Research Article

Lin Zheng*, Liang Chen, and Yangxin Tang

Convergence rate of the modified Levenberg-Marquardt method under Hölderian local error bound

https://doi.org/10.1515/math-2022-0485
received November 21, 2021; accepted July 21, 2022

Abstract: In this article, we analyze the convergence rate of the modified Levenberg-Marquardt (MLM) method under the Hölderian local error bound condition and the Hölderian continuity of the Jacobian, which are more general than the local error bound condition and the Lipschitz continuity of the Jacobian. Under special circumstances, the convergence rate of the MLM method coincides with the results presented by Fan. A globally convergent MLM algorithm by the trust region technique will also be given.

Keywords: nonlinear equation, Levenberg-Marquardt method, convergence rate, Hölderian local error bound

MSC 2020: 65K05, 90C30

1 Introduction

We consider the system of nonlinear equations

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

where $F(x) : R^n \to R^n$ is continuously differentiable. Denote by X^* and $\|\cdot\|$, the solution set of (1.1) and 2-norms, respectively. Throughout the article, we assume that X^* is nonempty because equation (1.1) may have no solutions for the nonlinearity of $F(x)$.

Nonlinear equations play an important role in many fields of science, and many numerical methods are developed to solve nonlinear equations [1–3]. Many efficient solution techniques such as the Newton method, quasi-Newton methods, the Gauss-Newton method, trust region methods, and the Levenberg-Marquardt method are available for this problem [1–22].

The most common method to solve (1.1) is the Newton method. At every iteration, it computes the trial step

$$d_k^N = -J_k^{-1}F_k, \tag{1.2}$$

where $F_k = F(x_k)$ and $J_k = F'(x_k)$ is the Jacobian. If $J(x)$ is Lipschitz continuous and nonsingular at the solution, then the convergence of the Newton method is quadratic. However, the Newton method has some disadvantages, especially when the Jacobian matrix J_k is singular or near singular. To overcome the

* **Corresponding author: Lin Zheng,** School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, China; Institute of Quantitative Economics, Anhui University of Finance and Economics, Bengbu 233030, China, e-mail: hbzhenglin@126.com
Liang Chen: School of Sciences, Changzhou Institute of Technology, Changzhou, Jiangsu 213032, China
Yangxin Tang: School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, China; Institute of Quantitative Economics, Anhui University of Finance and Economics, Bengbu 233030, China

Open Access. © 2022 Lin Zheng *et al.*, published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

difficulties caused by the possible singularity of J_k, the Levenberg-Marquardt method [2,3] computes the trial step by

$$d_k^{\mathrm{LM}} = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k, \tag{1.3}$$

where $\lambda_k > 0$ is the LM parameter that is updated in every iteration.

However, it is too strong to assume that the Jacobian is nonsingular. The local error bound requirement is weaker than the nonsingularity condition. It is necessary that

$$c\ \mathrm{dist}(x, X^*) \le \|F(x)\|, \quad \forall x \in N(x^*) \tag{1.4}$$

holds for some constant $c > 0$, where $\mathrm{dist}(x, X^*)$ is the distance from x to X^* and $N(x^*)$ is some neighborhood of $x^* \in X^*$.

Under the local error bound condition, Yamashita and Fukushima [4] and Fan and Yuan [5] show that the LM method has quadratic convergence if the LM parameter was chosen as $\lambda_k = \|F_k\|^2$ and $\lambda_k = \|F_k\|^\alpha$ with $\alpha \in [1, 2]$, respectively. Interested readers are referred to [6–8] for related work.

Inspired by the two-step Newton's method, Fan [9] presented a modified Levenberg-Marquardt (MLM) method with an approximate LM step

$$d_k^{\mathrm{MLM}} = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k), \tag{1.5}$$

where $y_k = x_k + d_k^{\mathrm{LM}}$, and the trial step is

$$s_k = d_k^{\mathrm{LM}} + d_k^{\mathrm{MLM}}. \tag{1.6}$$

The MLM method has cubic convergence under the local error bound condition. For more general cases, Fan [19] gave an accelerated version of the MLM method. She also extended the LM parameter $\lambda_k = \|F(x_k)\|^\alpha$ from $\alpha \in [1, 2]$ to $\alpha \in [0, 2]$. The convergence order of the accelerating MLM method is $\min\{1 + 2\alpha, 3\}$, which is a continuous function of α.

To save more Jacobian calculations and achieve a fast convergence rate, Zhao and Fan [10] and Chen [11] presented a higher-order Levenberg-Marquardt method by computing the approximate step twice, and the method has biquadratic convergence under the local error bound condition.

At present, the Levenberg-Marquardt method is very widely used. It is a classical method for solving nonlinear least squares problems, and it can be used to solve financial problems [23,24]. In real applications, some nonlinear equations may not satisfy the local error bound condition but satisfy the Hölderian local error bound condition defined as follows.

Definition 1.1. We say that $F(x)$ provides a Hölderian local error bound of order $\gamma \in (0, 1]$ in some neighborhood of $x^* \in X^*$, if there exists a constant $c > 0$ such that

$$c\mathrm{dist}(x, X^*) \le \|F(x)\|^\gamma, \quad \forall x \in N(x^*). \tag{1.7}$$

We can see that the Hölderian local error bound condition is more generalized from (1.4) and (1.7); when $\gamma = 1$, the local error bound condition is included as a special case. Hence, the local error bound condition is stronger. For example, the Powell singular function [25]

$$h(x_1, x_2, x_3, x_4) = (x_1 + 10x_2, \sqrt{5}(x_3 - x_4), (x_2 - 2x_3)^2, \sqrt{10}(x_1 - x_4)^2)^T$$

satisfies the Hölderian local error bound condition of order $\frac{1}{2}$ around the zero point but does not satisfy the local error bound condition [12]. In a biochemical reaction network, the problem of finding the moiety conserved steady state can be formulated as a system of nonlinear equations, which satisfies the Hölderian local error bound condition [12]. Recently, some scholars discussed the convergence results of the LM method under the Hölderian local error bound condition and the Hölderian continuity of the Jacobian [12–14,22]. In this article, we will investigate the convergence rate of the MLM method under the Hölderian local error bound condition and Hölderian continuity of the Jacobian, which are more general than the local error bound condition and the Lipschitz continuity of the Jacobian.

This article is organized as follows. In Section 2, we propose an MLM algorithm and show that it converges globally under the Hölderian continuity of the Jacobian. In Section 3, we study the convergence rate of the algorithm under the Hölderian local error bound condition and the Hölderian continuity of the Jacobian. We finish the work with some conclusions and references.

2 A globally convergent MLM algorithm

In this section, we propose an MLM algorithm by the trust region technique and then prove that it converges globally under the Hölderian continuity of the Jacobian.

We take

$$\Phi(x) = \|F(x)\|^2 \tag{2.1}$$

as the merit function for (1.1). We define the actual reduction of $\Phi(x)$ at the kth iteration as

$$\mathrm{Ared}_k = \|F_k\|^2 - \|F(x_k + d_k^{\mathrm{LM}} + d_k^{\mathrm{MLM}})\|^2. \tag{2.2}$$

The predicted reduction needs to be nonnegative.

