修改 Broyden 族在一类非精确线搜索下的全局收敛性*

Global Convergence of the Modified Broyden's Family with a Class of Inexact Line Searches

韦增欣，谢品杰
WEI Zeng-xin，XIE Pin-jie

(广西大学数学与信息科学学院，广西南宁 530004)
(Department of Mathematics and Information Science，Guangxi University，Nanning，Guangxi，530004，China)

摘要：将一类 Wolfe 类线搜索模型的 LS 搜索模型与文献[10]提出的修改 Broyden 族(MBC1和 MBC2)相结合，得到 MBC1算法和 MBC2算法，并证明 MBC1算法和 MBC2算法在 LS 搜索模型下具有全局收敛性.

关键词：无约束最优化　拟牛顿方程　Broyden 族　全局收敛性
中图法分类号：O242.23　**文献标识码**：A　**文章编号**：1005-9164(2006)01-0012-05

Abstract：Using the combinations of a class Wolfe-type inexact line search and a modified Broyden's family proposed by Reference[10]，we derive a MBC1 algorithm and a MBC2 algorithm. Under suitable conditions，we prove that this two algorithms are global convergence with inexact linesearch.

Key words：unconstrained optimization，quasi-Newton equation，Broyden's family，global convergence

求解无约束最优化问题：

$$\min\{f(x) \mid x \in \mathscr{R}^n\}, \tag{0.1}$$

其中 $f(x)$：$\mathscr{R}^n \to \mathscr{R}$ 是连续可微函数，拟牛顿法是应用最广、理论最为成熟的方法之一，而 Broyden 族又是其中最著名的方法，其迭代形式为：

$$x_{k+1} = x_k + \alpha_k d_k, \tag{0.2}$$

其中 α_k 是步长，d_k 是搜索方向并且满足：

$$g_k + B_k d_k = 0, \tag{0.3}$$

B_k 是 Hessian 阵 $G(x)$ 的近似，它有以下迭代形式产生：

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} + \phi_k w_k w_k^T, \tag{0.4}$$

其中 $g_k = \bigtriangledown f(x_k)$，$y_k = g_{k+1} - g_k$，$s_k = x_{k+1} - x_k = \alpha_k d_k$，$w_k = (s_k^T B_k s_k)^{\frac{1}{2}}[\frac{y_k^*}{s_k^T y_k^*} - \frac{B_k s_k}{s_k^T B_k s_k}]$，$\phi_k$ 是常数.

当 $\phi_k \in [0,1]$ 时，称为 Broyden 凸族，特别地 $\phi_k = 1$ 或者 0 时，即分别为著名的 DFP 算法与 BFGS 算法；而当 $\phi_k < 0$ 时，则称为 Broyden 非凸族[6]. 20 世纪 70 年代以来，对于 Broyden 族的收敛性研究取得了相当的成果[1~6]. 近年来，基于修改拟牛顿方程的牛顿类算法的研究也成为无约束最优化算法理论研究的一个新的热点[7~10]. 最近，韦增欣等[10] 提出一类修正拟牛顿方程：

$$B_{k+1} s_k = y_k^*, \tag{0.5}$$

其中 $y_k^* = y_k + A_k s_k$，A_k 为一矩阵，并给出了两种类型的修改 Broyden 族(MBC1 和 MBC2) 的迭代形式.

MBC1：

$$B_{k+1}(1) = B_k(1) - \frac{B_k(1) s_k s_k^T B_k(1)}{s_k^T B_k(1) s_k} + \frac{y_k^* y_k^{*T}}{y_k^{*T} s_k} + \phi_k w_k w_k^T, \tag{0.6}$$

其中 $w_k = (s_k^T B_k(1) s_k)^{\frac{1}{2}}[\frac{y_k^*}{s_k^T y_k^*} - \frac{B_k(1) s_k}{s_k^T B_k(1) s_k}]$，$y_k^* = y_k + A_k(1) s_k$，$A_k(1) = \mu_k \frac{|g(x_{k+1})^T (g(x_{k+1}) - g(x_k))|}{\|s_k\|} I$.

