清華大學深圳国际研究生院
Tsinghua Shenzhen International Graduate School

工程硕士数学

第三章 线性方程组的迭代解法

胡振中 副教授

海洋大楼 702-3，15999640239（微信同号）

邮箱：huzhenzhong@tsinghua.edu.cn

个人网站：http://www.huzhenzhong.net

内容提纲

3.1 概论及背景知识

3.2 Jacobi迭代法和Gauss-Seidel迭代法

3.3 超松弛迭代法

3.4 共轭梯度法*

3.5 编程实践（穿插在两周中）

本章研究的问题

对于第二章所定义的有唯一解的线性方程组$\mathbf{A}\vec{x}=\vec{b}$，其中

$$\mathbf{A}\in\mathbb{R}^{n\times n}\qquad \det|\mathbf{A}|\neq 0\qquad \vec{b}=\mathbb{R}^{n}$$

我们把方程改写成 $\vec{x}=\mathbf{B}\vec{x}+\vec{f}$

设一个初始值 $\vec{x}^{(0)}$，以此构造迭代法

迭代矩阵

$$\vec{x}^{(1)}=\mathbf{B}\vec{x}^{(0)}+\vec{f}\ \Rightarrow\ \vec{x}^{(2)}=\mathbf{B}\vec{x}^{(1)}+\vec{f}\ \Rightarrow\ \cdots\cdots\ \Rightarrow\ \vec{x}^{(k)}=\boxed{\mathbf{B}}\vec{x}^{(k-1)}+\vec{f}$$

迭代公式会产生一个序列$\left\{\vec{x}^{(k)}\right\}$，与方程的解 $\vec{x}^{*}$ 关系？

本章研究的问题

即，我们要研究：$\lim\limits_{k\to\infty}\vec{x}^{(k)}=\vec{x}^*$ 是否成立？

或者说，我们要研究：$\lim\limits_{k\to\infty}\{\vec{x}^{(k)}-\vec{x}^*\}=0$ 是否成立？

由方程的确定解 $\vec{x}^*$ 带入改造以后的方程，可得：$\vec{x}^*=\mathbf{B}\vec{x}^*+\vec{f}$

$$\vec{x}^{(k)}-\vec{x}^*=\mathbf{B}\vec{x}^{(k-1)}+\vec{f}-\vec{x}^*=\mathbf{B}\vec{x}^{(k-1)}-\mathbf{B}\vec{x}^*=\mathbf{B}\left(\vec{x}^{(k-1)}-\vec{x}^*\right)=\mathbf{B}^2\left(\vec{x}^{(k-2)}-\vec{x}^*\right)=\ldots$$

$$\boxed{\vec{x}^{(k)}-\vec{x}^*=\boxed{\mathbf{B}^k}\left(\vec{x}^{(0)}-\vec{x}^*\right)}$$

$\forall\vec{x}^{(0)}$ 该式是否趋近于0向量？即$\mathbf{B}^k$是否趋近于**0矩阵**！

本章研究的问题

- 构造方法 $\vec{x} = \mathbf{B}\vec{x} + \vec{f}$ (**构造B**)
  - ✓ Jacobi方法
  - ✓ Gauss-Seidel方法
  - ✓ 超松弛方法
  - ✓ 共轭梯度法
- 研究迭代法的**收敛性**
  - ✓ 即矩阵**B**满足什么条件可以使得 $\{\vec{x}^{(k)}\}$ 收敛
- **收敛速度**
  - ✓ 特别是针对大规模的稀疏方程组

向量序列的收敛

定义3.1.1（教材编号）

在 $\mathbb{R}^n$ 中，有一个向量序列 $\{\vec{x}^{(0)},\vec{x}^{(1)},\vec{x}^{(2)},\ldots\}$，记为 $\{\vec{x}^{(k)}\}_{k=0}^{\infty}$，或简记为 $\{\vec{x}^{(k)}\}$

若 $\mathbb{R}^n$ 定义了范数 $\|\cdot\|$，且存在 $\vec{x}^* \in \mathbb{R}^n$，满足：

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

则称 $\{\vec{x}^{(k)}\}$ **收敛于** $\vec{x}^*$，简记为 $\lim_{k\to\infty}\vec{x}^{(k)} = \vec{x}^*$

向量序列的收敛

$\vec{x}^{(k)} = \left(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}\right)^T \quad \vec{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)^T$ 定义3.1.1中，选取∞-范数

$$\lim_{k\to\infty} \vec{x}^{(k)} = \vec{x}^* \Rightarrow \lim_{k\to\infty} \max_{1\le i\le n} \left|x_i^{(k)} - x_i^*\right| = 0$$

$$\Rightarrow \lim_{k\to\infty} \left|x_i^{(k)} - x_i^*\right| = 0, i = 1,2,\ldots,n$$

$$\Rightarrow \lim_{k\to\infty} x_i^{(k)} = x_i^*, i = 1,2,\ldots,n$$

注意

1.有限维线性空间的范数相互等价，因而定义中的范数可任意选择

2.$\left\{\vec{x}^{(k)}\right\}$的**每一个分量**都应该收敛

由于范数的等价性

向量的收敛问题 $\xLeftrightarrow{\text{等价于}}$ 每一个**分量的数列收敛**问题！

矩阵序列的收敛

定义3.1.2（教材编号）

在 $\mathbb{R}^{n\times n}$ 中，有一个矩阵序列，记为 $\{\mathbf{A}^{(k)}\}_{k=0}^{\infty}$ ，或简记为 $\{\mathbf{A}^{(k)}\}$

若 $\mathbb{R}^{n\times n}$ 定义了范数 $\|\cdot\|$，且存在 $\mathbf{A}^* \in \mathbb{R}^{n\times n}$ ，满足：

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

则称 $\{\mathbf{A}^{(k)}\}$ **收敛于** $\mathbf{A}^*$，简记为 $\lim\limits_{k\to\infty}\mathbf{A}^{(k)}=\mathbf{A}^*$

同样注意

1.范数可以任意选择

2.记 $\mathbf{A}^{(k)}=\left[a_{ij}^{(k)}\right]$ $\mathbf{A}^*=\left[a_{ij}^*\right]$ 则同样可得

$$\lim_{k\to\infty}\mathbf{A}^{(k)}=\mathbf{A}^* \Leftrightarrow \lim_{k\to\infty}a_{ij}^{(k)}=a_{ij}^*, i,j=1,2,\ldots,n$$

矩阵序列的收敛

举例

$$\mathbf{A}=\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix} \Rightarrow \mathbf{A}^2=\begin{bmatrix}\lambda^2 & 2\lambda\\ 0 & \lambda^2\end{bmatrix} \Rightarrow \cdots\cdots \Rightarrow \mathbf{A}^k=\begin{bmatrix}\lambda^k & k\lambda^{k-1}\\ 0 & \lambda^k\end{bmatrix}$$

若 $|\lambda|<1$，则 $\lim\limits_{k\to\infty} A^{(k)}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}=\mathbf{0}$

按照每一个元素来处理，对于小规模的矩阵可以“看出来”

对于较大规模的矩阵$\mathbf{B}$，如何得出关于$\mathbf{B}^k$的结论？

矩阵序列的收敛

定理3.1.1（教材编号）

设 $\lim\limits_{k\to\infty}\mathbf{A}^{(k)}=\mathbf{0}$ 的充分必要条件是：$\lim\limits_{k\to\infty}\mathbf{A}^{(k)}\vec{x}=\vec{0},\forall\vec{x}\in\mathbb{R}^n$

其中 $\mathbf{0}$ 代表0矩阵，$\vec{0}$ 代表0向量

充分性

矩阵从属范数的特性 $\forall\vec{x}\in\mathbb{R}^n\quad\|\mathbf{A}^{(k)}\vec{x}\|\le\|\mathbf{A}^{(k)}\|\|\vec{x}\|$，如果 $\lim\limits_{k\to\infty}\mathbf{A}^{(k)}=\mathbf{0}$ 成立，由定义3.1.2可知 $\lim\limits_{k\to\infty}\|\mathbf{A}^{(k)}-\mathbf{0}\|=\lim\limits_{k\to\infty}\|\mathbf{A}^{(k)}\|=0\Rightarrow\lim\limits_{k\to\infty}\|\mathbf{A}^{(k)}\vec{x}\|=0$

再由定义3.1.1可知 $\lim\limits_{k\to\infty}\mathbf{A}^{(k)}\vec{x}=\vec{0}$

必要性

第j个元素

取 $e_j=(0,\ldots,1,\ldots,0)^T$，则由 $\lim\limits_{k\to\infty}\mathbf{A}^{(k)}e_j=\vec{0}$ 可得矩阵**A**的第j列元素收敛于0

$j=1,2,\ldots,n$ 时，可得矩阵A的所有元素都收敛于0，则 $\Rightarrow\lim\limits_{k\to\infty}\mathbf{A}^{(k)}=\mathbf{0}$

矩阵序列的收敛

定理3.1.2（教材编号）

设 $\mathbf{B}\in\mathbb{R}^{n\times n}$，则下面三个命题等价：

(1) $\lim\limits_{k\to\infty}\mathbf{B}^k=\mathbf{0}$　　(2) $\rho(\mathbf{B})<1$　　(3) 至少存在一种从属范数满足 $\|B\|<1$

(1)→(2)：反证法！假定**B**有一个特征值 $\lambda\geq 1$，则存在 $\vec{x}\neq 0$ 使得

$$\mathbf{B}\vec{x}=\lambda\vec{x} \Rightarrow \|\mathbf{B}^k\vec{x}\|=|\lambda|^k\|\vec{x}\|$$

显然 $k\to\infty$时，$\{\mathbf{B}^k\vec{x}\}$ 不趋近于0向量，与(1)矛盾。得证！

(2)→(3)：根据上一章定理2.4.3 $\|\mathbf{B}\|\leq\rho(\mathbf{B})+\varepsilon,\varepsilon>0$，(2)成立时只要选择合适的 ε，即可得(3)

