International Journal of Recent Engineering Science (IJRES),
ISSN: 2349-7157, *Volume 2 Issue 6 November to December 2015*

A Non-monotone Trust-Region-Based Approach for Symmetric Nonlinear Systems

Baowei Liu
College of Mathematics and Statistics, Cangzhou Normal University, China

Abstract:

This paper puts forward a new trust-region process for solving symmetric systems of equations having several variables. The proposed approach the efficiency and robustness of the trust-region framework. The global convergence and the quadratic convergence of the proposed approach are established.

Keywords- *adaptive radius, non-monotone technique, nonlinear equations, trust-region method*

I. INTRODUCTION

Let F : $R^n \to R^n$ be a continuously differentiable image in the following form $F(x) = (F_1(x), F_2(x), \ldots, F_n(x))^T$, and then consider the following symmetric nonlinear systems of equations:

$$F(x) = 0, \quad x \in R^n \qquad (1)$$

This class of problems is incessantly close to both constrained and unconstrained optimization problems. It is also worth noting that a feasible approach to Equation (1) consists in reformulating it as a nonlinear unconstrained least-squares problem:

$$\min_{x \in R^n} f(x) = \frac{1}{2}\|F(x)\|^2 \qquad (2)$$

where $\|\cdot\|$ denote the Euclidean norm. Nonlinear least-squares problems have been comprehensively studied by more and more authors such that various iterative procedures have been proposed to solve this problem, for example [1-4].

The trust region framework for solving systems of nonlinear equations (1) is such a popular of iterative procedure. In each iterate, it can generate

makes use of the combination of both an efficient adaptive trust-region radius and a non-monotone method, and this strategy can improve

a trial step d_k by computing an approximate solution of the following sub-problem:

$$\begin{aligned} &\min m_k(x_k + d_k) \\ &= \frac{1}{2}\|F_k + J_k d_k\|^2 \\ &= f_k + d_k^T J_k^T F_k + \frac{1}{2} d_k^T J_k^T J_k d_k, \\ &s.t.\, d_k \in R^n \ \ and \ \ \|d_k\| \le \Delta_k, \end{aligned} \qquad (3)$$

where $x_{k+1} = x_k + d_k$, $f_k = f(x_k)$, $F = F(x_k)$, $J_k = F^{'}(x_k)$ is the Jacobian of $F(x)$, and $\Delta_k > 0$ is the trust-region radius.

It is clear that the calculation of the exact Jacobian J_k at each iterate is a common drawback of sub-problem (3) that straightway increases the entire computational cost of solving a problem. BFGS and Broyden family updated formulas instead of the exact Jacobian matrix J_k in Equation (3), the sub-problem can be rewritten as

$$\begin{aligned} &\min \hat{m}_k(x_k + d_k) \\ &= f_k + d_k^T B_k F_k + \frac{1}{2} d_k^T B_k B_k d_k \\ &s.t.\, d_k \in R^n \ \ and \ \ \|d_k\| \le \Delta_k \end{aligned} \qquad (4)$$

The B_k is satisfied with the following in-equation [5]

$$\left\|B_k^{-1}\right\| \geq \frac{\|d_{k-1}\|}{\|F_k - F_{k-1}\|} \qquad (5)$$

and

$$\Delta_k = c^{p_k} \min\left\{\frac{\|d_{k-1}\|}{\|F_k - F_{k-1}\|}, M\right\}\|F_k\| \qquad (6)$$

where $c \in (0,1)$, M is a constant and p_k is the smallest non-negative integer p ensuring that the trust-region ratio is greater than a real-valued parameter $\mu \in (0,1)$.

Recently, non-monotone techniques are generally used in the trust region methods. In 1982, the first non-monotone technique that is the so-called watchdog technique was proposed by Chamberlain et al. [6] 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 [7]. In 1993, Deng et al. [8] 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 solving optimization problems [9]. Especially, nowadays some researchers are focused on utilizing non-monotone techniques in adaptive trust region method and good numerical results have been achieved [10].

The general non-monotone form is as follows:

$$f(x_{l(k)}) = \max_{0 \leq j \leq m_k} f(x_{k-j})$$

$m(0) = 0,$

$0 \leq m(k) \leq \min\{M, m(k-1)+1\}$ and

$M \geq 0$ is an integer constant.

