第3章　解线性方程组的直接法

主讲：纪庆革副教授

中山大学数据科学与计算机学院

E-Mail: 1024180018@qq.com

在实践中存在大量的解线性方程组的问题。

很多数值方法到最后也会涉及到线性方程组的求解问题：

如样条插值，曲线拟合，方程组的Newton迭代等问题。

对线性方程组：

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{cases}$$

或者：$Ax = b$

我们有**Gram**法则：当且仅当 $\det(A) \neq 0$ 时，有唯一的解，而且解为：

$$x_i = \frac{D_i}{D}, D = \det(A), D_i = \det\begin{pmatrix} a_{11} & \cdots & a_{1i-1} & b_1 & a_{1i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni-1} & b_n & a_{ni+1} & \cdots & a_{nn} \end{pmatrix}$$

但**Gram**法则不能用于计算方程组的解，如n＝100，10^{33}次/秒的计算机要算10^{120}年

解线性方程组的方法可以分为2类：

①直接法：准确，可靠，理论上得到的解是精确的

②迭代法：速度快，但有误差

本部分讲解直接法

主要内容

3.1 高斯消元法

3.2 高斯主元素消元法

3.3 直接分解法

3.3 矩阵范数和条件数

3.4 条件数和病态矩阵

3.1 高斯消元法

我们知道，下面有3种方程的解我们可以直接求出：

①

n次运算

$$A = diag(a_{11}, a_{22}, \cdots, a_{nn}) \Rightarrow x_i = \frac{b_i}{a_{ii}}, i = 1, \cdots, n$$

②

（n＋1）n/2次运算

$$A = \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \Rightarrow x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij} x_j}{l_{ii}}, i = 1, \cdots, n$$

③

（n＋1）n/2次运算

$$A=\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix} \Rightarrow x_i=\frac{b_i-\sum_{j=i+1}^{n}u_{ij}x_j}{u_{ii}},\ i=n,\cdots,1$$

对方程组，作如下的变换，解不变

①交换两个方程的次序

②一个方程的两边同时乘以一个非0的数

③一个方程的两边同时乘以一个非0数，加到另一个方程

因此，对应的对增广矩阵***(A，b)***，作如下的变换，解不变

①交换矩阵的两行

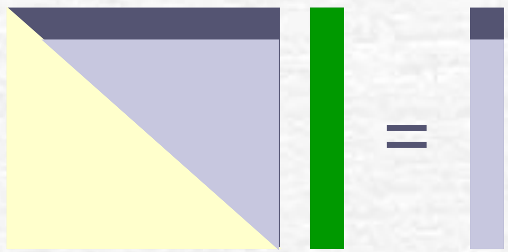
②某一行乘以一个非0的数

③某一个乘以一个非0数，加到另一行

消元法就是对增广矩阵作上述行的变换，变为我们已知的3种类型之一，而后求根

➢ 高斯消元法：

思路 首先将*A*化为上三角阵，再回代求解 。



$$
\begin{pmatrix}
a_{11} & a_{12} & \Lambda & a_{1n} & b_1 \\
a_{21} & a_{22} & \Lambda & a_{2n} & b_2 \\
M & M & O & M & M \\
a_{n1} & a_{n2} & \Lambda & a_{nn} & b_n
\end{pmatrix}
\Rightarrow
\begin{pmatrix}
a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & L & a_{1n}^{(1)} & b_1^{(1)} \\
0 & a_{22}^{(2)} & a_{23}^{(2)} & L & a_{2n}^{(2)} & b_2^{(2)} \\
0 & 0 & a_{33}^{(3)} & L & a_{3n}^{(3)} & b_3^{(3)} \\
M & M & M & O & M & M \\
0 & 0 & 0 & L & a_{nn}^{(n)} & b_n^{(n)}
\end{pmatrix}
$$

步骤如下：

第一步：$$第1行\times\frac{-a_{i1}}{a_{11}}+第i行, i=2,\Lambda\ ,n$$

$$\begin{pmatrix} a_{11} & a_{12} & \Lambda & a_{1n} & b_1 \\ a_{21} & a_{22} & \Lambda & a_{2n} & b_2 \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} & \mathrm{M} \\ a_{n1} & a_{n2} & \Lambda & a_{nn} & b_n \end{pmatrix} \Rightarrow \begin{pmatrix} a_{11} & a_{12} & \Lambda & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & \Lambda & a_{2n}^{(2)} & b_2^{(2)} \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} & \mathrm{M} \\ 0 & a_{n2}^{(2)} & \Lambda & a_{nn}^{(2)} & b_n^{(2)} \end{pmatrix}$$