(3)→(1)：若存在一种从属范数满足 $\|\mathbf{B}\|<1$，由 $\|\mathbf{B}^k\|\leq\|\mathbf{B}\|^k$，可得，即可得(1) $\lim\limits_{k\to\infty}\|\mathbf{B}^k\|=0$

矩阵序列的收敛

定理3.1.3（教材编号）

设 $\mathbf{B}\in\mathbb{R}^{n\times n}$，$\|\cdot\|$ 为任意一种矩阵范数，则：

$$\lim_{k\to\infty}\left\|\mathbf{B}^k\right\|^{\frac{1}{k}}=\rho(\mathbf{B})$$

根据上一章定理2.4.3 $\rho(\mathbf{B})\le\|\mathbf{B}\|$ ⇨ $\boxed{\rho(\mathbf{B})=\rho\left(\mathbf{B}^k\right)^{\frac{1}{k}}\le\left\|\mathbf{B}^k\right\|^{\frac{1}{k}}}$

$\forall\varepsilon>0$ 记 $\mathbf{B}_\varepsilon=[\rho(\mathbf{B})+\varepsilon]^{-1}\mathbf{B}$，则 $\rho(\mathbf{B}_\varepsilon)=\dfrac{\rho(\mathbf{B})}{\rho(\mathbf{B})+\varepsilon}<1$ ⇨ $\lim\limits_{k\to\infty}\mathbf{B}_\varepsilon^k=\mathbf{0}$

当k足够大时 $\left\|\mathbf{B}_\varepsilon^k\right\|=\dfrac{\left\|\mathbf{B}^k\right\|}{[\rho(\mathbf{B})+\varepsilon]^k}<1$ ⇨ $\boxed{\left\|\mathbf{B}^k\right\|^{\frac{1}{k}}<\rho(\mathbf{B})+\varepsilon}$

构造迭代矩阵B

已如第二章所定义的有唯一解的线性方程组 $\mathbf{A}\vec{x}=\vec{b}$，其中

$$\mathbf{A}\in\mathbb{R}^{n\times n}\qquad \det|\mathbf{A}|\neq 0\qquad \vec{b}=\mathbb{R}^n$$

$$\mathbf{A}=\mathbf{M}-\mathbf{N}\ \Rightarrow\ (\mathbf{M}-\mathbf{N})\vec{x}=\vec{b}\ \Rightarrow\ \mathbf{M}\vec{x}=\mathbf{N}\vec{x}+\vec{b}\ \Rightarrow\ \vec{x}=\mathbf{M}^{-1}\mathbf{N}\vec{x}+\mathbf{M}^{-1}\vec{b}$$

$$\mathbf{B}=\mathbf{M}^{-1}\mathbf{N}=\mathbf{M}^{-1}(\mathbf{M}-\mathbf{A})=\mathbf{I}-\mathbf{M}^{-1}\mathbf{A}\qquad \vec{f}=\mathbf{M}^{-1}\vec{b}\qquad \vec{x}^{(k+1)}=\mathbf{B}\vec{x}^{(k)}+\vec{f}$$

定义3.1.3（教材编号）

若存在向量 $\vec{x}^*\in\mathbb{R}^n$，使得如上迭代法获得的向量序列 $\{\vec{x}^{(k)}\}$ 满足

$$\lim_{k\to\infty}\vec{x}^{(k)}=\vec{x}^*,\forall\vec{x}^{(0)}\in\mathbb{R}^n$$

则称如上迭代法是**收敛**的

收敛性分析

$\mathbf{A}\vec{x}^* = \vec{b}$ $\quad$ $\vec{x}^* = \mathbf{B}\vec{x}^* + \vec{f}$ $\quad$ $\vec{x}^{(k+1)} = \mathbf{B}\vec{x}^{(k)} + \vec{f}$

记每次迭代的误差 $\vec{e}^{(k)} = \vec{x}^{(k)} - \vec{x}^*$

我们已经知道： $\vec{e}^{(k)} = \mathbf{B}\vec{e}^{(k-1)} = \mathbf{B}^2\vec{e}^{(k-2)} = \ldots = \mathbf{B}^k\vec{e}^{(0)}$

定理3.1.4（教材编号）

迭代法 $\vec{x}^{(k+1)} = \mathbf{B}\vec{x}^{(k)} + \vec{f}$ 收敛的充分必要条件为 $\rho(\mathbf{B}) < 1$

如上迭代法收敛 ⇨ $\lim\limits_{k\to\infty} \vec{e}^{(k)} = \lim\limits_{k\to\infty} \mathbf{B}^k\,\vec{e}^{(0)} = \vec{0}, \forall \vec{e}^{(0)} \in \mathbb{R}^n$

⇨ $\lim\limits_{k\to\infty} \mathbf{B}^k = \mathbf{0}$

⇨ $\rho(\mathbf{B}) < 1$

ρ(B)往往比较难判别，怎么办？

用从属范数！

收敛性分析

定理3.1.5（教材编号）

设 $\vec{x}^*$ 是方程组 $\vec{x} = \mathbf{B}\vec{x} + \vec{f}$ 唯一解，$\|\cdot\|$ 是一种向量范数，其从属矩阵范数满足

$\|\mathbf{B}\| = q < 1$ 则上述迭代法收敛，且

$$\left\|\vec{e}^{(k)}\right\| = \left\|\vec{x}^{(k)} - \vec{x}^*\right\| \le \frac{q}{1-q}\boxed{\left\|\vec{x}^{(k)} - \vec{x}^{(k-1)}\right\|}$$

我们习惯使用的中止条件

$$\left\|\vec{e}^{(k)}\right\| = \left\|\vec{x}^{(k)} - \vec{x}^*\right\| \le \frac{q^k}{1-q}\boxed{\left\|\vec{x}^{(1)} - \vec{x}^{(0)}\right\|}$$

可以预先估计k步以后的误差

注意：若q非常接近1时，我们习惯的中止条件会怎么样？

收敛性分析

定理的证明

1、定理中的第二式到底有什么意义？
2、定理是充分必要条件吗？

$$\rho(\mathbf{B}) \le \|\mathbf{B}\| = q < 1 \Rightarrow \text{迭代法必定收敛} \lim_{k\to\infty} \vec{x}^{(k)} = \vec{x}^*$$

$$\vec{x}^{(k)} - \vec{x}^* = \mathbf{B}\left(\vec{x}^{(k-1)} - \vec{x}^*\right) = \mathbf{B}\left(\vec{x}^{(k-1)} - \vec{x}^{(k)}\right) + \mathbf{B}\left(\vec{x}^{(k)} - \vec{x}^*\right)$$

$$\Rightarrow \boxed{\|\vec{x}^{(k)} - \vec{x}^*\|} \le \|\mathbf{B}\left(\vec{x}^{(k-1)} - \vec{x}^{(k)}\right)\| + \|\mathbf{B}\left(\vec{x}^{(k)} - \vec{x}^*\right)\| \le q\|\vec{x}^{(k-1)} - \vec{x}^{(k)}\| + q\boxed{\|\vec{x}^{(k)} - \vec{x}^*\|}$$

$$\Rightarrow \|\vec{x}^{(k)} - \vec{x}^*\| \le \frac{q}{1-q}\|\vec{x}^{(k)} - \vec{x}^{(k-1)}\|$$

$$\boxed{\begin{aligned} \vec{x}^{(k)} &= \mathbf{B}\vec{x}^{(k-1)} + \vec{f} \\ \vec{x}^{(k-1)} &= \mathbf{B}\vec{x}^{(k-2)} + \vec{f} \end{aligned}}$$

再反复使用

$$\boxed{\|\vec{x}^{(k)} - \vec{x}^{(k-1)}\| = \|\mathbf{B}\left(\vec{x}^{(k-1)} - \vec{x}^{(k-2)}\right)\|} \le q\|\vec{x}^{(k-1)} - \vec{x}^{(k-2)}\| \Rightarrow \left\|\vec{x}^{(k)} - \vec{x}^*\right\| \le \frac{q^k}{1-q}\left\|\vec{x}^{(1)} - \vec{x}^{(0)}\right\|$$

收敛率

几步能收敛满意?
k与$\rho(\mathbf{B})$关系?

已知 $\left\|\vec{e}^{(k)}\right\| = \left\|\mathbf{B}^k \vec{e}^{(0)}\right\| \le \left\|\mathbf{B}^k\right\| \left\|\vec{e}^{(0)}\right\| \Rightarrow \dfrac{\left\|\vec{e}^{(k)}\right\|}{\left\|\vec{e}^{(0)}\right\|} \le \left\|\mathbf{B}^k\right\|$

回想一下从属范数的定义!

$$\left\|\mathbf{B}^k\right\| = \max_{\vec{e}^{(0)} \ne 0} \frac{\left\|\mathbf{B}^k \vec{e}^{(0)}\right\|}{\left\|\vec{e}^{(0)}\right\|} = \max_{\vec{e}^{(0)} \ne 0} \frac{\left\|\vec{e}^{(k)}\right\|}{\left\|\vec{e}^{(0)}\right\|}$$

这个可以看作迭代k次后误差比的最大值

平均到每次迭代的误差可以看作: $\left\|\mathbf{B}^k\right\|^{\frac{1}{k}}$

若希望 $\dfrac{\left\|\vec{e}^{(k)}\right\|}{\left\|\vec{e}^{(0)}\right\|} \le \boxed{\varepsilon = 10^{-s}} \ll 1 \Rightarrow \left\|\mathbf{B}^k\right\| \le \varepsilon \Rightarrow \left\|\mathbf{B}^k\right\|^{\frac{1}{k}} \le \varepsilon^{\frac{1}{k}}$

$$\Rightarrow \ln\left\|\mathbf{B}^k\right\|^{\frac{1}{k}} \le \frac{1}{k}\ln\varepsilon \Rightarrow k \ge \frac{-\ln\varepsilon}{-\ln\left\|\mathbf{B}^k\right\|^{\frac{1}{k}}}$$

