Send Orders for Reprints to reprints@benthamscience.ae

218 *The Open Electrical & Electronic Engineering Journal*, 2014, *8*, 218-221

Open Access

A Non-Monotone Line Search Combination Technique for Unconstrained Optimization

Ping Hu* and Zong-yao Wang

Faculty of Mathematics and Physics, Huaiyin Institute of Technology, Huaian 223003, P.R. China

Abstract: We propose a non-monotone line search combination rule for unconstrained optimization problems, the corresponding non-monotone search algorithm is established and its global convergence can be proved. Finally, we use some numerical experiments to illustrate the new combination of non-monotone search algorithm's effectiveness.

Keywords: BFGS algorithm, combination rule, global convergence, non-monotone line search algorithm, unconstrained optimization.

1. INTRODUCTION

Unconstrained optimization problems

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

Has wide application background. Many authors have devoted a lot of time and effort to the algorithm research of unconstrained optimization problems [1]. The algorithm is generally solved by iterative method. At the current iteration point x_k, $g_k = \nabla f(x_k) \neq 0$, then step length α_k is determined along the search direction d_k, the next iteration point obtained by (1.1).

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.1}$$

In (1.1), there are many algorithms for d_k, Since the BFGS is currently the best quasi-newton method, and gives a convenient algorithm program. In this paper we use BFGS algorithm :

$$d_k = -H_k g_k \text{, and if } g_k^T d_k \geq 0 \text{, let } d_k = -g_k \tag{1.2}$$

Where, $g_k = \nabla f(x_k) \neq 0$, H_k are generated by the BFGS correction formula.

In (1.1), α_k can be produced by the Arimijo rule, Goldstein rule and Wolfe rule [1] that are all classic monotone line search rules. Recent research [2-4] indicates that the monotone line search method may considerably reduce the rate of convergence when the iteration point x_k is trapped near a narrow curved valley, so the non-monotone line search method was proposed by Grippo *et al.* [4]. As a result, it is helpful to overcome this drawback. Some numerical experiments also show that the non-monotone line search method is effective. From non-monotone line search rule [2-10], the inequality $f(x_{k+1}) > f(x_k)$ may hold for any k, which play the non-monotone search role. However, the above non-monotone line search rules required conditions are $f(x_k) \leq f(x_0)$. Under this condition, if x_0 is near bottom of the valley, x_k is difficult to get out of this valley and search for a better point. Therefore, we improve non-monotone line search methods, and prove the global convergence property by virtue of [2, 3, 11]. With the help of numerical experiments it is shown that the proposed method is very effective.

2. NONMONOTONE COMBINATION RULE FOR LINE SEARCHES

In this section we establish a non-monotone combination rules. First of all, we give the assumption and definition as follows.

Assumption 1. The function $f: R^n \to R$ is bounded and differentiable on the level set

$$\Omega = \left\{ x \,\middle|\, f(x) \leq c\ f(x_0) \text{, constant } c \geq 1 \right\}$$

And $f(x) > 0$.

If $\min f_1(x) = -\infty$, it is meaningless to solve the problem $\min f_1(x)$. So we assume $\inf f_1(x) = w$. If $f_1(x) \leq 0$, take fixed constant $\beta > |w|$, such that $f_1(x) + \beta = f(x) > 0$. Obviously, $f(x)$ and $f_1(x) + \beta$ have the same extreme value point, derivative and continuous roundedness. Therefore, it is feasible to set $f(x) > 0$.

Definition 1. The function $\sigma: [0, +\infty) \to [0, +\infty)$ is a forcing function (F- function), if for any sequence $\{t_i\} \subset [0, +\infty)$,

$$\lim_{i\to\infty} \sigma(t_i) = 0 \text{ implies } \lim_{i\to\infty} t_i = 0. \tag{2.1}$$

Now we give the non-monotone combination rule for line searches as follows. Let $\alpha_k \geq 0$ be bounded and satisfy

$$f(x_{k+1}) = f(x_k + \alpha_k d_k) \leq \sum_{r=o}^{m(k)} \lambda_{kr} f(x_{k-r}) - \sigma(t_k), \tag{2.2}$$

where $m(k) = \min[k, M-1]$, with $M \geq 1$, a positive integer,

$$\sum_{r=o}^{m(k)} \lambda_{kr} = \mu_k \geq 1\ ,\ \prod_{k=0}^{\infty} \mu_k = c \ \ (c \text{ is a limited constant}),$$

*Address correspondence to this author at the MeiCheng east road no. 1, Huaian, P.R.China. Postcard: 223003; Tel: +86-517-83559093; Fax: +86-517-83559053; E-mail: hyhuping@163.com

$\lambda_{kr} \geq \lambda$, $\lambda \in (0,1)$, σct_k is a forcing function [2], with $t_k = -g_k^T d_k / \|d_k\| \geq 0$.

