Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 7(1): 1241–1256.
DOI: 10.3934/math.2022073
Received: 11 July 2021
Accepted: 11 October 2021
Published: 21 October 2021

Research article

A variant of the Levenberg-Marquardt method with adaptive parameters for systems of nonlinear equations

Lin Zheng[1,4], Liang Chen[2,*] and Yanfang Ma[3,5]

[1] School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

[2] School of Sciences, Changzhou Institute of Technology, Changzhou, Jiangsu 213032, China

[3] School of Computer Science and Information Engineering, Changzhou Institute of Technology, Changzhou, Jiangsu 213032, China

[4] Institute of Quantitative Economics, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

[5] School of Computer Science and Technology, Huaibei Normal University, Huaibei, Anhui 235000, China

* **Correspondence:** Email: clmyf2@163.com.

Abstract: The Levenberg-Marquardt method is one of the most important methods for solving systems of nonlinear equations and nonlinear least-squares problems. It enjoys a quadratic convergence rate under the local error bound condition. Recently, to solve nonzero-residue nonlinear least-squares problem, Behling et al. propose a modified Levenberg-Marquardt method with at least superlinearly convergence under a new error bound condtion [3]. To extend their results for systems of nonlinear equations, by choosing the LM parameters adaptively, we propose an efficient variant of the Levenberg-Marquardt method and prove its quadratic convergence under the new error bound condition. We also investigate its global convergence by using the Wolfe line search. The effectiveness of the new method is validated by some numerical experiments.

Keywords: systems of nonlinear equations; Levenberg-Marquardt method; global convergence; Wolfe line search; error bound condition
Mathematics Subject Classification: 65K05, 90C30

1. Introduction

We consider the numerical solution of the following system of nonlinear equations

$$F(x) = 0, \tag{1.1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^m$ is a continuously differentiable function.

Nonlinear equations of the form (1.1) are often solved as a key ingredient in simulations of many real-world problems. Classic methods for solving (1.1) include the Gauss-Newton method, the inexact Newton method, the Broyden's method and the trust region method [9, 13, 15]. In actual computations, however, the Gauss-Newton method becomes less competitive when the Jacobian is (nearly) rank-deficient. By adopting a trust-region approach in place of the line search in the Gauss-Newton method, the Levenberg-Marquardt (LM) method circumvents this shortcoming even though it uses the same Hessian approximations as in the Gauss-Newton method.

In the trial step of the LM method, one needs to solve per step the following linear system

$$(J_k^T J_k + \lambda_k I)d_k = -J_k^T F_k, \tag{1.2}$$

where $\lambda_k \geq 0$, $F_k = F(x_k)$, $J_k = J(x_k)$ is the Jacobian and $I \in \mathbb{R}^{n\times n}$ stands for the identity matrix. If J_k is nonsingular and Lipschitz continuous for the case $m = n$, the initial guess x_0 is close enough to the solution x^* of (1.1) and the LM parameter λ_k is updated recursively, then the LM method has a quadratic convergence rate.

For some applications, the need for a nonsingular Jacobian J_k can be rather stringent. Therefore, it is necessary to come up with numerical methods in the absence of a nonsingular Jacobian. To this end, some efforts have been made recently; for instance, Yamashita et al. propose a local error bound condition which does not requires nonsingularity of the Jacobian [19]. In what follows, we denote by X^* the nonempty solution set of (1.1) and use $\|\cdot\|$ to represent the 2-norm of vectors or matrices if there is no ambiguity. Let $N(x^*, b) = \{x \mid \|x-x^*\| \leq b\}$ be a subset of the n-dimensional vector space such that the intersection $X^* \cap N(x^*, b)$ is nonempty. The LM method is shown to have a quadratic convergence rate if there exists a positive constant c satisfying the following *local error bound condition* [2, 6, 19]

$$c\ \mathrm{dist}(x, X^*) \leq \|F(x)\|, \quad \text{for } x \in N(x^*, b), \tag{1.3}$$

where $\mathrm{dist}(x, X^*)$ is the distance from x to X^*.

In spite of the advantage of avoiding nonsingularity of the Jacobian, the local error bound condition (1.3) is not always applicable for some ill-conditioned nonlinear equations from application fields like biochemical systems. In light of this, Guo et al. present the Hölderian error bound condition that is more applicable than (1.3) [8]. The Hölderian error bound condition is given by

$$c\ \mathrm{dist}(x, X^*) \leq \|F(x)\|^\gamma, \quad \text{for } x \in N(x^*, b), \tag{1.4}$$

where $c > 0$ and $\gamma \in (0, 1]$. Obviously, the Hölderian error bound condition (1.4) includes the local error bound condition (1.3) as a special case. In fact, the bound (1.4) reduces to (1.3) when $\gamma = 1$. It should be noted that the Hölderian local error bound condition is also called Hölder metric subregularity which is closely related with the Łojasiewica inequalities; see [14, 17] for detail. With the assumption (1.4), the LM method converges at least superlinearly when γ and the LM parameter satisfy certain conditions [1, 8, 18, 21].

Apart from its application in solving systems of nonlinear equations, the LM method also finds its way into numerical solution of nonlinear least squares problems. To investigate the local convergence of the LM method for the nonlinear least-squares problem with possible nonzero residue, Behling et al. [3] present a local error bound condition characterized by $\|J(x)^T F(x)\|$, i.e.,

$$c\ \mathrm{dist}(x, X^*) \le \|J(x)^T F(x)\|, \quad \text{for } x \in N(x^*, b), \tag{1.5}$$

where $c > 0$. We stress that the local error bound condition (1.5) can also be derived from the bound (1.3) [10, Lemma 5.4]. However, the former is more practical than the latter in that it does not require the nonsingularity of the Jacobian. With the assumption (1.5), the LM method is shown to have at least linearly convergence order with suitable choices of the LM parameter [3].

As observed from (1.2), the LM parameter λ_k is introduced in case that $J_k^T J_k$ is (nearly) singular. Such practice not only guarantees the uniqueness of solution of (1.2) but also helps to reduce the iteration steps. In this sense, the LM parameter plays a key role in the LM method. Some promising candidates of the LM parameter have been proposed recently; for instance, Yamashita et al. [19] select $\lambda_k = \|F_k\|^2$ and show that the LM method has quadratically convergence with the assumption (1.3). Fan and Yuan [6] generalize it with the LM parameter $\lambda_k = \|F_k\|^\delta$ with $\delta \in [1, 2]$. It is shown that the quadratic convergence is still retained with the assumption (1.3). Dennis and Schnable consider the choice $\lambda_k = O(\|J_k^T F_k\|)$ [4]. Following this reasoning, Fischer employs $\lambda_k = \|J_k^T F_k\|$ in [7] which is further generalized to the form $\lambda_k = \|J_k^T F_k\|^\delta$ with $\delta \in (0, 1]$ in [3]. With the assumtion (1.5), Behling et al. conclude that the LM method converges at least linearly to some solution of (1.1) when $\delta \in (0, 1)$ and quadratically when $\delta = 1$ [3]. More recent progress in choosing the LM parameter λ_k can be found in [5, 10, 11].

Instead of adopting the choice used in [3], we propose to use the LM parameter $\lambda_k = \|J_k^T F_k\|^\delta$ with $\delta \in [1, 2]$ in this work. The motivation of our work is clarified as follows. Intuitively, the step size $\|d_k\|$ is small if $\|J_k^T F_k\|$ is too large, which may hamper a fast convergence. Fortunately, it poses no difficulty by considering the following choice, i.e.,

$$\lambda_k = \begin{cases} \|J_k^T F_k\|^\delta, & \text{if } \|J_k^T F_k\| \le 1, \\ \|J_k^T F_k\|^{-\delta}, & \text{Otherwise}, \end{cases} \qquad \delta \in [1, 2]. \tag{1.6}$$

From the convergence theory, we know $\|J_k^T F_k\|$ always converges to 0, hence $\|J_k^T F_k\| > 1$ only occurs at beginning finite iterate steps and it is a special case for the numerical method. Since the choice of λ_k in (1.6) is adaptive, then the variant LM method is called an adaptive Levenberg-Marquardt method (ALMM) in this paper.