Note that the step d_k^{LM} in (1.3) is the minimizer of the convex minimization problem

$$\min_{d\in R^n} \|F_k + J_k d\|^2 + \lambda_k\|d\|^2 \triangleq \varphi_{k,1}(d). \tag{2.3}$$

If we let

$$\Delta_{k,1} = \|d_k^{\mathrm{LM}}\| = \|-(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k\|, \tag{2.4}$$

then it can be verified that d_k^{LM} is also a solution of the trust region subproblem

$$\min_{d\in R^n} \|F_k + J_k d\|^2 \quad \text{s.t. } \|d\| \le \Delta_{k,1}. \tag{2.5}$$

We obtain from the result given by Powell [15] that

$$\|F_k\|^2 - \|F_k + J_k d_k^{\mathrm{LM}}\|^2 \ge \|J_k^T F_k\| \min\left\{\|d_k^{\mathrm{LM}}\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}. \tag{2.6}$$

In the same way, the step d_k^{MLM} in (1.5) is not only the minimizer of the problem

$$\min_{d\in R^n} \|F(y_k) + J_k d\|^2 + \lambda_k\|d\|^2 \triangleq \varphi_{k,2}(d), \tag{2.7}$$

but also the solution of the trust region problem

$$\min_{d\in R^n} \|F(y_k) + J_k d\|^2 \quad \text{s.t. } \|d\| \le \Delta_{k,2}, \tag{2.8}$$

where

$$\Delta_{k,2} = \|d_k^{\mathrm{MLM}}\| = \|-(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k)\|. \tag{2.9}$$

Thus, we also obtain

$$\|F(y_k)\|^2 - \|F(y_k) + J_k d_k^{\mathrm{MLM}}\|^2 \ge \|J_k^T F(y_k)\| \min\left\{\|d_k^{\mathrm{MLM}}\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|}\right\}. \tag{2.10}$$

We define the newly predicted reduction from (2.6) and (2.10) as follows:

$$\mathrm{Pred}_k = \|F_k\|^2 - \|F_k + J_k d_k^{\mathrm{LM}}\|^2 + \|F(y_k)\|^2 - \|F(y_k) + J_k d_k^{\mathrm{MLM}}\|^2, \tag{2.11}$$

which satisfies

$$\mathrm{Pred}_k \geq \|J_k^T F_k\| \min\left\{\|d_k^{\mathrm{LM}}\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\} + \|J_k^T F(y_k)\| \min\left\{\|d_k^{\mathrm{MLM}}\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|}\right\}. \tag{2.12}$$

The ratio of the actual reduction to the predicted reduction

$$r_k = \frac{\mathrm{Ared}_k}{\mathrm{Pred}_k} \tag{2.13}$$

is used in deciding whether to accept the trial step and how to update the MLM parameter λ_k.

The algorithm is presented as follows.

Algorithm 2.1. Given $x_1 \in R^n$, $0 < \alpha \leq 2$, $\mu_1 > m > 0$, $0 < p_0 \leq p_1 \leq p_2 < 1$, $a_1 > 1 > a_2 > 0$. Set $k := 1$.

Step 1. If $\|J_k^T F_k\| = 0$, then stop. Solve

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F_k \quad \text{with } \lambda_k = \mu_k \|F_k\|^\alpha \tag{2.14}$$

to obtain d_k^{LM} and set

$$y_k = x_k + d_k^{\mathrm{LM}}.$$

Solve

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F(y_k) \tag{2.15}$$

to obtain d_k^{MLM} and set

$$s_k = d_k^{\mathrm{LM}} + d_k^{\mathrm{MLM}}.$$

Step 2. Compute $r_k = \mathrm{Ared}_k / \mathrm{Pred}_k$. Set

$$x_{k+1} = \begin{cases} x_k + s_k, & \text{if } r_k \geq p_0, \\ x_k, & \text{otherwise}. \end{cases} \tag{2.16}$$

Step 3. Choose μ_{k+1} as

$$\mu_{k+1} = \begin{cases} a_1\mu_k, & \text{if } r_k < p_1, \\ \mu_k, & \text{if } r_k \in [p_1, p_2], \\ \max\{a_2\mu_k, m\}, & \text{if } r_k > p_2. \end{cases} \tag{2.17}$$

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

We give some assumptions before studying the global convergence of Algorithm 2.1.

Assumption 2.2.

(a) The Jacobian $J(x)$ is Hölderian continuous of order $\nu \in (0, 1]$, i.e., there exists a positive constant κ_{hj} such that

$$\|J(x) - J(y)\| \leq \kappa_{hj}\|x - y\|^\nu, \quad \forall x, y \in R^n. \tag{2.18}$$

(b) $J(x)$ is bounded above, i.e., there exists a positive constant κ_{bj} such that

$$\|J(x)\| \leq \kappa_{bj}, \quad \forall x \in R^n. \tag{2.19}$$

From (2.18), we can obtain

$$\|F(y) - F(x) - J(x)(y - x)\| = \left\| \int_0^1 J(x + t(y - x))(y - x)\mathrm{d}t - J(x)(y - x) \right\| \tag{2.20}$$

$$\le \|y-x\| \int_0^1 \|J(x+t(y-x)) - J(x)\| \mathrm{d}t$$

$$\le \kappa_{hj} \|y-x\|^{1+v} \int_0^1 t^v \mathrm{d}t$$

$$= \frac{\kappa_{hj}}{1+v} \|y-x\|^{1+v}.$$

Theorem 2.3. *Let Assumption* 2.2 *hold. Then Algorithm* 2.1 *terminates in finite iterations or satisfies*

$$\lim_{k\to\infty} \|J_k^T F_k\| = 0. \tag{2.21}$$

Proof. We prove the theorem by contradiction. Suppose that (2.21) is not true, then there exists a positive constant τ and infinitely many k such that

$$\|J_k^T F_k\| \ge \tau. \tag{2.22}$$

Let S_1, S_2 be the sets of the indices as follows:

$$S_1 = \{k \| \|J_k^T F_k\| \ge \tau\},$$
$$S_2 = \left\{k \| \|J_k^T F_k\| \ge \frac{\tau}{2} \quad \text{and} \quad x_{k+1} \ne x_k\right\}.$$

Then, S_1 is an infinite set. In the following, we will derive the contradictions whether S_2 is finite or infinite.

Case I: S_2 is finite. Then, the set

$$S_3 = \{k \| \|J_k^T F_k\| \ge \tau \quad \text{and} \quad x_{k+1} \ne x_k\}$$

is also finite. Let $\tilde{k}$ be the largest index of S_3. Then, $x_{k+1} = x_k$ holds for all $k \in \{k > \tilde{k} | k \in S_1\}$. Define the indices set

$$S_4 = \{k > \tilde{k} \| \|J_k^T F_k\| \ge \tau \quad \text{and} \quad x_{k+1} = x_k\}.$$

If $k \in S_4$, we can deduce that $\|J_{k+1}^T F_{k+1}\| \ge \tau$ and $x_{k+2} = x_{k+1}$. Hence, we have $k+1 \in S_4$. By induction, we know that $\|J_k^T F_k\| \ge \tau$ and $x_{k+1} = x_k$ hold for all $k > \tilde{k}$, which implies that $r_k < p_0$. Therefore, we have

$$\mu_k \to \infty \quad \text{and} \quad \lambda_k \to \infty \tag{2.23}$$

due to (2.14). Hence, we obtain

$$d_k^{\mathrm{LM}} \to 0. \tag{2.24}$$

Moreover, it follows from (2.8), (2.20), and (2.23) that

$$\begin{aligned}
\|d_k^{\mathrm{MLM}}\| &= \|-(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k)\| \\
&\le \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k\| + \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T J_k d_k^{\mathrm{LM}}\| + \frac{\kappa_{hj}}{1+v} \|d_k^{\mathrm{LM}}\|^{1+v} \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \\
&\le \|d_k^{\mathrm{LM}}\| + \|d_k^{\mathrm{LM}}\| + \frac{\kappa_{hj} \cdot \kappa_{bj}}{\lambda_k (1+v)} \|d_k^{\mathrm{LM}}\|^{1+v} \\
&\le c_1 \|d_k^{\mathrm{LM}}\|
\end{aligned} \tag{2.25}$$

holds for all sufficiently large k, where c_1 is a positive constant. Therefore, we have

$$\|s_k\| = \|d_k^{\mathrm{LM}} + d_k^{\mathrm{MLM}}\| \le (1 + c_1) \|d_k^{\mathrm{LM}}\|. \tag{2.26}$$

