ON THE GLOBAL CONVERGENCE OF A PARAMETER-ADJUSTING LEVENBERG-MARQUARDT METHOD

LIYAN QI
School of Mathematical Sciences
Dalian University of Technology
Dalian 116023, China
and
School of Sciences
Dalian Ocean University
Dalian 116023, China

XIANTAO XIAO AND LIWEI ZHANG
School of Mathematical Sciences
Dalian University of Technology
Dalian 116023, China

ABSTRACT. The Levenberg-Marquardt (LM) method is a classical but popular method for solving nonlinear equations. Based on the trust region technique, we propose a parameter-adjusting LM (PALM) method, in which the LM parameter $\mu_k$ is self-adjusted at each iteration based on the ratio between actual reduction and predicted reduction. Under the level-bounded condition, we prove the global convergence of PALM. We also propose a modified parameter-adjusting LM (MPALM) method. Numerical results show that the two methods are very efficient.

1. Introduction. In this paper, we aim to solve the system of nonlinear equations

$$F(x) = 0,$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuously differentiable. Define

$$f(x) := \frac{1}{2} \|F(x)\|^2,$$

then

$$\nabla f(x) = JF(x)^T F(x).$$

It is obvious that solving system (1) is equivalent to find the stationary point of the unconstrained optimization problem,

$$\text{minimize}_{x \in \mathbb{R}^n} \ f(x).$$

2010 Mathematics Subject Classification. Primary: 90C30, 65K05.

Key words and phrases. Nonlinear equations, Levenberg-Marquardt method, global convergence.

This work is supported by the National Natural Science Foundation of China under projects No. 11071029, 11101064 and 91130007 and the Fundamental Research Funds for the Central Universities under project No. DUT13LK42.
A classical but very popular method for solving nonlinear equations (1) is the Levenberg-Marquardt method [5, 6]. We do not attempt to survey the literature on LM method, which is vast. For the $k$–th iterate point $x_k$, denote

$$f_k = f(x_k), \quad F_k = F(x_k), \quad \nabla f_k = \nabla f(x_k), \quad B_k = JF(x_k)^TJF(x_k).$$

At the $k$–th iteration, the LM method calculates the direction $p_k$ by solving the following linear equations

$$[B_k + \mu_k I]p_k + \nabla f_k = 0,$$  \hspace{1cm} (3)

where $\mu_k > 0$ is the LM parameter which is updated at each iteration. Note that, if $B_k$ is nonsingular and $\mu_k = 0$, the the LM direction $p_k$ is reduced to the traditional Newton step. Thus, the LM method has global convergence and local quadratic convergence rate if the LM parameter $\mu_k$ is chosen suitably and the Jacobian $B_k$ is nonsingular at the solution. However, the nonsingularity condition of the Jacobian is too strong. Yamashita and Fukushima [10] proved that, under the local error bound condition, the LM method maintain the quadratic convergence. Fan [1] proposed a modified Levenberg-Marquardt method, which uses the addition of an LM step and an approximate LM step as the LM direction. At every iteration, the modified LM method first obtains an LM step $d_k$ by solving the linear equations

$$[B_k + \mu_k I]p_k + \nabla f_k = 0,$$

then solves the linear equations

$$[B_k + \mu_k I]p_k + \nabla f(y_k) = 0 \text{ with } y_k = x_k + d_k$$

to obtain the approximate LM step $\hat{d}_k$, and the LM direction is $p_k = d_k + \hat{d}_k$. The author showed that, under the local error bound condition, the modified LM method achieves the cubic convergence. The modified LM method with line search was recently introduced in [2].

This paper focuses on the choice of the LM parameter $\mu_k$. It is well known that, if $\nabla f(x)$ is Lipschitz continuous and nonsingular at the solution, the LM method converges quadratically if $\mu_k$ is chosen as $O(\|\nabla f_k\|)$. Fan and Yuan [4] chose $\mu_k = \|F_k\|^\delta$ with $\delta \in (0, 2]$ and showed that, under the local error bound condition, the LM method converges superlinearly when $\delta \in (0, 1)$ and quadratically when $\delta \in [1, 2]$. Moré [7] reviewed the relationship between Levenberg-Marquardt method and the trust region method. Inspired by the trust region method, we proposed a new LM method, in which the parameter $\mu_k$ is self-adjusted based on the ratio between actual reduction and predicted reduction. A similar parameter self-adjusted LM method was studied in [3]. However, the Lipschitz of $F(x)$ and $JF(x)$ is assumed to obtain the global convergence. In this paper, based on the traditional trust region method we proved that, the parameter-adjusting LM method we construct is convergent globally if $F(x)$ is continuously differentiable and $f(x)$ is level bounded.

