第八讲：非线性方程组数值解法

教师: 胡俊

北京大学数学科学学院

May 7, 2019

目录

目录

应用数学中很多问题可以归结为: 对于给定的非线性函数$f:\mathbb{R}^n \to \mathbb{R}^n$,

$$求x^* \in \mathbb{R}^n,\ 使得f(x^*) = 0.$$

当$n = 1$时, 这就是单个变量非线性方程求解问题; 当$n > 1$时, 这就是非线性方程组的求解问题.

Definition 1.1

设序列$\{x_k\}$以x^*为极限. 记$e_k = x_k - x^*$, 如果存在正数r和非负常数C, 使得

$$\lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|^r} = C,$$

则称序列$\{x_k\}$收敛到x^*的收敛速度是r阶的. 常见情形

(1) $r = 1$, 此时称为一阶收敛或线性收敛. 显然, 此时必有$0 < C < 1$.

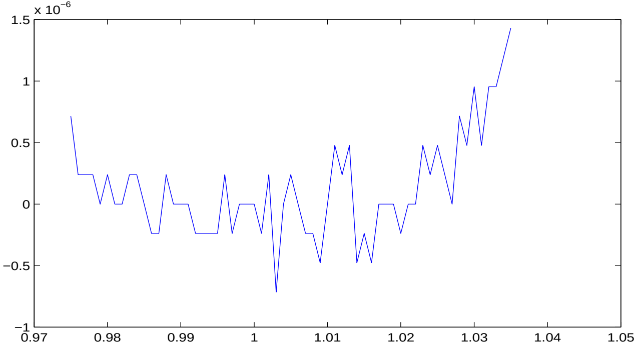
(2) $r > 1$或$r = 1, C = 0$, 此时称为超线性收敛.

(3) $r = 2$, 此时称为二阶收敛.

考虑多项式

$$P_4(x) = x^4 - 4x^3 + 6x^2 - 4x + 1$$

由$P_4(x) = (x-1)^4$, 易知1是唯一零点(4重). 如果用单精度计算, 在区间[0.975, 1.035]内间隔0.001取点, 会发现在很多地方会变号, 如下图所示， 由介值定理应该有许多近似零点. [0.981, 1.026]之间的值都可以看成是真解的近似. 产生这个问题的原因, 一方面是方法的问题, 一方面是精度限制还有舍入误差.



目录

二分法

设$f(x) \in C[a,b]$, 且$f(a)f(b) < 0$, 由介值定理则必有$x^* \in (a,b)$, 使得$f(x^*) = 0$.

令$a_0 = a, b_0 = b, c_0 = a_0 + (b_0 - a_0)/2$, 若$f(a_0)f(c_0) = 0$, 则$x^* = c_0$ 即为所求; 若$f(a_0)f(c_0) > 0$, 则令$a_1 = c_0, b_1 = b_0$; 若$f(a_0)f(c_0) < 0$, 则$a_1 = a_0, b_1 = c_0$. 这样得到的区间(a_1, b_1)含有原方程的根, 且区间长度是原来的一半.

如此进行下去, 假设到第$n-1$步, 令$c_{n-1} = a_{n-1} + (b_{n-1} - a_{n-1})/2$, 若$f(a_{n-1})f(c_{n-1}) = 0$, 则$x^* = c_{n-1}$即为所求; 若$f(a_{n-1})f(c_{n-1}) > 0$, 则令$a_n = c_{n-1}, b_n = b_{n-1}$; 若$f(a_{n-1})f(c_{n-1}) < 0$, 则$a_n = a_{n-1}, b_n = c_{n-1}$. 再如此进行下去. 这样得到的序列$c_n(k = 0, 1, 2, \cdots)$的极限$x^*$是原方程的根.

