第五章 解线性方程组的直接方法

内容提要

5.1 引言与预备知识

5.2 高斯消去法

5.3 高斯列主元消去法

5.4 矩阵三角分解法

5.5 向量与矩阵的范数

（5.6 误差分析）

5.1 引言

线性方程组（标量方程组形式）

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

矩阵形式（矩阵向量形式）

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{bmatrix}$$

简记为

$$Ax = b$$

通常考虑*m*=*n*情况

上万维，零元素很多，
非零元素很少。

Existence and uniqueness of the solution?

$$solution\ \exists! \Leftrightarrow |A| \neq 0$$

How to get the solution?

经过有限步算术运算直接求得精确解（在没有舍入误差的情况下），但实际上机器总存在舍入误差，因此求得的是近似解。

Cramer rule: $x_i = \cancel{D_i / D}$

Computation cost: (n+1)!

通过某种迭代系统（公式）求得近似解，优点：编程简单 缺点：存在收敛性和收敛速度问题

Coefficient matrix A

非零元素较多，零元素较少

低维稠密阵 small dense matrix	高维稀疏阵 large sparse matrix
Gaussian elimination 列/行/完全主元素(pivoting)消去法 **Square root/improved square root methods**平方根法 追赶法	**Jaccobi iteration** **Gauss-Seidel iteration** **SOR**
直接法	迭代法

Gauss elimination 1810

关于线性方程组的数值解法一般有两类。

1、直接解法：经过有限次的算术运算，可求得方程组精确解的方法（若计算过程中没有舍入误差）。但实际计算中由于舍入误差的存在和影响，这种方法也只能求得线性方程组的近似解。本章主要研究此类问题的解法。

2、迭代法：用某种极限过程去逐步逼近线性方程组精确解的方法。迭代法具有需要计算机的存储单元较少、程序设计简单、原始系数矩阵在计算过程中始终不变等优点。

例5-1 用消去法解方程组

$$\begin{cases} x_1 + x_2 + x_3 = 6 \\ 4x_2 - x_3 = 5 \\ 2x_1 - 2x_2 + x_3 = 1 \end{cases}$$

解:

$$(A|b)=\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 2 & -2 & 1 & 1 \end{array}\right] \qquad (消去第一列)$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 0 & -4 & -1 & -11 \end{array}\right] \qquad (消去第二列)$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 0 & 0 & -2 & -6 \end{array}\right] \qquad (回代、求解)$$

解得

$$x^* = (1,2,3)^{\mathrm{T}}$$

行梯形row-echelon form (REF)

5.2 高斯消去法

(1)

$$\left[A^{(1)}:b^{(1)}\right]=\left[\begin{array}{cccc|c} \boxed{a_{11}^{(1)}} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{array}\right] \xrightarrow[(i=2,3,\cdots,n)]{m_{i1}=\frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \quad (a_{11}^{(1)}\neq 0)}$$

$$\left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right]=\left[A^{(2)}:b^{(2)}\right]$$

$$\begin{cases} a_{ij}^{(2)}=a_{ij}^{(1)}-m_{i1}a_{1j}^{(1)} & (i=2,\cdots,n;\ j=2,\cdots,n) \\ b_i^{(2)}=b_i^{(1)}-m_{i1}b_1^{(1)} & (i=2,\cdots,n) \end{cases}$$

(2) $$\left[A^{(2)}:b^{(2)}\right]=\left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & \boxed{a_{22}^{(2)}} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right] \xrightarrow[(i=3,\cdots,n)]{m_{i2}=\frac{a_{i2}^{(2)}}{a_{22}^{(2)}} \quad (a_{22}^{(2)}\neq 0)}$$

$$\left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(3)} & b_n^{(3)} \end{array}\right]=\left[A^{(3)}:b^{(3)}\right]$$

