第五章 解线性方程组的直接法

5.1 直接法与三角形方程组求解

5.1.1 线性方程组的分类

不同类型的方程组由不同的解法:

按**系数矩阵零元素**多少分:稠密（元素满）、稀疏（占生活中的绝大多数 80%，如推荐系统）

按**系数矩阵阶数**分:高阶 4000 阶以上、**低阶**

按**系数矩阵形状**分:**对称正定(对称$A^T=A$,正定$X^TAX=E$)、对角线**

对角占优（对角阵元素的绝对值比其他元素大，一定有解且唯一）、**三角形**方程组

5.1.2 三角形线性方程组的解法

下三角形　**形式**: $\begin{cases} l_{11}x_1 & = b_1 \\ l_{21}x_1 + l_{22}x_2 & = b_2 \\ \cdots\cdots & \cdots \\ l_{n1}x_1 + l_{n2}x_2 + \cdots + l_{nn}x_n & = b_n \end{cases}$ 简写为 $\boldsymbol{Lx=b}$

解: $\begin{cases} x_1 = \frac{b_1}{l_{11}} \\ x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij}x_j}{l_{ii}} \quad (i=2,3,\ldots,n) \end{cases}$ 先解出x_1，不断迭代解出x_i（**对角阵元素不能为零**）

上三角形　**形式**: $\begin{cases} u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n & = b_1 \\ u_{22}x_2 + \cdots + u_{2n}x_n & = b_2 \\ \cdots\cdots & \cdots \\ u_{nn}x_n & = b_n \end{cases}$ 简写为 $\boldsymbol{Ux=b}$

解: $\begin{cases} x_n = b_n/u_{nn} \\ x_i = \frac{b_i - \sum_{j=i+1}^{n} u_{ij}x_j}{u_{ii}} \quad (i=n-1,n-2,\ldots,2,1) \end{cases}$

5.1.3 直接法概述

高斯消去法　设方程组$Ax=b$，其中$A=(a_{ij})_{n\times n}$，$x=(x_1,x_2,\ldots,x_n)^T$　$b=(b_1,b_2,\ldots,b_n)^T$

增广矩阵: $(A|b) \xrightarrow{\text{行变换}} (A^{(n)}|b^{(n)})$　其中$A^{(n)}$为**上三角阵**，则$A^{(n)}x=b^{(n)}$与原方程组同解

称此方法为**高斯消去法**

三角分解法　若有$A=LU$（任意矩阵皆可如此变换)，其中L,U为三角阵，则有

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

通常L为下三角(将原阵对角线下元素消为零的**倍数**)，U为上三角。此法为三角分解法

5.2 高斯消去法

5.2.1 消元与回代计算

消元　例如：$(A|b)=\left(\begin{array}{ccc|c}1&1&1&6\\1&3&-2&1\\2&-2&1&1\end{array}\right)\xrightarrow[r_3-2r_1]{r_2-r_1}\left(\begin{array}{ccc|c}1&1&1&6\\0&2&-3&-5\\0&-4&-1&-11\end{array}\right)\xrightarrow{r_3+2r_2}\left(\begin{array}{ccc|c}1&1&1&6\\0&2&-3&-5\\0&0&-7&-21\end{array}\right)$ **对角行变换**

将其变为**上三角阵**。一般情况，设$\det A\neq 0$

① **第一次消元**：$\left(A^{(1)}|b^{(1)}\right)=\begin{pmatrix}a_{11}^{(1)}&a_{12}^{(1)}&\cdots&a_{1n}^{(1)}&b_1^{(1)}\\a_{21}^{(1)}&a_{22}^{(1)}&\cdots&a_{2n}^{(1)}&b_2^{(1)}\\\cdots&\cdots&\cdots&\cdots&\cdots\\a_{n1}^{(1)}&a_{n2}^{(1)}&\cdots&a_{nn}^{(1)}&b_n^{(1)}\end{pmatrix}\to\begin{pmatrix}a_{11}^{(1)}&a_{12}^{(1)}&\cdots&a_{1n}^{(1)}&b_1^{(1)}\\0&a_{22}^{(2)}&\cdots&a_{2n}^{(2)}&b_2^{(2)}\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&a_{n2}^{(2)}&\cdots&a_{nn}^{(2)}&b_n^{(2)}\end{pmatrix}=\left(A^{(2)}|b^{(2)}\right)$

