Research Journal of Applied Sciences, Engineering and Technology 5(15): 3968-3974, 2013
DOI:10.19026/rjaset.5.4461
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: October 22, 2012 Accepted: November 19, 2012 Published: April 25, 2013

Research Article

Global Convergence of a New Nonmonotone Algorithm

Jing Zhang
Basic Courses Department of Beijing Union University, Beijing 100101, China

Abstract: In this study, we study the application of a kind of nonmonotone line search in BFGS algorithm for solving unconstrained optimization problems. This nonmonotone line search is belongs to Armijo-type line searches and when the step size is being computed at each iteration, the initial test step size can be adjusted according to the characteristics of objective functions. The global convergence of the algorithm is proved. Experiments on some well-known optimization test problems are presented to show the robustness and efficiency of the proposed algorithms.

Keywords: Global convergence, nonmonotone line search, unconstrained optimization

INTRODUCTION

Unconstrained optimization problems:

$$\min f(x), \quad x \in \mathbf{R}^n \tag{1}$$

The quasi-Newton algorithm BFGS method because of its stable numerical results and fast convergence is recognized as one of the most effective methods to solve the unconstrained problem (1). Iterative formula of this method is as follows:

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

$$d_k = -B_k^{-1} g_k \quad (k > 1)$$

$$d_1 = -g_1 \tag{2}$$

where, α_k is step length, $g_k = \nabla f(x_k), d_k$ is the search direction:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} \tag{3}$$

where, $s_k = x_{k+1} - x_k$, $y_k = g_{k+1} - g_k$. Inexact line search, especially in the monotone line search method global convergence of many results. Byrd *et al*. (1987) proved in addition to the DFP method of Broyden family Wolfe line search, global convergence for solving convex minimization problem. Byrd and Nocedal (1989) proved the global convergence of the BFGS method Armijo line search for solving convex minimization problem. Sun and Yuan (2006) constructed a counter-example to show that the BFGS method under the Wolfe line search for non-convex minimization problem does not have global convergence. Since 1986, Grippo *et al*. (1986) first proposed a non-monotonic linear search technology has broader means non-exact line search. One benefit of the technology of non-monotonic is does not require the function value decreases, so that the step the selection of a more flexible, even with step as large as possible. Panier and Tits (1991) proved that a nonmonotonic search technology to avoid Maratos effect. A large number of numerical results show that non-monotonic search is better than the monotonous search numerical performance; in particular, it helps to overcome along the bottom of narrow winding produces slow convergence of iterative sequence (Dai, 2002a; Hüther, 2002). Quasi-Newton method to introduce the nonmonotonic technology also has its practical significance, but not more discussion on global convergence. Han and Liu (1997) prove that the nonmonotone Wolfe modified linear search, BFGS method global convergence of convex objective function. Since the beginning of this century, new non-monotone line search methods continue to put forward, such as the Zhang and Hager (2004) and Zhen-Jun and Jie (2006) proposed a new non-monotone line search method.

NONMONOTONIC LINE SEARCH

This study a class of non-monotonic linear search is belongs to the Armijo type of linear search the ideological sources. Dai (2002c) proposed a class of monotone line search him and conjugate gradient method combined study. Dai (2002b) monotonous line search is: find α_k, so that the following two formulas:

$$f(x_k + \alpha_k d_k) - f(x_k) \le \delta \alpha_k g_k^T d_k$$

and

$$0 \neq g_{k+1}^T d_{k+1} \leq -\sigma \| d_{k+1} \|^2$$

At the same time set up. This study $\|.\|$ refers to the Euclidean norm. In this study, the two equations improvements for weaker conditions, further transformed into a linear search of non-monotonic and the BFGS algorithm combined into a class of quasi-Newton algorithm. Nonmonotone linear search of the text of the study also draws (Zhen-Jun and Jie, 2006), the line search in each step to calculate the step length factor α_k. When to introduce timely changes in the initial test step r_k, instead follows the Grippo-Lampariello-Lucidi search in the fixed initial test step. If the initial test step is fixed, the non-monotonic class of linear search is essentially a class of linear search without derivative. Its monotonous situation, initially by Leone *et al*. (1984) study, but there in the form of relatively complex; such line search form research this study are concise and full of operability.

