数值计算方法

线性方程组的迭代解法

张晓平

2019 年 10 月 12 日

武汉大学数学与统计学院

Table of contents

线性方程组的性态

线性方程组的性态

向量范数

向量范数

定义：向量范数

向量范数是一个 $\|\cdot\|:\mathbb{R}^n \to \mathbb{R}$ 的非负函数，它满足：

(1) 正定性：

$$\|\boldsymbol{x}\| \geqslant 0, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^n, \quad \text{且} \quad \|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = 0$$

(2) 齐次性：

$$\|\alpha \boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^n,\ \alpha \in \mathbb{R}$$

(3) 三角不等式：

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|, \quad \forall\ \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$$

由 (2) 和 (3) 易知，$\forall\ \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ 有

$$\big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| \leqslant \max_{1 \leqslant i \leqslant n} \|e_i\| \sum_{i=1}^{n} |x_i - y_i|.$$

这说明$\|\cdot\|$ 作为 $\mathbb{R}^n$ 的实函数是连续的。

向量范数

定义：p 范数

$$\|\boldsymbol{x}\|_p = (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}, \quad p \geqslant 1$$

- 1 范数

$$\|\boldsymbol{x}\|_1 = |x_1| + \cdots + |x_n|$$

- 2 范数

$$\|\boldsymbol{x}\|_2 = \left(|x_1|^2 + \cdots + |x_n|^2\right)^{\frac{1}{2}} = \sqrt{\boldsymbol{x}^T\boldsymbol{x}}$$

- ∞ 范数

$$\|\boldsymbol{x}\|_\infty = \max_{i=1,\cdots,n} \{|x_i|\}$$

向量范数

定理：范数等价性

设 $\|\cdot\|_\alpha$ 和 $\|\cdot\|_\beta$ 是 $\mathbb{R}^n$ 的任意两个范数，则存在正常数 c_1 和 c_2 使得 $\forall\ \boldsymbol{x} \in \mathbb{R}^n$ 有

$$c_1\|\boldsymbol{x}\|_\alpha \leqslant \|\boldsymbol{x}\|_\beta \leqslant c_2\|\boldsymbol{x}\|_\alpha$$

向量范数

定理：范数等价性

设 $\|\cdot\|_\alpha$ 和 $\|\cdot\|_\beta$ 是 $\mathbb{R}^n$ 的任意两个范数，则存在正常数 c_1 和 c_2 使得 $\forall\ \boldsymbol{x} \in \mathbb{R}^n$ 有

$$c_1\|\boldsymbol{x}\|_\alpha \leqslant \|\boldsymbol{x}\|_\beta \leqslant c_2\|\boldsymbol{x}\|_\alpha$$

请自行验证

$$\begin{aligned}
\|\boldsymbol{x}\|_2 &\leqslant \|\boldsymbol{x}\|_1 \leqslant \sqrt{n}\|\boldsymbol{x}\|_2,\\
\|\boldsymbol{x}\|_\infty &\leqslant \|\boldsymbol{x}\|_2 \leqslant \sqrt{n}\|\boldsymbol{x}\|_\infty,\\
\|\boldsymbol{x}\|_\infty &\leqslant \|\boldsymbol{x}\|_1 \leqslant n\|\boldsymbol{x}\|_\infty,
\end{aligned}$$

向量范数

定理

设 $\boldsymbol{x}^{(k)}, \boldsymbol{x} \in \mathbb{R}^n$，则

$$\lim_{k\to\infty} \|\boldsymbol{x}^{(k)} - \boldsymbol{x}\| = 0 \iff \lim_{k\to\infty} |x_i^{(k)} - x_i| = 0, \ \ i = 1, \cdots, n$$

线性方程组的性态

矩阵范数

矩阵范数

性质

矩阵范数是一个 $\|\cdot\|:\mathbb{R}^{n\times n}\to\mathbb{R}$ 的非负函数，它满足：

(1) 正定性：

$$\|\boldsymbol{A}\|\geqslant 0,\quad \forall\ \boldsymbol{A}\in\mathbb{R}^{n\times n},\ \text{且}\ \|\boldsymbol{A}\|=0\Longleftrightarrow \boldsymbol{A}=0$$

(2) 齐次性：

$$\|\alpha\boldsymbol{A}\|=|\alpha|\cdot\|\boldsymbol{A}\|,\quad \forall\ \boldsymbol{A}\in\mathbb{R}^{n\times n},\ \alpha\in\mathbb{R}$$

(3) 三角不等式：

$$\|\boldsymbol{A}+\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|+\|\boldsymbol{B}\|,\quad \forall\ \boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times n}$$

(4) 相容性：

$$\|\boldsymbol{A}\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|\cdot\|\boldsymbol{B}\|,\quad \forall\ \boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times n}$$

矩阵范数

性质

1° $\mathbb{R}^{n\times n}$ 上的任意两个范数等价

2° 矩阵序列的范数收敛等价于元素收敛，即

$$\lim_{k\to\infty}\|\boldsymbol{A}_k-\boldsymbol{A}\|=0 \iff \lim_{k\to\infty} a_{ij}^{(k)}=a_{ij},\ \ i,j=1,\cdots,n,$$

其中 $\boldsymbol{A}_k=(a_{ij}^{(k)})$.

矩阵范数

定义：矩阵范数与向量范数的相容性

若矩阵范数 $\|\cdot\|_M$ 和向量范数 $\|\cdot\|_v$ 满足

$$\|\boldsymbol{A}\boldsymbol{x}\|_v \leqslant \|\boldsymbol{A}\|_M \|\boldsymbol{x}\|_v, \quad \boldsymbol{A} \in \mathbb{R}^{n\times n}, \quad \boldsymbol{x} \in \mathbb{R}^n,$$

则称矩阵范数 $\|\cdot\|_M$ 和向量范数 $\|\cdot\|_v$ 相容。

矩阵范数

定义：矩阵范数与向量范数的相容性

若矩阵范数 $\|\cdot\|_M$ 和向量范数 $\|\cdot\|_v$ 满足

$$\|\boldsymbol{A}\boldsymbol{x}\|_v \leqslant \|\boldsymbol{A}\|_M \|\boldsymbol{x}\|_v, \quad \boldsymbol{A} \in \mathbb{R}^{n\times n}, \quad \boldsymbol{x} \in \mathbb{R}^n,$$

则称矩阵范数 $\|\cdot\|_M$ 和向量范数 $\|\cdot\|_v$ 相容。

若无特别说明，总假定矩阵范数和向量范数相容。

矩阵范数

定义：从属范数

设 $\|\cdot\|$ 是 $\mathbb{R}^n$ 上的一个向量范数，若定义

$$|||\boldsymbol{A}||| = \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\|, \quad \boldsymbol{A} \in \mathbb{R}^{n\times n},$$

则称 $|||\cdot|||$ 是 $\mathbb{R}^{n\times n}$ 上的一个矩阵范数。

矩阵范数

定义：从属范数

设 $\|\cdot\|$ 是 $\mathbb{R}^n$ 上的一个向量范数，若定义

$$|||\boldsymbol{A}||| = \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\|, \quad \boldsymbol{A} \in \mathbb{R}^{n\times n},$$

则称 $|||\cdot|||$ 是 $\mathbb{R}^{n\times n}$ 上的一个矩阵范数。

这样的矩阵范数 $|||\cdot|||$ 称为从属于向量范数 $\|\cdot\|$ 的矩阵范数，也称由向量范数 $\|\cdot\|$ 诱导出的算子范数。

矩阵范数

由 $\mathbb{R}^n$ 上的 p 范数可诱导出 $\mathbb{R}^{n\times n}$ 上的算子范数 $\|\cdot\|_p$:

$$\|\boldsymbol{A}\|_p = \max_{\|x\|_p=1} \|\boldsymbol{A}\boldsymbol{x}\|_p, \quad \boldsymbol{A} \in \mathbb{R}^{n\times n}.$$

矩阵范数

推论

设 $\boldsymbol{A}=(a_{ij})\in\mathbb{R}^{n\times n}$，则

- **列范数**

$$\|\boldsymbol{A}\|_1=\max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|$$

- **行范数**

$$\|\boldsymbol{A}\|_\infty=\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}|$$

- **谱范数**

$$\|\boldsymbol{A}\|_2=\sqrt{\lambda_{max}(\boldsymbol{A}^T\boldsymbol{A})}$$

其中 $\lambda_{max}(\boldsymbol{A}^T\boldsymbol{A})$ 表示 $\boldsymbol{A}^T\boldsymbol{A}$ 的最大特征值。

矩阵范数

推论

设 $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$，则

- 列范数

$$\|\boldsymbol{A}\|_1 = \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}|$$

- 行范数

$$\|\boldsymbol{A}\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}|$$

- 谱范数

$$\|\boldsymbol{A}\|_2 = \sqrt{\lambda_{max}(\boldsymbol{A}^T\boldsymbol{A})}$$

其中 $\lambda_{max}(\boldsymbol{A}^T\boldsymbol{A})$ 表示 $\boldsymbol{A}^T\boldsymbol{A}$ 的最大特征值。

注

当 $\boldsymbol{A} = 0$ 时定理显然成立，以下证明只考虑 $\boldsymbol{A} \neq 0$。

矩阵范数

证明（列范数）：

设 $\boldsymbol{A} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$，且 $\delta = \|\boldsymbol{a}_{j_0}\| = \max_{1 \leqslant j \leqslant n} \|\boldsymbol{a}_j\|_1$.

证明（列范数）：

设 $\boldsymbol{A}=[\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n]$，且 $\delta=\|\boldsymbol{a}_{j_0}\|=\max_{1\leqslant j\leqslant n}\|\boldsymbol{a}_j\|_1$．对任意满足 $\|\boldsymbol{x}\|_1=1$ 的 $\boldsymbol{x}\in\mathbb{R}^n$，有

$$
\begin{aligned}
\|\boldsymbol{A}\boldsymbol{x}\|_1 &= \left\|\sum_{j=1}^n x_j\boldsymbol{a}_j\right\|_1 \leqslant \sum_{j=1}^n |x_j|\|\boldsymbol{a}_j\|_1 \\
&\leqslant \sum_{j=1}^n |x_j| \max_{1\leqslant j\leqslant n}\|\boldsymbol{a}_j\|_1 = \|\boldsymbol{a}_{j_0}\|_1 = \delta.
\end{aligned}
$$

证明（列范数）：

设 $\boldsymbol{A} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$，且 $\delta = \|\boldsymbol{a}_{j_0}\| = \max_{1\leqslant j\leqslant n} \|\boldsymbol{a}_j\|_1$．对任意满足 $\|\boldsymbol{x}\|_1 = 1$ 的 $\boldsymbol{x} \in \mathbb{R}^n$，有

$$\begin{aligned}\|\boldsymbol{A}\boldsymbol{x}\|_1 &= \left\|\sum_{j=1}^{n} x_j \boldsymbol{a}_j\right\|_1 \leqslant \sum_{j=1}^{n} |x_j| \|\boldsymbol{a}_j\|_1 \\ &\leqslant \sum_{j=1}^{n} |x_j| \max_{1\leqslant j\leqslant n} \|\boldsymbol{a}_j\|_1 = \|\boldsymbol{a}_{j_0}\|_1 = \delta.\end{aligned}$$

因 $\|\boldsymbol{e}_{j_0}\| = 1$ 且 $\|\boldsymbol{A}\boldsymbol{e}_{j_0}\|_1 = \|a_{j_0}\|_1 = \delta$，故

$$\|\boldsymbol{A}\|_1 = \max_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1 = \delta = \max_{1\leqslant j\leqslant n} \|\boldsymbol{a}_j\|_1 = \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{ij}|.$$

证明（行范数）：

设 $\eta = \max\limits_{1\leqslant i\leqslant n} \sum\limits_{j=1}^{n} |a_{ij}|$，对任意满足 $\|x\|_\infty = 1$ 的 $\boldsymbol{x} \in \mathbb{R}^n$，有

$$\|\boldsymbol{A}\boldsymbol{x}\|_\infty = \max_{1\leqslant i\leqslant n} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}||x_j| \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}| = \eta.$$

证明（行范数）：

设 $\eta = \max\limits_{1\leqslant i\leqslant n} \sum\limits_{j=1}^{n} |a_{ij}|$，对任意满足 $\|x\|_\infty = 1$ 的 $\boldsymbol{x} \in \mathbb{R}^n$，有

$$\|\boldsymbol{A}\boldsymbol{x}\|_\infty = \max_{1\leqslant i\leqslant n} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}||x_j| \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{ij}| = \eta.$$

设 $\boldsymbol{A}$ 的第 k 行的 1 范数最大，即 $\eta = \sum_{j=1}^{n} |a_{kj}|$。令

$$\boldsymbol{x}^* = (\mathrm{sgn}(a_{k1}),\ \cdots,\ \mathrm{sgn}(a_{kn}))^T,$$

则

证明（行范数）：

设 $\eta = \max\limits_{1 \leqslant i \leqslant n} \sum\limits_{j=1}^{n} |a_{ij}|$，对任意满足 $\|x\|_\infty = 1$ 的 $\boldsymbol{x} \in \mathbb{R}^n$，有

$$\|\boldsymbol{A}\boldsymbol{x}\|_\infty = \max_{1 \leqslant i \leqslant n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}| |x_j| \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}| = \eta.$$

