A nonmonotone adaptive trust-region algorithm for symmetric nonlinear equations

Gong-Lin Yuan¹, Cui-Ling Chen², Zeng-Xin Wei¹

¹College of Mathematics and Information Science, Guangxi University, Nanning, China; glyuan@gxu.edu.cn
²College of Mathematics Science, Guangxi Normal University, Guilin, China

Received 23 December 2009; revised 1 February 2010; accepted 3 February 2010.

ABSTRACT

In this paper, we propose a nonmonotone adaptive trust-region method for solving symmetric nonlinear equations problems. The convergent result of the presented method will be established under favorable conditions. Numerical results are reported.

Keywords: Trust Region Method; Global Convergence; Symmetric Nonlinear Equations

1. INTRODUCTION

Consider the following system of nonlinear equations:

\[ g(x) = 0, x \in \mathbb{R}^n \]  \hspace{1cm} (1)

where \( g : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is continuously differentiable, the Jacobian \( \nabla g(x) \) of \( g \) is symmetric for all \( x \in \mathbb{R}^n \).

Define a norm function by \( \phi(x) = \frac{1}{2} \| g(x) \|^2 \). It is not difficult to see that the nonlinear equations problem Eq.1 is equivalent to the following global optimization problem

\[
\min \phi(x), \ x \in \mathbb{R}^n \hspace{1cm} (2)
\]

Here and throughout this paper, we use the following notations.

- \( \| \cdot \| \) denote the Euclidian norm of vectors or its induced matrix norm.
- \( \{ x_k \} \) is a sequence of points generated by an algorithm, and \( g(x_k) \) and \( \phi(x_k) \) are replaced by \( g_k \) and \( \phi_k \) respectively.
- \( B_k \) is a symmetric matrix which is an approxim-a-
Eq.1. More precisely, we solve Eq.1 by the method of iteration and the main step at each iteration of the following method is to find the trial step $d_k$. Let $x_k$ be the current iteration. The trial step $d_k$ is a solution of the following trust region subproblem

$$\min q_k(d) = \nabla\phi(x_k)^T d + \frac{1}{2} d^TB_kd$$

s. t. $\|d\| \leq \Delta_k$, $d \in \mathbb{R}^n \tag{4}$

where $\nabla\phi(x_k) = \nabla g(x_k)g(x_k)$, $\Delta_k = \varepsilon \|\nabla\phi(x_k)\| M_s$, $0 < \varepsilon < 1$, $M_s = \|B_k^{-1}\|$, $p$ is a nonnegative integer, and matrix $B_k$ is an approximation of $\nabla g(x_k)^T g(x_k)$ which is generated by the following BFGS formula [31]:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \tag{5}$$

where $s_k = x_{k+1} - x_k$, $y_k = g(x_k + \delta_k) - g_k$, $\delta_k = g_{k+1} - g_k$. By $y_k = g(x_k + \delta_k) - g_k$, we have the approximate relations

$$y_k = g(x_k + \delta_k) - g_k \approx \nabla g_k \delta_k \approx \nabla g_k \nabla g_{k+1} s_k$$

Since $B_{k+1}$ satisfies the secant equation $B_{k+1} r_k = y_k$ and $\nabla g_k$ is symmetric, we have approximately

$$B_{k+1} \approx \nabla g_k \nabla g_{k+1} s_k = \nabla g_k B_k s_k$$

This means that $B_{k+1}$ approximates $\nabla g_k \nabla g_{k+1}$ along direction $s_k$. We all know that the update Eq.5 can ensure the matrix $B_{k+1}$ inherits positive property of $B_k$ if the condition $s_k^T y_k > 0$ is satisfied. Then we can use this way to insure the positive property of $B_k$.

This paper is organized as follows. In the next section, the new algorithm for solving Eq.1 is represented. In Section 3, we prove the convergence of the given algorithm. The numerical results of the method are reported in Section 4.

2. THE NEW METHOD

In this section, we give our algorithm for solving Eq.1. Firstly, one definition is given. Let

$$\varphi_{(k)} = \max_{0 \leq j < n(k)} \{\varphi_{i-j}\}, \quad k = 0, 1, 2, \cdots \tag{6}$$

where $n(k) = \min\{M, k\}, \quad M \geq 0$ is an integer constant. Now the algorithm is given as follows.