Actually, the most common non-monotone ratio is defined as follows:

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

However, although the non-monotone technique has many advantages, Zhang et al. [11] found that it still has some drawbacks and they proposed a new non-monotone

$$\rho_k = \frac{C_k - f(x_k + s_k)}{m_k(x_k) - m_k(x_k + s_k)} \qquad (7)$$

$$C_k = \begin{cases} f(x_k), & k = 0 \\ \frac{\eta_{k-1}Q_{k-1}C_{k-1} + f(x_k)}{Q_k}, & k \geq 1 \end{cases} \qquad (8)$$

where $Q_k = \begin{cases} 1, & k = 0 \\ \eta_{k-1}Q_{k-1} + 1, & k \geq 1 \end{cases}$

$\eta_{\min} \in [0,1)$, $\eta_{\max} \in [\eta_{\min}, 1]$ and

$\eta_{k-1} \in [\eta_{\min}, \eta_{\max}]$, are two given constants.

The rest of this paper is organized as follows. In Section 2, we introduce the new non-monotone trust-region-based approach. In Section 3, we analyze the new method and prove the global convergence. Some conclusions are given in Section 4.

II. ALGORITHM 1: A FRESH NON-MONOTONE TRUST-REGION-BASED ALGORITHM

Step 0**:** An initial point $x_0 \in R^n$, a symmetric positive definite matrix $B_0 \in R^{n \times n}$, $k_{\max}$, c, $\mu \in (0,1)$, $\mathrm{M} > 0$, $\mathrm{N} > 0$ and $\varepsilon > 0$. the original value $\|F_0\| = \Delta_0$, $\mathrm{R}_0 = \frac{1}{2}\|F_0\|^2$, $p = 0$, $\hat{r}_0 = 0$, $k = 0$ then compute $B_0 F_0$ and $\|F_k\| \geq \varepsilon$, $k \leq k_{max}, \hat{r}_k < \mu$. (Start of outer loop).

Step 1: Specify the trial point d_k by solving sub-problem (5); $\hat{x}_{k+1} = x_k + d_k$.

Step 2: $F(\hat{x}_{k+1}) = \hat{F}_{k+1}, \frac{1}{2}\left\|\hat{F}_{k+1}\right\|^2 = \hat{f}_{k+1}$; and calculate $Ared = C_k - \hat{f}_{k+1}$,

$\mathrm{P}red = \hat{m}(x_k) - \hat{m}(\hat{x}_{k+1})$, and determine the trust-region ratio $\hat{r}_k$ by using $\hat{r}_k = \dfrac{Ared}{\mathrm{P}red}$.

If $\hat{r}_k < \mu$, $p = p+1$ and update the trust-region radius Δ_k with (6); else $x_{k+1} = \hat{x}_{k+1}$. (End of inner loop)

Step 3: $p \leftarrow 0$; and determine Δ_k using (6).

Step 4: Update B_{k+1} by (5), and define the $m(k+1) = min\{m(k)+1, \mathrm{N}\}$; while calculate the C_{k+1} and generate η_{k+1} by (8).

Step 5: Set $F_{k+1} = \hat{F}_{k+1}$, $k = k+1$, go to step 1.

In order to testifying the convergence analysis of the proposed algorithm, the following assumptions are required:

(A1) Let the level set $L(x_0) = \{x \in R^n \mid f(x) \le f(x_0)\}$ be bounded for any given $x_0 \in R^n$ and $F(x)$ be continuously differentiable on the compact convex set Ω containing the level set $L(x_0)$.

(A2) Assume that the following condition holds:

$$\left\|(\mathrm{J}_k - \mathrm{B}_k)^T F_k\right\| = \mathrm{O}(\|d_k\|) \qquad (9)$$

(A3) The matrix $\{B_k\}$ is uniformly bounded above; that is to say, there exists a constant $M_1 > 0$ such that

$$\|B_k\| \le M_1 \qquad (10)$$

for all $k \in N \cup \{0\}$.

Moreover, we assume that $M_0 \|F_k\| \le \|B_k F_k\|$, for a constant $M_0 > 0, k \in N \cup \{0\}$.

(A4) The abatement of the model $\hat{m}_k$ is at least as much as a fraction of that obtained by the Cauchy point; in other words, there exists a constant $\beta \in (0,1)$ such that

$$\begin{aligned} &\hat{m}_k(x_k) - \hat{m}_k(x_k + d_k^p) \\ &\ge \beta \|B_k F_k\| min\left[\Delta_k, \frac{\|B_k F_k\|}{\|\hat{B}_k\|}\right] \end{aligned} \qquad (11)$$

for all $k \in N \cup \{0\}$.