In (2.2), when $M = 1$ and $\lambda_{k0} = 1$, we have

$f(x_k + \alpha_k d_k) \leq f(x_k) - \sigma(t_k)$,

Where $\sigma ct_k) \leq \alpha_k \rho \; -g_k^T d_k$, combination rule become Arimijo rule.

If $f(x_{k-l}) = \max\limits_{0 \leq r \leq m(k)} \{f(x_{k-r})\}$, take $\lambda_{kl} = 1, \lambda_{kr} = 0$ ($r \neq l$), then from (2.2) were:

$f(x_k + \alpha_k d_k) \leq \max\limits_{0 \leq j \leq m(k)} \{f(x_{k-j})\} - \sigma(t_k)$,

Which was the search rule proposed by Grippo *et al.* [4]. It is visible that combination rule is more universal and meanwhile it has the global convergence.

3. THE GLOBAL CONVERGENCE

In this section, we investigate the global convergence of the non-monotone combination search BFGS algorithm.

Lemma 1. If $\mu_k \geq 1$ and $\prod\limits_{k=0}^{\infty} \mu_k = c$ then $s(n) = \prod\limits_{k=0}^{n} \mu_k$ is a monotonically non-decreasing function about n, and $s(n) \leq c$.

Theorem 1. If α_k satisfies combination rule (2.2),

$f(x) > 0$, $\mu_k \geq 1$ and $\prod\limits_{k=0}^{\infty} \mu_k = c$, then

$$f(x_{k+1}) \leq f(x_0) \prod_{n=0}^{k} \mu_n - \lambda \sum_{r=o}^{k-1} \sigma(t_r) - \sigma(t_k) . \qquad (3.1)$$

Proof. We will use mathematical induction to prove this theorem. In fact, for the case of $k = 1$, we have

If M=1 , then $m(k) = 0$. From (2.2), we have

$f(x_2) \leq \lambda_{10} f(x_1) - \sigma(t_1)$

$=\mu_1 f(x_1) - \sigma(t_1)$,

Which combined with

$f(x_1) \leq \lambda_{00} f(x_0) - \sigma(t_0)$

$=\mu_0 \; f(x_0) - \sigma(t_0)$,

Gives $f(x_2) \leq \mu_1 \{\mu_0 f(x_0) - \sigma(t_0)\} - \sigma(t_1)$

$= \mu_1 \; \mu_0 \; f(x_0) - \mu_1 \sigma(t_0) \; - \sigma(t_1)$

$\leq \mu_1 \; \mu_0 \; f(x_0) - \lambda \sigma(t_0) \; - \sigma(t_1)$.

If $M > 1$, it is concluded from（2.2）that

$f(x_2) \leq \lambda_{10} f(x_1) + \lambda_{11} f(x_0) - \sigma(t_1)$

$\leq \lambda_{10} \; \{\mu_0 f(x_0) - \sigma(t_0)\} \; + \lambda_{11} f(x_0) - \sigma(t_1)$

$=(\lambda_{10}\mu_0 + \lambda_{11}) f(x_0) \; - \lambda_{10} \; \sigma(t_0) \; - \sigma(t_1)$

$\leq \mu_0 (\lambda_{10} + \lambda_{11}) f(x_0) \; - \lambda_{10} \; \sigma(t_0) \; - \sigma(t_1)$

$\leq \mu_0 \mu_1 \; f(x_0) \; - \lambda \; \sigma(t_0) \; - \sigma(t_1)$;

That is (3.1) holds for $k = 1$. Let us now assume

$$f(x_k) \leq f(x_0) \prod_{n=0}^{k-1} \mu_n - \lambda \sum_{r=o}^{k-2} \sigma(t_r) - \sigma(t_{k-1}) \; .$$

