

# 第5章 非线性方程(组)

非线性方程(组)是科学技术与工程计算中的重要问题之一,求解方法远比线性方程(组)复杂,通常都要使用数值迭代技术。本章主要介绍二分法、牛顿法、拟牛顿法和不动点法,并讨论它们的收敛性和计算复杂性。这些迭代技术在实践中非常重要,应用十分广泛。阅读本章的数学基础是初等微分学。

## 5.1 非线性方程

### 5.1.1 二分法

**定义 5.1.1** 如果  $f(\xi)=0$ , 则称  $\xi$  是方程  $f(x)=0$  的一个根(解), 或者说是函数  $y=f(x)$  的一个零点。若  $f(x)$  是非线性函数, 则称  $f(x)=0$  为非线性方程。

二分法是求解非线性方程最简单的实用方法, 其理论基础源于微分学中的介值定理和实数连续性的闭区间套定理。

介值定理: 如果  $f(x)$  在闭区间  $[a, b]$  上连续, 则对  $f(a)$  和  $f(b)$  之间的任何一个数  $y$ , 都存在  $\xi \in [a, b]$  使得  $f(\xi)=y$ , 即  $f(x)$  能取到  $f(a)$  和  $f(b)$  之间的所有数。

闭区间套定理: 设  $[a_n, b_n] \supset [a_{n+1}, b_{n+1}], n=1, 2, \dots$ , 是一个闭区间套, 若  $\lim_{n \rightarrow \infty} (b_n - a_n) = 0$ , 则存在唯一  $\xi \in [a_n, b_n], n=1, 2, \dots$ , 即长度趋于零的闭区间套的交是单点集,  $\{\xi\} = \bigcap_{n=1}^{\infty} [a_n, b_n]$ 。

若  $f(x)$  是闭区间  $[a, b]$  上的连续函数, 且在端点异号  $f(a) \cdot f(b) < 0$ , 则 0 必在  $f(a)$  和  $f(b)$  之间, 根据介值定理存在  $\xi \in [a, b]$  使得  $f(\xi)=0$ , 即方程  $f(x)=0$  在  $[a, b]$  上有根。

在  $[a, b]$  上找根: 首先取中点  $c=(a+b)/2$ , 如果  $f(c)=0$  则找到一个根  $c$ ; 否则比较  $f(c)$  和  $f(a)$  的符号: 若  $f(a) \cdot f(c) < 0$ , 令  $a_1=a, b_1=c$ , 否则  $a_1=c, b_1=b$ 。于是, 得到一个新区间  $[a_1, b_1] \subset [a, b]$ , 且

$$b_1 - a_1 = \frac{b-a}{2}, \quad f(a_1) \cdot f(b_1) < 0$$

在  $[a_1, b_1]$  上重复前面的操作, 并将这个过程继续下去, 要么得到一个根  $c$ ; 要么得到一个区间套

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

$$b_n - a_n = \frac{b-a}{2^n} \rightarrow 0 (n \rightarrow \infty)$$

$$f(a_n) \cdot f(b_n) < 0$$

在第二种情况下, 由闭区间套定理存在  $\xi \in [a_n, b_n]$ , 且  $f(\xi)$  在  $f(a_n)$  和  $f(b_n)$  之间, 因此必

有  $f(\xi)=0$ 。不难看出

$$\left| \frac{a_n + b_n}{2} - \xi \right| \leq \frac{b-a}{2^{n+1}} \quad (5.1)$$

因此,根是闭区间套中点的极限

$$\xi = \lim_{n \rightarrow \infty} \frac{a_n + b_n}{2}$$

这就是二分法的寻根过程,首先找有根区间,然后二分区间将寻根范围缩小到原区间的二分之一,继续这个过程直至达到期望精度,得到近似根  $\xi \approx (a_n + b_n)/2$ 。在寻根过程中,不需要观察函数曲线的性态,只须计算函数值  $y_n = f((a_n + b_n)/2)$ ,算法极为简单。以下是二分法的概述。

### 二分法

给定初始区间  $[a, b]$ ,  $f(a) \cdot f(b) < 0$ ; 控制误差  $\delta$

(1)  $(b-a)/2 > \delta$ , 置  $c = (a+b)/2$ 。若  $f(c) = 0$ , 输出根  $c$ ; 否则

若  $f(a) \cdot f(c) < 0$ , 令  $b = c$ , 否则  $a = c$ ;

(2)  $(b-a)/2 \leq \delta$ , 输出近似根  $c = (a+b)/2$ , 否则返回(1)。

**例 5.1.1** 利用二分法求方程  $f(x) = x^3 - x - 1 = 0$  在  $[1, 2]$  上的根。

**解**  $f(1) \cdot f(2) = -5 < 0$ , 因此在  $[1, 2]$  上方程有根, 表 5.1 给出了二分法前 9 次迭代的计算结果, 图 5.1 是函数  $f(x) = x^3 - x - 1$  曲线和最初 4 次迭代区间端点的图示。

表 5.1 二分法寻方程  $x^3 - x - 1 = 0$  的根: 前 9 次迭代

$n$	$a_n$	$b_n$	$c_n = (a_n + b_n)/2$	$f(c_n)$
0	1.0	2.0	1.5	0.125
1	1.0	1.5	1.25	-0.609375
2	1.25	1.5	1.375	-0.291016
3	1.375	1.5	1.4375	-0.0959473
4	1.4375	1.5	1.46875	0.0112
5	1.4375	1.46875	1.45313	-0.0431938
6	1.45313	1.46875	1.46094	-0.0162034
7	1.46094	1.46875	1.46484	-0.00255352
8	1.46484	1.46875	1.4668	0.00431024
9	1.46484	1.4668	1.46582	0.00087512

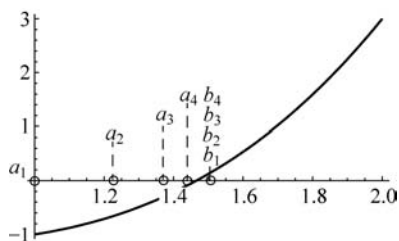


图 5.1 曲线  $f(x) = x^3 - x - 1$  和二分法前 4 次迭代的区间端点

二分法第  $n$  次迭代的近似根  $c_n = (a_n + b_n)/2$ , 与真值之间的误差为

$$|c_n - \xi| \leq (b - a)/2^{n+1} \quad (5.2)$$

迭代过程中只需计算  $(n+1)$  次函数值, 因而计算量极小。每次迭代都以常数因子  $1/2$  减小误差, 因此二分法是线性收敛的, 或者说二分法线性收敛。

如果取控制误差  $\delta = 0.5 \times 10^{-p}$ , 即期望根精确到小数点后  $p$  位, 则从  $(b-a)/2^{n+1} < 0.5 \times 10^{-p}$  估算出所需要的迭代次数

$$n > p \log_2 10 + \log_2 (b - a) \approx 3.3219p + \log_2 (b - a) \quad (5.3)$$

实际迭代次数要小于这个估计数。例如: 在例 5.1.1 中, 如果期望根精确到小数点后 3 位, 则迭代次数  $n > 3.3219p + \log_2 1 = 9.99657$ , 即迭代 10 次后可以使根精确到小数点后 3 位。

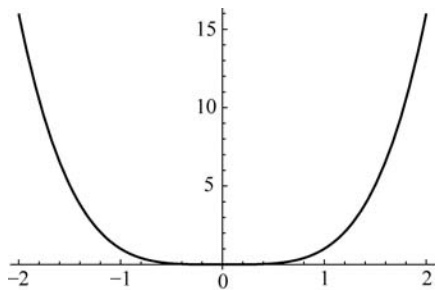


图 5.2 二分法失效

事实上, 这个方程的根是 1.46557 (精确到小数点后 5 位), 从表 5.1 可以看出第 9 次迭代根就已经精确到小数点后 3 位。

二分法的困难部分也许是确定端点异号的初始区间, 一旦这种区间被找到二分法就保证迭代一定收敛到根。值得指出的是, 二分法不适合重根的情况, 对图像如  $y = x^{2m}$  的一类问题失效, 如图 5.2 所示。下面介绍的牛顿法可解决这类问题, 收敛速度也比二分法快。

### 5.1.2 牛顿法

牛顿法, 也称牛顿-拉夫逊 (Newton-Raphson) 方法, 它的收敛速度比二分法快, 求根的迭代过程有明显的几何意义, 其原理是用一系列线性方程的根逼近非线性方程的根。

为求方程  $f(x) = 0$  的根, 先给定一个初始猜测值 (简称初始值)  $x_0$ , 然后过点  $(x_0, f(x_0))$  作曲线  $y = f(x)$  的切线

$$y = f(x_0) + f'(x_0)(x - x_0)$$

它与  $x$ -轴的交点  $x_1$ , 如图 5.3 所示, 是线性方程  $f(x_0) + f'(x_0)(x - x_0) = 0$  的解

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

再以  $x_1$  代替  $x_0$  重复上面的操作, 得到

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

将这个过程继续下去, 就得到牛顿迭代点列

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

可见, 它是一系列线性方程  $f(x_k) + f'(x_k)(x - x_k) = 0$ ,  $k = 0, 1, 2, \dots$  的解。

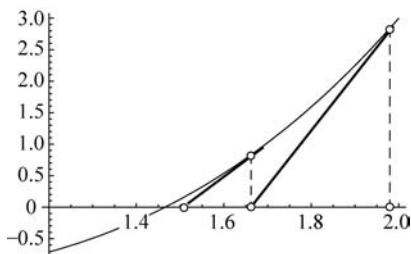


图 5.3 牛顿迭代几何

#### 牛顿法

解方程  $f(x) = 0$

给定初始值:  $x_0$

$$\text{迭代: } x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, k=0, 1, 2, \dots$$

由 Taylor 公式也可导出牛顿法。由 Taylor 定理, 在  $x$  和  $x_0$  之间存在  $\eta_x$  使得

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

当  $x$  在  $x_0$  附近时, 二次项很小可忽略不计, 即

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

当  $x_0$  在方程  $f(x)=0$  的根附近时, 自然希望下次迭代  $x_1$  更靠近方程的根, 因此希望  $x_1$  满足

$$f(x_0) + f'(x_0)(x_1 - x_0) = 0 \quad (5.4)$$

于是

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

满足式(5.4)的  $x_1$  只是比  $x_0$  更接近于根  $\xi$ , 但未必有  $f(x_1)=0$ 。为了找到更精确的解, 需要继续这个过程, 因此就产生了逐渐趋近于根  $\xi$  的牛顿迭代,  $x_{k+1} = x_k - f(x_k)/f'(x_k)$ ,  $k=0, 1, 2, \dots$ 。

和二分法一样, 牛顿法也是通过迭代逼近方程的根, 不同的是牛顿法将导数用在迭代过程中, 从而充分利用了函数的局部变化信息(导数是函数的瞬时变化率), 因此收敛速度比二分法更快。很可惜, 牛顿迭代不总是收敛的。

**例 5.1.2** 从给定的初始值开始对下述方程作 8 次牛顿迭代, 并观察收敛性:

(i)  $x^2 - 2 = 0, x_0 = 2.5$

(ii)  $\sin x = 0, x_0 = 0.8$

(iii)  $x^3 - x - 3 = 0, x_0 = 0.0$

(iv)  $x^{1/3} = 0, x_0 = 1.0$

**解** 这四个方程的牛顿迭代分别为

(i)  $x_{k+1} = x_k - \frac{2x_k^2 - 1}{2x_k}, x_0 = 2.5$

(ii)  $x_{k+1} = x_k - \frac{\sin x_k}{\cos x_k}, x_0 = 0.8$

(iii)  $x_{k+1} = x_k - \frac{x_k^3 - x_k - 3}{3x_k^2 - 1}, x_0 = 0.0$

(iv)  $x_{k+1} = x_k - \frac{x_k^{1/3}}{x_k^{-2/3}/3}, x_0 = 1.0$

表 5.2 给出了 8 次迭代结果, 可以看出: 对于  $x^2 - 2 = 0$  和  $\sin x = 0$  牛顿法收敛, 且第 4 次迭代就分别给出精确到小数点后 8 位和 24 位的根。对于方程  $x^3 - x - 3 = 0$ , 从初始值  $x_0 = 0.0$  开始, 迭代 3 次后进入循环状态, 因此对给定的初始值牛顿迭代不收敛。对于  $x^{1/3} = 0$ , 牛顿迭代对给定的初始值是发散的。

