Volume : 4 | Issue : 8 | Aug 2015 • ISSN No 2277 - 8179

Research Paper

A Novel Non-Monotone Self-Adaptive Trust Region Methodfor Unconstrained Optimization

Mathematics

KEYWORDS : non-monotonestrategy; self-adaptive trust region method; unconstrained optimization; global convergence.

Liran Yang Collegeof Mathematics andInformationScience,HebeiUniversity, Baoding,071002,HebeiProvince,China

Qinghua Zhou Collegeof Mathematics andInformationScience,HebeiUniversity, Baoding,071002,HebeiProvince,China *Corresponding author

ABSTRACT

In this paper, we propose a new non-monotone self-adaptive trust region method for solvingunconstrained optimization problem. Different fromthe usual trust region methods, our newalgorithm uses a new rule to update the trust region radius.Theoretical analysis indicates that the new method preserves the global convergence under some reasonable assumptions.

1. Introduction

Consider the following large unconstrained optimization problem:

$$\min f(x), \quad x \in R^n \qquad (1.1)$$

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

For a given iteration point x_k, the trust region methodscompute a trial step d_k by solving the following quadratic sub-problem:

$$\begin{aligned} \min \quad & q_k(d) = f_k + g_k^T d + \tfrac{1}{2} d^T B_k d, \\ s.t. \quad & \|d\| \le \Delta_k, \end{aligned} \qquad (1.2)$$

Where $f_k = f(x_k)$, $g_k = \nabla f(x_k)$, $B_k \in R^{n\times n}$ is a symmetric matrix which is the Hessian matrix or its approximation of $f(x)$ at the current point x_k, $\Delta_k > 0$ is the trust radius and $\|\cdot\|$ denotes to the Euclidean norm.The ratio r_k between the actual reduction $f(x_k) - f(x_{k+1})$ and the predicted reduction $q_k(0) - q_k(d_k)$ plays a key role to decide whether d_k is acceptable or not and how to adjust the trust region radius. The iteration is said to be successful whenever r_k greater than a positive constant η_1.This leads us to the new point x_{k+1}, and the trust region radius is updated.If not, the iteration is unsuccessful, and the trail point is rejected.Generally, trust region radius update rule can be described as follow:

$$\Delta_{k+1} = \begin{cases} [\gamma_1\Delta_k, \gamma_2\Delta_k], & if\ r_k < \eta_1; \\ [\gamma_2\Delta_k, \Delta_k], & if\ \eta_1 \le r_k \le \eta_2; \\ [\Delta_k, \infty), & if\ r_k \ge \eta_2. \end{cases}$$

(1.3)

Where the constants γ_1, γ_2, η_1 and η_2 satisfy

$$0 \le \eta_1 < \eta_2 < 1, \quad 0 < \gamma_1 \le \gamma_2 < 1 \qquad (1.4)$$

In the trust region method, the difficulty is how to adjust the trust region radius. In order to choose an adaptive trust region radius at each iteration, many adaptive trust region methods have

been Studiedin
In1997,Sartenaer [3]introducedastrategy that canautomatically determine the initial trust regionradius.Thestrategyrequiresad ditionale valuations of the objective function.Zhangetal. [4]presented anotherstrategyofupdating the trust regionradius.Their basic idea is originated from the following sub - problemin [5].

$$\begin{aligned} \min \quad & q_k(d) = g_k^T d + \tfrac{1}{2} d^T B_k d, \\ s.t. \quad & -\Delta_k \le d_i \le \Delta_k, \quad i = 1,2,\ldots,n, \end{aligned} \tag{1.5}$$

where $\Delta_k = c^p (\|g_k\|/\gamma)$, $0 < c < 1$, $\gamma = \min(\|B_k\|, 1)$ and p is a nonnegative integer.

Hei [10] and Walmget al.[11] proposed self-adaptive update method,respectively.The trust region radius are product of so-called R-function and the Λ-function, that is, $\Delta_{k+1} = R(r_k)\Delta_k$ and $\Delta_{k+1} = \Lambda(r_k)\Delta_k$.Recently, the L-function is introduced by Lu et al. in [12]. They presented a new self-adaptive trust region method, in which the radius is $\Delta_{k+1} = L(r_k)\Delta_k$.

