Vol. 34 (2014)
No. 3

数 学 杂 志
J. of Math. (PRC)

一个新的 MBFGS 信赖域算法

景书杰 [1]，苗 荣 [1]，李少娟 [2]
(1. 河南理工大学数学与信息科学学院, 河南 焦作 454000)
(2. 河南护理职业学院, 河南 安阳 455000)

摘 要: 本文研究了无约束最优化问题. 利用 MBFGS 信赖域算法的基本思想, 通过对 BFGS 校正公式的改进, 并结合线搜索技术, 提出了一种新的 MBFGS 信赖域算法, 拓宽了信赖域算法的适用范围, 并在一定条件下证明了该算法的全局收敛性和超线性收敛性.

关 键 词: 无约束最优化; 信赖域算法; BFGS(MBFGS) 方法; 线搜索
MR(2010) **主题分类号**: 90C30 **中图分类号**: O224
文献标识码: A **文章编号**: 0255-7797(2014)03-0569-08

1 引言

考虑无约束最优化问题

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

其中 $f(x):\mathbf{R^n}\longrightarrow\mathbf{R}$ 是二次连续可微函数. 信赖域算法 [1–3] 是求解问题 (1.1) 的一个重要方法. 其基本思想如下：在每一个迭代点, 试探步 s_k 一般是子问题：

$$\min_{\|s\|\le\Delta_k} q_k(s)=g_k^Ts+\frac{1}{2}s^TB_ks \tag{1.2}$$

的解. 其中 $g_k=\nabla f(x_k)$,$B_k\in R^{n\times n}$ 是 f 在 x_k 处的 Hessian 矩阵 $G(x)$ 或其近似矩阵, Δ_k 是信赖域半径,$\|\cdot\|$ 是 R^n 中的 Euclid 范数.

众所周知, BFGS 方法是拟牛顿方法中解问题 (1.1) 的重要方法之一, 其校正公式是：

$$B_{k+1}=B_k-\frac{B_ks_ks_k^TB_k}{s_k^TB_ks_k}+\frac{y_ky_k^T}{s_k^Ty_k}, \tag{1.3}$$

其中 $s_k=x_{k+1}-x_k$,$y_k=g_{k+1}-g_k$;g_{k+1} 和 g_k 分别是 $f(x)$ 在 x_{k+1} 和 x_k 处的梯度值. 文献 [4–8] 给出了一些修改的 BFGS 方法, 并分析了其收敛性. 韦等 [4,5] 给出了新的公式 (M_3BFGS)：

$$B_{k+1}=B_k-\frac{B_ks_ks_k^TB_k}{s_k^TB_ks_k}+\frac{y_k^ly_k^{lT}}{s_k^Ty_k^l},$$

其中 $y_k^l=y_k+A_k(2)s_k$,$y_k=g_{k+1}-g_k$, $A_k(2)=\frac{[2(f_k-f_{k+1})+(g_{k+1}+g_k)^Ts_k]I}{\|s_k\|^2}$, $s_k=x_{k+1}-x_k$.

*收稿日期: 2012-06-04 接收日期: 2012-09-14
基金项目：国家自然科学基金项目 (10671057); 河南理工大学运筹学与控制论重点学科资助项目 (10671057).
作者简介：景书杰 (1965–), 男, 河南长恒, 副教授, 研究方向：最优化理论与应用.

受到上面公示的启发, 我们做如下假设:

$$B_{k+1}=B_k-\frac{B_ks_ks_k^TB_k}{s_k^TB_ks_k}+\frac{y_k^*y_k^{*T}}{s_k^Ty_k^*}, \tag{1.4}$$

其中 $y_k^*=\frac{\delta_k^Ts_k}{|\delta_k^Ts_k|}\delta_k$,$\delta_k=\theta y_k+(1-\theta)A_k(2)s_k$, $A_k(2)=\frac{[2(f_k-f_{k+1})+(g_{k+1}+g_k)^Ts_k]I}{\|s_k\|^2}$, $\theta\in[0,1]$, $y_k=g_{k+1}-g_k$. 当且仅当 $s_k^Ty_k>0$ 时, B_{k+1} 继承 B_k 的正定性. 由 y_k^* 的定义, 有

$$s_k^Ty_k^*=\frac{(\delta_k^Ts_k)^2}{|\delta_k^Ts_k|}=|\delta_k^Ts_k|=|s_k^TB_ks_k+o(\|s_k\|^2)|,$$

若 B_k 正定, 当 $\|s_k\|$ 充分小时, 则有

$$|s_k^TB_ks_k+o(\|s_k\|^2)|>0, \tag{1.5}$$

因此由公式 (1.4) 所得的 B_{k+1} 总具有正定性.