表 5.2 牛顿迭代

$x^2 - 2 = 0$	$\sin x = 0$	$x^3 - x - 3 = 0$	$x^{1/3} = 0$
2.5	0.8	0.00000	1.00000
1.65	-0.2296385570503640	-3.00000	-2.00000
1.431060606060606	0.00412357916974798	-1.96153	4.00000
1.414312727593564	-0.000000233724753	-0.00657	16.0000
1.414213565849603	$3.308722450212 \times 10^{-24}$	-3.00038	-32.0000
1.414213562373095	0.	-1.96181	64.0000
1.414213562373095	0.	-1.14743	-128.000
1.414213562373095	0.	-0.00725	256.000
1.414213562373095	0.	-3.00047	-512.000

方程  $x^3 - x - 3 = 0$  精确到小数点后 10 位的根是  $\xi = 1.6716998816$ , 选择充分靠近根  $\xi$  的点作为初始值, 比如  $x_0 = 1.0$ , 你会发现牛顿迭代收敛到根  $\xi$ , 即从初始值  $x_0 = 1.0$  开始牛顿法收敛。这说明牛顿法的收敛性依赖于初始值的选择, 有时在充分靠近根的情况下才收敛, 也就是说牛顿法局部收敛。最特别的是方程  $x^{1/3} = 0$ , 无论如何选取初始值, 无论多么靠近方程的根(只要不等于根  $\xi$ ), 经过若干次迭代后将逐渐偏离方程的根, 也就是说牛顿法对这类方程失效。

### 1. 二阶收敛性

牛顿法何时收敛, 收敛速度如何? 这是下面要讨论的问题, 先给出收敛速度的定义。

**定义 5.1.2** 假定  $\xi$  是方程  $f(x) = 0$  的一个根,  $\{x_k : k = 1, 2, \dots\}$  是收敛到  $\xi$  的迭代点列。如果存在常数  $c$  使得

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^p} = c \quad (5.5)$$

则称迭代最终收敛速度是  $p$  阶的, 或者说迭代  $p$  阶收敛。特别地,  $p = 1$ , 称为线性收敛;  $p = 2$ , 称为平方收敛;  $1 < p < 2$ , 称为超线性收敛。  $p$  越大收敛速度就越快。

关于牛顿迭代的收敛性与收敛速度有下述结论:

**定理 5.1.1** 假定  $f(x)$  有连续的二阶导数。若  $x_0$  充分接近方程  $f(x) = 0$  的根  $\xi$ , 且  $f'(\xi) \neq 0$ , 则牛顿迭代二阶收敛于  $\xi$ 。

**证明** 记  $c_\xi = \left| \frac{f''(\xi)}{2f'(\xi)} \right|$  ( $f'(\xi) \neq 0$ ), 由  $f''(x)$  的连续性, 当  $c > c_\xi$  时必存在  $\delta$  使得,

$$\forall x, \eta \in (\xi - \delta, \xi + \delta), \quad \left| \frac{f''(\eta)}{2f'(x)} \right| < c \quad (5.6)$$

根据 Taylor 定理,  $f(x_k) + f'(x_k)(\xi - x_k) + f''(\eta_k)(\xi - x_k)^2/2 = f(\xi) = 0$ , 其中  $\eta_k$  位于  $x_k$  和  $\xi$  之间, 于是

$$\xi = x_k - \underbrace{\frac{f(x_k)}{f'(x_k)}}_{x_{k+1}} - \frac{f''(\eta_k)}{2f'(x_k)}(\xi - x_k)^2$$

即

$$x_{k+1} - \xi = \frac{f''(\eta_k)}{2f'(x_k)}(x_k - \xi)^2 \quad (5.7)$$

因  $x_0$  充分接近  $\xi$ , 不妨假定  $|x_0 - \xi| < \min\{\delta, 1/c\}$ 。由式(5.6)和式(5.7),

$$|x_{k+1} - \xi| = \left| \frac{f''(\eta_k)}{2f'(x_k)} (x_k - \xi)^2 \right| < c |x_k - \xi|^2$$

于是

$$|x_{k+1} - \xi| < (c |x_0 - \xi|)^{k+1} |x_0 - \xi|$$

由于  $c |x_0 - \xi| < 1$ , 必有  $x_{k+1} - \xi \rightarrow 0 (k \rightarrow \infty)$ , 故牛顿迭代收敛。

因  $\eta_k$  在  $x_k, \xi$  之间, 所以  $\eta_k$  也收敛到  $\xi$ 。由式(5.7),

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^2} = \lim_{k \rightarrow \infty} \left| \frac{f''(\eta_k)}{2f'(x_k)} \right| = c_\xi$$

按定义 5.1.2, 牛顿法最终收敛速度是二阶的。证毕。

应该指出的是, 定理中的条件“ $x_0$  充分接近  $\xi$ ”一般不可检验, 因为事先不知道根  $\xi$  的位置, 我们正想找到根的位置。另外, “充分接近”也很模糊, 根据定理的证明过程可认为  $x_0$  在  $\xi$  的某个邻域, 比如  $|x_0 - \xi| < \min\{\delta, 1/c\}$ , 当然这也不可检验。最后, 定理只是说最终收敛速度是二阶的, 未指出达到这个速度之前需要多少次迭代。因此, 人们常说牛顿法的收敛性和收敛速度依赖于给定的初始值。尽管如此, 定理 5.1.1 仍有十分重要的理论意义, 因为它给出了牛顿迭代局部收敛的充分条件和最终收敛速度。

继续考虑例 5.1.2 中的方程  $x^2 - 2 = 0$ , 它的精确解

$$\xi = \sqrt{2}, \quad c_\xi = \left| \frac{f''(\sqrt{2})}{2f'(\sqrt{2})} \right| \approx 0.353553$$

记  $e_k \triangleq x_k - \sqrt{2}$ , 它是第  $k$  次迭代误差, 从表 5.2 得到最初几次的迭代误差

$$\begin{aligned} e_1 &= 0.235786 \\ e_2 &= 0.016847 & e_2/e_1^2 &= 0.30303 \\ e_3 &= 0.000099 & e_3/e_2^2 &= 0.34939 \\ e_4 &= 3.47651 \times 10^{-9} & e_4/e_3^2 &= 0.35352 \approx c_\xi \end{aligned}$$

注意, 对于这个方程, 初始取  $x_0 = 0$ , 牛顿法失败; 取  $x_0 > 0$ , 收敛到  $\sqrt{2}$ ; 取  $x_0 < 0$ , 收敛到  $-\sqrt{2}$ 。

牛顿法不像二分法那样, 可以通过估计绝对误差来终止迭代, 在实际问题中需要设置终止条件。给定容差  $\epsilon$ , 常用的终止条件有

绝对终止条件:  $|x_{k+1} - x_k| \leq \epsilon$ 。

相对终止条件:  $\frac{|x_{k+1} - x_k|}{|x_{k+1}|} \leq \epsilon$  (根不在 0 附近)。

绝对/相对终止条件:  $\frac{|x_{k+1} - x_k|}{\max\{|x_{k+1}|, \theta\}} \leq \epsilon$ , 其中  $\theta > 0$ , 常用于根在 0 附近。

## 2. 线性收敛性

在定理 5.1.1 中,  $f'(\xi) \neq 0$  是牛顿法收敛的充分条件, 但不是必要条件。例如方程  $x^p = 0$ , 零是它的  $p$  重根, 即  $f(0) = f'(0) = \dots = f^{(p-1)}(0) = 0, f^{(p)}(0) = p \neq 0$ , 牛顿迭代

$$x_{k+1} = x_k - \frac{x_k^p}{p x_k^{p-1}} \left( = \frac{p-1}{p} x_k \right)$$

以 $(p-1)/p$ 的因子减小,因此收敛到方程的根。从 $(x_{k+1}-0)/(x_k-0)=(p-1)/p$ 知,此时牛顿法的收敛速度是线性的,不再有二阶收敛性。下述定理概括了这类问题。

**定理 5.1.2** 假定  $f(x)$  有连续的  $p+1$  阶导数,  $\xi$  是  $f(x)=0$  的  $p+1$  重根, 即

$$f(\xi) = f'(\xi) = \cdots = f^{(p)}(\xi) = 0, \quad f^{(p+1)}(\xi) \neq 0$$

若  $x_0$  充分接近  $\xi$ , 则牛顿迭代收敛于  $\xi$ , 且

$$\lim_{k \rightarrow \infty} \left| \frac{x_{k+1} - \xi}{x_k - \xi} \right| = \frac{p}{p+1} \quad (5.8)$$

即牛顿迭代的收敛速度是线性的。

**证明** 在牛顿迭代  $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$  的两边同减去  $\xi$ , 得到

$$x_{k+1} - \xi = x_k - \xi - \frac{f(x_k)}{f'(x_k)} \quad (5.9)$$

由重根条件, 对  $f(x_k)$  和  $f'(x_k)$  在点  $\xi$  作 Taylor 展开

$$f(x_k) = \frac{f^{(p+1)}(\eta_k)}{(p+1)!} (x_k - \xi)^{(p+1)}, \quad f'(x_k) = \frac{f^{(p+1)}(\eta'_k)}{p!} (x_k - \xi)^p$$

其中  $\eta_k, \eta'_k$  位于  $\xi, x_k$  之间。将上式代入式(5.9), 得到

$$x_{k+1} - \xi = (x_k - \xi) \left( 1 - \frac{1}{p+1} \cdot \frac{f^{(p+1)}(\eta_k)}{f^{(p+1)}(\eta'_k)} \right) \quad (5.10)$$

由于  $f(x)$  是连续  $p+1$  阶可导的, 且  $f^{(p+1)}(\xi) \neq 0$ , 所以取  $0 < \epsilon < 1/(p+1)$ , 一定存在  $\delta > 0$  使得当  $\eta, \eta' \in (\xi - \delta, \xi + \delta)$  时,

$$\left| 1 - \frac{1}{p+1} \cdot \frac{f^{(p+1)}(\eta)}{f^{(p+1)}(\eta')} \right| < 1 - \epsilon \quad (5.11)$$

其中  $0 < 1 - \epsilon < 1$ 。因  $x_0$  充分接近  $\xi$ , 不妨假定  $x_0 \in (\xi - \delta, \xi + \delta)$ , 根据式(5.10)和式(5.11)得到

$$|x_1 - \xi| = \left| (x_0 - \xi) \left( 1 - \frac{1}{p+1} \cdot \frac{f^{(p+1)}(\eta_k)}{f^{(p+1)}(\eta'_k)} \right) \right| < (1 - \epsilon) |x_0 - \xi| < \delta$$

从而  $x_1 \in (\xi - \delta, \xi + \delta)$ , 同理可推知所有  $x_k \in (\xi - \delta, \xi + \delta)$ , 因此

$$|x_{k+1} - \xi| < (1 - \epsilon) |x_k - \xi| < \cdots < (1 - \epsilon)^{k+1} |x_0 - \xi|$$

因  $(1 - \epsilon)^{k+1} \rightarrow 0 (k \rightarrow \infty)$ , 必有  $x_{k+1} - \xi \rightarrow 0$ 。再由式(5.10),

$$\lim_{k \rightarrow \infty} \left| \frac{x_{k+1} - \xi}{x_k - \xi} \right| = \lim_{k \rightarrow \infty} \left| 1 - \frac{1}{p+1} \cdot \frac{f^{(p+1)}(\eta_k)}{f^{(p+1)}(\eta'_k)} \right| = \frac{p}{p+1}$$