Given σ>0, $\beta \in (0,1), \delta \in (0,1)$, M is a non-negative integer and to let $r_k = -\frac{\sigma g_k^T d_k}{\|d_k\|^2}$. Take $\alpha_k = \beta^{m(k)r_k}, m(k) = 0, 1, 2, \ldots m(k)$makes holds the smallest non-negative integer:

$$f(x_k + \alpha_k d_k) \leq \max_{0 \leq j \leq l(k)} f(x_{k-j}) - \delta \| \alpha_k d_k \|^2 \tag{4}$$

where, $l(0) = 0$, $0 \leq l(k) \leq \min\{l(k-1)+1, M\}$, $k \geq 1$. Obviously, if d_k is a descent direction, that $d_k^T g_k < 0$, then, when the m(k) sufficiently large, the inequality (4) is always true, thus satisfying the conditions α_k exist. The line search in each iteration, the initial test step length is no longer maintained constant, but can be automatically adjusted to r_k. Global convergence proof which will see the reasonableness of this proposal. In magnitude, change the initial test step length of practice better results can be obtained to calculate the larger step length factor α_k, thereby reducing the number of iterations.

ALGORITHM

Based on this non-monotonous line search technique, we give the nonmonotonic following BFGS algorithm.

1° given initial point x_0, initial matrix B_0 = I (Unit matrix). Given constant σ>0, $\beta \epsilon (0,1), \delta \epsilon (0,1)$ and non-negative integer M. Given iteration terminate error ε. Let k: = 0. Calculate g_k. If $\|g_k\| < \varepsilon$, it is terminated, x_k is what we seek; otherwise, go to 2°.

2°solution of linear equations calculates the direction of the search d_k:

$$B_k d_k + g_k = 0 \tag{5}$$

3°step factor α_k calculated according to the non-monotonic linear search NLS.

4° let $x_{k+1} = x_k + \alpha_k d_k$

5° calculated g_{k+1}, if $\|g_{k+1}\| < \varepsilon$, it is terminated, x_{k+1} is the demand; otherwise, according to the BFGS correction formula (3) to give B_{k+1}.

6° let $k \leftarrow k+1$, go to 2°

Remark:

- In order to step factor calculated in Step 3 to take advantage of the non-monotonic linear search NLS α_k must make the search direction d_k descent direction, by (5), only to meet $g_k^T = d_k = -g_k^T B_k^{-1} g_k < 0$, This requires nonmonotonic line search NLS based on research in this chapter, every step BFGS correction formula B_{k+1} is positive definite, it see Theorem 1.
- In order to be able to calculate a larger step length factor α_k, we can consider a class of mixed non-monotone line search that (4) can be rewritten as:

$$f(x_k + \alpha_k d_k) \leq \max_{0 \leq j \leq l(k)} f(x_{k-j}) - \max\{\delta_1 \| \alpha_k d_k \|^2, \delta_2 \alpha_k g_k^T d_k\}$$

where, $\delta_1, \delta_2 \in (0, 1)$

Theorem and proof: Nonmonotonic line search global convergence proof often needs to meet:

(i) **The Sufficient descent conditions:** $g_k^T d_k \leq -C_1 \| g_k \|^2$

(ii) **Boundedness conditions:** $\|d_k\| \leq C_2 \|g_k\|$, where C_1 and C_2 is a positive number. These two conditions are strong, difficult to meet the general quasi-Newton method. This is also the problem of study difficulty.

This study is not BFGS formula to make any changes and non-monotone line search NLS1, does not require the search direction d_k satisfy the conditions (i)-(ii) of the premise, to prove the global convergence. The general assumption in this section is given below:

H1: $f(x)$ order continuous differentiable.

