线性方程组的数值解法

Lucian Xu (app1eDog)

2023 年 6 月 1 日

线性方程组的数值解法

Gauss 消去法, Gauss 列主元消去法

Cramer 法则

考虑一般的线性方程组

$$Ax = b,$$

如果 $\det A \neq 0$ 则上述方程有唯一解 $x_i = \frac{|A_i|}{|A|}$.

Gauss 消去法

将线性方程组

$$Ax = b,$$

用增广矩阵表示, 记为

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

接着进行 $n-1$ 步消元将其化为上三角矩阵. 例如

$$\bar{\mathbf{A}}^{(1)} \to \bar{\mathbf{A}}^{(2)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{i2}^{(2)} & \cdots & a_{in}^{(2)} & a_{i,n+1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & a_{n,n+1}^{(2)} \end{pmatrix}.$$

最后得到

$$\bar{\mathbf{A}}^{(1)} \to \bar{\mathbf{A}}^{(2)} \to \bar{\mathbf{A}}^{(3)} \to \cdots \to \bar{\mathbf{A}}^{(k)} \to \cdots \to \bar{\mathbf{A}}^{(n)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,n-1}^{(1)} & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2,n-1}^{(2)} & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ 0 & 0 & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} & a_{n-1,n+1}^{(n-1)} \\ 0 & 0 & 0 & 0 & a_{nn}^{(n)} & a_{n,n+1}^{(n)} \end{pmatrix}.$$

如果记

$$U=\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,n-1}^{(1)} & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2,n-1}^{(2)} & a_{2n}^{(2)} \\ 0 & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} \\ 0 & 0 & 0 & 0 & a_{nn}^{(n)} \end{pmatrix}, y=\begin{pmatrix} a_{1,n+1}^{(1)} \\ a_{2,n+1}^{(2)} \\ \vdots \\ a_{n-1,n+1}^{(n-1)} \\ a_{n,n+1}^{(n)} \end{pmatrix},$$

那么原方程等价于 $Ux=y$, 回代可求解.

整个计算过程时间复杂度为 $O(n^3)$.

Th 3.1

对于给定的线性方程组 $Ax=b$, 如果 A 的各阶顺序主子式非零, 那么 Gauss 消去法中的各阶主元 $a_{kk}^{(k)}(k=1,2,\cdots,n)$ 均非零.

三对角方程组的追赶法

考虑三对角方程组

$$\begin{pmatrix} b_1 & c_1 & & & & \\ a_2 & b_2 & c_2 & & & \\ & a_3 & b_3 & c_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & & a_n & b_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}=\begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_{n-1} \\ d_n \end{pmatrix},$$

其中系数矩阵方程满足

1. $|b_1|>|C_1|>0$;
2. $b_i\geqslant|a_i|+|c_i|, a_ic_i\neq0(i=2,3,\cdots,n-1)$;
3. $|b_n|>|a_n|>0$.

这是系数矩阵非奇异, 利用 Gauss 消去法每一步消元只需要消去一个元素, 计算量较小.

Gauss 列主元消去法

在第 k 次消元

$$\bar{A}^{(k)} \to \bar{A}^{(k+1)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1,k}^{(1)} & a_{1,k+1}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2,k}^{(2)} & a_{2,k+1}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3,k}^{(3)} & a_{3,k+1}^{(3)} & \cdots & a_{3n}^{(3)} & a_{3,n+1}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \cdots & a_{k,n}^{(k)} & a_{k,n+1}^{(k)} \\ 0 & 0 & 0 & \cdots & 0 & a_{k+1,k+1}^{(k+1)} & \cdots & a_{k+1,n}^{(k+1)} & a_{k+1,n+1}^{(k+1)} \\ 0 & 0 & 0 & \cdots & 0 & a_{k+2,k+1}^{(k+1)} & \cdots & a_{k+2,n}^{(k+1)} & a_{k+2,n+1}^{(k+1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & a_{i,k+1}^{(k+1)} & \cdots & a_{i,n}^{(k+1)} & a_{i,n+1}^{(k+1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & a_{n,k+1}^{(k+1)} & \cdots & a_{nn}^{(k+1)} & a_{n,n+1}^{(k+1)} \end{pmatrix}$$

时, 从第 k 列位于对角线一下的元素中选绝对值最大者作为主元, 例如, 如果 $|a_{sk}^{(k)}| = \max\limits_{k \leqslant i \leqslant n} |a_{ik}^{(k)}|$, 则交换第 s 行和第 k 行在进行校园.

矩阵的 LU 分解

Gauss 消去法第一步消元相当于计算 $L_1\bar{A}^{(1)} = \bar{A}^{(2)}$, 其中矩阵

$$\mathbf{L}_1 = \begin{pmatrix} 1 & & & & \\ -I_{21} & 1 & & & \\ -I_{31} & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \\ -I_{n1} & 0 & & \cdots & 1 \end{pmatrix}, I_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}, 2 \leqslant i \leqslant n.$$

以此类推有 $L_{n-1}L_{n-2}\cdots L_1\bar{A}^{(1)} = \bar{A}^{(n)}$, 记 $L = L_1^{-1}\ldots L_{n-1}^{-1}$ 可得 $A = LU$.

