科学计算/计算方法(Scientific Computing)

第十章　非线性方程求根

任课老师: Xiao-liang Cheng(程晓良)

Department of Mathematics, Zhejiang University, Hangzhou 310027, P.R. China.

2023. 5 .23

非线性方程(组)求根

1. 求实根的二分法(对半区间法)

非线性方程(组)求根

1. 求实根的二分法(对半区间法)
2. 迭代法

非线性方程(组)求根

1. 求实根的二分法(对半区间法)
2. 迭代法
3. 迭代收敛的加速

非线性方程(组)求根

非线性方程(组)求根

非线性方程(组)求根

1. 求实根的二分法(对半区间法)
2. 迭代法
3. 迭代收敛的加速
4. Newton法
5. 弦位法
6. 抛物线法

非线性方程(组)求根

1. 求实根的二分法(对半区间法)
2. 迭代法
3. 迭代收敛的加速
4. Newton法
5. 弦位法
6. 抛物线法
7. 解非线性方程组的Newton法和拟Newton法

非线性方程(组)求根

为什么要求解非线性方程的根?

例1: 求多项式的根

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

例2: 求超越方程的根

$$e^{-x} - \sin(\frac{\pi x}{2}) = 0.$$

根?单根?m重根?

基本方法:

(1) 二分法
(2) 迭代法
(3) 牛顿法
(4) 弦截法

根存在性

定理: 若函数$f(x)$ 在区间$[a,b]$ 上是连续的，且$f(a)f(b)<0$，则$f(x)=0$ 在区间$[a,b]$上至少存在一个根.

这只是一个充分条件，也是二分法求根的基础.

二分法算法

1) 计算$f(x)$ 有解区间$[a,b]$ 端点处的值$f(a),f(b)$，若$f(a)=0$ 或$f(b)=0$，则终止，若$f(a)f(b)>0$，则重新选择a或b.

2) 计算$f(x)$ 区间$[a,b]$ 中点处的值$f(x_1)$.

3) 若$f(x_1)=0$，则x_1 即是根，否则检验：

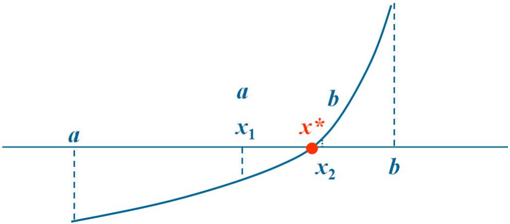
 (i) 若$f(x_1)$ 与$f(a)$ 异号，则知解位于区间 $[a,x_1], b_1=x_1, a_1=a$.

 (ii) 若$f(x_1)$ 与$f(b)$ 异号，则知解位于区间 $[x_1,b], b_1=b, a_1=x_1$.

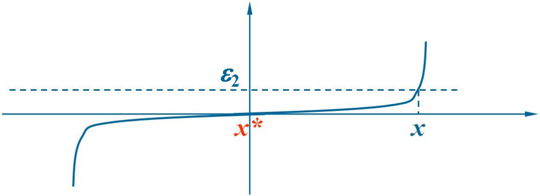
重复执行步骤2) 3), 可得：$(a,b),(a_1,b_1),\ldots,(a_k,b_k),\ldots$

4) 当$|b_{k+1}-a_{k+1}|<\varepsilon$ 时，取$x_{k+1}=\frac{1}{2}(a_k+b_k)$即为根的近似.

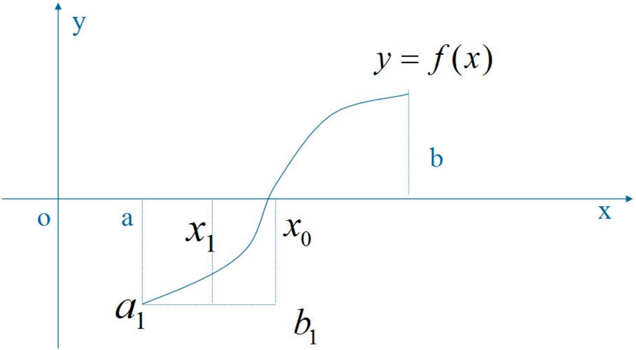
二分法



$$|x_{k+1} - x_k| < \varepsilon_1, \quad or \quad |f(x)| < \varepsilon_2.$$



二分法误差估计



$$|x^* - x_k| \leqslant \frac{1}{2^{k+1}}(b-a).$$

误差估计、对分次数

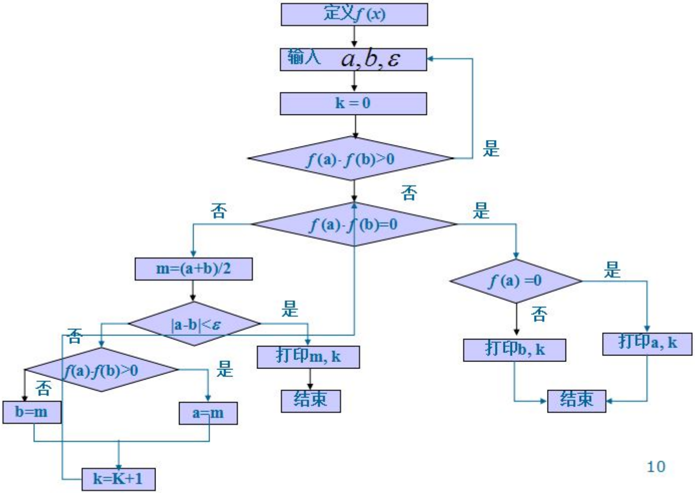
对二分法，因为数值分析的结果允许含有一定的误差。所以对给定的误差精度ε，只要选择足够大的k，就可使得