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

$$
(k) \qquad \left[A^{(k)} : b^{(k)}\right] = \left[\begin{array}{cccccc|c}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\
 & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\
 & & \ddots & \vdots & & \vdots & \vdots \\
 & & & \boxed{a_{kk}^{(k)}} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\
 & & & \vdots & & \vdots & \vdots \\
 & & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & b_n^{(k)}
\end{array}\right]
\xrightarrow[(i=k+1,\cdots,n)]{m_{ik}=\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \quad (a_{kk}^{(k)}\neq 0)}
$$

$$
\left[\begin{array}{cccccc|c}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\
 & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\
 & & \ddots & \vdots & & \vdots & \vdots \\
 & & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\
 & & & \vdots & & \vdots & \vdots \\
 & & & 0 & \cdots & a_{nn}^{(k+1)} & b_n^{(k+1)}
\end{array}\right] = \left[A^{(k+1)} : b^{(k+1)}\right]
$$

$$
\begin{cases}
a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik}a_{kj}^{(k)} & (i = k+1,\cdots,n;\ j = k+1,\cdots,n) \\
b_i^{(k+1)} = b_i^{(k)} - m_{ik}b_k^{(k)} & (i = k+1,\cdots,n)
\end{cases}
$$

(n)　　继续上述过程, 直到完成消元计算。

最后得到与原方程组等价的简单方程组 $A^{(n)}x = b^{(n)}$.

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

对应方程组为　$A^{(n)}x = b^{(n)}$

在求解三角方程组, 得

$$\begin{cases} x_n = b_n^{(n)} \Big/ a_{nn}^{(n)} \\ x_k = (b_k^{(k)} - \sum\limits_{j=k+1}^{n} a_{kj}^{(k)} x_j) \Big/ a_{kk}^{(k)} \end{cases} \qquad (**)$$

elementary transform

高斯消去法的条件

定理5-1　设 $Ax=b$, 其中$A\in R^{n\times n}$

(1)　如果$a_{kk}^{(k)}\neq 0\ (k=1,2,\cdots,n)$，则可通过高斯消去法将 $Ax=b$ 约化为等价的三角方程组(*)，且计算公式为(**)　。

(2)　如果A为非奇异矩阵，则可通过高斯消去法（及交换两行的初等变换）将$Ax=b$ 约化为(*)。

定理5-2　约化的主元素$a_{ii}^{(i)}\neq 0\ (i=1,2,\cdots,k)$ 的充要条件是矩阵A 的顺序主子式$D_i\neq 0\ (i=1,2,\cdots,k)$. 即

$$D_1=a_{11}\neq 0,$$

$$D_i=\begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} \end{vmatrix}\neq 0\quad (i=1,2,\cdots,k)$$

sequential principal minor

Gauss消去法求解*n*元线性方程组的乘除运算量是：

$$\frac{1}{3}(n^3 + 3n^2 - n)$$

n=20时, Gauss消去法需3060次乘除法运算。

n=40时, Gauss消去法约需2万3千次乘除法运算。

n=80时, Gauss消去法约需17万7千次乘除法运算。

n=160时, Gauss消去法约需139万次乘除法运算。

n=320时, Gauss消去法约需1102万次乘除法运算。

n=640时, Gauss消去法约需8800万次乘除法运算。

n=1280时, Gauss消去法约需7亿次乘除法运算。

注：Cramer法则是一种不实用的直接法。

5.3 高斯主元素消去法

由高斯消去法知道：

1）在消元过程中可能出现 $a_{kk}^{(k)}=0$ 的情况，这时消去法将无法进行；

2）即使主元素 $a_{kk}^{(k)}\neq 0$ 但很小时，用其作除数，会导致其它元素数量级的严重增长和舍入误差的扩散，最后也使得计算解不可靠。

对同一个数值问题，用不同的计算方法，得到的结果的精度大不一样。

一个计算方法，如果用此方法的计算过程中舍入误差得到控制，对计算结果影响较小，称此方法为数值稳定的。

否则，如果用此计算方法的计算过程中舍入误差增长迅速，计算结果受舍入误差影响较大称此方法为数值不稳定。

因此，我们解数值问题时，应选择和使用数值稳定的计算方法。否则，如果使用数值不稳定的计算方法去解数值计算问题，就可能导致计算失败。

例　解线性方程组（有效数字最多4位）

$$\begin{cases} 0.0001x_1 + 1.00x_2 = 1.00 \\ 1.00x_1 + 1.00x_2 = 2.00 \end{cases}$$

(用Cramer法则可得精确解x_1^*=1.0001 , x_2^*=0.9999)

