Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 5608195, 11 pages
https://doi.org/10.1155/2021/5608195

Research Article

A New Modified Efficient Levenberg–Marquardt Method for Solving Systems of Nonlinear Equations

Zhenxiang Wu,[1] Tong Zhou,[1] Lei Li,[1] Liang Chen ,[1,2] and Yanfang Ma[3,4]

[1]*School of Mathematical Sciences, Huaibei Normal University, Huaibei 235000, China*
[2]*School of Sciences, Changzhou Institute of Technology, Changzhou, Jiangsu 213032, China*
[3]*School of Computer Science and Technology, Huaibei Normal University, Huaibei 235000, China*
[4]*School of Computer Science and Information Engineering, Changzhou Institute of Technology, Changzhou, Jiangsu 213032, China*

Correspondence should be addressed to Liang Chen; clmyf2@163.com

Received 7 June 2021; Accepted 14 August 2021; Published 31 August 2021

Academic Editor: Lei Liu

For systems of nonlinear equations, a modified efficient Levenberg–Marquardt method with new LM parameters was developed by Amini et al. (2018). The convergence of the method was proved under the local error bound condition. In order to enhance this method, using nonmonotone technique, we propose a new Levenberg–Marquardt parameter in this paper. The convergence of the new Levenberg–Marquardt method is shown to be at least superlinear, and numerical experiments show that the new Levenberg–Marquardt algorithm can solve systems of nonlinear equations effectively.

1. Introduction

Consider the system of nonlinear equations

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

where the function $F(x)$: $\mathbf{R}^n \longrightarrow \mathbf{R}^m$ is continuously differentiable. In this paper, we assume that the solution set of (1) (denoted by X^*) is nonempty, with $\|\cdot\|$ referring to the 2-norm.

Newton method is an important method to solve system (1) in [1]. At each iteration, it uses the trial step

$$d_k^N = -J_k^{-1} F_k, \tag{2}$$

where $F(x_k) = F_k$ and $J_k = F'(x_k)$ is Jacobian matrix. When $J(x)$ is Lipschitz continuous and nonsingular, then the convergence of this method is quadratic at the solution. However, trial step d_k^N may not exist and J_k is singular or near singular. Newton method may not be well defined. To overcome this difficulty, Levenberg–Marquardt (LM) method was created by Levenberg [2] and Marquardt [3] which uses the trial step d_k^{LM} at each iteration, where

$$d_k^{\mathrm{LM}} = -\left(J_k^T J_k + \lambda_k I\right)^{-1} J_k^T F_k, \tag{3}$$

and λ_k is a nonnegative constant. By introducing a nonnegative parameter λ_k, LM method overcomes the problem that J_k is singular or near singular; furthermore, excessive step size $\|d_k\|$ is avoided. In this case, where $\lambda_k = 0$ and Jacobian matrix J_k is nonsingular, the LM method is reduced to Newton method.

The efficiency of the LM method is affected by the parameter λ_k. For example, let $\lambda_k = \|F(x_k)\|^2$, under the local error bound condition, the LM method is shown to have quadratic convergence by Yamashita and Fukushima in [4]. However, when the sequence $\{x_k\}$ is far away from the set X^*, $\|F_k\|$ may be very big which may lead to large λ_k. It will result in a smaller LM step size, further reducing the efficiency of the algorithm. In [5], Fan used $\lambda_k = \mu_k \|F(x_k)\|^\delta$, $\delta \in (1, 2]$, where μ_k is updated with trust region technology in each iteration, the LM method also has quadratic convergence under some suitable conditions, and $\lambda_k = \mu_k \|F(x_k)\|^\delta$ can alleviate the effect of the initial point being far away from the set X^*.

To avoid this trouble, Amini used $\lambda_k = \mu_k\|F_k\|/(1+\|F_k\|)$ in [6]; when the sequence $\{x_k\}$ is far from the solution set and $\|F_k\|$ is very large, λ_k is close to μ_k, which effectively controls the range of λ_k. Umar proposed some new LM parameters $\lambda_k = (\|J_k\|/\|J_k\|^2)$, $\lambda_k = (\|J_k^T J_k\|/\|J_k^T J_k\|^2)$ in [7]. Wang used $\lambda_k = \eta_k\|J_k^T F_k\|^{\alpha}$ with η_k updated by trust region techniques from iteration to iteration in [8].

Ma introduced $\|J_k^T F_k\|$ into LM method and used a new LM parameter $\lambda_k = \theta\|F_k\| + (1-\theta)\|J_k^T F_k\|$ in [9], where $0 \le \theta \le 1$. It is noticeable that λ_k is a convex combination of $\|F_k\|$ and $\|J_k^T F_k\|$, and the quadratic convergence of this method is proved. There are numerous other various LM methods to solve (1); interested readers are referred to [10–12] for related work. In order to discuss the range of parameter λ_k, inspired by [6, 8, 9], in this paper, we choose a new LM parameter as follows:

$$\lambda_k = \mu_k\left(\theta\frac{\|F_k\|^{\delta}}{1+\|F_k\|^{\delta}} + (1-\theta)\frac{\|J_k^T F_k\|^{\delta}}{1+\|J_k^T F_k\|^{\delta}}\right), \quad \delta \in (0,2], \theta \in [0,1], \tag{4}$$

where λ_k is a convex combination of $\|F_k\|^{\delta}/(1+\|F_k\|^{\delta})$ and $\|J_k^T F_k\|^{\delta}/(1+\|J_k^T F_k\|^{\delta})$ and μ_k is updated with trust region technology in each iteration.

Now, we set

$$\phi(x) = \|F(x)\|^2 \tag{5}$$

as the merit function for (1). We define the actual reduction and the predicted reduction of $\phi(x)$ at the kth iteration as follows:

$$\text{Ared}_k = \|F_k\|^2 - \|F(x_k + d_k)\|^2, \tag{6}$$

$$\text{Pred}_k = \|F_k\|^2 - \|F_k + J_k d_k\|^2, \tag{7}$$

where d_k is computed by (3). The following ratio is

$$r_k = \frac{\text{Ared}_k}{\text{Pred}_k}. \tag{8}$$

Grippo applied the nonmonotone line search technique to Newton's method in [13]; some authors have extended the nonmonotone techniques to trust region algorithm and proposed a lot of effective nonmonotone trust region methods in [14, 15]. And Amini proposed nonmonotone line search technique for the LM method in [6]. Numerical experiments show that the algorithm with the nonmonotone technique is more efficient than the algorithm without the nonmonotone technique. Inspired by these theories, we apply a nonmonotone strategy to LM method in this paper. Let us replace actual reduction (6) with the following actual reduction:

$$\overline{\text{A}}\text{red}_k = F_{l(k)}^2 - \|F(x_k + d_k)\|^2, \tag{9}$$

where

$$F_{l(k)} = \max_{0\le j\le n(k)}\left\{\|F_{k-j}\|\right\}, \quad k = 0,1,\ldots, \tag{10}$$

$n(k) = \min\{N_0, k\}$, and N_0 is a positive integer constant. Obviously, by this change, $\|F_{k+1}\|$ will be compared with the $\max_{0\le j\le n(k)}\{\|F_{k-j}\|\}$ in each iteration, further leading to affect the ratio. The ratio after the change is

$$\overline{r}_k = \frac{\overline{\text{A}}\text{red}_k}{\text{Pred}_k}. \tag{11}$$

It can be used to decide whether the trial step is accepted and update the trust region parameter μ_k.