定义3.1.4（教材编号）

以上迭代法的**平均收敛率**定义为

$$R_k(\mathbf{B}) = -\ln\left\|\mathbf{B}^k\right\|^{\frac{1}{k}}$$

依赖于从属范数和迭代次数！

定义3.1.5（教材编号）

$$R(\mathbf{B}) = \lim_{k\to\infty} R_k(\mathbf{B}) = \lim_{k\to\infty} -\ln\left\|\mathbf{B}^k\right\|^{\frac{1}{k}} = -\ln\rho(\mathbf{B})$$

定义为以上迭代法的**渐进收敛率**或**渐进收敛速度**

内容提纲

Jacobi迭代法

$$\mathbf{A}=\mathbf{M}-\mathbf{N} \Rightarrow \vec{x}=\mathbf{M}^{-1}\mathbf{N}\vec{x}+\mathbf{M}^{-1}\vec{b}$$

取 $\mathbf{A}=\mathbf{D}-\mathbf{L}-\mathbf{U}$

$\mathbf{D}=diag\left(a_{11},a_{22},\ldots,a_{nn}\right)$ 即以A的所有对角元构成的对角矩阵

$$\mathbf{L}=-\begin{bmatrix} 0 & & & \\ a_{21} & 0 & & \\ \vdots & & \ddots & \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{bmatrix} \qquad \mathbf{U}=-\begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & 0 & & \vdots \\ & & \ddots & a_{n-1,n} \\ & & & 0 \end{bmatrix}$$

注意和LU分解的区别

此处是直接将**A**的上、下三角元直接赋予一个上三角阵和一个下三角阵！

Jacobi迭代法

假定所有对角元素不为0，即**D**非奇异，取

$$\mathbf{M}=\mathbf{D} \qquad \mathbf{N}=\mathbf{L}+\mathbf{U} \qquad \mathbf{B}_J=\mathbf{I}-\mathbf{D}^{-1}\mathbf{A}=\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})$$

迭代方程为： $\vec{x}^{(k+1)}=\mathbf{B}_J\vec{x}^{(k)}+\vec{f}_J$ 其中 $\vec{f}_J=\mathbf{D}^{-1}\vec{b}$

进一步拆解每一个分量

$$\vec{x}^{(k+1)}=\mathbf{D}^{-1}\left[(\mathbf{L}+\mathbf{U})\vec{x}^{(k)}+\vec{b}\right]$$

第i行除了对角元以外所有元素

$$\Rightarrow \quad x_i^{(k+1)}=\frac{1}{a_{ii}}\left(b_i-\underbrace{\sum_{j=1}^{i-1}a_{ij}x_j^{(k)}}_{\text{下对角部分}}-\underbrace{\sum_{j=i+1}^{n}a_{ij}x_j^{(k)}}_{\text{上对角部分}}\right)$$

这称为**Jacobi迭代法**，简称**J法**

Jacobi迭代法—收敛性

定理3.2.1（教材编号）-1

J法收敛的充分必要条件是 $\rho(\mathbf{B}_J)<1$

J法收敛的充分条件是 $\|\mathbf{B}_J\|<1$

定理3.2.2（教材编号）-1

严格对角占优 $|a_{ii}|>\sum_{j=1,i\neq j}^{n}|a_{ij}|, i=1,2,\ldots,n$

若矩阵**A**为严格对角占优矩阵或为不可约弱对角占优矩阵，则A非奇异，且方程组的J法收敛

第二章定理2.3.1

定理3.2.3（教材编号）-1

若矩阵**A**对称，且对角元都大于0，则J法收敛的充分必要条件是**A**及2**D**-**A**均为正定矩阵，其中 $\mathbf{D}=diag(a_{11},a_{22},\ldots,a_{nn})$

Gauss-Seidel迭代法

从J法的分量形式开始 $x_i^{(k+1)} = \dfrac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right), i = 1,2,\ldots,n$

如果在k+1步，已经算完分量 $x_{i-1}^{(k+1)}$，则 $x_i^{(k+1)}$ 分量计算的过程中，之前的分量都用**新的**结果

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \boxed{\sum_{j=1}^{i-1}} a_{ij}\boxed{x_j^{(k+1)}} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right), i = 1,2,\ldots,n$$

第i行 $\begin{bmatrix} 0 & & & \\ a_{21} & 0 & & \\ \vdots & & \ddots & \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{bmatrix} \begin{bmatrix} \vdots \\ x_{i-1} \\ x_i \\ \vdots \end{bmatrix}$ 已经更新！

这称为**Gauss-Seidel迭代法**，简称**GS法**

Gauss-Seidel迭代法

GS法分量形式 $$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right), i = 1, 2, \ldots, n$$

写回矩阵的形式

$$\text{J法}: \vec{x}^{(k+1)} = \mathbf{D}^{-1}\left[(\mathbf{L}+\mathbf{U})\vec{x}^{(k)} + \vec{b}\right] = \mathbf{D}^{-1}\left(\mathbf{L}\vec{x}^{(k)} + \mathbf{U}\vec{x}^{(k)} + \vec{b}\right)$$

$$\text{GS法}: \vec{x}^{(k+1)} = \mathbf{D}^{-1}\left(\mathbf{L}\vec{x}^{(k+1)} + \mathbf{U}\vec{x}^{(k)} + \vec{b}\right)$$

$$\Rightarrow \quad (\mathbf{D}-\mathbf{L})\vec{x}^{(k+1)} = \mathbf{U}\vec{x}^{(k)} + \vec{b} \qquad \vec{x}^{(k+1)} = \mathbf{B}_G\vec{x}^{(k)} + \vec{f}_G$$

$$\Rightarrow \quad \mathbf{B}_G = (\mathbf{D}-\mathbf{L})^{-1}\mathbf{U} = \mathbf{I} - (\mathbf{D}-\mathbf{L})^{-1}\mathbf{A} \qquad \vec{f}_G = (\mathbf{D}-\mathbf{L})^{-1}\vec{b}$$

写成矩阵形式可以在矩阵编程环境中方便操作
但是更重要的是为了理论分析！

J法和GS法的收敛性

定理3.2.1（教材编号）-完整

J法收敛的充分必要条件是

$$\rho(\mathbf{B}_J)<1$$

GS法收敛的充分必要条件是

$$\rho(\mathbf{B}_{GS})<1$$

J法收敛的充分条件是

$$\|\mathbf{B}_J\|<1$$

GS法收敛的充分条件是

$$\|\mathbf{B}_J\|<1$$

一个关于对角占优阵的定理

定理3.2.2（教材编号）-完整

若矩阵**A**为严格对角占优矩阵或为不可约弱对角占优矩阵，
则A非奇异，且方程组的J法和GS法都收敛

第二章定理2.3.1

严格对角占优 $|a_{ii}| > \sum_{j=1, i\neq j}^{n} |a_{ij}|, i = 1, 2, \ldots, n$

反证法：如果**A**奇异！

则 $\mathbf{A}\vec{x} = \mathbf{0}$ 有非0解 $\vec{x}^*$ ⇨ $\|\vec{x}^*\|_\infty = \max_{1\le i\le n} |\vec{x}_i^*| = |x_k^*| \neq 0$ ⇨ 第k行方程 $\sum_{j=1}^{n} a_{kj} x_j^* = 0$

$$|a_{kk} x_k^*| = \left| -\sum_{j=1, j\neq k}^{n} a_{kj} x_j^* \right| \le \sum_{j=1, j\neq k}^{n} |a_{kj}| |x_j^*| \Rightarrow |a_{kk}| \le \sum_{j=1, j\neq k}^{n} |a_{kj}| \left| \frac{x_j^*}{x_k^*} \right| \le \sum_{j=1, j\neq k}^{n} |a_{kj}|$$

和严格对角占优矛盾！

一个关于对角占优阵的定理

证明J法严格对角占优的情况 $\mathbf{A}=\mathbf{D}-\mathbf{L}-\mathbf{U}$ $\mathbf{B}_J=\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})$

$$-\begin{bmatrix} 0 & \frac{a_{12}}{a_{11}} & \cdots & \frac{a_{1n}}{a_{11}} \\ \frac{a_{21}}{a_{22}} & 0 & & \vdots \\ \vdots & & \ddots & \frac{a_{n-1,n}}{a_{n-1,n-1}} \\ \frac{a_{n1}}{a_{nn}} & \cdots & \frac{a_{n,n-1}}{a_{nn}} & 0 \end{bmatrix} = \begin{cases} 0 & ,i=j \\ \boxed{-\frac{a_{ij}}{a_{ii}}} & ,i\neq j \end{cases} \quad i,j=1,2,\ldots,n$$

$$\Rightarrow \|\mathbf{B}_J\|_\infty=\max_{1\le i\le n}\sum_{j=1}^{n}|b_{ij}|=\max_{1\le i\le n}\sum_{j\neq i}\left|\frac{a_{ij}}{a_{ii}}\right|<1 \quad \Rightarrow \rho(\mathbf{B}_J)\le\|\mathbf{B}_J\|_\infty<1$$

J法收敛！

一个关于对角占优阵的定理

证明GS法严格对角占优的情况 $\mathbf{A}=\mathbf{D}-\mathbf{L}-\mathbf{U}$

$$\mathbf{C}=\left[c_{ij}\right]=\begin{bmatrix}\lambda a_{11} & a_{12} & \cdots & a_{1n}\\ \lambda a_{21} & \lambda a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn}\end{bmatrix}$$

$$\mathbf{B}_G=(\mathbf{D}-\mathbf{L})^{-1}\mathbf{U}=\mathbf{I}-(\mathbf{D}-\mathbf{L})^{-1}\mathbf{A}$$

记作C

$$\det(\lambda\mathbf{I}-\mathbf{B}_G)=\det\left((\mathbf{D}-\mathbf{L})^{-1}\right)\underline{\det(\lambda(\mathbf{D}-\mathbf{L})-\mathbf{U})}=0$$

