

一个新的 BFGS 信赖域算法

景书杰, 于俊霞

(河南理工大学数学与信息科学学院, 河南 焦作 454000)

摘 要: 本文对于无约束最优化问题提出了一个新的 BFGS 信赖域算法. 利用 BFGS 方法和信赖域方法, 提出了改进的 BFGS 信赖域方法. 推广了文献 [3, 5] 中的两种算法, 得到一个新的 BFGS 信赖域算法, 在适当条件下证明了算法的全局收敛性.

关键词: BFGS 修正公式; 信赖域算法; 全局收敛性

MR(2010) 主题分类号: 90C30

中图分类号: O221.2

文献标识码: A

文章编号: 0255-7797(2015)01-0131-04

1 引言

考虑无约束优化问题

$$\min_{x \in R^n} f(x), \quad (1.1)$$

其中 $f(x): R^n \rightarrow R$ 是连续可微函数. 信赖域方法是一类数值计算方法, 其具有较好的可靠性和强收敛性, 因此, 受到了许多研究者的关注. 信赖域方法是一种迭代的方法, 其基本思想是, 每次迭代需求解信赖域子问题

$$\min \varphi_k(d) = g_k^T d + \frac{1}{2} d^T B_k d \quad \text{s.t.} \quad \|d\| \leq \Delta_k, \quad (1.2)$$

其中 $g_k = \nabla f(x_k)$; $B_k \in R^{n \times n}$ 是函数 $f(x)$ 在 x_k 处的 Hessian 矩阵或其近似矩阵; $\Delta_k > 0$ 是信赖域半径, $\|\cdot\|$ 是 R^n 中 Euclid 范数. 目前 BFGS 方法是拟牛顿方法中求解无约束最优化问题 (1.1) 的重要方法之一 [4].

针对著名的 BFGS 校正公式:

$$B_{k+1} = B_k + \frac{y_k y_k^T}{s_k^T y_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}, \quad (1.3)$$

其中 $s_k = x_{k+1} - x_k$, $y_k = g_{k+1} - g_k$, g_k 和 g_{k+1} 分别是 $f(x)$ 在 x_k 和 x_{k+1} 出的梯度值. 文献 [1-3] 给出一些修改的 BFGS 方法及其收敛性分析.

文献 [3] 中 Li 和 Fukushima 给出一个修改的 BFGS 校正公式

$$B_{k+1} = B_k + \frac{y_k^* y_k^{*T}}{s_k^T y_k^*} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}, \quad (1.4)$$

*收稿日期: 2013-01-17

接收日期: 2013-03-20

基金项目: 国家自然科学基金 (10671057); 河南理工大学运筹学与控制论重点学科资助项目; 河南省数学一级重点学科资助项目.

作者简介: 景书杰 (1965-), 男, 河南长垣, 教授, 博士, 主要研究方向: 最优化理论与应用. E-mail: yujunxia1980@163.com.

其中 $y_k^* = y_k + t_k \|\nabla f(x_k)\| s_k$, $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$, $t_k = C + \max\{0, -\frac{y_k^T s_k}{\|s_k\|^2}\} \|\nabla f(x_k)\|^{-1}$, $C > 0$ 是一个常数.

文献 [5] 在式 (1.3) 基础上引入可调整的参数 θ : 给出了一个修改的 BFGS 公式

$$B_{k+1} = B_k + \frac{y_k^* y_k^{*T}}{s_k^T y_k^*} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}, \quad (1.5)$$

其中 $y_k^* = \theta y_k + (1 - \theta) t_k \|\nabla f(x_k)\| s_k$, $\theta \in [0, 1]$, $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$,

$$t_k = C + \max\{0, -\frac{y_k^T s_k}{\|s_k\|^2}\} \|\nabla f(x_k)\|^{-1},$$

$C > 0$ 是一个常数.

受上面公式启发. 本文作如下修改:

$$B_{k+1} = B_k + \frac{y_k^* y_k^{*T}}{s_k^T y_k^*} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}, \quad (1.6)$$

这里 $y_k^* = \alpha y_k + \beta t_k \|\nabla f(x_k)\| s_k$, $\alpha, \beta \in [0, 1]$, $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$, $t_k = C + \max\{0, -\frac{y_k^T s_k}{\|s_k\|^2}\} \|\nabla f(x_k)\|^{-1}$, $C > 0$ 是一个常数.

显然, 当 $\alpha = 1, \beta = 0$ 时, 式 (1.6) 退化为式 (1.3); 当 $\alpha = 1, \beta = 1$ 时, 式 (1.6) 退化为式 (1.4); 当 $\alpha + \beta = 1$ 时, 式 (1.6) 退化为式 (1.5).