解　用顺序Gauss消去法, 消元得

$$\begin{cases} 0.0001x_1 + 1.00x_2 = 1.00 \\ -1000x_2 = -1000 \end{cases}$$

回代得解: x_2=1.00 , x_1=0.00

若将方程组改写成

$$\begin{cases} 1.00x_1 + 1.00x_2 = 2.00 \\ 0.0001x_1 + 1.00x_2 = 1.00 \end{cases}$$

用顺序Gauss消去法, 消元得

$$\begin{cases} 1.00x_1 + 1.00x_2 = 2.00 \\ 1.00x_2 = 1.00 \end{cases}$$

回代得解: x_2=1.00 , x_1=1.00

为了提高计算的数值稳定性, 在消元过程中采用选择主元的方法. 常采用的是列主元消去法.

例5-2 用高斯列主元法解线性方程组

列主元消去法

$$
\begin{bmatrix} 0 & 2 & 0 & 1 \\ 2 & 2 & 3 & 2 \\ 4 & -3 & 0 & 1 \\ 6 & 1 & -6 & -5 \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
=
\begin{bmatrix} 0 \\ -2 \\ -7 \\ 6 \end{bmatrix}
$$

解:

$$
[A|b] = \left[\begin{array}{cccc|c} 0 & 2 & 0 & 1 & 0 \\ 2 & 2 & 3 & 2 & -2 \\ 4 & -3 & 0 & 1 & -7 \\ \boxed{6} & 1 & -6 & -5 & 6 \end{array}\right]
\to
\left[\begin{array}{cccc|c} 6 & 1 & -6 & -5 & 6 \\ 2 & 2 & 3 & 2 & -2 \\ 4 & -3 & 0 & 1 & -7 \\ 0 & 2 & 0 & 1 & 0 \end{array}\right]
\to
\left[\begin{array}{cccc|c} 6 & 1 & -6 & -5 & 6 \\ 0 & \frac{5}{3} & 5 & \frac{11}{3} & -4 \\ 0 & \boxed{-\frac{11}{3}} & 4 & \frac{13}{3} & -11 \\ 0 & 2 & 0 & 1 & 0 \end{array}\right]
$$

$$
\to
\left[\begin{array}{cccc|c} 6 & 1 & -6 & -5 & 6 \\ 0 & -\frac{11}{3} & 4 & \frac{13}{3} & -11 \\ 0 & 0 & \frac{75}{11} & \frac{186}{33} & -9 \\ 0 & 0 & \frac{24}{11} & \frac{111}{33} & -6 \end{array}\right]
\to
\left[\begin{array}{cccc|c} 6 & 1 & -6 & -5 & 6 \\ 0 & -\frac{11}{3} & 4 & \frac{13}{3} & -11 \\ 0 & 0 & \frac{75}{11} & \frac{62}{11} & -9 \\ 0 & 0 & 0 & \frac{429}{275} & -\frac{78}{25} \end{array}\right]
$$

故

$$
\begin{cases} x_1 = -\frac{1}{2} \\ x_2 = 1 \\ x_3 = \frac{1}{3} \\ x_4 = -2 \end{cases}
$$

LU 1940 (130年后)

5.4 矩阵三角分解法

$Ax=b$是线性方程组，A是$n\times n$方阵，并设A的各阶顺序主子式不为零。令 $A^{(1)}=A$，当高斯消元法进行第一步后，相当于用一个初等矩阵左乘$A^{(1)}$。不难看出，这个初等矩阵为

elementary matrix

$$L_1=\begin{bmatrix} 1 & & & & \\ -m_{21} & 1 & & & \\ -m_{31} & & 1 & & \\ \vdots & & & \ddots & \\ -m_{n1} & & & & 1 \end{bmatrix}$$

$$L_1A^{(1)}=A^{(2)},\quad L_1b^{(1)}=b^{(2)}$$

Gauss elimination 1810

一般地，

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

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

重复这个过程，最后得到

$$\left.\begin{aligned} L_{n-1} \cdots L_2 L_1 A^{(1)} = A^{(n)} \\ L_{n-1} \cdots L_2 L_1 b^{(1)} = b^{(n)} \end{aligned}\right\}$$