The paper is organized as follows. In Section 2, we present a new algorithm and then prove the global convergence of the new algorithm under some conditions. In Section 3, under the local error bound condition, the convergence of the new Levenberg–Marquardt method is shown to be at least superlinear. In Section 4, the new algorithm is an effective algorithm, which is demonstrated by numerical results. At last, we give some conclusions in Section 5.

2. The Efficient Algorithm and Global Convergence

In this section, firstly, we present the new efficient LM algorithm and then prove the global convergence of the new algorithm.

When sequence $\{x_k\}$ is close to the solution, the steps may be too large, so we require

$$\mu_k \ge m, \tag{12}$$

in the new algorithm, where m is a positive constant, and this is implemented by Step 5.

Lemma 1. *For all $k \in N$, we have*

$$\text{Pred}_k \ge \|J_k^T F_k\| \min\left\{\|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}. \tag{13}$$

Proof. This proof is directly derived from the important theory given by Powell in [16].

From literature [6], the following lemma can be obtained. □

Lemma 2. *Suppose the sequence $\{x_k\}$ be generated by Algorithm 1, then the sequence $\{F_{l(k)}\}$ converges.*

Next we present some of the assumptions needed in the following content.

Assumption 1

(a) $F(x)$ is continuously differentiable and Lipschitz continuous; i.e., there exists a positive constant L_2 that makes

$$\|F(y) - F(x)\| \le L_2\|y - x\|. \tag{14}$$

(b) $J(x)$ is Lipschitz continuous; i.e., there exists a positive constant L_1 such that

$$\|J(y)-J(x)\| \le L_1\|y-x\|. \tag{15}$$

Lemma 3. *If Assumption 1 holds, then we have*

$$\|F(y)-F(x)-J(x)(y-x)\| \le L_1\|y-x\|^2, \tag{16}$$

$$J(x) \le L_2. \tag{17}$$

Proof. The proof of (17) can be found in [17]. So, we only prove (16). Using mean value theorem, there exists $z \in [x, y]$ that makes

$$F(y) = F(x) + J(x)(y-x), \tag{18}$$

and hence,

$$F(y)-F(x)-J(x)(y-x) = (J(z)-J(x))(y-x). \tag{19}$$

According to the last equation, we can obtain

$$\begin{aligned}\|F(y)-F(x)-J(x)(y-x)\| &\le \|(J(z)-J(x))\|\|(y-x)\| \\ &\le L_1\|(z-x)\|\|(y-x)\| \\ &\le L_1\|(y-x)\|^2.\end{aligned} \tag{20}$$

So (16) is true. □

Theorem 1. *Suppose that Assumption 1 is true. Then, Algorithm 1 terminates in finite iterations or satisfies*

$$\liminf_{k\longrightarrow\infty}\left\|J_k^T F_k\right\| = 0. \tag{21}$$

Proof. Assume that the theorem is incorrect; then, there exist a positive constant ε_0 and a constant $k_0 \in N$ that makes

$$\left\|J_k^T F_k\right\| \ge \varepsilon_0, \quad \forall k \ge k_0. \tag{22}$$

Firstly, we prove that

$$\lim_{k\longrightarrow\infty}\left\|F\left(x_{l(k)}\right)\right\| = \lim_{k\longrightarrow\infty}\left\|F\left(x_k\right)\right\|. \tag{23}$$

Since d_k is accepted by the algorithm, we have

$$F_{l(k)}^2 - \left\|F\left(x_k+d_k\right)\right\|^2 \ge p_0 \text{Pred}_k. \tag{24}$$

This, along with (17), (22), and Lemma 1, for all $k \ge k_0$, that means

$$\begin{aligned}F_{l(k)}^2 - \left\|F_{k+1}\right\|^2 &\ge p_0\left\|J_k^T F_k\right\|\min\left\{\left\|d_k\right\|, \frac{\left\|J_k^T F_k\right\|}{\left\|J_k^T J_k\right\|}\right\} \\ &\ge p_0\varepsilon_0 \min\left\{\left\|d_k\right\|, \frac{\varepsilon_0}{L_2^2}\right\}.\end{aligned} \tag{25}$$

Replacing k with $l(k)-1$, for all sufficiently large k, there is

$$F_{l(l(k)-1)}^2 - \left\|F_{l(k)}\right\|^2 \ge p_0 \min\left\{\left\|d_{l(k)-1}\right\|, \frac{\varepsilon_0}{L_2^2}\right\}. \tag{26}$$

From Lemma 2

$$\lim_{k\longrightarrow\infty}\left(\left\|F_{l(l(k)-1)}^2\right\| - \left\|F_{l(k)}\right\|^2\right) = 0, \tag{27}$$

which together with the last inequality yields

$$\lim_{k\longrightarrow\infty} \min\left\{\left\|d_{l(k)-1}\right\|, \frac{\varepsilon_0}{L_2^2}\right\}. \tag{28}$$

ε_0/L_1^2 is a positive constant, so

$$\lim_{k\longrightarrow\infty}\left\|d_{l(k)-1}\right\| = 0. \tag{29}$$

Using Assumption 1, the last equality implies that

$$\lim_{k\longrightarrow\infty}\left\|F\left(x_{l(k)}\right)\right\| = \lim_{k\longrightarrow\infty}\left\|F\left(x_{l(k)-1}\right)\right\|. \tag{30}$$

Let $\widehat{l}(k) = l(k+N_0+2)$. Using induction, for all $j \ge 1$, we can show that

$$\lim_{k\longrightarrow\infty}\left\|d_{\widehat{l}(k)-j}\right\| = 0. \tag{31}$$

For $j = 1$, we can have from (31) that (29) is true. Assuming that (29) is true for given j, we show that (29) holds for given $j+1$. Let k be large enough such that $\widehat{l}(k)-(j+1) > 0$. Substituting k with $\widehat{l}(k)-j-1$ and using (24), we obtain

$$F_{l(\widehat{l}(k)-j-1)}^2 - \left\|F\left(x_{\widehat{l}(k)-j}\right)\right\|^2 \ge p_0 \text{Pred}_{\widehat{l}(k)-j-1}. \tag{32}$$

Similarly, we can deduce that

$$\lim_{k\longrightarrow\infty}\left\|d_{\widehat{l}(k)-j-1}\right\| = 0. \tag{33}$$

Therefore (31) holds. Along with Assumption 1 we imply that

$$\lim_{k\longrightarrow\infty}\left\|F\left(x_{\widehat{l}(k)-j-1}\right)\right\| = \lim_{k\longrightarrow\infty}\left\|F\left(x_{\widehat{l}(k)-j}\right)\right\|. \tag{34}$$

Similarly, for any given $j \ge 1$, we have $\lim_{k\to\infty}\|F(x_{\widehat{l}(k)-j})\| = \lim_{k\to\infty}\|F(x_{l(k)})\|$. On the one hand, for any k, we have

$$x_{k+1} = x_{\widehat{l}(k)} - \sum_{j=1}^{\widehat{l}(k)-k-1} d_{\widehat{l}(k)-j}, \tag{35}$$

Using (31) and the fact that $\widehat{l}(k)-j-1 \le N_0+1$, we have

$$\lim_{k\longrightarrow\infty}\left\|x_{k+1} - x_{\widehat{l}(k)}\right\| = 0, \tag{36}$$

With Assumption 1, we conclude

$$\lim_{k\longrightarrow\infty}\left\|F\left(x_{l(k)}\right)\right\| = \lim_{k\longrightarrow\infty}\left\|F\left(x_{\widehat{l}(k)}\right)\right\| = \lim_{k\longrightarrow\infty}\left\|F\left(x_k\right)\right\|. \tag{37}$$

