中国农业大学学报　1998,3(6):16～18
Journal of China Agricultural University

一维优化问题的一种Q-2阶收敛混合算法①

王兆智②　邓乃扬
(中国农业大学工程基础科学部)

摘　要　利用牛顿法和不精确牛顿法构造了1维优化问题的混合算法。用该算法连续2次迭代只计算1次2阶导数值。在一合理的假设下证明了该算法具有Q-2阶收敛速率。

关键词　1维优化；不精确牛顿法；Q-2阶收敛性
中图分类号　O 221.2

A Hybrid Algorithm With Q-2 Order Convergence Rate for One Dimension Optimization Problems

Wang Zhaozhi　Deng Naiyang
(College of Applied Engineering Sciences, CAU)

Abstract　A hybrid algorithm exploiting Newton method and inexact Newton method for one dimension optimization problems is presented. The second derivative value is computed one time in two successive iterations. The Q-2 order convergence is proved under a reasonable assumption.

Key words　1-dimension optimization; inexact Newton method; Q-2 order convergence

考虑1维优化问题

$$\min f(x), x\in \mathbf{R} \tag{1}$$

其中$f(x)$是连续可微函数。对于问题(1)，利用函数导数信息的常用算法有牛顿法

$$x^{k+1}=x^k-\frac{f'(x^k)}{f''(x^k)} \tag{2}$$

其收敛速度很快。在$f'(x^*)=0, f''(x^*)\neq 0$时，只要初始点充分接近$x^*$，牛顿法2阶收敛(见文献[1])。

由于用牛顿法需要计算2阶导数，在$f''(x)$计算复杂时很不方便，因此用差商

$$A^k=\frac{f'(x^k)-f'(x^{k-1})}{x^k-x^{k-1}}$$

替换式(2)中的2阶导数$f''(x^k)$，得割线法

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

只要初始点x^0, x^1充分接近x^*，其收敛速率是Q-$(\sqrt{5}+1)/2$阶(见文献[1])。为了充分利用1阶导数和函数值信息，文献[2]用

收稿日期:1998-04-10
①国家自然科学基金资助项目
②王兆智，北京清华东路17号 中国农业大学(东校区)71信箱，100083

$$B^k=\left[4f'(x^k)+2f'(x^{k-1})-6\frac{f'(x^k)-f'(x^{k-1})}{x^k-x^{k-1}}\right](x^k-x^{k-1})^{-1} \tag{3}$$

替换 $f''(x_k)$,得近似牛顿法如下:

$$x_{k+1}=x_k-\frac{f'(x^k)(x^k-x^{k-1})}{4f'(x^k)+2f'(x^{k-1})-6[f(x^k)-f(x^{k-1})](x^k-x^{k-1})^{-1}} \tag{4}$$

文献①提出的优化方法的 1 维情况也属迭代公式(4)。文献[2]讨论了式(4)产生的点列在一些情形具有 Q-2阶收敛速率,但在 $f'''(x^*)\neq 0$ 时不易得到。文献[3]和[4]利用牛顿法和不精确牛顿法对单和优化问题和 1 维问题分别构造了混合算法,但理论上也只得到了 2 步 4 阶收敛性结论。

有无一种 1 维优化方法比牛顿法计算量少,在 $f'''(x^*)\neq 0$ 这一较弱假设下,与牛顿法一样具有 Q-2阶收敛速率?笔者构造出交替使用式(2)和(4)的混合迭代算法,并证明了其 Q-2阶收敛性。

求解问题(1)的混合迭代算法的步骤如下:

1)给定初始点 x^0,误差 $\varepsilon>0,k=0$;

2)若 $|f'(x^k)|<\varepsilon$,停止;

3)当 k 是偶数时,x^{k+1}有式(2)产生;当 k 是奇数时,x^{k+1}有式(4)产生;

