线性方程组的古典迭代解法

邓建松

2018 年 10 月 29 日

直接法的局限

- 计算机的存储量日益增大，计算速度迅速提高，直接法（如Gauss消去法、平方根法）在计算机上可以求解的线性方程组的规模也越来越大

直接法的局限

- 计算机的存储量日益增大，计算速度迅速提高，直接法（如Gauss消去法、平方根法）在计算机上可以求解的线性方程组的规模也越来越大
- 在实际应用中，特别是偏微分方程的数值求解时，通常遇到的就是大型稀疏线性方程组

直接法的局限

- 计算机的存储量日益增大，计算速度迅速提高，直接法（如Gauss消去法、平方根法）在计算机上可以求解的线性方程组的规模也越来越大
- 在实际应用中，特别是偏微分方程的数值求解时，通常遇到的就是大型稀疏线性方程组
- 而直接法在对矩阵进行分解的时候，会破坏矩阵的稀疏性

直接法的局限

- 计算机的存储量日益增大，计算速度迅速提高，直接法（如Gauss消去法、平方根法）在计算机上可以求解的线性方程组的规模也越来越大
- 在实际应用中，特别是偏微分方程的数值求解时，通常遇到的就是大型稀疏线性方程组
- 而直接法在对矩阵进行分解的时候，会破坏矩阵的稀疏性
- 寻求能够保持稀疏性的有效算法是数值线性代数中一个重要的研究课题

求解稀疏线性方程组的方法

主要有两类

- 迭代法：按照某种规则构造一个向量序列，其极限是方程组的精确解

求解稀疏线性方程组的方法

主要有两类

- 迭代法：按照某种规则构造一个向量序列，其极限是方程组的精确解
- 稀疏直接法：是直接法与某些稀疏矩阵技巧有机结合的结果，利用矩阵的特点，使得分解结果尽可能保持稀疏性

求解稀疏线性方程组的方法

主要有两类

- 迭代法：按照某种规则构造一个向量序列，其极限是方程组的精确解
- 稀疏直接法：是直接法与某些稀疏矩阵技巧有机结合的结果，利用矩阵的特点，使得分解结果尽可能保持稀疏性
- 本课程只讲迭代法

迭代法中的问题

- 如何构造迭代序列？

迭代法中的问题

- 如何构造迭代序列？
- 构造的序列是否收敛？在什么情况下收敛？

迭代法中的问题

- 如何构造迭代序列？
- 构造的序列是否收敛？在什么情况下收敛？
- 如果收敛，收敛速度如何？（收敛速度的定量刻划）

迭代法中的问题

- 如何构造迭代序列？
- 构造的序列是否收敛？在什么情况下收敛？
- 如果收敛，收敛速度如何？（收敛速度的定量刻划）
- 迭代有限步停止。需要对近似解进行误差估计和舍入误差分析

迭代法有效性

- 方法是否有效要看得到具有某个精度的近似解而付出的代价如何

迭代法有效性

- 方法是否有效要看得到具有某个精度的近似解而付出的代价如何
- 这通常是以运算量和存储量的要求为标准

迭代法有效性

- 方法是否有效要看得到具有某个精度的近似解而付出的代价如何
- 这通常是以运算量和存储量的要求为标准
- 在这个标准下，很多时候直接法要比迭代法好

迭代法有效性

- 方法是否有效要看得到具有某个精度的近似解而付出的代价如何
- 这通常是以运算量和存储量的要求为标准
- 在这个标准下，很多时候直接法要比迭代法好
- 但对大型稀疏方程组来说，迭代法更实用

Jacobi迭代法

- 考虑非奇异线性方程组$Ax = b$

Jacobi迭代法

- 考虑非奇异线性方程组$Ax = b$
- 令$A = D - L - U$, 其中D为A的对角元构成的对角阵，$-L$为A的下三角阵，$-U$为A的上三角阵（均不含对角元）

Jacobi迭代法

- 考虑非奇异线性方程组$Ax = b$
- 令$A = D - L - U$, 其中D为A的对角元构成的对角阵，$-L$为A的下三角阵，$-U$为A的上三角阵（均不含对角元）
- 则$Ax = b$可写为$x = Bx + g$, 其中$B = D^{-1}(L + U)$, $g = D^{-1}b$

迭代格式

- 取初始向量

$$x_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})^T$$

迭代格式

- 取初始向量

$$x_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})^T$$

- 代入$x = Bx + g$的右边，得到新向量

$$x_1 = Bx_0 + g$$

迭代格式

- 取初始向量

$$x_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})^T$$

- 代入$x = Bx + g$的右边，得到新向量

$$x_1 = Bx_0 + g$$

- 再把x_1代入右边，又得一个新向量x_2

迭代格式

- 取初始向量

$$x_0 = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})^T$$

- 代入$x = Bx + g$的右边，得到新向量

$$x_1 = Bx_0 + g$$

- 再把x_1代入右边，又得一个新向量x_2
- 依此类推，我们有$x_k = Bx_{k-1} + g$，$k = 1, 2, \ldots$

迭代矩阵

- 这就是Jacobi迭代法，由C.G.J. Jacobi提出

迭代矩阵

- 这就是Jacobi迭代法，由C.G.J. Jacobi提出
- 其中B称为Jacobi迭代法的迭代矩阵，其对角元全是零

迭代矩阵

- 这就是Jacobi迭代法，由C.G.J. Jacobi提出
- 其中B称为Jacobi迭代法的迭代矩阵，其对角元全是零
- g称为Jacobi迭代法的常数项

例

- 设$A = \begin{pmatrix} 2 & 1 \\ 5 & 7 \end{pmatrix}$, $b = \begin{pmatrix} 11 \\ 13 \end{pmatrix}$, $x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

例

- 设$A = \begin{pmatrix} 2 & 1 \\ 5 & 7 \end{pmatrix}$, $b = \begin{pmatrix} 11 \\ 13 \end{pmatrix}$, $x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
- 则$D^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/7 \end{pmatrix}$, $L = \begin{pmatrix} 0 & 0 \\ -5 & 0 \end{pmatrix}$,
  $U = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$

例

- 设$A = \begin{pmatrix} 2 & 1 \\ 5 & 7 \end{pmatrix}$, $b = \begin{pmatrix} 11 \\ 13 \end{pmatrix}$, $x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
- 则$D^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/7 \end{pmatrix}$, $L = \begin{pmatrix} 0 & 0 \\ -5 & 0 \end{pmatrix}$,
$U = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$
- 迭代矩阵和常数项分别为

$$B = \begin{pmatrix} 0 & -1/2 \\ -5/7 & 0 \end{pmatrix}, \quad g = \begin{pmatrix} 11/2 \\ 13/7 \end{pmatrix}$$

例

- 迭代得到

$$x_1 = \begin{pmatrix} 5 \\ 8/7 \end{pmatrix} \approx \begin{pmatrix} 5 \\ 1.143 \end{pmatrix},$$

$$x_2 = \begin{pmatrix} 69/14 \\ -12/7 \end{pmatrix} \approx \begin{pmatrix} 4.929 \\ -1.714 \end{pmatrix}$$

例

- 迭代得到

$$x_1 = \begin{pmatrix} 5 \\ 8/7 \end{pmatrix} \approx \begin{pmatrix} 5 \\ 1.143 \end{pmatrix},$$

$$x_2 = \begin{pmatrix} 69/14 \\ -12/7 \end{pmatrix} \approx \begin{pmatrix} 4.929 \\ -1.714 \end{pmatrix}$$

- 25次迭代后，得到$x \approx \begin{pmatrix} 7.111 \\ -3.222 \end{pmatrix}$，这约等于方程的准确解$\begin{pmatrix} 64/9 \\ -29/9 \end{pmatrix}$

Jacobi迭代法代码片段


```
for i=1 to n
  y[i]=0.0
  for j=1 to n
    y[i]=y[i]+B[i][j]*x[j]
  y[i]=y[i]+g[i]
x=y
```

Jacobi迭代法中分量计算顺序

- 在Jacobi迭代法中各分量的计算顺序是没有关系的

Jacobi迭代法中分量计算顺序

- 在Jacobi迭代法中各分量的计算顺序是没有关系的
- 先算哪个分量，结果都不变

Jacobi迭代法中分量计算顺序

- 在Jacobi迭代法中各分量的计算顺序是没有关系的
- 先算哪个分量，结果都不变
- 我们做一下改变：在计算x_k的第一个分量用x_{k-1}的各个分量计算，但当计算x_k的后面分量时，采用已算出的新分量$x_1^{(k)}$代替$x_1^{(k-1)}$, 而其它分量仍用$x_i^{(k-1)}$

新代码片段


```
for i=1 to n
  x[i]=0.0
  for j=1 to n
    x[i]=x[i]+B[i][j]*x[j]
  x[i]=x[i]+g[i]
```

新格式

- 经上述改变后，我们有

$$x_k = D^{-1}Lx_k + D^{-1}Ux_{k-1} + g, \quad k = 1, 2, \ldots$$

新格式

- 经上述改变后，我们有

$$x_k = D^{-1}Lx_k + D^{-1}Ux_{k-1} + g, \quad k = 1, 2, \ldots$$

- 这称为Gauss-Seidel迭代法，简称为G-S迭代法

新格式

- 经上述改变后，我们有

$$x_k = D^{-1}Lx_k + D^{-1}Ux_{k-1} + g, \quad k = 1, 2, \ldots$$

- 这称为Gauss-Seidel迭代法，简称为G-S迭代法
- 如此变化，编程时存储量减少了

新格式

- 经上述改变后，我们有

$$x_k = D^{-1}Lx_k + D^{-1}Ux_{k-1} + g, \quad k = 1, 2, \ldots$$

- 这称为Gauss-Seidel迭代法，简称为G-S迭代法
- 如此变化，编程时存储量减少了
- 该格式1823年C.F. Gauss在给其学生C.L. Gerling的信中提到，P.L. von Seidel在1874年发表了这一方法

迭代矩阵

- 如果$(D-L)^{-1}$存在，那么迭代格式为

$$x_k = (D-L)^{-1}Ux_{k-1} + (D-L)^{-1}b$$

迭代矩阵

- 如果$(D-L)^{-1}$存在，那么迭代格式为

$$x_k = (D-L)^{-1}Ux_{k-1} + (D-L)^{-1}b$$

- 我们称$L_1 = (D-L)^{-1}U$为G-S迭代法的迭代矩阵，其第一列全是零，$(D-L)^{-1}b$称为G-S迭代法的常数项

迭代矩阵

- 如果$(D-L)^{-1}$存在，那么迭代格式为

$$x_k = (D-L)^{-1}Ux_{k-1} + (D-L)^{-1}b$$

- 我们称$L_1 = (D-L)^{-1}U$为G-S迭代法的迭代矩阵，其第一列全是零，$(D-L)^{-1}b$称为G-S迭代法的常数项
- 此时分量的计算次序是不能改变的

例

- 设$A = \begin{pmatrix} 16 & 3 \\ 7 & -11 \end{pmatrix}$, $b = \begin{pmatrix} 11 \\ 13 \end{pmatrix}$

例

- 设$A = \begin{pmatrix} 16 & 3 \\ 7 & -11 \end{pmatrix}$, $b = \begin{pmatrix} 11 \\ 13 \end{pmatrix}$
- 则$(D-L)^{-1} = \begin{pmatrix} 1/16 & 0 \\ 7/176 & -1/11 \end{pmatrix}$