The rest of this paper is organized as follows. In Section 2, the adaptive Levenberg-Marquardt method is introduced. Its convergence rate under the assumption (1.5) is examined. In section 3, the adaptive Levenberg-Marquardt method with Wolfe line search rule as well as its global convergence are investigated. In Section 4, some numerical experiments are used to verify the effectiveness of the new method. Finally, some conclusions are given in Section 5.

2. Local Convergence of the adaptive LM method

In this section, we consider the adaptive LM method with unit step size and investigate its local convergence near a solution.

To begin with our discussion, we present the following adaptive LM method:

$$d_k = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k, \tag{2.1}$$

$$x_{k+1} = x_k + d_k,$$

where the LM parameter is defined in (1.6).

To establish the local convergence results for the adaptive LM algorithm, we need the following assumptions throughout the paper.

Assumption 2.1. *(a) The Jacobian $J(x)$ is Lipschitz continuous in a neighborhood $N(x^*, b)$, i.e., there exists a constant $L_1 > 0$ such that*

$$\|J(x) - J(y)\| \le L_1\|x - y\|, \quad \forall x, y \in N(x^*, b). \tag{2.2}$$

(b) We said that $\|J(x)^T F(x)\|$ provides a local error bound on $N(x^*, b)$ if there exists a constant $c > 0$ such that

$$c\ \mathrm{dist}(x, X^*) \le \|J(x)^T F(x)\|, \quad \forall x \in N(x^*, b). \tag{2.3}$$

To guarantee the initial point x_0 is sufficiently close to x^*, we assume $b > 0$ is sufficient small.

From Assumption 2.1(a), we note that

$$\|F(x) - F(y) - J(y)(x - y)\| \le L_1\|x - y\|^2, \quad \forall x, y \in N(x^*, b). \tag{2.4}$$

By compactness, we have

$$\|J(x)\| \le L_2 \text{ and } \|F(x)\| \le \beta, \quad \forall x \in N(x^*, b), \tag{2.5}$$

where constants $L_2 > 0$ and $\beta > 0$. Therefore, it follows from the mean value inequality that

$$\|F(x) - F(y)\| \le L_2\|x - y\|, \quad \forall x, y \in N(x^*, b). \tag{2.6}$$

Denote by $\bar{x}_k \in X^*$ which satisfies

$$\|\bar{x}_k - x_k\| = \mathrm{dist}(x_k, X^*).$$

Lemma 2.1. *Let the sequence $\{x_k\}$ be generate by the adaptive LM method and Assumptions 2.1 hold. There exists some positive constants $c_1, \tilde{c}_1$, such that*

$$\tilde{c}_1 \mathrm{dist}(x_k, X^*)^{\pm\delta} \le \lambda_k \le \min\{1, c_1^\delta \mathrm{dist}(x_k, X^*)^\delta\}. \tag{2.7}$$

Proof. We derive the proof in two cases.

Case I: $\|J_k^T F_k\| \le 1$. Then $\lambda_k = \|J_k^T F_k\|^\delta$. From Assumption 2.1 (b), the inequality in the left-hand side (2.7) is obtained, i.e.,

$$c^\delta \mathrm{dist}(x_k, X^*)^\delta \le \lambda_k = \|J_k^T F_k\|^\delta.$$

Now, we verify the right-hand side inequality in (2.7).

It follows from (2.5) and (2.6) that

$$\begin{aligned}
\|J(x)^T F(x) - &J(y)^T F(y)\| \\
&= \|J(x)^T F(x) - J(x)^T F(y) + J(x)^T F(y) - J(y)^T F(y)\| \\
&\le \|J(x)^T\|\ \|F(x) - F(y)\| + \|F(y)\|\ \|J(x)^T - J(y)^T\|
\end{aligned}$$

$$\leq L_2^2\|x-y\| + \beta L_1\|x-y\| = c_1\|x-y\|, \tag{2.8}$$

where $c_1 = L_2^2 + \beta L_1$. Since $\lambda_k = \|J_k^T F_k\|^\delta$, then we obtain

$$\lambda_k \leq c_1^\delta \text{dist}(x_k, X^*)^\delta.$$

Case II: $\|J_k^T F_k\| > 1$. Then $\lambda_k = \|J_k^T F_k\|^{-\delta} < 1$. From (2.8), we also have

$$c_1^{-\delta}\text{dist}(x_k, X^*)^{-\delta} \leq \lambda_k = \|J_k^T F_k\|^{-\delta}.$$

Summarizing the above two cases, we obtain the inequality (2.7) with $\tilde{c}_1 = \min\{c^\delta, c_1^{-\delta}\}$. The proof is completed.

Lemma 2.2. *Let the sequence $\{x_k\}$ be generate by the adaptive LM method and Assumptions 2.1 hold. If $x_k \in N(x^*, b/2)$, there exists a constant $c_2 > 0$ such that*

$$\|d_k\| \leq c_2 \text{dist}(x_k, X^*). \tag{2.9}$$

Proof. From the assumption, we have

$$\|\bar{x}_k - x^*\| \leq \|\bar{x}_k - x_k\| + \|x_k - x^*\| \leq \|x_k - x^*\| + \|x_k - x^*\| \leq b,$$

which indicates that $\bar{x}_k \in N(x^*, b)$. Define

$$\varphi_k(d) = \|F_k + J_k d\|^2 + \lambda_k\|d\|^2. \tag{2.10}$$

From (2.1) and the convexity of $\varphi_k(d)$, we note that d_k is not only a stationary point but also a minimizer of $\varphi_k(d)$. By using the fact that $x_k, \bar{x}_k \in N(x^*, b)$, we have from (2.4) and Lemma 2.1 that

$$\begin{aligned}\|d_k\|^2 \leq &\frac{\varphi_k(d_k)}{\lambda_k} \leq \frac{\varphi_k(\bar{x}_k - x_k)}{\lambda_k} = \frac{\|F_k + J_k(\bar{x}_k - x_k)\|^2 + \lambda_k\|d\|^2}{\lambda_k}\\ \leq &L_1^2\tilde{c}_1\|\bar{x}_k - x_k\|^{4\mp\delta} + \|\bar{x}_k - x_k\|^2 \leq (L_1^2\tilde{c}_1 + 1)\|\bar{x}_k - x_k\|^2.\end{aligned}$$

It implies that

$$\|d_k\| \leq c_2 \text{dist}(x_k, X^*),$$

where $c_2 = \sqrt{L_1^2\tilde{c}_1 + 1}$. The proof is completed.

Lemma 2.3. *Let the sequence $\{x_k\}$ be generate by the adaptive LM method and Assumptions 2.1 hold. Assume $x_k, x_{k+1} \in N(x^*, b/2)$, then*

$$\begin{aligned}c\ \text{dist}(x_{k+1}, X^*) \leq &L_1L_2(2 + 3c_2 + 2c_2^2)\|\bar{x}_k - x_k\|^2\\ &+ L_1L_2^2(2 + c_2)(1 + c_2)^2\|\bar{x}_k - x_k\|^3 + L_2c_2\lambda_k\|\bar{x}_k - x_k\|.\end{aligned}$$

Proof. For all $x_k, x_{k+1} \in N(x^*, b/2)$, we get from (2.4) and (2.5) that

$$\begin{aligned}\|J_k^T F(x_{k+1}) - J_k^T F_k - J_k^T J_k(x_{k+1} - x_k)\| \leq &L_1\|J_k\|\ \|x_{k+1} - x_k\|^2\\ \leq &L_1L_2\ \|x_{k+1} - x_k\|^2,\end{aligned}$$

and

$$\begin{aligned}&\|J_k^T F(x_{k+1}) - J(x_{k+1})^T F(x_{k+1})\\&\quad + J(x_{k+1})^T F(x_{k+1}) - J_k^T F_k - J_k^T J_k(x_{k+1} - x_k)\| \le L_1 L_2 \|x_{k+1} - x_k\|^2.\end{aligned}$$

