第2章 解线性方程组的直接法

本章讨论n元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \tag{2.1}$$

的直接解法. 方程组(2.1)的矩阵形式为

$$\boldsymbol{Ax}=\boldsymbol{b},$$

其中

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix}.$$

若矩阵$\boldsymbol{A}$非奇异, 即$\det(\boldsymbol{A})\neq 0$, 则方程组(2.1)有唯一解.

所谓直接解法是指，若不考虑计算过程中的舍入误差，经过有限次算术运算就能求出线性方程组的精确解的方法. 但由于实际计算中舍入误差的存在，用直接解法一般也只能求出方程组的近似解.

Cramer法则是一种不实用的直接法，下面介绍几种实用的直接法.

§1 Gauss消去法

Gauss消去法的基本思想是通过一种规则化的逐次加减消元计算，把一般线性方程组的求解转化为等价的上三角形方程组进行求解.

§1.1 顺序Gauss消去法

为了清楚起见，先看一个简单的例子. 考虑线性方程组：

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_1 - 3x_2 - 3x_3 = -1 \\ 4x_1 + 2x_2 + 2x_3 = 3 \end{cases}$$

消去后两个方程中的x_1得

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

再消去最后一个方程的x_2得

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ \quad\; -5x_2 - 2x_3 = -2 \\ \quad\quad\quad \dfrac{42}{5}x_3 = \dfrac{7}{5} \end{cases}$$

消元结束, 经过回代得解: $x_3 = \dfrac{1}{6}, x_2 = \dfrac{1}{3}, x_1 = \dfrac{1}{2}$

上述求解的消元过程可用矩阵表示为:

$$(\boldsymbol{A}, \boldsymbol{b})= \begin{pmatrix} 2 & 4 & -2 & 2 \\ 1 & -3 & -3 & -1 \\ 4 & 2 & 2 & 3 \end{pmatrix}$$

$$\underset{r_3-2r_1}{\overset{r_2-\frac{1}{2}r_1}{\sim}} \begin{pmatrix} 2 & 4 & -2 & 2 \\ 0 & -5 & -2 & -2 \\ 0 & -6 & 6 & -1 \end{pmatrix} \overset{r_3-\frac{6}{5}r_2}{\sim} \begin{pmatrix} 2 & 4 & -2 & 2 \\ 0 & -5 & -2 & -2 \\ 0 & 0 & \frac{42}{5} & \frac{7}{5} \end{pmatrix}$$

这是Gauss消去法的计算形式. 新的增广矩阵对应的线性方程组就是上三角形方程组, 可进行回代求解.

现在介绍求解线性方程组(2.1)的顺序Gauss消去法:

记 $\boldsymbol{A}^{(1)}=\boldsymbol{A}$, $\boldsymbol{b}^{(1)}=\boldsymbol{b}$, $a^{(1)}_{ij} = a_{ij}$, $b^{(1)}_{i} = b_i$,则线性方程组(2.1)的增广矩阵为

$$(\mathbf{A}^{(1)},\mathbf{b}^{(1)}) = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \ldots & a_{1n}^{(1)} & b_{1}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} & \ldots & a_{2n}^{(1)} & b_{2}^{(1)} \\ a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} & \ldots & a_{3n}^{(1)} & b_{3}^{(1)} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & a_{n3}^{(1)} & \ldots & a_{nn}^{(1)} & b_{n}^{(1)} \end{pmatrix}$$

第一步: 设$a_{11}^{(1)} \neq 0$,依次用 $-l_{i1} = -\dfrac{a_{i1}^{(1)}}{a_{11}^{(1)}},\qquad (i = 2,3,\ldots,n)$

乘矩阵的第1行加到第i行, 得到矩阵:

$$(\mathbf{A}^{(2)},\mathbf{b}^{(2)}) = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \ldots & a_{1n}^{(1)} & b_{1}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \ldots & a_{2n}^{(2)} & b_{2}^{(2)} \\ 0 & a_{32}^{(2)} & a_{33}^{(2)} & \ldots & a_{3n}^{(2)} & b_{3}^{(2)} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & a_{n2}^{(2)} & a_{n3}^{(2)} & \ldots & a_{nn}^{(2)} & b_{n}^{(2)} \end{pmatrix}$$

其中

$$a_{ij}^{(2)} = a_{ij}^{(1)} - l_{i1}a_{1j}^{(1)},\quad i,j = 2,3,\ldots,n$$

$$b_i^{(2)} = b_i^{(1)} - l_{i1}b_1^{(1)},\quad i = 2,3,\ldots,n$$

第二步: 设$a_{22}^{(2)} \neq 0$, 依次用$-l_{i2} = -\dfrac{a_{i2}^{(2)}}{a_{22}^{(2)}}$, $\qquad (i = 3,4,\ldots,n)$

乘矩阵的第2行加到第i行, 得到矩阵:

$$(\mathbf{A}^{(3)},\mathbf{b}^{(3)}) = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \ldots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \ldots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \ldots & a_{3n}^{(3)} & b_3^{(3)} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & a_{n3}^{(3)} & \ldots & a_{nn}^{(3)} & b_n^{(3)} \end{pmatrix}$$

其中

$$a_{ij}^{(3)} = a_{ij}^{(2)} - l_{i2}a_{2j}^{(2)},\quad i,j = 3,4,\ldots,n$$

$$b_i^{(3)} = b_i^{(2)} - l_{i2}b_2^{(2)},\quad i = 3,4,\ldots,n$$

如此继续消元下去, 第n-1步结束后得到矩阵:

$$(\mathbf{A}^{(\mathbf{n})},\mathbf{b}^{(\mathbf{n})})=\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \ldots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \ldots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \ldots & a_{3n}^{(3)} & b_3^{(3)} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{nn}^{(n)} & b_n^{(n)} \end{pmatrix}$$

这就完成了消元过程. 对应的方程组变成:

$$\begin{cases} a_{11}^{(1)}x_1+a_{12}^{(1)}x_2+\ldots+a_{1n}^{(1)}x_n=b_1^{(1)} \\ \qquad\qquad a_{22}^{(2)}x_2+\ldots+a_{2n}^{(2)}x_n=b_2^{(2)} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad\qquad\qquad a_{nn}^{(n)}x_n=b_n^{(n)} \end{cases}$$

对此方程组进行回代, 就可求出方程组的解. 回代的具体算法是:

$$\begin{cases} x_n = b_n^{(n)} / a_{nn}^{(n)}, \\ x_i = \left(b_i^{(i)} - \sum_{j=i+1}^{n} a_{ij}^{(i)} x_j \right) \Big/ a_{ii}^{(i)}, \quad i = n-1, n-2, \dots, 1. \end{cases}$$

顺序Gauss消去法求解n元线性方程组的乘除运算量是:

$$n^2 - 1 + (n-1)^2 - 1 + \cdots + 2^2 - 1 + 1 + 2 + \cdots + n$$

$$= \sum_{k=1}^{n} (k^2 - 1) + \sum_{k=1}^{n} k = \frac{n(n+1)(2n+1)}{6} - n + \frac{n(n+1)}{2}$$

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

n=20时, 顺序Gauss消去法只需3060次乘除法运算.

顺序Gauss消去法通常也简称为**Gauss消去法**.

顺序Gauss消去法中的$a_{kk}^{(k)}$, $(k=1, 2, \ldots, n)$称为**主元素**(**简称主元**).

注: 一般而言, 对某矩阵来说, 主元≠对角元.

命题: **对于矩阵A,**

A的主元素都不为零 $\Leftrightarrow$ A的各阶顺序主子式都不为零.

请思考原因.

§1.2 主元Gauss消去法

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**和**全主元消去法**.

给定线性方程组$\boldsymbol{Ax}=\boldsymbol{b}$，记$\boldsymbol{A}^{(1)}=\boldsymbol{A}$，$\boldsymbol{b}^{(1)}=\boldsymbol{b}$，**列主元Gauss消去法**的具体过程如下:

首先在增广矩阵$\boldsymbol{B}^{(1)}=(\boldsymbol{A}^{(1)},\ \boldsymbol{b}^{(1)})$的第一列元素中选取$a_{k1}^{(1)}$, 使

$$\left|a_{k1}^{(1)}\right| = \max_{1\leq \mathrm{k}\leq n}\left|a_{i1}^{(1)}\right|,$$

并交换第1行与第k行元素, 即$r_k \leftrightarrow r_1$. ($a_{k1}^{(1)}$成为新的主元.)

