**Scholars Bulletin**
(A Multidisciplinary Journal)
An Official Publication of "Scholars Middle East Publishers",
Dubai, United Arab Emirates
Website: http://scholarsbulletin.com/

**ISSN 2412-9771 (Print)**
**ISSN 2412-897X (Online)**

# A non-monotone trust region method

**Liran Yang[1], Qinghua Zhou[1]***
[1]*Key Lab. of Machine Learning and Computational Intelligence, College of Mathematics and Information Science, Hebei University, Baoding, 071002, Hebei Province, China*

***Corresponding Author:**
Qinghua Zhou
Email: qinghua.zhou@gmail.com

**Abstract:** In this paper, we propose a non-monotone trust region algorithm based on conic model for unconstrained optimization problems. The method combines a non-monotone technique and a trust region method. Different from the usual trust region methods, the sub-problem of our new algorithm is the conic minimization sub-problem. The theoretical analysis indicates that the new method has the global convergence under some reasonable conditions.
**Keywords:** non-monotone technique; trust region method; line search; unconstrained optimization; conic model; global convergence.

## INTRODUCTION

Consider the following large unconstrained optimization problem:

$$\min f(x), \quad x \in R^n \tag{1.1}$$

where $f(x): R^n \to R$ is a continuously differentiable function.

There are many methods to solve problem (1.1) such as line search and trust region method. Trust region methods for unconstrained optimization problem have been studied by many researchers [1-3] because it has strong convergence and robustness, and it does not need to require the approximate Hessian matrix of the trust region sub-problem to be positive definite.

In recent years, a variety of trust region methods have been proposed in the literatures. Nocedal and Yuan [4] and Gertz [5] presented methods which combine line search and trust region method.

The non-monotone trust region methods listed above are all based on quadratic model, but we have not seen any non-monotone trust-region methods based on conic model. The traditional non-monotone trust region methods sometimes can not deal with the problems very well, especially when the objective function has strong non-quadratic.

In 1980, Davidon [6] presented a conic model methods for unconstrained optimization. A typical conic model as follow:

$$\min \; c_k(s) = f_k + \frac{g_k^T s}{1-h_k^T s} + \frac{1}{2}\frac{s^T B_k s}{(1-h_k^T s)^2} \tag{1.2}$$

In 2008, Qu *et al.* [7] proposed a new trust region sub-problem based on conic model for unconstrained optimization:

$$\begin{aligned} \min \; & c_k(s) = f_k + \frac{g_k^T s}{1-h_k^T s} + \frac{1}{2}\frac{s^T B_k s}{(1-h_k^T s)^2}, \\ s.t. \; & 1 - h_k^T s > 0, \\ & \|s\| \le \Delta_k, \end{aligned} \tag{1.3}$$

where $c_k(s)$ is called conic model which is an approximation to $f(x_k + s_k) - f(x_k)$ and $h_k$ is the associated vector for conic model and it is usually called horizontal vector, $\Delta_k$ is conic trust region radius.

Recently, non-monotone techniques have been studied by many authors since Grippo et al. [8]. Zhang et al. [9] proposed another non-monotone line search method. In detail, their method finds a step-size $\alpha_k$ satisfying the following condition:

$$f(x_k+\alpha_k d_k)\le C_k+\beta\alpha_k\nabla f(x_k)^T d, \tag{1.4}$$

where

$$C_k=\begin{cases} f(x_k), & k=0,\\ \frac{\eta_{k-1}Q_{k-1}C_{k-1}+f(x_k)}{Q_k}, & \text{k}\ge 1\ , \end{cases}\qquad Q_k=\begin{cases}1, & k=0,\\ \eta_{k-1}Q_{k-1}+1, & k\ge 1,\end{cases} \tag{1.5}$$

and $\eta_{k-1}\in[\eta_{\min},\eta_{\max}]$, $\eta_{\min}\in[0,1)$ and $\eta_{\max}\in[\eta_{\min},1)$ are two chosen parameters.

Many authors have generalized the non-monotone strategy to trust region methods and presented non-monotone trust region methods [10-11].

In our paper, we combine the sub-problem (1.3) with non-monotone technique proposed in [9] and the new line search technique proposed by Shi [12] to propose the new algorithm. This paper is organized as follows. In the next section, we describe our new non-monotone trust region method with line search based on conic model. The properties of this new algorithm and the global convergence property are given in Section 3. Finally, some concluding remarks are addressed in Section 4.

**New algorithm**

In this section, we describe our new algorithm. We obtain the trial step $s_k$ by solving the conic model sub-problem (1.3).