By the triangle inequality, the above inequality yields

$$\begin{aligned}&\|J(x_{k+1})^T F(x_{k+1}) - J_k^T F_k - J_k^T J_k(x_{k+1} - x_k)\|\\&\le L_1 L_2 \|x_{k+1} - x_k\|^2 + \|(J_k - J(x_{k+1}))^T F(x_{k+1})\|.\end{aligned} \tag{2.11}$$

For all $\bar{x}_k \in X^* \cap N(x^*, b)$, we obtain

$$\begin{aligned}&\|(J_k - J(x_{k+1}))^T F(x_{k+1})\|\\&=\|(J_k - J(\bar{x}_k) + J(\bar{x}_k) - J(x_{k+1}))^T F(x_{k+1})\|\\&\le\|(J_k - J(\bar{x}_k))^T F(x_{k+1})\| + \|(J(\bar{x}_k) - J(x_{k+1}))^T F(x_{k+1})\|\\&\le\|(J_k - J(\bar{x}_k))^T\| \left(\|F(\bar{x}_k) + J(\bar{x}_k)(x_{k+1} - \bar{x}_k)\| + L_2^2\|x_{k+1} - \bar{x}_k\|^2\right)\\&\quad + \|(J(\bar{x}_k) - J(x_{k+1}))^T\| \left(\|F(\bar{x}_k) + J(\bar{x}_k)(x_{k+1} - \bar{x}_k)\|\right.\\&\quad \left.+L_2^2\|x_{k+1} - \bar{x}_k\|^2\right)\\&\le L_1 L_2\|x_k - \bar{x}_k\| \|x_{k+1} - \bar{x}_k\| + L_1 L_2^2\|x_k - \bar{x}_k\| \|x_{k+1} - \bar{x}_k\|^2\\&\quad + L_1 L_2\|x_{k+1} - \bar{x}_k\|^2 + L_1 L_2^2\|x_{k+1} - \bar{x}_k\|^3.\end{aligned} \tag{2.12}$$

Similarly, using the triangle inequality yields

$$\begin{aligned}&\|J(x_{k+1})^T F(x_{k+1}) - J_k^T F_k - J_k^T J_k(x_{k+1} - x_k)\|\\&\ge\|J(x_{k+1})^T F(x_{k+1})\| - \|J_k^T F_k + J_k^T J_k(x_{k+1} - x_k)\|.\end{aligned} \tag{2.13}$$

It follows from (1.2), (2.11) and (2.13) that

$$\begin{aligned}&\|J(x_{k+1})^T F(x_{k+1})\|\\&\le\|J(x_{k+1})^T F(x_{k+1}) - J_k^T F_k - J_k^T J_k(x_{k+1} - x_k)\|\\&\quad + \|J_k^T F_k + J_k^T J_k(x_{k+1} - x_k)\|\\&\le L_1 L_2 \|d_k\|^2 + \|(J_k - J(x_{k+1}))^T F(x_{k+1})\| + \|J_k^T F_k + J_k^T J_k d_k\|\\&\le L_1 L_2 \|d_k\|^2 + \|(J_k - J(x_{k+1}))^T F(x_{k+1})\| + L_2\lambda_k\|d_k\|.\end{aligned} \tag{2.14}$$

From Lemma 2.2, we have $\|d_k\| \le c_2\|\bar{x}_k - x_k\|$, which implies that

$$\|x_{k+1} - \bar{x}_k\| \le \|x_{k+1} - x_k\| + \|x_k - \bar{x}_k\| \le (1 + c_2)\|\bar{x}_k - x_k\|. \tag{2.15}$$

Since $\bar{x}_k \in X^* \cap N(x^*, b)$ and $\delta \in [1, 2]$, together with Assumption 2.1 (b), (2.12), (2.14) and (2.15), we obtain

$$\begin{aligned}c \operatorname{dist}(x_{k+1}, X^*) \le&\|J(x_{k+1})^T F(x_{k+1})\|\\\le&L_1 L_2 c_2^2\|\bar{x}_k - x_k\|^2 + L_1 L_2(1 + c_2)\|x_k - \bar{x}_k\|^2\end{aligned}$$

$$
\begin{aligned}
&+ L_1L_2^2(1+c_2)^2\|\bar{x}_k - x_k\|^3 + L_1L_2(1+c_2)^2\|\bar{x}_k - x_k\|^2\\
&+ L_1L_2^2(1+c_2)^3\|\bar{x}_k - x_k\|^3 + L_2c_2\lambda_k\|\bar{x}_k - x_k\|\\
\le& L_1L_2(2+3c_2+2c_2^2)\|\bar{x}_k - x_k\|^2\\
&+ L_1L_2^2(2+c_2)(1+c_2)^2\|\bar{x}_k - x_k\|^3 + L_2c_2\lambda_k\|\bar{x}_k - x_k\|.
\end{aligned}
$$

The proof is completed.

Henceforth, according to the choices of the LM parameter, namely $\|J_k^T F_k\| \le 1$ and $\|J_k^T F_k\| > 1$, we divide the convergence analysis in two cases.

Case 1: $\|J_k^T F_k\| \le 1$

Firstly, we consider the convergence rate of the adaptive LM method with the LM paramter $\|J_k^T F_k\| \le 1$ in this subsection.

Lemma 2.4. *Let the sequence* $\{x_k\}$ *be generate by the adaptive LM method and Assumptions 2.1 hold. If* $x_k, x_{k+1} \in N(x^*, b/2)$ *and* $\|J_k^T F_k\| \le 1$*, then there exists a positive constant* c_3 *such that*

$$
\mathrm{dist}(x_{k+1}, X^*) \le c_3 \mathrm{dist}(x_k, X^*)^2. \tag{2.16}
$$

Proof. From Lemmas 2.1 and 2.3, we have

$$
\begin{aligned}
c\ \mathrm{dist}(x_{k+1}, X^*) \le& L_1L_2(2+3c_2+2c_2^2)\|\bar{x}_k - x_k\|^2\\
&+ L_1L_2^2(2+c_2)(1+c_2)^2\|\bar{x}_k - x_k\|^3 + L_2c_2\lambda_k\|\bar{x}_k - x_k\|\\
\le& L_1L_2(2+3c_2+2c_2^2)\|\bar{x}_k - x_k\|^2\\
&+ L_1L_2^2(2+c_2)(1+c_2)^2\|\bar{x}_k - x_k\|^3 + L_2c_2c_1^{\delta}\|\bar{x}_k - x_k\|^{1+\delta}.
\end{aligned}
$$

Since $\delta \in [1,2]$, then Lemma 2.4 holds with $c_3 = c^{-1}(L_1L_2(2+3c_2+2c_2^2) + L_2c_2c_1^{\delta} + L_1L_2^2(2+c_2)(1+c_2)^2)$. The proof is completed.