MBC2：

$$B_{k+1}(2) = B_k(2) - \frac{B_k(2) s_k s_k^T B_k(2)}{s_k^T B_k(2) s_k} + \frac{y_k^* y_k^{*T}}{y_k^{*T} s_k} + \phi_k w_k w_k^T, \tag{0.7}$$

其中 $w_k = (s_k^T B_k(2) s_k)^{\frac{1}{2}}[\frac{y_k^*}{s_k^T y_k^*} - \frac{B_k(2) s_k}{s_k^T B_k(2) s_k}]$，$y_k^* = y_k + A_k(2) s_k$，$A_k(2) = \frac{2[f(x_k) - f(x_{k+1})] + [g(x_{k+1}) + g(x_k)]^T s_k}{\|s_k\|^2}$.

收稿日期：2005-04-18
作者简介：韦增欣(1963-)，男，广西武鸣人，教授，主要从事最优化理论与方法研究。
* 国家自然科学基金(10161002)和广西自然科学基金(0135004)联合资助。

对于修改 Broyden 族我们有如下引理：

引理 0.1 设 B_k 正定，对于修改 Broyden 族校正公式(0.6)和(0.7)，B_{k+1} 仍然正定的充要条件是

$$s_k^T y_k^* > 0 \tag{0.8}$$

且

$$\phi_k > \phi_k^{*c} \equiv \frac{1}{1-\mu_k^*}, \mu_k^* = \frac{(y_k^{*T} H_k y_k^*)(s_k^T B_k s_k)}{(y_k^{*T} s_k)^2}. \tag{0.9}$$

证明 参见文献[11]定理 5.5.2 的证明过程.

文献[9,10]证明在适当条件下，结合 Wolfe 搜索的修改 BFGS(MBFGS)算法具有全局收敛性和超线性收敛性，并且数据试验表明，修改 BFGS 算法要优于 BFGS 算法. 本文将 LS 搜索模型与修改 Broyden 族(MBC1 和 MBC2)相结合，其中

$$\phi_k \in [(1-v)\phi_k^{*c}, 1), v \in (0,1). \tag{0.10}$$

并证明在适当的条件下，MBC1 和 MBC2 具有全局收敛性，从而推广了已有的结论.

1 LS 搜索模型和算法

为方便介绍，引入以下符号：

$$\cos\theta_k = \frac{s_k^T B_k s_k}{\|s_k\| \|B_k s_k\|}, r_k = -\frac{g_k^T d_k}{\|d_k\|}, a_k = g_k^T H_k g_k. \tag{1.1}$$

同时引入一个重要的函数定义：

定义 1.1[4] 映射 $\sigma: [0,\infty) \to [0,\infty)$ 称为强迫函数，如果 $\forall\, t_i \geqslant 0$ 有

$$\lim_{i\to\infty} \sigma(t_i) = 0 \Rightarrow \lim_{t\to\infty} t_i = 0.$$

本文使用的线搜索模型为：

LS 搜索模型[4] 给定 $b \in [0,1)$，函数 $\psi_1: R^1 \to R^1$ 满足

$$\lim_{t\to 0} \psi_1(t) = 0,$$

σ_1, σ_2 都是强迫函数，步长 α_k 满足

$$f(x_k + \alpha_k d_k) \leqslant f(x_k) - \min\{\sigma_1(r_k), \sigma_2(\alpha_k a_k)\}, \tag{1.2}$$

$$\frac{g(x_k+\alpha_k d_k)^T d_k}{\|d_k\|} \geqslant -br_k - \psi_1(\|s_k\|),$$

$$\limsup_{k\to\infty} \frac{\psi_1(\|s_k\|)}{r_k} < 1-b. \tag{1.3}$$

将 LS 搜索模型分别与 MBC1 和 MBC2 相结合，我们将得到以下两个算法：

MBC1算法

步骤 1：给出 $x_0 \in \mathscr{R}^n$，$B_0 \in \mathscr{R}^{n\times n}$，$0 \leqslant \varepsilon < 1$，$k := 0$.

步骤 2：若 $\|x_k\| \leqslant \varepsilon$ 则停止，否则计算 $B_k d_k + g_k = 0$，求得搜索方向 d_k.

