An efficient $m$–step Levenberg–Marquardt method for nonlinear equations

Liang Chen\textsuperscript{a,}\textsuperscript{*}, Yanfang Ma\textsuperscript{b}, Chencheng Su\textsuperscript{a}

\textsuperscript{a}School of Mathematical Sciences, Huaibei Normal University, Huaibei, Anhui 235000, PR China
\textsuperscript{b}School of Computer Science and Technology, Huaibei Normal University, Huaibei, Anhui 235000, PR China

Abstract

In this paper, we construct and analyze an efficient $m$–step Levenberg–Marquardt method for nonlinear equations. The main advantage of this method is that the $m$–step LM method could save more Jacobian calculations with frozen $(J_k^T J_k + \lambda_k I)^{-1} J_k^T$ at every iteration. Under the local error bound condition which is weaker than nonsingularity, the $m$–step LM method has been proved to have $(m + 1)$th convergence order. The global convergence has also been given by trust region technique. Numerical results show that the $m$–step LM method is efficient and could save many calculations of the Jacobian especially for large scale problems.

Keywords: Unconstrained optimization, Trust region, Systems of nonlinear equations, Levenberg-Marquardt method, Local error bound

2010 MSC: 65K05, 90C30

1. Introduction

It’s a well-known problem in science and engineering that is to find the solutions of nonlinear equations

$$F(x) = 0,$$  \hspace{1cm} (1)

where $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuously differentiable function. Due to the nonlinearity of $F(x)$, (1) may have no solutions. Throughout the paper, we let that the solution set of (1) is nonempty and denote it by $X^*$, and in all cases $\|\cdot\|$ refers to the 2-norm.

There are many numerical methods to approximate the solutions of (1) because the exact solutions is difficult to find. A classical numerical method is Newton method which computes the trial step

$$d_k^N = -J_k^{-1} F_k$$

at every iteration, where $F_k = F(x_k)$ and $J_k = F'(x_k)$ is the Jacobian. And the Newton method has quadratic rate of convergence under the condition that $J(x)$ is Lipschitz continuous and nonsingular at the solution of (1). However, the Newton method will be failed when $J_k$ is singular or near singular. To overcome these disadvantages, a large number of researchers have presented many modifications of Newton method \cite{1, 2, 3}. One of them is the Levenberg–Marquardt method (LM) \cite{4, 5}, which is a famous numerical method with computing the trail step

$$d_k^{LM} = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k$$  \hspace{1cm} (2)

at every iteration, where $\lambda_k \geq 0$ is the LM parameter. It is well-known that the Levenberg–Marquardt has quadratic convergence as the Newton method if the Jacobian matrix is nonsingular and Lipschitz continuous at the solution.

However, the condition of nonsingularity is too strong. Yamashita and Fukushima \cite{6} proposed a local error bound condition which is weaker than nonsingularity. Under the local error bound condition with choosing the LM parameter
\( \lambda_k = \| F_k \|^2 \), the Levenberg–Marquardt method also has quadratic convergence. Fan and Yuan [7] obtained a similar result when LM parameter \( \lambda_k = \| F_k \| \). Furthermore, Fan and Pan [8] chose \( \lambda_k = \| F_k \|^\delta \) with \( \delta \in (0, 2] \) and showed that, under the same conditions, the Levenberg–Marquardt method has quadratic convergence when \( \delta \in [1, 2] \) and superlinear convergence when \( \delta \in (0, 1) \) respectively. Some others gave more general choices of the LM parameter [9, 10, 11].

As we all known, the cost of Jacobian computations is expensive when \( F(x) \) is complicated or \( n \) is quite large. Recently, to save Jacobian calculations and achieve a fast convergence rate, Fan [12] presented a modified Levenberg–Marquardt method (MLM) with cubic convergence. At every iteration, the MLM method solves not only the linear equations

\[
( J_k^T J_k + \lambda_k I ) d = - J_k^T F_k
\]
to obtain the LM step (2), where \( \lambda_k = \mu_k \| F_k \|, \mu_k > 0 \) and \( \delta \in [1, 2] \), but also the linear equations

\[
( J_k^T J_k + \lambda_k I ) d = - J_k^T F(y_k)
\]
to obtain the approximate LM step

\[
d_k^{\text{MLM}} = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k)
\]
with \( y_k = x_k + d_k^{\text{LM}} \), and the trial step is

\[
s_k^{\text{MLM}} = d_k^{\text{LM}} + d_k^{\text{MLM}}.
\]

Fan use \((J_k^T J_k + \lambda_k I)^{-1} J_k^T \) in stead of \((J_k^T J_k + \mu_k I + \|F(y_k)\|^\delta I)^{-1} J_k^T \) in (3), which does not involve the calculation of \( J_k^T y_k \). Since \( J_k \) has been used in (2), the cost of Jacobian calculations will be saved.

Similarly, to save more Jacobian calculations, Yang [13] presented a high-order Levenberg–Marquardt method (HLM) with biquadratic convergence by solving another linear equations

\[
( J_k^T J_k + \lambda_k I ) d = - J_k^T F(z_k)
\]
to obtain another approximate LM step

\[
d_k^{\text{HLM}} = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(z_k)
\]
with \( z_k = y_k + d_k^{\text{LM}} \). Yang still use \((J_k^T J_k + \lambda_k I)^{-1} J_k^T \) in stead of \((J_k^T J_k + \mu_k I + \|F(z_k)\|^\delta I)^{-1} J_k^T \) in (4), which does not need to compute \( J_k^T z_k \). The trial step of the HLM method is

\[
s_k^{\text{HLM}} = d_k^{\text{LM}} + d_k^{\text{MLM}} + d_k^{\text{HLM}}.
\]

Furthermore, some other literatures give more general cases [14, 15, 16].

If we consider the MLM method as two-step Levenberg–Marquardt method and the HLM method as three-step Levenberg–Marquardt method respectively, then, it is easy to see that \((J_k^T J_k + \lambda_k I)^{-1} J_k^T \) is computed in all of the first LM step (2), the second LM step (3) and the third LM step (4). So, we can consider \((J_k^T J_k + \lambda_k I)^{-1} J_k^T \) is frozen in the two-step LM method and three-step LM method.

Now, in order to save more Jacobian calculations and achieve a faster convergence rate, we will compute many approximate LM steps

\[
d_k^{(i)} = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(x_k^{(i)}), \quad i = 0, \ldots, m - 1 \quad \text{and} \quad m \geq 1
\]
with frozen \((J_k^T J_k + \lambda_k I)^{-1} J_k^T \), where \( x_k^{(0)} = x_k, d_k^{(0)} = d_k^{\text{LM}}, x_k^{(i+1)} = x_k^{(i)} + d_k^{(i)} \) and \( x_k^{(m)} \). And the trial step of the new LM method will be computed by

\[
s_k = d_k^{(0)} + \cdots + d_k^{(m-1)}.
\]
In every iteration, we only need to compute the Jacobian matrix \( J_k \) once since the matrix in the \( m \) steps is the same. It is quite clear that the above new LM method will reduce to the Levenberg–Marquardt method while \( m = 1 \), the two-step LM method while \( m = 2 \) and the three-step LM method while \( m = 3 \) respectively. We wonder whether the new LM method could save more Jacobian calculations and whether it could achieve a faster convergence rate under
the local error bound condition. So, we will analyze the convergence properties of the above method and do some numerical experiments to test its efficiency in this paper.

We will organize this paper as follow: In Section 2, we first give a new modified Levenberg–Marquardt method which is called $m$–step Levenberg–Marquardt algorithm. In Section 3, we derive the global convergence of the algorithm by using trust region technique. Then we derive the convergence order of the algorithm under the local error bound condition in Section 4. Finally, some numerical results of the new algorithm are given in Section 5.

2. The $m$–step Levenberg–Marquardt algorithm

We first present the $m$–step Levenberg–Marquardt algorithm by using trust region technique, then prove the global convergence in next section.