The main differences between the modified parameter-adjusting LM method we proposed and the existing Levenberg-Marquardt method can be stated as follows. First, the choice of LM parameter $\mu_k$ is different. In most of the existing LM methods, the parameter $\mu_k$ is chosen as $\|F_k\|^\delta$ with $\delta \in (0, 2]$. However, when the iterative sequence is close to the solution set, the LM parameter $\mu_k = \|F_k\|^\delta$ may be too small to lose its role. In our new LM method, the parameter $\mu_k$ is automatically updated according to the ratio between actual reduction and predicted reduction, which is inspired by the success of traditional trust region method. Second, the assumption to obtain global convergence is relaxed. Fan and Pan [3] assumed that
both $F(x)$ and $JF(x)$ are Lipschitz continuous. We show in this paper that the
global convergence can also be proved without this assumption, which will expand
the applicability of LM method.

The paper is organized as follows. In Section 2, we proposed a parameter-
adjusting LM method and obtain its global convergence under some assumptions.
A modified parameter-adjusting LM method is proposed in Section 3. Finally, some
numerical results are presented in Section 4.

2. A parameter-adjusting LM method. Consider the quadratic model for ap-
proximating $f$ around $x_k$ defined as follows

$$m_k(p) = f_k + \nabla f_k^T p + \frac{1}{2}p^T B_k p.$$  

For $p_k$, which will be calculated in the following method, define the ratio between
actual reduction and predicted reduction

$$\rho_k = \frac{f_k - f(x_k + p_k)}{m_k(0) - m_k(p_k)}. \quad (4)$$

We present the parameter-adjusting Levenberg-Marquardt (PALM) method as
follows.

**Algorithm 1 (PALM).** Given $\mu_0 > 0$, $\eta \in [0, 1/4)$:

for $k = 0, 1, 2, \ldots$

Calculate $p_k = -[B_k + \mu_k I]^{-1}\nabla f_k$;

Evaluate $\rho_k$ from (4);

if $\rho_k < 1/4$

$\mu_{k+1} = 4\mu_k$;

else

if $\rho_k > 3/4$

$\mu_{k+1} = \mu_k/2$;

else

$\mu_{k+1} = \mu_k$;

if $\rho_k > \eta$

$x_{k+1} = x_k + p_k$;

else

$x_{k+1} = x_k$;

end (for).

The following lemma, which gives a lower bound of the predicted reduction, is a
key to the proof of global convergence theorem.

**Lemma 2.1.**

$$m_k(0) - m_k(p_k) \geq \frac{1}{2} \|\nabla f_k\| \min \left[ \|p_k\|, \frac{\|\nabla f_k\|}{\|B_k\|} \right]. \quad (5)$$

**Proof.** Noting that $(p, \lambda) = (p_k, \mu_k)$ satisfies the following relations

$$\nabla f_k + B_k p + \lambda p = 0, \quad 0 \leq \lambda \leq \|p_k\| - \|p\| \geq 0,$$

which is indeed the Karush-Kuhn-Tucker system of the following problem

$$\min m_k(p)$$

s.t. $\|p\|^2 \leq \|p_k\|^2.$  

(6)
Since $m_k(\cdot)$ is a convex function, we have that $p_k$ is the optimal solution to the problem (6).

If we take the Cauchy step defined as in [9, Chapter 4]

$$p_k^c = \tau_k \frac{\nabla f_k}{\| \nabla f_k \|},$$

where

$$\tau_k = \begin{cases} 
1, & \text{if } \nabla f_k^T B_k \nabla f_k = 0, \\
\min \left[ 1, \frac{\| \nabla f_k \|^3}{\| p_k \| \| \nabla f_k \|} \right], & \text{if } \nabla f_k^T B_k \nabla f_k > 0,
\end{cases}$$

then the following inequality follows from [9, Lemma 4.3]

$$m_k(0) - m_k(p_k^c) \geq \frac{1}{2} \| \nabla f_k \| \min \left[ \| p_k \|, \frac{\| \nabla f_k \|}{\| p_k \|} \right].$$

Noticing that $\| p_k^c \|^2 \leq \| p_k \|^2$, one has

$$m_k(0) - m_k(p_k) \geq m_k(0) - m_k(p_k^c),$$

so the inequality (5) is obtained.