H2: The level set $L_0 = \{x | f(x) \leq f(x_0), x \in R^n\}$ is convex sets and there is c_1>0:

$$c_1 \| z \|^2 \le z^T G(x) z\ ,\ \forall x \in L_0\ ,\ \forall z \in \mathbf{R}^n \tag{6}$$

where, $G(x) = \nabla^2 f(x)$.

The above assumptions with the references (Byrd and Nocedal, 1989) the same, paper (Liu *et al*., 1995) weakened, in particular, is to remove the literature (Grippo *et al*., 1986) desired search direction d_k satisfy the sufficient descent condition and boundedness conditions.

Assumptions (H1) and (H2) conditions, we easily obtain the following results:

- The level set $L_0 = \{x | f(x) \le f(x_0),\ x \varepsilon R^n\}$ is bounded closed set. (Proof may see (Sun and Yuan, 2006))
- The function $f(x)$ in the level set L_0 bounded and uniformly continuous
- $[g(x)-g(y)]^T(x-y)$

$$\ge c_1 \| x - y \|^2\ ,\ \forall x\ ,\ y \in L_0\ ,$$

where, $g(x) = \nabla f(x), c_1 > 0$ is a constant, such as assuming that (H2) as defined.

Proof: To be a vector-valued function g use of the integral form of the mean value theorem may:

$$g(x) - g(y) = \int_0^1 G(y + \theta(x-y))(x-y)\mathrm{d}\theta \cdot$$

Thus, by (6), there exists $c_1>0$, making the:

$$\begin{aligned}&[g(x)-g(y)]^T (x-y)\\ &= \int_0^1 (x-y)^T G(y+\theta(x-y))\mathrm{d}\theta \cdot (x-y)\\ &\ge \int_0^1 c_1 \| x-y \|^2\ \mathrm{d}\theta\\ &= c_1 \| x-y \|^2 .\end{aligned}$$

Lemma: 1 Let B_k is symmetric positive definite matrix, $s_k^T y_k > 0$, Broyden family formula:

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

where, $v_k = \frac{y_k}{y_k^T s_k} - \frac{B_k s_k}{s_k^T B_k s_k}$, when $\emptyset_k \ge 0, B_{k+1}^{\emptyset}$ maintaining positive definite.

Theorem 1: On the assumption that (H1), (H2) Conditions, the step factor α_k by nonmonotonic line search NLS B_0 is symmetric positive definite matrix, then when $\emptyset_k \ge 0$, by the correction formula of Broyden family B_k also maintained positive definite.

Proof: Proved by induction. When $k = 0$, B_0 is symmetric positive definite matrix, the resulting search direction d_0 downward direction. By nonmonotonic line search can find α_0, resulting in x_1, it is seen from (4):

$$f(x_1) \le f(x_0) - \delta \| \alpha_0 d_0 \|^2 < f(x_0)$$

As a result, $x_1 \in L_0$. Thus, by Lemma 2.3:

$$\begin{aligned}&s_0^T y_0\\ &= (g_1 - g_0)^T (x_1 - x_0)\\ &\ge c_1 \| x_1 - x_0 \|^2\\ &> 0\end{aligned}$$

By Lemma 1, B_1 also maintained positive definite:

Assume B_k is positive definite, so the resulting search direction d_k is down direction. Can be found by the non - monotone line search NLS α_k, resulting x_{k+1}, it is seen from (4):

$$\begin{aligned}&f(x_2) \le \max_{0\le j\le l(1)} f(x_{1-j}) - \delta \| \alpha_1 d_1 \|^2 < f(x_0)\ ,\\ &f(x_3) \le \max_{0\le j\le l(2)} f(x_{2-j}) - \delta \| \alpha_2 d_2 \|^2 < f(x_0)\ ,\\ &\quad \cdots\cdots\\ &f(x_{k+1}) \le \max_{0\le j\le l(k)} f(x_{k-j}) - \delta \| \alpha_k d_k \|^2 < f(x_0) \cdot\end{aligned}$$

As a result, $x_k \in L_0$. Thus, by (c),

$$\begin{aligned}&s_k^T y_k\\ &= (g_{k+1} - g_k)^T (x_{k+1} - x_k)\ .\\ &\ge c_1 \| x_{k+1} - x_k \|^2\\ &> 0\end{aligned}$$

By Lemma 1, B_{k+1} also maintained positive definite.