证毕。

**注释** 对于  $p+1$  重根牛顿法线性收敛, 如果利用下面的迭代

$$x_{k+1} = x_k - \frac{(p+1)f(x_k)}{f'(x_k)}, \quad k = 1, 2, \dots$$

则恢复到牛顿法的二阶收敛性。因事先难以确定根的重数, 这种迭代方式仅有理论意义而无实用价值, 对此不作详细讨论。

### 5.1.3 拟牛顿法

牛顿法对单根二阶收敛, 对重根线性收敛, 但在迭代过程中必须计算导数。实际问题的

函数可能非常复杂,或者根本没有解析表达式,此时牛顿法需要花费很大代价来计算导数,自然希望有不需计算导数的迭代方法来解决此类问题。拟牛顿法是避免计算导数的一类方法,它的一般迭代形式为

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

这里  $T_k$  是  $f'(x_k)$  的近似  $T_k \approx f'(x_k)$ 。选择不同的  $T_k$  导致不同的拟牛顿法,其中最著名的是下面将要介绍的割线法,其收敛速度介于线性与二阶之间。

### 1. 割线法

众所周知,导数是切线的斜率,而切线在几何上是割线的极限,因此导数可用割线的斜率来近似。通过  $(x_{k-1}, f(x_{k-1}))$  和  $(x_k, f(x_k))$  两点的割线斜率

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

替代牛顿法中  $f'(x_k)$ , 得到割线法的迭代

$$x_{k+1} = x_k - \frac{f(x_k)}{f[x_{k-1}, x_k]} \quad (5.12)$$

事实上,  $x_{k+1}$  是割线与  $x$ -轴的交点,如图 5.4 所示。与牛顿法不同,实施割线法需要两个初始值。

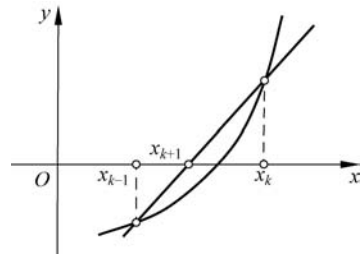


图 5.4 割线法迭代几何

#### 割线法

解方程  $f(x) = 0$

给定初始值:  $x_0, x_1$

迭代:  $x_{k+1} = x_k - \frac{f(x_k)}{f[x_{k-1}, x_k]}, k = 1, 2, \dots$

在讨论割线法的收敛性和收敛速度之前,先看一个例子。继续考察方程  $x^2 - 2 = 0$ , 此次使用割线法,初始值取  $x_0 = 2.5, x_1 = 1.5$ 。按迭代公式(5.12),最初 6 次迭代的近似值及其误差如下:

$$\begin{aligned} x_2 &= 1.4375 & e_2 &= 0.0232864 \\ x_3 &= 1.41489361 & e_3 &= 0.0006800 \\ x_4 &= 1.41429114 & e_4 &= 5.551 \times 10^{-6} \\ x_5 &= 1.41421356 & e_5 &= 1.334 \times 10^{-9} \\ x_6 &= 1.41421356 & e_6 &= 2.664 \times 10^{-15} \end{aligned}$$

观察误差可以发现,割线法的收敛速度快于线性收敛速度,但比二阶收敛速度慢。根据下面的定理,准确收敛阶  $p = (1 + \sqrt{5})/2 \approx 1.618$ , 位于线性与二阶中间偏上一些。

**定理 5.1.3** 假定  $f(x)$  有连续的二阶导数,  $\xi$  是方程  $f(x) = 0$  的根,  $f'(\xi) \neq 0$ 。若  $x_0, x_1$  充分接近  $\xi$ , 则割线法收敛到  $\xi$ , 且

$$\lim_{k \rightarrow \infty} \left| \frac{e_{k+1}}{e_k e_{k-1}} \right| = c_\xi \quad \left( c_\xi \triangleq \left| \frac{f''(\xi)}{2f'(\xi)} \right| \right) \quad (5.13)$$

**证明** 由式(5.12), 得到

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

因此

$$\begin{aligned} e_{k+1} &= e_k - \frac{f(x_k)e_k - f(x_k)e_{k-1}}{f(x_k) - f(x_{k-1})} = e_k e_{k-1} \left( \frac{1}{e_{k-1}} - \frac{\frac{f(x_k)}{e_{k-1}} - \frac{f(x_k)}{e_k}}{f(x_k) - f(x_{k-1})} \right) \\ &= e_k e_{k-1} \cdot \underbrace{\frac{f(x_k)/(x_k - \xi) - f(x_{k-1})/(x_{k-1} - \xi)}{x_k - x_{k-1}}}_{A(x_{k-1}, x_k)} \cdot \underbrace{\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}}_{B(x_{k-1}, x_k)} \end{aligned} \quad (5.14)$$

当  $x_k, x_{k-1} \rightarrow \xi$  时

$$B(x_k, x_{k-1}) \rightarrow \frac{1}{f'(\xi)} \quad (5.15)$$

考虑  $A(x_k, x_{k-1})$  在  $x_k, x_{k-1} \rightarrow \xi$  时的极限: 定义  $h(x) = \frac{f(x)}{x - \xi}$ , 则

$$A(x_k, x_{k-1}) = \frac{h(x_k) - h(x_{k-1})}{x_k - x_{k-1}}$$

当  $x_k, x_{k-1} \rightarrow \xi$  时,  $A(x_k, x_{k-1}) \rightarrow h'(\xi)$ 。下面计算  $h'(\xi)$ :

$$h'(x) = \frac{f'(x)(x - \xi) - f(x)}{(x - \xi)^2}$$

根据求极限的洛必达法则

$$h'(\xi) = \lim_{x \rightarrow \xi} \frac{f'(x)(x - \xi) - f(x)}{(x - \xi)^2} = \lim_{x \rightarrow \xi} \frac{f''(x)(x - \xi)}{2(x - \xi)} = \frac{f''(\xi)}{2}$$

因此

$$A(x_k, x_{k-1}) \rightarrow \frac{f''(\xi)}{2} \quad (5.16)$$

由式(5.15)和式(5.16),

$$|A(x_k, x_{k-1}) \cdot B(x_k, x_{k-1})| \rightarrow \left| \frac{f''(\xi)}{2f'(\xi)} \right| = c_\xi \quad (5.17)$$

于是, 给定  $\epsilon > 0$  必存在  $\delta > 0$  使得当  $x_k, x_{k-1} \in (\xi - \delta, \xi + \delta)$  时,

$$\begin{aligned} |A(x_k, x_{k-1}) \cdot B(x_k, x_{k-1})| &< c_\xi + \epsilon \\ |e_{k+1}| &< (c_\xi + \epsilon) |e_k e_{k-1}| \end{aligned}$$

由于  $x_0, x_1$  充分接近  $\xi$ , 不妨假定  $x_0, x_1 \in (\xi - \delta', \xi + \delta')$ , 其中  $\delta' = \min \left\{ \delta, \frac{1}{2(c_\xi + \epsilon)} \right\}$ ,

于是

$$|e_2| < (c_\xi + \epsilon) |e_1 e_0| < \frac{1}{2} |e_0| < \delta'$$

一般,

$$|e_{k+1}| < \frac{1}{2} |e_{k-1}| < \delta'$$

因此, 所有  $x_k$  都落在  $(\xi - \delta', \xi + \delta')$  内, 且  $|e_{k+1}| < \frac{1}{2^k} |e_0| \rightarrow 0$ , 故割线法收敛到  $\xi$ , 从而由

式(5.14)和式(5.17)得到

$$\lim_{k \rightarrow \infty} \left| \frac{e_{k+1}}{e_k e_{k-1}} \right| = c_\xi$$

证毕。

根据定理 5.1.3, 可推知割线法的收敛阶  $p = \frac{1+\sqrt{5}}{2} \approx 1.618$ ; 由式(5.13), 对充分大的  $k$

$$|e_{k+1}| \approx c_\xi |e_k| \cdot |e_{k-1}| \quad (5.18)$$

假定存在常数  $c, p$  使得对充分大的  $k, |e_k| \approx c |e_{k-1}|^p$ , 即

$$|e_{k-1}| \approx \left( \frac{|e_k|}{c} \right)^{1/p}$$

将上式和  $|e_{k+1}| \approx c |e_k|^p$  代入式(5.18), 得到

$$c |e_k|^p \approx c_\xi |e_k| \cdot \left( \frac{|e_k|}{c} \right)^{1/p} = c_\xi c^{-1/p} |e_k|^{1+1/p}$$

$$c_\xi^{-1} c^{1+1/p} \approx |e_k|^{1-p+1/p}$$

上式左端是无关  $k$  的常数, 因此右端指数必须为零, 即

$$1 - p + \frac{1}{p} = 0$$

于是  $p = (1 \pm \sqrt{5})/2$ 。阶是正值, 故割线法的收敛阶  $p = (1 + \sqrt{5})/2 \approx 1.618$ 。

## 2. 抛物线法

不难看出, 割线法是用一系列线性插值函数(割线)的零点逼近方程的根, 它的直接推广是以二次插值函数的零点逼近方程的根, 如图 5.5 所示, 由此产生解方程的抛物线法, 也称为密勒(Müller)法。

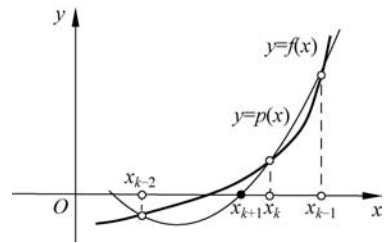


图 5.5 抛物线法迭代几何

### 抛物线法

解方程  $f(x) = 0$

给定初始值:  $x_0, x_1, x_2$

迭代:

$$x_{k+1} = x_k - \frac{2f(x_k)}{b(x_k, x_{k-1}, x_{k-2}) + \text{sign}(b) \sqrt{b^2(x_k, x_{k-1}, x_{k-2}) - 4f(x_k)f[x_k, x_{k-1}, x_{k-2}]}} \quad k=2, 3, \dots$$

假定  $x_k, x_{k-1}, x_{k-2}$  是当前迭代的三个近似根, 以它们为节点的二次多项式插值

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

有两个零点

$$x_\pm = x_k - \frac{2f(x_k)}{b(x_k, x_{k-1}, x_{k-2}) \pm \sqrt{b^2(x_k, x_{k-1}, x_{k-2}) - 4f(x_k)f[x_k, x_{k-1}, x_{k-2}]}}$$

(5.19)

其中

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

为了确定第  $k+1$  次迭代  $x_{k+1}$ , 需要选择根式前的正负号。在三个近似根  $x_k, x_{k-1}, x_{k-2}$  中, 自然假定  $x_k$  更接近方程的精确解, 因而选择最接近于  $x_k$  的零点作为  $x_{k+1}$ , 此时根式前符号与  $b$  的符号相同。这样就得到了抛物线法的第  $k+1$  次迭代。

**例 5.1.3** 用割线法和抛物线法求方程  $xe^x = 1$  的前 6 次迭代的近似根, 其中割线法的两个初始值  $x_0 = 2, x_1 = 1.5$ ; 抛物线法初始值  $x_2 = 1.144809$  是割线法的第一次迭代。

**解** 计算结果见表 5.3, 可以看出割线法第 6 次迭代才达到小数点后 4 位的精度; 抛物线法以  $x_0, x_1$  和  $x_2 = 1.144809$  为初始值时, 首次迭代就达到了小数点后 14 位的精度, 抛物线法比割线法收敛更快。

表 5.3 割线法和抛物线法解方程  $xe^x = 1$

迭代次数	割线法	抛物线法
1	1.144809274721835	
2	0.849721510613437	0.567143290409784
3	0.668652203745926	0.567143290409783
4	0.587755224926128	0.567143290409783
5	0.568784819013826	0.567143290409783
6	0.567170784669773	0.567143290409783

事实上, 在一定的条件下, 可以证明抛物线法的迭代误差有如下渐近关系:

$$\lim_{k \rightarrow \infty} \frac{|e_{k+1}|}{|e_k|^{1.840}} = \left| \frac{f'''(\xi)}{2f'(\xi)} \right|^{0.42} \quad (5.20)$$