- **Algorithm 1.**
  - **Initial:** Given constants $\rho, \varepsilon \in (0, 1)$, $M = 0$, $x_0 \in \mathbb{R}^n$, $B_0 \in \mathbb{R}^{n \times n}$. Let $k = 0$;
  - **Step 1:** If $\|\nabla\phi_k\| < \varepsilon$, stop. Otherwise, go to step 2;
  - **Step 2:** Solve the problem Eq.4 with $\Delta = \Delta_k$ to get $d_k$;
  - **Step 3:** Calculate $n(k)$, $\varphi_{i(k)}$ and the following $r_k$:

$$r_k = \frac{\varphi_{i(k)} - \varphi(x_k + d_k)}{q_k(0) - q_k(d_k)} \tag{7}$$

If $r_k < \rho$, then we let $p = p + 1$, go to step 2. Otherwise, go to step 4;
  - **Step 4:** Let $x_{k+1} = x_k + d_k$, $\delta_k = g_{k+1} - g_k$, $y_k = g(x_k + \delta_k) - g_k$. If $d_k^T y_k > 0$, update $B_{k+1}$ by Eq.5, otherwise let $B_{k+1} = B_k$.

**Remark.** i) In this algorithm, the procedure of “Step 2-Step 3-Step 2” is named as inner cycle. ii) The Step 4 in Algorithm 1 ensures that the matrix sequence $\{B_k\}$ is positive definite.

In the following, we give some assumptions.

**Assumption A.** i) Let $\Omega$ be the level set defined by

$$\Omega = \{x : \|g(x)\| \leq \|g(x_k)\|\} \tag{8}$$

is bounded and $g(x)$ is continuously differentiable in $\Omega$ for all any given $x_k \in \mathbb{R}^n$.

ii) The matrices $\{B_k\}$ are uniformly bounded on $\Omega$, which means that there exists a positive constant $M$ such that

$$\|B_k\| \leq M, \quad \forall k \tag{9}$$

Based on Assumption A and Remark (ii), we have the following lemma.

**Lemma 2.1.** Suppose that Assumption A(jj) holds. If $d_k$ is the solution of Eq.4, then we have

$$-q_k(d_k) \geq -1 \frac{1}{2} \|\nabla\phi(x_k)\| \min\{\Delta_k, \|\nabla\phi(x_k)\|\} \tag{10}$$

**Proof.** Using $d_k$ is the solution of Eq.4, for any $\alpha \in [0, 1]$, we get

$$-q_k(d_k) \geq -q_k(-\alpha \frac{\Delta_k}{\|\nabla\phi(x_k)\|}) \|\nabla\phi(x_k)\|$$

$$= \alpha \Delta_k \|\nabla\phi(x_k)\| - \frac{1}{2} \alpha^2 \Delta_k^2 \left\|\nabla\phi(x_k)\right\|^2 B_k \nabla\phi(x_k) \|\nabla\phi(x_k)\| \right\|$$

$$\geq \alpha \Delta_k \|\nabla\phi(x_k)\| - \frac{1}{2} \alpha^2 \Delta_k^2 \|B_k\|$$

Then, we have

$$-q_k(d_k) \geq \max \{\alpha \Delta_k \|\nabla\phi(x_k)\| - \frac{1}{2} \alpha^2 \Delta_k^2 \|B_k\|\} \|\nabla\phi(x_k)\| \min\{\Delta_k, \|\nabla\phi(x_k)\|\} \tag{11}$$

$$\geq \frac{1}{2} \|\nabla\phi(x_k)\| \min\{\Delta_k, \|\nabla\phi(x_k)\|\} \tag{12}$$

Copyright © 2010 SciRes.

OPEN ACCESS
The proof is complete.
In the next section, we will concentrate on the convergence of Algorithm 1.

3. CONVERGENCE ANALYSIS

The following lemma guarantees that Algorithm 1 does not cycle infinitely in the inner cycle.

Lemma 3.1. Let the Assumption A hold. Then Algorithm 1 is well defined, i.e., Algorithm 1 does not cycle in the inner cycle infinitely.