For the sake of simplicity, we have introduced the notation:

$$h(k) = \max\{ i \mid 0 \le k - i \le l(k)\ ,\ f(x_i) = \max_{0\le j\le l(k)} f(x_{k-j})\}$$

Namely h(k) is a non-negative integer and satisfies the following two formulas:

$$k - l(k) \le h(k) \le k\ , \tag{7}$$

$$f(x_{h(k)}) = \max_{0\le j\le l(k)} f(x_{k-j}) \cdot \tag{8}$$

Thus, the nonmonotonic line search NLS (4) can be rewritten as:

$$f(x_{k+1}) \le f(x_{h(k)}) - \delta \| \alpha_k d_k \|^2 . \tag{9}$$

Lemma 2: Under the conditions of assumption (H1), the sequence $\{f(x_{h(k)})\}$ decreases monotonically.
Proof: By (9), knowledge of all k:

$$f(x_{k+1}) \le f(x_{h(k)}) \tag{10}$$

have been established.

Nonmonotonic line search NLS, 0≥l(k)≤l(k–1)+1, therefore:

$$\begin{aligned} & f(x_{h(k)}) \\ &= \max_{0\le j\le l(k)} f(x_{k-j}) \\ &\le \max_{0\le j\le l(k-1)+1} f(x_{k-j}) \\ &= \max\{\max_{0\le j\le l(k-1)} f(x_{k-1-j}),\, f(x_k)\} \\ &= \max\{f(x_{h(k-1)}),\, f(x_k)\} \end{aligned}$$

Then from (10), to give:

$$\begin{aligned} & f(x_{h(k)}) \\ &\le \max\{f(x_{h(k-1)}),\, f(x_k)\} \\ &= f(x_{h(k-1)}) \end{aligned}$$

Lemma 3: Assuming (H1) holds, then the limit $\lim_{k\to\infty}$ f(xh(k)) exists and

$$\lim_{k\to\infty} \alpha_{h(k)-1} \| d_{h(k)-1} \| = 0 \tag{11}$$

Proof By (b) knowledge f(x) in the level set L_0 on the lower bound, $\{x_k\} \subset L_0$ (See the proof of Theorem 1) and the sequence $\{f(x_{h(k)})\}$ monotonically decreasing, $\lim_{k\to\infty}$ f(xh(k)) exist. From (9), there are

$$f(x_{h(k)}) \le f(x_{h(h(k)-1)}) - \delta \| \alpha_{h(k)-1} d_{h(k)-1} \|^2$$

On both sides so that k→ and notes δ>0, so

$$\lim_{k\to\infty} \| \alpha_{h(k)-1} d_{h(k)-1} \|^2 = 0$$

namely formula (11).

Lemma 4: Under the assumptions (H1), (H2) of the condition, $\lim_{k\to\infty} \alpha_{k\|d_k\|} = 0$).