设 $\boldsymbol{A}$ 的第 k 行的 1 范数最大，即 $\eta = \sum_{j=1}^{n} |a_{kj}|$。令

$$\boldsymbol{x}^* = (\operatorname{sgn}(a_{k1}), \; \cdots, \; \operatorname{sgn}(a_{kn}))^T,$$

则

$$\boldsymbol{A} \neq 0 \implies \|\boldsymbol{x}^*\|_\infty = 1 \implies \|\boldsymbol{A}\boldsymbol{x}^*\|_\infty = \eta$$

证明（行范数）：

设 $\eta = \max\limits_{1 \leqslant i \leqslant n} \sum\limits_{j=1}^{n} |a_{ij}|$，对任意满足 $\|x\|_\infty = 1$ 的 $\boldsymbol{x} \in \mathbb{R}^n$，有

$$\|\boldsymbol{A}\boldsymbol{x}\|_\infty = \max_{1 \leqslant i \leqslant n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}||x_j| \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}| = \eta.$$

设 $\boldsymbol{A}$ 的第 k 行的 1 范数最大，即 $\eta = \sum_{j=1}^{n} |a_{kj}|$。令

$$\boldsymbol{x}^* = (\operatorname{sgn}(a_{k1}), \ \cdots, \ \operatorname{sgn}(a_{kn}))^T,$$

则

$$\boldsymbol{A} \neq 0 \implies \|\boldsymbol{x}^*\|_\infty = 1 \implies \|\boldsymbol{A}\boldsymbol{x}^*\|_\infty = \eta$$

从而

$$\|\boldsymbol{A}\|_\infty = \eta = \max_{1 \leqslant j \leqslant n} \sum_{j=1}^{n} |a_{ij}|.$$

矩阵范数

证明（谱范数）：

$\|\boldsymbol{A}\|_2$ 可表示为

$$\|\boldsymbol{A}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{A}\boldsymbol{x}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \sqrt{\boldsymbol{x}^T(\boldsymbol{A}^T\boldsymbol{A})\boldsymbol{x}}$$

显然 $\boldsymbol{A}^T\boldsymbol{A}$ 对称半正定，故可设其特征值和特征向量分别为

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_2 \geqslant 0, \quad \boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in \mathbb{R}^n \text{且} \|\boldsymbol{v}_i\| = 1.$$

矩阵范数

证明（谱范数）：

$\|\boldsymbol{A}\|_2$ 可表示为

$$\|\boldsymbol{A}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{A}\boldsymbol{x}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \sqrt{\boldsymbol{x}^T(\boldsymbol{A}^T\boldsymbol{A})\boldsymbol{x}}$$

显然 $\boldsymbol{A}^T\boldsymbol{A}$ 对称半正定，故可设其特征值和特征向量分别为

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_2 \geqslant 0, \quad \boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in \mathbb{R}^n \text{且} \|\boldsymbol{v}_i\| = 1.$$

一方面，对任意满足 $\|\boldsymbol{x}\|_2 = 1$ 的 $\boldsymbol{x} \in \mathbb{R}^n$，有 $\boldsymbol{x} = \sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i$ 和 $\sum_{i=1}^{n} \alpha_i^2 = 1$，于是

$$\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{n} \lambda_i \alpha_i^2 \leqslant \lambda_1.$$

矩阵范数

证明（谱范数）：

另一方面，取 $\boldsymbol{x} = \boldsymbol{v}_1$，则有

$$\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{v}_1^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v}_1 = \boldsymbol{v}_1^T\lambda_1\boldsymbol{v}_1 = \lambda_1$$

矩阵范数

证明（谱范数）：

另一方面，取 $\boldsymbol{x} = \boldsymbol{v}_1$，则有

$$\boldsymbol{x}^T \boldsymbol{A}^T \boldsymbol{A} \boldsymbol{x} = \boldsymbol{v}_1^T \boldsymbol{A}^T \boldsymbol{A} \boldsymbol{v}_1 = \boldsymbol{v}_1^T \lambda_1 \boldsymbol{v}_1 = \lambda_1$$

于是

$$\|\boldsymbol{A}\|_2 = \max_{\|\boldsymbol{x}\|_2 = 1} \|\boldsymbol{A}\boldsymbol{x}\|_2 = \sqrt{\lambda_1} = \sqrt{\lambda_{max}(\boldsymbol{A}^T \boldsymbol{A})}$$

矩阵范数

定义：Frobenius 范数

$$\|\boldsymbol{A}\|_F = \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2}$$

它是向量 2 范数的自然推广。

定义：谱半径

设 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$，称

$$\rho(\boldsymbol{A}) = \max\{|\lambda| : \lambda \in \lambda(\boldsymbol{A})\}$$

为 $\boldsymbol{A}$ 的谱半径，其中 $\lambda(\boldsymbol{A})$ 表示 $\boldsymbol{A}$ 的特征值的全体。

矩阵范数

定理：谱半径与矩阵范数的关系

设 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$，则

(1) 对 $\mathbf{C}^{n\times n}$ 上的任意矩阵范数 $\|\cdot\|$，有

$$\rho(\boldsymbol{A}) \leqslant \|\boldsymbol{A}\|$$

(2) $\forall \epsilon > 0$，存在 $\mathbf{C}^{n\times n}$ 上的矩阵范数 $\|\cdot\|$ 使得

$$\|\boldsymbol{A}\| \leqslant \rho(\boldsymbol{A}) + \epsilon$$

矩阵范数

定理：谱半径与矩阵范数的关系

设 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$，则

(1) 对 $\mathbf{C}^{n\times n}$ 上的任意矩阵范数 $\|\cdot\|$，有

$$\rho(\boldsymbol{A}) \leqslant \|\boldsymbol{A}\|$$

(2) $\forall \epsilon > 0$，存在 $\mathbf{C}^{n\times n}$ 上的矩阵范数 $\|\cdot\|$ 使得

$$\|\boldsymbol{A}\| \leqslant \rho(\boldsymbol{A}) + \epsilon$$

证明

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}, \quad \boldsymbol{x} \neq \boldsymbol{0} \quad \Rightarrow \quad |\lambda|\|\boldsymbol{x}\| = \|\boldsymbol{A}\boldsymbol{x}\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{x}\|$$
$$\Rightarrow \quad |\lambda| \leqslant \|\boldsymbol{A}\|.$$

矩阵范数

定理

设 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$，则

$$\lim_{k\to\infty} \boldsymbol{A}^k = 0 \iff \rho(\boldsymbol{A}) < 1$$

定理

设 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$，则

$$\lim_{k\to\infty} \boldsymbol{A}^k = 0 \iff \rho(\boldsymbol{A}) < 1$$

证明

$\Rightarrow$ 令 $|\lambda_0| = \rho(\boldsymbol{A})$，由 $\lim_{k\to\infty} \boldsymbol{A}^k = 0$ 知

$$\rho(\boldsymbol{A})^k = |\lambda_0|^k = |\lambda_0^k| \leqslant \rho(\boldsymbol{A}^k) \leqslant \|\boldsymbol{A}^k\|_2 < 1 \quad (k\to\infty)$$

从而有 $\rho(\boldsymbol{A}) < 1$.

定理

设 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$，则

$$\lim_{k\to\infty} \boldsymbol{A}^k = 0 \iff \rho(\boldsymbol{A}) < 1$$

证明

$\Rightarrow$ 令 $|\lambda_0| = \rho(\boldsymbol{A})$，由 $\lim_{k\to\infty} \boldsymbol{A}^k = 0$ 知

$$\rho(\boldsymbol{A})^k = |\lambda_0|^k = |\lambda_0^k| \leqslant \rho(\boldsymbol{A}^k) \leqslant \|\boldsymbol{A}^k\|_2 < 1 \quad (k\to\infty)$$

从而有 $\rho(\boldsymbol{A}) < 1$.

$\Leftarrow$

$$\begin{aligned} \rho(\boldsymbol{A}) < 1 \quad &\Longrightarrow \quad \text{存在范数}\|\cdot\|\text{使得}\|\boldsymbol{A}\| < 1 \\ &\Longrightarrow \quad 0 \leqslant \|\boldsymbol{A}^k\| \leqslant \|\boldsymbol{A}\|^k \to 0, \quad k\to\infty. \end{aligned}$$

矩阵范数

推论

设 $A \in \mathbf{C}^{n\times n}$，$\mathbf{C}^{n\times n}$ 上的 $\|\cdot\|$ 满足 $\|I\| = 1$，则

$$\|A\| < 1 \implies \begin{cases} I - A\text{可逆} \\ \left\|(I - A)^{-1}\right\| \leqslant \dfrac{1}{1 - \|A\|} \end{cases}$$

证明

若 $\boldsymbol{I}-\boldsymbol{A}$ 奇异，则 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=0$ 有非零解，即存在 $\boldsymbol{x}^*\neq 0$ 使得

$$\boldsymbol{x}^*=\boldsymbol{A}\boldsymbol{x}^*$$

从而

$$\|\boldsymbol{x}^*\|=\|\boldsymbol{A}\boldsymbol{x}^*\|\leqslant\|\boldsymbol{A}\|\|\boldsymbol{x}^*\|$$

因 $\|\boldsymbol{x}^*\|>0$，故 $\|\boldsymbol{A}\|\geqslant 1$，这与已知矛盾，从而 $\boldsymbol{I}-\boldsymbol{A}$ 必非奇异。

证明

若 $\boldsymbol{I}-\boldsymbol{A}$ 奇异，则 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=0$ 有非零解，即存在 $\boldsymbol{x}^*\neq 0$ 使得

$$\boldsymbol{x}^* = \boldsymbol{A}\boldsymbol{x}^*$$

从而

$$\|\boldsymbol{x}^*\| = \|\boldsymbol{A}\boldsymbol{x}^*\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{x}^*\|$$

因 $\|\boldsymbol{x}^*\|>0$，故 $\|\boldsymbol{A}\|\geqslant 1$，这与已知矛盾，从而 $\boldsymbol{I}-\boldsymbol{A}$ 必非奇异。

由 $(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}-\boldsymbol{A})^{-1}=\boldsymbol{I}$ 得

$$(\boldsymbol{I}-\boldsymbol{A})^{-1} = \boldsymbol{I}+\boldsymbol{A}(\boldsymbol{I}-\boldsymbol{A})^{-1}$$

证明

若 $\boldsymbol{I}-\boldsymbol{A}$ 奇异，则 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=0$ 有非零解，即存在 $\boldsymbol{x}^*\neq 0$ 使得

$$\boldsymbol{x}^* = \boldsymbol{A}\boldsymbol{x}^*$$

从而

$$\|\boldsymbol{x}^*\| = \|\boldsymbol{A}\boldsymbol{x}^*\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{x}^*\|$$

因 $\|\boldsymbol{x}^*\|>0$，故 $\|\boldsymbol{A}\|\geqslant 1$，这与已知矛盾，从而 $\boldsymbol{I}-\boldsymbol{A}$ 必非奇异。

由 $(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}-\boldsymbol{A})^{-1}=\boldsymbol{I}$ 得

$$\begin{aligned}
& (\boldsymbol{I}-\boldsymbol{A})^{-1} = \boldsymbol{I}+\boldsymbol{A}(\boldsymbol{I}-\boldsymbol{A})^{-1} \\
\Rightarrow\quad & \|(\boldsymbol{I}-\boldsymbol{A})^{-1}\| \leqslant \|\boldsymbol{I}\| + \|\boldsymbol{A}\|\|(\boldsymbol{I}-\boldsymbol{A})^{-1}\| \\
\Rightarrow\quad & \|(\boldsymbol{I}-\boldsymbol{A})^{-1}\| \leqslant \frac{1}{1-\|\boldsymbol{A}\|}
\end{aligned}$$

线性方程组的性态

线性方程组的敏度分析

线性方程组的敏度分析

线性性方程组的敏感性问题

考察线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

若给 $\boldsymbol{A}$ 和 $\boldsymbol{b}$ 以微小的扰动，其解会有何影响。

例

$$\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} \quad \rightarrow \quad \boldsymbol{x} = (1\ 1)^T.$$

$$\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4+0.0002 \\ 4-0.0002 \end{bmatrix} \quad \rightarrow \quad \boldsymbol{x}^* = (1.5\ 0.5)^T.$$

线性方程组的敏度分析

例

$$\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} \quad \rightarrow \quad \boldsymbol{x} = (1\ 1)^T.$$

$$\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4+0.0002 \\ 4-0.0002 \end{bmatrix} \quad \rightarrow \quad \boldsymbol{x}^* = (1.5\ 0.5)^T.$$

$$\frac{\|\delta \boldsymbol{b}\|_\infty}{\|\boldsymbol{b}\|_\infty} = \frac{1}{20000}$$