其中：$a_{ij}^{(2)}=a_{ij}^{(1)}-\underbrace{\frac{a_{i1}^{(1)}}{a_{11}^{(1)}}}_{\text{比值}}\underbrace{a_{1j}^{(1)}}_{\text{第一行}}\quad j=1,2,\dots,n\qquad b_i^{(2)}=b_i^{(1)}-\frac{a_{i1}^{(1)}}{a_{11}^{(1)}}b_1^{(1)}\quad i=2,3,\dots,n$

注：若$a_{11}^{(1)}$，则在第一列中至少有一个元素不等于 0，可交换该行后再消元

② **第k次消元**：$\left(A^{(k)}|b^{(k)}\right)=\begin{pmatrix}a_{11}^{(1)}&\cdots&a_{1k}^{(1)}&\cdots&a_{1n}^{(1)}&b_1^{(1)}\\&\ddots&\cdots&\cdots&\cdots&\cdots\\&&a_{kk}^{(k)}&\cdots&a_{kn}^{(k)}&b_k^{(k)}\\&&\cdots&\cdots&\cdots&\cdots\\&&a_{nk}^{(k)}&\cdots&a_{nn}^{(k)}&b_n^{(k)}\end{pmatrix}$

$$\to\begin{pmatrix}a_{11}^{(1)}&\cdots&a_{1k}^{(1)}&a_{1(k+1)}^{(1)}&\cdots&a_{1n}^{(1)}&b_1^{(1)}\\&\ddots&\vdots&\vdots&\cdots&\vdots&\vdots\\&&a_{kk}^{(k)}&a_{k(k+1)}^{(k)}&\cdots&a_{kn}^{(k)}&b_k^{(k)}\\&&0&a_{(k+1)(k+1)}^{(k+1)}&\cdots&a_{(k+1)n}^{(k+1)}&b_{k+1}^{(k+1)}\\&&\vdots&\vdots&\cdots&\vdots&\vdots\\&&0&a_{n(k+1)}^{(k+1)}&\cdots&a_{nn}^{(k+1)}&b_n^{(k+1)}\end{pmatrix}=\left(A^{(k+1)}|b^{(k+1)}\right)$$

其中：$a_{ij}^{(k+1)}=\underbrace{a_{ij}^{(k)}}_{\text{自己行}}-\underbrace{\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}}_{\text{倍数}}\underbrace{a_{kj}^{(k)}}_{\text{对应行}}\quad j=k+1,k+2,\dots,n$

$b_i^{(k+1)}=b_i^{(k)}-\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}b_k^{(k)}\qquad i=k+1,k+2,\dots,n$

为了**减少计算量**，令单独变量倍数$m_{ik}=\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$　则$a_{ij}^{(k+1)}=a_{ij}^{(k)}-m_{ik}\,a_{kj}^{(k)}\quad j=k+1,k+2,\dots,n$

$b_i^{(k+1)}=b_i^{(k)}-m_{ik}\,b_k^{(k)}\qquad i=k+1,k+2,\dots,n$

③ **主元素**：　当$k=n-1$时，得到**上三角阵** $\left(A^{(n)}|b^{(n)}\right)=\begin{pmatrix}a_{11}^{(1)}&a_{12}^{(1)}&\cdots&a_{1n}^{(1)}&b_1^{(1)}\\&a_{22}^{(2)}&\cdots&a_{2n}^{(2)}&b_2^{(2)}\\&&\ddots&\cdots&\cdots\\&&&a_{nn}^{(n)}&b_n^{(n)}\end{pmatrix}$