然后进行第一步消元得增广矩阵$\boldsymbol{B}^{(2)}=(\boldsymbol{A}^{(2)},\ \boldsymbol{b}^{(2)})$. 再在矩阵$\boldsymbol{B}^{(2)}=(\boldsymbol{A}^{(2)}, \boldsymbol{b}^{(2)})$的第二列元素中选取$a_{k2}^{(2)}$, 使

$$\left|a_{k2}^{(2)}\right| = \max_{2\leq \mathrm{k}\leq n}\left|a_{i2}^{(2)}\right|,$$

并交换第2行与第k行元素, 即$r_k \leftrightarrow r_2$. ($a_{k2}^{(2)}$成为新的主元.)
然后进行第二步消元得增广矩阵$\boldsymbol{B}^{(3)}=(\boldsymbol{A}^{(3)}, \boldsymbol{b}^{(3)})$.

按此方法继续进行下去, 经过n-1步选主元和消元运算, 得到增广矩阵$\boldsymbol{B}^{(n)}=(\boldsymbol{A}^{(n)}, \boldsymbol{b}^{(n)})$. 则方程组$\boldsymbol{A}^{(n)}\boldsymbol{x}=\boldsymbol{b}^{(n)}$是与原方程组等价的上三角形方程组, 可进行回代求解.

易证: 只要$|\boldsymbol{A}|\neq 0$, 列主元Gauss消去法就可顺利进行.

例2 采用十进制四位浮点计算, 分别用顺序Gauss消去法和列主元Gauss消去法求解线性方程组:

$$\begin{cases} 0.0120x_1 + 0.0100x_2 + 0.1670x_3 = 0.6781 \\ 1.000x_1 + 0.8334x_2 + 5.910x_3 = 12.10 \\ 3200x_1 + 1200x_2 + 4.200x_3 = 983.3 \end{cases}$$

方程组具有四位有效数字的精确解为

$$x_1^*=17.46,\ x_2^*=-45.76,\ x_3^*=5.546.$$

解1. 用顺序Gauss消去法求解，消元过程为

$$\begin{pmatrix} 0.0120 & 0.0100 & 0.167 & 0.6781 \\ 1.000 & 0.8334 & 5.910 & 12.10 \\ 3200 & 1200 & 4.200 & 983.3 \end{pmatrix}$$

$$\sim\begin{pmatrix} 0.0120 & 0.0100 & 0.1670 & 0.6781 \\ 0 & 0.0001 & -8.007 & -44.41 \\ 0 & -1467 & -4453\times10 & -1798\times10^2 \end{pmatrix}$$

$$\sim\begin{pmatrix} 0.0120 & 0.0100 & 0.1670 & 0.6781 \\ 0 & 0.0001 & -8.007 & -44.41 \\ 0 & 0 & -1175\times10^5 & -6517\times10^5 \end{pmatrix}$$

回代得: $x_1=5.810,\ x_2=-31.78,\ x_3=5.546.$

2. 用列主元Gauss消去法求解, 消元过程为

$$
\begin{pmatrix}
0.0120 & 0.0100 & 0.1670 & 0.6781 \\
1.000 & 0.8334 & 5.910 & 12.10 \\
3200 & 1200 & 4.200 & 983.3
\end{pmatrix}
$$

$$
\underset{r_1 \leftrightarrow r_3}{\overset{\text{选主元}}{\sim}}
\begin{pmatrix}
3200 & 1200 & 4.200 & 983.3 \\
1.000 & 0.8334 & 5.910 & 12.10 \\
0.0120 & 0.0100 & 0.1670 & 0.6781
\end{pmatrix}
$$

$$
\sim
\begin{pmatrix}
3200 & 1200 & 4.200 & 983.3 \\
0 & 0.4584 & 5.909 & 11.79 \\
0 & 0.55\times 10^{-2} & 0.1670 & 0.6744
\end{pmatrix}
$$

$$
\sim
\begin{pmatrix}
3200 & 1200 & 4.200 & 983.3 \\
0 & 0.4584 & 5.909 & 11.79 \\
0 & 0 & 0.0961 & 0.5329
\end{pmatrix}
$$

$$\sim\begin{pmatrix} 3200 & 1200 & 4.200 & 983.3 \\ 0 & 0.4584 & 5.909 & 11.79 \\ 0 & 0 & 0.0961 & 0.5329 \end{pmatrix}$$

回代得: x_1=17.46, x_2= -45.76, x_3=5.545.

可见, 列主元Gauss消去法是在每一步消元前, 在主元所在的一列选取绝对值最大的元素作为主元素.

而全主元Gauss消去法是在每一步消元前, 在所有元素中选取绝对值最大的元素作为主元素, 但由于计算程序更为复杂, 实际应用中并不经常使用.

§2 直接三角分解法

§2.1 Gauss消去法的矩阵表示

对矩阵

$$\boldsymbol{A}^{(1)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix},$$

若 $a_{11}^{(1)} \neq 0$, 令 $l_{i1} = a_{i1}^{(1)}/a_{11}^{(1)}, i = 2,3,\ldots,n.$

$$\boldsymbol{L}_1 = \begin{pmatrix} 1 & & & & \\ -l_{21} & 1 & & & \\ -l_{31} & 0 & 1 & & \\ \cdots & \cdots & \cdots & \cdots & \\ -l_{n1} & 0 & \cdots & \cdots & 1 \end{pmatrix},$$

则有

$$\boldsymbol{A}^{(2)}=\boldsymbol{L}_1\boldsymbol{A}^{(1)}=\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)} \end{pmatrix}.$$

若 $a_{22}^{(2)}\neq 0$，令 $l_{i2}=a_{i2}^{(2)}/a_{22}^{(2)}, i=3,4,\dots,n.$

$$\boldsymbol{L}_2=\begin{pmatrix} 1 & & & \\ & 1 & & \\ & -l_{32} & 1 & & \\ & \dots & & \dots & \\ & -l_{n2} & & & 1 \end{pmatrix},$$

则有

$$A^{(3)} = L_2 A^{(2)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \dots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \dots & a_{3n}^{(3)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & a_{n3}^{(3)} & \dots & a_{nn}^{(3)} \end{pmatrix}.$$

如此进行下去, (最多)在第n-1步:

若 $a_{n-1,n-1}^{(n-1)} \neq 0$, 令 $l_{i,n-1} = a_{i,n-1}^{(n-1)} / a_{n-1,n-1}^{(n-1)}$, $i = n$.

$$L_{n-1} = \begin{pmatrix} 1 & & & & \\ & \dots & & & \\ & & 1 & & \\ & & & 1 & \\ & & & -l_{n,n-1} & 1 \end{pmatrix},$$

则有

$$A^{(n)} = L_{n-1}A^{(n-1)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ & & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ & & & \cdots & \cdots \\ & & & & a_{nn}^{(n)} \end{pmatrix},$$

上述过程可以描述为:

$$A^{(n)}=L_{n-1}A^{(n-1)}=L_{n-1}L_{n-2}A^{(n-2)}=\ldots=L_{n-1}L_{n-2}\ldots L_2L_1A^{(1)},$$

所以有:

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

若记: $L=L_1^{-1}L_2^{-1}\ldots L_{n-1}^{-1}$, $U=A^{(n)}$, 则有

$$A=LU.$$

以下讨论矩阵L, U的性质. 先从L_k, k=1, 2,…, n, 的性质开始讨论.

性质1: 对每个矩阵$\boldsymbol{L}_k$, k=1, 2,…, n-1,

$$\boldsymbol{L}_\mathrm{k} = \begin{pmatrix} 1 & & & & & \\ & \cdots & & & & \\ & & 1 & & & \\ & & -l_{\mathrm{k+1,k}} & 1 & & \\ & & \cdots & & & \\ & & -l_{\mathrm{n,k}} & & \cdots & \\ & & & & & 1 \end{pmatrix},$$

其逆矩阵为

$$\boldsymbol{L}_k^{-1} = \begin{pmatrix} 1 & & & & & \\ & \cdots & & & & \\ & & 1 & & & \\ & & l_{\mathrm{k+1,k}} & 1 & & \\ & & \cdots & & & \\ & & l_{\mathrm{n,k}} & & \cdots & \\ & & & & & 1 \end{pmatrix},$$

性质2: 若每个矩阵$\boldsymbol{L}_k^{-1}$, k=1,2,…, n-1, 形为

$$\boldsymbol{L}_k^{-1}=\begin{pmatrix} 1 & & & & \\ & \cdots & & & \\ & & 1 & & \\ & & l_{k+1,k} & 1 & \\ & & \cdots & & \\ & & l_{n,k} & & \cdots & \\ & & & & & 1 \end{pmatrix},$$

则它们的乘积

$$\boldsymbol{L}=\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\ldots\boldsymbol{L}_{n-1}^{-1}=\begin{pmatrix} 1 & & & & & \\ l_{2,1} & \cdots & & & & \\ l_{3,1} & \cdots & 1 & & & \\ & & l_{k+1,k} & \cdots & & \\ \vdots & \vdots & \cdots & \cdots & 1 & \\ l_{n,1} & \cdots & l_{n,k} & \cdots & l_{n,n-1} & 1 \end{pmatrix}.$$

注意此次序不能随意颠倒, 否则得不到右边结果.

