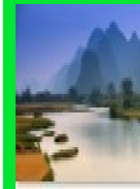
现代数值计算方法

第二章 解线性方程组的迭代法



第二章 解线性方程组的迭代法

在工程计算和科学研究中, 经常会遇到求解 n 阶线性代数方程组的问题. 这种方程组具有如下形式:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \quad\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \tag{2.1}$$

为了讨论方便, 仅考虑实系数的方程组, 即系数 a_{ij} 和常数项 b_i 均为实数, x_i 称为未知数. 记 $A = (a_{ij})_{n\times n}$, $b = (b_1, b_2, \cdots, b_n)^T$, 则可将 (2.1) 写成矩阵形式

$$Ax = b. \tag{2.2}$$

线性方程组的解法大致分为精确解法 (直接法) 和迭代解法 (间接法) 两大类. 本章首先介绍迭代解法, 这类算法的一个突出优点是算法简单, 因而编制程序比较容易. 计算实践表明, 迭代法对于求解大型稀疏方程组是十分有效的, 因为它可以保持系数矩阵稀疏的优点, 从而节省大量的存储量和计算量.

本章的目的就是介绍求解线性方程组 (2.1) 的迭代解法. 主要介绍雅可比 (Jacobi) 迭代法, 高斯-赛德尔 (Gauss-Seidel) 迭代法, 逐次超松弛 (SOR) 迭代法, 并讨论各类迭代算法的收敛性.

§2.1 迭代法的一般理论

§2.1.1 向量范数和矩阵范数

首先介绍向量范数和矩阵范数, 它们是迭代法理论的基础.

1. 向量范数

Back
Close

定义 2.1 设 $x=(x_1,x_2,\cdots,x_n)^T$, $y=(y_1,y_2,\cdots,y_n)^T$, 称

$$(x,y)=x^Ty=\sum_{i=1}^{n}x_iy_i$$

为向量 x 和 y 的内积.

向量内积具有如下性质

(1) $(x,x)\geq 0$, 当且仅当 $x=0$ 时, $(x,x)=0$;　　（非负性）

(2) $(x,y)=(y,x)$;　　（对称性）

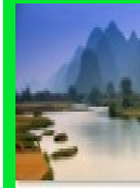
(3) $(\alpha x,y)=\alpha(x,y)$, 其中 α 为某一实数;　　（齐次性）

(4) $(x+y,z)=(x,z)+(y,z)$.　　（可加性）

定义 2.2 若对 $x,y\in R^n$ 有

(1) $\|x\|\geq 0$, 且 $\|x\|=0$ 当且仅当 $x=0$;

(2) $\|\alpha x\|=|\alpha|\cdot\|x\|$;



Back
Close

(3) $\|x+y\| \le \|x\| + \|y\|$;

则称 $\|x\|$ 为向量 x 的范数. 定义了范数的线性空间称为赋范线性空间.

常用的向量范数有

(i) $\|x\|_1 = \sum_{i=1}^{n} |x_i|$, （1-范数）

(ii) $\|x\|_2 = \sqrt{(x,x)} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$, （2-范数）

(iii) $\|x\|_\infty = \max_{1\le i\le n} |x_i|$. （无穷范数）

不难验证, 上述三种范数均满足范数的定义.

2. 矩阵范数

将 $m\times n$ 矩阵 A 看作线性空间 $R^{m\times n}$ 中的元素, 则完全可以按照定义 2.2 的方式引入矩阵的范数. 其中最常用的是与向量 2-范数相

对应的范数

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}$$

称为矩阵 A 的 F-范数.

关于矩阵范数与向量范数之间的关系, 引入相容性的概念.

定义 2.3 对于给定 R^n 上的一种范数 $\|x\|$ 和 $R^{m\times n}$ 上的一种范数 $\|A\|$, 若有

$$\|Ax\| \leq \|A\| \cdot \|x\|, \quad \forall x \in R^n,\ A \in R^{m\times n},$$

则称上述矩阵范数和向量范数是相容的.

这样, 可以利用相容性定义矩阵的范数, 定义

$$\|A\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\| \tag{2.3}$$

Back
Close

为矩阵 A 的范数, 称为相容性范数.

按 (2.3) 可以求出矩阵常用的三种相容性范数:

(1) $\|A\|_1 = \max\limits_{1\le j\le n} \sum\limits_{i=1}^{m} |a_{ij}|,$ （列和范数）

(2) $\|A\|_2 = \sqrt{\lambda_{\max}(A^TA)}$; （谱范数）

(3) $\|A\|_\infty = \max\limits_{1\le i\le m} \sum\limits_{j=1}^{n} |a_{ij}|,$ （行和范数）

其中 $\lambda_{\max}(A^TA)$ 表示矩阵 A^TA 的最大特征值.

