第6章 解线性方程组的迭代法

6.1 迭代法的基本概念

6.1.1 引言

定义 1 迭代法，$\boldsymbol{x}=\boldsymbol{B}\boldsymbol{x}+\boldsymbol{f}$（一阶定常迭代法，这里$\boldsymbol{B}$与$k$无关）

迭代法收敛

$$\lim_{k\to\infty}\boldsymbol{\varepsilon}^{(k)}=\mathbf{0}\iff\lim_{k\to\infty}\boldsymbol{B}^{k}=\mathbf{0}$$

6.1.2 向量序列与矩阵序列的极限

定义 2 向量序列的收敛

定义 3 矩阵序列的收敛

<u>矩阵序列极限的概念可以用矩阵算子范数来描述。</u>

定理 1 $\lim_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}\iff\lim_{k\to\infty}\|\boldsymbol{A}_k-\boldsymbol{A}\|=0$

定理 2 $\lim\boldsymbol{A}_k=0\iff\lim\boldsymbol{A}_k\boldsymbol{x}=0,\ \forall\boldsymbol{x}\in\mathbb{R}^n$

定理 3 以下三种说法等价

(1) $\lim_{k\to\infty}\boldsymbol{B}^k=0$

(2) $\rho(\boldsymbol{B})<1$

(3) 至少存在一种从属的矩阵范数$\|\cdot\|_\varepsilon$，使得$\|\boldsymbol{B}\|_\varepsilon<1$

定理 4 矩阵范数极限和谱半径的关系，课上没讲

6.1.3 迭代法及其收敛性

分裂矩阵

定理 5 迭代法（1.11）收敛的充要条件是$\rho(\boldsymbol{B})<1$

定理5是定常迭代法的基本定理

> **若当标准型**
>
> 每个n阶的复数矩阵A都与一个若尔当形矩阵相似，这个若尔当形矩阵除去其中若尔当块的排列次序是被矩阵A唯一确定的，它称为矩阵A的若尔当标准型。首先，Jordan标准型由主对角线为特征值，主对角线上方相邻斜对角线为1的若当块按对角排列组成的矩阵称为Jordan形矩阵，而主对角线上的小块方阵Ji称为Jordan块；其次，每个n阶的复数矩阵A都与一个若尔当形矩阵相似，这个若尔当形矩阵除去其中若尔当块的排列次序是被矩阵A唯一确定的，它成为矩阵A的若尔当标准型。——百度百科
>
> **若当块**

$$
\boldsymbol{J}_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}
$$

定理 6（迭代法收敛的充分条件）设有线性方程组及一阶定常迭代法，如果有B的某种算子范数$\|\boldsymbol{B}\| = q < 1$，则

(1) 迭代法收敛，即对任取$\boldsymbol{x}^0$有

$$
\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^*, \text{ and } \boldsymbol{x}^* = \boldsymbol{B}\boldsymbol{x}^* + \boldsymbol{f}.
$$

(2) $\|\boldsymbol{x}^* - \boldsymbol{x}^{(k)}\| \le q^k \|\boldsymbol{x}^* - \boldsymbol{x}^{(0)}\|$

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

(4) $\|\boldsymbol{x}^* - \boldsymbol{x}^{(k)}\| \le \frac{q^k}{1-q} \|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\|$

> ρ较大时，$\|\boldsymbol{B}\|$可能大于1，但还是收敛，不过速度会很慢
>
> 由迭代矩阵$\|\boldsymbol{B}\| < 1$不但可判别收敛性，还可估计迭代的精度
>
> 可用$|\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_i^{(k-1)}| < \varepsilon$作为迭代的停止准则，需要注意当$\|\boldsymbol{B}\| < 1$且充分接近1时，$\frac{\|\boldsymbol{B}\|}{1-\|\boldsymbol{B}\|}$很大，尽管$|\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_i^{(k-1)}|$很小，$|\boldsymbol{x}_i^{(k)} - \boldsymbol{x}_i^*|$仍然较大，迭代收敛可能很慢，中止有风险

定义 4 迭代法的平均收敛速度，课上没讲

6.2 雅可比迭代法与高斯-赛德尔迭代法

注：以下定义中

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

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

不是下三角/上三角矩阵

6.2.1 雅可比迭代法

Jacobi迭代法式

$$\begin{cases} \boldsymbol{x}^{(0)}, \text{ initial vector} \\ \boldsymbol{x}^{(k+1)} = \boldsymbol{B}\boldsymbol{x}^{(k)} + \boldsymbol{f} \end{cases}$$

其中$\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{D}^{-1}\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{D}^{-1}(\boldsymbol{L}+\boldsymbol{U}) \equiv \boldsymbol{J},\ \boldsymbol{f} = \boldsymbol{D}^{-1}\boldsymbol{b}.$ 称J为解$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$的雅可比迭代法的迭代矩阵。

（注：D是对角矩阵）

6.2.2 高斯-赛德尔迭代法

选取分裂矩阵$\boldsymbol{M}$为$\boldsymbol{A}$的下三角矩阵，即选取

$$\boldsymbol{M} = \boldsymbol{D} - \boldsymbol{L}$$

$$\boldsymbol{A} = \boldsymbol{M} - \boldsymbol{N}$$

于是得到GS迭代法

$$\begin{cases} \boldsymbol{x}^{(0)}, \text{ initial vector} \\ \boldsymbol{x}^{(k+1)} = \boldsymbol{B}\boldsymbol{x}^{(k)} + \boldsymbol{f} \end{cases}$$