反证法，若：$|\lambda|\geq 1$

$$|c_{ii}|=|\lambda a_{ii}|>|\lambda|\left(\sum_{j=1}^{i-1}|a_{ij}|+\sum_{j=i+1}^{n}|a_{ij}|\right)\geq\boxed{\sum_{j=1}^{i-1}|\lambda a_{ij}|}+\boxed{\sum_{j=i+1}^{n}|a_{ij}|}=\sum_{\substack{j=1\\ j\neq i}}^{n}|c_{ij}|$$

此时，C也同样是严格对角占优，则det(C)无法等于0

因而λ必须小于1；得证！

一个关于对角占优阵的定理

证明GS法弱对角占优不可约的情况

$$\mathbf{C}' = \begin{bmatrix} a_{12} & \lambda^{-1}a_{12} & \cdots & \lambda^{-1}a_{1n} \\ a_{21} & a_{22} & \cdots & \lambda^{-1}a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$\det\left(\lambda\mathbf{I}-\mathbf{B}_G\right)=\det\left(\left(\mathbf{D}-\mathbf{L}\right)^{-1}\right)\underline{\det\left(\lambda\left(\mathbf{D}-\mathbf{L}\right)-\mathbf{U}\right)}$$

$$\det\left(\lambda\left(\mathbf{D}-\mathbf{L}\right)-\mathbf{U}\right)=\lambda^n\det\left(\mathbf{D}-\mathbf{L}-\lambda^{-1}\mathbf{U}\right)$$

$$\mathbf{A}=\mathbf{D}-\mathbf{L}-\mathbf{U} \qquad \mathbf{C}'=\mathbf{D}-\mathbf{L}-\lambda^{-1}\mathbf{U}$$

A不可约，则**C**’也不可约，同样反证：若 $|\lambda|\geq 1$，则**C**’将保持弱对角占优，det(**C**’)同样无法等于0，因而λ**必须**小于1；得证！

J法弱对角占优不可约略！

一个关于对称正定矩阵的定理

定理3.2.3（教材编号）-完整

若矩阵**A**对称，且对角元都大于0，则

（1）如上述解线性方程组的J法收敛的**充分必要条件**是**A**及**2D-A**均为正定矩阵，其中 $\mathbf{D} = diag\left(a_{11}, a_{22}, \ldots, a_{nn}\right)$

（2）如上述解线性方程组的GS法收敛的**充分条件**是**A**正定

SOR收敛性中的一部分！

编程实践

✓ 通过Jacobi迭代求解线性方程组

- 在LinearEquations中添加网络学堂下载的JacobiIteration类文件
- 将Program中的Main函数改成如下内容

```
public class Program
{
    public static void Main(string[] args)
    {
        JacobiIteration.Sample();

        Console.ReadKey();
    }
}
```

- 运行发现程序抛出异常，请补全JacobiIteration.Solve方法

✓ JacobiIteration类解读（Sample）

一个用于测试的方法，与我们前面写过的代码类似

```
public static void Sample()
{
    Matrix A = new Matrix(new double[,] {
        { 6, 2, 1, -1 },
        { 2, 4, 1, 0 },
        { 1, 1, 4, -1 },
        { -1, 0, -1, 3 }
    });

    Vector b = new Vector(6, 1, 5, -5);

    Vector x = Solve(A, b);

    Console.WriteLine("x为\n" + x);
    Console.WriteLine("偏差Ax-b为");
    Console.WriteLine(A * x - b);
}
```

系数矩阵

常数项

计算解向量，并求偏差Ax-b

✓ JacobiIteration类解读（Solve）

```
public static Vector Solve(Matrix A, Vector b, int maxCount = 1000)
{
    //此处需构造迭代所需的BJ和fJ

    Vector x = new Vector(b.Length);

    int count = 0;
    while (Norm.One(A * x - b) > 1e-10)
    {
        if (count >= maxCount)
            throw new Exception("未在指定迭代次数内收敛，Jacobi方法可能不收敛");

        //此处需进行Jacobi迭代

        count++;
    }
    return x;
}
```

初始化迭代向量x为零向量

增加了最大迭代数，避免不收敛时陷入死循环

计算Ax-b的1-范数，以1e-10作为误差限

记录循环次数

✓ JacobiIteration类缺失代码参考（Solve）

```
public static Vector Solve(Matrix A, Vector b, int maxCount = 1000)
{
    // ……
    Matrix DInv = new Matrix(A.RowCount, A.ColumnCount);

    for (int i = 0; i < A.RowCount; i++)
    {
        DInv[i, i] = 1 / A[i, i];
    }

    Matrix BJ = Matrix.Identity(A.RowCount)- DInv * A;

    Vector fJ = DInv * b;
    // ……
}
```

$$A = D - L - U$$

对角矩阵的逆

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

$$f_J = D^{-1}b$$

循环体中填：x = BJ * x + fJ;

$$x = B_J x + f_J$$

编程实践

✓ 通过Gauss-Seidel迭代求解线性方程组

- 在LinearEquations文件夹下新建一个GaussSeidelIteration类
- 记得①去掉命名空间后的.LinearEquations；②添加public修饰符
- 直接拷贝JacobiIteration的Solve和Sample方法过来（两者相似）
- 将Program的Main方法改为

```
public class Program
{
    public static void Main(string[] args)
    {
        JacobiIteration.Sample();
        GaussSeidelIteration.Sample();

        Console.ReadKey();
    }
}
```

编程实践

✓ 修改Gauss-Seidel类的Solve方法

思路1（直观）
计算B_G和f_G，替换掉B_J与f_J即可

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

$$f_G = (D-L)^{-1}b$$

需计算新的B_G和f_G，涉及求三角矩阵的逆

只要改这一句就可以！

循环体中填：
x = BJ * x + fJ;

思路2（简单？）
按照定义，每次采用"最新"算出的计算值

x的每个分量需从前往后计算，计算后一分量时采用最新的分量值

```
for (int i = 0; i < x.Length; i++)
{
    x[i] = BJ.GetRow(i) * x + f[i];
}
```

计算x[i]时，x[i-1]、x[i-2]、…、x[0]已经被更改为最新值√

✓ 两种迭代法对比

➢ 在两个类的Solve方法中添加一句代码（return之前），输出迭代数

```
// 迭代不断进行直到Ax-b接近于0
while (Norm.One(A * x - b) > 1e-10)...

Console.WriteLine("Jacobi方法迭代次数: " + count);

return x;
```

```
// 迭代不断进行直到Ax-b接近于0
while (Norm.One(A * x - b) > 1e-10)...

Console.WriteLine("Gauss-Seidel方法迭代次数: " + count);

return x;
```

E:\我的编程\VS_workspace\NumNet\NumNet\bin\Debug\NumNet.exe

```
Jacobi方法迭代次数: 75
x为
[0.790575916233429, -0.361256544499281, 0.863874345553061, -1.11518324607631]
偏差Ax-b为
[3.1377567211166E-11, 2.27962093646283E-11, 2.26982876938564E-11, -1.54063428681184E-11]
Gauss-Seidel方法迭代次数: 20
x为
[0.790575916235836, -0.361256544507004, 0.863874345551619, -1.11518324607085]
偏差Ax-b为
[2.34763319895137E-11, -4.72555328201452E-12, 6.15685280536127E-12, 0]
```

学习重点

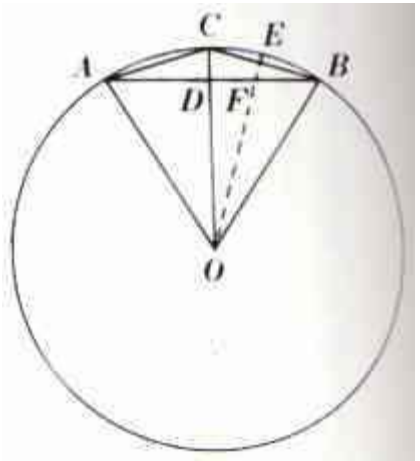
- 掌握迭代法的**收敛性**分析和**收敛率**概念
- 理解**雅克比迭代法**
- 掌握**高斯—赛德尔迭代法**
- 理解J法和GS法求解相关问题的**编程**

作业

教材P89-1、P89-2(1)、 P90-4

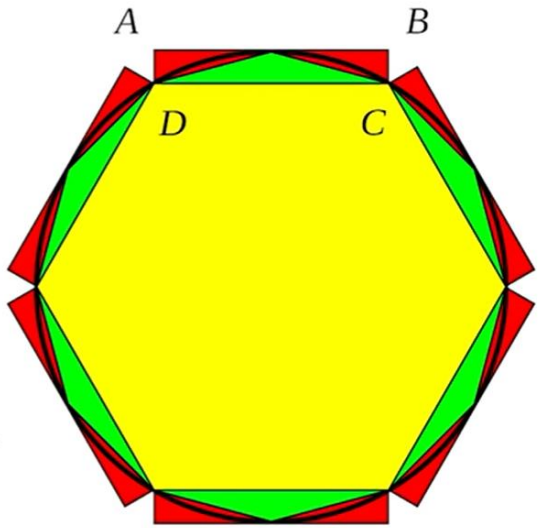
内容提纲

松弛的概念—从割圆说起



$$S = \pi r^2$$

估一个近似的面积 $\pi \approx \dfrac{S}{r^2}$



若取$r=10$　96边形 $S_{96} = 313 + \dfrac{584}{625}$　192边形 $S_{192} = 314 + \dfrac{64}{625}$

再往下手算太难了，怎么办？　$\tilde{S} = (1+\omega)S_{192} - \omega S_{96}$ 变成了找ω的问题

我们无法想象，刘徽当年是如何找到：$\omega = \dfrac{36}{105}$　$\tilde{S} = 314.16 \approx S_{3072}$