运算量： ***(n-1)*(1+n)***

第二步： $$第2行\times\frac{-a_{i2}^{(2)}}{a_{22}^{(2)}}+第i行, i=3,\Lambda\ ,n$$

$$\begin{pmatrix} a_{11} & a_{12} & \Lambda & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & \Lambda & a_{2n}^{(2)} & b_2^{(2)} \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} & \mathrm{M} \\ 0 & a_{n2}^{(2)} & \Lambda & a_{nn}^{(2)} & b_n^{(2)} \end{pmatrix} \Rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} & \Lambda & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \Lambda & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \Lambda & a_{3n}^{(3)} & b_3^{(3)} \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} & \mathrm{M} \\ 0 & 0 & a_{n3}^{(3)} & \Lambda & a_{nn}^{(3)} & b_n^{(3)} \end{pmatrix}$$

运算量： ***(n-2)*(1+n-1)=(n-2)n***

类似的做下去，我们有：

第k步：$第k行\times\frac{-a_{ik}^{(k)}}{a_{kk}^{(k)}}+第i行, i=k+1,\Lambda\ ,n$

运算量： ***(n－k)*(1＋n－k＋1)=(n－k)(n－k＋2)***

n－1步以后，我们可以得到变换后的矩阵为：

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \Lambda & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \Lambda & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \Lambda & a_{3n}^{(3)} & b_3^{(3)} \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} & \mathrm{M} \\ 0 & 0 & 0 & \Lambda & a_{nn}^{(n)} & b_n^{(n)} \end{pmatrix}$$

因此，总的运算量为：

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

加上 解上述上三角阵的运算量***(n+1)n/2***，总共为：

$$\frac{n^3}{3}+n^2-\frac{n}{3}=O(n^3)$$

注意到，计算过程中$a_{kk}^{(k)}$处在被除的位置，因此整个计算过程要保证它不为0

所以，Gauss消元法的可行条件为：$a_{kk}^{(k)} \neq 0$

就是要求A的所有顺序主子式均不为0，即

$$\det\begin{pmatrix} a_{11} & \Lambda & a_{1i} \\ \mathrm{M} & \mathrm{O} & \mathrm{M} \\ a_{i1} & \Lambda & a_{ii} \end{pmatrix} \neq 0, i = 1, \Lambda\ , n$$

因此，有些有解的问题，不能用Gauss消元求解

另外，如果某个 $a_{kk}^{(k)}$ 很小的话，会引入大的误差

例：单精度解方程组 $\begin{cases} 10^{-9}x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases}$

/* 精确解为 $x_1 = \dfrac{1}{1-10^{-9}} = 1.\overbrace{00\ldots0}^{8个}100\ldots$ 和 $x_2 = 2 - x_1 = 0.\overbrace{99\ldots9}^{8个}899\ldots$ */

用Gaussian 消元法计算：

$$m_{21} = a_{21}/a_{11} = 10^9$$

$$a_{22} = 1 - m_{21}\times 1 = 0.\underbrace{0\ldots0}_{8个}1\times 10^9 - 10^9 \doteq -10^9$$

$$b_2 = 2 - m_{21}\times 1 \doteq -10^9$$

$$\Rightarrow \begin{bmatrix} 10^{-9} & 1 & \\ 0 & -10^9 & -10^9 \end{bmatrix}$$

小主元可能导致计算失败。

$$\Rightarrow \quad x_2 = 1, \quad x_1 = 0$$

3.2　高斯主元素消元法

1、问题的提出

在高斯消元过程中，如果出现 $a_{kk}^{(k-1)}=0$ 的情况，
这时消元法将无法进行；

即使主元素 $a_{kk}^{(k-1)}\neq 0$ ，**但当其值很小时，用其作除数，会导致不可靠结果。**

例1：P40

2、完全主元素消元法

在A中选取绝对值最大的元素作为主元素，如

$$\left|a_{i_1 j_1}\right| = \max_{\substack{1\le i\le n\\ 1\le j\le n}} \left|a_{ij}\right|$$