Lemma 2.4 shows that if $x_k \in N(x^*, b/2)$ for all k, then $\{\mathrm{dist}(x_k, X^*)\}$ converges to zero quadratically. Next, we show that the latter theory holds if x_0 is sufficiently close to x^*. Let

$$
r = \min\left\{\frac{b}{2(1+2c_2)}, \frac{1}{2c_3}\right\}. \tag{2.17}
$$

Lemma 2.5. *Let the sequence* $\{x_k\}$ *be generate by the adaptive LM method and Assumptions 2.1 hold. If* $x_0 \in N(x^*, r)$ *with* r *given by (2.17), then for all* k*, we have* $x_k \in N(x^*, b/2)$*.*

Proof. We show the proof by induction. It follows from Lemma 2.2 that

$$
\|x_1 - x^*\| \le \|x_0 - x^*\| + \|d_0\| \le \|x_0 - x^*\| + \|x_0 - \bar{x}_0\| \le (1+c_2)r \le b/2.
$$

It indicates that $x_1 \in N(x^*, b/2)$. Assume for $i = 2, \cdots, k$, $x_i \in N(x^*, b/2)$. It follows from Lemma 2.4 that

$$
\mathrm{dist}(x_i, X^*) \le c_3 \mathrm{dist}(x_{i-1}, X^*)^2 \le \cdots \le c_3^{2^i-1}\|x_0 - x^*\|^{2^i} \le r\left(\frac{1}{2}\right)^{2^i-1},
$$

where the last inequality is derived from $\|x_0 - x^*\| \le r$ and $r \le 1/2c_3$. Therefore, we have from

Lemma 2.2

$$\|d_i\| \leq c_2 \mathrm{dist}(x_i, X^*) \leq c_2 r \left(\frac{1}{2}\right)^{2^i-1} \leq c_2 r \left(\frac{1}{2}\right)^{2i-1}, \tag{2.18}$$

for $i = 1, \cdots, k$. It then follows from (2.17) that

$$\begin{aligned}\|x_{k+1} - x^*\| \leq & \|x_1 - x^*\| + \sum_{i=1}^{k} \|d_i\| \leq (1 + c_2) r + c_2 r \sum_{i=1}^{k} \left(\frac{1}{2}\right)^{2i-1} \\ \leq & (1 + c_2) r + c_2 r \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^{i} \leq (1 + 2c_2) r \leq \frac{b}{2},\end{aligned}$$

which indicates that $x_{k+1} \in N(x^*, b/2)$. The proof is completed.

Theorem 2.1. *Let Assumption 2.1 hold and $\{x_k\}$ be the LM sequence which is generated by the adaptive LM method with $x_0 \in N(x^*, r)$, where r is given by (2.17). If $\|J_k^T F_k\| \leq 1$, then the sequence $\{\mathrm{dist}(x_k, X^*)\}$ converges to zero quadratically. Moreover, $\{x_k\}$ converges to a solution of (1.1).*

Proof. Lemma 2.4 and 2.5 indicates that the sequence $\{\mathrm{dist}(x_k, X^*)\}$ converges to 0 quadratically. So, we only have to prove the second part.

According to the assumption, we have $x_k \in N(x^*, b/2)$ for all k. Then we only have to prove that $\{x_k\}$ converges to some solution $\bar{x} \in X^*$. In fact, for any $p, q \in \mathbb{N}_+$ (let $p \geq q$, we also obtain the same result for $p < q$), from (2.18), we have

$$\|x_p - x_q\| \leq \sum_{i=q}^{p-1} \|d_i\| \leq \sum_{i=q}^{\infty} \|d_i\| \leq c_2 r \sum_{i=q}^{\infty} c_2 r \left(\frac{1}{2}\right)^{2i-1} = \frac{4}{3} c_2 r \left(\frac{1}{2}\right)^{2q-1}. \tag{2.19}$$

The above inequality indicates that the sequence $\{x_k\}$ is a Cauchy sequence, and hence $\{x_k\}$ converges. The proof is completed.

Theorem 2.1 shows that the sequence $\{\mathrm{dist}(x_k, X^*)\}$ converges to zero quadratically and $\{x_k\}$ converges to the solution set X^*. However, little is known about the behaviour of the sequence $\{x_k\}$. In the following theorem, we will see that the sequence $\{x_k\}$ converges to a solution $\bar{x}$ of (1.1), and that the rate of convergence is also locally quadratic.

Theorem 2.2. *Let Assumption 2.1 hold, $\{x_k\}$ be the LM sequence which is generated by the adaptive LM method with $x_0 \in N(x^*, r)$ where r is given by (2.17), and limit point $\hat{x}^* \in X^* \cap N(x^*, b/2)$. If $\|J_k^T F_k\| \leq 1$, then the sequence $\{x_k\}$ converges to $\hat{x}^*$ quadratically.*

Proof. In view of Theorem 2.1, we have $\mathrm{dist}(x_{k+1}, X^*) \leq \frac{1}{2}\mathrm{dist}(x_k, X^*)$ for all sufficiently large k. By letting $p \to \infty$ in (2.19), we deduce from Lemma 2.2 and 2.4 that

$$\begin{aligned}\|\hat{x}^* - x_q\| \leq & \sum_{i=q}^{\infty} \|d_i\| \leq c_2 \sum_{i=q}^{\infty} \mathrm{dist}(x_i, X^*) \leq c_2 \sum_{i=q}^{\infty} \left(\frac{1}{2}\right)^{i-q} \mathrm{dist}(x_q, X^*) \\ \leq & 2c_2 \mathrm{dist}(x_q, X^*) \leq 2c_2 c_3 \mathrm{dist}(x_{q-1}, X^*)^2 \leq 2c_2 c_3 \|\hat{x}^* - x_{q-1}\|^2,\end{aligned}$$

where the last inequality follows from the definition of $\mathrm{dist}(x_k, X^*)$. Hence, the sequence $\{x_k\}$ converges to $\hat{x}^*$ quadratically. The proof is completed.

Case 2: $\|J_k^T F_k\| > 1$

Now, we consider the convergence rate of adaptive LM method with the LM paramter $\|J_k^T F_k\| > 1$.

Lemma 2.6. *Let the sequence* $\{x_k\}$ *be generate by the adaptive LM method and Assumptions 2.1 hold. Assume* $\|J_k^T F_k\| > 1$, *if* $x_k, x_{k+1} \in N(x^*, b/2)$ *and* $\mathrm{dist}(x_k, X^*) \le r < 1$, *where*

$$r = \min\left\{\frac{b}{2(1+\frac{c_2}{1-c_4})}, \frac{c - L_2c_2}{L_1L_2(2+3c_2+2c_2^2) + L_1L_2^2(2+c_2)(1+c_2)^2}\right\} \tag{2.20}$$

with $c > L_2c_2$, *then there exists a positive constant* $c_4 \in (0,1)$ *such that*

$$\mathrm{dist}(x_{k+1}, X^*) \le c_4 \mathrm{dist}(x_k, X^*). \tag{2.21}$$

Proof. From Lemma 2.1, we have $\lambda_k < 1$. Together with Lemma 2.3, we obtain

$$\begin{aligned}
&c\ \mathrm{dist}(x_{k+1}, X^*)\\
\le &L_1L_2(2+3c_2+2c_2^2)\|\bar{x}_k - x_k\|^2\\
&+ L_1L_2^2(2+c_2)(1+c_2)^2\|\bar{x}_k - x_k\|^3 + L_2c_2\lambda_k\|\bar{x}_k - x_k\|\\
\le &\Big(\big(L_1L_2(2+3c_2+2c_2^2) + L_1L_2^2(2+c_2)(1+c_2)^2\big) r + L_2c_2\Big)\|\bar{x}_k - x_k\|,
\end{aligned}$$

which indicates that Lemma 2.6 holds with $c_4 = c^{-1}(L_1L_2(2+3c_2+2c_2^2)+L_1L_2^2(2+c_2)(1+c_2)^2)r + c^{-1}L_2c_2$.

The proof is completed.

Lemma 2.7. *Let the sequence* $\{x_k\}$ *be generate by the adaptive LM method and Assumptions 2.1 hold. If* $x_0 \in N(x^*, r)$ *with r given by (2.20), then for all k, we have* $x_k \in N(x^*, b/2)$ *and* $\mathrm{dist}(x_k, X^*) \le r$.

Proof. Since the proof is analogous to the one of Lemma 2.5, we only verify the inductive step, i.e., assume Lemma 2.7 holds with $i = k$ and consider the next step.

