Journal of Applied Analysis and Computation
Volume 14, Number 4, August 2024, 1959–1976
Website:http://www.jaac-online.com
DOI:10.11948/20220441

LEVENBERG-MARQUARDT METHOD WITH A GENERAL LM PARAMETER AND A NONMONOTONE TRUST REGION TECHNIQUE*

Luyao Zhao[1] and Jingyong Tang[1,†]

Abstract We propose a new Levenberg-Marquardt (LM) method for solving the nonlinear equations. The new LM method takes a general LM parameter $\lambda_k = \mu_k[(1-\theta)\|F_k\|^\delta + \theta\|J_k^T F_k\|^\delta]$ where $\theta \in [0,1]$ and $\delta \in (0,3)$ and adopts a nonmonotone trust region technique to ensure the global convergence. Under the local error bound condition, we prove that the new LM method has at least a superlinear convergence rate with the order $\min\{1+\delta, 4-\delta, 2\}$. We also apply the new LM method to solve the nonlinear equations arising from the weighted linear complementarity problem. Numerical experiments indicate that the new LM method is efficient and promising.

Keywords Nonlinear equations, Levenberg-Marquardt method, nonmonotone technique, local error bound, weighted linear complementarity problem.

MSC(2010) 65K05, 90C30.

1. Introduction

We consider the system of nonlinear equations

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

where $F(x) : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function. Throughout the paper, we write the Jacobian $F'(x)$ as $J(x)$ and use the notions $F_k = F(x_k)$ and $J_k = J(x_k)$.

As it is well-known, the Levenberg-Marquardt (LM) method is one of the most effective methods for solving the nonlinear equations (1.1). At every iteration, the LM method computes the LM step

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

where λ_k is the LM parameter updated from iteration to iteration. The LM parameter λ_k has a great influence on the numerical performance and theoretical

[†]The corresponding author.
[1]College of Mathematics and Statistics, Xinyang Normal University, 464000 Xinyang, China
[*]This work was supported by the Henan Province Natural Science Foundation (222300420520) and the Key Scientific Research Projects of Higher Education of Henan Province (22A110020).
Email: zly976557281@163.com(L. Zhao), tangjy@xynu.edu.cn(J. Tang)

results of the LM method. Yamashita and Fukushima [26] showed, under the local error bound condition which is weaker than nonsingularity, the LM method has quadratic convergence if the LM parameter is chosen as $\lambda_k = \|F_k\|^2$. Under the same condition, Fan and Yuan [12] proved that the LM method taking $\lambda_k = \|F_k\|$ has the quadratic convergence. Although the numerical results in [12] show that the choice of $\lambda_k = \|F_k\|$ performs better than that of $\lambda_k = \|F_k\|^2$, it does not perform very well when the sequence $\{x_k\}$ is far from the solution set. To overcome this difficulty, Fan [9] used $\lambda_k = \mu_k\|F_k\|$ with μ_k being updated from iteration to iteration by trust region techniques. In [13], Fan and Yuan extended the results in [9, 26] and proved that the LM method taking $\lambda_k = \|F_k\|^\delta$ where $\delta \in [1, 2]$ still achieves the quadratic convergence under the local error bound condition. Besides $\lambda_k = O(\|F_k\|)$, many researchers studied the convergence properties of the LM method with $\lambda_k = O(\|J_k^T F_k\|)$ (e.g., [24, 25, 30]). Ma and Jiang [15] took $\lambda_k = (1-\theta)\|F_k\| + \theta\|J_k^T F_k\|$ where $\theta \in [0, 1]$ and proved that the LM method has quadratic convergence under the local error bound condition. Some other choices of the LM parameter are given [2, 10, 11].

On the other hand, many LM methods used the trust region technique to ensure the global convergence (e.g., [9, 10, 24, 28]). Define the actual reduction and the predicted reduction of $\|F(x)\|^2$ at the k-th iteration as

$$Ared_k = \|F_k\|^2 - \|F(x_k + d_k)\|^2, \tag{1.2}$$

and

$$Pred_k = \|F_k\|^2 - \|F_k + J_k d_k\|^2. \tag{1.3}$$

The ratio of the actual reduction to the predicted reduction

$$r_k = \frac{Ared_k}{Pred_k} \tag{1.4}$$

has been used in the LM methods to decide whether to accept the LM step and how to adjust the parameter μ_k. Recently, many researchers generalized the nonmonotone techniques to trust region methods and proposed some efficient nonmonotone trust region methods (e.g., [1, 8, 23, 27]). A lot of numerical experiments show that the algorithms with nonmonotone strategies are more efficient than the algorithms with monotone strategies.

Motivated by all of the work cited above, in this paper we aim to propose a new LM method which takes the LM parameter

$$\lambda_k = \mu_k[(1-\theta)\|F_k\|^\delta + \theta\|J_k^T F_k\|^\delta], \quad \text{where } \theta \in [0, 1] \text{ and } \delta \in (0, 3). \tag{1.5}$$

This new LM parameter is very general which includes the LM parameters used in [9, 13, 15] as special cases. Moreover, the new LM method adopts a nonmonotone trust region technique to ensure its global convergence. Under the local error bound condition, we prove that the new LM method has at least a superlinear convergence rate with the order $\min\{1+\delta, 4-\delta, 2\}$. This convergence result is more general than those obtained in [9, 13, 15]. We also apply the new LM method to solve the nonlinear equations arising from the weighted linear complementarity problem. Numerical experiments show the local fast convergence rate and the advantages of the new LM method.

The paper is organized as follows. In Section 2, we give a detailed description of the new LM method and establish its global convergence. In Section 3, we derive the convergence order of the new LM method under the local error bound condition. In Section 4, we apply the new LM method to solve some nonlinear equations and report some numerical results. Finally, we deliver some conclusions in Section 5.

2. The new LM method and its global convergence

In this section, we first give a detailed description of the new LM method and then prove its global convergence.

Algorithm 2.1 (A new LM method for nonlinear equations)
Step 0: Choose $\mu_0 > m_0 > 0$, $0 < p_0 \leq p_1 \leq p_2 < 1$, $\theta \in [0,1]$, $\tau \in (0,1]$ and $\delta \in (0,3)$. Choose $x_0 \in \mathbb{R}^n$ and set $\mathcal{W}_0 = \|F_0\|^2$. Set $k := 0$.
Step 1: If $\|J_k^T F_k\| = 0$, then stop. Otherwise, set

$$\lambda_k = \mu_k[(1-\theta)\|F_k\|^\delta + \theta\|J_k^T F_k\|^\delta]. \tag{2.1}$$

Step 2: Compute d_k by solving the following system

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F_k. \tag{2.2}$$

Step 3: Compute $Pred_k$ by (1.3) and

$$\widetilde{Ared}_k = \mathcal{W}_k - \|F(x_k + d_k)\|^2. \tag{2.3}$$

Set

$$\tilde{r}_k = \frac{\widetilde{Ared}_k}{Pred_k}. \tag{2.4}$$

Step 4: Set

$$x_{k+1} = \begin{cases} x_k + d_k & \text{if } \tilde{r}_k \geq p_0, \\ x_k & \text{otherwise.} \end{cases} \tag{2.5}$$

Set

$$\mathcal{W}_{k+1} = (1-\tau)\mathcal{W}_k + \tau\|F_{k+1}\|^2. \tag{2.6}$$

Step 5: Choose μ_{k+1} as

$$\mu_{k+1} = \begin{cases} 4\mu_k & \text{if } \tilde{r}_k < p_1, \\ \mu_k & \text{if } \tilde{r}_k \in [p_1, p_2], \\ \max\{\frac{\mu_k}{4}, m\} & \text{otherwise.} \end{cases} \tag{2.7}$$

Set $k = k + 1$ and go to Step 1.

