International Journal of Scientific Engineering and Applied Science (IJSEAS) – Volume-6, Issue-10, October 2020
ISSN: 2395-3470
www.ijseas.com

An Improved Trust-Region Method for Nonlinear Equations

Fenghua Liu[1] , Lujing Huo[2] , Meihong Zhou[3] , Baowei Liu[1] & Qinghua Zhou[2,*]

[1]College of Computer Science and Engineering, Cangzhou Normal University, Hebei, China

[2]School of Applied Mathematics, Beijing Normal University, Zhuhai, China

[3]Xinghua primary school, Shijiazhuang, Hebei, China

zhouqinghua@bnuz.edu.cn (*Correspondence author)

Abstract In this paper, we propose an improved trust region method for solving unconstrained optimization problems. Our algorithm resolve the subproblem within the trust region centered at some point located in the direction of the negative gradient, while the current iteration point is on the boundary set. Moreover, a nonmonotone technique is used to improve the algorithm efficiency. Theoretical analysis indicates that the new method preserves the global convergence under mild conditions.

Keywords: nonlinear equations; trust region method; nonmonotone strategy; negative gradient direction; global convergence

1. Introduction

Consider the following nonlinear system of equations:

$$F(x)=0,\ x\in R^n. \tag{1}$$

where $F: R^n \to R^n$ is a continuously differentiable mapping in the form $F(x)=\left(F_1(x),F_2(x),\cdots,F_n(x)\right)^T$. Suppose that $F(x)$ has a zero, then the nonlinear system (1) is equivalent to the following nonlinear unconstrained least-squares problem

$$\begin{aligned} &\min\ f(x)=\frac{1}{2}\|F(x)\|^2 \\ &s.t.\ \ x\in R^n. \end{aligned} \tag{2}$$

where $\|.\|$ denotes the Euclidean norm.

Nonlinear equations not only have great importance in theory but also have wide applications in reality. Many problems arising from chemical technology, economy, and communications and so on, can be formulated as nonlinear equations. There are various methods for solving system of nonlinear equations, such as the Newton and the quasi-Newton methods [1-4], the spectral method [5,6], the Levenberg-Marquardt method [11,12], trust-region-based methods [7-10]. The trust-region method is a very popular way for nonlinear equations and have a number of attractive features.

Traditionally, the trust region methods are iterative algorithms. At each iterative point x_k, the traditional TR methods obtain the trial step d_k using the following subproblem model:

$$\min q_k(d) = \frac{1}{2}\|F_k + J_k d\|^2 \quad s.t.\ d \in R^n \text{ and } \|d\| \le \Delta_k. \tag{3}$$

where $f_k = f(x_k), F_k = F(x_k), J_k = \nabla F(x_k)$, and $\Delta_k > 0$ is trust region radius. Typically, the trust region is a ball centered at the current iteration point x_k. But by a simple analysis, we find that not all of the function values at the points within the traditional trust region area are smaller than the function values at the current point. Actually, the objective function increases in the area where the angle between the search direction and the negative gradient direction is an obtuse one. It is distinct that only when the angle between the search direction and the negative gradient one is an acute one, then the objective function will descend. It means that about half of the trust region does not work.

To overcome this disadvantage of the traditional trust regions, Zhou [13] raised a new center of the trust region and improved the trust region method. Similar to [13], we give a new trust region subproblem as follows:

$$\min_{d \in R^n} q_k(d) = \frac{1}{2}\|F_k + J_k d\|^2 = \frac{1}{2}\|F_k\|^2 + d^T F_k J_k + \frac{1}{2} d^T J_k^T J_k d \quad s.t.\ \left\| d + \frac{\Delta_k J_k}{\|J_k\|} \right\| \le \Delta_k. \tag{4}$$

where d is the trial step, $x_k - \frac{\Delta_k J_k}{\|J_k\|}$ is the center, Δ_k is the radius of the improved trust region.