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

这时即可用x_k 作为x^* 的近似值。上式也称为二分法的误差估计式。对于事先给定的精度ε，易求得二分次数：

$$k > ln(\frac{b-a}{\varepsilon})/ln2 - 1$$

二分法程序框图


定义$f(x)$
输入 a, b, ε
k = 0
$f(a)\cdot f(b)>0$
是
否
$f(a)\cdot f(b)=0$
否
是
m=(a+b)/2
$f(a)=0$
是
否
$|a-b|<\varepsilon$
是
否
打印m, k
打印b, k
打印a, k
$f(a)\cdot f(b)>0$
是
否
b=m
a=m
结束
结束
k=K+1

例题分析

例1： 求下列方程位于$[1, 1.5]$ 内的一个根。

$$f(x) = x^3 - x - 1 = 0$$

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

例题分析

例2 利用二分法求方程$2\sin\pi x+\cos\pi x=0$，在$[0,1]$内的根($\varepsilon=0.01$)。

解： 令$f(x)=2\sin\pi x+\cos\pi x$，取$a_0=0, b_0=1$，有$f(0)>0, f(1)<0$。利用上述二分法，可以得到如下的计算结果，可以看出，当$k=7$时，

$$|b_7-a_7|=\frac{1}{128}<\varepsilon$$

所以可以将$[a_7,b_7]$ 的中点作为所求根x^*的近似。

例题分析

k	a_k	b_k	x_k	$f(x_k)$的符号
0	0	1	0.5	$+$
1	0.5	1	0.75	$+$
2	0.75	1	0.875	-
3	0.75	0.875	0.8125	$+$
4	0.8125	0.875	0.84375	$+$
5	0.84375	0.875	0.85937	-
6	0.84375	0.85937	0.85156	$+$
7	0.85156	0.85937	0.85547	

取$x^* \approx 0.8555$，即满足精度要求。

思考题

函数

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

在区间$[0, 3]$中只有一个零点$f(1) = 0$.

$$f(0) = -2, \qquad f(3) = 1.$$

但

$$f(1.5) = -\frac{1}{5}.$$

为什么二分法找不到根$x = 1$?

二分法小结

- 二分法是求实根的近似计算中行之有效的最简单方法，易于在计算机上实现;
- 对于函数的性质要求不高，仅仅要求它在有根区间上连续，且区间端点的函数值异号即可;
- 缺点是不能求偶数重根，也不能求复根，收敛速度与以 $\frac{1}{2}$ 为公比的等比数列相同，不算太快.

因此一般在求方程近似根时，不太单独使用，常用它来为其它方法求方程近似根提供好的初始值.

迭代法

迭代法是数值计算中一种典型的重要方法，尤其是计算机的普遍使用，使迭代法的应用更为广泛.

所谓迭代法就是用某种收敛于所给问题的精确解的极限过程，逐步逼近的一种计算方法，从而可以用有限个步骤算出精确解的具有指定精度的近似解. 简单说迭代法是一种逐步逼近的方法.

将方程$f(x)=0$改写成等价形式(不唯一)

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

于是可用迭代

$$x_{n+1}=\phi(x_n),\ \ n=0,1,2,\cdots.$$

迭代法

例： 求方程$x^3-x-1=0$在$x=1.5$附近的一个根（用六位有效数字计算）.

解： 改写方程$x=\sqrt[3]{x+1}$
将$x_0=1.5$ 代入，$x_1=\sqrt[2]{x_0+1}=1.35721$
将$x_1=1.35721$ 代入，$x_2=\sqrt[3]{x_1+1}=1.33086$
将$x_2=1.33086$ 代入，$x_3=\sqrt[3]{x_2+1}=1.32588$
重复步骤，$x_{k+1}=\sqrt[3]{x_k+1}, k=0,1,2,\ldots$

迭代法

k	x_k	k	x_k
0	1.5	5	1.32476
1	1.35721	6	1.32473
2	1.33086	7	1.32472
3	1.32588	8	1.32472
4	1.32494		

迭代法

对于一般形式的方程$f(x)=0$，先将方程化为$x=\phi(x)$，再从某一数x_0 出发，作序列$\{x_n\}$，$x_{n+1}=\phi(x_n), n=0,1,2,\ldots$，若序列有极限，即$\lim_{n\to\infty} x_n=a$，则可得$a=\phi(a)$，即$f(a)=0$，即$a$是方程的根. 称$x_0$ 为初始近似，称x_n为n次近似，称$\phi(x)$ 为迭代函数，称$x_{n+1}=\phi(x_n)$ 为迭代公式.

迭代法的结束条件:

$$|x_k-x_{k-1}|<\varepsilon.$$

迭代法例题

例: 求方程$f(x)=x^3-x-1=0$ 在$x_0=1.5$ 附近的根x^*。
解:
(1) 将方程改写为$x=\sqrt[3]{x+1}$，由此建立迭代公式$x_{k+1}=\sqrt[2]{x_k+1}, k=0,1,2\ldots,$

k	0	1	2	...	7	8
x_k	1.5	1.357217	1.33086	...	1.32472	1.32472

迭代收敛。
(2) 若将方程改写为$x=x^3-1$，建立迭代公式$x_{k+1}=x_k^3-1$，

k	0	1	2	...
x_k	1.5	2.375	12.39	...