Proof Let $\hat{h}(k) = h(k + M + 2$, First proved by mathematical induction, for any i≥1, the following holds:

$$\lim_{k\to\infty} \alpha_{\hat{h}(k)-i} \| d_{\hat{h}(k)-i} \| = 0 \tag{12}$$

$$\lim_{k\to\infty} f(x_{\hat{h}(k)-i}) = \lim_{k\to\infty} f(x_{h(k)}) \tag{13}$$

When I = 1, by $\hat{h}$ definition, apparently $\{\hat{h}(k)\} \subset \{h(k)\}$. Thus, by Lemma 3, $\lim_{k\to\infty} f(x_{\hat{h}(k)})$ exists and

$$\lim_{k\to\infty} f(x_{\hat{h}(k)}) = \lim_{k\to\infty} f(x_{h(k)}) \tag{14}$$

By (11), known (12) was established.

$$\begin{aligned} & x_{\hat{h}(k)} - x_{\hat{h}(k)-1} \\ &= s_{\hat{h}(k)} \\ &= \alpha_{\hat{h}(k)-1} d_{\hat{h}(k)-1} \end{aligned}$$

This indicates that:

$$\| x_{\hat{h}(k)} - x_{\hat{h}(k)-1} \| \to 0 \, (k \to \infty)$$

and then by f(x) in L_0 uniformly continuous, so:

$$\begin{aligned} & \lim_{k\to\infty} f(x_{\hat{h}(k)-1}) \\ &= \lim_{k\to\infty} f(x_{\hat{h}(k)}) \\ &= \lim_{k\to\infty} f(x_{h(k)}) \end{aligned}$$

i.e., (13) established for i = 1:

Now suppose for a given i, (12) and (13). From (9), there are

$$f(x_{\hat{h}(k)-i}) \le f(x_{h(\hat{h}(k)-i-1)}) - \delta \| \alpha_{\hat{h}(k)-i-1} d_{\hat{h}(k)-i-1} \|^2,$$

On both sides so that k→∞, by (13) and:

$$\lim_{k\to\infty} f(x_{h(\hat{h}(k)-i-1)}) = \lim_{k\to\infty} f(x_{h(k)}),$$

And notes δ>0, so:

$$\lim_{k\to\infty} \alpha_{\hat{h}(k)-(i+1)} \| d_{\hat{h}(k)-(i+1)} \| = 0\,. \tag{15}$$

This indicates that, for any i≥1, (12) was established.
The (15) also implies:

$$\| x_{\hat{h}(k)-i} - x_{\hat{h}(k)-(i+1)} \| \to 0 \ (k \to \infty),$$

Due to f(x) in the level set L_0 is uniformly continuous and thus

$$\begin{aligned} & \lim_{k\to\infty} f(x_{\hat{h}(k)-(i+1)}) \\ &= \lim_{k\to\infty} f(x_{\hat{h}(k)-i}) \\ &= \lim_{k\to\infty} f(x_{\hat{h}(k)}) \\ &= \lim_{k\to\infty} f(x_{h(k)}) \end{aligned}$$

This shows that, for any i≥1, (13) is also true.
By definition of $\hat{h}$ and (7) can be obtained:

$$\hat{h}(k) = h(k+M+2) \le k+M+2$$

Namely:

$$\hat{h}(k)-k-1 \le M+1 \tag{16}$$

Thus, for any k, do deformation:

$$x_{k+1} = x_{\hat{h}(k)} - \sum_{i=1}^{\hat{h}(k)-k-1}(x_{\hat{h}(k)-i+1} - x_{\hat{h}(k)-i})$$

$$= x_{\hat{h}(k)} - \sum_{i=1}^{\hat{h}(k)-k-1}\alpha_{\hat{h}(k)-i} d_{\hat{h}(k)-i}\,. \tag{17}$$

On where $x_{\hat{h}(k)}$ transposition and notes (16), was:

$$\| x_{k+1} - x_{\hat{h}(k)} \| = \| -\sum_{i=1}^{\hat{h}(k)-k-1}\alpha_{\hat{h}(k)-i} d_{\hat{h}(k)-i} \| \le \sum_{i=1}^{M+1}\| \alpha_{\hat{h}(k)-i} d_{\hat{h}(k)-i} \| \tag{18}$$

On both sides so that k→∞, by (12):

$$\lim_{k\to\infty}\| x_{k+1} - x_{\hat{h}(k)} \| = 0\,.$$

Then and then by f(x) consistent continuity:

$$\lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} f(x_{\hat{h}(k)})$$

By (14), we can see:

$$\lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} f(x_{h(k)}) \tag{19}$$

Nine On both sides of order k→∞, by (19) and noted that δ>0, Lemma 4 holds.

