ON CONVERGENCE PROPERTIES OF THE MODIFIED TRUST REGION METHOD UNDER HÖLDERIAN ERROR BOUND CONDITION

JIRUI MA
School of Mathematical Sciences, Shanghai Jiao Tong University
Shanghai 200240, China

JINYAN FAN*
School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University
Shanghai 200240, China

(Communicated by Liqun Qi)

Abstract. Trust region method is one of the important methods for nonlinear equations. In this paper, we show that the modified trust region method converges globally under the Hölderian continuity of the Jacobian. The convergence order of the method is also given under the Hölderian error bound condition.

1. Introduction. In this paper, we consider the system of nonlinear equations

\[ F(x) = 0, \]

where \( F: \mathbb{R}^n \rightarrow \mathbb{R}^n \) is continuously differentiable. Throughout the paper, we assume that the solution set of (1) is nonempty and denote it by \( X^* \). Nonlinear equations have wide applications in chemical technology, industrial engineering, economy and so on. There are many existing methods for nonlinear equations, such as the Newton method, the Levenberg-Marquardt method, the trust region method, the subspace method, etc. (cf. [7, 8, 9, 16, 17]).

At the \( k \)-th iteration, the trust region method for (1) solves the subproblem

\[
\begin{align*}
\min_{d \in \mathbb{R}^n} & \quad \| F_k + J_k d \|^2 \\
\text{s.t.} & \quad \| d \| \leq \Delta_k
\end{align*}
\]

to obtain the trial step \( d_k \), where \( F_k = F(x_k) \), \( J_k \) is the Jacobian at \( x_k \), \( \Delta_k > 0 \) is trust region radius and \( \| \cdot \| \) refers to the 2-norm.

Define the actual reduction and the predicted reduction of the merit function \( \phi(x) = \| F(x) \|^2 \) as \( \text{Ared}_k = \| F_k \|^2 - \| F(x_k + d_k) \|^2 \) and \( \text{Pred}_k = \| F_k \|^2 - \| F_k + J_k d_k \|^2 \), respectively. The ratio \( r_k \) of \( \text{Ared}_k \) to \( \text{Pred}_k \) plays a key role in deciding whether the trial step \( d_k \) is satisfactory and how to update the trust region radius.

2020 Mathematics Subject Classification. 65K05, 65K10, 90C26, 90C30.

Key words and phrases. Nonlinear equations, modified trust region method, Hölderian continuity, Hölderian error bound, convergence order.

The authors are supported by the National Natural Science Foundation of China grant 11971309.

* Corresponding author: Jinyan Fan.
Generally, the step $d_k$ is accepted if the ratio is larger than a small positive constant, and rejected otherwise. The trust region radius is enlarged if the trial step is satisfactory, and reduced otherwise \cite{15}. That is, we set

$$x_{k+1} = \begin{cases} x_k + d_k, & \text{if } r_k \geq c_0, \\ x_k, & \text{otherwise} \end{cases} \quad (3)$$

and

$$\Delta_{k+1} = \begin{cases} [p_0\|d_k\|, p_1\Delta_k], & \text{if } r_k < c_1, \\ [\Delta_k, p_2\Delta_k], & \text{otherwise}, \end{cases} \quad (4)$$

where $0 < c_0 < c_1 < 1$ and $0 < p_0 < p_1 < 1 < p_2$ are constants.

In \cite{2}, Fan proposed a new kind of trust region method for (1) with the trust region radius converges to zero. It takes the trust region radius as

$$\Delta_{k+1} = \mu_{k+1}\|F_{k+1}\|^\delta, \quad (5)$$

where $\delta \in (\frac{1}{2}, 1)$ is a given constant and $\mu_{k+1}$ is updated by

$$\mu_{k+1} = \begin{cases} c_4\mu_k, & \text{if } r_k < c_1, \\ \mu_k, & \text{if } r_k \in [c_1, c_2], \\ \min\{c_4\mu_k, M\}, & \text{otherwise}, \end{cases} \quad (6)$$

where $0 < c_1 < c_2 < 1$ and $0 < c_3 < 1 < c_4$ are constants. It was shown that the method converges to the solution set with order $2\delta$ under the local error bound condition, which is weaker than the nonsingularity of the Jacobian.

Further, Fan and Lu proposed a modified trust region method for (1) (cf. \cite{5}). At the $k$-th iteration, it first solves (2) to obtain the trust region step $d_k$ and sets $y_k = x_k + d_k$, then computes an inexact trust region step $\bar{d}_k$ by solving the subproblem

$$\begin{aligned}
\min_{d \in \mathbb{R}^n} & \quad \|F(y_k) + J_kd\|^2 \\
\text{s.t.} & \quad \|d\| \leq \bar{\Delta}_k = \mu_k\|F(y_k)\|^\delta.
\end{aligned} \quad (7)$$

The trial step is set as $s_k = d_k + \bar{d}_k$. Define the actual reduction of the merit function $\phi(x) = \|F(x)\|^2$ at the $k$-th iteration as usual:

$$\text{Ared}_k = \|F_k\|^2 - \|F(x_k + d_k + \bar{d}_k)\|^2. \quad (8)$$

Since $s_k$ may not be a descent direction of the merit function $\phi(x)$ at $x_k$, we define the new prediction as

$$\text{Pred}_k = \|F_k\|^2 - \|F_k + J_kd_k\|^2 + \|F(y_k)\|^2 - \|F(y_k) + J_kd_k\|^2. \quad (9)$$

