第六章　方程求根

引言、从多项式方程求根说起

$ax+b=0 \quad a\neq 0$ 根：$x=-\dfrac{b}{a}$

$ax^2+bx+c=0 \quad a\neq 0$

◆ 20世纪考古发现，公元前1700年，美索不达米亚人已经有了解二次方程的成法；用现代的代数语言来叙述就是：

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

$$ax^3+bx^2+cx+d=0 \quad a\neq 0$$

- 意大利数学家费罗(S. d. Ferro，1465～1526)首先得到了该方程的一般求根公式，没有公开他的解法，按当时的习俗作为挑战对手的秘密武器；

- Ferro在临终前将解法传给了他的学生安东尼奥·菲奥尔(Antonio M. Fior)；

- 费罗去世后，菲奥尔向当时意大利最大的数学家之一塔尔塔利亚(Tartaglia，1500—1557)提出挑战，要他解出30个三次方程,塔尔塔利亚用8天时间解出了全部30个方程，得到了解缺项三次方程的一般方法。

◆ 米兰的数学和物理教授卡尔达诺(Cardano，1501—1576)获悉该事后央求塔尔塔利亚将密诀告诉他，并发誓保密，在卡尔达诺的恳求下，塔尔塔利亚把他的方法写成一首晦涩的诗告诉了卡尔达诺；

◆ 1545年卡尔达诺出版著作《大法》(Arsmagna)，公布了一般三次方程求根公式，称为卡尔达诺公式；

◆《大法》同时公布了意大利数学家费拉里（Ferrali，1522-1565）仿照一般三次方程求根思想，推导的一般四次方程求根公式；

◆ 从二次方程到三、四次方程求根公式历经至少3245年；

◆ 意大利数学家的成功促使当时的众多数学家开始寻求更高次方程的解法；

◆ 量变引起了质变，数学家们徒劳了两个多世纪，没有成功；

◆ 1771年法国数学家拉格朗日在论文《关于代数方程解法的思考》中指出，用代数运算解一般的(n>4)次方程是不可能的，或者这个问题超出了人类的智力范围，或者是根的表达方式不同于当时所知道的一切；

◆ 1824年，天才的挪威数学家阿贝尔(Abel，1802—1829)在其出版的著作中证明：如果方程的次数n≥5，并且将方程的系数看成字母，那么任何一个由这些字母组成的根式都不可能是方程的根；

◆ 近300年的努力果然是徒劳的；

◆ 阿贝尔之后，不少人找到了特殊高次方程的求根方法，得到了有理根式形式的解；

埃尔米特：阿贝尔留下的思想可供数学家们工作150年

◆ 到底哪些方程可以得到有理根式形式的解？

◆ 1831年天才的法国数学家伽罗瓦(E.Galois，1811—1832)给出了高次方程存在根式解的充分必要条件。

天才的伽罗华

- 1829年，伽罗华中学毕业前，把关于群论的初步研究结果的论文提交给法国科学院，科学院委托当时法国最杰出的数学家柯西审核论文。
- 在1830年1月18日柯西计划对伽罗华的研究成果在科学院举行一次全面的意见听取会。他在一封信中写道：“今天我应当向科学院提交一份关于年轻的伽罗华的工作报告……但因病在家，我很遗憾未能出席今天的会议，希望安排我参加下次会议，讨论已指明的议题。”

E. Galois (法国)
1811－1832

◆ 第二周，柯西向科学院宣读他自己的一篇论文时，忘记了原来的议题。

◆ 1830年2月，伽罗华将论文寄给当时的科学院终身秘书傅立叶，傅立叶于当年5月去世，在他的遗物中未发现伽罗华的手稿。伽罗华递交的两次数学论文均被遗失。

◆ 1831年1月，伽罗华将包含新成果的论文提交给法国科学院，负责审查的数学家泊松（Possion),四个月后，以“完全不能理解”，建议科学院退稿。

◆ 1831年1月8日，因伽罗华揭发校长的政治两面派行为，被皇家国民教育委员会批准开除出巴黎师范大学；

- ◆ 1831年5月10日，伽罗华以“企图暗杀国王”的罪名被捕，关押在圣佩拉吉监狱；
- ◆ 1832年3月16日伽罗华获释后不久，为了一个舞女决定为“爱情与荣誉”决斗；
- ◆ 1832年5月29日夜，伽罗华仓促地把自己生平的数学研究心得扼要写出，附以论文手稿，并在给朋友舍瓦利叶的信中说：“我在分析方面做出了一些新发现。有些是关于方程论的；有些是关于整函数的……。公开请求雅可比或高斯，不是对这些定理的正确性，而是对这些定理的重要性发表意见。我希望将来有人发现，这些对于消除所有有关的混乱是有益的。”
- ◆ 1832年5月30日上午，伽罗华死于决斗；

◆ 伽罗华死后，舍瓦利叶把他的信发表在《百科评论》中。

◆ 1846年，法国数学家刘维尔领悟到伽罗华的天才思想，他花了几个月的时间将伽罗华手稿中的部分内容发表在他的极有影响的《纯粹与应用数学杂志》上，并向数学界推荐。

◆ 1870年法国数学家约当根据伽罗华的思想，写了《论置换与代数方程》一书，向人类展示了跨越世纪的伽罗华思想：关于群和域的理论。这套理论创立了抽象代数学，把代数学的研究推向了一个新的里程，并标志着数学发展现代阶段的开始。

开拓者总是孤独的

例如代数方程 $x^5-x^3+24x+1=0$,

超越方程 $\sin(5x^2)+e^{-x}=0$.

方程是在科学研究中不可缺少的工具, $f(x)=0$

方程求解是科学计算中一个重要的研究对象.

几百年前就已经找到了代数方程中二次至四次方程的求解公式;

但是,对于更高次数的代数方程目前仍无有效的精确解法;

对于无规律的非代数方程的求解也无精确解法.

因此, 研究非线性方程的数值解法成为必然.

主要研究单根区间上方程求根的各种近似算法.

一、方程求根与二分法

本章主要讨论单变量非线性方程

$$f(x)=0 \tag{1.1}$$

的求根问题，这里$x\in R$, $f(x)\in C[a, b]$. 在科学与工程计算中有大量方程求根问题，

1. 一类特殊的问题是多项式方程

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n \quad (a_0\neq 0), \tag{1.2}$$

其中系数$a_i(i=0,1,\cdots,n)$为实数.

2. 另一类就是超越方程的求根。

$$cos(x)cosh(x)+1=0$$

方程$f(x)=0$的根x^*，又称为函数$f(x)$的零点，它使得$f(x^*)=0$，若$f(x)$可分解为

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

其中m为正整数，且$g(x^*)\neq 0$. 当$m=1$时，则称x^*为单根，若$m>1$称x^*为(1.1)的m重根，或x^*为函数$f(x)$的m重零点. 若x^*是$f(x)$的m重零点，且$g(x)$充分光滑，则

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

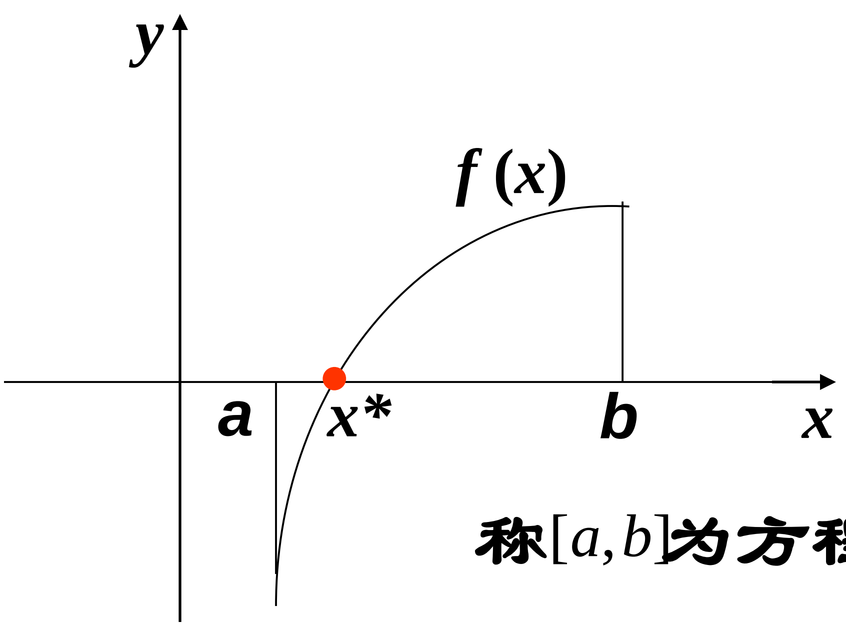
当$f(x)$为代数多项式(1.2)时，根据代数基本定理可知，n次代数方程$f(x)=0$在复数域有且只有n个根(含复根，m重根为m个根).

$n=1,2$时方程的根是大家熟悉的，$n=3,4$时虽有求根公式但比较复杂，可在数学手册中查到，但已不适合数值计算，而$n\geq 5$时就不能用公式表示方程的根.因此，通常对$n\geq 3$的多项式方程求根与一般连续函数方程(1.1)一样都可采用迭代法求根.

迭代法要求给出根x^*的一个近似，若$f(x)\in C[a, b]$且$f(a)f(b)<0$，根据连续函数性质中的介值定理可知方程$f(x)=0$在(a, b)内至少有一个实根，这时称$[a, b]$为方程(1.1)的有根区间，通常可通过逐次搜索法求得方程(1.1)的有根区间.

求 $f(x) = 0$ 的根

原理：若 $f \in C[a, b]$，且 $f(a) \cdot f(b) < 0$，则 f 在 (a, b) 上必有一根。



称$[a,b]$为方程的有根区间。

求隔根区间的一般方法

若 $f(x)$在$[a,b]$内连续, 且 $f(a) \cdot f(b)<0$, 则 $f(x)=0$ 在$[a,b]$内必有根; 若$f(x)$在$[a,b]$内还严格单调, 则$f(x)=0$在$[a,b]$内只有一根, 据此可得求隔根区间的两种方法.

1. (描)做图法

画出 $y=f(x)$ 的草图, 由$f(x)$与横轴交点的大概位置来确定隔根区间; 或者利用导函数$f'(x)$的正、负与函数$f(x)$的单调性的关系确定根的大概位置.

若$f(x)$比较复杂, 还可将方程$f(x)=0$化为一个等价方程$\varphi(x)=\psi(x)$, 则曲线$y=\varphi(x)$与$y=\psi(x)$之交点$A(x^*,y^*)$的横坐标 x^*即为原方程之根, 据此也可通过作图求得x^*的隔根区间.

例1 判别下列方程有几个实根，并求隔根区间.

(1) $f(x)=x^3-x-1=0$, (2) $f(x)=x^4-4x^3+1=0$.

解 (1)将方程变形为

$$x^3=x+1$$

绘曲线图 $y=x^3$ 及 $y=x+1$

由图可知, 方程只有一个实根$x^*\in(1, 1.5)$，所以(1, 1.5)即为其隔根区间.

(2) 方程 $f(x)=x^4-4x^3+1=0$.

由$f'(x)= 4x^2(x-3)=0$ 得驻点$x_1=0$, $x_2=3$.

该二点将实轴分为三个区间:

$$(-\infty, 0),\quad (0, 3),\quad (3, +\infty)$$

$f'(x)$在此三个区间上的符号分别为"−"、"−"、"+"，

又知 $f(-\infty)>0$, $f(0)=1>0$, $f(3)=-26<0$, $f(+\infty)>0$.

可见$f(x)$仅有两个实根, 分别位于$(0, 3)$, $(3, +\infty)$,
又$f(4)=1>0$, 所以第二根的隔根区间可缩小为$(3, 4)$.

以上分析可用下表表示

x	$(-\infty,0)$	0	$(0,3)$	3	$(3,4)$	4	$(4,+\infty)$
$f'(x)$ $f(x)$	− ↘	0 +	− ↘	0 −	+ ↗	+ +	+ ↗
隔根区间			$(0,3)$		$(3,4)$		

2. 逐步搜索法

从区间$[a, b]$的左端点 a 出发，按选定的步长h一步步向右搜索，若

$$f(a+jh)\, f(a+(j+1)h)<0 \qquad (j=0,1,2,\cdots)$$

则区间$[a+jh, a+(j+1)h]$内必有根. 搜索过程也可从b开始，这时应取步长 $h<0$.

