

# A Non-Monotone Trust Region Method with Non-Monotone Wolfe-Type Line Search Strategy for Unconstrained Optimization

Changyuan Li\*, Qinghua Zhou\*, Xiao Wu

College of Mathematics and Information Science, Hebei University, Baoding, China  
Email: \*[mbydlcy@126.com](mailto:mbydlcy@126.com), \*[qinghua.zhou@gmail.com](mailto:qinghua.zhou@gmail.com), [wuxiao616@163.com](mailto:wuxiao616@163.com)

Received 21 May 2015; accepted 27 June 2015; published 30 June 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In this paper, we propose and analyze a non-monotone trust region method with non-monotone line search strategy for unconstrained optimization problems. Unlike the traditional non-monotone trust region method, our algorithm utilizes non-monotone Wolfe line search to get the next point if a trial step is not adopted. Thus, it can reduce the number of solving sub-problems. Theoretical analysis shows that the new proposed method has a global convergence under some mild conditions.

## Keywords

Unconstrained Optimization, Non-Monotone Trust Region Method, Non-Monotone Line Search, Global Convergence

---

## 1. Introduction

Consider the following unconstrained optimization problem:

$$\min f(x), \quad x \in R^n, \quad (1)$$

where  $f: R^n \rightarrow R$  is a twice continuously differentiable function. Throughout this paper, we use the following notation:

- $\|\cdot\|$  is the Euclidean norm.
- $g(x) = \nabla f(x) \in R^n$  and  $H(x) \in R^{n \times n}$  are the gradient and Hessian matrix of  $f$  evaluated at  $x$ , respectively.

---

\*Corresponding authors.

- $f_k = f(x_k)$ ,  $g_k = g(x_k)$ ,  $H_k = \nabla^2 f(x_k)$  and  $B_k$  is a symmetric matrix which is either  $H_k$  or an approximation of  $H_k$ .

For solving (1), trust region methods usually compute  $d_k$  by solving the quadratic sub-problem:

$$\min m_k(d) = f_k + g_k^T d + \frac{1}{2} d^T B_k d, \quad \|d\| \leq \Delta_k, \quad (2)$$

where  $\Delta_k > 0$  is a trust region radius. Some criteria are used to determine whether a trial step  $d_k$  is accepted. If not, the sub-problem (2) may be computed several times at one iterate until an acceptable step is found. There is no doubt that the repetitive process will increase the cost of solving the problem.

To improve the performance of the algorithm, Nocedal and Yuan [1] put forward an algorithm which combine trust region algorithm and line search method for the first time in 1998, then Gertz [2] proposed a new trust region algorithm that use Wolfe line search at each iteration to obtain a new iteration point regardless of whether  $d_k$  is accepted. Both of them improved the computational efficiency by fully using the advantages of two kinds of algorithm.

Algorithms mentioned above are monotonic algorithm. In 1982, Chamberlain *et al.* in [3] proposed the watchdog technique for constrained optimization to overcome the Maratos effect. Motivated by this idea, Grippo *et al.* first introduced a non-monotone line search technique for Newton's method in [4]. In 1993, Deng *et al.* [5] proposed a non-monotone trust region algorithm in which they combined non-monotone term and trust region method for the first time. Due to the high efficiency of non-monotone techniques, many authors are interested in working on the non-monotone techniques for unconstrained optimization problem [6]-[15]. Especially, some researchers have combined non-monotone trust region method with non-monotone Armijo-type line search method and good numerical results have been achieved [16] [17]. Besides, Zhang and Hager [18] proposed a non-monotone Wolfe-type condition. Inspired by [16]-[18], we present a non-monotone trust region algorithm with non-monotone Wolfe-type line search strategy. To be specific, the algorithm first solve sub-problem (2) to compute the trial step  $d_k$ , if the trial step is accepted, set  $x_{k+1} = x_k + d_k$ . Otherwise, the algorithm performs the non-monotone Wolfe-type line search strategy to find an iterative point instead of resolving the sub-problem.