将上三角矩阵$A^{(n)}$记为U, 得到

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

（因$A^{(1)} = A$且$L_{n-1}\cdots L_2L_1A^{(1)} = A^{(n)} = U$）其中，

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

为单位下三角矩阵。

这就是说，高斯消去法实质上产生了一个将$\boldsymbol{A}$分解为两个三角形矩阵相乘的因式分解，于是我们得到如下重要定理。

定理5-3 (矩阵的LU分解) 设A 为n 维矩阵，如果A 的顺序主子式 $D_i \neq 0 \ (i = 1, 2, \cdots, n-1)$，则$A$ 可分解为一个单位下三角矩阵L 和一个上三角矩阵U 的乘积，且这种分解是唯一的。

证明 只证唯一性, 设有两种分解

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U} = \bar{L}\bar{U}$$

则有 $$\underbrace{\bar{L}^{-1}L}_{\text{单位下三角阵}} = \underbrace{\bar{U}U^{-1}}_{\text{上三角形矩阵}} = I$$

所以得 $L = \bar{L},\ U = \bar{U}.$

证明不做要求

当***A***进行***LU***分解后，***Ax=b***就容易解了.

即***Ax=b***等价于

$$\begin{cases} Ly = b \\ Ux = y \end{cases}$$

验证：$Ax = LUx = L(Ux) = Ly = b$

具体计算公式为

forward substitution

$$(1) \quad \begin{cases} y_1 = b_1 \\ y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k \end{cases} \qquad (i = 2, 3, \cdots, n)$$

back substitution

$$(2) \quad \begin{cases} x_n = y_n / u_{nn} \\ x_i = (y_i - \sum_{k=i+1}^{n} u_{ik} x_k) / u_{ii} \end{cases} \qquad (i = n-1, n-2, \cdots, 1)$$

Example 3.21. Use Gaussian elimination to construct the triangular factorization of the matrix

$$A = \begin{bmatrix} 4 & 3 & -1 \\ -2 & -4 & 5 \\ 1 & 2 & 6 \end{bmatrix}.$$

The matrix L will be constructed from an identity matrix placed at the left. For each row operation used to construct the upper-triangular matrix, the multipliers m_{ij} will be put in their proper places at the left. Start with

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 & -1 \\ -2 & -4 & 5 \\ 1 & 2 & 6 \end{bmatrix}.$$

Row 1 is used to eliminate the elements of A in column 1 below a_{11}. The multiples $m_{21} = -0.5$ and $m_{31} = 0.25$ of row 1 are subtracted from rows 2 and 3, respectively. These multipliers are put in the matrix at the left and the result is

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -0.5 & 1 & 0 \\ 0.25 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 & -1 \\ 0 & -2.5 & 4.5 \\ 0 & 1.25 & 6.25 \end{bmatrix}.$$

课堂演算

Row 1 is used to eliminate the elements of $\boldsymbol{A}$ in column 1 below a_{11}. The multiples $m_{21} = -0.5$ and $m_{31} = 0.25$ of row 1 are subtracted from rows 2 and 3, respectively. These multipliers are put in the matrix at the left and the result is

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ -0.5 & 1 & 0 \\ 0.25 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 & -1 \\ 0 & -2.5 & 4.5 \\ 0 & 1.25 & 6.25 \end{bmatrix}.$$

Row 2 is used to eliminate the elements in column 2 below the diagonal of the second factor of $\boldsymbol{A}$ in the above product. The multiple $m_{32} = -0.5$ of the second row is subtracted from row 3, and the multiplier is entered in the matrix at the left and we have the desired triangular factorization of $\boldsymbol{A}$.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ -0.5 & 1 & 0 \\ 0.25 & -0.5 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 & -1 \\ 0 & -2.5 & 4.5 \\ 0 & 0 & 8.5 \end{bmatrix}. \tag{8}$$

■

计算公式：

1） $u_{1i} = a_{1i} (i = 1,2,\cdots,n), l_{i1} = {}^{a_{i1}}/_{u_{11}}, i = 2,3,\cdots,n$

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

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

课堂演算

例5-3 用矩阵直接三角分解法解

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

解: 用分解公式计算得

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & -5 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -4 \\ 0 & 0 & -24 \end{bmatrix} = LU$$