The ratio of the actual reduction to the predicted reduction is then defined as

$$r_k = \frac{\text{Ared}_k}{\text{Pred}_k} = \frac{\|F_k\|^2 - \|F(x_k + d_k + \bar{d}_k)\|^2}{\|F_k\|^2 - \|F_k + J_kd_k\|^2 + \|F(y_k)\|^2 - \|F(y_k) + J_kd_k\|^2}. \quad (10)$$

The new iterative point is set as

$$x_{k+1} = \begin{cases} x_k + d_k + \bar{d}_k, & \text{if } r_k \geq c_0, \\ x_k, & \text{otherwise}, \end{cases} \quad (11)$$

and the new trust region radius are computed by (5) with $\mu_k$ being updated by (6). The modified trust region method above converges globally under the Lipschitz continuity of the Jacobian and its convergence rate was also given under the local error bound condition \cite{5}. 

\text{JIRUI MA AND JINYAN FAN}
As we know, the local error bound condition is weaker than the nonsingularity of the Jacobian. Interested readers are referred to [3, 4, 6, 14, 18] for details and examples. However, in real applications some problems may not satisfy the local error bound condition but satisfy the more general Hölderian error bound condition.

**Definition 1.1.** Let $N(x^*)$ be some neighbourhood of $x^* \in X^*$. We say $F(x)$ provides a Hölderian error bound on $N(x^*)$ for (1), if there exist constants $\kappa_{\text{hleb}} > 0$ and $0 < \gamma \leq 1$ such that

$$\kappa_{\text{hleb}} \cdot \text{dist}(x, X^*) \leq \|F(x)\|^\gamma, \quad \forall x \in N(x^*).$$

(12)

For example, the function $h(x_1, \ldots, x_n) = (s_1 x_1^{t_1}, \ldots, s_n x_n^{t_n})$ with $t_i > 1, s_i > 0 (i = 1, \ldots, n)$ satisfies the Hölderian error bound condition with $\gamma = \max\{t_i\}$ in some neighbourhood of zero point, but does not satisfy the local error bound condition.

Obviously, the Hölderian error bound condition includes the local error bound condition as a special case when $\gamma = 1$. The convergence properties of the Levenberg-Marquardt method and the inexact Levenberg-Marquardt method under the Hölderian error bound condition were studied in [1, 12, 13]. In this paper, we investigate the convergence properties of the modified trust region method under the Hölderian continuity of the Jacobian and the Hölderian error bound condition.

The paper is organized as follows. In Section 2, we discuss the global convergence of the modified trust region method under the Hölderian continuity of the Jacobian. In Section 3, we study the convergence rate of the method under the Hölderian continuity of the Jacobian and the Hölderian error bound condition. Finally, we conclude the paper in Section 4.

2. The algorithm and global convergence. In this section, we first present the modified trust region algorithm for nonlinear equations given in [5], then discuss the global convergence of the algorithm under the Hölderian continuity condition.

**Algorithm 2.1.** (The modified trust region algorithm)

*Step 0.* Given $x_1 \in \mathbb{R}^n, 0 < c_0 < c_1 < c_2 < 1, 0 < c_3 < 1 < c_4, \mu_1 > 0, M > c_4 \mu_1, \delta > 0, \Delta_1 = \mu_1 \|F_1\|^\delta, k := 1.$

*Step 1.* If $\|J_k^T F_k\| = 0$, then stop;

Solve (2) to obtain $d_k$;

Set $y_k = x_k + d_k$;

Solve (7) to obtain $\tilde{d}_k$;

*Step 2.* Compute $r_k = A_{\text{red}} k / \text{Pred}_k$ by (10);

Update $x_{k+1}$ by (11);

Update $\mu_{k+1}$ by (6).

*Step 3.* Set $k := k + 1$, go to Step 1.

We make the following assumption.

**Assumption 2.2.** $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable, and the Jacobian $J(x)$ is Hölderian continuous and bounded above, i.e. there exist constants $0 < v \leq 1, \kappa_{\text{hj}} > 0$ and $\kappa_{\text{bj}} > 0$ such that

$$\|J(x) - J(y)\| \leq \kappa_{\text{hj}} \|x - y\|^v, \quad \forall x, y \in \mathbb{R}^n,$$

$$\|J(x)\| \leq \kappa_{\text{bj}}, \quad \forall x \in \mathbb{R}^n.$$
It follows from (13) that for all \( x, y \in \mathbb{R}^n \),
\[
\| F(y) - F(x) - J(x)(y - x) \| = \left\| \int_0^1 J(x + t(y - x))(y - x) dt - J(x)(y - x) \right\|
\leq \|y - x\| \int_0^1 \| J(x + t(y - x)) - J(x) \| dt \\
\leq \kappa h_j \|y - x\|^{1+v} \int_0^1 t^v dt \\
= \frac{\kappa h_j}{1 + v} \|y - x\|^{1+v}.
\] (15)

Due to Powell’s result given in [11], the trust region step \( d_k \) obtained by (2) satisfies
\[
\| F_k \|^2 - \| F_k + J_k d_k \|^2 \geq \| J_k^T F_k \| \min \left\{ \Delta_k, \frac{||J_k^T F_k||}{\|J_k J_k\|} \right\}.
\] (16)

Similarly,
\[
\| F(y_k) \|^2 - \| F(y_k) + J_k d_k \|^2 \geq \| J_k^T F(y_k) \| \min \left\{ \Delta_k, \frac{||J_k^T F(y_k)||}{\|J_k J_k\|} \right\}.
\] (17)

Hence, the predicted reduction at the \( k \)-th iteration satisfies
\[
\text{Pred}_k \geq \| J_k^T F_k \| \min \left\{ \Delta_k, \frac{||J_k^T F_k||}{\|J_k J_k\|} \right\} + \| J_k^T F(y_k) \| \min \left\{ \Delta_k, \frac{||J_k^T F(y_k)||}{\|J_k J_k\|} \right\}.
\] (18)

The inequality (18) plays an important role in the global convergence of Algorithm 2.1.

