

数值计算方法

第七章 非线性方程与方程组的 数值解法

智能与计算学部



第七章 非线性方程与方程组的数值解法



- ▶ 7.1 非线性方程求根与二分法
- ▶ 7.2 不动点迭代法及其收敛性
- ▶ 7.3 迭代收敛的加速方法
- ▶ 7.4 牛顿法、简化牛顿法与牛顿下山法
- ▶ 7.5 弦截法
- ▶ 7.6 非线性方程组的数值解法

7.1 非线性方程求根与二分法

7.1.1 非线性方程的求根问题

7.1.2 二分法



7.1.1 非线性方程求根 (1)



- ▶ 非线性问题是实际问题中经常遇到的问题。非线性方程的求根是其中一个重要的课题。
- ▶ 方程 $\sin(x) - \frac{1}{2} = 0$ 在 \mathbb{R} 上有无穷多根
- ▶ 方程 $e^x - \cos(\pi x) - 1 = 0$ 在 $(0, 1)$ 上有一个根, 但无法用初等函数表示
- ▶ 方程 $\begin{cases} y = x^2 + c \\ x = y^2 + c \end{cases}$ 的根与 c 的值相关
 - ▶ $c = 0$, 方程有2个根 $x = 0, 1$
 - ▶ $c = 0$, 方程有1个根 $x = 1/2$
 - ▶ $c = 0$, 方程有3个根 $x = -1, 0, (1 + \sqrt{5})/2$

7.1.1 非线性方程求根 (2)



- ▶ 非线性问题相比较于线性问题要复杂得多，大部分非线性问题的解没有解析形式，只能通过数值方法求近似解。
- ▶ 现有的理论大部分只讨论非线性问题的局部性质。因此数值上，只能在局部求解非线性方程的根。
- ▶ 本章中讨论二分法（对分区间法）和迭代法

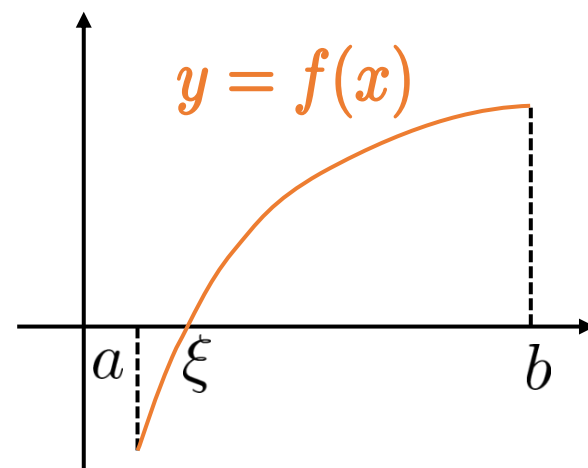
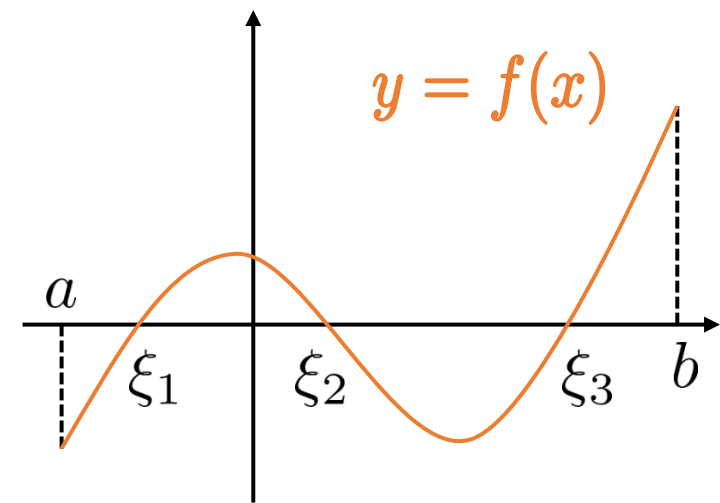
7.1.2 二分法—零点定理



► 若 $f(x) \in \mathbb{C}^0[a, b]$, $f(a)f(b) < 0$, 则至少有一点 $\xi \in (a, b)$, 使得 $f(\xi) = 0$, 即方程 $f(x) = 0$ 在 (a, b) 内至少存在一个实根。 $[a, b]$ 被称为 $f(x)$ 的**有根区间**。

► 设 $f(x) \in \mathbb{C}^0[a, b]$ 且**单调**, 若有 $f(a)f(b) < 0$, 则有且仅有一点 $\xi \in (a, b)$, 使得 $f(\xi) = 0$ 。

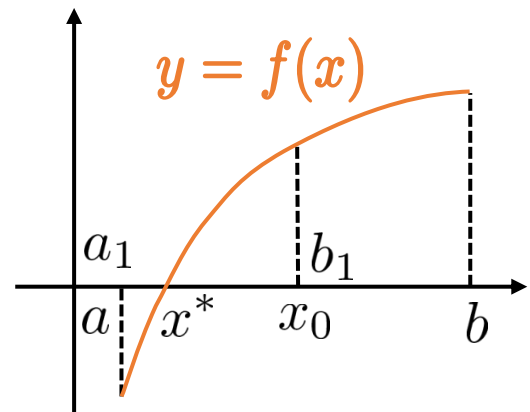
► 零点定理的几何解释





7.1.2 二分法的求解过程

- 令 $x_0 = (a + b)/2$,
若 $f(a)f(x_0) > 0$, 则 $a_1 \leftarrow x_0, b_1 \leftarrow b$;
若 $f(a)f(x_0) < 0$, 则 $a_1 \leftarrow a, b_1 \leftarrow x_0$;
不管哪种情况, 必有 $b_1 - a_1 = (b_0 - a_0)/2$



- 对 $[a_1, b_1]$ 进行同样操作, 可得到序列
 $[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \cdots [a_k, b_k] \supset \cdots$

且有 $b_k - a_k = (b - a)/2^k$,

此过程无限继续, 则这些区间必收缩于一点 x^* 。

- 取 $x_k = (a_k + b_k)/2$, 则 $x_0, x_1, \cdots, x_k, \cdots$ 必趋向于 x^* 。

7.1.2 二分法的精度控制



►二分法的精度

$$|x^* - x_k| \leq (b_k - a_k)/2 = (b - a)/2^{k+1}$$

给定精度 ϵ , 只需使 $(b - a)/2^{k+1} \leq \epsilon$, 可求得k值。即有

$$k \geq \log_2 \frac{b - a}{\epsilon} - 1$$

7.1.2 二分法的伪代码



Algorithm: 求解非线性方程 $f(x) = 0$ 的二分法

Input : a, b ; // 初始有根区间 $[a, b]$
 fx ; // 函数 $f(x)$
 ϵ ; // 迭代终止判断条件

Output: x^* ; // $f(x)$ 在 $[a, b]$ 内的根值

while $b - a \geq \epsilon$ **do**

if $fx(\frac{a+b}{2}) = 0$ **then**

return $\frac{a+b}{2}$;

else if $fx(a)fx(\frac{a+b}{2}) < 0$ **then**

$a \leftarrow a$;

$b \leftarrow \frac{a+b}{2}$;

else

$a \leftarrow \frac{a+b}{2}$;

$b \leftarrow b$;

end if

end while

return $\frac{a+b}{2}$; // 达到循环结束条件



7.1.2 二分法例题

►例：求方程 $f(x) = x^3 - x - 1 = 0$ 在区间 $[1.0, 1.5]$ 内的一个实根，要求准确到小数点后的第2位。

►解： $a=1.0, b=1.5$ ，令 $(b-a)/2^{k+1} \leq 0.5 \times 10^{-2}$ ，得 $k \geq 6$

表7-2

k	a_k	b_k	x_k	$f(x_k)$ 符号
0	1.0	1.5	1.25	—
1	1.25		1.375	+
2		1.375	1.3125	—
3	1.3125		1.3438	+
4		1.3438	1.3281	+
5		1.3281	1.3203	—
6	0.3203		1.3242	—

不动点迭代及其收敛性

7.2.1 不动点与不动点迭代

7.2.2 不动点存在性与迭代收敛性

7.2.3 局部收敛性与收敛阶





7.2.1 不动点迭代法引例

► **例1** 求解方程 $3x - e^x = 0$.