We take $\Phi(x) = \|F(x)\|^2$ as the merit function for (1). Define the actual reduction of $\Phi(x)$ at the $k$th iteration as

$$\text{Ared}_k = \|F_k\|^2 - \|F(x_k + d_k^0 + \cdots + d_k^{m-1})\|^2,$$

where $d_k^{l(0)}$ are computed by (5). Note that the predicted reduction cannot be defined as usual definition $\|F_k\|^2 - \|F_k + J_k(d_k^{m(0)} + \cdots + d_k^{m-1})\|^2$, because it cannot be proven to be nonnegative, which is required for the global convergence in the trust region method. Hence, similar to [12, 13], we use the modified predicted reduction as follows:

$$\text{Pred}_k = \|F_k\|^2 - \|F_k + J_k d_k^{m(0)}\|^2 + \cdots + \|F(x_k^{(m-1)})\|^2 - \|F(x_k^{(m-1)} + J_k d_k^{m-1})\|^2, \quad m \geq 1. \tag{8}$$

Lemma 2.1. Let the predicted reduction is defined by (8), then

$$\text{Pred}_k \geq \sum_{i=1}^{m} \|J_k^T F(x_k^{(i-1)})\| \min \left\{|d_k^{(i-1)}||, \frac{\|J_k^T F(x_k^{(i-1)})\|}{\|J_k^T J_k\|}\right\} \tag{9}$$

where $m \geq 1$.

Proof. It is easy to see that $d_k^{l(0)}, i = 0, \cdots, m - 1$ is not only the minimizer of the convex minimization problem

$$\min_{d \in \mathbb{R}^n} \|F(x_k^{(0)}) + J_k d\|^2 + \lambda_k \|d\|^2 \triangleq \varphi_{x_k}(d), \tag{10}$$

but also a solution to the trust region problem

$$\min_{d \in \mathbb{R}^n} \|F(x_k^{(0)}) + J_k d\|^2, \quad \text{s.t. } \|d\| \leq \Delta_k, \tag{11}$$

where $\Delta_k = \| - (J_k^T J_k + \lambda_k I)^{-1} J_k^T F(x_k^{(0)})\| = \|d_k^{(0)}\|$. We have

$$\|F(x_k^{(0)})\|^2 - \|F(x_k^{(0)} + J_k d_k^{(0)})\|^2 \geq \|J_k^T F(x_k^{(0)})\| \min \left\{|d_k^{(0)}||, \frac{\|J_k^T F(x_k^{(0)})\|}{\|J_k^T J_k\|}\right\}.$$

Then (9) holds with $i = 0, \cdots, m - 1$.

Algorithm 2.2.

Step 1. Given $x_0 \in \mathbb{R}^n, \mu_1 > M > 0, 0 < p_0 \leq p_1 \leq p_2 < 1, 1 \leq \delta \leq 2, \varepsilon > 0, k := 0$ and $m \geq 1$. 

3
Step 2. Set \( x_k^{(0)} = x_k \) and \( d_k^{(0)} = d_k^M \). Compute \( F_k = F(x_k^{(0)}) \), \( J_k = J(x_k^{(0)}) \) and
\[
\Gamma_k = (J_k^T J_k + \lambda_k I)^{-1} J_k^T \quad \text{with} \quad \lambda_k = \mu_k \|F_k\|^2.
\] (12)
Step 3. If \( \|J_k^T F_k\| < \epsilon \), then stop. Compute
\[
d_k^{(0)} = -\Gamma_k F(x_k^{(0)}).
\] (13)
Set \( x_k^{(1)} = x_k^{(0)} + d_k^{(0)} \). Compute
\[
d_k^{(1)} = -\Gamma_k F(x_k^{(1)}).
\]
Set \( x_k^{(2)} = x_k^{(1)} + d_k^{(1)} \).
\[
\ldots
\]
Compute
\[
d_k^{(m-1)} = -\Gamma_k F(x_k^{(m-1)}).
\]
Set the trial step as
\[
s_k = d_k^{(0)} + \ldots + d_k^{(m-1)}.
\] (14)
Step 4. Compute \( r_k = \text{Ared}_k / \text{Pred}_k \). Set
\[
x_{k+1} = \begin{cases} x_k + s_k, & \text{if } r_k \geq p_0, \\ x_k, & \text{otherwise}. \end{cases}
\]
Step 5. Choose \( \mu_{k+1} \) as
\[
\mu_{k+1} = \begin{cases} 4\mu_k, & \text{if } r_k < p_1, \\ \mu_k, & \text{if } r_k \in [p_1, p_2], \\ \max\{\frac{\mu_k}{\tau}, \tilde{m}\}, & \text{if } r_k > p_2. \end{cases}
\] (15)
Set \( k = k + 1 \), and go to Step 2.

Notice that, \( \mu_k \) should be no less than a positive constant \( M \) to prevent the steps from being too large when the sequence \( \{x_k\} \) is near the solution.

3. The global convergence

To study the global convergence of Algorithm 2.2, we need the following assumptions.

**Assumption 3.1.** Let \( F(x) \) is continuously differentiable, and both \( F(x) \) and its Jacobian \( J(x) \) are Lipschitz continuous, i.e., there exist positive constant \( L_1 \) and \( L_2 \) such that
\[
\|J(y) - J(x)\| \leq L_1 \|y - x\|, \quad \forall x, y \in \mathbb{R}^n
\] (16)
and
\[
\|F(y) - F(x)\| \leq L_2 \|y - x\|, \quad \forall x, y \in \mathbb{R}^n.
\] (17)

By the Lipschitzness of the Jacobian proposed by (16), we have
\[
\|F(y) - F(x) - J(x)(y - x)\| \leq L_1 \|y - x\|^2, \quad \forall x, y \in \mathbb{R}^n.
\] (18)

**Theorem 3.2.** Under the conditions of Assumption 3.1, Algorithm 2.2 will terminates in finite iterations or satisfies
\[
\lim_{k \to \infty} \|J_k^T F_k\| = 0.
\] (19)
Proof. By contradiction, suppose there exist a positive \( \tau \) and infinite many \( k \) such that 
\[
\|J_k^T F_k\| \geq \tau. \tag{20}
\]

Let \( T_1, T_2 \) be the sets of the indices as follow:
\[
T_1 = \{ k \mid \|J_k^T F_k\| \geq \tau \},
\]
\[
T_2 = \{ k \mid \|J_k^T F_k\| \geq \frac{\tau}{2} \text{ and } x_{k+1} \neq x_k \}.
\]
It is easy to see that \( T_1 \) is infinite. In the following, we will derive the contradictions whether \( T_2 \) is finite or infinite. 

Case 1: \( T_2 \) is finite. Then the set
\[
T_3 = \{ k \mid \|J_k^T F_k\| \geq \tau \text{ and } x_{k+1} \neq x_k \}
\]
is also finite. Let \( \bar{k} \) be the largest index of \( T_3 \). Then it is easy to see that \( x_{\bar{k}+1} = x_{\bar{k}} \) holds for all \( k \in \{ k > \bar{k} \mid k \in T_1 \} \). Define the indices set
\[
T_4 = \{ k > \bar{k} \mid \|J_k^T F_k\| \geq \tau \text{ and } x_{k+1} = x_k \}
\]
If \( k \in T_4 \), we can deduce that \( \|J_k^T F_{k+1}\| \geq \tau \) and \( x_{k+2} = x_{k+1} \). Hence, we have \( x_{k+1} \in T_2 \). By induction, we know that \( \|J_k^T F_k\| \geq \tau \) and \( x_{k+1} = x_k \) hold for all \( k > \bar{k} \), which means \( r_k < p_0 \).

