Mathematics

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

AIMS Mathematics, 8(8): 18649–18664.
DOI: 10.3934/math.2023950
Received: 17 April 2023
Revised: 13 May 2023
Accepted: 18 May 2023
Published: 02 June 2023

Research article

Convergence properties of a family of inexact Levenberg-Marquardt methods

Luyao Zhao and Jingyong Tang*

College of Mathematics and Statistics, Xinyang Normal University, Xinyang 464000, China

* **Correspondence:** Email: tangjy@xynu.edu.cn.

Abstract: We present a family of inexact Levenberg-Marquardt (LM) methods for the nonlinear equations which takes more general LM parameters and perturbation vectors. We derive an explicit formula of the convergence order of these inexact LM methods under the Höderian local error bound condition and the Höderian continuity of the Jacobian. Moreover, we develop a family of inexact LM methods with a nonmonotone line search and prove that it is globally convergent. Numerical results for solving the linear complementarity problem are reported.

Keywords: nonlinear equations; inexact Levenberg-Marquardt method; global convergence; convergence rate; Höderian local error bound
Mathematics Subject Classification: 90C33, 65K05

1. Introduction

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. In the paper, we assume that the solution set of (1.1) is nonempty and denote it by X^*. Moreover, we denote the Jacobian $F'(x)$ as $J(x)$ and use the notations $F_k = F(x_k)$, $J_k = J(x_k)$ for simplification.

The Levenberg-Marquardt (LM) method is one of the most important algorithms for solving (1.1). At every iteration, the LM method computes the trial step d_k by solving the following linear system

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

where μ_k is the LM parameter which plays an important role in analyzing the convergence rate of the LM method. For example, Yamashita and Fukushima [13] proved that the LM method taking

$\mu_k = \|F_k\|^2$ has quadratic convergence under the local error bound condition which is weaker than the nonsingularity. Fan and Yuan [6] proved that the LM method taking $\mu_k = \|F_k\|^\delta$ with $\delta \in [1, 2]$ still achieves the quadratic convergence under the local error bound condition. More researches on the LM method can be found in [1, 11, 14–16] and references therein.

The LM method solves the linear system (1.2) exactly at every iteration which may be very expensive when solving a large-scale nonlinear equation. The inexact approach is one way to overcome this difficulty. In the inexact LM method, the direction d_k is given by the solution of the system

$$(J_k^T J_k + \mu_k I)d_k = -J_k^T F_k + p_k, \tag{1.3}$$

where $p_k \in \mathbb{R}^n$ is a perturbation vector which measures how inexactly the linear system (1.2) is solved. Under the nonsingularity, Facchinei and Kanzow [3] proved that if $\mu_k \to 0$ and $\|p_k\| \le o(\|J_k^T F_k\|)$, then the inexact LM method has superlinear convergence rate and if $\mu_k = O(\|J_k^T F_k\|)$ and $\|p_k\| = O(\|J_k^T F_k\|^2)$, then its convergence rate is quadratic. Suppose

$$\mu_k = O(\|F_k\|^\alpha) \ \text{ and } \ \|p_k\| = O(\|F_k\|^{\alpha+\theta}),$$

where $\alpha > 0$ and $\theta > 0$ are constants. Under the local error bound condition, many researchers (e.g., [2, 4, 5, 7]) investigated the convergence rate of the inexact LM method for different values of α and θ respectively. Lately, Wang and Fan [12] studied the convergence rate of the inexact LM method taking $\mu_k = \|F_k\|^\alpha$ with $\|p_k\| = \|F_k\|^{\alpha+\theta}$ and $\mu_k = \|J_k^T F_k\|^\alpha$ with $\|p_k\| = \|J_k^T F_k\|^{\alpha+\theta}$ respectively under the Höderian local error bound condition and the Höderian continuity of the Jacobian, which are more general than the local error bound condition and the Lipschitz continuity of the Jacobian used in [1, 2, 4, 5, 7].

In this paper, we study the convergence rate of a family of inexact LM methods with more general LM parameters and perturbation vectors. We consider

$$\mu_k = \sigma\|F_k\|^\alpha + (1-\sigma)\|J_k^T F_k\|^\alpha, \tag{1.4}$$

$$\|p_k\| = \tau\|F_k\|^{\alpha+\theta} + (1-\tau)\|J_k^T F_k\|^{\alpha+\theta}, \tag{1.5}$$

where $\sigma, \tau \in [0, 1]$ and $\alpha, \theta > 0$. We derive an explicit formula of the convergence order under the Höderian local error bound condition and the Höderian continuity of the Jacobian. Moreover, we develop a family of inexact LM methods with a nonmonotone line search and prove its global convergence. We also investigate the numerical performances of these inexact LM methods by solving the nonlinear equations arising in the linear complementarity problem.

The organization of this paper is as follows. In the next section, we investigate the convergence order under the Höderian local error bound condition and the Höderian continuity of the Jacobian. In Section 3, we present a family of inexact LM methods with a nonmonotone line search and prove that it is globally convergent. In Section 4, we apply these inexact LM methods to solve the nonlinear equations arising from the linear complementarity problem and report some numerical results.

2. Convergence rate of the inexact LM methods

In this section, we study the convergence rate of the inexact LM methods with the iteration

$$x_{k+1} = x_k + d_k, \tag{2.1}$$

where d_k is obtained by (1.3) with μ_k, p_k satisfying (1.4) and (1.5). We suppose that the generated sequence $\{x_k\}$ converges to the solution set X^* and lies in some neighbourhoods of $x^* \in X^*$.

Assumption 2.1. (a) $F(x)$ provides a Höderian local error bound of order $\gamma \in (0,1]$ in some neighbourhoods of $x^* \in X^*$, i.e., there exist constants $\kappa > 0$ and $0 < r < 1$ such that

$$\kappa \text{dist}(x, X^*) \leq \|F(x)\|^\gamma, \quad \forall x \in N(x^*, r) = \{x \in \mathbb{R}^n | \, \|x - x^*\| \leq r\}. \tag{2.2}$$

(b) $J(x)$ is Höderian continuous of order $\upsilon \in (0,1]$, i.e., there exists a constant $L > 0$ such that

$$\|J(x) - J(y)\| \leq L\|x - y\|^\upsilon, \quad \forall x, y \in N(x^*, r). \tag{2.3}$$

It is worth pointing out that if $\gamma = \upsilon = 1$, then Assumption 2.1 (a) is the local error bound condition and Assumption 2.1 (b) is the Lipschitz continuity of the Jacobian. Moreover, by (2.3), we have (see [12])

$$\|F(y) - F(x) - J(x)(y - x)\| \leq \frac{L}{1+\upsilon}\|y - x\|^{1+\upsilon}, \quad \forall x, y \in N(x^*, r). \tag{2.4}$$

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

$$\|F(y) - F(x)\| \leq M\|y - x\|, \quad \forall x, y \in N(x^*, r). \tag{2.5}$$

In the following, we denote by $\overline{x}_k$ the vector in X^* that is closest to x_k, i.e.,

$$\|\overline{x}_k - x_k\| = \text{dist}(x_k, X^*). \tag{2.6}$$

Now we suppose the singular value decomposition (SVD) of $J(\bar{x}_k)$ is

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

where $\bar{\Sigma}_{k,1} = \text{diag}(\bar{\sigma}_{k,1}, ..., \bar{\sigma}_{k,r})$ with $\bar{\sigma}_{k,1} \geq \cdots \geq \bar{\sigma}_{k,r} > 0$. The corresponding SVD of J_k is