于是得到: $A=A^{(1)}=L_1^{-1}L_2^{-1}\ldots L_{n-1}^{-1}A^{(n)}=LU$, 其中

$$L=\begin{pmatrix} 1 & & & & & \\ l_{2,1} & \cdots & & & & \\ l_{3,1} & \cdots & 1 & & & \\ \vdots & \vdots & l_{k+1,k} & \cdots & & \\ \vdots & \vdots & \cdots & \cdots & 1 & \\ l_{n,1} & \cdots & l_{n,k} & \cdots & l_{n,n-1} & 1 \end{pmatrix},$$

L称为**单位下三角矩阵**.

$$U=\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,k}^{(1)} & \cdots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} & \cdots & a_{2n}^{(2)} \\ & & \cdots & \cdots & \cdots & \cdots \\ & & & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ & & & & \cdots & \cdots \\ & & & & & a_{nn}^{(n)} \end{pmatrix}.$$

U是**上三角矩阵**.

返回第35页

总结

若矩阵$\boldsymbol{A}$存在分解

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

其中$\boldsymbol{L}$为**单位**下三角阵, $\boldsymbol{U}$为上三角阵, 则称此分解为**$\boldsymbol{A}$的杜立特尔(Doolittle)分解或直接三角分解**.

由上述讨论可知：只要矩阵$\boldsymbol{A}$的顺序Gauss消去过程中每步主元 $a_{kk}^{(k)}\neq 0$, 则$\boldsymbol{A}$必存在Doolittle分解.

注：这种关于矩阵$\boldsymbol{A}$的“三角型分解”有很多种，分解结果也未必惟一.

然而......

§ 2.2 直接三角分解法

定理2.1 设n阶方阵A的各阶顺序主子式不为零, 则存在唯一单位下三角矩阵L和上三角矩阵U使$A=LU$.

证明: 存在性. 只须证明对于A的k阶顺序主子式成立

$$\det(A_k) = a_{11}^{(1)} a_{22}^{(2)} \dots a_{kk}^{(k)}. \tag{2.9}$$

显然, $\det(A_1) = a_{11}^{(1)}$, 即式(2.9)对k=1成立.

以下假设式(2.9)对$k=m$-1成立, 此时可完成Gauss消去法的第m-1步消元, 于是可得

$$\begin{pmatrix} L_m & O \\ M & I_{n-m} \end{pmatrix} \begin{pmatrix} A_m & B \\ C & D \end{pmatrix} = \begin{pmatrix} U_m & V \\ W & X \end{pmatrix}.$$

根据分块矩阵运算法则可知

$$L_m A_m = U_m,\ \ det(L_m)det(A_m) = det(U_m).$$

由此可知式(2.9)对$k=m$也成立. 根据归纳法原理, 式(2.9)成立. 于是三角分解的存在性得证.

唯一性: 设有两种分解

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

则有

$$\bar{\boldsymbol{L}}^{-1}\boldsymbol{L} = \bar{\boldsymbol{U}}\boldsymbol{U}^{-1} = \boldsymbol{E}.$$

观察这个等式两端矩阵的形状

所以得

$$\boldsymbol{L} = \bar{\boldsymbol{L}},$$

$$\boldsymbol{U} = \bar{\boldsymbol{U}}.$$

证毕.

Doolittle分解的用处:

如果$\boldsymbol{A}$存在Doolittle分解$\boldsymbol{A}=\boldsymbol{LU}$, 则

$$\boldsymbol{Ax}=\boldsymbol{b} \;\Rightarrow\; \boldsymbol{LUx}=\boldsymbol{b}$$

令$\boldsymbol{Ux}=\boldsymbol{y}$得

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

于是对方程组

$$\boldsymbol{Ax}=\boldsymbol{b}$$

的求解转化为: 先对方程组$\boldsymbol{Ly}=\boldsymbol{b}$求解, 再对$\boldsymbol{Ux}=\boldsymbol{y}$求解.

由于$\boldsymbol{L}$和$\boldsymbol{U}$都是三角形矩阵, 因此对这两个方程组的计算一般要容易得多.

由

$$\begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \cdots & \cdots & \cdots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix},$$

$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \cdots & \cdots \\ & & & u_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}=\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}.$$

可得

$$\begin{cases} y_1 = b_1, \\ y_k = b_k - \sum_{i=1}^{k-1} l_{ki} y_i, \end{cases} \quad k = 2,3,\ldots,n.$$

$$\begin{cases} x_n = y_n / u_{nn}, \\ x_i = \left(y_i - \sum_{j=i+1}^{n} u_{ij} x_j \right) \Big/ u_{ii}, \end{cases} \quad i = n-1, n-2, \ldots, 1.$$

这就是求解方程组$\boldsymbol{Ax}=\boldsymbol{b}$的直接三角分解方法.

问题: 矩阵A的Doolittle分解是否只能由Gauss消元法得到?

答： 矩阵A的Doolittle分解也可以由其他途径得到.

矩阵三角分解的直接三角分解方法

思想：根据矩阵乘法

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

关于具体算法流程, 参见教材P24. 或详见教师讲解.

直接三角分解可在分解过程中实时地进行压缩存储. 参见教材P23.

具体算例：对以下矩阵做直接三角分解.

$$\mathbf{A}=\begin{pmatrix} 2 & 1 & 1 \\ 4 & 4 & 3 \\ 6 & 7 & 7 \end{pmatrix}$$

例3 利用直接三角分解方法解线性方程组

$$\begin{cases} 2x_1 + x_2 + x_3 = 3 \\ 4x_1 + 4x_2 + 3x_3 = 7 \\ 6x_1 + 7x_2 + 7x_3 = 13 \end{cases}$$

解 因为

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 4 & 3 \\ 6 & 7 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

所以

先解 $\begin{pmatrix} 1 & & \\ 2 & 1 & \\ 3 & 2 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 13 \end{pmatrix}$，得 $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$

再解 $\begin{pmatrix} 2 & 1 & 1 \\ & 2 & 1 \\ & & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$，得 $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

解线性方程组$\boldsymbol{Ax}=\boldsymbol{b}$的直接三角分解法的计算量约为**(1/3)$n^3$**，与Gauss消去法基本相同. **其优点在于求一系列同系数的线性方程组$\boldsymbol{Ax}=\boldsymbol{b}_k$, ($k$=1, 2,…, m)时，可大大节省运算量.**

例: 求矩阵$\boldsymbol{A}$的逆矩阵, 这里

$$\boldsymbol{A}=\begin{pmatrix}1 & 2 & -3\\ 2 & -1 & 3\\ 3 & -2 & 2\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 3 & \frac{8}{5} & 1\end{pmatrix}\begin{pmatrix}1 & 2 & -3\\ 0 & -5 & 9\\ 0 & 0 & -\frac{17}{5}\end{pmatrix}$$

解题思路: 可分别取常向量

$$\boldsymbol{b}_1=(1,0,0)^{\mathrm{T}},\ \boldsymbol{b}_2=(0,1,0)^{\mathrm{T}},\ \boldsymbol{b}_3=(0,0,1)^{\mathrm{T}},$$

令 $\boldsymbol{X}=\boldsymbol{A}^{-1}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)$, 则

$$\boldsymbol{A}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)=(\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3).$$

逐个解方程组

$$\boldsymbol{A}\boldsymbol{x}_k=\boldsymbol{b}_k,\ k=1,2,3,$$

即可求得$\boldsymbol{A}^{-1}$.

具体地, 由

$$\begin{pmatrix}1&&\\2&1&\\3&\frac{8}{5}&1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}1&2&-3\\&-5&9\\&&-\frac{17}{5}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix},\text{得}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}\frac{4}{17}\\\frac{5}{17}\\-\frac{1}{17}\end{pmatrix}$$

$$\begin{pmatrix}1&&\\2&1&\\3&\frac{8}{5}&1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}1&2&-3\\&-5&9\\&&-\frac{17}{5}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix},\text{得}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}\frac{2}{17}\\\frac{11}{17}\\\frac{8}{17}\end{pmatrix}$$

$$\begin{pmatrix}1&&\\2&1&\\3&\frac{8}{5}&1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\begin{pmatrix}0\\0\\1\end{pmatrix},\begin{pmatrix}1&2&-3\\&-5&9\\&&-\frac{17}{5}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix},\text{得}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}\frac{3}{17}\\-\frac{9}{17}\\-\frac{5}{17}\end{pmatrix}$$

所以 $$\mathbf{A}^{-1}=\begin{pmatrix}\frac{4}{17}&\frac{2}{17}&\frac{3}{17}\\\frac{5}{17}&\frac{11}{17}&-\frac{9}{17}\\-\frac{1}{17}&\frac{8}{17}&-\frac{5}{17}\end{pmatrix}=\frac{1}{17}\begin{pmatrix}4&2&3\\5&11&-9\\-1&8&-5\end{pmatrix}$$

对于n阶矩阵的求逆问题

如果在刚才的计算过程中采用Gauss消去法, 则

全部计算量=n次Gauss消去法=$n\cdot O(n^3)=O(n^4)$.