也就是说, 抛物线法的收敛阶  $p = 1.840$  ( $\alpha^3 - \alpha^2 - \alpha - 1 = 0$  的根), 因此也是超线性收敛的, 与割线法收敛阶  $p = 1.618$  相比, 抛物线法收敛更快, 几乎接近牛顿法的二阶收敛性。

**注释** 抛物线法在迭代过程中, 二次插值多项式可能与  $x$  轴不相交, 即得到的迭代  $x_{k+1}$  可能是复数, 此时可得到方程的近似复根。也就是说, 抛物线法也适用于求多项式的复根。若问题只关心实根, 可以应用逆二次插值法使每次迭代  $x_{k+1}$  都是实数。所谓逆二次插值法是指, 用  $y$  的二次多项式  $x = ay^2 + by + c$  而不是用  $x$  的二次多项式, 进行插值, 如图 5.6 所示。它与  $x$ -轴有唯一交点  $x = c(y = 0)$ , 因此  $x_{k+1} = c$ , 读者可自行写出逆插值二次法的迭代公式。

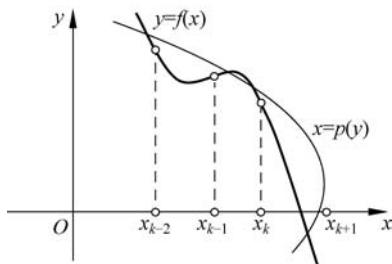


图 5.6 逆二次插值法迭代几何

### 5.1.4 不动点法

**定义 5.1.3** 如果  $\xi$  使得  $\phi(\xi) = \xi$ , 则称  $\xi$  为函数  $y = \phi(x)$  的不动点。在几何上, 不动点是曲线  $y = \phi(x)$  与直线  $y = x$  交点的  $x$ -坐标。

不动点常用来分析动力系统的性态。动力系统的行为在不动点处达到平衡状态, 既不增长也不衰减。非线性方程  $f(x) = 0$  的求解也可归结为求一个函数的不动点, 比如: 令

$$\phi(x) = x - \frac{f(x)}{f'(x)} \quad (f'(x) \neq 0)$$

定义迭代  $x_{k+1} = \phi(x_k), k=0, 1, 2, \dots$ 。如果迭代收敛于  $\xi$ , 则它是  $\phi(x)$  的不动点  $\xi = \phi(\xi)$ , 同时  $\xi$  也是方程  $f(x) = 0$  的根, 因为从  $\xi = \phi(\xi)$  得到

$$\xi = \xi - \frac{f(\xi)}{f'(\xi)} \quad (f'(\xi) \neq 0)$$

故  $f(\xi) = 0$ 。事实上, 此处  $x_{k+1} = \phi(x_k)$  就是牛顿迭代。

一般, 将方程  $f(x) = 0$  化为同解方程  $x = \phi(x)$  后, 就可以从给定初值开始迭代

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

称  $\phi$  为迭代函数。若迭代收敛于  $\xi$  且迭代函数连续, 则  $\xi$  是迭代函数的不动点, 同时也是方程  $f(x) = 0$  的根。这就是解方程的不动点法。

### 不动点法

解方程  $f(x) = 0$

构造同解方程:  $x = \phi(x)$

给定初始值:  $x_0$

迭代:  $x_{k+1} = \phi(x_k), k = 0, 1, 2, \dots$

同解方程  $\phi(x) = x$  有多种多样的形式, 不同形式会有不同的性态, 因此相应迭代的收敛性也不相同。

**例 5.1.4** 考虑方程  $x^3 + x - 1 = 0$  的四种同解方程:

(i)  $x = 1 - x^3$

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

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

(iv)  $x = \frac{1 + 2x^3}{1 + 3x^2}$

其中(iv)是在(i)两边同时加  $3x^3$  得到的。它们相应的迭代函数分别为

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

$$\phi_2(x) = (1 - x)^{1/3}$$

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

$$\phi_4(x) = \frac{1 + 2x^3}{1 + 3x^2}$$

从初始值  $x_0 = 0.75$  开始迭代, 表 5.4 给出了最初 16 次的迭代结果。方程  $x^3 + x - 1 = 0$  的精确解

$$\xi = \frac{\sqrt[3]{2(9 + \sqrt{93})^2} - 2\sqrt[3]{3}}{\sqrt[3]{36} \sqrt[3]{9 + \sqrt{93}}} \approx 0.682327803828019$$

表 5.4 不动点迭代

$x_{k+1} = \phi_1(x_k)$	$x_{k+1} = \phi_2(x_k)$	$x_{k+1} = \phi_3(x_k)$	$x_{k+1} = \phi_4(x_k)$
0.578125	0.629960524947436	0.64	0.686046511627907
0.806774139404296	0.717930965249187	0.709421112372304	0.682339582597314
0.474883207970300	0.655820725748032	0.665212815374884	0.682327803946512
0.892907159169156	0.700801309750398	0.693237013444374	0.682327803828019
0.288100127392519	0.668836392902901	0.675412102122535	0.682327803828019
0.976087204437633	0.691853594202033	0.686727578064001	0.682327803828019
0.070036595169878	0.675438329507026	0.679534898317339	0.682327803828019
0.999656461769720	0.687225200495086	0.684103236806338	0.682327803828019
0.001030260675835	0.678803258036743	0.681200191457860	0.682327803828019
0.999999998906443	0.684841984588895	0.683044385812104	0.682327803828019
$3.2806707350 \times 10^{-9}$	0.680522965212763	0.681872592311510	0.682327803828019
1.	0.683617568033420	0.682617045997094	0.682327803828019
0.	0.681403123516969	0.682144045988890	0.682327803828019
1.	0.682989202492196	0.682444557594172	0.682327803828019
0.	0.681853935624274	0.682253626695813	0.682327803828019
1.	0.682666908956669	0.682374932546297	0.682327803828019

从表 5.4 可以看出,  $\phi_1$  迭代若干次后进入周期状态, 循环取 0, 1 两个值, 因此不收敛;  $\phi_2$  收敛但收敛速度缓慢, 迭代到 16 次才达到小数点后 3 位的精度;  $\phi_3$  收敛速度稍快;  $\phi_4$  收敛速度非常快, 迭代 4 次就达到小数点后 15 位的精度。

同解方程为什么会表现出如此不同的性质? 为此, 分析迭代函数在不动点局部邻域  $\Delta = [0.5, 0.8]$  的性态。考虑迭代函数在  $\Delta$  上的最小、最大值, 以及导数在  $\Delta$  上的最大值:

$$\phi_1(\Delta) = [0.488, 0.875] \not\subset \Delta, \max_{x \in \Delta} |\phi_1'(x)| = 1.92 > 1$$

$$\phi_2(\Delta) = [0.5848, 0.7937] \subset \Delta, \max_{x \in \Delta} |\phi_2'(x)| = 0.9746 < 1$$

$$\phi_3(\Delta) = [0.6097, 0.8] \subset \Delta, \max_{x \in \Delta} |\phi_3'(x)| = 0.6495 < 1$$

$$\phi_4(\Delta) = [0.6823, 0.7142] \subset \Delta, \max_{x \in \Delta} |\phi_4'(x)| = 0.1756 < 1$$

可见, 收敛的迭代函数都有共同的质性  $\phi(\Delta) \subset \Delta, \max |\phi'(x)| < 1$ , 发散的迭代函数没有这个特性;  $\max |\phi'(x)|$  越小收敛越快。这一现象可由后面的定理得到进一步证实。

### 1. 整体收敛性

**定义 5.1.4** 如果存在常数  $\rho < 1$  使得

$$|\phi(x) - \phi(y)| \leq \rho |x - y|, \quad \forall x, y \in [a, b]$$

则称  $\phi(x)$  是  $[a, b]$  上压缩映射。

如果  $M \triangleq \max_{a \leq x \leq b} |\phi'(x)| < 1$ , 则  $\phi(x)$  必是  $[a, b]$  上压缩映射。压缩映射缩小任意两点间的距离, 因此在  $[a, b]$  上一致连续。

**定理 5.1.4 (整体收敛的充分条件)** 如果  $\phi(x)$  是  $[a, b]$  上的压缩映射, 且  $\phi([a, b]) \subset [a, b]$ , 则

(i) 在  $[a, b]$  上  $\phi(x)$  存在唯一的不动点  $\xi$

(ii) 对任意  $x_0 \in [a, b]$  为初始值的迭代  $x_{k+1} = \phi(x_k) \in [a, b]$ , 且

$$|x_{k+1} - \xi| \leq \frac{\rho^{k+1}}{1-\rho} |x_1 - x_0| \quad (5.21)$$

(iii)  $x_k \rightarrow \xi (k \rightarrow \infty)$

**证明** 令  $\Phi(x) = x - \phi(x)$ 。因  $\phi([a, b]) \subset [a, b]$ , 即  $\forall x \in [a, b], a \leq \phi(x) \leq b$ , 所以  $\Phi(a) \leq 0, \Phi(b) \geq 0$ 。由压缩映射的连续性和介值定理, 在  $[a, b]$  上存在  $\xi$  使  $\Phi(\xi) = 0$ , 即  $\xi$  是  $\phi(x)$  的不动点  $\xi$ 。如果还存在一个不动点  $\eta \neq \xi$ , 则

$$|\xi - \eta| = |\phi(\xi) - \phi(\eta)| \leq \rho |\xi - \eta| \Rightarrow \rho \geq 1$$

与  $\rho < 1$  矛盾, 故不动点是唯一的。

$$|x_{k+1} - \xi| = |\phi(x_k) - \phi(\xi)| \leq \rho |x_k - \xi| \leq \rho (|x_{k+1} - x_k| + |x_{k+1} - \xi|)$$

由此可推知,

$$|x_{k+1} - \xi| \leq \frac{\rho}{1-\rho} |x_{k+1} - x_k| \quad (5.22)$$

此外

$$|x_{k+1} - x_k| = |\phi(x_k) - \phi(x_{k-1})| \leq \rho |x_k - x_{k-1}| \Rightarrow |x_{k+1} - x_k| \leq \rho^k |x_1 - x_0|$$

代入式(5.22),

$$|x_{k+1} - \xi| \leq \frac{\rho^{k+1}}{1-\rho} |x_1 - x_0|$$

最后, 因  $\rho < 1$ , 得到  $\lim_{k \rightarrow \infty} |x_{k+1} - \xi| = 0$ 。证毕。

称式(5.21)和式(5.22)为迭代误差的事先估计和事后估计。事先估计是迭代开始时对第  $k+1$  迭代误差的一个粗略估计, 以便为终止迭代提供信息; 事后估计是当前迭代的误差估计, 比事先估计准确。给定精度  $\varepsilon$ , 由事先估计

$$\frac{\rho^{k+1}}{1-\rho} |x_1 - x_0| = \varepsilon \Rightarrow k+1 = \rho^{-1} \ln \left( \frac{1-\rho}{|x_1 - x_0|} \right)$$

得到需要迭代的次数。当然, 这必须以已知  $\rho$  为前提, 在  $\rho$  未知情况下仍使用 5.1.2 节介绍的终止条件。

**定理 5.1.5 (整体发散的充分条件)** 假定存在常数  $\rho > 1$  使得  $|\phi(x) - \phi(y)| \geq \rho |x - y|$  ( $\forall x, y \in [a, b]$ ), 或者  $\min_{a \leq x \leq b} |\phi'(x)| > 1$ 。若初始点  $x_0$  不是不动点, 则迭代发散。

证明略, 留给读者。

**例 5.1.5** 下述三个方程在  $[2, 3]$  上都与方程  $x^3 - 2x - 5 = 0$  同解, 试分析不动点迭代的收敛性:

(i)  $x = x^3 - x - 5$

(ii)  $x = (2x + 5)^{1/3}$

(iii)  $x = (2 + 5/x)^{1/2}$

**解**

(i)  $\phi(x) = x^3 - x - 5, \phi'(x) = 3x^2 - 1, \min_{2 \leq x \leq 3} |\phi'(x)| = |\phi'(2)| = 11 > 1$ , 由定理 5.1.5, 迭代发散。