**Theorem 2.3.** Under Assumption 2.2, the sequence \( \{x_k\} \) generated by Algorithm 2.1 satisfies
\[
\liminf_{k \to +\infty} \| J_k^T F_k \| = 0.
\] (19)

**Proof.** We prove by contradiction. Suppose that (19) is not true. Then there exists a constant \( \tau > 0 \) such that
\[
\| J_k^T F_k \| \geq \tau, \quad \forall k.
\] (20)

Denote the index set \( \mathbb{I} = \{ k \mid r_k \geq c_1 \} \). Since \( \{\|F_k\|\} \) is monotonically non-increasing and bounded below, by (14), (18) and (20), we have
\[
+\infty > \sum_{k=1}^{\infty} (\|F_k\|^2 - \|F_{k+1}\|^2) \\
\geq \sum_{k \in \mathbb{I}} (\|F_k\|^2 - \|F_{k+1}\|^2) \\
\geq \sum_{k \in \mathbb{I}} c_1 \text{Pred}_k \\
\geq \sum_{k \in \mathbb{I}} c_1 \|J_k^T F_k\| \min \left\{ \Delta_k, \frac{||J_k^T F_k||}{\|J_k J_k\|} \right\} \\
\geq \sum_{k \in \mathbb{I}} c_1 \tau \min \left\{ \Delta_k, \frac{\tau}{\kappa h_j} \right\}.
\] (21)
If $\mathbb{I}$ is finite, then $\mu_{k+1} = c_3\mu_k$ for all large $k$. In view of $0 < c_3 < 1$, we have
\[
\lim_{k \to \infty} \mu_k = 0. \tag{22}
\]
If $\mathbb{I}$ is infinite, then $(21)$ gives
\[
\lim_{k \in \mathbb{I}, k \to \infty} \Delta_k = \lim_{k \in \mathbb{I}, k \to \infty} \mu_k\|F_k\|^\delta = 0. \tag{23}
\]
By (14) and (20),
\[
\|F_k\| \geq \frac{\|J_k^TF_k\|}{\|J_k\|} \geq \frac{\tau}{\kappa_{hj}}. \tag{24}
\]
Combining (23) with (24), we obtain
\[
\lim_{k \in \mathbb{I}, k \to \infty} \mu_k = 0. \tag{25}
\]
Because $\mu_{k+1} = c_3\mu_k$ and $0 < c_3 < 1$ for $k \notin \mathbb{I}$, (22) also holds true when $\mathbb{I}$ is infinite.

Note that $\|F_k\| \leq \|F_1\|$ and $\|d_k\| \leq \Delta_k$. It follows from (22) that
\[
\lim_{k \to \infty} \Delta_k = 0, \tag{26}
\]
and
\[
\lim_{k \to \infty} d_k = 0. \tag{27}
\]
These, together with (15), imply that there exist constants $\bar{\tau}$ and $\tilde{\tau}$ such that, for all sufficiently large $k$,
\[
\|F(y_k)\| \leq \|F_k + J_kd_k\| + \frac{\kappa_{hj}}{1 + v}\|d_k\|^{1+v} \leq \|F_k\| + \frac{\kappa_{hj}}{1 + v} \leq \bar{\tau} \tag{28}
\]
and
\[
\|J_k^TF_k\| \geq \|J_k^TF_k\| - \|J_k^TJ_kd_k\| - \frac{\kappa_{hj}}{1 + v}\|J_k\||d_k\|^{1+v} \geq \tilde{\tau}. \tag{29}
\]
Hence,
\[
\lim_{k \to \infty} \Delta_k = \lim_{k \to \infty} \mu_k\|F(y_k)\|^\delta = 0, \tag{30}
\]
which gives
\[
\lim_{k \to \infty} \tilde{d}_k = 0. \tag{31}
\]

