LL^T$ 对称正定	$\frac{1}{3}n^3$
修正平方根法	$Ax = b$, $A = LDL^T$ 对称正定	$\frac{1}{3}n^3$

目 录

§3.4.1 向量范数

定义 (向量范数)

设 $\|\cdot\|$ 是 $R^n \to R$ 的一个映射，若对 $x \in R^n$ 满足

(1) 正定性： $\|x\| \geq 0, \forall x \in R^n$ 且 $\|x\| = 0 \Leftrightarrow x = 0;$

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

(3) 三角不等式： $\|x + y\| \leq \|x\| + \|y\|, \forall x, y \in R^n$

则称 $\|\cdot\|$ 是定义在 R^n 的一个范数.

常用 p 范数

设 $x = (x_1, \cdots, x_n)^T \in R^n$, 常用的 p 范数 (Holder 范数)

$$\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}, \quad 1 \leq p < +\infty$$

例如:$p=1,2,\infty$ 时，分别称为 1 范数、2 范数及无穷范数

$$\begin{aligned} \|x\|_1 &= |x_1|+\cdots+|x_n| \\ \|x\|_2 &= (|x_1|^2+\cdots+|x_n|^2)^{1/2}=\sqrt{x^Tx} \\ \|x\|_\infty &= \max\{|x_i| : i=1,2,\cdots,n\} \end{aligned}$$

三个重要不等式

1. Holder 不等式

$$|x^Ty| \le \|x\|_p\|y\|_q, \quad x,y\in R^n, \quad \frac{1}{p}+\frac{1}{q}=1, \quad 1\le p,q\le +\infty.$$

2. Cauchy-Schwartz 不等式

$$|x^Ty| \le \|x\|_2\|y\|_2, \quad x,y\in R^n.$$

3. Minkowshi 不等式

$$\|x+y\|_p \le \|x\|_p+\|y\|_p, \quad x,y\in R^n, \quad 1\le p\le +\infty..$$

定理（范数等价性）

设 $\|\cdot\|_\alpha$ 和 $\|\cdot\|_\beta$ 是定义在 R^n 上的任意两个范数，则存在正常数 c_1, c_2 使得对一切 $x \in R^n$

$$c_1\|x\|_\alpha \le \|x\|_\beta \le c_2\|x\|_\alpha$$

比如常见的三种范数满足如下关系：

$$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2$$
$$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$$
$$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$$

我们给出第一个等价关系的证明. 设 $x = (x_1, \cdots, x_n)^T \in R^n$,

$$\|x\|_1^2 = (\sum_{i=1}^n |x_i|)^2 \ge \sum_{i=1}^n |x_i|^2 = \|x\|_2^2$$

记 $y=(1,\cdots,1)^T$, 由 Cauchy-Schwartz 不等式知

$$\|x\|_1=\sum_{i=1}^{n}|x_i|=\sum_{i=1}^{n}|x_i|\cdot 1=\sum_{i=1}^{n}|x_i|\cdot y_i\le\|x\|_2\|y\|_2=\sqrt{n}\|x\|_2$$

定理（向量序列的范数极限）

设 $x_k\in R^n$, 则 $\lim\limits_{k\to+\infty}\|x_k-x\|=0$ 的充要条件是 $\lim\limits_{k\to+\infty}|(x_k)_i-(x)_i|=0$.
即向量序列的范数收敛等价于向量分量收敛.

重要性质

(1) 向量范数的等价性具有传递性;

(2) R^n 的所有向量范数是彼此等价的.

§3.4.2 矩阵范数

定义 (矩阵范数)

设 $\|\cdot\|$ 是 $R^{n\times n} \to R$ 的一个映射，若对 $A \in R^{n\times n}$, 存在唯一的实数 $\|A\|$ 与之相对应，满足

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

(2) 齐次性：$\|\alpha A\| = |\alpha| \|A\|, \forall A \in R^{n\times n}, \alpha \in R$;

(3) 三角不等式：$\|A + B\| \leq \|A\| + \|B\|, \forall A, B \in R^{n\times n}$;

(4) 相容性：$\|AB\| \leq \|A\| \, \|B\|, \forall A, B \in R^{n\times n}$;

则称 $\|A\|$ 是定义在 $R^{n\times n}$ 中矩阵 A 的范数.