(ii)  $\phi(x) = (2x + 5)^{1/3}, \phi'(x) = 2(2x + 5)^{-2/3}/3$

$$\min_{2 \leq x \leq 3} \phi(x) = \phi(2) = 9^{1/3} > 2, \max_{2 \leq x \leq 3} \phi(x) = \phi(3) = 11^{1/3} < 3$$

$$\Downarrow$$

$$\phi([2,3]) \subset [2,3]$$

$$\max_{2 \leq x \leq 3} |\phi'(x)| = \phi'(2) = 2 \cdot 9^{-2/3} / 3 \approx 0.15408 < 1$$

根据定理 5.1.4, 在区间  $[2,3]$  上迭代整体收敛。

$$(iii) \phi(x) = (2+5/x)^{1/2}, \phi'(x) = -5(2+5/x)^{-1/2} / 2x$$

$$\min_{2 \leq x \leq 3} \phi(x) = \phi(3) = (11/2)^{1/2} < 2, \max_{2 \leq x \leq 3} \phi(x) = \phi(2) = (9/2)^{1/2} < 3$$

$$\Downarrow$$

$$\phi([2,3]) \not\subset [2,3]$$

因此, 不满足定理 5.1.4 条件。如果将区间缩小到  $[2,2.5]$ , 就有

$$\phi([2,2.5]) \subset [2,2.5]$$

$$\max_{2 \leq x \leq 2.5} |\phi'(x)| = \phi'(2) = 5 \cdot 2^{-1/2} / 12 < 1$$

由定理 5.1.4, 迭代在区间  $[2,2.5]$  上整体收敛。事实上, 在区间  $[2,3]$  上迭代也是整体收敛的, 这与定理 5.1.4 不矛盾, 因它给出的仅是收敛性的充分条件。

## 2. 局部收敛性

如果初始值充分靠近不动点, 则迭代的收敛性仅依赖于迭代函数的局部性质, 对此有如下定理:

**定理 5.1.6 (局部收敛的充分条件)** 若  $\phi(x)$  在不动点  $\xi$  的某个邻域有连续的导数, 且  $|\phi'(\xi)| < 1$ , 则迭代局部收敛于  $\xi$ , 也就是说当  $x_0$  充分靠近不动点时, 迭代  $x_{k+1} = \phi(x_k)$  必收敛到  $\xi$ 。

**证明** 因  $\phi'(x)$  在包含  $\xi$  的邻域内连续且  $\phi'(\xi) < 1$ , 所以存在  $\delta > 0$  使得

$$\phi'(x) < 1, \quad \forall x \in [\xi - \delta, \xi + \delta]$$

记

$$\rho = \max\{|\phi'(x)| : x \in [\xi - \delta, \xi + \delta]\} < 1$$

由微分中值定理,  $\forall x, y \in [\xi - \delta, \xi + \delta]$ ,

$$\phi(x) - \phi(y) = \phi'(\eta)(x - y) \Rightarrow |\phi(x) - \phi(y)| \leq \rho |x - y|$$

因此

$$|\phi(x) - \xi| = |\phi(x) - \phi(\xi)| \leq \rho |x - \xi| < \delta \Rightarrow \phi([\xi - \delta, \xi + \delta]) \subset [\xi - \delta, \xi + \delta]$$

根据定理 5.1.4, 迭代  $x_{k+1} = \phi(x_k)$  在  $[\xi - \delta, \xi + \delta]$  上收敛到  $\xi$ 。证毕。

**定理 5.1.7 (局部发散的充分条件)** 假定  $\phi(x)$  在不动点  $\xi$  的某个邻域有连续的导数且  $|\phi'(\xi)| > 1$ , 则若  $x_1 \neq x_0$  (即  $x_0$  不是不动点), 则迭代局部发散。

**例 5.1.6** 不难验证: 迭代函数  $\phi_1(x) = x^2 - 1$ ,  $\phi_2(x) = 2x - (1 + \sqrt{5})/2$ ,  $\phi_3(x) = \sqrt{1+x}$  有同一个不动点

$$\xi = (1 + \sqrt{5})/2 \approx 1.61803$$

由于  $\phi_1'(\xi) = 1 + \sqrt{5} > 1$ ,  $\phi_2'(\xi) = 2 > 1$ ,  $|\phi_3'(\xi)| = 1/2 < 1$ , 根据定理 5.1.7, 对于  $\phi_1(x)$ ,  $\phi_2(x)$  迭代都是局部发散的; 根据定理 5.1.6 对于  $\phi_3(x)$  迭代收敛。两个发散迭代的性态不相同, 对于  $\phi_1(x)$  迭代最终进入周期状态; 而对于  $\phi_2(x)$  随着迭代次数的增加离不动点越来越远。表 5.5 给出了最初 15 次的迭代结果。

表 5.5 不动点局部迭代(初始值: 1.60)

$x_{k+1} = \phi_1(x_k)$	$x_{k+1} = \phi_2(x_k)$	$x_{k+1} = \phi_3(x_k)$
1.5600000000000000	1.581966011250105	1.612451549659710
1.4336000000000001	1.545898033750315	1.616307999627456
1.0552089600000004	1.473762078750736	1.617500540843018
0.113465949264290	1.329490168751577	1.617869135883066
-0.987125478357553	1.040946348753260	1.617983045610511
-0.025583289977371	0.463858708756626	1.618018246377497
-0.999345495273933	-0.690316571236641	1.618029124081979
-0.001308581075696	-2.998667131223177	1.618032485484138
-0.999998287615568	-7.615368251196250	1.618033524215162
-3.4247659311814 $\times 10^{-6}$	-16.84877049114239	1.618033845200761
-0.999999999988270	-35.31557497103468	1.618033944390772
-2.3458124331909 $\times 10^{-11}$	-72.24918393081927	1.618033975042172
-1.	-146.1164018503884	1.618033984513975
0.	-293.8508376895267	1.618033987440923
-1.	-589.3197093678034	1.618033988345400

最后,以不动点迭代的收敛阶定理结束本节。

**定理 5.1.8** 如果  $\phi(x)$  在不动点  $\xi$  的某个邻域有连续的  $p$  阶导数,且

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

则迭代  $p$  阶局部收敛于  $\xi$ 。

**证明** 由  $\phi'(\xi) = 0$ , 迭代必局部收敛于  $\xi$ 。对  $\phi(x)$  在点  $\xi$  应用 Taylor 定理:

$$x_k = \phi(x_{k-1}) = \phi(\xi) + \frac{\phi'(\xi)}{2}(x_{k-1} - \xi) + \cdots + \frac{\phi^{(p-1)}(\xi)}{(p-1)!}(x_{k-1} - \xi)^{p-1} + \frac{\phi^{(p)}(\eta_{k-1})}{p!}(x_{k-1} - \xi)^p$$

于是

$$x_k = \xi + \frac{\phi^{(p)}(\eta_{k-1})}{p!}(x_{k-1} - \xi)^p$$

因此

$$\frac{|x_k - \xi|}{|x_{k-1} - \xi|^p} = \left| \frac{\phi^{(p)}(\eta_{k-1})}{p!} \right| \rightarrow \left| \frac{\phi^{(p)}(\xi)}{p!} \right| \quad (k \rightarrow \infty)$$

故迭代  $p$  阶收敛。证毕。

## 5.2 非线性方程组

非线性方程组是指由  $n$  个  $n$  元非线性方程构成的方程组,一般形式为

$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0 \\ f_2(x_1, x_2, \cdots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \cdots, x_n) = 0 \end{cases} \quad (5.23)$$

记

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))^T, \quad \mathbf{x} = (x_1, x_2, \dots, x_n)^T$$

它是  $\mathbf{R}^n$  到  $\mathbf{R}^n$  的映射, 称为向量值函数。利用这个记号, 将非线性方程组写成简洁的向量形式

$$\mathbf{f}(\mathbf{x}) = 0 \quad (5.24)$$

求解非线性方程组远比求解一元非线性方程困难, 本节的主要目的是将上节中的牛顿法、拟牛顿法和不动点迭代法推广到非线性方程组。

### 5.2.1 多元牛顿法

#### 1. 向量函数的微分

为了将牛顿法推广到非线性方程组, 需要类似一元函数导数的向量值函数的导数概念, 尤其是向量值函数的一阶 Taylor 展开。

如果对每个  $1 \leq i \leq n$ , 函数  $f_i(\mathbf{x})$  在点  $\mathbf{x}_0$  可微, 则称向量值函数  $\mathbf{f}(\mathbf{x})$  在点  $\mathbf{x}_0$  可微。矩阵

$$\mathbf{D}_f(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}_0) & \frac{\partial f_n}{\partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}_0) \end{bmatrix}$$

称为  $\mathbf{f}(\mathbf{x})$  在点  $\mathbf{x}_0$  的 Jacobi 矩阵, 在向量值函数分析中它相当于一元函数的导数。由 Jacobi 矩阵定义的线性映射

$$L(\mathbf{d}) = \mathbf{D}_f(\mathbf{x}_0)\mathbf{d} \quad (5.25)$$

称为  $\mathbf{f}(\mathbf{x})$  在点  $\mathbf{x}_0$  的微分, 它是  $\mathbf{f}(\mathbf{x}_0 + \mathbf{d}) - \mathbf{f}(\mathbf{x}_0)$  的局部线性近似。也就是说, 在  $\mathbf{x}_0$  的邻域内

$$\mathbf{f}(\mathbf{x}_0 + \mathbf{d}) - \mathbf{f}(\mathbf{x}_0) \approx L(\mathbf{d}) = \mathbf{D}_f(\mathbf{x}_0)\mathbf{d}$$

即

$$\mathbf{f}(\mathbf{x}_0 + \mathbf{d}) \approx \mathbf{f}(\mathbf{x}_0) + \mathbf{D}_f(\mathbf{x}_0)\mathbf{d} \quad (5.26)$$

例如: 令  $\mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_1^2 + 4x_2^2 - 4 \\ 4x_1^2 + x_2^2 - 4 \end{bmatrix}$ , ( $f_1(\mathbf{x}) = x_1^2 + 4x_2^2 - 4$ ,  $f_2(\mathbf{x}) = 4x_1^2 + x_2^2 - 4$ ), 它在点  $\mathbf{x}$  的 Jacobi 矩阵是

$$\mathbf{D}_f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} 2x_1 & 8x_2 \\ 8x_1 & 2x_2 \end{bmatrix}$$

在点  $\mathbf{x}_0 = (1, 1)^T$  的微分

$$L(\mathbf{d}) = \mathbf{D}_f(1, 1)\mathbf{d} = \begin{bmatrix} 2 & 8 \\ 8 & 2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

由式(5.26),得到  $f(1.1, 1.1)$  的近似值

$$f(1.1, 1.1) \approx f(1, 1) + \begin{bmatrix} 2 & 8 \\ 8 & 2 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

而精确值  $f(1.1, 1.1) = (2.05, 2.05)^T$ 。可见,  $f(x)$  在点  $x_0$  邻域内可由该点微分得到很好的线性近似。

回忆  $n$  元函数在点  $x_0$  的一阶 Taylor 展开

$$f_i(x) = f_i(x_0) + \nabla_{f_i}^T(x_0)(x - x_0) + r_i(x_0, x)$$

其中

$$\nabla_{f_i}(x_0) = \left( \frac{\partial f_i}{\partial x_1}(x_0), \frac{\partial f_i}{\partial x_2}(x_0), \dots, \frac{\partial f_i}{\partial x_n}(x_0) \right)^T$$