本文在文献 [4–10] 的基础上, 给出了一个改进的带线搜索的信赖域方法, 其中 B_k 由公式 (1.4) 产生. 在一定条件下证明了该方法的全局收敛性和超线性收敛性.

2 算法

算法 1(MBFGS 信赖域算法)

步 0: 给定常数 $\eta\in(0,1)$, $\beta\in(0,1)$, $r_1\in(0,1)$, $\rho\in(0,1)$, $\varepsilon>0$, $0<\Delta_{\min}<\Delta_0$, $r_2>1$, $x_0\in R^n$, $B_0\in R^{n\times n}$ 是对称正定矩阵, $k:=0$;

步 1: 计算 g_k, 若 $\|g_k\|\le\varepsilon$, 则停; 否则, 转步 2;

步 2: 求子问题 (1.2) 的解 s_k;

步 3: 计算 $r_k=\frac{f(x_k)-f(x_k+s_k)}{q_k(0)-q_k(s_k)}$, 若 $r_k\ge\eta$, 令 $\alpha_k=1$, $\Delta_k=\max\{r_2\Delta_k,\Delta_{\min}\}$, 转步 5: 否则, 转步 4;

步 4: 若 $r_k<\eta$, 取 α_k 为 $\{1,\beta,\beta^2,\cdots\}$ 中满足下式的最大数:

$$f(x_k+\alpha s_k)-f_k\le\rho\alpha g_k^Ts_k. \tag{2.1}$$

令 $\Delta_k=\max\{r_1\Delta_k,\Delta_{\min}\}$, 转步 5;

步 5: 计算下一个迭代点 $x_{k+1}=x_k+\alpha_ks_k$, 利用公式 (1.4) 修正 B_{k+1}, 令 $k:=k+1$, 转步 1.

由于 $MBFGS$ 方法保证 $y_k^{*T}s_k>0$, 所以 B_k 总是对称正定的. 因此子问题 (1.2) 有唯一解 s_k. 由约束条件的最优性条件知, 存在乘子 $\alpha_k\ge0$, 使得

$$\begin{cases}B_ks_k+\alpha_ks_k=-g_k,\\ \alpha_k(\|s_k\|-\Delta_k)=0.\end{cases} \tag{2.2}$$

由此式可知,

$$\alpha_k=\frac{-g_k^Ts_k-s_k^TB_ks_k}{\|s_k\|^2}, \tag{2.3}$$

即 $g_k^T s_k + s_k^T B_k s_k = -\alpha_k \|s_k\|^2 \le 0$, 故有

$$q_k(s_k) = g_k^T s_k + \frac{1}{2} s_k^T B_k s_k \le \frac{1}{2} g_k^T s_k \le -\frac{1}{2} s_k^T B_k s_k \le 0. \tag{2.4}$$

3 算法的全局收敛性

为了证明算法 1 的全局收敛性, 我们给出如下假设条件:

假设 (i):

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

B) f 在 $L(x_0)$ 上连续可微, 其导数满足 Lipschitz 条件, 即存在一个常数 $L > 0$, 使得

$$\|g(x) - g(y)\| \le L\|x - y\|, \forall x, y \in L(x_0). \tag{3.1}$$

由线搜索和公式 (2.4) 知, 该算法为下降算法. 因此由假设 A) 知, 存在一个常数 f^*, 满足 $\lim\limits_{k\to\infty} f(x_k) = f^*$. 根据 $\{x_k\}$ 有界这一事实, 利用假设 B), 知存在一个常数 M_0, 使得当 k 充分大时, 有 $\|g_k\| \le M_0$.