The global convergence requires the following assumption.

**Assumption 1.** The mapping $F$ is continuously differentiable and the level set

$$L(x_0) = \{ x \in \mathbb{R}^n : f(x) \leq f(x_0) \}$$

is bounded.

Next, we show that the sequence generated by PALM method with $\eta = 0$ converges to a stationary point of the merit function $f(x)$.

**Theorem 2.2.** Let Assumption 1 be satisfied and $\eta = 0$ in Algorithm 1. Suppose that $\{ x_k \}$ is generated by Algorithm 1, then we have

$$\lim_{k \to \infty} \| \nabla f_k \| = 0. \quad (7)$$

**Proof.** We proceed the proof by contradiction. Suppose that there exist $\varepsilon > 0$ and an infinite set of positive integers $\mathcal{N}$ such that

$$\| \nabla f_k \| \geq \varepsilon, \quad \forall k \in \mathcal{N}. \quad (8)$$

Note that

$$| \rho_k - 1 | = \left| \frac{m_k(p_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)} \right|. \quad (9)$$

From Taylor’s theorem, we have that

$$| m_k(p_k) - f(x_k + p_k) | = \left| \frac{1}{2} p_k^T B_k p_k - \int_0^1 [\nabla f(x_k + tp_k) - \nabla f(x_k)]^T p_k \, dt \right| \leq \frac{1}{2} \beta \| p_k \|^2 + c(x_k, p_k) \| p_k \|,$$

where

$$\beta = \max_{x \in L(x_0)} \| JF(x)^T JF(x) \|_2, \quad \text{(spectral norm)}$$

$$c(x_k, p_k) = \max_{0 \leq t \leq 1} \| \nabla f(x_k + tp_k) - \nabla f(x_k) \|. \quad (9)$$
In view of (5) in Lemma 2.1, we have for $k \in \mathcal{N}$ that
\[ m_k(0) - m_k(p_k) \geq \frac{1}{2} \| \nabla f_k \| \min \left[ \| p_k \|, \frac{\| \nabla f_k \|}{\| B_k \|} \right] \geq \frac{1}{2} \varepsilon \min \left[ \| p_k \|, \varepsilon \beta \right]. \tag{10} \]
Using (9) and (10), we obtain for $k \in \mathcal{N}$ that
\[ |\rho_k - 1| \leq \frac{\| p_k \| [\beta \| p_k \| + 2c(x_k, p_k)]}{\varepsilon \min \left[ \| p_k \|, \varepsilon \beta \right]} \tag{11} \]
Let $\overline{\mu} > \mu_N$, where $\mathcal{N}$ is the first index in $\mathcal{N}$, be sufficiently large such that when $\mu_k \geq \overline{\mu}$, $k \in \mathcal{N}$, $p_k = -[B_k + \mu_k I]^{-1} \nabla f_k$ is small enough satisfying
\[ \| p_k \| \leq \frac{\varepsilon}{8\beta}, c(x_k, p_k) \leq \frac{\varepsilon}{16}, \]
then
\[ \beta \| p_k \| + 2c(x_k, p_k) \leq \frac{\varepsilon}{4}. \]
We have from (11) for $k \in \mathcal{N}$ with $p_k$ small enough that
\[ |\rho_k - 1| \leq \frac{\frac{\varepsilon}{4} \| p_k \|}{\| p_k \|} = \frac{1}{4}. \]
Therefore, $\rho_k \geq 3/4$ for $k \in \mathcal{N}$ with $p_k$ small enough, and by Algorithm 1, we have $\mu_{k+1} \leq \mu_k$ when $\mu_k \geq \overline{\mu}$.

