DOI: 10.13656/j.cnki.gxkx.2010.01.011
广西科学 Guangxi Sciences 2010, 17(1): 36~ 38, 42

一个新的解无约束优化问题的信赖域算法*

A New Trust Region Algorithm for Unconstrained Optimization

朱光军[1],韦增欣[1],陆 莎[1,2]
ZHU Guang-jun[1], WEI Zeng-xin[1], LU Sha[1,2]

(1.广西大学数学与信息科学学院,广西南宁 530004; 2.广西师范学院数学与计算机科学系,广西南宁 530001)

(1. College of Mathematics and Information Science, Guangxi University, Nanning, Guangxi, 530004, China; 2. Department of Mathematics and Computer Science, Guangxi Teachers Education University, Nanning, Guangxi, 530001, China)

摘要:为了减少求解信赖域子问题的次数,通过对当前目标函数下降量与成功迭代的目标函数下降量最小值的比较,提出一个新的解无约束优化问题的信赖域算法,证明了该算法的全局收敛性,并用数值实验说明新算法是有效的.

关键词:无约束优化 信赖域算法 全局收敛性
中图法分类号: O221.2 文献标识码: A 文章编号: 1005-9164(2010)01-0036-03

Abstract In order to reduce the number of solving the trust-region subproblem, a new trust region algorithm is proposed by comparing the current reduction of the objective function with the minimum reduction of the objective function for all successful iterations. the proof of the global convergence of the algorithm is given. Numerical tests show the algorithm is effective.

Key words unconstrained optimization, trust-region algorithm, global convergence

本文考虑一般的无约束极小化问题

$$\min_{x\in R^n} f(x), \qquad (0.1)$$

其中,$f: R^n \to R$ 二次连续可微.

信赖域方法是解问题(0.1)的一类有效方法,它包括基本信赖域算法(BTR)[1],非单调算法[2]和过滤子算法[3,4]等,其中非单调算法和过滤子算法都是在BTR算法上发展起来的.

在BTR算法中,迭代步长 s_k 通过解如下信赖域子问题得到,

$$\min m(x_k+s_k) = m(x_k) + g(x_k)^T s_k + \frac{1}{2} s_k^T H_k s_k,$$

$$\text{s.t.} \| s_k \| \leqslant \Delta_k, \qquad (0.2)$$

其中 $m(x_k) = f(x_k)$, $g(x_k) = \nabla f(x_k)$, H_k 是 $f(x)$ 在 x_k 处的海色阵 $\nabla_{xx} f(x_k)$ 的一个对称近似阵, $\|\cdot\|$ 为欧氏模,Δ_k 为信赖域半径,称点 $x_k^+ = x_k + s_k$ 为试验点.

记 $d_k = \dfrac{f(x_k) - f(x_k+s_k)}{m(x_k) - m(x_k+s_k)}$,对于 $0 < Z < 1$,若 $d_k \geqslant Z$,则称迭代成功,称试验点 $x_k^+ = x_k + s_k$ 被接受,令 $x_{k+1} = x_k^+$;若 $d_k < Z$,则称迭代失败.此时在BTR算法中通常是重解子问题(0.2).

在被称为失败的迭代点中可能存在对极小化目标函数 $f(x)$ 有用的点,如何加大这些点被接受的可能性值得探讨.当试验点 x_k^+ 失败时,在文献[3,4]中利用过滤技术,在文献[5,6]中利用线搜索,产生下一个迭代点,得到了较为有效的算法.与利用过滤技术和线搜索技术不同的是,本文通过分析 d_k、目标函数下降量和模型下降量的状态,寻找加大试验点被接受的可能性途径.

1 解无约束优化问题的新算法

记 $W_k = f(x_k) - f(x_k+s_k)$, $e_k = m_k(x_k) - m_k(x_k+s_k)$,分别称为目标函数的下降量和模型下降

收稿日期: 2009-04-30
作者简介:朱光军(1965-),男,讲师,主要从事优化与管理研究。
* 国家自然基金项目(10761001)资助。

量.注意到当 $d_k = W_k / e_k \geqslant Z$时,可能出现 W_k, e_k 都很小,而当 $d_k < Z$时 W_k有可能较大,它可能比“成功”迭代时的目标函数的下降量要大,但此时却被认为是“失败”的,这显然是不合理的.我们对此情形给出改进的方法.

