计算方法

令 丹

数学与统计学院

邮箱：danling@xjtu.edu.cn

第七章
非线性方程 (组) 的迭代解法

非线性方程

- 代数方程

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0, \quad a_n \neq 0$$

- 超越方程

$$3x^2 - \mathrm{e}^x = 0, \quad x + 2\sin x - \frac{1}{2} = 0$$

非线性方程：$n > 1$ 的代数方程、超越方程

非线性方程

设函数 $f(x)$ 可以分解为

$$f(x) = (x - x^*)^m g(x),$$

其中 m 为正整数, $g(x)$ 满足 $g(x^*) \neq 0$, 则称 x^* 为 $f(x)$ 的 m 重零点, 或称 x^* 为 $f(x) = 0$ 的 m 重根.

若函数 $f(x)$ 在区间 $[a, b]$ 上有 m 阶连续导数, 则 x^* 是 $f(x)$ 的 m 重零点等价于

$$f(x^*) = f'(x^*) = \cdots = f^{(m-1)}(x^*) = 0, f^{(m)}(x^*) \neq 0.$$

非线性方程

若函数 $f(x)$ 有多个零点, 则可根据已求出的零点 $x_1^*, x_2^*, \cdots,$ x_m^*, 通过函数

$$f_m(x) = \frac{f(x)}{\prod_{i=1}^{m}(x - x_i^*)},$$

求 $f(x)$ 的第 $m+1$ 个零点

非线性方程

若方程 $f(x)=0$ 在区间 $[a,b]$ 上至少有一个根,则称 $[a,b]$ 是方程的一个有根区间.

根的隔离：假定 $f(x)=0$ 在区间 $[a,b]$ 上存在多个根, 则把区间 $[a,b]$ 分成若干个子区间, 使得在每个子区间上有且只有一个根.

主要内容

1. 二分法

2. 简单迭代法（不动点法）

3. 牛顿类迭代法

4. 非线性方程组的迭代解法

1. 二分法

非线性方程

对于方程

$$f(x) = 0$$

如果在区间 $[a, b]$ 上至少有一个根, 就称 $[a, b]$ 是方程的一个有根区间.

根的隔离：假定 $f(x) = 0$ 在区间 $[a, b]$ 上存在多个根, 则把区间 $[a, b]$ 分成若干个子区间, 使得在每个子区间有且只有一个根.

零点存在定理

如果 $f(x)$ 在区间 $[a, b]$ 上连续, 且 $f(a)f(b) < 0$, 则在 (a, b) 内存在一点 x^*, 使得 $f(x^*) = 0$.

二分法

设方程 $f(x) = 0$ 的一个有根区间为 $[a, b]$, 且在区间 $[a, b]$ 只有一个根, 满足

$$f(a)f(b) < 0,$$

则可以用区间对分的方法形成有根区间序列 $\{[a_k, b_k]\}$.

二分法

设方程 $f(x)=0$ 的一个有根区间为 $[a,b]$, 且在区间 $[a,b]$ 只有一个根, 满足

$$f(a)f(b)<0,$$

则可以用区间对分的方法形成有根区间序列 $\{[a_k,b_k]\}$.

令 $[a_0,b_0]=[a,b]$, 对区间 $[a_k,b_k]$ 中点 $x_k=\frac{1}{2}(a_k+b_k)$, 判断 $f(x_k)$ 的符号. 若 $f(a_k)f(x_k)<0$, 则取新的有根区间为 $[a_k,x_k]$; 否则取为 $[x_k,b_k]$. 由此产生一系列含根区间

$$[a,b]\supset[a_1,b_1]\supset\cdots\supset[a_k,b_k]\supset[a_{k+1},b_{k+1}]\supset\cdots,$$

而每个区间的中点 $\{x_k\}$ 即为方程的根 x^* 的近似序列.

二分法

算法 (二分法)

给定 $\varepsilon > 0, \delta > 0$, $a_0 := a, b_0 := b, k := 0$.

(1) 令 $x_k := \frac{1}{2}(a_k + b_k)$.

(2) 若 $|f(x_k)| < \varepsilon$ 或 $b_k - a_k < \delta$, 则取 $x^* \approx x_k$, 停止;

(3) 若 $f(a_k)f(b_k) < 0$, 则 $a_{k+1} := a_k, b_{k+1} := x_k$;

否则取 $a_{k+1} := x_k, b_{k+1} := b_k$

(4) 令 $k := k + 1$, 转 (1).

误差估计

定理

设 $f(x) \in C[a,b], f(a)f(b) < 0$, 则由二分法产生的序列 $\{x_k\}$ 满足

$$\boxed{|x^* - x_k| \leqslant \frac{b-a}{2^{k+1}},}$$

其中 x^* 为方程 $f(x) = 0$ 的根.

证 由于 $[a_k, b_k]$ 是由 $[a_{k-1}, b_{k-1}]$ 一分为二得到, 所以

$$b_k - a_k = \frac{b_{k-1} - a_{k-1}}{2} = \cdots = \frac{b_0 - a_0}{2^k} = \frac{b-a}{2^k},$$

而 $x^* \in [a_k, b_k]$, 且 $x_k = (a_k + b_k)/2$, 因此

$$|x^* - x_k| = \left|x^* - \frac{a_k + b_k}{2}\right| \leqslant \frac{b_k - a_k}{2} = \frac{b-a}{2^{k+1}},$$

即有当 $k \to \infty$ 时, 有 $x_k \to x^*$.

二分法

例 1：已知 $f(x)=x^3+4x^2-10$ 在 $[1,2]$ 上有一个零点，且 $f(1)=-5, f(2)=14$，则由二分法计算结果如下:

k	$[a_k, b_k]$	x_k	$\operatorname{sgn}\big(f(a_k)f(x_k)\big)$
0	[1,2]	1.5	-1
1	[1.0,1.5]	1.25	$+1$
2	[1.25,1.5]	1.375	-1
3	[1.25,1.375]	1.3125	$+1$
4	[1.3125,1.375]	1.34375	$+1$
5	[1.34375,1.375]	1.359375	$+1$
6	[1.359375,1.375]	1.3671875	-1
7	[1.359375,1.3671875]	1.36328125	$+1$

二分法

当 $k=7$ 时, 得到 $x_7=1.36328125$, 由误差估计式有

$$\left|x^*-x_7\right| \leqslant 2^{-8}=3.90625\times 10^{-3}.$$

若要求 $\left|x^*-x_k\right|\leqslant \varepsilon$, 则可根据估计式确定必要的二分次数, 即有

$$\left|x^*-x_k\right| \leqslant \frac{b-a}{2^{k+1}}=2^{-(k+1)}\leqslant \varepsilon \quad \Longrightarrow \quad k \geqslant -\frac{\lg \varepsilon}{\lg 2}-1.$$

例如取 $\varepsilon=10^{-5}$, 则

$$k \geqslant \frac{5}{\lg 2}-1\approx 15.6096,$$

即至少需要计算到 $k=16$ 才能满足精度要求.

二分法的优缺点

Bisection method

- **优点**　方法简单、可靠且总收敛; 对函数要求不高, 只需连续即可.
- **缺点**　运算量较大, 需要在每一步计算函数值; 不能用于求复数根和偶数重根.

2. 简单迭代法（不动点法）

简单迭代法

Simpel Iterative Method

将方程 $f(x)=0$ 改写为

$$x=\phi(x),$$

若 x^* 满足 $f(x^*)=0$ 则满足 $x^*=\phi(x^*)$. 此时称 x^* 为 $\phi(x)$ 的不动点, 即映射关系 ϕ 将 x^* 映射到 x^* 自身.

设 $\phi(x)$ 连续, 构造迭代格式

$$\boldsymbol{x_{k+1}=\phi(x_k),\quad k=0,1,\cdots}$$

其中 $\phi(x)$ 称为**迭代函数**.

简单迭代法

给定初始近似值 x_0, 按迭代格式逐次迭代, 得到迭代序列 $\{x_k\}$. 若 $\lim\limits_{k\to\infty} x_k = x^*$, 则

$$x^* = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} \phi(x_k) = \phi\big(\lim_{k\to\infty} x_k\big) = \phi(x^*)$$

即 x^* 是 $\phi(x)$ 的不动点, 也即 $f(x)$ 的零点.

实际计算时, 若 $\big|f(x_{k+1})\big| < \varepsilon$ 或 $|x_{k+1} - x_k| < \delta$, 则可取 x_{k+1} 作为 x^* 的近似值.

简单迭代法

例 1：已知方程 $x^3+4x^2-10=0$ 在区间 $[1,2]$上有一个根，试用简单迭代法求之.

解法 1　$x^2=\frac{10}{x+4}, x=\sqrt{10/(x+4)}$，即 $\phi(x)=\sqrt{10/(x+4)}$.

解法 2　$4x^2=10-x^3, x=\frac{1}{2}\sqrt{10-x^3}$，即 $\phi(x)=\frac{1}{2}\sqrt{10-x^3}$.

解法 3　$x^3=10-4x^2, x=\sqrt[3]{10-4x^2}$，即 $\phi(x)=\sqrt[3]{10-4x^2}$.

解法 4　$4x=\frac{10}{x}-x^2, x=\frac{1}{4}\left(\frac{10}{x}-x^2\right)$，即 $\phi(x)=\frac{1}{4}\left(\frac{10}{x}-x^2\right)$.

取 $x_0=1.5$，用上述四种方法分别构造迭代格式进行计算，得结果如下：

简单迭代法

k	解法 1	解法 2	解法 3	解法 4
0	1.50000000	1.50000000	1.50000000	1.50000000
1	1.34839972	1.37371432	1.00000000	1.10416666
2	1.36737637	1.36412898	1.81712059	1.95935493
3	1.36495701	1.36536974	-1.4747949	0.31616218
4	1.36526474	1.36521223	1.09137019	7.88234428
5	1.36522559	1.36523227	1.73642773	-15.21567333
6	1.36523057	1.36522972	-1.27254559	-58.04348303
7	1.36522994	1.36523005	1.52154258	-842.30455175
8	1.36523002	1.36523000	0.90435447	-177369.242443923
9	1.36523001	1.36523001	1.88787962	-7864962041.28281
10	1.36523001	1.36523001	-1.62061323	$-1.54644069 \times 10^{19}$

简单迭代法

一般理论

(1) 如何选取合适的迭代函数 $\phi(x)$?

(2) 迭代函数 $\phi(x)$ 应满足什么条件, 迭代序列 $\{x_k\}$ 才能收敛?

(3) 如果 $\{x_k\}$ 收敛, 收敛速度如何? 怎样加速 $\{x_k\}$ 的收敛?

几何意义: x^* 是直线 $y = x$ 与曲线 $y = \phi(x)$ 的交点的横坐标.

折线法: 当 $|\phi'(x)| < 1$ 时收敛.

简单迭代法

定理 (不动点定理)

设 $\phi(x) \in C[a,b]$, 且满足

$$a \leqslant \phi(x) \leqslant b, \quad \forall x \in [a,b],$$

则 $\varphi(x)$ 在 $[a,b]$ 上一定有不动点. 进一步, 若存在常数 $L \in (0,1)$ 使得

$$\big|\phi(x) - \phi(y)\big| \leqslant L|x-y|, \quad \forall x,y \in [a,b], \qquad (**)$$

则 $\phi(x)$ 在 $[a,b]$ 上的不动点是唯一的.

条件 $(**)$ 称为 Lipschitz (利普希茨) 条件, L 称为 Lipschitz 常数.