It follows that increase of $\mu_k$ (by a factor 4) can only occur if $\mu_k < \overline{\mu}$. Therefore, we conclude that
\[ \mu_k \leq 4 \overline{\mu} \text{ for all } k \in \mathcal{N}. \tag{12} \]
Suppose now that there exists an infinite index set $\mathcal{K} \subset \mathcal{N}$ such that $\rho_k \geq 1/4$ for $k \in \mathcal{K}$. If $k \in \mathcal{K}$, we have from (4) and (5) that
\[ f(x_k) - f(x_{k+1}) \geq \frac{1}{4} [m_k(0) - m_k(p_k)] \geq \frac{1}{8} \varepsilon \min \| p_k \|, \varepsilon / \beta]. \]
Since $f$ is bounded from below, we have
\[ \lim_{k \in \mathcal{K}, k \to \infty} \| p_k \| = 0, \]
which implies $\mu_k \to \infty$, $k \in \mathcal{K}$, $k \to \infty$, this contradicts (12). Hence no such infinite index set $\mathcal{K}$ can exist, and we must have $\rho_k < 1/4$ for all $k \in \mathcal{N}$ large enough. In this case, $\mu_k, k \in \mathcal{N}$ will eventually be enlarged by a factor 4 at every iteration, and we have $\mu_k \to \infty$, which again contradicts (12). Therefore, (8) is incorrect, giving (7).

Finally in this section, we show that the PALM method also globally converges to a stationary point when $\eta \in (0, 1/4)$.

**Theorem 2.3.** Let Assumption 1 be satisfied and $\eta \in (0, 1/4)$ in Algorithm 1. Suppose that $\{x_k\}$ is generated by Algorithm 1, then we have
\[ \lim_{k \to \infty} \| \nabla f_k \| = 0. \tag{13} \]
Proof. We consider a particular positive index \( m \) with \( \nabla f_m \neq 0 \). Define the scalar 
\[ \varepsilon := \| \nabla f_m \| > 0. \]
If the entire sequence \( \{ x_k \}_{k \geq m} \) satisfies \( \| \nabla f_k \| \geq \varepsilon \), by using the reasoning in the proof of Theorem 2.2, we can show that this scenario does not occur. Therefore, we can choose the index \( l \geq m \) such that \( x_{l+1} \) is the first iterate after \( x_m \) with \( \| \nabla f_{l+1} \| < \varepsilon \). From Assumption 1, \( \nabla f(x) \) is continuous, then there exists a positive scalar \( R > 0 \) such that 
\[ \| x_{l+1} - x_m \| \geq R. \]
From the framework of Algorithm 1, we can write 
\[ x_{l+1} - x_m = \sum_{k=m}^{l} x_{k+1} - x_k = \sum_{k=m, x_k \neq x_{k+1}}^{l} p_k, \]
where we have limited the sum to the iterations \( k \) for which \( x_k \neq x_{k+1} \), that is, those iterations on which a step was actually taken. Note that those iterations occur only under the condition \( \rho_k \geq \eta \), by using (10) we have
\[ f(x_m) - f(x_{l+1}) = \sum_{k=m, x_k \neq x_{k+1}}^{l} f(x_k) - f(x_{k+1}) \]
\[ \geq \sum_{k=m, x_k \neq x_{k+1}}^{l} \eta [ m_k(0) - m_k(p_k) ] \]
\[ \geq \sum_{k=m, x_k \neq x_{k+1}}^{l} \frac{1}{2} \eta \varepsilon \min \left[ \| p_k \|, \frac{\varepsilon}{\beta} \right]. \]
If \( \| p_k \| \leq \varepsilon / \beta \) for all \( k = m, m + 1, \ldots, l \), we have
\[ f(x_m) - f(x_{l+1}) \geq \frac{1}{2} \eta \varepsilon \sum_{k=m}^{l} \| p_k \| \geq \frac{1}{2} \eta \varepsilon R. \quad (14) \]
Otherwise, we have \( \| p_k \| > \varepsilon / \beta \) for some \( k = m, m + 1, \ldots, l \), and so 
\[ f(x_m) - f(x_{l+1}) \geq \frac{1}{2} \eta \varepsilon \frac{\varepsilon}{\beta}. \quad (15) \]
Since the sequence \( \{ f(x_k) \}_{k=0}^{\infty} \) is decreasing and bounded below, we have that 
\[ f(x_k) \downarrow f^* \]
for some \( f^* > -\infty \). Therefore, using (14) and (15), we can see 
\[ f(x_m) - f^* \geq f(x_m) - f(x_{l+1}) \geq \frac{1}{2} \eta \varepsilon \min \left[ R, \frac{\varepsilon}{\beta} \right] \]
\[ = \frac{1}{2} \eta \| \nabla f_m \| \min \left[ R, \frac{\| \nabla f_m \|}{\beta} \right] > 0, \]
which implies 
\[ \| \nabla f_m \| \to 0. \]
\[ \square \]
3. A modified parameter-adjusting LM method. In this section, we proposed a modified parameter-adjusting LM (MPALM) method. One and only difference between PALM and MPALM is that the calculation of LM direction in MPALM is altered to

\[ p_k = -[B_k + \mu_k\|F_k\|I]^{-1}\nabla f_k. \]

The proof of global convergence for MPALM is almost the same as PALM.