二分法

设$f(x)$在区间$[a, b]$上连续, $f(a)\ f(b)<0$, 则在$[a, b]$内有方程的根. 取$[a, b]$的中点 $x_0=\frac{1}{2}(a+b)$,

将区间一分为二. 若 $f(x_0)=0$, 则x_0就是方程的根, 否则判别根 x^*在 x_0 的左侧还是右侧.

若$f(a)\ f(x_0)<0$, 则$x^*\in(a, x_0)$, 令 $a_1=a$, $b_1=x_0$;

若$f(x_0)\ f(b)<0$, 则$x^*\in(x_0, b)$, 令 $a_1=x_0$, $b_1=b$.

不论出现哪种情况, (a_1, b_1)均为新的有根区间, 它的长度只有原有根区间长度的一半, 达到了压缩有根区间的目的.

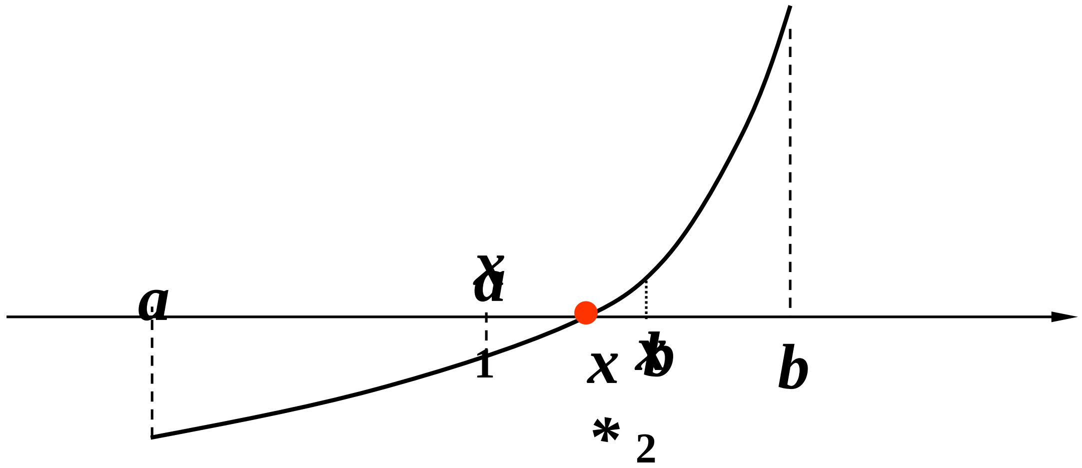
对压缩了的有根区间, 又可实行同样的步骤, 再压缩. 如此反复进行, 即可得一系列有根区间套

$$[a,b]\supset[a_1,b_1]\supset\cdots\supset[a_n,b_n]\supset\cdots$$

由于每一区间都是前一区间的一半，因此区间 $[a_n,b_n]$ 的长度为

$$b_n-a_n=\frac{1}{2^n}(b-a)$$

若每次二分时所取区间中点都不是根，则上述过程将无限进行下去. 当 $n\to\infty$ 时，区间必将最终收缩为一点 x^*，显然 x^* 就是所求的根.



给定有根区间 $[a, b]$ $(f(a) \cdot f(b) < 0)$ 和 精度 ε 或 δ

1. 令 $x = (a+b)/2$

2. 如果 $b - a < \varepsilon$ 或 $f(x) < \delta$, 停机，输出 x

3. 如果 $f(a) f(x) < 0$，则令 $b = x$，否则令 $a = x$, 返回第1步

用二分法求根，通常先给出 $f(x)$ 草图以确定根的大概位置。

若取区间$[a_n, b_n]$的中点

$$x_n = \frac{1}{2}(a_n + b_n)$$

作为x^*的近似值，则有下述误差估计式

$$\left| x^* - x_n \right| \le \frac{1}{2}(b_n - a_n) = \frac{1}{2^{n+1}}(b - a)$$

$$x^*, x_n \in (a_{n+1}, b_{n+1})$$

只要 n 足够大, (即区间二分次数足够多)，误差就可足够小.

由于在偶重根附近曲线 $y=f(x)$ 为上凹或下凸，即 $f(a)$与$f(b)$的符号相同，因此不能用二分法求偶重根.

记 $a_1 = a,\ b_1 = b$，第 k 步的有根区间为 $[\,a_k, b_k\,]$

$$\left|x_k - x^*\right| = \left|\frac{b_k + a_k}{2} - x^*\right| \le \frac{b_k - a_k}{2} \le \frac{b_{k-1} - a_{k-1}}{4} \le \cdots \le \frac{b_1 - a_1}{2^k}$$

对于给定的精度 ε,可估计二分法所需的步数 k：

$$\frac{b-a}{2^k} < \varepsilon \quad \Rightarrow \quad k > \log_2 \frac{b-a}{\varepsilon}, \qquad \text{取}\ \ k = \left\lceil \log_2 \frac{b-a}{\varepsilon} \right\rceil$$

- ✓ 简单易用
- ✓ 对 $f(x)$ 要求不高，只要连续即可收敛

- ✓ 收敛速度慢
- ✓ 无法求复根及偶重根

例2 用二分法求例1中方程 $f(x)=x^3-x-1=0$的实根, 要求误差不超过0.005.

解 由例1可知$x^* \in (1, 1.5)$, 要想满足题意，即：

$$|x^*-x_n|\le 0.005$$

则要

$$\frac{1}{2^{n+1}}(b-a)=\frac{1}{2^{n+1}}(1.5-1)=\frac{1}{2^{n+2}}\le 0.005$$

由此解得$n\ge\frac{2}{\lg 2}-1\approx 5.6$, 取$n=6$, 按二分法计算过程见下表, $x_6=1.3242$ 为所求之近似根.

n	a_n	b_n	x_n	$f(x_n)$	说明
0	1.0	1.5	1.25	−	(1) $f(a)<0$,
1	1.25	1.5	1.375	+	$f(b)>0$
2	1.25	1.375	1.3125	−	(2) 根据精
3	1.3125	1.375	1.3438	+	度要求，
4	1.3125	1.3438	1.3281	+	取到小数
5	1.3125	1.3281	1.3203	−	点后四位
6	1.3203	1.3281	1.3242	−	即可.

二分法的优点是算法简单，且总是收敛的，缺点是收敛得太慢，故一般不单独将其用于求根，只是用其为根求得一个较好的近似值.

二分法的计算步骤:

步骤1 准备 计算函数$f(x)$在区间$[a, b]$端点处的值$f(a)$, $f(b)$.

步骤2 二分 计算函数$f(x)$在区间中点$(a+b)/2$处的值$f((a+b)/2)$.

步骤3 判断 若$f((a+b)/2)=0$，则$(a+b)/2$即是根，计算过程结束，否则检验.

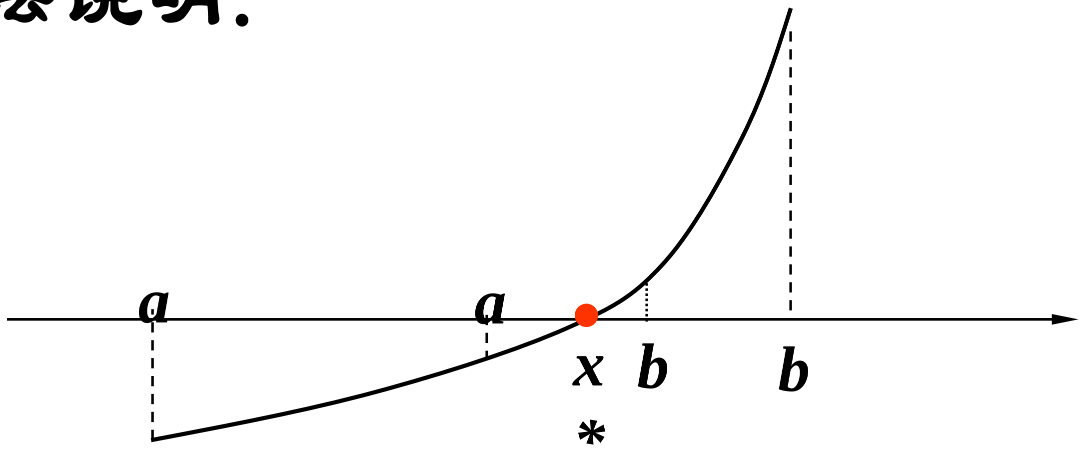
若$f(a)\ f((a+b)/2)<0$, 则以$(a+b)/2$代替b，否则以$(a+b)/2$代替a.

反复执行步骤2和步骤3，直到区间$[a, b]$长度小于允许误差ε，此时中点$(a+b)/2$即为所求近似根.

例： 求 $f(x)=x^3-x-1=0$ 在 (1, 1.5) 的实根，要求误差不超过0.005。

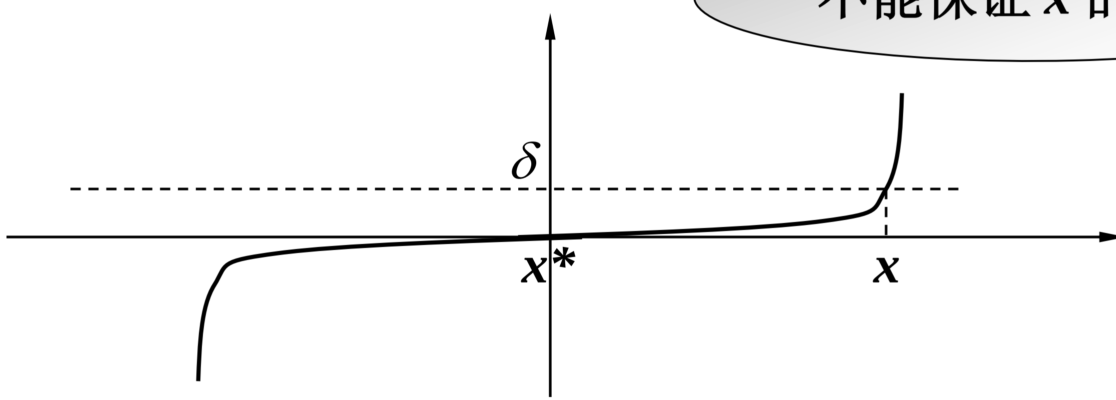
STEP 0	输入 $a, b, eps, delta$, $fa=f(a)$, $fb=f(b)$
STEP 1	$x=(a+b)/2$, $fx=f(x)$
STEP 2	判断：$\lvert b-a\rvert<eps$ or $\lvert f\,x\rvert< delta$ 若是，goto step 4 ;否则，执行下一步
STEP 3	若 $fb*fx<0$,则 $a=x$ 否则 $b=x$. goto step 1
STEP 4	输出 x, fx, 停机.