松弛的概念—再比较下J法和GS法

$$\mathbf{A}=\mathbf{M}-\mathbf{N} \Rightarrow (\mathbf{M}-\mathbf{N})\vec{x}=\vec{b} \Rightarrow \mathbf{M}\vec{x}=\mathbf{N}\vec{x}+\vec{b} \Rightarrow \vec{x}=\mathbf{M}^{-1}\mathbf{N}\vec{x}+\mathbf{M}^{-1}\vec{b}$$

J法： $\mathbf{M}=\boxed{\mathbf{D}}$ $\mathbf{N}=\boxed{\mathbf{L}+\mathbf{U}}$ $\mathbf{B}_J=\mathbf{I}-\mathbf{D}^{-1}\mathbf{A}=\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})$

GS法： $\mathbf{M}=\boxed{\mathbf{D}-\mathbf{L}}$ $\mathbf{N}=\boxed{\mathbf{U}}$ $\mathbf{B}_G=(\mathbf{D}-\mathbf{L})^{-1}\mathbf{U}=\mathbf{I}-(\mathbf{D}-\mathbf{L})^{-1}\mathbf{A}$

$$x_i^{(k+1)}=\frac{1}{a_{ii}}\left(b_i-\sum_{j=1}^{i-1}a_{ij}\boxed{x_j^{(k)}}-\sum_{j=i+1}^{n}a_{ij}x_j^{(k)}\right) \quad \text{J法}$$

$$x_i^{(k+1)}=\frac{1}{a_{ii}}\left(b_i-\sum_{j=1}^{i-1}a_{ij}\boxed{x_j^{(k+1)}}-\sum_{j=i+1}^{n}a_{ij}x_j^{(k)}\right),i=1,2,\ldots,n \quad \text{GS法}$$

从之前的实验可以得出，似乎GS法收敛较好的结论；从这里对比也可以看到：GS法在迭代的过程中，尽量使用更新过的值！

从GS法出发，如果要进一步加快迭代，如何?

松弛的概念—直接给出定义

令松弛因子$\omega>0$，利用松弛的基本思路 $\boxed{\vec{x}^{(k+1)}}=\omega\boxed{\tilde{x}^{(k+1)}}+(1-\omega)\vec{x}^{(k)}$

松弛后的结果　　GS法k+1的结果

$$\Rightarrow\ x_i^{(k+1)}=\omega\frac{1}{a_{ii}}\left(b_i-\sum_{j=1}^{i-1}a_{ij}x_j^{(k+1)}-\sum_{j=i+1}^{n}a_{ij}x_j^{(k)}\right)+(1-\omega)x_i^{(k)},i=1,2,\ldots,n$$

这一方法称为逐次松弛(Successive over-Relaxation, SOR) 法

- ✓ 显然$\omega=1$，SOR法退化为GS法
- ✓ 当 $\omega>1$时，称为**超松弛**方法；反之称为低松弛方法
- ✓ **按元素**编程实现时，不需要任何新的技巧，在GS法的基础上加一步加权平均即可

松弛的概念—直接给出定义

同样的道理，为了研究其性能，我们需要写出矩阵的形式

$$x_i^{(k+1)} = \omega \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right) + (1-\omega) x_i^{(k)}, i = 1, 2, \ldots, n$$

$$\boxed{\vec{x}^{(k+1)}} = \omega \mathbf{D}^{-1} \left(\vec{b} + \mathbf{L}\boxed{\vec{x}^{(k+1)}} + \boxed{\mathbf{U}\vec{x}^{(k)}} \right) + (1-\omega)\boxed{\vec{x}^{(k)}}$$

$$\Rightarrow (\mathbf{D} - \omega \mathbf{L}) \vec{x}^{(k+1)} = \omega \vec{b} + \left[\omega \mathbf{U} + (1-\omega) \mathbf{D} \right] \vec{x}^{(k)}$$

$$\Rightarrow \vec{x}^{(k+1)} = (\mathbf{D} - \omega \mathbf{L})^{-1} \left[\omega \mathbf{U} + (1-\omega) \mathbf{D} \right] \vec{x}^{(k)} + \underline{\omega (\mathbf{D} - \omega \mathbf{L})^{-1}} \vec{b}$$

$$\mathbf{M} = \frac{1}{\omega} (\mathbf{D} - \omega \mathbf{L})$$

$$\mathbf{N} = \frac{1}{\omega} \left[\omega \mathbf{U} + (1-\omega) \mathbf{D} \right]$$

松弛的概念—直接给出定义

$$\vec{x}^{(k+1)} = (\mathbf{D} - \omega\mathbf{L})^{-1}[\omega\mathbf{U} + (1-\omega)\mathbf{D}]\vec{x}^{(k)} + \omega(\mathbf{D} - \omega\mathbf{L})^{-1}\vec{b}$$

$$\Rightarrow \quad \vec{x}^{(k+1)} = \mathbf{L}_\omega \vec{x}^{(k)} + \vec{f} \qquad L_\omega = (\mathbf{D} - \omega\mathbf{L})^{-1}[\omega\mathbf{U} + (1-\omega)\mathbf{D}]$$

我们只需要讨论$\mathbf{L}_\omega$的性能!

显然，SOR法收敛的充分必要条件是：$\rho(\mathbf{L}_\omega) < 1$

一个充分条件是：$\|\mathbf{L}_\omega\| < 1$

所以SOR收敛性能的问题转化为：松弛因子如何影响$\mathbf{L}_\omega$谱半径和范数?

SOR法的收敛性

定理3.3.1（教材编号）

设 $\mathbf{A}\in\mathbb{R}^{n\times n}$，其对角元 $a_{ii}\neq 0, i=1,2,\ldots,n$，则对所有的实数 ω 有

$$\rho(\mathbf{L}_\omega)\geq|\omega-1|$$

证明： $L_\omega=(\mathbf{D}-\omega\mathbf{L})^{-1}\left[\omega\mathbf{U}+(1-\omega)\mathbf{D}\right]$

若 $\mathbf{L}_\omega$ 有 n 个特征值：$\lambda_1,\lambda_2,\ldots,\lambda_n$，则：$\lambda_1,\lambda_2,\ldots,\lambda_n$

$$\det(\mathbf{L}_\omega)=\det\left((\mathbf{D}-\omega\mathbf{L})^{-1}\left[\omega\mathbf{U}+(1-\omega)\mathbf{D}\right]\right)=\det\Big(\underbrace{(\mathbf{D}-\omega\mathbf{L})^{-1}}_{\text{下三角阵}}\Big)\det\Big(\underbrace{\omega\mathbf{U}+(1-\omega)\mathbf{D}}_{\text{上三角阵}}\Big)$$

$$=\det\left(\mathbf{D}^{-1}\right)\det\left((1-\omega)\mathbf{D}\right)=\det\left((1-\omega)\mathbf{I}\right)=(1-\omega)^n$$

$$\rho(\mathbf{L}_\omega)=\max_{1\leq i\leq n}|\lambda_i|\geq|\lambda_1\lambda_2\ldots\lambda_n|^{\frac{1}{n}}=|1-\omega|$$ 得证！

所以，若要SOR法收敛：$|\omega-1|<1$，即：$0<\omega<2$ **必要条件！**

SOR法的收敛性

定理3.3.2（教材编号）

设 $\mathbf{A} \in \mathbb{R}^{n\times n}$为对称正定矩阵，且 $0 < \omega < 2$ ，则SOR法收敛

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

证明：假定λ是$\mathbf{L}_\omega$的一个特征值，对应的特征向量 $\vec{x} \neq 0$

$$(\mathbf{D} - \omega\mathbf{L})^{-1}[\omega\mathbf{U} + (1-\omega)\mathbf{D}]\vec{x} = \lambda\vec{x} \Rightarrow [\omega\mathbf{U} + (1-\omega)\mathbf{D}]\vec{x} = \lambda(\mathbf{D} - \omega\mathbf{L})\vec{x}$$

$$\left(\left[\omega\mathbf{U} + (1-\omega)\mathbf{D}\right]\vec{x}, \vec{x}\right) = \left(\lambda(\mathbf{D} - \omega\mathbf{L})\vec{x}, \vec{x}\right)$$

$$\Rightarrow \omega\boxed{(\mathbf{U}\vec{x}, \vec{x})} + (1-\omega)(\mathbf{D}\vec{x}, \vec{x}) = \lambda\left[(\mathbf{D}\vec{x}, \vec{x}) - \omega\boxed{(\mathbf{L}\vec{x}, \vec{x})}\right] \qquad \vec{x} \neq 0$$

>0，记作p

SOR法的收敛性

证明（续）

$$\omega(\mathbf{U}\vec{x},\vec{x})+(1-\omega)(\mathbf{D}\vec{x},\vec{x})=\lambda\left[(\mathbf{D}\vec{x},\vec{x})-\omega(\mathbf{L}\vec{x},\vec{x})\right]$$

$(\mathbf{U}\vec{x},\vec{x})=(\mathbf{L}^T\vec{x},\vec{x})$ $\vec{x}$ 是一个复向量，记：$(\mathbf{L}\vec{x},\vec{x})=\alpha+i\beta$

$$(\mathbf{U}\vec{x},\vec{x})=(\mathbf{L}^T\vec{x},\vec{x})=(\vec{x},\mathbf{L}\vec{x})=\overline{(\mathbf{L}\vec{x},\vec{x})}=\alpha-i\beta$$

以上证明包含了定理3.2.3的（2）：$\omega=1$

实际为充要条件

$$\lambda=\frac{(1-\omega)p+\omega\alpha-i\omega\beta}{p-\omega\alpha-i\omega\beta},\ |\lambda|^2=\frac{\left[p-\omega(p-\alpha)\right]^2+\omega^2\beta^2}{(p-\omega\alpha)^2+\omega^2\beta^2}$$