**Algorithm 2 (MPALM).** Given \( \mu_0 > 0, \eta \in [0, 1/4) \):

```plaintext
for \( k = 0, 1, 2, \ldots \)
    Calculate \( p_k = -[B_k + \mu_k\|F_k\|I]^{-1}\nabla f_k; \)
    Evaluate \( \rho_k \) from (4);
    if \( \rho_k < 1/4 \)
        \( \mu_{k+1} = 4\mu_k; \)
    else
        if \( \rho > 3/4 \)
            \( \mu_{k+1} = \mu_k/2; \)
        else
            \( \mu_{k+1} = \mu_k; \)
    if \( \rho_k > \eta \)
        \( x_{k+1} = x_k + p_k; \)
    else
        \( x_{k+1} = x_k; \)
end (for).
```

The following key lemma can be proved as Lemma 2.1.

**Lemma 3.1.**

\[
m_k(0) - m_k(p_k) \geq \frac{1}{2}\|\nabla f_k\| \min \left[ \|p_k\|, \frac{\|\nabla f_k\|}{\|B_k\|} \right]. \tag{16}\]

**Theorem 3.2.** Let \( \eta = 0 \) be chosen in Algorithm 2 and Assumption 1 be satisfied. Let \( \{x_k\} \) be generated by Algorithm 2. Then we have

\[
\lim_{k \to \infty} \|\nabla f_k\| = 0. \tag{17}\]

**Proof.** The proof is very similar to Theorem 2.2. Define

\[
\beta = \max_{x \in L(x_0)} \|JF(x)\|^T JF(x)\|_2, \text{ (spectral norm)}
\]

\[
\beta_1 = \max_{x \in L(x_0)} \|JF(x)\|_2, \text{ (spectral norm)}
\]

\[
c(x_k, p_k) = \max_{0 \leq t \leq 1} \|\nabla f(x_k + tp_k) - \nabla f(x_k)\|.
\]

We proceed the proof by contradiction. Suppose that there exists \( \varepsilon > 0 \) and an infinite set of positive integers \( N \) such that

\[
\|\nabla f_k\| \geq \varepsilon \quad \forall k \in N. \tag{18}\]

As in the proof of Theorem 2.2, by using Lemma 3.1 we obtain for \( k \in N \) that

\[
|\rho_k - 1| \leq \frac{\|p_k\|}{\|p_k\|, \|\nabla f_k\|} + 2c(x_k, p_k).
\]

Since \( \|F_k\| \geq \|\nabla f_k\|/\beta_1 \), so we have

\[
\lambda_{\min}(B_k + \mu_k\|F_k\|I) \geq \mu_k\|F_k\| \geq \mu_k\varepsilon/\beta_1,
\]
where \( \lambda_{\text{min}}(\cdot) \) denotes the minimal eigenvalue of a matrix. Thus
\[
\| (B_k + \mu_k\|F_k\|I)^{-1} \|_2 = [\lambda_{\text{min}}(B_k + \mu_k\|F_k\|I)]^{-1} \leq [\mu_k\varepsilon/\beta_1]^{-1}.
\]
Recall that \( p_k = -[B_k + \mu_k\|F_k\|I]^{-1}\nabla f_k \), we obtain
\[
\| p_k \| \leq [\mu_k\varepsilon/\beta_1]^{-1}\| \nabla f_k \| = \beta_1\| \nabla f_k \|/[\mu_k\varepsilon].
\]
Let \( \bar{\mu} > \mu_N \), where \( N \) is the first index in \( \mathcal{N} \) be sufficiently large such that when \( \mu_k \geq \bar{\mu}, k \in \mathcal{N} \), \( p_k \) is small enough satisfying
\[
\| p_k \| \leq \frac{\varepsilon}{8\beta}, \quad c(x_k, p_k) \leq \frac{\varepsilon}{16},
\]
then
\[
\beta\| p_k \| + 2c(x_k, p_k) \leq \frac{\varepsilon}{4}.
\]
We have from (19) for \( k \in \mathcal{N} \) with \( p_k \) small enough that
\[
|\rho_k - 1| \leq \frac{1}{4}.
\]
The remainder is similar to the proof of Theorem 2.2.