迭代不收敛。

迭代过程的收敛性

定理10.1： 设函数$\varphi(x) \in C[a,b]$ 满足如下条件，
(1) 当$x \in [a,b]$ 时，$\varphi(x) \in [a,b]$;
(2) $\varphi(x)$ 在$[a,b]$ 上满足李普希斯条件，即对任意$x_1, x_2 \in [a,b]$，成立$|\varphi(x_1) - \varphi(x_2)| \leqslant L|x_1 - x_2|$，其中$0 < L < 1$. 称L为李普希斯常数.
则方程$x = \phi(x)$在$[a,b]$上存在唯一解x^*，并且迭代格式$x_{k+1} = \phi(x_k)$收敛，并有误差估计

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

推论: 证明过程中有

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

收敛性证明

证明: 先证x^* 的存在性. 由条件可知，函数$\varphi(x)$在$[a,b]$ 上是连续函数，令$g(x)=x-\varphi(x)$，则$g(x)$ 在$[a,b]$ 上也连续，由条件(1)可得，

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

由连续函数的性质可知，在$[a,b]$ 上必至少存在一点x^*，使$g(x^*)=0$，即$x^*=\varphi(x^*)$。

收敛性证明

再证其唯一性. 设有x_1^*, x_2^*，满足$x_1^* = \varphi(x_1^*), x_2^* = \varphi(x_2^*)$，则由Lipschitz连续性质得，

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

即$x_1^* - x_2^* = 0$，因此x^* 唯一。

收敛性证明

最后来证明$\{x_k\}$ 的收敛性. 由LIpschitz连续性质，

$$|x^* - x_{k+1}| = |\varphi(x^*) - \varphi(x_k)| \leq L|x^* - x_k|, \ k = 0, 1, 2, \cdots .$$

由归纳法可得，

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

因为$L < 1$，所以
有$\lim_{k\to\infty} |x^* - x_k| \leqslant \lim_{k\to\infty} L^k|x^* - x_0| = 0$，
即$\lim_{k\to\infty} x_k = x^*$，迭代收敛.

收敛性证明

误差估计式:

因为

$$\begin{aligned} |x_k - x^*| &= |x_k - x_{k+1} + x_{k+1} - x^*| \\ &\le |x_{k+1} - x_k| + |x_{k+1} - x^*| \\ &\le L|x_k - x_{k-1}| + L|x_k - x^*|. \end{aligned}$$

即

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

对k递推.

收敛定理的另一种形式

定理: 设迭代函数$\varphi(x)$ 在$[a,b]$ 上具有连续的一阶导数，且

(1) 当$a \leqslant x \leqslant b$ 时，有$a \leqslant \varphi(x) \leqslant b$.

(2) 存在正数$L<1$，使对任意的$x \in [a,b]$，有$|\varphi'(x)| \leqslant L < 1$ 成立，则$\varphi(x)$ 在$[a,b]$ 上存在唯一的解，并且对任意选取的初值$x_0 \in [a,b]$，迭代过程$x_{k+1} = \varphi(x_k)$ 所产生的迭代数列收敛于x^*.

例题分析

例: 求方程$9x^2-\sin x-1=0$ 在$[0,1]$ 内的根.

解: 将方程写成迭代格式$x_{n+1}=\frac{1}{3}\sqrt{\sin x_n+1}$，由于

$$|\varphi'(x)|=\frac{1}{6}\frac{|\cos x|}{\sqrt{\sin x+1}}\leqslant\frac{1}{6}$$

迭代法收敛。任取初值，如$x_0=0.4$，利用程序进行计算，得到一列近似值。当$k=14$ 时，有$x_{14}=0.391846907002651$8，可以看做方程很好的近似根。

迭代过程的局部收敛性

在实际应用迭代法时，通常首先在根x^* 的邻近考察。称一种迭代过程在根x^* 邻近收敛，如果存在邻域$\triangle : |x - x^*| \leqslant \delta$，使迭代过程对于任意初值$x_0 \in \triangle$ 均收敛，这种在根的邻近所具有的收敛性被称为局部收敛性.

定理: 设$\varphi(x)$ 在$x = \varphi(x)$ 的根x^* 邻近有连续导数，且成立

$$|\varphi'(x^*)| < 1$$

则迭代过程$x_{k+1} = \varphi(x_k)$ 在x^* 邻近具有局部收敛性。

迭代法小结

迭代法是一种逐次逼近法，这种方法使用某个固定公式—所谓迭代公式反复校正根的近似值，使之逐步精确化，直至满足精度要求的结果。

迭代法的求根过程分成两步，第一步先提供根的某个猜测值，即所谓迭代初值，然后将迭代初值逐步加工成满足精度要求的根.

迭代法的设计思想是，将隐式方程$x = \varphi(x)$ 归结为计算一组显式公式$x_{k+1} = \varphi(x_k)$, 也就是说，迭代过程实质上是一个逐步显式化的过程.

收敛速度的概念

一个具有实用价值的迭代法，不断要求它收敛，而且还要求它收敛比较快。迭代收敛的速度是指迭代误差的下降速度. 在定理10.1中我们知道L 越小收敛越快，但要准确反映收敛速度，还需要引进收敛阶的概念，它是衡量迭代好坏的标志之一.

迭代过程的收敛速度

定义： 设由某方法确定的序列$\{x_k\}$ 收敛于方程的根x^*，如果存在正实数p，使得

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

C为非零常数. 则称序列$\{x_k\}$ 收敛于x^* 的收敛速度是p阶的，或称该方法具有p 阶敛速. 当$p = 1$ 时，称该方法为线性（一次）收敛；当$p = 2$ 时，称方法为平方（二次）收敛; 当$1 < p < 2$ 时，称方法为超线性收敛.

迭代过程的收敛速度

定理: 若$\varphi(x)$ 在x^* 某个邻域内有$p(p \geqslant 1)$阶连续导数，且

$\varphi(x^*) = x^*, \quad \varphi^{(k)}(x^*) = 0, \quad k = 1, 2, 3, ..., p-1. \quad \varphi^{(p)}(x^*) \neq 0$

则对一个任意靠近x^* 的初始值，迭代公式

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

是p阶收敛的，且有

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

如果$p = 1$，要求$|\varphi'(x^*)| < 1$.