Remark 2.1. There are two notable differences of the new LM method from existing LM methods. First, the LM parameter defined by (2.1) allows $\delta \in (0,3)$ which is more general than those used in existing LM methods where one usually requires $\delta \in (0,2]$. Second, Algorithm 2.1 adopts a nonmonotone trust region technique. It is noticeable that $\mathcal{W}_k$ is a convex combination of $\mathcal{W}_{k-1}$ and $\|F_k\|^2$. Since $\mathcal{W}_0 = \|F_0\|^2$, it follows that $\mathcal{W}_k$ is a convex combination of $\|F_0\|^2, \|F_1\|^2, ..., \|F_k\|^2$.

Lemma 2.1. *The predicted reduction $Pred_k$ defined by* (1.3) *satisfies*

$$Pred_k \geq \|J_k^T F_k\| \min\left\{\|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}. \tag{2.8}$$

Proof. The result can be found in [9, Lemma 3.1]. □

Lemma 2.2. *The sequence $\{x_k\}$ generated by Algorithm* 2.1 *satisfies $\|F_k\|^2 \leq \mathcal{W}_k$, $\mathcal{W}_{k+1} \leq \mathcal{W}_k$ and $\|F_k\| \leq \|F_0\|$ for all $k \geq 0$.*

Proof. First, we prove $\|F_k\|^2 \leq \mathcal{W}_k$ for all $k \geq 0$. Suppose that $\|F_k\|^2 \leq \mathcal{W}_k$ holds for some k. If $\tilde{r}_k < p_0$, then by (2.5) we have $x_{k+1} = x_k$ and so

$$\mathcal{W}_k \geq \|F_k\|^2 = \|F_{k+1}\|^2. \tag{2.9}$$

Otherwise, $\tilde{r}_k \geq p_0$ and by (2.5) we have $x_{k+1} = x_k + d_k$. Then it follows from (2.3) and (2.4) that

$$\tilde{r}_k = \frac{\mathcal{W}_k - \|F(x_k + d_k)\|^2}{Pred_k} = \frac{\mathcal{W}_k - \|F_{k+1}\|^2}{Pred_k} \geq p_0,$$

which together with (2.8) implies

$$\mathcal{W}_k \geq \|F_{k+1}\|^2 + p_0 Pred_k \geq \|F_{k+1}\|^2. \tag{2.10}$$

Thus, we have $\|F_{k+1}\|^2 \leq \mathcal{W}_k$ which together with (2.6) yields

$$\|F_{k+1}\|^2 \leq (1-\tau)\mathcal{W}_k + \tau\|F_{k+1}\|^2 = \mathcal{W}_{k+1}. \tag{2.11}$$

Since $\|F_0\|^2 = \mathcal{W}_0$, by induction on k, we obtain $\|F_k\|^2 \leq \mathcal{W}_k$ for all $k \geq 0$. Moreover, by (2.9) and (2.10), it holds that $\|F_{k+1}\|^2 \leq \mathcal{W}_k$ for all $k \geq 0$. This together with (2.6) gives for all $k \geq 0$,

$$\mathcal{W}_{k+1} \leq (1-\tau)\mathcal{W}_k + \tau\mathcal{W}_k = \mathcal{W}_k.$$

Furthermore, we have $\|F_k\|^2 \leq \mathcal{W}_k \leq \mathcal{W}_0 = \|F_0\|^2$ for all $k \geq 0$. The proof is completed. □

To establish the global convergence of Algorithm 2.1, we make the following assumption.

Assumption 2.1. *The Jacobian $J(x)$ is bounded and Lipschitz continuous on $\mathbb{R}^n$, i.e., there exist positive constants L_1 and L_2 such that*

$$\|J(x)\| \leq L_1, \ \forall\, x \in \mathbb{R}^n, \tag{2.12}$$

and

$$\|J(y) - J(x)\| \leq L_2\|y - x\|, \quad \forall\, x, y \in \mathbb{R}^n. \tag{2.13}$$

By (2.13), we have

$$\|F(y) - F(x) - J(x)(y-x)\| \leq L_2\|y - x\|^2, \quad \forall\, x, y \in \mathbb{R}^n. \tag{2.14}$$

Theorem 2.1. *Under the conditions of Assumption* 2.1, *the sequence $\{x_k\}$ generated by Algorithm* 2.1 *satisfies*

$$\liminf_{k\to\infty} \|J_k^T F_k\| = 0. \tag{2.15}$$

Proof. If the theorem is not true, then there exist a constant $\eta > 0$ and an index $\bar{k}$ such that

$$\|J_k^T F_k\| \geq \eta, \quad \forall\ k \geq \bar{k}. \tag{2.16}$$

By the second result in Lemma 2.2, the sequence $\{\mathcal{W}_k\}$ is monotonically decreasing and bounded below. Thus, there exists a constant $\mathcal{W}^* \geq 0$ such that $\lim_{k\to\infty} \mathcal{W}_k = \mathcal{W}^*$. Furthermore, by (2.6) we have

$$\lim_{k\to\infty} \|F_k\|^2 = \lim_{k\to\infty} \frac{\mathcal{W}_k - (1-\tau)\mathcal{W}_{k-1}}{\tau} = \mathcal{W}^*.$$

Define the set of successful iterations as

$$K = \{k | \tilde{r}_k \geq p_0\}.$$

We derive the contradictions in two cases.

Case 1. K is infinite. In this case, by (2.8), (2.12) and (2.16), we have for all $k \in K$ and $k \geq \bar{k}$,

$$\begin{aligned} \mathcal{W}_k - \|F_{k+1}\|^2 &= \mathcal{W}_k - \|F(x_k + d_k)\|^2 \\ &= \widetilde{Ared}_k \\ &\geq p_0 Pred_k \\ &\geq p_0 \|J_k^T F_k\| \min\left\{\|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\} \\ &\geq p_0 \eta \min\left\{\|d_k\|, \frac{\eta}{L_1^2}\right\}. \end{aligned}$$

It follows from $\lim_{k\to\infty} \|F_k\|^2 = \lim_{k\to\infty} \mathcal{W}_k = \mathcal{W}^*$ that $\lim_{(K\ni)k\to\infty} d_k = 0$. Note that $x_{k+1} - x_k = 0$ if $k \notin K$. Thus, we have

$$\lim_{k\to\infty} d_k = 0. \tag{2.17}$$

This together with (2.2) and (2.16) yields

$$\lim_{k\to\infty} \lambda_k = +\infty. \tag{2.18}$$

Due to the third result in Lemma 2.2 and (2.12), we have

$$(1-\theta)\|F_k\|^\delta + \theta \|J_k^T F_k\|^\delta \leq (1-\theta)\|F_0\|^\delta + \theta L_1^\delta \|F_0\|^\delta.$$

So, by (2.1) and (2.18) we have

$$\lim_{k\to\infty} \mu_k = +\infty. \tag{2.19}$$

Moreover, by the result given in the proof of [24, Theorem 2.4], we have

$$|\|F(x_k + d_k)\|^2 - \|F_k + J_k d_k\|^2| \leq \|F_k + J_k d_k\| O(\|d_k\|^2) + O(\|d_k\|^4). \tag{2.20}$$