二分法

二分法

输入$a, b, n, \delta, \epsilon$

$fa = f(a)$

$fb = f(b)$

$e = b - a$

如果$\text{sign}(fa) = \text{sign}(fb)$，则停止

对$k = 0$到n, 进行下面操作

  $e = e/2$

  $c = a + e$

  $fc = f(c)$

  如果$|e| < \delta$或者$|fc| < \epsilon$，则停止

  如果$\text{sign}(fc) \neq \text{sign}(fa)$, 则

  $b = c$

  $fb = fc$

  不然$a = c; fa = fc$

结束

二分法

在上面算法中有三个迭代停止准则，一个是达到最大迭代步数n，另外两个是误差很小或者函数值很小，用δ 和ϵ控制. 下面两个图是有可能其中一个停止准则满足，另外一个不满足.

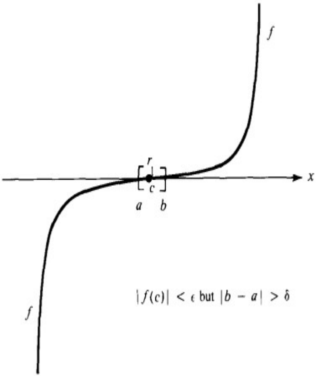


Figure 3.2(a) Criterion $|b - a| < \delta$ failure

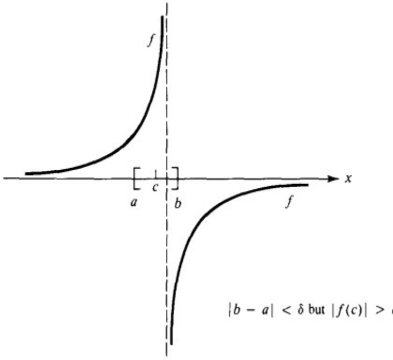


Figure 3.2(b) Criterion $|f(c)| < \varepsilon$ failure

误差分析:

根据二分法的算法, 迭代区间为$[a_0, b_0], [a_1, b_1], \cdots$. 有

$$a_0 \le a_1 \le a_2 \le \cdots \le b_0$$
$$b_0 \ge b_1 \ge b_2 \ge \cdots \ge a_0$$
$$b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n) \ (n \ge 0)$$

因此$\{a_n\}, \{b_n\}$收敛, 有

$$b_n - a_n = 2^{-n}(b_0 - a_0)$$
$$\lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = 0$$

对$0 \ge f(a_n)f(b_n)$取极限, 有$0 \ge [f(x^*)]^2$得到$f(x^*) = 0$.

二分法误差分析

当第n步二分法停止时, 最好的逼近是区间$[a_n, b_n]$的中点. 令$c_n = (b_n + a_n)/2$, 误差有

$$|x^* - c_n| \le \frac{b_n - a_n}{2} = 2^{-(n+1)}(b_0 - a_0)$$

二分法是线性收敛的

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|} = \frac{1}{2}.$$

综上可得到如下定理.

Theorem 2.1

设$[a_0, b_0], [a_1, b_1], \cdots, [a_n, b_n], \cdots$是二分法的迭代区间, 则$\lim_{n\to\infty} a_n$ 和$\lim_{n\to\infty} b_n$存在且相等, 是f 的一个零点. 如果$x^* = \lim_{n\to\infty} c_n$且$c_n = (b_n + a_n)/2$, 则

$$|x^* - c_n| \le 2^{-(n+1)}(b_0 - a_0)$$

Example 2.2

设二分法的开始区间是[50,63], 需要多少迭代步能使得相对误差是10^{-12}

$$\frac{|x^* - c_n|}{|x^*|} \leq 10^{-12}$$

知道$r \geq 50$, 收敛的充分条件是

$$\frac{|x^* - c_n|}{50} \leq 10^{-12}$$

由上面定理, 收敛的充分条件是

$$\frac{2^{-(n+1)} \times 13}{50} \leq 10^{-12}$$

因此, 需要$n \geq 37$.