It follows that

$$f(x_{k+1}) \leq \sum_{r=o}^{m(k)} \lambda_{k\,r} f(x_{k-r}) - \sigma(t_k)$$

$$\leq \sum_{r=o}^{m(k)} \lambda_{k\,r} \; \{f(x_0) \prod_{n=0}^{k-r-1} \mu_n - \lambda \sum_{i=o}^{k-r-2} \sigma(t_i) - \sigma(t_{k-r-1})\} \; - \sigma(t_k)$$

$$\leq \sum_{r=o}^{m(k)} \lambda_{k\,r} \; \{f(x_0) \prod_{n=0}^{k-1} \mu_n - \lambda \sum_{i=o}^{k-m(k)-2} \sigma(t_i) - \sigma(t_{k-r-1})\} \; - \sigma(t_k)$$

$$= f(x_0) \; (\sum_{r=o}^{m(k)} \lambda_{k\,r}) \prod_{n=0}^{k-1} \mu_n \; - \lambda \; (\sum_{r=o}^{m(k)} \lambda_{k\,r}) \sum_{i=o}^{k-m(k)-2} \sigma(t_i)$$

$$- \sum_{r=o}^{m(k)} \lambda_{k\,r} \; \sigma(t_{k-r-1}) \; - \sigma(t_k)$$

$$= f(x_0) \; \mu_k \prod_{n=0}^{k-1} \mu_n \; - \lambda \; \mu_k \sum_{i=o}^{k-m(k)-2} \sigma(t_i) \; - \sum_{r=o}^{m(k)} \lambda_{k\,r} \; \sigma(t_{k-r-1})$$

$$- \sigma(t_k)$$

$$\leq f(x_0) \prod_{n=0}^{k} \mu_n \; - \lambda \sum_{i=o}^{k-m(k)-2} \sigma(t_i) \; - \lambda \sum_{i=k-m(k)-1}^{k-1} \sigma(t_i) \; - \sigma(t_k)$$

$$\leq f(x_0) \prod_{n=0}^{k} \mu_n \; - \lambda \sum_{r=o}^{k-1} \sigma(t_r) \; - \sigma(t_k) \, .$$

This means that (3.1) holds for any $k \in Z, k \geq 1$.

The proof is completed.

Next, we will prove the global convergence of the non-monotone combination search algorithm.

Theorem 2. Assume Assumption 1 hold, the search direction d_k satisfies BFGS algorithm, line search length α_k satisfies rule (2.2), $\{x_k\}$ is generated sequence, then $\{x_k\} \in \Omega$, and $\lim\limits_{k \to \infty} \|g_k\| = 0$.

Proof. From Theorem 1 we have

$$f(x_{k+1}) \leq f(x_0) \prod_{n=0}^{k} \mu_n - \lambda \sum_{r=o}^{k-1} \sigma(t_r) - \sigma(t_k) \leq c \; f(x_0) \, ,$$

According to the definition of Ω , we know that $\{x_k\} \in \Omega$.

From (3.1) we have

$$f(x_{k+1}) \leq f(x_0) \prod_{n=0}^{k} \mu_n - \lambda \sum_{r=o}^{k-1} \sigma(t_r) - \sigma(t_k)$$

$$\leq f(x_0) \prod_{n=0}^{k} \mu_n - \lambda \sum_{r=o}^{k} \sigma(t_r) \, .$$

That is

$$0 \leq \lambda \sum_{r=0}^{k} \sigma(t_r) \leq c \; f(x_0) - f(x_{k+1}) \, . \qquad (3.2)$$

From Assumption 1 we know that $f(x_{k+1})$ are bounded in the Ω , so according to (3.2), when $k \to \infty$,we have

$\lambda \sum\limits_{r=o}^{\infty} \sigma(t_r) < \infty$, or equivalently $\lim\limits_{k \to \infty} \sigma(t_k) = 0$.

From Definition 1, we have

$\lim\limits_{k \to \infty} t_k = \lim\limits_{k \to \infty} (-g_k^T d_k / \|d_k\|) = 0$,

That is

$\lim\limits_{k \to \infty} \|g_k\| \cos(-g_k, d_k) = 0$, From (1.2), we have $\cos(-g_k, d_k) > 0$, so $\lim_{k\to\infty}\|g_k\| = 0$.

This completes the proof of Theorem 2.

Theorem **2** illustrates the global convergence of the non-monotone combination search algorithm. To test the effectiveness of the non-monotone combination rule, we give some numerical experiments.