► **解析** 方程变形 $x = e^x / 3$

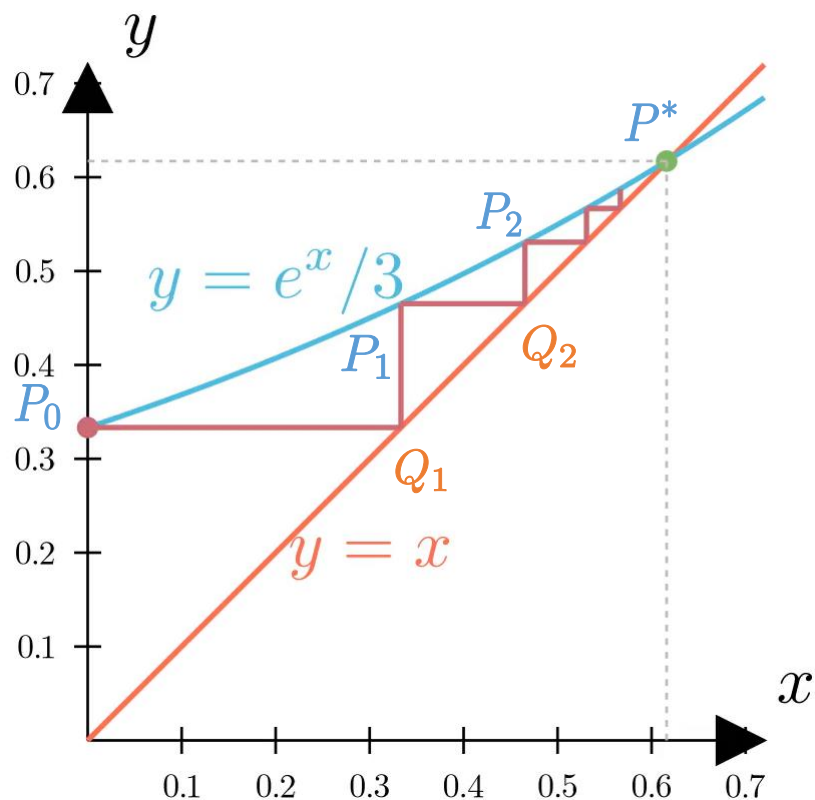
方程 $3x - e^x = 0$ 的迭代求解表

$$x = e^x / 3$$

序号	0	1	2	3	4
左边	0	0.333	0.465	0.531	0.567
右边	0.333	0.465	0.531	0.567	0.588

5	6	7	8	9	10
0.588	0.599	0.607	0.612	0.615	0.616
0.599	0.607	0.612	0.615	0.616	0.616

迭代过程如下图所示



7.2.1 不动点迭代法



- ▶ 将方程 $f(x) = 0$ 变换为其等价形式 $x = \varphi(x)$ ，若所求的 x^* 满足 $f(x^*) = 0$ ，则 $x^* = \varphi(x^*)$ ，反之亦然。 x^* 为函数 $\varphi(x)$ 的一个不动点。
- ▶ 举例： $f(x) = 0 \iff x = x + af(x), a \neq 0$
$$\iff \sqrt[3]{x^3 + f(x)}$$
- ▶ 构造迭代方程 $x_{k+1} = \varphi(x_k)$ ，若对任何 $x \in [a, b]$ ，序列 $\{x_k\}$ 有极限 $\lim_{k \rightarrow \infty} x_k = x^*$ ，则称该迭代方程收敛。称以上迭代法为不动点迭代法。



7.2.2 不动点迭代法的收敛条件

► **例2** 用迭代法求解方程

$$x^3 - x^2 - x - 1 = 0.$$

► **解析：** 第一步构造迭代函数

$$x = x^3 - x^2 - 1 = \phi_1(x)$$

$$x = \sqrt[3]{x^2 + x + 1} = \phi_2(x)$$

$$x = 1 + \frac{1}{x} + \frac{1}{x^2} = \phi_3(x)$$

对于给定的方程 $f(x) = 0$ ，
有多种方式将它改写成等
价的形式 $x = \phi(x)$ 。但如
何改写使得序列收敛？

$x_0 = 1$	$\phi_1(x)$	$\phi_2(x)$	$\phi_3(x)$
1	-1	1.4422	3.0000
2	-3	1.6537	1.4444
3	-37	1.7532	2.1716
4	-52023	1.7995	1.6725
5		1.8209	1.9554
6		1.8308	1.7730
7		1.8354	1.8822
8		1.8175	1.8136
9		1.8385	1.8554
10		1.8389	1.8294
11		1.8391	1.8454
12		1.8392	1.8355
13		1.8392	1.8416

精确解： $x^* = 1.8393$

7.2.2 不动点迭代的收敛条件



►定理1 (压缩映象原理)

►设 $\varphi(x) \in C[a, b]$ 满足以下条件:

1. 对任意 $x \in [a, b]$, 有 $a \leq \varphi(x) \leq b$;

2. 存在常数 $0 < L < 1$, 对 $\forall x, y \in [a, b]$ 有

$$|\varphi(x) - \varphi(y)| \leq L|x - y|$$

则 $\varphi(x)$ 在 $[a, b]$ 存在唯一不动点 x^* ,

且对任意 $x_0 \in [a, b]$, 由 $x_{k+1} = \varphi(x_k)$ 得到的序列 $\{x_k\}$ 有

$$\lim_{k \rightarrow \infty} x_k = x^*, \text{ 并有误差估计 } |x_k - x^*| \leq \frac{L^k}{1 - L} |x_1 - x_0|.$$



7.2.2 收敛条件说明 (1)

► (1) 对条件 $|\varphi(x) - \varphi(y)| \leq L|x - y|$ 的说明

► 设 $\varphi(x) \in C^1[a, b]$ 且对任意 $x \in [a, b]$ 都有

$$|\varphi'(x)| \leq L < 1,$$

则根据拉格朗日中值定理, 对 $\forall x, y \in [a, b]$, 有

$$|\varphi(x) - \varphi(y)| = |\varphi'(\xi)(x - y)| \leq L|x - y|, \xi \in (a, b)$$

故上述定理中的条件(2)常被表述成 $|\varphi'(x)| < 1$

7.2.2 收敛条件说明 (2)



► (2) 对误差估计的说明

► 由于:

$$\begin{aligned} |x_{k+p} - x_{k+p-1}| &= |\varphi(x_{k+p-1}) - \varphi(x_{k+p-2})| \leq L|x_{k+p-1} - x_{k+p-2}| \\ &\leq L^2|x_{k+p-2} - x_{k+p-3}| \\ &\vdots \\ &\leq L^{p-1}|x_{k+1} - x_k| \end{aligned}$$

所以:

$$\begin{aligned} |x_{k+p} - x_k| &\leq |x_{k+p} - x_{k+p-1}| + |x_{k+p-1} - x_{k+p-2}| + \cdots + |x_{k+1} - x_k| \\ &\leq (L^{p-1} + L^{p-2} + \cdots + 1)|x_{k+1} - x_k| \\ &\leq \frac{1}{1-L}|x_{k+1} - x_k| \end{aligned}$$

$$p \rightarrow \infty, \therefore |x^* - x_k| \leq \frac{1}{1-L}|x_{k+1} - x_k|$$

结论: 只要相邻两次计算结果的偏差足够小, 即可保证 x_k 有足够精度。

7.2.2 不动点迭代伪代码



Algorithm: 求解非线性方程 $f(x) = 0$ 的不动点迭代法

Input : φ ; // 函数 $\varphi(x)$
 x_0 ; // 迭代初始值
 K ; // 循环最大次数
 ϵ ; // 迭代终止判断条件

Output: k, x^* ; // 迭代次数及根值

for $k \leftarrow 0$ **to** K **do**

$x_1 \leftarrow \varphi(x_0)$;

if $|x_1 - x_0| < \epsilon$ **then**
 return k, x_1 ;
 end if

$x_0 \leftarrow x_1$;

end for

return ∞, null ; // 超出最大循环次数

if $|\varphi(x_1) - x_1| < \epsilon$ **then**
 return k, x_1 ;
end if



7.2.2 不动点迭代收敛性例题



►例 (课后习题2): 为求方程 $x^3 - x^2 - 1 = 0$ 在 $x_0 = 1.5$ 附近的一个根, 设将方程改写成下列等价形式, 并建立相应的迭代公式,