Newton迭代法

设$f(x)$在其零点x^*附近充分光滑, x_k, x_{k+1}都在x^*附近. 由Taylor展开公式

$$f(x^*) = f(x_k) + f'(x_k)(x^* - x_k) + \frac{1}{2}f''(\xi_k)(x^* - x_k)^2$$

(ξ_k介于x_k与x^*之间), 当$|x^* - x_k|$很小时, 忽略右端最后一项高阶小量, 从而有

$$f(x_k) + f'(x_k)(x^* - x_k) \approx 0.$$

这样我们有$x^* \approx x_k - \frac{f(x_k)}{f'(x_k)}$, 即可以构造迭代序列

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \ k = 0, 1, 2, \cdots .$$

Newton迭代法

输入x_0, n, ϵ, τ

$y = f(x_0)$

如果$|y| < \tau$，则停止

对$k = 1$到n，进行下面操作

　$x_1 = x_0 - y/f'(x_0)$

　$y = f(x_1)$

　如果$|x_1 - x_0| < \epsilon$或$|y| < \tau$，则停止

结束

Newton法的几何解释

$$f(x) = f(c) + f'(c)(x-c) + \frac{1}{2}f''(c)(x-c)^2 + \cdots .$$

在c点对f线性化有

$$l(x) = f(c) + f'(c)(x-c)$$

则l是f在c处的切线, 与f有同样的函数值和斜率, Newton法是以切线的零点作为下一步近似零点, 参见图3.3. 此外, 初始点选取的不好有可能导致Newton迭代法失败, 参见图3.4. 所以Newton法收敛需要初始点x_0尽可能靠近零点或者f 的形状比较好.

Newton法的几何解释

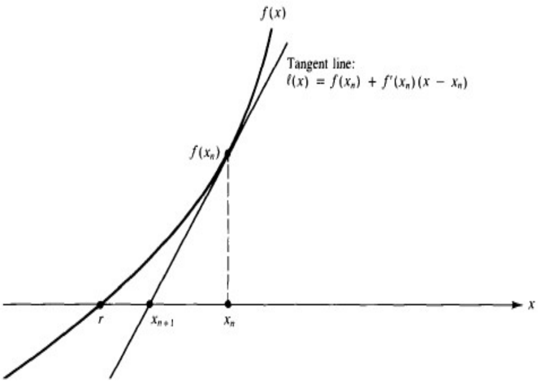


Figure 3.3 Geometric interpretation of Newton's method

Newton法的几何解释

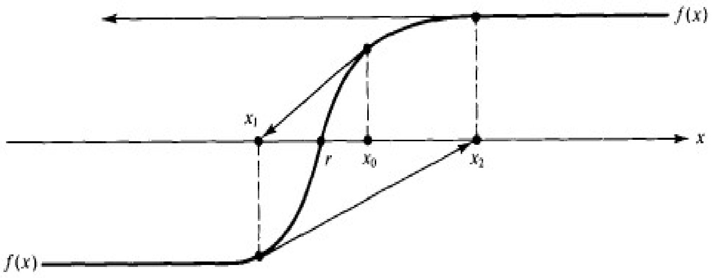


Figure 3.4 Example of nonconvergence of Newton's method

Newton法误差分析

误差分析:

设

$$e_n = x_n - x^*$$

假设f''连续, x^*是f的单重零点, 则$f(x^*) = 0 \neq f'(x^*)$. 根据Newton迭代, 有

$$\begin{aligned} e_{n+1} = x_{n+1} - x^* &= x_n - \frac{f(x_n)}{f'(x_n)} - x^* = e_n - \frac{f(x_n)}{f'(x_n)} \\ &= \frac{e_n f'(x_n) - f(x_n)}{f'(x_n)} = \frac{e_n f'(x_n) - (f(x_n) - f(x^*))}{f'(x_n)} \end{aligned}$$

由Taylor定理, 有

$$f(x_n)-f(x^*)=f'(x_n)(x_n-x^*)-\frac{1}{2}e_n^2f''(\xi_n)$$

$$e_{n+1}=\frac{1}{2}\frac{f''(\xi_n)}{f'(x_n)}e_n^2\approx\frac{f''(x^*)}{f'(x^*)}e_n^2=Ce_n^2 \tag{1}$$

因此, Newton迭代法是二阶收敛的. 关于收敛性的严格证明可参见《David Kincaid ., Ward Cheney, Numerical Analysis: Mathematics of Scientific Computing, Third Edition, 机械工业出版社》.