Thus, from (2.8), (2.16) and (2.20), we have for $k \geq \bar{k}$,

$$|r_k - 1| = \left|\frac{Ared_k - Pred_k}{Pred_k}\right|$$

$$
\begin{aligned}
&\le \frac{|\|F(x_k+d_k)\|^2 - \|F_k + J_k d_k\|^2|}{\|J_k^T F_k\| \min\left\{\|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}} \\
&\le \frac{\|F_k + J_k d_k\| O(\|d_k\|^2) + O(\|d_k\|^4)}{\eta \min\left\{\|d_k\|, \frac{\eta}{L_1^2}\right\}} \\
&= \frac{\|F_k + J_k d_k\| O(\|d_k\|^2) + O(\|d_k\|^4)}{\|d_k\|}.
\end{aligned} \tag{2.21}
$$

Since $\|F_k + J_k d_k\| \le \|F_0\| + L_1\|d_k\|$, the inequality (2.21) yields $r_k \to 1$ as $k \to \infty$. Since $\mathcal{W}_k \ge \|F_k\|^2$ for all $k \ge 0$, it holds that $\widetilde{Ared}_k \ge Ared_k$ and so

$$
\tilde{r}_k = \frac{\widetilde{Ared}_k}{Pred_k} \ge \frac{Ared_k}{Pred_k} = r_k \to 1.
$$

In view of the updating rule of μ_k, there exists a positive constant $\tilde{m} > m$ such that $\mu_k < \tilde{m}$ holds for all large k, which is a contradiction to (2.19).

Case 2. K is finite. In this case, there exists an index $\hat{k}$ such that $\tilde{r}_k < p_0$ for all $k > \hat{k}$. By Step 5 of Algorithm 2.1, we have $\mu_{k+1} = 4\mu_k$ for all $k > \hat{k}$, which yields

$$
\lim_{k\to\infty} \mu_k = +\infty. \tag{2.22}
$$

By (2.12) and (2.16), we have

$$
\|F_k\| \ge \frac{\|J_k^T F_k\|}{L_1} \ge \frac{\eta}{L_1}, \quad \forall\, k \ge \bar{k}.
$$

It follows that

$$
(1-\theta)\|F_k\|^\delta + \theta\|J_k^T F_k\|^\delta \ge \frac{(1-\theta)\eta^\delta}{L_1^\delta} + \theta\eta^\delta > 0, \quad \forall\, k \ge \bar{k}.
$$

Thus, by (2.1) and (2.22) we have

$$
\lim_{k\to\infty} \lambda_k = +\infty, \tag{2.23}
$$

which together with (2.2) and (2.16) gives

$$
\lim_{k\to\infty} d_k = 0.
$$

By the same analysis as Case 1, we have $\tilde{r}_k \ge r_k \to 1$ as $k \to \infty$. Thus, there exists a constant $\hat{m} > m$ such that $\mu_k \le \hat{m}$ holds for all large k, which is a contradiction to (2.22).

Summarizing Case 1 and Case 2, we have (2.15) and complete the proof. □

3. Convergence rate of Algorithm 2.1

In this section, we analyze the convergence rate of Algorithm 2.1. We assume that the sequence $\{x_k\}$ generated by Algorithm 2.1 converges to the solution set X^* of the nonlinear equations (1.1) and lies in some neighbourhood of $x^* \in X^*$. We make the following assumption.

Assumption 3.1. (a) *$F(x)$ is continuously differentiable and $\|F(x)\|$ provides a local error bound on some neighbourhood of $x^* \in X^*$, i.e., there exist positive constants $\kappa > 0$ and $\varepsilon > 0$ such that*

$$\|F(x)\| \geq \kappa \text{dist}(x, X^*), \quad \forall\, x \in N(x^*, \varepsilon) = \{x \| \|x - x^*\| \leq \varepsilon\}. \tag{3.1}$$

(b) *The Jacobian $J(x)$ is Lipschitz continuous on $N(x^*, \varepsilon)$, i.e., there exists a constant $L > 0$ such that*

$$\|J(y) - J(x)\| \leq L\|y - x\|,\ \forall\, x, y \in N(x^*, \varepsilon). \tag{3.2}$$

By the Lipschitzness of Jacobian given in (3.2), we have

$$\|F(y) - F(x) - J(x)(y - x)\| \leq L\|y - x\|^2,\ \forall\, x, y \in N(x^*, \varepsilon). \tag{3.3}$$

Thus, there exists a constant $M > 0$ such that

$$\|F(y) - F(x)\| \leq M\|y - x\|,\ \forall\, x, y \in N(x^*, \varepsilon). \tag{3.4}$$

Moreover, by (3.4) we have (see [31])

$$\|J(x)\| \leq M,\ \forall\, x \in N(x^*, \varepsilon). \tag{3.5}$$

Due to the result given by Behling and Iusem in [4, Theorem 1], if $\|F(x)\|$ provides a local error bound, then there exists a positive constant $\zeta > 0$ such that

$$\text{rank}(J(\bar{x})) = \text{rank}(J(x^*)), \quad \forall\, \bar{x} \in N(x^*, \zeta) \cap X^*.$$

We assume without loss of generality that $\text{rank}(J(\bar{x})) = r$ for all $\bar{x} \in N(x^*, \zeta) \cap X^*$. Suppose that the singular value decomposition (SVD) of $J(\bar{x}_k)$ is

$$\bar{J}_k = \bar{U}_k \bar{\Sigma}_k \bar{V}_k^T = (\bar{U}_{k,1}, \bar{U}_{k,2}) \begin{pmatrix} \bar{\Sigma}_{k,1} & \\ & 0 \end{pmatrix} \begin{pmatrix} \bar{V}_{k,1}^T \\ \bar{V}_{k,2}^T \end{pmatrix} = \bar{U}_{k,1} \bar{\Sigma}_{k,1} \bar{V}_{k,1}^T,$$

where $\bar{\Sigma}_{k,1} = \text{diag}(\bar{\sigma}_{k,1}, \cdots, \bar{\sigma}_{k,r}) > 0$, and correspondingly the SVD of J_k is

$$\begin{aligned} J_k &= U_k \Sigma_k V_k^T \\ &= (U_{k,1}, U_{k,2}) \begin{pmatrix} \Sigma_{k,1} & \\ & \Sigma_{k,2} \end{pmatrix} \begin{pmatrix} V_{k,1}^T \\ V_{k,2}^T \end{pmatrix} \\ &= U_{k,1} \Sigma_{k,1} V_{k,1}^T + U_{k,2} \Sigma_{k,2} V_{k,2}^T, \end{aligned}$$

where $\Sigma_{k,1} = \text{diag}(\sigma_{k,1}, \cdots, \sigma_{k,r}) > 0$ and $\Sigma_{k,2} = \text{diag}(\sigma_{k,r+1}, \cdots, \sigma_{k,n}) \geq 0$. In the following, if the context is clear, we neglect the subscription k in $\Sigma_{k,i}$ and $U_{k,i}$,$V_{k,i} (i = 1, 2)$ and write J_k as

$$J_k = U_1 \Sigma_1 V_1^T + U_2 \Sigma_2 V_2^T.$$

In the following, we denote $\bar{x}_k$ as the vector in X^* that satisfies

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

The following lemma gives the estimations of $\|U_1 U_1^T F_k\|$ and $\|U_2 U_2^T F_k\|$ whose proof can be found in [24, Lemma 3.4].