## 2. Non-Monotone Term and Wolfe-Type Line Search Condition

The general non-monotone form is as follows:

$$f_{l(k)} = f(x_{l(k)}) = \max_{0 \leq j \leq m(k)} \{f_{k-j}\}, \quad k = 0, 1, 2, \dots \quad (3)$$

where  $m(0) = 0$ ,  $0 \leq m(k) \leq \min\{M, m(k-1)+1\}$  and  $M \geq 0$  is an integer constant. Actually, the most common non-monotone ratio is defined as follows:

$$r_k = \frac{f_{l(k)} - f(x_k + d_k)}{m_k(0) - m_k(d_k)}.$$

Some researchers showed that utilizing non-monotone techniques may improve both the possibility of finding the global optimum and the rate of convergence [4] [18]. However, although the non-monotone technique has many advantages, it contains some drawbacks [11] [17] [18]. To overcome those disadvantages, Ahookhosh *et al.* in [11] proposed a new non-monotone technique to replace (3). They define

$$R_k = \eta_k f_{l(k)} + (1 - \eta_k) f_k, \quad (4)$$

where  $\eta_{\min} \in [0, 1)$ ,  $\eta_{\max} \in [\eta_{\min}, 1]$  and  $\eta_k \in [\eta_{\min}, \eta_{\max}]$ . At the same time, they have the new non-monotone ratio:

$$\rho_k = \frac{R_k - f(x_k + d_k)}{pred_k}. \quad (5)$$

In 2004, Zhang and Hager [18] presented a modified non-monotone Wolfe-type condition. In order to keep

the consistency of non-monotone term and simplify the form of Wolfe-type condition, we define the new modified non-monotone Wolfe-type condition as follows:

$$f(x_k + \alpha_k d_k) \leq R_k + \beta \alpha_k g_k^T d_k, \quad (6)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \gamma g_k^T d_k. \quad (7)$$

The rest of this paper is organized as follows. In Section 3, we introduce the algorithm of non-monotone trust region method with line search strategy. In Section 4, we analyze the new method and prove the global convergence. Some conclusions are given in Section 5.

### 3. New Algorithm

In this paper, we consider the following assumptions that will be used to analyze the convergence properties of the below algorithm:

(H1) The level set  $L_0 = \{x \in R^n \mid f(x) \leq f(x_0)\} \subset \Omega$ , where  $\Omega \in R^n$  is a closed, bounded set.

(H2) The matrix  $B_k$  is a uniformly bounded matrix, *i.e.* there exists a constant  $b > 0$  such that  $\|B_k\| \leq b$  for all  $k$ .

(H3)  $\nabla f(x)$  is a Lipschitz continuous function, *i.e.* there exists a constant  $L > 0$  such that  $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$ ,  $\forall x, y \in R^n$ .

(H4)  $\{\Delta_k\}$  has the upper bound  $\Delta$ .

The new algorithm can be described as follows:

Algorithm 0

**Step 1** An initial point  $x_0 \in R^n$  and a symmetric positive definite matrix  $B_0 \in R^{n \times n}$  are given. The constants  $0 < \mu < 1$ ,  $0 < \beta < \gamma < 1$ ,  $0 < c_1 < 1 < c_2$ ,  $0 \leq \eta_{\min} \leq \eta_{\max} \leq 1$ ,  $M \geq 0$ ,  $\tau > 0$ ,  $\Delta_0 > 0$  and  $\varepsilon > 0$  are also given. Compute  $f(x_0)$  and set  $k = 0$ .