Grippo et al. [13] firstly proposed a non-monotone line search for Newton's method. This algorithm accepts thestep-size α_k whether

$$f(x_k + \alpha_k d_k) \le f(x_{l(k)}) + \beta \alpha_k \nabla f(x_k)^T d, \tag{1.6}$$

where $\beta \in (0, \tfrac{1}{2})$, $f(x_{l(k)}) = \max_{0 \le j \le m_k} f(x_{k-j}), m_0 = 0, 0 \le m_k \le \min\{m_{k-1} + 1, M\} (k \ge 1)$, and $M \ge 0$ is an integer. It has been proved that the sequence $\{f(x_k)\}$ is not increasing. Since then, many researchers [4-5] have exploited the non-monotone technique and a lot of numerical tests have showed that the non-monotone technique proposed by Grippo et al. [13] is efficient at some extent. In 1993, Deng et al. in [14] made some changes and applied it to the trust region method, and proposed a non-monotone trust region method for unconstrained optimization. Theoretical analysis and numerical results show that algorithms with non-monotone strategy are more effective than algorithms without it.

Zhang and Hager [15] proposed another non-monotone line search method, they replaced the maximum function value with an average of function values. In detail, their method finds a step-size α_k satisfying the following condition:

$$f(x_k + \alpha_k d_k) \le C_k + \beta \alpha_k \nabla f(x_k)^T d, \tag{1.7}$$

where

$$C_k = \begin{cases} f(x_k), & k = 0, \\ \frac{\eta_{k-1} Q_{k-1} C_{k-1} + f(x_k)}{Q_k}, & k \geq 1, \end{cases} \qquad Q_k = \begin{cases} 1, & k = 0, \\ \eta_{k-1} Q_{k-1} + 1, & k \ge 1, \end{cases} \tag{1.8}$$

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. Numerical results showed that this non-monotone technique was superior to (1.6). Then, this non-monotone was applied to the trust region methods[16,17].In addition, many non-monotone adaptive trust region methods have been proposed in [6-9].

Inspired by the ideas introducedabove, we use the L-function to update the radius, then appliedit to the trust region method with non-monotone strategy proposed byZhang and Hager [15].The purpose of this paper is to present a new non-monotone adaptive trust region method.

Therest of the paper is organized as follows. In Section 2,we describe L-function and our new non-monotone self-adaptive trust region algorithm.Theglobal convergence properties of this novel algorithmare given in Section 3.Finally, some conclusions are summarized in Section 4.

2. L-function and Algorithm

The L-function rule proposed by Lu [14] can be described as follows:

$$\Delta_{k+1} = L(r_k)\Delta_k \tag{2.1}$$

where the L-function $L(r_k)$ is chosen as

$$L(r_k) = \begin{cases} c_1 + (c_2 - c_1)\exp(r_k), & if\ r_k \le 0, \\ \frac{1-\beta_1 \exp(\eta_2)}{1-\exp(\eta_2)} - \frac{(1-\beta_1)\exp(\eta_2)}{1-\exp(\eta_2)}\exp((r_k - \eta_2)), & if 0 < r_k < \eta_2, \\ \beta_2, & if\ \eta_2 \le r_k < 2-\eta_2, \\ \beta_3 + (\beta_2 - \beta_3)\exp(-(\frac{r_k+\eta_2-2}{\eta_2-2})^2), & if\ r_k > 2-\eta_2, \end{cases} \tag{2.2}$$

where β_1, β_2, β_3, c_1, c_2 and η_2 are constants.

Now describe the new non-monotone self-adaptive trust region algorithm with the new radius update rule.

When weobtain d_k ,then the ratio r_k is computed by

$$r_k = \frac{Ared_k}{\Pr ed_k} = \frac{C_k - f(x_k + d_k)}{q_k(0) - q_k(d_k)}, \tag{2.3}$$

Algorithm 2.1

Step 1. Given $x_0 \in R^n$, $\Delta_0 > 0$, $B_0 \in R^{n\times n}$, $0 < \eta_1 < \eta_2 < 1$, $0 < c_1 < c_2 < 1$, $\varepsilon \ge 0$,

$0 < \beta_1 \le \beta_3 < 1 \le \beta_2$,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 (1.2) for d_k. Compute C_k, $Ared_k$, $Pred_k$ and r_k.

Step 4. If $r_k > \eta_1$, set $x_{k+1} = x_k + d_k$, otherwise, $x_{k+1} = x_k$.

Step 5. Update the trust region radius Δ_{k+1} by (2.1) and (2.2).

Step 6. Compute g_{k+1} and B_{k+1}, and set $k := k+1$, go to Step 2.

3. Convergence

In this section, we will prove the global convergence propertiesof Algorithm 2.1. The following assumptionsarenecessary to analyze the convergence properties.