备注

注意到，$R^{n\times n}$ 上的矩阵可以看成是 R^{n^2} 上的向量，所以矩阵范数具有向量范数的一切性质，比如

(1) 任意两个矩阵范数是等价的;

(2) 矩阵序列的范数收敛等价于矩阵分量收敛

$$\lim_{k\to+\infty}\|A_k - A\| = 0 \Leftrightarrow \lim_{k\to+\infty}|a_{ij}^{(k)} - a_{ij}| = 0.$$

定义 (向量范数与矩阵范数相容性)

若矩阵范数 $\|\cdot\|_M$ 和向量范数 $\|\cdot\|_v$ 满足

$$\|Ax\|_v \le \|A\|_M\|x\|_v, \quad \forall A \in R^{n\times n}, x \in R^n$$

则称矩阵范数 $\|\cdot\|_M$ 和向量范数 $\|\cdot\|_v$ 是相容的.

事实上，对于任意给定的向量范数，都可以构造一个与之相容的矩阵范数.

定理

设 $\|\cdot\|$ 是 R^n 中的任意一种向量范数，若定义

$$|||A||| = \max_{\|x\|=1} \|Ax\|, \quad A \in R^{n\times n}$$

则 $|||\cdot|||$ 是 $R^{n\times n}$ 上的一个矩阵范数，又称其为从属于向量 $\|\cdot\|$ 的矩阵范数（算子范数）．

证明：1）矩阵范数与向量范数的相容性：

$$\frac{\|Ax\|}{\|x\|} = \left\| A\frac{x}{\|x\|} \right\| \le |||A||| \Rightarrow \|Ax\| \le |||A|||\,\|x\|$$

2) 证明 $|||\cdot|||$ 满足矩阵范数的四条性质：
(1) 正定性：设 $A \neq 0$, 则存在 e_i, $Ae_i \neq 0$, 由矩阵范数与向量范数的相容性可知

$$0 < \|Ae_i\| \leq |||A|||\|e_i\| \Rightarrow |||A||| > 0.$$

(2) 齐次性:

$$|||\alpha A||| = \max_{\|x\|=1} \|\alpha Ax\| = |\alpha|\, |||A|||.$$

(3) 三角不等式:

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

(4) 相容性:

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

矩阵范数的一般定义形式:

$$\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p, \quad A \in R^{n\times n}.$$

定理

设 $A \in R^{n\times n}$，则有

- 列范数: $\|A\|_1 = \max\limits_{1\leq j\leq n} \sum\limits_{i=1}^{n} |a_{ij}|$,
- 行范数: $\|A\|_\infty = \max\limits_{1\leq i\leq n} \sum\limits_{j=1}^{n} |a_{ij}|$,
- 谱范数: $\|A\|_2 = \sqrt{\lambda_{\max}(A^TA)}$, 其中 $\lambda_{\max}(A^TA)$ 表示 A^TA 的最大特征值.

证明：$A=0$，显然结论成立. 下面的证明都是假定 $A\neq 0$.
1) 列范数 (1)： 记矩阵 $A=[a_1,a_2,\cdots,a_n]$, $\delta=\|a_{j_0}\|_1=\max_{1\leq j\leq n}\|a_j\|_1$.
则对任意 $x\in R^n$, $\|x\|_1=\sum_{i=1}^n |x_i|=1$, 有

$$\|Ax\|_1=\|\sum_{j=1}^n x_j a_j\|_1\leq\sum_{j=1}^n |x_j|\|a_j\|_1\leq(\sum_{j=1}^n |x_j|)\max_{1\leq j\leq n}\|a_j\|_1=\|a_{j_0}\|_1=\delta$$

注意到

$$\|Ae_{j_0}\|_1=\|a_{j_0}\|_1=\delta.$$

因此

$$\|A\|_1=\max_{\|x\|_1=1}\|Ax\|_1=\delta=\max_{1\leq j\leq n}\|a_j\|_1=\max_{1\leq j\leq n}\sum_{i=1}^n |a_{ij}|$$

2) 行范数 (∞)：记 $\eta = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$，则对任一满足 $\|x\|_\infty = 1$ 的 $x \in R^n$ 有

$$\|Ax\|_\infty = \max_{1\le i\le n}|\sum_{j=1}^{n} a_{ij}x_j| \le \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}||x_j| \le \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}| = \eta.$$