引理 3.1 [5] 令序列 B_k 由标准的 BFGS 公式产生, 其中 B_0 是对称正定矩阵. 若存在常数 $M > m > 0$, 使得对所有 $k \ge 0$ 均满足

$$\frac{y_k^T s_k}{\|s_k\|^2} \ge m, \frac{\|y_k\|^2}{y_k^T s_k} \le M, \tag{3.2}$$

则存在常数 $\mu_i > 0, (i = 1, 2, 3)$, 使得对 $\forall k \ge 0$, 至少存在 $[\frac{k}{2}]$ 个 i 值 (其中 $i \le k$) 满足不等式

$$\mu_1\|s_i\|^2 \le s_i^T B_i s_i \le \mu_2\|s_i\|^2, \|B_i s_i\| \le \mu_3\|s_i\|. \tag{3.3}$$

将 $[1, k]$ 中满足不等式 (3.3) 的下标从小到大依次记为: $K_k = \{i_1 < i_2 < \cdots < i_k\}$, $K = \bigcup\limits_{k=1}^{\infty} K_k$, 由引理 2.1, 有

$$\mu_1\|s_k\|^2 \le s_k^T y_k \le s_k^T G(\xi^*) s_k \le \mu_2\|s_i\|^2, \tag{3.4}$$

其中 $\xi^* = x_k + \tau(x_{k+1} - x_k), \tau \in (0, 1)$.

引理 3.2 [5] 对所有的试探步 α_k 满足

$$\alpha_k \ge \min\{1, \frac{(1-\rho)\beta}{L} \cdot \frac{s_k^T B_k s_k}{\|s_k\|^2}\}, \tag{3.5}$$

其中 $L > 0$ 是 g 的 Lipschitz 常数.

定理 3.1 设 $\{x_k\}$ 是由标准 BFGS 信赖域线搜索方法产生的点列, f 二次连续可微且一致凸, 则序列 $\{x_k\}$ 收敛到 (1.1) 的唯一解.

证 因为算法保证了 $\{f(x_k)\}$ 单调递减, 且 f 一致凸. 所以只需证明

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{3.6}$$

因为 f 是一个二次连续可微的凸函数, 则存在常数 $M>0, m>0$, 使得对所有 k 不等式 (3.2) 成立. 因此, 上面定义的指标集 K 是无限集, 另外由式 (3.3) 和式 (3.4) 可得, 对所有 $i \in K$ 有

$$\alpha_i \geq \min\{1,(1-\rho)\beta\mu_1 L^{-1}\} \equiv \overline{\alpha} > 0.$$

分两种情况讨论:

如果 $r_k \geq \eta$, 则对所有 $i \in K$, 有 $f(x_i)-f(x_{i+1}) \geq \eta(q_k(0)-q_k(s_i)) \geq \frac{\mu_1\eta\|s_i\|^2}{2}$.

如果 $r_k < \eta$, 则 $f(x_i)-f(x_{i+1}) \geq -\rho\alpha_i g_i^T s_i \geq \rho\alpha_i s_i^T B_i s_i \geq \rho\overline{\alpha}\mu_1\|s_i\|^2$, 所以存在一个常数 $\mu>0$, 对所有 $i \in K$ 满足

$$f(x_i)-f(x_{i+1}) \geq \mu\|s_i\|^2. \tag{3.7}$$

又因为序列 $\{f(x_k)\}$ 单调递减且有界, 所以极限存在. 因此

$$\lim_{i\in K, i\to\infty}\|s_k\| = 0. \tag{3.8}$$

否则 $\Delta_i \geq \Delta_{\min}$ 对所有 i 成立, 对 $\forall i \in K, \|s_i\| < \Delta_i$ 对充分大 i 成立. 由式 (2.2) 可得 $\alpha_i = 0$, 由式 (3.3) 知 $\|g_i\| = \|B_i s_i\| \leq \mu_3\|s_i\|$, 结合式 (3.8) 得 $\lim\limits_{k\to\infty}\inf\|g_k\| = 0$.