例

- 设$A=\begin{pmatrix}16 & 3\\ 7 & -11\end{pmatrix}$, $b=\begin{pmatrix}11\\ 13\end{pmatrix}$
- 则$(D-L)^{-1}=\begin{pmatrix}1/16 & 0\\ 7/176 & -1/11\end{pmatrix}$
- 迭代矩阵和常数项分别为

$$L_1=\begin{pmatrix}0 & -3/16\\ 0 & -21/176\end{pmatrix},\quad g=\begin{pmatrix}11/16\\ -131/176\end{pmatrix}$$

例

- 取初值$x_0 = (1, 1)^T$

例

- 取初值$x_0 = (1, 1)^T$
- 迭代得到各向量依次为

$$\begin{pmatrix} 0.5 \\ -0.8636 \end{pmatrix} \to \begin{pmatrix} 0.8494 \\ -0.6413 \end{pmatrix} \to \begin{pmatrix} 0.8077 \\ -0.6678 \end{pmatrix}$$

例

- 取初值$x_0 = (1, 1)^T$
- 迭代得到各向量依次为
$$\begin{pmatrix} 0.5 \\ -0.8636 \end{pmatrix} \to \begin{pmatrix} 0.8494 \\ -0.6413 \end{pmatrix} \to \begin{pmatrix} 0.8077 \\ -0.6678 \end{pmatrix}$$
- 经六次迭代，得到向量$\begin{pmatrix} 0.8122 \\ -0.6650 \end{pmatrix}$

例

- 取初值$x_0 = (1,1)^T$
- 迭代得到各向量依次为

$$\begin{pmatrix} 0.5 \\ -0.8636 \end{pmatrix} \to \begin{pmatrix} 0.8494 \\ -0.6413 \end{pmatrix} \to \begin{pmatrix} 0.8077 \\ -0.6678 \end{pmatrix}$$

- 经六次迭代，得到向量$\begin{pmatrix} 0.8122 \\ -0.6650 \end{pmatrix}$
- 方程的准确解为

$$\begin{pmatrix} 160/197 \\ -131/197 \end{pmatrix} \approx \begin{pmatrix} 0.812183 \\ -0.664975 \end{pmatrix}$$

两种方法的共性

- 在这两种方法中，新的近似解x_k是已知近似解x_{k-1}的线性函数，并且只与x_{k-1}有关，即它们都可以写为

$$x_k = Mx_{k-1} + g$$

两种方法的共性

- 在这两种方法中，新的近似解x_k是已知近似解x_{k-1}的线性函数，并且只与x_{k-1}有关，即它们都可以写为

$$x_k = Mx_{k-1} + g$$

  - Jacobi迭代法：$M = D^{-1}(L+U)$, $g = D^{-1}b$

两种方法的共性

- 在这两种方法中，新的近似解x_k是已知近似解x_{k-1}的线性函数，并且只与x_{k-1}有关，即它们都可以写为

$$x_k = Mx_{k-1} + g$$

  - Jacobi迭代法：$M = D^{-1}(L+U)$, $g = D^{-1}b$
  - G-S迭代法：$M = (D-L)^{-1}U$, $g = (D-L)^{-1}b$

定义

- 我们把$x_k = Mx_{k-1} + g$的迭代法称为**单步线性定常迭代法**

定义

- 我们把$x_k = Mx_{k-1} + g$的迭代法称为单步线性定常迭代法
  - $M \in \mathbb{R}^{n\times n}$称为迭代矩阵

定义

- 我们把$x_k = Mx_{k-1} + g$的迭代法称为单步线性定常迭代法
  - $M \in \mathbb{R}^{n\times n}$称为迭代矩阵
  - $g \in \mathbb{R}^n$称为常数项

定义

- 我们把$x_k = Mx_{k-1} + g$的迭代法称为单步线性定常迭代法
  - $M \in \mathbb{R}^{n\times n}$称为迭代矩阵
  - $g \in \mathbb{R}^n$称为常数项
  - $x_0 \in \mathbb{R}^n$称为初始向量

收敛与发散

- 如果迭代格式对任意的初始向量产生的向量序列$\{x_k\}$（称为迭代序列）都有极限，则称该迭代格式是收敛的

收敛与发散

- 如果迭代格式对任意的初始向量产生的向量序列$\{x_k\}$（称为迭代序列）都有极限，则称该迭代格式是收敛的
- 否则称它是不收敛的，或者是发散的

收敛与发散

- 如果迭代格式对任意的初始向量产生的向量序列$\{x_k\}$（称为迭代序列）都有极限，则称该迭代格式是收敛的
- 否则称它是不收敛的，或者是发散的
- 若收敛，记迭代序列的极限为x_*，则有$x_* = Mx_* + g$

迭代法与方程组的相容性

- x_*是方程组$(I - M)x = g$的解

迭代法与方程组的相容性

- x_*是方程组$(I-M)x=g$的解
- 如果存在非奇异矩阵$G\in\mathbb{R}^{n\times n}$使得

$$G(I-M)=A,\quad Gg=b$$

则迭代序列也收敛到$Ax=b$的解，此时两个方程称为等价方程组，也称迭代格式与方程组$Ax=b$是相容的

迭代法与方程组的相容性

- x_*是方程组$(I-M)x=g$的解
- 如果存在非奇异矩阵$G\in\mathbb{R}^{n\times n}$使得

$$G(I-M)=A,\quad Gg=b$$

则迭代序列也收敛到$Ax=b$的解，此时两个方程称为等价方程组，也称迭代格式与方程组$Ax=b$是相容的

- 已有的两种格式是相容的

误差向量

- 设x_*是$Ax = b$的解

误差向量

- 设x_*是$Ax = b$的解
- $\{x_k\}$是由相容格式$x_k = Mx_{k-1} + g$产生的迭代序列

误差向量

- 设x_*是$Ax = b$的解
- $\{x_k\}$是由相容格式$x_k = Mx_{k-1} + g$产生的迭代序列
- 定义$y_k = x_k - x_*$, 称为x_k的误差向量

误差向量

- 设x_*是$Ax = b$的解
- $\{x_k\}$是由相容格式$x_k = Mx_{k-1} + g$产生的迭代序列
- 定义$y_k = x_k - x_*$，称为x_k的误差向量
- 那么

$$\begin{aligned} y_{k+1} = x_{k+1} - x_* &= Mx_k + g - (Mx_* + g) \\ &= My_k = M^k y_0 \end{aligned}$$

收敛条件

- 对任意y_0都有$y_k \to 0$（即$x_k \to x_*$）当且仅当$M^k \to 0$

收敛条件

- 对任意y_0都有$y_k \to 0$（即$x_k \to x_*$）当且仅当$M^k \to 0$
- 而$M^k \to 0$的充要条件是$\rho(M) < 1$

收敛条件

- 对任意y_0都有$y_k \to 0$（即$x_k \to x_*$）当且仅当$M^k \to 0$
- 而$M^k \to 0$的充要条件是$\rho(M) < 1$

定理

解方程组$Ax = b$的单步线性定常迭代法$x_k = Mx_{k-1} + g$收敛的充要条件是迭代矩阵M的谱半径小于*1*

注解

- 迭代序列收敛取决于迭代矩阵的谱半径，而与初始向量的选取和常数项无关

注解

- 迭代序列收敛取决于迭代矩阵的谱半径，而与初始向量的选取和常数项无关
- 相同的方程组，Jacobi迭代矩阵和G-S迭代矩阵的谱半径不一定相同，而且无包含关系

注解

- 迭代序列收敛取决于迭代矩阵的谱半径，而与初始向量的选取和常数项无关
- 相同的方程组，Jacobi迭代矩阵和G-S迭代矩阵的谱半径不一定相同，而且无包含关系
- 例子: `Mathematica sec4.2.1.nb`

- 用谱半径判断迭代法是否收敛，这是很不方便的：谱半径的计算相当困难

- 用谱半径判断迭代法是否收敛，这是很不方便的：谱半径的计算相当困难
- 需要一些容易计算的充分条件

- 用谱半径判断迭代法是否收敛，这是很不方便的：谱半径的计算相当困难
- 需要一些容易计算的充分条件

定理

如果范数$\|M\| = q < 1$，则我们有

$$\|x_k - x_*\| \leqslant \frac{q^k}{1-q}\|x_1 - x_0\|$$

- 用谱半径判断迭代法是否收敛，这是很不方便的：谱半径的计算相当困难
- 需要一些容易计算的充分条件

定理

如果范数$\|M\| = q < 1$，则我们有

$$\|x_k - x_*\| \leqslant \frac{q^k}{1-q}\|x_1 - x_0\|$$

- 此定理结论的右边与精确解无关

定理证明

- 在$y_k = M^k y_0$的两边取范数，则有

$$\|y_k\| \leqslant \|M\|^k \|y_0\| = q^k \|y_0\|$$

定理证明

- 在$y_k = M^k y_0$的两边取范数，则有

$$\|y_k\| \leqslant \|M\|^k \|y_0\| = q^k \|y_0\|$$

- 根据y_0的定义

$$\begin{aligned}\|y_0\| &= \|x_0 - x_*\| \leqslant \|x_0 - x_1\| + \|x_1 - x_*\| \\ &= \|x_0 - x_1\| + \|My_0\| \leqslant \|x_0 - x_1\| + q\|y_0\|\end{aligned}$$

所以有$\|y_0\| \leqslant \dfrac{1}{1-q}\|x_1 - x_0\|$

注解

- 从上述近似解的误差估计可以计算出为了达到精度要求，需要进行多少次迭代

注解

- 从上述近似解的误差估计可以计算出为了达到精度要求，需要进行多少次迭代
- 这种估计一般是偏高的

注解

- 从上述近似解的误差估计可以计算出为了达到精度要求，需要进行多少次迭代
- 这种估计一般是偏高的

定理

若$\|M\| = q < 1$, 则我们有

$$\|x_k - x_*\| \leqslant \frac{q}{1-q}\|x_{k-1} - x_k\|$$

注解

- 从上述近似解的误差估计可以计算出为了达到精度要求，需要进行多少次迭代
- 这种估计一般是偏高的

定理

若$\|M\| = q < 1$, 则我们有

$$\|x_k - x_*\| \leqslant \frac{q}{1-q}\|x_{k-1} - x_k\|$$

- 此定理结论是用刚得到的两个迭代向量估计最新结果的精度

定理证明

- 因为

$$\begin{aligned}\|x_k - x_*\| =& \|M(x_{k-1} - x_*)\| \leqslant q\|x_{k-1} - x_*\| \\ &\leqslant q\|x_{k-1} - x_k\| + q\|x_k - x_*\|\end{aligned}$$

所以

$$\|x_k - x_*\| \leqslant \frac{q}{1-q}\|x_{k-1} - x_k\|$$

注解

- 只要M的范数不是很接近1，当相邻两次迭代向量很接近时，那么x_k与x_*也就很接近

注解

- 只要M的范数不是很接近1，当相邻两次迭代向量很接近时，那么x_k与x_*也就很接近
- 实际计算时可以用量$\|x_{k-1} - x_k\|$是否适当小来判别迭代法是否应该终止

注解

- 只要M的范数不是很接近1，当相邻两次迭代向量很接近时，那么x_k与x_*也就很接近
- 实际计算时可以用量$\|x_{k-1}-x_k\|$是否适当小来判别迭代法是否应该终止
  - 若$\|M\|=0.9$, $\|x_{k-1}-x_k\|=10^{-8}$, 则$\|x_k-x_*\|\leqslant 9\times 10^{-8}$