在 BTR算法中,仅当 $d_k \geqslant Z$时,试验点 $x_k^+ = x_k + s_k$ 才被接受,为了提高试验点 x_k^+ 被接受的可能性,我们提如下处理方法:

(1)当 $d_k \geqslant Z$时,试验点 $x_k^+ = x_k + s_k$ 被接受;

(2)当 $d_k < Z$时,记 $\_k$ 为前 k次迭代中成功迭代的目标函数下降量的最小值,即

$$\_k = \min_{1 \leqslant i \leqslant k-1} \{W_i \mid d_i = \frac{W_i}{e_i} \geqslant Z\}, \qquad (1.1)$$

若 $W_k = f(x_k) - f(x_k + s_k) \geqslant \_k$,则接受试验点 $x_k^+ = x_k + s_k$.

根据上述的处理方法,我们对 BTR算法进行修改,给出如下新算法:

算法 1

步骤 0 取初始值 $x_0 \in R^n$,初始信赖域半径 $\Delta_0 > 0$,常数 Z和 V_1, V_2, V_3 分别满足条件 $0 < Z < 1$, $0 < V_1 \leqslant V_2 < 1 \leqslant V_3$,计算 $f(x_0)$,记 $k = 0$,取 $_0$为足够大的常数.

步骤 1 计算 $g(x_k)$,若满足终止条件,则终止.

步骤 2 解信赖域子问题(0.2),得到 s_k,记 $x_k^+ = x_k + s_k$.

步骤 3 记 $W_k = f(x_k) - f(x_k^+)$,计算 $d_k = \frac{W_k}{m(x_k) - m(x_k^+)}$.

步骤 4 确定试验点的可接授性:

若 $d_k \geqslant Z$,则(i) $x_{k+1} = x_k^+$;(ii)若 $W_k < \_k$ 则 $\_k = W_k$;

若 $d_k < Z$, $W_k \geqslant \_k$,则 $x_{k+1} = x_k^+$;否则 $x_{k+1} = x_k$.

步骤 5 更新信赖域半径,取

$$\Delta_{k+1} \begin{cases} \in [V_1\Delta_k, V_2\Delta_k], \text{ if } d_k < Z; \\ = \Delta_k, \text{ if } d_k \geqslant Z \text{ and } \|s_k\| < \Delta_k; \\ \in [\Delta_k, V_3\Delta_k], \text{ if } d_k \geqslant Z \text{ and } \|s_k\| = \Delta_k. \end{cases}$$

产生对称矩阵 H_{k+1},记 $k = k + 1$,转步骤 1.

2 算法的全局收敛性

为证明算法 1收敛性的需要,作如下假定:

A1 f 在 R^n 上二次连续可微.

A2 算法 1的所有迭代点 x_k都在 R^n中的一个有界闭区域内.

A3 对所有的 k, H_k一致有界,即存在 $k_{umb} > 0$,使得 $\|H_k\| \leqslant k_{umb}$.

设 $j = j_c$是满足

$$x_k(j) = x_k - k_{bck}^j \frac{\Delta_k}{\|g_k\|} g_k, \; m_k(x_k(j)) \leqslant m_k(x_k) + k_{ubs} g_k^T (x_k(j) - x_k)$$

的最小正整数,其中 $k_{bck} \in (0,1)$, $k_{ubs} \in (0, \frac{1}{2})$ 是常数.记 $x_k^{AC} = x_k(j_c)$,称 x_k^{AC}为信赖域子问题(0.2)的近似柯西点.对于 x_k^{AC}存在如下结论:

引理 2.1 设假定 A1成立, x_k^{AC}是信赖域子问题(0.2)的近似柯西点,则

$$m_k(x_k) - m_k(x_k^{AC}) \geqslant k_{dcp}\|g_k\| \min[\frac{\|g_k\|}{\|H_k\|}, \Delta_k], \qquad (2.1)$$

其中 $k_{dcp} \in (0,1)$是常数.