步骤 3：利用 LS 搜索模型求得步长 α_k.

步骤 4：令 $x_{k+\alpha} = x_k + \alpha_k d_k$，利用公式(0.6)计算 $A_k(1)$ 和 B_{k+1}.

步骤 5：$k := k+1$.

MBC2算法

将 MBC1 算法中的步骤 4 替换成“利用公式(0.7)计算 $A_k(2)$ 和 B_{k+1}”，即得 MBC2 算法.

2 全局收敛性分析

2.1 MBC2算法的全局收敛性

先作出以下假设：

假设 1 水平集 $\Omega = \{x \mid f(x) \leqslant f(x_0)\}$ 包含在有界凸集 D 内.

假设 2 f 在 D 上连续可导，并且存在常数 $L > 0$，使得

$$\|g(x) - g(y)\| \leqslant L\|x - y\|, \forall\, x, y \in D.$$

假设 3 f 是一致凸的，即存在正数 m_1, m_2 使得

$$m_1\|z\|^2 \leqslant z^T G(x) z \leqslant m_2 \|z\|^2.$$

由假设 1 和假设 2，我们可以推出存在常数 $M > 0$ 使得

$$\|y_k\| \leqslant M, k = 1,2,3,\cdots \tag{2.1}$$

成立.

下面我们先给出几个引理.

引理 2.1[5] 若非负序列 $\{m_k\}$ 满足

$$\prod_{i=0}^{k} m_i \geqslant c_0^k, k = 1,2,3,\cdots$$

其中 c_0 为正常数，则 $\limsup_k m_k > 0$.

引理 2.2[4] 设 $\{x_k\}$ 由 MBC2 算法产生，则

$$\sum_{k=0}^{\infty} \min\{\sigma_1(r_k), \sigma_2(\alpha_k a_k)\} < \infty. \tag{2.2}$$

引理 2.3 设 $\{x_k\}$ 由 MBC2 算法产生，其中 ϕ_k 满足(0.10)，则有

$$Det(B_{k+1}) \geqslant Det(B_k) v \frac{y_k^{*T} s_k}{s_k^T B_k s_k}, \tag{2.3}$$

其中 $Det(B_k)$ 表示 B_k 的行列式.

证明 首先证明

$$Det(B_{k+1}) = [1 - \phi_k(1-\mu_k^*)] Det(B_k) \frac{y_k^{*T} s_k}{s_k^T B_k s_k} \tag{2.4}$$

成立. 由(0.7)知

$$B_{k+1} = B_{k+1}^{MBFGS} + \phi_k w_k w_k^T.$$

利用公式 $Det(I + uv^T) = 1 + u^T v$[13]，有

$$Det(B_{k+1}) = Det\{B_{k+1}^{MBFGS}[I + \phi_k (B_{k+1}^{MBFGS})^{-1} w_k w_k^T]\} = Det(B_{k+1}^{MBFGS})[1 + \phi_k w_k^T (B_{k+1}^{MBFGS})^{-T} w_k].$$

利用公式 $Det(I + u_1 u_2^T + u_3 u_4^T) = (1 + u_1^T u_2)(1 + u_3^T u_4) - (u_1^T u_4)(u_2^T u_3)$[13]，我们可以求得

$$Det(B_{k+1}^{MBFGS}) = Det(B_k)\frac{y_k^{*T}s_k}{s_k^TB_ks_k}.$$

注意到

$$(B_{k+1}^{MBFGS})^{-1} = H_{k+1}^{MBFGS} = H_k + \frac{(s_k - H_ky_k^*)s_k^T + s_k(s_k - H_ky_k^*)^T}{y_k^{*T}s_k} - \frac{y_k^*(s_k - H_ky_k^*)}{(y_k^{*T}s_k)^2}s_ks_k^T = (I - \frac{s_k y_k^{*T}}{y_k^{*T}s_k})H_k(I - \frac{y_k^*s_k^T}{y_k^{*T}s_k}) + \frac{s_ks_k^T}{y_k^{*T}s_k}.$$

代入 $w_k^T(B_{k+1}^{MBFGS})^{-T}w_k$ 容易求得

$$w_k^T(B_{k+1}^{MBFGS})^{-T}w_k = \frac{(y_k^{*T}H_ky_k^*)(s_k^TB_ks_k)}{(y_k^{*T}s_k)^2} - 1 = \mu_k^* - 1.$$

故(2.4)式成立.再根据(0.10)式即得(2.3)式成立.