而如果在刚才的计算过程中采用直接三角分解法, 则

全部计算量=1次直接三角分解+$2n$次回代

$=O(n^3)+2n\cdot O(n^2)=O(n^3)$.

可见对于这类问题两种不同的方法的计算量相差了一个数量级.

为了提高数值稳定性, 可考虑列主元三角分解法. 具体算法参见教材P23.

但要注意： **应用列主元三角分解法时， 方程组右端的常数项也要做相应的交换.**

问题： **当矩阵A的顺序Gauss消元法无法执行时，A的Doolittle分解是否还存在?**

请自己思考. (可能比想象中复杂)

§ 2.3 平方根法

设$\boldsymbol{A}$为对称正定矩阵, 则有唯一分解$\boldsymbol{A=LU}$, 其中$\boldsymbol{U}=(u_{ij})$且其每个对角元u_{kk}均必满足条件: $u_{kk}>0$. 因此还可得到

Why?

想想$\boldsymbol{U}$的对角元,

正定矩阵与顺序主子式的关系

$$
\boldsymbol{U}=\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \cdots & \cdots \\ & & & u_{nn} \end{pmatrix}
$$

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

则有 $\boldsymbol{A=LDM}$.

由 $\boldsymbol{A}=\boldsymbol{LDM}$，并由$\boldsymbol{A}$的对称性可得:

$$(\boldsymbol{LDM})^{\mathrm{T}}=\boldsymbol{M}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}=\boldsymbol{LDM},$$

比较后一个等式的两端, 并由此时Doolittle分解的唯一性知:

$$\boldsymbol{M}=\boldsymbol{L}^{\mathrm{T}}.$$

再令

$$\boldsymbol{D}=\underset{\substack{\uparrow\\ \boldsymbol{D}^{1/2}}}{\begin{pmatrix}\sqrt{u_{11}} & & & \\ & \sqrt{u_{22}} & & \\ & & \cdots & \\ & & & \sqrt{u_{nn}}\end{pmatrix}}\underset{\substack{\uparrow\\ \boldsymbol{D}^{1/2}}}{\begin{pmatrix}\sqrt{u_{11}} & & & \\ & \sqrt{u_{22}} & & \\ & & \cdots & \\ & & & \sqrt{u_{nn}}\end{pmatrix}}$$

则有

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{D}^{1/2}\boldsymbol{D}^{1/2}\boldsymbol{L}^{\mathrm{T}}=(\boldsymbol{L}\boldsymbol{D}^{1/2})(\boldsymbol{L}\boldsymbol{D}^{1/2})^{\mathrm{T}}=\boldsymbol{G}\boldsymbol{G}^{\mathrm{T}},\ 其中\,\boldsymbol{G}=\boldsymbol{L}\boldsymbol{D}^{1/2}.$$

上述分解$A=GG^T$称为对称正定矩阵A的**Cholesky(乔列斯基)分解**，也叫**平方根分解**.

注1：当矩阵A对称正定时，必存在Cholesky分解$A=GG^T$，其中G为满秩的下三角矩阵.

注2：严格来说，对称正定矩阵A的形如$A=GG^T$的分解是不唯一的．而Cholesky分解(一般指限制矩阵G的对角元都为正实数的情形)$A=GG^T$是唯一的.

注3：对任意的对称正定矩阵A，也可以不通过Doolittle分解而直接做Cholesky分解. (具体算法见以后内容)

注4：另有所谓“改进的平方根法”，其思想和原理就是基于分解$A=LDL^T$，具体算法参见教材P27.

求解系数矩阵A为对称正定矩阵的线性方程组的平方根法

步骤1: 利用平方根分解将$\boldsymbol{Ax=b}$ 转换为两个方程组

$$\boldsymbol{Gy=b},\ \boldsymbol{G}^{\mathrm{T}}\boldsymbol{x=y}.$$

对A的平方根分解$A=GG^{\mathrm{T}}$的具体算法为:

若记$\boldsymbol{G}=(g_{ij})$, 则有: 对$k=1, 2,\ldots, n.$

$$\begin{cases} g_{kk} = \left(a_{kk} - \sum_{m=1}^{k-1} g_{km}^2 \right)^{1/2}, \\ g_{ik} = \left(a_{ik} - \sum_{m=1}^{k-1} g_{im} g_{km} \right) \Big/ g_{kk}, \end{cases}$$

$$i = k+1, \ldots, n.$$

步骤2: 先后解三角型方程组 $\boldsymbol{Gy}=\boldsymbol{b}$, $\boldsymbol{G}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{y}$, 可得

$$\begin{cases} y_k = \left(b_k - \sum_{m=1}^{k-1} g_{km} y_m\right) \Big/ g_{kk}, \quad k = 1,2,\ldots,n. \\ x_k = \left(y_k - \sum_{m=i+1}^{n} g_{mk} x_m\right) \Big/ g_{kk}, \quad k = n, n-1, \ldots, 1. \end{cases}$$

具体算例: 对以下矩阵做Cholesky分解(平方根分解).

$$\boldsymbol{A}=\begin{pmatrix}4 & -2 & 4\\ -2 & 5 & 0\\ 4 & 0 & 6\end{pmatrix}.$$

具体计算过程, 见教师黑板讲解.

例4 解线性方程组 $\begin{cases} 4x_1 - 2x_2 + 4x_3 = 2 \\ -2x_1 + 5x_2 = 1 \\ 4x_1 + 6x_3 = 0 \end{cases}$

解

$$\boldsymbol{A} = \boldsymbol{G}\boldsymbol{G}^{\mathrm{T}} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 2 & 0 \\ 2 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & -1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

解方程 $\begin{pmatrix} 2 & & \\ -1 & 2 & \\ 2 & 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$，得 $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix}$

再解方程 $\begin{pmatrix} 2 & -1 & 2 \\ & 2 & 1 \\ & & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix}$，得 $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4.5 \\ 2 \\ -3 \end{pmatrix}$

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的LU分解的一半. (理论上)Cholesky分解具有数值稳定性。

§ 2.4 追 赶 法

追赶法是求如下三对角线性方程组的三角分解法.

$$\begin{pmatrix} a_1 & c_1 & & & & \\ d_2 & a_2 & c_2 & & & \\ & \cdots & \cdots & \cdots & & \\ & & d_{n-1} & a_{n-1} & c_{n-1} \\ & & & d_n & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_{n-1} \\ b_n \end{pmatrix}$$

三对角矩阵$\boldsymbol{A}$的各阶顺序主子式都不为零的一个充分条件是:

$|a_1|>|c_1|>0$; $|a_n|>|d_n|>0$; $|a_i|\geq|c_i|+|d_i|$, $c_i d_i \neq 0$, $i=2, 3,\dots , n-1$.

当$\boldsymbol{A}$的各阶顺序主子式都不为零时，存在分解$\boldsymbol{A}=\boldsymbol{TM}$,其中$\boldsymbol{T}$为下三角阵, $\boldsymbol{M}$为单位上三角阵. 称此分解为矩阵$\boldsymbol{A}$的**Crout(克劳特)分解.**

对三对角矩阵$\boldsymbol{A}$进行Crout分解, 有

$$\boldsymbol{A}=\begin{pmatrix}\alpha_1 & & & & \\ \gamma_2 & \alpha_2 & & & \\ & \cdots & \cdots & & \\ & & \gamma_{n-1} & \alpha_{n-1} & \\ & & & \gamma_n & \alpha_n\end{pmatrix}\begin{pmatrix}1 & \beta_1 & & & \\ & 1 & \beta_2 & & \\ & & \cdots & \cdots & \\ & & & 1 & \beta_{n-1} \\ & & & & 1\end{pmatrix}$$

其中

$$\begin{cases}\alpha_1=a_1,\beta_1=c_1/\alpha_1,\gamma_i=d_i,\ \ i=2,3,\dots,n.\\ \alpha_i=a_i-d_i\beta_{i-1}, i=2,3,\dots,n.\\ \beta_i=c_i/\alpha_i,\ i=2,3,\dots,n-1.\end{cases}$$

解三角方程 $\boldsymbol{Ty}=\boldsymbol{b}$, $\boldsymbol{Mx}=\boldsymbol{y}$ 可得

$$\begin{cases}y_1=b_1/\alpha_1, y_i=(b_i-\gamma_i y_{i-1})/\alpha_i,\ \ i=2,3,\dots,n.\\ x_n=y_n, x_i=y_i-\beta_i x_{i+1},\ \ \ i=n-1,n-2,\dots,1.\end{cases}$$

称之为解三对角方程组的**追赶法**.

