Journal of Computational Mathematics, Vol.22, No.1, 2004, 11–20.

A CLASS OF REVISED BROYDEN ALGORITHMS WITHOUT EXACT LINE SEARCH *1)

Ding-guo Pu Sheng-hua Gui Wei-wen Tian
(*Department of Mathematics, Shanghai Tongji University, Shanghai 200331, China*)

Abstract

In this paper, we discuss the convergence of the Broyden algorithms with revised search direction. Under some inexact line searches, we prove that the algorithms are globally convergent for continuously differentiable functions and the rate of local convergence of the algorithms is one-step superlinear and n-step second-order for uniformly convex objective functions.

Key words: Variable metric algorithms, Line search, Convergence, Convergence rate.

1. Introduction

The Broyden family of algorithms remains a standard workhorse for minimization. These methods share the properties of finite termination on strictly convex quadratic functions, a superlinear rate of convergence on general strictly convex functions, and no need to store or evaluate the second derivative matrix. (see [2, 4, 1, 5, 6, 7]). However, there are several unsolved problems for the Broyden algorithms. In this paper we propose a new class of variable metric algorithms with revised search directions. We prove that the algorithms are convergent for the continuously differentiable objective functions. Also the new algorithms are superlinear and n-step second order convergent for uniformly convex functions when the line searches are inexact, but satisfy some search conditions.

These algorithms are iterative. Given a starting point x_1 and an initial positive definite matrix B_1, they generate a sequence of points $\{x_k\}$ and a sequence of matrices of $\{B_k\}$ which are given by following (1) and (2)

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

where $\alpha_k > 0$ is the step factor, d_k is the search direction satisfying

$$-d_k = H_k g_k + \|Q_k H_k g_k\| R_k g_k,$$

where g_k is the gradient of $f(x)$ at x_k, H_k is the inverse of B_k, $\{Q_k\}$ and $\{R_k\}$ are two sequences of positive definite matrices which are uniformly bounded. All eigenvalues of these matrices are included in $[q, r]$, $0 < q \leq r$, *i.e.*, for all k and $x \in R^n, x \neq 0$

$$q\|x\|^2 \leq x^T Q_k x \leq r\|x\|^2; \quad q\|x\|^2 \leq x^T R_k x \leq r\|x\|^2.$$

If $g_k = 0$, the algorithms terminate, otherwise let

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} + \phi(s_k^T B_k s_k) v_k v_k^T \tag{2}$$

* Received July 1, 2001; final revised March 16, 2003.
1) This research is supported by National Natural Science Foundation of China, Ministry of Education P. R. C and the Research and development foundation of Shanghai Education Commission.

where $y_k = g_{k+1} - g_k, v_k = y_k(s_k^T y_k)^{-1} - B_k s_k(s_k^T B_k s_k)^{-1}$ and $\phi \in [0,1]$. In the above algorithms, if $\phi = 0$ we call it revised BFGS algorithm, or RBFGS algorithm and if $\phi = 1$ we call it revised DFP algorithm, or RDFP algorithm.

The matrix H_{k+1} denotes the inverse of B_{k+1}, the recurrence formula of H_{k+1} is

$$H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{s_k s_k^T}{s_k^T y_k} + \frac{\rho \mu_k \mu_k^T}{y_k^T H_k y_k}, \tag{3}$$

where

$$\mu_k = H_k y_k - \frac{y_k^T H_k y_k}{s_k^T y_k} s_k \tag{4}$$

and $\rho \in [0,1]$, the relationship of ρ and ϕ is

$$\phi = \frac{(1-\rho)(s_k^T y_k)^2}{(1-\rho)(s_k^T y_k)^2 + \rho y_k^T H_k y_k s_k^T B_k s_k}.$$

In this paper, the line searches are not required to be exact. In order to guarantee descentness of the objective function values and the convergence of the algorithms, we must give some conditions for determining α_k. We use Wolfe conditions on line searches,

$$f(x_k) - f(x_{k+1}) \geq \zeta_0(-g_k^T s_k) \tag{5}$$

and

$$|g_{k+1}^T s_k| \leq \theta_0(-g_k^T s_k), \tag{6}$$

where ζ_0 and θ_0 be two constants satisfying $0 < \zeta_0 \leq \theta_0 < 1/2$. We always try $\alpha_k = 1$ first in choosing the step length.

