第六章 线性方程组的迭代解法

第二节 向量和矩阵的范数

向量范数

定义 对 $\forall x \in R^n$，若存在对应的非负实数 $\|x\|$，满足

1) $\|x\| \geq 0$，且等号当且仅当 $x=0$ 时成立；（正定性）

2) 对任意实数 α，有 $\|\alpha x\| = |\alpha| \cdot \|x\|$；（齐次性）

3) 对任意 x 和 y，有 $\|x+y\| \leq \|x\| + \|y\|$；（三角不等式）

则称 $\|x\|$ 为向量 x 的范数。

- 常见向量范数：

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

$$\boxed{\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}}$$

$$\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

定理1

定理 1 对于任意向量 x，

$$\lim_{p\to\infty}\|\mathbf{x}\|_p = \|\mathbf{x}\|_\infty \qquad (15)$$

证 因

$$\left(\max_{1\leqslant i\leqslant n}|x_i|^p\right)^{1/p} \leqslant \left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}$$

$$\leqslant \left(n\cdot\max_{1\leqslant i\leqslant n}|x_i|^p\right)^{1/p}$$

故有

$$\|\mathbf{x}\|_\infty \leqslant \|\mathbf{x}\|_p \leqslant n^{1/p}\|\mathbf{x}\|_\infty \qquad (16)$$

令$p\to\infty$，注意到 $n^{1/p}\to 1$，即得式（15）. 证毕.

向量序列的收敛

定义 若存在常数 $C_1, C_2>0$ 使得 $C_1\|x\|_\alpha \le \|x\|_\beta \le C_2\|x\|_\alpha$ 对任意 $x\in R^n$ 都成立，则称 $\|\cdot\|_\alpha$ 和 $\|\cdot\|_\beta$ 是 **等价** 的。

R^n 上的所有向量范数都是等价的。

定义 设向量序列 $\{x^{(k)}\}$ 和向量 x^*，若

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

则称 $\{x^{(k)}\}$ 收敛到 x^*，记作 $\lim_{k\to\infty} x^{(k)} = x^*$。

定理 $$\lim_{k\to\infty} x^{(k)} = x^* \Leftrightarrow \lim_{k\to\infty}\|x^{(k)}-x^*\|=0$$

其中 $\|\cdot\|$ 为任一向量范数。

❖主要性质

性质1: $\|-x\| = \|x\|$

性质2: $|\ \|x\| - \|y\|\ | \leqslant \|x-y\|$

性质3: 向量范数$\|x\|$是R^n上向量x的连续函数.

矩阵范数

定义 对 $\forall A \in R^{m\times n}$，若存在对应的非负实数 $\|A\|$，满足

1) $\|A\| \geq 0$，且等号当且仅当 $A=0$ 时成立；（正定性）

2) 对任意实数 α，有 $\|\alpha A\| = |\alpha|\cdot\|A\|$；（齐次性）

3) 对任意 A 和 B，有 $\|A+B\| \leq \|A\| + \|B\|$；（三角不等式）

4) 对任意 A 和 B，有 $\|AB\| \leq \|A\|\cdot\|B\|$；（相容性）

则称 $\|A\|$ 为矩阵 A 的范数。

定义 设 A 是 n 阶方阵，则称

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

为 A 的谱半径，其中 λ_i 为 A 的特征值。

常见的矩阵范数

- 算子范数：（诱导范数）

由向量范数 $\|\cdot\|_p$ 导出关于矩阵 $A \in R^{n\times n}$ 的 p 范数：

$$\|A\|_p = \max_{x\neq 0}\frac{\|Ax\|_p}{\|x\|_p} = \max_{\|x\|_p=1}\|Ax\|_p$$

典型代表：

$$\|A\|_1 = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}| \quad \text{（1-范数，列和范数）}$$

$$\|A\|_\infty = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}| \quad \text{（}\infty\text{-范数，行和范数）}$$

$$\|A\|_2 = \sqrt{\rho(A^T A)} \quad \text{（2-范数，谱范数）}$$

- Frobenius 范数：$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2}$ （F-范数）

是向量 $\|\cdot\|_2$ 的直接推广，但不是算子范数。

证明

$$\|A\| = \max_{\mathbf{x}\neq 0}\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{\mathbf{x}\neq 0}\left\|A\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right)\right\|$$