And then (22) is proved. By using (22) and (25), we can deduce that

$$\lim_{k\longrightarrow\infty}\|d_k\| = 0. \tag{38}$$

Then, it follows from (3) in Algorithm 1, (17), and (22) that

$$\mu_k \longrightarrow \infty, \text{ as } k \longrightarrow \infty. \tag{39}$$

On the other hand, by (6), (8), (17), (22), (38), and Lemma 1, we can deduce that

$$\begin{aligned} r_k &= \frac{\text{Ared}_k}{\text{Pred}_k} \\ &= \frac{\|F_k\|^2 - \|F_{k+1}\|^2}{\text{Pred}_k} \\ &= 1 + \frac{\|F_k + J_k d_k\| O\left(\|d_k\|^2\right) + O\left(\|d_k\|^4\right)}{\text{Pred}_k} \\ &\le 1 + \frac{\|F_k + J_k d_k\| O\left(\|d_k\|^2\right) + O\left(\|d_k\|^4\right)}{\|J_k^T F_k\| \min\left(\|d_k\|, \left(\|J_k^T F_k\| / \|J_k^T J_k\|\right)\right)} \\ &\le 1 + \frac{O\left(\|d_k\|^2\right)}{\|d_k\|} \longrightarrow 1. \end{aligned} \tag{40}$$

And then combined with (8) and (10), we can see that

$$\begin{aligned} \overline{r}_k &= \frac{\overline{\text{Ared}}_k}{\text{Pred}_k} = \frac{F_{l(k)}^2 - \|F(x_k + d_k)\|^2}{\text{Pred}_k} \ge \frac{F_k^2 - \|F_{k+1}\|^2}{\text{Pred}_k} \\ &= r_k \longrightarrow 1. \end{aligned} \tag{41}$$

Considering Algorithm 1, obviously, for all large k, there exists a positive constant $\overline{\mu} > m$ such that $\mu_k < \overline{\mu}$, which conflicts with (22) and so Theorem 1 is true. □

3. Local Convergence

Definition 1. Let N be a subset of $\mathbf{R}^n$ that makes $N \cap X^* \neq \varnothing$. We say $\|F(x)\|$ provides a local error bounded on N for (1), where $\|x_k - \overline{x}_k\| = \text{dist}(x_k, X^*)$, if there exists a positive constant C_1 such that

$$\|F(x)\| \ge C_1 \text{dist}(x, X^*), \quad \forall x \in N. \tag{42}$$

Assumption 2

(a) $F(x)$ is continuously differentiable, and $\|F(x)\|$ provides a local error bound on subset $N(x^*, b)$ for problem (1), where

$$N(x^*, b) = \left\{x \in \mathbf{R}^n \mid \|x - x^*\| \le b\right\}, \quad 0 < b < 1. \tag{43}$$

(b) *$F(x)$ and $J(x)$ are both Lipschitz continuous on $N(x^*, b)$; that is, there exist two positive constants L_1, L_2 that make*

$$\|J(y) - J(x)\| \le L_1 \|y - x\|, \quad \forall x, y \in N(x^*, b), \tag{44}$$

$$\|F(y) - F(x)\| \le L_2 \|y - x\|, \quad \forall x, y \in N(x^*, b), \tag{45}$$

which implies

$$\begin{aligned} &\|F(y) - F(x) - J(x)(y - x)\| \le L_1 \|y - x\|^2, \\ &\forall x, y \in N(x^*, b). \end{aligned} \tag{46}$$

Lemma 4. *Suppose that Assumption 2 is true. Then, for all sufficiently large k, we have the following.*

(1) There exists a positive constant $M > m$ that makes

$$\mu_k \le M. \tag{47}$$

(2) $\lambda_k \ge C_3 \|x_k - \overline{x}_k\|^\delta$, where $C_3 = (C_1^\delta/2) m\theta + (C_2^\delta/2) m(1 - \theta)$.

Proof. The proof process of (1) is the same as that of Lemma 3.2 in [6], so we are not going to prove it here and only give the proof of (2).

$$\|F_k\|^2 = F_k^T F_k = F_k^T \left(F(\overline{x}_k) + J_k(x_k - \overline{x}_k)\right) + F_k^T V_k, \tag{48}$$

where $V_k = F_k - \overline{F}_k - J_k(x_k - \overline{x}_k)$, obviously,

$$F_k^T J_k (x_k - \overline{x}_k) = \|F_k\|^2 - F_k^T V_k. \tag{49}$$

We can obtain from Definition 1 and (45) and (46) that

$$\begin{aligned} \|J_k^T F_k\| \|x_k - \overline{x}_k\| &\ge C_1^2 \|x_k - \overline{x}_k\|^2 - L_1 L_2 \|x_k - \overline{x}_k\|^3, \\ \|J_k^T F_k\| &\ge C_1^2 \|x_k - \overline{x}_k\| - L_1 L_2 \|x_k - \overline{x}_k\|^2 \\ &\ge C_2 \|x_k - \overline{x}_k\|. \end{aligned} \tag{50}$$

Thus we obtain

$$\|J_k^T F_k\|^\delta \ge C_2^\delta \|x_k - \overline{x}_k\|^\delta. \tag{51}$$

Then we show the following inequalities:

$$\begin{aligned} \frac{\|F_k\|^\delta}{1 + \|F_k\|^\delta} &\ge \frac{C_1^\delta}{2} \|x_k - \overline{x}_k\|^\delta, \\ \frac{\|J_k^T F_k\|^\delta}{1 + \|J_k^T F_k\|^\delta} &\ge \frac{C_2^\delta}{2} \|x_k - \overline{x}_k\|^\delta. \end{aligned} \tag{52}$$

If $\|F_k\| \le 1$, then the following holds:

$$\frac{\|F_k\|^\delta}{1 + \|F_k\|^\delta} \ge \frac{C_1^\delta}{2} \|x_k - \overline{x}_k\|^\delta, \tag{53}$$

and if $\|F_k\| > 1$, then

$$\frac{\|F_k\|^\delta}{1+\|F_k\|^\delta} > \frac{1}{2}. \quad (54)$$

Through the above two inequalities, we have

$$\frac{\|F_k\|^\delta}{1+\|F_k\|^\delta} \geq \max\left\{\frac{1}{2}, \frac{C_1^\delta}{2}\|x_k - \overline{x}_k\|^\delta\right\}. \quad (55)$$

Similarly, if $\|J_k^T F_k\| \leq 1$, then the following holds

$$\frac{\|J_k^T F_k\|^\delta}{1+\|J_k^T F_k\|^\delta} \geq \frac{\|J_k^T F_k\|^\delta}{2} \geq \frac{C_2^\delta}{2}\|x_k - \overline{x}_k\|^\delta, \quad (56)$$

and if $\|J_k^T F_k\| > 1$, we have

$$\frac{\|J_k^T F_k\|^\delta}{1+\|J_k^T F_k\|^\delta} > \frac{1}{2}. \quad (57)$$

Hence, we obtain

$$\frac{\|J_k^T F_k\|^\delta}{1+\|J_k^T F_k\|^\delta} \geq \max\left\{\frac{1}{2}, \frac{C_2^\delta}{2}\|x_k - \overline{x}_k\|^\delta\right\}. \quad (58)$$