不动点定理

证 由于 $\phi(x) \in C[a,b]$ 且当 $x \in [a,b]$ 时, 有 $a \leqslant \phi(x) \leqslant b$. 作辅助函数

$$g(x) = \phi(x) - x,$$

则有 $g(x) \in C[a,b]$ 且满足

$$g(a) = \phi(a) - a \geqslant 0, \quad g(b) = \phi(b) - b \leqslant 0.$$

当 $g(a) = 0$ 或 $g(b) = 0$ 成立时, 有 $\phi(a) = a$ 或 $\phi(b) = b$, 即 a 或 b 为 $\phi(x)$ 的不动点.

零点存在定理ß

当 $g(a) > 0, g(b) < 0$ 时, 根据连续函数的零点存在定理, 存在 $x^* \in (a,b)$ 使得 $g(x^*) = 0$, 即 $x^* = \phi(x^*)$, x^* 为 $\phi(x)$ 的不动点.

不动点定理

进一步, 假设 $\phi(x)$ 满足 Lipschitz 条件, 若 $\phi(x)$ 有两个不同的不动点 $x_1^*, x_2^* \in [a, b]$, 则有

$$\left|x_1^* - x_2^*\right| = \left|\phi(x_1^*) - \phi(x_2^*)\right| \leqslant L\left|x_1^* - x_2^*\right| < \left|x_1^* - x_2^*\right|,$$

产生矛盾, 故 $\phi(x)$ 的不动点是唯一的.

简单迭代法

推论

设 $\phi(x) \in C^1[a,b]$, 对一切的 $x \in [a,b]$ 成立 $a \leqslant \phi(x) \leqslant b$, 且存在常数 $L \in (0,1)$ 使得

$$\left|\phi'(x)\right| \leqslant L, \qquad \forall x \in [a,b],$$

则 $\phi(x)$ 在 $[a,b]$ 上存在唯一的不动点.

迭代法的收敛性

定义

设 x^* 为 方程 $f(x)=0$ 的根, $[a,b]$ 为含根区间, $\delta>0$,

$$N_\delta(x^*)=\left\{x\Big||x-x^*|\leqslant\delta\right\},$$

为 x^* 的一个邻域.

- 局部收敛 $\forall \boldsymbol{x_0}\in \boldsymbol{N_\delta(x^*)}$, $\lim\limits_{k\to\infty} x_k=x^*$.
- 全局收敛 $\forall \boldsymbol{x_0}\in [\boldsymbol{a},\boldsymbol{b}]$, $\lim\limits_{k\to\infty} x_k=x^*$.

简单迭代法的收敛性

定理 (迭代法全局收敛定理)

设迭代函数 $\phi(x) \in C^1[a,b]$ 且满足下列条件:

(1) 当 $x \in [a,b]$ 时, 有 $\phi(x) \in [a,b]$;

(2) 存在常数 L $(0 < L < 1)$, 使得 $\forall x \in [a,b], \left|\phi'(x)\right| \leqslant L < 1$.

则对任意的初始点 $x_0 \in [a,b]$, 由迭代格式 $x_{k+1} = \phi(x_k)$ 产生的迭代序列 $\{x_k\}$ 收敛于 $\phi(x)$ 在区间 $[a,b]$ 上的唯一不动点 x^*, 并且有如下误差估计式

$$\left|x^* - x_k\right| \leqslant \frac{1}{1-L}\left|x_{k+1} - x_k\right| \leqslant \frac{L}{1-L}\left|x_k - x_{k-1}\right|,$$

$$\left|x^* - x_k\right| \leqslant \frac{L^k}{1-L}\left|x_1 - x_0\right|.$$

简单迭代法的收敛性

证 由定理条件 (1) 和 (2), 可知迭代函数 $\phi(x)$ 在区间 $[a,b]$ 上存在唯一的不动点 x^*. 下证迭代序列 $\{x_k\}$ 收敛于该不动点.

由条件 (2) 可知

$$\begin{aligned}\left|x^*-x_k\right| &= \left|\phi(x^*)-\phi(x_{k-1})\right| = \left|\phi'(\xi_{k-1})(x^*-x_{k-1})\right| \\ &\leqslant L\left|x^*-x_{k-1}\right| \leqslant L^2\left|x^*-x_{k-2}\right| \leqslant \cdots \leqslant L^k\left|x^*-x_0\right|,\end{aligned}$$

由于 $0<L<1$, 令 $k\to\infty$, 两端取极限得

$$0 \leqslant \lim_{k\to\infty}\left|x^*-x_k\right| \leqslant \lim_{k\to\infty} L^k\left|x^*-x_0\right| = 0,$$

即有

$$\lim_{k\to\infty}\left|x^*-x_k\right| = 0 \quad \Longrightarrow \quad \lim_{k\to\infty} x_k = x^*.$$

简单迭代法的收敛性

接下来证误差估计式.

$$\begin{aligned}\left|x_{k+n}-x_k\right| &= \left|x_{k+n}-x_{k+n-1}+x_{k+n-1}-x_{k+n-2}+\cdots+x_{k+1}-x_k\right| \\ &\leqslant \sum_{j=1}^{n}\left|x_{k+j}-x_{k+j-1}\right|,\end{aligned}$$

而

$$\begin{aligned}\left|x_{k+j}-x_{k+j-1}\right| &= \left|\phi(x_{k+j-1})-\phi(x_{k+j-2})\right| \\ &= \left|\phi'(\xi)(x_{k+j-1}-x_{k+j-2})\right| \\ &\leqslant L\left|x_{k+j-1}-x_{k+j-2}\right| \\ &\leqslant \cdots \leqslant L^{j-1}\left|x_{k+1}-x_k\right|,\end{aligned}$$

简单迭代法的收敛性

因此有

$$\begin{aligned} |x_{k+n} - x_k| &= \sum_{j=1}^{n} |x_{k+j} - x_{k+j-1}| \\ &\leqslant \sum_{j=1}^{n} L^{j-1} |x_{k+1} - x_k| = \frac{1 - L^n}{1 - L} |x_{k+1} - x_k|. \end{aligned}$$

两端令 $n \to \infty$ 取极限, 注意到 $L \in (0, 1)$, 可得

$$|x^* - x_k| \leqslant \frac{1}{1 - L} |x_{k+1} - x_k| \leqslant \frac{L}{1 - L} |x_k - x_{k-1}|.$$

而根据 $|x_{k+1} - x_k| \leqslant L^k |x_1 - x_0|$, 则可得

$$|x^* - x_k| \leqslant \frac{L^k}{1 - L} |x_1 - x_0|.$$

简单迭代法的收敛性

(1) L 越小, $\{x_k\}$ 收敛越快.

(2) L 不是很接近 1 时, 若 $\left|x_{k+1}-x_k\right| \leqslant \varepsilon$, 则迭代终止, $x^* \approx x_{k+1}$.

(3) 对给定的精度 ε, 可用误差估计式预估迭代次数

$$\left|x^*-x_k\right| \leqslant \frac{L^k}{1-L}\left|x_1-x_0\right| \leqslant \varepsilon \quad \Longrightarrow \quad L^k \leqslant \frac{\varepsilon(1-L)}{\left|x_1-x_0\right|}$$

注意到 $L \in (0,1)$, 于是有

$$k \ln L \leqslant \ln\left(\frac{\varepsilon(1-L)}{\left|x_1-x_0\right|}\right) \quad \Longrightarrow \quad k \geqslant \ln\left(\frac{\varepsilon(1-L)}{\left|x_1-x_0\right|}\right) / \ln L.$$

简单迭代法的收敛性

收敛性：迭代函数的一阶导 < 1

定理 (迭代法局部收敛定理)

设迭代函数 $\phi(x)$ 在其不动点 x^* 的某个邻域 $[x^* - \delta, x^* + \delta]$ 内连续可微, 并且存在常数 L $(0 < L < 1)$, 使得

$$|\phi'(x)| \leqslant L < 1, \quad \forall x \in [x^* - \delta, x^* + \delta]$$

成立, 则对任意的初始点 $x_0 \in [x^* - \delta, x^* + \delta]$, 由迭代格式 $x_{k+1} = \phi(x_k)$ 产生的迭代序列 $\{x_k\}$ 收敛于 $\phi(x)$ 的唯一不动点 x^*.

迭代法的收敛速度 convergence velocity

定义

设迭代序列 $\{x_k\}$ 收敛于 x^*, 若存在常数 $p \geqslant 1$ 和 $c > 0$, 使得

$$\lim_{k\to\infty} \frac{\left|x^* - x_{k+1}\right|}{\left|x^* - x_k\right|^p} = c \neq 0,$$

则称迭代序列是 p 阶收敛的, c 称为渐近误差常数. 特别地,

- $p = 1\ (0 < c < 1)$ 线性收敛
- $p > 1$ 超线性收敛
- $p = 2$ 平方收敛或二阶收敛.

如果迭代法产生的迭代序列 $\{x_k\}$ 是 p 阶收敛的, 则称该迭代法具有 p 阶收敛速度.

迭代法的收敛速度

定理 (收敛阶定理)

设迭代函数 $\phi(x)$ 在其不动点 x^* 的邻域内具有连续的 p $(p>1)$ 阶导数, 则由迭代格式 $x_{k+1}=\phi(x_k)$ 产生的迭代序列 $\{x_k\}$ p 阶收敛的充要条件是

$$\phi'(x^*)=\phi''(x^*)=\cdots=\phi^{(p-1)}(x^*)=0, \quad \phi^{(p)}(x^*)\neq 0.$$

证 充分性. 根据定理条件, 对 $\phi(x_k)$ 在 x^* 处进行泰勒展开得到

$$\begin{aligned} x_{k+1}=\phi(x_k)&=\phi(x^*)+\phi'(x^*)(x_k-x^*)+\cdots \\ &\quad+\frac{1}{(p-1)!}\phi^{(p-1)}(x^*)(x^*-x_k)^{p-1}+\frac{1}{p!}\phi^{(p)}(\xi_k)(x^*-x_k)^p \end{aligned}$$

迭代法的收敛速度

$$
\begin{aligned}
x_{k+1} &= \phi(x^*) + \frac{1}{p!}\phi^{(p)}(\xi_k)(x^* - x_k)^p \\
&= x^* + \frac{1}{p!}\phi^{(p)}(\xi_k)(x^* - x_k)^p, \quad \text{其中 } \xi_k \text{ 介于 } x_k \text{ 与 } x^* \text{ 之间.}
\end{aligned}
$$

由此可得

$$
\left|x^* - x_{k+1}\right| = \frac{1}{p!}\left|\phi^{(p)}(\xi_k)(x^* - x_k)^p\right|,
$$

所以

$$
\lim_{k\to\infty} \frac{\left|x^* - x_{k+1}\right|}{\left|x^* - x_k\right|^p} = \lim_{k\to\infty} \frac{1}{p!}\left|\phi^{(p)}(\xi_k)\right| = \frac{1}{p!}\left|\phi^{(p)}(x^*)\right| \neq 0,
$$

即迭代序列 $\{x_k\}$ p 阶收敛.

迭代法的收敛速度

必要性. 反证法. 设迭代序列 $\{x_k\}$ p 阶收敛, 而条件

$$\phi'(x^*) = \phi''(x^*) = \cdots = \phi^{(p-1)}(x^*) = 0, \quad \phi^{(p)}(x^*) \neq 0.$$

不成立, 则必然存在最小正整数 $p_0 < p$, 使得

$$\phi^{(i)}(x^*) = 0 \ (i = 1, 2, \cdots, p_0 - 1), \qquad \phi^{(p_0)}(x^*) \neq 0.$$

则根据充分性的证明过程可知迭代序列 $\{x_k\}$ p_0 阶收敛, 产生矛盾. 故成立

$$\phi'(x^*) = \phi''(x^*) = \cdots = \phi^{(p-1)}(x^*) = 0, \quad \phi^{(p)}(x^*) \neq 0.$$

简单迭代法的收敛速度

当 $0 < \left|\phi'(x^*)\right| < 1$ 时, 简单迭代法线性收敛.