(1) $x = 1 + 1/x^2$, 迭代公式 $x_{k+1} = 1 + 1/x_k^2$

(2) $x^3 = 1 + x^2$, 迭代公式 $x_{k+1} = \sqrt[3]{1 + x_k^2}$

(3) $x^2 = \frac{1}{x-1}$, 迭代公式 $x_{k+1} = 1/\sqrt{x_k - 1}$

试分析每种迭代公式的收敛性, 并选取一种公式求出具有四位有效数字的近似根。

7.2.2 不动点迭代收敛性例题 (2)



►解：考虑 $x_0 = 1.5$ 的邻域 $[1.3, 1.6]$

► (1) 当 $x \in [1.3, 1.6]$, 有 $\varphi(x) = 1 + \frac{1}{x^2} \in [1.3, 1.6]$,

$$|\varphi'(x)| = \left| -\frac{2}{x^3} \right| \leq \frac{2}{1.3^2} \approx 0.910 = L < 1, \text{ 故迭代}$$

$x_{k+1} = x + 1/x_k^2$ 在 $[1.3, 1.6]$ 上整体收敛;

(2) 当 $x \in [1.3, 1.6]$, 有 $\varphi(x) = (1 + x^2)^{1/3} \in [1.3, 1.6]$

$$|\varphi'(x)| = \frac{2}{3} \left| \frac{x}{(1 + x^2)^{2/3}} \right| < \frac{2}{3} \frac{1.6}{(1 + 1.3^2)^{2/3}} \approx 0.522 = L < 1$$

故迭代 $x_{k+1} = \sqrt[3]{x + x_k^2}$ 在 $[1.3, 1.6]$ 上整体收敛;



7.2.2 不动点迭代收敛性例题 (3)

(3) 当 $x \in [1.3, 1.6]$, $\varphi(x) = \frac{1}{\sqrt{x-1}} \in [1.29, 1.83] \not\subset [1.3, 1.6]$

$$|\varphi'(x)| = \left| \frac{-1}{2(x-1)^{3/2}} \right| > \frac{1}{2(1.6-1)} > 1$$

故迭代 $x_{k+1} = 1/\sqrt{x_k-1}$ 不一定收敛。

由于(2)中的L较小, 故取(2)的迭代公式计算, 结果具有4位有效数字, 只需:

$$|x_k - x^*| \leq \frac{L}{1-L} |x_k - x_{k-1}| < 0.5 \times 10^{-3}$$
$$\Rightarrow |x_k - x_{k-1}| < \frac{1-L}{L} \times 0.5 \times 10^{-3} < 0.5 \times 10^{-3}$$

取 $x_0 = 1.5$
计算表为:

k	x_k	k	x_k	k	x_k
1	1.481 248 034	3	1.468 817 314	5	1.466 243 010
2	1.472 705 730	4	1.467 047 973	6	1.465 876 820

由于 $|x_6 - x_5| < 0.5 \times 10^{-3}$, 故可取 $x^* \approx x_6 = 1.466$

7.2.3 不动点迭代的局部收敛



►定义：局部收敛

►设 $\varphi(x)$ 有不动点 x^* , 如果存在 x^* 的某个邻域 $R: |x - x^*| \leq \delta$ 对任意的 $x_0 \in R$, $x_{k+1} = \varphi(x_k)$ 产生的序列 $\{x_k\} \subset R$, 且收敛到 x^* , 则称迭代法 $x_{k+1} = \varphi(x_k)$ **局部收敛**。

►定理2 (局部收敛定理)

►设 x^* 为 $\varphi(x)$ 的不动点, $\varphi'(x)$ 在 x^* 的某个邻域连续, 且 $|\varphi'(x^*)| < 1$, 则迭代法 $x_{k+1} = \varphi(x_k)$ 局部收敛。

7.2.3 收敛阶 (1)



►如何判断收敛的速度？

►L在一定程度上可以反映收敛速度

►L的取值与邻域范围选择相关，不适合判定收敛快慢

►定义：设迭代过程 $x_{k+1} = \varphi(x_k)$ 收敛于方程 $x = \varphi(x)$ 的根 x^* ，迭代误差 $e_k = x_k - x^*$ 满足

$$\lim_{k \rightarrow \infty} \frac{e_{k+1}}{(e_k)^p} = c, c \neq 0$$

则称该迭代过程是 p 阶收敛的。特别地，当 $p = 1$ 时，称为线性收敛， $p > 1$ 时，称为超线性收敛， $p = 2$ 时，称为平方收敛。

7.2.3 收敛阶 (2)



►定理：对于迭代过程 $x_{k+1} = \varphi(x_k)$ 及正整数 p , 如果 $\varphi^{(p)}(x)$ 在所求根 x^* 的邻近连续, 并且

$$\varphi'(x^*) = \varphi''(x^*) = \cdots = \varphi^{(p-1)}(x^*) = 0$$

$$\varphi^{(p)}(x^*) \neq 0$$

则该迭代过程在点 x^* 邻近是 p 阶收敛的。

7.2.3 收敛阶 (3)



►证明：由于 $\varphi'(x^*) = 0$ ，由局部收敛定理可以迭代过程

$x_{k+1} = \varphi(x_k)$ 具有局部收敛性。

再将 $\varphi(x_k)$ 在根 x^* 处做泰勒展开，利用定理中条件，有：

$$\varphi(x_k) = \varphi(x^*) + \frac{\varphi^{(p)}(\xi)}{p!} (x_k - x^*)^p, \quad \xi \text{ 在 } x_k \text{ 和 } x^* \text{ 之间。}$$

注意到 $\varphi(x_k) = x_{k+1}$, $\varphi(x^*) = x^*$, 所以：

$$x_{k+1} - x^* = \frac{\varphi^{(p)}(\xi)}{p!} (x_k - x^*)^p$$

因此对于迭代误差，当 $k \rightarrow \infty$ 时有

$$\frac{e_{k+1}}{e_k^p} \rightarrow \frac{\varphi^{(p)}(x^*)}{p!} \neq 0$$

故原迭代过程为 p 阶收敛。

7.2.3 收敛阶例题



►例：为求方程 $x^3 + 2x - 5 = 0$ 在 $[1, 2]$ 内的根，产生如下迭代格式 $x_{k+1} = \sqrt[3]{5 - 2x_k}$ ，给出该格式的收敛性和收敛阶。

►解：由题知 $\varphi(x) = \sqrt[3]{5 - 2x}$ ，在区间 $[1, 2]$ 内，有

$$|\varphi'(x)| = \left| \frac{1}{3} \frac{-2}{(5 - 2x)^{2/3}} \right| \leq \frac{2}{3} < 1$$

所以，该迭代格式是收敛的。

在根 x^* 处，有 $\varphi'(x^*) = -\frac{2}{3(5 - 2x^*)^{2/3}} = -\frac{2}{3(x^*)^2} \neq 0$

所以迭代格式是线性收敛的。

7.3 迭代收敛的加速方法

7.3.1 斯特芬森迭代法





7.3.1 斯特芬森迭代法 (1)

► 若迭代格式是线性收敛的, 即有

$$\lim_{k \rightarrow \infty} \frac{x_{k+1} - x^*}{x_k - x^*} = C$$

则, 当 k 充分大时, 有

$$x_{k+1} - x^* \approx C(x_k - x^*)$$

$$x_{k+2} - x^* \approx C(x_{k+1} - x^*)$$

因此

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

可以得到:

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

7.3.1 斯特芬森迭代法 (2)



► 斯特芬森加速迭代法 (Steffensen)