分子减分母

$$\left[p-\omega(p-\alpha)\right]^2-(p-\omega\alpha)^2=(2\omega\alpha-\omega p)(2p-\omega p)=\omega p\boxed{(2\alpha-p)}(2-\omega)<0$$

$$(\mathbf{A}\vec{x},\vec{x})=(\mathbf{D}\vec{x},\vec{x})-(\mathbf{L}\vec{x},\vec{x})-(\mathbf{U}\vec{x},\vec{x})=p-2\alpha>0$$

$\Rightarrow$ $|\lambda|^2<1$ 得证！

最优松弛因子

回想一下收敛速度　　什么是最好的松弛因子？

$$R(\mathbf{B})=\lim_{k\to\infty}R_k(\mathbf{B})=\lim_{k\to\infty}-\ln\left\|\mathbf{B}^k\right\|^{\frac{1}{k}}=-\ln\rho(\mathbf{B})\quad k\geq\frac{-\ln\varepsilon}{R(\mathbf{B})}=\frac{s\ln 10}{R(\mathbf{B})}$$

谱半径取得最小！

但是，对任意矩阵，这一讨论是十分困难的，我们仅仅能针对一些特殊的矩阵

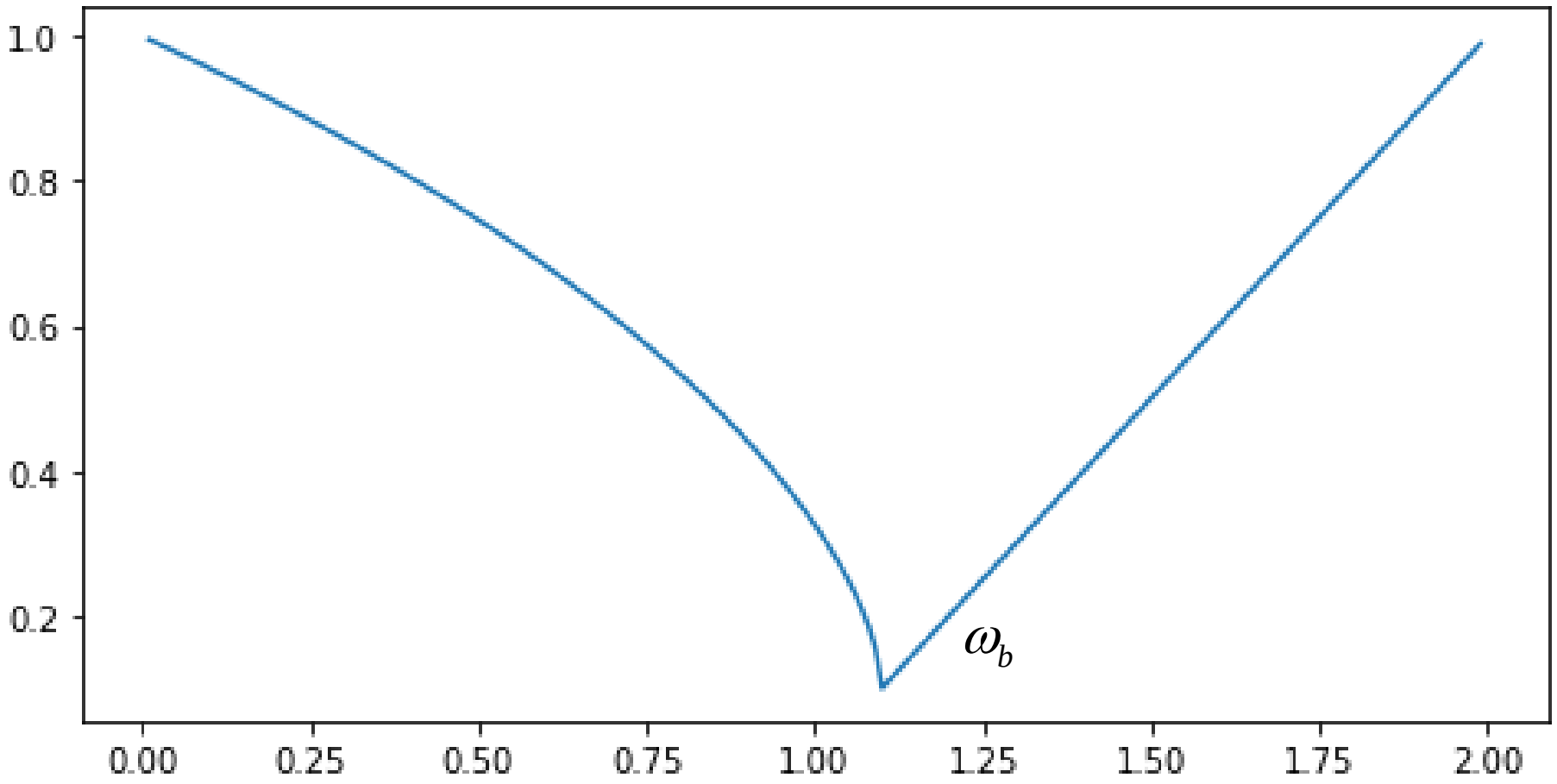
定理3.3.3（教材编号）

若方程组中的**A**是**对称正定三对角**矩阵，则

$$\rho(\mathbf{B}_G)=\left[\rho(\mathbf{B}_J)\right]^2<1\qquad \omega_b=\frac{2}{1+\sqrt{1-\left[\rho(\mathbf{B}_J)\right]^2}}\qquad \rho(\mathbf{L}_{\omega_b})=\omega_b-1$$

ω_b 为最优松弛因子！

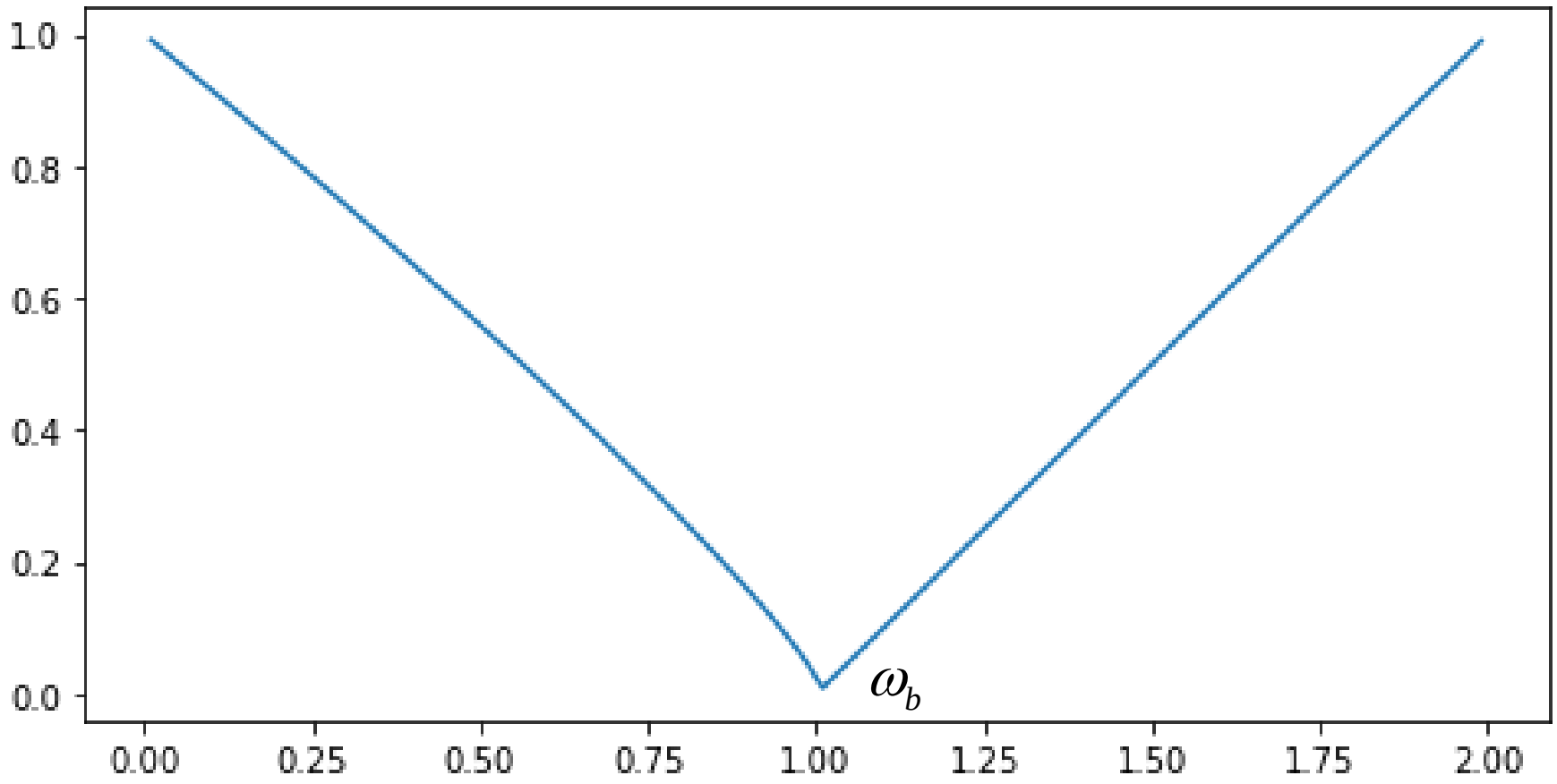
最优松弛因子



一般情况下，$\rho(\mathbf{L}_\omega)$ 在左侧变化剧烈，而在右侧成线性变化

最优松弛因子

将上题的矩阵对角线加大一些

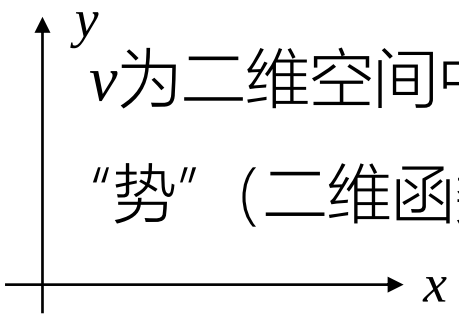


最优松弛因子退化至GS法，即 $\omega_b \approx 1$

块迭代*

泊松方程

$$\frac{\partial^2 v(x,y)}{\partial x^2}+\frac{\partial^2 v(x,y)}{\partial y^2}=\rho(x,y)$$



v为二维空间中待求的一个“势”（二维函数）

计算空间任何一个离散点v_{ij}，需要其邻域离散点

$v_{i-1,j-1}$	$v_{i-1,j}$	$v_{i-1,j+1}$
$v_{i,j-1}$	$v_{i,j}$	$v_{i,j+1}$
$v_{i+1,j-1}$	$v_{i+1,j}$	$v_{i+1,j+1}$