Now, due to (12), (13) and (15), we obtain
\[
\lambda_k \to +\infty \text{ and } \mu_k \to +\infty \tag{21}
\]
and
\[
d_k^{(0)} \to 0.
\]
Moreover, it follows from (17) and (18) that
\[
\|d_k^{(i)}\| = \| - \Gamma_k F(x_k^{(i)}) \|
\leq \|\Gamma_k F(x_k^{(i)})\| + \|\Gamma_k J_k (d_k^{(0)} + \cdots + d_k^{(i-1)})\| + L_1 \|d_k^{(0)} + \cdots + d_k^{(i-1)}\|^2 \|\Gamma_k\|
\leq \|\Gamma_k F(x_k^{(i)})\| + \|d_k^{(0)} + \cdots + d_k^{(i-1)}\| + L_1 \|d_k^{(0)} + \cdots + d_k^{(i-1)}\|^2.
\]
with \( i = 1, \cdots, m - 1 \). Hence, by induction, we obtain
\[
\|d_k^{(i)}\| \leq \varepsilon \|d_k^{(0)}\| \tag{22}
\]
for some \( \varepsilon > 0 \). Note that
\[
\|F(x_k^{(i)})\|^2 - \|F(x_k^{(i-1)}) + J_k d_k^{(i-1)}\|^2 = \left(\|F(x_k^{(i)})\|^2 + \|F(x_k^{(i-1)}) + J_k d_k^{(i-1)}\|^2 - 2\|F(x_k^{(i)})\| \|F(x_k^{(i-1)}) + J_k d_k^{(i-1)}\| \right)
\leq \left(2\|F_k + J_k (d_k^{(0)} + \cdots + d_k^{(i-1)})\| + L_1 \|d_k^{(0)} + \cdots + d_k^{(i-1)}\|^2 + L_1 \|d_k^{(0)} + \cdots + d_k^{(i-1)}\|^2 \right)
\]
with \( i = 2, \cdots, m \), and
\[
\|F(x_k^{(i)})\|^2 - \|F_k + J_k d_k^{(0)}\|^2 \leq 2L_1 \|F_k + J_k d_k^{(0)}\| \|d_k^{(0)}\|^2 + L_1^2 \|d_k^{(0)}\|^4. \tag{23}
\]
It follows from (7), (9), (17), (23) and (24) that
\[
|r_k - 1| = \left| \frac{\text{Ared}_k - \text{Pred}_k}{\text{Pred}_k} \right| \leq \frac{\sum_{i=1}^m \|F(x_k^{(i)})\|^2 - \|F(x_k^{(i-1)}) + J_k d_k^{(i-1)}\|^2}{\sum_{i=1}^m \|J_k^T F(x_k^{(i)})\| \min \left\{ \|d_k^{(i-1)}\|, \frac{\|J_k^T F(x_k^{(i-1)})\|}{\|J_k^{(i-1)}\|} \right\}} \leq O\left(\frac{\|d_k^{(0)}\|^2}{\|d_k^{(0)}\|}\right) \to 0,
\]
\[
5
\]
which implies that \( r_k \to 1 \). In view of the updating rule of \( \mu_k \), we know that there exists a positive constant \( \bar{\mu} > M \) such that \( \mu_k < \bar{\mu} \) holds for all sufficiently large \( k \), which is a contradiction to (21).

**Case 2:** \( T_2 \) is infinite. It follows from (9) and (17) that

\[
\|F\|_2^2 \geq \sum_{k \in T_2} (\|F_k\| - \|F_{k+1}\|) \geq \sum_{k \in T_2} \left( \|F_k\| - \|F_{k+1}\| \right) \geq \sum_{k \in T_2} p_0 \text{Pred}_k
\]

\[
\geq \sum_{k \in T_2} p_0 \left[ \sum_{i=1}^m \|J_k^T F(x_k^{(i-1)})\| \min \left\{ \|d_k^{(i-1)}\|, \|J_k^T F(x_k^{(i-1)})\| \right\} \right]
\]

\[
\geq \frac{\sum_{k \in T_2} p_0 \tau}{2} \min \left\{ \|d_k\|, \frac{\tau}{2L^2} \right\}.
\]

which implies

\[
\|d_k\| \to 0, \quad k \to \infty, \quad k \in T_2.
\]

Then the definition of \( d_k \) gives

\[
\mu_k \to +\infty, \quad k \in T_2.
\]

Moreover, it follows from (16), (17), (22) and (25) that

\[
\sum_{k \in T_2} \|J_k^T F_k\| - \|J_{k+1}^T F_{k+1}\| \leq \sum_{k \in T_2} \left( \|J_k^T F_k\| - \|J_{k+1}^T F_{k+1}\| - \|J_k^T F_{k+1}\| \right)
\]

\[
\leq \sum_{k \in T_2} \left[ L_2 \|J_k^T\| \|d_k^0\| + \cdots + d_k^{(m-1)}\| - L_1 \|F_{k+1}\| \|d_k^0\| + \cdots + d_k^{(m-1)}\| \right]
\]

\[
\leq L_1 L_2 \hat{c} \sum_{k \in T_2} \|d_k\| \leq +\infty,
\]

with some constants \( \hat{c} > 0 \), which together with (20) implies there exists a sufficiently large \( \hat{k} \) such that

\[
\|J_{\hat{k}}^T F_{\hat{k}}\| \geq \tau \quad \text{and} \quad \sum_{k \in T_2} \|J_k^T F_k\| - \|J_{k+1}^T F_{k+1}\| < \frac{\tau}{2}.
\]

Hence we can derive that \( \|J_{\hat{k}}^T F_{\hat{k}}\| \geq \frac{\tau}{2} \) for all \( k \geq \hat{k} \). Combining (26) with (27), we have

\[
\|d_k\| \to 0 \quad \text{and} \quad \mu_k \to +\infty.
\]

In the same way as proved in Case 1, we can also obtain that

\[
r_k \to 1.
\]

Hence, there exists a positive constant \( \bar{\mu} \) such that \( \mu_k < \bar{\mu} \) holds for all sufficiently large \( k \), which is contradicted to (28). The proof is completed.

### 4. The local convergence

In this section, we assume that \( x_k \to x^* \in X^* \) and the sequence \( \{x_k\} \) lies on some neighbourhood of \( x^* \), i.e., there exist a positive constant \( b_1 < 1 \) such that \( x \in N(x^*, b_1) \). We give some assumptions which the local convergence theory required.

**Assumption 4.1.** (a) \( F(x) \) is continuously differentiable, and Jacobian \( J(x) \) is Lipschitz continuous on \( N(x^*, b_1) \), i.e., there exist a positive constant \( L_1 \) such that

\[
\|J(y) - J(x)\| \leq L_1 \|y - x\|, \quad \forall x, y \in N(x^*, b_1) = \{ x \mid \|x - x^*\| \leq b_1 \}.
\]
(b) $|F(x)|$ provides a local error bound on some neighborhood of \(x^* \in X^*\), i.e., there exist a positive constant \(c > 0\) such that
\[
|F(x)| \geq c \text{ dist}(x, X^*), \quad \forall x \in N(x^*, b_1).
\] (30)

Since the condition of nonsingularity of \(J(x)\) is too strong, the Assumption 4.1 (b) provides a weak local error bound condition, which implies that the converse is not necessarily true [6].

By (29), we have
\[
|F(y) - F(x) - J(x)(y - x)| \leq L_1|y - x|^2, \quad \forall x, y \in N(x^*, b_1),
\] (31)
and
\[
|F(y) - F(x)| \leq L_2|y - x|, \quad \forall x, y \in N(x^*, b_1),
\] (32)
where \(L_2\) is a positive constant.

There exists a positive constant \(\omega > 0\) if \(F(x)\) provides a local error bound which proposed by Behling and Iusem in [17], then
\[
\text{rank}(J(\tilde{x})) = \text{rank}(J(x^*)), \quad \forall \tilde{x} \in N(x^*, \omega) \cap X^*.
\]

Let \(b \in (0, 1)\) and \(b_1 = \min\{\omega, b\}\). Without loss of generality, we further assume that \(x^{(i)}_k, i = 0, 1, \cdots, m - 1\) lie in \(N(x^*, \frac{b_1}{2})\).

In the following, we denote \(\tilde{x}_k \in X^*\) such that
\[
|\tilde{x}_k - x_k| = \text{dist}(x_k, X^*).
\]
Hence, we have
\[
|\tilde{x}_k - x^*| \leq |\tilde{x}_k - x_k| + |x_k - x^*| \leq 2 |x_k - x^*| \leq b_1,
\]
which implies that \(\tilde{x}_k \in N(x^*, b_1)\).