是  $f_i(x)$  在点  $x_0$  的梯度;

$$r_i(x_0, x) = o(\|x_0 - x\|_2)$$

是  $\|x_0 - x\|_2$  的无穷小量, 即

$$\lim_{\|x-x_0\|_2 \rightarrow 0} \frac{r_i(x_0, x)}{\|x - x_0\|_2} = 0$$

因此

$$\begin{aligned} f(x) &= f(x_0) + \begin{bmatrix} \nabla_{f_1}^T(x_0) \\ \nabla_{f_2}^T(x_0) \\ \vdots \\ \nabla_{f_n}^T(x_0) \end{bmatrix} (x - x_0) + \begin{bmatrix} r_1(x_0, x) \\ r_2(x_0, x) \\ \vdots \\ r_n(x_0, x) \end{bmatrix} \\ &= f(x_0) + D_f(x_0)(x - x_0) + r(x_0, x) \end{aligned} \quad (5.27)$$

其中  $r(x_0, x) = o(\|x_0 - x\|_2)$  是  $\|x_0 - x\|_2$  的无穷小向量, 即

$$\lim_{\|x-x_0\|_2 \rightarrow 0} \frac{\|r(x_0, x)\|_2}{\|x - x_0\|_2} = 0$$

公式(5.27)称为向量值函数的一阶 Taylor 展开式。

## 2. 多元牛顿法

求解非线性方程组的牛顿法是一元方程牛顿思想的自然推广。用 Jacobi 矩阵  $D_f(x_k)$  替代一元牛顿法中的导数  $f'(x_k)$ , 得到多元牛顿迭代

$$x_{k+1} = x_k - D_f^{-1}(x_k)f(x_k), \quad k = 0, 1, 2, \dots \quad (5.28)$$

---

### 多元牛顿法

解方程  $f(x) = 0$

给定初始值:  $x_0$

迭代:  $D_f(x_k)d_k = -f(x_k)$

$x_{k+1} = x_k + d_k, k = 0, 1, 2, \dots$

---

已知第  $k$  次迭代  $x_k$ , 自然希望下次迭代  $x_{k+1}$  更靠近方程组  $f(x) = 0$  的解。由一阶 Taylor 展开式

$$\begin{aligned} 0 &\approx \mathbf{f}(\mathbf{x}_{k+1}) = \mathbf{f}(\mathbf{x}_k) + \mathbf{D}_f(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k) + \mathbf{r}(\mathbf{x}_{k+1}, \mathbf{x}_k) \\ &\approx \mathbf{f}(\mathbf{x}_k) + \mathbf{D}_f(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k) \end{aligned}$$

于是,我们希望

$$\mathbf{f}(\mathbf{x}_k) + \mathbf{D}_f(\mathbf{x}_k)(\mathbf{x}_{k+1} - \mathbf{x}_k) = 0$$

故,第  $k+1$  次迭代

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{D}_f^{-1}(\mathbf{x}_k)\mathbf{f}(\mathbf{x}_k)$$

在求解大型方程组中,为避免计算 Jacobi 矩阵的逆,可先解线性方程组(称为牛顿方程)

$$\mathbf{D}_f(\mathbf{x}_k)\mathbf{d}_k = -\mathbf{f}(\mathbf{x}_k) \quad (\mathbf{d}_k = \mathbf{x}_{k+1} - \mathbf{x}_k)$$

得到  $\mathbf{d}_k$ ,再实现牛顿迭代  $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$ 。牛顿法在每步迭代的计算量是:计算  $n$  个分量函数的值, $n^2$  个偏导数(Jacobi 矩阵)和解一次牛顿方程,因此有  $O(n^3)$  次运算。

**例 5.2.1** 用牛顿法解方程组

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_1^2 + 4x_2^2 - 4 \\ 4x_1^2 + x_2^2 - 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (5.29)$$

**解** 计算  $\mathbf{f}(\mathbf{x})$  在点  $\mathbf{x}$  的 Jacobi 矩阵

$$\mathbf{D}_f(\mathbf{x}) = \begin{bmatrix} 2x_1 & 8x_2 \\ 8x_1 & 2x_2 \end{bmatrix}$$

初始值取  $\mathbf{x}_0 = \begin{bmatrix} 6.5 \\ 5.5 \end{bmatrix}$ ,  $\mathbf{D}_f(\mathbf{x}_0) = \begin{bmatrix} 13 & 44 \\ 52 & 11 \end{bmatrix}$ ,  $\mathbf{f}(\mathbf{x}_0) = \begin{bmatrix} 159.25 \\ 195.25 \end{bmatrix}$ 。从牛顿方程  $\mathbf{D}_f(\mathbf{x}_0)\mathbf{d}_0 = -\mathbf{f}(\mathbf{x}_0)$  得到

$$\mathbf{d}_0 = \begin{bmatrix} -829/260 \\ -101/110 \end{bmatrix}$$

因此

$$\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{d}_0 = \begin{bmatrix} 6.5 \\ 5.5 \end{bmatrix} + \begin{bmatrix} -829/260 \\ -101/110 \end{bmatrix} = \begin{bmatrix} 3.311538461538461 \\ 2.822727272727272 \end{bmatrix}$$

表 5.6 给出了最初 10 次迭代结果。这个方程组有四个解  $x_1 = \pm 2/\sqrt{5}$ ,  $x_2 = \pm 2/\sqrt{5}$ , 分别以初始值  $\mathbf{x}_0 = (-6.5, 5.5)$ 、 $(-6.5, -5.5)$  和  $(6.5, -5.5)$  进行迭代可得到另外三个解,但这样做没有必要,因为根据方程组的平方项特性,从一个解就可确定其他三个解。

**表 5.6** 多元牛顿迭代

$k$	$\mathbf{x}_k$
0	{6.5, 5.5}
1	{3.3115384615384613, 2.8227272727272723}
2	{1.7765590100955955, 1.5530705606792563}
3	{1.1134338611524544, 1.0340895796272630}
4	{0.9159659295118101, 0.9038584739280965}
5	{0.8946804303627222, 0.8944763962130312}
6	{0.8944272268396262, 0.8944271923533070}
7	{0.8944271909999166, 0.8944271909999160}
8	{0.8944271909999159, 0.8944271909999159}
9	{0.8944271909999159, 0.8944271909999159}
10	{0.8944271909999159, 0.8944271909999159}

从表 5.6 可以看出,第 6 次迭代得到精确到小数点后 6 位的解。此后,收敛速加快,第 8 次迭代精度准确到小数点后 16 位,呈现出局部二阶收敛性特征。

事实上,如果  $f(\mathbf{x})$  可微且在解处的 Jacobi 矩阵  $\mathbf{D}_f(\mathbf{x}_*)$  可逆,则牛顿法超线性收敛;如果  $\mathbf{D}_f(\mathbf{x})$  在解的局部邻域内还满足李普希兹条件  $\|\mathbf{D}_f(\mathbf{x}) - \mathbf{D}_f(\mathbf{y})\|_2 < L \|\mathbf{x} - \mathbf{y}\|_2$ , 则牛顿法二阶收敛,证明请参考[9]。

### 3. 不精确牛顿法

在每次迭代中,牛顿法需要求解牛顿方程  $\mathbf{D}_f(\mathbf{x}_k)\mathbf{d}_k = -f(\mathbf{x}_k)$ 。对于大规模问题,求此方程精确解的代价十分高昂。为了节省计算量,通常采用迭代法解牛顿方程获得近似解,这样就产生了不精确牛顿法。这里的不精确是相对牛顿方程的,并非说原始非线性方程组的解不精确。事实上,在每步迭代中没有必要获得牛顿方程的精确解,只要将精度控制在某个范围内就可以。具体地说,在不精确牛顿法的每步迭代中,先将求解线性方程组的某种迭代法用于牛顿方程,得到满足下述条件的迭代解  $\mathbf{d}_k$  :

$$\|\mathbf{D}_f(\mathbf{x}_k)\mathbf{d}_k + f(\mathbf{x}_k)\|_2 \leq \eta_k \|f(\mathbf{x}_k)\|_2$$

其中  $\eta_k \in (0, 1)$  是强制项,用来控制近似解  $\mathbf{d}_k$  的精度。

#### 牛顿-雅可比法

解方程  $f(\mathbf{x})=0$

给定初始值:  $\mathbf{x}_0$

迭代 内迭代: 应用雅可比法迭代求解牛顿方程  $\mathbf{D}_f(\mathbf{x}_k)\mathbf{d}_k = -f(\mathbf{x}_k)$  使得

$$\|\mathbf{D}_f(\mathbf{x}_k)\mathbf{d}_k + f(\mathbf{x}_k)\|_2 \leq \eta_k \|f(\mathbf{x}_k)\|_2$$

外迭代:  $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k, k=0, 1, 2, \dots$

可以看出,不精确牛顿法是一种内外迭代法。外层迭代是牛顿迭代;内层迭代是求解线性方程组的某种迭代,比如 Jacobi、Gauss-Seidel 和逐步松弛(SOR)等迭代法,使用不同的线性迭代法导致不同版本的不精确牛顿法。

## 5.2.2 多元拟牛顿法

拟牛顿法是对牛顿法修正所得到的,其目的一是为了避免牛顿法每迭代步都要计算 Jacobi 矩阵和求解牛顿方程,二是为了适应向量值函数不可微的一类方程组。主要思想是以矩阵  $\mathbf{A}_k$  简单近似(或替代)Jacobi 矩阵  $\mathbf{D}_f(\mathbf{x}_k)$ ,一般迭代格式为

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{A}_k^{-1}f(\mathbf{x}_k), & k=0, 1, 2, \dots \\ \mathbf{A}_{k+1}\mathbf{d}_k = \mathbf{D}_k \\ \mathbf{A}_{k+1} = \mathbf{A}_k + \Delta\mathbf{A}_k, & \text{rank}(\Delta\mathbf{A}_k) = m \end{cases} \quad (5.30)$$

其中  $\mathbf{d}_k = \mathbf{x}_{k+1} - \mathbf{x}_k, \mathbf{D}_k = f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)$ ; 第二个方程称为拟牛顿方程,它相当于 Jacobi 矩阵要满足的方程;  $\Delta\mathbf{A}_k$  称为  $\mathbf{A}_k$  的秩为  $m$  的修正矩阵,在实践中常见的是  $m=1, 2$  两种情形,分别称为秩 1 拟牛顿法和秩 2 拟牛顿法。

### 1. 布洛伊登法

布洛伊登法(Broyden)是秩 1 拟牛顿法。布洛伊登注意到拟牛顿方程  $\mathbf{A}_{k+1}\mathbf{d}_k = \mathbf{D}_k$  没有利用  $\mathbf{d}_k$  的正交补空间  $\mathcal{O}_k = \{\mathbf{u}; \mathbf{d}_k^T \mathbf{u} = 0\}$  的任何信息,于是要求  $\mathbf{A}_{k+1}$  对任意  $\mathbf{u} \in \mathcal{O}_k$  有

$\mathbf{A}_{k+1}\mathbf{u}=\mathbf{A}_k\mathbf{u}$ 。也就是说,要求  $\mathbf{A}_{k+1}$  满足

$$\begin{cases} \mathbf{A}_{k+1}\mathbf{d}_k = \mathbf{D}_k \\ \mathbf{A}_{k+1}\mathbf{u} = \mathbf{A}_k\mathbf{u}, \quad \forall \mathbf{u} \in \mathcal{O}_k \end{cases} \quad (5.31)$$

且

$$\text{rank}\Delta\mathbf{A}_k = \text{rank}(\mathbf{A}_{k+1} - \mathbf{A}_k) = 1 \quad (5.32)$$

根据式(5.32)和式(5.31)的第二式

$$\mathbf{A}_{k+1} - \mathbf{A}_k = \mathbf{w}\mathbf{d}_k^T \quad (\text{其中 } \mathbf{w} \text{ 为待定向量})$$

由式(5.31)的第一式

$$\mathbf{w}\mathbf{d}_k^T\mathbf{d}_k = (\mathbf{A}_{k+1} - \mathbf{A}_k)\mathbf{d}_k = \mathbf{D}_k - \mathbf{A}_k\mathbf{d}_k \Rightarrow \mathbf{w} = \frac{\mathbf{D}_k - \mathbf{A}_k\mathbf{d}_k}{\mathbf{d}_k^T\mathbf{d}_k}$$