In order to give simple algorithms for solving sub-problem (1.3), we will first simplify the sub-problem (1.3). We consider $\chi(x_k)I$ as an approximation of $B_k$, where $\chi(x_k)\in R$. And the $\chi(x_k)$ can be expressed as [12]

$$\chi(x_k)=\begin{cases}\frac{2}{s_{k-1}^T s_{k-1}}[f(x_{k-1})-f(x_k)+g_k^T s_{k-1}], & if\ \ f(x_{k-1})-f(x_k)+g_k^T s_{k-1}>0,\\ \frac{2\delta}{s_{k-1}^T s_{k-1}}, & otherwise.\end{cases} \tag{2.1}$$

So, the sub-problem (1.3) can be modified as

$$\begin{aligned} \min\quad & c_k(s)=f_k+\frac{g_k^T s}{1-h_k^T s}+\frac{1}{2}\frac{\chi(x_k)s^T s}{(1-h_k^T s)^2},\\ s.t.\quad & 1-h_k^T s>0,\\ & \|s\|\le\Delta_k, \end{aligned} \tag{2.2}$$

**Algorithm 2.1**

Step 1. Given $x_0\in R^n$, $\Delta_0>0$, $0<c_0<1$, $0<c_1<c_2<1<c_3$, $\delta\in(\frac{1}{2},1)$, $\varepsilon>0$, $h_0=0$, $\chi_0=1$, $\theta>0$, $\beta\in(0,1)$. Choose parameters $\eta_{\min}\in[0,1)$ and $\eta_{\max}\in[\eta_{\min},1)$. Set $k=0$.

Step 2. Compute $g_k$. If $\|g_k\|\le\varepsilon$, stop. Otherwise, go to Step 3.

Step 3. Solve the sub-problem (2.2) for $s_k$.

Step 4. Compute $C_k$, $Ared_k=C_k-f_{k+1}$, $Pred_k=c_k(0)-c_k(s_k)$ and $r_k=\frac{Ared_k}{Pred_k}$.

Step 5. If $r_k\ge c_0$, set $x_{k+1}=x_k+s_k$ and go to Step 8. Otherwise go to Step 6.

Step 6. Set $v_k=-\frac{g_k^T s_k}{\chi_k\|s_k\|^2}$, select $\alpha_k$, which is the largest number in $\{v_k,\beta v_k,\beta^2 v_k,\ldots\}$ such that

$$f(x_k+\alpha s_k)\le C_k+\delta\alpha(g_k^T s_k+\tfrac{1}{2}\alpha\chi_k s_k^T s_k) \tag{2.3}$$

Set $x_{k+1}=x_k+\alpha_k s_k$.

Step 7. Compute $\chi(x_{k+1})$. If $\chi(x_{k+1})\le\varepsilon$ or $\chi(x_{k+1})\ge\frac{1}{\varepsilon}$, set $\chi(x_{k+1})=\theta$, and $B_{k+1}=\chi(x_{k+1})I$.

Step 8. Update the trust region radius $\Delta_{k+1}$ by

$$\Delta_{k+1} \begin{cases} \in [c_1\|s_k\|, c_2\Delta_k], & if\ r_k < c_0 \\ \Delta_k, & if\ r_k \ge c_0, and\ \|s_k\| < \Delta_k \\ \in [\Delta_k, c_3\Delta_k], & if\ r_k \ge c_0, and\ \|s_k\| = \Delta_k \end{cases} \tag{2.4}$$

Update $h_k$ by

$$h_k = (1-\gamma) g_k / g_k^T s_k \tag{2.5}$$

where $\gamma = \begin{cases} \frac{f_k - f_{k+1} + \sqrt{\rho}}{-g_k^T s_k}, & \rho \ge 0; \\ 0, & \rho < 0, \end{cases}$ and $\rho = (f_k - f_{k+1})^2 - (g_{k+1}^T s_k)(g_k^T s_k)$.

Choose $\eta_k \in [\eta_{\min}, \eta_{\max}]$. Set $k = k+1$, go to Step 2.

**Remark:** 1) It is obvious that for all $k$,

$$0 < m_1 = \min(\varepsilon, \theta) \le \chi(x_k) \le \max(\tfrac{1}{\varepsilon}, \theta) = m_2 \tag{2.6}$$

2) In order to guarantee the global convergence of the algorithm, we select a small enough constant $\sigma > 0$ and we have $\Delta_k \|h_k\| \le \sigma$ for all $k$.

3) In order to ease of reference, we define two index sets as below:

$I = \{k | r_k \ge c_0\}$ and $J = \{k | r_k < c_0\}$.

## CONVERGENCE ANALYSIS

In this section, we prove the global convergence property of Algorithm 2.1. The following assumptions are necessary to analyze the convergence.