引理 2.4 设$\{x_k\}$由MBC2算法产生,则存在$M_1 > 0$使得

$$\frac{y_k^Ts_k}{\|s_k\|^2} \leqslant m_2, \quad (2.5)$$

$$\frac{\|y_k^*\|^2}{y_s^{*T}s_k} \leqslant M_1. \quad (2.6)$$

证明 利用假设3,有

$$y_k^Ts_k = s_k^T[\int_0^1 G(x_k + \eta s_k)d\eta]s_k \leqslant m_2\|s_k\|^2,$$

即(2.5)式成立.下面再证明(2.6)式.一方面利用Taylor展开式,有

$$s_k^Ty_k^* = s_k^Ty_k + 2[f(x_k) - f(x_{k+1})] + [g(x_{k+1}) + g(x_k)]^Ts_k = 2[f(x_k) - f(x_{k+1})] + 2g(x_{k+1})^Ts_k = 2[- g(x_{k+1})^Ts_k + \frac{1}{2}s_k^TG(x_k + \theta(x_{k+1} - x_k))s_k] + 2g(x_{k+1})^Ts_k = s_k^TG(x_k + \theta(x_{k+1} - x_k))s_k, \theta \in (0,1).$$

由假设3可得

$$m_1\|s_k\|^2 \leqslant s_k^Ty_k^* \leqslant m_2\|s_k\|^2. \quad (2.7)$$

另一方面,利用y_k^*的定义和假设2、假设3有

$$\|y_k^*\| \leqslant \|y_k\| + \frac{|2[f(x_k) - f(x_{k+1})] + [g(x_{k+1}) + g(x_k)]^Ts_k|}{\|s_k\|} \leqslant 2\|y_k\| + \frac{|s_k^TG(x_k + \theta(x_{k+1} - x_k))s_k|}{\|s_k\|} \leqslant 2L\|s_k\| + m_2\|s_k\| = (2L + m_2)\|s_k\|.$$

所以(2.6)式成立,其中$M_1 = \frac{(2L + m_2)^2}{m_1}$.

引理 2.5 设$\{x_k\}$由MBC2算法产生,其中ϕ_k满足(0.10)式,则存在正数m和k_1,使得$k \geqslant k_1$时有

$$\frac{s_k^Ty_k^*}{s_k^TB_ks_k} \geqslant \frac{m}{\alpha_k}. \quad (2.8)$$

证明 由LS搜索模型的第二条不等式(1.3)可知,存在正数$\varepsilon_1 < 1 - b, k_1$,使得$k \geqslant k_1$时有

$$\psi_1(\|s_k\|) \leqslant \varepsilon_1 r_k.$$

故有

$$y_k^Ts_k = \alpha_k[g(x_k + \alpha_kd_k) - g(x_k)]^Td_k \geqslant \alpha_k[- br_k\|d_k\| - \|d_k\|\psi_1(\|s_k\|) - g_k^Td_k] \geqslant \alpha_k[- b(- g_k^Td_k) - \varepsilon_1r_k\|d_k\| - g_k^Td_k] = (1 - b - \varepsilon_1)(- g_k^Ts_k), \quad k \geqslant k_1. \quad (2.9)$$

再利用$s_k = - \alpha_kB_k^{-1}g_k$,有

$$(1 - b - \varepsilon_1)s_k^TB_ks_k = - (1 - b - \varepsilon_1)\alpha_ks_k^Tg_k \leqslant \alpha_ky_k^Ts_k = \alpha_ks_k^T[\int_0^1 G(x_k + \tau s_k)d\tau]s_k, k \geqslant k_1.$$

由假设3可知

$$s_k^TB_ks_k \leqslant \frac{\alpha_k}{1 - b - \varepsilon_1}m_2\|s_k\|^2, \quad k \geqslant k_1,$$

结合(2.7)式即得

$$\frac{s_k^Ty_k^*}{s_k^TB_ks_k} \geqslant \frac{m}{\alpha_k}, k \geqslant k_1,$$

其中$m = \frac{m_1(1 - b - \varepsilon_1)}{m_2}$.