程序算法说明：



$$\left|x_{k+1}-x_k\right|<\varepsilon \qquad \text{或} \qquad \left|f(x)\right|<\delta$$

不能保证 *x* 的精度



二、迭代法及其收敛性

不动点迭代法

将方程$f(x)=0$改写为等价方程形式

$$x=\varphi(x). \qquad (2.1)$$

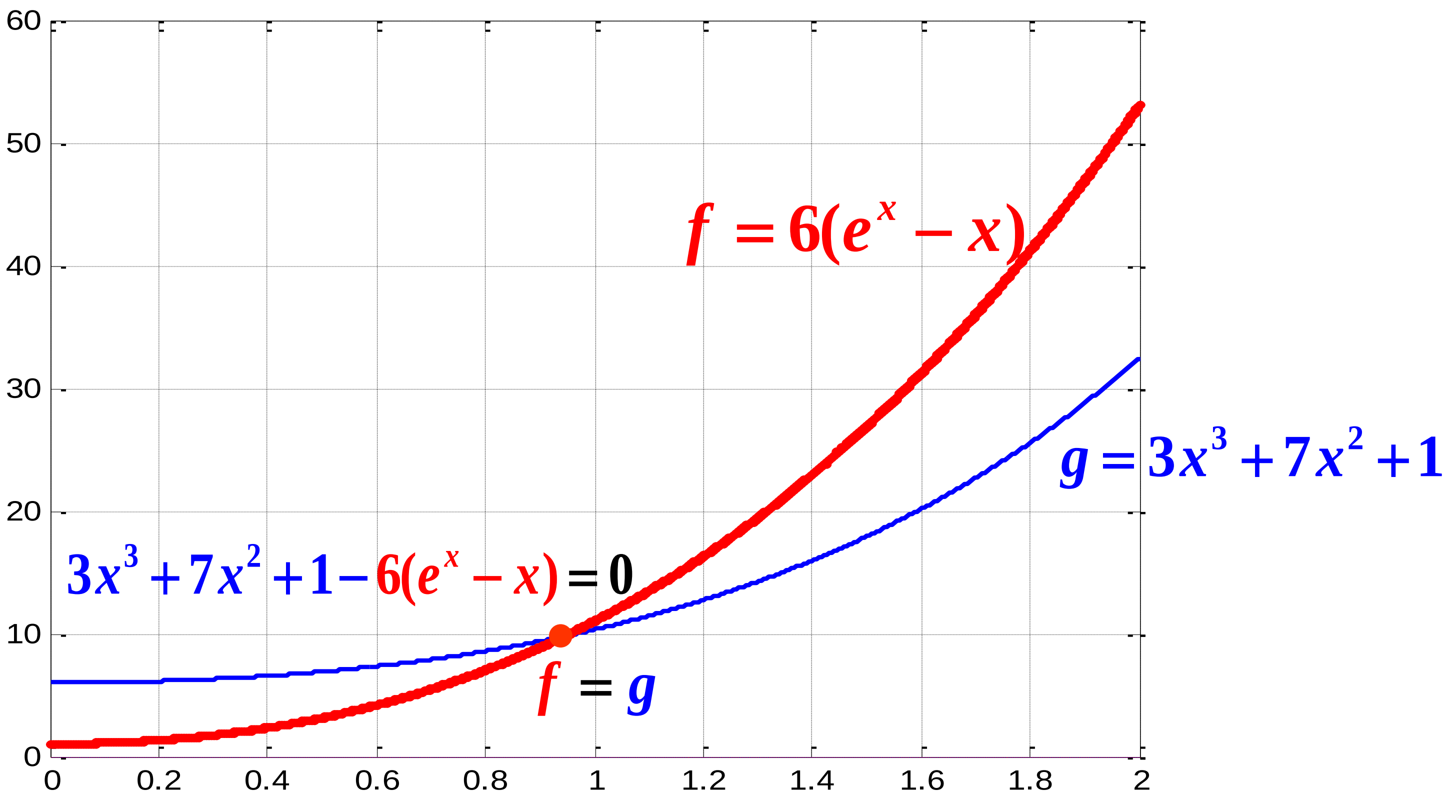
若要求x^*满足$f(x^*)=0$，则$x^*=\varphi(x^*)$；反之亦然，称x^*为函数$\varphi(x)$的一个不动点. 求$f(x)$的零点就等于求$\varphi(x)$的不动点。选择一个初始近似值x_0，将它代入(2.1)右端，即可求得

$$x_1=\varphi(x_0).$$

基本原理

$$f(x)=0 \xleftrightarrow{\text{等价变换}} x=\varphi(x)$$

$$f(x)\text{ 的根} \longleftrightarrow \varphi(x)\text{ 的不动点}$$



基本原理

$$f(x)=0 \xleftrightarrow{\text{等价变换}} x=\varphi(x)$$

$f(x)$ 的根 $\Longleftrightarrow$ $\varphi(x)$的不动点

思路　　从一个初值 x_0 出发，计算 $x_1=\varphi(x_0)$，$x_2=\varphi(x_1), \ldots,$ $x_{k+1}=\varphi(x_k), \ldots$ 若 $\{x_k\}_{k=0}^{\infty}$ 收敛，即存在 x^* 使得 $\lim\limits_{k\to\infty} x_k = x^*$，且 φ 连续，则由 $\lim\limits_{k\to\infty} x_{k+1} = \lim\limits_{k\to\infty} \varphi(x_k)$ 可知 $x^*=\varphi(x^*)$，即x^* 是 φ 的不动点，也就是 f 的根。

可以如此反复迭代计算

$$x_{k+1}=\varphi(x_k) \quad (k=0,1,2,\cdots). \qquad (2.2)$$

$\varphi(x)$称为迭代函数. 如果对任何$x_0\in[a, b]$，由(2.2)得到的序列$\{x_k\}$有极限

$$\lim_{k\to\infty} x_k = x^*.$$

则称迭代方程(2.2)收敛. 且$x^*=\varphi(x^*)$为$\varphi(x)$的不动点，故称(2.2)为不动点迭代法.

上述迭代法是一种逐次逼近法，其基本思想是将隐式方程(2.1)归结为一组显式的计算公式(2.2)，迭代过程实质上是一个逐步显式化过程.

$$x_{k+1}=\varphi(x_k)\quad (k=0,1,2,\cdots)$$

$$\lim_{k\to\infty} x_k = x^*.$$

当$\varphi(x)$连续时，显然x^*就是方程$x=\varphi(x)$之根(不动点).

于是可以从数列$\{x_k\}$中求得满足精度要求的近似根.

这种求根方法称为不动点迭代法,

$$x_{k+1}=\varphi(x_k)\quad (k=0,1,2,\cdots)$$

称为迭代格式, $\varphi(x)$称为迭代函数, x_0 称为迭代初值,

数列$\{x_k\}$称为迭代序列. 如果迭代序列收敛, 则称迭代格式收敛,否则称为发散.

将 $f(x)=0$ 转化为 $x=\varphi(x)$ 的方法有多种多样，

例： $f(x)=x^3+4x^2-10=0$ 在 $[1,2]$ 上可有以下方法：

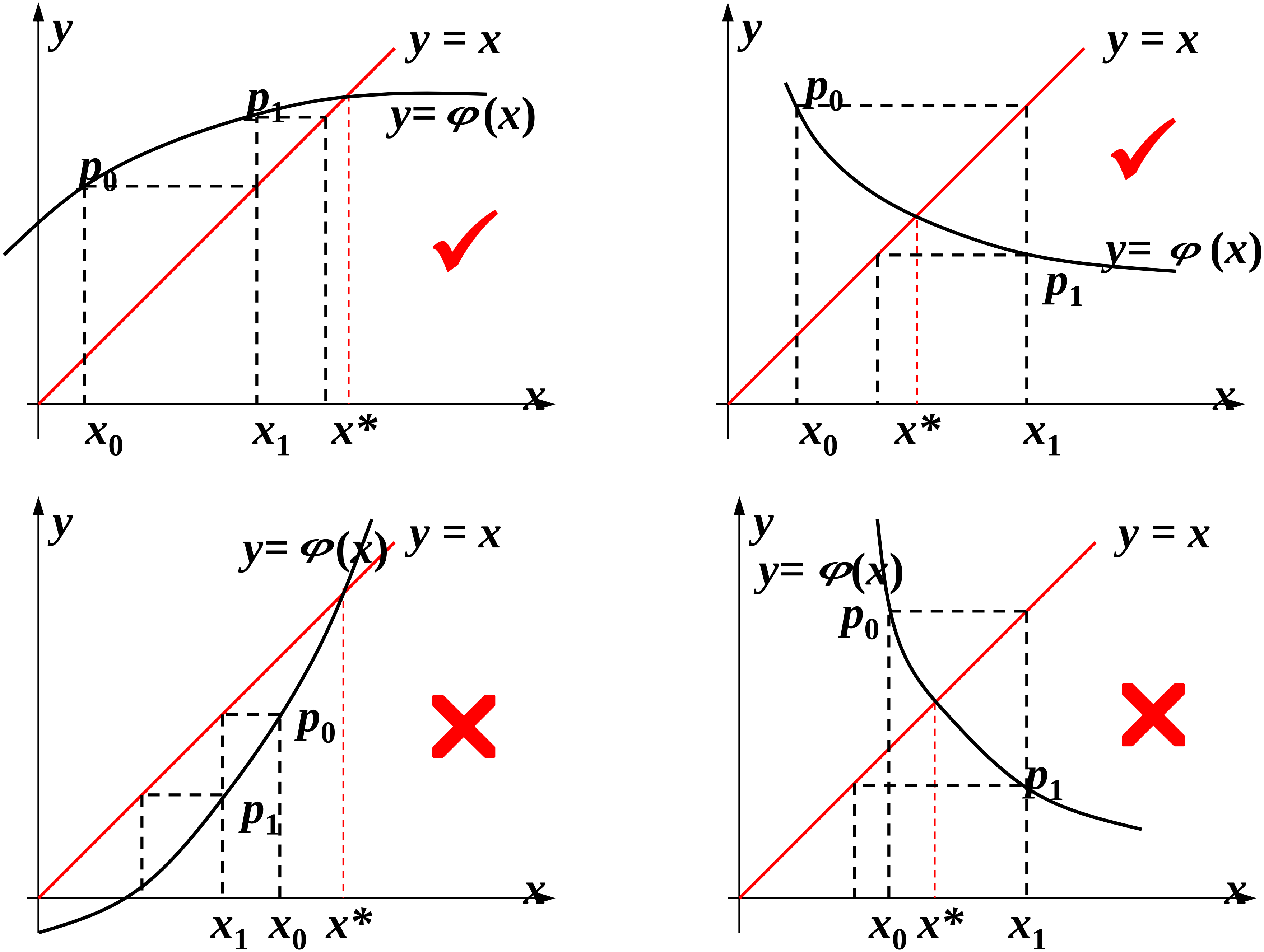
(1) $x=x-x^3-4x^2+10$

(2) $x=(1/2)(10-x^3)^{1/2}$

(3) $x=(10/x-4x)^{1/2}$

(4) $x=[10/(4+x)]^{1/2}$

取 $x_0=1.5$，有的收敛、有的发散、有的快、有的慢。



y
y = x
y= φ(x)
p1
p0
x
x0
x1
x*
y
y = x
p0
y= φ (x)
p1
x
x0
x*
x1
y
y= φ(x)
y = x
p0
p1
x
x1
x0
x*
y
y = x
y= φ(x)
p0
p1
x
x0
x*
x1

例3 用迭代法求方程$x^4+2x^2-x-3=0$ 在区间$[1, 1.2]$内的实根.

解 对方程进行如下三种变形:

$$x^4+2x^2-x-3=0 \Rightarrow \begin{cases} x=\varphi_1(x)=(3+x-2x^2)^{\frac{1}{4}} \\ x=\varphi_2(x)=\sqrt{\sqrt{x+4}-1} \\ x=\varphi_3(x)=x^4+2x^2-3 \end{cases}$$

分别按以上三种形式建立迭代公式，并取$x_0=1$进行迭代计算，结果如下:

$$x_{k+1}=\varphi_1(x_k)=(3+x_k-2x_k^2)^{\frac{1}{4}}$$

$$x_{26}=x_{27}=1.124123$$

$$x_{k+1}=\varphi_2(x_k)=\sqrt{\sqrt{x_k+4}-1}$$

$$x_6=x_7=1.124123$$

$$x_{k+1}=\varphi_3(x_k)=x_k^4+2x_k^2-3$$

$$x_3=96,\quad x_4=8.495307\times10^7$$

准确根 $x^*=1.124123029$, 可见迭代公式不同, 收敛情况也不同. 第二种公式比第一种公式收敛快得多, 而第三种公式不收敛.

迭代法——算例分析

例如：用迭代法求解方程 $2x^3 - x - 1 = 0$

解1： 将原方程化为等价方程

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

取初值 $x_0 = 0$

$$x_1 = 2x_0^3 - 1 = -1$$

$$x_2 = 2x_1^3 - 1 = -3$$

$$x_3 = 2x_2^3 - 1 = -55$$

………

显然迭代法发散

（2）如果将原方程化为等价方程

$$2x^3 - x - 1 = 0 \longleftrightarrow x = \sqrt[3]{\frac{x+1}{2}}$$

仍取初值

$$x_0 = 0 \qquad x_1 = \sqrt[3]{\frac{x_0 + 1}{2}} = \sqrt[3]{\frac{1}{2}} \approx 0.7937$$

$$x_2 = \sqrt[3]{\frac{x_1 + 1}{2}} = \sqrt[3]{\frac{1.7937}{2}} \approx 0.9644$$

依此类推, 得

$$\begin{aligned} x_2 &= 0.9644 \\ x_3 &= 0.9940 \\ x_4 &= 0.9990 \\ x_5 &= 0.9998 \\ x_6 &= 1.0000 \\ x_7 &= 1.0000 \end{aligned}$$

同样的方程
不同的迭代格式
有不同的结果

迭代函数的构造有关

已经收敛, 故原方程的解为

$$x = 1.0000$$

什么形式的迭代法
能够收敛呢?