假设r是单重零点, 令$\Delta : |x - x^*| \le \delta$, 定义$c(\delta)$依赖于$\delta$

$$c(\delta) := \frac{1}{2} \max_{x \in \Delta} |f''(x)| / \min_{x \in \Delta} |f'(x)| \quad (\delta > 0) \tag{2}$$

因为当$\delta \to 0$, $c(\delta) \to |\frac{1}{2} f''(x^*)/f'(x^*)|$, 故$\delta c(\delta) \to 0$. 选择$\delta$ 使得(2)中的分母大于零, 且$\delta c(\delta) < 1$. 固定δ, 令$\rho = \delta c(\delta)$. 假设初值$x_0 \in \Delta$, 则$|e_0| \le \delta$和$|\xi_0 - x^*| \le \delta$. 因此有

$$\frac{1}{2} |f''(\xi_0)/f'(x_0)| \le c(\delta).$$

因此(1)表明

$$\begin{aligned} |x_1 - x^*| = |e_1| &\le e_0^2 c(\delta) \le |e_0| \delta c(\delta) = \rho |e_0| \\ &< |e_0| \le \delta. \end{aligned}$$

因此对于下一个迭代点有$x_1 \in \Delta$, 重复上述过程

$$|e_1| \le \rho|e_0|,\ |e_2| \le \rho|e_1| \le \rho^2|e_0|, \cdots$$

因而

$$x_n \in \Delta,\ |e_n| \le \rho^n|e_0|.$$

因为$0 \le \rho < 1$, 我们有$\lim\limits_{n\to\infty} \rho^n = 0$, 故$\lim\limits_{n\to\infty} e_n = 0$. 综上得到下面的定理.

Theorem 2.3

设f在其零点x^*附近二阶连续可微, x^*是单重零点, 则存在x^*的邻域$\Delta : |x - x^*| \le \delta$, 则对任意属于$\Delta$的初值, *Newton* 迭代法产生的序列$\{x_k\}$都收敛, 而且收敛速度是二阶的.

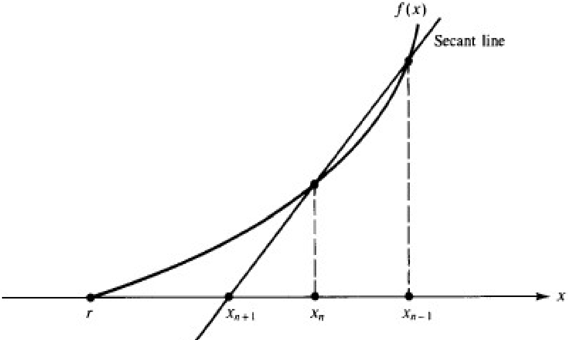
割线法

为了避免在Newton迭代法中每步迭代都要计算导函数的值，用相邻两步的函数值作差商来逼近导数值，即

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

从而迭代割线法

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n), \ n = 1, 2, \cdots .$$



割线法

输入a, b, n, ϵ, τ

$fa = f(a); fb = f(b)$

对$k = 2$到n, 进行下面操作

  如果$|fa| > |fb|$, 则$a \leftrightarrow b; fa \leftrightarrow fb$

  $s = (b - a)/(fb - fa)$

  $b = a$

  $fb = fa$

  $a = a - fa \cdot s$

  如果$|b - a| < \epsilon$或$|fa| < \tau$, 则停止

结束

在程序中有时交换端点a, b, 为保证$|f(a)| < |f(b)|$, 从而对点x_n, x_{n-1}有$|f(x_n)| \le |f(x_{n-1})|$, 接下来对点$x_{n+1}, x_n$有$|f(x_{n+1})| \le |f(x_n)|$. 这样确保点列从第二项开始是递减的, 所以数值稳定.