因此

$$\mathbf{A}_{k+1} = \mathbf{A}_k + \frac{(\mathbf{D}_k - \mathbf{A}_k\mathbf{d}_k)\mathbf{d}_k^T}{\mathbf{d}_k^T\mathbf{d}_k} \quad (5.33)$$

这是布洛伊登法  $\mathbf{A}_k$  的更新公式。

#### 布洛伊登法 I

解方程  $f(\mathbf{x})=0$

给定初始值:  $\mathbf{x}_0; \mathbf{A}_0$

迭代: 解拟牛顿方程  $\mathbf{A}_k\mathbf{d}_k = -f(\mathbf{x}_k), k=0,1,2,\dots$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$$

$$\mathbf{D}_k = f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)$$

$$\mathbf{A}_{k+1} = \mathbf{A}_k + \frac{(\mathbf{D}_k - \mathbf{A}_k\mathbf{d}_k)\mathbf{d}_k^T}{\mathbf{d}_k^T\mathbf{d}_k}$$

布洛伊登法需要一个近似的 Jacobi 矩阵  $\mathbf{A}_0$  开始。若无法获得初值  $\mathbf{x}_0$  的 Jacobi 矩阵,通常取  $\mathbf{A}_0$  是单位矩阵。布洛伊登法的计算比牛顿法简单,在每步迭代中计算  $n$  个分量函数在点  $\mathbf{x}_k$  的值,然后解一次拟牛顿方程  $\mathbf{A}_k\mathbf{d}_{k+1} = -f(\mathbf{x}_k)$  得到  $\mathbf{x}_{k+1}$ ,再计算  $n$  个分量函数在点  $\mathbf{x}_{k+1}$  的值,最后计算  $\mathbf{A}_{k+1}$ 。从表面上看,需要  $O(n^3)$  次运算,若利用下面 Sherman-Morrison 求逆引理,计算量可降低到  $O(n^2)$ 。

#### Sherman-Morrison 求逆引理

若  $\mathbf{A}$  为非奇异矩阵,则  $\mathbf{A} = \mathbf{A} + \mathbf{u}\mathbf{v}^T$  也是非奇异的,当  $s \triangleq 1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} \neq 0$  时,

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{A}^{-1} - s^{-1}\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} \quad (5.34)$$

证明留给读者做练习。

由此引理,

$$\begin{aligned} \mathbf{B}_{k+1} &= \mathbf{A}_{k+1}^{-1} = \left( \mathbf{A}_k + \frac{(\mathbf{D}_k - \mathbf{A}_k\mathbf{d}_k)\mathbf{d}_k^T}{\mathbf{d}_k^T\mathbf{d}_k} \right)^{-1} \\ &= \mathbf{B}_k - \left( 1 + \frac{\mathbf{d}_k^T\mathbf{B}_k(\mathbf{D}_k - \mathbf{A}_k\mathbf{d}_k)}{\mathbf{d}_k^T\mathbf{d}_k} \right)^{-1} \left( \frac{\mathbf{B}_k(\mathbf{D}_k - \mathbf{A}_k\mathbf{d}_k)\mathbf{d}_k^T\mathbf{B}_k}{\mathbf{d}_k^T\mathbf{d}_k} \right) \\ &= \mathbf{B}_k + \frac{(\mathbf{d}_k - \mathbf{B}_k\mathbf{D}_k)\mathbf{d}_k^T\mathbf{B}_k}{\mathbf{d}_k^T\mathbf{B}_k\mathbf{D}_k} \end{aligned}$$

这样就得到一种不用解拟牛顿方程的布洛伊登法,称为第二种布洛伊登法。与布洛伊登法 I 不同,这里的初始矩阵  $\mathbf{B}_0 = \mathbf{D}_f^{-1}(\mathbf{x}_0)$ ,若 Jacobi 矩阵不可计算也可取单位矩阵。

### 布洛伊登法 II

解方程  $f(\mathbf{x})=0$

给定初始值:  $\mathbf{x}_0; \mathbf{B}_0$

迭代:  $\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{B}_k f(\mathbf{x}_k), k=0, 1, 2, \dots$

$$\mathbf{D}_k = f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)$$

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{(\mathbf{d}_k - \mathbf{B}_k \mathbf{D}_k) \mathbf{d}_k^T \mathbf{B}_k}{\mathbf{d}_k^T \mathbf{B}_k \mathbf{D}_k}$$

### 例 5.2.2 用布洛伊登法 I 和 II 解方程组

$$\begin{cases} x^3 - y^3 + x = 0 \\ x^2 + y^2 - 1 = 0 \end{cases} \quad ((x_0, y_0) = (1, 1); \mathbf{A}_0 = \mathbf{B}_0 = \mathbf{I})$$

解 考虑布洛伊登法 I, 由

$$\mathbf{A}_0 = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}, \quad f(\mathbf{x}_0) = (1, 1)^T$$

从  $\mathbf{A}_0 \mathbf{d}_0 = -f(\mathbf{x}_0)$  得到  $\mathbf{d}_0 = (-1, -1)^T$ , 因此

$$\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{d}_0 = (0, 0)^T$$

$$f(\mathbf{x}_1) = (0, -1)^T, \quad \mathbf{D}_0 = f(\mathbf{x}_1) - f(\mathbf{x}_0) = (-1, -2)^T$$

$$\mathbf{A}_1 = \mathbf{A}_0 + \frac{(\mathbf{D}_0 - \mathbf{A}_0 \mathbf{d}_0) \mathbf{d}_0^T}{\mathbf{d}_0^T \mathbf{d}_0} = \begin{bmatrix} 1 & 0 \\ 0.5 & 1.5 \end{bmatrix}$$

重复上面的过程得到  $\mathbf{x}_2 = \mathbf{x}_0 + \mathbf{d}_0 = (0, 0.666667)^T$ 。表 5.7 给出了布洛伊登法 I 和 II 的最初 15 次迭代结果。可以看出两种方法,每次迭代结果一致,具有很快的收敛速度。事实上,在一定的条件下布洛伊登法是超线性收敛的。

表 5.7 布洛伊登迭代

	布洛伊登法 I	布洛伊登法 II
0	{1., 1.}	{1., 1.}
1	{0., 0.}	{0., 0.}
2	{0., 0.6666666666666666}	{0., 0.6666666666666666}
3	{0.5, 1.125}	{0.5, 1.125}
4	{0.706278271598889, 0.502845255939402}	{0.706278271598889, 0.502845255939402}
5	{0.468495884132446, 0.809419903629634}	{0.468495884132446, 0.809419903629634}
6	{0.542857470718812, 0.841055958859420}	{0.542857470718812, 0.841055958859420}
7	{0.511338131697325, 0.867764710168130}	{0.511338131697325, 0.867764710168130}
8	{0.507607106104817, 0.860440707336289}	{0.507607106104817, 0.860440707336289}
9	{0.507993931920313, 0.861355824683596}	{0.507993931920313, 0.861355824683596}
10	{0.507992049859365, 0.861361723650724}	{0.507992049859365, 0.861361723650724}
11	{0.507992000676276, 0.861361786301928}	{0.507992000676276, 0.861361786301928}

续表

	布洛伊登法 I	布洛伊登法 II
12	{0.507992000407849, 0.861361786662122}	{0.507992000407849, 0.861361786662122}
13	{0.507992000407952, 0.861361786661985}	{0.507992000407952, 0.861361786661985}
14	{0.507992000407951, 0.861361786661985}	{0.507992000407951, 0.861361786661985}
15	{0.507992000407951, 0.861361786661985}	{0.507992000407951, 0.861361786661985}

## 2. 秩 2 拟牛顿法

在理解了布洛伊登法的基础上,不难构造秩 2 拟牛顿法。仿照布洛伊登法 II,由于任何一个秩 2 矩阵可表示为  $\mathbf{u}^{(1)}(\mathbf{v}^{(1)})^T + \mathbf{u}^{(2)}(\mathbf{v}^{(2)})^T$ , 其中  $\mathbf{u}^{(i)}, \mathbf{v}^{(i)}$  是  $n$  维向量,因此要使  $\mathbf{B}_{k+1} - \mathbf{B}_k$  的秩等于 2,则它必为下述形式

$$\mathbf{B}_{k+1} - \mathbf{B}_k = \mathbf{u}_k^{(1)}(\mathbf{v}_k^{(1)})^T + \mathbf{u}_k^{(2)}(\mathbf{v}_k^{(2)})^T, \quad \mathbf{u}_k^{(i)}, \mathbf{v}_k^{(i)} \in \mathcal{R}^n$$

此外,在布洛伊登法 II 中,  $\mathbf{B}_{k+1}\mathbf{D}_k = \mathbf{d}_k$ , 这是拟牛顿方程的逆形式,在这里当然也要求成立:

$$(\mathbf{B}_k + \mathbf{u}_k^{(1)}(\mathbf{v}_k^{(1)})^T + \mathbf{u}_k^{(2)}(\mathbf{v}_k^{(2)})^T)\mathbf{D}_k = \mathbf{d}_k$$

即

$$\mathbf{u}_k^{(1)}(\mathbf{v}_k^{(1)})^T\mathbf{D}_k + \mathbf{u}_k^{(2)}(\mathbf{v}_k^{(2)})^T\mathbf{D}_k = \mathbf{d}_k - \mathbf{B}_k\mathbf{D}_k \quad (5.35)$$

若  $(\mathbf{v}_k^{(1)})^T\mathbf{D}_k \neq 0, (\mathbf{v}_k^{(2)})^T\mathbf{D}_k \neq 0$ , 取

$$\mathbf{u}_k^{(1)} = \frac{\mathbf{d}_k}{(\mathbf{v}_k^{(1)})^T\mathbf{D}_k}, \quad \mathbf{u}_k^{(2)} = -\frac{\mathbf{B}_k\mathbf{D}_k}{(\mathbf{v}_k^{(2)})^T\mathbf{D}_k}$$

则(5.35)成立。于是

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{\mathbf{d}_k(\mathbf{v}_k^{(1)})^T}{(\mathbf{v}_k^{(1)})^T\mathbf{D}_k} - \frac{\mathbf{B}_k\mathbf{D}_k(\mathbf{v}_k^{(2)})^T}{(\mathbf{v}_k^{(2)})^T\mathbf{D}_k} \quad (5.36)$$

在式(5.36)中,若  $\mathbf{v}_k^{(1)}, \mathbf{v}_k^{(2)}$  线性相关,则修正矩阵是秩 1 的。在式(5.36)中选择不同的线性无关向量  $\mathbf{v}_k^{(1)}, \mathbf{v}_k^{(2)}$  就得到不同的秩 2 拟牛顿法的更新矩阵  $\mathbf{B}_{k+1}$ 。两个较著名的算法是 DFP(Davidon-Fletcher-Powell)和 BFS(Broyden-Fletcher-Shanno)法。

在 DFP 法中,取  $\mathbf{v}_k^{(1)} = \mathbf{d}_k, \mathbf{v}_k^{(2)} = \mathbf{B}_k^T\mathbf{D}_k$ , DFP 的更新公式

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{\mathbf{d}_k\mathbf{d}_k^T}{\mathbf{d}_k^T\mathbf{D}_k} - \frac{\mathbf{B}_k\mathbf{D}_k(\mathbf{B}_k\mathbf{D}_k)^T}{(\mathbf{B}_k\mathbf{D}_k)^T\mathbf{D}_k}, \quad k = 0, 1, 2, \dots$$

### DFP 法

解方程  $f(\mathbf{x}) = 0$

给定初始值:  $\mathbf{x}_0; \mathbf{B}_0$

迭代:  $\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{B}_k\mathbf{f}(\mathbf{x}_k), k = 0, 1, 2, \dots$

$$\mathbf{D}_k = \mathbf{f}(\mathbf{x}_{k+1}) - \mathbf{f}(\mathbf{x}_k)$$

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{\mathbf{d}_k\mathbf{d}_k^T}{\mathbf{d}_k^T\mathbf{D}_k} - \frac{\mathbf{B}_k\mathbf{D}_k(\mathbf{B}_k\mathbf{D}_k)^T}{(\mathbf{B}_k\mathbf{D}_k)^T\mathbf{D}_k}$$