令

$$\tilde{x} = [\mathrm{sign}(a_{k1}),\cdots,\mathrm{sign}(a_{kn})]^T.$$

由 $A \ne 0, \|\tilde{x}\|_\infty = 1,$ 且 $\|A\tilde{x}\|_\infty = \eta,$ 从而证明了

$$\|A\|_\infty = \eta = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|.$$

2) 谱范数 (2):

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \max_{\|x\|_2=1} [(Ax)^T Ax]^{1/2} = \max_{\|x\|_2=1} [x^T(A^T A)x]^{1/2}$$

因为 $A^T A$ 是半正定的对称阵, 可设其特征值为

$$\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n \geq 0,$$

其对应的正交规范特征向量为 $v_1, \cdots, v_n \in R^n$, 则对任一满足 $\|x\|_2 = 1$ 的向量 $x \in R^n$ 有

$$x = \sum_{i=1}^{n} \alpha_i v_i, \quad \sum_{i=1}^{n} \alpha_i^2 = 1$$

于是，有

$$x^T(A^T A)x = \sum_{i=1}^{n} \lambda_i \alpha_i^2 \leq \lambda_1$$

另一方面, 取 $x = v_1$， 则

$$x^T(A^TA)x = v_1^T(A^TA)v_1 = v_1^T\lambda_1 v_1 = \lambda_1.$$

所以

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \sqrt{\lambda_1} = \sqrt{\lambda_{\max}(A^TA)}.$$

定理得证.

例： 求下矩阵 A 的 $1, \infty, 2$ 范数

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

解： $\|A\|_p = 3$, $p = 1, \infty, 2$.

定理（谱范数的常用性质）

设 $A \in R^{n\times n}$，则

(1)

$$\|A\|_2 = \max\{|y^T Ax| : x, y \in R^n, \|x\|_2 = \|y\|_2 = 1\}$$

(2)

$$\|A\|_2 = \|A^T\|_2 = \sqrt{\|A^T A\|_2}$$

(3) 对于任意的正交矩阵 U, V 有

$$\|UA\|_2 = \|AV\|_2 = \|A\|_2$$

谱范数不易于计算，因此定义常用的且易于计算的 F 范数 (Frobenius 范数)

$$\|A\|_F = \left(\sum_{i,j=1}^{n} |a_{i,j}|^2 \right)^{\frac{1}{2}}.$$

定义 (谱半径)

设 $A \in C^{n\times n}$，则称

$$\rho(A) = \max\{|\lambda| : \lambda \in \lambda(A)\}$$

为 A 的谱半径，这里 $\lambda(A)$ 表示 A 的所有特征值.

谱半径与矩阵范数的关系

定理

设 $A \in C^{n\times n}$，则有

(1) 对 $C^{n\times n}$ 上的任意矩阵范数 $\|\cdot\|$, 有 $\rho(A) \leq \|A\|$.

(2) 对任给的 $\varepsilon > 0$, 存在 $C^{n\times n}$ 上的某种算子范数 $\|\cdot\|$，使得 $\|A\| \leq \rho(A) + \varepsilon$.

证明：(1) 设 λ 是 A 的任一特征值，使得

$$Ax = \lambda x (x \neq 0)$$

从而有

$$|\lambda|\|x\| = \|\lambda x\| = \|Ax\| \leq \|A\|\|x\| \Rightarrow \rho(A) \leq \|A\|.$$

(2) 由 Jordan 分解定理可知，存在非奇异矩阵 $X \in C^{n\times n}$ 使得

$$X^{-1}AX = \begin{bmatrix} \lambda_1 & \delta_1 & & & \\ & \lambda_2 & \delta_2 & & \\ & & \ddots & \ddots & \\ & & & \lambda_{n-1} & \delta_{n-1} \\ & & & & \lambda_n \end{bmatrix}$$

其中 $\delta_i = 0, 1$. 对任意给定的 $\varepsilon > 0$, 令

$$D_\varepsilon = diag(1, \varepsilon, \varepsilon^2, \cdots, \varepsilon^{n-1})$$

则有

$$D_\varepsilon^{-1}X^{-1}AXD_\varepsilon = \begin{bmatrix} \lambda_1 & \varepsilon\delta_1 & & & \\ & \lambda_2 & \varepsilon\delta_2 & & \\ & & \ddots & \ddots & \\ & & & \lambda_{n-1} & \varepsilon\delta_{n-1} \\ & & & & \lambda_n \end{bmatrix}$$