引理 2.6 设$\{x_k\}$由MBC2算法产生,记

$$\eta_k = \phi_k(\frac{\|y_k^*\|\|s_k\|\cos\theta_k}{y_k^{*T}s_k})^2 - 2\phi_k\frac{y_k^{*T}B_ks_k\|s_k\|\cos\theta_k}{y_k^{*T}s_k\|B_ks_k\|},$$

则

$$\lim_{k\to\infty}\cos\theta_k = 0 \Rightarrow \lim_{k\to\infty}\eta_k\|g_k\| = 0.$$

证明

$$|\eta_k\|g_k\|| \leqslant \phi_kM_1\frac{\|s_k\|^2\cos\theta_k\|g_k\|}{y_k^{*T}s_k} + 2\phi_k\frac{\|B_ks_k\|\|s_k\|}{\|B_ks_k\|}\frac{\|y_k^{*T}\|\|g_k\|\cos\theta_k}{y_k^{*T}s_k\|B_ks_k\|} \leqslant \phi_kM_1\frac{\|g_k\|cos\theta_k}{m_1} + 2\phi_k\sqrt{M_1}\frac{\|s_k\|\|g_k\|cos\theta_k}{\sqrt{y_k^{*T}s_k}} \leqslant \phi_k\frac{M_1\|g_k\|\cos\theta_k}{m_1} + 2\phi_k\sqrt{\frac{M_1}{m_1}}\|g_k\|\cos\theta_k = \theta_k(\frac{M_1}{m_1} + 2\sqrt{\frac{M_1}{m_1}})\|g_k\|\cos\theta_k.$$

其中第一不等号利用了(2.6)式,第二和第三个不等号则是由(2.7)式所得到.

再由(2.1)式,得

$$0 \leqslant |\eta_k\|g_k\|| \leqslant \phi_k(\frac{M_1}{m_1} + 2\sqrt{\frac{M_1}{m_1}})M\cos\theta_k.$$

显然,引理得证.

注意到不等式(2.5)和(2.9)成立,类似于文献[4]中定理1的证明过程,我们有以下结论:

引理 2.7 设$\{x_k\}$由MBC2算法产生,其中ϕ_k满足(0.10)式,则

$$\lim_{k\to\infty}\sup\cos\theta_k > 0 \Rightarrow \lim_{k\to\infty}\inf\|g_k\| = 0.$$

下面我们将给出本文的主要结论,即在LS搜索模型下,MBC2算法具有全局收敛性($\phi_k \in [(1 - v)\phi_k^{*c}, 1), v \in (0,1)$).

定理 2.1 设$\{x_k\}$由MBC2算法产生,其中ϕ_k满足(0.10)式,则有

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

证明 由引理 2.7 知,我们只需考虑 $\lim_{k\to\infty} \cos\theta_k = 0$ 的情况.

下面我们用反证法来证明定理成立,否则总存在常数 $c > 0$ 使得

$$\|g_k\| \geqslant c > 0, k = 1,2,3,\cdots \tag{2.10}$$

对(0.7)式两边同时求迹,有

$$Tr(B_{k+1}) = Tr(B_k) + \frac{\|y_k^*\|^2}{y_k^{*T}s_k} - \frac{\|B_ks_k\|^2}{s_k^TB_ks_k} + \phi_k\|w_k\|^2 = Tr(B_k) + \frac{\|y_k^*\|^2}{y_k^{*T}s_k} + \phi_k\frac{\|y_k^*\|^2}{y_k^{*T}s_k}\frac{s_k^TB_ks_k}{y_k^{*T}s_k} - (1-\phi_k)\frac{\|B_ks_k\|^2}{s_k^TB_ks_k} - 2\phi_k\frac{y_k^{*T}B_ks_k}{y_k^{*T}s_k} = Tr(B_k) + \frac{\|y_k^*\|^2}{y_k^{*T}s_k} + \eta_k\frac{s_k^TB_ks_k}{\|s_k\|^2\cos^2\theta_k} - (1-\phi_k)\frac{\|B_ks_k\|^2}{s_k^TB_ks_k} = Tr(B_k) + \frac{\|y_k^*\|^2}{y_k^{*T}s_k} + [\eta_k\|g_k\| - (1-\phi_k)\|g_k\|]\frac{\|g_k\|}{a_k}, \tag{2.11}$$