然后交换$\tilde{A}^{(0)}$中的第1行与第i1行，第1列与第j1列，经第1次消元计算，得

$$\tilde{A}^{(0)} \to \tilde{A}^{(1)}$$

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

其次，在$\tilde{A}^{(1)}$中的第2行至第n行及第2列至第n列选取绝对值最大的元素作为主元素，如

$$\left|a_{i_2 j_2}\right| = \max_{\substack{2 \le i \le n \\ 2 \le j \le n}} \left|a_{ij}\right|$$

然后交换$\tilde{A}^{(1)}$中的第2行与第 i_2行，第2列与第j_2列，经第2次消元计算，得

$$\tilde{A}^{(1)} \to \tilde{A}^{(2)}$$

$$\tilde{A}^{(2)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \Lambda & a_{1n} & a_{1(n+1)} \\ & a_{22} & a_{23} & \Lambda & a_{2n} & a_{2(n+1)} \\ & & a_{33} & \Lambda & a_{3n} & \\ & & \mathrm{M} & & \mathrm{M} & \mathrm{M} \\ & & a_{n3} & \Lambda & a_{nn} & a_{n(n+1)} \end{bmatrix}$$

重复上述过程，假设已完成了第k-1次消元，则在$\widetilde{A}^{(k-1)}$的第k行到第n行，第k列到第n列中选取绝对值最大的元素作为主元素，如

$$\left|a_{i_k j_k}\right| = \max_{\substack{k \le i \le n \\ k \le j \le n}} \left|a_{ij}\right|$$

然后交换$\widetilde{A}^{(1)}$中的第k行与第 i_k行，第k列与第j_k列，进行第k次消元计算。最后得

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

通过回代求得原方程组的解。该方法称为高斯完全主元素消元法

如何节省存储空间？

在消元过程中，可用约化后新的a_{ij}冲掉约化前旧的a_{ij}，在回代过程中，同样可用代入后新的常数项冲掉代入前旧的常数项，并以此表示未知量。

这种做法与普通消元法相同。

2、列主元消元法

在 ***Gauss*** 消元第 **k** 步之前，做如下的事情：

若 $\max\limits_{k\le i\le n} |a_{ik}^{(k)}| = |a_{jk}^{(k)}|$ 交换k行和j行

行的交换，不改变方程组的解，同时又有效地克服了 ***Gauss*** 消元的缺陷

例：$\begin{bmatrix} 10^{-9} & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 2 \\ 10^{-9} & 1 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$

$$\Rightarrow \quad x_2 = 1, \quad x_1 = 1 \checkmark$$

全主元消去法与列主元消去法的优缺点？

3、**Gauss-Jordan消元法**

将在***Gauss***消元第***k***步，变为

$$第k行\times\frac{-a_{ik}^{(k)}}{a_{kk}^{(k)}}+第i行, i=1,\Lambda\ ,k-1,k+1,\Lambda\ ,n$$

将该行上三角地部分也变为0

最后变为一个对角阵。

它的运算次数比***Gauss***消元多。使用于计算多个系数一样的方程组，如

$$AX=B$$

X*，*B均为矩阵

3.3 直接分解法

***Gauss*消元法的第*k*步：**

$$第k行\times\frac{-a_{ik}^{(k)}}{a_{kk}^{(k)}}+第i行, i=k+1,\cdots,n$$

从矩阵理论来看，相当于左乘矩阵

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & \vdots & \vdots & \vdots & \\ & & 1 & & & \\ & & l_{k+1k}^{(k)} & 1 & & \\ & & \vdots & & \ddots & \vdots \\ & & l_{nk}^{(k)} & & & 1 \end{pmatrix}, l_{ik}^{(k)}=\frac{-a_{ik}^{(k)}}{a_{kk}^{(k)}}, i=k+1,\cdots,n$$

因此，整个Gauss消元法相当于左乘了一个单位下三角阵

$$L=\begin{pmatrix} 1 & & & & & \\ l_{21} & \ddots & & & & \\ & & 1 & & & \\ \vdots & & l_{k+1k}^{(k)} & 1 & & \\ & & \vdots & & \ddots & \\ l_{n1} & \cdots & l_{nk}^{(k)} & \cdots & l_{nn-1}^{(n)} & 1 \end{pmatrix}$$