Lemma 3.1. *Under the conditions of Assumption* 3.1*, for all sufficiently large* k,
(a) $\|U_1U_1^TF_k\| \le M\|\bar{x}_k - x_k\|$;
(b) $\|U_2U_2^TF_k\| \le 2L\|\bar{x}_k - x_k\|^2$,
where M *and* L *are given in* (3.4) *and* (3.2) *respectively.*

Lemma 3.2. *Under the conditions of Assumption* 3.1*, there exists a constant* $c > 0$ *such that for all sufficiently large* k,

$$c\|\bar{x}_k - x_k\| \le \|J_k^TF_k\| \le M^2\|\bar{x}_k - x_k\|, \tag{3.6}$$

where M *is given in* (3.4).

Proof. For all $x_k \in N(x^*, \varepsilon/2)$, we have

$$\|\bar{x}_k - x^*\| \le \|\bar{x}_k - x_k\| + \|x_k - x^*\| \le 2\|x_k - x^*\| \le \varepsilon,$$

which implies that $\bar{x}_k \in N(x^*, \varepsilon)$. Then, by (3.4) and (3.5), we have for all sufficiently large k,

$$\|J_k^TF_k\| \le \|J_k\|\|F_k\| \le M^2\|\bar{x}_k - x_k\|,$$

which proves the right inequality in (3.6). Moreover, by (3.3) and (3.4), we can further obtain that for all sufficiently large k,

$$F_k^T(F_k - J_k(x_k - \bar{x}_k)) \le ML\|\bar{x}_k - x_k\|^3. \tag{3.7}$$

It follows from (3.1) and (3.7) that for all sufficiently large k,

$$\begin{aligned}
\|F_k^TJ_k\| &\ge \frac{F_k^TJ_k(x_k - \bar{x}_k)}{\|x_k - \bar{x}_k\|} \\
&= \frac{\|F_k\|^2 - F_k^T(F_k - J_k(x_k - \bar{x}_k))}{\|x_k - \bar{x}_k\|} \\
&\ge \frac{\kappa\|\bar{x}_k - x_k\|^2 - ML\|\bar{x}_k - x_k\|^3}{\|x_k - \bar{x}_k\|} \\
&= \kappa\|\bar{x}_k - x_k\| - ML\|\bar{x}_k - x_k\|^2 \\
&\ge c\|\bar{x}_k - x_k\|,
\end{aligned}$$

where $c > 0$ is some constant. This proves the left inequality in (3.6). □

Lemma 3.3. *Under the conditions of Assumption* 3.1*, there exists a constant* $\tilde{c} > 0$ *such that for all sufficiently large* k,

$$\|d_k\| \le \tilde{c}\|\bar{x}_k - x_k\|^{\min\{2-\frac{\delta}{2},1\}}. \tag{3.8}$$

Proof. By $\mu_k \ge m$, (3.1) and the left inequality in (3.6), we have from (2.1) that

$$\lambda_k \ge m_0[(1-\theta)\kappa^\delta + \theta c^\delta]\|\bar{x}_k - x_k\|^\delta. \tag{3.9}$$

For any $k \ge 0$, since d_k is a solution of the following minimization problem

$$\min_{d\in\mathbb{R}^n} \varphi_k(d) := \|F_k + J_kd\|^2 + \lambda_k\|d\|^2, \tag{3.10}$$

by (3.3) and (3.9), we have for all sufficiently large k,

$$\|d_k\|^2 \le \frac{\varphi_k(d_k)}{\lambda_k}$$

$$\begin{aligned}
&\le \frac{\varphi_k(\bar{x}_k - x_k)}{\lambda_k} \\
&= \frac{\|F_k + J_k(\bar{x}_k - x_k)\|^2}{\lambda_k} + \|\bar{x}_k - x_k\|^2 \\
&\le \frac{L^2}{m_0[(1-\theta)\kappa^\delta + \theta c^\delta]}\|\bar{x}_k - x_k\|^{4-\delta} + \|\bar{x}_k - x_k\|^2 \\
&\le \left(\frac{L^2}{m_0[(1-\theta)\kappa^\delta + \theta c^\delta]} + 1\right)\|\bar{x}_k - x_k\|^{\min\{4-\delta,2\}}.
\end{aligned}$$

By letting $\tilde{c} = \sqrt{L^2/m_0[(1-\theta)\kappa^\delta + \theta c^\delta] + 1}$, we have (3.8). $\square$

Lemma 3.4. *Under the conditions of Assumption* 3.1*, there exists a positive constant* $\Theta > m$ *such that*

$$\mu_k \le \Theta \tag{3.11}$$

holds for all sufficiently large k.

Proof. First we prove that for all sufficiently large k, the predicted reduction $Pred_k$ satisfies

$$Pred_k \ge \bar{c}\|F_k\|\|d_k\|^{\max\{1,\frac{2}{4-\delta}\}}, \tag{3.12}$$

where $\bar{c} > 0$ is some constant. We consider two cases. If $\|\bar{x}_k - x_k\| \le \|d_k\|$, then by (3.1), (3.3), (3.8) and the fact that d_k is the solution of (3.10), we have

$$\begin{aligned}
\|F_k\| - \|F_k + J_k d_k\| &\ge \|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\| \\
&\ge \kappa\|\bar{x}_k - x_k\| - L\|\bar{x}_k - x_k\|^2 \\
&\ge \bar{c}_1\|\bar{x}_k - x_k\| \\
&\ge \bar{c}_2\|d_k\|^{\max\{1,\frac{2}{4-\delta}\}},
\end{aligned} \tag{3.13}$$

where $\bar{c}_1, \bar{c}_2 > 0$ are some constants. Otherwise, $\|\bar{x}_k - x_k\| < \|d_k\|$. In this case, by the third inequality of (3.13), we have

$$\begin{aligned}
&\|F_k\| - \|F_k + J_k d_k\| \\
&\ge \|F_k\| - \left\|F_k + \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}J_k(\bar{x}_k - x_k)\right\| \\
&\ge \|F_k\| - \left\|\left(1 - \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}\right)F_k + \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}(F_k + J_k(\bar{x}_k - x_k))\right\| \\
&\ge \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}(\|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\|) \\
&\ge \bar{c}_1\|d_k\|.
\end{aligned} \tag{3.14}$$

Thus, by (3.13) and (3.14), for all sufficiently large k,

$$\begin{aligned}
Pred_k &= \|F_k\|^2 - \|F_k + J_k d_k\|^2 \\
&= (\|F_k\| + \|F_k + J_k d_k\|)(\|F_k\| - \|F_k + J_k d_k\|) \\
&\ge \bar{c}\|F_k\|\|d_k\|^{\max\{1,\frac{2}{4-\delta}\}},
\end{aligned}$$

where $\bar{c} > 0$ is some constant. Since $\delta \in (0, 3)$, we have $\max\{1, \frac{2}{4-\delta}\} < 2$. Also note that $\|F_k + J_k d_k\| \le \|F_k\|$ by (3.13) and (3.14). Thus, by (2.20) and (3.12), for all sufficiently large k,

$$
\begin{aligned}
|r_k - 1| &= \left| \frac{Ared_k - Pred_k}{Pred_k} \right| \\
&\leq \frac{|\|F(x_k + d_k)\|^2 - \|F_k + J_k d_k\|^2}{\bar{c}\|F_k\|\|d_k\|^{\max\{1, \frac{2}{4-\delta}\}}} \\
&\leq \frac{\|F_k + J_k d_k\| O(\|d_k\|^2) + O(\|d_k\|^4)}{\bar{c}\|F_k\|\|d_k\|^{\max\{1, \frac{2}{4-\delta}\}}} \\
&\to 0.
\end{aligned} \tag{3.15}
$$

Furthermore, we have

$$
\tilde{r}_k = \frac{\widetilde{Ared_k}}{Pred_k} \geq \frac{Ared_k}{Pred_k} = r_k \to 1.
$$

Hence, there exists a positive constant $\Theta > m$ such that $\mu_k < \Theta$ holds for all sufficiently large k. The proof is completed. □

Now we give the convergence order of Algorithm 2.1 as follows.