Using the mathematical induction it is easy to imply that B_k and H_k are positive definite matrices if H_1 and B_1 are positive definite matrices.

If no ambiguities are arisen we may drop the subscript of the characters, for example, g, x, R denote g_k, x_k, R_k, and use subscript * to denote the amounts obtained by the next iteration, *i.e.*, g_*, x_*, R_* denote g_{k+1}, x_{k+1}, R_{k+1}, respectively.

For simplicity we let

$$\begin{aligned} U_k &= \frac{-g_k^T H_k y_k}{y_k^T H_k y_k};\ V_k = \frac{y_k^T H_k y_k}{s_k^T y_k}; \qquad W_k = \frac{-g_k^T d_k}{y_k^T d_k} = \frac{-g_k^T s_k}{s_k^T y_k}; \\ Z_k &= H_k g_k + \frac{-g_k^T H_k y_k}{s_k^T y_k} s_k \\ &= \frac{\|Q_k H_k g_k\| y_k^T R_k g_k}{s_k^T y_k} s_k - \|Q_k H_k g_k\| R_k g_k. \end{aligned} \tag{7}$$

The paper is outlined as follows: Section 2 gives several convergence results without the convexity assumption. Section 3 gives some results for convex objective functions. In Sections 4, we prove that the algorithms are linearly convergent for $\phi \in [0,1)$ in detail. In Section 5, we prove that our algorithms are one-step superlinearly convergent, then give the quadratical convergence of the algorithms without detail proof.

Throughout this paper the vector norms are Euclidian.

2. Results Without Convexity Assumption

In this section, we assume:

1. $f(x) \in C^{1,1}$, *i.e.*, there exists an $L > 1$ such that for any $x, y \in R^n$, $\|g(x) - g(y)\| \leq L\|x - y\|$.

2. For any $x_1 \in R^n$, the level set $S(x_1) = \{x | \, f(x) \leq f(x_1)\}$ is bounded.

3. We may assume, for simplicity (cf. [9]), $f(0) = \min f(x) = 0$.

By the properties of R and Q, these is a c_1 such that for all k

$$\|d\|(1 - c_1\|x\|) \leq \|Hg\| \leq \|d\|(1 + c_1\|x\|). \tag{8}$$

The following holds for all k.

$$-g^T s = \alpha[g^T Hg + \|QHg\| g^T Rg] \geq \frac{q^2\|g\|^2\|s\|}{1 + r^2\|g\|}. \tag{9}$$

Then the following theorem can be given (cf. [9]).

Theorem 2.1. *The algorithms are globally convergent, i.e.,*

$$\lim_{k\to\infty} g_k = 0. \tag{10}$$

By taking the trace of both sides of (2), we get

$$\mathrm{tr}(B_*) = \mathrm{tr}(B) + \frac{\|y\|^2}{s^T y} + \frac{\phi\|y\|^2 s^T Bs}{(s^T y)^2} - \frac{2\phi y^T Bs}{s^T y} - (1 - \phi)\frac{\|Bs\|^2}{s^T Bs}. \tag{11}$$

By taking the trace of both sides of (3), we obtain

$$\mathrm{tr}(H_*) = \mathrm{tr}(H) - \frac{\|Hy\|^2}{y^T Hy} + \frac{\|s\|^2}{s^T y} + \frac{\rho\|\mu\|^2}{y^T Hy}. \tag{12}$$

By calculating, we get (see [6])

$$\begin{aligned} H_* g_* &= [U + \rho(W - U)]\mu + Z + (1 - W)s \\ &= [(1 - \rho)U + \rho W]\mu + Z + (1 - W)s. \end{aligned} \tag{13}$$

Let

$$\eta_{1k}(\rho) = [(1 - \rho)U_k + \rho W_k]^{-1}, \tag{14}$$

then we get

$$\mu = \eta_1(\rho)[H_* g_* - Z - (1 - W)s], \tag{15}$$

where the subscript k of $\eta_{1k}(\rho)$ is dropped, and

$$Hy = \mu + Vs = \eta_1(\rho)[H_* g_* - Z - (1 - W)s] + Vs. \tag{16}$$

3. Results for Convex Functions

In this section, we assume:

1. The objective function $f(x)$ is uniformly convex and there exist M and m, $M \geq m > 0$, such that, for all $x, y \in R^n$, $m\|x\|^2 \leq x^T G(y)x \leq M\|x\|^2$, where $G(y)$ is the Hessian of $f(x)$ at y.