设 x^* 是 $\phi(x)$ 的不动点, 即有 $x^* = \phi(x^*)$,

$$x^* - x_{k+1} = \phi(x^*) - \phi(x_k) = \phi'(\xi_k)(x^* - x_k),$$

其中 ξ_k 介于 x^* 与 x_k 之间, 从而有

$$\lim_{k\to\infty} \frac{\left|x^* - x_{k+1}\right|}{\left|x^* - x_k\right|} = \lim_{k\to\infty} \left|\phi'(\xi_k)\right| = \left|\phi'(x^*)\right|.$$

简单迭代法难收敛?

加速收敛技术

1. 松弛加速法

对 $x=\phi(x)$ 做同解变形

$$x-\omega x=\phi(x)-\omega x \quad \Longrightarrow \quad x=\frac{\phi(x)-\omega x}{1-\omega}$$

其中 $\omega \neq 1$ 为松弛因子. 记

$$\psi(x)=\frac{\phi(x)-\omega x}{1-\omega},$$

则 $x=\psi(x)$. 由此得迭代格式

$$x_{k+1}=\psi(x_k)=\frac{\phi(x_k)-\omega x_k}{1-\omega}.$$

若要该迭代格式二阶收敛, 则需 $\psi'(x^*)=0$, 即需 $\omega=\phi'(x^*)$. 由于 x^* 未知, 故 ω 无法准确求解.

加速收敛技术

取 $\overline{x}$ 作为 x^* 的一个好的近似值, 令 $\omega = \phi'(\overline{x})$, 代入迭代格式得

$$\boxed{x_{k+1} = \psi(x_k) = \frac{\phi(x_k) - \phi'(\overline{x})x_k}{1 - \phi'(\overline{x})}}$$

由于

$$\psi'(x^*) = \frac{\phi'(x^*) - \phi'(\overline{x})}{1 - \phi'(\overline{x})} \quad \Longrightarrow \quad |\psi'(x^*)| << 1, \ |\psi'(x^*)| < |\phi'(x^*)|$$

因此大大提高了收敛速度.

2. Aitken (艾特肯) 加速法

设 $\{x_k\}$ 线性收敛于 x^*, 即有

$$\frac{x^* - x_{k+1}}{x^* - x_k} \approx \frac{x^* - x_k}{x^* - x_{k-1}}$$

由此解得

$$x^* \approx \frac{x_{k-1}x_{k+1} - x_k^2}{x_{k+1} - 2x_k + x_{k-1}} = x_{k+1} - \frac{(x_{k+1} - x_k)^2}{x_{k+1} - 2x_k + x_{k-1}},$$

取上式右端作为 x^* 的近似值, 即令

$$\overline{x}_{k+1} = x_{k+1} - \frac{(x_{k+1} - x_k)^2}{x_{k+1} - 2x_k + x_{k-1}}$$

则 $\overline{x}_{k+1}$ 比 x_{k+1} 更接近 x^*.

Steffensen (斯特芬森) 迭代法

$$\begin{cases} y_k = \phi(x_k), \\ z_k = \phi(y_k), \\ x_{k+1} = z_k - \dfrac{(z_k - y_k)^2}{z_k - 2y_k + x_k} \end{cases}$$

简单迭代法

例 2：对例 1 中的解法 1 和解法 2 进行改进，使收敛更快．对解法 3 和解法 4 进行改善使其收敛．

加速收敛技术

回顾前面的四种解法：

解法 1 $x^2=\frac{10}{x+4}, x=\sqrt{10/(x+4)}$，即 $\phi(x)=\sqrt{10/(x+4)}$.

解法 2 $4x^2=10-x^3, x=\frac{1}{2}\sqrt{10-x^3}$，即 $\phi(x)=\frac{1}{2}\sqrt{10-x^3}$.

解法 3 $x^3=10-4x^2, x=\sqrt[3]{10-4x^2}$，即 $\phi(x)=\sqrt[3]{10-4x^2}$.

解法 4 $4x=\frac{10}{x}-x^2, x=\frac{1}{4}\left(\frac{10}{x}-x^2\right)$，即 $\phi(x)=\frac{1}{4}\left(\frac{10}{x}-x^2\right)$.

简单迭代法

k	解法 1	解法 2	解法 3	解法 4
0	1.50000000	1.50000000	1.50000000	1.50000000
1	1.34839972	1.37371432	1.00000000	1.10416666
2	1.36737637	1.36412898	1.81712059	1.95935493
3	1.36495701	1.36536974	-1.4747949	0.31616218
4	1.36526474	1.36521223	1.09137019	7.88234428
5	1.36522559	1.36523227	1.73642773	-15.21567333
6	1.36523057	1.36522972	-1.27254559	-58.04348303
7	1.36522994	1.36523005	1.52154258	-842.30455175
8	1.36523002	1.36523000	0.90435447	-177369.242443923
9	1.36523001	1.36523001	1.88787962	-7864962041.28281
10	1.36523001	1.36523001	-1.62061323	$-1.54644069 \times 10^{19}$

简单迭代法

解法 1 和解法 2 收敛, 用松弛法进行加速. 取 $\overline{x} = \frac{3}{2}$.

改进解法 1 由于 $\phi(x) = \sqrt{10/(x+4)}$, 于是有

$$\omega = \phi'(\overline{x}) = -\frac{2\sqrt{5}}{11\sqrt{11}} \implies \psi(x) = \frac{\phi(x) - \omega x}{1-\omega}.$$

据此得到迭代格式

$$x_{k+1} = \psi(x_k) = \frac{\phi(x_k) - \omega x}{1-\omega}.$$

改进解法 2 由于 $\phi(x) = \frac{1}{2}\sqrt{10 - x^3}$,

$$\omega = \phi'(\overline{x}) = -\frac{27}{4\sqrt{106}} \implies \psi(x) = \frac{\phi(x) - \omega x}{1-\omega}.$$

由此得到迭代格式

$$x_{k+1} = \psi(x_k) = \frac{\phi(x_k) - \omega x}{1-\omega}.$$

改进解法 3 由于 $\phi(x) = \sqrt[3]{10 - 4x^2}$,

$$\phi'(x) = -\frac{8x}{3\sqrt[3]{(10 - 4x^2)^2}}, \quad \left|\phi'(x)\right| > 1$$

因此

$$\left|\phi(x) - x^*\right| = \left|\phi'(\xi)(x - x^*)\right| > \left|x - x^*\right|, \quad \forall x \in [1, 2]$$

即该迭代格式不收敛. 对此进行改善. 取 $\overline{x} = \frac{3}{2}$.

$$\omega = \phi'(\overline{x}) = -4 \implies \psi(x) = \frac{\phi(x) - \omega x}{1 - \omega}.$$

由此得到迭代格式

$$x_{k+1} = \psi(x_k) = \frac{\phi(x_k) - \omega x}{1 - \omega}.$$

改进解法 4　由于 $\phi(x)=\frac{1}{4}\left(\frac{10}{x}-x^2\right)$,

$$\phi'(x)=-\frac{1}{2}\left(\frac{5}{x^2}+x\right),\quad \left|\phi'(x)\right|\geqslant\frac{13}{8}>1$$

因此

$$\left|\phi(x)-x^*\right|=\left|\phi'(\xi)(x-x^*)\right|>\left|x-x^*\right|,\quad \forall x\in[1,2]$$

即迭代格式不收敛. 对此进行改善. 取 $\overline{x}=\frac{3}{2}$.

$$\omega=\phi'(\overline{x})=-\frac{67}{36}\quad\Longrightarrow\quad\psi(x)=\frac{\phi(x)-\omega x}{1-\omega}.$$

由此得到迭代格式

$$x_{k+1}=\psi(x_k)=\frac{\phi(x_k)-\omega x}{1-\omega}.$$

简单迭代法

用改进后的四种迭代格式重新计算，仍取 $x_0 = \frac{3}{2}$，结果如下：

k	解法 1	解法 2	解法 3	解法 4
0	1.50000000	1.50000000	1.50000000	1.50000000
1	1.36495391	1.36520603	1.40000000	1.36165048
2	1.36523115	1.36523011	1.37853216	1.36543699
3	1.36523000	1.36523001	1.37054669	1.36521824
4	1.36523001	1.36523001	1.36738626	1.36523068
5	1.36523001		1.36610938	1.36522997
6			1.36558943	1.36523001
7			1.36537705	1.36523001
$\vdots$			$\vdots$	
17			1.36523003	
18			1.36523002	
19			1.36523001	
20			1.36523001	

例 3：已知方程 $x^3-3x+1=0$ 在区间 $[1,2]$ 上有一个根. 给出一种简单迭代法的迭代格式, 讨论其收敛性. 若收敛, 对迭代格式进行改进, 使其加速收敛; 若不收敛, 对迭代格式进行改善, 使其收敛.

解法 1 $3x=x^3+1, x=\frac{1}{3}(x^3+1)$, 即 $\phi(x)=\frac{1}{3}(x^3+1)$. 由于

$$\phi'(x)=x^2 \quad\Longrightarrow\quad \left|\phi'(x)\right|\geqslant 1,\quad \forall x\in[1,2]$$

因此

$$\left|\phi(x)-x^*\right|=\left|\phi'(\xi)(x-x^*)\right|\geqslant\left|x-x^*\right|,\quad \forall x\in[1,2]$$

即迭代格式不收敛. 对此进行改善.

简单迭代法

取 $\overline{x} = \frac{3}{2}$.

$$\omega = \phi'(\overline{x}) = \frac{9}{4} \implies \psi(x) = \frac{\phi(x) - \omega x}{1 - \omega}.$$

由此得到迭代格式

$$x_{k+1} = \psi(x_k) = \frac{\phi(x_k) - \omega x}{1 - \omega}.$$

简单迭代法

解法 2 $x^3 = 3x - 1, x = \sqrt[3]{3x-1}$, 即 $\phi(x) = \sqrt[3]{3x-1}$.

$$\phi'(x) = (3x-1)^{-\frac{2}{3}}, \quad \phi'(x) \in [\frac{1}{\sqrt[3]{25}}, \frac{1}{\sqrt[3]{4}}].$$

于是

$$\left|\phi'(x)\right| \leqslant \frac{1}{\sqrt[3]{4}} \approx 0.629961 < 1$$

因此迭代法收敛. 可以利用松弛法进行加速. 取 $\overline{x} = \frac{3}{2}$.

$$\omega = \phi'(\overline{x}) = \left(\frac{2}{7}\right)^{\frac{2}{3}} \implies \psi(x) = \frac{\phi(x) - \omega x}{1-\omega}.$$

由此得到改进后迭代格式 $x_{k+1} = \psi(x_k)$. 由于

$$\left|\psi'(x)\right| \leqslant \frac{\frac{1}{\sqrt[3]{4}} - \omega}{1-\omega} \approx 0.346453$$

因此收敛加速.

简单迭代法

例 4：利用迭代法的思想证明

$$\lim_{k\to\infty} \underbrace{\sqrt{2+\sqrt{2+\cdots+\sqrt{2+\sqrt{2+\sqrt{2}}}}}}_{\text{开方 } k \text{ 次}} = 2$$

解 令

$$x_{k+1} = \underbrace{\sqrt{2+\sqrt{2+\cdots+\sqrt{2+\sqrt{2+\sqrt{2}}}}}}_{\text{开方 } k \text{ 次}}$$

则 x_{k+1} 可看作是初值为 $x_0 = \sqrt{2}$ 的迭代格式 $x_{k+1} = \sqrt{2+x_k}$ 生成的迭代序列，其中迭代函数为 $\phi(x) = \sqrt{2+x}$.