注解

- 只要M的范数不是很接近1，当相邻两次迭代向量很接近时，那么x_k与x_*也就很接近
- 实际计算时可以用量$\|x_{k-1}-x_k\|$是否适当小来判别迭代法是否应该终止
  - 若$\|M\|=0.9$, $\|x_{k-1}-x_k\|=10^{-8}$, 则$\|x_k-x_*\|\leqslant 9\times 10^{-8}$
- 若$\|M\|$很接近1, 则不能断定精度

注解

- 只要M的范数不是很接近1，当相邻两次迭代向量很接近时，那么x_k与x_*也就很接近
- 实际计算时可以用量$\|x_{k-1}-x_k\|$是否适当小来判别迭代法是否应该终止
  - 若$\|M\|=0.9$, $\|x_{k-1}-x_k\|=10^{-8}$, 则$\|x_k-x_*\|\leqslant 9\times 10^{-8}$
- 若$\|M\|$很接近1, 则不能断定精度
- 实际一般应用1范数和∞范数：容易计算

分析

- Jacobi迭代法的迭代矩阵B容易得到，因此前面的判别法基本令人满意

分析

- Jacobi迭代法的迭代矩阵B容易得到，因此前面的判别法基本令人满意
- G-S迭代法中的迭代矩阵需要求$L_1 = (D - L)^{-1}U$, 不太容易得到

分析

- Jacobi迭代法的迭代矩阵B容易得到，因此前面的判别法基本令人满意
- G-S迭代法中的迭代矩阵需要求$L_1 = (D - L)^{-1}U$，不太容易得到
- 因此我们需要给出一些辅助判别

分析

- Jacobi迭代法的迭代矩阵B容易得到，因此前面的判别法基本令人满意
- G-S迭代法中的迭代矩阵需要求$L_1 = (D - L)^{-1}U$，不太容易得到
- 因此我们需要给出一些辅助判别
  - 能否从B的性质判断L_1的性质？

分析

- Jacobi迭代法的迭代矩阵B容易得到，因此前面的判别法基本令人满意
- G-S迭代法中的迭代矩阵需要求$L_1 = (D - L)^{-1}U$, 不太容易得到
- 因此我们需要给出一些辅助判别
  - 能否从B的性质判断L_1的性质？
  - 能否直接利用A的性质进行判断？

B与L_1的关系

- 回忆

B与L_1的关系

- 回忆
  - Jacobi迭代：$y = D^{-1}Lx + D^{-1}Ux$

B与L_1的关系

- 回忆
  - Jacobi迭代：$y = D^{-1}Lx + D^{-1}Ux$
  - G-S迭代：$y = D^{-1}Ly + D^{-1}Ux$

B与L_1的关系

- 回忆
  - Jacobi迭代：$y = D^{-1}Lx + D^{-1}Ux$
  - G-S迭代：$y = D^{-1}Ly + D^{-1}Ux$
  - $D^{-1}L$为下三角矩阵，对角元为零；$D^{-1}U$为上三角矩阵，对角元为零

B与L_1的关系

- 回忆
  - Jacobi迭代：$y = D^{-1}Lx + D^{-1}Ux$
  - G-S迭代：$y = D^{-1}Ly + D^{-1}Ux$
  - $D^{-1}L$为下三角矩阵，对角元为零；
    $D^{-1}U$为上三角矩阵，对角元为零
- 所以我们完全有理由根据B的性质推断L_1的性质

∞范数

- 设$\|B\|_\infty < 1$

∞范数

- 设$\|B\|_\infty < 1$
- 定义

$$\mu = \max_i \frac{\sum_{j=i+1}^{n} |b_{ij}|}{1 - \sum_{j=1}^{i-1} |b_{ij}|}$$

∞范数

- 设$\|B\|_\infty < 1$
- 定义

$$\mu = \max_i \frac{\sum\limits_{j=i+1}^{n} |b_{ij}|}{1 - \sum\limits_{j=1}^{i-1} |b_{ij}|}$$

引理

$$\mu \leqslant \|B\|_\infty < 1$$

证明

- 令$\ell_i = \sum_{j=1}^{i-1} |b_{ij}|,\ u_i = \sum_{j=i+1}^{n} |b_{ij}|$

证明

- 令$\ell_i = \sum_{j=1}^{i-1} |b_{ij}|$, $u_i = \sum_{j=i+1}^{n} |b_{ij}|$
- 则$\ell_i + u_i \leqslant \|B\|_\infty < 1$

证明

- 令$\ell_i = \sum_{j=1}^{i-1} |b_{ij}|$, $u_i = \sum_{j=i+1}^{n} |b_{ij}|$
- 则$\ell_i + u_i \leqslant \|B\|_\infty < 1$
- 注意到$\forall i$, $b_{ii} = 0$, 所以存在i, $\ell_i + u_i = \|B\|_\infty$

- 注意到

$$\begin{aligned}\ell_i + u_i - \frac{u_i}{1-\ell_i} &= \frac{1}{1-\ell_i}((\ell_i + u_i)(1-\ell_i) - u_i) \\ &= \frac{\ell_i}{1-\ell_i}(1-\ell_i-u_i) \geqslant 0\end{aligned}$$

- 注意到

$$\begin{aligned}\ell_i + u_i - \frac{u_i}{1-\ell_i} &= \frac{1}{1-\ell_i}((\ell_i + u_i)(1-\ell_i) - u_i) \\ &= \frac{\ell_i}{1-\ell_i}(1-\ell_i-u_i) \geqslant 0\end{aligned}$$

- 所以我们有 $\dfrac{u_i}{1-\ell_i} \leqslant \ell_i + u_i$

- 注意到

$$\begin{aligned}\ell_i + u_i - \frac{u_i}{1-\ell_i} &= \frac{1}{1-\ell_i}((\ell_i+u_i)(1-\ell_i)-u_i)\\ &= \frac{\ell_i}{1-\ell_i}(1-\ell_i-u_i) \geqslant 0\end{aligned}$$

- 所以我们有$\dfrac{u_i}{1-\ell_i} \leqslant \ell_i + u_i$
- 两边对i取最大值，我们得到

$$\mu = \max_i \frac{u_i}{1-\ell_i} \leqslant \max_i(\ell_i+u_i) = \|B\|_\infty < 1$$

$\|B\|_\infty$与$\|L_1\|_\infty$

定理

设$\|B\|_\infty < 1$，则$\|L_1\|_\infty \leqslant \|B\|_\infty < 1$，而且由G-S迭代法产生的近似解$x_k$与准确解$x_*$之间满足

$$\|x_k - x_*\|_\infty \leqslant \frac{\mu^k}{1-\mu}\|x_1 - x_0\|_\infty$$

定理证明

- 存在满足$\|x\|_\infty = 1$的$x \in \mathbb{R}^n$使得$\|L_1\|_\infty = \|L_1 x\|_\infty$

定理证明

- 存在满足$\|x\|_\infty = 1$的$x \in \mathbb{R}^n$使得$\|L_1\|_\infty = \|L_1 x\|_\infty$
- 令$y = L_1 x$, $|y_i| = \|y\|_\infty$

定理证明

- 存在满足$\|x\|_\infty = 1$的$x \in \mathbb{R}^n$使得$\|L_1\|_\infty = \|L_1 x\|_\infty$
- 令$y = L_1 x$, $|y_i| = \|y\|_\infty$
- 根据$L_1 = (D-L)^{-1}U$的定义，我们有$y = D^{-1}Ly + D^{-1}Ux$

定理证明

- 存在满足$\|x\|_\infty = 1$的$x \in \mathbb{R}^n$使得$\|L_1\|_\infty = \|L_1 x\|_\infty$
- 令$y = L_1 x$, $|y_i| = \|y\|_\infty$
- 根据$L_1 = (D-L)^{-1}U$的定义，我们有$y = D^{-1}Ly + D^{-1}Ux$
- 由于$B = D^{-1}L + D^{-1}U$, 所以$D^{-1}L$为B的下三角部分，$D^{-1}U$为B的上三角部分

- 比较两边的第i个分量，我们有

$$y_i = \sum_{j=1}^{i-1} b_{ij} y_j + \sum_{j=i+1}^{n} b_{ij} x_j$$

- 比较两边的第i个分量，我们有

$$y_i = \sum_{j=1}^{i-1} b_{ij} y_j + \sum_{j=i+1}^{n} b_{ij} x_j$$

- 两边取绝对值，可得

$$\|y\|_\infty \leqslant \|y\|_\infty \ell_i + u_i$$

- 比较两边的第i个分量，我们有

$$y_i = \sum_{j=1}^{i-1} b_{ij} y_j + \sum_{j=i+1}^{n} b_{ij} x_j$$

- 两边取绝对值，可得

$$\|y\|_\infty \leqslant \|y\|_\infty \ell_i + u_i$$

- 由此可得

$$\|L_1\|_\infty = \|y\|_\infty \leqslant \frac{u_i}{1-\ell_i} \leqslant \mu < 1$$

估计式的证明

- 根据$\|L_1\|_\infty \leqslant \mu < 1$可得

$$\begin{aligned}\|x_k - x_*\|_\infty &\leqslant \frac{\|L_1\|_\infty^k}{1-\|L_1\|_\infty}\|x_1 - x_0\|_\infty \\ &\leqslant \frac{\mu^k}{1-\mu}\|x_1 - x_0\|_\infty\end{aligned}$$

1范数

- 设$\|B\|_1 < 1$

1范数

- 设$\|B\|_1 < 1$
- 定义

$$\tilde{\mu} = \max_j \frac{\sum_{i=1}^{j-1} |b_{ij}|}{1 - \sum_{i=j+1}^{n} |b_{ij}|}$$

1范数

- 设$\|B\|_1 < 1$
- 定义

$$\tilde{\mu} = \max_j \frac{\sum_{i=1}^{j-1} |b_{ij}|}{1 - \sum_{i=j+1}^{n} |b_{ij}|}$$

- 则类似前面对$\mu < 1$的证明可证$\tilde{\mu} \leqslant \|B\|_1 < 1$

$\|B\|_1$与$\rho(L_1)$

定理

若$\|B\|_1 < 1$, 则$\rho(L_1) \leqslant \|B\|_1 < 1$, 而且由$G$-$S$迭代法所得近似解$x_k$与准确解$x_*$满足

$$\|x_k - x_*\|_1 \leqslant \frac{\tilde{\mu}^k}{(1-\tilde{\mu})(1-s)}\|x_1 - x_0\|_1$$

其中$s = \max_j \sum_{i=j+1}^{n} |b_{ij}|$, 这是$B$的下三角阵$D^{-1}L$的1范数

$\rho(L_1) < 1$的证明

- 因为

$$(D-L)^{-1}U = (I-D^{-1}L)^{-1}(D^{-1}U(I-D^{-1}L)^{-1})(I-D^{-1}L)$$

所以L_1与$\tilde{L}_1 = D^{-1}U(I-D^{-1}L)^{-1}$相似，从而$\rho(L_1) = \rho(\tilde{L}_1)$

$\rho(L_1) < 1$的证明

- 因为

$$(D-L)^{-1}U = (I-D^{-1}L)^{-1}(D^{-1}U(I-D^{-1}L)^{-1})(I-D^{-1}L)$$

所以L_1与$\tilde{L}_1 = D^{-1}U(I-D^{-1}L)^{-1}$相似，从而$\rho(L_1) = \rho(\tilde{L}_1)$