2. $G(x)$ satisfies the Lipschitz condition, *i.e.*, there exists an $L > 1$ such that, for all $x, y \in R^n$, $\|G(x) - G(y)\| \leq L\|x - y\|$.

For simplicity, we assume (cf. [9])

3. $f(0) = \min f(x) = 0$ *and* $G(0) = I_{n\times n}$, *i.e.*, the n^{-th} order unit matrix.

Let

$$c_2 = L\sqrt{\frac{M}{m}}(1+\frac{1}{m}), \qquad G_k = \int_0^1 G(x_k + ts_k)dt, \tag{17}$$

and let $(G)^{-1}$ denote the inverse of G. We get (see [6])

$$\max\{m;\ 1 - c_2\|x\|\} \le \frac{\|y\|^2}{s^T y} \le \min\{M;\ 1 + c_2\|x\|\}. \tag{18}$$

and

$$\max\left\{\frac{1}{M};\ 1 - c_2\|x\|\right\} \le \frac{\|s\|^2}{s^T y} \le \min\left\{\frac{1}{m};\ 1 + c_2\|x\|\right\}. \tag{19}$$

The Quasi-Newton $H_* y = s$ implies that $g_*^T s = g_*^T H_* y$ and

$$|g_*^T H_* s - (1-W)\|s\|^2| \le 2c_2\|x\|\|H_* g_*\|\|s\|. \tag{20}$$

(7) and (8) imply that there is a constant $c_3 > 0$ such that for all k,

$$\|Z\| \le c_3\|d\|\|x\|. \tag{21}$$

(15), (20) and (21) imply

$$|s^T\mu| \le \eta_1(\rho)\|x\|\|s\|(2c_2\|H_* g_*\| + c_3\|d\|). \tag{22}$$

By (16) and (22), we obtain

$$\begin{aligned}\frac{\|Hy\|^2 - \rho\|\mu\|^2}{y^T H y} &= \frac{V\|s\|^2}{s^T y} + \frac{(1-\rho)\|\mu\|^2}{y^T H y} + \frac{2s^T\mu}{s^T y} \\ &\ge V(1 - 2c_2\|x\|) + \frac{(1-\rho)\|\mu\|^2}{y^T H y} - \frac{2\eta_1(\rho)\|x\|\|s\|(2c_2\|H_* g_*\| + c_3\|d\|)}{s^T y}\end{aligned} \tag{23}$$

(21) implies $2|Z^T H_* g_*| \le c_3\|x\|(\|H_* g_*\|^2 + \|d\|^2)$, (15) and (22) imply that there exists a $c_4 > 0$ such that for all k,

$$\begin{aligned}\|\mu\|^2 \ \ge\ & \eta_1^2(\rho)[\|H_* g_*\|^2(1 - c_4\|x\|) - c_4\|x\|\|d\|^2 \\ & -(1-W)g_*^T s(1 + c_4\|x\|)].\end{aligned} \tag{24}$$

We may get the follows

$$d^T y(1 - c_5\|x\|) \le -g^T H y \le d^T y(1 + c_5\|x\|), \tag{25}$$

and

$$y^T H y \ge (1-\theta_0)^2(1 - 2c_5\|x\|)y^T d. \tag{26}$$

Without loss of generality, we may assume that (25)-(26) hold for all k. Substituting (24) and (26) into (23), then given any $\rho \in [0,1)$ and $\gamma \in (0, 1-\rho]$, we know that there exists $c_6 = c_6(\gamma,\rho) > 0$ such that for all k,

$$\begin{aligned}\frac{\|Hy\|^2 - \rho\|\mu\|^2}{y^T H y} \ &\ge V(1 - c_6\|x\|) + \frac{\gamma\eta_1(\rho)^2\|H_* g_*\|^2(1 - c_6\|x\|)}{y^T H y} \\ & - \frac{\gamma\eta_1(\rho)^2(1-W)g_*^T s(1 + c_6\|x\|)}{y^T H y} - c_6\|x\|.\end{aligned} \tag{27}$$