引理 3.3 [4] 假设 $(x_k, x_{k+1}, g_{k+1}, s_k)$ 由算法 1 产生, 则 $\forall k$ 都有

$$f(x_k) = f(x_{k+1}) + g_{k+1}^T(x_k - x_{k+1}) + \frac{(x_k-x_{k+1})^T B_{k+1}(x_k-x_{k+1})}{2}.$$

引理 3.4 设 $\{x_k\}$ 是由算法 1 产生的点列, B_0 是对称正定矩阵. 若存在常数 $M>m>0$ 和 $P \geq 0$ 使得不等式

$$m\|s_k\|^2 \leq y_k^{*T}s_k \leq M\|s_k\|^2, \|y_k^*\| \leq P\|s_k\|, \tag{3.9}$$

对所有 $k \geq 0$ 成立, 则存在常数 $\mu_i > 0, i=1,2,3$, 使得 $\forall k \geq 0$, 至少 $[\frac{k}{2}]$ 个 i 值满足式 (3.3), 其中 $i \leq k$.

证 由 y_k^* 的定义可知

$$\begin{aligned}
y_k^{*T}s_k &= \frac{\delta_k^T s_k \cdot \delta_k^T s_k}{|\delta_k^T s_k|} = |\delta_k^T s_k| = |[\theta y_k + (1-\theta)A_k(2)s_k]^T s_k]| \\
&= |\theta y_k^T s_k + (1-\theta)[2(f_k - f_{k+1}) + (g_{k+1}+g_k)^T s_k]| \\
&= |g_{k+1}^T s_k + (1-2\theta)g_k^T s_k + (2-2\theta)(f_k - f_{k+1})| \\
&= |(g_{k+1}^T s_k + f_k - f_{k+1}) + (1-2\theta)(g_k^T s_k + f_k - f_{k+1})| \\
&= |\frac{s_k^T G(\xi_1)s_k}{2} + (2\theta-1)\frac{s_k^T G(\xi_2)s_k}{2}| \\
&\leq |\frac{s_k^T G(\xi_1)s_k}{2}| + |(2\theta-1)\frac{s_k^T G(\xi_2)s_k}{2}| \\
&\leq (\frac{1}{2} + \frac{|(2\theta-1)|}{2}))|s_k^T G(\xi_3)s_k| \\
&\leq (\frac{1}{2} + \frac{|(2\theta-1)|}{2})\mu_2\|s_k\|^2,
\end{aligned}$$

其中 $s_k^TG(\xi_3)s_k=\max\{s_k^TG(\xi_1)s_k, s_k^TG(\xi_2)s_k\}$, $\xi_1=x_k+\theta_1(x_{k+1}-x_k)$, $\xi_2=x_k+\theta_2(x_{k+1}-x_k)$, $\xi_3=x_k+\theta_3(x_{k+1}-x_k)$, $\theta_1,\theta_2,\theta_3\in(0,1)$, 令 $M=(\frac{1}{2}+\frac{|(2\theta-1)|}{2})\mu_2$, 则 $y_k^{*T}s_k\le M\|s_k\|^2$.

用同样的方法可以得到 $m\|s_k\|^2\le y_k^{*T}s_k$, 得到式 (3.9) 的第一个不等式. 下证式 (3.9) 的第二个不等式.

$$\begin{aligned}
\|y_k^*\| &= \|\theta y_k+(1-\theta)A_k(2)s_k\|=\|\theta y_k+(1-\theta)\frac{2(f_k-f_{k+1})+(g_{k+1}+g_k)^Ts_k}{\|s_k\|^2}Is_k\|\\
&= \|\theta y_k+(1-\theta)\frac{-2g_{k+1}^Ts_k+s_k^TG(\xi_4)s_k+g_{k+1}^Ts_k+g_k^Ts_k}{\|s_k\|^2}Is_k\|\\
&\le \theta\|y_k\|+(1-\theta)\|y_k\|+\frac{(1-\theta)|s_k^TG(\xi_4)s_k|}{\|s_k\|}\le\|y_k\|+(1-\theta)\mu_2\|s_k\|\\
&\le [L+(1-\theta)\mu_2]\|s_k\|,
\end{aligned}$$

其中 $\xi_4=x_k+\theta_4(x_{k+1}-x_k)$, $\theta_4\in(0,1)$, 最后一个不等式利用了假设 B). 令 $P=L+(1-\theta)\mu_2$, 得到式 (3.9) 的第二个不等式. 类似引理 3.1 的证明, 得出引理 3.4 成立.