It then follows from (15) and (26) that
\[
\|F(x_k + d_k)\|^2 - \|F_k + J_kd_k\|^2 \leq \|F(x_k + d_k) - F_k - J_kd_k\|^2 + 2\|F(x_k + d_k) - F_k - J_kd_k\|\|F_k + J_kd_k\| \leq O(\|d_k\|^{2+2v}) + \|F_k + J_kd_k\|O(\|d_k\|^{1+v}) \leq O(\|d_k\|^{1+v}) \tag{32}
\]
Similarly, by (15) and (31),
\[
\begin{align*}
\left\| F(x_k + d_k + \bar{d}_k) \right\|^2 & - \left\| F(x_k + d_k) + J_k \bar{d}_k \right\|^2 \\
= & \left\| F(x_k + d_k + \bar{d}_k) - F_k - J_k(d_k + \bar{d}_k) + F_k + J_k(d_k + \bar{d}_k) \right\|^2 \\
& - \left\| F(x_k + d_k) + J_k \bar{d}_k - F_k - J_k(d_k + \bar{d}_k) + (F_k + J_k(d_k + \bar{d}_k)) \right\|^2 \\
= & \left\| F(x_k + d_k + \bar{d}_k) - F_k - J_k(d_k + \bar{d}_k) \right\|^2 \\
& + 2\left\| F(x_k + d_k + \bar{d}_k) - F_k - J_k(d_k + \bar{d}_k) \right\| \left\| F_k + J_k(d_k + \bar{d}_k) \right\| \\
& + \left\| F(x_k + d_k) + J_k \bar{d}_k - F_k - J_k(d_k + \bar{d}_k) \right\|^2 \\
& + 2\left\| F(x_k + d_k) - F_k - J_k d_k \right\| \left\| F_k + J_k(d_k + \bar{d}_k) \right\| \\
\leq & O(\left\| d_k + \bar{d}_k \right\|^{2+2v} + \left\| d_k \right\|^{2+2v}) \\
& + \left\| F_k + J_k(d_k + \bar{d}_k) \right\| O(\left\| d_k + \bar{d}_k \right\|^{1+v} + \left\| d_k \right\|^{1+v}) \\
\leq & O(\left\| d_k + \bar{d}_k \right\|^{1+v}) + O(\left\| d_k \right\|^{1+v}).
\end{align*}
\]
This implies that
\[ r_k \to 1. \] (34)

Thus, by (6), there exists a constant \( \bar{\mu} > 0 \) such that
\[ \mu_k \geq \bar{\mu} \] (35)
holds for all large \( k \), which leads a contradiction to (22). So (20) cannot be true. The proof is completed. \hfill \Box

Theorem 2.3 indicates that at least one accumulation point of the sequence generated by Algorithm 2.1 is a stationary point of the merit function \( \phi(x) \). Actually, the strong global convergence of Algorithm 2.1 can be further obtained, that is, each accumulation point of the sequence is a stationary point of the merit function.

**Theorem 2.4.** Under Assumption 2.2, the sequence \( \{x_k\} \) generated by Algorithm 2.1 satisfies
\[
\lim_{k \to +\infty} \left\| J_k^T F_k \right\| = 0.
\] (36)

**Proof.** We prove by contradiction. Suppose that (36) is not true. Then there exists a constant \( \epsilon > 0 \) such that \#\{k : \left\| J_k^T F_k \right\| \geq \epsilon \} = \infty. Denote the index sets
\[
K = \{k_i : \left\| J_k^T F_k \right\| \geq \epsilon \}
\] (37)
and
\[
L = \{l : \left\| J_l^T F_l \right\| < \frac{\epsilon}{3} \}.
\] (38)

It follows from Theorem 2.3 that \#L is also infinity.
For each $k_i \in K$, let $l_i = \min\{l \in L : l > k_i\}$. Hence, we have

$$
\|J_k^T F_k\| \geq \frac{\epsilon}{3}, \quad \forall k_i \leq k < l_i.
$$

(39)

If the $k$-th step is successful ($k_i \leq k < l_i$), by (18) and $c_0 < 1$,

$$
\|F_k\|^2 - \|F_{k+1}\|^2 \geq c_0 \|J_k^T F_k\| \min \left\{ \Delta_k, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|} \right\}
\geq \frac{c_0 \epsilon}{3} \min \left\{ \|d_k\|, \frac{\epsilon}{3\kappa_{kj}} \right\}.
$$

(40)

Since $\|F_k\|$ is non-increasing and bounded below, we have, for $k_i \leq k < l_i$,

$$
+\infty > \sum_{j=1}^{\infty} (\|F_j\|^2 - \|F_{j+1}\|^2)
\geq \sum_{i=1}^{\infty} (\|F_k\|^2 - \|F_{k+1}\|^2)
\geq \sum_{i=1}^{\infty} \frac{c_0 \epsilon}{3} \min \left\{ \|d_k\|, \frac{\epsilon}{3\kappa_{kj}} \right\}.
$$

Thus, for $k_i \leq k < l_i$, $d_k \to 0$ as $i \to \infty$. This, together with (40), gives

$$
\|F_k\|^2 - \|F_{k+1}\|^2 \geq \frac{c_0 \epsilon}{3} \|x_{k+1} - x_k\| \quad \text{for} \quad k_i \leq k < l_i \quad \text{as} \quad i \to \infty.
$$

(41)

Obviously, (41) also holds true when the $k$-th step is unsuccessful ($k_i \leq k < l_i$).

Thus,

$$
c_0 \epsilon \|x_{k_i} - x_{l_i}\| \leq c_0 \epsilon (\|x_{k_i} - x_{k_i+1}\| + \|x_{l_i-1} - x_{l_i}\|)
\leq 3(\|F_{k_i}\|^2 - \|F_{k_i+1}\|^2) + \ldots + 3(\|F_{l_i-1}\|^2 - \|F_{l_i}\|^2)
= 3\|F_{k_i}\|^2 - 3\|F_{l_i}\|^2.
$$

(42)

Since $\{\|F_k\|^2\}$ is monotonically nonincreasing and bounded below, $\{\|F_k\|^2\}$ converges. Hence,

$$
\|x_{k_i} - x_{l_i}\| \to 0 \quad \text{as} \quad i \to \infty.
$$

(43)

This, together with the fact that $\|J(x)^T F(x)\|$ is continuous, gives

$$
\|J_{k_i}^T F_{k_i} - J_{l_i}^T F_{l_i}\| < \frac{1}{3} \epsilon
$$

(44)

for sufficiently large $i$.