例3表明原方程化为(2.1)的形式不同，有的收敛，有的不收敛，有的发散，只有收敛的的迭代过程(2.2)才有意义，为此我们首先要研究$\varphi(x)$的不动点的存在性及迭代法(2.2)的收敛性.

不动点的存在性与迭代法的收敛性

首先考察$\varphi(x)$在$[a, b]$上不动点的存在唯一性.

定理1 设$\varphi(x)\in C[a, b]$满足以下两个条件:

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

2° 存在正数$L<1$，使对任意$x,y\in[a, b]$都有

$$|\varphi(x)-\varphi(y)|\le L|x-y|. \qquad (2.4)$$

则$\varphi(x)$在$[a, b]$上存在唯一的不动点x^*.

证明 先证不动点的存在性. 若$\varphi(a)=a$或$\varphi(b)=b$，显然$\varphi(x)$在$[a, b]$上存在不动点. 因为$a\leqslant\varphi(x)\leqslant b$，以下设$\varphi(a)>a$及$\varphi(b)<b$定义函数

$$f(x)=\varphi(x)-x.$$

显然$f(x)\in C[a,b]$,且满足$f(a)=\varphi(a)-a>0$, $f(b)=\varphi(b)-b<0$,

由连续函数性质可知存在 $x^*\in(a,b)$ 使 $f(x^*)=0$，即

$x^*=\varphi(x^*)$，x^*即为$\varphi(x)$的不动点.

再证不动点的唯一性. 设$x_1^*, x_2^*\in[a,b]$都是$\varphi(x)$

的不动点，则由(2.4)得

$$|x_1^*-x_2^*|=|\varphi(x_1^*)-\varphi(x_2^*)|\le L|x_1^*-x_2^*|<|x_1^*-x_2^*|.$$

引出矛盾，故$\varphi(x)$的不动点只能是唯一的. 证毕.

在$\varphi(x)$的不动点存在唯一的情况下，可得到迭代法(2.2)收敛的一个充分条件.

定理2 设$\varphi(x)\in C[a,b]$满足定理1中的两个条件，则对任意$x_0\in[a,b]$，由(2.2)得到的迭代序列$\{x_k\}$收敛到的不动点x^*，并有误差估计式

$$\left|x^*-x_k\right|\le\frac{L^k}{1-L}\left|x_1-x_0\right|. \tag{2.5}$$

$$\left|x^*-x_k\right|\le\frac{L}{1-L}\left|x_{k+1}-x_k\right|. \tag{2.6}$$

证明 设$x^*\in[a,b]$是$\varphi(x)$在$[a,b]$上的唯一不动点，由条件1°，可知$\{x_k\}\in[a,b]$，再由(2.4)得

$$\left|x^*-x_k\right|=\left|\varphi(x^*)-\varphi(x_{k-1})\right|$$

$$\le L\left|x^*-x_{k-1}\right|\le\cdots\le L^k\left|x^*-x_0\right|.$$

因$0<L<1$，故当$k\to\infty$时序列$\{x_k\}$收敛到x^*.

下面证明估计式(2.5)，由(2.4)有

$$|x_{k+1}-x_k|=|\varphi(x_k)-\varphi(x_{k-1})|\le L|x_k-x_{k-1}|$$

$$\le L^2|x_{k-1}-x_{k-2}|\le\cdots\le L^k|x_1-x_0|.$$

于是对任意正整数p有

$$|x_{k+p}-x_k|\le|x_{k+p}-x_{k+p-1}|+|x_{k+p-1}-x_{k+p-2}|$$

$$+\cdots+|x_{k+1}-x_k|$$

$$\le(L^{k+p-1}+L^{k+p-2}+\cdots+L^k)|x_1-x_0|$$

$$=\frac{L^k(1-L^p)}{1-L}|x_1-x_0|\le\frac{L^k}{1-L}|x_1-x_0|.$$

上述令$p\to\infty$, 注意到$\lim x_{k+p}=x^*\ (p\to\infty)$即得(2.5)式.

又由于对任意正整数p有

$$
\begin{aligned}
\left|x_{k+p}-x_k\right| &\le \left|x_{k+p}-x_{k+p-1}\right|+\left|x_{k+p-1}-x_{k+p-2}\right| \\
&\quad +\cdots+\left|x_{k+1}-x_k\right| \\
&\le (L^{p-1}+L^{p-2}+\cdots+L)\left|x_{k+1}-x_k\right| \\
&= \frac{1-L^p}{1-L}\left|x_{k+1}-x_k\right| \le \frac{1}{1-L}\left|x_{k+1}-x_k\right|.
\end{aligned}
$$

上述令$p\to\infty$, 及$\lim x_{k+p}=x^*\,(p\to\infty)$即得(2.6)式.　证毕.

迭代过程是个极限过程. 在用迭代法进行时，必须按精度要求控制迭代次数. 误差估计式(2.5)原则上确定迭代次数，但它由于含有信息L而不便于实际应用. 而误差估计式(2.6)是实用的，只要相邻两次计算结果的偏差足够小即可保证近似值x_k具有足够精度.

对定理1和定理2中的条件2°可以改为导数，即在使用时如果$\varphi(x)\in C[a, b]$且对任意$x\in[a, b]$有

$$\left|\varphi'(x)\right|\le L<1. \tag{2.7}$$

则由微分中值定理可知对任意$x, y\in[a, b]$有

$$\left|\varphi(x)-\varphi(y)\right|=\left|\varphi'(\xi)(x-y)\right|\le L\left|(x-y)\right|,\ \xi\in(a,b).$$

故定理中的条件2°是成立的.

定理指出， 只要构造的迭代函数满足

$$|\varphi'(x)| \le L < 1$$

对于预先给定的误差ε

迭代法 $x_{k+1} = \varphi(x_k)$ 就收敛，即要求 $|x_k - x^*| < \varepsilon$

只要

$$\frac{L}{1-L}\left| x_k - x_{k-1} \right| < \varepsilon$$

因此，当 $\left| x_k - x_{k-1} \right| < \frac{1-L}{L}\varepsilon \approx \varepsilon$

迭代就可以终止， x_k 可以作为方程的近似解

例： 用迭代法求方程的近似解, 精确到小数点后6位

$$e^x + 10x - 2 = 0$$

解： 由于 $e^x > 0,$ 则 $2 - 10x > 0$ $\quad x < 0.2$

$$x < 0\text{时} \quad ,0 < e^x < 1, \quad 2 - 10x > 2$$

因此 $[0, 0.2]$ 为有根区间

本题迭代函数有两种构造形式

$$x = \varphi_1(x) = \frac{2 - e^x}{10} \qquad x = \varphi_2(x) = \ln(2 - 10x)$$

由于 $|\varphi_1'(x)| = \dfrac{e^x}{10} < \dfrac{e^{0.2}}{10} < 1$ $\quad |\varphi_2'(x)| = \dfrac{10}{2 - 10x} \geq 5$

因此采用迭代函数 $x = \varphi_1(x) = \dfrac{2 - e^x}{10}$

$$e^x + 10x - 2 = 0 \qquad x_{k+1} = \varphi(x_k) = (2 - e^{x_k})/10$$

取初值 $x_0 = 0$ $\quad x_1 = \dfrac{2 - e^{x_0}}{10} = 0.1$

x_1 = 0.1000000	d_1 = 0.1000000
x_2 = 0.0894829	d_2 = -0.0105171
x_3 = 0.0906391	d_3 = 0.1156e-002
x_4 = 0.0905126	d_4 = -0.1265e-003
x_5 = 0.0905265	d_5 = 0.1390e-004
x_6 = 0.0905250	d_6 = -0.1500e-005
x_7 = 0.0905251	d_7 = 0.1000e-006

由于 $|d_7|$ =0.1000e-006<1e-6

因此原方程的解为 $\quad x^* \approx x_7 =$ 0.090525

局部收敛性与收敛阶

上面给出了迭代序列$\{x_k\}$在区间$[a, b]$上的收敛性，通常称为**全局收敛性**. 有时不易检验定理的条件，实际应用时通常只在不动点x^*的邻近考察其收敛性，即**局部收敛性**.

定义1 设$\varphi(x)$有不动点x^*，如果存在x^*的某个邻域R: $|x-x^*|\leqslant \delta$，对任意$x_0\in R$，迭代公式(2.2)产生的序列$\{x_k\}\in R$，且收敛到x^*，则称迭代法(2.2)**局部收敛**.

定理3 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(2.2)局部收敛.

证明 由连续函数的性质，存在不动点x^*的某个邻域R: $|x-x^*|\leqslant \delta$，使对于任意$x\in R$成立

$$|\varphi'(x)| \le L < 1.$$

此外，对于任意$x\in R$，总有$\varphi(x)\in R$，这时因为

$$|\varphi(x)-x^*| = |\varphi(x)-\varphi(x^*)| = |\varphi'(\xi)||x-x^*|$$

$$\le L|x-x^*| \le |x-x^*|.$$

于是依据定理2可以断定迭代过程$x_{k+1}=\varphi(x_k)$对于任意初值$x_0\in R$均收敛. 证毕.

全局收敛与局部收敛

- 定理的条件保证了不动点迭代的全局收敛性。即迭代的收敛性与初始点的选取无关。

定理中的条件 $|\varphi'(x)| \le L < 1$ 可以适当放宽

$\varphi'(x)$ 在 x^* 的某个邻域内连续，且 $|\varphi'(x^*)| < 1$

由 $\varphi'(x)$ 的连续性及 $|\varphi'(x^*)| < 1$ 即可推出：

存在 x^* 的某个 δ 邻域 $N(x^*) = [x^* - \delta,\ x^* + \delta]$，使得对 $\forall\, x \in N(x^*)$ 都有 $|\varphi'(x)| \le L < 1$，则由 $\forall x_0 \in N(x^*)$ 开始的迭代都收敛。

- 这种在 x^* 的邻域内具有的收敛性称为局部收敛性。

例4 用不同迭代法求方程$x^2-3=0$的根 $x^* = \sqrt{3}$.

解 这里$f(x)= x^2-3$，可以改写为各种不同的等价形式$x=\varphi(x)$，其不动点为 $x^* = \sqrt{3}$ ，由此构造不同的迭代法.

$$(1)\ x_{k+1} = x_k^2 + x_k - 3, \varphi(x) = x^2 + x - 3, \varphi'(x) = 2x + 1,$$

$$\varphi'(x^*) = \varphi'(\sqrt{3}) = 2\sqrt{3} + 1 > 1.$$

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

$$(3)\ x_{k+1} = x_k - \frac{1}{4}(x_k^2 - 3), \varphi(x) = x - \frac{1}{4}(x^2 - 3),$$

$$\varphi'(x) = 1 - \frac{1}{2}x,\ \varphi'(x^*) = 1 - \frac{\sqrt{3}}{2} \approx 0.134 < 1.$$

(4) $x_{k+1}=\frac{1}{2}\left(x_k+\frac{3}{x_k}\right), \varphi(x)=\frac{1}{2}\left(x+\frac{3}{x}\right),$

$$\varphi'(x)=\frac{1}{2}\left(1-\frac{3}{x^2}\right), \varphi'(x^*)=\varphi'(\sqrt{3})=0.$$

取$x_0=2$, 对上式4种迭代法, 计算三步所得结果入下表.

k	x_k	迭代法(1)	迭代法(2)	迭代法(3)	迭代法(4)
0	x_0	2	2	2	2
1	x_1	3	1.5	1.75	1.75
2	x_2	9	2	1.73475	1.732143
3	x_3	87	1.5	1.732361	1.732051
⋮	⋮	⋮	⋮	⋮	⋮

注意 $\sqrt{3}=1.7320508\cdots$，从计算结果看到迭代法(1)及(2)均不收敛，且它们均不满足定理3中的局部收敛条件，迭代法(3)和(4)均满足局部收敛条件，且迭代法(4)比(3)收敛快，因在迭代法(4)中 $\varphi'(x^*)=0$. 为了衡量迭代法(2.2)收敛速度的快慢可给出以下定义.