定理 3.2 假设 (i) 成立, 点列 $\{x_k\}$ 由 MBFGS 信赖域算法产生, 则

$$\lim_{k\to\infty}\inf\|g_k\|=0. \tag{3.10}$$

证 由 $f(x_k)$ 的下降性, 得 $\{x_k\}\subset L(x_0)$ 有界. 如果式 (3.10) 非真, 则存在一个正常数 δ, 对所有 k 使 $\|g_k\|\ge\delta$ 成立. 根据引理 3.1, 类似定理 3.1 的证明, 推出矛盾, 从而结论成立.

4 超线性敛性

为了证明算法 1 的超线性收敛性, 给出如下假设:

假设 (ii):

C) 算法 1 产生的序列 $\{x_k\}$ 收敛到 x^* 且 $\nabla^2f(x^*)$ 正定;

D) 函数 f 在 x^* 的邻域内二次连续可微, 且 $\nabla f(x^*)=0$;

E) Hessian 矩阵 $G(x)$ 在 x^* 处 Hölder 连续, 即存在正常数 υ,Ψ 使得不等式

$$\|G(x)-G(x^*)\|\le\Psi\|x-x^*\|^{\upsilon}, \tag{4.1}$$

对所有的 $x\in U(x^*)$ 成立.

引理 4.1 [4] 设点列 $\{x_k\}$ 由算法 1 产生, 若假设 (ii) 成立, 则有

$$\lim_{k\to\infty}\|A_k(2)\|=0. \tag{4.2}$$

引理 4.2 [4] 设点列 $\{x_k\}$ 由算法 1 产生, 若假设 (i) 和假设 (ii) 成立, 则有

$$\sum_{k=0}^{\infty}\|x_k-x^*\|<\infty,\sum_{k=0}^{\infty}\phi_k<\infty,\sum_{k=0}^{\infty}A_k(2)<\infty,$$

其中 $\phi_k=\max\{\|x_k-x^*\|^{\upsilon},\|x_{k+1}-x^*\|^{\upsilon}\}$.

引理 4.3　设 $\{x_k\}$ 是算法 1 产生的点列, 定义 $Q=G(x^*)^{-\frac{1}{2}}$,$H_k=B_k^{-1}$, 则存在正常数 $b_i, i=1,2,3,\cdots,7$, 及 $\gamma\in(0,1)$, 使得对充分大 k,

$$\|B_{k+1}-G(x^*)\|_{Q,F}\le(1+b_1\phi_k)\|B_k-G(x^*)\|_{Q,F}+b_2\phi_k+b_3\|A_k(2)\|, \tag{4.3}$$

$$\begin{aligned}&\|H_{k+1}-G(x^*)^{-1}\|_{Q^{-1},F}\\ \le\ &(\sqrt{1-\gamma\omega_k^2}+b_4\phi_k+b_5\|A_k(2)\|)\|H_k-G(x^*)^{-1}\|_{Q^{-1},F}+b_6\phi_k+b_7\|A_k(2)\|,\end{aligned} \tag{4.4}$$

其中 $\|A\|_{Q,F}=\|Q^TAQ\|_F$, $\|.\|_F$ 是 Frobenius 范数. ω_k 由下式定义

$$\omega_k=\frac{\|Q^{-1}(H_k-G(x^*)^{-1})y_k^*\|}{\|H_k-G(x^*)^{-1}\|_{Q^{-1},F}\|Qy_k^*\|}, \tag{4.5}$$

且 $\{\|B_k\|\}_F$ 和 $\{\|H_k\|\}_F$ 有界.