现定义

$$\|G\|_\varepsilon = \|D_\varepsilon^{-1}X^{-1}GXD_\varepsilon\|_\infty, \quad G \in C^{n\times n}$$

这容易验证这样定义的函数 $\|\cdot\|_\varepsilon$ 是有如下定义的向量范数

$$\|x\|_{XD_\varepsilon} = \|(XD_\varepsilon)^{-1}x\|_\infty, \quad x \in C^n$$

诱导出的算子范数，而且有

$$\|A\|_\varepsilon = \|D_\varepsilon^{-1}X^{-1}AXD_\varepsilon\|_\infty = \max_{1\le i\le n}(|\lambda_i| + |\varepsilon\delta_i|) \le \rho(A) + \varepsilon.$$

其中 $\delta_n = 0$.

定理 (矩阵范数的等价性)

设 $\|\cdot\|_\alpha, \|\cdot\|_\beta$ 是定义在 $R^{n\times n}$ 上的两种范数，则存在正数 m, M，使得

$$m\|A\|_\alpha \le \|A\|_\beta \le M\|A\|_\alpha, \quad \forall A \in R^{n\times n}.$$

易证

$$\|A\|_2 \le \|A\|_F \le \sqrt{n}\|A\|_2$$
$$\frac{1}{n}\|A\|_2 \le \|A\|_\infty \le \sqrt{n}\|A\|_2$$

定理

设 $A \in C^{n\times n}$，则

$$\lim_{k\to\infty} A^k = 0 \Leftrightarrow \rho(A) < 1.$$

证明：必要性. 设 $\lim_{k\to\infty} A^k = 0$, $\lambda \in \lambda(A)$ 满足 $Ax = \lambda x (x \neq 0)$. 则

$$A^k x = \lambda^k x$$

因

$$\rho(A)^k = \rho(A^k) \leq \|A^k\| \to 0$$

故而可知

$$\rho(A) < 1.$$

充分性. 设 $\rho(A) < 1$, 由前面定理可知，必有算子范数 $\|\cdot\|$, 使得 $\|A\| < 1$, 从而

$$0 \le \|A^k\| \le \|A\|^k \to 0, \quad k \to \infty,$$

故而可得

$$\lim_{k\to\infty} A^k = 0.$$

定理

设 $A \in C^{n\times n}$，则

(1) $\sum\limits_{k=1}^{\infty} A^k$ 收敛的充分必要条件是 $\rho(A) < 1$;

(2) 当 $\sum\limits_{k=1}^{\infty} A^k$ 收敛时，有

$$\sum_{i=k}^{\infty} A^k = (I-A)^{-1}$$

且存在从属矩阵范数 $\|\cdot\|$ 满足

$$\|(I-A)^{-1} - \sum_{k=1}^{m} A^k\| \le \frac{\|A\|^{m+1}}{1-\|A\|}, \quad \forall m \in Z.$$

从上定理可以得到如下结论

定理

设 $\|\cdot\|$ 是 $C^{n\times n}$ 上的一个满足条件 $\|I\|=1$ 的矩阵范数，并假设 $A\in C^{n\times n}$ 满足 $\|A\|<1$, 则 $I-A$ 可逆且有

$$\|(I-A)^{-1}\|\le\frac{1}{1-\|A\|}$$

证明：反证法. 假设 $I-A$ 不可逆, 则存在一个非零 $x\in C^n$ 使得 $(I-A)x=0$，从而有 $Ax=x$，这意味着 1 是 A 的一个特征值，故 $1\le\rho(A)\le\|A\|$, 与假设矛盾.

$$\begin{aligned}\|I\| &= \|(I-A)(I-A)^{-1}\|=\|(I-A)^{-1}-A(I-A)^{-1}\|\\ &\ge \|(I-A)^{-1}\|(1-\|A\|)\Rightarrow\|(I-A)^{-1}\|\le\frac{1}{1-\|A\|}\end{aligned}$$

目 录

§3.5.1 线性方程组的敏度分析

问题

对于线性方程组 $Ax=b$，A, b 的微小扰动将对线性方程组的解有何影响?