Now, we suppose rank\((J(\tilde{x})) = r\) for all \(\tilde{x} \in N(x^*, b_1) \cap X^*\) and the SVD of \(J(\tilde{x})\) is
\[
J(\tilde{x}) = \tilde{U}_k\tilde{\Sigma}_k\tilde{V}_k^T = (\tilde{U}_{k,1}, \tilde{U}_{k,2}) \begin{pmatrix} \tilde{\Sigma}_{k,1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{V}_{k,1} \\ \tilde{V}_{k,2} \end{pmatrix} = \tilde{U}_{k,1}\tilde{\Sigma}_{k,1}\tilde{V}_{k,1}^T + \tilde{U}_{k,2}\tilde{\Sigma}_{k,2}\tilde{V}_{k,2}^T,
\]
where \(\tilde{\Sigma}_{k,1} = \text{diag}(\tilde{\sigma}_{k,1}, \tilde{\sigma}_{k,2}, \cdots, \tilde{\sigma}_{k,r})\) with \(\tilde{\sigma}_{k,1} \geq \tilde{\sigma}_{k,2} \geq \cdots \geq \tilde{\sigma}_{k,r} > 0\). The corresponding SVD of \(J_k\) is
\[
J_k = U_k\Sigma_k V_k^T = (U_{k,1}, U_{k,2}, U_{k,3}) \begin{pmatrix} \Sigma_{k,1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_{k,1} \\ V_{k,2} \end{pmatrix} = U_{k,1}\Sigma_{k,1}V_{k,1}^T + U_{k,2}\Sigma_{k,2}V_{k,2}^T,
\]
where \(\Sigma_{k,1} = \text{diag}(\sigma_{k,1}, \sigma_{k,2}, \cdots, \sigma_{k,r})\) with \(\sigma_{k,1} \geq \sigma_{k,2} \geq \cdots \geq \sigma_{k,r} > 0\), and \(\Sigma_{k,2} = \text{diag}(\sigma_{k,r+1}, \sigma_{k,r+2}, \cdots, \sigma_{k,r+q})\) with \(\sigma_{k,r+1} \geq \sigma_{k,r+2} \geq \cdots \geq \sigma_{k,r+q} > 0\). We will neglect the subscript \(k\) if the context is clear in the following, and write \(J_k\) as
\[
J_k = U_1\Sigma_1 V_1^T + U_2\Sigma_2 V_2^T.
\] (33)

By the theory of matrix perturbation [18] and the Lipschitzness of \(J_k\), we have
\[
||\text{diag}(\tilde{\Sigma}_1 - \tilde{\Sigma}_1, \Sigma_2, 0)|| \leq ||J_k - \tilde{J}_k|| \leq L_1||\tilde{x}_k - x_k||, \quad (34)
\]
which yields
\[
||\Sigma_1 - \tilde{\Sigma}_1|| \leq L_1||\tilde{x}_k - x_k|| \quad \text{and} \quad ||\Sigma_2|| \leq L_1||\tilde{x}_k - x_k||. \quad (35)
\]
Hence
\[
||A_k^{-1}\Sigma_2|| = \frac{||\Sigma_2||}{\mu_k||F_k||^p} \leq \frac{L_1||\tilde{x}_k - x_k||}{mc^p||\tilde{x}_k - x_k||^p} = L_1m^{-1}c^{-p}||\tilde{x}_k - x_k||^{1-p}
\]
Since for any positive \(\sigma_i (i = 1, 2, \cdots, r)\), we have
\[
\frac{\sigma_i}{\sigma_i^2 + A_k} \leq \frac{\sigma_i}{2\sigma_i \sqrt{A_k}} = \frac{1}{2 \sqrt{A_k}}.
\]
Combining (4) and (36) with $\delta$, under the condition of Assumption 4.1, we have

$$\left\| \Sigma_1^2 + \lambda_k J_k^{-1} \right\| \leq \frac{1}{2 \sqrt{\mu_k} \|F\|} \leq \frac{1}{2} m^{-\frac{1}{2}} c^{-\frac{1}{2}} \|\bar{x}_k - x_k\|^{-\frac{1}{2}}. \quad (36)$$

Thus from (31), we get

$$P = \frac{\|\Sigma_1^2 + \lambda_k J_k^{-1} \Sigma_1\|}{\|\Sigma_1^2 + \lambda_k J_k^{-1} \Sigma_2\|} = \frac{\|\Sigma_1^2 + \lambda_k J_k^{-1} \Sigma_1\|}{\|\Sigma_1^2 + \lambda_k J_k^{-1} \Sigma_2\|}. \quad (37)$$

### 4.1. Properties of the trial step

Firstly, we investigate the properties of $d_k^{(0)}$, and hence $s_k$.

**Lemma 4.2.** Under the condition of Assumption 4.1, for sufficiently large $k$, we have

$$\|d_k^{(0)}\| \leq c_0 \text{dist}(x_k, X^*), \quad i = 0, 1, \ldots, m - 1,$$

where $c_0$ are some positive constants.

**Proof.** From (10), it is easy to see that $d_k^{(0)}$ is the minimizer of the following convex optimization problem:

$$\min_{d \in \mathbb{R}^k} \varphi_{d,0}(d) = \|F(x_k^{(0)}) + J_k d\|^2 + \lambda_k \|d\|^2.$$

Thus from (31), we get

$$\|d_k^{(0)}\|^2 \leq \frac{\varphi_{d,0}(d_k^{(0)})}{\lambda_k} \leq \frac{\varphi_{d,0}(\bar{x}_k - x_k)}{\lambda_k} = \frac{\|F(x_k^{(0)}) + J_k (\bar{x}_k - x_k)\|^2}{\lambda_k} + \|\bar{x}_k - x_k\|^2$$

$$\leq L_1 c^{-1} m^{-\frac{1}{2}} \|\bar{x}_k - x_k\|^2 \delta + \|\bar{x}_k - x_k\|^2,$$

with $\delta \in [1, 2]$, which implies that

$$\|d_k^{(0)}\| \leq c_0 \text{dist}(x_k, X^*). \quad (38)$$

Besides, from (5), (31), (37) and (38), we obtain

$$\|d_k^{(0)}\| = \| - (J_k^T J_k + \lambda_k J_k^{-1})^{-1} F_k (x_k^{(0)}) \|$$

$$\leq \| (J_k^T J_k + \lambda_k J_k^{-1})^{-1} F_k \| + \| (J_k^T J_k + \lambda_k J_k^{-1})^{-1} J_k^{(0)} + \cdots + d_k^{(i-1)} \|$$

$$+ L_1 \| d_k^{(0)} + \cdots + d_k^{(i-1)} \|^2 \| (J_k^T J_k + \lambda_k J_k^{-1})^{-1} J_k \|$$

$$\leq 2 \| d_k^{(0)} \| + \| d_k^{(1)} \| + \cdots + \| d_k^{(i-1)} \| + L_1 \| d_k^{(0)} + \cdots + d_k^{(i-1)} \|^2 \text{O} \| \bar{x}_k - x_k \|^{-\frac{1}{2}}$$

$$\leq 2 \| d_k^{(0)} \| + \cdots + \| d_k^{(i-1)} \| + \| d_k^{(i-1)} \| \leq c_0 \| \bar{x}_k - x_k \|^{-\frac{1}{2}},$$

with $i = 1, \ldots, m - 1$, for some positive constant $c_0$. The proof is completed.

Lemma 4.2 indicates that the trail step

$$\|s_k\| = \|d_k^{(0)} + \cdots + d_k^{(i-1)}\| \leq \|d_k^{(0)}\| + \cdots + \| d_k^{(i-1)} \| \leq \varepsilon \text{dist}(x_k, X^*),$$

for some positive constants $\varepsilon$. 

8
4.2. Boundedness of the LM parameter

**Lemma 4.3.** Under the conditions of Assumption 4.1, there exists a positive \( \tilde{M} > M \) such that \( \mu_k \leq \tilde{M} \) holds for all sufficiently large \( k \).