求解

$$Ly = (14, 18, 20)^{\mathrm{T}},\ 得 y = (14, -10, -72)^{\mathrm{T}}$$

$$Ux = (14, -10, -72)^{\mathrm{T}},\ 得 x = (1, 2, 3)^{\mathrm{T}}$$

课堂演算

了解即可，可略

例5-4 求A 的LU分解，并利用分解结果求A^{-1}.

$$A=\begin{bmatrix}1 & 0 & 2\\ 1 & 2 & 3\\ 2 & 6 & 9\end{bmatrix}$$

解: A 的LU分解为 $L=\begin{bmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 2 & 3 & 1\end{bmatrix}$, $U=\begin{bmatrix}1 & 0 & 2\\ 0 & 2 & 1\\ 0 & 0 & 2\end{bmatrix}$

从而 $L^{-1}=\begin{bmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ 1 & -3 & 1\end{bmatrix}$, $U^{-1}=\begin{bmatrix}1 & 0 & -1\\ 0 & 1/2 & -1/4\\ 0 & 0 & 1/2\end{bmatrix}$

故 $$A^{-1}=U^{-1}L^{-1}=\begin{bmatrix}1 & 0 & -1\\ 0 & 1/2 & -1/4\\ 0 & 0 & 1/2\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ 1 & -3 & 1\end{bmatrix}=\begin{bmatrix}0 & 3 & -1\\ -3/4 & 5/4 & -1/4\\ 1/2 & -3/2 & 1/2\end{bmatrix}$$

LU分解：存储在矩阵的原来位置，
且不影响计算

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

称矩阵的三角分解 $A = LU$ 为杜利特尔(**Doolittle**)分解。其中

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

若L为下三角阵, U为单位上三角阵。称矩阵的这种分解为**Crout**分解。其中

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

Example 3.25. Use the MATLAB command `[L,U,P]=lu(A)` on the matrix A in Example 3.22. Verify that $A \doteq P^{-1}LU$ (equivalent to showing that $PA = LU$).

```
>>A=[1 2 6 ;4 8 -1;-2 3 -5];
>>[L,U,P]=lu(A)
```

```
L=
     1.0000   0        0
    -0.5000   1.0000   0
     0.2500   0        1.0000
```

```
U=
     4.0000   8.0000   -1.0000
     0        7.0000    4.5000
     0        0         6.2500
```

```
P=
   0  1  0
   0  0  1
   1  0  0
```

```
>>inv(P)*L*U
   1   2  6
   4   8 -1
  -2   3  5
```

课堂Matlab复现

平方根法

设A为对称正定矩阵, 则有唯一分解$A=LU$, 且$u_{kk}>0$.

而
$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix} = \underbrace{\begin{pmatrix} u_{11} & & & \\ & u_{22} & & \\ & & \ddots & \\ & & & u_{nn} \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} 1 & \frac{u_{12}}{u_{11}} & \cdots & \frac{u_{1n}}{u_{11}} \\ & 1 & \cdots & \frac{u_{2n}}{u_{22}} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}}_{M}$$

则有 $A=LDM=LDL^{\mathrm{T}}$

【因为$(LDM)^{\mathrm{T}}=M^{\mathrm{T}}DL^{\mathrm{T}}=LDM$

所以 $M=L^{\mathrm{T}}$ 】

其中, $G=LD^{\frac{1}{2}}$

令$$D=\begin{pmatrix}\sqrt{u_{11}} & & & \\ & \sqrt{u_{22}} & & \\ & & \ddots & \\ & & & \sqrt{u_{nn}}\end{pmatrix}\begin{pmatrix}\sqrt{u_{11}} & & & \\ & \sqrt{u_{22}} & & \\ & & \ddots & \\ & & & \sqrt{u_{nn}}\end{pmatrix}$$

$$=D^{\frac{1}{2}}D^{\frac{1}{2}}$$

则有 $A=LD^{\frac{1}{2}}D^{\frac{1}{2}}L^{\mathrm{T}}=(LD^{\frac{1}{2}})(LD^{\frac{1}{2}})^{\mathrm{T}}=GG^{\mathrm{T}}$

分解$\boldsymbol{A}=\boldsymbol{G}\boldsymbol{G}^{\mathrm{T}}$称为对称正定矩阵的楚列斯基**(Cholesky)**分解**.**

平方根法是求对称正定系数线性方程组的三角分解法, 对称正定矩阵的Cholesky分解的计算量和存贮量均约为一般矩阵**LU**分解的一半. 且Cholesky分解具有数值稳定性.