简单迭代法

在区间 $[1,3]$ 上考虑该迭代格式的收敛性. 由于

$$\phi'(x)=\frac{1}{2\sqrt{2+x}} \quad\Longrightarrow\quad \left|\phi'(x)\right|<1,\ \forall x\in[1,2]$$

因此以 $\phi(x)$ 为迭代函数的简单迭代法 $x_{k+1}=\phi(x_k)$ 在区间 $[1,3]$ 上收敛, 且收敛于的 $\phi(x)$ 的不动点 x^*. 易知 $\phi(2)=2$, 即 $x^*=2$,

Ø(x) = x

$$\lim_{k\to\infty}\underbrace{\sqrt{2+\sqrt{2+\cdots+\sqrt{2+\sqrt{2+\sqrt{2}}}}}}_{\text{开方 } k \text{ 次}}=x^*=2.$$

3. 牛顿类迭代法

牛顿迭代法

设函数 $f(x)$ 在区间 $[a,b]$ 上二阶连续可微, 且 $f'(x)\neq 0$, x_k 是方程 $f(x)=0$ 的根 x^* 的一个近似值. 于是有泰勒展开

$$0=f(x^*)=f(x_k)+f'(x_k)(x^*-x_k)+\frac{1}{2}f''(\xi_k)(x^*-x_k)^2.$$

舍去二次项, 得近似线性方程

$$f(x_k)+f'(x_k)(x^*-x_k)\approx 0 \quad\Longrightarrow\quad x^*\approx x_k-\frac{f(x_k)}{f'(x_k)}.$$

牛顿法

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

牛顿法

几何意义　x_{k+1} 是 $y = f(x)$ 在点 (x_k, y_k) 处的切线

$$y = f(x_k) + f'(x_k)(x - x_k)$$

与 x 轴的交点.

切线法　若 x_0 充分靠近 x^*, 则 x_k 收敛于 x^*.

改进牛顿法 1

设 $f(x)$ 在区间 $[a,b]$ 上三阶连续可微, 且 $f'(x)\neq 0, f''(x)\neq 0$, x_k 是方程 $f(x)=0$ 的根 x^* 的一个近似值, 则

$$0=f(x^*)=f(x_k)+f'(x_k)(x^*-x_k)+\frac{1}{2!}f''(x_k)(x^*-x_k)^2 \\ +\frac{1}{3!}f'''(\xi_k)(x^*-x_k)^3.$$

舍去三次项, 得近似二次方程

$$f''(x^*-x_k)^2+2f'(x_k)(x^*-x_k)+2f(x_k)\approx 0,$$

解得

$$x^*\approx x_k+\frac{-f'(x_k)\pm\sqrt{f'^2(x_k)-2f(x_k)f''(x_k)}}{f''(x_k)}.$$

改进牛顿法 1

记

$$\widetilde{x}_{k+1} = x_k - \frac{f'(x_k) + \mathsf{sgn}\big(f'(x_k)\big)\sqrt{f'^2(x_k) - 2f(x_k)f''(x_k)}}{f''(x_k)},$$
$$\overline{x}_{k+1} = x_k - \frac{2f(x_k)}{f'(x_k) + \mathsf{sgn}\big(f'(x_k)\big)\sqrt{f'^2(x_k) - 2f(x_k)f''(x_k)}}.$$

取 $\widetilde{x}_{k+1}, \overline{x}_{k+1}$ 中靠近 x_k 的作为 x_{k+1}.

改进牛顿法 2

设 $y=f(x)$ 的反函数 $x=g(y)$ 三阶连续可微, 将 $g(y)$ 在 y_k 处展开

$$g(y)=g(y_k)+g'(y_k)(y-y_k)+\frac{1}{2!}g''(y_k)(y-y_k)^2 \\ +\frac{1}{3!}g'''(\eta_k)(y-y_k)^3.$$

由 $f(x^*)=0$ 知 $x^*=g(0)$, 则

$$x^*=g(0)=g(y_k)-g'(y_k)y_k+\frac{1}{2!}g''(y_k)y_k^2-\frac{1}{3!}g'''(\eta_k)y_k^3.$$

取

$$x^*=g(0)\approx g(y_k)-g'(y_k)y_k+\frac{1}{2}g''(y_k)y_k^2,$$

改进牛顿法 2

由于 $y_k = f(x_k)$, $x_k = g(y_k)$, $g'(y_k) = 1/f'(x_k)$ 及

$$
\begin{aligned}
g''(y_k) &= \left.\frac{dg'(y)}{dy}\right|_{y=y_k} = \left.\frac{d}{dy}\left(\frac{1}{f'(g(y))}\right)\right|_{y=y_k} \\
&= \left.\frac{-f''(g(y))}{f'^2(g(y))}\cdot g'(y)\right|_{y=y_k} = \left.\frac{-f''(g(y))}{f'^3(g(y))}\right|_{y=y_k} \\
&= \frac{-f''(g(y_k))}{f'^3(g(y_k))} = -\frac{f''(x_k)}{f'^3(x_k)}
\end{aligned}
$$

$$
\Longrightarrow \quad x^* = g(0) \approx x_k - \frac{f(x_k)}{f'(x_k)} - \frac{f''(x_k)}{2\big[f'(x_k)\big]^3} f^2(x_k)
$$

$$
\boxed{x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} - \frac{f''(x_k)}{2\big[f'(x_k)\big]^3} f^2(x_k), \quad k = 0, 1, 2, \cdots}
$$

简化牛顿法

实际计算中，$f'(x)$ 难以计算或计算量较大，可用 $f'(x_0)$ 或常数 c 代替牛顿法中的 $f'(x_k)$，得到

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)}, \qquad k = 0, 1, 2, \cdots$$

$$\text{或} \quad x_{k+1} = x_k - \frac{f(x_k)}{c}, \qquad k = 0, 1, \cdots$$

几何意义　用过点 $(x_k, f(x_k))$ 且斜率为 $f'(x_0)$ 或常数 c 的平行线与 x 轴的交点作为 x^* 的新近似点.

牛顿下山法

为了放宽 x_0 的选择范围, 可将牛顿法修改为

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

其中 λ 称为下山因子. λ 的选取应满足

$$\left|f(x_{k+1})\right| < \left|f(x_k)\right|, \qquad k = 0, 1, 2, \cdots$$

实际计算中 λ 采用试算的办法，即从 $\lambda = 1$ 开始, 逐次减半进行计算, 直至上式成立.

弦割法

在牛顿迭代法中, 用曲线上的两点连线的斜率来代替 $y = f(x)$ 在点 $\big(x_k, f(x_k)\big)$ 处的切线斜率, 即

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \quad \text{或} \quad f'(x_k) \approx \frac{f(x_k) - f(x_0)}{x_k - x_0}.$$

- 两点弦割法

$$x_{k+1} = x_k - \frac{f(x_k)/f(x_{k-1})}{f(x_k)/f(x_{k-1}) - 1}(x_k - x_{k-1}), \quad k = 1, 2, \cdots$$

- 单点弦割法

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_0)}{f(x_k) - f(x_0)}, \quad k = 1, 2, \cdots$$

弦割法

- **几何意义** 取过两点 $\big(x_{k-1}, f(x_{k-1})\big), \big(x_k, f(x_k)\big)$ 或 $\big(x_0, f(x_0)\big), (x_k, f(x_k))$ 的直线与 x 轴的交点作为 x^* 的近似值 x_{k+1}.
- **优缺点**
  ① 无需计算导数 $f'(x_k)$
  ② 由于 $f(x_{k-1}), f(x_k)$ 相近, 需高精度计算

改进弦割法

设 $y=f(x)$ 的反函数为 $x=g(y)$, 对 $g(y)$ 作二次牛顿插值 $N_2(y)$,

$$
\begin{aligned}
x=g(y)\approx N_2(y)=g(y_x)&+g[y_k,y_{k-1}](y-y_k)\\
&+g[y_k,y_{k-1},y_{k-2}](y-y_k)(y-y_{k-1}).
\end{aligned}
$$

取 $y=0$, 则得

$$
x^*=g(0)\approx g(y_k)-g[y_k,y_{k-1}]y_k+g[y_k,y_{k-1},y_{k-2}]y_ky_{k-1}.
$$

改进弦割法

由于

$$g[y_k, y_{k-1}] = \frac{g(y_k) - g(y_{k-1})}{y_k - y_{k-1}},$$

$$g[y_k, y_{k-1}, y_{k-2}] = \frac{g[y_k, y_{k-1}] - g[y_{k-1}, y_{k-2}]}{y_k - y_{k-2}}$$

$$= \left(\frac{g(y_k) - g(y_{k-1})}{y_k - y_{k-1}} - \frac{g(y_{k-1}) - g(y_{k-2})}{y_{k-1} - y_{k-2}} \right) / (y_k - y_{k-2})$$

改进弦割法

由于

$$g[y_k, y_{k-1}] = \frac{g(y_k) - g(y_{k-1})}{y_k - y_{k-1}},$$

$$g[y_k, y_{k-1}, y_{k-2}] = \frac{g[y_k, y_{k-1}] - g[y_{k-1}, y_{k-2}]}{y_k - y_{k-2}}$$

$$= \left(\frac{g(y_k) - g(y_{k-1})}{y_k - y_{k-1}} - \frac{g(y_{k-1}) - g(y_{k-2})}{y_{k-1} - y_{k-2}} \right) / (y_k - y_{k-2})$$

因此

$$x^* \approx g(y_x) - g[y_k, y_{k-1}]y_k + g[y_k, y_{k-1}, y_{k-2}]y_k y_{k-1}$$

$$= x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k)$$

$$+ \left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} - \frac{x_{k-1} - x_{k-2}}{f(x_{k-1}) - f(x_{k-2})} \right) \frac{f(x_k) f(x_{k-1})}{f(x_k) - f(x_{k-2})}$$

改进弦割法

给定 x_0, x_1, x_2,

$$
\begin{aligned}
x_{k+1} = x_k &- \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k) \\
&+ \left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} - \frac{x_{k-1} - x_{k-2}}{f(x_{k-1}) - f(x_{k-2})} \right) \frac{f(x_k) f(x_{k-1})}{f(x_k) - f(x_{k-2})}, \\
& k = 2, 3, \cdots
\end{aligned}
$$

牛顿法的收敛性

牛顿法相当于迭代函数 $\phi(x) = x - \frac{f(x)}{f'(x)}$ 的简单迭代法.

设 $f(x)$ 在其零点 x^* 附近二阶连续可微, 且 $f'(x^*) \neq 0$,

$$\phi'(x) = 1 - \frac{f'^2(x) - f(x)f''(x)}{f'^2(x)} = \frac{f(x)f''(x)}{f'^2(x)} \implies \phi'(x^*) = 0.$$

当 x 充分靠近 x^* (x 属于 x^* 的某邻域) 时, 存在常数 $L \in (0,1)$ 使得 $\left|\phi'(x)\right| \leqslant L < 1$.

定理 (牛顿法的局部收敛性)

设 $f(x)$ 在其零点 x^* 附近二阶连续可微且 $f'(x^*) \neq 0$, 则当初始点 x_0 充分靠近 x^*时, 由牛顿法产生的迭代序列 $\{x_k\}$ 收敛于 x^*.

牛顿法的收敛性

定理 (牛顿法的全局收敛性)

设 $f(x)$ 在 $[a,b]$ 上二阶连续可微, 并且满足

(1) $f(a)f(b) < 0$;

(2) $\forall x \in [a,b],\ f'(x) \neq 0$;

根的存在性

(3) $\forall x \in [a,b]$, $f''(x)$ 不变号;

(4) 初始点 $x_0 \in [a,b]$, 且 $f(x_0)f''(x_0) > 0$;

则由牛顿法产生的迭代序列 $\{x_k\}$ 单调地收敛于 $f(x)=0$ 在 $[a,b]$ 上的唯一根 x^*.

证 条件 (1) 和 (2) 保证了根的存在唯一性. 下证收敛性.