From Algorithm 1, (52), and (58), we have

$$\begin{aligned}\lambda_k &= \mu_k\left(\theta\frac{\|F_k\|^\delta}{1+\|F_k\|^\delta} + (1-\theta)\frac{\|J_k^T F_k\|^\delta}{1+\|J_k^T F_k\|^\delta}\right)\\ &\geq \frac{C_1^\delta}{2} m\theta\|x_k - \overline{x}_k\|^\delta + \frac{C_2^\delta}{2} m(1-\theta)\|x_k - \overline{x}_k\|^\delta \geq C_3\|x_k - \overline{x}_k\|^\delta,\end{aligned} \quad (59)$$

where $C_3 = (C_1^\delta/2)m\theta + (C_2^\delta/2)m(1-\theta)$. So when $\|x_k - \overline{x}_k\|$ is sufficiently small,

$$\lambda_k \geq C_3\|x_k - \overline{x}_k\|^\delta. \quad (60)$$

□

Lemma 5. *Suppose that Assumption 2 is true, for all sufficiently large k. $\|d_k\|$ computed by (3) satisfies*

$$\|d_k\| \leq O(\|x_k - \overline{x}_k\|). \quad (61)$$

Proof. If we set $\varphi_k(d) = \|F_k + J_k d\|^2 + \lambda_k\|d\|^2$, then we have from (3) that d_k is a minimizer of $\varphi_k(d)$, so it follows from Algorithm 1, (12), (45), (46), and $\|F(\overline{x}_k)\| = 0$ that

$$\begin{aligned}\|d_k\|^2 &\leq \frac{1}{\lambda_k}\varphi_k(x_k - \overline{x}_k)\\ &= \frac{1}{\lambda_k}\left(\|F_k + J_k(x_k - \overline{x}_k)\|^2 + \lambda_k\|x_k - \overline{x}_k\|^2\right)\\ &\leq \frac{1}{C_3\|x_k - \overline{x}_k\|^\delta}\left(\|F_k + J_k(x_k - \overline{x}_k)\|^2 + \lambda_k\|x_k - \overline{x}_k\|^2\right)\\ &\leq \frac{L_1^2\|x_k - \overline{x}_k\|^4}{C_3\|x_k - \overline{x}_k\|^\delta} + \|x_k - \overline{x}_k\|^2\\ &= O\left(\|x_k - \overline{x}_k\|^2\right).\end{aligned} \quad (62)$$

So there is

$$\|d_k\| \leq O(\|x_k - \overline{x}_k\|). \quad (63)$$

Lemma 6. *Suppose that Assumption 2 is true, for all sufficiently large k. So, we have*

$$\lambda_k \leq C_4\|x_k - \overline{x}_k\|^\delta. \quad (64)$$

Proof. Firstly, we deal with these two equations:

$$\begin{aligned}&\frac{\|F_k\|^\delta}{1+\|F_k\|^\delta} \leq \|F_k\|^\delta \leq L_2^\delta\|x_k - \overline{x}_k\|^\delta,\\ &\|J_k^T F_k\| \leq \|J_k^T\|\|F_k\| \leq L_2\|F(x_k) - F(\overline{x}_k)\| \leq L_2^2\|x_k - \overline{x}_k\|.\end{aligned} \quad (65)$$

On the other hand,

$$\frac{\|J_k^T F_k\|^\delta}{1+\|J_k^T F_k\|^\delta} \leq \|J_k^T F_k\|^\delta \leq L_2^{2\delta}\|x_k - \overline{x}_k\|^\delta. \quad (66)$$

We conclude

$$\begin{aligned}\lambda_k &= \mu_k\left(\theta\frac{\|F_k\|^\delta}{1+\|F_k\|^\delta} + (1-\theta)\frac{\|J_k^T F_k\|^\delta}{1+\|J_k^T F_k\|^\delta}\right)\\ &\leq M\left(\theta L_2^\delta\|x_k - \overline{x}_k\|^\delta + (1-\theta)L_2^{2\delta}\|x_k - \overline{x}_k\|^\delta\right)\\ &\leq C_4\|x_k - \overline{x}_k\|^\delta,\end{aligned} \quad (67)$$

where $C_4 = M\theta L_2^\delta + M(1-\theta)L_1^\delta L_2^\delta$.

Without loss of generality, for all $\overline{x} \in N(x^*, b) \cap X^*$, suppose that $\operatorname{rank}(J(\overline{x})) = r$, and we prove the local convergence of Algorithm 1 by singular value decomposition (SVD) of $J(\overline{x})$.

$$J(\overline{x}) = [\,\overline{U}_1 \;\; \overline{U}_2\,] \begin{bmatrix} \overline{\Sigma}_1 & \\ & 0 \end{bmatrix} [\,\overline{V}_1^T,\; \overline{V}_2^T\,], \tag{68}$$

where $\overline{\Sigma}_1 = \operatorname{diag}(\overline{\sigma}_1, \ldots, \overline{\sigma}_r)$ with $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_r > 0$, $\operatorname{rank}(\overline{\Sigma}_1) = r$, $\overline{U} = [\overline{U}_1, \overline{U}_2]$, and $\overline{V} = [\overline{V}_1, \overline{V}_2]$. And assume the SVD of $J(x)$ are as follows:

$$\begin{aligned} J(x) = U\Sigma V^T &= [U_1, U_2, U_3] \begin{bmatrix} \Sigma_1 & & \\ & \Sigma_2 & \\ & & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \\ V_3^T \end{bmatrix} \\ &= U_1 \Sigma_1 V_1^T + U_2 \Sigma_2 V_2^T, \end{aligned} \tag{69}$$

where $U = [U_1, U_2, U_3]$ and $V = [V_1, V_2, V_3]$ are two orthogonal matrixes. $\Sigma_1 = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and $\Sigma_2 = \operatorname{diag}(\sigma_{r+1}, \sigma_{r+2}, \ldots, \sigma_{r+q})$, $\sigma_{r+1} \geq \sigma_{r+2} \geq \cdots \geq \sigma_{r+q} > 0$.

Since $J(x)$ is Lipschitz continuous, by the theory of matrix perturbation [18], we have

$$\left\| \operatorname{diag}(\Sigma_1 - \overline{\Sigma}_1, \Sigma_2, 0) \right\| \leq \left\| J_k - \overline{J}_k \right\| \leq L_1 \left\| x_k - \overline{x}_k \right\|. \tag{70}$$

So there is

$$\begin{aligned} \left\| \Sigma_1 - \overline{\Sigma}_1 \right\| &\leq L_1 \left\| (x_k - \overline{x}_k) \right\|, \\ \left\| \Sigma_2 \right\| &\leq \left\| J_k - \overline{J}_k \right\| \leq L_1 \left\| (x_k - \overline{x}_k) \right\|. \end{aligned} \tag{71}$$

Since $\{x_k\}$ converges to the set X^*, then we have $L_1 \|(x_k - \overline{x}_k)\| \leq (\overline{\sigma}_r/2)$ hold for all sufficiently large k. So combined with (71), there is

$$\left\| \Sigma_1^{-1} \right\| \leq \frac{1}{\overline{\sigma}_r - L_1 \left\| (x_k - \overline{x}_k) \right\|} \leq \frac{2}{\overline{\sigma}_r}. \tag{72}$$