线性方程组的敏度分析

例

$$\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} \quad \rightarrow \quad \boldsymbol{x} = (1\ 1)^T.$$

$$\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4+0.0002 \\ 4-0.0002 \end{bmatrix} \quad \rightarrow \quad \boldsymbol{x}^* = (1.5\ 0.5)^T.$$

$$\frac{\|\delta \boldsymbol{b}\|_\infty}{\|\boldsymbol{b}\|_\infty} = \frac{1}{20000} \quad \Rightarrow \quad \frac{\|\boldsymbol{x}^* - \boldsymbol{x}\|_\infty}{\|\boldsymbol{x}\|_\infty} = \frac{1}{2}$$

线性方程组的敏度分析

考察非奇异线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\left.\begin{array}{l} \boldsymbol{A} \to \boldsymbol{A} + \delta\boldsymbol{A} \\[2ex] \boldsymbol{b} \to \boldsymbol{b} + \delta\boldsymbol{b} \end{array}\right\} \Longrightarrow \boldsymbol{x} \to \boldsymbol{x} + \delta\boldsymbol{x}.$$

线性方程组的敏度分析

考察非奇异线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\left.\begin{array}{l} \boldsymbol{A} \to \boldsymbol{A} + \delta\boldsymbol{A} \\[2ex] \boldsymbol{b} \to \boldsymbol{b} + \delta\boldsymbol{b} \end{array}\right\} \implies \boldsymbol{x} \to \boldsymbol{x} + \delta\boldsymbol{x}.$$

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b}$$

$$\implies (\boldsymbol{A} + \delta\boldsymbol{A})\delta\boldsymbol{x} = \delta\boldsymbol{b} - \delta\boldsymbol{A}\boldsymbol{x}$$

线性方程组的敏度分析

考察非奇异线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\left.\begin{array}{l} \boldsymbol{A} \to \boldsymbol{A} + \delta\boldsymbol{A} \\ \boldsymbol{b} \to \boldsymbol{b} + \delta\boldsymbol{b} \end{array}\right\} \implies \boldsymbol{x} \to \boldsymbol{x} + \delta\boldsymbol{x}.$$

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b}$$

$$\implies (\boldsymbol{A} + \delta\boldsymbol{A})\delta\boldsymbol{x} = \delta\boldsymbol{b} - \delta\boldsymbol{A}\boldsymbol{x}$$

只要$\|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\| < 1$，就有$\boldsymbol{A} + \delta\boldsymbol{A}$ 可逆，并且

$$\|(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A})^{-1}\| \leqslant \frac{1}{1 - \|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\|}.$$

考察非奇异线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\left.\begin{array}{l} \boldsymbol{A} \to \boldsymbol{A} + \delta\boldsymbol{A} \\ \boldsymbol{b} \to \boldsymbol{b} + \delta\boldsymbol{b} \end{array}\right\} \implies \boldsymbol{x} \to \boldsymbol{x} + \delta\boldsymbol{x}.$$

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b}$$

$$\implies (\boldsymbol{A} + \delta\boldsymbol{A})\delta\boldsymbol{x} = \delta\boldsymbol{b} - \delta\boldsymbol{A}\boldsymbol{x}$$

只要 $\|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\| < 1$，就有 $\boldsymbol{A} + \delta\boldsymbol{A}$ 可逆，并且

$$\|(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A})^{-1}\| \leqslant \frac{1}{1 - \|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\|}.$$

于是

$$\delta\boldsymbol{x} = (\boldsymbol{A} + \delta\boldsymbol{A})^{-1}(\delta\boldsymbol{b} - \delta\boldsymbol{A}\boldsymbol{x}) = \left(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{-1}(\delta\boldsymbol{b} - \delta\boldsymbol{A}\boldsymbol{x}),$$

线性方程组的敏度分析

定义：条件数

称 $\kappa(\boldsymbol{A}) = \|\boldsymbol{A}^{-1}\| \cdot \|\boldsymbol{A}\|$ 为矩阵 $\boldsymbol{A}$ 的条件数。

线性方程组的敏度分析

由

$$\delta x = \left(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{-1}(\delta\boldsymbol{b} - \delta\boldsymbol{A}\boldsymbol{x})$$

可得

$$\|\delta\boldsymbol{x}\| \quad \leqslant \quad \|(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A})^{-1}\| \cdot \|\boldsymbol{A}^{-1}\| \cdot (\|\delta\boldsymbol{b}\| + \|\delta\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|)$$

线性方程组的敏度分析

由

$$\delta \boldsymbol{x} = \left(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta \boldsymbol{A}\right)^{-1} \boldsymbol{A}^{-1}(\delta \boldsymbol{b} - \delta \boldsymbol{A}\boldsymbol{x})$$

可得

$$\begin{aligned} \|\delta \boldsymbol{x}\| &\leqslant \|(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta \boldsymbol{A})^{-1}\| \cdot \|\boldsymbol{A}^{-1}\| \cdot (\|\delta \boldsymbol{b}\| + \|\delta \boldsymbol{A}\| \cdot \|\boldsymbol{x}\|) \\ &\leqslant \end{aligned}$$

线性方程组的敏度分析

由

$$\delta x = \left(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{-1}(\delta\boldsymbol{b} - \delta\boldsymbol{A}\boldsymbol{x})$$

可得

$$\begin{aligned} \|\delta\boldsymbol{x}\| &\leqslant \|(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A})^{-1}\| \cdot \|\boldsymbol{A}^{-1}\| \cdot (\|\delta\boldsymbol{b}\| + \|\delta\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|) \\ &\leqslant \frac{\|\boldsymbol{A}^{-1}\|}{1 - \|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\|}(\|\delta\boldsymbol{b}\| + \|\delta\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|). \end{aligned}$$

线性方程组的敏度分析

由

$$\delta x = \left(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{-1}(\delta\boldsymbol{b} - \delta\boldsymbol{A}\boldsymbol{x})$$

可得

$$\begin{aligned}
\|\delta\boldsymbol{x}\| &\leqslant \|(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A})^{-1}\| \cdot \|\boldsymbol{A}^{-1}\| \cdot (\|\delta\boldsymbol{b}\| + \|\delta\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|) \\
&\leqslant \frac{\|\boldsymbol{A}^{-1}\|}{1 - \|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\|}(\|\delta\boldsymbol{b}\| + \|\delta\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|).
\end{aligned}$$

两端同时除以 $\|\boldsymbol{x}\|$，并由 $\|\boldsymbol{b}\| \leqslant \|\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|$ 可知

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant \frac{\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|}{1 - \|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\|}\left(\frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{A}\|\|\boldsymbol{x}\|} + \frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}\right)$$

线性方程组的敏度分析

由

$$\delta x = \left(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{-1}(\delta\boldsymbol{b} - \delta\boldsymbol{A}\boldsymbol{x})$$

可得

$$\begin{aligned}\|\delta\boldsymbol{x}\| &\leqslant \|(\boldsymbol{I} + \boldsymbol{A}^{-1}\delta\boldsymbol{A})^{-1}\| \cdot \|\boldsymbol{A}^{-1}\| \cdot (\|\delta\boldsymbol{b}\| + \|\delta\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|) \\ &\leqslant \frac{\|\boldsymbol{A}^{-1}\|}{1 - \|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\|}(\|\delta\boldsymbol{b}\| + \|\delta\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|).\end{aligned}$$

两端同时除以 $\|\boldsymbol{x}\|$，并由 $\|\boldsymbol{b}\| \leqslant \|\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|$ 可知

$$\begin{aligned}\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} &\leqslant \frac{\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|}{1 - \|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\|}\left(\frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{A}\|\|\boldsymbol{x}\|} + \frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}\right) \\ &\leqslant \frac{\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|}{1 - \|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\|}\left(\frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|} + \frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}\right)\end{aligned}$$

线性方程组的敏度分析

定理

条件：

- $\|\cdot\|$ 满足 $\|\boldsymbol{I}\|=1$，$\boldsymbol{A}\in\mathbb{R}^{n\times n}$ 非奇异，$\boldsymbol{b}\in\mathbb{R}^{n}$ 非零
- $\delta\boldsymbol{A}\in\mathbb{R}^{n\times n}$ 满足 $\|\boldsymbol{A}^{-1}\|\cdot\|\delta\boldsymbol{A}\|<1$
- $$\left\{\begin{array}{rcl} \boldsymbol{A}\boldsymbol{x} & = & \boldsymbol{b} \\ (\boldsymbol{A}+\delta\boldsymbol{A})(\boldsymbol{x}+\delta\boldsymbol{x}) & = & \boldsymbol{b}+\delta\boldsymbol{b} \end{array}\right.$$

线性方程组的敏度分析

定理

条件:

- $\|\cdot\|$ 满足 $\|\boldsymbol{I}\| = 1$, $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 非奇异, $\boldsymbol{b} \in \mathbb{R}^n$ 非零
- $\delta\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 满足 $\|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\| < 1$
- $$\begin{cases} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} \\ (\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) &= \boldsymbol{b} + \delta\boldsymbol{b} \end{cases}$$

结论:

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant \frac{\kappa(\boldsymbol{A})}{1 - \kappa(\boldsymbol{A})\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \left(\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|} + \frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}\right).$$

线性方程组的敏度分析

$\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}$较小

线性方程组的敏度分析

$\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}$较小

$$\Longrightarrow \quad \frac{\kappa(\boldsymbol{A})}{1-\kappa(\boldsymbol{A})\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \approx \kappa(\boldsymbol{A})$$

线性方程组的敏度分析

$\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}$较小

$$\Longrightarrow \quad \frac{\kappa(\boldsymbol{A})}{1-\kappa(\boldsymbol{A})\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \approx \kappa(\boldsymbol{A})$$

$$\Longrightarrow \quad \frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant \color{red}{\kappa(\boldsymbol{A})} \left(\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|} + \frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|}\right)$$

线性方程组的敏感度分析

$\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}$较小

$$\Longrightarrow \quad \frac{\kappa(\boldsymbol{A})}{1-\kappa(\boldsymbol{A})\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \approx \kappa(\boldsymbol{A})$$

$$\Longrightarrow \quad \frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant {\color{red}\kappa(\boldsymbol{A})} \left(\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|} + \frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|} \right)$$

结论

- $\boldsymbol{x}$ 的相对误差由 $\boldsymbol{b}$ 和 $\boldsymbol{A}$ 的相对误差之和放大 $\kappa(\boldsymbol{A})$ 倍得来

线性方程组的敏度分析

$$\frac{\|\delta A\|}{\|A\|}\text{较小}$$

$$\Longrightarrow \quad \frac{\kappa(A)}{1-\kappa(A)\frac{\|\delta A\|}{\|A\|}} \approx \kappa(A)$$

$$\Longrightarrow \quad \frac{\|\delta x\|}{\|x\|} \leqslant \kappa(A)\left(\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right)$$

结论

- x 的相对误差由 b 和 A 的相对误差之和放大 $\kappa(A)$ 倍得来
- 若 $\kappa(A)$ 不大，则扰动对解的影响也不会太大

线性方程组的敏度分析

$$\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}\text{较小}$$

$$\Longrightarrow \quad \frac{\kappa(\boldsymbol{A})}{1-\kappa(\boldsymbol{A})\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \approx \kappa(\boldsymbol{A})$$

$$\Longrightarrow \quad \frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant \kappa(\boldsymbol{A})\left(\frac{\|\delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}+\frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|}\right)$$

结论

- $\boldsymbol{x}$ 的相对误差由 $\boldsymbol{b}$ 和 $\boldsymbol{A}$ 的相对误差之和放大 $\kappa(\boldsymbol{A})$ 倍得来
- 若 $\kappa(\boldsymbol{A})$ 不大，则扰动对解的影响也不会太大
- 若 $\kappa(\boldsymbol{A})$ 很大，则扰动对解的影响可能会太大

线性方程组的敏度分析

条件数在一定程度上刻画了扰动对解的影响程度。

线性方程组的敏度分析

条件数在一定程度上刻画了扰动对解的影响程度。

- 若$\kappa(\boldsymbol{A})$ 很大，则称线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 求解问题是病态的，或者说 $\boldsymbol{A}$ 是病态的；

线性方程组的敏度分析

条件数在一定程度上刻画了扰动对解的影响程度。

- 若$\kappa(\boldsymbol{A})$ 很大，则称线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 求解问题是病态的，或者说 $\boldsymbol{A}$ 是病态的；
- 若$\kappa(\boldsymbol{A})$ 很小，则称该线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 求解问题是良态的，或者说 $\boldsymbol{A}$ 是良态的。

线性方程组的敏度分析

条件数与范数有关，$\mathbb{R}^{n\times n}$ 上任意两种范数下的条件数 $\kappa_\alpha(\boldsymbol{A})$ 与 $\kappa_\beta(\boldsymbol{A})$ 都是等价的，即存在常数 c_1 和 c_2，使得

$$c_1\kappa_\alpha(\boldsymbol{A}) \leqslant \kappa_\beta(\boldsymbol{A}) \leqslant c_2\kappa_\alpha(\boldsymbol{A}).$$