On the other hand, due to $\|J_{k_i}^T F_{k_i}\| \geq \epsilon$ and $\|J_{l_i}^T F_{l_i}\| < \frac{1}{3} \epsilon$, we have

$$
\|J_{k_i}^T F_{k_i} - J_{l_i}^T F_{l_i}\| > \frac{2}{3} \epsilon,
$$

(45)

which contradicts (44). So, (36) holds true. The proof is completed. \qed
3. Convergence rate of the algorithm. In this section, we study the convergence rate of Algorithm 2.1 under the Hölderian continuity of the Jacobian and the Hölderian error bound condition. We assume that the sequence \(\{x_k\}\) generated by Algorithm 2.1 converges to the solution set \(X^*\) of (1) and \(\{x_k\}\) lies in some neighbourhood of \(x^* \in X^*\).

Assumption 3.1. (i) \(F : \mathbb{R}^n \to \mathbb{R}^n\) is continuously differentiable, and \(F(x)\) provides a Hölderian error bound of order \(\gamma\) at \(x^*\), i.e., there exist constants \(0 < \gamma \leq 1\) and \(0 < b < \frac{1}{2}\) such that
\[
\kappa_{\text{hleb}} \cdot \text{dist}(x, X^*) \leq \|F(x)\|^\gamma, \quad \forall x \in N(x^*, b).
\] (46)

(ii) The Jacobian of \(F(x)\) is Hölderian continuous, i.e., there exist constants \(0 < v \leq 1\) and \(\kappa_{hj} > 0\) such that
\[
\|J(x) - J(y)\| \leq \kappa_{hj} \|x - y\|^v, \quad \forall x, y \in N(x^*, b).
\] (47)

Similarly to (15), we have
\[
\|F(y) - F(x) - J(x)(y - x)\| \leq \frac{\kappa_{hj}}{1 + v} \|y - x\|^{1 + v}, \quad \forall x, y \in N(x^*, b).
\] (48)

Hence, there exists a constant \(\kappa_{lf} > 0\) such that
\[
\|F(y) - F(x)\| \leq \kappa_{lf} \|y - x\|, \quad \forall x \in N(x^*, b).
\] (49)

Denote by \(\bar{x}_k \in X^*\) the vector that satisfies
\[
\|\bar{x}_k - x_k\| = \text{dist}(x_k, X^*).
\] (50)

Lemma 3.2. Under Assumption 3.1, if \(\frac{1}{1 + v} < \delta < 1\), then there exists a positive constant \(\bar{\mu} > 0\) such that for all large \(k\),
\[
\mu_k \geq \bar{\mu}.
\] (51)

Proof. We first give an estimation of the predicted reduction \(\text{Pred}_k\). If \(\|\bar{x}_k - x_k\| \leq \|d_k\|\), then \(\bar{x}_k - x_k\) is a feasible solution of (2). By (46), (48) and \(\frac{1}{\gamma} = 1 + v\),
\[
\|F_k\| - \|F_k + J_k d_k\| \geq \|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\|
\geq \kappa_{hleb} \|\bar{x}_k - x_k\|^{\frac{1}{\gamma}} - \frac{\kappa_{hj}}{1 + v} \|\bar{x}_k - x_k\|^{1 + v}
\geq O(\|\bar{x}_k - x_k\|^{\frac{1}{\gamma}}).
\] (52)

If \(\|\bar{x}_k - x_k\| > \|d_k\|\), then it follows from (52) that
\[
\|F_k\| - \|F_k + J_k d_k\|
\geq \|F_k\| - \|F_k - \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} J_k(\bar{x}_k - x_k)\|
= \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} \|F_k\| - \left( \|F_k - \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} J_k(\bar{x}_k - x_k)\| - \left( 1 - \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} \right) \|F_k\| \right)
\geq \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} \|F_k\| - \|F_k - \left( 1 - \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} \right) F_k - \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} J_k(\bar{x}_k - x_k)\|
\geq \|d_k\| \left( \|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\| \right)
\geq O(\|d_k\| \|\bar{x}_k - x_k\|^{\frac{1}{\gamma} - 1}).
\] (53)
Combining (52) with (53), we obtain
\[ ||F_k||^2 - ||F_k + J_k d_k||^2 \geq ||F_k|| \left( ||F_k|| - ||F_k + J_k d_k|| \right) \]
\[ \geq ||F_k|| \min \{ O(||\bar{x}_k - x_k||^{\frac{1}{2}}), O(||d_k|| ||\bar{x}_k - x_k||^{\frac{1}{2}}) \}. \] (54)

Similarly, we have
\[ ||F(y_k)||^2 - ||F(y_k) + J_k \bar{d}_k||^2 \geq ||F(y_k)|| \min \{ O(||\bar{y}_k - y_k||^{\frac{1}{2}}), O(||\bar{d}_k|| ||\bar{y}_k - y_k||^{\frac{1}{2}}) \}. \] (55)