Th 3.2

对于给定的线性方程组 $Ax = b$, 如果 A 的各阶顺序主子式非零, 那么可以对 A 作唯一的 LU 分解.

方程组的性态与误差分析

Def 3.1

范数.

常用的三个范数:

1. 1 - 范数: $\|x\|_1 = \sum\limits_{i=1}^{n} |x_i|$;
2. ∞ - 范数: $\|x\|_\infty = \max\limits_{1 \leqslant i \leqslant n} |x_i|$;
3. 2 - 范数: $\|x\|_2 = \left(\sum\limits_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}}$.

Th 3.3

范数连续.

Def 3.2

范数等价.

Th 3.4

$\mathbb{R}^n$ 上任意两个范数等价.

Def 3.3

设 $\|\cdot\|$ 是 $\mathbb{R}^n$ 上的范数, x^* 是精确值, x 是近似值. 称 $\|x^* - x\|$ 为 x^* 绝对误差, $\frac{\|x^*-x\|}{\|x^*\|}$ 或 $\frac{\|x^*-x\|}{\|x\|}$ 为 x^* 绝对误差.

Def 3.4

(依范数) 收敛.

矩阵范数

称

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

为矩阵 A 的范数, 记为 $\|A\|$.

没有使用 $\sup$ 的原因是一定能保证最大值能取到. 事实上 $\|A\| = \max\limits_{x \in \mathbb{R}^n, \|x\|=1} \|Ax\|$, 根据范数的连续性, 在闭球上一定能取到最值.

矩阵范数的性质:

1. $\forall A \in \mathbb{R}^{n \times n}, \|A\| \geqslant 0$, 且 $\|A\| = 0 \Leftrightarrow A = 0$,
2. $\forall A \in \mathbb{R}^{n \times n}, \lambda \in \mathbb{R}$, 有 $\|\lambda A\| = |\lambda| \|A\|$,
3. $\forall A, B \in \mathbb{R}^{n \times n}$, 有 $\|A + B\| \leqslant \|A\| + \|B\|$,
4. $\forall A, B \in \mathbb{R}^{n \times n}$, 有 $\|AB\| \leqslant \|A\| \|B\|$,
5. $\forall A \in \mathbb{R}^{n \times n}, x \in \mathbb{R}^n$, 有 $\|Ax\| \leqslant \|A\| \|x\|$.

Def 3.6

对于 $A \in \mathbb{R}^{n\times n}, \lambda_1, \lambda_2, \cdots, \lambda_n$ 为 A 的 n 个特征值. 称

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

为矩阵 A 的谱半径.

Th 3.5

设 $A \in \mathbb{R}^{n\times n}$, 则

1. $\|A\|_1 = \max\limits_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max\limits_{1\leqslant j\leqslant n} \sum\limits_{i=1}^{n} |a_{ij}|$,
2. $\|A\|_\infty = \max\limits_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max\limits_{1\leqslant i\leqslant n} \sum\limits_{j=1}^{n} |a_{ij}|$,
3. $\|A\|_2 = \max\limits_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sqrt{\rho(AA^T)}$.

Th 3.6

$\mathbb{R}^{n\times n}$ 中的矩阵的谱半径被其范数控制.

Th 3.7

$\mathbb{R}^{n\times n}$ 中的矩阵的谱半径等于其 2 - 范数.

Def 3.7

矩阵范数等价.

Th 3.8

$\mathbb{R}^{n\times n}$ 上任意两个矩阵范数都等价.

Def 3.8

矩阵依范数的距离.

Def 3.9

矩阵的 (依范数) 收敛.

Th 3.9

设 $A \in \mathbb{R}^{n\times n}$, 则 $\lim\limits_{k\to\infty} A^k = 0 \Leftrightarrow \rho(A) < 1$.

(条件数: 先挖坑, 过几天再填)

线性方程组的迭代法

迭代格式的构造

将方程 $Ax = b$ 改写成等价的方程

$$x = Bx + f,$$

进而产生迭代格式

$$x^{(k+1)} = Bx^{(k)} + f,$$

如果产生的序列 $\left\{x^{(k)}\right\}_{k=1}^{\infty}$ 收敛于 $\bar{x}$, 对迭代格式两侧取极限即得 $\bar{x}$ 是原方程的解.

接着问题就是如何构造迭代格式, 以及迭代格式何时收敛.

