سیزدهمین کنفرانس بین المللی انجمن ایرانی تحقیق در عملیات
۳ و ۴ اردیبهشت ۱۳۹۹
13th International Conference of Iranian Operations Research Society
22-23 April 2020

A nonmonotone adaptive method for solving the unconstrained optimization problems

Z. Saeidian[1]

[1] Assistant Professor, Department of Mathematics, University of Kashan; Saeidian@Kashanu.ac.ir.

ABSTRACT

In this paper, we propose a new technique for solving the large scale unconstrained optimization problems. Using a simple cubic regularization in the trust region subproblem, an efficient adaptive nonmonotone scheme is provided. Under some suitable and standard assumptions, we proof the global convergence properties of the new algorithm. Numerical experiments shows the efficiency of the new proposed algorithm.

Keywords : Trust region methods; Nonmonotone technique; Adaptive cubic regularization method; Global convergence.

1. INTRODUCTION

Consider the unconstrained optimization problem as follows:

$$min_{x \in R^n} f(x) \qquad (۱)$$

in which $f: R^n \to R$ is a twice continuously differentiable function. For solving (1) many effective iterative procedures is provided, however we can devided these schemes into two classes of methods, i.e., line search and trust region methods [1]. Moving along a direction that implies a sufficient reduction in the objective function, is the base of the line search methods. Whereas in the standard trust region method, the trial step is achived by minimizing a (quadratic or (recently) cubic) model of the objective function over a region around the current point. Then, we evaluate the suitability of the model and the objective function at the new point by trust region ratio that this value leads us for accepting or rejecting the new point.

Furthermore, the trust region radius is updated appropriately. Practically, Given x_k, the trial step d_k is computed by

$$\min q_k(d) = f_k + g_k^T d + \frac{۱}{۲} d^T B_k d \qquad (2)$$

$$s.t. \|d\| \le \Delta_k,$$

where $f_k = f(x_k)$, $g_k = \nabla f(x_k)$, B_k is the exact Hessian $\nabla^2 f(x_k)$, or it's approximation, and the trust region radius is $\Delta_k > 0$. Also, the trust region ratio is defined as follows:

$$r_k = \frac{f(x_k) - f(x_k + d_k)}{q_k(0) - q_k(d_k)} \tag{3}$$

where it is used for deciding whether the trial step is accepted or rejected. Then, the trust region radius is updeted appropriately. Studies Shows that if the sequence of the objective values is monotonically decreasing, it frequently cuases the low convergance rate. In order to overcome this difficulty, the nonmonotone trust region methods is proposed [11, 12].

Ahookhosh et al. [3,4} proposed a new efficient nonmonotone term as follows:

$$R_k = \varepsilon_k f_{l(k)} + (1 - \varepsilon_k) f_k, \tag{٤}$$

where $f_k = f(x_k)$, $\varepsilon_k \in [\varepsilon_{min}, \varepsilon_{max}] \subset [0,1]$ and

$$f_{l(k)} = \max_{\{0 \le j \le M(k)\}} f_{k_j}, \tag{٥}$$

that is the Grippo's nonmonotone term and $M(0) = 0$ and, for $k \geq 1$, $M(k) = \min\{k, M\}$, for given positive integer M.

Another important procedure in the trust region method is the process of updating the trust region radius at every iteration. Also, selecting the initial trust region radius is crucial, for more information see [2,17,18,19]. As some important factors in updating the trust region radius, we can denote to; using the more information of the objective function and avoiding to enforcement the costly calculation.

Becuase of the advantage of nonmonotone and adaptive schemes, we employ these techniques in the structures of the trust region algorithm. Using the full information of the last two iterates, Sang et al. [16] proposed the nonmonotone adaptive trust region method. The framework of this method is based on a simple subproblem and it is efficient for the large scale unconstriained optimization problems.

Hereupon, some applicable subproblems are introduced, see [9,15,21]. Recently, in order to use Cartis et al. ,in order to use the more information of the objective function, suggested the adaptive regularization method using cubic model of the objective function(ARC) [10]. Bacuase of the results that shows the efficieny of ARC method in comparison with the standard trust region method, many researchers attend to study about this manner [7,8]. In this way, using a real positive definite scalar matrix to approximate the exact Hessian, Zheng et al. modified the adaptive cubic regularization method for large-scale unconstrained optimization problem [20].
In this paper, we introduce a new nonmonotone adaptive trust region method based on the cubic regularization subproblem. Our method is equiped the nonmonotone technique of [3,4], and for constructing an approximation of the Hessian at the current point, it uses a slight modification of the secant condition in [13].
The global convergence is established under some standard assumptions.
The reminder of this paper is organized as follows: In Section 2, we present the structure of the new nonmonotone adaptive trust region method in details. The global convergence property is established in Section 3.