DFP 法不仅可以解方程组也可用求解无约束最优化问题,是一个较为成功的算法,但在实际问题中稳定性稍逊下面的 BFS 法。

取  $\mathbf{v}_k^{(1)} = s_k \mathbf{d}_k - \mathbf{B}_k^T \mathbf{D}_k (s_k = 1 + \mathbf{D}_k^T \mathbf{B}_k \mathbf{D}_k / \mathbf{d}_k^T \mathbf{D}_k)$ ,  $\mathbf{v}_k^{(2)} = \mathbf{d}_k$ , 代入式(5.36)得到 BFS 的更新公式

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{s_k \mathbf{d}_k \mathbf{d}_k^T - \mathbf{d}_k \mathbf{D}_k^T \mathbf{B}_k - \mathbf{B}_k \mathbf{D}_k \mathbf{d}_k^T}{\mathbf{d}_k^T \mathbf{D}_k}, \quad k = 0, 1, 2, \dots$$

### BFS 法

解方程组  $f(\mathbf{x}) = 0$

给定初始值:  $\mathbf{x}_0; \mathbf{B}_0$

迭代:  $\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{B}_k f(\mathbf{x}_k), k = 0, 1, 2, \dots$

$$\mathbf{D}_k = f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)$$

$$s_k = 1 + \mathbf{D}_k^T \mathbf{B}_k \mathbf{D}_k / \mathbf{d}_k^T \mathbf{D}_k$$

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{s_k \mathbf{d}_k \mathbf{d}_k^T - \mathbf{d}_k \mathbf{D}_k^T \mathbf{B}_k - \mathbf{B}_k \mathbf{D}_k \mathbf{d}_k^T}{\mathbf{d}_k^T \mathbf{D}_k}$$

### 5.2.3 多元不动点法

解方程组  $f(\mathbf{x}) = 0$  的不动点法与解方程的不动点法类似,也是先将它化为不动点形式的同解方程组  $\mathbf{x} = \mathbf{g}(\mathbf{x})$ , 其中  $\mathbf{g}(\mathbf{x})$  称为迭代映射, 然后从初始点  $\mathbf{x}_0$  开始迭代

$$\mathbf{x}_{k+1} = \mathbf{g}(\mathbf{x}_k), \quad k = 0, 1, 2, \dots$$

若迭代收敛于  $\mathbf{x}_*$  且  $\mathbf{g}(\mathbf{x})$  连续, 则  $\mathbf{x}_*$  是  $\mathbf{g}(\mathbf{x})$  不动点, 同时是  $f(\mathbf{x}) = 0$  的解。

#### 不动点法

解方程组  $f(\mathbf{x}) = 0$

构造同解方程组  $\mathbf{x} = \mathbf{g}(\mathbf{x})$

给定初始值:  $\mathbf{x}_0$

迭代:  $\mathbf{x}_{k+1} = \mathbf{g}(\mathbf{x}_k), k = 0, 1, 2, \dots$

#### 例 5.2.3 利用不动点迭代解方程组

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_1^2 + x_2^2/4 - 1 \\ x_1^2/8 + x_2^2 - 1 \end{bmatrix} = 0 \quad \left( \mathbf{x}_0 = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix} \right) \quad (5.37)$$

解 构造不动点形式的方程组

$$\mathbf{x} = \begin{bmatrix} (1 - x_2^2/4)^{1/2} \\ (1 - x_1^2/8)^{1/2} \end{bmatrix} \triangleq \mathbf{g}(\mathbf{x})$$

然后, 开始迭代

$$\mathbf{x}_1 = \mathbf{g}(\mathbf{x}_0) = \begin{bmatrix} 0.661438 \\ 0.847791 \end{bmatrix}, \quad \mathbf{x}_2 = \mathbf{g}(\mathbf{x}_1) = \begin{bmatrix} 0.905711 \\ 0.972272 \end{bmatrix}, \quad \dots$$

表 5.8 给出了最初 20 次的迭代结果。方程组(5.37)在点  $\mathbf{x}_0$  附近的精确解是

$$x_1 = 2\sqrt{6/31} \approx 0.8798826901281201, \quad x_2 = 2\sqrt{6/31} \approx 0.9503819266229829$$

从表 5.8 中可以看出, 第 19 次迭代得到精确到小数点后 13 位的解。从方程组(5.37)的平方项特性, 容易得到另外三个解。

表 5.8 不动点迭代

$k$	$\mathbf{x}_k$
0	{1.5, 1.5}
1	{0.661437827766147, 0.847791247890658}
2	{0.905711046636839, 0.972271824131502}
3	{0.873883215881847, 0.947344149451507}
4	{0.880701291940122, 0.951073612095825}
5	{0.879695825893103, 0.950287143076699}
6	{0.879908282962367, 0.950403549413764}
7	{0.879876851221481, 0.950378964780248}
8	{0.879883489915457, 0.950382602342641}
9	{0.879882507662873, 0.950381834065537}
10	{0.879882715121485, 0.950381947739229}
11	{0.879882684426081, 0.950381923730562}
12	{0.879882690909162, 0.950381927282865}
13	{0.879882689949931, 0.950381926532594}
14	{0.879882690152527, 0.950381926643604}
15	{0.879882690122551, 0.950381926620158}
16	{0.879882690128882, 0.950381926623627}
17	{0.879882690127945, 0.950381926622894}
18	{0.879882690128143, 0.950381926623003}
19	{0.879882690128114, 0.950381926622980}
20	{0.879882690128120, 0.950381926622983}

关于不动点迭代的收敛性,有下述定理:

**定理 5.2.1** 如果映射  $\mathbf{g}(\mathbf{x})$  在闭集  $D$  上满足: 存在  $\rho < 1$  使得  $\forall \mathbf{x}, \mathbf{y} \in D$  有

$$\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\|_2 \leq \rho \|\mathbf{x} - \mathbf{y}\|_2 \quad (5.38)$$

且  $\mathbf{g}(D) \subset D$ , 则  $\mathbf{g}(\mathbf{x})$  在  $D$  上有唯一不动点  $\mathbf{x}_* \in D$ , 且对任意  $\mathbf{x}_0 \in D$  迭代都收敛于  $\mathbf{x}_*$ , 即在  $D$  上整体收敛于  $\mathbf{x}_*$ 。

**证明** 因  $\mathbf{g}(D) \subset D$ , 所以  $\mathbf{x}_{k+1} \in D, k=0, 1, 2, \dots$ 。由式(5.38),

$$\begin{aligned} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2 &= \|\mathbf{g}(\mathbf{x}_k) - \mathbf{g}(\mathbf{x}_{k-1})\|_2 \\ &\leq \rho \|\mathbf{x}_k - \mathbf{x}_{k-1}\|_2 \leq \dots \leq \rho^k \|\mathbf{x}_1 - \mathbf{x}_0\|_2 \end{aligned}$$

于是

$$\begin{aligned} \|\mathbf{x}_{k+N} - \mathbf{x}_k\|_2 &\leq \sum_{i=1}^N \|\mathbf{x}_{k+i} - \mathbf{x}_{k+i-1}\|_2 \leq \sum_{i=1}^N \rho^{i-1} \cdot \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2 \\ &= \frac{1 - \rho^N}{1 - \rho} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2 \leq \frac{(1 - \rho^N)\rho^k}{1 - \rho} \|\mathbf{x}_1 - \mathbf{x}_0\|_2 \end{aligned}$$

由  $\rho < 1$ ,  $\{\mathbf{x}_k\}$  是闭集  $D$  上的 Cauchy 点列, 因而存在  $\mathbf{x}_* \in D$  使得  $\mathbf{x}_k \rightarrow \mathbf{x}_* (k \rightarrow \infty)$ 。式(5.38)表明映射  $\mathbf{g}(\mathbf{x})$  在  $D$  上连续, 因此  $\mathbf{x}_*$  是  $\mathbf{g}(\mathbf{x})$  的不动点。若还存在一个不动点  $\mathbf{y}_* \neq \mathbf{x}_*$ , 则

$$\|\mathbf{x}_* - \mathbf{y}_*\|_2 = \|\mathbf{g}(\mathbf{x}_*) - \mathbf{g}(\mathbf{y}_*)\|_2 \leq \rho \|\mathbf{x}_* - \mathbf{y}_*\|_2 \Rightarrow \rho \geq 1$$

与  $\rho < 1$  矛盾, 故不动点是唯一的。证毕。

不难看出,不动点迭代误差的事先估计和事后估计分别为

$$\| \mathbf{x}_{k+1} - \mathbf{x}_* \|_2 \leq \frac{\rho^{k+1}}{1-\rho} \| \mathbf{x}_1 - \mathbf{x}_0 \|_2 \quad (5.39)$$

$$\| \mathbf{x}_{k+1} - \mathbf{x}_* \|_2 \leq \frac{\rho}{1-\rho} \| \mathbf{x}_{k+1} - \mathbf{x}_k \|_2 \quad (5.40)$$

## 习 题

以下习题必要时可上机作业。

1. 用二分法求下列方程的根,精确到小数点后 8 位:

(i)  $x^3 = 9$

(ii)  $x^5 + x = 0$

(iii)  $\sin x - 6x - 5 = 0$

2. 用下列方法求方程  $x^3 - 3x - 1 = 0$  在  $x_0 = 2.0$  附近的根,精确到小数点后 6 位:

(i) 牛顿法;

(ii) 割线法,初始值:  $x_0 = 2.0, x_1 = 1.9$ ;

(iii) 抛物线法,初始值:  $x_0 = 1.0, x_1 = 3.0, x_2 = 2.0$ 。

3. 求下列函数的所有不动点:

(i)  $\varphi(x) = 3/x$

(ii)  $\varphi(x) = x^2 - 2x + 2$

(iii)  $\varphi(x) = (6+x)/(3x-2)$

4. 下列哪个不动点迭代收敛到  $\sqrt{2}$ ? 依收敛速度从慢到快进行排序。

(i)  $x_{n+1} = \frac{1}{2}x_n + \frac{1}{x_n}$

(ii)  $x_{n+1} = \frac{2}{3}\left(x_n + \frac{1}{x_n}\right)$

(iii)  $x_{n+1} = \frac{3}{4}x_n + \frac{1}{2x_n}$

5. 用不动点法解下列方程,精确到小数点后 8 位:

(i)  $x^5 + x = 0$

(ii)  $\sin x - 6x - 5 = 0$

(iii)  $\ln x + x^2 - 3 = 0$

6. 令  $\varphi(x) = x - p(x)f(x) - q(x)f^2(x)$ , 试确定函数  $p(x)$  和  $q(x)$ , 使得求解方程  $f(x) = 0$  以  $\varphi(x)$  为迭代函数的不动点法至少三阶收敛。

7. 证明 Sherman-Morrison 求逆引理: 若  $\mathbf{A}$  为非奇异矩阵, 则  $\mathbf{A} + \mathbf{u}\mathbf{v}^T$  也是非奇异的, 当  $s \triangleq 1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u} \neq 0$  时,

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

8. 用牛顿法找下述方程组的所有解

$$\begin{cases} x^2 + y^2 = 4 \\ x^2 - y^2 = 1 \end{cases}$$

9. 用布洛伊登法 I 和 II 求下列方程组的解, 初始值  $x_0 = y_0 = 1$ :

$$(i) \begin{cases} x^2 + y^2 = 1 \\ (x-1)^2 + y^2 = 1 \end{cases}$$

$$(ii) \begin{cases} x^2 + 4y^2 = 4 \\ 4x^2 + y^2 = 4 \end{cases}$$

$$(iii) \begin{cases} x^2 - y^2 = 4 \\ (x-1)^2 + y^2 = 4 \end{cases}$$

10. DFP 法和 BFS 法解习题 9 中的方程组。