误差分析:

下面给出一个不严格的收敛性分析, 严格的证明可参见《数值分析-张平文、李铁军, 北京大学出版社》

割线法误差分析

$$
\begin{aligned}
e_{n+1} = x_{n+1} - x^* &= \frac{f(x_n)x_{n-1} - f(x_{n-1})x_n}{f(x_n) - f(x_{n-1})} - x^* \\
&= \frac{f(x_n)e_{n-1} - f(x_{n-1})e_n}{f(x_n) - f(x_{n-1})} \\
&= \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\left[\frac{f(x_n)/e_n - f(x_{n-1})/e_{n-1}}{x_n - x_{n-1}}\right]e_n e_{n-1}
\end{aligned}
$$

由Taylor展开, 有

$$
f(x_n) = f(x^* + e_n) = f(x^*) + e_n f'(x^*) + \frac{1}{2}e_n^2 f''(x^*) + O(e_n^3)
$$

$$
f(x_{n-1}) = f(x^*) + e_{n-1} f'(x^*) + \frac{1}{2}e_{n-1}^2 f''(x^*) + O(e_{n-1}^3)
$$

因为$x_n - x_{n-1} = e_n - e_{n-1}$, 得

$$
\frac{f(x_n)/e_n - f(x_{n-1})/e_{n-1}}{x_n - x_{n-1}} \approx \frac{1}{2} f''(x^*)
$$

$$\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \approx \frac{1}{f'(x^*)}$$

因此

$$e_{n+1} \approx \frac{f''(x^*)}{f'(x^*)} e_n e_{n-1} = C e_n e_{n-1}$$

若设$|e_{n+1}| \sim A|e_n|^\alpha$, 则

$$|e_n| \sim A|e_{n-1}|^\alpha \text{ 和} |e_{n-1}| \sim (A^{-1}|e_n|)^{1/\alpha}$$

$$A|e_n|^\alpha \sim |C||e_n|(A^{-1}|e_n|)^{1/\alpha}$$

可以写成

$$A^{1+1/\alpha}|C|^{-1} \sim |e_n|^{1-\alpha+1/\alpha}$$

割线法误差分析

当$e_n \to 1$时，上式左边是非零常数，因此$1-\alpha+1/\alpha=0$，即$\alpha=(1+\sqrt{5})/2 \approx 1.62$. 因此割线法是超线性收敛的.

与Newton迭代法相比，割线法不用计算导函数的值，计算量降低；其收敛速度比Newton迭代法稍慢.

不动点迭代法

$$把方程f(x)=0写成等价形式x=F(x),\ 于是问题转化为$$

$$求x^*,使得x^*=F(x^*).$$

称为函数$F(x)$的不动点问题, x^*称为函数$F(x)$的不动点.

为了逼近不动点问题, 我们可以采用如下迭代法:

$$\begin{cases} 给定初值x_0 \\ x_{k+1}=F(x_k),\ k=0,1,2,\cdots. \end{cases} \tag{3}$$

这种迭代法称为不动点迭代法, x_0为初值, 连续函数$F(x)$称为迭代函数.

不动点迭代法

Definition 2.4

设$C \subset \mathbb{R}$, 一个映射(或函数)$F: C \to C$称为压缩映射, 如果存在$\lambda < 1$, 使得

$$|F(x) - F(y)| \leq \lambda |x - y|$$

对任意$x, y \in C$均成立.