Proof. First, we prove that the following relation holds when $p$ is sufficiently large

$$
\varphi_k - \varphi(x_{k+1}) \geq \rho
$$

(11)

Obviously, $\| \nabla \varphi(x_k) \| \geq \epsilon$ holds, otherwise, Algorithm 1 stops. Hence

$$
\Delta_k = \frac{e^p \| \nabla \varphi(x_k) \|}{|| B_k ||} \to 0, \ p \to \infty
$$

(12)

By Lemma 2.1, we conclude that

$$
-q_k(d_k) \geq \frac{1}{2} \| \nabla \varphi(x_k) \| \min \{ \Delta_k, \frac{\psi_k(\varphi(x_k))}{\| B_k \|} \} \geq \frac{1}{2} \epsilon \Delta_k,
$$

(13)

as $p \to \infty$

Consider

$$
\varphi_k - \varphi(x_{k+1}) - q_k(d_k) = O(\| d_k \|^2)
$$

(14)

By Eqs.12-14, and $\| d_k \| \leq \Delta_k$, we get

$$
-q_k(d_k) - \frac{1}{2} \| \nabla \varphi(x_k) \| \min \{ \Delta_k, \frac{\psi_k(\varphi(x_k))}{\| B_k \|} \} \leq 20(\| d_k \|^2) \frac{\epsilon \Delta_k}{\Delta_k} \to 0
$$

(15)

Therefore, for $p$ sufficiently large, which implies Eq.11. The definition of the algorithm means that

$$
r_k = \frac{\varphi_{(k)} - \varphi(x_{k+1})}{-q_k(d_k)} \geq \frac{\varphi_k - \varphi(x_{k+1})}{-q_k(d_k)} \geq \rho.
$$

This implies that Algorithm 1 does not cycle in the inner cycle infinitely. Then we complete the proof of this lemma.

Lemma 3.2. Let Assumption A hold and $\{ x_k \}$ be generated by the Algorithm 1. Then we have $\{ x_k \} \subset \Omega$.

Proof. We prove the result by induction. Assume that $\{ x_k \} \subset \Omega$, for all $k \geq 0$. By using the definition of the algorithm, we have

$$
r_{(k)} \geq \rho > 0
$$

(16)

Then we get

$$
\varphi_{(k)} \geq \varphi_{k+1} - \rho q_k(d_k) > \varphi_{k+1}
$$

By $l(k) \leq k$, $\varphi_{(k)} \leq \varphi_0$, from Eq.16, we have

$$
\varphi_{k+1} \leq \varphi_0
$$

this implies

$$
\| g_{k+1} \| \leq \| g_0 \|
$$

i.e.,

$$
x_{k+1} \in \Omega
$$

which completes the proof.

Lemma 3.3. Let Assumption A hold. Then $\{ \varphi_{(k)} \}$ is not increasing monotonically and is convergent.

Proof. By the definition of the algorithm, we get

$$
\varphi_{(k)} \geq \varphi_{k+1}, \ \forall k
$$

(17)

We proceed the proof in the following two cases.

1) $k \geq M$ . In this case, from the definition of $\varphi_{(k)}$ and Eq.17, it holds that

$$
\varphi_{(k)} = \max_{0 \leq j \leq M} \{ \varphi_{k+1-j} \}
$$

(18)

$$
\leq \varphi_{(k)}
$$

2) $k < M$ . In this case, using induction, we can prove that

$$
\varphi_{(k)} = \varphi_0
$$

Therefore, the sequence $\{ \varphi_{(k)} \}$ is not increasing monotonically. By Assumption A(j) and Lemma 3.2, we know that $\{ \varphi_k \}$ is bounded. Then $\{ \varphi_{(k)} \}$ is convergent.

In the following theorem, we establish the convergence of Algorithm 1.

Theorem 3.1. Let the conditions in Assumption A hold. If $\epsilon = 0$, then the algorithm either stops finitely or generates an infinite sequence $\{ x_k \}$ such that

$$
\liminf_{k \to \infty} \varphi_k = 0
$$

(19)

Proof. We prove the theorem by contradiction. Assume that the theorem is not true. Then there exists a constant $\epsilon_i > 0$ satisfying

$$
\varphi_k \geq \epsilon_i, \ \forall k
$$

(20)

By Assumption A(j) and the definition of $B_k$, there exists a constant $m > 0$ such that

$$
\| B_k \| \geq m
$$

(21)

Therefore, according to Assumption A(j), Lemma 2.1, Eq.20, and Eq.21, there is a constant $h_i > 0$ such that

$$
-q_k(d_k) \geq h_i e^{\rho_k}
$$

(22)

where $p_k$ is the value of $p$ at which the algorithm
gets out of the inner cycle at the point $x_k$. By step 2, step 3, step 4, and Eq. 22, we know

$$
\varphi_{i(k)} \geq \varphi_{i+1} + \rho h \varphi^p
$$

(23)