Furthermore, it follows from (2.12), (2.19), (2.22), (2.24), and (2.26) that

$$
\begin{aligned}
|r_k - 1| &= \left| \frac{\mathrm{Ared}_k - \mathrm{Pred}_k}{\mathrm{Pred}_k} \right| \\
&\le \left| \frac{\|F(x_k + d_k^{\mathrm{LM}} + d_k^{\mathrm{MLM}})\|^2 - \|F_k + J_k d_k^{\mathrm{LM}}\|^2 + \|F(y_k)\|^2 - \|F(y_k) + J_k d_k^{\mathrm{MLM}}\|^2}{\|J_k^T F_k\| \min\left\{\|d_k^{\mathrm{LM}}\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\} + \|J_k^T F(y_k)\| \min\left\{\|d_k^{\mathrm{MLM}}\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|}\right\}} \right| \\
&\le \frac{\|F_k + J_k s_k\| O(\|d_k^{\mathrm{LM}}\|^{1+\nu}) + O(\|d_k^{\mathrm{LM}}\|^{2+2\nu}) + \|F_k + J_k d_k^{\mathrm{LM}}\| O(\|d_k^{\mathrm{LM}}\|^{1+\nu})}{\|J_k^T F_k\| \min\left\{\|d_k^{\mathrm{LM}}\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}} \\
&\le \frac{\|F_k + J_k d_k\| O(\|d_k^{\mathrm{LM}}\|^{1+\nu}) + O(\|d_k^{\mathrm{LM}}\|^{2+\nu}) + O(\|d_k^{\mathrm{LM}}\|^{2+2\nu})}{\|J_k^T F_k\| \min\left\{\|d_k^{\mathrm{LM}}\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}} \to 0,
\end{aligned} \tag{2.27}
$$

which implies that $r_k \to 1$. In view of the updating rule of μ_k, we know that there exists a positive constant $\tilde{m} > m$ such that $\mu_k < \tilde{m}$ holds for all sufficiently large k, which is a contradiction to (2.23).

Case II: S_2 is infinite. It follows from (2.12) and (2.19) that

$$
\begin{aligned}
\|F_1\|^2 &\ge \sum_{k \in S_2} (\|F_k\|^2 - \|F_{k+1}\|^2) \\
&\ge \sum_{k \in S_2} p_0 \mathrm{Pred}_k \\
&\ge \sum_{k \in S_2} p_0 \left\{ \|J_k^T F_k\| \min\left\{\|d_k^{\mathrm{LM}}\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\} + \|J_k^T F(y_k)\| \min\left\{\|d_k^{\mathrm{MLM}}\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|}\right\}\right\} \\
&\ge \sum_{k \in S_2} \frac{p_0 \tau}{2} \min\left\{\|d_k^{\mathrm{LM}}\|, \frac{\tau}{2\kappa_{bj}^2}\right\},
\end{aligned} \tag{2.28}
$$

which implies

$$\lim_{k \to \infty, k \in S_2} d_k^{\mathrm{LM}} = 0. \tag{2.29}$$

Then, from definition of d_k^{LM}, we have

$$\lambda_k \to +\infty, \; k \in S_2. \tag{2.30}$$

Similarly to (2.25), there exists a positive c_2 such that

$$\|d_k^{\mathrm{MLM}}\| \le c_2 \|d_k^{\mathrm{LM}}\| \tag{2.31}$$

holds for all sufficiently large $k \in S_2$. From (2.28), we obtain

$$\|s_k\| \le \|d_k^{\mathrm{LM}} + d_k^{\mathrm{MLM}}\| \le (1 + c_2)\|d_k^{\mathrm{LM}}\|. \tag{2.32}$$

So, we derive that

$$\sum_{k \in S_2} \|s_k\| = \sum_{k \in S_2} \|d_k^{\mathrm{LM}} + d_k^{\mathrm{MLM}}\| < +\infty. \tag{2.33}$$

Furthermore, it follows from (2.18) and (2.19) that

$$\sum_{k \in S_2} \big| \|J_k^T F_k\| - \|J_{k+1}^T F_{k+1}\| \big| < +\infty.$$

Since (2.22) holds for infinitely many k, there exists a large $\hat{k}$ such that $\|J_k^T F_k\| \ge \tau$ and

$$\sum_{k \in S_2, k \ge \hat{k}} \big| \|J_k^T F_k\| - \|J_{k+1}^T F_{k+1}\| \big| < \frac{\tau}{2}.$$

From (2.28) to (2.31), we can deduce that $\lim_{k \to \infty} x_k$ exists and

$$d_k^{\mathrm{LM}} \to 0, \quad d_k^{\mathrm{MLM}} \to 0. \tag{2.34}$$

Therefore, we can obtain

$$\mu_k \to +\infty. \tag{2.35}$$

In the same way as proved in case I, we can also have

$$r_k \to 1.$$

Hence, there exists a positive constant $\bar{m} > m$ such that $\mu_k < \bar{m}$ holds for all sufficiently large k, which is a contradiction to (2.35). The proof is completed. □

3 Convergence rate of Algorithm 2.1

In this section, we analyze the convergence rate of Algorithm 2.1 under the Hölderian local error bound condition and the Hölderian continuity of the Jacobian. We assume that the sequence $\{x_k\}$ generated by the MLM method converges to the solution set X^* of (1.1) and lies in some neighborhood of $x^* \in X^*$.

First, we will make the following assumption for studying the local convergence theory.

Assumption 3.1.

(a) $F(x)$ provides a Hölderian local error bound of order $\gamma \in (0, 1]$ in some neighborhood of $x^* \in X^*$, i.e., there exist constants $c > 0$ and $0 < b < 1$ such that

$$c\operatorname{dist}(x, X^*) \le \|F(x)\|^{\gamma}, \quad \forall x \in N(x^*, b), \tag{3.1}$$

where $N(x^*, b) = \{x \in R^n | \|x - x^*\| \le b\}$.

(b) $J(x)$ is Hölderian continuous of order $v \in (0, 1]$, i.e., there exists a positive constant κ_{hj} such that

$$\|J(x) - J(y)\| \le \kappa_{hj}\|x - y\|^{v}, \quad \forall x, y \in N(x^*, b). \tag{3.2}$$

Similar to (2.20), we have

$$\|F(y) - F(x) - J(x)(y - x)\| \le \frac{\kappa_{hj}}{1 + v}\|y - x\|^{1+v}, \quad \forall x, y \in N(x^*, b). \tag{3.3}$$

Moreover, there exists a constant $\kappa_{bf} > 0$ such that

$$\|F(y) - F(x)\| \le \kappa_{bf}\|y - x\|, \quad \forall x, y \in N(x^*, b). \tag{3.4}$$

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

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

3.1 Properties of d_k^{LM} and d_k^{MLM}

In the section, we investigate the relationship among $\|d_k^{\mathrm{LM}}\|$, $\|d_k^{\mathrm{MLM}}\|$, and $\operatorname{dist}(x_k, X^*)$.

Suppose 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 $\bar{\Sigma}_{k,1} = \operatorname{diag}(\bar{\sigma}_{k,1}, \ldots, \bar{\sigma}_{k,r})$ with $\bar{\sigma}_{k,1} \ge \bar{\sigma}_{k,2} \ge \cdots \ge \bar{\sigma}_{k,r} > 0$. The corresponding SVD of J_k is

$$\begin{aligned}J_k &= U_k\Sigma_k V_k^T\\ &= (U_{k,1}, U_{k,2}, U_{k,3})\begin{pmatrix}\Sigma_{k,1} & & \\ & \Sigma_{k,2} & \\ & & 0\end{pmatrix}\begin{pmatrix}V_{k,1}^T\\ V_{k,2}^T\\ V_{k,3}^T\end{pmatrix}\\ &= U_{k,1}\Sigma_{k,1}V_{k,1}^T + U_{k,2}\Sigma_{k,2}V_{k,2}^T,\end{aligned}$$

where $\Sigma_{k,1} = \operatorname{diag}(\sigma_{k,1},\ldots,\sigma_{k,r})$ with $\sigma_{k,1} \geq \sigma_{k,2} \geq \cdots \geq \sigma_{k,r} > 0$, and $\Sigma_{k,2} = \operatorname{diag}(\sigma_{k,r+1},\ldots,\sigma_{k,r+q})$ with $\sigma_{k,r} \geq \sigma_{k,r+1} \geq \cdots \geq \sigma_{k,r+q} > 0$. In the following, if the context is clear, we will omit the subscription k in $\Sigma_{k,i}$, $U_{k,i}$ and $V_{k,i}$ $(i = 1, 2, 3)$ and write J_k as

$$J_k = U_1\Sigma_1V_1^T + U_2\Sigma_2V_2^T.$$

Lemma 3.2. *Under the conditions of Assumption* 3.1, *if* $x_k, y_k \in N(x^*, b/2)$, *then there exists a constant* $c_3 > 0$ *such that*

$$\|s_k\| \leq c_3\operatorname{dist}(x_k, X^*)^{\min(1,1+\nu-\alpha/2\gamma,(1+\nu)(1+\nu-\alpha/2\gamma)+\nu-\alpha/\gamma)} \tag{3.5}$$

holds for all sufficiently large k.