From (A3), it is distinct that the matrix $B_k B_k$ is uniformly bounded above as well. Meanwhile, condition (11) can be easily achieved by approximately solving the trust-region sub-problem (4), on account of some effective procedures [12, 13]. The detailed results can be generalized in the subsequence lemma.

III. GLOBAL CONVERGENCE

Lemma 3.1. [14] Suppose that d_k is a solution of sub-problem (4) such that

$$\begin{aligned} &\hat{m}_k(x_k) - \hat{m}_k(x_k + d_k) \\ &\ge \hat{m}_k(x_k) - \hat{m}_k(x_k + d_k^{CP}) \end{aligned} \qquad (12)$$

where d_k^{CP} is the Cauchy point. Then, condition (11) holds.

Lemma 3.2. [15] Suppose that (A2) and (A3) hold, the sequence $\{x_k\}$ is generated by Algorithm 1 and d_k is a solution of sub-problem (4). Then, we have

$$|f(x_k + d_k) - \hat{m}_k(x_k + d_k)| \le \mathrm{O}(\|d_k\|^2) \qquad (13)$$

Lemma 3.3. Suppose that the sequence $\{x_k\}$ is generated by Algorithm 1. Then the inequality

$$f(x_{k+1}) \le C_{k+1} \le C_k$$

$\forall k$ holds, where C_k is defined in (8).

Proof: We prove $C_k \le f(x_0)$ by induction. Obviously, when $j=0$, we have $C_0 = f(x_0)$. Assume that $C_j \le f(x_0)$ holds when $j=1,\ldots,k-1$. When $j=k$, from (8), that we have

$$\begin{aligned} C_k &= \frac{\eta_{k-1}Q_{ik-1}C_{k-1} + f(x_k)}{Q_k} \\ &\le \frac{\eta_{k-1}Q_{k-1}f(x_0) + f(x_0)}{Q_k} \\ &= f(x_0) \end{aligned} \tag{14}$$

then we obtain that $f(x_{k+1}) \le C_{k+1} \le C_k \le f(x_0)$.

We complete the proof.

Lemma 3.4. Suppose that (A2), (A3) and (A4) hold and the sequence $\{x_k\}$ is generated by Algorithm 1. Then, the inner cycle of Algorithm 1 is well defined.

Proof: Assume that the inner cycle of Algorithm 1 cycles infinity; that is, $\Delta_k^p \to 0$ as $p \to \infty$. Using the fact that x_k is not the optimum of Equation (2), it can be concluded that there exists a constant $\varepsilon > 0$ such that $\|F_k\| \ge \varepsilon_0$. These facts, (A3) and (11) suggest that

$$\begin{aligned} &\hat{m}_k(x_k) - \hat{m}_k(x_k + d_k^p) \\ &\ge \beta\|B_kF_k\| min\left[\Delta_k^p, \frac{\|B_kF_k\|}{\|\hat{B}_k\|}\right] \\ &\ge \beta M_0\varepsilon min\left[\Delta_k^p, \frac{M_0\varepsilon}{M_1^2}\right] \\ &\ge \beta M_0\varepsilon\Delta_k^p, \end{aligned} \tag{15}$$

Where d_k^p is a solution of sub-problem (4) corresponding to p in the k-th iterate. Now the Lemma 3.2, In-equation (15) and lead to

$$\begin{aligned} &\left|\frac{f_k - f(x_k + d_k^p)}{\hat{m}_k(x_k) - \hat{m}_k(x_k + d_k^p)} - 1\right| \\ &\le \left|\frac{f(x_k + d_k^p) - \hat{m}_k(x_k + d_k^p)}{\hat{m}_k(x_k) - \hat{m}_k(x_k + d_k^p)}\right| \\ &\le \frac{O(\|d_k^p\|^2)}{\beta M_0\varepsilon\Delta_k^p} \le \frac{O((\Delta_k^p)^2)}{\beta M_0\varepsilon\Delta_k^p} \to 0, \end{aligned}$$

as $p \to \infty$. Therefore, there exists a sufficiently large p, which is called p_k, such that

$$\frac{f_k - f(x_k + d_k^{p_k})}{\hat{m}_k(x_k) - \hat{m}_k(x_k + d_k^{p_k})} \ge \mu.$$

Besides, from the definition of C_k, and from Lemma 3.3, we have that $C_k \ge f_k$. Therefore, this fact along with the previous inequality immediately implies that $\hat{r} \ge \mu$, which means that p_k is a finite integer number, so the inner cycle of Algorithm 1 is well defined.