**Proof.** We first prove the following inequalities for all sufficiently large \( k \),

\[
\|F(x_k^{(0)})\|^2 - \|F(x_k^{(0)}) + J_k d_k^{(0)}\|^2 > \varepsilon(0)\|F(x_k^{(0)})\| \min \left\{ \|d_k^{(0)}\|, \|x_k^{(0)} - x_k^{(0)}\| \right\},
\]

where \( i = 0, 1, \ldots, m - 1 \).

Consider two cases. (a) \( \|\bar{x}_k^{(0)} - x_k^{(0)}\| \leq \|d_k^{(0)}\| \). By (30), (31) and the fact that \( d_k^{(0)} \) is the solution of (11), we have

\[
\|F(x_k^{(0)})\| - \|F(x_k^{(0)} + J_k d_k^{(0)})\| > \|F(x_k^{(0)}) - \|F(x_k^{(0)}) + J_k (\bar{x}_k^{(0)} - x_k^{(0)})\| - \|J_k - J_k(x_k^{(0)})\| \|x_k^{(0)} - x_k^{(0)}\|
\]

\[
> c\|\bar{x}_k^{(0)} - x_k^{(0)}\| - L_1\|\bar{x}_k^{(0)} - x_k^{(0)}\|^2 - L_1(||x_k^{(0)}|| + \cdots + ||d_k^{(i-1)}||)||\bar{x}_k^{(0)} - x_k^{(0)}||
\]

\[
> \varepsilon(0)||x_k^{(0)} - x_k^{(0)}||, \quad (39)
\]

for some \( \varepsilon(0) > 0 \) when \( k \) is sufficiently large.

(b) \( \|\bar{x}_k^{(0)} - x_k^{(0)}\| > \|d_k^{(0)}\| \). Similarly, we have

\[
\|F(x_k^{(0)})\| - \|F(x_k^{(0)} + J_k d_k^{(0)})\| > \|F(x_k^{(0)}) - \|F(x_k^{(0)}) + J_k (\bar{x}_k^{(0)} - x_k^{(0)})\| - \|J_k - J_k(x_k^{(0)})\| \|x_k^{(0)} - x_k^{(0)}\|
\]

\[
> \frac{\|d_k^{(0)}\|}{\|\bar{x}_k^{(0)} - x_k^{(0)}\|} \left( \|F(x_k^{(0)}) - \|F(x_k^{(0)}) + J_k (\bar{x}_k^{(0)} - x_k^{(0)})\| - \|J_k - J_k(x_k^{(0)})\| \|x_k^{(0)} - x_k^{(0)}\| \right)
\]

\[
> \frac{\|d_k^{(0)}\|}{\|\bar{x}_k^{(0)} - x_k^{(0)}\|} \varepsilon(0)||\bar{x}_k^{(0)} - x_k^{(0)}||
\]

\[
> \varepsilon(0)||d_k^{(0)}||. \quad (40)
\]

Combining (39) with (40), we obtain

\[
\|F(x_k^{(0)})\|^2 - \|F(x_k^{(0)} + J_k d_k^{(0)})\|^2 = (\|F(x_k^{(0)})\| + \|F(x_k^{(0)} + J_k d_k^{(0)})\|)(\|F(x_k^{(0)})\| - \|F(x_k^{(0)} + J_k d_k^{(0)})\|)
\]

\[
> \varepsilon(0)||F(x_k^{(0)})\| \min \{\|d_k^{(0)}\|, \|\bar{x}_k^{(0)} - x_k^{(0)}\| \}.
\]

Hence, it follows from (8) and Lemma 4.2, we have

\[
\text{Pred}_k \geq O(||\bar{x}_k - x_k||d_k^{(0)}||).
\]

Since \( d_k^{(0)} \) is a minimizer of (10), we have the following results from (32) and Lemma 4.2 that

\[
\|F_k + J_k d_k^{(0)} + \cdots + d_k^{(i-1)}\| \leq \|F_k + J_k d_k^{(0)} + \cdots + d_k^{(i-1)}\| + ||J_k d_k^{(0)}||
\]

\[
\leq \|F_k + J_k d_k^{(0)} + ||J_k d_k^{(0)}|| + \cdots + ||J_k d_k^{(0)}||
\]

\[
\leq ||F_k|| + ||J_k d_k^{(0)}|| + \cdots + ||J_k d_k^{(0)}||
\]

\[
\leq L_2||\bar{x}_k - x_k|| + \cdots + \varepsilon(0) L_2||\bar{x}_k - x_k||
\]

\[
> \varepsilon(0) ||\bar{x}_k - x_k||,
\]

with \( i = 0, \cdots, m - 1 \) for some positive constants \( \varepsilon(0) > 0 \). Also, follows from (23), we have

\[
\|F(x_k^{(0)})\|^2 - \|F(x_k^{(0)} + J_k d_k^{(0)})\|^2 \leq O(||\bar{x}_k - x_k||d_k^{(0)}||^2)
\]
which implies that
\[
|r_k - 1| = \frac{|\text{Ared}_k - \text{Pred}_k|}{\text{Pred}_k} \leq \frac{\mathcal{O}(\|x_k - x_i\|d_k^{(0)}||^2)}{\mathcal{O}(\|x_k - x_i\|d_k^{(0)})} \to 0
\]
holds for sufficiently large \(k\). Hence
\[r_k \to 1.\]
Therefore there exists a positive \(\tilde{M} > M\) such that \(\mu_k \leq \tilde{M}\) holds for all sufficiently large \(k\). The proof is completed.

4.3. Convergence order of \(m\)-step Levenberg-Marquardt algorithm

We now prove the convergence order of \(m\)-step LM algorithm based on the results obtained in the above two subsections.

By the SVD of \(J_k\) proposed in (33), we have
\[
d_k^{(0)} = -\bar{V}(\Sigma^2 + \alpha_k I)^{-1} \Sigma_1 U_1^T F(x_k^{(0)}) - \bar{V}_2(\Sigma^2 + \alpha_k I)^{-1} \Sigma_2 U_2^T F(x_k^{(0)}),
\]
with \(i = 0, \ldots, m - 1\).

**Lemma 4.4.** Under the condition of Assumption 4.1, if \(x_k^{(i)} \in N(x^*, b_1/2)\), then we have

(a) \(\|U_1 U_1^T F(x_k^{(i)})\| \leq \mathcal{O}(\|x_k - x_i\|^{i+1})\);
(b) \(\|U_2 U_2^T F(x_k^{(i)})\| \leq \mathcal{O}(\|x_k - x_i\|^{i+2})\);
(c) \(\|U_3 U_3^T F(x_k^{(i)})\| \leq \mathcal{O}(\|x_k - x_i\|^{i+2})\);

with \(i = 0, \ldots, m - 1\).

**Proof.** We will prove this lemma by an induction process. For \(i = 1, 2\), the results have been shown by Fan and Yang, and we have (see [12, 13])
\[
\|d_k^{(1)}\| \leq \mathcal{O}(\|x_k - x_i\|^{i+1}), \quad \|F(x_k^{(1)}) + J_k d_k^{(1)}\| \leq \mathcal{O}(\|x_k - x_i\|^{i+1}),
\]
\[
\|d_k^{(2)}\| \leq \mathcal{O}(\|x_k - x_i\|^{i}), \quad \|F(x_k^{(2)}) + J_k d_k^{(2)}\| \leq \mathcal{O}(\|x_k - x_i\|^{i+1}).
\]

Assuming the truth for some \(i - 1\), we obtain the induction hypothesis:
\[
\|d_k^{(i-1)}\| \leq \mathcal{O}(\|x_k - x_i\|^{i-1}), \quad \|F(x_k^{(i-1)}) + J_k d_k^{(i-1)}\| \leq \mathcal{O}(\|x_k - x_i\|^{i}).
\]