3. 谱半径

定义 2.4 设 $A \in R^{n\times n}$, 其特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则称

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

为矩阵 A 的谱半径.

由上述定义, $\|A\|_2$ 可定义为

$$\|A\|_2 = \sqrt{\rho(A^T A)}.$$

特别地, 当 A 为对称矩阵时

$$\|A\|_2 = \rho(A).$$

对于一般情况, 有如下定理.

定理 2.1 设 $A \in R^{n\times n}$, 则 A 的谱半径不超过 A 的任何（相容性）范数, 即

$$\rho(A) \leq \|A\|.$$

证　设 λ 是 A 的一个特征值, 且 A 的谱半径 $\rho(A) = |\lambda|$, u 是对应于 λ 的特征向量, 即 $Au = \lambda u$, 所以对任何一种向量范数, 有

Back
Close

$\|Au\| = |\lambda| \cdot \|u\|$, 故有

$$\rho(A) = |\lambda| = \frac{\|Au\|}{\|u\|} \leq \max_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} = \|A\|.$$

$\square$

定理 2.2 对任意的 $A \in R^{n\times n}$ 和任意正数 ε, 一定存在某种矩阵范数 $\|A\|_\alpha$, 使得

$$\|A\|_\alpha \leq \rho(A) + \varepsilon.$$

(证明略去).

关于矩阵范数, 还有下述结论:

定理 2.3 若 $\|A\| < 1$, 则矩阵 $I - A$ 非奇异, 且满足

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

Back
Close

证 用反证法. 设 $\det(I-A)=0$, 则方程 $(I-A)x=0$ 有非零解, 即存在 $x_0\in R^n$, $x_0\neq 0$, 使得 $(I-A)x_0=0$. 故

$$\|A\|=\max_{\|x\|\neq 0}\frac{\|Ax\|}{\|x\|}\geq\frac{\|Ax_0\|}{\|x_0\|}=1.$$

这与 $\|A\|<1$ 矛盾.

进一步, 由于 $(I-A)(I-A)^{-1}=I$, 则 $(I-A)^{-1}=I+A(I-A)^{-1}$, 从而

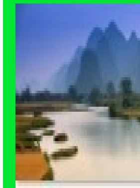
$$\|(I-A)^{-1}\|\leq\|I\|+\|A\|\cdot\|(I-A)^{-1}\|.$$

将上式整理即得要证的结论. □

§2.1.2 迭代格式的构造

首先将方程 (2.2) 的系数矩阵 A 分裂为

$$A=N-P, \tag{2.4}$$



Back
Close

这里要求 N 非奇异, 于是方程 (2.2) 等价地可写成

$$x = N^{-1}Px + N^{-1}b.$$

构造迭代公式如下

$$x^{(k+1)} = Mx^{(k)} + f, \tag{2.5}$$

其中 $M = N^{-1}P$, $f = N^{-1}b$. 称 M 为迭代格式 (2.5) 的迭代矩阵.

对于方程组 (2.2) 和迭代格式 (2.5), 若存在非奇异矩阵 Q, 使

$$M = I - QA, \quad f = Qb, \tag{2.6}$$

则称 Q 为分裂矩阵.

当任选一个解的初始近似 $x^{(0)}$ 后, 即可由 (2.5) 产生一个向量序列 $\{x^{(k)}\}$, 如果它是收敛的, 即

$$\lim_{k\to\infty} x^{(k)} = x^*,$$

对 (2.5) 两边取极限得

$$x^* = Mx^* + f.$$

若 (2.6) 式成立, 可得 $Ax^* = b$, 故 x^* 满足方程组 (2.2), 即迭代格式 (2.5) 和方程组 (2.2) 相容.

对迭代格式 (2.5), 定义误差向量

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

则误差向量有如下的递推关系:

$$e^{(k)} = Me^{(k-1)} = M^2 e^{(k-2)} = \cdots = M^k e^{(0)}, \tag{2.7}$$

这里, $e^{(0)} = x^{(0)} - x^*$ 是解的初始近似 $x^{(0)}$ 与精确解的误差.

引进误差向量后, 迭代的收敛问题就等价于误差向量序列收敛于 0 的问题.

§2.1.3 迭代的收敛性

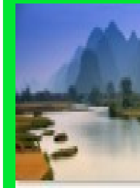
欲使迭代格式 (2.5) 对任意的初始向量 $x^{(0)}$ 都收敛, 误差向量 $e^{(k)}$ 应对任意的初始误差 $e^{(0)}$ 都收敛于零向量. 于是迭代格式 (2.5) 对任意的初始向量都收敛的充分必要条件是

$$\lim_{k\to\infty} M^k = 0. \tag{2.8}$$

定理 2.4 迭代格式 (2.5) 对任意的初始向量 $x^{(0)}$ 都收敛的充分必要条件是 $\rho(M) < 1$, 这里 M 是迭代矩阵, $\rho(M)$ 表示 M 的谱半径.