定义2 设迭代过程 $x_{k+1}=\varphi(x_k)$ 收敛于方程 $x=\varphi(x)$ 的根 x^*，如果迭代误差 $e_k=x_k-x^*$ 当 $k\to\infty$ 时成立下列渐近关系式

$$\frac{e_{k+1}}{e_k^p}\to C \qquad (\text{常数} C\neq 0).$$

则称该迭代法是 p 阶收敛的. 特别地，$p=1$ 时称线性收敛，$p>1$ 时称超线性收敛，$p=2$ 时称平方收敛.

定理4 对于迭代过程$x_{k+1}=\varphi(x_k)$，如果$\varphi^{(p)}(x)$在所求根x^*的邻近连续，并且

$$\varphi'(x^*)=\varphi''(x^*)=\cdots=\varphi^{(p-1)}(x^*)=0,\varphi^{(p)}(x^*)\neq 0. \quad (2.8)$$

则该迭代过程在x^*的邻近是p阶收敛的.

证明 由于$\varphi'(x^*)=0$，根据定理3立即可以断定迭代过程$x_{k+1}=\varphi(x_k)$具有局部收敛性.

再将$\varphi(x_k)$在根x^*处做泰勒展开, 利用条件(2.4), 则有

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

注意到$\varphi(x_k)=x_{k+1}$，$\varphi(x^*)=x^*$，由上式得

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

因此对迭代误差，令$k\to\infty$时有

$$\frac{e_{k+1}}{e_k^p}\to\frac{\varphi^{(p)}(x^*)}{p!}.$$

这表明迭代过程$x_{k+1}=\varphi(x_k)$确实为p阶收敛.　　证毕.

上述定理告诉我们，迭代过程的收敛速度依赖于迭代函数$\varphi(x)$的选取. 如果$x\in[a, b]$但$\varphi'(x)\neq 0$时，则该迭代过程只可能是线性收敛.

例子 证明迭代公式 $x_{k+1}=x_k(x_k^2+3a)/(3x_k^2+a)$是求 $\sqrt{a}\,(a>0)$ 的三阶方法. 假设 x_0 充分靠近 x^*, 求

$$\lim_{k\to\infty}\frac{\sqrt{a}-x_{k+1}}{(\sqrt{a}-x_k)^3}$$

证明 首先由泰勒展式可得

$$\lim_{k\to\infty}\frac{\sqrt{a}-x_{k+1}}{(\sqrt{a}-x_k)^3}=\lim_{k\to\infty}\frac{-e_{k+1}}{(-e_k)^3}=\frac{1}{3!}\varphi'''(\sqrt{a})=\frac{1}{4a}$$

而$1/4a\neq 0$，故此迭代公式是三阶方法.

三、迭代收敛的加速方法

埃特金加速收敛方法

对于收敛的迭代过程，只要迭代足够多次，就可以使结果达到任意的精度，但是有时迭代过程收敛较慢，从而使计算量变得很大，因此迭代过程的加速是个重要的课题.

设x_0是根x^*的某个近似值, 用迭代公式校正一次得

$$x_1=\varphi(x_0)$$

而由微分中值定理，有

$$x_1-x^*=\varphi(x_0)-\varphi(x^*)=\varphi'(\xi)(x_0-x^*),$$ ξ在x^*与x_0之间.

假设$\varphi'(x)$改变不大，近似地取某个近似值L，则有

$$x_1 - x^* \approx L(x_0 - x^*). \qquad (3.1)$$

若将校正值$x_1=\varphi(x_0)$再校正一次，又得

$$x_2=\varphi(x_1)$$

由于 $x_2-x^*\approx L(x_1-x^*)$.

将它与(3.1)式联立，消去未知的L，有

$$\frac{x_1 - x^*}{x_2 - x^*} \approx \frac{x_0 - x^*}{x_1 - x^*}$$

由此推知

$$x^* \approx \frac{x_0 x_2 - x_1^2}{x_2 - 2x_1 + x_0} = x_0 - \frac{(x_1 - x_0)^2}{x_2 - 2x_1 + x_0}.$$

在计算了x_1及x_2之后，可用上式右端作为x^*的新近似，记作$\bar{x}_1$，一般情形是由x_k计算x_{k+1}，x_{k+2}，记

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

$$= x_k - \frac{(\Delta x_k)^2}{\Delta^2 x_k} \qquad (k = 0,1,\cdots). \qquad (3.2)$$

(3.2)式称为埃特金(Aitken) △²加速方法.

可以证明

$$\lim_{k\to\infty} \frac{\bar{x}_{k+1} - x^*}{x_k - x^*} = 0.$$

它表明序列$\{\bar{x}_k\}$的收敛速度比$\{x_k\}$的收敛速度快.

这个加速迭代法也可写成下面格式

$$\begin{cases} x_{k+1}^{(1)} = \varphi(x_k) \\ x_{k+1}^{(2)} = \varphi(x_{k+1}^{(1)}) \\ x_{k+1} = x_{k+1}^{(2)} - \dfrac{(x_{k+1}^{(2)} - x_{k+1}^{(1)})^2}{x_{k+1}^{(2)} - 2x_{k+1}^{(1)} + x_k} \end{cases}$$

也称为埃特金 (**Aitken**) 外推法. 可以证明:

若 $x_{k+1} = \varphi(x_k)$为线性收敛,则埃特金法为平方收敛;

若 $x_{k+1} = \varphi(x_k)$ 为 p ($p > 1$)阶收敛，$\varphi(x)$ 的 p 阶导数连续，则埃特金法为 $2p-1$ 阶收敛.

例题 求方程 $x = e^{-x}$ 在 $x=0.5$ 附近的根.

解 取 $x_0=0.5$, 迭代格式

$$x_{k+1} = e^{-x_k}$$

得 $x_{25}=x_{26}=0.5671433$

若对此格式用埃特金法, 则

$$x_{k+1}^{(1)} = e^{-x_k} \qquad x_{k+2}^{(2)} = e^{-x_{k+1}^{(1)}}$$

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

$$x_{k+1}^{(1)} = e^{-x_k} \qquad x_{k+2}^{(2)} = e^{-x_{k+1}^{(1)}}$$

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

仍取 $x_0=0.5$, 得

$$x_1^{(1)} = 0.6065307 \quad x_1^{(2)} = 0.5452392 \quad x_1 = 0.5676279$$

$$x_2^{(1)} = 0.5668708 \quad x_2^{(2)} = 0.5672979 \quad x_2 = 0.5671433$$

$$x_3^{(1)} = 0.5671433 \quad x_3^{(2)} = 0.5671433 \quad x_3 = 0.5671433$$

由此可见, 埃特金法加速收敛效果是相当显著的.

斯蒂芬森(Steffensen)迭代法

埃特金方法不管原序列$\{x_k\}$是怎样产生的，对$\{x_k\}$进行加速计算，得到序列$\{\bar{x}_k\}$. 如果把埃特金加速技巧与不动点迭代结合，则可得到如下的迭代法：

$$y_k = \varphi(x_k),\quad z_k = \varphi(y_k),$$

$$x_{k+1} = x_k - \frac{(y_k - x_k)^2}{z_k - 2y_k + x_k} \qquad (k = 0,1,\cdots). \qquad (3.3)$$

称为斯蒂芬森(Steffensen)迭代法. 它可以这样理解，我们要求$x=\varphi(x)$的根x^*，令误差$\varepsilon(x)=\varphi(x)-x$，有等式$\varepsilon(x^*)=\varphi(x^*)-x^*=0$，已知$x^*$的近似值$x_k$及$y_k$，其误差分别为

$$\varepsilon(x_k)=\varphi(x_k)-x_k=y_k-x_k,$$
$$\varepsilon(y_k)=\varphi(y_k)-y_k=z_k-y_k.$$

把误差$\varepsilon(x)$“外推到零”，即过$(x_k,\varepsilon(x_k))$及$(y_k,\varepsilon(y_k))$两点做线性插值函数，它与x轴交点就是(3.3)中的x_{k+1}，即方程

$$\varepsilon(x_k)+\frac{\varepsilon(y_k)-\varepsilon(x_k)}{y_k-x_k}(x-x_k)=0.$$

的解

$$x=\varepsilon(x_k)-\frac{\varepsilon(x_k)}{\varepsilon(y_k)-\varepsilon(x_k)}(y_k-x_k)$$
$$=x_k-\frac{(y_k-x_k)^2}{z_k-2y_k+x_k}=x_{k+1}.$$

实际上(3.3)是将不定点迭代法(2.2)计算两步合并成一步得到的，可将它写成另一种不动点迭代

$$x_{k+1}=\psi(x_k)\quad (k=0,1,\cdots), \tag{3.4}$$

其中

$$\psi(x)=x-\frac{[\varphi(x)-x]^2}{\varphi(\varphi(x))-2\varphi(x)+x}. \tag{3.5}$$

对不动点迭代(3.5)有以下局部收敛性定理.

定理5 若x^*为(3.5)定义的迭代函数$\Psi(x)$的不动点，则x^*为$\varphi(x)$的不定点. 反之，若x^*为$\varphi(x)$的不动点，设$\varphi''(x)$存在，$\varphi'(x)\neq 1$，则x^*是$\Psi(x)$的不动点，且斯蒂芬森迭代法(3.3)是2阶收敛的.

四、牛 顿 法

牛顿法及其收敛性

对于方程$f(x)=0$，如果$f(x)$是线性函数，则它的求根是容易的．牛顿法实质上是一种线性化方法，其基本思想是将非线性方程$f(x)=0$逐步归结为某种线性方程来求解.

设已知方程$f(x)=0$有近似根x_0，且在 x_0附近$f(x)$可用一阶泰勒多项式近似，表示为

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

当$f'(x_0)\neq 0$时，方程$f(x)=0$可用线性方程(切线) 近似代替，即

$$f(x_0)+f'(x_0)(x-x_0)=0. \qquad (4.1)$$

解此线性方程得

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

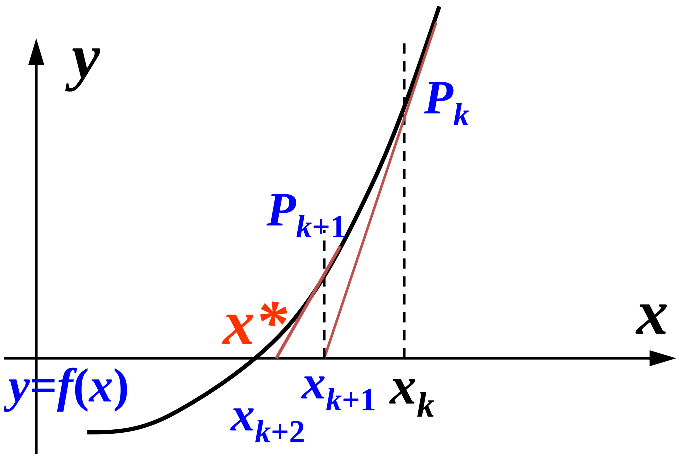
得迭代公式

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad (k=0,1,\cdots), \qquad (4.2)$$

此式称为牛顿(Newton)迭代公式.

牛顿法有显然的几何意义，方程$f(x)=0$的根x^*可解释为曲线$y=f(x)$与x轴交点的横坐标. 设x_k是根x^*的某个近似值，过曲线$y=f(x)$上横坐标为x_k的点P_k引切线，并将该切线与x轴交点的横坐标x_{k+1}作为x^*的新的近似值. 注意到切线方程为

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



这样求得的值x_{k+1}必满足(4.1)，从而就是牛顿公式(4.2)的计算结果. 由于这种几何背景，所以牛顿迭代法也称切线法.

牛顿法的算法构造

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

[牛顿迭代法]

1: 初始化. x_0, M,δ,ε,置i:=0

2: 如果$|f(x_i)| \le \delta$，则停止.