Theorem 2.5

设C是$\mathbb{R}$上的闭区间, $F: C \to C$是压缩映射, 则F在C上存在唯一的不动点. 且$\forall x_0 \in C$, 方程(3)产生的序列$\{x_k\}$都收敛于不动点.

Proof: 由压缩映射的性质

$$|x_k - x_{k-1}| = |F(x_{k-1}) - F(x_{k-2})| \leq \lambda |x_{k-1} - x_{k-2}|$$

重复上述过程有

$$|x_k - x_{k-1}| \leq \lambda |x_{k-1} - x_{k-2}| \leq \lambda^2 |x_{k-2} - x_{k-3}| \leq \cdots \leq \lambda^{k-1} |x_1 - x_0|$$

因此

$$|x_{n+k} - x_n| \leq \sum_{m=1}^{k} |x_{n+m} - x_{n+m-1}| \leq \lambda^n \frac{1-\lambda^k}{1-\lambda} |x_1 - x_0|$$

显然$\{x_n\}$是Cauchy列, 因为C是闭集, 因此$\{x_n\}$在C中有极限. 记$x^* = \lim_{k\to\infty} x_k$. 则$F(x^*) = x^*$(注意到, 压缩映射说明函数是连续的). 对于唯一性, 设x, y是不动点, 则

$$|x - y| = |F(x) - F(y)| \leq \lambda |x - y|$$

由于$\lambda < 1$, 因此$|x - y| = 0$, 不动点唯一. 证明完毕.

Theorem 2.6

设x^*是$F(x)$的不动点. 如果$F(x)$在x^*的某个邻域中是连续可微的, 而且$|F'(x^*)| < 1$, 则一定存在$\delta > 0$, 只要初值x_0 满足$|x_0 - x^*| < \delta$, 不动点迭代序列$\{x_k\}$就收敛于x^*.

Proof.

由$F'(x)$连续的性质可知, 存在x^*的邻域$\Delta : |x - x^*| \le \delta$, 使得$\forall x \in \Delta$, 有$|F'(x)| \le \lambda < 1$, 于是

$$|F(x) - x^*| = |F(x) - F(x^*)| \le \lambda |x - x^*| \le |x - x^*| \le \delta,$$

即$F(x)$可以看成是Δ上的压缩映射. 因此由定理2.5可知, 对任意初值$x_0 \in \Delta$不动点迭代收敛, 定理得证. □

不动点收敛速度:

设存在整数q, 使得

$$F^{(k)}(x^*) = 0,\ 1 \le k < q,\ F^{(q)}(x^*) \neq 0.$$

$$\begin{aligned} e_{n+1} &= x_{n+1} - x^* = F(x_n) - F(x^*) = F(x^* + e_n) - F(x^*) \\ &= e_n F'(x^*) + \frac{1}{2} e_n^2 F''(x^*) + \cdots + \frac{1}{(q-1)!} e_n^{q-1} F^{(q-1)}(x^*) + \frac{1}{q!} e_n^q F^{(q)}(\xi_n) \end{aligned}$$

不动点收敛速度

因此

$$e_{n+1} = \frac{1}{q!} e_n^q F^{(q)}(\xi_n)$$

因为$\lim_{n\to\infty} \xi_n = x^*$, 因此

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^q} = \frac{1}{q!} |F^{(q)}(x^*)|.$$

不动点迭代至少是q阶收敛的.

目录

考虑如下两变量的方程组

$$\begin{cases} f_1(x_1, x_2) = 0 \\ f_2(x_1, x_2) = 0 \end{cases}$$

假设$(x_1^{(0)}, x_2^{(0)})$离真解不远, 计算校正(h_1, h_2)使得$(x_1^{(1)}, x_2^{(1)}) = (x_1^{(0)}, x_2^{(0)}) + (h_1, h_2)$