迭代过程的收敛速度

证明： 显然迭代格式是局部收敛的. 应用泰勒展开式有

$$
\begin{aligned}
x_{k+1} = \varphi(x_k) =& \varphi(x^*) + \varphi'(x^*)(x_k - x^*) + \ldots \\
& + \frac{\varphi^{(p-1)}(x^*)}{(p-1)!}(x_k - x^*)^{p-1} + \frac{\varphi^{(p)}(\xi)}{p!}(x_k - x^*)^p
\end{aligned}
$$

其中，ξ 位于x^* 与x_k 之间，再由条件(3.18)得

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

迭代过程的收敛速度

即

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

两边取极限得

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

如果记$e_k = x_k - x^*$，则上式还可以记为

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

迭代的加速

对于$x = \phi(x)$, 收敛速度和L(一阶导数的界或Lipschitz常数)有关.
如何加速?
设$\lambda \neq -1$, 令

$$x = \frac{\lambda}{1+\lambda}x + \frac{1}{1+\lambda}\phi(x) =: \psi(x).$$

而

$$\psi'(x) = \frac{1}{1+\lambda}(\lambda + \phi'(x)).$$

假定$\phi'(x)$连续，如果选取$\lambda_n = -\phi'(x_n) \neq -1$, 则

$$x_{n+1} = (1-\omega_n)x_n + \omega_n\phi(x_n), \quad \omega_n = \frac{1}{1+\lambda_n},$$

就是松弛方法.

Aitken加速

假如求得x_n后，

$$x_{n+1}^{(1)} = \phi(x_n), \qquad x_{n+1}^{(2)} = \phi(x_{n+1}^{(1)}),$$

记

$$x_n - x^* = \varepsilon,$$

则

$$x_{n+1}^{(1)} - x^* = \phi(x_n) - \phi(x^*) \approx \phi'(x^*)\varepsilon = c\varepsilon,$$

$$x_{n+1}^{(2)} - x^* = \phi(x_n^{(1)}) - \phi(x^*) \approx \phi'(x^*)(x_{n+1}^{(1)} - x^*) = c^2\varepsilon,$$

于是

$$x_{n+1}^{(1)} - x_n \approx (c-1)\varepsilon, \quad x_{n+1}^{(2)} - x_{n+1}^{(1)} \approx c(c-1)\varepsilon.$$

$$c^2\varepsilon \approx \frac{(x_{n+1}^{(2)} - x_{n+1}^{(1)})^2}{(x_{n+1}^{(2)} - x_{n+1}^{(1)}) - (x_{n+1}^{(1)} - x_n)}.$$

得到一个改善的解

$$x^* \approx x_{n+1}^{(2)} - \frac{(x_{n+1}^{(2)} - x_{n+1}^{(1)})^2}{x_{n+1}^{(2)} - 2x_{n+1}^{(1)} + x_n}.$$

即

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

还有一个形式

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

数值例子说明加速的有效性. P.218-219

设迭代法$x_{k+1}=\phi(x_k)$，则Aitken加速$x_{k+1}=\psi(x_k)$，其中

$$\psi(x)=x-\frac{(\phi(x)-x)^2}{\phi(\phi(x))-2\phi(x)+x}.$$

假设x_*是准确解，于是

$$\begin{aligned}\phi(x) &= \phi(x_*)+\phi'(x_*)(x-x_*)+O((x-x_*)^2)\\ &= x_*+\phi'(x_*)(x-x_*)+O((x-x_*)^2);\end{aligned}$$

$$\begin{aligned}\phi(\phi(x)) &= \phi(x_*+\phi'(x_*)(x-x_*)+O((x-x_*)^2))\\ &= x_*+(\phi'(x_*))^2(x-x_*)+O((x-x_*)^2).\end{aligned}$$

则

$$\psi(x)-x_*=(x-x_*)-\frac{((1-\phi'(x_*))(x-x_*)+O((x-x_*)^2))^2}{(1-\phi'(x_*))^2(x-x_*)+O((x-x_*)^2}.$$

$$\psi(x)-x_*=O((x-x_*)^2).$$

Newton法

一般迭代法$x_{k+1} = \phi(x_k)$, 如何选择合适的迭代函数$\varphi(x)$，是提高迭代数列$\{x_k\}$ 的收敛速度的关键.

本节介绍一种确定迭代函数$\varphi(x)$ 的方法——牛顿法。牛顿法是求解方程$f(x) = 0$ 的一种重要方法，它的最大优点是方程在单根附近具有较高的收敛速度，它还可以用于求代数方程的重根、复根；也可以拓广用于求解非线性方程组的问题。

牛顿法的导出: 线性化

取$x_0 \approx x^*$，将$f(x)$ 在x_0 做一阶Taylor展开：

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi)}{2!}(x - x_0)^2.$$

ξ 在x_0 和x 之间。将$(x^* - x_0)^2$ 看成高阶小量，则有：

$$0 = f(x^*) \approx f(x_0) + f'(x_0)(x^* - x_0) \Rightarrow x^* \approx x_0 - \frac{f(x_0)}{f'(x_0)}$$

牛顿公式: $x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)}$

只要$f \in C^1$ 每一步迭代都有$f'(x_k) \neq 0$，而且$\lim_{k\to\infty} x_k = x^*$，则$x^*$ 就是$f(x)$ 的根。

牛顿法的几何表示

从几何角度来分析一下牛顿公式的直观结构。方程$f(x)=0$ 的根就是曲线$y=f(x)$ 与x 轴的交点。设x_k 为x^* 的一个近似值，过曲线$y=f(x)$ 上横坐标为x_k 的点$p(x_k,f(x_k))$，引一条切线，其方程为