4. NUMERICAL EXPERIMENTS

In this section we use the non-monotone combination line search rule and BFGS algorithm to test a few standard test problems, where line search steps meet non-monotone search combination rule, the following are the basic steps of algorithms:

Step 1: A given initial value $x_0 \in R^n$, $\alpha_0 = 1$, $\varepsilon \ge 0$, integer $M \ge 1$, $\rho \in (0, 0.5)$, $H_0 \in R^{n\times n}$ (symmetric positive), let $k = 0$;

Step 2: Testing the termination condition, we calculate $g_k = \nabla f(x_k)$, if $\|g_k\| \le \varepsilon$, $x^* = x_k$, then stop;

Step 3: Calculating the search direction, $d_k = -H_k g_k$; if $g_k^T d_k > 0$, then $d_k = -g_k$;

Step 4: Determine the line search step length α_k. Order $m(k) = \min\{k, M-1\}$, λ_{kr} meet $\sum_{r=o}^{m(k)} \lambda_{kr} = \mu_k \ge 1$ and $\prod_{k=0}^{\infty} \mu_k = c$. If α eligible

$$f(x_k + \alpha\, d_k) \le \sum_{r=o}^{m(k)} \lambda_{kr} f(x_r) + \rho\alpha g_k^T d_k \ ,$$

Then $\alpha_k = \alpha$, otherwise, using two points two times interpolation shorten α;

Step 5: Calculating new points. Order $s_k = \alpha_k d_k$, $x_{k+1} = x_k + s_k$, calculating $f(x_{k+1})$, $g_{k+1} = \nabla f(x_{k+1})$;

Step 6: Calculating $y_k = g_{k+1} - g_k$,

$$v_k = (y_k^T H_k y_k)^{\frac{1}{2}} [\frac{s_k}{s_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k}],$$

$$H_{k+1} = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + v_k v_k^T,$$

$k = \mathrm{k}+1$, go to step 2.

We choose the following tests.

Question 1. *GROSEN* function

$$f(x) = \sum_{i=2}^{n} [100(x_i - x_{i-1}^2)^2 + (1 - x_{i-1})^2],$$

$$x_0 = (-1.2\ ,1\ ,-1.2\ ,1\ ,\ \cdots\ ,1\ )^T.$$

Question 2. Wood function

$$f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_3^2 - x_4)^2 + 10.1[(x_2 - 1)^2 + (x_4 - 1)^2] + 19.8(x_2 - 1)(x_4 - 1),$$

$$x_0 = [-3, -1,\ -3,\ -1]^T.$$

In the above two questions, the parameters involved are taken as $\lambda_{kr} = \frac{1}{1+m(k)} \beta^{\frac{1}{(1+k)^{1.2}}}$ $(r = 0,1,\cdots,m(k))$, $\varepsilon = 10^{-6}$, $\rho = 10^{-3}$, $H_0 = I_{n\times n}$, $\beta \ge 1$ and $M \ge 1$ are controlled constants for $\beta \in R$ and $M \in Z$. Note that the rule reduces to the monotone line search method when $M = 1$ and $\beta = 1$, while the rule is similar to the non-monotone line search of F-rule when $M > 1$ and $\beta = 1$ In the following three tables, n_g denotes the outside loop iterations, n_f denotes the function evaluations, while $f(x^*)$ stands for the function value of the approximate solution x^*.

From the results of above three questions, we can see that it is effective for the non-monotone combination line search algorithm.

For question 1, see Table **1**, the iterations of non-monotone line search algorithm are less than monotone line search algorithm, while the iterations of combination line search algorithm are much less and the accuracy is higher. From Table **2**, we can see that non-monotone line search

Table 1. The calculation results of question 1 ($n = 2$).

M	β	n_g	n_f	$f(x^*)$	M	β	n_g	n_f	$f(x^*)$
1	1	39	96	1.5647e-017	1	5	37	91	7.3465e-019
2	1	38	92	3.5514e-023	2	5	31	75	7.8019e-020
3	1	38	92	3.5514e-023	3	5	31	75	7.8019e-020
4	1	38	90	1.2758e-017	4	5	31	75	7.8019e-020
5	1	38	90	1.2758e-017	5	5	31	75	7.8019e-020

Table 2. The calculation results of question 1 , ($n = 100$).

M	β	n_g	n_f	$f(x^*)$	M	β	n_g	n_f	$f(x^*)$
1	1	1001	14781	2.0205e-012	1	5	1001	14040	6.6312e-012
2	1	1001	14314	5.2899e-014	2	5	1001	13689	5.8931e-013
3	1	1001	13083	1.2005e-015	3	5	1001	12950	1.7840e-015
4	1	655	1807	4.5093e-016	4	5	1001	12286	1.6415e-015
5	1	659	1814	5.2653e-016	5	5	681	1833	5.2825e-016

Table 3. The calculation results of question 2.