Theorem 3.1. *Under the conditions of Assumption* 3.1*, the sequence* $\{x_k\}$ *converges to the solution set* X^* *at least superlinearly with the order* $\min\{1+\delta, 4-\delta, 2\}$

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

$$
\|\mathrm{diag}(\Sigma_1 - \bar{\Sigma}_1, \Sigma_2)\| \leq \|J_k - \bar{J}_k\| \leq L\|\bar{x}_k - x_k\|,
$$

which gives

$$
\|\Sigma_1 - \bar{\Sigma}_1\| \leq L\|\bar{x}_k - x_k\| \quad \text{and} \quad \|\Sigma_2\| \leq L\|\bar{x}_k - x_k\|. \tag{3.16}
$$

Since $\{x_k\}$ converges to the solution set X^*, we assume that $L\|\bar{x}_k - x_k\| \leq \bar{\sigma}_r/2$ holds for all sufficiently large k. Then it follows from (3.16) that for all sufficiently large k

$$
\|(\Sigma_1^2 + \lambda_k I)^{-1}\| \leq \|\Sigma_1^{-2}\| \leq \frac{1}{(\bar{\sigma}_r - L\|\bar{x}_k - x_k\|)^2} \leq \frac{4}{\bar{\sigma}_r^2}. \tag{3.17}
$$

Moreover, by Lemma 3.4, (3.4) and the right inequality in (3.6), we have from (2.1) that

$$
\lambda_k \leq \Theta[(1-\theta)M^\delta + \theta M^{2\delta}]\|\bar{x}_k - x_k\|^\delta. \tag{3.18}
$$

By the SVD of J_k, we compute

$$
d_k = -V_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F_k - V_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F_k.
$$

So, we have

$$
\begin{aligned}
F_k + J_k d_k &= F_k - U_1\Sigma_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F_k - U_2\Sigma_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F_k \\
&= \lambda_k U_1(\Sigma_1^2 + \lambda_k I)^{-1} U_1^T F_k + \lambda_k U_2(\Sigma_2^2 + \lambda_k I)^{-1} U_2^T F_k,
\end{aligned}
$$

which together with Lemma 3.1, (3.17), (3.18) and $\|(\Sigma_2^2 + \lambda_k I)^{-1}\| \leq \lambda_k^{-1}$ yields

$$
\|F_k + J_k d_k\| \leq \lambda_k \|(\Sigma_1^2 + \lambda_k I)^{-1}\| \|\|U_1^T F_k\| + \|U_2^T F_k\|\|
$$

$$\leq \frac{4\Theta[(1-\theta)M^{\delta}+\theta M^{2\delta}]}{\bar{\sigma}_r^2}\|\bar{x}_k-x_k\|^{1+\delta}+2L\|\bar{x}_k-x_k\|^2$$
$$\leq C\|\bar{x}_k-x_k\|^{\min\{1+\delta,2\}}, \tag{3.19}$$

where $C = 4\Theta[(1-\theta)M^{\delta}+\theta M^{2\delta}]/\bar{\sigma}_r^2+2L$. Furthermore, by (3.1), (3.3), (3.8) and (3.19), we have

$$\begin{aligned}
\mathrm{dist}(x_{k+1},X^*) &= \|\bar{x}_{k+1}-x_{k+1}\| \\
&\leq \frac{1}{\kappa}\|F(x_{k+1})\| = \frac{1}{\kappa}\|F(x_k+d_k)\| \\
&\leq \frac{1}{\kappa}(\|F_k+J_kd_k\|+L\|d_k\|^2) \\
&\leq \frac{1}{\kappa}\big(C\|\bar{x}_k-x_k\|^{\min\{1+\delta,2\}}+L\tilde{c}^2\|\bar{x}_k-x_k\|^{\min\{4-\delta,2\}}\big) \\
&\leq \frac{C+L\tilde{c}^2}{\kappa}\|\bar{x}_k-x_k\|^{\min\{1+\delta,4-\delta,2\}} \\
&= O(\mathrm{dist}(x_k,X^*)^{\min\{1+\delta,4-\delta,2\}}).
\end{aligned}$$

The proof is completed. □

Remark 3.1. (a) Theorem 3.1 indicates that the sequence $\{x_k\}$ converges to the solution set X^* superlinearly for $\delta \in (0,1)$, and quadratically for $\delta \in [1,2]$. These results are the same as those obtained for the LM method (e.g., [10,13]). However, Theorem 3.1 also shows that $\{x_k\}$ converges to X^* superlinearly with the order $4-\delta$ when $\delta \in (2,3)$. Therefore, Theorem 3.1 generalizes existing convergence results of the LM method.
(b) By (3.1), (3.4) and Theorem 3.1, we have

$$\begin{aligned}
\|F_{k+1}\| = O(\|\bar{x}_{k+1}-x_{k+1}\|) &= O(\mathrm{dist}(x_k,X^*)^{\min\{1+\delta,4-\delta,2\}}) \\
&= O(\|F_k\|^{\min\{1+\delta,4-\delta,2\}}).
\end{aligned}$$

This indicates that the sequence $\{\|F_k\|\}$ converges to zero at least superlinearly with the order $\min\{1+\delta,4-\delta,2\}$.

4. Application to weighted linear complementarity problems

Numerical performances of the LM method for solving some singular problems have been done in [2,10,13,15,30] which clearly show the efficiency of the LM method. In this section, we pay particular attention to the performances of the LM method for solving nonlinear equations arising from the weighted linear complementarity problem (wLCP).

4.1. Nonlinear equations arising from wLCP

The weighted linear complementarity problem (wLCP) was introduced by Potra [16] which is to find vectors $x \in \mathbb{R}^n, s \in \mathbb{R}^n, y \in \mathbb{R}^m$ such that

$$\text{(wLCP)} \quad x \geq 0, \quad s \geq 0, \quad Px+Qs+Ry=a, \quad xs=w. \tag{4.1}$$

Here $P \in \mathbb{R}^{(n+m)\times n}, Q \in \mathbb{R}^{(n+m)\times n}, R \in \mathbb{R}^{(n+m)\times m}$ are given matrices, $a \in \mathbb{R}^{n+m}$ is a given vector, $w \geq 0$ is a given weight vector (the data of the problem) and xs is the componentwise product of the vectors x and s. The significance of studying the wLCP lies in the fact that a lot of equilibrium problems in economics can be formulated in a natural way as wLCP [16]. Moreover, those formulations lend themselves to the development of highly efficient algorithms for solving the corresponding equilibrium problems [16]. For example, the Fisher market equilibrium problem, which can be modelled as a nonlinear CP, can also be formulated as a wLCP that can be efficiently solved by interior-point methods [16]. In recent years, the wLCP has received considerable attention from researchers (see, [3, 6, 7, 14, 17, 19–22, 29]).