- 类似∞范数时的情况可证$\|\tilde{L}_1^T\|_\infty \leqslant \tilde{\mu} < 1$, 从而有

$$\rho(L_1) \leqslant \|\tilde{L}_1^T\|_\infty \leqslant \tilde{\mu} < 1$$

- 实际上，令$C = B^T$. $\|B\|_1 < 1 \Longrightarrow \|C\|_\infty < 1$

- 实际上，令$C = B^T$. $\|B\|_1 < 1 \Longrightarrow \|C\|_\infty < 1$
- C的下三角部分为$(D^{-1}U)^T = U^T D^{-1}$, 上三角部分为$(D^{-1}L)^T = L^T D^{-1}$

- 实际上，令$C = B^T$. $\|B\|_1 < 1 \Longrightarrow \|C\|_\infty < 1$
- C的下三角部分为$(D^{-1}U)^T = U^T D^{-1}$, 上三角部分为$(D^{-1}L)^T = L^T D^{-1}$
- $\tilde{L}_1^T = (D^{-1}U(I - D^{-1}L)^{-1})^T = (I - L^T D^{-1})^{-1} U^T D^{-1}$

- 实际上，令$C = B^T$. $\|B\|_1 < 1 \Longrightarrow \|C\|_\infty < 1$
- C的下三角部分为$(D^{-1}U)^T = U^T D^{-1}$, 上三角部分为$(D^{-1}L)^T = L^T D^{-1}$
- $\tilde{L}_1^T = (D^{-1}U(I - D^{-1}L)^{-1})^T = (I - L^T D^{-1})^{-1} U^T D^{-1}$
- $\tilde{L}_1^T x = y \Longrightarrow y = L^T D^{-1} y + U^T D^{-1} x$

误差估计的证明

- 由G-S迭代法的格式，可得

$$x_k - x_{k-1} = D^{-1}L(x_k - x_{k-1}) + D^{-1}U(x_{k-1} - x_{k-2})$$

误差估计的证明

- 由G-S迭代法的格式，可得

$$x_k - x_{k-1} = D^{-1}L(x_k - x_{k-1}) + D^{-1}U(x_{k-1} - x_{k-2})$$

- 分量表示即为

$$x_i^{(k)} - x_i^{(k-1)} = \sum_{j=1}^{i-1} b_{ij}(x_j^{(k)} - x_j^{(k-1)}) + \sum_{j=i+1}^{n} b_{ij}(x_j^{(k-1)} - x_j^{(k-2)})$$

- 两边取绝对值后对i求和，再交换求和顺序

$$
\begin{aligned}
\sum_{i=1}^{n}\left|x_i^{(k)}-x_i^{(k-1)}\right| &\leqslant \sum_{i=1}^{n}\left(\sum_{j=1}^{i-1}|b_{ij}|\left|x_j^{(k)}-x_j^{(k-1)}\right|\right. \\
&\qquad \left.+\sum_{j=i+1}^{n}|b_{ij}|\left|x_j^{(k-1)}-x_j^{(k-2)}\right|\right) \\
&\leqslant \sum_{j=1}^{n}\left(\sum_{i=j+1}^{n}|b_{ij}|\left|x_j^{(k)}-x_j^{(k-1)}\right|\right. \\
&\qquad \left.+\sum_{i=1}^{j-1}|b_{ij}|\left|x_j^{(k-1)}-x_j^{(k-2)}\right|\right)
\end{aligned}
$$

- 令

$$\tilde{u}_j = \sum_{i=j+1}^{n} |b_{ij}|, \quad \tilde{\ell}_j = \sum_{i=1}^{j-1} |b_{ij}|$$

- 令

$$\tilde{u}_j = \sum_{i=j+1}^{n} |b_{ij}|, \quad \tilde{\ell}_j = \sum_{i=1}^{j-1} |b_{ij}|$$

- 则接前推导我们有

$$\sum_{i=1}^{n} \left| x_i^{(k)} - x_i^{(k-1)} \right| \leqslant \sum_{j=1}^{n} \left(\tilde{u}_j \left| x_j^{(k)} - x_j^{(k-1)} \right| + \tilde{\ell}_j \left| x_j^{(k-1)} - x_j^{(k-2)} \right| \right)$$

从而有

$$\begin{aligned}
\sum_{j=1}^{n}(1-\tilde{u}_j)\left|x_j^{(k)}-x_j^{(k-1)}\right| &\leqslant \sum_{j=1}^{n}\tilde{\ell}_j\left|x_j^{(k-1)}-x_j^{(k-2)}\right| \\
&\leqslant \tilde{\mu}\sum_{j=1}^{n}(1-\tilde{u}_j)\left|x_j^{(k-1)}-x_j^{(k-2)}\right| \\
&\leqslant \cdots\cdots \\
&\leqslant \tilde{\mu}^{k-1}\sum_{j=1}^{n}(1-\tilde{u}_j)\left|x_j^{(1)}-x_j^{(0)}\right|
\end{aligned}$$

- 根据$\tilde{\mu}$和s的定义，$1 - s \leqslant 1 - \tilde{u}_j < 1$

- 根据$\tilde{\mu}$和s的定义，$1-s \leqslant 1-\tilde{u}_j < 1$
- 所以

$$(1-s)\sum_{j=1}^{n}\left|x_j^{(k)}-x_j^{(k-1)}\right| \leqslant \tilde{\mu}^{k-1}\sum_{j=1}^{n}\left|x_j^{(1)}-x_j^{(0)}\right|$$

- 根据$\tilde{\mu}$和s的定义，$1-s \leqslant 1-\tilde{u}_j < 1$
- 所以

$$(1-s)\sum_{j=1}^{n}\left|x_j^{(k)}-x_j^{(k-1)}\right| \leqslant \tilde{\mu}^{k-1}\sum_{j=1}^{n}\left|x_j^{(1)}-x_j^{(0)}\right|$$

- 这就是$(1-s)\|x_k-x_{k-1}\|_1 \leqslant \tilde{\mu}^{k-1}\|x_1-x_0\|_1$

- 注意到$x_k - x_* = \sum_{i=k}^{\infty}(x_i - x_{i+1})$

- 注意到$x_k - x_* = \sum_{i=k}^{\infty}(x_i - x_{i+1})$
- 所以

$$\begin{aligned}\|x_k - x_*\|_1 &\leqslant \sum_{i=k}^{\infty} \|x_i - x_{i+1}\|_1 \\ &\leqslant \frac{1}{1-s}\sum_{i=k}^{\infty} \tilde{\mu}^i \|x_1 - x_0\|_1 \\ &\leqslant \frac{\tilde{\mu}^k}{(1-\tilde{\mu})(1-s)}\|x_1 - x_0\|_1\end{aligned}$$

从A直接判断：正定矩阵

定理

如果A是对称的，而且对角元全是正数，则*Jacobi*迭代法收敛的充要条件是A和$2D-A$都正定

定理证明

- 迭代矩阵$B = D^{-1}(L + U) = I - D^{-1}A$

定理证明

- 迭代矩阵$B = D^{-1}(L+U) = I - D^{-1}A$
- 由于对角元全是正数，所以

$$B = D^{-1/2}(I - D^{-1/2}AD^{-1/2})D^{1/2}$$

定理证明

- 迭代矩阵$B = D^{-1}(L+U) = I - D^{-1}A$
- 由于对角元全是正数，所以

$$B = D^{-1/2}(I - D^{-1/2}AD^{-1/2})D^{1/2}$$

- A对称，所以$I - D^{-1/2}AD^{-1/2}$也是对称的，而且它与B相似，所以B的特征值全为实数

- Jacobi迭代法收敛$\Longleftrightarrow \rho(B) < 1$. 而$B$的特征值全为实数，所以这又等价于$I - B$和$I + B$的特征值均为正实数

- Jacobi迭代法收敛$\Longleftrightarrow \rho(B) < 1$. 而$B$的特征值全为实数，所以这又等价于$I - B$和$I + B$的特征值均为正实数
- 注意到

$$I - B = D^{-1/2}(D^{-1/2}AD^{-1/2})D^{1/2}$$
$$I + B = D^{-1/2}(2I - D^{-1/2}AD^{-1/2})D^{1/2}$$

- Jacobi迭代法收敛$\Longleftrightarrow \rho(B)<1$. 而$B$的特征值全为实数，所以这又等价于$I-B$和$I+B$的特征值均为正实数
- 注意到

$$I-B=D^{-1/2}(D^{-1/2}AD^{-1/2})D^{1/2}$$

$$I+B=D^{-1/2}(2I-D^{-1/2}AD^{-1/2})D^{1/2}$$

- 所以$\rho(B)<1$等价于A与$2D-A$均正定

- Jacobi迭代法收敛$\Longleftrightarrow$ $\rho(B) < 1$. 而B的特征值全为实数，所以这又等价于$I - B$和$I + B$的特征值均为正实数
- 注意到

$$I - B = D^{-1/2}(D^{-1/2}AD^{-1/2})D^{1/2}$$

$$I + B = D^{-1/2}(2I - D^{-1/2}AD^{-1/2})D^{1/2}$$

- 所以$\rho(B) < 1$等价于A与$2D - A$均正定
  - $D^{-1/2}AD^{-1/2}$相合于A;
    $2I - D^{-1/2}AD^{-1/2}$相合于$2D - A$

对称正定：G-S迭代

- 正定理

对称正定：G-S迭代

- 正定理

定理

若系数矩阵A对称正定，则G-S迭代法收敛

对称正定：G-S迭代

- 正定理

定理

若系数矩阵A对称正定，则G-S迭代法收敛

- 定理证明见第4节

对称正定：G-S迭代

- 正定理

定理

若系数矩阵A对称正定，则G-S迭代法收敛

- 定理证明见第4节

- 逆定理(习题第7题)

对称正定：G-S迭代

- 正定理

定理

若系数矩阵A对称正定，则G-S迭代法收敛

- 定理证明见第4节

- 逆定理(习题第7题)

定理

若系数矩阵A是具有正对角元的对称矩阵，G-S迭代法对任意初值收敛，则A必是正定的

对角占优矩阵

- 矩阵$A = (a_{ij}) \in \mathbb{R}^{n\times n}$. 若对所有的$i = 1, \dots, n$,

$$|a_{ii}| \geqslant \sum_{j=1, j\neq i}^{n} |a_{ij}|$$

并且至少有一个对i有严格不等号成立，则称A是弱严格对角占优的。如果对所有i都有严格不等号成立，则称A是严格对角占优的

可约和不可约矩阵

- 矩阵$A=(a_{ij})\in\mathbb{R}^{n\times n}$. 若存在$n$阶排列方阵$P$使得

$$PAP^T=\begin{pmatrix} A_{11} & 0 \\ A_{12} & A_{22}\end{pmatrix}$$

其中A_{11}是r阶方阵，A_{22}是$n-r$阶方阵，则称A是可约的或可分的；反之称为不可约的或不可分的

可约矩阵的意义

- 如果A可约，可以把$Ax=b$化为

$$PAP^TPx = Pb$$

可约矩阵的意义

- 如果A可约，可以把$Ax=b$化为

$$PAP^TPx = Pb$$

- 记$Px=y$, $Pb=f$, 则有

$$\begin{pmatrix} A_{11} & 0 \\ A_{12} & A_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$$

可约矩阵的意义

- 如果A可约，可以把$Ax = b$化为

$$PAP^T Px = Pb$$

- 记$Px = y$, $Pb = f$, 则有

$$\begin{pmatrix} A_{11} & 0 \\ A_{12} & A_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$$