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

where $\Sigma_{k,1} = \text{diag}(\sigma_{k,1}, ..., \sigma_{k,r})$ with $\sigma_{k,1} \geq \cdots \geq \sigma_{k,r} > 0$ and $\Sigma_{k,2} = \text{diag}(\sigma_{k,r+1}, ..., \sigma_{k,n})$ with $\sigma_{k,r+1} \geq \cdots \geq \sigma_{k,n} \geq 0$. For simplicity, we neglect the subscript k in $U_{k,i}, \Sigma_{k,i}, V_{k,i} (i = 1, 2)$ and write $J(\bar{x}_k)$ and J_k as

$$J(\bar{x}_k) = \bar{U}_1\bar{\Sigma}_1\bar{V}_1^T, \quad J_k = U_1\Sigma_1V_1^T + U_2\Sigma_2V_2^T.$$

By the matrix perturbation theory [8] and (2.3), we have

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

which gives

$$\|\Sigma_1 - \bar{\Sigma}_1\| \leq L\|\bar{x}_k - x_k\|^\upsilon, \quad \|\Sigma_2\| \leq L\|\bar{x}_k - x_k\|^\upsilon. \tag{2.7}$$

Moreover, by (1.3) we have

$$\begin{aligned} d_k = &- (J_k^T J_k + \mu_k I)^{-1} J_k^T F_k + (J_k^T J_k + \mu_k I)^{-1} p_k \\ = &- V_1(\Sigma_1^2 + \mu_k I)^{-1}\Sigma_1 U_1^T F_k - V_2(\Sigma_2^2 + \mu_k I)^{-1}\Sigma_2 U_2^T F_k \\ &+ V_1(\Sigma_1^2 + \mu_k I)^{-1} V_1^T p_k + V_2(\Sigma_2^2 + \mu_k I)^{-1} V_2^T p_k, \end{aligned} \tag{2.8}$$

and

$$\begin{aligned} F_k + J_k d_k = &F_k - J_k(J_k^T J_k + \mu_k I)^{-1} J_k^T F_k + J_k(J_k^T J_k + \mu_k I)^{-1} p_k \\ = &\mu_k U_1(\Sigma_1^2 + \mu_k I)^{-1} U_1^T F_k + \mu_k U_2(\Sigma_2^2 + \mu_k I)^{-1} U_2^T F_k \\ &+ U_1\Sigma_1(\Sigma_1^2 + \mu_k I)^{-1} V_1^T p_k + U_2\Sigma_2(\Sigma_2^2 + \mu_k I)^{-1} V_2^T p_k. \end{aligned} \tag{2.9}$$

In the following, we suppose without loss of generality that x_k lies in $N(x^*, \frac{r}{2})$.

Lemma 2.1. *Under the conditions of Assumption* 2.1*, if* $\upsilon > \frac{2}{\gamma} - 2$*, then there exist positive constants* a_1, a_2, b_1, b_2 *such that*

$$a_1\|\bar{x}_k - x_k\|^{(\frac{2}{\gamma}-1)\alpha} \leq \mu_k \leq a_2\|\bar{x}_k - x_k\|^{\alpha}, \tag{2.10}$$

$$b_1\|\bar{x}_k - x_k\|^{(\frac{2}{\gamma}-1)(\alpha+\theta)} \leq \|p_k\| \leq b_2\|\bar{x}_k - x_k\|^{\alpha+\theta}. \tag{2.11}$$

Proof. Since $\|\bar{x}_k - x^*\| \leq \|\bar{x}_k - x_k\| + \|x_k - x^*\| \leq 2\|x_k - x^*\| \leq r$, we have $\bar{x}_k \in N(x^*, r)$. Hence, by (2.2) and (2.5),

$$\kappa^{\frac{1}{\gamma}}\|\bar{x}_k - x_k\|^{\frac{1}{\gamma}} \leq \|F_k\| \leq M\|\bar{x}_k - x_k\|. \tag{2.12}$$

By (2.5), we have

$$\|J_k^T F_k\| \leq \|J_k\|\|F_k - F(\bar{x}_k)\| \leq M^2\|\bar{x}_k - x_k\|. \tag{2.13}$$

Let $T_k := F_k - F(\bar{x}_k) - J_k(x_k - \bar{x}_k)$. Then,

$$F_k^T J_k(x_k - \bar{x}_k) = \|F_k\|^2 - F_k^T T_k. \tag{2.14}$$

It follows from (2.4), (2.12) and (2.14) that

$$\begin{aligned} \|F_k^T J_k\|\|\bar{x}_k - x_k\| &\geq \|F_k\|^2 - F_k^T T_k \\ &\geq \|F_k\|^2 - \|F_k\|\|F_k - F(\bar{x}_k) - J_k(x_k - \bar{x}_k)\| \\ &\geq \kappa^{\frac{2}{\gamma}}\|\bar{x}_k - x_k\|^{\frac{2}{\gamma}} - \frac{LM}{1+\upsilon}\|\bar{x}_k - x_k\|^{2+\upsilon}, \end{aligned}$$

which together with $\upsilon > \frac{2}{\gamma} - 2$ gives

$$\begin{aligned} \|F_k^T J_k\| &\geq \kappa^{\frac{2}{\gamma}}\|\bar{x}_k - x_k\|^{\frac{2}{\gamma}-1} - \frac{LM}{1+\upsilon}\|\bar{x}_k - x_k\|^{1+\upsilon} \\ &\geq (\kappa^{\frac{2}{\gamma}} - \frac{LM}{1+\upsilon}\|\bar{x}_k - x_k\|^{2+\upsilon-\frac{2}{\gamma}})\|\bar{x}_k - x_k\|^{\frac{2}{\gamma}-1} \\ &\geq \tilde{c}\|\bar{x}_k - x_k\|^{\frac{2}{\gamma}-1}, \end{aligned} \tag{2.15}$$

where $\tilde{c} > 0$ is some constant. By (2.13) and (2.15), we have

$$\tilde{c}\|\bar{x}_k - x_k\|^{\frac{2}{\gamma}-1} \leq \|J_k^T F_k\| \leq M^2\|\bar{x}_k - x_k\|. \tag{2.16}$$

Since $\mu_k = \sigma\|F_k\|^\alpha + (1-\sigma)\|J_k^T F_k\|^\alpha$, by (2.12) and (2.16), we have

$$a_1\|\overline{x}_k - x_k\|^{\max\{\frac{\alpha}{\gamma},(\frac{2}{\gamma}-1)\alpha\}} \le \mu_k \le a_2\|\overline{x}_k - x_k\|^\alpha,$$

where $a_1 := \sigma\kappa^{\frac{\alpha}{\gamma}} + (1-\sigma)\tilde{c}^\alpha$ and $a_2 := \sigma M^\alpha + (1-\sigma)M^{2\alpha}$, which together with $\frac{2}{\gamma} - 1 \ge \frac{1}{\gamma}$ gives

$$a_1\|\overline{x}_k - x_k\|^{(\frac{2}{\gamma}-1)\alpha} \le \mu_k \le a_2\|\overline{x}_k - x_k\|^\alpha.$$

This proves (2.10). Moreover, since $\|p_k\| = \tau\|F_k\|^{\alpha+\theta} + (1-\tau)\|J_k^T F_k\|^{\alpha+\theta}$, according to (2.12) and (2.16), we have

$$b_1\|\overline{x}_k - x_k\|^{\max\{\frac{\alpha+\theta}{\gamma},(\frac{2}{\gamma}-1)(\alpha+\theta)\}} \le \|p_k\| \le b_2\|\overline{x}_k - x_k\|^{\alpha+\theta},$$

where $b_1 := \tau\kappa^{\frac{\alpha+\theta}{\gamma}} + (1-\tau)\tilde{c}^{\alpha+\theta}$ and $b_2 := \tau M^{\alpha+\theta} + (1-\tau)M^{2(\alpha+\theta)}$, which together with $\frac{2}{\gamma} - 1 \ge \frac{1}{\gamma}$ gives

$$b_1\|\overline{x}_k - x_k\|^{(\frac{2}{\gamma}-1)(\alpha+\theta)} \le \|p_k\| \le b_2\|\overline{x}_k - x_k\|^{\alpha+\theta}.$$