当$|A|\neq 0$时，$a_{ii}^{(i)}\neq 0$ 称为主元素

回代　对于主元素$\left(A^{(n)}|b^{(n)}\right)=\begin{pmatrix}a_{11}^{(1)}&a_{12}^{(1)}&\cdots&a_{1n}^{(1)}&b_1^{(1)}\\&a_{22}^{(2)}&\cdots&a_{2n}^{(2)}&b_2^{(2)}\\&&\ddots&\cdots&\cdots\\&&&a_{nn}^{(n)}&b_n^{(n)}\end{pmatrix}$　$\begin{cases}x_n=\frac{b_n^{(n)}}{a_{nn}^{(n)}}\\x_i=\frac{b_i^{(i)}-\sum_{j=i+1}^{n}a_{ij}^{(i)}x_j}{a_{ii}^{(i)}}\end{cases}(i=n-1,n-2,\dots,2,1)$

注意：　若 A 非奇异，则上述过程可行，但**无法事先判断**。故可能

① 某一步找不到非零元

② 到最后一步得$a_{nn}^{(n)}=0$，无法回代，需在计算中中断

5.2.2 gauss 消去法的运算量

$$m_{ik}=\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$$

消元过程　在**第k步**中：$a_{ij}^{(k+1)}=a_{ij}^{(k)}-\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}a_{kj}^{(k)}\quad j=k+1,k+2,\dots,n$

$b_i^{(k+1)}=b_i^{(k)}-\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}b_k^{(k)}\quad i=k+1,k+2,\dots,n$

除法：$n-k$　　**乘法**：$(n-k)(n-k+1)$　　**共**：$(n-k)(n-k+2)$次

共$n-1$步，所以有$\sum_{k=1}^{n-1}(n-k)(n-k+2)=\boldsymbol{\frac{n^3}{3}}+\frac{n^2}{2}-\frac{5}{6}n$

回代过程

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

求x_i中，**乘法**$n-i$次，**除**1次，**共**$n-i+1$次，所以**共有** $\sum_{k=1}^{n-1}(n-i+1)+1=\boldsymbol{\frac{n^2}{2}}-\frac{n}{2}$

总和次数　$\boldsymbol{MD}=\frac{n^3}{3}+n^2-\frac{n}{3}=\boldsymbol{\frac{n^3}{3}+O(n^2)}$

可以看到，当$n=20$时，总次数约为 2670 次，比使用克莱姆法则9.7×10^{20}次极端减少。

5.3 高斯列主元素消去法

5.3.1 主元素的作用

作用　在消元过程中以主元素$a_{kk}^{(k)}$依次作除法计算：$m_{ik}=\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$

注意　① 若消元过程中对角线元素出现$a_{kk}^{(k)}=0$，则消元无法继续

② 若$a_{kk}^{(k)}\neq0$，但其值较小，则会产生较大误差

③ 为**避免小主元作除数**，在$A^{(k)}$的第k列主对角线以下元素$a_{ik}^{(k)}(k\le i\le n)$中**挑选绝对值最大者**，并通过行变换，使之位于主对角线上作为主元素，仍记为$a_{kk}^{(k)}$，然后在进行消元计算。即$a_{kk}^{(k)}=\max_{k\le i\le n}|a_{ik}^{(k)}|$

第k列主对角元以下元素绝对值最大者作主元（该行与第k行**整体对调**），称每一步都按列选主元的消去法为 **gauss 列主元素消去法**

例题　1. 用舍入三位有效数字求解线性方程组 $\begin{cases}0.0300x_1+58.9x_2=59.2\\5.31x_1-6.10x_2=47.0\end{cases}$

① **不选主元**的 Gauss 消去法计算结果：$x_1=-10.0,x_2=1.01$，此解无效

② **按列选主元**的 Gauss 消去法计算结果：$x_1=10.0,x_2=1.00$，此解有效

5.3.2 算法设计

本节略去

详细内容参阅教学课件

5.3.3 消元过程与系数矩阵的分解

不带行变换的消元过程

第k步：$\left(A^{(k)}\middle|b^{(k)}\right)\xrightarrow[i=k+1,k+2,\dots,n]{r_i-m_{ik}r_k}\left(A^{(k+1)}\middle|b^{(k+1)}\right)$ 相当于用**一系列的系数矩阵左乘**：$L_k\left(A^{(k)}\middle|b^{(k)}\right)=\left(A^{(k+1)}\middle|b^{(k+1)}\right)$