例如，

$$\begin{aligned}
\frac{1}{n}\kappa_2(\boldsymbol{A}) &\leqslant \kappa_1(\boldsymbol{A}) \leqslant n\kappa_2(\boldsymbol{A}),\\
\frac{1}{n}\kappa_\infty(\boldsymbol{A}) &\leqslant \kappa_2(\boldsymbol{A}) \leqslant n\kappa_\infty(\boldsymbol{A}),\\
\frac{1}{n^2}\kappa_1(\boldsymbol{A}) &\leqslant \kappa_\infty(\boldsymbol{A}) \leqslant n^2\kappa_2(\boldsymbol{A}).
\end{aligned}$$

推论

条件:

- $\|\cdot\|$ 满足条件 $\|\boldsymbol{I}\| = 1$， $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 非奇异
- $\delta\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 满足 $\|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\| < 1$

结论

- $\boldsymbol{A} + \delta\boldsymbol{A}$ 非奇异
- $$\frac{\|(\boldsymbol{A}+\delta\boldsymbol{A})^{-1} - \boldsymbol{A}^{-1}\|}{\|\boldsymbol{A}^{-1}\|} \leqslant \frac{\kappa(\boldsymbol{A})}{1-\kappa(\boldsymbol{A})\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}.$$

推论

条件：

- $\|\cdot\|$ 满足条件 $\|\boldsymbol{I}\| = 1$，$\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 非奇异
- $\delta\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 满足 $\|\boldsymbol{A}^{-1}\| \cdot \|\delta\boldsymbol{A}\| < 1$

结论

- $\boldsymbol{A} + \delta\boldsymbol{A}$ 非奇异
- $$\frac{\|(\boldsymbol{A}+\delta\boldsymbol{A})^{-1} - \boldsymbol{A}^{-1}\|}{\|\boldsymbol{A}^{-1}\|} \leqslant \frac{\kappa(\boldsymbol{A})}{1 - \kappa(\boldsymbol{A})\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}.$$

这表明 $\kappa(\boldsymbol{A})$ 可当做矩阵求逆问题的条件数。

Jacobi 迭代与 Gauss-Seidel 迭代

Jacobi 迭代与 Gauss-Seidel 迭代

Jacobi 迭代法

Jacobi 迭代法

设 $a_{ii} \neq 0$,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

Jacobi 迭代法

设 $a_{ii} \neq 0$,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

$$\Longrightarrow$$

Jacobi 迭代法

设 $a_{ii} \neq 0$，

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

$$\Longrightarrow \begin{cases} x_1 = \dfrac{1}{a_{11}}\left(\qquad\qquad - a_{12}x_2 - a_{13}x_3 + b_1 \right) \\ x_2 = \dfrac{1}{a_{22}}\left(- a_{21}x_1 \qquad\qquad - a_{23}x_3 + b_2 \right) \\ x_3 = \dfrac{1}{a_{33}}\left(- a_{31}x_1 - a_{32}x_2 \qquad\qquad + b_3 \right) \end{cases}$$

Jacobi 迭代法

取 $\boldsymbol{x}^{(0)} = \left(x_1^{(0)},\ x_2^{(0)},\ x_3^{(0)}\right)^T$，代入上式右端得

$$
\begin{cases}
x_1^{(1)} = \dfrac{1}{a_{11}}\left(\qquad\qquad -a_{12}x_2^{(0)} - a_{13}x_3^{(0)} + b_1\right) \\
x_2^{(1)} = \dfrac{1}{a_{22}}\left(-a_{21}x_1^{(0)} \qquad\qquad - a_{23}x_3^{(0)} + b_2\right) \\
x_3^{(1)} = \dfrac{1}{a_{33}}\left(-a_{31}x_1^{(0)} - a_{32}x_2^{(0)} \qquad\qquad + b_3\right)
\end{cases}
$$

Jacobi 迭代法

$$
\begin{cases}
x_1^{(k+1)} = \dfrac{1}{a_{11}}\Big(& & -a_{12}x_2^{(k)} & - a_{1,3}x_3^{(k)} & +b_1\Big) \\
x_2^{(k+1)} = \dfrac{1}{a_{22}}\Big(& - a_{21}x_1^{(k)} & & - a_{23}x_3^{(k)} & +b_2\Big) \\
x_3^{(k+1)} = \dfrac{1}{a_{33}}\Big(& - a_{31}x_1^{(k)} & -a_{32}x_2^{(k)} & & +b_3\Big)
\end{cases}
$$

Jacobi 迭代法

线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{1}$$

线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{1}$$

令

$$\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{L} + \boldsymbol{U}$$

其中

$$\begin{aligned}
\boldsymbol{D} &= \operatorname{diag}(a_{11},\ a_{22},\ \cdots,\ a_{nn}), \quad a_{ii} \neq 0 \\
\boldsymbol{L} &= \begin{bmatrix} 0 & & & \\ a_{21} & 0 & & \\ \vdots & \ddots & \ddots & \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{bmatrix}, \quad \boldsymbol{U} = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & 0 & \ddots & \vdots \\ & & \ddots & a_{n-1,n} \\ & & & 0 \end{bmatrix}
\end{aligned}$$

Jacobi 迭代法

线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{1}$$

方程 (1) 可写成

$$\boldsymbol{x} = \boldsymbol{M}_J\boldsymbol{x} + \boldsymbol{g}$$

其中

$$\begin{aligned} \boldsymbol{M}_J &= -\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U}), \\ \boldsymbol{g} &= -\boldsymbol{D}^{-1}\boldsymbol{b}. \end{aligned}$$

Jacobi 迭代法

Jacobi 迭代格式

Jacobi 迭代格式为

$$\forall \boldsymbol{x}^{(0)}, \quad \boldsymbol{x}^{(k+1)} = \boldsymbol{M}_J \boldsymbol{x}^{(k)} + \boldsymbol{g}, \quad k = 0, 1, 2, \cdots,$$

其中

$$\begin{aligned} \boldsymbol{M}_J &= -\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U}) \quad \rightarrow \text{Jacobi 迭代矩阵}, \\ \boldsymbol{g} &= -\boldsymbol{D}^{-1}\boldsymbol{b}. \end{aligned}$$

Jacobi 迭代法

Jacobi 迭代格式的分量形式

任给 $x_i^{(0)}\,(i=1,2,\cdots,n)$,

$$
\begin{aligned}
x_i^{(k+1)} &= \frac{1}{a_{ii}}\left(-\sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i\right) \\
&= x_i^{(k)} + \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{n} a_{ij}x_j^{(k)}\right) \\
&\qquad i=1,2,\cdots,n, \quad k=0,1,2,\cdots
\end{aligned}
$$

Jacobi 迭代与 Gauss-Seidel 迭代

Gauss-Seidel 迭代

Gauss-Seidel 迭代

Jacobi 迭代

$$\begin{cases} x_1^{(k+1)} = \dfrac{1}{a_{11}}(& & -a_{12}x_2^{(k)} & - a_{1,3}x_3^{(k)} & +b_1) \\ x_2^{(k+1)} = \dfrac{1}{a_{22}}(& -a_{21}x_1^{(k)} & & - a_{23}x_3^{(k)} & +b_2) \\ x_3^{(k+1)} = \dfrac{1}{a_{33}}(& -a_{31}x_1^{(k)} & -a_{32}x_2^{(k)} & & +b_3) \end{cases}$$

Gauss-Seidel 迭代

Jacobi 迭代

$$\begin{cases} x_1^{(k+1)} = \dfrac{1}{a_{11}}(& & -a_{12}x_2^{(k)} - a_{1,3}x_3^{(k)} & +b_1) \\ x_2^{(k+1)} = \dfrac{1}{a_{22}}(& -a_{21}x_1^{(k)} & - a_{23}x_3^{(k)} & +b_2) \\ x_3^{(k+1)} = \dfrac{1}{a_{33}}(& -a_{31}x_1^{(k)} & -a_{32}x_2^{(k)} & +b_3) \end{cases}$$

Gauss-Seidel 迭代

$$\begin{cases} x_1^{(k+1)} = \dfrac{1}{a_{11}}(& & -a_{12}x_2^{(k)} - a_{13}x_3^{(k)} & +b_1) \\ x_2^{(k+1)} = \dfrac{1}{a_{22}}(& -a_{21}x_1^{(k+1)} & - a_{23}x_3^{(k)} & +b_2) \\ x_3^{(k+1)} = \dfrac{1}{a_{33}}(& -a_{31}x_1^{(k+1)} & -a_{32}x_2^{(k+1)} & +b_3) \end{cases}$$

Gauss-Seidel 迭代格式

Gauss-Seidel 迭代格式为

$$\forall \boldsymbol{x}^{(0)}, \quad \boldsymbol{D}\boldsymbol{x}^{(k+1)} = -\boldsymbol{L}\boldsymbol{x}^{(k+1)} - \boldsymbol{U}\boldsymbol{x}^{(k)} + \boldsymbol{b}, \quad k = 0, 1, 2, \cdots,$$

即

$$\forall \boldsymbol{x}^{(0)}, \quad \boldsymbol{x}^{(k+1)} = \boldsymbol{M}_{GS}\boldsymbol{x}^{(k)} + \boldsymbol{g}, \quad k = 0, 1, 2, \cdots,$$

其中

$$\begin{aligned} \boldsymbol{M}_{GS} &= -(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{U}, \\ \boldsymbol{g} &= (\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{b}. \end{aligned}$$

Jacobi 与 Gauss-Seidel 迭代的收敛性分析

Jacobi 与 Gauss-Seidel 迭代的收敛性分析

收敛的充分必要条件

收敛的充分必要条件

定义

对于线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{2}$$

给定初值 $\boldsymbol{x}^{(0)}$，称迭代格式

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{M}\boldsymbol{x}^{(k)} + \boldsymbol{g}, \quad k = 0, 1, 2, \cdots, \tag{3}$$

为单步线性定常迭代，其中

- $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ 为迭代矩阵，
- $\boldsymbol{g} \in \mathbb{R}^{n}$ 为常数项，
- $\boldsymbol{x}^{(0)} \in \mathbb{R}^{n}$ 为初始向量.

收敛的充分必要条件

定义

对于线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{2}$$

给定初值 $\boldsymbol{x}^{(0)}$，称迭代格式

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{M}\boldsymbol{x}^{(k)} + \boldsymbol{g}, \quad k = 0, 1, 2, \cdots, \tag{3}$$

为单步线性定常迭代，其中

- $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ 为迭代矩阵，
- $\boldsymbol{g} \in \mathbb{R}^n$ 为常数项，
- $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ 为初始向量.

若对任意初始向量，由 (3) 产生的迭代序列都有极限，则称该迭代法是收敛的；否则称为发散的。

收敛的充分必要条件

- Jacobi 迭代

$$\boldsymbol{M} = -\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U}), \quad \boldsymbol{g} = \boldsymbol{D}^{-1}\boldsymbol{b}$$

- Gauss-Seidel 迭代

$$\boldsymbol{M} = -(\boldsymbol{D} + \boldsymbol{L})^{-1}\boldsymbol{U}, \quad \boldsymbol{g} = (\boldsymbol{D} + \boldsymbol{L})^{-1}\boldsymbol{b}$$

收敛的充分必要条件

若 (3) 收敛，记其极限为 x^*，则两端取极限得

$$x^* = Mx^* + g \;\Rightarrow\; (I - M)x^* = g. \tag{4}$$

若该式与 (2) 等价，则存在可逆矩阵 $G \in \mathbb{R}^{n\times n}$ 使得

$$G(I - M) = A, \quad Gg = b. \tag{5}$$

当上式成立时，称迭代法 (3) 与方程组 (2)相容。

收敛的充分必要条件

称

$$\boldsymbol{e}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x}^*$$

为 $\boldsymbol{x}^*$ 在第 k 步的误差向量。

收敛的充分必要条件

称

$$\boldsymbol{e}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x}^*$$

为 $\boldsymbol{x}^*$ 在第 k 步的误差向量。由

$$\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b} \quad \xLeftrightarrow{\text{同解方程}} \quad \boldsymbol{x}^* = \boldsymbol{M}\boldsymbol{x}^* + \boldsymbol{g}$$

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{M}\boldsymbol{x}^{(k)} + \boldsymbol{g}$$

知

收敛的充分必要条件

称

$$\boldsymbol{e}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x}^*$$

为 $\boldsymbol{x}^*$ 在第 k 步的误差向量。由

$$\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b} \quad \xLeftrightarrow{\text{同解方程}} \quad \boldsymbol{x}^* = \boldsymbol{M}\boldsymbol{x}^* + \boldsymbol{g}$$

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{M}\boldsymbol{x}^{(k)} + \boldsymbol{g}$$

知

$$\boldsymbol{e}^{(k+1)} = \boldsymbol{M}\boldsymbol{e}^{(k)}, \quad k = 0, 1, 2, \cdots .$$

$$\Rightarrow \quad \boldsymbol{e}^{(k)} = \boldsymbol{M}^k \boldsymbol{e}^{(0)}.$$

收敛的充分必要条件

推论

迭代法 (3) 收敛的充分必要条件是

$$\boldsymbol{M}^k \to 0.$$

收敛的充分必要条件

定理

$M^k \to 0$ 的充分必要条件是 $\rho(M) < 1$.