不妨设 $f'(x) > 0, f''(x) > 0$. 根据条件 (4) 有

$$0 < f(x_0) = f(x_0) - f(x^*) = f'(\xi)(x_0 - x^*) \implies x_0 - x^* > 0,$$

即 $x^* < x_0$, 从而有

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \implies x_1 - x_0 = -\frac{f(x_0)}{f'(x_0)} < 0 \implies x_1 < x_0.$$

牛顿法的收敛性

另一方面,

$$
\begin{aligned}
x_1 - x^* &= x_0 - \frac{f(x_0)}{f'(x_0)} - x^* = x_0 - x^* - \frac{f(x_0) - f(x^*)}{f'(x_0)} \\
&= x_0 - x^* - \frac{f'(\xi)(x_0 - x^*)}{f'(x_0)} = \frac{f'(x_0) - f'(\xi)}{f'(x_0)}(x_0 - x^*) \\
&= \frac{f''(\eta)}{f'(x_0)}(x_0 - \xi)(x_0 - x^*) > 0,
\end{aligned}
$$

即有 $x^* < x_1 < x_0$. 同理可证 $x^* < x_2 < x_1$. 依次类推, 得到

$$
x^* < x_k < x_{k-1} < \cdots < x_2 < x_1 < x_0
$$

即 $\{x_k\}$ 为单调下降有界数列, 因此收敛.

牛顿法的收敛性

设收敛于 a, 由于

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

根据 $f(x), f'(x)$ 的连续性, 两边令 $k \to \infty$ 取极限, 得

$$a = a - \frac{f(a)}{f'(a)} \implies f(a) = 0 \implies a = x^* \implies \lim_{k\to\infty} x_k = x^*.$$

简化牛顿法的收敛性 看作 $\phi(x) = x - f(x)/c$ 的简单迭代法, 若

$$\left|\phi'(x)\right| = \left|1 - \frac{f'(x)}{c}\right| < 1 \implies 0 < \frac{f'(x)}{c} < 2$$

在 x^* 附近成立, 则简化牛顿法局部收敛.

例 1：用牛顿法求 $\sqrt{7}$ 的近似值.

解　由于 $\sqrt{7}$ 满足方程 $x^2-7=0$. 可设 $f(x)=x^2-7$. 根据牛顿法构造迭代格式

$$x_{k+1}=x_k-\frac{f(x_k)}{f'(x_k)}=\frac{1}{2}\left(x_k+\frac{7}{x_k}\right).$$

同时可选取迭代区间为 $[2,3]$. 由于

$$\phi(x)=\frac{1}{2}\left(x+\frac{7}{x}\right)\implies\left|\phi'(x)\right|=\frac{1}{2}\left|1-\frac{7}{x^2}\right|\leqslant\frac{3}{4}<1,\ \forall x\in[2,3]$$

因此迭代法收敛.

牛顿法

选取 $x_0 = 2.5$, 根据牛顿法计算结果如下:

k	x_k
0	2.50000000000000
1	2.65000000000000
2	2.64575471698113
3	2.64575131106678
4	2.64575131106459
5	2.64575131106459

牛顿法的应用举例

对于任意给定的正数 C, 导出 $\sqrt{C}$ 的计算公式

$$x_{k+1} = \frac{1}{2}\left(x_k + \frac{C}{x_k}\right).$$

现证这种迭代公式对于任意初始值 $x_0 > 0$ 都是收敛的.

$$x_{k+1} - \sqrt{C} = \frac{1}{2x_k}\left(x_k - \sqrt{C}\right)^2,$$

$$x_{k+1} + \sqrt{C} = \frac{1}{2x_k}\left(x_k + \sqrt{C}\right)^2,$$

两式相除得到

$$\frac{x_{k+1} - \sqrt{C}}{x_{k+1} + \sqrt{C}} = \left(\frac{x_k - \sqrt{C}}{x_k + \sqrt{C}}\right)^2.$$

牛顿法的应用举例

依次递推有

$$\frac{x_k - \sqrt{C}}{x_k + \sqrt{C}} = \left(\frac{x_0 - \sqrt{C}}{x_0 + \sqrt{C}}\right)^{2^k} \triangleq q^{2^k}.$$

由此得到

$$x_k - \sqrt{C} = 2\sqrt{C}\frac{q^{2^k}}{1 - q^{2^k}}.$$

对于任意的 $x_0 > 0$, 总有 $|q| < 1$. 因此当 $k \to \infty$ 时有 $x_k \to \sqrt{C}$,
即迭代过程总收敛.

弦割法的收敛性

定理 (弦割法的局部收敛性)

设 $f(x)$ 在包含方程 $f(x)=0$ 的根 x^* 的某个邻域上二阶连续可微, 且 $f'(x^*)\neq 0$, 初始点 x_0, x_1 充分接近 x^*, 则由弦割法迭代格式

$$x_{k+1}=x_k-\frac{x_k-x_{k-1}}{f(x_k)-f(x_{k-1})}f(x_k)$$

产生的迭代序列 $\{x_k\}$ 收敛于 x^*.

证 设函数 $y=f(x)$ 的反函数为 $x=g(y)$, 由牛顿插值公式有

$$x=g(y)=g(y_k)+g[y_k,y_{k-1}](y-y_k)+\frac{1}{2}g''(\eta_k)(y-y_k)(y-y_{k-1})$$

弦割法的收敛性

取 $y=0$, 则得

$$
\begin{aligned}
x^* = g(0) &= g(y_k) - g[y_k, y_{k-1}]y_k + \frac{1}{2}g''(\eta_k)y_k y_{k-1} \\
&= x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}f(x_k) - \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)}f(x_k)f(x_{k-1}) \\
&= x_{k+1} - \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)}f(x_k)f(x_{k-1}) \\
&= x_{k+1} - \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)}\big(f(x_k) - f(x^*)\big)\big(f(x_{k-1}) - f(x^*)\big) \\
&= x_{k+1} - \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)}f'(\xi_k^*)(x_k - x^*)f'(\xi_{k-1}^*)(x_{k-1} - x^*).
\end{aligned}
$$

弦割法的收敛性

因此

$$\boxed{x_{k+1} - x^* = \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)} f'(\xi_k^*)(x_k - x^*) f'(\xi_{k-1}^*)(x_{k-1} - x^*)}$$

当 x_{k-1}, x_k 充分靠近 x^* 时, 即 x_{k-1}, x_k 属于 x^* 的某个邻域 $N_\delta = [x^* - \delta, x^* + \delta]$, 可以选取合适的 $\delta > 0$ 使得

$$\left|x_{k+1} - x^*\right| < \left|x_k - x^*\right|,$$

即 $\{|x_k - x^*|\}$ 是单调下降数列且有下界, 因此迭代序列 $\{x_k\}$ 收敛且收敛于 x^*.

弦割法的收敛性

定理 (弦割法的全局收敛性)

设 $f(x)$ 满足牛顿法全局收敛性定理的条件, 取 $x_1 \in [a,b]$, 使得 $f(x_1)f''(x_1) < 0$, 则由弦割法迭代格式

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k)$$

产生的迭代序列 $\{x_k\}$ 收敛于方程 $f(x) = 0$ 在 $[a,b]$ 上的唯一根 x^*.

牛顿法的收敛速度

设 $f(x)$ 在含根区间 $[a,b]$ 上二阶连续可微且 $f'(x)\neq 0$, x_k 是 $f(x)=0$ 的根 x^* 的近似值, 由 Taylor 展开有

$$0=f(x^*)=f(x_k)+f'(x_k)(x^*-x_k)+\frac{1}{2}f''(\xi_k)(x^*-x_k)^2,$$

因此有

$$\begin{aligned}0&=x^*-x_k+\frac{f(x_k)}{f'(x_k)}+\frac{1}{2}\frac{f''(\xi_k)}{f'(x_k)}(x^*-x_k)^2\\&=x^*-x_{k+1}+\frac{1}{2}\frac{f''(\xi_k)}{f'(x_k)}(x^*-x_k)^2,\end{aligned}$$

$$\Longrightarrow\quad \lim_{k\to\infty}\frac{\left|x^*-x_{k+1}\right|}{\left|x^*-x_k\right|^2}=\lim_{k\to\infty}\frac{1}{2}\frac{\left|f''(\xi_k)\right|}{\left|f'(x_k)\right|}=\frac{1}{2}\frac{\left|f''(x^*)\right|}{\left|f'(x^*)\right|}\neq 0$$

即牛顿法二阶收敛.

两点弦割法的收敛速度

由弦割法的局部收敛性定理分析过程可知,

$$x_{k+1} - x^* = \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)} f'(\xi_k^*)(x_k - x^*) f'(\xi_{k-1}^*)(x_{k-1} - x^*),$$

于是

$$\frac{\left|x^* - x_{k+1}\right|}{\left|x^* - x_k\right|^p} = \frac{1}{2}\left|\frac{f''(\xi_k) f'(\xi_k^*) f'(\xi_{k-1}^*)}{f'^3(\xi_k)}\right| \cdot \left(\frac{\left|x^* - x_k\right|}{\left|x^* - x_{k-1}\right|^p}\right)^{1-p} \cdot \left|x^* - x_{k-1}\right|^{1+p(1-p)}$$

两点弦割法的收敛速度

两边令 $k \to \infty$, 取极限得

$$\lim_{k\to\infty} \frac{\left|x^* - x_{k+1}\right|}{\left|x^* - x_k\right|^p} = c = c^{1-p} \cdot \frac{1}{2}\left|\frac{f''(x^*)}{f'(x^*)}\right| \cdot \lim_{k\to\infty} \left|x^* - x_{k-1}\right|^{1+p(1-p)}$$

于是有

$$1 + p - p^2 = 0 \quad \Longrightarrow \quad p = \frac{1+\sqrt{5}}{2} \in (1, 2)$$

即两点弦割法超线性收敛.

单点弦割法的收敛速度

可看作简单迭代法, 迭代函数为

$$\phi(x) = x - \frac{f(x)}{f(x) - f(x_0)}(x - x_0),$$

因此有

$$\phi'(x) = 1 + \frac{f'(x)f(x_0)(x - x_0) - f(x)\big(f(x) - f(x_0)\big)}{\big(f(x) - f(x_0)\big)^2},$$

$$\phi'(x^*) = 1 + \frac{f'(x^*)f(x_0)(x^* - x_0)}{\big(f(x^*) - f(x_0)\big)^2} = 1 - \frac{f'(x^*)}{f'(\xi)},$$

其中 ξ 介于 x_0 与 x^* 之间.

单点弦割法的收敛速度

$$
\begin{aligned}
x_{k+1} = \phi(x_k) &= \phi(x^*) + \phi'(x^*)(x_k - x^*) + \frac{1}{2}\phi''(\xi_k^*)(x^* - x_k)^2 \\
&= x^* + \phi'(x^*)(x_k - x^*) + \frac{1}{2}\phi''(\xi_k^*)(x^* - x_k)^2,
\end{aligned}
$$

$$
\Longrightarrow \quad \left|\frac{x^* - x_{k+1}}{x^* - x_k}\right| = \left|\phi'(x^*) + \frac{1}{2}\phi''(\xi_k^*)(x^* - x_k)\right|
$$

两边令 $k \to \infty$, 取极限得

$$
\begin{aligned}
\lim_{k\to\infty}\left|\frac{x^* - x_{k+1}}{x^* - x_k}\right| &= \lim_{k\to\infty}\left|\phi'(x^*) + \frac{1}{2}\phi''(\xi_k^*)(x^* - x_k)\right| \\
&= \left|\phi'(x^*)\right| = \left|1 - \frac{f'(x^*)}{f'(\xi)}\right|.
\end{aligned}
$$

即单点弦割法在一般情况下线性收敛.

牛顿类迭代法

例 2：用牛顿法和弦割法求解方程 $x = \cos x$, 初值取 $x_0 = 0.5$, $x_1 = \frac{\pi}{4}$.