**Step 2** Compute  $g_k$ . If  $\|g_k\| \leq \varepsilon$  then stop, else go to Step 3.

**Step 3** Similar to [1], solve (2) inaccurately to determine  $d_k$ , satisfying

$$m_k(0) - m_k(d_k) \geq \tau \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}, \quad (8)$$

$$g_k^T d_k \leq -\tau \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}. \quad (9)$$

**Step 4** Compute  $m(k)$ ,  $f_{l(k)}$ ,  $R_k$  and  $\rho_k$ . If  $\rho_k \geq \mu$ , go to Step 5. Otherwise, find the step-length  $\alpha_k$  satisfying (6) and (7), then set  $x_{k+1} = x_k + \alpha_k d_k$  and update  $\Delta_{k+1} \in [\|x_{k+1} - x_k\|, c_1 \Delta_k]$ , go to step 6.

**Step 5** Set  $x_{k+1} = x_k + d_k$ , and

$$\Delta_{k+1} \begin{cases} = \Delta_k, & \text{if } \|d_k\| < \Delta_k \\ \in [\Delta_k, c_2 \Delta_k], & \text{if } \|d_k\| = \Delta_k \end{cases}.$$

**Step 6** Update the symmetric matrix  $B_k$  by a quasi-Newton formula, set  $k = k + 1$ , go to step 2.

### 4. Convergence Analysis

For the convenience of expression, we Let  $I = \{k \in K \mid \rho_k \geq \mu\}$  and  $J = \{k \in K \mid \rho_k < \mu\}$ . Obviously,  $K = I \cup J$  is an infinite subset of the set  $\{0, 1, 2, \dots\}$ .

We need the following lemmas in order to prove the convergence of the new algorithm.

**Lemma 1** Assume that Algorithm 0 generates an infinite sequence  $\{x_k\}$ , (H2), (H3) and (H4) hold, and there is a  $\varepsilon > 0$  such that  $\|g_k\| \geq \varepsilon$ . Then for all  $k \in J$ , there exists a constant  $\bar{\alpha} > 0$  such that  $\alpha_k > \bar{\alpha}$ .

**Proof.** From (7) and (H3), we have

$$L\alpha_k \|d_k\|^2 \geq (g(x_k + \alpha_k d_k) - g_k)^T d_k \geq (\gamma - 1) g_k^T d_k > 0. \tag{10}$$

Thus, we can conclude that

$$\alpha_k \geq \frac{(\gamma - 1) g_k^T d_k}{L \|d_k\|^2} = \frac{(1 - \gamma) |g_k^T d_k|}{L \|d_k\|^2}. \tag{11}$$

This inequality, together with (H2), (H4) and (9), lead us to have

$$\alpha_k \geq \frac{(1 - \gamma) \tau \|g_k\| \min\{\Delta_k, \|g_k\|/\|B_k\|\}}{L \Delta_k^2} \geq \frac{(1 - \gamma) \tau \varepsilon}{L} \min\left\{\frac{1}{\Delta}, \frac{\varepsilon}{\Delta^2 b}\right\}. \tag{12}$$

Let  $\bar{\alpha} = \frac{(1 - \gamma) \tau \varepsilon}{L} \min\left\{\frac{1}{\Delta}, \frac{\varepsilon}{\Delta^2 b}\right\}$ , we complete the proof.

**Lemma 2** (See Lemma 2.1 and Corollary 3.1 in [17]) Suppose that the sequence  $\{x_k\}$  is generated by Algorithm 0. Then, we have  $x_k \in L_0$ ,  $\{f_{l(k)}\}$  is a decreasing sequence and

$$\lim_{k \rightarrow \infty} R_k = \lim_{k \rightarrow \infty} f(x_k). \tag{13}$$

**Lemma 3** Suppose that (H1) holds, if sequence  $\{x_k\}$  does not converge to a stationary point, i.e. there exists a constant  $\varepsilon > 0$  such that  $\|g_k\| \geq \varepsilon$ . Then

$$\lim_{k \rightarrow \infty} \min\{\Delta_k, \varepsilon/M_k\} = 0, \tag{14}$$

where  $M_k = 1 + \max_{1 \leq i \leq k} \|B_i\|$ .