Thus,
\[ \text{Pred}_k \geq ||F_k|| \min \{ O(||\bar{x}_k - x_k||^{\frac{1}{2}}), O(||d_k|| ||\bar{x}_k - x_k||^{\frac{1}{2}}) \} \]
\[ + ||F(y_k)|| \min \{ O(||\bar{y}_k - y_k||^{\frac{1}{2}}), O(||\bar{d}_k|| ||\bar{y}_k - y_k||^{\frac{1}{2}}) \}. \] (56)

Next we prove (51). Since \(d_k\) and \(\bar{d}_k\) are solutions of (2) and (7) respectively, we have
\[ ||F_k + J_k d_k|| \leq ||F_k||, \quad ||F(y_k) + J_k \bar{d}_k|| \leq ||F(y_k)||. \] (57)

By (48) and (57),
\[ \left( ||F(y_k)||^2 - ||F_k + J_k d_k||^2 \right) = \left( ||F(y_k)|| + ||F_k + J_k d_k|| \right) \left( ||F(y_k)|| - ||F_k + J_k d_k|| \right) \]
\[ \leq (2 ||F_k + J_k d_k|| + O(||d_k||^{1+v})) O(||d_k||^{1+v}) \]
\[ \leq ||F_k|| O(||d_k||^{1+v}) + O(||d_k||^{2+2v}). \] (58)

Meanwhile, by (47), (48) and (57), we have
\[ ||F(x_k + d_k + \bar{d}_k)||^2 - ||F(y_k)||^2 \]
\[ = ||F(x_k + d_k + \bar{d}_k) - F(y_k) - J(y_k)\bar{d}_k + (J(y_k) - J_k)\bar{d}_k + (F(y_k) + J_k \bar{d}_k)||^2 \]
\[ \leq ||F(x_k + d_k + \bar{d}_k) - F(y_k) - J(y_k)\bar{d}_k||^2 + ||J(y_k) - J_k)\bar{d}_k||^2 \]
\[ + 2 ||F(x_k + d_k + \bar{d}_k) - F(y_k) - J(y_k)\bar{d}_k|| ||(J(y_k) - J_k)\bar{d}_k|| \]
\[ + 2 ||(J(y_k) - J_k)\bar{d}_k|| ||F(y_k) + J_k \bar{d}_k|| \]
\[ \leq O(||\bar{d}_k||^{2+2v} + ||d_k||^{2v} ||\bar{d}_k||^2 + ||\bar{d}_k||^{2+v} ||d_k||) \]
\[ + O(||\bar{d}_k||^{2+2v} + ||d_k||^{2v} ||\bar{d}_k||^2 + ||\bar{d}_k||^{2+v} ||d_k||) \]
\[ + ||F(y_k)|| O(||\bar{d}_k||^{1+v} + ||d_k||^v ||\bar{d}_k||). \] (59)

It then follows from (58) and (59) that
\[ |r_k - 1| = \frac{|A_{\text{red}} - \text{Pred}_k|}{\text{Pred}_k} \]
\[ \leq \frac{||F(y_k)||^2 - ||F_k + J_k d_k||^2 + ||F(x_k + d_k + \bar{d}_k)||^2 - ||F(y_k) + J_k \bar{d}_k||^2}{\text{Pred}_k} \]
\[ = \frac{F_k O(||d_k||^{1+v}) + ||F(y_k)|| O(||\bar{d}_k||^{1+v} + ||d_k||^v ||\bar{d}_k||)}{\text{Pred}_k} \]
\[ + \frac{O(||\bar{d}_k||^{2+2v} + ||\bar{d}_k||^{2+v} ||d_k||^v + ||\bar{d}_k||^{2} ||d_k||^{2+v} + ||d_k||^{2+2v})}{\text{Pred}_k}. \] (60)
By (2), (6), (7) and (49),

\[
\|d_k\| \leq \Delta_k = \mu_k \|F_k\|^\delta \leq M \kappa^\delta_{ij} \|\bar{x}_k - x_k\|^\delta, \\
\|\tilde{d}_k\| \leq \bar{\Delta}_k = \mu_k \|F(y_k)\|^\delta \leq M \kappa^\delta_{ij} \|\bar{y}_k - y_k\|^\delta.
\]

(61) \hspace{3cm} (62)

In view of \(\frac{1}{\gamma(1+\nu)} < \delta \leq 1, 0 < \gamma \leq 1\), (46) and (56), we have

\[
\frac{\|F_k\|O(\|d_k\|^{2+2v})}{\text{Pred}_k} \leq \frac{\|F_k\|O(\|d_k\|^{1+v})}{\text{Pred}_k} \leq \|F_k\| \min\{O(\|\bar{x}_k - x_k\|^{\frac{v}{\gamma}}), O(\|d_k\| \|\bar{x}_k - x_k\|^{\frac{1}{\gamma} - 1})\} \\
\quad \leq \max\{O(\|\bar{x}_k - x_k\|^{\delta(1+v) - \frac{1}{\gamma}}), O(\|\bar{x}_k - x_k\|^{\delta(1+v) - \frac{4}{\gamma} + 1})\} \\
\quad \leq O(\|\bar{x}_k - x_k\|^{\delta(1+v) - \frac{1}{\gamma}}) \to 0,
\]