Proof. Since $x_k \in N(x^*, b/2)$, we obtain

$$\|\bar{x}_k - x^*\| \leq \|\bar{x}_k - x_k\| + \|x_k - x^*\| \leq 2\|x_k - x^*\| \leq b,$$

which implies that $\bar{x}_k \in N(x^*, b)$. From (3.1) and (2.17), we have

$$\lambda_k = \mu_k\, \|F_k\|^\alpha \geq mc^{\alpha/\gamma}\, \|\bar{x}_k - x_k\|^{\alpha/\gamma}. \tag{3.6}$$

From (3.3), we can obtain

$$\|F_k + J_k(\bar{x}_k - x_k)\|^2 = \|F(\bar{x}_k) - F_k - J_k(\bar{x}_k - x_k)\|^2 \leq \left(\frac{\kappa_{hj}}{1+\nu}\right)^2 \|\bar{x}_k - x_k\|^{2+2\nu}.$$

Since d_k^{LM} is the minimizer of $\varphi_{k,1}(d)$, we have

$$\begin{aligned}\|d_k^{\mathrm{LM}}\|^2 &\leq \frac{\varphi_{k,1}(d_k^{\mathrm{LM}})}{\lambda_k}\\ &\leq \frac{\varphi_{k,1}(\bar{x}_k - x_k)}{\lambda_k}\\ &= \frac{\|F_k + J_k(\bar{x}_k - x_k)\|^2 + \lambda_k\, \|\bar{x}_k - x_k\|^2}{\lambda_k}\\ &\leq \frac{\kappa_{hj}^2c^{-\alpha/\gamma}}{m(1+\nu)^2}\, \|\bar{x}_k - x_k\|^{2+2\nu-\alpha/\gamma} + \|\bar{x}_k - x_k\|^2\\ &\leq c_4^2\, \|\bar{x}_k - x_k\|^{2\min(1,1+\nu-\alpha/2\gamma)},\end{aligned} \tag{3.7}$$

where $c_4 = \sqrt{\kappa_{hj}^2c^{-\alpha/\gamma}/m(1+\nu)^2 + 1}$. Then

$$\|d_k^{\mathrm{LM}}\| \leq c_4\, \|\bar{x}_k - x_k\|^{\min(1,1+\nu-\alpha/2\gamma)}. \tag{3.8}$$

It follows from (3.3), we have

$$\begin{aligned}\|d_k^{\mathrm{MLM}}\| &= \|-(J_k^TJ_k + \lambda_kI)^{-1}J_k^TF(y_k)\|\\ &\leq \|(J_k^TJ_k + \lambda_kI)^{-1}J_k^TF_k\| + \|(J_k^TJ_k + \lambda_kI)^{-1}J_k^TJ_kd_k^{\mathrm{LM}}\| + \frac{\kappa_{hj}}{1+\nu}\|d_k^{\mathrm{LM}}\|^{1+\nu}\,\|(J_k^TJ_k + \lambda_kI)^{-1}J_k^T\|\\ &\leq 2\|d_k^{\mathrm{LM}}\| + \frac{\kappa_{hj}}{1+\nu}\|d_k^{\mathrm{LM}}\|^{1+\nu}\,\|(J_k^TJ_k + \lambda_kI)^{-1}J_k^T\|.\end{aligned} \tag{3.9}$$

Now, using the SVD of J_k, we can obtain

$$\|(J_k^TJ_k+\lambda_kI)^{-1}J_k^T\|=\left\|(V_1,V_2,V_3)\begin{pmatrix}\Sigma_{k,1}&&\\&\Sigma_{k,2}&\\&&0\end{pmatrix}\begin{pmatrix}U_1^T\\U_2^T\\U_3^T\end{pmatrix}\right\| \tag{3.10}$$
$$\leq\left\|\begin{pmatrix}(\Sigma_1^2+\lambda_kI)^{-1}\Sigma_1&&\\&(\Sigma_2^2+\lambda_kI)^{-1}\Sigma_2&\\&&0\end{pmatrix}\right\|$$
$$\leq\left\|\begin{pmatrix}\Sigma_1^{-1}&\\&\lambda_k^{-1}\Sigma_2\end{pmatrix}\right\|.$$

By the theory of matrix perturbation [26], we have

$$\|\operatorname{diag}(\Sigma_1-\bar\Sigma_1,\Sigma_2,0)\|\leq\|J_k-J(\bar x_k)\|\leq\kappa_{hj}\,\|\bar x_k-x_k\|^{\nu}.$$

The above inequalities imply

$$\|\Sigma_1-\bar\Sigma_1\|\leq\kappa_{hj}\,\|\bar x_k-x_k\|^{\nu},\ \|\Sigma_2\|\leq\kappa_{hj}\,\|\bar x_k-x_k\|^{\nu}. \tag{3.11}$$

Since $\{x^k\}$ converges to x^*, without loss of generality, we assume that $\kappa_{hj}\,\|\bar x_k-x_k\|^{\nu}\leq\frac{\bar\sigma_r}{2}$ holds for all large k. From (3.11), we have

$$\|\Sigma_1^{-1}\|\leq\frac{1}{\bar\sigma_r-\kappa_{hj}\,\|\bar x_k-x_k\|^{\nu}}\leq\frac{2}{\bar\sigma_r}. \tag{3.12}$$

From (3.6), we can derive

$$\|\lambda_k^{-1}\Sigma_2\|=\frac{\|\Sigma_2\|}{\mu_k\|F(x_k)\|^{\alpha}}\leq\frac{\kappa_{hj}}{mc^{\alpha/\gamma}}\|\bar x_k-x_k\|^{\nu-\frac{\alpha}{\gamma}}. \tag{3.13}$$

From (3.9) and (3.10), we have that there exist positive c_5 and $\bar c$ such that

$$\|d_k^{\mathrm{MLM}}\|\leq2\|d_k^{\mathrm{LM}}\|+c_5\|d_k^{\mathrm{LM}}\|^{1+\nu}\|\bar x_k-x_k\|^{\nu-\alpha/\gamma}\leq\bar c\|\bar x_k-x_k\|^{\min\{1,1+\nu-\alpha/2\gamma,(1+\nu)(1+\nu-\alpha/2\gamma)+\nu-\alpha/\gamma\}} \tag{3.14}$$

holds for all sufficiently large k. Therefore, we can obtain

$$\|s_k\|\leq\|d_k^{\mathrm{LM}}+d_k^{\mathrm{MLM}}\|\leq\|d_k^{\mathrm{LM}}\|+\|d_k^{\mathrm{MLM}}\|\leq c_3\|\bar x_k-x_k\|^{\min\{1,1+\nu-\alpha/2\gamma,(1+\nu)(1+\nu-\alpha/2\gamma)+\nu-\alpha/\gamma\}}, \tag{3.15}$$

where c_3 is a positive constant. The proof is completed. □

The updating rule of μ_k indicates that μ_k is bounded below. Next, we show that μ_k is also bounded above.