► 若有迭代格式 $x_{k+1} = \varphi(x_k)$, 则格式

$$\begin{cases} y_k = \varphi(x_k), z_k = \varphi(y_k) \\ x_{k+1} = x_k - \frac{(y_k - x_k)^2}{z_k - 2y_k + x_k} \end{cases} \quad k = 0, 1, 2, \dots$$

称为斯特芬森迭代法。

► 定理：设 $x^* = \varphi(x^*)$, $\varphi(x)$ 在包含 x^* 的某个开区间内具有连续的二阶导数, 并且 $\varphi'(x^*) \neq 1$, 则如上的迭代格式至少二阶收敛。

7.3.1 斯特芬森迭代伪代码



Algorithm: 求解非线性方程 $f(x) = 0$ 的斯特芬森迭代法

Input : φ ; // 函数 $\varphi(x)$

x_0 ; // 迭代初始值

K ; // 循环最大次数

ϵ ; // 迭代终止判断条件

Output: k, x^* ; // 迭代次数及根值

for $k \leftarrow 0$ **to** K **do**

$y \leftarrow \varphi(x_0)$;

$z \leftarrow \varphi(y)$;

$x_1 \leftarrow x_0 - (y - x_0)^2 / (z - 2y + x_0)$;

if $|x_1 - x_0| < \epsilon$ **then**

return k, x_1 ;

end if

$x_0 \leftarrow x_1$;

end for

return ∞, null ; // 超出最大循环次数

7.3.1 斯特芬森迭代法例题(1)



►例 (课后习题2): 为求方程 $x^3 - x^2 - 1 = 0$ 在 $x_0 = 1.5$ 附近的一个根, 设将方程改写成下列等价形式, 并建立相应的迭代公式,

(1) $x = 1 + 1/x^2$, 迭代公式 $x_{k+1} = 1 + 1/x_k^2$

(2) $x^3 = 1 + x^2$, 迭代公式 $x_{k+1} = \sqrt[3]{1 + x_k^2}$

(3) $x^2 = \frac{1}{x-1}$, 迭代公式 $x_{k+1} = 1/\sqrt{x_k - 1}$

将史蒂芬森加速迭代法应用到以上迭代格式中。



7.3.1 斯特芬森迭代法例题(2)

►解：令 $f(x) = x^3 - x^2 - 1$ ，对所有迭代格式，迭代终止条件设置为 $|f(x_k)| < \epsilon = 10^{-5}$ ，迭代结果如下：

(1)

13	21
1.5000000	1.5000000
1.4658586	1.4444444
1.4577077	1.4792899
1.4656993	1.4569760
1.4675345	1.4710806
1.4656148	1.4620905
.....
1.4655769	1.4655561
1.4655447	1.4655808
1.4655755	1.4655651
1.4655795	1.4655751
1.4655717	1.4655688

(2)

3	12
1.5000000	1.5000000
1.4655585	1.4812480
1.4757171	1.4727057
1.4655686	1.4688173
	1.4670480
	1.4662430

	1.4656345
	1.4656000
	1.4655843
	1.4655772
	1.4655739

(3)

13	inf
1.5000000	1.5000000
1.4673423	1.4142136
1.4526935	1.5537740
1.4875698	1.3437972
1.4876030	1.7054887
1.4663162	1.1905701
.....
1.4657236	0.5+0.866j
1.4655981	0.5-0.866j
1.4655177	0.5+0.866j
1.4655164	0.5-0.866j
1.4655708

7.4 牛顿迭代法

7.4.1 牛顿迭代法原理

7.4.2 牛顿迭代法的几何意义

7.4.3 牛顿迭代法的收敛性

7.4.4 简化的牛顿法与牛顿下山法





7.4.1 牛顿迭代法原理

► 基本思路：将非线性方程 $f(x) = 0$ 线性化

假设方程 $f(x) = 0$ 有根 x_k ，将 $f(x)$ 在 x_k 处展开，有：

$$f(x) \approx f(x_k) + f'(x_k)(x - x_k)$$

所以 $f(x) = 0$ 可近似表示为 $f(x_k) + f'(x_k)(x - x_k) = 0$

化简得到 $x = x_k - \frac{f(x_k)}{f'(x_k)}$ ，

所以迭代公式为

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

7.4.1 牛顿迭代法的伪代码



Algorithm: 求解非线性方程 $f(x) = 0$ 的牛顿迭代法

Input : \mathbf{fx} ; // 函数 $f(x)$
 \mathbf{dfx} ; // $f(x)$ 的导函数 $f'(x)$
 x_0 ; // 迭代初始值
 K ; // 循环最大次数
 ϵ ; // 迭代终止判断条件