解 由于 $f(x) = x - \cos x$, 所以 $f'(x) = 1 + \sin x$, 牛顿法的迭代格式为

$$x_{k+1} = x_k - \frac{x_k - \cos x_k}{1 + \sin x_k}.$$

两点弦割法和单点弦割法的迭代公式分别为

$$\left.\begin{aligned} x_{k+1} &= x_k - \frac{(x_k - x_{k-1})(x_k - \cos x_k)}{(x_k - \cos x_k) - (x_{k-1} - \cos x_{k-1})}, \\ x_{k+1} &= x_k - \frac{(x_k - x_0)(x_k - \cos x_k)}{(x_k - \cos x_k) - (x_0 - \cos x_0)}, \end{aligned}\right. \quad k = 1, 2, \cdots$$

牛顿类迭代法

k	牛顿法	两点弦割法	单点弦割法
0	0.500000000000	0.500000000000	0.500000000000
1	0.755222417105	0.785398163397	0.785398163397
2	0.739141666149	0.736384138836	0.736384138836
3	0.739085133920	0.739058139213	0.739246689466
4	0.739085133215	0.739085149337	0.739075484183
5	0.739085133215	0.739085133215	0.739085709559
6		0.739085133215	0.739085098789
7			0.739085135271
$\vdots$			$\vdots$
12			0.739085133215
13			0.739085133215

4. 非线性方程组的迭代解法

非线性方程组

$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0, \\ f_2(x_1, x_2, \cdots, x_n) = 0, \\ \qquad \cdots \quad \cdots \\ f_n(x_1, x_2, \cdots, x_n) = 0. \end{cases}$$

其中 $f_i(x_1, x_2, \cdots, x_n) : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, 且 $f_i(x_1, x_2, \cdots, x_n)$ 中至少有一个非线性函数.

非线性方程组

$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0, \\ f_2(x_1, x_2, \cdots, x_n) = 0, \\ \qquad \cdots \quad \cdots \\ f_n(x_1, x_2, \cdots, x_n) = 0. \end{cases}$$

其中 $f_i(x_1, x_2, \cdots, x_n) : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, 且 $f_i(x_1, x_2, \cdots, x_n)$ 中至少有一个非线性函数.

记

$$x = \left(x_1, x_2, \cdots, x_n\right)^{\mathrm{T}}, \; f(x) = \left(f_1(x), f_2(x), \cdots, f_n(x)\right)^{\mathrm{T}}$$

则上述非线性方程组可以表示为

$$f(x) = 0.$$

简单迭代法

- 同解形式

$$x_i = \phi_i(x_1, x_2, \cdots, x_n), \qquad i = 1, 2, \cdots, n$$

- 迭代格式

$$x_i^{(k+1)} = \phi_i(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}), \qquad i = 1, 2, \cdots, n$$

选取 $x^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}\right)^{\mathrm{T}}$, 令 $k = 1, 2, \cdots$, 由迭代格式可得向量迭代序列 $\left\{x^{(k)}\right\}$.

简单迭代法

- 同解形式

$$x_i = \phi_i(x_1, x_2, \cdots, x_n), \qquad i = 1, 2, \cdots, n$$

- 迭代格式

$$x_i^{(k+1)} = \phi_i(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}), \qquad i = 1, 2, \cdots, n$$

选取 $x^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}\right)^{\mathrm{T}}$, 令 $k = 1, 2, \cdots$, 由迭代格式可得向量迭代序列 $\left\{x^{(k)}\right\}$.

若方程组存在唯一解 x^*, 并且 $\left\{x^{(k)}\right\}$ 收敛, 则 $\left\{x^{(k)}\right\}$ 收敛于 x^*.

简单迭代法

记

$$x^{(k)} = \left(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}\right)^{\mathrm{T}},\ \phi(x) = \left(\phi_1(x), \phi_2(x), \cdots, \phi_n(x)\right)^{\mathrm{T}}.$$

$\phi(x)$ 称为迭代函数. 迭代格式的向量形式

$$x^{(k+1)} = \phi(x^{(k)}), \qquad k = 0, 1, 2, \cdots$$

不动点: 若 x^* 满足 $x^* = \phi(x^*)$, 则 x^* 为 $\phi(x)$ 的不动点.

高斯-赛德尔迭代格式:

$$x_i^{(k+1)} = \phi_i\left(x_1^{(k+1)}, \cdots, x_{i-1}^{(k+1)}, x_i^{(k)}, \cdots, x_n^{(k)}\right), \quad i = 1, 2, \cdots, n$$

简单迭代法

例 1：分别用简单迭代法和高斯-赛德尔迭代法求解如下方程组

$$\begin{cases} 4x_1 - x_2 + 0.1e^{x_1} = 1, \\ -x_1 + 4x_2 + 0.125x_2^2 = 0. \end{cases}$$

解 简单迭代法：

$$\begin{cases} x_1^{(k+1)} = 0.25\left(1 + x_2^{(k)} - 0.1e^{x_1^{(k)}}\right), \\ x_2^{(k+1)} = 0.25\left(x_1^{(k)} - 0.125\left(x_2^{(k)}\right)^2\right). \end{cases} \quad k = 0, 1, \cdots$$

高斯-赛德尔迭代法：

$$\begin{cases} x_1^{(k+1)} = 0.25\left(1 + x_2^{(k)} - 0.1e^{x_1^{(k)}}\right), \\ x_2^{(k+1)} = 0.25\left(x_1^{(k+1)} - 0.125\left(x_2^{(k)}\right)^2\right). \end{cases} \quad k = 0, 1, \cdots$$

简单迭代法

取 $x^{(0)} = (0, 0)^{\mathrm{T}}$, 计算结果如下

$x^{(k)}$	不动点	G-S
$x^{(0)}$	$(0.000000,0.000000)^{\mathrm{T}}$	$(0.000000,0.000000)^{\mathrm{T}}$
$x^{(1)}$	$(0.225000,0.000000)^{\mathrm{T}}$	$(0.225000,0.054667)^{\mathrm{T}}$
$x^{(2)}$	$(0.218691,0.054667)^{\mathrm{T}}$	$(0.218691,0.053178)^{\mathrm{T}}$
$x^{(3)}$	$(0.232555,0.053178)^{\mathrm{T}}$	$(0.232555,0.056448)^{\mathrm{T}}$
$x^{(4)}$	$(0.231749,0.056448)^{\mathrm{T}}$	$(0.231749,0.056258)^{\mathrm{T}}$
$\vdots$	$\vdots$	$\vdots$
$x^{(8)}$	$(0.232564,0.056452)^{\mathrm{T}}$	$(0.232564,0.056450)^{\mathrm{T}}$

简单迭代法的收敛性

定理 (压缩映射原理)

设 $\phi : D \subset \mathbf{R}^n \longrightarrow \mathbf{R}^n$, 在闭区域上 $D_0 \subset D$ 满足

(i) $\forall x \in D_0,\ \phi(x) \in D_0$; (ii) 存在常数 $L\ (0 < L < 1)$ 使得对任意的 $x, y \in D_0$, $\big\|\phi(x) - \phi(y)\big\| \leqslant L\|x - y\|$, 则以下结论成立:

(1) ϕ 在 D_0 上存在唯一的不动点 x^*;

(2) 对任意的初始向量 $x^{(0)} \in D_0$, 由简单迭代法产生的迭代序列 $\left\{x^{(k)}\right\} \subset D_0$ 线性收敛于 x^*, 并且有误差估计式

$$\|x^* - x^{(k)}\| \leqslant \frac{L^k}{1 - L}\|x^{(1)} - x^{(0)}\|,$$

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

简单迭代法

(1) 若 ϕ 是压缩的, 则它是连续的.

(2) 压缩常数 L 与范数有关.

(3) 压缩条件有时不易验证.

定理

设 $\phi:\ D\subset \mathbf{R}^n \longrightarrow \mathbf{R}^n$, 如果 ϕ 在凸区域 $D_0\subset D$ 内是可微的, $\phi(x)=\big(\phi_1(x),\cdots,\phi_n(x_i)\big)^{\mathrm{T}}$ 满足: 存在 $L\in(0,1)$ 使得

$$\left|\frac{\partial\phi_i(x)}{\partial x_j}\right|\leqslant\frac{L}{n},\quad \forall x\in D_0,\quad i,j=1,2,\cdots,n$$

则 ϕ 在 D_0 中对 p $(p=1,2,\infty)-$ 范数是压缩的.

简单迭代法

设向量函数 $\phi(x)$ 的所有分量函数 $\phi_i(x)$ 在点 x 处可微, 矩阵

$$J_\phi(x) = \begin{pmatrix} \frac{\partial\phi_1(x)}{\partial x_1} & \frac{\partial\phi_1(x)}{\partial x_2} & \cdots & \frac{\partial\phi_1(x)}{\partial x_n} \\ \frac{\partial\phi_2(x)}{\partial x_1} & \frac{\partial\phi_2(x)}{\partial x_2} & \cdots & \frac{\partial\phi_2(x)}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial\phi_n(x)}{\partial x_1} & \frac{\partial\phi_n(x)}{\partial x_2} & \cdots & \frac{\partial\phi_n(x)}{\partial x_n} \end{pmatrix}$$

称为向量函数 $\phi(x)$ 在 x 处的雅可比矩阵.

简单迭代法的收敛性

若存在常数 $0 < L < 1$ 使得 $\forall x \in D_0,\quad i,j = 1,2,\cdots,n$

$$\left|\frac{\partial \phi_i(x)}{\partial x_j}\right| \leqslant \frac{L}{n} \quad \Longrightarrow \quad \rho\big(J_\phi(x)\big) < 1 \quad .$$

定理 (局部收敛性定理)

设映射 ϕ 在其定义域内有不动点 x^*, ϕ 的分量 $\phi_1,\cdots,\phi_n$ 有连续偏导数且 $J_\phi(x^*)$ 的谱半径 $\rho\big(J_\phi(x^*)\big) < 1$, 则存在 x^* 的一个邻域

$$D_\delta = \{x \,|\, \|x - x^*\| < \delta,\ \delta > 0\} \subset D_0,$$

对任意的 $x^{(0)} \in D_\delta$, 由简单迭代法产生的迭代序列 $\{x^{(k)}\} \subset D_\delta$ 收敛于 x^*.

迭代法的收敛性

定义

设向量序列 $\{x^{(k)}\}$ 收敛于 x^*, 若存在常数 $p \geqslant 1$ 和 $c > 0$, 使得

$$\lim_{k\to\infty} \frac{\|x^* - x^{(k+1)}\|}{\|x^* - x^{(k)}\|^p} = c,$$

则称 $\{x^{(k)}\}$ p 阶收敛于 x^*.

牛顿法

思想: 将非线性函数通过 Taylor 展开逐次线性化, 从而形成一个迭代过程.

将每个方程在 $x^{(k)}$ 处 Taylor 展开

$$
\begin{aligned}
f_i(x) &= f_i(x^{(k)}) + \sum_{j=1}^{n} \frac{\partial f_i(x^{(k)})}{\partial x_j}(x_j - x_j^{(k)}) + \cdots \\
&= f_i(x^{(k)}) + \nabla f_i(x^{(k)}) \cdot (x - x^{(k)}) + \cdots \\
&= f_i(x^{(k)}) + \nabla f_i(x^{(k)})^{\mathrm{T}}(x - x^{(k)}) + \cdots \quad i = 1, 2, \cdots, n
\end{aligned}
$$

牛顿法

取线性项, 由 $f_i(x)=0$ 得近似方程组

$$f_i(x^{(k)})+\nabla f_i(x^{(k)})^{\mathrm{T}}(x-x^{(k)})=0,\quad i=1,2,\cdots,n$$

矩阵向量形式

$$J_f(x^{(k)})(x-x^{(k)})=-f(x^{(k)})$$

令 $\Delta x^{(k)}=x-x^{(k)}$, 则方程组为

$$J_f(x^{(k)})\Delta x^{(k)}=-f(x^{(k)}).$$

求解时先解出 $\Delta x^{(k)}$, 再令 $x^{(k+1)}=x^{(k)}+\Delta x^{(k)}$, 即得

牛顿法 $$\begin{cases} J_f(x^{(k)})\Delta x^{(k)}=-f(x^{(k)}) \\ x^{(k+1)}=x^{(k)}+\Delta x^{(k)} \end{cases}$$

牛顿法

算法:

(1) 给定 $\varepsilon > 0, \delta > 0, x^{(0)} \in D$, 以及最大迭代次数 K. 令 $k = 1$.