收敛的充分必要条件

定理

$M^k \to 0$ 的充分必要条件是 $\rho(M) < 1$.

定理

迭代法 (3) 收敛的充分必要条件是 $\rho(M) < 1$.

收敛的充分必要条件

定理

$M^k \to 0$ 的充分必要条件是 $\rho(M) < 1$.

定理

迭代法 (3) 收敛的充分必要条件是 $\rho(M) < 1$.

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。

收敛的充分必要条件

解同一方程组时，Jacobi 迭代矩阵与 Gauss-Seidel 迭代矩阵的谱半径不一定相同，且无包含关系。

例

$$\boldsymbol{A}^1 = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}, \quad \boldsymbol{A}^2 = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

收敛的充分必要条件

对于 $\boldsymbol{A}^1$，

$$\boldsymbol{M}_J^1 = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix}, \quad \boldsymbol{M}_{GS}^1 = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{bmatrix}$$

- 因 $\rho(\boldsymbol{M}_J^1) = 1.081 \times 10^{-5}$，故 Jacobi 迭代收敛；
- 因 $\rho(\boldsymbol{M}_{GS}^1) = 2$，故 Gauss-Seidel 迭代不收敛。

收敛的充分必要条件

对于 $\boldsymbol{A}^2$，

$$\boldsymbol{M}_J^2 = \begin{bmatrix} 0 & 1/2 & -1/2 \\ -1 & 0 & -1 \\ 1/2 & 1/2 & 0 \end{bmatrix}, \quad \boldsymbol{M}_{GS}^2 = \begin{bmatrix} 0 & 1/2 & -1/2 \\ 0 & -1/2 & -1/2 \\ 0 & 0 & -1/2 \end{bmatrix}$$

- 因 $\rho(\boldsymbol{M}_J^2) = 1.118$，故 Jacobi 迭代不收敛；
- 因 $\rho(\boldsymbol{M}_{GS}^2) = 1/2$，故 Gauss-Seidel 迭代不收敛。

Jacobi 与 Gauss-Seidel 迭代的收敛性分析

收敛的充分条件及误差估计

收敛的充分条件及误差估计

- 很显然，计算谱半径非常困难，用它来判断迭代格式是否收敛很不方便。
- 能否找一些比较容易计算的条件呢？

收敛的充分条件及误差估计

定理

若 $\|\boldsymbol{A}\| < 1$，则 $\boldsymbol{I} - \boldsymbol{A}$ 非奇异，且

$$\|(\boldsymbol{I} - \boldsymbol{A})^{-1}\| \leqslant \frac{1}{1 - \|\boldsymbol{A}\|}.$$

收敛的充分条件及误差估计

定理

对于迭代格式

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{M}\boldsymbol{x}^{(k)} + \boldsymbol{g}, \quad k = 0, 1, \cdots .$$

若 $\|\boldsymbol{M}\| = q < 1$，则

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

收敛的充分条件及误差估计

定理

对于迭代格式

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{M}\boldsymbol{x}^{(k)} + \boldsymbol{g}, \quad k = 0, 1, \cdots .$$

若 $\|\boldsymbol{M}\| = q < 1$，则

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

注

由该定理可知，给定近似解的精度，可计算出迭代次数，但实际计算时用起来并不方便。

收敛的充分条件及误差估计

证明

由 $\boldsymbol{e}^{(k)} = \boldsymbol{M}^k \boldsymbol{e}^{(0)}$ 可知

$$\|\boldsymbol{e}^{(k)}\| = \|\boldsymbol{M}^k \boldsymbol{e}^{(0)}\| \leqslant \|\boldsymbol{M}\|^k \|\boldsymbol{e}^{(0)}\| = q^k \|\boldsymbol{e}^{(0)}\|$$

证明

由 $\boldsymbol{e}^{(k)}=\boldsymbol{M}^k\boldsymbol{e}^{(0)}$ 可知

$$\|\boldsymbol{e}^{(k)}\|=\|\boldsymbol{M}^k\boldsymbol{e}^{(0)}\|\leqslant\|\boldsymbol{M}\|^k\|\boldsymbol{e}^{(0)}\|=q^k\|\boldsymbol{e}^{(0)}\|$$

因 $\|\boldsymbol{M}\|<1$，故 $(\boldsymbol{I}-\boldsymbol{M})^{-1}$ 存在且 $\boldsymbol{x}^*=(\boldsymbol{I}-\boldsymbol{M})^{-1}\boldsymbol{g}$。

收敛的充分条件及误差估计

证明

由 $\boldsymbol{e}^{(k)} = \boldsymbol{M}^k \boldsymbol{e}^{(0)}$ 可知

$$\|\boldsymbol{e}^{(k)}\| = \|\boldsymbol{M}^k \boldsymbol{e}^{(0)}\| \leqslant \|\boldsymbol{M}\|^k \|\boldsymbol{e}^{(0)}\| = q^k \|\boldsymbol{e}^{(0)}\|$$

因 $\|\boldsymbol{M}\| < 1$，故 $(\boldsymbol{I} - \boldsymbol{M})^{-1}$ 存在且 $\boldsymbol{x}^* = (\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{g}$。于是

$$\begin{aligned}
\boldsymbol{e}^{(0)} &= \boldsymbol{x}^{(0)} - \boldsymbol{x}^* = \boldsymbol{x}^{(0)} - (\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{g} \\
&= (\boldsymbol{I} - \boldsymbol{M})^{-1}\left[\boldsymbol{x}^{(0)} - (\boldsymbol{M}\boldsymbol{x}^{(0)} + \boldsymbol{g})\right] = (\boldsymbol{I} - \boldsymbol{M})^{-1}\left[\boldsymbol{x}^{(0)} - \boldsymbol{x}^{(1)}\right]
\end{aligned}$$

从而有

$$\|\boldsymbol{e}^{(0)}\| \leqslant \|(\boldsymbol{I} - \boldsymbol{M})^{-1}\| \, \|\boldsymbol{x}^{(0)} - \boldsymbol{x}^{(1)}\|.$$

证明

由 $\boldsymbol{e}^{(k)} = \boldsymbol{M}^k \boldsymbol{e}^{(0)}$ 可知

$$\|\boldsymbol{e}^{(k)}\| = \|\boldsymbol{M}^k \boldsymbol{e}^{(0)}\| \leqslant \|\boldsymbol{M}\|^k \|\boldsymbol{e}^{(0)}\| = q^k \|\boldsymbol{e}^{(0)}\|$$

因 $\|\boldsymbol{M}\| < 1$，故 $(\boldsymbol{I} - \boldsymbol{M})^{-1}$ 存在且 $\boldsymbol{x}^* = (\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{g}$。于是

$$\begin{aligned}
\boldsymbol{e}^{(0)} &= \boldsymbol{x}^{(0)} - \boldsymbol{x}^* = \boldsymbol{x}^{(0)} - (\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{g} \\
&= (\boldsymbol{I} - \boldsymbol{M})^{-1}\left[\boldsymbol{x}^{(0)} - (\boldsymbol{M}\boldsymbol{x}^{(0)} + \boldsymbol{g})\right] = (\boldsymbol{I} - \boldsymbol{M})^{-1}\left[\boldsymbol{x}^{(0)} - \boldsymbol{x}^{(1)}\right]
\end{aligned}$$

从而有

$$\|\boldsymbol{e}^{(0)}\| \leqslant \|(\boldsymbol{I} - \boldsymbol{M})^{-1}\| \; \|\boldsymbol{x}^{(0)} - \boldsymbol{x}^{(1)}\|.$$

注意到当 $\|\boldsymbol{M}\| = q < 1$ 时，有

$$\|(\boldsymbol{I} - \boldsymbol{M})^{-1}\| \leqslant \frac{1}{1 - \|\boldsymbol{M}\|} = \frac{1}{1 - q}.$$

从而定理得证。

收敛的充分条件及误差估计

定理

若 $\|M\| = q < 1$，则

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

收敛的充分条件及误差估计

定理

若 $\|\boldsymbol{M}\| = q < 1$，则

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

注

该定理表明，可用相邻两步近似值的差来判别迭代法是否应该终止，这在实际编程时非常好用。

证明

由

$$\begin{aligned}
\boldsymbol{x}^{(k)} - \boldsymbol{x}^{*} &= \boldsymbol{M}\boldsymbol{x}^{(k-1)} + \boldsymbol{g} - (\boldsymbol{M}\boldsymbol{x}^{*} + \boldsymbol{g}) \\
&= \boldsymbol{M}\boldsymbol{x}^{(k-1)} - \boldsymbol{M}\boldsymbol{x}^{*} \\
&= \boldsymbol{M}\boldsymbol{x}^{(k-1)} - \boldsymbol{M}(\boldsymbol{I} - \boldsymbol{M})^{-1}\boldsymbol{g} \\
&= \boldsymbol{M}(\boldsymbol{I} - \boldsymbol{M})^{-1}(\boldsymbol{x}^{(k-1)} - \boldsymbol{x}^{(k)})
\end{aligned}$$

可得

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

定理得证。

收敛的充分条件及误差估计

注

- 用范数判定迭代法是否收敛虽然只是一个充分条件，但用起来比较方便。
- 常用 1 范数和 ∞ 范数来进行判定。
- 迭代矩阵 M_J 比较容易计算，故范数判别法对 Jacobi 迭代而比较好用。
- 迭代矩阵 M_{GS} 涉及到求下三角阵的逆，计算比较麻烦，故范数判别法不太适用于 G-S 迭代。

收敛的充分条件及误差估计

定理

若 $\boldsymbol{A}$ 对称且 $\forall i,\ a_{ii}>0$，则

$$\text{Jacobi 迭代收敛} \iff \boldsymbol{A},\ 2\boldsymbol{D}-\boldsymbol{A}\text{皆正定}$$

收敛的充分条件及误差估计

证明

因 $a_{ii} > 0$，故 $\boldsymbol{D}^{-1}$ 存在。

- 由

$$\begin{aligned} \boldsymbol{M}_J &= -\boldsymbol{D}^{-1}(\boldsymbol{L}+\boldsymbol{U}) = -\boldsymbol{D}^{-1}(\boldsymbol{A}-\boldsymbol{D}) \\ &= \boldsymbol{D}^{-1/2}(\boldsymbol{I}-\boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2})\boldsymbol{D}^{1/2} \end{aligned}$$

  可知

$$\boldsymbol{M}_J \sim \boldsymbol{I}-\boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}.$$

- 由 $\boldsymbol{A}$ 对称可知 $\boldsymbol{I}-\boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{1/2}$ 也对称，从而其特征值为实数。

于是，$\boldsymbol{M}_J$ 的特征值为实数。

收敛的充分条件及误差估计

证明（充分性）：

由

Jacobi 迭代收敛 $\iff \rho(\boldsymbol{M}_J) < 1 \iff \rho(\boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}) < 1$

知

证明（充分性）：

由

$$\text{Jacobi 迭代收敛} \iff \rho(\boldsymbol{M}_J) < 1 \iff \rho(\boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}) < 1$$

知

$$\begin{cases} \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}\text{的特征值} \in (0,2) \implies \boldsymbol{A}\text{正定} \\ 2\boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}\text{的特征值} \in (0,2) \end{cases}$$

收敛的充分条件及误差估计

证明（充分性）：

由

Jacobi 迭代收敛 $\iff \rho(\boldsymbol{M}_J) < 1 \iff \rho(\boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}) < 1$

知

$$\begin{cases} \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}\text{的特征值} \in (0,2) \implies \boldsymbol{A}\text{正定} \\ 2\boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}\text{的特征值} \in (0,2) \end{cases}$$

又因 $2\boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2} = \boldsymbol{D}^{-1/2}(2\boldsymbol{D} - \boldsymbol{A})\boldsymbol{D}^{-1/2}$，从而知$2\boldsymbol{D} - \boldsymbol{A}$ 正定。

收敛的充分条件及误差估计

证明（必要性）：

$$\left.\begin{array}{c} \boldsymbol{D}^{1/2}(\boldsymbol{I}-\boldsymbol{M}_J)\boldsymbol{D}^{-1/2} = \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2} \\ \boldsymbol{A}\text{正定} \end{array}\right\}$$

$$\Longrightarrow \quad \lambda(\boldsymbol{I}-\boldsymbol{M}_J) > 0 \quad \Longrightarrow \quad \lambda(\boldsymbol{M}_J) < 1.$$

收敛的充分条件及误差估计

证明（必要性）：

$$\left.\begin{array}{c} \boldsymbol{D}^{1/2}(\boldsymbol{I}-\boldsymbol{M}_J)\boldsymbol{D}^{-1/2}=\boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2} \\ \boldsymbol{A}\text{正定} \end{array}\right\}$$

$$\Longrightarrow \quad \lambda(\boldsymbol{I}-\boldsymbol{M}_J)>0 \quad \Longrightarrow \quad \lambda(\boldsymbol{M}_J)<1.$$

$$\left.\begin{array}{l} \boldsymbol{D}^{1/2}(\boldsymbol{I}+\boldsymbol{M}_J)\boldsymbol{D}^{-1/2}=\boldsymbol{D}^{-1/2}(2\boldsymbol{D}-\boldsymbol{A})\boldsymbol{D}^{-1/2} \\ 2\boldsymbol{D}-\boldsymbol{A}\text{正定} \end{array}\right\}$$

$$\Longrightarrow \quad \lambda(\boldsymbol{I}+\boldsymbol{M}_J)>0 \quad \Longrightarrow \quad \lambda(\boldsymbol{M}_J)>-1.$$