追赶法

在一些实际问题中， 例如解常微分方程边值问题，热传导方程以及船体数学放样中建立三次样条函数等，都会要求解系数矩阵为对角占优的三对角线方程组

$$\begin{bmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{bmatrix}$$

其中$|i-j|>1$时, $a_{ij}=0$, 且满足如下的对角占优条件:

(1) $|b_1|>|c_1|>0$, $|b_n|>|a_n|>0$

(2) $|b_i|\ge|a_i|+|c_i|$, $a_ic_i\ne0$, $i=2,3,\ldots,n-1$.

了解即可，可略

了解即可，可略

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

$$=\begin{bmatrix} \alpha_1 & & & \\ \gamma_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \gamma_n & \alpha_n \end{bmatrix}\begin{bmatrix} 1 & \beta_1 & & \\ & 1 & \ddots & \\ & & \ddots & \beta_{n-1} \\ & & & 1 \end{bmatrix}$$

其中

α_i、β_i、γ_i为待定系数。

了解即可，可略

计算公式

$$\beta_1 = c_1 / b_1$$

$$\beta_i = c_i / (b_i - a_i\beta_{i-1}) \qquad (i = 2,3,\cdots,n-1)$$

$$\alpha_1 = b_1, \quad \alpha_i = b_i - a_i\beta_{i-1} \qquad (i = 2,3,\cdots,n)$$

$$\gamma_i = a_i$$

解 $Ly = f$

$$y_1 = f_1 / b_1$$

$$y_i = (f_i - a_i y_{i-1}) / (b_i - a_i\beta_{i-1}) \qquad (i = 2,3,\cdots,n)$$

解 $Ux = y$

$$x_n = y_n$$

$$x_i = y_i - \beta_i x_{i+1} \qquad (i = n-1, n-2, \cdots, 2, 1)$$

计算系数 $\beta_1 \to \beta_2 \to \cdots \to \beta_{n-1} \to \beta_n$ 及 $y_1 \to y_2 \to \cdots \to y_{n-1} \to y_n$ 的过程称为追的过程。

计算方程组的解 $x_n \to x_{n-1} \to \cdots \to x_2 \to x_1$ 的过程称为赶的过程。

例5-5 用追赶法解方程组

了解即可，可略

$$\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}=\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

解: $$A=\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}=\begin{bmatrix} 2 & & & & \\ -1 & 3/2 & & & \\ & -1 & 4/3 & & \\ & & -1 & 5/4 & \\ & & & -1 & 6/5 \end{bmatrix}\begin{bmatrix} 1 & -1/2 & & & \\ & 1 & -2/3 & & \\ & & 1 & -3/4 & \\ & & & 1 & -4/5 \\ & & & & 1 \end{bmatrix}=\bar{L}\bar{U},$$

$$\bar{L}y=b\Rightarrow y=\begin{bmatrix} 1/2 \\ 1/3 \\ 1/4 \\ 1/5 \\ 1/6 \end{bmatrix},\ \bar{U}x=y\Rightarrow x=\begin{bmatrix} 5/6 \\ 2/3 \\ 1/2 \\ 1/3 \\ 1/6 \end{bmatrix}.$$

5.5 向量和矩阵的范数

定义5-1 (向量范数) x 和 y 是 R^n 中的任意向量，向量范数‖•‖是定义在 R^n 上的实值函数，它满足：

(1) $\|x\|\geq 0$，并且当且仅当 $x=0$ 时，$\|x\|=0$； 正定性，零零性

(2) $\|kx\|=|k|\,\|x\|$，k 是一个实数； 齐次性（比例性）

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

常使用的向量范数有三种，设 $x=(x_1, x_2, \ldots, x_n)^T$：

$$\|x\|_\infty=\max_{1\leq i\leq n}|x_i|$$ 向量的∞-范数

$$\|x\|_1=\sum_{i=1}^{n}|x_i|$$ 向量的1-范数

$$\|x\|_2=\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}}$$ 向量的2-范数

Report: 向量乘法及范数不等式？

定义5-2 (矩阵的范数) 如果矩阵$A \in R^{n\times n}$ 的某个非负的实值函数$N(A)=\|A\|$，满足条件

(1) $\|A\| \geq 0$ ($\|A\|=0 \Leftrightarrow A=0$) (正定条件)