证　由式 (1.4) 得

$$\begin{aligned}\|B_{k+1}-G(x^*)\|_{Q,F}&=\|B_k-G(x^*)+\frac{B_ks_ks_k^TB_k}{s_k^TB_ks_k}+\frac{y_k^*y_k^{*T}}{s_k^Ty_k^*}\|_{Q,F}\\ &\le\|B_k-G(x^*)+\frac{B_ks_ks_k^TB_k}{s_k^TB_ks_k}+\frac{y_ky_k^T}{s_k^Ty_k}\|_{Q,F}+\|\frac{y_k^*y_k^{*T}}{s_k^Ty_k^*}-\frac{y_ky_k^T}{s_k^Ty_k}\|_{Q,F}\\ &\le(1+b_1\phi_k)\|B_k-G(x^*)\|_{Q,F}+b_2\phi_k+\|\frac{y_k^*y_k^{*T}}{s_k^Ty_k^*}-\frac{y_ky_k^T}{s_k^Ty_k}\|_{Q,F}.\end{aligned}$$

最后一个不等式由文献 [6] 得到

$$\begin{aligned}&\|\frac{y_k^*y_k^{*T}}{s_k^Ty_k^*}-\frac{y_ky_k^T}{s_k^Ty_k}\|_{Q,F}=\|\frac{\delta_k\delta_k^T}{|\delta_k^Ts_k|}-\frac{y_ky_k^T}{s_k^Ty_k}\|_{Q,F}\\ =\ &\|\frac{[\theta y_k+(1-\theta)A_k(2)s_k][\theta y_k+(1-\theta)A_k(2)s_k]^T}{s_k^T[\theta y_k+(1-\theta)A_k(2)s_k]}-\frac{y_ky_k^T}{s_k^Ty_k}\|_{Q,F}\\ =\ &\|\frac{s_k^Ty_k[\theta y_k+(1-\theta)A_k(2)s_k][\theta y_k+(1-\theta)A_k(2)s_k]^T-s_k^T[\theta y_k+(1-\theta)A_k(2)s_k]y_ky_k^T}{s_k^Ty_ks_k^T[\theta y_k+(1-\theta)A_k(2)s_k]}\|_{Q,F}\\ \le\ &(1-\theta)\|A_k(2)\|\frac{\|2\theta s_k^Ty_ky_ks_k^T\|_{Q,F}+\|s_k^Ts_ky_ky_k^T\|_{Q,F}+\|(1-\theta)s_k^Ty_kA_k(2)s_ks_k^T\|_{Q,F}}{s_k^Ty_ks_k^T\delta_k}\\ \le\ &(1-\theta)\|A_k(2)\|\frac{(2\theta+1)\|s_k\|^2\|y_k\|^2\|Q\|_F^2+(1-\theta)\|s_k\|^3\|y_k\|\|A_k(2)\|\|Q\|_F^2}{s_k^Ty_ks_k^T\delta_k}\\ \le\ &\|A_k(2)\|\frac{(1-\theta)\|Q\|_F^2[(2\theta+1)\|y_k\|^2+(1-\theta)\|s_k\|\|y_k\|\|A_k(2)\|]}{m\mu_1\|s_k\|^2}.\end{aligned}$$

最后一个不等式利用了式 (3.4) 和式 (3.9).

由假设 (ii) 可知, 存在常数 b_3, 使得式 (4.3) 成立. 下证式 (4.4). 易知式 (1.4) 的逆

$$\begin{aligned}H_{k+1}&=H_k+\frac{(s_k-H_ky_k^*)s_k^T+s_k(s_k-H_ky_k^*)^T}{y_k^{*T}s_k}-\frac{y_k^{*T}(s_k-H_ky_k^*)s_ks_k^T}{(y_k^{*T}s_k)^2}\\ &=(I-\frac{s_ky_k^{*T}}{y_k^{*T}s_k})H_k(I-\frac{y_k^*s_k^T}{y_k^{*T}s_k})+\frac{s_k^*s_k^T}{y_k^{*T}s_k}.\end{aligned}$$