所以有 $\exists L \quad s.t. \quad LA=U$, **L**为单位下三角阵，其中$L=L_{n-1}\cdots L_2L_1$。

U为上三角阵 $\Rightarrow \exists L \quad s.t. \quad A=LU$ ，其中 $L=L_1^{-1}L_2^{-1}\cdots L_{n-1}^{-1}$

因此 $Ax=b\Rightarrow LUx=b\Rightarrow\begin{cases} Ly=b \\ Ux=y \end{cases}$

我们可以通过2次反代过程求解方程组

定义3.1 设A 为n 阶矩阵（$n\geq2$）。称$A=LU$为矩阵A的三角分解，其中L是下三角矩阵，U是上三角矩阵。

定义3.2 如果L是单位下三角矩阵， U是上三角矩阵，则称$A=LU$为 Doolittle 分解；如果L是下三角矩阵，U是单位上三角矩阵，则称$A=LU$为 Crout 分解。

定理3.1 如果n阶（$n\geq2$）矩阵A的前n-1个顺序主子式不为零，则A有惟一 Doolittle 分解和惟一的 Crout 分解。

例1: 求矩阵 $A=\begin{bmatrix}1 & 1 & 1\\ 0 & 4 & -1\\ 2 & -2 & -1\end{bmatrix}$

的LU分解。（p49）

解：由高斯消元法

$$l_{21}=\frac{a_{21}}{a_{11}}=0, l_{31}=\frac{a_{31}}{a_{11}}=2$$

且 $A=A^{(0)}\rightarrow\begin{bmatrix}1 & 1 & 1\\ 0 & 4 & -1\\ 0 & -4 & -3\end{bmatrix}=A^{(1)}$

接下来有$l_{32} = \dfrac{a_{32}^{(1)}}{a_{22}^{(1)}} = \dfrac{-4}{4} = -1$，且

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

注意：

分解的理论由***Gauss***消元得出，因此分解能够进行的条件与***Gauss***消元一样

1、**Doolittle（杜利特尔）分解**

L为单位下三角，U为上三角

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

比较第1行： $a_{1j} = u_{1j} \quad j = 1, \cdots, n \quad \Rightarrow u_{1j} = a_{1j}$

比较第1列： $a_{i1} = l_{i1} \mathrm{u}_{11} \quad i = 2, \cdots, n \quad \Rightarrow l_{i1} = \dfrac{a_{i1}}{u_{11}}$

比较第2行：$a_{2j} = l_{21}u_{1j} + u_{2j} \quad j = 2,\Lambda\ ,n \quad \Rightarrow u_{2j} = a_{1j} - l_{21}u_{1j}$

比较第2列：$a_{i2} = l_{i1}\mathrm{u}_{12} + l_{i2}\mathrm{u}_{22} \quad i = 3,\Lambda\ ,n \quad \Rightarrow l_{i2} = \dfrac{a_{i2} - l_{i1}\mathrm{u}_{12}}{u_{22}}$

$$\begin{pmatrix} a_{11} & a_{12} & \mathrm{K} & a_{1n} \\ a_{21} & a_{22} & \Lambda & a_{2n} \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} \\ a_{n1} & a_{n2} & \Lambda & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \\ l_{n1} & l_{n2} & \Lambda & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \mathrm{K} & u_{1n} \\ & u_{22} & \Lambda & u_{2n} \\ & & \mathrm{O} & \mathrm{M} \\ & & & u_{nn} \end{pmatrix}$$

比较第k行：$a_{kj} = \sum_{r=1}^{k-1} l_{kr}u_{rj} + u_{kj} \quad j = k,\Lambda\ ,n \quad \Rightarrow u_{kj} = a_{kj} - \sum_{r=1}^{k-1} l_{kr}u_{rj}$

K-1次

比较第k列：

$$a_{ik} = \sum_{r=1}^{k-1} l_{ir}u_{rk} + l_{ik}u_{kk} \quad i = k+1,\Lambda\ ,n \quad \Rightarrow l_{ik} = \frac{a_{ik} - \sum_{r=1}^{k-1} l_{ir}u_{rk}}{u_{kk}}$$