其中：$L_k=\begin{pmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -m_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -m_{n,k} & 0 & \cdots & 1\end{pmatrix}$，$k=1,2,\dots,n$

可得到$L_{n-1}L_{n-2}\dots L_2L_1A^{(1)}=A^{(n)}\Rightarrow A=A^{(1)}=(L_{n-1}L_{n-2}\dots L_2L_1)^{-1}A^{(n)}=LU$

其中$L=L_1^{-1}L_2^{-1}\dots L_{n-1}^{-1}=\begin{pmatrix}1 & & & & \\ m_{21} & 1 & & & \\ m_{31} & m_{32} & 1 & & \\ \dots & \dots & \dots & \ddots & \\ m_{n1} & m_{n2} & m_{n3} & \dots & 1\end{pmatrix}$称为**单位下三角阵**

$U=A^{(n)}=\begin{pmatrix}a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ & & \ddots & \dots \\ & & & a_{nn}^{(n)}\end{pmatrix}$称为**上三角阵**。 显然有$|A|=|U|$，称$\boldsymbol{A=LU}$**为矩阵**$\boldsymbol{A}$**的**$\boldsymbol{LU}$**分解**

注意：由归纳法可得：$a_{ii}^{(i)}\neq 0\iff A$的顺序主子式$D_1,D_2,\dots,D_k\neq 0$

定理 3.1 基本定理 若A的顺序主子式$D_1,D_2,\dots,D_{n-1}\neq 0$，则$A$的$LU$分解存在且唯一。

带行交换的消元过程

换行： 记 $I_{ik}=\begin{pmatrix}1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 0 & \cdots & 1 & & \\ & & & \vdots & \ddots & \vdots & & \\ & & & 1 & \cdots & 0 & & \\ & & & & & & 1 & \\ & & & & & & & \ddots \\ & & & & & & & & 1\end{pmatrix}\begin{array}{l} \\ \\ \\ 第k行 \\ \\ 第i行 \\ \\ \\ \\ \end{array}$ $(i>k)$则交换$\left(A^{(k)}\middle|b^{(k)}\right)$中的$k$行和$i$行再消元可以表示为

$L_kI_{ik}\left(A^{(k)}\middle|b^{(k)}\right)=\left(A^{(k+1)}\middle|b^{(k+1)}\right)$

设当$n=4$时，有$L_3I_{i_33}L_2I_{i_22}L_1I_{i_11}A=A^{(4)}=U\Rightarrow L_3\left(I_{i_33}L_2I_{i_33}\right)\left(I_{i_33}I_{i_22}L_1I_{i_22}I_{i_33}\right)\left(I_{i_33}I_{i_22}I_{i_11}\right)A=U$

令 $\tilde{L}_2=I_{i_33}L_2I_{i_33}$，$\tilde{L}_1=I_{i_33}I_{i_22}L_1I_{i_22}I_{i_33}$，$P=I_{i_33}I_{i_22}I_{i_11}$ 称为**排列阵**，可以**表示所有的行变换**。

易知当$i,k>k$时，$\tilde{L}_k=I_{ij}L_kI_{ij}$与$L_k$形状相同，只是交换了第$k$列对角线以下的$i$、$j$行元素，故$\tilde{L}=L_3\widetilde{L_2}\widetilde{L_1}$仍为单位下三角阵。

一般地有$L_{n-1}\tilde{L}_{n-2}\dots\widetilde{L_2}\widetilde{L_1}PA=U$，从而$PA=LU$，其中$L=\left(L_{n-1}\tilde{L}_{n-2}\dots\widetilde{L_2}\widetilde{L_1}\right)^{-1}$为单位下三角阵

定理 3.2 若$|A|\neq 0$，则存在排列阵P，以及单位下三角阵L和非奇异上三角阵U，使得$PA=LU$