□

Lemma 7. *Suppose Assumption 2 holds; for all sufficiently large k, we have*

(1) $\|U_1 U_1^T F_k\| \leq O(\|x_k - \overline{x}_k\|)$

(2) $\|U_2 U_2^T F_k\| \leq O(\|x_k - \overline{x}_k\|^2)$

Proof. The proof process is similar to Lemma 7 in [9], so we omit it here. □

Theorem 2. *Under Assumption 2, let $\{x_k\}$ be a sequence generated by Algorithm 1 with trust region technique. If $\delta \in (0,1)$, then $\{x_k\}$ converges superlinearly to the solution. If $\delta \in [1,2]$, then sequence $\{x_k\}$ converges quadratically to the solution.*

Proof. Using (3) and (69), we obtain

$$\begin{aligned} d_k &= -V_1 \left(\Sigma_1^2 + \lambda_k I \right)^{-1} \Sigma_1 U_1^T F_k - V_2 \left(\Sigma_2^2 + \lambda_k I \right)^{-1} \Sigma_2 U_2^T F_k, \\ F_k + J_k d_k &= F_k - U_1 \Sigma_1 \left(\Sigma_1^2 + \lambda_k I \right)^{-1} \Sigma_1 U_1^T F_k - U_2 \Sigma_2 \left(\Sigma_2^2 + \lambda_k I \right)^{-1} \Sigma_2 U_2^T F_k \\ &= \lambda_k U_1 \left(\Sigma_1^2 + \lambda_k I \right)^{-1} U_1^T F_k + \lambda_k U_2 \left(\Sigma_2^2 + \lambda_k I \right)^{-1} U_2^T F_k. \end{aligned} \tag{73}$$

From (47), (71), Definition 1, and Lemma 5, we have

$$\begin{aligned} \left\| F_k + J_k d_k \right\| &\leq \lambda_k \left\| U_1 U_1^T F_k \right\| + \left\| U_2 U_2^T F_k \right\| \\ &\leq C_4 \left\| x_k - \overline{x}_k \right\|^{\delta} O\left(\left\| x_k - \overline{x}_k \right\| \right) + O\left(\left\| x_k - \overline{x}_k \right\|^2 \right) \\ &\leq C_4 \left\| x_k - \overline{x}_k \right\|^{1+\delta} + O\left(\left\| x_k - \overline{x}_k \right\|^2 \right) \\ &\leq O\left(\left\| x_k - \overline{x}_k \right\|^{1+\delta} \right) + O\left(\left\| x_k - \overline{x}_k \right\|^2 \right). \end{aligned} \tag{74}$$

From (44), (46), (74), Definition 1, and Lemma 5, we conclude that

$$\begin{aligned} C_1 \left\| (x_{k+1} - \overline{x}_{k+1}) \right\| &\leq \left\| F(x_{k+1}) \right\| = \left\| F(x_k + d_k) \right\| \\ &\leq O\left(\left\| (x_k - \overline{x}_k) \right\|^{1+\delta} \right) + O\left(\left\| (x_k - \overline{x}_k) \right\|^2 \right). \end{aligned} \tag{75}$$

On the other hand, it is obvious that

$$\left\| x_k - \overline{x}_k \right\| = \operatorname{dist}(x_k, X^*) \leq \left\| x_{k+1} - \overline{x}_k \right\| \leq \left\| x_{k+1} - \overline{x}_{k+1} \right\| + d_k. \tag{76}$$

It follows from (75) and Lemma 5 that $\|x_k - \overline{x}_k\| \leq 2\|d_k\| \leq O(\|x_k - \overline{x}_k\|)$ holds for all sufficiently large k. So, $\|d_k\| = O(\|x_k - \overline{x}_k\|)$, and this is related to (75). We deduce that if $\delta \in (0,1)$, $\|d_{k+1}\| = O(\|d_k\|^{1+\delta})$. If $\delta \in [1,2]$, $\|d_{k+1}\| = O(\|d_k\|^2)$. So if $\delta \in (0,1)$, $\{x_k\}$ converges superlinearly to the solution; otherwise, if $\delta \in [1,2]$, $\{x_k\}$ converges quadratically to the solution. □

4. Numerical Experiments

In Section 5, we compare the performance of Algorithm 1 with Algorithm 2.1 (writing Algorithm 2.1 as AELM) in [6] through some numerical experiments. The test function $F(\overline{x})$ is improved by the method in [19]. The form is as follows:

$$F(\overline{x}) = F(x) - J(x^*) A \left(A^T A \right)^{-1} A^T (x - x), \tag{77}$$

Input: given $x_0 \in \mathbf{R}^n, N_0 > 0, \mu_0 > m > 0, \varepsilon > 0, 0 < p_0 < p_1 < p_2 < 1, k := 0$
Output:
Step 1. If $J_k^T F_k \leq \varepsilon$, stop. Otherwise, set
$\lambda_k = \mu_k (\theta \|F_k\|^\delta / (1 + \|F_k\|^\delta) + (1-\theta)\|J_k^T F_k\|^\delta / (1 + \|J_k^T F_k\|^\delta))$,
where $\theta \in [0,1], \delta \in (0,2]$.
Step 2. Compute the search direction
$d_k = (J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k$.
Step 3. By (7), (8), and (10), compute $\mathrm{Pred}_k, \overline{\mathrm{A}}\mathrm{red}_k, \overline{r}_k$.
Step 4. Set
$$x_{k+1} = \begin{cases} x_k + d_{k,} & \overline{r}_k \geq p_0 \\ x_{k,} & \overline{r}_k < p_0 \end{cases}$$
Step 5. Choose μ_{k+1} as
$$\mu_{k+1} = \begin{cases} 4\mu_{k,} & \overline{r}_k < p_1 \\ \mu_{k,} & \overline{r}_k \in [p_1, p_2] \\ \max\{(\mu_k/4), m\}, & \overline{r}_k > p_2 \end{cases}$$
Step 6. Set $k := k+1$, and go to Step 1.

Algorithm 1: A modified efficient Levenberg–Marquardt algorithm.

Table 1: Results on singular nonlinear equations with rank$(F'(x^*)) = n - 1$.

			NF/NJ					
Problem	n	x_0	AELM	Algorithm with $\delta = 1$		Algorithm with $\delta = 2$		
				$\theta = 0$	$\theta = 0.5$	$\theta = 0$	$\theta = 0.5$	$\theta = 1$
Helical valley function	3	−100	7/7	7/7	8/8	9/9	7/7	7/7
		−10	6/6	6/6	7/7	7/7	6/6	6/6
		−1	1/1	1/1	1/1	1/1	1/1	1/1
		1	9/9	9/9	9/9	9/9	9/9	9/9
		10	8/8	8/8	8/8	9/9	8/8	8/8
		100	8/8	8/8	8/8	8/8	8/8	8/8
Discrete boundary value	10	−100	14/14	14/14	14/14	14/14	14/14	14/14
		−10	11/11	10/10	11/11	10/10	10/10	10/10
		−1	8/8	8/8	8/8	7/7	7/7	7/7
		1	3/3	3/3	3/3	4/4	4/4	4/4
		10	11/11	10/10	10/10	9/9	9/9	9/9
		100	9/9	9/9	9/9	9/9	9/9	9/9
	100	−100	14/14	13/13	14/14	13/13	14/14	14/14
		−10	11/11	9/9	11/11	8/8	9/9	9/9
		−1	12/12	5/5	5/5	4/4	4/4	4/4
		1	4/4	4/4	4/4	3/3	3/3	3/3
		10	7/7	7/7	7/7	7/7	5/5	8/8
		100	12/12	11/11	12/12	11/11	12/12	12/12
	500	−100	15/15	13/13	15/15	12/12	14/14	15/15
		−10	8/8	6/6	7/7	5/5	6/6	6/6
		−1	6/6	5/5	5/5	4/4	4/4	4/4
		1	1/1	1/1	1/1	1/1	1/1	1/1
		10	8/8	6/6	7/7	5/5	6/6	6/6
		100	15/15	12/12	14/14	11/11	13/13	14/14

Table 1: Continued.