Turning now to the case for \(i\). It follows from above induction hypothesis that
\[
\|F(x_k^{(i)})\| = \|F(x_k^{(i-1)} + d_k^{(i-1)})\| \leq \|F(x_k^{(i-1)}) + J(x_k^{(i-1)})d_k^{(i-1)}\| + L_1 \|d_k^{(i-1)}\|^2
\]
\[
\leq \|F(x_k^{(i-1)}) + J_k d_k^{(i-1)}\| + \|J(x_k^{(i-1)}) - J_k\|\|d_k^{(i-1)}\| + L_1 \|d_k^{(i-1)}\|^2
\]
\[
\leq \mathcal{O}(\|x_k - x_i\|^{i+1}) + L_1 \|x_k - x_i\| \mathcal{O}(\|x_k - x_i\|^{i}) + L_1 \|x_k - x_i\|^{i+2}
\]
\[
\leq \mathcal{O}(\|x_k - x_i\|^{i+1}).
\]

So, we have
\[
\|U_1 U_1^T F(x_k^{(i)})\| \leq \|F(x_k^{(i)})\| \leq \mathcal{O}(\|x_k - x_i\|^{i+1}).
\]

Moreover, the local error bound condition implies that
\[
\|x_k^{(i)} - x_k^{(0)}\| \leq c^{-1}\|F(x_k^{(i)})\| \leq \mathcal{O}(\|x_k - x_i\|^{i+1}).
\]
Let \( \tilde{q}_k = -J_k^*F(x_k^{(0)}) \). Then \( \tilde{q}_k \) is the least squares solution of \( \| \min F(x_k^{(0)}) + J_kq \| \). It follows from (29), (31), (43) and Lemma 4.2 that

\[
\|U_3U_3^T F(x_k^{(0)}) - J_k \tilde{q}_k\| = \|F(x_k^{(0)}) + J_k \tilde{q}_k\| \\
\leq \|F(x_k^{(0)}) + J_k(x_k^{(0)} - x_k^{(0)})\| + \|J(x_k^{(0)} - J(x_k^{(0)}))(x_k^{(0)} - x_k^{(0)})\| + \cdots
\]

\[
+ \|J(x_k^{(0)} - J(x_k^{(0)}))(x_k^{(0)} - x_k^{(0)})\| \\
\leq L_1\|x_k^{(0)} - x_k^{(0)}\|^2 + L_2\|d_k^{(0)}\|\|x_k^{(0)} - x_k^{(0)}\| + \cdots + L_1\|d_k^{(0)}\|\|x_k^{(0)} - x_k^{(0)}\| \\
= O(\|x_k^{(0)} - x_k^{(0)}\|^2).
\]

(44)

Let \( \tilde{J}_k = U_1^*V_1^T \) and \( \tilde{q}_k = -\tilde{J}_k^*F(x_k^{(0)}) \). Since \( \tilde{q}_k \) is the least squares solution of \( \| \min F(x_k^{(0)}) + \tilde{J}_kq \| \), deducing from (29), (31), (35), (43) and Lemma 4.2 that

\[
\|U_2U_2^T + U_3U_3^T\| F(x_k^{(0)}) = \|F(x_k^{(0)}) + \tilde{J}_k(x_k^{(0)} - x_k^{(0)})\| \\
\leq \|F(x_k^{(0)}) + J(x_k^{(0)})(x_k^{(0)} - x_k^{(0)})\| + \|\tilde{J}_k(x_k^{(0)} - x_k^{(0)})\| \\
\leq L_1\|x_k^{(0)} - x_k^{(0)}\|^2 + \|\tilde{J}_k - J(x_k^{(0)})(x_k^{(0)} - x_k^{(0)})\| \\
\leq L_1\|x_k^{(0)} - x_k^{(0)}\|^2 + L_2\|d_k^{(0)}\|\|x_k^{(0)} - x_k^{(0)}\| + \cdots + L_1\|d_k^{(0)}\|\|x_k^{(0)} - x_k^{(0)}\| + L_1\|\tilde{x}_k - x_k\|\|x_k^{(0)} - x_k^{(0)}\| \\
= O(\|x_k^{(0)} - x_k^{(0)}\|^2).
\]

(45)

Due to the orthogonality of \( U_2 \) and \( U_3 \), combining (44) and (45), we know that

\[
\|U_2U_2^T F(x_k^{(0)})\| = O(\|\tilde{x}_k - x_k\|^2).
\]

The proof is completed.

Now, we are ready to give the estimations of \( d_k^{(m-1)} \) and \( F(x_k^{(m-1)}) + J_kd_k^{(m-1)} \).

**Lemma 4.5.** Under the condition of Assumption 4.1, for sufficiently large \( k \), we have

(a) \( \|d_k^{(m-1)}\| = O(\|\tilde{x}_k - x_k\|^m) \);

(b) \( \|F(x_k^{(m-1)}) + J_kd_k^{(m-1)}\| = O(\|\tilde{x}_k - x_k\|^m) \).

**Proof.** By (34), we have

\[
\|\Sigma_k\|^{-1} = \frac{1}{\sigma_r} \leq \frac{1}{\sigma_r - L_1\|\tilde{x}_k - x_k\|},
\]

which implies

\[
\|\Sigma_k\|^{-1} \leq \frac{2}{\sigma_r}.
\]

When \( \delta \in [1, 2] \), it then follows from Lemma 4.3, Lemma 4.4, (4), (41) and (42) that

\[
\|d_k^{(m-1)}\| = \|V_1(\Sigma_k^2 + \lambda_kI)^{-1}\Sigma_kU_1^T F(x_k^{(m-1)}) - V_2(\Sigma_k^2 + \lambda_kI)^{-1}\Sigma_kU_2^T F(x_k^{(m-1)})\| \\
\leq \|\Sigma_k\|^{-1} \|U_1^T F(x_k^{(m-1)})\| + \|\Sigma_k\|^{-1} \|U_2^T F(x_k^{(m-1)})\| \\
= O(\|\tilde{x}_k - x_k\|^m) + O(\|\tilde{x}_k - x_k\|^m) \\
\leq O(\|\tilde{x}_k - x_k\|^m).
\]

\[
\|F(x_k^{(m-1)}) + J_kd_k^{(m-1)}\| = \|\lambda_k U_1(\Sigma_k^2 + \lambda_kI)^{-1}U_1^T F(x_k^{(m-1)}) + \lambda_k U_2(\Sigma_k^2 + \lambda_kI)^{-1}U_2^T F(x_k^{(m-1)}) + U_3U_3^T F(x_k^{(m-1)})\| \\
\leq \lambda_k \|\tilde{x}_k - x_k\|^m + \|\tilde{x}_k - x_k\|^m + \|\tilde{x}_k - x_k\|^m + O(\|\tilde{x}_k - x_k\|^m) \\
\leq O(\|\tilde{x}_k - x_k\|^m).
\]

The proof is completed.
Based on the results above, we obtain the convergence rate of Algorithm 2.2.

**Theorem 4.6.** Under the conditions of Assumptions 4.1, the convergence rate of Algorithm 2.2 is $(m + 1)^{th}$.

**Proof.** It follows from Lemma 4.2 and Lemma 4.5 that

$$
\|\tilde{x}_{k+1} - x_{k+1}\| \leq \|F(x_{k+1})\| = \|F(x_k + s_k)\| = \|F(x_k^{(m-1)} + d_k^{(m-1)})\|
$$

$$
\leq \|F(x_k^{(m-1)}) + J(x_k^{(m-1)})d_k^{(m-1)}\| + L_1\|d_k^{(m-1)}\|^2
$$

$$
\leq \|F(x_k^{(m-1)}) + J(x_k^{(m-1)})d_k^{(m-1)}\| + \|\|J(x_k^{(m-1)}) - J_k\|d_k^{(m-1)}\| + L_1\|d_k^{(m-1)}\|^2 + L_1\|d_k^{(m-1)}\|^2
$$

$$
\leq \|F(x_k^{(m-1)}) + J(x_k^{(m-1)})d_k^{(m-1)}\| + L_1(\|d_k^{(0)}\| + \cdots + \|d_k^{(m-2)}\|)\|d_k^{(m-1)}\| + L_1\|d_k^{(m-1)}\|^2
$$

$$
\leq O(\|\tilde{x}_k - x_k\|^{m+1}) + L_1(\|\tilde{x}_k - x_k\| + \cdots + \|\tilde{x}_k - s_k\|)O(\|\tilde{x}_k - x_k\|^{m}) + O(\|\tilde{x}_k - x_k\|^{2m})
$$

with $m \geq 1$. Hence we have

$$
\|\tilde{x}_{k+1} - x_{k+1}\| \leq O(\|\tilde{x}_k - x_k\|^{m+1}),
$$

which means that $\{x_k\}$ generated by $m$-step LM method converges to the solution set $X^*$ with $(m + 1)^{th}$ order. The proof is completed.