这是 DFP 校正公式的对偶形式, 对应关系 $H_k \to B_k, H_{k+1} \to B_{k+1}, s_k \to y_k$. 由文献 [7] 中的引理 4.1 知, 存在常数 $b_8, b_9 \geq 0$, 使得

$$\begin{aligned}\|H_{k+1}-G(x^*)^{-1}\|_{Q^{-1},F} \quad &\leq \quad (\sqrt{1-\gamma\omega_k^2}+b_8\frac{\|Q^{-1}s_k-Qy_k^*\|}{\|Qy_k^*\|})\|H_k-G(x^*)^{-1}\|_{Q^{-1},F}\\ &\quad +b_9\frac{\|s_k-G(x^*)^{-1}y_k^*\|}{\|Qy_k^*\|},\end{aligned}\tag{4.6}$$

其中 ω_k 由式 (4.5) 产生.

$$\begin{aligned}\|Qy_k^*-Q^{-1}s_k\| \quad &\leq \quad \|Q\|\|y_k^*-Q^{-2}s_k\|=\|Q\|\|y_k^*-G(x^*)s_k\|=\|Q\|\|\frac{\delta_k^Ts_k}{|\delta_k^Ts_k|}\delta_k-G(x^*)s_k\|\\ &= \quad \|Q\|\|\theta y_k+(1-\theta)A_k(2)s_k-G(x^*)s_k\|\\ &\leq \quad \|Q\|(\|\theta\int_0^1 G(x_k+\gamma s_k)s_kd\gamma-\theta G(x^*)s_k\|+\|(1-\theta)A_k(2)s_k\|)\\ &\leq \quad \|Q\|\|s_k\|[\theta\Psi\int_0^1\|x_k-x^*-\gamma s_k\|^{\upsilon}d\gamma+(1-\theta)\|A_k(2)\|]\\ &\leq \quad \|Q\|\|s_k\|[\theta\Psi\phi_k+(1-\theta)\|A_k(2)\|].\end{aligned}$$

又因为 $\phi_k \to 0$, $A_k(2) \to 0$, 则 k 充分大时, 有 $\|Qy_k^*-Q^{-1}s_k\| \leq q\|Q\|\|s_k\|$ 成立, 其中 $q \in (0,1)$ 为常数. 此外, 存在常数 b_{10}, 使得对充分大的 k, 下式成立.

$$\begin{aligned}\|Qy_k^*\| \quad &= \quad \|Q[\theta y_k+(1-\theta)A_k(2)s_k]\| \geq \|Q\theta(g_{k+1}-g_k)\|-(1-\theta)\|A_k(2)\|\|Qs_k\|\\ &\geq \quad b_{10}\|x_{k+1}-x_k\|-(1-\theta)\|A_k(2)\|\|Q\|\|s_k\| \geq [b_{10}-(1-\theta)\|A_k(2)\|\|Q\|]\cdot\|s_k\|.\end{aligned}$$

利用 $\|A_k(2)\| \to 0$, 上述不等式隐含存在常数 c, 使得当 k 充分大时, $\|Qy_k^*\| \geq c\|s_k\|$. 因此得到

$$\frac{\|Qy_k^*-Q^{-1}s_k\|}{\|Qy_k^*\|} \leq c^{-1}\|Q\|[\theta\Psi\phi_k+(1-\theta)\|A_k(2)\|],\tag{4.7}$$

$$\begin{aligned}&\frac{\|s_k-G(x^*)^{-1}y^*\|}{\|Qy_k^*\|}=\frac{\|s_k-Q^2y^*\|}{\|Qy_k^*\|}=\frac{\|Q(Qy^*-Q^{-1}s_k)\|}{\|Qy_k^*\|}\leq\frac{\|Q\|\|Qy^*-Q^{-1}s_k\|}{\|Qy_k^*\|}\\ \leq \quad &c^{-1}\|Q\|^2[\theta\Psi\phi_k+(1-\theta)\|A_k(2)\|].\end{aligned}\tag{4.8}$$

由式 (4.6), (4.7), (4.8) 得到式 (4.4), 由 $\sum\limits_{k=0}^{\infty}\phi_k \leq 0$, $\sum\limits_{k=0}^{\infty}A_k(2) \leq 0$ 和式 (4.3), (4.4), 得到 $\|B_k-G(x^*)\|_{Q,F}$ 和 $\|H_k-G(x^*)^{-1}\|_{Q^{-1},F}$ 收敛, 且 $\{\|B_k\|\}_F$ 和 $\{\|H_k\|\}_F$ 有界.