This approach is to remove half of the ball whose intersection angle with the direction of the negative gradient is an obtuse one. It chooses some point located in the direction of the negative gradient as the center of the improved trust region and chooses the distance between the center and the current point as the radius of the improved trust region.

It is well known that the nonmonotone technique is one of the most interesting techniques for improving the iterative algorithms in optimization. The classical nonmonotone line search technique by Grippo et al. [14] replaces the usual (monotone) Armijo rule by the test

$$f(x_k+\alpha_k d_k) \le f(x_{l(k)}) + \beta\alpha_k \nabla f(x_k)^T d. \tag{5}$$

where

$$f_{l(k)} = \max_{0\le j\le m(k)} \left\{ f_{k-j} \right\}, \quad k=0,1,2,\ldots,\ \beta\in(0,1/2),$$
$$m(0)=0,\ \ 0\le m(k)\le \min\left\{ m(k-1)+1, N \right\}, N\ge 0.$$

In 1993, Deng et al. in [15] 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 [16] 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:

$$f(x_k+\alpha_k d_k) \le C_k + \beta\alpha_k \nabla f(x_k)^T d. \tag{6}$$

where

$$C_k = \begin{cases} f(x_k), & k=0. \\ \dfrac{\eta_{k-1}Q_{k-1}C_{k-1}+f(x_k)}{Q_k}, & \mathrm{k}\ge 1 \end{cases}. \qquad Q_k = \begin{cases} 1, & k=0. \\ \eta_{k-1}Q_{k-1}+1, & k\ge 1. \end{cases} \tag{7}$$

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. This strategy has been applied to a nonmonotone trust-region framework by Ahookhosh and Amini [17], where the ratio is

$$r_k = \frac{ared_k}{pred_k} = \frac{f(x_k+d_k)-C_k}{q_k(d_k)-q_k(0)}. \tag{8}$$

If $r_k \ge \mu_1$, the trial step d_k is accepted and it is called as a successful iteration. 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 trial point is rejected.

The rest of this paper is organized as follows. In Section 2, the new algorithm will be introduced. The convergence analysis is investigated in Section 3. Finally, some conclusions are addressed in Section 4.

2. Algorithm

Inspired by the ideas introduced above, we use an improved trust region method, which the trust region centered at some point located in the direction of the negative gradient, and we use the non-monotone strategy proposed by Zhang and Hager [15].

Now, we outline our algorithm as follows:

Algorithm 2.1

Initial: Choose a starting point $x_0 \in R^n$, an initial trust region radius Δ_0, and constants $\varepsilon > 0, \mu_1 \geq 0, 0 < \mu_2 < 1, 0 < \beta_2 < \beta_3 < 1 < \beta_1$. Set $k = 0$.

Step 1: If $\|F_k\| < \varepsilon$ holds, stop; otherwise, go to step 2.

Step2: Solve trust region subproblem (4) and obtain d_k.

Step3: Compute C_k by (7) and r_k by (8). If $r_k \geq \mu_1$, let $x_{k+1} = x_k + d_k$;Otherwise, let $x_{k+1} = x_k$. Set

$$\Delta_{k+1} \in \begin{cases} \left(\beta_2 \dfrac{\|d_k\|\|J_k\| + \Delta_k J_k\|}{\|J_k\|}, \beta_3 \Delta_k \right), & \textit{if } r_k < \mu_2. \\ \left(\Delta_k, \beta_1 \Delta_k \right), & \textit{otherwise.} \end{cases} \tag{9}$$

Step 4: increment k by 1 and go to Step 1.

3. Convergence analysis

This section gives some convergence results under the following assumptions.

(H1) Let the level set $\Omega = \left\{ x \middle| f(x) \leq f(x_0) \right\}$ be bounded.

(H2) $F(x)$ is continuously differentiable on an open convex set Ω_1 containing Ω, $\left\{ \|F_k\| \right\}$ is bounded.