例：线性方程组

$$\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 4 \\ 4 \end{bmatrix} \Rightarrow x=\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

若方程组右端有扰动 $\delta b=[2\times 10^{-4}, -2\times 10^{-4}]^T$, 则原方程组变为

$$\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix}\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix}=\begin{bmatrix} 4.0002 \\ 3.9998 \end{bmatrix} \Rightarrow \tilde{x}=\begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix}$$

解的相对误差与右端项的相对误差分别为

$$\frac{\|x-\tilde{x}\|}{\|x\|}=\frac{1}{2}, \quad \frac{\|\delta b\|}{\|b\|}=\frac{1}{20000}$$

由此可以看到，解的相对误差是右端项的相对误差的 10000 倍，说明该方程组的解对初始元素的扰动非常敏感。
下面就一般的非奇异线性方程组 $Ax=b$ 讨论其敏感性问题. 假定该方程组经微小扰动后变为

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

从而可知

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

定理

设 $\|\cdot\|$ 是 $R^{n\times n}$ 上的一个满足条件 $\|I\|=1$ 的矩阵范数，并假定 $A\in R^{n\times n}$ 非奇异，$b\in R^n$ 非零. 再假定 $\delta A\in R^{n\times n}$ 满足 $\|A^{-1}\|\|\delta A\|<1$. 若 x 和 $x+\delta x$ 分别为线性方程组

$$Ax=b,\quad (A+\delta A)(x+\delta x)=(b+\delta b)$$

的解，则

$$\frac{\|\delta x\|}{\|x\|}\le\frac{\kappa(A)}{1-\kappa(A)\frac{\|\delta A\|}{\|A\|}}\left(\frac{\|\delta b\|}{\|b\|}+\frac{\|\delta A\|}{\|A\|}\right)$$

这里 $\kappa(A)=\|A\|\|A^{-1}\|$.

证明：由 $(A+\delta A)(x+\delta x)=(b+\delta b)$ 可得

$$\delta x=(A+\delta A)^{-1}(\delta b-\delta Ax)=(I+A^{-1}\delta A)^{-1}A^{-1}(\delta b-\delta Ax)$$

两边取范数可得

$$\begin{aligned}\|\delta x\| &\leq \|(I+A^{-1}\delta A)^{-1}\|\|A^{-1}\|(\|\delta b\|+\|\delta A\|\|x\|)\\ &\leq \frac{\|A^{-1}\|}{1-\|A^{-1}\|\|\delta A\|}(\|\delta b\|+\|\delta A\|\|x\|)\end{aligned}$$

注意到 $\|b\|\leq\|A\|\|x\|$, 从而可得相对误差

$$\frac{\|\delta x\|}{\|x\|}\leq\frac{\|A\|\|A^{-1}\|}{1-\|A^{-1}\|\|\delta A\|}\left(\frac{\|\delta b\|}{\|b\|}+\frac{\|\delta A\|}{\|A\|}\right)$$

定义 (条件数)

设 $A \in R^{n\times n}$ 为可逆矩阵, 则称 $\kappa(A) = \|A\|\|A^{-1}\| = cond(A)$ 为矩阵 A 的条件数.

注

条件数在一定程度上刻画了扰动对方程组的解的影响程度.

- 若线性方程组的条件数 $\kappa(A)$ 很大，则称方程组 $Ax = b$ 为病态方程组, 矩阵 A 为病态矩阵;
- 反之，如果 $\kappa(A)$ 很小，则称方程组 $Ax = b$ 为良态方程组, 矩阵 A 为良态矩阵。

$\boxed{\text{病态方程组对任何算法都将产生数值不稳定性.}}$

条件数与范数有关，给定具体的范数得到特定的条件数，如

- 若矩阵范数取 2-范数，则得到谱条件数：

$$\kappa_2(A) = \|A\|_2\|A^{-1}\|_2$$

- 若矩阵范数取 1-范数，则得到 1 条件数：

$$\kappa_1(A) = \|A\|_1\|A^{-1}\|_1$$

- 若矩阵范数取 ∞-范数，则得到 ∞ 条件数：

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

问题

一个方程在一种范数下是病态的，在另一范数下是否病态？

由矩阵范数的等价性，很容易推证 $C^{n\times n}$ 上任意两个范数下的条件数都是等价的，即

$$c_1\kappa_\alpha(A) \le \kappa_\beta(A) \le c_2\kappa_\alpha(A)$$

如

$$\frac{1}{n}\kappa_2(A) \le \kappa_1(A) \le n\kappa_2(A)$$
$$\frac{1}{n}\kappa_\infty(A) \le \kappa_2(A) \le n\kappa_\infty(A)$$
$$\frac{1}{n^2}\kappa_1(A) \le \kappa_\infty(A) \le n^2\kappa_1(A)$$