- 如此我们可以先求解r阶方程组$A_{11}y_1 = f_1$, 得到y_1, 代入$A_{12}y_1 + A_{22}y_2 = f_2$ 后就可以解出y_2

可约矩阵的等价定义

- 设A为n $(n \geqslant 2)$阶方阵，$\mathcal{W} = \{1, \ldots, n\}$. 如果存在$\mathcal{W}$的两个非空的子集$\mathcal{S}$和$\mathcal{T}$满足

$$\mathcal{S} \cup \mathcal{T} = \mathcal{W}, \mathcal{S} \cap \mathcal{T} = \emptyset$$

使得$a_{ij} = 0$, $(i \in \mathcal{S}, j \in \mathcal{T})$, 则称$A$为可约的；否则称$A$是不可约的

不可约对角占优

- 如果一个矩阵不可约，并且是弱严格对角占优的，则称该矩阵为不可约对角占优的

不可约对角占优

- 如果一个矩阵不可约，并且是弱严格对角占优的，则称该矩阵为不可约对角占优的
- 下述三对角阵为不可约对角占优矩阵

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

不可约对角占优

- 如果一个矩阵不可约，并且是弱严格对角占优的，则称该矩阵为不可约对角占优的
- 下述三对角阵为不可约对角占优矩阵

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

- 弱严格对角占优矩阵，有可能一行全是零；而不可约对角占优，就排除了这种情形，从而任意对角元非零

定理

- 对角占优与非奇异的关系：

定理

- 对角占优与非奇异的关系：

定理

如果矩阵A是严格对角占优的，或者不可约对角占优的，则A非奇异

定理

- 对角占优与非奇异的关系：

定理

如果矩阵A是严格对角占优的，或者不可约对角占优的，则A非奇异

- 可约对角占优的，矩阵有可能奇异：

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 3 \end{pmatrix}$$

严格对角占优情形的证明

- 若A奇异，则$Ax=0$有非零解x. 不妨设$|x_i|=\|x\|_\infty=1$, 则有

$$|a_{ii}|=|a_{ii}x_i|=\left|\sum_{j=1,j\neq i}^{n} a_{ij}x_j\right|\leqslant\sum_{j=1,j\neq i}^{n}|a_{ij}|$$

严格对角占优情形的证明

- 若A奇异，则$Ax = 0$有非零解x. 不妨设$|x_i| = \|x\|_\infty = 1$, 则有

$$|a_{ii}| = |a_{ii}x_i| = \left|\sum_{j=1,j\neq i}^{n} a_{ij}x_j\right| \leqslant \sum_{j=1,j\neq i}^{n} |a_{ij}|$$

- 这与A严格对角占优矛盾

不可约对角占优情形的证明

- 同样，若A奇异，则存在x满足$\|x\|_\infty = 1$使得$Ax = 0$

不可约对角占优情形的证明

- 同样，若A奇异，则存在x满足$\|x\|_\infty = 1$使得$Ax = 0$
- 定义$\mathcal{S} = \{i : |x_i| = 1\}$, $\mathcal{T} = \{k : |x_k| < 1\}$

不可约对角占优情形的证明

- 同样，若A奇异，则存在x满足$\|x\|_\infty = 1$使得$Ax = 0$
- 定义$\mathcal{S} = \{i : |x_i| = 1\}$, $\mathcal{T} = \{k : |x_k| < 1\}$
- 显然$\mathcal{S} \cup \mathcal{T} = \mathcal{W}$, $\mathcal{S} \cap \mathcal{T} = \emptyset$, 而且$\mathcal{S}$非空

T为空集

- 假设T为空集

T为空集

- 假设T为空集
- 那么x的各分量的绝对值均为1，从而$\forall i \in \mathcal{S}$,

$$|a_{ii}| \leqslant \sum_{j=1, j\neq i}^{n} |a_{ij}|$$

T为空集

- 假设T为空集
- 那么x的各分量的绝对值均为1，从而$\forall i \in \mathcal{S}$,

$$|a_{ii}| \leqslant \sum_{j=1, j\neq i}^{n} |a_{ij}|$$

- 这与A弱严格对角占优矛盾

$\mathcal{T}$非空：利用不可约完成证明

- 因为A不可约，所以存在$i \in \mathcal{S}$, $k \in \mathcal{T}$使得$a_{ik} \neq 0$

$\mathcal{T}$非空：利用不可约完成证明

- 因为A不可约，所以存在$i \in \mathcal{S}$, $k \in \mathcal{T}$使得$a_{ik} \neq 0$
- 于是$|a_{ik}x_k| < |a_{ik}|$, 并且

$$\begin{aligned}|a_{ii}| = |a_{ii}x_i| &\leqslant \sum_{j\in\mathcal{S},j\neq i} |a_{ij}||x_j| + \sum_{j\in\mathcal{T}} |a_{ij}||x_j| \\ &< \sum_{j\in\mathcal{S},j\neq i} |a_{ij}| + \sum_{j\in\mathcal{T}} |a_{ij}| \\ &= \sum_{j\neq i} |a_{ij}|\end{aligned}$$

这也与A为弱严格对角占优矛盾

推论

推论

若A是严格对角占优的或者不可约对角占优的对称矩阵，且A的对角线元素均为正数，则A正定

证明

- A对称，所以特征值均为实数。为证A正定，只需要证明所有特征值均为正数

证明

- A对称，所以特征值均为实数。为证A正定，只需要证明所有特征值均为正数
- 若有一个特征值$\lambda \leqslant 0$，考虑矩阵$A - \lambda I$

证明

- A对称，所以特征值均为实数。为证A正定，只需要证明所有特征值均为正数
- 若有一个特征值$\lambda \leqslant 0$，考虑矩阵$A - \lambda I$
- 矩阵$A - \lambda I$只是在A的对角线元素上增加了一些，所以$A - \lambda I$和A一样是严格对角占优或者不可约对角占优的，从而是非奇异的，这与λ为A的特征值矛盾

定理

定理

若A是严格对角占优的，或者不可约对角占优的，则Jacobi迭代法和G-S迭代法都收敛

证明：Jacobi迭代

- $\forall i,\ a_{ii} \neq 0$

证明：Jacobi迭代

- $\forall i,\ a_{ii} \neq 0$
- 所以矩阵D可逆

证明：Jacobi迭代

- $\forall i,\ a_{ii} \neq 0$
- 所以矩阵D可逆
- 假设Jacobi迭代法的迭代矩阵B的某个特征值λ的模长不小于1

证明：Jacobi迭代

- $\forall i,\ a_{ii} \neq 0$
- 所以矩阵D可逆
- 假设Jacobi迭代法的迭代矩阵B的某个特征值λ的模长不小于1
- 考虑矩阵$\lambda D - L - U$, 这也是严格对角占优的或者不可约对角占优的，从而非奇异

证明：Jacobi迭代

- $\forall i,\ a_{ii} \neq 0$
- 所以矩阵D可逆
- 假设Jacobi迭代法的迭代矩阵B的某个特征值λ的模长不小于1
- 考虑矩阵$\lambda D - L - U$, 这也是严格对角占优的或者不可约对角占优的，从而非奇异
- 由于$\lambda I - B = D^{-1}(\lambda D - L - U)$, 所以$\det(\lambda I - B) \neq 0$, 矛盾

证明：G-S迭代

- G-S迭代法的迭代矩阵$L_1 = (D - L)^{-1}U$

证明：G-S迭代

- G-S迭代法的迭代矩阵$L_1 = (D - L)^{-1} U$
- $\lambda I - L_1 = (D - L)^{-1}(\lambda D - \lambda L - U)$

证明：G-S迭代

- G-S迭代法的迭代矩阵$L_1 = (D-L)^{-1}U$
- $\lambda I - L_1 = (D-L)^{-1}(\lambda D - \lambda L - U)$
- 若$|\lambda| \geqslant 1$, 则$D-L$和$\lambda D - \lambda L - U$也是严格对角占优的，或者不可约对角占优的，从而它们均可逆

证明：G-S迭代

- G-S迭代法的迭代矩阵$L_1 = (D - L)^{-1}U$
- $\lambda I - L_1 = (D - L)^{-1}(\lambda D - \lambda L - U)$
- 若$|\lambda| \geqslant 1$, 则$D - L$和$\lambda D - \lambda L - U$也是严格对角占优的，或者不可约对角占优的，从而它们均可逆
- 这就证明了G-S迭代法收敛

平均收敛速度

- 单步线性定常迭代法：$x_k = Mx_{k-1} + g$

平均收敛速度

- 单步线性定常迭代法：$x_k = Mx_{k-1} + g$
- 误差向量$y_k = x_k - x_*$满足$y_k = M^k y_0$，从而有

$$\|y_k\| \leqslant \|M^k\| \|y_0\|$$

平均收敛速度

- 单步线性定常迭代法：$x_k = Mx_{k-1} + g$
- 误差向量$y_k = x_k - x_*$满足$y_k = M^k y_0$，从而有

$$\|y_k\| \leqslant \|M^k\| \|y_0\|$$

- 初始误差$\|y_0\|$一般不知道，通常用$\|M^k\|$的大小刻画迭代法收敛的速度

平均收敛速度

线性方程组的古典迭代解法

邓建松

单步线性定常迭代法

Jacobi迭代法

Gauss-Seidel迭代法

形式推广

收敛性理论

收敛的充要条件

收敛的充分条件及误差估计

Jacobi迭代法和G-S迭代法的收敛性

收敛速度

平均收敛速度和渐近收敛速度

模型问题

Jacobi迭代法和G-S迭代法的渐近收敛速度

超松弛迭代法

- 单步线性定常迭代法：$x_k = Mx_{k-1} + g$
- 误差向量$y_k = x_k - x_*$满足$y_k = M^k y_0$，从而有

$$\|y_k\| \leqslant \|M^k\|\|y_0\|$$

- 初始误差$\|y_0\|$一般不知道，通常用$\|M^k\|$的大小刻画迭代法收敛的速度
- 定义$R_k(M) = \dfrac{-\ln\|M^k\|}{k}$为$k$次迭代的平均收敛速度

误差缩减的比例因子

- 若迭代法收敛，则当$k \to \infty$时，$\|M^k\| \to 0$

误差缩减的比例因子

- 若迭代法收敛，则当$k \to \infty$时，$\|M^k\| \to 0$
- 所以当k充分大时，总有$R_k(M) > 0$.