To equivalently reformulate the wLCP as a system of nonlinear equations, we consider the following weighted complementarity function

$$\phi^c(a,b) = (a+b)^3 - \left(\sqrt{a^2+b^2+2c}\right)^3,\ \forall (a,b) \in \mathbb{R}^2,$$

where $c \geq 0$ is a constant.

Lemma 4.1. (a) *The function ϕ^c satisfies*

$$\phi^c(a,b) = 0 \Longleftrightarrow a \geq 0,\ b \geq 0,\ ab = c.$$

(b) *The function ϕ^c is continuously differentiable at any $(a,b) \in \mathbb{R}^2$ with*

$$\nabla \phi^c(a,b) = \begin{pmatrix} 3[(a+b)^2 - a\sqrt{a^2+b^2+2c}] \\ 3[(a+b)^2 - b\sqrt{a^2+b^2+2c}] \end{pmatrix}.$$

Let $z := (x,s,y)$. Then, due to Lemma 4.1, solving the wLCP is equivalent to computing a solution of the following nonlinear equations

$$F(z) = \begin{pmatrix} Px + Qs + Ry - a \\ \phi^{w_1}(x_1, s_1) \\ \vdots \\ \phi^{w_n}(x_n, s_n) \end{pmatrix} = 0, \tag{4.2}$$

where $w = (w_1, ..., w_n)^T$ is the weight vector given in the wLCP. Since the function $F(z)$ is continuously differentiable at any $z \in \mathbb{R}^{2n+m}$, we can apply the LM method to solve the nonlinear equations (4.2) so that a solution of the wLCP can be obtained.

By Lemma 4.1 (b), the Jacobian of $F(z)$ is given as

$$J(z) = \begin{bmatrix} P & Q & R \\ \mathrm{diag}\left(\frac{\partial \phi^{w_i}}{\partial x_i}\right) & \mathrm{diag}\left(\frac{\partial \phi^{w_i}}{\partial s_i}\right) & 0 \end{bmatrix}, \tag{4.3}$$

where

$$\frac{\partial \phi^{w_i}}{\partial x_i} = 3\left[(x_i+s_i)^2 - x_i\sqrt{x_i^2+s_i^2+2w_i}\right],$$

$$\frac{\partial \phi^{w_i}}{\partial s_i} = 3\big[(x_i + s_i)^2 - s_i\sqrt{x_i^2 + s_i^2 + 2w_i}\big].$$

In the following, we show that the Jacobian $J(z)$ satisfies the Lipschitz continuity, i.e., Assumption 3.1 (b) holds for the nonlinear equations (4.2).

Theorem 4.1. *The Jacobian $J(z)$ given in* (4.3) *is Lipschitz continuous on the closed and convex set $N(z) = \{z \in \mathbb{R}^{2n+m} | \|z\| \le \varrho\}$ for any $\varrho > 0$.*

Proof. Obviously, we only need to prove that the gradient $\nabla\phi^c(a,b)$ is Lipschitz continuous on the closed and convex set $\Omega := \{(a,b) \in \mathbb{R}^2 | \|(a,b)\| \le \zeta\}$ for any $\zeta > 0$. Let $h^c(a,b) = \sqrt{a^2 + b^2 + 2c}$. It is easy to see that

$$h^c(a,b) \le \sqrt{\zeta^2 + 2c}, \quad \forall\ (a,b) \in \Omega. \tag{4.4}$$

We consider the following three cases.

Case 1. $c > 0$. In this case, $h^c(a,b) > 0$ for any $(a,b) \in \Omega$. Thus, ϕ^c is twice continuously differentiable at any $(a,b) \in \Omega$ with

$$\nabla^2\phi^c(a,b) = \begin{bmatrix} \dfrac{\partial^2\phi^c}{\partial a^2} & \dfrac{\partial^2\phi^c}{\partial a\partial b} \\[2ex] \dfrac{\partial^2\phi^c}{\partial b\partial a} & \dfrac{\partial^2\phi^c}{\partial b^2} \end{bmatrix},$$

where

$$\begin{aligned} \frac{\partial^2\phi^c}{\partial a^2} &= 3\big\{2(a+b) - \big(a^2/h^c(a,b) + h^c(a,b)\big)\big\}, \\ \frac{\partial^2\phi^c}{\partial b^2} &= 3\big\{2(a+b) - \big(b^2/h^c(a,b) + h^c(a,b)\big)\big\}, \\ \frac{\partial^2\phi^c}{\partial a\partial b} &= \frac{\partial^2\phi^c}{\partial b\partial a} = 3\big\{2(a+b) - ab/h^c(a,b)\big\}. \end{aligned}$$

By (4.4), we have for any $(a,b) \in \Omega$,

$$\max\{a^2/h^c(a,b),\ b^2/h^c(a,b),\ ab/h^c(a,b)\} \le h^c(a,b) \le \sqrt{\zeta^2 + 2c}.$$

Thus, there exists a constant $C > 0$ independent of $(a,b) \in \Omega$ such that

$$\|\nabla^2\phi^c(a,b)\| \le C, \quad \forall\ (a,b) \in \Omega.$$

By Mean Value Theorem, we have that

$$\|\nabla\phi^c(a_1,b_1) - \nabla\phi^c(a_2,b_2)\| \le C\|(a_1,b_1) - (a_2,b_2)\|$$

holds for any $(a_1,b_1), (a_2,b_2) \in \Omega$ and prove the desired result.

Case 2. $c = 0$ and $(0,0) \notin \Omega$. In this case, $h^0(a,b) = \sqrt{a^2 + b^2} > 0$ for any $(a,b) \in \Omega$. Thus, ϕ^0 is twice continuously differentiable at any $(a,b) \in \Omega$. By following exactly the same steps as in the Case 1, we can prove the desired result.

Case 3. $c = 0$ and $(0,0) \in \Omega$. Then, similarly as Case 1, we can prove that there exists a constant $\bar{C} > 0$ independent of (a,b) such that

$$\|\nabla^2\phi^0(a,b)\| \le \bar{C}, \quad \forall\ (a,b) \ne (0,0) \in \Omega.$$

Then, by [5, Lemma 2.6], we have

$$\|\nabla\phi^0(a_1,b_1)-\nabla\phi^0(a_2,b_2)\|\leq\bar{C}\|(a_1,b_1)-(a_2,b_2)\| \tag{4.5}$$

holds for all $(a_1,b_1),(a_2,b_2)\in\Omega$ with $(0,0)\notin[(a_1,b_1),(a_2,b_2)]$. Moreover, since $\nabla\phi^0(0,0)=(0,0)$, the inequality (4.5) also holds in case $(a_1,b_1)=(a_2,b_2)=(0,0)$. Therefore, we can assume $(a_1,b_1)\neq(0,0)\in\Omega$. Since ϕ^0 is continuously differentiable for all $(a,b)\in\mathbb{R}^2$ with $\nabla\phi^0(0,0)=(0,0)$, by using a continuity argument, we obtain that the inequality (4.5) remains true for all $(a_2,b_2)\in\Omega$. Thus, the inequality (4.5) holds for all $(a_1,b_1),(a_2,b_2)\in\Omega$ which proves the desired result. □