It follows from Lemma 2.6 that

$$\mathrm{dist}(x_{k+1}, X^*) \le c_4\mathrm{dist}(x_k, X^*) \le c_4 r < r \tag{2.22}$$

and

$$\mathrm{dist}(x_{k+1}, X^*) \le c_4\mathrm{dist}(x_k, X^*) \le \cdots \le c_4^{k+1}\mathrm{dist}(x_0, X^*) \le c_4^{k+1} r < r. \tag{2.23}$$

Thus, from Lemma 2.2 and (2.20), we have

$$\begin{aligned}
\|x_{k+1} - x^*\| \le &\|x_1 - x^*\| + \sum_{i=1}^{k}\|d_i\| \le \|x_1 - x^*\| + \sum_{i=1}^{k} c_2\mathrm{dist}(x_i, X^*)\\
\le &(1+c_2)r + c_2 r\sum_{i=1}^{\infty} c_4^i \le (1 + \frac{c_2}{1-c_4})r \le \frac{b}{2},
\end{aligned}$$

which indicates that $x_{k+1} \in N(x^*, b/2)$. The proof is completed.

Theorem 2.3. *Let Assumption 2.1 hold and* $\{x_k\}$ *be the LM sequence which is generated by the adaptive LM method with* $x_0 \in N(x^*, r)$, *where r is given by (2.20). If* $\|J_k^T F_k\| > 1$, *then the sequence* $\{\mathrm{dist}(x_k, X^*)\}$ *converges to zero linearly. Moreover, the sequence* $\{x_k\}$ *converges to a solution* $\hat{x}^* \in X^* \cap N(x^*, b/2)$ *linearly.*

Proof. The proof is similar to the proofs of Theorems 2.1 and 2.2.

3. Global convergence of the adaptive LM method

To establish the global convergence of the adaptive LM method, we employ some line search rules such as Armijo rule, Goldstein rule and Wolfe rule [15] etc. Consider the merit function

$$\Phi(x) = \frac{1}{2}\|F(x)\|^2.$$

At iteration k, the next step is computed by

$$x_{k+1} = x_k + \alpha_k d_k,$$

where d_k is a direction from (2.1) and α_k is a step size satisfying certain line search conditions. The Wolfe line search is one of commonly used inexact line search which requires $\alpha_k > 0$ satisfies

$$\|F(x_k + \alpha_k d_k)\|^2 \leq \|F(x)\|^2 + \sigma_1 \alpha_k F_k^T J_k d_k$$

and

$$F(x_k + \alpha_k d_k)^T J(x_k + \alpha_k d_k) d_k \geq \sigma_2 F_k^T J_k d_k. \tag{3.1}$$

Here $\sigma_1 \leq \sigma_2$ are two constants in $(0, 1)$.

Algorithm 3.1 (The adaptive LM method with Wolfe line search)**.**

Step 1: Given $x_0 \in \mathbb{R}^n$, $\delta \in [1, 2]$, $\eta \in (0, 1)$, $\sigma_1 \in (0, 1/2)$, $\sigma_2 \in (\sigma_1, 1)$, $k := 0$.

Step 2: If $\|J_k^T F_k\| = 0$, *stop. Set* λ_k *as (1.6); determine* d_k *by computing (2.1).*

Step 3: If d_k *satisfies*

$$\|F(x_k + d_k)\| \leq \eta \|F(x_k)\|, \tag{3.2}$$

set $x_{k+1} = x_k + d_k$, *and go to step 5. Otherwise, go to step 4.*

Step 4: Set $x_{k+1} = x_k + \alpha_k d_k$, *where* α_k *is determined by Wolfe line search.*

Step 5: Set $k := k + 1$*; go to Step 2.*

Theorem 3.1. *Assume* $F(x)$ *is continuously differentiable. Let* $\{x_k\}$ *be a sequence generated by Algorithm 3.1. Then any accumulation point* x^* *of* $\{x_k\}$ *is a stationary point of* Φ.

Proof. From [20, Eq (2.10)], the inequality (3.1) implies that

$$\|F(x_{k+1})\|^2 \leq \|F_k\|^2 - \sigma_1 \sigma_3 \frac{(F_k^T J_k d_k)^2}{\|d_k\|^2}, \tag{3.3}$$

where σ_3 is some positive constant. Together with Steps 3 of Algorithm 3.1, the sequence $\{\|f(x_k)\|\}$ is monotonically decreasing and bounded from below, and thus converges to zero. Hence $\{x_k\}$ converges to a stationary point x^* of Φ. The proof is completed.

Theorem 3.2. *Under Assumption 2.1, let* $\{x_k\}$ *be a sequence generated by Algorithm 3.1 and has an accumulation point* x^*. *If* x^* *is a solution of system of nonlinear Eq (1.1), then the sequence* $\{x_k\}$ *converges to* x^* *at least linearly.*

Proof. It is sufficient to show that $\|F(x_k + d_k)\| \le \eta \|F(x_k)\|$ holds for all large k.

If $\|J_k^T F_k\| \le 1$. Since the sequence $\{x_k\}$ converges to a stationary point x^* which is a solution of system of nonlinear Eq (1.1), we have that

$$\|F(x_K)\| \le \frac{c^2\eta}{L_2^3 c_3} \tag{3.4}$$

and

$$\|x_K - x^*\| \le r,$$

hold for all sufficiently large $K \in \mathbb{N}$, where r is defined by (2.17), and c, c_3 and L_2 are given in Section 2.

Let sequence $\{y_k\}$ be generated by the adaptive LM method with unit step size and $y_0 = x_K$. Then, by the result of Theorem 2.1, the sequence $\mathrm{dist}(y_l, X^*)$ quadratic converges to zero. Hence, we only have to prove that $x_{K+l} = y_l$ for all $l \in \mathbb{N}$, i.e., the sequence $\{y_l\}$ satisfies

$$\|F(y_{l+1})\| \le \eta \|F(y_l)\|.$$

Let $\bar{y}_{l+1} \in X^*$ such that $\mathrm{dist}(y_{l+1}, X^*) = \|\bar{y}_{l+1} - y_{l+1}\|$. Then we obtain from Assumption 2.1(b), Lemma 2.4, (2.6) and (3.4) that

$$\begin{aligned}
\|F(y_{l+1})\| =& \|F(\bar{y}_{l+1}) - F(y_{l+1})\| \le L_2 \mathrm{dist}(y_{l+1}, X^*) \\
\le& L_2 c_3 \mathrm{dist}(y_l, X^*)^2 \le \frac{L_2 c_3}{c^2} \|J(y_l)^T F(y_l)\|^2 \\
\le& \frac{L_2^3 c_3}{c^2} \|F(y_l)\|^2 \le \frac{L_2^3 c_3 \|F(y_l)\|}{c^2} \|F(y_l)\| \\
\le& \eta \|F(y_l)\|
\end{aligned}$$

holds for $\eta \in (0, 1)$ and all l. The above inequality indicates that the step size $\alpha_k = 1$ holds for all large k in Algorithm 3.1. We conclude that (3.2) holds for all $k \ge K$. Consequently, by mathematical induction, Algorithm 3.1 reduces to the adaptive LM method for all $k \ge K$. Thus, we have that $\{x_k\}$ converges to the solution x^* quadratically.

Similar to the above process, when $\|J_k^T F_k\| > 1$, we obtain that $\{x_k\}$ converges to the solution x^* linearly.

The proof is completed.

4. Numerical examples

In this section, we carry out some numerical experiments to verify the effectiveness of the proposed adaptive Levenberg-Marquardt method (ALMM). The Levenberg-Marquardt method (LMM) given by Behling et al. [3] is used for comparison. The first test is a nonlinear least squares problem while the second are some systems of nonlinear equations.

Example 4.1. *Consider the nonlinear least squares problem [3]*

$$\min_{x \in \mathbb{R}^n} \quad \Phi(x) = \frac{1}{2}\|F(x)\|^2, \tag{4.1}$$

where $F(x) = (x_1^3 - x_1x_2 + 1, x_1^3 + x_1x_2 + 1)^T$.