(H3) The Jacobian of $F(x)$ is symmetric, bounded and positive definite on Ω_1, i.e., there exist positive constants $M \geq m > 0$ such that

$$\|\nabla F(x)\| \leq M, \quad \forall x \in \Omega. \tag{10}$$

and

$$m\|d\|^2 \leq d^T \nabla F(x) d, \quad \forall x \in \Omega, d \in R^n. \tag{11}$$

Lemma 3.1 If d_k is the solution of (4), then

$$-pred_k(d_k) \geq \frac{1}{2}\|J_k F_k\| \min\left\{\Delta_k, \frac{\|J_k F_k\|}{\|J_k^T J_k\|}\right\} \tag{12}$$

holds.

Proof. The proof is similar to Lemma 3.1 in [18].

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

$$f_{k+1} \leq C_{k+1} \leq C_k \tag{13}$$

Proof. Let $k \geq 0$ be an arbitrary fixed integer. By the definition of r_k and $r_k > \mu_1$, we have

$$\begin{aligned} C_k - f_{k+1} &\geq \mu_1\left(q_k(0) - q_k(d_k)\right) = \mu_1\left(-pred_k(d_k)\right) \\ &\geq \frac{\mu_1}{2}\|J_k F_k\| \min\left\{\Delta_k, \frac{\|J_k F_k\|}{\|J_k^T J_k\|}\right\} \geq 0. \end{aligned} \tag{14}$$

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

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

So

$$C_k \geq f_k. \tag{15}$$

On the other hand, we have

$$C_{k+1} = \frac{\eta_k Q_k C_k + f_{k+1}}{Q_{k+1}} \leq \frac{\eta_k Q_k C_k + C_k}{Q_{k+1}} = C_k. \tag{16}$$

From (15) and (16), Lemma 3.2 holds.

Lemma 3.3 Suppose that (H1)-(H3) hold and the sequence $\{x_k\}$ be generated by Algorithm

2.1. Then the sequence $\{C_k\}$ is convergent.

Proof. The proof is similar to Corollary 3.1 in [17].

Lemma 3.4 Suppose that (H1)-(H3) hold and the sequence $\{x_k\}$ be generated by Algorithm 2.1. Then we have

$$\lim_{k\to\infty} C_k = \lim_{k\to\infty} f_k. \tag{17}$$

Proof. The proof is similar to Lemma 3.1 in [17].

(H4) Let x^* stands for the unique solution of Equation (1) in Ω_1, ∇F is Holder continuous at x^*, i.e. there are constants M_1 and γ such that for every x in a neighborhood of x^*,

$$\left\|\nabla F(x) - \nabla F(x^*)\right\| \le M_1 \left\|x - x^*\right\|^{\gamma}. \tag{18}$$

Theorem 3.1 Let (H1)-(H4) hold, Algorithm 2.1 either terminates in finite steps or generates a sequence of iterations $\{x_k\}$, such that

$$\lim_{k\to\infty} \inf \|F_k\| = 0. \tag{19}$$

Proof. We assume that Algorithm 2.1 does not terminate in finite steps and $\lim_{k\to\infty} \inf \|F_k\| \neq 0$, that is, there exists a positive constant $\delta > 0$, such that

$$\|F_k\| \ge \delta . \tag{20}$$

for all k.

We define I to be the set of integers k, which satisfies $r_k \ge \mu_2$.

We make the connections (10), (11), (12), (15), (17), (20) and $f(x)$ being bounded below, and we have

$$\begin{aligned} +\infty &> \sum_{k=1}^{+\infty}[C_k - f_{k+1}] \ge \sum_{k\in I}[f_k - f_{k+1}] \ge \sum_{k\in I}\mu_2(-pred_k) \\ &\ge \sum_{k\in I}\frac{\mu_2}{2}\|J_k F_k\| \min\left\{\Delta_k, \frac{\|J_k F_k\|}{\|J_k^T J_k\|}\right\} \\ &\ge \sum_{k\in I}\frac{\mu_2}{2}\|J_k F_k\| \min\left\{\Delta_k, \frac{m\delta}{M^2}\right\} \end{aligned}$$