证明（必要性）：

$$\left.\begin{array}{c} \boldsymbol{D}^{1/2}(\boldsymbol{I}-\boldsymbol{M}_J)\boldsymbol{D}^{-1/2} = \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2} \\ \boldsymbol{A}\text{正定} \end{array}\right\}$$

$$\Longrightarrow \quad \lambda(\boldsymbol{I}-\boldsymbol{M}_J) > 0 \quad \Longrightarrow \quad \lambda(\boldsymbol{M}_J) < 1.$$

$$\left.\begin{array}{l} \boldsymbol{D}^{1/2}(\boldsymbol{I}+\boldsymbol{M}_J)\boldsymbol{D}^{-1/2} = \boldsymbol{D}^{-1/2}(2\boldsymbol{D}-\boldsymbol{A})\boldsymbol{D}^{-1/2} \\ 2\boldsymbol{D}-\boldsymbol{A}\text{正定} \end{array}\right\}$$

$$\Longrightarrow \quad \lambda(\boldsymbol{I}+\boldsymbol{M}_J) > 0 \quad \Longrightarrow \quad \lambda(\boldsymbol{M}_J) > -1.$$

联立可得 $\rho(\boldsymbol{M}_J) < 1$，从而 Jacobi 迭代收敛。

收敛的充分条件及误差估计

定理

$$A\text{对称正定} \quad \Longrightarrow \quad \text{Gauss-Seidel 迭代收敛}$$

证明

设 λ 与 $\boldsymbol{v}$ 为 $M_{GS}=-(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{U}$ 的特征值与特征向量，故有

$$-(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{v}=\lambda\boldsymbol{v} \quad \xLeftrightarrow{\boldsymbol{U}=\boldsymbol{L}^T} \quad -(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{L}^T\boldsymbol{v}=\lambda\boldsymbol{v}$$

$$\Longrightarrow \quad -\lambda(\boldsymbol{D}+\boldsymbol{L})\boldsymbol{v}=\boldsymbol{L}^T\boldsymbol{v} \Longrightarrow \quad -\lambda\underbrace{\boldsymbol{v}^*\boldsymbol{D}\boldsymbol{v}}_{\delta}-\lambda\underbrace{\boldsymbol{v}^*\boldsymbol{L}\boldsymbol{v}}_{\alpha+i\beta}=\underbrace{\boldsymbol{v}^*\boldsymbol{L}^T\boldsymbol{v}}_{\alpha-i\beta}$$

收敛的充分条件及误差估计

证明

设 λ 与 $\boldsymbol{v}$ 为 $\boldsymbol{M}_{GS} = -(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{U}$ 的特征值与特征向量，故有

$$-(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{v} = \lambda\boldsymbol{v} \quad \xLeftrightarrow{\boldsymbol{U}=\boldsymbol{L}^T} \quad -(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{L}^T\boldsymbol{v} = \lambda\boldsymbol{v}$$

$$\Longrightarrow \quad -\lambda(\boldsymbol{D}+\boldsymbol{L})\boldsymbol{v} = \boldsymbol{L}^T\boldsymbol{v} \Longrightarrow \quad -\lambda\underbrace{\boldsymbol{v}^*\boldsymbol{D}\boldsymbol{v}}_{\delta} - \lambda\underbrace{\boldsymbol{v}^*\boldsymbol{L}\boldsymbol{v}}_{\alpha+i\beta} = \underbrace{\boldsymbol{v}^*\boldsymbol{L}^T\boldsymbol{v}}_{\alpha-i\beta}$$

$$-\lambda\delta - \lambda(\alpha+i\beta) = \alpha - i\beta \quad \xRightarrow{\text{取模}} \quad |\lambda|^2\left((\delta+\alpha)^2+\beta^2\right) = \alpha^2+\beta^2$$

$$\Longrightarrow \quad |\lambda|^2 = \frac{\alpha^2+\beta^2}{(\delta+\alpha)^2+\beta^2} \quad < 1?$$

证明

设 λ 与 $\boldsymbol{v}$ 为 $\boldsymbol{M}_{GS}=-(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{U}$ 的特征值与特征向量，故有

$$-(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{v}=\lambda\boldsymbol{v} \quad \overset{\boldsymbol{U}=\boldsymbol{L}^T}{\Longleftrightarrow} \quad -(\boldsymbol{D}+\boldsymbol{L})^{-1}\boldsymbol{L}^T\boldsymbol{v}=\lambda\boldsymbol{v}$$

$$\Longrightarrow \quad -\lambda(\boldsymbol{D}+\boldsymbol{L})\boldsymbol{v}=\boldsymbol{L}^T\boldsymbol{v} \Longrightarrow \quad -\lambda\underbrace{\boldsymbol{v}^*\boldsymbol{D}\boldsymbol{v}}_{\delta}-\lambda\underbrace{\boldsymbol{v}^*\boldsymbol{L}\boldsymbol{v}}_{\alpha+i\beta}=\underbrace{\boldsymbol{v}^*\boldsymbol{L}^T\boldsymbol{v}}_{\alpha-i\beta}$$

$$-\lambda\delta-\lambda(\alpha+i\beta)=\alpha-i\beta \quad \xrightarrow{\text{取模}} \quad |\lambda|^2\left((\delta+\alpha)^2+\beta^2\right)=\alpha^2+\beta^2$$

$$\Longrightarrow \quad |\lambda|^2=\frac{\alpha^2+\beta^2}{(\delta+\alpha)^2+\beta^2} \quad <1?$$

事实上，

$$0<\boldsymbol{v}^*\boldsymbol{A}\boldsymbol{v}=\boldsymbol{v}^*(\boldsymbol{D}+\boldsymbol{L}+\boldsymbol{L}^T)\boldsymbol{v}=\delta+2\alpha$$

故

$$(\delta+\alpha)^2+\beta^2=\delta^2+2\alpha\delta+\alpha^2+\beta^2=\delta(\delta+2\alpha)+\alpha^2+\beta^2>\alpha^2+\beta^2.$$

定义：对角占优矩阵

设 $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$，对于不等式组

$$|a_{ii}| \geqslant \sum_{j=1,\ j\neq i}^{n} |a_{ij}|, \quad \forall i = 1, 2, \cdots, n \tag{6}$$

- 若至少存在一个 i，不等式严格成立，则称 $\boldsymbol{A}$弱严格对角占优；
- 若对所有 i，不等式均严格成立，则称 $\boldsymbol{A}$严格对角占优。

定理

严格对角占优矩阵是非奇异的。

证明

若 $\boldsymbol{A}$ 奇异，则 $\boldsymbol{A}\boldsymbol{x}=0$ 有非零解 $\boldsymbol{x}$。不妨设 $|x_i|=\|\boldsymbol{x}\|_\infty=1$，则

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

这与 $\boldsymbol{A}$ 严格对角占优矛盾。

推论

若 A 为对角元皆为正的对称矩阵，且严格对角占优，则 A 正定。

推论

若 $\boldsymbol{A}$ 为对角元皆为正的对称矩阵，且严格对角占优，则 $\boldsymbol{A}$ 正定。

证明

设 $\lambda \leqslant 0$，考察矩阵 $\boldsymbol{A}-\lambda \boldsymbol{I}$。

$$
\begin{aligned}
& \boldsymbol{A}-\lambda \boldsymbol{I} \text{只是在 } \boldsymbol{A} \text{ 的对角线上加上了一些正数} \\
\Longrightarrow \quad & \boldsymbol{A}-\lambda \boldsymbol{I} \text{与} \boldsymbol{A} \text{一样，是严格对角占优的} \\
\Longrightarrow \quad & \boldsymbol{A}-\lambda \boldsymbol{I} \text{非奇异}
\end{aligned}
$$

注意到 $\lambda \leqslant 0$ 的任意性，知 $\boldsymbol{A}$ 的特征值均大于 0，从而 $\boldsymbol{A}$ 正定。

收敛的充分条件及误差估计

定理

若 A严格对角占优，则 Jacobi 迭代和 Gauss-Seidel 迭代收敛。

证明

$$A\text{严格对角占优} \quad \Longrightarrow \quad D\text{可逆}$$

收敛的充分条件及误差估计

证明（Jacobi 迭代）：

若 M_J 的某个特征值 $|\lambda| \geqslant 1$，考察 $\lambda \boldsymbol{D} + \boldsymbol{L} + \boldsymbol{U}$.

证明（Jacobi 迭代）：

若 M_J 的某个特征值 $|\lambda| \geqslant 1$，考察 $\lambda \boldsymbol{D}+\boldsymbol{L}+\boldsymbol{U}$.

$$
\begin{aligned}
& \boldsymbol{A}\text{严格对角占优} \\
\Longrightarrow \quad & \lambda \boldsymbol{D}+\boldsymbol{L}+\boldsymbol{U}\text{也严格对角占优} \\
\Longrightarrow \quad & \lambda \boldsymbol{D}+\boldsymbol{L}+\boldsymbol{U}\text{非奇异}
\end{aligned}
$$

证明（Jacobi 迭代）：

若 M_J 的某个特征值 $|\lambda| \geqslant 1$，考察 $\lambda \boldsymbol{D} + \boldsymbol{L} + \boldsymbol{U}$.

$$\begin{aligned} & \boldsymbol{A}\text{严格对角占优} \\ \Longrightarrow\quad & \lambda \boldsymbol{D} + \boldsymbol{L} + \boldsymbol{U}\text{也严格对角占优} \\ \Longrightarrow\quad & \lambda \boldsymbol{D} + \boldsymbol{L} + \boldsymbol{U}\text{非奇异} \end{aligned}$$

而

$$\lambda \boldsymbol{I} - \boldsymbol{M}_J = \lambda \boldsymbol{I} + D^{-1}(\boldsymbol{L} + \boldsymbol{U}) = \boldsymbol{D}^{-1}(\lambda \boldsymbol{D} + \boldsymbol{L} + \boldsymbol{U})$$

$$\Longrightarrow \quad \det(\lambda \boldsymbol{I} - \boldsymbol{M}_J) = \det(\boldsymbol{D}^{-1}) \cdot \det(\lambda \boldsymbol{D} + \boldsymbol{L} + \boldsymbol{U}) \neq 0$$

这与 λ 是 $\boldsymbol{M}_J$ 的特征值矛盾，从而 $\rho(\boldsymbol{M}_J) < 1 \quad \Longrightarrow \quad$ Jacobi 迭代收敛。

收敛的充分条件及误差估计

证明（Gauss-Seidel 迭代）：

若 M_{GS} 的某个特征值 $|\lambda| \geqslant 1$，考察 $\lambda \boldsymbol{D} + \lambda \boldsymbol{L} + \boldsymbol{U}$

收敛的充分条件及误差估计

证明（Gauss-Seidel 迭代）：

若 M_{GS} 的某个特征值 $|\lambda| \geqslant 1$，考察 $\lambda \boldsymbol{D} + \lambda \boldsymbol{L} + \boldsymbol{U}$

$$\begin{aligned} &\boldsymbol{A}\text{严格对角占优} \\ &\implies \lambda \boldsymbol{D} + \lambda \boldsymbol{L} + \boldsymbol{U}\text{也严格对角占优} \\ &\implies \lambda \boldsymbol{D} + \lambda \boldsymbol{L} + \boldsymbol{U}\text{非奇异} \end{aligned}$$

收敛的充分条件及误差估计

证明（Gauss-Seidel 迭代）：

若 M_{GS} 的某个特征值 $|\lambda| \geqslant 1$，考察 $\lambda \boldsymbol{D} + \lambda \boldsymbol{L} + \boldsymbol{U}$

$$\begin{aligned} &\boldsymbol{A}\text{严格对角占优} \\ \Longrightarrow\quad &\lambda \boldsymbol{D} + \lambda \boldsymbol{L} + \boldsymbol{U}\text{也严格对角占优} \\ \Longrightarrow\quad &\lambda \boldsymbol{D} + \lambda \boldsymbol{L} + \boldsymbol{U}\text{非奇异} \end{aligned}$$

而

$$\lambda \boldsymbol{I} - \boldsymbol{M}_{GS} = \lambda \boldsymbol{I} + (\boldsymbol{D} + \boldsymbol{L})^{-1} \boldsymbol{U} = (\boldsymbol{D} + \boldsymbol{L})^{-1} (\lambda \boldsymbol{D} + \lambda \boldsymbol{L} + \boldsymbol{U})$$

$$\Longrightarrow \quad \det(\lambda \boldsymbol{I} - \boldsymbol{M}_{GS}) = \det((\boldsymbol{D} + \boldsymbol{L})^{-1}) \ \cdot\ \det(\lambda \boldsymbol{D} + \lambda \boldsymbol{L} + \boldsymbol{U}) \neq 0$$