Remark: If (4) becomes a monotonous line search conditions, can obviously be seen $f(x_k)$ is monotonically decreasing, if $f(x)$ lower bound, easy to get $\sum_{k=1}^{\infty}\alpha_k^2\|d_k\|^2 < +\infty$, in particular, have Lemma 4. Here the weak non-monotonic search, to prove Lemma 4 fee to a lot of twists and turns.

Lemma 5: Under the assumptions (H1), (H2) of the condition, $\frac{y_k^T s_k}{s_k^T s_k} \ge c_1$ and $\frac{\|y_k\|^2}{s_k^T s_k} \le \frac{c_2^2}{c_1}$ established. Among them, $c_1 > 0$ with (H2), as defined in C_2>0 a constant.

Proof: Prove modeled in Sun and Yuan (2006), Lemma 5.3.2.

Lemma 6: Set up B_k BFGS formula (3) obtained, B_0 is symmetric positive definite. If the presence of a positive constant number *m*, *M* Such that for any k≥0, y_k and S_k meet $\frac{y_k^T s_k}{s_k^T s_k} \ge m$ and $\frac{\|y_k\|^2}{s_k^T s_k} \le M$. Then for any $p \in (0,1)$, The presence of a positive constant number β_1, β_2, β_3 make any k≥0 inequality:

$$\beta_1 \le \frac{\| B_i s_i \|}{\| s_i \|} \le \beta_2$$

and

$$\beta_1 \le \frac{s_i^T B_i s_i}{\| s_i \|^2} \le \beta_3,$$

the i∈ $\{0, 1, 2, \ldots, k\}$ at least [pk] indicators established. [x] is not less than x the smallest integer.

Theorem 2: Under the assumptions (H1), (H2) of the condition, for Algorithm 1, or the existence of k, making the

$$\|g_k\| = 0$$

Or

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

Proof: Use reduction ad absurdum. Assume that the conclusion is not established, there is a constant $\varepsilon > 0$, such that for any k≥0, there is:

$$\|g_k\| \ge \varepsilon \tag{20}$$

Lemma 5: Shows that the algorithm is in line with the conditions of Lemma 6. Thus, for any $p \in (0,1)$, the presence of a positive constant number β_1, β_2, β_3 make any k≥0 , the inequality:

$$\beta_1 \| s_i \| \le \| B_i s_i \| \le \beta_2 \| s_i \|$$

and

$$\beta_1 \| s_i \|^2 \le s_i^T B_i s_i \le \beta_3 \| s_i \|^2$$

To i = ∈ $\{0, 1, 2, \ldots, k\}$ at least [pk] indicators established. Noted that $S_i = \alpha_i d_i$ equations, can be written as:

$$\beta_1 \| d_i \| \le \| B_i d_i \| \le \beta_2 \| d_i \|$$

and

$$\beta_1 \| d_i \|^2 \le d_i^T B_i d_i \le \beta_3 \| d_i \|^2$$

by (5), two equations that

$$\beta_1 \| d_i \| \le \| g_i \| \le \beta_2 \| d_i \|$$

and

$$\beta_1 \| d_i \|^2 \le -d_i^T g_i \le \beta_3 \| d_i \|^2$$

Based on the above discussion, we define the index set J_t and J are as follows (out of habit , in the above formula , the subscript i replaced by k) :

$$J_t = \{k \le t \mid \beta_1 \| d_k \| \le \| g_k \| \le \beta_2 \| d_k \| \quad \beta_1 \| d_k \|^2 \le -d_k^T g_k \le \beta_3 \| d_k \|^2\}$$