K-1＋1次

分解过程完毕，加上两次反代过程

$$y_i = b_i - \sum_{j=1}^{i-1} l_{ij} y_j \quad , \ i = 1,\Lambda\ ,n$$

$$x_i = \frac{y_i - \sum\limits_{j=i+1}^{n} u_{ij} x_j}{u_{ii}} \ , \ i = n,\Lambda\ ,1$$

总运算量为：

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

存储在矩阵的原来位置，且不影响计算

$$\begin{pmatrix} u_{11} & u_{12} & \mathrm{K} & u_{1n} \\ l_{21} & u_{22} & \Lambda & u_{2n} \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} \\ l_{n1} & l_{n2} & \Lambda & u_{nn} \end{pmatrix}$$

2、Crout 分解

L为下三角，U为单位上三角

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

比较第k列：$a_{ik} = \sum_{r=1}^{k-1} l_{ir}u_{rk} + l_{ik} \quad i = k,\Lambda\ ,n \quad \Rightarrow l_{ik} = a_{ik} - \sum_{r=1}^{k-1} l_{ir}u_{rk}$

比较第k行：

$$a_{kj} = \sum_{r=1}^{k-1} l_{kr}u_{rj} + l_{kk}u_{kj} \quad j = k+1,\Lambda\ ,n \quad \Rightarrow u_{kj} = \frac{a_{kj} - \sum_{r=1}^{k-1} l_{kr}u_{rj}}{l_{kk}}$$

两次反代过程

$$y_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij}y_j}{l_{ii}} \quad , \ i = 1,\Lambda\ ,n$$

$$x_i = y_i - \sum_{j=i+1}^{n} u_{ij}x_j \quad , \ i = n,\Lambda\ ,1$$

下面，我们对一下特殊的矩阵，提出一些特定的分解法

3. 三对角阵的追赶法

$$
\begin{pmatrix}
a_1 & b_1 & & & \\
c_2 & a_2 & & & \\
 & \ddots & \ddots & & b_{n-1} \\
 & & & c_n & a_n
\end{pmatrix}
=
\begin{pmatrix}
\alpha_1 & & & & \\
\gamma_2 & \alpha_2 & & & \\
 & \ddots & \ddots & & \\
 & & & \gamma_n & \alpha_n
\end{pmatrix}
\begin{pmatrix}
1 & \beta_1 & & & \\
 & 1 & \ddots & & \\
 & & \ddots & & \beta_{n-1} \\
 & & & & 1
\end{pmatrix}
$$

$$
\begin{cases}
\gamma_i = c_i \quad , \quad i = 2, \cdots, n \\
\alpha_i = a_i - c_i \beta_{i-1} \quad , \quad i = 1, \cdots, n \quad , \quad c_1 = 0 \\
\beta_i = b_i / \alpha_i \quad , \quad i = 1, \cdots, n
\end{cases}
$$

$$
\begin{cases}
y_i = (f_i - c_i y_{i-1}) / \alpha_i \quad , \quad i = 1, \cdots, n \\
x_i = y_i - \beta_i x_{i+1} \quad , \quad i = n, \cdots, 1 \quad (\beta_n = 0)
\end{cases}
$$

所以，有计算过程如下：

$$\begin{cases} \alpha_i = a_i - c_i\beta_{i-1} \\ \beta_i = b_i / \alpha_i \\ y_i = (f_i - c_i y_{i-1}) / \alpha_i \end{cases} \qquad i = 1, \Lambda\ , n$$

$$x_k = y_k - \beta_k x_{k+1} \quad , \qquad k = n, \Lambda\ , 1$$

3. 对称正定阵的LDL^T分解 (平方根法，也称Cholesky分解法)

若A对称正定，则有下三角阵L，使得

$$A = LL^T$$

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

所以有：

$$\begin{cases} l_{kk} = (a_{kk} - \sum_{r=1}^{k-1} l_{kr}^{\ 2})^{1/2} \\ l_{ik} = (a_{ik} - \sum_{r=1}^{k-1} l_{ik} l_{kr}) \Big/ l_{kk}, \quad i = k+1, \cdots, n \end{cases}$$