这与 λ 是 $\boldsymbol{M}_{GS}$ 的特征值矛盾，从而 $\rho(\boldsymbol{M}_{GS}) < 1 \quad \Longrightarrow \quad$ Gauss-Seidel 迭代收敛。

超松弛迭代法

超松弛迭代法

超松弛迭代 (SOR)

超松弛迭代 (SOR)

超松弛迭代可看做是 Gauss-Seidel 迭代的加速。

超松弛迭代 (SOR)

超松弛迭代可看做是 Gauss-Seidel 迭代的加速。

Gauss-Seidel 迭代：

$$\boldsymbol{x}^{(k+1)} = -\boldsymbol{D}^{-1}\boldsymbol{L}\boldsymbol{x}^{(k+1)} - \boldsymbol{D}^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)} + \boldsymbol{D}^{-1}\boldsymbol{b}, \quad k = 0, 1, 2, \cdots .$$

超松弛迭代可看做是 Gauss-Seidel 迭代的加速。

Gauss-Seidel 迭代:

$$\boldsymbol{x}^{(k+1)} = -\boldsymbol{D}^{-1}\boldsymbol{L}\boldsymbol{x}^{(k+1)} - \boldsymbol{D}^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)} + \boldsymbol{D}^{-1}\boldsymbol{b}, \quad k = 0, 1, 2, \cdots .$$

由

$$\begin{aligned}
\boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \boldsymbol{r}^{(k)}, \\
\boldsymbol{r}^{(k)} &= -\boldsymbol{x}^{(k)} - \boldsymbol{D}^{-1}\boldsymbol{L}x^{(k+1)} - \boldsymbol{D}^{-1}\boldsymbol{U}x^{(k)} + \boldsymbol{D}^{-1}\boldsymbol{b} \\
&= \boldsymbol{D}^{-1}\left[\boldsymbol{b} - \boldsymbol{L}\boldsymbol{x}^{(k+1)} - (\boldsymbol{D} + \boldsymbol{U})\boldsymbol{x}^{(k)}\right]
\end{aligned}$$

可看出,

超松弛迭代 (SOR)

超松弛迭代可看做是 Gauss-Seidel 迭代的加速。

Gauss-Seidel 迭代:

$$\boldsymbol{x}^{(k+1)} = -\boldsymbol{D}^{-1}\boldsymbol{L}\boldsymbol{x}^{(k+1)} - \boldsymbol{D}^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)} + \boldsymbol{D}^{-1}\boldsymbol{b}, \quad k = 0, 1, 2, \cdots .$$

由

$$\begin{aligned} \boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \boldsymbol{r}^{(k)}, \\ \boldsymbol{r}^{(k)} &= -\boldsymbol{x}^{(k)} - \boldsymbol{D}^{-1}\boldsymbol{L}x^{(k+1)} - \boldsymbol{D}^{-1}\boldsymbol{U}x^{(k)} + \boldsymbol{D}^{-1}\boldsymbol{b} \\ &= \boldsymbol{D}^{-1}\left[\boldsymbol{b} - \boldsymbol{L}\boldsymbol{x}^{(k+1)} - (\boldsymbol{D} + \boldsymbol{U})\boldsymbol{x}^{(k)}\right] \end{aligned}$$

可看出，对 Gauss-Seidel 迭代来说，$\boldsymbol{x}^{(k+1)}$ 可看作对 $\boldsymbol{x}^{(k)}$ 加上修正项 $\boldsymbol{r}^{(k)}$而得到。

超松弛迭代 (SOR)

若在$\boldsymbol{r}^{(k)}$前加一个参数 ω，便得松弛法

$$
\begin{aligned}
\boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \omega \boldsymbol{r}^{(k)} \\
&= (1-\omega)\boldsymbol{x}^{(k)} + \omega(-\boldsymbol{D}^{-1}\boldsymbol{L}x^{(k+1)} - \boldsymbol{D}^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)} + \boldsymbol{D}^{-1}\boldsymbol{b}),
\end{aligned} \tag{7}
$$

其中ω 叫做松弛因子。

超松弛迭代 (SOR)

若在$\boldsymbol{r}^{(k)}$前加一个参数 ω，便得松弛法

$$\begin{aligned}\boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \omega \boldsymbol{r}^{(k)} \\ &= (1-\omega)\boldsymbol{x}^{(k)} + \omega(-\boldsymbol{D}^{-1}\boldsymbol{L}x^{(k+1)} - \boldsymbol{D}^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)} + \boldsymbol{D}^{-1}\boldsymbol{b}),\end{aligned} \tag{7}$$

其中ω 叫做松弛因子。

注

- $\omega > 1 \quad \rightarrow$ 超松弛 (Successive Over-Relaxation, SOR)
- $\omega < 1 \quad \rightarrow$ 低松弛 (Successive Under-Relaxation, SUR)
- $\omega = 1 \quad \rightarrow$ Gauss-Seidel 迭代

超松弛迭代 (SOR)

定义：松弛法

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{M}_\omega x^{(k)} + \omega(\boldsymbol{D} + \omega\boldsymbol{L})^{-1}\boldsymbol{b}$$

其中

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

超松弛迭代 (SOR)

注

- 超松弛迭代法是解大型方程组，特别是大型稀疏矩阵方程组的有效方法之一。
- 具有计算公式简单、程序设计容易、占用计算机内存单元较少等优点。
- 只要松弛因子 ω 选择得好，其收敛速度就会加快。

超松弛迭代 (SOR)

例

分别用 Jacobi、Gauss-Seidel 和 SOR 求解

$$\begin{pmatrix} 5 & 1 & -1 & -2 \\ 2 & 8 & 1 & 3 \\ 1 & -2 & -4 & -1 \\ -1 & 3 & 2 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2 \\ -6 \\ 6 \\ 12 \end{pmatrix}$$

- 终止条件:

$$\max|\Delta x_i| < \max|x_i^{(k+1)} - x_i^{(k)}| < 10^{-5}$$

- 精确解:

$$x^\star = (1,\ -2,\ -1,\ 3)^T$$

超松弛迭代 (SOR)

解 (Jacobi)：

$$\left\{\begin{array}{lllllllll}
x_1^{(k+1)} & = & x_1^{(k)} & +\frac{1}{5}(& -2 & -5x_1^{(k)} & -x_2^{(k)} & +x_3^{(k)} & +2x_4^{(k)}) \\
x_2^{(k+1)} & = & x_2^{(k)} & +\frac{1}{8}(& -6 & -2x_1^{(k)} & -8x_2^{(k)} & -x_3^{(k)} & -3x_4^{(k)}) \\
x_3^{(k+1)} & = & x_3^{(k)} & -\frac{1}{4}(& 6 & -x_1^{(k)} & +2x_2^{(k)} & +4x_3^{(k)} & +x_4^{(k)}) \\
x_4^{(k+1)} & = & x_4^{(k)} & +\frac{1}{7}(& 12 & +x_1^{(k)} & -3x_2^{(k)} & -2x_3^{(k)} & -7x_4^{(k)})
\end{array}\right.$$

超松弛迭代 (SOR)

解 (Jacobi)：

$$\left\{\begin{array}{lllllllll}
x_1^{(k+1)} & = & x_1^{(k)} & +\frac{1}{5}(& -2 & -5x_1^{(k)} & -x_2^{(k)} & +x_3^{(k)} & +2x_4^{(k)}) \\
x_2^{(k+1)} & = & x_2^{(k)} & +\frac{1}{8}(& -6 & -2x_1^{(k)} & -8x_2^{(k)} & -x_3^{(k)} & -3x_4^{(k)}) \\
x_3^{(k+1)} & = & x_3^{(k)} & -\frac{1}{4}(& 6 & -x_1^{(k)} & +2x_2^{(k)} & +4x_3^{(k)} & +x_4^{(k)}) \\
x_4^{(k+1)} & = & x_4^{(k)} & +\frac{1}{7}(& 12 & +x_1^{(k)} & -3x_2^{(k)} & -2x_3^{(k)} & -7x_4^{(k)})
\end{array}\right.$$

迭代 24 次后，近似解为

$$x^{(24)} = (0.9999941, \ -1.9999950, \ -1.0000040, \ 2.9999990)^T$$

超松弛迭代 (SOR)

解 (GS)：

$$
\begin{cases}
x_1^{(k+1)} = x_1^{(k)} + \frac{1}{5}\left(-2 - 5x_1^{(k)} - x_2^{(k)} + x_3^{(k)} + 2x_4^{(k)} \right) \\
x_2^{(k+1)} = x_2^{(k)} + \frac{1}{8}\left(-6 - 2x_1^{(k+1)} - 8x_2^{(k)} - x_3^{(k)} - 3x_4^{(k)} \right) \\
x_3^{(k+1)} = x_3^{(k)} - \frac{1}{4}\left(6 - x_1^{(k+1)} + 2x_2^{(k+1)} + 4x_3^{(k)} + x_4^{(k)} \right) \\
x_4^{(k+1)} = x_4^{(k)} + \frac{1}{7}\left(12 \; x_1^{(k+1)} - 3x_2^{(k+1)} - 2x_3^{(k+1)} - 7x_4^{(k)} \right)
\end{cases}
$$

超松弛迭代 (SOR)

解 (GS)：

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \frac{1}{5}(-2 -5x_1^{(k)} -x_2^{(k)} +x_3^{(k)} +2x_4^{(k)}) \\ x_2^{(k+1)} = x_2^{(k)} + \frac{1}{8}(-6 -2x_1^{(k+1)} -8x_2^{(k)} -x_3^{(k)} -3x_4^{(k)}) \\ x_3^{(k+1)} = x_3^{(k)} - \frac{1}{4}(6 -x_1^{(k+1)} +2x_2^{(k+1)} +4x_3^{(k)} +x_4^{(k)}) \\ x_4^{(k+1)} = x_4^{(k)} + \frac{1}{7}(12 \quad x_1^{(k+1)} -3x_2^{(k+1)} -2x_3^{(k+1)} -7x_4^{(k)}) \end{cases}$$

迭代 14 次后，近似解为

$$x^{(14)} = (0.9999966, \ -1.9999970, \ -1.0000040, \ 2.9999990)^T$$

超松弛迭代 (SOR)

解 (SOR)：

$$
\begin{cases}
x_1^{(k+1)} = x_1^{(k)} + \frac{\omega}{5}\left(-2 - 5x_1^{(k)} - x_2^{(k)} + x_3^{(k)} + 2x_4^{(k)} \right) \\
x_2^{(k+1)} = x_2^{(k)} + \frac{\omega}{8}\left(-6 - 2x_1^{(k+1)} - 8x_2^{(k)} - x_3^{(k)} - 3x_4^{(k)} \right) \\
x_3^{(k+1)} = x_3^{(k)} - \frac{\omega}{4}\left(6 - x_1^{(k+1)} + 2x_2^{(k+1)} + 4x_3^{(k)} + x_4^{(k)} \right) \\
x_4^{(k+1)} = x_4^{(k)} + \frac{\omega}{7}\left(12 \quad x_1^{(k+1)} - 3x_2^{(k+1)} - 2x_3^{(k+1)} - 7x_4^{(k)} \right)
\end{cases}
$$

超松弛迭代 (SOR)

解 (SOR)：

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \frac{\omega}{5}\left(-2 - 5x_1^{(k)} - x_2^{(k)} + x_3^{(k)} + 2x_4^{(k)}\right) \\ x_2^{(k+1)} = x_2^{(k)} + \frac{\omega}{8}\left(-6 - 2x_1^{(k+1)} - 8x_2^{(k)} - x_3^{(k)} - 3x_4^{(k)}\right) \\ x_3^{(k+1)} = x_3^{(k)} - \frac{\omega}{4}\left(6 - x_1^{(k+1)} + 2x_2^{(k+1)} + 4x_3^{(k)} + x_4^{(k)}\right) \\ x_4^{(k+1)} = x_4^{(k)} + \frac{\omega}{7}\left(12 \quad x_1^{(k+1)} - 3x_2^{(k+1)} - 2x_3^{(k+1)} - 7x_4^{(k)}\right) \end{cases}$$

取 $\omega = 1.15$，迭代 8 次后，近似解为

$$x^{(8)} = (0.9999965,\ -1.9999970,\ -1.0000010,\ 2.9999990)^T$$

超松弛迭代法

收敛性分析

收敛性分析

定义：超松弛迭代法

SOR 迭代格式为

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{M}_\omega \boldsymbol{x}^{(k)} + \boldsymbol{g}$$

其中

$$\begin{aligned}\boldsymbol{M}_\omega &= (\boldsymbol{D} + \omega \boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{D} - \omega \boldsymbol{U}] \\ \boldsymbol{g} &= \omega(\boldsymbol{D} + \omega \boldsymbol{L})^{-1}\boldsymbol{b}.\end{aligned}$$

收敛性分析

定理：充分必要条件

$$\text{SOR 收敛} \quad \Longleftrightarrow \quad \rho(\boldsymbol{M}_\omega) < 1$$

收敛性分析

定理：必要条件

$$\text{SOR 收敛} \quad \Longrightarrow \quad 0 < \omega < 2.$$

收敛性分析

定理：必要条件

$$\text{SOR 收敛} \quad \Longrightarrow \quad 0 < \omega < 2.$$

该定理表明，对任何系数矩阵，若要 SOR 收敛，必须选取松弛因子 $\omega \in (0, 2)$。

证明

设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\boldsymbol{M}_\omega$ 的 n 个特征值, 由 SOR 收敛可知 $\rho(\boldsymbol{M}_\omega) < 1$，从而有 $|\det(\boldsymbol{M}_\omega)| = |\lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n| < 1$.