误差缩减的比例因子

- 若迭代法收敛，则当$k \to \infty$时，$\|M^k\| \to 0$
- 所以当k充分大时，总有$R_k(M) > 0$.
- 设$R_k = R_k(M) > 0$, 数量

$$\sigma = \left(\frac{\|y_k\|}{\|y_0\|}\right)^{1/k} \approx \|M^k\|^{1/k} = e^{-R_k}$$

就表示误差范数在k次迭代中平均每次迭代所缩减的比例因子

对称矩阵情形

- 若M是对称矩阵，或者Hermite矩阵，或者正规矩阵，则显然有

$$\|M^k\|_2 = (\rho(M))^k$$

对称矩阵情形

- 若M是对称矩阵，或者Hermite矩阵，或者正规矩阵，则显然有

$$\|M^k\|_2 = (\rho(M))^k$$

- 所以$R_k(M) = -\ln\rho(M)$, 这是与k无关的量

对称矩阵情形

- 若M是对称矩阵，或者Hermite矩阵，或者正规矩阵，则显然有

$$\|M^k\|_2 = (\rho(M))^k$$

- 所以$R_k(M) = -\ln \rho(M)$, 这是与k无关的量
- 但在一般情况下，R_k是依赖于k的，这时计算R_k就非常复杂

收敛速度比较

- 如果两个迭代矩阵G和H满足$R_k(H) > R_k(M) > 0$，我们就说，对于k次迭代来讲，对应于H的迭代法比对应于G的迭代法的收敛速度快

收敛速度比较

- 如果两个迭代矩阵G和H满足$R_k(H) > R_k(M) > 0$, 我们就说，对于k次迭代来讲，对应于H的迭代法比对应于G的迭代法的收敛速度快
- 有可能对某些k, $R_k(G) < R_k(H)$, 而对另一些k, $R_k(G) > R_k(H)$

收敛速度比较

- 如果两个迭代矩阵G和H满足$R_k(H) > R_k(M) > 0$，我们就说，对于k次迭代来讲，对应于H的迭代法比对应于G的迭代法的收敛速度快
- 有可能对某些k，$R_k(G) < R_k(H)$，而对另一些k，$R_k(G) > R_k(H)$
- 所以我们考虑$k \to \infty$时$R_k(M)$的极限

渐近收敛速度

- 为了刻画整个迭代过程的收敛速度，我们定义

$$R_\infty(M) = \lim_{k\to\infty} R_k(M)$$

这称为渐近收敛速度

渐近收敛速度

- 为了刻画整个迭代过程的收敛速度，我们定义

$$R_\infty(M) = \lim_{k\to\infty} R_k(M)$$

这称为渐近收敛速度

- 我们有

渐近收敛速度

- 为了刻画整个迭代过程的收敛速度，我们定义

$$R_\infty(M) = \lim_{k\to\infty} R_k(M)$$

这称为渐近收敛速度

- 我们有

定理

$$R_\infty(M) = -\ln \rho(M)$$

渐近收敛速度

- 为了刻画整个迭代过程的收敛速度，我们定义

$$R_\infty(M) = \lim_{k\to\infty} R_k(M)$$

这称为**渐近收敛速度**

- 我们有

定理

$$R_\infty(M) = -\ln \rho(M)$$

- 定理的成立不依赖于范数的选取

证明

- 只需证明 $\lim_{k\to\infty} \|M^k\|^{1/k} = \rho(M)$ 即可

证明

- 只需证明 $\lim_{k\to\infty} \|M^k\|^{1/k} = \rho(M)$ 即可
- 因为$(\rho(M))^k = \rho(M^k) \leqslant \|M^k\|$, 所以$\rho(M) \leqslant \|M^k\|^{1/k}$

证明

- 只需证明 $\lim_{k\to\infty} \|M^k\|^{1/k} = \rho(M)$ 即可
- 因为$(\rho(M))^k = \rho(M^k) \leqslant \|M^k\|$, 所以$\rho(M) \leqslant \|M^k\|^{1/k}$
- 另一方面，对$\forall \varepsilon > 0$, 考虑矩阵

$$B_\varepsilon = \frac{1}{\rho(M) + \varepsilon} M$$

证明

- 只需证明$\lim\limits_{k\to\infty}\|M^k\|^{1/k}=\rho(M)$即可
- 因为$(\rho(M))^k=\rho(M^k)\leqslant\|M^k\|$, 所以$\rho(M)\leqslant\|M^k\|^{1/k}$
- 另一方面，对$\forall\varepsilon>0$, 考虑矩阵
$$B_\varepsilon=\frac{1}{\rho(M)+\varepsilon}M$$
- 显然$\rho(B_\varepsilon)<1$

- 于是 $\lim_{k\to\infty} B_\varepsilon^k = 0$

- 于是$\lim_{k\to\infty} B_\varepsilon^k = 0$
- 所以存在自然数K，当$k \geqslant K$时有$\|B_\varepsilon^k\| \leqslant 1$

- 于是$\lim\limits_{k\to\infty} B_\varepsilon^k = 0$
- 所以存在自然数K，当$k \geqslant K$时有$\|B_\varepsilon^k\| \leqslant 1$
- 这就是$\|M^k\| \leqslant (\rho(M)+\varepsilon)^k$

- 于是$\lim\limits_{k\to\infty} B_\varepsilon^k = 0$
- 所以存在自然数K，当$k \geqslant K$时有$\|B_\varepsilon^k\| \leqslant 1$
- 这就是$\|M^k\| \leqslant (\rho(M)+\varepsilon)^k$
- 至此我们证明了：对任意的$\varepsilon > 0$，存在自然数K，当$k \geqslant K$时，

$$\rho(M) \leqslant \|M^k\|^{1/k} \leqslant \rho(M) + \varepsilon$$

这就完成了证明

模型问题

- 为了具体说明迭代法的收敛速度，我们考虑用五点差分格式求解$[0,1]^2$上的Poisson方程第一边值问题

$$\begin{cases} \triangle u = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = -f(x,y), (x,y) \in (0,1)^2 \\ u|_\Gamma = 0 \end{cases}$$

这里Γ表示正方形定义域的边界

离散化

- 把正方形的每边n等分，令$h = 1/n$, 用等分线把正方形$[0,1]^2$分割成n^2个小正方形

离散化

- 把正方形的每边n等分，令$h = 1/n$, 用等分线把正方形$[0, 1]^2$分割成n^2个小正方形
- 记小正方形的顶点为(x_i, y_j)

离散化

- 把正方形的每边n等分，令$h = 1/n$, 用等分线把正方形$[0, 1]^2$分割成n^2个小正方形
- 记小正方形的顶点为(x_i, y_j)
- 用二阶差商

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{(x_i,y_j)} = \frac{1}{h^2}(u_{i-1,j} - 2u_{i,j} + u_{i+1,j}) + O(h^2)$$

$$\left.\frac{\partial^2 u}{\partial y^2}\right|_{(x_i,y_j)} = \frac{1}{h^2}(u_{i,j-1} - 2u_{i,j} + u_{i,j+1}) + O(h^2)$$

代替二阶偏微分，这里$u_{i,j} = u(x_i, y_j)$

矩阵形式

- 如此得到方程组

$$\begin{cases} 4u_{i,j} - (u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) = h^2 f_{ij} \\ \qquad i,j = 1,\dots,n-1 \\ u_{i,0} = u_{i,n} = u_{0,j} = u_{n,j} = 0, \quad i,j = 0,\dots,n \end{cases}$$

矩阵形式

- 如此得到方程组

$$\begin{cases} 4u_{i,j} - (u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) = h^2 f_{ij} \\ \qquad i,j = 1,\ldots,n-1 \\ u_{i,0} = u_{i,n} = u_{0,j} = u_{n,j} = 0, \quad i,j = 0,\ldots,n \end{cases}$$

- 写成矩阵形式：

$$T_{n-1}U + UT_{n-1} = h^2 F,$$

其中$U = (u_{i,j})$, $F = (f_{ij})$, T_{n-1}为$n-1$阶三对角矩阵矩阵，主对角元素为2, 上下次对角元素为-1

矩阵拉直：自然顺序排列

- 为了得到通常的$Ax = b$形式的线性方程组，我们对U和F中的元素进行“拉直”：先按j由小到大排列，j相同的按i由小到大排列

矩阵拉直：自然顺序排列

- 为了得到通常的$Ax = b$形式的线性方程组，我们对U和F中的元素进行“拉直”：先按j由小到大排列，j相同的按i由小到大排列
- 如此我们得到方程组

$$Au = h^2 f$$

其中向量u和f是矩阵U和F拉直的结果

系数矩阵A

$$A = \begin{pmatrix} T_{n-1} + 2I_{n-1} & -I_{n-1} & & \\ -I_{n-1} & \ddots & \ddots & \\ & \ddots & \ddots & -I_{n-1} \\ & & -I_{n-1} & T_{n-1} + 2I_{n-1} \end{pmatrix}$$

- A是$(n-1)^2$阶块三对角阵，五条对角线上有非零元；它也是不可约对角占优的对称正定阵

系数矩阵A

$$A = \begin{pmatrix} T_{n-1} + 2I_{n-1} & -I_{n-1} & & \\ -I_{n-1} & \ddots & \ddots & \\ & \ddots & \ddots & -I_{n-1} \\ & & -I_{n-1} & T_{n-1} + 2I_{n-1} \end{pmatrix}$$

- A是$(n-1)^2$阶块三对角阵，五条对角线上有非零元；它也是不可约对角占优的对称正定阵
- 每个对角元的左、右各有两个非零元素，据离对角元远近，分别对应对角元上下和左右邻居

A的特征值和特征向量

- T_{n-1}的特征值为$\lambda_j = 2 - 2\cos\dfrac{j\pi}{n}$, 对应的单位特征向量为

$$z_j = \left(\sqrt{\frac{2}{n}}\sin\frac{j\pi}{n}, \sqrt{\frac{2}{n}}\sin\frac{2j\pi}{n}, \ldots, \sqrt{\frac{2}{n}}\sin\frac{(n-1)j\pi}{n}\right)^T$$

A的特征值和特征向量

- T_{n-1}的特征值为$\lambda_j = 2 - 2\cos\dfrac{j\pi}{n}$, 对应的单位特征向量为

$$z_j = \left(\sqrt{\frac{2}{n}}\sin\frac{j\pi}{n}, \sqrt{\frac{2}{n}}\sin\frac{2j\pi}{n}, \ldots, \sqrt{\frac{2}{n}}\sin\frac{(n-1)j\pi}{n}\right)^T$$

- 利用“拉直”操作，可以证明A的特征值为$\lambda_{pq} = \lambda_p + \lambda_q$, 对应特征向量为$z_p z_q^T$“拉直”的结果

Jacobi迭代法

- 针对上述模型问题，Jacobi迭代法的迭代矩阵为$B = D^{-1}(L+U) = I - \frac{1}{4}A$

Jacobi迭代法

- 针对上述模型问题，Jacobi迭代法的迭代矩阵为$B = D^{-1}(L+U) = I - \frac{1}{4}A$
- B的对角元为0，其左右分别有两个非零元素（值为1/4），对应于对角元的左右和上下邻居