其中 $Tr(B_k)$ 表示 B_k 的迹.

记 $J_1 = \{k|0 \leqslant \phi_k < 1, k\in N\}$, $J_2 = N - J_1 = \{k|(1-v)\phi_k^{*C} \leqslant \phi_k < 0, k \in N\}$.

一方面,当 $k \in J_1$ 时,由引理 2.6 和(2.10)式知

$$\limsup_{k\to\infty}[\eta_k\|g_k\| - (1-\phi_k)\|g_k\|] \leqslant \limsup_{k\to\infty}[\eta_k\|g_k\| - c(1-\phi_k)] = -c(1-\phi_k),$$

从而得知,当 k 充分大的时候,成立

$$\eta_k\|g_k\| - (1-\phi_k)\|g_k\| < -\frac{c(1-\phi_k)}{2}, \tag{2.12}$$

不失一般性,我们不妨假设(2.12)式对一切的 $k \in J_1$ 均成立,将之代入(2.11)式,并结合引理 2.4,有

$$Tr(B_{k+1}) \leqslant Tr(B_k) + M_1 - \frac{c^2(1-\phi_k)}{a_k}. \tag{2.13}$$

当 $k \in J_2$ 时,由(2.6),(2.10)及(2.11)式的第一个等式,得

$$Tr(B_{k+1}) \leqslant Tr(B_k) + M_1 - \frac{c^2}{a_k}. \tag{2,14}$$

所以,不等式(2.13)对 $k \in N$ 均成立.从而,我们总有

$$Tr(B_{k+1}) \leqslant Tr(B_1) - \sum_{i=1}^k \frac{c_0^2}{a_i} + M_1k, \tag{2.15}$$

其中 $c_0 = -c^2(1-\phi_k)$.

由于 B_k 保持正定,所以上式可以推得

$$Tr(B_{k+1}) \leqslant Tr(B_1) + M_1k \leqslant [Tr(B_1) + M_1]k, \tag{2.16}$$

$$\sum_{i=1}^k \frac{1}{a_i} \leqslant \frac{Tr(B_1) + M_1k}{c_0^2} \leqslant \frac{[Tr(B_1) + M_1]k}{c_0^2} = c_1k. \tag{2.17}$$

根据几何 - 算术均值公式[12],从(2.17)式可以得到

$$\prod_{i=1}^k \frac{1}{a_i} \leqslant [\frac{\sum_{i=1}^k \frac{1}{a_i}}{k}]^k \leqslant [\frac{c_1k}{k}]^k = c_1^k,$$

所以,有

$$\prod_{i=1}^k a_i \geqslant (\frac{1}{c_1})^k = c_2^k. \tag{2.18}$$

再根据不等式 $Det(B_{k+1}) \leqslant [\frac{Tr(B_{k+1})}{n}]^n$[12],由(2.17)式得

$$Det(B_{k+1}) \leqslant [\frac{Tr(B_k) + M_1}{n}k]^n \leqslant c_3^k. \tag{2.19}$$

另一方面,将(2.8)式代入(2.3)式,有

$$Det(B_{k+1}) \geqslant Det(B_k)\frac{vm}{\alpha_k} \geqslant \cdots \geqslant Det(B_{k_1})\prod_{i=k_1}^k \frac{vm}{\alpha_i}, k > k_1,$$

亦即

$$\prod_{i=k_1}^k \alpha_i \geqslant \frac{Det(B_{k_1})}{Det(B_{k+1})}(vm)^{k-k_1+1}. \tag{2.20}$$