Consider $X^* = \{(0, \xi), \xi \in \mathbb{R}\}$ *be the non-isolated set of minimizers such that* $\operatorname{dist}(x, X^*) = |x_1|$. *Then the rank of the Jacobian will be* 0 *at the origin,* 1 *at* x *with* $x_1 = 0$ *for* $x_2 \neq 0$, *and* 2 *for* $x_1 \neq 0$. *Thus the Jacobian is not always of full rank at the stationary points. The starting point is set to be* $x_0 = (0.008, 2)^T$. *All methods terminate if* $\|J_k^T F_k\| < 10^{-10}$. *The results are tabulated in Table 1.*

As illustrated, ALMM generally converges to the required accuracy with less iterations than LMM. Besides, distances between x_k *obtained from ALMM and the solution set* X^* *are shorter than those from LMM.*

Table 1. Numerical results for nonlinear least-squares problem.

LMM				ALMM			
δ	Iters	$\operatorname{dist}(x_k, X^*)$	$\|J_k^T F_k\|$	δ	Iters	$\operatorname{dist}(x_k, X^*)$	$\|J_k^T F_k\|$
10^{-4}	0	8.0000e-03	6.4385e-02	1	0	8.0000e-03	6.4385e-02
	2	9.3495e-05	7.4799e-04		1	1.6286e-05	1.3029e-04
	5	1.2786e-07	1.0228e-06		2	6.6308e-11	5.3046e-10
	8	1.7465e-10	1.3972e-09		3	1.0899e-17	0
	10	2.1481e-12	1.7185e-11				
0.5	0	8.0000e-03	6.4385e-02	1.5	0	8.0000e-03	6.4385e-02
	1	1.9951e-04	1.5963e-03		1	3.1845e-05	2.5477e-04
	2	9.6178e-07	7.6941e-06		2	7.7713e-10	6.2174e-09
	3	3.3268e-10	2.6613e-09		3	7.4217e-17	8.8818e-16
	4	2.1187e-15	1.6875e-14				
1	0	8.0000e-03	6.4385e-02	2	0	8.0000e-03	6.4385e-02
	1	1.6286e-05	1.3029e-04		1	4.5185e-05	3.6159e-04
	2	6.6308e-11	5.3046e-10		2	1.5793e-09	1.2639e-08
	3	1.0899e-17	0		3	5.0847e-18	0

Example 4.2. *Consider systems of nonlinear equations adapted from the nonsingular problems given in [12, 16]*

$$\hat{F}(x) = F(x) - J(x^*)A(A^TA)^{-1}A^T(x - x^*) = 0,$$

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* $1 \leq k \leq n$. *It is easy to check that* $\hat{F}(x^*) = 0$ *and the rank of* $\hat{J}(x^*) = J(x^*)(I - A(A^TA)^{-1}A^T)$ *is* $n - k$. *A disadvantage of these problems is that* $\hat{F}(x)$ *may have roots that are not roots of* $F(x)$. *We present two sets of singular problems with the rank of* $\hat{J}(x^*)$ *being* $n - 1$ *and* $n - 2$, *respectively. The corresponding matrices of A and* A^T *are given by*

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

and

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

Note that the size of the original problem which has $n + 2$ *equations in* n *unknowns is reduced by eliminating the* $(n - 1)$*st and the nth equations.*

Several choices of the LM parameter are considered in the two LM methods. In accordance with the range of δ defined in LMM and ALMM, we use $\delta = 10^{-4}$, 0.5 and 1 associated with $\lambda_k = \|J_k^T F_k\|^\delta$ for LMM and employ $\delta = 1$, 1.5 and 2 for ALMM. All algorithms are terminated if $\|J_k^T F_k\| < 10^{-6}$ or the number of the iterations exceeds $100(n + 1)$. Numerical results for the rank $n - 1$ case and the rank $n - 2$ case are listed in Table 2 and in Table 3, respectively. The values 1, 10 and 100 in the third column associate with starting points with x_0, $10x_0$, and $100x_0$, where x_0 is the option suggested in [12]. The symbol "–" is used if the corresponding method fails to reach the required accuracy within the prescribed maximum iterations. To ensure the numerical stability, we use the MATLAB function `pcg` *(the preconditioned conjugate gradient method) to solve the inner linear system (1.2).*

Some remarks are in order. In all tests, ALMM converges to the required accuracy within the maximum iterations while LMM fails for some cases; see, for instance, Powell badly scaled problem in Table 2 and Discrete integral equation problem in Table 3. Furthermore, the number of iteration step required by ALMM is less than that by LMM. For this reason, we conclude that ALMM is a competitive variant of the Levenberg-Marquardt method.

Table 2. Numerical results of the first singular test with rank$(F'(x^*)) = n - 1$.

Function	n	x_0	LMM			ALMM		
			$\delta = 10^{-4}$	$\delta = 0.5$	$\delta = 1$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$
			Iters/Fun./Times	Iters/Fun./Times	Iters/Fun./Times	Iters/Fun./Times	Iters/Fun./Times	Iters/Fun./Times
Rosenbrock	2	1	–	145/1.0994e-04/0.05	31/5.2121e-05/0.01	21/6.3319e-05/0.02	13/8.1986e-05/0.01	11/1.6249e-04/0.01
		10	–	165/1.0996e-04/0.05	64/5.9032e-05/0.02	17/3.5338e-04/0.01	15/1.3638e-04/0.00	14/1.8112e-04/0.00
		100	–	215/1.1005e-04/0.06	291/3.1076e-04/0.07	24/4.0893e-05/0.01	19/1.4436e-04/0.01	17/1.6579e-04/0.01
Powell badly scaled	2	1	–	–	46/2.1150e-05/0.02	16/1.6882e-05/0.01	17/2.0346e-06/0.01	22/3.0992e-08/0.01
		10	–	–	43/5.9893e-05/0.01	3/4.3848e-08/0.00	3/4.3848e-08/0.00	3/4.3848e-08/0.00
		100	–	–	–	3/4.1815e-08/0.00	3/4.1815e-08/0.00	3/4.1814e-08/0.00
Wood	4	1	–	68/8.2258e-05/0.03	26/2.5450e-04/0.01	16/1.0639e-04/0.01	22/2.6387e-07/0.01	19/3.4455e-07/0.01
		10	–	73/8.4198e-05/0.03	79/1.1002e-04/0.02	19/9.3086e-05/0.01	25/2.0022e-07/0.01	22/3.7076e-07/0.01
		100	–	94/8.6162e-05/0.03	–	23/1.5300e-04/0.01	27/4.8468e-07/0.02	25/4.7057e-07/0.01
Helical valley	3	1	395/8.2335e-05/0.18	36/1.7042e-05/0.02	22/1.6542e-05/0.01	14/2.2764e-07/0.01	11/2.2832e-08/0.00	10/2.9028e-11/0.00
		10	396/8.2458e-05/0.15	39/2.3758e-05/0.02	37/1.7814e-05/0.01	13/3.8763e-09/0.00	10/1.2850e-08/0.00	9/2.3258e-09/0.00
		100	386/8.3458e-05/0.25	40/1.0493e-05/0.01	138/6.9960e-09/0.02	13/1.2005e-09/0.01	9/3.8526e-06/0.00	9/7.9799e-10/0.00
Brown almost-linear	10	1	323/1.7755e-04/0.09	11/1.4159e-04/0.00	9/1.3099e-04/0.00	7/1.3034e-04/0.00	7/9.2295e-05/0.00	7/8.2906e-05/0.00
		10	327/1.7624e-04/0.06	25/1.3093e-04/0.01	35/1.0117e-04/0.01	22/1.2089e-04/0.00	22/9.7275e-05/0.01	22/9.1952e-05/0.01
		100	349/1.7616e-04/0.05	47/1.2040e-04/0.01	200/1.1852e-04/0.03	44/9.6552e-05/0.01	44/7.6970e-05/0.01	44/7.2340e-05/0.01
Discrete boundary value	10	1	59/1.7216e-04/0.03	6/1.6987e-04/0.01	3/1.6852e-04/0.00	3/1.6852e-04/0.00	4/1.3377e-05/0.00	2/1.2224e-04/0.00
		10	–	306/1.7513e-03/0.14	21/3.2163e-04/0.02	19/2.3639e-04/0.01	11/1.1817e-05/0.01	9/5.4119e-06/0.01
		100	–	77/7.0660e-05/0.03	62/8.1935e-07/0.02	20/4.6172e-05/0.01	14/5.9162e-09/0.01	11/6.4234e-06/0.01
Discrete integral equation	30	1	–	31/9.2503e-04/0.05	7/1.1033e-04/0.02	7/1.1033e-04/0.02	6/1.1846e-05/0.02	5/1.3736e-05/0.01
		10	–	109/9.2782e-04/0.15	24/9.2831e-05/0.04	22/6.1841e-05/0.05	14/8.2706e-06/0.03	11/1.2210e-05/0.02
		100	49/1.5445e-05/0.07	22/4.7434e-07/0.03	97/1.1979e-06/0.12	12/1.2250e-08/0.03	10/2.3696e-06/0.02	10/3.0014e-09/0.02
Variably dimensioned	10	1	30/4.3266e-05/0.02	13/3.0661e-05/0.00	16/1.0323e-05/0.00	13/2.2903e-05/0.01	13/2.2553e-05/0.01	13/2.2472e-05/0.01
		10	44/1.9588e-04/0.02	15/1.2677e-04/0.01	35/2.4191e-05/0.01	15/1.1615e-05/0.01	15/1.1407e-05/0.00	15/1.1345e-05/0.01
		100	–	29/2.0406e-04/0.02	249/1.3347e-05/0.05	18/3.8117e-05/0.01	18/3.7443e-05/0.01	18/3.7241e-05/0.01
Broyden tridiagonal	30	1	1676/4.0928e-04/1.49	25/3.5494e-04/0.03	12/1.9277e-05/0.02	10/2.9073e-05/0.01	9/1.6273e-05/0.02	9/1.1863e-05/0.02
		10	1681/4.0933e-04/1.48	31/3.7621e-04/0.03	66/3.0745e-05/0.04	15/2.8837e-05/0.02	14/1.4125e-05/0.01	14/9.5072e-06/0.02
		100	1685/4.0919e-04/1.50	35/3.6087e-04/0.03	564/2.0700e-05/0.15	18/3.8010e-05/0.03	17/1.7068e-05/0.01	17/1.0588e-05/0.02
Broyden banded	30	1	468/2.1508e-04/0.37	17/1.0301e-04/0.02	15/3.3592e-06/0.02	13/2.6804e-06/0.03	12/1.7694e-06/0.02	12/1.3777e-06/0.02
		10	474/2.1511e-04/0.37	23/1.1049e-04/0.02	71/6.8476e-06/0.04	19/3.0285e-06/0.04	18/2.0737e-06/0.02	18/1.6308e-06/0.02
		100	480/2.1493e-04/0.38	29/9.8317e-05/0.02	571/5.1853e-06/0.19	24/5.9445e-06/0.03	23/4.5273e-06/0.02	23/3.4937e-06/0.02