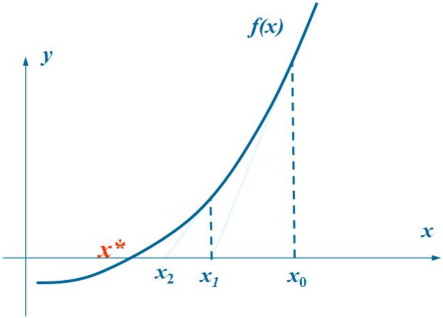
$$y=f(x_k)+f'(x_k)(x-x_k)$$

令其为零，得切线l 与x 轴的交点为

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

牛顿法的几何表示

将此式$x = x_k - \frac{f(x_K}{f'(x_k)}$ 与上面所求得的牛顿公式$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$ 进行比较即可知道：牛顿公式实际上就是用曲线$y = f(x)$ 在点$p(x_k, f(x_k))$ 处的切线与x 轴的交点作为曲线$y = f(x)$ 与x 轴交点的近似，如下图所示。



牛顿法的另一个导出方法

令$x = x + h(x)f(x)$, $h(x^*) \neq 0$, 则$\phi(x) = x + h(x)f(x)$,

$$x_{n+1} = \phi(x_n);$$

计算

$$\phi'(x) = 1 + h(x)f'(x) + h'(x)f(x).$$

取$h(x) = -\frac{1}{f'(x)}$时， $\phi'(x^*) = 0$.

即Newton迭代法:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \cdots .$$

牛顿法例题

例 用牛顿法求解方程$x = e^{-x}$ 在$x_0 = 0.5$ 附近的根($\varepsilon = 10^{-5}$)。
解： 将方程$x = e^{-x}$ 转化为等价方程$xe^x - 1 = 0$，
令$f(x) = xe^x - 1$，则牛顿迭代公式为

$$x_{k+1} = x_k - \frac{x_k e^{x_k} - 1}{e^{x_k}(x_k + 1)} = x_k - \frac{x_k - e^{-x_k}}{1 + x_k}, \quad k = 0, 1, 2, \ldots$$

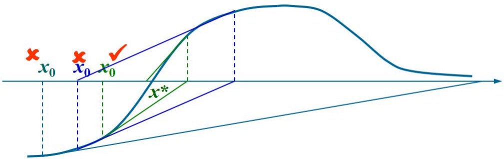
$x_0 = 0.5$，迭代结果如下表。

k	0	1	2	3	4
x_k	0.5	0.5710204	0.5671555	0.5671433	0.5671432

$x^* \approx 0.567143$，与前面的迭代结果进行比较可见，牛顿迭代公式的收敛速度是相当快的。

牛顿法收敛条件

疑问：牛顿法是否永远收敛呢？如果不是，那么与谁有直接关系呢？



结论：牛顿法的收敛性依赖于x_0 的选取.

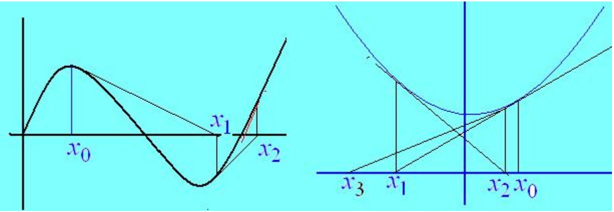
疑问：牛顿法是否有可能永远不收敛呢？
例. $f(x) = x^{\frac{1}{3}}$有唯一零点$x = 0$, Newton法

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

对所有初值$x_0 \neq 0$都不收敛.

牛顿法的缺点

(1) 牛顿迭代法存在从一个根跳到另一个根的情况。
(2) 如果$f(x)=0$ 没有实根，则牛顿迭代序列不收敛。
(3) 每次迭代都要求导，程序易中断。
(4) 对初始值的敏感性。混沌的例子，上机作业.



上机作业四

在复平面上，用Newton法求$z^3-1=0$的根

$$z_{n+1}=z_n-\frac{z_n^3-1}{3z_n^2},\quad n=0,1,2,\cdots$$

取初值($n=30,40$等)

$$z_0=\frac{k}{n}+i\frac{l}{n},\quad k,l=0,\pm1,\pm2,\cdots,\pm n.$$

这里需要去掉原点$z_0=0$. 根据应用Newton法时收敛到三个不同根的情况，用三种不同颜色标出z_0的区域.

牛顿法的收敛性

定理: 设$f(x)$ 在$[a,b]$ 上满足下列条件:

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

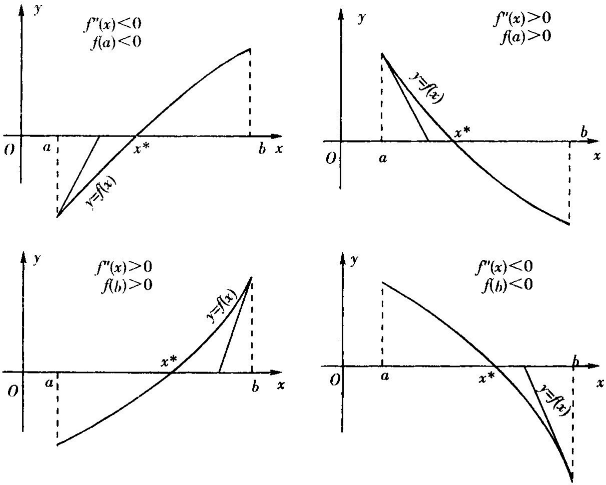
(2) $f'(x)\neq 0$;

(3) $f''(x)$ 存在且不变号;

(4) 取$x_0\in[a,b]$,使得$f''(x)\cdot f(x_0)>0$.