This proves (2.11). □ □

Lemma 2.2. *Under the conditions of Assumption* 2.1*, if* $\upsilon > \frac{2}{\gamma} - 2$*,* $0 < \alpha < \frac{2\gamma(1+\upsilon)}{2-\gamma}$ *and* $\theta > \frac{(2-2\gamma)\alpha}{\gamma}$*, there exists a constant* $c > 0$ *such that*

$$\|d_k\| \le c\|\overline{x}_k - x_k\|^{\min\{1, 1+\upsilon-\frac{(2-\gamma)\alpha}{2\gamma}, \frac{(2\gamma-2)\alpha}{\gamma}+\theta\}}. \tag{2.17}$$

Proof. Let

$$\bar{d}_k := -(J_k^T J_k + \mu_k I)^{-1} J_k^T F_k. \tag{2.18}$$

Then $\bar{d}_k$ is the LM step computed by solving (1.2). Moreover, by (1.3) we have

$$d_k = \bar{d}_k + (J_k^T J_k + \mu_k I)^{-1} p_k. \tag{2.19}$$

Now we define

$$\varphi_k(d) := \|F_k + J_k d\|^2 + \mu_k\|d\|^2. \tag{2.20}$$

Then, the LM step $\overline{d}_k$ defined by (2.18) is the minimizer of $\varphi_k(d)$. By (2.4) and the left inequality in (2.10), we have

$$\begin{aligned}
\|\overline{d}_k\|^2 \le \frac{\varphi_k(\overline{d}_k)}{\mu_k} &\le \frac{\varphi_k(\overline{x}_k - x_k)}{\mu_k} = \frac{\|F_k + J_k(\overline{x}_k - x_k)\|^2}{\mu_k} + \|\overline{x}_k - x_k\|^2 \\
&\le \left(\frac{L}{1+\upsilon}\right)^2 a_1^{-1}\|\overline{x}_k - x_k\|^{2+2\upsilon-(\frac{2}{\gamma}-1)\alpha} + \|\overline{x}_k - x_k\|^2 \\
&\le c_1\|\overline{x}_k - x_k\|^{2\min\{1+\upsilon-\frac{\alpha}{\gamma}+\frac{\alpha}{2}, 1\}},
\end{aligned} \tag{2.21}$$

where $c_1 := ((L/(1+\upsilon))^2 a_1^{-1} + 1)$. Thus, by (2.19), (2.21) and the left inequality in (2.10) and the right inequality in (2.11), we have

$$\begin{aligned}
\|d_k\| &\le \|\overline{d}_k\| + \|d_k - \overline{d}_k\| \\
&= \|\overline{d}_k\| + \|(J_k^T J_k + \mu_k I)^{-1} p_k\|
\end{aligned}$$

$$\leq \|\bar{d}_k\| + \frac{\|p_k\|}{\mu_k}$$
$$\leq c_1\|\bar{x}_k - x_k\|^{\min\{1+\upsilon-\frac{\alpha}{\gamma}+\frac{\alpha}{2},1\}} + \frac{b_2}{a_1}\|\bar{x}_k - x_k\|^{\alpha+\theta-(\frac{2}{\gamma}-1)\alpha}$$
$$\leq c\|\bar{x}_k - x_k\|^{\min\{1,1+\upsilon-\frac{(2-\gamma)\alpha}{2\gamma},\frac{(2\gamma-2)\alpha}{\gamma}+\theta\}}, \tag{2.22}$$

where $c = \sqrt{c_1} + b_2/a_1$. □ □

Lemma 2.3. [12, Lemma 2.3] *Under the conditions of Assumption* 2.1*, we have*
(i)$\|U_1U_1^TF_k\| \leq M\|\bar{x}_k - x_k\|$; (ii)$\|U_2U_2^TF_k\| \leq 2L\|\bar{x}_k - x_k\|^{1+\upsilon}$.

Theorem 2.1. *Under the conditions of Assumption* 2.1*, if* $\upsilon > \frac{2}{\gamma} - 2$*,* $0 < \alpha < \frac{2\gamma(1+\upsilon)}{2-\gamma}$ *and* $\theta > \frac{(2-2\gamma)\alpha}{\gamma}$*, then the sequence* $\{x_k\}$ *converges to the solution set* X^* *with the order* $h(\alpha,\theta,\gamma,\upsilon)$ *where*

$$h(\alpha,\theta,\gamma,\upsilon) = \gamma\min\Big\{1+\alpha, 1+\upsilon, \alpha+\theta, \frac{(2\gamma-2)\alpha}{\gamma}+\theta+\upsilon,$$
$$(1+\upsilon)\Big(1+\upsilon-\frac{(2-\gamma)\alpha}{2\gamma}\Big), (1+\upsilon)\Big(\frac{(2\gamma-2)\alpha}{\gamma}+\theta\Big)\Big\}. \tag{2.23}$$

Proof. Since x_k converges to X^*, we assume that $L\|\bar{x}_k - x_k\|^\nu \leq \frac{\bar{\sigma}}{2}$ holds for all sufficiently large k. Then, it follows from (2.7) that

$$\|(\Sigma_1^2 + \mu_k I)^{-1}\| \leq \|\Sigma_1^{-2}\| \leq \frac{1}{(\bar{\sigma} - L\|\bar{x}_k - x_k\|^\nu)^2} < \frac{4}{\bar{\sigma}^2}. \tag{2.24}$$

On the other hand, by (2.7) and the left inequality in (2.10), for all sufficiently large k,

$$\|\Sigma_2(\Sigma_2^2 + \mu_k I)^{-1}\| \leq \frac{\|\Sigma_2\|}{\mu_k} \leq La_1^{-1}\|\bar{x}_k - x_k\|^{\upsilon-(\frac{2}{\gamma}-1)\alpha}. \tag{2.25}$$

Hence, it follows from (2.9)–(2.11), (2.24), (2.25), $\|(\Sigma_2^2 + \mu_k I)^{-1}\| \leq \mu_k^{-1}$ and Lemma 2.3 that

$$\|F_k + J_kd_k\| \leq \|\mu_kU_1(\Sigma_1^2+\mu_kI)^{-1}U_1^TF_k\| + \|\mu_kU_2(\Sigma_2^2+\mu_kI)^{-1}U_2^TF_k\|$$
$$+ \|U_1\Sigma_1(\Sigma_1^2+\mu_kI)^{-1}V_1^Tp_k\| + \|U_2\Sigma_2(\Sigma_2^2+\mu_kI)^{-1}V_2^Tp_k\|$$
$$\leq \frac{4Ma_2}{\bar{\sigma}^2}\|\bar{x}_k - x_k\|^{1+\alpha} + 2L\|\bar{x}_k - x_k\|^{1+\upsilon}$$
$$+ \frac{2}{\bar{\sigma}}b_2\|\bar{x}_k - x_k\|^{\alpha+\theta} + La_1^{-1}b_2\|\bar{x}_k - x_k\|^{\frac{(2\gamma-2)\alpha}{\gamma}+\theta+\upsilon}$$
$$\leq \bar{c}\|\bar{x}_k - x_k\|^{\min\{1+\alpha,1+\upsilon,\alpha+\theta,\frac{(2\gamma-2)\alpha}{\gamma}+\theta+\upsilon\}}, \tag{2.26}$$

where $\bar{c} = 4Ma_2/\bar{\sigma}^2 + 2L + \frac{2}{\bar{\sigma}}b_2 + La_1^{-1}b_2$. Therefore, by (2.2), (2.4), (2.26) and Lemma 2.2, we have