而

$$\left\|\frac{\mathbf{x}}{\|\mathbf{x}\|}\right\| = 1$$

故矩阵范数亦可等价地定义为

$$\|A\| = \max_{\|\mathbf{x}\| = 1}\|A\mathbf{x}\|$$

矩阵范数的性质证明

矩阵范数的性质证明

迭代过程的收敛性

迭代法的收敛条件

$$X^{(k+1)} = GX^{k} + d$$

定理1：对任意初始向量$X^{(0)}$及常向量d，上述迭代格式

收敛的充分必要条件是迭代矩阵G的谱半径$\rho(G) < 1$。

定理2：若迭代矩阵G的某种范数 $\|G\| < 1$则上述

确定的迭代法对任意初值$X^{(0)}$均收敛于方程组

$X = GX + d$的唯一解x^{*}。

迭代收敛的充分条件

定理 3

对给定方阵 G，若 $\|G\|<1$，则矩阵 I-G 为非奇异.

证　用反证法.

若 I-G 为奇异阵，则存在非零向量 x，使

$$(I-G)x=0$$

即有

$$x=Gx$$

于是据式（17）得

$$\|\mathbf{x}\|=\|G\mathbf{x}\|\leqslant\|G\|\|\mathbf{x}\|$$

由于 $x\neq 0$，又按题设 G <1，故上式不可能成立.

命题得证.

收敛性的证明

定理 4 若迭代矩阵 G 满足

$$\|G\| < 1$$

则迭代公式（23）对于任意初值 $x^{(0)}$ 均收敛.

证

由于 ， 据定理 3知 I-G 为非奇异阵，

因此 方程组（22）有唯一解 x*：

$$x^* = Gx^* + d$$

得

$$x^{(k+1)} - x^* = G(x^{(k)} - x^*)$$

据此反复递推，并利用条件（24）知

收敛性的证明

$$\left\|x^{(k+1)}-x^{*}\right\| \le \|G\| \left\|(x^{(k)}-x^{*})\right\|$$

因之，有

$$\left\|x^{(k)}-x^{*}\right\| \le \|G\|^{k} \left\|(x^{(0)}-x^{*})\right\| \to 0$$

$$(k \to \infty)$$

故迭代过程收敛.

对角占优矩阵

定义1： 如果矩阵的每一行中，不在主对角线上的所有元素绝对值之和小于主对角线上元素的绝对值，

即

$$\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| < |a_{ii}| \qquad i = 1, 2, \cdots, n$$

则称矩阵A按行严格对角占优，类似地，也有按列严格对角占优。

定理5

若 A 为对角占优阵，则它是非奇异的.

证： 因 A 为对角占优，其主对角线元素 a_{ii} 全不为 0，对角阵 D =diag（a ）为非奇异的

得：

$$I - D^{-1}A = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \cdots & -\frac{a_{2n}}{a_{22}} \\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \cdots & 0 \end{bmatrix} \tag{27}$$

定理5的证明

利用对角占优条件（26）知

$$\|I - D^{-1}A\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{\substack{j=1\\ j\neq i}}^{n} \frac{|a_{ij}|}{|a_{ii}|} < 1$$

故利用定理 3可以断定 $D^{-1}A$ 为非奇异，从而 A 为非奇异.

定理6

若线性方程组$AX=b$的系数矩阵A按行严格对角占优，则雅克比迭代法和高斯——赛得尔迭代法对任意给定初值均收敛。

设　A=D+L+U

证　雅可比公式（11）的迭代矩阵为

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

由前面的证明可知

$$\|G\|_\infty<1$$

所以，雅克比迭代是收敛的

高斯-塞德尔公式的证明

高斯-塞德尔迭代公式为

$$\tilde{G} = -(D+L)^{-1}U$$

令 $y=\tilde{G}x$，则有

$$y = -(D+L)^{-1}Ux$$

$$(D+L)y = -Ux$$

$$Dy = -Ly - Ux$$

$$y = -D^{-1}Ly - D^{-1}Ux$$

高斯-塞德尔公式的证明

写出分量形式有

$$y_i = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} y_j - \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{ii}} x_j ,$$

$$i = 1,2,\cdots,n \qquad (29)$$

设 $\|\mathbf{x}\|_\infty = \max\limits_{1\leqslant i\leqslant n} |x_i| = 1$ 且 $\|\mathbf{y}\|_\infty = \max\limits_{1\leqslant i\leqslant n} |y_i| = |y_k|, 1 \leqslant k \leqslant n$

得

$$\|\mathbf{y}\|_\infty = |y_k| \leqslant \sum_{j=1}^{k-1} \frac{|a_{kj}|}{|a_{kk}|} \|\mathbf{y}\|_\infty + \sum_{j=k+1}^{n} \frac{|a_{kj}|}{|a_{kk}|}$$