(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) Thematrix B_k isuniformlybounded,i.e.,thereexistsaconstant $M_0 > 0$,suchthat,forall k ,

$$\|B_k\| \le M_0 .$$

Lemma 3.1. If d_k is the solution to sub-problem (1.2), then

$$Pred_k = q_k(0) - q_k(d_k) \ge \tfrac{1}{2}\|g_k\| \min\{\Delta_k, \tfrac{\|g_k\|}{\|B_k\|}\} . \tag{3.1}$$

$$g_k^T d_k \le -\tfrac{1}{2}\|g_k\| \min\{\Delta_k, \tfrac{\|g_k\|}{\|B_k\|}\} \quad (3.2)$$

Proof. From Lemma 13.3.1 in [18], we know (3.1) holds. And from (3.1), we can see

$$Pred_k = -g_k^T d_k - \tfrac{1}{2} d_k^T B_k d_k \ge \tfrac{1}{2}\|g_k\| \min\{\Delta_k, \tfrac{\|g_k\|}{\|B_k\|}\}$$

Considerthe above inequality and the fact $d_k^T B_k d_k > 0$, (3.2) obviously holds. Therefore, the lemma is true.

Lemma 3.2. Let $\{x_k\}$ be the sequence generated by Algorithm 2.1. For any fixed $k \ge 0$, we have

$$f_{k+1} \le C_{k+1} \le C_k \quad (3.3)$$

Proof. Let $k \ge 0$ be an arbitrary fixed integer.By the definition of r_k and $r_k > \eta_1$, we have

$$\begin{aligned} C_k - f_{k+1} &\ge \eta_1 Pred_k = \eta_1(q_k(0) - q_k(d_k)) \\ &\ge \tfrac{\eta_1}{2}\|g_k\| \min\{\Delta_k, \tfrac{\|g_k\|}{\|B_k\|}\} \ge 0 \end{aligned} \quad (3.4)$$

Thus, $C_k \ge f_{k+1}$. Then, by the definition of C_k, we obtain that

$$C_k = \tfrac{\eta_{k-1}Q_{k-1}C_{k-1} + f_k}{Q_k} \ge \tfrac{\eta_{k-1}Q_{k-1}f_k + f_k}{Q_k} = f_k .$$

So

$$C_k \ge f_k . \tag{3.5}$$

On the other hand, we have

$$C_{k+1} = \tfrac{\eta_k Q_k C_k + f_{k+1}}{Q_{k+1}} \le \tfrac{\eta_k Q_k C_k + C_k}{Q_{k+1}} = C_k . \tag{3.6}$$

From (3.5) and (3.6), Lemma 3.2 holds.

Lemma 3.3. Suppose that the sequence $\{x_k\}$ is generated by Algorithm 2.1. The algorithm is well defined.

Proof .The process is similar to Lemma 3.4 in [19].

Lemma 3.4. Suppose that(H1)-(H2) hold. Then

$$|f(x_k+d_k)-q(d_k)|=O(\|d_k\|^2) \qquad (3.7)$$

Proof. From the Taylor expansion, the lemma is true.

Lemma 3.5.(See Lemma 3.6 in [19]) Suppose that Assumption 1 holds, and there is a positive number $\tau>0$ such that $\|g_k\|\geq\tau$ for all k, then there exists a $\overline{\Delta}>0$, such that for all k, we have $\Delta_k\geq\overline{\Delta}$.

Theorem 3.6. Suppose that (H1)-(H2) hold. Let the sequence $\{x_k\}$ generated by Algorithm 2.1, then we have

$$\lim_{k\to\infty}\inf\|g_k\|=0. \quad (3.8)$$

Proof.We assume that Formula (3.8) is not true,that is, there exists a positive constant $\tau>0$, such that

$$\|g_k\|\geq\tau \quad \text{for all } k. \qquad (3.9)$$

By Lemma 3.1, Lemma 3.2, we have

$$f_{k+1}\leq C_k-\eta_1 Pred_k\leq C_k-\tfrac{1}{2}\eta_1\|g_k\|\min\{\Delta_k,\tfrac{\|g_k\|}{\|B_k\|}\} \qquad (3.10)$$

Then

$$\begin{aligned}C_{k+1}=\frac{\eta_kQ_kC_k+f_{k+1}}{Q_{k+1}}&\leq\frac{\eta_kQ_kC_k+C_k-\frac{1}{2}\eta_1\|g_k\|\min\{\Delta_k,\frac{\|g_k\|}{\|B_k\|}\}}{Q_{k+1}}\\&=C_k-\frac{\frac{1}{2}\eta_1\|g_k\|\min\{\Delta_k,\frac{\|g_k\|}{\|B_k\|}\}}{Q_{k+1}}\\&\leq C_k-\frac{\frac{1}{2}\eta_1\tau\min\{\overline{\Delta},\frac{\tau}{M_0}\}}{Q_{k+1}}\end{aligned} \qquad (3.11)$$