Lemma 3.5. [15] Suppose that (A3) and (A4) hold, the sequence $\{x_k\}$ is generated by Algorithm 1 and d_k is a solution of sub-problem (4). Then, we have

$$\hat{m}_k(x_k) - \hat{m}_k(x_k + d_k) \ge L_k\|F_k\|^2 \tag{16}$$

where $L_k = \beta min\left\{c^{p_k} min\left\{\frac{1}{M_1}, M\right\}, \frac{M_0}{M_1^2}\right\}$.

Theorem 3.6. Suppose that (A1)-(A4) hold. Then, Algorithm 1 either stops at a stationary point of $f(x)$ or generates an infinite sequence $\{x_k\}$ such that

$$\lim_{k\to\infty}\|F_k\| = 0 \tag{17}$$

Proof: By contradiction, for all sufficiently large k, assume that there exist a constant $\varepsilon > 0$ and an infinite subset $K \subseteq N \cup \{0\}$ satisfying

$$\|F_k\| > \varepsilon \qquad (18)$$

for all $k \in K$.

Using in-equation (14) and $\hat{r}_k > \mu$, it can be written that

$$\begin{aligned} & C_k - f(x_k + d_k) \\ & \geq \mu\left[\hat{m}_k(x_k) - \hat{m}_k(x_k + d_k)\right] \\ & \geq \mathrm{L}_k \|F_k\|^2 \\ & \geq \mu\varepsilon^2 L_k. \end{aligned}$$

Now, by taking a limit from both sides of this inequality, as $k \to \infty$, we have that $lim_{k\to\infty} L_k = 0$; that is $p_k \to \infty$ for sufficiently large $k \in K$. But this is possible only if $p_k \to \infty$, as $k \to \infty$ and $k \in K$. This clearly results in a contradiction with Lemma3.3. Therefore, hypothesis (18) is not true, and the result of the theorem is given.

IV. CONCLUSIONS

This paper is involved with the introduction and analysis of a trust-region-based algorithm for solving systems of equation by developing an effective adaptive trust-region radius and a credible non-monotone strategy. Practical utilization of the trust-region framework has indicated that the application of adaptive techniques can decline the number of sub-problem's count, and the employment of non-monotone strategy increases the efficiency and robustness of the algorithm. As a result, the two methods are combined in a trust-region framework to construct a more promising algorithm for solving symmetric non-linear systems.

ACKNOWLEDGMENTS

The author would be grateful to the anonymous referees for their valuable comments and suggestions.

REFERENCES

[1] C. G. Broyden, The convergence of an algorithm for solving sparse nonlinear systems, Math. Comp., 25(114), 1971, 285-294.

[2] H. Dan, N. Yamashita, and M. Fukushima, Convergerce properties of the inexact Levenberg-Marquardt method under local error bound condition, Optimization Methods Software, 17, 2002, 605-626.

[3] J. E. Dennis, On the convergence of Broyden's method for nonlinear systems of equations, Mathematics of Computatiion., 25(115), 1971, 559-567.

[4] J. Y. Fan, Convergence rate of the trust region method for nonlinear equations under local error bound condition, Computer Optimization Application, 34, 2005, 215-227.

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

[6] R. M. Chamberlain, M. J. D. Powell, The watchdog technique for forcing convergence in algorithm for constrained optimization, Mathematical Programming Study, 16, 1982, 1-17.

[7] L. Grippo, F. Lamparillo, S. Lucidi, A nonmonotone line search technique for Newton's method, SIAM J. Numer. Anal., 23(4), 1986, 707-716.

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

[9] Ph. L. Toint, Non-monotone trust-region algorithm for nonlinear optimization subject to convex constraints, Mathematical Programming 77, 1997, 69-94.

[10] Y. Yang, W. Y. Sun, A new nonmonotone self-adaptive trust region algorithm with line search, Chinese Journal of Engineering Mathematics, 24(5), 2007, 788-794.

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

[12] A. R. Conn, N. M. Gould, and Ph. L. Toint, Trust-region Methods (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2000).

[13] J. Nocedal and S. J. Wright, Numerical Optimization (Springer, New York, 2006).

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

[15] Masoud Ahookhosh, Hamid Esmaeili, Morteza Kimiaei. An effective trust-region-based approach for symmetric nonlinear systems, International Journal of Computer Mathematics, 90(3), 2013, 671-690.