证明 参见文献[1]中定理 6.3.3的证明.

用 $\lambda_{\min}[H_k]$表示 H_k 的最小特征值,当 $f_k = \lambda_{\min}[H_k] < 0$时,假设 u_k满足条件:

$$u_k^T g_k \leqslant 0, \|u_k\| = \Delta_k, u_k^T H_k u_k \leqslant k_{snc} f_k \Delta_k^2, k_{snc} \in (0,1]. \qquad (2.2)$$

设 $j = j_c$是满足条件

$$t_j = k_{bck}^j, x_k(j) = x_k + t_j u_k, m_k(x_k(j)) \leqslant m_k(x_k) + k_{ubc} f_k t_j^2 \|u_k\|^2$$

的最小正整数,其中 $b_{bck} \in (0,1)$, $k_{ubc} \in (0, \frac{1}{2}k_{snc})$是常数.记 $t_k^{AE} = t_{j_c}$, $x_k^{AE} = x(t_k^{AE}) = x_k + t_k^{AE} u_k$,称 x_k^{AE}为信赖域子问题(0.2)的近似特征点.

引理 2.2 设假定 A1成立, $f_k = \lambda_{\min}[H_k] < 0$, u_k满足(2.2)式, x_k^{AE}为问题(0.2)的近似特征点,则

$$m_k(x_k) - m_k(x_k^{AE}) \geqslant -k_{sod} f_k \min[f_k^2, \Delta_k^2], \qquad (2.3)$$

其中 $k_{sod} > 0$是常数.

证明 参见文献[1]中定理 6.6.2的证明.

如同文献[1]中的(6.6.10)式,可要求信赖域子问题(0.2)的解 s_k满足条件:

$$m_k(x_k) - m_k(x_k + s_k) \geqslant k_{umd}(m_k(x_k) - \min[m_k(x_k^{AE}), m_k(x_k^{AE})]) \qquad (2.4)$$

其中 $k_{umd} \in (0,1]$是常数.

结合(2.1)式,(2.3)式和(2.4)式,得到问题(0.2)的解 s_k满足如下充分下降性条件:

$$m_k(x_k) - m_k(x_k + s_k) \geqslant k_{umd}\max\{k_{dcp}\|g_k\| \min[\frac{\|g_k\|}{\|H_k\|}, \Delta_k], -k_{sod} f_k \min[f_k^2, \Delta_k^2]\}. \qquad (2.5)$$

引理 2.3 设假定 A1, A2, A3成立,若存在常数 $k_{lbg} > 0$,对全部的 k满足 $\|g_k\| \geqslant k_{lbg}$,则存在常数 $k_{lbd} > 0$,使得 $\Delta_k \geqslant k_{lbd}$.

证明 参见文献[1]中定理 6.4.3的证明.

引进记号: $S = \{k \mid x_{k+1} = x_k + s_k\}$, $C = \{k \mid d_k <$

$Z, W_k \geqslant \_ _k \}$.

定理 **2.4** 设假定 A1、A2、A3成立,信赖域子问题(0.2)的解 s_k 满足(2.5)式,且算法 1只有有限步成功迭代,则必存在某个 k,使得 $x_k = x^*$, $g(x^*) = 0$.

证明 设 k_0 是最后一步成功迭代,则对任意的 $j > 0, x^* = x_{k_0+1} = x_{k_0+j}$,有 $d_{k_0+j} < Z$.由算法 1的步骤 5得: $0 < \Delta_{k_0+j} \leqslant V_2 \Delta_{k_0}$,可以推出

$$\lim_{k \to \infty} \Delta_k = 0. \qquad (2.6)$$

反之,若存在正数 X,对全部的 j 满足 $\| g_{k_0+j} \| \geqslant X > 0$,根据引理 2.3得到 $\Delta_{k_0+j} \geqslant k_{lbd}$,这与(2.6)矛盾.故存在 $\| g_{k_0+j} \| = 0$.