$$\|\bar{x}_{k+1} - x_{k+1}\| \leq \frac{1}{\kappa}\|F_{k+1}\|^\gamma$$
$$\leq \frac{1}{\kappa}(\|F_k + J_kd_k\| + \frac{L}{1+\upsilon}\|d_k\|^{1+\upsilon})^\gamma$$
$$\leq \frac{1}{\kappa}\Big(\bar{c}\|\bar{x}_k - x_k\|^{\min\{1+\alpha,1+\upsilon,\alpha+\theta,\frac{(2\gamma-2)\alpha}{\gamma}+\theta+\upsilon\}}$$

$$+\frac{L}{1+\upsilon}c^{1+\upsilon}\|\overline{x}_k - x_k\|^{\min\{1+\upsilon,(1+\upsilon)(1+\upsilon-\frac{(2-\gamma)\alpha}{2\gamma}),(1+\upsilon)(\frac{(2\gamma-2)\alpha}{\gamma}+\theta)\}}\Big)^{\gamma}$$
$$\leq O(\|\overline{x}_k - x_k\|^{h(\alpha,\theta,\gamma,\upsilon)},$$

where $h(\alpha,\theta,\gamma,\upsilon)$ is given in (2.23). □ □

Corollary 2.1. *Under the conditions of Assumption* 2.1 *and* $\gamma = \upsilon = 1$*, if* $0 < \alpha < 4$*, then the sequence* $\{x_k\}$ *converges to the solution set* X^* *with the order* $h(\alpha,\theta)$ *where*

$$h(\alpha,\theta) = \begin{cases} \min\{\alpha+\theta, 4-\alpha, 2\theta\}, & \text{if } 0 < \theta < 1, \\ \min\{2, 1+\alpha, 4-\alpha\}, & \text{if } \theta \geq 1. \end{cases}$$

More precisely,

$$h(\alpha,\theta) = \begin{cases} \alpha+\theta & \text{if } \alpha \in (0,\theta], \\ 2\theta & \text{if } \alpha \in (\theta, 4-2\theta], \quad \text{if } \ 0 < \theta < 1, \\ 4-\alpha & \text{if } \alpha \in (4-2\theta, 4], \end{cases}$$

and

$$h(\alpha,\theta) = \begin{cases} 1+\alpha & \text{if } \alpha \in (0,1], \\ 2 & \text{if } \alpha \in (1,2], \quad \text{if } \ \theta \geq 1, \\ 4-\alpha & \text{if } \alpha \in (2,4]. \end{cases}$$

As we know from [5] that for any $\alpha \in (0,2]$ and $\theta \geq 1$, the sequence generated by the inexact LM method (1.3) converges to some solution of (1.1), which is a stronger result than the convergence to the solution set. We show that this convergence result also holds true for the inexact LM methods studied in this paper.

Theorem 2.2. *Under the conditions of Assumption* 2.1 *and* $\gamma = \upsilon = 1$*, if* $\alpha \in (0,2]$ *and* $\theta \geq 1$*, then the sequence* $\{x_k\}$ *converges to a solution of* (1.1) *with the order* $g(\alpha)$ *where*

$$g(\alpha) = \begin{cases} 1+\alpha & \text{if } \alpha \in (0,1], \\ 2 & \text{if } \alpha \in (1,2]. \end{cases} \tag{2.27}$$

Proof. By Corollary 2.1, when $\alpha \in (0,2]$ and $\theta \geq 1$, it holds that

$$\|\overline{x}_{k+1} - x_{k+1}\| \leq O(\|\overline{x}_k - x_k\|^{g(\alpha)}), \tag{2.28}$$

where $g(\alpha)$ is defined by (2.27). Then, for all sufficiently large k,

$$\|\bar{x}_k - x_k\| \leq \|x_k - \bar{x}_{k+1}\| \leq \|x_{k+1} - \bar{x}_{k+1}\| + \|d_k\| \leq O(\|\overline{x}_k - x_k\|^{g(\alpha)}) + \|d_k\|,$$

which together with $g(\alpha) > 1$ gives

$$\|\bar{x}_k - x_k\| \leq 2\|d_k\|. \tag{2.29}$$

Hence, we deduce from (2.17), (2.28) and (2.29) that

$$\begin{aligned} \|d_{k+1}\| &= O(\|\bar{x}_{k+1} - x_{k+1}\|^{\min\{1,2-\frac{\alpha}{2},\theta\}}) \\ &= O(\|\bar{x}_{k+1} - x_{k+1}\|) \\ &= O(\|\bar{x}^k - x_k\|^{g(\alpha)}) \\ &= O(\|d_k\|^{g(\alpha)}). \end{aligned}$$

This implies that the sequence $\{x_k\}$ converges to some solution of (1.1) with the order $g(\alpha)$. □ □

3. Globally convergent inexact LM methods

In this section, we study a family of globally convergent inexact LM methods. We define the merit function $\psi : \mathbb{R}^n \to \mathbb{R}$ as

$$\psi(x) := \frac{1}{2}\|F(x)\|^2. \tag{3.1}$$

Obviously, $\psi(x)$ is continuously differentiable at any $x \in \mathbb{R}^n$ with $\nabla\psi(x) = J(x)^T F(x)$. Our method is described as follows.

Algorithm 3.1. *Choose parameters* $\sigma, \tau \in [0, 1]$, $\alpha \in (0, 4]$, $\theta > 0$, $\rho, \xi, \chi, \delta, \zeta \in (0, 1)$ *and* $x_0 \in \mathbb{R}^n$. *Let* $\Theta_0 := \psi(x_0)$. *Set* $k := 0$.
Step 1: *If* $\|\nabla\psi(x_k)\| = 0$, *then stop.*
Step 2: *Set*

$$\mu_k := \sigma\|F_k\|^\alpha + (1 - \sigma)\|J_k^T F_k\|^\alpha, \tag{3.2}$$

$$w_k := \tau\|F_k\|^{\alpha+\theta} + (1 - \tau)\|J_k^T F_k\|^{\alpha+\theta}. \tag{3.3}$$

Find a search direction $d_k \in \mathbb{R}^n$ *which satisfies*

$$(J_k^T J_k + \mu_k I)d_k = -\nabla\psi(x_k) + p_k, \tag{3.4}$$

where

$$\|p_k\| \leq \min\{\rho\|\nabla\psi(x_k)\|, w_k\}. \tag{3.5}$$

If d_k *satisfies*

$$\|F(x_k + d_k)\| \leq \xi\|F(x_k)\|, \tag{3.6}$$

then set $\lambda_k := 1$ *and go to Step 4.*
Step 3: *If the descent condition*

$$\nabla\psi(x_k)^T d_k \leq -\chi\|d_k\|^2 \tag{3.7}$$

is not satisfied, then set $d_k := -\nabla\psi(x_k)$. *Let* l_k *be the smallest nonnegative integer* l *satisfying*

$$\psi(x_k + \delta^l d_k) \leq \Theta_k - \zeta\|\delta^l d_k\|^2. \tag{3.8}$$

Set $\lambda_k := \delta^{l_k}$ *and go to Step 4.*
Step 4: *Set* $x_{k+1} := x_k + \lambda_k d_k$ *and*

$$\Theta_{k+1} := \frac{(\Theta_k + 1)\psi(x_{k+1})}{\psi(x_{k+1}) + 1}. \tag{3.9}$$

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

Algorithm 3.1 is designed based on the inexact LM method [2] and the nonmonotone smoothing Newton method [10]. The main feature of Algorithm 3.1 is that it takes more general LM parameter μ_k and perturbation vector p_k and adopts a nonmonotone line search technique to ensure the global convergence.

Theorem 3.1. *Algorithm* 3.1 *generates an infinite sequence* $\{x_k\}$ *which satisfies* $\psi(x_k) \leq \Theta_k$ *for all* $k \geq 0$.