其中$\boldsymbol{B} = \boldsymbol{I} - (\boldsymbol{D} - \boldsymbol{L}^{-1})\boldsymbol{A} = (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{U} \equiv \boldsymbol{G},\ \boldsymbol{f} = (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{b}.$ 称$\boldsymbol{G} = (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{U}$为解$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$的高斯-赛德尔迭代法的迭代矩阵。

注：G=tril(A)^-1 * (D - triu(A))，即下三角矩阵的逆乘以diag与上三角矩阵的差

> ”串行计算，不适合多线程，实际应用的时候要考虑“

6.2.3 雅可比迭代法与高斯-赛德尔迭代法的收敛性

定理 7 两种迭代法的收敛条件，$\rho < 1$，很显然

> “迭代法的构造方式有很多种，课上举例的只是最经典的做法”

定义 6 （对角占优矩阵）

$$\text{严格对角占优矩阵：} |a_{ij}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$$

$$\text{弱对角占优矩阵：} |a_{ij}| \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$$

（注意：比较的对象是对角元素的绝对值与对应行（/列）其他元素绝对值的**和**）

> **置换矩阵**
>
> 在数学上，特别是在矩阵理论中，置换矩阵是一个方形二进制矩阵，它在每行和每列中只有一个1，而在其他地方则为0。

> 设P 是一个 m×n 的 (0,1) 矩阵，如果 m≤n且 PP′=E，则称 P为一个 m×n的置换矩阵。其中P′是P的转置矩阵，E是m阶单位方阵。

定义 7（可约与不可约矩阵）设$\boldsymbol{A}=(a_{ij})_{n\times n}(n\geq 2)$，如果存在置换矩阵$\boldsymbol{P}$使

$$\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{0} & \boldsymbol{A}_{22}\end{pmatrix}$$

其中$\boldsymbol{A}_{11}$为r阶方阵，$\boldsymbol{A}_{22}$为$n-r$阶方阵$(1\leq r<n)$，则称$\boldsymbol{A}$为**可约矩阵**，否则，如果不存在这样的置换矩阵$\boldsymbol{P}$使上式成立，则称$\boldsymbol{A}$为**不可约矩阵。**

（联系合同变换）

定理 8（对角占优定理）如果$\boldsymbol{A}=(a_{ij})_{n\times n}$为严格对角占优矩阵或$\boldsymbol{A}$为不可约弱对角占优矩阵，则$\boldsymbol{A}$为非奇异矩阵.

定理 9 设$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$，如果：

(1) $\boldsymbol{A}$为严格对角占优矩阵，则解$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$的雅可比迭代法，高斯-赛德尔迭代法均收敛

(2) $\boldsymbol{A}$为弱对角占优矩阵，且$\boldsymbol{A}$为不可约矩阵，则解$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$的雅可比迭代法，高斯-赛德尔迭代法均收敛

定理 10 对称矩阵，且对角元大于0，雅可比收敛的**充要**条件是矩阵正定，高斯-赛德尔的**充分**条件是矩阵正定

6.3 超松弛迭代法

6.3.1 逐次超松弛迭代法

选取分裂矩阵$\boldsymbol{M}$为带参数的下三角矩阵

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

其中$\omega>0$为可选择的松弛因子.

SOR迭代法

迭代矩阵为

$$\boldsymbol{L}_\omega=\boldsymbol{I}-\omega(\boldsymbol{D}-\omega\boldsymbol{L})^{-1}\boldsymbol{A}=(\boldsymbol{D}-\omega\boldsymbol{L})^{-1}((1-\omega)\boldsymbol{D}+\omega\boldsymbol{U})$$

解$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$的SOR方法为

$$\begin{cases}\boldsymbol{x}^{(0)},\ \text{initial vector}\\ \boldsymbol{x}^{(k+1)}=\boldsymbol{L}_\omega\boldsymbol{x}^{(k)}+\boldsymbol{f}\end{cases}$$

> “是一种加权平衡的思想“
>
> ”求聚类中心也是加权平衡“

(1) 显然，当$\omega=1$时，SOR方法即为高斯-赛德尔迭代法

(2) SOR方法每迭代一次主要运算量是计算一次矩阵与向量的乘法

(3) 当$\omega > 1$时，称为超松弛法；当$\omega < 1$时，称为低松弛法

(4) 在计算机实现时可用

$$\max_{1\le i\le n} |\Delta x_i| = \max_{1\le i\le n} |x_i^{(k+1)} - x_i^{(k)}| < \varepsilon$$

控制迭代中止，或用$\|\boldsymbol{r}^{(k)}\|_\infty = \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(k)}\|_\infty < \varepsilon$控制迭代中止（后者更难算）

> ”ω在1到2之间，想知道哪个比较好，那就试呗“

6.3.2 SOR迭代法的收敛性

定理 11（SOR迭代法收敛的必要条件）$0 < \omega < 2$

定理 12 设$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$，如果：

(1) $\boldsymbol{A}$为对称正定矩阵，$\boldsymbol{A} = \boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U}$

(2) $0 < \omega < 2$

则解$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$的SOR迭代法收敛

定理 13 设$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$，如果：

(1) $\boldsymbol{A}$为严格对角占优矩阵（或$\boldsymbol{A}$为弱对角占优不可约矩阵）

(2) $0 < \omega \le 1$

则解$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$的SOR迭代法收敛