And

$$J = \bigcup_{t=1}^{\infty} J_t$$

We can put it another way, there is a positive constant number β_1, β_2, β_3 and infinite indicators set J, such that for any $k \in J$, to meet

$$\beta_1 \| d_k \| \le \| g_k \| \le \beta_2 \| d_k \| \tag{21}$$

$$\beta_1 \| d_k \|^2 \le -d_k^T g_k \le \beta_3 \| d_k \|^2 \tag{22}$$

Nonmonotonic line search NLS (4), for any $k \in J$, there is

$$\begin{aligned} & f(x_k + \frac{\alpha_k d_k}{\beta}) \\ & > \max_{0 \le j \le l(k)} f(x_{k-j}) - \delta(\frac{\alpha_k}{\beta})^2 \|d_k\|^2 \\ & \ge f(x_k) - \delta(\frac{\alpha_k}{\beta})^2 \|d_k\|^2, \end{aligned} \tag{23}$$

$f(x_k)$ f(x_k) transpose,

$$f(x_k + \frac{\alpha_k d_k}{\beta}) - f(x_k) > -\delta(\frac{\alpha_k}{\beta})^2 \|d_k\|^2, \tag{24}$$

the use of the upper - left of the mean value theorem and finishing,

$$g(u_k)^T d_k \ge -\delta(\frac{\alpha_k}{\beta}) \|d_k\|^2, \tag{25}$$

where $u_{k=x_k} \frac{\omega_k \alpha_k d_k}{\beta} \omega_k \in (0,1), k \in J$.

Notice $\{x_k\} \subset L_0$ (See the proof of Theorem 1), L_0 bounded, i.e. $\{x_k\}$ also bounded. Therefore, you can always find a convergent subsequence $\{x_k | k \in J'\} \subseteq \{x_k | k \in J\} \subseteq \{x_k\}$. For subseries $\{x_k | k \in J'\}$, the corresponding sequence $\{d_k | k \in J'\}$, may be constructed according to algorithm. Impossible in sequence $\{x_k | k \in J'\}$, found an infinite number of points makes $\|d_k\| = 0$, otherwise known from (21), (20) contradictions and assumptions. So, can find a convergent subsequence $\left\{\frac{d_k}{\|d_k\|} | k \in J'' \subseteq J'\right\}$. In this way, we find convergent subsequence $\{x_k | k \in J''\} \subseteq \{x_k | k \in J\}$, so that at the same time meet the:

$$\lim_{k \to \infty,\, k \in J''} x_k = \hat{x}, \tag{26}$$

$$\lim_{k \to \infty,\, k \in J''} \frac{d_k}{\|d_k\|} = \hat{d} \cdot \tag{27}$$

Lemma 4 and (26), we obtain $\lim_{k \to \infty, k \in \infty} u_{k=\hat{x}}$. By the continuity of g, it is found $\lim_{k \to \infty, k \in J''} g(u_k)$ exists and

$$\lim_{k \to \infty,\, k \in J''} g(u_k) = g(\hat{x}) \cdot \tag{28}$$

For (25), let $k \in J''$ and divide both sides by $\|d_k\|$, let $k \to \infty$,

$$\begin{aligned} & \lim_{k \to \infty,\, k \in J''} g(u_k)^T \frac{d_k}{\| d_k \|} \\ & = \lim_{k \to \infty,\, k \in J''} g(u_k)^T \lim_{k \to \infty,\, k \in J''} \frac{d_k}{\| d_k \|} \\ & \ge \lim_{k \to \infty,\, k \in J''} -\frac{\delta}{\beta} \alpha_k \|d_k\| \end{aligned}$$

by (28), (27) and Lemma 4:

$$g(\hat{x})^T \hat{d} \ge 0 \,. \tag{29}$$

The following inequality (22) on the left to take the same means, so $k \in J''$ and divide both sides by $\|d_k\|$,

$$0 < \beta_1 \| d_k \| \le -g_k^T \frac{d_k}{\| d_k \|}$$

let $k \to \infty$, by the continuity of g, (26), (27) and (29) can be obtained $\lim_{k \to \infty, k \in J''} \|d_k\| = 0$. Then from (21), $\lim_{k \to \infty, k \in J''} \|g_k\| = 0$ contradiction of this hypothesis (20).