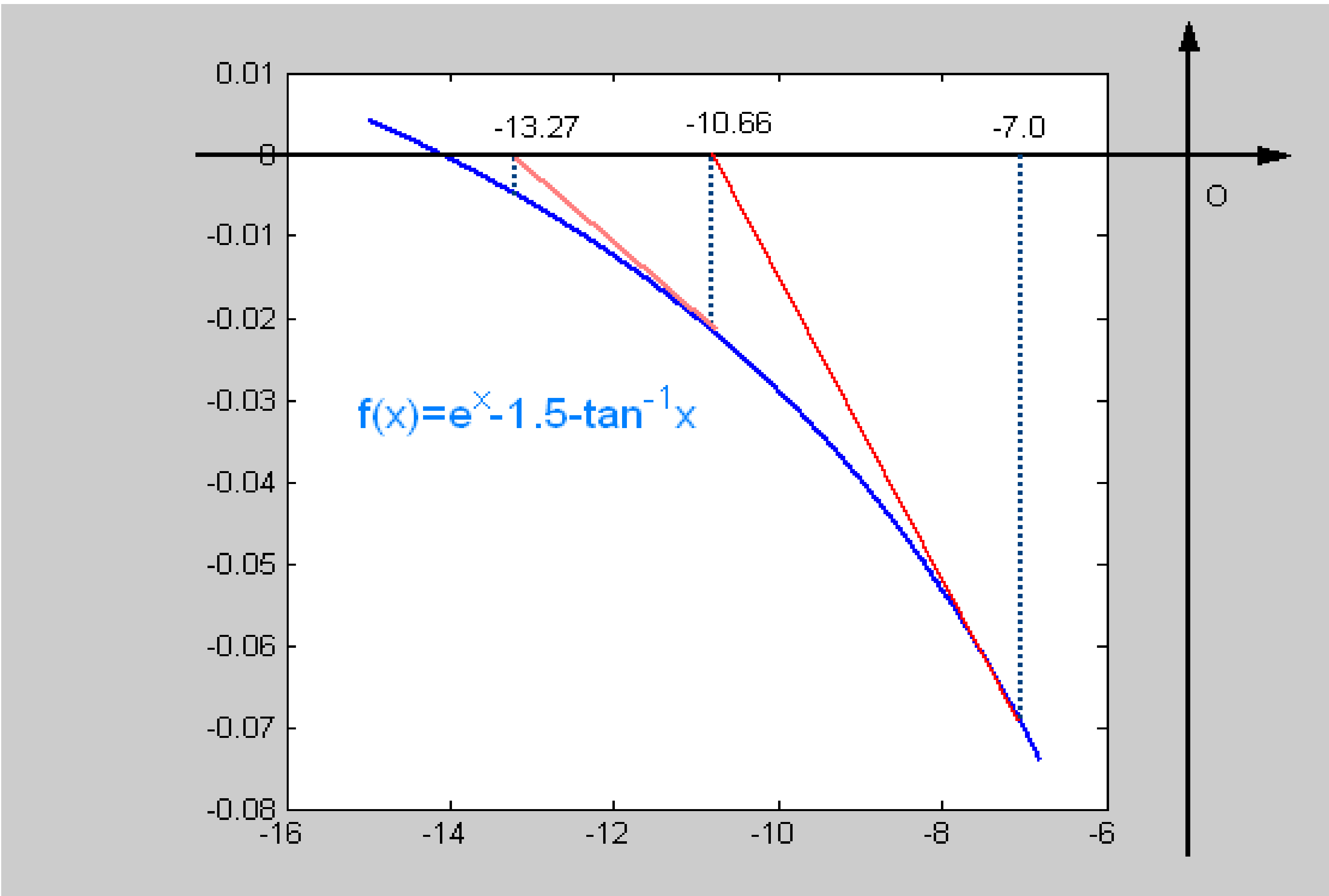
3: 计算 $x_{i+1} := x_i - f(x_i) / f'(x_i)$

4: 如果$|x_{i+1}-x_i| < \varepsilon$ or $|f(x_i)| \le \delta$，则停止.

5: $i:=i+1$, 转至3.

例： 利用牛顿迭代法求解

$f(x)=e^x-1.5-tan^{-1}x$ 的零点。初始点$x_0=-7.0$



例1： 利用牛顿迭代法求

$f(x)=e^x-1.5-tan^{-1}x$ 的零点。初始点x_0=-7.0

解： $f(x_0)=-0.702\times 10^{-1}$，$f'(x)=e^x-(1+x^2)^{-1}$

计算迭代格式： $$x_{k+1}=x_k-\frac{f(x_k)}{f'(x_k)}$$

计算结果如下表：(取$|f(x)|<=10^{-10}$)

k	x	$f(x)$
0	-7.0000	-0.0701888
1	-10.6771	-0.0225666
2	-13.2792	-0.00436602
3	-14.0537	-0.00023902
4	-14.1011	-7.99585e-007
5	-14.1013	-9.00833e-012

Newton' Method收敛的充分条件

设 $f \in C^2[a, b]$，若

有根

根唯一

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

(2) 在整个$[a, b]$上 f''不变号且 $f'(x) \neq 0$;

(3) 选取 $x_0 \in [a, b]$ 使得 $f(x_0)\,f''(x_0) > 0$;

则Newton's Method产生的序列$\{x_k\}$ 收敛到$f(x)$ 在 $[a, b]$ 的唯一根。

产生的序列单调有界，保证收敛。

牛顿迭代法的收敛性

牛顿迭代法的迭代函数为

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

设x^*是$f(x)$的一个单根，即$f(x^*)=0$，$f'(x^*)\neq 0$, 有

$$\varphi'(x^*)=\frac{f(x^*)f''(x^*)}{[f'(x^*)]^2}=0,$$

$$\varphi''(x^*)=\frac{f''(x^*)}{f'(x^*)}\neq 0.$$

由定理4的(2.9)式可得　　(4.3)式

$$\lim_{k\leftarrow\infty}\frac{x_{k+1}-x^*}{(x_k-x^*)^2}=\lim_{k\leftarrow\infty}\frac{\frac{1}{2!}\varphi''(\xi)(x_k-x^*)^2}{(x_k-x^*)^2}=\frac{1}{2!}\varphi''(x^*)=\frac{f''(x^*)}{2f'(x^*)}\neq 0.$$

即有下面的局部收敛性定理.

定理(局部收敛性) 设$f\in C^2[a, b]$, 若x^*为$f(x)$在$[a, b]$上的根，且$f'(x^*)\neq 0$，则存在x^*的邻域U，使得任取初值$x_0\in U$，牛顿法产生的序列$\{x_k\}$收敛到x^*，且满足

$$\lim_{k\leftarrow\infty}\frac{x_{k+1}-x^*}{(x_k-x^*)^2}=\frac{f''(x^*)}{2f'(x^*)}.$$

由此得到，当x^*为单根时，牛顿迭代法在根x^*的邻近是二阶(平方)收敛的.

例7 用牛顿迭代法求方程$x=e^{-x}$在$x=0.5$附近的根.

解 将原方程化为$x-e^{-x}=0$，则

$$f(x)=x-e^{-x},\quad f'(x)=1+e^{-x},$$

牛顿迭代公式为

$$x_{k+1}=x_k-\frac{x_k-e^{-x_k}}{1+e^{-x_k}}$$

取 $x_0=0.5$，迭代得

$x_1=0.566311$, $x_2=0.5671431$, $x_3=0.5671433$.

简化牛顿法与牛顿下山法

牛顿法的优点是收敛快，缺点①每步迭代要计算 $f(x_k)$ 及 $f'(x_k)$，计算量较大，且有时 $f'(x_k)$ 计算较困难；②初始近似值 x_0 只在根 x^* 附近才能保证收敛，如 x_0 给的不合适可能不收敛. 为克服这两个缺点，通常可用下述方法.

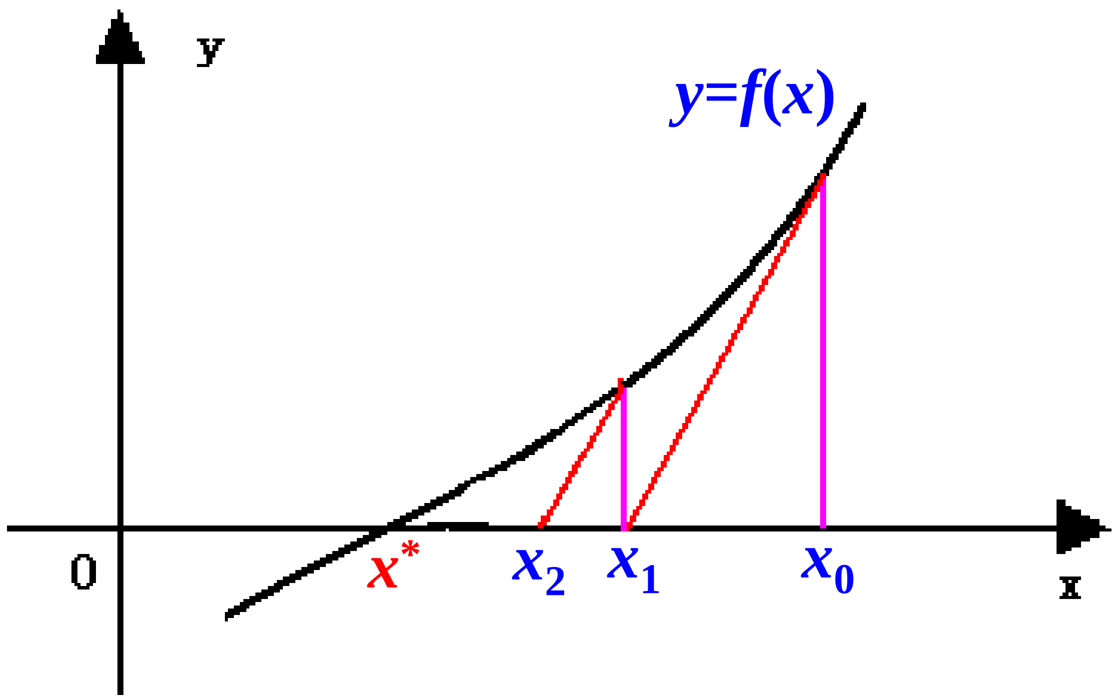
(1) 简化牛顿法，也称平行弦法，其迭代公式为

$$x_{k+1} = x_k - Cf(x_k) \quad C \neq 0, k = 0,1,\cdots. \qquad (4.7)$$

迭代函数为 $\varphi(x)=x-Cf(x)$.

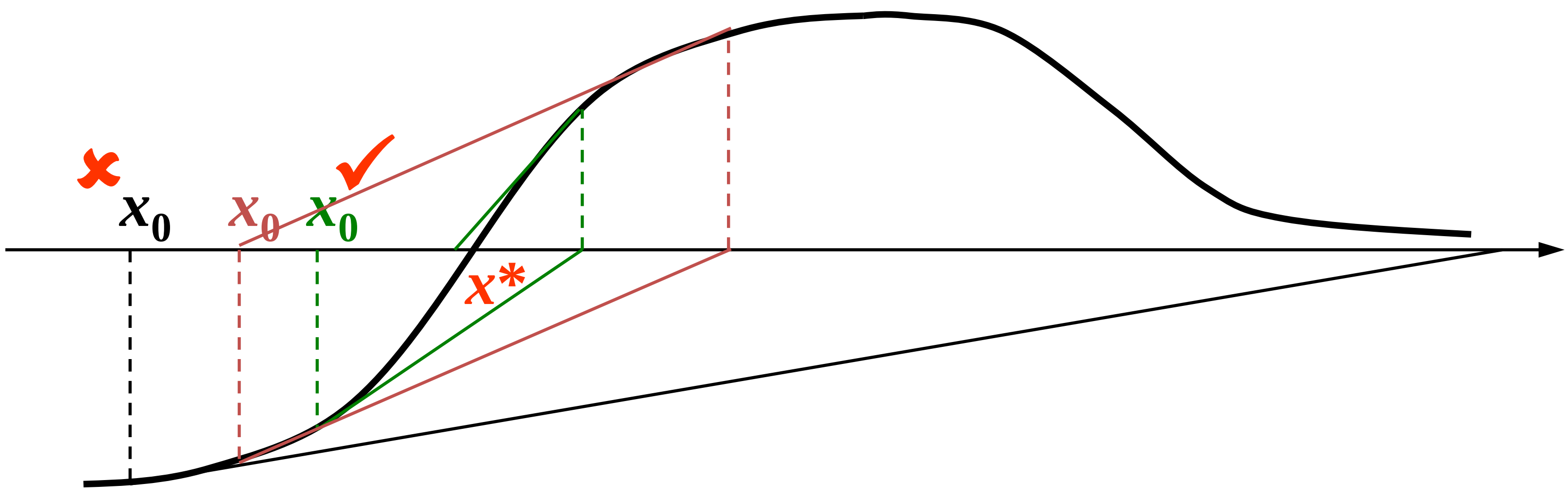
若$|\varphi'(x_k)|=|1-Cf'(x)|<1$，即取$0<Cf'(x)<2$. 在根$x^*$附近成立，则迭代法(4.7)局部收敛.