收敛性分析

证明

设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\boldsymbol{M}_\omega$ 的 n 个特征值, 由 SOR 收敛可知 $\rho(\boldsymbol{M}_\omega) < 1$，从而有 $|\det(\boldsymbol{M}_\omega)| = |\lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n| < 1$.

由

$$\left.\begin{aligned} \boldsymbol{M}_\omega &= (\boldsymbol{I} + \omega \boldsymbol{D}^{-1} L)^{-1} \cdot [(1-\omega)\boldsymbol{I} - \omega \boldsymbol{D}^{-1} \boldsymbol{U}] \\ \det\left((\boldsymbol{I} + \omega \boldsymbol{D}^{-1} \boldsymbol{L})^{-1}\right) &= 1 \\ \det\left((1-\omega)\boldsymbol{I} - \omega \boldsymbol{D}^{-1} \boldsymbol{U}\right) &= (1-\omega)^n \end{aligned}\right\}$$

可知

收敛性分析

证明

设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\boldsymbol{M}_\omega$ 的 n 个特征值，由 SOR 收敛可知 $\rho(\boldsymbol{M}_\omega) < 1$，从而有 $|\det(\boldsymbol{M}_\omega)| = |\lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n| < 1$.

由

$$\left.\begin{aligned} \boldsymbol{M}_\omega = \left(\boldsymbol{I} + \omega \boldsymbol{D}^{-1} L\right)^{-1} \cdot \left[(1-\omega)\boldsymbol{I} - \omega \boldsymbol{D}^{-1}\boldsymbol{U}\right] \\ \det\left(\left(\boldsymbol{I} + \omega \boldsymbol{D}^{-1}\boldsymbol{L}\right)^{-1}\right) = 1 \\ \det\left((1-\omega)\boldsymbol{I} - \omega \boldsymbol{D}^{-1}\boldsymbol{U}\right) = (1-\omega)^n \end{aligned}\right\}$$

可知

$$|\det(\boldsymbol{M}_\omega)| = |(1-\omega)^n|$$

证明

设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\boldsymbol{M}_\omega$ 的 n 个特征值，由 SOR 收敛可知 $\rho(\boldsymbol{M}_\omega) < 1$，从而有 $|\det(\boldsymbol{M}_\omega)| = |\lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n| < 1$.

由

$$\left.\begin{array}{c} \boldsymbol{M}_\omega = (\boldsymbol{I} + \omega \boldsymbol{D}^{-1} L)^{-1} \cdot \left[(1-\omega)\boldsymbol{I} - \omega \boldsymbol{D}^{-1} \boldsymbol{U}\right] \\ \det\left((\boldsymbol{I} + \omega \boldsymbol{D}^{-1} \boldsymbol{L})^{-1}\right) = 1 \\ \det\left((1-\omega)\boldsymbol{I} - \omega \boldsymbol{D}^{-1} \boldsymbol{U}\right) = (1-\omega)^n \end{array}\right\}$$

可知

$$|\det(\boldsymbol{M}_\omega)| = |(1-\omega)^n|$$

从而

$$|(1-\omega)^n| < 1 \quad \Longrightarrow \quad |1-\omega| < 1 \quad \Longrightarrow \quad 0 < \omega < 2.$$

定理

$$\left.\begin{array}{c}\text{严格对角占优}\\[2ex] \omega\in(0,1)\end{array}\right\}\Longrightarrow\text{SOR 收敛}$$

收敛性分析

证明

设 M_ω 的某个特征值 $|\lambda| \geqslant 1$，考察 $(\lambda - 1 + \omega)\boldsymbol{D} + \lambda\omega\boldsymbol{L} + \omega\boldsymbol{U}$.

证明

设 M_ω 的某个特征值 $|\lambda| \geqslant 1$，考察 $(\lambda - 1 + \omega)\boldsymbol{D} + \lambda\omega\boldsymbol{L} + \omega\boldsymbol{U}$.

$$\boldsymbol{A}\text{严格对角占优且}\lambda - 1 + \omega > \lambda\omega$$

证明

设 M_ω 的某个特征值 $|\lambda| \geqslant 1$，考察 $(\lambda - 1 + \omega)\boldsymbol{D} + \lambda\omega\boldsymbol{L} + \omega\boldsymbol{U}$.

$$\boldsymbol{A}\text{严格对角占优且}\lambda - 1 + \omega > \lambda\omega$$

$$\Longrightarrow \quad (\lambda - 1 + \omega)\boldsymbol{D} + \lambda\omega\boldsymbol{L} + \omega\boldsymbol{U}\text{也严格对角占优}$$

收敛性分析

证明

设 M_ω 的某个特征值 $|\lambda| \geqslant 1$，考察 $(\lambda - 1 + \omega)\boldsymbol{D} + \lambda\omega\boldsymbol{L} + \omega\boldsymbol{U}$.

$$\boldsymbol{A}\text{严格对角占优且}\lambda - 1 + \omega > \lambda\omega$$

$$\Longrightarrow \quad (\lambda - 1 + \omega)\boldsymbol{D} + \lambda\omega\boldsymbol{L} + \omega\boldsymbol{U}\text{也严格对角占优}$$

$$\Longrightarrow \quad (\lambda - 1 + \omega)\boldsymbol{D} + \lambda\omega\boldsymbol{L} + \omega\boldsymbol{U}\text{非奇异}$$

证明

设 $\boldsymbol{M}_\omega$ 的某个特征值 $|\lambda| \geqslant 1$，考察 $(\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}$.

$$\boldsymbol{A}\text{严格对角占优且}\lambda-1+\omega>\lambda\omega$$
$$\Longrightarrow \quad (\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}\text{也严格对角占优}$$
$$\Longrightarrow \quad (\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}\text{非奇异}$$

而

$$\begin{aligned} \lambda\boldsymbol{I}-\boldsymbol{M}_\omega &= \lambda\boldsymbol{I}-(\boldsymbol{D}+\omega\boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{D}-\omega\boldsymbol{U}] \\ &= (\boldsymbol{D}+\omega\boldsymbol{L})^{-1}\left[(\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}\right] \end{aligned}$$

收敛性分析

证明

设 M_ω 的某个特征值 $|\lambda| \geqslant 1$，考察 $(\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}$.

$$\boldsymbol{A}\text{严格对角占优且}\lambda-1+\omega>\lambda\omega$$

$$\Longrightarrow \quad (\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}\text{也严格对角占优}$$

$$\Longrightarrow \quad (\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}\text{非奇异}$$

而

$$\begin{aligned}
\lambda\boldsymbol{I}-\boldsymbol{M}_\omega &= \lambda\boldsymbol{I}-(\boldsymbol{D}+\omega\boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{D}-\omega\boldsymbol{U}] \\
&= (\boldsymbol{D}+\omega\boldsymbol{L})^{-1}\left[(\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}\right] \\
\Longrightarrow \quad & \det(\lambda\boldsymbol{I}+\boldsymbol{M}_\omega)\neq 0
\end{aligned}$$

证明

设 M_ω 的某个特征值 $|\lambda| \geqslant 1$，考察 $(\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}$.

$$\boldsymbol{A}\text{严格对角占优且}\lambda-1+\omega>\lambda\omega$$

$$\Longrightarrow\quad(\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}\text{也严格对角占优}$$

$$\Longrightarrow\quad(\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}\text{非奇异}$$

而

$$\begin{aligned}\lambda\boldsymbol{I}-\boldsymbol{M}_\omega &= \lambda\boldsymbol{I}-(\boldsymbol{D}+\omega\boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{D}-\omega\boldsymbol{U}]\\ &= (\boldsymbol{D}+\omega\boldsymbol{L})^{-1}\left[(\lambda-1+\omega)\boldsymbol{D}+\lambda\omega\boldsymbol{L}+\omega\boldsymbol{U}\right]\\ \Longrightarrow\quad & \det(\lambda\boldsymbol{I}+\boldsymbol{M}_\omega)\neq 0\end{aligned}$$

这与 λ 是 $\boldsymbol{M}_\omega$ 的特征值矛盾，从而 $\rho(\boldsymbol{M}_\omega)<1 \quad\Longrightarrow\quad$ 松弛法收敛。

收敛性分析

定理

对于实对称正定矩阵，

$$\omega \in (0,2) \iff \text{SOR 收敛}$$

收敛性分析

证明

设 λ 与 v 是 M_ω 的特征值与特征向量，故

$$(\boldsymbol{D}+\omega \boldsymbol{L})^{-1}\left[(1-\omega)\boldsymbol{D}-\omega \boldsymbol{U}\right]\boldsymbol{v}=\lambda \boldsymbol{v}.$$

证明

设 λ 与 $\boldsymbol{v}$ 是 M_ω 的特征值与特征向量，故

$$(\boldsymbol{D}+\omega\boldsymbol{L})^{-1}\left[(1-\omega)\boldsymbol{D}-\omega\boldsymbol{U}\right]\boldsymbol{v}=\lambda\boldsymbol{v}.$$

因 $\boldsymbol{U}=\boldsymbol{L}^T$，故

$$\begin{aligned}
&(\boldsymbol{D}+\omega\boldsymbol{L})^{-1}\left[(1-\omega)\boldsymbol{D}-\omega\boldsymbol{L}^T\right]\boldsymbol{v}=\lambda\boldsymbol{v}\\
\Longrightarrow\quad &\left[(1-\omega)\boldsymbol{D}-\omega\boldsymbol{L}^T\right]\boldsymbol{v}=\lambda(\boldsymbol{D}+\omega\boldsymbol{L})\boldsymbol{v}\\
\Longrightarrow\quad &\lambda\left[\boldsymbol{v}^*\boldsymbol{D}\boldsymbol{v}+\omega\boldsymbol{v}^*\boldsymbol{L}\boldsymbol{v}\right]=(1-\omega)\boldsymbol{v}^*\boldsymbol{D}\boldsymbol{v}-\omega\boldsymbol{v}^*\boldsymbol{L}^T\boldsymbol{v}
\end{aligned}$$

令

$$\boldsymbol{v}^*\boldsymbol{D}\boldsymbol{v}=\delta$$

$$\boldsymbol{v}^*\boldsymbol{L}\boldsymbol{v}=\alpha+i\beta\quad\Longrightarrow\quad\boldsymbol{v}^*\boldsymbol{L}^T\boldsymbol{v}=\alpha-i\beta.$$

收敛性分析

证明（续）：

$$\lambda[\delta + \omega(\alpha + i\beta)] = (1 - \omega)\delta - \omega(\alpha - i\beta)$$

收敛性分析

证明（续）：

$$\lambda[\delta + \omega(\alpha + i\beta)] = (1-\omega)\delta - \omega(\alpha - i\beta)$$

$$\xRightarrow{\text{取模}} \quad |\lambda|^2 = \frac{[(1-\omega)\delta - \omega\alpha]^2 + \omega^2\beta^2}{(\delta - \omega\alpha)^2 + \omega^2\beta^2} \quad < 1?$$

收敛性分析

证明（续）：

$$\lambda[\delta + \omega(\alpha + i\beta)] = (1-\omega)\delta - \omega(\alpha - i\beta)$$

$$\xRightarrow{\text{取模}} \quad |\lambda|^2 = \frac{[(1-\omega)\delta - \omega\alpha]^2 + \omega^2\beta^2}{(\delta - \omega\alpha)^2 + \omega^2\beta^2} \quad {\color{red}< 1?}$$

事实上，

$$\begin{aligned}
& [(1-\omega)\delta - \omega\alpha]^2 + \omega^2\beta^2 - (\delta - \omega\alpha)^2 - \omega^2\beta^2 \\
= \quad & (1-\omega)^2\delta^2 - 2(1-\omega)\omega\delta\alpha - \delta^2 + 2\omega\delta\alpha \\
= \quad & [(1-\omega)\delta - \omega\alpha]^2 - (\delta - \omega\alpha)^2 \\
= \quad & \omega\delta(\delta - 2\alpha)(\omega - 2)
\end{aligned}$$

收敛性分析

证明（续）：

又因

$$\left.\begin{array}{r} \boldsymbol{A}\text{正定} \quad \Longrightarrow \quad \delta > 0, \quad \delta - 2\alpha > 0 \\ 0 < \omega < 2 \end{array}\right\} \Longrightarrow \omega\delta(\delta-2\alpha)(\omega-2) < 0$$

于是有

$$|\lambda|^2 < 1,$$

定理得证.