$$\begin{cases} 0 = f_1(x_1^{(0)} + h_1, x_2^{(0)} + h_2) \approx f_1(x_1^{(0)}, x_2^{(0)}) + (h_1 \frac{\partial f_1}{\partial x_1} + h_2 \frac{\partial f_1}{\partial x_2})(x_1^{(0)}, x_2^{(0)}) \\ 0 = f_2(x_1^{(0)} + h_1, x_2^{(0)} + h_2) \approx f_2(x_1^{(0)}, x_2^{(0)}) + (h_1 \frac{\partial f_2}{\partial x_1} + h_2 \frac{\partial f_2}{\partial x_2})(x_1^{(0)}, x_2^{(0)}) \end{cases}$$

$$J\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = -\begin{bmatrix} f_1(x_1^{(0)}, x_2^{(0)}) \\ f_2(x_1^{(0)}, x_2^{(0)}) \end{bmatrix}$$

得

$$\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = -J^{-1}\begin{bmatrix} f_1(x_1^{(0)}, x_2^{(0)}) \\ f_2(x_1^{(0)}, x_2^{(0)}) \end{bmatrix}$$

因此, Newton迭代法如下

$$\begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} + \begin{bmatrix} h_1^{(k)} \\ h_2^{(k)} \end{bmatrix}$$

非线性方程组的迭代解法

对于n阶非线性方程组

$$f_i(x_1, \cdots, x_n) = 0,\ 1 \le i \le n$$

令$X = (x_1, \cdots, x_n)^T$, $F = (f_1, \cdots, f_n)^T$,

$$F(X) = 0$$

类似地有

$$0 = F(X + H) \approx F(X) + F'(X)H$$

其中$H = (h_1, \cdots, h_n)^T$, $F'(X)$是Jacobi矩阵

$$F'(X) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

非线性方程组的迭代解法

如果$F'(X)$可逆, 有

$$H = -F'(X)F(X)$$

多变量的Newton迭代格式

$$F'(X^{(k)})H^{(k)} = -F(X^{(k)})$$
$$X^{(k+1)} = X^{(k)} + H^{(k)}$$

目录

Continuation method

构造函数

$$h(t,x) = tf(x) + (1-t)[f(x) - f(x_0)] \quad t \in [0,1]$$

给定t, 得到$h(t,x)$的零点$x(t)$.

$$h(t,x(t)) = 0 \Rightarrow h_t + h_x x'(t) = 0 \Rightarrow x'(t) = -\frac{h_t}{h_x}$$

得到如下Cauchy初值问题

$$\begin{cases} x'(t) = -h_x^{-1} h_t \\ x(0) = x_0 \end{cases}$$

Continuation method

Example 4.1

$$f(x) = \begin{bmatrix} \xi_1^2 - 3\xi_2^2 + 3 \\ \xi_1\xi_2 + 6 \end{bmatrix} \quad x = (\xi_1, \xi_2) \in \mathbb{R}^2$$

$$h_x = f'(x) = \begin{bmatrix} 2\xi_1 & -6\xi_2 \\ \xi_2 & \xi_1 \end{bmatrix}, \quad h_t = f(x_0)$$

$$h_x^{-1} = \frac{1}{\Delta} \begin{bmatrix} \xi_1 & 6\xi_2 \\ -\xi_2 & 2\xi_1 \end{bmatrix} \quad \Delta = 2\xi_1^2 + 6\xi_2^2$$

目录

作业

1. 设f充分光滑, 已知根的重数k, 即

$$f(x)=f'(x)=\cdots=f^{(k-1)}(x)=0$$
$$f^{(k)}(x)\neq 0$$

证明

$$x_{n+1}=x_n-\frac{kf(x_n)}{f'(x_n)},\ f'(x_n)\neq 0$$

至少具有二阶局部收敛性.

2. 证明(Steffen迭代)

$$x_{n+1}=x_n-\frac{f(x_n)}{g(x_n)}$$

具有二阶收敛性. 其中

$$g(x)=\frac{f(x+f(x))-f(x)}{f(x)}$$

谢谢！