Proof. For some k, we assume that $\psi(x_k) \leq \Theta_k$. If $\nabla\psi(x_k) \neq 0$, then $F(x_k) \neq 0$ and hence $\mu_k > 0$. So, the matrix $J_k^T J_k + \mu_k I$ is positive definite and the search direction d_k in Step 2 is always obtained. Notice that the obtained $d_k \neq 0$. In fact, if $d_k = 0$, then by (3.4) we have $\|p_k - \nabla\psi(x_k)\| = 0$. Since $\|p_k\| \leq \rho\|\nabla\psi(x_k)\|$, it follows that $\|\nabla\psi(x_k)\| = \|p_k\| = 0$ which contradicts $\nabla\psi(x_k) \neq 0$. So, in Step 3, if the descent condition (3.7) holds, then $\nabla\psi(x_k)^T d_k \leq -\chi\|d_k\|^2 < 0$. Otherwise, $d_k = -\nabla\psi(x_k)$ which gives $\nabla\psi(x_k)^T d_k = -\|\nabla\psi(x_k)\|^2 < 0$. Thus, the direction d_k used in the line search (3.8) is always a descent direction of ψ. Next we show that there exists at least a nonnegative integer l satisfying (3.8). On the contrary, we suppose that for all nonnegative integer l,

$$\psi(x_k + \delta^l d_k) > \Theta_k - \zeta\|\delta^l d_k\|^2,$$

which together with $\psi(x_k) \leq \Theta_k$ gives

$$\frac{\psi(x_k + \delta^l d_k) - \psi(x_k)}{\delta^l} + \zeta\delta^l\|d_k\|^2 > 0. \tag{3.10}$$

By letting $l \to \infty$ in (3.10), we have $\nabla\psi(x_k)^T d_k \geq 0$ which contradicts $\nabla\psi(x_k)^T d_k < 0$. So, we can obtain x_{k+1} in Step 4. Now we show $\psi(x_{k+1}) \leq \Theta_{k+1}$. In fact, if the condition (3.6) holds, then

$$\psi(x_{k+1}) \leq \xi^2\psi(x_k) < \psi(x_k) \leq \Theta_k.$$

Otherwise, by Step 3, we also have $\psi(x_{k+1}) \leq \Theta_k$. Hence, from (3.9) it holds that

$$\Theta_{k+1} = \frac{(\Theta_k + 1)\psi(x_{k+1})}{\psi(x_{k+1}) + 1} \geq \frac{(\psi(x_{k+1}) + 1)\psi(x_{k+1})}{\psi(x_{k+1}) + 1} = \psi(x_{k+1}).$$

Therefore, we can conclude that if $\psi(x_k) \leq \Theta_k$ and $\nabla\psi(x_k) \neq 0$ for some k, then x_{k+1} can be generated by Algorithm 2.1 with $\psi(x_{k+1}) \leq \Theta_{k+1}$. Since $\psi(x_0) = \Theta_0$, by induction on k, we prove the theorem. □

Theorem 3.2. *Every accumulation point* x^* *of a sequence* $\{x_k\}$ *generated by Algorithm* 3.1 *is a stationary point of* $\psi(x)$, *i.e.,* $\nabla\psi(x^*) = 0$.

Proof. By Steps 2 and 3, we have $\psi(x_{k+1}) \leq \Theta_k$ for all $k \geq 0$. This and (3.9) yield that

$$\Theta_{k+1} = \frac{\Theta_k\psi(x_{k+1}) + \psi(x_{k+1})}{\psi(x_{k+1}) + 1} \leq \frac{\Theta_k\psi(x_{k+1}) + \Theta_k}{\psi(x_{k+1}) + 1} = \Theta_k.$$

Thus there exists $\Theta^* \geq 0$ such that $\lim_{k\to\infty} \Theta_k = \Theta^*$. Further, by (3.9) we have

$$\lim_{k\to\infty} \psi(x_k) = \lim_{k\to\infty} \frac{\Theta_k}{\Theta_{k-1} - \Theta_k + 1} = \Theta^*,$$

and so

$$\lim_{k\to\infty} \|F(x_k)\| = \sqrt{2\Theta^*}. \tag{3.11}$$

So, if there are infinitely many k for which $\|F(x_k + d_k)\| \leq \xi\|F(x_k)\|$ holds, then $\sqrt{2\Theta^*} \leq \xi\sqrt{2\Theta^*}$ which together with $\xi \in (0, 1)$ yields $\Theta^* = 0$, i.e., $\lim_{k\to\infty} F(x_k) = 0$. By the continuity, we have the desired result. Next, we assume that there exists an index $\bar{k}$ such that $\|F(x_k + d_k)\| > \xi\|F(x_k)\|$ for all $k \geq \bar{k}$, i.e., λ_k is determined by (3.8) for all $k \geq \bar{k}$. Since x^* is the accumulation point of $\{x_k\}$, there exists a subsequence $\{x_k\}_{k\in K}$ where $K \subset \{0, 1, ...\}$ such that $\lim_{(K\ni)k\to\infty} x_k = x^*$. We assume that $\nabla\psi(x^*) \neq 0$ and will derive a contradiction. Since $\nabla\psi(x^*) = J(x^*)^T F(x^*) \neq 0$, we have $\|J(x^*)\| > 0$ and $\|F(x^*)\| > 0$. Moreover, by the continuity, we have

$$\lim_{(K\ni)k\to\infty} \mu_k = \sigma\|F(x^*)\|^\alpha + (1-\sigma)\|J(x^*)^T F(x^*)\|^\alpha := \mu^*.$$

Obviously, $\mu^* > 0$. So, there exists a positive constant $\bar{\mu}$ such that $\mu_k \geq \bar{\mu} > 0$ for all $k \in K$. Since $\{\nabla\psi(x_k)\}$ is bounded on any convergent subsequence $\{x_k\}_{k\in K}$, for any $k \in K$, either

$$\|d_k\| \leq \|(J_k^T J_k + \mu_k I)^{-1}\|(\|\nabla\psi(x_k)\| + \|p_k\|) \leq \frac{(1+\rho)\|\nabla\psi(x_k)\|}{\bar{\mu}} < \infty,$$

or $\|d_k\| = \|-\nabla\psi(x_k)\| < \infty$. Hence, the sequence $\{\|d_k\|\}_{k\in K}$ is bounded. By passing to the subsequence, we suppose $\lim_{(K_1\ni)k\to\infty} d_k = d^*$ where $K_1 \subset K$ is an infinite subset. In the following, we prove $\nabla\psi(x^*)^T d^* = 0$. By (3.8) we have

$$\zeta\lambda_k^2\|d_k\|^2 \leq \Theta_k - \psi(x_{k+1}).$$

This and $\lim_{k\to\infty} \psi(x_k) = \lim_{k\to\infty} \Theta_k = \Theta^*$ yield $\lim_{k\to\infty} \lambda_k\|d_k\| = 0$. Hence, if $\lambda_k \geq \bar{\lambda} > 0$ for any $k \in K_1$ where $\bar{\lambda} > 0$ is a constant, then $\lim_{(K_1\ni)k\to\infty} d_k = d^* = 0$ and hence $\nabla\psi(x^*)^T d^* = 0$. Otherwise, $\{\lambda_k\}_{k\in K_1}$ has a subsequence converging to zero and we suppose $\lim_{(K_2\ni)k\to\infty} \lambda_k = 0$ where $K_2 \subset K_1$ is an infinite set. From (3.8), for all $k \geq \bar{k}$ and $k \in K_2$,

$$\psi(x_k + \delta^{-1}\lambda_k d_k) > \Theta_k - \zeta\|\delta^{-1}\lambda_k d_k\|^2 \geq \psi(x_k) - \zeta\|\delta^{-1}\lambda_k d_k\|^2,$$

which gives

$$\frac{\psi(x_k + \delta^{-1}\lambda_k d_k) - \psi(x_k)}{\delta^{-1}\lambda_k} > -\zeta\delta^{-1}\lambda_k\|d_k\|^2. \tag{3.12}$$