将上式与(2.18)式相乘,有

$$\prod_{i=1}^k \alpha_ia_i \geqslant (\prod_{i=1}^{k_1-1}\alpha_i)\frac{Det(B_{k_1})}{Det(B_{k+1})}(vm)^{k-k_1+1}c_2k,$$

然后再将(2.19)式代入,有

$$\prod_{i=1}^k \alpha_ia_i \geqslant (\prod_{i=1}^{k_1-1}\alpha_i)Det(B_{k_1})(vm)^{k-k_1+1}(\frac{c_2}{c_3})^k \geqslant c_4^k. \tag{2.21}$$

由引理 2.1,可知

$$\limsup_{k\to\infty} \alpha_ka_k > 0,$$

即存在 $\kappa \subseteq N, \varepsilon_2 > 0$ 使得

$$\alpha_ka_k \geqslant \varepsilon_2, \forall\ k \in \kappa. \tag{2.22}$$

由于 σ_2 是强迫函数,根据强迫函数的定义,易知存在常数 $\varepsilon_3 > 0$,使得成立

$$\sigma_2(\alpha_ka_k) \geqslant \varepsilon_3, \forall\ k \in \kappa.$$

根据引理 2.2 有

$$\sum_{k=0}^\infty \min\{\sigma_1(r_k), \varepsilon_3\} < \infty,$$

因此

$$\lim_{k\in\kappa}\sigma_1(r_k) = 0,$$

同样由于 σ_1 也为强迫函数,故有

$$\lim_{k\in\kappa} r_k = 0, \tag{2.23}$$

而由假设 $x_k \in D$,可知存在 M_2,使得 $\|s_k\| \leqslant \|x_{k+1}\| + \|x_k\| \leqslant M_2$ 及 $a_k = r_k\|d_k\|$,由此可得到

$$0 < \alpha_ka_k = \alpha_kr_k\|d_k\| = \|s_k\|r_k \leqslant M_2r_k,$$

结合(2.23)式,有结论

$$\lim_{k\in\kappa}\alpha_ka_k = 0,$$

与(2.22)式矛盾,故

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

定理 2.1 得证.

2.2 MBC1算法的全局收敛性

MBC2 算法是建立在目标函数 $f(x)$ 一致凸的前提下分析其收敛性的,对于 MBC1 算法,假设可以减弱,仅需假设 1、假设 2 及假设 4.

假设 4 公式(0.6)$A_k(1)$ 中的 μ_k 满足不等式

$$\lambda_1\|s_k\|^2 \leqslant s_k^T y_k + \mu_k|g_{k+1}^T(g_{k+1} - g_k)|\|s_k\| \leqslant \lambda_2\|s_k\|^2. \qquad (2.24)$$

实际上假设 4 是完全可行,例如:对于正数 λ_1,λ_2 定义

$$\phi = \{\lambda_1\|s_k\|^2 \leqslant \varphi|g_{k+1}^T(g_{k+1} - g_k)| \leqslant \lambda_2\|s_k\|,$$

选取

$$\mu_k = \begin{cases} 1, & \text{if } k \in \phi \\ (\omega\lambda_1 + (1-\omega)\lambda_2)\dfrac{\|s_k\|^2}{|g_{k+1}^T(g_{k+1} - g_k)|}, & \text{if } k \notin \phi' \end{cases}$$

其中 $\omega \in [0,1]$. 容易验证上述 μ_k 的取法满足假设 4.

对于 MBC1 算法我们有:

引理 2.8 设$\{x_k\}$由 MBC1 算法产生,则存在数 $M_3 > 0$ 使得

$$\frac{\|y_k^*\|^2}{y_k^{*T}s_k} \leqslant M_3. \qquad (2.25)$$

证明 由 y_k^* 的定义及(2.24)式,有

$$\|y_k^*\| = \|y_k + \mu_k\frac{|g_{k+1}^T(g_{k+1} - g_k)|}{\|s_k\|}s_k\| \leqslant L\|s_k\| + \mu_k|g_{k+1}^T(g_{k+1} - g_k)| \leqslant L\|s_k\| + \lambda_2\|s_k\| = (L+\lambda_2)\|s_k\|.$$

由(2.24)式容易知道

$$\lambda_1\|s_k\|^2 \leqslant s_k^T y_k^* \leqslant \lambda_2\|s_k\|^2.$$

因此(2.25)式成立,其中 $M_3 = \dfrac{(L+\lambda_2)^2}{\lambda_1}$.