**(H1)** The level set $L(x_0) = \{x \in R^n \,|\, f(x) \le f(x_0)\}$ is bounded for any given $x_0 \in R^n$.

**(H2)** There exists a constant $K_1 > 0$, such that $K_1 \|s\|^2 \le s^T B_k s$ for all $k$.

**(H3)** There exists a constant $\sigma \in (0,1)$, such that, for all $k$, $\|h_k\| \Delta_k \le \sigma$.

**Remark:** There is a positive number $\tau > 0$ such that $\|g_k\| \ge \tau$ for all $k$, then there exists a $\bar{\Delta} > 0$, such that for all $k$, we have $\Delta_k \ge \bar{\Delta}$.

**Lemma 3.1.** If $s_k$ is the solution of sub-problem (1.2) and suppose that (H1-H3) hold. Then there exist a positive scalar $v$ such that

$$Pred_k \ge v\|g_k\| \min\{\Delta_k, \tfrac{\|g_k\|}{\chi_k}\} \ge v\|g_k\| \min\{s_k, \tfrac{\|g_k\|}{\chi_k}\} \tag{3.1}$$

$$g_k^T s_k \le -(1-\sigma) v \|g_k\| \min\{\Delta_k, \tfrac{\|g_k\|}{\chi_k}\} \le 0 \tag{3.2}$$

hold for all $k$.

Proof. The proof is analogous to Theorem 3.1 in [11].

**Lemma 3.2.** Let $\{x_k\}$ be the sequence generated by Algorithm 2.1. Then, we have

$$f_{k+1} \le C_{k+1} \le C_k \tag{3.3}$$

Proof. If $k \in I$, i.e., $r_k \ge c_0$. By the definition of $r_k$ and $r_k \ge c_0$, we have

$$C_k - f_{k+1} \ge c_0 Pred_k \ge c_0 v \|g_k\| \min\{\Delta_k, \tfrac{\|g_k\|}{\chi_k}\} \ge 0 \tag{3.4}$$

Therefore, we have $C_k \ge f_{k+1}$.

If $k \in J$, i.e., $r_k < c_0$. From (2.3), we have

$$\begin{aligned} f_{k+1} - C_k &\le \delta\alpha_k (g_k^T s_k + \tfrac{1}{2}\alpha_k \chi_k s_k^T s_k) \\ &\le \delta\alpha_k (g_k^T s_k + \tfrac{1}{2}\nu_k \chi_k s_k^T s_k) \\ &= \tfrac{1}{2}\delta\alpha_k g_k^T s_k < 0 \end{aligned} \tag{3.5}$$

So, we have $C_k \ge f_{k+1}$.

From (1.5), (3.4) and (3.5), we have

$$C_k = \frac{\eta_{k-1}Q_{k-1}C_{k-1} + f_k}{Q_k} \ge \frac{\eta_{k-1}Q_{k-1}f_k + f_k}{Q_k} = f_k \tag{3.6}$$

These inequalities yield $f_{k+1} \le C_{k+1} \le C_k$.

**Lemma 3.3.** Suppose that assumptions hold, the sequence $\{x_k\}$ generated by Algorithm 2.1 is contained in level set $L(x_0)$.

Proof. The assertion follows from lemma 3.2, (H1-H3) and $C_0 = f_0$.

**Lemma 3.4.** Algorithm 2.1 is well defined, i.e., it could not cycle infinitely between step 6 and step 7.

Proof . Suppose that Algorithm 2.1 cycles infinitely between step 6 and step 7. From Taylor expansion and the lemma 2, we get

$$\begin{aligned} f(x_k + \alpha_k s_k) &= f_k + \alpha_k g_k^T s_k + \tfrac{1}{2} o(\alpha_k) \\ &\le C_k + \delta\alpha_k g_k^T s_k \\ &\le C_k + \delta\alpha_k (g_k^T s_k + \tfrac{1}{2}\alpha_k \chi_k s_k^T s_k) \end{aligned} \tag{3.7}$$

From (3.7), we know the step 6 and step 7 of the algorithm terminates in a finite number of the steps.

**Lemma 3.5.** Suppose that Assumption (H1-H3) hold, there exist $\overline{\omega}$ for all $k \in J$, we have

$$\alpha_k \ge \overline{\omega}. \tag{3.8}$$

Proof. The proof is similar to lemma 3.4 in [13].

**Lemma 3.6.** (See Lemma 3.6 in [13]) Suppose that (H1-H3) hold, and the sequence $\{x_k\}$ is generated by Algorithm 2.1. Then

$$(1-\eta_{\max})\gamma \|g_k\| \min\{\Delta_k, \tfrac{\|g_k\|}{\chi_k}\} \le C_k - C_{k+1} \tag{3.9}$$

**Theorem 3.7.** Suppose that (H1-H3) hold. Algorithm 2.1 either terminates in finite iterations, or generates an infinite sequence $\{x_k\}$ which satisfies

$$\liminf_{k\to\infty} \|g_k\| = 0. \tag{3.10}$$

Proof. If Algorithm 2.1 terminates in finite iterations, the theorem is obviously true. Assume Algorithm 2.1 generates an infinite sequence $\{x_k\}$ in the following proof.