Lemma 3.3. *Under the conditions of Assumption* 3.1, *if* $x_k,y_k\in N(x^*,b/2)$ *and*

$$\nu>\max\left\{\frac{1}{\gamma}-1,\ \frac{1}{\gamma(1+\nu)-\frac{\alpha}{2}}-1,\ \frac{1-\gamma}{\gamma(1+\nu)-\frac{\alpha}{2}}\right\},$$

then there exists a constant $M>m$ *such that*

$$\mu_k\leq M \tag{3.16}$$

holds for all sufficiently large k.

Proof. First, we prove that for all sufficiently large k

$$\mathrm{Pred}_k\geq\check c\|F_k\|\|d_k^{\mathrm{LM}}\|^{\max\{1/\gamma,1/(\gamma(1+\nu)-\alpha/2),(1-\gamma)/(\gamma(1+\nu)-\alpha/2)+1\}}, \tag{3.17}$$

where $\check c$ is a positive constant.

We consider two cases:

Case 1: $\|\bar{x}_k - x_k\| \le d_k^{\mathrm{LM}}$. It follows from (3.1), (3.3), (3.8), and $v > 1/\gamma - 1$ that

$$\begin{aligned}\|F_k\| - \|F_k + J_k d_k^{\mathrm{LM}}\| &\ge \|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\|\\ &\ge c^{1/\gamma}\,\|\bar{x}_k - x_k\|^{1/\gamma} - \frac{\kappa_{hj}}{1+v}\,\|\bar{x}_k - x_k\|^{1+v}\\ &\ge c^{1/\gamma}\,\|\bar{x}_k - x_k\|^{1/\gamma}\\ &\ge c_6\,\|d_k^{\mathrm{LM}}\|^{\max\{1/\gamma,1/(\gamma(1+v)-\alpha/2)\}}\end{aligned} \tag{3.18}$$

holds for some $c_6 > 0$.

Case 2: $\|\bar{x}_k - x_k\| > d_k^{\mathrm{LM}}$. From (3.18), we can obtain

$$\begin{aligned}\|F_k\| - \|F_k + J_k d_k^{\mathrm{LM}}\| &\ge \|F_k\| - \left\|F_k + \frac{\|d_k^{\mathrm{LM}}\|}{\|\bar{x}_k - x_k\|}J_k(\bar{x}_k - x_k)\right\|\\ &\ge \|F_k\| - \left\|\left(1 - \frac{\|d_k^{\mathrm{LM}}\|}{\|\bar{x}_k - x_k\|}\right)F_k + \frac{\|d_k^{\mathrm{LM}}\|}{\|\bar{x}_k - x_k\|}(F_k + J_k(\bar{x}_k - x_k))\right\|\\ &\ge \frac{\|d_k^{\mathrm{LM}}\|}{\|\bar{x}_k - x_k\|}(\|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\|)\\ &\ge c_7\|d_k^{\mathrm{LM}}\|\,\|\bar{x}_k - x_k\|^{1/\gamma-1}\\ &\ge \check{c}\,\|d_k^{\mathrm{LM}}\|^{\max\{1/\gamma,(1-\gamma)/(\gamma(1+v)-\alpha/2)+1\}}\end{aligned} \tag{3.19}$$

holds for some $c_7, \check{c} > 0$.

From (3.18) and (3.19), we have

$$\begin{aligned}\|F_k\|^2 - \|F_k + J_k d_k^{\mathrm{LM}}\|^2 &= (\|F_k\| + \|F_k + J_k d_k^{\mathrm{LM}}\|)(\|F_k\| - \|F_k + J_k d_k^{\mathrm{LM}}\|)\\ &\ge \|F_k\|(\|F_k\| - \|F_k + J_k d_k^{\mathrm{LM}}\|)\\ &\ge \check{c}\|F_k\|\,\|d_k^{\mathrm{LM}}\|^{\max(1/\gamma,1/(\gamma(1+v)-\alpha/2),(1-\gamma)/(\gamma(1+v)-\alpha/2)+1\}}.\end{aligned} \tag{3.20}$$

Since d_k^{MLM} is a solution of (2.8), we know that $\|F(y_k)\|^2 - \|F(y_k) + J_k d_k^{\mathrm{MLM}}\|^2 \ge 0$. Hence, we obtain

$$\begin{aligned}\mathrm{Pred}_k &= \|F_k\|^2 - \|F_k + J_k d_k^{\mathrm{LM}}\|^2 + \|F(y_k)\|^2 - \|F(y_k) + J_k d_k^{\mathrm{MLM}}\|^2\\ &\ge \|F_k\|^2 - \|F_k + J_k d_k^{\mathrm{LM}}\|^2\\ &\ge \check{c}\|F_k\|\,\|d_k^{\mathrm{LM}}\|^{\max(1/\gamma,1/(\gamma(1+v)-\alpha/2),(1-\gamma)/(\gamma(1+v)-\alpha/2)+1\}}.\end{aligned}$$

It follows from (3.3), (3.8), and (3.17) that

$$\begin{aligned}|r_k - 1| &= \left|\frac{\mathrm{Ared}_k - \mathrm{Pred}_k}{\mathrm{Pred}_k}\right|\\ &= \left|\frac{\|F(x_k + d_k^{\mathrm{LM}} + d_k^{\mathrm{MLM}})\|^2 - \|F_k + J_k d_k^{\mathrm{LM}}\|^2 + \|F(y_k)\|^2 - \|F(y_k) + J_k d_k^{\mathrm{MLM}}\|^2}{\mathrm{Pred}_k}\right|\\ &\le \frac{\|F_k + J_k s_k\|O(\|d_k^{\mathrm{LM}}\|^{1+v}) + O(\|d_k^{\mathrm{LM}}\|^{2+2v}) + \|F_k + J_k d_k^{\mathrm{LM}}\|O(\|d_k^{\mathrm{LM}}\|^{1+v})}{\check{c}\|F_k\|\,\|d_k^{\mathrm{LM}}\|^{\max\{1/\gamma,1/(\gamma(1+v)-\alpha/2),(1-\gamma)/(\gamma(1+v)-\alpha/2)+1\}}}.\end{aligned}$$

In view of (3.4), (3.8), (3.9), and (3.14), we have

$$\|F_k + J_k d_k^{\mathrm{LM}}\| \le \|F_k\| \tag{3.21}$$

and

$$\begin{aligned}\|F_k + J_k s_k\| &\le \|F_k + J_k d_k^{\mathrm{LM}}\| + \|J_k d_k^{\mathrm{MLM}}\|\\ &\le \|F_k\| + \kappa_{bf}\|d_k^{\mathrm{MLM}}\|\\ &\le O(\|\bar{x}_k - x_k\|^{\min\{1,1+v-\alpha/2\gamma,(1+v)(1+v-\alpha/2\gamma)+v-\alpha/\gamma\}}).\end{aligned} \tag{3.22}$$

Since

$$v > \max\left\{\frac{1}{\gamma} - 1, \frac{1}{\gamma(1+v) - \frac{\alpha}{2}} - 1, \frac{1-\gamma}{\gamma(1+v) - \frac{\alpha}{2}}\right\},$$

then, we can obtain

$$r_k \to 1.$$

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

Lemma 3.3 together with (3.4) indicates that the MLM parameter satisfies

$$\lambda_k = \mu_k \|F_k\|^\alpha \le M\kappa_{bf}^\alpha \|\bar{x}_k - x_k\|^\alpha. \tag{3.23}$$

Hence, the MLM parameter is also bounded above.