在(4.7)中取$C=1/f'(x_0)$，则称为简化牛顿法，这类方法计算量省，但只有线性收敛，其几何意义是用平行弦与x轴交点作为x^*的近似，见下图.



(2) 牛顿下山法, 牛顿法收敛性依赖初值x_0的选取, 如果x_0偏离所求根x^*较远, 则牛顿法可能发散.

注：Newton’s Method收敛性依赖于x_0的选取.



例如，用牛顿法求解方程 $x^3-x-1=0$. (4.8)

此方程在$x=1.5$附近的一个根x^*.

设取迭代初值$x_0=1.5$，用牛顿迭代法公式

$$x_{k+1}=x_k-\frac{x_k^3-x_k-1}{3x_k^2-1}. \tag{4.9}$$

计算得 $x_1=1.34783$, $x_2=1.32520$, $x_3=1.32472$.

迭代3次得到的结果x_3有6位有效数字.

但是，如取$x_0=0.6$，用(4.9)式迭代1次得

计算得 $x_1=17.9$.

这个结果反而比$x_0=0.6$更偏离了所求的根$x^*=1.32472$.

为了防止迭代发散，我们对迭代过程再附加一项要求，即具有单调性.

$$\left| f(x_{k+1}) \right| < \left| f(x_k) \right|. \qquad (4.10)$$

满足这项要求的算法称为下山法.

我们将牛顿法与下山法结合起来使用，即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. 为此，我们将牛顿法的结果

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

与前一项的近似值x_k适当加权平均作为新的改进值

$$x_{k+1} = \lambda \bar{x}_{k+1} + (1-\lambda) x_k, \qquad (4.11)$$

其中$\lambda(0<\lambda\leqslant 1)$称为下山因子，(4.11)即为

$$x_{k+1} = x_k - \lambda \frac{f(x_k)}{f'(x_k)} \quad (k = 0,1,\cdots). \qquad (4.12)$$

称为牛顿下山法.

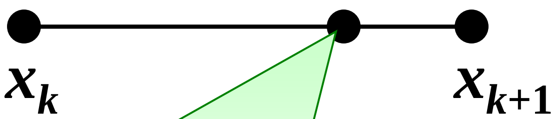
选择下山因子时从$\lambda=1$开始，逐次将λ减半进行试算，直到能使下降条件(4.10)成立为止.

若用此法解方程(4.8)，当$x_0=0.6$时由(4.9)式求得$x_1=17.9$，它不满足条件(4.10)，通过λ逐次减半进行试算,当$\lambda=1/32$时可求得$x_1=1.140625$. 有$f(x_0)=-1.384$, $f(x_1)=-0.656643$, 显然$|f(x_1)|<|f(x_0)|$. 计算$x_2,x_3,\cdots$时，均能使条件(4.10)成立. 得到$x_4=1.32472$即为x^*的近似.

➢ 下山法 /* Descent Method */

——Newton's Method 局部微调：

原理：若由 x_k 得到的 x_{k+1} 不能使 $|f|$ 减小，则在 x_k 和 x_{k+1} 之间找一个更好的点 $\overline{x_{k+1}}$，使得 $|f(\overline{x_{k+1}})| < |f(x_k)|$。



$$\lambda x_{k+1} + (1-\lambda)x_k \,,\ \lambda \in [0, 1]$$

$$\overline{x_{k+1}} = \lambda\left[x_k - \frac{f(x_k)}{f'(x_k)}\right] + (1-\lambda)x_k$$

$$= x_k - \lambda \frac{f(x_k)}{f'(x_k)}$$

注：$\lambda = 1$ 时就是Newton's Method 公式。

当 $\lambda = 1$ 代入效果不好时，将 λ 减半计算。

一般情况下，只要能使条件选择(4.10)成立，则可得到极限

$$\lim_{k\to\infty} f(x_k) = 0,$$

从而使数列$\{x_k\}$收敛.

重根情形

当x^*为$f(x)$的$m(m>0)$重根时，则$f(x)$可表为

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

其中$g(x^*)\neq 0$，此时用牛顿迭代法(4.2)求x^*仍然收敛，只是收敛速度将大大减慢．事实上，因为迭代公式

$$x_{k+1}=x_k-\frac{f(x_k)}{f'(x_k)}=x_k-\frac{(x_k-x^*)g(x_k)}{mg(x_k)+(x_k-x^*)g'(x_k)}$$

令$e_k=x_k-x^*$，则

$$e_{k+1}=x_{k+1}-x^*=e_k-\frac{e_k g(x_k)}{mg(x_k)+e_k g'(x_k)}$$

从而有

$$\lim_{k\to\infty}\frac{e_{k+1}}{e_k}=\lim_{k\to\infty}\left[1-\frac{g(x_k)}{mg(x_k)+e_k g'(x_k)}\right]=1-\frac{1}{m}\neq 0.$$

可见用牛顿法求方程的重根时仅为线性收敛.

两种提高求重根的收敛速度的方法:

1） 取如下迭代函数

$$\varphi(x)=x-m\frac{f(x)}{f'(x)},\quad 则\ \varphi'(x^*)=0.$$

得到迭代公式

$$x_{k+1}=x_k-m\frac{f(x_k)}{f'(x_k)}\quad (k=0,1,\cdots).\quad (4.13)$$

求m重根具有2阶收敛. 但要知道x^*的重数m.

下面介绍一个求重数m的方法，令

$$\lambda_k = \frac{x_k - x_{k-1}}{x_{k-1} - x_{k-2}}$$

则

$$\lambda_k = \frac{e_k - e_{k-1}}{e_{k-1} - e_{k-2}} = \frac{e_k}{e_{k-1}} \cdot \frac{1 - \frac{e_{k-1}}{e_k}}{1 - \frac{e_{k-2}}{e_{k-1}}}$$

由式

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k} = 1 - \frac{1}{m}$$

得

$$\lim_{k\to\infty} \lambda_k = 1 - \frac{1}{m} = \frac{m-1}{m}.$$

因此得估计m的式子为 $m \approx \dfrac{1}{1-\lambda_k}.$

2）将求重根问题化为求单根问题.

对$f(x)=(x-x^*)^m g(x), g(x^*)\neq 0$，令函数

$$\mu(x)=\frac{f(x)}{f'(x)}=\frac{(x-x^*)g(x)}{mg(x)+(x-x^*)g'(x)}.$$

则为求$\mu(x)=0$的单根x^*的问题，对它用牛顿法是二阶(平方)收敛的. 其迭代函数为

$$\varphi(x)=x-\frac{\mu(x)}{\mu'(x)}=x-\frac{f(x)f'(x)}{[f'(x)]^2-f(x)f''(x)}.$$

从而构造出迭代方法为

$$x_{k+1}=x_k-\frac{f(x_k)f'(x_k)}{[f'(x_k)]^2-f(x_k)f''(x_k)} \quad (k=0,1,\cdots). \quad (4.14)$$

例8 用牛顿迭代法求函数

$f(x)=(x-1)[\sin(x-1)+3x]-x^3+1=0$ 在0.95附近之根.

解 取$x_0 = 0.95$

用牛顿迭代法求得的x_k见右表. 可见x_k收敛很慢.

k	x_k	λ_k	m
0	0.95		
1	0.9744279		
2	0.9870583		
3	0.9934878	0.5090	2.0369
4	0.9967328	0.5047	2.0190
5	0.9983576	0.5007	2.0028
6	0.9991901	0.5125	2.0511

由重根数$m=2$，用(4.13)式加速法，作

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

求得 $x_0=0.95$, $x_1=0.9988559$, $x_2=x_3=1$.

收敛速度大大加快于直接用牛顿迭代公式.

五、弦截法与抛物线法

用牛顿法求方程$f(x)=0$的根，每步除计算$f(x_k)$外还要算$f'(x_k)$，当函数$f(x)$比较复杂时，计算$f'(x)$往往比较困难，为此可以利用已求函数值$f(x_k),f(x_{k-1}),\cdots$来回避导数值$f'(x_k)$的计算. 这类方法是建立在插值原理基础上的，下面介绍两种常用方法.

弦截(割线)法

设x_k,x_{k-1}是$f(x)=0$的近似根，我们利用$f(x_k),f(x_{k-1})$构造一次插值多项式$p_1(x)$，并用$p_1(x)=0$的根作为方程$f(x)=0$的新的近似根x_{k+1}，由于

$$p_1(x)=f(x_k)+\frac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}(x-x_k). \qquad (5.1)$$

因此有

$$x_{k+1}=x_k-\frac{f(x_k)}{f(x_k)-f(x_{k-1})}(x_k-x_{k-1}). \qquad (5.2)$$

这样导出的迭代公式(5.2)可以看做牛顿公式

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

中的导数 $f'(x_k)$用差商 $\dfrac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}$ 取代的结果.

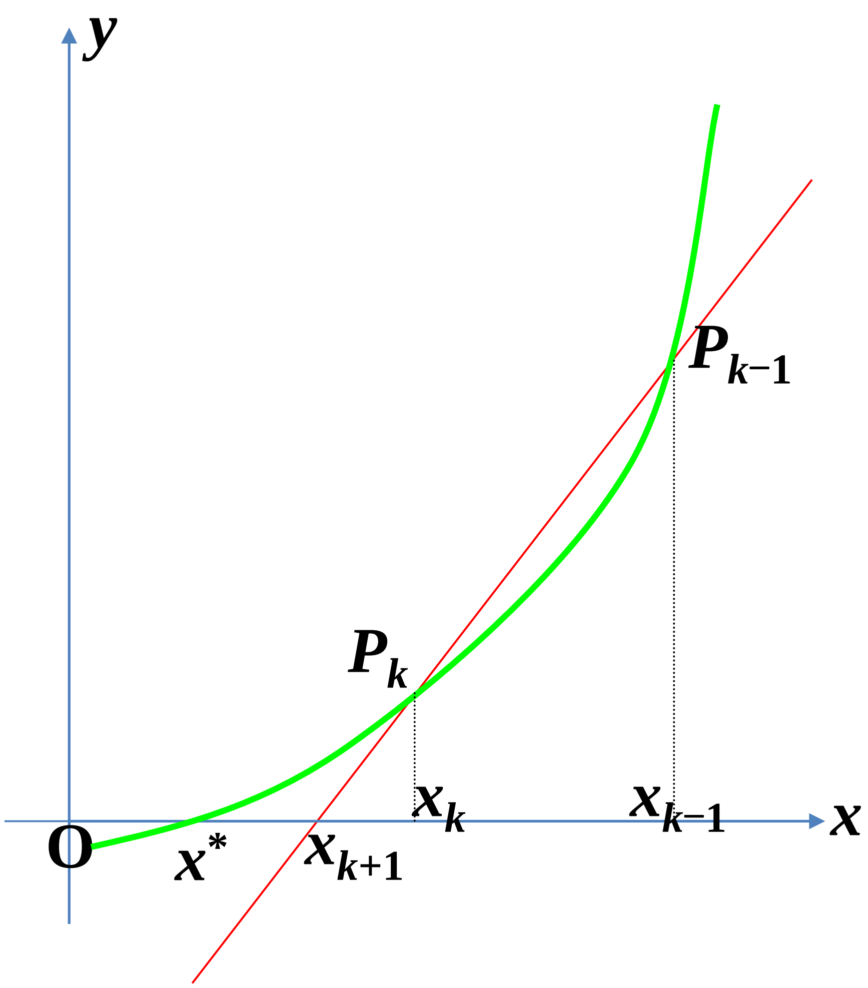
(5.2)式有明显的几何意义:

设曲线$y=f(x)$上横坐标为x_{k-1}和x_k的点分别为P_0和P_k, 则差商 $\dfrac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}$ 表示弦 $\overline{P_{k-1}P_k}$ 的斜率,

弦$\overline{P_{k-1}P_k}$的方程为

$$y=f(x_k)+\frac{f(x_k)-f(x_0)}{x_k-x_0}(x-x_k)$$

因此，按(5.2)式求得 x_{k+1}实际上是两点弦线 $\overline{P_{k-1}P_k}$ 与x轴交点的横坐标(令$y=0$解出x即可).这种算法因此而形象地称为弦截(割线)法.