Since ψ is continuously differentiable at x^*, by letting $k \to \infty$ with $k \in K_2$ in (3.12), we have $\nabla\psi(x^*)^T d^* \geq 0$. On the other hand, since d_k is a sufficient descent direction of ψ, we have $\nabla\psi(x^*)^T d^* = \lim_{(K_2\ni)k\to\infty} \nabla\psi(x_k)^T d^k \leq 0$. These give $\nabla\psi(x^*)^T d^* = 0$. Hence, we can conclude that $\nabla\psi(x^*)^T d^* = 0$. Let $\bar{K} := \{k \in K_1 | d_k = -\nabla\psi(x_k)\}$. If $\bar{K}$ is an infinite set, then

$$\|\nabla\psi(x^*)\|^2 = \lim_{(\bar{K}\ni)k\to\infty} \|\nabla\psi(x_k)\|^2 = \lim_{(\bar{K}\ni)k\to\infty} -\nabla\psi(x_k)^T d_k = -\nabla\psi(x^*)^T d^* = 0,$$

which contradicts the assumption $\nabla\psi(x^*) \neq 0$. Otherwise, $\bar{K}$ is a finite set and d_k satisfies (3.7) for all sufficiently large $k \in K_1$. Then, by (3.7) we have

$$\chi\|d^*\|^2 = \lim_{(K_1\ni)k\to\infty} \chi\|d_k\|^2 \leq \lim_{(K_1\ni)k\to\infty} -\nabla\psi(x_k)^T d_k = -\nabla\psi(x^*)^T d^* = 0,$$

which gives $d^* = 0$. By (3.4), we have for all $k \in K_1$,

$$\|p_k - \nabla\psi(x_k)\| \le \|J_k^T J_k + \mu_k I\| \|d_k\|. \quad (3.13)$$

Since $\lim_{(K_1 \ni) k\to\infty} (J_k^T J_k + \mu_k I) = J(x^*)^T J(x^*) + \mu^* I$, by (3.13) and $d^* = 0$, we have

$$\lim_{(K_1 \ni) k\to\infty} \|p_k - \nabla\psi(x_k)\| = 0. \quad (3.14)$$

Since $\|p_k\| \le \rho \|\nabla\psi(x_k)\|$, we can deduce from (3.14) that

$$\|\nabla\psi(x^*)\| = \lim_{(K_1 \ni) k\to\infty} \|\nabla\psi(x_k)\| = \lim_{(K_1 \ni) k\to\infty} \|p_k\| = 0,$$

which also contradicts the assumption $\nabla\psi(x^*) \neq 0$. We complete the proof. □

Next, we analyze the convergence rete of Algorithm 3.1. Suppose that the generated iteration sequence $\{x_k\}$ has an accumulation point x^* such that $F(x^*) = 0$ and Assumption 2.1 holds at x^* for $\gamma = \upsilon = 1$. We will show that the whole sequence $\{x_k\}$ converges to x^* at leat superlinearly for any $\alpha \in (0, 2]$ and $\theta > 1$.

Lemma 3.1. *Assume that Assumption* 2.1 *holds for* $\gamma = \upsilon = 1$. *Let* $\alpha \in (0, 2]$ *and* $\theta > 1$. *If* $x_k, x_k + d_k \in N(x^*, r/2)$*, then there exists* $\hat{c} > 0$ *such that*

$$\mathrm{dist}(x_k + d_k, X^*) \le \hat{c}\,\mathrm{dist}(x_k, X^*)^{\min\{\frac{\alpha}{2}+1, \theta\}}. \quad (3.15)$$

Proof. Since $\bar{d}_k$ defined by (2.18) is the minimizer of $\varphi_k(d)$ in (2.20), by (2.4) and (2.10),

$$\begin{aligned} \varphi_k(\bar{d}_k) &\le \varphi_k(\bar{x}_k - x_k) \\ &= \|F_k + J_k(\bar{x}_k - x_k)\|^2 + \mu_k \|\bar{x}_k - x_k\|^2 \\ &\le L^2/4 \|\bar{x}_k - x_k\|^4 + a_2 \|\bar{x}_k - x_k\|^{\alpha+2} \\ &\le (L^2/4 + a_2) \|\bar{x}_k - x_k\|^{\alpha+2}. \end{aligned} \quad (3.16)$$

It holds from (2.20) and (3.16) that

$$\|F_k + J_k \bar{d}_k\| \le \sqrt{\varphi_k(\bar{d}_k)} \le \sqrt{L^2/4 + a_2}\, \|\bar{x}_k - x_k\|^{\frac{\alpha}{2}+1}. \quad (3.17)$$

Since $d_k = \bar{d}_k + (J_k^T J_k + \mu_k I)^{-1} p_k$, we have from (2.5), (2.10), (2.11) and (3.17) that

$$\begin{aligned} \|F_k + J_k d_k\| &= \|F_k + J_k \bar{d}_k + J_k (J_k^T J_k + \mu_k I)^{-1} p_k\| \\ &\le \|F_k + J_k \bar{d}_k\| + \frac{M\|p_k\|}{\mu_k} \\ &\le \sqrt{L^2/4 + a_2}\, \|\bar{x}_k - x_k\|^{\frac{\alpha}{2}+1} + \frac{M b_2}{a_1} \|\bar{x}_k - x_k\|^{\theta} \\ &\le \tilde{C} \|\bar{x}_k - x_k\|^{\min\{\frac{\alpha}{2}+1, \theta\}}, \end{aligned} \quad (3.18)$$

where $\tilde{C} = \sqrt{L^2/4 + a_2} + M b_2 a_1^{-1}$. Moreover, by (2.17), $\alpha \in (0, 2]$ and $\theta > 1$, we have

$$\|d_k\| \le c \|\bar{x}_k - x_k\|. \quad (3.19)$$

Thus, by (2.4), (2.17), (3.18) and (3.19), we have

$$\begin{aligned}\|F(x_k+d_k)\| &\le \|F_k+J_kd_k\|+L/2\|d_k\|^2\\ &\le \tilde{C}\|\bar{x}_k-x_k\|^{\min\{\frac{\alpha}{2}+1,\theta\}}+Lc^2/2\|\bar{x}_k-x_k\|^2\\ &\le (\tilde{C}+Lc^2/2)\|\bar{x}_k-x_k\|^{\min\{\frac{\alpha}{2}+1,\theta\}},\end{aligned}$$

which together with (2.2) gives

$$\mathrm{dist}(x_k+d_k,X^*)\le\kappa^{-1}\|F(x_k+d_k)\|\le(\tilde{C}+Lc^2/2)\kappa^{-1}\mathrm{dist}(x_k,X^*)^{\min\{\frac{\alpha}{2}+1,\theta\}}.$$

Letting $\hat{c}:=(\tilde{C}+Lc^2/2)\kappa^{-1}$, we complete the proof. □ □

Lemma 3.2. *Under the same conditions in Lemma* 3.1*, there exists an index* $\bar{k}$ *such that for all* $k\ge\bar{k}$ *it holds:* (i) $x_k, x_k+d_k\in N(x^*,r/2)$; (ii) $\|F(x_k+d_k)\|\le\xi\|F(x_k)\|$.