(2) $\|cA\|=|c|\|A\|$ c为实数 (齐次条件)

(3) $\|A+B\| \leq \|A\|+\|B\|$ (三角不等式)

(4) $\|AB\| \leq \|A\|\|B\|$

Report: 矩阵其它乘法及范数不等式？

则称$N(A)$ 是$R^{n\times n}$的一个矩阵范数（或模）

常使用的矩阵范数有三种:

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

1列∞行

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

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^{\mathrm{T}}A)}$$ 矩阵的2-范数（谱范数）

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

可否$\|A\|_2 = \sqrt{\lambda_{\max}(AA^{\mathrm{T}})}$?

$\lambda_{\max}(AA^{\mathrm{T}}) = \lambda_{\max}(A^{\mathrm{T}}A)$?

矩阵的**F-**范数: $\|\boldsymbol{A}\|_{\mathrm{F}} = \sqrt{\sum_{i,j=1}^{n} a_{ij}^2}$

例　设矩阵 $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$

求矩阵$\boldsymbol{A}$的范数$\|\boldsymbol{A}\|_p$, p=1,2, ∞,F.

解　$\|\boldsymbol{A}\|_1=4$, $\|\boldsymbol{A}\|_\infty=5$, $\|\boldsymbol{A}\|_{\mathrm{F}} = \sqrt{15}$　　　1列∞行

$$A^{\mathrm{T}}A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 5 & 10 \end{pmatrix}$$

令 $\begin{vmatrix} 5-\lambda & 5 \\ 5 & 10-\lambda \end{vmatrix} = 0$, 得 $\lambda_1 = \dfrac{15+5\sqrt{5}}{2}$, $\lambda_2 = \dfrac{15-5\sqrt{5}}{2}$

所以 $\|A\|_2 = \sqrt{\dfrac{15+5\sqrt{5}}{2}}$

矩阵的范数与矩阵的特征值之间也有密切的联系。

设λ是矩阵$\boldsymbol{A}$的特征值，$\boldsymbol{x}$是对应的特征向量，则有 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$
利用向量和矩阵范数的相容性，则得

$$|\lambda|\|\boldsymbol{x}\|=\|\lambda\boldsymbol{x}\|=\|\boldsymbol{Ax}\|\leq\|\boldsymbol{A}\|\|\boldsymbol{x}\|$$

于是$|\lambda|\leq\|\boldsymbol{A}\|$。

设n维矩阵$\boldsymbol{A}$的n个特征值为$\lambda_1, \lambda_2, \ldots, \lambda_n$，则称

$$\rho(A)=\max_{1\leq i\leq n}|\lambda_i|$$

为矩阵**A**的谱半径. 对矩阵的任何一种相容范数都有

$$\rho(\boldsymbol{A})\leq\|\boldsymbol{A}\|$$

另外，$\forall\varepsilon>0$，$\exists$一种相容范数，使 $\|\boldsymbol{A}\|\leq\rho(\boldsymbol{A})+\varepsilon$

例5-6 设$A=\begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix}$，计算$A$的各种范数。

解：$\|A\|_1 = 6,$　　　1列∞行

$$\|A\|_\infty = 7$$

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

…

也Matlab计算：

命令行窗口

```
>> A=[1 -2; -3 4]

A =

     1    -2
    -3     4

>> norm(A,1)

ans =

     6

>> norm(A,inf)

ans =

     7

>> norm(A,2)

ans =

    5.4650

>> norm(A,'fro')

ans =

    5.4772
```

知识结构图

- 直接法解方程组
  - 高斯消去法
    - 高斯消去法
    - 列主元消去法
  - 矩阵三角分解法
    - *LU*分解
    - *LDL*ᵀ分解/平方根分解/Cholesky分解
  - 追赶法解三对角方程组
  - 向量和矩阵的范数
    - 定义
    - 常用范数
    - 范数的性质
  - （矩阵条件数及迭代改善法）

复习与思考题(无需提交)

P175: 3, 6, 7, 8

习题(需提交)

P176: 7, 8, 12