Lemma 3.1. *Let $\{D_k\}$ be a sequence of positive numbers, and let t_1, t_2, t_3, t_4 and t_5 be positive numbers. If the following holds for all k:*

$$t_1 + \sum_{j=1}^{k} D_j(1 - t_2\|x_j\|) \le t_3 + t_4 k + \sum_{j=1}^{k} t_5\|x_j\|,$$

then there exists a positive number $t_6 > 0$ such that for all k,

$$t_1 + \sum_{j=1}^{k} D_j \le t_4 k + \sum_{j=1}^{k} t_6\|x_j\|. \tag{28}$$

4. Linear Convergence

In this section, we assume the assumptions 1-3 in Section 3 hold. We discuss the linear convergence of our algorithms for $\phi \in [0, 1)$. The following Lemma 4.1 holds for any $\phi \in [0, 1)$.

Lemma 4.1. *Given any $\phi \in [0, 1)$, there exists $c_7 > 0$ such that, for all k,*

$$tr(B_{k+1}) + \sum_{j=1}^{k} \frac{(1-\phi)\|B_j s_j\|^2}{s_j^T B_j s_j} + \frac{\phi s_j^T B_j s_j}{s_j^T y_j} \le k + \sum_{j=1}^{k} c_7\|x_j\|. \tag{29}$$

Proof. we have

$$2y^T Bs - \frac{\phi\|y\|^2 s^T Bs}{s^T y} \ge s^T Bs - \frac{3c_2\|x\|\|Bs\|^2 s^T y}{s^T Bs}, \tag{30}$$

then the following is implied by (11) and (30).

$$\mathrm{tr}(B_*) + \frac{(1-\phi-3c_2\|x\|)\|Bs\|^2}{s^T Bs} + \frac{\phi s^T Bs}{s^T y} \le \mathrm{tr}(B) + 1 + c_2\|x\|. \tag{31}$$

Adding both sides of (31) over $j = 1, 2, \cdots, k$, we get

$$\begin{aligned} & \mathrm{tr}(B_{k+1}) + \sum_{j=1}^{k} \left[\frac{(1-\phi-2c_2\|x_j\|)\|B_j s_j\|^2}{s_j^T B_j s_j} + \frac{\phi s_j^T B_j s_j}{s_j^T y_j} \right] \\ \le \;& \mathrm{tr}(B_1) + k + \sum_{j=1}^{k} c_2\|x_j\|. \end{aligned} \tag{32}$$

Lemma 4.1 can be implied by Lemma 3.1 and (32).

Lemma 4.2. *For ϕ=0, there exists a constant $c_8 > 0$ such that for all k,*

$$tr(H_{k+1})/2 + \sum_{j=1}^{k} V_j \le k + \sum_{j=1}^{k} c_8\|x_j\|. \tag{33}$$

Proof. (8) implies that there exists a constant $c_9 > 0$ such that for all k and $\phi = 0$ $(\rho = 1)$,

$$\begin{aligned} & \alpha_*\|H_* g_*\|^2 \le \alpha_*(1 + c_0\|x\|)^2\|d_*\|^2 \le m^{-1}(1 + c_0\|x\|)^2 y_*^T d_* \\ \le \;& c_9(g_*^T H_* g_* + \|H_* g_*\|\|s\|). \end{aligned} \tag{34}$$

So, we get

$$\|H_*g_*\|\|s\| \le \frac{g_*^T H_* g_*}{2} + \frac{\|H_*g_*\|\|s\|}{2} + \frac{c_9\|s\|^2}{2\alpha_*} \le g_*^T H_* g_* + \frac{c_9\|s\|^2}{\alpha_*}. \tag{35}$$

On the other hand, taking the inner product of both sides of (13) with g_* for $\rho = 1$, we obtain, by (6) and (25), that

$$\begin{aligned}\frac{g_*^T H_* g_*}{s^T y} &= \frac{-g^T s g_*^T H y}{(s^T y)^2} - \frac{(-g^T s) y^T H y g_*^T s}{(s^T y)^3} + \frac{(g_*^T s)^2}{(s^T y)^2} + \frac{Z^T g_*}{s^T y} \\ &\le (1-\theta_0)^{-1}(1 + c_5\|x\|)\alpha^{-1} + (1-\theta_0)^{-2} V + 1 + \frac{Z^T g_*}{s^T y}. \end{aligned}\tag{36}$$