4.2. Computational experiments

In this subsection, we apply Algorithm 2.1 to solve the nonlinear equations (4.2) with

$$P=\begin{pmatrix}A\\ M\end{pmatrix},\quad Q=\begin{pmatrix}0\\ -I\end{pmatrix},\quad R=\begin{pmatrix}0\\ -A^T\end{pmatrix},\quad a=\begin{pmatrix}b\\ -f\end{pmatrix}, \tag{4.6}$$

where $A\in\mathbb{R}^{m\times n}$ is a full row rank matrix with $m<n$, $b\in\mathbb{R}^m$, $f\in\mathbb{R}^n$ and $M\in\mathbb{R}^{n\times n}$ is an symmetric positive semidefinite matrix. It is worth pointing out that the wLCP (4.1) with (4.6) is the optimality conditions of the quadratic programming and weighted centering problem [16, Theorem 2.1]. In experiments, we generate a random matrix $A\in\mathbb{R}^{m\times n}$ with full row rank and set $M=BB^T/\|BB^T\|$ with $B=\mathbf{rand}(n,n)$. Then we choose $\hat{x}=\mathbf{rand}(n,1)$, $f=\mathbf{rand}(n,1)$ and set $b=A\hat{x}$, $\hat{s}=M\hat{x}+f$ and $w=\hat{x}\hat{s}$. The parameters used in Algorithm 2.1 are chosen as $p_0=10^{-4},p_1=0.25,p_2=0.75,m_0=10^{-8},\tau=0.5$ and μ_0,θ,δ are specified in the experiments.

First, to observe the local convergence behavior, we generate one test problem with $n=100$ and $m=50$ and solve it by Algorithm 2.1 with $\mu_0=10^{-4}$. We test the following LM parameters:

(i) $\theta=0$, i.e., $\lambda_k=\mu_k\|F_k\|^\delta$.

(ii) $\theta=0.5$, i.e., $\lambda_k=\mu_k\frac{\|F_k\|^\delta+\|J_k^TF_k\|^\delta}{2}$.

(iii) $\theta=1$, i.e., $\lambda_k=\mu_k\|J_kF_k\|^\delta$.

The starting point is chosen as $x_0=s^0=(1,...,1)^T$ and $y_0=(0,...,0)^T$. Table 1 gives the value of $\|F(z_k)\|$ at the k-th iteration.

From Table 1, three observations can be made here.

(a) Algorithm 2.1 has at least superlinear convergence rate for $\delta\in(0,3)$.

(b) Algorithm 2.1 taking $\delta\in[1,2)$ converges faster than that taking $\delta\in(0,1)\cup[2,3)$.

(c) The efficiency of Algorithm 2.1 is reduced in initial steps when $\delta\in[2,3)$.

These observations confirm the theoretical results of the new LM method.

Next, we further investigate the influences the the parameter θ on Algorithm 2.1. We test Algorithm 2.1 with $\delta=1$, i.e., $\lambda_k=\mu_k[(1-\theta)\|F_k\|+\theta\|J_k^TF_k\|]$. For each problem with different sizes $n(=2m)$, we generate five instances and solve them by Algorithm 2.1. For the purpose of comparison, we also apply the LM method studied by Fan [9] to solve these problems. It is worth pointing out that Fan's LM method [9] took $\lambda_k=\mu_k\|F_k\|$ with μ_k being updated by the trust region technique. The starting point is chosen as before. We use $\|F(z_k)\|<10^{-6}$ and $iter<30$

Table 1. The value of $\|F(z_k)\|$ at the k-th iteration

		$\delta = 0.6$	$\delta = 1.0$	$\delta = 1.5$	$\delta = 2$	$\delta = 2.2$
$\theta = 0$	$k = 1$	7.3940	7.3615	7.1490	44.9536	53.5820
	$k = 2$	1.4243	1.4095	1.5719	21.3399	51.9793
	$k = 3$	0.3603	0.2656	0.1937	3.6613	45.8181
	$k = 4$	0.0705	0.0427	0.0111	1.5099	26.2801
	$k = 5$	0.0084	0.0037	2.7974e-04	0.1865	4.7418
	$k = 6$	1.7980e-04	3.6944e-05	3.0876e-07	0.0064	2.5236
	$k = 7$	8.7850e-08	3.7522e-09	3.9399e-13	2.3145e-05	0.6067
	$k = 8$	3.4460e-14	2.9348e-14	0	7.1417e-10	0.0652
	$k = 9$	0	0	0	2.5806e-14	0.0018
	$k = 10$	0	0	0	0	4.8512e-06
	$k = 11$	0	0	0	0	4.0442e-11
$\theta = 0.5$	$k = 1$	7.3739	7.1787	7.4973	36.6905	52.5572
	$k = 2$	1.4149	1.3341	1.9771	9.7323	47.0581
	$k = 3$	0.2936	0.1584	0.2898	3.4634	29.1759
	$k = 4$	0.0510	0.0113	0.0139	1.6859	5.9672
	$k = 5$	0.0050	2.5241e-04	2.8185e-04	0.3429	3.2019
	$k = 6$	6.5933e-05	2.4935e-07	3.1004e-07	0.0293	1.3536
	$k = 7$	1.1888e-08	2.5763e-13	3.9592e-13	8.8565e-04	0.4265
	$k = 8$	2.6622e-14	0	0	1.3332e-06	0.0662
	$k = 9$	0	0	0	3.0416e-12	0.0052
	$k = 10$	0	0	0	0	4.5166e-05
	$k = 11$	0	0	0	0	3.4861e-09
	$k = 12$	0	0	0	0	2.6096e-14
$\theta = 1$	$k = 1$	7.3626	7.0494	7.8009	18.3519	31.2565
	$k = 2$	1.4098	1.2873	2.3134	3.0913	7.0409
	$k = 3$	0.2663	0.1559	0.4035	0.8192	3.3569
	$k = 4$	0.0429	0.0111	0.0227	0.0626	1.5570
	$k = 5$	0.0037	2.4747e-04	3.0924e-04	8.2097e-04	0.4177
	$k = 6$	3.7527e-05	2.3615e-07	3.1045e-07	4.0869e-07	0.0577
	$k = 7$	3.8708e-09	2.2738e-13	3.9534e-13	1.9931e-13	0.0040
	$k = 8$	2.5699e-14	0	0	0	2.7297e-05
	$k = 9$	0	0	0	0	1.2737e-09
	$k = 10$	0	0	0	0	3.1010e-14

as the stopping criterion where *iter* denotes the number of iterations. Numerical results are listed in Table 2 where **AIT** and **ACPU** denote the average number of iterations and the average CPU time in seconds respectively, and $*$ stands for that the algorithm fails to solve some instances as the iteration number is greater than 30 and the average is based on the successful instances through our numerical report.