5. Conclusions

We present a Levenberg-Marquardt method with an adaptive LM parameter for solving systems of nonlinear equations. We have analyzed its local and global convergence under a new error bound condition of function, which can be derived from the local error bound condition, and Lipschitz continuity of the Jacobian. These properties hold in many applied problems, as they are satisfied by any real analytic function. The effectiveness of the adaptive Levenberg-Marquardt method is validated by the numerical examples.

Table 3. Numerical results of the second singular test with rank$(F'(x^*)) = n - 2$.

Function	n	x_0	LMM			ALMM		
			$\delta = 10^{-4}$	$\delta = 0.5$	$\delta = 1$	$\delta = 1$	$\delta = 1.5$	$\delta = 2$
			Iters/Fun./Times	Iters/Fun./Times	Iters/Fun./Times	Iters/Fun./Times	Iters/Fun./Times	Iters/Fun./Times
Rosenbroc	2	1	191/1.3540e-04/0.04	12/7.5794e-05/0.00	12/1.2508e-04/0.00	10/6.1241e-05/0.03	10/5.2300e-05/0.00	10/4.9886e-05/0.01
		10	194/1.3508e-04/0.03	14/1.2086e-04/0.00	27/3.6441e-05/0.01	12/1.3200e-04/0.01	12/1.1282e-04/0.00	12/1.0763e-04/0.00
		100	197/1.3524e-04/0.03	18/6.8028e-05/0.00	139/6.3285e-05/0.02	16/4.4471e-05/0.01	16/3.7792e-05/0.00	16/3.5998e-05/0.01
Powell badly scaled	2	1	–	–	24/2.1152e-03/0.01	15/1.8965e-03/0.01	9/2.0698e-03/0.00	9/1.8361e-03/0.01
		10	2/3.3652e-05/0.00	2/3.3652e-05/0.00	2/3.3652e-05/0.00	2/3.3652e-05/0.00	2/3.3648e-05/0.00	2/3.3541e-05/0.00
		100	2/9.9781e-03/0.00	2/9.9781e-03/0.00	2/9.9781e-03/0.00	4/8.8941e-03/0.00	3/6.0819e-05/0.00	3/4.0610e-05/0.00
Wood	4	1	244/1.5339e-04/0.10	15/6.9391e-05/0.01	20/2.7220e-06/0.01	13/6.7353e-06/0.01	13/4.0105e-06/0.01	13/3.7030e-06/0.02
		10	247/1.5336e-04/0.09	18/6.7864e-05/0.01	62/3.7183e-06/0.02	16/6.2995e-06/0.02	16/3.7303e-06/0.01	16/3.4399e-06/0.02
		100	250/1.5354e-04/0.08	21/7.8180e-05/0.01	448/6.9661e-06/0.07	19/9.5795e-06/0.01	19/5.9123e-06/0.01	19/5.5402e-06/0.02
Helical valley	3	1	–	74/9.6343e-05/0.03	26/7.3446e-05/0.01	16/6.9973e-05/0.01	19/1.1355e-06/0.01	16/5.2337e-07/0.01
		10	–	80/9.8653e-05/0.03	40/1.7523e-04/0.01	15/1.3215e-05/0.01	11/3.3338e-09/0.01	10/2.0070e-06/0.01
		100	–	97/9.6417e-05/0.03	166/1.3437e-04/0.03	11/3.5474e-07/0.01	10/1.0864e-07/0.01	10/1.1739e-07/0.01
Brown almost-linear	10	1	323/1.7755e-04/0.05	11/1.4159e-04/0.00	9/1.3099e-04/0.00	7/1.3034e-04/0.01	7/9.2295e-05/0.00	7/8.2906e-05/0.01
		10	327/1.7624e-04/0.05	25/1.3093e-04/0.01	35/1.0117e-04/0.01	22/1.2089e-04/0.01	22/9.7275e-05/0.01	22/9.1952e-05/0.02
		100	349/1.7616e-04/0.05	47/1.2040e-04/0.01	200/1.1852e-04/0.03	44/9.6552e-05/0.01	44/7.6970e-05/0.02	44/7.2340e-05/0.02
Discrete boundary value	10	1	52/1.7225e-04/0.02	6/1.7033e-04/0.01	3/1.6895e-04/0.00	3/1.6895e-04/0.00	4/1.3268e-05/0.01	2/1.2316e-04/0.00
		10	–	307/1.7442e-03/0.15	21/2.9159e-04/0.02	18/3.3749e-04/0.03	11/1.0832e-05/0.02	9/5.5811e-06/0.02
		100	–	96/1.7953e-04/0.04	63/5.2469e-06/0.02	21/1.0307e-04/0.03	14/4.2442e-06/0.01	12/1.0017e-07/0.02
Discrete integral equation	30	1	–	31/9.2504e-04/0.05	7/1.1033e-04/0.02	7/1.1033e-04/0.03	6/1.1846e-05/0.02	5/1.3736e-05/0.02
		10	–	109/9.2784e-04/0.16	24/9.2835e-05/0.04	22/6.1844e-05/0.10	14/8.2708e-06/0.04	11/1.2211e-05/0.04
		100	–	98/4.5114e-03/0.15	112/1.5727e-03/0.13	22/2.3434e-03/0.06	19/1.6518e-05/0.06	14/2.9746e-05/0.05
Variably dimensioned	10	1	30/4.2924e-05/0.02	13/3.0590e-05/0.01	16/1.0322e-05/0.01	13/2.2897e-05/0.01	13/2.2549e-05/0.01	13/2.2469e-05/0.01
		10	32/4.1881e-05/0.01	15/2.2471e-05/0.01	35/1.6634e-05/0.01	15/1.1612e-05/0.01	15/1.1406e-05/0.02	15/1.1344e-05/0.02
		100	36/4.1593e-05/0.01	22/1.1535e-05/0.01	246/1.7497e-05/0.04	18/3.8108e-05/0.01	18/3.7440e-05/0.01	18/3.7239e-05/0.02
Broyden tridiagonal	30	1	1676/4.0925e-04/1.45	25/3.5485e-04/0.03	12/1.9268e-05/0.02	10/2.9527e-05/0.02	9/1.6388e-05/0.03	9/1.1923e-05/0.03
		10	1681/4.0930e-04/1.48	31/3.7612e-04/0.03	66/3.0634e-05/0.04	15/2.8724e-05/0.03	14/1.4120e-05/0.03	14/9.5108e-06/0.04
		100	1685/4.0916e-04/1.49	35/3.6079e-04/0.03	564/2.0621e-05/0.15	18/3.7836e-05/0.02	17/1.7045e-05/0.03	17/1.0580e-05/0.03
Broyden banded	30	1	468/2.1499e-04/0.40	17/1.0294e-04/0.03	15/3.3561e-06/0.02	13/2.6796e-06/0.02	12/1.7692e-06/0.04	12/1.3776e-06/0.04
		10	474/2.1502e-04/0.38	23/1.1041e-04/0.02	71/6.8405e-06/0.04	19/3.0266e-06/0.03	18/2.0731e-06/0.03	18/1.6305e-06/0.04
		100	480/2.1484e-04/0.39	29/9.8244e-05/0.02	571/5.1799e-06/0.22	24/5.9415e-06/0.03	23/4.5260e-06/0.04	23/3.4930e-06/0.04