再注意到不等式

$$y_k^{*T}s_k = y_k^T s_k + \mu_k|g_{k+1}^T(g_{k+1} - g_k)|\|s_k\| \geqslant y_k^T s_k \geqslant \cdots \geqslant (1-b-\varepsilon_1)(-g_k^T s_k)$$

成立,类似于引理 2.6 可得到:

引理 2.9 设$\{x_k\}$由 MBC1 算法产生,记

$$\eta_k = \phi_k(\frac{\|y_k^*\|\|s_k\|\cos\theta_k}{y_k^{*T}s_k})^2 - 2\phi_k\frac{y_k^{*T}B_k s_k\|s_k\|\cos\theta_k}{y_k^{*T}s_k\|B_k s_k\|},$$

则

$$\lim_{k\to\infty}\cos\theta_k = 0 \Rightarrow \lim_{k\to\infty}\eta_k\|g_k\| = 0.$$

利用以上 2 个引理,采用类似证明 MBC2 算法全局收敛性的方法,将得到 MBC1 算法的全局收敛性.

定理 2.2 设$\{x_k\}$由 MBC1 算法产生,其中 ϕ_k 满足(0.10)式,则有

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

3 结束语

本文利用一类 Wofle 类线搜索的 LS 搜索模型,与修改 Broyden 族相结合,得到 MBC1算法和 MBC2 算法。在适当条件下,MBC1算法和 MBC2具有全局收敛性,这推广了已有的结论。

参考文献:

[1] POWELL M J D. Some global convergence properties of a variable metric algorithm for minimization without exact line searches, in Nolinear Programming, SIAM-AMS proceedings: Vol IX [C]. R W COTTLE, C E LEMKE, eds. New York: American Mathematical Society Providence, RI, 1976.

[2] BYRD R, NOCEDAL J, YUAN Y. Global convergence of a class of quasi-Newton methods on convex problems [J]. SIAM Journal on Numerical Analysis, 1987(24): 1171-1189.

[3] BYRD R, NOCEDAL J. A tool for the analysis of quasi-Newton methods with application to unconstrained minimization[J]. SIAM Journal on Numerical Analysis, 1989(26): 727-739.

[4] 刘光辉,韩继业.带一类非精确搜索的 Broyden 族的全局收敛性[J].计算数学,1996(3):233-240.

[5] 刘光辉,尹红婷.BFGS 算法的全局收敛性分析[J].曲阜师范大学学报,1994(1):1-8.

[6] 柯小伍.Broyden 非凸族的收敛性[J].北京师范大学学报,1995(1):6-9.

[7] YUAN Ya-xiang, RICHARD, BYRD H. Non-quasi-Newton unconstrained optimization[J]. IMA Journal of Computational Mathematics, 1995(2): 95-107.

[8] LI D, FUKUSHIMA M. Amodified BFGS method and its global convergence in nonconvex minimization [J]. Journal of Computer Applied Mathematics, 2001(129): 15-35.

[9] WEI Z, YU G, YUAN G, et al. The superlinear convergence of a modified BFGS type method for unconstrained optimization[J]. Computational Optimization and Applications, 2004(25): 315-332.

[10] WEI Z, YU G. Some recent progress in unconstrained nonlinear optimization numerical linear algebra and optimization: proceedings of the 2003's International Conference on Numerical Optimization and Numerical Linear Algebra[C]. Yuan Y, eds. Beijing/New York: Science Press, 2004. 110-141.

[11] 王宜举,修乃华.非线性规划理论和算法[M].西安:陕西科学技术出版社,2004.

[12] 袁亚湘,孙文瑜.最优化理论和方法[M].北京:科学出版社,1997.

(责任编辑:韦廷宗 邓大玉)