则牛顿迭代序列$\{x_k\}$ 收敛于$f(x)$ 在$[a,b]$ 上的唯一根x^*。

牛顿法收敛性示意图


y
f''(x)<0
f(a)<0
O
a
x*
b
x
y=f(x)
y
f''(x)>0
f(a)>0
y=f(x)
x*
b
O
a
x
y
f''(x)>0
f(b)>0
y=f(x)
x*
O
a
b
x
y
f''(x)<0
f(b)<0
O
a
x*
b
x
y=f(x)

牛顿法的收敛速度

迭代过程的收敛速度依赖于迭代函数$\varphi(x)$ 的选取. 如果当$x \in [a,b]$ 时$\varphi'(x) \neq 0$，则该迭代过程只可能是线性收敛的。

对牛顿公式$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$

其迭代函数为$\varphi(x) = x - \frac{f(x)}{f'(x)}$

由于$\varphi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$，假定$x^*$ 是$f(x)$ 的一个单根，即$f(x^*) = 0$，$f'(x^*) \neq 0$，则由上式知$\varphi'(x^*) = 0$，由上述定理知，牛顿法在根x^* 的邻近至少是平方收敛的。

重根情形

若x^*是$f(x)=0$的m重根，则Newton法的收敛速度?

计算$\varphi'(x)=\frac{f(x)f''(x)}{[f'(x)]^2}$，设$f(x)=(x-x^*)^m h(x)$, $h(x^*)\neq 0$, 于是

$$f'(x)=m(x-x^*)^{m-1}h(x)+O((x-x^*)^m);$$

$$f''(x)=m(m-1)(x-x^*)^{m-2}h(x)+O((x-x^*)^{m-1}).$$

因此

$$\varphi'(x^*)=\lim_{x\to x^*}\frac{f(x)f''(x)}{(f'(x))^2}=\frac{m-1}{m}<1.$$

即重根时，Newton法收敛，但只有一阶.

重根情形

修正I: 已知重根的重数$m \geq 2$,

$$x_{n+1} = x_n - m\frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \cdots .$$

修正II: 令$h(x) = \frac{f(x)}{f'(x)}$, x^*就是$h(x)$的单根，将Newton法应用到$h(x)$上，得

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

牛顿法的应用

1. $\sqrt{2}$的计算:

问题转化为解二次方程$x^2-2=0$的正根. 应用Newton法

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

现证此迭代公式对于初值$x_0>0$ 都是收敛的. 由迭代公式得

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

$$\Rightarrow\frac{x_{k+1}-\sqrt{2}}{x_{k+1}+\sqrt{2}}=(\frac{x_k-\sqrt{2}}{x_k+\sqrt{2}})^2\Rightarrow\frac{x_k-\sqrt{2}}{x_k+\sqrt{2}}=(\frac{x_0-\sqrt{2}}{x_0+\sqrt{2}})^{2^k}$$

记$q=\frac{x_0-\sqrt{2}}{x_0+\sqrt{2}}\Rightarrow x_k-\sqrt{2}=2\sqrt{2}\frac{q^{2^k}}{1-q^{2^k}}$ 对任意$x_0>0$，总有$|q|<1$，故当$k\to\infty$时，$x_k\to\sqrt{2}$。

牛顿法的应用

2. 不用除法，计算$\frac{1}{a}$, $a > 0$:

转化为方程$\frac{1}{x} - a = 0$的根，应用Newton法

$$x_{k+1} = x_k - \frac{\frac{1}{x_k} - a}{-\frac{1}{x_k^2}} = 2x_k - ax_k^2, \ \ k = 0, 1, 2, \cdots .$$

则

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

牛顿法的优缺点

Newton法具有收敛快，稳定性好，精度高等优点，是求解非线性方程的最有效方法之一. 但它每次迭代均需计算函数值与导数值，故计算量较大. 而且当导数值提供有困难时，Newton法无法进行.

弦位法(弦截法)(割线法)

基本思想：利用一些函数值$f(x_k), f(x_{k-1}), ...$来回避导数值$f'(x_k)$ 的计算。

设$x_k, x_{k-1}, ..., x_{k-r}$是$f(x) = 0$ 的一组近似根，利用函数值$f(x_k), f(x_{k-1}), ..., f(x_{k-r})$，构造插值多项式$p_r(x)$，并适当选取$p_r(x) = 0$ 的一个根作为$f(x) = 0$ 的新的近似根x_{k+1}，这就确定了一个迭代过程，记迭代函数为φ:

$$x_{k+1} = \varphi(x_k, x_{k-1}, ..., x_{k-r})$$

当$r = 1$ 时为弦截法，当$r = 2$ 为抛物线法.

弦截法

设x_k, x_{k-1} 是$f(x)=0$ 的近似根，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$f(x)=0$ 的新的近似根x_{k+1}. 由

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

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

此迭代公式可看做牛顿公式$x_{k+1}=x_k-\frac{f(x_k)}{f'(x_k)}$ 中的导数$f'(x)$ 用差商$\frac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}$取代的结果.

弦截法的又一推导思想

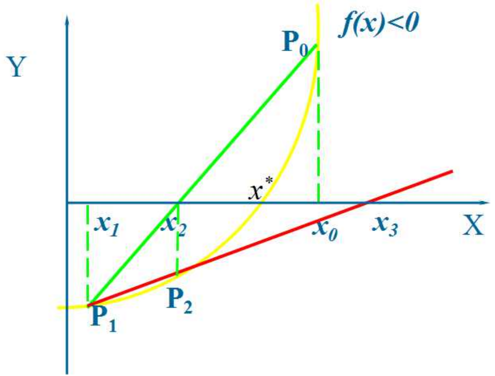
在牛顿迭代格式中，将曲线$y=f(x)$ 上点$(x_k, f(x_k))$ 的切线斜率$f'(x_k)$，改为其两点连线（弦）的斜率

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

则得双点或单点弦割法迭代格式

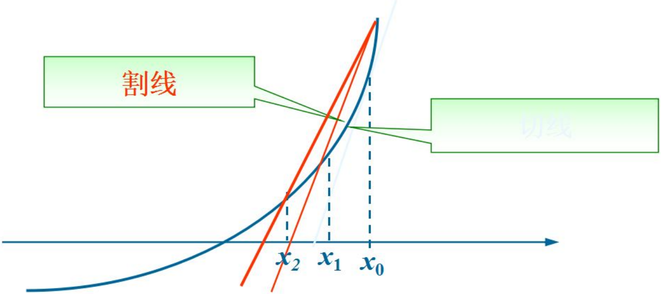
$$x_{k+1}=x_k-\frac{f(x_k)(x_k-x_{k-1})}{f(x_k)-f(x_{k-1})}$$

弦截法的几何表示



弦截法在求x_{k+1}时要用到前面两步的结果x_k, x_{k-1}，需要两个初值x_0, x_1，而牛顿切线法在计算x_{k+1}时，只用到前一步的x_k值.