众所周知, 在式 (1.3) 中当 $y_k^T s_k > 0$ 时, B_{k+1} 继承 B_k 的正定性. 所以当 $y_k^{*T} s_k > 0$ 时, 利用 (1.6) 式更新 B_{k+1} , 反之不更新. 本文借鉴文献 [5–8] 的思想, 将 (1.6) 式与信赖域算法相结合, 给出一个新的 BFGS 信赖域算法, 该算法能够保证修正矩阵的正定性, 同时在一定条件下证明了算法的全局收敛性.

2 算法及性质

通过求解信赖域子问题 (1.2) 得 d_k , 同时定义实际下降量 $Ared_k = f(x_k) - f(x_k + d_k)$; 预估计下降量 $Pred_k = \varphi_k(0) - \varphi_k(d_k)$; 定义比值 $r_k = \frac{Ared_k}{Pred_k}$.

算法 1

步骤 0 取 $\Delta_m > 0$ 为迭代过程中最大的信赖域半径, $x_0 \in R^n$, $B_0 \in R^{n \times n}$ 为对称正定矩阵, $\varepsilon \geq 0$, $0 < \Delta_0 < \Delta_m$, $0 < \eta < 1$, $\lambda_2 > 1 > \lambda_1 > 0$, $k = 0$;

步骤 1 计算 g_k , 若 $\|g_k\| \leq \varepsilon$, 则停止, 否则转步 2;

步骤 2 求解子问题 (1.2) 得 d_k ;

步骤 3 计算 $r_k = \frac{Ared_k}{Pred_k}$. 更新

$$x_{k+1} = \begin{cases} x_k + d_k, & r_k > 0, \\ x_k, & r_k \leq 0, \end{cases} \quad \Delta_{k+1} = \begin{cases} \min\{\lambda_2 \Delta_k, \Delta_m\}, & r_k \geq \eta, \\ \lambda_1 \Delta_k, & r_k < \eta; \end{cases} \quad (2.1)$$

步骤 4 如果 $r_k > 0$, 利用式 (1.6) 修正 B_k 产生 B_{k+1} , 令 $k := k + 1$, 转步 1.

3 算法的全局收敛性

为证明算法的全局收敛性, 现给出如下假设 H

H₁ 水平集 $L(x_0) = \{x | f(x) \leq f(x_0)\}$ 有界;

H₂ 函数 $f(x)$ 在水平集 $L(x_0)$ 上二阶连续可微有下界;

H₃ $\{B_k\}$ 一致有界, 即存在 $M > 0$ 使得对任意的 k 有 $\|B_k\| \leq M$.

引理 3.1^[4] 设 d_k 是信赖域子问题 (1.2) 的解, 则有

$$Pred_k = \varphi_k(0) - \varphi_k(d_k) \geq \frac{1}{2} \|g_k\| \min\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\}.$$

引理 3.2 $y_k^{*T} s_k > 0$.

证 由 (1.6) 式得到

$$\begin{aligned} y_k^{*T} s_k &= (\alpha y_k + \beta t_k \|\nabla f(x_k)\| s_k)^T s_k = \alpha y_k^T s_k + \beta t_k \|\nabla f(x_k)\| s_k^T s_k \\ &= \alpha y_k^T s_k + \beta t_k \|\nabla f(x_k)\| \|s_k\|^2 \geq \beta C \|\nabla f(x_k)\| \|s_k\|^2 > 0. \end{aligned}$$

显然, 当 B_k 正定时, B_{k+1} 也正定.

定理 3.3 (收敛性定理) 若假设 H₁, H₂, H₃ 均成立, 由算法 1 产生的点列为 x_k , 则

$$\lim_{k \rightarrow \infty} \inf \|g_k\| = 0. \quad (3.1)$$

证 由 r_k 定义知

$$|r_k - 1| = \left| \frac{f_k - f(x_k + d_k)}{\varphi_k(0) - \varphi_k(d_k)} - 1 \right| = \left| \frac{\varphi_k(d_k) - f(x_k + d_k) + f_k}{\varphi_k(d_k)} \right|, \quad (3.2)$$

利用 Taylor 展开式 $f(x_k + d_k) = f_k + g_k^T d_k + \int_0^1 d_k^T (g(x_k + \tau d_k) - g_k) d\tau$. 所以在 $\Delta_k > 0$ 充分小时, 有

$$\begin{aligned} |\varphi_k(d_k) - f(x_k + d_k) + f_k| &= \left| \frac{1}{2} d_k^T B_k d_k - \int_0^1 d_k^T (g(x_k + \tau d_k) - g_k) d\tau \right| \\ &\leq \frac{M}{2} \|d_k\|^2 + o(\|d_k\|) \leq \frac{M}{2} \Delta_k^2 + o(\Delta_k). \end{aligned} \quad (3.3)$$