**Proof.** We consider two cases:

Case 1.  $k \in I$ . The proof is similar to Lemma 3.3 in [17], we can obtain that

$$f(x_k + d_k) \leq R_k - \varphi_1 \min\{\Delta_k, \varepsilon/M_k\}.$$

Case 2.  $k \in J$ . From (6), (9) and (12), we have

$$\begin{aligned} f(x_k + \alpha_k d_k) &\leq R_k + \beta \alpha_k g_k^T d_k \leq R_k - \beta \alpha_k \tau \|g_k\| \min\left\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\right\} \\ &\leq R_k - \frac{\beta(1 - \gamma) \tau^2 \varepsilon^2 \min\left\{\frac{1}{\Delta}, \frac{\varepsilon}{\Delta^2 b}\right\}}{L} \min\left\{\Delta_k, \frac{\|g_k\|}{1 + \max_{1 \leq i \leq k} \|B_i\|}\right\} \\ &\leq R_k - \frac{\beta(1 - \gamma) \tau^2 \varepsilon^2 \min\left\{\frac{1}{\Delta}, \frac{\varepsilon}{\Delta^2 b}\right\}}{L} \min\left\{\Delta_k, \frac{\varepsilon}{M_k}\right\} \\ &= R_k - \varphi_2 \min\left\{\Delta_k, \frac{\varepsilon}{M_k}\right\}, \end{aligned}$$

where  $\varphi_2 = \frac{\beta(1 - \gamma) \tau^2 \varepsilon^2 \min\left\{\frac{1}{\Delta}, \frac{\varepsilon}{\Delta^2 b}\right\}}{L}$ .

Let  $\varphi = \min\{\varphi_1, \varphi_2\}$ , we can conclude

$$f_{k+1} \leq R_k - \varphi \min\{\Delta_k, \varepsilon/M_k\}, \quad \forall k \in K. \quad (15)$$

Considering Lemma 2 and (15), we get (14).

**Lemma 4** Suppose that all conditions of Lemma 1 and Lemma 3 hold. Then for all sufficiently large  $k$ , we have

$$\Delta_k > \varepsilon/M_k. \quad (16)$$

**Proof.** The proof is similar to Lemma 3.6 and 3.7 in [16], we omit it for convenience.

**Lemma 5** Suppose that (H2) and (H3) hold, if there exists a constant  $\varepsilon > 0$  such that  $\|g_k\| \geq \varepsilon$ , then for a sufficiently large integer  $\bar{k}$ , we have  $\sum_{k=\bar{k}}^{+\infty} \frac{1}{M_k} < +\infty$ .

**Proof.** Assume that for  $k \geq \bar{k}$ , (16) holds. Using the fact and (15), we obtain

$$f_{l(k)} = \eta_k f_{l(k)} + (1 - \eta_k) f_{l(k)} \geq R_k \geq f_{k+1} + \varphi \min\left\{\Delta_k, \frac{\varepsilon}{M_k}\right\} \geq f_{k+1} + \frac{\varphi\varepsilon}{M_k}.$$

Then, the monotonicity of  $\{M_k\}$  imply that

$$\begin{aligned} f_{l(k)} &\geq f_{k+1} + \frac{\varphi\varepsilon}{M_k} \geq f_{k+1} + \frac{\varphi\varepsilon}{M_{k+M+1}} \\ f_{l(k+1)} &\geq f_{k+2} + \frac{\varphi\varepsilon}{M_{k+1}} \geq f_{k+2} + \frac{\varphi\varepsilon}{M_{k+M+1}} \\ &\vdots \\ f_{l(k+M)} &\geq f_{k+M+1} + \frac{\varphi\varepsilon}{M_{k+M}} \geq f_{k+M+1} + \frac{\varphi\varepsilon}{M_{k+M+1}}. \end{aligned}$$

Thus, for  $s = 0, 1, \dots, M$ , we have

$$f_{l(k+s)} \geq f_{k+s+1} + \frac{\varphi\varepsilon}{M_{k+s}} \geq f_{k+s+1} + \frac{\varphi\varepsilon}{M_{k+M+1}}.$$