Problem	n	x_0	NF/NJ					
			AELM	Algorithm with $\delta = 1$		Algorithm with $\delta = 2$		
				$\theta = 0$	$\theta = 0.5$	$\theta = 0$	$\theta = 0.5$	$\theta = 1$
Discrete integral equation	30	−100	16/16	16/16	16/16	16/16	16/16	16/16
		−10	11/11	11/11	11/11	11/11	11/11	11/11
		−1	9/9	9/9	9/9	9/9	9/9	9/9
		1	7/7	7/7	7/7	6/6	6/6	6/6
		10	11/11	11/11	11/11	11/11	11/11	11/11
		100	10/10	10/10	10/10	10/10	10/10	10/10
	100	−100	16/16	16/16	16/16	16/16	16/16	16/16
		−10	12/12	12/12	12/12	12/12	12/12	12/12
		−1	9/9	10/10	10/10	9/9	9/9	9/9
		1	8/8	7/7	8/8	7/7	7/7	7/7
		10	12/12	11/11	11/11	11/11	11/11	12/12
		100	10/10	10/10	10/10	10/10	10/10	10/10
	500	−100	17/17	17/17	17/17	17/17	17/17	17/17
		−10	13/13	13/13	13/13	13/13	13/13	13/13
		−1	11/11	11/11	11/11	11/11	10/10	10/10
		1	9/9	8/8	9/9	8/8	8/8	8/8
		10	12/12	12/12	12/12	12/12	12/12	12/12
		100	10/10	10/10	10/10	10/10	10/10	10/10

Table 2: Numerical results for singular nonlinear equations with $\operatorname{rank}(F'(x^*) = n - 1)$.

Problem	n	x_0	Iters/fnorm/trust/EI		
			AELM (times: 3.63)	$\delta = 1$ (times: 7.47)	$\delta = 2$ (times: 3.76)
Rosenbrock	2	1	16/1.5876e − 07/1/1.044	16/2.3984e − 07/1/1.044	16/2.3152e − 07/1/1.071
		10	18/1.199e − 07/1/1.038	18/1.211e − 07/1/1.038	18/1.2102e − 07/1/1.0628
		100	21/1.3601e − 07/1/1.034	21/1.3604e − 07/1/1.034	21/1.3603e − 07/1/1.053
Powell badly	2	1	40/2.0905e − 07/1/1.017	40/6.4701e − 07/1/1.017	40/4.8829e − 07/1/1.027
		10	25/3.4159e − 07/1/1.028	25/3.6737e − 07/1/1.028	25/3.0308e − 07/1/1.045
		100	30/3.4445e − 07/1/1.023	30/3.1265e − 07/1/1.023	30/3.0929e − 07/1/1.037
Wood	4	1	16/1.5876e − 07/1/1.044	16/2.0905e − 07/1/1.044	16/2.0897e − 07/1/1.071
		10	19/1.926e − 07/1/1.037	19/1.9262e − 07/1/1.037	19/1.9262e − 07/1/1.059
		100	22/2.8391e − 07/1/1.031	22/2.8391ee − 07/1/1.032	22/2.8391e − 07/1/1.050
Helical valley	3	1	9/9.8336e − 07/0/1.007	9/1.5533e − 09/0/1.007	96/1.4071e − 09/0/1.128
		10	8/2.7959e − 07/0/1.090	8/7.0428e − 07/0/1.090	9/8.3791e − 14/0/1.128
		100	8/2.0685e − 07/0/1.090	8/1.2902e − 08/0/1.090	8/1.1432e − 08/0/1.013
Brown almost-linear	10	1	8/7.2842e − 07/1/1.090	8/7.3653e − 05/1/1.090	8/7.4202e − 05/1/1.013
		10	23/7.9915e − 07/1/1.030	23/7.9915e − 05/1/1.030	23/7.9915e − 05/1/1.048
		100	45/6.5326e − 07/1/1.015	45/6.5326e − 05/1/1.015	45/6.5326e − 05/11.024
Discrete boundary value	10	1	3/0.00016727/0/1.259	3/0.00016758/0/1.259	4/1.1606e − 05/0/1.316
		10	11/5.6012e − 07/0/1.064	10/1.3863e − 05/0/1.071	9/1.1763e − 05/0/1.129
		100	9/9.2906e − 07/0/1.079	9/8.9129e − 06/0/1.079	9/8.7568e − 06/0//1.129
Discrete integral equation	30	1	7/9.5122e − 07/1/1.103	7/8.4521e − 06/1/1.103	6/1.1785e − 05/1/1.001
		10	11/9.8059e − 07/1/1.064	11/8.1669e − 06/1/1.064	11/7.3177e − 06/1/1.002
		100	10/1.0601e − 07/0/1.071	10/1.0609e − 09/0/1.071	10/1.0606e − 09/0/1.116
Trigonometric	10	1	8/5.2977e − 07/0/1.008	8/2.4187e − 07/0/1.008	804/6.2122e − 09/1/1.001
		10	19/2.3813e − 07/0/1.037	–/19.064/0/1	540/2.7155e − 09/1/1.002
		100	–/46.227e − 07/0/1	–/29.583/0/1	19/62.146/1/1
Variably dimensioned function	10	1	14/2.2087e − 07/0/1.050	14/2.2087e − 05/1/1.050	14/2.2087e − 05/1/1.081
		10	16/1.1269e − 07/1/1.044	16/1.1269e − 05/1/1.044	16/1.1269e − 05/1/1.071
		100	19/3.702e − 07/1/1.037	19/3.702e − 05/1/1.037	19/3.702e − 05/1/1.059
Broyden tridiagonal	30	1	9/2.2916e − 07/1/1.007	9/2.4133e − 05/1/1.007	9/2.3888e − 05/1/1.129
		10	14/1.419e − 05/1/1.050	14/1.419e − 05/1/1.050	14/1.419e − 05/1/1.081
		100	17/2.5157e − 05/1/1.041	17/2.5157e − 07/1/1.041	17/2.5157e − 05/1/1.066
Broyden banded	30	1	12/4.2271e − 07/1/1.059	12/4.2313e − 06/1/1.059	12/4.2305e − 06/1/1.095
		10	18/4.5992e − 07/1/1.039	18/4.5992e − 06/1/1.039	18/4.5992e − 06/1/1.062
		100	24/2.7182e − 07/1/1.029	24/2.7182e − 06/1/1.029	24/2.7182e − 06/1/1.046

Table 3: Results on singular nonlinear equations with rank $(F'(x^*) = n - 2$.