Output: k, x^* ; // 迭代次数及根值

for $k \leftarrow 0$ **to** K **do**

$x_1 \leftarrow x_0 - \mathbf{fx}(x_0)/\mathbf{dfx}(x_0);$

if $|x_1 - x_0| < \epsilon$ **then**

return $k, x_1;$

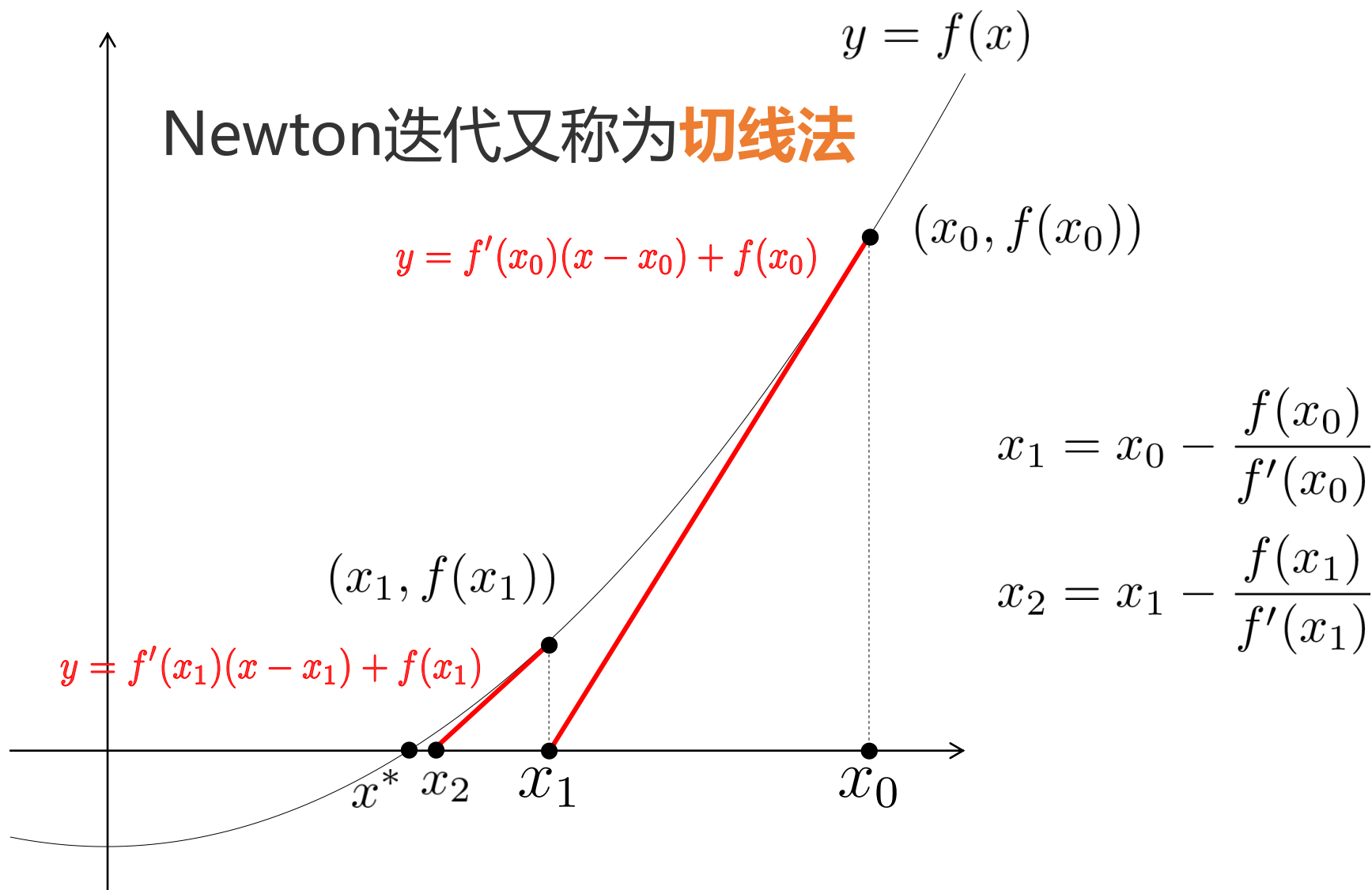
end if

$x_0 \leftarrow x_1;$

end for

return ∞, null ; // 超出最大循环次数

7.4.2 牛顿迭代的几何意义



7.4.3 牛顿迭代的收敛性



►由牛顿迭代格式知，对应的等价方程为

$$x = x - \frac{f(x)}{f'(x)}$$

记：

$$\varphi(x) = x - \frac{f(x)}{f'(x)}$$

则：

$$\varphi'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}$$

设 $f'(x^*) \neq 0$ ，由于 $f(x^*) = 0$ ，所以 $\varphi'(x^*) = 0$ 。

►定理：牛顿迭代格式在 x^* 附近至少平方收敛。



7.4.3 牛顿迭代法例题

►例：根据牛顿法写出求解 \sqrt{a} , ($a > 0$) 的计算公式，并计算 $\sqrt{115}$ 。

►解：设 $f(x) = x^2 - a$ ，求 \sqrt{a} 即计算 $f(x)$ 的根。

由牛顿迭代公式可得：

$$\begin{aligned}x_{k+1} &= x_k - \frac{f(x_k)}{f'(x_k)} = x_k - \frac{x_k^2 - a}{2x_k} \\&= \frac{1}{2} \left(x_k + \frac{a}{x_k} \right)\end{aligned}$$

令 $a = 115$, 有以下迭代结果：

k	0	1	2	3	4
x_k	10	10.750 000	10.723 837	10.723 805	10.723 805



7.4.4 牛顿迭代法存在的问题

- ▶ 1. 每步计算都需要计算 $f'(x)$ 。计算量大，且当 $f'(x)$ 没有解析形式时，计算更困难。

简化牛顿法（平行弦法）

▶ 例如 $f(x) = x^{x/2} - x$

- ▶ 2. 对初始值 x_0 的选择敏感。只有 x_0 很接近 x^* 时，才能够保证收敛。

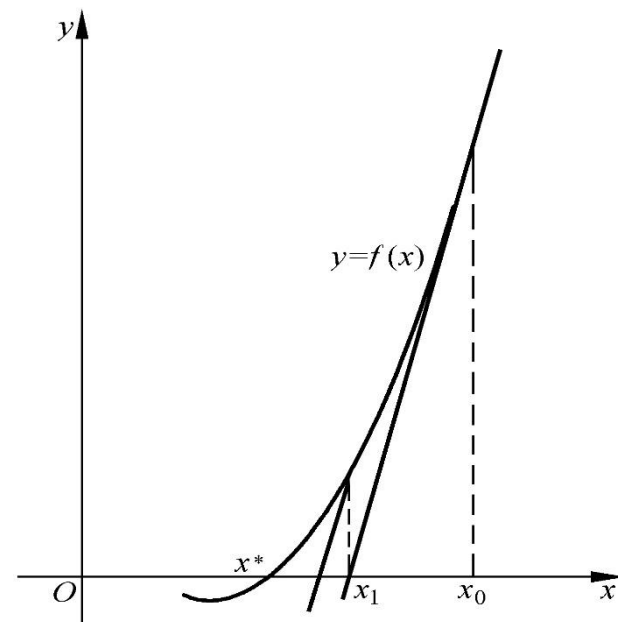
牛顿下山法

7.4.4 简化牛顿法



- ▶ 将牛顿迭代 $\varphi(x) = x - f(x)/f'(x)$ 中的 $f'(x)$ 用常数 C 代替
- ▶ 为在 x^* 附近有更好收敛性, 令 $|\varphi'(x^*)| = \left| 1 - \frac{f(x^*)}{C} \right| = 0$ 得 $C = f'(x^*)$
- ▶ 这种情况很难发生, 在实际应用中, 取 $C = f'(x_0)$
- ▶ 也被称为**平行弦法**
- ▶ 不同于牛顿法, 简化牛顿法为线性收敛

简化牛顿法的几何意义





7.4.4 简化牛顿法的伪代码

Algorithm: 求解非线性方程 $f(x) = 0$ 的简化牛顿法

Input : \mathbf{fx} ; // 函数 $f(x)$
 \mathbf{dfx} ; // $f(x)$ 的导函数 $f'(x)$
 x_0 ; // 迭代初始值
 K ; // 循环最大次数
 ϵ ; // 迭代终止判断条件

Output: k, x^* ; // 迭代次数及根值

$C \leftarrow \mathbf{dfx}(x_0)$;

for $k \leftarrow 0$ **to** K **do**

$x_1 \leftarrow x_0 - \mathbf{fx}(x_0)/C$;

if $|x_1 - x_0| < \epsilon$ **then**

return k, x_1 ;

end if

$x_0 \leftarrow x_1$;

end for

return ∞, null ; // 超出最大循环次数



7.4.4 牛顿迭代不收敛实例

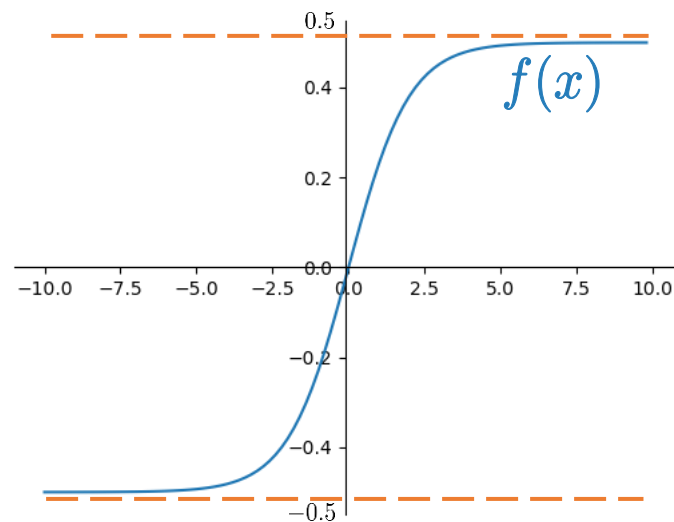
►例：用牛顿迭代求解函数 $f(x) = \text{sig}(x) - 0.5$ 的根，其中，

$\text{sig}(x) = \frac{e^x}{1 + e^x}$ 被称为 sigmoid 函数。

►解：由图像知， $x^* = 0$ 是 $f(x)$ 的根。

牛顿迭代格式为

$$\begin{aligned} x_{k+1} &= x_k - \frac{f(x_k)}{f'(x_k)} \\ &= x_k - (e^{x_k} + 1) \end{aligned}$$