(21) implies that there exists a constant $c_{10} > 0$ such that for all k

$$|Z^T g_*| \le c_3\|x\|\|d\|\|g_*\| \le c_{10}\|s\|\|d\|.$$

By (26), we may obtain the following for sufficiently large k

$$V \ge \frac{1}{4\alpha} \tag{37}$$

and $\eta_1(1) = W^{-1} < 1 + \theta_0$ (cf. the definition in (14)). Combining (35)-(37), we get for all k,

$$\frac{2\eta_1(1)\|x\|\|s\|(2c_2\|H_*g_*\| + c_3\|d\|)}{s^T y} \le c_{11}\|x\|(V + V_* + 1), \tag{38}$$

where $c_{11} > 0$ is a constant. Substituting (23) and (38) into (12), we obtain for all k,

$$\mathrm{tr}(H_*) + V(1 - 2(c_2 + c_{11})\|x\|) - c_{11}\|x\| - \frac{c_{11}\|x\|}{\alpha^*} \le \mathrm{tr}(H) + 1 + c_2\|x\|. \tag{39}$$

The following holds for sufficiently large k:

$$\mathrm{tr}(H) \ge \frac{y^T H y}{\|y\|^2} \ge \frac{V}{1 + c_2\|x\|} \ge \frac{1}{4\alpha(1 + c_2\|x\|)} \ge \frac{1}{8\alpha}. \tag{40}$$

Without loss of generality, we may assume that (40) and $1 - 4(2c_2 + 4c_{11})\|x\| > 0$ hold for all k. Adding both sides of (39) for $j = 1, 2, \cdots, k$, then, by (39) and (40), we obtain

$$\begin{aligned} & \mathrm{tr}(H_{k+1})/2 + \sum_{j=1}^{k} V_j[1 - 2c_2\|x_j\| - 2c_{11}(\|x_j\| + \|x_{j+1}\|)] \\ \le\ & \mathrm{tr}(H_1) + \sum_{j=1}^{k}(c_2 + 4c_{11})\|x_j\|. \end{aligned}\tag{41}$$

Lemma 3.1 and (41) implies that the lemma is true.

Lemma 4.3. *There exists a $c_{12} > 0$ such that for all k, and $\phi \in [0, 1)$*

$$\sum_{j=1}^{k} \alpha_j^{-1} \le c_{12} k.$$

Proof. Lemma 4.2 implies that this lemma holds for $\rho = 1$ (or $\phi = 0$). So we only need to prove this result for $\rho \in (0,1)$ or $\phi \in (0,1)$. By the definition (14) of $\eta_1(\rho)$, (6), (25) and (26), we know that there exists a constant $c_{13} > 0$ such that for sufficiently large k and given $0 < \rho < 1$,

$$0 < \frac{\eta_1(\rho)(-g^T Hy)}{y^T Hy} \le \frac{1}{1-\rho} \le c_{13} \tag{42}$$

and

$$0 < \eta_1(\rho) \le \frac{s^T y}{\rho(-g^T s)} \le \frac{(1+\theta_0)}{\rho} \le c_{13}. \tag{43}$$

(25) (42) and (43) imply that there exists a constant $c_{14} > 0$ such that for sufficiently large k and any given $\rho \in (0,1)$,

$$\frac{\eta_1^2(\rho)(1-W)g_*^T s(1+c_6\|x\|)}{y^T Hy} \le \frac{(1+\theta_0)(-g^T s)^2(1+c_6\|x\|)}{\rho(1-\rho)(-g^T Hy)(s^T y)} \le c_{14}\alpha.$$

So for given $\rho \in (0,1)$ and $\gamma \le 1/(4c_{14})$, (27) implies that there exists a $c_6 > 0$ such that

$$\frac{\|Hy\|^2 - \rho\|\mu\|^2}{y^T Hy} \ge V(1-c_6\|x\|) - c_6\|x\| - \alpha/4. \tag{44}$$