称为**平方根法**，
因为带了开方运算，
因此不常用

又
$$\begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & & & \\ l'_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l'_{n1} & l'_{n2} & \cdots & 1 \end{pmatrix}}_{\tilde{L}} \underbrace{\begin{pmatrix} l_{11} & & & \\ & l_{22} & & \\ & & \ddots & \\ & & & l_{nn} \end{pmatrix}}_{\tilde{D}}$$

则有

$$A = \tilde{L}\tilde{D}\tilde{D}^T\tilde{L}^T = LDL^T \qquad L = \begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{pmatrix} \qquad D = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

$$
\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{pmatrix}
=
\begin{pmatrix}
1 & & & \\
l_{21} & 1 & & \\
\vdots & \vdots & \ddots & \\
l_{n1} & l_{n2} & \cdots & 1
\end{pmatrix}
\begin{pmatrix}
d_1 & d_1 l_{21} & \cdots & d_1 l_{n1} \\
 & d_2 & \cdots & d_2 l_{n2} \\
 & & \ddots & \vdots \\
 & & & d_n
\end{pmatrix}
$$

比较等号两边后，有

$$
\begin{cases}
d_k = a_{kk} - \sum_{r=1}^{k-1} l_{kr}{}^2 d_r \\
l_{ik} = \left. \left(a_{ik} - \sum_{r=1}^{k-1} l_{ik} l_{kr} d_r\right) \middle/ d_k \right. , i = k+1, \cdots, n
\end{cases}
$$

改进的**平方根法**，

$$A = \begin{pmatrix} -6 & 3 & 2 \\ 3 & 5 & 1 \\ 2 & 1 & 6 \end{pmatrix} = \begin{pmatrix} 1 & & \\ l_{21} & 1 & \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{pmatrix} \begin{pmatrix} 1 & l_{21} & l_{31} \\ & 1 & l_{32} \\ & & 1 \end{pmatrix}$$

$$d_1 = a_{11} = -6 \qquad l_{21} = a_{21} / d_1 = -1/2 \qquad l_{31} = a_{31} / d_1 = -1/3$$

$$d_2 = a_{22} - l_{21} d_1 l_{21} = 13/2 \qquad l_{32} = (a_{32} - d_1 l_{31} l_{21}) / d_2 = 4/13$$

$$d_3 = a_{33} - l_{31} d_1 l_{31} - l_{32} d_2 l_{32} = 6\frac{2}{39}$$

为了提高数值稳定性，可考虑列主元三角分解法，设已完成 ***A=LU***的***k-1***步分解计算，矩阵分解成

$$
\begin{pmatrix}
u_{11} & u_{12} & \Lambda & u_{1k} & \Lambda & u_{1n} \\
l_{21} & u_{22} & \Lambda & u_{2k} & \Lambda & u_{2n} \\
M & M & O & M & & M \\
l_{k-11} & l_{k-12} & \Lambda & u_{k-1k} & \Lambda & u_{k-1n} \\
l_{k1} & l_{k2} & \Lambda & a_{kk} & \Lambda & a_{kn} \\
M & M & & M & & M \\
l_{n1} & l_{n2} & \Lambda & a_{nk} & \Lambda & a_{nn}
\end{pmatrix}
$$

设 $a_{ik} - \sum_{j=1}^{k-1} l_{ij} u_{jk} = \max_{k \le t \le n} (a_{tk} - \sum_{j=1}^{k-1} l_{tj} u_{jk})$ 令 $r_k \leftrightarrow r_i$

相当于取 $u_{kk} = a_{ik} - \sum_{j=1}^{k-1} l_{ij} u_{jk}$ 为第k步分解的主元素.

但要注意方程组的常数项也要相应变换.

3.3’ 向量的范数

定义1 设$x=(x_1,x_2,\cdots,x_n)^T, y=(y_1,y_2,\cdots,y_n)^T\in R^n$,

记 $$(x,y)=x^T\cdot y=y^T\cdot x=\sum_{i=1}^{n}x_i y_i \qquad (3.5.1)$$

为向量x与y的内积。记非负实数

$$\|x\|_2=\sqrt{(x,x)}=\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}} \qquad (3.5.2)$$