弦截法与切线法(牛顿法)都是线性化分法，但两者有本质的区别. 切线法在计算x_{k+1}时只用到前一步的值x_k，而弦截法要用到前面两步的结果x_{k-1}, x_k，因此使用这种方法必须先给出两个开始值x_0, x_1.

定理6 假设$f(x)$在根x^*的邻域内$\Delta: |x-x^*| \leqslant \delta$具有二阶连续导数，且对任意$x \in \Delta$有$f'(x) \neq 0$，所取的初值$x_0, x_1 \in \Delta$，那么当邻域$\Delta$充分小时，弦截法(5.2)将按阶

$$p = \frac{1+\sqrt{5}}{2} \approx 1.618.$$

收敛到x^*. 这里p是方程$\lambda^2-\lambda-1=0$的正根.

因为(5.2)式用到前两点x_{k-1}和x_k的值，故此方法又称为双点割线法.

如果把(5.2)式中的x_{k-1}改为x_0，即迭代公式为

$$x_{k+1} = x_k - \frac{f(x_0)}{f(x_k) - f(x_0)}(x_k - x_0).$$

每步只用一个新点x_k的值，此方法称为单点割线法.

例题 用牛顿迭代法和割线法求方程

$$f(x)=x^4+2x^2-x-3=0,$$

在区间$(1, 1.5)$内之根(误差为10^{-9}).

解 取$x_0=1.5$，用牛顿法, 可得$x_6=1.12412303030$;

取$x_0=1.5$, $x_1=1$，用双点割线法，迭代6次得到同样的结果，而采用单点割线法，则迭代18次得

$x_{18}=1.124123029$.

例题 用快速弦截法求方程$xe^x-1=0$的根. 设方程的两个初始近似根为$x_0=0.5$, $x_1=0.6$.

计算结果表

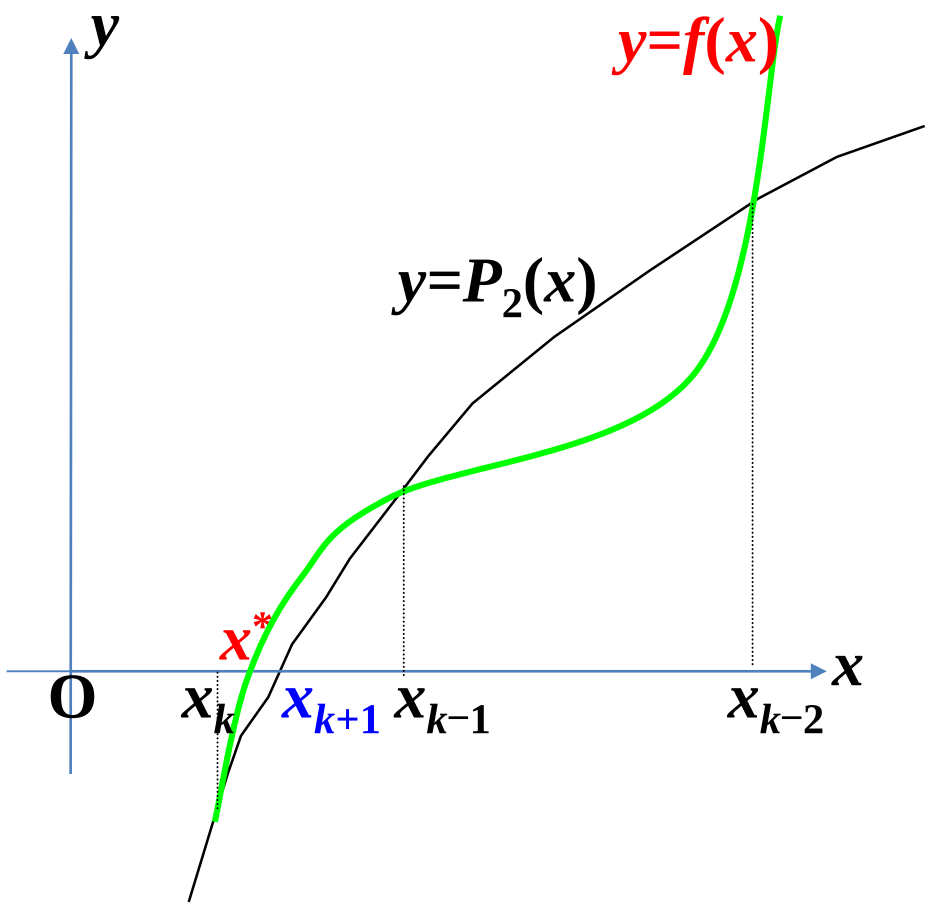
k	x_k	x_k-x_{k-1}
0	0.5	
1	0.6	0.1
2	0.56532	−0.03468
3	0.56709	0.00177
4	0.56714	0.00005

与例7中牛顿法的计算结果相比较，可以看出快速弦截法的收敛速度也是相当快的，迭代到第4步就得到精度 $\varepsilon=10^{-4}$ 的结果.

抛物线法

设已知方程$f(x)=0$的三个近似根x_k,x_{k-1},x_{k-2}，我们以这三点为节点构造二次插值多项式$p_2(x)$，并适当选取$p_2(x)$的一个零点x_{k+1}作为新的近似根，这样确定的迭代过程称为**抛物线法**，亦称为**密勒(Müller)法**. 在几何图形上, 这种方法的基本思想是用抛物线$y=p_2(x)$与x轴的交点x_{k+1}作为所求根x^*的近似位置.

抛物线法的几何意义见下面图形.



现在推导抛物线法的计算公式. 插值多项式

$$p_2(x) = f(x_k) + f[x_k, x_{k-1}](x - x_k) + f[x_k, x_{k-1}, x_{k-2}](x - x_k)(x - x_{k-1}).$$

有两个零点

$$x_{k+1} = x_k - \frac{2f(x_k)}{\omega \pm \sqrt{\omega^2 - 4f(x_k)f[x_k, x_{k-1}, x_{k-2}]}}. \qquad (5.3)$$

式中 $\omega = f[x_k, x_{k-1}] + f[x_k, x_{k-1}, x_{k-2}](x_k - x_{k-1}).$

因了在(5.3)式定出一个值x_{k+1}，我们需要讨论根式前正负号的取舍问题.

在x_k, x_{k-1}, x_{k-2}三个近似值中，自然假定x_k更接近所求的根x^*，这时，为了保证精度，我们选(5.3)式中接近x_k的一个值作为新的近似根x_{k+1}. 为此，只要取根式前的符号与ω的符号相同.

例11 用抛物线法求解方程$f(x)=xe^x-1=0$.

解 取$x_0=0.5$, $x_1=0.6$, $x_2=0.56532$开始，计算得

$f(x_0)=-0.175639$, $f(x_1)=0.093271$, $f(x_2)=-0.005031$.

$f[x_1,x_0]=2.68910$, $f[x_2,x_1]=2.83373$, $f[x_2,x_1,x_0]=2.21418$.

故

$$\omega = f[x_2,x_1]+f[x_2,x_1,x_0](x_2-x_1)=2.75694.$$

代入(5.3)式求得

$$x_3 = x_2 - \frac{2f(x_2)}{\omega+\sqrt{\omega^2-4f(x_2)f[x_2,x_1,x_0]}} = 0.56714.$$

以上计算表明，抛物线法比弦截法收敛更快.

事实上，在一定条件下可以证明，对于抛物线法，迭代误差有下列渐近关系式

$$\frac{|e_{k+1}|}{|e_k|^{1.840}} \to \left|\frac{f'''(x^*)}{6f'(x^*)}\right|^{0.42}.$$

由此式可见抛物线法也是超线性收敛的，其收敛的阶是$p=1.840$(是方程$\lambda^3-\lambda^2-\lambda-1=0$的根)，收敛速度比弦截法更接近于牛顿法.

从(5.3)式看到，即使 x_k, x_{k-1}, x_{k-2} 均为实数，x_{k+1} 也可以是复数，所以抛物线法适用于求多项式的实根和复根.

六、解非线性方程组的牛顿迭代法

考察方程组

$$\begin{cases} f_1(x_1,\cdots,x_n)=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_n(x_1,\cdots,x_n)=0. \end{cases} \tag{6.1}$$

其中$f_1,\cdots,f_n$均为$(x_1,\cdots,x_n)$的多元函数. 若用向量记号记$x=(x_1,\cdots,x_n)^T\in R^n$, $F=(f_1,\cdots,f_n)^T$, (6.1)就可写成

$$F(x)=0. \tag{6.2}$$

当$n\geq 2$，且$f_1,\cdots,f_n$中至少有一个是自变量$x_1,\cdots,x_n$的非线性函数，则称方程组(6.1)为非线性方程组．非线性方程组求根问题是前面介绍的方程(即$n=2$)求根的直接推广，实际上只要把前面介绍的单变量函数$f(x)$看成向量函数$F(x)$，则可得向量方程(6.2)的一个近似根$x^{(k)}=(x_1^{(k)},\cdots,x_n^{(k)})^{\mathrm{T}}$，将函数$F(x)$的分量$f_i(x)(i=1,\cdots,n)$在$x^{(k)}$用多元函数泰勒展开，并取其线性部分，则可表示为.

$$F(x)\approx F(x^{(k)})+F'(x^{(k)})(x-x^{(k)}).$$

令上式右端为零，得到线性方程组

$$F'(x^{(k)})(x-x^{(k)})=-F(x^{(k)}). \tag{6.3}$$

其中

$$F'(x)=\begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \cdots & \frac{\partial f_2(x)}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_n} & \cdots & \frac{\partial f_n(x)}{\partial x_n} \end{pmatrix}. \tag{6.4}$$

称为$F(x)$的雅可比(Jacobi)矩阵. 求解线性方程组(6.3), 并记解为$x^{(k+1)}$, 则得

$$x^{(k+1)}=x^{(k)}-F'(x^{(k)})^{-1}F(x^{(k)})\quad (k=0,1,\cdots).\quad (6.5)$$

这就是解非线性方程组(6.2)的牛顿迭代法.

例12 求解方程组

$$\begin{cases} f_1(x_1,x_2)=x_1+2x_2-3=0, \\ f_2(x_1,x_2)=2x_1^2+x_2^2-5=0. \end{cases}$$

给定初值$x^{(0)}=(1.5, 1.0)^T$, 用牛顿法求解.

解 先求Jacobi矩阵

$$F'(x)=\begin{pmatrix} 1 & 2 \\ 4x_1 & 2x_2 \end{pmatrix}, F'(x)^{-1}=\frac{1}{2x_2-8x_1}\begin{pmatrix} 2x_2 & -2 \\ -4x_1 & 1 \end{pmatrix}.$$

用牛顿法(6.5)得

$$x^{(k+1)}=x^{(k)}-\frac{1}{2x_2^{(k)}-8x_1^{(k)}}\begin{pmatrix}2x_2^{(k)} & -2\\ -4x_1^{(k)} & 1\end{pmatrix}\begin{pmatrix}x_1^{(k)}+2x_2^{(k)}-3\\ 2(x_1^{(k)})^2+(x_2^{(k)})^2-5\end{pmatrix}.$$

即

$$\begin{cases}x_1^{(k+1)}=x_1^{(k)}-\dfrac{(x_2^{(k)})^2-2(x_1^{(k)})^2-x_1^{(k)}x_2^{(k)}-3x_2^{(k)}+5}{x_2^{(k)}-4x_1^{(k)}},\\ x_2^{(k+1)}=x_2^{(k)}-\dfrac{(x_2^{(k)})^2-2(x_1^{(k)})^2-8x_1^{(k)}x_2^{(k)}+12x_2^{(k)}-5}{2(x_2^{(k)}-4x_1^{(k)})},\end{cases}$$

$$(k=0,1,\cdots).$$

由$x^{(0)}=(1.5, 1.0)^T$逐次迭代得到

$$x^{(1)}=(1.5,\ 0.75)^T,$$

$$x^{(2)}=(1.488095,\ 0.755952)^T,$$

$$x^{(3)}=(1.488034,\ 0.755983)^T.$$

$x^{(3)}$的每一位都是有效数字.

第六章习题

P.161.　第1（迭代5次）、第3题、第7题
P.162.　第12题、第13题