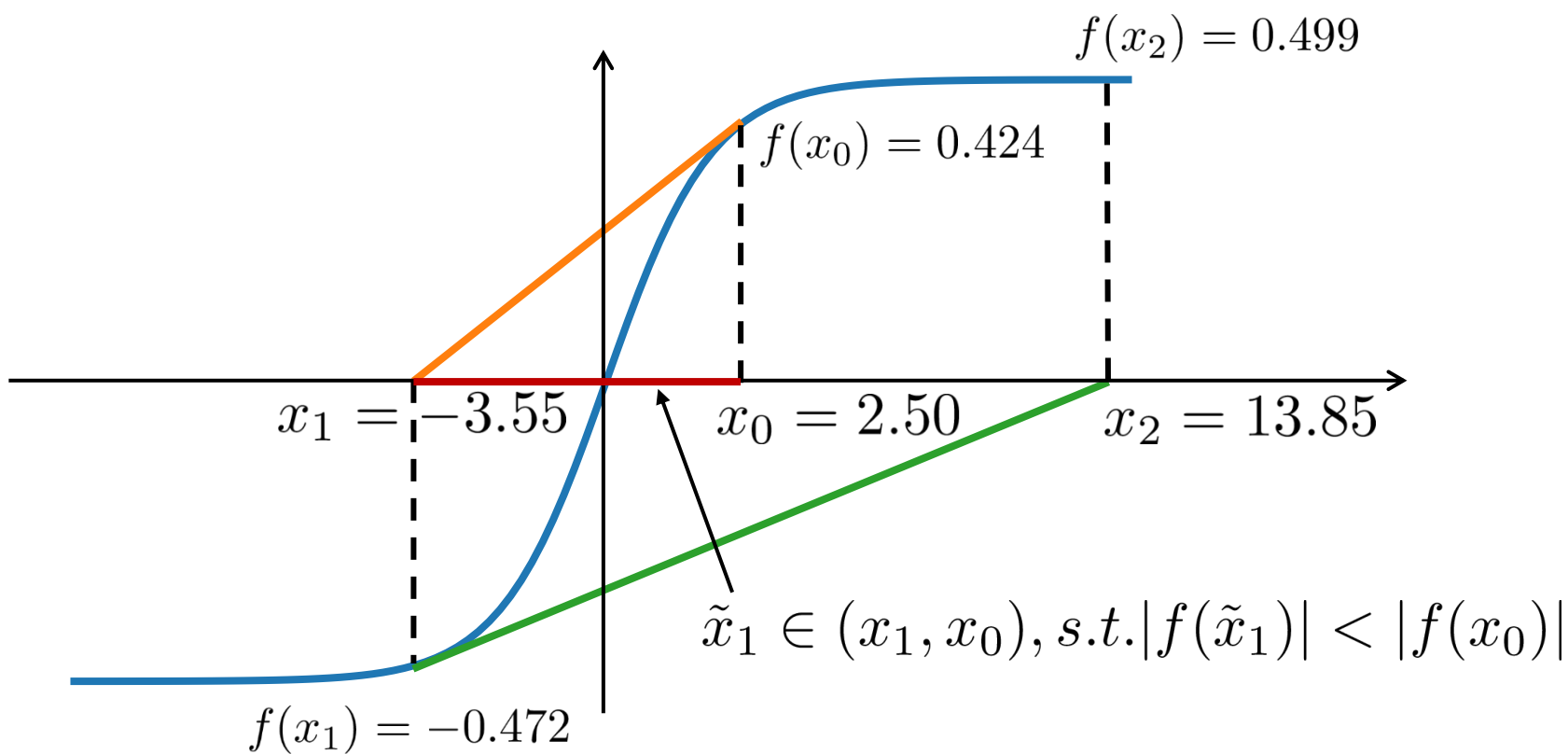
选不同初值迭代结果如下：

初值/k	0	1	2	3	4
$x_0 = 1.5$	1.500 000	-0.629 280	0.042 362	-0.000 013	0.000 000
$x_0 = 2.5$	2.500 000	-3.550 205	13.845 65	-515287.6	inf

7.4.4 牛顿下山法(1)



牛顿下山法的基本思想



7.4.4 牛顿下山法(2)



- ▶ 在下山法保证函数值稳定下降的前提下，用牛顿法加快收敛速度。
- ▶ 先使用牛顿法计算 \bar{x}_{k+1} ，再对 x_k 和 \bar{x}_{k+1} 加权平均得到新的 x_{k+1} 。

$$\bar{x}_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

$$x_{k+1} = \lambda \bar{x}_{k+1} + (1 - \lambda)x_k, 0 < \lambda \leq 1$$

$$|f(x_{k+1})| < |f(x_k)|$$

- ▶ λ 如何确定？

初始 $\lambda = 1$ ，逐次将 λ 减半，直到 $|f(x_{k+1})| < |f(x_k)|$ 成立

7.4.4 牛顿下山法的伪代码



Algorithm: 求解非线性方程 $f(x) = 0$ 的牛顿下山法

Input : $\mathbf{fx}, \mathbf{dfx}$; // 函数 $f(x)$, $f(x)$ 的导函数 $f'(x)$
 x_0, K, ϵ ; // 迭代初始值, 循环最大次数, 迭代终止判断条件

Output: k, x^* ; // 迭代次数及根值

for $k \leftarrow 0$ **to** K **do**

$\bar{x}_1 \leftarrow x_0 - \mathbf{fx}(x_0) / \mathbf{dfx}(x_0)$;

$\lambda \leftarrow 1$;

$x_1 \leftarrow \lambda \bar{x}_1 + (1 - \lambda)x_0$;

while $|\mathbf{fx}(x_1)| \geq |\mathbf{fx}(x_0)|$ **do**

$\lambda \leftarrow \frac{1}{2}\lambda$;

$x_1 \leftarrow \lambda \bar{x}_1 + (1 - \lambda)x_0$;

end while

if $|x_1 - x_0| < \epsilon$ **then**

return k, x_1 ;

end if

$x_0 \leftarrow x_1$;

end for

return ∞, null ; // 超出最大循环次数

思考: 是否存在 $\forall \lambda \in (0, 1], |f(x_1)| \geq |f(x_0)|$ 的情况?

7.5 弦截法

7.5.1 弦截法原理

7.5.2 弦截法的几何意义

7.5.3 弦截法与牛顿法的对比





7.5.1 弦截法原理

- ▶ 牛顿法：计算 $f'(x_k)$ 需要用到导函数的解析形式，而且计算量大
- ▶ 简化牛顿法：收敛性为线性收敛
- ▶ 使用差商近似 $f'(x_k)$

差商：
$$f[x_{k-1}, x_k] = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

- ▶ 弦截法的迭代公式：

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

牛顿法

弦截法

7.5.1 弦截法的伪代码



Algorithm: 求解非线性方程 $f(x) = 0$ 的弦截法

Input : \mathbf{fx} ; // 函数 $f(x)$

x_0, x_1 ; // 2个迭代初始值

K, ϵ ; // 循环最大次数, 迭代终止判断条件

Output: k, x^* ; // 迭代次数及根值

for $k \leftarrow 0$ **to** K **do**

$x_2 \leftarrow x_1 - \mathbf{fx}(x_1) \cdot \frac{x_1 - x_0}{\mathbf{fx}(x_1) - \mathbf{fx}(x_0)};$

if $|x_2 - x_1| < \epsilon$ **then**

return k, x_2 ;