Jacobi迭代法

- 针对上述模型问题，Jacobi迭代法的迭代矩阵为$B = D^{-1}(L+U) = I - \frac{1}{4}A$
- B的对角元为0，其左右分别有两个非零元素（值为1/4），对应于对角元的左右和上下邻居
- 所以迭代格式为

$$u_{ij}^{(k)} = \frac{1}{4}\left(u_{i+1,j}^{(k-1)} + u_{i-1,j}^{(k-1)} + u_{i,j+1}^{(k-1)} + u_{i,j-1}^{(k-1)}\right) + \frac{h^2}{4} f_{ij},$$

$$u_{i0} = u_{in} = u_{0j} = u_{nj} = 0, i,j = 1,\ldots,n-1$$

Jacobi迭代法的渐近收敛速度

- B的特征值为$\mu_{pq} = 1 - \frac{1}{4}\lambda_{pq}$

Jacobi迭代法的渐近收敛速度

- B的特征值为$\mu_{pq} = 1 - \dfrac{1}{4}\lambda_{pq}$
  - 从而若μ为B的特征值，则$-\mu$也是B的特征值

Jacobi迭代法的渐近收敛速度

- B的特征值为$\mu_{pq} = 1 - \dfrac{1}{4}\lambda_{pq}$
  - 从而若μ为B的特征值，则$-\mu$也是B的特征值
- 所以$\rho(B) = \cos\dfrac{\pi}{n} = \cos h\pi$, 从而可知渐近收敛速度为

$$R_\infty(B) = -\ln\rho(B) = -\ln\cos h\pi$$

Jacobi迭代法的渐近收敛速度

- B的特征值为$\mu_{pq} = 1 - \dfrac{1}{4}\lambda_{pq}$
  - 从而若μ为B的特征值，则$-\mu$也是B的特征值
- 所以$\rho(B) = \cos\dfrac{\pi}{n} = \cos h\pi$, 从而可知渐近收敛速度为

$$R_\infty(B) = -\ln\rho(B) = -\ln\cos h\pi$$

- 从而当$h \to 0$时，我们有

$$R_\infty(B) \sim \frac{1}{2}\pi^2 h^2$$

B对应的特征值问题

- 我们这里要基于B的特征值问题给出L_1的特征值

B对应的特征值问题

- 我们这里要基于B的特征值问题给出L_1的特征值
- B的特征值问题为$B\eta = \mu\eta$

B对应的特征值问题

- 我们这里要基于B的特征值问题给出L_1的特征值
- B的特征值问题为$B\eta=\mu\eta$
- 回忆：B的对角元全是零，每个对角元的左、右各有两个非零元素（值为1/4），所以采用二重指标表示B对应的特值问题为

$$\begin{cases} \mu\eta_{ij} = \dfrac{1}{4}\left(\eta_{i+1,j}+\eta_{i-1,j}+\eta_{i,j+1}+\eta_{i,j-1}\right) \\ \eta_{i0}=\eta_{in}=\eta_{0j}=\eta_{nj}=0 \end{cases}$$

L_1对应的特征值问题

- L_1对应的特征值问题为

$$L_1\xi = \lambda\xi \to \lambda D\xi = \lambda L\xi + U\xi$$

L_1对应的特征值问题

- L_1对应的特征值问题为

$$L_1\xi = \lambda\xi \to \lambda D\xi = \lambda L\xi + U\xi$$

- 所以采用二重指标表示L_1的特征值问题为

$$\begin{cases} \lambda\xi_{ij} = \dfrac{1}{4}(\lambda\xi_{i-1,j} + \lambda\xi_{i,j-1} + \xi_{i,j+1} + \xi_{i+1,j}) \\ \xi_{i0} = \xi_{in} = \xi_{0j} = \xi_{nj} = 0 \end{cases}$$

- 设$\lambda \neq 0$, 作变换$\xi_{ij} = \lambda^{(i+j)/2}\eta_{ij}$ 得

$$\begin{aligned}\lambda^{1+\frac{i+j}{2}}\eta_{ij} =& \frac{1}{4}\Big(\lambda^{1+(i+j-1)/2}\eta_{i-1,j} + \lambda^{1+(i+j-1)/2}\eta_{i,j-1} \\ &+\lambda^{(i+j+1)/2}\eta_{i+1,j} + \lambda^{(i+j+1)/2}\eta_{i,j+1}\Big) \\ =& \frac{1}{4}\lambda^{\frac{i+j+1}{2}}\left(\eta_{i-1,j} + \eta_{i,j-1} + \eta_{i+1,j} + \eta_{i,j+1}\right)\end{aligned}$$

- 设$\lambda \neq 0$, 作变换$\xi_{ij} = \lambda^{(i+j)/2}\eta_{ij}$ 得

$$\begin{aligned}\lambda^{1+\frac{i+j}{2}}\eta_{ij} =&\frac{1}{4}\left(\lambda^{1+(i+j-1)/2}\eta_{i-1,j} + \lambda^{1+(i+j-1)/2}\eta_{i,j-1}\right.\\&\left.+\lambda^{(i+j+1)/2}\eta_{i+1,j} + \lambda^{(i+j+1)/2}\eta_{i,j+1}\right)\\=&\frac{1}{4}\lambda^{\frac{i+j+1}{2}}\left(\eta_{i-1,j}+\eta_{i,j-1}+\eta_{i+1,j}+\eta_{i,j+1}\right)\end{aligned}$$

- 由边值为零，可知若λ是L_1的非零特征值当且仅当$\lambda^{1/2}$是B的非零特征值

- 设$\lambda \neq 0$, 作变换$\xi_{ij} = \lambda^{(i+j)/2}\eta_{ij}$ 得

$$\begin{aligned}\lambda^{1+\frac{i+j}{2}}\eta_{ij} =& \frac{1}{4}\left(\lambda^{1+(i+j-1)/2}\eta_{i-1,j} + \lambda^{1+(i+j-1)/2}\eta_{i,j-1}\right.\\ &\left.+\lambda^{(i+j+1)/2}\eta_{i+1,j} + \lambda^{(i+j+1)/2}\eta_{i,j+1}\right)\\ =& \frac{1}{4}\lambda^{\frac{i+j+1}{2}}\left(\eta_{i-1,j} + \eta_{i,j-1} + \eta_{i+1,j} + \eta_{i,j+1}\right)\end{aligned}$$

- 由边值为零，可知若λ是L_1的非零特征值当且仅当$\lambda^{1/2}$是B的非零特征值
- 因此L_1的特征值非负

G-S迭代法的渐近收敛速度

- 从而我们可知

$$R_\infty(L_1) = -\ln\rho(L_1) = -2\ln\rho(B) \sim \pi^2 h^2, \quad h \to 0$$

G-S迭代法的渐近收敛速度

- 从而我们可知

$$R_\infty(L_1) = -\ln\rho(L_1) = -2\ln\rho(B) \sim \pi^2 h^2, \quad h \to 0$$

- 这说明G-S迭代法的渐近收敛速度是Jacobi迭代法的渐近收敛速度的两倍

超松弛迭代法

- 这是一种新的迭代法

超松弛迭代法

- 这是一种新的迭代法
- 它是G-S迭代法的引申和推广

超松弛迭代法

- 这是一种新的迭代法
- 它是G-S迭代法的引申和推广
- 也可以看作是G-S迭代法的加速

迭代格式的重新认识

- 在迭代格式中令$x_{k+1}-x_k=\Delta x$，即$x_{k+1}=x_k+\Delta x$

迭代格式的重新认识

- 在迭代格式中令$x_{k+1} - x_k = \Delta x$，即$x_{k+1} = x_k + \Delta x$
- 这可以看作是在向量x_k上加上修正项Δx得到x_{k+1}

迭代格式的重新认识

- 在迭代格式中令$x_{k+1} - x_k = \Delta x$，即$x_{k+1} = x_k + \Delta x$
- 这可以看作是在向量x_k上加上修正项Δx得到x_{k+1}
- 对于已有的迭代格式，Δx是由格式完全确定的

迭代格式的重新认识

- 在迭代格式中令$x_{k+1} - x_k = \Delta x$，即$x_{k+1} = x_k + \Delta x$
- 这可以看作是在向量x_k上加上修正项Δx得到x_{k+1}
- 对于已有的迭代格式，Δx是由格式完全确定的
- 我们可以在修正项的前面加上一个参数ω以得到松弛迭代格式

G-S迭代法对应的松弛格式

- G-S迭代格式：

$$x_{k+1} = D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b$$

G-S迭代法对应的松弛格式

- G-S迭代格式:

  $$x_{k+1} = D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b$$

- $\Delta x = D^{-1}Lx_{k+1} + (D^{-1}U - I)x_k + D^{-1}b$

G-S迭代法对应的松弛格式

- G-S迭代格式：
  $$x_{k+1} = D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b$$
- $\Delta x = D^{-1}Lx_{k+1} + (D^{-1}U - I)x_k + D^{-1}b$
- 松弛迭代格式

$$\begin{aligned} x_{k+1} &= x_k + \omega\Delta x \\ &= (1-\omega)x_k + \omega(D^{-1}Lx_{k+1} + D^{-1}Ux_k + D^{-1}b) \end{aligned}$$

术语

- ω叫做松弛因子

术语

- ω叫做松弛因子
- 当$\omega > 1$时，对应的格式叫做超松弛迭代法，简记为SOR迭代法

术语

- ω叫做松弛因子
- 当$\omega > 1$时，对应的格式叫做超松弛迭代法，简记为SOR迭代法
  - SOR — Successive Over–Relaxation

术语

- ω叫做松弛因子
- 当$\omega > 1$时，对应的格式叫做超松弛迭代法，简记为SOR迭代法
  - SOR — Successive Over–Relaxation
- 当$\omega < 1$时，对应的格式叫做低松弛迭代法

术语

- ω叫做松弛因子
- 当$\omega > 1$时，对应的格式叫做超松弛迭代法，简记为SOR迭代法
  - SOR — Successive Over–Relaxation
- 当$\omega < 1$时，对应的格式叫做低松弛迭代法
- 当$\omega = 1$时，对应的格式就是G-S迭代法

迭代矩阵

- 如果$(I-\omega D^{-1}L)^{-1}$存在，松弛迭代格式可以写为

$$x_{k+1}=L_\omega x_k+\omega(D-\omega L)^{-1}b$$

其中

$$L_\omega=(D-\omega L)^{-1}((1-\omega)D+\omega U)$$

称为松弛迭代格式的迭代矩阵

迭代矩阵

- 如果$(I-\omega D^{-1}L)^{-1}$存在，松弛迭代格式可以写为
$$x_{k+1}=L_\omega x_k+\omega(D-\omega L)^{-1}b$$
其中
$$L_\omega=(D-\omega L)^{-1}((1-\omega)D+\omega U)$$
称为松弛迭代格式的迭代矩阵
- 注意：$(D-\omega L)^{-1}$为下三角矩阵，$((1-\omega)D+\omega U$为上三角矩阵

收敛定理

定理

*SOR*迭代法收敛的充分必要条件是

$$\rho(L_\omega) < 1$$

定理

*SOR*迭代法收敛的必要条件是$0 < \omega < 2$

第一个定理显然。下面证明第二个

$0 < \omega < 2$的证明

- 迭代法收敛，所以$\rho(L_\omega) < 1$

$0 < \omega < 2$的证明

- 迭代法收敛，所以$\rho(L_\omega) < 1$
- 若λ_i为L_ω的n个特征值，则

$$|\det L_\omega| = |\lambda_1 \cdots \lambda_n| < 1$$

$0 < \omega < 2$的证明

- 迭代法收敛，所以$\rho(L_\omega) < 1$
- 若λ_i为L_ω的n个特征值，则

$$|\det L_\omega| = |\lambda_1 \cdots \lambda_n| < 1$$

- 注意到

$$L_\omega = (I - \omega D^{-1}L)^{-1}((1-\omega)I + \omega D^{-1}U),$$

$$\det((1-\omega)I + \omega D^{-1}U) = (1-\omega)^n,$$

$$\det(I - \omega D^{-1}L)^{-1} = 1$$

- 从而有

$$|\det L_\omega| = |(1-\omega)^n| < 1$$

- 从而有
$$|\det L_\omega| = |(1-\omega)^n| < 1$$
- 即$|1-\omega| < 1$，也就是$0 < \omega < 2$

- 从而有
$$|\det L_\omega| = |(1-\omega)^n| < 1$$
- 即$|1-\omega| < 1$，也就是$0 < \omega < 2$
- 这个结果说明，要使SOR迭代法收敛，必须选取收敛因子$\omega \in (0,2)$