Acknowledgements

The work of Lin Zheng was supported by the Natural Science Foundation of the Higher Education Institutions of Anhui Province grant KJ2020A0017. The work of Liang Chen was supported by the Abroad Visiting of Excellent Young Talents in Universities of Anhui Province grant GXGWFX2019022 and the Natural Science Foundation of the Higher Education Institutions of Anhui Province grant KJ2020ZD008. The work of Yanfang Ma was supported by the Natural Science Foundation of Anhui Province grant 2108085MF204 and the Natural Science Foundation of the Higher Education Institutions of Anhui Province grant KJ2019A0604. Part of this work was done while Liang Chen and Yanfang Ma were visiting scholars at Department of Mathematics, the University of Texas at Arlington from August 2019 to August 2020. They would like to thank Prof. Ren-Cang Li for his hospitality during the visit.

Conflict of interest

The authors declare no conflict of interest.

References

1. M. Ahookhosh, F. J. A. Artacho, R. M. T. Fleming, P. T. Vuong, Local convergence of the Levenberg–Marquardt method under Hölder metric subregularity, *Adv. Comput. Math.*, **45** (2019), 2771–2806. doi: 10.1007/s10444-019-09708-7.

2. K. Amini, F. Rostami, G. Caristi, An efficient Levenberg-Marquardt method with a new LM parameter for systems of nonlinear equations, *Optimization*, **67** (2018), 637–650. doi: 10.1080/02331934.2018.1435655.

3. R. Behling, D. S. Gonçalves, S. A. Santos, Local convergence analysis of the Levenberg–Marquardt framework for Nonzero-Residue nonlinear least-squares problems under an error bound condition, *J. Optim. Theory Appl.*, **183** (2019), 1099–1122. doi: 10.1007/s10957-019-01586-9.

4. J. E. Dennis, R. B. Schnable, *Numerical methods for unconstrained optimization and nonlinear equations*, 1983. doi: 10.1137/1.9781611971200.

5. J. Y. Fan, J. Y. Pan, A note on the Levenberg-Marquardt parameter, *Appl. Math. Comput.*, **207** (2009), 351–359. doi: 10.1016/j.amc.2008.10.056.

6. J. Y. Fan, Y. X. Yuan, On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption, *Computing*, **74** (2005), 23–39. doi: 10.1007/s00607-004-0083-1.

7. F. Andreas, Local behavior of an iterative framework for generalized equations with nonisolated solutions, *Math. Program.*, **94** (2002), 91–124. doi: 10.1007/s10107-002-0364-4.

8. L. Guo, G. H. Lin, J. J. Ye, Solving mathematical programs with equilibrium constraints, *J. Optim. Theory Appl.*, **166** (2015), 234–256. doi: 10.1007/s10957-014-0699-z.

9. M. Heydari, T. D. Niri, M. M. Hosseini, A new modified trust region algorithm for solving unconstrained optimization problems, *J. Math. Ext.*, **12** (2018), 115–135.

10. E. W. Karas, S. A. Santos, B. F. Svaiter, Algebraic rules for computing the regularization parameter of the Levenberg–Marquardt method, *Comput. Optim. Appl.*, **65** (2016), 723–751. doi: 10.1007/s10589-016-9845-x.

11. C. F. Ma, L. H. Jiang, Some research on Levenberg-Marquardt method for the nonlinear equations, *Appl. Math. Comput.*, **184** (2007), 1032–1040. doi: 10.1016/j.amc.2006.07.004.

12. J. J. Moré, B. S. Garbow, K. E. Hillstrom, Testing unconstrained optimization software, *ACM T. Math. Software*, **7** (1981), 17–41. doi: 10.1145/355934.355936.

13. T. D. Niri, M. Heydari, M. M. Hosseini, Correction of trust region method with a new modified Newton method, *Int. J. Comput. Math.*, **97** (2020), 1118–1132. doi: 10.1080/00207160.2019.1607844.

14. H. V. Ngai, Global error bounds for systems of convex polynomials over polyhedral constraints, *SIAM J. Optim.*, **25** (2015), 521–539. doi: 10.1137/13090599X.

15. J. Nocedal, S. J. Wright, *Numerical optimization*, New York: Springer, 1999. doi: 10.1007/978-3-540-35447-5.

16. R. B. Schnabel, P. D. Frank, Tensor methods for nonlinear equations, *SIAM J. Numer. Anal.*, **21** (1984), 815–843. doi: 10.1137/0721054.

17. H. H. Vui, Global Hölderian error bound for nondegenerate polynomials, *SIAM J. Optim.*, **23** (2013), 917–933. doi: 10.1137/110859889.

18. H. Y. Wang, J. Y. Fan, Convergence rate of the Levenberg-Marquardt method under Hölderian local error bound, *Optim. Method. Softw*, **35** (2020), 767–786. doi: 10.1080/10556788.2019.1694927.

19. N. Yamashita, M. Fukushima, On the rate of convergence of the Levenberg-Marquardt method, In: *Topics in numerical analysis*, Vienna: Springer, **15** (2001), 239–249. doi: 10.1007/978-3-7091-6217-0_18.

20. Y. X. Yuan, Problems on convergence of unconstrained optimization algorithms, In: *Numerical linear algebra and optimization*, Beijing: Science Press, 1999, 95–107.

21. X. D. Zhu, G. H. Lin, Improved convergence results for a modified Levenberg-Marquardt method for nonlinear equations and applications in MPCC, *Optim. Method. Software*, **31** (2016), 791–804. doi: 10.1080/10556788.2016.1171863.

AIMS Press