Proof. By Lemma 3.1, similarly as the proof of [9, Lemma 11], we can prove the result. □ □

Theorem 3.3. *Under the same conditions in Lemma* 3.1*, the whole sequence* $\{x_k\}$ *converges to* x^* *with*

$$\|x_{k+1}-x^*\|=O(\|x_k-x^*\|^{\min\{\frac{\alpha}{2}+1,\theta\}}).$$

Proof. By Lemma 3.2 and Step 2 of Algorithm 3.1, we have $x_{k+1}=x_k+d_k$ and $\|F(x_{k+1})\|\le\xi\|F(x_k)\|$ for all $k\ge\bar{k}$. It follows that $\lim_{k\to\infty}\|F(x_k)\|=0$, which together with (2.2) yields $\lim_{k\to\infty}\mathrm{dist}(x_k,X^*)=0$. Thus, from Lemma 3.1, for all sufficiently large k,

$$\mathrm{dist}(x_{k+1},X^*)\le\hat{c}\mathrm{dist}(x_k,X^*)^{\min\{\frac{\alpha}{2}+1,\theta\}}=\hat{c}\mathrm{dist}(x_k,X^*)^{\min\{\frac{\alpha}{2}+1,\theta\}-1}\mathrm{dist}(x_k,X^*).$$

Since $\min\{\frac{\alpha}{2}+1,\theta\}>1$ and $\lim_{k\to\infty}\mathrm{dist}(x_k,X^*)=0$, we have $\hat{c}\mathrm{dist}(x_k,X^*)^{\min\{\frac{\alpha}{2}+1,\theta\}-1}<\frac{1}{2}$ for all sufficiently large k. It follows that for all sufficiently large k,

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

This implies that for all sufficiently large k,

$$\begin{aligned}\mathrm{dist}(x_k,X^*) &\le \|x_k-\bar{x}_{k+1}\|=\|x_{k+1}-\bar{x}_{k+1}-d_k\|\\ &\le \mathrm{dist}(x_{k+1},X^*)+\|d_k\|\\ &\le \frac{1}{2}\mathrm{dist}(x_k,X^*)+\|d_k\|,\end{aligned}$$

that is

$$\mathrm{dist}(x_k,X^*)\le2\|d_k\|. \tag{3.20}$$

By (3.19) and Lemma 3.1, for all sufficiently large k,

$$\|d_{k+1}\|\le c\mathrm{dist}(x_{k+1},X^*)\le c\hat{c}\mathrm{dist}(x_k,X^*)^{\min\{\frac{\alpha}{2}+1,\theta\}}, \tag{3.21}$$

which together with $\lim_{k\to\infty}\mathrm{dist}(x_k,X^*)=0$ gives $\lim_{k\to\infty}d_k=0$ and

$$\|d_{k+1}\|\le\frac{1}{4}\mathrm{dist}(x_k,X^*). \tag{3.22}$$

By (3.20) and (3.22), for all sufficiently large k,

$$\|d_{k+1}\| \le \frac{1}{2}\|d_k\|. \tag{3.23}$$

So, when k is sufficiently large, (3.23) gives

$$\|x_{k+1} - x^*\| = \|\sum_{j=k+1}^{\infty} d_j\| \le \sum_{j=k+1}^{\infty} \|d_j\| \le 2\|d_{k+1}\|. \tag{3.24}$$

This and $\lim_{k\to\infty} d_k = 0$ yield $\lim_{k\to\infty} x_k = x^*$. Further, by (3.21) and (3.24) we have

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^{\min\{\frac{\alpha}{2}+1,\theta\}}} \le \lim_{k\to\infty} \frac{2\|d_{k+1}\|}{\mathrm{dist}(x_k, X^*)^{\min\{\frac{\alpha}{2}+1,\theta\}}} < \infty.$$

The proof is completed. □ □

4. Numerical results

We apply Algorithm 3.1 to solve the nonlinear equations arising in the well-known linear complementarity problem (LCP):

$$\text{(LCP)}\;\; u \ge 0,\; v \ge 0,\; u = Mv + q,\; u^T v = 0, \tag{4.1}$$

where $M \in \mathbb{R}^{n\times n}$ and $q \in \mathbb{R}^n$ are given matrix and vector. To reformulate the LCP into an equivalent system of equations, we define the function $\phi : \mathbb{R}^2 \to \mathbb{R}$ as

$$\phi(a, b) = a^2 + b^2 - \mathrm{sgn}(a + b)(a + b)^2, \;\; \forall (a, b) \in \mathbb{R}^2, \tag{4.2}$$

where $\mathrm{sgn}(t) := \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{if } t = 0. \\ -1 & \text{if } t < 0 \end{cases}$

Proposition 4.1. (i) $\phi(a, b) = 0 \Longleftrightarrow a \ge 0,\; b \ge 0,\; ab = 0.$
(ii) *ϕ is continuously differentiable at any $(a, b) \in \mathbb{R}^2$ whose gradient is given by*

$$\nabla\phi(a, b) = 2\begin{bmatrix} a - |a + b| \\ b - |a + b| \end{bmatrix}.$$

Proof. Let $f(t) := \mathrm{sgn}(t)t^2$. Since f_q is a bijective function, it follows that

$$\begin{aligned} \phi(a, b) = 0 \;\; &\Longleftrightarrow \;\; f(\sqrt{a^2 + b^2}) - f(a + b) = 0 \\ &\Longleftrightarrow \;\; f(\sqrt{a^2 + b^2}) = f(a + b) \\ &\Longleftrightarrow \;\; \sqrt{a^2 + b^2} = a + b \\ &\Longleftrightarrow \;\; a \ge 0,\; b \ge 0,\; ab = 0. \end{aligned}$$

The result (ii) holds because $f(t)$ is continuously differentiable everywhere with $f'(t) = 2|t|$. □

Let $x := (u, v)$. By using the function ϕ, we may have that solving the LCP is equivalent to computing a solution of the smooth nonlinear equations

$$F(x) = \begin{pmatrix} Mv + q - u \\ \phi(u_1, v_1) \\ \vdots \\ \phi(u_n, v_n) \end{pmatrix} = 0. \tag{4.3}$$

In the following, we apply Algorithm 3.1 to solve the nonlinear equations (4.3). The parameters are chosen as $\rho = 10^{-3}, \xi = 0.5, \gamma = 10^{-5}, \zeta = 10^{-5}, \delta = 0.8, \theta = 1, \tau = 0.5$ and σ, α are given the specific experiments. In Step 2, GMRES is used as the linear solver to find the inexact direction d_k. Moreover, we use $\|F(x_k)\| \leq 10^{-5}$ as the stopping criterion.

We test two classes of LCPs defined as follows:

LCP (i) Let M be the block diagonal matrix with $\frac{N_1^T N_1}{\|N_1^T N_1\|}, ..., \frac{N_4^T N_4}{\|N_4^T N_4\|}$ as block diagonals, i.e., $M = \mathbf{diag}(\frac{N_i^T N_i}{\|N_i^T N_i\|})$ with $N_i = \mathbf{rand}(\frac{n}{4}, \frac{n}{4})$ for $i = 1, ..., 4$. Take $q = \mathbf{rand}(n, 1)$. Obviously, the matrix M is positive semidefinite.

LCP (ii) Let $M = \mathbf{diag}(\frac{N_i}{\|N_i\|} - \mathbf{eye}(n/4))$ with $N_i = \mathbf{rand}(\frac{n}{4}, \frac{n}{4})$ for $i = 1, ..., 4$. Take $q = \mathbf{rand}(n, 1)$.