例5 解线性方程组

$$\begin{pmatrix} 3 & 2 & & \\ -1 & 3 & 2 & \\ & -1 & 3 & 2 \\ & & -1 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ 11 \\ 15 \\ 9 \end{pmatrix}$$

解:

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & & \\ -1 & 3 & 2 & \\ & -1 & 3 & 2 \\ & & -1 & 3 \end{pmatrix} \to \begin{pmatrix} 3 & \frac{2}{3} & & \\ -1 & \frac{11}{3} & \frac{6}{11} & \\ & -1 & \frac{39}{11} & \frac{22}{39} \\ & & -1 & \frac{139}{39} \end{pmatrix}$$

所以

$$\mathbf{A} = \begin{pmatrix} 3 & & & \\ -1 & \frac{11}{3} & & \\ & -1 & \frac{39}{11} & \\ & & -1 & \frac{139}{39} \end{pmatrix}\begin{pmatrix} 1 & \frac{2}{3} & & \\ & 1 & \frac{6}{11} & \\ & & 1 & \frac{22}{39} \\ & & & 1 \end{pmatrix}$$

解方程 $\begin{pmatrix} 3 & & & \\ -1 & \frac{11}{3} & & \\ & -1 & \frac{39}{11} & \\ & & -1 & \frac{139}{39} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 7 \\ 11 \\ 15 \\ 9 \end{pmatrix}$，得 $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} \frac{7}{3} \\ \frac{40}{11} \\ \frac{205}{39} \\ 4 \end{pmatrix}$

再解方程 $\begin{pmatrix} 1 & \frac{2}{3} & & \\ & 1 & \frac{6}{11} & \\ & & 1 & \frac{22}{39} \\ & & & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \frac{7}{3} \\ \frac{40}{11} \\ \frac{205}{39} \\ 4 \end{pmatrix}$，得 $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$

当满足条件

$|a_1|>|c_1|>0$; $|a_n|>|d_n|>0$; $|a_i|\ge|c_i|+|d_i|$, $c_i d_i \neq 0$, $i=2, 3,\ldots, n-1.$

时, 追赶法是数值稳定的. 追赶法具有计算程序简单、存贮量少、计算量小等优点.

§2.4 向量和矩阵的范数

为了研究线性方程组近似解的误差估计和迭代法的收敛性，我们需要对R^n (n维向量空间)中的向量或者$\mathrm{R}^{n\times n}$ ($n\times n$实矩阵空间)中的矩阵的“大小”引进某种度量.

向量范数概念可以看作实数域上绝对值函数以及三维欧式空间中向量长度概念的推广，在数值分析中起着重要作用.

回忆：

若$N(x)=|x|$代表实数域R上的绝对值函数，或者$N(\boldsymbol{x})=|\boldsymbol{x}|$代表三维向量空间$\mathrm{R}^3$上的欧氏模(长度)函数, 此时

$$\boldsymbol{x}=(x_1, x_2, x_3)^{\mathrm{T}},$$

则这样的函数$|\boldsymbol{x}|$均满足以下条件:

(1) 对任何$\boldsymbol{x}$均有$|\boldsymbol{x}|\geq 0$, 且$|\boldsymbol{x}|=0$当且仅当$\boldsymbol{x}=\boldsymbol{0}$;

(2) 对任何$\boldsymbol{x}$和实数α, 均有

$$|\alpha\boldsymbol{x}|=|\alpha|\,|\boldsymbol{x}|;$$

(3) 对任何$\boldsymbol{x}, \boldsymbol{y}\in\mathrm{R}$ (或R^3), 均有

$$|\boldsymbol{x}+\boldsymbol{y}|\leq|\boldsymbol{x}|+|\boldsymbol{y}|.$$

以下来构造R^n空间上的度量函数.

§ 2.4.1 向量的范数

定义2.1 设$\|\bullet\|$是向量空间R^n上的实值函数, 若满足条件:

(1)非负性: 对任何向量$\boldsymbol{x}\in\mathrm{R}^n$, 成立$\|\boldsymbol{x}\|\ge 0$; 且$\|\boldsymbol{x}\|=0$当且仅当$\boldsymbol{x}=\boldsymbol{0}$;

(2)齐次性: 对任何向量$\boldsymbol{x}\in\mathrm{R}^n$和实数$\alpha$, 成立

$$\|\alpha\boldsymbol{x}\|=|\alpha|\,\|\boldsymbol{x}\|;$$

(3)三角不等式: 对任何向量$\boldsymbol{x}, \boldsymbol{y}\in\mathrm{R}^n$, 成立

$$\|\boldsymbol{x}+\boldsymbol{y}\|\le\|\boldsymbol{x}\|+\|\boldsymbol{y}\|,$$

则称函数$\|\bullet\|$为空间R^n上的一个**向量范数**, $\|\boldsymbol{x}\|$为向量$\boldsymbol{x}$的范数.

补充

根据以上基本性质，还会得到一些比较基本的性质，比如：

对任何向量$\boldsymbol{x}, \boldsymbol{y} \in \mathrm{R}^n$，成立

$$\big|\, \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \,\big| \le \|\boldsymbol{x} - \boldsymbol{y}\|.$$

思考.

记$\boldsymbol{x}=(x_1, x_2,\dots, x_n)^T$, 常用的向量范数有:

向量的1-范数: $\|\boldsymbol{x}\|_1=|x_1|+|x_2|+\dots+|x_n|$

向量的2-范数: $\|\boldsymbol{x}\|_2=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}$

欧氏

范数

向量的∞-范数: $\|\boldsymbol{x}\|_\infty=\max\limits_{1\le i\le n}|x_i|$.

例 设向量$\boldsymbol{x}=(2, -4, 3, 1)^T$, 求向量范数$\|\boldsymbol{x}\|_p$, $p=1, 2, \infty$.

解: 可得$\|\boldsymbol{x}\|_1=10$, $\|\boldsymbol{x}\|_\infty=4$, $\|\mathbf{x}\|_2=\sqrt{30}$.

例题

对任给$\boldsymbol{x}=(x_1, x_2, x_3)\in R^3$, 判断下列函数是否是向量范数.

(1) $\|\boldsymbol{x}\|=\lvert x_1\rvert+\lvert x_2\rvert-\lvert x_3\rvert$.	不满足非负性和三角不等式
(2) $\|\boldsymbol{x}\|=\lvert x_1\rvert+\lvert 2x_2+x_3\rvert$.	不满足非负性.
(3) $\|\boldsymbol{x}\|=\lvert x_1\rvert^2+\lvert x_2\rvert^2+\lvert x_3\rvert^2$.	不满足齐次性和三角不等式.
(4) $\|\boldsymbol{x}\|=\lvert x_1\rvert+2\lvert x_2\rvert+3\lvert x_3\rvert$.	是向量范数.
(5) $\|\boldsymbol{x}\|=\lvert x_1+x_2\rvert+\lvert x_2+x_3\rvert+\lvert x_3+x_1\rvert$.	是向量范数
(6) $\|\boldsymbol{x}\|=\lvert x_1-x_2\rvert+\lvert x_2-x_3\rvert+\lvert x_3-x_1\rvert$.	不满足非负性.

补充: 加权范数

设$\boldsymbol{x}\in\mathrm{R}^n$, $\boldsymbol{W}\in\mathrm{R}^{n\times n}$, 且为**非奇异的**对角阵(即对角线不为0的对角阵), 定义$\boldsymbol{x}$的**加权范数**为:

$$\|\boldsymbol{x}\|_{\boldsymbol{W}}=\|\boldsymbol{W}\boldsymbol{x}\|,$$

其中$\|\boldsymbol{x}\|$是$\boldsymbol{x}$的某种向量范数.

注: $\boldsymbol{W}$的对角元素相当于权系数.

例: 对于$\boldsymbol{x}=(x_1, x_2, x_3)^{\mathrm{T}}\in\mathrm{R}^3$,

$$\boldsymbol{W}=\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & -3\end{pmatrix}\in\mathrm{R}^{3\times 3},$$

其加权1-范数为:

$$\|\boldsymbol{x}\|_{\boldsymbol{W}}=\|\boldsymbol{W}\boldsymbol{x}\|_1=\left\|\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & -3\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}\right\|_1=|x_1|+2|x_2|+3|x_3|$$

引申:

命题: 设$\boldsymbol{x}\in \mathrm{R}^n$, $\boldsymbol{W}\in \mathrm{R}^{n\times n}$且为**非奇异**阵(或**列满秩**阵), $\|\boldsymbol{x}\|$是$\boldsymbol{x}$的某种向量范数, 则

$$\|\boldsymbol{x}\|_W=\|\boldsymbol{W}\boldsymbol{x}\|$$

必为$\boldsymbol{x}$的一个向量范数.