end if

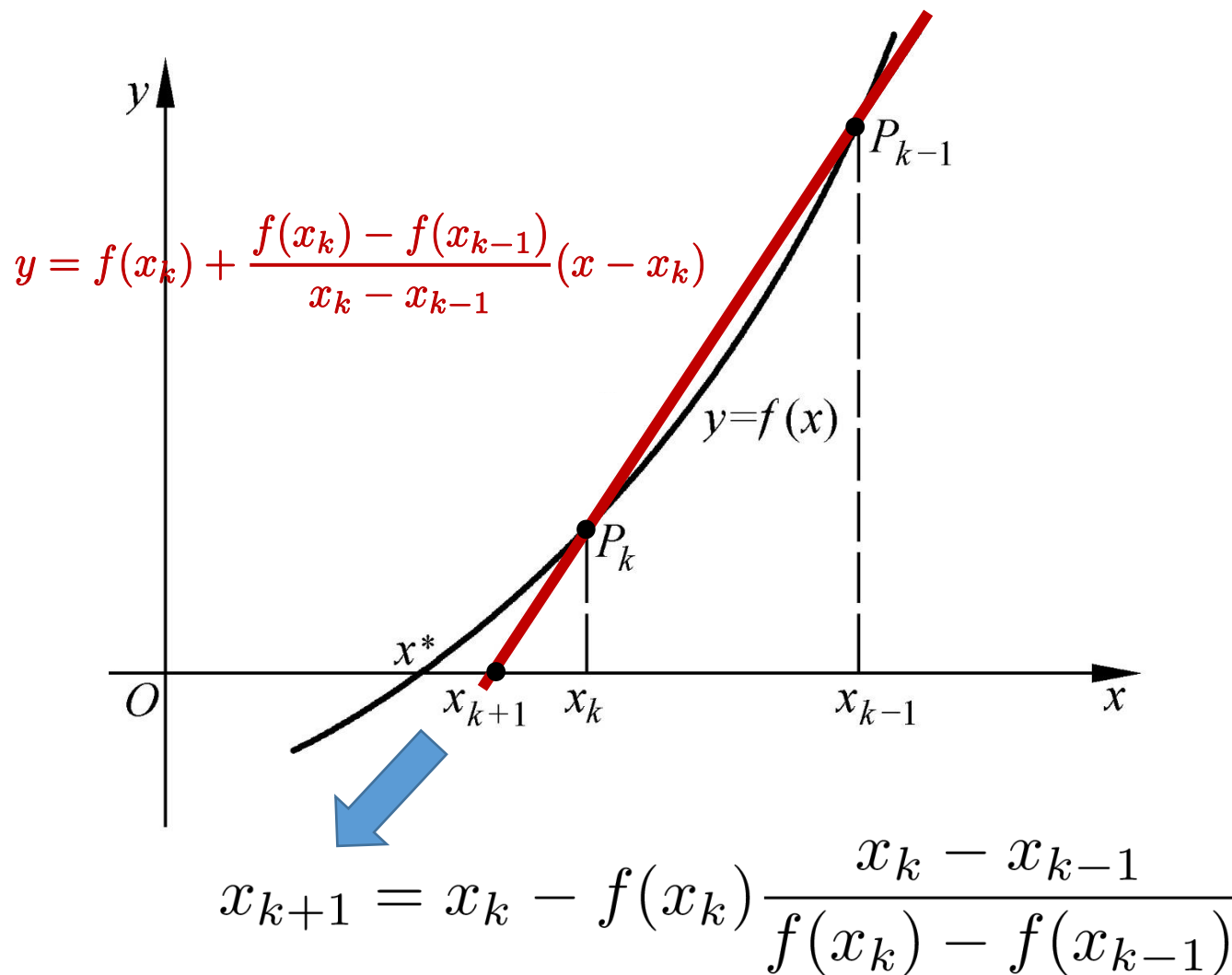
$x_0 \leftarrow x_1$;

$x_1 \leftarrow x_2$;

end for

return ∞, null ; // 超出最大循环次数

7.5.2 弦截法的几何解释



7.5.3 弦截法与牛顿法对比 (1)



- ▶ 相对于牛顿迭代来说，弦截法不需要导数的显示表达式，计算上相对简单。
- ▶ 弦截法需要2个初值才可以开始计算。
- ▶ 弦截法在计算 x_{k+1} 时，除了跟 x_k 相关外，还跟前一步的值 x_{k-1} 相关，称这种格式为**多步格式**。牛顿迭代中，计算 x_{k+1} 只与 x_k 相关，称为**单步格式**。
- ▶ 弦截法的收敛阶要低于牛顿迭代，但高于简化的牛顿迭代，为1.618。

7.5.3 弦截法与牛顿法对比 (2)



- ▶ 从插值近似的角度来看，牛顿迭代和弦截法都是用一个近似多项式的根来作为函数 $f(x)$ 的根的估计
 - ▶ 牛顿迭代是用满足 $f(x)$ 和 $f'(x)$ 的埃尔米特型插值多项式
 - ▶ 弦截法是用满足 $f(x_{k-1})$ 和 $f(x_k)$ 的线性插值多项式
- ▶ 是否可以用3个点，如 $f(x_k)$, $f(x_{k-1})$, $f(x_{k-2})$ 做2次插值多项式来近似原方程的根吗？

可以，该格式被称为**抛物线法**

7.6 非线性方程组的牛顿迭代法

7.6.1 非线性方程组

7.6.2 非线性方程组的牛顿迭代

7.6.3 非线性方程组的不动点迭代





7.6.1 非线性方程组

► 方程组
$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0 \\ f_2(x_1, x_2, \cdots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \cdots, x_n) = 0 \end{cases}$$

► f_1, f_2, \cdots, f_n 均为 (x_1, x_2, \cdots, x_n) 的多元函数

► 记 $\mathbf{x} = [x_1, x_2, \cdots, x_n]^\top \in \mathbb{R}^n$, $\mathbf{F} = [f_1, f_2, \cdots, f_n]^\top$

上述方程组可写为 $\mathbf{F}(\mathbf{x}) = \mathbf{0}$

► 若 f_1, f_2, \cdots, f_n 中至少有一个是自变量 $x_i, i = 1, \cdots, n$ 的非线性函数, 则称该方程组为**非线性方程组**

7.6.2 非线性方程组的牛顿迭代 (1)



► 将 f_1, f_2, \dots, f_n 在 $\mathbf{x}^{(k)} = [x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}]^\top$ 附近作泰勒一阶展开, 有

$$\left\{ \begin{array}{l} f_1(\mathbf{x}^{(k)}) + \underbrace{(x_1 - x_1^{(k)})}_{\text{wavy}} \frac{\partial f_1(\mathbf{x}^{(k)})}{\partial x_1} + \dots + \underbrace{(x_n - x_n^{(k)})}_{\text{wavy}} \frac{\partial f_1(\mathbf{x}^{(k)})}{\partial x_n} = 0 \\ f_2(\mathbf{x}^{(k)}) + \underbrace{(x_1 - x_1^{(k)})}_{\text{wavy}} \frac{\partial f_2(\mathbf{x}^{(k)})}{\partial x_1} + \dots + \underbrace{(x_n - x_n^{(k)})}_{\text{wavy}} \frac{\partial f_2(\mathbf{x}^{(k)})}{\partial x_n} = 0 \\ \vdots \\ f_n(\mathbf{x}^{(k)}) + \underbrace{(x_1 - x_1^{(k)})}_{\text{wavy}} \frac{\partial f_n(\mathbf{x}^{(k)})}{\partial x_1} + \dots + \underbrace{(x_n - x_n^{(k)})}_{\text{wavy}} \frac{\partial f_n(\mathbf{x}^{(k)})}{\partial x_n} = 0 \end{array} \right.$$

$$\Delta x_i^{(k)} = x_i - x_i^{(k)}, i = 1 \dots n$$



7.6.2 非线性方程组的牛顿迭代 (1)

► 将 f_1, f_2, \dots, f_n 在 $\mathbf{x}^{(k)} = [x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}]^\top$ 附近作泰勒一阶展开, 有

$$\left\{ \begin{array}{l} f_1(\mathbf{x}^{(k)}) + \underbrace{\Delta x_i^{(k)}}_{\text{wavy}} \frac{\partial f_1(\mathbf{x}^{(k)})}{\partial x_1} + \dots + \underbrace{\Delta x_n^{(k)}}_{\text{wavy}} \frac{\partial f_1(\mathbf{x}^{(k)})}{\partial x_n} = 0 \\ f_2(\mathbf{x}^{(k)}) + \underbrace{\Delta x_i^{(k)}}_{\text{wavy}} \frac{\partial f_2(\mathbf{x}^{(k)})}{\partial x_1} + \dots + \underbrace{\Delta x_n^{(k)}}_{\text{wavy}} \frac{\partial f_2(\mathbf{x}^{(k)})}{\partial x_n} = 0 \\ \vdots \\ f_n(\mathbf{x}^{(k)}) + \underbrace{\Delta x_i^{(k)}}_{\text{wavy}} \frac{\partial f_n(\mathbf{x}^{(k)})}{\partial x_1} + \dots + \underbrace{\Delta x_n^{(k)}}_{\text{wavy}} \frac{\partial f_n(\mathbf{x}^{(k)})}{\partial x_n} = 0 \end{array} \right.$$

$$\Delta x_i^{(k)} = x_i - x_i^{(k)}, i = 1 \dots n$$



7.6.2 非线性方程组的牛顿迭代 (2)

► 令 $\Delta \mathbf{x}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)} = [\Delta x_1, \Delta x_2 \cdots \Delta x_n]^\top$

$$\mathbf{F}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{x})}{\partial x_1} & \frac{\partial f_n(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_n(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