The formula above indicates that

$$\sum_{k\in I}\min\left\{\Delta_k,\frac{m\delta}{M^2}\right\}<+\infty. \tag{21}$$

If the set I is finite, then from (9) we know that $\Delta_{k+1}\le\beta_3\Delta_k$ for all sufficiently large k. Because of $\beta_3<1$, the sequence $\{\Delta_k\}$ converges to zero. If the set I is not finite, then from Formula (21) we have

$$\lim_{k\in I}\Delta_k=0. \tag{22}$$

We also know that

$$\begin{aligned}\|d_k\|&=\left\|d_k+\frac{\Delta_k J_k}{\|J_k\|}-\frac{\Delta_k J_k}{\|J_k\|}\right\|\\&\le\left\|d_k+\frac{\Delta_k J_k}{\|J_k\|}\right\|+\left\|\frac{\Delta_k J_k}{\|J_k\|}\right\|\\&\le\Delta_k+\Delta_k\\&=2\Delta_k.\end{aligned} \tag{23}$$

From Lemma 3.2 in [9], there exist constants $M_3\ge m_3>0$ such that

$$m_3\|F_k\|\le\|d_k\|\le M_3\|F_k\|. \tag{24}$$

We apply (8), (10), (15), (18), (24) that

$$\begin{aligned}|ared_k-pred_k|&=|f(x_k+d_k)-C_k-pred_k|\\&\le|f(x_k+d_k)-f_k-pred_k|\\&=\left|\frac{1}{2}\|F_k+\nabla F_\xi d_k\|^2-\frac{1}{2}\|F_k\|^2-pred_k\right|\\&=\Big|d_k^T(\nabla F_\xi-J_k)F_k+\frac{1}{2}d_k^T\nabla F_\xi^T(\nabla F_\xi-J_k)d_k\\&\quad+\frac{1}{2}d_k^T(\nabla F_\xi^T-J_k^T)J_kd_k\Big|\\&\le\|d_k\|\|F_k\|\|\nabla F_\xi-J_k\|+\frac{1}{2}M\|d_k\|^2\|\nabla F_\xi-J_k\|\\&\quad+\frac{1}{2}M\|d_k\|^2\|\nabla F_\xi-J_k\|\\&=O\left(\|d_k\|^{2+\gamma}\right).\end{aligned} \tag{25}$$

where $\nabla F_\xi=\nabla F(x_k+\xi_1 d_k),\ \xi_1\in(0,1)$.

Because of Formula (22), the formula

$$\min\left\{\Delta_k, \frac{m\delta}{M^2}\right\} = \Delta_k. \tag{26}$$

By the definition of $pred_k$, Lemma 3.1, (20), (23), (24) and (26), we obtain

$$\begin{aligned} O\left(\|d_k\|^2\right) &\le \frac{1}{4} m \|F_k\| \|d_k\| \le \frac{1}{2}\|J_k F_k\| \Delta_k \\ &\le \frac{1}{2}\|J_k F_k\| \min\left\{\Delta_k, \frac{m\delta}{M^2}\right\} \\ &\le \frac{1}{2}\|J_k F_k\| \min\left\{\Delta_k, \frac{\|J_k F_k\|}{\|J_k^T J_k\|}\right\} \\ &\le \left|-pred_k\left(d_k\right)\right| \\ &\le \|d_k\|\|J_k\|\|F_k\| + \|J_k d_k\|^2 \\ &= O\left(\|d_k\|^2\right). \end{aligned} \tag{27}$$

From (25) and (27), it is not difficult to get

$$|r_k - 1| = \frac{|ared_k - pred_k|}{|pred_k|} \to 0. \tag{28}$$

for sufficient large k.

From Formula (9), we know that there exists a positive constant Δ^* such that $\Delta_k > \Delta^*$ for all large k. This contradicts (22).