From Table 2, we may see that Algorithm 2.1 with $\theta = 0$ always outperforms

Table 2. Comparison of Algorithm 2.1 with different values of θ

μ_0		$\theta = 0$		$\theta = 0.5$		$\theta = 1$		Fan-LM	
	n	**AIT**	**ACPU**	**AIT**	**ACPU**	**AIT**	**ACPU**	**AIT**	**ACPU**
10^{-4}	100	6.8	0.04	6.6	0.02	6.6	0.02	6.8	0.02
	300	7.2	0.22	7.0	0.24	7.0	0.24	7.2	0.22
	500	7.2	0.78	7.0	0.67	7.0	0.69	7.4	0.70
	700	7.0*	1.42	7.0	1.39	7.0	1.46	7.0	1.51
	900	7.0	2.52	7.0	2.54	7.0*	2.97	7.5*	2.75
	1100	7.4	4.53	7.2	4.68	8.8*	5.11	7.4	4.25
	1300	7.2	6.64	8.4	7.89	10.2	9.26	7.2	6.24
	1500	7.8	10.04	7.4	9.88	10.3*	14.11	7.7*	10.83
10^{-2}	100	6.4	0.02	6.4	0.02	6.4	0.02	6.6	0.03
	300	6.8	0.21	6.6	0.19	7.2	0.21	6.6	0.17
	500	7.0	0.81	7.0	0.77	7.6	0.82	7.0	0.70
	700	7.0	1.50	7.8	1.73	8.6	1.77	7.0	1.51
	900	8.0	3.10	8.6	3.27	8.4	3.26	7.4	2.70
	1100	7.2	4.46	8.4	5.17	8.2*	5.37	8.2	4.81
	1300	8.0	7.51	9.4	8.86	8.0*	7.47	8.0*	7.98
	1500	7.6	10.20	9.2*	12.32	8.6	11.49	8.2	10.64

or at least performs as well as it with $\theta = 0.5$ or $\theta = 1$ in most cases. Moreover, we may observe that Algorithm 2.1 taking $\theta = 0$, i.e., $\lambda_k = \mu_k\|F_k\|$, performs better than Fan's LM method [9] which also took $\lambda_k = \mu_k\|F_k\|$. This indicates that the nonmonotone trust region technique introduced in this paper improves the numerical performance of the LM method. We have tested Algorithm 2.1 with other values of δ and the computation effect is similar.

5. Conclusions

In this paper we have improved the Levenberg-Marquardt method by taking a general LM parameter and adopting a nonmonotone trust region technique. We have proved that the new LM method has global convergence and its convergence order is $\min\{1+\delta, 4-\delta, 2\}$ where $\delta \in (0,3)$ under the local error bound condition. We have also applied the new LM method to solve the nonlinear equations arising from the weighted linear complementarity problem where the associated mapping satisfies the Lipschitz continuity of the Jacobian. The numerical results showed that the new LM method is efficient and promising.

References

[1] M. Ahookhosh and K. Amini, *An efficient nonmonotone trust-region method for unconstrained optimization*, Numer. Algor., 2012, 59(4), 523–540.

[2] K. Amini, F. Rostami and G. Caristi, *An efficient Levenberg-Marquardt method with a new LM parameter for systems of nonlinear equations*, Optim., 2018, 67(5), 637–650.

[3] S. Asadi, Z. Darvay and G. Lesaja, et al., *A full-Newton step interior-point method for monotone weighted linear complementarity problems*, J. Optim. Theory Appl., 2020, 186(3), 864–878.

[4] R. Behling and A. Iusem, *The effect of calmness on the solution set of systems of nonlinear equations*, Math. Program., 2013, 137(1–2), 155–165.

[5] J. S. Chen, *The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem*, J. Global Optim., 2006, 36(4), 565–580.

[6] X. N. Chi, Z. P. Wan and Z. J. Hao, *A full-modified-Newton step $O(n)$ infeasible interior-point method for the special weighted linear complementarity problem*, J. Ind. Manag. Optim., 2022, 18(4), 2579–2598.

[7] X. N. Chi and G. Q. Wang, *A full-Newton step infeasible interior-point method for the special weighted linear complementarity problem*, J. Optim. Theory Appl., 2021, 190(1), 108–129.

[8] N. Y. Deng, Y. Xiao and F. J. Zhou, *Nonmonotonic trust region algorithm*, J. Optim. Theory Appl., 1993, 76(2), 259–285.

[9] J. Y. Fan, *A modified Levenberg-Marquardt algorithm for singular system of nonlinear equations*, J. Comput. Math., 2003, 21(5), 625–636.

[10] J. Y. Fan and J. Y. Pan, *Convergence properties of a self-adaptive Levenberg-Marquardt algorithm under local error bound condition*, Comput. Optim. Appl., 2006, 34(1), 47–62.

[11] J. Y. Fan and J. Y. Pan, *A note on the Levenberg-Marquardt parameter*, Appl. Math. Comput., 2009, 207(2), 351–359.

[12] J. Y. Fan and Y. X. Yuan, *On the convergence of a new Levenberg-Marquardt method*, AMSS, Chinese Academy of Sciences, 2001, Report NO. 005, 1–11.

[13] J. Y. Fan and Y. X. Yuan, *On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption*, Comput., 2005, 74(1), 23–39.

[14] M. S. Gowda, *Weighted LCPs and interior point systems for copositive linear transformations on Euclidean Jordan algebras*, J. Global Optim., 2019, 74(2), 285–295.

[15] C. F. Ma and L. H. Jiang, *Some research on Levenberg-Marquardt method for the nonlinear equations*, Appl. Math. Comput., 2007, 184(2), 1032–1040.

[16] F. A. Potra, *Weighted complementarity problems-a new paradigm for computing equilibria*, SIAM J. Optim., 2012, 22(4), 1634–1654.

[17] F. A. Potra, *Sufficient weighted complementarity problems*, Comput. Optim. Appl., 2016, 64(2), 467–488.

[18] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*, Academic Press, Inc., Boston, MA, San Diego, CA, 1990.

[19] J. Y. Tang, *A variant nonmonotone smoothing algorithm with improved numerical results for large-scale LWCPs*, Comput. Appl. Math., 2018, 37(3), 3927–3936.

[20] J. Y. Tang and H. C. Zhang, *A nonmonotone smoothing Newton algorithm for weighted complementarity problems*, J. Optim. Theory Appl., 2021, 189(3), 679–715.

[21] J. Y. Tang and J. C. Zhou, *A modified damped Gauss-Newton method for non-monotone weighted linear complementarity problems*, Optim. Methods Softw., 2022, 37(3), 1145–1164.

[22] J. Y. Tang and J. C. Zhou, *Quadratic convergence analysis of a nonmonotone Levenberg-Marquardt type method for the weighted nonlinear complementarity problem*, Comput. Optim. Appl., 2021, 80(1), 213–244.

[23] Ph. L. Toint, *Non-monotone trust-region algorithm for nonlinear optimization subject to convex constraints*, Math. Program., 1997, 77(1), 69–94.

[24] H. Y. Wang and J. Y. Fan, *Convergence rate of the Levenberg-Marquardt method under Hölderian local error bound*, Optim. Methods Softw., 2020, 35(4), 767–786.

[25] H. Y. Wang and J. Y. Fan, *Convergence properties of inexact Levenberg-Marquardt method under Hölderian local error bound*, J. Ind. Manag. Optim., 2021, 17(4), 2265–2275.

[26] N. Yamashita and M. Fukushima, *On the rate of convergence of the Levenberg-Marquardt method*, Comput. Suppl., 2001, 15, 239–249.

[27] Y. X. Yuan, *Recent advances in trust region algorithms*, Math. Program., 2015, 151(1), 249–281.

[28] M. L. Zeng and G. H. Zhou, *Improved convergence results of an efficient Levenberg-Marquardt method for nonlinear equations,* J. Appl. Math. Comput., 2022, 68(6), 3655–3671.

[29] J. Zhang, *A smoothing Newton algorithm for weighted linear complementarity problem*, Optim. Lett., 2016, 10(3), 499–509.

[30] L. Zheng, L. Chen and Y. F. Ma, *A variant of the Levenberg-Marquardt method with adaptive parameters for systems of nonlinear equations*, AIMS Math., 2022, 7(1), 1241–1256.

[31] W. J. Zhou, *On the convergence of the modified Levenberg-Marquardt method with a nonmonotone second order Armijo type line search*, J. Comput. Appl. Math., 2013, 239, 152–161.