5. **Numerical results**

We will compute some singular problems, which come from [19] with the same forms as in [20], to test Algorithm 2.2, and compare it with the general LM algorithm (LM), the two-step LM method which has presented in [12] and the three-step LM method which has presented in [13].

We compute these test problems with different initial points and different size,$$
\tilde{F}(x) = F(x) - J(x)A(ATA)^{-1}A^T(x - x^*),
$$
where $F(x)$ is the standard nonsingular test function, $x^* $ is its root, and $A \in \mathbb{R}^{n \times k}$ has full column rank with $1 \leq k \leq n$. Obviously, $\tilde{F}(x^*) = 0$ and

$$
\tilde{J}(x^*) = J(x^*)(I - A(ATA)^{-1}A^T)
$$

has rank $n - k$. A disadvantage of these problems is that $\tilde{F}(x)$ may have roots that are not roots of $F(x)$. We chose the rank of $\tilde{J}(x^*)$ to be $n - 1$ and $n - 2$, respectively, by using

$$
A \in \mathbb{R}^{n \times 1}, \quad A^T = (1, 1, \cdots, 1)
$$

and

$$
A \in \mathbb{R}^{n \times 2}, \quad A^T = \begin{pmatrix}
1 & 1 & 1 & 1 & \cdots & 1 \\
1 & -1 & -1 & -1 & \cdots & \pm 1
\end{pmatrix}.
$$

Set $p_0 = 0.0001$, $p_1 = 0.25$, $p_2 = 0.75$, $\bar{m} = 10^{-6}$, $\mu_1 = 1$, $\delta = 1$. The stopping criteria for the Algorithm is $\|J_kF_k\| < 10^{-5}$ or the iteration number exceeds $100(n + 1)$. The points $x_0, 10x_0, 100x_0$ in the third column of the tables are the starting points, where $x_0$ was suggested by Moré et. al in [19]. “NF” and “NJ” represent the number of function calculations and Jacobian calculations, respectively. We use “NF+n*NJ” to measure the total calculations since the Jacobian calculations are usually $n$ times of the function calculations. If the method failed to find the solution in $100(n + 1)$ iterations, we denoted it by the sign “-”, and if the iterations had underflows or overflows, we denoted it by “OF”.

Here, we set $m = 5$. From table 1 and table 2, we can see that the $m$-step LM method almost always outperforms or at least performs as well as the two-step LM method or the three-step LM method whether on the first singular test set or on the second test set. The Jacobian calculations of the $m$-step LM method are much less than those of the two-step LM method and the three-step LM method, although the function calculations are more, which contributes to fewer calculations of the $m$-step LM method than the two-step LM method and the three-step LM method. It implies that the calculations of Jacobian will be saved depending on the number of step $m$. That would be great helpful for the real application of the method and especially useful for the large scale problems.

12
6. Conclusions

In this work, we presented an efficient $m$–step LM method for nonlinear equations. At every iteration, we compute $m$ approximate LM steps with frozen $(J_k^T J_k + A_k I)^{-1} J_k^T$. The $m$–step LM method have been proved to have $(m + 1)$th convergence order under the local error bound condition which is weaker than nonsingularity. Numerical results show that the $m$–step LM method saved more Jacobian calculations although the calculations of function are more.

Acknowledgments

The work is supported by the NSFC (61300048), the Anhui Provincial Natural Science Foundation (1508085MA14), the Key Natural Science Foundation of Universities of Anhui Province (KJ2014A223), the Major Teaching Reform Project of Anhui Higher Education Revitalization Plan (2014ZDY058) and the Excellent Young Talents in Universities of Anhui Province.