We assume that there exists a positive constant $\tau > 0$, such that

$$\|g_k\| \ge \tau \quad \text{for all } k. \tag{3.11}$$

By (3.9), we have

$$C_k - C_{k+1} \ge (1-\eta_{\max})\gamma \|g_k\| \min\{\overline{\Delta}, \tfrac{\tau}{\chi_k}\} \tag{3.12}$$

It follows that Lemma 3.2 that $f_k \le C_k$ for all $k$ and $\{C_k\}$ is decreasing. Thus, by assumptions and Lemma 3.3 we have that $\{f_k\}$ is bounded below. Therefore, $\{C_k\}$ is convergent. From (3.12) that

$$\sum_{k=0}^{\infty} \frac{\min\{\overline{\Delta}, \frac{\tau}{\max(\frac{1}{\varepsilon},\theta)}\}}{Q_{k+1}} < \infty \tag{3.13}$$

By (1.5), we know $Q_0 = 1$ and $\eta_k \in [0,1)$, we have

$$Q_{k+1} = 1 + \sum_{i=0}^{k} \prod_{m=0}^{i} \eta_{k-m} \le 1 + \sum_{i=0}^{k} \eta_{\max}^{i+1} \le \sum_{i=0}^{\infty} \eta_{\max}^{i} = \frac{1}{1-\eta_{\max}} \tag{3.14}$$

Set $\min\{\bar{\Delta}, \frac{\tau}{\max(\frac{1}{\varepsilon},\theta)}\} = \lambda$. Formula (3.13) can be written as

$$\sum_{k=0}^{\infty} k(1-\eta_{\max})\lambda < \infty \tag{3.15}$$

We know $\sum_{k=0}^{\infty} k(1-\eta_{\max})\lambda$ is not convergent. This is a contradiction with Formula (3.15). Theorem 3.7 has been proved.

## CONCLUSIONS

In this paper, we give a new non-monotone conic trust region method with line search for unconstrained optimization. In the algorithm, we use a new search technique that is called large step search criteria, and it increase line search step length factor. Under some mild conditions, we establish the global convergence result of the proposed method.

## ACKNOWLEDGMENTS

This work is supported by the Natural Science Foundation of Hebei Province (Grant No. A2014201003, A2014201100) and the Post-graduate's Innovation Fund Project of Hebei University (No.X2015060).

## REFERENCES

1. Fletcher, F., (1980). Practical Methods of Optimization: *Unconstrained Optimization*, vol. 1.
2. Powell, M. J. D. (1984). On the global convergence of trust region algorithms for unconstrained minimization. *Mathematical Programming*, *29*(3), 297-303.
3. Y. Yuan, Y., Sun, W., (1997). Optimization Theory and Methods, *Science Press of China*.
4. Nocedal, J., & Yuan, Y. X. (1998). *Combining trust region and line search techniques* (pp. 153-175). Springer US.
5. Gertz, E. M. (1999). *Combination trust-region line-search methods for unconstrained optimization*. University of California, San Diego.
6. W. C. Davidono, Conic approximation and collinear scaling for optimizers, SIAM Journal on Numerical Analysis, 17(2), 1980, 268-281.
7. Qu, S. J., Zhang, K. C., & Zhang, J. (2008). A nonmonotone trust-region method of conic model for unconstrained optimization. *Journal of Computational and Applied Mathematics*, *220*(1), 119-128.
8. Grippo, L., Lampariello, F., & Lucidi, S. (1986). A nonmonotone line search technique for Newton's method. *SIAM Journal on Numerical Analysis*, *23*(4), 707-716.
9. Zhang, H., & Hager, W. W. (2004). A nonmonotone line search technique and its application to unconstrained optimization. *SIAM Journal on Optimization*, *14*(4), 1043-1056.
10. Deng, N., Xiao, Y., Zhou, F., (1993). Nonmontonic trust region algorithm, *Journal of Optimization Theory and Application*, 76(2), 259-285.
11. Mo, J., Liu, C., & Yan, S. (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*(1), 97-108.
12. Shi, Z. J. (2006). Convergence of quasi-Newton method with new inexact line search. *Journal of mathematical analysis and applications*, *315*(1), 120-131.
13. Cui, Z., Wu, B., & Qu, S. (2011). Combining nonmonotone conic trust region and line search techniques for unconstrained optimization. *Journal of computational and applied mathematics*, *235*(8), 2432-2441.