(2) 计算 $f(x^{(k)})$, $J_f(x^{(k)})$.

(3) 求解 $\Delta x^{(x)}$ 的线性方程组 $J_f(x^{(k)})\Delta x^{(k)} = -f(x^{(k)})$.

(4) 令 $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$.

(5) 若 $\frac{\|\Delta x^{(k)}\|}{\|x^{(k)}\|} < \delta$ 或 $\|f(x^{(k+1)}\| < \varepsilon$ 成立, 则取$x^* \approx x^{(k+1)}$;

否则

若 $k = K$, 则输出 K 次迭代不满足精度要求的信息.

否则取 $k := k + 1$ 转 (2).

牛顿法

例 2：用牛顿法求解例 1.

解　根据题目有

$$\begin{cases} f_1(x) = 4x_1 - x_2 + 0.1\mathrm{e}^{x_1} - 1 \\ f_2(x) = -x_1 + 4x_2 + 0.125x_1^2 \end{cases}$$

由此得到

$$J_f(x) = \begin{pmatrix} 4 + 0.1\mathrm{e}^{x_1} & -1 \\ -1 + 0.25x_1 & 4 \end{pmatrix} \Longrightarrow \begin{cases} J_f(x^{(k)})\Delta x^{(k)} = -f(x^{(k)}), \\ x^{(k+1)} = x^{(k)} + \Delta x^{(k)}. \end{cases}$$

牛顿法

取 $x^{(0)} = (0,0)^{\mathrm{T}}$, 计算结果如下

k	$x_1^{(k)}$	$x_2^{(k)}$
1	0.000000000	0.000000000
1	0.233766233	0.058441558
2	0.232567040	0.056451572
3	0.232567005	0.056451519
4	0.232567005	0.056451519

因此在 $\|x^{(k+1)} - x^{(k)}\| < 10^{-8}$ 的情况下, 可取 $x^* \approx x^{(4)}$.

牛顿法的收敛性

定理

设 $x^* \in \text{int}(D)$ 是方程组 $f(x)=0$ 的解, $f(x)$ 在 x^* 的某个开邻域 $D_0 \subset D$ 内连续可微, 且 $J_f(x^*)$ 非奇异, 则存在以 x^* 为中心, 正数 δ 为半径的闭球

$$\overline{B}(x^*;\delta)=\{x|\ \|x-x^*\|\leqslant\delta,\delta>0\}\subset D_0$$

使得对任意的 $x^{(0)}\in\overline{B}(x^*;\delta)$, 由牛顿法产生的迭代序列 $\{x^{(k)}\}\subset\overline{B}(x^*;\delta)$ 超线性收敛于 x^*.

进而, 若 $f(x)$ 在 D_0 上二阶可微, 则序列 $\{x^{(k)}\}$ 至少二阶收敛.

牛顿法的收敛性

证 由于 $J_f(x^*)$ 非奇异, 记 $\alpha = \|J_f^{-1}(x^*)\| > 0$. 由于 $f(x)$ 在 D_0 内连续可微, 则对任意的 $\varepsilon \in \left(0, \frac{1}{2\alpha}\right)$, 存在 $\delta > 0$ 使得

$$\|J_f(x) - J_f(x^*)\| < \varepsilon, \quad \forall x \in \overline{B}(x^*;\delta) \subset D_0$$

其中 $\overline{B}(x^*;\delta)$ 是以 x^* 为中心, 以 δ 为半径的闭球. 由于 $\alpha\varepsilon < \frac{1}{2}$, 因此 $J_f(x)$ 可逆, 并且对任意的 $x \in \overline{B}(x^*;\delta)$ 有

$$\|J_f^{-1}(x)\| \leqslant \frac{\alpha}{1 - \alpha\varepsilon} < 2\alpha.$$

于是由牛顿迭代格式产生的迭代序列为

$$x^{(k+1)} = x^{(k)} - J_f^{-1}(x^{(k)}) f(x^{(k)}).$$

牛顿法的收敛性

若 $x^{(k)} \in \overline{B}(x^*;\delta)$, 则

$$\begin{aligned} x^{(k+1)} - x^* &= x^{(k)} - x^* - J_f^{-1}(x^{(k)})f(x^{(k)}) \\ &= x^{(k)} - x^* - J_f^{-1}(x^{(k)})f(x^{(k)}) + J_f^{-1}(x^*)f(x^*) \\ &= -J_f^{-1}(x^{(k)})\Big[f(x^{(k)}) - f(x^*) - J_f(x^*)(x^{(k)} - x^*) \\ &\qquad + \big(J_f(x^*) - J_f(x^{(k)})\big)(x^{(k)} - x^*)\Big] \end{aligned}$$

选取合适的 $\delta > 0$ 使得

$$\|f(x^{(k)}) - f(x^*) - J_f(x^*)(x^{(k)} - x^*)\| < \varepsilon\|x^{(k)} - x^*\|, \ \forall x \in \overline{B}(x^*;\delta).$$

牛顿法的收敛性

因此有

$$
\begin{aligned}
\|x^{(k+1)}-x^*\| &\leqslant 2\alpha\Big(\big\|f(x^{(k)})-f(x^*)-J_f(x^*)(x^{(k)}-x^*)\big\| \\
&\qquad + \big\|\big(J_f(x^*)-J_f(x^{(k)})\big)(x^{(k)}-x^*)\big\|\Big) \\
&< 2\alpha\big(\varepsilon\|x^{(k)}-x^*\|+\varepsilon\|x^{(k)}-x^*\|\big) \\
&= 2\alpha\varepsilon\|x^{(k)}-x^*\| < \|x^{(k)}-x^*\|.
\end{aligned}
$$

即序列 $\{x^{(k)}\}\in\overline{B}(x^*;\delta)$. 同时可知 $\{\|x^{(k)}-x^*\|\}$ 单调下降且有下界, 因此收敛, 即 $\lim\limits_{k\to\infty}x^{(k)}=x^*$.

简化牛顿法

为减少计算量, 将牛顿法中的 $J_f(x^{(k)})$ 用常数矩阵 $J_f(x^{(0)})$ 代替, 得到简化牛顿法

$$\begin{cases} J_f(x^{(0)})\Delta x^{(k)} = -f(x^{(k)}) \\ x^{(k+1)} = x^{(k)} + \Delta x^{(k)} \end{cases}$$

由于 $J_f(x^{(0)})$ 为常数矩阵, 可先对其作三角分解, 从而在每次迭代中只需解两个三角方程组.

相比牛顿法, 计算量减少但收敛速度低.

弦割法

牛顿法中的方程组

$$f_i(x^{(k)}) + \sum_{j=1}^{n} \frac{\partial f_i(x^{(k)})}{\partial x_j}(x_j - x_j^{(k)}) = 0, \quad i = 1, 2, \cdots, n$$

用差商代替偏导数

$$\frac{\partial f_i(x^{(k)})}{\partial x_j} \approx \frac{f_i(x^{(k)} + e_j h) - f_i(x^{(k)})}{h}$$

$$f_i(x^{(k)}) + \frac{1}{h}\sum_{j=1}^{n}\left[f_i(x^{(k)} + e_j h) - f_i(x^{(k)})\right](x_j - x_j^{(k)}) = 0, \quad i = 1, 2, \cdots, n$$

其中 $\Delta x_j^{(k)} = x_j - x_j^{(k)}$, $e_j = (0, \cdots, 0, 1, 0, \cdots)^{\mathrm{T}} \in R^n$.

弦割法

记 $\boxed{f_{ij}^{(k)} = f_i(x^{(k)} + e_j h),\ f_i^{(k)} = f_i(x^{(k)}),}$ 则

$$f_i^{(k)} + \frac{1}{h}\sum_{j=1}^{n}\left(f_{ij}^{(k)} - f_i^{(k)}\right)\Delta x_j^{(k)} = 0, \quad i = 1, 2, \cdots, n$$

$$\Longrightarrow \left(1 - \frac{1}{h}\sum_{j=1}^{n}\Delta x_j^{(k)}\right) f_i^{(k)} + \frac{1}{h}\sum_{j=1}^{n} f_{ij}^{(k)}\Delta x_j^{(k)} = 0,$$

$$\Longrightarrow \quad \sum_{j=1}^{n} f_{ij}^{(k)} \cdot \frac{\frac{1}{h}\sum\limits_{j=1}^{n}\Delta x_j^{(k)}}{\frac{1}{h}\sum\limits_{j=1}^{n}\Delta x_j^{(k)} - 1} = f_i^{(k)}.$$

弦割法

令 $\boxed{z_j = (\sum_{j=1}^{n} \Delta x_j^{(k)})/(\sum_{j=1}^{n} \Delta x_j^{(k)} - h),}$ 则上式可写为

$$\sum_{j=1}^{n} f_{ij}^{(k)} z_j = f_i^{(k)} \quad i = 1, 2, \cdots, n.$$

因此有

$$\Delta x_j^{(k)} = h z_j \left(\frac{1}{h} \sum_{j=1}^{n} \Delta x_j^{(k)} - 1 \right)$$

$$\Longrightarrow \quad \sum_{j=1}^{n} \Delta x_j^{(k)} = \sum_{j=1}^{n} h z_j \cdot \left(\frac{1}{h} \sum_{j=1}^{n} \Delta x_j^{(k)} - 1 \right)$$

故

$$\sum_{j=1}^{n} \Delta x_j^{(k)} = \frac{\sum_{j=1}^{n} h z_j}{\sum_{j=1}^{n} z_j - 1} \Longrightarrow \Delta x_j^{(k)} = \frac{h z_j}{\sum_{j=1}^{n} z_j - 1}$$

弦割法

$$\begin{cases} \text{线性方程组:} \quad \sum_{j=1}^{n} f_{ij}^{(k)} z_j = f_i^{(k)}, \qquad i = 1, 2, \cdots, n \\ x_j^{(k+1)} = x_j^{(k)} + \Delta x_j^{(k)} = x_j^{(k)} + \dfrac{h z_j}{\sum\limits_{j=1}^{n} z_j - 1}, \quad j = 1, 2, \cdots, n \end{cases}$$

给定 $\delta > 0, \varepsilon > 0$, 若 $\|x^{(k+1)} - x^{(k)}\| < \delta$ 或 $\|f(x^{(k+1)})\| < \varepsilon$, 则迭代终止, $x^* \approx x^{(k+1)}$.