注: 参见教材P46的问题2-11. 请思考其具体证明.

虽然同一向量的不同范数的值可能不同, 但是……

范数的连续性和等价性

事实上

向量的*p*-范数: $\|\mathbf{x}\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}.$

（$p \geq 1$）

可以证明:

$$\lim_{p \to +\infty} \|\mathbf{x}\|_p = \max_{1 \leq i \leq n} |x_i| = \|\mathbf{x}\|_\infty.$$

定理2.3(范数的等价性): 对于 R^n 上的任何两种向量范数 $\|\bullet\|_\alpha$ 和 $\|\bullet\|_\beta$，存在正的常数 m 和 M, 使得

$$m\|\boldsymbol{x}\|_\alpha \le \|\boldsymbol{x}\|_\beta \le M\|\boldsymbol{x}\|_\alpha,\ \forall \boldsymbol{x}\in R^n.$$

证明：对于 R^n 中的有界闭集 $S=\{\boldsymbol{x}|\ \boldsymbol{x}\in R^n,\ \|\boldsymbol{x}\|_2=1\}$，因 R^n 上的任意范数 $N(\boldsymbol{x})$ 是连续函数，故 $N(\boldsymbol{x})$ 必在S上有最大值 M 和最小值 m. (此性质适用于有限维空间中有界闭集上的连续函数). 此时, 对于任意给定的范数 $N(\boldsymbol{x})$, 存在正的常数 M, m, 使

$$0<m \le N(\boldsymbol{x}) \le M,\ \ \forall \boldsymbol{x}\in S.$$

以下主要证明 R^n 上的 2-范数 $\|\bullet\|_2$ 与任意范数 $\|\bullet\|_\alpha$ 是等价的.

当 $\boldsymbol{x}=\boldsymbol{0}$ 时，定理结论中的不等式显然成立，事实上此时不等号可取等号, 也即: 对于任何两种向量范数 $\|\bullet\|_\alpha$ 和 $\|\bullet\|_\beta$，对于任意正的常数 m 和 M, 均有

$$m\|\boldsymbol{0}\|_\alpha = \|\boldsymbol{0}\|_\beta = M\|\boldsymbol{0}\|_\alpha. \qquad (*)$$

而当$\boldsymbol{x}\neq\boldsymbol{0}$时, 必有

$$\frac{1}{\|\boldsymbol{x}\|_2}\boldsymbol{x}\in S.$$

因此, 对于$N(\boldsymbol{x})=\|\bullet\|_\alpha$, 必存在正的常数$M_1$, m_1,使

$$m_1\le\left\|\frac{1}{\|\boldsymbol{x}\|_2}\boldsymbol{x}\right\|_\alpha\le M_1.$$

也即: $m_1\|\boldsymbol{x}\|_2\le\|\boldsymbol{x}\|_\alpha\le M_1\|\boldsymbol{x}\|_2$, $\forall\boldsymbol{x}\in\mathrm{R}^n$, $\boldsymbol{x}\neq\boldsymbol{0}$.

同理, 对于$N(\boldsymbol{x})=\|\bullet\|_\beta$, 必存在正的常数$M_2$, m_2,使

$$m_2\|\boldsymbol{x}\|_2\le\|\boldsymbol{x}\|_\beta\le M_2\|\boldsymbol{x}\|_2,\ \forall\boldsymbol{x}\in\mathrm{R}^n,\ \boldsymbol{x}\neq\boldsymbol{0}.$$

因此可知: 对于R^n上的任意两种向量范数$\|\bullet\|_\alpha$和$\|\bullet\|_\beta$, 必有

$$\frac{m_2}{M_1}\|\boldsymbol{x}\|_\alpha\le\|\boldsymbol{x}\|_\beta\le\frac{M_2}{m_1}\|\boldsymbol{x}\|_\alpha,\quad\forall\boldsymbol{x}\in R^n,\boldsymbol{x}\neq\boldsymbol{0}.$$

再结合(*)式, 可得定理命题成立. 证毕.

注1：范数的等价性定理结论中，只要给定R^n上任意的两个向量范数$\|\bullet\|_\alpha$和$\|\bullet\|_\beta$. 即可确定常数M, m, 而这组常数M, m适用于所有的$\boldsymbol{x}\in\mathrm{R}^n$.

注2：范数的定价性定理说明：对于R^n 上任意一个给定的向量$\boldsymbol{x}$, 尽管关于$\boldsymbol{x}$的不同范数的取值大小不同, 但这些不同的范数值“相差不大”.

R^n上常用的三种向量范数满足如下等价关系:

$$\|\boldsymbol{x}\|_\infty \le \|\boldsymbol{x}\|_1 \le n\|\boldsymbol{x}\|_\infty, \quad \forall \boldsymbol{x} \in R^n.$$

$$\|\boldsymbol{x}\|_\infty \le \|\boldsymbol{x}\|_2 \le \sqrt{n}\|\boldsymbol{x}\|_\infty, \quad \forall \boldsymbol{x} \in R^n.$$

$$\|\boldsymbol{x}\|_2 \le \|\boldsymbol{x}\|_1 \le \sqrt{n}\|\boldsymbol{x}\|_2, \quad \forall \boldsymbol{x} \in R^n.$$

R^n上的向量序列$\boldsymbol{x}^{(k)}=(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})^T$, $k=1, 2, \ldots$ 是指以下这些向量所形成的序列:

$$\begin{array}{cccccc}
\boldsymbol{x}^{(1)}, & \boldsymbol{x}^{(2)}, & \boldsymbol{x}^{(3)}, & \cdots\cdots & \boldsymbol{x}^{(k)}, & \cdots\cdots \\
\parallel & \parallel & \parallel & & \parallel & \\
\begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \\ \cdots \\ x_n^{(1)} \end{pmatrix}, & \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ \cdots \\ x_n^{(2)} \end{pmatrix}, & \begin{pmatrix} x_1^{(3)} \\ x_2^{(3)} \\ \cdots \\ x_n^{(3)} \end{pmatrix}, & \cdots\cdots & \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \cdots \\ x_n^{(k)} \end{pmatrix}, & \cdots\cdots
\end{array}$$

定义2.2 设有向量序列$\boldsymbol{x}^{(k)}=(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})^{\mathrm{T}}, k=1, 2,\ldots$ 以及常向量$\boldsymbol{x}^*=(x_1^*, x_2^*, \ldots, x_n^*)^{\mathrm{T}}$. 如果

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

则称**向量序列$\{\boldsymbol{x}^{(k)}\}$收敛于向量$\boldsymbol{x}^*$**, 记作

$$\lim_{k\to\infty}\boldsymbol{x}^{(k)}=\boldsymbol{x}^*, \quad \text{或}\ \boldsymbol{x}^{(k)}\to\boldsymbol{x}^*, (k\to\infty)$$

思考: 以下命题必成立.

$$\boldsymbol{x}^{(\mathrm{k})}\to\boldsymbol{x}^*\Leftrightarrow x_{\mathrm{i}}^{(\mathrm{k})}\to x_{\mathrm{i}}^*, \mathrm{i}=1,2,\ldots,\mathrm{n}.$$

补充定理(范数的连续性): 设$N(\boldsymbol{x})=\|\boldsymbol{x}\|$是$\mathrm{R}^n$上的任意向量范数, 则$N(\boldsymbol{x})$是$\boldsymbol{x}$的(分量$x_1, x_2, \ldots, x_n$的)连续函数.

证明: 对于R^n中任意范数$N(\boldsymbol{x})=\|\boldsymbol{x}\|$以及任意两个向量$\boldsymbol{x}, \boldsymbol{y}$, 均有

$$|N(\boldsymbol{x})-N(\boldsymbol{y})| = |\|\boldsymbol{x}\|-\|\boldsymbol{y}\|| \le \|\boldsymbol{x}-\boldsymbol{y}\|$$

由三角不等式的推论

故当$\boldsymbol{x}\to\boldsymbol{y}$时(见定义2.2), 也即$\|\boldsymbol{x}-\boldsymbol{y}\|\to 0$时, 必有

$$|N(\boldsymbol{x})-N(\boldsymbol{y})|\to 0.$$

证毕.

注: 这说明当$\boldsymbol{x}$变化不大时, 其范数值$\|\boldsymbol{x}\|$也变化不大.