证　必要性. 设对初始向量 $x^{(0)}$ 迭代格式 (2.5) 是收敛的, 那么 (2.8) 成立. 由定理 2.1, 对于任意的矩阵范数, 成立关系式

$$\rho(M) \le \|M\|.$$



Back
Close

若 $\rho(M) < 1$ 不成立, 即 $\rho(M) \geq 1$, 则

$$\|M^k\| \geq \rho(M^k) = [\rho(M)]^k \geq 1,$$

这与 (2.8) 成立矛盾.

充分性. 若 $\rho(M) < 1$, 则存在一个正数 ε, 使得

$$\rho(M) + 2\varepsilon < 1.$$

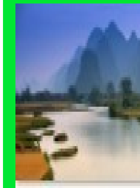
根据定理 2.2, 存在一种矩阵范数 $\|M\|$, 使

$$\|M\| < \rho(M) + \varepsilon < 1 - \varepsilon.$$

故得

$$\|M^k\| \leq \|M\|^k < (1-\varepsilon)^k,$$

从而当 $k \to \infty$ 时, $\|M^k\| \to 0$, 即 $M^k \to 0$, 充分性得证. $\square$



Back
Close

由此可见, 迭代是否收敛仅与迭代矩阵的谱半径有关, 即仅与方程组的系数矩阵和迭代格式的构造有关, 而与方程组的右端向量 b 及初始向量 $x^{(0)}$ 无关.

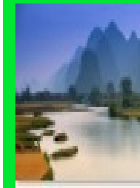
如果迭代格式是收敛的, 我们还可以给出近似解与准确解的误差估计.

定理 2.5 设 M 为迭代矩阵, 若 $\|M\| = q < 1$, 则对迭代 (2.5), 有误差估计式

$$\|x^{(k)} - x^*\| \le \frac{q^k}{1-q}\|x^{(0)} - x^{(1)}\|. \tag{2.9}$$

证 由 (2.7) 有

$$\|x^{(k)} - x^*\| = \|e^{(k)}\| \le \|M^k\| \cdot \|e^{(0)}\| \le q^k\|e^{(0)}\|.$$



Back
Close

注意到 $x^* = (I-M)^{-1}f$, 于是

$$
\begin{aligned}
\|e^{(0)}\| &= \|x^{(0)} - x^*\| = \|x^{(0)} - (I-M)^{-1}f\| \\
&= \|(I-M)^{-1}[(I-M)x^{(0)} - f]\| \\
&= \|(I-M)^{-1}(x^{(0)} - x^{(1)})\| \\
&\le \|(I-M)^{-1}\| \cdot \|x^{(0)} - x^{(1)}\|.
\end{aligned}
$$

因 $\|M\| < 1$, 根据定理 2.3 有

$$\|(I-M)^{-1}\| \le \frac{1}{1-q},$$

于是

$$\|x^{(k)} - x^*\| \le \frac{q^k}{1-q}\|x^{(0)} - x^{(1)}\|. \qquad \square$$

在理论上, 可用上述定理来估计近似解达到某一精度所需要的迭代次数, 但由于 q 不易计算, 故计算实践中很少使用.

Back
Close

定理 2.6 若 $\|M\| < 1$, 则对任一初始近似 $x^{(0)}$, 由迭代格式 (2.5) 产生的向量序列 $\{x^{(k)}\}$ 收敛, 且有估计式

$$\|x^{(k)} - x^*\| \leq \frac{\|M\|}{1 - \|M\|}\|x^{(k)} - x^{(k-1)}\|. \tag{2.10}$$

证 收敛性由定理 2.4 是显然的. 下证误差估计式 (2.10). 由于

$$\begin{aligned}
e^{(k)} &= x^{(k)} - x^* \\
&= (Mx^{(k-1)} + f) - (Mx^* + f) \\
&= Mx^{(k-1)} - Mx^* \\
&= Mx^{(k-1)} - M(I - M)^{-1}f \\
&= M(I - M)^{-1}[(I - M)x^{(k-1)} - f] \\
&= M(I - M)^{-1}(x^{(k-1)} - x^{(k)}),
\end{aligned}$$

利用定理 2.3, 对上式两边取范数即得定理的结论. □

Back
Close

由上述定理可知, 只要 $\|M\|$ 不很接近于 1, 则可用 $\{x^{(k)}\}$ 的相邻两项之差的范数 $\|x^{(k)}-x^{(k-1)}\|$ 来估计 $\|x^{(k)}-x^*\|$ 的大小.

作业： P31: 2.2; P33: 2.10.