Using Lemma 2, we get

$$f_{l(k)} \geq \max\{f_{k+1}, f_{k+2}, \dots, f_{k+M+1}\} + \frac{\varphi\varepsilon}{M_{k+M+1}} = f_{l(k+M+1)} + \frac{\varphi\varepsilon}{M_{k+M+1}}.$$

This inequality and the fact that  $\{f_{l(k)}\}$  is convergent (see [4]) show

$$\varphi\varepsilon \sum_{k=\bar{k}}^{+\infty} \frac{1}{M_{k+M+1}} \leq \sum_{k=\bar{k}}^{+\infty} (f_{l(k)} - f_{l(k+M+1)}) = \sum_{k=\bar{k}}^{+\infty} \sum_{s=0}^M (f_{l(k+s)} - f_{l(k+s+1)}) \leq \sum_{s=0}^M \sum_{k=1}^{+\infty} (f_{l(k+s)} - f_{l(k+s+1)}).$$

Therefore,  $\sum_{k=\bar{k}}^{+\infty} \frac{1}{M_{k+M+1}} \leq \sum_{s=0}^M \sum_{k=1}^{+\infty} (f_{l(k+s)} - f_{l(k+s+1)}) / \varphi\varepsilon < +\infty$ . Then we have  $\sum_{k=\bar{k}}^{+\infty} \frac{1}{M_k} < +\infty$ .

**Theorem 6** Suppose that (H1)-(H4) hold, then sequence  $\{x_k\}$  generated by Algorithm 0 satisfies

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (17)$$

**Proof.** Assume that (17) does not hold, then there exists a constant  $\varepsilon > 0$  such that  $\|g_k\| \geq \varepsilon$ . From (H2) and the definition of  $M_k$ , we have

$$\sum_{k=\bar{k}}^{+\infty} \frac{1}{M_k} = \sum_{k=\bar{k}}^{+\infty} \frac{1}{1 + \max_{1 \leq i \leq k} \|B_k\|} \geq \sum_{k=\bar{k}}^{+\infty} \frac{1}{b+1} \rightarrow +\infty.$$

Obviously, this contradicts Lemma 5. The proof is completed.

## 5. Conclusion

In this paper, we introduce the algorithm of non-monotone trust region method with non-monotone Wolfe-type line search strategy for unconstrained optimization problems based on (5), (6) and (7). When compared with (3), it is obviously that we fully employ the current objective function value  $f_k$  and we can derive the better convergence results by choosing an adaptive  $\eta_k$ . Besides, with the help of line search strategy, new algorithm can reduce the number of solving sub-problems. We analyzed the properties of the algorithm and proved the global convergence theory under some mild conditions.

## Acknowledgements

This work is supported by the National Natural Science Foundation of China (61473111) and the Natural Science Foundation of Hebei Province (Grant No. A2014201003, A2014201100).