M	β	n_g	n_f	$f(x^*)$	M	β	n_g	n_f	$f(x^*)$
1	1	101	2579	6.0152e-013	1	5	101	2579	6.0152e-013
2	1	101	2577	1.3019e-014	2	5	38	103	3.5936e-016
3	1	35	97	1.1446e-017	3	5	42	111	5.2348e-020
4	1	42	111	5.2348e-020	4	5	42	111	5.2348e-020
5	1	64	161	6.3740e-018	5	5	64	161	6.3740e-018

algorithm is better than monotone line search algorithm, and the results of non-monotone combination line search algorithm are basically the same as ordinary non-monotone line search algorithm, which indicates the efficiency of non-monotone combination line search algorithm.

For question 2, see Table **3**, the results of non-monotone combination line search algorithm are basically the same as ordinary non-monotone line search algorithm, that are all better than monotone line search algorithm.

CONCLUSION

From (3.2), we can obtain $c\,f(x_0)-f(x_k)\geq 0$, that is $f(x_k)\leq c\,f(x_0)$. From $\prod_{k=0}^{\infty}\mu_k=c$ and $\mu_k\geq 1$, we have $c\geq 1$, so $f(x_k)>f(x_0)$ can be achieved, which is our biggest breakthroughs to make x_k out a nearby valley of x_0 to search a better solution.

Rule (2.2) we proposed is easy to implement and emulated a lot of μ_k, example $\mu_k=\beta^{\frac{1}{(k+1)^p}}$, constant $\beta\geq 1$, $p>1$, now p-series $\sum_{k=0}^{\infty}\frac{1}{(k+1)^p}$ converge on a limited number s, then

$$\prod_{k=0}^{\infty}\mu_k\leq\prod_{k=0}^{\infty}\beta^{\frac{1}{(1+k)^p}}=\beta^{\sum_{k=0}^{\infty}\frac{1}{(1+k)^p}}=\beta^{s}=c\,.$$

In order to prove easily, we assume $f(x)>0$ in rule (2.2). If $f(x)\leq 0$, μ_k should satisfy $0<\mu_k\leq 1$, now theoretical proof is more troublesome, therefore we need further research.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Research supported by the Natural Science Foundation of Jiangsu Province (BK20141409) and the Natural Science Foundation of China (Grant No.: 61074188).

REFERENCES

[1] Q. Ni, *Optimization Method and Program Design*, Beijing: Science Press, 2009.
[2] W. Y. Sun, J. Y. Han, and J. Sun, “Global convergence of non-monotone descent methods for unconstrained optimization problems,” *J. Computat. Appl. Math.*, vol. 146, pp. 89-98, 2002.
[3] Z. S. Yu, and D. G. Pu, “A new non-monotone line search technique for unconstrained optimization,” *J. Computat. Appl. Math.*, vol. 219, pp. 134-144, 2008.
[4] L. Grippo, F. Lampariello, and S. Lucidi, “A non-monotone line search technique for Newton ’s Methods,” *SIAM J. Numer. Anal.*, vol. 23 , pp.707-716, 1986.
[5] P. P. Liu, and L. P. Chen, “A Quasi-Newton non-monotonic trust Region algorithm and its convergence,” *Adv. in Math.*, vol. 37, pp.92-100, 2008.
[6] Y. H. Xiao, H. J. Sun, and Z. G. Wang, “A globally convergent BFGS method non-monotone line search for non-convex minimization,” *J. Computat. Appl. Math.*, vol. 230, pp. 95-106, 2009.
[7] W. Y. Sun, and Q. Y. Zhou, “An unconstrained optimization method using non-monotone second order Goldstein’s line search,” *Sci. China Ser. A Math.*, vol. 50, pp. 1389-1400, 2007.
[8] Q. Y. Zhou, and W. Y. Sun, “Adaptive non-monotone line search method for unconstrained optimization,” *Front. Math. China*, vol. 3, pp. 133-148, 2008.
[9] M. A. Diniz-Ehrhardt, J. M. Martinez, and M. Raydan, “Aderivative-free non-monotone line search technique for unconstrained optimization,” *J. Computat. Appl. Math.*, vol. 219, pp. 383-397, 2008.
[10] L. P. Chen, J. J. Yu, and B. C. Jiao, “A class of hybrid conjugate gradient methods with non-monotone line search,” *Math. Pract. Theory*, vol. 38, pp.101-107, 2008.
[11] P. Hu, X. Q. Liu, “A nonmonotone line search slackness technique for unconstrained optimization,” *J. Optim. Theory Appl.*, vol. 158 pp. 773-786, 2013.

Received: June 09, 2014 Revised: September 22, 2014 Accepted: November 24, 2014