充分条件：对角占优

定理

若系数矩阵A是严格对角占优的，或者不可约对角占优的，且松弛因子$\omega \in (0, 1)$，则SOR迭代法收敛

证明关键

- 取λ, $|\lambda| \geqslant 1$,

$$\lambda I - L_\omega = (D-\omega L)^{-1}((\lambda+\omega-1)D - \lambda\omega L - \omega U)$$

证明关键

- 取λ, $|\lambda| \geqslant 1$,

$$\lambda I - L_\omega = (D - \omega L)^{-1}((\lambda + \omega - 1)D - \lambda\omega L - \omega U)$$

- 根据三角不等式:

$$\begin{aligned} |\lambda + \omega - 1| &\geqslant |\lambda| - (1 - \omega) \\ &= |\lambda|\omega + (|\lambda| - 1)(1 - \omega) \geqslant |\lambda|\omega \end{aligned}$$

证明关键

- 取λ, $|\lambda| \geqslant 1$,

$$\lambda I - L_\omega = (D - \omega L)^{-1}((\lambda + \omega - 1)D - \lambda\omega L - \omega U)$$

- 根据三角不等式:

$$\begin{aligned}|\lambda + \omega - 1| &\geqslant |\lambda| - (1 - \omega) \\ &= |\lambda|\omega + (|\lambda| - 1)(1 - \omega) \geqslant |\lambda|\omega\end{aligned}$$

- 所以$D - \omega L$, $(\lambda + \omega - 1)D - \lambda\omega L - \omega U$ 都是严格对角占优的，或者不可约对角占优的，从而是非奇异的

充分条件：对称正定

定理

若系数矩阵A是实对称的正定矩阵，则当$0 < \omega < 2$时，SOR方法收敛

根据该定理可知，当A对称正定时，G-S迭代法收敛

证明

- 设λ是L_ω的任一特征值（可以是复数），x为对应的特征向量，则(注意$L^T = U$)

$$((1-\omega)D + \omega L^T)x = \lambda(D - \omega L)x$$

证明

- 设λ是L_ω的任一特征值（可以是复数），x为对应的特征向量，则(注意$L^T = U$)

$$((1-\omega)D+\omega L^T)x = \lambda(D-\omega L)x$$

- 左乘x的共轭转置向量x^*:

$$x^*((1-\omega)D+\omega L^T)x = \lambda x^*(D-\omega L)x$$

证明

- 设λ是L_ω的任一特征值（可以是复数），x为对应的特征向量，则(注意$L^T = U$)

$$((1-\omega)D + \omega L^T)x = \lambda(D - \omega L)x$$

- 左乘x的共轭转置向量x^*:

$$x^*((1-\omega)D + \omega L^T)x = \lambda x^*(D - \omega L)x$$

- 令$x^*Dx = \delta$, $x^*Lx = \alpha + \mathrm{i}\beta$, 则有$x^*L^Tx = \alpha - \mathrm{i}\beta$

- 我们有

$$(1-\omega)\delta+\omega(\alpha-\mathrm{i}\beta)=\lambda(\delta-\omega(\alpha+\mathrm{i}\beta))$$

- 我们有

$$(1-\omega)\delta+\omega(\alpha-\mathrm{i}\beta)=\lambda(\delta-\omega(\alpha+\mathrm{i}\beta))$$

- 两边取模得到

$$|\lambda|^2=\frac{((1-\omega)\delta+\omega\alpha)^2+\omega^2\beta^2}{(\delta-\omega\alpha)^2+\omega^2\beta^2}$$

- 我们有

$$(1-\omega)\delta+\omega(\alpha-\mathrm{i}\beta)=\lambda(\delta-\omega(\alpha+\mathrm{i}\beta))$$

- 两边取模得到

$$|\lambda|^2=\frac{((1-\omega)\delta+\omega\alpha)^2+\omega^2\beta^2}{(\delta-\omega\alpha)^2+\omega^2\beta^2}$$

- 右侧的分子和分母相减，得到$\omega\delta(\delta-2\alpha)(\omega-2)$

- 我们有
$$(1-\omega)\delta+\omega(\alpha-\mathrm{i}\beta)=\lambda(\delta-\omega(\alpha+\mathrm{i}\beta))$$
- 两边取模得到
$$|\lambda|^2=\frac{((1-\omega)\delta+\omega\alpha)^2+\omega^2\beta^2}{(\delta-\omega\alpha)^2+\omega^2\beta^2}$$
- 右侧的分子和分母相减，得到$\omega\delta(\delta-2\alpha)(\omega-2)$
- 当A正定时$\delta>0$，$\delta-2\alpha=x^*Ax>0$，所以当$0<\omega<2$时$|\lambda|<1$，即SOR迭代法收敛

最佳收敛因子

- SOR迭代法的谱半径依赖于ω

最佳收敛因子

- SOR迭代法的谱半径依赖于ω
- 如何选取恰当的ω, 从而使得收敛速度最快?

最佳收敛因子

- SOR迭代法的谱半径依赖于ω
- 如何选取恰当的ω，从而使得收敛速度最快？
- 为此我们考虑特征值问题$L_\omega x = \lambda x$，即

$$((\lambda - 1 + \omega)D - \lambda\omega L - \omega U)x = 0$$

最佳收敛因子

- SOR迭代法的谱半径依赖于ω
- 如何选取恰当的ω，从而使得收敛速度最快？
- 为此我们考虑特征值问题$L_\omega x = \lambda x$，即

$$((\lambda - 1 + \omega)D - \lambda\omega L - \omega U)x = 0$$

- 回忆：L_ω的所有特征值相乘等于$(1-\omega)^n$，所以当$\omega \neq 1$时无零特征值

仍是那个模型问题

- 对于模型问题，上述问题为

$$(\lambda+\omega-1)u_{ij} = \frac{\omega}{4}(\lambda u_{i-1,j} + \lambda u_{i,j-1} + u_{i+1,j} + u_{i,j+1}$$

$$u_{i0} = u_{in} = u_{0j} = u_{nj} = 0,$$

仍是那个模型问题

- 对于模型问题，上述问题为

$$(\lambda+\omega-1)u_{ij}=\frac{\omega}{4}(\lambda u_{i-1,j}+\lambda u_{i,j-1}+u_{i+1,j}+u_{i,j+1}$$

$$u_{i0}=u_{in}=u_{0j}=u_{nj}=0,$$

- 下面分两步讨论问题：

仍是那个模型问题

- 对于模型问题，上述问题为

$$(\lambda+\omega-1)u_{ij} = \frac{\omega}{4}(\lambda u_{i-1,j} + \lambda u_{i,j-1} + u_{i+1,j} + u_{i,j+1}$$

$$u_{i0} = u_{in} = u_{0j} = u_{nj} = 0,$$

- 下面分两步讨论问题：
  - B与L_ω的特征值之间的关系

仍是那个模型问题

- 对于模型问题，上述问题为

$$(\lambda+\omega-1)u_{ij}=\frac{\omega}{4}(\lambda u_{i-1,j}+\lambda u_{i,j-1}+u_{i+1,j}+u_{i,j+1}$$

$$u_{i0}=u_{in}=u_{0j}=u_{nj}=0,$$

- 下面分两步讨论问题：
  - B与L_ω的特征值之间的关系
  - $\rho(L_\omega)$随ω变化的情况

B与L_ω的特征值之间的关系

- 当$\lambda \neq 0$时，作变换$u_{ij} = (\pm\lambda^{1/2})^{i+j} v_{ij}$, 则有

$$\mu v_{ij} - \frac{1}{2}(v_{i-1,j} + v_{i,j-1} + v_{i+1,j} + v_{i,j+1}) = 0,$$

其中$\mu = \pm\dfrac{\lambda + \omega - 1}{\omega\lambda^{1/2}}$

B与L_ω的特征值之间的关系

- 当$\lambda \neq 0$时，作变换$u_{ij} = (\pm\lambda^{1/2})^{i+j} v_{ij}$, 则有

$$\mu v_{ij} - \frac{1}{2}(v_{i-1,j} + v_{i,j-1} + v_{i+1,j} + v_{i,j+1}) = 0,$$

其中$\mu = \pm\dfrac{\lambda + \omega - 1}{\omega\lambda^{1/2}}$

- 当$\omega \neq 1$时，若λ是L_ω的特征值，则由

$$(\lambda + \omega - 1)^2 = \mu^2\omega^2\lambda$$

确定的两个μ都是B的特征值；反之亦然

- 当$\omega=1$时，前式简化为$\lambda^2=\mu^2\lambda$

- 当$\omega = 1$时，前式简化为$\lambda^2 = \mu^2\lambda$
- 当B的特征值为$\pm\mu_i$时，对应于L_1的特征值为0和μ_i^2

- 当$\omega = 1$时，前式简化为$\lambda^2 = \mu^2\lambda$
- 当B的特征值为$\pm\mu_i$时，对应于L_1的特征值为0和μ_i^2
- 回忆：若μ是B的特征值，则$-\mu$也是B的特征值

$\rho(L_\omega)$随ω变化的情况

- 设$0 \leqslant \mu < 1$是B的一个特征值，$0 < \omega < 2$, 则由方程
$$\lambda + \omega - 1 = \pm\mu\omega\lambda^{1/2}$$
可得L_ω的两个特征值分别是
$$\lambda_+(\omega, \mu) = \left(\frac{\mu\omega}{2} + \sqrt{\left(\frac{\mu\omega}{2}\right)^2 - (\omega - 1)}\right)^2,$$
$$\lambda_-(\omega, \mu) = \left(\frac{\mu\omega}{2} - \sqrt{\left(\frac{\mu\omega}{2}\right)^2 - (\omega - 1)}\right)^2,$$

- 为了研究$\rho(L_\omega)$随ω的变化，我们考虑

$$M(\omega,\mu)=\max(|\lambda_+(\omega,\mu)|,|\lambda_-(\omega,\mu)|)$$

- 为了研究$\rho(L_\omega)$随ω的变化，我们考虑

$$M(\omega,\mu)=\max(|\lambda_+(\omega,\mu)|,|\lambda_-(\omega,\mu)|)$$

- 分析细节忽略

- 为了研究$\rho(L_\omega)$随ω的变化，我们考虑

$$M(\omega,\mu) = \max(|\lambda_+(\omega,\mu)|, |\lambda_-(\omega,\mu)|)$$

- 分析细节忽略
- 结论：随着ω从0开始增加，$\rho(L_\omega)$减少，直至

$$\omega = \frac{2}{1+\sqrt{1-\rho(B)^2}}$$

时$\rho(L_\omega)$达到极小，然后开始变大。这个ω称为最佳松弛因子

- 当ω取上述最佳松弛因子时，我们有

$$
\begin{aligned}
R_\infty(L_\omega) &= -\ln \rho(L_\omega) \\
&= -\ln \frac{1-\sqrt{1-\rho(B)^2}}{1+\sqrt{1-\rho(B)^2}} \\
&= -\ln \frac{1-\sin h\pi}{1+\sin h\pi} \sim 2h\pi,\ h \to 0
\end{aligned}
$$

- 当ω取上述最佳松弛因子时，我们有

$$
\begin{aligned}
R_\infty(L_\omega) &= -\ln \rho(L_\omega) \\
&= -\ln \frac{1-\sqrt{1-\rho(B)^2}}{1+\sqrt{1-\rho(B)^2}} \\
&= -\ln \frac{1-\sin h\pi}{1+\sin h\pi} \sim 2h\pi, \ h \to 0
\end{aligned}
$$

- SOR迭代法要比Jacobi迭代法和G-S迭代法快得多！