Substituting (44) into (12), then add both sides of the resulting expressions for $j = 1, 2, \cdots, k$ we get

$$\begin{aligned} & \operatorname{tr}(H_{k+1}) + \sum_{j=1}^{k}[V_j(1-c_6\|x_j\|) - c_7\|x_j\| - \alpha_j/4] \\ \le \quad & \operatorname{tr}(H_1) + \sum_{j=1}^{k} c_2\|x_j\| + k. \end{aligned} \tag{45}$$

Because $\|B_j s_j\|\|s_j\| \ge s_j^T B_j s_j$ and $(s_j^T B_j s_j/\alpha_j)^2 \ge s_j^T B_j s_j g_j^T H_j g_j \ge (s_j^T y_j/2)^2$, add both sides of (29) and (45) respectively, the following is obtained

$$\begin{aligned} & \sum_{j=1}^{k}[V_j(1-c_6)\|x_j\| + (1/2 - c_2\|x_j\| - 1/4)\alpha_j] \\ \le \quad & \operatorname{tr}(H_1 + B_1) + 2k + \sum_{j=1}^{k}(c_2 + c_6 + c_7)\|x_j\|. \end{aligned} \tag{46}$$

Clearly, this lemma holds by Lemma 3.1 and (46).

Because the method which gives the recurrence formula of B_{k+1} from B_k is the same as the Broyden update formula, we may use the proof of Lemma 4.2 in Byrd et al (1987) almost word by word almost to prove the following Theorem 4.1 (cf. [8, 9]).

Theorem 4.1. *There exists a constant δ, $0 < \delta < 1$, such that, for $\phi \in [0,1)$ and sufficiently large k,*

$$f(x_{k+1}) \le \delta^k f(x_1). \tag{47}$$

5. Superlinear Convergence

In this section we assume the assumptions 1-3 in Section 3 hold. First, we discuss the superlinear convergence of the algorithms for $\phi \in [0,1]$. Lemma 4.1 implies that the following lemma holds for any $\phi \in [0,1)$ immediately.

Lemma 5.1. *Given any there exists a constant $c_{15} > 0$ such that, for all k,*

$$tr(B_{k+1}) + \sum_{j=1}^{k} \frac{s_j^T B_j s_j}{s_j^T y_j} \le k + c_{15}. \tag{48}$$

Lemma 5.2. *There exists a $c_{17} > 0$ such that for all k,*

$$tr(H_{k+1}) + \sum_{j=1}^{k} V_j \le k + c_{17}. \tag{49}$$

Proof. (40) infers that, for sufficiently large k, $\|d\| \le 8\text{tr}(H)\|s\|$ and $\|H_* g_*\| \le \text{tr}(H_*)\|g_*\| \le \text{tr}(H_*)\text{tr}(B)(M/m)^{3/2}\|Hg\|$. So Lemma 4.1 implies that given $\rho \in [0,1)$, there exists a constant $c_{18} > 0$ such that for all k,

$$\sum_{j=1}^{\infty} \frac{\eta_{1j}(\rho)\|x_j\|\|s_j\|(2c_2\|H_{j+1}g_{j+1}\| + c_3\|d_j\|)}{s_j^T y_j} = c_{18}. \tag{50}$$

By (23) and (50), then adding both sides over $j = 1, 2, 3, \cdots, k$, we get

$$\text{tr}(H_{k+1}) + \sum_{j=1}^{k} V_j(1 - c_2\|x_j\|) - c_{18} \le \text{tr}(H_1) + k + \sum_{j=1}^{k} c_2\|x_j\|.$$

Now it is easy to see that the lemma holds.

Theorem 5.1. *The algorithms presented in this paper are one-step superlinearly convergent for uniformly objective functions, i.e.,*

$$\lim_{k\to\infty} \|g_{k+1}\|/\|g_k\| = 0.$$

Proof. Adding both sides of (48) and (49) respectively, we get

$$\text{tr}(B_{k+1} + H_{k+1}) + 2\sum_{j=1}^{k} \left[\frac{(y_j^T H_j y_j)^{1/2}(s_j^T B_j s_j)^{1/2}}{s_j^T y_j} - 1 \right]$$
$$+ \quad \frac{[(y_j^T H_j y_j)^{1/2} - (s_j^T B_j s_j)^{1/2}]^2}{s_j^T y_j} \le c_{15} + c_{17}, \tag{51}$$