在点i,j附近进行泰勒展开，消去高阶项

$$\frac{\partial^2 v(x,y)}{\partial x^2}\approx\frac{v_{i,j+1}-2v_{i,j}+v_{i,j-1}}{h^2}$$

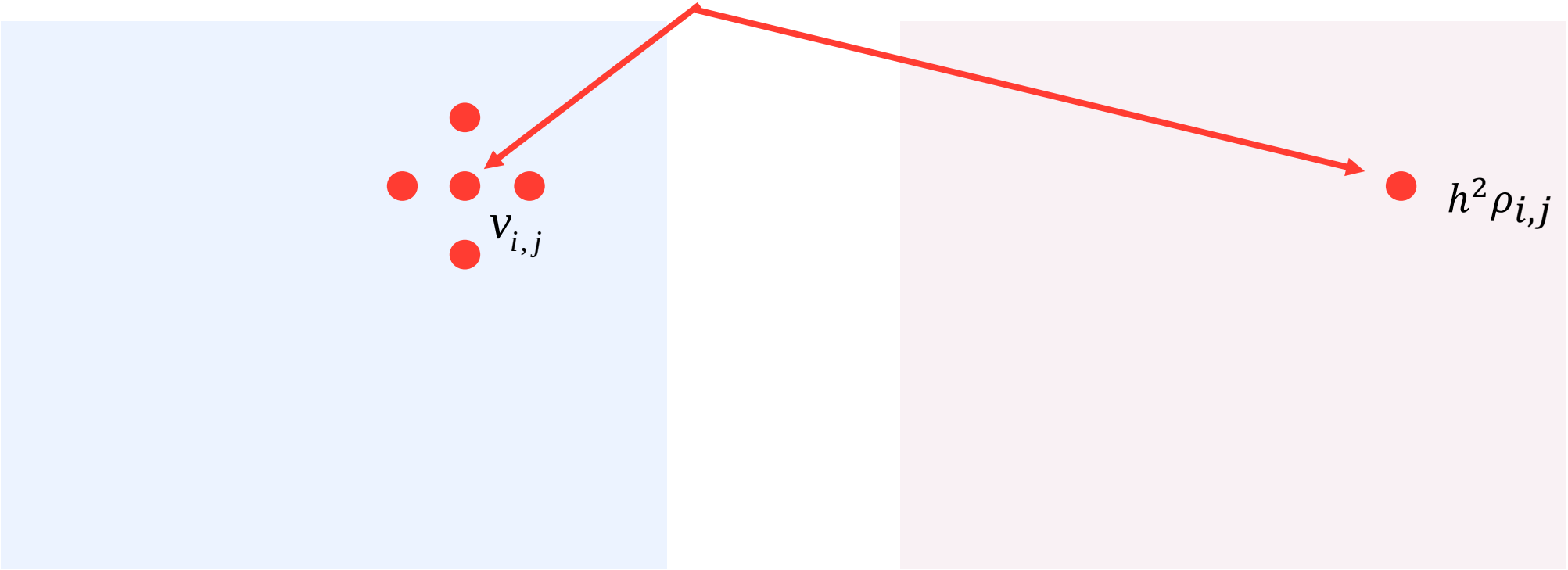
$$\frac{v_{i+1,j}-2v_{i,j}+v_{i-1,j}}{h^2}+\frac{v_{i,j+1}-2v_{i,j}+v_{i,j-1}}{h^2}=\rho_{i,j}$$

泊松方程的五点差分格式 ⇨

$$v_{i+1,j}+v_{i-1,j}+v_{i,j+1}+v_{i,j-1}-4v_{i,j}=h^2\rho_{i,j}$$

块迭代*

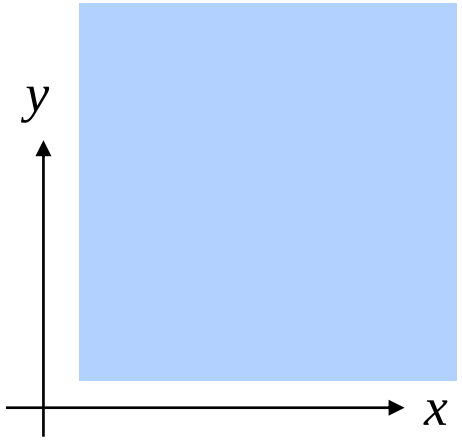
任意一个离散格点 $v_{i,j}$ 都满足上一页的差分方程



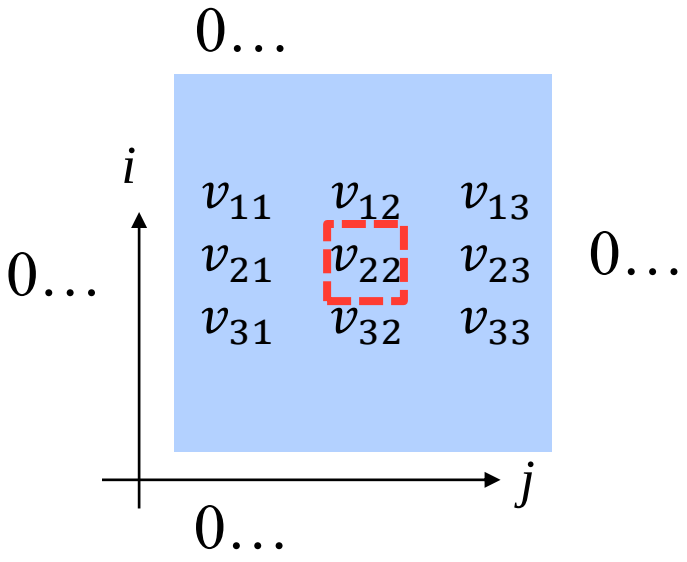
待求解的 $v(x,y)$ ⇐ 已知的密度分布 $\rho(x,y)$

已知的浓度分布 $v(x,y)$ ⇒ 待求解的粒子运动 $\rho(x,y)$

块迭代*



假设离散成3*3网格



$$\frac{v_{i+1,j}-2v_{i,j}+v_{i-1,j}}{\cancel{h^2}}+\frac{v_{i,j+1}-2v_{i,j}+v_{i,j-1}}{\cancel{h^2}}=\rho_{i,j}$$

$$\begin{matrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{matrix}$$

块迭代*

泊松方程组 $\mathbf{A}\vec{v} = \vec{\rho}$

$$\begin{array}{ccc} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{array}$$

$$\begin{array}{ccc} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{array}$$

按行展开

$$\left[\begin{array}{ccc|ccc|ccc}
-4 & 1 & \square & 1 & \square & \square & \square & \square & \square \\
1 & -4 & 1 & & 1 & & & & \\
\square & 1 & -4 & 0 & & 1 & & & \\
\hline
1 & & 0 & -4 & 1 & & 1 & & \\
\square & 1 & & 1 & -4 & 1 & & 1 & \\
\square & & 1 & & 1 & -4 & 0 & & 1 \\
\hline
\square & & & 1 & & 0 & -4 & 1 & \\
\square & & & & 1 & & 1 & -4 & 1 \\
\square & & & & & 1 & & 1 & -4
\end{array}\right]
\begin{bmatrix} v_{11} \\ v_{12} \\ v_{13} \\ v_{21} \\ v_{22} \\ v_{23} \\ v_{31} \\ v_{32} \\ v_{33} \end{bmatrix}
=
\begin{bmatrix} \rho \\ \\ \vdots \\ \\ \rho \end{bmatrix}$$

块迭代*

分块矩阵

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{A}_{22} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \mathbf{A}_{33} \end{bmatrix}$$

$n*n$的离散网格

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{I} & & & \\ \mathbf{I} & \mathbf{A}_{22} & \mathbf{I} & & \\ & \mathbf{I} & \ddots & \ddots & \\ & & \ddots & \ddots & \mathbf{I} \\ & & & \mathbf{I} & \mathbf{A}_{nn} \end{bmatrix}$$

$$\mathbf{A}_{ii} = \begin{bmatrix} -4 & 1 & & & \\ 1 & -4 & 1 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & -4 \end{bmatrix}_{n\times n} \qquad \mathbf{I} = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}_{n\times n}$$

块迭代*

不失一般性，如果有分块矩阵构成的方程组

$$\mathbf{A}=\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1n}\\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ \mathbf{A}_{n1} & \mathbf{A}_{n2} & \cdots & \mathbf{A}_{nn}\end{bmatrix}$$

可以按照迭代法做相同的拆解

$$\mathbf{D}=\begin{bmatrix}\mathbf{A}_{11} & & & \\ & \mathbf{A}_{22} & & \\ & & \ddots & \\ & & & \mathbf{A}_{nn}\end{bmatrix}\quad \mathbf{L}=\begin{bmatrix}\mathbf{0} & & & \\ -\mathbf{A}_{21} & \mathbf{0} & & \\ \vdots & \vdots & \ddots & \\ -\mathbf{A}_{n1} & -\mathbf{A}_{n2} & \cdots & \mathbf{0}\end{bmatrix}\quad \mathbf{U}=\begin{bmatrix}\mathbf{0} & -\mathbf{A}_{12} & \cdots & -\mathbf{A}_{1n}\\ & \mathbf{0} & \cdots & -\mathbf{A}_{2n}\\ & & \ddots & \vdots\\ & & & \mathbf{0}\end{bmatrix}$$