References

<table>
<thead>
<tr>
<th>Problem</th>
<th>n</th>
<th>x₀</th>
<th>LM</th>
<th>Two-step LM</th>
<th>Three-step LM</th>
<th>m–step LM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>NF/NJ/NJT/F</td>
<td>NF/NJ/NJT/F</td>
<td>NF/NJ/NJT/F</td>
<td>NF/NJ/NJT/F</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>15/15/45</td>
<td>21/11/43</td>
<td>25/9/43</td>
<td>36/8/52</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>17/17/51</td>
<td>25/13/51</td>
<td>31/11/53</td>
<td>41/9/59</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>21/21/63</td>
<td>29/15/59</td>
<td>37/13/63</td>
<td>46/10/66</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>–</td>
<td>OF</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>35/18/71</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1</td>
<td>16/16/80</td>
<td>23/12/71</td>
<td>28/10/68</td>
<td>36/8/68</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>19/19/95</td>
<td>27/14/83</td>
<td>34/12/82</td>
<td>46/10/86</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>22/22/110</td>
<td>31/16/95</td>
<td>40/14/96</td>
<td>51/11/95</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1</td>
<td>8/8/32</td>
<td>11/6/29</td>
<td>13/5/28</td>
<td>16/4/28</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>8/8/32</td>
<td>9/5/24</td>
<td>16/6/34</td>
<td>41/8/65</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>8/8/32</td>
<td>11/6/29</td>
<td>22/8/46</td>
<td>36/8/60</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>1</td>
<td>44/25/819</td>
<td>75/15/540</td>
<td>142/20/762</td>
<td>136/9/415</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>1</td>
<td>8/8/88</td>
<td>11/6/71</td>
<td>13/5/63</td>
<td>16/4/56</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>23/23/253</td>
<td>33/17/203</td>
<td>40/14/180</td>
<td>56/12/176</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>7/7/77</td>
<td>11/6/71</td>
<td>13/5/63</td>
<td>16/4/56</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>9/9/99</td>
<td>13/7/83</td>
<td>16/6/76</td>
<td>21/5/71</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>7/7/217</td>
<td>11/6/191</td>
<td>13/5/163</td>
<td>21/5/171</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>10/10/310</td>
<td>13/7/223</td>
<td>16/6/196</td>
<td>21/5/171</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>7/7/357</td>
<td>11/6/311</td>
<td>13/5/263</td>
<td>21/5/271</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>10/10/510</td>
<td>13/7/363</td>
<td>16/6/316</td>
<td>21/5/271</td>
</tr>
<tr>
<td>11</td>
<td>30</td>
<td>1</td>
<td>35/19/605</td>
<td>45/9/315</td>
<td>106/18/646</td>
<td>146/13/536</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>60/46/1440</td>
<td>85/23/775</td>
<td>205/42/1465</td>
<td>286/28/1126</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>45/35/1095</td>
<td>99/24/819</td>
<td>187/45/1537</td>
<td>366/42/1626</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td></td>
<td>19/7/369</td>
<td>55/14/755</td>
<td>82/11/632</td>
<td>96/6/396</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>35/35/1235</td>
<td>163/45/2413</td>
<td>118/24/1318</td>
<td>281/26/1581</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>41/27/1391</td>
<td>121/34/1821</td>
<td>148/27/1498</td>
<td>221/22/1321</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>1</td>
<td>14/14/154</td>
<td>19/10/119</td>
<td>25/9/115</td>
<td>31/7/101</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>16/16/176</td>
<td>21/11/131</td>
<td>28/10/128</td>
<td>36/8/116</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>19/19/209</td>
<td>27/14/167</td>
<td>34/12/154</td>
<td>46/10/46</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td></td>
<td>20/20/1020</td>
<td>27/14/727</td>
<td>34/12/634</td>
<td>46/10/546</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>21/21/1071</td>
<td>31/16/831</td>
<td>37/13/687</td>
<td>51/11/601</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>25/25/1275</td>
<td>35/18/935</td>
<td>43/15/793</td>
<td>61/3/1711</td>
</tr>
<tr>
<td>13</td>
<td>30</td>
<td>1</td>
<td>9/9/279</td>
<td>13/7/223</td>
<td>16/6/196</td>
<td>21/5/171</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>14/14/434</td>
<td>19/10/319</td>
<td>25/9/295</td>
<td>31/7/241</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>17/17/527</td>
<td>25/13/415</td>
<td>31/11/361</td>
<td>41/9/311</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td></td>
<td>9/9/459</td>
<td>13/7/363</td>
<td>16/6/316</td>
<td>21/5/271</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>14/14/714</td>
<td>19/10/519</td>
<td>25/9/475</td>
<td>31/7/381</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>17/17/867</td>
<td>25/13/675</td>
<td>31/11/581</td>
<td>41/9/491</td>
</tr>
<tr>
<td>14</td>
<td>30</td>
<td>1</td>
<td>12/12/372</td>
<td>17/9/287</td>
<td>22/8/262</td>
<td>31/7/241</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>18/18/558</td>
<td>27/14/447</td>
<td>31/11/361</td>
<td>41/9/311</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>24/24/744</td>
<td>35/18/575</td>
<td>43/15/493</td>
<td>56/12/416</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td></td>
<td>12/12/612</td>
<td>17/9/467</td>
<td>22/8/422</td>
<td>31/7/381</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>18/18/918</td>
<td>27/14/727</td>
<td>31/11/581</td>
<td>41/9/491</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>24/24/1224</td>
<td>35/18/935</td>
<td>43/15/793</td>
<td>56/12/656</td>
</tr>
<tr>
<td>Problem</td>
<td>n</td>
<td>$x_0$</td>
<td>LM</td>
<td>Two-step LM</td>
<td>Three-step LM</td>
<td>m-step LM</td>
</tr>
<tr>
<td>---------</td>
<td>----</td>
<td>-------</td>
<td>----</td>
<td>-------------</td>
<td>--------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NF/NI/NT/F</td>
<td>NF/NI/NT/F</td>
<td>NF/NI/NT/F</td>
<td>NF/NI/NT/F</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>11/11/33</td>
<td>25/9/31</td>
<td>19/7/33</td>
<td>26/6/38</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>13/13/39</td>
<td>19/10/39</td>
<td>22/8/38</td>
<td>31/7/45</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>17/17/51</td>
<td>23/12/47</td>
<td>28/10/48</td>
<td>36/8/52</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>35/25/85</td>
<td>57/17/91</td>
<td>82/17/116</td>
<td>231/23/277</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>59/54/167</td>
<td>91/34/159</td>
<td>148/32/212</td>
<td>166/24/214</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>25/18/61</td>
<td>51/15/81</td>
<td>67/14/95</td>
<td>86/11/108</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1</td>
<td>14/14/70</td>
<td>19/10/59</td>
<td>25/9/61</td>
<td>31/7/59</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>17/17/85</td>
<td>23/12/71</td>
<td>28/10/68</td>
<td>41/9/77</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>20/20/100</td>
<td>29/15/89</td>
<td>34/12/82</td>
<td>46/10/86</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1</td>
<td>13/13/52</td>
<td>19/10/49</td>
<td>25/9/52</td>
<td>31/7/52</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>14/14/56</td>
<td>19/10/49</td>
<td>25/9/52</td>
<td>31/7/52</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>24/18/78</td>
<td>21/11/54</td>
<td>25/9/52</td>
<td>31/7/52</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>1</td>
<td>3200/1793/58783</td>
<td>3201/801/28032</td>
<td>3202/731/25863</td>
<td>3201/576/21057</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>1</td>
<td>8/8/88</td>
<td>11/6/71</td>
<td>13/5/63</td>
<td>16/4/56</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>23/23/253</td>
<td>33/17/203</td>
<td>40/14/180</td>
<td>56/12/176</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>7/7/77</td>
<td>11/6/71</td>
<td>13/5/63</td>
<td>16/4/56</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>10/10/110</td>
<td>13/7/83</td>
<td>16/6/76</td>
<td>21/5/71</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1</td>
<td>2/2/102</td>
<td>3/2/103</td>
<td>4/2/104</td>
<td>6/2/106</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>6/6/306</td>
<td>9/5/259</td>
<td>13/5/263</td>
<td>16/4/216</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>9/9/459</td>
<td>13/7/363</td>
<td>16/6/316</td>
<td>21/5/271</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>7/7/217</td>
<td>11/6/191</td>
<td>13/5/163</td>
<td>21/5/171</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>10/10/310</td>
<td>15/8/255</td>
<td>19/7/229</td>
<td>26/6/206</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>7/7/357</td>
<td>11/6/311</td>
<td>13/5/263</td>
<td>21/5/271</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>11/11/561</td>
<td>15/8/415</td>
<td>19/7/369</td>
<td>26/6/326</td>
</tr>
<tr>
<td>11</td>
<td>30</td>
<td>1</td>
<td>19/9/289</td>
<td>41/9/311</td>
<td>70/12/430</td>
<td>101/8/341</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>62/4/1292</td>
<td>65/17/575</td>
<td>115/23/805</td>
<td>516/55/2166</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>43/35/1093</td>
<td>87/23/777</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1</td>
<td>25/14/725</td>
<td>61/15/811</td>
<td>100/14/800</td>
<td>106/7/456</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>50/34/1750</td>
<td>75/20/1075</td>
<td>133/22/1233</td>
<td>211/19/1161</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>51/35/1801</td>
<td>117/37/1967</td>
<td>133/26/1433</td>
<td>226/21/1276</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>1</td>
<td>14/14/154</td>
<td>19/10/119</td>
<td>25/9/115</td>
<td>31/7/101</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>16/16/176</td>
<td>21/11/131</td>
<td>28/10/128</td>
<td>36/8/116</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>19/19/209</td>
<td>27/14/167</td>
<td>34/12/154</td>
<td>46/10/146</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1</td>
<td>20/20/1020</td>
<td>27/14/727</td>
<td>34/12/634</td>
<td>46/10/546</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>21/21/1071</td>
<td>31/16/831</td>
<td>37/13/687</td>
<td>51/11/601</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>25/25/1275</td>
<td>35/18/935</td>
<td>43/15/793</td>
<td>61/3/711</td>
</tr>
<tr>
<td>13</td>
<td>30</td>
<td>1</td>
<td>9/9/279</td>
<td>13/7/223</td>
<td>16/6/196</td>
<td>21/5/171</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>14/14/434</td>
<td>19/10/319</td>
<td>25/9/295</td>
<td>31/7/241</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>17/17/527</td>
<td>25/13/415</td>
<td>31/11/361</td>
<td>41/9/311</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1</td>
<td>9/9/459</td>
<td>13/7/363</td>
<td>16/6/316</td>
<td>21/5/271</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>14/14/714</td>
<td>19/10/519</td>
<td>25/9/475</td>
<td>31/7/381</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>17/17/867</td>
<td>25/13/675</td>
<td>31/11/581</td>
<td>41/9/491</td>
</tr>
<tr>
<td>14</td>
<td>30</td>
<td>1</td>
<td>12/12/372</td>
<td>17/9/287</td>
<td>22/8/262</td>
<td>31/7/241</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>18/18/558</td>
<td>27/14/447</td>
<td>31/11/361</td>
<td>41/9/311</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>24/24/744</td>
<td>35/18/575</td>
<td>43/15/493</td>
<td>56/12/416</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1</td>
<td>12/12/612</td>
<td>17/9/467</td>
<td>22/8/422</td>
<td>31/7/381</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>18/18/918</td>
<td>27/14/727</td>
<td>31/11/581</td>
<td>41/9/491</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
<td>24/24/1124</td>
<td>35/18/935</td>
<td>43/15/793</td>
<td>56/12/656</td>
</tr>
</tbody>
</table>