§ 2.4.2 矩阵的范数

定义2.3 设$\|\bullet\|$是以n阶方阵为变量的实值函数，且满足条件：对于任意的n阶方阵$\boldsymbol{A}$与$\boldsymbol{B}$，均成立

(1) **非负性**：$\|\boldsymbol{A}\|\geq 0$，且$\|\boldsymbol{A}\|=0$当且仅当$\boldsymbol{A}=\boldsymbol{O}$；

(2) **齐次性**：$\|\alpha\boldsymbol{A}\| = |\alpha|\,\|\boldsymbol{A}\|$，$\alpha\in\mathrm{R}$；

(3) **三角不等式**：$\|\boldsymbol{A}+\boldsymbol{B}\|\leq\|\boldsymbol{A}\|+\|\boldsymbol{B}\|$；

(4) **相容性**：$\|\boldsymbol{AB}\|\leq\|\boldsymbol{A}\|\|\boldsymbol{B}\|$，

则称$\|\boldsymbol{A}\|$为矩阵$\boldsymbol{A}$的**矩阵范数**.

记$\boldsymbol{A}=(a_{\rm ij})_{n\times n}$, 即

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

常用的矩阵范数有:

矩阵的1-范数: $\|\boldsymbol{A}\|_1=\max\limits_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|$, 也称矩阵的**列范数**.

矩阵的2-范数: $\|\boldsymbol{A}\|_2=\sqrt{\max\lambda(\boldsymbol{A}^{\rm T}\boldsymbol{A})}$, 也称矩阵的**谱范数**.

矩阵的∞-范数: $\|\boldsymbol{A}\|_\infty=\max\limits_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$, 也称矩阵的**行范数**.

矩阵的F-范数:

$$\|\boldsymbol{A}\|_{\rm F}=\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2}=\sqrt{\operatorname{tr}(\boldsymbol{A}^{\rm T}\boldsymbol{A})}.$$

Frobenius 范数

也是矩阵范数中的**欧氏范数**

例 设矩阵

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

求矩阵A的范数$\|A\|_p$, (p =1, 2, ∞)以及F-范数$\|A\|_F$.

解: $\|A\|_1=4$, $\|A\|_\infty=5$, $\|A\|_F=\sqrt{15}$

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

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

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

设$\|\bullet\|$是一种向量范数, 则称

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

为由向量范数派生的**矩阵算子范数**.

矩阵的算子范数满足形如下面的不等式:

$$\|Ax\| \le \|A\|\,\|x\|,\ \forall x \in \mathrm{R}^n.$$

把满足此式的矩阵范数称为**与向量范数相容的矩阵范数.**

注: 矩阵算子范数必定与某个向量范数相容, 但**与某向量范数相容的矩阵范数未必是算子范数(原因见后面).**

命题: 矩阵范数

$$\|\boldsymbol{A}\|_p = \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p}, \quad p = 1, 2, \infty,$$

是由向量范数$\|\bullet\|_p$ (p=1, 2, ∞)分别派生的矩阵算子范数.

需严格证明. 参见有关文献.

因此, 对于任意n阶方阵$\boldsymbol{A}$与n维向量$\boldsymbol{x}$, 必有

$$\|\boldsymbol{A}\boldsymbol{x}\|_1 \le \|\boldsymbol{A}\|_1 \|\boldsymbol{x}\|_1,$$

$$\|\boldsymbol{A}\boldsymbol{x}\|_2 \le \|\boldsymbol{A}\|_2 \|\boldsymbol{x}\|_2,$$

$$\|\boldsymbol{A}\boldsymbol{x}\|_\infty \le \|\boldsymbol{A}\|_\infty \|\boldsymbol{x}\|_\infty,$$

所以$\|\boldsymbol{A}\|_p$是与$\|\boldsymbol{x}\|_p$ (p=1, 2, ∞)相容的矩阵范数.

性质: 设$\|\bullet\|$是任意算子范数, 则对于单位矩阵$\boldsymbol{E}$必有

$$\|\mathbf{E}\| = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{E}\mathbf{x}\|}{\|\mathbf{x}\|} = 1$$

但是, 对于$\mathrm{R}^{n\times n}$上的矩阵F-范数$\|\bullet\|_{\mathrm{F}}$, 根据其定义可知

$$\|\mathbf{E}\|_{\mathrm{F}} = \sqrt{n}.$$

这说明: $\|\bullet\|_{\mathrm{F}}$不是一种算子范数.

$\|\bullet\|_{\mathrm{F}}$的性质:

1. $\|\bullet\|_{\mathrm{F}}$不是一种算子范数;
2. $\|\bullet\|_{\mathrm{F}}$与向量的2-范数相容, 即对任意n阶方阵$\boldsymbol{A}$与n维向量$\boldsymbol{x}$, 必有

$$\|\boldsymbol{A}\boldsymbol{x}\|_2 \le \|\boldsymbol{A}\|_{\mathrm{F}}\,\|\boldsymbol{x}\|_2.$$

关于矩阵的F-范数与2-范数的一个重要性质:

对于任意方阵$\boldsymbol{A}$以及与它同阶的正交矩阵$\boldsymbol{Q}$, 必有:

(1) $\|\boldsymbol{QA}\|_F = \|\boldsymbol{AQ}\|_F = \|\boldsymbol{A}\|_F$;

(2) $\|\boldsymbol{QA}\|_2 = \|\boldsymbol{AQ}\|_2 = \|\boldsymbol{A}\|_2$.

以上性质留作思考题.

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

$$|\lambda|\|\boldsymbol{x}\|=\|\lambda\boldsymbol{x}\|=\|\boldsymbol{Ax}\|\le\|\boldsymbol{A}\|\|\boldsymbol{x}\|,$$

于是 $|\lambda|\le\|\boldsymbol{A}\|$.

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

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

为矩阵$\boldsymbol{A}$的**谱半径**.

由上述推导过程可知, 对于**与向量范数相容的矩阵范数**, 有

$$\rho(\boldsymbol{A})\le\|\boldsymbol{A}\|.$$

补充: 事实上还可以推得更一般的结论, 即:

命题: 对于满足性质(4)(**矩阵范数的相容性: $\|AB\|\le\|A\|\,\|B\|$**)的**任意**矩阵范数$\|\bullet\|$, 以及**任意**的方阵A, 必成立:

$$\rho(A)\le\|A\|.$$

注: 根据本教材的定义, 所有的矩阵范数必然满足性质(4).

以上命题的具体推导详见教师讲解.

此外还存在另一结论, 即

命题: 对于任意给定的矩阵A, $\forall\varepsilon>0$, 必存在某相容矩阵范数$\|\bullet\|_\beta$ (此范数一般与A和ε有关), 使

$$\|A\|_\beta\le\rho(A)+\varepsilon.$$

具体证明请参见相关文献.

请根据上述结论思考矩阵谱半径与矩阵范数之间的关系.

任何两种矩阵范数也具有等价性, 即

命题: 对于$\mathbb{R}^{n\times n}$上的任何两种矩阵范数$\|\bullet\|_\alpha$和$\|\bullet\|_\beta$, 存在正的常数m和M, 使得

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

矩阵序列的收敛可也定义为

$$\lim_{k\to\infty}\mathbf{A}^{(k)}=\mathbf{A}^{*} \stackrel{\Delta}{=} \lim_{k\to\infty}\left\|\mathbf{A}^{(k)}-\mathbf{A}^{*}\right\|=0$$

$$\Leftrightarrow\ a_{ij}^{(k)}\longrightarrow a_{ij}^{*}, 1\le i,j\ \le n.$$

例题练习:

设$\|\cdot\|$是$\mathrm{R}^{n\times n}$上的一个算子范数. 如果$\boldsymbol{A}\in \mathrm{R}^{n\times n}$, $\|\boldsymbol{A}\|<1$, 则矩阵$\boldsymbol{I}\pm\boldsymbol{A}$可逆, 且有

$$\|(\boldsymbol{I}\pm\boldsymbol{A})^{-1}\|\leq\frac{1}{1-\|\boldsymbol{A}\|}.$$

思考, 或见黑板讨论.

§ 2.5 线性方程组固有性态与误差分析

§ 2.5.1 线性方程组的固有性态

考虑线性方程组

$$\begin{pmatrix} 1 & 0.99 \\ 0.99 & 0.98 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

其精确解为 $\boldsymbol{x}^*=(-9800b_1+9900b_2, 9900b_1-10000b_2)^T$.

若把线性方程组变为

$$\begin{pmatrix} 1 & 0.99 \\ 0.99 & 0.98 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1+\varepsilon \\ b_2-\varepsilon \end{pmatrix}$$

解为$\boldsymbol{x}=(-9800b_1+9900b_2-19700\varepsilon, 9900b_1-10000b_2+19900\varepsilon)^T$

可见: $\boldsymbol{x}-\boldsymbol{x}^*=(-19700\varepsilon, 19900\varepsilon)^T$.