3.2 Convergence rate of Algorithm 2.1

By the SVD of J_k, we can obtain

$$d_k^{\mathrm{LM}} = -V_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F_k - V_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F_k, \tag{3.24}$$

$$d_k^{\mathrm{MLM}} = -V_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F(y_k) - V_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F(y_k), \tag{3.25}$$

$$\begin{aligned} F_k + J_k d_k^{\mathrm{LM}} &= F_k - U_1\Sigma_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F_k - U_2\Sigma_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F_k \\ &= \lambda_k U_1(\Sigma_1^2 + \lambda_k I)^{-1}U_1^T F_k + \lambda_k U_2(\Sigma_2^2 + \lambda_k I)^{-1}U_2^T F_k + U_3U_3^T F_k, \end{aligned} \tag{3.26}$$

$$\begin{aligned} F(y_k) + J_k d_k^{\mathrm{MLM}} &= F(y_k) - U_1\Sigma_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F(y_k) - U_2\Sigma_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F(y_k) \\ &= \lambda_k U_1(\Sigma_1^2 + \lambda_k I)^{-1}U_1^T F(y_k) + \lambda_k U_2(\Sigma_2^2 + \lambda_k I)^{-1}U_2^T F(y_k) + U_3U_3^T F(y_k). \end{aligned} \tag{3.27}$$

In the following, we will give the estimations of $\|U_1U_1^TF_k\|$, $\|U_2U_2^TF_k\|$, $\|U_3U_3^TF_k\|$ as well as $\|U_1U_1^TF(y_k)\|$, $\|U_2U_2^TF(y_k)\|$, and $\|U_3U_3^TF(y_k)\|$.

Lemma 3.4. *Under the conditions of Assumption* 3.1, *if* $x_k \in N(x^*, b/2)$, *then we obtain*
(*a*) $\|U_1U_1^TF_k\| \le \kappa_{bf}\|\bar{x}_k - x_k\|$,
(*b*) $\|U_2U_2^TF_k\| \le 2\kappa_{hj}\|\bar{x}_k - x_k\|^{1+v}$,
(*c*) $\|U_3U_3^TF_k\| \le \kappa_{hj}\|\bar{x}_k - x_k\|^{1+v}$,

where κ_{bf}, κ_{hj} *are given in* (3.2) *and* (3.4), *respectively*.

Proof. We derive the result (a) from (3.4) immediately.

Let $\bar{s}_k = -J_k^+F_k$, where J_k^+ is the pseudo-inverse of J_k. It is easy to see that $\bar{s}_k$ is the least squares solution of $\min_{s\in R^n}\|F_k + J_ks\|$. Hence, we have from (3.3) that

$$\|U_3U_3^TF_k\| = \|F_k + J_k\bar{s}_k\| \le \|F_k + J_k(\bar{x}_k - x_k)\| \le \kappa_{hj}\|\bar{x}_k - x_k\|^{1+v}. \tag{3.28}$$

Let $\tilde{J}_k = U_1\Sigma_1V_1^T$ and $\tilde{s}_k = -\tilde{J}_k^+F_k$. Since $\tilde{s}_k$ is the least squares solution of $\min_{s\in R^n}\|F_k + \tilde{J}_ks\|$, it follows from (3.3) and (3.11)

$$\begin{aligned} \|(U_2U_2^T + U_3U_3^T)F_k\| &= \|F_k + \tilde{J}_k\tilde{s}_k\| \\ &\le \|F_k + \tilde{J}_k(\bar{x}_k - x_k)\| \\ &\le \|F_k + J_k(\bar{x}_k - x_k)\| + \|(\tilde{J}_k - J_k)(\bar{x}_k - x_k)\| \end{aligned} \tag{3.29}$$

$$\begin{aligned}&\le \frac{\kappa_{hj}}{1+v}\|\bar{x}_k - x_k\|^{1+v} + \|U_2\Sigma_2V_2^T(\bar{x}_k - x_k)\|\\&\le \frac{\kappa_{hj}}{1+v}\|\bar{x}_k - x_k\|^{1+v} + \kappa_{hj}\|\bar{x}_k - x_k\|^{v}\|\bar{x}_k - x_k\|\\&\le 2\kappa_{hj}\|\bar{x}_k - x_k\|^{1+v}.\end{aligned}$$

Due to the orthogonality of U_2 and U_3, we can obtain result (b). ☐

Lemma 3.5. *Under the conditions of Assumption* 3.1, *if* $x_k, y_k \in N(x^*, b/2)$, $v > 1/\gamma - 1$ *and* $1/\gamma - 1 < \alpha \le 2\gamma v$, *then we obtain*

(*a*) $\|U_1U_1^TF(y_k)\| \le c_8\|\bar{x}_k - x_k\|^{\min\{1+\alpha,1+v\}}$,

(*b*) $\|U_2U_2^TF(y_k)\| \le c_9\|\bar{x}_k - x_k\|^{\min\{v+\gamma+\gamma\alpha,v+\gamma+\gamma v\}}$,

(*c*) $\|U_3U_3^TF(y_k)\| \le c_{10}\|\bar{x}_k - x_k\|^{\min\{v+\gamma+\gamma\alpha,v+\gamma+\gamma v\}}$,

where c_8, c_9, c_{10} *are positive constants.*

Proof. It follows from (3.12), (3.26), and Lemma 3.4 that

$$\|F_k + J_kd_k^{\mathrm{LM}}\| \le \left(\frac{4M\kappa_{bf}^{1+\alpha}}{\bar{\sigma}_r^2} + 3\kappa_{hj}\right)\|\bar{x}_k - x_k\|^{\min\{1+\alpha,1+v\}}, \tag{3.30}$$

which together with (3.3) and (3.8) imply that

$$\begin{aligned}\|F(y_k)\| &= \|F(x_k + d_k^{\mathrm{LM}})\|\\&\le \|F_k + J_kd_k^{\mathrm{LM}}\| + \frac{\kappa_{hj}}{1+v}\|d_k^{\mathrm{LM}}\|^{1+v}\\&\le \left(\frac{4M\kappa_{bf}^{1+\alpha}}{\bar{\sigma}_r^2} + 2\kappa_{hj} + \frac{\kappa_{hj}}{1+v}c_{11}\right)\|\bar{x}_k - x_k\|^{\min\{1+\alpha,1+v\}}\\&\le c_8\|\bar{x}_k - x_k\|^{\min\{1+\alpha,1+v\}},\end{aligned} \tag{3.31}$$

for some $c_8, c_{11} > 0$, which gives result (a).

From (3.1), we can obtain

$$\|\bar{y}_k - y_k\| \le c^{-1}\|F(y_k)\|^{\gamma} \le c_{12}\|\bar{x}_k - x_k\|^{\min\{\gamma(1+\alpha),\gamma(1+v)\}}, \tag{3.32}$$

where c_{12} is a positive constant.

Let $\bar{p}_k = -J_k^+F_k$, then $\bar{p}_k$ is the least squares solution of $\min_{p\in R^n}\|F_k + J_kp\|$. It follows from (3.1), (3.2), and (3.32) that

$$\begin{aligned}\|U_3U_3^TF(y_k)\| &= \|F(y_k) + J_k\bar{p}_k\|\\&\le \|F(y_k) + J_k(\bar{y}_k - y_k)\|\\&\le \|F(y_k) + J(y_k)(\bar{y}_k - y_k)\| + \|(J_k - J(y_k))(\bar{y}_k - y_k)\|\\&\le \frac{\kappa_{hj}}{1+v}\|\bar{y}_k - y_k\|^{1+v} + \kappa_{hj}\|d_k^{\mathrm{LM}}\|^{v}\|\bar{y}_k - y_k\|\\&\le \frac{\kappa_{hj}}{1+v}c_{12}\|\bar{x}_k - x_k\|^{\min\{\gamma(1+\alpha)(1+v),\gamma(1+v)^2\}} + c_{13}\|\bar{x}_k - x_k\|^{\min\{v+\gamma(1+\alpha),v+\gamma(1+v)\}}\\&\le c_{10}\|\bar{x}_k - x_k\|^{\min\{v+\gamma+\gamma\alpha,v+\gamma+\gamma v\}},\end{aligned} \tag{3.33}$$

for some positive c_{10} and c_{13}.