假若式 (3.1) 不成立, 则存在 $\varepsilon_0 > 0$ 和 k_0 使得 $\forall k \geq k_0$, 有 $\|g_k\| \geq \varepsilon_0$, 由引理 3.1, 对 $k \geq k_0$, 有

$$\varphi_k(0) - \varphi_k(d_k) = -\varphi_k(d_k) \geq \frac{1}{2} \|g_k\| \min\left\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\right\} \geq \frac{1}{2} \varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{M}\right\}, \quad (3.4)$$

利用 (3.2)–(3.4) 式得

$$|r_k - 1| \leq \frac{2(\frac{M}{2} \Delta_k^2 + o(\Delta_k))}{\varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{M}\right\}} = \frac{M_0 \Delta_k^2 + 2o(\Delta_k)}{\varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{M}\right\}} = \frac{M_0 \Delta_k^2 + o(2\Delta_k)}{\varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{M}\right\}}. \quad (3.5)$$

假若存在自然数列 N 的一个无穷子列 N_0 , 使得对 $\forall k \in N_0$, 有 $r_k \geq \eta$, 则对 $k \in N_0, k \geq k_0$, 由 (3.4) 及算法 1 步 3 中的 (2.1) 式得

$$f_k - f_{k+1} \geq \eta[\varphi_k(0) - \varphi_k(d_k)] \geq \frac{1}{2} \eta \varepsilon_0 \min\left\{\Delta_k, \frac{\varepsilon_0}{M}\right\}. \quad (3.6)$$

由数列 $\{f(x_k)\}$ 单调不增有下界. 故它有极限, 因此有 $\lim_{k \rightarrow \infty} (f_k - f_{k+1}) = 0$, 从而对充分大的 k , 由 (3.6) 式有

$$\lim_{k \rightarrow \infty} \Delta_k = 0. \quad (3.7)$$

又由 (3.5) 式, 有 $\lim_{k \rightarrow \infty} r_k = 1$. 据算法 1, 对充分大的 k 应有 $\Delta_{k+1} = \min\{\lambda_2 \Delta_k, \Delta_m\} \geq \Delta_k$, 这与式 (3.7) 式矛盾, 因此, 对充分大的 k , $r_k < \eta$, 在这种情况下, Δ_k 将以 λ_1 的比例压缩, 又得到 $\lim_{k \rightarrow \infty} \Delta_k = 0$. 这又与由 (3.5) 得出的 $r_k \rightarrow 1$ 时, $\Delta_{k+1} \geq \Delta_k$ 矛盾. 所以前面的假设成立, 定理结论得证.

参 考 文 献

- [1] Powell M J D. A new alogorithm for unconstrained optimation [A]. Rose J B, Mangasarian OL, Ritter k, et al. Nonlinear progamming [C]. New York: Academic Press, 1970.
- [2] Wei Z, Yu G, Yuan G, et al. The superlinear convergence of a modified BFGS-type method for unconstrainedoptimization [J]. Computational Optimization and Applications, 2004, 29: 315–332.
- [3] LI D, Fukushima M. A modified BFGS method and its global convergence in non-converminimization [J]. Journal of Computational and Applied Mathematics, 2002, 129: 15–35.
- [4] 袁亚湘, 孙文瑜. 最优化理论与方法 [M]. 北京: 科学出版社, 1997.
- [5] 景书杰, 李少娟. 一个改进的 BFGS 信赖域算法及收敛性分析 [J]. 河南理工大学学报 (自然科学版), 2012, 113(4): 1673–9787.
- [6] 袁功林, 吴燕林, 韦增欣. 一个改进的 BFGS 信赖域算法 [J]. 广西科学, 2009, 16(4): 397–399.
- [7] 吴红梅. 无约束优化问题的一个改进的 BFGS 信赖域算法 [J]. 西安工业大学学报, 2009, 299(3): 1673–9965.
- [8] 袁亚湘. 信赖域算法的收敛性 [J]. 计算数学, 1994, 16: 333–346.
- [9] 王宜举, 修乃华. 非线性规划理论与方法 [M]. 西安: 陕西科学技术出版社, 2008.

A NEW BFGS RUST REGION METHOD

JING Shu-jie , YU Jun-xia

(*Institute of Mathematics and Information Science, Henan Polytechnic University,
Jiaozuo 454000, China*)

Abstract: A new BFGS rust region algorithm for unconstrained optimization problems is presented. By using BFGS method and trust region method, a class of new BFGS trust region algorithm that reference [3] and [5] promoted was obtained, and the global convergence of the algorithm is proved under appropriate conditions.

Keywords: modified BFGS formula; trust region algorithm; global convergence

2010 MR Subject Classification: 90C30