定理 **2.5** 设假定 A1、A2、A3成立,且集合 C 的元素个数 $|C| = +\infty$,则

$$\lim_{k \to \infty} \inf \| g_k \| = 0. \qquad (2.7)$$

证明 设存在 $X > 0$,对任意的 k 下列不等式成立,

$$\| g_k \| \geqslant X. \qquad (2.8)$$

记集合 C 中所有指标为 $\{k_i\}$,则由(1.1)式可知

$$W_i = f(x_{k_i}) - f(x_{k_i+1}) \geqslant \_ _{k_i}. \qquad (2.9)$$

设在前 k_i 次成功迭代中的第 p_i 次迭代的目标函数的下降量最小,即

$$\_ _{k_i} = W_{p_i} = \min_{\leqslant \leqslant k_i - 1} \{ W \mid d = \frac{W}{e_i} \geqslant Z \}. \qquad (2.10)$$

结合(2.5)式,(2.8)式,(2.9)式和(2.10)式有

$f(x_{k_i}) - f(x_{k_i+1}) = f(x_{k_i}) - f(x_{k_i+1}) + f(x_{k_i+1}) - f(x_{k_i+2}) + \cdots + f(x_{k_{i+1}-1}) - f(x_{k_i+1}) = W_i + W_{i+1} + \cdots W_{k_{i+1}-1} \geqslant W_i \geqslant \_ _{k_i} = W_{p_i} \geqslant Z[m_{p_i}(x_{p_i}) - m_{p_i}(x_{p_i} + s_{p_i})] \geqslant Z k_{umd} k_{dcp} \| g_{p_i} \| \cdot \min[\frac{\| g_{p_i} \|}{\| H_{p_i} \|}, \Delta_{p_i}] \geqslant Z k_{umd} k_{dcp} X \min[\frac{X}{k_{umd}}, k_{lbd}]$, (2.11)

由(2.11)式得到

$$f(x_{k_1}) - f(x_{k_{i+1}}) = \sum_{j=1}^{i} [f(x_{k_j}) - f(x_{k_{j+1}})] \geqslant \sum_{j=1}^{i} Z k_{umd} k_{dcp} X \min[\frac{X}{k_{umh}}, k_{lbd}] \to +\infty,$$

从而 $\lim_{i \to \infty} f(x_{k_i}) = -\infty$,即 $f(x)$ 无下界,这与假定 A1、A2矛盾.定理 2.5得证.

定理 **2.6** 设假定 A1、A2、A3成立,且 $|S| = +\infty$, $|C| < +\infty$,则

$$\lim_{k \to \infty} \inf \| g_k \| = 0.$$

证明 设存在 $X > 0$,对任意的 k 有下列不等式成立,

$$\| g_k \| \geqslant X.$$

由 $|S| = +\infty$, $|C| < +\infty$,存在充分大的 k_0,当 $k > k_0$ 时,有 $k \in S, k \notin C$,则

$f(x_k) - f(x_{k+1}) \geqslant Z[m_k(x_k) - m_k(x_{k+1})] \geqslant Z k_{umd} k_{dcp} \| g_k \| \min[\frac{\| g_k \|}{\| H_k \|}, \Delta_k] \geqslant Z k_{umd} k_{dcp} X \min[\frac{X}{k_{umh}}, k_{lbd}]$, (2.12)

由(2.12)式得到

$$f(x_{k_0}) - f(x_{k_{k+1}}) = \sum_{j=k_0}^{k} [f(x_j) - f(x_{j+1})] \geqslant (k - k_0) Z k_{ump} k_{dcp} X \min[\frac{X}{k_{umh}}, k_{lbd}] \to +\infty.$$

这表明 $f(x)$ 无下界,与假定 A1和 A2矛盾.定理 2.6得证.