4)$k=k+1$,转步 2)。

此算法连续 2 次迭代只计算 1 次 2 阶导数,计算量比牛顿法少。

为了证明收敛性理论,首先给出 2 个引理,其中引理 1 为文献[1]中定理 10.2.2 的 1 维情形,引理 2 见文献[2]。

引理 1　若 $f'(x^*)=0,f''(x^*)\neq 0$,且 $f'''(x^*)\neq 0$,则当 x^0 充分接近于 x^* 时,牛顿法(2)产生的点列 $\{x^k\}$ 精确地 2 阶收敛于 x^*,即存在 $M_1,M_2>0$,使

$$M_1|x^{k-1}-x^*|^2\leqslant|x^k-x^*|\leqslant M_2|x^{k-1}-x^*|^2 \tag{5}$$

引理 2　当 x^{k-1},x^k 充分接近于 x^* 时,由式(3)得 $f''(x^k)$ 的近似值 B^k,有误差估计

$$f''(x^k)-B^k=O(x^k-x^{k-1})^2 \tag{6}$$

下面是本文的主要结论。

定理　若 $f'(x^*)=0,f''(x^*)\neq 0$,且 $f'''(x^*)\neq 0$,则当 x^0,x^1 充分接近 x^* 时,上述算法产生的点列 $\{x^k\}$ 收敛于 x^*,收敛速率为 Q-2阶,即存在 $M_3>0$,使

$$|x^{k+1}-x^*|\leqslant M_3|x^k-x^*|^2 \tag{7}$$

证明　当 k 为偶数时,x^{k+1}由式(2)产生。由式(5)得

$$M_1|x^k-x^*|^2\leqslant|x^{k+1}-x^*|\leqslant M_2|x^k-x^*|^2 \tag{8}$$

当 k 为奇数时,x^k 由式(2)产生,从而

$$|x^{k-1}-x^*|\leqslant(|x^k-x^*|M_1^{-1})^{1/2} \tag{9}$$

$$|x^k-x^*|\leqslant M_2|x^{k-1}-x^*|^2 \tag{10}$$

由式(6),(9)和(10),存在 $M_4>0$,使

①Zhang J Z, Deng N Y, Chen L H. New equasi-Newton equation and related methods for unconstrained optimization. Report MA-96-05, Hong Kong: City University of Hong Kong, 1996

$$|f''(x^k)-B^k|=O(x^k-x^{k-1})^2\leqslant$$
$$O(|x^k-x^*|+|x^{k-1}-x^*|)^2=$$
$$O(x^{k-1}-x^*)^2\leqslant$$
$$M_4|x^{k-1}-x^*|^2\leqslant$$
$$M_4|x^k-x^*|M_1^{-1} \tag{11}$$

由式(4)及(11)得

$$|x^{k+1}-x^*|=|x^k-x^*-f'(x^k)B^{-k}|=$$
$$\left|\frac{f''(x_k)(x^k-x^*)-f'(x_k)-[f''(x^k)-B^k](x^k-x^*)}{B^k}\right|=$$
$$\left|\frac{f'''(\theta)(x^k-x^*)^2-[f''(x_k)-B^k](x^k-x^*)}{B^k}\right|\leqslant$$
$$\frac{4|f'''(x^*)|+2M_4/M_1}{|f''(x^*)|}(x^k-x^*)^2 \tag{12}$$

其中,θ 是 x^k 与 x^* 之间的数,且式(11)中最后1个不等式只需注意到对于充分接近 x^* 的 x,有 $|f'''(x)|<2|f'''(x^*)|$ 及 $|f''(x)|>|f''(x^*)|/2$ 即可。令 $M_3\geqslant M_2$,且

$$M_3\geqslant\frac{4|f'''(x^*)|+2M_4/M_1}{|f''(x^*)|}$$

再结合式(8)和(12),可知$\{x^k\}$收敛于 x^*,且式(7)成立。

注 定理中条件 $f'''(x^*)\neq0$ 非常弱,是牛顿法精确2阶收敛的充要条件。利用引理1,可类似地证明文献[3]和[4]中的混合算法在此条件下也是Q-2阶收敛。

参 考 文 献

1 Ortega J M, Rheinbolt W C. Iterative Solution of Nonlinear Equations in Several Variables. London:Academic Press,1970. 299~357

2 王慧娟,袁亚湘.一维优化的一个二阶收敛算法.运筹学杂志,1992,11(2):1~11

3 王兆智.牛顿与近似牛顿混合算法的收敛速率.中国农业大学学报,1997,2(2):25~28

4 林正华.牛顿与二阶拟牛顿混合迭代算法.高等学校计算数学学报,1994,16(3):217~224