**Theorem 3.3.** Let Assumption 1 be satisfied and \( \eta \in (0, \frac{1}{4}) \) in Algorithm 2. Suppose that \( \{x_k\} \) is generated by Algorithm 2, then we have
\[
\lim_{k \to \infty} \| \nabla f_k \| = 0.
\]

4. **Numerical Results.** In our numerical experiments, we would like to illustrate the behavior of PALM and MPALM on the following test functions that are obtained from \([8]\). These functions are defined in the following general format:
(a) Dimensions;
(b) Function definition;
(c) Initial point;
(d) Solution.

**Problem 1** (Rosenbrock function).  
(a) \( n = 2, m = 2 \)
(b) \( F_1(x) = 10(x(2) - x(1)^2), F_2(x) = 1 - x(1) \)
(c) \( x_0 = (-1.2, 1) \)
(d) \( x^* = (1, 1) \)

**Problem 2** (Helical valley function).  
(a) \( n = 3, m = 3 \)
(b)
\[
F(x_1, x_2, x_3) = \begin{bmatrix}
10|x_3 - 10\varphi(x_1, x_2)| \\
10\sqrt{x_1^2 + x_2^2 - 1}
\end{bmatrix},
\]
where
\[
\varphi(x_1, x_2) = \begin{cases}
\frac{1}{\pi} \arctan(\frac{x_2}{x_1}), & x_1 > 0 \\
\frac{1}{\pi} \arctan(\frac{x_2}{x_1}) + 0.5, & x_1 < 0
\end{cases}
\]
(c) \( x_0 = (-1, 0, 0) \)
(d) \( x^* = (1, 0, 0) \)

**Problem 3** (Powell badly scaled function).  
(a) \( n = 2, m = 2 \)
(b) \( F_1(x) = 10^4x(1)x(2) - 1, F_2(x) = \exp(-x(1)) + \exp(-x(2)) - 1.0001 \)
(c) \( x_0 = (0, 0.1) \)
(d) \( x^* = (1.098 \cdot 10^{-5}, 9.106) \)
Problem 4 (Powell singular function).  
(a) $n = 4, m = 4$
(b) $F_1(x) = x(1) + 10x(2), \ F_2(x) = \sqrt{5}(3x - x(4))$,  
$F_3(x) = (2x - 2x(3))^2$, $F_4(x) = \sqrt{10}(1 - x(4))^2$
(c) $x_0 = (3, -1, 0, 1)$
(d) $x^* = 0$

Problem 5 (Wood function).  
(a) $n = 4, m = 6$
(b) $F_1(x) = 10x(1) - x(1)^2, \ F_2(x) = 1 - x(1)$,  
$F_3(x) = \sqrt{90}(x(4) - x(3)^2), \ F_4(x) = 1 - x(3)$,  
$F_5(x) = \sqrt{10}(x(2) + x(4) - 2), F_6(x) = (x(2) - x(4))/\sqrt{10}$
(c) $x_0 = (-3, -1, -3, -1)$
(d) $x^* = (1, 1, 1, 1)$

Problem 6 (Trigonometric function).  
(a) $n = 100, m = 100$
(b) $F_k(x) = n + k(1 - \cos(x(k))) - \sin(x(k)) - \sum_{j=1}^n \cos(x(j))$
(c) $x_0 = (1/n, \ldots, 1/n)$
(d) $x^* = 0$

Problem 7 (Brown almost-linear function).  
(a) $n = 200, m = 200$
(b) $F_k(x) = x(k) - (n + 1) + \sum_{j=1}^n x(j), k = 1, \ldots, (m - 1)$,  
$F_m(x) = (\prod_{j=1}^n x(j)) - 1$
(c) $x_0 = (0.005, 0.005, \ldots, 0.005)$
(d) $x^* = (\alpha, \ldots, \alpha, \alpha^{1-n})$, where $\alpha$ satisfies  
$$n \alpha^n - (n + 1)\alpha^{n-1} + 1 = 0.$$  

Problem 8 (Discrete boundary value function).  
(a) $n = 10, m = 10$
(b) $F_k(x) = 2x(k) - x(k - 1) - x(k + 1) + h^2(x(k) + t(k + 1))$ \[ -\frac{1}{2} \]
where  
$$h = 1/(n + 1), \ t(k) = kh, \ x(0) = x(n + 1) = 0$$
(c) $x_0 = (\xi_k)$ with $\xi_k = t(k)(t(k) - 1)$
(d) $x^* = -10^{-4} \cdot (432, 816, 1145, 1410, 1599, 1699, 1691, 1552, 1254, 754)$