由定理 2.4,定理 2.5,定理 2.6知,算法 1产生的点列 $\{x_k\}$ 中或者存在一个一阶稳定点或者存在一个聚点是一阶稳定点.

下面考虑算法 1收敛子二阶稳定点的条件.增加假定:

A4 当 $\lim_{k \to \infty} \| g_k \| = 0$ 时,

$$\lim_{k \to \infty} \| \nabla_{xx} f(x_k) - H_k \| = 0.$$

引理 **2.7** 设假定 A1~A4成立,对充分大的 k,存在常数 $k_{mqd} > 0$,有 $m_k(x_k) - m_k(x_k + s_k) \geqslant k_{mqd} \| s_k \|^2$,再设 $\lim_{k \to \infty} \| s_k \| = 0$,则存在 $k_0 > 0, T > 0$,当 $k > k_0$ 时,$d_k > Z, \Delta_k > T$.

证明 参见文献[1]中引理 6.5.4的证明.

定理 **2.8** 设假定 A1~A4成立,x^* 是算法 1产生的点列 $\{x_k\}$ 的唯一聚点,则 x^* 是二阶稳定点.

证明 参见文献[3]中定理 3.7的证明.

3 算法的数值实验

用 Matlab编程来实现算法 1,参数选择如下: $Z = 0.25, \Delta_0 = 0.5$,设定的终止准则是 $\| \nabla f(x_k) \| \leqslant 10^{-6}$.程序的运行过程中,利用 BFGS公式修正得到 H_{k+1},试验的结果见表 1,其中 N是变量的维数,NIT为迭代的次数,NF为目标函数的计算次数,NG为目标函数梯度的计算次数.

从表 1的数值结果看,对于 Extended Rosenbrock问题,当 N为 128和 2048,算法 1比 BTR有所改善,对于 Extended penalty1问题,当 N为 32,124,512和 1024,算法 1在迭代的次数,目标函数和梯度的计算次数比 BTR小,因此算法 1是有效的.

(下转第 42页 Continue on page 42)

$$\sum_{n=1}^{\infty} k_n^{U} P(\sum_{i=1}^{k_n} a_{ni}^{+} X_{ni} > X) < \infty, \forall X > 0, \quad (2.11)$$

$$\sum_{n=1}^{\infty} k_n^{U} P(\sum_{i=1}^{k_n} a_{ni}^{+} X_{ni} < -X) < \infty, \forall X > 0. \quad (2.12)$$

首先证明(2.11)式.由引理1(iii)得,$\{a_{ni}^{k}X_{ni}; 1\leqslant i\leqslant k_n, n\geqslant 1\}$仍然是两两NQD随机阵列,又由Markov不等式以及引理1的(i),定理2的条件$\sum_{i=1}^{k_n} a_{ni}^2 = O(\frac{1}{k_n^{U+2}})$,得

$$\sum_{n=1}^{\infty} k_n^{U} P(\sum_{i=1}^{k_n} a_{ni}^{+} X_{ni} > X) \leqslant$$

$$\sum_{n=1}^{\infty} k_n^{U} X^{2} E(\sum_{i=1}^{k_n} a_{ni}^{+} X_{ni})^2 \leqslant \sum_{n=1}^{\infty} k_n^{U} X^{-2} \sum_{i=1}^{k_n} (a_{ni}^{+})^2 EX_{ni}^2 \leqslant$$

$$c\sum_{n=1}^{\infty} k_n^{U} \sum_{i=1}^{k_n} (a_{ni}^{+})^2 \leqslant c\sum_{n=1}^{\infty} k_n^{U} \sum_{i=1}^{k_n} a_{ni}^2 \leqslant c\sum_{n=1}^{\infty} k_n^{U} \frac{1}{k_n^{U+2}} =$$

$$c\sum_{n=1}^{\infty} \frac{1}{k_n^2} \leqslant c\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

将上面证明过程中的X_{ni}用$-X_{ni}$替换,又由引理1的(iii)知,$\{a_{ni}^{+}(-X_{ni}); 1\leqslant i\leqslant k, n\geqslant 1\}$仍然是两两NQD随机阵列,则对$\forall X > 0$,有

$$\sum_{n=1}^{\infty} k_n^{U} P(\sum_{i=1}^{k_n} a_{ni}^{+} X_{ni} < -X) < \infty.$$

由此,(2.9)式得到证明.综上所述,定理2证明完毕.