(63)

\[
\frac{O(\|d_k\|^{2+2v})}{\text{Pred}_k} \leq \frac{O(\|d_k\|^{2+2v})}{\text{Pred}_k} \leq \|F_k\| \min\{O(\|\bar{x}_k - x_k\|^{\frac{v}{\gamma}}), O(\|d_k\| \|\bar{x}_k - x_k\|^{\frac{1}{\gamma} - 1})\} \\
\quad \leq \max\{O(\|\bar{x}_k - x_k\|^{2\delta(1+v) - \frac{1}{\gamma}}), O(\|\bar{x}_k - x_k\|^{\delta(2(1+v) - \frac{4}{\gamma} + 1)})\} \\
\quad \leq O(\|\bar{x}_k - x_k\|^{2\delta(1+v) - \frac{2}{\gamma}}) \to 0,
\]

(64)

\[
\frac{\|F(y_k)\|O(\|d_k\|^{1+v})}{\text{Pred}_k} \leq \frac{\|F(y_k)\|O(\|d_k\|^{1+v})}{\text{Pred}_k} \leq \|F(y_k)\| \min\{O(\|\bar{y}_k - y_k\|^{\frac{v}{\gamma}}), O(\|d_k\| \|\bar{y}_k - y_k\|^{\frac{1}{\gamma} - 1})\} \\
\quad \leq \max\{O(\|\bar{y}_k - y_k\|^{\delta(1+v) - \frac{1}{\gamma}}), O(\|\bar{y}_k - y_k\|^{\delta(1+v) - \frac{4}{\gamma} + 1})\} \\
\quad \leq O(\|\bar{y}_k - y_k\|^{\delta(1+v) - \frac{2}{\gamma}}) \to 0,
\]

(65)

\[
\frac{O(\|d_{k}^{2+2v}\|)}{\text{Pred}_k} \leq \frac{O(\|d_{k}^{2+2v}\|)}{\text{Pred}_k} \leq \|F(y_k)\| \min\{O(\|\bar{y}_k - y_k\|^{\frac{v}{\gamma}}), O(\|d_k\| \|\bar{y}_k - y_k\|^{\frac{1}{\gamma} - 1})\} \\
\quad \leq \max\{O(\|\bar{y}_k - y_k\|^{2\delta(1+v) - \frac{2}{\gamma}}), O(\|\bar{y}_k - y_k\|^{\delta(2(1+v) - \frac{4}{\gamma} + 1)})\} \\
\quad \leq O(\|\bar{y}_k - y_k\|^{2\delta(1+v) - \frac{3}{\gamma}}) \to 0,
\]

(66)

Hence, by (63)–(66),

\[
\frac{\|F(y_k)\|O(\|d_{k}^{1+v}\|)}{\text{Pred}_k} \leq \frac{\|F_k + J_k d_k\| + O(\|d_{k}^{1+v}\|)}{\text{Pred}_k} O(\|d_{k}^{1+v}\|) \\
\quad \leq \frac{\|F_k\|O(\|d_{k}^{1+v}\|)}{\text{Pred}_k} + \frac{O(\|d_{k}^{2+2v}\|)}{\text{Pred}_k} \\
\quad \to 0,
\]

(67)

\[
\frac{\|F(y_k)\|O(\|d_{k}^{n}\|^{\|d_{k}\|}}{\text{Pred}_k} = \left(\frac{\|F(y_k)\|O(\|d_{k}^{1+v}\|)}{\text{Pred}_k}\right)^{\frac{1}{1+n}} \left(\frac{\|F(y_k)\|O(\|d_{k}^{1+v}\|)}{\text{Pred}_k}\right)^{\frac{1}{1+n}} \\
\quad \to 0,
\]

(68)
Combining (46), (47), (48), (61), (62), (76) with 
\[ \delta \gamma v \]
the proof is completed. 

Thus, Theorem 3.3.

Under Assumption 3.1, if the Hölderian continuity of the Jacobian and the Hölderian error bound condition.

and

So

which yields

The above inequalities give

\[ |r_k - 1| \to 0. \]

So \( r_k \to 1 \). This implies that there exists a positive constant \( \bar{\mu} < M \) such that

\[ \mu_k \geq \bar{\mu}. \]

The proof is completed.

Based on the above result, we derive the convergence order of Algorithm 2.1 under the Hölderian continuity of the Jacobian and the Hölderian error bound condition.

Theorem 3.3. Under Assumption 3.1, if \( \frac{1}{\gamma(1+v)} < \delta < \gamma \), then the convergence order of Algorithm 2.1 is \( \delta \gamma v + \delta^2 \gamma^2 (1 + v) \).

Proof. By (46) and Lemma 3.2, we have

\[ \Delta_k = \mu_k ||F_k||^\delta \geq \bar{\mu} \kappa_{hleb} ||x_k - x||^{\frac{1}{v}} \]

and

\[ \bar{\Delta}_k = \mu_k ||F(y_k)||^\delta \geq \bar{\mu} \kappa_{hleb} ||y_k - y||^{\frac{1}{v}}. \]