Let $\tilde{J}_k = U_1\Sigma_1V_1^T$ and $\tilde{p}_k = -\tilde{J}_k^+F_k$. Since $\tilde{p}_k$ is the least squares solution of $\min_{p\in R^n}\|F_k + \tilde{J}_kp\|$, deducing from (3.2), (3.3), and (3.32), we have

$$\begin{aligned}\|(U_2U_2^T + U_3U_3^T)F(y_k)\| &= \|F(y_k) + \tilde{J}_k\tilde{p}_k\|\\&\le \|F(y_k) + \tilde{J}_k(\bar{y}_k - y_k)\|\end{aligned} \tag{3.34}$$

$$
\begin{aligned}
&\le \|F(y_k) + J_k(\bar{y}_k - y_k)\| + \|(\tilde{J}_k - J(y_k))(\bar{y}_k - y_k)\| \\
&\le \frac{\kappa_{hj}}{1+v} \|\bar{y}_k - y_k\|^{1+v} + \|(\tilde{J}_k - J(y_k) - U_2\Sigma_2 V_2^T)(\bar{y}_k - y_k)\| \\
&\le \frac{\kappa_{hj}}{1+v} \|\bar{y}_k - y_k\|^{1+v} + \|(\tilde{J}_k - J(y_k))(\bar{y}_k - y_k)\| + \|U_2\Sigma_2 V_2^T(\bar{y}_k - y_k)\| \\
&\le \frac{\kappa_{hj}}{1+v} \|\bar{y}_k - y_k\|^{1+v} + \kappa_{hj} \|d_k^{\mathrm{LM}}\|^v \|\bar{y}_k - y_k\| + \kappa_{hj} \|\bar{x}_k - x_k\|^v \|\bar{y}_k - y_k\| \\
&\le \frac{\kappa_{hj}}{1+v} c_{12} \|\bar{x}_k - x_k\|^{\min\{\gamma(1+\alpha)(1+v), \gamma(1+v)^2\}} + c_{14} \|\bar{x}_k - x_k\|^{\min\{v+\gamma+\gamma\alpha, v+\gamma+\gamma v\}} \\
&\le c_9 \|\bar{x}_k - x_k\|^{\min\{v+\gamma+\gamma\alpha, v+\gamma+\gamma v\}},
\end{aligned}
$$

for some positive c_9 and c_{14}. Due to the orthogonality of U_2 and U_3, we can obtain result (b). □

Theorem 3.6. *Under the conditions of Assumption* 3.1, *if* $v > 1/\gamma - 1$ *and* $1/\gamma - 1 < \alpha \le 2\gamma v$, *the sequence* $\{x_k\}$ *generated by Algorithm* 2.1 *converges to some solution of* (1.1) *with order* ϑ

$$
\vartheta = \min\{\gamma(1+2\alpha), \gamma(1+2v), \gamma(1+\alpha+v), \gamma(v+\gamma+\gamma\alpha), \gamma(v+\gamma+\gamma v)\}. \tag{3.35}
$$

Proof. From (3.6), (3.11), (3.12), (3.25), (3.27), and Lemma 3.5, we have

$$
\begin{aligned}
\|d_k^{\mathrm{MLM}}\| &= \|-V_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F(y_k) - V_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F(y_k)\| \\
&\le \|\Sigma_1^{-1}\| \|U_1^T F(y_k)\| + \|\lambda_k^{-1}\Sigma_2\| \|U_2^T F(y_k)\| \\
&\le \frac{2c_8}{\bar{\sigma}_r} \|\bar{x}_k - x_k\|^{\min\{1+\alpha, 1+v\}} + c_{15} \|\bar{x}_k - x_k\|^{\min\left\{\gamma+\gamma\alpha+2v-\frac{\alpha}{\gamma}, \gamma+\gamma v+2v-\frac{\alpha}{\gamma}\right\}} \\
&\le c_{16} \|\bar{x}_k - x_k\|^{\min\left\{1+\alpha, 1+v, \gamma+\gamma\alpha+2v-\frac{\alpha}{\gamma}, \gamma+\gamma v+2v-\frac{\alpha}{\gamma}\right\}}
\end{aligned} \tag{3.36}
$$

and

$$
\begin{aligned}
\|F(y_k) + J_k d_k^{\mathrm{MLM}}\| &= \|\lambda_k U_1(\Sigma_1^2 + \lambda_k I)^{-1} U_1^T F(y_k) + \lambda_k U_2(\Sigma_2^2 + \lambda_k I)^{-1} U_2^T F(y_k) + U_3 U_3^T F(y_k)\| \\
&\le \lambda_k \|\Sigma_1^{-2}\| \|U_1^T F(y_k)\| + \|U_2^T F(y_k)\| + \|U_3^T F(y_k)\| \\
&\le \frac{4Mc_8\kappa_{bf}^{\alpha}}{\bar{\sigma}_r^2} \|\bar{x}_k - x_k\|^{\min\{1+2\alpha, 1+v+\alpha\}} + (c_9 + c_{10})\|\bar{x}_k - x_k\|^{\min\{v+\gamma+\gamma\alpha, v+\gamma+\gamma v\}} \\
&\le c_{17} \|\bar{x}_k - x_k\|^{\min\{1+2\alpha, 1+v+\alpha, v+\gamma+\gamma\alpha, v+\gamma+\gamma v\}},
\end{aligned} \tag{3.37}
$$

where c_{15}, c_{16}, c_{17} are positive constants.

Combining (3.1), (3.2), (3.8)–(3.10), (3.36), and (3.37), we obtain

$$
\begin{aligned}
(c\|\bar{x}_{k+1} - x_{k+1}\|)^{1/\gamma} &\le \|F(x_{k+1})\| = \|F(y_k + d_k^{\mathrm{MLM}})\| \\
&\le \|F(y_k) + J(y_k)d_k^{\mathrm{MLM}}\| + \frac{\kappa_{hj}}{1+v} \|d_k^{\mathrm{MLM}}\|^{1+v} \\
&\le \|F(y_k) + J_k d_k^{\mathrm{MLM}}\| + \|(J(y_k) - J_k)d_k^{\mathrm{MLM}}\| + \frac{\kappa_{hj}}{1+v} \|d_k^{\mathrm{MLM}}\|^{1+v} \\
&\le c_{17} \|\bar{x}_k - x_k\|^{\min\{1+2\alpha, 1+v+\alpha, v+\gamma+\gamma\alpha, v+\gamma+\gamma v\}} + \kappa_{hj} \|d_k^{\mathrm{LM}}\|^v \|d_k^{\mathrm{MLM}}\| + \frac{\kappa_{hj}}{1+v} \|d_k^{\mathrm{MLM}}\|^{1+v} \\
&\le c_{18} \|\bar{x}_k - x_k\|^{\min\{1+2\alpha, 1+2v, 1+\alpha+v, v+\gamma+\gamma\alpha, v+\gamma+\gamma v\}},
\end{aligned} \tag{3.38}
$$

where c_{18} is the positive constant.

Therefore,

$$
c\|\bar{x}_{k+1} - x_{k+1}\| \le c_{18}^{\gamma} \|\bar{x}_k - x_k\|^{\min\{\gamma(1+2\alpha), \gamma(1+2v), \gamma(1+\alpha+v), \gamma(v+\gamma+\gamma\alpha), \gamma(v+\gamma+\gamma v)\}}. \tag{3.39}
$$

The proof is completed. □

Corollary 3.7. *Under the conditions of Assumption* 3.1.