注 2.1 设$\{k_n; n\geqslant 1\}$是正整数序列,$\{X_{ni}; 1\leqslant i\leqslant k_n, n\geqslant 1\}$是两两NQD随机阵列,且$EX_{ni}=0$,存在某一随机变量$X$和正常数$c$,使得对所有的$1\leqslant i\leqslant k_n, n\geqslant 1$和$x>0$,有$P(|X_{ni}|>x)\leqslant cP(|X|>x)$,$EX^2<\infty$.假设$\{b_{ni}; 1\leqslant i\leqslant k_n, n\geqslant 1\}$是实数阵列,满足$\lim\sup_{n\to\infty}\sum_{i=1}^{k_n} b_{ni}^2 < \infty$,则

$$\sum_{n=1}^{\infty} k_n^{U} P(|\sum_{i=1}^{k_n} b_{ni} X_{ni}| > X\log k_n) < \infty, \forall X > 0.$$

证明　令$a_{ni}=\frac{b_{ni}}{\log k_n}$,由定理2的证明过程,即可得以证明.

参考文献:

[1] Lehmann E L. Some concepts of dependence[J]. Ann Math Stat, 1966, 43: 1137-1153.

[2] Matula P. A note on the almost sure convergence of sums of negatively dependent random variables[J]. Statist Probab Lett, 1992, 15: 209-213.

[3] 吴群英.两两NQD列的收敛性质[J].数学学报,2002,45(3): 617-624.

[4] Baek J I, Park S T. Convergence of weighted sums for arrays of negatively dependent random variables and its applications[J]. Theoretical Probability, 2008, 37(1): 73-80.

[5] 吴群英.混合序列的概率极限理论[M].北京: 科学出版社,2006.

(责任编辑: 韦廷宗)

(上接第38页 Continue from page 38)

表 1　数值实验结果

Table 1　Nunerical results

Problem	N	BTR (NIT/NF/NG)	Algorithm 1 (NIT/NF/NG)
Extended Rosenbrock	32	42/43/43	42/43/43
	64	43/44/44	43/44/44
	128	46/47/47	42/43/43
	256	49/50/50	49/50/50
	512	46/47/47	46/47/47
	1024	52/53/53	52/53/53
	2048	57/58/58	51/52/52
Extended penaltyl	32	141/142/142	122/123/123
	64	116/117/117	116/117/117
	128	124/125/125	120/121/121
	256	122/123/123	122/123/123
	512	133/134/134	132/133/133
	1024	160/161/161	155/156/156
	2048	173/174/174	173/174/174

参考文献:

[1] Conn A R, Gould N I M, Toint Ph L. Trust region methods[M]. Philadelphia: SIAM, 2000.

[2] 李正峰,邓乃扬.一类新的非单调信赖域算法及其收敛性[J].应用数学学报,1999,22(3): 458-465.

[3] 缪卫华,孙文瑜.一个解无约束优化问题的过滤信赖域方法[J].高等学校计算数学学报,2007(3): 88-89.

[4] Gould N I M, Sainvitu C, Toint Ph L. A filter-trust-region method for unconstrained Optimization[J]. SIAM Journal on Optimization, 2006, 16(2): 341-357.

[5] Mo Jiangtao, Zhang Kecun, Wei Zengxin. A nonmonotone trust region method for unconstrained optimization[J]. Applied Mathematics and Computation, 2005, 171: 371-384.

[6] 姚升保,施保昌,彭叶辉.一类带线搜索的非单调信赖域算法[J].数学杂志,2003,23(3): 290-294.

(责任编辑: 韦廷宗)