4. Conclusion

In this paper, we propose an improved trust region method for solving unconstrained optimization problems. Different with traditional trust region methods, our algorithm does not resolve the subproblem within the trust region centered at the current iteration point, but within an improved one centered at some point located in the direction of the negative gradient, while the current iteration point is on the boundary set. Moreover, to improve the algorithm efficiency, we use a nonmonotone technique. Under mild conditions, we obtain the global convergence.

For further research, we should study this trust-region method where the Jacobian matrix is replaced by other update matrix (such as quasi-Newton). Moreover, numerical experiments for large practical problems should be done in the future.

Acknowledgements

This work is supported by the Natural Science Foundation of Guangdong province (Grant No. 2018A0303130026) And the Teacher Research Capacity Promotion Program of Beijing Normal University, Zhuhai.

References

[1] RH Byrd. Global convergence of a class of quasi-Newton methods on convex problems. SIAM Journal on Numerical Analysis. 1987, 24(5): 1171–1189.

[2] M Albaali, E Spedicato, F Maggioni. Broyden's quasi-Newton methods for a nonlinear system of equations and unconstrained optimization: a review and open problems. Optimization Methods and Software. 2014, 29(5): 937-954.

[3] JE Dennis. Quasi-Newton methods, motivation and theory. Siam Review. 1977, 19(1): 46-89.

[4] GZ Gu, DH Li, L Qi, SZ Zhou. Descent directions of quasi-Newton methods for symmetric nonlinear equations. Siam Journal on Numerical Analysis. 2002, 40(5): 1763–1774.

[5] W Cheng, DH Li. A derivative-free nonmonotone line search and its application to the spectral residual method. Ima Journal of Numerical Analysis. 2009, 29(3): 814-825(12).

[6] WL Cruz, M Raydan. Nonmonotone spectral methods for large-scale nonlinear systems. Optimization Methods and Software. 2003, 18(5): 583- 599.

[7] J Fan, J Pan. An improved trust region algorithm for nonlinear equations. Computational. Optimization and Applications. 2011, 48(1): 59–70.

[8] J Fan, J Pan. A modified trust region algorithm for nonlinear equations with new updating rule of trust region radius. International Journal of Computer Mathematics. 2010, 87(14): 3186–3195.

[9] G Yuan, S Lu, Z Wei. A new trust-region method with line search for solving symmetric nonlinear equations. Int. J. Comput. Math. 2011, 88(10): 2109–2123.

[10] G Yuan, Z Wei, X Lu. A BFGS trust-region method for nonlinear equations. Computing. 2011, 92(4): 317–333.

[11] JY Fan. The modified Levenberg–Marquardt method for nonlinear equations with cubic convergence. Math. Comput. 2012, 81: 447–466.

[12] JY Fan, YX Yuan. On the quadratic convergence of the Levenberg–Marquardt method without nonsingularity assumption. Computing. 2005, 74(1): 23–39.

[13] QH Zhou, YR Zhang. An improved trust region method for unconstrained optimization. Science China Mathematics. 2013, 56(2): 425-434.

[14] L Grippo, F Lamparillo, S Lucidi. A nonmonotone line search technique for Newton's method. Siam Journal on Numerical Analysis. 1986, 23(4): 707-716.

[15] N Deng, Y Xiao, F Zhou. Nonmontonic trust region algorithm. Journal of Optimization Theory and Application. 1993, 76(2): 259-285.

[16] H Zhang, WW Hager. A nonmonotone line search technique and its application to unconstrained optimization. Siam Journal on Optimization. 2004, 14(4): 1043-1056.

[17] K Amini, M Ahookhosh. A hybrid of adjustable trust-region and nonmonotone algorithms for unconstrained optimization. Applied Mathematics Modelling. 2014, 38(9-10): 2601-2612.

[18] G Yuan, S Lu, Z Wei. A new trust-region method with line search for solving symmetric nonlinear equations. International of Computer Mathematics. 2011, 88(10): 2109– 2123.