Problem	n	x_0	NF/NJ					
			AELM	Algorithm with $\delta = 1$		Algorithm with $\delta = 2$		
				$\theta = 0$	$\theta = 0.5$	$\theta = 0$	$\theta = 0.5$	$\theta = 1$
Powell badly scaled	2	−100	852/717	275/250	169/158	1802/1521	2585/2168	366/318
		−10	76/64	163/144	69/61	1595/1334	3865/3226	585/484
		−1	183/167	97/79	118/103	2754/2296	2585/2168	341/279
		1	1288/1076	128/111	127/105	–	996/833	711/591
		10	3/3	92/76	3/3	7/4	7/4	1043/873
		100	3/3	3/3	3/3	3/3	3/3	1954/1622
Discrete boundary value	10	−100	14/14	14/14	14/14	14/14	14/14	14/14
		−10	11/11	10/10	11/11	10/10	10/10	10/10
		−1	8/8	8/8	8/8	7/7	7/7	7/7
		1	3/3	3/3	3/3	4/4	4/4	4/4
		10	11/11	10/10	10/10	9/9	9/9	9/9
		100	10/10	10/10	10/10	10/10	10/10	10/10
	100	−100	14/14	13/13	14/14	13/13	14/14	14/14
		−10	11/11	9/9	11/11	8/8	9/9	9/9
		−1	6/6	5/5	5/5	4/4	4/4	4/4
		1	4/4	4/4	4/4	3/3	3/3	3/3
		10	7/7	9/9	7/7	7/7	8/8	8/8
		100	13/13	12/12	13/13	11/11	12/12	13/13
Discrete integral equation	30	−100	16/16	16/16	16/16	16/16	16/16	16/16
		−10	11/11	11/11	11/11	11/11	11/11	11/11
		−1	9/9	9/9	9/9	9/9	9/9	9/9
		1	7/7	7/7	7/7	6/6	6/6	6/6
		10	11/11	11/11	11/11	11/11	11/11	11/11
		100	12/12	12/12	12/12	12/12	12/12	12/12
	100	−100	16/16	16/16	16/16	16/16	16/16	16/16
		−10	12/12	12/12	12/12	12/12	12/12	12/12
		−1	10/10	10/10	10/10	9/9	9/9	9/9
		1	8/8	7/7	8/8	7/7	7/7	7/7
		10	12/12	11/11	11/11	11/11	11/11	12/12
		100	14/14	14/14	14/14	14/14	14/14	14/14

Table 4: Numerical results for singular nonlinear equations with rank $(F'(x^*) = n - 2)$.

Problem	n	x_0	Iters/fnorm/trust/EI		
			AELM (times: 4.655)	$\delta = 1$ (times: 3.374)	$\delta = 2$ (times: 3.614)
Rosenbrock	2	1	11/4.6197e − 05/0/1.065	11/4.6202e − 05/0/1.065	11/4.6202e − 05/0/1.105
		10	13/0.00010074/0/1.054	13/0.00010074/0/1.054	13/0.00010074/0/1.088
		100	17/3.4089e − 05/0/1.041	17/3.4089e − 05/0/1.041	21/63/3.4089e − 05/0/1.066
Powell badly	2	1	1288/1.0433e − 07/0/1.000	127/4.9595e − 07/1/1.008	996/41.8112e − 07/0/1.001
		10	3/3.3583e − 05/0/1.259	3/3.3562e − 05/0/1.028	7/1.6053e − 05/0/1.169
		100	3/0.0099779/0/1.259	3/0.009978/0/1.023	3/0.009978/0/1.442
Wood	4	1	14/3.2876e − 06/0/1.050	14/3.2877e − 06/0/1.081	14/3.2877e − 06/0/1.081
		10	17/3.0842e − 06/0/1.041	17/3.0842e − 06/0/1.041	17/3.0842e − 06/1/1.066
		100	20/4.5433e − 06/0/1.035	20/4.5433e − 06/0/1.035	20/4.5433e − 06/0/1.056
Helical valley	3	1	13/5.9421e − 06/1/1.054	13/5.9553e − 06/0/1.054	13/5.9694e − 06/0/1.088
		10	14/2.087e − 06/1/1.050	14/2.0897e − 06/0/1.050	14/2.0934e − 06/1/1.081
		100	18/3.0231e − 06/1/1.039	18/3.1777e − 06/0/1.039	18/3.3522e − 06/1/1.062
Brown almost-linear	10	1	8/7.2842e − 05/1/1.147	8/7.3653e − 05/1/1.090	8/7.4202e − 05/0/1.147
		10	23/7.9915e − 05/1/1.030	23/7.9915e − 05/1/1.030	23/7.9915e − 05/1/1.048
		100	45/6.5326e − 05/1/1.015	45/6.5326e − 05/1/1.015	45/6.5325e − 05/1/1.024
Discrete boundary value	10	1	3/0.0001676/0/1.259	3/0.00016795/0/1.259	4/1.1572e − 05/0/1.316
		10	11/5.5963e − 06/0/1.065	10/48/1.383e − 05/0/1.065	9/1.2359e − 05/0/1.129
		100	10/6.7629e − 08/0/1.071	10/6.3555e − 08/0/1.071	10/6.1475e − 08/0//1.116
Discrete integral equation	30	1	7/9.5122e − 06/1/1.104	7/8.4521e − 06/1/1.104	6/1.1785e − 05/1/1.200
		10	11/9.806e − 06/1/1.065	11/8.167e − 06/1/1.065	11/7.3178e − 06/1/1.105
		100	12/1.3149e − 07/0/1.059	12/11.8382e − 07/0/1.059	12/1.137e − 09/0/1.095

Table 4: Continued.

Problem	n	x_0	Iters/fnorm/trust/EI AELM (times: 4.655)	$\delta = 1$ (times: 3.374)	$\delta = 2$ (times: 3.614)
Trigonometric	10	1	13/2.2087e – 05/1/1.054	13/9.9955e – 06/0/1.054	14/7.3496e – 06/1/1.001
		10	31/1.113e – 09/0/1.022	22/2.4389e – 08/0/1.032	16/1.8831e – 07/0/1.002
		100	–/15.261/0/1	1/120.79/0/1	19/48/170.57/0/1
Variably dimensioned function	10	1	14/2.2087e – 05/1/1.050	14/2.2087e – 05/1/1.050	14/2.2087e – 05/1/1.081
		10	16/1.1269e – 05/0/1.044	16/1.1269e – 05/0/1.044	16/1.1269e – 05/0/1.071
		100	19/3.702e – 05/0/1.037	19/3.702e – 05/0/1.037	19/3.702e – 05/0/1.059
Broyden tridiagonal	30	1	9/2.294e – 05/1/1.080	9/2.4133e – 05/1/1.007	9/2.3903e – 05/1/1.129
		10	14/1.4189e – 05/1/1.050	14/1.419e – 05/1/1.050	14/1.419e – 05/1/1.081
		100	17/2.5157e – 05/1/1.041	17/2.5157e – 07/1/1.041	17/2.5157e – 05/1/1.066
Broyden banded	30	1	12/4.2271e – 07/1/1.059	12/4.2313e – 06/1/1.059	12/4.2305e – 06/1/1.095
		10	18/4.5992e – 07/1/1.039	18/4.2297e – 06/11.039	18/4.5984e – 06/1/1.062
		100	24/2.7182e – 07/1/1.029	24/2.7182e – 06/1/1.029	24/2.7178e – 06/1/1.046

where $A \in \mathbf{R}^{n\times k}\,(1 \le k \le n)$ has full column rank, and $F(x^*) = 0$. It is certain that

$$\overline{J}(x^*) = J(x^*)\left(I - A\left(A^T A\right)^{-1} A^T\right), \tag{78}$$

where test problems $F(x)$ are nonsingular test functions from [20]. We take $A = [1, 1, \ldots, 1]^T \in \mathbf{R}^{n\times 1}$, rank of $\overline{J}(x^*)$ as $n-1$, and choose

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & -1 & \cdots & \pm 1 \end{bmatrix}, \tag{79}$$

rank of $\overline{J}(x^*)$ as $n-2$.