单点弦截法

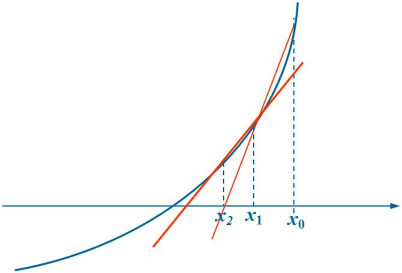


切线斜率≈割线斜率

$$\Rightarrow f'(x_k) \approx \frac{f(x_k) - f(x_0)}{x_k - x_0}$$

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

双点弦截法



切线斜率≈割线斜率，需要两个初值x_0, x_1.

$$\Rightarrow f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

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

误差

如果$f(x)$在零点x^*附近有连续的二阶微商，$f'(x^*) \neq 0$, 且初始值x_0, x_1充分接近x^*, 那么弦位法的迭代过程是收敛的，且有阶$p = \frac{1+\sqrt{5}}{2} \approx 1.618$.

$$|x_{k+1} - x^*| \approx |f''(x^*)/(2f'(x^*))|^{0.618} |x_k - x^*|^{1.618}.$$

Newton法的"单步割线法"

设x_*是$f(x)=0$的单根，$f(x)$在x_*的某邻域内三次连续可微，则

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

$$f'(x)\simeq\frac{f(x+f(x))-f(x)}{f(x)},$$

得到

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

可以证明：对充分接近x_*的初值，方法收敛且至少是2阶的.

例题分析

例12 用快速弦截法求方程$xe^x-1=0$ 在$x=0.5$ 附近的根($\varepsilon=10^{-4}$).

解： 取$x_0=0.5, x_1=0.6$，由迭代公式求得下表

k	x_k	x_k-x_{k-1}
0	0.5	
1	0.6	
2	0.56754	0.03246
3	0.56715	-0.00036
4	0.56714	-0.00001

故$x^*\approx 0.56714$，满足精度要求。

抛物线法

思想: 前面的弦截法$r=2$的情形. 已知$(x_0, f(x_0))$, $(x_1, f(x_1))$, $(x_2, f(x_2))$三点，构造一个二次插值多项式$p(x)$, 求解$p(x)=0$得到点x_3, 再利用$(x_i, f(x_i))$, $i=1,2,3$插值计算x_4, 依次进行.

当$f(x)$在单根附近有三次导数，且初值接近解，则有

$$|x_{n+1}-x^*| \approx K|x_n-x^*|^{1.84}.$$

解非线性方程组的Newton法和拟Newton法

二元非线性方程组

$$\begin{cases} f_1(x,y)=0, \\ f_2(x,y)=0. \end{cases}$$

解非线性方程的Newton法是把非线性问题线性化，用二元Taylor展开，并取其线性部分.

$$\begin{cases} f_1(x,y)\approx f_1(x_0,y_0)+\dfrac{\partial f_1(x_0,y_0)}{\partial x}(x-x_0)+\dfrac{\partial f_1(x_0,y_0)}{\partial y}(y-y_0), \\ f_2(x,y)\approx f_2(x_0,y_0)+\dfrac{\partial f_2(x_0,y_0)}{\partial x}(x-x_0)+\dfrac{\partial f_2(x_0,y_0)}{\partial y}(y-y_0). \end{cases}$$

$$\begin{cases} f_1(x,y)\approx 0, \\ f_2(x,y)\approx 0. \end{cases}$$

解上面的线性方程组，得Newton法

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} \frac{\partial f_1(x_0,y_0)}{\partial x} & \frac{\partial f_1(x_0,y_0)}{\partial y} \\ \frac{\partial f_2(x_0,y_0)}{\partial x} & \frac{\partial f_2(x_0,y_0)}{\partial y} \end{pmatrix}^{-1} \begin{pmatrix} f_1(x_0,y_0) \\ f_2(x_0,y_0) \end{pmatrix}.$$

和一维形式相同

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

一般情况：考虑$\mathbf{f}(\mathbf{x})=\mathbf{0}$,

$$\mathbf{f}(\mathbf{x})=\begin{pmatrix} f_1(x_1,x_2,\cdots,x_n) \\ f_2(x_1,x_2,\cdots,x_n) \\ \vdots \\ f_n(x_1,x_2,\cdots,x_n) \end{pmatrix}.$$

记

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

Newton法:

$$\mathbf{x}^{(\mathbf{k+1})} = \mathbf{x}^{(\mathbf{k})} - [\mathbf{f}'(\mathbf{x}^{(\mathbf{k})})]^{-1}\mathbf{f}(\mathbf{x}^{(\mathbf{k})}), \ \ k = 0, 1, 2, \cdots .$$

这里$\mathbf{x}^{(\mathbf{0})}$是初始猜测值，假设矩阵$\mathbf{f}'(\mathbf{x})$可逆.