Jacobi 迭代格式

$$\begin{aligned}
x_1^{(k+1)} &= \frac{1}{a_{11}}\left(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)}\right), \\
x_2^{(k+1)} &= \frac{1}{a_{22}}\left(b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)}\right), \\
x_3^{(k+1)} &= \frac{1}{a_{33}}\left(b_3 - a_{31}x_1^{(k)} - a_{32}x_2^{(k)} - \cdots - a_{3n}x_n^{(k)}\right), \\
&\cdots \\
x_n^{(k+1)} &= \frac{1}{a_{nn}}\left(b_n - a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - \cdots - a_{n,n-1}x_{n-1}^{(k)}\right).
\end{aligned}$$

给出其矩阵形式, 记

$$L=\begin{pmatrix}0&&&&&\\a_{21}&0&&&&\\a_{31}&a_{32}&0&&&\\\vdots&\vdots&\vdots&\ddots&&\\a_{n-1,1}&a_{n-1,2}&a_{n-1,3}&\cdots&0&\\a_{n1}&a_{n2}&a_{n3}&\cdots&a_{n,n-1}&0\end{pmatrix},$$

$$D=\begin{pmatrix}a_{11}&&&&\\&a_{22}&&&\\&&a_{33}&&\\&&&\ddots&\\&&&&a_{nn}\end{pmatrix},$$

$$U=\begin{pmatrix}0&a_{12}&a_{13}&\cdots&a_{1n}\\&0&a_{23}&\cdots&a_{2n}\\&&\ddots&\ddots&\vdots\\&&&0&a_{n,n-1}\\&&&&0\end{pmatrix}.$$

有

$$x^{(k+1)}=Jx^{(k)}+f_J,\ 其中J=-D^{-1}(L+U),f_J=D^{-1}b.$$

Gauss - Seidel 迭代格式

在 Jacobi 迭代中将已经求出的分量直接参与下一个分量的计算, 得到的就是 Gauss - Seidel 迭代.

$$\begin{aligned}
x_1^{(k+1)}&=\frac{1}{a_{11}}\left(b_1-a_{12}x_2^{(k)}-a_{13}x_3^{(k)}-\cdots-a_{1n}x_n^{(k)}\right),\\
x_2^{(k+1)}&=\frac{1}{a_{22}}\left(b_2-a_{21}x_1^{(k+1)}-a_{23}x_3^{(k)}-\cdots-a_{2n}x_n^{(k)}\right),\\
x_3^{(k+1)}&=\frac{1}{a_{33}}\left(b_3-a_{31}x_1^{(k+1)}-a_{32}x_2^{(k+1)}-\cdots-a_{3n}x_n^{(k)}\right),\\
&\cdots\\
x_n^{(k+1)}&=\frac{1}{a_{nn}}\left(b_n-a_{n1}x_1^{(k+1)}-a_{n2}x_2^{(k+1)}-\cdots-a_{n,n-1}x_{n-1}^{(k+1)}\right).
\end{aligned}$$

给出其矩阵形式

$$x^{(k+1)}=Gx^{(k)}+f_G,\ 其中G=-(D+L)^{-1}U,f_G=(D+L)^{-1}b.$$

SOR 迭代格式

将 Gauss - Seidel 迭代得到的 $x^{(k+1)}$ 与 $x^{(k)}$ 加权平均, 得到 SOR 迭代.

$$
\begin{aligned}
x_1^{(k+1)} &= (1-\omega)x_1^{(k)} + \frac{\omega}{a_{11}}\left(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)}\right), \\
x_2^{(k+1)} &= (1-\omega)x_2^{(k)} + \frac{\omega}{a_{22}}\left(b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)}\right), \\
x_3^{(k+1)} &= (1-\omega)x_3^{(k)} + \frac{\omega}{a_{33}}\left(b_3 - a_{31}x_1^{(k+1)} - a_{32}x_2^{(k+1)} - \cdots - a_{3n}x_n^{(k)}\right), \\
&\cdots \\
x_n^{(k+1)} &= (1-\omega)x_n^{(k)} + \frac{\omega}{a_{nn}}\left(b_n - a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - \cdots - a_{n,n-1}x_{n-1}^{(k+1)}\right).
\end{aligned}
$$

其中 ω 为松弛因子.

给出其矩阵形式

$$x^{(k+1)} = Sx^{(k)} + f_S,\ \text{其中} S = -(D+\omega L)^{-1}((1-\omega)D - \omega U), f_S = \omega(D+\omega L)^{-1}b.$$

迭代格式的收敛性

Th 3.11

迭代格式收敛等价于迭代矩阵的谱半径小于 1.

Def 3.12

按行严格对角占优, 按列严格对角占优, 严格对角占优.

Lemma

严格对角占优矩阵的行列式不为 0.

一些迭代格式的判别方法.

1. 如果 A 是严格对角占优矩阵, 那么 Jacobi 迭代格式和 Gauss - Seidel 迭代格式收敛.
2. 如果 A 是对称正定且 $\omega \in (0,2)$, 那么 SOR 迭代格式收敛.
3. 如果 A 是对称正定 Gauss - Seidel 迭代格式收敛.

幂法与反幂法

幂法可用于求模最大的特征值以及其对应的特征向量. 当特征值非零时, 反幂法可用于求模最小的特征值以及其对应的特征向量.

(先挖坑, 之后再填.)