Thus, \( \bar{x}_k - x_k \) and \( \bar{y}_k - y_k \) are feasible solutions of (2) and (7), respectively. By (46) and (48),

\[ \kappa_{hleb} ||y_k - y||^{\frac{1}{v}} \leq ||F(x_k + d_k)|| \]

\[ \leq ||F_k + J_k d_k|| + O(||d_k||^{1+v}) \]

\[ \leq ||F_k + J_k (\bar{x}_k - x_k)|| + O(\Delta_k^{1+v}) \]

\[ \leq O(||\bar{x}_k - x_k||^{1+v}) + O(||\bar{x}_k - x_k||^{\delta(1+v)}) \]

\[ \leq O(||\bar{x}_k - x_k||^{\delta(1+v)}), \]

which yields

\[ ||\bar{y}_k - y_k|| \leq O(||\bar{x}_k - x_k||^{\delta(1+v)}). \]

Combining (46), (47), (48), (61), (62), (76) with \( \frac{1}{\gamma(1+v)} < \delta \), we have

\[ \kappa_{hleb} ||\bar{x}_{k+1} - x_{k+1}||^{\frac{1}{v}} \]

\[ \leq ||F(x_k + d_k + \bar{d}_k)|| \]

\[ \leq ||F(y_k) + J_k \bar{d}_k|| + ||J(y_k) - J_k|| ||\bar{d}_k|| + O(||\bar{d}_k||^{1+v}) \]

\[ \leq ||F(y_k) + J_k \bar{d}_k|| + ||J(y_k) - J_k|| ||\bar{d}_k|| + O(||\bar{d}_k||^{1+v}) \]

\[ \leq ||F(y_k) + J_k (\bar{y}_k - y_k)|| + ||J(y_k) - J_k|| ||\bar{y}_k - y_k|| \]
Therefore,
\[
\|\bar{x}_{k+1} - x_{k+1}\| \leq O(\|\bar{x}_k - x_k\|^{\delta \gamma v + \delta^2 \gamma^2(1+v)}).
\]
(78)

So, the sequence \(\{x_k\}\) converges to the solution set \(X^*\) of (1) with order \(\delta \gamma v + \delta^2 \gamma^2(1+v)\). \(\square\)

From Theorem 3.3, we can see that when \(\gamma = v = 1\), \(\{x_k\}\) converges to the solution set with order \(2\delta^2 + \delta\), which is exactly the result given in [5] under the Lipschitz continuity of the Jacobian and the local error bound condition.

4. Conclusions. We studied the convergence properties of the modified trust region method under the Hölderian continuity of the Jacobian and the Hölderian error bound condition, which are more general than the Lipschitz continuity of the Jacobian and the local error bound condition. The results obtained in this paper include those given in [5] as special cases.

REFERENCES

[1] M. Ahookhosh, F. J. Aragón, R. M. T. Fleming and Phan T. Vuong, Local convergence of Levenberg-Marquardt methods under Hölderian metric subregularity, Adv. Comput. Math., 45 (2019), 2771–2806.
[2] J. Fan, Convergence rate of the trust region method for nonlinear equations under local error bound condition, Comput. Optim. Appl., 34 (2006), 215–227.
[3] J. Fan, The modified Levenberg-Marquardt method for nonlinear equations with cubic convergence, Math. Comp., 81 (2012), 447–466.
[4] J. Fan, J. Huang and J. Pan, An adaptive multi-step Levenberg-Marquardt method, J. Sci. Comput., 78 (2019), 531–548.
[5] J. Fan and N. Lu, On the modified trust region algorithm for nonlinear equations, Optim. Methods Softw., 30 (2015), 478–491.
[6] J. Fan and Y. Yuan, On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption, Computing, 74 (2005), 23–39.
[7] C. T. Kelley, Solving Nonlinear Equations with Newton’s Method, SIAM, Philadelphia, 2003.
[8] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, Quart. Appl. Math., 2 (1944), 164–168.
[9] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear inequalities, J. Soc. Indust. Appl. Math., 11 (1963), 431–441.
[10] J. J. Moré, The Levenberg-Marquardt algorithm: Implementation and theory, in G. A. Watson, ed., Lecture Notes in Mathematics 630: Numerical Analysis, Springer-Verlag, Berlin, (1978), 105–116.
[11] M. J. D. Powell, Convergence properties of a class of minimization algorithms, Nonlinear Programming, 2 (1974), 1–27.
[12] H. Wang and J. Fan, Convergence rate of the Levenberg-Marquardt method under Hölderian error bound, Optim. Methods Softw., 35 (2020), 767–786.
[13] H. Wang and J. Fan, Convergence properties of inexact Levenberg-Marquardt method under Hölderian error bound, J. Ind. Manag. Optim., 17 (2021), 2265–2275.
[14] N. Yamashita and M. Fukushima, On the rate of convergence of the Levenberg-Marquardt method, Topics in Numerical Analysis, 15 (2001), 239–249.
[15] Y. Yuan, Trust region algorithms for nonlinear equations, Information, 1 (1998), 7–20.
[16] Y. Yuan, Subspace methods for large scale nonlinear equations and nonlinear least squares, \textit{Optim. Eng.}, 10 (2009), 207–218.

[17] Y. Yuan, Recent advances in numerical methods for nonlinear equations and nonlinear least squares, \textit{Numer. Algebra Control Optim.}, 1 (2011), 15–34.

[18] X. Zhu and G.-H. Lin, Improved convergence results for a modified Levenberg-Marquardt method for nonlinear equations and applications in MPCC, \textit{Optim. Methods Softw.}, 31 (2016), 791–804.

Received May 2021; revised September 2021; early access December 2021.

\textit{E-mail address:} majirui@sjtu.edu.cn

\textit{E-mail address:} jyfan@sjtu.edu.cn