## References

- [1] Nocedal, J. and Yuan, Y.X. (1998) Combining Trust Region and Line Search Techniques. In: Yuan, Y., Ed., *Advanced in Nonlinear Programming*, Kluwer Academic Publishers, Dordrecht, 153-175. [http://dx.doi.org/10.1007/978-1-4613-3335-7\\_7](http://dx.doi.org/10.1007/978-1-4613-3335-7_7)
- [2] Michael Gertz, E. (2004) A Quasi-Newton Trust Region Method. *Mathematical Programming*, **100**, 447-470. <http://dx.doi.org/10.1007/s10107-004-0511-1>
- [3] Chamberlain, R.M. and Powell, M.J.D. (1982) The Watchdog Technique for Forcing Convergence in Algorithm for Constrained Optimization. *Mathematical Programming Study*, **16**, 1-17. <http://dx.doi.org/10.1007/BF0120945>
- [4] Grippo, L., Lampariello, F. and Lucidi, S. (1986) A Nonmonotone Line Search Technique for Newton's Method. *Society for Industrial and Applied Mathematics*, **23**, 707-716.
- [5] Deng, N.Y., Xiao, Y. and Zhou, F.J. (1993) Nonmonotone Trust Region Algorithm. *Journal of Optimization Theory and Application*, **76**, 259-285. <http://dx.doi.org/10.1007/BF00939608>
- [6] Toint, Ph.L. (1996) An Assessment of Nonmonotone Linesearch Technique for Unconstrained Optimization. *Society for Industrial and Applied Mathematics*, **17**, 725-739.
- [7] Toint, Ph.L. (1997) Non-Monotone Trust-Region Algorithm for Nonlinear Optimization Subject to Convex Constraints. *Mathematical Programming*, **77**, 69-94. <http://dx.doi.org/10.1007/BF02614518>
- [8] Sun, W.Y. (2004) Nonmonotone Trust Region Method for Solving Optimization Problems. *Applied Mathematics and Computation*, **156**, 159-174. <http://dx.doi.org/10.1016/j.amc.2003.07.008>
- [9] Mo, J.T., Zhang, K.C. and Wei, Z.X. (2005) A Nonmonotone Trust Region Method for Unconstrained Optimization. *Applied Mathematics and Computation*, **171**, 371-384. <http://dx.doi.org/10.1016/j.amc.2005.01.048>
- [10] Mo, J.T., Liu, C.Y. and Yan, S.C. (2007) A Nonmonotone Trust Region Method Based on Nonincreasing Technique of Weighted Average of the Successive Function Values. *Journal of Computational and Applied Mathematics*, **209**, 97-108. <http://dx.doi.org/10.1016/j.cam.2006.10.070>
- [11] Ahoosh, M. and Amini, K. (2012) An Efficient Nonmonotone Trust-Region Method for Unconstrained Optimization. *Numerical Algorithms*, **59**, 523-540. <http://dx.doi.org/10.1007/s11075-011-9502-5>
- [12] Ahoosh, M., Amini, K. and Bahrami, S. (2012) A Class of Nonmonotone Armijo-Type Line Search Method for Unconstrained Optimization. *Optimization*, **61**, 387-404. <http://dx.doi.org/10.1080/02331934.2011.641126>
- [13] Zhou, Q.Y., Chen, J. and Xie, Z.W. (2014) A Nonmonotone Trust Region Method Based on Simple Quadratic Models. *Journal of Computational and Applied Mathematics*, **272**, 107-115. <http://dx.doi.org/10.1016/j.cam.2014.04.026>
- [14] Huang, S., Wan, Z. and Chen, X.H. (2015) A New Nonmonotone Line Search Technique for Unconstrained Optimization. *Numerical Algorithms*, **68**, 671-689. <http://dx.doi.org/10.1007/s11075-014-9866-4>
- [15] Zhou, Q.Y. and Hang, D. (2015) Nonmonotone Adaptive Trust Region Method with Line Search Based on New Diagonal Updating. *Applied Numerical Mathematics*, **91**, 75-88. <http://dx.doi.org/10.1016/j.apnum.2014.12.009>
- [16] Gu, N.Z. and Mo, J.T. (2008) Incorporating Nonmonotone Strategies into the Trust Region for Unconstrained Optimization. *Computers and Mathematics with Applications*, **55**, 2158-2172. <http://dx.doi.org/10.1016/j.camwa.2007.08.038>
- [17] Ahoosh, M., Amini, K. and Peyghami, M.R. (2012) A Nonmonotone Trust-Region Line Search Method for Large-Scale Unconstrained Optimization. *Applied Mathematical Modelling*, **36**, 478-487. <http://dx.doi.org/10.1016/j.apm.2011.07.021>
- [18] Zhang, H.C. and Hager, W.W. (2004) A Nonmonotone Line Search Technique and Its Application to Unconstrained Optimization. *SIAM Journal on Optimization*, **14**, 1043-1056. <http://dx.doi.org/10.1137/S1052623403428208>