Problem 9 (Discrete integral equation function).  
(a) $n = 10, m = 10$
(b)  
$$F_k(x) = x(k) + 0.5h \left[ (1 - t(k))(h) \sum_{j=1}^k t(j)(x(j) + t(j) + 1)^3 
+ t(k) \sum_{j=k+1}^n (1 - t(j))(x(j) + t(j) + 1)^3 \right]$$
where  
$$h = 1/(n + 1), \ t(k) = kh, \ x(0) = x(n + 1) = 0$$
(c) $x_0 = (\xi_k)$ with $\xi_k = t(k)(t(k) - 1)$
(d) $x^* = -10^{-4} \cdot (432, 816, 1145, 1410, 1599, 1699, 1691, 1552, 1254, 754)$

Problem 10 (Broyden tridiagonal function).  
(a) $n = 20, m = 20$
(b) $F_k(x) = (3 - 2x(k))x(k) - x(k - 1) - 2x(k + 1) + 1$ with  
$$x(0) = x(n + 1) = 0$$
(c) $x_0 = (-1, -1, \ldots, -1)$
(d)  
$$x^* = -10^{-4} \cdot (5708, 6819, 7025, 7063, 7071, 7071, 7071, 7071, 7071, 7071, 7071, 7070, 7068, 7064, 7051, 7015, 6919, 6658, 5960, 4164)$$
We set \( \mu_0 = 10^{-3}\|F_0\| \), \( \eta = 0 \). The algorithms are terminated when the norm of derivative of \( f \) at \( x_k \) is less than \( 10^{-5} \), i.e., \( \|\nabla f_k\| \leq 10^{-5} \), or the number of the iterations exceeds \( 100 \cdot n \). The numerical experiments were carried out with MATLAB 2014b on a laptop with 2.3GHz CPU and 4GB memory. The numerical results for PALM and MPALM on those ten numerical problems are presented in Table 1. The second column \((n,m)\) of the table denotes the dimensions of the problems. The third column shows that the initial point is \( x_0, 10x_0 \) or \( 100x_0 \), where \( x_0 \) is defined in the problems. “NF”, “NJ” and “IT” stand for the number of function calculations, Jacobian calculations and the number of iterations, respectively. Note that, for general nonlinear equations, the number of calculations of a Jacobian is usually \( n \) times of that of a function. The sign “-” denotes that the iterations had underflows or overflows.

**Table 1. Numerical results for PALM and MPALM**

| Problem (n,m) | \( x_0 \) | NF | NJ | IT | NF | NJ | IT |
|--------------|-----------|----|----|----|----|----|----|
| 1 (2,2)      | 1         | 43 | 19 | 24 | 41 | 18 | 23 |
|              | 10        | 56 | 27 | 29 | 60 | 30 | 34 |
|              | 100       | 104| 51 | 53 | 117| 56 | 61 |
| 2 (3,3)      | 1         | 16 | 8  | 8  | 17 | 9  | 8  |
|              | 10        | 17 | 9  | 8  | 28 | 12 | 16 |
|              | 100       | 44 | 20 | 24 | 32 | 16 | 16 |
| 3 (2,2)      | 1         | 204| 95 | 109| 221| 96 | 125|
|              | 10        | 174| 84 | 90 | 148| 71 | 77 |
|              | 100       | -  | -  | -  | 85 | 40 | 45 |
| 4 (4,4)      | 1         | 19 | 10 | 9  | 19 | 10 | 9  |
|              | 10        | 25 | 13 | 12 | 27 | 14 | 13 |
|              | 100       | 37 | 19 | 18 | 49 | 25 | 24 |
| 5 (4,6)      | 1         | 159| 67 | 92 | 154| 65 | 89 |
|              | 10        | 167| 71 | 96 | 54 | 25 | 29 |
|              | 100       | 92 | 44 | 48 | 176| 85 | 91 |
| 6 (100,100)  | 1         | 61 | 26 | 35 | 33 | 12 | 21 |
|              | 10        | 55 | 25 | 30 | 61 | 27 | 34 |
|              | 100       | 43 | 21 | 22 | 57 | 25 | 32 |
| 7 (200,200)  | 1         | 7  | 4  | 3  | 9  | 5  | 4  |
|              | 10        | 7  | 4  | 3  | 9  | 5  | 4  |
|              | 100       | 7  | 4  | 3  | 11 | 6  | 5  |
| 8 (10,10)    | 1         | 5  | 3  | 2  | 5  | 3  | 2  |
|              | 10        | 7  | 4  | 3  | 7  | 4  | 3  |
|              | 100       | 19 | 10 | 9  | 19 | 10 | 9  |
| 9 (10,10)    | 1         | 5  | 3  | 2  | 5  | 3  | 2  |
|              | 10        | 7  | 4  | 3  | 7  | 4  | 3  |
|              | 100       | 19 | 10 | 9  | 19 | 10 | 9  |
| 10 (20,20)   | 1         | 9  | 5  | 4  | 9  | 5  | 4  |
|              | 10        | 17 | 9  | 8  | 17 | 9  | 8  |
|              | 100       | 25 | 13 | 12 | 35 | 18 | 17 |