2. THE NEW ALGORITHM

In this section, we describe our proposed method. Consider the following cubic regularization model [10] of (1):

$$\min q_k(d) = g_k^T d + \frac{1}{2} d^T \gamma(x_k) d + \frac{1}{3}\kappa_k \|d\|^3 \tag{6}$$

$$s.t. \ \|d\| \leq \Delta_k$$

$$\nabla q_k(d) = g_k + \gamma_k d + \kappa_k d \|d\|.$$

where

As the trial step d_k is parallel to $-g_k$ we have:

$$d_k = -\alpha_k g_k \tag{7}$$

and by considering (6) we conclude that [13]:

$$\alpha_k = \frac{2}{\gamma_k + \sqrt{\gamma_k^2 + 4\kappa \|g_k\|}} \tag{8}$$

Using the trial step d_k, the nonmonotone ratio is computed by [5]:

$$r_k = \frac{(1+\Omega_k)R_k - f(x_k+d_k)}{Pred_k} \tag{9}$$

$$\Delta_k = \min\{\nu_k \frac{\|g_k\|}{\gamma_k}, \Delta_{max}\} \quad (۱۰)$$

where $\Delta_{max} > 0$ is a bounded value for the trust region radius [14] and

$$\nu_{k+1} = \begin{cases} \sigma_0 \nu_k & r_k < \mu_1 \\ \nu_k & \mu_1 \le r_k \le \mu_2 \\ \min\{\sigma_1 \nu_k, \nu_{max}\} & r_k > \mu_2 \end{cases} \quad (۱۱)$$

where $0 < \sigma_0 < 1 < \sigma_1$, $0 < \mu_1 < \mu_2 \le 1$ and $\nu_{max} > 0$ are given numbers. Now, if $r_k \ge \mu$ then the new point is introduced by $x_{k+1} = x_k + d_k$, otherwise, we set $x_{k+1} = x_k$.

In order to calculate the parameter γ_k we have [6,13]:

$$\gamma_k = \arg min_{\gamma} \left\| d_{k-1} - \frac{1}{\gamma} y_{k-1}{}^{\pi_k} \right\| = \frac{d_k^T y_k{}^{\pi_k}}{d_k^T d_k} \quad (۱۲)$$

where, $y_k^{\pi_k} = y_k + \frac{\pi}{2}\|d_k\| d_k$, $y_k = g_{k+1} - g_k$ and π_k is updated at each iteration as follows:

$$\pi_{k+1} = \begin{cases} \sigma_0 \pi_k & r_k < \mu_1 \\ \pi_k & \mu_1 \le r_k \le \mu_2 \\ \sigma_1 \pi_k & r_k > \mu_2 \end{cases} \quad (۱۳)$$

Now, we describe the structure of our algorithm:

A new algorithm

Input: $x_0 \in R^n, 0 < \mu < \mu_1 < \mu_2 \leq 1$, $0 < \sigma_0 < 1 < \sigma_1$, $0 < \varepsilon_{min} < \varepsilon_{max} < 1$,

$\varepsilon, \epsilon, M, \nu_{max}, \Delta_{max} > 0, 0 < \theta_1 < \theta_2$ and $\Delta_{max} > 0, 0 < \theta_1 < \theta_2, \delta > 0$.

Step 0: Set $k = 0$, $\gamma_0 := \gamma(x_0) = 1, g_0 = g(x_0), \nu_0 = 1$ and $\Delta_0 = \min\left\{\nu_0 \frac{\|g_0\|}{\gamma_0}, \Delta_{max}\right\}$.

Step 1: If $\|g_k\| \leq \epsilon,$ Then Stop.

Step 2: Determine d_k by solving $q_k(d_k) \leq q_k(d_k^C)$ where $d_k^C = -\alpha_k g_k$ is the Cauchy point and α_k is computed by (7). Also, compute r_k using (8).

Step 3: If $r_k < \mu$, Then set $\Delta_k = \sigma_0 \Delta_k$ and go to Step 2.

Step 4: Set $x_{k+1} = x_k + d_k$.

Step 5: Compute γ_{k+1} by (12).

If $\gamma_{k+1} \leq \varepsilon$, Then set $\gamma_{k+1} = \theta_1$. If $\gamma_{k+1} \geq \frac{1}{\varepsilon}$,Then set $\gamma_{k+1} = \theta_2$.

Step 6: Update ν_{k+1} using (11) and set $\Delta_{k+1} = \min\{\nu_{k+1} \frac{\|g_{k+1}\|}{\gamma_{k+1}}, \Delta_{max}\}$.

Set $k =: k + 1$ and go to Step 1.

.

3. Convergence analysis

For considering the global convergence, we assume some standard assumptions as:

A1. The set $\Omega = \{x \in R^n | f(x) \leq f(x_0)\}$ is a closed bounded set.

A2. $f(x)$ is a twice continuously differentiable function in Ω.

A3. The function $\nabla f(x)$ is a Lipschitz continuous function on Ω.