Newton法每次需要计算导数，一维时用差商代替导数的办法来得到割线法. 对多维也可以类似处理:

取$\mathbf{A}_1 \approx \mathbf{f}'(\mathbf{x})$满足

$$\mathbf{A}_1\mathbf{s} = \mathbf{y}, \quad \mathbf{s} = \mathbf{x}^{(1)} - \mathbf{x}^{(0)}, \quad \mathbf{y} = \mathbf{f}(\mathbf{x}^{(\mathbf{1})}) - \mathbf{f}(\mathbf{x}^{(\mathbf{0})}).$$

若$\mathbf{A}_1$可逆，就取

$$\mathbf{x}^{(\mathbf{2})} = \mathbf{x}^{(\mathbf{1})} - \mathbf{A}_1^{-1}\mathbf{f}(\mathbf{x}^{(\mathbf{1})}).$$

注意到，$\mathbf{A}_1$的确定不是唯一的，这就给我们选择. 这类方法通称拟Newton法. 下面介绍一种称为Broyden秩1修正法.

Broyden秩1修正法

对$\mathbf{A}_0$作适当的修正就得到$\mathbf{A}_1$:

设$\mathbf{u}$待定，取

$$\mathbf{A}_1 - \mathbf{A}_0 = \mathbf{u}\mathbf{s}^T.$$

$$(\mathbf{A}_1 - \mathbf{A}_0)\mathbf{s} = \mathbf{u}\mathbf{s}^T\mathbf{s}.$$

由拟Newton法可得

$$\mathbf{u} = \frac{\mathbf{y} - \mathbf{A}_0\mathbf{s}}{\mathbf{s}^T\mathbf{s}}.$$

$$\mathbf{A}_1 = \mathbf{A}_0 + \frac{(\mathbf{y} - \mathbf{A}_0\mathbf{s})\mathbf{s}^T}{\mathbf{s}^T\mathbf{s}}.$$

Sherman-Morrison公式

设$\mathbf{A}$是非奇异n阶矩阵，$\mathbf{u}, \mathbf{v}$是n维向量，$\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} \neq -1$, 则

$$(\mathbf{A}+\mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}$$

$$\mathbf{H}_1 = \mathbf{A}_1^{-1} = \mathbf{H}_0 - \frac{(\mathbf{H}_0\mathbf{y}-\mathbf{s})\mathbf{s}^T\mathbf{H}_0}{\mathbf{s}^T\mathbf{H}_0\mathbf{y}}, \qquad \mathbf{H}_0 = \mathbf{A}_0^{-1}.$$

最速下降法

令

$$\Phi(x,y) = (f_1(x,y))^2 + (f_2(x,y))^2,$$

则通过求$\Phi(x,y)$的零极小值点来求非线性方程组的解.

假定Φ的解附近具有二阶连续导数，并且它的Hessian矩阵主子式皆非负

$$\frac{\partial^2\Phi}{\partial x^2} \geq 0, \quad \frac{\partial^2\Phi}{\partial y^2} \geq 0, \quad \begin{vmatrix} \frac{\partial^2\Phi}{\partial x^2} & \frac{\partial^2\Phi}{\partial x\partial y} \\ \frac{\partial^2\Phi}{\partial y\partial x} & \frac{\partial^2\Phi}{\partial y^2} \end{vmatrix} \geq 0.$$

函数$\Phi(x,y)$在几何上是一空间曲面，它与xOy面相切的点即是它的零极小值点.

从定义域D内的某一点(x_0, y_0), 沿着使$\Phi(x, y)$值下降的方向逐步下降Φ值，一直降到它的零极小值，那么就可以得到所求问题的解.

函数$\Phi(x, y)$在(x, y)处的梯度方向

$$g = (\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y})^T$$

是使Φ值上升最快的方向，因此其反方向就是使Φ值下降最快的方向.

设(x_0, y_0)是解的一个近似值，计算Φ在此点的梯度值$\mathbf{g}_0 = (g_{10}, g_{20})^T$,

$$\begin{cases} g_{10} = \frac{\partial \Phi}{\partial x}|_{(x_0,y_0)} = 2\Big((\frac{\partial f_1}{\partial x})f_1 + (\frac{\partial f_2}{\partial x})f_2\Big)|_{(x_0,y_0)} \\ g_{20} = \frac{\partial \Phi}{\partial y}|_{(x_0,y_0)} = 2\Big((\frac{\partial f_1}{\partial y})f_1 + (\frac{\partial f_2}{\partial y})f_2\Big)|_{(x_0,y_0)} \end{cases}$$

从(x_0, y_0)出发，沿负梯度方向$-\mathbf{g}_0$跨一适当的步长λ,

$$\begin{cases} x_1 = x_0 - \lambda g_{10} \\ y_1 = y_0 - \lambda g_{20} \end{cases}$$

选取λ使得

$$\Phi(x_1, y_1) = \min_{\lambda \in \mathbb{R}^1} \Phi(x_0 - \lambda g_{10}, y_0 - \lambda g_{20}).$$

将$f_1(x_0 - \lambda g_{10}, y_0 - \lambda g_{20})$和$f_2(x_0 - \lambda g_{10}, y_0 - \lambda g_{20})$在$(x_0, y_0)$处展开，略去$\lambda^2$项及高级项，把$\Phi$对$\lambda$求导得

$$\lambda = \frac{(g_{10}\frac{\partial f_1}{\partial x} + g_{20}\frac{\partial f_1}{\partial y})f_1 + (g_{10}\frac{\partial f_2}{\partial x} + g_{20}\frac{\partial f_2}{\partial y})f_2}{(g_{10}\frac{\partial f_1}{\partial x} + g_{20}\frac{\partial f_1}{\partial y})^2 + (g_{10}\frac{\partial f_2}{\partial x} + g_{20}\frac{\partial f_2}{\partial y})^2}\Bigg|_{(x_0, y_0)}$$