定理

设 $\|\cdot\|$ 是 $R^{n\times n}$ 上的一个满足条件 $\|I\|=1$ 的矩阵范数，并假定 $A\in R^{n\times n}$ 非奇异，$\delta A\in R^{n\times n}$ 满足 $\|A^{-1}\|\|\delta A\|<1$. 则 $A+\delta A$ 也是非奇异的，且

$$\frac{\|(A+\delta A)^{-1}-A^{-1}\|}{\|A^{-1}\|}\le\frac{\kappa(A)}{1-\kappa(A)\frac{\|\delta A\|}{\|A\|}}\frac{\|\delta A\|}{\|A\|}.$$

此定理表明，条件数可以作为矩阵求逆问题的条件数.

定理

设 $A\in R^{n\times n}$ 非奇异，则

$$\min\left\{\frac{\|\delta A\|_2}{\|A\|_2}:A+\delta A\text{奇异}\right\}=\frac{1}{\|A\|_2\|A^{-1}\|_2}=\frac{1}{\kappa_2(A)}.$$

即：在谱范数下，一个矩阵的条件数的倒数正好等于该矩阵与全体奇异矩阵所成集合的相对距离.

证明：只需证明下等式成立即可

$$\min\left\{\|\delta A\|_2 : A+\delta A\text{奇异}\right\} = \frac{1}{\|A^{-1}\|_2}.$$

注意到当 $\|A^{-1}\|_2\|\delta A\|_2 < 1$ 时，$A+\delta A$ 可逆，则

$$\min\left\{\|\delta A\|_2 : A+\delta A\text{奇异}\right\} \geq \frac{1}{\|A^{-1}\|_2}.$$

因谱范数是由向量 2 范数诱导的算子范数，故比存在向量 $x \in R^n$ 满足 $\|x\|_2 = 1$, 使得 $\|A^{-1}x\|_2 = \|A^{-1}\|_2$. 令

$$y = \frac{A^{-1}x}{\|A^{-1}x\|_2}, \quad \delta A = -\frac{xy^T}{\|A^{-1}\|_2},$$

则有 $\|y\|_2 = 1$, 且

$$(A+\delta A)y = Ay + \delta Ay = \frac{x}{\|A^{-1}x\|_2} - \frac{xy^T y}{\|A^{-1}\|_2} = \frac{x}{\|A^{-1}x\|_2} - \frac{x}{\|A^{-1}x\|_2} = 0$$

故而有

$$\|\delta A\|_2 = \max_{\|z\|_2=1} \left\| \frac{xy^T}{\|A^{-1}\|_2} z \right\|_2 = \frac{\|x\|_2}{\|A^{-1}\|_2} \max_{\|z\|_2=1} |y^T z| = \frac{1}{\|A^{-1}\|_2}$$

§3.5.2 病态方程的解法

判定方程组为病态的经验方法

(1) 当 $|det(A)|$ 相对来说很小时，或者矩阵 A 的某些行（列）近似线性相关时，可能为病态；

(2) 矩阵在采用选主元消去法求解方程组时，在消元过程中出现很小的主元，可能为病态；

(3) 解方程组得到了一个很大的解，或者特征值相差大数量级，可能为病态；

(4) 当系数矩阵的元素间数量级相差很大，且无一定规则时，可能为病态。

求解病态方程组时，常用的几种处理原则：

(1) 采用高精度的算术运算；

(2) 采用预处理方法；

$$Ax = b \Leftrightarrow PAQQ^{-1}x = Pb \Leftrightarrow \overline{A}\bar{x} = \bar{b}$$

可逆矩阵 P, Q 的选择要求满足 $cond(PAQ) \ll cond(A)$.

(3) 采用某些特殊的数值方法求解；

(4) 重新寻找出现病态的原因，改变原问题的提法。

知识小结

主要内容

三角分解法 Gauss 消去法 列主元消去法 平方根法 范数 敏感度分析 病态问题求解

重点难点

列主元 Gauss 消去法 范数 敏感度分析

Many thanks for your attention！

中国石油大学(华东)
CHINA UNIVERSITY OF PETROLEUM