Table 1: Results of number of numerical experiments to space limitations

Test	Rosenbrock		Penalty	
M	n_i	n_f	n_i	n_f
0	562	1195	83	174
1	238	538	49	327
2	257	806	196	507
3	411	1089	196	723
4	333	896	115	519
5	442	1721	196	833
6	606	1566	196	104
7	659	1663	27	41

CONCLUSION

The author of this study, a number of numerical experiments to space limitations. The procedures Matlab6.5 been prepared on a normal PC, taken as parameters unified $\sigma = 1$, $\beta = 0.2$, $\delta = 0.9$, $\varepsilon = 10^{-6}$. We calculated for different values of M, n_i represents the number of iterations, n_f represents the number of calculations of the function value calculation times, gradient. Results from the numerical point of view, the proposed line search method has the following advantages (Table 1):

- When the correct initial testing step according to the formula of this study, is usually better than the initial test step is fixed
- For non-monotone line search method, the number of iterations, the function value calculation times are reduced
- The Nonmonotone strategy is effective for most functions, especially in the case of the high-dimensional, or initial test step length is fixed (Table 1).

ACKNOWLEDGMENT

The study is supported by Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (PHR (IHLB)) (PHR201108407). Thanks for the help.

REFERENCES

Byrd, R. and J. Nocedal, 1989. A tool for analysis of quasi-newton methods with application to unconstrained minimization. SIAM J. Numer. Anal., 26(3): 727-739.

Byrd, R., J. Nocedal and Y.X. Yuan, 1987. Global convergence of a class of quasi-Newton methods on convex problems. SIAM J. Numer. Anal., 24(4): 1171-1190.

Dai, Y.H., 2002a. Convergence properties of the BFGS algorithm. SIAM J. Optimiz., 13(3): 693-701.

Dai, Y.H., 2002b. On the nonmonotone line search. J. Optimiz. Theory App., 112(2): 315-330.

Dai, Y.H., 2002c. Conjugate gradient methods with Armijo-type Line Searches. Acta Math. Appl. Sin-E., 18(1): 123-130.

Grippo, L., F. Lampariello and S. Lucidi, 1986. A nonmonotone line search technique for newton's method. SIAM J. Numer. Anal., 23(4): 707-716.

Han, J.Y. and G.H. Liu 1997. Global convergence analysis of a new nonmonotone BFGS algorithm on convex objective functions. Comput. Optimiz. App., 7(2): 277-289.

Hüther, B., 2002. Global convergence of algorithms with nonmonotone line search strategy in unconstrained optimization. Results Math., 41: 320-333.

Leone, R.D., M. Gaudioso and L. Grippo, 1984. Stopping criteria for line search methods without derivatives. Math. Program., 30(3): 285-300.

Liu, G., J. Han and D. Sun, 1995. Global convergence of the BFGS algorithm with nonmonotone line search. Optimization, 34(2): 147-159.

Panier, E.R. and A.L. Tits 1991. Avoiding the maratos effect by means of nonmonotone line search. SIAM J. Numer. Anal., 28(4): 1183-1195.

Sun, W. and Y.X. Yuan, 2006. Optimization Theory and Methods: Nonlinear Programming. Springer, New York, pp: 47-48.

Zhang, H.C. and W.W. Hager, 2004. A Nonmonotone line search technique and its application to unconstrained optimization. SIAM J. Optimiz., 14(4): 1043-1056.

Zhen-Jun, S. and S. Jie, 2006. Convergence of nonmonotone line search method. J. Comput. Appl. Math., 193(2): 397-412.