We set the following parameters in the algorithms: $p_0 = 10^{-4}, p_1 = 0.25, p_1 = 0.75, N_0 = 5, \mu_1 = 1, m = 10^{-8}$, $\theta = 0, 0.5$ or 1, and $\delta = 1$ or 2. The algorithms are terminated when $\|J_k^T F_k\| \le 10^{-5}$ or the number of iterates exceeds 10^4. When $\theta = 1$ and $\delta = 1$, Algorithm 1 is reduced to the AELM.

By numerical experiments, we find the numerical results of Algorithm 1 are the same as the numerical results of AELM in some functions. So we only list the results of other experiments in the following tables. Further, we adopt the efficiency index defined as EI in [21] to compare the performance of algorithm AELM and Algorithm 1. The results of the four experiments with rank$J(x^*) = n-1$ are shown in Tables 1 and 2, and the results of the four experiments with rank$J(x^*) = n-2$ are shown in Tables 3 and 4, respectively. We use six starting points $\pm 100x_0$, $\pm 10x_0$, and $\pm x_0$ for each test problem, where x_0 is suggested in [20].

(i) NF stands for the quantity of function calculations

(ii) NJ stands for the quantity of Jacobian calculations

(iii) '–' indicates that the iteration number is more than 10^4

(iv) $E.I. = \rho^{1/NF}$, where ρ is the convergence order of algorithm.

It is shown in Table 1 that when $\theta = 0.5$ and $\delta = 2$, the effect of Algorithm 1 is obviously better than that of AELM. Algorithm 1 wins 40.5% of the numerical results while AELM wins 2.38%, and 57.1% of the two algorithms have the same results. The advantage of Algorithm 1 is not obvious when $\theta = 0.5$ and $\delta = 1$. Algorithm 1 can win 19% of the numerical results while AELM win 7.14%, and 73.8% of two algorithms have the same result.

Table 3 shows that when $\theta = 0.5$ and $\delta = 1$, Algorithm 1 and AELM have the best experimental results. Algorithm 1 win 23.3% of the numerical results, and 76.6% of the two algorithms has the same result. The advantage of Algorithm 1 is not obvious when $\theta = 0.5$ and $\delta = 2$. Algorithm 1 wins 40% of the numerical results while AELM wins 20%, and 40% of the two algorithms has the same results.

Further, we adopt the EI and let $\theta = 0.5$ in the experiment. Tables 2 and 4 show that when $\delta = 1$, the experimental data EI of AELM and Algorithm 1 are similar, but when $\delta = 1$, the EI of Algorithm 1 is obviously larger than that of AELM. In addition, in terms of the experimental time, except when ranking $(F'(x^*) = n-2)$ and $\delta = 1$, the execution time of Algorithm 1 is longer than that of AELM. In other cases, the execution time of Algorithm 1 is close to or less than that of AELM.

In general, it is shown that for most test problems, Algorithm 1 performs better than AELM. So it can indicate that Algorithm 1 is more efficient than AELM to solve systems of nonlinear equations.

5. Conclusion

In this paper, we propose a new LM algorithm by modifying the LM parameter for systems of nonlinear equations. Through numerical experiments, we find the calculation amounts of Algorithm 1 smaller than AELM in the case where θ and δ take some suitable value, which shows the effectiveness of the new Algorithm 1. Under some conditions, the global convergence of the new LM method is proved, and the local convergence of the new LM method is shown to be at least superlinear. Numerical results show that the new algorithm is efficient.

Data Availability

No data were used to support the findings of this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Natural Science Foundation of the Anhui Higher Education Institutions under Grant nos. KJ2020ZD008 and KJ2019A0604, by the Natural Science Foundation of Anhui Province under no. 2108085MF204, and by the Abroad Visiting of Excellent Young Talents in Universities of Anhui Province under Grant no. GXGWFX2019022.

References

[1] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, NY, USA, 1999.

[2] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, no. 2, pp. 164–168, 1944.

[3] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[4] N. Yamashita and M. Fukushima, "On the rate of convergence of the levenberg-marquardt method," in *Topics in Numerical Analysis*, pp. 239–249, Springer, Vienna, AS, USA, 2001.

[5] J.-Y. Fan and Y.-X. Yuan, "On the quadratic convergence of the levenberg-marquardt method without nonsingularity assumption," *Computing*, vol. 74, no. 1, pp. 23–39, 2005.

[6] K. Amini, F. Rostami, and G. Caristi, "An efficient Levenberg-Marquardt method with a new LM parameter for systems of nonlinear equations," *Optimization*, vol. 67, no. 5, pp. 637–650, 2018.

[7] A. O. Umar, I. M. Sulaiman, M. Mamat, M. Y. Waziri, and N. Zamri, "On damping parameters of Levenberg-Marquardt algorithm for nonlinear least square problems," *Journal of Physics: Conference Series*, vol. 1734, Article ID 012018, 2021.

[8] H. Wang and J. Fan, "Convergence rate of the Levenberg-Marquardt method under Hölderian local error bound," *Optimization Methods and Software*, vol. 35, no. 4, pp. 767–786, 2020.

[9] C. Ma and L. Jiang, "Some research on Levenberg-Marquardt method for the nonlinear equations," *Applied Mathematics and Computation*, vol. 184, no. 2, pp. 1032–1040, 2007.

[10] J. Fan, J. Huang, and J. Pan, "An adaptive multi-step levenberg-marquardt method," *Journal of Scientific Computing*, vol. 78, no. 1, pp. 531–548, 2019.

[11] K. Amini and F. Rostami, "A modified two steps Levenberg-Marquardt method for nonlinear equations," *Journal of Computational and Applied Mathematics*, vol. 288, pp. 341–350, 2015.

[12] L. Chen and Y. Ma, "Shamanskii-like Levenberg-Marquardt method with a new line search for systems of nonlinear equations," *Journal of Systems Science and Complexity*, vol. 33, no. 5, pp. 1694–1707, 2020.

[13] L. Grippo, F. Lampariello, and S. Lucidi, "A nonmonotone line search technique for Newton's method," *SIAM Journal on Numerical Analysis*, vol. 23, no. 4, pp. 707–716, 1986.

[14] M. Ahookhosh and K. Amini, "An efficient nonmonotone trust-region method for unconstrained optimization," *Numerical Algorithms*, vol. 59, no. 4, pp. 523–540, 2012.

[15] M. Ahookhosh and K. Amini, "A Nonmonotone trust region method with adaptive radius for unconstrained optimization problems," *Computers and Mathematics with Applications*, vol. 60, no. 3, pp. 411–422, 2010.

[16] M. J. D. Powell, "An iterative method for finding stationary values of a function of several variables," *The Computer Journal*, vol. 5, no. 2, pp. 147–151, 1962.

[17] W. Zhou, "On the convergence of the modified Levenberg-Marquardt method with a nonmonotone second order Armijo type line search," *Journal of Computational and Applied Mathematics*, vol. 239, pp. 152–161, 2013.

[18] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, Cambridge, MA, USA, 1990.

[19] R. B. Schnabel and P. D. Frank, "Tensor methods for nonlinear equations," *SIAM Journal on Numerical Analysis*, vol. 21, no. 5, pp. 815–843, 1984.

[20] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, "Testing unconstrained optimization software," *ACM Transactions on Mathematical Software*, vol. 7, no. 1, pp. 17–41, 1981.

[21] S. Qasim, Z. Ali, F. Ahmad et al., "Solving systems of nonlinear equations when the nonlinearity is expensive," *Computers and Mathematics with Applications*, vol. 71, no. 7, pp. 1464–1478, 2016.