Lemma 3.1. Assume that d_k is a solution of the problem (7). Then,one has [14]:

$$Pred_k := q_k(0) - q_k(d_k) \geq \frac{1}{2}\|g_k\| \min\left\{\Delta_k, \frac{\|g_k\|}{\gamma_k}\right\}.$$

Proof. [14]

Theorem 3.2. Suppose that Assumption A1 to A3 holds and x_k is the sequence generated by the new algorithm. Then, the new algorithm either stops at a stationary point or

$\lim\ inf_{k\to\infty}\|g_k\|$ =0.

Proof. [13, 14].

10. ACKNOWLEDGMENT

The author would like to thank the Research Council of University of Kashan for supporting this work.

11. REFERENCES

[1] Nocedal, J. and Wright, S.J. ; *Numerical Optimization*, Springer, NewYork, 2006.

[2] Ahookhosh, M., Amini, K. ; "*A nonmonotone trust region method with adaptive radius for*

unconstrained optimization", Comput. Math.Appl., No.60, pp. 411-422, 2010.

[3] Ahookhosh, M., Amini, K.; "*An efficient nonmonotone trust-region method for unconstrained optimization*", Numer. Algorithms, No. 59, pp. 523-540, 2011.

[4] Ahookhosh, M., Amini, K., Peyghami, M.Reza. ; "*A nonmonotone trust-region line search method for large-scale unconstrained optimization*", Appl. Math. Model., No. 36, pp. 478-487, 2012.

[5] Ataee Tarzanagh, D., Peyghami, M. Reza., Mesgarani, H.;"*A new nonmonotone trust region method for unconstrained optimization equipped by an efficient adaptive radius"*}, Optim. Methods Softw., No. 29(4), pp. 819-836, 2014.

[6] Barzilai, J., Borwein, J.; "*Two-point step size gradient methods"*, IMA J. Numer.Anal., No.8 , pp. 141–148, 1988.

[7] Bergou, E., Diouane, Y. and Gratton, S.; "*On the use of the energy norm in trust region and adaptive cubic regularization subproblems"*, Comput. Optim. Appl., No.68, pp. 533–554, 2017.

[8] Bianconcini, T., Sciandrone, M.; "*A cubic regularization algorithm for unconstrained optimization using line search and nonmonotone techniques"*, Optim. Methods Softw. No. 31, pp.1008–1035, 2016.

[9] Biglari,F., Solimanpur, M.; "*Scaling on the Spectral Gradient Method"*, J. Optim. Theory Appl., No.158(2),pp. 626-635, 2018.

[10] Cartis, C., Gould, N.I.M. and Toint, P.L. ; "*Adaptive cubic regularisation methods for unconstrained optimization"*, Part I: motivation, convergence and numerical results, Math. Program., No.127, pp. 245–295, 2011.

[11] Grippo, L., Lampariello,F., Lucidi, S.; "*A nonmonotone line search technique for Newton's method"*, SIAM J. Numer. Anal., No.23, pp. 707-716, 1986.

[12] Grippo, L., Lampariello, F., Lucidi, S.; "*A truncated Newton method with nonmonotone line-search for unconstrained optimization*", J. Optim. Theory Appl., No. 60, pp. 401-419, 1989.

[13] Li, Q., Zheng, B., Zheng, Y.; "*An efficient nonmonotone adaptive cubic regularization method with line search for unconstrained optimization problem"*, J. Appl. math. Let., No. 98, pp. 74-80, 2019.

[14] Saeidian, Z., Peyghami, M.Reza.; "*An adaptive nonmonotone trust region method for unconstriained optimization problems based on a simple subproblem"*, Iranain Journal of Numerical Analysis and Optimization, 2015.

[15] Saeidian, Z.; "*A Subsampled cubic regularization trust region method"*, The 11 th International conference of Iranian Operations Research, Kermanshah, 2018.

[16] Sang, Z., Sun, Q. ; "*A self-adaptive trust region method with line search based on a simple subproblem model"*, J. Appl. Math. Comput., No.232, pp. 514-522, 2009.

[17] Sartenaer, A.; "*Automatic determination of an initial trust region in nonlinear programming*", SIAM J. Sci. Comput., No.18(6),pp.1788-1803, 1997.

[18] Shi, Z., Guo, J.; "*A new trust region method for unconstrained optimization"*, J. Comput. Appl. Math., No. 213, pp. 509-520, 2008.

[19]

Shi, Z.J., Wang, H.Q.; "*A new self-adaptive trust region method for unconstrained optimization*", Technical Report, College of Operations Research and Management, Qufu Normal University, 2004.

[20] Zheng, Y., Zheng, B.; "*A modified adaptive cubic regularization method for large-scale unconstrained optimization problem*", https://arxiv.org/pdf/1904.07440.pdf.

[21] Zhou, Q., Zhang, C.; "*A new nonmonotone adaptive trust region method based on simple quadratic models*", J. Appl. Math. Comput., No.40, pp.111-123, 2012.