布洛依登法

用常数矩阵 A_k 代替牛顿法中的雅可比矩阵, 可得布洛依登法

$$x^{(k+1)} = x^{(k)} - A_k^{-1} f(x^{(k)})$$

(1) A_k 应满足的条件

取 $f(x) \approx f(x^{(k)}) + J_f(x^{(k)})(x - x^{(k)})$, 则

$$f(x^{(k-1)}) \approx f(x^{(k)}) + J_f(x^{(k)})(x^{(k-1)} - x^{(k)}).$$

用 A_k 代替上式中的 $J_f(x^{(k)})$, 并令

$$f(x^{(k-1)}) = f(x^{(k)}) + A_k(x^{(k-1)} - x^{(k)}),$$

$$\Longrightarrow \quad A_k\left(x^{(k)} - x^{(k-1)}\right) = f(x^{(k)}) - f(x^{(k-1)}).$$

布洛依登法

(2) A_k 的构造

记
$$s^{(k)} = x^{(k)} - x^{(k-1)}, \quad y^{(k)} = f(x^{(k)}) - f(x^{(k-1)}),$$

则有 $A_k s^{(k)} = y^{(k)}$. 令

$$A_k = A_{k-1} + \Delta A_{k-1} = A_{k-1} + u s^{(k)\mathrm{T}}, \quad \text{其中 } u \in \mathbf{R}^n.$$

由于

$$y^{(k)} = A_k s^{(k)} = \left(A_{k-1} + u s^{(k)\mathrm{T}}\right) s^{(k)} = A_{k-1} s^{(k)} + u \|s^{(k)}\|_2^2.$$

$$\Longrightarrow \quad u = \frac{y^{(k)} - A_{k-1} s^{(k)}}{\|s^{(k)}\|_2^2}, \quad A_k = A_{k-1} + \frac{\left(y^{(k)} - A_{k-1} s^{(k)}\right) s^{(k)\mathrm{T}}}{\|s^{(k)}\|_2^2}$$

布洛依登法

定理 (Sherman-Morrison)

设 A 是 n 阶可逆矩阵, $u, v \in \mathbf{R}^n$ 为任意向量. 若 $1 + v^{\mathrm{T}} A^{-1} u \neq 0$, 则 $A + uv^{\mathrm{T}}$ 是可逆矩阵, 且其逆矩阵为

$$\left(A + uv^{\mathrm{T}}\right)^{-1} = A^{-1} - \frac{A^{-1} u v^{\mathrm{T}} A^{-1}}{1 + v^{\mathrm{T}} A^{-1} u}$$

取 $v = s^{(k)}$, 则有

$$\begin{aligned} A_k^{-1} = \left(A_{k-1} + u s^{(k)\mathrm{T}}\right)^{-1} &= A_{k-1}^{-1} - \frac{A_{k-1}^{-1} u s^{(k)\mathrm{T}} A_{k-1}^{-1}}{1 + s^{(k)\mathrm{T}} A_{k-1}^{-1} u} \\ &= A_{k-1}^{-1} + \frac{\left(s^{(k)} - A_{k-1}^{-1} y^{(k)}\right) s^{(k)\mathrm{T}} A_{k-1}^{-1}}{s^{(k)\mathrm{T}} A_{k-1}^{-1} y^{(k)}} \end{aligned}$$

布洛依登法

(1) 给定 $x^{(0)} \in D, \delta > 0, \varepsilon > 0$ 以及最大迭代次数 K.

(2) 计算 $A_0 = J_f(x^{(0)}), x^{(1)} = x^{(0)} - A_0^{-1} f(x^{(0)})$. 取 $k = 1$.

(3) 计算

$$s^{(k)} = x^{(k)} - x^{(k-1)}, \qquad y^{(k)} = f(x^{(k)}) - f(x^{(k-1)}),$$

$$A_k^{-1} = A_{k-1}^{-1} + \frac{\left(s^{(k)} - A_{k-1}^{-1} y^{(k)}\right) s^{(k)\mathrm{T}} A_{k-1}^{-1}}{s^{(k)\mathrm{T}} A_{k-1}^{-1} y^{(k)}},$$

$$x^{(k+1)} = x^{(k)} - A_k^{-1} f(x^{(k)}).$$

(4) 若 $\|x^{(k+1)} - x^{(k)}\| < \delta$ 或 $\|f(x^{(k+1)})\| < \varepsilon$, 停止计算, 取 $x^* \approx x^{(k+1)}$. 否则

若 $k = K$ 停止运算, 输出 K 次迭代不满足精度要求的信息;

否则令 $k = k + 1$ 转 (3).

布洛依登法

例 3：用布洛依登法求解方程组

$$\begin{cases} x_1^2 + x_2^2 - 5 = 0, \\ (x_1 + 1)x_2 - 3x_1 = 1. \end{cases}$$

解　首先根据题设, 可得

$$\begin{cases} f_1(x) = x_1^2 + x_2^2 - 5, \\ f_2(x) = (x_1 + 1)x_2 - 3x_1 - 1, \end{cases} \qquad J_f(x) = \begin{pmatrix} 2x_1 & 2x_2 \\ x_2 - 3 & x_1 + 1 \end{pmatrix}.$$

布洛依登法

取 $(x_1^{(0)}, x_2^{(0)}) = (1,1)$, 计算得

$$f(x^{(0)}) = \begin{pmatrix} -3 \\ -2 \end{pmatrix}, \quad A_0 = \begin{pmatrix} 2 & 2 \\ -2 & 2 \end{pmatrix}, \quad A_0^{-1} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

于是有

$$x^{(1)} = x^{(0)} - A_0^{-1} f(x^{(0)}) = \frac{1}{4}(5,9)^{\mathrm{T}}, \quad f(x^{(1)}) = \frac{1}{16}(26,5)^{\mathrm{T}},$$

$$s^{(1)} = x^{(1)} - x^{(0)} = \frac{1}{4}(1,5)^{\mathrm{T}},$$

$$y^{(1)} = f(x^{(1)}) - f(x^{(0)}) = \frac{1}{16}(74,37)^{\mathrm{T}},$$

$$A_1^{-1} = A_0^{-1} + \frac{\left(s^{(1)} - A_0^{-1} y^{(1)}\right) s^{(1)\mathrm{T}} A_0^{-1}}{s^{(1)\mathrm{T}} A_0^{-1} y^{(1)}} = \begin{pmatrix} \frac{233}{1184} & -\frac{338}{1184} \\ \frac{203}{1184} & \frac{234}{1184} \end{pmatrix}.$$

布洛依登法

从而可依次计算 $x^{(2)}, s^{(2)}, y^{(2)}, A_2^{-1}, x^{(3)}, \cdots,$ 数值计算结果如下

k	$x_1^{(k)}$ $x_2^{(k)}$	$f_1(x^{(k)})$ $f_2(x^{(k)})$	A_k^{-1}	
0	1.000000	-3.000000	0.250000	-0.250000
	1.000000	-2.000000	0.250000	0.250000
1	1.250000	1.625000	0.196790	-0.285472
	2.250000	0.312500	0.171452	0.197635
2	1.019425	-0.314090	0.194649	-0.285502
	1.909628	-0.201924	0.123278	0.196963
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

本章总结

- 二分法
  - 方法简单易实现，收敛
  - 收敛慢，运算量较大
  - $\Longrightarrow$ 用于根的隔离
- 简单迭代法
  - $x_{k+1}=\phi(x_k)$. $\phi(x)$ 满足一定条件时局部 (全局) 收敛
  - 线性收敛 $\longrightarrow$ 松弛加速法，艾特肯加速 $\longrightarrow$ steffesen 迭代法
- 牛顿类迭代法
  - 牛顿法 $x_{k+1}=x_k-\dfrac{f(x_k)}{f'(x_k)}$ 　$\boxed{\text{牛顿下山法 } f'(x_k)\rightarrow \dfrac{1}{\lambda}f'(x_k)}$
  - 简化牛顿法 $f'(x_k)\rightarrow c$ 或 $f'(x_0)$
  - 弦割法 $f'(x_k)\rightarrow \dfrac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}$ 超线性收敛
- 非线性方程组迭代法
  - 简单迭代法 $\longrightarrow$ 高斯-赛德尔迭代法
  - 牛顿法 $\Longrightarrow$
    - 弦割法 $\dfrac{\partial f_i(x^{(k)})}{\partial x_j}\rightarrow \dfrac{f_i(x^{(k)}+e_jh)-f_i(x^{(k)})}{h}$
    - 布洛依登法 $J_f(x^{(k)})\longrightarrow A_k$ 超线性收敛

要求

1. 熟练掌握求解非线性方程的几种基本迭代法：二分法、简单迭代法、牛顿法、弦割法
2. 迭代法的收敛性
3. 掌握求解非线性方程组的几种迭代法：简单迭代法、牛顿法
4. 了解求解非线性方程组的弦割法、Broyden法

1. 简单迭代法线性收敛

简单迭代法$x_{k+1} = \phi(x_k)$当$0 < |\phi'(x^*)| < 1$ 时是线性收敛的.

事实上，设x^*是$x = \phi(x)$的不动点，即$x^* = \phi(x^*)$. 该式减去$x_{k+1} = \phi(x_k)$, 则得

$$x^* - x_{k+1} = \phi(x^*) - \phi(x_k) = \phi'(\xi_k)(x^* - x_k).$$

其中ξ_k介于x^*与x_k之间. 从而

$$\lim_{k\to\infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|} = \lim_{k\to\infty} |\phi'(\xi_k)| = |\phi'(x^*)|.$$

2. 牛顿法二阶收敛

设$f(x)$在含根区间$[a,b]$上二阶连续可微且$f'(x)\neq 0$. x_k是方程$f(x)=0$的根x^*的一个近似值. 则由泰勒公式有

$$0=f(x^*)=f(x_k)+f'(x_k)(x^*-x_k)+\frac{1}{2!}f''(\xi_k)(x^*-x_k)^2.$$

上式两端同除以$f'(x_k)$得

$$\begin{aligned}&x^*-x_k+\frac{f(x_k)}{f'(x_k)}+\frac{1}{2}\frac{f''(\xi_k)}{f'(x_k)}(x^*-x_k)^2\\&=x^*-x_{k+1}+\frac{1}{2}\frac{f''(\xi_k)}{f'(x_k)}(x^*-x_k)^2=0.\end{aligned}$$

所以,

$$\lim_{k\to\infty}\frac{|x^*-x_{k+1}|}{|x^*-x_k|^2}=\lim_{k\to\infty}\frac{1}{2}\frac{|f''(\xi_k)|}{|f'(x_k)|}=\frac{1}{2}\frac{|f''(x^*)|}{|f'(x^*)|}\neq 0.$$

3. 弦割法超线性收敛

由牛顿插值公式有，

$$x_{k+1}-x^*=\frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)}f'(\xi_k^*)(x_k-x^*)f'(\xi_{k-1}^*)(x_{k-1}-x^*).$$

$$\frac{|x^*-x_{k+1}|}{|x^*-x_k|^p}=\frac{1}{2}\left|\frac{f''(\xi_k)f'(\xi_k^*)f'(\xi_{k-1}^*)}{f'^3(\xi_k)}\right|$$

$$\times\left(\frac{|x_k-x^*|}{|x^*-x_{k-1}|^p}\right)^{1-p}|x^*-x_{k-1}|^{1+p(1-p)}.$$

在上式两端令$k\to\infty$，取极限得

$$c=\frac{1}{2}\left|\frac{f''(x^*)}{f'(x^*)}\right|c^{1-p}\lim_{k\to\infty}|x_{k-1}-x^*|^{1+p-p^2}.$$

得$p=\frac{1+\sqrt{5}}{2}\approx 1.618$，即弦割法超线性收敛.

4. 单点弦割法线性收敛

单点弦割法实质上是简单迭代法，其迭代函数

$$\phi(x) = x - \frac{f(x)}{f(x) - f(x_0)}(x - x_0)$$

在点x^*处的导数 $\phi'(x^*) = 1 - \frac{f'(x^*)}{f'(\xi)}$, 其中$\xi$介于$x_0$与$x^*$之间，由此可见单点弦割法一般线性收敛，但当$f'(x)$变化不大时$\phi'(x^*) \approx 0$,收敛可能很快.

弦割法的收敛阶虽然低于牛顿法，但每次迭代只需计算一个函数值$f(x_k)$，不需计算导数值$f'(x_k)$，效率高，实际问题中经常采用.