称为向量x的欧氏范数。

设 $x, y \in R^n, k$ 为实数。

(1) 非负性 $\|x\|_2 \geq 0$ 且 $\|x\|_2 = 0$ 当且仅当*x=0*时成立。

(2) 齐次性 $\|k \cdot x\|_2 = |k| \cdot \|x\|_2$

(3) 柯西-施瓦茨不等式 $|(x, y)| \leq \|x\|_2 \|y\|_2$

等式当且仅当*x*与*y*线性相关时成立

(4) 三角不等式 $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$

定义2 设$\|x\|$ 是R^n 上定义的一个实值函数，如果对任意的 $x, y \in R^n$，$k \in R$, 满足

(1) 非负性 $\|x\| \geq 0$ 且 $\|x\| = 0$ 当且仅当$x=0$时成立。

(2) 齐次性 $\|k \cdot x\| = |k| \cdot \|x\|$

(3) 三角不等式 $\|x + y\| \leq \|x\| + \|y\|$

则称$\|x\|$ 是向量x的一个范数（或模）。

由(3)可推出 $\big|\|x\| - \|y\|\big| \leq \|x - y\|$

几种常用范数：

(1) ∞-范数，也称为最大范数或切比雪夫范数

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

(2) 1-范数，也称为绝对范数 $\|x\|_1 = \sum_{i=1}^{n} |x_i|$

(3) 2-范数，也称为欧几里得范数 $\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}$

(4) p-范数 $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$ ，其中 $p \in [1,+\infty)$

上述前三种范数都是p-范数的特殊情况，

并且满足下列关系：

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

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

$$\frac{1}{\sqrt{n}}\|x\|_1 \le \|x\|_2 \le \|x\|_1$$

引理（向量范数的连续性） 设非负函数$\|x\|$为R^n上的任一向量范数，则$\|x\|$是x的分量$x_1, x_2, \ldots, x_n$的连续函数。

定理1（向量范数的等价性定理）

设$\|x\|_s, \|x\|_t$为R^n上的任意两个向量范数，则存在常数 $C_1, C_2 > 0,$ 使得对一切$x \in R^n$的分量恒有

$$C_1\|x\|_s \le \|x\|_t \le C_2\|x\|_s$$

注1：此定理不能推广到无穷维空间

注2：对于某一个向量x来说，如果它的某一种范数小（或大），那么它的任一种范数也不会很大（或很小）

定义3 设 $\left\{x^{(k)}\right\}$ 为R^n中一向量序列，$x^* \in R^n$, 如果

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

则称 x^k 依范数$\|\cdot\|$收敛于 x^* 。

注2：如果在某种范数意义下向量序列收敛，则在任何一种范数意义下该向量序列也收敛。

注2：一般按计算的需要采用不同的范数，把向量序列 $\left\{x^{(k)}\right\}$收敛于向量 x^k 记为 $\lim\limits_{k\to\infty} x^{(k)} = x^*$

而不强调是在哪种范数意义下收敛。

定理2 设A为$m \times n$阶矩阵，其列向量为线性无关的，如果$\|\cdot\|$是R^m中范数，则

$$N(x)=\|Ax\|, x \in R^n$$

是R^n中的一种范数。

3.3 矩阵范数和条件数

定义 设 ‖•‖ 是以n阶方阵为变量的实值函数, 且满足条件:

(1)**非负性:** ‖**A**‖≥0 ,且‖**A**‖=0当且仅当**A=0**

(2)**齐次性:** ‖α**A**‖=|α |‖**A**‖, α∈R

(3)三角不等式: ‖**A+B**‖≤‖**A**‖+‖**B**‖

(4)相容性: ‖**AB**‖≤‖**A**‖‖**B**‖

则称 ‖ **A** ‖ 为矩阵**A**的范数.

定义: 设‖•‖是R^n一种向量范数, $A \in R^{n\times n}$, 记

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

称之为由向量范数派生的**矩阵算子范数**.