定理 4.1　设 $\{x_k\}$ 是算法 1 产生的序列, 若假设条件 (i) 和 (ii) 成立, 则序列 $\{x_k\}$ 超线性收敛到 x^*.

证　由引理 4.3 知 $\{B_k\}$ 和 $\{B_k^{-1}\}$ 有界. 与文献 [8] 中引理 3.9 的证明相似, 容易得到

$$\lim_{k\to\infty}\frac{\|(B_k-G(x_*))s_k\|}{\|s_k\|}=0.\tag{4.9}$$

因为函数 f 在 $U(x^*)$ 内强凸, 且 $\{B_k^{-1}\}$ 有界, 则当 $k\to\infty$ 时, $\|B_k^{-1}g_k\| \to 0$. 事实上, 对所有充分大的 k, $\|B_k^{-1}g_k\| \leq \Delta_{\min}$. 因此, 当 k 分大时, $s_k=-B_k^{-1}g_k$ 是式 (1.2) 的唯一解. 算法 1 同单位步长的 BFGS 一致. 因而, 结合式 (4.9) 得出点列 $\{x_k\}$ 的超线性收敛.

参 考 文 献

[1] Shultz G A, Schnabel R B, Byrd R H. A family of trust-region-based algorithms for unconstrained minimization with strong global convergence properties[J]. SIAM Journal on Numerical analysis, 1985, 22: 47–67.

[2] Buleau J P, Vial J Ph. Curvilinear path and trust region in unconstrained optimization, a convergence analysis[J]. Math. Prog. Study, 1987, 30: 82–101.

[3] Rendl F, Wolkowicz H. A semidefinite framework for trust region subproblems with applications to large scale minimization[J]. Math. Prog., 1997, 77: 273–299.

[4] Wei Z, Li G, Qi L. New quasi-Newton methods for unconstrained optimization problems[J]. Applied Mathmatics and Computation, 2006, 175(2): l156–1188.

[5] Wei Z, Zhou Y, Deng X. New trust region method for line search[J]. Journal of Chongqing Institute of Technology, 2007, 21: 1–7.

[6] A. Griewank, Ph.L. Toint. Local convergence analysis for partitioned for quasi-Newton update[J]. Numer. Math., 1982, 39: 429–448.

[7] Dai Y. Convergence properties of the BFGS algorithm [J]. SIAM J. Optimization, 2003, 13: 693–701

[8] Wei Z, Yu G, Yuan G, Lian Z. The superlinear convergence of a modified BFGS-type method for unconstrained optimization[J]. Computational Optimization and Applications, 2004, 29: 315–332

[9] Nocedal J, Yuan Y. Combining trust-region and line search techniques[J]. Advance in Nonlinear Programming, 1998, 14: 153–175.

[10] Byrd R, Nocedal J. A tool for the analysis of Quasi-Newton methods with application to unconstrained minimization[J]. SIAM J. Numerical Analysis, 1989, 26(3): 727–739.

A NEW MBFGS OF TRUST REGION ALGORITHM

JING Shu-jie[1] , MIAO Rong[1] , LI Shao-juan[2]

(*1.School of Mathematics and Information Science, Henan Polytechnic University, Jiaozuo 454000, China*)

(*2.Henan Care Vocational College, Anyang 455000, China*)

Abstract: In this paper, we study unconstrained optimization problems. By using the basic idea of the MBFGS trust region algorithm, we improve BFGS correction formula, combine with line search technique and put forward a new MBFGS trust region algorithm which broadens the scope of application of the trust region algorithm. Under certain conditions, the global convergence and superlinear convergence of the algorithm is proved.

Keywords: unconstrained optimization; trust region; BFGS(MBFGS) modification; line search

2010 MR Subject Classification: 90C30