As $y_j^T H_j y_j s_j^T B_j s_j \ge (s_j^T y_j)^2$, we have

$$\text{tr}(H_{k+1} + B_{k+1}) \le c_{15} + c_{17}$$

and

$$\lim_{k\to\infty} \left| \frac{y_k^T H_k y_k}{s_k^T y_k} - \frac{s_k^T B_k s_k}{s_k^T y_k} \right| = 0. \tag{52}$$

By (52), we can get (cf. [8])

$$d_k^T g(x_k + d_k) = o(\|d_k\|^2) = o(-g_k^T d_k) \tag{53}$$

and

$$\begin{aligned} & f(x_k) - f(x_k + d_k) \\ \geq \; & \frac{\|d_k\|^2 - L\|d_k\|^3}{2} - d_k^T g(x_k + d_k) \geq \zeta_0 g_k^T H_k g_k. \end{aligned} \tag{54}$$

(53) and (54) show that $\alpha_k = 1$ must satisfy (5) and (6) for sufficiently large k. So we can take $\alpha \equiv 1$ for sufficiently large k. (53) and (52) imply

$$\lim_{k\to\infty} \frac{\|g_{k+1}\|}{\|g_k\|} = 0.$$

This completes the proof of Theorem 5.1.

We list the superlinear convergence of the RDFP algorithm and the quadratic convergence of the RBroyden algorithms without detail proof.

Theorem 5.2. *If we replace η_0 by $\eta_k = \eta_0 \min\{1, 1/\alpha_k\}$ in line search conditions (5) and (6) then the RDFP algorithm is one step superlinearly convergent.*

Theorem 5.3. *If we replace η_0 by $\eta_k = \eta_0 \min\{1, \|g_k\|\}$ in line search conditions (5) and (6) then the algorithms presented in this paper are n-step quadratically convergent.*

We have done some preliminary computational experiments for the revised Broyden algorithms, which indicates that these algorithms are quite promising. Some testing results are listed in Table 1, in which LS=the number of line searches, NF= the number of function evaluation, MGH is the function in [3]. Parameters and other testing results are referred in [8, 9].

Table 1

Problem	Reference	BFGS		RBFGS	
		LS	NF	LS	NF
Helical and Valley3	MGH1	31	35	27	32
Powell badly Scaled2	MGH4	159	193	120	149
Wood4	MGH17	88	107	69	93
Box three-dimensional	MGH5	18	38	20	32
Waton6	MGH7	37	39	25	38
Trigonometric of Spedicato10	MGH13	55	67	36	61
Gaussian3	MGH3	8	8	9	10
Chebyquad8	MGH18	23	31	20	27

References

[1] Byrd R. H., J. Nocedal and Y. Yuan, Global convergence of a class of quasi-Newton methods on convex problems, *SIAM. J. Numer. Anal.*, **24** (1987), 1171-1190.

[2] Powell M. J. D., On the convergence of the variable metric algorithm, *Inst. Math. Appl.*, **7** (1971), 21-36.

[3] More, J. J., B. S. Garbow and K. E. Hillstrom, Testing unconstrained optimization software, *ACM transactions on mathematical software,* **7** (1981), 17-41.

[4] Powell M. J. D., Some global convergence properties of the variable metric algorithm for minimization without exact line searchers, SIAM-AMS Proceedings, Vol 6 R.W. Cottle and C.E. Lemke eds. Nonlinear Programming, SIAM Publications (1976).

[5] Pu D. and W. Yu, On the convergence property of the DFP algorithm, *J. of Ann. of Oper. Res.,* **24** (1990), 175-184.

[6] Pu D, A class of DFP algorithm without exact linear Search , *Asia-Pacific J. of Oper. Res. ,* **9** (1992), 207-220.

[7] Pu D., The convergence of DFP algorithm without exact linear search, *J. Opt. Theory Appl.,* **112** (2002), 187-211.

[8] Pu D. and W. Tian, A class of modified Broyden algorithms , *J. of Computational Mathematics,* **8** (1994), 366-379.

[9] Pu D. and W. Tian, A class of modified Broyden algorithms without exact line search, *Applied Mathematics - A J. of Chinese Universities,* **10** (1995), 313-322.