Which implies

$$C_k-C_{k+1}\geq\frac{\frac{1}{2}\eta_1\tau\min\{\overline{\Delta},\frac{\tau}{M_0}\}}{Q_{k+1}} \qquad (3.12)$$

It follows that lemma 3.2 that $f_k\leq C_k$ for all k and $\{C_k\}$ is decreasing. Thus, by assumptions, lemma 3.3 and Formula (3.21), we have that $\{f_k\}$ is bounded below. Therefore, $\{C_k\}$ is convergent. From (3.12) that

$$\sum_{k=0}^{\infty}\frac{\min\{\overline{\Delta},\frac{\tau}{M_0}\}}{Q_{k+1}}<\infty \quad (3.13)$$

By (1.8), we know $Q_0=1$ and $\eta_k\in[0,1)$, we have

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

Set $\min\{\overline{\Delta}, \frac{\tau}{M_1}\} = \lambda$. Formula (3.13) can be written as

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

we know $\sum_{k=0}^{\infty} k(1-\eta_{\max})\lambda$ is not convergent. This is a contradiction with Formula (3.15). Theorem 3.6 has been proved.

4. Conclusions

In this paper, we present a new non-monotoneself-adaptive trust region method. And the form of the new method is very simple. Under some mild conditions, we proved the global convergence result of the proposed method.

Acknowledgments

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).

REFERENCE

[1]. J.Y.Fan,W. B.Ai,Q.Y.Zhang,Alinesearchandtrustregionalgorithmwithtrust regionradiusconvergingtozero,JournalofComputationalMathematics,22(6),2004, 865–872. | [2]. E.M.Gertz,Aquasi-Newtontrust-regionmethod,MathematicalProgrammingSeriesA,100, 2004, 447–470. | [3]. A.Sartenaer,Automaticdeterminationofaninitialtrust regioninnonlinearprogramming, SIAMJournalonScientificComputing, 18(6), 1997,1788–1803. | [4]. X.Zhang,J.Zhang,L.Liao,Anadaptivetrustregionmethodanditsconvergence, ScienceinChinaA. 45(5), 2002,620–631. | [5]. X.Zhang,Trustregionmethodinneuralnetwork,ActaMathematicaeApplicataeSinica, 12, 1997,1–10. | [6]. M. Ahookhosh, K. Amini,An efficient nonmonotone trust-region method for unconstrained optimization,Journal of Numerical Algorithms, 59(4), 2012,523-540. | [7]. M.Ahookhosh, K.Amini,Anonmonotonetrustregionmethodwithadaptiveradiusforunconstrainedoptimizationproblems,Computers&MathematicswithApplications, 60(3), 2010,411–422. | [8]. Z.Shi, S. Wang,Nonmonotoneadaptivetrustregionmethod, EuropeanJournalofOperationalResearch, 208(1), 2011,28–36. | [9]. Z.Cui, B.Wu,Anewmodifiednonmonotoneadaptivetrustregionmethodforunconstrainedoptimization, ComputationalOptimizationandApplication,53(3), 2012,795–806. | [10]. L.Hei,Aself-adaptivetrustregionalgorithm, JournalofComputationalMathematics, 21(2), 2003,229–236. | [11]. J.M.B.Walmag, E.J.M.Delhez, Anoteontrust-regionradiusupdate, SIAMJournalonOptimization,16(2), 2005,548–562. | [12]. Y. Lu, W.Li, M. Cao, Y. Yang, A novel self-adaptivetrustregionalgorithm for unconstrained optimization,Journalof Applied Mathematics,2014(2014), 2014. | [13]. L. Grippo, F. Lamparillo, S. Lucidi, A nonmonotone line search technique for Newton's method, SIAM J. Numer. Anal. 23(4), 1986, 707-716. | [14]. N. Deng, Y. Xiao, F. Zhou, Nonmontonic trust region algorithm, Journal of Optimization Theory and Application, 76(2), 1993, 259-285. | [15]. H. Zhang, W. Hager, A nonmonotone line search technique and its application to unconstrained optimization,SIAM Journal on Optimization, 14(4), 2004, 1043-1056. | [16]. J. Mo, C. Liu, S. Yan, A nonmonotone trust region method based on nonincreasing technique of weighted average of the successive function values, Journal of Computation and AppliedMathematics, 209(1), 2007, 97-108. | [17]. Q. Wu, Nonmonotone trust region algorithm for unconstrained optimization problems, Applied Mathematicsand Computation, 217(8), 2010, 4274-4281. | [18]. Y. Yuan, W. Sun, Optimization Theory and Methods, Science Press of China, 1997. | [19]. Q. Zhou, J. Chen, Z. Xie, A nonmonotone trust region method based on simple quadratic models, Journal of Computational and Applied Mathematics, 272(1), 2014, 107-115.