高斯-塞德尔公式的证明

得

$$\|\mathbf{y}\|_\infty \leqslant \frac{\sum_{j=k+1}^{n} \frac{|a_{kj}|}{|a_{kk}|}}{1-\sum_{j=1}^{k-1} \frac{|a_{kj}|}{|a_{kk}|}}$$

利用对角占优条件知

$$\|G\|_\infty = \max_{\|\mathbf{x}\|_\infty = 1} \|\mathbf{y}\|_\infty < 1$$

命题得证

线性方程组的性态问题

考虑线性方程组: $Ax = b$

由于系数矩阵和右端项都是通过计算或观察得来的，通常都带有一定的误差，即受到了一些（相对）微小的扰动。那么这些扰动对方程组的解会产生什么样的影响?

例: $\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$ 真解: $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

如果右端项加上一个小扰动: $\delta b = [0.0002, -0.0002]$

则得 $\begin{bmatrix} 2.0002 & 1.9998 \\ 1.9998 & 2.0002 \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} 4.0002 \\ 3.9998 \end{bmatrix} \Rightarrow \tilde{x} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix}$

$\Rightarrow \dfrac{\| x - \tilde{x} \|_\infty}{\| x \|_\infty} = \dfrac{1}{2} = 10000 \times \dfrac{\| \delta \boldsymbol{b} \|_\infty}{\| \boldsymbol{b} \|_\infty} = \dfrac{1}{20000}$

条件数

- 理论分析：

(1) 由于右端项的扰动而引起的解的变化

设 $A(x+\delta x)=b+\delta b \Rightarrow \delta x = A^{-1}\cdot\delta b$

$$\left.\begin{array}{l}\Rightarrow \|\delta x\|\le\|A^{-1}\|\cdot\|\delta b\| \\ \text{又}\quad \|b\|\le\|A\|\cdot\|x\|\end{array}\right\} \Rightarrow \frac{\|\delta x\|}{\|x\|}\le\boxed{\|A\|\cdot\|A^{-1}\|}\frac{\|\delta b\|}{\|b\|}$$

(2) 由于系数矩阵的扰动而引起的解的变化

设 $(A+\delta A)\cdot(x+\delta x)=b \Rightarrow \delta x = A^{-1}\cdot\delta A\cdot(x+\delta x)$

$\Rightarrow \|\delta x\|\le\|A^{-1}\|\cdot\|\delta A\|\cdot\|x+\delta x\|$

$\Rightarrow \dfrac{\|\delta x\|}{\|x+\delta x\|}\le\boxed{\|A\|\cdot\|A^{-1}\|}\dfrac{\|\delta A\|}{\|A\|}$

$Ax=b$ 的条件数
矩阵A 的条件数

条件数性质

$$\boxed{\mathrm{Cond}(A) = \|A\| \cdot \|A^{-1}\|}$$

- 常用的条件数有： $\mathrm{Cond}_\infty(A) = \|A\|_\infty \cdot \|A^{-1}\|_\infty$

  $\mathrm{Cond}_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2$

- 如果 A 是对称正定矩阵，则： $\mathrm{Cond}_2(A) = \dfrac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$

- 条件数的性质：

  (1) $\mathrm{Cond}(A) \geq 1$

  (2) $\mathrm{Cond}(cA) = \mathrm{Cond}(A)$，其中 c 为非零常数

  (3) 当 A 是正交矩阵时，$\mathrm{Cond}_2(A) = 1$

  (4) $\mathrm{Cond}_2(PA) = \mathrm{Cond}_2(AP) = \mathrm{Cond}_2(A)$，其中 P 为正交矩阵

举例

- 良态：系数矩阵的条件数相对较小
- 病态：系数矩阵的条件数相对较大

例：**Hilbert 矩阵**

$$H_n = \begin{bmatrix} 1 & 1/2 & \cdots & 1/n \\ 1/2 & 1/3 & \cdots & 1/(n+1) \\ \vdots & \vdots & \ddots & \vdots \\ 1/n & 1/(n+1) & \cdots & 1/(2n-1) \end{bmatrix}$$

$$\mathrm{Cond}_\infty(H_3) = 748$$

$$\mathrm{Cond}_\infty(H_6) = 2.907 \times 10^7$$

$$\mathrm{Cond}_\infty(H_9) = 1.0996 \times 10^{12}$$