解的误差可能放大到常数项的误差的近2万倍.

这种由于原始数据微小变化而导致解严重失真的线性方程组称为**病态方程组**, 相应的系数矩阵称为**病态矩阵.**

以下我们将在两种情况下导出可以反映方程组病态程度的一个观测量——条件数.

设线性方程组$\boldsymbol{Ax}=\boldsymbol{b}$的系数矩阵是精确的，右端常数项有误差$\Delta\boldsymbol{b}$，此时记解为$\boldsymbol{x}+\Delta\boldsymbol{x}$，则

$$\boldsymbol{A}(\boldsymbol{x}+\Delta\boldsymbol{x})=\boldsymbol{b}+\Delta\boldsymbol{b}.$$

于是

$$\boldsymbol{A}\Delta\boldsymbol{x}=\Delta\boldsymbol{b}.$$

所以

$$\|\Delta\boldsymbol{x}\|=\|\boldsymbol{A}^{-1}\Delta\boldsymbol{b}\|\le\|\boldsymbol{A}^{-1}\|\,\|\Delta\boldsymbol{b}\|.$$

又由于 $\|\boldsymbol{b}\|=\|\boldsymbol{Ax}\|\le\|\boldsymbol{A}\|\,\|\boldsymbol{x}\|$，因此

$$\|\Delta\boldsymbol{x}\|\|\boldsymbol{b}\|\le\|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\|\|\Delta\boldsymbol{b}\|\|\boldsymbol{x}\|,$$

即得

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

再设$\boldsymbol{b}$是精确的, $\boldsymbol{A}$有误差$\Delta\boldsymbol{A}$, 此时记解为$\boldsymbol{x}+\Delta\boldsymbol{x}$, 则

$$(\boldsymbol{A}+\Delta\boldsymbol{A})(\boldsymbol{x}+\Delta\boldsymbol{x})=\boldsymbol{b}.$$

则有

$$\boldsymbol{A}\Delta\boldsymbol{x}+\Delta\boldsymbol{A}(\boldsymbol{x}+\Delta\boldsymbol{x})=\mathbf{0}.$$

所以

$$\Delta\boldsymbol{x}=-\boldsymbol{A}^{-1}\Delta\boldsymbol{A}(\boldsymbol{x}+\Delta\boldsymbol{x}).$$

于是

$$\|\Delta\boldsymbol{x}\|\le\|\boldsymbol{A}^{-1}\|\|\Delta\boldsymbol{A}\|\|\boldsymbol{x}+\Delta\boldsymbol{x}\|.$$

也就是

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

定义: 当矩阵A可逆时, 记$\mathrm{Cond}(A)=\|A\|\|A^{-1}\|$, 称其为方程组$Ax=b$或矩阵$A$的**条件数**.

经常使用的条件数有

$$\mathrm{Cond}_p(A)=\|A\|_p\|A^{-1}\|_p,\quad p=1,2,\infty.$$

当A为**对称矩阵**时, 有

$$\mathrm{Cond}_2(A)=|\lambda_1|/|\lambda_n|,$$

其中λ_1, λ_n分别是A的按绝对值最大和最小特征值.

请思考其证明过程. (参见黑板的讨论)

推论: 若A为**对称正定矩阵**, 则$\mathrm{Cond}_2(A)=\lambda_1/\lambda_n$, 其中$\lambda_1$, λ_n分别是A的最大和最小特征值.

例如, 对前面方程组的系数矩阵A有

$$\mathbf{A}=\begin{pmatrix}1 & 0.99\\ 0.99 & 0.98\end{pmatrix},\mathbf{A}^{-1}=-\frac{1}{0.0001}\begin{pmatrix}0.98 & -0.99\\ -0.99 & 1\end{pmatrix}$$

$$\mathrm{Cond}_1(A)=\mathrm{Cond}_\infty(A)=39601,\ \mathrm{Cond}_2(A)\approx 39206.$$

由于计算条件数运算量较大, 实际计算中若遇到下述情况之一, 方程组就有可能是病态的:

(1) 矩阵元素间数量级差很大, 且无一定规律;

(2) 矩阵的行列式值相对来说很小 ;

(3) 列主元消去法求解过程中出现量级很小的主元素;

(4) 数值求解过程中, 计算解x的剩余向量$r=b-Ax$已经很小, 但x仍不符合要求.

§ 2.5.2 预条件和迭代改善

1. 线性方程组的预条件处理

对病态方程组$\boldsymbol{Ax}=\boldsymbol{b}$,

方案(I) 考虑线性方程组 $\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}=\tilde{\boldsymbol{b}}$

其中 $\tilde{\boldsymbol{A}}=\boldsymbol{PA},\ \ \tilde{\boldsymbol{x}}=\boldsymbol{x},\ \ \tilde{\boldsymbol{b}}=\boldsymbol{Pb}.$

方案(II)考虑线性方程组 $\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}=\tilde{\boldsymbol{b}}$

其中 $\tilde{\boldsymbol{A}}=\boldsymbol{PAQ},\ \tilde{\boldsymbol{x}}=\boldsymbol{Q}^{-1}\boldsymbol{x},\ \tilde{\boldsymbol{b}}=\boldsymbol{Pb}.$

称 $\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}=\tilde{\boldsymbol{b}}$ 为**预条件方程组**. 显然它与原方程组$\boldsymbol{Ax}=\boldsymbol{b}$等价. 可逆矩阵$\boldsymbol{P}$或$\boldsymbol{Q}$称为**预条件矩阵**. 它们应满足条件:

(1) $\mathrm{Cond}(\tilde{\boldsymbol{A}})\ll\mathrm{Cond}(\boldsymbol{A}).$

(2) 预条件矩阵$\boldsymbol{P}$, $\boldsymbol{Q}$具有某些性质使得相应的方程组$\boldsymbol{Pz}=\boldsymbol{d}$或 $\boldsymbol{Qz}=\boldsymbol{d}$ 容易求解.

2. 线性方程组的迭代改善

设已求得方程组$\boldsymbol{Ax}=\boldsymbol{b}$的近似解$\boldsymbol{x}^{(1)}$, 计算剩余向量

$$\boldsymbol{r}^{(1)}=\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}^{(1)}.$$

再求解余量方程组$\boldsymbol{Ax}=\boldsymbol{r}^{(1)}$, 得到解 $\tilde{\boldsymbol{x}}^{(1)}$,

$\boldsymbol{x}^{(1)}$的迭代改善解为:

$$\boldsymbol{x}^{(2)}=\boldsymbol{x}^{(1)}+\tilde{\boldsymbol{x}}^{(1)}.$$

补充: 关于矩阵的2-范数、F-范数和谱半径的一些性质

矩阵的2-范数: $\|\boldsymbol{A}\|_2 = \sqrt{\max \lambda(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})}$.

由于$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$必为对称半正定矩阵(原因详见教师的解释), 故其全体特征值均为非负实数.

因此, 矩阵$\boldsymbol{A}$的2-范数也可以写为:

$$\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})}.$$

其中ρ(.)表示矩阵的谱半径.

而当$\boldsymbol{A}$为对称矩阵时, 易见.

$$\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^2)} = \sqrt{[\rho(\boldsymbol{A})]^2} = \rho(\boldsymbol{A}).$$

$$\boldsymbol{A}^{\mathrm{T}}=\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix},\ \boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

矩阵的F-范数：

$$\|\boldsymbol{A}\|_{\mathrm{F}}=\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^{2}}$$

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{bmatrix} \sum_{k=1}^{n}a_{k1}^{2} & * & \cdots & * \\ * & \sum_{k=1}^{n}a_{k2}^{2} & \cdots & * \\ \cdots & \cdots & \ddots & \cdots \\ * & * & \cdots & \sum_{k=1}^{n}a_{kn}^{2} \end{bmatrix}.$$

若记 tr(.) 为矩阵的迹(即矩阵的对角线元素之和), 则由上述讨论可知:

$$\operatorname{tr}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{kj}^2.$$

因此可得:

$$\|\boldsymbol{A}\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2} = \sqrt{\operatorname{tr}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})}.$$

再由矩阵的迹与特征值之间的关系

$$= \sqrt{\sum_{\mathrm{k}=1}^{\mathrm{n}} \lambda_k(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})},$$

其中$\lambda_k(.)$代表矩阵的第k个实特征值.

关于矩阵的谱半径ρ(.)的一些性质

1. 对于任意的$A\in R^{n\times n}$, $\rho(A)$是矩阵A的函数.

2. 对于一般的矩阵$A\in R^{n\times n}$, $\rho(A)$不是A的矩阵范数.

请思考反例.

3. 当矩阵A为对称矩阵时，由于此时有

$$\rho(A) = \|A\|_2,$$

因而此时$\rho(A)$ 是A的一个矩阵范数.