(*i*) *If* $\gamma = 1$ *and* $v \in \left(0, \frac{1}{2}\right)$, *then*

$$\|\bar{x}_{k+1} - x_{k+1}\| \le \begin{cases} O(\|\bar{x}_k - x_k\|^{1+2\alpha}), & \text{if } \alpha \in (0, v), \\ O(\|\bar{x}_k - x_k\|^{1+2v}), & \text{if } \alpha \in [v, 2v]. \end{cases} \tag{3.40}$$

(*ii*) *If* $v = 1$ *and* $\gamma \in \left(\frac{1}{2}, 1\right]$, *then*

$$\|\bar{x}_{k+1} - x_{k+1}\| \le \begin{cases} O(\|\bar{x}_k - x_k\|^{\gamma(1+\gamma+\gamma\alpha)}), & \text{if } \alpha \in \left(\frac{1}{\gamma} - 1, 1\right), \\ O(\|\bar{x}_k - x_k\|^{\gamma(1+2\gamma)}), & \text{if } \alpha \in [1, 2\gamma]. \end{cases} \tag{3.41}$$

Particularly, if $\gamma = v = 1$, *then*

$$\|\bar{x}_{k+1} - x_{k+1}\| \le \begin{cases} O(\|\bar{x}_k - x_k\|^{1+2\alpha}), & \text{if } \alpha \in (0, 1), \\ O(\|\bar{x}_k - x_k\|^{3}), & \text{if } \alpha \in [1, 2]. \end{cases} \tag{3.42}$$

Corollary 3.7 indicates that, if $\gamma = v = 1$, i.e., if $F(x)$ satisfies the local error bound condition and the Jacobian $J(x)$ is Lipschitz continuous, then the sequence $\{x_k\}$ generated by Algorithm 2.1 converges to some solution of (1.1) superlinearly with order $1 + 2\alpha$ for any $\alpha \in (0, 1)$ and cubic for any $\alpha \in [1, 2]$. This coincides with the results in [9].

4 Conclusion

In this article, we study the convergence rate of the MLM method under the Hölderian local error bound condition and the Hölderian continuity of the Jacobian, which are more general than the local error bound condition and the Lipschitz continuity of the Jacobian. If $\gamma = v = 1$, i.e., if $F(x)$ satisfies the local error bound condition and the Jacobian $J(x)$ is Lipschitz continuous, then the sequence $\{x_k\}$ generated by Algorithm 2.1 converges to some solution of (1.1) superlinearly with order $1 + 2\alpha$ for any $\alpha \in (0, 1)$ and cubic for any $\alpha \in [1, 2]$. This coincides with the results in [9].

Acknowledgements: The authors thank the referees for valuable comments and suggestions, which improved the presentation of this manuscript.

Funding information: This work was supported by the Natural Science Foundation of Education Bureau of Anhui Province (KJ2020A0017, KJ2017A432).

Conflict of interest: The authors state no conflict of interest.

References

[1] C. T. Kelley, *Solving nonlinear equations with Newton's method*, Fundamentals of Algorithms, SIAM, Philadelphia, 2003.
[2] K. Levenberg, *A method for the solution of certain nonlinear problems in least squares*, Quart. Appl. Math. **2** (1944), no. 2, 164–166.
[3] D. W. Marquardt, *An algorithm for least-squares estimation of nonlinear inequalities*, SIAM J. Appl. Math. **11** (1963), no. 2, 431–441, DOI: https://doi.org/10.1137/0111030.
[4] N. Yamashita and M. Fukushima, *On the rate of convergence of the Levenberg-Marquardt method*, in: G. Alefeld and X. Chen (eds), Topics in Numerical Analysis, Computing Supplementa, vol. 15, Springer, Vienna, 2001, DOI: https://doi.org/10.1007/978-3-7091-6217-0_18.

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

[6] J. Y. Fan, *A modified Levenberg-Marquardt algorithm for singular system of nonlinear equations*, J. Comput. Math. **21** (2003), no. 5, 625–636.

[7] J. J. Moré, *The Levenberg-Marquardt algorithm: implementation and theory*, In: G. A. Watson, (eds), Numerical Analysis. Lecture Notes in Mathematics, vol. 630, Springer, Berlin, Heidelberg, 1978, DOI: https://doi.org/10.1007/BFb0067700.

[8] J. Y. Fan, J. C. Huang, and J. Y. Pan, *An adaptive multi-step Levenberg-Marquardt method*, J. Sci. Comput. **78** (2019), no. 1, 531–548, DOI: https://doi.org/10.1007/s10915-018-0777-8.

[9] J. Y. Fan, *The modified Levenberg-Marquardt method for nonlinear equations with cubic convergence*, Math. Comp. **81** (2012), no. 277, 447–466, DOI: https://doi.org/10.1090/S0025-5718-2011-02496-8.

[10] X. Zhao and J. Y. Fan, *On the multi-point Levenberg-Marquardt method for singular nonlinear equations*, Comput. Appl. Math. **36** (2017), no. 1, 203–223, DOI: https://doi.org/10.1007/s40314-015-0221-8.

[11] L. Chen, *A high-order modified Levenberg-Marquardt method for systems of nonlinear equations with fourth-order convergence*, Appl. Math. Comput. **285** (2016), 79–93, DOI: https://doi.org/10.1016/j.amc.2016.03.031.

[12] M. Ahookhosh, F. J. Aragn, R. M. T. Fleming, and P. T. Vuong, *Local convergence of Levenberg-Marquardt methods under Hölder metric subregularity*, Adv. Comput. Math. **45** (2019), no. 5, 2771–2806, DOI: https://doi.org/10.1007/s10444-019-09708-7.

[13] H. Y. Wang and J. Y. Fan, *Convergence rate of the Levenberg-Marquardt method under Hölderian local error bound*, Optim. Methods Softw. **35** (2020), no. 4, 767–786, DOI: https://doi.org/10.1080/10556788.2019.1694927.

[14] X. D. Zhu and G. H. Lin, *Improved convergence results for a modified Levenberg-Marquardt method for nonlinear equations and applications in MPCC*, Optim. Methods Softw. **31** (2016), no. 4, 791–804, DOI: https://doi.org/10.1080/10556788.2016.1171863.

[15] M. J. D. Powell, *Convergence properties of a class of minimization algorithms*, Nonlinear Program. **2** (1975), 1–27, DOI: https://doi.org/10.1016/B978-0-12-468650-2.50005-5.

[16] W. Y. Sun and Y. X. Yuan, *Optimization Theory and Methods: Nonlinear Programming*, Springer, New York, 2006.

[17] Y. X. Yuan, *Recent advances in trust region algorithms*, Math. Program. **151** (2015), 249–281, DOI: https://doi.org/10.1007/s10107-015-0893-2.

[18] Y. X. Yuan, *Recent advances in numerical methods for nonlinear equations and nonlinear least squares*, Numer. Algebra Control Optim. **1** (2011), no. 1, 15–34, DOI: https://doi.org/10.3934/naco.2011.1.15.

[19] J. Y. Fan, *Accelerating the modified Levenberg-Marquardt method for nonlinear equations*, Math. Comput. **83** (2014), no. 287, 1173–1187, DOI: https://doi.org/10.1090/S0025-5718-2013-02752-4.

[20] X. H. Miao, K. Yao, C. Y. Yang, and J. S. Chen, *Levenberg-Marquardt method for absolute value equation associated with second-order cone*, Numer. Algebra Control Optim. **12** (2022), no. 1, 47–61, DOI: https://doi.org/10.3934/naco.2021050.

[21] H. Y. Wang and J. Y. Fan, *Convergence properties of inexact Levenberg-Marquardt method under Hölderian local error bound*, J. Ind. Manag. Optim. **17** (2021), no. 4, 2265–2275, DOI: https://doi.org/10.3934/jimo.2020068.

[22] Z. F. Dai, T. Li, and M. Yang, *Forecasting stock return volatility: The role of shrinkage approaches in a data-rich environment*, J. Forecast. **41** (2022), no. 5, 980–996, DOI: https://doi.org/10.1002/for.2841.

[23] Z. F. Dai and H. Y. Zhu, *Time-varying spillover effects and investment strategies between WTI crude oil, natural gas and Chinese stock markets related to belt and road initiative*, Energy Econ. **108** (2022), 105883, DOI: https://doi.org/10.1016/j.eneco.2022.105883.

[24] J. Moré, B. Garbow, and K. Hillstrom, *Testing unconstrained optimization software*, ACM Trans. Math. Software **7** (1981), no. 1, 17–41.

[25] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory, Computer Science and Scientific Computing*, Academic Press, Boston, 1990.

[26] L. Zheng, L. Chen, and Y. H. Ma, *A variant of the Levenberg-Marquardt method with adaptive parameters for systems of nonlinear equations*, AIMS Math. **7** (2021), no. 1, 1241–1256, DOI: https://doi.org/10.3934/math.2022073.