构造块SOR法（BSOR），形式完全类似

$$\vec{x}^{(k+1)}=(\mathbf{D}-\omega\mathbf{L})^{-1}[\omega\mathbf{U}+(1-\omega)\mathbf{D}]\vec{x}^{(k)}+\omega(\mathbf{D}-\omega\mathbf{L})^{-1}\vec{b}$$

块迭代*

1. 当**A**为对称正定时，同样取 $0<\omega<2$， BSOR法收敛
2. 当A为类似于泊松方程差分得到矩阵时，称为T-矩阵
   若D非奇异，且 $\mathbf{J}=\mathbf{I}-\mathbf{D}^{-1}\mathbf{A}$

   同样有

$$\omega_b=\frac{2}{1+\sqrt{1-[\rho(\mathbf{J})]^2}} \qquad \rho\left(\mathbf{L}_{\omega_b}\right)=\omega_b-1$$

   且

$$\rho(\mathbf{L}_\omega)=\begin{cases}\frac{1}{4}\left[\omega\rho(\mathbf{J})+\sqrt{\omega^2\rho(\mathbf{J})^2-4(\omega-1)}\right]^2, & 0<\omega<\omega_b\\ \omega-1, & \omega_b\le\omega<2\end{cases}$$

内容提纲

线性方程组等价的变分问题

已如第二章所定义的有唯一解的线性方程组 $\mathbf{A}\vec{x}=\vec{b}$， 其中

$\mathbf{A}\in\mathbb{R}^{n\times n}$ 且对称正定 $\vec{b}=\mathbb{R}^n$

建立一个关于$\vec{x}$的二次函数（泛函）

$$\varphi(\vec{x})=\varphi(x_1,x_2,\ldots,x_n)=\frac{1}{2}(\mathbf{A}\vec{x},\vec{x})-\left(\vec{b},\vec{x}\right)=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j-\sum_{j=1}^{n}b_j\,x_j$$

显然

$$\nabla\varphi(\vec{x})=\left[\frac{\partial\varphi}{\partial x_i}\right]=\mathbf{A}\vec{x}-\vec{b}$$

线性方程组在此时等价于一个泛函的最值问题

$\nabla\varphi(\vec{x})=0$ 最大还是最小?

线性方程组等价的变分问题

定理3.4.1（教材编号）

如果**A**是**对称正定矩阵**，则方程组的解 $\vec{x}^*$ 满足

$$\varphi(\vec{x}^*) = \min_{\vec{x}\in\mathbb{R}^n} \varphi(\vec{x})$$

最小值问题！

不加证明的给出

$$\varphi(\vec{x})-\varphi(\vec{x}^*)=\frac{1}{2}\left(\mathbf{A}\left(\vec{x}-\vec{x}^*\right),\vec{x}-\vec{x}^*\right)\geq 0$$ 对一切 $\vec{x}\in\mathbb{R}^n$ 均满足以上条件

若有任意其他一个 $\tilde{x}$ 使等号成立，则

$$(\mathbf{A}(\tilde{x}-\vec{x}^*),\tilde{x}-\vec{x}^*) = 0$$

再由**A**的正定性可得：$\tilde{x}=\vec{x}^*$

最速下降法

$$\varphi(\vec{x}^*) = \min_{\vec{x}\in\mathbb{R}^n} \varphi(\vec{x})$$

取初始值

$$\vec{x}^{(0)}$$

第一次迭代的残余向量

$$\vec{r}^{(0)} = \vec{b} - \mathbf{A}\vec{x}^{(0)}$$

每次迭代的残余向量即是梯度下降方向

$$-\nabla\varphi\left(\vec{x}^{(0)}\right) = \vec{b} - \mathbf{A}\vec{x}^{(0)} = \vec{r}^{(0)}$$

沿该方向搜索

$$\varphi\left(\vec{x}^{(1)}\right) = \varphi\left(\vec{x}^{(0)} + \alpha^{(0)}\vec{r}^{(0)}\right)$$

α为待定的参数

依次迭代 …

具体步骤（不加推导和证明）

取初始值 $\vec{x}^{(0)} \in \mathbb{R}^n$

对 $k = 0, 1, \ldots$

$$\vec{r}^{(k)} = \vec{b} - \mathbf{A}\vec{x}^{(k)} \Rightarrow \alpha^{(k)} = \frac{\left(\vec{r}^{(k)}, \vec{r}^{(k)}\right)}{\left(\mathbf{A}\vec{r}^{(k)}, \vec{r}^{(k)}\right)} \Rightarrow \vec{x}^{(k+1)} = \vec{x}^{(k)} + \alpha^{(k)}\vec{r}^{(k)}$$

说明

1. 所有的 $\vec{r}^{(k)}$ 互相正交
2. $\left\{\varphi\left(\vec{x}^{(k)}\right)\right\}$ 序列单调下降，且有下界满足 $\lim_{k\to\infty} \vec{x}^{(k)} = \vec{x}^*$

缺点：接近解的地方很慢！

共轭梯度法

定义3.4.1（教材编号）

$\mathbf{A} \in \mathbb{R}^{n\times n}$且对称正定，若一组向量$\{\vec{p}^{(0)}, \vec{p}^{(1)}, \ldots, \vec{p}^{(l)}\}$，满足

$$\left(\mathbf{A}\vec{p}^{(i)}, \vec{p}^{(j)}\right) = 0, i \neq j$$

则称为一个A-共轭向量组

由**A**对称正定的任取性，当取单位阵**I**时，这组向量即退化成正交向量；而一般情况下，这是一组线性无关的向量

我们将搜索方向从梯度下降改为：$\vec{x}^{(k+1)} = \vec{x}^{(k)} + \alpha^{(k)}\vec{p}^{(k)}$，即为共轭梯度法（CG法）

对比 $\vec{x}^{(k+1)} = \vec{x}^{(k)} + \alpha^{(k)}\vec{r}^{(k)}$

具体步骤（不加推导和证明）

取初始值 $\vec{x}^{(0)} \in \mathbb{R}^n$ ⇨ $\vec{r}^{(0)} = \vec{b} - \mathbf{A}\vec{x}^{(0)}$, $\vec{p}^{(0)} = \vec{r}^{(0)}$

对 $k = 0, 1, \ldots$

$$\alpha^{(k)} = \frac{\left(\vec{r}^{(k)}, \vec{r}^{(k)}\right)}{\left(\mathbf{A}\vec{p}^{(k)}, \vec{p}^{(k)}\right)} \Rightarrow \vec{x}^{(k+1)} = \vec{x}^{(k)} + \alpha^{(k)}\vec{p}^{(k)}$$

$$\Rightarrow \vec{r}^{(k+1)} = \vec{b} - \alpha^{(k)}\mathbf{A}\vec{p}^{(k)} \Rightarrow \beta^{(k)} = \frac{\left(\vec{r}^{(k+1)}, \vec{r}^{(k+1)}\right)}{\left(\vec{r}^{(k)}, \vec{r}^{(k)}\right)}$$

$$\Rightarrow \vec{p}^{(k+1)} = \vec{r}^{(k)} + \beta^{(k)}\vec{p}^{(k)}$$

内容提纲

3.1 概论及背景知识

3.2 Jacobi迭代法和Gauss-Seidel迭代法

3.3 超松弛迭代法

3.4 共轭梯度法*

3.5 编程实践（穿插在两周中）

✓ 通过SOR迭代求解线性方程组

Jacobi迭代

```
x = BJ * x + fJ;
```

每次采用"最新"算出的计算值

Gauss-Seidel迭代

```
for (int i = 0; i < x.Length; i++)
{
    x[i] = BJ.GetRow(i) * x + fJ[i];
}
```

通过旧值和新值进行插值

SOR迭代

```
for (int i = 0; i < x.Length; i++)
{
    x[i] += w * (BJ.GetRow(i) * x + fJ[i] - x[i]);
}
```

$$B_J = I - D^{-1}A \qquad f_J = D^{-1}b$$

$$B_G = I - (D-L)^{-1}A \qquad f_G = (D-L)^{-1}b$$

$$B_S = (D-\omega L)^{-1}[(1-\omega L)D + \omega U]$$

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

$$x_i^{(k+1)} = x_i^{(k)} + \omega\left(\boxed{\bar{x}_i^{(k+1)}} - x_i^{(k)}\right)$$

$\omega < 0$ $\quad$ $\omega \in (0,1)$ $\quad$ $\omega > 1$

$\omega = 0$ $\quad$ $\omega = 1$

$x_i^{(k)}$ $\quad$ $\bar{x}_i^{(k+1)}$

✓ 通过CG方法求解线性方程组

```
Vector x = new Vector(b.Length);
Vector r = b, p = b, rOld;
double alpha, beta;
int count = 0;
while (Norm.One(r) > 1e-10 && !p.IsZero())
{
    if (count >= maxCount)
        throw new Exception("……");
    alpha = r * r / (A * p * p);
    x += alpha * p;
    rOld = r;
    r -= alpha * A * p;
    beta = r * r / (rOld * rOld);
    p = r + beta * p;
    count++;
}
```

- 任取x，取为零向量
- $r = b - Ax = b$，$p = r = b$
- 迭代

$$\alpha = \frac{(r,\ r)}{(Ap,\ p)}$$

$$x = x + \alpha p$$

$$\boxed{\begin{aligned} r_{old} &= r \\ r &= r - \alpha A p \end{aligned}}$$

$$\beta = \frac{(r,\ r)}{(r_{old},\ r_{old})}$$

$$p = r + \beta p$$

学习重点

- 掌握迭代法的**收敛性**分析和**收敛率**概念
- 理解**雅克比迭代法**
- 掌握**高斯—赛德尔迭代法**
- 掌握**超松弛方法**及**最优松弛因子**
- 理解**共轭梯度法**

作业

教材P90-8、P91-11