Then

$$
\varphi_{i(k+1)} \leq \varphi_{i(k)} - \rho h \varphi^p
$$

(24)

By Lemma 3.3 and Eq. 24, we deduce that

$$
\lim_{k \to \infty} \varphi_{i(k)} = \infty
$$

(25)

The definition of the algorithm implies that $d_{i(k)}^*$ which corresponds to the following subproblem is unacceptable:

$$
\min_{d \in \mathbb{R}^n} \varphi_{i(k)}^* d + \frac{1}{2} d^T B_{i(k)} d = q_{i(k)}(d),
$$

s.t. \[ \|d\| \leq e^{\rho i(k-1)} M_{i(k)} \varphi_{i(k)} = \frac{\Delta_{i(k)}}{c} \] (26)

i.e.,

$$
\frac{\varphi_{i(k)} - \varphi(x_{i(k)} + d_{i(k)}^*)}{-q_{i(k)}(d_{i(k)}^*)} < \rho
$$

(27)

By the definition of $\varphi_{i(k)}$, we have

$$
\frac{\varphi_{i(k)} - \varphi(x_{i(k)} + d_{i(k)}^*)}{-q_{i(k)}(d_{i(k)}^*)} \geq \frac{\varphi_{i(k)} - \varphi(x_{i(k)} + d_{i(k)}^*)}{-q_{i(k)}(d_{i(k)}^*)}
$$

(28)

By step 2, step 3, and step 4, we have that when $k$ is sufficiently large, the following formula holds:

$$
\frac{\varphi_{i(k)} - \varphi(x_{i(k)} + d_{i(k)}^*)}{-q_{i(k)}(d_{i(k)}^*)} \geq \rho
$$

(29)

This combines with Eq. 28 will contradicts Eq. 27. The contradiction shows that the theorem is true. The proof is complete.

Remark. Theorem 3.1 shows that the iterative sequence $\{x_k\}$ generated by Algorithm 1 such that $\nabla g(x_k) g(x_k) \to 0$. If $x^*$ is a cluster point of $\{x_k\}$ and $\nabla g(x^*)$ is nonsingular, then we have $\|g(x_k)\| \to 0$. This is a standard convergence result for nonlinear equations. At present, there is no method that can satisfy $\|g(x_k)\| \to 0$ without the assumption that $\nabla g(x^*)$ is nonsingular.

4. NUMERICAL RESULTS

In this section, results of some preliminary numerical experiments are reported to test our given method.

**Problem.** The discretized two-point boundary value problem is the same to the problem in [39]

$$
g(x) = Ax + \frac{1}{(n+1)^2} F(x) = 0
$$

where $A$ is the $n \times n$ tridiagonal matrix given by

$$
A = \begin{bmatrix}
3 & -1 & & & \\
-1 & 3 & -1 & & \\
& -1 & 3 & \ddots & -1 \\
& & \ddots & \ddots & \ddots \\
& & & -1 & 3
\end{bmatrix}
$$

and $F(x) = (F_1(x), F_2(x), \ldots, F_n(x))^T$, with

$$
F_j(x) = \sin x_j - 1, \quad i=1,2,\ldots,n
$$

In the experiments, the parameters were chosen as $c = 0.01$, $M = 10$, and $\rho = 0.8$, $B_0$ is the unit matrix. Solving the subproblem Eq. 4 to get $d_k$ by Dogleg method. The program was coded in MATLAB 7.0. We stopped the iteration when the condition $\|g_k\| \leq 10^{-5}$ was satisfied. The columns of the tables have the following meaning:

- **Dim:** the dimension of the problem.
- **NG:** the number of the function evaluations.
- **NI:** the total number of iterations.
- **GG:** the norm of the function evaluations.

**Discussion.** In this paper, based on [23], a modified algorithm for solving symmetric nonlinear equations is presented. The convergent result is established and the numerical results are reported. We hope that the proposed method can be a topic of further research for symmetric nonlinear equations.

**Table 1.** Test results for problem.

<table>
<thead>
<tr>
<th>$x_0$</th>
<th>(2, … ,2)</th>
<th>(10, … ,10)</th>
<th>(50, … ,50)</th>
<th>(-10, … , -10)</th>
<th>(-2, … , -2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dim</td>
<td>NI/NG/GG</td>
<td>NI/NG/GG</td>
<td>NI/NG/GG</td>
<td>NI/NG/GG</td>
<td>NI/NG/GG</td>
</tr>
</tbody>
</table>
REFERENCES