We use $v_0 = (1, 0, ..., 0)^T$ and $u_0 = Mv_0 + q$ as the initial point. Tables 1 and 2 show numerical experimental results of Algorithm 3.1 with different values of σ and α, in which **IT** denotes the iteration number, **CPU** denotes the CPU time in seconds, **Fx** denotes the value of $\|F(x_k)\|$ at the final iteration point and "–" stands for that the algorithm fails to find the solution. These numerical results show that Algorithm 3.1 is efficient for solving LCPs. It can find a solution point meeting the desired accuracy in very few iteration numbers and in short CPU time. Moreover, from our numerical implementations, we may find that Algorithm 3.1 with $\sigma = 1$, i.e., $\mu_k = \|F_k\|^\alpha$, has advantage over it with $\sigma = 0$, i.e., $\mu_k = \|J_k^T F_k\|^\alpha$. At last, we point out that we have tested Algorithm 3.1 with different values of τ and found that the numerical performances are same. □

Table 1. Numerical results for LCP (i).

		$\sigma = 0$			$\sigma = 0.5$			$\sigma = 1$		
α	n	IT	CPU	Fx	IT	CPU	Fx	IT	CPU	Fx
1	1000	7	2.94	1.176e-06	7	2.86	1.668e-06	6	2.58	6.607e-06
	1300	7	5.65	2.905e-07	7	5.65	3.887e-07	6	4.56	2.253e-06
	1500	7	7.86	8.645e-08	7	7.58	1.247e-07	6	6.84	3.491e-06
	1700	–	–	–	–	–	–	–	–	–
	2000	6	13.93	9.715e-06	6	13.72	6.021e-06	6	13.59	1.957e-06
	2500	7	28.52	2.317e-07	6	23.35	7.880e-06	6	23.41	1.507e-06
2	1000	–	–	–	6	2.73	1.753e-06	6	2.44	3.739e-07
	1300	–	–	–	6	4.76	1.936e-06	6	4.55	3.479e-09
	1500	7	8.62	2.622e-06	7	8.02	5.073e-10	5	5.75	8.161e-06
	1700	–	–	–	–	–	–	–	–	–
	2000	7	15.30	1.478e-07	7	16.38	2.749e-11	5	11.40	8.913e-06
	2500	–	–	–	6	24.52	1.018e-06	6	24.18	7.721e-10

Table 2. Numerical results for LCP (ii).

		$\sigma = 0$			$\sigma = 0.5$			$\sigma = 1$		
α	n	IT	CPU	Fx	IT	CPU	Fx	IT	CPU	Fx
1	1000	5	2.77	2.895e-06	5	2.17	4.086e-06	5	2.55	4.250e-06
	1300	5	4.02	1.334e-07	5	4.11	5.072e-07	4	3.19	9.426e-06
	1500	4	4.82	3.402e-08	3	3.50	8.312e-06	3	3.34	5.988e-06
	1700	5	7.72	4.276e-06	5	7.67	2.346e-06	5	7.87	1.505e-06
	2000	4	10.24	5.374e-07	4	9.72	3.970e-07	3	8.73	9.130e-06
	2500	5	21.22	2.845e-07	5	20.35	4.013e-07	4	16.23	8.123e-06
2	1000	4	1.77	3.116e-06	4	1.88	3.262e-06	4	1.82	2.224e-06
	1300	4	3.27	2.005e-07	4	3.26	1.156e-07	3	3.52	1.381e-08
	1500	3	3.56	5.093e-08	3	3.42	3.387e-08	3	3.75	1.226e-08
	1700	4	6.02	7.728e-06	4	6.23	1.625e-06	4	6.27	1.532e-07
	2000	3	7.15	1.553e-07	3	7.18	8.826e-08	3	7.40	4.989e-08
	2500	4	16.81	1.474e-07	4	16.52	1.299e-07	4	16.78	1.053e-07

5. Conclusions

We have presented a family of inexact Levenberg-Marquardt methods for the nonlinear equations. The presented LM method takes more general LM parameters and perturbation vectors which are convex combinations of $\|F_k\|^\alpha$ and $\|J_k^T F_k\|^\alpha$ and $\|F_k\|^{\alpha+\theta}$ and $\|J_k^T F_k\|^{\alpha+\theta}$. Under the Höderian local error bound condition and the Höderian continuity of the Jacobian, we have derived an explicit formula of the convergence order of these inexact LM methods. Moreover, we have developed a family of globally convergent inexact LM methods and showed its effectiveness by some numerical experiments.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

Research of this paper was supported by Natural Science Foundation of Henan Province (222300420520) and Key scientific research projects of Higher Education of Henan Province (22A110020).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. 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. https://doi.org/10.1080/02331934.2018.1435655

2. H. Dan, N. Yamashita, M. Fukushima, Convergence properties of the inexact Levenberg-Marquardt method under local error bound, *Optim. Method. Softw.*, **17** (2002), 605–626. http://dx.doi.org/10.1080/1055678021000049345
3. F. Facchinei, C. Kanzow, A nonsmooth inexact Newton method for the solution of large scale nonlinear complementarity problems, *Math. Program.*, **76** (1997), 493–512. https://doi.org/10.1007/bf02614395
4. J. Y. Fan, J. Y. Pan, Inexact Levenberg-Marquardt method for nonlinear equations, *Discrete Cont. Dyn.-B*, **4** (2004), 1223–1232. http://dx.doi.org/10.3934/dcdsb.2004.4.1223
5. J. Y. Fan, J. Y. Pan, On the convergence rate of the inexact Levenberg-Marquardt method, *J. Ind. Manag. Optim.*, **7** (2011), 199–210. http://dx.doi.org/10.3934/jimo.2011.7.199
6. J. Y. Fan, Y. X. Yuan, On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption, *Computing*, **74** (2005), 23–39. https://doi.org/10.1007/s00607-004-0083-1
7. A. Fischera, P. K. Shuklaa, M. Wang, On the inexactness level of robust Levenberg-Marquardt methods, *Optimization*, **59** (2010), 273–287. https://doi.org/10.1080/02331930801951256
8. G. W. Stewart, J. G. Sun, *Matrix Perturbation Theory*, San Diego: Academic Press, 1990.
9. J. Y. Tang, J. C. Zhou, Quadratic convergence analysis of a nonmonotone Levenberg-Marquardt type method for the weighted nonlinear complementarity problem, *Comput. Optim. Appl.*, **80** (2021), 213–244. http://dx.doi.org/10.1007/S10589-021-00300-8
10. J. Y. Tang, H. C. Zhang, A nonmonotone smoothing Newton algorithm for weighted complementarity problems, *J. Optim Theory Appl.*, **189** (2021), 679–715. http://dx.doi.org/10.1007/S10957-021-01839-6
11. H. Y. Wang, J. Y. Fan, Convergence rate of the Levenberg-Marquardt method under Hölderian local error bound, *Optim. Methods Softw.*, **35** (2020), 767–786. http://dx.doi.org/10.1080/10556788.2019.1694927
12. H. Y. Wang, J. Y. Fan, Convergence properties of inexact Levenberg-Marquardt method under Hölderian local error bound, *J. Ind. Manag. Optim.*, **17** (2021), 2265–2275. http://doi.org/10.3934/jimo.2020068
13. N. Yamashita, M. Fukushima, On the rate of convergence of the Levenberg-Marquardt method, *Computing*, **15** (2001), 239–249. http://doi.org/10.1007/978-3-7091-6217-0-18
14. M. Zeng, G. Zhou, Improved convergence results of an efficient Levenberg-Marquardt method for nonlinear equations, *J. Appl. Math. Comput.*, **68** (2022), 3655–367. http://doi.org/10.1007/S12190-021-01599-6
15. L. Zheng, L. Chen, Y. F. Ma, A variant of the Levenberg-Marquardt method with adaptive parameters for systems of nonlinear equations, *AIMS Math.*, **7** (2021), 1241–1256. http://doi.org/10.3934/math.2022073
16. L. Zheng, L. Chen, Y. X. Tang, Convergence rate of the modified Levenberg-Marquardt method under Hölderian local error bound, *Open Math.*, **20** (2022), 998–1012. http://dx.doi.org/10.1515/MATH-2022-0485

AIMS Press