From Table 1, we can see that the performances of PALM and MPALM are almost in the same level. For every numerical problems, the two methods can
always obtain the solutions with different initial points. If we change the stopping
criterion to \( f(x_k) = \frac{1}{2} \| F(x_k) \|^2 \leq 10^{-5} \), the number of iterations will be fewer for
most of the problems.

5. Conclusions. Inspired by the traditional trust region method, we proposed a
parameter-adjusting LM (PALM) method. The main difference between classical
LM method and PALM is that, the LM parameter \( \mu_k \) in PALM is self-adjusted
at each iteration according to the ratio between actual reduction and predicted
reduction. An analogous method can be found in the literature. However, the
global convergence of PALM is proved under a weaker condition.

Next, we will consider the local quadratic convergence of PALM. In the literature,
a local error bound condition is necessary to prove the local quadratic convergence.
A natural question is that, is it possible to relax the local error bound condition?
Another question is, how to accelerate the PALM method to a higher order conver-
gence rate. The papers [1, 2] provide some insights for finding the answers.

Another very important question is about solving nonsmooth (or semismooth)
nonlinear equations by Levernberg-Marquardt method. In the literature, almost all
of the references on LM method focus on solving smooth nonlinear equations. It
is well known that the nonsmooth nonlinear equations have various applications.
Therefore, it is of significant importance for us to study the LM method for non-
smooth equations. Based on the existing results of nonsmooth Newton method, it
is very likely to construct a LM method for solving nonsmooth equations and prove
the corresponding convergence results.

Acknowledgments. We would like to thank the referees for the helpful comments
and suggestions.

REFERENCES

[1] J. Fan, The modified Levenberg-Marquardt method for nonlinear equations with cubic con-
vergence, Math. Comp., 81 (2012), 447–466.
[2] J. Fan, Accelerating the modified Levenberg-Marquardt method for nonlinear equations,
Math. Comp., 83 (2014), 1173–1187.
[3] J. Fan and J. Pan, Convergence properties of a self-adaptive Levenberg-Marquardt algorithm
under local error bound condition, Comput. Optim. Appl., 34 (2006), 47–62.
[4] J. Fan and Y. Yuan, On the quadratic convergence of the Levenberg-Marquardt method
without nonsingularity assumption, Computing, 74 (2005), 23–39.
[5] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, Quart.
Appl. Math., 2 (1944), 164–166.
[6] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear inequalities, SIAM
J. Appl. Math., 11 (1963), 431–441.
[7] J. J. Moré, Recent developments in algorithms and software for trust region methods, in
Mathematical Programming: the state of the art (Bonn, 1982), Springer, Berlin, (1983),
258–287.
[8] J. J. Moré, B. S. Garbow and K. E. Hillstrom, Testing unconstrained optimization software,
ACM Trans. Math. Software, 7 (1981), 17–41.
[9] J. Nocedal and S. J. Wright, Numerical optimization, 2nd edition, Springer Series in Operations
Research and Financial Engineering, Springer, New York, 2006.
[10] N. Yamashita and M. Fukushima, On the rate of convergence of the Levenberg-Marquardt method, in Topics in numerical analysis, Comput. Suppl., Springer, Vienna, 15 (2001), 239–249.

Received January 2015; 1st revision February 2015; final revision March 2015.

E-mail address: qiliyan@dlou.edu.cn
E-mail address: xtxiao@dlut.edu.cn
E-mail address: lwzhang@dlut.edu.cn