\mathbf{F} 的雅可比矩阵

上述方程可化为:

$$\mathbf{F}(\mathbf{x}^{(k)}) + \mathbf{F}'(\mathbf{x}^{(k)})\Delta \mathbf{x}^{(k)} = \mathbf{0}$$

7.6.2 非线性方程组的牛顿迭代 (3)



► 求解出 $\Delta \mathbf{x}^{(k)}$ 后, 再令

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta \mathbf{x}^{(k)}$$

► 每步包括计算 (1) 向量函数 $\mathbf{F}(\mathbf{x}^{(k)})$; (2) 矩阵 $\mathbf{F}'(\mathbf{x}^{(k)})$

► 推导过程也可直接将牛顿法应用于方程组

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



$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mathbf{F}'(\mathbf{x}^{(k)})^{-1} \mathbf{F}(\mathbf{x}^{(k)}), k = 0, 1, \dots$$

$$\Leftrightarrow \mathbf{F}(\mathbf{x}^{(k)}) + \mathbf{F}'(\mathbf{x}^{(k)}) \Delta \mathbf{x}^{(k)} = \mathbf{0}, k = 0, 1, \dots$$

7.6.2 非线性方程组例题



►例：用牛顿法求解方程组

$$\begin{cases} x_1^2 - 10x_1 + x_2^2 + 8 = 0, \\ x_1x_2^2 + x_1 - 10x_2 + 8 = 0. \end{cases}$$

►解：

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} x_1^2 - 10x_1 + x_2^2 + 8 \\ x_1x_2^2 + x_1 - 10x_2 + 8 \end{bmatrix} \quad \mathbf{F}'(\mathbf{x}) = \begin{bmatrix} 2x_1 - 10 & 2x_2 \\ x_2^2 + 1 & 2x_1x_2 - 10 \end{bmatrix}$$

选 $\mathbf{x}^{(0)} = (0,0)^T$ ，解线性方程组 $\mathbf{F}'(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} = -\mathbf{F}(\mathbf{x}^{(0)})$

$$\text{即：} \quad \begin{bmatrix} -10 & 0 \\ 1 & -10 \end{bmatrix} \begin{bmatrix} \Delta x_1^{(0)} \\ \Delta x_2^{(0)} \end{bmatrix} = \begin{bmatrix} -8 \\ -8 \end{bmatrix},$$

7.6.2 非线性方程组例题



解得 $\Delta \mathbf{x}^{(0)} = (0.8, 0.88)^T$, $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \Delta \mathbf{x}^{(0)}$

按牛顿迭代法, 有计算结果

表7-12 计算结果

	$\mathbf{x}^{(0)}$	$\mathbf{x}^{(1)}$	$\mathbf{x}^{(2)}$	$\mathbf{x}^{(3)}$	$\mathbf{x}^{(4)}$
$x_1^{(k)}$	0	0.80	0.9917872	0.9999752	1.00000000
$x_2^{(k)}$	0	0.88	0.9917117	0.9999685	1.00000000

7.6.3 非线性方程组的不动点迭代



- ▶ 为求方程组 $F(x) = 0$ ，将它改为迭代形式 $x = \Phi(x)$ ，其中向量函数 $\Phi \in D \subset \mathbb{R}^n$ ，且在定义域 D 上连续，如果有 $x^* = \Phi(x^*)$ ，则称 x^* 为 Φ 的不动点。
- ▶ 根据 $x = \Phi(x)$ 可构造迭代格式 $x^{(k+1)} = \Phi(x^{(k)})$, $k = 0, 1, \dots$

关于其收敛性，类似于一元非线性函数的压缩映象原理，有如下定理。

7.6.3 非线性方程组压缩映象原理



►定理：函数 Φ 定义在区域 $D \subset \mathbb{R}^n$ ，假设：

1. 存在 $D_0 \subset D$ ，对任意 $x \in D_0$ ，有 $\Phi(x) \in D_0$ ；
2. 存在常数 $0 < L < 1$ ，对 $\forall x, y \in D_0$ 有

$$\|\Phi(x) - \Phi(y)\| \leq L\|x - y\|$$

则 $\Phi(x)$ 在 D_0 存在唯一不动点 x^* ，

且对任意 $x^{(0)} \in D_0$ ，由 $x^{(k+1)} = \Phi(x)^{(k)}$ 得到的序列 $\{x^{(k)}\}$ 有 $\lim_{k \rightarrow \infty} x^{(k)} = x^*$ ，并有误差估计

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



7.6.3 非线性方程组的局部收敛条件

►定理：设 Φ 在定义域内有不动点 x^* , Φ 的分量函数有连续偏导且

$$\rho(\Phi'(x^*)) < 1,$$

则存在 x^* 的一个邻域 S , 对任意的 $x^{(0)} \in S$, 迭代法 $x^{(k+1)} = \Phi(x)^{(k)}$ 产生的序列收敛于 x^* .

7.6.3 非线性方程组的不动点迭代



例 用不动点迭代法求解方程组

$$\begin{cases} x_1^2 - 10x_1 + x_2^2 + 8 = 0, \\ x_1x_2^2 + x_1 - 10x_2 + 8 = 0. \end{cases}$$

解 将方程组化为 $\mathbf{x} = \Phi(\mathbf{x})$ 的形式, 其中

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \Phi(\mathbf{x}) = \begin{bmatrix} \varphi_1(\mathbf{x}) \\ \varphi_2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \frac{1}{10}(x_1^2 + x_2^2 + 8) \\ \frac{1}{10}(x_1x_2^2 + x_1 + 8) \end{bmatrix}$$

设 $D = \{(x_1, x_2) | 0 \leq x_1, x_2 \leq 1.5\}$, 不难验证 $0.8 \leq \varphi_1(\mathbf{x}) \leq 1.25$
 $0.8 \leq \varphi_2(\mathbf{x}) \leq 1.2875$, 故有 $\mathbf{x} \in D$ 时 $\Phi(\mathbf{x}) \in D$.

又对一切 $\boldsymbol{x}, \boldsymbol{y} \in D$,

$$|\varphi_1(\boldsymbol{y}) - \varphi_1(\boldsymbol{x})| = \frac{1}{10}|y_1^2 - x_1^2 + y_2^2 - x_2^2| \leq \frac{3}{10}(|y_1 - x_1| + |y_2 - x_2|)$$

$$|\varphi_2(\boldsymbol{y}) - \varphi_2(\boldsymbol{x})| = \frac{1}{10}|y_1y_2^2 - x_1x_2^2 + y_1 - x_1| \leq \frac{4.5}{10}(|y_1 - x_1| + |y_2 - x_2|)$$

由于

$$\Phi'(\mathbf{x}) = \begin{bmatrix} \frac{1}{5}x_1 & \frac{1}{5}x_2 \\ \frac{1}{10}(x_2^2 + 1) & \frac{1}{5}x_1x_2 \end{bmatrix}$$

对一切 $\mathbf{x} \in D$ 都有 $|\frac{\partial \varphi_i(\mathbf{x})}{\partial x_j}| \leq \frac{0.9}{2}$, 故 $\|\Phi'(\mathbf{x})\|_1 \leq 0.9$,

从而有 $\rho(\Phi'(\mathbf{x})) < 1$, 满足定理7的条件.

此外还可看到 $\Phi(\mathbf{x}^*) = \begin{bmatrix} 0.2 & 0.2 \\ 0.2 & 0.2 \end{bmatrix}$, $\|\Phi(\mathbf{x}^*)\|_1 = 0.4 < 1$

故 $\rho(\Phi(\mathbf{x}^*)) \leq 0.4$, 即满足局部收敛定理条件.

小结



- ▶ 7.1 非线性方程求根与二分法
- ▶ 7.2 不动点迭代法及其收敛性
- ▶ 7.3 迭代收敛的加速方法
- ▶ 7.4 牛顿法、简化牛顿法与牛顿下山法
- ▶ 7.5 弦截法
- ▶ 7.6 非线性方程组的数值解法