对应于3种常见的向量范数，有3种矩阵范数

$$\|A\|_1 = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|$$ 列和的最大值

$$\|A\|_\infty = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$$ 行和的最大值

$$\|A\|_2 = \sqrt{\lambda_1}$$ λ_1 是$\boldsymbol{A^TA}$的最大特征值，也称为谱范数

矩阵范数的一些性质：

① $\|A\| \ge 0,\&,\|A\| = 0 \Leftrightarrow A = 0$

② $\|\lambda A\| = |\lambda| \cdot \|A\|, \forall \lambda \in R$

③ $\|A+B\| \le \|A\| + \|B\| \quad , \quad \forall \ A, B \in R^n$

④ $\|AB\| \le \|A\| \cdot \|B\| \quad , \quad \forall \ A, B \in R^n$

⑤ $\|Ax\| \le \|A\| \cdot \|x\|, \forall x \in R^n$

定理： 若 λ 为 A 的特征值，则$|\lambda| \leq \|A\|$

证： $\lambda \cdot x = A \cdot x$ x为A的特征值

$$\|\lambda \cdot x\| = \|A \cdot x\|$$

$$\|\lambda \cdot x\| = |\lambda| \cdot \|x\| = \|A \cdot x\| \leq \|A\| \cdot \|x\|$$

$$\|A\| \geq \frac{\|A \cdot x\|}{\|x\|} = |\lambda|$$ 证毕

定义5.2： 谱半径 $\rho(A) = \max\limits_{1 \leq r \leq n} |\lambda_r|$

易知： $\rho(A) \leq \|A\|$

3.4 条件数和病态矩阵

定义5.3：(条件数) $Cond_p(A)=\|A\|_p \cdot \|A^{-1}\|_p$　　$\|\bullet\|_p$ 表示某种范数

设 $Ax=b$，A 引入误差δA 后，解引入误差 δx ，则

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

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

$$\delta x = -(A+\delta A)^{-1}\delta A\cdot x$$

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

$$= \left\|(I+A^{-1}\delta A)^{-1}A^{-1}\delta A\right\|$$

$$\left(A(I+A^{-1}\delta A)\right)^{-1} = (A+\delta A)^{-1}$$

$$\le \left\|(I + A^{-1}\delta A)^{-1}\right\| \cdot \left\|A^{-1}\delta A\right\|$$

$$\le \left\|(I + A^{-1}\delta A)^{-1}\right\| \cdot \left\|A^{-1}\right\| \cdot \left\|\delta A\right\|$$

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

因为: $$D = (I + B)^{-1} \quad \Rightarrow \quad 1 = \|I\| = \|(I + B)D\| = \|D + BD\|$$

$$\ge \|D\| - \|B\| \cdot \|D\| = \|D\|(1 - \|B\|)$$

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

$$\le \frac{\|A^{-1}\| \cdot \|\delta A\|}{1-\|A^{-1}\| \cdot \|\delta A\|} = \frac{\|A^{-1}\| \cdot \|A\| \frac{\|\delta A\|}{\|A\|}}{1-\|A^{-1}\| \cdot \|A\| \frac{\|\delta A\|}{\|A\|}} \approx \underbrace{\|A^{-1}\| \cdot \|A\|}_{\text{条件数}} \frac{\|\delta A\|}{\|A\|}$$

很小

条件数表示了对误差的放大率

同样，类似有

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

$$\Rightarrow A\delta x=\delta b$$

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

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$cond\,(A)_2$。

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

解：考察 A 的特征根

$$\det(\lambda I - A) = 0 \quad \Rightarrow \quad \begin{aligned} \lambda_1 &= 1.980050504 \\ \lambda_2 &= -0.000050504 \end{aligned}$$

$$cond\,(A)_2 = \left|\frac{\lambda_1}{\lambda_2}\right| \approx 39206 >> 1$$

为对称矩阵

测试病态程度：

给 $\vec{b}$ 一个扰动 $\delta\vec{b} = \begin{pmatrix} -0.97\times 10^{-4} \\ 0.106\times 10^{-3} \end{pmatrix}$，其相对误差为

$$\frac{\|\delta\vec{b}\|_2}{\|\vec{b}\|_2} \approx 0.513\times 10^{-4} < 0.01\%$$

此时精确解为 $\vec{x}^* = \begin{pmatrix} 3 \\ -1.0203 \end{pmatrix}$

$$\delta\vec{x} = \vec{x}^* - \vec{x} = \begin{pmatrix} 2 \\ -2.0203 \end{pmatrix} \Rightarrow \frac{\|\delta\vec{x}\|_2}{\|\vec{x}\|_2} \approx 2.0102 > 200\%$$

- 本课件是在中国科技大学张瑞老师课件基础上修改而成的，在此表示感谢！